Isolation of the Gene Encoding the Hin Recombinational Enhancer Binding Protein

Reid C. Johnson, Catherine A. Ball, Diana Pfeffer, and Melvin I. Simon

PNAS 1988;85;3484-3488
doi:10.1073/pnas.85.10.3484

Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 3484–3488, May 1988
Genetics

Isolation of the gene encoding the Hin recombinational enhancer binding protein

(site-specific DNA recombination/Fis/*Escherichia coli*)

REID C. JOHNSON*†, CATHERINE A. BALL†, DIANA PFEFFER*, AND MELVIN I. SIMON‡

*Department of Biological Chemistry, School of Medicine, and †Molecular Biology Institute, University of California at Los Angeles, Los Angeles, CA 90024; and ‡Division of Biology, California Institute of Technology, Pasadena, CA 91125

Contributed by Melvin I. Simon, December 23, 1987

ABSTRACT **_In vitro_ DNA inversion mediated by the protein Hin requires the presence of a recombinational enhancer sequence located _in cis_ relative to the recombination sites and a protein, Fis, which binds to the enhancer. We have cloned and determined the primary sequence of the gene encoding Fis. The deduced amino acid sequence of Fis indicates that the protein is 98 amino acids long and contains a potential helix-turn-helix DNA binding motif at its carboxyl terminus. The gene encoding Fis maps at 72 min on the _Escherichia coli_ chromosome. The construction of mutant strains of _E. coli_ that lack a functional _fis_ gene demonstrates that Fis is not essential for cell growth under laboratory conditions but is required for high rates of Hin-mediated site-specific inversion _in vivo_.**

A site-specific DNA inversion reaction controls flagellar gene expression in *Salmonella* by switching the orientation of its promoter (1, 2). This reaction has been studied in an *in vitro* system derived from *Escherichia coli* (3–5). In addition to the product of the *hin* gene, which is encoded within the invertible segment, two other host proteins have been shown to be required for high rates of inversion. One of these proteins is the nonspecific double-stranded DNA binding protein HU found in many prokaryotes, and the other is a protein referred to as factor II or Fis (*f*actor for *i*nversion *s*timulation). Gin-mediated inversion of the G segment from bacteriophage Mu also requires the function of the Fis protein (6).

Fis migrates on $NaDodSO_4$ polyacrylamide gels with a M_r of ≈12,000 (4, 6). It binds to two sites within the 60-base-pair (bp) Hin recombinational enhancer sequence (7). The enhancer sequence can function at many different locations and in either orientation to stimulate strand exchange at the recombination sites (3). Plasmid substrates containing mutant enhancers that have lost the ability to bind Fis at either of the two binding sites recombine at rates less than 0.5% of the wild-type substrate. Thus, Fis is intimately involved in enhancer-mediated stimulation of the DNA inversion reaction. The relative position of the two Fis binding sites on the DNA helix is critical for enhancer activity, suggesting that Fis may participate in the formation or stabilization of the synaptic complex by interacting with Hin molecules bound to the DNA at the recombination sites (8).

In addition to its role in site-specific inversion, Fis has been shown to bind to DNA within the bacteriophage λ attachment site, *attP*, and to stimulate excision *in vitro* of bacteriophage λ under conditions of limiting excisionase (9). Furthermore, Fis activity, as measured by sequence-specific binding assays (9) and stimulation of Hin-mediated inversion *in vitro* (R.C.J., unpublished results), decreases markedly as cells enter stationary phase. This growth regulation is opposite to that observed with integration host factor, a heterodimeric protein required for λ-integrase (Int)-mediated site-specific recombination (10). In order to investigate the role and regulation of Fis in *E. coli* and to facilitate genetic and biochemical studies on the mechanism of Fis-enhanced stimulation of site-specific DNA inversion, we have cloned the gene encoding the Fis protein. We present the primary sequence of *fis*§ and preliminary analysis of mutant *E. coli* strains that are deficient in Fis expression.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. Table 1 lists the *E. coli* strains used in this paper. Hfr mapping strains containing Tn*10* transposons 15–25 min from the origin of transfer were from M. Singer and C. Gross (University of Wisconsin, Madison).

Plasmids containing *fis* are depicted in Fig. 1. pRJ741 is the original plasmid isolated after screening an *E. coli* library. Tn*10*mini-*kan* was transposed from λNK1105 onto pRJ741 as described (16). Restriction mapping located the defective transposon on pRJ794 between the *Asu* II and *Mlu* I sites as shown in Fig. 1. pRJ753 was derived from pRJ741 by deleting the DNA between the *Hin*dIII sites in the insert and the vector. pRJ767 was constructed by digesting pRJ753 with *Bst*EII, filling in the protruding ends with T4 DNA polymerase and dNTPs, and ligating in the presence of pRZ102 (colE1::Tn*5*; ref. 17), which had been digested with *Hin*dIII, treated with T4 DNA polymerase plus dNTPs, and then digested with *Sma* I. This resulted in the 1327-bp fragment containing the *neo* gene with its promoter from Tn*5* (18) substituted between nucleotides +68 and +215 in the *fis* coding region.

pMS658 has an 1100-bp *Eco*RV fragment containing the *hin*$^+$ inversion region from pJZ143 (19) cloned between the *Pvu* II sites of pACYC184 (20). pMS21 has an *Eco*RI fragment from λ*fla378* (2) containing the *H* inversion region and the *H2* flagellin gene cloned into the *Eco*RI site of pBR322 (from M. Silverman, Agouron Institute, La Jolla, CA). It is *hin*$^-$ and *H2off*. pRJ792 is pUC18 containing the *hin* gene from pMS621 (3) between the *Eco*RI and the *Hin*dIII sites such that the *hin* gene is transcribed from the *lac* promoter.

The *zhc*-794::Tn*10*mini-*kan* mutation was recombined into the *E. coli* chromosome by transforming pRJ794 into RZ221 (*polA*) and selecting a kanamycin-resistant (Kan^r), ampicillin-resistant (Amp^r) transformant (RJ1562). A deletion that removed the vector sequences but retained the *zhc-794*::Tn*10*mini-*kan* mutation [ampicillin-sensitive (Amp^s), Kan^r;

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Kan^r, kanamycin resistant; Amp^r, Amp^s, ampicillin resistant and sensitive, respectively.

§The sequence reported in this paper is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03245).

Table 1. Bacterial strains used in this study

Strain	Genotype	Source
KL708 (CGSC4248)	F'141 *leuB6 tonA2 lacY1 supE44 gal-6 hisG1 recA1 argG6 rpsL104 malA1 xyl-7 mtl-2 metB1* λ^-	Ref. 11 via B. Bachmann*
DV9 (CGSC6932)	*zhc-9*::Tn*10 panD2 gyrA216 relA1 panF11 spoT1 metB1* λ^-	Ref. 12 via B. Bachmann*
XAC*supF*	*supF* Δ(*pro–lac*) *araD metB argE rif nal*	J. Beckwith†
Hfl	*hfl-1 ser trp leu ilv lys rpsL*	Ref. 13
Ymel	F^+ *supE pro mel*	J. Yin‡
MC1000	Δ*lacX74 araD139* Δ(*ara–leu*)*7697 galU galK strA*	Ref. 14
NO1247	F^- *aroE* λ^-	M. Nomura§
CAG1714	*zhe-7*::Tn*10 cysG* HfrH *thi-1 mal-18 relA1 spoT1* λ^-	C. Gross¶
CAG18457	*zhe-69*::Tn*10 aroE gal-13*	C. Gross¶
RZ211	Δ(*lac–pro*) *ara str recA56 srl*	Ref. 15
RZ221	*polAam* Δ(*lac–pro*) *ara str nal*	T. McNeil‖
RJ1519	Δ(*lac–pro*) *his rpsE cysG zhe-7*::Tn*10 mal* (λ^r)	This paper
RJ1522	MC1000 *cysG zhe*::Tn*10* F'141	This paper
RJ1541	MC1000 *aroE zhe-69*::Tn*10*	This paper
RJ1542	RJ1541 F'141	This paper
RJ1548	RJ1519 F'141 *fis-767*	This paper
RJ1555	*fis-767 hag1427 lacZΔS20* HfrN *rel-1 thi his recA56*	This paper
RJ1561	RZ211 *fis-767*	This paper
RJ1562	RZ221::pRJ794	This paper
RJ1563	RZ221 *zhc-794*::Tn*10*mini-*kan*	This paper
RJ1564	Δ(*lac–pro*) *his rpsE zhc-794*::Tn*10*mini-*kan*	This paper
RJ1580	NO1247 *rpsL zhc-9*::Tn*10*	This paper
RJ2451	RZ211 pRJ792 F'*pro lacI*^{sq}*Z*$_{u118}$ *fzz*::Tn*5*-320(Tet^r)	This paper
RJ2452	RJ2451 *fis-767*	This paper

Tet^r, tetracycline resistant.
*Yale University, New Haven, CT.
†Harvard Medical School, Cambridge, MA.
‡University of Wisconsin, Madison.
§University of California, Irvine.
¶University of Wisconsin, Madison.
‖University of Wisconsin, Madison.

RZ1563] was then obtained. The *fis-767* substitution was transferred to the *E. coli* chromosome as follows: λ*Plac-5cI857nin5S7* was grown on XAC*supF* containing pRJ767. Amp^r, Kan^r lysogens, which were the result of a recombination event between the phage and pRJ767, were obtained by transducing the high-frequency lysogenizing strain Hfl. A lysogen was induced and a phage (Kan^r, Amp^s, Lac$^-$; λRJ796) in which the plasmid vector sequences had been removed by a recombination event was isolated. Restriction analysis confirmed that λRJ796 contained the *fis-767* DNA. RJ1522 λRJ796 was mated with RJ1519, and a Kan^r, Cys$^+$, spectinomycin-resistant (Spc^r) exconjugant (RJ1548) was selected. P1*vir* was grown on RJ1548 and used to transduce the *fis-767* mutation to the desired strain by selecting for Kan^r.

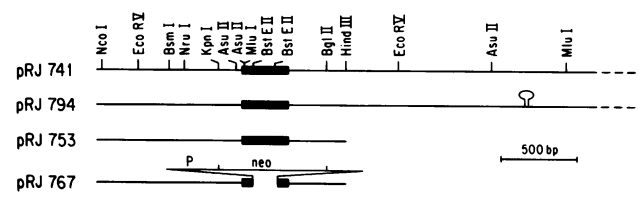

FIG. 1. Physical map of the *E. coli fis* region and plasmids used in this study. The location of the Fis coding sequence is denoted by the thick line. pRJ741 contains an additional ≈1850 bp beyond the right endpoint shown. ♀, location of the Tn*10*mini-*kan* in pRJ794.

Construction and Screening of an *E. coli* Library. Ten micrograms of *E. coli* (Ymel) DNA was digested with 2.5 units of *Sau*3A for 5 min at 37°C such that the average DNA fragment length was greater than 10 kilobases (kb). Two-tenths of a microgram of this *Sau*3A partially digested DNA was ligated with 0.2 μg of pUC9, which was digested with *Bam*HI and treated with calf intestinal phosphatase. The ligation mix was transformed into RZ211 and plated onto Luria–Bertani agar (LB) plates containing ampicillin. The resulting colonies were replicated onto duplicate nitrocellulose filters and prepared for hybridization essentially as described (21). The hybridization was performed in 0.9 M NaCl/0.09 M sodium citrate/10× Denhardt's solution (1× = 0.02% Ficoll/0.02% bovine serum albumin/0.02% polyvinylpyrrolidone)/0.1% NaDodSO$_4$/100 μg of herring testes DNA per ml at 30°C. The filters were successively washed in the same solution without the Denhardt's solution and the DNA at 30°C, 42°C, 45°C, 48°C, and 50°C, followed by autoradiography.

DNA Sequence Analysis. DNA sequencing was performed by the dideoxynucleotide chain-termination method (22) using single-stranded M13 templates or alkaline-denatured plasmid DNA templates (23). Synthetic 17–20 base oligonucleotides were used as primers. The entire sequence, with the exception noted below, was read from both strands using both avian myeloblastosis virus (AMV) reverse transcriptase (Promega Biotec, Madison, WI) and the Klenow fragment (Promega Biotec) of DNA polymerase. Because the sequence from +265 to +285 prevented efficient elongation with both enzymes, the chemical degradation method (24) was used in this region. DNA was end-labeled at either the 5′ or 3′ end of the *Bst*EII site (+214), and the sequence was read from both strands to beyond nucleotide +380.

***In Vivo* Hin Inversion Assays.** Assays using λ*fla406* were performed essentially as described (24). Briefly, RJ2451 (*lacP-hin, fis*$^+$) and RJ2452 (*lacP-hin, fis-767*) were grown in LB medium containing 0.1% maltose and 1 mM isopropyl-β-D-thiogalactoside (IPTG) to a density of 2 × 10^8 cells per ml and adsorbed with λ*fla406off* (previously grown from a single plaque on RJ1561) at a multiplicity of infection of ≈0.5. After 20 min the infected cells were washed twice and grown for an additional 100 min in the same medium. The resulting phage were plated on RJ1561 in the presence of 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal) to distinguish between phage in the "on" (blue plaques) and "off" (white plaques) orientation. The rare Lac$^+$ phage generated after growth on RJ2452 were tested as follows to determine if they represent correct inversion events. Phage from isolated blue plaques were grown overnight on RJ2451 in the presence of IPTG and then plated on RJ1561 as above. Phage that gave rise to >20% white plaques and thus were capable of switching back to the off orientation at the wild-type frequency were considered the result of correct Hin-mediated inversion events.

Inversions in a plasmid substrate were obtained after cotransforming pMS658 and pMS21 into RJ1555. The transformants were stabbed into motility agar [0.5% tryptone (Difco)/0.5% NaCl/0.35% Bactoagar (Difco)], and motile bacteria from "flares" were picked for further analysis.

***In Vitro* Hin Inversion Assays.** Extracts were prepared from 100 ml of cells grown in LB medium and harvested at an OD$_{600}$ = 0.5 essentially as described previously for Fis (4). The extracts were incubated at 80°C for 10 min, and the remaining soluble protein was dialyzed into 20 mM Tris·HCl, pH 7.5/0.2 M NaCl/0.1 mM ethylenediaminetrichloroacetic acid/1 mM dithiothreitol/20% glycerol. Protein concentration was measured by the method of Bradford (25) using bovine serum albumin as the standard. Inversion assays

were performed as described (4) using pMS551 (3) as the substrate DNA and substituting the extracts for purified Fis.

Protein Sequence Analysis. Approximately 1 nmol of purified Fis (4) was subjected to automated Edman degradation on an Applied Biosystems (Foster City, CA) model 470A gas-phase protein Sequenator by the USC Microchemical Core Laboratory.

RESULTS

Cloning of the *fis* Gene. The sequence of the first 29 amino acids (with two unidentified residues) of Fis is shown in Fig. 2. Two sets of 17-base mixed oligonucleotide probes were synthesized corresponding to amino acids 2–7 (probes 1 and 2) and 15–20 (probes 3 and 4) (Fig. 2). These oligonucleotides were labeled with ^{32}P and used to probe Southern blots of restriction endonuclease-digested *E. coli* DNA. Probes 1 and 3 were found to hybridize to a 1.8-kb *Eco*RV fragment, whereas probes 2 and 4 failed to specifically hybridize (data not shown). Since probe 1 gave a significantly stronger hybridization signal than probe 3, we used this probe to hybridize to duplicate filters containing ≈5000 transformants of an *E. coli Sau*3A library in pUC9. Crude preparations of plasmid DNA from seven clones that gave significant hybridization using probe 1 were rescreened with probes 1 and 3. pRJ741, which contains a 5.1-kb insert with an internal 1.8-kb *Eco*RV restriction fragment (Fig. 1), was found to hybridize strongly with both probes, suggesting that it contains the coding region for Fis.

In order to confirm that pRJ741 contains the gene encoding Fis, the plasmid was used as a dideoxy sequencing template after alkaline denaturation. Probe 1 was used to prime synthesis by reverse transcriptase at 42°C. An unambiguous sequence was read beginning at position +27 (see Fig. 3) and extending over 100 bases. Translation of this DNA sequence matches the Fis protein sequence from amino acid 10 to 29 (Fig. 2) and thus demonstrates that the clone contains the *fis* gene. Genetic studies described below further confirm this assignment.

DNA Sequence of the *fis* Gene. The sequence of the *fis* coding region and surrounding DNA is presented in Fig. 3. Translation of the DNA sequence indicates that Fis contains 98 amino acids with a calculated M_r of 11,239 and a pI of 9.4,

```
 1                5                  10
MET-PHE-GLU-GLN-ARG-VAL-ASN-SER-ASP-VAL-LEU-THR-VAL
LEU

 1  TTT GAA CAA CGT GTN AA
      C   G   G   C

 2  TTT GAA CAA CGA GTN AA
      C   G   G A G

    15                  20                  25
SER-THR-VAL-ASN-SER-GLN-ASP-X-VAL-THR-GLN-LYS-X-LEU-ARG-ASP
                                             TRP

 3  ACN GTN AAC TCN CAA GA
               T       G

 4  ACN GTN AAC AGC CAA GA
               T   T   G
```

FIG. 2. Amino-terminal amino acid sequence of Fis and corresponding oligonucleotide probes (1–4). The sequence of the first 29 amino acids of Fis is given. Peaks corresponding to the two phenylthiohydantoin amino acids shown for residues 1 and 25 were obtained. Residue 26 could not be unambiguously determined, and residue 21 was lost prior to chromatography. The nucleotide compositions of the probes are given below the corresponding amino acids. All possible codons are included; however, approximately 90% of the potential arginine codons used in *E. coli* are the two represented in probe 1; the other four are represented in probe 2 (26). The complexity of the probes are as follows: probe 1, 64-fold; probe 2, 128-fold; probe 3, 256-fold; and probe 4, 128-fold degenerate. N indicates that all four bases are present.

```
                -50
GCG TTG GAG GCA TAC TTC GAA AAT TTT GCG TAA ACAGAAATAAAGAGC
ala leu glu ala tyr phe glu asn phe ala ter

          -1 +1
TGACAGAACT ATG TTC GAA CAA CGC GTA AAT TCT GAC GTA CTG ACC
           met phe glu gln arg val asn ser asp val leu thr
            1                                       10
                    50
GTT TCT ACC GTT AAC TCT CAG GAT CAG GTA ACC CAA AAA CCC CTG
val ser thr val asn ser gln asp gln val thr gln lys pro leu
                        20
                           100
CGT GAC TCG GTT AAA CAG GCA CTG AAG AAC TAT TTT GCT CAA CTG
arg asp ser val lys gln ala leu lys asn tyr phe ala gln leu
      30                                            40
                              150
AAT GGT CAG GAT GTG AAT GAC CTC TAT GAG CTG GTA CTG GCT GAA
asn gly gln asp val asn asp leu tyr glu leu val leu ala glu
                           50
                                 200
GTA GAA CAG CCC CTG TTG GAC ATG GTG ATG CAA TAC ACC CGT GGT
val glu gln pro leu leu asp met val met gln tyr thr arg gly
      60                                            70
                                       250
AAC CAG ACC CGT GCT GCG CTG ATG ATG GGC ATC AAC CGT GGT ACG
asn gln thr arg ala ala leu met met gly ile asn arg gly thr
    ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ 80 ▬▬▬                ▬▬▬▬▬▬▬▬▬▬▬▬
                                                 300
CTG CGT AAA AAA TTG AAA AAA TAC GGC ATG AAC TAA TTCAGGTTAGCT
leu arg lys lys leu lys lys tyr gly met asn ter
▬▬▬▬▬▬▬ 90 ▬▬▬▬▬▬▬▬▬▬▬▬
                                          350
AAATGCTTGATTAAAAAGGCGCTACTCGGCATGGGGAAGCGCCTTTTTTATAGGTGTCAC
                                    400
AAAGGGAGTGACC ATG AGA ACA GGA TGT GAA CCG ACC CGG TTT GGT
              met arg thr gly cys glu pro thr arg phe gly
```

FIG. 3. DNA sequence of the *fis* gene. The sequence of the DNA in the *fis* region is given along with the predicted protein sequence. The nucleotides are numbered above the sequence beginning with the adenosine (+1) of the initiating methionine of Fis. The amino acids of Fis are numbered below the sequence. The carboxyl-terminal and amino-terminal portions of the open reading frames located before and after *fis*, respectively, are also shown. The hatched bars indicate the location of the potential helix-turn-helix DNA binding motif in Fis.

which are consistent with what has been observed with the purified protein (refs. 4 and 6 and R.C.J., unpublished results). There are no cysteine, histidine, or tryptophan codons represented in the sequence. The initiating methionine is preceded by a poor Shine–Delgarno translation initiation sequence that does not fit the rules of Stormo *et al.* (27). In addition, there is an unusually high number of rare codons [11% calculated from the list of Konigsberg and Godson (28)], which may also contribute to the relatively low abundance of Fis in *E. coli* (4). An open reading frame of greater than 300 amino acids terminates 25 bp 5′ of the initiating methionine of *fis*. Beginning 86 bp from the 3′ end of *fis* is another open reading frame of greater than 100 amino acids. Preliminary experiments indicate that the upstream open reading frame is cotranscribed with *fis*, whereas the downstream gene is transcribed at least in part from its own promoter (data not shown). The functions of these two open reading frames, if any, are not known.

Location of *fis* on the *E. coli* Chromosome. To facilitate mapping of *fis*, a 1.8-kb Tn*10*mini-*kan* transposon was inserted approximately 1.5-kb downstream of *fis* in pRJ741 (Fig. 1). This plasmid (pRJ794) was integrated into the chromosome, and a segregant (RJ1563) that had lost the vector sequences but retained the kanamycin-resistance determinant was obtained. Hybridization experiments with DNA prepared from RJ1563 demonstrated that the wild-type 3.2-kb *Eco*RV fragment was replaced by a 5.0-kb fragment containing Tn*10*mini-*kan*, indicating that the plasmid had inserted into the correct locus. *fis* was initially localized within the chromosomal interval between min 66 (KL14; ref. 11) and 84 (KL228; ref. 11) by Hfr-mediated crosses. P1*vir*-mediated transduction established that *fis* is located at 72 min, between *aroE* and *fabE*. Three- and four-factor crosses

with markers in the region (e.g., Fig. 4*B*) established the order of genes as that depicted in Fig. 4*A*.

Effect of a *fis* Null Mutation on *E. coli* Growth and Hin Inversion. DNA encoding amino acids 22–73 in *fis* was substituted with the *neo* gene from transposon Tn*5* (pRJ767; Fig. 1). The mutated *fis* gene (*fis-767*) was recombined onto F'141 and transferred to RJ1519 to give RJ1548 (see *Materials and Methods*). A P1*vir* lysate was prepared on RJ1548 and used to transduce *fis-767* into RJ1542 (containing F'141 and thus diploid for the *fis* locus) and RJ1541 (haploid for the *fis* locus). Equivalent numbers of Kan^r transductants were obtained (4.4×10^{-6} and 3.8×10^{-6} per plaque-forming unit, respectively), demonstrating that *fis* is not essential for cell growth. Hybridization experiments with DNA prepared from RJ1561 (*fis::767*) demonstrated that the wild-type *fis* gene was replaced by the *fis::767* substitution (data not shown).

In order to confirm that the *fis* gene was inactivated by the *fis-767* mutation, extracts were prepared from RZ211 (*fis*$^+$) and RJ1561 (*fis-767*) and used to complement *in vitro* Hin-inversion assays in the absence of purified Fis. No detectable Fis activity was observed in the RJ1561 extract, even when 40 times the amount of protein that gave 30% inversion products with the RZ211 extract was added (Table 2).

A λ derivative (λ*fla406*) in which *lacZ* expression is controlled by the *hin* inversion region was grown on RJ2451 (*lacP-hin, fis*$^+$) and RJ2452 (*lacP-hin, fis-767*) under conditions of Hin overproduction. After 2 hr of growth on RJ2451, 27% of the phage were switched to the on orientation (Lac$^+$), whereas only 0.01% of the phage produced in RJ2452 contained inversions (Table 2). Thus in this assay, Hin inversion *in vivo* is decreased by ≈4000-fold in the absence of Fis.

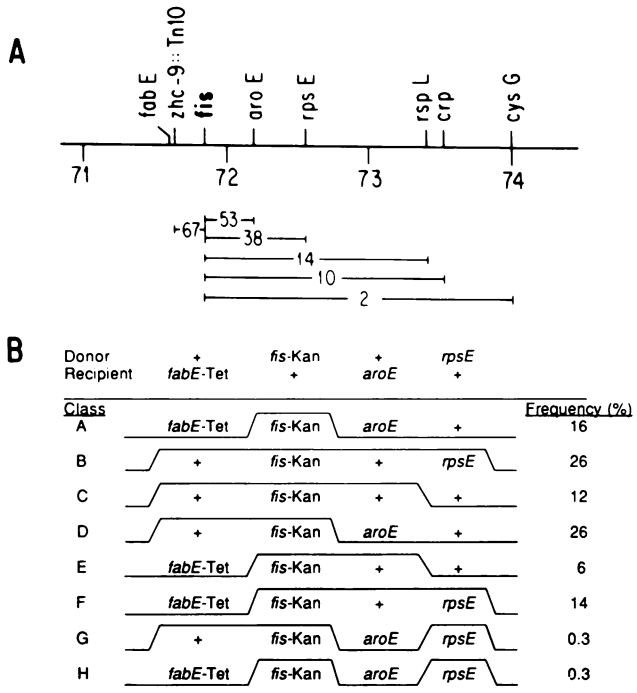

FIG. 4. Location of *fis* in the *E. coli* chromosome. (*A*) The region of the *E. coli* chromosome between 71 and 74 min is shown with the linkage values determined by P1*vir* transduction between markers in the region and *fis* denoted. (*B*) Four-factor cross determining gene order in the *fis* region. P1*vir* was grown on RJ1564 (donor) and used to transduce RJ1580 (recipient). Three hundred seven Kan^r (zhc-794::Tn*10*mini-*kan*; 1.5 kb 3′ of *fis*) transductants were screened for spectinomycin resistance (*rpsE*), *aroE*, and tetracycline resistance (*zhc-9*::Tn*10*). The location of *zhc-9*::Tn*10*, which is 95% cotransducible with *fabE* was reported by Valleri and Rock (12). The crossover events required to generate the recombinants are denoted.

Table 2. Effect of the *fis-767* mutation on Hin-mediated inversions *in vitro* and *in vivo*

Strain	Protein*, μg	% recombinants	Inversions per molecule
		In vitro	
RZ211 (*fis*$^+$)	0.5	31†	0.48
RJ1561 (*fis-767*)	0.5	<1†	<0.01
RJ1561 (*fis-767*)	20.0	<1†	<0.01
		In vivo	
RJ2451 (*fis*$^+$)		27‡	0.39
RJ2451 (*fis-767*)		0.01‡	0.0001

The number of inversions per molecule was calculated from the percent recombinants as previously described (29).
*Amount of protein extract added to the *in vitro* inversion reaction in place of purified Fis.
†Obtained by scanning photographic negatives of gels.
‡Number of phage $\times 10^{-2}$ switched to the on orientation and capable of additional inversions per total plaque-forming units.

Inversions generated in a plasmid substrate (pMS21) in the absence of Fis (RJ1555) were also analyzed. The restriction patterns were identical to the inversions generated in the presence of Fis (data not shown), suggesting that Fis enhances the rate of inversion and not the specificity of site selection.

DISCUSSION

Fis is a small basic protein (98 amino acids, pI = 9.4), which is required for high rates of Hin-mediated DNA inversion *in vitro* (4) and *in vivo* (this paper). It functions to stimulate DNA strand exchange when bound to a DNA segment that can be located at numerous positions relative to the recombination sites. Thus, Fis must contain at least two activities: it must recognize and bind to specific sites on DNA and it must somehow stimulate recombination.

Analysis of the Fis sequence suggests a mode of DNA binding. There are no cysteines or histidines in the sequence, making the presence of a "zinc finger" motif unlikely, and there is no homology to the type II DNA binding proteins such as HU or integration host factor. However, located at the carboxyl terminus (amino acids 74–93) of Fis is a region that bears similarity to the helix-turn-helix DNA binding motifs, which have been characterized for repressors and other DNA binding proteins (for a review, see ref. 30). The carboxyl-proximal helix is highly amphipathic, and the entire 20-amino-acid bihelical region displays 25–45% identity to other known (e.g., Cro and cAMP receptor protein) and presumed [Hin, InsA, NtrC (NR_1), and TnpR] helix-turn-helix DNA binding motifs (Fig. 5).

The homology with NtrC (NR_1) is particularly striking as there is a 45% identity between the carboxyl-terminal 29 amino acids, which includes the proposed DNA binding region for both proteins (Fig. 5). NtrC (NR_1) is another example of a bacterial regulatory protein that functions in a position-independent manner (35). It activates transcription of RpoN (σ^{54})-dependent promoters found associated with nitrogen assimilation and fixation operons. There is little similarity, however, between the DNA sequence of the binding sites for Fis and the binding sites for NtrC (NR_1). Extended homology is also seen between Fis and InsA, an insertion sequence 1 (IS*1*)-encoded 91 amino acid protein required for transposition (33, 34). Alignment of these two proteins indicates a 31% match over the carboxyl-terminal 39 amino acids, with two gaps of one amino acid in the InsA sequence (Fig. 5) (or one gap of two amino acids in InsA generating a 28% match). InsA binds to the ends of IS*1*, but its precise role in transposition is not yet known.

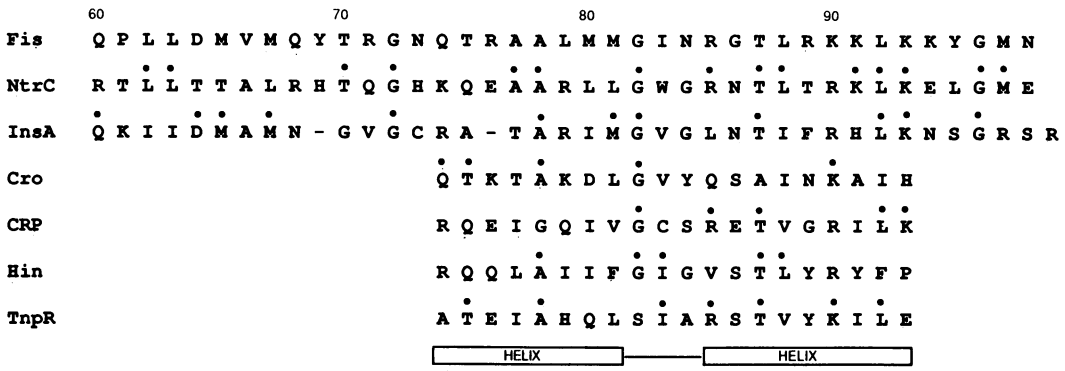

FIG. 5. Sequence of the carboxyl-terminal 39 amino acids of Fis and their relationship to the sequence of other DNA binding proteins. Below the sequence of Fis are the carboxyl-terminal residues from 442 to 470 of NtrC (NR_1) from *Klebsiella pneumoniae* (31, 32) and from 54 to 91 of InsA from IS*1* (33, 34). The helix-turn-helix regions from Cro, cAMP receptor protein (CRP), Hin, and TnpR are taken from Pabo and Sauer (30). Dots indicate identical amino acids, and dashes designate gaps used to maximize alignment.

If Fis is interacting with DNA via a helix-turn-helix motif, it is likely to display some differences from those characterized to date, as Fis seems to recognize sites on DNA that differ greatly in sequence. The sequences at the two Fis binding sites of the Hin enhancer display little resemblance to the site overlapping the Xis binding site in λ*attP*, where Fis also binds (9). Other Fis binding sites show little similarity to either of these sites (M. Bruist and M.I.S., unpublished results and ref. 9). A feature in common between the enhancer and λ*att* sites is that they both appear to contain "bent" DNA (8, 9), suggesting that Fis may be recognizing some structural feature in addition to the nucleotide sequence of its DNA binding site.

Like integration host factor (IHF), Fis is not essential for the growth of *E. coli* under normal laboratory conditions in minimal or rich media. Indeed, strains containing substitution mutations in both *fis* and *hip* (*himD*, one of the subunits of IHF) are still viable. The ability to propagate strains lacking Fis will allow the investigation of the role of Fis in other reactions such as transposition or adaptation to different growth conditions. Strains lacking Fis generate Hin-mediated inversions at an extremely low rate, even in the presence of high levels of Hin, consistent with that observed *in vitro*. The cloned gene along with host cells lacking *fis* will allow the isolation and characterization mutations in *fis* with respect to their effect on the inversion reaction.

Note Added in Proof. R. Weisberg (National Institutes of Health) has pointed out that the NtrC-like protein from "*Bradyrhizobium parasponiae*" (36) displays greater homology to the *E. coli* Fis protein than does the *K. pneumoniae* NtrC. In the carboxyl-terminal region of the two proteins, there are 16 identities over a 22-amino acid region, and statistically significant homology extends over the entire Fis sequence.

We thank B. Bachmann and C. Gross for strains, D. Glitz and S. Horvath for oligonucleotide synthesis, L. Williams for protein sequencing, and A. Glasgow and K. Hughes for critically reading the manuscript. This work was supported by a grant from the California Institute for Cancer Research, a Basil O'Connor Starter Scholar Research Award No. 5-623 from the March of Dimes Birth Defects Foundation, the Searle Scholars Program/The Chicago Community Trust, and grant GM38509 from the National Institutes of Health to R.C.J. C.A.B. was supported in part by National Research Service Award GM07104 from the National Institutes of Health. M.I.S. was supported by a grant from the National Science Foundation.

1. Zieg, J., Silverman, M., Hilman, M. & Simon, M. (1977) *Science* **196**, 170–172.
2. Silverman, M. & Simon, M. (1980) *Cell* **19**, 845–854.
3. Johnson, R. C. & Simon, M. I. (1985) *Cell* **41**, 781–789.
4. Johnson, R. C., Bruist, M. F. & Simon, M. I. (1986) *Cell* **46**, 531–539.
5. Johnson, R. C. & Simon, M. I. (1987) *Trends Genet.* **3**, 262–267.
6. Koch, C. & Kahmann, R. (1986) *J. Biol. Chem.* **261**, 15673–15678.
7. Bruist, M. F., Glasgow, A. C., Johnson, R. C. & Simon, M. I. (1987) *Genes Dev.* **1**, 762–772.
8. Johnson, R. C., Glasgow, A. C. & Simon, M. I. (1987) *Nature (London)* **329**, 462–465.
9. Thompson, J. F., de Vargas, L. M., Koch, C., Kahmann, R. & Landy, A. (1987) *Cell* **50**, 901–908.
10. Bushman, W., Thompson, J. F., Vargas, L. & Landy, A. (1985) *Science* **230**, 906–911.
11. Low, K. B. (1972) *Bacterial Rev.* **36**, 587–607.
12. Vallari, D. S. & Rock, C. O. (1985) *J. Bacteriol.* **164**, 136–142.
13. Belfort, M. & Wulff, D. L. (1971) *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 739–742.
14. Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* **138**, 179–207.
15. Johnson, R. C., Yin, J. C. P. & Reznikoff, W. S. (1982) *Cell* **30**, 873–882.
16. Way, J. C., Davis, M. A., Morisato, D., Roberts, D. E. & Kleckner, N. (1984) *Gene* **32**, 369–379.
17. Jorgensen, R. A., Rothstein, S. J. & Reznikoff, W. S. (1979) *Mol. Gen. Genet.* **177**, 65–72.
18. Beck, E., Ludwig, G., Auerswald, E. A., Reiss, B. & Schaller, H. (1982) *Gene* **19**, 327–336.
19. Zieg, J. & Simon, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4196–4200.
20. Chang, A. C. Y. & Cohen, S. N. (1978) *J. Bacteriol.* **134**, 1141–1156.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Mierendorf, R. C. & Pfeffer, D. (1987) *Methods Enzymol.* **152**, 556–562.
24. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 497–560.
25. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254.
26. Ikemura, T. (1985) *Mol. Biol. Evol.* **2**, 13–34.
27. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10**, 2971–2996.
28. Konigsberg, W. & Godson, G. N. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 687–691.
29. Bruist, M. F. & Simon, M. I. (1984) *J. Bacteriol.* **159**, 71–79.
30. Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293–321.
31. Buikema, W. J., Szeto, W. W., Lemley, P. V., Orme-Johnson, W. H. & Ausubel, F. M. (1985) *Nucleic Acids Res.* **13**, 4539–4555.
32. Drummond, M., Whitty, P. & Wootton, J. (1986) *EMBO J.* **5**, 441–447.
33. Machida, Y., Machida, D. & Ohtsubo, E. (1984) *J. Mol. Biol.* **177**, 229–246.
34. Zerbib, D., Jakowec, M., Prentki, P., Galas, D. J. & Chandler, M. (1987) *EMBO J.* **6**, 3163–3169.
35. Reitzer, L. J. & Magasanik, B. (1986) *Cell* **45**, 785–792.
36. Nixon, B. T., Ronson, C. W. & Ausubel, F. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7850–7854.