

# Binding Specificity of an $\alpha$ -Helical Protein Sequence to a Full-Length Hsp70 Chaperone and Its Minimal Substrate-Binding Domain<sup>†</sup>

Carolina A. Vega,<sup>‡,§,||</sup> Neşe Kurt,<sup>‡,§</sup> Zhongjing Chen,<sup>§</sup> Stefan Rüdiger,<sup>⊥</sup> and Silvia Cavagnero<sup>\*,§</sup>

Department of Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, and Cellular Protein Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

Received July 14, 2006; Revised Manuscript Received September 13, 2006

**ABSTRACT:** Hsp70 chaperones are involved in the prevention of misfolding, and possibly the folding, of newly synthesized proteins. The members of this chaperone family are capable of interacting with polypeptide chains both co- and posttranslationally, but it is currently not clear how different structural domains of the chaperone affect binding specificity. We explored the interactions between the bacterial Hsp70, DnaK, and the sequence of a model all- $\alpha$ -helical globin (apoMb) by cellulose-bound peptide scanning. The binding specificity of the full-length chaperone was compared with that of its minimal substrate-binding domain, DnaK- $\beta$ . Six specific chaperone binding sites evenly distributed along the apoMb sequence were identified. Binding site locations are identical for the full-length chaperone and its substrate-binding domain, but relative affinities differ. The binding specificity of DnaK- $\beta$  is only slightly decreased relative to that of full-length DnaK. DnaK's binding motif is known to comprise hydrophobic regions flanked by positively charged residues. We found that the simple fractional mean buried area correlates well with Hsp70's binding site locations along the apoMb sequence. In order to further characterize the properties of the minimal binding host, the stability of DnaK- $\beta$  upon chemical denaturation by urea and protons was investigated. Urea unfolding titrations yielded an apparent folding  $\Delta G^\circ$  of  $3.1 \pm 0.9$  kcal mol<sup>-1</sup> and an  $m$  value of  $1.7 \pm 0.4$  kcal mol<sup>-1</sup> M<sup>-1</sup>.

Hsp70 chaperones participate in different fundamental biological processes, such as the folding of newly synthesized proteins (1), the refolding of stress-denatured proteins (2), the translocation of proteins across membranes (3), and the disaggregation of large aggregates (4, 5). The Hsp70 family of chaperones is found in most prokaryotic and all eukaryotic organisms. The members of the Hsp70 family have a similar general architecture, comprising an N-terminal ATPase domain and a substrate-binding domain. The recently published structure of a nearly full-length eukaryotic Hsp70 and NMR residual dipolar coupling data on a bacterial Hsp70 chaperone provide information on the relative orientation of the two domains (6, 7). The substrate-binding domain has a fairly conserved three-dimensional fold (5). The structure of the substrate-binding domain of the *Escherichia coli* Hsp70, DnaK, was elucidated by X-ray crystallography (8). It consists of a  $\beta$ -sandwich subdomain followed by  $\alpha$ -helical segments. Remarkably, the  $\beta$ -sandwich portion of the molecule is highly conserved among Hsp70 chaperones.

DnaK, the best characterized Hsp70 chaperone, associates with 9–18% of the newly synthesized proteins in *E. coli*, including incomplete ribosome-bound chains (5). This indicates that DnaK may prevent protein misfolding both co-

and posttranslationally. While it is still not entirely clear how DnaK assists in co- and posttranslational protein folding (5), it is known that its action involves the ability to transiently bind hydrophobic segments of polypeptide chains (8, 9). The bound substrate interacts directly only with the  $\beta$ -sandwich subdomain. Two elements of the binding cavity are known to play an important role in substrate binding: namely, a hydrophobic arch formed by two residues (Met<sub>404</sub> and Ala<sub>429</sub>) over the binding channel and a hydrophobic pocket accommodating a single hydrophobic side chain of the substrate (8, 9; Figure 1A). Systematic work based on cellulose-bound peptide scanning led to the realization that the DnaK binding motif typically consists of four or five nonpolar residues and two flanking regions enriched in basic residues (10).

Hsp70's ability to recognize substrates is important for protein folding and the prevention of conformational misbehavior during translation. Yet, it is not well-established how chaperones affect the secondary and tertiary structure of polypeptides and proteins, as they achieve their native state from a chaperone-bound state (5). We recently showed that bacterial Hsp70, DnaK, has the ability to behave as a "holder" upon interaction with incomplete N-terminal polypeptides and as a "folding enhancer" when interacting with full-length proteins (11) in the absence of its regulatory ATPase domain. The folding enhancer function of DnaK is likely a consequence of transient binding, leading to increased folding yields even in the absence of any significant interactions at equilibrium (11). This study was performed with DnaK's substrate-binding domain DnaK- $\beta$  (Figure 1A), regarded as a minimal chaperone model containing the entire active site

<sup>†</sup> This work was supported by NIH Grant GM068535.

<sup>\*</sup> Corresponding author. Phone: 608-262-5430. Fax: 608-262-9918. E-mail: cavagnero@chem.wisc.edu.

<sup>‡</sup> Joint first authors.

<sup>§</sup> University of Wisconsin—Madison.

<sup>||</sup> Present address: Potentia Pharmaceuticals, Louisville, KY 40202.

<sup>⊥</sup> Utrecht University.

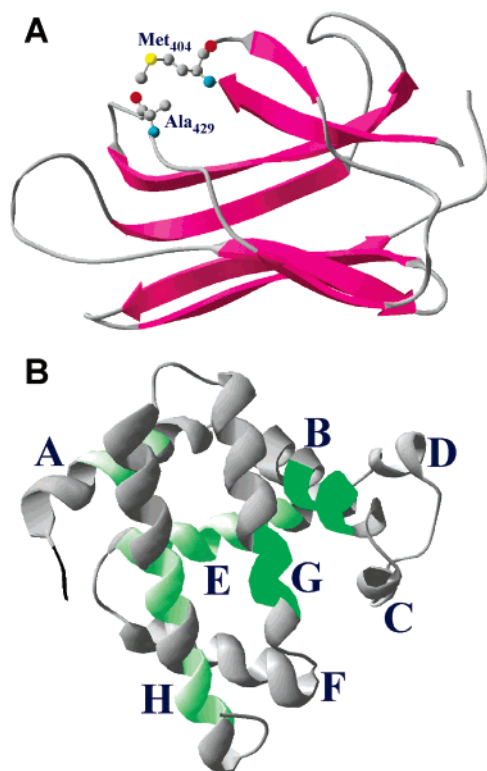


FIGURE 1: Ribbon representation of DnaK- $\beta$  and apoMb structures. (A) Three-dimensional structure of DnaK(393–507), here denoted as DnaK- $\beta$  (12). Met<sub>404</sub> and Ala<sub>429</sub>, which form a hydrophobic arch over the substrate-binding site, are shown at atomic resolution. (B) Three-dimensional structure of sperm whale myoglobin. The specific DnaK binding sites identified by cellulose-bound peptide scanning experiments are mapped on the structure in green. The two highest affinity sites are shown in dark green. The  $\alpha$ -helices in the native structure are labeled A through H. Coordinates were derived from the X-ray structure of carbonmonoxymyoglobin (47) (PDB code 1MBC). Images were created with the Swiss-PdbViewer software package (version 3.7) (48).

and retaining functionality but lacking ATP-mediated regulation. The basic structural features of DnaK- $\beta$  and its conformational changes associated with the binding of small peptides have been examined before (8, 12). However, quite surprisingly, the overall stability and thermodynamic properties of this chaperone model remain undefined. In order to provide a better DnaK- $\beta$  characterization in aqueous solution, the stability of this chaperone toward chemical denaturation and its response to pH changes have been investigated here.

The all- $\alpha$ -helical model protein apomyoglobin (apoMb,<sup>1</sup> Figure 1B) was chosen as a chaperone substrate. Its typical globin fold is extensively represented across all three kingdoms of life (13). Sperm whale apoMb is the best characterized member of the globin family, in terms of its structure, folding mechanisms, and interaction with chaperones (11, 14–19). This protein is routinely produced in bacterial expression systems (20, 21), and it is a suitable model for investigating substrate interactions with chaperones, because portions of its sequence are known to severely self-associate in the absence of chaperones (19). Interaction with DnaK is known to enhance folding yields by preventing

apoMb's misfolding (11). However, the number and location of the potential DnaK binding sites for the apoMb sequence were not identified in prior investigations. Yet, this information is of key importance to understand how chaperone binding affects polypeptide tertiary structure, dynamics, and folding.

Here, we report a comparative analysis of the binding specificity of both the full-length DnaK chaperone and its substrate-binding domain, DnaK- $\beta$ , for the apoMb sequence. In order to further characterize the properties of the minimal binding host, the secondary structure and stability of DnaK- $\beta$  upon chemical denaturation by urea and protons were also investigated. A comparison between the minimal substrate-binding domain and full-length DnaK is of key importance to understand the role of the regulatory ATPase domain on binding specificity. The cellulose-bound peptide library scanning technique, which exploits the screening of substrate segments for chaperone binding activity, was employed to evaluate the interactions between apoMb and the two forms of DnaK. The all- $\alpha$ -helical protein apoMb is employed as a model substrate. Chaperone binding sites are reported to be less common in segments corresponding to helices in native proteins (10). We have, however, found that there are relatively strong binding sites for DnaK along the apoMb sequence. Substrate binding is discussed in the context of the nonpolar character of the substrate sequence and the predictions by an algorithm developed to estimate DnaK binding (10). In addition, we provide a quantitative assessment, by 2D NMR spectroscopy, of DnaK- $\beta$ 's affinity for a peptide corresponding to the strongest apoMb binding site.

## EXPERIMENTAL PROCEDURES

**Preparation of Chaperones.** The substrate-binding domain of DnaK, DnaK- $\beta$  (comprising residues 393–505 of the parent chaperone), was prepared as described by Kurt et al. (11). DnaK- $\beta$  (pI 7.0) was prepared in His-tagged form (11). Its molecular mass is 14579 Da. This value takes into account the loss of the N-terminal methionine, which takes place upon *E. coli* expression. After protein purification, DnaK- $\beta$  was refolded in 10 mM sodium acetate adjusted to pH 6.0. Stock solutions were flash-frozen with liquid N<sub>2</sub> and stored at –80 °C. Lack of Trp and other aromatic amino acid side chains in the DnaK- $\beta$  sequence prevented the use of intrinsic absorbance and the extinction coefficient for concentration determination. The DnaK- $\beta$  concentration was determined spectrophotometrically according to the procedure by Bradford (22) based on calibrations of protein concentration carried out by amino acid analysis. The Bradford reagent used here (Coomassie Bradford protein kit; Pierce, Rockford, IL) led to a net absorbance for DnaK- $\beta$  solutions corresponding to an apparent extinction coefficient of  $22200 \pm 600 \text{ M}^{-1} \text{ cm}^{-1}$  at 595 nm. This value is reliable over the linear range of 0.2–0.4 net absorbance (at 595 nm). The overall  $\beta$ -sheet structure of DnaK- $\beta$  and its ability to cooperatively fold/unfold were verified by far-UV circular dichroism (CD) and urea titrations, respectively (see below). DnaK was purchased from Stressgen (Victoria, British Columbia, Canada).

**Circular Dichroism Spectroscopy.** CD experiments were performed with an MOS-450 spectropolarimeter from BioLogic (BioLogic Science Instruments, Claix, France). CD

<sup>1</sup> Abbreviations: apoMb, apomyoglobin; CD, circular dichroism; HPLC, high-performance liquid chromatography; MRE, mean residue ellipticity; pI, isoelectric point; IAANS, 2-[4'-(iodoacetamido)anilino]-naphthalene-6-sulfonic acid.

data were collected on a 30.8  $\mu\text{M}$  DnaK- $\beta$  sample. A 0.1 cm path length quartz cuvette was used (Hellma, Müllheim, Germany). Wavelength scanning was performed in the 200–250 nm far-UV spectral region. Data were collected in 1 nm steps, with acquisition duration of 10 s/nm. Mean residue ellipticities (MRE, in units of  $\text{deg cm}^2 \text{dmol}^{-1}$ ) were calculated according to

$$\text{MRE} = \theta / (10CN_A l) \quad (1)$$

where  $\theta$  is the ellipticity in millidegrees,  $C$  is the concentration of the protein in moles per liter,  $N_A$  is the number of amino acids ( $N_A = 134$  for DnaK- $\beta$ ), and  $l$  is the cuvette path length in centimeters.

**pH Titrations.** The effect of solution pH on DnaK- $\beta$ 's secondary structure was probed by pH titrations. Data were collected by far-UV circular dichroism at 210 nm on a 3.4  $\mu\text{M}$  DnaK- $\beta$  solution in 10 mM sodium acetate buffer, pH 6.0. Solution pH was varied by stepwise addition of 0.1 M HCl. A 1 cm path length quartz cuvette was used (Hellma, Müllheim, Germany). Data were collected at room temperature and baseline-corrected against a buffer blank.

**Urea Titrations.** The thermodynamic stability of DnaK- $\beta$  was tested by urea titration, upon progressively increasing urea concentration at constant temperature. Data were collected by far-UV CD at 235 nm, with an integration rate of 10 s/nm on a 12.4  $\mu\text{M}$  DnaK- $\beta$  solution in 10 mM sodium acetate buffer, pH 6.0. A 1 cm path length quartz cuvette was used (Hellma, Müllheim, Germany). Accurate concentrations of urea stock solutions were determined by refractometry (23). Data were collected at room temperature and baseline-corrected against a buffer blank. Standard unfolding free energies in aqueous solution ( $\Delta G^\circ_{\text{H}_2\text{O}}$ ) and  $m$  values were calculated via the method of Santoro and Bolen (24), according to

$y =$

$$\frac{[(b_n + m_n[\text{U}]) + (b_u + m_u[\text{U}])] \exp\left[-\left(\frac{\Delta G^\circ_{\text{H}_2\text{O}}}{RT} - \frac{m[\text{U}]}{RT}\right)\right]}{1 + \exp\left[-\left(\frac{\Delta G^\circ_{\text{H}_2\text{O}}}{RT} - \frac{m[\text{U}]}{RT}\right)\right]} \quad (2)$$

where  $y$  is the observed variable parameter and  $b_n$  and  $m_n$ , and  $b_u$  and  $m_u$ , are the slope and intercept of the pre- and posttransition baselines, respectively. The  $m$  value reflects the dependence of  $\Delta G^\circ_{\text{H}_2\text{O}}$  on denaturant concentration ( $[\text{U}]$ ), and it is reported to be proportional to the fractional exposure of nonpolar surface area upon protein unfolding (25). Accuracies on  $\Delta G^\circ_{\text{H}_2\text{O}}$  and  $m$  were calculated as the difference between the errors determined in two independent experiments.

**Peptide Synthesis.** The QRKLFFNLRKTKQC peptide, whose sequence encompasses the Q<sub>132</sub>–Q<sub>144</sub> residues of the  $\sigma^{32}$  protein (26) with an additional C-terminal cysteine (here denoted as  $\sigma^{32}\text{Q}_{132}$ –Q<sub>144</sub>-C), was synthesized and purified at the local Peptide Synthesis Facility (Biotechnology Center, University of Wisconsin–Madison). The peptide was labeled at its C-terminal Cys with the IAANS fluorophore (Molecular Probes–Invitrogen, Carlsbad, CA) and purified according to published procedures (26). Purity was assessed by reverse-phase chromatography on a C<sub>18</sub> column (Grace Vydac,

Hesperia, CA). The HGQDILIRLFKSH and IIVLHSRH peptides, corresponding to apoMb's amino acids 24–36 and 111–119, respectively, were synthesized and purified as described elsewhere (27). Briefly, the peptides were prepared by solid-phase peptide synthesis in an Applied Biosystems model 432A synthesizer using Fmoc chemistry and purified by reverse-phase HPLC with a C<sub>18</sub> column. Both peptides were blocked with a primary amide group at the C-terminus. ApoMb<sub>24–36</sub> was <sup>15</sup>N-labeled at G<sub>25</sub>, I<sub>28</sub>, I<sub>30</sub>, F<sub>33</sub>, and S<sub>35</sub> and apoMb<sub>111–119</sub> was <sup>15</sup>N-labeled at L<sub>115</sub>. Fmoc-<sup>15</sup>N-labeled G, I, F, and S were purchased from Isotec Inc. (Miamisburg, OH), and Fmoc-<sup>15</sup>N-L was purchased from Cambridge Isotope Laboratories (Andover, MA).

**Peptide Binding by Fluorescence Titrations.** The binding of IAANS-labeled  $\sigma^{32}\text{Q}_{132}$ –Q<sub>144</sub>-C to DnaK- $\beta$  was assessed by fluorescence emission spectroscopy. Binding titrations were performed by stepwise addition of DnaK- $\beta$  to a solution containing a fixed amount of labeled peptide (0.2  $\mu\text{M}$  initial concentration). The solution was equilibrated for at least 5 min under gentle stirring after each addition. The intensity of the fluorescence signal does not change over time, indicating that this equilibration time is sufficiently long. Experiments were run at room temperature in 100 mM sodium acetate and 25 mM Hepes at pH 6.0. Fluorescence emission by the IAANS label was monitored at 475 nm, upon excitation at 326 nm. Slit widths of 1 and 2 mm for excitation and emission channels, respectively, were employed. A 420 nm cut-on filter was used in the emission channel. Background corrections for residual light scattering were carried out on all titration points by subtracting the apparent fluorescence signal resulting from parallel control titrations from the emission signal at 475 nm of the titration sample. The control titrations were carried out under identical conditions except for the fact that unlabeled peptide was used. Observed fluorescence signals ( $F_{\text{obs}}$ ) were fit to the relation:

$$F_{\text{obs}} = F_{\text{Pep}} + \frac{(F_{\text{PepCh}} - F_{\text{Pep}})}{2[\text{Pep}]_{\text{T}}} \{K_d + [\text{Ch}]_{\text{T}} + [\text{Pep}]_{\text{T}} - [(K_d + [\text{Ch}]_{\text{T}} - [\text{Pep}]_{\text{T}})^2 + 4K_d[\text{Pep}]_{\text{T}}]^{1/2}\} \quad (3)$$

where Pep, Ch, and PepCh represent the peptide, the chaperone, and the peptide–chaperone complex, respectively.  $F$  denotes the fluorescence emission intensity of the species of interest. The subscript T stands for total. The observed fluorescence  $F_{\text{obs}}$  results from the sum of fluorescence from the IAANS-labeled free peptide and the chaperone–polypeptide complex. Equation 3, expressing  $F_{\text{obs}}$  in terms of known total concentrations and the unknown dissociation constant, was derived starting from the conventional definition of dissociation constant and mass balance relations. Data for each titration were fit to eq 3 with three adjustable parameters ( $F_{\text{PepCh}}$ ,  $F_{\text{Ch}}$ , and  $K_d$ ), yielding the dissociation constant. The experimental accuracy of the equilibrium dissociation constant  $K_d$  reflects the absolute value of the curve-fitting error averaged over two independent experiments.

**Peptide Binding by NMR Titrations.** The binding of apoMb<sub>24–36</sub> to DnaK- $\beta$  was tested by NMR spectroscopy via 2D <sup>1</sup>H–<sup>15</sup>N HSQC-detected titrations. The binding kinetics of apoMb<sub>24–36</sub> to DnaK- $\beta$  is slow on the chemical shift time scale. Experiments were carried out by adding concentrated



aliquots (2 mM) of selectively  $^{15}\text{N}$ -labeled apoMb<sub>24–36</sub> peptide to a solution containing a fixed amount of unlabeled DnaK- $\beta$  ( $\sim 0.1$  mM), up to a 12.5-fold excess. Sensitivity-enhanced  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments (28) were performed at 30 °C on a Varian INOVA600 MHz spectrometer equipped with a  $^1\text{H}\{^{13}\text{C},^{15}\text{N}\}$  triple resonance probe with a triple axis gradient. Each peptide addition was followed by thermal equilibration in the spectrometer for 10 min. HSQC spectra were acquired with 96 increments in the indirect dimension in States mode over a 1208 Hz sweep width. A total of 2048 complex points were collected in the direct dimension, with a 6000 Hz sweep width. NMR data were processed with the NMRPipe (29) and NMRView (30) software packages. Proton chemical shifts were referenced relative to internal 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS).  $^{15}\text{N}$  chemical shifts were referenced indirectly as described (31). Resonance assignments are reported elsewhere (27). ApoMb<sub>24–36</sub> NH resonance intensities for the chaperone-bound form of the peptide were followed as a function of peptide concentration. Intensities were then converted to concentrations, based on the assumption that peptide binding to DnaK- $\beta$  has reached saturation at 10-fold excess of peptide over DnaK- $\beta$ . Concentrations of bound apoMb<sub>24–36</sub> were used to calculate the fraction of bound chaperone,  $\theta$ , according to

$$\theta = [\text{PepCh}]/[\text{Ch}]_{\text{T}} \quad (4)$$

where  $[\text{Ch}]_{\text{T}}$  is the total DnaK- $\beta$  concentration and  $[\text{PepCh}]$  is the concentration of the peptide–chaperone complex, coinciding with the concentration of bound apoMb<sub>24–36</sub>. The resulting  $\theta$  values were then plotted as a function of free apoMb<sub>24–36</sub> (i.e.,  $[\text{Pep}]$ ) and then fit to the expression

$$\theta = [\text{Pep}]/(K_{\text{d}} + [\text{Pep}]) \quad (5)$$

to derive the dissociation constant for the DnaK–apoMb<sub>24–36</sub> complexation event,  $K_{\text{d}}$ . Uncertainties in  $K_{\text{d}}$  were estimated from the curve-fitting error.

**Cellulose-Bound Peptide Scanning.** The DnaK and DnaK- $\beta$  binding affinity for different regions of apoMb's amino acid sequence was tested by the cellulose-bound peptide scanning method. Analysis was performed as described elsewhere (10, 26), with minor modifications. Briefly, a covalently cellulose-bound series of peptides was synthesized on a cellulose solid support (JPT Peptide Technologies GmbH, Berlin, Germany). The peptide library comprises a linear series of 13-mer peptides with 10-residue overlaps. The DnaK and DnaK- $\beta$  affinity for apoMb was tested by overnight incubation at room temperature of the peptide-containing cellulose membrane with either DnaK [150 nM in incubation buffer: 170 mM NaCl, 6.4 mM KCl, 0.05% (v/v) Tween 20, 5.0% (w/v) sucrose, 31 mM Tris-HCl, pH 7.6] or DnaK- $\beta$  (150 or 1500 nM in incubation buffer at pH 6.0). After incubation, the bound chaperone was transferred to a PVDF membrane using a semidry blotter (Owl Separation systems, Portsmouth, NH). The solutions used during the electroblotting steps do not contain methanol. The electroblotting was carried out as four individual short transfers ( $1 \times 20$  min and  $3 \times 30$  min), leading to four separate blots. DnaK and DnaK- $\beta$  were detected with anti-DnaK mouse monoclonal antibodies (Stressgen, Victoria, British Columbia, Canada) and with a

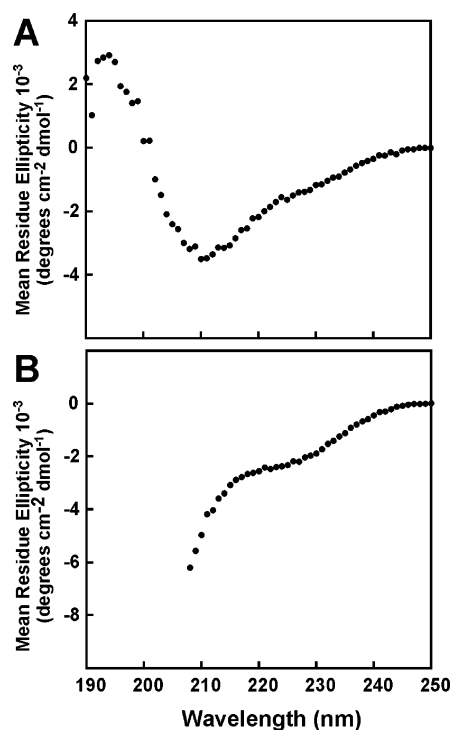


FIGURE 2: Far-UV CD spectrum of DnaK- $\beta$ . Data have been collected at room temperature for a 30.8  $\mu\text{M}$  DnaK- $\beta$  solution in 10 mM sodium acetate at (A) pH 6.0 and (B) pH 2.5.

chemifluorescent ECF substrate on a Typhoon Trio imager (both ECF substrate and imager are from Amersham Biosciences, Little Chalfont, U.K.). All of the peptide scanning experiments were performed in duplicate.

Images were processed using Adobe Photoshop (version 5.5, Adobe Systems Inc.). Background subtraction was applied to each blot. Then, the four blots from each peptide scan were superimposed and averaged, obtaining one image for each experiment. For the DnaK peptide scan, the first blot of the four was not averaged since it presented very high background. After intensities were averaged over two independent experiments, a contrast of 75 was applied to the final image. Blot intensities were analyzed using the ImageQuant software (version 5.2, Molecular Dynamics, 1999), the background was subtracted, and intensities were obtained for each spot. The data sets were normalized by imposing that the summation of the values for each experimental condition is equal to a given constant (set to 10 in Figures 7 and 10).

## RESULTS

**DnaK- $\beta$  Secondary Structure and Stability.** The overall stability of DnaK- $\beta$  (i.e., DnaK's substrate-binding domain) was investigated by probing the change in its secondary structure upon pH titration and chemical denaturation. Figure 2 displays the far-UV circular dichroism (CD) spectrum of DnaK- $\beta$  at pH 6.0, which reveals the secondary structure of a predominantly  $\beta$ -sheet protein. This is in agreement with the published NMR solution structure of a similar DnaK- $\beta$  construct (12) (Figure 1A) and with the X-ray crystal structure of the substrate-binding domain containing the helical lid (8). A minimum in MRE is observed close to 208 nm. The protein's secondary structure is progressively perturbed upon decreasing the solution pH. The far-UV CD

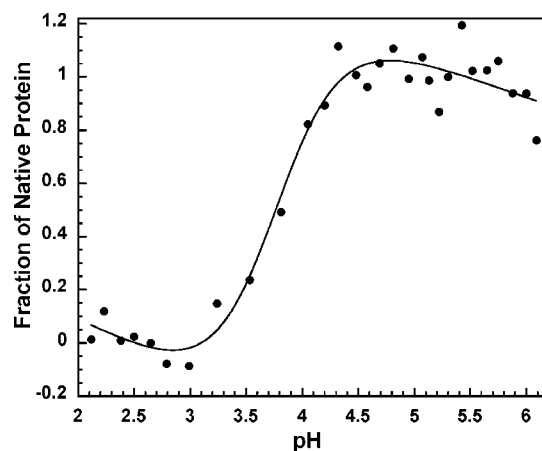


FIGURE 3: Equilibrium pH titration of DnaK- $\beta$ . Data were collected by far-UV CD at 210 nm at room temperature. Experimental points were fit to a generic two-state expression similar to eq 2 to optimize data visualization.

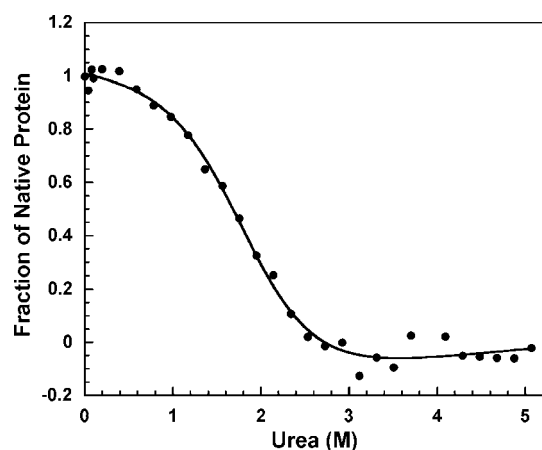


FIGURE 4: Equilibrium urea titration of DnaK- $\beta$ . Protein unfolding was followed by far-UV CD at 235 nm at room temperature in a 10 mM sodium acetate solution.

spectrum of the acid-unfolded protein at pH 2.5 is displayed in Figure 2B. Unfolding proceeds according to an apparent two-state transition (Figure 3). The chaperone begins unfolding at pH = 4.5 and it is fully unfolded at pH = 3.0. The transition midpoint is at pH  $\sim$ 3.8.

The overall thermodynamic stability of DnaK- $\beta$  was investigated by urea titration. At urea concentrations greater than 1.0 M, DnaK- $\beta$  begins to lose its secondary structure, and it unfolds according to an apparent two-state transition (Figure 4). The lack of intrinsic fluorescence by the DnaK- $\beta$  construct (DnaK- $\beta$  contains no Trp nor aromatic amino acids) prohibited observation of tertiary structure variations upon unfolding. While the two-state character of the unfolding transition was not rigorously tested (32, 33), the sharp character of the transition in both urea and pH titrations shows that equilibrium unfolding intermediates are not significantly populated. The reversibility of the unfolding transition was confirmed by the complete regain of secondary structure content upon refolding the urea-unfolded species at the end of the titration. Fitting the urea titration data to eq 2 (Figure 4), which utilizes both pre- and posttransition baselines, yields an apparent unfolding  $\Delta G^\circ$  value of  $3.1 \pm 0.9$  kcal/mol and a corresponding  $m$  value of  $1.7 \pm 0.4$  kcal mol $^{-1}$  M $^{-1}$ .

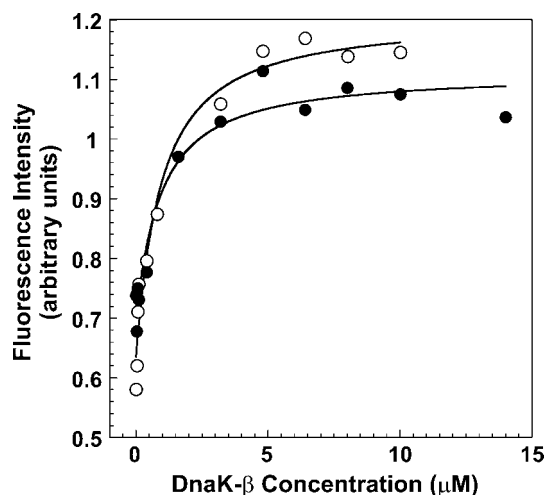


FIGURE 5: Fluorescence titrations illustrating the binding of  $\sigma^{32}$ Q<sub>132</sub>-Q<sub>144</sub>-C control peptide to DnaK- $\beta$ . The open and closed symbols illustrate the data for two independent experiments. The continuous lines denote the curves resulting from data fitting to eq 3, employed to derive the equilibrium dissociation constant  $K_d$ .

**Identification of ApoMb Binding Sites.** The substrate-binding domain of DnaK is known to bind substrates more weakly than the full-length chaperone. Previous studies have shown that about 25-fold reduction in affinity takes place for the binding of NRLLLTG to a construct similar to DnaK- $\beta$ , relative to the corresponding binding to full-length DnaK in the presence of ATP (12). We assessed the reduction in substrate-binding affinity upon going from full-length DnaK to our DnaK- $\beta$  construct by determining the equilibrium dissociation constant ( $K_d$ ) for the  $\sigma^{32}$ Q<sub>132</sub>-Q<sub>144</sub>-C peptide. This sequence is derived from the  $\sigma^{32}$  transcription factor, a natural chaperone substrate. The  $K_d$  value for the binding of this peptide to full-length DnaK has been determined before (26). We followed a procedure similar to the published one to quantify the affinity of  $\sigma^{32}$ Q<sub>132</sub>-Q<sub>144</sub>-C for DnaK- $\beta$  by fluorescence. The binding isotherm yielded a  $K_d$  of  $840 \pm 340$  nM (Figure 5). This value is 10-fold larger than the corresponding  $K_d$  for binding to full-length DnaK (26). Therefore, the experiments below were carried out with two different DnaK- $\beta$  concentrations, i.e., 150 nM (this value matches the concentration used for full-length DnaK) and 1.5  $\mu$ M.

The cellulose-bound peptide scanning approach was selected to identify the DnaK and DnaK- $\beta$  binding sites along the apoMb sequence (10). The cellulose membrane is divided into 48 squares, each containing a spot covalently linked to a member of a peptide library comprising the entire apoMb sequence. The library is composed of 13-mer peptides with a 10-residue overlap. Peptide scanning enabled the identification of at least six apoMb regions of primary structure displaying affinity for DnaK and DnaK- $\beta$  (Figure 6). The binding site cores were determined as the amino acids common to adjacent peptides displaying a chaperone binding response (Table 1). The data show the presence of a few well-localized specific binding sites, distributed approximately evenly along the apoMb sequence.

Two of the sites show evidence for stronger chaperone binding (Figure 6A). Spot darkness is regarded here as directly proportional to binding affinity for the corresponding peptide sequence, within each data set. As illustrated in Figure 1B, one of these sites maps onto apoMb's native B

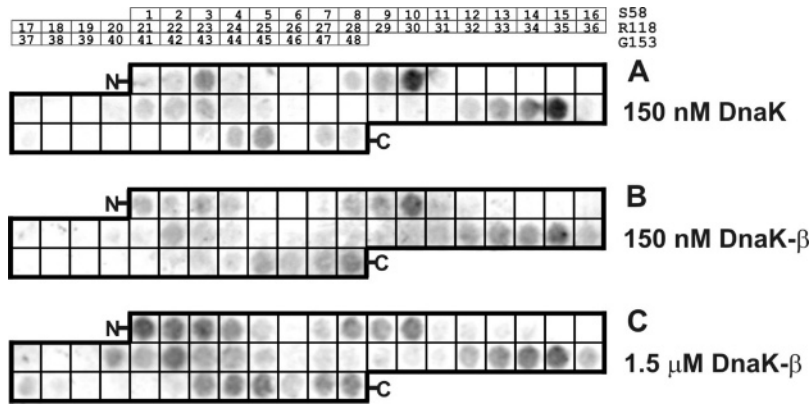


FIGURE 6: DnaK and DnaK- $\beta$  binding to apoMb detected by cellulose-bound peptide scanning. This experiment was performed as described in the Experimental Procedures section. The grid above the figure panels provides a numerical scheme to guide the identification of specific squares. The C-terminal residues of the peptides in the last square of each row are indicated on the right-hand side of the grid. Squares 1–48 comprise the entire apoMb sequence from the N- to C-terminus. Each square illustrates the behavior of a 13-mer peptide with a 10-residue sequence overlap with the peptide in the adjacent square. Data resulting from a cellulose membrane containing covalently bound apoMb peptides incubated with (A) 150 nM DnaK, (B) 150 nM DnaK- $\beta$ , and (C) 1.5  $\mu$ M DnaK- $\beta$  are shown. Six specific binding sites evenly distributed along the apoMb sequence are present for both DnaK and DnaK- $\beta$ .

Table 1: Amino Acid Sequence of Binding Site Cores for ApoMb Binding to the DnaK and DnaK- $\beta$  Chaperones Determined by Cellulose-Bound Peptide Scanning<sup>a</sup>

helix	sequence and binding site
A (S <sub>3</sub> –E <sub>18</sub> )	SEGEWQLVLHVWAKVE
B (D <sub>20</sub> –S <sub>35</sub> )	DVAGHGQDILIRLFKS
E (S <sub>58</sub> –K <sub>77</sub> )	SEDLKKHGVTVLTALGAILK
G (P <sub>100</sub> –R <sub>118</sub> )	PIKYLEFISEAIIHVLHSR
H (G <sub>124</sub> –E <sub>148</sub> )	GADAQGAMNKALELFRKDIAAKYKE

<sup>a</sup> The name of apoMb's native helices involved in chaperone binding and the corresponding residue numbers are listed in the first column. Binding site cores are highlighted in bold type and underlined.

helix (squares 8–10) and the other maps onto the G helix (squares 32–36). Squares 20–25 provide evidence for two or three overlapping weaker binding sites in the E helix. Finally, three additional relatively weak binding sites are located in the A (squares 1–4) and H (squares 43–45 and 47–48) helices.

Peptide scanning in the presence of DnaK- $\beta$  shows binding sites similar to those of the corresponding experiment in the presence of full-length DnaK (Figure 6B). Some of the apoMb peptide scanning squares for the data collected in the presence of 150 nM DnaK- $\beta$  are characterized by weak or no intensity. This suggests that overall binding affinity may be weaker for DnaK- $\beta$  than for the full-length chaperone. This is in agreement with the known weaker affinity of DnaK- $\beta$  for peptide substrates. In order to ensure visualization of all DnaK- $\beta$  binding sites, we repeated the peptide scanning experiment in the presence of a 10-fold higher DnaK- $\beta$  concentration (Figure 6C). Under these conditions, we find that additional squares show evidence for chaperone binding (compare panels A and C of Figure 6), pointing to the presence of some degree of specificity loss. In order to provide a more quantitative comparison among the different cases, the relative intensities for peptide scanning in the presence of DnaK and two different concentrations of DnaK- $\beta$  are plotted in Figure 7. Data were normalized to facilitate comparisons among different conditions. The *residue number* label corresponds to the central amino acid of the 13-residue cellulose-bound peptide. Data in Figure 7 indicate a very good general agreement between the apoMb

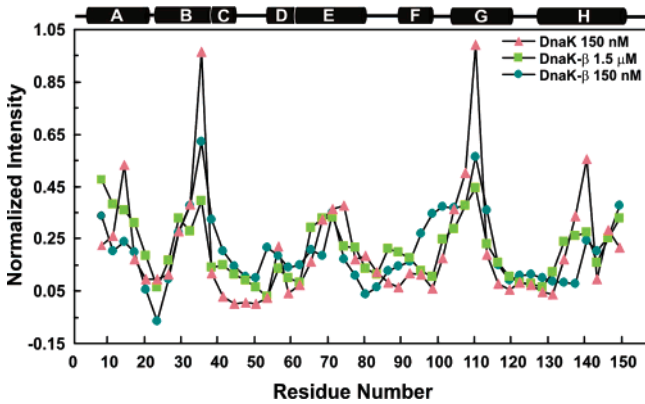


FIGURE 7: Mapping of DnaK chaperone binding affinities for the apoMb sequence. Intensity of peptide scanning signals for the binding of apoMb 13-mer peptides to full-length DnaK (150 nM, orange triangles) and two different concentrations of DnaK- $\beta$  (150 nM, dark green circles; 1.5  $\mu$ M, light green squares). Spot intensities were calculated with the Image-Quant software (version 5.2, Molecular Dynamics, 1999). Binding intensities are expressed in arbitrary units, as the raw intensities were normalized to enable optimal visualization of binding site distribution for the three different conditions. The residue number label denotes the amino acid corresponding to the center of each 13-mer peptide. The native apoMb helices are mapped above the graph.

binding site locations for the two forms of the chaperone. However, an outlier region corresponding to residues 90–100 (see also squares 28–30 of Figure 6) displays higher binding affinity at low DnaK- $\beta$  concentration. Unlike full-length DnaK, which displays two binding sites with much higher affinity than the other sites, a more uniform distribution of binding strength is observed in the case of DnaK- $\beta$ . This result may be due to some loss in specificity upon interaction with DnaK- $\beta$ . Alternatively, this finding may be due to a better penetration of DnaK- $\beta$  through the cellulose-supported hydrated peptide array (each peptide is linked via its C-terminus), due to its smaller size. The latter hypothesis is supported by the fact that the two binding sites displaying exceptionally high binding strength to DnaK (squares 10 and 35 in Figure 6A) have the putative core binding motif located at the N-terminal region of the 13-mer peptide, presumably optimizing accessibility to the smaller DnaK- $\beta$  chaperone. However, such a position dependence was not reported for

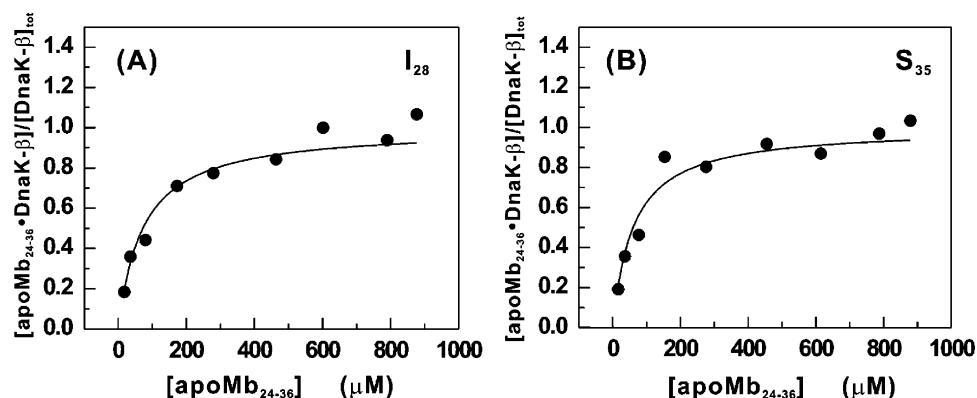


FIGURE 8: Binding isotherms derived from NMR titrations of apoMb<sub>24–36</sub> peptide binding to DnaK-β. Filled circles denote experimental data from <sup>1</sup>H–<sup>15</sup>N HSQC NMR ligand-binding titrations performed in the presence of increasing amounts of apoMb<sub>24–36</sub> peptide. Samples were prepared in 10 mM sodium acetate at pH 6.0, and temperature was kept constant at 30 °C. The (A) I<sub>28</sub> and (B) S<sub>35</sub> resonances, yielding the best resolved bound state signals, were followed independently. The solid line shows curve fitting by nonlinear regression according to eq 5, while the horizontal axis is the concentration of apoMb<sub>24–36</sub>, and the vertical axis is the fraction of complex per chaperone. The best-fit values for the dissociation constants were  $K_d = 72 \pm 11 \mu\text{M}$  for I<sub>28</sub> and  $K_d = 60 \pm 9 \mu\text{M}$  for S<sub>35</sub>.

previous peptide scanning experiments (9, 10, 34), and additional targeted experiments need to be performed to further test the above hypothesis.

In order to more rigorously quantify binding affinities, we carried out NMR titrations with a selectively <sup>15</sup>N-labeled peptide comprising the amino acid sequence of one of the strongest binding sites identified in this work. The isotherms obtained for the binding of apoMb<sub>24–36</sub> to DnaK-β, focusing on two well-resolved resonances for the peptide's chaperone-bound state, are shown in Figure 8. Curve fitting of the experimental data led to an average equilibrium dissociation constant ( $K_d$ ) of  $66 \pm 14 \mu\text{M}$ . Similar titrations in the presence of full-length DnaK were prevented by the large size of the complete Hsp70 protein.

**Chaperone Affinity and Chain Hydrophobicity.** The binding motif for DnaK has been shown to consist of three to five central hydrophobic residues flanked by positively charged amino acids (10). It is not known, however, if the degree of substrate nonpolar character correlates quantitatively with binding affinity and which hydrophobicity scale best describes the nonpolar nature in relation to chaperone binding.

In order to assess the correlation between degree of apoMb nonpolar content and interaction with chaperone, the relative binding affinities assessed by peptide scanning were compared to the hydrophobic/hydrophilic nature of apoMb's polypeptide chain (Figure 9). Different criteria to describe amino acid polarity were employed in this analysis. These include the hydropathy scores by Kyte and Doolittle (35) and two versions of the buried surface area scale by Rose (36). One scale by Rose (Figure 9D) considers the difference between the total surface area of each amino acid type in its reference state and the corresponding average solvent-accessible surface area, evaluated over a set of native protein crystal structures. The reference state corresponds to the solvent-accessible surface area of an amino acid, X, in a Gly-X-Gly tripeptide preserving the distribution of dihedral angles observed in actual proteins (36). Thus, the standard state reflects the degree by which an amino acid is buried solely due to the covalently bonded neighbor atoms, without any surface burial arising from secondary or tertiary interactions. The resulting value, here denoted as *mean buried area*, relates directly to the tendency by each given amino acid

type to be shielded from the solvent. This is a measure of nonpolar character. The second scale by Rose (Figure 9C) is obtained by dividing the mean buried area by the total surface area. The purpose of this normalization is to eliminate any bias due to amino acid size. We define the resulting value as *fractional mean buried area*.

Given that the core of DnaK binding motif is known to comprise five nonpolar residues (10), we assessed the hydrophobicity of the five central residues of all possible 13-mer peptides along the apoMb sequence. This ensures that the hydrophobicity of all binding sites is calculated, whether or not they are located in the middle of the experimentally studied peptides. The residue number label in Figure 9 corresponds to the central amino acid of the 13-residue peptide. The peptides employed in the scanning experiment are indicated by vertical bars. The binding affinity to full-length DnaK is in fairly good agreement with the nonpolar content of the apoMb chain according to the scores of panels B and C of Figure 9. However, the mean buried area plot reported in panel D does not predict significant DnaK binding to apoMb's residues corresponding to the native E helix. On the other hand, experimental evidence (Figure 9A) is consistent with the presence of a low-affinity DnaK binding site in this helix. In addition, the scores of panel D suggest that, based on the nonpolar nature of the chain alone, there should be a binding site for the residues corresponding to the CD loop. This site is not observed experimentally. Nevertheless, the Kyte–Doolittle and fractional mean buried area scores correlate best with the location of the experimentally determined binding sites.

A predictive algorithm, based on cellulose-bound peptide scanning experiments on multiple proteins (10), was used to estimate the DnaK binding locations in the apoMb sequence (Figure 9E). Negative scores denote higher probability of binding DnaK. The algorithm is successful at qualitatively predicting nearly all of the binding sites in the apoMb sequence. This algorithm, however, only predicts 50–80% likelihood of binding for the site corresponding to the G helix, as it is an atypical binding site for DnaK with a negatively charged Glu residue. The binding prediction algorithm suggests that the apoMb<sub>111–119</sub> peptide has good affinity for DnaK (Figure 9E). However, peptide scanning experiments show no evidence for binding (Figures 6 and



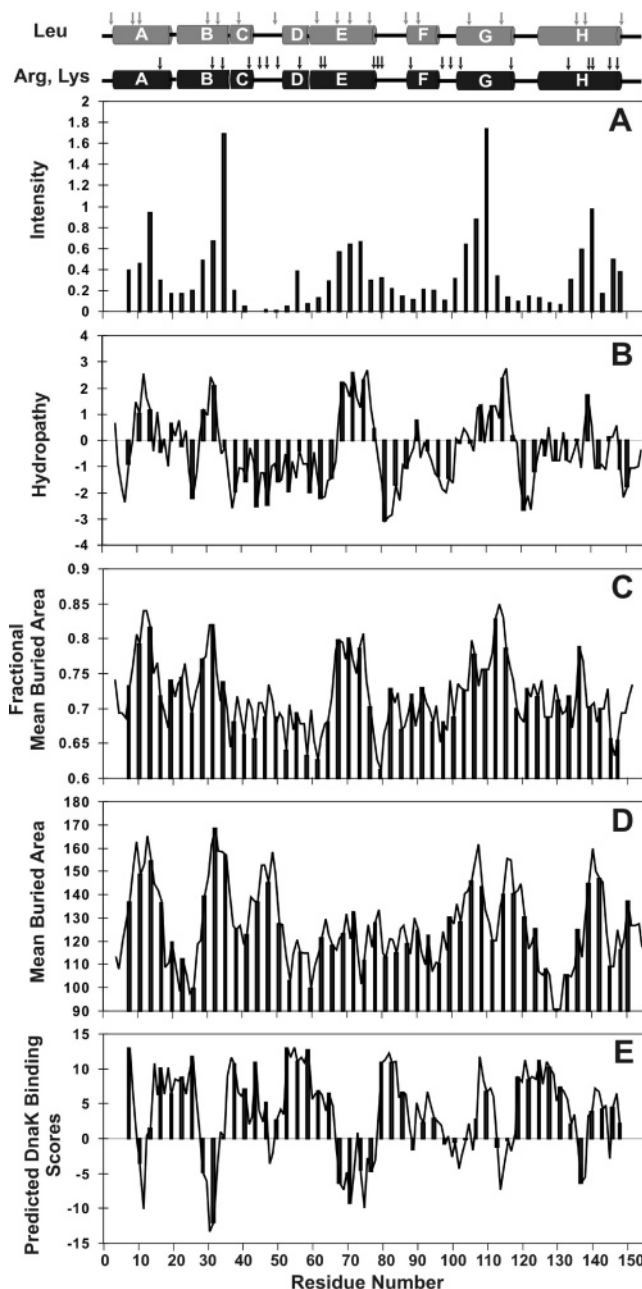


FIGURE 9: Comparison of experimental DnaK binding affinities with computed polypeptide chain properties. (A) Graphical representation of cellulose-bound peptide scanning intensities, reflecting the site-specific interaction between full-length DnaK and 13-mer apoMb-derived peptides scanning the whole protein sequence. Each model peptide differs from the previous one by a three amino acid shift along the apoMb sequence (from the N- to C-terminus). The location of native apoMb helices along the sequence is mapped above the figure. ApoMb's Leu and positively charged residues (i.e., Arg and Lys) are mapped as gray and black arrows, respectively, above the helix-map diagrams. The nonpolar character of the model peptide sequences is illustrated by three different calculated parameters in panels B–D. A window size of five residues was used in the calculations. Hence, the values reflect the nonpolar character of the five central residues of the 13-mer peptides. (B) Hydropathy scores for the apoMb peptide sequences according to Kyte and Doolittle (35). (C) Fraction of mean buried surface area upon folding according to Rose (36), determined from the normalized difference between the calculated solvent-accessible surface area and average solvent-accessible surface area for each residue according to the folded protein structures in the database. Normalization has been achieved by dividing the above difference by the calculated solvent-accessible surface area values. (D) Unnormalized mean buried surface area illustrating the difference between calculated and average solvent-accessible surface area (36). (E) Probability scores for the binding of DnaK to the apoMb protein sequence using the known DnaK prediction algorithm (10). Negative scores denote higher binding probabilities. The residue number label denotes the central amino acid of the peptide corresponding to each bar.

9A). The titrations with the apoMb<sub>111–119</sub> peptide did not yield any detectable resonance for the bound form (data not shown), suggesting either that this peptide is in fast exchange with its chaperone-bound form (on the NMR chemical shift time scale) or that the peptide does not exhibit any significant affinity for the chaperone.

The positively charged residues are known to contribute to substrate binding through favorable electrostatic interactions with the chaperone. The vertical arrows on top of Figure 9 correspond to the positively charged residues in the apoMb sequence. There are a total of 35 such residues distributed throughout the sequence. However, the high number and



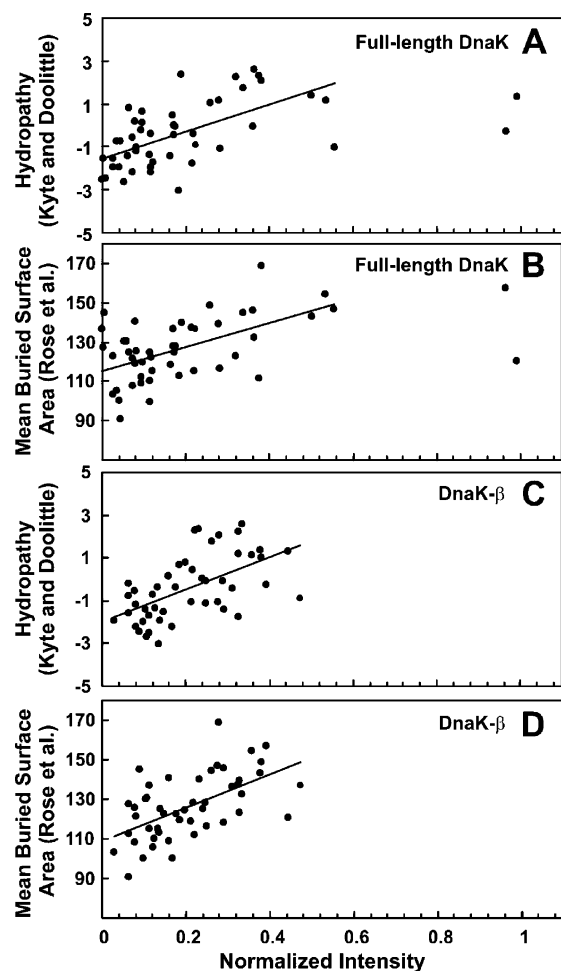


FIGURE 10: Correlation between experimental chaperone binding affinities and degree of nonpolar content of substrate. DnaK and DnaK- $\beta$  peptide binding intensities (assessed by cellulose-bound peptide scanning) are plotted against hydrophobicity of the peptide sequence. The upper panels illustrate the relation between DnaK (150 nM) peptide scanning intensities and (A) the hydropathy scale of Kyte and Doolittle (35) and (B) mean buried surface area scores by Rose et al. (36). The lower panels describe the correlation between DnaK- $\beta$  (1.5  $\mu$ M) peptide scanning intensities and (C) the hydropathy scale of Kyte and Doolittle (35) and (D) mean buried surface area scores by Rose et al. (36).

wide dispersion of the positively charged residues along the sequence defer establishing any apparent correlation between the location of these residues and experimentally observed binding sites.

In order to more directly assess the correlation between nonpolar character of the substrate and affinity for the DnaK chaperone, we plotted the nonpolar content of the different apoMb residues, according to the scales employed in Figure 9, against experimentally observed binding affinities (Figure 10). A general agreement between the two parameters was observed, in the case of both DnaK and DnaK- $\beta$ . However, linear regression curve fitting points to the presence of a relatively weak correlation. This result shows that the degree of nonpolar content of a polypeptide chain is only a qualitative indicator of binding affinity for the DnaK chaperone. The data for the two strongest experimentally observed binding sites for full-length DnaK behave as clear outliers (panels A and B of Figure 10) and were not used for curve fitting. This supports the idea that factors directly related to the nature of the peptide scanning experiment,

including chaperone accessibility to different regions of the membrane-bound 13-mer peptides, may, in some instances, play a nonnegligible role.

## DISCUSSION

The substrate-binding domain of bacterial Hsp70, DnaK- $\beta$ , has previously been used as a model system to investigate chaperone-peptide interactions. However, a detailed characterization of this domain has not been available so far. The DnaK- $\beta$  biophysical characterization provided here enables the quantitative assessment of the response to pH variations and thermodynamic stability under physiologically relevant conditions. Equilibrium unfolding urea titrations (Figure 4) yielded an apparent standard free energy change  $\Delta G^\circ = 3.09 \pm 0.89$  kcal/mol for unfolding and a corresponding  $m$  value of  $1.71 \pm 0.42$  kcal mol $^{-1}$  M $^{-1}$ . Our experimental  $m$  value is consistent with that of typical single domain globular proteins (37). These thermodynamic data on *E. coli*'s DnaK- $\beta$  are in qualitative agreement with the results by Fuertes et al. (38) on the substrate-binding domain of the human inducible Hsp70 ( $\Delta G^\circ$  of  $3.1 \pm 0.1$  kcal/mol upon GdmHCl denaturation at pH 6.5). As expected for a multidomain protein, the full-length DnaK chaperone differs from DnaK- $\beta$  in that it unfolds via an equilibrium intermediate (39).

We also investigated how the lack of the ATPase domain and the helical C-terminal subdomain affect substrate-binding specificity. The peptide scanning experiments on DnaK- $\beta$  show that this chaperone domain binds specifically to the same sites along the apoMb sequence as the full-length chaperone. This result unambiguously proves that the specificity of DnaK for its substrates is fully captured by this minimal binding domain model, with little or no contributions from the ATPase domain or the C-terminal helical "lid" subdomain. It has been reported that the helical subdomain is not required for substrate binding (12, 40). Our results corroborate and expand this observation by showing that DnaK's substrate-binding domain lacking the helical lid is fully necessary and sufficient to maintain binding specificity.

The affinity of DnaK- $\beta$  for peptide substrates is significantly lower relative to the affinity of the full-length chaperone. This minimal chaperone construct lacks the ATPase domain and, hence, the regulatory role of the nucleotide state on substrate affinity (41, 42). Moreover, constructs containing the ATPase domain but lacking the helical subdomain have been reported to have a decreased affinity for their substrates (9, 43). The minimal binding domain examined here lacks both of the above structural units. Hence, the lack of these units decreases binding affinity without significantly affecting specificity.

One alternative explanation for the apparent weaker signal obtained with DnaK- $\beta$  in the peptide scanning experiment is the possibility of a weaker response by the antibody employed for signal detection. However, this is somewhat unlikely, given that we employed monoclonal anti-DnaK antibodies. The extent of antibody binding to the substrate-binding domain should be insensitive to the presence of an additional domain, under the blotting conditions used. Most importantly, the fluorescence-based binding experiments with IAANS-labeled  $\sigma^{32}$ Q $_{132}$ -Q $_{144}$ -C (Figure 5) show that this

peptide has a lower affinity for DnaK- $\beta$  than for the full-length chaperone. This indicates that a decrease in binding affinity upon switching from full-length DnaK to DnaK- $\beta$  is present regardless of the antibody response.

The high-resolution binding affinity profiles for DnaK and DnaK- $\beta$  are compared in Figure 7. The plot provides a comparison of the relative binding affinities at each site. While the specificity is the same in the sense that both versions of the chaperone respond to the same binding sites, the overall distribution of the specificity at each site is broader in the case of DnaK- $\beta$ . Therefore, the presence of the helical lid and/or the regulatory ATPase domain may be involved in secondary effects such as the fine-tuning of the local sharpness in site specificity.

Even though no fluorescence signal could be observed for the binding of  $\sigma^{32}$ Q<sub>132</sub>–Q<sub>144</sub>–C-IAANS to the DnaK(1–538) lid-truncated construct (9), we did obtain a fluorescence signal for the binding of this peptide to DnaK- $\beta$ . DnaK(1–538) binds substrates similarly to wild-type chaperone with slightly lower affinity (9, 34). Even though the truncated ~100 DnaK C-terminal amino acids do not contact the substrate nor affect specificity, the fluorescence signal is lost when they are absent in the construct. As mentioned above, we do observe a fluorescence signal upon binding of the peptide to DnaK- $\beta$ . It is possible that the  $\beta$ -sandwich becomes more flexible in the absence of the ATPase domain and the helical elements of the substrate-binding domain. This might result into accommodation of more than one bound substrate conformation. This hypothesis is consistent with the slight loss of specificity observed for DnaK- $\beta$  in the peptide scanning experiments.

We examined a number of physical properties of the substrate and their relation to the ability to support binding. The first property is hydrophobicity, which is an essential substrate feature for DnaK binding (8, 10). The nonpolar character of each amino acid was ranked according to different scales to explore whether the results depend on the choice of the scale and how the nonpolar amino acid nature correlates with chaperone binding affinity (Figure 9). The binding motif for DnaK is expected to have a three to five residue hydrophobic core (10). Therefore, we assessed the nonpolar character of the five central residues of each peptide. The Kyte–Doolittle (Figure 9B) and the fractional mean buried area scale by Rose (Figure 9C) yield very similar predictions and are in fairly good agreement with affinities obtained from peptide scanning intensities. On the other hand, the mean buried area plot of Figure 9D shows a somewhat different profile, and it predicts that the residues corresponding to the E helix are much less nonpolar. However, there is a DnaK binding site for the residues encoding this helix. Thus the mean buried area scores for this region do not correlate effectively with affinity for DnaK.

Fractional mean buried area correlates with experimentally observed binding affinities better than mean buried area. It follows that fractional nonpolar surface area serves as an effective descriptor of chaperone binding to this  $\alpha$ -helical model protein. Interestingly, the fundamental difference between the above two scales by Rose is that the mean buried area scale does not factor out the size of the nonpolar surface. Therefore, the residues with nonbulky side chains, i.e., Ala and Gly, in the region corresponding to the native E helix contribute much less according to this scale, due to their

relatively small side chain size. These data suggest that the size of the side chains is not important in DnaK binding, as long as the residues directly involved in binding are nonpolar. However, it has been demonstrated that Hsp70 chaperones do not recognize their substrates based solely on hydrophobicity criteria. Contacts with specific hydrophobic side chains are required (8, 10, 44, 45). For instance, Leu contributes significantly more to DnaK binding relative to Ile or Val, although it is not significantly more hydrophobic. Similarly, when the three Leu residues are replaced by Ala in the peptide NRLLLTG, no binding to the chaperone can be detected (46). The specific binding sites along the apoMb sequence identified here are entirely consistent with this criterion in that they all contain Leu (except for the weak binding site at helix E). Hence, one long hydrophobic side chain that fits to the binding site pocket of the chaperone is needed for strong binding, while the nearby (long or short) hydrophobic residues and the flanking positive charges contribute to increasing binding affinity. The direct correspondence between nonpolar nature of the chain and chaperone binding affinity observed here may be a consequence of the fact that the other important known determinants for chaperone binding, i.e., Leu residues and positive charges, are abundant and approximately evenly spread across apoMb's amino acid sequence (Figure 9).

The correlation between different scores and binding to DnaK and DnaK- $\beta$  is best visualized by the plots in Figure 10. Although the hydropathy scores are a good predictor of binding site location, they cannot be employed as a quantitative prediction tool. The mathematical correlation between hydropathy scores and experimental chaperone binding intensities is rather weak. In the case of binding to DnaK, the Kyte–Doolittle hydropathy scores exhibit a slightly better correlation, relative to the mean buried surface values. The degree of correlation between the two parameters is comparable in the case of DnaK- $\beta$ . We also repeated the hydropathy calculations for all of the possible 5-mer binding cores in each 13-mer peptide. We considered the maximum of these hydropathy scores as the score of that peptide. Evaluating the hydropathy of all of the 5-mers did not result in any significant change in the weak correlations between hydropathy scores and experimental peptide scanning intensities of Figure 10. The NMR and X-ray crystal structures of a substrate peptide bound to DnaK's binding domain show no interactions beyond the five central nonpolar residues of the peptide (8, 40). Nevertheless, systematic analysis of sequences that bind DnaK has revealed a binding motif with a hydrophobic core flanked by positively charged residues (10). In order to test the importance of positively charged residues, we examined their location along the apoMb sequence (black vertical arrows on top of Figure 9). There are a large number of positively charged residues distributed along the sequence of this high-pI protein. This may be related to the relatively good agreement of hydropathy scores with DnaK binding affinity for this particular protein, without the need to consider the positive charges. The abundance of positively charged residues makes it difficult to establish an apparent correlation between the positively charged residues and the experimentally determined binding site locations. However, the crystal structure of DnaK binding domain shows a largely negative electrostatic potential at the surface of the binding channel, flanking the central hydrophobic

region (8). In accordance with this, it has been reported that the DnaK binding affinity for the NRLLLTG depends on salt concentration and that changing the positively charged Arg to Ala reduces the binding capacity (46). Thus, the positively charged residues are expected to contribute to binding, even though they are not as important as the central hydrophobic contacts (10). In the cell, the affinity of DnaK for positively charged residues may also help in diverting the nascent chain from favorable interactions with the largely negatively charged ribosome surface.

Finally, we compared the experimental results to the predictions of a recently developed DnaK binding algorithm (10) based on the systematic peptide scanning analysis of several proteins (Figure 9E). The algorithm successfully predicts the location of most binding sites. Upon comparing the algorithm scores with experimental binding affinities, an agreement similar to that between experiment and predicted substrate nonpolar content was observed (Figure 9, panels B and C). As mentioned above, we propose that this is due to the abundance of positively charged residues in this high-*pI* protein.

The location of apoMb's binding sites for DnaK are mapped on the ribbon representation of the native protein in Figure 1B. It has been reported that only a minority of DnaK binding sites are in sequences populating  $\alpha$ -helical conformation in the native protein, and no binding sites are found in relatively long and stable helical sequences at dimerization interfaces (10). All of the apoMb binding sites detected here correspond to native helices for this protein. Furthermore, none of the sites are in the loops connecting the helices. The X-ray and NMR structure of the binding domain in complex with a substrate peptide shows that the peptide binds in an extended conformation (8, 40). This finding and additional recent evidence (27) indicate that it is unlikely that the binding cleft is able to accommodate the apoMb binding sites in their native helical conformation. The two strongest binding sites are close in space in the native protein structure, and they both belong to apoMb's strong nonpolar core. They are not accessible for the chaperone to bind when the protein is in its native state. Hence, we conclude that apoMb needs to be in a nonnative conformation for DnaK binding to occur. The residues corresponding apoMb's binding sites for DnaK may be exposed during protein biosynthesis, upon apoMb expression in *E. coli* (routinely performed during the production of this protein for biophysical studies), when the chain is incomplete and cannot properly fold. In agreement with the above, purified N-terminal fragments of apoMb interact with DnaK- $\beta$ , while the full-length protein does not, at equilibrium (11). It is interesting to note that these N-terminal fragments self-associate and adopt non-native  $\beta$ -sheet structure, when taken out of the context of the native protein (19). For a sequence with  $\alpha$ -helical conformation in the native protein, chaperone binding is still needed if the local helical propensity is low, and tertiary contacts are needed to stabilize the helix. In such a case, chaperone binding prevents the misfolding routes accessible to the chain before its elongation is complete. The helical structure becomes stabilized by intramolecular tertiary contacts once the polypeptide chain reaches its full length and the native protein is generated in solution.

In conclusion, the results presented here provide a detailed characterization of the binding specificity of the DnaK

chaperone and its minimal substrate-binding domain for the sequence of the all- $\alpha$ -helical protein apoMb. This work supports the use of the DnaK- $\beta$  substrate-binding domain as a chaperone model retaining a very similar substrate specificity to that of the full-length chaperone. The data are consistent with the ATPase domain and the helical lid having a regulatory effect on binding affinity, without significantly affecting substrate specificity.

## ACKNOWLEDGMENT

We thank Elizabeth A. Craig and Willy Walter for sharing laboratory equipment.

## REFERENCES

1. Bukau, B., Deuerling, E., Pfund, C., and Craig, E. A. (2000) Getting newly synthesized proteins into shape, *Cell*, 119–122.
2. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: From nascent chain to folded protein, *Science* 295, 1852–1858.
3. Gething, M.-J. (1999) Role and regulation of the ER chaperone BiP, *Semin. Cell Dev. Biol.* 10, 465–472.
4. Erbse, A., Mayer, M., and Bukau, B. (2004) Mechanism of substrate recognition by Hsp70 chaperones, *Biochem. Soc. Trans.* 32, 617–621.
5. Deuerling, E., and Bukau, B. (2004) Chaperone-assisted folding of newly synthesized proteins in the cytosol, *Crit. Rev. Biochem. Mol. Biol.* 39, 261–277.
6. Jiang, J. W., Prasad, K., Lafer, E. M., and Sousa, R. (2005) Structural basis of interdomain communication in the Hsc70 chaperone, *Mol. Cell* 20, 513–524.
7. Revington, M., Zhang, Y. B., Yip, G. N. B., Kurochkin, A. V., and Zuiderweg, E. R. P. (2005) NMR investigations of allosteric processes in a two-domain *Thermus thermophilus* Hsp70 molecular chaperone, *J. Mol. Biol.* 349, 163–183.
8. Zhu, X. T., Zhao, X., Burkholder, W., Gragerov, A., Otaga, C., Gottesman, M., and Hendrickson, W. (1996) Structural analysis of substrate binding by the molecular chaperone DnaK, *Science* 272, 1606–1614.
9. Mayer, M. P., Schroder, H., Rüdiger, S., Paal, K., Laufen, T., and Bukau, B. (2000) Multistep mechanism of substrate binding determines chaperone activity of Hsp70, *Nat. Struct. Biol.* 7, 586–593.
10. Rüdiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries, *EMBO J.* 16, 1501–1507.
11. Kurt, N., Rajagopalan, S., and Cavagnero, S. (2006) Effect of Hsp70 chaperone on the folding and misfolding of polypeptides modeling and elongating protein chain, *J. Mol. Biol.* 355, 809–820.
12. Pellicchia, M., Montgomery, D., Stevens, S., Vander Kooi, C., Feng, H., Gierasch, L., and Zuiderweg, E. (2000) Structural insights into substrate binding by the molecular chaperone DnaK, *Nat. Struct. Biol.* 7, 298–303.
13. Vinogradov, S., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Guertin, M., Gough, J., Dewilde, S., Moens, L., and Vanfleteren, J. (2005) Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life, *Proc. Natl. Acad. Sci. U.S.A.* 102, 11385–11389.
14. Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., and Kutyschenko, V. P. (1988) Thermodynamic study of the apomyoglobin structure, *J. Mol. Biol.* 202, 127–138.
15. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate, *Science* 249, 1544–1548.
16. Jennings, P. A., and Wright, P. E. (1993) Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin, *Science* 262, 892–896.
17. Cavagnero, S., Dyson, H. J., and Wright, P. E. (1999) Effect of H-helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin, *J. Mol. Biol.* 285, 269–282.
18. Garcia, C., Nishimura, C., Cavagnero, S., Dyson, H. J., and Wright, P. E. (2000) Changes in the apomyoglobin folding pathway caused



- by mutation of the distal histidine residue, *Biochemistry* 39, 11227–11237.
19. Chow, C. C., Chow, C., Raghunathan, V., Huppert, T. J., Kimball, E. B., and Cavagnero, S. (2003) The chain length dependence of apomyoglobin folding: Structural evolution from misfolded sheets to native helices, *Biochemistry* 42, 7090–7099.
  20. Bakke, C., Jungbauer, L., and Cavagnero, S. (2006) In vitro expression and characterization of native apomyoglobin under low molecular crowding conditions, *Protein Expression Purif.* 45, 381–392.
  21. Jennings, P., Stone, M., and Wright, P. E. (1995) Overexpression of myoglobin and assignment of its amide C $\alpha$  and C $\beta$  resonances, *J. Biomol. NMR* 6, 271–276.
  22. Bradford, M. M. (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding, *Anal. Biochem.* 72, 248–254.
  23. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol.* 131, 266–280.
  24. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. Unfolding of phenylmethanesulfonyl-chymotrypsin using different denaturants, *Biochemistry* 27, 8063–8068.
  25. Schellman, J. (1978) Solvent denaturation, *Biopolymers* 17, 1305–1322.
  26. McCarty, J. S., Rüdiger, S., Schönfeld, H.-J., Schneider-Mergener, J., Nakahigashi, K., Yura, T., and Bukau, B. (1996) Regulatory region of the *E. coli* heat shock transcription factor,  $\sigma^{32}$ , constitutes a DnaK binding site and is conserved among eubacteria, *J. Mol. Biol.* 256, 829–837.
  27. Chen, Z., Kurt, N., Rajagopalan, S., and Cavagnero, S. (2006) Secondary structure mapping of DnaK-bound protein fragments: Chain helicity and local helix unwinding at the binding site, *Biochemistry* 45, 12325–12333.
  28. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc.* 114, 10663–10665.
  29. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe—a multidimensional spectral processing system based on Unix pipes, *J. Biomol. NMR* 6, 277–293.
  30. Johnson, B. A., and Blevins, R. A. (1994) NMRView—a computer-program for the visualization and analysis of NMR data, *J. Biomol. NMR* 4, 603–614.
  31. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995)  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shift referencing in biomolecular NMR, *J. Biomol. NMR* 6, 135–140.
  32. Schindler, T., Herler, M., Marahiel, M., and Schmid, F. (1995) Extremely rapid protein folding in the absence of intermediates, *Nat. Struct. Biol.* 2, 663–673.
  33. Baldwin, R. L. (1996) On-pathway versus off-pathway folding intermediates, *Folding Des.* 1, R1–R8.
  34. Rüdiger, S., Mayer, M., Schneider-Mergener, J., and Bukau, B. (2000) Modulation of substrate specificity of the DnaK chaperone by alteration of a hydrophobic arch, *J. Mol. Biol.* 304, 245–251.
  35. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157, 105–132.
  36. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., and Zehfus, M. H. (1985) Hydrophobicity of amino acid residues in globular proteins, *Science* 29, 834–838.
  37. Scholtz, J., Barrick, D., York, E., Stewart, J., and Baldwin, R. L. (1995) Urea unfolding of peptide helices as a model for interpreting protein unfolding, *Proc. Natl. Acad. Sci. U.S.A.* 92, 185–189.
  38. Fuertes, M. A., Pérez, J. M., Soto, M., Menéndez, M., and Alonso, C. (2004) Thermodynamic stability of the C-terminal domain of the human inducible heat shock protein 70, *Biochim. Biophys. Acta* 1699, 45–56.
  39. Palleros, D. R., Shi, L., Reid, K. L., and Fink, A. L. (1993) Three-state denaturation of DnaK induced by guanidine hydrochloride. Evidence for an expandable intermediate, *Biochemistry* 32, 4314–4321.
  40. Stevens, S., Cai, S., Pellicchia, M., and Zwietering, E. (2003) The solution structure of the bacterial Hsp70 chaperone protein domain DnaK(393–507) in complex with the peptide NRLLLTG, *Protein Sci.* 12, 2588–2596.
  41. McCarty, J., Buchberger, A., Reinstein, J., and Bukau, B. (1995) The role of ATP in the functional cycle of the DnaK chaperone system, *J. Mol. Biol.* 249, 126–137.
  42. Mayer, M., and Bukau, B. (2005) Hsp70 chaperones: Cellular functions and molecular mechanism, *Cell. Mol. Life Sci.* 62, 670–684.
  43. Buczyński, G., Slepnev, S., Sehorn, M., and Witt, S. (2001) Characterization of a lidless form of the molecular chaperone DnaK, *J. Biol. Chem.* 276, 27231–27236.
  44. Gragerov, A., Li, Z., Xun, Z., Burkholder, W., and Gottesman, M. (1994) Specificity of DnaK peptide binding, *J. Mol. Biol.* 235, 848–854.
  45. Rüdiger, S., Buchberger, A., and Bukau, B. (1997) Interaction of Hsp70 chaperones with substrates, *Nat. Struct. Biol.* 4, 342–349.
  46. Liu, W., Bratko, D., Prausnitz, J., and Blanch, H. (2003) Electrostatic interactions between peptides and the molecular chaperone DnaK, *J. Phys. Chem. B* 107, 11563–11569.
  47. Kuriyan, J., Wilz, S., Karplus, M., and Petsko, G. A. (1986) X-ray structure and refinement of carbon-monoxide (FeII)-myoglobin at 1.5 Å resolution, *J. Mol. Biol.* 192, 133–154.
  48. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis* 18, 2714–2723.

BI061432A