Supporting Information for

Selective promiscuity in binding of the *E. coli* Hsp70 chaperone to an unfolded protein

Eugenia M. Clerico[a,#], Alexandra K. Pozhidaeva[a,#], Rachel Jansen[a,&,#], Can Özden[a,c], Joseph M. Tilitsky[b,%] and Lila M. Gierasch[a,b,*]

Departments of [a]Biochemistry & Molecular Biology and [b]Chemistry and [c]Graduate Program in Molecular and Cellular Biology
University of Massachusetts Amherst
Amherst, MA 01003

*To whom correspondence may be addressed. Email: gierasch@biochem.umass.edu

[#] These authors contributed equally to this paper.

[&]Current address: Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA 94709.

[%]Current address: Anokion, 50 Hampshire St, Cambridge, MA 02139.

This PDF file includes:

SI Text

Design of proPhoA^{S4}

We sought to use fully reduced proPhoA as a full-length unfolded protein substrate for DnaK and thus needed to ensure that it lacked structure and remained in this state over the time necessary for biophysical studies. The original proPhoA preparation containing the native four cysteine residues when kept in reducing conditions has been successfully used as a model for an unfolded substrate by Kalodimos and coworkers (1, 2). However, proPhoA purified from *E. coli* in the presence of urea and reductant and then exchanged into aqueous buffer consists of a mixture of unfolded and partially folded species and must be fractionated by size exclusion chromatography to isolate the unfolded protein. The resulting protein retains a tendency to take up some structure over time, as indicated by dispersed NMR signals (Fig. S1B). In order to have a homogeneous sample of unfolded proPhoA we created a cysteine-less mutant of proPhoA (proPhoA^{S4}) by site directed mutagenesis of the proPhoA gene. This variant of proPhoA indeed showed no indication of residual structure (Fig. S1B). Clones were fully sequenced, and proPhoA^{S4} was overexpressed and purified as described for wild type proPhoA (1) with one exception: Once eluted from the Ni-NTA column, proPhoA^{S4}-containing fractions were pooled and concentrated using Amicon® Ultra Centrifugal Filter Units. The resulting preparation was further purified by size exclusion using HiLoad 16/600 Superdex 200 pg (Cytiva) in buffer containing 50 mM Tris pH 8.0, 50 mM NaCl and 6 M urea. The fractions containing pure proPhoA^{S4} were concentrated and frozen at -20 °C. The urea was removed immediately before each experiment using micro BioSpin 6 columns (BioRad). Protein purity was determined by SDS-PAGE and the concentration by UV absorption. ProPhoA^{S4} has to be kept cold in aqueous buffer to prevent aggregation, as it readily aggregates at concentrations above 40 μM.

Validation of the experimental approach

We validated our combined experimental approach using two well-known Hsp70 binding peptides: NR (NRLLLTG) and NR^{Ile} (NRLILTG). In crystals, NR binds the DnaK SBD in

the N- to C- orientation with L4 in the central pocket, and NR^{Ile} binds in a C- to N-orientation with L3 in the 0^{th} site (Fig. S6A and B) (3, 4). Binding of NR and NR^{Ile} bearing a N-terminal cysteine (CysNR or $CysNR^{Ile}$) resulted in a single resonance for SBD residues Ile401 and Ile438 in ^{1}H-^{13}C HMQC spectra, consistent with the binding modes observed in the crystal structures (Fig. S6F). Intriguingly, despite the overall sequence similarity between the peptides and the fact that leucine occupies the SBD central pocket in both cases, the Ile401 and Ile438 chemical shifts were very different for the two complexes. This observation suggests that the positions of the Ile401 and Ile438 resonances might report on the binding orientation. To confirm that the opposing binding orientations for these two peptides are preserved in solution, we deployed PRE and cross-linking methods. For these experiments, we used longer versions of the NR peptides, CysNR* and $CysNR^{Ile}$* (Fig. S6C), since the original peptides were too short to cross-link to either SBD^{Cys425} or SBD^{Cys458}. The Ile401 and Ile438 resonances for SBD bound to CysNR* overlapped with the ones seen for SBD/CysNR complex indicating that the addition of the linker did not affect the mode of binding (Fig. S6F). Consistent with the crystal structure and our methyl-TROSY data, CysNR* formed a disulfide bond with the SBD^{Cys425} variant, and its labeling with MTSL resulted in resonance broadening on the side of the SBD expected for the N- to C- binding mode (Fig. S6D and E). For $CysNR^{Ile}$*, the major Ile401 and Ile438 resonances overlapped with the ones seen for $CysNR^{Ile}$ (Fig. S6F), but we also observed an additional set of Ile401 and Ile438 resonances suggesting that the presence of the linker resulted in an additional binding mode. [Interestingly, the major Ile438 resonance was split, most likely because two distinct rotameric conformations are populated (5).] The additional Ile401 and Ile438 resonances were located in the spectral region that would suggest N- to C- binding orientation if our predictions were correct. We speculate that second arginine in the linker allowed for an appearance of a binding mode in which L7 occupies the central pocket and R4 is placed in the -3 position. Indeed, $CysNR^{Ile}$* cross-linked to both SBD^{Cys425} and SBD^{Cys458} indicating that this peptide binds in both orientations in solution and providing evidence that the positions of SBD Ile401 and Ile438 resonances in ^{1}H-^{13}C-HMQC spectra report on the orientation of the bound peptide (Fig. S6D). Overall, the results on these two peptides validated the approaches designed to determine 1) the binding mode and 2)

SBD-bound peptide orientation, which we then used for all proPhoA peptides as described in the main text.

SI Materials and Methods

Protein expression and purification

The gene encoding wild type *E. coli* DnaK (residues 1-638, cloned into the pMSK plasmid) (6) and the gene encoding the DnaK SBD (residues 389-607 cloned into pSCG plasmid carrying an N-terminal His_6-tag followed by a Tobacco Etch Virus (TEV) protease digestion sequence) were used to overexpress the proteins in *E. coli* BL21(*DE3*). DnaK SBD (389-607) variants carrying cysteine residues at positions 425 (SBD^{Cys425}) and 458 (SBD^{Cys458}) were created by site directed mutagenesis of the wild type gene. Cells carrying the appropriate plasmids were typically grown in 2 L of Luria Bertani (LB) media with 100 μg/ml ampicillin at 37 °C until OD_{600} ~0.6. Protein overexpression was induced by addition of 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) to 0.5 mM, and the cultures were incubated for an additional 3-4 h at 30 °C, when the cells were harvested by centrifugation. To prepare uniformly ^{15}N- or $^{15}N,^{13}C$-labeled proteins for NMR experiments, cells carrying the appropriate plasmids were grown in M9 medium in H_2O containing $^{15}NH_4Cl$ and/or ^{13}C-glucose as a source of nitrogen and carbon, respectively. Cultures were grown at 37 °C until OD_{600} ~0.8, and protein expression was induced by addition of IPTG at concentration of 0.5 mM. After induction, cells were allowed to grow for an additional 16 h at 18-20 °C. To prepare proteins with selectively labeled Ile, Leu and Val methyl groups (*i.e.*, [U-$^{2}H,^{12}C,^{15}N$]; Ileδ1-[$^{13}CH_3$]; Leu,Val-[$^{13}CH_3,^{12}CD_3$]), cells were grown in D_2O-based M9 minimal medium containing $^{15}NH_4Cl$ and ^{2}H-^{12}C-glucose. The same protocol was used for bacterial growth and additionally, sodium salts of α-ketobutyric (Methyl-^{13}C, 3,3-D2) and α-ketoisovaleric (3-Methyl-^{13}C, 3,4,4,4-D4) acids were added (70 mg/L and 120 mg/L, respectively) 1 h before induction (7). All isotopically labeled compounds were obtained from Cambridge Isotope Laboratories, Inc.

All proteins were purified at 4 °C unless otherwise noted. DnaK was purified by anion-exchange followed by ATP-affinity chromatography as previously described (6). Cells carrying the overexpressed SBD constructs were resuspended in His-binding buffer (50

mM Tris pH 8.0, 500 mM sodium chloride), supplemented with protease inhibitors leupeptin, phenylmethylsulfonyl fluoride (PMSF) and pepstatin, and lysed by passage through a cell disruptor (Microfluidics). The lysate was cleared by centrifugation at 48,300 $\times$ *g* at 4 °C for 45 min. The soluble fraction was loaded in a 15 ml Ni-NTA column equilibrated in His-binding buffer, and the column was washed with 10 column volumes of His-binding buffer supplemented with 6 M urea to unfold the SBD and remove any bound substrates. The Ni-NTA-bound DnaK SBD was allowed to refold by washing the column with His-binding buffer. The protein was eluted by applying a linear gradient between 0 and 200 mM imidazole in His-binding buffer. The fractions containing the SBD were treated with His-tagged TEV protease for 16 h at 4 °C to remove the His_6-tag sequence. The digestion mixture was exchanged into His-binding buffer without imidazole and then passed through the Ni-NTA column again to remove the undigested protein, digested His_6-tag, and His-tagged TEV. TEV-digested SBD was buffer-exchanged to 20 mM Hepes pH 7.6, 10 mM $MgCl_2$, 100 mM KCl (HMK buffer). The purity of the protein was determined by SDS-PAGE and the concentration by quantitative amino acid analysis. Proteins were flash frozen and stored at -80 °C. For NMR experiments, the SBD protein samples were buffer-exchanged into water, aliquoted, flash frozen in liquid N_2 and lyophilized.

The plasmid encoding the precursor of the *E. coli* alkaline phosphatase (proPhoA) labeled at the C-terminus with a His_6-tag was generously provided by C. Kalodimos. The wild type, reduced proPhoA was purified as described before (1). Briefly, cells containing the plasmid were grown in 1 L LB-ampicillin cultures at 37 °C until OD_{600} ~0.4. Then, IPTG was added to 0.2 mM and induction allowed to proceed for 16 h at 16-20 °C. Cells were harvested and protein purified by Ni-NTA affinity chromatography in the presence of 6 M urea in His-binding buffer supplemented with 1 mM DTT. proPhoA was eluted from the Ni NTA column by applying a linear gradient between 10 and 200 mM imidazole in His-binding buffer containing urea. The protein was buffer exchanged by ultrafiltration into 20 mM Tris, pH 8.0, 1 mM DTT and further purified by size exclusion (as described above) in aqueous buffer. Protein purity was assayed by SDS-PAGE and concentration determined by UV absorption.

Peptides

All peptides used in this work (proPhoA peptides, FITC-p5 (ALLLSAPRR), drosocin (CSHPRPIRV) and $E_{sh}V$ (RGSQVRIASR)) were purchased from Biomatik, Ruifu Chemical or Genscript at >99 % purity. For PRE and cross-linking experiments, peptides containing N-terminal cysteine including CysNR (CNRLLLTG), CysNR* (CSRSRNRLLLTG), $CysNR^{Ile}$ (CNRLILTG), $CysNR^{Ile*}$ (CSRSRNRLILTG), CysA (CRKQSTIALALLPLLFTPRR), CysC (CEANQQKPLLGLFADG) and CysD (CAKKEGNTLVIVTADH) were obtained from Biomatik and GenScript. The peptides were resuspended in H_2O, spun at 20,000 x g for 15 min, and the concentration of the stock solutions was verified by quantitative amino acid analysis at the Protein Chemistry Laboratory (Texas A&M University) or UC Davis Proteomics Core facility (UCDavis). The pH of the peptide solution was adjusted to 7.4 using potassium hydroxide if necessary. The purity of the peptides was confirmed by HPLC: peptides in water were loaded in a reverse phase C18 analytical column (VYDAC) and eluted with a linear gradient of acetonitrile 0.1% trifluoroacetic acid. The mass of each peptide was confirmed by MALDI analysis at the University of Massachusetts Mass Spectrometry Center.

NMR spectroscopy

NMR experiments were carried out at 25 °C (unless stated otherwise) on a 600 MHz Bruker Avance III spectrometer equipped with a CryoProbe. Data were processed with NMRPipe (8) and analyzed in ccpNMR Analysis (9). For Methyl-TROSY experiments of the complexes between DnaK or DnaK SBD with proPhoA peptides, the peptides were resuspended in H_2O, adjusted to pH 7.4 and then lyophilized. Each lyophilized peptide was mixed with the protein in NMR buffer (10 mM potassium phosphate pD 7.4, 5 mM DTT, 0.02% sodium azide, 0.1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) in D_2O) and incubated for 16 h at 4 °C. The concentration of ILV-$^{13}CH_3$ DnaK SBD and peptides in NMR experiments was 40 μM each, unless otherwise indicated. For complexes between DnaK SBD and $proPhoA^{S4}$, $proPhoA^{S4}$ was buffer exchanged into NMR buffer at 4 °C immediately before mixing with the SBD and incubated for 16 h at 4 ºC. For titration of $proPhoA^{S4}$ with ILV-$^{13}CH_3$ DnaK SBD, a solution of 40 μM $proPhoA^{S4}$ in NMR buffer was used to resuspend the lyophilized SBD and then

incubated at 4 °C for 10 h before spectra taken. Subsequently, the same sample was used to resuspend increasing amounts of lyophilized SBD. All ^{1}H-^{13}C correlation spectra were collected using a standard Bruker heteronuclear multiple quantum coherence (HMQC) pulse sequence (hmqcphpr) with a relaxation delay of 1.5 s. Assignments of SBD Ile401 and Ile438 resonances were obtained previously (10). For ^{1}H-^{15}N HSQC-TROSY (10), samples of ^{15}N-labeled proPhoA and proPhoA^{S4} proteins at concentrations of 40 μM were used in NMR buffer containing 90% H_2O/10% D_2O. Spectra were collected at 15 °C using a standard Bruker pulse sequence (trosyetfpf3gpsi).

For paramagnetic relaxation experiments (PREs), lyophilized Cys-containing proPhoA peptides and CysNR* were resuspended in water without adjusting the pH to prevent disulfide bond formation. Immediately prior to incubating with 10X molar equivalents of MTSL (S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate, Sant Cruz, Cayman), pH was adjusted to 8 by addition of Tris to a final concentration of 20 mM. The labeling reaction proceeded for two hours at 22 ºC followed by an additional 16 h incubation at 4 °C in presence of ^{15}N-labeled SBD at 2:1 peptide/protein ratio (note that wild-type SBD of DnaK does not contain cysteine residues in its sequence). The next day the SBD/MTSL-peptide complexes were buffer exchanged to remove the excess MTSL and the unbound peptide, and transferred to NMR buffer containing 10 mM potassium phosphate pH 7.4 and 10% D_2O. The presence of the labeled peptide was confirmed by mass spectrometry. Diamagnetic samples were prepared by adding 10X molar excess of fresh sodium ascorbate. Initial chemical shift assignments of ^{1}H-^{15}N HSQC-TROSY spectra were transferred from the ones available for the shorter SBD construct (386-552^{YE}) bound to NR (12). To confirm the assignments and assign the SBD backbone resonances corresponding to the α-helical lid (residues 507-607), SBD bound to peptide A was used to collect 3D NMR experiments including TROSY-based HNCA, HNCO, and HN(CA)CO. HNCA experiments were collected for SBD bound to all other proPhoA peptides to facilitate the transfer of NMR chemical shift assignments between the complexes. PRE data were obtained by recording ^{1}H-^{15}N HSQC-TROSY spectra of the SBD in the presence of peptides in paramagnetic and diamagnetic states, and intensity ratios of well-resolved peaks were extracted with the following calculation of I_{para}/I_{dia} ratio for each SBD/peptide complex (13). Peak intensity

loss in the paramagnetic state was different for each SBD/peptide complex depending on the MTSL labeling efficiency for the particular peptide. In each case, we designated the 15-20 SBD residues most affected by MTSL presence as the ones experiencing a significant PRE effect.

Cross-linking experiments

Disulfide cross-linking between SBD and peptides

In order to determine the binding orientation of each peptide bound to DnaK SBD in solution, we developed an approach where a cysteine residue introduced at the N-terminus of a given substrate peptide was evaluated for its ability to form a disulfide bond with a cysteine residue introduced into the SBD. We introduced a single cysteine at opposite sides of the SBD at positions 425 or 458 (SBD^{Cys425} and SBD^{Cys458}, respectively) (Fig. 5A) by site directed mutagenesis and at the N-terminus of peptides, CysNR (CNRLLLTG), CysNR* (CSRSRNRLLLTG), $CysNR^{Ile}$ (CNRLILTG), $CysNR^{Ile}$* (CSRSRNRLILTG), CysA (CRKQSTIALALLPLLFTPRR), CysC (CEANQQKPLLGLFADG), CysD (CAKKEGNTLVIVTADH) and CysE (CEHTGSQLRIAAYGP). The introduction of a cysteine to the SBD did not perturb its binding to proPhoA peptides and p5 (Fig. S5C and D), and the introduction of N-terminal cysteine to the peptides did not perturb their mode of binding to the SBD (Fig. S5B).

20 μM SBD SBD^{Cys425} and SBD^{Cys458} were incubated with 40 μM Cys-peptide for 2 h at 22 °C in HMK buffer with 1 mM DTT or 50 mM Tris-HCl pH 8.0, 1 mM DTT. Excess peptide and DTT were removed by loading 50 μl of the sample in a desalting spin column (Zeba Spin, Thermo Scientific). The samples were kept at 22 °C for 15 min to allow the cysteine in the bound peptide to form a disulfide bond with the cysteine in the SBD (with Cys425 if the peptide is bound on the N- to C- orientation, and with Cys458 if it is bound in the C- to N-orientation). The samples were then supplemented with iodoacetamide (Thermo Fisher Scientific) at a concentration of 100 mM to block the unreacted thiol groups, and kept at 22 °C for 15 mins followed by addition of 5X SDS-PAGE loading buffer (without reducing agent). The results were visualized by SDS-PAGE using 12% acrylamide gels (GenScript).

SBD cross-linking to bound proPhoA^{S4}

To evaluate the stoichiometry of the complex between the DnaK SBD and proPhoA^{S4} we used Sulfo-GMBS (N-γ-maleimidobutyryl-oxysulfosuccinimide ester, Thermo Fisher Scientific) to covalently cross-link a single cysteine residue in SBD^{Cys425} with bound proPhoA^{S4}. Sulfo-GMBS is a hetero bifunctional reagent that contains NHS-ester and maleimide reactive groups, allowing the cross-linking of amine-to-sulfhydryl groups. First, we pre-formed complexes between 20 μM proPhoA^{S4} and SBD Cys^{425} at 0, 10 and 120 μM for 16 h at 4 °C in ~50 μl buffer potassium phosphate 20 mM, pH 7.2, 1 mM DTT and Halt Protease inhibitor cocktail (Thermo Fisher Scientific). Next, DTT and residual Tris buffer were removed using a desalting spin column (as described above), and the mix was supplemented with a 20X molar excess over protein of Sulfo-GMBS (from a 10 mM stock), or buffer for noncross-linked control. The reaction was allowed to proceed for 15 mins at 22 °C and then quenched by the addition of 5X SDS-PAGE loading buffer in the presence of reducing agent. The samples were loaded in SDS-PAGE gels (Any kD, BioRad) and stained with Coomassie blue for visualization.

Substrate-binding assay

The apparent affinities of DnaK SBD binding for all proPhoA peptides were measured by fluorescence anisotropy using competition assays at 25 °C in a Biotek Synergy 2 micro plate reader (Biotek), with excitation at 485 nm and emission at 516 nm. Protein-peptide mixtures were incubated in 384 well plates for 16 h at 4 °C, then allowed to reach room temperature, after which the fluorescence anisotropy was measured. The affinity of the SBD for fluorescein isothiocyanate (FITC)-labeled peptide p5 (ALLLSAPRR) was measured first by direct binding as previously described (6, 14); briefly 0-100 μM SBD were added to 35 or 50 nM FITC-p5 in HMK buffer, 1 mM DTT in triplicate. The affinity of the DnaK SBD (389-607) for FITC-p5 is 0.008 ± 0.003 μM. For the experiments where proPhoA^{S4} was used as a competitor for FITC-p5 bound to SBD or DnaK, proPhoA^{S4} was buffer exchanged into HMK buffer with 1 mM DTT and added from 0-35 μM to solutions containing SBD or DnaK (100 nM) and FITC-p5 (50 nM). The measured anisotropy values were converted to fraction bound (f_b) using the equation:

$$fb = (r - r_0)/(r_b - r_0)$$

where *r* is the anisotropy value at any point, r_0 the anisotropy value of FITC-labeled p5 in the absence of SBD, and r_b the anisotropy value of FITC-p5 when fully bound to the SBD. Using Kaleidagraph (Synergy Software) the curve of f_b vs. concentration of SBD was fitted to the equation:

$$fb = \frac{\left\{[P+S+K_D] - \sqrt{[(P+S+K_D)^2 - 4*P*S]}\right\}}{2*S}$$

where f_b is fraction bound, *S* is the concentration of FITC-p5, *P* is the total concentration of added SBD, and K_D is the dissociation constant of the complex (15).

For peptide competition assays, 0-100 μM of the competitor peptide was added to 100 nM SBD and 50 nM FITC-p5 to assay the ability of each peptide to compete off the FITC-p5 from the SBD binding site. To obtain the IC_{50}, each competition curve of f_b vs. concentration of added peptide was fitted to:

$$fb = fb_0 + (fb_f - fb_0) * x/(x + IC_{50})$$

Where f_{b0} and f_{bf} are the initial and final f_b values, respectively. Using the Cheng and Prusoff Equation (16) IC_{50} was converted to apparent K_D for each peptide tested using the equation:

$$K_i = IC_{50}/[1 + \left(\frac{R}{K_D}\right)]$$

where K_i is the dissociation constant of the SBD and the competitor peptide, *R* is the concentration of FITC-p5 and K_D corresponds to the affinity between FITC-p5 and SBD (0.008±0.003 uM), obtained with the direct binding assay described above. A more accurate cubic fit was also applied to analyze the binding of competing ligands to the SBD (17, 18), but in this case did not yield good fits to the competition curves, especially for

peptides with higher affinity such as A and $A_{sh}C$. Because our goal was to compare the relative affinities of the SBD for the proPhoA peptides, we used the simplified Hill equation described in the text to obtain IC_{50}, and then estimated the K_D using the Cheng and Prusoff equation, and we call the obtained values “apparent affinity” or “apparent K_D”.
The affinities of SBD^{Cys425} and SBD^{Cys458} for FITC-p5 were measured by a direct binding assay (Fig. S5D).

Crystallization

Crystals of DnaK SBD (389-607) bound to proPhoA-derived peptides were grown using similar conditions to previous work (3, 4). The DnaK SBD was concentrated to 4 mg/mL in 20 mM Tris-HCl pH 8.0, 5% glycerol and mixed with lyophilized peptides at a ratio of 1:5 or 1:7.5, protein to peptide. Crystals were grown at ~20 °C using the hanging drop diffusion method. The precipitant in both the reservoir and drop contained 1.8-2.8 M ammonium sulfate pH 6.9-7.1, and 0.1 M potassium phosphate, and the reservoir contained 600 µl of precipitant solution. In some cases, seeding with crystals of SBD in complex with peptide E was used to obtain crystals. Exact crystallization conditions for each complex can be found in Table S1.

Structure determination and refinement

A cryo-protectant containing ~1.5-1.8 M ammonium sulfate and 20% glycerol in liquid nitrogen was used for data collection. Diffraction data were collected at the UMass Amherst Biophysical Characterization Facility, at wavelength 1.5418 Å using a Rigaku MicroMax-007 HF x-ray source, coupled to a Rigaku VariMax HF optic system. The data were integrated, merged and scaled using HKL-2000 (19). The first structure was determined via molecular replacement by Phaser (20) using the coordinates from the SBD/NR structure (PDB ID: 1DKZ, (3)) to solve the phase problem. After obtaining the first structure, that model was used for molecular replacement of the other structures. The SBD-bound peptides were built into the structures using Coot, and refinement was performed with Refmac (21). The programs Phaser, Coot and Refmac are part of the CCP4 software suite (22). A variety of calculations were done to analyze the SBD-peptide interaction. φ and ψ, the torsion angles of the peptide, were calculated in COOT. The

average main-chain temperature factor, <B>, of each residue was calculated using Baverage software. The degree of solvent exposure was calculated using Areaimol (23). COOT, Baverage and Areaimol software are part of CCP4 software suite (22). The upper limit (r_{limit}) for van der Waals contacts was calculated using the methods from Zahn *et al.* (4). All structures were represented using PyMol (Schrodinger LLC).

Table S1. Crystallization conditions, diffraction and lattice statistics for SBD Structures in complex with proPhoA Peptides.

Peptide	Crystallization Conditions	Seeding (Y/N)	Growing Time (months)	Crystal Shape	Backbone Direction	Resolution (Å)	R Value	Free R Value	Unit Cell Dimensions	Space Group
A	Precipitant: 2.4 M $(NH_4)_2SO_4$, 0.1 M K_3PO_4 pH 7.0 100 µM SBD, 0.75 mM peptide	N	4	Pyramidal	N- to C-/ C- to N-	2.00	0.232	0.293	a = 47.3 b = 57.6 c = 181.5	I 2 2 2
$B_{sh}N$	Precipitant: 2.4 M $(NH_4)_2SO_4$, 0.1 M K_3PO_4 pH 7.0 100 µM SBD, 0.75 mM peptide	N	3.5	Pyramidal	C- to N-	2.56	0.220	0.321	a = 38.1 b = 95.5 c = 117.4	I 2 2 2
C	Precipitant: 2.8 M $(NH_4)_2SO_4$, 0.1 M K_3PO_4 pH 7.0 100 µM SBD, 0.75 mM peptide	N	6	Pyramidal	N- to C-	2.42	0.256	0.340	a = 37.6 b = 95.4 c = 117.4	I 2 2 2
D_{sh}	Precipitant: 2.8 M $(NH_4)_2SO_4$, 0.1 M K_3PO_4 pH 7.0 150 µM SBD, 0.75 mM peptide	N	6	Pyramidal	C- to N-	3.09	0.189	0.290	a = 38.0 b = 95.5 c = 117.5	I 2 2 2

E	Precipitant: 2.4 M $(NH_4)_2SO_4$, 0.1 M K_3PO_4 pH 7.0 150 µM SBD, 0.75 mM peptide	Y	1.5	Pyramidal	N- to C-	2.40	0.232	0.300	a = 37.3 b = 95.6 c = 117.0	I 2 2 2

Table S2. Hydrogen bonds between peptides containing the DnaK binding sites in proPhoA and the DnaK SBD.

Peptide/Number of hydrogen bonds	SBD pocket	Peptide Residue	H-bond
Peptide A (N- to C-) Total H-bonds: 6	-3	L	-
	-2	L	Ser427 amide to carbonyl of L(-11)
	-1	P	Met404 amide to carbonyl of P(-10)
	0	L	Gln433 side chain NH_2 to carbonyl of L(-9) Ser427 carbonyl to amide of L(-9)
	+1	L	Thr437 amide to carbonyl of L(-8)
	+2	F	-
	+3	T	Thr437 side chain OH to amide of T(-6)
Peptide A (C- to N-) Total H-bonds: 7	-2	L	-
	-1	A	Ser427 amide to carbonyl of A(-15) Ser427 carbonyl to amide of A(-15)
	0	I	Met404 amide to carbonyl of I(-16)
	+1	T	Gln433 side chain NH_2 to carbonyl of T(-17)
	+2	S	Thr437 amide to carbonyl of S(-18) Thr437 side chain OH to amide of S(-18)
	+3	Q	Thr437 carbonyl to side chain of Q(-19)
Peptide $B_{sh}N$ (C- to N-) Total H-bonds: 7	-2	L	-
	-1	I	Ser427 amide to carbonyl of I47 Ser427 carbonyl to amide of I47
	0	I	Met404 amide to carbonyl of I46
	+1	N	Gln433 side chain NH_2 to carbonyl of N45 Ala429 amide to carbonyl of N45
	+2	K	Thr437 amide to carbonyl of K44 Thr437 side chain OH to amide of K44
Peptide C (N- to C-) Total H-bonds: 6	-2	P	Ser427 amide to carbonyl of P255
	-1	L	Met404 amide to carbonyl of L256
	0	L	Ala429 amide to carbonyl of L257 Gln433 side chain NH_2 to carbonyl of L257 Ser427 carbonyl to amide of L257
	+1	G	Thr437 amide to carbonyl of G258
Peptide D_{sh} (C- to N-) Total H-bonds: 4	-2	I	-
	-1	V	Ser427 amide to carbonyl of V365 Ser427 carbonyl to amide of V365
	0	L	Met404 amide to carbonyl of L364
	+1	T	Gln433 side chain NH_2 to carbonyl of T363
Peptide E (N- to C-) Total H-bonds: 6	-2	L	Ser427 amide to carbonyl of L418
	-1	R	Met404 amide to carbonyl of R419
	0	I	Ala429 amide to carbonyl of I420 Gln433 side chain NH_2 to carbonyl of I420 Ser427 carbonyl to amide of I420
	+1	A	Thr437 amide to carbonyl of A421
	+2	A	-

Figure S1

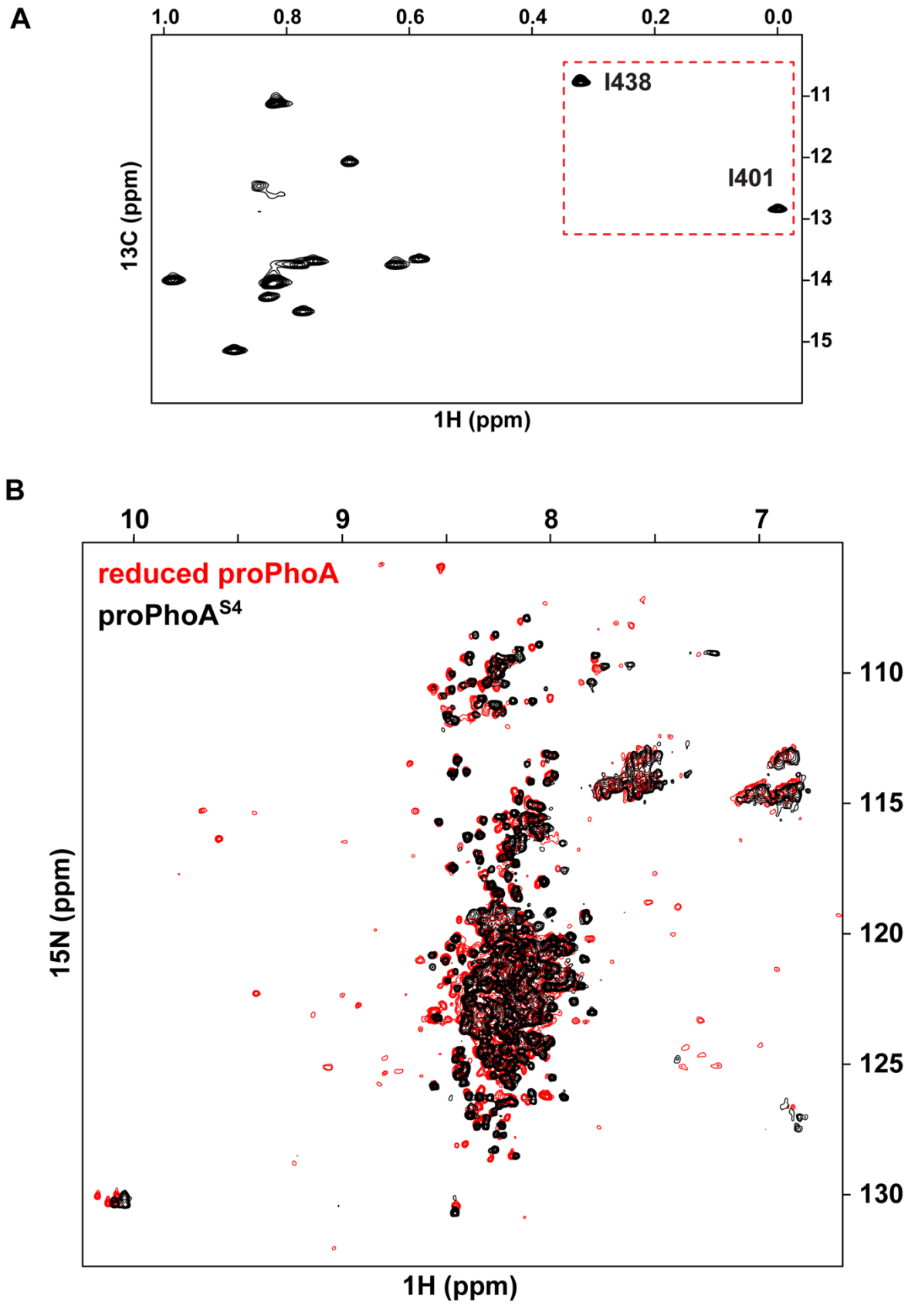
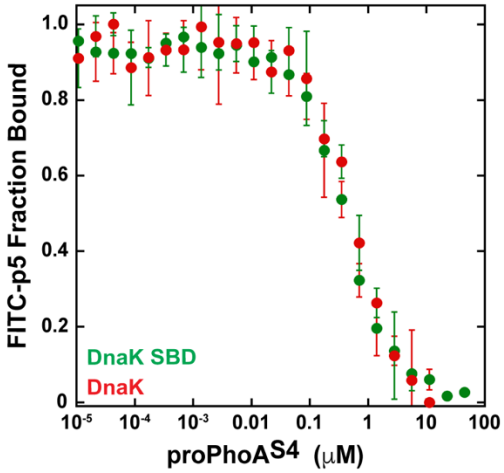

A
1.0
0.8
0.6
0.4
0.2
0.0
I438
I401
11
12
13
14
15
13C (ppm)
1H (ppm)
B
10
9
8
7
reduced proPhoA
proPhoA^S4
110
115
120
125
130
15N (ppm)
1H (ppm)

Figure S1. (*A*) Isoleucine region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) where the resonances for Ile401 and Ile438 are in a dashed box. (*B*) ^{1}H-^{15}N TROSY-HSQC spectra of 40 μM reduced proPhoA purified from *E. coli* in the presence of urea and diluted into aqueous buffer as directed in (1) (red), and 40 μM proPhoA^{S4} where all four cysteine residues in proPhoA were mutated to serines, purified as described in SI Materials and Methods (black). Spectra taken at 15 °C. The narrow dispersion of the proPhoA^{S4} resonances in the ^{1}H dimension indicates that the protein is unfolded.

Figure S2

A


FITC-p5 Fraction Bound
1.0
0.8
0.6
0.4
0.2
0
DnaK SBD
DnaK
10-5
10-4
10-3
0.01
0.1
1
10
100
proPhoAS4 (μM)

B

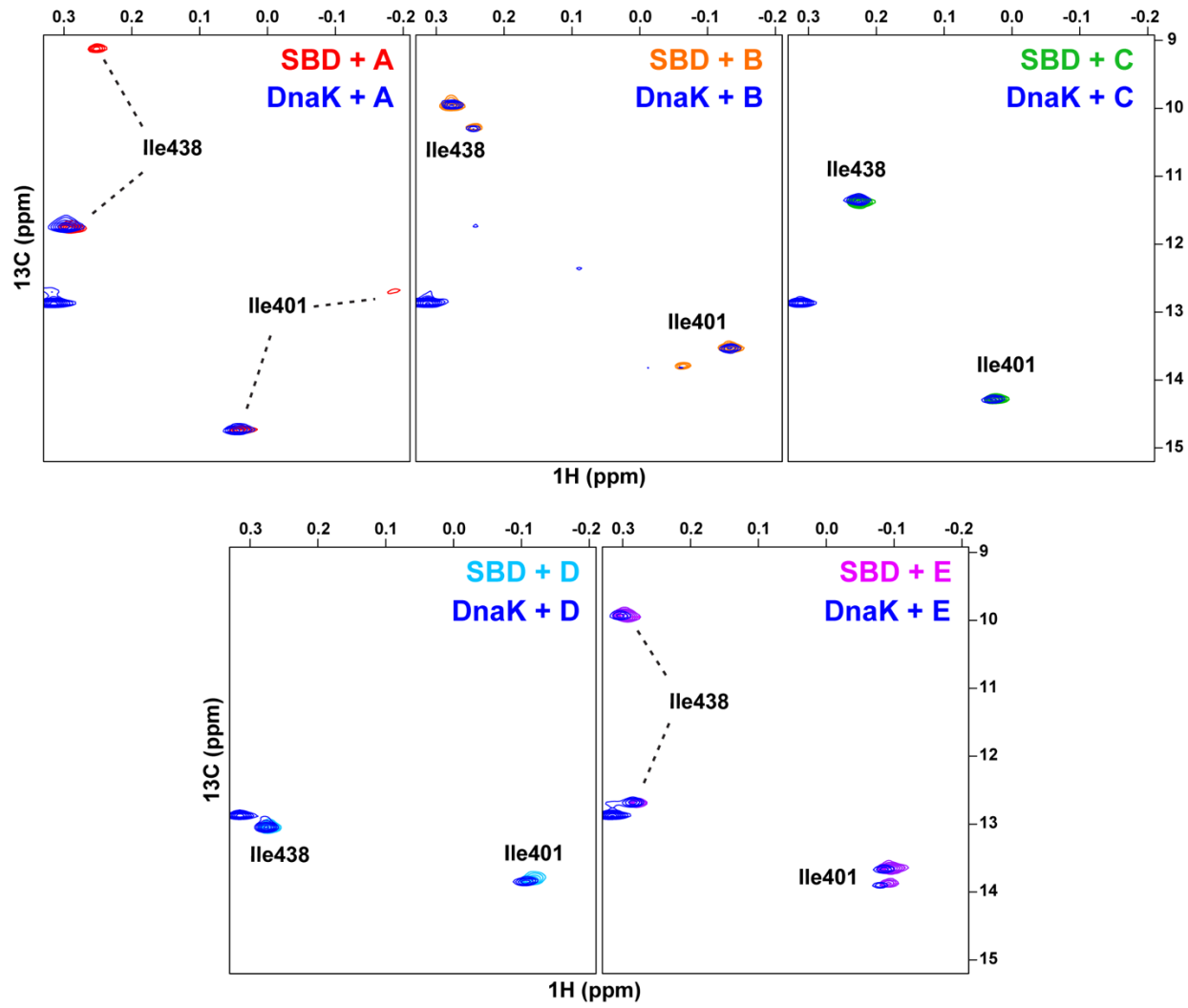

SBD + A
DnaK + A
SBD + B
DnaK + B
SBD + C
DnaK + C
SBD + D
DnaK + D
SBD + E
DnaK + E
Ile438
Ile401
13C (ppm)
1H (ppm)

Figure S2. (*A*) Binding of proPhoA^{S4} to DnaK and DnaK SBD. The decrease of the fluorescence anisotropy of FITC-labeled peptide p5 bound to the canonical binding site of DnaK (red circles) or SBD (green circles) is monitored as the competitor ProPhoA^{S4} is added in increasing concentrations and competes the peptide out of the binding site. Fluorescence anisotropy data were normalized between 0 and 1 for visualization purposes only. (*B*) Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) bound to 40 μM of each extended peptide (A, red; B, orange; C, green; D, blue; E, magenta) overlapped with the spectra of ILV-$^{13}CH_3$ full-length DnaK bound to the same peptides (blue). The resonances of Ile401 and Ile438 for peptide-bound SBD overlap those for peptide-bound, full length DnaK indicating that all peptides bind to both DnaK constructs in the same way.

Figure S3

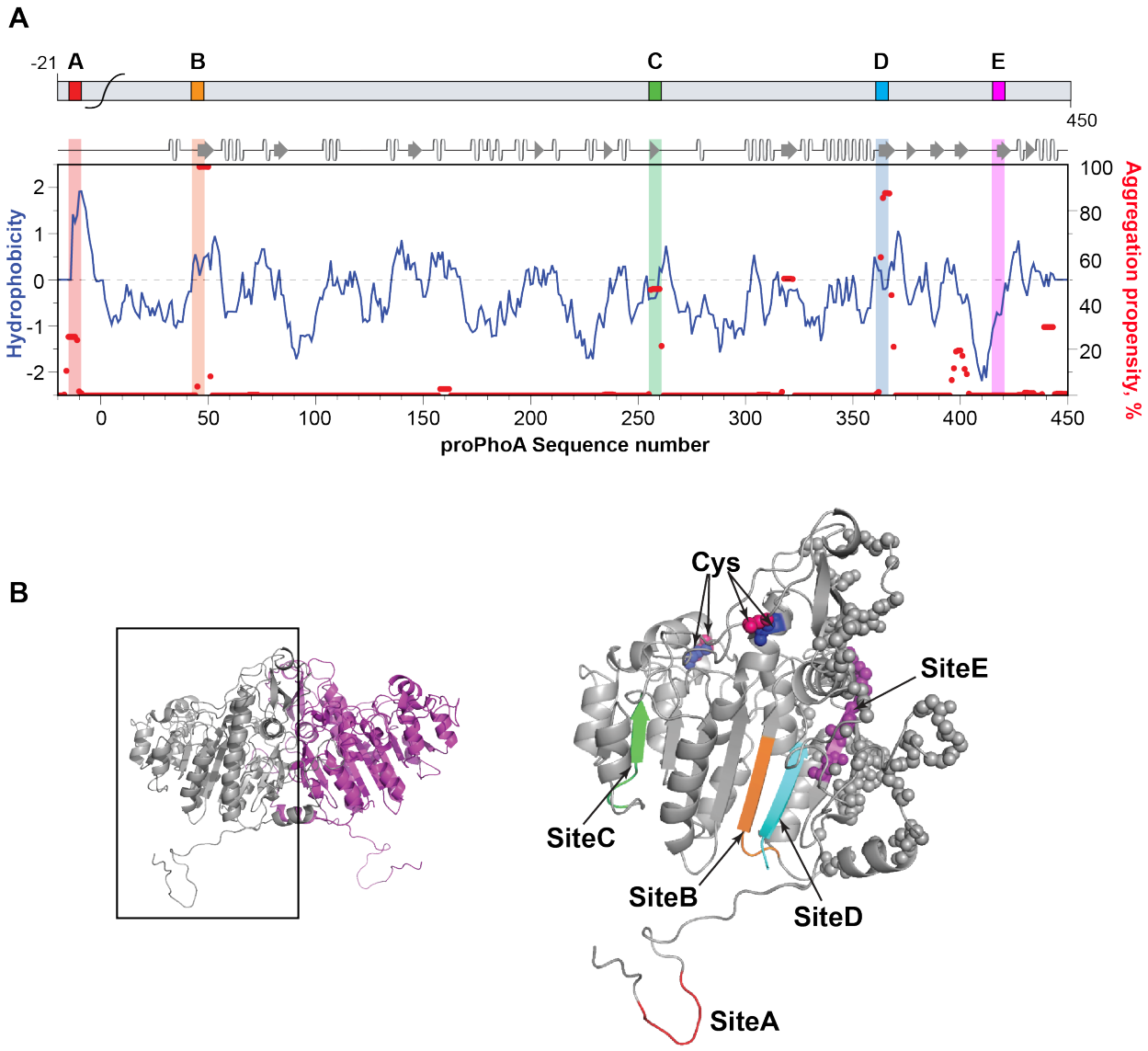


Figure S3. (*A*) Cartoon representation of the proPhoA sequence showing the sites identified as strong DnaK binders in the proPhoA peptide array (A-E) (17). The secondary structure of native PhoA is plotted along the proPhoA sequence. Below, the aggregation propensity, calculated using TANGO (24-26) (in red dots), and the hydrophobicity (27) (blue line) of the proPhoA sequence are plotted. (*B*) The structure of PhoA dimer (left) and monomer (right) (PDB ID 1Y6V) (28). The DnaK binding sites are colored on the monomer as indicated. Cysteine residues are depicted in blue or magenta spheres. Residues located at the dimer interface are represented as spheres, this region involves DnaK binding site E. Residues 336-355 were removed to allow for the clear visualization of DnaK binding sites B and D. The signal sequence was modelled in the structure using Modeller (29).

Figure S4

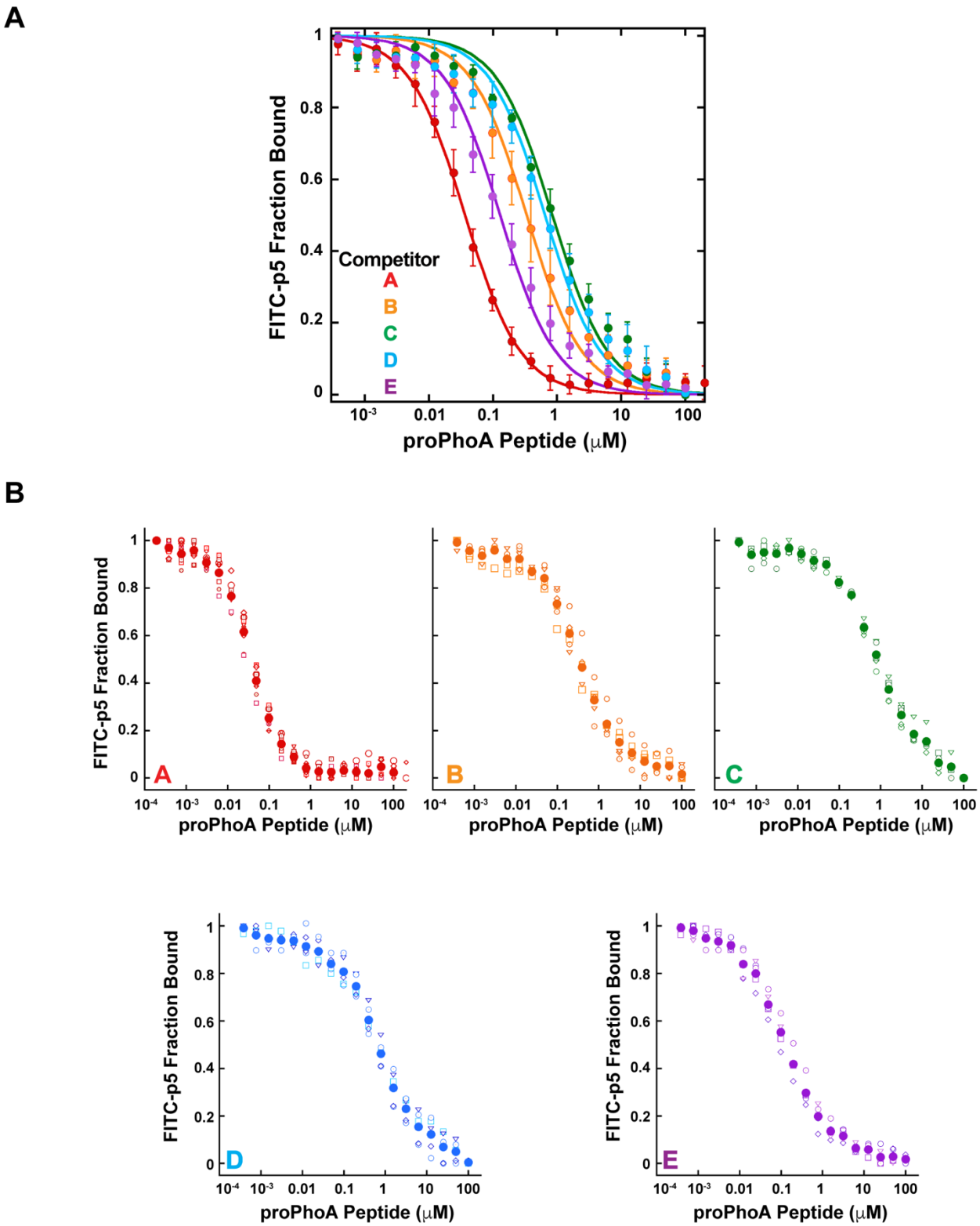


Figure S4. (*A*) Apparent binding affinity of SBD for all proPhoA peptides measured by fluorescence anisotropy. The fraction of FITC-p5 (ALLLSAPRR) bound to the DnaK SBD

is plotted against the concentration of competitor peptide added. The curves show the average data obtained from at least two independent experiments with at least three replicates each. Individual data points were fitted to the equation described in SI Materials and Methods (*B*) Data from individual experiments used to build the average curves shown in panel A (empty symbols). The average data are represented in filled symbols.

Figure S5

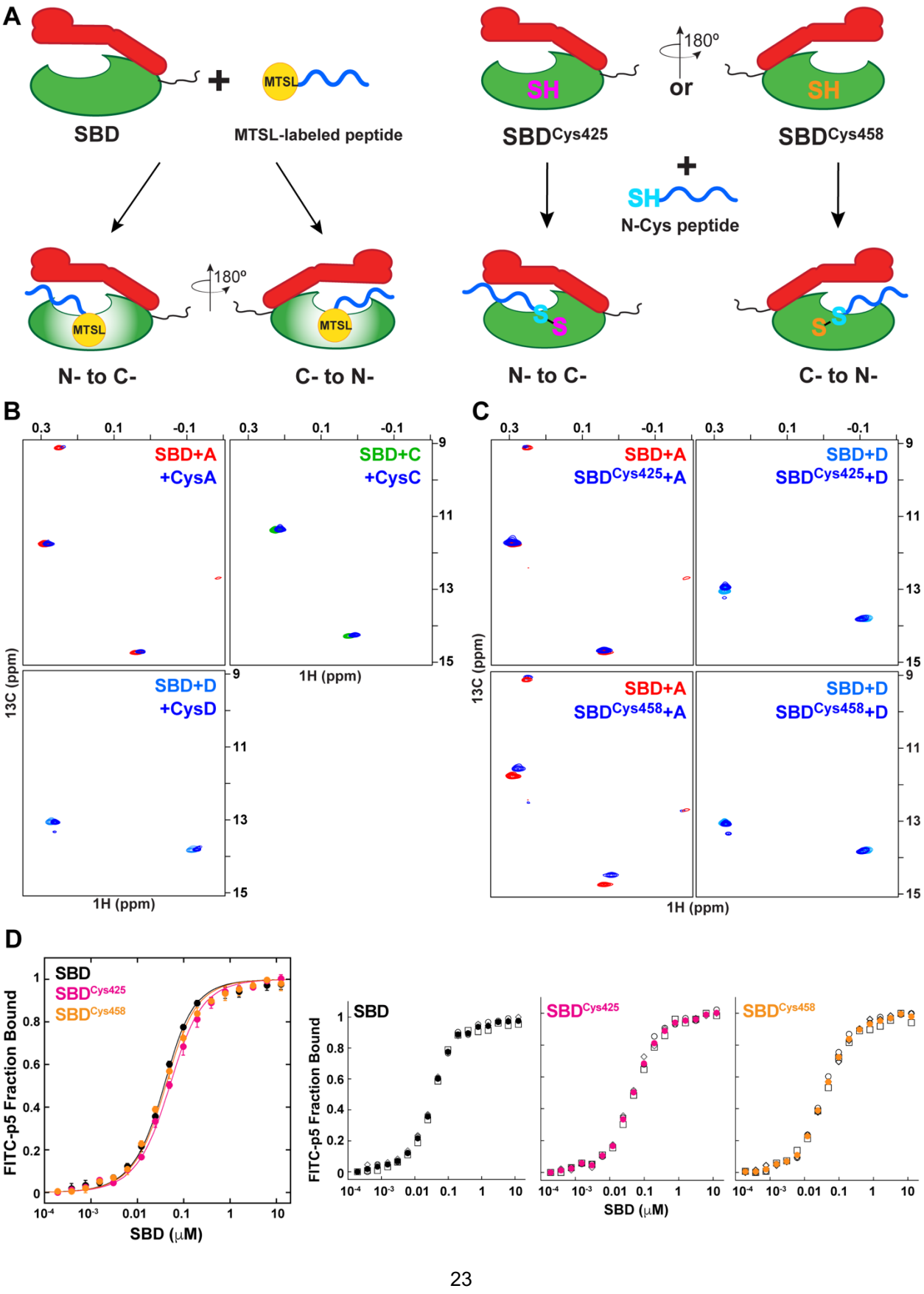

A
SBD
MTSL
MTSL-labeled peptide
SBDCys425
180°
or
SBDCys458
SH
N-Cys peptide
N- to C-
C- to N-
N- to C-
C- to N-
B
SBD+A
+CysA
SBD+C
+CysC
SBD+D
+CysD
13C (ppm)
1H (ppm)
C
SBD+A
SBDCys425+A
SBD+D
SBDCys425+D
SBD+A
SBDCys458+A
SBD+D
SBDCys458+D
13C (ppm)
1H (ppm)
D
SBD
SBDCys425
SBDCys458
FITC-p5 Fraction Bound
SBD (μM)

Figure S5. (*A*) Cartoon representation of the two approaches used to determine the orientation of the SBD-bound peptides in solution. Left: PRE, where broadening of SBD resonances in ^{1}H-^{15}N TROSY-HSQC spectra corresponding to residues on either side of the domain is caused by the spatial proximity of an MTSL label attached to the N-terminal cysteine residue on a substrate peptide. Right: disulfide cross-linking, where a cysteine placed at the N-terminus of a peptide is able to form a disulfide bond with the SBD^{Cys425} or SBD^{Cys458} (at opposite sides of the SBD) depending on the binding orientation. (*B*) Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) bound to the proPhoA peptides A, C and D (in corresponding colors) overlaid with the resonances of the SBD bound to the peptides CysA, CysC and CysD (in blue). (*C*) ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD bound to peptides A (red) and D (light blue) overlaid on the spectra of ILV-$^{13}CH_3$ DnaK SBD^{Cys425} (top) and DnaK SBD^{Cys458} (bottom) bound to peptides A and D (blue). The concentration of all SBD variants was 40 μM. Overall, the presence of Cys425 or Cys458 in the SBD does not affect binding to the peptides; we attribute small changes in the positions of Ile401 and Ile438 resonances seen for SBD^{Cys458} to subtle local changes in the electron environment of the isoleucine side chains caused by substitution of asparagine to cysteine. (*D*) The binding affinity of DnaK SBD, SBD^{Cys425} and SBD^{Cys458} for FITC-p5 peptide was assayed by fluorescence anisotropy. The fraction of FITC-p5 bound to the SBDs is plotted against the concentration of SBD added. The curve on the left shows the average data obtained from an experiment with three replicates. Panels on the right show the individual data points (empty symbols) used to build the average curves shown in the left panel (average data represented in filled symbols). The binding curves indicate that Cys425 and Cys458 mutations do not change the SBD substrate-binding affinity.

Figure S6

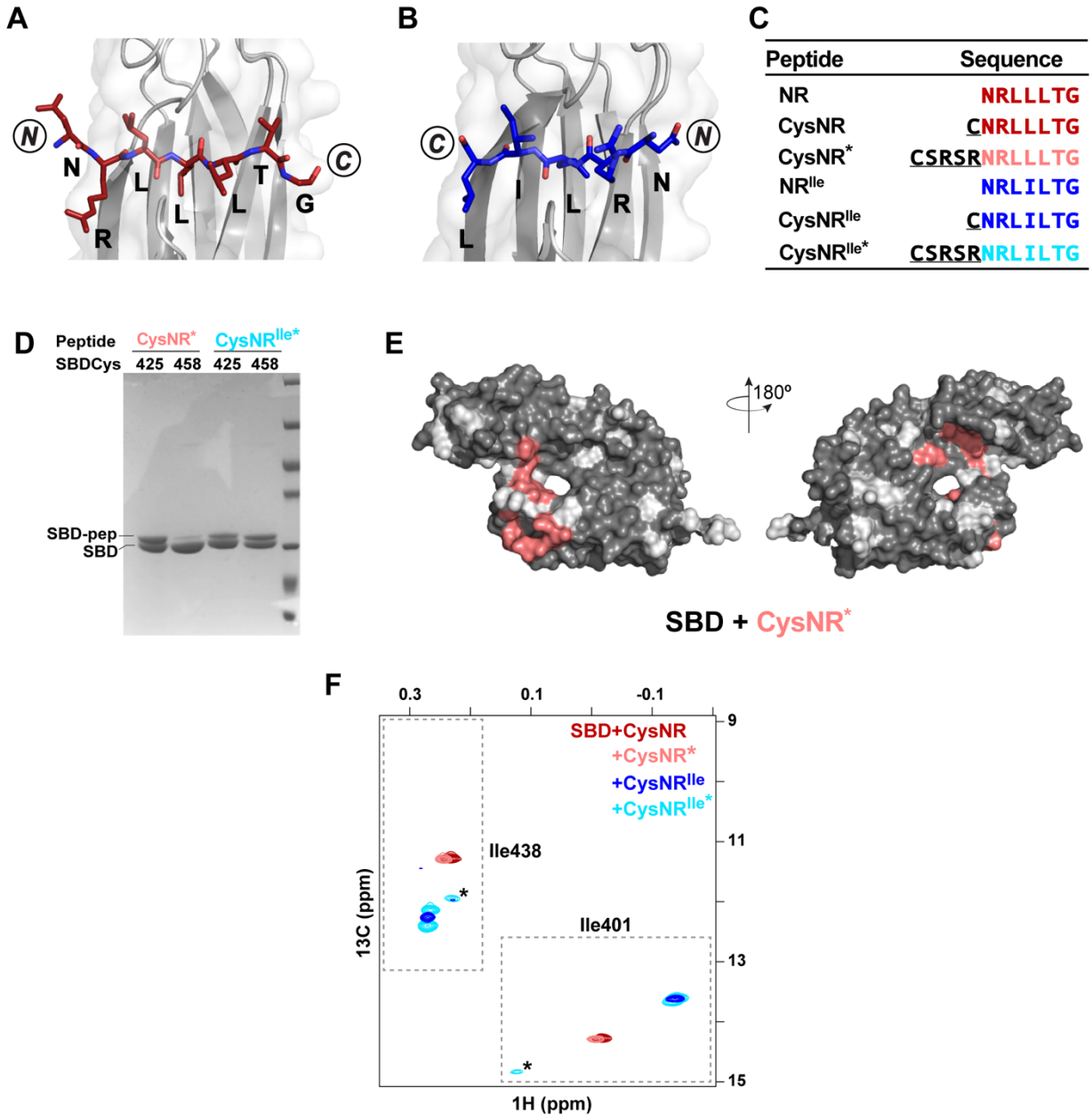


Figure S6. (*A*) Crystal structure (top view) of SBD (gray) in complex with peptide NR (in maroon sticks) which binds in a N- to C- orientation and (*B*) Crystal structure (top view) of SBD (gray) in complex with peptide NR^{Ile} (blue sticks) which binds in a C- to N-orientation (blue) (3, 4). The α-helical lid (residues 507-603) was removed to better visualize the bound peptides. (*C*) List of sequences of the peptides derived from NR and used in this study. Underlined residues were added and are not part of the original peptide sequence. (*D*) SDS-PAGE showing the ability of SBD^{Cys425} and SBD^{Cys458} to form a disulfide bond with peptides CysNR* and $CysNR^{Ile}$*. SBD-bound CysNR* forms a disulfide bond almost exclusively with Cys425 in the SBD, while $CysNR^{Ile}$* can form disulfide bonds with Cys425 or Cys458 indicating that the peptide can bind to the SBD in both orientations. (*E*) PRE NMR of SBD in complex with CysNR*. SBD residues significantly

broadened in ^{1}H-^{15}N TROSY-HSQC spectra in the presence of the spin labeled peptide are mapped on the SBD structure and colored in salmon. Residues for which no data is available are indicated in white. (*F*) Ile401 and Ile438 region of the ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) bound to 150 μM of CysNR (maroon), CysNR* (salmon), $CysNR^{Ile}$ (blue) or $CysNR^{Ile*}$ (cyan). The presence of the linker in $CysNR^{Ile*}$ resulted in appearance of additional Ile401 and Ile438 resonances in the spectra (marked with *) indicative of an additional binding mode.

Figure S7

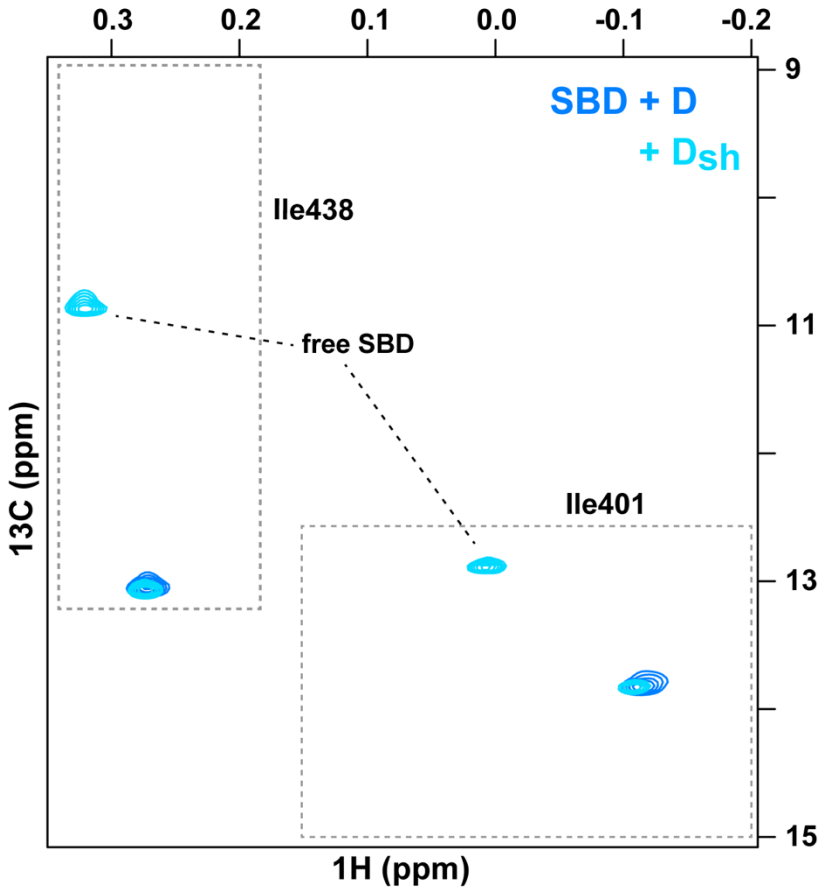


Figure S7. Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD bound to peptides D (blue) and D_{sh} (cyan). No difference in SBD binding was observed for the two versions of the peptide. SBD and peptides were at 40 μM. Peptide D_{sh} has low solubility and as a consequence, its concentration in the sample is not enough to saturate the binding sites of all SBD molecules in the sample. Thus, the Ile401 and Ile438 resonances for empty ILV-$^{13}CH_3$ DnaK SBD are observed.

Figure S8

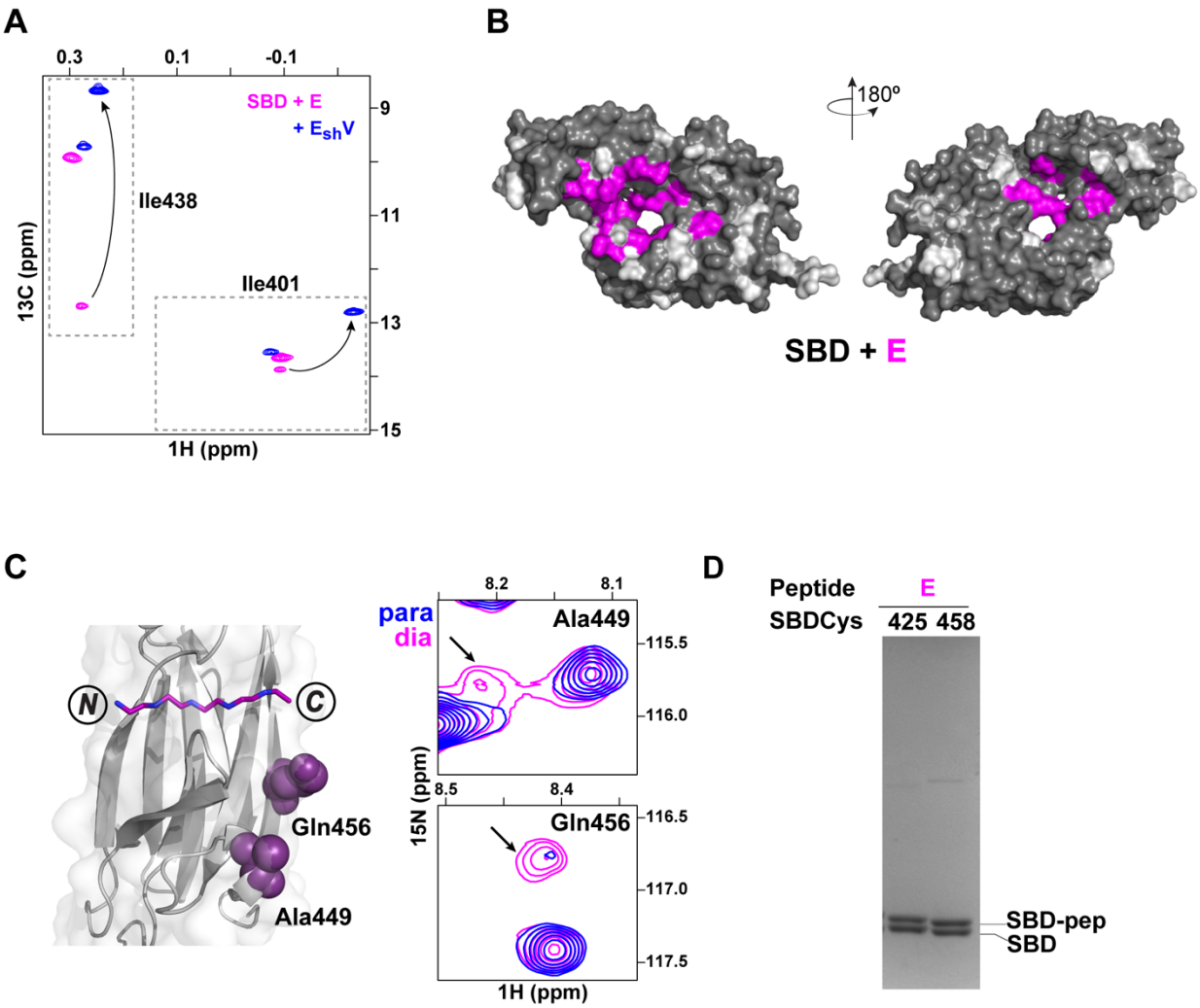


Figure S8. (*A*) Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) bound to peptides E (in magenta) and E_{sh}V (R^{415}GSQVRIA^{421}SR with L418 mutated to V, in blue). The resonances corresponding to L418 in the SBD central pocket that shift as a consequence of the mutation are marked with arrows. (*B*) and (*C*) PRE NMR of SBD complex with peptide E. In (*B*), SBD residues significantly broadened in ^{1}H-^{15}N TROSY-HSQC spectra in the presence of MTSL-labeled peptide are mapped on the SBD structure and colored in magenta. Residues for which no data is available are depicted in white. (*C*) Left: Gln456 and Ala449 are shown in purple spheres in the structure of peptide E-bound DnaK SBD. Right: ^{1}H-^{15}N TROSY-HSQC resonances of ILV-$^{13}CH_3$ DnaK SBD residues Ala449 and Gln456 bound to MTSL-labeled peptide E in paramagnetic (blue) and diamagnetic (magenta) states. Minor species are marked with arrows. Please note that the major species for these residues are not broadened in the paramagnetic state. In contrast, the minor species experiences a strong PRE effect indicating that the minor population binds in C- to N- orientation. *(D)* SDS-PAGE of SBD^{Cys425} and SBD^{Cys458} cross-linked to peptide E via disulfide bond. The peptide cross-

links to SBD^{Cys425} and SBD^{Cys458} confirming that it binds in both N- to C and C- to N-orientations.

Figure S9

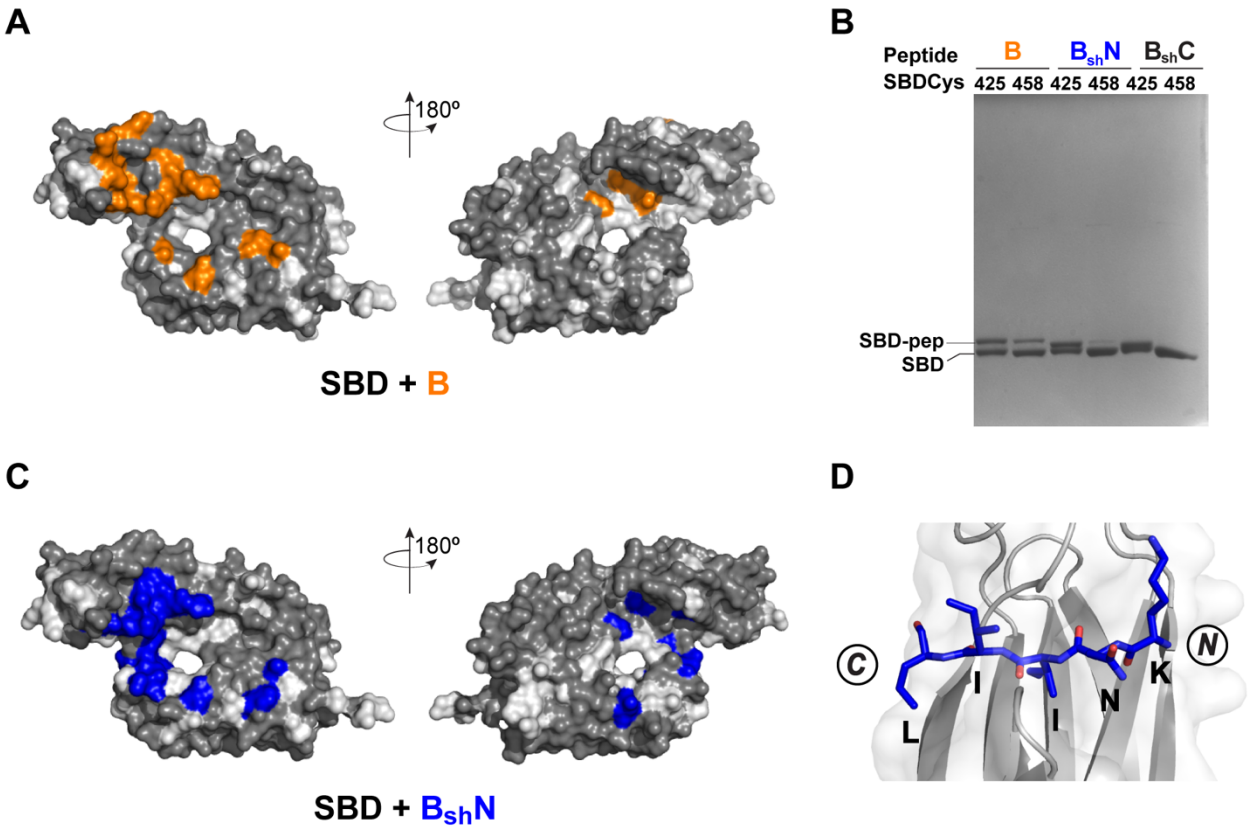


Figure S9. (*A*) and (*C*) PRE NMR of SBD complex with peptides B and $B_{sh}N$. SBD residues significantly broadened in ^{1}H-^{15}N TROSY-HSQC spectra in the presence of the spin-labeled peptide are mapped on the SBD structure and colored in orange for peptide B (*A*) and in blue for peptide $B_{sh}N$ (*C*). Residues for which no data is available are shown in white. (*B*) SDS-PAGE of SBD^{Cys425} and SBD^{Cys458} cross-linked to peptides B, $B_{sh}N$ and $B_{sh}C$ via disulfide bonding. Peptides B and $B_{sh}N$ cross-link to SBD^{Cys425} and SBD^{Cys458} confirming that they bind in both, N- to C and C- to N- orientation, respectively, whereas peptide $B_{sh}C$ can form a disulfide only with SBD^{Cys425}, indicative of N- to C- binding orientation. (*D*) Crystal structure (top view) of the SBD (gray) in complex with peptide $B_{sh}N$ (in blue sticks), which binds in a C- to N- orientation. The α-helical lid (residues 507-603) was removed to better visualize the bound peptide.

Figure S10

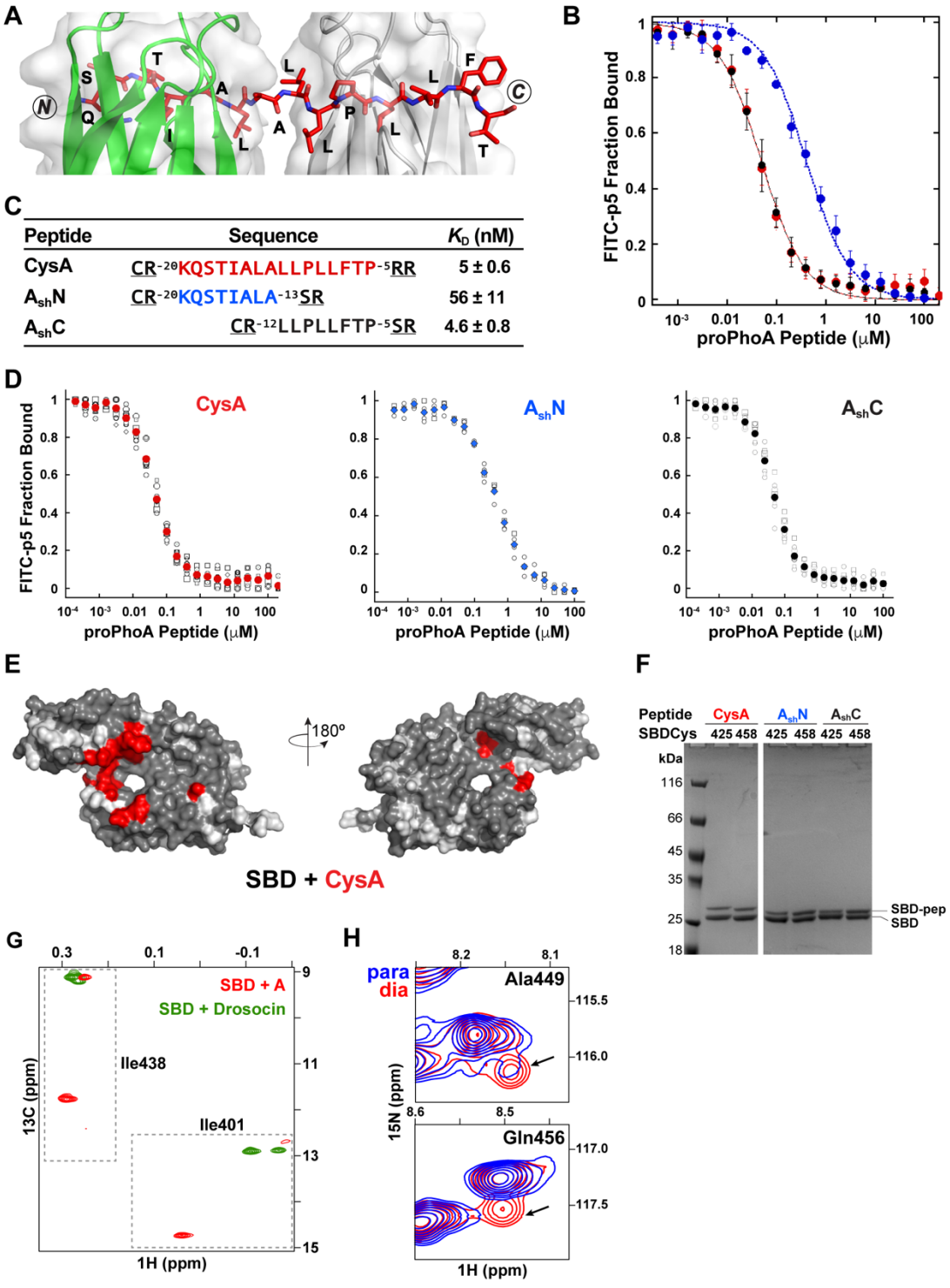


Figure S10. (*A*), Crystal structure (top view) of SBD (gray) in complex with peptide A (in red sticks) showing two SBD molecules aligned at the two-fold axis and binding to peptide A simultaneously. The sequence "LLPLLFT" binds in the N- to C- orientation, while the sequence "QSTIALA" binds in the C- to N- orientation. (*B*), Binding of the DnaK SBD to peptides CysA, A_{sh}N and A_{sh}C tested using a competition assay. The decrease of the

fluorescence anisotropy of FITC-labeled peptide p5 bound to the canonical binding site of SBD is monitored as the competitor peptides (CysA, red curve; $A_{sh}N$, blue curve; $A_{sh}C$, black curve) are added in increasing concentrations and compete the labeled peptide out of the binding site. The fraction of FITC-p5 bound to the SBDs is plotted against the concentration of competitor peptide added. The plotted data are the average of at least two experiments with three replicates each and were fitted to the equations described in SI Materials and Methods. (*C*), Table with peptide sequences reporting the apparent affinity values obtained from the curves presented in (*B*). Note that the apparent affinity of the DnaK SBD for peptide CysA is the same as that reported for binding of the DnaK SBD to peptide A (5 ±1 nM), indicating that the addition of a cysteine residue does not affect the binding. (*D*), Individual competition curves (empty symbols) used to obtain the average values (in filled symbols). (*E*), PRE NMR of SBD complex with peptide A. SBD residues significantly broadened in ^{1}H-^{15}N TROSY-HSQC spectra in the presence of the spin-labeled peptide are mapped on the SBD structure and colored in red. Residues for which no data is available are shown in white. (*F*), SDS-PAGE of SBD^{Cys425} and SBD^{Cys458} cross-linked to peptides CysA, $A_{sh}N$ and $A_{sh}C$ via disulfide bond. All three peptides cross-link to SBD^{Cys425} and SBD^{Cys458} confirming that they bind in both, N- to C and C- to N-orientations. (*G*), Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of ILV-$^{13}CH_3$ DnaK SBD (40 μM) bound to peptide A (in red) and drosocin (green), which places proline in the SBD central pocket (4). The Ile401 and Ile438 resonances at ~12.8 ppm and 9 ppm in ^{13}C dimension, respectively are assigned to proline at the 0^{th} site. (*H*), ^{1}H-^{15}N TROSY-HSQC resonances of ILV-$^{13}CH_3$ DnaK SBD residues Ala449 and Gln456 bound to MTSL-labeled peptide A in paramagnetic (blue) and diamagnetic (red) states. Minor species are marked with arrows. Please note that major species for these residues are not broadened in the paramagnetic state. In contrast, minor species experiences a strong PRE effect indicating that the minor population binds in C- to N- orientation. The positions of Gln456 and Ala449 in the SBD are shown in Figure S8C.

Figure S11

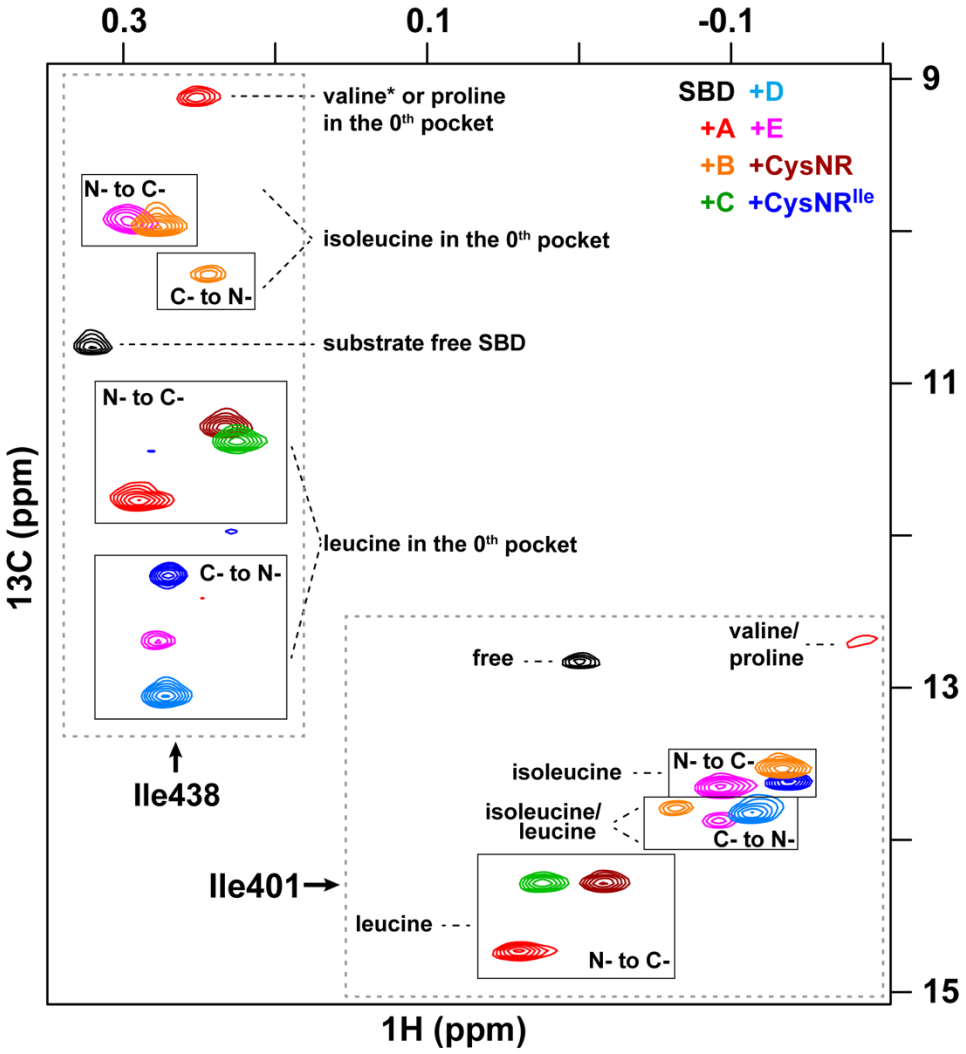


Figure S11. Ile401 and Ile438 region of ^{1}H-^{13}C HMQC spectra of 40 μM ILV-$^{13}CH_3$ DnaK SBD alone (black) and bound to peptides A (red), B (orange), C (green), D (light blue), E (magenta), CysNR (maroon) and CysNR^{Ile} (blue). Please note that the positions of SBD Ile401 and Ile438 resonances in the spectra depend on the residue of the bound substrate that occupies the SBD 0th pocket and on orientation of the binding. * – Substrate valine in the central pocket of SBD has been observed previously by Rosenzweig *et al*. (5). The boxes grouping peptides according to binding orientation are based on experimental observations and shown for illustration purposes only; at this point we cannot definitively say where the border between the positions of resonances corresponding to substrates bound in N- to C- or C- to N- modes is located.

Figure S12

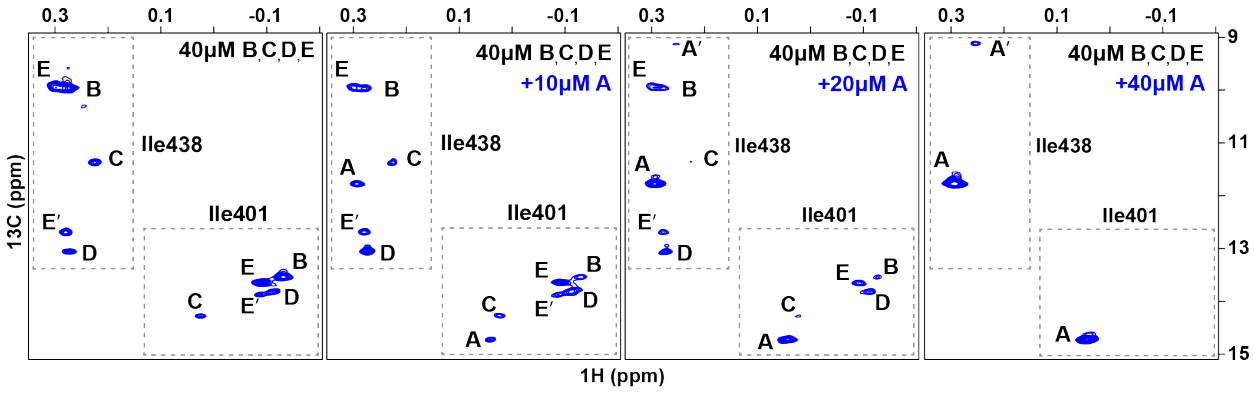


Figure S12. Titration of an equimolar (40 μM) mixture of Ile-^{13}CH3 DnaK SBD bound to peptides B, C, D and E with increasing amounts of peptide A. The Ile401 and Ile438 region of ^{1}H-^{13}C TROSY-HMQC spectra is shown. The peptide mixture:peptide A ratios are indicated.

Figure S13

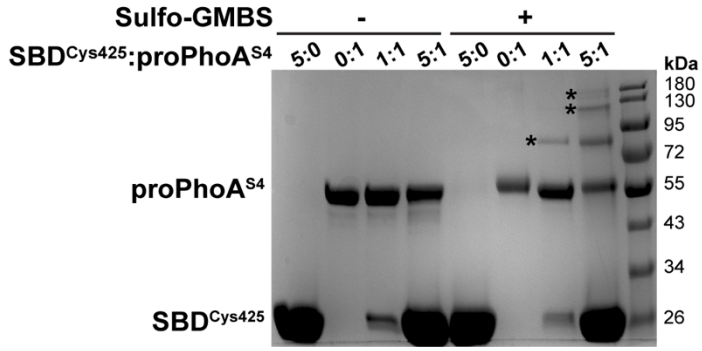


Figure S13. A complete version of SDS-PAGE presented in Figure 5B, where the first four lines correspond to SBD (lane 1), proPhoA^{S4} (lane 2) and SBD/proPhoA^{S4} complexes at 1:1 (lane 3) and 5:1 (lane 4) ratios in absence of sulfo-GMBS. Addition of the cross-linker to either SBD or proPhoA^{S4} alone does not result in higher molecular weight species (lanes 5 and 6). At 1:1 ratio of SBD to proPhoA^{S4}, 1:1 complex is observed (lane 7) while at 5:1 ratio of SBD to proPhoA^{S4}, complexes containing multiple SBD bound to one proPhoA^{S4} molecule are seen.

SI References

1. T. Saio, X. Guan, P. Rossi, A. Economou, C. G. Kalodimos, Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* **344**, 1250494 (2014).
2. M. F. Sardis *et al.*, Preprotein conformational dynamics drive bivalent translocase docking and secretion. *Structure* **25**, 1056-1067 e1056 (2017).
3. X. Zhu *et al.*, Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **272**, 1606-1614 (1996).
4. M. Zahn *et al.*, Structural studies on the forward and reverse binding modes of peptides to the chaperone DnaK. *J Mol Biol* **425**, 2463-2479 (2013).
5. R. Rosenzweig, A. Sekhar, J. Nagesh, L. E. Kay, Promiscuous binding by Hsp70 results in conformational heterogeneity and fuzzy chaperone-substrate ensembles. *eLife* **6** (2017).
6. D. L. Montgomery, R. I. Morimoto, L. M. Gierasch, Mutations in the substrate binding domain of the *Escherichia coli* 70 kda molecular chaperone, DnaK, which alter substrate affinity or interdomain coupling1. *J Mol Biol* **286**, 915-932 (1999).
7. V. Tugarinov, L. E. Kay, An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* **28**, 165-172 (2004).
8. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
9. S. P. Skinner *et al.*, CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR* **66**, 111-124 (2016).
10. A. Zhuravleva, E. M. Clerico, L. M. Gierasch, An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell* **151**, 1296-1307 (2012).
11. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**, 12366-12371 (1997).
12. J. F. Swain, E. G. Schulz, L. M. Gierasch, Direct comparison of a stable isolated Hsp70 substrate-binding domain in the empty and substrate-bound states. *J Biol Chem* **281**, 1605-1611 (2006).
13. J. L. Battiste, G. Wagner, Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355-5365 (2000).
14. W. Meng, E. M. Clerico, N. McArthur, L. M. Gierasch, Allosteric landscapes of eukaryotic cytoplasmic Hsp70s are shaped by evolutionary tuning of key interfaces. *Proc Natl Acad Sci U S A* **115**, 11970-11975 (2018).
15. T. D. Pollard, A guide to simple and informative binding assays. *Mol Biol Cell* **21**, 4061-4067 (2010).
16. Y. Cheng, W. H. Prusoff, Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22**, 3099-3108 (1973).
17. T.H. Thrall, J. Reinstein, B.M. Wöhrl, R.S. Goody, Evaluation of human immunodeficiency virus type 1 reverse transcriptase primer tRNA binding by fluorescence spectroscopy: specificity and comparison to primer/template binding. *Biochemistry* **35**(14), 4609-4618 (1996).
18. M.H. Roehrl, J.Y. Wang, G.A. Wagner, General framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* **43**(51),16056-16066 (2004).

19. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
20. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
21. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255 (1997).
22. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).
23. B. Lee, F. M. Richards, The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* **55**, 379-400 (1971).
24. A. M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, L. Serrano, Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* **22**, 1302-1306 (2004).
25. R. Linding, J. Schymkowitz, F. Rousseau, F. Diella, L. Serrano, A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J Mol Biol* **342**, 345-353 (2004).
26. F. Rousseau, J. Schymkowitz, L. Serrano, Protein aggregation and amyloidosis: confusion of the kinds? *Curr Opin Struct Biol* **16**, 118-126 (2006).
27. J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132 (1982).
28. J. Wang, K. A. Stieglitz, E. R. Kantrowitz, Metal specificity is correlated with two crucial active site residues in *Escherichia coli* alkaline phosphatase. *Biochemistry* **44**, 8378-8386 (2005).
29. B. Webb, A. Sali, Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **54**, 5.6.1-5.6.37 (2016).