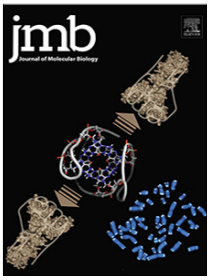
Research Article



Two Bacterial Small Heat Shock Proteins, IbpA and IbpB, Form a Functional Heterodimer

Artur Piróg [1], Francesca Cantini [2], Łukasz Nierzwicki [3], Igor Obuchowski [1], Bartłomiej Tomiczek [1], Jacek Czub [3*] and Krzysztof Liberek [1*]

*1 - **Intercollegiate Faculty of Biotechnology UG-MUG,** University of Gdansk, Abrahama 58, 80-307 Gdansk, Poland*
*2 - **Magnetic Resonance Center and Department of Chemistry,** University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy*
*3 - **Department of Physical Chemistry,** Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland*

***Correspondence to Jacek Czub and Krzysztof Liberek:** jacek.czub@pg.edu.pl (J. Czub), krzysztof.liberek@ug.edu.pl (K. Liberek)*
https://doi.org/10.1016/j.jmb.2021.167054
Edited by Prof. J. Buchner

Abstract

Small heat shock proteins (sHsps) are a conserved class of ATP-independent chaperones which in stress conditions bind to unfolded protein substrates and prevent their irreversible aggregation. Substrates trapped in sHsps-containing aggregates are efficiently refolded into native structures by ATP-dependent Hsp70 and Hsp100 chaperones. Most γ-proteobacteria possess a single sHsp (IbpA), while in a subset of *Enterobacterales*, as a consequence of *ibpA* gene duplication event, a two-protein sHsp (IbpA and IbpB) system has evolved. IbpA and IbpB are functionally divergent. Purified IbpA, but not IbpB, stably interacts with aggregated substrates, yet both sHsps are required to be present at the substrate denaturation step for subsequent efficient Hsp70-Hsp100-dependent substrate refolding. IbpA and IbpB interact with each other, influence each other's expression levels and degradation rates. However, the crucial information on how these two sHsps interact and what is the basic building block required for proper sHsps functioning was missing. Here, based on NMR, mass spectrometry and crosslinking studies, we show that IbpA-IbpB heterodimer is a dominating functional unit of the two sHsp system in *Enterobacterales*. The principle of heterodimer formation is similar to one described for homodimers of single bacterial sHsps. β-hairpins formed by strands β5 and β7 of IbpA or IbpB crystallin domains associate with the other one's β-sandwich in the heterodimer structure. Relying on crosslinking and molecular dynamics studies, we also propose the orientation of two IbpA-IbpB heterodimers in a higher order tetrameric structure.

Introduction

Exposure to stress results in cellular proteostasis imbalance and intracellular protein aggregation.[1,2] The proteostasis is restored by the protein quality control network.[3] Aggregated polypeptides are either degraded by cellular proteases or separated from aggregates and refolded into native structures by chaperones. The disaggregation and refolding process depends on small heat shock proteins (sHsps), the Hsp70 chaperone system and an Hsp100 disaggregase.[4,5] Small Hsps are first to act in this process by binding to partially unfolded client proteins to form reactivation-prone assemblies, a specific type of aggregates. These are then processed by both Hsp70 and Hsp100 which leads

to the extraction of single polypeptides and their refolding to native structures.[6,7]

Small Hsps family members are characterized by low molecular mass (12–43 kDa) and share universal structural organization consisting of a ~90 aa α-crystallin domain organized in highly conserved β-sandwich fold flanked by unstructured, not conserved N- and C-termini of variable length.[8,9] The central α-crystallin domain, which defines the sHsp, is composed of two antiparallel sheets of three and four β-strands connected by a short loop.[9] Isolated α-crystallin domains form stable dimeric structures.[10,11] Dimerization of bacterial, fungal, plant and archaeal α-crystallins occur via swapping the β6-strands of one α-crystallin domain into the β-sandwich of the neighboring one.[9,11] The dimerization of metazoan α-crystallins occurs due to formation of a dimer interface by their characteristic fused β6 and β7-strands.[9] Dimers of α-crystallin domains do not possess chaperone activity.[8] The full length sHsps form large, often polydisperse oligomeric species.[12–14] Formation of such big oligomeric structures depends on the presence of N- and C-terminal parts of sHsps.[15] In particular, universally conserved IXI motif present in C terminus of all sHps is critical for oligomerization events.[9] Increasing temperature triggers dissociation of sHsps oligomers into smaller forms capable of binding to substrate proteins.[12,16–18], recruiting them in assemblies containing both substrates and sHsps. Such assemblies are smaller than amorphous aggregates and store substrates in near-native conformations, shielding them from further aggregation.[19–21] Subsequently, when stress conditions are released, the substrate proteins from the assemblies are refolded into native structures by the Hsp70-Hsp100 chaperone system.[20,22]

The number of sHsps coding genes varies among organisms, ranging from 10 in human and even 19 in *Arabidopsis thaliana* to substantially lower numbers, usually one or two genes, in Prokaryotes and Archea.[8] Bacteria from the majority of γ-proteobacteria taxons possess one sHsp (IbpA), while in the ancestor of *Enterobacterales,* sHsp gene underwent a duplication and in consequence the bacteria in this order have two sHsps (IbpA and IbpB). This additional sHsp copy (*ibpB*) evolved faster than *ibpA*, suggesting its functional divergence.[23] In *Escherichia coli*, a member of *Enterobacterales* clade, *ibpA* and *ibpB* genes form an operon under heat shock sigma 32 factor control.[24] Their expression is strongly induced by heat stress and is regulated both at the transcriptional and translational level.[24,25] IbpA and IbpB were originally identified as inclusion body-associated proteins[26] and later were found to be associated with heat shock-induced protein aggregates.[27] *E. coli* IbpA and IbpB possess structural and oligomeric features characteristic for sHsps and share 58 % aa sequence similarity (47% identity), yet their functions are diverse. IbpA is specialized in efficient substrate binding upon aggregation (holdase activity)[19,28] while the presence of IbpB substantially facilitates dissociation of both sHsps from assemblies without compromising their formation.[23] IbpA and IbpB not only functionally cooperate in formation of refolding-competent assemblies with substrates but were also shown to interact with each other *in vitro* and *in vivo*, forming mixed complexes.[29] It was also reported that IbpA and IbpB influence each other's expression levels[24,25] as well as *in vivo* and *in vitro* degradation rates.[30,31]

The ability of IbpA and IbpB to interact and form mixed complexes[29,32] is probably the key factor determining their activity as a two protein sHsp system. However, it is not known what the basic building blocks of such IbpA-IbpB complexes are. The two most likely scenarios involve IbpA and IbpB forming either heterodimers or homodimers that further hetero-oligomerize. In both cases the formation of such complexes results in the incorporation of both IbpA and IbpB into the assemblies with substrates upon proteotoxic stress.

In this study we analyzed the functional subunit formed by IbpA and IbpB using diversified experimental approaches. Our results suggest that IbpA-IbpB heterodimer is the basic building block of *Enterobacterales* small Hsps chaperone system.

Results

Biochemical experiments reveal the preference for IbpA- IbpB heterodimers formation

IbpA and IbpB form large oligomeric structures of polymorphic nature[33,34] (Figure 1(A)), which makes the analysis of their basic functional units highly challenging. Since the formation of stable sHsps dimers is solely dependent on the α-crystallin domain (ACD), we decided to reduce the complexity of the system by using only α-crystallin domains of IbpA ($IbpA^{ACD}$) and IbpB ($IbpB^{ACD}$) (Figure S1 (A)). The α-crystallin domains of IbpA and IbpB were expressed and purified (Figure S1(B)). The sizes of the structures formed by purified domains were calculated from diffusion coefficients measured using the dynamic light scattering (DLS). Both $IbpA^{ACD}$ and $IbpB^{ACD}$ domains as well as the equimolar mixture of these show a similar, monodisperse peaks of approximately 4 nm in DLS measurements (Figure 1(A)). Assuming the globular shape of analyzed α-crystallin domains, we found that hydrodynamic radius of entities formed by $IbpA^{ACD}$, $IbpB^{ACD}$ and their mixture suggests the existence of dimeric structures in solution.

Next, to confirm that dimers of α-crystallin domains are present in solution, we analyzed the ability of these domains to form homo- or heterodimers using DMS crosslinking approach. Initially, it was not possible to resolve the crosslinked dimers of $IbpA^{ACD}$ and $IbpB^{ACD}$ using

SDS-PAGE (Figure S2). However, addition of 6 M Urea to the gel allowed for separation of respected dimers. When the equimolar mixture of IbpA and IbpB α-crystallin domains was crosslinked, an additional dominant band migrating between IbpA^{ACD} and IbpB^{ACD} dimers appeared on a gel, suggesting that heterodimeric structure is formed (Figure 1(B)). Correspondingly, the intensity of bands corresponding to IbpA^{ACD} and IbpB^{ACD} dimers became lower as compared to control lanes (Figure 1(B)). Next, blue native electrophoresis for the samples containing different ratio of IbpA^{ACD} and IbpB^{ACD} domains was performed. Similarly to crosslinking approach, the formation of heterodimer migrating in between IbpA^{ACD} and IbpB^{ACD} dimers was evident. The band corresponding to heterodimer was the most abundant when equimolar concentrations of IbpA^{ACD} and IbpB^{ACD} were used (Figure 1(C)). To directly distinguish between homo- and heterodimeric structures formed by IbpA^{ACD} and IbpB^{ACD}, we used native mass spectrometry. Consistently with the above result, the peak corresponding to heterodimer (IbpA^{ACD}-IbpB^{ACD}) was dominant over peaks corresponding to two possible homodimers (IbpA^{ACD}-IbpA^{ACD} and IbpB^{ACD}-IbpB^{ACD}) when IbpA^{ACD} and IbpB^{ACD} were present in equimolar concentrations (Figure 2). Heterodimer (IbpA^{ACD}-IbpB^{ACD}) was also evident when different ratios of both analyzed α-crystallins were used (Figure 2).

All the above experiments point to the preference for the formation of IbpA^{ACD}-IbpB^{ACD} heterodimers over two possible homodimers. This suggests that IbpA-IbpB heterodimer is a basic functional unit defining the mode of cooperation between IbpA and IbpB. To further analyze the structural determinants of IbpA^{ACD}-IbpB^{ACD} heterodimer formation, we performed solution NMR studies.

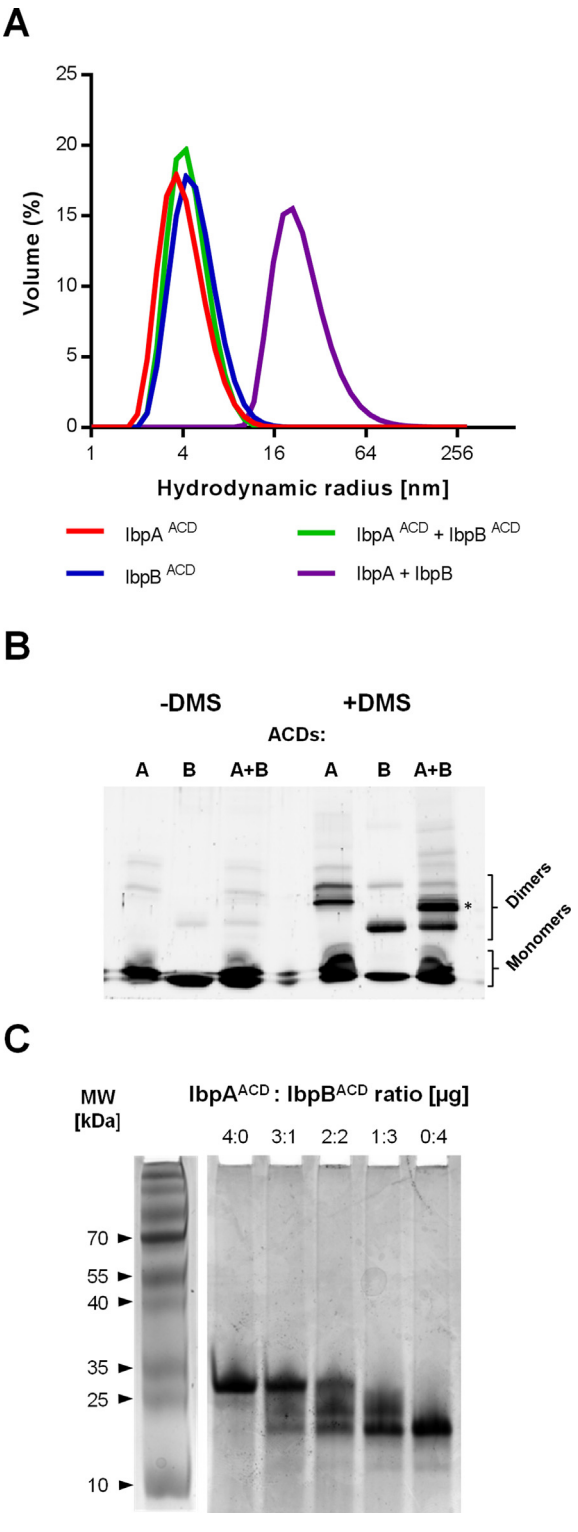


Figure 1. Isolated α-crystallin domains of IbpA and IbpB form dimers. (A) Particle size distributions measured by dynamic light scattering. The hydrodynamic radius of IbpA^{ACD}, IbpB^{ACD}, their mixture and the mixture of wild type IbpA and IbpB was measured by DLS. IbpA^{ACD} and IbpB^{ACD} were present at 5 μg/μl, IbpA^{ACD}, IbpB^{ACD} mixture at 10 μg/μl, and WT IbpA and IbpB complex at 0.4 μg/μl concentration. The measurements were performed at 25 °C. (B) DMS crosslinking of IbpA and IbpB α-crystallin domains and their mixture suggests heterodimer formation. Samples following DMS crosslinking were resolved using SDS-PAGE-Urea gel electrophoresis system which allowed to separate dimer populations. 0.5 μg of respective α-crystallin domains were loaded on each lane. Asterisk marks IbpA^{ACD}-IbpB^{ACD} heterodimer. **C**. Blue native electrophoresis of IbpA^{ACD} and IbpB^{ACD}, and their mixtures suggests that IbpA^{ACD}-IbpB^{ACD} heterodimer is present in solution. 4 μg of respective α-crystallin domains were loaded on each lane.

NMR combined with molecular dynamics provide insight into IbpA^{ACD}-IbpB^{ACD} structure

To investigate the IbpA^{ACD}-IbpB^{ACD} heterodimer formation, particularly to obtain information on its structure and binding interface, we acquired ^{1}H-^{15}N heteronuclear single quantum correlation

(HSQC) spectra on both IbpA^{ACD} and IbpB^{ACD}. Initially, we acquired NMR spectra at 298 K and pH 6.8, but due to the severe line broadening and protein precipitation, we shifted to pH 7.2. The ^{1}H-^{15}N HSQC spectra of both IbpA^{ACD} and IbpB^{ACD} acquired at 298 K and pH 7.2 still showed severe line broadening beyond detection of several amide signals (Figure S3). However, upon increasing the temperature to 313 K, ^{1}H-^{15}N HSQC spectra of both IbpA^{ACD} and IbpB^{ACD} displayed an increase number of well-dispersed resonances indicative of folded proteins (Figure S3). At this temperature, 82 of the expected 92 ^{15}N backbone amide resonances were detected for IbpA^{ACD}, and 77 out of the expected 91 for IbpB^{ACD} protein. In the case of IbpB^{ACD}, the quality of the ^{1}H-^{15}N HSQC and of the 3D triple resonance spectra allowed a partial resonance backbone assignment. We were able to assign 63% of backbone resonances of IbpB^{ACD} (i.e. stretches Y35-E38, I47-Q62, T66-T73, Q86-L88, F93-S94, V106-T109, V111-L114, and H116-N122; numbers refer to full length IbpB) (Figure S4). Assignment of the backbone resonances allowed to predict the presence of secondary structure elements in IbpB^{ACD} from ^{15}N, ^{13}C', ^{13}Cα and ^{13}Cβ chemical shifts (Figure 3 (C)). On the contrary, despite the same experimental conditions, the 3D triple resonance experiments acquired on IbpA^{ACD} showed the lack of many Cα and Cβ resonances. This behavior, may indicate that IbpA^{ACD}, and partly IbpB^{ACD}, sample multiple conformations, which exchange with each other with an intermediate exchange rate on the NMR time scale.

The dynamic properties of IbpA^{ACD} and IbpB^{ACD} domains at 0.5 mM concentration were experimentally characterized through ^{15}N relaxation measurements at 313 K and 500 MHz. The average backbone ^{15}N longitudinal R_1 and transversal R_2 relaxation rates, and ^{15}N^{1}H-NOEs values are 1.8 ± 0.2 s^{-1}, 18.7 ± 2.1 s^{-1} and 0.6 ± 0.1, for IbpA^{ACD} and 1.6 ± 0.4 s^{-1}, 18.7 ± 1.5 s^{-1} and 0.7 ± 0.1 for IbpB^{ACD}, respectively (Figure S5, Table S1). The correlation times for overall rotational tumbling (τ_c), as estimated from the R_2/R_1 ratios, are 12.2 ± 1.3 ns for IbpA^{ACD}, and 12.9 ± 0.8 ns for IbpB^{ACD}. The τ_c value calculated by the HYDRONMR program[35] is about 6.8 ns for the monomeric protein. Therefore, the experimental τ_c values are consistent with both IbpA^{ACD} and IbpB^{ACD} domains being in a dimeric state.

The formation of IbpA^{ACD}-IbpB^{ACD} heterodimer was further examined with NMR titration

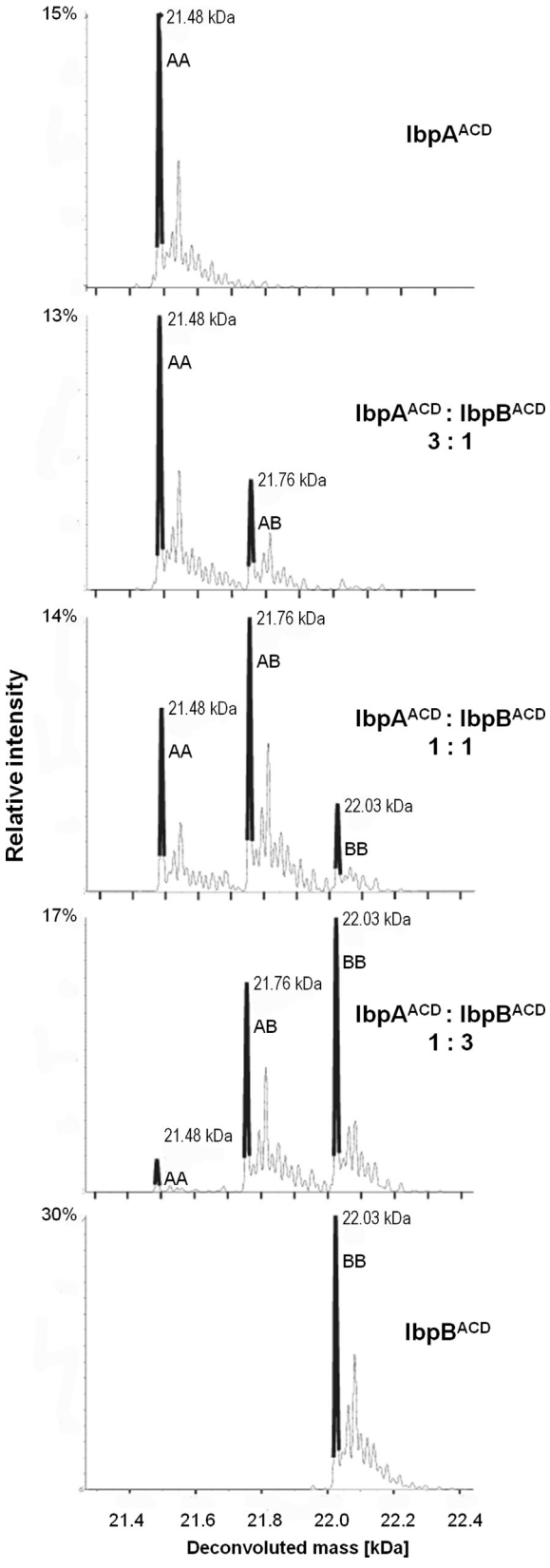


Figure 2. Native mass spectrometry of IbpA and IbpB α-crystallin domains show heterodimer formation. Native ESI mass spectra of purified IbpA^{ACD}, IbpB^{ACD} domains and its mixtures of indicated ratio. Deconvoluted spectra show indicated peaks corresponding to IbpA^{ACD}-IbpA^{ACD} homodimer, IbpB^{ACD}-IbpB^{ACD} homodimer and IbpA^{ACD}-IbpB^{ACD} heterodimer.

experiments, where unlabeled IbpA^{ACD} was added to a solution of ^{15}N-labeled IbpB^{ACD} (Figure 3(A)) and vice versa, where unlabeled IbpB^{ACD} was added to ^{15}N-labeled IbpA^{ACD} (Figure S6). In both cases, a second set of signals in a slow exchange regime on the NMR timescale appeared, indicating the formation of a stable IbpA^{ACD}-IbpB^{ACD} complex. The overlay of the 2D ^{1}H–^{15}N HSQC spectra acquired on ^{15}N-labeled IbpB^{ACD} alone and in the presence of one IbpA^{ACD} equivalent and analogous spectra of ^{15}N-labeled IbpA^{ACD} alone and in the presence of one IbpB^{ACD} equivalent showed that some residues exhibited different chemical shift values (Figure 3 (A) and Figure S6). IbpB^{ACD} residues with different chemical shift values are presented in Figure 3(A). Due to the lack of chemical shift assignment for IbpA^{ACD} residues, the same was not possible for IbpA^{ACD}. The ^{15}N longitudinal R_1 relaxation rates acquired at 313 K for IbpB^{ACD} in the presence of unlabeled IbpA^{ACD} confirmed the formation of a heterodimeric IbpA^{ACD}-IbpB^{ACD} complex (the average backbone ^{15}N longitudinal R_1 relaxation rate of the IbpA^{ACD}-IbpB^{ACD} complex is 1.4 ± 0.6 s^{-1}).

Next, we used the set of amide chemical shifts of IbpB^{ACD} affected by the presence of IbpA^{ACD} (Figure 3(B)) as structural restraints in replica-averaged steered molecular dynamics (MD) to predict the structure of IbpA^{ACD}-IbpB^{ACD} dimer. To this end, we first obtained a homology model of both IbpA^{ACD} and IbpB^{ACD} monomers with I-Tasser using IbpA homolog from *Xanthomonas axonopodis* γ-proteobacterium (Hsp20; pdb ID 3GLA) as a template. This sHsp was selected based on high sequence similarity, consistent with the well-known strong structural conservation among the bacterial α-crystallin domains. To obtain the initial structure of the IbpA^{ACD}-IbpB^{ACD} heterodimer for MD, we superimposed IbpA^{ACD} and IbpB^{ACD} monomers on individual subunits of the Hsp20 dimer and initially relaxed the model with 1 μs conventional MD run. Next, we refined the IbpA^{ACD}-IbpB^{ACD} structure using replica-averaged steered MD to reproduce the chemical shifts measured for IbpB^{ACD} in the heterodimer. During this simulation the sum of square deviation between the model-predicted and experimental chemical shifts improved from 66.7 to 28.3 ppm^2. Clustering of the MD ensemble revealed the existence of two similar binding modes, in which the long flexible β-hairpins formed by strands β5 and β7 of both IbpA^{ACD} and IbpB^{ACD} associate with the β-sandwich of their binding partner. In both these modes, the IbpB^{ACD} residues, which were identified to change their chemical shifts upon addition of IbpA^{ACD} are found at the interface of IbpA^{ACD}-IbpB^{ACD} heterodimer, mainly in the IbpB^{ACD} β-sandwich region, where they interact with the flexible β-hairpin of IbpA^{ACD} and in the IbpB^{ACD} β-hairpin - interacting with IbpA^{ACD} β-sandwich region (Figure 4). Notably, by predicting the structure of the IbpB^{ACD} homodimer (Figure S7) using the same approach as for the heterodimer, we found that the observed change in IbpB^{ACD} chemical shifts arise mainly from the exchange of a binding partner to IbpA^{ACD}. In fact, the average conformations adopted by IbpB^{ACD} in the IbpA^{ACD}-IbpB^{ACD} and IbpB^{ACD}-IbpB^{ACD} dimers are very similar and thereby the internal conformation does not largely differentiate the chemical shifts of IbpB^{ACD} (Figures 4 and S7). Unfortunately, we were not able to do the same analysis for the IbpA^{ACD} homodimer as the resonances in its spectra were not assigned.

Overall, our NMR measurements confirm that IbpA^{ACD} and IbpB^{ACD} tend to form stable heterodimers whose structure resembles typical bacterial α-crystallin homodimers with long flexible β-hairpins (β5 and β7), and possessing in its structure short β6 strand which interacts with the β-sandwich of the partner subunit.

IbpA-IbpB herodimers form higher order structures

Knowing that α-crystallin domains of IbpA and IbpB preferentially form heterodimers, we turned to the analysis of full-length proteins. Since both IbpA and IbpB form polydisperse and highly dynamic oligomeric structures, we decided to use crosslinking approach. When the equimolar mixture of IbpA and IbpB was crosslinked with DMS, not only IbpA and IbpB homodimers were observed but also additional bands migrating in between appeared on the polyacrylamide gel, suggesting that a heterodimeric structure is formed (Figure S8). Next, we applied specific crosslinking approach combined with mass spectrometry. We *in vivo* incorporated UV-crosslinkable unnatural amino acid, p-benzoyl-phenylalanine (Bpa) into several positions in both IbpA and IbpB using genetic code expansion technology.[36] The sites of Bpa incorporation (Figure 5(A)) were selected considering several factors. Firstly, highly conserved sites were ruled out due to high probability of protein structure disruption and activity loss. Secondly, we focused more on sites in the N- and C-termini, as these parts of sHsps are involved in oligomerization and were not analyzed in our NMR studies. Finally, some additional sites were selected based on results from 37 – considering *in vivo* activity or distinctive crosslinking pattern (e.g. IbpB Y45Bpa). Sites of Bpa incorporation in IbpA and IbpB were designed to be homologous (Figure 5(A)). We purified respective IbpA and IbpB variants and analyzed their ability, when present during luciferase thermal denaturation step, to stimulate luciferase refolding by bacterial DnaK-ClpB chaperone machinery. Initially, the activity of respective IbpA and IbpB Bpa variants was assayed, titrating them together with their wild type

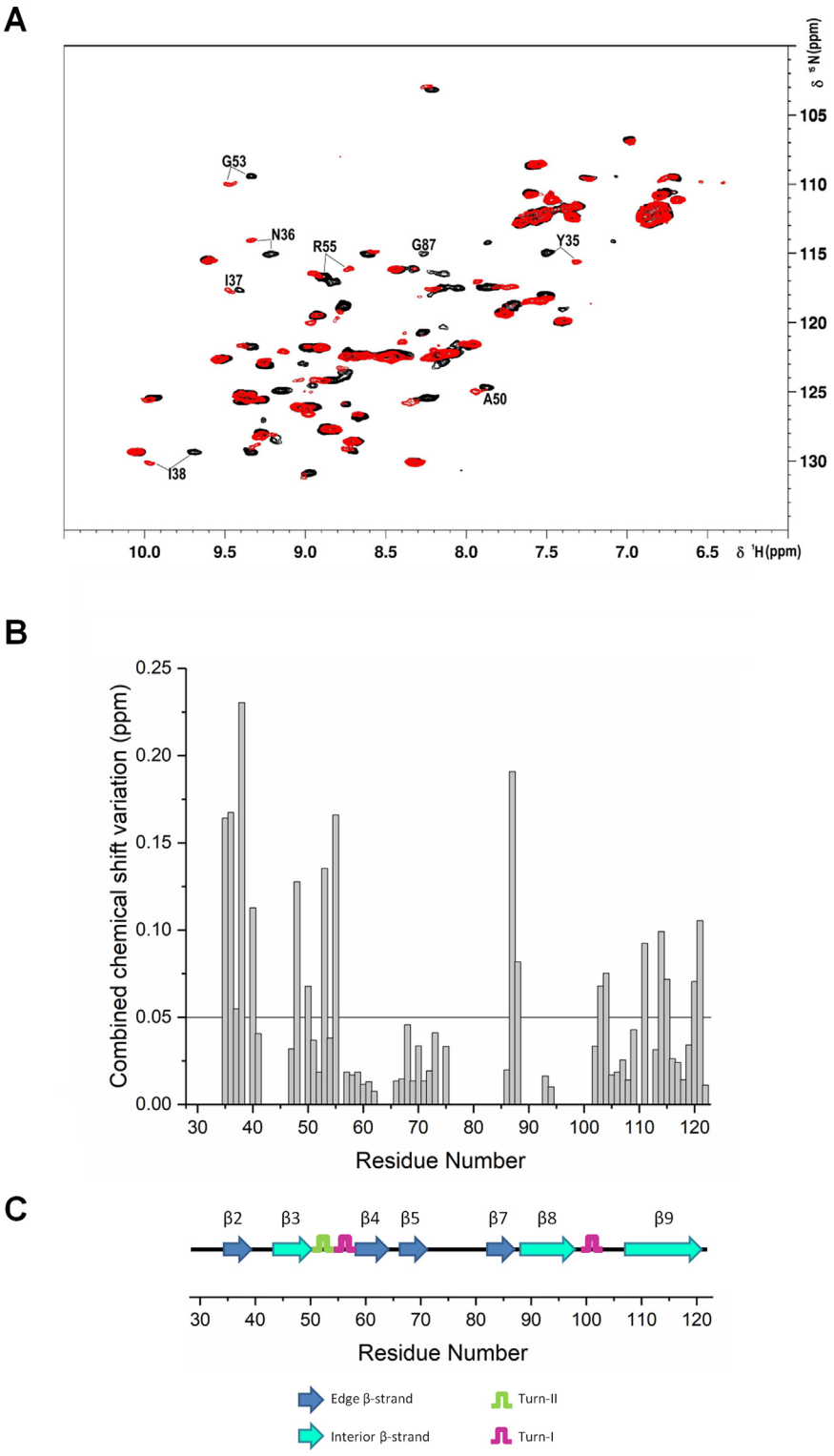


Figure 3. NMR studies show that IbpA and IbpB α-crystallin domains preferentially interact with each other. (A) Overlay of 2D ^{1}H-^{15}N HSQC spectra acquired at 500 MHz at 313 K of ^{15}N IbpB^{ACD} alone (black) and ^{15}N IbpB^{ACD}-IbpA^{ACD} complex (red). (B) Combined chemical shift variations of signals between ^{15}N IbpB^{ACD} and ^{15}N IbpB^{ACD} in the presence of unlabelled IbpA^{ACD}. These values are calculated from the experimental ^{1}H and ^{15}N chemical shift changes ($\Delta\delta(^{1}H)$ and $\Delta\delta(^{15}N)$, respectively) between corresponding peaks in the two forms as described in 62. Missing assignment of a number of amide protons prevented the determination of their chemical shifts perturbations. **C**. Analysis of the secondary structure elements from the backbone chemical shifts. Secondary structure elements were predicted from ^{15}N, ^{13}C', ^{13}Cα and ^{13}Cβ chemical shifts using TALOS + program.

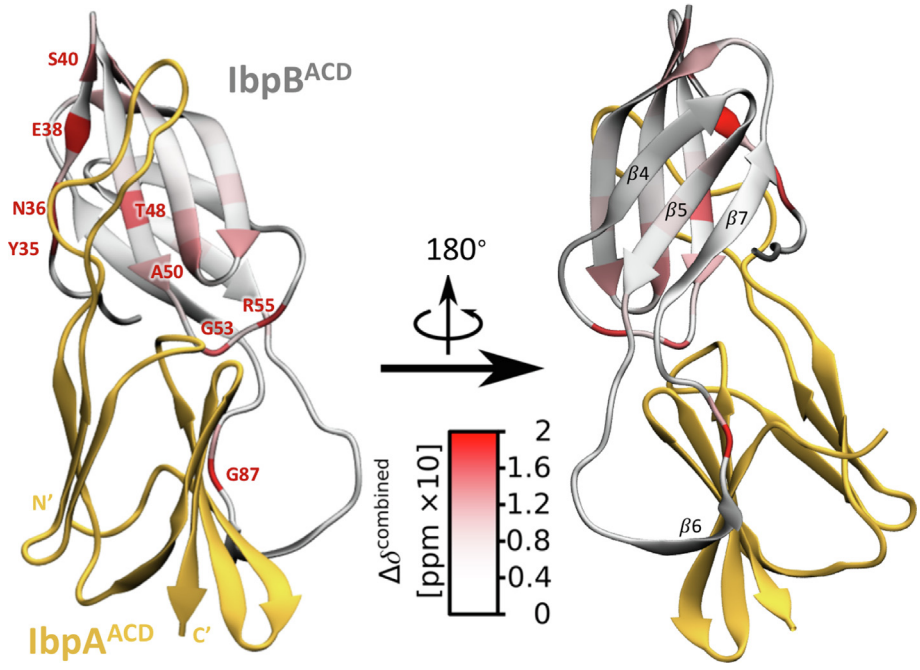


Figure 4. Model of IbpA α-crystallin - IbpB α-crystallin domain heterodimer. Representative structure of $IbpA^{ACD}$-$IbpB^{ACD}$ heterodimer obtained by NMR-guided MD simulations. Differences in each $IbpB^{ACD}$ residue's experimental chemical shifts upon addition of $IbpA^{ACD}$ are indicated by color scale. The residues undergoing large chemical shift changes are additionally marked. Nine $IbpB^{ACD}$ residues with highest combined chemical shift variation are visualised on the structure. The distance of marked residue in $IbpB^{ACD}$ to the closest residue in $IbpA^{ACD}$ were determined ($IbpB^{ACD}$ Y35 to $IbpA^{ACD}$ R93 0.18 nm; N36 to I90 0.22 nm; E38 to Q88 0.17 nm; S40 to Y85 0.25 nm; T48 to R93 0.23 nm; A50 to A54 0.24 nm; G53 to A52 0.26 nm; R55 to N114 0.75 nm; G87 to N38 0.22 nm).

partner in refolding experiment (Figures S9 and S10) and then at fixed saturating concentration (Figure 5(B)). Only those variants that showed at least 40% activity of wild type IbpA and IbpB in refolding stimulation assay (Figure 5(B)), that is two-fold higher than the activity of luciferase refolded from aggregates formed in the absence of sHsps (approx. 20%), were subjected to crosslinking analysis.

Complexes of sHsps variants were UV-irradiated and the resulting crosslinking products were further separated by SDS-PAGE and stained with Coomassie Brilliant Blue. Bands corresponding to monomer and dimer were excised and subjected to in-gel tryptic digestion followed by NanoLC/MS analysis using Orbitrap Velos mass spectrometer operating in HCD fragmentation mode. To identify crosslinked peptides, we used StavroX,[38] a software for analyzing crosslinked products in protein interaction studies. The search was performed against IbpA and IbpB sequences. With this approach, we detected 38 either intra- or interchain crosslinks, which are listed in Table S2. We found that ten of the identified crosslinks were formed either within or between the α-crystallin domains. Six of these α-crystallin crosslinks, two pairs located within each of the domains and two interdomain ones (Table S2), can be mapped onto our predicted structure of the $IbpA^{ACD}$-$IbpB^{ACD}$ heterodimer within the maximum distance constraint of 2 nm for the Bpa crosslinker[39] (Figure 6), thereby supporting the $IbpA^{ACD}$- $IbpB^{ACD}$ NMR-based model.

The four remaining α-crystallin crosslinks (Table S2) were found to be located at the exposed surface of the $IbpA^{ACD}$-$IbpB^{ACD}$ structure, suggesting that they may in fact form between individual heterodimers in higher-order oligomers. Assuming that all these inter-dimer crosslinks exist between a pair of IbpA-IbpB heterodimers, their positions indicate that the heterodimers may further self-assemble into centrosymmetric tetramers with all exposed crosslink sites buried at the dimer-dimer interface (Figure 6). To predict the structure of such a hypothetical higher-order oligomer, we used MD simulations. In particular, since the inter-dimer crosslinks unequivocally determine the relative orientation of the dimers, we first used them as distance restraints in steered MD simulations to drive the association of two initially separated IbpA-IbpB heterodimers into a tetramer. Subsequently, after 500 ns of an additional unrestrained MD refinement, we clustered the conformational ensemble of the tetramer based on backbone similarity, which led us to identify a single tetrameric binding mode composed of 82% of the trajectory (Figure 6). To validate this structure, we examined whether the experimental crosslinks involving the N- and C-terminal tails of

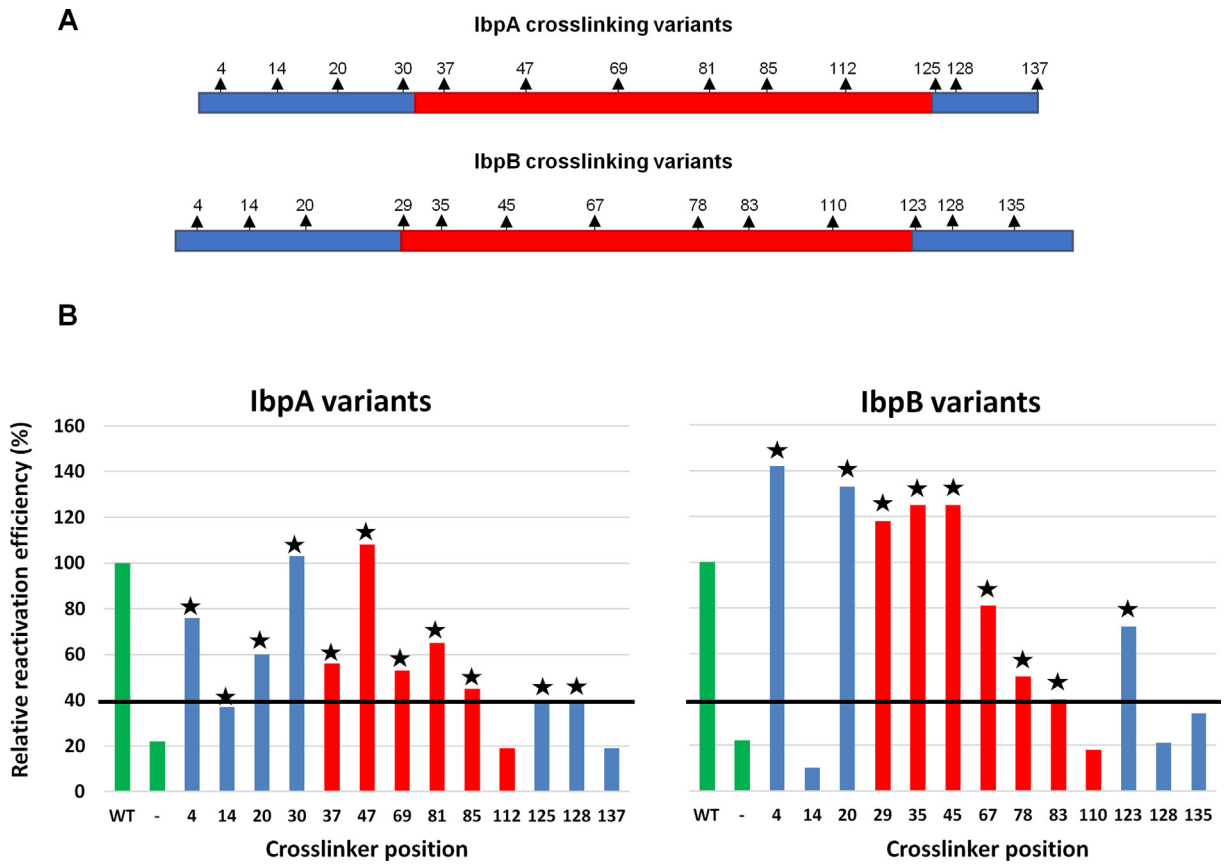


Figure 5. Crosslinking variants of IbpA and IbpB and their activities. (A) Location of Bpa crosslinker in IbpA and IbpB. C- and N- terminal tails (blue) and α-crystallin domains (red). (B) Activity of obtained variants (blue and red, for variants with crosslinker located in tails or in α-crystallin domains, respectively) relative to the activity of the wild-type IbpA and IbpB (green). Variants used in further analyses are indicated with asterisk. Luciferase (1.5 μM) was denatured in the presence of indicated sHsps present at 10 μM concentration (sum of IbpA or its variant present at 3 μM and IbpB or its variant at 7 μM), diluted to 42 nM and subsequently refolded for 60 min in the presence of DnaK (0.7 μM), DnaJ (0.3 μM), GrpE (0.21 μM) and ClpB (2.4 μM).

IbpA and IbpB are consistent with the model. We assumed that to allow for a crosslink formation between the two sites, the minimal Cα-Cα distance between these sites should be less than 2.0 nm and, at the same time, the average Cα-Cα distance should not exceed 2.5 nm.

With these criteria, we found an excellent agreement between the crosslinking experiments and the MD-identified tetrameric structure, with 13 out of 16 experimentally identified crosslinks consistent with the model. 3 of these crosslinks were identified between IbpA and IbpB forming the heterodimer building blocks, 5 were formed between the two IbpA-IbpB heterodimers and 5 could be formed either within the heterodimer (Table S3) or between the heterodimer blocks (Table S4). The three remaining experimental crosslinks involve residues 40–46 and 123–131 of IbpB that in our tetramer model form a well-defined continuous surface exposed to a solvent (Figure S11). Therefore, we postulate that this surface might be responsible for the formation of higher-order assemblies.

Notably, our tetramer structure is in agreement with previous results indicating that the C-terminal tail of IbpA is crucial for oligomerization.[32] Indeed, in the predicted structure the two isoleucine residues of the conserved IXI motif of the IbpA C-terminal tail spontaneously insert into the hydrophobic cleft in the α-crystallin domain of IbpB in the neighboring heterodimer (Figure S12) and the C-terminal arginines 132 and 133 of IbpA form salt bridges with glutamates 104 and 60 or 64, respectively, in IbpB (Figure S13).

Taken together, our crosslinking-guided MD simulations allowed us to propose a model of the sHsps tetramer in which the interaction between two IbpA-IbpB heterodimer building blocks is mediated predominantly by the N- and C-terminal tails.

Discussion

By interacting with misfolded proteins and protecting them from irreversible aggregation, small Hsps act as a first line of defense against

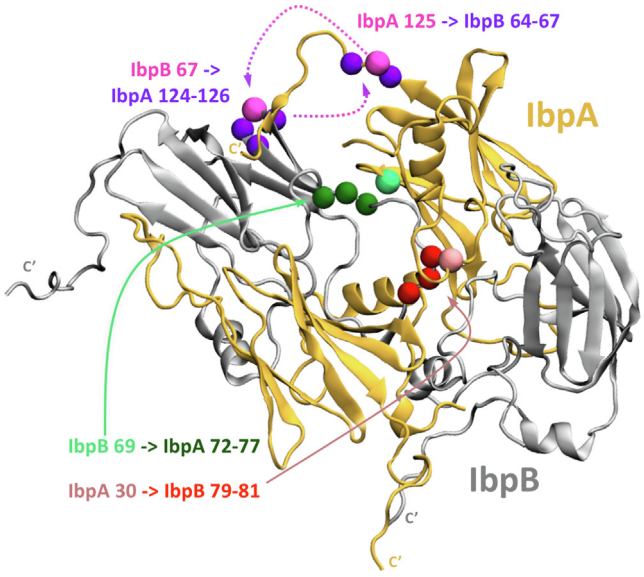


Figure 6. MD-predicted structure of IbpA-IbpB tetramer. The color spheres denote crosslinks between IbpA and IbpB that were used to drive the initial association of the IbpA-IbpB heterodimers. Light colors define the crosslinker position. Dark colors define the most probable site of crosslink.

proteotoxic stress. In *Enterobacterales,* as a result of gene duplication event, two sHsp system consisting of IbpA and IbpB has evolved. Several reports showed that IbpA and IbpB cooperate closely to protect proteins both *in vivo* and *in vitro*,[23,29,32] yet the crucial information on how these two sHsps interact and what is the basic building block of higher order oligomeric structures was missing. Here, using multiple methods, we showed that IbpA-IbpB heterodimer is a dominating functional unit of the two sHsp system in *Enterobacterales*.

NMR studies combined with molecular dynamics simulations, native mass spectrometry, crosslinking approach and blue native electrophoresis gave complementary and consistent results which revealed that α-crystallin domains of IbpA and IbpB form heterodimer in different experimental conditions. The principle of $IbpA^{ACD}$-$IbpB^{ACD}$ heterodimer formation is very similar to the one described for homodimers of single bacterial, fungal, plant and archaeal sHsps.[9] NMR studies showed that β-hairpins formed by strands β5 and β7 of both $IbpA^{ACD}$ and $IbpB^{ACD}$ associate with the other one's β-sandwich in the heterodimer structure. Mass spectrometry, and NMR titration experiments strongly suggested that $IbpA^{ACD}$-$IbpB^{ACD}$ heterodimer formation is preferred over $IbpA^{ACD}$ and $IbpB^{ACD}$ homodimers that are formed in the absence of the proper interacting partner. Of course, when analyzing these experiments, one has to additionally take into account the stability of preexisting isolated homodimers - the time scale and conditions at which the $IbpA^{ACD}$ and $IbpB^{ACD}$ homodimers efficiently exchange into heterodimer is not known.

It should also be noticed that *ibpA* and *ibpB* genes form an operon and are translated from a bicistronic mRNA.[24,40] This might allow for the efficient formation of heterodimers during posttranslational folding events which would require no dimer subunits mobility at all. Such scenario would strongly promote the heterodimer over homodimer formation *in vivo*. Indeed, it was previously reported that in bacteria the assembly of heterodimeric luciferase is an organized cotranslational process that is facilitated by spatially confined translation of the subunits encoded on a polycistronic mRNA.[41] One has to also have in mind that IbpA and IbpB might not be expressed at the same level. There was a report suggesting that IbpB concentration in the cell is higher than IbpA.[42] The excess of IbpB over IbpA in bacterial cell might be required to keep the majority of IbpA molecules in a heterodimeric form for some reason. Potentially, the excess of IbpB might be beneficial for the cell since IbpA alone binds unfolded substrates so tightly that *in vitro* DnaK-ClpB-dependent substrate refolding is inhibited.[23] Formation of heterodimers results in the incorporation of IbpB molecules, alongside IbpA, into assemblies formed at stress conditions. Since IbpB alone possesses low substrate binding potential,[19,23] its presence in heterodimers with IbpA weakens the overall interaction of heterodimer with the substrate. Thus, the dissociation of sHsps from the assemblies, required to initiate the DnaK-ClpB dependent disaggregation and refolding process,[20] becomes easier which is beneficial for the process. In agreement, it was reported that *in vitro* several IbpA proteins from different bacteria, both from single and two protein sHsps systems, were similarly efficient in formation of assemblies with protein substrates.

However, addition of IbpB partner to those forming two protein system, allowed for easier sHsps dissociation and in consequence lowered the demand for DnaK required for refolding.[23] Additionally, in *E. coli* Δ*ibpB* strain, IbpA was maintained in the aggregates in contrast to the wild-type strain where it was cleared significantly faster.[23] Thus, the presence of a two protein sHsps system, composed of IbpA and IbpB heterodimers, is beneficial since it allows for fast and efficient binding of sHsps to an aggregating substrate and at the same time allows for easier dissociation of sHsps from their assemblies.

Our results also suggest that IbpA-IbpB heterodimers show a clear tendency to form higher order assemblies, i.e., tetramers. In accordance with a dimeric nature of isolated α-crystallin domains, the predicted structure of tetramers suggests that they are mostly held together by the interactions between the N-terminal tails and by the conserved C-terminal IXI motif interacting with the α-crystallin domain. In fact, we found the IXI motif to bind precisely into the location predicted by.[32] It was shown there that the C-terminal tail of IbpA docks into the hydrophobic groove in α-crystallin of neighbouring IbpA dimer and arginine in position 133 forms salt bridge with glutamic acid in this dimer. This *in silico* prediction was also confirmed by *in vitro* experimental data showing that both the IXI motif and arginine 133 are required for proper IbpA oligomerization and functional cooperation with IbpB in luciferase refolding assay.[32] In our tetramer model, the IXI motif of IbpA C terminus from one of the heterodimers interacts with the other heterodimer's IbpB α-crystallin domain in a very similar manner. Isoleucine residues from IXI motif locate inside the hydrophobic cleft in the α-crystallin domain and C-terminal arginine residues in positions 132 and 133 form salt bridges with acidic residues in neighbouring heterodimer, forming the tetrameric structure. N-terminal ends in turn were shown to be flexible and to be involved in intermolecular interactions with the other heterodimer N-terminal ends, which contributes to the tetramer stability. Taking into account the solid tetramer model support, it seems to be a likely building block for even higher order oligomers. These are suggested to be the "storage form" of sHsps, keeping the small Hsps inactive which protects the cell from unwanted action of sHsps in non-stress conditions.

IbpA-IbpB heterodimers, described here as a minimal oligomeric structure formed as a consequence of *ibpA* gene duplication in the ancestor of *Enterobacterales*, is an example of a relatively rare event in evolution of prokaryotes. It was reported that in procaryotes paralogous pairs of proteins form heterooligomers substantially less frequently than homooligomeric structures.[43] This tendency is reversed in eukaryotes in which heteromeric complexes dominate over homomeric.[43,44] It was suggested that formation of heterodimeric structures formed by human sHsps may exert regulatory effects and allow the precise tuning of their chaperone properties.[45] Similarly the formation of heterodimeric structure by IbpA and IbpB in *Enterobacterales* may allow to tune their chaperone properties, especially substrate selection, binding and dissociation.

Materials and methods

Proteins

Purification of WT IbpA. WT IbpA was expressed in *E. coli* BL21DE3Δ*ibpAB* cells from pET15b plasmid. Cells were lysed in buffer A (50 mM Tris pH 7.4, 150 mM KCl, 10% glycerol and 5 mM β-ME) using a French press. Lysate was centrifuged at 81,000g for 45 min. Supernatant was collected and pellet was extracted with the same buffer with addition of 2 M Urea and centrifuged again in the same conditions. Resulting pellet was extracted again using 6 M Urea as additive. The most IbpA-enriched supernatant was dialyzed against buffer A supplemented with 6 M urea. The dialysate was loaded on Q-Sepharose column and eluted using gradient of 150–500 mM KCl. Most abundant fractions were dialyzed against 20 mM phosphate buffer pH 6.8, 150 mM KCl, 10% glycerol, 6 M Urea and 5 mM β-ME and subsequently incubated overnight with hydroxyapatite resin, washed and eluted using 20–500 mM phosphate gradient. Finally, selected fractions were dialyzed against buffer B (50 mM Tris pH 8.2, 150 mM KCl, 10% glycerol, 6 M Urea and 5 mM β-ME). Next, protein was refolded by stepwise dialysis against buffer B without Urea and subsequently to buffer A in the last step.

Purification of WT IbpB. WT IbpB was expressed in *E. coli* BL21DE3Δ*ibpAB* cells from pET28b plasmid. Cells were lysed in buffer containing 50 mM Tris pH 7.6; 10% glycerol; 100 mM KCl; 5 mM β-ME using a French press. Proteins were fractionally precipitated using 20% and 50% saturated ammonium sulfate. Second precipitated fraction contained IbpB protein and were dissolved and subsequently dialyzed against 50 mM Tris pH 8.9, 6 M Urea, 10% glycerol, 5 mM β-ME. The dialysate was loaded onto Q-Sepharose column equilibrated with the same buffer. The column was washed with the same buffer supplemented with 100 mM NaCl and the proteins were eluted using 100–500 mM NaCl. Selected fractions were dialyzed against buffer A supplemented with 6 M Urea. Subsequently, IbpB was refolded by a stepwise dialysis into buffer A.

Purification of IbpA-α-crystallin and IbpB-α-crystallin domains. α-Crystallin domains ($IbpA^{ACD}$ or $IbpB^{ACD}$) with N-terminal His-tags were expressed in *E. coli* BL21DE3Δ*ibpAB* cells. Cells were disrupted by sonication in buffer D (50 mM Tris pH 7.6, 6 M Urea, 10% glycerol, 500 mM NaCl, 5 mM β-ME). Lysate was clarified by centrifugation at 81,000*g* for 45 min. Supernatant was loaded on Ni-NTA resin (Macherey-Nagel Protino NTA Agarose), and washed with buffer D supplemented with 20 mM imidazole. Subsequently, proteins were eluted by a gradient of 20–360 mM imidazole in the same buffer. Proteins were subjected to additional rounds of Ni-NTA purifications, until purity level was satisfactory. Proteins were dialyzed into buffer B with 1 mM $CaCl_2$. Proteins were refolded by a stepwise dialysis into buffer C (50 mM Tris pH 8.2, 150 mM KCl, 10% glycerol, 5 mM β-ME) with 1 mM $CaCl_2$. His-tags were removed by using 0.5 u/ml of recombinant thrombin (Serva). Digestion was carried out overnight at room temperature. Thrombin was inactivated by 1 mM PMSF. As a result of cloning and thrombin cleave procedure three additional aa GSH were present at the N terminus of $IbpA^{ACD}$ and $IbpB^{ACD}$. Subsequently, α-crystallin domains were dialyzed back to buffer B, and incubated with Ni-NTA resin to remove residual undigested protein and subsequently refolded by stepwise dialysis to buffer C, and then to buffer A. After that, protein were dialyzed into 50 mM Tris pH 7.4, 150 mM KCl, and finally into 50 mM NH_4HCO_3. For samples intended to be measured by NMR 2 mM KCl was added to facilitate solubilization after freeze-drying. In samples intended for mass spectrometry measurements KCl was omitted due to possible ionization suppression. Protein concentration was determined by UV absorbance at 280 nm and 214 nm wavelengths. For 280 nm wavelength, extinction coefficients were predicted using ProtParam, and for 214 nm, coefficients were calculated according to.[46] The calculated molecular mass of purified $IbpA^{ACD}$ and $IbpB^{ACD}$ domains is 10,743 Da and 11,011 Da, respectively.

Other proteins

Published protocols were used for the purification of *E. coli* DnaK, DnaJ, GrpE,[47] and ClpB.[48] Firefly luciferase was purchased from Promega. Protein concentrations were determined with the Bradford (Bio-Rad) assay system, using bovine serum albumin as a standard. Molar concentrations are given assuming a monomeric structure of the proteins.

DMS crosslinking

α-Crystallin domains or full length sHsps were dialyzed to 50 mM potassium phosphate buffer pH 6.8 and 150 mM KCl. 30 mM stock solution of DMS were prepared freshly in the same buffer. DMS was added to protein solution to the final concentration of 3 mM. Crosslinking reaction was quenched by the addition of Laemmli sample buffer. Urea-SDS electrophoresis was performed similarly to classic Laemmli protocol.[49] Gels were casted manually, including 6 M urea in resolving gel, stacking gel and sample buffer.

Blue native electrophoresis

Blue native PAGE gel (12% acrylamide, 10% glycerol, 50 mM TRIS-HCl pH 8.4) polymerization was started by addition of 32 μl of freshly dissolved 10% APS and 8 μl of TEMED per 8 ml of the gel mixture. After initial polymerization (15 min, room temp.) gels were incubated for 2 h in 37 °C, another 2 h in room temp and run immediately after the last incubation. Samples were loaded after short pre-electrophoresis (10 min, 80 V) and the electrophoresis was performed for ~90 min at 180 V in cooling tap water bath. Cathode buffer: 40 mM TRIS-TRICINE pH 8.4, 0.005% Coomassie G-250; anode buffer: 50 mM TRIS-HCl pH 8.4; final sample buffer (with the dye): 75 mM KCl, 25 mM TRIS-HCl pH 8.4, 10% glycerol, 0.025% Coomassie Brilliant Blue G-250. Molar weight marker was Thermo Scientific™ PageRuler™ Prestained Protein Ladder, 10 to 180 kDa.

Native mass spectrometry

Lyophilized α-crystallin domains were reconstituted in 50 mM ammonium acetate. IbpA and IbpB α-crystallin domains were mixed in the selected ratios and incubated for at least 30 min at room temperature. Samples were subjected to native ESI mass spectrometry on TripleTOF 5600 + instrument.

Incorporation of p-benzoyl-phenylalanine (Bpa)

Method for unnatural amino acid incorporation was adapted from 36.

E. coli BL21(DE3) strain was used for the incorporation of p-benzoyl-phenylalanine into N-terminal parts of protein. IbpA variants were expressed from pET15b plasmid, and IbpB variants from pET28b plasmid. For other variants, due to low efficiency of incorporation and co-purification of truncated product an *E. coli* B95Δ*A*Δ*fabR* strain[50] lacking RF-1 release factor was used. Cells were grown in minimal media supplemented with casamino amino acids mixture, of exact composition: 42 mM Na_2HPO_4 * $2H_2O$, 22 mM KH_2PO_4 9.4 mM NH_4Cl, 8.6 mM NaCl, 340 mg/l thiamine, 0.2% casamino amino acids, 1% glycerol, 2 mM MgSO4, 0.1 mM CaCl2, ampicillin or kanamycin and pH adjusted to 7.4. Protein synthesis was induced by 1 mM IPTG in presence of 0.5 mM Bpa. Bpa incorporation machinery was

induced from the plasmid pEVOL-pBpF by 0.2% arabinose at the same time.

Purification of Bpa-containing sHsp variants

Cells were resuspended in 50 mM Tris pH 7.6, 6 M urea, 10% glycerol, 500 mM NaCl, 5 mM β-ME and disrupted by sonication. Lysate was clarified by centrifugation at 81,000*g* for 45 min. Supernatant was loaded onto a Ni-NTA column (Macherey-Nagel Protino NTA Agarose). The column was washed with 50 mM Tris pH 7.6, 6 M urea, 10% glycerol, 500 mM NaCl, 5 mM β-ME, 20 mM imidazole and the proteins were eluted with the 20–360 mM imidazole gradient in the same buffer. In case of unsatisfactory purity proteins were subjected to second round of purification on Ni-NTA resin in identical conditions, additionally purified with the use of Q-sepharose in a way similar to WT IbpA (for IbpA variants), or contaminants were removed by binding to Q-Sepharose resin, and pure protein was collected from flow-through fraction (for IbpB variants). Proteins were dialyzed to 50 mM Tris pH 7.6, 6 M urea, 10% glycerol, 500 mM NaCl, 5 mM β-ME prior to refolding. Protein variants were mixed with WT counterpart in 1:2 (IbpA:IbpB) molar ratio and dialyzed against buffer B with 1 mM $CaCl_2$. Proteins were refolded by stepwise dialysis into buffer C with 1 mM $CaCl_2$. His-Tag tags were removed similarly as for α-crystallin domains. Protein concentration was determined by SDS-PAGE electrophoresis with Coomassie blue staining using respective wild-type protein as a standard.

Analysis of Bpa-containing sHsp variants activity

Luciferase (1.5 μM) was denatured at 48 °C for 10 min in buffer E (50 mM Tris pH 7.4, 150 mM KCl, 20 mM Mg acetate, 5 mM DTT) in the presence or absence of 10 μM sHsps (3.3 μM IbpA and 6.7 μM IbpB, or analysed Bpa-containing IbpA or IbpB variants). Luciferase refolding was started by 35-fold dilution of denatured luciferase in the DnaK-ClpB chaperone cocktail (DnaK 0.7 μM, DnaJ 0.3 μM, GrpE 0.21 μM, ClpB 2.4 μM) in buffer E. All assays were performed in the presence of an ATP-regenerating system (18 mM creatine phosphate, 0.1 mg/ml creatine kinase, 5 mM ATP). The disaggregation reaction was carried out for 1 hour at 25 °C. The luciferase activity was measured in a Sirius Luminometer using Luciferase Assay system (Promega).

DLS measurements

Dynamic light scattering was measured using ZetaSizer NanoS instrument (Malvern). sHSP and α-crystallin domain samples were prepared in buffer A. Measurements were performed at 25 °C. At least three runs of 10 × 10 s measurements each were collected. From each run, 5 best measurements were averaged, and particle size distribution were calculated as percent volume of 70 discrete bins between 0.4 and 10,000 nm.

^{15}N and $^{13}C^{15}N$ labeling of α-crystallin domains

BL21(DE3)/pLys cells expressing $IbpA^{ACD}$ or $IbpB^{ACD}$ from pET28b plasmid were inoculated from overnight LB culture in 1:100 ratio into 400 ml of M9 minimal media (48 mM Na_2HPO_4; 22 mM KH_2PO_4; 8.6 mM NaCl; 7.6 mM $(NH_4)_2SO_4$; 2 mM $MgSO_4$; 0.3 mM $CaCl_2$; 50 μM $FeCl_3$; 10 μM $MnCl_2$; 10 μM $ZnSO_4$; 2 μM $CoCl_2$; 2 μM $CuCl_2$; 2 μM $NiCl_2$; 2 μM Na_2MoO_4; 2 μM Na_2SeO_3; 2 μM H_3BO_3, 0.5% (w/v) glucose (D-Glucose U-$^{13}C6$, 99% ^{13}C for carbon labelled samples); 1× BME vitamin stock and 0.1% (w/v) 99% ^{15}N $(NH_4)_2SO_4$. Cells were grown with shaking at 210 rpm at 37 °C for 8 hours in 2 l gauze-covered baffled flasks. After that, protein expression was induced by the addition of 1 mM IPTG. Temperature were reduced to 28 °C, and induction was carried out overnight.

NMR spectra acquisition

Lyophilized proteins were resuspended in 20 mM potassium phosphate buffer pH 6.8 or 7.2 (leading to additional presence of 20 mM KCl due to presence of KCl in lyophilized samples). Protein concentration was 5 mg/ml. NMR spectra were acquired at 298 K and 313 K on Avance 700 and 500 MHz Bruker spectrometers, all equipped with a triple resonance cryoprobe. The NMR spectra acquired for the backbone resonances assignment of $IbpB^{ACD}$ are summarized in Table S1 of Supplementary Materials. ^{15}N heteronuclear relaxation experiments on ^{15}N-labeled $IbpA^{ACD}$ and $IbpB^{ACD}$ were collected at 500 MHz at 313 K to measure ^{15}N backbone longitudinal and transverse relaxation rates and heteronuclear ^{15}N ^{1}H NOEs in order to provide information on overall reorientational correlation time of the protein. NMR titration experiments between ^{15}N-labeled $IbpA^{ACD}$ and $IbpB^{ACD}$ proteins and vice versa were performed at 313 K at 500 MHz. All NMR spectra were processed using the standard BRUKER software (Topspin 2.1), and analyzed through the CARA program. Prediction of the secondary structure elements were obtained from ^{15}N, $^{13}C'$, $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, using TALOS + program.[51] NMR chemical shift assignments data were deposited in the BioMagResBank (BMRB; accession number: 50799).

Mass spectrometry for crosslink identification and MS data analysis

Crosslinked and blank protein samples were run on 1.5 mm thick 15% glycerol-polyacrylamide

gels. Proteins were subjected to in-gel digestion with trypsin. Trypsin digestion and mass spectrometry were performed as service by Environmental Laboratory of Mass Spectrometry at Institute of Biophysics and Biochemistry of Polish Academy of Sciences. 1 h nanoLC runs with detection using Orbitrap Velos mass spectrometer operated in HCD mode were used for sample analysis.

All data were obtained as *.mgf peak lists and were analyzed using StavroX software.[38] Search database consisted of sequences of IbpA and IbpB protein sequences with appropriate additional residues resulting from His-tag removal. A score cutoff was adjusted to 100.

Simulation protocol

Simulations were performed using Gromacs 2018 package[52] with Plumed 2.5.2 plugin.[53] Charm-m36m force field[54] was used to represent proteins and TIP3P model was used for water. Simulations were performed in NPT ensemble with the temperature kept at 310 K with the v-rescale thermostat[55] and the pressure kept at 1 bar with the Parrinello-Rahman barostat.[56] The particle Mesh Ewald method[57] was used to calculate long-range electrostatic interactions with a cut-off radius of 1 nm and a grid spacing of ~0.1 nm. Van der Waals interactions were computed with Lennard-Jones potential with cut-off radius of 1 nm. LINCS algorithm was used to constrain all bonds involving hydrogen. The equations of motion were integrated with the leap-frog Verlet algorithm with a time step of 2 fs.

Structure prediction of the α-crystallin dimers

The initial models of the α-crystallin domains of IbpA (residues 32–123) and IbpB (residues 29–122) were obtained with I-TASSER[58] using a high-resolution crystal structure of Hsp20 (pdb ID 3GLA)[59] as a template. The initial structures of the IbpA^{ACD}-IbpB^{ACD} and IbpB^{ACD}-IbpB^{ACD} dimers were obtained by superimposing either IbpA^{ACD} and IbpB^{ACD} or two IbpB^{ACD} monomers on individual protomers of the Hsp20 dimer. Steric clashes were manually removed and the proteins were solvated with approx. 22,000 water molecules as well as Na^+ and Cl^- ions to neutralize the systems and provide physiological ionic strength. Then, the dimer structures were relaxed with energy-minimization and subsequent 1 μs run of conventional MD simulation. Next, the initially equilibrated structure of the dimer was refined by an integrative approach combining MD simulations with chemical shift data obtained titration of ^{15}N labelled IbpB^{ACD} with unlabeled IbpA^{ACD}.[60] To this end, we performed MD simulation with ensemble averaged restraint for 3 replicas with total ~1 μs and ~600 ns of trajectory length for IbpA^{ACD}-IbpB^{ACD} heterodimer and IbpB^{ACD}-IbpB^{ACD} homodimer, respectively. The additional restraining potential of the form $k\sum_i(\delta_i - \delta_{i,exp})^2$ was used, where δ_i is a replica-averaged chemical shift of the *i*-th nucleus computed based on the CamShift model,[61] $\delta_{i,exp}$ is the corresponding experimental value, and the summation runs over all ^{15}N and ^1H backbone resonances assigned for IbpB^{ACD}. During the first 280 ns, in case of the IbpA^{ACD}-IbpB^{ACD} dimer, the force constant k was gradually increased from 1 kJ ppm^{-2} to 13 kJ ppm^{-2} during the first 280 ns. For the IbpB^{ACD}-IbpB^{ACD} dimer, the force constant k was kept at 1 kJ ppm^{-2}. For each dimer, the last 50 ns of each replica trajectory were merged and clustered based on the backbone root mean square deviation (RMSD) with a cut-off of 0.25 nm. The structural model of IbpA^{ACD}-IbpB^{ACD} heterodimer was deposited in the ModelArchive (ma-2iypv).

Structure prediction of the IbpA-IbpB heterotetramer

The models of complete IbpA and IbpB proteins were prepared with I-Tasser by adding N- and C-terminal tails to the crystalline domains in the NMR-refined IbpA-IbpB dimer. The dimer was solvated with 25,236 water molecules and 83Na^+ and 74Cl^- ions. The system was equilibrated by energy minimization and 1 μs of conventional MD simulation. The resulting relaxed heterodimer was used in steered MD simulations to predict the structure of IbpA-IbpB heterotetramer by driving the distances corresponding to the experimentally determined crosslinks between crystalline domains in the complete IbpA and IbpB proteins. Initially, two spatially separated IbpA-IbpB heterodimers were placed in the simulation box such that the center-of-mass (COM) distances between the backbone atoms of residues 30, 60 and 124–126 of IbpA from the first heterodimer and residues 79–81, 72–77 and 65–69 of IbpB from the second heterodimer were equal to 1.4, 1.4 and 3.8 nm, respectively. The proteins were solvated with 106,676 water molecules and 326Na^+ and 308Cl^- ions. Next, 200 ns of steered MD simulation were performed, during which the centers of one-sided harmonic potentials with a spring constant of 2000 kJ nm^{-2} were moved at a constant speed from their initial COM distance values to 0.8 nm. At the same time, the conformation of the crystalline domains of individual protomers was kept by restraining their backbone RMSD to the initial conformations with a harmonic potential with a spring constant of 8000 kJ nm^{-2}. Next, all restraints were removed the system was subject to 500 ns of unbiased MD simulations. The first 100 ns of this simulation was discarded from further analysis as equilibration of the heterotetramer structure after steered-MD simulation. The remaining trajectory was clustered based on the protein backbone atoms with a cut-off of 0.3 nm, resulting in one dominant binding mode (82%) of IbpA-IpbB heterotetramer. The structural ensemble of the identified binding mode

was then used to compute inter-residue distances for each crosslinking pair identified in the experimental studies.

CRediT authorship contribution statement

Artur Piróg: Conceptualization, Investigation, Visualization, Writing. **Francesca Cantini:** Conceptualization, Investigation, Visualization, Writing. **Łukasz Nierzwicki:** Investigation, Visualization. **Igor Obuchowski:** Investigation, Visualization. **Bartłomiej Tomiczek:** Conceptualization, Investigation, Visualization. **Jacek Czub:** Conceptualization, Supervision, Writing. **Krzysztof Liberek:** Conceptualization, Supervision, Funding acquisition, Writing.

Acknowledgements

We thank dr. Katarzyna Macur for native mass spectrometry measurements and dr. Szymon Zietkiewicz and Igor Grochowina for discussions.

This work was supported by the grants of the Polish National Science Centre (2013/08/A/NZ1/00683) and (2019/33/B/NZ1/00352) to KL. NMR measurements were supported by iNEXT, grant number 653706, funded by the Horizon 2020 programme of the European Union, providing access to the CERM research infrastructure (PID 4552). Molecular dynamics simulations were supported by PL-Grid Infrastructure (http://www.plgrid.pl/en). I.O. is a recipient of START program (63.2020) from the Foundation for Polish Science.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.167054.

Received 3 February 2021;
Accepted 12 May 2021;
Available online 20 May 2021

Keywords:
bacterial small heat shock proteins;
chaperones;
modification of protein aggregation;
protein refolding;
heterodimer of sHsps

References

1. Ostankovitch, M., Buchner, J., (2015). The network of molecular chaperones: insights in the cellular proteostasis machinery. *J. Mol. Biol.*, **427**, 2899–2903.
2. Hartl, F.U., Bracher, A., Hayer-Hartl, M., (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, **475**, 324–332.
3. Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., Hartl, F.U., (2013). Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.*, **82**, 323–355.
4. Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E., Bukau, B., (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol. Microbiol.*, **50**, 585–595.
5. Liberek, K., Lewandowska, A., Zietkiewicz, S., (2008). Chaperones in control of protein disaggregation. *EMBO J.*, **27**, 328–335.
6. Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., et al., (2004). Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell*, **119**, 653–665.
7. Glover, J.R., Lindquist, S., (1998). Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell*, **94**, 73–82.
8. Haslbeck, M., Vierling, E., (2015). A first line of stress defense: small heat shock proteins and their function in protein homeostasis. *J. Mol. Biol.*, **427**, 1537–1548.
9. Haslbeck, M., Weinkauf, S., Buchner, J., (2019). Small heat shock proteins: Simplicity meets complexity. *J. Biol. Chem.*, **294**, 2121–2132.
10. Studer, S., Obrist, M., Lentze, N., Narberhaus, F., (2002). A critical motif for oligomerization and chaperone activity of bacterial alpha-heat shock proteins. *Eur. J. Biochem.*, **269**, 3578–3586.
11. Kim, K.K., Kim, R., Kim, S.H., (1998). Crystal structure of a small heat-shock protein. *Nature*, **394**, 595–599.
12. Benesch, J.L., Aquilina, J.A., Baldwin, A.J., Rekas, A., Stengel, F., Lindner, R.A., et al., (2010). The quaternary organization and dynamics of the molecular chaperone HSP26 are thermally regulated. *Chem. Biol.*, **17**, 1008–1017.
13. Shi, J., Koteiche, H.A., McHaourab, H.S., Stewart, P.L., (2006). Cryoelectron microscopy and EPR analysis of engineered symmetric and polydisperse Hsp16.5 assemblies reveals determinants of polydispersity and substrate binding. *J. Biol. Chem.*, **281**, 40420–40428.
14. Stengel, F., Baldwin, A.J., Painter, A.J., Jaya, N., Basha, E., Kay, L.E., et al., (2010). Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 2007–2012.
15. Delbecq, S.P., Klevit, R.E., (2013). One size does not fit all: the oligomeric states of alphaB crystallin. *FEBS Lett.*, **587**, 1073–1080.
16. Lentze, N., Aquilina, J.A., Lindbauer, M., Robinson, C.V., Narberhaus, F., (2004). Temperature and concentration-controlled dynamics of rhizobial small heat shock proteins. *Eur. J. Biochem.*, **271**, 2494–2503.
17. Painter, A.J., Jaya, N., Basha, E., Vierling, E., Robinson, C. V., Benesch, J.L., (2008). Real-time monitoring of protein complexes reveals their quaternary organization and dynamics. *Chem. Biol.*, **15**, 246–253.
18. Haslbeck, M., Walke, S., Stromer, T., Ehrnsperger, M., White, H.E., Chen, S., et al., (1999). Hsp26: a temperature-regulated chaperone. *EMBO J.*, **18**, 6744–6751.
19. Ratajczak, E., Zietkiewicz, S., Liberek, K., (2009). Distinct activities of Escherichia coli small heat shock proteins IbpA

and IbpB promote efficient protein disaggregation. *J. Mol. Biol.*, **386**, 178–189.
20. Zwirowski, S., Klosowska, A., Obuchowski, I., Nillegoda, N. B., Pirog, A., Zieztkiewicz, S., et al., (2017). Hsp70 displaces small heat shock proteins from aggregates to initiate protein refolding. *EMBO J.*, **36**, 783–796.
21. Ungelenk, S., Moayed, F., Ho, C.T., Grousl, T., Scharf, A., Mashaghi, A., et al., (2016). Small heat shock proteins sequester misfolding proteins in near-native conformation for cellular protection and efficient refolding. *Nature Commun.*, **7**, 13673.
22. Mogk, A., Schlieker, C., Friedrich, K.L., Schonfeld, H.J., Vierling, E., Bukau, B., (2003). Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK. *J. Biol. Chem.*, **278**, 31033–31042.
23. Obuchowski, I., Pirog, A., Stolarska, M., Tomiczek, B., Liberek, K., (2019). Duplicate divergence of two bacterial small heat shock proteins reduces the demand for Hsp70 in refolding of substrates. *PLoS Genet.*, **15**, e1008479.
24. Gaubig, L.C., Waldminghaus, T., Narberhaus, F., (2011). Multiple layers of control govern expression of the Escherichia coli ibpAB heat-shock operon. *Microbiology (Reading)*, **157**, 66–76.
25. Miwa, T., Chadani, Y., Taguchi, H., (2021). Escherichia coli small heat shock protein IbpA is an aggregation-sensor that self-regulates its own expression at post-transcriptional levels. *Mol. Microbiol.*, **115**, 142–156.
26. Allen, S.P., Polazzi, J.O., Gierse, J.K., Easton, A.M., (1992). Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in Escherichia coli. *J. Bacteriol.*, **174**, 6938–6947.
27. Laskowska, E., Wawrzynow, A., Taylor, A., (1996). IbpA and IbpB, the new heat-shock proteins, bind to endogenous Escherichia coli proteins aggregated intracellularly by heat shock. *Biochimie*, **78**, 117–122.
28. Kitagawa, M., Miyakawa, M., Matsumura, Y., Tsuchido, T., (2002). Escherichia coli small heat shock proteins, IbpA and IbpB, protect enzymes from inactivation by heat and oxidants. *Eur. J. Biochem.*, **269**, 2907–2917.
29. Matuszewska, M., Kuczynska-Wisnik, D., Laskowska, E., Liberek, K., (2005). The small heat shock protein IbpA of Escherichia coli cooperates with IbpB in stabilization of thermally aggregated proteins in a disaggregation competent state. *J. Biol. Chem.*, **280**, 12292–12298.
30. Bissonnette, S.A., Rivera-Rivera, I., Sauer, R.T., Baker, T. A., (2010). The IbpA and IbpB small heat-shock proteins are substrates of the AAA plus Lon protease. *Mol. Microbiol.*, **75**, 1539–1549.
31. Shi, X., Yan, L., Zhang, H., Sun, K., Chang, Z., Fu, X., (2014). Differential degradation for small heat shock proteins IbpA and IbpB is synchronized in Escherichia coli: implications for their functional cooperation in substrate refolding. *Biochem. Biophys. Res. Commun.*, **452**, 402–407.
32. Strozecka, J., Chrusciel, E., Gorna, E., Szymanska, A., Zietkiewicz, S., Liberek, K., (2012). Importance of N- and C-terminal regions of IbpA, Escherichia coli small heat shock protein, for chaperone function and oligomerization. *J. Biol. Chem.*, **287**, 2843–2853.
33. Ratajczak, E., Strozecka, J., Matuszewska, M., Zietkiewicz, S., Kuczynska-Wisnik, D., Laskowska, E., et al., (2010). IbpA the small heat shock protein from Escherichia coli forms fibrils in the absence of its cochaperone IbpB. *FEBS Lett.*, **584**, 2253–2257.
34. Shearstone, J.R., Baneyx, F., (1999). Biochemical characterization of the small heat shock protein IbpB from Escherichia coli. *J. Biol. Chem.*, **274**, 9937–9945.
35. Garcia de la Torre, J., Huertas, M.L., Carrasco, B., (2000). HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.*, **147**, 138–146.
36. Chin, J.W., Martin, A.B., King, D.S., Wang, L., Schultz, P. G., (2002). Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 11020–11024.
37. Fu, X., Shi, X., Yan, L., Zhang, H., Chang, Z., (2013). In vivo substrate diversity and preference of small heat shock protein IbpB as revealed by using a genetically incorporated photo-cross-linker. *J. Biol. Chem.*, **288**, 31646–31654.
38. Gotze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C. H., Krauth, F., et al., (2012). StavroX–a software for analyzing crosslinked products in protein interaction studies. *J. Am. Soc. Mass Spectrom.*, **23**, 76–87.
39. Forne, I., Ludwigsen, J., Imhof, A., Becker, P.B., Mueller-Planitz, F., (2012). Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. *Mol. Cell. Proteomics*, **11** (M111), 012088.
40. Kuczyńska-Wisńik, D., Laskowska, E., Taylor, A., (2001). Transcription of the ibpB heat-shock gene is under control of sigma(32)- and sigma(54)-promoters, a third regulon of heat-shock response. *Biochem. Biophys. Res. Commun.*, **284**, 57–64.
41. Shieh, Y.W., Minguez, P., Bork, P., Auburger, J.J., Guilbride, D.L., Kramer, G., et al., (2015). Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science*, **350**, 678–680.
42. Han, M.J., Park, S.J., Park, T.J., Lee, S.Y., (2004). Roles and applications of small heat shock proteins in the production of recombinant proteins in Escherichia coli. *Biotechnol. Bioeng.*, **88**, 426–436.
43. Mallik, S., Tawfik, D.S., (2020). Determining the interaction status and evolutionary fate of duplicated homomeric proteins. *PLoS Comput. Biol.*, **16**, e1008145.
44. Hochberg, G.K.A., Shepherd, D.A., Marklund, E.G., Santhanagoplan, I., Degiacomi, M.T., Laganowsky, A., et al., (2018). Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions. *Science*, **359**, 930–934.
45. Mymrikov, E.V., Riedl, M., Peters, C., Weinkauf, S., Haslbeck, M., Buchner, J., (2020). Regulation of small heat-shock proteins by hetero-oligomer formation. *J. Biol. Chem.*, **295**, 158–169.
46. Kuipers, B.J.H., Gruppen, H., (2007). Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *J. Agric. Food Chem.*, **55**, 5445–5451.
47. Zietkiewicz, S., Krzewska, J., Liberek, K., (2004). Successive and synergistic action of the Hsp70 and Hsp100 chaperones in protein disaggregation. *J. Biol. Chem.*, **279**, 44376–44383.
48. Woo, K.M., Kim, K.I., Goldberg, A.L., Ha, D.B., Chung, C. H., (1992). The heat-shock protein ClpB in Escherichia coli is a protein-activated ATPase. *J. Biol. Chem.*, **267**, 20429–20434.

49. Laemmli, U.K., (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
50. Mukai, T., Hoshi, H., Ohtake, K., Takahashi, M., Yamaguchi, A., Hayashi, A., et al., (2015). Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. *Sci. Rep.*, **5**, 9699.
51. Cornilescu, G., Delaglio, F., Bax, A., (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
52. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al., (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.*, **1–2**, 19–25.
53. Bonomi, M., Bussi, G., Camilloni, C., Tribello, G.A., Banáš, P., Barducci, A., et al., (2019). Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods*, **16**, 670–673.
54. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., et al., (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, **14**, 71–73.
55. Bussi, G., Donadio, D., Parrinello, M., (2007). Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 014101.
56. Parrinello, M., Rahman, A., (1981). Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
57. Darden, T., York, D., Pedersen, L., (1993). Particle Mesh Ewald - an N. Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.*, **98**, 10089–10092.
58. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., (2015). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, **12**, 7–8.
59. Hilario, E., Teixeira, E.C., Pedroso, G.A., Bertolini, M.C., Medrano, F.J., (2006). Crystallization and preliminary X-ray diffraction analysis of XAC1151, a small heat-shock protein from Xanthomonas axonopodis pv. citri belonging to the alpha-crystallin family. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, **62**, 446–448.
60. Camilloni, C., Robustelli, P., De Simone, A., Cavalli, A., Vendruscolo, M., (2012). Characterization of the conformational equilibrium between the two major substates of RNase A using NMR chemical shifts. *J. Am. Chem. Soc.*, **134**, 3968–3971.
61. Kohlhoff, K.J., Robustelli, P., Cavalli, A., Salvatella, X., Vendruscolo, M., (2009). Fast and accurate predictions of protein NMR chemical shifts from interatomic distances. *J. Am. Chem. Soc.*, **131**, 13894–13895.
62. Garrett, D.S., Seok, Y.J., Peterkofsky, A., Clore, G.M., Gronenborn, A.M., (1997). Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the Escherichia coli phosphotransferase system. *Biochemistry*, **36**, 4393–4398.