CMLS, Cell. Mol. Life Sci. 60 (2003) 1998–2008
1420-682X/03/091998-11
DOI 10.1007/s00018-003-3176-7

CMLS Cellular and Molecular Life Sciences

Research Article

Identification of the region in *Escherichia coli* DnaA protein required for specific recognition of the DnaA box

Y. Yoshida[a], T. Obita[a,‡], Y. Kokusho[a], T. Ohmura[b], T. Katayama[c], T. Ueda[a,*] and T. Imoto[a,‡‡]

[a] Department of Immunology, Graduate School of Pharmaceutical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582 (Japan), Fax: + 81 92 642 6667, e-mail: ueda@phar.kyushu-u.ac.jp
[b] Accelrys K. K. Shinkawa 2-8-4, Chuo-Ku, Tokyo 104-0033 (Japan)
[c] Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582 (Japan)

Received 6 May 2003; received after revision 18 June 2002; accepted 4 July 2003

Abstract. DnaA protein binds specifically to a 9-base-pair motif called the DnaA box. Domain IV comprises 94 amino acid residues and is required for DNA binding. Using nuclear magnetic resonance analysis, we investigated the interaction between DnaA domain IV and both a DnaA box and a non-specific oligonucleotide that has a reduced affinity for DnaA. The ^{1}H-^{15}N HSQC spectrum of DnaA domain IV showed prominent chemical shift perturbations on six residues (Arg399, Ala404, Leu422, Asp433, Thr435 and Thr436) in the presence of the DnaA box. Through homology modeling, we located all of these residues on one side surface of the DnaA domain IV molecule. Moreover, we compared the chemical shift perturbation of the ^{1}H-^{15}N HSQC spectrum in the presence of the DnaA box with that in the presence of a non-specific oligonucleotide, and the results suggested that Leu422 imparts specificity in binding with the DnaA box.

Key words: DnaA domain IV; DnaA box; homology modeling; interaction; mutation analysis; nuclear magnetic resonance.

DnaA protein plays a key role in the initiation of chromosomal replication in *Escherichia coli* [1, 2]. This protein binds specifically to an asymmetric 9-bp replication sequence, TTATNCACA, which is called the DnaA box. The chromosomal replication origin, *oriC,* contains five DnaA boxes, and binding of DnaA to these sites is a first step in the formation of a specific nucleoprotein complex that causes the initiation reaction of replication. In the *oriC*-DnaA complex, a part of the *oriC* DNA duplex is opened and DnaB helicase is loaded onto the single-stranded DNA, which leads to the loading of DnaG primase and DNA polymerase III holoenzyme. The DNA-loaded form of the β subunit of DNA polymerase III holoenzyme immediately interacts with DnaA in the presence of Hda protein. This protein represses extra initiations by inactivating the DnaA initiation activity [3–6]. Essential to an understanding of the functional mechanisms of DnaA in replicational initiation is to elucidate the interaction between DnaA and the DnaA box. However, *E. coli* DnaA protein has a strong tendency to aggregate [7, 8], which makes it difficult to crystallize not only the full-length DnaA but also the complex of DnaA and the DnaA box.

The DnaA structure is functionally subdivided into four domains [9–11]. The C-terminal domain, domain IV (DAD-IV), contains 94 amino acids, from Val374 to Ser467, and DNA-binding activity [10]. Analysis of the

* Corresponding author.
‡ Present address: Medical Institute of Bioregulation, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582 (Japan)
‡‡ Present address: Department of Engineering, Sojo University, Ikeda 4-22-1, Kumamoto 860-0082 (Japan)

interaction between DAD-IV and the DnaA box should give us information as to how DnaA recognizes the DnaA box.

Recently, we overproduced and purified a DnaA segment containing only DAD-IV [12]. The specific affinity of this protein for the DnaA box is comparable to the affinity of DnaA for the DnaA box. Next, we assigned ^{1}H, ^{13}C and ^{15}N resonances of this DAD-IV by nuclear magnetic resonance (NMR) analyses and identified the secondary structure in solution [12]. As NMR is the most powerful tool to detect a protein-DNA interaction in solution at the atomic level, further investigation using NMR analysis would enable us to detect the fine interaction between DAD-IV and the DnaA box. In the present paper, we use NMR and mutation analysis to report the first detection of the specific interaction between DAD-IV and the DnaA box.

Materials and methods

Preparation of $^{13}C/^{15}N$- and ^{15}N-labeled DAD-IV

$^{13}C/^{15}N$- and ^{15}N-labeled DAD-IV were prepared separately according to our previous paper [12].

Preparation of oligonucleotides

The 13-mer duplex DNA (5′-TGTTATCCACAGG-3′), which is contained in DnaA box R1, and the 13-mer duplex DNA (5′-TGAACTATATCGG-3′), which was shown to interact with DnaA non-specifically [13], were purchased from Invitrogen K. K. (Tokyo, Japan).

Analysis of DNA binding to DAD-IV using NMR spectroscopy

All NMR experiments were performed at 30 °C on a Varian Unity INOVA 600 MHz spectrometer equipped with a pulsed-field gradient unit and triple-resonance probe with an actively shielded Z-gradient. All spectra were processed with the NMRPipe package [14] on a Sun Microsystems workstation or a Silicon Graphics workstation. ^{1}H and ^{13}C chemical shifts were referenced to the methyl proton resonance of an external DSS standard.

To detect the binding sites of DNA duplexes (DnaA box R1 and a non-specific oligonucleotide) on DAD-IV, ^{15}N-labeled protein was dissolved in 400 µl 90% H_2O/10% D_2O, and the unlabeled DNA duplex was added to the protein solution. With addition of the unlabeled DNA duplex to the protein solution, the chemical shifts of many cross-peaks in NMR spectra changed. Unlabeled DNA duplex was added until the chemical shifts ceased. Under these conditions, more than 99% of the protein-DNA complexes were formed in each of the DNA duplexes. The ^{1}H-^{15}N HSQC spectrum of the DAD-IV-13-mer oligonucleotide duplex complexes was compared with that of the free protein. The backbone assignments of DAD-IV in the complexes were determined from the results of CBCANH, HNCA and ^{1}H-^{15}N NOESY-HSQC (150-ms mixing time) experiments, using the $^{13}C/^{15}N$- or ^{15}N-labeled protein.

Homology modeling

The homology study was performed using the homology module of the INSIGHT II molecular modeling package (MSI/Biosym, San Diego, Calif.). Because the primary and secondary structures of domain IV of *Aquifex aeolicus* DnaA protein are similar to those of *E. coli* DnaA protein [15], the atomic coordinates for the former, whose tertiary structure was previously solved by X-ray crystallographic study, were obtained from the Brookhaven Protein Data Bank. For structurally conserved regions, which are located in α-helices common to DAD-IV and domain IV of the *A. aeolicus* DnaA protein, the corresponding coordinates in the domain IV of *A. aeolicus* DnaA protein were used as a structural template. Loop structures connecting structurally conserved regions were selected from ten loop coordinates in the homology module. The initial DAD-IV structure was refined according to the method of our previous work [16]. The DAD-IV structural image shown here was generated using the MOLSCRIPT program [17].

Construction and purification of mutant DAD-IV protein

Mutant DAD-IVs were prepared according to the method of our previous work [12]. The structures of the mutagenic primers used for site-directed mutagenesis to replace Arg399 with Ala399, Ala404 with Gly404, Leu422 with Ala422, Asp433 with Ala433, Thr435 with Ala435 and Thr436 with Ala 436 were 5′-CTTTCCAAGC GTGCGTCCCGCTG-3′, 5′-GCTCGGTGGGCCGTCCG CGC-3′, 5′-CTAACCACAGTGCGCCGGAGATTG-3′, 5′-GTGGCCGTGCGCACACGACG-3′), (5′-CGTGAC CACGCGACGGTGCTTC-3′ and 5′-GACCACACGG CGGTGCTTCATG-3′, respectively. The mutations in the DAD-IV gene were confirmed using a DNA sequence analyzer.

Mutant DAD-IVs were overexpressed from plasmid pIU1-pET21d in the BL21(DE3) cell line. Each pIU1-pET21d was constructed using the pET22b expression vector (Novagen) and the DNA coding sequence for mutant DAD-IV. Mutant DAD-IVs were produced from a cell culture grown in LB medium. Production of mutant DAD-IVs was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 1 mM, when cultures grown at 37 °C in LB medium had reached an $A_{600\ nm}$ of 0.6 units. Soluble fractions of the proteins were obtained by sonication of suspended cells on ice and were purified by cation exchanger (P11) chromatography. No impurities were observed on SDS-PAGE. The proteins were identified by amino acid analy-

sis (Hitachi L-8500) and MALDI-TOF mass spectrometry (PerSeptive Biosystems). The yields of purified proteins were 5–6 mg per liter of culture.

SDS-PAGE of DAD-IVs

Polyacrylamide (15%) gel electrophoresis was conducted according to Laemmli's method [18].

Circular dichroism spectra of DAD-IVs

Purified DAD-IV proteins were dissolved in 0.05 M phosphate buffer containing 0.15 M NaCl at pH 7.2. Protein concentrations were calculated by measuring absorption at 275 nm using an extinction coefficient of 2780 l/mol per centimeter. Circular dichroism spectra were recorded at room temperature using quartz cells of 1-mm optical path length with a circular dichroism spectropolarimeter, JASCO J-720W.

Immobilization of the DnaA box in the cuvette

Biospecific interaction analysis was carried out using a biosensor, IAsys (Fisons), based on the principle of surface plasmon resonance [19, 20]. First, PBS/T buffer [0.05 M phosphate buffer, pH 7.2, containing 0.15 M NaCl and 0.05% (v/v) Tween] was added to the cuvette placed in the instrument, and the solution was kept at 20 °C for 10 min. After the solution had been removed, 3′ biotinated DnaA box R1 (dissolved in 10 mM Tris-HCl buffer containing 1 mM EDTA and 0.1 M NaCl at pH 7.4) was added to the cuvette where streptavidine is linked to the detecting surface. After the solution was again removed, the cuvette was washed with PBS/T buffer.

Kinetic measurement of the interaction between DAD-IVs and the DnaA box

The interaction between DAD-IVs and the DnaA box R1 was monitored at 22 °C as an indication of the change in the surface plasmon resonance response. The association rate for the binding between DAD-IVs and the immobilized DnaA box R1 can be expressed by the following equation,

$$dR/dt = k_{ass}[C](R_{max} - R) - k_{diss}R \quad (1)$$

where k_{ass} is the association rate constant, k_{diss} is the dissociation rate constant, R_{max} is the maximum binding capacity (in resonance units) of the immobilized DnaA box R1 surface (as determined by saturation with DAD-IVs), R is the amount of bound DAD-IV measured as the surface plasmon resonance response (resonance units) at time t, and [C] is the concentration of DAD-IV added to the cuvette (the immobilized DnaA box R1). A linear plot of dR/dt versus R yields

$$\text{slope} = -(k_{ass}[C] + k_{diss}) \quad (2)$$

$$\text{y intercept} = k_{ass}[C]R_{max} \quad (3)$$

with dR/dt obtained from measurements of the slope at multiple time points on the real-time association curve. A new line is obtained by plotting the slopes of the lines for the dR/dt versus R plots as a function of the DAD-IV concentrations [C], and the association rate constant, k_{ass}, can be obtained as the slope. At the same time, the dissociation constant, k_{diss}, can be obtained as the intercept. The dissociation constant, K_d, is then calculated from k_{diss}/k_{ass}.

Protocol for evaluation of k_{ass} and k_{diss} on the interaction between DAD-IVs and an immobilized DnaA box R1

The cuvette was filled with 200 µl of 0.05 M phosphate buffer containing 0.15 M NaCl at pH 7.2 and thermostated at 22 °C for 15 min. DAD-IV solution (0.1–7 µM), in which DAD-IV was dissolved in 0.05 M phosphate buffer containing 0.15 M NaCl at pH 7.2 and pre-incubated at 22 °C, was added to the cuvette (association reaction). The association reaction was monitored for 15 min. The DAD-IV solution was then removed with a sucker, and 200 µl of 0.05 M phosphate buffer containing 0.15 M NaCl at pH 7.2, which had been pre-incubated at 22 °C, was immediately added to the cuvette (dissociation reaction). The dissociation reaction was monitored for 10 min. To remove DAD-IVs that had bound to the matrix on the cuvette surface, the immobilized DnaA box was regenerated by replacement with a 10 mM HCl solution after the association and dissociation reactions. After the DnaA box was regenerated, the cuvette was refilled as soon as possible with 0.05 M phosphate buffer containing 0.15 M NaCl at pH 7.2. DnaA box R1 solution was added to the cuvette again before the next measurement.

Molecular dynamics simulation

For the simulations of the extended conformations of the respective tripeptides, the values of the phi and psi angles were set to 180 and 180°. The molecular modeling software, Discovery Studio Modeling version 1.0 (Accelrys Inc., San Diego, Calif.), was used to construct and display the molecules. All force-field calculations were performed with DS-CHARMm using a non-bonded cut-off distance of 12 Å. All molecular structures were minimized in vacuum using the Adopted-Basis Newton-Raphson algorithm until energy tolerance was zero. All molecular dynamics simulations were performed in vacuum at 300 K using 50,000 time steps of 0.001 ps. The coordinates were recorded every 100 steps.

Results

Binding of duplex of DnaA box R1 to DAD-IV

The DnaA box R1 (5′-TCCTTGTTATCCACAGGG-CAG) is located in the *oriC* region and is known to have a most potent affinity for DnaA with an dissociation con-

stant (K_d) of 0.9 nM [13]. DNA-IV also binds to the DnaA box R1 as a monomer with high specificity and high affinity [10, 12]. To identify the binding site of DAD-IV to DnaA box R1, we performed chemical shift perturbation experiments using NMR spectroscopy. The ^{1}H-^{15}N HSQC spectrum has been frequently used to map the binding surfaces of protein-protein, protein-nucleic acid and protein-ligand interactions. In this spectrum, resonances are usually well resolved because of large ^{15}N chemical shift dispersion. Because each amino acid residue gives one peak in this spectrum, monitoring the chemical shift perturbation of each amino acid residue is easy. In the absence of DNA, we measured the ^{1}H-^{15}N HSQC spectrum of DAD-IV and found that the resonances were usually well resolved (fig. 1, in red) [12]. A ^{15}N-labeled DAD-IV was titrated with an unlabeled double-stranded oligonucleotide containing DnaA box R1. Selective chemical shift changes were observed for signals in the ^{1}H-^{15}N HSQC spectrum upon protein-DNA complex formation. As is often seen in strong interactions, a slow exchange mode was exhibited at the NMR time scale.

To detect the binding sites of DNA duplexes (DnaA box R1) on DAD-IV, the unlabeled DNA duplex was added to the protein solution. The final concentration was 0.5 mM for the protein and 1.2 mM for the DNA duplex. Figure 1 (in blue) shows the ^{1}H-^{15}N HSQC spectrum obtained by mixing DAD-IV and the DnaA box DNA in a 1:2.4 stoichiometry. The spectrum assignment of the DAD-IV-duplex DNA complex was determined from the results of HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, HNCACB and ^{1}H-^{15}N NOESY-HSQC spectra, using either the ^{15}N- or the $^{13}C/^{15}N$-labeled DAD-IV-DnaA box complex. Of the 92 expected ^{1}H-^{15}N cross-peaks, 91 were identified in the ^{1}H-^{15}N HSQC spectrum of the complex. This backbone assignment provided a good database for comparing the structural differences between the free and complexed forms of DAD-IV. Thus, the chemical shift of $^{13}C\alpha$ atoms exhibited by DAD-IV in its complex with the DnaA box indicated that elements of the secondary structure of the protein in the complex were essentially maintained as those of the protein alone (data not shown). This finding, when taken together with sequential and short-range NOE connectivities of the protein in complex with

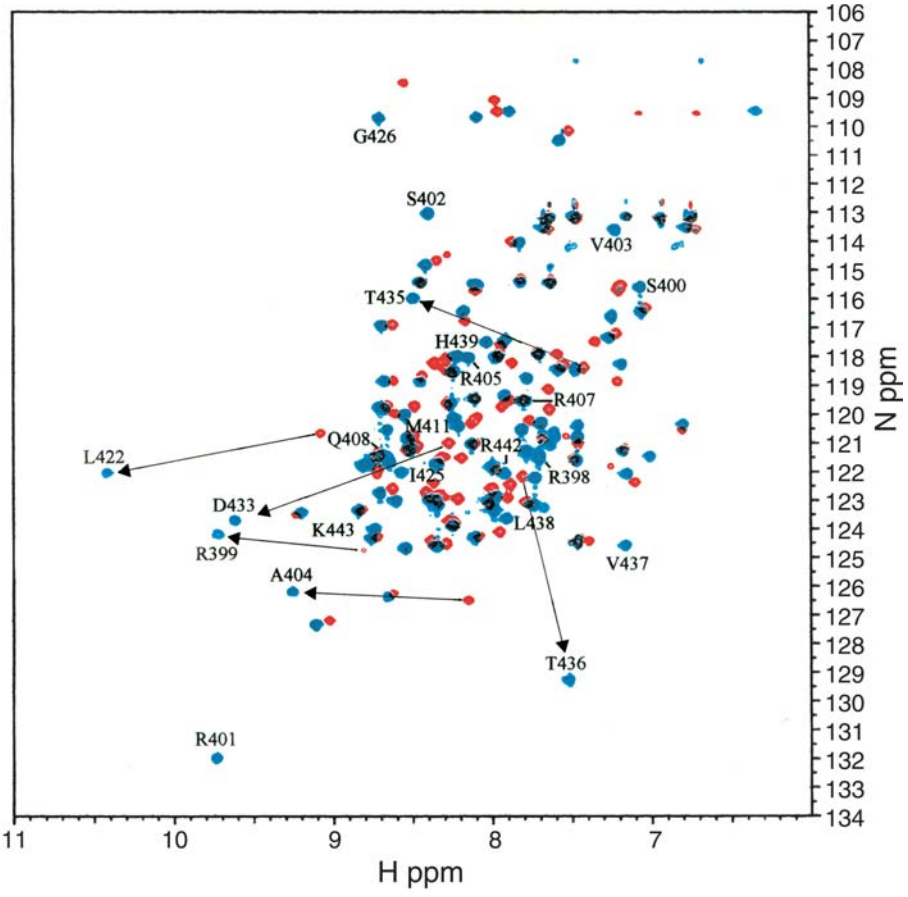

Figure 1. Superposition of ^{1}H-^{15}N HSQC spectra of free ^{15}N-labeled DAD-IV, and complexed with unlabeled DNA box R1 duplex. The protein/DNA ratio in the complex is 1:2.4. The cross-peaks of free protein are shown in red, and those in the complex are shown in blue. Only peaks that were affected by complex formation are labeled. Prominent chemical shift changes of cross-peaks are marked by arrows.

the DnaA box DNA, led us to conclude that the strong interaction between the protein and its ligand DNA had no significant effects on the secondary structure of the protein.

Upon the formation of a protein-DNA complex, specific chemical shift perturbation was observed in the ^{1}H-^{15}N HSQC spectrum of DAD-IV. When a protein and a ligand form a complex, their interactions change the amino acid environment at the interfaces, resulting in chemical shift changes. Any small additional conformational changes near the direct-contact surfaces will cause additional chemical shift perturbations. Thus, the surface mapped by chemical shift perturbation contained and extended beyond the direct binding surface. In addition, the residues that had considerable chemical shift changes were usually located within the binding interface.

As shown in fig. 2 A, the chemical shifts of most residues of DAD-IV were little affected, indicating that complex formation between the protein and DNA did not cause large conformational changes in the protein. However, for a few residues, the complex formation caused significant chemical shifts. The shifts at residues Arg399, Ala404,

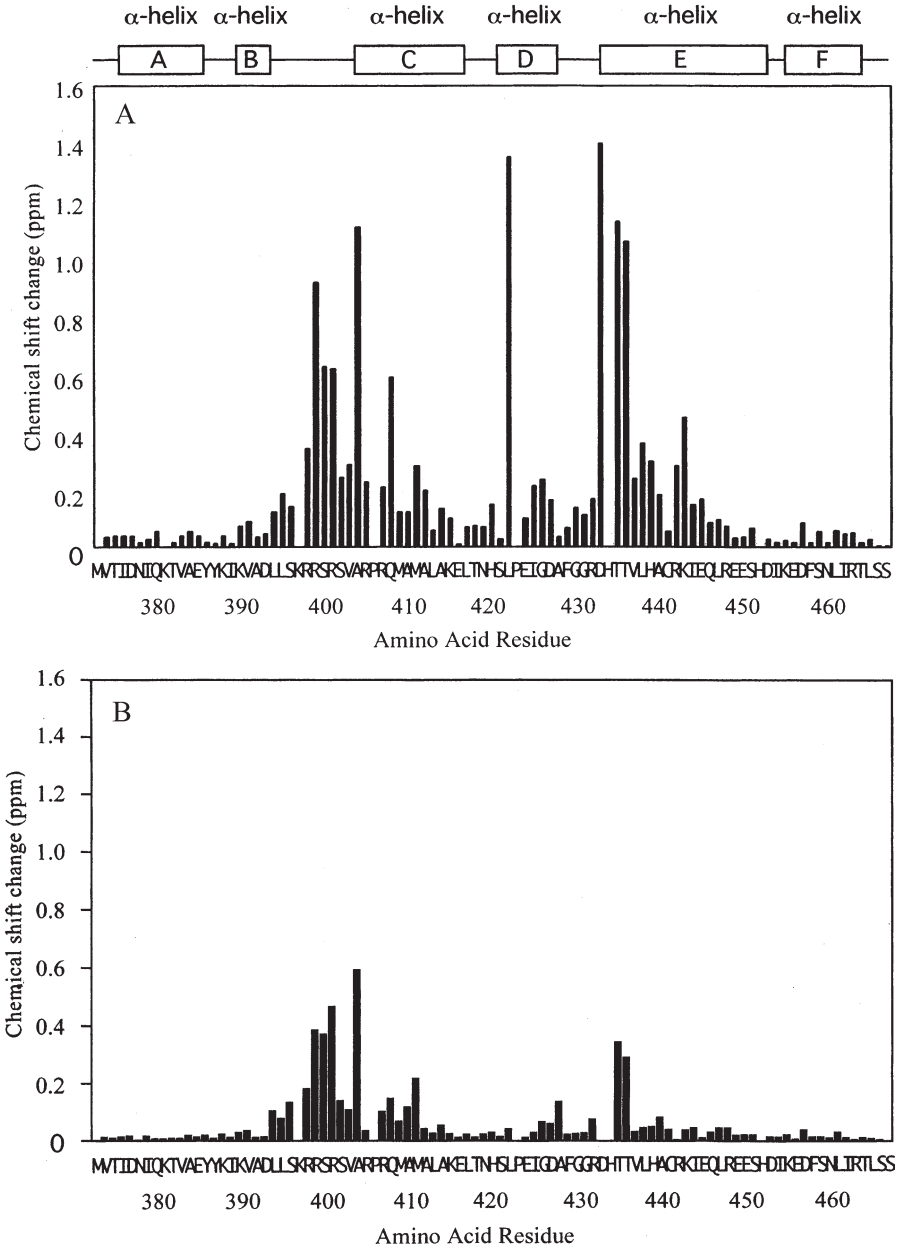

Figure 2. Chemical shift changes of backbone amides of DAD-IV upon binding to DNA duplex as a function of the amino acid residue of the protein. A combined index of the chemical shift change of each amide cross-peak was calculated according to the equation, $[\Delta\delta(^{1}H)^{2} + (\Delta\delta(^{15}N)/7)^{2}]^{1/2}$. (*A*) Binding of DAD-IV to DnaA box R1, (*B*) Binding of DAD-IV to non-specific oligonucleotide.

Leu422, Asp433, Thr435 and Thr436 were each more than 0.8 ppm (fig. 2A). Therefore, these residues were expected to be located in the binding interface with the DnaA box.

Homology modeling of DAD-IV

To predict the locations of the residues that would be highly affected by complex formation in the three-dimensional structure of DAD-IV, we constructed a DAD-IV model. We found that the primary amino acid sequence and the DAD-IV secondary structure, both of which were obtained by weight-average secondary chemical shift analysis [12], showed significant homology with the corresponding region (domain IV) from the *A. aeolicus* DnaA protein (fig. 3). The tertiary structure of domains III and IV of this protein has been solved by X-ray crystallography [15]. For our DAD-IV model, we therefore used homology modeling based on the tertiary structure of domain IV of the *A. aeolicus* DnaA protein.

The residues that were highly affected by complex formation (Arg399, Ala404, Leu422, Asp433, Thr435 and Thr436) were plotted on the predicted structure (fig. 4). These residues were all located on the same side surface, suggesting that this region was the binding interface and that these residues were involved in the interaction with the DnaA box DNA.

Mutation analysis on binding between DAD-IV and the DnaA box

To confirm that the highly affected residues (Arg399, Ala404, Leu422, Asp433, Thr435 and Thr436) were involved in the interaction between DAD-IV and the DnaA box DNA, we introduced amino acid substitution into DAD-IV. The following substitutions were prepared as described in Materials and methods: R399A (in which Arg399 is mutated to Ala), A404G (Ala404 mutated to Gly), L422A (Leu422 mutated to Ala), D433A (Asp433 mutated to Ala), T435A (Thr435 mutated to Ala) and T436A (Thr436 mutated to Ala). Figure 5 shows that wild-type DAD-IV and its mutant proteins were highly purified in the final preparations. Amino acid compositions and TOF-mass spectrometry also indicated that the mutant DAD-IVs were homogeneous (data not shown). By measuring their circular dichroism spectra (data not shown), we confirmed that all mutations barely affected the global structures of the mutant DAD-IVs. We determined the kinetic parameters of the interactions between the DAD-IVs and the immobilized duplex DNA carrying the DnaA Box R1.

```
A. aeolicus 306 M---QIVEFVANYYAVKVEDILSDKRNKRTSEARK
                h   hhhhhhhhh    hhhh         hhhhh
E. coli     374 VTIDNIQKTVAEYYKIKVADLLSKRRSRSVARPRQ
                  hhhhhhhhhhh  hhhhhhh        hhhhh

            338 IAMYLCHKVCSASLIEIARAFKRKDHTTVIHAIRS
                hhhhhhhhhh   hhhhh          hhhhhhh
            409 MAMALAKELTNHSLPEIGDAFGGRDHTTVLHACRK
                hhhhhhhhh    hhhhhhh     hhhhhhhhhh

            373 VEEEKKKDRKFKHLVGFLEKQAFDKIC
                hhh             hhhhhhhhhh
            444 IEQLREESHDIKEDFSNLIRTLSS
                hhhhhhh  hhhhhhhhhhhhh
```

Figure 3. Homology between domain IV of DnaA from *A. aeolicus* (upper sequence) and domain IV of DnaA from *E. coli* (lower sequence). The residue shown by the letter h consists of α-helical structure.

Using the FAST FIT program, we analyzed the response curves obtained when various amounts of DAD-IV were added to the cuvette, and found that all association or dissociation phases were attributable to monophasic reactions. Figure 6A–G plots the slope (dR/dt vs R) against the concentrations of the wild-type and mutant DAD-IV added to the cuvette. These plots show good straight lines. The slope and the y intercept give k_{ass} and k_{diss} , respectively, of a DAD-IV with the immobilized DnaA box R1. The determined values are shown in table 1. The table also shows the dissociation constants, K_d, of the DAD-IV-immobilized DnaA box R1 complexes, which were obtained as the ratios of k_{ass} to k_{diss}. The dissociation constants, K_d, of the mutant DAD-IV-immobilized DnaA box R1 complexes were all larger than that of the wild-type DAD-IV-immobilized DNA, indicating that the most affected residues in NMR analysis were involved in the interface between DAD-IV and DnaA box R1. Among all mutant DAD-IVs, the mutation of R399A gave the largest K_d (table 1).

Non-specific DNA binding of DAD-IV

A 13-mer DNA (5′-TGAACTATATCGG-3′) was used as a representative non-specific DNA. This DNA is known to bind to DnaA protein very weakly ($K_d > 200$ nM) [13]. The interaction between DAD-IV and this non-specific oligonucleotide was investigated through NMR. The unlabeled non-specific oligonucleotide was added to the protein solution. The final concentration of protein was 0.3 mM, and that of the duplex DNA was 1.5 mM. Under these conditions, almost all of the protein formed protein-DNA complexes. After the addition of the unlabeled non-specific oligonucleotide, chemical shifts of many cross-peaks in ^{1}H-^{15}N HSQC spectrum changed. The chemical shift change at each residue were calculated according to the equation, $[\Delta\delta(^{1}H)^2 + (\Delta\delta(^{15}N)/7)^2]^{1/2}$. The residues with a chemical shift change more than 0.2 ppm, which is nearly an average of the chemical shift changes in the experiment, were considered to be involved in the interaction with the non-specific oligonucleotide. Residues of Arg399, Ala404, Thr435 and Thr436, which were significantly perturbed upon complex formation with DnaA box R1, showed chemical shift changes upon complex formation with this DNA, whereas the cross-peak from Asp433 disappeared (fig. 2B). Thus, these residues were

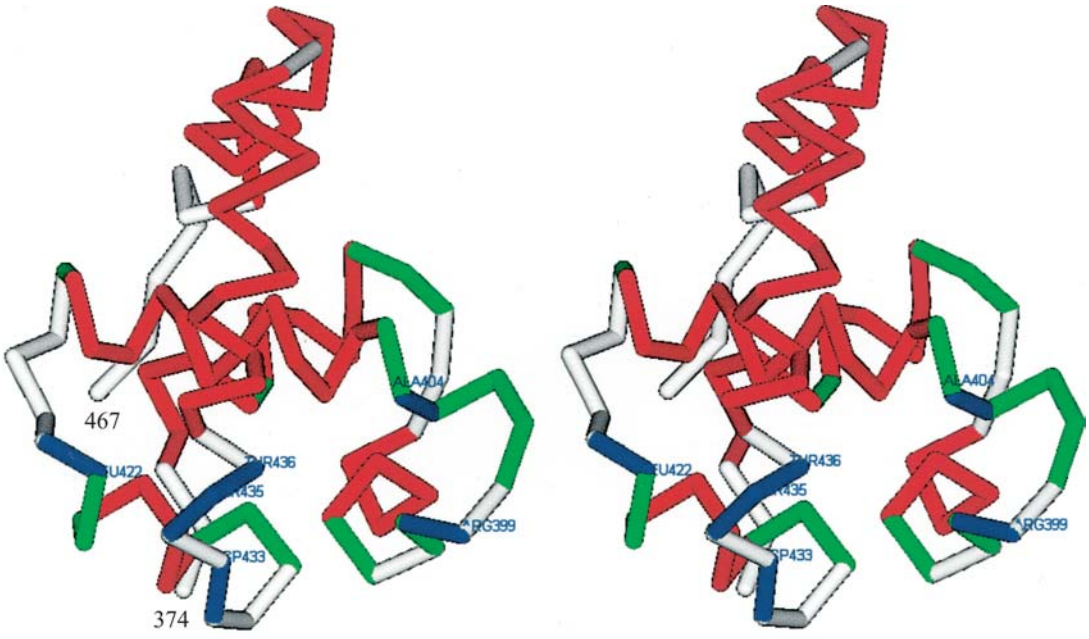

Figure 4. Stereoview of ribbon drawing of the structure of DAD-IV from *E. coli,* obtained by homology modeling methods (see Materials and methods). Helix regions are colored in red and turn regions are colored in green. The residues involved in the interaction with the DNA box R1 duplex are colored in blue.

Table 1. Thermodynamic and kinetic parameters for the interaction between DAD-IVs and immobilized DnaA box R1 DNA at pH 7.2 and 22 °C.

	Wild-type	R399A	A404G	L422A	D433A	T435A	T436A
$k_{ass} \times 10^{-3}$ ($s^{-1}M^{-1}$)	15.0	0.86	3.3	4.7	3.0	3.4	3.5
$k_{diss} \times 10^{3}$ (s^{-1})	13	25	12	12	17	21	16
$K_d \times 10^{6}$ (M)	0.86	32.0	3.5	2.5	5.2	6.3	3.8
$\Delta\Delta G^a$ (kcal/mol)	–	2.1	0.8	0.6	1.1	1.2	0.9

[a] The difference in free energy change between $\Delta G_{wild\text{-}type}$ and ΔG_{mutant}, which is calculated from K_d. The deviations of kinetic parameters were estimated to be within 10%.

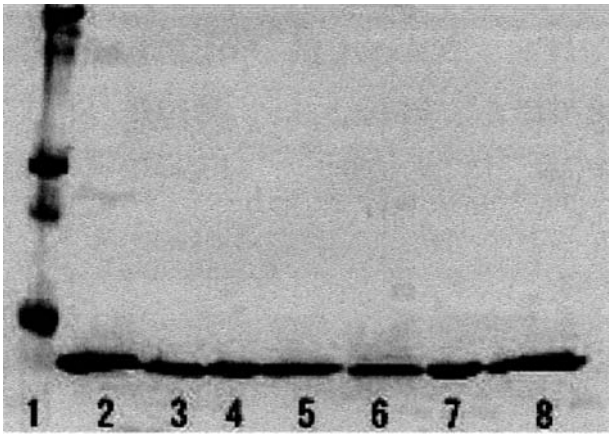

Figure 5. SDS-PAGE of DAD-IV and its mutants. Lane 1, marker; lane 2, wild-type; lane 3, R399A; lane 4, A404G; lane 5, L422A; lane 6, D433A; lane 7, T435A; lane 8, T436A.

likely involved in non-specific oligonucleotide binding. However, Leu422, which was one of the most perturbed residues in specific binding for DnaA box R1, was hardly affected with the non-specific oligonucleotide. Thus, Leu422 probably imparted specificity in binding with DnaA box R1.

Molecular dynamics simulation of tripeptides

To further examine the mutation from Leu422 to Ala, we conducted the following experiment. As in the primary sequence (fig. 3), the leucine residue was followed by a proline residue. The backbone conformation at any given residue followed by the prolyl residue is known to be restricted [21]. Therefore, we conducted molecular dynamic simulation of the tripeptide (Ser-Leu-Pro) including the preceding residue of the leucine residue and that of the tripeptide (Ser-Ala-Pro) as a control (see Materials

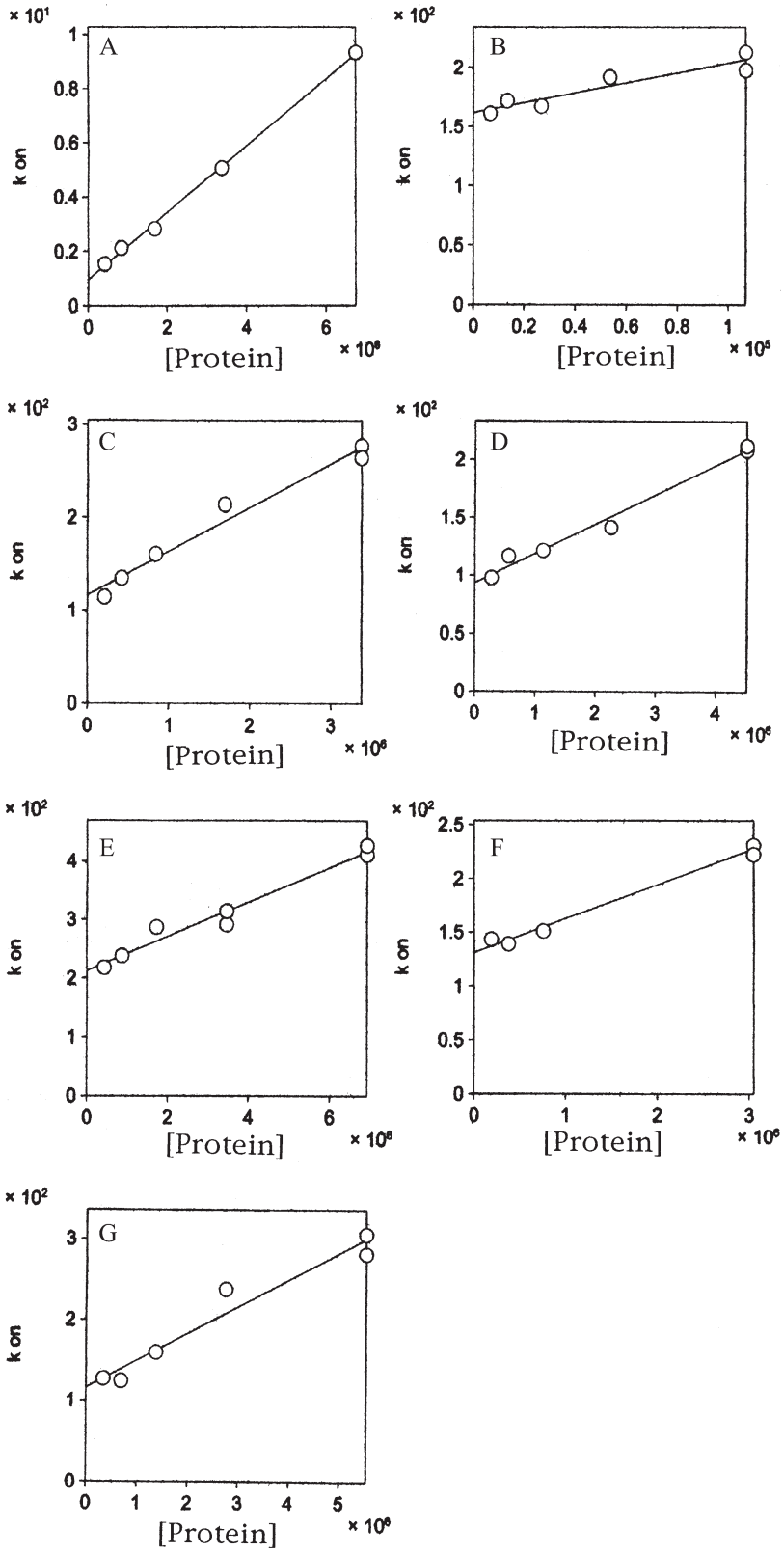

Figure 6. Kinetic measurement of the interactions of DAD-IVs with the immobilized DnaA box R1. The relationship between the slopes and the concentrations of DAD-IV added to the cuvette. The details of the procedure for the kinetic analysis are given in Materials and methods. (*A*) Wild-type. (*B*) R399A. (*C*) A404G. (*D*) L422A. (*E*) D433A. (*F*) T435A. (*G*) T436A.

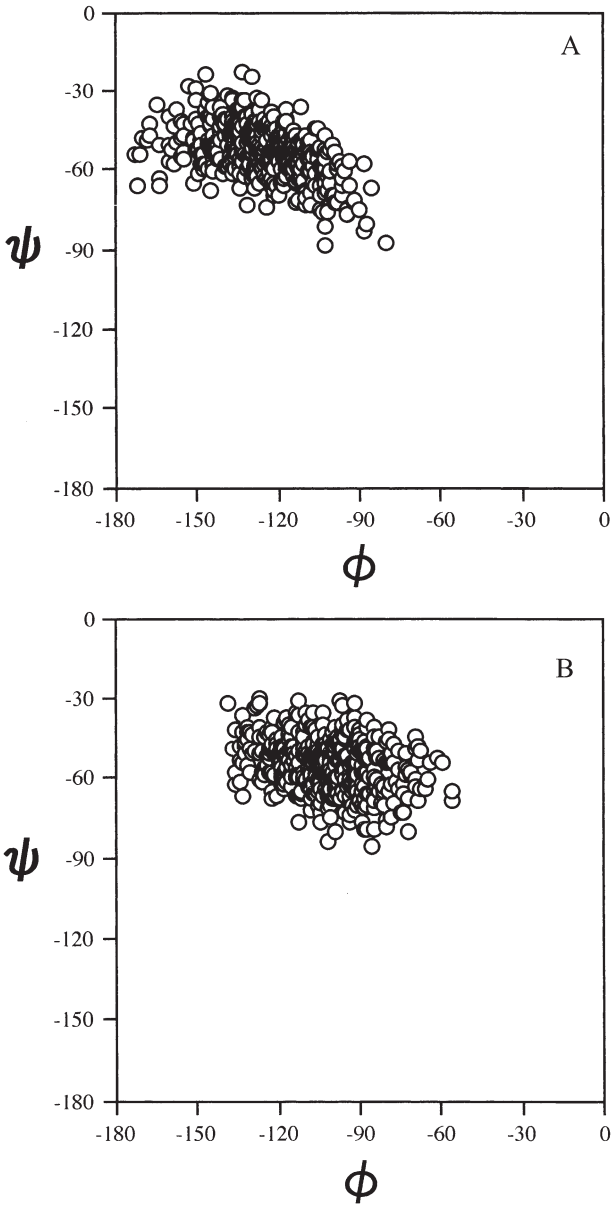

Figure 7. Positions of conformation of the center residue in the tripeptide at each 100 steps during molecular dynamic simulation are shown on a Ramachandran plot. The details of the procedure are given in Materials and methods. (*A*) Ser-Ala-Pro. (*B*) Ser-Leu-Pro.

and methods). From the resulting 500 conformations in each tripeptide, we plotted the phi and psi angles at the center residues of the Ser-Ala-Pro and Ser-Leu-Pro peptides as shown in figure 7A and 7B, respectively. The averages of these possible phi and psi angles were –125.2 ± 10.2 and –53.1 ± 7.7, respectively, at that of the Ser-Ala-Pro peptide and –101.7 ± 10.9 and –55.4 ± 8.0, respectively, at the center residue of the Ser-Leu-Pro peptide. This result suggested that the possible phi angle at the center residue differed from tripeptide to tripeptide, which would mean that the backbone conformation at Leu in Ser-Leu-Pro could be different from that at Ala in Ser-Ala-Pro.

Discussion

Interaction of DAD-IV with DnaA box R1

By using NMR experiments and mutation analysis, the residues Arg399, Ala404, Leu422, Asp433, Thr435 and Thr436 in DAD-IV are nominated as candidates for involvement in the interaction with DnaA box R1. Based on the secondary structure (fig. 2) and homology modeling (fig.4) of DAD-IV, we located Arg399 and Ala404 at the loop between the B helix and C helix and at the N-terminal of the C helix (region I); Leu422 is located at the N-terminal of the D helix (region II), and Asp433, Thr435 and Thr436 are located at the N-terminal of the E helix (region III). These residues are located at three regions that are spatially separated but that are also located on the same side surface of the DAD-IV molecule (fig. 4). Therefore, these residues are those most likely to be involved in the interaction with DnaA box R1.

Previous genetic and biochemical analyses identified some of the residues required for DnaA to bind with the DnaA box [10, 11, 22, 23]. Consistent with this study, Arg399 and Val403 in region I as well as Thr435 and Thr436 in region III were previously shown to be crucial to DnaA box binding. Asp433 is a member of a highly conserved motif among bacterial DnaA homologues and was suggested to be important in DnaA box binding [14, 23]. In the present study, using D433A DAD-IV, we obtained the first biochemical evidence to support the importance of Asp433. Moreover, we suggest for the first time that Leu422 in region II plays a unique role in interaction with the DnaA box.

As for surface plasmon resonance analysis, the association rate constant for the interaction between R399A and the immobilized DnaA box R1 was much slower than the rate constants of the wild-type and the other mutants with the immobilized DnaA box R1 (table 1). This indicates that the positive charge of the side chain of Arg399 may interact with phosphate ion in the sugar-phosphate backbone. On the other hand, Asp433 has negative charges. This residue is unlikely to interact with the negatively charged phosphate backbone of a DNA. The side chain carbonyl group of Asp is capable of forming hydrogen bonds with a cytidine base [24]. Mutation from Asp433 to Ala in DAD-IV decreased binding affinity by about 1 kcal/mol against DnaA box R1, indicating that Asp433 may be responsible for recognition of a base, e.g. the cytosine, in DnaA box R1.

Mutation Leu422 to Ala in DAD-IV gave a slight but distinct reduction in binding affinity between DnaA box R1 and DAD-IV (table 1). The extent of the reduction in binding affinity in L422A was smaller than that in the

other mutants (table 1). To analyze this small discrepancy, we simulated the molecular dynamics of tripeptides (Ser-Ala-Pro and Ser-Leu-Pro). The average of the possible phi angles at the center residue in each tripeptide clearly changed by replacing a leucine residue with an alanine residue (fig. 7). Therefore, the slight reduction in binding affinity in the mutation from Leu422 to Ala in DAD-IV can be attributed to a conformational change in the backbone at Leu422 in DAD-IV. These findings, taken together with the finding that severe chemical shift perturbation in the ^{1}H-^{15}N HSQC spectrum in the presence of DnaA box R1, indicate that DnaA box R1 does not interact with the side chain but rather with the backbone at Leu422 in DAD-IV.

The mechanism of DAD-IV binding differs between non-specific DNA and DnaA box R1

DAD-IV preferentially binds DnaA box R1 DNA, but binding to non-specific oligonucleotide is extremely weak [20, 23]. In chemical shift perturbation experiments of DAD-IV in the presence of non-specific oligonucleotides, we found that the residues Arg399, Ala404, Thr435 and Thr436 were involved in this DNA interaction, although the extent of this chemical shift perturbation was smaller than that seen in the presence of DnaA box R1. In the present measurement, the cross-peak of Asp433, which is located at region III, disappeared in the ^{1}H-^{15}N HSQC spectrum. This may be due to multiple conformations in its binding to DNA. And since we observed the small chemical shift perturbations of Thr435 and Thr436, which are also located in region III, we can say that the region III in DAD-IV weakly interacted with the non-specific oligonucleotide. In contrast, the chemical shift perturbation of Leu422 in DAD-IV, which was severely affected in the presence of DnaA box R1, was not observed in the presence of non-specific oligonucleotide. These results suggested that non-specific oligonucleotide bound weakly to regions I and III, which include Arg399, Ala404, Thr435 and Thr436, but that it did not bind to region II, which includes Leu422. In a homology model, Leu422 is located at the N-terminal of the D helix, which is spatially apart from regions I and III, where residues Arg399, Ala404, Asp433, Thr435 and Thr436 are located (fig. 4). Therefore, the interaction involved in residue 422 could be unique among interactions with the DnaA box R1. Yoshikawa and Ogasawara [9] showed the alignment of the primary sequences of DAD-IV among various bacteria. Although those bacteria were genetically separate from each other, the leucine residue in question was highly conserved. This finding supported the idea that Leu422 is important in the interaction between DnaA and the DnaA box.

Biological meaning of the interaction of DAD-IV with DnaA box R1

A structural model of the *A. aeolicus* DnaA homologue-DNA complex has been constructed, based on the structure of the *trp* repressor-DNA complex [14]. The tertiary structure of the *trp* repressor has a certain similarity to that of DAD-IV. This model suggests that regions I and III of DAD-IV interact with DNA and their finding is consistent with the findings of the present study. Thus, these features may be common among DnaA homologues in interaction with DNA. In contrast, the model does not predict the interaction of region II with DNA. This may reflect a difference in biological function between *trp* repressor and DnaA. Specifically, the *trp* repressor is a mere DNA-binding protein, whereas DnaA binds to DNA, followed by DNA-unwinding reactions.

We identified the amino acid residues which interacted with DnaA box R1 (fig. 2A), mutated the amino acid residues and evaluated the dissociation constants of mutant DnaA domain IVs and DnaA box R1. As a result, the dissociation constants between them were all larger than the dissociation constant between the wild-type DnaA domain IV and DnaA box R1 (table 1). Thus, DnaA box R1 does appear to interact with DnaA domain IV at residues Arg399, Ala404, Leu422, Asp433, Thr435 and Thr436. Arg399 and Ala404 are located in region I, Leu422 is located in region II and Asp433, Thr435 and Thr436 are located in region III. Based on the homology model (fig. 4), when DnaA box R1 simultaneously bound to region I-III, it was suggested that DnaA box R1 should bend to some extent. Schaper and Messer demonstrated that one DnaA monomer was found to bind to a DnaA box and to induce a bend of about 40° [13]. The present finding was consistent with the result of the previous reports [13, 25]. In the present study, DAD-IV recognized DnaA box R1 duplex at three regions (I, II and III) on the molecule but recognized oligonucleotide including Nonsence box duplex at two regions (I and III) on the molecule. A stereoview of the binding sites on the DAD-IV molecule indicates that a DnaA box R1 duplex may kink when it is bound simultaneously to three regions on the DAD-IV molecule, whereas a oligonucleotide including Nonsence box duplex is unable to because it is bound to two regions on the DAD-IV molecule (fig. 4). Therefore, the specific interaction between region II, containing Leu422, and DnaA may participate in the kinking of the DNA duplex in the initiation of DNA replication.

Blaesing et al. [22] have already shown that Asp433 and Thr436 are involved in DNA binding from results of the mutations of Asp433 to Ala and Thr436 to Ala [22]. However, to date, residues around Leu422 had not been shown to be involved in a specific DNA contact. Thus, the finding that Leu422 is involved in a specific DNA contact may contribute to understanding the specific DNA binding of DnaA domain IV.

Acknowledgements. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan. We also thank KN-international (USA) for improvement of our English.

1 Kornberg A. and Baker T. A. (1992) DNA Replication, 2nd edn, Freeman, New York
2 Messer W. and Weigel C. (1996) Initiation of chromosome replication. In: *Escherichia coli and Salmonella*: Cellular and Molecular Biology, 2nd edn, pp. 1579–1601, Neidhardt F. C., Curtiss R. III, Ingraham J. L., Lin E. C. C., Low K. B., Magasanik B. et al. (eds), ASM Press, Washington, DC
3 Katayama T., Kubota T., Kurokawa K., Crook E. and Sekimizu K. (1998) The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase. Cell **94:** 61–71
4 Kurokawa K., Nishida S., Emoto A., Sekimizu K. and Katayama T. (1999) Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. EMBO J. **18:** 6642–6652
5 Kato J. and Katayama T. (2001) Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. EBMO J. **20:** 4253–4262
6 Katayama T. (2001) Feedback controls restrain the initiation of *Escherichia coli* chromosomal replication. Mol. Microbiol. **41:** 9–17
7 Sekimizu K., Yung B. Y.-M. and Kornberg A. (1988) The *dnaA* protein of *Escherichia coli*: abundance, improved purification, and membrane binding. J. Biol. Chem. **263:** 7136–7140
8 Kubota T., Katayama T., Ito Y., Mizushima T. and Sekimizu K. (1997) 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:** 130–135
9 Yoshikawa H. and Ogasawara N. (1991) Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication on origins. Mol. Microbiol. **5:** 2589–2597
10 Roth A. and Messer W. (1995) The DNA binding domain of the initiator protein DnaA. EMBO J. **14:** 2106–2111
11 Sutton M. D. and Kaguni J. M. (1997) The *Escherichia coli dnaA* gene: four functional domains. J. Mol. Biol. **274:** 546–561
12 Obita T., Iwura T., Su'etsugu M., Yoshida Y., Tanaka Y., Katayama T. et al. (2002) Determination of the secondary structure in solution of the *Escherichia coli* DnaA DNA-binding domain. Biophys. Biochem. Res. Commun. **299:** 42–48
13 Schaper S. and Messer W. (1995) Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. J. Biol. Chem. **270:** 17622–17626
14 Delaglio F., Grzesiek S., Vuister G. W., Zhu G., Pfeifer J. and Bax A. (2000) NMR Pipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR **6:** 277–293
15 Erzberger J. P., Pirrucello M. M. and Berger J. M. (2002) The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. EMBO J. **21:** 4763–4773
16 Nishida S., Fujimitsu K., Sekimizu K., Ohmura T., Ueda T. and Katayama T. (2002) 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:** 14986–14995
17 Koradi R., Billeter M. and Wuthrich K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. **14:** 51–55
18 Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685
19 Gorgani N. N., Parish C. R., Easterbrook Smith S. B. and Altin J. G. (1997) Histidine-rich glycoprotein binds to human IgG and C1q and inhibits the formation of insoluble immune complexes. Biochemistry **36:** 6653–6662
20 Krebs B., Griffin H., Winter G. and Rose-John S. (1998) Recombinant human single chain Fv antibodies recognizing human interleukin-6. J. Biol. Chem. **273:** 2858–2865
21 Schimmel P. R. and Flory P. J. (1968) Conformational energies and configuration statistics of copolypeptides containing L-proline. J. Mol. Biol. **34:** 105–120
22 Bleasing F., Weigel C., Welzeck M. and Messer W. (2000) Analysis of the DNA-binding domain of *Escherichia coli* DnaA protein. Mol. Microbiol. **36:** 557–569
23 Sutton M. D. and Kaguni J. M. (1997) Threonine 435 of *Escherichia coli* DnaA protein confers sequence-specific DNA binding activity. J. Biol. Chem. **272:** 23017–23024
24 Suzuki M., Brenner S. E., Gerstein M. and Yagi N. (1995) DNA recognition and superstructure formation by helix-turn-helix proteins. Protein Eng. **4:** 319–328
25 Messer W. (2002) The bacterial replication initiator DnaA. DnaA and *ori*C, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev. **26:** 355–374