J. Mol. Biol. (1996) **260**, 236–250

JMB

NMR Structure of the J-domain and the Gly/Phe-rich Region of the *Escherichia coli* DnaJ Chaperone

Maurizio Pellecchia[1], Thomas Szyperski[1], Daniel Wall[2] Costa Georgopoulos[2] and Kurt Wüthrich[1]*

[1]*Institut für Molekularbiolgie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg CH-8093, Zürich Switzerland*

[2]*Département de Biochimie Médicale, Centre Médical Universitaire, 1, rue Michel-Servet, CH-1221 Genève 4, Switzerland*

**Corresponding author*

The recombinant N-terminal 107-amino acid polypeptide fragment 2–108 of the DnaJ molecular chaperone of *Escherichia coli*, which contains the J-domain (residues 2 to 76) and the Gly/Phe-rich region (residues 77 to 108), was uniformly labeled with nitrogen-15 and carbon-13. The complete NMR solution structure of the J-domain was determined with the program DIANA on the basis of 682 nuclear Overhauser enhancement (NOE) upper distance limits and 180 dihedral angle constraints. It contains three well-defined helices comprising residues 6 to 10, 18 to 32 and 41 to 57, and a fourth helix, consisting of residues 61 to 68, which is well defined as a regular secondary structure but for which the location relative to the remainder of the molecule is not precisely determined. The helices II and III form an antiparallel helical coiled-coil. Helix I is approximately parallel to the plane defined by the helices II and III and runs from the carboxy-terminal end of the helix III to the center of helix II. Helix IV is positioned near the carboxy-terminal end of helix III and is on the same side of the coiled coil as helix I, but it is oriented approximately perpendicular to the plane of the helices II and III. This novel α-protein topology leads to formation of a hydrophobic core involving side-chains of all four helices. A strong correlation is seen between the extent of sequence-conservation of hydrophobic residues in the family of J-domain homologues, and the structural organization of the hydrophobic core in these proteins. The residues which have key roles for the specificity of the interaction of DnaJ-like proteins with their corresponding Hsp70 counterparts are located on the outer surfaces of the helices II and III, and in the loop connecting these two helices. Measurements of backbone amide proton exchange rates, ^{15}N spin relaxation times and heteronuclear ^{15}N {^{1}H} NOEs provided additional insights into local conformational equilibria and internal rate processes in DnaJ(2–108). In the Gly/Phe-rich region, which is poorly ordered in the NMR solution structure and does not form a globular core, the polypeptide segment 90 to 103 differs from the segments 77 to 89 and 104 to 108 by reduced local flexibility. Considering that this same segment shows sequence conservation with corresponding segments in the Gly/Phe-rich regions of other DnaJ-like proteins, its

Present address: D. Wall, Department of Biochemistry, Stanford University Medical Center, Stanford, CA 94305, USA.

Abbreviations used: DnaJ/DnaK/GrpE, proteins of the Hsp70 chaperone complex from *E. coli*; ATP, adenosine-5'-triphosphate; Hsp, heat shock protein; DnaJ(2–108), N-terminal 107-residue polypeptide fragment from DnaJ; J-domain, the N-terminal 76-residue polypeptide segment of DnaJ; Gly/Phe-rich region, polypeptide segment 77 to 108 of DnaJ; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; 2D, 3D, 2-, 3-dimensional; NOESY, nuclear Overhauser enhancement spectroscopy; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; MEXICO, measurement of fast proton exchange rates in isotopically labeled compounds; *ct*, constant-time; ppm, parts per million; $^{3}J_{HN\alpha}$, vicinal spin–spin coupling constant between the backbone amide proton and the α proton; $^{3}J_{N\beta}$, vicinal spin–spin coupling constant between the backbone amide nitrogen and one of the β protons; $^{3}J_{\alpha\beta}$, vicinal spin-spin coupling constant between the α proton and one of the β protons; r.m.s.d., root-mean-square deviation; T_1, longitudinal relaxation time; T_2, transverse relaxation time; $T_{1\rho}$, relaxation time in the rotating frame; REDAC, use of redundant dihedral angle constraints.

0022–2836/96/270236–15 $18.00/0

reduced flexibility may be directly linked to the formation of the ternary DnaJ-DnaK-polypeptide complex.

Keywords: DnaJ; molecular chaperone; heat shock protein; NMR structure; protein dynamics

Introduction

The DnaK, DnaJ and GrpE molecular chaperone machine participates in a variety of biological processes such as protein folding, protein disaggregation, survival at high temperature, negative regulation of the heat shock response, modulation of *in vivo* proteolysis rates, and replication of several plasmids (reviewed by Georgopoulos *et al*., 1990; Georgopoulos & Welch, 1993). Central to the mechanism of action of this chaperone machine is that DnaJ stimulates the hydrolysis of DnaK-bound ATP (Liberek *et al*., 1991) which, in turn, regulates the affinity of DnaK for substrates such as σ^{32} (Liberek & Georgopoulos, 1993). Furthermore, Kimura *et al*. (1995) have recently shown that Ydj1, a eukaryotic DnaJ-like protein, interacts with both Hsp70 and Hsp90 in establishing several signal pathways, presumably by facilitating critical changes in protein conformation/oligomerization that are required for signal transduction.

In the full length *Escherichia coli* DnaJ protein there are four domains formed by the 376-amino acid sequence. An amino-terminal 108-residue fragment of DnaJ, represented in this study by DnaJ(2–108), is sufficient to stimulate the ATPase activity of DnaK, regulate the conformational state and substrate binding of DnaK in the presence of ATP, support λ phage replication as well as down-regulate the heat shock response *in vivo* (Liberek *et al*., 1995; Wall *et al*., 1994, 1995). However, this fragment does not bind polypeptide substrates. The amino-terminal 76 residues of DnaJ form an evolutionarily highly conserved motif, the ''J-domain'', which basically defines the family of DnaJ-like proteins (Bork *et al*., 1992; Silver & Way, 1993). The adjoining Gly/Phe-rich region, comprising residues 77 to 108, contains a large number of glycyl and phenylalanyl residues. Deletion of residues 77 to 108 leads to a mutant protein that still binds to DnaK and stimulates its ATPase activity at wild-type levels, but is defective in activating DnaK to bind σ^{32}. This indicates that the Gly/Phe-rich region has an essential role in modulating the substrate-binding activity of DnaK (Wall *et al*., 1995). Carboxy-terminal to these modules, there are a ''zinc-binding'' and a ''low homology'' region, which are also present in many other DnaJ-like proteins.

To obtain further insight into the structural basis of DnaJ functions, we determined the three-dimensional structure of DnaJ(2–108) by NMR spectroscopy. Preliminary data showed that the J-domain consists of four α-helices, where the global fold is characterized by an antiparallel arrangement of the helices II and III, with the helices I and IV oriented approximately perpendicular to II and III (Szyperski *et al*., 1994a). The same fold was also reported for a different construct, DnaJ(1–78) (Hill *et al*., 1995), which represents the complete J-domain but is not biochemically active. In contrast to the J-domain, the Gly/Phe-rich region was found to be flexibly disordered in solution (Szyperski *et al*., 1994a). Here we describe the complete, high-quality NMR solution structure determination of the J-domain supplemented by studies of its internal mobility through measurements of the ^{15}N spin relaxation times. In view of the important role of the Gly-Phe-rich region for the biological function of DnaJ, we further include a characterization of the structural and dynamic properties of residues 77 to 108 by measurements of the amide proton exchange rates and ^{15}N spin relaxation parameters.

Results

Resonance assignments

Sequence-specific NMR assignments for DnaJ(2–108) were obtained based in part on establishing sequential NOE connectivities (Billeter *et al*., 1982; Wagner & Wüthrich, 1982; Wüthrich, 1986) with 3D ^{15}N-resolved [^1H, ^1H] NOESY (Fesik & Zuiderweg, 1988; Messerle *et al*., 1989) and 3D ^{15}N-resolved [^1H, ^1H] TOCSY (Fesik & Zuiderweg, 1988), and in part on heteronuclear scalar coupling connectivities established using 3D $\underline{H}^{\alpha/\beta}\underline{C}^{\alpha/\beta}$(CO)NHN (Szyperski *et al*., 1994b) and 3D CBCANHN (Grzesiek & Bax, 1992) spectra. Combination of the results from the two approaches yielded nearly complete assignments for the back bone $^1H^N$ and ^{15}N resonances and the αCH-βCH_n fragments. The only missing backbone assignments are those of the NH moieties of the N-terminal Ala2, and of Lys3, and all resonances of the five glycyl residues 81, 82, 86, 104 and 105, which could not be assigned because of degeneracy of the ^1H, ^{15}N and $^{13}C^{\alpha}$ chemical shifts. The chemical shifts of the αCH-βCH_n fragments provided the starting point for the ^1H and ^{13}C assignments of the aliphatic side-chains with 3D HCCH-TOCSY (Bax *et al*., 1990) and 3D HCCH-COSY (Kay *et al*., 1990). Complete assignments for the aliphatic CH_n moieties were obtained, except that Arg22 $\delta^{13}C$, Arg27 $\delta^{13}C$, Arg63 $\gamma^{13}C$, Arg108 $\delta^{13}C$, and all ^{13}C resonances and the γ- and δ-proton resonances of Arg106 could not be identified due to spectral overlap. Among the labile side-chain

protons, the amide groups of all Asn and Gln, and the Arg108 ϵ-proton resonance were assigned. Stereospecific assignments were obtained for 13 out of 35 βCH_2 groups with non-degenerate β-proton shifts, and for the isopropyl methyls of Val12, Val56 and Leu57. The ^1H spin systems of the aromatic rings of the tyrosyl and phenylalanyl residues in the J-domain were identified using 2D $^{13}C(\omega_2)$-half-filtered [^1H, ^1H] TOCSY (Otting & Wüthrich, 1990). Subsequently, these resonances were sequentially assigned *via* NOEs between the aromatic ring and the β- and/or α-protons (Wüthrich, 1986), using 3D ^{13}C-resolved [^1H, ^1H] NOESY (Ikura *et al*., 1990) and 2D $^{13}C(\omega_2)$-half-filtered [^1H, ^1H] NOESY (Otting & Wüthrich, 1990). Likewise, the δ-protons of the phenylalanyl residues in the Gly/Phe-rich region (residues 77 to 108) were assigned using 3D ^{13}C-resolved [^1H, ^1H] NOESY. Since all these resonances have virtually identical chemical shifts, assignments of the corresponding ϵ and ζ-protons could not be obtained. The ^{13}C assignments for the tyrosyl rings were obtained from the proton chemical shifts using a 2D *ct*-[^{13}C, ^1H] COSY spectrum (Vuister & Bax, 1992). The ^{13}C resonances of the phenylalanyl residues in the Gly/Phe-rich region could not be assigned because of spectral overlap. The aromatic resonances of His33 and His71 were identified using 2D *ct*-[^{13}C, ^1H] COSY, but sequential assignments were not obtained because the NOEs between the rings and the αCH-βCH_2 moieties could not be detected. This may be rationalized on the basis of the NMR solution structure (see below): both histidine imidazole moieties are solvent-exposed and flexibly disordered. Overall, the analysis of the two and three-dimensional spectra for resonances of the structured J-domain was not severely affected by the presence of the more flexible Gly/Phe-rich region. Most of the resonance lines of residues 77 to 108 are not as sharp and intense as one would expect for a fully flexible random coil polypeptide segment, so that there were no pronounced truncation or cancellation artifacts in the NMR spectra.

DnaJ(2–108) exhibits only very limited proton chemical shift dispersion, as is typically observed for α-helical proteins. Because of this and the fact that the Gly-Phe-rich region is predominantly disordered, $^{13}C/^{15}N$ labeling was mandatory for obtaining complete assignments. Outstanding proton chemical shifts are those of H^α of Lys14 and Ala53, which deviate by more than 1.0 ppm from their random coil values. The calculation of ring current effects arising from aromatic side-chains (Perkins & Wüthrich, 1979) showed that these large shifts originate almost entirely from close spatial proximity of the two α-protons to the rings of Tyr7 and Tyr6, respectively.

Collection of conformational constraints

As outlined by Szyperski *et al*. (1994a) the Gly-Phe-rich region is disordered in the NMR solution structure, and only intraresidual and sequential NOEs were observed for this polypeptide segment. Therefore, we restricted the structure calculation to the J-domain of residues 2 to 76. Within the J-domain, 548 NOEs from ^{15}N-resolved [^1H, ^1H] NOESY (100 ms mixing time) and 1142 NOEs from ^{13}C-resolved [^1H, ^1H] NOESY (70 ms mixing time) were assigned. In addition 55 $^3J_{HN\alpha}$ coupling constants were extracted by inverse Fourier transformation of in-phase multiplets from a 2D [^{15}N, ^1H] COSY spectrum (Szyperski *et al*., 1992), 14 $^3J_{N\beta}$ coupling constants were estimated from 3D *ct*-HNNHB (Archer *et al*., 1991; Chary *et al*., 1991), and 16 $^3J_{\alpha\beta}$ coupling constants were estimated from 3D ^{15}N-resolved [^1H, ^1H] TOCSY (50 ms mixing time) for residues with $^3J_{HN\alpha} < 6.0$ Hz

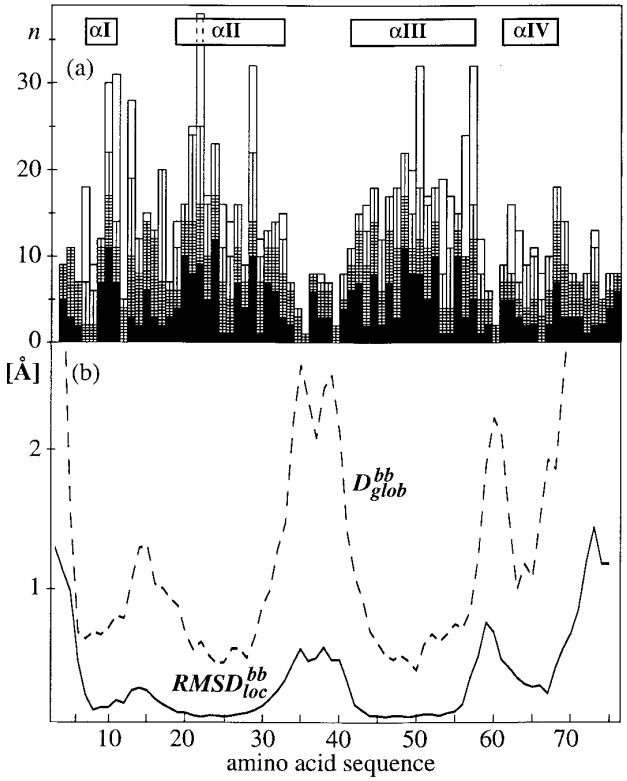

Figure 1. (a) Number and types of NOE constraints per residue used in the calculations of the solution structure of the J-domain (residues 2 to 76) *versus* the amino acid sequence. The following code is used to define the different NOE constraints: black, intraresidual; cross-hatched, sequential; vertically hatched, medium-range; white, long-range. (b) Plots *versus* the amino acid sequence of the mean of the global backbone displacements per residue, D^{bb}_{glob}, of the 20 energy-minimized DIANA conformers relative to the mean NMR structure calculated after superposition of the backbone heavy atoms N, C^α and C' of the helical residues for minimal r.m.s.d. (broken line), and the mean local r.m.s. deviation, $RMSD^{bb}_{loc}$, for the backbone superposition of all tripeptide segments along the sequence relative to the mean NMR structure (continuous line). The r.m.s.d. values for the tripeptide segments are plotted at the position of the central residue. The locations of the α-helices are indicated at the top of the Figure.

Table 1. Analysis of the 20 best DIANA conformers of the J-domain (residue 2 to 76) in the polypeptide fragment 2–108 from *E. coli* DnaJ after restrained energy minimization with the program OPAL

Quantity	Average value ± standard deviation	Range
DIANA target function (Å²)[a]	0.90 ± 0.23	(0.45. . .1.27)
AMBER energy (kcal/mol)	–2876.65 ± 95.48	(–3010.13. . . –2668.16)
Residual NOE distance constraint violations:		
Sum (Å)	7.10 ± 0.59	(6.22. . .8.80)
Maximum (Å)	0.10 ± 0.00	(0.09. . .0.11)
Residual dihedral angle constraint violations:		
Sum (°)	24.05 ± 3.32	(18.90. . .29.80)
Maximum (°)	2.02 ± 0.23	(1.59. . .2.47)

[a] Before energy minimization.

(Clore *et al*., 1991). The resulting input for the final DIANA structure calculation contained 682 NOE upper distance limits (310 intraresidual, 80 sequential backbone–backbone, 34 medium-range and long-range backbone–backbone, 258 interresidual with side-chain protons) and 180 dihedral angle constraints. Figure 1(a) shows the sequence distribution of the NOEs observed for the J-domain. Although a rather uniform distribution of long-range NOE constraints was observed for the helices I, II and III, only a few long-range constraints were obtained for helix IV. In addition, no long-range NOEs were observed either for the loop comprising residues 34 to 40, or for the N-terminal and C-terminal residues of the J-domain.

The NMR solution structure of the J-domain

The tertiary fold of the J-domain (residues 2 to 76) is essentially that described previously by Szyperski *et al*. (1994a), except that the lengths of the helical secondary structures are slightly different. The relative orientation of the first three α-helices, comprising residues 6 to 10, 18 to 32 and 41 to 57, is well-defined, but for the fourth helix, comprising residues 61 to 68, the spatial orientation with respect to the remainder of the molecule could not be precisely determined (Figure 1(b)). The helices II and III form an antiparallel helical coiled-coil. Helix I is approximately parallel to a plane containing the helices II and III and runs from the carboxy-terminal end of helix III toward the center of helix II. Helix IV near the carboxy-terminal end of helix III is on the same side of the helices II and III, as helix I, but is oriented approximately perpendicular to the plane containing the helices II and III (Figure 3). Helices I and II are connected by a tight turn; the amide proton of Ala16 exchanges slowly and a hydrogen bond between this amide group and the carbonyl oxygen of Ser13 has been identified (see below), but the structural determination for this polypeptide segment is not sufficiently precise (Figure 1(b)) to unambiguously identify the type of the turn. Similarly, the loop of residues 34 to 40, which connects the helices II and III, and the turn of residues 58 to 60, which links the helices III and IV, are not defined with high precision (Figure 1(b)).

An overview of the results of the structure calculation, after energy minimization of the DIANA conformers, is provided by Table 1. The small size and number of residual constraint violations show that the input data represent a self-consistent set, and that the constraints are well satisfied in the calculated conformers. The mean of the local backbone displacements (Figure 1(b)), as well as a plot of the ϕ and ψ-angles observed in the 20 energy-refined DIANA conformers *versus* the sequence (Figure 2), shows that all four helices of the J-domain are locally well defined in the NMR structure. However, calculation of the r.m.s.d.

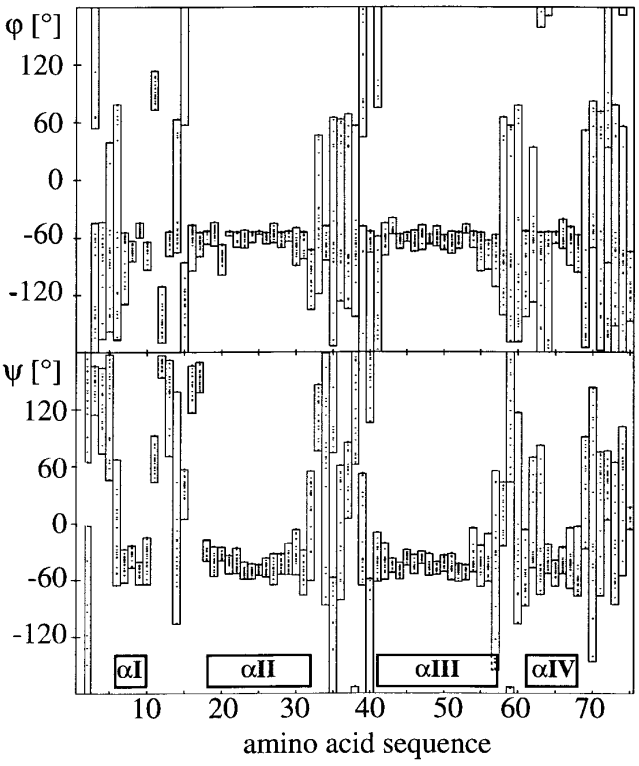

Figure 2. Plots of the values of the dihedral angles φ and ψ *versus* the amino acid sequence of the J-domain. The dots represent the φ and ψ values in the individual energy-minimized DIANA conformers. The boxes are drawn to cover the angular range that contains all the values of the 20 conformers that are used to represent the solution structure. The locations of the four helices are indicated at the bottom.

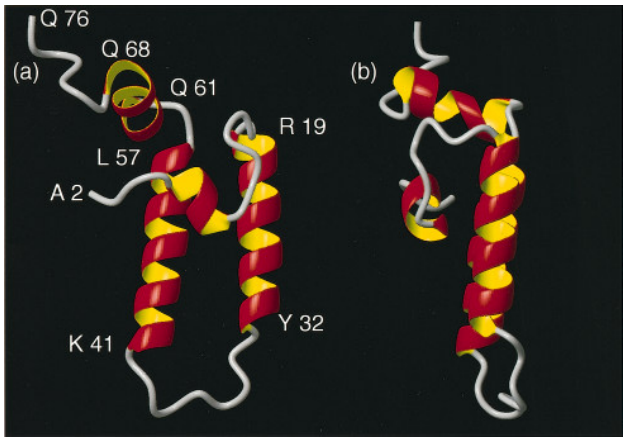

Figure 3. Two views of a schematic drawing of the polypeptide segment comprising the residues 2 to 76 of the DIANA conformer with the lowest residual target function value. The four α-helices are drawn as red and yellow ribbons, and the intervening turns and loops as grey tubes. The same structure is shown in (a) and (b); the two views differ by a rotation of approximately 90° about a vertical axis in the projection plane. In (a) some α-carbon positions are identified by the one-letter amino acid code and the sequence number. The Figure was generated with the program MOLMOL (Koradi *et al*., 1996).

values for the different atom selections (Table 2) shows that the location of helix IV is globally significantly less well defined when compared with the first three helices. Technically, the increased global disorder for helix IV is due to the fact that we could observe long-range NOEs only for the three residues Lys62, Arg63 and Tyr66 in this helix (Figure 1(a)). This scarcity of long-range NOEs is due to the particular packing geometry of helix IV with respect to the remainder of the molecule, as is evidenced by the fact that in the final structure there are no distances shorter than 4.0 Å with hydrogen atoms of helix IV, for which the corresponding NOEs were not observed. Furthermore, increased local disorder (Figures 1(b) and 2) is observed for all turns and loops connecting the helices, and for the N and C-terminal polypeptide segments.

Inspection of the dihedral angle distribution (Figure 2) and the hydrogen bonds (Table 3) observed for the J-domain reveals that Asp5, Glu17 and Asp40 form N-caps (Richardson & Richardson, 1988) of helices I, II and III, respectively. For Asp5 the side-chain hydrogen bond is formed with the backbone amide proton of the sequentially following residue, which is rarely observed for longer helices (Richardson & Richardson, 1988). The carboxyl oxygen of Glu17 is hydrogen bonded with the backbone amide proton of Arg19, thus interacting with the second residue of helix II. A remarkable conformation is observed for the N-cap of helix III, Asp40, in that one of its side-chain carboxyl oxygen atoms forms a bifurcated hydrogen bond with the backbone amide protons of Lys41 and Ala43 (Figure 4). The NMR structure determination leads to a static picture; therefore there remains the open question whether the side-chain carboxylate of Asp40 might actually form a ''time-shared'' hydrogen bond (Szyperski *et al*., 1994c) with the amide protons of Lys41 and Ala43. Similarly, the side-chain carboxylates of Asp5 as well as Asp59, which is properly located to act as the N-cap of helix IV, might form time-shared hydrogen bonds with the amide protons of Tyr7 and Glu8, and Gln61 and Lys62, respectively. Unfortunately, recording of 2D [^{15}N, ^{1}H] COSY spectra at different pH values revealed that DnaJ(2–108) is highly sensitive to acid denaturation and irreversibly unfolds at pH values below 4.5. It was thus not possible to further document these putative time-shared hydrogen bonds by pH-titration experiments.

The antiparallel arrangement of the helices II and III resembles a two-stranded α-helical coiled-coil. The angle between the axes of two cylinders representing the helices II and III is about 20° (Figure 3), which is typical of a super-coil in which the hydrophobic contact surface between the two

Table 2. Average r.m.s.d. values calculated for different atom selections in the solution structure of the J-domain (residues 2 to 76) in the polypeptide fragment 2–108 from *E. coli* DnaJ

Atoms used for the comparison[a]	r.m.s.d. (Å) ± standard deviation[b]
Backbone atoms N, C^{α}, C' (2-76)	2.16 ± 0.43
All heavy atoms (2-76)	2.80 ± 0.38
Backbone atoms (6-10; 18-32; 41-57; 61-68)	1.00 ± 0.19
Backbone atoms (6-10; 18-32; 41-57)	0.65 ± 0.13
Same + best-defined side-chains[c]	0.79 ± 0.10
Same + hydrophobic core side-chains[d]	1.10 ± 0.18
Backbone atoms (6-57)	0.99 ± 0.22

[a] The numbers in parentheses denotes the protein segments considered in the comparison.
[b] Averages and standard deviations are given for the pairwise r.m.s.d. values between each of the energy-refined DIANA conformers and the mean structure.
[c] Best-defined side-chains are those with global displacements smaller than 1.6 Å and include the following residues: 9, 10, 12, 21, 24, 28, 29, 43, 47, 50 53, 54, 56, 57.
[d] Includes the backbone atoms N, C^{α} and C' of residues 6 to 10, 18 to 32 and 41 to 57 and the side-chain heavy atoms of the 18 residues forming the hydrophobic core of the molecule, i.e. residues 6, 7, 9, 10, 12, 21, 24, 25, 28, 29, 32 ,43, 47, 50, 53, 54, 56, 57.

Table 3. Hydrogen bonds identified in the 20 energy-minimized DIANA conformers of the J-domain (residues 2 to 76) in the fragment 2–108 from *E. coli* DnaJ

Donor[a]	Acceptor[a]	Number of conformers[b]
Tyr6 NH	Asp5 O^{δ}	9
Leu10 NH	Tyr6 O'	11
Gly11 NH	Tyr7 O'	19
Val12 NH	Tyr7 O'	9
Ala16 NH	Ser13 O'	11
Arg19 NH	Glu17 O^{ϵ}	9
Glu20 NH*	Glu17 O'	18
Ile21 NH	Glu17 O'	18
Arg22 NH	Glu18 O'	20
Ala24 NH	Glu20 O'	20
Tyr25 NH	Ile21 O'	20
Lyr26 NH	Arg22 O'	18
Arg27 NH	Lys23 O'	20
Leu28 NH	Ala24 O'	20
Ala29 NH	Tyr25 O'	20
Met30 NH*	Lys26 O'	16
Lys31 NH*	Arg27 O'	11
Tyr32 NH*	Leu28 O'	13
Lys41 NH*	Asp40 O^{δ}	14
Ala43 NH*	Asp40 O^{δ}	11
Ala45 NH*	Lys41 O'	15
Lys46 NH*	Glu42 O'	20
Phe47 NH*	Ala43 O'	20
Lys48 NH*	Glu44 O'	18
Glu49 NH	Ala45 O'	17
Ile50 NH	Lys46 O'	18
Lys51 NH	Phe47 O'	20
Glu52 NH	Lys48 O'	20
Ala53 NH	Glu49 O'	20
Tyr54 NH	Ile50 O'	19
Glu55 NH	Lys51 O'	16
Val56 NH	Glu52 O'	19
Leu57 NH	Ala53 O'	18
Ala65 NH	Gln61 O'	14
Tyr66 NH	Lys62 O'	9
Asp67 NH	Arg63 O'	12
Gln68 NH*	Ala64 O'	19
Tyr69 NH*	Ala65 O'	14
Gly70 NH*	Tyr66 O'	9
Ala72 NH	Tyr69 O'	9

[a] Hydrogen bonds are listed if they are identified in at least 9 of the 20 energy-minimized DIANA conformers. NH and O' identify the backbone amide proton and the carbonyl oxygen atom, respectively, O^{δ} and O^{ϵ} identify the side-chain carboxyl oxygen atoms of Asp and Glu, respectively. Backbone amide protons that exchange with a rate $<2 \times 10^{-3}$ s^{-1} with the solvent are underlined, while amide protons exchanging in the intermediate range from 1×10^{-2} to 1.0 s^{-1} are marked with an asterisk (see also Figure 7).

[b] This column lists the number of energy-refined DIANA conformers in which the hydrogen bond was identified. The criteria used for the identification of a hydrogen bond are that the proton-acceptor distance must be less than 2.4 Å, and the angle between the donor-proton bond and the line connecting the acceptor and donor heavy atoms less than 35°. This criterion is sufficiently permissive to detect bifurcated hydrogen bonds.

helices is maximized (Schultz & Schirmer, 1979) (Figure 5 and 6). In particular, the coiled-coil is stabilized by interchain hydrophobic interactions involving the residues *a* and *d* of the coiled-coil heptad repeats denoted as *a b c d e f g* in Figure 5. Exposed residues located in the positions *b*, *b*', *c* and *c*' (Figure 5) possess charged side-chains, with the sole exception of Val56, which forms hydrophobic contacts with helix I. Except for Met30, the amino acid residues in the positions *e*, *e*', *f*, *f*', *g* and *g*' possess charged side-chains or are alanyl residues.

The spatial association of helix I with the helices II and III leads to the formation of a well defined hydrophobic core (Figure 6): the average r.m.s.d. value relative to the mean structure calculated after superposition of the helices I to III amounts to 1.10 Å for all heavy atoms (Table 2). The side-chains of Ile9, Leu10, Val12, Ile21, Ala53 and Leu57 are located at the intersection of the helices I, II and III, and constitute the central part of the hydrophobic core. Directly associated are Tyr25 and Tyr54 at the interface between the helices II and III, and Tyr6, Tyr7 and Val56, which are oriented towards helix IV. Furthermore, Ala24, Ala 29 and Ala45 are peripherally attached to the hydrophobic cluster, and the helix IV contacts the core of the molecule near Tyr6 and Val56 through Tyr66.

Studies of conformational equilibria and internal rate processes in DnaJ(2–108) by measurement of backbone amide proton exchange rates and ^{15}N spin relaxation parameters

Exchange rates of slowly exchanging amide protons were obtained from a series of 2D [^{15}N, ^{1}H] COSY spectra recorded immediately after fully protonated, ^{15}N-labeled DnaJ(2–108) was dissolved in $^{2}H_2O$. In addition, exchange rates for more rapidly exchanging backbone amide protons were determined with a series of MEXICO experiments (Gemmecker *et al*., 1993). The combined use of both approaches allowed measurement of exchange rates covering the two ranges 7.7×10^{-6} to 2.1×10^{-3} s^{-1} and 1.4 to 32.7 s^{-1}, and all the rates in DnaJ(2–108) which were not detected by either of the two experimental approaches could be attributed to the range from 1×10^{-2} s^{-1} to 1 s^{-1} (Figure 7). Following the general considerations by Liepinsh *et al*. (1992) in connection with protein hydration studies, special care was used in the analysis of seryl and threonyl residues to account for exchange effects between the amide and the hydroxyl groups. For example, the backbone amide proton of Thr58 manifests slow exchange in the $^{1}H/^{2}H$ exchange measurements, but also gives rise to a cross-peak in the MEXICO experiments. The peak in the MEXICO experiment must thus be due to an intraresidual NOE between the backbone amide proton and the hydroxyl proton, which exchanges more rapidly with the solvent water (Wüthrich, 1986). We could thus not *a priori* exclude that the amide protons of Ser13, Thr15, and Ser91, all of which are observable in the MEXICO experiments but could not be observed immediately after dissolving the protein in $^{2}H_2O$, actually exchange with a rate in the intermediate range 1×10^{-2} s^{-1} to 1 s^{-1}. For tyrosyl residues the occurrence of a NOE between the hydroxyl group and its own amide protein is sterically unlikely, but additional spurious rapid exchange could be due to

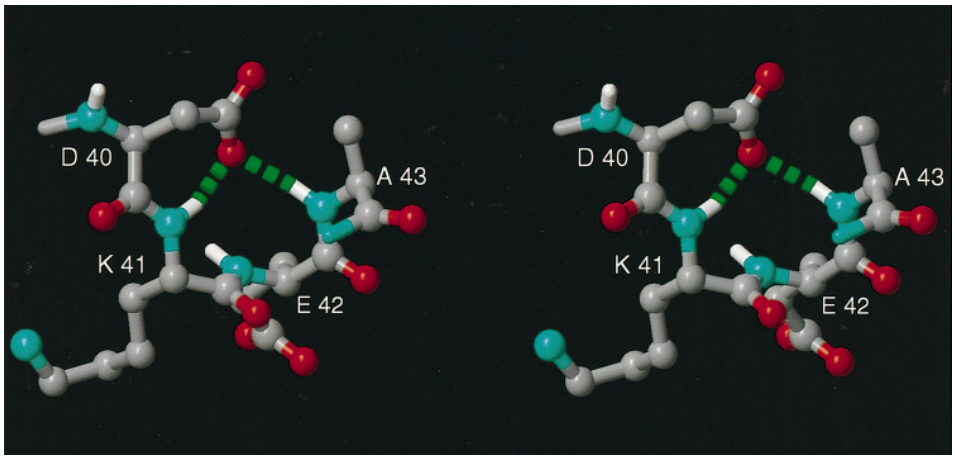

Figure 4. Stereo view of the polypeptide segment Asp40-Lys41-Glu42-Ala43. Oxygen, nitrogen and backbone amide protons are displayed in red, blue and white, respectively. Other atoms are grey. One of the carboxylate oxygen atoms of the N-cap of helix III, Asp40, forms a bifurcated hydrogen bond, depicted in green, with the backbone amide protons of Lys41 and Ala43.

close spatial proximity between amide and hydroxyl groups in the three-dimensional protein structure. Analysis of the 20 energy-refined DIANA conformers shows that only the backbone amide protons of Lys14 and Gln61 are located within 4.0 Å of hydroxyl protons, i.e. those of Ser13 and Ser60, respectively. However, these two amide protons exchange sufficiently slowly to be observed by $^1H/^2H$ exchange, and no signals indicative of NOEs with the hydroxyl groups were observed in the MEXICO experiments.

The backbone amide protons of the residues in the helices I to III which form the hydrophobic core of the J-domain, exhibit the slowest exchange rates (Figure 7). In comparison, relatively fast exchange rates are observed for the amide protons of helix IV, the residues 29 to 33 of helix II and the residues 41 to 44 of helix III. The backbone amide protons of the loop connecting the helices II and III exchange rapidly, whereas in the turn between the helices I and II the backbone amide protons of Lys14 and Ala16 exchange slowly. The backbone amide proton of Ala16 is involved in a hydrogen bond with the carbonyl oxygen of Ser13 (Table 3), and the protection observed for the amide proton of Lys14 is probably due to formation of a hydrogen bond

					Helix II						
d	*e*	*f*	*g*	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*a*
I21	R22	K23	A24	Y25	K26	R27	L28	A29	M30	K31	Y32
L57	V56	E55	Y54	A53	E52	K51	I50	E49	K48	F47	K46
d'	*c'*	*b'*	*a'*	*g'*	*f'*	*e'*	*d'*	*c'*	*b'*	*a'*	*g'*
					Helix III						

Figure 5. Schematic presentation of the coiled-coil heptad repeats present at the interface of the helices II and III. The residues of a given heptad repeat (McLachlan & Stewart, 1975) are denoted *a* to *g* for helix II, and *a*' to *g*' fro helix III. The residues in the shaded boxes form the hydrophobic interface between the helices II and III. Solvent-exposed, charged residues are underlined.

with the side-chain oxygen of Ser13, which, however, was identified in only 4 of the 20 energy-refined DIANA conformers and is therefore not listed in Table 3. Comparison of Figures 7 and 8 shows that the amide proton exchange rates correlate with the local precision of the NMR structure determination in that slow exchange prevails for the segments which are determined with the highest precision (see also Figure 1(b)). There is also good agreement with the hydrogen bonding pattern observed for the NMR structure of the J-domain (Table 3), since with the sole exceptions of Tyr6, Arg19, and Ala72 all hydrogen-bonded amide protons have readily measurable protection factors.

The analysis of the MEXICO data on rapidly exchanging amide protons provides further insight into the conformational properties of the Gly / Phe-rich region. Figure 7 shows that we could not detect significant amide proton protection factors for any of the residues in this region (the protection factors are smaller than 7 throughout), which is in agreement with the earlier observation that the Gly / Phe-rich region adopts predominantly an extended disordered conformation (Szyperski *et al*., 1994a).

To complement the implicit indication of local conformational equilibria by the variable precision of the structure determination along the polypeptide chain and by the amide proton exchange rates (Figures 1(b), 2, 7 and 8) with information on the frequencies of associated rate processes, we further measured T_1, T_2 and $T_{1\rho}$ spin relaxation times for ^{15}N, and steady-state $^{15}N\{^1H\}$ NOEs for ^{15}N-labeled DnaJ(2–108). From the T_1/T_2 ratio of the helical residues of the J domain (Figure 9), an overall rotational correlation time for DnaJ(2–108) of approximately 6 ns was obtained, using the assumption that the protein reorients isotropically in solution. Within the J-domain (residues 2 to 76), some increased values of T_2 in the loop linking the helices II and III, and for the chain-terminal

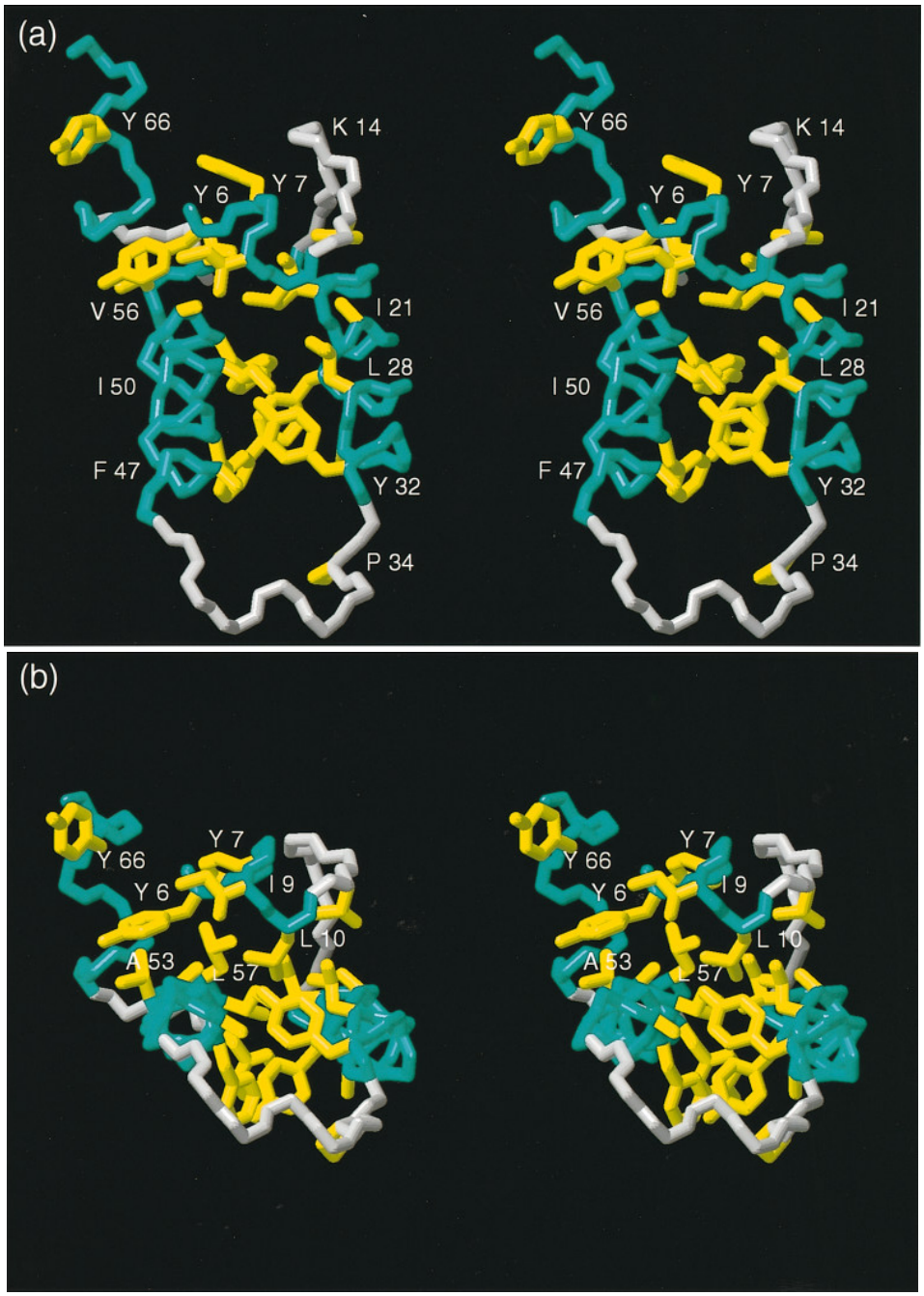

Figure 6. Two views of a stereo drawing of the hydrophobic core of the J-domain. The backbone of the polypeptide segment 6 to 69 and the side-chain heavy atoms of the residues forming the core of the molecule are shown. Some side-chain positions are identified with the one-letter amino acid code and the sequence numbers. The backbone atoms of the four α-helices are colored in blue. Side-chain heavy atoms are shown in yellow. (b) is the same as (a) after a rotation of approximately 90° about a horizontal axis in the projection plane.

segments reveal the presence of higher-frequency rate processes characterized by correlation times in the ns to ps time range, which is also supported by the observed $^{15}N\{^{1}H\}$ NOE values (Figure 9). Furthermore, a very short T_2 value for Asp35 shows that additional slow conformational exchange processes are present in the loop between the helices II and III. Although the positioning of helix IV in the NMR structure is less well defined than that of the helices I to III (Figure 1(b) and 8, Table 2), the relaxation data provide no evidence for local rate processes involving helix IV. We conclude that the poor precision of the structure determination near helix IV arises as a consequence of the scarcity of long-range NOE distance constraints (Figure 1(a)) rather than because of dynamic disorder.

The ^{15}N relaxation data (Figure 9) show that the Gly / Phe-rich region reorients with an effective correlation time which is significantly shorter than the correlation time for overall rotational tumbling

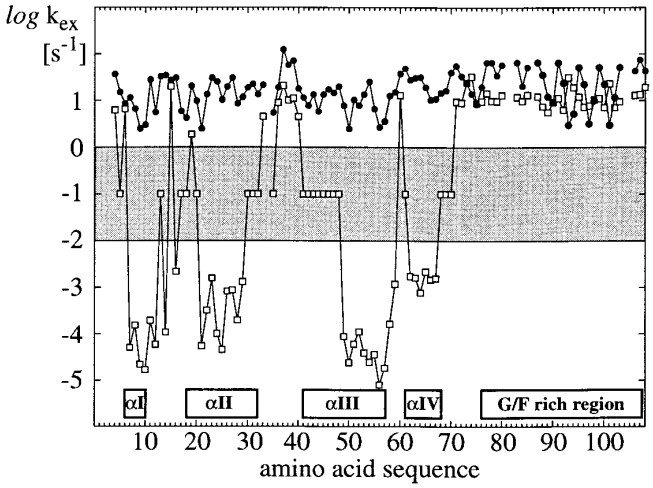

Figure 7. The measured rates of backbone amide proton exchange with the solvent at pH 6.2 and 28°C (open squares) are plotted *versus* the amino acid sequence of DnaJ2(2–108). The intrinsic backbone amide proton exchange rates estimated as described by Bai *et al*. (1993). are similarly displayed as filled circles. Data points at log $k_{ex} = -1$ inside the shaded area represent residues exhibiting an ''intermediate exchange rate'' in the range from 1×10^{-2} s^{-1} to 1 s^{-1} (see the text). The locations of the α-helices in the J-domain and the Gly / Phe-rich region are indicated at the bottom.

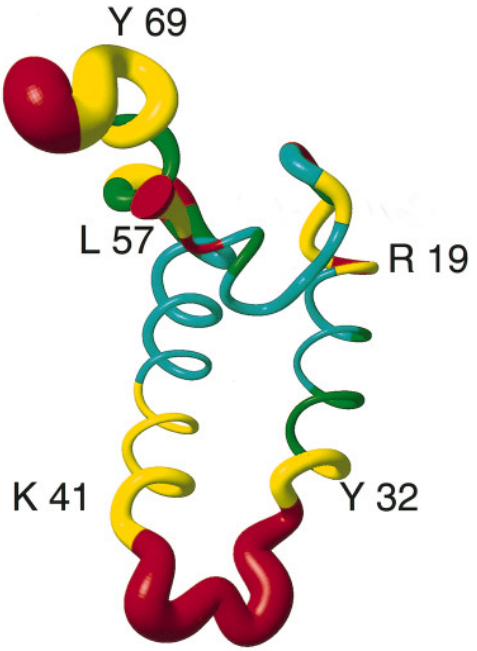

Figure 8. Backbone of the NMR solution structure of the J-domain present in DnaJ(2–108) of *E. coli*. A spline function was drawn through the C^{α} positions, and the thickness of the cylindrical rod is proportional to the mean of the global displacements, D^{bb}_{glob}, of the 20 energy-minimized DIANA conformers calculated after superposition of the heavy atom N, C^{α} and C' of all helical residues for minimal r.m.s.d. (Figure 1(b)). The colours indicate the protection factors *P* calculated as the ratio between the estimated intrinsic backbone amide proton exchange rates (Bai *et al*., 1993) and the corresponding measured values. The color code used is log $P > 4.75$, cyan; $3.75 < \log P < 4.74$, green; $1.75 < \log P < 3.75$, yellow; log $P < 1.75$, red. The Figure was generated with the program MOLMOL (Koradi *et al*., 1996).

of the molecule. The data also show that the polypeptide segments of residues 83 to 90 and 104 to 108 reorient with an effective correlation time that approaches the extreme motional narrowing limit. In contrast, the polypeptide segment of residues 90 to 103 has a significantly longer correlation time. Using the assumption that the backbone amide groups reorient isotropically in solution, the T_1 / T_2 ratios measured for the residues 90 to 103 yield effective correlation times of about 1–2 ns. This polypeptide segment thus does not have the properties of a typical unstructured ''random coil'', but must adopt preferential conformations in solution which reorient with increased, ''non-random'' effective correlation times in solution. The fact that these implicated conformational preferences are not evidenced either by the NOE data or by the chemical shifts leads further to the following conclusions: (1) These preferred conformations are either predominantly extended, or all except the intraresidual and sequential NOEs are effectively quenched by the local mobility. (2) There must be fast exchange on the chemical shift time scale between multiple preferred conformations, so that conformation-dependent shifts are averaged out. To establish a lower limit on the rate of these exchange processes, we performed additional measurements of rotating frame ^{15}N spin relaxation times with two different spin-lock power levels of 6978 rad s^{-1} and 12,868 rad s^{-1}. The fact that no variation of $T_{1\rho}$ with the spin-lock power could be observed establishes an upper limit of 100 μs for effective correlation times that would be dominated by the conformation interconversions (Szyperski *et al*., 1993).

Discussion

The three-dimensional structure of the J-domain represents a novel folding topology for an α-protein. Sequence alignment of the J-domain with the sequences of all three-dimensional protein structures available from the Brookhaven Protein Data Bank (Bernstein *et al*., 1977) did not identify any sequence that exhibits more than 25% homology. Moreover, the distance matrix algorithm of Holm & Sander (1993) implemented in the program DALI revealed no topological similarities with any of the 599 protein chains currently used to represent the contents of the Protein Data Bank in this context (Holm & Sander, 1994).

The J-domain from the *E. coli* DnaJ chaperone very likely represents the three-dimensional fold of a large number of homologous J-domain from widely different origins. A sequence alignment with 50 J-domains (21 from bacteria, 9 from yeasts, 20 from other eukaryotes) revealed a striking

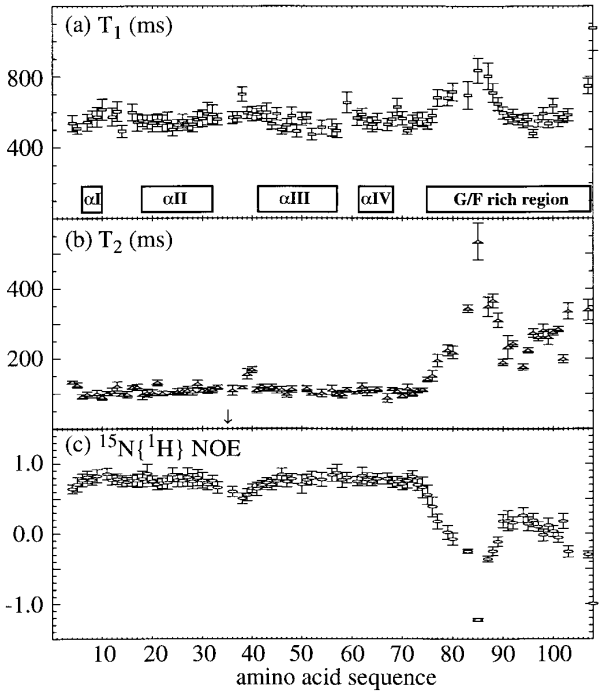

Figure 9. Relaxation times and steady-state $^{15}N\{^1H\}$ NOEs measured for the backbone amide nitrogen atoms of DnaJ(2–108) at a ^1H frequency of 600 MHz. (a) Longitudinal relaxation time, T_1; (b) transverse relaxation time, T_2. The arrow indicates the very short transverse relaxation time of Asp35; (c) $^{15}N\{^1H\}$ NOEs. The locations of the α-helices in the J-domain and of the Gly-Phe-rich region are indicated at the bottom of (a).

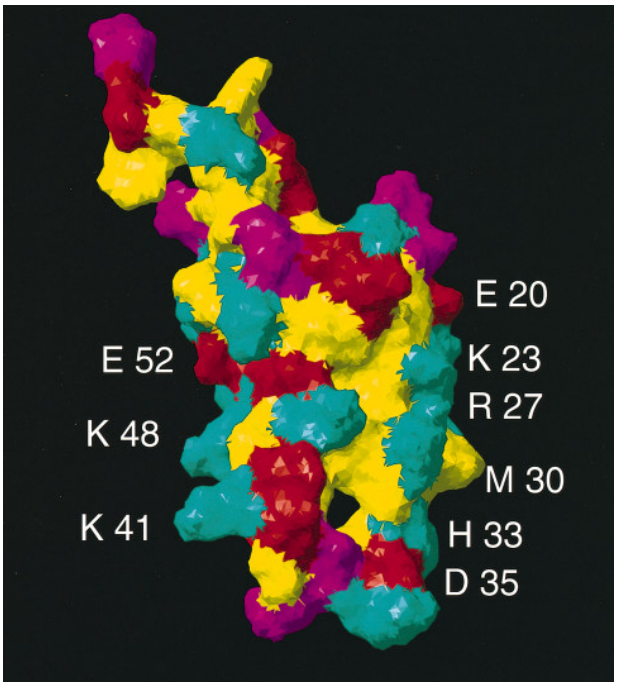

Figure 10. Plot of the Conolly surface of the J-domain present in *E. coli* DnaJ(2–108). Some amino acid residues are identified using the one-letter code and the sequence positions. Positively and negatively charged amino acid residues are depicted in cyan and red, respectively, polar residues are shown in magenta, and hydrophobic residues are yellow. The Figure was generated using the program MOLMOL (Koradi *et al*., 1996).

correlation between the conservation of hydrophobic residues and their structural roles in the hydrophobic cluster (Figure 6; see also Szyperski *et al*., 1994a). The packing of the conserved residues Tyr6, Tyr7, Ile9, Leu10 and Leu57 leaves a small cavity that accommodates the methyl group of the highly conserved residue Ala53 (Figure 6(b)). In Sec63-1, a J-domain-like protein which interacts with Kar2 (Hsp70) in protein transport and folding in the lumen of the endoplasmatic reticulum of yeast, the substitution Ala179 → Thr, which corresponds to a mutation at position 53 in *E. coli* DnaJ and likely results in a perturbation of the hydrophobic cluster, leads to a thermolabile phenotype (Nelson *et al*., 1993). Furthermore, due to the spatial restriction, other substitutions at position 53 provoke size-compensatory mutations for the five surrounding residues (Figure 6(b)) in order to preserve the scaffold of the J-domain. For example, in the J-domain of the SV40 T-antigen, the substitution Ala53 → Leu occurred in concert with Tyr6 → Leu and Leu57 → Met (the residue numbers are those for the DnaJ J-domain of *E. coli*).

The antiparallel coiled-coil formed by the helices II and III is stabilized by interchain hydrophobic interactions between apolar residues situated at positions *a* and *d* of the heptad repeats of Figure 5. Accordingly, Tyr25, Tyr32, Phe47 and Tyr54 at positions *a* and *a'* are almost invariably aromatic, and Ile21, Leu28, Ile50 and Leu57 at positions *d* and *d'* are invariably hydrophobic in the different J-domain sequences. High levels of conservation are also observed for Lys62, Arg63 and Tyr66, which are located in helix IV and constitute its interface to the remainder of the J-domain, and for Val12, which contacts the helices I and II. In contrast, Ala24, Ala29, Met30 and Ala45 of DnaJ, which are not conserved as hydrophobic residues in other J-domain sequences, are only peripherally associated with the hydrophobic cluster (Figure 6; see also Figure 5 of Szyperski *et al*., 1994a). Furthermore, lack of conservation in the polypeptide segment 41 to 45, which constitutes the amino-terminal part of helix III, reflects the fact this segment is at best marginally involved in the formation of the hydrophobic core (Figure 6(a)).

Recent mutational studies provide evidence that some of the solvent-exposed side-chains occupying the positions *b*, *c* and *f* of the heptad repeats (Figure 5) have crucial roles for the interaction of J-domains with Hsp70-like proteins. Schlenstedt *et al*. (1995) have shown that swapping of the J-domain between Sis1p or Mdj1p and Sec63p leads to non-functional chimeric proteins. However, by additional mutations of the J-domain of Sis1p in the positions corresponding to Arg19, Lys23 and Lys48 of the *E. coli* DnaJ-domain (Figure 10), a fully functional Sec63p fusion protein containing the mutant Sis1p J-domain could be obtained. Remark-

ably, we find that Met30 in *E. coli* DnaJ represents a solvent-exposed hydrophobic residue in an *f* position of helix II (Figures 5 and 10), although a charge or polar side-chain would be energetically more favorable. This observation suggests that Met30 could also play a significant role for the specificity of the DnaJ-DnaK interaction.

There is ample evidence that the tripeptide segment His33–Pro34–Asp35, which is universally conserved among the J-domain sequences, plays an essential role in the interactions of DnaJ-like proteins with their Hsp70 counterparts (Bork *et al*., 1992; Kelley & Landry, 1994): First, these interactions are completely abolished by the point mutation His33 → Gln in *E. coli* DnaJ (Wall *et al*., 1994). Second, Feldheim *et al*. (1992) showed for Sec63 that amino acid substitutions in positions which are homologous to Pro34 and Asp35 of *E. coli* DnaJ result in an inactive protein. Third, mutations in the His–Pro–Asp segment of the SV40 T-antigen result in mutant phenotypes (Kelley & Landry, 1994). In the NMR structure of the *E. coli* DnaJ J-domain this tripeptide segment is located in a flexibly disordered loop (Figure 1(b), 8 and 9) which is not part of the structural core of the protein. The implication then is that the high degree of conservation of this tripeptide segment is related to its functional role for the specific protein-protein interactions of DnaJ-like proteins with their Hsp70 counterparts (Szyperski *et al*., 1994a). The slow rate processes observed for this loop (Figure 9(b)) would make this structure element a good candidate for an ''induced-fit'' mechanism in the interactions with the Hsp70 proteins.

Overall, the aforementioned apparent structure–function correlations show that both affinity and selectivity of a particular J-domain to its Hsp70 partner (DnaK in *E. coli*) are mainly determined by specific interactions that involve residues located in the loop connecting the helices II and III and on the outer surfaces of the helices II and III (Figure 10). In this context it is then of interest that the helices II and III are too short to independently form a typical, stable antiparallel coiled-coil, since formation of this structure type would require longer helices consisting of at least three heptad repeats (see Lau *et al*., 1984). It is thus tempting to speculate that the main function of the helices I and IV is to stabilize the antiparallel coiled-coil arrangement of the helices II and III, which in turn represent the scaffold for proper spatial positioning of residues that determine and modulate the interactions with Hsp70. The comparatively low precision of the NMR structure determination for helix IV (Figures 1(b) and 8) is due to the fact that it is connected with the hydrophobic core of the molecule almost solely *via* Tyr66. This rather loose ''docking'' indicates an alternative possible mechanism for fine regulation of the function of the J-domain in that, depending on the positioning of helix IV, the directly adjoining domains might adopt different spatial arrangements relative to the core of the J-domain. This could be particularly important for those J-domain proteins which do not possess the Gly/Phe-rich region located immediately carboxy-terminal to the J-domain.

Ample evidence exists to suggest that the Gly-Phe-rich region may have auxilary functions to those of the J-domain. For example, it is known that deletion of residues 77 to 108 in *E. coli* DnaJ results in a mutant protein that interacts with DnaK and polypeptide substrates with wild-type affinities, but is defective in its ability to activate DnaK to bind substrates (Wall *et al*., 1995). Within the Gly/Phe-rich region, the polypeptide segment comprising residues 92 to 102 shares high sequence homology with corresponding segments in the Gly-Phe-rich regions of other DnaJ-like proteins (Figure 1 of Wall *et al*., 1995). In the case of *E. coli* DnaJ, the characteristic tripeptide sequence, Asp-Ile/Val-Phe, is repeated three times in this 11-amino acid segment. In view of the absence of a structural role in the molecular architecture, the high conservation of residues 92 to 102 suggests that they must be important for the function of DnaJ (Wall *et al*., 1995). In this context it is intriguing that the present NMR investigation shows that the polypeptide segment 90 to 103, which covers the conserved polypeptide segment, does not correspond to a flexible extended ''random coil'' but undergoes rapid exchange between different locally preferred conformations. This finding is of special interest with regard to the hypothesis that the Gly-Phe-rich region participates directly in the formation of a ternary DnaJ-DnaK-polypeptide complex (Wall *et al*., 1995). Conformational preferences in the unbound state, which could also include transient contacts of this polypeptide segment with the J-domain, would reduce the partition function of this region, and the correspondingly reduced entropy loss upon binding of DnaJ to DnaK would eventually contribute to the stability of the ternary complex.

Materials and Methods

Sample preparation

$^{13}C/^{15}N$ doubly-labeled DnaJ(2–108) was prepared as described by Szyperski *et al*. (1994a). To obtain uniformly ^{15}N-labeled DnaJ(2–108) the *E. coli* expression strain DW531 grown in 100 ml Isogro medium culture was used to inoculate seven liters of M9 minimal media (Sambrook *et al*., 1989) containing $^{15}NH_4Cl$ (1 g/l, Isotec Inc.) as the sole nitrogen source. The ^{15}N-labeled protein was purified as described previously (Wall *et al*., 1994).

NMR spectroscopy

NMR measurements were performed on a Bruker AMX-600 spectrometer equipped with four channels, using $^{15}N/^{13}C$ doubly-labeled and ^{15}N-labeled samples of DnaJ(2–108) at a concentration of about 1 mM in a solvent of 90% H_2O/10% ^2H_2O. All spectra were recorded at pH 6.2 and at $T = 28°C$. Quadrature detection in the indirect dimensions was achieved using States-TPPI (Marion *et al*., 1989). For data processing and analysis we

used the programs PROSA (Güntert *et al*., 1992) and XEASY (Bartels *et al*., 1995), respectively.

The following spectra, were recorded to obtain sequence-specific resonance assignments: 3D $\underline{H}^{\alpha/\beta}\underline{C}^{\alpha/\beta}$(CO)NHN (Szyperski *et al*., 1994b) with a time domain data size of $82 \times 30 \times 512$ complex points; $t_{1,max}(^{13}C) = 6.5$ ms and, since the ^1H chemical shifts were scaled up by a factor 1.65, $t_{1,max}(^1H) = 10.8$ ms, $t_{2,max}(^{15}N) = 21.6$ ms, $t_{3,max}(^1H) = 65.5$ ms. 3D CBCANHN (Grzesiek & Bax, 1992) with a time domain data size of $56 \times 31 \times 512$ complex points, with $t_{1,max}(^{13}C) = 6.5$ ms, $t_{2,max}(^{15}N) = 22.3$ ms, $t_{3,max}(^1H) = 65.5$ ms; 2D [^{13}C, ^{1}H] COSY (Bodenhausen & Ruben, 1980), with a time domain data size of 150×1024 complex points, with $t_{1,max}(^{13}C) = 14.1$ ms, $t_{2,max}(^1H) = 131.1$ ms; 3D HCCH-TOCSY (Bax *et al*., 1990) with a time domain data size of $100 \times 30 \times 512$ complex points and a mixing time of 12 ms, with $t_{1,max}(^1H) = (^1H) = 17.4$ ms, $t_{2,max}(^{13}C) = 9.6$ ms, $t_{3,max}(^1H) = 74.7$ ms; 3D HCCH-COSY (Kay *et al*., 1990) with a time domain data size of $100 \times 20 \times 512$ complex points, with $t_{1,max}(^1H) = 17.4$ ms, $t_{2,max}(^{13}C) = 6.4$ ms, $t_{3,max}(^1H) = 74.7$ ms; 2D *ct*-[^{13}C,^{1}H] COSY (Vuister & Bax, 1992) recorded for aromatic side-chain assignments, with a time domain data size 512×1024 complex points, and $t_{1,max}(^{13}C) = 26$ ms, $t_{2,max}(^1H) = 149.5$ ms; $^{13}C(\omega_2)$-half-filtered 2D [^{1}H,^{1}H] TOCSY and $^{13}C(\omega_2)$-half-filtered 2D [^{1}H, ^{1}H] NOESY (Otting & Wüthrich, 1990) with mixing times of 50 ms and 100 ms, respectively, a time domain data size of 200×512 complex points, and $t_{1,max}(^1H) = 34.8$ ms, $t_{2,max}(^1H) = 74.8$ ms. Unless specified otherwise, the carrier positions were set to 4.72 ppm for ^{1}H, 117 ppm for ^{15}N, 56 ppm, 43 ppm, 35 ppm and 128 ppm for $^{13}C^{\alpha}$, $^{13}C^{\alpha/\beta}$, $^{13}C^{aliph.}$ and $^{13}C^{arom}$, respectively, and 177 ppm for carbonyl carbons. The ^{1}H chemical shifts are relative to internal 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS). ^{15}N and ^{13}C chemical shifts are relative to internal DSS using the conversion factors proposed by Wishart *et al*. (1995).

The conformational constraints for the structure determination were collected from 3D ^{15}N-resolved [^{1}H, ^{1}H] NOESY (Fesik *et al*., 1988) recorded with 100 ms mixing time, a time domain data size of $180 \times 32 \times 512$ complex points, and $t_{1,max}(^1H) = 31.3$ ms, $t_{2,max}(^{15}N) = 27.4$ ms and $t_{3,max}(^1H) = 74.8$ ms and from 3D ^{13}C-resolved [^{1}H, ^{1}H] NOESY (Ikura *et al*., 1990) performed with a mixing time of 70 ms, a time domain data size of $160 \times 30 \times 512$ complex points, and $t_{1,max}(^1H) = 27.8$ ms, $t_{2,max}(^{13}C) = 9.6$ ms and $t_{3,max}(^1H) = 74.8$ ms. The vicinal $^3J_{HN\alpha}$ coupling constants were obtained by inverse Fourier transformation of in-phase multiplets (Szyperski *et al*., 1992) from a 2D [^{15}N ^{1}H] COSY spectrum recorded with a data size of 170×2048 complex points, and $t_{1,max}(^{15}N) = 122.4$ ms and $t_{2,max}(^1H) = 131.1$ ms. Vicinal $^3J_{N\beta}$ scalar coupling constants were estimated from a 3D *ct*-HNNHB spectrum (Archer *et al*., 1991) performed with a time domain data size of $90 \times 38 \times 512$ complex points, and $t_{1,max}(^1H) = 31.3$ ms, $t_{2,max}(^{15}N) = 27.4$ ms and $t_{3,max}(^1H) = 74.8$ ms $^3J_{\alpha\beta}$ coupling constants were derived from a ^{15}N-resolved [^{1}H, ^{1}H] TOCSY spectrum recorded with 50 ms mixing time, a data size of $180 \times 38 \times 512$ complex points, and $t_{1,max}(^1H) = 31.3$ ms, $t_{2,max}(^{15}N) = 27.4$ ms and $t_{3,max}(^1H) = 74.8$ ms.

For the measurement of rapid backbone amide proton exchange rates a series of seven 2D MEXICO experiments (Gemmecker *et al*., 1993) was recorded with ^{15}N/^{13}C doubly-labeled DnaJ(2–108). The mixing times were 50, 100, 150, 200, 250, 300 and 400 ms, respectively, with time domain data size of 512×1024 complex points, and $t_{1,max}(^{15}N) = 92.2$ ms and $t_{2,max}(^1H) = 131.1$ ms. The total measuring time was 30 hours. From these data exchange rates were obtained by a least square fit of the cross-peak intensities of the function $I(t) = I_o(1 - e^{-kt})$, where I_o and $I(t)$ denote the cross-peak volumes at mixing time zero and mixing time t, respectively, and k is the first order exchange rate constant. For the determination of slower exchange rates, a sample of 500 µl of a fully protonated 1 mM H_2O solution of uniformly ^{15}N labeled DnaJ(2-108) was lyophilized. The protein was then redissolved in the same amount of ^2H_2O, and a series of twelve 2D [^{15}N, ^{1}H] COSY spectra were recorded, using a time domain data size of 200×512 complex points, and $t_{1,max}(^{15}N) = 122.4$ ms and $t_{2,max}(^1H) = 131.1$ ms. The individual measurements were started 4, 34, 64 ,94, 124, 184, 244, 364, 484, 604, 844 and 1084 minutes after preparation of the ^2H_2O solution, which added up to a total recording time of 24 hours. The rate constants were obtained from a least-square fit of the cross-peak volumes to a single exponential function.

^{15}N spin relaxation time measurements were performed with uniformly ^{15}N-labeled DnaJ(2–108) at a ^{1}H resonance frequency of 600 MHz as described by Szyperski *et al*. (1993). All [^{15}N, ^{1}H] COSY spectra needed for these measurements were recorded with a time domain data size of 200×512 complex points, with $t_{1,max}(^{15}N) = 122.4$ ms and $t_{2,max}(^1H) = 131.1$ ms. For the measurement of longitudinal ^{15}N spin relaxation times, T_1, a series of 11 spectra was recorded. The following relaxation delays were used: $t_{rel} = 10$, 40, 200, 400, 550, 700 and 900 ms, 1.0, 1.5, 2.0 and 2.5 seconds. The total measurement time was 22 hours. Transverse ^{15}N spin relaxation times, T_2, were obtained from a series of nine 600 MHz spectra with the following relaxation delays: $t_{rel} = 25$, 40, 50, 60, 80, 100, 140, 180 and 260 ms. The total measuring time was 18 hours. Rotating frame ^{15}N spin relaxation times, $T_{1\rho}$, were measured with a series of seven spectra and relaxation delays of $t_{rel} = 30$, 50, 70, 90, 130, 170 and 240 ms. Two different series of experiments were performed, with spin-lock powers of 12,868 rad s^{-1} and 6978 rad s^{-1}, respectively. In order to avoid possible small off-resonance effects for glycyl residues, the series with 6978 rad s^{-1} was measured twice, setting the ^{15}N carrier position to 109 and 126 ppm, respectively. The individual ^{15}N–$^1H^N$ cross-peaks were analyzed in the experiment with the smaller offset to the ^{15}N carrier position. The total measuring time was 63 hours. All ^{15}N spin relaxation times were obtained from a least-squares fit of the experimental data to a single exponential function.

$^{15}N\{^1H\}$ steady-state heteronuclear Overhauser effects were determined as described by Skelton *et al*. (1993), using the sensitivity enhancement scheme of Palmer *et al*. (1992) modified to the extent that proton saturation was achieved by application of high-power 120 degree pulse spaced at 20 ms intervals for three seconds prior to the first pulse on ^{15}N (Kay *et al*., 1989). A relaxation delay of six seconds was used in order to reduce effects arising from amide proton exchange. The total measuring time was 88 hours for the two spectra recorded with and without proton-presaturation. NOE values were determined as the ratio of corresponding cross-peak volumes measured from the spectra acquired with and without ^{1}H saturation during the recycle delay.

Structure calculation

The input for the distance geometry calculations with the program DIANA consisted of upper distance limits derived from the NOESY cross-peak intensities with the

program CALIBA (Güntert *et al*., 1991), and of dihedral angle constraints derived from $^3J_{HN\alpha}$, $^3J_{\alpha\beta}$, $^3J_{N\beta}$ coupling constants and from local NOEs using the program HABAS (Güntert *et al*., 1989). HABAS also provided a number of stereospecific assignments of β-methylene protons. Several rounds of structure calculations with DIANA (Güntert *et al*., 1991) and NOESY cross-peak assignments with ASNO (Güntert *et al*., 1993) were performed. Backbone dihedral angle constraints derived from conformation-dependent $^{13}C^{\alpha}$ chemical shifts (Luginbühl *et al*., 1995) were initially used to accelerate this iterative cycle, but were not used in the final structure calculation. Additional stereospecific assignments were obtained using the program GLOMSA (Güntert *et al*., 1991). No hydrogen bond constraints were included in the input for DIANA. The final round of DIANA structure calculations was started with 50 randomized conformers and included two REDAC cycles, where the maximal target function value per residue for locally acceptable segments was set to $C^{(1)} = 0.6$ Å and $C^{(2)} = 0.4$ Å, respectively (Güntert & Wüthrich, 1991). The resulting 20 DIANA conformers with the smallest residual target function values were subjected to restrained energy minimization using the AMBER force field (Weiner *et al*., 1986) as implemented in the program OPAL (P. Luginbühl, P. Güntert, M. Billeter & K. Wüthrich, unpublished), where the pseudo-potential for NMR constraints was adjusted such that violations of 0.1 Å for distance constraints and 2.5° for dihedral angle constraints corresponded to $k_BT/2$ at room temperature. The energy minimization was carried out in a shell of water molecules with a minimal thickness of 6 Å. A total of 1500 steps of conjugate gradient minimization were performed for each conformer. The resulting 20 energy-refined conformers are used to represent the solution structure of the J-domain. R.m.s.d. values were calculated with the program XAM (Xia, 1992), color Figures and stereo views were generated with the program MOLMOL (Koradi *et al*., 1996).

Acknowledgements

Financial support was obtained from the Schweizerischer Nationalfonds (Projects 31.32035.91 and 31.31129.91). M.P. is indebted to Universita' degli studi di Napoli ''Federico II'' for a fellowship. The energy minimizations were performed on the NEC SX-3/22 of the Centro Svizzero di Calcolo Scientifico in Manno, Ticino, Switzerland. We thank William Kelley for providing unpublished alignments of J-domain sequences, and R. Marani for the careful processing of the manuscript.

References

Archer, S. J., Ikura, M., Torchia, D. A. & Bax, A. (1991). An alternative 3D NMR technique for correlating backbone ^{15}N with side chain Hβ resonances in larger proteins. *J. Magn. Reson.* **95**, 636–641.

Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 75–86.

Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.

Bax, A., Clore, G. M. & Gronenborn, A. M. (1990). ^1H-^1H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson.* **88**, 425–431.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimamouchi, T. & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

Billeter, M., Braun, W. & Wüthrich, K. (1982). Sequential resonance assignments in protein ^1H nuclear magnetic resonance spectra: computation of sterically allowed proton–proton distances and statistical analysis of proton–proton distances in single crystal protein conformations. *J. Mol. Biol.* **155**, 321–346.

Bondenhausen, G. & Ruben, D. (1980). Natural abundance ^{15}N NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Letters*, **69**, 185–188.

Bork, P., Sander, C., Valencia, A. & Bukau, D. (1992). A module of the DnaJ heat shock proteins found in malaria parasites. *Trends Biochem. Sci.* **17**, 129.

Chary, K. V. R., Otting, G. & Wüthrich, K. (1991). Measurements of small heteronuclear ^1H–^{15}N coupling constants in ^{15}N-labeled proteins by 3D H_NNH_{AB}-COSY. *J. Magn. Reson.* **93**, 218–224.

Clore, G. M., Bax, A. & Gronenborn, A. M. (1991). Stereospecific assignment of β-methylene protons in larger proteins using 3D ^{15}N-separated Hartmann-Hahn and ^{13}C-separated rotating frame Overhauser spectroscopy. *J. Biomol. NMR*, **1**, 13–22.

Feldheim, D., Rothblatt, J. & Schekman, R. (1992). Topology and functional domains of Sec63p, an endoplasmic reticulum membrane protein required for secretory protein translocation. *Mol. Cell. Biol.* **12**, 3288–3296.

Fesik, S. W. & Zuiderweg, E. R. P. (1988). Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *J. Magn. Reson.* **78**, 588–593.

Gemmeker, G., Jahnke, W. & Kessler, H. (1993). Measurement of fast proton exchange rates in isotopically labeled compounds. *J. Am. Chem. Soc.* **115**, 11620–11621.

Georgopoulos, C. & Welch, W. J. (1993). Role of the major heat shock proteins as molecular chaperones. *Annu. Rev. Cell. Biol.* **9**, 601–634.

Georgopoulos, C., Ang, D., Liberek, K. & Zylicz, M. (1990). In *Stress Proteins in Biology and Medicine* (Morimoto, R., Tissières, A. & Georgopoulos, C., eds), pp. 191–221, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Grzesiek, S. & Bax, A. (1992). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201–207.

Güntert, P. & Wüthrich, K. (1991). Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR*, **1**, 447–456.

Güntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989). Automated stereospecific ^1H NMR assignments and their impact on the precision of protein structure determinations in solution *J. Am. Chem. Soc.* **111**, 3997–4004.

Güntert, P., Braun, W. & Wüthrich, K. (1991). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting

programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.

Güntert, P., Dötsch, V., Wider, G. & Wüthrich, K. (1992). Processing of multi-dimensional *NMR* data with the new software PROSA. *J. Biomol. NMR*, **2**, 619–629.

Güntert, P., Berndt, K. D. & Wüthrich, K. (1993). The program ASNO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J. Biomol. NMR*, **3**, 601–606.

Hill, B. R., Flanagan, J. M. & Prestegard, J. H. (1995). ^{1}H and ^{15}N magnetic resonance assignments, secondary structure, and tertiary fold of *Escherichia coli* DnaJ(1-78). *Biochemistry*, **34**, 5587–5596.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.

Holm, L. & Sander, C. (1994). The FSSP database of structurally aligned protein fold families. *Nucl. Acids Res.* **22**, 3600–3609.

Ikura, M., Kay, L. E., Tschudin, R. & Bax, A. (1990). Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein. *J. Magn. Reson.* **86**, 204–209.

Kay, L. E., Ikura, M. & Bax, A. (1990). Proton-proton correlation via carbon-carbon couplings: a three-dimensional NMR approach for the assignment of aliphatic resonance in proteins labeled with carbon-13. *J. Am. Chem. Soc.* **112**, 888–889.

Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972–8979.

Kelley, W. & Landry, S. J. (1994). Chaperone power in a virus? *Trends Biochem. Sci.* **19**, 277–278.

Kimura, Y., Yahara, I. & Lindquist, S. (1995). Role of the protein chaperone YDJ1 in establishing Hsp90-mediated signal transduction pathways. *Science*, **268**, 1362–1365.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, in the press.

Lau, S. Y. H., Taneja, A. K. & Hodges, R. S. (1984). Effects of high-performance liquid chromatographic solvents and hydrophobic matrices on the secondary and quaternary structure of a model protein. Reversed-phase and size-exclusion HPLC. *J. Chromatogr.* **317**, 129–140.

Liberek, K. & Georgopoulos, C. (1993). Autoregulation of the *Escherichia coli* heat shock response by the DnaK and DnaJ heat shock proteins. *Proc. Natl Acad. Sci. USA*, **90**, 11019–11023.

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. & Zylicz, M. (1991). *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc. Natl Acad Sci. USA*, **88**, 2874–2878.

Liberek, K., Wall, D. & Georgopoulos, C. (1995). The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the σ^{32} heat shock transcriptional regulator. *Proc. Natl Acad. Sci. USA*, **92**, 6224–6228.

Liepinsh, E., Otting, G. & Wüthrich, K. (1992). NMR spectroscopy of hydroxyl protons in aqueous solutions of peptides and proteins. *J. Biomol. NMR*, **2**, 447–465.

Luginbühl, P., Szyperski, T. & Wüthrich, K. (1995). Statistical basis for the use of $^{13}C^{\alpha}$ chemical shifts in protein structure determination. *J. Magn. Reson.* **109B**, 229–233.

Marion, D., Ikura, K., Tschudin, R. & Bax, A. (1989). Rapid recording of 2D NMR spectra without phase cycling: application of the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393–399.

McLachlan, A. D. & Stewart, M. (1975). Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol.* **98**, 293–304.

Messerle, B. A., Wider, G., Otting, G., Weber, C. & Wüthrich, K. (1989). Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with H_2O solutions. *J. Magn. Reson.* **85**, 608–613.

Nelson, M. K., Kurihara, T. & Silver, P. A. (1993). Extragenic suppressors of mutation in the cytoplasmic C terminus of Sec63 define five genes in *Saccharomyces cerevisae*. *Genetics*, **134**, 159–173.

Otting, G. & Wüthrich, K. (1990). Heteronuclear filters in two-dimensional [^{1}H,^{1}H]-NMR spectroscopy: combined use with isotope labelling for studies of macromolecular conformation and intermolecular interaction. *Quart. Rev. Biophys.* **23**, 39–96.

Palmer, A. G. III, Canavagh, J., Wright, P. E. & Rance, M. (1992). Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **93**, 151–170.

Perkins, S. J. & Wüthrich, K. (1979). Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. *Biochem. Biophys. Acta*, **576**, 409–423.

Richardson, J. S. & Richardson, D. C. (1988). Amino acid preferences for specific locations at the end of α helices. *Science*, **240**, 1648–1652.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schlenstedt, G. , Harris, S., Risse, B., Lill, R. & Silver, P. A. (1995). A yeast DnaJ homologue, Scj1p, can function in the endoplasmic reticulum with Bip/Kar2p via a conserved domain that specifies interactions with Hsp70s. *J. Cell Biol.* **129**, 979–988.

Schultz, G. E. & Schirmer, R. H. (1979). *Principles of Protein Structure*, pp. 79–81, Springer-Verlag, New York.

Silver, P. A. & Way, J. C. (1993). Eukaryotic DnaJ homologs and the specificity of Hsp70 activity. *Cell*, **74**, 5–6.

Skelton, N. J., Palmer, A. G., III, Akke, M., Ködel, J., Rance, M. & Chazin, W. J. (1993). Practical aspects of two-dimensional proton-detected ^{15}N spin relaxation measurements. *J. Magn. Reson.* **102B**, 253–264.

Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992). Determination of scalar coupling constants by inverses Fourier transformation of in-phase multiplets. *J. Magn. Reson.* **99**, 552–560.

Szyperski, T., Luginbühl, P., Otting, P., Güntert, P. & Wüthrich, K. (1993). Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J. Biomol. NMR*, **3**, 151–164.

Szyperski, T., Pellecchia, M., Wall, D., Georgopoulos, C. & Wüthrich, K. (1994a). NMR structure determination of the *Escherichia coli* DnaJ molecular chaperone: secondary structure and backbone fold of the N-terminal region (residues 2-108) containing the highly conserved J domain. *Proc. Natl Acad. Sci. USA*, **91**, 11343–11347.

Szyperski, T., Pellecchia, M. & Wüthrich, K. (1994b). 3D $\underline{H}^{\alpha/\beta}\underline{C}^{\alpha/\beta}$(CO)NHN, a projected 4D NMR experiment for sequential correlation of polypeptide $^{1}H^{\alpha/\beta}$,$^{13}C^{\alpha/\beta}$ and backbone ^{15}N and $^{1}H^{N}$ chemical shifts. *J. Magn. Reson.* **105B**, 188–191.

Szyperski, T., Antuch, W., Schick, M., Betz, A., Stone, S. R. & Wüthrich, K. (1994c). Transient hydrogen bonds identified on the surface of the NMR solution structure of Hirudin. *Biochemistry*, **33**, 9303–9310.

Vuister, G. W. & Bax, A. (1992). Resolution enhancement and spectral editing of uniformly ^{13}C enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn Reson.* **98**, 428–435.

Wagner, G. & Wüthrich, K. (1982). Sequential resonance assignments in protein ^{1}H nuclear magnetic resonance spectra: basic pancreatic trypsin inhibitor. *J. Mol. Biol.* **155**, 347–366.

Wall, D., Zylicz, M. & Georgopoulos, C. (1994). The NH_2-terminal 108 amino acids of the *Escherichia coli* DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for λ replication. *J. Biol. Chem.* **269**, 5446–5451.

Wall, D., Zylicz, M. & Georgopoulos, C. (1995). The conserved G/F motif of the DnaJ chaperone is necessary for the activation of the substrate binding properties of DnaK chaperone. *J. Biol. Chem.* **270**, 2139–2144.

Weiner, P. K., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* **7**, 230–252.

Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. & Sykes, B. D. (1995). ^{1}H,^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Xia, T. (1992). Software for determination and visual display of NMR structures of proteins: the distance geometry program DGPLAY and the computer graphics programs CONFOR and XAM. PhD. thesis Nr. 9831, ETH Zürich, Switzerland.

Edited by P. E. Wright

(Received 29 February 1996; accepted 25 April 1996)

Supplementary material, comprising one table, is available from JMB Online.