# Structural insights into substrate binding by the molecular chaperone DnaK

Maurizio Pellecchia[1,2], Diana L. Montgomery[3,4], Shawn Y. Stevens[1], Craig W. Vander Kooi[5], Hwa-ping Feng[3], Lila M. Gierasch[3] and Erik R.P. Zuiderweg[1,5,6]

*[1]Biophysics Research Division, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, USA. [2]Present address: TRIAD Biotechnology, 5820 Nancy Ridge Road, San Diego, California 92121, USA. [3]Departments of Chemistry and Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts 01003, USA. [4]Present address: Drug Metabolism and Pharmacokinetics, Schering Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, New Jersey 07033, USA. [5]Department of Chemistry, and [6]Department of Biological Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, USA.*

**How substrate affinity is modulated by nucleotide binding remains a fundamental, unanswered question in the study of 70 kDa heat shock protein (Hsp70) molecular chaperones. We find here that the *Escherichia coli* Hsp70, DnaK, lacking the entire α-helical domain, DnaK(1–507), retains the ability to support λ phage replication *in vivo* and to pass information from the nucleotide binding domain to the substrate binding domain, and *vice versa*, *in vitro*. We determined the NMR solution structure of the corresponding substrate binding domain, DnaK(393–507), without substrate, and assessed the impact of substrate binding. Without bound substrate, loop L3,4 and strand β3 are in significantly different conformations than observed in previous structures of the bound DnaK substrate binding domain, leading to occlusion of the substrate binding site. Upon substrate binding, the β-domain shifts towards the structure seen in earlier X-ray and NMR structures. Taken together, our results suggest that conformational changes in the β-domain itself contribute to the mechanism by which nucleotide binding modulates substrate binding affinity.**

The 70 kDa heat shock protein (Hsp70) family of molecular chaperones participates in a variety of biological processes including protein folding, disruption and rescue of protein aggregates, modulation of the heat shock response and protein translocation across membranes (for a recent review see ref. 1). The Hsp70 proteins consist of a 44 kDa nucleotide binding domain (residues 1–392) and a 26 kDa substrate binding domain (also known as the substrate binding unit). The latter was found to consist of a domain of β-sheet topology that harbors the substrate binding cleft (residues 393–507, here referred to as the β-domain) and a 14 kDa domain from residues 508 to 638, of which the first 100 residues are α-helical and appear to act as a lid covering the substrate binding cleft[2-6].

Allosteric modulation of substrate affinity, in which ATP binding to the nucleotide binding domain stimulates release of substrate from the substrate binding domain, is central to the function of the Hsp70 chaperones[7], but the mechanism of this process on a molecular scale remains unknown. An X-ray study of the substrate binding domain of DnaK (residues 389–607) suggested that the α-helical domain may be involved in the allosteric control of substrate binding[3]. However, a recent study of the Hsp70 protein BiP[8] showed that a large part of this helical domain could be removed, resulting in a partially truncated construct (residues 1–542 in DnaK-numbering) that maintains the allosteric functions of the wild type protein.

Here we report the striking result that removal of the entire domain from position 508 to the C-terminus of Dnak, results in a protein that retains the ATP-induced peptide substrate release and the peptide stimulation of ATP hydrolysis activities, and also supports the replication of bacteriophage λ *in vivo*. Intrigued by this observation, we have determined by NMR the solution structure of the corresponding β-domain, DnaK(393–507). The absence of the helical domain in this construct enabled us to observe the structure of the substrate binding domain for the first time without a bound substrate or an intramolecularly bound polypeptide. Surprisingly, loops L1,2 and L4,5 in this apo-β-domain have the same conformation as that in the longer construct (389–607), despite the fact that they are in intimate contact with the N-terminal region of the α-helical domain in this longer construct[3]. Other regions of the domain undergo substantial rearrangements with respect to the peptide-bound forms, which may arise from either the loss of the helical domain or the absence of a substrate. Our

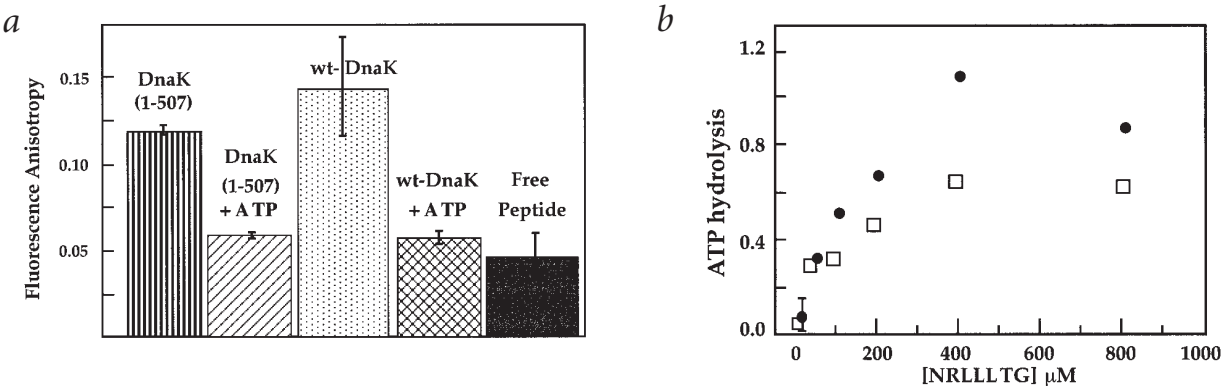


**Fig. 1.** *In vitro* studies of DnaK(1–507) allosteric function. ***a***, ATP-induced release of peptide F-APPY in DnaK(1–507) measured by fluorescence anisotropy. The first bar represents the anisotropy value for peptide bound to 1.1 μM DnaK(1–507). The second bar represents the anisotropy value 5 min after addition of 0.44 mM ATP. The third and fourth bars represent the values for wtDnaK under comparable conditions, and the last bar indicates the anisotropy value of free peptide. Error bars reflect the standard deviation from a mean of three measurements. ***b***, Peptide stimulation of ATPase activity of DnaK(1–507) (●) and wtDnaK (□). As DnaK(1–507) is titrated with the peptide NRLLLTG, the ATPase activity is stimulated in a manner similar to that of wtDnaK. The hydrolysis rate is reported as moles of ATP hydrolyzed per minute per mole of DnaK(1–507) or wtDnaK. The error bar on the first point reflects the standard deviation from a mean of three measurements and is valid for both assays.

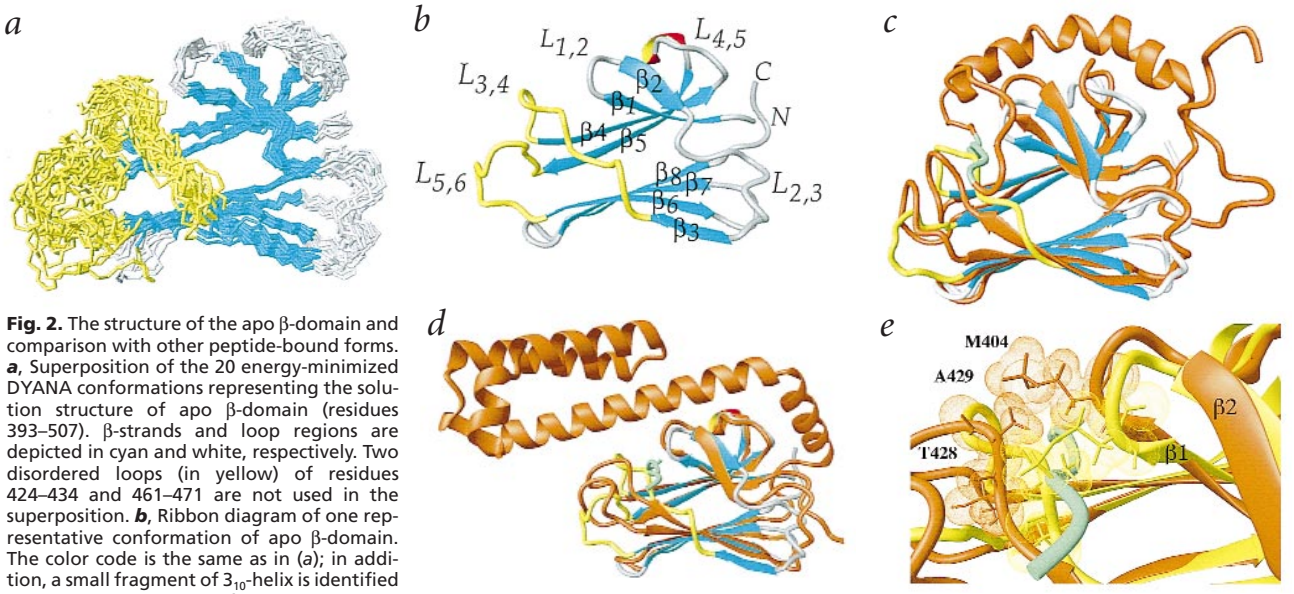


**Fig. 2.** The structure of the apo β-domain and comparison with other peptide-bound forms. ***a***, Superposition of the 20 energy-minimized DYANA conformations representing the solution structure of apo β-domain (residues 393–507). β-strands and loop regions are depicted in cyan and white, respectively. Two disordered loops (in yellow) of residues 424–434 and 461–471 are not used in the superposition. ***b***, Ribbon diagram of one representative conformation of apo β-domain. The color code is the same as in (*a*); in addition, a small fragment of $3_{10}$-helix is identified in red. ***c***, Superposition of the apo β-domain (a representative conformation) and the NMR structure of the DnaK substrate binding domain[4] containing the β-domain and part of the α-helical domain (DnaK(386–561); in orange; PDB code 1BPR). The substrate binding cleft in the latter structure is occupied by residues Asp 540-His 541-Leu 542-Leu 543-His544-Ser 545-Thr 546-Arg 547 (in green). ***d***, Superposition of the apo β-domain with the X-ray structure of the DnaK substrate binding domain containing the β-domain and the α-helical domain (in orange; DnaK(389–607); PDB code 1DKZ) in complex with the peptide NRLLLTG (green)[3]. The heavy atoms of the strands of apo β-domain were used for the superpositions. The r.m.s. deviations for the superposition are reported in Table 1. ***e***, A close up view of the substrate binding cleft of the superposition shown in (*d*). The X-ray structure is in orange, the bound peptide in green and the apo β-domain in yellow. The side chains of Met 404, Thr 428 and Ala 429 are drawn for both structures in the corresponding colors.

finding that substrate binding drives the conformation toward that observed for the larger construct supports the latter interpretation. The substrate-induced changes affect both the conformation and the dynamics of the β-domain and extend from the substrate-binding cleft to the remote loop L2,3.

## DnaK (1–507) can function in vivo and in vitro

Several assays were carried out to determine which properties, if any, of DnaK would be affected by truncation at position 507. First, differential scanning calorimetric experiments on DnaK(1–507) show three cooperative transitions at 41.3, 48.1, and 71.5 °C. These values are comparable to wild type DnaK (wtDnaK) under the same conditions[9], demonstrating that the overall protein stability is not compromised. Second, since DnaK is required to support the propagation of bacteriophage λ[10], we have used this *in vivo* activity as a screen to detect functional DnaK variants carrying mutations[9]. When the truncated DnaK(1–507) is expressed in a DnaK-deficient strain[11], the number of bacteriophage λ plaques is ~1/4 the number observed with wtDnaK. This surprising result indicates that the severely truncated DnaK(1–507) retains the capacity to function *in vivo*, albeit with lower activity than wt Dna K (see Methods).

Third, from the many available *in vitro* assays for DnaK, we characterized peptide binding, ATP modulation of peptide affinity and peptide-stimulation of ATPase activity. Using a fluorescence assay we find that DnaK(1–507) binds either of the labeled peptides, fluoroscein-labeled CALLQSRLLLSAPRRAAATARY (F-APPY)[9] or acrylodan-labeled NRLLLTG, with about five-fold lower affinity than does wtDnaK[9,12]. In the presence of ATP, we observe that the affinity of DnaK(1–507) for F-APPY (monitored by fluorescence anisotropy as described previously[9]) decreases ~100-fold (from a $K_d$ value of 0.33 ± 0.18 μM to 36.6 ± 9.5 μM). This difference in binding affinities in the absence or presence of saturating ATP is comparable to what has been reported for wtDnaK binding to acrylodan-labeled NRLLLTG[12]. Fig. 1*a* illustrates the ATP-induced release of F-APPY from DnaK(1–507). The effect of ATP on peptide binding is considered a hallmark of Hsp70 chaperone allosteric function, and is here found to be approximately equal for DnaK(1-507) and *wt*DnaK. In reciprocal experiments, addition of NRLLLTG stimulates the ATPase activity of DnaK(1–507) (Fig. 1*b*). The observed enhancement is comparable to the stimulation of wtDnaK ATPase by NRLLLTG, confirming that interdomain communication in DnaK(1–507) is functional in both directions. Taken together, the results of these assays provide strong evidence that DnaK(1–507) retains allosteric function despite removal of the entire α-helical domain.

## The NMR solution structure of the DnaK apo β-domain

The three-dimensional solution structure of the apo β-domain of DnaK (DnaK(393–507)) was obtained by multi-nuclear multi-dimensional NMR spectroscopy (Table 1, Fig. 2*a,b*). This is the first reported structure of an apo form of any of the Hsp70s, as previously studied constructs containing part or all of the α-helical domain had a high propensity either to interact intramolecularly[4,5] or to oligomerize in the absence of peptide[3,13]. The fold of the β-domain is similar to that of the peptide-bound substrate binding domains[3–5] (Fig. 2*c,d*), consisting of two sheets with four antiparallel β-strands each: a first sheet of strands β3, β6, β7, β8, and a second highly irregular sheet of strands β5, β4, β1, β2 (Fig. 2*b*). Surprisingly, loops L1,2 and L4,5 have retained their structure despite removal of the helix, which is in direct contact with these loops in the X-ray structure (Fig. 2*d*). The retention of the conformation of these loops likely stems from a hydrophobic cluster formed by residues Met 404 in loop L1,2 and Leu 441 and Ala 448 in loop L4,5.

Despite the overall similar fold, there are several very significant differences between the structure of the apo β-domain and the structures of this domain with the binding site occupied[3–5]. These structural differences are localized primarily in the C-terminal half of strand β3 and in the loop regions (Fig. 2*a–d*). Particularly noteworthy is the shift observed for loop L3,4 and, to a smaller extent, for loop L1,2, where the side chains of Thr 428, Ala 429, and Met 404 protrude to partially occlude the space that is taken up by the substrate in the peptide-bound structures of DnaK (Fig. 2*e*). In the crystal structure[3] of DnaK(389–607) and in the solution structure[4] of DnaK(386–561), the region Gln 424–Ser 427 of strand β3 is an integral part of the lower β-sheet and is hydrogen bonded to strand β6 (Fig. 2*c*,*d*). Here, no characteristic cross strand NOEs can be found for these residues, indicating that this part of the sheet is not formed. Intriguingly, the regions with large structural differences are also characterized by extensive line broadening of their $^{1}$HN resonances (Fig. 3). This broadening is most dramatic for residues in the regions of Gln 424, Val 425, Phe 426 and Ser 427, and for the amide protons of Gly 405, Thr 417 and Ile 418, which have completely disappeared from the amide proton detected NMR spectra. The $^{1}$HN line broadening is much more severe at 800 MHz than at 500 MHz; moreover, the resonances become narrower at higher temperatures. We therefore conclude that the broadening arises from intermediate exchange among multiple conformations at the millisecond–microsecond time scale. The conformational flexibility of apo β-domain is not restricted to the substrate binding cleft, but extends to loop L2,3, which is remote from both the cleft and the α-helical domain (Fig. 3).

One might attribute the observed structural changes and conformational flexibility to a destabilizing effect caused by the removal of the helical region. However, the chemical shift dispersion in the $^{15}$N HSQC spectrum is identical to that of well-folded constructs that contain part of the helical domain[4,5]. Furthermore, the spectrum does not change noticeably between 15 and 45 °C. The construct precipitates quantitatively between 45 and 50 °C, a temperature close to the 55.6 °C transition assigned to the thermal unfolding of the substrate binding domain in wtDnaK[14], suggesting that the isolated β-domain folds with similar stability.

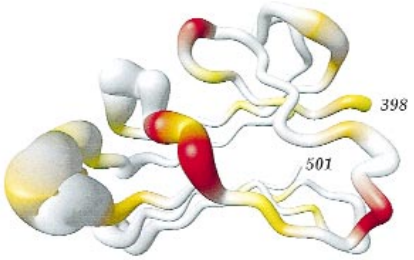


**Fig. 3.** Backbone conformation flexibility in the apo β-domain. A cylinder was drawn through the Cα positions, and the thickness of the cylindrical rod is proportional to the mean of the global displacement of the 20 energy-minimized DYANA conformations, calculated after superimposing all heavy atoms of the β-strands to minimize r.m.s. deviations. The color code represents the differences in $^{1}$HN line width (ΔLW) in spectra measured at 800 MHz and 500 MHz proton frequencies. The color code used is: |ΔLW| < 5 Hz, white; 5 Hz < |ΔLW| < 10 Hz, yellow; |ΔLW| > 10 Hz, orange. In red are residues that are exchange-broadened beyond detection and do not exhibit crosspeaks in absence of peptide.

### Peptide binding to apo β-domain monitored by NMR

Titration of $^{15}$N/$^{13}$C labeled apo β-domain with the peptide NRLLLTG[3,12,15] was followed with $^{15}$N,$^{1}$H and $^{13}$C,$^{1}$H correlation spectroscopy (Fig. 4). Upon addition of the peptide, new crosspeaks appear at the expense of those assigned to the apo β-domain, indicating that the kinetics of binding are slow on the chemical shift time scale. By quantifying the intensity changes of the NMR resonances corresponding to either the apo β-domain or the peptide-bound β-domain in the titration, we estimate a $K_d$ of 600 ± 200 μM for the NRLLLTG–β-domain complex. This binding is substantially weaker than that reported for the binding of this peptide to wtDnaK in the presence of ATP (43 μM)[12]. Interestingly, despite the 15-fold weaker binding, the rate constants for association and dissociation of NRLLLTG to DnaK(393–507) are of the same order of magnitude as the values for binding of this peptide to wtDnaK in the presence of ATP[12]. The lack of line broadening for any of the NMR crosspeaks corresponding to either apo or peptide-bound β-domain throughout the entire titration allowed us to estimate an upper limit of $2 \times 10^{4}$ $M^{-1}$ $s^{-1}$ for the bimolecular association rate constant, which in combination with the measured $K_d$ yields 12.5 $s^{-1}$ for the dissociation rate constant (see Methods). By comparison, binding of acrylodan-NRLLLTG to wtDnaK–ATP is characterized by a bimolecular association rate constant of $1 \times 10^{5}$ $M^{-1}s^{-1}$ and a dissociation rate constant of 4.5 $s^{-1}$ (ref. 12).

Loss of binding affinity of the β-domain could be a consequence of either removal of the helical lid or absence of the ATPase domain. The latter is consistent with the finding that DnaK(1–507) bound to ATP displays a higher affinity for peptide (37 μM) than does the isolated β-domain DnaK(393–507) (600 μM). However, the entire C-terminal domain (384–638) binds NRLLLTG tightly, with a $K_d$ of 11 μM[16] (midway between the $K_d$ for binding of peptide to wtDnaK in the absence (0.2 μM)[12] and presence of ATP), arguing that it is not the lack of the ATPase domain that compromised the β-domain binding affinity. Since the affinity of the entire C-terminal domain is still considerably tighter than that of the β-domain DnaK(393–507), it is possible that residues 384–393 influence peptide binding affinity in the absence of the ATPase domain. This conserved linker region may play a structural or allosteric role in influencing access to the peptide binding cleft in the substrate binding domain. Supporting this role are recently reported results of mutations in this region, in which allosteric defects were noted[17].

Changes in the β-domain NMR spectra upon saturation with peptide suggest that the conformational exchange in the apo β-domain of the protein is caused by the lack of substrate, and not by the lack of the α-helical domain. The $^{15}$N,$^{1}$H correlation spectrum of the apo β-domain becomes devoid of excess line broadening, and the $^{1}$HN resonances for the C-terminus of strand β3 reappear. Large chemical shift changes occur in strand β3 and in loops L1,2, L2,3, L3,4 and L5,6 (Fig. 4*a*). Similar shifts occurred when a different peptide (KLIGVLCSLFRPK; ref. 18) was added. The final $^{1}$HN, $^{15}$N, and $^{13}$C chemical shifts of the peptide-bound β-domain closely resemble those of the intramolecularly bound substrate binding domain[4] (Fig. 4*b*), with the exception that differences are observed in loops L1,2 and L4,5, which lie at the interface between the α-helical and the β-domains and would therefore be expected to have different local environments. This result indicates that the peptide-bound β-domain must closely resemble the previously studied substrate binding domains[3,4].

Significantly, chemical shift differences upon complex formation are not limited to the binding site region but also involve residues along strand β3 and loop L2,3 at the N-terminal end of the molecule (Fig. 4*d*). Together with the aforementioned

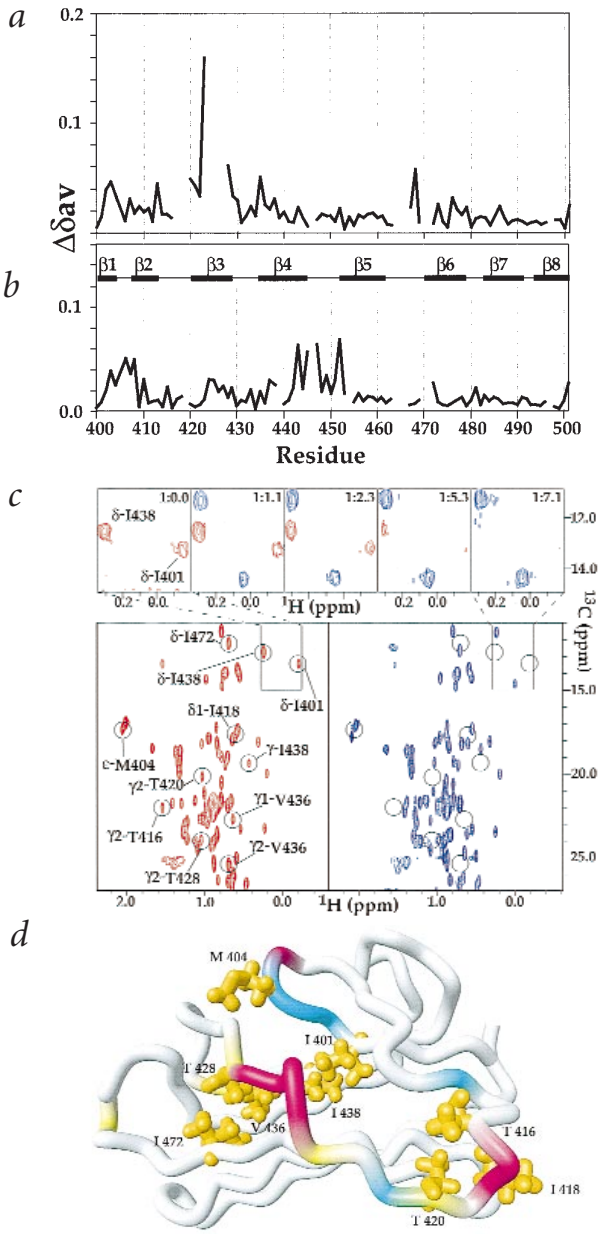


**Fig. 4.** NMR titration of DnaK(393–507) with the peptide NRLLLTG. Panels ***a*** and ***b*** show plots of the weighted averaged $^{1}$HN, $^{15}$N, and $^{13}$Cα, chemical shift differences (Δδav). To account for the intrinsic differences in chemical shift dispersion of $^{1}$HN, $^{15}$N, and $^{13}$Cα nuclei, Δδav was defined as: $\Delta\delta av = \{[(\Delta\delta^{1}HN / 3.5)^2 + (\Delta\delta^{15}N / 30)^2 + (\Delta\delta^{13}C\alpha / 20)^2]/3\}^{1/2}$. ***a***, Chemical shift differences between apo β-domain and β-domain in complex with the peptide NRLLLTG (0.66 mM β domain and 4.66 mM peptide). ***b***, Chemical shift differences between NRLLLTG-bound β-domain and the DnaK substrate binding domain consisting of the β-domain and part of the α-helical domain (DnaK386–561), which binds intramolecularly the residues Asp 540–Arg 547. ***c***, Methyl regions of the $^{13}$C,$^{1}$H correlation spectra of the β-domain in the apo (bottom panel, red) and peptide-bound form (bottom panel, blue). The complex formation of apo β-domain with NRLLLTG is slow on the chemical shift time scale, as indicated by the disappearance of the NMR resonances corresponding to the apo state and appearance of the resonances corresponding to the peptide-bound state. Several crosspeaks that clearly undergo these changes upon peptide binding are labeled, and their positions are indicated with circles. The top panels show the titration of the side chain methyl groups of Ile 401 and Ile 438 at different protein:peptide ratios (indicated on the top of each panel). A solution of 0.7mM $^{13}$C/$^{15}$N labeled β-domain was titrated with a 27.6 mg ml$^{-1}$ peptide solution. The concentrations of the reactants were determined using simple 1D NMR experiments. For each titration point a [$^{15}$N,$^{1}$H] HSQC and a [$^{13}$C,$^{1}$H] HSQC spectrum was recorded at 500 MHz $^{1}$H frequency. ***d***, Polypeptide backbone of the NMR structure of apo β-domain represented by a tube drawn through the Cα positions. Chemical shift variations due to the peptide NRLLLTG binding, Δδav (see Fig. 3*a*) are indicated by the following color code: white, Δδav < 0.035; cyan, 0.035 < Δδav < 0.049; yellow, Δδav > 0.049. In magenta are residues for which no assignments could be obtained for apo β-domain due to conformational exchange broadening (see text and Fig. 3). Side chains atoms of residues for which large methyl group chemical shift differences were observed (as from the spectra in *c*) are also displayed.

dynamics for these regions of the molecule, our results clearly point to a long range influence of peptide binding.

### Implications for the allosteric mechanism of DnaK

Our data establish that interaction between the α-domain and β-domain is not absolutely required for several of the functions of the DnaK chaperone protein. DnaK(1–507) can support bacteriophage λ propagation *in vivo*. This truncated construct can bind peptide substrates with a binding affinity only five-fold less than wtDnaK. Furthermore, the addition of ATP lowers the peptide binding affinity by a similar magnitude (100-fold) as occurs for wtDnaK. The ATP-induced modulation of peptide binding affinity indicates that DnaK(1–507) retains allosteric communication. Peptide-induced stimulation of ATPase activity in DnaK(1–507) shows that the allosteric interface is functional in both directions. These data, together with previous observations that partial truncations of the helical domain do not abolish Hsp70 functions[8], argue that the β-domain itself controls a significant portion of the ATP-induced allosteric signal.

We have observed substantial structural changes in the apo β-domain in comparison to the peptide-bound forms, which appear to arise from lack of a bound substrate and not from loss of the helix. Therefore, one may investigate these effects to learn about potential allosterically important conformational changes. A very significant difference between the apo form and substrate-bound form of the β-domain is found at the C-terminal end of strand β3. In the apo form, it does not participate in β-sheet formation, occludes the substrate binding cleft and exhibits conformational exchange dynamics. In the peptide-bound form, these residues shift conformation and participate in the β-sheet as their resonances become narrow and coincide with those of the construct that has this region as a very regular β-sheet (Figs 4*b*, 2*c*). In addition, cross strand NOEs between β3 and β6 are observed for these residues in the peptide-bound form (S.Y.S., M.P. and E.R.P.Z., unpublished results). As strand β3 (Thr 420-Lys 421-His 422-Ser 423-Gln 424-Val 425-Phe 426-Ser 427) does not possess a strong propensity to adopt a β-structure[19], it is tempting to speculate that this region serves as a conformational switch. Our observations also indicate that peptide binding induces structural and dynamic changes that are propagated from the binding cleft towards loop L2,3. Conversely, substrate binding properties could then be modulated through long range effects that originate from structural adjustments in the area of loop L2,3.

It is very interesting to note that several mutations that affect the allosteric function of DnaK have been mapped to the same loop L2,3 region (residues 410–420). The DnaK mutant K414I cannot support λ replication *in vivo* and also shows an absence of allosteric functions *in vitro*[9]. In addition, proteolytic cleavage at Lys 414 in DnaK is much more pronounced in the ATP state than in the ADP state, implicating this site in nucleotide-induced conformational changes[16]. Two recent studies[20,21] reported that

**Table 1 NMR structure statistics of apo DnaK(393–507)**

| | |
|---|---|
| DYANA target function (Å²) | 0.75 ± 0.22 (0.37–1.15) |
| Residual distance constraint violations[1] | |
| Number ≥ 0.2 Å | 1 ± 1 (0–3) |
| Maximum (Å) | 0.22 ± 0.06 (0.14–0.33) |
| Residual dihedral angle constraint violations[1] | |
| Number ≥ 2.5° | 0 ± 0 (0–0) |
| Maximum (°) | 0.03 ± 0.02 (0.01–0.08) |
| FANTOM energy (kcal mol⁻¹) | -421 ± 48 (-362– -506) |
| PROCHECK Ramachandran statistics[2] (%) | 78.9, 13.7, 6.3, 1.1 |
| R.m.s. deviations to the averaged coordinates (Å) | |
| Secondary structure elements[3] | |
| Backbone atoms | 0.58 |
| Backbone and heavy side chain atoms | 1.07 |
| Full structure excluding the two ill defined loops[4] | 0.88 |
| Pairwise r.m.s. deviations between the mean structure and X-ray or NMR coordinates for different constructs (Å)[3] | |
| X-ray (PDB code 1DKZ) | 1.52 |
| NMR (PDB code 1BPR) | 1.71 |

[1]Before energy minimization.
[2]Statistics for the mean coordinates after energy minimization determined with the software PROCHECK NMR[30] for residues in most favored region, additionally allowed region, generously allowed region, and disallowed region respectively.
[3]Sequence including residues 399–403, 407–412, 420–423, 435–442, 452–460, 472–478, 484–489 and 496–501.
[4]Sequence including residues 398–423, 435–460 and 472–501.

mutation of Pro 419 (DnaK numbering) in Hsp70s leads to defects in function *in vivo*, despite retention of substrate binding capacity *in vitro*.

In summary, our findings lead to the following conclusions that are relevant for the allosteric mechanism of DnaK: (i) the role of the helix is considerably less prominent than previously believed; (ii) loop L3,4 can occlude the binding groove and compete with substrate; (iii) loop L2,3, previously shown to be important for allosteric control, communicates with the substrate binding area. Further evaluation of these conclusions in terms of a full fledged hypothesis awaits mutagenesis studies and structural characterization of Hsp70 chaperones that include both substrate and ATP binding areas.

## Methods

**Construction and purification of the β-domain and DnaK(1–507).** DnaK(1–507) and the β-domain (residues 393–507) were constructed by PCR amplification of segments of the wild type *dnaK* gene and subcloned into pMS-EH[9] and pET15b (Novagen), respectively. DNA sequences introduced by PCR amplification were confirmed by automated DNA sequencing at the University of Massachusetts facility. The β-domain was expressed in the *E. coli* strain BL21(DE3) and contains a His-tag (MGSSHHHHHHGLVPRGSHM)[393]DnaK[507]. The protein was purified using a Ni-chelating column (Qiagen). The N-terminal His-tag was not cleaved from the protein. DnaK(1–507) was expressed in the *E. coli* strain BB1553 that is DnaK deficient[11]. The protein was purified as reported[9].

**Characterization of the β-domain and DnaK(1–507).** Both the apo β-domain and the peptide-bound β-domain are monomeric in solution as indicated by $^{15}$N NMR relaxation experiments ($\tau_c$ = 10 ns at 30 °C). Bacteriophage λ propagation assays for DnaK(1–507) were performed as described for wtDnaK[9]. Briefly, the pMS vector is IPTG-inducible, although in the absence of IPTG it is not tightly repressed and expresses wtDnaK at a level (as determined by Western blot) similar to that of a normal (not heat-induced) *E. coli* strain DH5α[9]. λ propagation assays with DnaK(1–507) were performed in the absence of added IPTG in the BB1553 strain at 30 °C and assayed side-by-side with the pMS-wtDnaK construct under the same conditions. Peptide binding of DnaK(1–507) was measured in 20 mM HEPES, 5 mM $MgCl_2$, 100 mM KCl, 5 mM DTT, pH 7.6 at 25 °C at peptide concentration of 40 nM. Binding affinity is determined by monitoring fluorescence anisotropy of the N-terminal fluoroscein-labeled peptide CALLQSRLLLSAPRRAAATARY (F-APPY) as described[9]. The ATPase activity of DnaK(1–507) was assayed at 1 μM in 40 mM HEPES, 50 mM KCl, 11 mM $Mg(OAc)_2$, 5 mM DTT, pH 7.6 at 30 °C, using the procedure as described[9]. Protein concentrations of DnaK(1–507) were measured using an extinction coefficient of $\varepsilon_{280} = 15.4 \times 10^3$ M$^{-1}$cm$^{-1}$.

**NMR spectroscopy, resonance assignments and structure calculation.** The NMR resonance assignments were obtained with 0.7 mM uniformly $^{15}$N labeled and uniformly $^{15}$N,$^{13}$C labeled apo β-domain in 90% $H_2O$/10% $D_2O$ or in 100% $D_2O$ containing 10 mM sodium phosphate buffer (pH 7.4) using double and triple resonance experiments. Unless otherwise noted, the spectra were measured on a Varian Inova spectrometer operating at 800 MHz $^1$H frequency. All measurements were performed at 30 °C. The chemical shifts for apo β-domain were obtained using HNCA, HN(CA)HA, $^{15}$N-resolved [$^1$H,$^1$H] NOESY (mixing time 70 ms), $^{15}$N-resolved [$^1$H,$^1$H] TOCSY, 3D H(C)CH-TOCSY (H)CCH-TOCSY, $^{13}$C-resolved [$^1$H, $^1$H] NOESY and 3D [$^{13}$C, $^{13}$C, $^1$H] HMQC-NOESY-HSQC. Backbone resonance assignments for the peptide bound β-domain were obtained with HNCA, HN(CA)HA and HNCACB (for a review of these methods see ref. 22). Stereospecific assignments for the methyls of Val and Leu were obtained using fractional $^{13}$C labeling[23]. All spectra were processed with the software PROSA[24]. Data and spectral analysis was performed with the program XEASY[25]. The chemical shifts of DnaK386–561 were obtained previously[4]. NOE distance restraints were obtained from 3D $^{15}$N-resolved [$^1$H,$^1$H] NOESY, 3D $^{13}$C-resolved [$^1$H, $^1$H] NOESY and 3D [$^{13}$C, $^{13}$C, $^1$H] HMQC-NOESY-HSQC recorded with a mixing time 70 ms. The final input for the torsion angle dynamics calculation with the program DYANA[26] consisted of 612 upper distance limits (283 intrareside, 132 sequential, and 239 long range) derived from NOE crosspeaks, hydrogen bond constraints (included after they were identified from an initial round of structure calculation), and 92 dihedral angle constraints (46 φ, 46 ψ) derived from $^{13}$Cα chemical shifts. The standard simulated annealing protocol was used and the final round of DYANA structure calculations was started with 50 randomized conformers. The 20 conformers with the lowest target function were subjected to restrained energy minimization *in vacuo* with a variable dielectric constant using the program FANTOM[27]. The statistics of the structure determination are listed in Table1. Structure analysis was carried out and color figures were made with the program MOLMOL[28].

**Peptide titration monitored by NMR.** The titration experiments were carried out with the peptide NRLLLTG using a solution of 0.66 mM β-domain. The peptide was synthesized by the University of Michigan Medical School Core Facility. A concentrated solution of NRLLLTG was added to the protein in peptide / protein ratios of 0, 0.1, 0.2, 0.4, 1.0, 2.1, 3.0 and 5.0. The binding was monitored by $^{15}$N-$^1$H and $^{13}$C-$^1$H HSQC spectroscopy at 11.7 T (500 MHz $^1$H). The chemical shifts for Fig. 4*a*,*b* were from the final titration sample. The kinetics of binding is slow with respect to the chemical shift difference. The dissociation constant was obtained by fitting a binding isotherm to the intensities of the NMR peaks corresponding to the either the free or complexed β-domain, and the standard deviation was estimated by comparing several cross peaks. The intensities of all resonances affected by the titration changed simultaneously upon peptide addition, indicating a single binding event. No detectable line broadening (that is, <5 Hz) occurred for the signals of the free protein during titration. According to a kinetic lineshape simulation[29], this leads to an average lifetime ($\tau_{free}$) of the free pro-

tein (the time before an association with ligand occurs) longer than approximately 50 ms at any stage of the titration where the free protein concentration is larger than 0.1 mM (sufficient signal of the free protein to allow a linewidth determination). The lifetime $\tau_{free}$ can then be used to obtain the bimolecular on-rate $k_{on}$ from the equation $1/\tau_{free} = k_{on}[L]$. The free ligand concentration [L] is easily calculated from the experimental conditions and the measured $K_d$. Our data leads to a maximum $k_{on}$ of 2 x $10^4$ $M^{-1}s^{-1}$. From this, and the $K_d$, we obtain an upper limit of 12.5 $s^{-1}$ for $k_{off}$.

**Coordinates.** The atomic coordinates of the β-domain have been deposited in the protein data bank (accession code 1DG4).

**Acknowledgments**
This work was supported by NIH grants to E.R.P.Z and to L.M.G., and a NIH fellowship to D.L.M.. The W.M. Keck Foundation, NIH, NSF and Parke-Davis/ Warner Lambert are gratefully acknowledged for financial support towards the 800 MHz NMR instrument. We thank J. Feltham for critical reading of the manuscript, and R. Sivendran for help with the assays of peptide-stimulated ATPase activity.

Correspondence should be addressed to E.R.P.Z. *email: zuiderwe@umich.edu* or to L.M.G. *email: gierasch@chem.umass.edu*

Received 16 September, 1999; accepted 18 February, 2000

1. Bukau, B. & Horwich, A. L. *Cell* **92**, 351–366 (1998).
2. Flaherty, K.M., DeLuca-Flaherty, C. & McKay, D.B. *Nature* **346**, 623–628 (1990).
3. Zhu, X. *et al. Science* **272**, 1606–1614 (1996).
4. Wang, H. *et al. Biochemistry* **37**, 7929–7940 (1998).
5. Morshauser, R.C. *et al. J. Mol. Biol.* **289**, 1387–1403 (1999).
6. Bertelsen, E.B., Zhou, H., Lowry, D.F., Flynn, G.C. & Dahlquist, F.W. *Protein Sci.* **8**, 343–354 (1999).
7. McCarty, J.S., Buchberger, A, Reinstein, J. & Bukau, B. *J. Mol. Biol.* **249**, 126–137 (1995).
8. Misselwitz, B., Staeck, O. & Rapoport, T.A. *Mol. Cell.* **2**, 593–603 (1998).
9. Montgomery, D.L., Morimoto, R.I. & Gierasch, L.M. *J. Mol. Biol.* **286**, 915–932 (1999).
10. Yochem, J. *et al. Mol. Gen. Genet.* **164**, 9–14 (1978).
11. Bukau, B. & Walker, G. C. *EMBO J.* **9**, 4027–4036 (1990).
12. Pierpaoli, E.V., Gisler, S.M. & Christen, P. *Biochemistry* **37**, 16741–16748 (1998).
13. Rüdiger, S., Buchberger, A. & Bukau, B. *Nature Struct. Biol.* **4,** 342–349 (1997).
14. Montgomery, D.L., Jordan, R., McMacken, R. & Freire, E. *J. Mol. Biol.* **232,** 680–692 (1993).
15. Gragerov, A., Zeng, L., Zhao, X., Burkholder, W. & Gottesman, M.E. *J. Mol. Biol.* **235**, 848–854 (1994).
16. Buchberger, A. *et al. J. Biol. Chem.* **270**, 16903–16910 (1995).
17. Laufen, T. *et al. Proc. Natl. Acad. Sci. USA* **96**, 5452–5457 (1999).
18. Flynn, G.C., Pohl, J., Flocco, M. T. & Rothman, J. E. *Nature* **353**, 726–730 (1991).
19. Rost, B. & Sander, C. *J.Mol. Biol.* **232**, 584–599 (1993).
20. Burkholder, W.F. *et al. Proc. Natl. Acad. Sci. USA* **93**, 10632–10637 (1996).
21. Voisine, C. *et al. Cell* **97**, 565–574 (1999).
22. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. III & Skelton, N. J. *Protein NMR spectroscopy* (Academic Press, San Diego, 1996).
23. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. *Biochemistry* **28**, 7510–7516 (1989).
24. Güntert, P., Dötsch, V., Wider, G. & Wüthrich, K. *J. Biomol. NMR* **2**, 619–629 (1992).
25. Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. *J. Biomol. NMR* **6**, 1–10 (1995).
26. Güntert, P., Mumenthaler, C. & Wüthrich, K. *J. Mol.Biol.* **273**, 283–298 (1997).
27. Schaumann, T., Braun, W. & Wüthrich, K. *Biopolymers* **29**, 679–694 (1990).
28. Koradi, R., Billeter, M. & Wüthrich, K. *J. Mol. Graphics* **14**, 52–55 (1996).
29. Carrington, A. & McLachlan, A. Introduction to magnetic resonance with applications to chemistry and chemical physics. (Harper & Row, New York, 1967).
30. Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. & Thornton, J.M. *J. Biomol. NMR* **8**, 477–486 (1996).



# Structural basis for the function of *Bacillus subtilis* phosphoribosyl-pyrophosphate synthetase

Tine A. Eriksen[1], Anders Kadziola[1], Ann-Kristin Bentsen[2], Kenneth W. Harlow[2,3] and Sine Larsen[1]

*[1]Center for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. [2]Center for Enzyme Research, Institute of Molecular Biology, University of Copenhagen, Sølvgade 83H, DK-1307 Copenhagen, Denmark. [3]Present Address: Zealand Pharmaceuticals, Smedeland 26 B, DK-2600 Glostrup, Denmark.*

**Here we report the first three-dimensional structure of a phosphoribosylpyrophosphate (PRPP) synthetase. PRPP is an essential intermediate in several biosynthetic pathways. Structures of the *Bacillus subtilis* PRPP synthetase in complex with analogs of the activator phosphate and the allosteric inhibitor ADP show that the functional form of the enzyme is a hexamer. The individual subunits fold into two domains, both of which resemble the type I phosphoribosyltransfereases. The active site is located between the two domains and includes residues from two subunits. Phosphate and ADP bind to the same regulatory site consisting of residues from three subunits of the hexamer. In addition to identifying residues important for binding substrates and effectors, the structures suggest a novel mode of allosteric regulation.**

Phosphoribosylpyrophosphate (PRPP) synthetase, catalyzing the reaction of ribose-5-phosphate (R5P) with ATP to yield PRPP (5-phosphoribosyl-α-1-pyrophosphate) and AMP, links the pentose phosphate pathway to the pyrimidine and purine nucleotide *de novo* and salvage pathways, the biosynthesis of histidine and tryptophan, and pyridine nucleotide coenzymes[1,2]. As PRPP is a metabolite required at all times in the cell, the function of PRPP synthetases is central to life. The family of PRPP synthetases includes sequences from 19 organisms (Fig. 1). Detailed kinetic[3–9] and structure-function[10–13] studies have been performed with some of these enzymes. Among the more salient features reported are their requirement for inorganic phosphate ($P_i$) and free $Mg^{2+}$ ions as activators[3,9,13–16], their allosteric behavior[7,13,17], their complex quaternary structure[13,18], their reaction mechanism[19], and identification of residues at their active sites[8,10–12,20].

ADP is generally the most potent inhibitor of PRPP synthetases[7,13,16,17,21]. Regulation by ADP can be either competitive or allosteric in nature, but allostery is only observed at saturating concentrations of R5P. Regulation of PRPP synthetases by $P_i$ is also complex as it has multiple effects on their activity and structure[18]. $P_i$ can be either stimulatory or inhibitory, depending on concentration and the presence of substrates and effectors[13,22,23]. Although sequence similarity is high within the PRPP synthetase family, its members have no striking overall similarity to other proteins. However, a short region of the PRPP binding motif[16] shows homology to the type I phosphoribosyltransferases, in which it interacts with PRPP[24–26].

Two crystal structures of the PRPP synthetase from *Bacillus subtilis* were determined to 2.3 Å and 2.2 Å resolution, respectively, from crystals prepared under different conditions (Table 1). In the $SO_4^{2-}$–PRPP synthetase structure, sulfate ions