Articles

Solvent Interaction of a Hsp70 Chaperone Substrate-Binding Domain Investigated with Water−NOE NMR Experiments†

Sheng Cai, Shawn Y. Stevens, Andrew P. Budor, and Erik R. P. Zuiderweg*

Biophysics Research Division and Departments of Biological Chemistry and Chemistry, The University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109

Received April 21, 2003; Revised Manuscript Received July 23, 2003

ABSTRACT: The interaction of solvent of the substrate binding domain of the bacterial heat shock 70 chaperone protein DnaK was studied in its apo form and with bound hydrophobic substrate peptide, using refined nuclear magnetic resonance experiments. Distinct differences between the two states of the protein were observed. According to our data, the apo form interacts more extensively with solvent than the peptide-bound form. Significantly, the open hydrophobic substrate binding cleft of DnaK in the apo form is found to contain several molecules of water which are displaced by the binding of the hydrophobic substrate, the peptide NRLLLTG. The solvent in the hydrophobic cleft has a residence time longer than 400 ps. It is predicted that the displacement of this trapped water must contribute to the binding free energy of the natural hydrophobic substrates of this class of protein-folding chaperone proteins.

The interaction of water with proteins is a major determinant of protein stability through its definition of the hydrophobic interaction and by its efficient competition for internal hydrogen bonding. In protein crystal structures, many water molecules are found to be closely associated with polar surface groups and may confer correlated dynamical behavior to these groups (*1*). Hydrodynamic studies of translational and rotational diffusion indicate that proteins in solution are effectively covered by a water layer of on average 1.5 Å thickness (*2*). The current paper is concerned with a nuclear magnetic resonance (NMR)[1] study of water binding to a Hsp70 chaperone protein domain.

Hsp70 chaperone systems (70 kDa heat shock proteins) are the main facilitators of protein folding, refolding, trafficking, and triage in mammalian cells (*3*−*7*). The Hsp70 chaperones bind to exposed hydrophobic patches in unfolded, partially folded, or misfolded protein substrates and release these proteins upon ATP binding, potentially further unfolding them in the process (*8*−*11*). We focus on the substrate binding domain of DnaK. DnaK is the Hsp70 molecular chaperone protein of *Escherichia coli* which shares the protein (re)folding tasks in these organisms with the GroEL/S machinery (*6*, *12*). DnaK consists of a 44 kDa ATPase domain (*13*) and a 26 kDa substrate binding domain (*14*). The latter contains the 12 kDa DnaKβ domain (residues 393−507), which harbors the substrate binding cleft (*14*, *15*). The substrate binding cleft is wide and deep and is lined with methyl groups. Hydrophobic peptides, which serve as a model for partially unfolded substrate proteins, bind in an extended conformation in this cleft (*14*, *16*−*18*). A pocket at the bottom of the cleft snuggly accommodates the side chain of a single substrate leucine; other hydrophobic substrate residues are complemented less specifically (*14*, *16*−*18*). Substrate backbone polar atoms are hydrogen-bonded by the polar protein backbone and side chain atoms (*14*).

Of particular interest to us is that the hydrophobic substrate binding cleft of DnaKβ is exposed to solvent in the ligand-free state. This is a somewhat uncommon situation; most folded proteins do not expose extended hydrophobic areas to solvent. According to the classical hydrophobic model, this area should contain clathrated water (*19*), and our question is if such water can be detected by NMR and whether this water becomes perturbed or released upon binding of the hydrophobic substrate. To this end, we have refined NMR experiments that selectively measure the NOE between water and protein, based on experiments developed in several groups (*20*−*24*). The general approach is to selectively excite the water magnetization and let its magnetization transfer to the protein ^{1}H nuclei via dipolar cross-relaxation. The apolar protein acceptor groups of the transferred magnetization (i.e., water interaction sites) are then identified by a 2D ^{1}H−^{13}C HSQC experiment.

† Supported by Grant GM52421 from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (734) 936-3850. Fax: (734) 764-3323. E-mail: zuiderwe@umich.edu.

[1] Abbreviations: Hsp70, 70 kDa heat shock protein; DnaKβ, the substrate binding domain (residues 393−507) of the Hsp70 chaperone DnaK of *Escherichia coli*; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; WMT, water magnetization transfer.

10.1021/bi030097c CCC: $25.00
Published on Web 09/05/2003

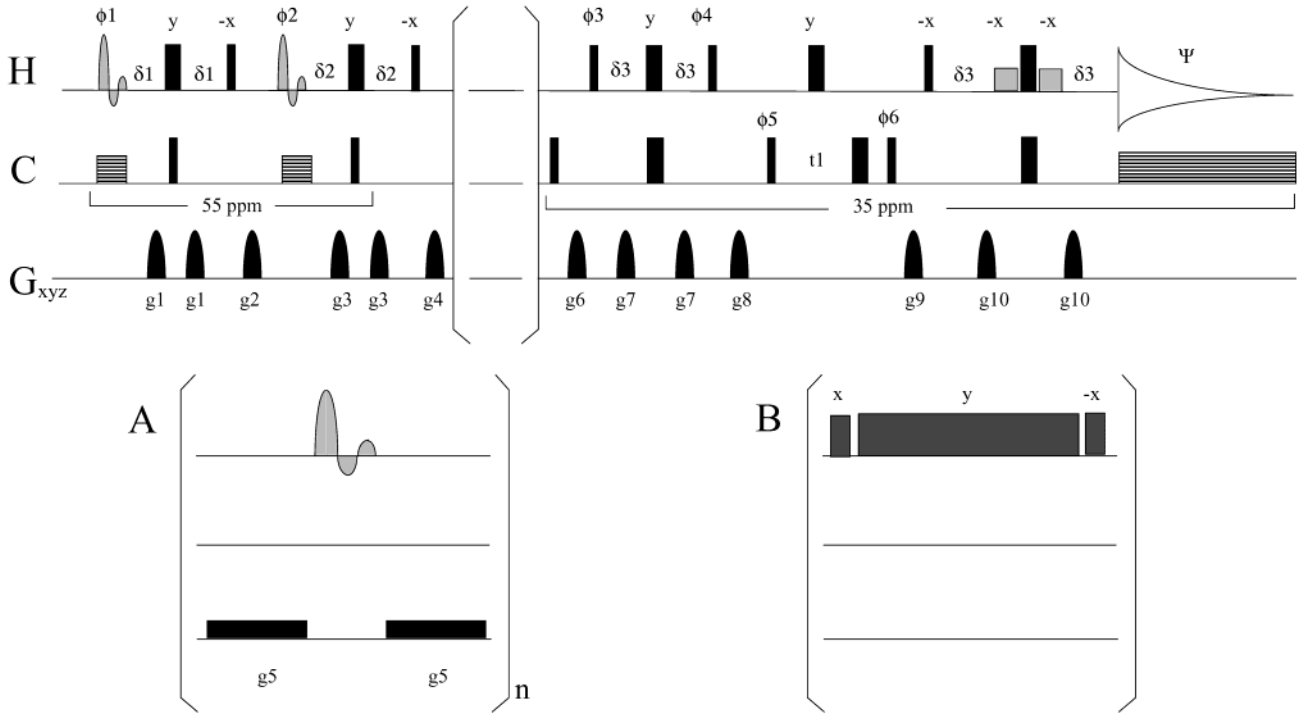

FIGURE 1: Pulse sequence for the water-magnetization-transfer (WMT) $^1H-^{13}C$ HSQC experiment. Insets A and B show laboratory-frame magnetization transfer and rotating frame magnetization transfer, respectively. The E-burp-2 pulses of 6 ms (with phases ϕ_1 and ϕ_2) were centered on the solvent resonance. The concomitant carbon decoupling was carried out with a GARP sequence centered at $^{13}C\alpha$, with a 3.12 kHz field strength. The adiabatic WURST decoupling pulses of 10 ms in inset A inverted the bandwidth of 6.2 ppm to 12.2 ppm. They were alternated with rectangular gradients g_5 of 5 ms with 0.3 G/cm along *Z*. The loop was carried out four times for a total mixing time of 80 ms. The power level for the proton pulses in the rotating frame experiment (inset B) was adjusted for a 80 μs 90° pulse. All other pulses were rectangular and had phases along *x*, unless indicated otherwise. The phase cycles are $\phi_1 = x, x, -x, -x$; $\phi_2 = x, x, x, x, -x, -x, -x, -x$; $\phi_3 = x, x, -x, -x, -x, -x, x, x$; $\phi_4 = 8(y), 8(-y)$; $\phi_5 = x, -x$; $\phi_6 = 8(x), 8(-x)$; $\Psi = x, -x$. Phase ϕ_5 was incremented for States−TPPI quadrature detection. Gradient g_{10} was a half-sine of 0.5 ms with 14 G/cm along *X*. All other gradients were 1 ms half-sines with strengths $g_1 = 12$ G/cm along *X*, $g_2 = -48$ G/cm along *Z*, $g_3 = 12$ G/cm along *Y*, $g_4 = 42$ G/cm along *Z*, $g_6 = 18$ G/cm along *X*, $g_7 = 6$ G/cm along *Z*, and $g_8 = g_9 = 2.4$ G/cm along *X*. The delays were $\delta_1 = 3.57$ ms, $\delta_2 = 3.3$ ms, and $\delta_3 = 1.5$ ms. The Watergate 90° pulse was 770 μs. The sequence is shown with a $^1H-^{13}C$ Watergate fHSQC (*36*) and returns the solvent resonance to $+Z$ during the detection period provided it remains in the *xy* plane during the ^{13}C t_1 period. This was the case for the 500 MHz instrument with conventional probe used. For systems with higher probe *Q*-factors, it is recommended to use a selective pulse to return the water resonance to $+Z$ during the t_1 period. For amide proton detection, the HSQC sequence can be replaced by a fast $^{15}N-^1H$ HSQC, a sensitivity-enhanced HSQC (*37*), or a TROSY (*38*) depending on need.

The refined water NOE experiments described in this paper indicate that the apo form of DnaKβ interacts more extensively with solvent than the peptide-bound form. Most significantly, we detect several molecules of water in the open hydrophobic cleft of apo DnaKβ. This water is found to be displaced by the binding of a hydrophobic substrate, the peptide NRLLLTG. The solvent in the open hydrophobic cleft has a residence time longer than 400 ps. It is predicted that the displacement of this bound water must contribute to the binding free energy of the natural hydrophobic substrates of this class of chaperone proteins. The nature of the bound water is discussed extensively.

MATERIALS AND METHODS

General. The expression and purification of the uniformly ^{15}N,^{13}C-labeled β-domain (residues 393−507) of DnaK were described previously (*13*). Both the apo β-domain sample (0.5 mM) and the peptide-bound sample (0.7 mM) were in 90% H_2O/10% D_2O containing 10 mM sodium phosphate buffer (pH 7.4). The peptide-bound sample was made by direct titration of the peptide NRLLLTG into the apoprotein in a peptide/protein ratio of 6:1, as described previously (*15*). The peptide NRLLLTG was synthesized by the University of Michigan Medical School Core Facility. All NMR spectra were acquired on a Bruker Avance 500 MHz NMR spectrometer using a sample temperature of 30 °C, were processed using NMRPipe (*25*), and were analyzed with NMRView (*26*).

The NMR Experiment. The water-magnetization-transfer $^1H-^{13}C$ HSQC (WMT ^{13}C HSQC) spectra were acquired using the pulse sequence shown in Figure 1. The scheme used combines ideas of Otting and Wüthrich (*20*, *21*, *27*), Gemmecker and Kessler (*28*), Grzesiek and Bax (*29*), Clore and co-workers (*22*), Boelens and co-workers (*23*), and Guéron and co-workers (*24*). Following an approach in ref *24*, two spin−echo blocks of three radio-frequency (rf) pulses each selectively return the solvent magnetization at the *Z*-axis while the protein proton magnetization has just experienced a net 270° pulse. The latter magnetization is subsequently dephased by a field gradient, leaving zero protein magnetization components. In practice, the sequence also retains protein H_α resonances close to the water resonance. The spin−echo filter suppresses this unwanted coherence in two ways: First, the relatively long delays (and the repeat of the entire spin−echo element) allow for R_2 relaxation during the filter; for resonances of large proteins this is quite effective. This implements an idea proposed in refs *24*, *30*, and *31*. Second, because a ^{13}C-labeled protein was used for the current studies,

one can selectively purge ^{13}C-coupled H_a magnetization following ideas originally proposed in refs *28* and *29*. The delay between the first selective 90° ^1H pulse and the 90° ^{13}C pulse is set to $^1/_2J_{CH}$. The partially excited H_α magnetization is at that point converted to two-spin coherence and dephased by the following gradient. Hence, all protein H_α *Z*-magnetization vanishes. Evolution of $2H_xC_z$ magnetization during the long selective pulse is suppressed by $^{13}C\alpha$ decoupling. The spin−echo sequence is repeated to obtain complete protein signal saturation and to cover a wider range of scalar couplings as the spin−echo delays are set differently.

Our sequence resembles an experiment proposed by Grzesiek and Bax (*32*), which has been referred to as WNOESY (*25*). That experiment also contains a combination of the spin−echo filter and soft pulses. However, in that sequence the protein magnetization remains along the positive *Z*-axis at all times, while the water resonance is selectively pulsed to $-Z$ or $+Z$. Hence, no magnetization transfer takes place in the subsequent mixing period in half of these experiments. In our proposed experiment, both subexperiments generate magnetization transfers since the protein magnetization is saturated. The basic idea for this approach was proposed earlier in ref *30* with the so-called WEX-II filter, which was combined with a 1D NMR readout. Further differences between the WNOESY experiment, and all other proposed water−NOE experiments, and our experiment occur during the mixing period.

The magnetization transfer (mixing) period of 80 ms is composed of several loops (see Figure 1, inset A). Each of these contains a band-selective 180° pulse covering the ^1H region from 6.2 to 12.2 ppm to suppress magnetization relay through resolved exchangeable protons, interleaved with a long (10 ms per loop) low-power pulsed field gradient which serves to prevent radiation damping-induced relaxation processes of the H_2O resonance. This NOE spin diffusion suppressing decoupling pulse sequence element is based on original ideas described in refs *33*−*35* and was in a more simple form applied using continuous wave decoupling to water−NOE experiments in the experiments of ref *23*. Applied to DnaKβ, it resulted in a significant reduction of the number of water cross-peaks. Its further merits are discussed in the main text. After the magnetization transfer period, the NOE-transferred protein magnetization was detected using a $^1H-^{13}C$ fast HSQC−Watergate sequence (*36*). A phase cycling was employed that returns the solvent resonance to $+Z$ during data acquisition for all scans, enhancing sensitivity. This was achieved by "following" the water resonance with the phase cycle of the first ^1H pulse of the HSQC sequence as shown.

Overnight 2D control experiments, one in which the two selective pulses were turned off and one in which the NOE mixing time was set to a few microseconds, showed zero signal, verifying the validity and selectivity of the pulse sequence. It was also verified that no NOEs were detected with this sequence for a protein dissolved in 100% ^2H_2O. It was confirmed that the band-selective amide proton pulses did not perturb the solvent resonance using a test sequence consisting of one of the band-selective pulses followed by a single 1° proton pulse.

An experiment where the ^{13}C HSQC was replaced by a ^{15}N Watergate−HSQC was successfully tested. The latter element can also be substituted with a sensitivity-enhanced ^{15}N HSQC (*37*) or ^{15}N TROSY (*38*) if so desired. In these amide proton detected experiments, the NOE mixing time consisted of a long low-power gradient pulse only. ROE forms of both the WMT ^{13}C and ^{15}N HSQC experiments were also tested successfully. Here the mixing time was replaced by a gradient−power switch−90x−LOCK(y)−90(−x)−power switch−gradient block as shown in inset B of Figure 1.

RESULTS AND DISCUSSION

NMR Spectra. The water-magnetization-transfer (WMT) NMR experiment (Figure 1) used in these studies (WMT ^{13}C HSQC) was redesigned from literature experiments, optimized, and calibrated. Our sequence saturates all protein resonances while selectively retaining the water magnetization, which subsequently donates its polarization to the saturated protein through dipolar (NOE) and chemical exchange magnetization transfer. The accepted magnetization is read out by a 2D $^{13}C-^1H$ HSQC sequence. Most of the enhancements made in the sequence are of a technical nature, except for the use of decoupling in the low-field region of the protein ^1H NMR spectrum during the magnetization-transfer period. This suppresses some of the interfering and uninteresting relayed transfer processes, extending ideas presented in ref *23*. Details are described in the Materials and Methods section and in the legend to Figure 1.

Panels a and b of Figure 2 show a comparison of the methyl region of the normal $^1H-^{13}C$ HSQC spectrum of apo (unliganded) DnaK-β with the results of the WMT ^{13}C HSQC experiment. The comparison of the NMR spectra shows that cross-peaks to only a select number of methyl groups result from the WMT experiment. This subset of cross-peaks is not merely a subset of the most intense peaks in the HSQC spectrum. With the use of a relatively short magnetization-transfer period (80 ms), the cross-peaks are also found to have very different relative intensities as compared to the HSQC reference peaks (Table 1). Similar observations of selectivity hold for the comparison of the reference and the WMT ^{13}C HSQC experiment carried out for DnaKβ with bound substrate (the peptide NRLLLTG; see Figure 2c,d and Table 1).

The great majority of the water cross-peaks in Figure 2b,d have the same sign as the peaks in the corresponding HSQC reference spectra, which for the current NMR experiment implies that they have the same sign as the magnetization-donating solvent resonance. This is of relevance to the mechanism of magnetization transfer, which we summarize in the following for convenience.

Three different processes are known to contribute to the transfer of magnetization between solvent water and protein protons, hence resulting in cross-peaks in the WMT NMR spectra: (i) dipolar cross-relaxation (direct NOE), (ii) chemical exchange, and (iii) exchange-relayed NOE (*20*, *23*, *27*).

The first process is the one of current interest as it identifies resident water. It gives rise to cross-peaks with the same sign as the donating water magnetization (positive cross-peaks) when the correlation time describing the reorientation of the local protein−water internuclear vector, τ_C, is longer than the inverse spectrometer frequency ω_0 ($\omega_0\tau_C > \sqrt{1.2}$, or $\tau_C > 356$ ps for the 500 MHz spectrometer

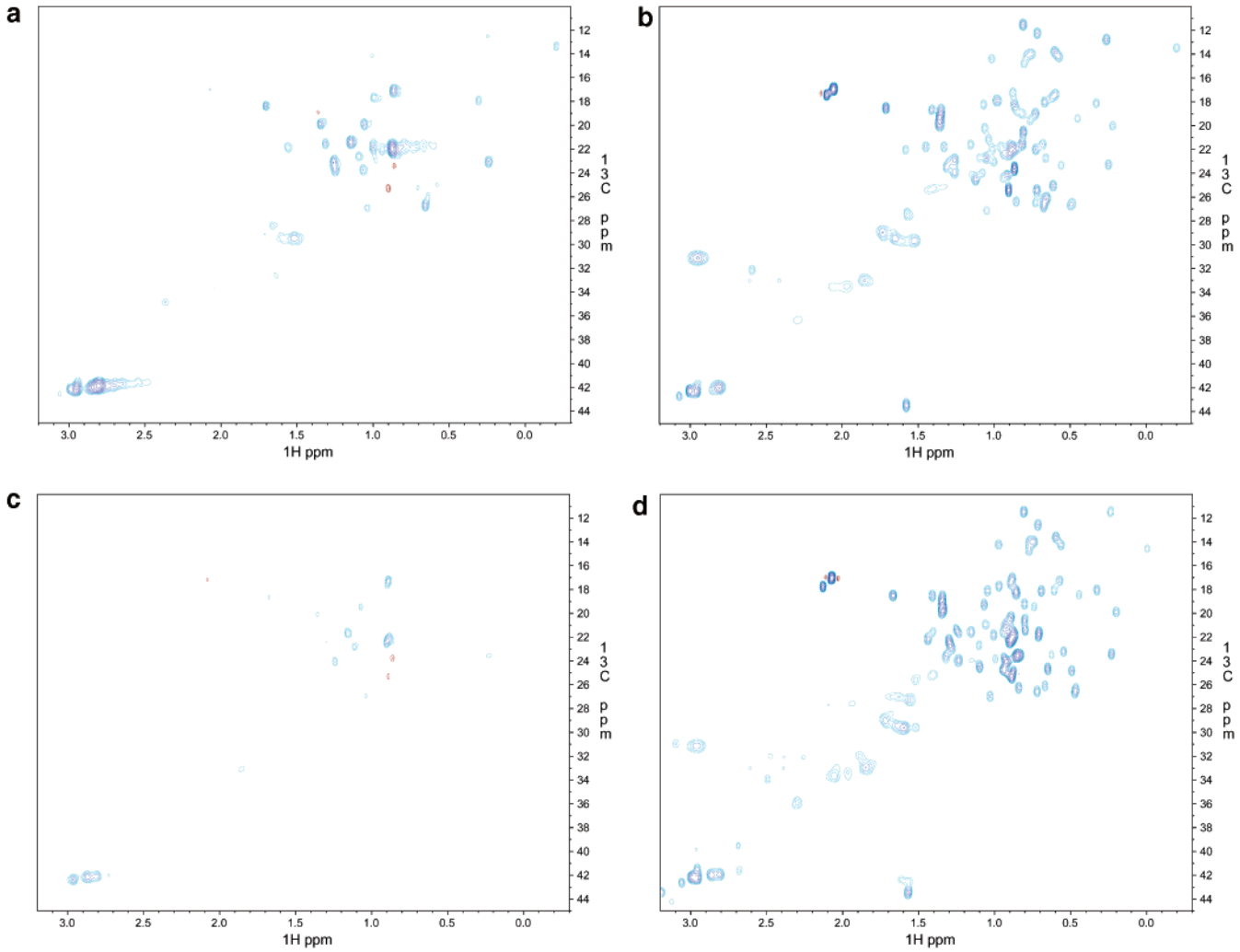

FIGURE 2: (a) WMT ^{13}C HSQC spectrum (mixing time 80 ms) of 0.5 mM apo-DnaKβ at 30 °C. The data were collected with 128 complex t_1 values and 512 scans in 44 h using a 500 MHz NMR system with conventional probe and the pulse sequence of Figure 1a. (b) $^1H-^{13}C$ correlation NMR spectrum of the same sample as in (a), using a Watergate $^1H-^{13}C$ HSQC sequence element as in (a), acquired with a 5 s recovery delay, 128 complex t_1 values, and eight scans. (c) WMT ^{13}C HSQC spectrum (mixing time 80 ms) of 0.7 mM DnaKβ complexed with a 6-fold excess of NRLLLTG at 30 °C. The data were collected with 128 complex t_1 values and 512 scans in 44 h. (d) $^1H-^{13}C$ correlation NMR spectrum of the same sample as in (c), using a Watergate $^1H-^{13}C$ HSQC sequence element as in (c), acquired with a 5 s recovery delay, 128 complex t_1 values, and eight scans.

used in this study). The correlation time τ_C is dominated by the shortest of the (local) rotational correlation time of the protein, τ_R, and the water residence time, τ_W, according to $\tau_C^{-1} = \tau_R^{-1} + \tau_W^{-1}$. The situation with positive cross-peaks corresponds to a negative dipolar cross-relaxation rate and is hence commonly referred to as a negative NOE. We will call these signals trapped water NOEs as they can only occur for ordered regions of the protein if $\tau_W > 400$ ps. The first process gives rise to cross-peaks with a sign opposite to that of the donating water magnetization when the water residence time at the site is shorter than 300 ps and/or if the protein is locally disordered at the water binding site.

The second process, direct chemical exchange, is not of relevance for the current investigation, since only resonances from nonexchangeable protons (methyls) are observed.

The third process is both relevant and problematic: it is a two-step magnetization-transfer process in which the water magnetization is first transferred via proton mass exchange to a labile protein proton (Thr, Ser, and Tyr hydroxyls, N-terminus, Arg, Asn, Gln, Lys, Trp side chains, exposed amide protons in general, and, at low pH, possibly also Asp, Glu, and C-terminal carboxyls) and then to other protein protons via macromolecular dipolar cross-relaxation. The rates of these relay processes depend on the exposure and hydrogen bonding of the labile protons and on the temperature and pH of the sample. It is hardly predictable and different from protein to protein. These less interesting exchange-relayed NOEs have the same cross-peak sign as trapped water NOEs. They cannot be distinguished from these through NOESY/ROESY comparisons either (*20*). However, the relayed NOEs can be suppressed if the resonances of the labile protons are resolved from the solvent resonance and can be decoupled during the magnetization-transfer period. This was first demonstrated by Boelens and co-workers (*23*), who used selective continuous wave decoupling fields directed at individual resolved Ser and Thr hydroxyl proton resonances. In this paper, this idea was extended by applying, during the mixing time, periodic band-selective decoupling pulses covering the entire ^1H chemical shift range from 12 to 6 ppm which encompasses most of the labile proton signals. This should eliminate all exchange relay via Asn and Gln side chain protons, via most of the

Table 1

	apo form			peptide-bound form			
assignment[a]	intensity[b]	δ_H	δ_C	intensity[c]	δ_H	δ_C	ratio[d]
L399CG	1.14	0.657	26.71				
I401CD1	1.13	−0.203	13.36				
T403CG2	1.15	1.094	22.60	1.02	1.105	22.87	0.740
T409CG2	0.96	1.313	21.55				
T410CG2	1.27	1.002	21.84				
L411CD1	0.39	0.58	24.99				
L411CD2	0.37	0.706	25.25				
I412CD1	0.65	1.008	14.17				
I412CG2	0.82	0.994	17.67	1.08	0.89	17.42	0.674
A413CB	0.74	1.346	19.88	0.28	1.342	19.86	0.240
T416CG2	1.14	1.563	21.85				
K421CE	2.15	2.806	41.88	0.84	2.816	42.17	0.981
T428CG2	1.15	1.067	23.76				
T437CG2	1.76	0.874	21.97				
I438CD1	0.43	0.247	12.53				
L441CD2	1.46	0.24	23.08	0.61	0.236	23.67	0.695
K452CE	1.20	2.951	42.16	0.74	2.962	42.42	0.948
Q456CB				0.51	1.85	33.15	0.312
I472CG1	1.03	1.039	26.93				
V474CG1				0.49	0.888	22.36	0.834
T475CG2	1.95	1.146	21.41	0.88	1.154	21.72	0.657
I483CG2	1.05	0.308	17.92				
K489CD	0.82	1.594	29.48				
K489CE				1.05	2.863	42.22	0.964
K491CD	1.16	1.663	28.41				
K498CG				0.63	1.239	24.14	0.561
K498CD	0.98	1.517	29.46				
I501CG1				0.74	1.036	27.20	0.829
K502CG	−0.18	0.861	23.43				
K502CD				−0.26	0.89	25.58	0.890
peak 1	2.16	0.865	17.10				
peak 2	−0.26	0.899	25.30				
peak 3	0.71	0.938	22.01				
peak 4	1.60	1.058	19.94				
peak 5	1.16	1.259	23.14				
peak 6	0.91	1.641	32.60				
peak 7	0.69	1.707	18.38				
peak 8	0.13	2.073	16.99				
peak 9	1.87	2.702	41.82				
peak 10	0.52	3.061	42.53				
peak 11				−0.20	0.868	23.71	1.214
peak 12				0.63	1.072	19.50	0.663
peak 13				0.36	1.671	18.71	0.937
average[e]	1.03			0.64			0.759

[a] Assignments of carbon atoms, in the form of residue type + residue number + carbon atom type. [b] Relative NOE intensity (percent) for the apoprotein, calculated from the peak intensities of the water−NOE spectrum divided by the corresponding peak intensities in the fully relaxed reference HSQC spectrum. [c] Relative NOE intensity (percent) for the peptide-bound protein using the same method. [d] The ratio of the NOE intensities with amide proton decoupling to the intensities without decoupling. [e] Calculated by changing all negative signs to positive.

Arg and Lys side chain protons, and via amide protons in general. A very significant difference between experiments with and without the band-selective decoupling was observed (results not shown). Many additional resonances appeared in the spectrum without decoupling, which could mistakenly have been identified as NOEs.

While this decoupling is clearly very useful, it still does not resolve the problem of exchange relay through labile protons which exchange very fast with solvent ($k_{ex} > 1000$ s^{-1}) and consequently resonate at the solvent frequency. These fast exchanging protons are likely also the most efficient relay conduits at the short NOE mixing times used. At the conditions of our experiments, these should include (some of) the Thr and Ser hydroxyl protons (DnaKβ has no Tyr residues) and some of the Arg and Lys labile side chain protons. Hence, all remaining trapped water−methyl cross-peaks in the decoupled water−NOE NMR spectra of Figure 2 must still be viewed with trepidation and must be culled from the water−NOE list if they belong to Thr methyls or if they belong to residues with protons that are within 5 Å ("NOE distance") of any of the very fast exchanging labile protons listed above.

Water Interactions with DnaKβ. The three-dimensional structure of the apo form of the substrate binding domain of DnaK, the Hsp70 protein of *E. coli*, was solved by NMR spectroscopy in our group (*15*). The substrate binding cleft is deep and is lined with hydrophobic residues, which serve to recognize the exposed hydrophobic side chains of partially unfolded proteins, constituting the main biological function of this chaperone protein. Hydrophobic peptides form extensive networks of contacts with hydrophobic residues in the binding cleft, especially with Ile 438 and Phe 426 (*14*, *16*−*18*). In the apo form of the protein, these hydrophobic residues line an open cleft (*15*). In addition, we observe that in the apo form residues Leu 399, Ile 401, and Leu 411 line a somewhat independent hydrophobic cavity of 41 Å^3, which is on one side open to the main cleft; this cavity is nonexistent in the peptide-bound form. However, even in the apo form, the substrate binding cleft is partially occupied by the dislodged extension of strand β3 (residues 423−427) of the β-sheet that forms the "bottom" of the protein in the view of Figure 3 (strands β3 and β6−β8). The position of this strand extension is not well-defined by NOE signals in the structural studies and fluctuates between different conformations as indicated by severe exchange broadening of the amide proton resonances in the area (*15*). The structural computation places the strand in the middle of the cleft, with room left open at both sides, that can potentially be occupied by solvent (see Figure 3a,b).

In the peptide-bound form, the dislodged strand participates in the β-sheet that forms the bottom of the protein both in the crystal (*14*) and in solution (*16*, *18*). Substrate binding also causes more widespread changes in the protein (*15*). Furthermore, the rotational correlation time of the protein as determined from ^{15}N relaxation studies changes from 15 ns in the apo form to 8 ns in the peptide-bound form (results not shown), potentially suggestive of a ligand binding-linked dimer−monomer equilibrium, currently under investigation in our laboratory. This potential dimer−monomer equilibrium is only of technical, but not of functional, relevance to the discussion of the water binding below.

The vast majority of the cross-peaks from the ^{13}C HSQC read-out experiments for both forms of the protein are positive; i.e., they have the same sign as the donating water resonance and would thus qualify as trapped water NOEs (see Table 1). The one assignable cross-peak with opposite sign unambiguously identifies Lys 502 as a water interaction site. This site is close to the (artificial) C-terminus of the protein and does not change between peptide-bound and peptide-free protein. The NMR data indicate that the Lys502 side chain by itself is very mobile with a fast correlation time (narrow NMR line width); this motion likely dominates the overall correlation time of the combined water−side chain system at this location and precludes a statement being made about the water residence time at this location (see

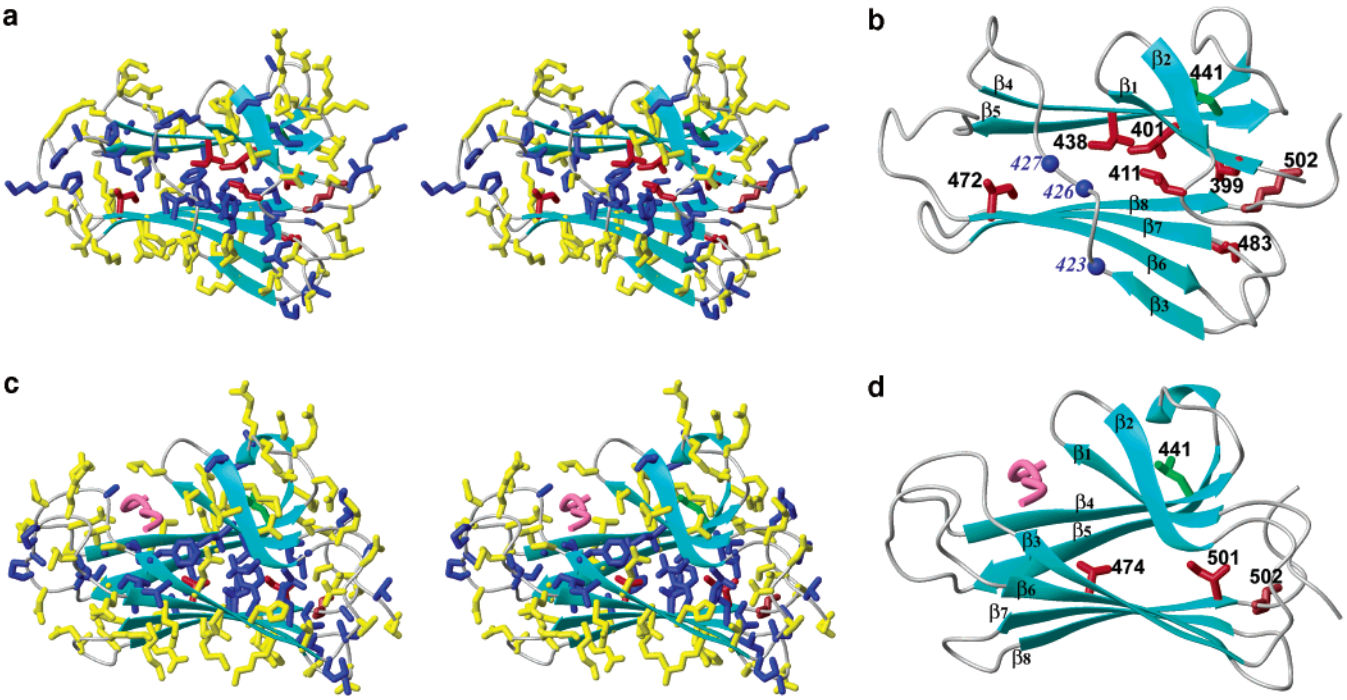

FIGURE 3: Water interaction data for the substrate binding domain of DnaK. (a) Stereo representation of the solution structure of apo-DnaKβ (15) color coded as follows: The apolar residues showing unambiguous NOEs with water (i.e., no labile proton within 5 Å; see text) that are unique to the apo form are colored red (Leu 399, Ile 401, Leu 411, Ile 438, Ile 472, and Ile 483). Apolar residues showing cross-peaks with water, with no labile proton within 5 Å, and that are in common with the peptide-bound form are colored green (Leu 441). All other apolar residues are colored blue. All polar residues, whether they display water interaction or not, are colored yellow, with the exception of Lys 502, which is colored brown (see text). Apolar residues showing cross-peaks with water assigned to relay processes are also colored yellow. (b) Simplified representation of (a) with labeling. (c) Stereo representation of the solution structure of DnaKβ complexed with the peptide NRLLLTG (*17*) color coded as follows: The apolar residues showing NOEs with water that are unique to the peptide-bound form, with no labile proton within 5 Å, are colored red (Val 474 and Ile 501). Other colors are as in (a). The bound peptide is colored pink. (d) Simplified representation of (c) with labeling. The figure was prepared with the program MOLMOL (*50*).

above). This site is shown as a brown-colored residue in Figure 3. It is not interesting from a functional point of view and will not be further discussed.

Figure 2 and Table 1 show that we detect many more water cross-peaks in the apo form than in the peptide-bound form. The first comment to be made about this observation is a technical one: because the bound peptide is not ^{13}C labeled, its Hα resonances resonating close to 4.8 ppm behave spectroscopically as the solvent resonance and cannot be purged in the WMT pulse sequence. Hence, it is to be anticipated that the WMT experiments on the peptide-bound form would also report on some of the known peptide−protein NOEs (*18*) and would thus intrinsically have more NOEs than the peptide-free form, especially in the substrate binding cleft. Since the opposite is found to be the case, it is concluded that the water excitation sequence is selective enough and/or that the relaxation occurring in the spin−echoes is long enough to suppress the peptide−protein NOE effects, except possibly one, as discussed below. It is also necessary to discuss whether the presence of more water transfer cross-peaks for the apo form can be due to the fact that the rotational correlation time of the apo form is longer than that of the peptide-bound form. On one hand, this would cause a more rapid NOE buildup and hence larger NOE cross-peaks at the short mixing times used, but on the other hand, this would reduce the efficiency of the HSQC sequence as well as spectral peak heights which are the main determinant of peak detectability. On balance these effects compensate each other; moreover, the concentration of the apo form sample is less than that of the peptide-bound sample, which would favor more cross-peaks for the latter. It is thus concluded that the NMR experiments indicate that more water is interacting with the apo form than with the peptide-bound form.

One-third of the water interaction peaks in the apo form could not be assigned, reflecting the fact that the assignment list of the protein in this state is not complete due to exchange broadening effects in the backbone resonances. From the WMT peaks that can be assigned, all but one (Lys 502; see above) indicate interaction between water and the protein with a long residence time (>400 ps) or chemical exchange-relay peaks. It turns out that 70% of these assigned cross-peaks belong to Thr, Lys, or Arg side chains mostly at the protein surface, which can be directly removed from the list of interesting NOEs. After this culling process, one is left with only few water-interacting side chains in both states of the protein that need to be further investigated for proximity to exchangeable protons. The most interesting of those occur in the area of the peptide binding cleft in the apo form of the protein. More significantly, there is a dramatic difference in WMT signals in this area between the protein in two forms (see Figure 3).

In the unliganded form of the protein, most methyl groups that uniquely accept WMT (as compared with the peptide-bound form) are located in the binding cleft. Significantly, we do not observe any WMT peaks associated with completely buried apolar groups in the protein (which are colored blue in Figure 3a,c).

The proximity of the (partially) exposed methyls that do report WMT signals to labile protons was investigated from

the coordinates of the NMR structure (*15*). The reported distances below are the heavy atom averages taken over the 20 structures deposited in the Protein Data Bank (1DG4) and include the standard deviation. In the DnaKβ apo form, the Ile 438 methyls are deeply inside the binding cleft. Their nearest labile proton neighbor site is Thr 437 OG at 6.4 $\pm$ 0.3 Å. The nearest labile proton neighbor site of the Leu 399 methyls is Thr 410 OG, at 8.7 $\pm$ 1.2 Å. The nearest labile proton site neighbor of Ile 401 CD is Thr 403 OG, at 5.6 $\pm$ 0.4 Å. The nearest labile proton neighbor site of the Leu 411 methyls is Thr 409 OG, at 7.3 $\pm$ 1.4 Å. It can be confidently stated that these water-magnetization-transfer cross-peaks are due to direct NOEs with captured water(s) because of the existence of internal controls consisting of methyl groups with closer labile neighbors that show no water interaction signals. For instance, the assigned methyls of Val 407 are within 4.1 $\pm$ 0.4 Å of Thr 403 OG and have no water cross-peaks. We conclude that, in the apo form of DnaK, at least one water molecule is captured deeply in the hydrophobic cleft to account for the observed NOEs with residue 438. Likely, at least one more water molecule interacts with residues 399, 401, and 411, which line a separate purely hydrophobic side pocket as described above.

Ile 472, also displaying a WMT signal unique to the apo form, is also at the binding cleft, but it is located at the other side of the dislodged β-strand. The water cross-peak is also undoubtedly due to a direct NOE interaction with water; the nearest labile neighbor site for this residue is Thr 428 OG at 7.0 $\pm$ 1.0 Å. The water molecule(s) interacting with Ile 472 is (are) likely located at the protein surface. If it were buried more deeply, a NOE with Ile 462 should have been observed, which is not the case. Parenthetically, the methyl groups of this Ile 462 are within 4.9 $\pm$ 1.1 Å of Thr 428 OG and yet do not show a NOE, which again proves that the WMT signals discussed so far must all be direct effects.

Ile 483 is located at the right of the domain in Figure 3a and is located in a shallow surface depression; its methyl group is solvent exposed. The nearest labile proton site to the methyl groups is Thr 500 OG at 6.6 $\pm$ 0.8 Å. Its water cross-peak is thus undoubtedly due to a direct NOE interaction. Interestingly, Ile 483 is not in the vicinity of the cleft area of the protein, yet displays only a NOE with water in the apo form. Potentially, this may shed light on the interface site for a possible ligand-linked dimer−monomer equilibrium (see above) or the long-range allosteric conformational changes upon ligand binding in the protein (*15*) or both.

Summarizing, in the apo form of DnaK, the methyl groups of residues 399, 401, 411, 438, 472, and 483 report NOE interactions with water molecules. These residues are indicated and colored red in Figure 3a,b.

Water-magnetization-transfer signals, common to both forms of the protein, are observed for the methyls of Ile 412, Ala 413, and Leu 441 (see Table 1). Ile 412 and Ala 413 are close to the surface of the protein and are not part of the hydrophobic side pocket lined by residues 399, 401, and 411. Both side chains are in the range 3.9−4.6 Å of Thr 416 OG in either or both of the structures; the water transfer signals are thus most likely to result from a relay process, and are of little interest. (Distances for the peptide-bound form are taken from the crystal structure, code 1DKZ.) Residue Leu 441 is also located at the protein surface but has as the closest labile proton site Ser 453 OG at a distance of 5.6 $\pm$ 0.9 Å (apo) and 5.3 Å (peptide bound). We thus propose that a water molecule interacts with this residue in both states of the molecule. The latter is colored green in Figure 3. Because this site does not change upon substrate binding, it is less interesting with respect to the function of the protein.

In the peptide-bound form, two methyl water peaks are observed that are unique to this form of the protein (Val 474 and Ile 501). Val 474 is very deeply buried in the substrate binding cleft and does not have any close labile proton neighbors (Thr 475 OG is closest at 6.2 Å). Val 474, however, does display a strong NOE with the side chain of the third leucine of the bound peptide (*18*). It may be possible that this NOE can be carried by spin diffusion from its α proton, as it is not saturated in our pulse sequence. We thus tentatively conclude that there is no water bound in the active site cleft of the peptide-bound form of DnaKβ. Nevertheless, Val 474 is highlighted in red in Figure 3c,d.

Ile 501 is at the C-terminus and at the right of the domain as shown in Figure 3c,d. The closest labile proton site is Ser 505 OG at 5.7 Å. It thus appears that this water transfer signal corresponds to a captured water NOE. Interestingly, this site is not very far away from one of the only three internal water molecules, coordinated by the backbone of residues 504 and 505, observed in the X-ray structure of DnaK (*14*). The other two internal water molecules in the X-ray structure of the substrate binding domain of the peptide-bound form of DnaK are coordinated by the backbone of residues 394−396 (*14*). These water molecules were not observed in the current NMR studies, presumably because the N-terminal part of the molecules is of different genetic construction. Significantly, there are no internal water molecules in the substrate−cleft area of the peptide-bound X-ray structure, corresponding with our observations that (most of) the water molecules are displaced by peptide.

Summarizing, the refined water-magnetization-transfer NMR methods described herein are found to be very selective and detect only very few water molecules that interact with or are trapped by either form of the substrate binding domain of the Hsp70 chaperone of *E. coli* DnaKβ. Remarkably, most of the trapped water molecules are located in the functionally important area of the protein, the hydrophobic substrate binding cleft, and are displaced upon ligand binding.

The Nature of the Bound Water. According to the classical hydrophobic model (*5*), the water contained in the hydrophobic cleft of apo-DnaKβ could be clathrated water, and its release upon binding of the hydrophobic ligand would result in a gain of solvent entropy and a net decrease in free energy. However, classical clathrate formation, consisting of waters with four hydrogen bonds each arranged in linked pentagons, is found to occur only around hydrophobic convex surface patches in molecular dynamics calculations [i.e., protrusions (*39*)]. In agreement with this, classical clathrate is not often observed in high-resolution protein structures which contain few exposed hydrophobic protrusions. It was detected in a crystal structure of β-galactosidase, but there it was found to encapsulate an electron density of unknown identity, likely one of the DMSO cosolvent molecules (*40*). Contrary to solvating hydrophobic surfaces, clathrate-like structures have been inferred around the hydrophilic hydroxyproline residues in the collagen triple helix (*41*).

Hydration of flat hydrophobic surfaces is very different (*39*): this situation is more dynamic, and individual water

molecules fluctuate evenly between clathrate forms and "inverted clathrate" forms, where one of the water hydrogen bonds is unfulfilled and points directly toward the hydrophobic surface. Release of such hydration should thus also give rise to a net free energy gain, here with a more enthalpic component. Water binding in (mostly) hydrophobic cavities presents a further deviation from the classical hydrophobic interaction: in a 200 ps molecular dynamics study of water binding to the hydrophobic substrate binding cavity of chymotrypsin, Rossky and co-workers observed that the hydrogen-bonding valence for the captured solvent was reduced to less than 50% of the bulk solvent value (*42*). This loss is partially, but not fully, compensated for by several hydrogen bonds of the encapsulated solvent with polar anchor points in the otherwise apolar surface of the cavity. The release of such bound water would thus result in a large gain in solvent enthalpy.

The hydrophobic substrate binding cleft of DnaKβ has a complex concave topology. Several polar anchor groups are available in form of the backbone atoms of the displaced β-strand that runs through the cleft and of several unfulfilled hydrogen bonds of the backbone of residues Ile 438 and Thr 437. According to the models discussed above, the water present in this cleft may interact with these polar groups, and its observed release upon ligand binding should result in mixed enthalpic and entropic contributions. The water−NOE data also reveal the presence of water in a purely hydrophobic side cavity. According to the model studies described above these molecules likely have a dearth of hydrogen bonding, and their release upon ligand binding should result in a gain of mostly enthalpy. Water trapped in purely hydrophobic cavities has been observed before for interleukin-1β (*22*, *43*) and for lysozyme (*44*).

The water molecules bound in the hydrophobic substrate binding cleft are displaced upon ligand binding. This process can in principle also affect the kinetics of ligand binding to DnaK, which is reported (*45*, *46*) to occur slower than the diffusion-controlled on-rate (10^3 and 5×10^5 M^{-1} s^{-1} for this peptide for the ADP and ATP states of DnaK, respectively). The upper limit of the water dissociation rate of $^1/_{400}$ ps (2.5×10^9 s^{-1}) as follows from the NMR data is much faster than the pseudo-first-order substrate association rate, even at very high concentrations of 1 mM substrate (1 and 5×10^2 s^{-1} for ADP and ATP states). Hence, the NMR experiments cannot determine if the kinetics of ligand binding is affected by the water release.

The residence time of the bound water in the substrate binding cleft and associated hydrophobic pocket is longer than 400 ps. This finding is in general agreement with results of other studies showing that internal water in such cavities is long-lived. In particular, the water molecules bound to the hydrophobic cavity of chymotrypsin did not change position over the duration of the 200 ps molecular dynamics simulation (*42*). Using magnetic relaxation dispersion methods to study the intestinal fatty acid binding protein, Halle and co-workers observed residence times of at least 1 ns for water molecules assigned to locate in the apolar cavity of this molecule (*47*, *48*). In NMR NOE studies these water molecules would give rise to macromolecular NOEs. Otting and Wüthrich concluded from their pioneering NMR studies that water molecules bound to the surface of BPTI have residence times shorter than 200 ps (*20*). Short residence times for individual surface waters were predicted by a follow-up computational molecular dynamics study of the same protein (*49*). In contrast, most of the (very few) surface waters detected in our study of both forms of DnaKβ that give rise to unambiguous NOEs have residence times longer than 400 ps. Another NMR group has reported similar findings for lysozyme (*23*): waters bound to the surface methyl groups of this protein also have long residence times. The discrepancy between these varied observations can be understood if the intrinsic flexibility of the exposed side chains is different for the proteins mentioned.

ACKNOWLEDGMENT

We thank Drs. Lila Gierasch (University of Massachusetts) and Maurizio Pellecchia (currently at the Burnham Institute) for the procedures to prepare the DnaKβ protein domain.

REFERENCES

1. Yamano, A., Heo, N. H., and Teeter, M. M. (1997) *J. Biol. Chem. 272*, 9597−9600.
2. García de la Torre, J., Huertas H. L., and Carrasco B. (2000) *Biophys. J. 78*, 719−730.
3. Hendrick, J. P., Langer, T., Davis, T. A., Hartl, F. U., and Wiedmann, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10216−10220.
4. Craig, E. A., Gambill, B. D., and Nelson, R. J. (1993) *Microbiol. Rev. 57*, 402−414.
5. Martin, J., and Hartl, F. U. (1997) *Curr. Opin. Struct. Biol. 7*, 41−52.
6. Bukau, B., and Horwich, A. L. (1998) *Cell 92*, 351−366.
7. Wickner, S., Maurizi, M., and Gottesman, S. (1999) *Science 286*, 1888−1893.
8. Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B., and Hartl, F. U. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10345−10349.
9. Mayer, M. P., Rudiger, S., and Bukau, B. (2000) *Biol. Chem. 381*, 877−885.
10. Hubbard, T. J., and Sander, C. (1991) *Protein Eng. 4*, 711−717.
11. Slepenkov S. V., and Witt, S. N. (2002) *Mol. Microbiol. 45*, 1197−1206.
12. Hartl, F. U. (1996) *Nature 381*, 571−580.
13. Flaherty, K. M., Deluca-Flaherty, C., and McKay, D. B. (1990) *Nature 346*, 623−628.
14. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) *Science 272*, 1606−1614.
15. Pellecchia, M., Montgomery, D. L., Stevens, S. Y., Vander Kooi, C. W., Feng, H., Gierasch, L. M., and Zuiderweg, E. R. P. (2000) *Nat. Struct. Biol. 7*, 298−303.
16. Wang, H., Kurochkin, A. V., Pang, Y., Hu, W., Flynn, G. C., and Zuiderweg, E. R. P. (1998) *Biochemistry 37*, 7929−7940.
17. Morshauser, R. C., Hu, W., Wang, H., Pang, Y., Flynn, G. C., and Zuiderweg, E. R. P. (1999) *J. Mol. Biol. 289*, 1387−1403.
18. Stevens, S. Y., Cai, S., Pellecchia, M., and Zuiderweg, E. R. P. (2003) *Protein Sci.* (in press).
19. Blokzijl, W., and Engberts, J. B. F. N. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 1545−1579.
20. Otting, G., Liepinsh, E., and Wüthrich, K. (1991) *Science 254*, 974−980.
21. Liepinsh, E., Rink, H., Otting, G., and Wüthrich, K. (1993) *J. Biomol. NMR 3*, 253−257.
22. Ernst, J. A., Clubb, R. T., Zhou, H. X., Gronenborn, A. M., and Clore, G. M. (1995) *Science 267*, 1813−1817.
23. Melacini, G., Kaptein, R., and Boelens, R. (1999) *J. Magn. Reson. 136*, 214−218.
24. Phan, A. T., Leroy, J.-L., and Guéron, M. (1999) *J. Mol. Biol. 286*, 505−519.
25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
26. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
27. Otting, G. (1997) *Prog. Nucl. Magn. Reson. Spectrosc. 31*, 259−285.

28. Gemmecker, G. Jahnke, W., and Kessler, H. (1993) *J. Am. Chem. Soc. 115*, 11620−11621.
29. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR 3*, 627−638.
30. Mori, S., Abeygunawardana, C., Van Zijl, P. C. M., and Berg, J. (1996) *J. Magn. Reson. B110*, 96−101.
31. Mori, S., Berg, J. M., and Van Zijl, P. C. M. (1996) *J. Biomol. NMR 7*, 77−81.
32. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
33. Olejniczak, E. T., Gampe, R. T., and Fesik, S. W. (1986) *J. Magn. Reson. 67*, 28−41.
34. Massefski, W., and Redfield, A. (1988) *J. Magn. Reson. 78*, 150−155.
35. Fejzo, J., Westler, W. M., Macura, S., and Markley, J. L. (1991) *J. Magn. Res. 92*, 195−203.
36. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. M. (1995) *J. Magn. Reson. B108*, 94−98.
37. Kay, L., Keifer, E. P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
38. Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
39. Cheng, Y. K., and Rossky, P. J. (1998) *Nature 392*, 696−699.
40. Juers, D. H., Jacobson, R. H., Wigley, D., Zhang, X. J., Huber, R. E., Tronrud, D. E., and Matthews, B. W. (2000) *Protein Sci. 9*, 1685−1699.
41. Melacini, G., Bonvin, A. M., Goodman, M., Boelens, R., and Kaptein, R. (2000) *J. Mol. Biol. 300*, 1041−1049.
42. Carey, C., Cheng, Y. K., and Rossky, P. J. (2000) *Chem. Phys. 258*, 415−425.
43. Yu, B., Blaber, M., Gronenborn, A. M., Clore, G. M., and Caspar, D. L. D. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 103−108.
44. Otting, G., Liepinsh, E., Halle, B., and Frey, U. (1997) *Nat. Struct. Biol. 4*, 396−404.
45. Buczynski, G., Slepenkov, S. V., Sehorn, M. G., and Witt, S. N. (2001) *J. Biol. Chem. 276*, 27231−27236.
46. Slepenkov, S. V., and Witt, S. N. (2002) *Biochemistry 41*, 12224−12235.
47. Scapin, G., Gordon, J. I., and Sacchettini, J. C. (1992) *J. Biol. Chem. 267*, 4253−4269.
48. Wiesner, S., Kurian, E., Prendergast, F. G., and Halle, B. (1999) *J. Mol. Biol. 286*, 233−246.
49. Brunne, R. M., Liepinsh, E., Otting, G, Wüthrich, K., and Van Gunsteren, W. F. (1993) *J. Mol. Biol. 231*, 1040−1048.
50. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI030097C