Hydroxyl radical footprints reveal novel structural features around the NF I binding site in adenovirus DNA

Haralabos Zorbas, Lars Rogge, Michael Meisterernst and Ernst-Ludwig Winnacker

Institut für Biochemie, Universität München, Karlstraße 23, D-8000 Munich 2, FRG

Received June 28, 1989; Revised and Accepted August 28, 1989

ABSTRACT

We have identified a number of as yet unknown structural abnormalities of the NF I-DNA binding site within the inverted terminal repetition of adenovirus DNA by probing it with a hydroxyl radical footprinting technique. NF I binding alters the accessibility of the deoxyribose moieties to hydroxyl radicals both at the 3' and at the 5' side of the recognition sequence 5'-TGG(N)₆GCCAA-3'. A smooth bend at the 5' side of the binding sequence is already present in naked linear DNA and it is further enhanced by protein binding. This could be demonstrated not only by hydroxyl radical footprinting but also by studying the temperature dependent mobility during gel electrophoresis of DNA fragments carrying the NF I binding site at circularly permuted positions. We propose that the bent conformation at this site is responsible for facilitating protein/DNA interactions.

INTRODUCTION

Nuclear factor I (NF I) was originally identified as a cellular protein which binds to the sequence 5'-TGG(N)₆GCCAA-3' within the inverted terminal repeats of adenovirus DNA and which stimulates viral DNA replication *in vitro* (1, 2). NFI-like activities occur in a variety of mammalian cells (3-8) and have been proposed to be identical with the 'CCAAT'-box transcription factors (CTF) which bind to the 5'-CCAAT-3' motif present within many eukaryotic promoter regions (9). From the structural characterization of cDNA species it became clear that several related but distinct NF I like proteins exist (4, 5, 10, 11). We have recently isolated an NF I protein species from porcine liver by virtue of its ability to bind to the original adenoviral DNA site and to stimulate viral DNA replication *in vitro* (3,12). This protein, however, turned out to bind only poorly to a variety of 'CCAAT' boxes with the possible exception of the α -globin promoter (12). As shown by the analysis of the corresponding cDNA, the 36 kDa porcine liver isolate is a protease-resistant fragment of the intact 55 kDa protein, which retains the entire DNA binding domain (11). In this communication we describe the interaction of this protein with its binding site in more detail by using Fe(EDTA)²⁻-generated hydroxyl radicals as footprinting agents (13). The use of this reagent revealed structural details around the NF I recognition site within the adenovirus inverted terminal repetition both in the absence and presence of the protein, which have not been detected in previous contact point analyses (e.g., ref. (14)).

MATERIALS AND METHODS*Bacterial strains and plasmids*

Plasmid pCMS has been described (12). Vector Bluescript M13(-) (49) was purchased from Stratagene, San Diego. All plasmids were propagated in *E.coli* strain XL 1-Blue

(Stratagene, San Diego) in 2×YT medium supplemented with 100 µg/ml ampicillin and 10 µg/ml tetracycline.

Enzymes and nucleotides

Enzymes XhoI, EcoRI, and calf intestine phosphatase were from Boehringer, Mannheim; HindIII, and NotI were from Biolabs, Schwalbach; KpnI, T4 polynucleotide kinase, and Mu-MLV reverse transcriptase were from BRL, Eggenstein. Incubation conditions were as recommended by the suppliers. γ -³²P-ATP and α -³²P-dATP (both at 3000 Ci/mmol) were from Amersham, Braunschweig.

Construction and in vitro manipulations of pTAd5

Fragment F146 carrying the left-hand terminus of adenovirus type 5 DNA (12) was re-cloned into the EcoRI and HindIII site of Bluescript M13(–) by conventional techniques (50) yielding pTAd5. For labeling the adenoviral (–) strand, 50 µg of pTAd5 were cleaved with XhoI, dephosphorylated at the 5' ends with calf intestine phosphatase, cleaved with KpnI, purified over a G50 Nick-Column (Pharmacia, Freiburg) and stored in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA. Whenever needed, aliquots of 1 µg (about 1 pmol 5'-OH ends) were labeled with T4 polynucleotide kinase and γ -³²P-ATP (10-fold molar excess over the 5'-OH ends) and purified from unincorporated ATP over a G50 Nick-Column, resulting in a yield of about 7×10^6 cpm. For labeling the (+) strand, an aliquot (1 µg) of the XhoI-cleaved DNA was filled-in with reverse transcriptase, α -³²P-dATP (10-fold molar excess over the 3' recessed ends) and 500 µM each of dGTP, dCTP and TTP, digested with KpnI and purified as above. For chemical sequencing, DNase I footprints, and hydroxyl radical footprints with purified NFI, the DNA labeled at either the 5' or the 3' end could be used directly; labeled DNA to be employed in gel retention assays was shortened by EcoRI cleavage prior to the binding reaction (see Figure 1).

For the generation of fragments 'E' and 'H' (see Results) pTAd5 was digested with EcoRI or HindIII, respectively, filled-in at the 3' ends with reverse transcriptase as above and cleaved either with KpnI or NotI, respectively.

DNA-sequencing

Chemical sequencing reactions were performed as described (51).

DNase I footprints

DNase I footprints have been described (12). The amount of protein necessary for complete binding of the DNA molecules was determined in analytical gel retention assays (52).

Hydroxyl radical footprints

Conditions for the binding of NF I were similar as described in (12) except that glycerol and dithiothreitol were omitted in the binding buffer (25 mM HEPES-KOH, pH 7.5, 150 mM NaCl). After 15 minutes of incubation at room temperature the reaction mixture was diluted with binding buffer to a final glycerol concentration of 0.2%. Reagents for the hydroxyl radical cleavage reaction (final concentrations: 1 mM ascorbate, 20 µM EDTA, 10 µM (NH₄)₂Fe(SO₄)₂, 0.3% H₂O₂) were added immediately. The higher concentration of H₂O₂ (cf. (16)) increases the yield without impairing the specific cleavage pattern. Incubation was carried out for 5 minutes at room temperature. Reactions for immediate analysis were stopped (16), extracted with phenol, concentrated with ethanol and analysed in a sequencing gel. When the complexes had to be separated from the unbound molecules, reactions were stopped by the addition of 10 mM thiourea, 2 mM EDTA, 5% glycerol and 0.125% bromophenol blue and the reaction products were run on non-denaturing polyacrylamide gels (53) with the omission of sodium-dodecyl sulfate. Bound complexes

Figure 1: Preparation of plasmid pTAd5 and the labeling strategy of its 5' and 3' ends. Plasmid pCMS contains a 146 bp long DNA fragment from the left-hand inverted terminal repetition of adenovirus type 5 DNA (open box) with the NF I binding site (hatched box). It was recombined into the multiple cloning site (black bar) of Bluescript M13(-). RI=EcoRI; H=HindIII; X=XhoI; K=KpnI; CIP=calf intestine phosphatase; PNK: T4 polynucleotide kinase. The end labeled DNA was cleaved with EcoRI (in brackets) only when the fragments had to be isolated from a native retention gel.

were recovered from the gel slices by isotachopheresis (54) and analysed in a sequencing gel after purification and concentration.

Analytical gel retention

For the analysis of bent DNA, labeled fragments were run in native retention gels as described (52) by using an SE 620 electrophoresis system (Hoefer Scientific Instruments, San Francisco) in tandem with a Brinkman RM 3 water bath, or alternatively, in non-denaturing sequencing gels (LKB, Freiburg) in TBE buffer (90 mM Tris-boric acid, pH 8.3, 2.5 mM EDTA) with the same results.

Sequencing gels and densitometric scanning

Sequencing gels were prepared and run as recommended (55). The buffer composition was $0.5 \times$ TBE (see above), 7 M urea. The results of hydroxyl radical protection experiments were quantitated by densitometric scanning of the films with an Ultrascan XL laser densitometer (LKB, Freiburg). Levels of protection were calculated as (band intensity in the absence of NF I)/(band intensity in the presence of NF I) (cf. (48)).

RESULTS

DNase I footprints

DNase I footprints of the porcine liver 36 kDa NF I isolate have been determined previously

by us for the (+) strand of the inverted terminal repetition of adenovirus type 5 DNA (12). As a reference for the hydroxyl radical footprints we analyzed the DNase I protected regions of both strands with increased sensitivity and by a modified, rapid procedure (see Materials and Methods). For this purpose and for all subsequent experimental steps a 146 bp long terminal fragment of adenovirus type 5 DNA comprising the NF I site was recloned into appropriate sites of the plasmid vector Bluescript M13(-) (Figure 1). For the analysis of the (-) strand the vector DNA was labeled at the 5' end of the XhoI site within the polylinker, whereas for the analysis of the (+) strand the same XhoI site was labeled at the 3' end (Figure 1).

NF I binding protects the phosphodiester bonds between G₁₈ and C₄₂ on the (-) strand (see Figures 2 and 6) and between T₂₀ and T₄₄ on the (+) strand (see Figures 3 and 6) from DNase I digestion. The binding pattern is independent of whether liver-extracts, the purified liver-protein fragment (36 kDa, see Figure 2, lanes 7, 8 and 9; ref. (12)), or the protein expressed *in vitro* from the full length cDNA clone pCTF2/NF I (about 55 kDa, see ref. (11); data not shown) are being used. The 36 kDa fragment thus contains the full and intact DNA binding domain of NF I (5, 11).

Hydroxyl radical footprints on the (-) strand

Free hydroxyl radicals generated by the Fenton reaction have been introduced as a footprinting agent by Tullius and coworkers (13). Their use results in a superb and conformation-sensitive resolution of naked DNA as well as of protein/DNA proximity at single nucleotide positions ((18, 19, 21); for reviews see (16,17)). Unfortunately, the yield of hydroxyl radicals, as generated by the Fenton reaction, is dramatically reduced by the presence of glycerol, which is often present in protein purification and storage buffers. We succeeded in footprinting NF I by working with a final glycerol concentration of only 0.2% (see Materials and Methods). Footprints with purified NF I could be analysed directly on a sequencing gel. With partially purified fractions (Mono Q extract, see ref. (12)), the reaction was slowed down dramatically, most probably due to the activity of H₂O₂-metabolizing enzymes. Footprints with these extracts could be obtained, however, by using an excess of input DNA and by purifying the protein/DNA complexes in a native retention gel prior to footprint analysis (see Materials and Methods). No footprints could be produced with full-length NF I translated *in vitro* in commercial reticulocyte extracts (11), since the Fenton reaction stops within seconds in the presence of these extracts.

The footprint results are shown in Figure 2 for the (-) strand and in Figure 3 for the (+) strand. Quantitative analyses by densitometric scanning are depicted in Figure 4; all data are summarized in Figure 6. From the experiments shown in Figure 2, lane 10, it is obvious that the binding site is protected from hydroxyl radical attack after protein binding at three regions along the (-) strand, around residues T₂₁, A₂₉, and T₃₇. The pattern obtained with pure NF I (lane 10) differs from that obtained with Mono Q extracts (lane 15) in that residues G₃₅ and G₃₆ appear protected only in the latter. However, after gel electrophoretic separation of bound and unbound DNA fragments in a retention gel (see above), these differences disappear (compare lanes 14 and 15). Therefore, the additional signals at G₃₅ and G₃₆ in lane 10 possibly originate from the presence of an excess of DNA fragments which bind NF I less tightly.

The protected regions around T₃₇ and A₂₉ (see Figure 6) correspond to the regions where contacts of HeLa NF I with the base and phosphate moieties of the DNA have been detected by different methods (14) and thus represent sites of proximity between the DNA and the protein.

Figure 2: Footprint analysis of the (-) strand of the left-hand inverted terminal repetition of adenovirus type 5 DNA. The analyses were performed on sequencing gels as described in Materials and Methods. Lane 1 (M): molecular weight marker (ϕ X RF I DNA, cleaved with HaeIII and 3' labeled). Lanes 2-5, 12, 13, 17, 18: Maxam and Gilbert sequencing ladders; G: guanines; Pu: purines; Py: pyrimidines; C: cytosines. Lanes 6-9: DNase I footprints; lane 6: no protein; lane 7: purified NF I; lanes 8 and 9: decreasing amounts of Mono Q extract. Lanes 10, 11, 14-16: Hydroxyl radical footprints; lanes 11 and 16: no protein; lanes 10 and 14: pure NF I; lane 15: Mono Q extract; lanes 14 and 15: protection pattern after gel separation of the bound complexes. The underlined sequences in the sequence insert mark the regions protected after protein binding; numbering begins at the terminal 3' G residue of the left-hand inverted terminal repetition of adenovirus type 5 DNA.

Figure 3: Footprint analysis of the (+) strand of the left-hand inverted terminal repetition of adenovirus type 5 DNA. The analyses were performed on sequencing gels as described in Materials and Methods. *Lanes 1, 2:* Maxam and Gilbert sequencing ladders of the (-) strand (5'-labeled). *Lane 3:* Hydroxyl radical reaction products of the (-) strand without protein. *Lane 4:* Hydroxyl radical reaction products of the (+) strand without protein. *Lanes 5, 6, 12-15:* Maxam and Gilbert sequencing ladders of the (+) strand (3'-labeled). G, Pu, Py, C: as in Figure 2. *Lanes 7, 8:* DNase I footprints of the (+) strand; *lane 7:* no protein; *lane 8:* Mono Q extract. *Lanes 10, 11, 16:* Hydroxyl radical reaction products of the (+) strand without protein; the bracket on the right side lane 16 marks the protected T residues on the naked DNA. *Lane 9:* Hydroxyl radical footprints of the (+) strand with Mono Q extract. Underlined sequences and numbering: as in Figure 2; asterisks mark residues of enhanced reactivity after protein binding.

However, no DNA/proteins contacts have been described previously for the third protected region around nucleotide T₂₁. Remarkably, this region already displays a slight but significant protection pattern even in the absence of protein (see lanes 11 or 16). The

protected residues A₂₂, T₂₁, and A₂₀ are located at the 3' end of the sequence 5'-CA₄TA₂-3' which is very reminiscent of 5'-CA₅₋₆T-3' known to be responsible for 'bending' of the kinetoplast DNA (K-DNA, ref. (20)). Hydroxyl radicals have been shown to cleave residues towards the 3' end of this sequence at lowest frequency with a smooth and progressive increase in cutting frequency of the nucleotides located further downstream (21). This is exactly what we observe at this region of the NF I binding site (see Figure 4, A and C). The distortion of the even cleavage pattern in this region is greatly enhanced after protein binding (visible in Figure 2, lanes 10, 14 and 15.) In analogy to the K-DNA, we therefore suggest a smooth bend to occur at this protected region of the DNA, the degree of which is increased by protein binding (see also below).

Hydroxyl radical footprints on the (+) strand

Also the (+) strand appears to be protected at three regions, namely around nucleotides T₂₅, A₃₂, and T₄₄ (see Figure 3, lane 9 and Figure 6).

The regions around T₂₅ and A₃₂ correspond to the known contacts of NF I with bases and phosphates (14). Although we also interpret the diminished reactivity around T₂₅ as protein proximity to the deoxyribose moieties, it is obvious that the corresponding region of the naked DNA is also protected from the hydroxyl radicals to a certain degree prior to protein binding (see Figure 3, lanes 4, 10, 11, and most clearly lane 16).

Region T₂₃, T₂₄, and T₂₅ is offset by 2 or 3 nucleotides in the 3' direction compared with the pattern of protected adenines on the complementary (-) strand (compare lanes 3 and 4 in Figure 3; compare A, C, and E in Figure 4; see also Figure 6). Again, this shift is very similar to the protection pattern of the kinetoplast-DNA bending unit (ref. (21), see also Discussion). We thus conclude that after protein binding the deoxyribose residues around T₂₅ are protected in part because of protein proximity and in part because of the enhanced alteration of the DNA conformation, *i.e.*, enhanced bending at this site.

We also observe changes in hydroxyl radical susceptibility around residues T₄₄ and T₃₉. The diminished reactivity around T₄₄ extends over the region protected from DNase I digestion (see Figure 6) and thus, by definition, cannot be caused by protein contacts. An increased susceptibility of residues A₃₈, T₃₉, and A₄₀ to the radicals is observed after protein binding (see Figure 3, lanes 9 and 10). We believe that both the increased reactivity around T₃₉ and the decreased reactivity around T₄₄ reflect induced changes in the DNA structure after protein binding rather than protein contacts (see Discussion).

Further support for DNA bending at the NF I binding site

In order to provide supporting evidence for the proposed bent structure of the DNA at the 5' border of the NF I recognition site by an independent method, we performed a mobility analysis by polyacrylamide gel electrophoresis of DNA fragments containing the NF I site. A DNA molecule containing a bent structure is retarded in the gel to a greater extent, if this structure is placed at the center of the fragment than if it is present at an end (ref. (20) and refs. (29, 30) therein). In order to circularly permute the NF I site we took advantage of the presence of many restriction sites surrounding fragment F146 in plasmid pTAd5, which carries the left-end terminus of adenovirus 5-DNA. pTAd5 was cleaved at appropriate sites as depicted in Figure 5A, giving rise to two fragments of about 190 bp each differing in length by only one nucleotide. The NF I site on the fragment designated 'E' is present close to one of the ends, whereas on fragment 'H' it is located more centrally. If the fragments are bent at the location suggested from the hydroxyl radical experiments (*i.e.*, NF I site), then fragment 'H' should be retarded stronger than fragment 'E' during electrophoresis. Because bending is stabilized at low temperatures (24–26), we compared the migration behaviour of the two fragments at a temperature of 0°C (see

Figure 5B). At this temperature fragment 'E' moves faster than 'H', as expected. If caused by a bend, these mobility differences should be reduced at higher temperatures (24–26). As shown in Figure 5B mobility differences can indeed be detected at 50°C only after long electrophoresis times; even then they are much less pronounced. At the physiological temperature of 37°C these mobility differences still persist (Figure 5B), and become much more obvious in the migrating protein-bound complexes. These data are consistent with the conclusions from the hydroxyl radical experiments in the previous sections, namely, a pre-existing bend at the NF I site, and an even stronger bending of the same site after NF I binding (see also Discussion).

DISCUSSION

We have investigated the interaction of porcine liver NF I with the NF I binding site within the inverted terminal repetition of adenovirus type 5 DNA. Sequencing and analysis of the DNA by DNase I and hydroxyl radical footprinting was facilitated by employing a labeling strategy which allowed immediate use of the DNA after labeling (see Fig. 1); this strategy accelerates the overall procedure and routinely gives fully reproducible results, with maximum fragment yield and specific radioactivity. This is of importance especially for the hydroxyl radical footprint analyses in which the amount, and hence the signal intensity, of input DNA is limited by the little amount of protein that can be used in order to keep the glycerol concentration as low as possible.

By comparing the DNase I protection pattern with the sequencing ladder we could accurately determine the regions protected by NF I on both strands. Due to differences in the chemical nature of the digestion steps, and thus the reaction products, the DNase I-derived bands are retarded by a 1.5 or a 1 band distance, depending on whether the DNA is labeled at the 5' or the 3' end, respectively, as compared to the cleavage products of the Maxam and Gilbert reactions (27). Taking this into account, binding of either the porcine liver '36 kDa' fragment (12) or the 55 kDa NF I full-length protein (11) protects exactly the same 24 phosphodiester bonds on both strands, that were determined for HeLa NF I (14), the prototype NF I species. This result further demonstrates the structural identity of the DNA binding domains of both NF I proteins (11). We therefore, suggest that the hydroxyl radical analysis performed here with the porcine NF I binding domain is directly relevant to that of the HeLa cell derived protein.

By using hydroxyl radical footprinting we observed protein proximity to the deoxyribose moieties around residues T₂₅, A₂₉, A₃₂, and T₃₇. The maxima of binding activities on either of the two strands are separated by 9–10 nucleotides (see Figure 4 and 6). This indicates that contacts occur at two consecutive DNA regions separated by a full helical turn and, thus, occur on one side of the DNA double strand only. The distance between adjacent maxima on the complementary strands is 6–7 nucleotides, which is indicative of contacts within the major groove of the DNA (see (16)). Phosphate and base contacts have been detected previously by others for the same regions of the protein binding site (14) and are in perfect agreement with the protein contacts identified in this study.

Two additional as yet unrecognized regions are affected *indirectly* by protein binding.

Figure 4: Densitometric analysis of the hydroxyl radical footprints on the left-hand inverted terminal repetition of adenovirus type 5 DNA. *A, B*: densitometric scans of Figure 2, lanes 11 and 10, respectively. *C, D*: densitometric scans of Figure 2, lanes 16 and 15, respectively. *E, F*: densitometric scans of Figure 3, lanes 10 and 9, respectively. Scans *A, C, and E* (solid lines) represent naked DNA and scans *B, D, and F* (dotted lines) protein bound DNA.

Figure 5: Analysis of DNA bending at the adenoviral NF I site. Fragment F146 was removed from pTAd5 by different restriction digests in order to obtain fragments 'E' and 'H' which are of identical length but in which the NF I site is located at different positions. *Panel A* : Generation of fragments 'E' and 'H'. Numbers '1 : 6.6' and '1 : 1.9' mark the ratios of the distances of the NF I site from the respective left or right-hand end of the fragments. Symbols: as in Figure 1. N= NotI; RI=EcoRI; H=HindIII; K=KpnI. *Panel B* : Gel electrophoretic analysis of fragments 'E' and 'H' at the indicated temperatures and at short (a) and longer (b) electrophoresis times. M= Molecular weight marker positions ('1 kb ladder' (BRL, Eggenstein), labeled at the 5' ends) are indicated. At 37°C, analyses were performed in the absence (-) and presence (+) of NF I.

Figure 6: Summary of the NF I footprint data. Coordinates: as in Figure 2. Bold letters: base contacts; arrows: phosphate contacts; base and phosphate contacts are taken from (14). Brackets: DNase I-protected region; solid circles: diminished deoxyribose reactivity before protein binding; small and large open circles: low and highly diminished deoxyribose reactivity after protein binding, respectively; rectangles (at G₃₅ and G₃₆): deoxyribose contacts visible only after gel separation (see text); asterisks: increased deoxyribose reactivity. Fully quantitative values of the relative levels of protection from hydroxyl radical cleavage after NF I binding are indicated by the histogram silhouettes; for the (+) strand the data from panels E and F, and for the (-) strand the data from panels C and D of Figure 4 were processed (see Materials and Methods).

One of them lies 5' to the NF I recognition site and has been identified as a preformed bend on the basis of a reduced susceptibility to hydroxyl radical digestion, of altered electrophoretic mobility of fragments containing a circularly permuted binding site, and of temperature dependence of the anomalous electrophoretic mobility.

Fragments 'E' and 'H' used in the mobility analysis possess the same size but have a partially different nucleotide composition, due to the different halves of the polylinker moiety (see Fig. 5A). Although similar approaches have been used by other authors in different systems (22, 23), a differential influence of these foreign sequences on the fragment mobility complicating the interpretation of the results cannot be totally excluded. From the experimental data in Fig. 5B, simulation of the bend at the NF I site would be possible only, if the polylinker were itself bent at the border to the Eco RI site on fragment 'H'. However, neither inspection of the polylinker at this site reveals the sequence requirements usually associated with a bend (oligoA, oligoT stretches), nor is there any indication from the hydroxyl radical cleavage pattern, that a bend of similar extent as that at the NF I site occurs or is enhanced after protein binding at this region. Therefore, an obscuring influence of the polylinker on the mobility analysis in the present case is highly improbable.

We consider the sequence 5'-CA₄TA₂-3' on the (-) strand to be responsible for the distortion of the DNA B-form. This sequence resembles the core sequence 5'-CA₅-6-T-3' of the kinetoplast (mitochondrial) minicircles, which is the first natural bent DNA analysed in detail (20, 25, 26, 28, 29; for a review see (30)). Although not identical, we assume these two sequences to adopt a similar conformation in solution. It has been pointed out, that 3' terminal thymidines in the core sequence do not impair the characteristic hydroxyl radical cleavage pattern (21) and that thymidines have the weakest effect among the other

nucleotides in interrupting the influence of an A stretch (31). The bend observed at the 5' side of the NF I recognition sequence is enhanced by protein binding.

What is the actual conformation of the DNA strands at the bent site? Wedges at ApA dinucleotides (32), heteronomous strands (33), or even cruciforms (23) have been suggested as the basis of sequence-directed DNA curvature. Although a cruciform structure can be drawn at the NF I site (2), we expect its stability to be extremely low, because of the linearity of the adenoviral DNA. Instead, similarly to the K-DNA case (see (21)), we consider a narrowing of the minor groove to be the structural determinant of the bent DNA axis. A reduced DNase I cleavage at this AT rich region prior to protein binding (see Figure 2, lane 6 and Figure 3, lane 7), indicative of a narrowness of the minor groove (34), further supports this hypothesis. We thus assume that binding of NF I in the major groove of the DNA increases the preformed bend by further contracting the adjacent minor groove.

The extent of bending at the NF I binding site is comparable to that of the 102 bp HindIII-AatII restriction fragment of pBR322 (21), the SV40 ORI region I (35), or the human β -actin-specific conserved promoter sequences (36), as is inferred from the hydroxyl radical cleavage pattern, or from comparisons of fragment mobility. Site I of the SV40 origin region was reported to be bent due to the CA₆ spacer sequence, and it was found that the bend persists or is even enhanced after large T-antigen binding (35). Static bent DNA structures have been found also in other replication origins, for example, in the λ -ORI (37, 38), in DNA fragments with ARS activity in yeast (39, 40) and in the pT181 ORI (41), as well as in several promoter regions (24, 36, 42, 43). Most of these depend on the presence of one or more oligo A stretches in frame with the DNA helix periodicity (35, 36, 37, 39, 40, 42).

DNA regions with bent character are critical for proper ORI function (see refs. (35, 37, 38, 39, 40)). Accordingly, there is evidence that NF I function is influenced by nucleotide positions just outside of the core binding sequence, exactly where the bend occurs: a mutant at nucleotide positions 22 and 23 converting the two AT base pairs to GC displays diminished binding and severe reduction of replication efficiency (44). In addition, the presence of two GC base pairs at positions 23 and 24 abolished binding of NF I due to structural alterations (15). The GC substitutions are expected to destroy bending by altering the specific sequence characterized above. In contrast, a mutant NF I site in which CA₄TA₂ is changed to CA₇ shows enhanced NF I binding (12) most probably because the lengthened A stretch favours bending (cf. (31)). An inverted NF I site (14) with similar protein/DNA interaction properties as the wildtype, changes the sequence CA₄TA₂ to CA₂TATA₂ but retains the AT rich character. Therefore, we propose this sequence to have characteristics similar to a bent DNA (cf. (31)).

Taken together, our data suggest that the pre-existing bend that is enhanced after NF I binding, facilitates the latter by providing the steric requirements for a successful DNA/protein interaction.

Why was bending at the NF I site not identified previously? NF I footprints have been performed with hydroxyl radicals only after intercalation of methidium-propyl-EDTA into the DNA (45). This intercalator is likely to disrupt specific conformations of the DNA back-bone in a way similar to that caused by the propidium ion (46). A study aimed at detecting unusual structures in natural DNA (47) could not be expected to detect the smooth bend at the 5' border of the NF I site, because of the terminal placement of the latter on the smallest analysed fragment of 196 bp (see Figure 3 in ref. (47)). Our investigations

thus demonstrate the possibility not only to analyse protein/DNA interactions with hydroxyl radicals but also to detect indigenous conformational features of naked DNA.

At the 3' border of the NF I binding site we observe distortions due both to enhanced and weakened hydroxyl radical reactivity. The reduced reactivity around residues T₄₄, A₄₅, and A₄₆ cannot be caused by direct protein contacts since this region extends beyond the DNase I-protected area (see Figure 6). An *increased* rather than a decreased reactivity around T₃₉ is likewise inconsistent with protein proximity to the attacked residues. Therefore, we believe both regions to obtain an altered *conformation* by protein binding that is detected by the free hydroxyl radicals.

The presence of an NF I 'arm' has previously been discussed to account for a contact at T₃₉ (14, 48). However, it is striking, that none of the surrounding bases (A₃₈, A₄₀), and especially none of the phosphate residues could be shown to contact the protein in this particular study (14). Our results also confirm the absence of contacts here; in all other positions, in which base and phosphate contacts of NF I were detected (around T₂₅, A₂₉, A₃₂, T₃₇; see ref. (14)), also sugar contacts have been revealed (this work). In addition, no contact could be observed at residue A₂₃ of the complementary strand; since it is located symmetrically to residue T₃₉, such a contact would be predicted and expected due to the dimeric nature of the NF I protein (11). Therefore we believe instead, that the observed altered reactivity of the entire region A₃₈–A₄₅, exposing the backbone sugars of residues A₃₈, T₃₉, and A₄₀ to hydroxyl radical attack, is due to a *structural* change induced indirectly by protein binding. This conformational change which we cannot ascribe in more detail as yet, is positioned symmetrically to the enhanced bend on the 5' border.

ACKNOWLEDGMENTS

We would like to thank H. Hartl and R. Föckler for technical assistance, M. Ihmsen and C. Mosch for typing parts of the manuscript, and H. Ibelgaufs for critically reading the manuscript. We greatly acknowledge the ideal support and critical interest of A. Arthur, B. Brenig, M. Müller, U. Müller, and the practical help of H. Domdey and E. Pöschl. The scans of the autoradiographs have been made under the instructions of W. Walter (LKB). This work was supported by the Deutsche Forschungsgemeinschaft (Forschergruppe Virus/Zell Wechselwirkungen, Fa 138/3-3).

REFERENCES

1. Nagata, K., Guggenheimer, R. A., Enomoto, T., Lichy, J.H. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA **79**, 6438–6442.
2. Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1983) Proc. Natl. Acad. Sci. USA **80**, 6177–6181.
3. Gander, I., Foeckler, R., Rogge, L., Meisterernst, M., Schneider, R., Mertz, R., Lottspeich, F. and Winnacker, E.-L. (1988) Biochim. Biophys. Acta **951**, 411–418.
4. Paonessa, G., Gounari, F., Frank, R. and Cortese, R. (1988) EMBO J. **7**, 3115–3123.
5. Gil, G., Smith, J. R., Goldstein, J. R., Slaughter, C. A., Orth, K., Brown, M. S. and Osborne, T. F. (1988) Proc. Natl. Acad. Sci. USA **85**, 8963–8967.
6. Borgmeyer, U., Nowock, J. and Sippel, A. E. (1984) Nucl. Acids Res. **12**, 4295–4311.
7. Rupp, R. A. W. and Sippel, A. E. (1987) Nucl. Acids Res. **15**, 9707–9726.
8. Bradshaw, M. S., Tsai, M.-J. and O'Malley, B. W. (1988) J. Biol. Chem. **263**, 8485–8490.
9. Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T.J. and Tjian, R. (1987) Cell **48**, 79–89.
10. Santoro, C., Mermod, N., Andrews, P. C. and Tjian, R. (1988) Nature **334**, 218–224.
11. Meisterernst, M., Rogge, L., Foeckler, R., Karaghiosoff, M. and Winnacker, E.-L. (1989) Biochemistry, in the press.
12. Meisterernst, M., Gander, I., Rogge, L. and Winnacker, E.-L. (1988) Nucl. Acids Res. **16**, 4419–4435.
13. Tullius, T. D. and Dombroski, B. A. (1986) Proc. Natl. Acad. Sci. USA **83**, 5469–5473.

14. de Vries, E., van Driel, W., van den Heuvel, S. J. L. and van der Vliet, P. C. (1987) *EMBO J.* **6**, 161–168.
15. Leegwater, P. A. J., van Driel, W. and van der Vliet, P. C. (1985) *EMBO J.* **4**, 1515–1521.
16. Tullius, T. D., Dombrosky, B. A., Churchill, M. E. A. and Kam, L. (1987) *Methods Enzymol.* **155**, 537–558.
17. Tullius, T. D. (1987) *TIBS* **12**, 297–300.
18. Tullius, T. D. and Dombroski, B. A. (1985) *Science* **230**, 679–681.
19. Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R. and Seeman, N. C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4653–4656.
20. Wu, H.-M. and Crothers, D. M. (1984) *Nature* **308**, 509–513.
21. Burghoff, A. M. and Tullius, T. D. (1987) *Cell* **48**, 935–943.
22. Morgan, J. G., Courtois, G., Fourel, G., Chodosh, L. A., Campbell, L., Evans, E. and Crabtree, R. G. (1988) *Mol. Cell. Biol.* **8**, 2628–2637.
23. Gough, G. W. and Lilley, D. M. J. (1985) *Nature* **313**, 154–156.
24. Bossi, L. and Smith, D. M. (1984) *Cell* **39**, 643–652.
25. Marini, J. C., Effron, P. N., Goodman, T. C., Singleton, C. K., Wells, R. D., Wartell, R. M. and Englund, P. T. (1984) *J. Biol. Chem.* **259**, 8974–8979.
26. Diekmann, S. and Wang, J. C. (1985) *J. Mol. Biol.* **189**, 1–11.
27. Thomas, G. H. and Elgin, S. C. R. (1988) *EMBO J.* **7**, 2191–2201.
28. Marini, J. C., Levene, S. D., Crothers, D. M. and Englund, P. T. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7664–7668.
29. Marini, J. C., Levene, S. D., Crothers, D. M. and Englund, P. T. (1983) *CSHSQB* **47**, 279–283.
30. Widom, J. (1985) *BioEssays* **2**, 11–14.
31. Koo, H.-S., Wu, H.-M. and Crothers, D. M. (1986) *Nature* **320**, 501–506.
32. Trifonov, E. N. and Sussman, J. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3816–3820.
33. Arnott, S., Chandrasekaran, R., Hall, I. H. and Puigjaner, L. C. (1983) *Nucl. Acids Res.* **11**, 4141–4155.
34. Drew, H. R. and Travers, A. A. (1984) *Cell* **37**, 491–502.
35. Ryder, K., Silver, S., DeLucia, A. L., Fanning, E. and Tegtmeyer, P. (1986) *Cell* **44**, 719–725.
36. Kawamoto, T., Makino, K., Orita, S., Nakata, A. and Kakunaga, T. (1989) *Nucl. Acids Res.* **17**, 523–537.
37. Zahn, K. and Blattner, F. R. (1985) *Nature* **317**, 451–453.
38. Zahn, K. and Blattner, F. R. (1987) *Science* **236**, 416–422.
39. Snyder, M., Buchman, A. R. and Davis, R. W. (1986) *Nature* **324**, 87–89.
40. Williams, J. S., Eckdahl, T. T. and Anderson, J. N. (1988) *Mol. Cell. Biol.* **8**, 2763–2769.
41. Koepsel, R. R. and Khan, S. A. (1986) *Science* **233**, 1316–1318.
42. Mizuno, T. (1987) *Gene* **54**, 57–64.
43. Inokuchi, K., Nakayama, A. and Hishinuma, F. (1988) *Nucl. Acids Res.* **16**, 6693–6710.
44. de Vries, E., van Driel, W., Tromp, M., van Boom, J. and van der Vliet, P. C. (1985) *Nucl. Acids Res.* **13**, 4935–4952.
45. van der Vliet, P. C., Claessens, J., de Vries, E., Leegwater, P. A. J., Puijn, G. J. M., van Driel, W. and van Miltenburg, R. T. (1988) In Kelly, T. and Stillman, B. (eds.), *Cancer Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Vol. **6**, pp. 61–70.
46. Wilson, W. D., Wang, Y.-H., Krishnamoorthy, C. R. and Smith, J. C. (1985) *Biochemistry* **24**, 3991–3999.
47. Anderson, J. N. (1986) *Nucl. Acids. Res.* **14**, 8513–8533.
48. Puijn, G. J. M., van Miltenburg, R. T., Claessens, J. A. J. and van der Vliet, P. C. (1988) *J. Virology* **62**, 3092–3102.
49. Mead, D. A. and Kemper, B. (1988) In Rodriguez, R.L. and Denhardt, D.T. (eds.), *Vectors, a survey of molecular cloning vectors and their uses*. Butterworths, Stoneham, pp. 85–102.
50. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
51. Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–559.
52. Schneider, R., Gander, I., Müller, U., Mertz, R. and Winnacker, E.-L. (1986) *Nucl. Acids Res.* **14**, 1303–1317.
53. Laemmli, U. K. (1970) *Nature* **227**, 680–685.
54. Öfverstedt, L. G., Hammarström, K., Balgobin, N., Hjerten, S., Pettersson, U. and Chattopadhyaya, J. (1984) *Biochem. Biophys. Acta* **782**, 120–126.
55. Heinrich, P. (1986) *Guidelines for quick and simple plasmid sequencing*. Boehringer, Mannheim.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.