COMMUNICATION

Effects of Base Substitutions on the Binding of a DNA-bending Protein

Christian Bailly¹, Michael J. Waring¹ and Andrew A. Travers²*

¹Department of Pharmacology
University of Cambridge
Tennis Court Road
Cambridge, CB2 1QJ, UK

²MRC Laboratory of
Molecular Biology, Hills Road
Cambridge, CB2 2QH, UK

In order to investigate whether the 2-amino group of guanine, which lies in the minor groove of the B-form helix, can directly influence DNA flexibility and major groove recognition by proteins we have examined the properties of DNA molecules containing inosine and/or 2,6-diaminopurine (DAP) residues. Appropriately substituted tyrT(A93) DNA fragments were prepared by the polymerase chain reaction. Their mobility in non-denaturing gels was affected, consistent with changed anisotropic flexibility leading to increased curvature due to G → I substitution and decreased curvature due to replacement of adenine with DAP. Band-shift assays of FIS protein binding revealed facilitated interaction with inosine-containing DNA and markedly reduced binding to DAP-containing DNA, attributable to altered bendability. DNase footprinting experiments confirmed that fewer sites would bind FIS in DAP-containing DNA at a given protein concentration, whereas higher levels of binding occurred with inosine-containing molecules. Thus base substitutions which affect the placement and presence of the purine 2-amino group in the minor groove can affect both the intrinsic curvature and the bendability of DNA.

© 1995 Academic Press Limited

Keywords: DNA flexibility; DNA bending; FIS; base substitutions; guanine 2-amino group

The problem of sequence specific DNA recognition is central to understanding the molecular mechanisms involved in control of cellular events such as transcription and replication. These processes require the harmonious functioning of complex machinery which implicates a network of DNA-binding proteins and their cofactors. The precise recognition of a defined nucleotide sequence in DNA by a given protein necessitates an optimal shape complementarity between the interacting species. In most cases, both the DNA and the ligand adapt their conformations so as to improve the fit between their respective recognition elements. The mutual conformational adjustment between the two binding components leads to stronger hydrogen-bonding, electrostatic contacts and van der Waals interactions and thereby affords more stringent sequence specificity.

In recent years numerous studies have been devoted to identifying the sequences and molecular features of DNA implicated in, or responsible for, the indirect recognition of DNA by proteins and drugs. It is now well established that both helix flexibility and bendability contribute significantly to that process (Travers, 1989, 1993). Certain alternating pyrimidine-purine elements such as CACA (Lyubchenko et al., 1993) are highly flexible and as such can be very sensitive to ligand-induced structural changes. It has become clear that the flexibility of DNA is dependent on its base composition (Hogan et al., 1989; Kahn et al., 1994) and that both the sequences at binding sites and regions flanking those sites contribute to the protein-DNA recognition (Koudelka et al., 1988; Lundin et al., 1994). However, the contribution of the individual bases in DNA and their substituents to the overall flexibility of the macromolecule remain unclear. The exocyclic 2-amino group of guanine is believed to represent an element of prime importance in DNA structure and recognition. Not only does it obstruct access to the floor of the minor groove in B-DNA, but it also disrupts the pattern of hydration, alters the electrostatic potential, and is the only hydrogen bond donor in the minor groove available to interact with a
ligand. It is believed to exert a significant influence on DNA bending and flexibility. Experiments with synthetic oligonucleotides containing modified bases have indicated that DNA curvature is modulated to a considerable extent by the purine 2-amino group (Diekmann et al., 1987, 1992; Koo & Crothers, 1987). In a recent study, we showed that the removal, addition or relocation of the purine 2-amino substituent strongly affects DNA cleavage by conformation-sensitive probes such as uranyl nitrate and DNase I, indicating that this element may play a key role in determining minor groove width (Bailly et al., 1995). Similarly, we have shown that the binding of small molecules and antibiotics to specific sequences in the minor groove is radically altered when the position of the 2-amino group is changed. Irrespective of their mode of interaction with DNA, the binding sites of all antibiotics tested were seen to follow strictly the placement of the purine 2-amino group, which accordingly serves as both a positive and negative effector within the minor groove of the B-form helix (Waring & Bailly, 1994; Bailly & Waring, 1995a,b). By the same token, the interaction with proteins which recognise DNA via the minor groove should be affected (Travers, 1995), for example the TATA-binding protein (Strubin & Struhl, 1992) and Hin recombinase (Feng et al., 1994).

Here we have examined the possibility that the purine 2-amino group could significantly affect DNA flexibility and major groove recognition. We explored this idea by examining the effects of base substitutions on the recognition of DNA by a major groove-binding protein, the *Escherichia coli* Factor for Inversion Stimulation (FIS). This protein is required for several processes including site-specific recombination, transcriptional activation and DNA replication (Finkel & Johnson, 1992). FIS binds to specific sites, each monomer having only a limited sequence preference mediated by a helix-turn-helix domain binding in the major groove (Kostrewa et al., 1991; Yuan et al., 1991). Each half-site in a 15 bp consensus sequence, 5'-GN13C, contains only a single highly conserved base (Hübner & Arber, 1989). However, the central core of the 15 bp consensus is AT-rich while YR base steps are usually found close to the extremities (Finkel & Johnson, 1992; Lazarus & Travers, 1993). These latter features are thought to confer flexibility allowing FIS to induce DNA bending. Accordingly, the binding of FIS to DNA should be highly sensitive to the bendability and flexibility of the target site and thus constitute an ideal system for studying DNA flexibility.

To test the effect of the 2-amino group of guanine on FIS binding (and indirectly DNA flexibility), we synthesized DNA molecules in which that group had been either removed from guanine, added to adenine, or both. The *Escherichia coli tyrT*(A93) promoter sequence retaining upstream sequences as far as position −98 relative to the *in vivo* transcription start (+1) was chosen as a substrate. This 160 bp fragment, which contains several FIS binding sites (Figure 1), differs from the well characterized wild-type *tyrT* promoter by a single A·T to T·A transversion at position −8, i.e. at position 93 relative to the in vivo transcription start (+1).
to the 5’ end. This transversion is a strong down-volatility mutation (Berman & Landy, 1979; Drew et al., 1985). We used the polymerase chain reaction to prepare homologous DNA samples having guanosine nucleotides replaced by inosines, adenine residues replaced by 2,6-diaminopurines (Figure 1), or both. The modified DNA species, as well as normal DNA prepared by the same route, were then studied for their ability to bind the FIS protein.

Figure 2 shows a autoradiogram from a non-denaturing gel electrophoresis experiment with the normal and substituted 160 bp tyrT(A93) DNA fragments. The addition, deletion or relocation of the purine 2-amino group exerts a significant effect on the electrophoretic mobility of these DNA molecules, attesting that the structure of the DNA has changed. Adenosine-substituted DNA migrates more slowly than normal DNA whereas DAP-substituted DNA migrates more rapidly. The A → DAP substitution has a marginally greater influence than the G → I substitution. The band corresponding to the doubly substituted I + DAP DNA is slightly retarded compared to normal DNA. These results can be interpreted in terms of the local conformation of the helix. Removal of the guanine 2-amino groups makes the former G-C base pairs, now I-C, behave like A-T base-pairs. This allows the resultant I·C pairs to adopt a high propeller twist (Diekmann et al., 1992) and permits short homopolymeric (dI)·(dC) tracts to confer intrinsic curvature on the DNA fragment (Diekmann et al., 1987, Koo & Crothers, 1987), which consequently migrates more slowly. This does not imply that the (dI)·(dC) tracts themselves are bent; only that they introduce additional bending by a mechanism presumably equivalent to that which operates at or close to (dA)·(dT) sequences. Conversely, substitution of DAP for adenine makes the 2-amino-A-T pairs behave like G-C pairs and abolishes the intrinsic curvature conferred by short (dA)·(dT) tracts. The effect of the double substitution is intermediate between that of the individual single substitutions.

In order to compare the interaction of FIS with natural and substituted DNA, we used the gel retardation technique which permits the separation of DNA-protein complexes from free DNA on polyacrylamide gels. The band-shift assays shown in Figure 3 reveal that the normal and modified tyrT(A93) DNA samples differ in their abilities to form complexes with FIS. Up to four retarded bands marked cI to cIV are readily apparent upon interaction of FIS with the different DNAs. These bands are believed to reflect the binding of FIS to the four sites centered at positions −71, −91, −51 and −19 of DNA was resuspended in the kinase buffer. The purified PCR products were 5’-end labelled with [γ-32P]ATP in the presence of T4 polynucleotide kinase according to a standard procedure for labelling blunt-ended DNA fragments. After completion the labelled DNA was again purified by 6% polyacrylamide gel electrophoresis and extracted from the gel as described above. Finally, the labelled DNA was resuspended in the binding buffer.
Figure 3. Gel retardation analysis of the interaction of FIS protein with normal and substituted tyrT(A93) DNA. PCR-generated DNA fragments were incubated with purified FIS and analysed on 6% non-denaturing polyacrylamide gels as described by Lazarus & Travers (1993). The stock solution of FIS protein was kept at −20°C in 50% (v/v) glycerol (kind gift from C. Koch); the (monomer) concentration was determined on the basis of a molecular weight value of 10,000. FIS was freshly diluted before each assay. The top three panels show the mobility of the normal, inosine- and DAP-containing DNAs in the absence (left) and presence (right) of 120 nM FIS. The bottom two panels show a concentration-dependence study for the binding of FIS to normal and I + DAP-containing DNA; in both cases, the 3P S-end labelled tyrT(A93) DNA was incubated with 0, 10, 25, 50, 100 and 150 nM FIS (lanes 0 to 5). Arrows indicate the different species that can be resolved: f is the free DNA; cI to cIV correspond to FIS-DNA complexes I to IV as depicted in Figure 1. Binding reactions were conducted in a total volume of 10 μl containing 25 mM Hepes buffer at pH 7.6, 0.1 mM EDTA, 1 mM DTT, 150 mM KCl and 10% glycerol.

Figure 4. Fraction of bound tyrT(A93) DNA as a function of the FIS concentration. Data were compiled from quantitative analysis of five gels and must be considered as a set of averaged values. A Molecular Dynamics 425E PhosphorImager was used to collect data from storage screens exposed to the dried gels overnight at room temperature. Scans were analysed by integrating the densities between two selected boundaries using ImageQuant version 3.3 software.

The FIS-DNA interaction. Complexes between FIS and DAP-containing tyrT(A93) DNA molecules can only be detected at much higher protein concentrations than with normal DNA. Conversely, the G → I substitution facilitates the binding of FIS to DNA. Binding of FIS to the weak affinity sites ending at position −58 and −26 (cIII and cIV, Figure 1) is more readily apparent with the inosine-containing DNA than with the normal DNA molecules. Complex formation with the normal and I + DAP DNA samples requires slightly higher concentrations of the protein to produce similar band intensities in the gel retardation assays. Titration of each species of the DNA with increasing concentrations of FIS provides some gauge of the relative affinity of the protein for the nucleic acid. The graph in Figure 4 showing the increase in the fraction of bound DNA as a function of protein concentration enables us to examine directly the capacity of the DNA-bending protein to form complexes with the different tyrT(A93) fragments. The relative affinities of the variously substituted DNA fragments for FIS fall in the order: inosine DNA > normal DNA ≥ I + DAP DNA ≥ DAP DNA. Three of the binding curves are clearly co-operative in character while the fourth, that for DAP-DNA, is insufficiently complete to determine whether there is any change in co-operativity resulting from the base substitution. These experiments were performed in a magnesium-free buffer but, in order to permit direct comparison with the DNase I footprinting data presented below, we repeated the gel retardation

the tyrT(A93) fragment (Figure 1). The band shift observed with normal DNA is identical to that reported by Lazarus & Travers (1993) using the same −98 construct. The ability of FIS to form complexes with the modified tyrT(A93) fragments varies significantly depending on the nature of the base substitutions, i.e. on the position of the purine 2-amino group. The A → DAP substitution markedly reduces
Figure 5. DNase I footprinting of FIS on the Watson strand of tyrT(A93) DNA containing the four natural nucleotides (Normal DNA), inosine residues in place of guanosine (Inosine DNA), DAP in place of adenines (DAP DNA) or inosine and DAP residues in place of guanine and adenine, respectively (I + DAP DNA). Each pair of lanes corresponds to digestion of the DNA in the absence (left) and presence (right) of FIS. In panel A the FIS concentration was 100 nM; in panel B it was 200 nM. The cleavage products resulting from DNase I digestion were resolved on an 8% polyacrylamide gel containing 8 M urea. Chemical identities of the digestion products were assigned by reference to the sequencing markers (lane G + A), taking into account the differences in mobility of the fragments due to the presence or absence of a 3'-phosphate group. The scale on the left corresponds to the numbering of the tyrT(A93) fragment as represented in Figure 1. Boxes on the right refer to the FIS binding sites cI to IV as shown in Figure 1. DNase I cleavage reactions were performed in 18 ml of the buffer used for the gel retardation assay (Figure 3) supplemented with Mn and Mg in order to potentiate the action of the nuclease. The digestion was initiated by the addition of 2 ml of a DNase I solution (Sigma Chemical Co.) whose concentration was adjusted to yield a final enzyme concentration of about 0.01 unit/ml in the reaction mixture. The enzyme was stored as 7200 units/ml solution in 20 mM NaCl, 2 mM MgCl₂, 2 mM MnCl₂ (pH 8.0) at −20°C, and was freshly diluted to the desired concentration immediately prior to use. The extent of digestion was limited to less than 30% of the starting material so as to minimize the incidence of multiple cuts in any strand (“single-hit” kinetic conditions). Optimal enzyme dilutions were established in preliminary calibration experiments. After three minutes, the digestion of the DNA was stopped by adding 80 ml of a stop solution containing 10 mM EDTA. Samples were immediately precipitated with three volumes of ethanol, washed once with 200 ml of 70% ethanol, dried and then resuspended in 4 ml of an 80% formamide solution containing tracking dyes. Samples were heated at 90°C for four minutes and chilled in ice for four minutes prior to electrophoresis under denaturing conditions (8% acrylamide, 8 M urea) in TBE buffer. Gels were soaked in 10% acetic acid for 15 minutes, transferred to Whatman 3MM paper, dried under vacuum at 80°C, and subjected to autoradiography at −70°C with an intensifying screen.
Communication

assays in a buffer supplemented with 2 mM MgCl₂. The same ranking order was observed for the different DNA fragments (data not shown).

Does the protein recognize the same sequences in normal and modified tyrT(A93) DNA? To answer this question we performed DNase I footprinting experiments aimed at defining the FIS binding sites. Comparison of the DNase I cleavage patterns in the presence and absence of the protein (Figure 5) shows that FIS binds to the same sites in all four DNA species, normal and modified. In each case, the cleavage by the nuclease is strongly inhibited at the sequence extending from position −66 to −83. The protected region encompasses the 15 bp sequence 5'-GGATGAAATTACGCG which perfectly matches the consensus sequence and probably corresponds to the complex cII in the gel retardation assay. The complex cII is believed to originate from binding of FIS to the −91 site which is not accessible to the footprinting analysis because it lies at the very 5'-end of the −98 tyrT(A93) construct. Footprints at sites III and IV are much more pronounced with the inosine DNA than with normal DNA. The same two low-affinity sites can also be identified with the doubly substituted I + DAP DNA although the footprints at these two regions are less pronounced than with normal DNA. With the DAP DNA, only the footprint at site I is reliably detectable. With this DNA, cleavage by the nuclease at sites III and IV in the presence of a high FIS concentration is little different from the protein-free DNA. Therefore, the footprinting data agree fully with the gel retardation assay to show that (1) the A → DAP substitution is detrimental to the binding of the DNA-bending protein, (2) conversely, the G → I substitution facilitates interaction between FIS and DNA, though possibly at a lower level of stringency compared to the natural molecule, and (3) the combined G → I and A → DAP substitutions exert a weak but noticeable effect on the capacity of the protein to bind DNA. From the gel retardation assay, it seems that the relative affinity of FIS for the I + DAP DNA is slightly less than that for the normal DNA containing all four canonical bases.

The results described in this paper show that alterations to the presence and location of the 2-amino group in tyrT(A93) DNA alter the electrophoretic mobility of the fragment and also its affinity for FIS protein. Both these observations can be interpreted on the basis of changes in curvature and anisotropic flexibility consequent upon the base substitutions.

The anisotropic flexibility of the tyrT(A93) DNA fragment was originally demonstrated by its adoption of a highly preferred rotational orientation in both minicircles and on the surface of the histone octamer (Drew & Travers, 1985). The fragment possesses only a small degree of intrinsic curvature (Drew & Travers, 1985; Plaskon & Wartell, 1987). The decrease in electrophoretic mobility on substitution of inosine for guanosine and the increase on substitution of diaminopurine for adenine indicate that these alterations respectively increase and decrease this intrinsic curvature. The fragment contains oligo(dA)ₙ·(dT)ₙ tracts at positions −71, −50 and +31, with oligo(dG)ₙ·(dC)ₙ tracts at positions −2, +40 and +49 (Figure 1). All these tracts are in approximate helical phase with each other. Since oligo(dI)·(dC) tracts confer intrinsic curvature in the same way as oligo(dA)·(dT) tracts (Diedkann et al., 1987, 1992) the observed increase in total curvature consequent upon G → I substitution can be explained by the effective addition of three in-phase sequences each of which contributes intrinsic curvature. By contrast the A → DAP substitution abolishes the ability of oligo(dA)·(dT) tracts to confer intrinsic curvature, consistent with the faster mobility observed here. The additive effect of the two substitutions confirms that their individual effects are independent.

The effects of base substitutions on FIS binding can be explained in a similar way. The fragment possesses two strong sites, centred at positions −72 and −92, respectively, together with at least two weaker sites (Figure 1). In general, at low FIS concentrations occupation of weaker sites by the protein is dependent on co-operative binding (Muskhelishvili et al., 1995). The data show that the occupation of the strong binding sites by FIS is not significantly affected by the G → I substitution suggesting either that recognition of the G-C base-pairs in the consensus is not dependent on the guanine moiety or that structural recognition of DNA is dominant. However, the substitution potentiates the occupation of the weaker sites in phase with the primary sites (Figure 3). These weaker sites are distant from the oligo(dG)·(dC) tracts and so it appears most likely that this effect is a consequence of local changes in anisotropic flexibility. The base substitution would act by giving a G + C-rich sequence, for example in the vicinity of −30, the potential to adopt a conformation with a narrow minor groove (Diedkann et al., 1992) in phase with the curvature induced by FIS dimers binding at the strong sites. By contrast the A → DAP substitution would reduce anisotropic flexibility, and hence the co-operativity of FIS binding, by limiting the ability of in-phase A + T-rich sequences to adopt a conformation with the high propeller twist normally associated with a narrow minor groove.

We therefore conclude that base substitutions that alter the placement and presence of the purine 2-amino group in the minor groove can affect both the intrinsic curvature and the bendability of DNA, and thereby modulate the interaction with proteins whose binding sites lie purely within the major groove of the double helix.

Acknowledgements

We thank Dean Gentle for his invaluable technical assistance. This work is a contribution from the Cambridge Centre for Molecular Recognition and was supported by grants (to M.J.W.) from the Cancer Research Campaign, the Wellcome Trust, and the Association for International Cancer Research. C.B. thanks the INSERM, the ARC and the Sir Halley Stewart Trust.
References


Edited by T. Richmond

(Received 24 May 1995; accepted 4 August 1995)