Volume 14 Number 22 1986 Nucleic Acids Research

Analysis of nuclear factor I binding to DNA using degenerate oligonucleotides

Richard M.Gronostajski

The Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M4X 1K9, Canada

Received 5 September 1986; Accepted 14 October 1986

ABSTRACT

Nuclear factor I (NFI) binds tightly to DNA containing the consensus sequence TGG$(N)_{6-7}$GCCAA. To study the role of the spacing between the TGG and GCCAA motifs, oligonucleotides homologous to the NFI binding site FIB-2 were synthesized and used for binding assays in vitro. The wild-type site (FIB-2.6) has a 6bp spacer region and binds tightly to NFI. When the size of this spacer was altered by ± 1 or 2bp the binding to NFI was abolished. To further assess the role of the spacer and bases flanking the motifs, two oligonucleotide libraries were synthesized. Each member of these libraries had intact TGG and GCCAA motifs, but the sequence of the spacer and the 3bp next to each motif was degenerate. The library with a 6bp spacer bound to NFI to 40-50% the level of FIB-2.6. The library with a 7bp spacer bound to NFI to only 4% the level of FIB-2.6 and some of this binding was weaker than that of FIB-2.6 DNA. This novel use of degenerate DNA libraries has shown that: 1) the structural requirements for FIB sites with a 7bp spacer are more stringent than for sites with a 6bp spacer and 2) a limited number of DNA structural features can prevent the binding of NFI to sites with intact motifs and a 6bp spacer region.

INTRODUCTION

Site-specific DNA binding proteins play a major role in the regulation of DNA metabolism in both prokaryotes and eukaryotes. Such proteins have been shown to be required for the transcription (1-4), replication (5-8) and recombination of DNA (9-11) in a variety of systems. Nuclear factor I (NFI) is a site-specific DNA binding protein, isolated from human HeLa cells, that is required for the efficient replication of adenovirus DNA in vitro (12-14) and in vivo (15-16). Analysis of the interaction of NFI with its binding sites on viral and host cell DNA should provide insight into its role in host cell and viral DNA metabolism.

NFI was first detected by its ability to stimulate the initiation of adenovirus DNA replication in vitro (12). The adenovirus genome is a 35,000 bp linear duplex DNA molecule with a virally-encoded 55,000 dalton terminal protein (TP) attached to the 5' end of each DNA chain (17-18). The initiation of replication occurs at either end of the genome by the covalent

attachment of dCMP to an 80,000-dalton precursor to this terminal protein (pTP) (19-20). The pTP-dCMP complex then acts as a primer for the synthesis of full length viral DNA (13,21). Optimal synthesis of the pTP-dCMP complex requires two other viral proteins, the adenovirus-encoded DNA polymerase (22-23) and the viral-encoded DNA binding protein (24). In addition to the viral proteins and DNA template, NFI is required for the efficient synthesis of the pTP-dCMP complex *in vitro*. NFI binds specifically to its target site within the adenovirus origin of replication present at each end of the genome (25). Several studies have correlated the binding of NFI to this site with its ability to stimulate the initiation reaction *in vitro* (26-27). The NFI binding site (FIB site) of adenovirus has also been shown to be essential for the *in vivo* replication of plasmids containing the viral origin of DNA replication (15-16).

Although the role of NFI in adenovirus DNA synthesis its well established, its function in host cell DNA metabolism is less well defined. Binding sites for NFI have been isolated from human DNA (28-31), and we have determined that a strong FIB site occurs about every 100 kilobases in the HeLa cell genome (28). FIB sites with different affinities for NFI have been identified near several eukaryotic genes (30-32) and in a number of other viral genomes (33-34). The number and location of FIB sites suggests some important role for NFI in DNA metabolism.

We (29) and others (30,34) have identified a consensus sequence common to the known FIB sites, $TGG(N)_{6-7}GCCAA$. Although this sequence is present within the known binding sites, it is also present on DNA molecules that do not bind to NFI (29,34). In this report, we have made novel use of libraries of partially degenerate oligonucleotides to examine the requirement of this consensus sequence for NFI-FIB site interaction.

MATERIALS AND METHODS

Synthetic Oligonucleotides

The synthetic oligonucleotides used in these studies were produced using an Applied Biosystems Model 280A DNA synthesizer. Each oligonucleotide was purified by excision of the appropriate size band from a 20% acrylamide-8M urea sequencing gel (35). The oligonucleotides are described in detail in the legend to Table 1. Primer-extension was performed by hybridization at 65°C for 30 min of 120 pmoles of the appropriate 12bp primer and 60 pmoles of the template DNA in a 10 μl reaction mixture containing 100 mM Tris-HCl, pH 7.4. To produce double-stranded DNA, an aliquot of the primed-template (30

pmoles of template) was incubated in a 20 µl reaction mixture containing 50 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 62.5 µM each of dATP, dGTP, dCTP and TTP, and 2 units of the large fragment of E. coli DNA polymerase I (Klenow fragment). The reactions were incubated at 23°C for 30-60 min, 10 µl of 50 mM EDTA was added and the reactions were extracted once with phenol: chloroform (1:1) and once with H_2O-saturated ether. Residual ether was removed by incubation at 37°C for 10 min and the repaired, double-stranded DNA was stored at 4°C until use. Non-repaired control DNA was prepared by incubation of the primer-template DNA under similar conditions except that the Klenow polymerase was omitted from the reaction. This control DNA was single-stranded except for the region of the template that hybridized to the primer. To generate ^{32}P-labelled duplex DNA, the reactions were performed as described above for double-stranded DNA except that the concentration of dCTP was reduced to 10 µM and 60-100 µCi of [α-^{32}P]dCTP (> 3000 Ci/mmole) was included in the reaction (final specific activity 6-10 µCi/µmole). Analysis of the products of the labelling reaction by electrophoresis on denaturing polyacrylamide gels (35) indicated that > 95% of the TCA-precipitable label was present in full-length DNA (not shown).

Preparation of Nuclear Factor I

NFI was purified from extracts of HeLa cell nuclei as previously described (12). Material purified through the denatured DNA-cellulose step was used unless otherwise noted. One unit of NFI activity is defined as the amount required to bind one fmole of FIB-site DNA in a nitrocellulose-filter binding assay. The units were determined by the number of fmoles of unlabelled FIB-site DNA required to reduce by 50% the binding of a labelled FIB-site DNA fragment. This method of calculating units of activity eliminates the error generated if the fraction of DNA-protein complex retained on the nitrocellulose filter is less than 100% at saturating levels of NFI (36).

Nitrocellulose-Filter Binding Assay

The retention of labelled DNA fragments on nitrocellulose filters was performed as previously described (28). All binding reactions contained 150 mM NaCl to repress non-specific binding. The amounts of DNA and NFI used are given in the legends of each Figure. In competition assays, both the labelled and unlabelled competitor DNA were present in the reaction mixtures prior to the addition of NFI.

Preparation of Plasmid DNA

pFIB-1 DNA contains the previously described 0.8kb FIB-1 site isolated from the HeLa cell genome cloned into the HindIII site of pBR322 (28-29).

All plasmid DNAs were propagated in the DH-1 strain of Esherichia coli (37) and were prepared by alkaline lysis followed by centrifugation to equilibrium in cesium chloride-ethidium bromide gradients (38). Plasmid DNAs used in the competition studies were in superhelical form.

RESULTS

Competition for Binding by Synthetic Oligonucleotides in vitro

NFI binds tightly and specifically to double-stranded DNAs containing cloned FIB sites (28). To determine if NFI interacts with FIB sites present on short duplex oligonucleotides, the ability of these synthetic oligonucleotides to compete with a ^{32}P-labelled FIB site was measured. Plasmid pFIB-1 contains a 0.8kb NFI binding site cloned into the HindIII site of pBR322 (28-29). Intact pFIB-1 DNA competes well with a 0.8kb ^{32}P-end labelled FIB-1 fragment for binding to NFI, while pBR322 DNA competes poorly (Fig. 1A). On the basis of mass of DNA, oligonucleotide FIB-2.6 competes

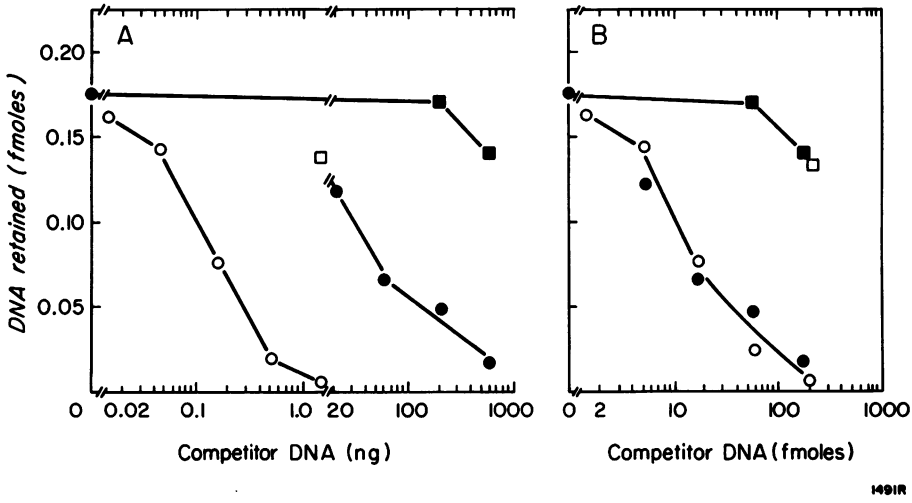

Fig. 1A and B. Competition for Binding to NFI by FIB-2.6
The filter binding competition assay was performed as described in Materials and Methods in 50 µl reactions containing 6 units of NFI, ≃ 1 fmole of a 0.8kb ^{32}P-end labelled fragment of FIB-1 DNA, (≃ 10000 cpm/fmole) and the indicated amounts of competitor DNA. In A, the X-axis is given in ng of competitor DNA while in B the X-axis is given in fmoles of competitor DNA, calculated by dividing the mass of DNA by its length in bp. The competitor DNAs are plasmid pFIB-1 DNA (-●-), duplex FIB-2.6 oligonucleotide (-O-), pBR322 DNA (-■-), and single-stranded, non-repaired FIB-2.6 oligonucleotide (-□-).

about 1000-fold better than pFIB-1 DNA for binding to NFI. When these data are expressed as moles of competitor DNA, FIB-2.6 and pFIB-1 bind to NFI with about equal efficiency (Fig. 1B). If the FIB-2.6 oligonucleotide is not repaired to form a duplex molecule, the partially single-stranded oligonucleotide does not compete effectively for binding to NFI. This result is expected since we previously demonstrated that NFI binds to FIB sites when the DNA is double-stranded but not when it is single-stranded (29). Because the previously cloned human NFI binding site FIB-1 had a 7bp spacer region, we made a 1bp insertion into the spacer region of FIB-2.6 to generate FIB-2.7 (Table 1). Surprisingly, this duplex oligonucleotide did not compete well for binding to NFI, and exhibited a level of competition similar to that of the control pBR322 or non-repaired FIB-2.6 DNA (Fig. 2). As a positive control in this same experiment, the oligonucleotide FIB-2.6 DNA was as effective as pFIB-1 DNA in binding to NFI.

Table 1. Sequences of Oligonucleotides Used in These Studies.*

Name	Length (bp)	Sequence
FIB-2.6	26	aggtc**TGG**ctttgg**GCCAA**gagccgc
FIB-2.6C2	26	aggtc**TcG**ctttgg**GCCAA**gagccgc
FIB-2.4	24	aggtc**TGG**ctttt**GCCAA**gagccgc
FIB-2.5	25	aggtc**TGG**cttgg**GCCAA**gagccgc
FIB-2.7	27	aggtc**TGG**ctttagg**GCCAA**gagccgc
FIB-2.8	28	aggtc**TGG**ctttatgg**GCCAA**gagccgc
FIB-2 primer	12	gcggctc**TTGGC**
FIB-N6	40	gcggatccnnnn**TGG**nnnnnn**GCCAA**nnnggtcgacggcga
FIB-N7	41	gcggatccnnnn**TGG**nnnnnnn**GCCAA**nnnggtcgacggcga
FIB-N primer	12	tcgccgtcgacc

*The oligonucleotides designated FIB-2.X (X = 4-8 or 6C2) are homologous to the previously described FIB-2 site cloned from HeLa cell DNA (29). FIB-2.6 is completely homologous to FIB-2 and contains a 6bp spacer region. In FIB-2.6C2 the G in position 2 of the TGG consensus motif is changed to a C. The TGG and GCCAA motifs are shown in bold capital letters. FIB-2.4, -2.5, -2.7 and -2.8 are homologous to FIB-2.6 but contain the indicated deletions or insertions to produce spacer regions of 4,5,7 or 8bp respectively. FIB-2 primer is a 12bp oligonucleotide which hybridizes to the 3' end of the FIB-2.X series DNAs and is used to prime synthesis of the second strand. FIB-N6 is a 40 mer containing TGG and GCCAA motifs but degenerate in the 6bp spacer region and for 3bp 5' and 3' of the motifs. FIB-N6 contains a BamHI site at its 5' end and SalI site at its 3' end. FIB-N7 is homologous to FIB-N6 but contains a 7bp degenerate spacer region. FIB-N primer is a 12bp primer which hybridizes to the 3' ends of FIB-N6 and FIB-N7 and is used to prime the synthesis of the second strand of these degenerate libraries.

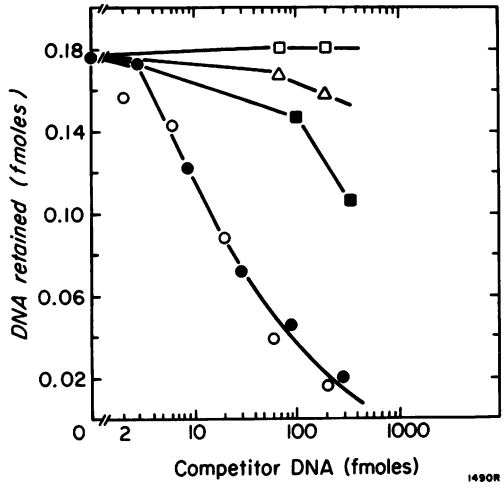

Fig. 2. Competition Assay Comparing FIB-2.6 and FIB-2.7 DNAs
A competition assay was performed as described in the legend to Figure 1 except that the competitor DNAs were pFIB-1 (-●-), duplex FIB-2.6 oligonucleotide (-O-), duplex FIB-2.7 oligonucleotide (-Δ-), single-stranded FIB-2.6 oligonucleotide (-□-) and pBR322 (-■-). The slightly greater competition seen here by pBR322 than by FIB-2.7 was not seen in other experiments.

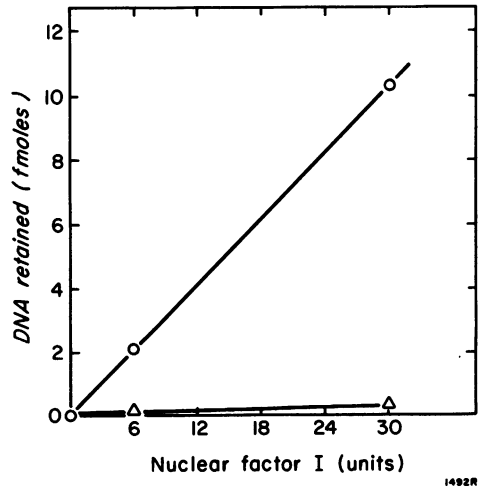

Fig. 3. Direct Filter Binding Assay of ^{32}P-Labelled Oligonucleotide
Nitrocellulose filter binding assays were performed in reactions containing 20 fmoles of the ^{32}P-labelled duplex oligonucleotides and the indicated amounts of NFI. ^{32}P-labelled FIB-2.6 (-O-), ^{32}P-labelled FIB-2.7 (-Δ-). The amount of FIB-2.7 DNA bound is equivalent to non-specific background binding seen in the assay (see Table 2 and Figure 4 for examples). The specific activity of the oligonucleotides was ≃ 350 cpm/fmole.

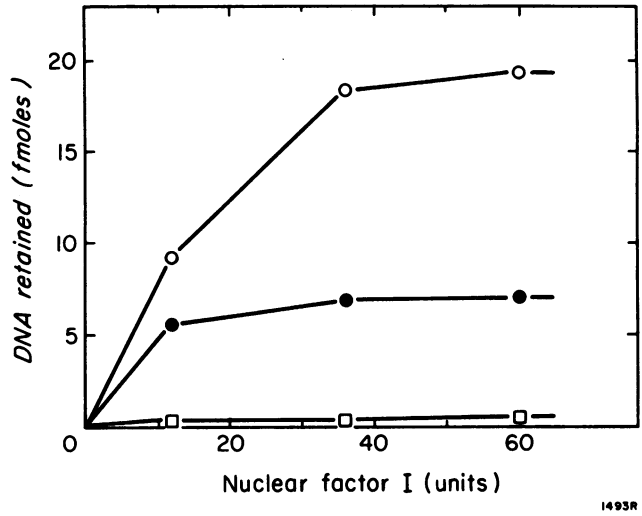

Fig. 4. Direct Filter Binding Assay of Oligonucleotide Library FIB-N6
Filter binding assays were performed using the indicated amounts of NFI and 20 fmoles of each ^{32}P-labelled DNA. FIB-2.6 oligonucleotide (-O-), FIB-N6 oligonucleotide library (-●-), ^{32}P-end labelled HindIII-digested pBR322 DNA (-□-). The specific activities of DNAs were ≃ 900 cpm/fmole for FIB-2.6 and FIB-N6 and ≃ 1000 cpm/fmole for pBR322.

Effect of "Spacer" Mutations on the Binding of NFI to Oligonucleotides

To further address the relative efficiency of FIB-2.6 and FIB-2.7 as binding sites, the oligonucleotides were labelled with [α-^{32}P]dCTP by primer-extension and the retention of the labelled molecules on nitrocellulose filters was measured. Labelled FIB-2.6 was retained on nitrocellulose filters in the presence of NFI while FIB-2.7 was not (Fig. 3). The difference in retention in this experiment was at least 30-fold. Thus, a single base insertion in the spacer region reduced the binding of NFI down to essentially background levels (see Fig. 4, and Table 2 for examples of background binding).

To assess further the effect of spacer size on FIB site binding, synthetic oligonucleotides homologous to FIB-2.6 but with spacer regions of 4,5 and 8bp were prepared, made duplex by repair in the presence of [α-^{32}P]dCTP and 3 unlabelled nucleotides, and tested for binding to NFI (Table 2). Of this series of oligonucleotides only FIB-2.6 was efficiently retained on nitrocellulose filters in the presence of NFI. As a negative control for binding a mutant of FIB-2.6 was synthesized (FIB-2.6C2) where the second nucleotide of the TGG motif was converted to a C (TCG, see Table 1). As expected, this point mutation within the consensus sequence eliminated binding of the DNA to

Table 2. Retention of ^{32}P-Labelled FIB-Sites by NFI.*

Labelled Oligonucleotide	Units of NFI	fmoles of DNA Retained
FIB-2.6	12	6.9
	24	15.2
FIB-2.4	12	<0.5
	24	<0.5
FIB-2.5	12	<0.5
	24	<0.5
FIB-2.7	12	<0.5
	24	<0.5
FIB-2.8	12	<0.5
	24	<0.5
FIB-2.6C2	12	<0.5
	24	<0.5

*Filter binding assays were performed in 50 μl reactions using 20 fmoles of each ^{32}P-labelled oligonucleotide and the indicated amounts of NFI. The specific activity of the ^{32}P-labelled oligonucleotides was $\simeq$ 1000 cpm/fmole. Background binding in the absence of NFI was about 0.1 fmole for each oligonucleotide.

NFI. Thus, the insertion or deletion of one or two base pairs in the spacer region of FIB-2.6 is as effective as a mutation in the consensus sequence in disrupting the binding of NFI.

Binding of Degenerate Oligonucleotide Libraries in vitro

We had previously suggested that some DNA sequence and/or structure in addition to the TGG(N)$_{6-7}$GCCAA consensus sequence is required for the binding of NFI. This suggestion was made due to the failure of the consensus sequence present in pBR322 DNA to function as a binding site for NFI (29). To examine how many additional structural features may be required for NFI binding to DNA, we synthesized a library of oligonucleotides containing intact TGG and GCCAA motifs but with complete degeneracy in the 6 base pair spacer region and for 3 base pairs 5' and 3' to the motifs (Table 1, FIB-N6). This library contains, in theory, 4^{12} or 16,777,216 different oligonucleotides, each possessing an intact consensus sequence. Thus, if the specificity of NFI binding is due only to the presence of the consensus sequence, all the members of this library should bind to NFI to the same extent as FIB-2.6 DNA. If, however, there are a large number of other

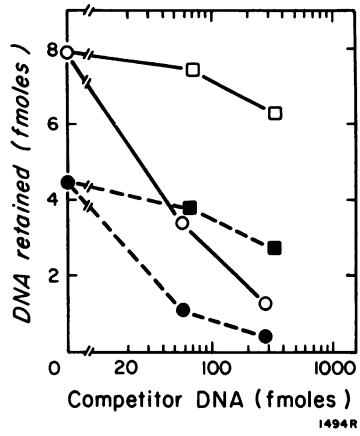

Fig. 5. Specificity of Binding of FIB-N6 DNA

Competition binding assays were performed in reactions containing 15 units of NFI, 20 fmoles of ^{32}P-labelled oligonucleotides, and the indicated amounts of unlabelled competitor DNA. The ^{32}P-labelled oligonucleotides were FIB-2.6 (-□-, -O-) and FIB-N6 (-■-, -●-). Competitor DNAs were pFIB-1 (-O-, -●-) and pBR322 (-□-, -■-). The specific activity of the ^{32}P-labelled oligonucleotides was ≃ 800 cpm/fmole.

structural features in addition to the consensus sequence that influence the binding of NFI, very few members of this library will be bound by the protein.

When ^{32}P-labelled FIB-N6 DNA was incubated with NFI and filtered, a substantial fraction of the labelled oligonucleotides were retained on the nitrocellulose filter (Fig. 4). At saturating levels of NFI, FIB-N6 DNA bound to NFI to about 40% the level of FIB-2.6 DNA. In the same experiment, pBR322 DNA was not retained by NFI (open squares). These data demonstrate that a large proportion of the members of this library can bind to NFI.

To determine whether this retention of FIB-N6 DNA on nitrocellulose filters was specific for NFI, a competition experiment was performed. Labelled FIB-2.6 and FIB-N6 DNAs were incubated separately with NFI in the absence and presence of unlabelled competitor DNA (Fig. 5). pFIB-1 DNA competed effectively with both FIB-N6 and FIB-2.6 DNA for binding to NFI while pBR322 DNA competed poorly. Thus the binding of FIB-N6 DNA to NFI appears specific and similar to the binding of FIB-2.6.

Since we had shown that a single base pair insertion into FIB-2.6 abolished its binding to NFI, we were interested in what fraction of FIB sites containing a seven base pair spacer region would bind to NFI. A partially degenerate oligonucleotide library identical to FIB-N6 but with a

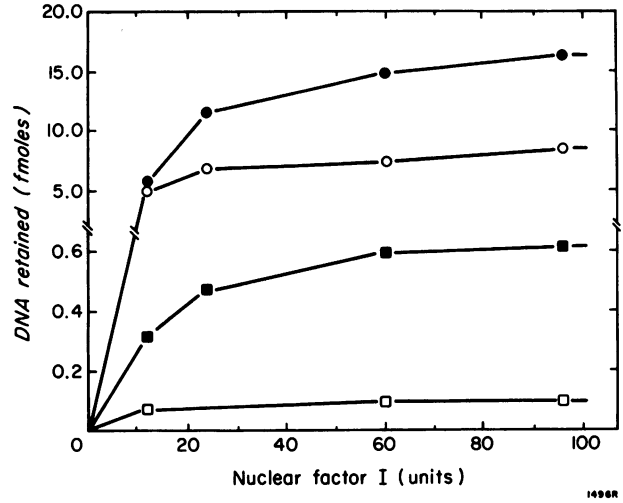

Fig. 6. Direct Filter Binding Assay of ^{32}P-Labelled FIB-N7 Library
Nitrocellulose filter binding assays were performed in reactions containing the indicated amounts of NFI and 20 fmoles of the ^{32}P-labelled oligonucleotides. FIB-2.6 (-●-), FIB-N6 (-○-), FIB-N7 (-■-) and FIB-2.6C2 (-□-). The specific activity of the oligonucleotides was ≃ 1000 cpm/fmole.

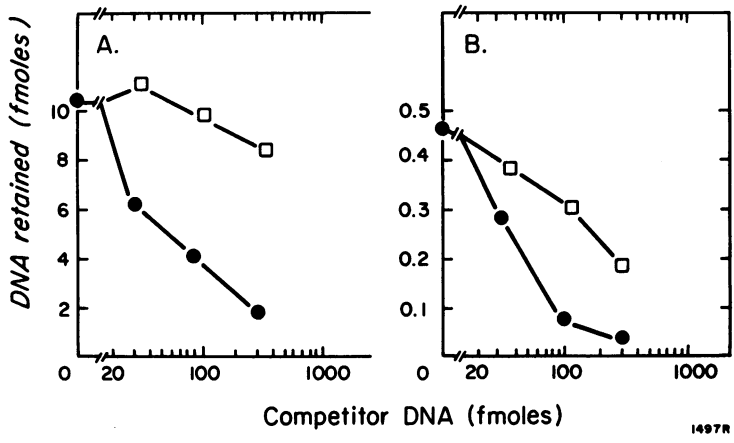

Fig. 7A and B. Specificity of NFI Binding to FIB-N7 DNA
Competition binding assays were performed in reactions containing 24 units of NFI, 20 fmoles of ^{32}P-labelled oligonucleotide and the indicated amounts of the competitor DNAs pFIB-1 (-●-) and pBR322 (-□-). Panel A, ^{32}P-labelled FIB-2.6 oligonucleotide. Panel B, ^{32}P-labelled FIB-N7 oligonucleotide library. The specific activity of the ^{32}P-labelled DNAs was ≃ 1000 cpm/fmole.

7bp spacer region (FIB-N7, Table 1) was synthesized and made double-stranded by primer-extension with [α-^{32}P]dCTP as described in Materials and Methods. ^{32}P-labelled FIB-N7 DNA was then compared with FIB-2.6, FIB-2.6C2 and FIB-N6 DNAs for binding to NFI (Fig. 6). The fraction of labelled FIB-N7 DNA retained on nitrocellulose filters increased with increasing amounts of NFI. The total amount of FIB-N7 DNA retained at saturating levels of NFI was at least 6-fold above the background level of binding exhibited by the consensus-sequence mutant FIB-2.6C2 DNA. However, the maximal percentage of FIB-N7 DNA bound was only about 4% of the amount of FIB-2.6 DNA retained (Fig. 6, note the break in the Y-axis). As expected, in the same experiment FIB-N6 DNA was retained to 40-50% the extent of FIB-2.6 DNA. Thus it appears that the frequency of FIB sites in the FIB-N7 library is an about order of magnitude lower than the frequency of sites in the FIB-N6 library.

To assess the specificity of NFI binding to FIB-N7 DNA, the ability of this binding to withstand competition by pBR322 and pFIB-1 DNA was measured (Fig. 7). As demonstrated previously, the binding of FIB-2.6 DNA was effectively competed by pFIB-1 DNA but not pBR322 DNA (Fig. 7A). The binding of FIB-N7 DNA was also effectively competed by pFIB-1 DNA, but in this instance pBR322 DNA also competed weakly for binding to NFI (Fig. 7B). These data demonstrate specific binding of FIB-N7 DNA to NFI, but suggest that the affinity of NFI for at least some of the binding sites present in FIB-N7 is somewhat weaker than for FIB-2.6. This is perhaps not surprising since several of the binding sites previously described in the human c-myc and chicken lysozyme genes are bound more weakly by NFI than are FIB-1 and FIB-2 (29,30,32).

DISCUSSION

Role of Consensus Motifs in Binding

These and other studies have demonstrated the importance of the TGG(N)$_{6-7}$GCCAA consensus sequence in the binding of NFI to DNA. For example, we (Table 2) and others (39-40) have shown that changes within the conserved motifs can reduce the binding of NFI. These results are not surprising since NFI has been shown to block the methylation of guanine residues within the motifs (29) and thus may make direct contacts with the DNA at these positions. However, the consensus sequence is not sufficient for the activity of a NFI binding site (29,34). The data presented here provide information concerning the number of parameters required, in addition to the consensus sequence, for NFI binding to DNA.

The best studied FIB sites are the ones present at nucleotides 25-38 within the origins of replication of adenovirus types 2 and 5. Initial studies using deletion mutants through this region of the viruses defined the requirement for the site in adenovirus replication *in vitro* (14,26-27) and *in vivo* (15-16). More recently, mutagenesis of the FIB site present in the origin of adenovirus type 2 has demonstrated directly the importance of the conserved TGG and GCCAA motifs in NFI binding and viral DNA replication (39). In addition, a variety of point mutants in the adenovirus type 5 FIB site show that the conserved motifs are essential for the tight binding of the NFI to this site (40).

Effect of Motif Spacing on Binding

The focus of the present study is on features required, in addition to the consensus motifs, for NFI binding to DNA. The data presented in Table 2 show clearly that altering the size of the spacer region by a single bp can disrupt the binding of NFI to the FIB-2 site. This finding is perhaps surprising since a FIB site cloned from the human genome (FIB-1) (29) and a site present upstream of the chicken lysozyme gene (29,32) have 7bp spacer regions.

One model that can reconcile these findings is that the nucleotide sequence in the spacer and/or flanking regions can determine whether a 6 or 7bp spacer region will permit the efficient binding of NFI to the site. It is possible, for instance, that the precise spacing between the TGG and GCCAA motifs is important for NFI binding and that particular combinations of bases yield 6 and 7bp chains of similar "apparent" lengths. A number of base sequence dependent features have been identified that can alter the apparent length of a DNA chain (41-44) including the "polyA" bend recently described in natural and synthetic DNAs (45-47). In addition, NFI itself may alter the spacing between the motifs during the process of binding to DNA, and some DNA sequences in the spacer region may promote or inhibit this process. Both of these examples provide mechanisms by which nucleotides that make no direct contact with NFI can still influence its binding to a FIB site. Such non-specific or "pseudo-specific" interactions may be important as a general mechanism for modulating the interaction of sequence-specific DNA binding proteins with their target sites.

The less than complete binding to NFI of the FIB-N6 and FIB-N7 library DNA has at least two possible explanations. One is that a subset of sequences from each library interacts strongly with NFI in a manner similar to FIB-2.6. An alternative is that some or all of the DNA sequences in the libraries

interact weakly with NFI. We have several lines of evidence against this latter possibility. For example, we (29) and others (34) have shown that some DNA fragments containing the consensus sequence do not bind detectably to NFI under the conditions used in the filter binding assay. Thus at least a fraction of each library must be deficient in binding ability. Also, the resistance of the binding to competition by pBR322 DNA (Fig. 6B) indicates an affinity within at least an order of magnitude of the affinity of NFI for FIB-2.6. Lastly, a very preliminary analysis of sites cloned from the FIB-N6 library indicates the presence of both strong binding sites and those with essentially no affinity for NFI (data not shown).

The finding that 40-50% of the degenerate FIB-N6 oligonucleotides bind tightly to NFI indicates that the conserved TGG and GCCAA motifs are primarily responsible for NFI-FIB site interaction. For example, this degree of binding would be consistent with only a single additional determinant, such as a purine or pyrimidine at a specified position, being required for NFI binding to DNA. However, it is also possible that certain combinations of bases at a number of positions outside the consensus motifs can influence binding, and the sum of these determinants permits 40-50% of the population to bind to NFI. These possibilities can be examined by sequence analysis of individual members of the library.

The low degree of binding of the FIB-N7 library of oligonucleotides to NFI suggests that the presence of a 7bp spacer region makes severe constraints on the other structural determinants of the FIB site. Since in both libraries the sequences flanking the motifs are degenerate, these data indicate that it is most likely the distance between the consensus motifs, rather than interactions in flanking sequences, that influence NFI binding to DNA. Since only a small subset of 7bp spacer regions allow NFI to interact with residues within the consensus sequence, it is somewhat surprising that sites with 7bp spacers have been identified in cellular DNA (29). One possibility is that some preference exists in the cell for sites with 7bp rather than 6bp spacer regions. Further work will be required to determine whether FIB sites with 7bp spacers and those with 6bp spacers are functionally equivalent in the genome. Also, the cloning and sequencing of the small subset of 7bp spacer regions that generate active FIB sites may shed light on the mechanism of binding of NFI to DNA.

Other Uses of Degenerate Libraries

Site-directed mutagenesis is a powerful tool in the analysis of the specificity of sequence-specific DNA binding proteins. The studies described

here demonstrate that in addition to the analysis of specific mutations, degenerate oligonucleotide libraries can be produced and used directly in vitro to yield information on structural requirements for binding sites on DNA. This technique should be particularly useful in addressing questions of the role of general features such as dyad symmetry (48) or A-T richness (49-50) in the interaction of proteins with DNA. Such partially degenerate oligonucleotide libraries can also be used to easily clone large numbers of variants in a binding site to permit a statistical analysis of the binding site. We hope to use such libraries in the future to examine questions about the binding of NFI to DNA that were not addressed in the current study such as 1) determinants of the relative affinity of NFI for different sites, 2) changes in DNA structure that occur as a consequence of NFI binding, 3) the correlation between binding and the activity of NFI in DNA replication and 4) the possible role of NFI in other aspects of cell metabolism, such as the regulation of RNA transcription.

ACKNOWLEDGEMENTS

We thank Drs. N. Miyamoto and C. Campbell for criticism of the manuscript and helpful discussions. We also thank Avril Jones for excellent secretarial assistance in preparing the manuscript. Special thanks go to Dr. K. Nagata who participated in the purification of NFI and to Dr. J. Hurwitz in whose laboratory NFI was first purified. R.M.G. is a scholar of the Medical Research Council of Canada. This work was supported in part by grants from the Banting Research Foundation and the Medical Research Council of Canada.

REFERENCES

1. Dynan, W.S. and Tjian, R. (1985) Nature **316**, 774-778.
2. Miyamoto, N.G., Moncollin, V., Egly, J.M. and Chambon, P. (1985) EMBO J. **4**, 3563-3570.
3. Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986) Cell **44**, 565-576.
4. Parker, C.S. and Topol, J. (1984) Cell **36**, 357-369.
5. Vocke, C. and Bastia, D. (1983) Cell **35**, 495-502.
6. Fuller, R.S., Funnell, B.E. and Kornberg, A. (1984) Cell **38**, 889-900.
7. Greenbaum, J.H. and Marians, K.J. (1984) J. Biol. Chem. **259**, 2594-2601.
8. Wobbe, C.R., Dean, F., Weissbach, L. and Hurwitz, J. (1985) Proc. Natl. Acad. Sci. USA **82**, 5710-5714.
9. Ross, W., Landy, A., Kikuchi, Y. and Nash, H. (1979) Cell **18**, 297-307.
10. Hoess, R.H. and Abremski, K. 1985 J. Mol. Biol. **181**, 351-362.
11. Gronostajski, R.M. and Sadowski, P.D. (1985) J. Biol. Chem. **260**, 12320-12327.
12. Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J.H. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA **79**, 6438-6442.
13. Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1983) Proc. Natl. Acad. Sci. USA **80**, 4266-4270.
14. Rawlins, D.R., Rosenfeld, P.J., Wides, R.J., Challberg, M.D. and Kelly, T.J. Jr. (1984) Cell **37**, 309-319.
15. Hay, R.T. (1985) EMBO J. **4**, 421-426.

16. Wang, K. and Pearson, G.D. (1985) Nucl. Acids Res. **13**, 5173-5187.
17. Challberg, M.D. and Kelly, T.J. Jr. (1982) Ann. Rev. Biochem. **51**, 901-934.
18. Stillman, B.W. (1983) Cell **35**, 7-9.
19. Challberg, M.D., Desiderio, S.V. and Kelly, T.J. Jr. (1980) Proc. Natl. Acad. Sci. USA **77**, 5105-5109.
20. Lichy, J.H., Horwitz, M.S. and Hurwitz, J. (1981). Proc. Natl. Acad. Sci. USA **78**, 2678-2682.
21. Friefeld, B.R., Lichy, J.H., Field, J., Gronostajski, R.M., Guggenheimer, R.A., Krevolin, M.D., Nagata, K., Hurwitz, J. and Horwitz, M.S. (1984) Curr. Topics in Microbiol. and Immunol. **110**, 221-255.
22. Lichy, J.H., Field, J., Horwitz, M.S. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA **79**, 5225-5229.
23. Stillman, B.W., Tamanoi, F. and Mathews, M.B. (1982) Cell **31**, 613-623.
24. van der Vleit, P.C., Zandberg, J. and Jansz, H.S. (1977) Virol. **80**, 98-110.
25. Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1983) Proc. Natl. Acad. Sci. USA **80**, 6177-6181.
26. Guggenheimer, R.A., Stillman, B.W., Nagata, K., Tamanoi, F. and Hurwitz, J. (1984) Proc. Natl. Acad. Sci. USA **81**, 3069-3073.
27. Leegwater, P.A.J., van Driel, W. and van der Vliet, P.C. (1985) EMBO J. **4**, 1515-1521.
28. Gronostajski, R.M., Nagata, K. and Hurwitz, J. (1984) Proc. Natl. Acad. Sci. USA **81**, 4013-4017.
29. Gronostajski, R.M., Adhya, S., Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1985) Mol. Cell. Biol. **5**, 964-971.
30. Siebenlist, U., Henninghausen, L., Battey, J. and Leder, P. (1984) Cell **37**, 381-391.
31. Henninghausen, L., Siebenlist, U., Danner, D., Leder, P., Rawlins, D., Rosenfeld, P. and Kelly, T.J. Jr. (1985) Nature (Lond.) **314**, 289-292.
32. Borgmeyer, U., Nowock, J. and Sippel, A.E. (1984) Nucl. Acids Res. **10**, 4295-4311.
33. Nowock, J., Borgmeyer, U., Puschel, A.W., Rupp, R.A.W. and Sippel, A.E. (1985) Nucl. Acids. Res. **13**, 2045-2061.
34. Henninghausen, L. and Fleckenstein, B. (1986) EMBO J. **5**, 1367-1371.
35. Maxam, A. and Gilbert, W. (1980) Meth. in Enzymol. **65**, 499-560.
36. Riggs, A.D., Suzuki, H. and Bourgeois, S. (1970) J. Mol. Biol. **48**, 67-83.
37. Hanahan, D. (1983) J. Mol. Biol **166**, 557-580.
38. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
39. de Vries, E., van Driel, W., Tromp, M., van Boom, J. and van der Vleit, P.C. (1985) Nucl. Acids. Res. **13**, 4935-4952.
40. Schneider, R., Gander, I., Muller, U., Mertz, R. and Winnacker, E.L. (1986) Nucl. Acids, Res. **14**, 1303-1317.
41. Calladine, C.R. (1982) J. Mol. Biol. **161**, 343-352.
42. Dickerson, R.E. (1983) J. Mol. Biol. **166**, 419-441.
43. Peck, L.J. and Wang, J.C. (1983) Proc. Natl. Acad. Sci. USA **80**, 6206-6210.
44. Nussinov, R., Shapiro, B., Lipkin, L.E. and Maizel, J.V. Jr. (1984) J. Mol. Biol. **177**, 591-607.
45. Koo, H.-S., Wu, H.-M. and Crothers, D.M. (1986) Nature **320**, 501-506.
46. Ulanovsky, L., Bodner, M., Trifonov, E.N. and Choder, M. (1986) Proc. Natl. Acad. Sci. USA **83**, 862-866.
47. Zahn, K. and Blattner, F.R. (1985) Nature **317**, 451-453.

48. Sadler, J.R., Sasmor, H. and Betz, J.L. (1983) Proc. Natl. Acad. Sci. USA **80**, 6785-6789.
49. Strauss, F. and Varshavsky, A. (1984) Cell **37**, 889-901.
50. Levinger, L.F. (1985) J. Biol. Chem. **260**, 14311-14318.