Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 11343–11347, November 1994
Biophysics

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

THOMAS SZYPERSKI†, MAURIZIO PELLECCHIA†, DANIEL WALL‡§, COSTA GEORGOPOULOS§, AND KURT WÜTHRICH†

†Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule–Hönggerberg, CH-8093 Zürich, Switzerland; ‡Department of Cellular, Viral, and Molecular Biology, University of Utah Medical Center, Salt Lake City, UT 84132; and §Département de Biochimie Médicale, Centre Médical Universitaire, 1, rue Michel-Servet, 1221 Genève 4, Switzerland

Contributed by Kurt Wüthrich, July 29, 1994

ABSTRACT **DnaJ from *Escherichia coli* is a 376-amino acid protein that functions in conjunction with DnaK and GrpE as a chaperone machine. The N-terminal fragment of residues 2–108, DnaJ-(2–108), retains many of the activities of the full-length protein and contains a structural motif, the J domain of residues 2–72, which is highly conserved in a superfamily of proteins. In this paper, NMR spectroscopy was used to determine the secondary structure and the three-dimensional polypeptide backbone fold of DnaJ-(2–108). By using $^{13}C/^{15}N$ doubly labeled DnaJ-(2–108), nearly complete sequence-specific assignments were obtained for ^{1}H, ^{15}N, $^{13}C^{\alpha}$, and $^{13}C^{\beta}$, and about 40% of the peripheral aliphatic carbon resonances were also assigned. Four α-helices in polypeptide segments of residues 6–11, 18–31, 41–55, and 61–68 in the J domain were identified by sequential and medium-range nuclear Overhauser effects. For the J domain, the three-dimensional structure was calculated with the program DIANA from an input of 536 nuclear Overhauser effect upper-distance constraints and 52 spin–spin coupling constants. The polypeptide backbone fold is characterized by the formation of an antiparallel bundle of two long helices, residues 18–31 and 41–55, which is stabilized by a hydrophobic core of side chains that are highly conserved in homologous J domain sequences. The Gly/Phe-rich region from residues 77 to 108 is flexibly disordered in solution.**

The *Escherichia coli dnaJ*, *dnaK*, and *grpE* heat shock genes encode a molecular chaperone machine that participates in a variety of biological processes. For example, they function in protein folding and transport, survival at high temperatures, negative autoregulation of the heat shock response, modulation of *in vivo* proteolysis rates, and the replication of bacteriophage λ and certain plasmids (for reviews, see refs. 1 and 2). The J domain, which is the N-terminal 71 residues of DnaJ (3, 4), is an evolutionarily highly conserved motif found in a superfamily of proteins that includes >50 members from prokaryotic and eukaryotic organisms (refs. 3–6; W. Kelley, personal communication). Likewise, DnaK is a member of the large family of highly conserved HSP70 proteins. Recently, a eukaryotic homologue of the GrpE protein has been identified in yeast as well (7).

Structure–function studies of DnaJ and its interactions with DnaK have demonstrated that the N-terminal region of amino acids 2–108, which includes the Gly/Phe-rich region (residues 77–108) in addition to the J domain, is necessary and sufficient to stimulate the ATPase activity of DnaK, regulate the conformational state of DnaK in the presence of ATP, activate DnaK to bind the heat shock protein σ^{32} in the presence of ATP, and elicit proper, albeit limited, chaperone function from DnaK for λ DNA replication (ref. 8; K. Liberek, D.W., and C.G., unpublished data). Thus, DnaJ-(2–108) possesses many of the properties of full-length DnaJ. To obtain further insights into the structural basis of DnaJ functions and, in particular, the DnaJ–DnaK interactions, we have started NMR structure determinations in solution. This paper reports sequence-specific resonance assignments and determination of the secondary structure and of the global polypeptide backbone fold of DnaJ-(2–108).

MATERIALS AND METHODS

Preparation of $^{13}C/^{15}N$ Doubly Labeled DnaJ-(2–108). DnaJ-(2–108) was purified to >95% purity from a strain carrying the expression vector pDW19dnaJ(2–108) as described (8). To doubly label DnaJ-(2–108), a 100-ml culture grown in Isogro medium (Isotec, Miamisburg, OH) was used to inoculate 7 liters of M9 minimal medium (9) supplemented with ampicillin (100 μg/ml), thiamine (2 μg/ml), $MgSO_4$ (1 mM), $MgCl_2$ (3 mM), $CaCl_2$ (0.1 mM), $FeCl_3$ (0.3 μM), [$^{13}C_6$]-D-glucose (0.3%, Isotec), and $^{15}NH_4Cl$ (1 g/liter, Isotec). The culture was induced with isopropyl β-D-thiogalactopyranoside (1 mM) at an A_{595} of 0.7 for 4 h before harvest. Protein sequence analysis revealed that Ala-2 represents the N-terminal residue.

NMR Spectroscopy. All NMR spectra were recorded on a Bruker AMX 600 spectrometer equipped with four channels, using a single sample of $^{15}N/^{13}C$ doubly labeled DnaJ-(2–108). The protein concentration was ≈1 mM in 90% H_2O/10% ^2H_2O at pH 6.2 and 28°C. The spectra that were recorded to obtain sequence-specific resonance assignments are in *Results*. For data processing and analysis, we used the programs PROSA (10) and XEASY (C. Bartels, T. H. Xia, M. Billeter, P. Güntert, and K.W., unpublished data). To collect the input for the structure calculations using the program DIANA (11), upper limit distance constraints were derived from three-dimensional (3D) ^{15}N-resolved [^{1}H,^{1}H]-nuclear Overhauser enhancement spectroscopy (NOESY) [ref. 12; mixing time τ_m = 100 ms; data size, 180 * 38 * 512 complex points, with $t_{1,max}$ (^{1}H) = 31.3 ms, $t_{2,max}$ (^{15}N) = 27.4 ms, $t_{3,max}$ (^{1}H) = 74.8 ms; digital resolution after zero-filling 11 Hz along ω_1, 22 Hz along ω_2, 5.7 Hz along ω_3] and 3D ^{13}C-resolved [^{1}H, ^{1}H]-NOESY [ref. 13; τ_m = 70 ms; ^{15}N-decoupling during t_1; data size 160

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: NOE, nuclear Overhauser effect; 2D and 3D, two and three dimensional, respectively; NOESY, NOE spectroscopy; COSY, correlated spectroscopy; TOCSY, total correlation spectroscopy; 3D CBCA NHN, spectrum correlating $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts with intraresidual and sequential amide proton and amide ^{15}N chemical shifts; 3D $H^{\alpha/\beta}C^{\alpha/\beta}$(CO)NHN, spectrum correlating $^{1}H^{\alpha/\beta}$ and $^{13}C^{\alpha/\beta}$ chemical shifts with sequential amide proton and amide ^{15}N chemical shifts.

* 30 * 512 complex points, with $t_{1,max}$ (^{1}H) = 27.8 ms, $t_{2,max}$ (^{13}C) = 9.6 ms, $t_{3,max}$ (^{1}H) = 74.8 ms; digital resolution after zero-filling 11 Hz along ω_1, 48 Hz along ω_2, 5.7 Hz along ω_3]. The vicinal scalar coupling constants $^3J_{HN\alpha}$ were extracted by inverse Fourier transformation of in-phase multiplets (14) from the two-dimensional (2D) [^{15}N,^{1}H]-correlated spectroscopy (COSY) spectrum [^{13}C=O and $^{13}C^{\alpha}$ decoupling during t_1; data size 170 * 1024 complex points, with $t_{1,max}$ (^{15}N) = 122.4 ms, $t_{2,max}$ (^{1}H) = 131.1 ms; digital resolution after zero-filling 2.7 Hz along ω_1, 1.9 Hz along ω_2]. Rapidly exchanging amide protons were identified in a 2D Mexico experiment (15) recorded with a mixing time of 300 ms.

RESULTS

For a general characterization of DnaJ-(2–108), the thermal denaturation was monitored by CD at 222 nm. A melting point of ≈75°C was obtained, demonstrating that DnaJ-(2–108) is thermally quite stable. Although full-length DnaJ is dimeric (16), ultracentrifugation experiments suggest that DnaJ-(2–108) is monomeric in dilute aqueous solution (50 μM protein/0.7 M K_2HPO_4, pH 7.0; L. Gonzalez and T. Alber, personal communication). Furthermore, the NMR linewidths are consistent with a monomeric species. Thus, the possibility that a homodimer might be formed was not further considered in the present NMR investigation.

Fig. 1 shows a 2D [^{15}N,^{1}H]-COSY spectrum of uniformly ^{15}N/^{13}C doubly labeled DnaJ-(2–108) recorded with decoupling of $^{13}C^{\alpha}$ and ^{13}C=O during t_1. The chemical shift dispersion in the ^{1}H dimension is rather limited, but the dispersion in the ^{15}N dimension is sufficient to resolve most of the cross-peaks. All of the aliphatic proton resonances are located between 0.2 ppm and 4.9 ppm, so that it would be virtually impossible to derive a high-quality NMR solution

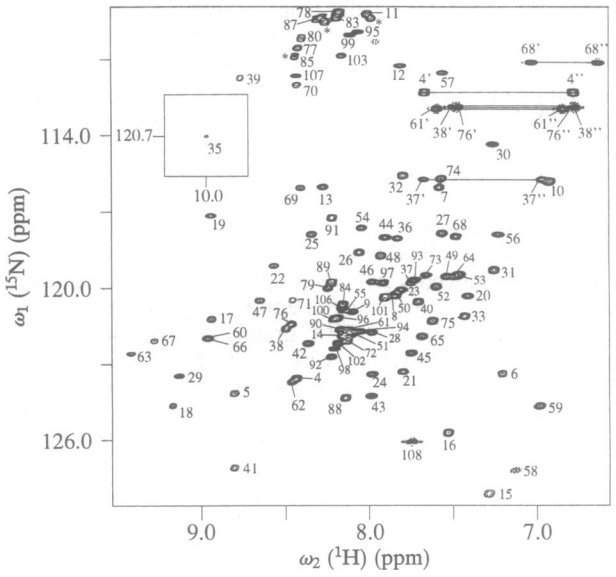

FIG. 1. Contour plot of a 2D [^{15}N,^{1}H]-COSY spectrum (17) of DnaJ-(2–108) uniformly labeled with ^{13}C and ^{15}N that was recorded with decoupling of $^{13}C^{\alpha}$ and ^{13}C=O during t_1 (^{1}H frequency, 600 MHz; protein concentration, 1 mM in 90% H_2O/10% ^2H_2O, pH 6.2, 28°C). The ^{1}H chemical shifts are relative to the water resonance at 4.72 ppm from [2,2,3,3-$^2H_4$]-trimethylsilylpropionate and the ^{15}N chemical shifts are referenced indirectly to external liquid NH_3 (18). The cross-peaks are marked with the sequence positions of the corresponding amino acid residues and the peaks that belong to one or several of the unassigned Gly residues in the polyglycine segments are marked with asterisks. The cross-peaks identified with primed and doubly primed numbers correspond to the pairs of side-chain amide protons of Asn and Gln. The cross-peak of Asp-35 is outside of the spectral region shown and has, therefore, been added as an insert. The cross-peaks belonging to Thr-15 and Thr-58 are folded along ω_1.

structure of DnaJ-(2–108) on the basis of ^{15}N-resolved and homonuclear ^{1}H NMR experiments alone. We therefore used ^{15}N/^{13}C doubly labeled DnaJ-(2–108) for the present project. Fig. 1 also shows that a homogeneous protein preparation was used for the NMR studies.

Sequence-Specific Resonance Assignments and Identification of Medium-Range Nuclear Overhauser Enhancements (NOEs). Fig. 2 presents an overview of the sequential and medium-range connectivities identified for DnaJ-(2–108). Sequence-specific polypeptide backbone assignments were obtained for the 101 residues 3–80, 83–85, 87–103, and 106–108 from 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY (12) and 3D ^{15}N-resolved [^{1}H,^{1}H]-TOCSY (total COSY) (12) or from 3D $H^{\alpha/\beta}C^{\alpha/\beta}$(CO)NHN (spectrum correlating $^1H^{\alpha/\beta}$ and $^{13}C^{\alpha/\beta}$ chemical shifts with sequential amide proton and amide ^{15}N chemical shifts) (19) and 3D CBCANHN (spectrum correlating $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts with intraresidual and sequential amide proton and amide ^{15}N chemical shifts) (20); at least one sequential connectivity was found for all pairs of assigned neighboring residues. While the chemical shifts of $C^{\alpha/\beta}$ and $H^{\alpha/\beta}$ of Lys-3 could be obtained from the 3D $H^{\alpha/\beta}C^{\alpha/\beta}$(CO)NHN experiment, the resonances of the backbone amide moiety of Lys-3 were not detected in the ^{15}N-resolved spectra. In addition, Asp-35 appears to be affected by a slow conformational exchange process on the millisecond time scale, since only a very weak signal was observed in the ^{15}N-resolved spectra (Fig. 1). For Pro-34, the observation of a strong $d_{\alpha/\delta}$ NOE indicates that it is present in the *trans* conformation (Fig. 2). Five Gly residues at positions 81, 82, 86, 104, and 105, which are located within the polyglycine segments in the Gly/Phe-rich region (Fig. 2), could not be sequentially assigned (Fig. 1).

Assignment of the Aliphatic Side-Chain Resonances and Identification of Long-Range NOEs. It is a special advantage of the 3D $H^{\alpha/\beta}C^{\alpha/\beta}$(CO)NHN experiment that it provides the chemical shifts of H^{α}, H^{β}, C^{α}, and C^{β} in a single measurement in conjunction with the sequential connectivities to ^{15}N and H^N of the following residue (19). The $^{13}C^{\alpha/\beta}$ and $^1H^{\alpha/\beta}$ chemical shifts thus obtained were used as the starting point for the ^{1}H assignments of the long aliphatic side chains in the 3D [HCCH]-TOCSY (21) and 3D [HCCH]-COSY (22) spectra. Overall, with the exception of the five aforementioned Gly residues, we obtained resonance assignments for all protons and for 77% (228 of 296) of the carbon positions in the aliphatic side chains. To identify medium-range $d_{\alpha\beta}(i,i+3)$ NOEs (Fig. 2) and long-range NOE connectivities involving side-chain protons, the peak positions from the 3D [HCCH]-COSY and 3D [HCCH]-TOCSY spectra were transferred to the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum.

Assignment of the Aromatic Spin Systems. The seven Tyr rings at positions 6, 7, 25, 32, 54, 66, and 69 of the J domain exhibit sufficiently large chemical shift dispersion to enable unambiguous identification of the NOE connectivities from the δ-protons to the α- and/or β-protons in 2D $^{13}C(\omega_2)$-half-filtered [^{1}H,^{1}H]-NOESY (23) and 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY. The complete ^{1}H assignments of the Tyr rings subsequently yielded numerous long-range NOE constraints within the J domain. Similarly, Phe-47 was assigned. For the Phe residues in the Gly/Phe-rich region (residues 77–108), the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum enabled us to obtain sequence-specific resonance assignments for the δ-protons of the Phe rings, which have nearly identical chemical shifts. Since no intraresidual NOE connectivities were observed between the $-\alpha CH-\beta CH_2-$fragment and the ring protons of the His residues 33 and 71, these rings remained unassigned.

Secondary Structure of DnaJ-(2–108). The sequential and medium-range connectivities identified in the 3D ^{15}N- and ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectra, the $^3J_{HN\alpha}$ coupling constants, and the chemical shifts of the $^{13}C^{\alpha}$ resonances (Fig. 2) resulted in the determination of four helices, by using

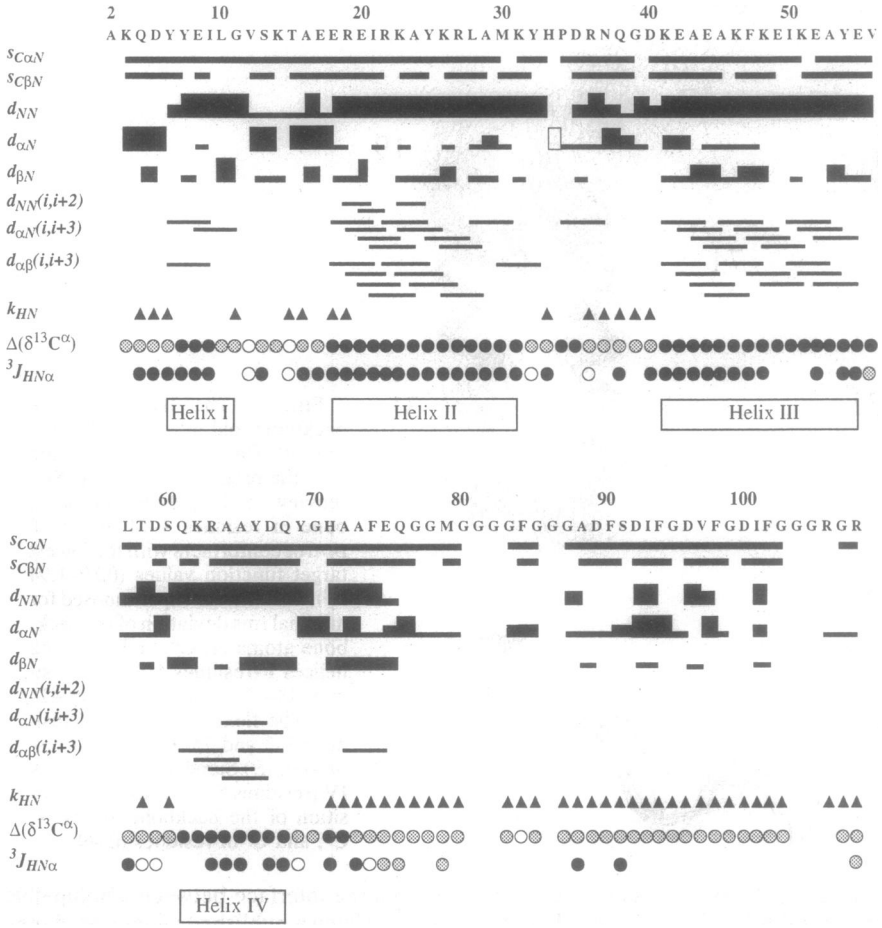

FIG. 2. Amino acid sequence of DnaJ-(2–108) and survey of the data collected for secondary structure identification. In rows $S_{C\alpha N}$ and $S_{C\beta N}$, a bar indicates that the sequential scalar coupling connectivities were identified using the 3D CBCANHN and 3D $\underline{H}^{\alpha/\beta}\underline{C}^{\alpha/\beta}$(CO)NHN experiments. Below these, the sequential NOE connectivities are indicated, where the thickness of the bars represents the relative NOE intensities. (For His-33–Pro-34, the sequential $d_{\alpha\delta}$ connectivity is indicated by an open bar.) The medium-range NOEs $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha\beta}(i,i+3)$ are represented by lines starting and ending at the positions of the connected residues. Triangles in row k_{HN} represent amide protons with exchange rate constants >1 min^{-1} at pH 6.2 and 28°C. In row $\Delta\delta(^{13}C^{\alpha})$, $\Delta(\delta^{13}C^{\alpha}) = \delta(^{13}C^{\alpha})_{obs} - \delta(^{13}C^{\alpha})_{rc}$, where $\delta(^{13}C^{\alpha})_{obs}$ and $\delta(^{13}C^{\alpha})_{rc}$ denote the observed and the random coil chemical shifts, respectively. Solid, shaded, and open circles indicate residues with $\Delta\delta(^{13}C^{\alpha}) > 1.5$ ppm, 1.5 ppm $\geq \Delta\delta(^{13}C^{\alpha}) \geq -1.5$ ppm, and $\Delta\delta(^{13}C^{\alpha}) < -1.5$ ppm, respectively. In row $^{3}J_{HN\alpha}$, solid, shaded, and open circles denote residues with $^{3}J_{HN\alpha} < 6.0$ Hz, 6.0 Hz $\leq {}^{3}J_{HN\alpha} \leq 8.0$ Hz, and $^{3}J_{HN\alpha} > 8.0$ Hz (some data are missing because of spectral overlap). The sequence locations of four helices identified from the data in this figure and from the structure calculations with the program DIANA (see text) are indicated at the bottom.

the following criteria for the identification of helical residues: (*i*) $^{3}J_{HN\alpha} < 6.0$ Hz (24); (*ii*) strong sequential d_{NN} NOEs (25); (*iii*) slow amide proton exchange rates from the fourth helical residue to the end of the helix (26); (*iv*) downfield chemical shifts of the $^{13}C^{\alpha}$ nuclei relative to their random coil values (27) of $\Delta\delta(^{13}C^{\alpha}) > 1.5$ ppm (28); and (*v*) presence of $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOE connectivities characteristic of α-helices (29). The thus identified helices I–IV are residues 6–11, 18–31, 41–55, and 61–68, respectively, and all are located in the J domain (residues 2–72; Fig. 2).

The Gly/Phe-rich region of residues 77–108 appears to be flexibly disordered in solution. Although spectral overlap in the 2D [^{15}N,^{1}H]-COSY spectrum (Fig. 1) prevented measurements of the $^{3}J_{HN\alpha}$ scalar coupling constants for most of the non-Gly residues in this region, this view is nonetheless supported by multiple independent observations (Fig. 2): (*i*) all backbone amide protons exchange rapidly; (*ii*) all $^{13}C^{\alpha}$ chemical shifts deviate from the random coil shifts by <1.6 ppm; (*iii*) all ^{15}N chemical shifts coincide to within 3 ppm with the sequence-corrected random coil values (30); and (*iv*) no medium-range or long-range NOE connectivities were detected that would be characteristic of regular secondary structure elements.

Calculation of the 3D Structure. For a preliminary calculation of the complete 3D structure of Dna-(2–108), we used an input of 620 NOE upper-distance constraints (279 intraresidual, 91 sequential, 31 medium-range and long-range backbone-backbone, and 219 interresidual constraints with side-chain protons) that were derived from the 3D ^{15}N-resolved and 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectra. In addition, 60 vicinal $^{3}J_{HN\alpha}$ scalar coupling constants were obtained from the 2D [^{15}N,^{1}H]-COSY spectrum (Figs. 1 and 2). Five-hundred-thirty-six of the NOE upper distance constraints (217, 75, 31, 213) and 52 vicinal scalar coupling constants are within the J domain, residues 2–72, which corresponds to >8 constraints per residue. The average pairwise rms deviation values relative to the mean structure calculated for the backbone atoms N, C^{α}, and C′ of the entire polypeptide segment considered (0.32 Å for helix I; 0.71 Å for helix II; 0.61 Å for helix III; 0.45 Å for helix IV) show that the four helices are locally well defined. The rms deviation value obtained after superposition of all helical residues is 2.05 Å, but the corresponding value after superposition of the first three helices is only 1.17 Å, which shows that the spatial orientation of helix IV is not well defined relative to the remainder of the J domain. Technically, this is due to the fact that there are only two long-range NOE connectivities involving protons of helix IV. Since residues 61 and 64–67 have been completely assigned and only the assignments of $^{13}C^{\gamma}$, $^{13}C^{\delta}$, $^{13}C^{\varepsilon}$ for the residues 62, 63, and 66 respectively, are missing, one can expect that additional NOEs with helix IV would have been detected. Therefore, the poorly defined orientation of helix IV (Fig. 3*A*) is probably due to increased flexibility relative to the core of the J domain. Nevertheless, after local superposition of residues 61–68, the precise definition of helix IV as a regular secondary structure element is readily apparent (Fig. 3*B*).

DISCUSSION

The 3D backbone fold of the DnaJ J domain, which likely represents a superfamily of structural motifs (3–7), describes a distinctive topology for the arrangement of α-helices I–III (Fig. 4). For example, a search of the amino acid sequences of the 3D protein structures available from the Protein Data Bank, Brookhaven National Laboratory (Upton, NY) did not identify any sequences that exhibit >25% sequence homology to the J domain, and no near-identity can be found with any of the folding types that have recently been surveyed by Thornton's

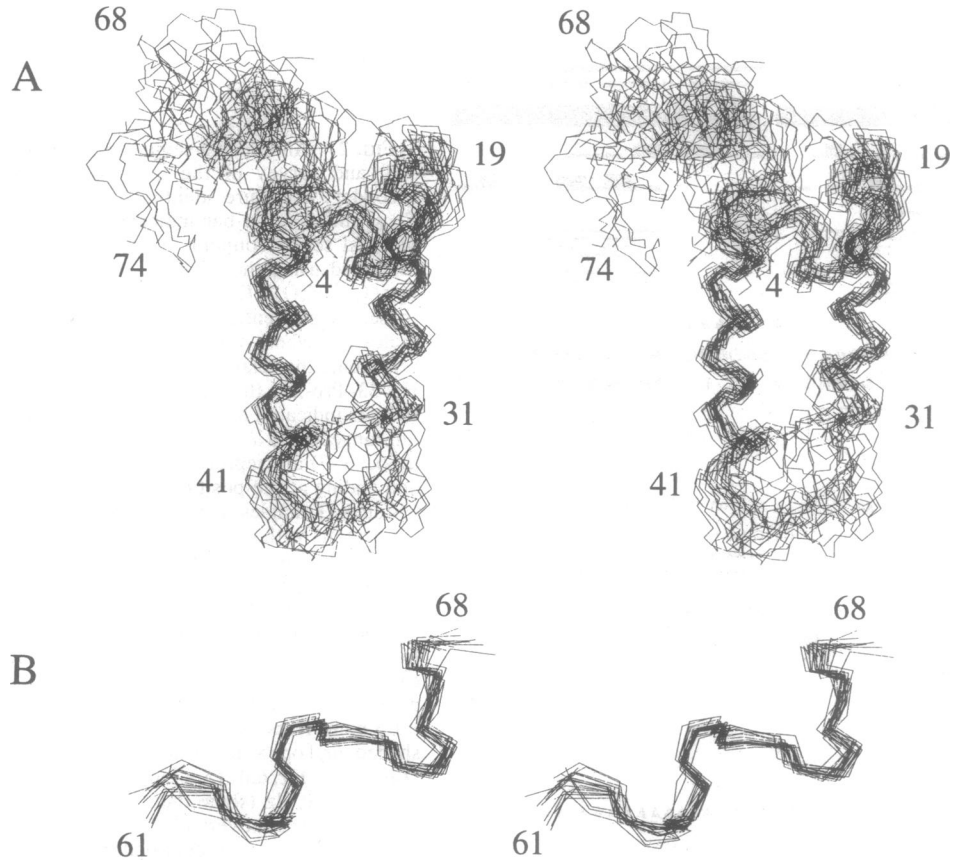

FIG. 3. The 3D polypeptide backbone fold in the NMR structure of DnaJ-(2–108) calculated with the program DIANA. (*A*) Stereoview of the polypeptide backbone of residues 4–74 of the 20 DIANA conformers with the lowest target function values (0.89–1.94 Å^2), which were superimposed for minimal rms deviation of the backbone atoms N, C$^\alpha$, and C′ of the helices I (residues 6–11), II (residues 18–31), and III (residues 41–55). The flexibly disordered residues 2–3 and 75–108 are not displayed. (*B*) Same as in *A* for helix IV (residues 61–68) after superposition of the backbone atoms N, C$^\alpha$, and C′ of residues 61–68.

group (32). Although the ongoing refinement of the NMR structure still has to properly elucidate the structural role of helix IV, it seems likely that helices I–III form a fold, which is most likely common to all J-domain sequences (3, 5); i.e., helices II and III are oriented in an antiparallel fashion, with helix I crossing from the C-terminal end of helix III to the central region of helix II (Figs. 3 and 4). This spatial arrangement of the three helices leads to the formation of a hydrophobic core that includes the interface between amphipathic helices II and III (Fig. 5). Using a published alignment of nine J-domain sequences (3), we found striking correlations be-

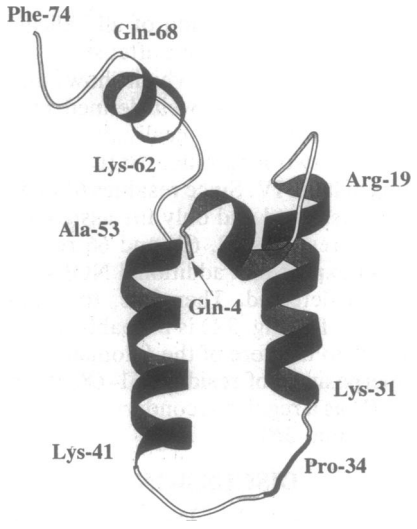

FIG. 4. Ribbon drawing of residues 4–74 of the DIANA conformer of DnaJ-(2–108) with the lowest residual target function value. Note that although helix IV (residues 61–68) is locally well determined (see Fig. 3*B*), its orientation relative to the remainder of the molecule is not defined by the NMR data (Fig. 3*A*). The coil region corresponding to the location of the highly conserved tripeptide segment His-33–Pro-34–Asp-35 is solid. The figure was generated with the program MOLSCRIPT (31).

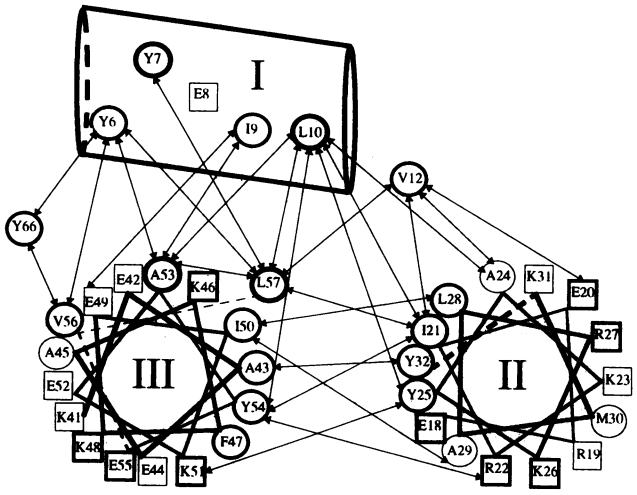

FIG. 5. Helix-wheel representation of the hydrophobic core of the DnaJ J domain in the NMR structure of DnaJ-(2–108). Interhelical NOEs are indicated by arrows, while circles and squares denote hydrophobic and charged residues, respectively. The thickness of the symbols corresponds to the degree of conservation of the amino acid residues in the alignment of nine J-domain sequences by Bork *et al.* (3). Thick circles denote strictly conserved hydrophobic residues, intermediate circles indicate residues with highly conserved aromatic or hydrophobic properties, and thin circles represent residues that are substituted by polar or charged side chains in some of the other eight J domains. Intermediate squares denote charged residues that are present in at least five of the nine sequences and in no case are replaced by residues with the opposite charge, and thin squares indicate positions with nonconservative mutations in the other J domains.

tween the conservation of hydrophobic residues in the J domain with their structural roles in the hydrophobic cluster (Fig. 5). The strictly conserved residues Tyr-7, Leu-10, Ala-53, and Leu-57 constitute the central part of this hydrophobic core. Interestingly, even conservative variations appear not to be tolerated for these residues. For example, the Ala-179→Thr substitution in Sec63-1, which corresponds to position 53 in DnaJ, leads to a thermolabile phenotype (33). In addition, the aforementioned sequence alignment (3) of nine J domains shows that residues Tyr-32, Phe-47, and Tyr-54 are invariably aromatic, residues Tyr-6, Val-12, Tyr-25, and Ile-50 are invariably hydrophobic; and Ile-9, Ile-21, Leu-28, Ala-43, and Val-56 have conserved hydrophobicity in all but one of the nine sequences. All of these residues are located within the hydrophobic core (Fig. 5). In contrast, residues Ala-24, Ala-29, Met-30, and Ala-45 in DnaJ, which are not conserved as hydrophobic residues in all J-domain sequences, are not at all or only peripherally associated with this hydrophobic cluster. On the hydrophilic sides of the amphipathic helices, charged residues Glu-18, Glu-20, Arg-22, Lys-26, Arg-27, Lys-46, Lys-48, and Glu-55 are present in at least five of the nine J-domain sequences aligned by Bork *et al.* (3) and are in no case replaced by residues with opposite charge. Therefore, the C-terminal segment of helix II containing residues 22–31 would contain an excess of positive charges in all nine J-domain sequences. Since this polypeptide segment is close to the tripeptide segment His-33–Pro-34–Asp-35 (Fig. 4), which plays an essential role in the DnaJ–DnaK interaction (8), it is tempting to speculate that interactions of the positively charged helix II with a negatively charged surface area of DnaK might contribute to the functionally important interactions between the two proteins.

Although we limited this discussion to the protein core formed by helices I–III (Fig. 4), it seems worthwhile to point out that position 66 in helix IV, which is invariably occupied by an aromatic residue in the aforementioned nine J domains (3), is involved in NOEs with two highly conserved core residues (Fig. 5). It remains to be seen whether helix IV is indeed in steady contact with this protein core or whether these NOEs are instead due to transient interactions of the two moieties.

The tripeptide segment His-33–Pro-34–Asp-35 (Fig. 4) is extremely conserved among J-domain sequences (3, 6). Its essential role in the DnaJ–DnaK interactions is directly shown by the fact that these interactions are completely abolished by the His-33→Gln point mutation (ref. 8; D.W., M. Zyclicz, and C.G., unpublished data). Similarly, Feldheim *et al.* (34) showed for Sec63, a protein that interacts with BiP (Hsp70) in protein transport and folding in the lumen of the endoplasmatic reticulum of yeast, that amino acid substitutions in positions that are homologous to Pro-34 and Asp-35 of DnaJ also result in an inactive protein. Furthermore, mutations in the His-Pro-Asp segment of simian virus 40 large tumor antigen result in mutant phenotypes (6). The fact that this invariable segment does not have an obvious structural role in the J domain (Fig. 4) is in agreement with the idea that the high degree of conservation is related with an essential functional role of the His-Pro-Asp segment, for example, in the specific protein–protein interactions with DnaK.

Many J-domain proteins contain a Gly/Phe-rich region C-terminal to the J domain, suggesting that this region may have auxiliary functions to those of the J domain. For example, an internal deletion of residues 77–108 in DnaJ results in a mutant protein that interacts with DnaK and with polypeptide substrates with wild-type affinities but is defective in its ability to activate DnaK to bind substrates (D.W., M. Zyclicz, and C.G., unpublished data). Our NMR investigation shows that the Gly/Phe-rich region is flexibly disordered in solution, a structural property that would match well with the proposed function as a linker between the J domain and the C-terminal substrate binding domain (4) and could help to orchestrate protein–protein interactions of DnaJ, DnaK, and the bound polypeptide substrate.

We thank Dr. W. Braun for performing the sequence alignments and Dr. W. Kelley for a critical reading of the manuscript and helpful discussions. We also thank Dr. T. Alber for help in the preliminary phase of this work and for permission to quote unpublished data. Financial support was obtained from the Schweizerischer Nationalfonds (Projects 31.32033.91 and 31.31129.91). M.P. is indebted to the Universitá degli studi di Napoli 'Federico II' for a fellowship.

1. Georgopoulos, C., Ang, D., Liberek, K. & Zylicz, M. (1990) in *Stress Proteins in Biology and Medicine*, eds. Morimoto, R., Tissières, A. & Georgopoulos, C. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 191–221.
2. Georgopoulos, C. & Welch, W. J. (1993) *Annu. Rev. Cell Biol.* **9**, 601–634.
3. Bork, P., Sander, C., Valencia, A. & Bukau, D. (1992) *Trends Biochem. Sci.* **17**, 129.
4. Silver, P. A. & Way, J. C. (1993) *Cell* **74**, 5–6.
5. Gething, M.-J. & Sambrook, J. (1992) *Nature (London)* **355**, 33–45.
6. Kelley, W. & Landry, S. J. (1994) *Trends Biochem. Sci.* **19**, 277–278.
7. Bollinger, L., Deloche, O., Glick, B. S., Georgopoulos, C., Jenö, P., Krenidou, N., Horst, M., Morishima, N. & Schatz, G. (1994) *EMBO J.* **13**, 1998–2008.
8. Wall, D., Zylicz, M. & Georgopoulos, C. (1994) *J. Biol. Chem.* **269**, 5446–5451.
9. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
10. Güntert, P., Dötsch, V., Wider, G. & Wüthrich, K. (1992) *J. Biomol. NMR* **2**, 619–629.
11. Güntert, P., Braun, W. & Wüthrich, K. (1991) *J. Mol. Biol.* **217**, 517–530.
12. Fesik, S. W. & Zuiderweg, E. R. P. (1988) *J. Magn. Reson.* **78**, 588–593.
13. Ikura, M., Kay, L. E., Tschudin, R. & Bax, A. (1990) *J. Magn. Reson.* **86**, 204–209.
14. Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992) *J. Magn. Reson.* **99**, 552–560.
15. Gemmecker, G., Jahnke, W. & Kessler, H. (1993) *J. Am. Chem. Soc.* **115**, 11620–11621.
16. Zylicz, M., Yamamoto, T., McKittrick, N., Sell, S. & Georgopoulos, C. (1985) *J. Biol. Chem.* **260**, 7591–7598.
17. Bodenhausen, G. & Ruben, D. (1980) *Chem. Phys. Lett.* **69**, 185–188.
18. Live, D. H., Davis, D. G., Agosta, W. C. & Cowburn, D. (1984) *J. Am. Chem. Soc.* **106**, 1939–1941.
19. Szyperski, T., Pellecchia, M. & Wüthrich, K. (1994) *J. Magn. Reson. B*, in press.
20. Grzesiek, S. & Bax, A. (1992) *J. Magn. Reson.* **99**, 201–207.
21. Bax, A., Clore, G. M. & Gronenborn, A. M. (1990) *J. Magn. Reson.* **88**, 425–431.
22. Kay, L. E., Ikura, M. & Bax, A. (1990) *J. Am. Chem. Soc.* **112**, 888–889.
23. Otting, G. & Wüthrich, K. (1990) *Q. Rev. Biophys.* **23**, 39–96.
24. Pardi, A., Billeter, M. & Wüthrich, K. (1984) *J. Mol. Biol.* **180**, 741–751.
25. Billeter, M., Braun, W. & Wüthrich, K. (1982) *J. Mol. Biol.* **155**, 321–346.
26. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
27. Richarz, R. & Wüthrich, K. (1978) *Biochemistry* **17**, 2263–2269.
28. Spera, S. & Bax, A. (1991) *J. Am. Chem. Soc.* **113**, 5490–5492.
29. Wüthrich, K., Billeter, M. & Braun, W. (1984) *J. Mol. Biol.* **180**, 715–740.
30. Braun, D., Wider, G. & Wüthrich, K. (1994) *J. Am. Chem. Soc.*, in press.
31. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
32. Orengo, C. A., Flores, T. P., Taylor, W. R. & Thornton, J. M. (1993) *Protein Eng.* **6**, 485–500.
33. Nelson, M. K., Kurihara, T. & Silver, P. A. (1993) *Genetics* **134**, 159–173.
34. Feldheim, D., Rothblatt, J. & Schekman, R. (1992) *Mol. Cell. Biol.* **12**, 3288–3296.