University of Groningen

MSc Biomolecular Sciences

Second research project

February-July 2022

Final report of Gaia Scilironi (s4500431)

Substrate recognition by DNAJB1

Bukau Lab (ZMBH, University of Heidelberg)

Supervisors: Prof. M.H.K. Linskens (University of Groningen) Prof. B. Bukau (University of Heidelberg)

Daily supervisors: Dr. A. Wentink, Dr. A. Mogk, T.L. Dang

Abstract

J domain proteins (JDPs) are crucial molecules in the context of the cellular chaperones system. Different JDPs present different cellular activities. In particular, despite their structural similarities, while class A JDPs target misfolded monomeric proteins and small aggregates, class B JPDs mostly interact with larger aggregates, including amyloid fibrils, and, in particular, the class B JDP DNAJB1 presents the unique feature of participating in the amyloid fibrils disaggregation mediated by the HSP70 chaperone machinery. DNAJB1 presents a N-terminal J-domain (JD) which is followed by the GF-rich region, two C-terminal client binding domains (CTDI and CTDII) and a dimerization domain (DD): therefore, the protein forms functional homodimers. Since client binding motifs or patterns have not been identified yet for the human JDPs, we aimed to investigate the substrate binding specificity of DNAJB1 through the screening of peptide libraries (α-synuclein, p53 and luciferase). We screened the libraries with both the full-length protein and smaller constructs carrying its different subdomains (JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI, His-SUMO-CTDII). Our aim was to identify binding motifs and patterns on DNAJB1 clients, confirming the important role of the C-terminal domains in the substrate binding. These data will be compared with human class A JDPs peptide library screening to identify possible differences in the interactions with clients. We designed, cloned, expressed and purified the above listed constructs and analyzed their oligomeric state and ability to perform α-synuclein amyloid fibrils disaggregation. We tested anti-DNAJB1 and anti-CTDI-CTDII antibodies recognition of the constructs through Western blots and performed the peptide libraries screening. We have been able to clone and purify the full-length protein, JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI, His-SUMO-CTDII. We determined the oligomeric state of the constructs, obtaining monomeric JD-GF and CTDI-CTDII and oligomeric CTDI-CTDII-DD and full-length DNAJB1 and we demonstrated that an efficient fibrils disaggregation is obtained only in the presence of the full-length protein. We obtained an efficient antibodies mediated recognition for all the constructs and, thus, we could proceed with the peptide libraries screening. We identified a preference to bind peptides enriched in acidic and aromatic residues for the full-length protein and the constructs carrying the substrate binding domains CTDI and CTDII. JD-GF did have different binding specificities preferring hydrophobic residues-enriched peptides, being in line with the different interactions these subdomains have been reported to have . We confirmed the DNAJB1 binding site on α-synuclein C-terminus which was already identified by NMR. Overall, these data illustrate a binding specificity for DNAJB1 which is slightly different from the one of the bacterial class A DNAJ, opening new questions about the different specificities JDPs may have.

Acknowledgments

I strongly **thank Prof. Bukau** for having hosted me in his lab.

Thanks to Prof. Linskens for their internal supervision at the University of Groningen.

Many thanks to Dr. Wentink, Dr. Mogk and Lieu Dang for their fantastic supervision.

Thanks to Svenja Jaeger and Panagiotis Katikaridis for having allowed me to use some of their proteins.

Many thanks to Lieu Dang, Svenja Jaeger, Lucia Svoboda, Lena Gottschalk and the whole Bukau group for their extreme support during this project.

Many thanks to my family, in particular to my grandfather Candido, for their support during my whole university studies.

Table of contents

1. Introduction

1.1. The HSP70 chaperone machinery

The HSP70 chaperone machinery represents one of the most effective tools in the cellular protein quality control network with activities ranging from protein folding and assembly of newly synthetized polypeptides to prevention of protein aggregation, aggregates solubilization and refolding as well as protein degradation (Mayer and Bukau, 2005; Meimaridou et al., 2019; Wentink et al., 2019).

Central to this chaperone machine is the 70-kDa heat shock protein HSP70 (Rosenzweig et al., 2019). It presents a ~ 44 kDa nucleotide binding domain (NBD), through which HSP70 binds and hydrolyzes ATP (Flaherty et al., 1990). The NBD is connected by a flexible and hydrophobic linker to a ~ 27 kDa substrate binding domain (SBD), which is responsible for the interaction with substrates. The SBD is further subdivided in a two layer β-sandwich subdomain (βSBD) and an α-helical lid (αSBD) (Zhu et al., 1996) (Figure 1-A). HSP70 is able to cycle between an ATP bound, open conformation, where the helical lid docks onto the NBD, which presents high rates of substrate association and dissociation (meaning that the affinity for the substrate is low) and a closed, ADP-bound, conformation with the helical lid docked over the βSBD binding pocked, where the affinity for the trapped substrate is higher (Mayer et al., 2000) (Figure 1-B).

J-domain proteins (JDPs), also known as HSP40 interact with HSP70 (Sugito et al., 1995) and are required for an efficient functioning of the HSP70 chaperone machinery. They stimulate HSP70 ATPase activity (Minami et al., 1996) through the induction of conformational changes (Wu et al., 2020, Tomiczek et al., 2020) upon interaction with the HSP70 flexible linker between NBD and SBD (Kityk et al., 2018). Both HSP70 and JDPs are responsible for the conformational changes occurring in the substrate molecule (Rodriguez et al., 2008).

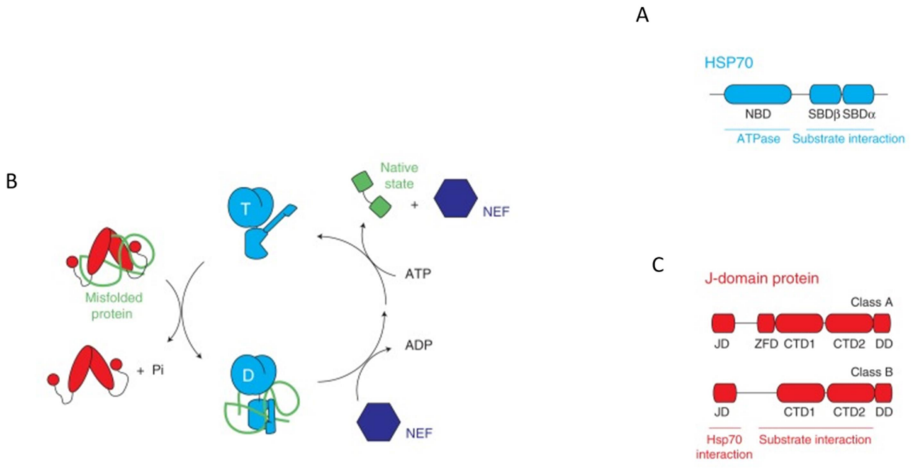


Figure 1: The HSP70 chaperone machinery. A) HSP70 (cyan) presents an NBD, essential for the ATP hydrolysis, and two subdomains αSBD and βSBD, which are responsible for the substrate binding B) The HSP70 chaperone (cyan) cycles between an ATP-bound open conformation with low substrate affinity and an ADP-bound close conformation with high affinity for substrates. JDPs (red) promote ATP hydrolysis while NEFs (violet) allow a fast ADP-ATP exchange. Substrates are shown in green. C) Class A and B JDPs (in red) are structurally similar: the N-terminal J-domain (JD) is followed by a GF-rich region, two substrate binding domains (CTDI and CTDII) and a dimerization domain. Class A CTDI protrudes in a Zinc finger domain (ZFD). Adapted from Wentink et al., 2019.

HSP110 proteins are nucleotide exchange factors (NEFs) which promote substrates release and HSP70 recycling, through an efficient ATP-ADP exchange (Rampelt et al., 2012; Mogk et al., 2018).

HSP70 identifies and bind sequence motifs of circa five amino acids enriched in hydrophobic residues flanked by positively charged residues. These sequences occur circa every 30-40 amino acids in proteins but are generally hidden in the proteins hydrophobic core: their exposure in unfolded or misfolded polypeptides allows HSP70 binding to a large range of different substrates (Rüdiger et al, 1997). JDPs recognize and bind substrates themselves, but interact with different binding sites when compared to HSP70 (Rüdiger et al, 2001). JDPs-substrates binding events precede or coincides with the substrate interaction with HSP70, thus JDPs work as substrate-targeting factors for HSP70 (Kampinga and Craig, 2010).

1.2. The J-domain protein families

J-domain proteins (JDPs) are chaperones (Kampinga and Craig, 2010) characterized by the conserved presence of a J domain, a helical domain of 70-75 residues, which is essential for the stimulation of HSP70 ATPase activity (Karzai and McMacken, 1996; Kityk et al., 2018). It presents a conserved HPD motif, essential for the domain function (Greene et al., 1998).

Three different classes of JDPs have been identified (Kelley, 1998; Qiu et al., 2006; Kampinga and Craig, 2010; Kampinga et al., 2018; Ayala Mariscal and Kirstein, 2021). Class A JDPs share the structural features of the bacterial DNAJ, presenting an N-terminal α-helical J domain (JD), followed by a structurally disordered Glycine/Phenylalanine (GF)- rich region, two C-terminal β-barrel domains (CTDI and CTDII), presenting each a client binding domain, and a C-terminal dimerization domain (DD). Class A JDPs present a Zinc finger-like region (ZFD) protruding from CTDI (Cheetham and Caplan, 1998; Li et al., 2003)(Figure 1-C). Class B JDPs share a similar domains architecture, but missing the Zinc finger domain (Borges et al., 2005; Hu et al, 2008) (Figure 1-C). Class C JDPs members present a J-domain within their structure, which can be placed anywhere in the protein and is accompanied by other different domains and functional motifs (Kampinga et al., 2018).

Class C JDPs are typically specific for a smaller set of substrates, while class A and B members interact with both unfolded, misfolded and aggregated proteins (Kampinga and Craig, 2010). Despite the structural similarities between the members of the A and B classes, they have evolved different features to drive specific functionalities in the context of the HSP70 chaperone machinery (Yu et al, 2015). While class A JDPs target small oligomeric species or monomeric misfolded proteins, being involved mainly in prevention of protein aggregation or refolding (Mattoo et al., 2013; Nillegoda et al., 2015; Nillegoda and Bukau, 2015; Irwin et al., 2022), class B JDPs interact primarily with larger oligomeric species, including also amyloid fibrils, and are involved in prevention of aggregates growth and in aggregates disaggregation (Nillegoda et al., 2015; Irwin et al., 2022). In fact, uniquely class B JDPs have been shown to play a role in the disaggregation of amyloid fibrils (Wentink et al., 2020; Faust et al., 2020).

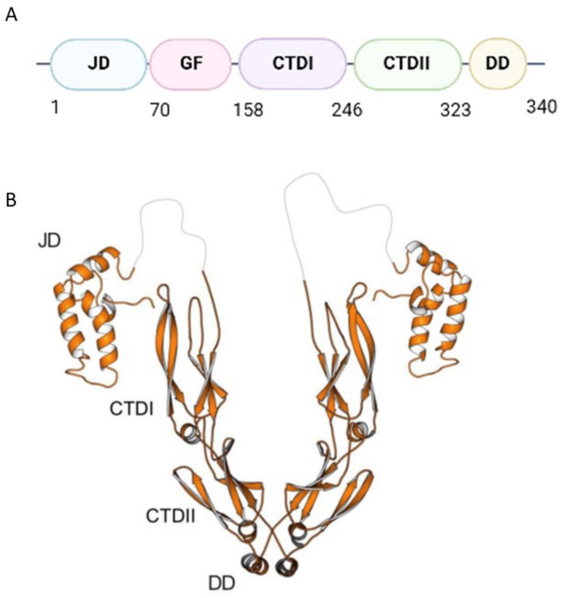


Figure 2: DNAJB1. A) DNAJB1 presents an N-terminal JD (light blue), interacting with HSP70 and stimulating its ATP hydrolysis, followed by the GF-rich region (pink), containing the helix V. CTDI (violet) and CTDII (green) contain each a substrate binding domain and DD (yellow) is required for dimerization. Made with Biorender.com. B) Structure of DNAJB1 homodimer: DNAJB1 forms homodimers due to the presence of the dimerization domain. The CTDs are structurally similar and β-barrel structured, the GF-rich region is mainly, but not completely, disordered, while the JD is mainly helical. Adapted from Wentink et al., 2020.

1.3.DNAJB1

DNAJB1 is a ~ 38 kDa human class B JDP which presents the structural features previously described for this class of proteins (Figure 2) (Hu et al., 2008).

This protein is able to form heterodimers with the class A JDP DNAJA2 acting on disordered aggregates (Nillegoda et al., 2018). Moreover, it has been demonstrated that DNAJB1 homodimers participate in amyloid fibrils disaggregation of α-synuclein (Duennwald et al., 2012; Gao et al. 2015; Wentink et al, 2020; Franco et al., 2021; Schneider et al., 2021; Beton et al., 2022), tau (Nachman et al., 2020), and Htt (Scior et al., 2018), concomitantly with the human constitutive HSP70 (HSC70) and the HSP110 NEF APG2 (Rampelt et al., 2012; Tittelmeier et al., 2020). In this context, DNAJB1 is able to recognize and bind amyloid fibrils, specifically targeting HSP70 to them (Pemberton et al., 2011; Wentink et al., 2020).

DNAJB1 presents a unique regulatory mechanism which has been demonstrated to be essential for amyloid fibrils disaggregation: a recently identified structural feature in the GF-rich region (helix V) docks onto the HSP70 binding site on DNAJB1 J domain, preventing DNAJB1 interaction with HSP70. An additional interaction between DNAJB1 CTDI and the C-terminal disordered region of HSP70 (EEVD peptide) is able to trigger conformational changes which release the helix V mediated J-domain inhibition. It has been demonstrated that this mechanism is strictly required for efficient fibril disaggregation: the proper clustering of HSP70 on the amyloid fibrils is essential to achieve the entropic pulling forces required to disassemble fibrils and depends on the presence of the helix V-mediated autoinhibition release mediated by EEVD-CTDI interactions (Faust et al., 2020).

DNAJB1 constitutively forms homodimers through the C-terminal dimerization domain (Figure 2-B). Dimers formation is crucial for DNAJB1 activity: within a dimer the number of substrates binding sites is doubled, compared to the monomeric species, and these multiple weak affinity binding events are crucial to overall stabilize the interaction with the substrates through the phenomenon of avidity. Thus, the presence of multiple aggregation prone motifs results in a higher chaperone affinity which may play a role in chaperone-substrate specificity (Wentink et al., 2020).

1.4. Aims and strategy

JDPs play an important role as targeting factors for HSP70. Therefore, the binding specificity of JDPs is a key player in the functioning of the whole HSP70 chaperone machinery. JDPs present at least four client binding sites in a dimer (one in each CTD domain) and it has been speculated that an additional binding region could be present in the GF-rich region of each monomer (Perales-Calvo et al., 2010; Kampinga et al., 2018). It has been demonstrated that multiple weak interaction in JDPs dimers stabilize the interaction with substrates by means of avidity (Wentink et al., 2020), but how the clients are selected and engaged remains

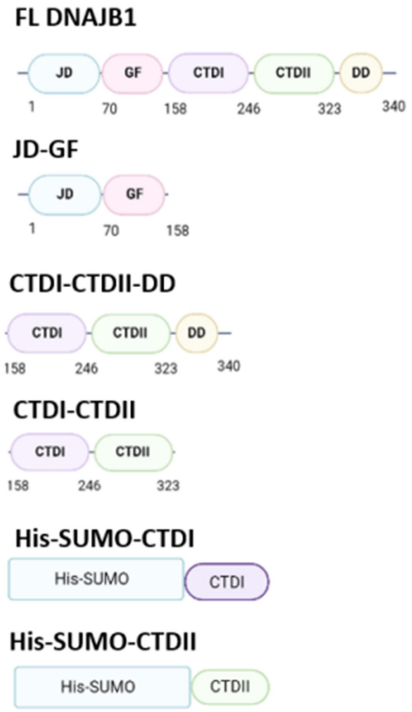


Figure 3: DNAJB1 constructs used in the peptide libraries screening. Six proteins were tested: FL DNAJB1, JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI, His-SUMO-CTDII. Made with Biorender.com

still an open question: substrate binding motifs or patterns have not been identified yet for the human JDPs. Furthermore, it has been speculated that the two CTDs have different binding specificities (Jiang et al., 2019). Moreover, the differences arisen between class A and B JDPs in their substrate binding specificity, despite the many structural features the two classes share, are still puzzling (Faust et al., 2020).

Therefore, analyzing JDPs substrates binding is meaningful. In particular, in this work, we analyzed the substrate binding specificity of DNAJB1 through screening of multiple peptides libraries. Peptide libraries are made of cellulose-bound, 13 residues-long peptides and represent a meaningful, easy-to-handle and quite high throughput tool to screen multiple peptides together for their binding to the chaperone (Rüdiger et al., 1997; Rüdiger et al., 2001).We screened three different peptide libraries: α-synuclein, p53 and luciferase (Supplementary information).

To efficiently describe the substrate binding properties of DNAJB1 we performed the libraries screening testing both the full-length DNAJB1 (FL DNAJB1) and several smaller constructs, carrying different subdomains (Table 3, Figure 3): apart from the full-length protein, we investigated the binding properties of the J-domain and GF-rich region (JD-GF), of the C-terminal domains with (CTDI-CTDII-DD) and without (CTDI-CTDII) the presence of the dimerization domain, to test the effects of dimerization and increase of client binding sites on DNAJB1 interactions with clients. Our hypothesis is that the presence of four binding sites, by means of avidity, strongly stabilizes the interaction with substrates. We also screened the single CTDI and CTDII domains (expressed with the tag His-SUMO).

In this study we first aimed to design, clone and efficiently express and purify DNAJB1 full-length and the smaller constructs carrying its subdomains and characterize their activity and oligomerization state. In order to perform our peptide libraries screening, we performed Western blots to control the efficiency of the antibodies-mediated recognition of DNAJB1 full-length and of its subdomains . Those preliminary experiments allowed us to eventually perform our peptide libraries screening. Using these screening we aimed to uncover motifs or patterns in the substrate sequence which are recognized and bound by DNAJB1. Testing the different DNAJB1 subdomains in the peptide libraries screening, we aimed to confirm the already identified binding sites of DNAJB1 (on CTDI and CTDII) (Hu et al., 2008) and assess their specificity.

Overall, in our work, we successfully cloned JD-GF, CTDI-CTDII-DD, CTDI-CTDII, CTDI and CTDII. We established efficient purification protocols for FL DNAJB1, JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI, His-SUMO-CTDII. We identified the oligomeric state of FL DNAJB1, CTDI-CTDII-DD, CTDI-CTDII and JD-GF: the full-length protein and CTDI-CTDII-DD eluted as oligomers, while JD-GF and CTDI-CTDII as monomers. We tested those four proteins for their ability to disassemble amyloid fibrils, which were efficiently disaggregated only in the presence of the full-length protein. Through our peptide libraries screening we identified a preference for binding hydrophobic peptides enriched in acidic and aromatic residues for the full-length DNAJB1. We confirmed the already identified (Wentink et al., 2020) DNAJB1

binding site on α-synuclein C-terminus. Moreover, we confirmed the substrate binding by CTDI and CTDII: CTDI-CTDII and CTDI-CTDII-DD preferentially interacted with hydrophobic peptides enriched in negatively charged and aromatic residues, similarly to FL DNAJB1.

2. Materials and methods

2.1. Constructs design

Constructs carrying JD-GF, CTDI-CTDII-DD, CTDI-CTDII, CTDI, CTDII domains were designed according to recent findings in DNAJB1 structure (Hu *et al.,* 2008). The domains boundaries are shown in Table 2. The constructs boundaries and expected molecular weights (calculated in Expasy) are shown in Table 3.

2.2. Cloning

JD-GF, CTDI-CTDII-DD, CTDI-CTDII, CTDI and CTDII constructs were cloned using standard pcr techniques (Optitaq polymerase, annealing temperature: 55°C, elongation temperature 68°C) followed by Gibson cloning (all reagents purchased from Invitrogen). The templates and primers used for the single constructs are shown in Table 1. All the cloning work was done using *E.coli* XL strains. Antibiotics (kanamycin) were purchased by Sigma-Aldricht and used with a final concentration of 50 µg/ml. Sequencing was provided by EurofinGenomics.

2.3. Protein overexpression and purification

Human DNAJB1 FL, DNAJB1 JD-GF, DNAJB1 CTDI-CTDII-DD, DNAJB1 CTDI-CTDII, DNAJB1 CTDI, DNAJB1 CTDII and HSP70 were expressed as His_6-Smt3 (His6-sumo) fusion proteins in *E. coli* Rosetta strains. Antibiotics were purchased by Sigma-Aldricht (kanamycin) and Carl Roth (chloramphenicol) used with a final concentration of 50 µg/ml. Protein expression was induced at OD_{600} of 0.7 with 1 M IPTG (Carl Roth). Constructs carrying DNAJB1 FL, DNAJB1 CTDI-CTDII-DD, DNAJB1 CTDI-CTDII, DNAJB1 CTDI, DNAJB1 CTDII were overexpressed at 18°C overnight, while DNAJB1 JD-GF and HSP70 constructs were overexpressed at 30°C for 5 hours.

Proteins were purified by nickel affinity purification (Ni-IDA, Macherey-Nagel). Buffers for DNAJB1 and its constructs purification were prepared as follow: lysis buffer (HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, imidazole pH 8 10 mM, β-mercaptoethanol 3 mM, 1 mM PMSF, Aprotinin, Leupeptin, Pepstatin, DNAse), wash buffer (HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, imidazole pH 8 10 mM, β-mercaptoethanol 3 mM), elution buffer (HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, imidazole pH 8 260 mM, β-mercaptoethanol 3 mM). The tag was cleaved by Ulp1 protease (4 mg/ml) and the excess of imidazole was removed by overnight dialysis (buffer: HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, β-mercaptoethanol 3 mM). HSP70 was purified with the same protocol but all the buffers had a KCl concentration of 150 mM. The cleaved tag was removed by reverse nickel affinity purification for HSP70, FL DNAJB1, JD-GF, CTDI-CTDII-DD, CTDI-CTDII. DNAJB1 FL and its constructs JD-GF, CTDI-CTDII-DD, CTDI-CTDII were further purified by SEC (Size exclusion chromatography: Superdex75, GE healthcare, buffer: HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, β-mercaptoethanol 3 mM) (Figure 4). The tag cleavage and the following steps were not performed in the case of His-SUMO-CTDI and His-SUMO-CTDII.

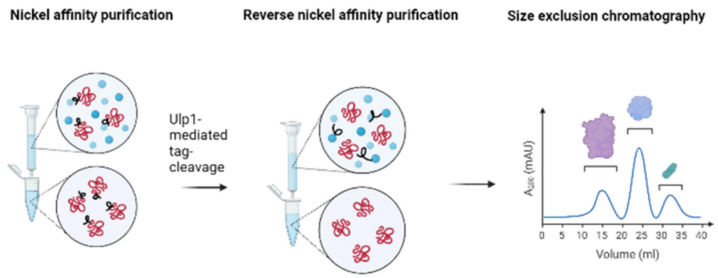


Figure 4: FL DNAJB1, JD-GF, CTDI-CTDII-DD, CTDI-CTDII, HSP70 purification set up. His-SUMO tagged proteins are purified with Nickel affinity purification. The tag is removed by Ulp1 cleavage and reverse nickel affinity purification. The samples are further purified by SEC. Made with Biorender.com.

2.4. Oligomeric state analyses

The oligomeric state of FL DNAJB1, CTDI-CTDII-DD, CTDI-CTDII and JD-GF was assessed by SEC (GE healthcare) using both Superdex75 (GE healthcare) and Superdex 200 (GE healthcare) for FL DNAJB1 and Superdex 75 for JD-GF, CTDI-CTDII-DD, CTDI-CTDII, CTDI, CTDII (buffer : HEPES-KOH pH 7.5 50 mM, KCl 500 mM, $MgCl_2$ 5 mM, glycerol 5%, β-mercaptoethanol 3 mM) . Standard curves were run using Gel Filtration Standard purchased by Bio-Rad.

2.5. Tht-based disaggregation assays

α-synuclein amyloid fibrils disaggregation was quantified detecting changes in ThT (thioflavin T) fluorescence over 8 hours at 30°C by FLUOrstar Omega plate reader (BMG LABTECH, excitation: 440 nm, emission: 480 nm). We incubated 2 µM of preformed WT α-synuclein fibrils with either 2 µM of DNAJB1 FL or its constructs JD-GF, CTDI-CTDII-DD, CTDI-CTDII, 4 µM of HSP70 (purified either by G. Scilironi or S. Jaeger) and 0.2 µM of APG2 (HSP110, purified by T.L. Dang), 2 mM ATP, 50 µM ThT in 50 mM HEPES-KOH pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 2 mM DTT.

2.6. Western blot

Different amounts of DNAJB1 FL, CTDI-CTDII-DD, CTDI-CTDII, JD-GF, His-SUMO-CTDI, His-SUMO-CTDII (20 ng, 40 ng, 80 ng, 160 ng, 320 ng) were run on 15% Express Plus PAGE in Tris-MOPS-SDS running buffer (GenScript). Proteins were transferred onto PVDF membranes (Thermofisher Scientific) by Trans-Blot Turbo RTA transfer kit (Bio-Rad). Proteins were immunoblotted with different anti-DNAJB1 polyclonal antibodies raised in rabbit: HSP40/Hdj1 polyclonal antibody (Enzo), HSP40 polyclonal antibody (Invitrogen), and anti our purified DNAJB1 CTDI-CTDII polyclonal antibodies , purchased by Davids Biotechnologie GMBH . Two different concentrations (1:1000 and 1:5000) of antibodies were tested. ECF substrate (GE healthcare) coupled with an anti-rabbit alkaline phosphatase-coupled secondary antibody (Invitrogen) were used to develop the membranes. Pictures of the blot were obtained on ImageQuant LAS-4000 (FUJIFILM Co). Signal was analyzed with Image Studio Lite Software (LI-COR Biosciences).

2.7. Peptide library screening

Peptide libraries of α-synuclein, p53 and luciferase were purchased from JPT Peptide Technologies GmbH. The libraries shared a similar occurrence of amino acids residues compared to the average found in natural proteins (Rüdiger et al., 2001). Peptide libraries were prepared and used as previously described by Rudiger *et al.* (1997, 2001). Libraries have been produced by automated spot synthesis. They have been C-terminally attached to cellulose by $(\beta\text{-Ala})_2$ spacer. Peptides are blocked through N-terminal acetylation. Each peptide presents a sequence of 13 amino acids and overlaps with the previous and following peptide by 10 amino acids. Libraries details are reported in supplementary information. The assay is done as follows: dry membranes are incubated with methanol at room temperature for 10 minutes, washed 3x10 minutes with Tris-buffered saline pH 7.6 (TBS: 10 mM Tris/HCl, 150 mM NaCl) at room temperature and washed 10 minutes with MP2 buffer (10 mM Tris/HCl pH 7.5, 150 mM KCl, 20 mM $MgCl_2$, 5% sucrose, 0.05% (v/v) Tween 20) at room temperature. Libraries are incubated with 1 µM of FL DNAJB1, CTDI-CTDII-DD, CTDI-CTDII, and 2 µM of JD-GF, His-SUMO-CTDI, and His-SUMO- CTDII in MP2 buffer for 30 minutes at room temperature, upon gentle shaking. Unbound proteins are removed washing at 4°C, 15 seconds, upon gentle shaking, with TBS. Bound proteins are electrotransferred with 4 consecutive blots on PVDF

membranes (Thermofisher scientific) using a semi-dry blotter (provided by Prof. Melchior lab). Each PVDF membrane is 0.5 cm larger than the cellulose library at each dimension. PVDF membranes are sandwiched between Whatman paper sheets activated by soaking with the anode buffers XA2 (231 mM Tris/HCl) and XA1 (69,2 mM Tris/HCl) and Whatman paper sheets soaked with catode buffer XK (57.7 mM Tris/HCl, 2.1 M amino hexan acid 0.1% (w/v) SDS). PVDF membranes are blotted at 4°C, for 30 minutes each, with a constant electrotransfer power 0.8 mA/cm^2 of the PVDF membranes. Blotted DNAJB1 constructs and FL are detected as described for the western blot. FL DNAJB1 was immunoblotted with 1:2500 anti-DNAJB1 plyclonal antibodies (Enzo). His-SUMO-CTDI, His-SUMO-CTDII, CTDI-CTDII-DD and CTDI-CTDII were immunoblotted using 1:1000 of specific anti-DNAJB1 CTDI-CTDII antibiotics purchased by Davids. JD-GF was detected by using 1:1000 anti-DNAJB1 antibodies (Invitrogen). The signal on the PVDF membrane is represented by dark spots on the white background. We identified the darkest, black spots as strong binder peptides. We identified the grey, less marked spots and the spots where only the borders were darker as weak binders. Where we did have white spots or no spots we identified non-binder peptides. The peptides were analyzed for their composition in terms of amino acids (polar, hydrophobic, aromatic, acidic or basic), their net charge at pH 7 and their hydrophilicity (Hopp and Woods, 1981), according to the calculations run in Bachem.com. The overall peptide libraries screening protocol is showed in Figure 5.

Table 1: Primers and templates used for JD-GF, CTDI-CTDII-DD, CTDI-CTDII, CTDI, CTDII cloning.

Construct	Forward primer	Reverse primer	Template
JD-GF	FW_universal_DNAJB1 5'TAGCTCGAGCACCACCACCACCACCACTG AGATCCGGCTGCTAAC 3'	RV_JD-GF_DNAJB1 5'TGGTGGTGGTGCTCGAGCTACTTTCGGGCG GGCTCTTGGGCAGAG 3'	SUMO-FL DNAJB1 pCA528
CTDI-CTDII-DD	FW_Ccd_DNAJB1 5'AAGCAAGATCCCCCAGTCACCCACGACCT TCGAGTCTCCCTTG 3'	RV_Ccd_DNAJB1 5'GTGACTGGGGGATCTTGCTTACCACCAATCT GTTCTCTGTGAG 3'	SUMO-FL DNAJB1 pCA528
CTDI-CTDII	FW_universal_DNAJB1 5'TAGCTCGAGCACCACCACCACCACCACTG AGATCCGGCTGCTAAC 3'	RV_Cc_DNAJB1 5'TGGTGGTGGTGCTCGAGCTAGGGGAAGATC ACTTCAAACTCAATAATG 3'	SUMO-CTDI-CTDII- DD pCA528
CTDI	FW_universal_DNAJB1 5'TAGCTCGAGCACCACCACCACCACCACTG AGATCCGGCTGCTAAC 3'	RV_CTD1_DNAJB1 5'TGGTGGTGGTGCTCGAGCTAGATATTGTGG GGCTTGTCCTTTAAAAC 3'	SUMO-CTDI-CTDII- DD pCA528
CTDII	FW_CTD2_DNAJB1 5'TTTAAGAGAGATGGCTCTGATGTCATTTAT CCTGCCAGGATCAGC 3'	RV_CTD2_DNAJB1 5'TCAGAGCCATCTCTCTTAAAACCACCAATCT GTTCTCTGTGAGCC 3'	SUMO-CTDI-CTDII pCA528

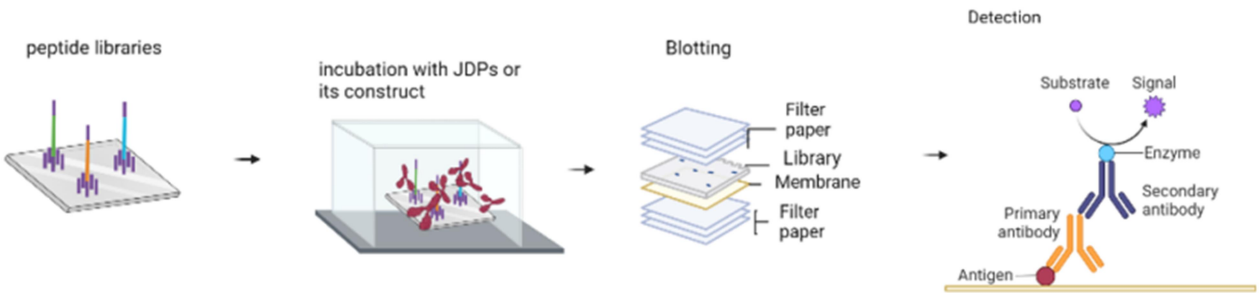


Figure 5: Peptide libraries screening set up. Libraries are activated and incubated with DNAJB1 constructs and blotted on PVDF membranes. The presence of the protein is detected by antibodies like in a standard western blot. Made with Biorender.com.

3. Results

3.1. Establishment of DNAJB1 domains purification

DNAJB1 subdomains constructs have been designed according to the domains boundaries mentioned in Hu et al., 2008; and Wentink et al., 2020 (Table 2).

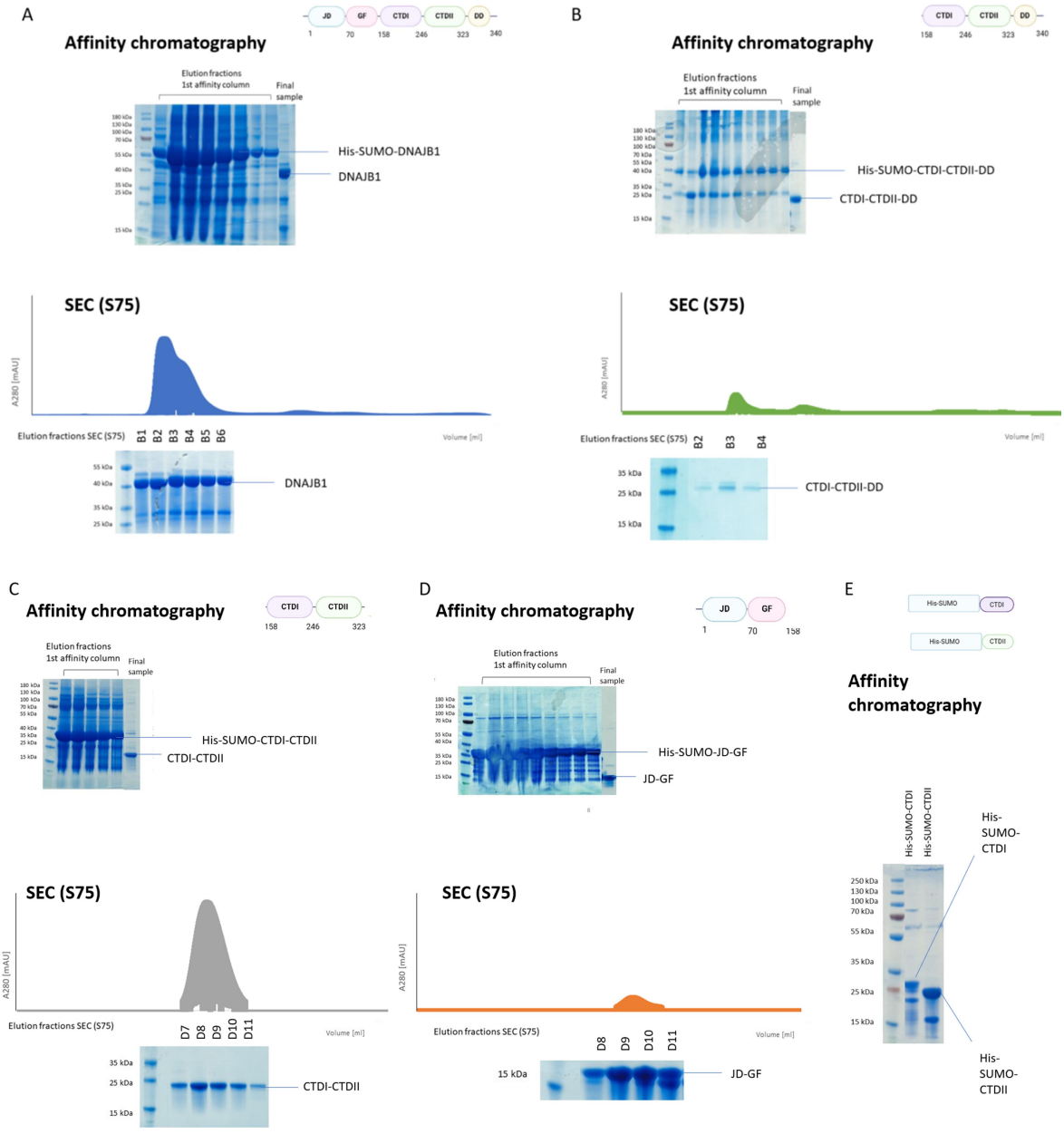


Figure 6: DNAJB1 full-length and constructs purification. A) Purification of FL DNAJB1 yielded an already quite pure DNAJB1 after the affinity chromatography (upper panel). A final pure protein was obtained after the SEC S75 (lower panel, chromatogram in blue). B) CTDI-CTDII-DD purification yielded already a pure protein after the affinity chromatography (upper panel). The result improved also after the SEC S75 (lower panel, chromatogram in green). C) CTDI-CTDII purification yielded already a pure protein after the affinity chromatography (upper panel). It was improved with the SEC S75 (lower panel, chromatogram in grey). D) JD-GF purification yielded a quite pure protein sample after the affinity chromatography (upper panel) which was strongly improved by the SEC S75 (lower panel, chromatogram in orange). E) Nickel affinity purification of His-SUMO-CTDI and His-SUMO-CTDII was quite successful.

We used standard pcr techniques to obtain vectors containing the desired constructs sequences, which have been confirmed by sequencing.

We overexpressed both the full-length DNAJB1 and its subdomains constructs as His-SUMO-tagged proteins. We, therefore, used a purification protocol based on nickel affinity chromatography and reverse nickel affinity chromatography after the tag cleavage by Ulp1, to obtain our proteins. The samples were further submitted to a gel filtration purification . This protocol gave pure proteins, suitable for the peptide libraries screening, in the case of DNAJB1 FL, CTDI-CTDII-DD, CTDI-CTDII and JD-GF (Figure 6-A, B, C, D). Unfortunately, neither this purification protocol nor other methods allowed an efficient purification of CTDI and CTDII (supplementary information). Since we have been able to obtain quite pure fragments using the nickel affinity purification on His-SUMO-CTDI and His-SUMO-CTDII, and the use of Ulp1 (even at high concentrations) cleavage resulted in a mixture of cleaved and uncleaved proteins (supplementary information), we decided to avoid the tag cleavage and work with the tagged substrates (Figure 6-E).

Table 2: Domains boundaries of DNAJB1.

Domain	Amino acids boundaries
J-domain (JD)	1-70
GF-rich region (GF)	71-158
C-terminal domain I (CTDI)	159-246
C-terminal domain II (CTDII)	247-323
Dimerization domain (DD)	324-340

3.2. FL DNAJB1 and CTDI-CTDII-DD elute as oligomers while CTDI-CTDII and JD-GF elute as monomers

An important quality control analysis for our proteins was represented by control of the proteins oligomeric state. In fact, due to the presence of the dimerization domain, the full-length DNAJB1 and CTDI-CTDII-DD should be able to form dimers as reported by structural analyses (Hu et al., 2008), while JD-GF and CTDI-CTDII are expected to be monomeric due to the absence of the dimerization domain. Confirming these hypothesis is, therefore, important, in the context of the peptide libraries screening, to elucidate the number of substrates binding domains present in each construct. In fact, the avidity phenomenon caused by the presence of at least four binding sites on a DNAJB1 dimer is responsible for a stabilization in the substrate-DNAJB1 interactions, which may lead to different results in the peptide libraries screening for monomeric and oligomeric chaperones.

To analyze the oligomeric states of the DNAJB1 constructs, we run a gel filtration column and compared the elution profile of the different proteins to standards (supplementary information). We determined the expected molecular weight for each construct in the monomeric state using Expasy.com. (Table 3). Due to the dimerization domain we expected to obtain dimeric FL DNAJB1 and CTDI-CTDII-DD, while the other constructs (CTDI-CTDII and JD-GF) were expected to be monomeric.

Table 3: Boundaries of the constructs used in our study and respective molecular weights (calculated with Expasy and experimentally determined by SEC).

Construct	Amino acids boundaries	Expected molecular weight calculated by Expasy (kDa)	Molecular weight of the SEC elution fractions (major species) (kDa)
Full-length (FL) DNAJB1	1-340	38.044	> 600
JD-GF	1-158	17.489	~ 18

CTDI-CTDII-DD	159-340	20.573	> 600
CTDI-CTDII	159-323	18.612	~ 18
CTDI	159-246	10.039	-
CTDII	247-323	8.591	-

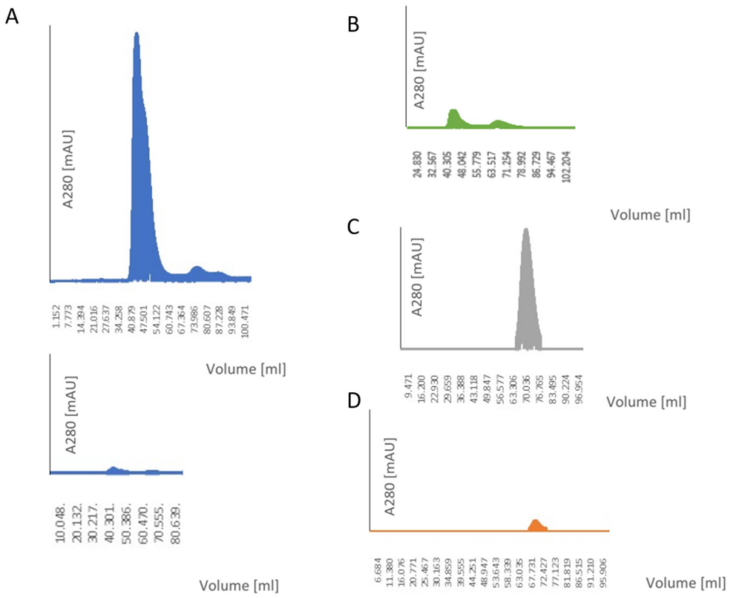


Figure 7: oligomeric state analyses of FL DNAJB1, CTDI-CTDII-DD, CTDI-CTDII, JD-GF. A) The elution profile of the FL DNAJB1 run on the S75 is shown in blue (upper panel, in blue) and shows the presence of different oligomeric states. Zooming on the elution profile obtained with the S200 (lower panel, in blue) a bigger early peak corresponding to multimers appears, followed by a minor peak representing dimers. B) The elution profile on SEC S75 of CTDI-CTDII-DD is represented in green and shows the presence of oligomers. C) The elution profile of CTDI-CTDII on SEC S75 is shown in grey and shows a sharp peak representing a monomeric species. D) The elution profile of JD-GF on SEC S75 is shown in orange and shows the presence of a monomeric protein.

We first run a Superdex75 (S75) column to characterize the oligomeric state of the full-length protein. From the chromatogram (Figure 7-A), we observed a peak in the early fractions corresponding to high molecular weight assembly. Moreover, the peak was not sharp, but presented a shoulder. Therefore, we hypothesized the presence of higher molecular weight species. To test our hypothesis, and gain a better resolution, we run a smaller amount of protein on a Superdex200 (S200) column. The elution profile (Figure 7-A) showed the presence of DNAJB1 multimers (molecular weight: higher than 600 kDa) which represented the most abundant species. In fact FL DNAJB1 started eluting at ~ 42 ml which represents also the beginning of the elution peak of the S200 standard curve (supplementary information) representing the elution of Thyroglobulin (molecular weight: 670 kDa). A small peak was also seen further ahead in the chromatogram: comparing it to the S200 standard curve we hypothesized it represented dimeric DNAJB1 species.

The elution profile on the S75 of CTDI-CTDII-DD clearly showed the presence of oligomers (molecular weight: circa 600 kDa) (Figure 7-B). The protein eluted mainly with a peak corresponding to the highest molecular weight proteins run in the standard curve (supplementary information), with a smaller peak corresponding to dimers.

Analyzing the chromatogram of CTDI-CTDII run on the S75, on the other hand, the protein eluted as a monomer, in line with the estimated molecular weight of circa 18 kDa (Figure 7-C): in fact a peak was seen in fractions slightly earlier than the ones where Myoglobin (17 kDa) runs in the S75 standard curve (supplementary information).

Lastly, JD-GF elution profile on the S75 showed the presence of monomers only (molecular weight circa 17-18 kDa) (Figure 7-D) eluting in similar fractions compared to Myoglobin (17 kDa) in the standard curve (supplementary information).

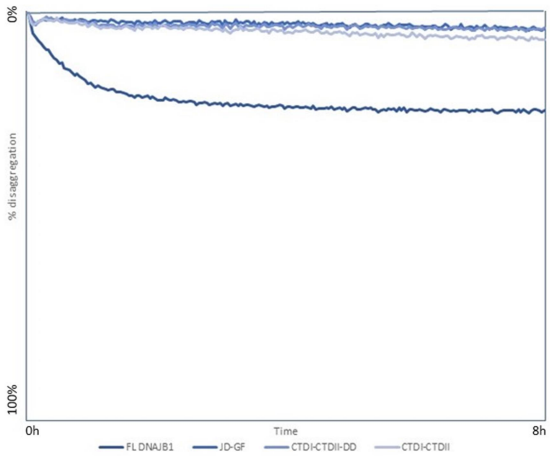


Figure 8: ThT mediated disaggregation assays. Efficient fibrils disaggregation (2µM of preformed WT α-synuclein fibrils either 2 µM of DNAJB1 FL or its constructs JD-GF, CTDI-CTDII-DD, CTDI-CTDII, 4 µM of HSP70 and 0.2 µM of, 2 mM ATP) only in the presence of the FL DNAJB1 (darkest blue). No disaggregation is seen for CTDI-CTDII-DD (light blue), CTDI-CTDII (grey) and JD-GF samples (blue). Data represent a duplicate.

Overall, the SEC analyses demonstrated that FL DNAJB1 and CTDI-CTDII-DD eluted mainly as oligomeric species while CTDI-CTDII and JD-GF as monomeric species. We could expect DNAJB1 oligomers as a major species: in fact, the tendency of being very sticky has been already reported for JDPs (Kampinga and Craig, 2010). This behavior has been reported to be even stronger in the absence of the J domain, and, indeed, the elution profile on the S75 of CTDI-CTDII-DD clearly showed the presence of oligomers. However, it is noteworthy that the depletion of the dimerization domain abolishes completely the presence of oligomeric species which are not found in CTDI-CTDII (and JD-GF).

3.3. Only DNAJB1 FL allows an efficient α-synuclein fibrils disaggregation

In order to understand precisely the role of each subdomain construct in the overall DNAJB1 activities, we investigated the ability of FL DNAJB1, CTDI-CTDII-DD, CTDI-CTDII and JD-GF to perform amyloid fibrils disaggregation in the presence of HSP70, the NEF APG2 and ATP. Knowing the importance of both the J domain and the DNAJB1 C-terminus in the interaction with HSP70 and their role in fibrils disaggregation (Wentink et al., 2020) we expected to obtain disaggregation only in the presence of the full-length protein.

We tested our hypothesis by monitoring the fluorescence of thioflavin T (ThT) over a 8-hours reaction incubating DNAJB1 full-length or its constructs, with HSP70, ATP and APG2. ThT detection allows to detect the presence and amount of amyloid fibrils over time. Our results clearly shows disaggregation (circa 20%) only in the presence of the full-length protein, while no significant decrease in the ThT signal was detected for the samples carrying CTDI-CTDII-DD, CTDI-CTDII and JD-GF (Figure 8). This result is consistent with the previously reported data which indicate the requirement of the full-length wild type DNAJB1 to efficiently stimulate fibrils disaggregation (Wentink et al., 2020).

3.4. Western blots establish different efficiency in antibodies-mediated recognition of the different DNAJB1 constructs

Detection through antibodies is essential for an efficient screening of the peptide libraries. That is why we tested two different anti-DNAJB1 polyclonal antibodies (purchased by Enzo and Invitrogen), at different dilutions and with different proteins concentrations, for their recognition of the different constructs we wanted to test for the libraries screening.

The FL DNAJB1 was well recognized by both the antibodies at all the tested concentrations (supplementary information). Nevertheless, the best results were achieved with Enzo antibodies, already at a 1:5000 dilution (Figure 9-A). The JD-GF construct was recognized well by both the antibodies, especially at 1:1000 dilution (supplementary information), but Invitrogen antibodies showed a better specificity for this fragment (Figure 9-B).

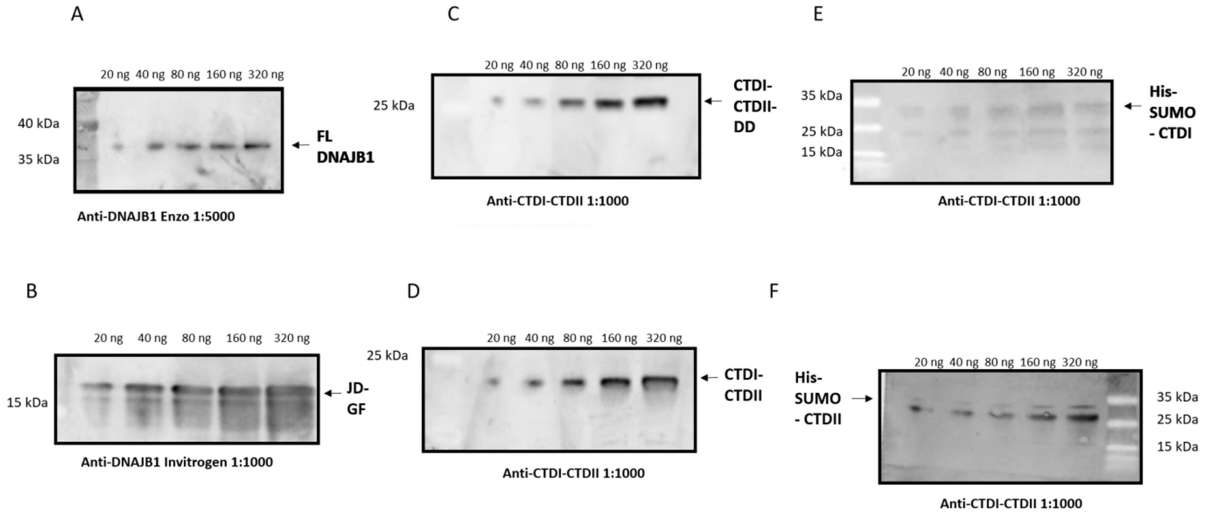


Figure 9: Western blots. A) Enzo polyclonal anti-DNAJB1 antibodies efficiently recognized different FL DNAJB1 concentrations at a dilution 1:5000. B) Invitrogen polyclonal anti-DNAJB1 polyclonal antibodies efficiently different concentrations of JD-GF at a dilution 1:1000. C) Davids anti-CTDI-CTDII polyclonal antibodies efficiently recognized CTDI-CTDII-DD at different concentration using a 1:1000 antibodies dilution. D) Davids anti CTDI-CTDII purchased polyclonal antibodies efficiently recognized different concentrations of CTDI-CTDII at a 1:1000 dilution. E) Davids anti-CTDI-CTDII polyclonal antibodies recognized His-SUMO-CTDI, in particular at higher protein concentration while testing a 1:1000 antibodies dilution. F) Davids anti-CTDI-CTDII-polyclonal antibodies efficiently recognized His-SUMO-CTDII at a 1:1000 dilution in particular at high protein concentrations.

Unfortunately, CTDI-CTDII-DD and CTDI-CTDII, and Ulp1 cleaved CTDI and CTDII were not strongly recognized by the polyclonal antibodies (neither by Enzo, nor by Invitrogen ones) at any dilutions and protein concentrations (supplementary information). We hypothesize that the main antigens recognized by those antibodies in DNAJB1 are located within the J domain and the GF-rich region. We, therefore, sent newly purified CTDI-CTDII to Davids Biotechnologie GMBH. The company provided us with rabbit serum containing anti-CTDI-CTDII polyclonal antibodies. We tested those antibodies on JD-GF (as a negative control), FL DNAJB1 (supplementary information), CTDI-CTDII-DD (Figure 9-C), CTDI-CTDII (Figure 9-D), His-SUMO-CTDI (Figure 9-E) and His-SUMO-CTDII (Figure 9-F) with different protein concentrations. The signal given for the constructs containing CTDII was generally good, in particular in the case of CTDI-CTDII-DD, CTDI-CTDII and the full-length protein. His-SUMO-CTDI was less well recognized by the antibody. We therefore hypothesize that the main antigen bound by the new polyclonal antibodies was located in CTDII. JD-GF, as expected, was not recognized by anti-CTDI-CTDII polyclonal antibodies (supplementary information).

Overall, we have been able to obtain generally a good signal from the different antibodies detection and we could, thus, proceed with the peptide libraries screening.

3.5. Peptide libraries screening identifies DNAJB1 full-length and its subdomains binding sites

In order to elucidate DNAJB1 client binding specificity, we performed a peptide libraries screening. In our work we screened three different peptide libraries: α-synuclein, p53 and luciferase (supplementary information).

α-synuclein is a small protein whose aberrant aggregation has been liked with neurodegenerative disorders such as Parkinson's disease (Bras et al., 2020). The relationship between DNAJB1 and α-synuclein has been extensively described due to the ability of the HSP70 chaperone machinery to bind and disaggregate α-synuclein fibrils (Pemberton et al., 2011; Duennwald et al., 2012; Gao et al. 2015; Wentink et al., Franco et al., 2021; Beton et al., 2022).

p53 is a transcription factor which is activated in response of cellular stress and represents the most frequently mutated protein in human tumors (Mantovani et al., 2019). Regulation of p53 mediated by DNAJB1 through the interaction with MDM2 has been reported (Qi et al., 2013), and, more importantly, the HSP70 chaperone machinery has been demonstrated to unfold p53 (Boysen et al., 2019).

Luciferase (from *R. Muelleri* in our study) is a commonly used substrate for protein refolding assays (Minami et al., 1996; Lüders et al., 1998) and has been, as well as p53, already screened in libraries binding experiments for the bacterial DNAK and DNAJ (Rüdiger et al., 1997; Rüdiger et al., 2001).

Each library was screened with FL DNAJB1; while only luciferase and p53 libraries were screened for JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI and His-SUMO-CTDII, due to the expiration of the α-synuclein library.

a. FL DNAJB1

First, we screened α-synuclein, p53 and luciferase libraries with FL DNAJB1 (Figure 10-A). To test the consistency of the results we repeated the experiment multiple times (supplementary information). Basing on Rüdiger's work on the bacterial class A DNAJ (2001), we expected to identify a shared binding specificity (meaning selectivity for hydrophobic residues) between DNAJB1 and DNAJ.

DNAJB1 showed the ability to discriminate between the peptide sequences, binding only some of them. Blotting results were visible on the third blot for the p53 and luciferase libraries and on the first blot for the α-synuclein library. Often clusters of binder peptides were found, indicating that a binding sequence was shared between the peptides.

Putting the data together, a total of 273 peptides were analyzed. Overall, 44 weak binders and 19 binders were identified. Analyzing the whole peptide sequences and classifying strong, weak and non-binder peptides, we reported a preference in binding hydrophobic peptides (Figure 10-B). Moreover, the net charge of the peptides which were engaged by FL DNAJB1 was negative (Figure 10-C). This is consistent with the preference of binding peptides enriched in acidic residues. Moreover, FL DNAJB1 showed a preference in binding aromatic residues enriched peptides. Leucine, Valine and Isoleucine (hydrophobic residues) were also abundant in strong binders, which underlines the importance of hydrophobicity in DNAJB1 substrate recognition. On the other hand, a decrease in the occurrence of basic residues in binder peptides was observed if compared to the non-binders or to the whole peptides pool (total)(Figure 10-D). The presence of hydrophobic and aromatic residues in the binding site is also consistent with the substrate binding sites identified for DNAJ (Rüdiger et al., 2001).

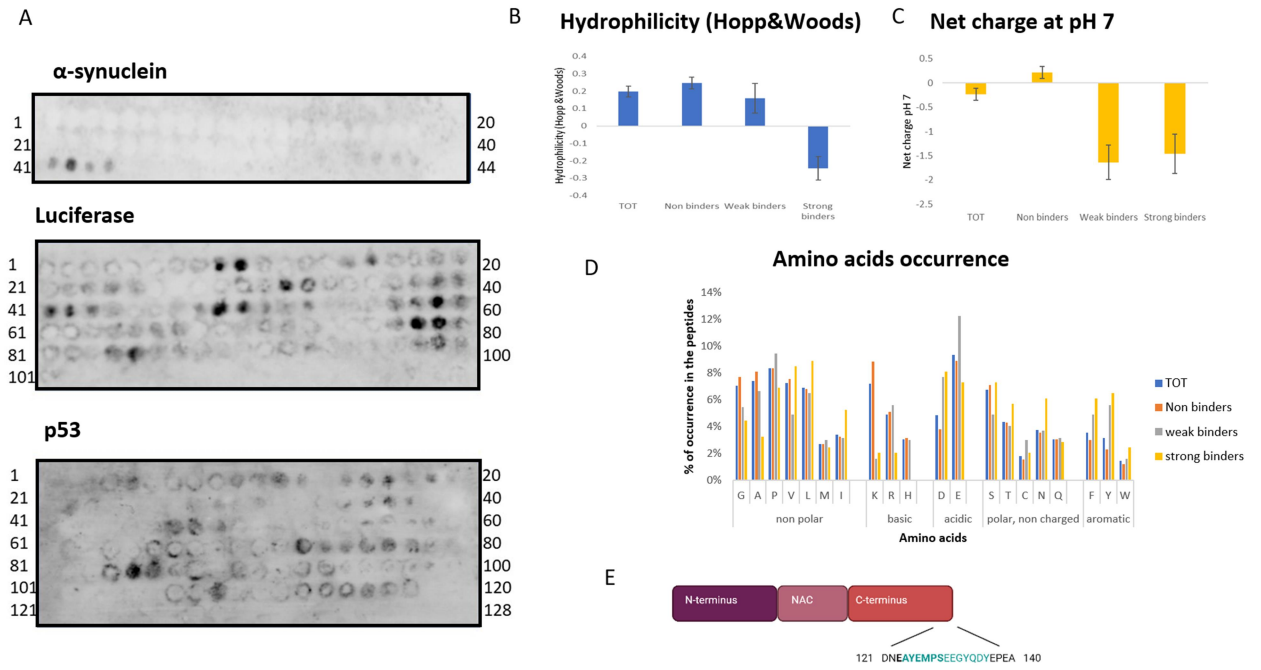


Figure 10: peptide libraries screening with FL DNAJB1. A) Images of the PVDF membranes representing the first blot of α-synuclein peptide library (upper panel), and the third blots of luciferase peptide library (medium panel), p53 library (lower panel). All the libraries were incubated with 1 µM of DNAJB1. We classified the strong binders peptides as the darkest spots, the weak binders as the grey spots or the spots with a dark border and the non-binders as the white spots or the spots which did not give any signal. B) Average hydrophilicity of the entire peptide libraries, of the non-binders, weak binders and strong binders for FL DNAJB1: the hydrophilicity diminishes while the binding becomes stronger. C) Net charge at pH 7 of the entire peptide libraries, of the non-binders, weak binders and strong binders for FL DNAJB1: FL DNAJB1 shows a preference for binding negatively charged peptides. D) Amino acids occurrence of the entire peptide libraries, in cyan, of the non-binders, in orange, weak binders, in grey, and strong binders, in yellow, for FL DNAJB1: the binder peptides show enrichment in negatively charged residues and aromatic residues, while binding to positively charged residues is avoided. E) Human α-synuclein domains (N-terminus, NAC and C-terminus) representation with a zoom on the last 20 residues of the C-terminal domain. The DNAJB1 binding site identified by our peptide library screening is highlighted in green in the sequence. The binding site identified by Wentink et al. (Nature, 2020), is highlighted in bold: the sequences found mostly overlap and the crucial Y125 residue is present in both the binding sequences found.

Of particular interest, the α-synuclein libraries shows a very clear binding pattern with a unique strong binder peptide (Figure 10-A, upper panel). It confirms the binding pattern described above (supplementary information), but, most importantly, it clearly identifies the DNAJB1 binding site due to the presence of the single strong binder peptide. This binding site (peptide 42, residues 124-137) almost perfectly overlaps with the DNAJB1 binding site on α-synuclein identified by NMR by Wentink et al., in 2020 (residues 123-129)(Figure 10-E). This result is meaningful since it confirms the important presence of acidic and aromatic amino acids in DNAJB1 binding motifs on the substrates and support the pivotal role of Y125 on α-synuclein for DNAJB1 binding described by Wentink et al., in 2020.

b. JD-GF

We then screened the p53 and luciferase peptide libraries for the binding with JD-GF. Unfortunately, we could not use the α-synuclein library anymore, since no signal was detected in previous replicate experiments with the full-length DNAJB1 protein. We expected a only partially shared binding specificity with the full-length protein, since the J domain does not present client engagement sites but interacts with HSP70 in the context of the whole chaperone machinery.

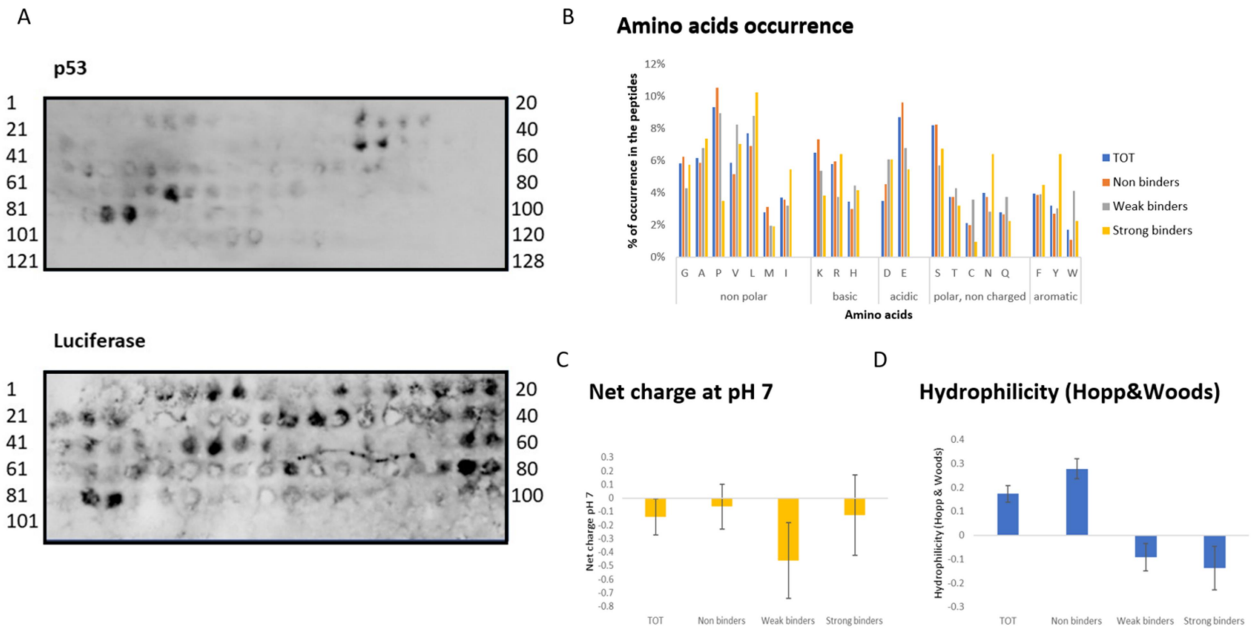


Figure 11: peptide libraries screening with JD-GF. A) Images of the PVDF membranes (third blots) of the p53 (upper panel) and luciferase (lower panel) libraries. All the libraries were incubated with 2 µM of JD-GF. We classified the strong binders peptides as the darkest spots, the weak binders as the grey spots or the spots with a dark border and the non-binders as the white spots or the spots which did not give any signal. B) Amino acids occurrence in the entire peptide libraries, in cyan, in the non-binders, in orange, weak binders, in grey, and strong binder peptides, in yellow. JD-GF shows a preference for aromatic residues, Leu and Ile. The binding of basic residues-enriched peptides increases, and the binding of acid amino acids-enriched peptides decreases, if compared to FL DNAJB1. C) Peptides net charge at pH 7 of the entire peptide libraries, of the non-binders, weak binders and strong binders for JD-GF: binders tend to have a lower charge, if compared to non-binders . D) Average hydrophilicity of the entire peptide libraries, of the non-binders, weak binders and strong binders for JD-GF: the protein tends to bind hydrophobic peptides.

Blotting results were visible on the third blot for the p53 and luciferase libraries. Using the p53 and luciferase libraries, we identified 43 weak binders and 24 strong binders for JD-GF in the 229 peptides screened. In fact, JD-GF showed the ability to discriminate between the peptide sequences, binding only some of them. Often clusters of binder peptides were found, indicating that a binding sequence was shared between the peptides. The blotting pattern seemed partially shared with the full-length protein (Figure 11-A). Nevertheless, analyzing the binders/non-binders peptides, some differences were found. In fact, despite the presence of aromatic residues, Valine, Leucine and Isoleucine in the binder sequences (which is shared with the full-length protein), more positively charged residues were bound and there was a less strong preference in binding peptides enriched in negatively charged amino acids (Figure 11-B). The net charge at pH 7 of the binders (in this case, in particular, considering the weak binders only) was still lower than the one shown by the non-binder peptides. Nevertheless, it was higher than the one showed screening the libraries with the full-length protein (Figure 11-C). Moreover, JD-GF showed the tendency of binding hydrophobic peptides resembling the FL DNAJB1. (Figure 11-D).

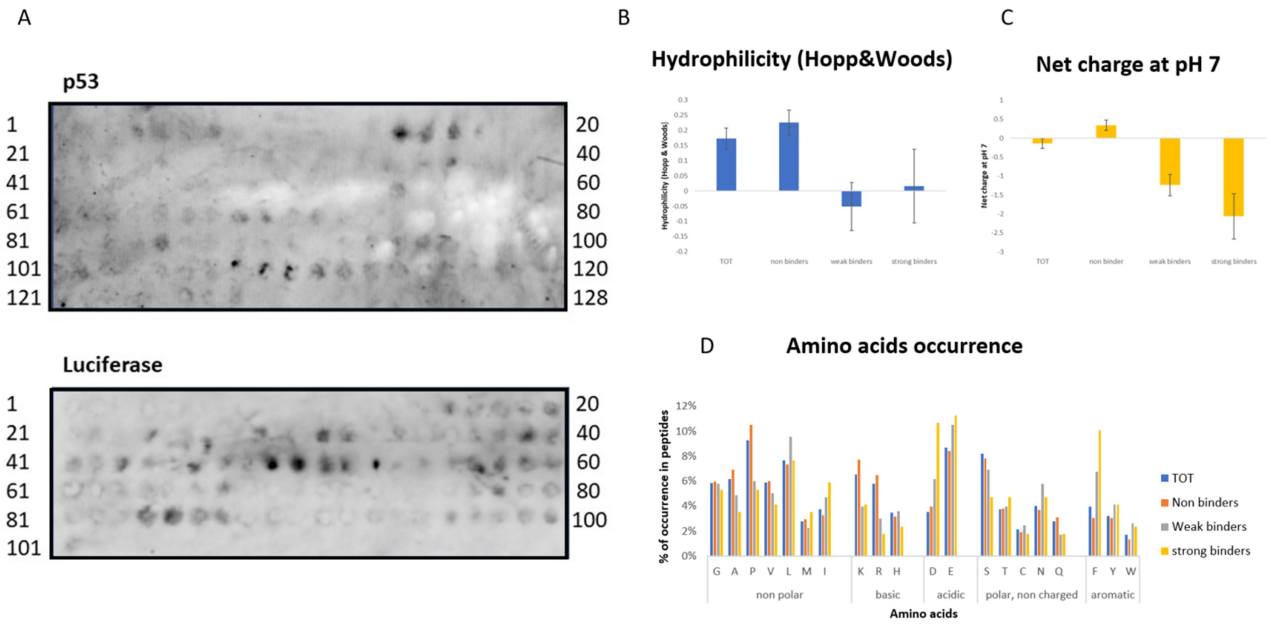


Figure 12: peptide libraries screening for CTDI-CTDII-DD. A) Images of the PVDF membranes (third blots) of the p53 (upper panel) and luciferase (lower panel) libraries. All the libraries were incubated with 1 µM of CTDI-CTDII-DD. We classified the strong binders peptides as the darkest spots, the weak binders as the grey spots or the spots with a dark border and the non-binders as the white spots or the spots which did not give any signal. B) Average hydrophilicity of the entire peptide libraries, of the non-binders, weak binders and strong binders for CTDI-CTDII-DD: the protein tends to bind less hydrophilic peptides. C) Peptides net charge at pH 7 of the entire peptide libraries, of the non-binders, weak binders and strong binders for CTDI-CTDII-DD: binder peptides (weak and strong) show overall a negative charge, while non-binders are slightly positively charged. D) Amino acids occurrence in the entire peptide libraries, in cyan, in the non-binders, in orange, weak binders, in grey, and strong binder peptides, in yellow. A strong preference for negatively charged and aromatic residues was detected.

A similar result was expected since the J domain is known not to interact with the substrate, but with the HSP70 hydrophobic linker region, and only speculations have been made regarding the possible presence of a substrate binding site in the GF-rich region.

c. CTDI-CTDII-DD

We screened p53 and luciferase libraries also with CTDI-CTDII-DD. We expect to obtain a similar binding pattern, compared to the FL DNAJB1, since CTDI-CTDII-DD has been reported to carry four client binding domains, two on the CTDs of each monomer within the dimer.

Spots were visible on the third blot for the p53 and luciferase libraries. The construct discriminated between the peptide sequences, binding only some of them. Often clusters of binder peptides were found, indicating that a binding sequence was shared between the peptides. The binding pattern resembled that of the full-length protein (Figure 12-A): we identified 48 weak and 19 strong binders. Similarly, the binder peptides show lower hydrophilicity (higher hydrophobicity) compared to the non-binders and whole peptides group (Figure 12-B). The net charge at pH 7 for the binders was negative (Figure 12-C), supporting the strong preference of CTDI-CTDII-DD in binding peptides enriched in acidic residues. Moreover, aromatic residues and Isoleucine residues were abundant in the weak and, in particular, in the strong binder peptides. CTDI-CTDII-DD showed a tendency of avoiding the binding to basic residues enriched peptides (Figure 12-D). These features are in line with the data found for the FL DNAJB1

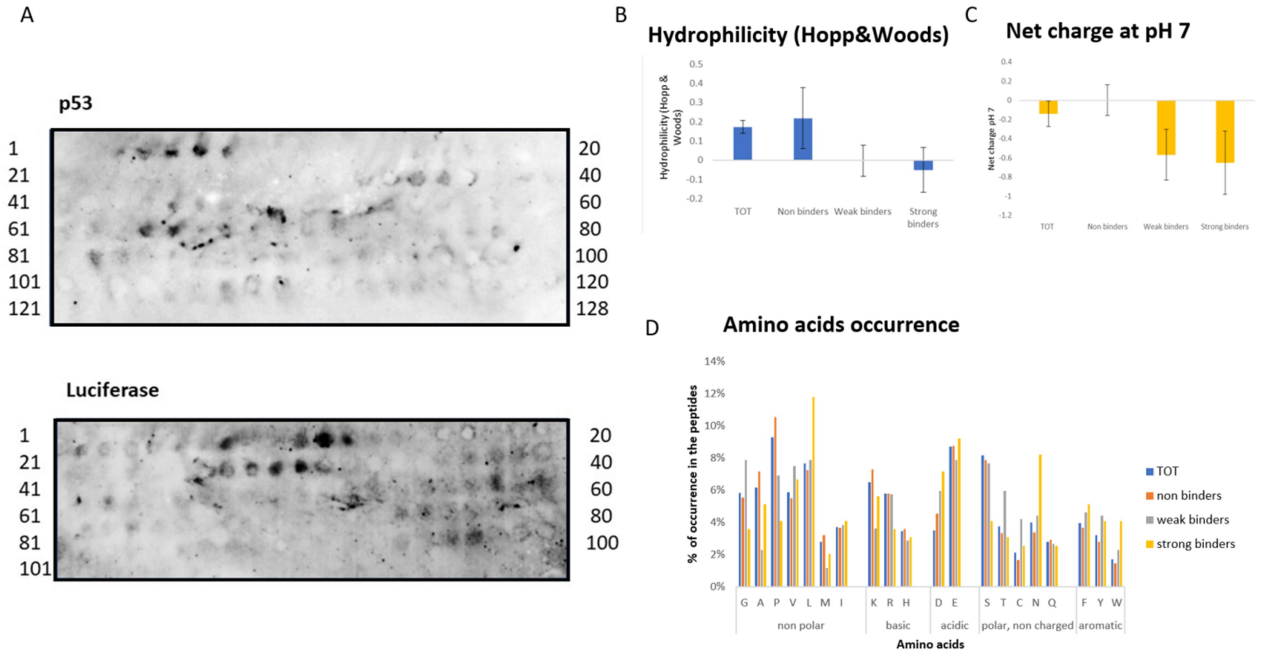


Figure 13: peptide libraries screening for CTDI-CTDII. A) Images of the PVDF membranes (third blots) of the p53 (upper panel) and luciferase (lower panel) libraries. All the libraries were incubated with 1 µM of CTDI-CTDII. We classified the strong binders peptides as the darkest spots, the weak binders as the grey spots or the spots with a dark border and the non-binders as the with spots or the spots which did not give any signal. B) Average hydrophilicity of the entire peptide libraries, of the non-binders, weak binders and strong binders for CTDI-CTDII: the protein tends to bind the most hydrophobic peptides . C) Peptides net charge at pH 7 of the entire peptide libraries, of the non-binders, weak binders and strong binders for CTDI-CTDII: binder peptides (weak and strong) show overall a negative charge, while non-binders are slightly positively charged. D) Amino acids occurrence in the entire peptide libraries, in cyan, in the non-binders, in orange, weak binders, in grey, and strong binder peptides, in yellow. A strong preference for negatively charged and aromatic residues, Leu, Ile and Asn was detected.

Spots were visible on the third blot for the p53 and luciferase libraries. The construct discriminated between the peptide sequences, binding only some of them. Often clusters of binder peptides were found, indicating that a binding sequence was shared between the peptides. The binding pattern resembled that of the full-length protein (Figure 12-A): we identified 48 weak and 19 strong binders. Similarly, the binder peptides show lower hydrophilicity (higher hydrophobicity) compared to the non-binders and whole peptides group (Figure 12-B). The net charge at pH 7 for the binders was negative (Figure 12-C), supporting the strong preference of CTDI-CTDII-DD in binding peptides enriched in acidic residues. Moreover, aromatic residues and Isoleucine residues were abundant in the weak and, in particular, in the strong binder peptides. CTDI-CTDII-DD showed a tendency of avoiding the binding to basic residues enriched peptides (Figure 12-D). These features are in line with the data found for the FL DNAJB1.

It is meaningful that the construct carrying the two substrate binding domains shows a similar pattern, compared to the full-length protein; in fact, this demonstrates that the selectivity in the substrate binding of FL DNAJB1 is dictated by CTDI and CTDII.

d. CTDI-CTDII

We screened p53 and luciferase libraries for CTDI-CTDII binding. In fact, we wanted to investigate the impact of the dimers formation in the substrate binding comparing the binding specificity and efficiency on the peptide libraries of CTDI-CTDII (monomeric) to the above presented CTDI-CTDII-DD.

Results were visible on the third blot for the p53 and luciferase libraries CTDI-CTDII discriminated between the peptide sequences, binding only some of them. Often clusters of binder peptides were found, indicating that a binding sequence was shared between the peptides. Through our libraries screening, we identified 40 weak binder peptides and 15 strong binders. The binding pattern resembles both the FL DNAJB1 and CTDI-CTDII-DD ones (Figure 13-A). In fact, the weak binders and the strong binders were less hydrophilic than the non-binders and showed lower hydrophilicity compared to average hydrophilicity of the whole libraries (Figure 13-B). We confirmed the preference of the substrate binding domains to engage negatively charged peptides, both in the case of weak and strong binder peptides (Figure 13-C). In fact, we identified a marked tendency of binding peptides enriched in acid and aromatic amino acids, Leucine and Isoleucine, with an overall behavior that resembles the one showed by CTDI-CTDII-DD and the full-length protein (Figure 13-D).

These data are noteworthy, since they confirm, once again that the specificity for the substrate binding is strongly directed by the two C-terminal domains of DNAJB1. Unfortunately, it was not possible to detect differences in affinity between CTDI-CTDII and CTDI-CTDII-DD using the libraries screening.

e. His-SUMO-CTDI and His-SUMO-CTDII

We screened p53 and luciferase libraries also for His-SUMO-CTDI and His-SUMO-CTDII constructs to identify possible differences in their substrate specificity. Unfortunately, in both the cases, we did not obtain any interpretable signals (Figure 14-A and B).

Multiple reasons could have concurred in causing this unsuccessful result: we first hypothesize that the libraries we were using were exhausted: in fact, a limited amount of experiments (circa 8-10) could be run on every library. Furthermore, it could be possible that the affinity showed by the single C-terminal domains for the peptides is not high enough to give detectable binding. In this case, using protein engineering techniques to produce obliged homodimers of CTDI and CTDII may, at least partially, solve the problem. Moreover, it is possible that CTDI and CTDII were not correctly folded in the tagged constructs we tested and, therefore, were not able to bind the libraries correctly. This hypothesis may be tested using circular dichroism techniques to check the folded structures. In the case of the presence of correctly folded β-barrel domains, extra attention would be required to repeat the experiments: the binding to the peptide libraries may be influenced by the presence of the tag, which is significantly large, compared to CTDI and CTDII. A negative control experiment, screening the libraries with His-SUMO and using anti-His tag antibodies would be required. We also demonstrated that, His-SUMO-CTDI detection through anti-CTDI-CTDII polyclonal antibodies (Figure 9-E) was not completely efficient. Thus, detection through anti-His antibodies may be required to increase and improve the signal.

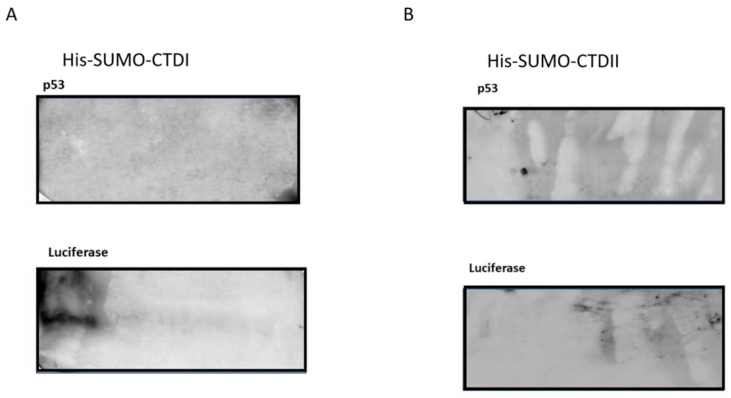


Figure 14: Peptide libraries screening for His-SUMO-CTDI and His-SUMO-CTDII. A) His-SUMO-CTDI screening of p53 library (upper panel) and luciferase (lower panel). All the libraries were incubated with 2 µM of protein. Both the images represents the third blot on PVDF membranes. Unfortunately no signal was detected. B) His-SUMO-CTDII screening of p53 library (upper panel) and luciferase (lower panel). Both the images represents the third blot on PVDF membranes. All the libraries were incubated with 2 µM of protein. Two spots could be seen on the p53 image, but it has not been possible to detect whether they were unspecific or not and which peptides they represented. Unfortunately no signal was detected on the luciferase library screening.

4. **Discussion**

We performed peptide libraries screening with full-length DNAJB1 and JD-GF, CTDI-CTDII-DD, CTDI-CTDII, His-SUMO-CTDI and His-SUMO-CTDII to identify binding motifs recognized by DNAJB1 and its client binding domains.

Overall, our data of the whole peptide libraries screening identified for the full-length DNAJB1 and the constructs carrying the substrate binding domains a preference in binding peptides enriched in acidic and aromatic residues. The weak and strong binders showed, generally, a negative charge, due to the presence of several acidic residues and the lack or scarcity of basic amino acids in the binder peptides and presented hydrophobicity features. The binder sequences we identified tended to be less hydrophilic (more hydrophobic) than non-binder ones and this is in line with slightly higher presence of hydrophobic residues such as Leucine, Valine and Isoleucine in the binders compared to the non-binder sequences. These data are in line with the substrate binding pattern identified by NMR in a recent work (Wentink et al., 2020).

Nevertheless, a different binding specificity emerges from our data comparing them to the peptide libraries screening run for the bacterial JDP DNAJ (Rüdiger et al., 2001) and HSP70 DNAK (Rüdiger et al., 1997). Rüdiger and co-workers in both their papers used both luciferase and p53 libraries: while the luciferase library they tested was bigger than ours, the p53 one resembled ours. However, the blots patterns and the bound peptides appear to be different. In particular, analyzing the sequences of the binder and non-binder peptides, we found out that even if binding to peptides enriched in aromatic, Leucine and Isoleucine residues was found to be a shared feature for DNAJB1 (FL, CTDI-CTDII, CTDI-CTDII-DD) and DNAJ (Rüdiger et al., 2001) and DNAJB1 and DNAK (Rüdiger et al., 1997), DNAJ and DNAK binding was disfavored to sequences enriched in negatively charged residues (Rüdiger et al., 1997; Rüdiger et al., 2001), while DNAJB1 (FL, CTDI-CTDII, CTDI-CTDII-DD) strongly prefers to interact with them. On the other hand, DNAJ already shows less disfavor for acidic amino acids (Rüdiger et al., 2001), if compared to DNAK (Rüdiger et al., 1997). Moreover, DNAK showed a preference for binding basic residues (Rüdiger et al., 1997, Wentink et al., 2020), which were disfavored by DNAJB1(FL, CTDI-CTDII, CTDI-CTDII-DD) . The presence of different binding specificities for HSP70 and J domain proteins is not surprising, in fact it has been demonstrated that they present different binding sites on α-synuclein analyzing DNAJB1 and HSP70 (Wentink et al., 2020). On the other hand, the differences in DNAJ and DNAJB1 (FL, CTDI-CTDII, CTDI-CTDII-DD) binding selectivity are quite puzzling. Differences may arise from the different JDPs classes they represent (DNAJ is a class A JDP), or from differences between the prokaryotic and eukaryotic systems. In the former case, the data we presented would be meaningful representing a first step in understanding the differences in substrates targeting for class A and B JDPs. Elucidations may come from peptide libraries screening with a human class A JDP. In fact, we strongly believe that peptide libraries screening with DNAJA2 or DNAJA1 (two of the major class A human JDPs) would be of interest: a comparison between our data obtained with DNAJB1 and data from a class A JDP (and its subdomains) may help in explaining the differences different classes have in binding specificities and cellular activities (Kampinga and Craig, 2010). On the other hand screening of the peptide libraries with different class B JDPs may be meaningful to elucidate better the features shared by the single protein class.

Unfortunately, we have not been able to detect differences in the binding strength or affinity comparing the results obtained with CTDI-CTDII and CTDI-CTDII-DD. Our hypothesis is that the presence of four binding sites, by means of avidity, strongly stabilizes the interaction with substrates, but using the peptide libraries screening we have not been able to detect this phenomenon. FRET or other techniques monitoring affinity, association and dissociation constants, may be more efficient for this purpose.

Our data from the libraries screening with JD-GF showed different binding specificities if compared to DNAJB1 FL, CTDI-CTDII and CTDI-CTDII-DD. This is not surprising, since the J domain is known to interact with the hydrophobic HSP70 NBD-SBD linker (Kytik et al., 2018). The linker is enriched in Leucine and Aspartic Acid residues (Chakafana et al., 2019), which are abundant in the peptides bound by JD-GF. Thus, our blots may represent this interaction, which might bias the whole libraries screening with JD-GF. However, it has been suggested that an additional substrate binding site may be present in the GF-rich region (Kampinga et al., 2019). To test this hypothesis, an efficient solution may be producing a DNAJB1 construct carrying only the GF region and screening the peptide libraries with it, comparing, eventually , the results with the ones obtained for DNAJB1 FL, CTDI-CTDII, CTDI-CTDII-DD.

Generally, in all the libraries which gave signals, it results difficult to identified a precise length in the substrate binding sequence. In fact, different lengths of binders stretches (meaning that binding occurs to different amounts of consequent peptides) could be seen. From the α-synuclein library screening we detected a single, unique binding site. This is remarkable: in our α-synuclein peptide libraries screening, we confirmed the presence of a binding site for DNAJB1 on α-synuclein C-terminus. Our data suggest that the binding site is comprised between residues 124 and 137: NMR data recently published on *Nature* (Wentink et al., 2020) identified the DNAJB1 binding site on α-synuclein between residues 123 and 129, emphasizing the importance of Y125 in mediating the interaction.

To use the libraries, we efficiently purified and determined the oligomeric state of DNAJB1 full-length and of the subdomains constructs JD-GF, CTDI-CTDII-DD and CTDI-CTDII. While JD-GF and CTDI-CTDII, lacking the dimerization domain, eluted from the gel filtration column as monomers, as expected, we observed the formation of higher molecular weight assemblies for FL DNAJB1 and CTDI-CTDII-DD. This is not surprising considering the sticky behavior JDPs often show. Nevertheless, smaller fractions of proteins presenting the expecting molecular weight of a dimer were observed.

Unfortunately, an efficient purification of CTDI and CTDII was not obtained. From our data (supplementary information) we can speculate that pure proteins may be obtained using an active Ulp1 able to completely remove the His-SUMO tag followed by cationic exchange chromatography. Further analyses determining the correct folding of the two domains will be required before proceeding with the experiments. In fact, we could not detect any specific signal on our libraries screening with the constructs carrying CTDI or CTDII. The peptide libraries represent an useful but fragile tool, which can be screened only a very limited amount of times and this may be why no specific signal was detected for CTDI and CTDII. This information would be crucial to understand how JDPs engage substrates. It has been reported that substrates mainly engage CTDII (Faust et al.,2020), or it has been speculated that CTDI and CTDII might have different specificities (Jiang et al., 2019). To elucidate the precise role of each CTD, it is crucial to efficiently purify them (even with a tag) and screen the peptide libraries with them. Building engineered CTDI and CTDII homodimers may represent an efficient solution to increase their affinity for the peptide libraries and obtain clearer data.

Lastly, we demonstrated that only the full-length DNAJB1 supports efficient α-synuclein fibrils disaggregation, while no disaggregation was detected adding either CTDI-CTDII, CTDI-CTDII-DD or JD-GF to HSP70, APG2 and ATP to the fibrils. This result is in line with the results published by Wentink et al. and Faust et al., in 2020. In fact, Faust and coworkers demonstrated that no binding to HSP70 was detected by investigating the interactions between DNAJB1 JD-GF and HSP70 by NMR. This phenomenon is dependent on the presence of helix V in the GF-rich region which inhibits the HSP70 binding, covering its interaction site on the J-domain and only the interaction of DNAJB1 CTDI with HSP70 C-terminal peptide EEVD releases

the helix V mediated inhibition (Faust et al., 2020). Therefore, we speculate that our JD-GF construct, lacking the second interaction site (CTDI) for HSP70, is not able to unlock the release of helix V mediated inhibition. Since the J domain-HSP70 interaction is missing, we hypothesize that JD-GF does not stimulate, through its J domain, the ATPase activity of HSP70. HSP70, in fact, show poor intrinsic ATPase activity, even in presence of the substrate, in absence of JDPs, and the absence of ATPase activity stimulation strongly impacts the HSP70 activity, which becomes mainly conceiving aggregation prevention (Kityk et al., 2018). It is unlikely that the absence of the C-terminal substrate binding domains in the JD-GF construct strongly affects the targeting function of DNAJB1 since a recent study demonstrated that the absence of DNAJB1 binding site on α-synuclein fibrils does not prevent HSP70 from fully interacting with fibrils (Wentink et al., 2020). Further analyses analyzing the behavior of JD-GF in the presence of HSP70 and fibrils may be required. On the other hand, the absence of the J domain mediated HSP70 conformational changes described in Wu et al., 2020, may represent the reason why we did not detect α-synuclein fibrils disaggregation in the reactions carried on with CTDI-CTDII or CTDI-CTDII-DD instead of the full-length DNAJB1: the intrinsic ATPase activity of HSP70 would be not sufficient to actively stimulate disaggregation. Moreover, it has been demonstrated that the presence of helix V is required for disaggregation (Faust et al., 2020), and in both the C-terminal subdomains constructs it is missing. Also, the lack of J domain may influence HSP70 binding to the substrate (Wentink et al., 2020). Moreover, it has been demonstrated that substrates and HSP70 C-terminal sequence EEVD can interact simultaneously with DNAJB1 CTDI (Faust et al., 2020): it is tempting to speculate that binding of HSP70 might occur in the presence of CTDI-CTDII-DD and CTDI-CTDII, which may also engage the substrate: testing this hypothesis, for instance using NMR approaches, would be meaningful.

.

5. **Conclusion**

In conclusion, our data demonstrate that the sequence DNAJB1 recognizes through the substrate binding domains CTDI and CTDII on its clients is enriched in acidic and aromatic amino acids, presents low hydrophilicity and negative charge at pH 7. These data are meaningful since they identify a target on the clients for their recognition and engagement by DNAJB1, and, in particular, by its C-terminal substrate binding domains CTDI and CTDII. The results are remarkable since they are also in line with data published in previous publications which identified the DNAJB1 binding site in α-synuclein by NMR. Nevertheless, the substrate specificity described for DNAJ, a bacterial, class A JDP, is slightly different, and this opens the question whether the differences which have been identified may arise from the different JDPs class DNAJ and DNAJB1 belong to. Thus, the data presented in this report will need to be confirmed for other JDPs from class B and their comparison with data from class A JDPs peptide libraries screening will be required to confirm this hypothesis. The use of other techniques, such as FRET or optical tweezers (which allow to test also structured or partially structured substrates), may be required to confirm the identified binding specificity, allowing a better comprehension of the chaperone system.

6. **References**

- Ayala Mariscal SM, Kirstein J. J-domain proteins interaction with neurodegenerative disease-related proteins. Exp Cell Res. 2021 Feb 15;399(2):112491. doi: 10.1016/j.yexcr.2021.112491. Epub 2021 Jan 16. PMID: 33460589.
- Beton JG, Monistrol J, Wentink A, Johnston EC, Roberts AJ, Bukau BG, Hoogenboom BW, Saibil HR. Cooperative amyloid fibre binding and disassembly by the Hsp70 disaggregase. EMBO J. 2022 Jun 13:e110410. doi: 10.15252/embj.2021110410. Epub ahead of print. PMID: 35698800.
- Borges JC, Fischer H, Craievich AF, Ramos CH. Low resolution structural study of two human HSP40 chaperones in solution. DJA1 from subfamily A and DJB4 from subfamily B have different quaternary structures. J Biol Chem. 2005 Apr 8;280(14):13671-81. doi: 10.1074/jbc.M408349200. Epub 2005 Jan 20. PMID: 15661747.
- Boysen M, Kityk R, Mayer MP. Hsp70- and Hsp90-Mediated Regulation of the Conformation of p53 DNA Binding Domain and p53 Cancer Variants. Mol Cell. 2019 May 16;74(4):831-843.e4. doi: 10.1016/j.molcel.2019.03.032. Epub 2019 Apr 23. PMID: 31027880.
- Brás IC, Xylaki M, Outeiro TF. Mechanisms of alpha-synuclein toxicity: An update and outlook. Prog Brain Res. 2020;252:91-129. doi: 10.1016/bs.pbr.2019.10.005. Epub 2019 Nov 23. PMID: 32247376.
- Chakafana G, Zininga T, Shonhai A. The Link That Binds: The Linker of Hsp70 as a Helm of the Protein's Function. Biomolecules. 2019 Sep 27;9(10):543. doi: 10.3390/biom9100543. PMID: 31569820; PMCID: PMC6843406
- Cheetham ME, Caplan AJ. Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. Cell Stress Chaperones. 1998 Mar;3(1):28-36. doi: 10.1379/1466-1268(1998)003<0028:sfaeod>2.3.co;2. PMID: 9585179; PMCID: PMC312945.
- Duennwald ML, Echeverria A, Shorter J. Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. PLoS Biol. 2012;10(6):e1001346. doi: 10.1371/journal.pbio.1001346. Epub 2012 Jun 19. PMID: 22723742; PMCID: PMC3378601.
- Faust O, Abayev-Avraham M, Wentink AS, Maurer M, Nillegoda NB, London N, Bukau B, Rosenzweig R. HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. Nature. 2020 Nov;587(7834):489-494. doi: 10.1038/s41586-020-2906-4. Epub 2020 Nov 11. PMID: 33177718.
- Flaherty KM, DeLuca-Flaherty C, McKay DB. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature. 1990 Aug 16;346(6285):623-8. doi: 10.1038/346623a0. PMID: 2143562.
- Franco A, Gracia P, Colom A, Camino JD, Fernández-Higuero JÁ, Orozco N, Dulebo A, Saiz L, Cremades N, Vilar JMG, Prado A, Muga A. All-or-none amyloid disassembly via chaperone-triggered fibril unzipping favors clearance of α-synuclein toxic species. Proc Natl Acad Sci U S A. 2021 Sep 7;118(36):e2105548118. doi: 10.1073/pnas.2105548118. PMID: 34462355; PMCID: PMC8433526.
- Gao X, Carroni M, Nussbaum-Krammer C, Mogk A, Nillegoda NB, Szlachcic A, Guilbride DL, Saibil HR, Mayer MP, Bukau B. Human Hsp70 Disaggregase Reverses Parkinson's-Linked α-Synuclein Amyloid Fibrils. Mol Cell. 2015 Sep 3;59(5):781-93. doi: 10.1016/j.molcel.2015.07.012. Epub 2015 Aug 20. PMID: 26300264; PMCID: PMC5072489.
- Greene MK, Maskos K, Landry SJ. Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proc Natl Acad Sci U S A. 1998 May 26;95(11):6108-13. doi: 10.1073/pnas.95.11.6108. PMID: 9600925; PMCID: PMC27593.

- Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A. 1981 Jun;78(6):3824-8. doi: 10.1073/pnas.78.6.3824. PMID: 6167991; PMCID: PMC319665.
- Hu J, Wu Y, Li J, Qian X, Fu Z, Sha B. The crystal structure of the putative peptide-binding fragment from the human Hsp40 protein Hdj1. BMC Struct Biol. 2008 Jan 22;8:3. doi: 10.1186/1472-6807-8-3. PMID: 18211704; PMCID: PMC2254625.
- Jiang Y, Rossi P, Kalodimos CG. Structural basis for client recognition and activity of Hsp40 chaperones. Science. 2019 Sep 20;365(6459):1313-1319. doi: 10.1126/science.aax1280. PMID: 31604242; PMCID: PMC7023980.
- Kampinga HH, Andreasson C, Barducci A, Cheetham ME, Cyr D, Emanuelsson C, Genevaux P, Gestwicki JE, Goloubinoff P, Huerta-Cepas J, Kirstein J, Liberek K, Mayer MP, Nagata K, Nillegoda NB, Pulido P, Ramos C, De Los Rios P, Rospert S, Rosenzweig R, Sahi C, Taipale M, Tomiczek B, Ushioda R, Young JC, Zimmermann R, Zylicz A, Zylicz M, Craig EA, Marszalek J. Function, evolution, and structure of J-domain proteins. Cell Stress Chaperones. 2019 Jan;24(1):7-15. doi: 10.1007/s12192-018-0948-4. Epub 2018 Nov 26. PMID: 30478692; PMCID: PMC6363617.
- Kampinga HH, Craig EA. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat Rev Mol Cell Biol. 2010 Aug;11(8):579-92. doi: 10.1038/nrm2941. Erratum in: Nat Rev Mol Cell Biol. 2010 Oct;11(10):750. PMID: 20651708; PMCID: PMC3003299.
- Karzai AW, McMacken R. A bipartite signaling mechanism involved in DnaJ-mediated activation of the Escherichia coli DnaK protein. J Biol Chem. 1996 May 10;271(19):11236-46. doi: 10.1074/jbc.271.19.11236. PMID: 8626673.
- Kelley WL. The J-domain family and the recruitment of chaperone power. Trends Biochem Sci. 1998 Jun;23(6):222-7. doi: 10.1016/s0968-0004(98)01215-8. PMID: 9644977.
- Kityk R, Kopp J, Mayer MP. Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. Mol Cell. 2018 Jan 18;69(2):227-237.e4. doi: 10.1016/j.molcel.2017.12.003. Epub 2017 Dec 28. PMID: 29290615.
- Li J, Qian X, Sha B. The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. Structure. 2003 Dec;11(12):1475-83. doi: 10.1016/j.str.2003.10.012. PMID: 14656432.
- Mantovani F, Collavin L, Del Sal G. Mutant p53 as a guardian of the cancer cell. Cell Death Differ. 2019 Jan;26(2):199-212. doi: 10.1038/s41418-018-0246-9. Epub 2018 Dec 11. PMID: 30538286; PMCID: PMC6329812.
- Mattoo RUH, Sharma SK, Priya S, Finka A, Goloubinoff P. Hsp110 is a bona fide chaperone using ATP to unfold stable misfolded polypeptides and reciprocally collaborate with Hsp70 to solubilize protein aggregates. J Biol Chem. 2013 Jul 19;288(29):21399-21411. doi: 10.1074/jbc.M113.479253. Epub 2013 Jun 4. PMID: 23737532; PMCID: PMC3774407.
- Mayer MP, Bukau B. Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci. 2005 Mar;62(6):670-84. doi: 10.1007/s00018-004-4464-6. PMID: 15770419; PMCID: PMC2773841.
- Mayer MP, Schröder H, Rüdiger S, Paal K, Laufen T, Bukau B. Multistep mechanism of substrate binding determines chaperone activity of Hsp70. Nat Struct Biol. 2000 Jul;7(7):586-93. doi: 10.1038/76819. PMID: 10876246.
- Meimaridou E, Gooljar SB, Chapple JP. From hatching to dispatching: the multiple cellular roles of the Hsp70 molecular chaperone machinery. J Mol Endocrinol. 2009 Jan;42(1):1-9. doi: 10.1677/JME-08-0116. Epub 2008 Oct 13. PMID: 18852216.

- Mogk A, Bukau B, Kampinga HH. Cellular Handling of Protein Aggregates by Disaggregation Machines. Mol Cell. 2018 Jan 18;69(2):214-226. doi: 10.1016/j.molcel.2018.01.004. PMID: 29351843.
- Nachman E, Wentink AS, Madiona K, Bousset L, Katsinelos T, Allinson K, Kampinga H, McEwan WA, Jahn TR, Melki R, Mogk A, Bukau B, Nussbaum-Krammer C. Disassembly of Tau fibrils by the human Hsp70 disaggregation machinery generates small seeding-competent species. J Biol Chem. 2020 Jul 10;295(28):9676-9690. doi: 10.1074/jbc.RA120.013478. Epub 2020 May 28. PMID: 32467226; PMCID: PMC7363153.
- Nillegoda NB, Bukau B. Metazoan Hsp70-based protein disaggregases: emergence and mechanisms. Front Mol Biosci. 2015 Oct 9;2:57. doi: 10.3389/fmolb.2015.00057. PMID: 26501065; PMCID: PMC4598581.
- Nillegoda NB, Kirstein J, Szlachcic A, Berynskyy M, Stank A, Stengel F, Arnsburg K, Gao X, Scior A, Aebersold R, Guilbride DL, Wade RC, Morimoto RI, Mayer MP, Bukau B. Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. Nature. 2015 Aug 13;524(7564):247-51. doi: 10.1038/nature14884. Epub 2015 Aug 5. PMID: 26245380; PMCID: PMC4830470.
- Nillegoda NB, Wentink AS, Bukau B. Protein Disaggregation in Multicellular Organisms. Trends Biochem Sci. 2018 Apr;43(4):285-300. doi: 10.1016/j.tibs.2018.02.003. Epub 2018 Feb 28. PMID: 29501325.
- Pemberton S, Madiona K, Pieri L, Kabani M, Bousset L, Melki R. Hsc70 protein interaction with soluble and fibrillar alpha-synuclein. J Biol Chem. 2011 Oct 7;286(40):34690-9. doi: 10.1074/jbc.M111.261321. Epub 2011 Aug 10. PMID: 21832061; PMCID: PMC3186418.
- Perales-Calvo J, Muga A, Moro F. Role of DnaJ G/F-rich domain in conformational recognition and binding of protein substrates. J Biol Chem. 2010 Oct 29;285(44):34231-9. doi: 10.1074/jbc.M110.144642. Epub 2010 Aug 20. PMID: 20729526; PMCID: PMC2962521.
- Qi M, Zhang J, Zeng W, Chen X. DNAJB1 stabilizes MDM2 and contributes to cancer cell proliferation in a p53-dependent manner. Biochim Biophys Acta. 2014 Jan;1839(1):62-9. doi: 10.1016/j.bbagrm.2013.12.003. Epub 2013 Dec 19. PMID: 24361594.
- Qiu XB, Shao YM, Miao S, Wang L. The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. Cell Mol Life Sci. 2006 Nov;63(22):2560-70. doi: 10.1007/s00018-006-6192-6. PMID: 16952052.
- Rampelt H, Kirstein-Miles J, Nillegoda NB, Chi K, Scholz SR, Morimoto RI, Bukau B. Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. EMBO J. 2012 Nov 5;31(21):4221-35. doi: 10.1038/emboj.2012.264. Epub 2012 Sep 18. PMID: 22990239; PMCID: PMC3492728
- Rodriguez F, Arsène-Ploetze F, Rist W, Rüdiger S, Schneider-Mergener J, Mayer MP, Bukau B. Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones. Mol Cell. 2008 Nov 7;32(3):347-58. doi: 10.1016/j.molcel.2008.09.016. PMID: 18995833.
- Rosenzweig R, Nillegoda NB, Mayer MP, Bukau B. The Hsp70 chaperone network. Nat Rev Mol Cell Biol. 2019 Nov;20(11):665-680. doi: 10.1038/s41580-019-0133-3. PMID: 31253954..
- Rüdiger S, Germeroth L, Schneider-Mergener J, Bukau B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. EMBO J. 1997 Apr 1;16(7):1501-7. doi: 10.1093/emboj/16.7.1501. PMID: 9130695; PMCID: PMC1169754.
- Rüdiger S, Schneider-Mergener J, Bukau B. Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. EMBO J. 2001 Mar 1;20(5):1042-50. doi: 10.1093/emboj/20.5.1042. PMID: 11230128; PMCID: PMC145471.

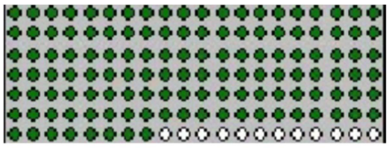
- Schneider MM, Gautam S, Herling TW, Andrzejewska E, Krainer G, Miller AM, Trinkaus VA, Peter QAE, Ruggeri FS, Vendruscolo M, Bracher A, Dobson CM, Hartl FU, Knowles TPJ. The Hsc70 disaggregation machinery removes monomer units directly from α-synuclein fibril ends. Nat Commun. 2021 Oct 14;12(1):5999. doi: 10.1038/s41467-021-25966-w. PMID: 34650037; PMCID: PMC8516981.
- Scior A, Buntru A, Arnsburg K, Ast A, Iburg M, Juenemann K, Pigazzini ML, Mlody B, Puchkov D, Priller J, Wanker EE, Prigione A, Kirstein J. Complete suppression of Htt fibrilization and disaggregation of Htt fibrils by a trimeric chaperone complex. EMBO J. 2018 Jan 17;37(2):282-299. doi: 10.15252/embj.201797212. Epub 2017 Dec 6. Erratum in: EMBO J. 2021 Oct 1;40(19):e109413. PMID: 29212816; PMCID: PMC5770855.
- Sugito K, Yamane M, Hattori H, Hayashi Y, Tohnai I, Ueda M, Tsuchida N, Ohtsuka K. Interaction between hsp70 and hsp40, eukaryotic homologues of DnaK and DnaJ, in human cells expressing mutant-type p53. FEBS Lett. 1995 Jan 23;358(2):161-4. doi: 10.1016/0014-5793(94)01417-y. PMID: 7828728.
- Tittelmeier J, Sandhof CA, Ries HM, Druffel-Augustin S, Mogk A, Bukau B, Nussbaum-Krammer C. The HSP110/HSP70 disaggregation system generates spreading-competent toxic α-synuclein species. EMBO J. 2020 Jul 1;39(13):e103954. doi: 10.15252/embj.2019103954. Epub 2020 May 25. PMID: 32449565; PMCID: PMC7327497.
- Tomiczek B, Delewski W, Nierzwicki L, Stolarska M, Grochowina I, Schilke B, Dutkiewicz R, Uzarska MA, Ciesielski SJ, Czub J, Craig EA, Marszalek J. Two-step mechanism of J-domain action in driving Hsp70 function. PLoS Comput Biol. 2020 Jun 1;16(6):e1007913. doi: 10.1371/journal.pcbi.1007913. PMID: 32479549; PMCID: PMC7289447.
- Wentink A, Nussbaum-Krammer C, Bukau B. Modulation of Amyloid States by Molecular Chaperones. Cold Spring Harb Perspect Biol. 2019 Jul 1;11(7):a033969. doi: 10.1101/cshperspect.a033969. PMID: 30755450; PMCID: PMC6601462.
- Wentink AS, Nillegoda NB, Feufel J, Ubartaitė G, Schneider CP, De Los Rios P, Hennig J, Barducci A, Bukau B. Molecular dissection of amyloid disaggregation by human HSP70. Nature. 2020 Nov;587(7834):483-488. doi: 10.1038/s41586-020
- Wu S, Hong L, Wang Y, Yu J, Yang J, Yang J, Zhang H, Perrett S. Kinetics of the conformational cycle of Hsp70 reveals the importance of the dynamic and heterogeneous nature of Hsp70 for its function. Proc Natl Acad Sci U S A. 2020 Apr 7;117(14):7814-7823. doi: 10.1073/pnas.1914376117. Epub 2020 Mar 20. PMID: 32198203; PMCID: PMC7148561.-2904-6. Epub 2020 Nov 11. Erratum in: Nature. 2021 Jan;589(7841):E2. PMID: 33177717.
- Yu HY, Ziegelhoffer T, Craig EA. Functionality of Class A and Class B J-protein co-chaperones with Hsp70. FEBS Lett. 2015 Sep 14;589(19 Pt B):2825-30. doi: 10.1016/j.febslet.2015.07.040. Epub 2015 Aug 3. PMID: 26247431; PMCID: PMC4570866.
- Zhu X, Zhao X, Burkholder WF, Gragerov A, Ogata CM, Gottesman ME, Hendrickson WA. Structural analysis of substrate binding by the molecular chaperone DnaK. Science. 1996 Jun 14;272(5268):1606-14. doi: 10.1126/science.272.5268.1606. PMID: 8658133; PMCID: PMC5629921.

7. **Supplementary information**

More supplementary information and raw data may be required to T.L. Dang and Dr. A. Mogk.

1. Peptide libraries peptides pattern and sequences:

- p53

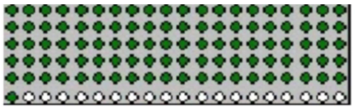


1 MEEPQSDPSVEPP	33 VPSQKTYQGSYGF	65 HLIRVEGNLRVEY	97 LRKKGEPHHELPP
2 PQSDPSVEPPLSQ	34 QKTYQGSYGFRLG	66 RVEGNLRVEYLDD	98 KGEPHHELPPGST
3 DPSVEPPLSQETF	35 YQGSYGFRLGFLH	67 GNLRVEYLDDRNT	99 PHHELPPGSTKRA
4 VEPPLSQETFSDL	36 SYGFRLGFLHSGT	68 RVEYLDDRNTFRH	100 ELPPGSTKRALPN
5 PLSQETFSDLWKL	37 FRLGFLHSGTAKS	69 YLDDRNTFRHSVV	101 PGSTKRALPNNTS
6 QETFSDLWKLLPE	38 GFLHSGTAKSVTC	70 DRNTFRHSVVVPY	102 TKRALPNNTSSSP
7 FSDLWKLLPENNV	39 HSGTAKSVTCTYS	71 TFRHSVVVPYEPP	103 ALPNNTSSSPQPK
8 LWKLLPENNVLSP	40 TAKSVTCTYSPAL	72 HSVVVPYEPPEVG	104 NNTSSSPQPKKKP
9 LLPENNVLSPLPS	41 SVTCTYSPALNKM	73 VVPYEPPEVGSDC	105 SSSPQPKKKPLDG
10 ENNVLSPLPSQAM	42 CTYSPALNKMFCQ	74 YEPPEVGSDCTTI	106 PQPKKKPLDGEYF
11 VLSPLPSQAMDDL	43 SPALNKMFCQLAK	75 PEVGSDCTTIHYN	107 KKKPLDGEYFTLQ
12 PLPSQAMDDLMLS	44 LNKMFCQLAKTCP	76 GSDCTTIHYNYMC	108 PLDGEYFTLQIRG
13 SQAMDDLMLSPDD	45 MFCQLAKTCPVQL	77 CTTIHYNYMCNSS	109 GEYFTLQIRGRER
14 MDDLMLSPDDIEQ	46 QLAKTCPVQLWVD	78 IHYNYMCNSSCMG	110 FTLQIRGRERFEM
15 LMLSPDDIEQWFT	47 KTCPVQLWVDSTP	79 NYMCNSSCMGGMN	111 QIRGRERFEMFRE
16 SPDDIEQWFTEDP	48 PVQLWVDSTPPPG	80 CNSSCMGGMNRRP	112 GRERFEMFRELNE
17 DIEQWFTEDPGPD	49 LWVDSTPPPGTRV	81 SCMGGMNRRPILT	113 RFEMFRELNEALE
18 QWFTEDPGPDEAP	50 DSTPPPGTRVRAM	82 GGMNRRPILTIIT	114 MFRELNEALELKD
19 TEDPGPDEAPRMP	51 PPPGTRVRAMAIY	83 NRRPILTIITLED	115 ELNEALELKDAQA
20 PGPDEAPRMPEAA	52 GTRVRAMAIYKQS	84 PILTIITLEDSSG	116 EALELKDAQAGKE
21 DEAPRMPEAAPPV	53 VRAMAIYKQSQHM	85 TIITLEDSSGNLL	117 ELKDAQAGKEPGG
22 PRMPEAAPPVAPA	54 MAIYKQSQHMTEV	86 TLEDSSGNLLGRN	118 DAQAGKEPGGSRA
23 PEAAPPVAPAPAA	55 YKQSQHMTEVVRR	87 DSSGNLLGRNSFE	119 AGKEPGGSRAHSS
24 APPVAPAPAAPTP	56 SQHMTEVVRRCPH	88 GNLLGRNSFEVRV	120 EPGGSRAHSSHLK
25 VAPAPAAPTPAAP	57 MTEVVRRCPHHER	89 LGRNSFEVRVCAC	121 GSRAHSSHLKSKK
26 APAAPTPAAPAPA	58 VVRRCPHHERCSD	90 NSFEVRVCACPGR	122 AHSSHLKSKKGQS
27 APTPAAPAPAPSW	59 RCPHHERCSDSDG	91 EVRVCACPGRDRR	123 SHLKSKKGQSTSR
28 PAAPAPAPSWPLS	60 HHERCSDSDGLAP	92 VCACPGRDRRTEE	124 KSKKGQSTSRHKK
29 PAPAPSWPLSSSV	61 RCSDSDGLAPPQH	93 CPGRDRRTEEENL	125 KGQSTSRHKKLMF
30 APSWPLSSSVPSQ	62 DSDGLAPPQHLIR	94 RDRRTEEENLRKK	126 STSRHKKLMFKTE
31 WPLSSSVPSQKTY	63 GLAPPQHLIRVEG	95 RTEEENLRKKGEP	127 RHKKLMFKTEGPD
32 SSSVPSQKTYQGS	64 PPQHLIRVEGNLR	96 EENLRKKGEPHHE	128 KKLMFKTEGPDSD

- α-synuclein

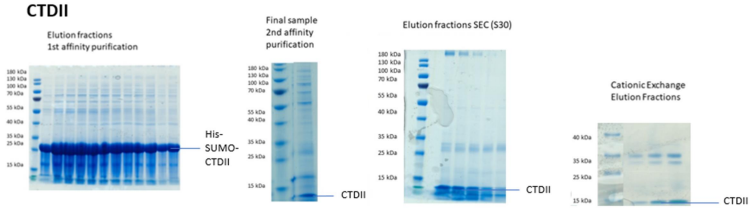
1 MDVFMKGLSKAKE	12 KEGVLYVGSKTKE	23 GGAVVTGVTAVAQ	34 LGKNEEGAPQEGI
2 FMKGLSKAKEGVV	13 VLYVGSKTKEGVV	24 VVTGVTAVAQKTV	35 NEEGAPQEGILED
3 GLSKAKEGVVAAA	14 VGSKTKEGVVHGV	25 GVTAVAQKTVEGA	36 GAPQEGILEDMPV
4 KAKEGVVAAAEKT	15 KTKEGVVHGVATV	26 AVAQKTVEGAGSI	37 QEGILEDMPVDPD
5 EGVVAAAEKTKQG	16 EGVVHGVATVAEK	27 QKTVEGAGSIAAA	38 ILEDMPVDPDNEA
6 VAAAEKTKQGVAE	17 VHGVATVAEKTKE	28 VEGAGSIAAATGF	39 DMPVDPDNEAYEM
7 AEKTKQGVAEAAG	18 VATVAEKTKEQVT	29 AGSIAAATGFVKK	40 VDPDNEAYEMPSE
8 TKQGVAEAAGKTK	19 VAEKTKEQVTNVG	30 IAAATGFVKKDQL	41 DNEAYEMPSEEGY
9 GVAEAAGKTKEGV	20 KTKEQVTNVGGAV	31 ATGFVKKDQLGKN	42 AYEMPSEEGYQDY
10 EAAGKTKEGVLYV	21 EQVTNVGGAVVTG	32 FVKKDQLGKNEEG	43 MPSEEGYQDYEPE
11 GKTKEGVLYVGSK	22 TNVGGAVVTGVTA	33 KDQLGKNEEGAPQ	44 PSEEGYQDYEPEA

- luciferase

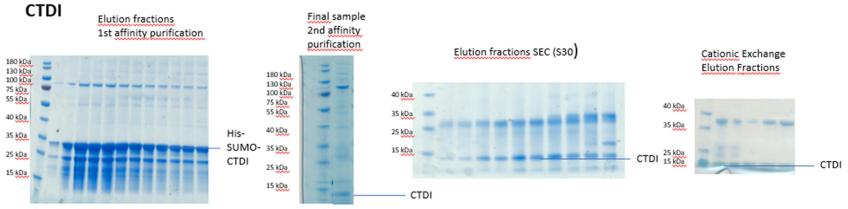


```
  1 MTSKVYDPELRKR      27 IGMGKSGKSGNGS      53 PDIEEDIALIKSE      79 EIVRNYNAYLRAS
  2 KVYDPELRKRMIT      28 GKSGKSGNGSYRL      54 EEDIALIKSEEGE      80 RNYNAYLRASHDL
  3 DPELRKRMITGPQ      29 GKSGNGSYRLLDH      55 IALIKSEEGEKMV      81 NAYLRASHDLPKM
  4 LRKRMITGPQWWA      30 GNGSYRLLDHYKY      56 IKSEEGEKMVLEN      82 LRASHDLPKMFIE
  5 RMITGPQWWARCK      31 SYRLLDHYKYLTE      57 EEGEKMVLENNFF      83 SHDLPKMFIESDP
  6 TGPQWWARCKQMN      32 LLDHYKYLTEWFK      58 EKMVLENNFFVET      84 LPKMFIESDPGFF
  7 QWWARCKQMNVLD      33 HYKYLTEWFKHLN      59 VLENNFFVETMLP      85 MFIESDPGFFSNA
  8 ARCKQMNVLDSFI      34 YLTEWFKHLNLPK      60 NNFFVETMLPSKI      86 ESDPGFFSNAIVE
  9 KQMNVLDSFINYY      35 EWFKHLNLPKKII      61 FVETMLPSKIMRK      87 PGFFSNAIVEGAK
 10 NVLDSFINYYDSE      36 KHLNLPKKIIFVG      62 TMLPSKIMRKLEP      88 FSNAIVEGAKKFP
 11 DSFINYYDSEKHA      37 NLPKKIIFVGHDW      63 PSKIMRKLEPEEF      89 AIVEGAKKFPNTE
 12 INYYDSEKHAENA      38 KKIIFVGHDWGAC      64 IMRKLEPEEFAAY      90 EGAKKFPNTEFVK
 13 YDSEKHAENAVIF      39 IFVGHDWGACLAF      65 KLEPEEFAAYLEP      91 KKFPNTEFVKVKG
 14 EKHAENAVIFLHG      40 GHDWGACLAFHYC      66 PEEFAAYLEPFKE      92 PNTEFVKVKGLHF
 15 AENAVIFLHGNAA      41 WGACLAFHYCYEH      67 FAAYLEPFKEKGE      93 EFVKVKGLHFSQE
 16 AVIFLHGNAASSY      42 CLAFHYCYEHQDR      68 YLEPFKEKGEVRR      94 KVKGLHFSQEDAP
 17 FLHGNAASSYLWR      43 FHYCYEHQDRIKA      69 PFKEKGEVRRPTL      95 GLHFSQEDAPDEM
 18 GNAASSYLWRHVV      44 CYEHQDRIKAVVH      70 EKGEVRRPTLSWP      96 FSQEDAPDEMGNY
 19 ASSYLWRHVVPHV      45 HQDRIKAVVHAES      71 EVRRPTLSWPREI      97 EDAPDEMGNYIKS
 20 YLWRHVVPHVEPV      46 RIKAVVHAESVVD      72 RPTLSWPREIPLV      98 PDEMGNYIKSFVE
 21 RHVVPHVEPVARC      47 AVVHAESVVDVIE      73 LSWPREIPLVKGG      99 MGNYIKSFVERVL
 22 VPHVEPVARCIIP      48 HAESVVDVIESWD      74 PREIPLVKGGKPD     100 YIKSFVERVLKNE
 23 VEPVARCIIPDLI      49 SVVDVIESWDEWP      75 IPLVKGGKPDVVE     101 IKSFVERVLKNEQ
 24 VARCIIPDLIGMG      50 DVIESWDEWPDIE      76 VKGGKPDVVEIVR
 25 CIIPDLIGMGKSG      51 ESWDEWPDIEEDI      77 GKPDVVEIVRNYN
 26 PDLIGMGKSGKSG      52 DEWPDIEEDIALI      78 DVVEIVRNYNAYL
```

2. **Purification attempts CTDII and CTDI**: affinity and reverse affinity purification were performed as described for the other constructs in materials and methods. SEC was performed with a Superdex 30 column, using the buffer previously described for the other constructs. Cationic exchange purification was performed with a Resource S column (Cytiva) with a salt concentration gradient ranging from 5 mM to 1 M KCl (in HEPES-KOH pH 7.5 50 mM, $MgCl_2$ 5 mM, glycerol 5%, β-mercaptoethanol 3 mM).
   a. CTDII

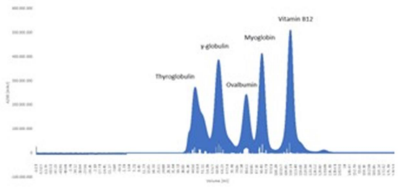


   b. CTDI



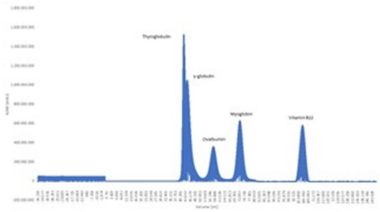
3. **Standard curves SEC Superdex200 and Superdex75.** The table below the curves shows the molecular weights of the proteins run in the standard curves.
   a. S200

S200 standard curve



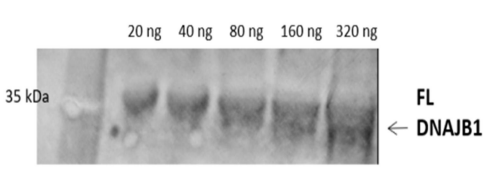
   b. S75

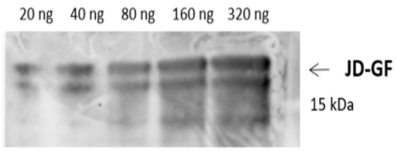
S75 standard curve

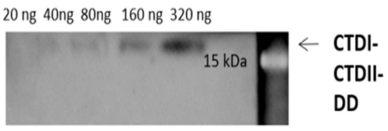


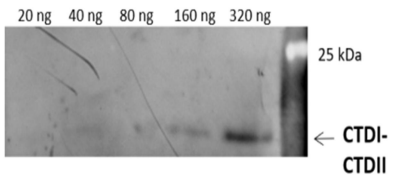
Gel filtration standard components	
Component	Molecular weight (Da)
Thyroglobulin	670000
γ-globulin	158000
Ovalbumin	44000
Myoglobin	17000
Vitamin B12	1350

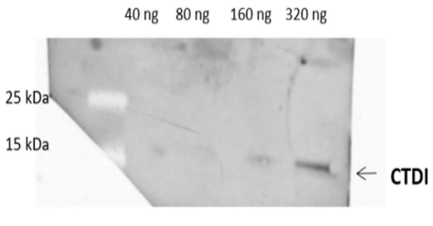
4. **Western blots**
   a. ENZO anti-DNAJB1 polyclonal antibodies, 1:1000 dilution

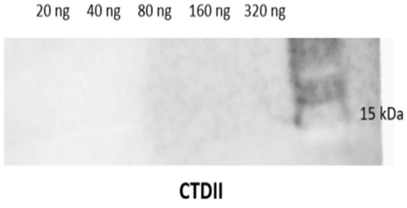




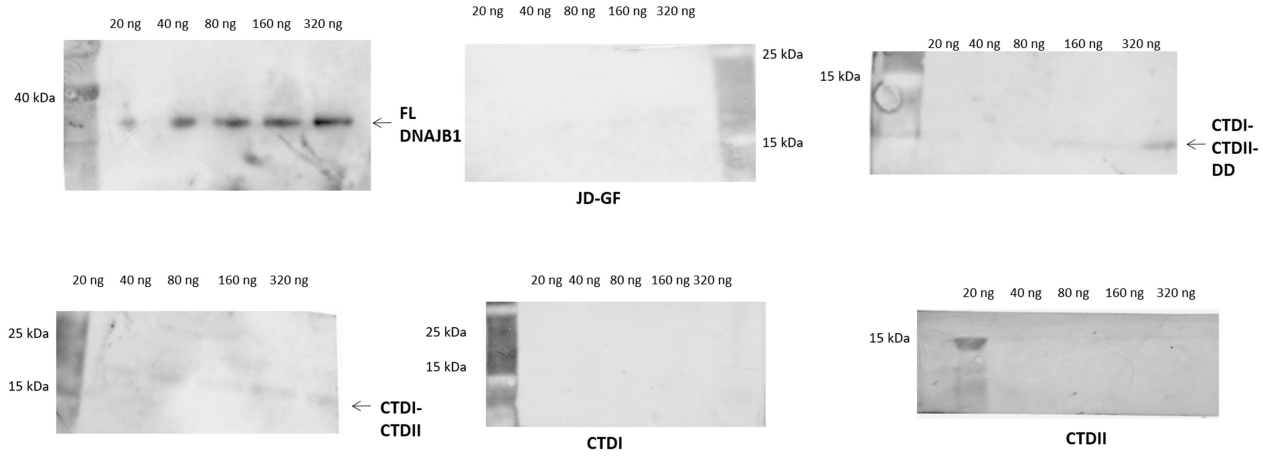




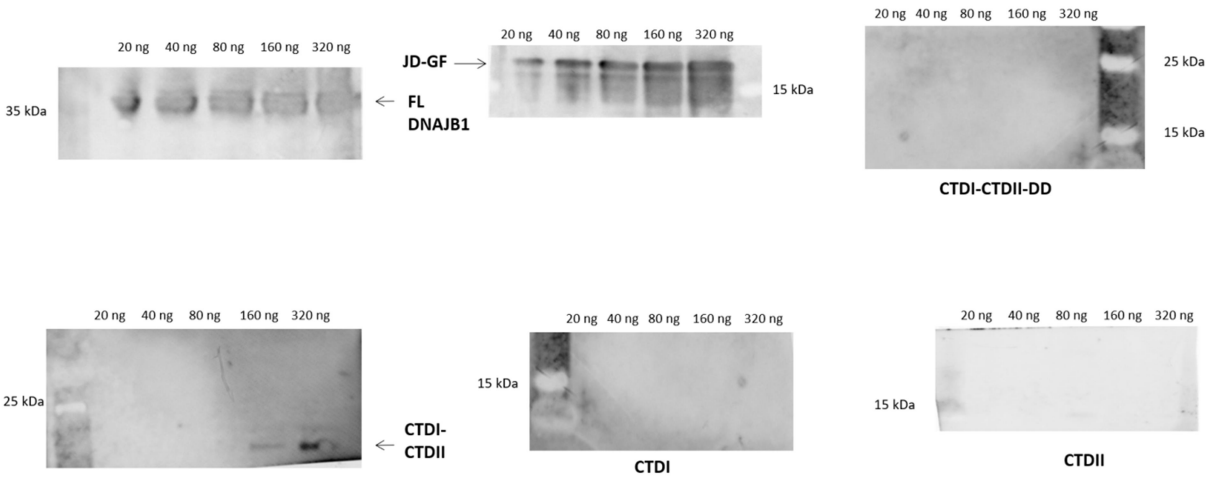




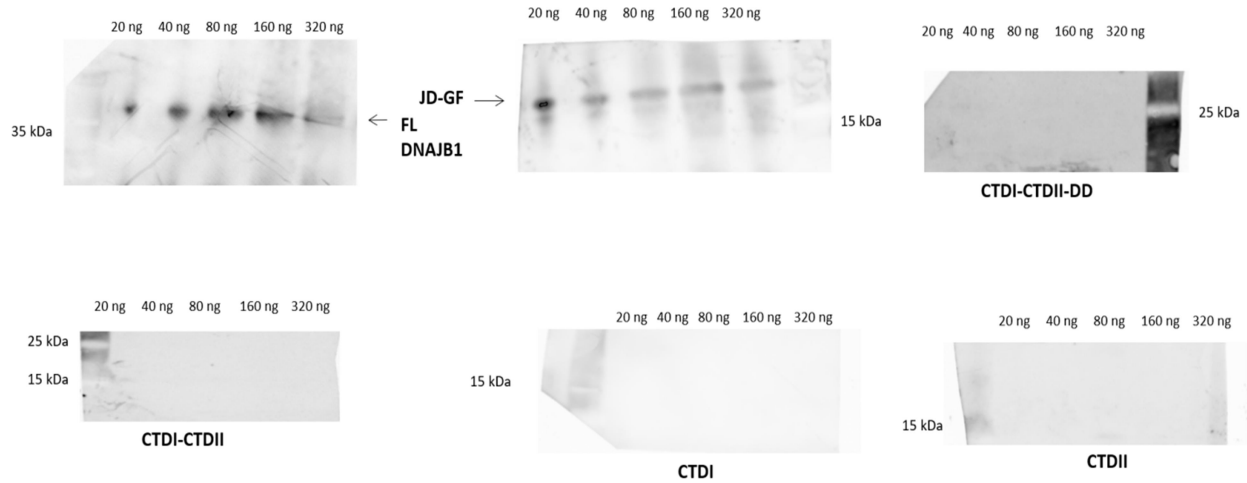
b. ENZO anti-DNAJB1 polyclonal antibodies 1:5000 dilution



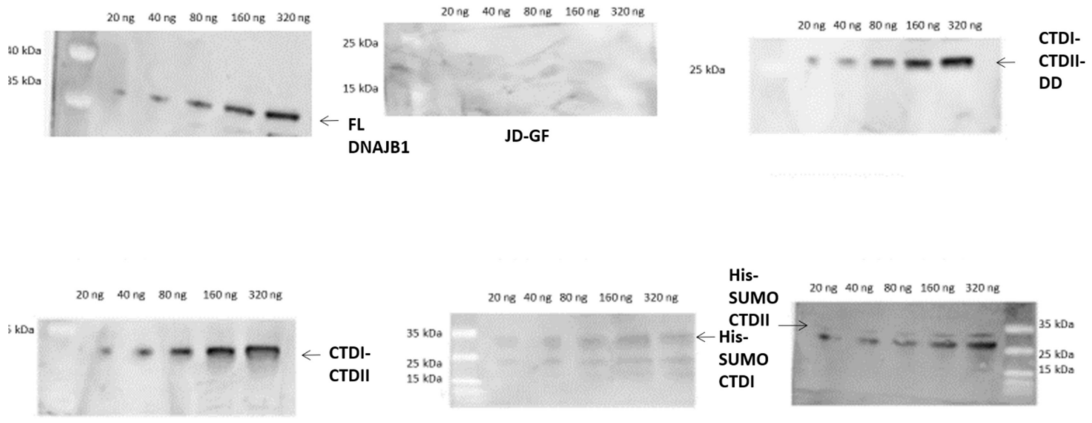
c. INVITROGEN anti-DNAJB1 polyclonal antibodies, 1:1000 dilution



d. INVITROGEN anti-DNAJB1 polyclonal antibodies, 1:5000 dilution



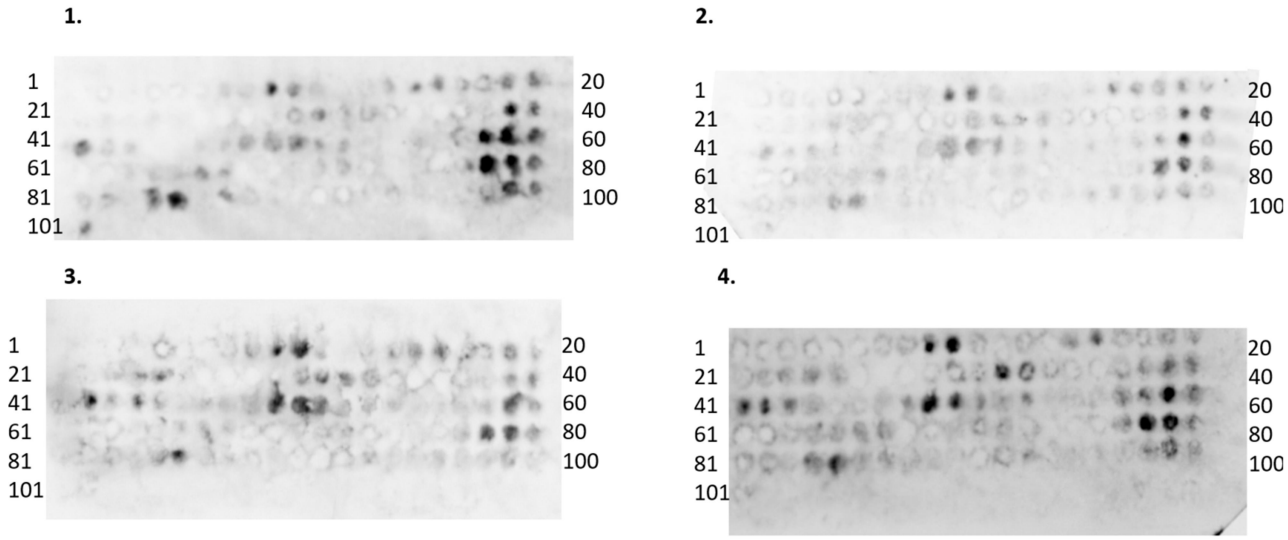
e. DAVIDS anti-CTDI-CTDII polyclonal antibodies, 1:1000 dilution



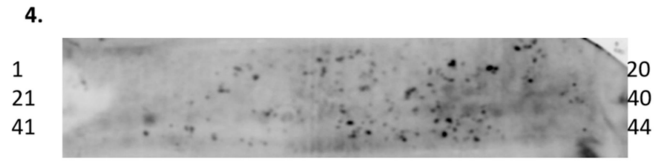
f. DAVIDS anti-CTDI-CTDII polyclonal antibodies, 1:5000 dilution

5. **FL DNAJB1 peptide libraries screening**

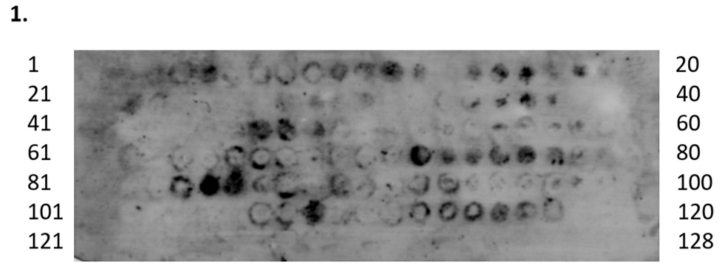
a. Luciferase library: all the pictures represents images of the third blot. The images were taken in four independent experiments.

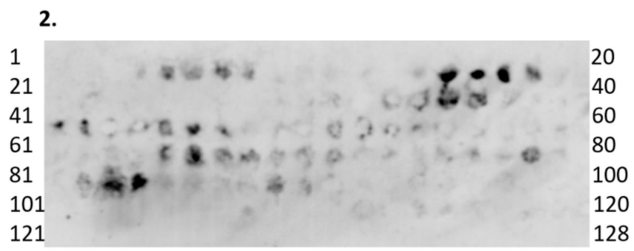


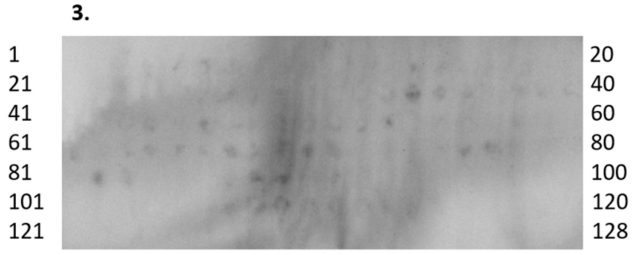
b. α-synuclein library: all the pictures represents images of the first blot. The images were taken in five independent experiments. Unfortunately, the signal was completely lost after the third experiments, maybe due to the very little amount of experiment which can be done with a library.



c. p53 library: all the pictures represents images of the third blot. The images were taken in three independent experiments.







6. **α-synuclein library analyses**: peptides were classified in: strong binders, weak binders and very weak binders. We analyzed them for the overall hydrophilicity, net charge at pH 7, amino acids occurrence exactly in the same way we analyzed the other libraries results. The data are comparable to the ones found screening luciferase and p53 libraries (the binding is occurring in the presence of aromatic and acid residues, the overall charge of the binders is negative). The only difference is in

the hydrophilicity, which tends to be higher in the very weak and weak binders compared to the non-binders.