Twelve Species of the Nucleoid-associated Protein from
Escherichia coli

SEQUENCE RECOGNITION SPECIFICITY AND DNA BINDING AFFINITY*

(Received for publication, June 28, 1999, and in revised form, August 2, 1999)

Talukder Ali Azam‡§ and Akira Ishihama‡¶

From the ‡Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan and the §Japan Science and Technology Corporation, Kawaguchi, Saitama 332-0012, Japan

The genome of Escherichia coli is composed of a single molecule of circular DNA with the length of about 47,000 kilobase pairs, which is associated with about 10 major DNA-binding proteins, altogether forming the nucleoid. We expressed and purified 12 species of the DNA-binding protein, i.e. CbpA (curved DNA-binding protein A), CbpB or Rob (curved DNA-binding protein B or right arm of the replication origin binding protein), DnaA (DNA-binding protein A), Dps (DNA-binding protein from starved cells), Fis (factor for inversion stimulation), Hfq (host factor for phage Qβ), H-NS (histone-like nucleoid structuring protein), HU (heat-unstable nucleoid protein), IciA (inhibitor of chromosome initiation A), IHF (integration host factor), Lrp (leucine-responsive regulatory protein), and StpA (suppressor of td− phenotype A). The sequence specificity of DNA binding was determined for all the purified nucleoid proteins using gel-mobility shift assays. Five proteins (CbpB, DnaA, Fis, HIF, and Lrp) were found to bind to specific DNA sequences, while the remaining seven proteins (CbpA, Dps, Hfq, H-NS, HU, IciA, and StpA) showed apparently sequence-nonspecific DNA binding activities. Four proteins, CbpA, Hfq, H-NS, and IciA, showed the binding preference for the curved DNA. From the apparent dissociation constant (Kd) determined using the sequence-specific or nonspecific DNA probes, the order of DNA binding affinity were determined to be: HU > HIF > Lrp > CbpB(Rob) > Fis > H-NS > StpA > CbpA > IciA > Hfq/Dps, ranging from 25 nM (HU binding to the non-curved DNA) to 250 nM (Hfq binding to the non-curved DNA), under the assay conditions employed.

The genome DNA of Escherichia coli is associated with a core set of 10–20 DNA-binding proteins, altogether forming the nucleoid (for review, see Ref. 1). These nucleoid-associated proteins, hereafter referred to as the “nucleoid proteins” in this paper, have long been considered to be structural proteins setting the overall DNA conformation, not only by wrapping or packaging of DNA but also by introducing bending or coiling (for example, see Ref. 2). The association of these nucleoid proteins, however, influences not only the conformation but also the functions of DNA such as replication, recombination, repair, and transcription. For instance, the template activity of DNA for transcription is either activated or repressed by the association of these nucleoid proteins to various extents depending on the gene (for example, see Ref. 3). The overall activities of the genome are considered to vary depending on cell growth conditions, because the composition of nucleoid proteins changes depending on cell growth conditions or growth phases (4). These observations taken together suggest the pleiotropic regulatory roles for the nucleoid proteins in global regulation of gene transcription.

In addition to these nucleoid proteins, a total of about 100 transcription factors interact with specific sequences near gene promoters and act as either activators or repressors of transcription of a gene or a set of genes (5). Recently, the molecular events leading to transcription regulation by at least some sequence-specific and mostly gene-specific transcription factors have been characterized in detail (for instance, see Ref. 5). In contrast, relatively little is known about transcription regulation by the nucleoid proteins. These group proteins have been thought to have the activities of sequence-nonspecific DNA-binding, but no systematic and comparative studies have been performed of the specificity and affinity of DNA binding for the nucleoid proteins.

In order to understand the overall configuration and physiological activities of the E. coli genome under various growth conditions and the role(s) of each nucleoid protein, we performed for the first time a systematic comparison of the recognition sequence specificity and the DNA-binding affinities among 12 species of the nucleoid protein from E. coli, including the abundant nucleoid proteins in growing E. coli cells, i.e. factor for inversion stimulation (Fis), H-NS (histone-like nucleoid structuring protein), HU (heat-unstable nucleoid protein), IHF (integration host factor), and Lrp (leucine-responsive regulatory protein), and a set of curved DNA-binding proteins, i.e. CbpA (curved DNA-binding protein A), CbpB (curved DNA-binding protein B; or Rob, right origin binding protein), and StpA (suppressor of td− phenotype A) (6–15). In addition, we extended our analysis to include two DNA-associated proteins, DnaA (DNA-binding protein A) and IciA (inhibitor of chromosome initiation A) with the regulatory activity of DNA replication (16, 17), a protein, called Hfq (host factor for phage Qβ replication), with the binding activity to both nucleoid and ribosomes (18), and a novel DNA-binding protein, Dps (DNA-binding protein from starved cells), which is synthesized only

* This work was supported by grants-in-aid from the Ministry of Education, Science and Culture of Japan, and CREST (Core Research for Evolutional Science and Technology) from the Japan Science and Technology Corporation (JST). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel: 81-559-81-6741; Fax: 81-559-6846; E-mail: aishihama@lab.nig.ac.jp.

1 The abbreviations used are: Fis, factor for inversion stimulation; CbpA, curved DNA-binding protein A; CbpB, curved DNA-binding protein B; DnaA, DNA-binding protein A; Dps, DNA-binding protein from starved cells; Hfq, host factor for phage Qβ; H-NS, histone-like nucleoid structuring protein; HU, heat-unstable nucleoid protein; IciA, inhibitor of chromosome initiation A; IHF, integration host factor; Lrp, leucine-responsive regulatory protein; StpA, suppressor of td− phenotype A; bp, base pair(s).
in the stationary phase of cell growth and plays a role in growth-dependent transformation of the nucleoid configuration (4). Among these 12 proteins, the specificity of DNA recognition has not been reported for Dps, Hfq, IciA, IpaH, and StpA. Here the affinity and sequence specificity of DNA binding were examined in parallel for all these 12 species of DNA-binding proteins under the same conditions by gel-mobility shift assays using various DNA probes.

**Experimental Procedures**

**Bacterial Strains, Plasmids, and Media—** *E. coli* strains and plasmids used for expression of the DNA-binding proteins are shown in Table I. Cells were grown in LB medium.

**Reagents and Chemicals—** Restriction endonucleases were purchased from Takara Shuzo (Japan). P11 phosphocellulose was purchased from Whatman, while WM-Sepharose, DEAE-Sepharose, heparin-Sepharose CL-68, Mono-Q, Mono-S, and Sepharose 12 were products of Amersham Pharmacia Biotech.

**Protein Purification—** All 12 DNA-binding proteins were purified from overexpressed protein. The plasmids and the conditions used for expression of the DNA-binding proteins are summarized in Table I. The steps of protein purification and the yield of proteins are summarized in Table II. The purity of proteins was analyzed by SDS-polyacrylamide gel electrophoresis followed by staining of gels with Coomassie Brilliant Blue R (Kodak). Protein concentration was measured by a staining method with Coomassie Blue dye (19) and using bovine serum albumin as a standard.

**Preparation of Synthetic DNA Probes—** Six different kinds of short DNA duplexes (see Fig. 2 for the sequence) were prepared by annealing two single-stranded oligonucleotides which were synthesized with a DNA synthesizer (Applied Biosystems, Model 394 DNA/RNA Synthesizer). For duplex formation, oligonucleotides in 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, and 1 mM EDTA were heated at 85 °C for 15 min and cooled slowly to room temperature. DNA was extracted with phenol/chloroform and then precipitated with ethanol. For radiolabeling, the oligonucleotide (SO) probe were synthesized, each carrying a 5′-32P-labeled free 3′-OH.

**Gel Mobility Shift Assays—** Gel mobility shift assays were carried out for the detection of complex formation between DNA-binding proteins and DNA probes. The procedure was essentially the same for 12 DNA-binding proteins. Mixtures of 32P-labeled (0.5–2.0 nM) or unlabeled probe DNA (5.0 nM) and various amounts of the test protein were incubated in 10 μl of binding buffer (10 mM Tris-HCl buffer, pH 7.5, containing 1 mM EDTA, 100 mM NaCl, and 1 mM dithiothreitol) for 25 min at 25 °C. The reaction mixtures were directly subjected to electrophoresis on 5% polyacrylamide gel (acrylamide:bisacrylamide, 38:2) in 0.5% TBE buffer consisting of 89 mM Tris borate (pH 8.0) and 1 mM EDTA. The electrophoresis was carried out at a constant voltage of 60 V at 25 °C (the same temperature as that used for DNA-protein complex formation). Gels were exposed to x-ray films or imaging plates. The amounts of unbound free and protein-bound DNA probes were determined by measuring the exposed films or plates with a NIH Image Software (version 1.61) or a PhosphorImager (Molecular Dynamics). The apparent dissociation constant \( K_d \) was determined by measuring the protein concentrations which gave 50% binding of the input probe (27).

Gel shift assays were also carried out using a mixture of HaeIII-treated pUC19. The reaction products were analyzed by electrophoresis on 1% agarose gels, and the amounts of unbound free and protein-bound DNA were determined by staining DNA with ethidium bromide.

**Results—** Purification of 12 Species of the *E. coli* DNA-binding Protein—As an initial effort for the systematic comparison of DNA binding properties among the nucleoid proteins in *E. coli*, we analyzed in this study 12 species of DNA-binding proteins, i.e. CbpA, CbpB (or Rob), DnaA, Dps, Fis, Hfq, H-NS, IciA, IpaH, Lrp, and StpA. The molecular properties so far identified for these proteins are summarized in Table III. For protein purification, the genes for these nucleoid proteins were transiently expressed at high levels under control of strong and inducible promoters (Table I). In the case of HU and IpaH, both consisting of two different subunits with similar sequences, the genes for two subunits are expressed simultaneously in the same cells, and thus the main components of HU and IpaH proteins used in this study were both heterodimers.

The overexpressed proteins were mostly recovered in soluble fractions of the respective cell lysates except Dps which was solubilized from inclusion bodies with a buffer containing 6 M urea and then renatured. Thus, the native conformation must have been retained for most of the proteins examined, suggesting that the specificities and activities herewith described represent those of native proteins. All the test proteins were purified to apparent homogeneity in two or three steps of chromatography essentially according to the published procedures (Table II; for details see “Experimental Procedures”). A total of 0.5–15 mg of proteins were obtained from 1 liter of the induced cultures. Fig. 1 shows the patterns of SDS-polyacrylamide gel electrophoresis for each step of the purification for all these 12 proteins. The purity at the final step was more than 95% for all these proteins except for DnaA.

Construction of DNA Probes for the Gel Shift Assay—To measure the DNA binding activities for all 12 DNA-binding proteins by the gel shift assay, six different types of synthetic oligonucleotide (SO) probe were synthesized, each carrying a unique sequence as described in Fig. 2. For simultaneous and quantitative comparison of the DNA binding properties for many proteins under the same condition, the size of the DNA probes were adjusted to be within the range of 40–64 bp. Two types of DNA probe were prepared for the assay of nonspecific DNA-binding proteins. SOA contains six consecutive (da)6 stretches, each being connected by a 4-bp (CGGC) interval, in a total length of 60 bp. This type DNA is known to form a curved configuration and has been used for binding assays of the curved DNA-binding proteins such as CbpA and H-NS (11, 28). SOB of 64 bp in length, consisting of equal amounts of GC and AT, contains two tandem repeats of the 26-bp long CbpB-

---

**Table I**

Expression of *E. coli* DNA-binding proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Plasmid</th>
<th>Host bacteria</th>
<th>Promoter</th>
<th>Induction&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Inducer</th>
<th>Time&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Plasmid source</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbpA</td>
<td>pCU180</td>
<td>C211L</td>
<td>tac</td>
<td>60</td>
<td>1 mM IPTG</td>
<td>2 h</td>
<td>T. Mizuno (Nagoya)</td>
</tr>
<tr>
<td>CbpB</td>
<td>pMK19</td>
<td>MV1184</td>
<td>tac</td>
<td>50</td>
<td>1 mM IPTG</td>
<td>4 h</td>
<td>T. Mizuno (Nagoya)</td>
</tr>
<tr>
<td>DnaA</td>
<td>pS5Y67</td>
<td>D33</td>
<td>lac</td>
<td>40</td>
<td>0.3 mM IPTG</td>
<td>4 h</td>
<td>S. Yasuda (Mishima)</td>
</tr>
<tr>
<td>Dps</td>
<td>pDPS1</td>
<td>BL21(DE3)</td>
<td>T7</td>
<td>30</td>
<td>1 mM IPTG</td>
<td>2 h</td>
<td>This laboratory</td>
</tr>
<tr>
<td>Fis</td>
<td>pF1077</td>
<td>BL21(DE3/fis)</td>
<td>T7</td>
<td>70</td>
<td>1 mM IPTG</td>
<td>1 h</td>
<td>R. C. Johnson (Los Angeles)</td>
</tr>
<tr>
<td>Hfq</td>
<td>pHFQ607</td>
<td>BL21(DE3)</td>
<td>T7</td>
<td>45</td>
<td>0.4 mM IPTG</td>
<td>5 h</td>
<td>This laboratory</td>
</tr>
<tr>
<td>H-NS</td>
<td>pHON11</td>
<td>CH252</td>
<td>tac</td>
<td>70</td>
<td>0.7 mM IPTG</td>
<td>4 h</td>
<td>T. Mizuno (Nagoya)</td>
</tr>
<tr>
<td>HU</td>
<td>pLupAhupB</td>
<td>N48301</td>
<td>αLp</td>
<td>50</td>
<td>42 °C</td>
<td>3 h</td>
<td>N. Goshima (Hiroshima)</td>
</tr>
<tr>
<td>IciA</td>
<td>pIC1S1</td>
<td>MC1061</td>
<td>araB</td>
<td>30</td>
<td>0.75% l-arabinose</td>
<td>4 h</td>
<td>D. S. Hwang (Seoul)</td>
</tr>
<tr>
<td>IpaH</td>
<td>pSA5hiphimA</td>
<td>A6740</td>
<td>αLp</td>
<td>50</td>
<td>42 °C</td>
<td>4 h</td>
<td>A. B. Oppenheim (Jerusalem)</td>
</tr>
<tr>
<td>Lrp</td>
<td>pMWD1</td>
<td>DL1552 (BL21/DE3)</td>
<td>T7</td>
<td>50</td>
<td>0.5 mM IPTG</td>
<td>3 h</td>
<td>D. Low (Salt Lake City)</td>
</tr>
<tr>
<td>StpA</td>
<td>pT7apa</td>
<td>BL21(DE3)</td>
<td>T7</td>
<td>40</td>
<td>0.75 mM IPTG</td>
<td>4 h</td>
<td>M. Belfort (Albany)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Turbidity measured with a Klett-Summerson photometer.

<sup>b</sup> Culture time after the induction.

<sup>c</sup> IPTG, isopropyl-1-thio-β-D-galactopyranoside.

---

*EXPERIMENTAL PROCEDURES*
binding site (13) connected by a 6-bp spacer with CCCGGG sequence (for the sequence see Fig. 2). This set of curved SOA and non-curved SOB was used for the gel mobility shift assay of the sequence nonspecific DNA-binding proteins (HU, IciA, and StpA) or those with no known specificity (CbpA, Dps, Hfq, and H-NS).

For the DNA binding assays of other proteins with sequence-specific recognition activities, probes SOC (60 bp), SOD (54 bp), SOE (40 bp), and SOF (62 bp) were synthesized, which contained three repeats of the DnaA box on the same DNA surface (29), three repeats of the 15-bp long Fis-binding site (6), a single 15-bp long IHF-binding site (30), and three repeats of the DnaA box on the same DNA surface (31), respectively.

**DNA Mobility Shift Assays in Vitro of 12 DNA-binding Proteins**—Gel mobility shift assays were performed for all 12 nucleoid proteins using both specific and nonspecific DNA probes with the chain length of 40–64 bp (for the probe sequences see Fig. 2). Mixtures of 0.5–3.0 nm $^{32}$P-labeled probe DNA and various amounts of one of the nucleoid proteins were incubated in 10 μl of the binding buffer (10 mM Tris-HCl buffer, pH 7.5, containing 1 mM EDTA, 100 mM NaCl, and 1 mM dithiothreitol), and then subjected to electrophoresis on native polyacrylamide gels. Typical patterns of the gel shift assays using 2.0 nM and then subjected to electrophoresis on native polyacrylamide gel.

**TABLE III**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Number of amino acids</th>
<th>Molecular mass kDa</th>
<th>Native protomer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbpA</td>
<td>297</td>
<td>33.4</td>
<td>Homodimer</td>
</tr>
<tr>
<td>CbpB (Rob)</td>
<td>289</td>
<td>33.0</td>
<td>Monomer</td>
</tr>
<tr>
<td>DnaA</td>
<td>467</td>
<td>53.0</td>
<td>Monomer</td>
</tr>
<tr>
<td>Dps</td>
<td>167</td>
<td>19.0</td>
<td>Monomer/dodecamer</td>
</tr>
<tr>
<td>Fis</td>
<td>98</td>
<td>11.2</td>
<td>Homodimer</td>
</tr>
<tr>
<td>Hfq</td>
<td>102</td>
<td>11.2</td>
<td>Monomer/hexamer</td>
</tr>
<tr>
<td>H-NS</td>
<td>137</td>
<td>15.4</td>
<td>Homodimer</td>
</tr>
<tr>
<td>HU</td>
<td>90</td>
<td>9.2</td>
<td>Heterodimer</td>
</tr>
<tr>
<td>HU-1</td>
<td>90</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>IciA</td>
<td>297</td>
<td>33.5</td>
<td>Homodimer</td>
</tr>
<tr>
<td>IHF</td>
<td>99</td>
<td>11.2</td>
<td>Heterodimer</td>
</tr>
<tr>
<td>IHFe</td>
<td>94</td>
<td>10.7</td>
<td></td>
</tr>
<tr>
<td>IHFβ</td>
<td>167</td>
<td>19.0</td>
<td>Homodimer</td>
</tr>
<tr>
<td>Lrp</td>
<td>133</td>
<td>15.3</td>
<td>Homodimer (?    )</td>
</tr>
</tbody>
</table>

$^a$ AS, ammonium sulfate; INC, inclusion bodies; PEI, polyethyleneimine; PC, phosphocellulose; CM-S, carboxymethyl-Sepharose; DEAE, DEAE cellulose; PCS, phosphocellulose slurry; G-DEAE, protein pak G-DEAE; H-S, heparin-Sepharose; S-12, Superose 12; M-S, Mono S; M-Q, Mono Q; CM-C, carboxymethyl-cellulose.
sequence-specific DNA-binding proteins in the assay system employed. No significant cooperativity was observed for this group of proteins in DNA binding, because the test proteins bound to the probe DNAs with multiple binding sites in a stepwise fashion.

The CbpB bound to the two CbpB-binding sites of SOB in a stepwise fashion, forming two retardation bands (Fig. 3A, lanes 3–12). The apparent dissociation constant (Kd) of the first site for CbpB was 67 nM (Figs. 4 and 5A), and approximately 120 nM CbpB was required for the shift of all input SOB probe (Fig. 3A, lane 10). Above 140 nM CbpB, all the input SOB were converted to complexes with two molecules of the CbpB protein (Fig. 3B, lane 11). When SOE without the CbpB site was used as a DNA probe, no DNA retardation was observed even at the concentration as high as 80 nM (Fig. 3A, lane 2). The apparent Kd of 213 nM for DnaA measured using probe SOC with three tandem DnaA-binding sequences (Fig. 3B, lanes 3–12) was the highest among the sequence-specific DNA-binding proteins tested (Figs. 4 and 5A), but this was partly attributed to aggregation of the purified DnaA protein, in particular at high protein.

FIG. 1. SDS-polyacrylamide gel electrophoresis of E. coli DNA-binding Proteins. Twelve species of the DNA-binding proteins were purified from E. coli as summarized in Table II. Each step fraction (1 μl aliquot) of the protein purification was analyzed by SDS-gel electrophoresis on 10% (DnaA), 12.5% (CbpA, CbpB, IciA), or 16.5% (Dps, Fis, Hfq, H-NS, HU, IHF, Lrp, and StpA) gels. Gels were stained with Coomassie Brilliant Blue (CBB). The following proteins were used as molecular mass markers: phosphorylase b, 97.4 kDa; bovine serum albumin, 66.2 kDa; ovalbumin, 45.0 kDa; carbonic anhydrase, 31.0 kDa; soybean trypsin inhibitor, 21.5 kDa; and lysozyme, 14.4 kDa. Sup, supernatant; AS, ammonium sulfate fraction; Ppt, precipitates; PC, phosphocellulose column fraction; PCS, phosphocellulose slurry fraction; G-DEAE, protein pak G-DEAE; and UCE, uninduced cell extract.

FIG. 2. DNA probes for gel shift assays. Six different kinds of duplex DNA probes of 40–64 bp in length were synthesized for gel mobility shift assays of the DNA-binding proteins. The complete sequence of each probe is shown. Probe SOA includes six repetitions of AT clusters, while probes SOB to SOF include the consensus sequences for the DNA-binding proteins (CbpB, DnaA, Fis, IHF, and Lrp) as underlined.
Fig. 3. Gel shift assays of the DNA-binding proteins. The reaction conditions for probe DNA-protein complex formation and electrophoresis are described under “Experimental Procedures.” The gel patterns are classified into three groups depending on the nature of DNA recognition: Panels A to E, sequence-specific DNA-binding proteins; Panels F to I, curved DNA-binding proteins; and Panels J to L, sequence nonspecific DNA-binding proteins. For each assay, the following 32P-labeled DNA probe was used at 2.0 nM. Panel A (CbpB), lanes 1 and 2, control SOE; and lanes 3–12, probe SOB. Panel B (DnaA), lanes 1 and 2, control SOB; and lanes 3–12, probe SOC. Panel C (Fis), lanes 1 and 2, control SOB; and lanes 3–12, probe SOD. Panel D (IHF), lanes 1–2, control SOB; and lanes 3–11, probe SOE. Panel E (Lrp), lanes 1 and 2, control SOB; and lanes 3–12, probe SOF. Panel F (CbpA), lanes 1–7, probe SOA; and lanes 8–14, probe SOB. Panel G (H-NS), lanes 1–12, probe SOA; and lanes 13–24, probe SOB. Panel H (Hfq), lanes 1–7, probe SOA; and lanes 8–14, probe SOB. Panel I (IciA), lanes 1–7, probe SOA; and lanes 8–14, probe SOB. Panel J (Dps), lanes 1–7, probe SOA; and lanes 8–14, probe SOB. Panel K (HU), lanes 1–7, probe SOA; and lanes 8–14, probe SOB. Panel L (StpA), lanes 1–7, probe SOA; and lanes 8–14, probe SOB.
concentrations. DnaA did not form complexes with a control probe (SOB) without the DnaA-binding sequence (Fig. 3B, lanes 1 and 2).

Fis also showed the stepwise formation of three types of complexes with SOD containing three repeats of the Fis-binding core sequence in protein in a concentration-dependent manner (Fig. 3C, lanes 3–12). The apparent $K_d$ of SOD for the Fis dimer was estimated to be about 114 nM (Figs. 4 and 5A). At 240 nM Fis, almost all the input probe was shifted to the complexes (Fig. 3B, lane 7). No retarded band was detected when SOB lacking the Fis site was used as probe (Fig. 3C, lanes 1 and 2).

IHF showed dose-dependent formation of a single complex with SOE containing a single IHF-binding sequence (Fig. 3D, lanes 3–11), giving the complete retardation at 80 nM (Fig. 3D, lane 6). The apparent $K_d$ of IHF for SOE was 37 nM (Figs. 4 and 5A). At high IHF concentrations, the IHF-SOE complex formed further aggregates, which migrated to discrete bands with slower migration rates (Fig. 3D, lanes 7–11), through protein-protein contacts between IHF molecules. IHF did not form complexes with SOB without the IHF site at least within the protein concentration range examined (Fig. 3D, lanes 1 and 2).

Lrp also formed three discrete bands with SOF containing three repeats of the Lrp-binding site in a concentration-dependent manner (Fig. 3E). At 140 nM Lrp, virtually all the input DNA was converted into Lrp complexes (Fig. 3E, lane 7). The apparent $K_d$ of SOF for Lrp was 60 nM (Figs. 4 and 5A). In contrast, the probe SOB without the Lrp-binding sequence did not form any complexes with Lrp (Fig. 3E, lanes 1 and 2).

Curved DNA-binding Proteins—The DNA-binding properties of the sequence nonspecific DNA-binding proteins were examined using a pair of curved SOA and non-curved SOB probes. Four proteins, CbpA, H-NS, Hfq, and IciA, were identified as the curved DNA-binding proteins. CbpA and H-NS have been recognized as the curved DNA-binding proteins, but the specificity of DNA binding has not yet been analyzed for Hfq and IciA.

CbpA showed a strong cooperativity in binding to the curved DNA. By adding small amounts of CbpA, certain fractions of
SOA were shifted to big complexes which were trapped on the gel top even though some DNA probes remained unbound. All the input SOA probe was shifted to complexes by adding 280–350 nM CbpA (Fig. 3F, lanes 5 and 6). The apparent Kd for CbpA was estimated to be 122 nM (Figs. 4 and 5B). When a non-curved probe (SOB) was used as a control, however, the affinity was found to be weaker than the curved DNA (Fig. 3F, lanes 7–14), giving the apparent Kd value of about 175 nM CbpA (Figs. 4 and 5B). The affinity was 1.43-fold higher for the curved DNA than that for the non-curved DNA. Under the same conditions, we examined the DNA-binding properties for H-NS using the same set of DNA probes, curved SOA (Fig. 3G, lanes 1–12) and non-curved SOB (Fig. 3G, lanes 13–24). The apparent Kd of SOA and SOB for H-NS was calculated to be 115 and 165 nM, respectively (Figs. 4 and 5B). The affinity for the curved SOA was 1.43-fold higher than that for the non-curved SOB. Since no intermediate bands were observed, the binding of H-NS to the probe SOA was again highly cooperative.

In addition to these two known curved DNA-binding proteins, we newly identified two DNA-binding proteins with the preference for curved DNA. Hfq showed dose-dependent complex formation with both SOA (Fig. 3H, lanes 1–7) and SOB (Fig. 3H, lanes 8–14) probes. The apparent Kd of SOA and SOB for Hfq was 128 and 250 nM, respectively (Figs. 4 and 5B), indicating that the Hfq protein has a binding preference (1.95-fold) for curved DNA. The DNA binding activity of Hfq to the non-curved DNA is the weakest (Kd 250 nM) among the 12 proteins examined. Complete band shift of the input DNA fragment was observed at the Hfq concentration as high as 350 nM (Fig. 3H, lane 6). Likewise, IciA preferred the AT-rich curved SOA probe for binding (Fig. 3I). The apparent Kd was 126 and 180 nM for SOA and SOB, respectively (Figs. 4 and 5B). Thus, the affinity to curved DNA was more than 1.43-fold higher than that to non-curved DNA. At high protein concentrations, complexes migrating slower than the fully saturated IciA-SOIA complex were identified, which may represent aggregates of DNA-protein complexes.

Sequence-nonspecific DNA-binding Proteins—HU has long been recognized as a curved DNA-binding protein with no sequence preference. HU formed a ladder of complexes with both SOA and SOB probes concomitantly with the increase in HU concentration (Fig. 3K), each corresponding to the binding of increasing amounts of the HU αβ heterodimer. Since HU forms tetramers or even oligomers as analyzed by protein-electrophoresis by StpA essentially to the same extent (Fig. 3L, lanes 1–7 for SOA, and lanes 8–14 for SOB). The apparent Kd of SOA for StpA was 118 nM (Figs. 4 and 5B) and approximately 280 nM StpA was required for the maximum mobility shift (Fig. 3L, lane 6), while the Kq value of SOB was 127 nM (Figs. 4 and 5B). The results suggest no binding preference of StpA for the curved DNA at least under the conditions employed.

Dps is one of the stationary-phase proteins in E. coli and its association with the genome DNA is considered to convert the genome DNA compact during the growth transition from the exponential to stationary phase, thereby leading to switch of the transcription pattern (4) and protecting the DNA from oxidative stress-induced damage (35). However, the DNA binding specificity has not been analyzed for the Dps protein. Here we carried out the gel shift assay using both the curved SOA and non-curved SOB probes (Fig. 3J). Both showed essentially the same levels of DNA binding with the apparent Kd of 172–178 nM (Figs. 4 and 5B). Thus, we concluded that the Dps protein binds nonspecifically to both the curved and non-curved DNA probes with similar affinity.

DISCUSSION

Comparison of the Recognition Specificity of DNA Sequences among 12 DNA-binding Proteins—Here we examined the sequence recognition properties for 12 E. coli DNA-binding proteins, including four proteins, Dps, Hfq, IciA, and StpA, of which the specificity of DNA binding has not yet been analyzed. Among the eight DNA-binding proteins, of which the sequence specificity has hitherto been reported, five, i.e. CbpB (Rob), DnaA, Fis, IHF, and Lrp, are known to be sequence-specific DNA-binding proteins. These five proteins indeed showed in our assay system the specific recognition of the respective unique sequence of 26 (13), 9 (29), 15 (6), 15 (30), and 15 bp (31) in length (see Fig. 2 for the recognition sequence and Fig. 3 for the gel shift assay). Within the protein concentration ranges examined, these proteins did not form stable complexes with DNA devoid of the specific recognition sequences (see Fig. 3, A–E).

CbpB was isolated as a curved DNA-binding protein from E. coli (13), but after cloning and sequencing CbpB was found to be identical with the Rob (right origin-binding protein) (36). The CbpB-binding site is often located near promoters as well as at the right arm of the replication origin, oriC. CbpB and its homologues, SoxS and MarA, regulate transcription of the genes involved in oxidative stress response and multiple antibiotic resistance (37, 38). DnaA was originally identified as a regulatory factor of the replication initiation (reviewed in Refs. 39 and 40). The DnaA protein binds to the DnaA boxes at the oriC region (16) and regulates the synthesis of RNA encoded by the oriC region. As in the case of CbpB, DnaA is known to be involved in transcription regulation of a set of genes including its own gene (41, 42).

Fis is the most abundant nucleoid-associated protein in growing E. coli cells (75). Fis plays a global transcription regulator of the genes highly expressed in growing E. coli (6, 43). Upon Fis dimer binding, the DNA helix is bent to various extents depending on the sequence flanking the 15-bp core recognition sequence (20). On the other hand, the IHF level increases in the stationary phase (75). IHF was originally identified as a factor for recombination, but is known to constitute one of the major nucleoid-associated proteins with the regulatory activity of genome functions (44, 45). IHF has 30% amino acid sequence identity with HU (2). Upon DNA binding, IHF facilitates DNA bending (30). Like other sequence-specific DNA-binding proteins, IHF regulates transcription initiation when it binds near the target promoters (7, 46). Lrp also has dual functions, one as a global transcription factor which regulates, positively or negatively, as many as 75 genes, and the other as a determinant of the nucleoid structure (for reviews, see Refs. 9, 47, and 48). The consensus sequence of Lrp recognition is composed of 15 bp in length, but Lrp protects DNA sequence of 80 to 100 bp against DNAase I digestion (31). Lrp binding to the specific site induces DNA bending (49). All these five sequence-specific DNA-binding proteins influence DNA functions such as replication and transcription. For transcription...
tion regulation of specific genes or specific sets of genes, the direct interaction with the RNA polymerase has been suggested at least for four proteins, CbpB (Rob) (38), DnaA (50), Fis (51), and IHF (52). CbpB, Fis, and IHF require the C-terminal domain of α subunit (class I factors) while DnaA interacts with the β subunit (class III factor).2

The seven other proteins showed apparently sequence nonspecific DNA binding activities. Five proteins, CbpA, H-NS, HU, IciA, and StpA, have been proposed to be curved DNA-binding proteins. The binding preference for curved DNA has been demonstrated for CbpA, H-NS, and HU (11, 12, 23), but no experimental evidence has been published for the specificity of curved DNA binding for IciA and StpA. The present study indicates that CbpA, H-NS, and IciA are indeed the curved DNA-binding proteins. H-NS is one of the major nucleoid-associated proteins with a sequence nonspecific DNA binding activity (53) with the preference for curved or bent DNA (11, 28). As in the case of HU and IHF, mutations in hns result in pleiotropic effects on E. coli cell growth and stress response (54, 55, 56), suggesting that H-NS is a global regulator of gene functions as well as a structural protein for compacting the genome DNA. CbpA has been identified as such a DNA-binding protein as H-NS with the recognition activity of a synthetic AT-rich curved DNA (11). This was confirmed in our assay. On the contrary, the same assay did not support the curved DNA binding nature for HU, which was considered to be a sequence-independent DNA-binding protein with a preference for binding bent or kinked DNA (57), curved DNA (58), or DNA containing single-strand breaks or gaps (59). Another unexpected result is no significant binding preference of StpA for the curved DNA. StpA was identified as a homologue of H-NS in amino acid sequence, and has been believed to bind to intrinsically curved DNA (34). The results herein described suggest that StpA has an as yet unidentified unique role different from that of H-NS.

The IciA protein was discovered as a protein that binds to three repeats of a 13-bp long AT-rich sequence within the ori region and inhibits the replication initiation by interfering with the action of DnaA protein (17). In agreement with the predicted function, the intracellular level of DnaA increases in exponentially growing E. coli cells but the IciA level increases in the stationary phase (60). The present study provided experimental evidence supporting the previous prediction of the curved DNA binding nature for IciA. Likewise, Hfq was identified for the first time as a curved DNA-binding protein. Hfq was originally identified as a host factor, designated as HF-I, required for replication of phage Qβ RNA (21, 61). Isolated Hfq has a weak binding activity of both RNA and DNA (18, 32). Mutant studies indicate that Hfq plays an important physiological role in uninfected E. coli (62). Among the 12 DNA-binding proteins examined, two proteins, Hfq and StpA, were originally identified as the proteins which interact with RNA. Hfq is associated with both the nucleoid and ribosomes in uninfected E. coli (18). In good agreement with its association with RNA and ribosomes, Hfq regulates translation of certain mRNA, including mRNA for RNA polymerase σ^38 subunit (63, 64). Here we found that Hfq is a curved DNA-binding protein with the highest preference for curved DNA, the K_d value for curved DNA being 50% the value for non-curved DNA. One attractive hypothesis for the physiological functions for Hfq and StpA is that these proteins with binding activities to both DNA and RNA play a role(s) in functional coordination between DNA and RNA such as transcription-translation coupling.

Dps is the only DNA-binding protein that is produced only in

---

2 F. Hansen and A. Ishihama, unpublished observation.
unidentified specific recognition sequences; (ii) the DNA binding affinity changes depending on the nature of DNA probes such as the target sequence, the sequence of flanking regions, the probe length, and the extent of DNA curvature; (iii) the DNA binding activities of proteins also vary depending on the reaction conditions such as the overall ionic strength, pH, and the species and concentrations of salts and metal ions.

In contrast to these sequence-specific DNA-binding proteins, the influence of primary DNA sequence on the DNA binding affinity is less for the nonspecific DNA-binding proteins. However, the overall conformation of DNA affects the binding affinity of this group of proteins. HU dimers bind to linear DNA fragments about every 9 bp regardless of their sequence or length with the apparent 

K_d value of 500 nM and the cooperativity constant (\( \omega \)) of 30 corresponding to a weak cooperativity (73). Later Bonnefoy et al. (74) reported that the HU dimers bind specifically to the cruciform DNA with an apparent dissociation constant (K_d) of 5 nM and the value of 1 corresponding to a non-cooperative nature. The K_d value of HU we estimated using the linear probe with or without DNA curvature was 51 and 25 nM, respectively (Figs. 4 and 5B). In this study, we used the HU purified from E. coli co-expressing both \( \alpha \) and \( \beta \) subunits. Tanaka et al. (23) found that the equilibrium dissociation constant (K_d) for various DNA fragments of the \( \alpha_2 \) or \( \beta_2 \) homodimers (1–2 \( \mu \)M) were not so different from that of the \( \alpha \beta \) heterodimer.

The apparent dissociation constants for DNA binding by the nucleoid-associated proteins examined in this study ranged from 25 to 250 nM, which are 10^{-2}–10^{-10} -fold higher than those of gene-specific transcription factors with the binding activities to specific regulatory sites generally arranged near promoters (5). The molecular basis of the difference in DNA binding affinity awaits results of the structural analysis of DNA-protein complexes for both group proteins. Based on the DNA binding affinities determined in this study, the protein composition of E. coli nucleoid can be estimated once the intracellular concentrations of these proteins be measured.

Acknowledgments—We thank Takeshi Mizuno (ChpA, ChpB, H-NS), Seiichi Yasuda (DnaA), Reid C. Johnson (Fis), Naoki Goshima (HU), Deog Sun Hwang (IclA), Amos B. Oppenheim (IHP), David Low (Lrp), and Marlene Belfort (StPa) for gifts of the expression plasmids for DNA-binding proteins and suggestions for the protein purification.

REFERENCES

Twelve Species of the Nucleoid-associated Protein from \textit{Escherichia coli}: SEQUENCE RECOGNITION SPECIFICITY AND DNA BINDING AFFINITY

Talukder Ali Azam and Akira Ishihama

doi: 10.1074/jbc.274.46.33105

Access the most updated version of this article at \url{http://www.jbc.org/content/274/46/33105}

Alerts:
\begin{itemize}
  \item When this article is cited
  \item When a correction for this article is posted
\end{itemize}

Click here to choose from all of JBC's e-mail alerts

This article cites 72 references, 25 of which can be accessed free at \url{http://www.jbc.org/content/274/46/33105.full.html#ref-list-1}