Volume 17 Number 14 1989 Nucleic Acids Research

Effects of all single base substitutions in the loop of *boxB* on antitermination of transcription by bacteriophage λ's N protein

Jed H.Doelling and Naomi C.Franklin

Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

Received March 21, 1989; Revised and Accepted June 9, 1989

ABSTRACT

The 'N' antitermination proteins of lambdoid bacteriophages are essential for overcoming multiple transcription terminators located within the major early operons of these phages (1). In order for N proteins to function, a genome sequence specifying N utilization, *nut*, must be located within an operon, between the promoter and the terminators (2). Two components have been identified within *nut*: 8-base *boxA*, conserved among different phages and implicated in the recognition of host NusA protein, required for N function (3); 15-base *boxB*, an interrupted palindrome (4), diverged in sequence among different lambdoid phages and hypothesized to be the site of recognition for different N proteins, also diverged in sequence (5). Here we apply a plasmid for testing termination and antitermination of transcription (6) to identify mutations at all positions in the 5−7 base loop of λ's *boxB*. Almost every base change at any position within the 5−7 base *boxB* loop was found to constrain antitermination of transcription by the N protein of bacteriophage λ. These observations extend previous mutational knowledge of *nut* (7) and are consistant with the hypothesis that the *boxB* loop is the direct site of recognition for N protein. Variations among the effects of different base changes suggest differential contacts between N protein and bases of the *boxB* loop, whether in DNA or RNA.

INTRODUCTION

Bacteriophage λ uses the unusual mechanism of transcription antitermination as one means of controlling gene expression (1). Because λ contains transcription termination signals at multiple locations in its genome, positioned between genes in both the left and right operons, its genome cannot be transcribed efficiently without a means of overcoming these stop signals.

When λ infects its host, *Escherichia coli*, host coded RNA polymerase recognizes the two early promoters *pL* and *pR*. Both the left and right early operons contain termination signals following the transcription of only one gene, *N* on the left and *cro* on the right. N protein, the product of gene *N*, is an essential component of λ's antitermination capability.

Mutants of λ have been found which no longer allow efficient transcription through terminators. In addition to recessive mutations within *N*, mutations blocking N function were located upstream of *N* and characterized as cis dominant (2). The DNA sequence of one such mutant revealed a single base change, located in the 'loop' between the two halves of an interrupted palindrome (4). A closely related palindrome was found between *cro* and terminator *tR1* in the early right operon (4), allowing the palindrome to be proposed as *nut*, a sequence required for N utilization (2). With increasing information about *nut* components affecting N function, the palindrome component of *nut* has become known as *boxB* and an 8-base upstream segment as *boxA* (1) (see Fig. 2).

All lambdoid phages tested, e.g. λ, Φ80, 21 and P22, have shown evidence of transcription antitermination (1, 5). The initial portions of the early operons of these phages show conserved genome arrangement despite very disparate nucleotide sequences (5, 8, 9). In particular, phages λ, 21 and P22 show interrupted palindromes at about the same distance downstream from the early promoters of left and right operons; these palindromes are very similar in the left and right operons of each phage, but differ markedly among the different phages (5). The upstream *boxA*, on the other hand, was found to be highly conserved in each of these positions (1,3,5). Because the N functions of λ, 21 and P22 do not normally complement each other, and because major differences in the amino acid sequences of their N proteins are accompanied by differences in nucleotide sequences of their *boxB*s, it is hypothesized that *boxB* may be a site of recognition for N (2,5,10). Mutations in *boxB* have focused attention on the 'loop' that interrupts its two symmetrical dyads: single mutations in the dyads were found to interfere with N function, but a double mutation that maintained base pairing between these stems did not (7); any change in the first position of the loop, however, effectively eliminated N function in vivo (2,7). These observations have prompted us to investigate the importance of each position of the *boxB* loop for the N antitermination function.

We approach this objective by an in vivo screen and assay of N function, based upon a plasmid contrived to reflect this function. The pTAT plasmids (6) allow termination and antitermination of transcription to be measured by the relative expression of *phoA* and *lacZ*, both reporter genes being expressed within a single operon but separated by a strong terminator (Fig. 1).

Between the promoter and the upstream reporter, *phoA*, are cloning sites usable for the introduction of a *nut* cassette, synthesized with near wild type sequence or with randomized bases in the *boxB* loop. When wild type *boxB* is introduced, as in pTAT6, and when N is supplied from a compatible plasmid, *lacZ* is expressed, causing colonies growing on XGal synthetic substrate to turn blue. When no *boxB* site is present, or when no N is supplied, *lacZ* is not expressed and colonies on XGal remain white. Defects in the *boxB* loop reduce or eliminate *lacZ* expression. The precise level of *lacZ* expression can be measured by assay of β-galactosidase, standardized by parallel assay of alkaline phosphatase synthesized from the same plasmid-carrying cells.

MATERIALS AND METHODS

Bacteria

The *E. coli* host strain used in these experiments, N567, has the following features : *sup*^{0} *trpE*(am)*9851* *laci* ^{q} Δ*lacZM15* Δ*phoA20* Tn*10*::Δ*recA* *thi* (6). Agar plates (5g NaCl, 5g tryptone and 15g agar per liter) were supplemented with 15mg/l chloramphenicol to select for the presence of pACYC184 plasmid derivatives, and 40mg/l ampicillin to select for pBR322 plasmid derivatives. XP (5-Bromo-4-Chloro-3-indolyl phosphate) at 0.08mM was added to screen for *phoA* expression. Alternatively, 0.17mM XGal (5-Bromo-4-Chloro-3-indolyl β-D-galactopyranoside) plus 0.025mM IPTG (isopropyl-β-D-thiogalactopyranoside) were added to provide an estimate of β-galactosidase activity.

Plasmids

Plasmids pAC*N* and pAC*ptacN* were derived from plasmid pACYC184(11) by cloning *N* or *ptacN* within the *Eco*RV site of the gene for tetracycline resistance (*Tc*). pTAT plasmids were derived from plasmid pBR322 for the in vivo testing of termination and antitermination (6). pTAT6 includes, in sequence, the strong IPTG-inducible promoter

ptac, cloning sites into which $nutL^+$ was introduced, intact *phoA* as upstream reporter, the strong transcription termination signal *t*1 from the end of *rrnB*, present in four copies, and translation-deficient *lacZ* as downstream reporter. Plasmid pTAT13 (Fig. 1) was derived from pTAT6 by replacing the *Hind*III-*Kpn*I fragment containing a prior cloning of $nutL^+$ with the *Hind*III-*Kpn*I portion of the pUC19 (12) cloning region. This introduces the two unique restriction targets, *Sal*I and *Bam*HI, usable for oriented insertion of oligonucleotide cassettes spanning *boxA*−*boxB* of λ*nut*.

Oligonucleotides

All deoxyribo-oligonucleotides were kindly provided by HHMI, synthesized by Diane Dunn and Bob Weiss on Applied Biosystems synthesizer #380B. Each was gel purified by electrophoresis on 5% acrylamide with TBM buffer (90mM Tris, 90mM H_3BO_3, 5mM $MgCl_2$). Two 20 mers were synthesized for use as primers in the sequencing reactions: the first is complementary to the template strand between the *ptac* promoter and the *nut* cloning region and the second (to enable sequencing in the opposite direction) is complementary to the coding strand at the extreme 5′ end of *phoA*.

Seven oligonucleotides of 50 bases each were designed with sequences based upon *nut* and its flanking regions in the left operon of λ (Fig. 2; discussed further below). Three oligonucleotides (J7, J8, J9) corresponding to the template strand, were designed to mutate the five bases in the *boxB* loop by scrambling to different degrees the four deoxynucleotides used in synthesizing all or part of the five positions corresponding to the *boxB* loop. A fourth oligonucleotide (J12), randomizes the two bases adjacent to the *boxB* loop. The fifth oligomer (J5), corresponding to the coding strand, is complementary to the other four and contains deoxyinosine (13) in each of the central 5 loop positions. Two additional 50-mers, J16 and J17, complementary to each other, were synthesized to test the effect of a transversion within the *boxB* stem, as discussed below. The oligomers were not phosphorylated, in order to avoid the possibility of multiple insertions into pTAT13. Cassettes were assembled by mixing at room temperature equimolar proportions of complementary oligonucleotides (each at about 500μg/ml). These mixtures were electrophoresed on 5% acrylamide gels with TBM buffer to purify the desired double stranded DNA.

All cassettes have cohesive ends matching the ends made by *Sal*I and *Bam*HI restriction endonucleases. They were ligated into pTAT13 pre-cut by these nucleases. Because the position adjacent to the *Bam*HI cohesive end in the cassette was made to be G rather than C, plasmids with the cassette insert are no longer cuttable by *Bam*HI. To increase the chances of finding pTAT*nut* plasmids, ligation mixes were digested with *Bam*HI, thereby selecting against parental plasmids with no *nut* insert. Ligation mixes were transformed into competent N567/pAC*N* and spread directly onto tryptone agar containing the indicator XP to assure the presence of *phoA*. This hardened agar top layer had been transferred immediately before the plating step onto a bottom agar layer containing the antibiotics ampicillin and chloramphenicol, ensuring that each drug resistant colony resulted from an independent transformation event (14). The initial screen for *phoA* was necessitated by the rather frequent deletion of *phoA*-*Bam*HI from pTAT13: about 50% of the ligated plasmids selected to have no *Bam*HI site were found to have a deletion spanning this segment. The basis for this deletion is under investigation.

$PhoA^+$ colonies were purified on plates containing ampicillin, chloramphenicol, XGal and the inducer IPTG. Single colonies exhibiting all levels of *lacZ* expression were cultured for further study. If the presence of the *nut* oligomer was indicated (by blue colony color

on XGal indicator plates or by restriction analysis), 40 ml cultures were grown to prepare plasmid DNA suitable for sequencing. This DNA was also used to transform N567, N567/pAC*N* and N567/pAC*ptacN* for use in assays of *phoA* and *lacZ* expression.

Plasmid DNA preparation

Plasmid DNA for restriction site analysis was prepared from 3 ml cultures (tryptone, yeast extract, ampicillin) by a modified alkaline lysis method (15).

To obtain sufficient double-stranded, supercoiled plasmid DNA suitable for sequencing, 40 ml cultures were grown (tryptone, yeast extract, ampicillin). Plasmid DNA from these cultures was purified by scaling up the boiling method, designed to yield supercoiled DNA (16). The cells from each culture were resuspended in 4 ml STET buffer (8% sucrose, 0.5% Triton X-100, 0.05M EDTA pH 8.0, and 0.05M Tris pH 8.0) and divided into 3 Eppendorf tubes. 70 μl of a fresh lysozyme solution (20mg/ml in 0.1M Tris pH 8.0) were added to each tube and mixed by repeated gentle inversions. The mixture was then boiled for one minute, centrifuged for ten minutes and the viscous pellet containing host DNA was removed with a toothpick. Approximately 400 μl of supernatant was mixed with 100 μl of 7.5M NH_4OAc and then with 0.5ml isopropanol. The solution was mixed gently but well and held at room temperature for 5 minutes before being centrifuged for 3 minutes; supernates were discarded. Each pellet was gently resuspended in 66 μl 0.05M Tris pH 8.0, 0.02M EDTA after which the contents of the three tubes were combined. The samples were centrifuged for 5 minutes to remove insoluble particles before digesting the supernatant with 80μl RNaseA (1mg/ml, boiled) for one hour at 37°C. Digests were extracted once with 0.1ml phenol/chloroform; DNA in the aqueous layer was mixed with NaOAc to 0.3M, precipitated twice with ethanol, washed with 70% ethanol and lyophilized. The DNA from each culture was resuspended in 40 μl H_2O, of which 10 μl were used for each set of sequencing reactions.

Sequencing

Double-stranded supercoiled plasmid DNA was sequenced (17) by Sanger's dideoxy method, using Sequenase™ and the procedure provided by United States Biochemical Corporation. One exception was to anneal the primer to the plasmid at 37°C. The 'dGTP' labelling mix was diluted 12 fold to enable the reading of sequence starting just 15–18 bases away from the primer; 0.17μl α-^{32}P dATP (Amersham #PB.10164) was used in each reaction. Following elongation and termination of the synthesis reaction, reaction mixes were electrophoresed on 8% acrylamide/urea denaturing gels. The gels were then autoradiographed to Kodak X-Omat RP film for 12–24 hours at −70°C with an intensifying screen. Care was taken to read the DNA sequence over an extended region, at least 10 base pairs (bp) upstream and 30 bp downstream from the *nut* oligomer.

Enzyme assays

Because phosphate inhibits the enzyme alkaline phosphatase, cultures to be assayed for both alkaline phosphatase and β-galactosidase were grown in 121 peptone medium supplemented with 0.5% glycerol as carbon source and ampicillin (40mg/l), as detailed previously (6). Chloramphenicol (15mg/l) was added when plasmid pAC*N* or pAC*ptacN* was present. Overnight cultures were diluted 50-fold into fresh 121 peptone medium and aerated for about 2 hours at 37°C to permit the cultures to enter logarithmic growth (OD600= 0.15). Cultures were then subdivided to assay for alkaline phosphatase and β-galactosidase, with and without induction. Cultures were induced or not by aerating at 37°C for an additional hour with or without 0.5mM IPTG. The cells were then permeablized

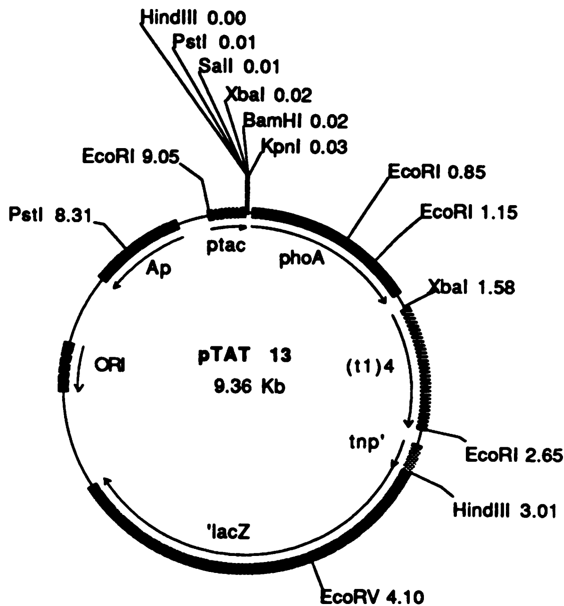

Figure 1. Genetic, physical and restriction target map of pTAT13. The map is to scale, with numbers giving distances in kilobases from the *Hin*dIII site in the cloning region. It was drawn with aid of the MacPlasmap computer program (25). This plasmid encodes a single major tester operon described more fully in reference 6: initiated from *ptac* inducible by IPTG, reading the upstream indicator *phoA*, encountering the tandem terminator $(t1)_4$ from the *rrn* operon and then reading the downstream indicator *lacZ*, whose translation start was crippled by fusion to a weakly translated segment, *tnp'*, from the *transposase* gene of TN10.

by vortexing with SDS/chloroform and assayed for alkaline phosphatase and β-galactosidase using the substrates Sigma 104 (*p*-nitrophenyl phosphate) and ONPG (*o*-nitrophenyl-β-D-galactopyranoside) respectively.

RESULTS

Rationale of the pTAT plasmids

N-mediated antitermination is dependent upon several factors that are provided by the bacteriophage λ and the *E. coli* host. λ supplies the N protein and the cis-required *nut* sequence while *E. coli* provides several proteins named Nus (N-utilization substance) (1, 3). The pTAT plasmids provide opportunity to investigate N-antitermination, because the *lacZ* gene of plasmid pTAT13 can only be expressed if transcription continues through the strong termination sequences between *phoA* and *lacZ* (Fig. 1). In vivo expression of the *lacZ* gene has been found to depend, just as does antitermination in λ, upon a *nut* sequence in cis and upon trans supplied phage N and *E. coli* Nus proteins (6). Here we report on the requirements of the *boxB* loop sequence of *nut* (Fig. 2) for antitermination in the presence of N^{λ} and all normal Nus factors of the *E. coli* host.

Plasmid pTAT13 was modified from pTAT6 to include the unique restriction targets *Sal*I and *Bam*HI. Its Lac^- phenotype in the absence of a *nut* sequence made it a good starting plasmid for the introduction of *nut* sequences to be tested.

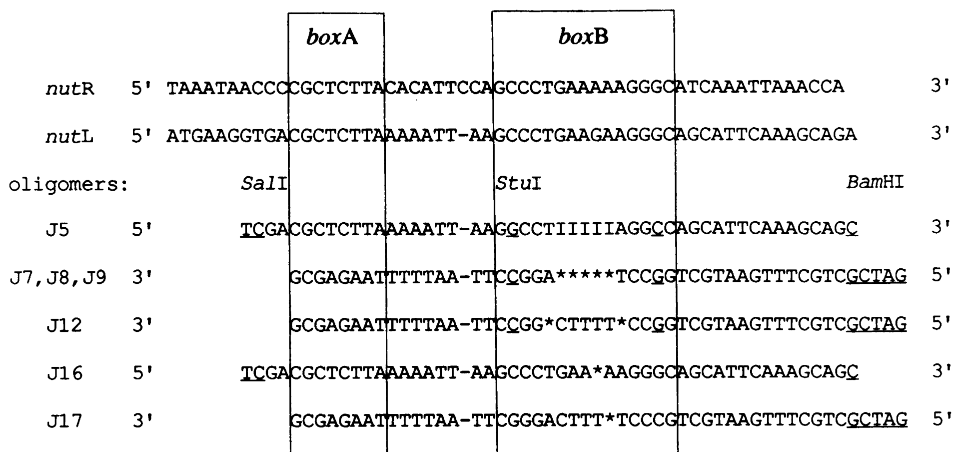

Figure 2. The DNA sequences of *nut*R, *nut*L and the synthetic oligonucleotides. Oligomer nucleotides that differ from *nut*L are underlined. Each star indicates a position where any one of two or more nucleotides is possible as explained in detail in the text.

Design of nut oligonucleotides

In order to introduce mutant *boxB* sequences into plasmid pTAT13, oligonucleotides spanning *nut* were designed to provide double-stranded mutant cassettes for unidirectional ligation into pTAT13. Plasmids obtaining these *nut* cassettes were named pTAT*nut*. The following features were included in the oligonucleotides (Fig. 2):

1) The nucleotide sequence of 50 mers was based upon the sequence spanning *nutL* of λ: 8-base *boxA*, 8-base interval and 15-base *boxB* . Upon insertion into pTAT13, *boxA* becomes located 44 bp downstream from the transcription start point, as compared with 33 bp in the early left operon of normal λ.

2) 50 mers with complementary sequences were designed to provide the cohesive ends of *Sal*1 and *Bam*H1 after hybridization. A single base substitution adjacent to the *Bam*HI cohesive end has the effect of eliminating the *Bam*HI restriction target following insertion into plasmid pTAT13. This provides a convenient direct selection for plasmids likely to have *nut* insertions.

3) A second pair of nucleotide alterations was introduced within the *boxB* palindrome changing a C-G stem pair to a G-C. This switch maintains stem stability while providing the unique *Stu*I restriction target between *boxA* and *boxB* (referred to as 'stu-stem'). Initial experience had shown about 10% spontaneous deletions in *boxA* of the pTAT*nut* plasmids, with or without the intended *boxB* changes. The introduction of the *Stu*l target allows the separation and recovery of such multiple changes.

4) The coding strand oligonucleotide (J5) was made to contain the unnatural nucleotide deoxyinosine in each of the 5 positions of the *boxB* loop. This allows limited base pairing with any of the four natural nucleotides of DNA (13) and prevents the J5 strand from contributing to any repair of the paired mutagenic oligomer. All *boxB* sequences recovered should therefore be derived from the sequence designated in the complementary oligonucleotide.

5) Three template strand oligonucleotides, the loop-mutagenic oligomers, were synthesized with different frequencies of incorrect bases at any or at specific positions in the five base *boxB* loop (designated '*' in Fig. 2). In J7 the frequency of incorrect bases was 15% at each position giving a 5% chance of finding any particular incorrect base at any particular position in the five base loop. In J9 the frequency of error was 50% at each position. In J8 the first, second and fifth positions were each given a 50% chance of being incorrect, correcting for a deficiency in mutations experienced at these three positions.
6) A fourth mutagenic oligonucleotide, J12, was synthesized with randomized nucleotides at both positions (−1 and +1) adjacent to the loop.
7) In order to test the effect of the 'stu-stem' transversion in all of the above plasmids, a pair of complementary oligonucleotides, J16 and J17, were designed to provide the wild type stem of the *boxB* palindrome. J16 and J17 were synthesized with ambiguity at two positions in the loop, allowing four different loop sequences to be recovered in association with the wild type stem.

Finding and analyzing pTATnut plasmids

The effects of mutations within *boxB* were readily observable following ligation of the oligonucleotide cassettes into pTAT13 and transformation into N567/pAC*N*. Individual $PhoA^{+}$ transformants, colony isolated on agar containing ampicillin + chloramphenicol + XGal + IPTG, showed a characteristic level of blueness, ranging from the white of the pTAT13 parent to the deep blue of pTAT6, the predecessor nut^{+} plasmid. Colonies with all levels of *lacZ* activity were examined for the DNA sequence in and around *nut* and for expression of *lacZ* relative to *phoA* when the plasmid was transferred into cells providing no, little or much wild type N protein. Among the $PhoA^{+}$ transformants, *nut* mutants were recovered at high frequency.

Approximately 100 pTAT*nut* plasmids were sequenced from the primer paired to *p*tac. Sequencing in the opposite direction with the second primer was used to clarify any ambiguous results. All 15 possible single base *boxB* loop mutants were found and are listed in Table 1. Many double and several triple mutants were also found within this loop. No multiple mutant had greater *lacZ* expression than its most crippling single component.

pTAT*nut* plasmids having single or multiple base changes in the *boxB* loop were transformed into strain N567 alone or in combination with plasmid pAC*N* or plasmid pAC*p*tac*N*. This permits one to measure *phoA* and *lacZ* activity in conjunction with no N protein, with little N, or much N. Enzyme activities differed by as much as 50% in different cultures of the same strain when tested on different days. Therefore multiple assays were made on independent cultures to increase reliability. The ratio of alkaline phosphatase to β-galactosidase was used as a means of standardizing the results, equalizing the effects of such extraneous factors as plasmid copy number or level of transcription from *p*tac, which may vary somewhat from culture to culture. No β-galactosidase was ever detectable in the absence of N. Table 1 shows the activities of both β-galactosidase and alkaline phosphatase for different *boxB* mutants, measured with little or much N supplied in trans. Figure 3 summarizes values of antitermination efficiency for these *boxB* mutants relative to the wild type *nut* sequence. In all cases there was good correlation between measured efficiency of antitermination and colony color on XGal.

Spontaneous nucleotide deletions in the *boxA* sequence or mutations in the spacer region between *boxA* and *boxB* were occasionally found with or without *boxB* loop mutations. Alkaline phosphatase and β-galactosidase enzyme activities were also measured in these

Table 1 . Summary of alkaline phosphatase and β-galactosidase activity values in mutant *boxB* plasmid strains induced with IPTG.

pTAT*nut* isolation #	*boxB* loop sequence	N567/pAC*N*/ pTAT*nut* colony color on XGal	Enzyme activity units per 1 OD_{600} unit in N567/pAC*N* β-gal	alk phos	in N567/pACptac*N* β-gal	alk phos	ratio*of β-gal to alk phos in N567/pAC*N*
Y1-1	GAAAA	+++++	61.0	6600	480	2000	92
F2B	GAAGA	+++++	60.2	5600	320	1900	110
A5-22	AAAAA	0	<1	3700	21	3000	<3
C2-1	TAAAA	0	<1	4700	15	3200	<3
A2-103	CAAAA	0	<1	4200	17	3400	<3
A3-4	GGAAA	+++++	59.1	6100	360	1800	97
Z1D9	GTAAA	++	6.7	6900	260	3200	10
B2-7	GCAAA	++	3.4	7800	3.1	3900	4
Z1B14	GAGAA	+++++	51.0	6100	430	1900	84
Z1D11	GATAA	+	<1	4200	225	2900	<3
M5	GACAA	++	2.4	4900	160	2500	5
Z1D8	GAATA	+++	18.4	5700	340	2400	32
Z1A28	GAACA	++	3.1	6300	110	2800	5
Z1B12	GAAAG	+	<1	4500	110	3000	<3
A4-20	GAAAT	+	<1	4300	95	3500	<3
A3B1	GAAAC	+	<1	4200	45	3900	<3
pTAT 13	no *nut*	0	<1	2400	20	1600	<3

The effect of single base *boxB* loop mutations in the pTAT*nut* plasmids on colony color in the presence of XGal + .025mM IPTG, and on alkaline phosphatase and β-galactosidase levels obtained in conjunction with little N protein (pAC*N*) and with much N protein (pACptac*N*) supplied in trans. Colony color: 0 = white, + = palest blue, +++++ = deep blue. Enzyme assays performed on cells lacking N protein showed less than 1 unit of β-galactosidase in every case, but high levels of alkaline phosphatase. β-galactosidase and alkaline phosphatase measured in uninduced cultures generally gave values 3–10% of the IPTG-induced level. The units calculated are the average of 3–4 independent cultures assayed on the same day.
* Values given are ratios multiplied by 10^4.

mutant plasmids if they were found in conjunction with the wild type *boxB* loop. These results are presented in Table 2 and show that positions in the *boxA* sequence as well as in the interval between *boxA* and *boxB* can be important for N-antitermination.

Because the loop-adjacent base pair in the *boxB* stem is T-A, a weaker pair than G-C, the possibility was tested that the −1 and +1 positions with respect to the loop might

G	A	A	A	A 100%
A <3%	G 105%	G 90%	G 120%	G <3%
T <3%	T 10%	T <3%	T 35%	T <3%
C <3%	C 5%	C 5%	C 5%	C <3%

Figure 3. Summary of all single base *boxB* loop mutations and their effect on N-antitermination in strain N567/pAC*N*. Each column represents one position of the 5 nucleotide *boxB* loop, reading 5′ to 3′. The top row gives the wild type loop sequence as found in the right operon of λ. The remaining 3 rows present all the single base mutations from that wild type sequence. The size of each nucleotide represents its functionality in antitermination. The figures given as percentages are antitermination efficiencies, calculated as the ratio of β-galactosidase to alkaline phosphatase activity units in *nut* mutants relative to *nut*$^+$.

Table 2. Summary of alkaline phosphatase and β-galactosidase activity values following IPTG induction in plasmid strains having *boxA* deletions or *boxA–boxB* interval mutations.

pTAT*nut* *isolation #*	*boxA* deletion	change in *boxA–boxB* interval	N567/pACN/ pTAT*nut* colony color on XGal	Enzyme activity units per 1 OD_{600} unit in N567/pAC*N* β-gal	alk phos	in N567/pAC*ptacN* β-gal	alk phos
Z1D13	ΔG	none	++	5.6	5600	196	3300
Z1D10	ΔG	G in pos 4[a]	++	1.7	5700	92	4500
Y3-2	ΔGC	none	+	1.2	5400	63	4500
Y1-10	ΔGCT	none	+	<1	6200	53	4900
B1B1	none	ΔTTAA	+	<1	6000	31	4100

Deletions found in the *boxA* sequence CGCTCTTA and mutations found in the interval sequence AAAATTAA between *boxA* and *boxB* are listed together with their effect on the function of the *nut* sequence. Colony color was observed on plates containing ampicillin, chloramphenicol, XGal and 0.025mM IPTG: the number of +'s indicate the degree of blueness. Cultures for enzyme assays were induced for one hour with 0.5mM IPTG.
[a] position 4 of the spacer is AAA<u>A</u>TTAA.

contribute to loop function. The effects of single or paired changes at these positions, adjacent to a wild type GAAAA loop, show that this is indeed the case (Table 3). This set of mutations (10 of the 15 possible) was not completed, because it became evident that individual nucleotides were determinative, independent of their potential to contribute, through pairing, to stem stability.

The *boxB* loop mutants assessed in Tables 1, 2 and 3 all were associated with a synthetic *boxB* stem, in which a C-G pair was substituted by a G-C pair in order to generate a StuI restriction target (Fig. 2). In order to test the effect of this 'stu-stem' alteration on *boxB* function, 4 of the loop mutations, representing four different classes of loop function, were generated independently in association with the wild type stem using complementary oligonucleotides J16 and J17 (Fig. 2). Comparison of *boxB* function for each loop associated with either wild type or synthetic stem (Table 4) shows that for each loop the wild type stem allows more antitermination function than does the synthetic stu-stem. The difference in activity was only 2-fold with the wild type loop, but the gap widened to more than 10-fold as the loop became increasingly suboptimal.

Table 3. The effect of *boxB* stem mutations in the two positions immediately adjacent to the 5 base *boxB* loop.

pTAT*nut* *isolation #*	upstream *boxB* stem nucleotide	downstream *boxB* stem nucleotide	N567/pACN/ pTAT*nut* colony color on XGal	Relative antitermination efficiency: β-galactosidase / alkaline phos as percentage of Y1-1
K10	A	A	+	
L1	A	C	+++	40
L48	A	T	+++	
K1-2	C	C	+++	30
L1-1	C	T	++	5
L1-5	G	C	++	10
K5	G	G	0	<3
Y1-1	T	A	++++	100
K1-4	T	C	+	
L7	T	G	++	
L1-7	T	T	+	<3

Colony color as in Tables 1 and 2.

Table 4. Comparison of *nut* function by four *boxB* loop sequences when associated with either the wild type or a synthetic *boxB* stem sequence.

boxB loop sequence	boxB stem sequence	Colony color on XGal	Enzyme activity units per 1 OD_{600} alk phosphatase −IPTG	+IPTG	β-galactosidase −IPTG	+IPTG	IPTG-induced β-gal-ase in mutant loop/WT
GAAAA	GCCCT-AGGGC	+++++	394	5995	9.8	82.4	
,,	GGCCT-AGGCC	+++++	523	6941	4.1	35.8	
GAATA	GCCCT-AGGGC	++++	497	5878	5.1	43.9	53%
,,	GGCCT-AGGCC	+++	577	6917	1.2	9.5	27%
GAACA	GCCCT-AGGGC	+++	475	6699	1.9	22.4	28%
,,	GGCCT-AGGCC	++	657	8095	0.6	2.3	7%
GAAAG	GCCCT-AGGGC	++	571	7417	1.1	10.1	12%
,,	GGCCT-AGGCC	+	596	6972	0.7	0.1	<1%

The effects shown were all measured in host N567/*p*AC*N*, which supplies N protein. The same *nut* plasmids tested in N567 with no N gave comparable levels of alkaline phosphatase expression but no measurable β-galactosidase. Colony color was assessed on tryptone agar with XGal and .025mM IPTG at 37°C: + = palest blue, +++++ = deep blue. The enzyme values given are the average of 2−3 independent cultures assayed on the same day. Thus the values for *boxB* loop mutants with the synthetic stem were obtained independently from those in Table 1.

DISCUSSION

Almost every base change at any position within the 5 base *boxB* loop was found to be deleterious to antitermination of transcription by the N protein of bacteriophage λ. Table 1 and Figure 3 show that the wild type sequences GAAAA of *nut*R or GAAGA of *nut*L have maximal activity, while mutants GGAAA and GAGAA are equally functional. Any change at position 1 or 5 causes major loss of function. Changes to T or C in the middle three positions may have strong or moderate influence. A generalized statement of *boxB* optimally functional with N^{λ} is therefore GPPPA, where P stands for purine. The importance of the entire loop sequence is thus substantiated.

Haphazard base changes outside of the loop were also encountered, always upstream from *boxB*. The effects of several *boxA* mutations (Table 2) reiterate the importance of that segment for N function (3). A mutation between the *A* and *B* boxes, changing A to G at the fourth position downstream from *boxA* (#Z1D10), also causes a reduction in antitermination; this mutation occurred together with a deletion of G in *boxA* (#Z1D13), a deletion that by itself reduces antitermination to 10% of wild type; antitermination in #Z1D10 is further reduced to 30% of that in #Z1D13. Although previous experience (7) had given little significance to the *A-B* interval, mutant Z1D10 shows that this interval must be considered further. The rather frequent (about 10% of pTAT*nut* recovered) unsolicited mutations upstream from *boxB* could have been favored for functional reasons: cells containing both pTAT*nut*$^+$ and N^+ plasmids show some sensitivity to IPTG above 0.10mM, forming smaller colonies. For this reason IPTG was held to 0.025mM in plates used to select pTAT-*nut* ligands, making it unlikely that overexpression from pTAT*nut* plasmids could be deleterious. Alternatively, the unsolicited mutations could have occurred as a consequence of errors during oligomer synthesis, since all occurred within the synthesized sequence. No extraneous mutations without phenotypic effects were

encountered, although many *nut* ligands without phenotype were sequenced. The great majority of plasmids sequenced, however, did have altered phenotypes.

Stem positions adjacent to the *boxB* loop have a major influence upon N function (Table 3). This influence is apparently independent of provision for base pairing adjacent to the loop, since a double mutant G-C combination allows little N function. Rather it is indicated that the normal T-A combination is frequently unpaired, allowing positions adjacent to the loop to contribute significantly to the configuration and character of the loop.

The stem sequence of *boxB* , GCCC(T)-(A)GGGC in the early left and right operons of wild type λ was found here to have unexpected influence over the antitermination function of *nut*. Assuming that the stem bases functioned only to present the *boxB* loop conformation, a C-G transversion was incorporated into most of the oligonucleotides used here, in order to create a convenient restriction target. Upon the advise of an astute editor, the transversion stem was subsequently compared to the wild type stem for four different classes of *boxB* loop sequences (Table 4). In each case, the wild type stem gave significantly more antitermination activity than the stem with the G-C transversion. The antitermination activities of the four loop sequences remained, however, in the same relative order, i.e. the order of loop sequence effectiveness with the wild type stem is the same as the order with the 'stu-stem', showing that loop defects are not differentially suppressed by stem sequence. It therefore seems to us unlikely that the stem nucleotides are having a direct influence on loop structure or on the putative interaction between N protein and loop. The stem may rather have a constant influence on another factor in the N-mediated event, for example the affinity of the *nut* structure for one or more of the Nus proteins. The additional stringency of *boxB* loop mutations when associated with the altered stem probably contributed inadvertantly to the recovery of the complete set of loop mutations (Table 1). This set gives a representation of the relative functionality of different loop sequences associated with the synthetic 'stu-stem', a representation that probably holds for the wild type stem as well.

There seems to have been little if any correction of the mismatches created by annealing different oligomers. The inclusion of a G-A-T-C sequence (non-methylated) in mutagenic oligomers J7, J8, J9, J12, J16 and J17 might have instigated methyl-directed mismatch repair (18) of the mutagenic nucleotides, yet mutations were recovered with apparent high efficiency, and in positions and frequencies that corresponded to designations in the synthesized oligomers. Possibly mismatches could not be corrected from the pairings with inosine, but even when the mutagenic oligomers were mismatched with standard nucleotides there was no evident repair. Somehow repair is not significantly operative in this situation.

The functionality of *boxB* in antitermination of transcription is seen to be strongly influenced by the amount of N protein supplied. Tables 1 and 2 show that operons with mutant *nut* sites may be non-functional at low levels of N and yet substantially functional at higher N levels. This reiterates previous experience with heterologous N-*nut* combinations which showed that specificity of N function is partially lost when N is oversupplied (3, 20). Under normal conditions of λ infection, N is supplied at a low level (21), both because of initial low copy number of the phage genome and also because N has been shown to be unstable during normal infection (22). When N is oversupplied from a strongly expressed plasmid operon, it is not as labile and its larger presence has been shown to supply the N requirements of N^- heterologous phages which are not complemented during coinfection with intact N^+ phage (3, 20). The *boxB* sequence requirements may therefore be even more stringent during normal infection than is indicated from the plasmid studies

reported here. The present studies reveal a curious specificity in the response to overproduced N: different *boxB* mutants are helped to different degrees by excess N, independent of their relative responses to limiting N. Combinatorial effects of mutations on nucleotide entities and on composite structure of the *boxB* loop could allow excess N to compensate differentially in different mutants.

We are gratified by the performance of the pTAT plasmids in detecting termination and antitermination of transcription. No *lacZ* function is ever detected in the absence of either *nut* or N. Expression of *lacZ* is detectable over a 60-fold range, as compared to the 10-fold range observed with the *galK* reporter system used in previous detector plasmids (1, 7). The measured levels of β-galactosidase show good agreement with the color of colonies grown in the presence of XGal (Tables 1, 2, 3 and 4), allowing colony color to be used at the screening level for assessment of mutations affecting termination and antitermination. Mutations in *boxB* having either stringent or mild effect were distinguishable, showing that it is possible to screen for partial defects and to observe the effects of mutations at each position in *boxB*. Previous selections based on loss of expression of functions downstream from *boxB* in intact λ had yielded *boxB* mutations at only one position, the more stringent of the two now seen to be most critical to *boxB* function; apparently mutations with mild effect could not be selected in the whole phage situation (2, 7).

The importance of the *boxB* loop sequence for N function, as demonstrated here, is consistent with the hypothesis that there is direct recognition between loop nucleic acid and N protein. A very similar relationship has been described for the interaction of R17 phage coat protein with a particular hairpin structure of R17's RNA genome: direct binding studies of pure coat protein to synthetic RNA oligomers showed the importance of the loop nucleotides, most particularly of the A adjacent to the stem, to protein binding (23, 24). The N protein is also expected to interact with RNA (1), but the known involvement of other proteins in antitermination by N (1,3) allows the possibility that the *boxB* loop in RNA is only indirectly contacted by N. Proof of direct contact will depend upon in vitro studies with pure components or upon the identification of N mutants that specifically compensate for the loop mutations characterized here.

ACKNOWLEDGEMENTS

We thank John S. Parkinson for careful reading of the manuscript, Peter Ames and Hua Ming Wang for help with sequencing from double stranded plasmid DNA, and Jindong Liu for the MacPlasmap computer program. Dale Poulter made the important suggestion that stem bases adjacent to the *boxB* loop be examined.

J.H.D. benefitted from a summer fellowship from the Genetics Society of America. Research was supported by National Science Foundation Grant #DMB-841628.

REFERENCES

1. Friedman, D.I. (1988) In Calendar, R. (ed) The Bacteriophages, Plenum Publishing Co., pp 263–319.
2. Salstrom, J.S. and Szybalski, W. (1978) J. Mol. Biol. **124**, 195–221.
3. Schauer, A.T., Carver, D.L., Bigelow, B., Baron, L.S. and Friedman, D.I. (1987) J. Mol. Biol. **194**, 679–690.
4. Rosenberg, M.m Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978). In Miller, J.H. and Reznikoff, W.S. (eds) The Operon, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 345–371.
5. Franklin, N.C. (1985) J. Mol. Biol. **181**, 75–84.
6. Franklin, N.C. (1989) Plasmid **21**, 31–42.
7. Szybalski, W., Brown, A.L., Hasan, N., Podhajska, A.J. and Somasekhar, G. (1987) In Reznikoff, W.S.,

Burgess, R.R., Dahlberg, J.E., Gross, C.A., Record, M.T. Jr. and Wickens, M.P. (eds) RNA Polymerase and the Regulation of Transcription, Elsevier, New York, pp 381–390.
8. Tanaka, S. and Matsushiro, A. (1985) Gene **38**, 119–129.
9. Masukata, H. and Tomizawa, J.-i. (1988) J. Mol. Biol. **202**, 551–563.
10. Friedman, D.I. and Gottesman, M. (1983) In Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (eds) Lambda II, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 21–51.
11. Chang, A.C.Y. and Cohen, S. (1978) J. Bacteriol. **134**, 1141–1156.
12. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene **33**, 103–119.
13. Reidhaar-Olson, J. F. and Sauer, R.T. (1988) Science **241**, 53–57.
14. Shortle, D., Koshland, D., Weinstock, G.M. and Botstein, D. (1980) Proc. Natl. Acad. Sci. USA **77**, 5375–5379.
15. Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. **7**, 1513–1523.
16. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 366–367.
17. Radman, M. (1988) In Kucherlapati, R. and Smith, G.R.(eds) Genetic Recombination, Amer. Soc. for Microbiol., Washington, D.C., pp 169–192.
18. Steege, D.A., Cone, K.C., Queen, C. and Rosenberg, M. (1987) J. Biol. Chem. **262**, 17651–17658.
19. Franklin, N.C. and Doelling, J.H. (1989) J. Bacteriol. **171**, in press.
20. Radding, C.M. and Echols, H. (1968) Proc. Natl. Acad. Sci. USA **60**, 707–712.
21. Gottesman, S., Gottesman, M., Shaw, J.E. and Pearson, M.L. (1981) Cell **24**, 225–233.
22. Carey, J., Cameron, V., deHaseth, P.L. and Uhlenbeck, O.C. (1983) Biochem. **22**, 2601–2610.
23. Liu, J. and Parkinson, J.S. (1989) Comp. Appl. Biosci., in press.

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