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, H-NS, 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 > IHF > 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 with 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 are 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: aishiham@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, 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 CM-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 *E. coli*. 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 at the final step was more than 90% for all these proteins except for DnaA. 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.

**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</th>
<th>Inducer</th>
<th>Time</th>
<th>Plasmid source</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbpA</td>
<td>pCU80</td>
<td>CU211</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>pRL1077</td>
<td>BL21(DE3)</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>pHOP11</td>
<td>CU252</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>λpL</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>IHF</td>
<td>pSAshiphA</td>
<td>A6740</td>
<td>λpL</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>DL1592 (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>p177tdA</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>

a Turbidity measured with a Klett-Summerson photometer.
b Culture time after the induction.
c IPTG, isopropyl-1-thio-β-D-galactosidase.

**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, HU, IciA, IHF, 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 IHF, 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 IHF 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 conditions, 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 non-specific 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 II. The purity of proteins was analyzed by SDS-polyacrylamide**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Plasmid</th>
<th>Host bacteria</th>
<th>Promoter</th>
<th>Induction</th>
<th>Inducer</th>
<th>Time</th>
<th>Plasmid source</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbpA</td>
<td>pCU80</td>
<td>CU211</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>pRL1077</td>
<td>BL21(DE3)</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>pHOP11</td>
<td>CU252</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>λpL</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>IHF</td>
<td>pSAshiphA</td>
<td>A6740</td>
<td>λpL</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>DL1592 (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>p177tdA</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>

a Turbidity measured with a Klett-Summerson photometer.
b Culture time after the induction.
c IPTG, isopropyl-1-thio-β-D-galactosidase.
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 15-bp long Lrp-binding site (31), respectively. Three repeats of the 15-bp long IHF-binding site (30), and three repeats of the DnaA box on the same DNA surface (29), three repeats of the 15-bp long Fis-binding site (6), a 15-bp long Lrp-binding site (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. The combinations of bound and unbound probes stayed constant, but in some cases, the protein-bound probes formed diffuse bands. In the latter cases, the measurement of protein-bound probes was relied only on the amount of unbound free DNA.

The results of gel shift assays using 2 nM radiolabeled probes are shown in Fig. 4. The measurements were repeated at least three times for each combination, and the values shown in Fig. 4 represent the averages, together with the fluctuation levels of the measurements. The apparent dissociation constants thus obtained are summarized in Fig. 5. The measurements were repeated at least three times for each combination, and the values shown in Fig. 4 represent the averages, together with the fluctuation levels of the measurements. The apparent dissociation constants thus obtained are summarized in Fig. 5. The order of binding affinity among the sequence-specific DNA-binding proteins was: HU > Lrp > CbpA > Fis > DnaA, and the dissociation constant ranged from 37 nM for IHF to 213 nM for DnaA.

On the other hand, the affinity order among sequence-nonspecific DNA-binding proteins was: HU > H-NS > StpA > CbpA > IciA > Hfq > Dps using a curved DNA (SOA) probe, with the dissociation constant ranging from 51 nM for HU to 172 nM for Dps (Figs. 4 and 5B); and HU > StpA > H-NS > CbpA > Dps > IciA > Hfq using a non-curved DNA (SOB) probe, with the dissociation constant ranging from 25 (HU) to 250 (Hfq) nM (Figs. 4 and 5B). The affinity ratio between the curved and non-curved DNA ranged from 0.49 (HU), 1.03 (Dps), 1.08 (StpA), 1.43 (CbpA), 1.43 (H-NS), 1.43 (IciA) to 1.95 (Hfq). One unexpected finding was that HU had a higher affinity to the non-curved DNA while Hfq had a higher affinity to the curved DNA.

We carried out the same type of measurements using both decreasing decreasing concentrations of the labeled probes and increasing concentrations of unlabeled DNA probes. Essentially the same specificities of DNA binding were observed for all the proteins examined (data not shown). The order of DNA binding affinity was also essentially the same at least among the nonspecific DNA-binding proteins (data not shown).

In order to confirm the binding specificity of the 12 species of E. coli DNA-binding protein, we also investigated the complex formation in vitro using a mixture of 11 DNA fragments generated by HaeIII digestion of pUC19 DNA. When nonspecific DNA-binding proteins were analyzed using such DNA fragment mixtures, the longer DNA fragments tended to shift at lower protein concentrations than the shorter fragments (32). The order of DNA binding affinity measured in this system was essentially the same at least among the nonspecific DNA-binding proteins (data not shown).

Sequence-specific DNA-binding Proteins—Five proteins, CbpB (Rob), DnaA, Fis, IHF, and Lrp, were confirmed to be the
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 ($K_d$) 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 $K_d$ 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 concentrations.

**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.

**Table II.** Purification of DNA-binding Proteins from *E. coli*.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Purification Schedule</th>
<th>Mass (kDa)</th>
<th>Source</th>
<th>Additional Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbpA</td>
<td>Step 1</td>
<td>50</td>
<td>Purified</td>
<td>Step 2</td>
</tr>
<tr>
<td>CbpB</td>
<td>Step 2</td>
<td>60</td>
<td>Purified</td>
<td>Step 3</td>
</tr>
<tr>
<td>DnaA</td>
<td>Step 3</td>
<td>70</td>
<td>Purified</td>
<td>Step 4</td>
</tr>
<tr>
<td>Fis</td>
<td>Step 4</td>
<td>80</td>
<td>Purified</td>
<td>Step 5</td>
</tr>
<tr>
<td>Hfq</td>
<td>Step 5</td>
<td>90</td>
<td>Purified</td>
<td>Step 6</td>
</tr>
<tr>
<td>H-NS</td>
<td>Step 6</td>
<td>100</td>
<td>Purified</td>
<td>Step 7</td>
</tr>
<tr>
<td>HU</td>
<td>Step 7</td>
<td>110</td>
<td>Purified</td>
<td>Step 8</td>
</tr>
<tr>
<td>IHF</td>
<td>Step 8</td>
<td>120</td>
<td>Purified</td>
<td>Step 9</td>
</tr>
<tr>
<td>Lrp</td>
<td>Step 9</td>
<td>130</td>
<td>Purified</td>
<td>Step 10</td>
</tr>
<tr>
<td>StpA</td>
<td>Step 10</td>
<td>140</td>
<td>Purified</td>
<td>Step 11</td>
</tr>
</tbody>
</table>
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; 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). 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 $K_d$ 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 $K_d$ 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 $K_d$ 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 $K_d$ 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 ($K_d$, 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 $K_d$ 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-SOA 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 $\alpha\beta$ heterodimer. Since HU forms tetramers or even oligomers as analyzed by protein-protein cross-linking (33), the multiple bands observed in the gel shift assay might also be due to the formation of probe-tetramer or probe-oligomer complexes. One unexpected result was that the apparent $K_d$ for HU was rather higher (25 nM) for the non-curved SOB than that (51 nM) for the curved SOA (Figs. 4 and 5B). We repeated the gel shift assay using 0.5–5.0 nM concentrations of probes, but always observed the higher affinity to non-curved DNA than curved DNA (data not shown). Thus, we concluded that HU was not a curved DNA-binding protein at least under the assay conditions employed. As in the case of HU, StpA was also found not to be a curved DNA-binding protein even though it was identified as a homologue of H-NS because of the sequence similarity (34). Both the curved SOA and non-curved SOB were retarded on polycrylamide gel electrophoresis by StpA essentially to the same extent (Fig. 3L, lanes 1–7 for SOA, and lanes 8–14 for SOB). The apparent $K_d$ 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 $K_d$ 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 switching 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 $K_d$ 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, 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 transcrip-
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).  

The other seven 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 σ38 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 the stationary phase E. coli (4) and plays a role in the structural conversion of nucleoid DNA into more compact and stress-resistant states (35). Dps is also considered to play a role in switching of the global pattern of gene expression during the growth transition from growing to stationary phase (65, 66). Overall repression of the template activity for transcription of the stationary phase nucleoid may be correlated with the binding of Dps protein. We demonstrated for the first time that the purified Dps indeed shows the sequence nonspecific DNA binding activity. No unique sequence is required for Dps binding to the genome DNA (see Fig. 3J). The results are in good agreement with the finding that Dps becomes the most abundant structural component of stationary phase nucleoid (75).

Comparison of the DNA Binding Affinity among 12 DNA-Binding Proteins—In this study, we also compared for the first time the binding affinity among 12 E. coli DNA-binding proteins under the same reaction conditions. Among these 12 DNA-binding proteins, the published values of DNA binding affinity are available for six proteins, i.e. CbpB (Rob), DnaA, Fis, HU, IHF, and Lrp, but no quantitative data have been published for the other six proteins, CbpA, Dps, Hfq, H-NS, IciA, and StpA.

Ariza et al. (37) reported that CbpB binds a DNA fragment containing the micF promoter with the dissociation constant of about 1 nM and more weakly to DNA containing the sodA, nfo, or zuf promoters with the dissociation constant of 10–100 nM. In our in vitro binding assay, it was rather difficult to determine the K_{d} value for DnaA, because the purified DnaA tended to form aggregates. However, DnaA oligomerization seems to be required for its regulatory function of transcription (67). Schaper and Messer (29) reported that the dissociation constant of DnaA for specific binding to the probe DNA with three repeats of DnaA box was between 1 and 50 nM. Pan et al. (20) measured the equilibrium dissociation constants of wild-type and mutant Fis proteins for the hin-D site, located in the distal domain of the hin recombinase enhancer, with the apparent dissociation constants ranging from about 1 nM for specific sites to as high as 400 nM for nonspecific sites. Martin and Rosner (68) reported that Fis binds the upstream regulatory region of the multiple antibiotic resistance marRAB operon, with an apparent dissociation constant of about 5 nM. Blomfield et al. (69) measured the affinity of IHF to two IHF sites upstream of the fimE-fim switch-fimA region for type 1 fimbriae expression, giving the K_{d} value for the probe DNA of 10 nM. The K_{d} value (1.5 nM) for the specific and cooperative binding of IHF to the attP site of phage λ DNA was observed to be approximately 1000-fold higher than the value for nonspecific binding (70). Wiese et al. (71) measured the binding in vitro of Lrp to DNA upstream of the gltBDF operon, which includes the genes specifying the large (GltB) and small (GltD) subunits of glutamate synthase, with the apparent K_{d} of higher than 100 nM, which increases 5–10-fold in the presence of leucine. On the other hand, Wang and Calvo (72) determined the K_{d} value of 0.5–2.0 μM for Lrp binding to the 200-bp region upstream of the ifeIH operon containing six repeats of the Lrp site.

The values we determined in this study are generally within the range of published determinations. Among the five sequence-specific DNA-binding proteins (CbpB, DnaA, Fis, IHF, and Lrp), IHF showed the highest affinity (or the lowest K_{d} value of 37 nM) to the IHF box and DnaA showed the lowest affinity (or the highest K_{d} value of 213 nM) to the DnaA box. The difference between our estimations and the published values may be due to the difference in either the DNA probes used or the reaction conditions employed. Along this line, it should be noted that: (i) some of the sequence-nonspecific DNA-binding proteins analyzed may bind with higher affinities to as yet

---

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 cooperative 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 nM, which are 102–103-fold higher than those of nucleoid-associated proteins examined in this study. The apparent dissociation constant ($K_d$) of 5 nM and the cooperative constant ($\omega$) of 5 nM and the cooperativity changes depending on the nature of DNA probes and suggests for the protein purification.

Acknowledgments—We thank Takeshi Mizuno (ChpA, ChpB, H-NS, Seichi Yasuda (DnA), Reid C. Johnson (Fis), Naoki Goshima (HU), Deog Sun Hwang (IciA), Amos B. Oppenheim (IHF), 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 *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 [http://www.jbc.org/content/274/46/33105](http://www.jbc.org/content/274/46/33105)

Alerts:
- When this article is cited
- When a correction for this article is posted

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 [http://www.jbc.org/content/274/46/33105.full.html#ref-list-1](http://www.jbc.org/content/274/46/33105.full.html#ref-list-1)