The transactivation region of the Fis protein that controls site-specific DNA inversion contains extended mobile β-hairpin arms

Martin K.Safo[1], Wei-Zen Yang, Leah Corselli[2], Sarah E.Cramton[2,3], Hanna S.Yuan[4] and Reid C.Johnson[2,4]

Institute of Molecular Biology, Academia Sinica, Taipei, Taiwan 11529, Republic of China and [2]Department of Biological Chemistry and Molecular Biology Institute, University of California at Los Angeles, Los Angeles CA 90095-1737, USA

[1]Present address: Institute of Structural Biology and Drug Discovery, Biotech Research Park, 800 E.Leigh Street, Richmond, VA 23298, USA
[3]Present address: Mikrobielle Genetik, Universität Tübingen, Waldhäuser Straße 70/8, D-72076 Tübingen, Germany

[4]Corresponding authors
e-mail: rjohnson@biochem.medsch.ucla.edu
e-mail: mbyuan@ccvax.sinica.edu.tw

The Fis protein regulates site-specific DNA inversion catalyzed by a family of DNA invertases when bound to a *cis*-acting recombinational enhancer. As is often found for transactivation domains, previous crystal structures have failed to resolve the conformation of the N-terminal inversion activation region within the Fis dimer. A new crystal form of a mutant Fis protein now reveals that the activation region contains two β-hairpin arms that protrude over 20 Å from the protein core. Saturation mutagenesis identified the regulatory and structurally important amino acids. The most critical activating residues are located near the tips of the β-arms. Disulfide cross-linking between the β-arms demonstrated that they are highly flexible in solution and that efficient inversion activation can occur when the β-arms are covalently linked together. The emerging picture for this regulatory motif is that contacts with the recombinase at the tip of the mobile β-arms activate the DNA invertase in the context of an invertasome complex.
*Keywords*: crystal structure/DNA inversion/Fis/site-specific DNA recombination/transactivation

Introduction

Control proteins are often composed of genetically and physically separate DNA-binding and regulatory regions. The DNA-binding domains are usually well ordered, which has led to a wealth of structural information on various DNA-binding motifs. On the other hand, regulatory regions that are involved in protein–protein communication are often disordered, which has resulted in the current lack of information on their structures. The *Escherichia coli* Fis protein is an example of a control protein with a defined helix–turn–helix DNA-binding domain but, heretofore, the region required for activation of DNA inversion by a family of DNA invertases had not

been visible in electron density maps obtained from crystals of wild-type and mutant proteins (Kostrewa *et al.*, 1991; Yuan *et al.*, 1991, 1994). We now demonstrate that this region consists of a flexible β-hairpin arm that extends from each subunit.

The Fis protein was first identified because of its critical role in regulating the Hin- and Gin-catalyzed site-specific DNA inversion reactions (Johnson *et al.*, 1986; Koch and Kahmann, 1986). Inversion of a 1 kb DNA segment of the *Salmonella* chromosome by Hin regulates the alternative synthesis of flagellin proteins (Glasgow *et al.*, 1989), while inversion of a 3 kb DNA segment in the Mu phage genome by Gin regulates the alternative synthesis of tail fiber proteins (Koch *et al.*, 1987). The Hin DNA invertase is catalytically inactive when bound to the specific recombination sites, *hixL* and *hixR*, at the boundaries of the invertible segment. Hin can assemble the *hix* sites into a synaptic complex, but these complexes are only weakly catalytically active under special solvent conditions *in vitro* (Johnson and Bruist, 1989; Heichman and Johnson, 1990). Efficient catalytic activation of Hin complexes requires the co-assembly of a recombinational enhancer bound by two dimers of Fis into an invertasome structure as shown in Figure 1A (Heichman and Johnson, 1990). Fis can be cross-linked specifically to Hin within this structure, as detected by immunoelectron microscopy. The three DNA segments, the 26 bp *hixL* and *hixR* recombination sites and the 65 bp enhancer, are organized within the invertasome in a precise topology at the base of a branch on plectonemically supercoiled DNA (Kanaar *et al.*, 1988; Heichman *et al.*, 1991). This assembly of the Fis-activated invertasome at the enhancer, which requires DNA supercoiling, overwhelmingly biases the recombination product towards inversion of the DNA segment between the two recombination sites. DNA invertase mutants which are catalytically active without Fis no longer require DNA supercoiling and are not topologically limited with respect to the types of recombination products they can generate (Haffter and Bickle, 1988; Klippel *et al.*, 1988; Crisona *et al.*, 1994). The locations and properties of these mutants, combined with other biochemical and structural studies on these enzymes and the homologous resolvases, suggest that the mechanism of Fis activation of the DNA invertase may involve an induced conformational change within the Hin dimer interface that occurs upon assembly of the invertasome (Hughes *et al.*, 1993; Klippel *et al.*, 1993; Haykinson *et al.*, 1996).

Fis is a homodimer of 98 amino acids in which each subunit contains a helix–turn–helix DNA-binding motif at its C-terminus. Crystal structures of wild-type or mutant Fis proteins were not able to resolve the N-terminal 25 amino acids but revealed the remaining 73 amino acids as a compact ellipsoid with each subunit containing four α-helices A–D (Kostrewa *et al.*, 1991, 1992; Yuan *et al.*,
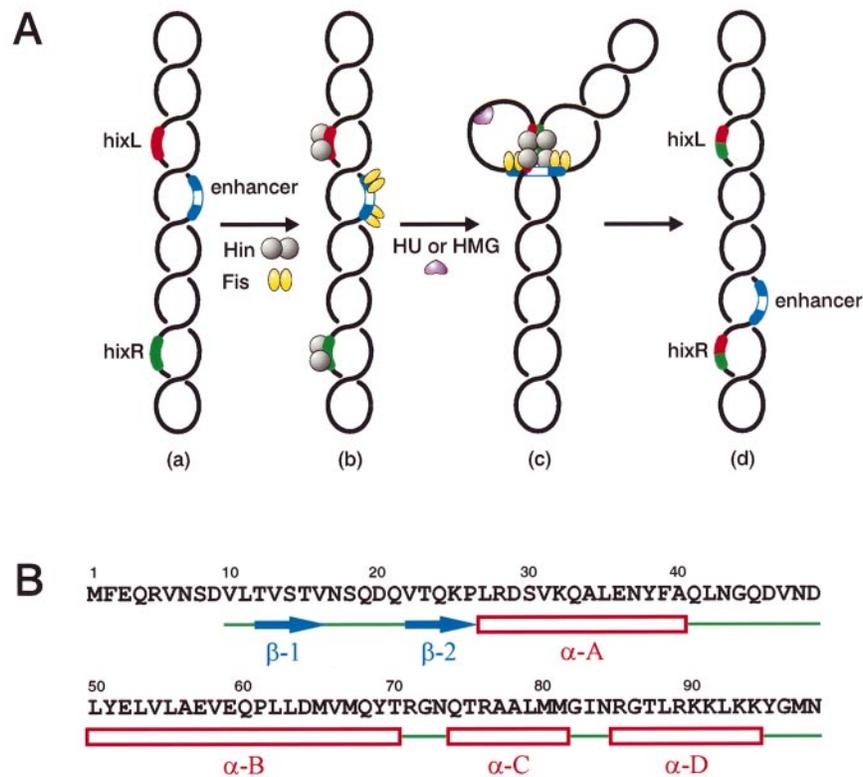
Fig. 1. (**A**) Schematic diagram of the Hin-catalyzed DNA inversion reaction. (a and b) Hin binds to the two recombination sites, *hixL* and *hixR*, and Fis binds to the two domains of the recombinational enhancer on a supercoiled DNA substrate. (c) An invertasome complex is assembled with the aid of HU or HMG1/2 proteins. This structure is catalytically competent for DNA cleavage and strand exchange. (d) After resolution, the DNA segment between *hixL* and *hixR* has been inverted. (**B**) Amino acid sequence and secondary structure as defined by PROCHECK (Laskowski *et al.*, 1993) of Fis mutant K36E.

1991, 1994). Previous mutational studies of Fis indicated that the N-terminal region between amino acids 10 and at least 34 was required for controlling DNA inversion (Koch *et al.*, 1991; Osuna *et al.*, 1991). Deletions within this segment, which is located within the N-terminal disordered region and extends into α-helix A, abolished the ability of Fis to stimulate inversion, and a few amino acid substitutions within this region affected DNA inversion to varying extents. DNA binding and other reactions regulated by Fis, such as phage λ excision and transcription, were relatively unaffected by these mutations (Koch *et al.*, 1991; Osuna *et al.*, 1991; Gosink *et al.*, 1993). Thus, the N-terminal region of Fis appears to be specific for controlling DNA inversion and probably contacts the DNA invertases to regulate their catalytic activity.

In this study we have performed extensive mutagenesis of the Fis inversion activation region to identify the functionally important residues. One of the mutants that was isolated crystallized into a form that now reveals the structure of the inversion activation region. The polypeptide chain from each subunit within the N-terminal region is organized into a β-hairpin arm that extends out from the molecule and is connected at its base by a network of hydrophobic interactions and hydrogen bonds to the body of the protein. The biochemical properties of amino acid substitutions and chemical modifications throughout the region identified those residues which most likely interact with Hin and those which perform important structural roles in organizing the β-arms. Disulfide cross-linking between cysteines engineered into the β-arms is

fully consistent with the functional relevance of the crystal structure and directly demonstrates the mobility of the arms.

Results and discussion

Mutational analysis of the Fis N-terminal inversion activation region

Mutagenesis and recombination assays. The N-terminal inversion activation domain of Fis was identified initially from the properties of internal deletions within the region (Koch *et al.*, 1991; Osuna *et al.*, 1991). These studies indicated that the segment between amino acids 10 and 34 was critical for stimulating DNA inversion, but the lack of single amino acid substitutions with significant effects (excluding proline or glycine replacements), combined with the lack of structural information prior to amino acid 25, prevented an understanding of the molecular nature of this region. To identify critical amino acids within the inversion activation domain, we subjected this region to localized mutagenesis combined with *in vivo* screening for mutants with decreased activity and to site-directed mutagenesis. Amino acid substitutions were obtained at nearly all positions within the disordered region, and residues within the A α-helix whose side chains were largely solvent exposed (see Table I). In many cases, the native amino acid was replaced by cysteine, which generated a mutant protein with a single cysteine. While the cysteine side chain is not as chemically inert as the more commonly used alanine replacement, it enabled

Table I. Effects of Fis N-terminal mutations on Hin-catalyzed DNA inversion

Mutant	DNA inversion		Sensitivity to chemical modification[c]	Comment[d]
	In vivo[a]	*In vitro*[b]		
WT	++++	100		Wild-type
L11C	−	0.9		disrupts hydrophobic interactions
V13C	+++	42	NEM: <2	disturbs van der Waal's contacts
V13I	+++	66		disturbs van der Waal's contacts
S14C	++++	135		not involved
T15C	++++	96		not involved
T15I	+++	90		not involved
V16C	+	17	NEM: <1	probable contact site for Hin
N17A	++	8		destroys H-bonds
N17C	+++	38	NEM: <2	destroys H-bonds
N17D	+++	88		alters H-bonds
S18C	++++	100		not involved
Q19C	+++	87		not involved
D20N	+	7		probable contact site for Hin
D20G	+	2		probable contact site for Hin
D20K	−	0.9		probable contact site for Hin
Q21C	+++	88		not involved
Q21L	+++	94		not involved
V22C	+	4		probable contact site for Hin
T23A	++++	nd		not involved
T23C	++++	96		not involved
Q24C	++++	99		not involved
K25C	+++	51		destroys H-bond
P26A	++	40		structural change
P26L	++	17		structural change
D29C	+++	81		not involved
S30C	++++	98		not involved
K32C	+++	55		destroys H-bond
Q33C	++	14	AE-8: 6	destroys H-bond and/or contact site for Hin
A34C	+++	29	NEM: <1	disrupts hydrophobic interactions
A34T	++	9		disrupts hydrophobic interactions
K36C	+++	54		destroys H-bond and/or contact site for Hin
K36E	+	14		alters H-bond and/or contact site for Hin
N37C	++++	87		not involved
Q41C	++++	53		not involved

[a]*In vivo* inversion rates as measured by color development on lactose MacConkey media: ++++ indicates red colonies developed after ~25 h incubation as observed for wild-type, +++ red color development was slightly delayed compared with wild-type, ++ red colonies developed after 30–34 h, + some red colonies after 40 h, − colonies remained white after 48 h as observed for the vector (no *fis*) control.
[b]*In vitro* inversion rates using purified proteins. The values for the mutants are expressed relative to reactions containing wild-type Fis (set at 100), which gave ~0.6 inversions/min/substrate DNA molecule. nd indicates not determined.
[c]Positions where a particular chemical modification of a cysteine had a detrimental effect that was significantly greater than the original cysteine substitution are noted along with the inversion rate relative to wild-type Fis. NEM indicates modification by *N*-ethylmaleimide; AE-8 indicates modification by *N*-[iodoethyl]trifluoroacetamide. Where applicable, the modification having the greatest detrimental effect is given.
[d]The probable structural consequence of the mutation or functional role of the amino acid in Hin-catalyzed DNA inversion is given.

further chemical modifications and cross-linking at the mutant position as described below.

The ability of each mutant to stimulate Hin-catalyzed recombination was qualitatively measured *in vivo* (Table I). Plasmids containing the *fis* mutant genes were transformed into a Δ*fis E.coli* strain in which inversion of the orientation of the promoter within the *H* control segment from *Salmonella* turns on *lacZ* expression (Osuna *et al.*, 1991). We confirmed that those mutants which displayed low *in vivo* inversion activity expressed normal levels of Fis by immunoblotting cell extracts with anti-Fis antibody (data not shown) and by their proficiency in stimulating other Fis-regulated reactions (see below). Purified preparations of mutant proteins were also assayed quantitatively *in vitro* for their ability to activate Hin-catalyzed inversion. An example of an *in vitro* inversion assay on a representative set of mutants is given in Figure 2, and initial rates of Fis-activated Hin inversion are given in Table I. In addition, each cysteine-containing mutant was alkylated

with *N*-ethylmaleimide (NEM) and *N*-[iodoethyl]trifluoroacetamide (AE-8), which introduces a primary amine onto the side group, and subsequently tested for stimulation of inversion. Those positions where a particular chemical modification resulted in a protein with significantly less activity than the original cysteine substitution are also noted in Table I.

Effect of side chain substitutions on Hin-catalyzed DNA inversion. Relatively few sites within the inversion activation region were found to be sensitive to mutation. The most sensitive positions were Leu11, Asp20 and Val22, which are all located in the disordered region. A cysteine substitution at positions 11 or 22 and a lysine or glycine substitution at position 20 essentially abolished inversion stimulation by Fis. The activity of the relatively conservative change D20N was decreased >10-fold. Cysteine substitutions at positions (Val)13, (Val)16 and (Asn)17 within the disordered region resulted in 2.5- to 6-fold
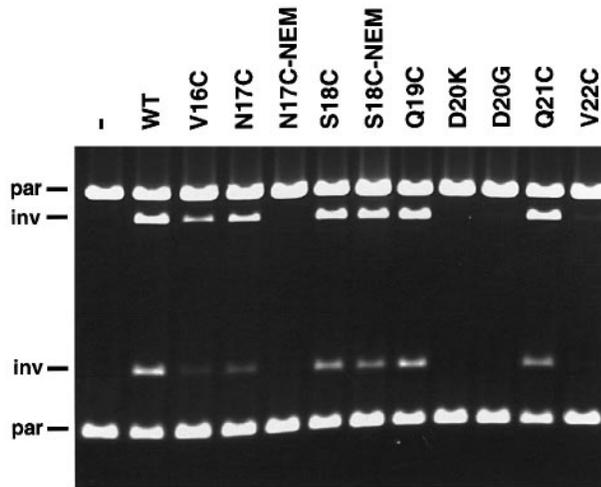
Fig. 2. Activation of Hin-catalyzed DNA inversion by Fis mutants with amino acid substitutions in the β-hairpin loop region. *In vitro* inversion reactions were incubated for 1 min followed by digestion of the DNA with restriction enzymes to determine the orientation of the invertible segment. (par) indicates the location of DNA fragments in the parental orientation; (inv) indicates the location of DNA fragments in the inverted orientation after agarose gel electrophoresis. The Fis mutants used in each reaction are listed on the top, with NEM designating after alkylation with *N*-ethylmaleimide. WT indicates wild-type and (–) indicates no Fis added.

decreases in initial rates of inversion, and modification of each of these residues with NEM abolished activity. Within the A α-helix, Pro26, Gln33, Ala34 and Lys36 were found to be important residues. An alanine or leucine substituted for the proline at position 26 lowered inversion rates 2.5- and 6-fold, respectively. The activity of Q33C was reduced 7-fold and that of Q33C-AE8 was reduced >16-fold. Inversion activation by A34T and A34C was 9–30% of wild-type, and A34C-NEM had no activity. The nature of the side chain at residue (Lys)36 is also important since a glutamic acid strongly decreased Fis activation of inversion, but a cysteine or its chemically modified forms affected activity only slightly. The structural consequences of each of these side chain substitutions and their probable functional consequences are discussed below in the context of the crystal structures.

Effect of N-terminal mutations on other Fis-regulated reactions. The N-terminal mutants were also assayed for their ability to stimulate phage λ excision *in vivo* (Ball and Johnson, 1991), to activate *proP* transcription *in vivo* (Xu and Johnson, 1995) and to bind and bend DNA *in vitro* (Wu and Crothers, 1984). Data for selected mutants which elicited a strong Hin inversion phenotype are given in Table II. None of these mutants were significantly affected for these other activities, except for K36E, which was partially defective in transcriptional activation and λ excision, and V22C, which was partially defective in λ excision. These results are consistent with other studies indicating that residues within the C-terminal third of Fis mediate these activities (Koch *et al.*, 1991; Osuna *et al.*, 1991; Bokal *et al.*, 1997; S.McLeod, S.E.Cramton and R.C.Johnson, unpublished).

Crystal structure of K36E reveals the structure of the inversion activation domain
Multiple crystal forms and structure quality. Crystallization trials were performed on a number of Fis mutants

Table II. Functional properties of selected N-terminal Fis mutants

Mutant	Inversion[a]	λ excision[b]	Transcription[c]	DNA binding[d]
WT	100	100	100	2×10^{-9}
Δ(17–21)[e]	<0.1	98	46	2×10^{-9}
L11C	0.9	73	69	2×10^{-9}
D20G	2	186	123	2×10^{-9}
D20K	0.9	141	80	2×10^{-9}
V22C	4	25	51	2×10^{-9}
Q33C	14	92	65	2×10^{-9}
A34C	29	62	61	2×10^{-9}
K36E	13	23	28	2×10^{-9}

[a]Inversion rates *in vitro* taken from Table I.
[b]Fold stimulation of λ excision *in vivo*. Values are the average of at least two experiments and are expressed relative to wild-type Fis (set at 100), which gave a 270- to 1150-fold increase in the number of plaque-forming units, per ml lysate, over the no Fis control (4×10^{6}–1.5×10^{7}).
[c]Fis activation of *proP* P2 transcription *in vivo*. Values are the average of at least two experiments and are expressed relative to wild-type (set at 100), which gave a 150- to 200-fold increase in ProP–LacZ activity over the no Fis control (4–7.5 U). *In vitro* transcription reactions with purified Fis mutants and RNA polymerase (σ^{38}) gave similar activation relative to wild-type for each of these mutants (S.McLeod and R.C.Johnson, unpublished).
[d]K_d (M) as determined by gel mobility shift assays.
[e]Δ(17–21) is an in-frame deletion of amino acids 17–21 (Osuna *et al.*, 1991).

in an attempt to obtain a structure of the N-terminal region. The Fis mutant K36E was found to crystallize in two different forms, orthorhombic and hexagonal, by the vapor diffusion method using different crystallization conditions. The orthorhombic crystals are isomorphous to those of the wild-type Fis (Kostrewa *et al.*, 1991; Yuan *et al.*, 1991) and the P61A (Yuan *et al.*, 1994) mutant; therefore, the orthorhombic K36E structure initially was refined using the higher resolution structure (1.9 Å) of the P61A mutant as a template. The structure of the hexagonal K36E was determined subsequently by the molecular replacement method using the refined orthorhombic K36E structure as a searching model. There are two molecules per asymmetric unit for both the orthorhombic and hexagonal forms. The Fis dimer is numbered from 1 to 98 for one subunit (monomer <u>a</u>) and from 101 to 198 for the other subunit (monomer <u>b</u>). The final hexagonal K36E model includes one Fis dimer (residues 10–98, 110–198) and 22 water molecules, yielding an *R*-factor of 22.2% and an R_{free} of 27.2% based on 6577 reflections ($F >1 \sigma F$) in the resolution range of 8.0–2.65 Å. The final orthorhombic K36E structure contains one Fis dimer (residues 10–13, 24–98, 110–113 and 124–198) and 74 water molecules, yielding an *R*-factor of 21.2% and an R_{free} of 28.2% based on 9786 reflections ($F >2 \sigma F$) in the resolution range of 8.0–2.1 Å.

The refined model of the orthorhombic K36E did not include residues 1–9 and 14–23 because no clearly interpretable electron density was observed. This disorder is typical of all Fis protein structures solved to date (Kostrewa *et al.*, 1991, 1992; Yuan *et al.*, 1991, 1994). The refined model of the hexagonal K36E consists of 1412 protein atoms and 22 water molecules. Similarly to the orthorhombic K36E, the first nine residues are not included for each monomer; however, residues 14–23 for both monomers are present as a result of stabilization by
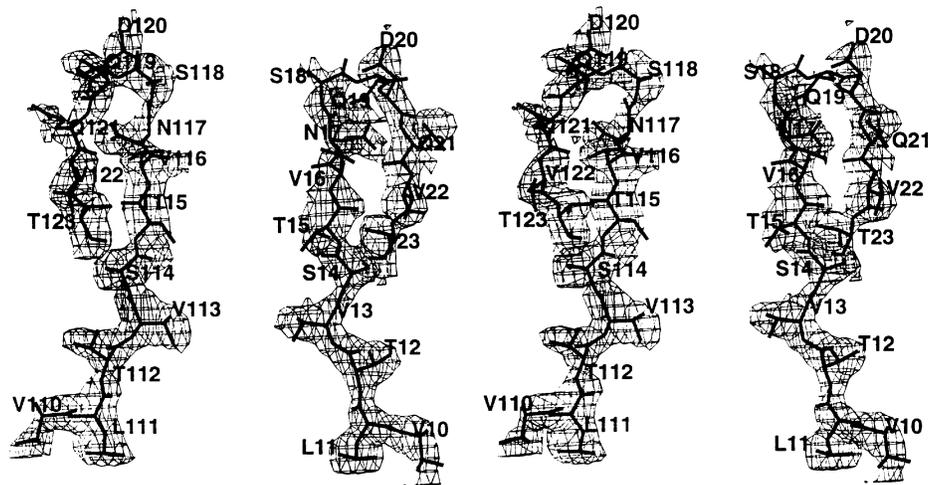
Fig. 3. The omit ($2F_o–F_c$) electron density map for hexagonal K36E at the N-terminal region. The map was calculated by omitting residues 10–24 in both of the Fis homodimer subunits in the hexagonal asymmetric unit. Simulated annealing refinement was done with a 3 Å spherical shell of fixed atoms surrounding the omitted regions, and the maps were contoured at 1.0 σ of the average electron density.

crystal packing interactions. The omitted electron density maps at the N-terminal ends (residues 10–23) for the two monomers are shown in Figure 3. The map contoured at 1.0 σ shows absence of electron density for Gln19 and the side chains of Val16 and Gln21 in monomer a. In monomer b, there is a small break in the density at Asn117 and also at the side chain of Gln121. The ($2F_o–F_c$) map calculated with the complete model contoured at 1.0 σ, however, does not show any such breaks in the main chains for these regions. Overall, the model has acceptable stereochemistry (see Table III). The average *B*-factor is high (42.1 Å^2); however, the ($2F_o–F_c$) map contoured at 1.0 σ is well defined for almost the entire polypeptide chain, including the newly revealed structures at the N-terminal region.

Overall structure of the hexagonal form of K36E. The overall structure of the hexagonal crystal form of K36E is shown in Figure 4A and B, and the secondary structure elements defined by PROCHECK are given in Figure 1B. The structural core consists of four α-helices, A (27–40), B (50–70), C (74–81) and D (85–94), for each subunit in the homodimer structure. This core domain, from residues 27 to 98, is almost structurally identical to that of wild-type Fis. The A and B helices within each subunit interlock and stabilize the folded structure. The C and D helices are the helix–turn–helix DNA-binding domains that are responsible for interaction with DNA at two consecutive major grooves.

The previously unresolved N-terminal structure is nicely revealed in the new hexagonal K36E structure. Two short β-strands located at residues 12–16 (β-1) and 22–26 (β-2) are joined by a hairpin loop (residues 17–21) and form an antiparallel β-sheet which protrudes from the surface of the otherwise compact molecule. The fact that the two β-sheets in the Fis homodimer have no surrounding residues suggests that they must be highly flexible in solution. In the context of the hexagonal crystal, these N-terminal β-hairpin units are stabilized by extensive intermolecular contacts (see section on crystal packing). Because of these contacts, they are well defined and can be resolved in the hexagonal K36E structure, whereas

Table III. X-ray diffraction statistics for the Fis mutant K36E in hexagonal and orthorhombic forms

	Hexagonal K36E	Orthorhombic K36E
Data collection and processing		
Space group	P6$_5$22	P2$_1$2$_1$2$_1$
Cell dimensions (Å)	a = 51.90 Å	78.33
	b = 51.90 Å	51.05
	c = 294.93 Å	48.02
Resolution (Å)	2.65	2.10
Observed reflections	79 532	36 090
Unique reflections	7139	11 107
Completeness—all data (%)	93.0	94.2
Completeness—last shell (%)	55.6	90.4
	(2.75–2.65 Å)	(2.20–2.10 Å)
R_{merge}—all data[a] (%)	6.9	7.7
R_{merge}—last shell	30.1	33.0
Refinement		
Resolution range (Å)	8.0–2.65	8.0–2.10
	($F>1$ σF)	($F>2$ σF)
Reflections	6577	9786
Non-hydrogen atoms		
protein	1416	1270
solvent molecules	22	74
R-factor (%)	22.2	21.2
R_{free} (%)	27.7	28.2
Model quality		
r.m.s. deviations in		
bond lengths (Å)	0.011	0.011
bond angles (°)	1.143	1.217
Average *B*-factor (Å^2)		
non-hydrogen atoms	42.1	34.5
protein atoms	42.1	33.8
solvent atoms	42.6	46.7
Ramachandran plot (%)		
most favored	93.8	91.2
additionally allowed	6.2	8.8
generously allowed	0.0	0.0
disallowed	0.0	0.0

$$^aR_{merge} = \frac{\sum_h \sum_i |Ih,i - \langle Ih \rangle|}{\sum_h \sum_i Ih,i}$$, where $\langle Ih \rangle$ is the mean intensity of the *i* observations for a given reflection *h*.
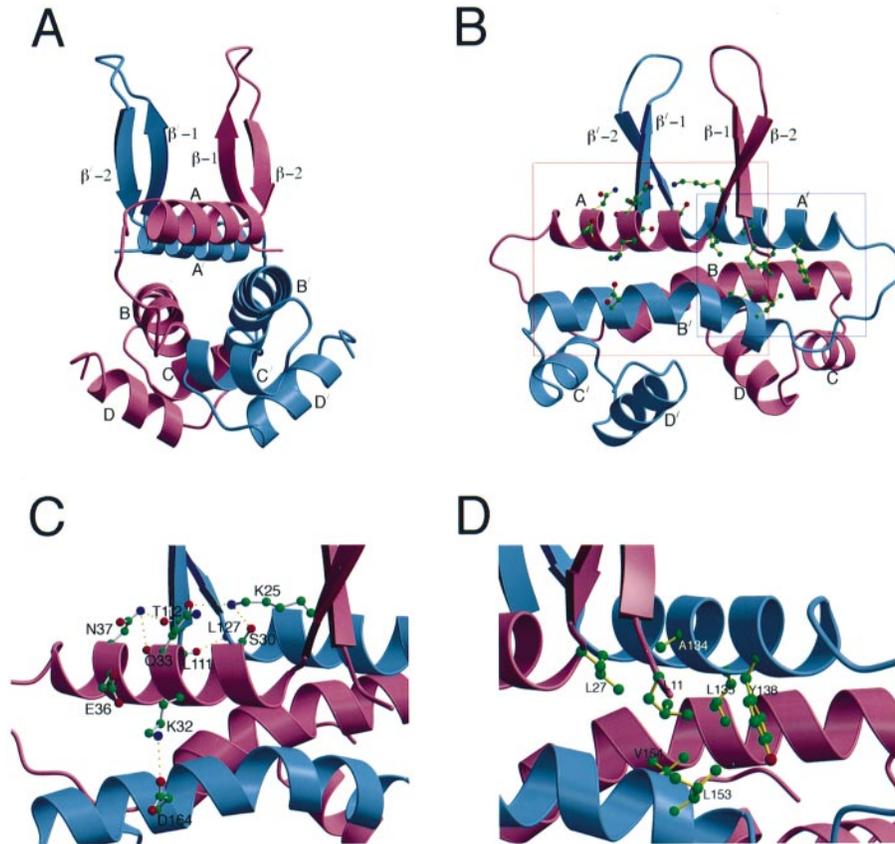
Fig. 4. The structure of the hexagonal Fis K36E homodimer represented in a ribbon diagram with monomer a in pink and monomer b in blue. (**A**) A view of the Fis K36E dimer. The four α-helices and two β-strands are labeled as A–D, β-1 and β-2 in monomer a and A′–D′, β′-1 and β′-2 in monomer b. It can be seen clearly that the regions involved in transactivation of Hin-catalyzed DNA inversion (β and β′ arms) and DNA binding (α-helices C–D and C′–D′) are located above and below the core structure of the A and B helices, respectively. (**B**) Another view of the Fis K36E dimer rotated ~90° apart from (A). The side chains involved in the hydrogen bonding network in monomer a and hydrophobic interactions around Leu11 are framed for enlargement in (C) and (D). (**C**) A close up of the hydrogen bonding network in the A helix that holds the two ends of the β-strands. (**D**) A close up of the hydrophobic interactions around Leu11. These interactions fix the base of the β-1 strand. The figure was prepared using Molscript (Kraulis, 1991) and Raster 3D (Merrit and Murphy, 1994).

residues 14–25 are disordered in the other orthorhombic Fis crystal structures.

The two β-hairpin arms are constrained at the regions where they attach to the core, but should be increasingly mobile the further out they extend. The N-terminus of the A helix is capped by Pro26, and at this position the protein backbone points upward, away from the main body. The side chain of Lys25 (NZ) hydrogen-bonds to Gln33 (OE1) and Ser30 (OG) to fix the conformation at the C-terminal end of the β-2 strand. The reduced Hin inversion activities of mutants K25C and Q33C may, in part, reflect a disruption of this network. On the other side, the backbone of Leu11 (O) hydrogen-bonds to Leu27 (N), and Thr12 (O) hydrogen-bonds to Asn37 (ND2), and these interactions anchor the N-terminal end of the β-1 strand (see Figure 4C). Moreover, the side chain of Leu11 inserts into a hydrophobic core made by several residues from the A and B helices, which are mostly in the partner subunit and include Leu27, Ala134, Leu135, Tyr138, Val154 and Leu153 (see Figure 4D). These hydrophobic interactions lock the base of the β-strands. The near complete absence of activity by L11C and the varying detrimental effects of substitutions and modifications at Ala34 provide strong support for the critical importance of this hydrophobic connection. Moreover, an arginine substitution at (Leu)27

inactivated all measured Fis activities and yielded an unstable protein *in vivo* (data not shown), consistent with the importance of this hydrophobic region to the Fis structure. A computer modeling study also predicted a hairpin structure at the Fis N-terminal region and previously pointed out these hydrophobic interactions surrounding Leu11 (Tzou *et al.*, 1997). There are no observed interactions between the two β-hairpin arms in the dimer, with distances between the backbone atoms of the two closest strands (β-1 and β′-1) ranging from 11.6 to 14.6 Å. The shortest distance of any two atoms between the two β-hairpins is 5.4 Å from Val13 (CG1) to Lys125 (NZ). The severe effect of the maleimide modification at (Val)13 reflects the close approach of this side chain to the other subunit in the dimer.

The two β-strands (β-1 and β-2) on each subunit are connected by a β-hairpin loop between Asn17 and Gln21. Figure 5A shows the detailed hydrogen bonding network within the β-hairpins. The loop region adopts a 3:5 β-hairpin conformation with a classic type I turn where the carbonyl of Asn17 hydrogen-bonds to the amide of Asp20 (Sibanda *et al.*, 1989). The Asn17 side chain is hydrogen bonded to Gln19 and Gln21 and is probably playing an important role in stabilizing the loop structure. This role of Asn17 is corroborated by the properties of
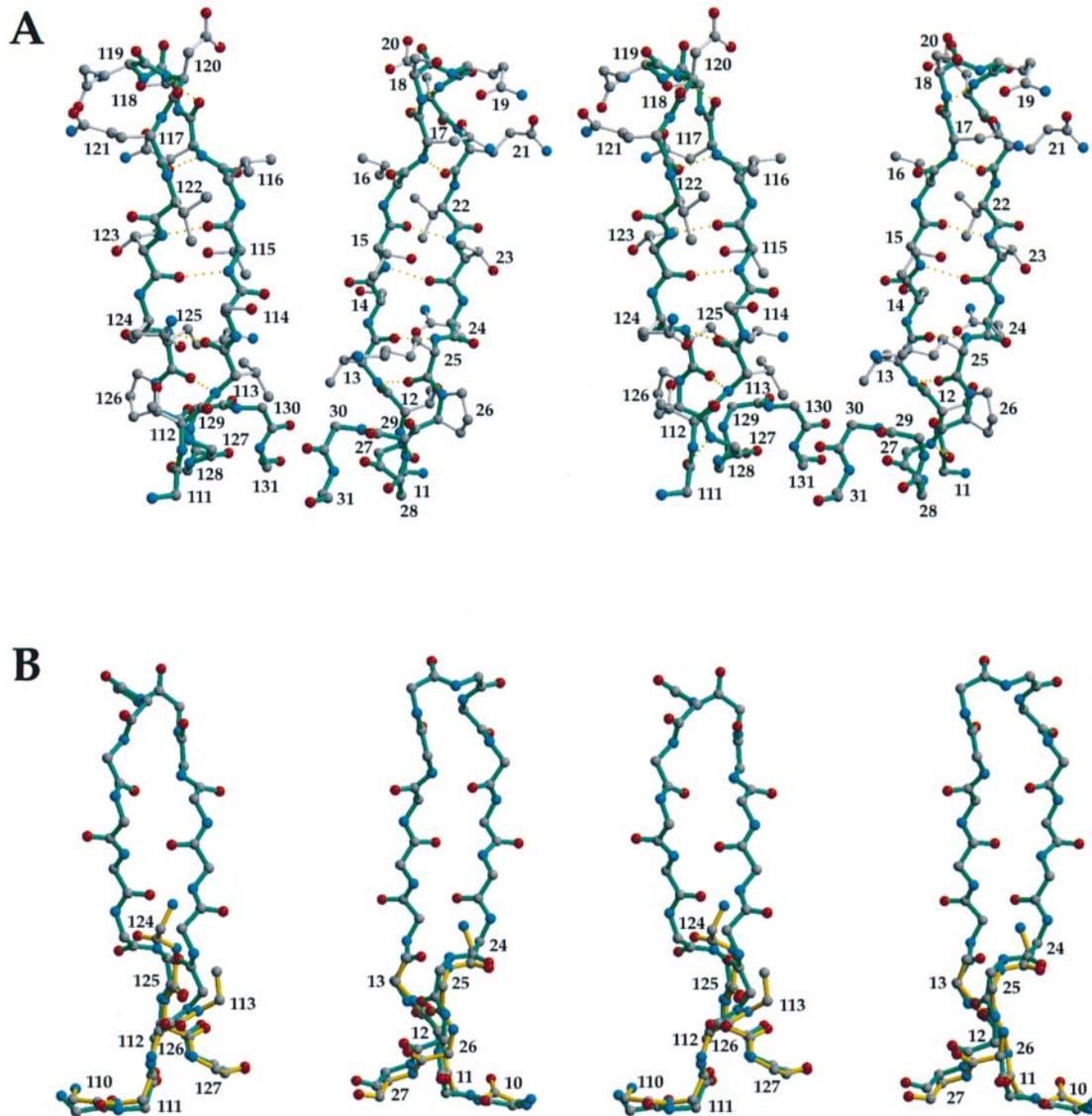
Fig. 5. (**A**) The detailed molecular structure of the extended β-hairpin arms in the Fis dimer. The putative hydrogen bonds between strands β-1 and β-2 in monomer a and b are displayed as dotted lines. For clarity, the main chain atoms are shown for residues 11–33 and 111–133, but only the side chain atoms are included for residues 12–26 and 112–126. (**B**) Cα superimposition of the N-terminal region for the orthorhombic (light yellow) and hexagonal K36E (dark green) structures. The tetrapeptide Val10–Val13 in the orthorhombic K36E structure matches the same position as the corresponding residues in the hexagonal structure. Thus, these four residues are constrained and can be revealed in both structures; however, most of the β-hairpin motifs are disordered in the orthorhombic unit cells due to their mobility. Wild-type Fis is likely to contain a similar β-hairpin structure at its N-terminus since it also crystallizes in an orthorhombic form that displays a tetrapeptide corresponding to Val10–Val13. These diagrams were drawn using Molscript (Kraulis, 1991).

mutations at this position: N17A displays only 8% of wild-type activity, N17C-NEM has essentially no activity, but N17D, which can support the hydrogen bonding, is very active.

The mutagenesis studies identified Asp20 within this loop as being of critical importance to Hin activation. The Asp20 side chain projects into the solvent, and thus is in an excellent position to contact Hin. Substitutions at Val16 and Val22 also result in severely reduced inversion activity. These two side chains, which are in van der Waals contact with each other, are located on the same side of the β-hairpin arm as Asp20 and may provide a hydrophobic patch for recombinase interactions.

β-Hairpin loops have been found to participate in protein–protein interactions in other contexts. A well-studied example is the antigen-binding site in antibodies

that are built from six β-hairpin loops that vary in length and amino acid sequence among different antibodies (Davies and Padlan, 1990). Other examples that use β-hairpin loops as protein–protein interaction sites include hormone and receptor interactions (Kossiakoff *et al.*, 1992; Livnah *et al.*, 1996), protease and inhibitor binding (Stubbs *et al.*, 1990) and proteins containing SH domains (Lee *et al.*, 1996) or PH domains (Ferguson *et al.*, 1994) found in signal transduction pathways. Thus, contacts mediated by flexible β-hairpin loop structures may be ideally suited for assembling higher order nucleoprotein complexes between different DNA-binding proteins.

Crystal packing. There is one Fis homodimer in the asymmetric unit of the hexagonal K36E that produces 12 layers of Fis molecules packed along the *c*-axis in the
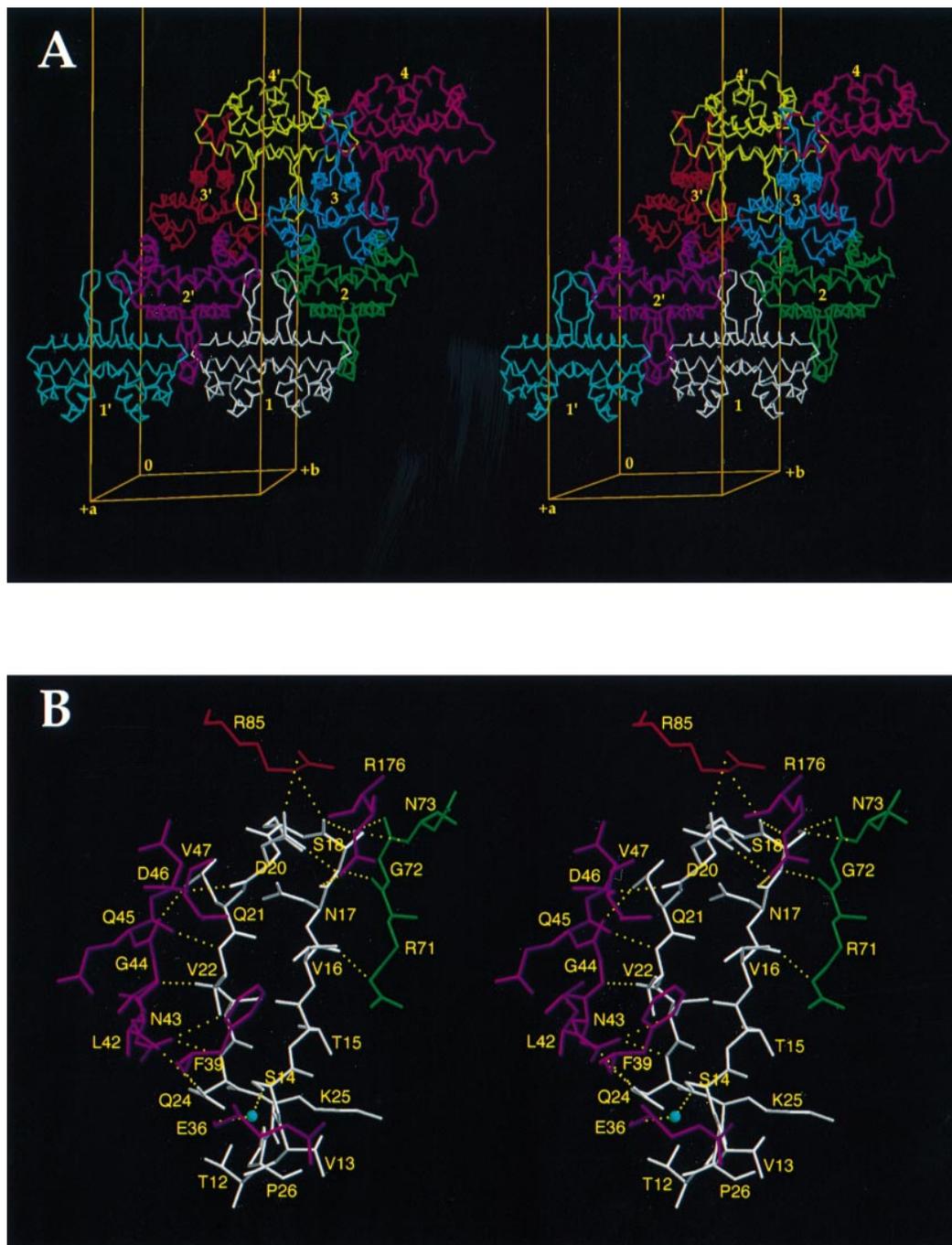
Fig. 6. Stereo view of the molecular packing for the K36E Fis mutant in the hexagonal cell. (**A**) A view along the *b*-axis on the *a,c*-plane. There are 12 layers of Fis molecules packed along the *c*-axis and, for clarity, only four layers of Fis molecules are shown. The conformation of the two β-hairpins is fixed by crystal packing interactions. The symmetry operators for the space group P6$_5$22 relating atom (*x*, y, *z*) to the corresponding one in the reference molecule are (*x*, −1+*y*, *z*) for 1′; (1−*x*, 1+*y*−*x*, 1/3−*z*) for 2; and (1−*x*, *y*−*x*, 1/3−*z*) for 2′. (**B**) Salt bridges and hydrogen bonds surrounding the β-hairpins of the monomer a in the reference molecule 1. The colors of the symmetry-related molecules are the same as those in (A). The small blue sphere between Ser14 (molecule 1) and Glu36 (molecule 2′) depicts a bridged water molecule.

entire unit cell. For clarity, Figure 6A only shows four layers of Fis molecules in the cell. Fis molecules are packed side by side in the same layer, and packed head-to-head (or tail-to-tail) to those in the next layer. The protruding β-hairpins face each other and are inserted between two neighboring Fis molecules in the next layers, similar to the way in which the teeth of two gears interlock. Thus, the presumably flexible structure of the β-hairpin is fixed by crystal packing forces. The height of the core

structure of Fis is ~25 Å, which adds up to 300 Å for 12 layers of Fis molecules packing along the *c*-axis of a dimension 294.9 Å. The width of a Fis molecule is ~50 Å, resulting in a dimension in the *a*- and *b*-axes of 51.9 Å.

The major packing interactions are centered around the two β-hairpins, which make extensive direct protein–protein hydrogen bonds and salt bridges to the surface loops in three crystallographically related molecules. These interactions, which surround the β-hairpin of monomer a,

are illustrated in Figure 6B. The five-residue β-hairpin loop and β-1 strand are also stabilized by a series of hydrogen bonds and salt bridges with three symmetry-related molecules. This extensive hydrogen bond network helps to stabilize the entire β-hairpin and may be relevant to how this β-hairpin with its charged and polar residues could interact with the DNA invertases. The β-hairpin of monomer b is involved in a similar hydrogen bonding network to that of monomer a, with only minor differences. There are a total of 14 direct contacts between the β-hairpin and the neighboring molecules in monomer b, whereas 17 are found in monomer a. The β-hairpin in monomer b is therefore less constrained than the β-hairpin in monomer a, resulting in slightly different conformations of their loop regions. A least-squares fitting between the Cα atoms of the subunits a and b gave an average root mean square (r.m.s.) difference of only 0.18 Å for the 74 residues in the core region where the surface loops of residues 17–21 and 40–49 were excluded for fitting (see details in Materials and methods). However, the average r.m.s. difference in the β-hairpin loop (residues 17–21) is as high as 2.79 Å. These differences demonstrate that the packing environments for the two subunits are not identical, although both subunits show similar β-hairpin structures in their N-terminal regions. Thus, the β-hairpin structure should reflect the intrinsic Fis structure that is not altered by crystal packing forces.

In contrast to the hexagonal packing environment, the orthorhombic K36E crystal has an empty space in the N-terminal region which could allow for mobility of the β-arms. The packing interactions are most extensive in the region of residues 69–77, which includes the loop between the B and C helices. Monomer a of K36E makes 20 direct protein–protein hydrogen bonds and monomer b makes 13 similar interactions with four symmetry-related molecules. In the orthorhombic wild-type Fis structure, the Lys36 residues also make water-bridged contacts to the neighboring Fis molecule. The replacement of Lys36 by Glu may have disturbed these interactions and may have contributed to the crystallization of K36E into the different hexagonal unit cell.

Comparison of K36E with the wild-type Fis structure. A least-squares fit between the Cα atoms (residues 26–98) in wild-type Fis and the orthorhombic K36E mutant gave an average difference of 0.48 Å. This small difference indicates that the overall conformation of the orthorhombic K36E structure is identical to wild-type Fis protein. In the wild-type Fis, there is a hydrogen bond network from Lys36 (NZ) to Gln33 (OE1), and from Gln33 (NE2) to Asn37 (OD1). In the orthorhombic K36E mutant, the amide groups of Gln33 and Asn37 are reversed so that the hydrogen bond network is from Glu36 (OE1) to Gln33 (NE2), and from Gln33 (OE1) to Asn37 (ND2). The rearrangement of the above side chains maintains the interactions between these charged and polar residues; however, Glu36 does not hydrogen-bond with its neighboring residues in the hexagonal K36E structure. Apparently the conformation of residue 36 can change in response to different environments, which may be important for function. Lys36 may be close to Hin in the invertasome since the introduction of the negative charge by K36E results in a large defect in inversion activation. Likewise,

the Lys36 side chain is located in the region of Fis that is responsible for activation of transcription and λ excision (Bokal *et al.*, 1997; S.E.Cramton, S.McLeod and R.C. Johnson, unpublished), thus accounting for the defect in K36E activity in these reactions. Alternatively, the change in charge and hydrogen bonding by K36E may destabilize the protein, which could contribute to the loss of activity.

The average difference in the Cα atoms between the hexagonal K36E (residues 26–98) and wild-type Fis is 1.03, with a comparable difference between the hexagonal and orthorhombic K36E of 0.93. These differences are located mainly in the regions of residues 40–46, 69–77, 84–90 and 97–98 which make contacts with the neighboring molecules either only in the hexagonal structure or in the orthorhombic one. Thus, the larger r.m.s. difference between the hexagonal K36E and the wild-type Fis structures, as compared with the orthorhombic K36E and wild-type, is largely the result of the different crystal lattice interactions found in the two crystal forms, and higher coordinate errors in the hexagonal K36E structure; the core structures are basically identical.

Figure 5B shows the superimposition of the N-terminal regions of the orthorhombic and hexagonal K36E structures after least-square fitting. The tetrapeptide Val10–Val13 in the orthorhombic structure matches well with the corresponding region in the hexagonal structure. The discontinuous tetrapeptides found in the wild-type Fis and P61A are also located at the same position (data not shown). This result demonstrates the flexibility of the β-hairpin arm at the top region between residues 14 and 24, and the stability at the base region between residues 10–13 and 24–26. It also suggests that wild-type Fis should have a similar β-hairpin structure at its N-terminal region; otherwise the position of its residues at the base (10–13 and 24–26) would not correspond with residues in the hexagonal K36E structure.

Oxidation of cysteines within the Fis N-terminus supports a flexible β-hairpin structure

Site-directed cysteine cross-linking was used to probe the structure of the N-terminal activation region under solution conditions similar to those used in inversion reactions. The purified cysteine-containing mutants were subjected to oxidation in the presence of oxidized glutathione for varying periods of time, and the percentage of disulfide-linked dimers was determined by SDS–PAGE (Figure 7A). The results are entirely consistent with the β-hairpin structure of the N-terminal region and provide direct evidence for the flexibility of the arms.

As summarized in Table IV, oxidation rates of the cysteine mutants varied greatly. Cysteines located at positions 15, 16, 17 and 18 formed covalent dimers very rapidly, with 25–50% of the dimers being covalently linked after a 1 min incubation with 10 mM oxidized glutathione. These positions are located within the adjacent β-1/β′-1 strands (residues 15/115 and 16/116) or within the proximal faces (residues 17/117 and 18/118) of the β-hairpin loops in the hexagonal K36E dimer structure (Figure 5A). The separation between Cβ atoms in these pairs of amino acids varies from 10.6 to 16.0 Å, respectively (see Table IV). Disulfide linkages formed at moderate efficiency between cysteines at positions 19/119, whose Cβ atoms are located near the top of the β-hairpin loop
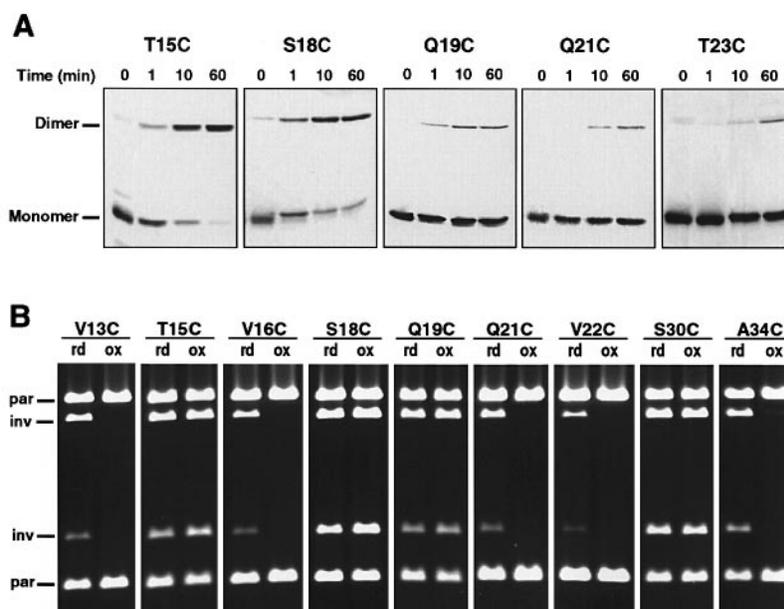
Fig. 7. Oxidation properties of selected Fis cysteine mutants. (**A**) Fis mutants were reduced and then exposed to 10 mM oxidized glutathione for the indicated times. The positions of the bands after SDS–PAGE corresponding to the monomer and disulfide-bridged dimer are noted. (**B**) *In vitro* inversion assays using reduced (rd) and oxidized (ox) forms of Fis mutants. Reaction times for each mutant are as follows: V13C, 2 min; T15C, 1 min; V16C, 2 min; S18C, 1 min; Q19C, 1 min; Q21C, 1 min; V22C, 5 min; S30C, 1 min; A34C, 2 min. The positions of the restriction fragments produced from molecules in the parental (par) and inverted (inv) orientation after agarose gel electrophoresis are marked.

Table IV. Rates of formation and inversion proficiency of disulfide-linked Fis cysteine mutants

Mutant	Location	%/1 min	%/60 min	Cα–Cα[a]	Cβ–Cβ[b]	% activity[c]
L11C	β-1 extension	12	34	18.0	16.5	not active
V13C	β-1	3	35	12.5	9.5	<1
S14C	β-1	3	17	15.5	18.1	18
T15C	β-1	27	85	11.6	10.6	110
V16C	β-1	46	84	13.1	13.3	<2
N17C	loop	50	86	13.5	16.0	18
S18C	loop	25	77	13.3	13.8	112
Q19C	loop	15	26	19.9	22.5	102
Q21C	loop	1	24	22.5	24.3	<1
V22C	β-2	5	22	21.4	22.9	<10
T23C	β-2	3	8	21.1	20.9	21
Q24C	β-2	<1	<1	19.8	22.1	not tested
K25C	β-2	<1	5	15.3	12.8	58
D29C	helix A	<1	<1	13.6	14.4	not tested
S30C	helix A	1	51	7.2	4.3	80
K32C	helix A	<1	<1	16.6	18.2	not tested
Q33C	helix A	<1	<1	16.4	14.5	not tested
A34C	helix A	<1	<1	14.8	12.4	6
K36E	helix A	<1	<1	25.0	25.1	not tested
N37C	helix A	<1	<1	24.8	22.5	not tested
Q41C	helix A	<1	<1	35.9	34.0	not tested

[a]Distances are given between the two Cα atoms of the same residue from each subunit in the Fis homodimer.
[b]Distances are given between the two Cβ atoms of the same residue from each subunit in the Fis homodimer.
[c]Rates of Hin-catalyzed DNA inversion promoted by disulfide-linked versus reduced forms of cysteine-containing Fis mutants. For each mutant, the inversion rate of the reduced form is set at 100%. In each case given, >90% of the dimers were covalently linked as assayed by non-reducing SDS–PAGE, except for V22C and A34C which were 70 and 55% disulfide-linked, respectively. L11C was not active under reducing or oxidizing conditions. No activity was obtained with oxidized V22C; however, because of its low activity under reducing conditions, only ~10% of the reduced activity could have been detected. Those mutants not tested formed insufficient amounts of disulfide-linked dimers to measure their activities.

and are separated by 22.5 Å. The ability of these cysteine pairs to form disulfide bonds readily is consistent with the β-hairpin structures in the hexagonal K36E crystal, provided that the arms can move freely in solution. An exception may be the efficient cross-linking between residues 17 and 117, since the Asn17/117 side chains are oriented away from each other. However, the cysteine side chains in the N17C mutant are probably no longer hydrogen bonded to Gln19/119 and Gln21/121 within the loop region and thus could easily reorient themselves. Cysteines at residues 13/113 and 14/114 near the base of the β-1/β′-1 arm region oxidize less efficiently since their movement would be more constrained. Cysteines along the outside β-2/β′-2 strands formed disulfide bonds at

much lower rates (21/121, 22/122, 23/123 and 25/125) or not at all (24/124). L11C, one of the most severely defective mutants, formed disulfide linkages relatively efficiently, but this probably reflects a major disruption of the arm structure due to an absence of the hydrophobic linkage to the core.

The only cysteine within the A helix to form disulfide linkages within the time frame measured in these experiments was at position 30. The Cβ atoms of the native Ser30/130 residues are positioned 4.3 Å across from each other. However, the initial rate of disulfide linkage of S30C in the presence of 10 mM oxidized glutathione was only 2–4% of the rate measured for cysteines located within the β-hairpin arms that oxidized efficiently. This lower rate presumably reflects the more rigidly fixed positions of the A helices within the body of the protein as compared with the β-arms. Moreover, once oxidized, S30C covalent dimers were much more difficult to reduce than those of cysteine mutants located in the β-arm region (data not shown). Cysteine mutants at positions 29 and 34 had extremely low rates of disulfide bond formation that were observed upon long-term storage. Linkage at these positions would cause large shifts in the interlocking A helices within the dimer.

Activity of disulfide-linked Fis mutants

The ability of those cysteine-containing Fis mutants that formed disulfide linkages to activate Hin-catalyzed inversion was determined. Examples of inversion assays using reduced and oxidized preparations of the Fis mutants are shown in Figure 7B, and the activities of the oxidized relative to the reduced preparations are quantitated in Table IV. As expected, disulfide-linked S30C, located across from each other in the A helix, was very active, but linkage at A34C, which would cause a major shift in the alignment of the A helices, inactivated the protein. The small amount of activity observed in the oxidized A34C reactions probably reflects incomplete oxidation. Within the β-hairpin region, oxidized forms of T15C, S18C and Q19C were completely active. The efficient activities of these disulfide-linked mutants are notable, particularly Q19C which immediately precedes the critical Hin-activating residue Asp20. On the other hand, disulfide linkage between residues 21 and 121 or 22 and 122, which would require severe twisting of the loop region, resulted in a protein that was unable to activate inversion. Likewise, linkage between positions 17 and 117, which would distort the loop region as described above, caused an even stronger inversion defect than the reduced form of N17C. S14C within β-1, and T23C and K25C within β-2, are able to form disulfide linkages at very low rates, but the linked forms are partially active. In these cases, the β-arms would be in a criss-crossed configuration. Disulfide linkage between residues 13 and 113 that are located in the N-terminal ends of the β-arms abolishes inversion activity, presumably because it would disrupt the non-covalent association of this end of the arms with the core. Oxidized preparations of V13C, S14C, V16C, N17C, Q21C, V22C and T23C bound DNA normally (data not shown), indicating that they were specifically defective in activating inversion.

The full activity of oxidized Q19C in which the β-arms are covalently linked at the amino acid preceding Asp20

is particularly significant with regard to probable Hin contacts by the β-arm region. This result implies either that the two β-arms are very closely associated in the functionally active form of Fis, or that a single arm is sufficient for full activation of Hin inversion. Preliminary experiments, in which the activity of heterodimers containing one active and one non-active β-arm was determined, are consistent with only a single functional β-arm being required (R.C.Johnson, unpublished). In either case, a model in which the two arms are interacting independently with a common site on each of the two subunits of a Hin dimer is unlikely.

Summary of the DNA invertase activation region of Fis

The crystal structure of the K36E Fis mutant revealed that the N-terminal region between amino acids 10 and 26 of each subunit consists of a mobile β-hairpin arm. Saturation mutagenesis demonstrated that the β-hairpin region together with several amino acids within the A α-helix constitutes the region responsible for regulation of site-specific DNA inversion. The most critical amino acid for Hin activation is an aspartic acid (Asp20) located near the tip of the β-hairpin loop. This amino acid presumably contacts the DNA invertase in an invertasome structure to induce a conformational change that potentiates its catalytic activity. Two other amino acids within the β-arms that are likely to be involved in direct invertase interactions are the valines at positions 16 and 22. The side chains of these residues are oriented in the same direction as Asp20 and thus may be involved in additional hydrophobic interactions. The effects of certain mutations at Asn17 are probably due to the important role of this amino acid in stabilizing the configuration of the β-hairpin loop. Since the two arms can be covalently linked together in a functional conformation, it is likely that only one arm within a dimer is required for activation of the invertase or that both arms associate together to form the active conformer. In this regard, the side chains of the Val16, Asp20, Val22 cluster in each arm are facing in opposite directions both in the K36E crystal structure and when the arms are covalently linked at positions 19/119.

Two solvent-accessible side chains within the A α-helix, Gln33 and Lys36, may also play a role in Hin interactions since mutations at these positions can have a strong effect on DNA inversion. The function of Gln33 may be, in part, to stabilize the base of the β-arm via a hydrogen bond with Lys125 on the other subunit, but K25C has only a relatively small effect on inversion rates. Lys36 is also involved in hydrogen bonding within the Fis dimer; however, K36C, which would also disrupt hydrogen bonding at residue 36, is only slightly defective in Hin inversion. The low activity of K36E may reflect primarily the detrimental effect of introducing a negatively charged side chain at this position.

The structure and mutation data also point to the critical importance of the connection between the N-terminal end of the β-arm, particularly Leu11, and the hydrophobic region of the Fis core surrounding Ala34. Mutations at these two residues disrupt the hydrophobic interactions that are essential for the β-arm fold. The main chain connection between the C-terminal end of the β-arm and the A α-helix at Pro26 is also sensitive to amino acid

changes, and such changes can completely alter the structure of the β-arm region (H.S.Yuan, unpublished).

A pair of extended β-hairpin arms that protrude from a compact core structure have been reported before in the HU (Tanaka *et al.*, 1984; Vis *et al.*, 1995) and IHF (Rice *et al.*, 1996) proteins. In these proteins, both arms are required for DNA binding. In the case of the HU structures determined in the absence of DNA, the eight residues at the ends of the β-hairpins were found to be flexible by NMR and not visible by X-ray crystallography. We emphasize that even though the β-arm motif in Fis is revealed by the crystal structure of a mutant protein, it is likely to be the functionally relevant conformer, given that the arms are mobile in solution. While the positions of the β-arms are determined by packing forces within the K36E crystal, these interactions are not identical between the two arms. Moreover, the structure is consistent with all the mutational and cysteine cross-linking data. Similar flexible β-arm motifs that function in protein–protein communication may be present in transactivation domains of other regulatory molecules whose three-dimensional structures appear disordered.

Materials and methods

Isolation and purification of Fis mutants

Some Fis mutants were isolated by error-prone PCR (Leung *et al.*, 1989) followed by cloning the resulting DNA fragment containing the N-terminal segment of Fis into pRJ1177, a derivative of pBR322 containing the *fis* gene with a set of unique restriction sites under the control of *tacP* (Pan *et al.*, 1996). Mutants displaying a reduced rate of Hin inversion were identified after transformation using the *in vivo* screen described in Osuna *et al.* (1991). Site-directed mutants derived from mutant oligonucleotides were obtained using the two-step PCR protocol of Landt *et al.* (1990). Alternatively, PCR was performed using a mutant oligonucleotide that overlapped a restriction site, and the resulting product was cloned directly into pRJ1177.

Fis mutant proteins were purified from cells containing derivatives of pRJ1177 or pRJ1077, in which *fis* was expressed from the T7 promoter as described (Pan *et al.*, 1996). For crystallography, large-scale preparations were obtained by chromatography on S-Sepharose followed by dialysis against a low salt buffer to precipitate Fis. Resolubilized Fis was stored at a concentration of 20 mg/ml in 20 mM Tris–HCl (pH 8.2), 1 M NaCl at –70°C.

Oxidation and chemical modification of Fis proteins

The ability of Fis dimers containing cysteines to form disulfide-linked dimers was determined as follows. Purified preparations were reduced by incubating for 5 min at 37°C with 5 mM dithiothreitol (DTT) followed by dialysis against 20 mM HEPES (pH 7.5), 0.5 M NaCl and 10% glycerol for 1–3 h at 4°C. The reduced preparations were then incubated with 10 mM oxidized glutathione at a concentration of 2 μM Fis in 20 μl of 20 mM HEPES (pH 7.5), 0.1 M NaCl, 10% glycerol for various times, ranging from 1 min to >1 h, at which point SDS sample buffer without reducing agent was added and the samples were frozen at –70°C. The amount of covalently linked dimer was determined by densitometry (Molecular Dynamics) after electrophoresis in an SDS–polyacrylamide gel and staining with Coomassie blue. To ensure that the disulfide linkages were not between Fis dimers, oxidized preparations were cross-linked with the homobifunctional reagent disuccinimidyl suberate under conditions that efficiently cross-linked the wild-type Fis dimer and then subjected to non-reducing SDS–PAGE. Only covalent dimers were obtained, indicating that disulfide bridges between dimers, which would result in higher order forms, were not formed. Oxidation reactions were also performed at a concentration of 50 nM Fis in the presence of 100 nM of the distal *hin* enhancer-binding site on a 33 bp annealed oligonucleotide. The amount of covalently linked dimer was determined by immunoblotting after SDS–PAGE. Qualitatively similar results were obtained, though the rates were somewhat faster than measured at the higher Fis concentrations without DNA (data not shown).

Cysteine mutants were chemically modified after reduction and

removal of the DTT as described above. Alkylation by NEM (Pierce) was performed by incubating 2 μM reduced Fis with 5 mM NEM for 5 min at 0°C in 20 mM HEPES (pH 7.0), 0.1 M NaCl and 10% glycerol. Alkylation with AE-8 (Pierce) was performed similarly except that the reactions were incubated overnight at 0°C in 20 mM HEPES (pH 8.2), 0.1 M NaCl and 10% glycerol. The reactions were quenched with the addition of an equal volume of 50% glycerol, 20 mM Tris–HCl (7.5), 0.1 NaCl, 20 mM DTT and 400 μg/ml bovine serum albumin (BSA). Similar treatments with wild-type Fis resulted in no significant loss of activity.

Hin inversion, phage λ excision, proP transcription and DNA-binding assays

For qualitative *in vivo* measurements of Fis activity, pRJ1177-derived plasmids were transformed into RJ2539 [λ*fla406*OFF, F′*proAB lac-I*q*Z*$_{U118}$, *fis::kan-767* Δ*(pro-lac) recA56 srl ara rpsL* pKH66 (*tacP-hin*, pSC101 *ori*, *spc*)] and plated on lactose MacConkey agar plates. The rate of red colony development due to inversion of the *H2* promoter into the orientation that transcribes *lacZ* was monitored relative to the wild-type *fis* and the no *fis* control. *In vitro* Hin-catalyzed DNA inversion reactions were performed and quantitated as described in Haykinson *et al.* (1996). Approximately 1 pmol of Fis was typically added to inversion reactions containing 0.1 pmol of substrate DNA, although higher amounts of those mutants or chemically modified preparations displaying lower activity were also tested to ensure that saturating amounts of Fis were present. Inversion reactions were typically incubated for 30 s, or longer for weakly active mutants, prior to quenching to enable measurements of initial rates as reported in Table I. Reactions employing Fis mutants containing cysteines were pre-incubated with 5 mM DTT. Oxidized preparations of Fis mutants for inversion assays were obtained by long term storage in ≤1 mM DTT, or in some cases by prolonged incubation in 10 mM oxidized glutathione, and 2 mM oxidized glutathione was included in the reaction. Immunoblotting with anti-Fis antibody confirmed that the mutants remained oxidized at the completion of the inversion reaction.

Phage λ excision assays were performed as described in Osuna *et al.* (1991). Fis activation of the *proP* P2 promoter was measured in transformants of RJ3147 [*fis::kan-767* Δ*lacX74* pRJ823 (pACYC184-lacIq) containing a λ prophage with a *proP–lacZ* fusion that has the *fis*-independent P1 promoter inactivated by a mutation at –12]. *In vivo* expression of this promoter is essentially dependent upon Fis (Xu and Johnson, 1995). Duplicate overnight cultures grown in LB were subcultured 1/100 into fresh LB, and β-galactosidase assays (Miller, 1972) were performed after 4 h of growth.

Fis binding to DNA *in vitro* was measured by polyacrylamide gel mobility shift assays. The electrophoretic mobility of the Fis–*hin-D* site complex was measured when the site was located near the middle or end of a 443 bp DNA fragment obtained from pRJ1213 to assess differences in Fis-induced DNA bending (Wu and Crothers, 1984; Pan *et al.*, 1996). No differences in apparent DNA bending by the mutants were observed, although a few mutants displayed small variations in the migration of the Fis–DNA complexes when bound to both DNA fragments.

Crystallization, X-ray data collection and processing

Crystals of the orthorhombic K36E were grown by the hanging drop vapor diffusion method from a solution of 20 mg/ml protein, 500 mM NaCl, 100 mM sodium acetate, 100 mM Tris–HCl, pH 8.5, and 20% w/v PEG4000, against 40% w/v PEG4000 in the reservoir. Crystals of the hexagonal form of K36E were also grown by the hanging drop vapor diffusion method from a solution of 20 mg/ml protein, 2 M sodium formate, against 4 M sodium formate in the reservoir. Both conditions gave X-ray diffraction quality crystals after 2 weeks. The orthorhombic K36E crystallized in the same space group P2$_1$2$_1$2$_1$ as the wild-type Fis protein, and the hexagonal K36E crystallized in the space group P6$_5$22.

X-ray diffraction intensities for all the crystals were collected at room temperature from an *R*-axis II imaging plate system equipped with a 300 kW rotating anode and double-focusing mirrors. Orthorhombic K36E diffracted to 2.1 Å with an R_{sym} of 7.7% based on the intensities between symmetry-related reflections, while hexagonal K36E crystals diffracted to 2.65 Å with an R_{sym} of 6.9%. All data sets were integrated, scaled and merged using the RAXIS-II data processing software. Table III summarizes the diffraction statistics.

Structure determination

The crystal structure of the hexagonal crystal form of the mutant K36E was solved by the molecular replacement method with the program

AMoRe (Navaza, 1994). A partially refined orthorhombic crystal form of the K36E structure (R-factor/R_{free} = 24.8/32.8%) was used as the searching model. This model included residues 24–98 of the two monomers and their corresponding tetrapeptide chains (residues 10–13). The tetrapeptide residues were changed to alanines, and temperature factor values of all protein atoms were left unaltered. Based on the solvent content of the unit cell (~53%), we expected one dimer per asymmetric unit. The cross rotation function was calculated using normalized structure factors with data from 10.0 to 3.0 Å. The top 10 rotation solutions were used for the translation function (10.0–3.0 Å), for the two possible space groups P6$_5$22 and P6$_1$22. The space group P6$_5$22 resulted in two distinct peaks for the top two cross rotation solutions, with correlation coefficients of 42.1 and 45.7 and R-factors of 53.1 and 51.7% respectively. These two solutions are related by a 2-fold axis (symmetry element 1–x, y–x, 1/3–z) and are therefore identical. AMoRe rigid body refinement resulted in a correlation coefficient of 50.4 and an R-factor of 49.1% for the highest peak. The next highest peak has a correlation coefficient of 30.3 and an R-factor of 56.9% that is substantially worse.

Structure refinement

All protein structures were refined with the XPLOR program (Brunger, 1992b). All non-protein atoms were deleted from the starting model coordinate files, and all B-factors were set to an arbitrary value of 20 before refinement. In all cases, a statistically random selection of 8% of the total reflection data was excluded from the refinement and used to calculate the free R-factor (R_{free}) as a monitor of model bias (Brunger, 1992a). Data with $F \geq 2.0$ σF for orthorhombic and $F \geq 1.0$ σF for hexagonal K36E were used in the refinements. Water molecules were included in the model if they had a peak above 3.0 σ (3.5 σ for hexagonal K36E) in the (F_o–F_c) and 1.0 σ in the ($2F_o$–F_c) maps, and showed a reasonable hydrogen bonding geometry. All model buildings and graphic operations were done with O (Jones, 1991) and TOM (Cambillau and Horjales, 1987). The models were also subjected to quality analyses in most stages of the refinement, using the programs O, XPLOR and PROCHECK. Refinement statistics are summarized in Table III.

The previously determined isomorphous Fis mutant protein structure, P61A (Yuan *et al.*, 1994), was used as a starting model for the orthorhombic K36E refinement. Only the model comprising the continuous polypeptide chains of 26–98 amino acid residues of the two monomers were used, and the Lys36 residue was changed to alanine before refinement. The model was subjected to a rigid body refinement with the two monomers as two independent groups to a crystallographic R-factor of 33.4% and an R_{free} of 33.3% using data between 8.0 and 3.0 Å. An NCS restraint (weight of 200) was applied to all non-hydrogen atoms, and a series of positional refinements to a resolution of 2.4 Å was undertaken. Two residues, Gln24 and Lys25, were then added to each of the two monomers, and Ala36 and Ala61 were changed to Glu36 and Pro61 respectively. In addition, two stretches of densities that were separated but adjacent to the N-termini of the two monomers were fitted with Val10–Leu11–Thr12–Val13, with the characteristic Y-shaped side chain at Leu11. A series of positional refinements and simulated annealings with gradual increments of the resolution resulted in an R-factor of 30.1% and an R_{free} of 37.0% with data from 8.0 to 2.1 Å. The NCS restraint weight, which at this stage was 50, was removed, followed by another round of simulated annealing. Addition of 74 waters (B-factors <65 Å^2) and B-factor refinement resulted in a final R-factor of 21.2% and an R_{free} of 28.2%.

For the refinement of the hexagonal K36E structure, the starting model comprised the continuous polypeptide chain of 24–98 amino acid residues of the two monomers and the separated polyalanine tetrapeptide chains (residues 10–13 and 110–113). The model was subjected to a rigid body refinement, with the two continuous chain monomers and the two polyalanine tetrapeptides as four independent groups. Data between 10.0 and 3.0 Å were used to yield an R-factor of 49.9% and an R_{free} of 49.4%. At this stage, a tight NCS restraint weight of 300 was applied to all non-hydrogen atoms and the model subjected to positional refinement with data from 10.0 to 3.0 Å. The R-factor and R_{free} dropped to 40.4 and 48.0% respectively. The ($2F_o$–F_c) density map was well defined along most of the entire input model. In addition, electron densities were observed between Gln24 of the main chain and Ala13 of the separated polyalanine tetrapeptide of both monomers. At this stage, the polyalanine tetrapeptides were assigned to their corresponding correct sequences, and residues Val13–Gln24 were assigned to the new density. The model was then subjected to several rounds of simulated annealing and positional refinements to a resolution of 2.65 Å, and the best model

was obtained with an NCS restraint weight of 100. At this stage, ($2F_o$–F_c) and (F_o–F_c) omit maps revealed that the regions between residues 17–21 and 40–49 were poorly defined, with large blobs of positive and negative densities, indicating that these areas do not obey the NCS symmetry imposed on the model. During all subsequent refinements, residues 17–21 and 40–49 were not restrained and this improved the R_{free} and the densities in these areas. The final restraint weight of 30 resulted in the best model, as indicated by both the R_{free} and the electron density maps. Addition of 22 waters (with B-factors >55 Å^2) and B-factor refinement resulted in a final R-factor of 22.2% and an R-free of 27.7%. The final coordinates of the hexagonal K36E model have been deposited in the Brookhaven Protein Data Bank with the access code of 1F36.

Acknowledgements

References

Ball,C.A. and Johnson,R.C. (1991) Efficient excision of phage lambda from the *Escherichia coli* chromosome requires the Fis protein. *J. Bacteriol.*, **173**, 4027–4031.

Bokal,A.J., Ross,W., Gaal,T., Johnson,R.C. and Gourse,R.L. (1997) Molecular anatomy of a transcriptional activation patch: Fis–RNA polymerase interactions at the *Escherichia coli rrn*B P1 promoter. *EMBO J.*, **16**, 154–162.

Brunger,A.T. (1992a) Free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.

Brunger,A.T. (1992b) *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Cambillau,C. and Horjales, E. (1987) TOM: a Frodo subpackage for protein–ligand fitting with interactive energy minimization. *J. Mol. Graph.*, **5**, 174–177.

Cowtan,K. (1994) DM: an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, **31**, 34–38.

Crisona,N.J., Kanaar, R., Gonzalez,T.N., Zechiedrich,E.L., Klippel,A. and Cozzarelli,N.R. (1994) Processive recombination by wild-type Gin and an enhancer-independent mutant. Insight into the mechanisms of recombination selectivity and strand exchange. *J. Mol. Biol.*, **243**, 437–457.

Davies,D.R. and Padlan,E.A. (1990) Antibody–antigen complexes. *Annu. Rev. Biochem.*, **59**, 439–473.

Ferguson,K.M., Lemmon,M.A., Schlessinger,J. and Sigler,P.B. (1994) Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human dynamin. *Cell*, **79**, 199–209.

Glasgow,A.C., Hughes,K.T. and Simon,M.I. (1989) Bacterial DNA inversion systems. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 637–659.

Gosink,K.K., Ross,W., Leirmo,S., Osuna,R., Finkel,S.E., Johnson,R.C. and Gourse,R.L. (1993) DNA binding and bending are necessary but not sufficient for Fis-dependent activation of *rrnB* P1. *J. Bacteriol.*, **175**, 1580–1589.

Haffter,P. and Bickle,T.A. (1988) Enhancer-independent mutants of the Cin recombinase have a relaxed topological specificity. *EMBO J.*, **7**, 3991–3996.

Haykinson,M.J., Johnson,L.M., Soong,J. and Johnson,R.C. (1996) The Hin dimer interface is critical for Fis-mediated activation of the catalytic steps of site-specific DNA inversion. *Curr. Biol.*, **6**, 163–177.

Heichman,K.A. and Johnson,R.C. (1990) The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. *Science*, **249**, 511–517.

Heichman,K.A., Moskowitz,I.P. and Johnson,R.C. (1991) Configuration of DNA strands and mechanism of strand exchange in the Hin invertasome as revealed by analysis of recombinant knots. *Genes Dev.*, **5**, 1622–1634.

Hughes,R.E., Rice,P.A., Steitz,T.A. and Grindley,N.D. (1993) Protein–protein interactions directing resolvase site-specific recombination: a structure–function analysis. *EMBO J.*, **12**, 1447–1458.

Johnson,R.C. and Bruist,M.F. (1989) Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombinational enhancer in the strand exchange reaction. *EMBO J.*, **8**, 1581–1590.

Johnson,R.C., Bruist,M.F. and Simon,M.I. (1986) Host protein requirements for *in vitro* site-specific DNA inversion. *Cell*, **46**, 531–539.

Jones,T.A. (1991) A graphic model building and refinement system for macromolecules. *J. Appl. Crystallogr.*, **11**, 268–272.

Kanaar,R., van de Putte,P. and Cozzarelli,N.R. (1988) Gin-mediated DNA inversion: product structure and the mechanism of strand exchange. *Proc. Natl Acad. Sci. USA*, **85**, 752–756.

Klippel,A., Cloppenborg,K. and Kahmann,R. (1988) Isolation and characterization of unusual *gin* mutants. *EMBO J.*, **7**, 3983–3989.

Klippel,A., Kanaar,R., Kahmann,R. and Cozzarelli,N.R. (1993) Analysis of strand exchange and DNA binding of enhancer-independent Gin recombinase mutants. *EMBO J.*, **12**, 1047–1057.

Koch,C. and Kahmann,R. (1986) Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. *J. Biol. Chem.*, **261**, 15673–15678.

Koch,C., Mertens,G., Rudt,F., Kahmann,R., Kanaar,R., Plasterk,R.H., van de Putte,P., Sandulache,R. and Kamp,D. (1987) The invertible G segment. In Symonds,N., Toussaint,A., van de Putte,P. and Howe,M.M. (eds), *Phage Mu*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 75–91.

Koch,C., Ninnemann,O., Fuss,H. and Kahmann,R. (1991) The N-terminal part of the *E.coli* DNA binding protein FIS is essential for stimulating site-specific DNA inversion but is not required for specific DNA binding. *Nucleic Acids Res.*, **19**, 5915–5922.

Kossiakoff,A.A., Ultsch,M. and de Vos,A.M. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306–312.

Kostrewa,D., Granzin,J., Koch,C., Choe,H.-W., Raghunathan,S., Wolf,W., Kahmann,R. and Saenger,W. (1991) Three-dimensional structure of the *E.coli* DNA-binding protein Fis. *Nature*, **349**, 178–180.

Kostrewa,D., Granzin,J., Stock,D., Choe,H.W., Labahn,J. and Saenger,W. (1992) Crystal structure of the factor for inversion stimulation FIS at 2.0 Å resolution. *J. Mol. Biol.*, **226**, 209–226.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Landt,O., Grunert,H.P. and Hahn,U. (1990) A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, **26**, 125–128.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Lee,C.-H., Saksela,K., Mirza,U.A., Chait,B.T. and Kyriyan,J. (1996) Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell*, **85**, 931–942.

Leung,D., Chen,E. and Goeddal,D. (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique*, **1**, 11–15.

Livnah,O., Stura,E.A., Johnson,D.L., Middleton,S.A., Mulcahy,L.S., Wrighton,N.C., Dower,W.J., Jolliffe,L.K. and Wilson,I.A. (1996) Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 Å. *Science*, **273**, 464–471.

Merrit,E.A. and Murphy,M.E.P. (1994) Raster 3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.

Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Osuna,R., Finkel,S.E. and Johnson,R.C. (1991) Identification of two functional regions in Fis: the N-terminus is required to promote Hin-mediated DNA inversion but not lambda excision. *EMBO J.*, **10**, 1593–1603.

Pan,C.Q., Finkel,S.E., Cramton,S.E., Feng,J.A., Sigman,D.S. and Johnson,R.C. (1996) Variable structures of Fis–DNA complexes determined by flanking DNA–protein contacts. *J. Mol. Biol.*, **264**, 675–695.

Rice,P.A., Yang,S.-W., Mizuuchi,K. and Nash,H.A. (1996) Crystal structure of an IHF–DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295–1306.

Sibanda,B.L., Blundell,T.L. and Thornton,J.M. (1989) Conformation of β-hairpins in protein structures—a systematic classification with application to modelling by homology, electron density fitting and protein engineering. *J. Mol. Biol.*, **206**, 759–777.

Stubbs,M.T., Laber,B., Bode,W., Huber,R., Jerala,R., Lenarcic,B. and Turk,V. (1990) The refined 2.4 Angstrom X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.*, **9**, 1939–1947.

Tanaka,I., Appelt,K., Dijk,J., White,S.W. and Wilson,K.S. (1984) 3-Å resolution structure of a protein with histone-like properties in prokaryotes. *Nature*, **310**, 376–381.

Tzou,W.-S. and Hwang,M.-J. (1997) A model for Fis N-terminus and Fis–invertase recognition. *FEBS Lett.*, **401**, 1–5.

Vis,H., Mariani,M., Vorgias,C.E., Wilson,K.S., Kaptein,R. and Boelens,R. (1995) Solution structure of the HU protein from *Bacillus stearothermophilus*. *J. Mol. Biol.*, **254**, 692–703.

Wu,H.M. and Crothers,D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.

Xu,J. and Johnson,R.C. (1995) Fis activates the RpoS-dependent stationary-phase expression of *proP* in *Escherichia coli*. *J. Bacteriol.*, **177**, 5222–5231.

Yuan,H.S., Finkel,S.E., Feng,J.A., Kaczor-Grezskowiak,M., Johnson,R.C. and Dickerson,R.E. (1991) The molecular structure of wild-type and a mutant Fis: relationship between mutational changes and recombinational enhancer function or DNA bending. *Proc. Natl Acad. Sci. USA*, **88**, 9558–9562.

Yuan,H.S., Wang,S.S., Yang,W.-Z., Finkel,S.E. and Johnson,R.C. (1994) The structure of Fis mutant Pro61–Ala illustrates that the kink within the long alpha-helix is not due to the presence of the proline residue. *J. Biol. Chem.*, **269**, 28947–28954.