Identification of New Fis Binding Sites by DNA Scission with Fis-1,10-phenanthroline−Copper(I) Chimeras†

Clark Q. Pan,‡ Reid C. Johnson,‡,§ and David S. Sigman*,‡,§,||

Molecular Biology Institute, Department of Biological Chemistry, School of Medicine, and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1570

Received August 28, 1995; Revised Manuscript Received February 8, 1996⊗

ABSTRACT: The chimeric nuclease Fis-OP has been used to identify novel Fis binding sites. Tethering the chemical nuclease OP-Cu^+ to position 73 of the protein with a newly developed longer acetyl-β-alanylamino spacer has facilitated the localization of two high-affinity Fis binding sequences in a 3 kb pUC19 plasmid. The shorter acetamido linker has allowed the chimeric nuclease to locate two strong Fis binding sites in the 50 kb phage λ genome. All four sites reside in biologically interesting loci and have been confirmed by gel-retardation and DNase I footprint analyses. A newly discovered site resides in the lac operon of *Escherichia coli*. The binding of Fis to this site may antagonize repression by the LacI repressor. These studies demonstrate the feasibility of applying chimeric chemical nucleases to the task of identifying functional protein binding sites of biological interest within genomes without any assumption about their sequence preference.

The *Escherichia coli* Fis protein regulates a diverse set of reactions including recombination, transcription, and replication [for review see Finkel and Johnson 1992)]. Studies on these systems have revealed a large number of specific Fis binding sites, though these are only poorly related at the primary sequence level. Approximately half of the known sites conform to a highly degenerate 15 bp sequence [5′ G N N C/T A/T N N A/T N N T/C G/A N N C 3′ where N can be any base (Hubner & Arber, 1989; Finkel & Johnson, 1992)] although only a fraction of sequences with this motif bind Fis. This lack of stringency precludes identification of specific Fis binding sites based on sequence information alone. In addition to binding specifically at a limited number of sites, Fis binds DNA at lower affinity sites, forming complexes which can be readily competed by excess nonspecific DNA.

We have previously shown that Fis linked to the chemical nuclease 1,10-phenanthroline−Cu attached at either of two different positions within the DNA binding region of Fis cleaves DNA at the edge of specific binding sites (Pan et al., 1994a). N73C-OP,[1] containing 1,10-phenanthroline linked via an acetamido spacer to amino acid 73, has been shown to be the most efficient DNA cleavage reagent of known sites cutting at all 26 of the Fis binding sites tested (Pan et al., 1994a). Since residue 73 is located at the periphery of the Fis−DNA interface, the DNA target is adjacent to the chemical nuclease OP−Cu^+ only in specific complexes when the DNA is appropriately wrapped around the protein. As a result, cleavage of a DNA sequence by the Fis-OP chimera provides a stringent criteria for distinguishing specific from nonspecific complexes.

In the present work, we have linked OP to the Fis protein at position 73 through a three-atom-long acetamido spacer (N73C-OP) and a seven-atom long acetyl-β-alanylamino spacer (N73C-AOP) whose longer tether permits the more efficient scission of known Fis sites in 3−6 kb plasmids. The N73C-AOP derivative facilitated the localization of two previously unknown Fis binding sites in a 3 kb pUC19 plasmid. We denote the first site as the "*lacP* site" because it resides within the *lac* promoter region of this plasmid and the second site as the "*oriE* site" because it is near the origin of replication of this colE1-derived plasmid. The derivative with the longer spacer was less useful when applied to larger pieces of DNA such as the 50 kb phage λ genome because of the high background cleavage resulting from less specific binding by Fis. The use of N73C-OP, on the other hand, enabled identification of a novel high-affinity Fis binding site in the left operator of the λ regulatory region. We name this site the "λ O_L Fis site I". We have performed gel-retardation assays and DNase I footprint studies to confirm that these three new Fis sites in fact bind tightly to Fis *in vitro*. The present report describes the first search for binding sites of a regulatory protein using OP conjugates without any previous information about the existence or sequence of such sites.

MATERIALS AND METHODS

Modeling of the Fis N73C-AOP−DNA Complex. Acetyl-β-alanylamino-OP was modeled onto the Fis N73C-OP−DNA complex (Pan et al., 1994a) by using the Insight II program (Biosym Technologies, San Diego).

Plasmids and Strains. The *Hin* distal and proximal enhancer sites are contained in the pUC9-derived plasmid pMS577 (Johnson & Simon, 1985). The pUCenB plasmid

† This work was supported by USPHS Grant GM38509 (R.C.J.) and USPHS Grant GM21199 (D.S.S.). C.Q.P. was a trainee of USPHS National Research Service Award GM07185.

* To whom correspondence should be addressed: Molecular Biology Institute, University of California, Los Angeles, CA 90095-1570. Phone: 310-825-8903. Fax: 310-206-7286. E-mail: sigman@ewald.mbi.ucla.edu.

‡ Molecular Biology Institute.

§ Department of Biological Chemistry, School of Medicine.

|| Department of Chemistry and Biochemistry.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1996.

[1] Abbreviations: OP-Cu^+, 1,10-phenanthroline−copper(I); Fis-OP, Fis-1,10-phenanthroline; N73C-OP or N73C-AOP, OP linked to position 73 of the Fis protein through an acetamido or an acetyl-β-alanylamino spacer; IPTG, isopropyl thio-β-galactoside.

0006-2960/96/0435-4326$12.00/0

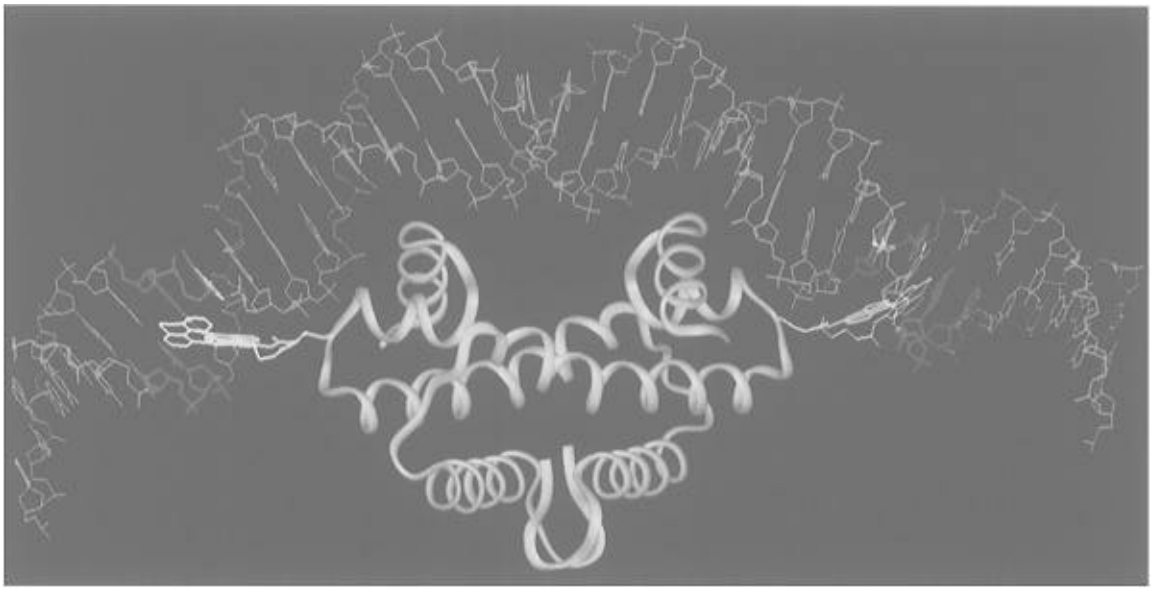

FIGURE 1: Fis N73C-OP and N73C-AOP interacting with a typical Fis binding site based on our working model of a Fis−DNA complex (Yuan et al., 1991; Feng et al., 1992). The DNA is colored blue, with the nucleotides cleaved by the Fis-OP derivatives in red. The longer acetyl-β-alanylamino linker (magenta) of N73C-AOP brings the copper atom up to 5 Å closer to the DNA than the shorter acetamido linker (yellow) of N73C-OP.

was constructed by cloning a 50 bp DNA that contains an engrailed homeodomain binding site into the *Sma*I site of pUC19 (Pan et al., 1995). The two pUC plasmids are identical except within the polycloning region from *Eco*RI to *Hin*dIII. The *Hin* enhancer distal and proximal sites are present individually in pMS609 and pMS529, respectively (Johnson & Simon, 1985). The λ *attR* site is located in pJT175 obtained from A. Landy (Brown University). The "dist21" and "dist27" variants of the distal site are contained in two derivatives of pCY4 (Prentki et al., 1987), named pRJ1129 and pRJ1128, respectively (R. Johnson, unpublished).

Derivatization of Fis N73C with Iodoacetamido-OP (IOP) and Iodoacetyl-β-alanylamino-OP (IAOP). The derivatization procedure was essentially the same as described previously (Pan et al., 1994a). The N73C mutant (typically 50−100 μM) was reduced for at least 30 min at 4 °C with a 5−10 times molar excess of dithiothreitol or β-mercaptoethanol, with the excess reducing agent removed by a G-25 spin column or dialysis. Within a few hours after the depletion of the reducing agent, a 5−10 times molar excess of IOP or IAOP was added. The derivatization reaction was allowed to proceed for at least 6 h at 4 °C, followed by the removal of the free OP derivatives by dialysis at 4 °C. A spin column can also be used, though it often results in greater loss of protein. The derivatization efficiency and the amount of the reducing agent were monitored by the DTNB assay with Ellman's reagent, 5,5′-dithiobis(2-nitrobenzoic acid) (Riddles et al., 1983). The concentration of the OP derivatives was followed by measuring the absorbance at 270 nm (ϵ = 30 000 $M^{-1}cm^{-1}$). The reducing and derivatization buffers were identical, typically at 20 mM Tris-HCl (pH 7.5) and 200 mM NaCl. The Fis-OP derivatives were stored in the above buffer plus 50% glycerol at −20 °C. Upon storage for greater than a month, the efficiency of scission was markedly reduced due to self-oxidation.

DNA Cleavage Reaction. DNA scission in the gel matrix was carried out essentially as described previously (Pan et al., 1994a) except that the gel-retardation assay was performed in 1 × TB (Tris−borate) buffer without EDTA. The bands in the gel-retardation assay were excised out of the gel and treated with an equal volume of 40 mM Tris-HCl (pH 7.5), 160 mM NaCl, and 100 μM $CuSO_4$ for 10 min at room temperature. The cleavage reaction was induced by the addition of 6 mM mercaptopropionic acid (MPA) and 6 mM H_2O_2, allowed to proceed for 2 h at room temperature, and quenched by the addition of 1 mM neocuproine. For cleavage in solution, the Fis-OP derivatives were mixed with ∼100 pM radiolabeled or ∼10 nM nonlabeled DNA in a buffer containing 20 mM Tris-HCl (pH 7.5), 80 mM NaCl, and 50 μM $CuSO_4$ for 10 min in a final volume of 5−20 μL. DNA cutting was initiated by the addition of 3 mM MPA, proceeded at room temperature for at least 1 h, and was stopped by the addition of 1 mM neocuproine. The products were analyzed on a 6% denaturing polyacrylamide gel except for Figure 6B, which was electrophoresed on a 0.6% agarose gel.

DNase I Footprint. The wild-type Fis protein was mixed with the radiolabeled DNA fragment (typically ∼100 pM) for 10 min in a buffer containing 20 mM Tris-HCl (pH 7.5), 80 mM NaCl, and 5 mM $MgCl_2$ in a total volume of 5−10 μL. Approximately 0.2 unit of DNase I was added to the mixture for 2 min, and the reaction was quenched by the addition of 50 mM EDTA in formamide. The DNase I footprint pattern was analyzed on a 6% denaturing polyacrylamide gel.

RESULTS

Acetyl-β-alanylamino Spacer. The bending angles induced in DNA by Fis range from 40° to 90° with most of the variation due to bending in the sequences flanking the 15 bp core (Thompson & Landy, 1988; Finkel & Johnson, 1992). The current model of a typical Fis−DNA complex proposed by Feng et al. (1992) includes an overall 60° bend in the DNA toward the protein (Figure 1). Without additional bending in the DNA flanking the core binding site, the model suggests that OP-Cu^+ linked to residue 73 of the protein is not close enough to the cleaved nucleotides when

the spacer is the acetamido group (Pan et al., 1994a). Therefore, for all those Fis binding sites with a 60° DNA bending angle or less, DNA cleavage by N73C-OP may be undetectable because the DNA is too far away from the OP-Cu^+. To increase the chance of identifying all possible specific Fis binding sites, we have chemically synthesized iodoacetyl-β-alanylamino-OP (IAOP). Molecular modeling indicates that derivatization of Fis N73C with IAOP (N73C-AOP) would bring the copper atom up to 5 Å closer to the DNA than the shorter acetamido linker (N73C-OP) (Figure 1).

To compare the ability of N73C-OP and N73C-AOP to cleave the DNA, we have performed DNA scission on a 200 bp DNA substrate that contains two Fis binding sequences: the *Hin* enhancer distal site and the *Hin* enhancer proximal site. The two OP-Cu^+ derivatives of Fis bound to these sites with approximately equal affinity (data not shown). In a gel-retardation assay on the 200 bp substrate, only one Fis-shifted band was observed at low Fis concentrations but a second band with a higher relative mobility is seen when more Fis was added. Presumably, the first shift was due to Fis occupying one of the two Fis binding sites, and the second shift resulted from occupancy at both sites**.** We isolated the free DNA bands as controls and the second Fis-shifted bands by N73C-OP and N73-AOP to ensure that both sites were saturated with Fis. DNA scission was then activated in the gel matrix by the addition of Cu^{2+} and mercaptopropionic acid. The cleaved products were resolved on a 10% denaturing polyacrylamide gel (Figure 2). The left flank of the *Hin* distal site was cleaved with equal efficiency by both N73C-OP and N73C-AOP while cleavage of the right flank of the distal site and both flanks of the proximal site was detectable with Fis N73C-AOP but not Fis N73-OP. Gel-retardation assays suggest that the left flank of the distal site bends sharply toward the protein while the other three flanks exhibit less bending. Thus, N73C-AOP can increase the chance of discovering those specific and biologically functional Fis sites that have lower DNA bending angles in plasmids 3−6 kb long.

Identification of Two New Fis Binding Sites in pUC19. To determine the reliability of the chimeric nuclease technique to correctly identify binding sites within large segments of DNA, we first tested the ability of N73C-OP to locate five known sites (*Hin* enhancer distal, *Hin* enhancer proximal, λ *attR*, and two *Hin* enhancer distal variants, dist21 and dist27) cloned into four different plasmids (pUC19, pBR322, pBR327, and pCY4) ranging from 3 to 5 kb in size. The predominant sites of double strand scission within the plasmid in each case were at these known sites. However, two weaker additional sites of cleavage were seen in the pUC19-derived plasmid which could be more readily detected with N73C-AOP. The positions of these previously unknown Fis sites were localized by cleaving the *Afl*III−*Hin*dIII and *Afl*III−*Ava*II fragments of pUC19. The first was denoted the *lacP* site because it is located in the middle of the *lac* promoter; the second, the *oriE* site because it lies near the origin of replication of the colE1-derived plasmid (data not shown).

To identify the precise positions of the two Fis binding sites, we performed DNase I protection and Fis-OP cleavage on DNA fragments from pUC19 that contain the *lacP* and *oriE* sequences (Figure 3). The DNase I protected regions by Fis include the sites of cleavage by N73C-AOP. An area

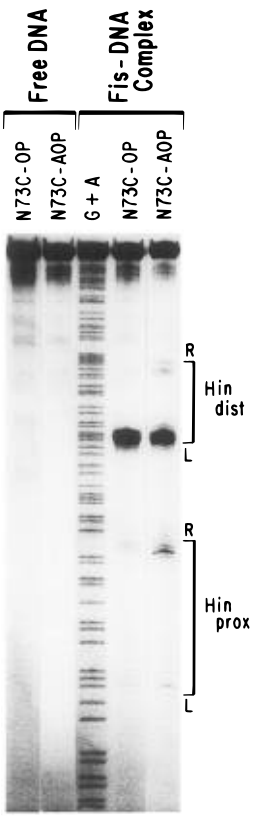

FIGURE 2: DNA cleavage pattern by Fis-OP derivatives in a polyacrylamide gel matrix. A 200 bp *Eco*RI−*Hin*dIII fragment of pMS577 (3′ radiolabeled at the *Eco*RI end) containing the *Hin* enhancer distal site (*Hin* dist) and the *Hin* enhancer proximal site (*Hin* prox) was incubated with either N73C-OP or N73C-AOP and electrophoresed in a polyacrylamide gel. The free DNA bands and the Fis-bound bands were excised out of the nondenaturing polyacrylamide gel and treated with copper and mercaptopropionic acid to induce DNA cleavage. L and R represent the left and right flanks of each Fis binding site (Pan et al., 1994a).

of protection by wild-type Fis from DNase I stretches from −62 to −39 on the upper strand and −66 to −44 on the lower strand with respect to the transcription initiation site (+1) for the *lacP* site and +38 to +62 of the top strand and +36 to + 57 of the bottom strand for the *oriE* site with respect to the RNA/DNA junction. With the 15 bp core consensus as a guide, the DNase I footprint and N73C-AOP cleavage data allowed us to deduce the exact location of the *lacP* Fis binding site as −60 to −46 for the core sequence and the *oriE* Fis binding site as +41 to +55 for the core (Figure 4). The large increase in cutting efficiency with the longer spacer suggests that both sites are weakly bent toward the protein. The asymmetrical pattern of cleavage further indicates that the flanking sequences are closer to Fis on the left side than the right for both sites.

Gel-retardation assays have also been used to characterize the relative affinity of the newly identified Fis binding sites. A 108 bp *Eco*RI−*Hin*dIII fragment derived from pUCenB containing a binding site for the *Drosophila* engrailed homeodomain protein (Pan et al., 1985) was used as a negative control while the *Hin* distal site residing on a 90 bp DNA generated from *Cla*I and *Hin*dIII digests of pMS577 was included as a positive control. The *lacP* and *oriE* sites were carried individually on a 95 and a 96 bp DNA fragment, respectively (see Figure 3 legend). The affinities of the *lacP* and *oriE* sites for Fis in the presence of 100 μg/mL nonspecific competitor DNA were within 2-fold of the *Hin*

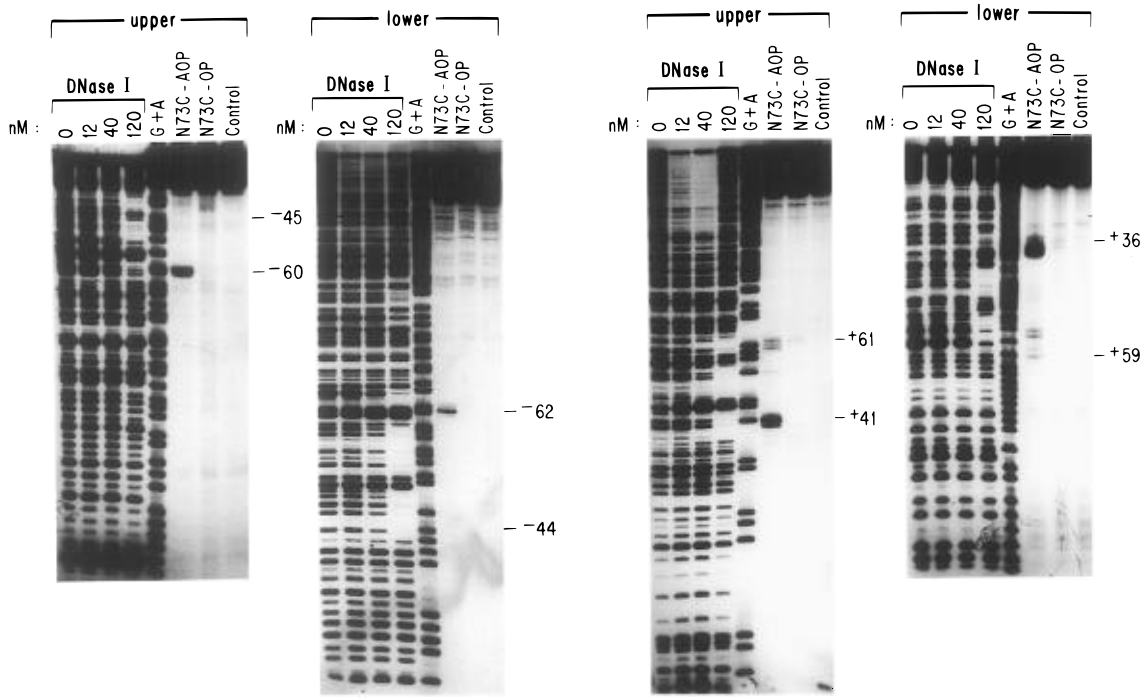

FIGURE 3: DNase I footprint and Fis-OP cleavage patterns of the *lacP* site (A, left) and the *oriE* site (B, right) generated in solution and electrophoresed on a 10% denaturing polyacrylamide gel. The polymerase chain reaction was used with primers pUCA (pUC coordinates 530−551) + pUCA′ (606−624) and pUCB′ (769−790) + pUCB (849−864) to generate the *lacP* and *oriE* containing fragments, respectively. The first four lanes of each panel depict the DNase I protection by wild-type Fis, with the amount of Fis added listed at the top of each lane. The DNA scission patterns by N73C-AOP and N73C-OP and the no protein control are illustrated in the last three lanes of each panel. All three of these lanes have also been treated with mercaptopropionic acid and copper. The *lacP* site (A) is numbered according to the coordinates of the *lac* operon with +1 as the start of transcription. The numbering of the *oriE* site (B) follows that of the origin of replication of the colE1 plasmid with +1 being the beginning of DNA synthesis. Both numbering systems and the designation of upper and lower strands correspond to those presented in Figure 4.

distal site, one of the strongest Fis binding sequences known. The negative control pUCenB fragment displayed no binding to Fis even at 10 times higher protein concentration than that required to generate complexes with the other three fragments (data not shown).

Other Fis Binding Sites in the pUC19 Plasmid. To date, over 35 Fis binding sites have been identified. About half of these sites perfectly match the 15 bp core consensus, with the other half containing a single bp mismatch. Double bp mismatches have also been found in a few cases. Computer-assisted examination of the 2.7 kb pUC19 vector sequence indicates that there are at least 177 potential Fis binding sequences (13 perfect matches and 164 single mismatches). However, the experiment with the Fis-OP derivatives revealed only two high-affinity Fis binding sites in the vector, each with a single bp mismatch relative to the consensus sequence. To test whether there may be other potential Fis binding sequences in the pUC19 plasmid not revealed by Fis-OP cleavage, we performed gel-retardation assays on the whole plasmid that has been digested by restriction enzymes which cut at multiple sites. The pUC18 plasmid (identical to pUC19 except for the orientation of the polycloning region) was digested with *Hpa*II to generate a series of defined DNA fragments that are small enough in size so that a change in electrophoretic mobility upon Fis binding could be detected. As shown in Figure 5, the mobility of only two DNA fragments released from the *Hpa*II digest were significantly altered at the lower concentrations of the wild-type Fis protein. The first one is 489 bp, corresponding to pUC19 coordinates 524−1013 and covers the two Fis binding sites identified by Fis-OP cleavage. The second is 242 bp, derived from pUC19 coordinates 1991−2233. As expected, at higher Fis concentrations, nearly all the fragments are bound by Fis. Gel-retardation assays on *Hae*III- and *Hha*I-digested pUC18 fragments were consistent with the presence of two high-affinity Fis binding sites at *lacP* and *oriE* and also point to the presence of an additional site between nucleotides 2088 and 2233, within the β-lactamase gene (data not shown). It was not detected initially with Fis-OP because *Sca*I, which cleaves within this region at nucleotide 2180, was used to linearize the plasmid prior to Fis-OP scission. N73C-OP digestion of pUC19 linearized with *Eco*RI revealed a scission site at approximately nucleotide 2200. *Therefore, there are only three strong Fis binding sites in pUC19 out of the 177 possible ones based on the consensus, and Fis-OP scission has identified each one of them.*

Fis Binding Sites in Phage λ Genome. Fis-OP cleavage was also used to probe the 48.5 kb bacteriophage λ genome for high-affinity Fis binding sites. In addition to a direct Fis role in stimulating λ excision by binding to a site in *attR* (Thompson et al., 1987; Ball & Johnson, 1991a), Fis also regulates lysogeny by a yet unknown mechanism (Ball & Johnson, 1991b). Thus, it is of interest to locate other Fis binding sites within the λ genome, particularly those within the regulatory regions.

Digestion of the whole λ genome with N73C-OP released an ~10 kb fragment that was barely visible above a high background of cleavage (data not shown), suggesting that there may be Fis site(s) near either λ coordinate 10 000 or 38 000 (Figure 6A). Since the latter is within the major regulatory region of the phage, DNA scission with N73C-

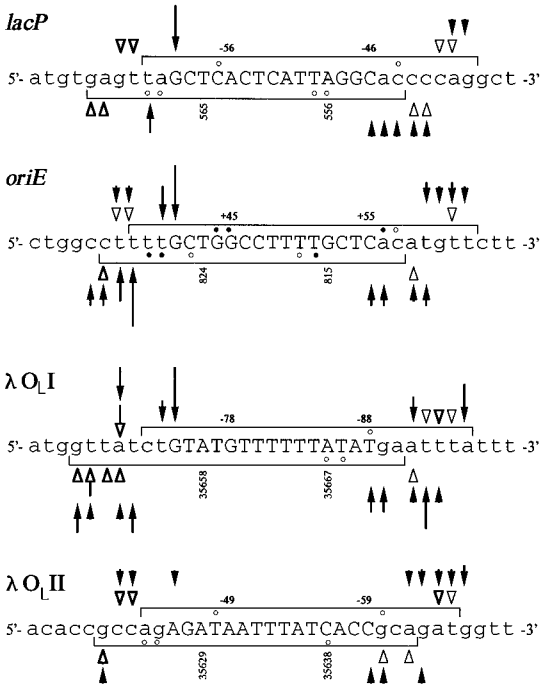

FIGURE 4: Summary of DNA scission patterns by Fis N73C-OP (opened arrows) and N73C-AOP (filled arrows) and the DNase I footprint of the wild-type Fis protein (brackets) on the four newly discovered Fis binding sites. The DNA sequence of the central 15 bp core consensus (uppercase letters) along with 10 bp of each flank (lowercase letters) is aligned according to the core. The bold numbering at the top of each sequence represents the coordinates of the promoters where the sites are derived for the *lacP* and the two λ O_L sites or the coordinates of the origin of replication for the *oriE* site. The numbers at the bottom correspond to that of the pUC19 vector for the first two sites or the phage λ genome for the last two. The major scission sites are depicted with the relative DNA cutting efficiencies approximately proportional to the length of each arrow and, in the case of N73C-OP, also to the thickness of the open arrowheads. Some of the scission sites shown for N73C-OP cleavage are too light to be seen in Figure 3 but have been observed on longer exposure of the autoradiograph or in other experiments. Filled dots represent those nucleotides that show an increase in DNase I sensitivity while open dots correspond to those with the same level of sensitivity upon the addition of the protein.

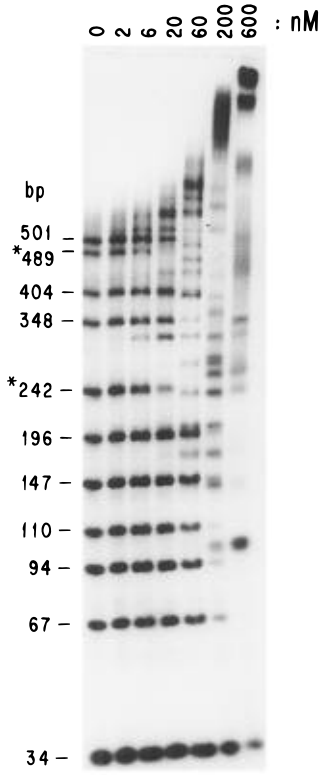

FIGURE 5: Gel retardation assay of 3′-radiolabeled *Hpa*II fragments of pUC18 by wild-type Fis. The length of each unbound fragment is shown on the left. Out of the 11 fragments shown, only two (242 and 489 bp in length, highlighted by asterisks) are significantly shifted at lower Fis concentrations. The amount of Fis added is denoted at the top of each lane.

OP on a 6.8 kb *Bgl*I−*Eco*RI fragment (nucleotides 32 323−39 168) was performed, yielding a fragment 3.5 kb from the *Eco*RI end, corresponding to λ coordinate ~35 700 (Figure 6B). Weaker scission has occasionally been observed at ~35 100 and ~36 000. Scission reactions of a PCR-generated 1.4 kb DNA (34705−36144) by N73C-OP analyzed on sequencing gels revealed cleavage sites around 35 040, 35 650, and 35 900 on the upper strand but only the 35 650 site on the lower strand, indicating that the central 35 650 site was the major site of scission.

The Fis-OP scission and DNase I footprint patterns of a 134 bp DNA (35 569−35 702) revealed two closely spaced Fis binding sites in the left promoter (P_L) of the phage. The data for the lower strand are shown in Figure 6C and summarized in Figure 4 along with that for the upper strand. The DNase I footprinted areas cover those nucleotides cleaved by Fis-OP. The wild-type Fis protein protects a region stretching from −72 to −95 and −43 to −65 for the top strand and −67 to −90 and −40 to −62 for the bottom strand relative to the P_L transcription start site, corresponding to a stronger Fis binding site, designated as λ O_L Fis site I, and a weaker site, designated as λ O_L Fis site II. With the 15 bp consensus Fis binding sequence as a guide, we were able to assign the core of site I with λ coordinates 35 656−35 670 (P_L −74 to −88) and that of site II with λ 35 627−35 641(P_L −45 to −59). The right flank of site I appears to be less bent than the left as demonstrated by the lower N73C-OP scission and the greater enhancement of cutting efficiency with the longer spacer (Figure 4). Site II, on the other hand, displays a much more symmetrical pattern of Fis-OP cleavage, suggesting that it is similarly bent on both sides. Site I exhibits approximately 5-fold higher affinity for Fis than site II on the basis of the footprint data. Gel-retardation analysis has confirmed that two binding sites are on the 134 bp fragment, with the stronger site having a similar affinity to Fis as the *Hin* distal site.

DISCUSSION

Spacer Length and DNA Scission Efficiency. The cleavage specificity of free DNA by the tetrahedral 2:1 1,10-phenanthroline−copper [$(OP)_2Cu^+$] depends on the binding affinity of the chelate to the minor groove (Sigman, 1990). In any DNA fragment, the cleavage pattern will vary according to sequence, but weak scission will be observed at all positions. To target the chemical nuclease, it has to be tethered to a ligand that interacts with high affinity to a specific site. DNA binding proteins can serve as effective carriers of the nuclease (Pan et al., 1994b). Protein-linked OP-Cu^+ yields highly localized DNA damage, suggesting that the oxidative species created during the cutting reaction must be liganded to the copper ion and not freely diffusible like a hydroxyl radical formed by ferrous EDTA. This localized cleavage would require an optimal placement of OP-Cu^+ relative to the DNA. Molecular modeling suggests

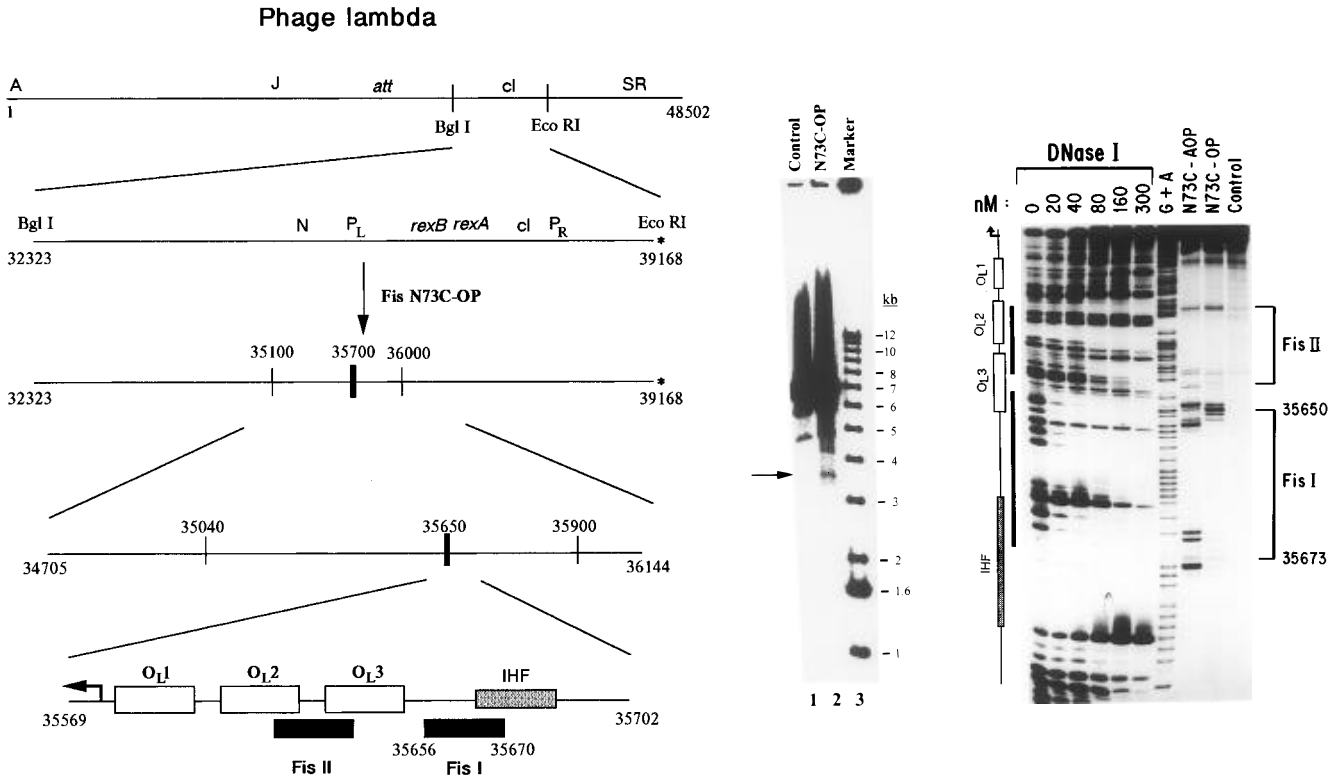

FIGURE 6: Identification of Fis binding sites in phage λ (l). (A, left) Map of the λ genome illustrating the four successive steps for localizing the two Fis binding sites in the left operator region of the phage. Line 1 gives the map of λ with some landmark genes and loci noted. Fis N73C-OP cleavage of the intact λ genome gave a band of approximately 10 kb over a high background. Cleavage of the major regulatory region (~10 kb from the right end) radiolabeled at the *Eco*RI end as denoted by the asterisk (line 2) by N73C-OP gave scission at approximately 35 700 and to lesser extent at 36 000 (line 3). More refined mapping of the scission sites within this region was obtained by cleavage of a PCR-generated fragment between 34 705 and 36 144 as depicted on line 4. Finally, the exact locations of the two Fis binding sites in the Pl region were determined by analysis of N73C-OP cleavage products and DNase I protection patterns on sequencing gels (line 5). Fis binds with high affinity to site I, which overlaps with an IHF binding site, and with lower affinity to site II, which spans the O_L2 and O_L3 operators. The start of transcription for Pl is designated by a leftward pointing arrow. (B, center) Cleavage by Fis N73C-OP of the 6.8 kb fragment containing the λ Pl promoter. The 6.8 kb *Bgl*I−*Eco*RI fragment (1,32 323−39 168 labeled at the *Eco*RI end) was treated with mercaptopropionic acid (MPA) and copper (lane 1) or MPA and copper plus Fis N73C-OP (lane 2). The scission products were then analyzed on a 0.6% agarose gel. Lane 3: 1 kb molecular weight ladder (Gibco BRL). The arrow highlights the 3.5 kb product released by N73C-OP cleavage. (C, right) DNase I footprint and Fis-OP scission patterns in solution of the lower strand of the 134 bp DNA (1, 35 569−35 702) after electrophoresis on a 6% denaturing polyacrylamide gel. The six left lanes are DNase I digestions of wild-type Fis binding with increasing amounts of Fis added to the fragment as listed at the top of each lane. The last three lanes show DNA cleavage by N73C-AOP, N73C-OP, and the no protein control.

that protein-targeted cleavage of DNA is most efficient when the copper ion is held within 4 Å of the C-1 H. This requirement of optimal geometry limits the number of amino acids available for OP attachment especially if one considers that the copper ion bound oxidant must be accessible to the C-1 H of both strands to accomplish double-stranded breaks. To broaden the number of residues that can be used, we have synthesized the acetyl-β-alanylamino linker to complement the previously developed acetamido spacer. Indeed, N73C-AOP has helped to locate the two strong Fis binding sites in the pUC19 vector, in particular the *lacP* and the *oriE* sites.

Although the longer spacer has proven to be useful for discovering Fis binding sequences in relatively short pieces of DNA (less than 5 kb), it was not as useful in larger DNAs (e.g., ~50 kb λ genomic DNA). One reason for its restricted utility is that Fis protein binds specific sequences only 2 orders of magnitude more tightly than nonspecific sequences. Since many potential binding sites will exist in a large DNA given the relaxed Fis consensus sequence, significant background cleavage may result from nonspecifically bound Fis molecules in which the longer tether would permit the close approach of the OP-Cu to the oxidatively sensitive minor groove. The random scission of DNA by nonspecifically bound protein has also been observed in the cleavage of plasmid by the Trp repressor−OP chimera in the absence of the corepressor L-tryptophan (data not shown). Pendergrast et al. (1994) have suggested that proteins which bend DNA, such as the catabolite activator protein (CAP), might be readily converted into efficient site-specific nuclease because longer tethers would only approach the DNA in a highly bent protein−DNA complex. Although applicable to CAP, this strategy is not useful with Fis since this protein not only has a relatively high nonspecific binding affinity for DNA but bends its recognition sites less than CAP at most binding sites. Faced with thousands of nonspecific sites in the λ genome, Fis N73C-AOP cleaves some of them, resulting in a smear of background trailing from the parent band. Fis N73C-OP cleavage, on the other hand, is more stringent probably because the short tether restricts the approach of the OP-Cu in nonspecific complexes and reduces background scission.

Mapping the Exact Location of Fis Binding Sites. Since the Fis protein lacks a clear consensus recognition sequence, the precise position of a Fis binding site is difficult to determine. DNase I footprint has traditionally been used to locate Fis binding sequences because they often display

hypersensitivity around positions 3 and 15 on each strand of the core (Finkel & Johnson, 1992). However, a number of them do not. The added advantage of applying the chimeric nuclease method is that the coordinates of the newly discovered site can be immediately obtained. DNA cleavage of 26 Fis binding sites by N73C-OP typically exhibits a pattern of scission at the third and fourth nucleotides 5′ to the 15 bp core and the fifth and sixth positions 3′ to the core. As illustrated in the N73C-OP scission pattern of the four new sites, this rule is obeyed in every case (Figure 4). Thus, Fis-OP cleavage provides another useful technique to define precisely a Fis binding site, and in those cases where no DNase I hypersensitivity can be detected, it may be superior to nuclease footprinting.

The consensus originally developed by Hubner and Arber (1989) is so degenerate that often two or more overlapping sequences that may satisfy the consensus exist in a footprinted region. For the *lacP* site, besides the one (*lac* promoter −60 to −46) determined on the basis of the rule established by the N73C-OP scission, two additional sequences that may be considered as potential Fis binding sequences are present within the DNase I footprinted area: one three nucleotides to the left of the site chosen by us and the other two nucleotides to the right (Figure 4). Each of these two alternative sequences possesses a double mismatch from the consensus whereas the one chosen has a single mismatch. Two bp mismatches have been found in a rare number of known Fis binding sites. For the origin of colE1 replication, in addition to the site defined by the N73C-OP cleavage pattern which has a single bp mismatch from the consensus, two sequences that have double mismatches are also present within the footprinted region. These two other sequences are two and three nucleotides to the right of the chosen *oriE* site. The λ O_L site I has only a single mismatch from the consensus, but a double mismatch sequence that is two nucleotides to the right shares the same region of protection from DNase I. Thus, we could not have been certain of the exact location of these three sites without the N73C-OP cleavage pattern. The placement of the λ O_L site II presents the greatest challenge. Four overlapping 15 bp sequences that have double mismatches from the consensus are bracketed by the same area of protection. Using the N73C-OP cleavage pattern as a guide, we believe that most likely this site occupies positions −46 to −59 of the left operator of the phage.

Potential Biological Roles of the Newly Discovered Fis Sites. We have identified four Fis binding sequences which all could be interesting biologically. The *oriE* site in pUC19 is located 48 bp downstream of the RNA/DNA junction in the colE1-derived plasmid origin of replication, but we have not noticed any instability of pUC19-derived plasmids in *fis* mutants. Hiasa and Marians (1994) have shown a modest inhibitory effect on colE1 replication *in vitro* by increasing concentrations of Fis, though other nucleoid-associated proteins like HU and IHF also inhibited at similar concentrations. Fis has been shown to bind to *oriC*, the origin of the *E. coli* chromosome, and enhance the stability of *oriC* minichromosomes *in vivo* (Gille et al., 1991; Filutowicz et al., 1992).

The λ O_L Fis sites may be related to the previous observation that λ lysogens encoding temperature-sensitive cI repressors are less stable in the absence of Fis (Ball & Johnson, 1991). Giladi et al. (1990) have demonstrated that integration host factor (IHF) stimulates the transcription of the phage P_L promoter by 3−4-fold by binding at two sites upstream of O_L. DNase I footprinting has located these sites to be between −67 to −105 and −154 to −203. The first region overlaps with the high-affinity Fis site I whose core binding region extends from −74 to −88. Therefore, it is possible that Fis may be interfering with IHF binding at the promoter proximal site and inhibiting activation of P_L by IHF. The lower affinity Fis site II overlaps both O_L2 and O_L3, the two cI repressor binding sites, and thus could play a role in controlling repression. Experiments are in progress to investigate the role of these Fis binding sites on the regulation of P_L activity and, in particular, the relationship between Fis, IHF, and cI in modulating its control.

Since the Fis binding site in the *lacP* region overlaps the CAP binding site, it seemed possible that Fis would negatively affect CAP activation. However, we did not observe a negative effect by Fis on *lacP* activity under inducing conditions but found instead that the absence of Fis produced a more completely repressed state (unpublished data). Possibly Fis-induced bending inhibits the tetrameric Lac repressor from forming a stable DNA loop between O1 and O3, a secondary operator centered at −83 of the lac upstream regulatory regions (Mossing & Record, 1986; Kramer et al., 1987; Borowiec et al., 1987).

In conclusion, a new paradigm for the discovery of binding sites for widely dispersed biological regulators has been presented. The assay for these sites is direct because specific scission relies on high-affinity binding of the protein. In the identification of new Fis sites on plasmids and in λ phage, it has been possible to use restriction mapping and PCR to localize the recognition sites at the sequence level. For larger genomes, ligation-mediated PCR techniques that have proven useful in genomic footprinting applications should play a major role (Mueller & Wold, 1989; Garrity & Wold, 1992).

ACKNOWLEDGMENT

We thank Young-Moon Cho for the synthesis of the iodoacetyl-β-alanylamino-OP, Steven Finkel for computer-assisted analysis of potential Fis binding sites in pUC19, and Ralf Landgraf for helpful suggestions throughout the course of this work.

REFERENCES

Ball, C. A., & Johnson, R. C. (1991b) *J. Bacteriol. 173*, 4032−4038.
Borowiec, J. A., Zhang, L., Dwight, S. S., & Gralla, J. D. (1987) *J. Mol. Biol. 196*, 101−111.
Feng, J. A., Yuan, H. S., Finkel, S. E., Johnson, R. C., Grzeskowiak, M. K., & Dickerson, R. E. (1992) in *Structure and Function, vol. 2: Proteins* (Sarma, R. H., & Sarma, M. H., Eds.) pp 1−9, Adenine Press, Schenectady, NY.
Filutowicz, M. F., Ross, W., Wild, J., & Gourse, R. L. (1992) *J. Bacteriol. 174*, 398−407.
Garrity, P. A., & Wold, B. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1021−1025.
Giladi, H., Gottesman, M., & Oppenheim, A. B. (1990) *J. Mol. Biol. 213*, 109−121.
Gille, H., Egan, J. B., Roth, A., & Messer, W. (1991) *Nucleic Acids Res. 19*, 4167−4172.
Hiasa, H., & Marians K. J. (1994) *J. Biol. Chem. 269*, 24999−25003.
Hubner, P., & Arber, W. (1989) *EMBO J. 8*, 557−585.
Johnson, R. C., & Simon, M. I. (1985) *Cell 41*, 781−791.
Kramer, H., Niemoller, M., Amouyal, M., Revet, B., Bergmann, B. W., & Muller-Hill, B. (1987) *EMBO J. 6*, 1481−1491.
Mossing, M. C., & Record, M. T. (1986) *Science 233*, 889−892.

Mueller, P. R., & Wold, B. (1989) *Science 246*, 780−786.
Pan, C. Q., Feng, J., Finkel, S. E., Landgraf, R., Johnson, R., & Sigman, D. S. (1994a) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1721−1725.
Pan, C. Q., Landgraf, R., & Sigman, D. S. (1994b) *Mol. Microbiol. 12*, 335−342.
Pan, C. Q., Landgraf, R., & Sigman, D. S. (1995) *J. Protein Sci. 4* 2279−2288.
Pendergrast, P. S., Ebright, Y. W., & Ebright, R. H. (1994) *Science 265*, 959−962.
Prentki, P., Pham, M.-H., & Galas, D. J. (1987) *Nucleic Acids Res. 15*, 10060.
Riddles, P., Blakeley, T., & Zerner, B. (1983) *Methods Enzymol. 91*, 49−60.
Sigman, D. S. (1990) *Biochemistry 29*, 9097−9105.
Thompson, J. F., & Landy, A. (1988) *Nucleic Acids Res. 16*, 9687−9705.
Thompson, J. F., Moitoso de Vargas, L., Koch, C., Kahmann, R., & Landy, A. (1987) *Cell 50*, 901−908.
Yuan, H. S. , Finkel, S. E., Feng, J.-A., Kaczor-Grzeskowiak, M., Johnson, R. C., & Dickerson, R. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9558−9562.

BI952040Z