Vol. 190, No. 12

Bacteriophage P22 Antitermination boxB Sequence Requirements Are Complex and Overlap with Those of λ[▽]

Alexis I. Cocozaki, Ingrid R. Ghattas, and Colin A. Smith*

Department of Biology, American University of Beirut, Beirut, Lebanon

Transcription antitermination in phages λ and P22 uses N proteins that bind to similar boxB RNA hairpins in regulated transcripts. In contrast to the λ N-boxB interaction, the P22 N-boxB interaction has not been extensively studied. A nuclear magnetic resonance structure of the P22 N peptide boxB$_{left}$ complex and limited mutagenesis have been reported but do not reveal a consensus sequence for boxB. We have used a plasmid-based antitermination system to screen boxBs with random loops and to test boxB mutants. We find that P22 N requires boxB to have a GNRA-like loop with no simple requirements on the remaining sequences in the loop or stem. U:A or A:U base pairs are strongly preferred adjacent to the loop and appear to modulate N binding in cooperation with the loop and distal stem. A few GNRA-like hexaloops have moderate activity. Some boxB mutants bind P22 and λ N, indicating that the requirements imposed on boxB by P22 N overlap those imposed by λ N. Point mutations can dramatically alter boxB specificity between P22 and λ N. A boxB specific for P22 N can be mutated to λ N specificity by a series of single mutations via a bifunctional intermediate, as predicted by neutral theories of evolution.

λ and P22 share a closely related system of regulating the expression of early lytic genes by allowing transcription past terminators in the P$_{left}$ and P$_{right}$ operons (47). The recognition of the boxB RNA hairpins of nut (*N-ut*ilization) sites by the binding domains of the viral N proteins (Fig. 1A, B, and C) initiates the assembly of an antitermination complex. This complex contains N, host factor NusA, RNA polymerase, and other host factors bound to the viral nut site boxB and the nonhairpin boxA, allowing transcription to proceed through downstream transcription termination signals (32). The four wild-type nut sites of λ and P22 are similar in sequence but differ even between boxB$_{left}$ and boxB$_{right}$ in the same virus. Notably, both P22 boxBs have a C in the loop where both λ boxBs have an A, P22 boxB stems are longer than those of λ by 1 base pair, and P22 boxB$_{right}$ appears to have a noncanonical C:C base pair. boxBs bind noncognate N peptides poorly in vitro (2, 8, 44). Likewise, noncognate N-nut interactions do not function in vivo (14, 28), and noncognate N proteins do not rescue N-deficient viruses (10).

The details of the λ N-boxB interaction have been revealed by extensive genetic and biochemical work (47) and by solution state nuclear magnetic resonance (NMR) structures of the arginine-rich domain of the N protein bound to boxB$_{left}$ (29) and boxB$_{right}$ (39) RNA. Genetic and biochemical studies of the P22 interaction are less complete (13, 44) but are supported by the solution state NMR structure of the arginine-rich domain of the N protein bound to P22 boxB$_{right}$ (5).

λ and P22 boxBs bind their N peptides as hairpins in which 4 of 5 bases adopt a GNRA-like tetraloop structure (Fig. 2). Tetraloops are frequently found in RNAs serving structural

roles as stable caps to stems and as motifs recognized by proteins; GNRA tetraloops are noted as particularly frequent and thermodynamically stable (9). The λ and P22 boxB loops adopt distinct conformations: the fourth nucleotide (nt) of the λ boxB pentaloop is extruded from the GNRA tetraloop (4-out), while in the P22 boxB it is the third nucleotide that is extruded (3-out). As in the canonical GNRA tetraloop, there is a nonstandard, sheared G:A base pair and extensive stacking. The bound peptides adopt helical conformations, occupy the major grooves, and make contacts with boxB stems and loops.

NMR data support a model in which the N peptides recognize a specific conformation of boxB, with little direct recognition of the sequence (5, 29, 39). Most observed contacts are not sequence specific but are made to the backbone or are hydrophobic in nature. The only certain, base-specific hydrogen bonds detected by NMR are to the guanine in the sheared G:A pair at the base of the loop. No other proposed peptide-base hydrogen bond is supported by mutagenesis. Cai et al. (5) report possible hydrogen bonding between a P22 N lysine and the base pair adjacent to the loop of P22 boxB$_{left}$, Legault et al. (29) propose hydrogen bonds between the λ N peptide and a base adjacent to the loop of λ boxB$_{left}$, while Scharpf et al. (39) have evidence only for hydrogen bonds to the guanine of the sheared G:A base pair of λ boxB$_{right}$. λ-P22 specificity may reflect conformational energetics, likely governed by subtle base-stacking interactions.

The regulatory role of RNA-protein recognition in important biological systems has elicited much interest (16). The diversity of known RNA-protein complexes raises the question as to which evolutionary mechanisms are capable of finding new recognition strategies. Kimura's neutral theory of molecular evolution contends that for any given genotype, there are sufficient mutants of neutral fitness to create smooth paths to different phenotypes (26). Though computational modeling of evolving RNA secondary structure strongly supports neutral theories (21, 38, 45), few experimental tests using biologically

* Corresponding author. Mailing address: Department of Biology, American University of Beirut, P.O. Box 11-0236, Riad El Solh 1107 2020, Beirut, Lebanon. Phone: 961 3 791313, ext. 3887. Fax: 961 1 744 461. E-mail: cs10@aub.edu.lb.
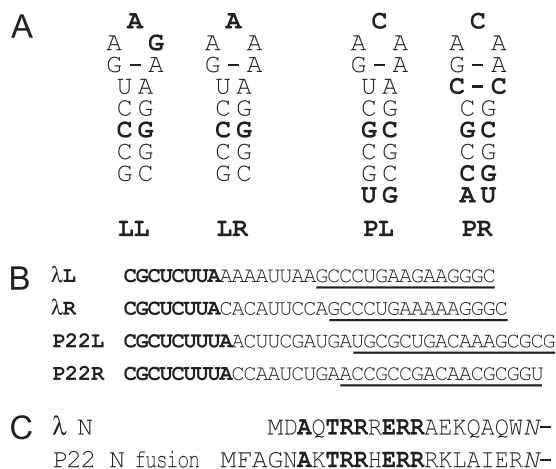
FIG. 1. Comparison of λ and P22 nut site RNAs and N protein RNA-binding domains. (A) The secondary structures of boxBs found in bacteriophages λ and P22 are represented from the beginning of the boxB stem. LL, λ boxB$_{left}$; LR, λ boxB$_{right}$; PL, P22 boxB$_{left}$; PR, P22 boxB$_{right}$. Lightface nucleotides are found in either λ or P22 boxB sequences, while boldface nucleotides are found only in λ boxBs or only in P22 boxBs. The C:C pair adjacent to the loop of P22 boxB$_{right}$ is presumed to be stacked into the helix, perhaps as a hydrogen-bonded, nonstandard base pair. (B) Sequences of λ and P22 nut sites, where boxA is bold and boxB is underlined. λL, λ nut$_{left}$; λR, λ nut$_{right}$; P22L, P22 nut$_{left}$; P22R, P22 nut$_{right}$. In this study, all boxBs were placed in the context of λ nut$_{left}$ to limit the effect of differences between boxAs and the boxA-boxB interval. (C) The sequences of the RNA-binding domains of λ N and the P22 N fusion plasmid used in this study are shown aligned from their amino termini to the fusion point with the common remainder of λ N. The italicized asparagine at the far right replaces a wild-type lysine in λ N to allow a restriction enzyme site in the N-supplying plasmids. Amino acids shown in bold are conserved between λ and P22 N.

FIG. 2. Structures and schematics of λ and P22 boxBs viewed from the minor groove. (A) The NMR structure of λ N-boxB$_{right}$ (Protein Data Bank accession no. 1QFQ [39]) (top), and that of P22 N- boxB$_{left}$ (Protein Data Bank accession no. 1A4T [5]) (bottom) are displayed in truncated form. The peptide is rendered in light-gray van der Waal radius spheres and the loop regions of boxBs are rendered as black lines. (B and C) Base identities and numbering superimposed on structural schematics, showing the λ 4-out and P22 3-out GNRA-like pentaloop conformations. Both have a GAAA tetralooplike structure (9). λ boxB$_{right}$ is above, and P22 boxB$_{left}$ is below. Open boxes represent bases; filled black boxes represent the contacting hydrophobic surface of the N protein.

active RNAs have been reported. Studies of small hairpin RNAs that bind arginine-rich peptides have shown that single substitutions and base pair changes are enough to create changes of specificity (23, 41), suggesting that neutral paths between distinct activities can be found when sequences are small. In the case of protein-binding RNAs, mutations leading from relaxed-specificity sequences to sequences with distinct binding preferences would be equivalent to speciation.

We examined the requirements of boxB recognition by P22 N using a plasmid-based reporter system that reconstitutes antitermination in *Escherichia coli* (14). boxBs active with P22 N were obtained from screens of two randomized loop libraries. Assays of selected and designed boxBs reveal that P22 N recognition requires a GNRA-like pentaloop with no simple requirements on the remaining sequences in the loop or stem. The base pair adjacent to the loop is restricted and appears to modulate N binding in cooperation with the loop and distal stem. A few GNRA-like hexaloops were found to have moderate activity. Single-nucleotide mutations alter boxB specificity and connect boxBs specific for P22 and λ N proteins, as predicted by neutral theories of evolution.

MATERIALS AND METHODS

General. Restriction enzymes, T4 polynucleotide kinase, and T4 DNA ligase were obtained from Roche (Germany). Bacterial-medium components were obtained from Oxoid 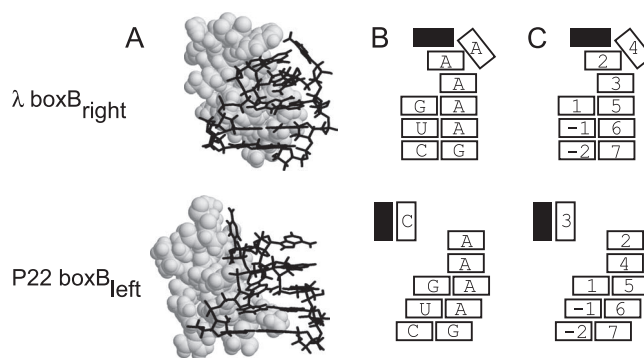(United Kingdom). Fine chemicals were obtained from Amersham (United Kingdom), Sigma (United States), and Amresco (United States).

Bacterial strains, plasmids, and viruses. *Escherichia coli* N567(pBR-ptac-N*λ) cells (11, 14), phage λ*imm*^{P22}Nam24, and phage λNam7am53 (13) were obtained from Naomi Franklin (University of Utah). DH5α cells, nut reporter cloning vector pACnutTAT13 (14), and control plasmids for the human immunodeficiency virus (HIV) Rev-Rev response element (RRE) interaction, namely, pBRN-HIVRev (17) and pAC-HIVRRE (17), were obtained from Kazuo Harada (Tokyo Gakugei University).

Construction of boxB reporter and library plasmids. Wild-type sequences of λ (accession no. NC_001416) and P22 (accession no. NC_002371) were obtained from GenBank. Each reporter plasmid was constructed with the boxB of interest replacing only the boxB of the λ nut$_{left}$ site, such that differences between reporters were confined to boxB. boxB cloning oligonucleotides were designed to form double strands with PstI and BamHI overhangs to replace the entire nut site of pACnutTAT13. The underlined regions in the oligonucleotides used to construct λ nut$_{left}$ indicate the boxB replacement sequences (LLf, 5'-GTCGAC GCTCTTAAAAATTAAGCCCTGAAGAAGGGCAGCATTCAAAGCAGG-3'; LLr, 5'-GATCCCTGCTTTGAATGCTGCCCTTCTTCAGGGCTTAATTT TTAAGAGCGTCGACTGCA-3'). These pairs of oligonucleotides were annealed and ligated into the pACnutTAT13 backbone previously digested with PstI and BamHI. Library inserts were constructed by primer extension with an antisense primer (5'-CGGGGATCCCTGCTTTGAATGC-3') on sense templates (the 7-nt library, 5'-GGGCTGCAGTCGACGCTCTTAAAAATTAATG CGCNNNNNNNGCGAGCATTCAAAGCAGGGATCCCCG-3', or the 6-nt library, 5'-GGGCTGCAGTCGACGCTCTTAAAAATTAATGCGCTNN NNNNAGCGCGAGCATTCAAAGCAGGGATCCCCG-3'), followed by digestion with PstI and BamHI. The digested double-stranded library oligonucleotide was ligated to the backbone, and the ligation mixture was used to transform competent DH5α bacteria by electroporation. The resulting colonies were scraped from the plates and used to prepare the plasmid for screening.

P22 N fusion-expressing plasmid. To construct the P22 N fusion supplier plasmid pBRNP22Nλ, a pair of synthetic DNA oligonucleotides (P22NF, 5'-C ATGTTTGCAGGCAATGCTAAAACTCGCCGTCATGAGCGGCGCAGAA AGCTAGCCATAGAGCGCAATGCA-3', and P22NR, 5'-CATTGCGCTCTA TGGCTAGCTTTCTGCGCCGCTCATGACGGCGAGTTTTAGCATTGCCT GCAAA-3') were designed with NcoI and BsmI overhangs to replace the N terminus of the N protein in pBR-ptac-N*λ with the RNA-binding domain of P22 N. These oligonucleotides were annealed and ligated into plasmid pBR-ptac-N*λ backbone to form pBRNP22Nλ. Ligations were transformed into N567 strains of *E. coli*, and minipreps were selected for each construct.

DNA preparation and sequencing. For each boxB reporter, mixtures of four or more minipreps were separately transformed into λ N, P22 N, and HIV Rev N

TABLE 1. Cognate N-boxB antitermination activities[d]

Reporter	Sequence[a]	β-Galactosidase units[b]			% of cognate's activity[c]		
		Wild-type λ N	P22 N	Rev N	Wild-type λ N	P22 N	Rev N
λ boxB$_{left}$	GCCCU GAAGA AGGGC	9,400 ± 2,100	24 ± 7.2	4.5 ± 0.47	110	25	4.8
λ boxB$_{right}$	GCCCU GAAAA AGGGC	8,300 ± 1,000	29 ± 2.9	4.2 ± 1.6	100	30	5.4
P22 boxB$_{left}$	UGCGCU GACAA AGCGCG	1,600 ± 290	97 ± 4.0	5.1 ± 0.93	20	100	5.5
P22 boxB$_{right}$	ACCGCC GACAA CGCGGU	3.5 ± 0.43	97 ± 8.2	3.3 ± 0.53	0.04	100	5.5
HIV-1 RRE IIB	–UGGGCGCA––UGACGGUACA–	3.5 ± 0.30	3.7 ± 0.55	93 ± 28	0.04	3.8	100

[a] All RNAs were assayed as boxB replacements in the context of the λ nut$_{left}$ site. Gaps flanking the loop are for clarity, and only the high-affinity site of the RRE boxB replacement is shown.
[b] β-Galactosidase units were assayed from saturated cultures in tryptone medium and with continuous 0.5 mM IPTG induction. Values represent averages ± standard deviations for three or more experiments.
[c] The percentage of the cognate's activity is calculated as the activity of the reporter relative to that of the cognate reporter (λ boxB$_{right}$ for wild-type λ N, P22 boxB$_{left}$ for P22 N, and HIV-1 RRE for Rev N).
[d] P22 N is a fusion of the RNA-binding domain of wild-type P22 N to the activation domain of λ N, and Rev N is a fusion of the RNA-binding domain of HIV-1 Rev to the activation domain of λ N.

supplier-containing cells and plated on tryptone agar containing X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) to reveal the general phenotype. At least three representative clones were restreaked, and fresh plasmid was prepared and retransformed to confirm the phenotype prior to large preparations for sequencing. boxB reporter plasmids were sequenced with the forward primer PACF2 (5′-AATCACTGCATAATTCGTGTC-3′), and the P22 N fusion supplier plasmid pBRNP22Nλ was sequenced with PBRNF2 (5′-ACTCCCGTTCTGGATA ATG-3′) and PBRNR2 (5′-GGCTTGCTGTACCATGTG-3′) using an ABI Prism BigDye Terminator v3.1 ready reaction cycle sequencing kit on an ABI 3100-Avant genetic analyzer. The entire insert was confirmed using Chromas Lite software from Technelysium (United Kingdom).

Library screening and X-Gal plate assay. Competent N567 host cells were transformed with the wild-type λ N supplier (pBR-ptac-N*λ), the P22 N fusion supplier (pBRNP22Nλ), or the HIV-Rev N fusion supplier (pBRN-HIVRev) by standard CaCl$_2$ procedures. For each N-boxB plate assay (36), individual boxB reporter plasmids (including the controls of λ boxBs, P22 boxBs, and HIV RRE reporters in parallel) were transformed into λ N, P22 N, and HIV Rev N supplier cells. Typically, 10 to 100 ng of plasmid per 100 μl of competent cells was transformed by heat shock and plated on tryptone plates containing 50 μg/ml ampicillin and 15 μg/ml chloramphenicol as antibiotics and 80 μg/ml X-Gal as the chromogenic substrate of the β-galactosidase reporter protein. Experiments with pBRNP22Nλ and pBRNHIVRev also included 0.05 mM IPTG (isopropyl-β-D-thiogalactoside) to induce the *tac* promoters expressing the N protein and the reporter transcript. The plates were viewed and scored after 1 and 2 days at 34°C. The intensity of the blue colonies was used to score the antitermination activity for selections and the preliminary assessment.

ONPG solution assay of antitermination. For each N-boxB interaction, representative colonies were picked from X-Gal plates for solution assays (31). At least three independent colonies were used for each interaction. For measurement of N-mediated antitermination, cultures were grown overnight at 37°C with aeration in tryptone with 50 μg/ml ampicillin and 15 μg/ml chloramphenicol as antibiotics and with 0.5 mM IPTG. The cells were then permeabilized and assayed for β-galactosidase activity using ONPG (*o*-nitrophenyl-β-D-galactoside), and units of β-galactosidase were calculated according to Miller (31). The percentages of activities are reported normalized to P22 boxB$_{left}$ for P22 N and λ boxB$_{right}$ for λ N because the absolute activities of controls sometimes varied up to threefold between experiments on different days. The levels of activation of specific reporters were calculated as the units of β-galactosidase observed with activation by P22 N or λ N divided by the units observed by N Rev activation.

Structure visualization. Protein Explorer (30) was used to view the solution state NMR structure of λ N peptide-λ boxB$_{right}$ (Protein Data Bank accession no. 1QFQ [39]), the NMR structure of P22 N peptide-P22 boxB$_{left}$ (Protein Data Bank accession no. 1A4T [5]), and the X-ray crystal structure of a *Haloarcula marismortui* large ribosomal subunit (Protein Data Bank accession no. 1JJ2 [27]).

RESULTS

We first confirmed N-boxB activity and specificity using a two-plasmid reporter system of Franklin (14). This reporter system reconstructs λ N-mediated antitermination in *E. coli* and has extensively been used for the study of λ N-boxB interaction and heterologous RNA-protein interactions. Four wild-type boxB reporters (Fig. 1A) were constructed in which λ and P22 boxBs were presented in the context of the λ nut$_{left}$ sites to avoid influence from variations in boxA and the remainder of the nut sites (Fig. 1B). λ N was expressed as the full-length protein, whereas the P22 N RNA-binding domain was expressed as a fusion to the λ activation domain (Fig. 1C). The P22 N fusion limited the differences between N proteins to the well-characterized RNA-binding domains and avoided the high level of background activity displayed by a full-length, wild-type P22 N construct (15; data not shown). The P22 N fusion supplier was tested and found to have low activity on a heterologous HIV RRE nut reporter and strong activity on its cognate boxB reporters as expected, though its absolute activity was substantially less than that of the λ N supplier. The P22-λ N fusion supplier plasmid rescues the P22 N-null phage (λimm^{P22}Nam24 [13]) but not the λ N-null phage (λNam7 am53 [13]; data not shown). Thus, while the P22 N fusion activity likely directly reflects boxB-peptide affinity, the lambda N activity likely reflects the ability of the boxB loop to bind lambda N in the proper 3-out conformation that recruits NusA and allows optimal antitermination (49). The P22-λ N fusion supplier was used for all further experiments.

We measured the activities of the four wild-type boxB reporters with the λ N supplier pBR-ptac-N*λ and with the P22 N fusion supplier pBRNP22Nλ using a solution assay with the β-galactosidase reporter gene (31) (Table 1). The boxBs display specificity for their cognate N proteins; the P22 boxB$_{right}$ reporter displays very high specificity. In vitro experiments indicate substantial affinity between λ boxB$_{left}$ and the P22 N peptide (44) and between P22 boxB$_{left}$ and the λ N peptide (7). Other work in vivo shows that λ boxB$_{right}$ has significant activity with P22 N in vivo (13, 19) and that P22 boxB$_{right}$ is highly specific (13, 28). Despite the uniform λ context of the boxB-peptide interaction, the specificities displayed by our boxB reporters largely agree with published data, though there are unresolved discrepancies between published boxB-N peptide specificities in vitro (2, 7, 43, 44).

Screening of a boxB library. In order to reveal the loop consensus for P22 boxB, we constructed a plasmid library

TABLE 2. Activities of boxB reporters selected from the library of seven random nucleotides

Reporter[a]	Loop sequence[b]	β-Galactosidase units		Fold activation[c]	% of P22$_{left}$'s activity[d]
		P22 N	Rev N		
L1	UGAAUAA	38 ± 20	3.0 ± 0.95	13 ± 6.6	120
L2	UGAGAAA	56 ± 7.9	4.0 ± 0.49	14 ± 2.0	130
L3	UGGUAAA	69 ± 14	6.8 ± 0.39	10 ± 1.99	77
L4	AGGCAAU	41 ± 3.5	3.7 ± 0.73	11 ± 0.95	96
L5	AGGCGAU	58 ± 3.8	6.2 ± 0.26	9.2 ± 0.61	65
L6	UGUUAAA	51 ± 14	6.95 ± 0.89	7.4 ± 2.0	50
L7	AGAUGAU	55 ± 11	7.5 ± 0.31	7.3 ± 1.5	54
L8	UGCAUAU	58 ± 12	7.95 ± 0.22	7.2 ± 1.5	57
L9	AGUCAAU	54 ± 6.6	8.4 ± 0.17	6.5 ± 0.78	51
L10	UGAGUAA	27 ± 2.8	4.3 ± 0.37	6.3 ± 0.64	64
L11	UGUUGAA	74 ± 14	14 ± 2.7	5.3 ± 1.0	72
L12	UGUGUAA	13 ± 4.5	2.8 ± 0.46	4.5 ± 1.6	40
L13	CGCCGAG	33 ± 7.96	6.97 ± 0.96	4.7 ± 1.1	37
L14	UGUACAA	25 ± 3.8	8.02 ± 0.32	3.1 ± 0.47	28

[a] Clones were selected from a randomized loop library presented on a P22 boxB$_{left}$ stem.
[b] Only loop and adjacent base pair sequences are shown; underlining indicates nucleotides different from those of P22 boxB$_{left}$.
[c] The n-fold activation is the ratio of β-galactosidase units of the reporter in the presence of P22 N divided by the β-galactosidase units in the presence of Rev N.
[d] The percentage of P22$_{left}$ activity is normalized to the activity of P22 boxB$_{left}$ assayed the same day.

based on P22 boxB$_{left}$ in which the 7 nt comprising the pentaloop and adjacent base pair were randomized. The NMR structure of P22 boxB bound to the P22 N peptide reveals boxB adopting a 3-out GNRA-like pentaloop, with a role for the extruded C at position 3 (C$_3$) in making hydrophobic contact with P22 N (5), suggesting a G$_1$N$_2$C$_3$R$_4$A$_5$ consensus. The library should contain 16,384 unique sequences, including P22 boxB$_{left}$. The library pool was sequenced and showed no bias in the randomized region. Approximately 40,000 colonies of N567 P22 N were screened on X-Gal plates, and 260 blue colonies were picked, pooled, and used for DNA preparation. To remove false positives, the reporter plasmid was separated from the N supplier plasmid and used to transform N567 Rev

N, where constitutively active reporters (presumably resulting from mutations outside boxB) show activity and reporters specific to P22 N are inactive. White colonies (500; 64%) were selected and used to isolate the reporter plasmid. After one additional round of selection on N567 P22 N yielding colonies that were 55% blue, plasmids were individually examined for activity on X-Gal plates. Nineteen clones with activities similar to that of P22 boxB$_{left}$ were sequenced, yielding 14 unique sequences. These were assayed for antitermination (Table 2). Neither P22 boxB$_{left}$ nor any sequence resembling P22 boxB$_{right}$ was recovered, indicating that our screen was not exhaustive. Indeed, only 1 of the 14 selected was a boxB$_{left}$ point mutant.

Only GNRA-like pentaloops are found in a randomized loop library. As expected of GNRA-like pentaloops, all selected sequences contained a G$_1$:A$_5$ pair (Table 2). All four nucleotides occurred at positions 2, 3, and 4, violating a strict GNRA-like consensus, assuming that all boxBs adopt the expected 3-out conformation when bound. Purines are common at positions 2 and 4 in strongly active boxBs, suggesting that base-stacking interactions are important. In the P22 boxB$_{left}$ NMR structure (5), A4 fulfils the purine role with an intramolecular hydrogen bond, but no prior mutagenesis in boxB had indicated its importance. Consistent with these data, GNRA-like pentaloops are known to exist without a purine in this role (12, 42).

A series of mutant boxBs were constructed to explore the determinants of P22 N activity (Table 3). The A$_2$→G mutation of P22 boxB$_{left}$ (clone D1) does not reduce activity, as was expected from published in vitro data (44). The C$_3$→A mutation (D2) reduces activity to 85%, while a G$_2$→A C$_3$→A double mutation (D3) reduces activity to 36%. In contrast, C$_3$→A A$_4$→G mutations (D4) maintain moderate to wild-type activity, as was also expected from published data (44). The low frequency of positive sequences, their similarity to GNRA tetraloops, and the lack of any observed hydrogen bonds between the peptide and bases are consistent with a model where the ability of boxB to adopt the required conformation depends on favorable base-stacking interactions (24), as well as

TABLE 3. Activities of designed boxB mutants

Clone[a]	boxB sequence[b]	β-Galactosidase units		Fold activation	% of P22$_{left}$'s activity
		P22 N	Rev N		
D1	UGCGCU GGCAA AGCGCG	120 ± 9.2	6.8 ± 0.55	18 ± 1.4	100
D2	UGCGCU GAAAA AGCGCG	90 ± 11	11 ± 0.52	8.3 ± 0.98	85
D3	UGCGCU GGAAA AGCGCG	38 ± 7.4	9.5 ± 1.4	4.0 ± 0.78	36
D4	UGCGCU GAAGA AGCGCG	87 ± 15.1	11 ± 1.0	7.9 ± 1.4	82
D5	UGCGCA GACAA UGCGCG	110 ± 4.3	8.0 ± 0.32	14 ± 0.54	100
D6	UGCGCA GAAAA UGCGCG	38 ± 9.4	4.2 ± 0.26	9.0 ± 2.2	64
D7	UGCGCA GGAAA UGCGCG	7.9 ± 1.1	4.2 ± 0.10	1.9 ± 0.27	13
D8	UGCGCC GACAA CGCGCG	11 ± 0.41	3.9 ± 0.21	2.9 ± 0.11	14
D9	UGCGCC GACAA GGCGCG	11 ± 1.7	4.2 ± 0.76	2.7 ± 0.39	15
D10	UGCGCG GACAA CGCGCG	4.4 ± 0.17	3.5 ± 0.28	1.2 ± 0.05	5.7
D11	UGCGCG GACAA GGCGCG	5.7 ± 0.94	3.8 ± 0.43	1.5 ± 0.25	7.3
D12	UGCGCG GACAA UGCGCG	3.3 ± 0.17	4.1 ± 0.22	0.80 ± 0.04	4.3
D13	ACCGCU GACAA AGCGGU	140 ± 9.7	3.4 ± 0.12	41 ± 2.9	140
D14	UCCGCU GACAA AGCGGG	110 ± 22	3.4 ± 0.61	31 ± 6.4	110
D15	AGCGCU GACAA AGCGCU	30 ± 6.7	3.1 ± 0.53	9.8 ± 2.2	31

[a] Reporters are designed boxB replacements in the context of the λ nut$_{left}$ site.
[b] Gaps flank the loop; underlining indicates nucleotides different from those of P22 boxB$_{left}$.

TABLE 4. Activities of boxBs with large loops

Clone[a]	Loop sequence[b]	β-Galactosidase units		Fold activation	% of P22$_{left}$'s activity
		P22 N	Rev N		
L15	GA<u>A</u>CAA	32 ± 5.2	3.4 ± 0.44	9.1 ± 1.5	54
L16	GA<u>AU</u>AA	19 ± 4.3	4.2 ± 0.65	4.7 ± 1.0	24
L17	GA<u>AUU</u>A	14 ± 3.5	3.5 ± 0.25	4.01 ± 0.98	17
D16	GAC<u>A</u>AA	6.0 ± 0.53	4.8 ± 0.54	1.2 ± 0.11	4.9
D17	GA<u>AC</u>AAA	4.4 ± 0.33	4.7 ± 0.34	0.94 ± 0.07	3.6

[a] Clones designated with L were selected from a randomized loop library, and those designated with D were designed boxB replacements. All boxBs are presented on a P22 boxB$_{left}$ stem in the context of the λ nut$_{left}$ site.
[b] Only loop and adjacent base pair sequences are shown; underlining indicates nucleotides different from those of P22 boxB$_{left}$.

hydrophobic interactions with the N protein (4) and ancillary factors (49).

A:U and U:A are preferred adjacent to the loop. The absence of a Watson-Crick base pair adjacent to the loop of P22 boxB$_{right}$ suggests either a tolerance of nonpaired bases or C:C pairing. The loop of phage 21 boxB adopts a structurally related conformation, and NMR-based molecular modeling suggests C:C pairing (7). Our library's random region encompassed these positions in order to examine their importance. Only U:A (wild type) and A:U base pairs were found in strongly positive clones, while the only boxB with a C:G base pair had reduced activity (Table 2, reporter L13). To address the role of the bases adjacent to the loop, we constructed a series of boxBs with substitutions at these positions (Table 3, D5 to D12). The placement of A:U into boxB$_{left}$ did not reduce activity (D5), though it did affect boxB$_{left}$ mutants (D6, D7). Interestingly, C:C, C:G, G:C, and wobble pair replacements strongly reduced activity to 15% or less in the context of boxB$_{left}$ (D8 to D12).

How does boxB$_{right}$ tolerate the apparent C:C pair? The U:A substitution adjacent to the loop of boxB$_{right}$ increased activity (D13), indicating that U:A is also preferred in this context and suggesting that the two base pair differences between boxB$_{left}$ and boxB$_{right}$ at the beginning of the stem contribute to boxB$_{right}$ activity. The separate placement of two boxB$_{right}$ base pairs into a boxB$_{left}$ context (D14, D15) increased and decreased activity, respectively. These data suggest that the contiguous stacking from the beginning of the stem through the loop allows distal mutations to modulate the backbone conformations required for binding P22 N.

Screening a hexaloop library yields few active boxBs. Since the GNRA motif is found imbedded in larger loops (1), we constructed a library in which the 5-nt loop of boxB$_{left}$ was replaced

with 6 random nt. This library should contain 4,096 unique sequences. Approximately 10,000 colonies were screened in N567 P22 N on X-Gal plates, and 287 blue colonies were picked, pooled, and screened against Rev N. Twelve colonies of the negative clones (90%) were collected and sequenced, yielding eight unique sequences. Of these, three were hexaloops and five were pentaloops. The unselected library pool was sequenced and showed no significant contamination with random pentaloops (data not shown), suggesting that active hexaloops are far rarer than pentaloops. Interestingly, the three selected hexaloops are all related by single insertions into high-activity boxBs (Table 4). A designed hexaloop (D16) and heptaloop (D17) were found to have minimal activity. We speculate that larger loops are more likely to accommodate extruded bases where they sterically hinder protein binding. Indeed, two hexaloops and one heptaloop with structures resembling GNRA tetraloops (20) found in a *Haloarcula marismortui* large ribosomal subunit (27) accommodated the extra nucleotides where they would interfere with N protein binding.

Mutations alter boxB specificity toward P22 N and λ N. The diversity of boxB variants active with P22 N suggested that some may also function with λ N. Though the recognition strategies of P22 and λ N are distinct, the origin of specificity is not obvious, and some sequences might be able to participate in both recognition strategies. Tan and Frankel (44) reported a boxB (equivalent to D4 here) that binds both P22 N and λ N peptides in vitro. We first designed and assayed some additional reporters with λ boxB stems (Table 5), which showed that the λ boxB stem did not abrogate P22 N activation. We measured the activities of all described boxB reporters with P22 and λ N proteins (Table 6). Using the ratio of the percentage of P22 N to the percentage of λ N activity to represent specificity, a wide range of specificities is apparent (Fig. 3A). While a few boxBs had higher specificity to λ N than to wild-type λ boxBs, none were found to be more specific to P22 N than to P22 boxB$_{right}$, possibly because most variants were in a P22 boxB$_{left}$ context and P22 boxB$_{right}$ has extremely high specificity. Remarkably, many boxBs have relaxed specificity and maintain at least moderate activity. Several boxBs selected for P22 N activity have more than 50% activity with λ N, suggesting that bifunctional boxBs may be relatively common.

Stem and loop position 3 make a clear contribution to specificity. To understand the basis for boxB specificity, nucleotide changes in different boxB contexts were examined. We observe that λ N tolerates λ boxBs with A$_3$→C loop mutations (D18, D22). In contrast to those of other workers, our results suggest

TABLE 5. Activities of boxBs with λ stems

Clone[a]	boxB sequence[b]	β-Galactosidase units		Fold activation	% of P22$_{left}$'s activity
		P22 N	Rev N		
D18	GC<u>C</u>CU GACAA AG<u>G</u>GC	54 ± 3.2	4.4 ± 0.17	12 ± 0.72	87
D19	GC<u>CCA</u> GA<u>AAA</u> <u>U</u>GGGC	5.8 ± 0.63	4.1 ± 0.30	1.4 ± 0.16	9.3
D20	GC<u>C</u>CU G<u>G</u>AAA AG<u>G</u>GC	14 ± 4.0	3.7 ± 0.09	3.7 ± 1.1	24
D21	GC<u>C</u>CU GA<u>G</u>AA AG<u>G</u>GC	15 ± 5.1	4.0 ± 0.44	3.7 ± 1.3	25
D22	GC<u>C</u>CU GAC<u>G</u>A AG<u>G</u>GC	35 ± 5.9	4.0 ± 0.51	8.8 ± 1.5	57

[a] All boxBs are represented as a boxB in the context of the λ nut$_{left}$ site.
[b] Gaps flank the loop; underlining indicates nucleotides different from those of P22 boxB$_{left}$.

TABLE 6. Specificities of boxB reporters with P22 N and λ N

Clone[a]	% of the cognate's activity[b]		Ratio of P22 N activity to λ N activity[c]
	P22 N	Wild-type λ N	
LL	25	110	0.22
LR	30	100	0.30
PL	100	20	5.0
PR	100	0.04	2,500
L1	120	93	1.3
L2	130	75	1.8
L3	77	26	3.0
L4	96	0.50	190
L5	65	0.17	380
L6	50	1.2	41
L7	54	62	0.87
L8	57	5.8	9.8
L9	51	0.06	850
L10	64	56	1.1
L11	72	0.76	96
L12	40	0.27	150
L13	37	29	1.3
L14	28	15	1.9
L15	54	0.26	210
L16	24	0.8	30
L17	17	0.38	46
D1	100	87	1.2
D2	85	110	0.77
D3	36	75	0.48
D4	82	120	0.69
D5	100	0.2	510
D6	64	43	1.5
D7	13	74	0.18
D8	14	0.07	210
D9	15	2.4	6.2
D10	5.7	0.06	98
D11	7.3	0.06	120
D12	4.3	0.05	84
D13	140	10	14
D14	110	17	6.2
D15	31	4.7	6.6
D16	4.9	0.57	8.7
D17	3.6	0.47	7.7
D18	87	100	0.87
D19	9.3	160	0.06
D20	24	120	0.20
D21	25	120	0.21
D22	57	100	0.57

[a] Clone identities are as described in Tables 1 to 5.

[b] The percentage of the cognate's activity is calculated as the activity of the reporter relative to that of the cognate reporter (P22 boxB$_{left}$ for P22 N and boxB$_{right}$ for λ N).

[c] The ratio of activity is the percentage of P22 N activity divided by the percentage of λ N activity and indicates specificity.

that the purine role can be fulfilled by C_3 in the 4-out GNRA-like pentaloop (6, 11). We cannot account for these discrepancies, but note that we have used a different reporter system and that under higher λ N expression, Doelling and Franklin report that the λ $A_3 \rightarrow C$ boxB$_{right}$ has moderate activity (11). In all other cases where comparisons can be made, our results agree well with published data. Interestingly, changing from C_3 to A_3 in every context shifted specificity from P22 to λ N activity, sometimes without a loss of the activity of one or the other N (Fig. 3A).

Published reports indicate that λ N is tolerant of mutations in the boxB stem so long as most of the base pairing is preserved (6, 11, 44). We find that exchanging a P22 boxB$_{left}$ stem

with that of λ boxB shifted specificity toward λ N, again without a necessary loss of activity of one or the other N (Fig. 3B). These coordinate shifts suggest that position 3 and the stem directly contribute to specificity. In contrast, though U:A-to-A:U changes adjacent to the loop are tolerated by P22 N and λ N in their cognate boxBs, the base pair switch causes no consistent alteration of specificity in other contexts (Fig. 3C). These data suggest that this base pair has a complex role coupled with the stem and loop sequence.

Single-nucleotide mutations can switch specificity between P22 and λ N. Point mutants of wild-type boxBs have a wide range of specificities, including relaxed specificity (Fig. 3D). Relaxed-specificity boxBs suggest that transitions between P22
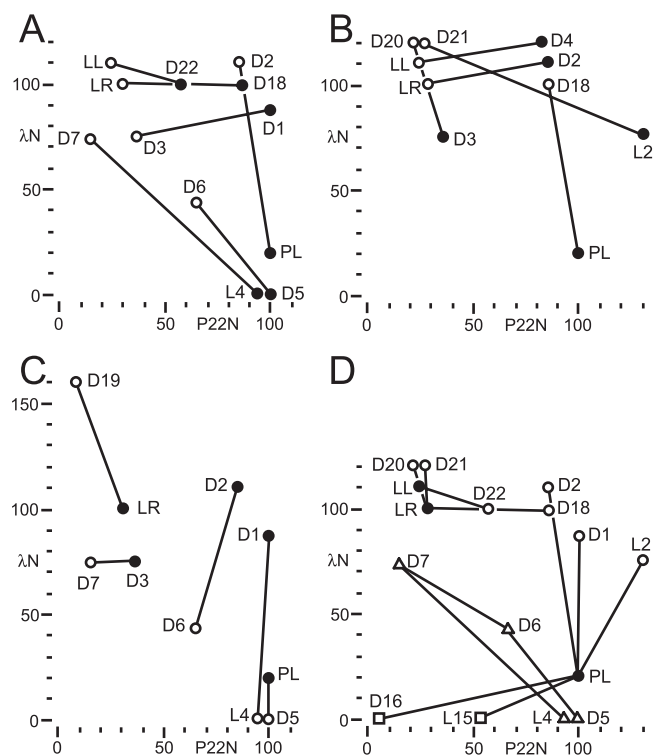

FIG. 3. Specificities of mutant boxBs. Representative boxB reporters in this study are plotted by their percentages of the activities with P22 N (abscissa) and λ N (ordinate) relative to those with P22 boxB$_{left}$ and λ boxB$_{right}$, respectively (Table 6). All clones are as in Tables 1 to 5. (A) boxBs related by C-to-A changes at position 3 are shown connected (D22 to λ boxB$_{left}$ [LL], D18 to λ boxB$_{right}$ [LR], P22 boxB$_{left}$ [PL] to D2, L4 to D7, D5 to D6). boxBs with C_3 are shown as filled circles, and boxBs with A_3 are shown as open circles. (B) boxBs related by P22 boxB$_{left}$-to-λ boxB stem changes are shown connected (D4 to LL, D2 to LR, PL to D18, L2 to D21, D3 to D20), where boxBs with P22 boxB$_{left}$ stems are shown as filled circles and boxBs with λ boxB stems are shown as open circles. (C) boxBs related by U:A-to-A:U changes adjacent to the loop are shown connected (LR to D19, D3 to D7, D2 to D6, D1 to L4, PL to D5), where boxBs with U:A pairs are shown as filled circles and boxBs with A:U pairs are shown as open circles. (D) Selected boxBs related by single-nucleotide changes are shown connected (LL to D22, LR to D18, LR to D20, LR to D21, PL to D1, PL to D2, PL to L2, PL to L15, PL to D16, L4 to D7, D5 to D6, D6 to D7). Wild-type boxBs are represented by filled circles, and single-substitution mutants are shown as open circles. In the case of P22 boxB$_{left}$, single-insertion mutants are shown as open squares. Four mutant boxBs related by single substitutions are shown as open triangles.

and λ recognition strategies may occur without an intermediate loss of function. We observed no series of single mutations connecting wild-type P22 and λ boxBs, because they differ by a base pair in the stem that was not mutagenized. However, a single mutation transforms P22 N-specific D5 into moderately bifunctional D6, which is a single mutation from λ N-specific D7. Indeed, a single substitution connects P22 N-specific L4 to λ N-specific D7.

DISCUSSION

Our results show that P22 N recognizes GNRA-like pentaloops, without strict adherence to the purine requirement. P22 N can also recognize GNRA-like hexaloops with reduced activity. Many boxB variants are recognized, but no simple pattern emerges. The subtle effects of base stacking likely dictate the ability of the pentaloop to adopt the conformation necessary for P22 N binding. P22 N prefers U:A and A:U adjacent to the loop and tolerates the C:C pair of boxB$_{right}$ with the cooperation of distal-stem base pairs. Single mutations can dramatically alter boxB specificity between P22 and λ N. Though we found no path between wild-type boxBs, we did find that a series of single mutations connects P22 and λ N-specific boxBs via a relaxed-specificity boxB, as predicted by neutral theories of evolution. Libraries were randomized only in the loop region, and screening was not exhaustive, suggesting that more bifunctional boxBs exist. Our results with P22 N illustrate the complexity of even small RNA-peptide interactions.

While our results relate to the P22 boxB-N peptide interaction, we have recapitulated the interaction in a somewhat heterologous λ antitermination system that has the added complexity of reflecting the ability of host factors to assemble on the boxB-N complex (as reported by Conant et al. [8a]). Mutations that do not affect N binding can affect antitermination (6, 32, 37), and heterologous RNA-protein interactions display substantially lower activities than would be predicted by their in vitro affinities (17). Lower absolute activities of non-λ interactions are consistent with the model where λ N tryptophan 18 stacks atop the λ boxB 4-out pentaloop, creating a precise conformation sensed by NusA, which then allows optimal antitermination (49). The P22 N fusion presumably does not permit proper NusA binding, since there is no residue to act as λ N tryptophan 18 and boxB adopts a 3-out conformation when bound to P22 N (4).

The sequence requirements for N binding may be particularly complex because of the relative absence of base-specific hydrogen bonds. In light of the NMR evidence of extensive contacts between P22 N peptide and the 5′ backbone of the lower stem (5), even minor conformational changes could drastically affect peptide binding. We speculate that the P22 loop consensus is obscure (besides the G:A pair), because the primary role of loop bases may be to promote a 3-out pentaloop for P22 N binding. Even the extruded C$_3$ makes only hydrophobic interactions with N protein. This model is similar to that proposed for λ boxB, in which the extent of λ peptide binding reflected the proportion of the 4-out pentaloop in unbound boxB (24). The marked preference for U:A and A:U base pairs adjacent to the loop also suggests subtle base-stacking interactions that influence the conformation of the loop

and stem. Other examples are known where base pairs influence the stability of proximal loops, such as that of λ boxB (24) and a related tetraloop (46). Interestingly, U:A and A:U base pairs adjacent to the loop of bovine immunodeficiency virus (BIV) trans-acting responsive element indirectly modulate the binding of the BIV Tat peptide to the stem (41). We speculate that the cytosines of P22 boxB$_{right}$ are tolerated because they are paired and stacked between the loop and the stem, as proposed for the homologous C:C base pair in phage 21 boxB (7). The complex thermodynamics of base stacking would be unlikely to yield a simple consensus.

What determines specificity between λ and P22 N-boxB interactions? Fluorescence studies support a model where any boxB bound by λ N adopts the 4-out conformation and any boxB bound by P22 N adopts a 3-out conformation (2). The thermodynamic costs of adopting λ- or P22-like loop and stem conformations likely dictate N binding, and specificity may be a simple consequence of disfavoring one conformation. Relaxed-specificity RNAs would be able to adopt either conformation, though at some cost to activity.

Neutral theories of evolution highlight the importance of genetic drift, asserting that mutational paths of neutral fitness exist between genotypes of distinct phenotypes (26, 34, 45). Computational studies of RNA secondary structure strongly support this assertion (21, 38, 45). However, experimental evidence with biologically active RNAs has been limited. Experimental support for the application of neutral theories to RNA activity has come from only a few examples (18, 22, 23, 40). Our findings that single substitutions connect a P22 N-specific boxB via a moderately bifunctional boxB to a λ N-specific boxB support the extension of neutral theories from computationally predictable RNA secondary structures to biologically active RNA phenotypes. However, the existence of bifunctional RNAs linking the N-dependent boxBs of P22 and λ boxBs to those of phage 21 or the structurally unrelated and N-independent RNAs of the HK022 polymerase utilization site (3) seems less likely. Other evolutionary mechanisms exist, such as RNA-protein coevolution and recombination between lambdoid phages (25). Nonetheless, we find that neutral theories apply in this small-model system, where conformation, rather than secondary structure as defined by base pairing, appears to define biological activity.

How common are RNAs that display multiple phenotypes? The intersection theorem of RNA secondary structure states that for any two secondary structures there is at least one sequence that is compatible with both structures (38). The phenotypes of biologically active RNA are typically more complex than the secondary structure, yet genetic drift along neutral paths may lead to new activities. Schultes and Bartel (40) found their intersection sequence able to adopt either of two distinct folds, albeit with severely reduced activity. Riboswitches are observed to undergo changes in secondary structure suggestive of intersection sequences, but few examples of RNAs that can bind different partners with no change in base pairing are known. NMR studies reveal that HIV RRE binds to its cognate peptide HIV Rev and to a selected peptide, RSG1.2, maintaining the same base pairing (50). A mutant hairpin RNA is able to function in either HIV or BIV transcriptional transactivation by binding homologous arginine-rich peptides in different recognition modes (41). These exam-

ples of RNAs able to bind multiple partners suggest that RNAs can serve as intersections of neutral paths between specific phenotypes.

Relaxed specificity may require the ability to adopt distinct conformations. Though plasticity comes with lowered thermodynamic stability, evolution presents many examples of conformational flexibility. Induced-fit interactions are commonly observed in RNA-protein interactions and in arginine-rich peptide-RNA interactions in particular (35). Indeed, the arginine-rich domain of λ N protein is entirely disordered until it is bound to boxB (33), and the boxB loop becomes ordered upon peptide binding (4, 24). The phenomenon of induced fit may increase the likelihood of sequences that can adopt more than one functional conformation, whether to signal binding, to interact with multiple partners, or to serve as intersection sequences. RNAs able to bind multiple partners seem more likely to facilitate the neutral coevolution of protein partners.

REFERENCES

1. **Abramovitz, D. L., and A. M. Pyle.** 1997. Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. J. Mol. Biol. **266:**493–506.
2. **Austin, R. J., T. Xia, J. Ren, T. T. Takahashi, and R. W. Roberts.** 2003. Differential modes of recognition in N peptide-boxB complexes. Biochemistry **42:**14957–14967.
3. **Banik-Maiti, S., R. A. King, and R. A. Weisberg.** 1997. The antiterminator RNA of phage HK022. J. Mol. Biol. **272:**677–687.
4. **Barrick, J. E., and R. W. Roberts.** 2003. Achieving specificity in selected and wild-type N peptide-RNA complexes: the importance of discrimination against noncognate RNA targets. Biochemistry **42:**12998–13007.
5. **Cai, Z., A. Gorin, R. Frederick, X. Ye, W. Hu, A. Majumdar, A. Kettani, and D. J. Patel.** 1998. Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. Nat. Struct. Biol. **5:**203–212.
6. **Chattopadhyay, S., J. Garcia-Mena, J. DeVito, K. Wolska, and A. Das.** 1995. Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. Proc. Natl. Acad. Sci. USA **92:**4061–4065.
7. **Cilley, C. D., and J. R. Williamson.** 2003. Structural mimicry in the phage phi21 N peptide-boxB RNA complex. RNA **9:**663–676.
8. **Cilley, C. D., and J. R. Williamson.** 1997. Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE). RNA **3:**57–67.
8a. **Conant, C. R., M. R. Van Gilst, S. E. Weitzel, W. A. Rees, and P. H. von Hippel.** 2005. A quantitative description of the binding states and in vitro function of antitermination protein N of bacteriophage lambda. J. Mol. Biol. **348:**1039–1057.
9. **Correll, C. C., and K. Swinger.** 2003. Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 Å resolution. RNA **9:**355–363.
10. **Dambly, C., and M. Couturier.** 1971. A minor Q-independent pathway for the expression of the late genes in bacteriophage lambda. Mol. Gen. Genet. **113:**244–250.
11. **Doelling, J. H., and N. C. Franklin.** 1989. Effects of all single base substitutions in the loop of boxB on antitermination of transcription by bacteriophage lambda's N protein. Nucleic Acids Res. **17:**5565–5577.
12. **Fountain, M. A., M. J. Serra, T. R. Krugh, and D. H. Turner.** 1996. Structural features of a six-nucleotide RNA hairpin loop found in ribosomal RNA. Biochemistry **35:**6539–6548.
13. **Franklin, N. C.** 2004. Morphing molecular specificities between Arm-peptide and NUT-RNA in the antitermination complexes of bacteriophages lambda and P22. Mol. Microbiol. **52:**815–822.
14. **Franklin, N. C.** 1993. Clustered arginine residues of bacteriophage lambda N protein are essential to antitermination of transcription, but their locale cannot compensate for boxB loop defects. J. Mol. Biol. **231:**343–360.
15. **Franklin, N. C., and J. H. Doelling.** 1989. Overexpression of N antitermination proteins of bacteriophages λ, 21, and P22: loss of N protein specificity. J. Bacteriol. **171:**2513–2522.
16. **Hall, K. B.** 2002. RNA-protein interactions. Curr. Opin. Struct. Biol. **12:**283–288.
17. **Harada, K., S. S. Martin, and A. D. Frankel.** 1996. Selection of RNA-binding peptides in vivo. Nature **380:**175–179.
18. **Held, D. M., S. T. Greathouse, A. Agrawal, and D. H. Burke.** 2003. Evolutionary landscapes for the acquisition of new ligand recognition by RNA aptamers. J. Mol. Evol. **57:**299–308.
19. **Hilliker, S., and D. Botstein.** 1975. An early regulatory gene of *Salmonella* phage P22 analogous to gene N of coliphage lambda. Virology **68:**510–524.
20. **Huang, H. C., U. Nagaswamy, and G. E. Fox.** 2005. The application of cluster analysis in the intercomparison of loop structures in RNA. RNA **11:**412–423.
21. **Huynen, M. A., P. F. Stadler, and W. Fontana.** 1996. Smoothness within ruggedness: the role of neutrality in adaptation. Proc. Natl. Acad. Sci. USA **93:**397–401.
22. **Ikawa, Y., K. Tsuda, S. Matsumura, S. Atsumi, and T. Inoue.** 2003. Putative intermediary stages for the molecular evolution from a ribozyme to a catalytic RNP. Nucleic Acids Res. **31:**1488–1496.
23. **Iwazaki, T., X. Li, and K. Harada.** 2005. Evolvability of the mode of peptide binding by an RNA. RNA **11:**1364–1373.
24. **Johnson, N. P., W. A. Baase, and P. H. von Hippel.** 2005. Low energy CD of RNA hairpin unveils a loop conformation required for lambda N antitermination activity. J. Biol. Chem. **280:**32177–32183.
25. **Juhala, R. J., M. E. Ford, R. L. Duda, A. Youlton, G. F. Hatfull, and R. W. Hendrix.** 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. J. Mol. Biol. **299:**27–51.
26. **Kimura, M.** 1991. Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. Proc. Natl. Acad. Sci. USA **88:**5969–5973.
27. **Klein, D. J., T. M. Schmeing, P. B. Moore, and T. A. Steitz.** 2001. The kink-turn: a new RNA secondary structure motif. EMBO J. **20:**4214–4221.
28. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59:**207–218.
29. **Legault, P., J. Li, J. Mogridge, L. E. Kay, and J. Greenblatt.** 1998. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell **93:**289–299.
30. **Martz, E.** 2002. Protein Explorer: easy yet powerful macromolecular visualization. Trends Biochem. Sci. **27:**107–109.
31. **Miller, J. H.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
32. **Mogridge, J., T. F. Mah, and J. Greenblatt.** 1995. A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. Genes Dev. **9:**2831–2845.
33. **Mogridge, J., P. Legault, J. Li, M. D. Van Oene, L. E. Kay, and J. Greenblatt.** 1998. Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. Mol. Cell **1:**265–275.
34. **Ohta, T.** 1998. Evolution by nearly-neutral mutations. Genetica *102*-**103:**83–90.
35. **Patel, D. J.** 1999. Adaptive recognition in RNA complexes with peptides and protein modules. Curr. Opin. Struct. Biol. **9:**74–87.
36. **Peled-Zehavi, H., C. A. Smith, K. Harada, and A. D. Frankel.** 2000. Screening RNA-binding libraries by transcriptional antitermination in bacteria. Methods Enzymol. **318:**297–308.
37. **Rees, W. A., S. E. Weitzel, A. Das, and P. H. von Hippel.** 1997. Regulation of the elongation-termination decision at intrinsic terminators by antitermination protein N of phage lambda. J. Mol. Biol. **273:**797–813.
38. **Reidys, C., C. V. Forst, and P. Schuster.** 2001. Replication and mutation on neutral networks. Bull. Math. Biol. **63:**57–94.
39. **Scharpf, M., H. Sticht, K. Schweimer, M. Boehm, S. Hoffmann, and P. Rosch.** 2000. Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. Eur. J. Biochem. **267:**2397–2408.
40. **Schultes, E. A., and D. P. Bartel.** 2000. One sequence, two ribozymes: implications for the emergence of new ribozyme folds. Science **289:**448–452.
41. **Smith, C. A., S. Crotty, Y. Harada, and A. D. Frankel.** 1998. Altering the context of an RNA bulge switches the binding specificities of two viral Tat proteins. Biochemistry **37:**10808–10814.
42. **Stefl, R., and F. H. Allain.** 2005. A novel RNA pentaloop fold involved in targeting ADAR2. RNA **11:**592–597.
43. **Su, L., J. T. Radek, L. A. Labeots, K. Hallenga, P. Hermanto, H. Chen, S. Nakagawa, M. Zhao, S. Kates, and M. A. Weiss.** 1997. An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. Genes Dev. **11:**2214–2226.
44. **Tan, R., and A. D. Frankel.** 1995. Structural variety of arginine-rich RNA-binding peptides. Proc. Natl. Acad. Sci. USA **92:**5282–5286.
45. **van Nimwegen, E., J. P. Crutchfield, and M. Huynen.** 1999. Neutral evolution of mutational robustness. Proc. Natl. Acad. Sci. USA **96:**9716–9720.

46. **Vecenie, C. J., and M. J. Serra.** 2004. Stability of RNA hairpin loops closed by AU base pairs. Biochemistry **43:**11813–11817.

47. **Weisberg, R. A., and M. E. Gottesman.** 1999. Processive antitermination. J. Bacteriol. **181:**359–367.

48. **Weiss, M. A., and N. Narayana.** 1998. RNA recognition by arginine-rich peptide motifs. Biopolymers **48:**167–180.

49. **Xia, T., A. Frankel, T. T. Takahashi, J. Ren, and R. W. Roberts.** 2003. Context and conformation dictate function of a transcription antitermination switch. Nat. Struct. Biol. **10:**812–819.

50. **Zhang, Q., K. Harada, H. S. Cho, A. D. Frankel, and D. E. Wemmer.** 2001. Structural characterization of the complex of the Rev response element RNA with a selected peptide. Chem. Biol. **8:**511–520.