

ARTICLE

Received 25 Aug 2015 | Accepted 2 Dec 2015 | Published 20 Jan 2016

DOI: 10.1038/ncomms10357

OPEN

Protein unfolding as a switch from self-recognition to high-affinity client binding

Bastian Groitl¹, Scott Horowitz^{1,2}, Karl A.T. Makepeace³, Evgeniy V. Petrotchenko³, Christoph H. Borchers³, Dana Reichmann^{1,†}, James C.A. Bardwell^{1,2} & Ursula Jakob¹

Stress-specific activation of the chaperone Hsp33 requires the unfolding of a central linker region. This activation mechanism suggests an intriguing functional relationship between the chaperone's own partial unfolding and its ability to bind other partially folded client proteins. However, identifying where Hsp33 binds its clients has remained a major gap in our understanding of Hsp33's working mechanism. By using site-specific Fluorine-19 nuclear magnetic resonance experiments guided by *in vivo* crosslinking studies, we now reveal that the partial unfolding of Hsp33's linker region facilitates client binding to an amphipathic docking surface on Hsp33. Furthermore, our results provide experimental evidence for the direct involvement of conditionally disordered regions in unfolded protein binding. The observed structural similarities between Hsp33's own metastable linker region and client proteins present a possible model for how Hsp33 uses protein unfolding as a switch from self-recognition to high-affinity client binding.

¹Department of Molecular, Cellular and Developmental Biology, University of Michigan, 830 N-University Avenue, Ann Arbor, Michigan 48109-1048, USA.

²Howard Hughes Medical Institute, University of Michigan, 830 N-University Avenue, Ann Arbor, Michigan 48109-1048, USA. ³Department of Biochemistry and Microbiology, Genome BC Proteomics Centre, University of Victoria, 4464 Markham Street #3101, Victoria, British Columbia, Canada V8Z5N3.

† Present address: Department of Biological Chemistry, Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Givat Ram Campus, Jerusalem 91904, Israel. Correspondence and requests for materials should be addressed to S.H. (email: horowsah@umich.edu) or to U.J. (email: ujakob@umich.edu).

A number of molecular chaperones contain extensive regions of native disorder^{1,2}. The realization that these disordered regions are often found in regions centric to protein–protein interactions fuelled the idea that these chaperones might engage their non-structured regions directly in client recognition^{3,4}. Intrinsically disordered structures are characterized by significant flexibility, which may provide a possible explanation for the plasticity that chaperones exhibit in client binding. However, intrinsically disordered regions in proteins also tend to be highly charged^{5,6}. It therefore remains unclear how these disordered regions in chaperones recognize and bind-folding intermediates, which are characterized by exposed hydrophobic surfaces that are normally buried in the interior of fully folded proteins^{7–9}. Unfortunately, experiments addressing these important questions are technically challenging given that both chaperone and client proteins are partially unfolded and thus difficult to structurally characterize. Moreover, promiscuous client binding as is observed with many of these chaperones, often results in only very small chemical shift changes as measured by nuclear magnetic resonance (NMR)¹⁰.

Over the past few years, a group of stress-specific chaperones have been identified that use post-translational unfolding as the mechanism for their activation¹¹. Under non-stress conditions, these proteins are folded and chaperone-inactive. However, upon exposure to specific protein-denaturing conditions, such as hypochlorous acid (HOCl) (for example, Hsp33)¹, low pH (for example, HdeA)² or elevated temperatures (for example, Hsp26)¹², the proteins rapidly undergo extensive unfolding and become active as chaperones. These chaperones, appropriately termed conditionally disordered^{11,13,14}, therefore, exist in at least two distinct conformations—a folded inactive conformation and a partially unfolded active conformation^{1,2}—making them ideal model systems for studying the role of protein disorder in chaperone function.

In this study, we focused on the redox-regulated chaperone Hsp33, which is rapidly activated upon oxidative stress-mediated protein unfolding¹⁵. Central to Hsp33's activation process is the folding status of a conditionally disordered ~40-amino-acid linker region (Fig. 1a, light pink), which connects the N-terminal linker-binding domain (Fig. 1a, cyan) with the C-terminal redox switch domain (Fig. 1a, purple). Under reducing conditions, the linker region is stably folded, making extensive contacts with a hydrophobic surface on Hsp33's N-terminal domain (Fig. 1a). Stabilization of the linker domain appears to be further mediated by the redox switch domain, where four absolutely conserved cysteine thiols coordinate zinc with high affinity¹⁵ (Fig. 1a). Under oxidative stress conditions, disulfide bonds form, zinc is released and the linker region unfolds. This unfolding converts Hsp33 into a high-affinity chaperone that preferentially binds to partially unfolded client proteins^{15,16}. A single linker-destabilizing mutation constitutively activates Hsp33, providing evidence that linker unfolding is necessary and sufficient for the activation of Hsp33's chaperone function¹⁷.

To directly investigate the role of conditionally disordered regions in chaperone–client interactions, we used a multipronged approach, combining *in vitro* and *in vivo* crosslinking with site-directed Fluorine-19 NMR (¹⁹F NMR)/spin-label experiments. We discovered that stress-induced unfolding of Hsp33 exposes a composite amphipathic client-binding site consisting of a stable hydrophobic surface and Hsp33's own conditionally disordered linker region. Structural similarities between Hsp33's own metastable region and the client proteins explain how these chaperones use unfolding to switch from self-recognition to high-affinity client binding.

Results

Identification of Hsp33-client interaction sites *in vivo*.

To determine the interaction sites between Hsp33 and client proteins in the cellular context, we decided to incorporate the non-canonical amino acid *p*-benzoyl-L-phenylalanine (BPA) into specific locations in Hsp33 and activate it as a zero-length crosslinker via UV irradiation¹⁸, thereby crosslinking Hsp33 to its client proteins *in vivo*. We reasoned that these experiments could (1) help determine those regions in Hsp33 that come close to its client proteins upon activation of its chaperone function and (2) guide subsequent structural studies *in vitro*.

Wild-type Hsp33 is normally activated by exposing cells to HOCl (ref. 1). However, since HOCl also mediates non-specific protein crosslinking, we decided to introduce BPA into our temperature-regulated variant Hsp33^{M172S} instead of wild-type Hsp33 (ref. 17). Hsp33^{M172S} behaves like wild-type Hsp33 under oxidative unfolding conditions. However, this variant no longer requires HOCl treatment for its activation because of a destabilizing mutation in its N-terminus that causes unfolding of its linker domain and activation of its chaperone function at temperatures as low as 40 °C (ref. 17). This variant, thus, allows us to activate Hsp33's chaperone function *in vivo* by exposing cells to a simple temperature shift and avoids the use of HOCl. Once activated by heat-shock temperatures, the Hsp33^{M172S} variant binds to a wide range of thermally unfolding proteins and sequesters them into insoluble aggregates¹⁷.

Eighteen different sites for BPA incorporation in Hsp33^{M172S} were chosen (Fig. 1a). This selection was primarily based on the rationale that the substitution of the phenylalanine-like BPA for an existing Phe or a residue that was found substituted for Phe in at least one known Hsp33 homologue would be unlikely to substantially affect Hsp33's tertiary structure. We individually replaced the selected codons with an amber (TAG) stop codon and co-expressed the respective Hsp33^{TAG}-containing plasmids together with the pEVOL single-vector construct, which contains two copies of *Methanocaldococcus jannaschii* aminoacyl-tRNA synthetases and an optimized suppressor tRNA^{CUA}, using a Hsp33-deletion mutant of *Escherichia coli*¹⁹. All mutant strains expressed full-length Hsp33 in a BPA-dependent manner (Supplementary Fig. 1a).

We exposed live cells that expressed the BPA-substituted Hsp33 variants to UV irradiation at either 30 °C (Supplementary Fig. 1b) or for various times after shift to heat-shock temperatures (Fig. 1b). Subsequent western blot analyses of whole-cell extracts using antibodies against Hsp33 were used to monitor which proteins crosslink with Hsp33 and to what extent. As depicted in Fig. 1b and summarized in Supplementary Table 1, we found the largest extent of Hsp33 crosslinking species with mutant variants that contained BPA in or adjacent to the flexible linker region (F157, M172, F187, L202, L203, W212 and Y223). Modest amounts of crosslinking products were found at some additional sites in the N-terminus (Y12 and Y39), whereas most of the remaining BPA incorporation sites did not undergo any substantial crosslinking (Supplementary Table 1). To test whether the observed higher migrating bands are indeed crosslinking products between Hsp33 and client proteins, or represent crosslinks between Hsp33 molecules in higher oligomeric states, we purified two of our Bpa mutant variants, Hsp33^{M172SY145BPA} (negative in *in vivo* crosslinking experiments) and Hsp33^{M172SL202BPA} (positive in *in vivo* crosslinking experiments). *In vitro* crosslinking experiments revealed no additional bands apart from Hsp33 dimers migrating at ~70 kDa (Supplementary Fig. 1c), strongly suggesting that the higher migrating bands that we observed in our *in vivo* crosslinking studies do indeed represent crosslinks between activated Hsp33 and cellular client proteins.

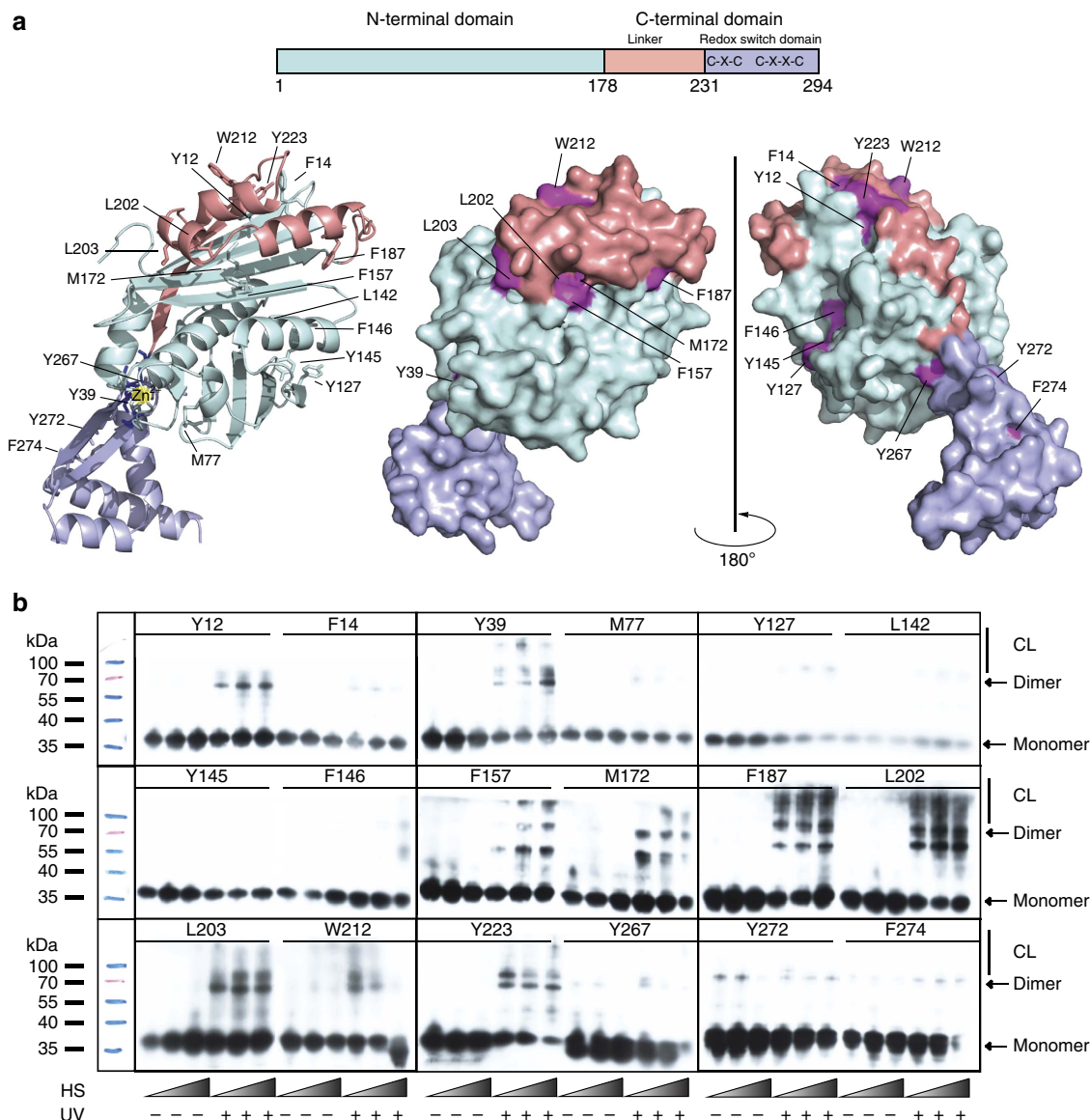


Figure 1 | Identification of the Hsp33-client-binding site *in vivo*. (a) Eighteen sites in Hsp33 were selected for the incorporation of unnatural amino acids as shown in the I-TASSER-based ribbon (left) and indicated in dark pink on the surface (right) models of reduced *E. coli* Hsp33. Hsp33 consists of an N-terminal domain (cyan), a metastable linker region (light pink) and a redox switch domain (purple), in which four thiolate anions arranged in a C₂₃₂-X-C₂₃₄-X₃₁-C₂₆₅-X-Y-C₂₆₈ motif coordinate one zinc ion (red sphere) under reducing conditions. Only one monomer of the dimeric crystal structure is shown. In solution, reduced Hsp33 is monomeric²⁸. (b) *E. coli* cells overexpressing the Hsp33^{M172S}-BPA variants were shifted from 30 °C to heat-shock conditions (43 °C for 5, 10 or 20 min) (HS). The cells were then either left untreated or exposed to UV irradiation for 10 min to induce crosslinking. Western blot analysis using anti-Hsp33 antibodies was used to visualize the crosslinking products (CL).

Note that most of the BPA mutant variants that crosslinked at all showed also a discrete crosslinking product, whose size corresponds with that of the crosslinked Hsp33 dimer (Fig. 1b; Supplementary Fig. 1c). As was the case with all other crosslinking species, this band only became detectable upon UV crosslinking (Fig. 1b and Supplementary Fig. 1b). This result was intriguing since, based on the available crystal structures of Hsp33 (PDB 1HW7), these sites are far away from each other, making crosslinking quite unlikely. The only published Hsp33 structure that places these crosslinked sites into close vicinity is the structure of an oxidized Hsp33¹⁻²⁵⁵ fragment (PDB 1HW7), in which the linker region of one subunit folds on top of the N-terminal linker-binding domain of the second subunit²⁰ (Supplementary Fig. 1d). This domain-swapped conformation could explain most of the Hsp33-Hsp33 crosslinks that we

observed *in vivo*. That the same Hsp33-BPA substitutions could crosslink either with each other or with client proteins suggests that the sites that Hsp33 uses for client binding under oxidative stress conditions overlap with those engaged in interactions between Hsp33's N-terminus and the linker region under non-stress conditions. Alternatively, Hsp33 crosslinks might be mediated by extended flexible loops coming transiently into sufficiently close vicinity. No substantial inter-Hsp33 or inter-client crosslinking bands were observed for Hsp33 variants containing BPA in a region of the N-terminus that lies opposite from the linker-binding surface (Y127, L142, Y145 and F146), or in the C-terminal redox switch domain (Y267, Y272 and F274). This finding agrees well with previous results, which indicated that the C-terminal redox switch domain has a primarily regulatory role and is not involved in substrate binding¹⁶.

These results provided a first clue as to the location of the client-binding site in Hsp33 and suggested that both the hydrophobic linker-binding surface as well as the flexible linker region might become directly involved in client binding following activation.

Monitoring conformational changes in Hsp33 by ^{19}F NMR.

Our *in vivo* crosslinking results suggested that Hsp33 can interact both with client proteins as well as with itself during heat-shock conditions *in vivo*. Crosslinking, however, is sensitive to regions of flexibility, non-equilibrium conditions and residue chemistry. To address these issues, we decided to use an independent approach to evaluate the interaction status of the residues we found using *in vivo* crosslinking. The strategy we employed is a novel application of *in vitro* ^{19}F NMR spectroscopy. We reasoned that by utilizing several of the same amber stop codon positions that we used for *in vivo* crosslinking, we could instead incorporate 4-(trifluoromethoxy)-phenylalanine (tFPA) (refs 21–23) at our preselected Hsp33 sites. Upon purification of these mutant variants, we would then be able to directly detect client binding by monitoring the ^{19}F NMR signal at these selected sites. ^{19}F NMR provides several advantages; ^{19}F is the natural isotope of fluorine and thus is present at 100% natural abundance, features high chemical shift dispersion and is one of the most sensitive NMR nuclei available in nature, with a sensitivity of 83% compared with ^1H , the most widely used isotope in protein NMR²⁴. Furthermore, ^{19}F is a strong reporter on changes in local chemical environments, including solvent exposure²⁵. These features, combined with the mobility of the trifluoromethoxy-group, make tFPA an extraordinarily sensitive chemical shift probe for monitoring binding or conformational changes^{26,27}. Most importantly, because of the site-selective nature of the amber stop codon strategy, there are no ambiguities in resonance assignment, as only one amino acid is labelled at a time, and it does not suffer from the same ambiguities as crosslinking arising from reactivity constraints. This combination of advantageous traits should make this method particularly suitable for disordered proteins, whose NMR spectra are typically extremely challenging to interpret²⁸ (Supplementary Fig. 2a,b). Use of a simple paramagnetic relaxation enhancement (PRE) experiment with our Hsp33-tFPA variants (see below for details) should be able to distinguish direct binding sites from areas of conformational changes within Hsp33.

We decided to focus on clearly positive *in vivo* crosslinking sites (Y12, Y39, F157, F187 and L202), one negative site (Y145) and one ambiguous site (W212) (Fig. 1b). We successfully purified all of the proteins except Hsp33 variants harbouring non-natural amino acids at either position Y12 or Y39. These two protein variants could not be expressed in sufficiently high quantities. We then tested the purified proteins for their *in vitro* chaperone function under both reducing and oxidizing conditions (Fig. 2a). All of the tested proteins showed wild-type-like activity when oxidized and disulfide bonded, indicating that client binding was not substantially affected by the introduction of tFPA at these sites. Moreover, although the activity of the reduced tFPA mutant variants was generally higher at 30 °C than the parental Hsp33^{M172S} mutant protein, all but one of our mutant proteins (Hsp33^{M172SY145tFPA}) maintained some degree of its temperature-induced chaperone activation. Their higher than parental activity at 30 °C suggests that the tFPA substitutions act to further destabilize the linker region. This idea is consistent with the finding that most of our crosslinking-positive Hsp33-BPA variants also showed some self-crosslinking in addition to client protein crosslinking at 30 °C (Supplementary Fig. 1b). One mutant variant, Hsp33^{M172SY145tFPA} was chaperone-inactive at both temperatures. This result agrees well with our *in vivo* data,

where we did not identify any crosslinking products at either 30 or 43 °C, despite the fact that the chaperone is active when oxidized (Fig. 1b and Supplementary Fig. 1b). This finding suggested that the Y145tFPA mutation suppresses the temperature regulation of the Hsp33^{M172S} mutant variant.

Only two of our mutant variants, Y145tFPA and W212tFPA, revealed a single peak in the ^{19}F NMR spectrum at 25 °C, as would be expected from a single labelled residue in a homogeneous environment (Fig. 2b). The three other variants, however, displayed two resonances in their 1D- ^{19}F NMR spectrum. These results strongly suggested that the three fluorinated residues, which reside in the linker-region docking surface of Hsp33's N-terminus (F157) or the linker region (F187 and L202), exist in at least two chemical environments that interchange slowly on the NMR timescale (millisecond exchange rate or slower). For all three variants, we found that the chemical shift of the up-field peaks was similar to that of the unbound tFPA (Fig. 2b). The simplest interpretation of these results is that the up-field peak corresponds to a conformation in which the tFPA is surface exposed, whereas the down-field peak corresponds to a more buried chemical environment (Fig. 2b). We thus hypothesized that the respective up-field resonances represent the unfolded, chaperone-active state of Hsp33, and the down-field peaks represent the folded, chaperone-inactive state of Hsp33. This hypothesis was further supported by quantum mechanics (QM) chemical shift calculations, which allowed us to quantify the effects of neighbouring residues on the chemical shifts. QM calculations demonstrated that small changes in the environment or the bond angle of our tFPA probe would result in measurable differences for the chemical shift (Supplementary Fig. 2c,d). These calculations revealed that in the case of F157 and F187, the protein interior indeed causes a down-field chemical shift change because of the close proximity to the side chains of R155 and R159 in the folded state (Fig. 2c, left). We tested this hypothesis by monitoring the ^{19}F NMR signals of F157tFPA and F187tFPA as a function of temperature and found that the area of the putative active-state peak indeed reversibly increased in size with increasing temperature and unfolding, whereas the area of the putative inactive state decreased (Fig. 2b, compare blue and red traces). No change in peak area or height upon shift to higher temperatures was detected for unbound tFPA. These results are in excellent agreement with our previous results¹⁷, which showed that the linker region in Hsp33^{M172S} unfolds in a temperature-dependent manner, leading to the activation of the protein. Thus, we concluded that for both F157 and F187, the up-field resonance peak represents the active Hsp33 chaperone with the linker region freed from the surface of Hsp33, and the down-field peak represents the closed, inactive state of Hsp33, in which the binding site is closed off from solution and brought in immediate contact with R155 and R159. For L202, two peaks were also detected in the NMR spectrum that changed size with temperature, potentially also representing the active, open state and the closed, inactive state, similar to F157 and F187. In contrast, increasing the temperature of the W212-tFPA mutant variant led to the appearance of an additional down-field resonance peak, which could potentially represent a second active chemical environment (Fig. 2b).

Monitoring Hsp33-client interactions *in vitro*. With the active and inactive states of Hsp33 identified using our tFPA substitutions, we proceeded to test how client binding affects the respective resonances in the ^{19}F NMR spectra. For these experiments, we used the soluble client peptide neuropeptide Y (SNPY) as a Hsp33 substrate¹⁶ (Supplementary Fig. 3a). Changes in chemical shift or line broadening upon addition of neuropeptide

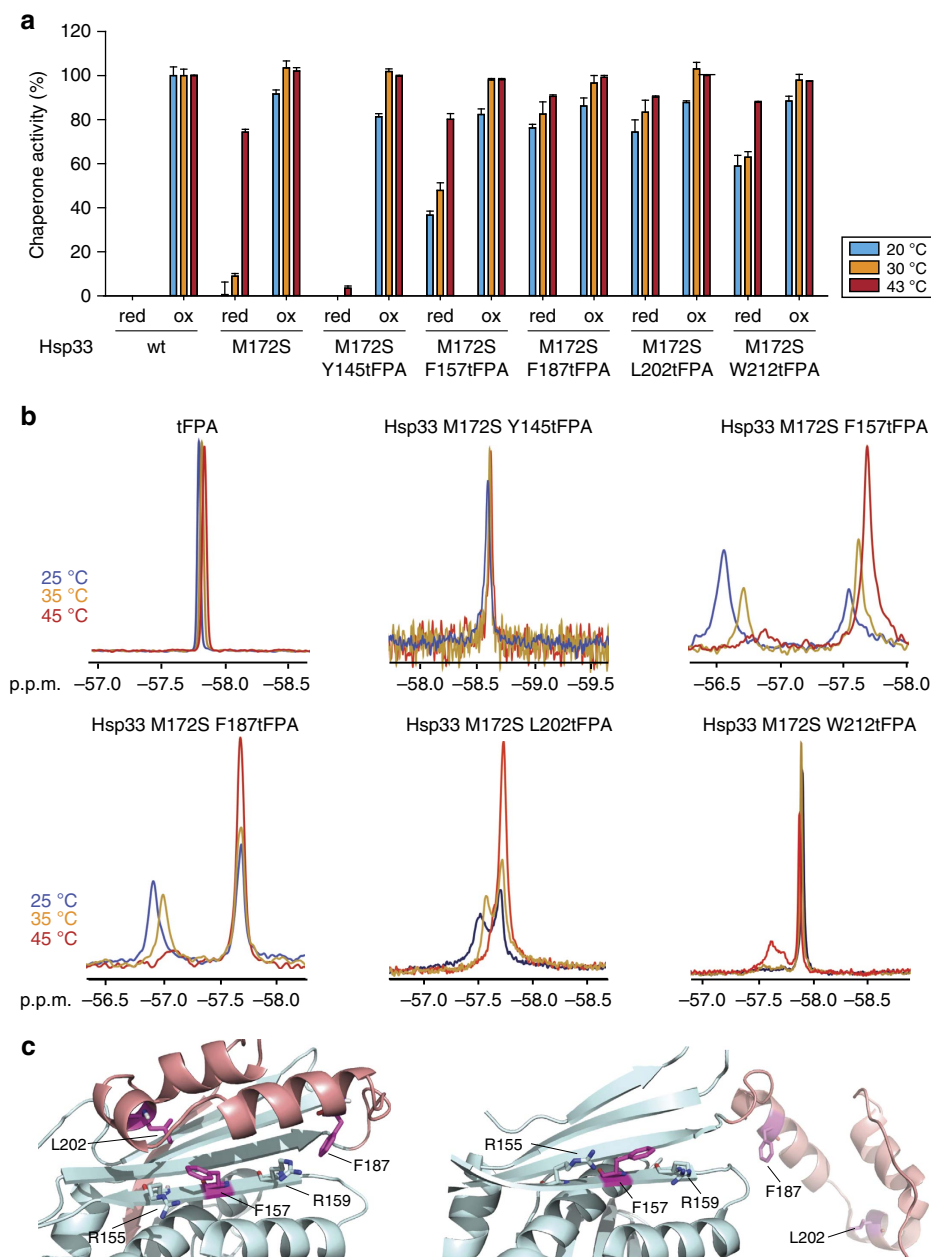


Figure 2 | Monitoring conformational rearrangements in purified Hsp33^{M172-tFPA} variants using ¹⁹F NMR. (a) Chaperone activity of reduced, zinc-reconstituted or HOCl-activated wild-type Hsp33, Hsp33^{M172S} or Hsp33^{M172S-tFPA} variants. Chaperone activity was measured by testing the influence of a four-fold molar excess of Hsp33 on the aggregation of chemically unfolded CS at either 20 °C (blue bars) or 30 °C (orange bars) or on thermally unfolded CS at 43 °C (red bars). Chaperone activity of 0% is defined as the light-scattering signal 4 min after addition of CS in the absence of chaperones. Activity of 100% corresponds to the light-scattering signal of CS in the presence of a four-fold molar excess of wild-type Hsp33 that had been activated for 2 min in 200 μM HOCl at 30 °C. All experiments were conducted at least 3–5 times and the s.e.m. is shown. (b) Temperature dependence of the ¹⁹F NMR signal in select Hsp33^{M172S-tFPA} mutants. ¹⁹F NMR spectra of tFPA alone or the indicated mutant variants were recorded at either 25 °C (blue), 35 °C (orange) or 45 °C (red). (c) N-terminal linker-docking surface (cyan) and the metastable linker region (pink) of *E. coli* Hsp33 in the inactive, closed state (left) (I-TASSER model) and in the activated, open state (right) (PDB 1HW7). The close proximity of F157 and F187 (and L202) to Arg155 and Arg159 in the closed state of Hsp33 is likely responsible for the distinctive down-field chemical shift change observed in the mutant variants under inactivating conditions.

Y (NPY) would indicate differences in the local environment of tFPA-labelled residues or in the dynamics of the protein, respectively. Upon titrating NPY into a solution containing the ¹⁹F-labelled chaperone variants, we observed distinct chemical shift changes and increased line broadening for the up-field active-state peaks of both F157 and F187 (Fig. 3a). In contrast, the inactive states (that is, down-field peaks) displayed comparatively little to no line broadening or chemical shift changes, consistent with the idea that clients bind to the more unfolded, active state

of Hsp33. Furthermore, upon addition of NPY, the peak area of the presumed inactive state of both F157 and F187 slightly decreased, whereas the peak size of the presumed active state increased, indicating that the addition of the chaperone substrate NPY shifts the equilibrium in the direction of active client-bound Hsp33. The Y145tFPA variant displayed no chemical shift or linewidth changes upon addition of NPY (Supplementary Fig. 3b), consistent with the finding that this mutant variant is inactive under these conditions (Supplementary Fig. 3a). Instead,

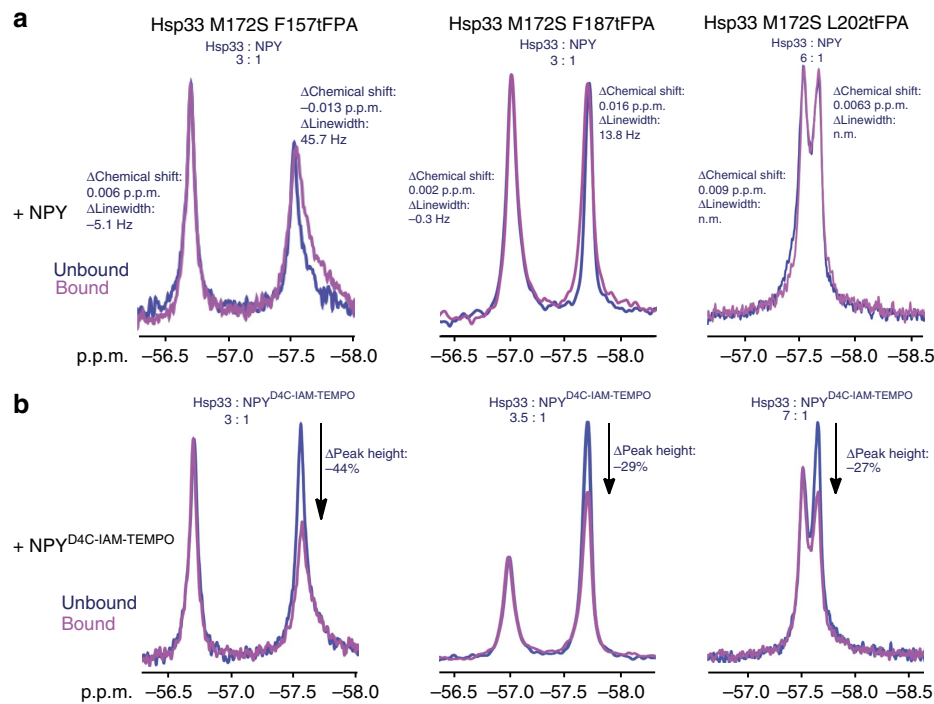


Figure 3 | Monitoring Hsp33-client-binding interactions using site-specific ^{19}F NMR. ^{19}F NMR spectra of select Hsp33^{M172StFPA} variants in the absence (blue) or presence (magenta) of (a) NPY or (b) NPY labelled with the paramagnetic spin-label TEMPO. The incubation temperature was set to 35 °C. All experiments were conducted at least three times.

upon addition of NPY, we observed significant precipitation of this mutant variant, which likely explains the decrease in peak size (Supplementary Fig. 3b). The peaks for the L202tFPA (Fig. 3a) and W212tFPA (Supplementary Fig. 3b) variants displayed only very small chemical shift changes and no detectable linewidth changes, leaving their role in binding still unclear.

To directly identify whether L202tFPA and W212tFPA are at client-binding sites, and to confirm that F157 and F187 indeed represent areas of direct client binding and do not instead show chemical changes and line broadening just because of conformational rearrangements in Hsp33, we employed a NPY analogue (NPY^{D4C}) labelled with a PRE tag. PRE tags line-broaden residues that are close in space to the tag, affecting residues in a distance-dependent fashion²⁹. For this purpose, we chose 4-(2-iodoacetamido)-TEMPO (IAM-TEMPO), which we conjugated to an engineered cysteine residue in NPY (NPY^{D4C}). The PRE tag did not substantially affect binding of NPY to our mutant Hsp33 variants, as tested by competition assays (Supplementary Fig. 3a). To evaluate the line-broadening effect of the PRE tag, we measured the relative peak-height change of the Hsp33-tFPA signals upon binding NPY^{D4C}-IAM-TEMPO versus binding unlabelled NPY. Analysis of peak heights is equivalent to evaluating line broadening as long as the peaks are properly normalized, and is commonly used in biomolecular NMR analysis of PREs. Notably, we found that addition of sub-stoichiometric concentrations of NPY^{D4C}-IAM-TEMPO to either Hsp33-F157tFPA or Hsp33-F187tFPA caused a reproducible loss in peak height of only the active-state resonances for each of the residues, implicating F157 and F187 as sites of direct client binding (Fig. 3b, Supplementary Fig. 3d,f). No pronounced change in the ratio between the active or inactive peak was observed when buffer alone was added to the proteins (Supplementary Fig. 3e), while titration of the NPY^{D4C}-IAM-TEMPO peptide revealed a peptide-concentration-dependent decrease in the peak height of the active-state

resonance (Supplementary Fig. 3f). Using the extent of PRE line broadening and the stoichiometry of binding (Fig. 3b)³⁰ we estimated an upper bound for the average interaction distance between NPY and residues F157 and F187 of ~13–15 Å. Similar levels of PRE line broadening (that is, decrease in peak height) were also observed for the presumed active-state peak of Hsp33-L202tFPA (Fig. 3b), whereas the Y145tFPA and W212tFPA substitutions did not reveal any line broadening upon addition of NPY^{D4C}-IAM-TEMPO, indicating that these residues are not directly involved in substrate interaction (Supplementary Fig. 3c). These results show that by using unnatural amino acid substitutions followed by *in vivo* crosslinking, ^{19}F NMR, and spin-label experiments, we obtained valuable insights into Hsp33's activation and client-binding process, and identified at least three sites in Hsp33 that are indeed bona fide client interaction sites—one in the linker-docking surface (F157) and two directly in the linker region of Hsp33 (F187 and L202).

***In vitro* crosslinking reveals client-binding sites in Hsp33.** The number and identity of residues that are chosen for substitution in a protein clearly limit the scope of *in vivo* crosslinking studies. To use an alternative approach that does not rely on amino acid substitution choices, we therefore conducted *in vitro* crosslinking studies followed by mass spectrometric analysis. We used crosslinking reagents that vary in their spacer distance (14, 7 and 0 Å) and performed crosslinking experiments with reduced inactive or oxidized active wild-type Hsp33 using NPY as the client peptide (Supplementary Fig. 4a). Digestion of the protein mixtures with either protease K or trypsin combined with tandem mass spectrometric (MS/MS) analysis was used to identify the crosslinking products. To facilitate the identification of Hsp33-NPY crosslinked peptides, we either used isotopically coded crosslinkers, such as collision-induced dissociation-cleavable amine-reactive 14 Å-length cyanur-biotin-dimercapto-propionyl-succinimide (CBDPS-H₈/D₈; ref. 31) or photoreactive 7 Å-length azido-benzoic-acid-succinimide (ABAS-¹²C₆/¹³C₆; ref. 32), or

worked with an equimolar mixture of ^{14}N - and ^{15}N -metabolically labelled Hsp33 in the case of the zero-length crosslinker (3-dimethylaminopropyl)carbo diimide (EDC) (flow chart shown in Supplementary Fig. 4b). Analysis of the crosslinking products by MS/MS analysis revealed a number of crosslinks between activated Hsp33_{ox} and NPY, which we did not observe when we tested inactive Hsp33_{red} (see Supplementary Data for the complete list of identified *in vitro* crosslinking products and MS/MS analyses). Several of the amino acids in Hsp33_{ox} that we found to crosslink with NPY *in vitro* were independently identified using at least two different crosslinkers (Fig. 4a and Table 1, in bold). Residues Y12, K44, K62 and K198 are of special

interest since they were either identified as *in vivo* crosslinking sites in this study (Fig. 4b,c), or positively identified as residues in Hsp33_{ox} that are more protected from tryptic digest when in complex with thermally unfolded luciferase¹⁶ (Fig. 5a). Moreover, K198, which is located in the linker region between Y187 and L202, was found to crosslink with the zero-length crosslinker EDC (Fig. 4c). This finding indicates a very close proximity between Hsp33's linker region and its client peptide, NPY. Most other crosslinks between Hsp33_{ox} and NPY were found with residues of the linker-docking surface of the N-terminal domain (Fig. 4b,c). Taken together, these experiments provide independent evidence that the chaperone Hsp33 uses residues of

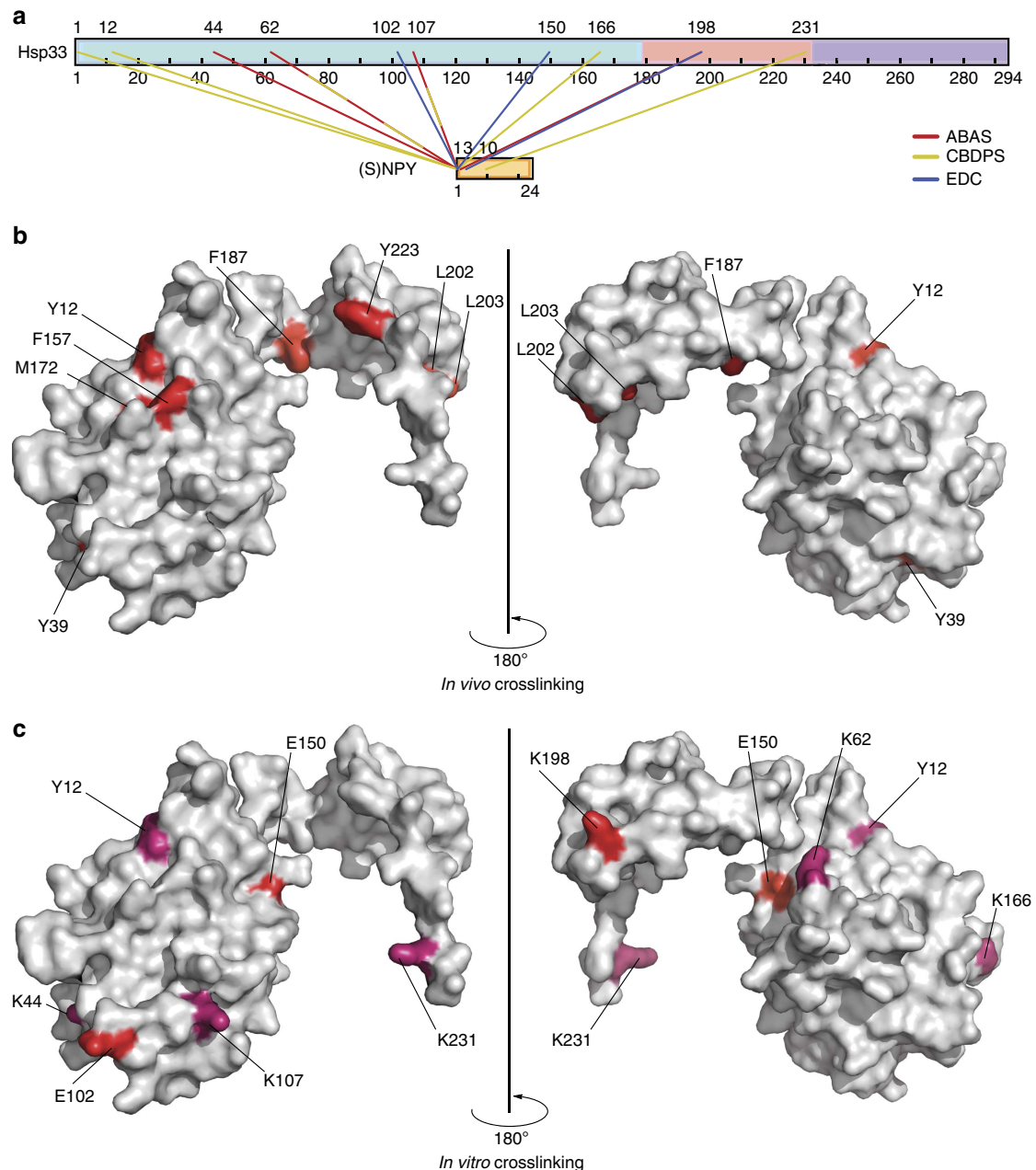


Figure 4 | Determining the Hsp33-client interaction sites using *in vitro* crosslinking. (a) *In vitro* crosslinks specifically between oxidized (activated) wild-type Hsp33 and (S)NPY-peptide are displayed in a linear representation created with Cross-Link Viewer⁵³ (SNPY is the NPY peptide containing one additional serine residue at the N-terminus). Residues that crosslinked with more than one crosslinker are shown as dashed lines in the respective colours (ABAS: red; CBDPS: yellow; and EDC: blue). (b) *In vivo* and (c) *in vitro* crosslinking sites are indicated on the crystal structure of oxidized, domain-swapped *E. coli* Hsp33¹⁻²⁵⁵ (PDB 1HW7). Only one of the two subunits that are found in the crystal structure is shown. Zero-length *in vitro* crosslinks with EDC are indicated in red; long- and medium-range crosslinks (CBDPS and ABAS) are shown in dark pink.

its linker-docking region as well as its natively disordered linker region to directly bind its unfolded client proteins. To verify that the linker-docking region indeed provides a favourable surface for the binding of client proteins, we performed protein–protein

docking simulations, in which the linker region of Hsp33 was removed, exposing the docking site to solvent (Fig. 5b). Notably, using a blind docking procedure in which no experimental or computational constraints were added to identify the binding site, all of the top 10 docking hits localized NPY to this linker-docking region. This docking simulation supports our ^{19}F NMR-binding data that the linker-docking region provides a favourable surface for binding partially folded substrates.

Table 1 | Summary of identified Hsp33_{ox}-NPY crosslinked peptides.

Crosslinker	Protease	Modified amino acid*	Hsp33 peptide sequence
ABAS (7 Å)	PK	44	E ₃₅ NHDYPQPVK _{N45}
	PK	62	L ₆₁ K FDGD ₆₆
	PK	62	L ₆₁ K FDGDIT ₆₈
	PK	107	G ₉₈ EIPENADL K ₁₀₇
	PK	198	T ₁₉₄ ETI K ₁₉₈
CBDPS (14 Å)	T	1	M ₁ IMPQHDQLHR ₁₁
	T	12	Y ₁₂ LFENFAVR ₂₀
	PK	62	L ₆₁ K FD ₆₄
	PK	107	G ₉₈ EIPENADL K ₁₀₇
	PK	166	G ₁₆₅ K PA ₁₆₈
	PK	231	D ₂₂₇ VEF K ₂₃₁
EDC (0 Å)	T	102	V ₉₆ QGEIP E NADL K ₁₀₇
	T	150	S ₁₄₉ E QLP T R ₁₅₅
	PK	198	T ₁₉₆ K TEELLTLPANEV ₂₁₀

ABAS, azido-benzoic-acid-succinimide; CBDPS, cyanur-biotin-dimercapto-propionyl-succinimide; EDC, (3-dimethylaminopropyl)carbo diimide; NPY, neuropeptide Y; PK, protease K; T, trypsin.
*Bold indicates residue was identified by at least two different crosslinkers. Bold and italicized alphabets indicate amino acids in Hsp33 peptides that were found crosslinked to NPY.

Discussion

A common feature of a growing group of chaperones, such as Hsp33, HdeA/HdeB and small heat-shock proteins, is their ability to specifically activate their chaperone function through stress-induced conformational rearrangements, unfolding and/or changes in oligomerization states^{2,12,33}. By using these mechanisms, these so-called stress-specific chaperones are activated only under stress conditions that cause protein unfolding and thus demand the presence of an additional chaperone activity¹¹. Once in their activated state, these chaperones bind potently and non-discriminatingly to a large number of protein folding intermediates. Recent evidence suggested that the metastable regions, whose folding status is central to their activation process, do not simply constitute molecular gates that control accessibility to potentially hydrophobic client-binding sites^{2,28,33}, but participate in the binding of unfolding client proteins. However, structural studies that would allow us and others to unambiguously address these questions rely on homogeneous and highly stable protein

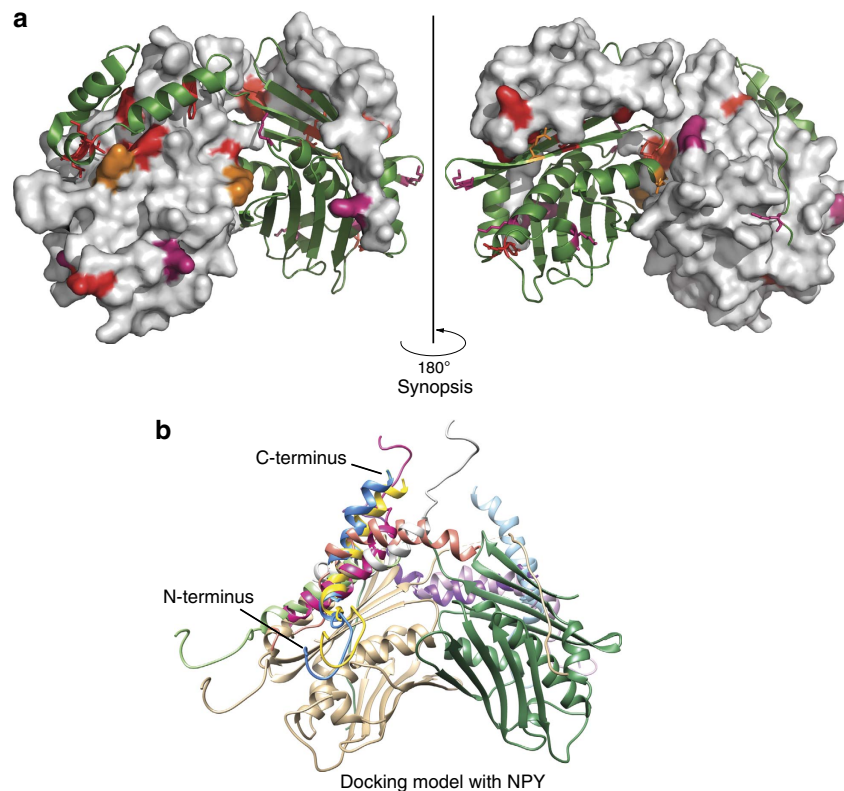


Figure 5 | The client-binding site in Hsp33. (a) Structure of domain-swapped *E. coli* Hsp33¹⁻²⁵⁵ (PDB 1HW7) with one monomer in surface representation and the other one in cartoon depiction. All *in vivo* crosslinking-positive residues, all zero-length *in vitro* crosslinking sites, and ^{19}F NMR-positive sites are highlighted in red. Long- and medium-range *in vitro* crosslinking sites are marked in dark pink, and sites previously suggested to be involved in client binding by limited proteolysis experiments¹⁶ are depicted in orange. Most identified client interaction sites overlap with interaction sites between Hsp33's linker-docking region and the linker region. (b) Docking model of NPY peptide (PDB 1RON) using the truncated oxidized Hsp33 model (residues 182-219 were removed from PDB 1HW7). The top ten docking models are shown in various colours. The Hsp33 subunits are depicted in green and gold, respectively.

complexes, and require participants that are folded and well-behaved. In the case of conditionally disordered chaperone–client complexes, both chaperones and client proteins are at least partially unstructured, and the folding status of neither partner is well-defined in the bound state. Methodologies that have been used in the past to track down client protein-binding sites in chaperones involve limited proteolysis and hydrogen/deuterium (H/D) exchange measurements combined with MS/MS analysis^{16,34,35}. With these approaches, changes in the accessibility of proteolytic sites or degrees/rates of H/D exchange in the absence and presence of client proteins are used as an indication of conformational changes in the chaperone that occur upon client binding. These studies, although very powerful in identifying sites of major conformational disturbance, do not distinguish between direct binding events and structural changes that take place in response to the client binding somewhere else in the protein. This limitation became apparent in recent studies on Hsp33, in which client-induced folding and unfolding events were observed in the linker region of Hsp33 but evidence for direct interactions between client proteins and the conditionally disordered linker region were missing¹⁶.

To address this question, which indeed would appear to pertain to many known chaperone–client protein complexes, we decided to employ a novel approach, incorporating unnatural amino acids into Hsp33 to perform both *in vivo* crosslinking and ¹⁹F NMR, using the same amber stop codon mutations. This method, followed by unbiased *in vitro* crosslinking experiments, unequivocally revealed a composite client-binding site in Hsp33 that consists of polar residues from the flexible linker region and nonpolar residues from the N-terminal linker-docking surface of Hsp33 (Fig. 5a). These results agreed in part with our previous limited proteolysis and H/D exchange studies, which demonstrated that client binding to activated Hsp33 limits access of proteolytic sites and significantly affects the thermodynamic stability of the linker region, suggesting that the linker region is involved in client binding¹⁶. In contrast, however, our previous study provided very little evidence for the involvement of the hydrophobic linker-binding platform in client binding, whereas the crosslinking experiments conducted in this study unequivocally pointed to a substantial involvement of this region in Hsp33–client interactions. The reason for this apparent discrepancy is likely found in the technical limitations of our previous techniques, including the absence of testable proteolytic cleavage sites in the hydrophobic linker-binding platform, and the possibility that H/D exchange measurements lack the necessary sensitivity to detect some of these interactions. In either case, our results that Hsp33 uses a composite binding site composed of a flexible polar linker region and a stable hydrophobic binding surface, are highly reminiscent of recent reports on the mechanism of substrate recognition by the canonical chaperone Tric/CCT (ref. 36). In their study, the authors postulate that charged–charged interactions between the chaperonin and its client protein confer specificity to binding, while nonpolar contacts allow for tight packing and additional stabilization of the complex. In analogy to this mechanism, it is tempting to speculate that Hsp33 uses its polar and highly flexible linker region to undergo initial electrostatic interactions with client proteins, thereby positioning and guiding them to the nonpolar N-terminal surface for additional hydrophobic interactions and increased stability. Subsequent rearrangement and refolding of the linker region¹⁶ might further contribute to the stabilization of the complex. These results suggest that unfolded protein recognition by chaperones might follow principles that go way beyond simple hydrophobicity, a force that has up to now been thought to drive most chaperone–client interactions³⁷.

Our studies suggest that Hsp33 uses many of the same interactions between its linker region and the linker-docking domain under reducing, non-stress conditions to engage with partially folded client proteins under oxidizing, activating conditions. These results could explain why activated Hsp33 binds preferentially to client proteins with substantial amounts of secondary structure elements¹⁶, since they could resemble the conformation of Hsp33's own linker region under reducing conditions. Intriguingly, the flexible linker region in the oxidized Hsp33 dimer shows partial helicity, similar to the peptide substrate model, NPY, both when free in solution and in the docking model. The similarity between these two sets of conformations and their binding modes to Hsp33 strongly suggests that the domain swap observed in oxidized Hsp33 dimers in the absence of client proteins recapitulates substrate binding. Furthermore, Hsp33's clients bind both the flexible linker as well as its docking platform, suggesting a potential model for how Hsp33 can effectively transfer partially folded clients to foldases such as the DnaK/DnaJ/GrpE system through multiple partially overlapping binding sites extending away from Hsp33's surface. As such, exposed hydrophobic residues of the unfolded protein could be protected by the docking platform, but then transit the flexible linker to the waiting refolding system.

These results uncover the power of the amber stop codon suppression approach to untangle difficult heterogeneous or dynamic experimental systems. The site-labelling strategy presented here combines *in vivo* and *in vitro* probing methods in a generalizable approach. This strategy could conceivably unravel many such experimentally difficult binding interactions in conditionally folded or partially folded proteins, where traditional techniques cannot overcome the complex barriers posed by the underlying biology.

Methods

Strains and plasmids. The incorporation of amber (TAG) stop codons at individual preselected sites in Hsp33^{M172S} was performed by site-directed mutagenesis using pET11a *hslO*^{M172S} as template¹⁷. After sequencing the plasmids, the inserts were recloned into respective pET21b to generate His-tagged versions of the Hsp33^{M172S} variants. All plasmids were transformed into the *E. coli* expression strain BL21 *hslO*⁻. A complete list of strains and plasmids used in this study can be found in Supplementary Table 2.

***In vivo* crosslinking.** pET11a plasmids containing the respective Hsp33^{M172S}-amber mutants were co-transformed with pEVOL (ref. 19), which contains two copies of *M. jannaschii* aminoacyl-tRNA synthetases and an optimized suppressor tRNA^{CUA}, into BL21 Δ *hslO* (a Hsp33-deleted mutant) of *E. coli*¹⁹. We then grew the strains in the absence and presence of BPA and confirmed that full-length Hsp33 mutant variants were only produced when BPA was supplied to the media (Supplementary Fig. 1a). These strains were plated on lysogeny broth (LB) plates supplemented with 200 mg ml⁻¹ ampicillin and 34 mg ml⁻¹ chloramphenicol. The next day, transformants were scraped off the plates and used to inoculate 5 ml of LB medium supplemented with 1 mM ZnCl₂, 1 mM BPA (Bachem Holding AG), 200 mg ml⁻¹ ampicillin and 34 mg ml⁻¹ chloramphenicol to an OD₆₀₀ of 0.4 (ref. 38). After 1 h of incubation at 37 °C, cells were shifted to 30 °C and protein expression was induced with 0.2 mM IPTG and 0.2% L-arabinose. Cells were harvested after 6 h, washed twice and resuspended in phosphate-buffered saline, pH 7.4 using 100 μ l of phosphate-buffered saline per 1 OD₆₀₀ unit. Cells were then incubated at either 30 °C or heat-shock temperature (43 °C) for the indicated times with shaking (400 r.p.m.). Immediately after incubation, 100 μ l cells were transferred into 96-well plates and placed on ice. The samples were crosslinked for 10 min using a 25 W, 365 nm UV lamp (UVGL-58; 115 V, 60 Hz, 0.16 A). UV irradiation was applied from the top, using a distance of ~2 cm from the 96-well plates. After UV crosslinking, cells were lysed with lysozyme (50 mg ml⁻¹), followed by sonication (15%, 15 s) and DNase treatment (1 mg ml⁻¹; ref. 39). Finally, samples were supplemented with \times 5 reducing SDS Laemmli buffer, separated on a 12% SDS–polyacrylamide gel electrophoresis gel, and Hsp33 was visualized by western blot using polyclonal antibodies against Hsp33 in a 1:5,000 dilution^{1,40}. Select Hsp33–Bpa variants were purified using the purification protocol described below.

***In vivo* incorporation of 4-trifluoromethoxy-phenylalanine.** pET21b plasmids containing a His-tagged version of the respective Hsp33^{M172S}-amber variants were

co-transformed with pDule2-pCNF (plasmid encoding an orthogonal aminoacyl-tRNA synthetase/tRNA pair)²⁷ into BL21 *hsIO*⁻. The transformants were grown overnight on LB plates, supplemented with 200 mg ml⁻¹ ampicillin and 100 mg ml⁻¹ spectinomycin. Freshly transformed cells were scraped off the plates and used to inoculate 1 l of protein expression medium (12 g l⁻¹ tryptone, 24 g l⁻¹ yeast extract, 4% glycerol (v/v), 2.1 g l⁻¹ potassium phosphate (monobasic) and 12.5 g l⁻¹ potassium phosphate dibasic), supplemented with 1 mM ZnCl₂, 1 mM tFPA (JRD Fluorochemicals), 200 mg ml⁻¹ ampicillin and 100 mg ml⁻¹ spectinomycin to an OD₆₀₀ of 0.4. After 1 h of incubation at 37 °C with shaking (200 r.p.m.), cells were shifted to 20 °C and proteins were induced with 0.2 mM IPTG and 0.2% (w/v) L-arabinose. Cultures were harvested the next day by centrifugation (4,000g, 20 min, 4 °C). Cells were resuspended in 40 mM potassium phosphate (KPi), 200 mM KCl, 10 mM imidazole and pH 7.5 supplemented with one tablet of protease inhibitor (Roche) and 1 mM phenylmethylsulphonyl fluoride (Sigma-Aldrich). Cell lysis was performed using a French press cell (3 × 1,400 p.s.i.). The cleared lysate was applied onto a nickel-NTA column (GE Healthcare) equilibrated in 40 mM KPi, 200 mM KCl, 10 mM imidazole, pH 7.5 and Hsp33 variants were eluted using a gradient from 10 to 250 mM imidazole in the same buffer. For further purification of the His-tagged Hsp33 variants, anion-exchange chromatography (Q-sepharose HP, GE Healthcare) was used. The column was equilibrated with 40 mM KPi, 200 mM KCl and Hsp33 was eluted using a gradient from 200 to 700 mM KCl in 40 mM KPi, pH 7.0. Purified protein fractions were pooled, dialyzed against storage buffer (40 mM KPi, pH 7.5) at 4 °C, concentrated using Amicon Ultra-15 centrifugal tubes (EMD Millipore) and stored at -80 °C.

Hsp33 purification. Unlabelled wild-type Hsp33 was expressed and purified⁴¹. In brief, Hsp33-expressing cells (BL21, pET11a, *hsIO*) were grown at 37 °C in the presence of 1 mM ZnCl₂ and harvested 6 h after induction with 1 mM IPTG. Cells were resuspended in buffer A (40 mM HEPES-KOH, 0.2 M KCl (pH 7.5), one tablet Roche complete protease inhibitor mix, 2 mM phenylmethylsulphonyl fluoride) and lysed (French Press, two cycles, 14,000 p.s.i.). The cleared supernatant was applied onto a Q-Sepharose column (Pharmacia) and eluted with a KCl gradient between 450 and 600 mM KCl in buffer A. The Hsp33-containing fractions were dialysed against 10 mM potassium phosphate (pH 6.8) buffer and applied onto a hydroxylapatite column. Hsp33 eluted between 10 and 70 mM potassium phosphate. The Hsp33-containing fractions were then loaded onto a Superdex 75 (Pharmacia) equilibrated in buffer A. Highly purified Hsp33 fractions from the Superdex column were dialyzed against storage buffer (40 mM potassium phosphate buffer (pH 7.5)). ¹⁵N-labelled wild-type Hsp33 and Hsp33-Y12E were expressed in M9 minimal medium supplemented with 0.1% (w/v) ¹⁵N-labelled ammonium chloride and purified as before.

Preparation of NPY^{D4C} IAM-TEMPO. NPY^{D4C}-peptide (GenScript) was dissolved in 40 mM HEPES, pH 7.5 and its concentration was determined at 280 nm ($\epsilon = 5,960 \text{ M}^{-1} \text{ cm}^{-1}$). A 10-fold molar excess of tris(2-carboxyethyl) phosphine (TCEP) was added and incubated for 30 min at 30 °C with shaking (400 r.p.m.) to fully reduce NPY^{D4C}. Then, 4-(2-Iodoacetamido)-2,2,6,6-tetramethyl-1-piperidinyloxy (4-(2-(iodoacetamido)-TEMPO or IAM-TEMPO, Sigma-Aldrich) was added in a stepwise manner (1/5 of the total volume added every minute) to yield a final molar ratio of peptide to label of 1:10 (ref. 42). This solution was incubated in the dark for 2 h at 30 °C with shaking (400 r.p.m.), followed by a dialysis against 40 mM KPi, pH 7.5 using a Spectra/Por Micro Float-A-Lyzer device (500 Da–1 kDa) (Spectrum Laboratories) to remove residual, non-reacted IAM-TEMPO. The concentration of NPY^{D4C} IAM-TEMPO was determined before use in NMR experiments.

Computational methods. Full-length models of Hsp33 for docking procedures and QM calculations were produced using I-TASSER (ref. 43) and ModRefiner (ref. 44). Docking models were generated by removing residues 182–219 from the Hsp33 model (PDB 1HW7), followed by Z-DOCK automated protein–protein docking⁴⁵ between the Hsp33 variant and the conformational ensemble of NPY (ref. 46) with no added constraints. tFPA-containing models for QM calculations were generated by mutating in tFPA into residue 187 in the I-TASSER Hsp33 model in UCSF Chimera 1.10 (ref. 47), followed by structure minimization using the default parameters. Residues were then trimmed to a minimal model surrounding tFPA, and hydrogens added to neutralize the system in Gaussian 5 (Gaussian). The tFPA side chain orientation was then optimized while holding the remainder of the structure constant using $wb97 \times d/6-31 + G(d,p)$ with implicit water solvation, followed by chemical shift calculation using gauge-independent atomic orbitals with the same calculation parameters in Gaussian 09 (ref. 48). The structure of trifluoroacetic acid was optimized by the same methods for chemical shift referencing, as were the calculations of chemical shifts and conformational energies for tFPA alone.

Hsp33 chaperone activity and competition assay. Preparation of reduced, inactive Hsp33 (Hsp33_{red}) or HOCl-oxidized active Hsp33 (Hsp33_{ox}), and chaperone-activity measurements using either chemically or thermally denatured

citrate synthase (CS) as client protein¹⁶. To test the effects of Hsp33 on chemically denatured clients, CS from porcine heart (Sigma-Aldrich) was denatured to a final concentration of 12 μM in 6.0 M guanidinium-hydrochloride (GdmCl), 40 mM HEPES (pH 7.5) overnight at room temperature. To initiate aggregation of CS, the unfolded enzyme was diluted 1:160 into 1,600 μl 40 mM HEPES (pH 7.5) at either 20 or 30 °C in the absence or presence of Hsp33. To analyse Hsp33's effects on thermally unfolding CS, 0.15 μM CS were incubated in 1,600 μl 40 mM HEPES (pH 7.5) at the indicated temperatures in the absence or presence of Hsp33. For both assays, light scattering was monitored using a Hitachi F4500 fluorimeter equipped with a thermostated cell holder and stirrer. Excitation and emission wavelengths were set to 360 nm, and the excitation and emission slit widths were set to 2.5 nm. For competition studies between CS and NPY/NPY D4C IAM-TEMPO, a 10-fold molar excess of peptide to CS was used and the assay performed as described.

¹⁹F NMR experiments. Between 300–700 μM of each Hsp33^{M172S}-tFPA variant was incubated in 40 mM KPi, 5 mM DTT, supplemented with 2 mM trifluoroacetic acid for intensity and chemical shift referencing, and 10% (v/v) D₂O. All ¹⁹F NMR experiments were performed on a 500 MHz Varian VNMRs spectrometer (Agilent Technologies), equipped with a 5 mm PFG OneNMR probe. ¹⁹F spectra were acquired using a single 90° pulse with carbon decoupling. ¹⁹F temperature-dependent spectra were acquired between 10 and 45 °C in steps of 5 °C. Peptide titration spectra of Hsp33 with NPY/NPY^{D4C}-IAM-TEMPO were performed at 30 or 35 °C. Data were processed with iNMR.

In vitro crosslinking procedure. Reaction mixtures containing 20 μM Hsp33_{red} or Hsp33_{ox} and 40 μM NPY or SNPY were crosslinked with either cyanur-biotin-dimercapto-propionyl-succinimide (CBDPS)-H8/D8 (Creative Molecules)³¹, ABAS-¹²C₆/¹³C₆ (Creative Molecules)³² or 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Sigma-Aldrich). To facilitate the identification of Hsp33-NPY zero-length crosslinked peptides (that is, EDC), 20 μM Hsp33 was prepared as an equimolar mixture of ¹⁴N- and ¹⁵N-metabolically labelled protein and incubated with 40 μM NPY.

For crosslinking with CBDPS, mixtures of wild-type Hsp33 and NPY were crosslinked by addition of 1 mM of an equimolar mixture of light and heavy isotopic forms of cyanur-biotin-dimercapto-propionyl-succinimide (CBDPS)-H8/D8 (Creative Molecules)³¹ dissolved in dimethyl sulfoxide and incubation at room temperature for 15 min. The crosslinking reaction was quenched by addition of 10 mM ammonium bicarbonate (ABC) (Sigma-Aldrich).

For crosslinking with ABAS, samples were incubated with 1 mM of an equimolar mixture of light and heavy isotopic forms of ABAS-¹²C₆/¹³C₆ (Creative Molecules)³² dissolved in dimethyl sulfoxide. To crosslink the proteins, the samples were exposed to UV irradiation from the top of an open 0.2 ml reaction tube with a 25 W 254 nm UV lamp (UVGL-58; 115 V, 60 Hz and 0.16A) (2 cm distance) for 10 min at room temperature. The crosslinking reaction was then quenched by addition of 10 mM ABC.

For crosslinking with EDC, samples were crosslinked by addition of 60 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (Sigma-Aldrich) at room temperature for 15 min. The crosslinking reaction was quenched by addition of 10 mM ABC.

Proteolytic digest of in vitro crosslinking products. Crosslinked proteins were digested with either trypsin (Promega) for 18 h at 37 °C or proteinase K (Worthington Biochemical) for 120 min at 37 °C. Both proteases were used at a 1:20 protease to protein ratio. Trypsin and proteinase K digestions were inhibited by the addition of 10 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (Sigma-Aldrich).

Peptide enrichment and sample desalting. CBDPS crosslinked peptides were enriched on monomeric avidin beads (Pierce Biotechnology), eluted from the beads with a solution of 0.1% (v/v) trifluoroacetic acid (Pierce Biotechnology), 50% (v/v) acetonitrile (ACN) (Caledon Laboratories), lyophilized until dry, resuspended and reduced with 10 mM tris-(2-carboxyethyl)-phosphine (TCEP) (Pierce Biotechnology) for 10 min at room temperature, and finally acidified to pH 2 with formic acid (FA) (Thermo Fisher Scientific). Protease digested ABAS and EDC crosslinked samples were desalted using ZipTip C18 (Merck Millipore) pipette tips, lyophilized until dry, resuspended and reduced with 10 mM TCEP for 10 min at room temperature, and subsequently acidified to pH 2 with FA.

LC/MS and data analysis of in vitro crosslinked peptides. Mass spectrometric (MS) analysis was carried out with an Easy-nLC II nano-HPLC system (LC) (Thermo Fisher Scientific) coupled to the ESI-source (Nanospray Flex Ion Source; Thermo Fisher Scientific) of an Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific), as described earlier⁴⁹. Samples were injected onto a 100 μm ID, 360 μm OD 1-cm trap column packed with Magic C18AQ (Bruker-Michrom), 100 Å, 5 μm pore size (prepared in-house) and desalted by washing for 15 min with 0.1% (v/v) FA. Peptides were separated with a 60 minute gradient (0–60 min: 4–40% B; 60–62 min: 40–80% B; 62–70 min: 80% B; with solvent B: 90% ACN, 10% water; and 0.1% FA) on a 75 μm ID, 360 μm OD 15 cm analytical column packed

(in-house) with Magic C18AQ, 100 Å, 5 µm pore size with IntegraFrit (New Objective) and equilibrated with 95% solvent A (2% (v/v) ACN, 98% water and 0.1% (v/v) FA). Mass spectrometry data were acquired with Xcalibur (version 2.1.0.1140) in data-dependent MS/MS mode. Dynamic exclusion was set to 60 s with a repeat count of 2 and a repeat duration of 15 s. Mass spectrometry scans (m/z 400–2,000 range) and MS/MS scans were acquired at 60,000 and 30,000 resolution, respectively. MS/MS fragmentation was performed by collision-induced dissociation at normalized collision energy of 35%. Fourier transform mass spectrometry full automatic gain control target was 1,000,000 and Fourier transform mass spectrometry MSn automatic gain control target was 100,000. For CBDPS-H8/D8 and ABAS-¹²C6/¹³C6 crosslinked samples, mass differences between light and heavy isotopic forms of 8.05 and 6.02 Da were used in the Mass Tags setting, respectively⁴⁹. For PICUP and EDC experiments involving ¹⁵N metabolically labelled Hsp33, a top 6 acquisition method was used. Proteome Discoverer (version 1.4.0.288) was used to generate.MGF files from.RAW files. Data analysis was performed using DXMSMS Match⁵⁰ of ICC-CLASS⁵¹ and ¹⁴N¹⁵N DXMSMS Match⁵².

¹H-¹⁵N HSQC protein NMR measurement. An amount of 300–500 µM Hsp33^{Y12E} was incubated in 40 mM KPi, 5 mM DTT, supplemented with 10% (v/v) D₂O. All ¹H-¹⁵N-NMR experiments were performed on an 800 MHz Varian VNMRs spectrometer (Agilent Technologies), equipped with a 5 mm HCN inverse probe with z -axis gradient. The NMR operating software was Vnmrj 3.2. Data were processed with iNMR.

References

- Winter, J., Ilbert, M., Graf, P. C., Ozcelik, D. & Jakob, U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell* **135**, 691–701 (2008).
- Foit, L., George, J. S., Zhang, B. W., Brooks, 3rd C. L. & Bardwell, J. C. Chaperone activation by unfolding. *Proc. Natl Acad. Sci. USA* **110**, E1254–E1262 (2013).
- Hsu, W. L. *et al.* Intrinsic protein disorder and protein-protein interactions. *Pac. Symp. Biocomput.* 116–127 (2012).
- Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272**, 5129–5148 (2005).
- Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl Acad. Sci. USA* **110**, 13392–13397 (2013).
- Muller-Spath, S. *et al.* From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **107**, 14609–14614 (2010).
- Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U. & Moarefi, I. Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. *Cell* **103**, 621–632 (2000).
- Goldberg, M. S. *et al.* Native-like structure of a protein-folding intermediate bound to the chaperonin GroEL. *Proc. Natl Acad. Sci. USA* **94**, 1080–1085 (1997).
- Wang, Z., Feng, H., Landry, S. J., Maxwell, J. & Gierasch, L. M. Basis of substrate binding by the chaperonin GroEL. *Biochemistry* **38**, 12537–12546 (1999).
- Burmann, B. M. & Hiller, S. Chaperones and chaperone-substrate complexes: dynamic playgrounds for NMR spectroscopists. *Prog. Nucl. Magn. Reson. Spectrosc.* **86–87C**, 41–64 (2015).
- Bardwell, J. C. & Jakob, U. Conditional disorder in chaperone action. *Trends Biochem. Sci.* **37**, 517–525 (2012).
- Franzmann, T. M., Menhorn, P., Walter, S. & Buchner, J. Activation of the chaperone Hsp26 is controlled by the rearrangement of its thermosensor domain. *Mol. Cell* **29**, 207–216 (2008).
- Varadi, M. *et al.* pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. *Nucleic Acids Res.* **42**, D326–D335 (2014).
- Sickmeier, M. *et al.* DisProt: the database of disordered proteins. *Nucleic Acids Res.* **35**, D786–D793 (2007).
- Ilbert, M. *et al.* The redox-switch domain of Hsp33 functions as dual stress sensor. *Nat. Struct. Mol. Biol.* **14**, 556–563 (2007).
- Reichmann, D. *et al.* Order out of disorder: working cycle of an intrinsically unfolded chaperone. *Cell* **148**, 947–957 (2012).
- Creemers, C. M., Reichmann, D., Hausmann, J., Ilbert, M. & Jakob, U. Unfolding of metastable linker region is at the core of Hsp33 activation as a redox-regulated chaperone. *J. Biol. Chem.* **285**, 11243–11251 (2010).
- Chin, J. W. & Schultz, P. G. *In vivo* photocrosslinking with unnatural amino acid mutagenesis. *ChemBiochem.* **3**, 1135–1137 (2002).
- Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* **395**, 361–374 (2010).
- Vijayalakshmi, J., Mukherjee, M. K., Graumann, J., Jakob, U. & Saper, M. A. The 2.2 Å crystal structure of Hsp33: a heat shock protein with redox-regulated chaperone activity. *Structure* **9**, 367–375 (2001).
- Jackson, J. C., Hammill, J. T. & Mehl, R. A. Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **129**, 1160–1166 (2007).
- Miyake-Stoner, S. J. *et al.* Generating permissive site-specific unnatural aminoacyl-tRNA synthetases. *Biochemistry* **49**, 1667–1677 (2010).
- Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C. & Mehl, R. A. Preparation of site-specifically labelled fluorinated proteins for 19F-NMR structural characterization. *Nat. Protoc.* **2**, 2601–2607 (2007).
- Marsh, E. N. & Suzuki, Y. Using (19)F NMR to probe biological interactions of proteins and peptides. *ACS Chem. Biol.* **9**, 1242–1250 (2014).
- Gerig, J. T. Fluorine NMR of Proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **26**, 293–370 (1994).
- Cellitti, S. E. *et al.* *In vivo* incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **130**, 9268–9281 (2008).
- Peeler, J. C. & Mehl, R. A. Site-specific incorporation of unnatural amino acids as probes for protein conformational changes. *Methods Mol. Biol.* **794**, 125–134 (2012).
- Graf, P. C. *et al.* Activation of the redox-regulated chaperone Hsp33 by domain unfolding. *J. Biol. Chem.* **279**, 20529–20538 (2004).
- Clore, G. M. & Iwahara, J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109**, 4108–4139 (2009).
- Gillespie, J. R. & Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **268**, 158–169 (1997).
- Petrotchenko, E. V., Serpa, J. J. & Borchers, C. H. An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell Proteomics* **10**, M110 001420 (2011).
- Brodie, N. I., Makepeace, K. A., Petrotchenko, E. V. & Borchers, C. H. Isotopically-coded short-range hetero-bifunctional photo-reactive crosslinkers for studying protein structure. *J. Proteomics* **118**, 12–20 (2015).
- Dahl, J. U. *et al.* HdeB functions as an acid-protective chaperone in bacteria. *J. Biol. Chem.* **290**, 65–75 (2015).
- Cheng, G., Basha, E., Wysocki, V. H. & Vierling, E. Insights into small heat shock protein and substrate structure during chaperone action derived from hydrogen/deuterium exchange and mass spectrometry. *J. Biol. Chem.* **283**, 26634–26642 (2008).
- Graf, C., Lee, C. T., Eva Meier-Andrejszki, L., Nguyen, M. T. & Mayer, M. P. Differences in conformational dynamics within the Hsp90 chaperone family reveal mechanistic insights. *Front. Mol. Biosci.* **1**, 4 (2014).
- Joachimski, L. A., Walzthoeni, T., Liu, C. W., Aebersold, R. & Frydman, J. The structural basis of substrate recognition by the eukaryotic chaperonin TRiC/CCT. *Cell* **159**, 1042–1055 (2014).
- Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. & Hartl, F. U. Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* **82**, 323–355 (2013).
- Lennon, C. W. *et al.* Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes Dev.* **26**, 2634–2646 (2012).
- Fu, X. *et al.* Small heat shock protein IbpB acts as a robust chaperone in living cells by hierarchically activating its multi-type substrate-binding residues. *J. Biol. Chem.* **288**, 11897–11906 (2013).
- Ruel, N. *et al.* Hsp33 controls elongation factor-Tu stability and allows *Escherichia coli* growth in the absence of the major DnaK and trigger factor chaperones. *J. Biol. Chem.* **287**, 44435–44446 (2012).
- Graumann, J. *et al.* Activation of the redox-regulated molecular chaperone Hsp33—a two-step mechanism. *Structure* **9**, 377–387 (2001).
- Holyoak, T. & Nowak, T. Structural investigation of the binding of nucleotide to phosphoenolpyruvate carboxykinase by NMR. *Biochemistry* **40**, 11037–11047 (2001).
- Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
- Xu, D. & Zhang, Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys. J.* **101**, 2525–2534 (2011).
- Pierce, B. G. *et al.* ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **30**, 1771–1773 (2014).
- Monks, S. A., Karagianis, G., Howlett, G. J. & Norton, R. S. Solution structure of human neurotrophin Y. *J. Biomol. NMR* **8**, 379–390 (1996).
- Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Frisch, M. J. *et al.* *Gaussian 09. Revision B.01 edn* (Wallingford, CT, 2009).

49. Petrotchenko, E. V., Makepeace, K. A., Serpa, J. J. & Borchers, C. H. Analysis of protein structure by cross-linking combined with mass spectrometry. *Methods Mol. Biol.* **1156**, 447–463 (2014).
50. Petrotchenko, E. V., Makepeace, K. A. & Borchers, C. H. DXMSMS match program for automated analysis of LC-MS/MS data obtained using isotopically coded CID-cleavable cross-linking reagents. *Curr. Protoc. Bioinformatics* **48**, 8 18 1–8 18 19 (2014).
51. Petrotchenko, E. V. & Borchers, C. H. ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics* **11**, 64 (2010).
52. Petrotchenko, E. V., Serpa, J. J., Makepeace, K. A., Brodie, N. I. & Borchers, C. H. (14)N(15)N DXMSMS Match program for the automated analysis of LC/ESI-MS/MS crosslinking data from experiments using (15)N metabolically labeled proteins. *J. Proteomics* **109**, 104–110 (2014).
53. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: cross-link network maps with residue resolution. *Mol. Cell Proteomics* **14**, 1137–1147 (2015).

Acknowledgements

We thank R. Mehl and the Unnatural Protein Facility for helpful advice and for providing the p-CNF vector, P. Schultz for providing the pEVOL vector, C. Lennon for helpful discussion on BPA crosslinking and C. Cremers for reading the manuscript, general technical advice and fruitful discussions. We thank the Genome Canada and Genome British Columbia Science and Technology Innovation Centre funding and The Natural Sciences and Engineering Research Council of Canada grant (to C.H.B.). This research was funded by the National Institutes of Health Grant GM065318 (to U.J.). J.C.A.B is a Howard Hughes Investigator.

Author contributions

B.G. performed experiments, performed data analysis and wrote the manuscript; S.H. conceived and performed experiments, performed data analysis and wrote the manuscript. K.A.T.M. and E.V.P. performed experiments and conducted data analysis under the supervision of C.H.B., J.C.A.B. conducted data analysis and wrote the paper. U.J. conceived experiments, conducted data analysis and analysed the data.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Groitl, B. *et al.* Protein unfolding as a switch from self-recognition to high-affinity client binding. *Nat. Commun.* **7**:10357 doi: 10.1038/ncomms10357 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>