Volume 15 Number 14 1987 Nucleic Acids Research

Site-specific DNA binding of nuclear factor I: effect of the spacer region

R.M.Gronostajski

Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

Received May 29, 1987; Accepted June 15, 1987

ABSTRACT

Nuclear factor I (NFI) is a site-specific DNA binding protein required for the replication of adenovirus type 2 DNA in vitro and in vivo. To study sequence requirements for the interaction of NFI with DNA, we have measured the binding of the protein to a variety of synthetic sites. Binding sites for NFI (FIB sites) were previously shown to contain a consensus sequence composed of 2 motifs, TGG (Motif 1), and GCCAA (Motif 2), separated by a 6 or 7bp spacer region. To assess conserved sequences in the spacer region and flanking sequences which affect NFI binding, we have isolated clones from oligonucleotide libraries that contain the two motifs flanked by 3 degenerate nucleotides and separated by degenerate spacer regions of 6 or 7 nucleotides. With a 6bp spacer region, a strong bias exists for a C or A residue in the first position of the spacer. Sites with a 7bp spacer region contain a G and C or A residue at the first and second positions, respectively, of the spacer, but also possess conserved residues at other positions of the site.

INTRODUCTION

Proteins that bind to specific sites on DNA play an essential role in cellular DNA and RNA metabolism. These proteins have been shown to function in the control of RNA transcription (1-5), DNA replication (6-9) and recombination (10-13) in both prokaryotes and eukaryotes. We are studying the interaction of one such protein, nuclear factor I (NFI), with its binding sites in cellular and viral DNA (14-16). An analysis of NFI interaction with a variety of binding sites should aid in our understanding of its physiological role in animal cells.

NFI was first isolated and characterized as a protein from the nuclei of HeLa cells that is required for the efficient replication of adenovirus DNA in vitro (17). Subsequent studies demonstrated that NFI binds to a specific site in the adenovirus origin of replication (18) and that the presence of this binding site (FIB site) is required for optimal virus replication in vitro (19-21) and in vivo (22-23). Extensive mutagenesis of the adenovirus FIB site indicates that this interaction of NFI is necessary but not sufficient for the replication of viral DNA (24-26).

Binding sites for NFI also exist in the genomes of many eukaryotic cells (14,15,27-30) and viruses (30-33), and we estimate that ~60,000 sites are present in the human genome (14). The functional similarity of the cellular and viral sites has been demonstrated by replacement of the adenovirus FIB site by the FIB-2 site isolated from the human genome (34). The exchange of the human site for the viral site generates a chimeric origin of replication that requires NFI for optimal initiation of DNA synthesis.

NFI also appears to function in RNA synthesis in animal cells. The presence of a viral FIB site near the promoter region of the Hepatitus B virus S antigen gene stimulates transcription in vivo by ~10-fold (33). In addition, NFI has been shown to be very similar or identical to the "CAAT" box transcription factor (CTF) which stimulates transcription from a number of eukaryotic promoters (30,35-36). Thus, NFI and FIB sites appear to function in both replication and transcription, and may be analogous to systems identified in prokaryotes (37-40) and eukaryotic viruses (41-42) which activate replication via the stimulation of transcription.

Binding sites for NFI contain the consensus sequence $TGG(N)_{6-7}GCCAA$ where N is at least partially degenerate (15-16,27-30,33). Although this sequence is present in FIB sites, it is also present in DNA fragments that do not bind to NFI (15,32). We have recently shown that ~50% of DNA fragments of the general sequence $TGG(N)_6GCCAA$ bind to NFI while only ~4% of fragments containing $TGG(N)_7GCCAA$ interact with the protein (16). In this report, we have analysed in detail the sequence requirements of the spacer regions of these two types of binding sites.

MATERIALS AND METHODS

Synthetic Oligonucleotides

Oligonucleotides were produced with an Applied Biosystems Model 280A DNA synthesizer. Each oligonucleotide was purified by electrophoresis through a 20% acrylamide-8M urea gel (43) and excision of the appropriate band. The oligonucleotides used are shown in Figure 1. Duplex oligonucleotides were generated by hybridizing the 12bp primer and appropriate template DNA (FIB-N6 or FIB-N7) and elongating the primer by incubating with the large fragment of DNA polymerase I in the presence of all 4 deoxynucleoside triphosphates. This procedure is described in detail in reference 16.

Cloning of Degenerate Oligonucleotides

1 nmole of each duplex, degenerate, oligonucleotide mixture (FIB-N6 or FIB-N7) was incubated for 2 hours at 37°C in 20 μl reactions containing 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 10 units of restriction

endonuclease BamH1, and 12 units of restriction endonuclease SalI. The reaction was stopped by addition of 3 μl of a solution containing 0.16M EDTA, pH 8.0, 1 μg of tRNA and 0.1 μg of the large fragment of gel-purified BamH1/SalI-digested M13mp19 DNA. The reaction was then extracted once with 30 μl of phenol:chloroform (1:1), once with 30 μl of chloroform, and the DNA was precipitated by adding 3 μl of 3M sodium acetate, 75 μl of 100% ethanol and incubating on dry ice for 30 min. The pellet was collected by centrifugation at 12,000 X g for 10 min at 4°C, rinsed once with 200 μl 70% ethanol, dried for 10 min in a Speed-Vac and resuspended in 18 μl of ligation mix (50 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 5 mM DTT). Ligation was initiated by adding 1 μl of 20 mM ATP and 1 μl of T4 DNA ligase (1 unit/μl) and continued for 18 hr at 16°C. 5 μl of the ligated DNA was used to transform the JM109 strain of E. coli (44) by the method of Hanahan (45) and plaques were isolated after 6 hr of growth at 37°C. Both duplex and single-stranded DNA were isolated from liquid cultures of individual plaques as described (46). DNA sequencing was performed by the chain-termination method (47).

Preparation of Nuclear Factor I

Nuclear factor I (NFI) was purified from extracts of HeLa cell nuclei as previously described (17). Material purified through the denatured DNA-cellulose step was used throughout the study. One unit of NFI has been defined as the amount of protein required to bind one fmol of labelled DNA in a nitrocellulose filter binding assay (16).

Nitrocellulose-Filter Binding Assay

To assess the binding of labelled DNA fragments to NFI ~0.5 μg of duplex DNA from each clone was digested for 4-12 hr at 37°C with 4 units of ClaI in 12 μl reactions containing 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$ and 1 mM DTT. The digested DNA (11 μl) was 3'-end labelled by addition of 1.5 μl of 10 X Klenow reaction buffer (500 mM Tris-HCl, pH 7.4, 100 mM $MgCl_2$, 10 mM DTT), 1μl 200 mM dGTP, 1 μl of α^{32}P-dCTP (3000 Ci/mmol) and 1 μl of the large fragment of DNA polymerase I (1 unit/μl). The reaction was incubated at room temperature for 15 min, 10 μl of 50 mM EDTA was added and the DNA was extracted once with 30 μl of phenol:chloroform and once with 30 μl of chloroform. The DNA was precipitated by addition of 6 μl of 5M ammonium acetate and 90 μl of 100% ethanol, incubation at -70°C overnight and centrifugation at 12,000 X g for 10 min at 4°C. The pellet was rinsed with 100 μl of 70% ethanol, dried for 10 min in a Speed-Vac and resuspended in 100 μl of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). 1 μl of end-labelled DNA (1-2 fmol) was incubated for 20 min at 4°C in 50 μl reactions containing buffer B (25 mM sodium Hepes, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 4 mM DTT, 200

μg/ml BSA) in the presence or absence of 6 units of NFI. The reaction mixtures were applied to nitrocellulose filters (Millipore, 0.45μ HAWP), washed twice with 1 ml of buffer B lacking BSA and the DNA retained on the filters was eluted and analysed by electrophoresis on 0.8% agarose gels as previously described (14-16).

Enrichment of FIB-N7 Library

A plate lysate was made from the FIB-N7 library and was used to infect a liquid culture of JM109 by standard techniques (46). Double-stranded DNA was isolated by alkaline-lysis and centrifugation to equilibrium on cesium chloride-ethidium bromide gradients (47). Superhelical FIB-N7 DNA (0.8 μg) was incubated in 50 μl reactions in the presence or absence of 30 units of NFI and processed as described under "Nitrocellulose-Filter Binding Assay". After elution from the filters the DNA was used to transform JM109 and individual plaques were picked and expanded. The presence of 30 units of NFI in the binding assay produced an ~10-fold increase in the number of plaques generated from the filter eluates.

Statistical Methods

Frequency tables for the DNA sequences were generated using the CUTOUT and GETFREQ programs of Staden (48). χ^2 analysis was performed on a DEC VAX 11/780 computer using the CHISQUARE routine of the MINITAB program package (Minitab, Inc.).

RESULTS

Analysis of N6 Clones

Binding sites on DNA for nuclear factor I (NFI) contain the consensus sequence TGG$(N)_{6-7}$GCCAA (15,27-33). To determine what other sequence or structural features in addition to this consensus region are important for binding, we synthesized a library of molecules that contains the consensus sequence but is degenerate in the 6-7bp spacer region and for 3bp 5' and 3' to the consensus motifs (Fig. 1). We had previously shown that 40-60% of these DNA fragments containing a 6bp spacer region bound to NFI in vitro (16). To determine the sequence specificity for NFI binding to these sites, individual members of this library were cloned into M13mp19, and tested for binding to NFI. A subset of representative clones is shown in Fig. 2. Analysis of 41 clones indicated that they could be divided conveniently into three classes: 1) Those that bound tightly to NFI (Fig. 2, lanes 3,6,18 and 21), 2) Those that bound only weakly to NFI (Fig. 2, lanes 9,12,15,24,30,36 and 39, note the weak intensity of the bands) and 3) Clones that show no apparent binding to the protein (Fig. 2, lanes 27,33,42 and 45). The relative

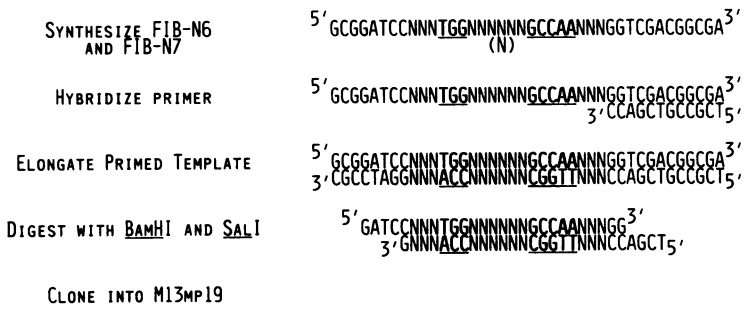

Fig. 1. Synthesis of Degenerate Oligonucleotide Libraries.
Two degenerate oligonucleotide libraries were synthesized using the FIB-N6 26bp DNA fragment, the FIB-N7 27bp DNA fragment and the 12bp primer shown in the figure. FIB-N6 consists of the following components listed in the 5' to 3' direction: an 8bp BamHI site (GCGGATCC), a 3bp degenerate sequence, a 3bp TGG motif, a 6bp degenerate spacer region, a 5bp GCCAA motif, a 3bp degenerate sequence, and a 12bp SalI site (GGTCGACGGCGA). FIB-N7 is identical except for an additional degenerate residue in the central spacer region. The 12bp primer is complementary to the 12bp SalI site on FIB-N6 and FIB-N7. The primer was hybridized to each oligonucleotide and elongated with the large subunit of *E. coli* DNA polymerase I as described (16). The degenerate libraries were constructed in M13mp19 as described in Materials and Methods.

binding ability of several of the clones was measured more precisely by a competition binding assay (15,16) which confirmed this relative ranking (data not shown). The individual clones were then sequenced and the DNA sequences and relative binding abilities are summarized in Figure 3. Several interesting features can be noted from these data: 1) All of the strong binders contain a C or A residue in the first position of the spacer region. This finding agrees with observations from several laboratories that FIB sites cloned from the DNA of several species contain a C or A residue at this position (15,27-33); 2) A preference for C or T residues at the first position 5' to the TGG motif in the strong binders suggests that this position may influence NFI binding to DNA; 3) Although a C or A is found in the first position of the spacer region in all of the strong binders, seven of the non-binders also share this attribute and yet do not interact detectably with NFI. These and other features are described in more detail in the Discussion section.

Analysis of N7 Clones

Some binding sites for NFI appear to contain a 7bp spacer region between the TGG and GCCAA motifs (see FIB-1 in Ref. 15). However, we had previously shown that only ~4% of molecules in an oligonucleotide library where the

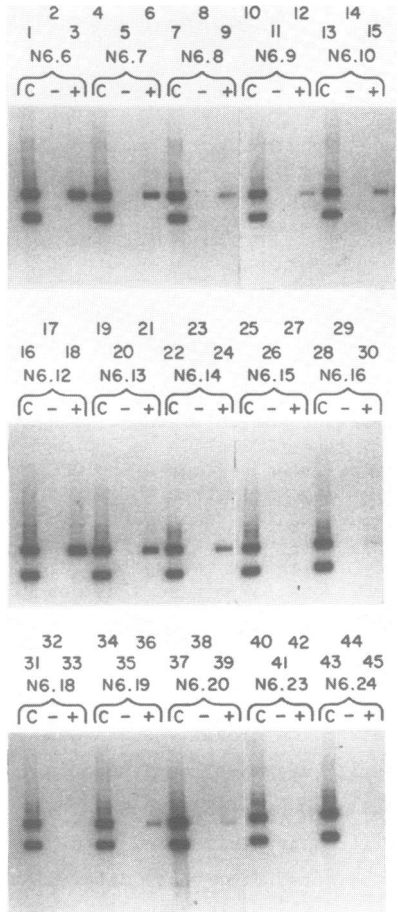

Fig. 2. Binding Assay on Clones from the FIB-N6 Library.
Superhelical DNA of clones from the FIB-N6 library was digested with ClaI, labeled with α^{32}P-dCTP and analysed for binding to NFI as described in Materials and Methods. The DNA was eluted from nitrocellulose filters and analysed by electrophoresis on a 0.8% agarose gel. Each clone is represented by 3 lanes C, -, and +; with the clone name shown above. Lanes marked C contain the total input DNA in the binding assay. Lanes marked - indicate DNA retained on nitrocellulose filters in the absence of NFI and lanes marked + indicate DNA retained in the presence of 6 units of NFI. Digestion with ClaI and end-labeling of the DNA produces 2 fragments; a 4.3kb fragment containing the cloned FIB site, and a 2.9kb fragment containing only M13 vector sequences. Note that in different clones, differing amounts of the 4.3kb fragment are retained on the filters in the presence of NFI. Data from these and similar analyses were used to assess the relative affinity of NFI for the cloned FIB sites.

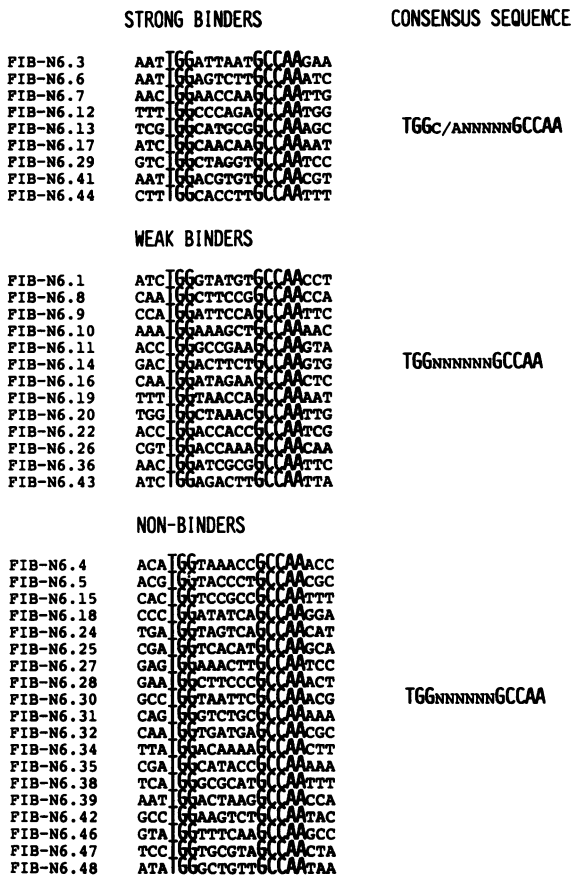

Fig. 3. Compilation of Binding Data for FIB-N6 Clones.

The clones analysed in Fig. 2 and other experiments were sequenced and grouped according to their binding to NFI. Strong binders were classified as those in which 20-60% of the input DNA was retained on nitrocellulose filters in the presence of NFI. Weak binders had < 20% of the DNA retained while non-binders had < 1% of DNA retained. Multiple exposures of autoradiograms were used to ensure that weak binding clones could be detected. The figure shows (left to right) the name of the clone, its sequence (between the BamHI and SalI sites) and the consensus sequences derived for each group. The nucleotides shown in large letters are those invariant bases in the library while the degenerate bases are shown in small letters. In the consensus sequences, N indicates that the 4 bases are used apparently at random.

motifs were separated by a 7bp degenerate spacer bound to NFI (16). Also, increasing the size of the spacer region of the human FIB site, FIB-2, by a single residue (from 6bp to 7bp) abolishes the tight binding of NFI to this DNA (16).

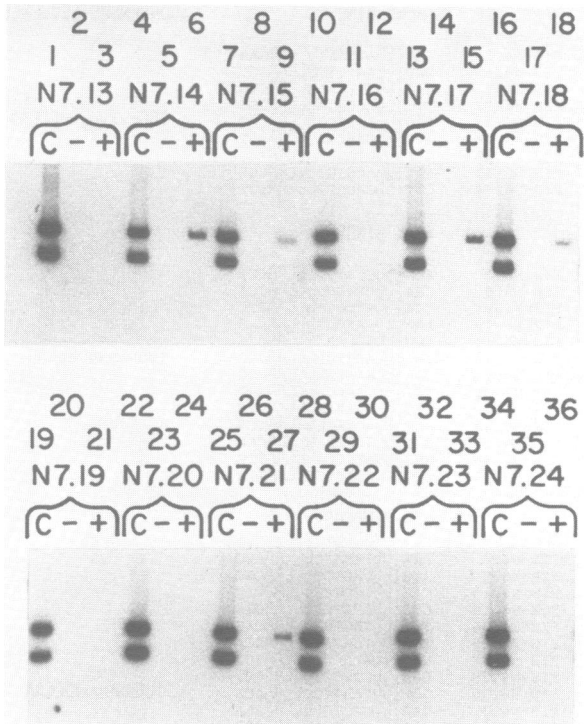

Fig. 4. Binding Assay on Clones from the FIB-N7 Library.
Superhelical DNA of clones from the FIB-N7 library was digested, labeled and analysed for binding to NFI as described in Materials and Methods. The figure is labeled and presented as described in the legend to Fig. 2. C, input DNA; -, DNA retained in the absence of NFI; +, DNA retained in the presence of 6 units of NFI. The data from this and additional experiments are summarized in Fig. 5.

It was therefore of interest to analyse several FIB sites that contain the TGG and GCCAA motifs separated by a 7 bp spacer region. This was achieved by cloning members of the aforementioned oligonucleotide library that contain the two motifs but are degenerate for 3bp 5' and 3' to the motifs and in the 7bp spacer region. To enrich for clones that bind NFI, double-stranded superhelical DNA from the cloned library was incubated with NFI, filtered through a nitrocellulose filter and the DNA-protein complexes retained on the filter were eluted and used to transform E. coli. Individual plaques from this transformation were then picked and double-stranded DNA from these clones was analysed for the binding to NFI (Fig. 4). Of 24 clones analysed 9 showed binding to NFI. This frequency of 37.5% is substantially higher than

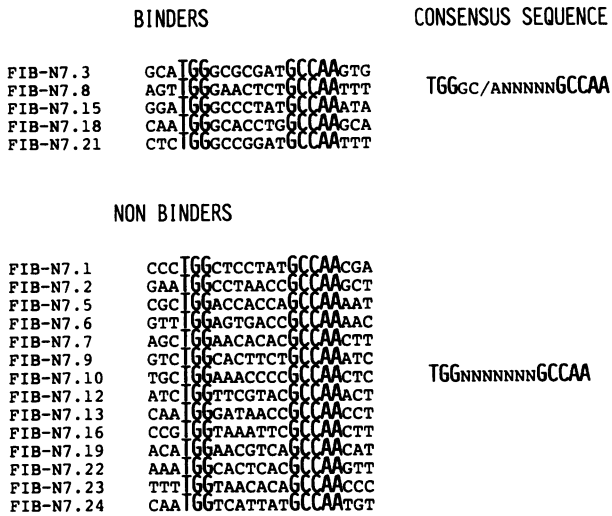

Fig. 5. Compilation of Binding Data for FIB-N7 Clones.
The clones analysed in Fig. 4 and other experiments were sequenced and grouped according to their binding to NFI. Binders had 1-20% of their DNA retained on nitrocellulose filters in the presence of NFI. These clones would be roughly equivalent to the weak binders shown in Figs. 2 and 3. Non-binders had < 1% of their DNA retained on filters in the presence of NFI. Nucleotides shown in large letters are those invariant in the library while small letters denote degenerate residues.

the ~4% of molecules in the uncloned oligonucleotide library that bind to NFI (16). Thus, it appears that the *in vitro* selection scheme produced about a 10-fold enrichment in clones that contain FIB sites. This level of enrichment is slightly less than the 20-fold increase obtained previously in the selection of binding sites from the human genome using a similar protocol (14).

The clones analysed appear to bind relatively weakly to NFI, as indicated by the low intensity of the DNA bands retained on nitrocellulose filters (Fig. 4). Sequence analysis of the clones demonstrated that all of the NFI-binding clones contained G and C or A residues in the first and second positions, respectively, of the 7bp spacer region (Fig. 5). Thus the consensus sequence TGGG$^{C}/_{A}$N$_5$GCCAA was produced (bold residues were apparently selected for in the degenerate spacer). Clone FIB-N7.3 was obtained four separate times from well separated plaques in the library, suggesting that it may represent a tighter binding clone than the other isolates. Since the 4 isolates of FIB-N7.3 were identical, they are treated as a single clone in

all subsequent analyses. In addition, clone FIB-N7.8 was obtained independently from two separate plaques, which indicates that we have probably isolated the majority of the binding sites present in this library.

DISCUSSION

The detailed interaction of NFI with the conserved residues of the $TGG(N)_{6-7}GCCAA$ consensus sequence have been well documented (15,49). The present study has generated information about potentially important interactions of NFI with residues outside of these conserved domains.

Analysis of the N6 Clones

Our previous studies using the uncloned degenerate FIB-N6 library indicated that ~50% of these molecules could interact with NFI (16). In our clonal analysis of this library, 22 of 43 members of the N6 library bound to NFI (51%) with 9 clones binding quite strongly and 13 binding with reduced affinity (Fig. 3). The distributions of bases at each position of the oligonucleotides for the Strong binders and the Non-binders is shown in Fig. 6. At each nucleotide position in the table (columns -9 to +10) the distributions of bases in the Strong Binders, Weak Binders, and Non-Binders were subjected to χ^2 analysis. At only 2 positions in the oligonucleotides (-7 and -3 of Fig. 6) was the distribution of bases between the 3 groups significantly different. At position -7, a comparison of the 3 groups gave a $\chi^2 = 14.42$ ($P<0.05$) which indicates that the distribution of bases are probably different. At position -3, $\chi^2 = 17.55$ ($P<0.01$) which strongly suggests that the distributions are different. Thus, the selection for a C or A residue at position -3 was the most significant deviation in the distribution of residues between the 3 classes. It should be noted that although such χ^2 analysis indicate significant deviations from randomness, the sample size of each population is quite small. We therefore place greater confidence in the most significant indication of the selection for C or A residues at position -3. Interestingly, three of the Weak Binders had G or T residues at this position of the site (Fig. 3, clones FIB-N6.1, -N6.11 and -N6.19). Analysis of other clones isolated from synthetic or cellular DNA should allow us to better assess if these residues are important for NFI - FIB site interaction.

Analysis of the N7 Clones

We had previously determined that only ~4% of DNA molecules containing the TGG and GCCAA motifs separated by a 7bp spacer region were bound by NFI (16). Therefore, we used *in vitro* selection with NFI to enrich for these clones in our library. The small number of clones obtained (five) preclude a

Frequency Table of Strong Binders

	N	N	N	T	G	G	C/A	N	N	N	N	N	G	C	C	A	A	N	N	N
P	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
T	2	4	5	9	0	0	0	2	3	1	2	5	0	0	0	0	0	4	3	3
C	1	1	3	0	0	0	5	2	3	4	1	0	0	9	9	0	0	1	1	3
A	5	4	0	0	0	0	4	4	2	2	3	3	0	0	0	9	9	3	2	1
G	1	0	1	0	9	9	0	1	1	2	3	1	9	0	0	0	0	1	3	2
N	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

Frequency Table of Non Binders

	N	N	N	T	G	G	N	N	N	N	N	N	G	C	C	A	A	N	N	N
P	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
T	4	3	1	19	0	0	9	4	5	6	4	6	0	0	0	0	0	5	4	5
C	6	7	5	0	0	0	2	6	4	6	8	6	0	19	19	0	0	6	7	6
A	4	6	10	0	0	0	5	7	7	4	5	6	0	0	0	19	19	5	5	7
G	5	3	3	0	19	19	3	2	3	3	2	1	19	0	0	0	0	3	3	1
N	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19

Fig. 6. Frequency Tables for FIB-N6 Strong Binders and Non-Binders.
Data from Fig. 3 were used to construct tables of nucleotide usage at each position of the FIB-N6 sites. The consensus sequence is given at the top of each table. The numbering of positions begins with the fourth base of the spacer region defined as 0 with positions to the left having negative values and those to the right having positive values. Thus, the TGG motif represents positions -6, -5 and -4, respectively, while the GCCAA motif represents positions 3,4,5,6 and 7, respectively. The symbols in the leftmost column are P, the position in the site; and T,C,A and G indicate the number of thymidyl, cytidyl, adenyl and guanyl residues, respectively, at each position. N indicates the total number of sequences used to compile the table. Underlined residues in the consensus sequences indicate invariant bases in the oligonucleotide library. Boxes are drawn around the invariant positions in the oligonucleotides and also around position -3 in strong binders to denote selection of C or A residues at this position. This position shows the largest deviation from randomness by χ^2 analysis of the FIB-N6 clones (see Discussion).

rigorous statistical analysis, however a striking similarity of the clones was noted. All of the clones that bind to NFI contain a G and C or A residues at the first and second positions, respectively, of the 7bp spacer region (Fig. 5). These same conserved residues are found in the 2 naturally

NAME	SEQUENCE
FIB-N7.3	GCA**TGGG**CGCGA**TGCCAA**GTG
FIB-N7.8	AGT**TGGG**AACTC**TGCCAA**TTT
FIB-N7.15	GGA**TGGG**CCCTA**TGCCAA**ATA
FIB-N7.18	CAA**TGGG**CACCT**GGCCAA**GCA
FIB-N7.21	CTC**TGGG**CCGGA**TGCCAA**TTT
FIB-1	ACC**TGGG**CACAG**TGCCAA**AGT
BS2	AGC**TGGG**CAGAT**TGCCAA**GGT
CONSENSUS	**TGGG**$^{C\ C\ G}_{A\ G\ T}$**GCCAA**

Fig. 7. Comparison of FIB Sites with 7bp Spacer Regions.
Clones from the FIB-N7 library that bind to NFI are aligned with two naturally occurring FIB sites that contain 7bp spacer regions. The large underlined letters are invariant bases in the oligonucleotide library while the large non-underlined letters indicate residues that are apparently selected for at the positions indicated. The overall consensus sequence for NFI binding sites containing a 7bp spacer region is shown at the bottom. FIB-1 is a site isolated from the human genome (15) and BS2 is a site located 5' to the chicken lysozyme gene (27).

occurring FIB sites that contain 7bp spacer regions (Fig. 7). In addition, a comparison of the synthetic and natural sites also suggests a selection for a C or G residue in the fourth position of the spacer and a T or G residue at the seventh position of the spacer (Fig. 7). Thus our previous prediction that only a highly restricted subclass of sites containing 7bp spacer regions can interact with NFI (16) is supported. Further analyses of these sites is needed to determine the requirement for these apparently conserved residues for the binding of NFI.

Overall Model of NFI Binding to DNA

The simplest model to explain the binding of NFI to FIB sites with 7bp spacer regions is that the observed selection for G and C or A residues at the first and second positions, respectively, in the spacer generates a binding site with GGG as the 1st motif of the site (i.e. $GGG^{C}/_{A}N_5GCCAA$ versus the usual $TGG^{C}/_{A}N_5GCCAA$). This explanation would suggest that the T in the 1st motif of these sites is not essential for binding activity.

Alternatively, the protein could be interacting with a "stretched" site of $TGGG^{C}/_{A}N_5GCCAA$ where the T in the first motif is important for binding. Site-directed mutagenesis of one or more of these clones may allow us to distinguish between these two models.

The importance of both of the two conserved motifs is directly addressed in these studies. It has recently been suggested that NFI is similar to or identical to a protein that binds to the "CAAT" box motif present in several eukaryotic promoters (26,30). Our findings would suggest that the previously determined "CAAT" motif (AGCCAAT, Ref. 30) is not sufficient for strong binding of NFI to these sites. Since the sequence GCCAA was surrounded by degenerate residues in our oligonucleotide library, and we selected only for sites that interacted with NFI, then if the AGCCAAT motif could bind strongly to NFI it should have been selected for. However, our results indicate that the selection for G and C or A residues at the first and second positions, respectively, of the spacer was much more frequent than the selection of A and T residues 5' and 3', respectively, to the GCCAA motif. Although these data may reflect some inherent bias of residues inserted during synthesis, we feel that it more likely represents a true requirement for both motifs ($TGG^{C}/_{A}$ and GCCAA) for the strong binding of NFI to DNA.

These results indicate several features of a FIB site that may be important in addition to the $TGG^{C}/_{A}$ and GCCAA motifs, for the binding of NFI. Clearly, the spacing between the motifs is critically important as previously indicated (16). Interestingly, when the spacing between the motifs is increased by 1bp, several other conserved residues appear to be selected for (Fig. 7). Further studies are needed to determine if these additional, apparently conserved, residues reflect true interactions of NFI with all of its binding sites, or rather aberrant interactions that only play a role in the protein binding to "stretched" sites that possess a longer than optimal spacer region.

Finally, the importance of the relative binding affinity of FIB sites to their physiological roles has yet to be addressed. NFI and its binding sites have been proposed to function in both adenovirus DNA replication (17-26) and cellular and viral RNA transcription (26,30,31,33). The relative affinity of NFI for its various binding sites differs by at least a factor of 10 (14,27-32). We have recently determined that both natural and synthetic FIB sites can stimulate RNA synthesis in vitro in chimeric promoter constructs (R.M. Gronostajski and N. Miyamoto, in preparation). We hope to use this assay system to assess the relationship between NFI binding to its sites and

this stimulation of transcription. Such studies, in conjunction with mutational analyses of the adenovirus binding site and the effect of NFI on viral replication (24-26), should generate important information on the physiological role of NFI in eukaryotic cells.

ACKNOWLEDGEMENTS

The author would like to thank Drs. N. Miyamoto, C. Campbell and S. Minkin for helpful discussions, and E. Giddens for excellent technical assistance. I also thank A. Jones for careful preparation of the manuscript. These studies were supported in part by grants from the Banting Research Foundation and the Medical Research Council of Canada.

REFERENCES

1. Ptashne, M. (1986) Nature (Lond.) **322**, 697-701.
2. McKnight, S. and Tjian, R. (1986) Cell **46**, 795-805.
3. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) Nature (Lond.) **319**, 154-158.
4. Moncollin, V., Miyamoto, N.G., Zheng, X.M. and Egly, J.M. (1986) EMBO J. **5**, 2577-2584.
5. Sive, H.L. and Roeder, R.G. (1986) Proc. Natl. Acad. Sci. USA **83**, 6382-6386.
6. Murakami, Y., Eki, T., Yamada, M.-A., Prives, C. and Hurwitz, J. (1986) Proc. Natl. Acad. Sci. USA **83**, 6347-6351.
7. Elias, P., O'Donnell, M.E., Mocarski, E.S. and Lehman, I.R. (1986) Proc. Natl. Acad. Sci. USA **83**, 6322-6326.
8. Vocke, C. and Bastia, D. (1983) Cell **35**, 495-502.
9. Fuller, R.S., Funnell, B.E. and Kornberg, A. (1984) Cell **38**, 889-900.
10. Gellert, M. and Nash, H. (1987) Nature (Lond.) **325**, 401-404.
11. Echols, H. (1986) Science **233**, 1050-1056.
12. Hoess, R.H. and Abremski, K. (1985) J. Mol. Biol. **181**, 351-362.
13. Gronostajski, R.M. and Sadowski, P.D. (1985) J. Biol. Chem. **260**, 12320-12327.
14. Gronostajski, R.M., Nagata, K. and Hurwitz, J. (1984) Proc. Natl. Acad. Sci. USA **81**, 4013-4017.
15. Gronostajski, R.M., Adhya, S., Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1985) Mol. Cell. Biol. **5**, 964-971.
16. Gronostajski, R.M. (1986) Nucl. Acids Res. **22**, 9117-9132.
17. Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J.H. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA **79**, 6438-6442.
18. Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1983) Proc. Natl. Acad. Sci. USA **80**, 6177-6181.
19. Guggenheimer, R.A., Stillman, B.W., Nagata, K., Tamanoi, F. and Hurwitz, J. (1984) Proc. Natl. Acad. Sci. USA **81**, 3069-3073.
20. Rawlins, D.R., Rosenfeld, P.J., Wides, R.J., Challberg, M.D. and Kelly, T.J. Jr. (1984) Cell **37**, 309-319.
21. Leegwater, P.A.J., van Driel, W. and van der Vliet, P.C. (1985) EMBO J. **4**, 1515-1521.
22. Hay, R.T. (1985) EMBO J. **4**, 421-426.
23. Wang, K. and Pearson, G.D. (1985) Nucl. Acids Res. **13**, 5173-5187.
24. de Vries, E., van Driel, W., Tromp, M., van Boom, J. and van der Vleit, P.C. (1985) Nucl. Acids. Res. **13**, 4935-4952.
25. Wides, R.J. Challberg, M.D., Rawlins, D.R. and Kelly, T.J. (1987) Mol. Cell. Biol. **7**, 864-874.

26. Rosenfeld, P.J., O'Neill, E.A., Wides, R.J. and Kelly, T.J. (1987) Mol. Cell. Biol. **7**, 875-886.
27. Borgmeyer, U., Nowock, J. and Sippel, A.E. (1984) Nucl. Acids Res. **10**, 4295-4311.
28. Siebenlist, U., Henninghausen, L., Battey, J. and Leder, P. (1984) Cell **37**, 381-391.
29. Henninghausen, L., Siebenlist, U., Danner, D., Leder, P., Rawlins, D., Rosenfeld, P. and Kelly, T.J. Jr. (1985) Nature (Lond.) **314**, 289-292.
30. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) Cell **48**, 79-89.
31. Nowock, J., Borgmeyer, U., Puschel, A.W., Rupp, R.A.W. and Sippel, A.E. (1985) Nucl. Acids. Res. **13**, 2045-2061.
32. Henninghausen, L. and Fleckenstein, B. (1986) EMBO J. **5**, 1367-1371.
33. Shaul, Y., Ben-Levy, R. and De-Medina, T. (1986) EMBO J. **5**, 1967-1971.
34. Adhya, S., Schneidman, P.S. and Hurwitz, J. (1986) J. Biol. Chem. **261**, 3339-3346.
35. Cohen, R.B., Sheffery, M. and Kinn, C.G. (1986) Mol. Cell. Biol. **6**, 821-832.
36. Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986) Cell **44**, 565-576.
37. Furth, M.E., Dove, W.F. and Meyer, B.S. (1982) J. Mol. Biol. **154**, 65-83.
38. Ogawa, T., Baker, T.A., van der Ende, A. and Kornberg, A. (1985) Proc. Natl. Acad. Sci. USA **82**, 3562-3566.
39. Minden, J.S. and Marians, K.J. (1985) J. Biol. Chem. **260**, 9316-9325.
40. Stuitje, A.R., de Wind, N., van der Spec, J.C., Pors, T.H. and Meijer, M. (1986) Nucl. Acids Res. **14**, 2334-2344.
41. Veldman, G.M., Lupton, S. and Kaman, R. (1985) Mol. Cell. Biol. **5**, 649-658.
42. Lusky, M. and Botchan, M.R. (1986) Proc. Natl. Acad. Sci. USA **83**, 3609-3613.
43. Maxam, A. and Gilbert, W. (1980) Meth. in Enzymol. **65**, 499-560.
44. Yanisch-Perron, C., Vieira, J. and J. Messing (1985) Gene **33**, 103-119.
45. Hanahan, D. (1983) J. Mol. Biol **166**, 557-580.
46. Zinder, N.D. and Broeke, J.D. (1982) Gene **19** 1-10.
47. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
48. Staden, R. (1986) Nucl. Acids Res. **14**, 217-231.
49. de Vries, E., van Driel, W., van den Heuvel, S.J.L. and van der Vliet, P.C. (1987) EMBO J. **6**, 161-168.