DNA Recognition by the NIT2 Nitrogen Regulatory Protein: Importance of the Number, Spacing, and Orientation of GATA Core Elements and Their Flanking Sequences upon NIT2 Binding†

Tso-Yu Chiang and George A. Marzluf*

Department of Biochemistry and the Program in Molecular, Cellular, and Developmental Biology, The Ohio State University, Columbus, Ohio 43210

Received September 2, 1993; Revised Manuscript Received November 5, 1993®

ABSTRACT: NIT2, a global positive-acting regulatory protein in *Neurospora crassa*, activates the expression of a series of unlinked structural genes in the nitrogen regulation circuit. NIT2 binding sites in the promoter region of the *nit-3*, *alc*, and *lao* genes are very different in sequence context except for the presence of at least two copies of a GATA core sequence. Changing a single nucleotide of only one of two closely spaced GATA core elements abolished NIT2 binding, demonstrating their importance for NIT2 binding. The effect of altering the number, orientation, or spacing of paired GATA elements and the importance of 5′- and 3′-flanking sequences upon NIT2 binding were examined. Strong binding sites for a NIT2–βGAL fusion protein appear to contain at least two GATA elements, which can have varied spacing but must be within a certain effective distance, approximately 30 bp, of each other. Surprisingly, the orientation of GATA elements and their flanking sequences have only modest effects upon NIT2 binding.

Neurospora crassa expresses a series of unlinked structural genes encoding various nitrogen catabolic enzymes which enable the utilization of secondary nitrogen sources, such as nitrate, purines, and amino acids, when favored nitrogen sources, such as ammonia or glutamate, are not available (Marzluf, 1981).

NIT2, a global positive-acting regulatory protein, activates the expression of this series of structural genes in the nitrogen regulatory circuit. NIT2 comprises 1036 amino acids with a molecular weight of 116 000 and possesses a single Cys-X_2-Cys-X_{17}-Cys-X_2-Cys type zinc finger motif that is responsible for sequence-specific DNA binding (Fu & Marzluf, 1990a). This zinc finger motif is highly homologous with the zinc fingers contained in transcription factors found in fungi, *Caenorhabditis elegans*, *Drosophila melanogaster*, birds, amphibians, and mammals (Orkin, 1992). DNA footprinting experiments have identified NIT2 binding sites upstream of three structural genes, *nit-3*, *alc*, and *lao*, which encode nitrate reductase, allantoicase, and L-amino-acid oxidase, respectively, all members of the *Neurospora* nitrogen regulatory circuit. These NIT2 binding sites have variable nucleotide sequences, except that they all possess two or more copies of a GATA core sequence. In addition to diverse flanking sequences, the number, orientation, and spacing of the GATA elements present in natural NIT2 binding sites also differ. In this work, we have examined the affinity of a NIT2–βGAL fusion protein for DNA fragments which contain two GATA elements separated by different numbers of bases. We also examined binding of the protein to DNA fragments which contained two equally spaced GATA elements which were oriented in either a head-to-tail or a head-to-head fashion or which had different flanking sequences. Fragments with one normal GATA element and one mutated GATA sequence, either GATC or TATA, were utilized to investigate whether a single GATA sequence element is sufficient for strong NIT2 binding. We present results which demonstrate that high-affinity binding sites for the NIT2–βGAL fusion protein contain at least two GATA core sequences, and these two elements must be within a certain distance, approximately 30-bp, of each other. Substitution of a single base in just one element of a pair of GATA sequences greatly reduced the binding of this protein. However, the orientation of the GATA elements as well as their flanking sequences appears to have only a modest effect on NIT2 binding.

MATERIALS AND METHODS

Plasmids and DNA Fragments. A series of DNA fragments with one GATA sequence element and a second GATA-related sequence element, with varied 5′-flanking sequences, were designed to investigate the effect of the orientation and number of GATA core elements and the influence of flanking sequences upon NIT2 binding ability. A 42-base single-stranded oligonucleotide, which contained one GATA and a second GATA-related element (G/TATA/C) and degenerate sequences (A/T and A/G) at the two bases immediately 5′ of both GATA elements, was rendered double stranded by annealing with a degenerate primer of sequences complementary to the 20 bases at the 3′ end of the template, which was extended by T4 DNA polymerase. The resulting synthesis mixture contained double-stranded DNA fragments with 64 different sequences, each with one GATA element and a second GATA-related element, which could be the correct sequence, GATA, the inverted sequence, TATC, or the mutated form TATA or GATC. Furthermore, the two bases 5′ of the GATA and GATA-related elements could be any combination of the four dinucleotides AA, AG, TA, or TG (Figure 1). This DNA mixture was cloned into the *Sma*I site of the Bluescript vector (Stratagene, La Jolla, CA) polylinker region and transformed into *Escherichia coli* strain XL1Blue. White colonies were picked from IPTG–XGAL–LB plates, and plasmid DNAs were extracted and screened by *Sty*I digestion, which rec-

† This research was supported by Grant GM 23367 from the National Institutes of Health.

* Corresponding author: Department of Biochemistry, The Ohio State University, 484 West 12th Ave., Columbus, OH 43210.

® Abstract published in *Advance ACS Abstracts*, January 1, 1994.

0006-2960/94/0433-576$04.50/0

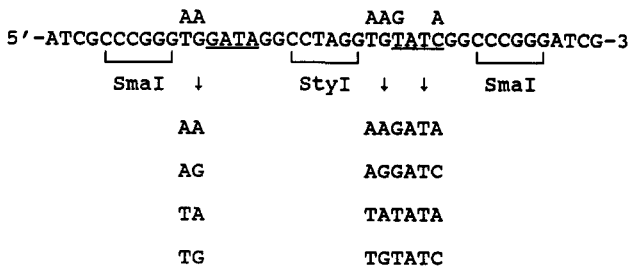

FIGURE 1: Degenerate oligonucleotides used to study the effect of changing GATA core and 5'-flanking sequences or the orientation of GATA elements on NIT2 binding. The GATA and GATA-related elements are underlined. G/T and A/C were introduced at the first and the fourth base of the GATA-related element, producing either normal GATA or TATC elements or mutated GATC or TATA sequences. A/T and A/G were introduced at the two bases 5' of each GATA or GATA-related element, so that the flanking sequences were a combination of any two of the following four dinucleotides: AA, AG, TA, or TG. The single-stranded oligonucleotides were converted to double-stranded form and cloned into Bluescript vector. *Sty*I does not cut Bluescript and was used to screen clones with an insert.

ognizes the 5' CCTAGG sequence in the insert (Figure 1). The precise sequence of positive clones was then determined by dideoxy sequencing. A total of 52 different DNA sequences out of the 64 possible were recovered. Each of these different fragments was excised from the plasmid by restriction at the *Pst*I and *Bam*HI sites of the Bluescript vector polylinker region, end labeled with ^{32}P by filling in with the Klenow fragment of DNA polymerase I, and analyzed for their NIT2–βGAL fusion protein affinity by gel-band mobility shift experiments.

A series of DNA fragments which contained either one GATA or two GATA elements separated by a different number of bases were prepared to examine the effect of spacing between GATA elements upon NIT2 binding. An 80-base single-stranded oligonucleotide, which contained two GATA sequence elements 45 bp apart, was made double stranded by annealing with a complementary 20-base primer, which was then extended by T4 DNA polymerase. The DNA fragment was then digested with *Kpn*I and *Not*I and cloned into the *Kpn*I and *Not*I sites of the Bluescript vector. A series of deletion clones was then obtained by site-directed mutagenesis (Kunkel, 1985) using primers designed to yield deletions of 40, 35, 30, 25, 15, and 5 bp between the two GATA elements of the original *Kpn*I–*Not*I clone. After mutagenesis, the potential deletion clones were first screened by restriction enzyme assays (Figure 2), and then their identities were verified by dideoxy sequencing. The spacing between GATA elements of the six deletion clones obtained is 5, 10, 15, 20, 30, or 40 bases (Figure 2). *Kpn*I–*Not*I fragments were then excised from these deletion clones, as well as from the original construct fragment with two GATA elements 45 bp apart; *Kpn*I–*Sal*I and *Nhe*I–*Not*I fragments, each with only a single GATA, and an *Xba*I–*Xho*I fragment between the two GATAs of the original clone (Figure 2) were also isolated. Each of these DNA fragments was end labeled with ^{32}P by filling in with the Klenow fragment of DNA polymerase I and tested for its affinity for the NIT2–βGAL fusion protein by gel-band mobility shift experiments.

Gel-Band Mobility Shift Experiments. The preparation of radioactive DNA fragments (probes) was described above. Expression and purification of the NIT2–βGAL fusion protein were carried out as described previously (Fu & Marzluf, 1990b). The fusion protein contains 217 amino acid residues of the NIT2 zinc finger domain fused to a short amino-terminal sequence (13 residues) encoded by the expression vector, and at its carboxyl terminus the finger domain is fused to the entire βGAL protein (Fu & Marzluf, 1990b). Mobility shift experiments were conducted as described (Eisen et al., 1988). The ^{32}P-labeled DNA fragments were incubated with or without the NIT2–βGAL fusion protein at room temperature for 20 min in a buffer containing 12 mM Hepes (pH 7.9), 2 mM DTT, 3.5 mM $MgCl_2$, 50 mM KCl, 15% (vol/vol) glycerol, and 1 μg of poly(dI)·poly(dC) before being loaded onto a 4% native polyacrylamide gel. Gels were run in 0.25× TBE (1× TBE is 0.09 M Tris–borate buffer, pH 8.2, and 0.002 M EDTA) without recirculation and then dried and scanned with a βetascope 603 blot analyzer (Betagen, Boston, MA) to quantify the intensity of shifted bands in comparison with that of the bands representing free probe. The data is presented as the fraction of the probe which was retarded as a DNA–protein complex and represents a quantitative measure of the NIT2 protein binding ability for each of the DNA fragments being examined.

RESULTS

Effect of Variable Spacing between GATA Elements on NIT2 Binding. In native NIT2 binding sites, GATA core elements are invariably present in two or more copies, and the spacing between these GATA elements varies from 3 bp up to 33 bp. To investigate the effect of different spacings between GATA elements on NIT2 binding, a series of DNA fragments with two GATA elements separated by 45, 40, 30, 20, 15, 10,

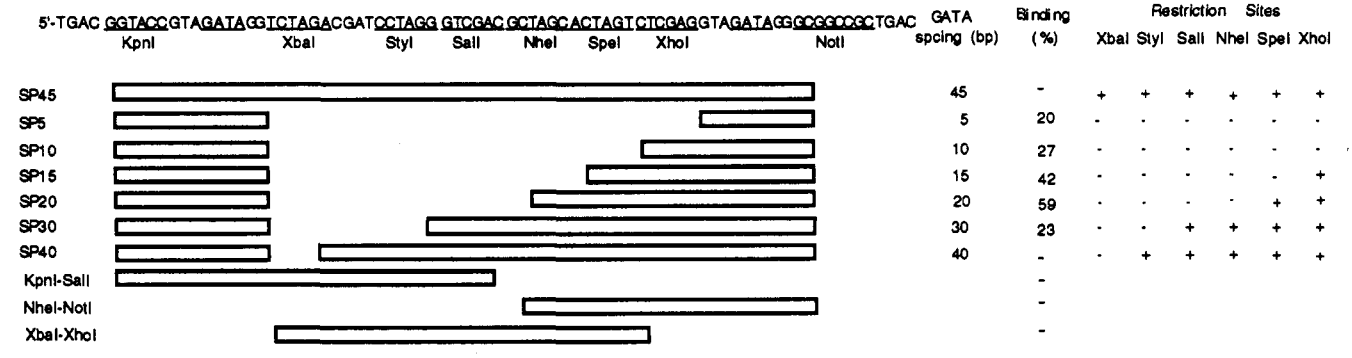

FIGURE 2: Series of DNA fragments used to study the effect of variable GATA element spacing upon NIT2 binding. The starting fragment, SP45, contains two GATA elements (underlined) 45 bp apart. Six deletion clones (derived from SP45), SP5, SP10, SP15, SP20, SP30, and SP40, which contain two GATA elements 5, 10, 15, 20, 30, and 40 bases apart, respectively, are shown. The *Kpn*I–*Sal*I fragment contains only the 5' GATA element, whereas the *Nhe*I–*Not*I fragment has only the 3' GATA element. The *Xba*I–*Xho*I fragment contains the DNA region between the GATA elements of the SP45 fragment. The deletion clones were checked by the loss of restriction sites as indicated. The quantitative results of the experiment shown in Figure 3 are also summarized. The percentage of the probe shifted was obtained by dividing the intensity of the band shift with the sum of the intensities of free and shifted probe for each fragment. –, trace only.

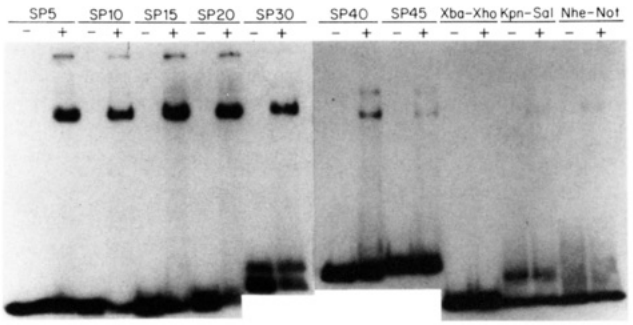

FIGURE 3: Gel-band mobility shift experiments of DNA fragments with different spacing between GATA elements. DNA fragments used for probes are indicated above each pair of lanes. Incubation of probes with or without NIT2–βGAL fusion protein is indicated by + or – above each lane. Refer to Figure 2 for DNA fragment nomenclature. SP20, whose GATA elements are separated by 20 bp, displayed the strongest shift. No detectable band shift was observed with the *Kpn*I–*Xho*I or *Xba*I–*Not*I fragments, which have only one GATA element, or with the *Xba*I–*Xho*I fragment, which represents the sequence between the GATA elements of SP45.

or 5 bp were tested for binding by the NIT2–βGAL fusion protein. DNA fragments with only a single GATA were also tested to determine whether one GATA is sufficent to form a NIT2 binding site. A DNA fragment lacking any GATA sequence served as a negative control. These various DNA fragments are described in Figure 2.

The results of gel-band mobility shift experiments are shown in Figure 3, and quantative results are summarized in Figure 2. Fragment SP20, with a spacing of 20 bp between GATAs, showed the highest affinity for the NIT2 fusion protein, i.e., it displayed the highest percentage of probe retarded as a DNA–protein complex (Figure 2). Other DNA fragments, which contained two GATA elements separated by 5–35 bp, were also bound strongly by NIT2. In contrast, when GATA elements were separated by 40 bp or more, binding by the NIT2 protein was almost totally eliminated (Figures 2 and 3). No significant gel-band shift was observed with fragments that contained a single GATA sequence, suggesting that a single GATA element is not sufficient for NIT2 binding. NIT2 did not bind to the negative control, a DNA fragment totally lacking any GATA sequence, as expected (Figures 2 and 3). These experiments revealed that the NIT2–βGAL fusion protein does not bind to DNA fragments that lack a GATA sequence or that contain a single GATA, or even two GATAs separated by 40 bp or more (Figure 2). These results suggest that at least two closely spaced GATA elements are required for the formation of a DNA–protein complex stable enough to be observed as a retarded band after gel electrophoresis.

Effect of GATA Element Orientation on NIT2 Binding. It was of interest to determine whether substitution of a single nucleotide in just one of the two paired GATA elements would affect NIT2 binding. Another important question to address was whether the 5′ sequences flanking the GATA elements or the orientation of the GATA elements with respect to each other would strongly affect the affinity of the NIT2 fusion protein for a binding site. A total of 52 different DNA fragments, some with two GATA elements either facing in the same direction (GATA–GATA) or toward each other (GATA–TATC), with different flanking sequences, and others with a substituted base in one of the two GATA sequences, were prepared as described in Materials and Methods. Each of these DNA fragments was tested for its affinity for the NIT2 protein by mobility shift experiments as shown in Figure 4 and summarized in Table 1.

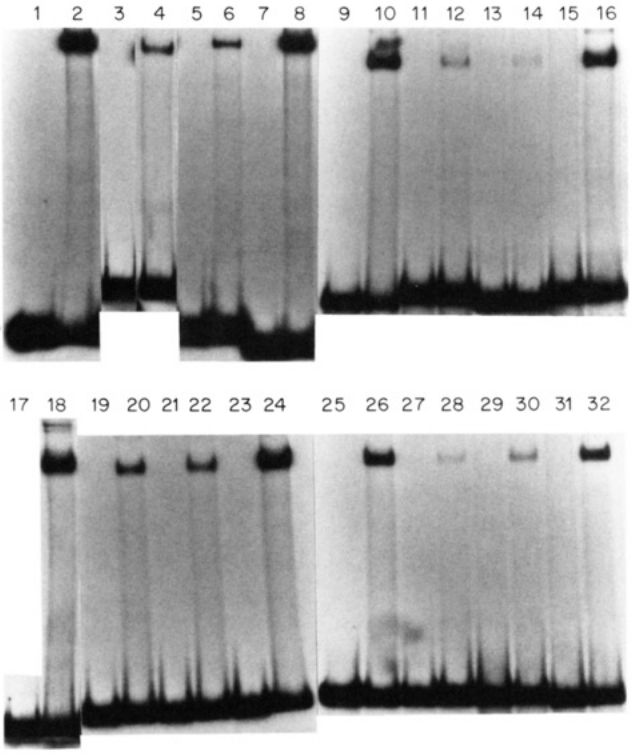

FIGURE 4: Selected gel-band mobility shift experiments of DNA fragments which have GATA elements with different orientations, different 5′-flanking sequences, or altered core sequences. DNA fragments in group II (lanes 1–8, with the 5′-flanking sequences TA and TG), group V (lanes 9–16, with the 5′-flanking sequences TG and TA), group IV (lanes 17–24, with 5′-flanking sequences TA and AG), and group VI (lanes 25–32, with 5′-flanking sequences TG and TG) were incubated with (all even-numbered lanes) or without (all odd-numbered lanes) NIT2–βGAL fusion protein, followed by gel electrophoresis.

In Table 1, the 52 different DNA sequences are divided into 16 groups in the order of decreasing NIT2 binding ability of the GATA–GATA fragment in each group. DNA fragments in each group contain the paired elements GATA–GATA, GATA–TATC, GATA–TATA, or GATA–GATC and have identical 5′-flanking sequences (Table 1). To examine the effect of orientation of two perfect GATA elements on NIT2 binding, we compared the NIT2 binding ability, i.e., the fraction of probe shifted, of GATA–GATA- and GATA–TATC-containing fragments in each group. In most cases, both the GATA–GATA and GATA–TATC fragments displayed good NIT2 protein binding, with differences in affinity of only approximately 2-fold or less; moreover, in some groups the fragment with the GATA–GATA-oriented elements showed stronger binding, whereas in other cases fragments with the inverted elements, GATA–TATC, showed a higher affinity for the protein. These results suggest that the orientation of paired GATA elements does not have a strong effect upon binding of the NIT2 fusion protein. However, in groups III and XII, the GATA–TATC fragment displayed noticeably low NIT2 binding ability, 1% and 10%, respectively, relative to the GATA–GATA fragment of the same group, presumably due to the effect of flanking sequences (Table 1).

Effect of an Altered GATA Core Sequence on NIT2 Binding. As revealed in Table 1, all GATA–GATC- or GATA–TATA-containing fragments lack signficant NIT2 binding ability. These results confirm the importance of the GATA core sequence and suggest that both elements of a pair must have a perfect GATA sequence. A potential binding site with one conserved GATA element and a second one with

Table 1: Summary of the Gel-Band Mobility Shift Analysis of 52 DNA Fragments That Have GATA Elements with Different Orientations or Different Flanking or Core Sequences[a]

group	DNA sequence				% of probe shifted
I	TA	GATA	TA	GATA	75
		GATA		GATC	3
		GATA		TATC	37
II	TA	GATA	AA	GATA	67
		GATA		GATC	2
III	AG	GATA	AA	GATA	62
		GATA		TATA	1
		GATA		GATC	0.5
		GATA		TATC	8
IV	TA	GATA	TG	GATA	56
		GATA		TATA	5
		GATA		GATC	2
		GATA		TATC	37
V	TG	GATA	TA	GATA	49
		GATA		TATA	2
		GATA		GATC	1
		GATA		TATC	30
VI	AA	GATA	TA	GATA	39
		GATA		GATC	1
VII	AG	GATA	TA	GATA	39
		GATA		TATA	1
		GATA		GATC	1
		GATA		TATC	30
VIII	TA	GATA	AG	GATA	37
		GATA		TATA	9
		GATA		GATC	8
		GATA		TATC	55
IX	AA	GATA	TG	GATA	37
		GATA		GATC	1
		GATA		TATC	32
X	AA	GATA	AG	GATA	27
		GATA		TATA	1
		GATA		GATC	2
XI	AA	GATA	AA	GATA	19
		GATA		TATC	43
XII	TG	GATA	AA	GATA	19
		GATA		TATA	0.2
		GATA		TATC	2
XIII	TG	GATA	TG	GATA	17
		GATA		TATA	1
		GATA		GATC	3
		GATA		TATC	21
XIV	AG	GATA	TG	GATA	17
		GATA		TATA	0.4
		GATA		GATC	0.5
		GATA		TATC	13
XV	TG	GATA	AG	GATA	2
		GATA		GATC	0.1
		GATA		TATC	5
XVI	AG	GATA	AG	TATA	0.2
		GATA		GATC	0.1
		GATA		TATC	5

[a] The percentage of probe shifted was obtained by dividing the intensity of the band shift by the sum of the intensities of the shifted and the free probe for each DNA fragment. The different groups are described in the text.

the sequence GATC or TATA is recognized very poorly, if at all, by the NIT2 protein. However, the sequence GATT may be able to partially replace a GATA element (see Discussion).

Effect of 5′-Flanking Sequence on NIT2 Binding. In Table 2, all GATA–GATA-containing fragments with different 5′-flanking sequences have been arranged in the order of decreasing affinity for NIT2 binding and separated into three groups. Group I contains five DNA fragments for which greater than approximately 50% of the probe was shifted; group II includes five DNA fragments that were shifted approximately 30–40%. Group III contains four DNA fragments with approximately 20% and a single fragment with only 2% of the probe shifted.

Table 2: Summary of Results of Mobility Shifts with DNA Fragments Containing Two GATA–GATA Elements with Varied 5′-Flanking Sequences[a]

	sequence					% of probe shifted
I	T A	GATA	T A	GATA		75
	T A		A A			67
	A G		A A			62
	T A		T G			56
	T G		T A			49
A	1 3		2 4			
T	4		3			
G	2		1			
II	A G		T A			39
	A A		T A			39
	T A		A G			37
	A A		T G			37
	A A		A G			27
A	4 4		2 2			
T	1		3			
G	1		3			
III	T G		A A			19
	A A		A A			19
	A G		T G			17
	T G		T G			17
	T G		A G			2
A	2 1		3 2			
T	3		2			
G	4		3			

[a] Results of experiments to determine the effect of 5′-flanking sequences of two GATA elements upon affinity for NIT2. Sequences are grouped according to their NIT2 affinity, and the frequency of appearance of each base at the two 5′ positions is summarized for each group.

The results shown in Table 2 show that variation in the two 5′ nucleotides flanking each of the GATA elements is compatible with significant NIT2 protein binding, although certain bases are preferred. Among the possible dinucleotide sequences in these 5′ positions which were examined (AA, AG, TA, and TG), TA appears to be preferred in a strong binding site, followed by AA, TG, and AG in decreasing order. In general, symmetrical pairs, e.g., TA–GATA AG–GATA and AG–GATA TA–GATA, show similar NIT2 binding affinity, although some discrepancies are found. With a single exception, the NIT2 binding affinities for all 15 DNA fragments with the various different 5′-flanking dinucleotides differed only by a factor of 4.

Effect of 3′-Flanking Sequence on NIT2 Binding. In Table 3, the GATA–TATC fragments with different 5′-flanking sequences are arranged in four groups in the order of decreasing NIT2 binding ability. Group I consists of two DNA fragments with over 40% probe being shifted, whereas group II contains five fragments which were shifted from 30 to 40%. It appears that the favored 5′-flanking sequence for the TATC element (inverted GATA sequence) is 5′-TA/G–TATC among TA, TG, AA, and AG dinucleotides (Figure 4). However, the 5′-flanking sequences of TATC can actually be viewed as the 3′-flanking nucleotides of GATA elements; thus T/CA would represent favored 3′-flanking dinucleotide sequences for the core GATA elements.

Table 3: Summary of Results with DNA Fragments Containing Head-to-Head GATA–TATC Elements To Determine the Effect of 5′- and 3′-Flanking Sequences upon Affinity for NIT2[a]

		sequence				% of probe shifted
I		T A	GATA	A G	TATC	55
		A A		A A		43
	A	1 2		2 1		
	T	1		0		
	G	0		1		
II		T A		T G		37
		T A		T A		37
		A A		T G		32
		T G		T A		30
		A G		T A		30
	A	2 3		0 3		
	T	3		5		
	G	2		2		
III		T G		T G		21
		A G		T G		13
	A	1 0		0 0		
	T	1		2		
	G	2		2		
IV		A G		A A		8
		A G		A G		5
		T G		A G		5
		T G		A A		2
	A	2 0		4 2		
	T	2		0		
	G	4		2		

[a] Results of experiments to determine the effect of flanking sequences of GATA–TATC elements upon affinity for NIT2. Sequences are grouped according to their NIT2 affinity, and the frequency of appearance of each base at the two flanking positions is summarized.

Table 4: Summary of Nucleotides That Occupy 5′- and 3′-Flanking Regions of GATA Elements of *Neurospora* NIT2 Binding Sites[a]

	−5	−4	−3	−2	−1	GATA	+1	+2	+3	+4	+5
A	6	5	5	2	8		3	5	2	3	5
T	3	3	5	11	5		3	4	4	3	4
C	5	3	4	3	5		5	4	7	6	3
G	3	7	5	3	1		7	5	5	3	3
cons	A/C	G/A	N	T	A		G/A	G/A	C	C	N

[a] These values record the frequency with which a nucleotide (A, T, C, or G) occupies a position in the 5′- or 3′-flanking sequence of a GATA core element in NIT2 binding sites of *Neurospora* structural genes *alc*(2), *lao*(4), *nit-3*(8), and *nit-6*(5). Cons, consensus sequence. The numbers in parentheses indicate the number of GATA elements for each gene. The *nit-6* sites have not yet been demonstrated to bind the NIT2 protein.

DISCUSSION

Binding sites for NIT2, the positive-acting regulatory protein in the nitrogen regulatory circuit in *Neurospora crassa*, have been identified in the 5′ promoter region of the *nit-3*, *alc*, and *lao* structural genes. Each NIT2 binding site contains two or more GATA sequence elements, yet the spacing between GATA elements, as well as their number and orientation, displays diversity. The 5′- and 3′-flanking sequences which surround the GATA elements in native NIT2 binding sites are different (Table 4), making it difficult to define a consensus flanking sequence. In the studies reported here, we found that DNA binding sites with different nucleotides in the two positions 5′ of GATA elements varied only approximately 4-fold in affinity for the NIT2 protein. Thus, it appears that the 5′- and 3′-flanking nucleotides play only a modest role in determining the affinity of a NIT2 binding site. However, our results suggest that the sequence TA is preferred at the two nucleotides immediately 5′ to GATA elements in high-affinity NIT2 binding sites, compared with either TG, AA, or AG. This experimental result agrees favorably with the characteristics of native NIT2 binding sites in the 5′ promoter region of *nit-3*, *alc*, *lao*, and *nit-6* genes (Exley et al., 1993), in which the consensus derived by examining 19 GATA elements is T at −2 (11/19) and A at −1 (8/19), relative to the GATA element (Table 4). The flanking sequences 3′ of GATA elements are also diverse (Table 4). In very recent studies of GATA binding proteins of mammals (Merika & Orkin, 1993) and chickens (Ko & Engel, 1993), diversity was found in the nucleotides 5′ and 3′ of the GATA recognition elements within strong binding sites. GATA-1 has a higher affinity for the sequence A–GATA–AAC than for T–GATA–GCA, indicating that bases flanking the core element do contribute to the strength of binding sites for the mammalian GATA-1 protein (Omichinski et al., 1993).

It appears that the GATA core elements are an important feature of strong NIT2 binding sites. An NMR structure of a DNA complex with the DNA binding domain of GATA-1 has been recently reported (Omichinski et al., 1993). Amino acid residues within the GATA-1 DNA binding motif were found to make hydrophobic contacts or hydrogen bonds with seven bases in the major groove and one in the minor groove of DNA; three contacts are with bases of the GATA sequence, and three contacts are with bases of the TATC sequence of the complementary strand. An interaction with GATA-1 was observed with the base immediately 5′ to the GATA sequence and with the nucleotide immediately 5′ to the TATC sequence on the complementary strand (Omichinski et al., 1993).

DNA fragments with one normal GATA element and a second mutated element, either GATC or TATA, displayed poor NIT2 protein binding, indicating that at least two intact GATA elements are required to form a high-affinity binding site for the NIT2 fusion protein. However, DNA fragments with two copies of the sequence GATT appear to have approximately 50% NIT2 binding ability compared with fragments with two GATA elements (T.-Y. Chiang et al., unpublished results). No natural NIT2 binding sites have been observed to possess a GATT sequence element. Moreover, a promoter which contained two copies of a UAS element from 5′ of the *DAL5* gene with GATT elements failed to support transcription activation in yeast (Bysani et al., 1991).

Preliminary studies of GATA elements in a natural NIT2 binding site in the 5′ promoter region of *nit-3* have revealed that when any two of the three GATA elements located there were mutated (to TATA), in vitro binding of the NIT2 fusion protein was almost totally eliminated (T.-Y. Chiang and G. A. Marzluf, unpublished results). Although these findings suggest that at least two GATA elements are required for good NIT2 binding, we cannot exclude the possibility that single GATA elements with special features might act as strong NIT2 binding sites. Avian GATA-1, -2, and -3 proteins have two zinc fingers which are highly homologous to the single zinc finger found in NIT2, GLN3, and DAL80 of yeast and in AREA of *Aspergillus* (Fu & Marzluf, 1990a; Caddick et al., 1986). The GATA-2 and -3 proteins recognized DNA fragments with GATA, GATC, or GATT sequences, whereas the avian GATA-1 protein only recognized DNA fragments with GATA elements (Ko & Engel, 1993). Ko and Engel (1993) found that the highest-affinity sites detected for the avian GATA factors were double sites in an inverted orientation.

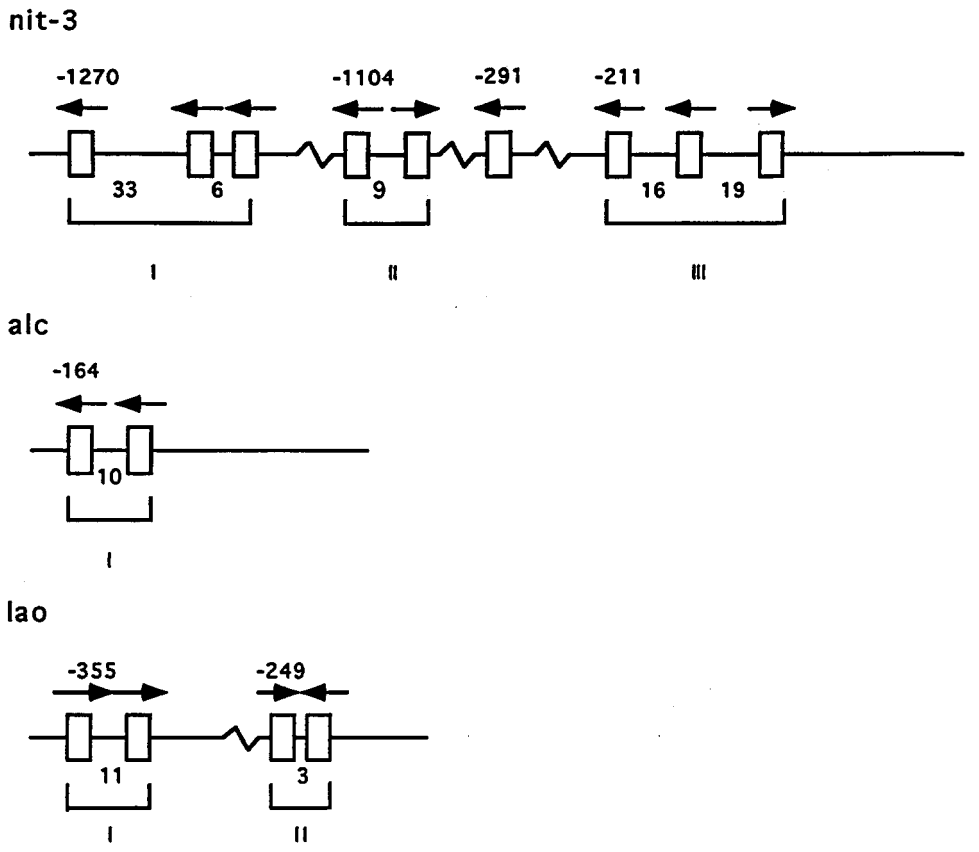

FIGURE 5: NIT2 binding sites in the promoter region of the *nit-3*, *alc*, and *lao* genes. Vertical bars represent GATA elements, and the arrows indicate their orientation (→, 5′-GATA-3′; ←, 5′-TATC-3′). The numbers above the arrowheads show the position of GATA elements relative to transcription start sites. The values between vertical bars indicate the spacing in base pairs between GATA elements. GATA sequences within brackets represent regions protected in DNA footprinting experiments. Three regions (I, II, and III) occur upstream of *nit-3*, two occur (I and II) for *lao*, and a single region occurs (I) for *alc*.

Variation is evident in the spacing between paired GATA elements in natural NIT2 binding sites. Hence, we carried out a systematic investigation of the effect that varied spacing between paired elements had upon the affinity of NIT2 binding. Surprisingly, the NIT2 fusion protein showed significant binding to DNA fragments containing a pair of GATA elements separated by 5, 10, 15, 20, or 30 bp, with an optimal spacing of 20 bp for strong NIT2 binding. In contrast, essentially no NIT2 binding was observed with DNA fragments containing two GATA elements separated by 40 or more bp; in fact, binding of the NIT2–βGAL fusion protein could barely be detected to fragments containing either a single GATA element or to those with two elements more distant than 40 bp from each other. Thus, it appears that the presence of at least two GATA elements within a certain critical distance of one another is required for effective binding of the NIT2 fusion protein. The spacing between GATA elements in native NIT2 binding sites in the promoters of the structural genes *alc*, *lao*, and *nit-3* and within the *nit-2* control gene itself ranges from 3 to 33 bp, but in no case is the interelement spacing greater than 40 bp, consistent with the experimental results obtained with the set of artificial constructs reported above. The proximal natural NIT2 binding site of the *nit-3* structural gene contains 3 GATA elements (Figure 5). When the central TATC element was mutated (to TATA), the interspace between the first and third GATA elements became 39 bp and protein binding was virtually eliminated (T.-Y. Chiang and G. A. Marzluf, unpublished results). Since the DNA helical structure makes a full turn every 10 bases, two GATA elements 5, 15, or 45 bp apart are on opposite sides of the helix, whereas two GATA elements 10, 20, 30, or 40 bp apart are on the same side of the helix. Our results indicate that when the spacing between GATA elements is within the 40-bp limit, their location on the same side or on opposite sides of the helix does not have a marked effect upon NIT2 binding.

One puzzling observation is that paired GATA elements in natural NIT2 binding sites occur in every possible orientation with respect to each other and the transcription start sites (Figure 5). We found that most DNA fragments with two core elements facing each other (GATA–TATC) possess about the same NIT2 protein binding ability as do otherwise identical fragments in which the elements both face in the same direction (GATA–GATA). Thus, the orientation of GATA elements does not appear to have a strong effect on NIT2 binding. All of the results reported here, as well as a simple examination of various natural NIT2 binding sites, imply that the NIT2 protein is very flexible in its ability to recognize its target sites in DNA. NIT2 binding sites clearly appear to consist of two or more GATA core elements, which however can occur in various orientations and in widely varied spacings up to a maximum critical distance of about 40 bp. Pertinent to these findings are the very recent observations that the yeast DAL80 protein, which has a zinc finger DNA binding domain homologous to that of NIT2, binds at paired GATAA sequences that must be separated by at least 15 bp and that can occur in a head-to-tail or a tail-to-tail orientation (Cunningham & Cooper, 1993). It is important to note that all of the studies reported here were carried out with a NIT2–βGAL fusion protein, and thus it is possible that the native NIT2 protein may show some differences in DNA binding specificity. With our increased understanding of the nature

of NIT2 binding sites, in vivo studies are now underway to relate NIT2 binding to its transcription activation and regulatory functions.

ACKNOWLEDGMENT

We thank Xiaodong Xiao, who provided some of the NIT2 fusion protein used in these studies.

REFERENCES

Bysani, N., Daughterty, J. R., & Cooper, T. C. (1991) *J. Bacteriol. 173*, 4977.

Caddick, M. X., Arst, H. N., Taylor, L. H., Johnson, R. I., & Brownlee, A. G. (1986) *EMBO J. 5*, 1087.

Cunningham, T. S., & Cooper, T. G. (1993) *J. Bacteriol. 175*, 5851.

Eisen, A., Taylor, W. E., Blumberg, H., & Young, E. T. (1988) *Mol. Cell. Biol. 8*, 4552.

Exley, G. E., Colandene, J. D., & Garrett, R. H. (1993) *J. Bacteriol. 175*, 2379.

Fu, Y. H., & Marzluf, G. A. (1990a) *Mol. Cell. Biol. 10*, 1056.

Fu, Y. H., & Marzluf, G. A. (1990b) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5331.

Ko, L. J., & Engel, J. D. (1993) *Mol. Cell. Biol. 13*, 4011.

Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488.

Marzluf, G. A. (1981) *Microbiol. Rev. 45*, 437.

Merika, M., & Orkin, S. H. (1993) *Mol. Cell. Biol. 13*, 3999.

Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., & Gronenborn, A. M. (1993) *Science 261*, 438.

Orkin, S. H. (1992) *Blood 80*, 575.