JB Journals.ASM.org

HK022 Nun Requires Arginine-Rich Motif Residues Distinct from λ N

Caroline S. Tawk,* Ingrid R. Ghattas, Colin A. Smith

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

ABSTRACT

Bacteriophage λ N protein binds boxB RNA hairpins in the nut (N utilization) sites of immediate early λ transcripts and interacts with host factors to suppress transcriptional termination at downstream terminators. In opposition to λ N, the Nun protein of HK022 binds the boxBs of coinfecting λ transcripts, interacts with a similar or identical set of host factors, and terminates transcription to suppress λ replication. Comparison of N-boxB and Nun-boxB nuclear magnetic resonance (NMR) structural models suggests similar interactions, though limited mutagenesis of Nun is available. Here, libraries of Nun's arginine-rich motif (ARM) were screened for the ability to exclude λ coinfection, and mutants were assayed for Nun termination with a boxB plasmid reporter system. Several Nun ARM residues appear to be immutable: Asp26, Arg28, Arg29, Arg32, Trp33, and Arg36. Asp26 and Trp33 appear to be unable to contact boxB and are not found at equivalent positions in λ N ARM. To understand if the requirement of Asp26, Trp33, and Arg36 indicated differences between HK022 Nun termination and λ N antitermination complexes, the same Nun libraries were fused to the activation domain of λ N and screened for clones able to complement N-deficient λ. Mutants were assayed for N antitermination. Surprisingly, Asp26 and Trp33 were still essential when Nun ARM was fused to N. Docking suggests that Nun ARM contacts a hydrophobic surface of the NusG carboxy-terminal domain containing residues necessary for Nun function. These findings indicate that Nun ARM relies on distinct contacts in its ternary complex and illustrate how protein-RNA recognition can evolve new regulatory functions.

IMPORTANCE

λ N protein interacts with host factors to allow λ nut-containing transcripts to elongate past termination signals. A competing bacteriophage, HK022, expresses Nun protein, which causes termination of λ nut transcripts. λ N and HK022 Nun use similar arginine-rich motifs (ARMs) to bind the same boxB RNAs in nut transcripts. Screening libraries of Nun ARM mutants, both in HK022 Nun and in a λ N fusion, revealed amino acids essential to Nun that could bind one or more host factors. Docking suggests that NusG, which is present in both Nun termination and N antitermination, is a plausible partner. These findings could help understand how transcription elongation is regulated and illustrate how subtle differences allow ARMs to evolve new regulatory functions.

The switch to delayed early gene expression in λ, P22, ϕ21, and other lambdoid bacteriophages depends on N proteins assembling antitermination complexes at nut (N utilization) sites on P_{left} and P_{right} transcripts that transcribe past downstream terminators (1–3). N proteins bind via their arginine-rich motifs (ARMs) to boxB hairpin RNAs in nut sites of P_{left} and P_{right} transcripts (Fig. 1) (4, 5) and interact with host factors in the transcription elongation complex, including NusA, NusB, NusE, and NusG (6–8). N antitermination has been extensively studied, and detailed biophysical (9–11), mechanistic (12–14), and structural models are published (15–17).

In competition with bacteriophage λ, HK022 uses its Nun protein to suppress the replication of coinfecting λ by premature termination of λ P_{left} and P_{right} transcripts (18–21). Similarly to N, Nun binds λ boxBs via its ARM in an elongation complex that also includes NusA, NusB, NusE, and NusG (6, 17, 20, 22–25). Based on Nun's conserved ARM sequence, its inability to exclude P22 and ϕ21 infections (18, 26), the affinity of Nun ARM-boxB *in vitro* (27, 28), and the similarity of Nun's ARM-boxB NMR structural model (29) to those of λ N (15, 16), the recognition strategy of the HK022 Nun ARM-boxB interaction has been assumed to be very similar or identical to that of the λ N-boxB interaction, though few Nun ARM mutants have been examined.

N antitermination in λ, P22, and ϕ21 is type specific: N proteins of one virus do not complement its absence in another (30, 31), and boxBs bind noncognate N ARM peptides poorly *in vitro* (32–34). The RNA binding ARMs of HK022 Nun, λ N, P22N, and ϕ21 N bind as α helices in the major grooves of boxBs and make contacts to the 5′-stems backbone and loops, but few base-specific contacts are made. λ, P22, and ϕboxBs hairpins adopt similar, GNRA tetraloop-like conformations in which loop bases stack upon the apical base pair of the stem. λ boxBs adopt a 4-out GNRA-like pentaloop (15, 16), P22 boxBs adopt a 3-out GNRA-like pentaloop (35), and ϕ21 boxBs adopt a U-turn that mimics the GNRA fold (36). Mutagenesis, biochemical, and nuclear magnetic resonance (NMR) data are consistent with type specificity arising from each N ARM recognizing a specific conformation of

Received 9 June 2015 **Accepted** 24 August 2015

Accepted manuscript posted online 8 September 2015

Citation Tawk CS, Ghattas IR, Smith CA. 2015. HK022 Nun requires arginine-rich motif residues distinct from λ N. J Bacteriol 197:3573–3582. doi:10.1128/JB.00466-15.

Editor: R. L. Gourse

Address correspondence to Colin A. Smith, colin.smith@aub.edu.lb.

* Present address: Caroline S. Tawk, Institute for Molecular Infection Biology, University of Würzburg, Würzburg, Germany.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.00466-15.

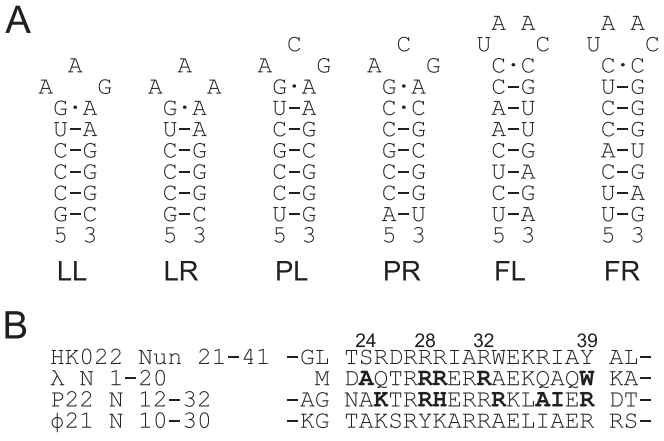

FIG 1 Comparison of lambdoid phage boxBs and partner ARMs. (A) The secondary structures of boxBs of λ, P22, and ϕ21 are shown from the base of the stem. LL, λ $boxB_{left}$; LR, λ $boxB_{right}$; PL, P22 $boxB_{left}$; PR, P22 $boxB_{right}$; FL, ϕ21 $boxB_{left}$; FR, ϕ21 $boxB_{right}$. Watson-Crick and wobble base pairs are connected with dashes and noncanonical base pairs connected with dots. (B) Alignment of ARMs of HK022 Nun, λ N, P22 N, and ϕ21 N proteins. The ARMs are separated by spaces for clarity. Numbers indicate residue positions in HK022 Nun. Residues indicated in bold type are those found to be important for function in previous studies (37–39, 46). Libraries of Nun ARM randomized at single residues Leu22 to Tyr39 were screened for active mutants with λ $boxB_{left}$ and λ $boxB_{right}$.

its cognate boxB loop. Mutational studies of λ N (37, 38) found λ Ala3, Arg7, Arg8, Arg11, and Trp18 critically important for N-boxB recognition (Fig. 1B). Mutational analysis of P22 N found a different pattern of requirements (39), and no mutational analysis of ϕ21 has been published.

The HK022 N ARM-boxB NMR model resembles λ N ARM-boxB NMR models (Fig. 2). Nun ARM makes few hydrogen bonds to bases and recognizes the 4-out GNRA-like pentaloop and backbone conformation of λ boxB to achieve specificity (29). Analogously to λ N at equivalent positions, Nun Ser24 nestles against the hydrophobic surface of C2 and C3 bases and riboses, Nun Arg28 hydrogen contacts the phosphate of U5 and base G6, Nun Arg29 hydrogen bonds to base A8, Nun Arg32 contacts phosphate of U5, and Nun Tyr39 stacks upon A7. Remaining residues make nonessential contacts to boxB, contact other Nun residues, or project away from the RNA. The interaction of the aromatic amino acid Nun Tyr39 on stacked bases in the boxB loop appears similar to that of λ N Trp18 (28, 29). Interestingly, the λ N Trp18 stacking interaction is not important for affinity *in vitro* but may stabilize a specific conformation in the elongation complex necessary for efficient antitermination (40–43).

Limited mutagenesis of Nun ARM has been reported (28), and the HK022 Nun-boxB structural model is less resolved than those of λ N-boxB (15, 16, 29). λ N ARM shares about equal numbers of identical residues with P22 and ϕ21 N ARMs as it does with HK022 Nun ARM (Fig. 1B). Faber et al. (29) note that HK022 Nun residues Leu22, Ile30, Trp33, Ile37, and Leu41 form a hydrophobic surface that is not present in λ N ARM, and they suggest it could be a recognition site for host factors. Thus, despite the similarity of N and Nun ARMs, Nun ARM may rely on a distinct recognition strategy to bind boxB productively (29).

To determine how the recognition strategy of HK022 Nun-boxB could differ from that of λ N-boxB, we constructed plasmid libraries expressing Nun with all single substitutions in its ARM, screened those libraries for members that exclude λ infection, identified functional mutants, and assayed Nun mutants in a termination reporter system (see Fig. S1 in the supplemental material). We found that Nun ARM requirements are similar to those known for N ARM, though Ser24 tolerates diverse substitutions. Unexpectedly, we found Nun Asp26, Trp33, and Arg36 to be immutable. These residues project from boxB and are aligned with other residues of the ridge identified by Faber et al. (29), support-

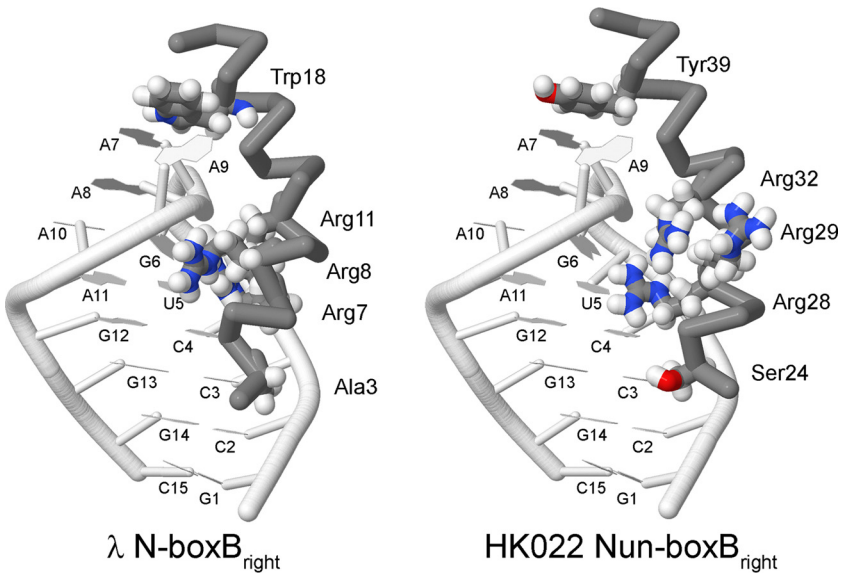

FIG 2 Comparison of NMR structural models highlighting important N residues and equivalent Nun residues. RNAs are rendered as white cartoons, and bases are labeled by single-letter code and position. ARMs are rendered with backbones in gray. Amino acids known to be important for λ N function (37, 38) and assumed by equivalence to be important in HK022 Nun are rendered as sticks, colored by atom type, and labeled by three-letter codes. λ N-boxB, PDB accession number 1QFQ (16); HK022 Nun-boxB, PDB accession number 1HJI (29).

ing the idea that N-boxB and Nun-boxB rely on different contacts to host factors (29). Screening the same Nun ARM libraries as λ N fusions revealed that a similar panel of functional mutants are able to complement N-deficient λ. Surprisingly, Nun ARM Asp26 and Trp33 are still immutable in the context of λ N fusions, which suggests that Nun ARM depends on putative host factors to recognize boxB. Rigid-body docking suggests that the host factor NusG carboxy-terminal domain (NusG-CTD) contacts Nun ARM, consistent with a NusG surface identified by mutagenesis (25). Our findings should contribute to understanding Nun termination, N antitermination, the mechanisms by which transcription elongation is regulated, and how protein-RNA recognition can evolve new regulatory functions.

MATERIALS AND METHODS

General. Laboratory chemicals were obtained from Acros (Belgium) and Affymetrix (United States). Disposable plasticware was obtained from Sarstedt (Germany). Fine chemicals were obtained from Amersham (United Kingdom), Amresco (United States), and Sigma (United States). Restriction enzymes and T4 DNA ligase were obtained from Roche (Germany). Bacterial medium components were obtained from Oxoid (United Kingdom) and HiMedia (India). General procedures, cloning, plaque assays, and reporter enzyme assays followed standard methods (44).

Bacterial strains, bacteriophages, plasmids, and DNA. *Escherichia coli* strain N567, supporting antitermination (45), lytic, N-deficient λ phages with immunity regions of λ (phage λ ClearNam7am53) (46) and P22 (phage λimm^{22}24am^{clr}) (46), and λ N-expressing plasmid pBR-ptac-N*λ (referred to below as λ N) (37) were obtained from Naomi Franklin (University of Utah). Wild-type λ phage (DSM no. 4499, ATCC 23724-B2), referred to below as λ, was obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (Germany). pBAD-GFPuv (GenBank accession number U62637.1) was obtained as pGLO from Bio-Rad (United States). pBRNP22N12-30 (referred to below as P22 N) (39), pBRNRev17 (referred to below as RevN) (47), and boxB reporters pACLL (referred to below as λ boxB$_{left}$) (48), pACLR (referred to below as λ boxB$_{right}$) (48), pACPL (referred to below as P22 boxB$_{left}$) (48), pACLR (referred to below as P22 boxB$_{right}$) (48), and pACRREIIB (referred to below as RRE IIB) (47) were available in-house. All boxB reporter plasmids are replacements of λ boxB in the λ nut$_{left}$ site (48). Long synthetic double-stranded DNA sequences were obtained from Epoch Biolabs (United States). DNA oligonucleotides were obtained from TIB Molbiol (Germany) and Sigma (United States).

Construction of Nun- and N-supplying plasmids and libraries. Plasmids expressing HK022 wild-type Nun (pBADwtNun, referred to below as wtNun) (see Fig. S1 in the supplemental material) and cassette Nun (pBADcasNun, referred to below as casNun) (see Fig. S1 and S2 in the supplemental material) were constructed in several steps from pBAD-GFPuv, an ampicillin-resistant, pBR322 ori plasmid with a P$_{BAD}$ promoter (49). The entire AgeI (bp 1074)-HindIII (bp 2114) region containing *araC*, P$_{BAD}$, and the green fluorescent protein (GFP) open reading frame (ORF) was replaced with synthetic sequences that ablated the BsmI site in the *araC* leader, replaced the NdeI-NheI at the amino-terminal methionine of pBAD-GFPuv with NcoI, and replaced the GFP ORF with that of HK022 Nun to make wtNun. casNun is identical to wtNun except for an NcoI-BsmI ARM mutagenesis cassette created by silent mutation of Ser18, insertion of a methionine after Ser18 and before Asp19 to form an NcoI site, and insertion of an asparagine and alanine after Ala40 and before Leu41 to form a BsmI site. All mutated regions of wtNun and casNun sequences were confirmed by sequencing.

The plasmid expressing casNun with its ARM replaced by HIV Rev ARM, pBADRevNun (referred to below as RevNun), was constructed by ligating a double-stranded DNA formed from Rev17NF (5′-C ATG GCA ACC CGC CAG GCC CGT CGT AAC CGT AGA CGT CGT TGG CGT GAG CGT CAG CGT GCA GCT GCG GCG AAT GCA-3′) and Rev17NR (5′-C ATT CGC CGC AGC TGC ACG CTG ACG CTC ACG CCA ACG ACG TCT ACG GTT ACG ACG GGC CTG GCG GGT TGC-3′) encoding the HIV RevARM as an NcoI-BsmI fragment into NcoI-BsmI-digested casNun.

To construct all possible Nun single mutants spanning residues 22 to 39, 18 double-stranded DNA libraries, each containing one completely randomized codon, were formed by primer extension and NcoI-BsmI digestion of synthetic oligonucleotides based on the sequence of NunF (5′-GCG CCC ATG GAT AGA GGT CTT ACA TCT CGA GAC AGG AGG AGA ATA GCG AGA TGG GAA AAA AGG ATA GCA TAT GCG AAT GCA AAT CCC-3′) and NunR (5′-GGG ATT TGC ATT CGC ATA TGC TAT CCT TTT TTC CCA TCT CGC TAT TCT CCT CCT GTC TCG AGA TGT AAG ACC TCT ATC CAT GGG CGC-3′), in which the sequence after digestion is underlined. The oligonucleotide libraries were ligated to NcoI-BsmI-cut RevNun and RevN backbones and transformed into chemically competent N567. Construction of specific mutants of casNun relied on annealing complementary oligonucleotides based on the sequence of NunF2 (5′-C ATG GAT AGA GGT CTT ACA TCT CGA GAC AGG AGG AGA ATA GCG AGA TGG GAA AAA AGG ATA GCA TAT GCG AAT GCA-3′) and NunR2 (5′-C ATT CGC ATA TGC TAT CCT TTT TTC CCA TCT CGC TAT TCT CCT CCT GTC TCG AGA TGT AAG ACC TCT ATC-3′).

Construction of the λN-Nun fusion and its mutant λN-T5D;A12W;Q15R (see Table 2) were based on the sequences of alignNF (5′-C ATG AGA GGT ATG GAT GCA CAA ACA AGG AGG AGA GAA AGG AGA GCA GAA AAA CAA GCA CAG TGG GCA-3′) and alignNR (5′-C CCA CTG TGC TTG TTT TTC TGC TCT CCT TTC TCT CCT CCT TGT TTG TGC ATC CAT ACC TCT-3′).

Construction of pBRNunN (referred to below as NunN) and its specific mutants relied on annealing complementary oligonucleotides based on the sequence of NunF2 (5′-C ATG GAT AGA GGT CTT ACA TCT CGA GAC AGG AGG AGA ATA GCG AGA TGG GAA AAA AGG ATA GCA TAT GCG AAT GCA-3′) and NunR2 (5′-C ATT CGC ATA TGC TAT CCT TTT TTC CCA TCT CGC TAT TCT CCT CCT GTC TCG AGA TGT AAG ACC TCT ATC-3′) ligated into λ N.

DNA preparation and sequencing. Plasmid DNA was prepared by standard procedures, purified by columns, and sent to Macrogen Inc. (Korea) for sequencing. Macrogen's primer pBAD-R (5′-GAT TTA ATC TGT ATC AGG-3′) was used to sequence all pBAD plasmids (expressing Nun mutants and fusions). Macrogen's primer pQE-F (5′-CCC GAA AAG TGC CAC CTG-3′) was used to sequence all pBRN plasmids (expressing λ N and fusions). Sequences of cloned regions were confirmed using Chromas Lite software from Technelysium (Australia).

Phage exclusion assays and screening libraries for Nun function. Competent N567 host cells were transformed with plasmid libraries or clones. Individual clones were used to grow overnight cultures in tryptone with ampicillin selection and aeration at 37°C. Arabinose was not used to induce pBAD expression, as weak growth resulted, presumably due to the known toxicity of Nun (50). Culture density at 600 nm was measured and adjusted to 2.0 with washing and resuspension in an appropriate volume of 10 mM $MgSO_4$. A 50-μl sample of each culture was mixed with approximately 100 PFU of λ phage in 50 μl SM (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-HCl [pH 7.5], 0.1 g/liter gelatin). Cells and virus were incubated at 37°C for 20 min for λ adsorption, then 1.2 ml tryptone top agar was added to each tube, and the content was poured onto 5-cm tryptone agar plates. The plates were incubated at 37°C for 9 to 11 h. Plaque formation or lack thereof was assessed by comparison to controls. During functional screening, at least 100 clones were tested for each library. Clones excluding λ were restreaked and retested, and individual clones were used to prepare plasmid, which was sequenced.

Reporter assays of Nun termination. Functional isolates, synthetic constructs, and control plasmids wtNun, casNun, and RevNun were used to transform N567 cells carrying λ boxB$_{left}$, λ boxB$_{right}$, P22 boxB$_{left}$, P22 boxB$_{right}$, and RRE IIB reporter plasmids. For each interaction, at least

TABLE 1 HK022 Nun excludes λ and terminates transcription

Nun supplier plasmid[a]	λ plaques[b]	*phoA* units[c] in:				
		λ $boxB_{left}$	λ $boxB_{right}$	P22 $boxB_{left}$	P22 $boxB_{right}$	RRE IIB[d]
wtNun	N	49 ± 11	45 ± 6	128 ± 14	5,200 ± 600	14,200 ± 500
casNun	N	310 ± 30	380 ± 20	3,800 ± 500	11,800 ± 1,300	14,000 ± 2,000
RevNun	Y	15,000 ± 1,200	14,700 ± 1,500	8,400 ± 900	15,000 ± 1,600	15,400 ± 800

[a] wtNun (pBADwtNun) expresses wild-type HK022 Nun protein; casNun (pBADcasNun) expresses Nun with an NcoI-BsmI ARM mutagenesis cassette created concomitantly with insertion of a methionine before Asp19 and an asparagine and alanine after Ala40 (see Fig. S1 in the supplemental material); RevNun has HIV Rev residues 34 to 50 flanked by alanines (A TRQARRNRRRRWRERQR AAAA) replacing the ARM of casNun and was used as a negative control. Arabinose induction was not used, because arabinose-containing cultures grew slowly, presumably due to the toxicity of Nun when overexpressed (50).

[b] *E. coli* N567 cells hosting Nun supplier plasmids were infected with ~100 PFU λ; Y, plaques were observed; N, no plaques were observed.

[c] Reporter plasmid RNAs are boxB replacements in the context of the λ nut_{left} site. Nun supplier plasmids were transformed into N567 cells hosting boxB reporter plasmids (37) in which *phoA* is downstream of nut sites. Alkaline phosphatase activity was measured with PNPP to determine *phoA* reporter gene expression after overnight growth in tryptone at 34°C with aeration and with 100 μM IPTG to induce reporter plasmids. *phoA* units represent the amount of transcription. HK022 Nun termination reduces *phoA* expression. Values are averages ± standard deviations for three or more replicates from a representative experiment.

[d] RRE is HIV RRE IIB in place of boxB and was used as a negative control. The RevNun-RRE interaction did not display Nun termination activity.

three separate colonies were used to inoculate separate cultures of 3 ml tryptone medium with 100 μg/ml ampicillin, 11 μg/ml chloramphenicol, and 100 μM IPTG (isopropyl-β-D-thiogalactopyranoside) (to induce boxB reporter plasmids). Cultures were grown at 34°C with aeration for 12 h. Each culture was diluted 2-fold with sterile water, washed twice with resuspension buffer (10 mM Tris-HCl [pH 8.0], 0.1 M NaCl), and resuspended in 1 ml resuspension buffer. Optical densities of washed cultures were measured, and 10-μl samples were permeabilized in reaction buffer (1 M Tris-HCl [pH 8.0]) with 60 μl chloroform and 30 μl 0.1% sodium dodecyl sulfate and vortexing. Alkaline phosphatase activity was measured using 4 mg/ml *p*-nitro-phenylphosphate (PPNP) at 28°C and 100 μl cold 1 M K_2HPO_4 for quenching. The number of alkaline phosphatase units was calculated according to Brickman and Beckwith (51).

Phage N complementation assays and library screening. N complementation assays were performed as described previously (39), using clear strains of N-deficient λ phage, in which the immunity region was from either λ or P22. Competent N567 host cells were transformed with N-supplier plasmids. As described above, cultures were assayed with approximately 100 PFU of N-deficient phage. When indicated, 300 μM IPTG was added to bottom agar to induce N expression. Plaque formation was assessed by comparison to controls, including λ N and NunN. During functional screening, 50 to 150 clones of each NunN library were tested. Those clones complementing N-deficient λ with IPTG induction were restreaked and retested, and individual clones were used to prepare plasmids for sequencing.

Reporter assays of N antitermination. For each ARM-RNA interaction, representative colonies were picked from X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) plates for use in assays. At least three independent colonies were used for each interaction. For measurement of N-mediated antitermination, cultures were grown overnight at 30°C with aeration in tryptone with 100 μg/ml ampicillin, 11 μg/ml chloramphenicol, and 100 μM IPTG (to induce N supplier and boxB reporter plasmids). The cells were then permeabilized, the β-galactosidase activity was assayed using *o*-nitrophenol-D-galactoside (ONPG), and the β-galactosidase activity was calculated by using the method of Miller (52). The activities were normalized using NunN for $boxB_{left}$ and $boxB_{right}$ reporters.

Structure visualization. Jmol, an open-source Java viewer for chemical structures in three dimensions (http://www.jmol.org/, accessed 1 January 2015), and the PyMOL Molecular Graphics System (Schrodinger, LLC) were used to view the solution state NMR models of λ N peptide-λ $boxB_{right}$ (PDB identifier 1QFQ) (16), λ N peptide-λ $boxB_{left}$ (15), HK022 Nun peptide-λ $boxB_{right}$ (PDB identifier 1HJI) (29), *E. coli* NusG carboxy-terminal domain (PDB identifier 2JVV) (25), and the results of docking.

Molecular docking. ZDOCK (53, 54) was used to dock NusG-CTD and the Nun ARM-boxB complex. The model was the result of 3 repeated unsupervised docking simulations using default parameters (see the supplemental material for the Protein Data Bank [PDB] file). PyDock (55), a rigid-body docking and scoring algorithm, was used to corroborate scoring of ZDOCK models. Protein surfaces are colored according to the Eisenberg hydrophobicity score (56). The W33A mutation in Nun was introduced using the crystallographic object-oriented toolkit (Coot-v0.82) (57).

RESULTS

We first constructed a Nun expression plasmid, wtNun (pBADwtNun), to express full-length, wild-type HK022 Nun. When transformed with this plasmid, *E. coli* N567 (45), a strain permissive to λ replication, excluded λ (Table 1). To facilitate cassette mutagenesis of the Nun ARM, restriction sites NcoI and BsmI were introduced to flank the Nun ARM, resulting in casNun (pBADcasNun), which has a methionine inserted before Asp19 and an asparagine and alanine after Ala40 (see Fig. S1 in the supplemental material). This plasmid also excluded λ.

To quantify Nun activity, we used an existing reporter plasmid series constructed for assaying N antitermination (37), in which the *tac* promoter drives a transcript with a λ nut_{left} site followed by a *phoA* (alkaline phosphatase) reporter gene, transcriptional terminators, and finally, a LacZ (β-galactosidase) reporter gene (see Fig. S1 in the supplemental material). The plasmid reports on N antitermination through downstream terminators by β-galactosidase expression, which can be quantified in cell extracts with ONPG. Nun termination activity was monitored by reduction of alkaline phosphatase expression in cell extracts with PNPP. wtNun expression reduced alkaline phosphatase activity about 300-fold, and casNun reduced alkaline phosphatase activity about 40-fold (Table 1). Neither wild-type Nun nor cassette Nun had significant activity on a reporter in which the noncognate HIV-1 RRE IIB replaced boxB in λ nut_{left}. Similar to λ N, wtNun has strong activity with P22 $boxB_{left}$ but weak activity with P22 $boxB_{right}$ (48).

Nun libraries D26X, R28X, R29X, R32X, W33X, R36X, and Y39X have few active members. Using synthetic oligonucleotides randomized at individual codons, 18 libraries from Leu22 to Tyr39 were constructed in casNun. The libraries were transformed into *E. coli* separately, the resulting colonies were pooled, and plasmid DNA libraries were prepared. The DNA libraries were transformed into host cells, and at least 100 individual colonies were grown as overnight cultures that were then tested for plaque formation with λ. The proportion of each library's clones with Nun function varied from a few percentage points to greater

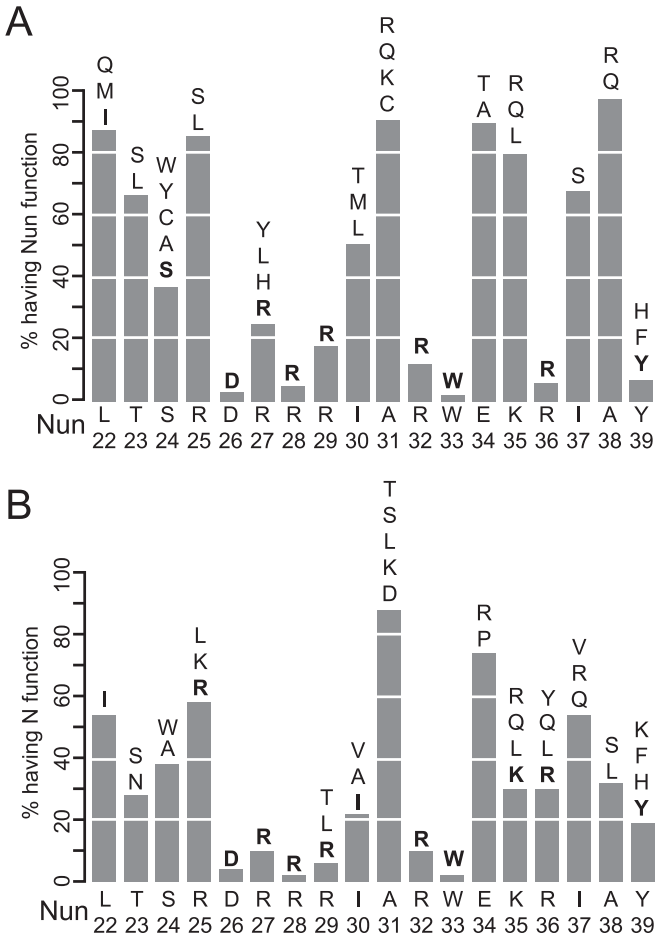

FIG 3 Proportion of active mutants in Nun ARM libraries. (A) At least 100 randomly selected clones of each Nun library were tested for the ability to exclude λ infection. Libraries are described by the amino acid identity and position in Nun. Bar height indicates the proportion of clones found to be functional in each library. The identities of sequenced, active mutants from selections are shown above each bar. Bold letters indicate that the wild-type residue was recovered. (B) Similar to panel A. At least 50 randomly selected clones of Nun ARM libraries fused to λ N activation domain were tested for the ability to complement N-deficient λ with 300 μM IPTG induction. Bold letters indicate that the Nun wild-type residue was recovered.

TABLE 2 Nun ARM mutants exclude virus and terminate transcription

Nun[a]	λ plaques[b]	Fraction termination[c]	
		λ $boxB_{left}$	λ $boxB_{right}$
casNun	N	0.979 ± 0.003	0.974 ± 0.003
RevNun	Y	0.00 ± 0.10	0.00 ± 0.14
L22I	N	0.971 ± 0.006	0.96 ± 0.006
L22 M	N	0.941 ± 0.015	0.89 ± 0.012
L22Q	N	0.91 ± 0.04	0.852 ± 0.016
T23L	N	0.971 ± 0.005	0.932 ± 0.010
T23S	N	0.961 ± 0.007	0.923 ± 0.006
S24A	N	0.979 ± 0.004	0.964 ± 0.004
S24W[d]	N	0.923 ± 0.011	0.86 ± 0.02
R25L	N	0.960 ± 0.005	0.891 ± 0.018
R25S	N	0.963 ± 0.008	0.940 ± 0.010
D26T[d]	Y	0.01 ± 0.08	0.08 ± 0.19
R27H	N	0.88 ± 0.02	0.84 ± 0.03
R27L	N	0.79 ± 0.10	0.79 ± 0.05
R27Y	N	0.908 ± 0.013	0.81 ± 0.03
R28K[d]	Y	−0.01 ± 0.12	−0.1 ± 0.2
R29K[d]	Y	0.67 ± 0.07	0.08 ± 0.15
R29Q[d]	Y	0.67 ± 0.05	0.01 ± 0.15
I30L	N	0.951 ± 0.018	0.937 ± 0.008
I30 M	N	0.942 ± 0.014	0.915 ± 0.012
I30T	N	0.90 ± 0.02	0.885 ± 0.013
A31C	N	0.960 ± 0.007	0.909 ± 0.011
A31K	N	0.983 ± 0.003	0.953 ± 0.006
A31Q	N	0.975 ± 0.004	0.947 ± 0.012
A31R	N	0.985 ± 0.002	0.968 ± 0.003
R32K[d]	Y	0.51 ± 0.08	−0.2 ± 0.2
W33A[d]	Y	0.34 ± 0.06	−0.1 ± 0.3
E34A	N	0.9742 ± 0.0016	0.964 ± 0.005
E34T	N	0.9706 ± 0.0012	0.958 ± 0.006
K35L	N	0.956 ± 0.004	0.951 ± 0.007
K35Q	N	0.928 ± 0.007	0.946 ± 0.006
K35R	N	0.9719 ± 0.0018	0.968 ± 0.006
R36K[d]	Y	0.72 ± 0.05	0.1 ± 0.2
R36Q[d]	N	0.83 ± 0.019	0.73 ± 0.04
I37S	N	0.970 ± 0.004	0.948 ± 0.011
A38Q	N	0.9582 ± 0.0016	0.942 ± 0.016
A38R	N	0.954 ± 0.006	0.934 ± 0.018
Y39A[d]	Y	−0.22 ± 0.09	−0.3 ± 0.4
Y39F	N	0.950 ± 0.005	0.93 ± 0.03
Y39G[d]	Y	0.76 ± 0.05	0.3 ± 0.3
Y39H	N	0.929 ± 0.011	0.86 ± 0.05
λ N[e]	Y	0.18 ± 0.15	0.20 ± 0.17
λ N-T5D;A12W;Q15R[f]	N	0.871 ± 0.018	0.40 ± 0.04

[a] All Nun mutants are of casNun and are listed by single-letter codes showing the wild-type residue and the mutant residue separated by the numerical position in Nun.
[b] As described in Table 1, N567 cells expressing Nun mutants were infected with λ; Y, plaques were observed; N, no plaques were observed.
[c] Nun constructs were assayed as described for Table 1. Fraction termination = 1 − (phoA units Nun/phoA units RevNun). Low values represent low Nun termination activity.
[d] This clone is a synthetic construct made for confirmation of activity or to assay values of mutants not isolated by selection.
[e] λ N ARM residues 1 to 18, MDAQTRRRERRAEKQAQW, replacing Nun ARM residues 19 to 39, DRGLTSRDRRRIARWEKRIAY, in pBADwtNun with a substitution of Nun D19S for cloning purposes.
[f] As explained in footnote *e*, with λ N ARM residues 1 to 18 containing T5D, A12W, Q15R, the equivalent of Nun D26, W33, and R36, MDAQ*D*RRRERR*W*EK*R*AQW (where italics indicate positions of T5D, A12W, and Q15R), replacing Nun ARM residues 19 to 39, DRGLTSRDRRRIARWEKRIAY, in pBADwtNun with a substitution of Nun D19S.

than 80% (Fig. 3A); the proportion should reflect the lack of importance of that residue. If HK022 Nun had the same requirements as λ N (37, 38), residues Ser24, Arg28, Arg29, Arg32, and Tyr39 would tolerate few substitutions. As expected, R28X, R29X, R32X, and Y39X libraries appeared to have few active members, yet surprisingly, D26X, W33X, and R36X libraries also had few active members and S24X had a relatively high proportion of active members.

Nun requires Asp26, Trp33, and Arg36. DNAs of several active clones from each library were prepared and sequenced (Fig. 3A). Sequencing results were consistent with the proportion of active library members: only wild-type residues were found as active members of D26X, R28X, R29X, R32X, W33X, and R36X, and tyrosine and similar aromatic residues capable of base stacking on the boxB loop were found in Y39X. A wide variety of residues, including tryptophan, were able to replace Ser24. Other libraries had unremarkable substitutions.

TABLE 3 HK022 ARM fused to the λ N activation domain complements N-deficient λ virus and leads to antitermination of transcription

N supplier plasmid[a]	λ N⁻ plaques[b]		P22 N⁻ plaques[b]		β-Galactosidase units[c]				
	Uninduced	Induced	Uninduced	Induced	λ $boxB_{left}$	λ $boxB_{right}$	P22 $boxB_{left}$	P22 $boxB_{right}$	RRE IIB[d]
λ N	Y	Y	N	N	1,530 ± 190	1,540 ± 110	160 ± 30	1.7 ± 0.7	2.06 ± 0.09
P22 N	N	N	Y	Y	3.6 ± 0.4	3.5 ± 0.8	11 ± 2	12.1 ± 1.0	1.98 ± 0.12
NunN	N	Y	N	N	188 ± 10	75 ± 9	4.3 ± 0.4	2.0 ± 0.5	2.1 ± 0.4
RevN[d]	N	N	N	N	2.23 ± 0.16	2.27 ± 0.16	2.31 ± 0.17	2.53 ± 0.08	85 ± 10

[a] N supplier plasmids are of pBR322 origin and express the indicated ARM as replacement of wild-type λ N residues 2 to 18 with a K19N mutation. P22 N expresses P22 N ARM residues 12 to 30 replacing λ N residues 2 to 18. NunN expresses Nun ARM residues 19 to 40 replacing λ N residues 2 to 18. RevN expresses HIV Rev ARM residues 34 to 50 flanked by alanines (A TRQARRNRRRRWRERQR AAAA) replacing λ N residues 2 to 18.

[b] N-deficient λ was used to infect N567 cells hosting the indicated N protein without or with 300 μM IPTG induction. Y, plaques were observed; N, no plaques were observed. N-deficient P22 is λ with its immunity region replaced with that of P22, and it includes an N amber mutant and nut sites.

[c] N supplier plasmids were transformed into N567 cells hosting the same boxB reporters as for Table 1 that have a LacZ reporter gene downstream from transcriptional terminators. β-Galactosidase units were assayed with ONPG from 16-h cultures grown in tryptone medium at 30°C from individual colonies, with aeration and continuous 100 μM IPTG induction. Values are averages ± standard deviations for three or more replicates from a representative experiment.

[d] The HIV Rev ARM-RRE IIB interaction is used as a heterologous control.

To confirm the restricted or relaxed mutability of interesting residues, specific mutants were constructed and assayed for λ exclusion and transcription termination, including those with the following substitutions: S24W, D26T, R28K, R29K, R29Q, R32K, W33A, R36K, R36Q, Y39A, and Y39G (Table 2). These substitutions were chosen on the basis of biochemical conservation (arginine to lysine or glutamine) or difference (serine to tryptophan and tyrosine to alanine or glycine), or the residue in question was replaced by the residue found at the equivalent position of λ N (D26T, W33A, and R36Q). The results confirm the interpretations illustrated in Fig. 2A. The absence or limited mutability of Arg28, Arg29, Arg32, and Tyr39 is consistent with Nun ARM interacting with boxB similarly to λ N. That Nun Ser24 is mutable to a much larger tryptophan suggests that the ARM-RNA interactions of Nun and N are not identical, consistent with average NMR structures showing the relative projection of Nun Ser24 away from the 5′ half of boxB compared to the intimate contact of λ N Ala3 with C2 and C3 of boxB. Intriguingly, mutants at Arg29, Arg32, Trp33, and Arg36 show a strong bias toward $boxB_{left}$.

Asp26, Trp33, and Arg36 are not conserved with λ N. Asp26 and Trp33 would not be expected to contribute directly to binding boxB, as the NMR model shows that they project from boxB and there are few likely RNA-binding roles for acidic and hydrophobic residues. Similarly to λ N Gln15, Arg36 is within hydrogen-bonding distance of the extruded base of the boxB loop, yet it also presents hydrogen-bond donors to the exterior of the complex. In particular, the large unburied hydrophobic surface of Trp33 suggests a complementary hydrophobic partner molecule. We considered that Asp26 and Arg36 could be part of the putative host factor recognition site composed of Nun Leu22, Ile30, Trp33, Ile37, and Leu41 that was identified by Faber et al. (29).

The affinities of Nun and N ARM peptides to boxB *in vitro* are similar (27, 28, 58). Reasoning that the importance of Nun's putative host factor binding site could be shown by replacing its ARM with that of λ N, we tested the activity of an N-Nun fusion (Table 2, λ N). Its activity dramatically increased with mutation to the equivalent of Nun Asp26, Trp33, and Arg36 (Table 2). This is consistent with Nun Asp26, Trp33, and Arg36 contributing to a putative host factor recognition site required for Nun function.

Nun ARM fused to the N activation domain complements N-deficient λ phage and functions in an N antitermination assay. Reasoning that distinct roles of Nun ARM residues in facilitating termination could be clarified by examining Nun ARM in the related context of λ N antitermination, we constructed plasmid NunN expressing Nun ARM as an amino-terminal fusion to λ N activation domain (see Fig. S1 in the supplemental material). This allows the assay of Nun ARM-boxB interaction by complementation of N-deficient λ and by antitermination reporter assay monitoring β-galactosidase activity with ONPG (37).

When induced with IPTG, NunN complemented N-deficient λ (Table 3) but not N-deficient P22 (phage λimm^{22}24am^{clr}). Though strongly reduced, NunN displayed specific, N-antitermination activity in the antitermination assay, similar to the activity of HIV Rev ARM-N fusion on a reporter in which HIV RRE IIB replaces boxB of λ nut_{left}. NunN did not have significant activity on P22 boxBs or HIV RRE reporters (Table 3).

Library screening reveals that Nun ARM preserves most requirements when fused to λ N. The same Nun ARM libraries as those prepared for Nun function were constructed as NunN libraries and screened for the ability to complement N-deficient λ (Fig. 3B). Unexpectedly, very similar patterns of proportions of functional library members were seen with NunN antitermination and with Nun termination, suggesting that the Nun ARM-boxB interaction relies on the same recognition strategy in its ternary complex with host factors regardless of whether part of a Nun termination or an N antitermination complex. Sequencing results were largely consistent with the proportion of active library members: only wild-type residues were found as active members of D26X, R27X, R28X, R32X, and W33X, and tyrosine and similar aromatic or hydrophobic substitutions capable of capping the boxB loop were found substituting for Tyr39 (Fig. 3B). Similar to Nun, other libraries had unremarkable diversities of active substitutions. To confirm Nun ARM requirements in the N fusion, the same specific Nun ARM mutants were constructed and tested in NunN. Though there are some differences between Nun ARM substitutions functional in Nun and NunN, the requirements of Asp26 and Trp33 and the tolerance to S24W are very similar to Nun ARM-boxB recognition.

Differences between Nun and N contexts. Despite the similarity in the panel of Nun ARM mutants having function whether as part of Nun or when fused to λ N activation domain, some differences suggest that the Nun ARM-boxB recognition strategy could be subtly altered by the N antitermination complex. Notably, Arg27 becomes immutable and Arg36 mutable in the context

TABLE 4 Complementation and antitermination assays of Nun ARM fused to the λ N activation domain

NunN fusion[a]	λ N⁻ plaques[b]	Fraction β-galactosidase activity[c]	
		λ $boxB_{left}$	λ $boxB_{right}$
NunN	Y	1.00 ± 0.07	1.00 ± 0.15
RevN	N	0.012 ± 0.001	0.03 ± 0.004
L22I	Y	2.14 ± 0.19	2.8 ± 0.3
T23S	Y	0.29 ± 0.07	0.63 ± 0.13
S24A	Y	0.93 ± 0.19	0.22 ± 0.10
S24W[d]	Y	0.27 ± 0.09	0.27 ± 0.06
R25K	Y	0.4 ± 0.2	0.62 ± 0.13
R25L	Y	0.46 ± 0.06	1.4 ± 0.2
D26T[d]	N	0.08 ± 0.01	0.07 ± 0.02
R28K[d]	N	0.05 ± 0.008	0.049 ± 0.011
R29K[d]	Y	0.26 ± 0.11	0.13 ± 0.03
R29L	Y	2.0 ± 0.5	0.39 ± 0.13
R29Q[d]	N	0.13 ± 0.03	0.10 ± 0.04
R29T	Y	0.21 ± 0.05	0.17 ± 0.06
I30A	Y	0.24 ± 0.06	0.79 ± 0.09
A31D	Y	0.26 ± 0.06	0.67 ± 0.06
A31K	Y	0.9 ± 0.3	1.5 ± 0.3
R32K[d]	N	0.098 ± 0.014	0.068 ± 0.019
W33A[d]	N	0.14 ± 0.02	0.14 ± 0.05
E34P	Y	0.5 ± 0.2	1.7 ± 0.6
E34R	Y	0.5 ± 0.2	1.25 ± 0.10
K35L	Y	0.15 ± 0.05	0.54 ± 0.05
K35R	Y	0.30 ± 0.11	0.75 ± 0.08
R36K[d]	N	0.016 ± 0.006	0.17 ± 0.03
R36Q[d]	Y	0.34 ± 0.09	2.1 ± 0.2
R36Y	Y	0.26 ± 0.05	0.7 ± 0.2
I37R	Y	0.52 ± 0.12	1.03 ± 0.13
I37V	Y	1.4 ± 0.2	1.2 ± 0.3
A38L	Y	0.4 ± 0.2	0.62 ± 0.2
A38S	Y	0.39 ± 0.12	1.0 ± 0.2
Y39A[d]	N	0.04 ± 0.02	0.15 ± 0.02
Y39F	Y	0.38 ± 0.14	1.0 ± 0.5
Y39G[d]	N	0.033 ± 0.012	0.118 ± 0.013
Y39H	Y	1.1 ± 0.3	0.60 ± 0.18
Y39K	Y	0.38 ± 0.19	0.8 ± 0.3

[a] NunN fusions are as described in Table 3 and are listed by single-letter codes showing the wild-type residue and the mutant residue separated by the numerical position in Nun.
[b] As described in Table 3, the ability of NunN mutants to complement N-deficient λ was assayed by observing plaque formation with 300 μM IPTG induction. Y, plaques were observed; N, no plaques were observed.
[c] Experiments were conducted as described in Table 3. Fraction β-galactosidase activity is calculated as the β-galactosidase units of the mutant NunN-boxB divided by the β-galactosidase units from NunN-boxB.
[d] This clone is a synthetic construct made for confirmation of activity or to assay values of mutants not found by selection.

of λ N, and Nun R29K, R29Q, R32K, W33A, and R36K, which showed strong preferences for $boxB_{left}$ over $boxB_{right}$ in termination assays (Table 2), have much less relative activity for $boxB_{left}$ in an N antitermination context (Table 4).

Rigid-body docking identifies a possible NusG-Nun ARM interaction. Faber et al. (29) include NusG as a possible Nun ARM-interacting host factor. Stagno et al. (17) propose a model that places NusG-CTD contacting N, and implicitly Nun, ARM. Mooney et al. (25) note that three of the seven Nun-defective substitutions, including hydrophobic Phe144 (24) and Phe165 (25), cluster on one face of NusG-CTD. Together, these observations prompted us to attempt docking the NusG-CTD structure to that of the Nun ARM-boxB structure. Using ZDOCK (53, 54) and PyDock (55), we find a predicted interaction that places Nun Trp33 nestled in a hydrophobic pocket on the surface of NusG-CTD (Fig. 4; see the supplemental material for the PDB file). Interestingly, in this model, Nun Trp33 contacts NusG Phe144 and Phe165, whose substitutions specifically abrogate Nun termination (24, 25), and Nun Arg36 contacts the NusG backbone at Gly139. No contacts are seen between NusG and Nun Asp26. Substituting an alanine at Nun Trp33 abolished docking of NusG-CTD and Nun ARM-boxB.

DISCUSSION

Structural roles of Nun residues. The sequence conservation of Nun ARM with λ N ARM combined with the overall similarity of Nun-boxB and λ N-boxB NMR structural models suggested that their recognition strategies were indistinguishable. The activity of wtNun and casNun on λ and P22 reporters (Table 1) indicates boxB recognition similar to that of λ N (Table 3) and is consistent with previous reports of Nun (26) activities.

Our mutagenesis results (Table 2) agree with the structural analysis of Faber et al. (29), in which Nun residues Arg28, Arg32, and Tyr39 interact with boxB using contacts very similar to the equivalent residues of λ N. We find Arg32, Arg36, and Tyr39 important: the proportions of active members in these libraries are low, only wild-type or conservative mutants were recovered from functional screening, mutants were unable to exclude λ infection, and mutants displayed low termination activity. In contrast, Burmann et al. (28) tested mutants of eight Nun residues and found all to be functional, including Arg32A, Arg36A, and Y39A. We note substantial differences between our materials and

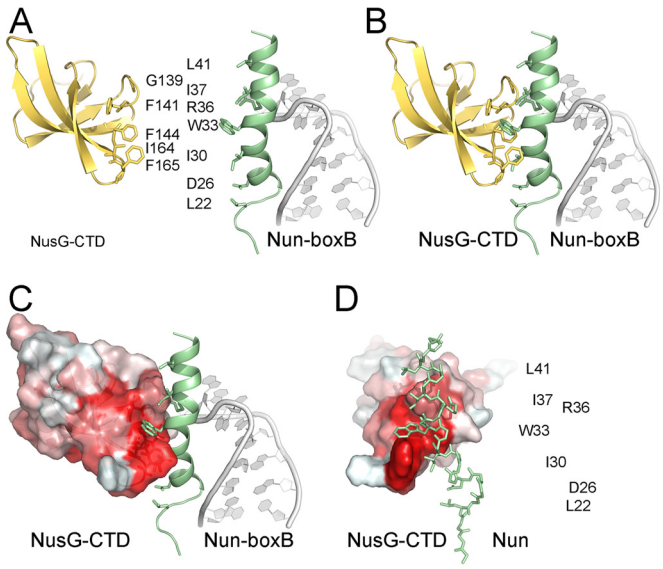

FIG 4 Rigid-body docking of NMR structural models of Nun ARM-boxB and NusG-CTD. (A) Left, NusG as a gold cartoon with residues shown as sticks next to labels. Right, Nun ARM rendered as a green cartoon and residues shown as sticks next to labels, and boxB rendered as a white cartoon. (B) Docked Nun ARM-boxB and NusG-CTD in the same rendering and orientation as in panel A. (C) As in panel A, with NusG colored according to hydrophobicity. (D) As in panel B, rotated about the vertical axis to look into the hydrophobic surface of NusG-CTD, without boxB shown, and Nun ARM rendered as a green backbone stick. Labels are displaced relative to side chain locations.

those of Burmann et al. (28): casNun is expressed from the tightly regulated *BAD* promoter (49) and has one residue amino-terminal and two residues carboxy-terminal to the Nun ARM that could amplify the effect of mutants, and our measurement of termination relies on reporter plasmids. Burmann et al. (28) expressed normal-length Nun from the less tightly regulated *tac* promoter (49) and monitored termination with a temperature-sensitive lysogenic reporter.

The importance of other conserved or similar residues, Ser24, Arg27, Lys35, and Arg36, appears more complex. Most clearly, the activity of Nun S24W argues that Nun Ser24 is not packed into a small hydrophobic pocket similarly to λ N Ala3. Our mutagenesis offers little support for Arg27 having importance: R27L and R27Y activities are reduced only moderately (Table 2). Nun Lys35 is mutable: K35L is active. This suggests that despite being conserved with λ N Lys14, the likely ionic interaction to a phosphate in the boxB backbone is not important to Nun. This contrast between which contacts may occur and which are important emphasizes the value of combining structural and functional analyses. Nun Arg36 appears to pack against the boxB loop near the extruded loop purine similarly to λ N Gln15. Interestingly, R36K was one of few mutants seen to display bias between $boxB_{left}$ and $boxB_{right}$, whose only difference is the identity of the extruded loop purine.

The mutagenesis offers strong support to the proposal of Faber et al. (29) that Nun ARM presents a recognition site for host factors. However, of the proposed contributing residues, we found only Trp33 to be essential, and we found that Asp26 and Arg36 are additional, essential residues. Other hydrophobic residues, Leu22, Ile30, and Ile37, were individually replaceable with a variety of residues, although radical substitutions were not specifically tested, and Leu41 was not mutagenized. These residues may be redundant or not absolutely required in our assay.

Nun ARM interactions with host factors. Nun termination and N antitermination interact with a similar or identical set of host factors, yet they must assemble structurally distinct elongation complexes in order to direct contrary outcomes. The striking similarity of requirements for Nun ARM residues Asp26 and Trp33 in both Nun termination and N antitermination contexts suggests that functional and stable Nun ARM-boxB recognition is dependent on host factors that are required for Nun termination and present in N antitermination. NusG, which may function largely as a scaffold protein for N antitermination, has point mutations that are defective for Nun termination, yet allow N antitermination (24, 25). Our rigid-body docking results, including the nestling of Nun Trp33 between NusG Phe144 and Phe165, which is the site of substitutions abrogating Nun function (24, 25), are consistent with contacts from Nun ARM Trp33 and Arg36 to NusG-CTD. Asp26 may contact other regions of NusG or other host factors. The proposed NusG-Nun ARM model would benefit from corroborating molecular dynamics studies, biochemical or biophysical studies of Nun termination complexes *in vitro*, or genetic screens for compensatory mutations in NusG.

Arginine-rich motifs are found in important regulatory complexes in which their primary, sometimes only, roles may be to recognize and attach a functional domain of a protein to an RNA (59, 60). Similar to the complex role that the stacking of λ N Trp18 on the boxB loop has in stabilizing the antitermination complex (9–11, 41–43, 58, 61), the Nun ARM host recognition site may offer more than simple binding (62–64). Thus, with subtle sequence changes, ARM-RNA recognition can evolve new functions.

ACKNOWLEDGMENTS

We gratefully acknowledge funding from the American University of Beirut University Research Board.

We thank Pascale Legault for λ N peptide-$boxB_{left}$ structure files and Yanjie Chao for assistance with docking. This work benefited from access to the Central Research Science Laboratory at the American University of Beirut.

REFERENCES

1. **Friedman DI, Court DL.** 1995. Transcription antitermination: the λ paradigm updated. Mol Microbiol **18:**191–200. http://dx.doi.org/10.1111/j.1365-2958.1995.mmi_18020191.x.
2. **Weisberg RA, Gottesman ME.** 1999. Processive antitermination. J Bacteriol **181:**359–367.
3. **Court DL, Oppenheim AB, Adhya SL.** 2007. A new look at bacteriophage λ genetic networks. J Bacteriol **189:**298–304. http://dx.doi.org/10.1128/JB.01215-06.
4. **Franklin NC.** 1985. "N" transcription antitermination proteins of bacteriophages λ, ϕ21 and P22. J Mol Biol **181:**85–91. http://dx.doi.org/10.1016/0022-2836(85)90326-2.
5. **Lazinski D, Grzadzielska E, Das A.** 1989. Sequence specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59:**207–218. http://dx.doi.org/10.1016/0092-8674(89)90882-9.
6. **Robledo R, Atkinson BL, Gottesman ME.** 1991. Escherichia coli mutations that block transcription termination by phage HK022 Nun protein. J Mol Biol **220:**613–619. http://dx.doi.org/10.1016/0022-2836(91)90104-E.
7. **DeVito J, Das A.** 1994. Control of transcription processivity in phage λ: Nus factors strengthen the termination-resistant state of RNA polymerase induced by N antiterminator. Proc Natl Acad Sci U S A **91:**8660–8664. http://dx.doi.org/10.1073/pnas.91.18.8660.
8. **Nudler E, Gottesman ME.** 2002. Transcription termination and antitermination in E coli. Genes Cells **7:**755–768. http://dx.doi.org/10.1046/j.1365-2443.2002.00563.x.
9. **Rees WA, Weitzel SE, Das A, von Hippel PH.** 1997. Regulation of the elongation-termination decision at intrinsic terminators by antitermination protein N of phage λ. J Mol Biol **273:**797–813. http://dx.doi.org/10.1006/jmbi.1997.1327.
10. **Conant CR, Van Gilst MR, Weitzel SE, Rees WA, von Hippel PH.** 2005. A quantitative description of the binding states and in vitro function of antitermination protein N of bacteriophage λ. J Mol Biol **348:**1039–1057. http://dx.doi.org/10.1016/j.jmb.2005.03.042.
11. **Conant CR, Goodarzi JP, Weitzel SE, von Hippel PH.** 2008. The antitermination activity of bacteriophage λ N protein is controlled by the kinetics of an RNA-looping-facilitated interaction with the transcription complex. J Mol Biol **384:**87–108. http://dx.doi.org/10.1016/j.jmb.2008.05.014.
12. **Muteeb G, Dey D, Mishra S, Sen R.** 2012. A multipronged strategy of an anti-terminator protein to overcome Rho-dependent transcription termination. Nucleic Acids Res **40:**11213–11228. http://dx.doi.org/10.1093/nar/gks872.
13. **Mishra S, Mohan S, Godavarthi S, Sen R.** 2013. The interaction surface of a bacterial transcription elongation factor required for complex formation with an antiterminator during transcription antitermination. J Biol Chem **288:**28089–28103. http://dx.doi.org/10.1074/jbc.M113.472209.
14. **Mishra S, Sen R.** 2015. N protein from lambdoid phages transforms NusA into an antiterminator by modulating NusA-RNA polymerase flap domain interactions. Nucleic Acids Res **43:**5744–5758. http://dx.doi.org/10.1093/nar/gkv479.
15. **Legault P, Li J, Mogridge J, Kay LE, Greenblatt J.** 1998. NMR Structure of the bacteriophage λ N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell **93:**289–299. http://dx.doi.org/10.1016/S0092-8674(00)81579-2.
16. **Schärpf M, Sticht H, Schweimer K, Boehm M, Hoffmann S, Rösch P.** 2000. Antitermination in bacteriophage λ. The structure of the N36 peptide-boxB RNA complex. Eur J Biochem **267:**2397–2408.
17. **Stagno JR, Altieri AS, Bubunenko M, Tarasov SG, Li J, Court DL, Byrd RA, Ji X.** 2011. Structural basis for RNA recognition by NusB and NusE in

the initiation of transcription antitermination. Nucleic Acids Res **39:** 7803–7815. http://dx.doi.org/10.1093/nar/gkr418.

18. **Robert J, Sloan SB, Weisberg RA, Gottesman ME, Robledo R, Harbrecht D.** 1987. The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. Cell **51:**483–492. http://dx.doi.org/10.1016/0092-8674(87)90644-1.
19. **Oberto J, Weisberg RA, Gottesman ME.** 1989. Structure and function of the nun gene and the immunity region of the lambdoid phage HK022. J Mol Biol **207:**675–693. http://dx.doi.org/10.1016/0022-2836(89)90237-4.
20. **Hung SC, Gottesman ME.** 1995. Phage HK022 Nun protein arrests transcription on phage λ DNA in vitro and competes with the phage λ N antitermination protein. J Mol Biol **247:**428–442. http://dx.doi.org/10.1006/jmbi.1994.0151.
21. **Vitiello CL, Kireeva ML, Lubkowska L, Kashlev M, Gottesman M.** 2014. Coliphage HK022 Nun protein inhibits RNA polymerase translocation. Proc Natl Acad Sci U S A **111:**E2368–E2375. http://dx.doi.org/10.1073/pnas.1319740111.
22. **Sullivan SL, Gottesman ME.** 1992. Requirement for E. coli NusG protein in factor-dependent transcription termination. Cell **68:**989–994. http://dx.doi.org/10.1016/0092-8674(92)90041-A.
23. **Watnick RS, Gottesman ME.** 1998. *Escherichia coli* NusA is required for efficient RNA binding by phage HK022 nun protein. Proc Natl Acad Sci U S A **95:**1546–1551. http://dx.doi.org/10.1073/pnas.95.4.1546.
24. **Burova E, Hung SC, Chen J, Court DL, Zhou JG, Mogilnitskiy G, Gottesman ME.** 1999. *Escherichia coli* nusG mutations that block transcription termination by coliphage HK022 Nun protein. Mol Microbiol **31:**1783–1793. http://dx.doi.org/10.1046/j.1365-2958.1999.01315.x.
25. **Mooney RA, Schweimer K, Rosch P, Gottesman M, Landick R.** 2009. Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. J Mol Biol **391:**341–358. http://dx.doi.org/10.1016/j.jmb.2009.05.078.
26. **Henthorn KS, Friedman DI.** 1996. Identification of functional regions of the Nun transcription termination protein of phage HK022 and the N antitermination protein of phage λ using hybrid nun-N genes. J Mol Biol **257:**9–20. http://dx.doi.org/10.1006/jmbi.1996.0142.
27. **Chattopadhyay S, Hung SC, Stuart AC, Palmer AG, III, Garcia-Mena J, Das A, Gottesman ME.** 1995. Interaction between the phage HK022 Nun protein and the nut RNA of phage λ. Proc Natl Acad Sci U S A **92:**12131–12135. http://dx.doi.org/10.1073/pnas.92.26.12131.
28. **Burmann BM, Uc-Mass A, Schweimer K, Gottesman ME, Rösch P.** 2008. The Y39A mutation of HK022 Nun disrupts a boxB interaction but preserves termination activity. Biochemistry **47:**7335–7341. http://dx.doi.org/10.1021/bi8004347.
29. **Faber C, Schärpf M, Becker T, Sticht H, Rösch P.** 2001. The structure of the coliphage HK022 Nun protein-λ-phage boxB RNA complex. J Biol Chem **276:**32064–32070. http://dx.doi.org/10.1074/jbc.M102975200.
30. **Dambly C, Couturier M.** 1971. A minor Q-independent pathway for the expression of the late genes in bacteriophage λ. Mol Gen Genet **113:**244–250. http://dx.doi.org/10.1007/BF00339545.
31. **Hilliker S, Botstein D.** 1976. Specificity of genetic elements controlling regulation of early functions in temperate bacteriophages. J Mol Biol **106:** 537–566. http://dx.doi.org/10.1016/0022-2836(76)90251-5.
32. **Austin RJ, Xia T, Ren J, Takahashi TT, Roberts RW.** 2003. Differential modes of recognition in N peptide-boxB complexes. Biochemistry **42:** 14957–14967. http://dx.doi.org/10.1021/bi0351312.
33. **Cilley CD, Williamson JR.** 1997. Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE). RNA **3:**57–67.
34. **Tan R, Frankel AD.** 1995. Structural variety of arginine-rich RNA-binding peptides. Proc Natl Acad Sci U S A **92:**5282–5286. http://dx.doi.org/10.1073/pnas.92.12.5282.
35. **Cai Z, Gorin A, Frederick R, Ye X, Hu W, Majumdar A, Kettani A, Patel DJ.** 1998. Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. Nat Struct Biol **5:**203–212. http://dx.doi.org/10.1038/nsb0398-203.
36. **Cilley CD, Williamson JR.** 2003. Structural mimicry in the phage ϕ21 N peptide-boxB RNA complex. RNA **9:**663–676. http://dx.doi.org/10.1261/rna.2189203.
37. **Franklin NC.** 1993. Clustered arginine residues of bacteriophage λ N protein are essential to antitermination of transcription, but their locale cannot compensate for boxB loop defects. J Mol Biol **231:**343–360. http://dx.doi.org/10.1006/jmbi.1993.1287.
38. **Su L, Radek JT, Hallenga K, Hermanto K, Chan G, Labeots LA, Weiss MA.** 1997. RNA recognition by a bent a-helix regulates transcriptional antitermination in phage λ. Biochemistry **36:**12722–12732. http://dx.doi.org/10.1021/bi971408k.
39. **Cocozaki AI, Ghattas IR, Smith CA.** 2008. The RNA-binding domain of bacteriophage P22 N protein is highly mutable, and a single mutation relaxes specificity toward λ. J Bacteriol **190:**7699–7708. http://dx.doi.org/10.1128/JB.00997-08.
40. **Chattopadhyay S, Garcia-Mena J, DeVito J, Wolska K, Das A.** 1995. Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage λ. Proc Natl Acad Sci U S A **92:**4061–4065. http://dx.doi.org/10.1073/pnas.92.9.4061.
41. **Xia T, Becker HC, Wan C, Frankel A, Roberts RW, Zewail AH.** 2003. The RNA–protein complex: direct probing of the interfacial recognition dynamics and its correlation with biological functions. Proc Natl Acad Sci U S A **100:**8119–8123. http://dx.doi.org/10.1073/pnas.1433099100.
42. **Xia T, Frankel A, Takahashi TT, Ren J, Roberts RW.** 2003. Context and conformation dictate the function of a transcriptional switch. Nat Struct Biol **10:**812–819. http://dx.doi.org/10.1038/nsb983.
43. **Xia T, Wan C, Roberts RW, Zewail AH.** 2005. RNA-protein recognition: single-residue ultrafast dynamical control of structural specificity and function. Proc Natl Acad Sci U S A **102:**13013–13018. http://dx.doi.org/10.1073/pnas.0506181102.
44. **Sambrook J, Fritsch EF, Maniatis T.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
45. **Doelling JH, Franklin NC.** 1989. Effects of all single base substitutions in the loop of boxB on antitermination of transcription by bacteriophage λ's N protein. Nucleic Acids Res **17:**5565–5577. http://dx.doi.org/10.1093/nar/17.14.5565.
46. **Franklin NC.** 2004. Morphing molecular specificities between Arm-peptide and NUT-RNA in the antitermination complexes of bacteriophages λ and P22. Mol Microbiol **52:**815–822. http://dx.doi.org/10.1111/j.1365-2958.2004.04018.x.
47. **Possik EJ, Bou Sleiman MS, Ghattas IR, Smith CA.** 2013. Randomized codon mutagenesis reveals that the HIV Rev arginine-rich motif is robust to substitutions and that double substitution of two critical residues alters specificity. J Mol Recognit **26:**286–296. http://dx.doi.org/10.1002/jmr.2272.
48. **Cocozaki AI, Ghattas IR, Smith CA.** 2008. Bacteriophage P22 antitermination boxB sequence requirements are complex and overlap with those of λ. J Bacteriol **190:**4263–4271. http://dx.doi.org/10.1128/JB.00059-08.
49. **Guzman LM, Belin D, Carson MJ, Beckwith J.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J Bacteriol **177:**4121–4130.
50. **Uc-Mass A, Khodursky A, Brown L, Gottesman ME.** 2008. Overexpression of phage HK022 Nun protein is toxic for *Escherichia coli*. J Mol Biol **380:**812–819. http://dx.doi.org/10.1016/j.jmb.2008.05.030.
51. **Brickman E, Beckwith J.** 1975. Analysis of the regulation of Escherichia coli alkaline phosphatase synthesis using deletions and ϕ80 transducing phages. J Mol Biol **96:**307–316. http://dx.doi.org/10.1016/0022-2836(75)90350-2.
52. **Miller JH.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria. Cold Spring Harbor Laboratory Press, Plainview, NY.
53. **Chen R, Li L, Weng Z.** 2003. ZDOCK: an initial-stage protein docking algorithm. Proteins **52:**80–87. http://dx.doi.org/10.1002/prot.10389.
54. **Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z.** 2014. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics **30:**1771–1773. http://dx.doi.org/10.1093/bioinformatics/btu097.
55. **Jiménez-Garcia B, Pons C, Fernández-Recio J.** 2013. pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. Bioinformatics **29:**1698–1699. http://dx.doi.org/10.1093/bioinformatics/btt262.
56. **Eisenberg D, Schwarz E, Komarony M, Wall R.** 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J Mol Biol **179:**125–142. http://dx.doi.org/10.1016/0022-2836(84)90309-7.
57. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr **66:**486–501. http://dx.doi.org/10.1107/S0907444910007493.

58. **Van Gilst MR, von Hippel PH.** 1997. Assembly of the N-dependent antitermination complex of phage λ: NusA and RNA bind independently to different unfolded domains of the N protein. J Mol Biol **274:**160–173. http://dx.doi.org/10.1006/jmbi.1997.1389.
59. **Weiss MA, Narayana N.** 1998. RNA recognition by arginine-rich peptide motifs. Biopolymers **48:**167–180. http://dx.doi.org/10.1002/(SICI)1097-0282(1998)48:2<167::AID-BIP6>3.0.CO;2-8.
60. **Draper DE.** 1999. Themes in RNA-protein recognition. J Mol Biol **293:**255–270. http://dx.doi.org/10.1006/jmbi.1999.2991.
61. **Mogridge J, Legault P, Li J, Van Oene MD, Kay LE, Greenblatt J.** 1998. Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage λ N protein. Mol Cell **1:**265–275. http://dx.doi.org/10.1016/S1097-2765(00)80027-1.
62. **Frankel AD, Smith CA.** 1998. Induced folding in RNA-protein recognition: more than a simple molecular handshake. Cell **92:**149–151. http://dx.doi.org/10.1016/S0092-8674(00)80908-3.
63. **Williamson JR.** 2000. Induced fit in RNA-protein recognition. Nat Struct Biol **7:**834–837. http://dx.doi.org/10.1038/79575.
64. **Mackereth CD, Sattler M.** 2012. Dynamics in multi-domain protein recognition of RNA. Curr Opin Struct Biol **22:**287–296. http://dx.doi.org/10.1016/j.sbi.2012.03.013.