Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 299 (2002) 42–48
BBRC
www.academicpress.com

Determination of the secondary structure in solution of the *Escherichia coli* DnaA DNA-binding domain

Takayuki Obita,[a,1,2] Takafumi Iwura,[b,1,3] Masayuki Su'etsugu,[b] Yoichiro Yoshida,[a] Yoshitsugu Tanaka,[a] Tsutomu Katayama,[b,*] Tadashi Ueda,[a,*] and Taiji Imoto[a]

[a] *Department of Immunology, Kyushu University Graduate School of Pharmaceutical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan*
[b] *Department of Molecular Biology, Kyushu University Graduate School of Pharmaceutical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan*

Received 29 September 2002

Abstract

DnaA protein binds specifically to a group of binding sites collectively called as DnaA boxes within the bacterial replication origin to induce local unwinding of duplex DNA. The DNA-binding domain of DnaA, domain IV, comprises the C-terminal 94 amino acid residues of the protein. We overproduced and purified a protein containing only this domain plus a methionine residue. This protein was stable as a monomer and maintained DnaA box-specific binding activity. We then analyzed its solution structure by CD spectrum and heteronuclear multi-dimensional NMR experiments. We established extensive assignments of the ^{1}H, ^{13}C, and ^{15}N nuclei, and revealed by obtaining combined analyses of chemical shift index and NOE connectivities that DnaA domain IV contains six α-helices and no β-sheets, consistent with results of CD analysis. Mutations known to reduce DnaA box-binding activity were specifically located in or near two of the α-helices. These findings indicate that the DNA-binding fold of DnaA domain IV is unique among origin-binding proteins.

Keywords: DnaA; DNA-binding domain; Circular dichroism; NMR; *Escherichia coli*; DNA replication; Origin-binding protein; Secondary structure; Chemical shift index; Solution structure

DnaA protein is a key protein in the initiation of chromosomal replication in *Escherichia coli* [1,2]. This protein recognizes the replication origin, *oriC*, and directs several proteins required for replication to this site. DnaA protein specifically binds the asymmetric 9-bp sequence $TTAT^{C}/_{A}CA^{C}/_{A}A$, which is called the 9-mer or DnaA box [3–5]. When 20–40 molecules of ATP-bound DnaA bind cooperatively to *oriC*, this protein causes local unwinding of the ATP-rich repeats followed by successive assembly of proteins at the *oriC* region to synthesize primer RNA and complementary DNA strands. DnaA protein thereby creates a specific nucleoprotein structure in which ordered structural DNA changes and protein–protein interactions take place. Questions about which fold of the DnaA protein directly interacts with DNA and how this affects DNA structure to cause unwinding of duplex DNA remain to be answered.

DnaA homologs are widely conserved in eubacterial species [6,7]. Based on structural similarity and specific functions, the structure of *E. coli* DnaA is subdivided into four domains [6–9]. The N-terminal domain I (amino acids 186) and domain II (amino acids 87–134) contain DnaA oligomerization activity [10] and DnaB helicase binding sites [11,12]. Domain III (amino acids 135–373) contains ATP-binding/hydrolysis modules that belong to the AAA^{+} ATPase family [13–16]. The C-terminal domain IV consists of an evolutionarily conserved structure, as do domains I and III, and bears DNA-binding activity [8]. This DnaA domain IV (DAD-IV) contains 94 amino acids from Val-374 to Ser-467.

* Corresponding authors. Fax: +81-92-642-6646 (T. Katayama); +81-92-642-6667 (T. Ueda).
E-mail addresses: katayama@phar.kyushu-u.ac.jp (T. Katayama); ueda@phar.kyushu-u.ac.jp (T. Ueda).
[1] These authors contributed equally to this study.
[2] Present address: Institute for Bioregulation, Kyushu University, Fukuoka, Japan.
[3] Present address: Kirin Brewery Co., Tokyo, Japan.

0006-291X/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0006-291X(02)02590-1

Structural analysis of DAD-IV is of crucial importance for understanding the mechanisms of DNA structural changes that take place during replication initiation. *E. coli* DnaA protein has a strong tendency to aggregate, which makes it difficult to crystallize the full-length protein. However, we performed the first successful analysis of the secondary structure of full-length DnaA by developing a new DnaA purification strategy and using circular dichroism (CD) spectrum analysis [17]. In the present study, we reveal for the first time the secondary structure of DAD-IV in solution using NMR analysis. We constructed and purified a protein that consists of the 94 amino acid residues of DAD-IV including the first methionine residue. This DAD-IV protein is stable as a monomer and has DnaA box-specific DNA-binding activity. The secondary structure of this DAD-IV protein was analyzed by CD spectrum and determined by NMR analysis of ^{1}H, ^{13}C, and ^{15}N resonance assignment. These results implicate a unique DNA-binding mode for DnaA protein.

Materials and methods

Construction of an overproducer of DAD-VI protein. Plasmid pHB9 bearing the *dnaA* region [18] was digested with *Bst*EII and *Xho*I to isolate the DAD-IV coding region (V374–S467). The cohesive end generated by *Bst*EII was filled in with DNA polymerase I large fragment and the resultant DNA fragment was ligated with an *Nco*I–*Xho*I fragment of pET-21d(+) (Novagen). The resultant plasmid termed pIU1 overexpresses a DAD-IV protein with methionine as the first N-terminal residue. pIU1 was introduced into BL21(DE3) cells bearing pLysS (Promega). However, we found that pLysS was not necessary for growth of this overproducing strain (see below).

Overexpression and purification of DAD-IV protein. For biochemical analysis, BL21(DE3) cells bearing pIU1 and pLysS were grown at 37 °C in LB medium and 1 mM IPTG was added when the optical density (A_{660}) of the culture reached 0.5. After further incubation for 2 h, cells were harvested, suspended in chilled buffer C (50 mM Hepes–KOH (pH 7.6), 1 mM EDTA, 2 mM DTT, and 15% glycerol), and incubated in the presence of 1 M KCl, 20 mM spermidine–HCl, and 0.2 mg/ml lysozyme for 30 min on ice and for 4 min at 37 °C with inversion. The following procedures were all performed at 0–6 °C. After chilling on ice, lysates were isolated in a Beckman Ti50-2 rotor by centrifugation at 40,000 rpm for 40 min (fraction I), proteins precipitated with 0.34 g/ml ammonium sulfate were eliminated by centrifugation, and the supernatants were pooled (fraction II). Dialysis was performed until the conductivity was equivalent to that of 50 mM KCl, and then proteins in fraction II were loaded on a Hi-Trap Heparin–agarose column (5 ml, Amersham Biotech) equilibrated with buffer C containing 50 mM KCl and eluted with a linear gradient from 50 to 800 mM KCl in the same buffer. Eluted DAD-IV protein was pooled (fraction III) and further fractionated through a Superose-12 PC3.2/30 column (Amersham Biotech) equilibrated with buffer C containing 400 mM KCl. Purified DAD-IV protein was pooled (fraction IV) and stored at −80 °C.

For NMR structural analysis, the same overproducing strain described above was used, except that pLysS was excluded. Excluding pLysS stimulated expression of DAD-IV protein several fold. The expression procedure was the same as described above. Soluble lysates were obtained, precipitates formed with polyethylenimine were removed, and DAD-IV protein in the soluble fraction was chromatographed on a P11 cation-exchange column (Whatman), yielding a purity >98%. The yield of purified protein was 40 and 15 mg per liter of LB medium and M9 medium, respectively. DAD-IV protein uniformly labeled with ^{13}C and ^{15}N was produced by growing the overproducing cells in M9 medium containing $^{15}NH_4Cl$ and [^{13}C-6]-D-glucose. Similarly, ^{15}N-labeled DAD-IV protein was obtained by culturing cells in M9 medium containing $^{15}NH_4Cl$.

Pull-down assay for DnaA box-binding activity. Streptavidin-coated magnet beads (Promega) were washed twice in an equal volume of buffer C containing 100 mM KCl, resuspended into an equal volume of buffer C containing 100 mM KCl and 200 μg/ml bovine serum albumin, and stored on ice. This suspension was used within 2 h for the following experiments. DnaA or DAD-IV protein was added to buffer C (50 μl) containing DNA oligomers (5 pmol) with or without the DnaA box and incubated for 5 min on ice, followed by incubation for 1 h on ice after the addition of a suspension (50 μl) of washed streptavidin-coated magnet beads. DNA and bound protein were recovered by magnetic force, washed in buffer C containing 100 mM KCl and 100 μg/ml bovine serum albumin, and analyzed by tricine–SDS–polyacrylamide (10%) gel electrophoresis [19].

The sequence of the upper strand of the double-stranded (ds) region (29-mer) of the DnaA box-containing oligomer is AAGCCGGA TCCTTGTTATCCACAGGGAG (DnaA box is underlined) [3,5]. The lower strand (29-mer) is complementary to this. When used for the pull-down assay, a 5′-biotin-tagged 30-mer T stretch was added to the 5′-end of the upper strand.

The sequence of the upper strand of 15-mer dsDNA containing the DnaA box is TTGTTATCCACAGGG. The lower strand is complementary to this.

In dsDNA oligomers lacking the DnaA box, the 9-mer DnaA box was replaced with the sequence AACTATATC in the upper strands [5]. The lower strands were complementary to these.

Gel-shift assay for DNA-binding activity. Experiments were generally done according to the method of Schaper and Messer [5]. The indicated amounts of the full-length DnaA or DAD-IV proteins were incubated for 30 min on ice in buffer (20 μl) containing 1 mM ATP and 5 pmol of 15-mer dsDNAs with or without a DnaA box. Sequences of these 15-mer dsDNAs were the same as above. The reaction mixtures were analyzed at 4 °C by native polyacrylamide (8% for the full-length DnaA or 12% for DAV-IV) gel electrophoresis. DNA was detected by staining with Gel-star (BioWhittaker Molecular Applications). Densitometric quantification was done using the computer program, NIH Image (NIH).

CD spectrum analysis. Purified DAD-IV protein was dissolved in buffer containing 5 mM Tris–HCl (pH 7.5), 0.1 mM EDTA, 10% glycerol, and 125 mM or 400 mM KCl, or in the same buffer used for NMR analysis (described below). Protein concentration was calculated by the Bradford method using bovine serum albumin as a standard, or by measuring absorption at 275 nm using an extinction coefficient of 2780 liters/mol/cm which was deduced from the amino acid composition of the DAD-IV protein. CD spectra were recorded at the indicated temperature using quartz cells of 1 mm optical path length with a CD spectropolarimeter JASCO J-720W [17]. The molar ellipticity (θ) expressed in degree × cm^2/dmol was calculated on the basis of a mean residue molecular weight of 110 Da. Dozens of spectra were measured for each sample. A program for the secondary structural analysis, SSE-338W (Japan Spectroscopy), which was based on a method developed by Yang et al. [20], was used for calculation of the structural composition of DAD-IV protein.

NMR analysis. NMR samples typically contained approximately 2 mM DAD-IV protein in 20 mM potassium phosphate buffer (pH 6.0) containing 100 mM KCl and 10 mM DTT in H_2O/D_2O (90%/10%, v/v). All NMR experiments were performed at 30 °C on a Varian Unity INOVA 600 MHz spectrometer equipped with a pulsed-field-gradient unit and a triple resonance probe with an actively shielded *Z*-gradient. Using the unlabeled DAD-IV protein samples, we recorded NOESY spectra, with mixing times of 80 and 150 ms, and TOCSY spectra, with

45 ms mixing times. For the ^{15}N-labeled DAD-IV protein, we acquired the following spectra for assignment purposes: ^{1}H–^{15}N HSQC [21], ^{1}H–^{15}N NOESY–HSQC (150 ms mixing time), and ^{1}H–^{15}N TOCSY–HSQC (50 ms mixing time). HNCA, HN(CO)CA, CBCA(CO)NH, CBCANH, HNCACB, HNCO, HN(CA)CO, C(CO)NH, and HC(CO)NH were used for assignments on the ^{13}C/^{15}N-labeled DAD-IV protein. All spectra were processed with the NMRPipe package [22] on a Sun Microsystems and Silicon Graphics workstation. ^{1}H and ^{15}C chemical shifts were referenced to the methyl proton resonance of an external DSS standard. Chemical shifts of the ^{1}H, ^{13}Cα, ^{13}Cβ, ^{13}C′, and ^{15}N resonance of DAD-IV have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) database under Accession No. 5200.

Results and discussion

Overproduction and purification of DAD-IV protein

For structural analysis of the DNA-binding domain of DnaA, we overproduced and purified a protein consisting of only DAD-IV plus the first methionine residue. The coding region of this module (V374–S467) was inserted into pET-21d(+) (see Materials and methods for details). The resulting plasmid, pIU1, was introduced into BL21(DE3) cells bearing pLysS. When these cells were grown at 37 °C and further incubated for 2 h in the presence of 1 mM ITPG, DAD-IV protein accumulated to make up about 3% of the total cellular protein as measured by SDS–polyacrylamide gel electrophoresis and Coomassie brilliant blue staining.

When wild-type DnaA protein is overproduced, large amounts (about 60–90%) are present as insoluble forms. In contrast, the majority of the DAD-IV protein that we recovered was in the soluble lysate of overexpressed cells. Even when ammonium sulfate was added to a saturating level in this lysate, the DAD-IV protein remained soluble. Through further purification using heparin–agarose affinity chromatography and a Superose 12 gel-filtration column, purity of this protein reached about 98% (Fig. 1A). Determination of the N-terminal sequence by Edman degradation and of the molecular mass using MALDI–TOF-MASS demonstrated that the entire DAD-IV protein was indeed obtained (data not shown). In the final fraction (Fr IV), this protein was present as a monomer (Fig. 1B). Similar results were obtained when this module was gel-filtrated in the presence of 125 mM KCl instead of 400 mM KCl (data not shown).

Activity of DAD-IV protein

Purified DAD-IV protein has a specific binding affinity for the DnaA Box (Fig. 2). In a pull-down assay that we developed for this study, a 29-mer dsDNA fragment with or without a DnaA box (9-mer) derived from the *oriC* DnaA box R1 was tagged with biotin [3,5]. When the tagged DNAs were recovered using streptavidin-coated beads, DAD-IV protein was co-recovered

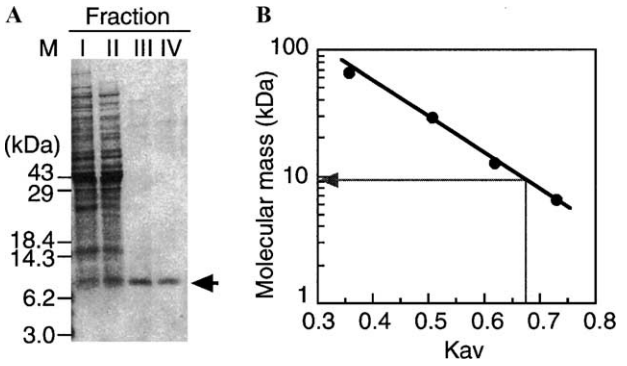

Fig. 1. Purification and gel-filtration analysis of DAD-IV protein. (A) DAD-IV protein was overproduced and purified as described in Materials and methods. Briefly, fraction I (Fr I) was a supernatant of the cell lysates, fraction II (Fr II) was a supernatant obtained from ammonium sulfate (0.34 g/ml) fractionation, fraction III (Fr III) was a pool of active fractions eluted from heparin–agarose column chromatography, and fraction IV was a pool of active fractions obtained from Superose-12 gel-filtration. DnaA box-specific DNA-binding activity of each fraction was assessed by a pull-down assay described in Fig. 2. Proteins (5.0 μg for Fr I and II, 0.5 μg for Fr III and IV) were analyzed by tricine–SDS–polyacrylamide (10%) gel electrophoresis and Coomasie brilliant blue staining. DAD-IV protein is indicated by an arrow. M, molecular mass markers (kDa). (B) DAD-IV protein in fraction III (7.1 μg) was analyzed by a Superose-12 PC3.2/30 gel-filtration column (Amersham Biosciences) equilibrated with buffer C containing 125 mM KCl (see Materials and methods). Eighty to 90% of DAD-IV protein in fraction III was recovered under these conditions. Similar results were obtained when buffer C containing 400 mM KCl or buffer used for CD analysis was used (data not shown). The elution peak of DAD-IV protein and the estimated molecular mass are indicated by a dark line and an arrowhead, respectively. Marker proteins were bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (12.6 kDa), and aprotinin (6.5 kDa).

in a DnaA box-dependent manner (Fig. 2A). Twenty to 25% of the input DAD-IV protein specifically bound to the DnaA box. Similar results were obtained using the wild-type DnaA protein (data not shown). Binding of DAD-IV protein to the DnaA box was specifically inhibited by addition of competitor DNA containing the same DnaA box (Fig. 2B) as well as that of wild-type DnaA (data not shown). This binding competition was seen also when a 15-mer dsDNA containing a 9-mer DnaA box and two 3-mer sequences at the 5′- and 3′-ends was used as a competitor, which also supports DnaA box-specific binding of DAD-IV protein (Fig. 2B).

Binding affinity for the DnaA box was further quantified by gel-mobility retardation assays [5]. The results suggest that the affinity of DAD-IV protein for the DnaA box is only slightly less than that of the full-length DnaA protein (Fig. 3). In the course of this study, Messer et al. [23] reported similar results using surface plasmon resonance analysis. As all of these data support the idea that the isolated DAD-IV protein maintains a conformation that confers specific DnaA-box recognition, we decided to use this protein for structural analyses.

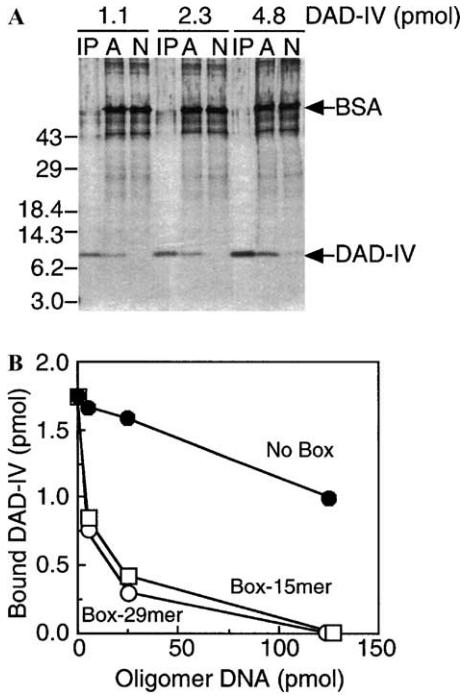

Fig. 2. Purified DAD-IV protein maintains DnaA box-specific DNA-binding activity. (A) DAD-IV protein was used for a pull-down assay using biotin-tagged oligo-DNAs. Indicated amounts of DAD-IV protein were incubated in buffer containing biotin-tagged DNA with or without a DnaA box (see Materials and methods). Streptavidin-coated magnet beads were added to recover the tagged DNA and bound protein. Proteins recovered were analyzed by tricine–SDS–polyacrylamide (10%) gel electrophoresis and silver-staining. IP, the same amount of input DAD-IV protein is shown. The biotin-tagged-DNA bearing a DnaA box (5.0 pmol) in the duplex region (29-mer) (A) or bearing non-specific sequence in the duplex region (29-mer) (N) was used. Migration positions of DAD-IV protein and bovine serum albumin (BSA) included in buffer are indicated. Several minor bands corresponding to 40–60 kDa were derived from the magnetic bead stock buffer and did not affect this assay. Similar results were obtained when 125 mM KCl was included in the reaction (data not shown). (B) Competition experiments showing that DAD-IV maintains DnaA box-specific binding. DAD-IV protein (8.0 pmol) was incubated in buffer containing the biotin-tagged-DNA bearing a DnaA box (5.0 pmol) in the duplex region (29-mer) and the indicated amounts (0, 5.0, 25, and 125 pmol) of 29-mer dsDNA (Box-29mer) with or without DnaA box (No Box). dsDNA (15-mer) bearing a DnaA box (Box-15mer) was also used as a competitor (see Materials and methods). The tagged DNA and bound protein were recovered and analyzed as above. Amounts of DAD-IV protein were determined using a silver-stained gel with quantitative standards.

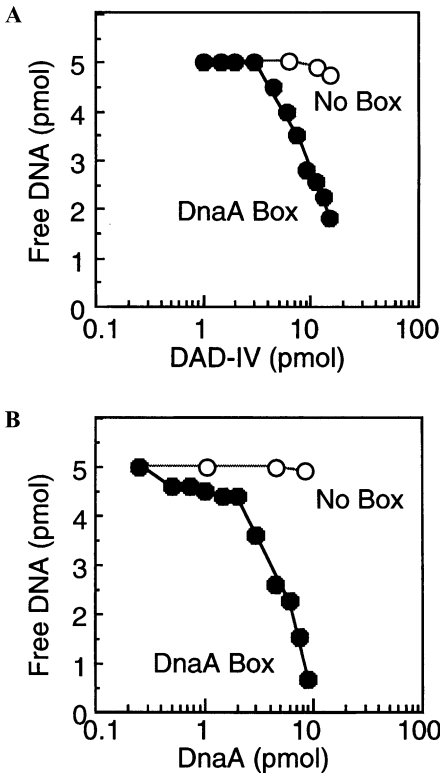

Fig. 3. Affinity of DAD-VI and full-length DnaA to a DnaA box. Indicated amounts of DAD-IV protein (A) or full-length DnaA protein (B) were incubated for 30 min on ice in buffer containing 5 pmol of 15-mer dsDNA with DnaA box (DnaA box) or without DnaA box (No Box) (see Materials and methods). Binding of proteins to DNA was assessed by a gel-mobility retardation assay. Amounts of DNA not bound by protein (Free DNA) were determined using a gel-star-stained gel with quantitative standards. These data are representative of several independent measurements.

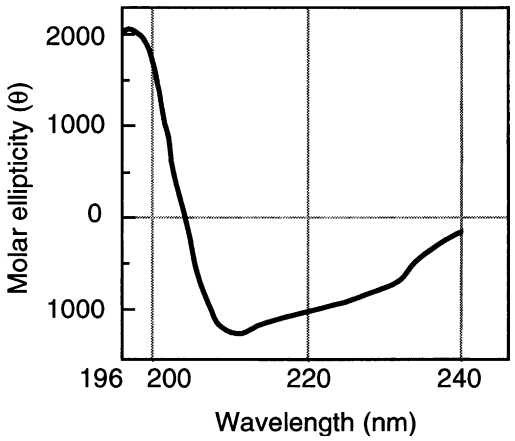

Fig. 4. CD spectra of DAD-IV. CD spectra of DAD-IV protein (1.2 μM) were assessed at 4 °C in buffer containing 125 mM KCl (see Materials and methods). Similar data were obtained when assessed in buffer incubated at 30 °C (data not shown).

CD analysis of the DAD-IV

The secondary structure of DAD-IV protein was analyzed by circular dichroism (CD) spectrum analysis (Fig. 4). Results revealed that domain IV of the DnaA protein has a high content of α-helices and no β-sheets. At 4 °C in the presence of 125 or 400 mM KCl, DAD-IV protein was composed of about 50–53% α-helices, 0% β-sheets, and 47–50% random structures. This composition was not changed by incubation at 30 °C in the presence of 125 mM KCl. In the presence of 400 mM KCl, the composition of α-helices slightly decreased to 41%, but that of β-sheets remained at 0%.

NMR assignment of ^{1}H, ^{13}C, and ^{15}N nuclei of the DAD-IV protein

In the NMR analysis of DAD-IV protein, pulse sequences including HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCA(CO)NH, CBCANH, and HNCACB were used to sequentially assign the backbone and Cβ- and Hβ-atoms (see Materials and methods). This procedure determined a sequential ^{1}H, ^{13}C, and ^{15}N resonance assignment for the majority of the DAD-IV protein backbone. This assignment accounts for most of the main-chain amide cross peaks in the ^{1}H–^{15}N HSQC spectrum (Fig. 5).

Ambiguities in the assignment were well resolved by inspection of the ^{1}H–^{15}N NOESY–HSQC spectrum where sequential NOEs pointed to the amide proton frequency of the adjacent residue. Additionally, NOE connectivities were used to bridge gaps in the assignment that occurred at Pro-406 and Pro-423. Only a few signals were missing in the triple-resonance spectra. It is likely that the lack of observable signals in the triple-resonance spectra primarily resulted from the combination of ^{1}H and ^{15}N line broadening. The chemical shifts of the other side chain atoms (except the Cβ side chain resonance that was assigned during the sequence-specific backbone assignment) were obtained from C(CO)NH and HC(CO)NH experiments recorded on the double-labeled protein. High separations of cross peaks in the ^{1}H–^{15}N HSQC spectrum will enable us to investigate the interaction between DAD-IV and its ligand as was shown in our previous study [24,25].

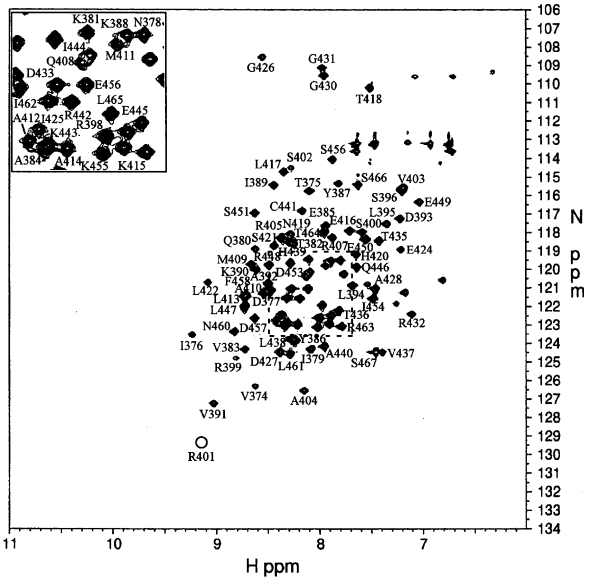

Fig. 5. ^{1}H–^{15}N HSQC spectrum of DAD-IV. ^{1}H–^{15}N HSQC spectrum of 2.0 mM DAD-IV in H_2O/D_2O (90%/10%, v/v), 20 mM potassium phosphate (pH 6.0), 100 mM KCl, and 10 mM DTT at 30 °C is shown. Backbone NH resonance data are labeled with their single letter codes and residue numbers. The region within the pointed square in the middle of the spectrum is magnified in the inset.

Determination of the secondary structure of the DAD-IV protein

The secondary structure of DAD-IV was inferred from a combination of weighted average secondary chemical shifts (WASS [26]) and characteristic NOEs [27]. Based on the $^{1}H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, and $^{13}C'$ chemical shift assignments, the WASS method allowed the delineation of the secondary structure elements. From this analysis, six α-helices were identified along the polypeptide chain (Fig. 6). Additional information about the secondary structure was provided by analysis of sequential and medium-range NOEs involving backbone Hα, side chain Hβ, and amide protons as observed in the ^{1}H–^{15}N NOESY–HSQC experiments (data not shown). These results further suggested that the dominant structural feature of DAD-IV is the presence of six major α-helices.

The combined evidence of the WASS analysis and networks of sequential NOEs between HN(i) and HN(i+1) and medium-range NOEs involving Hα(i), HN(i+3), and HN(i+4) thus determined the region of the six α-helices as follows: Ile-376–Tyr-386 (Helix-A), Lys-390–Leu-394 (Helix-B), Ala-404–Leu-417 (Helix-C), Ser-421–Ala-428 (Helix-D), Asp-433–Glu-450 (Helix-E), and Lys-455–Leu-465 (Helix-F) (Fig. 6). In the full-length DnaA protein, Helix-A may be connected to an amphipathic α-helix that is postulated to be immediately upstream [7].

Structure–function relationship of DAD-IV

Although the secondary structure of DAD-IV has only been postulated using several computer programs,

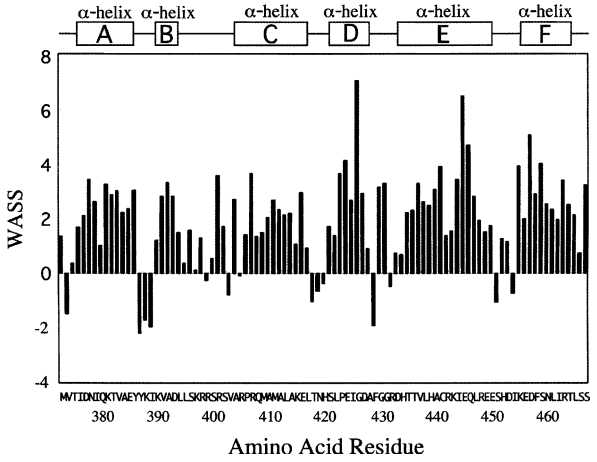

Fig. 6. Determination of DAD-IV secondary structure. Results of the WASS analysis of $^{1}H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, and $^{13}C'$ are shown. When the index number is >1, an α-helical secondary structure is predicted, and when the index number < −1, a β-sheet secondary structure is predicted [26]. The identified secondary structure elements are indicated at the top. Corresponding amino acid residues are shown at the bottom.

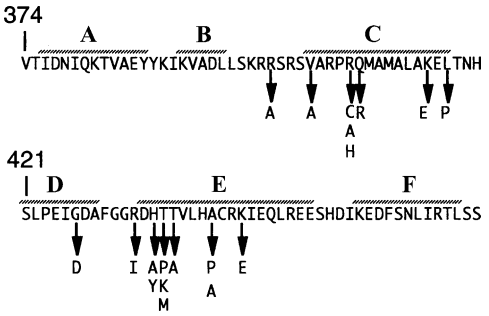

Fig. 7. Structure–function relationship of DAD-IV. Helices-A to -F (A–F) are indicated as lines over the corresponding amino acid sequences of DAD-IV. Single amino acid substitutions that severely reduce DnaA box-binding activity are indicated below the sequence. These mutations are: R339A, V403A, R407C, R407A, R407H, Q408R, K415E, L417P, G426D, R432I, T435P, T435K, T435M, T436A, A440P, A440V, and K443E [9,28]. Mutations that stimulate protein degradation are excluded.

here we have obtained concrete evidence using NMR analysis that DAD-IV contains six α-helices. In DnaA protein, mutational analyses revealed amino acid residues of DAD-IV that dramatically reduce DnaA box-binding activity when substituted with other residues [9,28]. These residues seem to be specifically located in or near Helices-C and -E (Fig. 7). Therefore, these helices most likely play crucial roles in DNA-binding. A group of DNA-binding proteins bears the helix–loop (or turn)–helix motif for direct interaction with DNA [29]. Two helix–loop (or turn)–helix motifs including Helices-C and -E could potentially form a DnaA-specific DNA-binding fold in DAD-IV.

Initiation activity of DnaA protein depends on ATP binding [30–32]. DnaA domain III contains an AAA^+ family-type ATP-binding fold. Binding of ATP, but not ADP, enables this protein to unwind duplex DNA in the *oriC* region. We recently found that in a complex of DnaA protein and pyridoxal–ATP, an ATP analog that contains a pyridoxal moiety bound to the γ-phosphate of ATP, the pyridoxal binds to Lys-415 residue of DnaA protein [33]. This suggests that the γ-phosphate of ATP is located near this residue, which is in Helix-C. As binding of ATP to DnaA protein affects the sequence specificity of DNA-binding [34,35], interaction of ATP could affect the conformation of this helix, leading to a change in the DNA-binding mode.

So far, only a few of the replication initiation protein structures have been solved. The crystal structure of RepE protein, an initiation factor for F plasmid, contains two DNA-binding domains that are topologically similar to each other [36]. Each domain forms a winged-helix motif that includes two or three β-sheets and three α-helices, which is also found as a conserved fold in the DNA-binding domains of Histone H5 and catabolite activating protein (CAP) [36,37] as well as in the putative DNA-binding domains of *Saccharomyces cerevisiae* Orc4, a subunit of the origin-binding protein complex, and of archeal *Pyrobaculum aerophilum* Cdc6/Orc1 protein, which is an Orc4 homolog [38]. DNA-binding cores of viral initiation proteins, including the papillomavirus E2 protein, Epstein–Barr virus EBNA1 protein, and the simian virus 40 large T-antigen, share structural similarity in folds that contain four conserved antiparallel β-sheets and two α-helices [39–42]. The DNA-binding domain of the papillomavirus E1 protein that forms a complex with the origin DNA and the E2 protein also contains four antiparallel β-sheets and a DNA-binding helix [43]. The secondary structure of the DNA-binding domain of DnaA protein, which has been elucidated in this study, is therefore unique overall, although the α-helix-including motif is conserved.

Just before submission of this manuscript, the crystal structure of domains III–IV of the thermophilic bacterium *Aquifex aeolicus* DnaA homolog was published [44]. This DnaA-homolog domain IV structure was basically consistent with the results of our present study on the *E. coli* DAD-IV solution structure. Determination of the three-dimensional structure of DAD-IV bound to DNA will further contribute to the understanding of the mechanisms involved in initiator function.

Acknowledgments

We thank Dr. Kohda for help with NMR processing and the PROI program, Mr. Nemoto for technical support, Dr. Delaglio for the NMRPipe program, and Dr. Kurumizaka for suggestions. This work was supported in part by research grants from the Japan Society for Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology of Japan. M.S. was a recipient of predoctoral fellowships from JSPS.

References

[1] A. Kornberg, T.A. Baker, DNA Replication, second ed., W.H. Freeman and Co., New York, 1992.
[2] W. Messer, C. Weigel, Initiation of chromosome replication, in: F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Schaechter, H.E. Umbarger (Eds.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, second ed., ASM Press, Washington, DC, 2, 1996, pp. 1579–1601.
[3] R.S. Fuller, B.E. Funnell, A. Kornberg, The dnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites, Cell 38 (1984) 889–900.
[4] C. Schaefer, W. Messer, DnaA protein/DNA interaction. Modulation of the recognition sequence, Mol. Gen. Genet. 226 (1991) 34–40.
[5] S. Schaper, W. Messer, Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target, J. Biol. Chem. 270 (1995) 17622–17626.
[6] H. Yoshikawa, N. Ogasawara, Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication origins, Mol. Microbiol. 5 (1991) 2589–2597.
[7] K. Skarstad, E. Boye, The initiator protein DnaA: evolution, properties and function, Biochim. Biophys. Acta 1217 (1994) 111–130.

[8] A. Roth, W. Messer, The DNA binding domain of the initiator protein DnaA, EMBO J. 14 (1995) 2106–2111.
[9] M.D. Sutton, J.M. Kaguni, The *Escherichia coli dnaA* gene: four functional domains, J. Mol. Biol. 274 (1997) 546–561.
[10] C. Weigel, A. Schmidt, H. Seitz, D. Tungler, M. Welzeck, W. Messer, The N-terminus promotes oligomerization of the *Escherichia coli* initiator protein DnaA, Mol. Microbiol. 34 (1999) 53–66.
[11] M.D. Sutton, K.M. Carr, M. Vicente, J.M. Kaguni, *Escherichia coli* DnaA protein: the N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin, J. Biol. Chem. 273 (1998) 34255–34262.
[12] H. Seitz, C. Weigel, W. Messer, The interaction of domains of the DnaA and DnaB replication proteins of *Escherichia coli*, Mol. Microbiol. 37 (2000) 1270–1279.
[13] A.F. Neuwald, L. Aravid, J.L. Spouge, E.V. Koonin, AAA^+: a class of chaperon-like ATPases associated with the assembly, operation, and disassembly of protein complexes, Genome Res. 9 (1999) 27–43.
[14] M. Su'etsugu, H. Kawakami, K. Kurokawa, T. Kubota, M. Takata, T. Katayama, DNA replication-coupled inactivation of DnaA protein in vitro: a role for DnaA arginine-334 of the AAA^+ Box VIII motif in ATP hydrolysis, Mol. Microbiol. 40 (2001) 376–386.
[15] J. Kato, T. Katayama, Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*, EMBO J. 20 (2001) 4253–4262.
[16] S. Nishida, K. Fujimitsu, K. Sekimizu, T. Ohmura, T. Ueda, T. Katayama, A nucleotide switch in the *Escherichia coli* DnaA protein initiates chromosomal replication: evidence from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo, J. Biol. Chem. 277 (2002) 14986–14995.
[17] T. Kubota, T. Katayama, Y. Ito, T. Mizushima, K. Sekimizu, Conformational transition of DnaA protein by ATP: structural analysis of DnaA protein, the initiator of the *Escherichia coli* chromosome replication, Biochem. Biophys. Res. Commun. 232 (1997) 130–135.
[18] H. Ohmori, K. Kimura, T. Nagata, Y. Sakakibara, Structural analysis of the *dnaA* and *dnaN* genes of *Escherichia coli*, Gene (Amst.) 28 (1984) 159–170.
[19] J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, P.T. Wingfield (Eds.), Current Protocols in Protein Science, John Wiley & Sons, Inc., New York, 1995.
[20] J.T. Yang, C.S. Wu, H.M. Martinez, Calculation of protein conformation from circular dichroism, Methods Enzymol. 130 (1986) 208–269.
[21] L.E. Kay, P. Keifer, T. Saarinen, Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, J. Am. Chem. Soc. 114 (1992) 10663–10665.
[22] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (2000) 277–293.
[23] W. Messer, F. Blaesing, D. Jakimowicz, M. Krause, J. Majka, J. Nardmann, S. Schaper, H. Seitz, C. Speck, C. Weigel, G. Wegrzyn, M. Welzeck, J. Zakrzewska-Czerwiska, Bacterial replication initiator DnaA. Roles for DnaA binding and roles of DnaA in origin unwinding and helicase loading, Biochimie 83 (2001) 5–12.
[24] S. Mine, T. Ueda, Y. Hashimoto, Y. Tanaka, T. Imoto, High-level expression of uniformly ^{15}N-labeled hen lysozyme in *Pichia pastoris* and identification of the site in hen lysozyme where phosphate ion binds using NMR measurements, FEBS Lett. 488 (1999) 33–37.
[25] S. Mine, S. Tate, T. Ueda, M. Kainosho, T. Imoto, Analysis of the relationship between enzyme activity and its internal motion using nuclear magnetic resonance: ^{15}N relaxation studies of wild type and mutant lysozyme, J. Mol. Biol. 286 (1999) 1547–1565.
[26] D.S. Wishert, B.D. Sykes, Chemical shifts as a tool for structure determination, Methods Enzymol. 239 (1994) 363–392.
[27] K. Wüthrich, NMR of Proteins and Nucleic acids, John Wiley & Sons, Inc., New York, 1986.
[28] F. Bleasing, C. Weigel, M. Welzeck, W. Messer, Analysis of the DNA-binding domain of *Escherichia coli* DnaA protein, Mol. Microbiol. 36 (2000) 557–569.
[29] R.G. Brennan, B.W. Matthews, The helix–turn–helix DNA-binding motif, J. Biol. Chem. 264 (1989) 1903–1906.
[30] K. Sekimizu, D. Bramhill, A. Kornberg, ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome, Cell 50 (1987) 259–265.
[31] T. Katayama, T. Kubota, K. Kurokawa, E. Crooke, K. Sekimizu, The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase, Cell 94 (1998) 61–71.
[32] K. Kurokawa, S. Nishida, A. Emoto, K. Sekimizu, T. Katayama, Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*, EMBO J. 18 (1999) 6642–6652.
[33] T. Kubota, Y. Ito, K. Sekimizu, M. Tagaya, T. Katayama, DnaA protein Lys-415 is close to the ATP-binding site: ATP–pyridoxal affinity labeling, Biochem. Biophys. Res. Commun. 288 (2001) 1141–1148.
[34] C. Speck, C. Weigel, W. Messer, ATP- and ADP-DnaA protein, a molecular switch in gene regulation, EMBO J. 18 (1999) 6169–6176.
[35] C. Speck, W. Messer, Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA, EMBO J. 20 (2001) 1469–1476.
[36] H. Komori, F. Matsunaga, Y. Higuchi, M. Ishiai, C. Wada, K. Miki, Crystal structure of a prokaryotic replication initiator bound to DNA at 2.6 Å resolution, EMBO J. 18 (1999) 4597–4607.
[37] S.A. MacNeil, Understanding the enzymology of archeal DNA replication: progress in form and function, Mol. Microbiol. 40 (2001) 520–529.
[38] R. Giraldo, R. Dîaz-Orejas, Similarities between the DNA replication initiators of Gram-negative bacteria plasmids (RepA) and eukaryotes (Orc4)/archea (Cdc6p), Proc. Natl. Acad. Sci. USA 98 (2001) 4938–4943.
[39] R.S. Hedge, S.R. Grossman, L.A. Laimins, P.B. Sigler, Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding protein bound to its DNA target, Nature 359 (1992) 505–512.
[40] A. Bochkarev, J. Barwell, R. Pfuetzner, E. Bochkareva, L. Frappier, A.M. Edwards, Crystal structure of the DNA-binding domain of the Epstein–Barr virus origin binding protein EBNA1, bound to DNA, Cell 84 (1996) 791–800.
[41] X. Luo, D.G. Sanford, P.A. Bullock, W.W. Bachovchin, Solution structure of the origin DNA-binding domain of SV40 T-antigen, Nat. Struct. Biol. 3 (1996) 1034–1039.
[42] A.M. Edwards, A. Bochkarev, L. Frappier, Origin DNA-binding proteins, Curr. Opin. Struct. Biol. 8 (1998) 49–53.
[43] E.J. Enemark, A. Stenlund, L. Joshua-Tor, Crystal structure of two intermediates in the assembly of the papillomavirus replication initiation complex, EMBO J. 21 (2002) 1487–1496.
[44] J.P. Erzberger, M.M. Pirruccello, J.M. Berger, The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation, EMBO J. 21 (2002) 4763–4773.