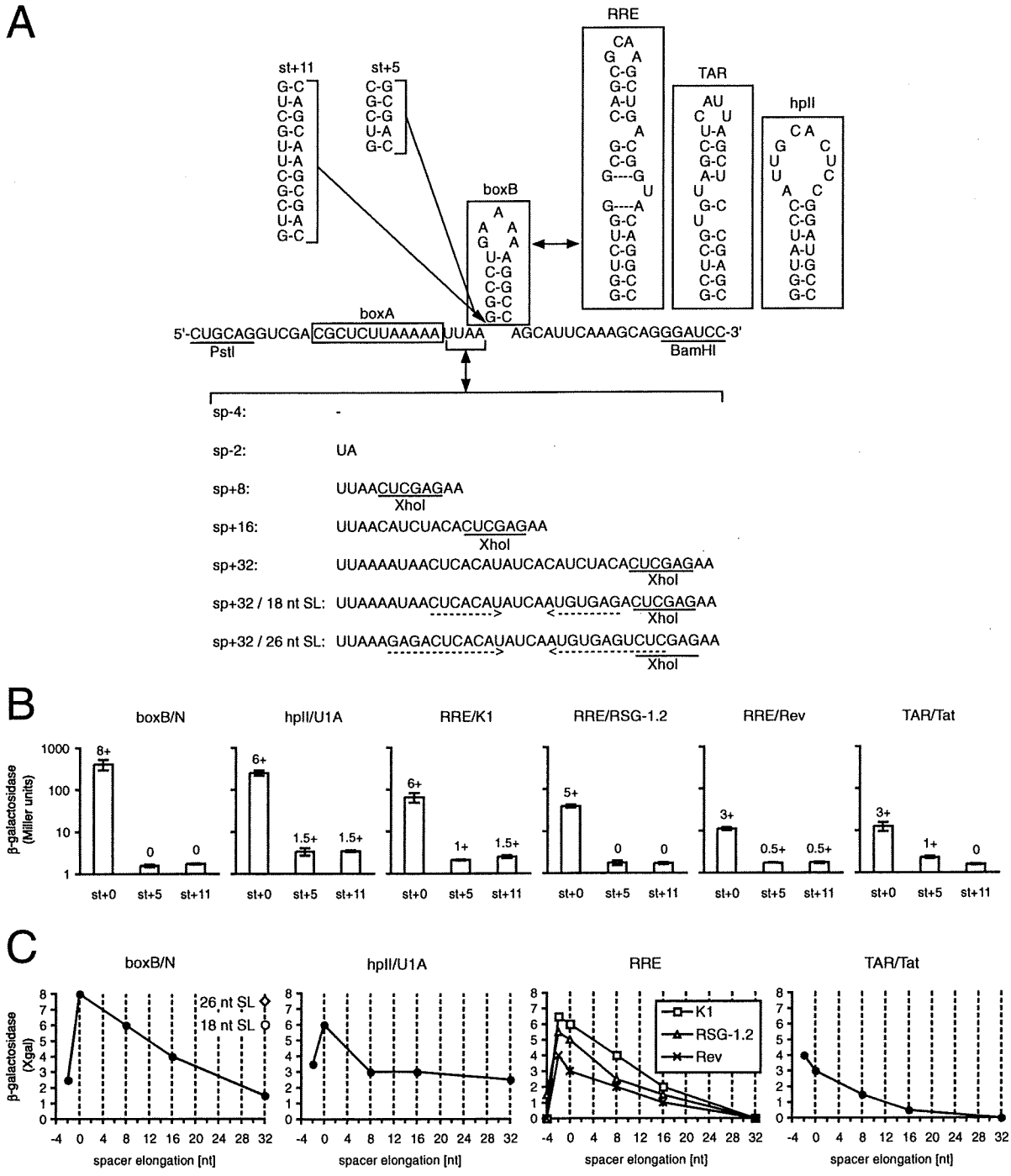

Fig. 2. The effect of the lengthening of the *nut* site linker and the stem of the boxB or heterologous RNA site on anti-termination activity.
A. The secondary structures of the *nut* site are shown with RNA sequence insertions indicated using one-sided arrows, and RNA sequence replacements indicated using two-sided arrows. The broken arrows under the sequence of sp+32/18nt SL and sp+32/26 nt SL indicate formation of an RNA stem.
B. The effect of lengthening the RNA stem of the boxB and heterologous RNA sites on the anti-termination activities of boxB/N and heterologous RNA–peptide interactions. The numbers of plusses above bars indicate the β-galactosidase activities assayed by the colony colour on Xgal plates.
C. The effect of lengthening the spacer length between the boxA element and the boxB or heterologous RNA sites on the anti-termination activities of boxB/N and heterologous RNA–peptide interactions.

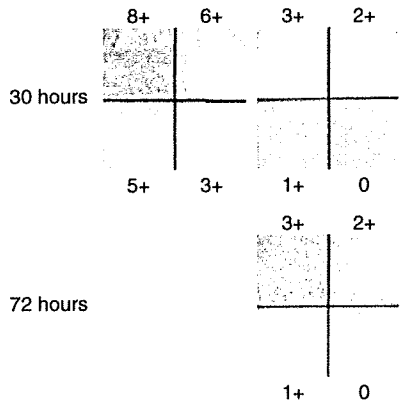

Fig. 3. Representative colony colours of standardized controls used in the β-galactosidase colony colour assays after incubation for 30 h (24 h at 37°C followed by 6 h at 24°C) or 72 h (an additional 42 h at 24°C) on Xgal/tryptone plates. The anti-termination activity was represented as the number of plusses; boxB/N: 8+ RRE/K1: 6+ RRE/RSG-1.2: 5+ RRE/Rev: 3+ RRE/Rev(A20): 2+ boxB/N(A20): 1+ boxB/Tat: 0.

a deletion of 2 or 4 nts (designated sp-2 or sp-4), and an insertion of 8, 16 or 32 nts (designated sp+8, sp+16 or sp+32), respectively, with respect to the wild-type *nut* spacer (Fig. 2A), and analysed their anti-termination activities on plates containing Xgal (Fig. 2C) and in solution using ONPG (Table S3). As a result, shortening of the linker by two bases (sp-2) in the boxB/N context led to a dramatic reduction in anti-termination activity to 2.5+ compared with 8+ for the wild-type context (sp+0), consistent with a previously report (Doelling and Franklin, 1989). A similar reduction in activity was observed in the hpII/U1A context upon shortening the linker by 2 nts. In contrast, shortening the linker by 2 nts (sp-2) in the RRE/Rev, RRE/RSG-1.2, RRE/K1 or TAR/Tat context resulted in anti-termination activities higher than those in the wild-type linker context (sp+0). However, further shortening of the linker by 2 nts (sp-4) in the case of the RRE led to a dramatic loss of activity. These results showed that the optimal length of the linker differs for the different RNA–peptide complex, and suggest that the wild-type linker has been evolutionarily optimized for the boxB/N interaction.

On the other hand, the elongation of the RNA linker led to a gradual decrease in anti-termination activity. In the case of the boxB/N and hpII/U1A interactions, a considerable amount of residual activity was observed even after insertion of a 32 nt linker, which is in contrast to the case of the RRE/K1, RRE/RSG-1.2, RRE/Rev and TAR/Tat interactions. This difference may reflect differences in the role that the boxA element plays in the formation of the wild-type and modified anti-termination complexes, and in the case of the hpII/U1A interaction, due to the high expression level of the U1A protein and a high background of anti-termination (Fig. S3A, lower).

In order to show that the reduction of anti-termination activity upon lengthening the spacer was due to an entropic effect, and not due to the disruption of anti-termination complex formation by steric clash, two additional constructs based on λ boxB (sp+32) were prepared. Both constructs contain 32 extra nts in the spacer, but one was capable of forming a 18 nt stem-loop [designated (sp+32/18 nt SL)], while the other can form a 26 nt stem-loop ([sp+32/26 nt SL]), resulting in an effective linker length corresponding to that of a 16 base linker (sp+16) and 8 base linker (sp+8) respectively (Fig. 2A). As might be expected, as the spacer region in H-19B *nut* site has been shown to contain similar stem-loop structures (Neely and Friedman, 1998; 2000), the activities of the two spacer constructs capable of forming stem-loops were dependent on the length of single-stranded region within the elongated spacer, and reported higher activities corresponding to the shortening of the effective length of the spacer (Fig. 2C, boxB/N, open circle and diamond).

Combination of boxB stem elongation and the spacer length of nut

As the dramatic loss of activity upon lengthening the stem region of the RNA site may be due in part to the disruption of interactions with factors linked to the boxA element (Fig. 1), it seemed possible that lengthening the spacer between boxA and the RNA site may suppress this loss of activity. In order to test this hypothesis, reporter constructs with combinations of both the elongated stems and spacers described above were prepared, and anti-termination activities were assayed on plates containing Xgal and in solution using ONPG (Fig. 4 and Table S3). As a result, while lengthening the spacer did not completely suppress the loss of function by stem lengthening, a significant recovery of activity was observed in some cases. In the case of the boxB/N interaction, simultaneous lengthening of the stem and spacer by 5 bp and 16 nts [boxB (sp+16; st+5)], respectively, resulted in a particularly high level of recovery to 4+ for the colony colour assay compared with an activity of 0 when only the stem of boxB was lengthened [boxB (sp+0; st+5)]. In contrast, when the stem of boxB was lengthened by 11 bp, the highest recovery was observed when the RNA spacer was lengthened by 8 nts [boxB (sp+8; st+11)], possibly reflecting the different spatial requirements imposed by lengthening the stem by 5 and 11 bp. Similarly, a moderate suppression of the loss of activity was observed in the case of the hpII/U1A, the RRE/Rev, and the TAR(st+11)/Tat interaction. In the case of the RRE/K1 and TAR(st+5)/Tat interactions, while suppression of the loss of activity was not observed by simultaneous lengthening of both the stem and linker, lengthening the stem appeared to result in a reduction of the decrease in activity upon lengthening the RNA linker.

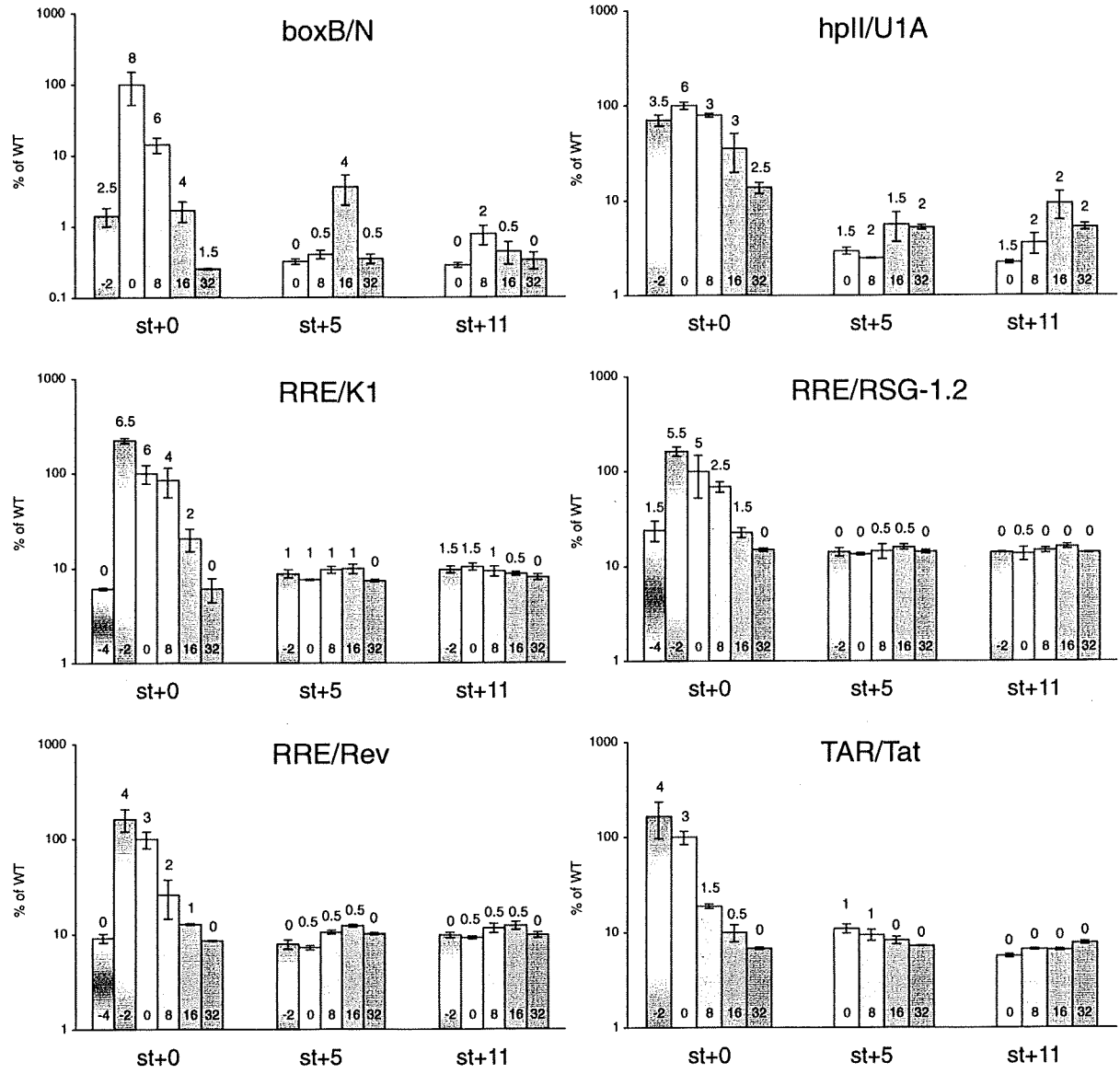

Fig. 4. The effect of the combination of lengthening the RNA stem of the boxB and heterologous RNA sites and the spacer length between the boxA element and the boxB or heterologous RNA sites on the anti-termination activities of boxB/N and heterologous RNA–peptide interactions. The numbers at the bottom of the bars indicate the nucleotide length inserted to the spacers. The percentage of wild-type (WT) is normalized to the average β-galactosidase Miller units of the wild-type reporter (sp+0; st+0) assayed the same day. The numbers above bars indicate the β-galactosidase activities assayed by the colony colour on Xgal plates.

Taken together, the above results suggest that the length of the RNA stem and linker are somehow linked in the formation of a functional anti-termination complex.

Insertion of peptide linkers following the NH_2-terminal RNA-binding domain of N

As the lengthening of the RNA spacer within the *nut* site did not completely suppress the loss of function by the lengthening of the stem of boxB or the other RNA sites as shown above, we next examined the effect of lengthening the space between the RNA-binding peptide (N_{1-20} or heterologous peptides) and the activation domain (N_{21-107}) of the N protein which is involved in NusA and RNA polymerase binding (Fig. 5 and Table S4). A tandem repeat of one, two or four linker peptides consisting of five alanine residues (AAAAA, or A5), or a combination of three glycines and two serines (GSGSG, or GS5), designated A5,

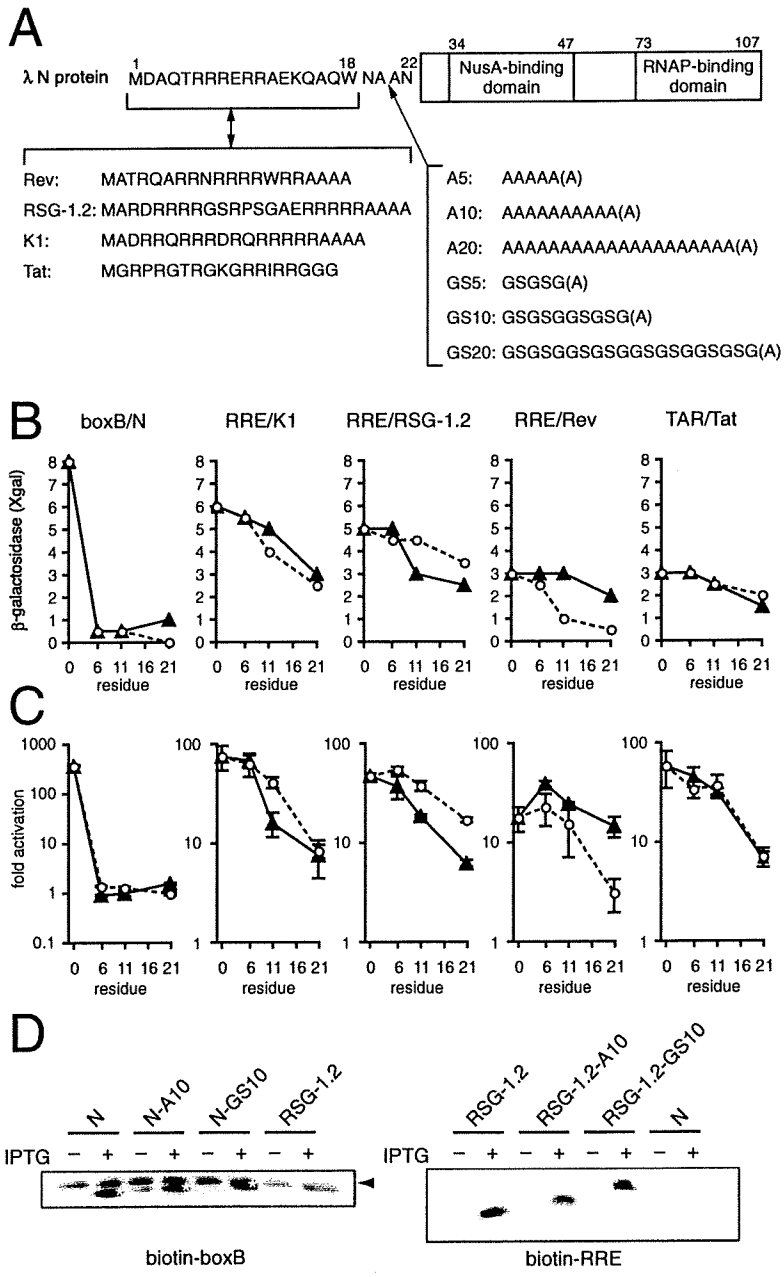

Fig. 5. The effect of the lengthening of the linker connecting the N peptide (N_{1-20}) and the remainder of N (N_{21-107}) on anti-termination activity.
A. Schematic representation of N and N-fusion protein and their NH_2-terminal sequences expressed by the N-expressor plasmids. The Lys19 of the N protein was changed to Asp19 to generate a BsmI site with a negligible effect on N function (Franklin, 1993). Two sided arrows indicate the substitution of the N peptide sequence with heterologous RNA-binding peptides, and one-sided arrows indicate the insertion of the peptide linker.
B and C. The effect of peptide linker insertion on the anti-termination activity of boxB/N and heterologous RNA–peptide interactions. Closed triangles indicate AAAAA linkers and open circles indicate GSGSG linkers. β-Galactosidase activities were measured on Xgal plates (B) and in solutions using ONPG (C). Fold activation (ONPG) is the ratio of β-galactosidase in Miller units of the cognate reporter in the presence of N or N-fusion divided by the β-galactosidase in Miller units of non-cognate reporter (TAR for N, K1, RSG-1.2 or Rev series and boxB for Tat series).
D. North Western blotting of N and RSG-1.2–N fusion protein with peptide linker insertions. N and RSG-1.2–N fusion proteins with peptide linker insertions were expressed in *E. coli* N567 with (+) or without (–) IPTG induction and subjected to 15% SDS-PAGE followed by North Western blotting with biotinylated RNA probes. Each lane contains the extract from 0.08 OD_{600} units of original culture (1 ml). An arrowhead and white asterisks in the left panel indicate the position of non-specific bands and the bands of N protein with or without peptide linkers respectively.

A10, A20 and GS5, GS10 GS20, respectively, in addition to a C-terminal alanine residue in the linkers for cloning purposes, was inserted into the BsmI site of the N expressor plasmid, resulting in the insertion of 6, 11 and 21 residues respectively (Fig. 5A). These polyalanine or glycine/serine linkers were expected to possess somewhat different structural features in that oligoalanines have a high α-helix propensity, while glycine/serine repeats are flexible, thereby possibly influencing the stability of the RNA–peptide complex as well as that of the anti-termination complex. We have previously shown that four alanines linked to the C-terminus of the α-helical Rev peptide led to a substantial increase in anti-termination activity towards the RRE, while a triglycine linker reduced activity (Harada and Frankel, 1998). On the contrary, when the triglycine linker was attached to the BIV Tat, which forms a β-hairpin, an increase in activity was observed (Harada and Frankel, 1998).

The results of the peptide linker insertions assayed in the wild-type *nut* context (sp+0; st+0) are shown in Fig. 5B and C and Table S4. Insertion of even a relatively short A5 or GS5 linker in the boxB/N context led to a dramatic decrease in activity (8+ to 0.5+), similar to when the boxB stem was lengthened. On the other hand, in the case of the heterologous RNA–peptide interactions, lengthening the peptide linker led to only a gradual decrease in anti-termination activity. Insertion of the peptide linkers between residues 20 and 21 of the N protein result in changes in the C-terminal region of what has been defined as the N peptide (residues 1–22). Although residues 20–22 of the N protein have been shown to be dispensable for specific binding to boxB *in vitro* (Cilley and Williamson, 1997; Su *et al.*, 1997a), in order to confirm that insertion of peptide linkers does not affect boxB binding or expression levels of N, Western and North Western blotting analyses of wild-type N as well as Rev, RSG-1.2 and K1 peptide N-fusion proteins without and with peptide linkers were carried out. Western blotting analysis showed that while protein levels were somewhat different depending on the peptide fused to $N_{21\text{-}107}$, and on the type of linker, these differences did not correlate with the differences observed in anti-termination activity (Fig. S3). In addition, North Western analysis of the N and the RSG-1.2–N fusion proteins, indicated that the insertion of peptides linkers did not affect boxB and RRE binding respectively (Fig. 5D).

In the case of the heterologous RNA–peptide interactions, no significant difference in the linker type was observed in the RRE/K1 and TAR/Tat context, while the flexible glycine linker was favoured in the RRE/RSG-1.2 interaction, and the alanine linkers with high α-helical propensity seemed to be preferred in the RRE/Rev context. These differences in the types of peptide linkers preferred may reflect the different constraints that the RNA–peptide complexes impose on the anti-termination complex, because it has been shown by nuclear magnetic resonance spectroscopy that the Rev and RSG-1.2 peptides bind to the RRE in completely different orientations (Battiste *et al.*, 1996; Gosser *et al.*, 2001; Zhang *et al.*, 2001), so that the remainder of the N protein (residues 21–107) would be positioned in a different direction.

As in the case of boxB RNA stem lengthening, the dramatic decrease in activity for the boxB/N interaction upon insertion of peptide linkers into N may reflect strict spatial requirements for the orientation of the boxB/N/NusA ternary complex (Mogridge *et al.*, 1995; Legault *et al.*, 1998). In the case of the heterologous RNA–peptide pairs, a significant relaxation in the stringency towards the length of the peptide linker was observed, presumably due to the loss of the interaction with NusA. As the loss of activity upon lengthening the boxB RNA stem could be partially suppressed by lengthening the *nut* linker RNA, we investigated the possibility that loss of activity upon lengthening the RNA stem could be suppressed by peptide linker insertion. However, unlike in the case of simultaneous lengthening of the RNA stem and *nut* linker (Fig. 4), no significant compensation of the decreased activity by stem lengthening was observed upon lengthening the peptide linker (Fig. S4 and Table S4). However, as in the case of the simultaneous RNA stem/*nut* spacer elongation, the decrease in activity observed upon lengthening the peptide linker in the st+0 context appeared to be more gradual when the RNA stem was lengthened by 5 or 11 bp (st+5 and st+11). For example, in the case of RRE/K1 context, the activity of (st+0)/(GS0) was down from 6+ to 3+ for (st+0)/(GS20), while the activity of (st+5)/(GS0) was as same as (st+5)/(GS20). This suggested that the RNA stem and peptide linker length are somewhat correlated even though the decrease in activity of stem elongation could not be recovered by peptide linker elongation.

Discussion

The spatial requirement for the wild-type nut *site RNA and N peptide within the anti-termination complex is optimized and strict*

It has been shown that high concentrations of N can induce anti-termination *in vitro*, even in the absence of boxB, demonstrating the central role of N in this process (Rees *et al.*, 1996). However, as the interaction of N with the transcription complex has been estimated to be relatively weak, the formation of a boxB/N/NusA complex with RNAP has been shown to be necessary for anti-termination *in vivo*, thereby allowing transcription through terminators located just downstream of the *nut* site (Whalen and Das, 1990; Mason *et al.*, 1992). Processive anti-termination through distant terminators requires NusB, NusG and S10 along with boxA RNA, suggesting that these proteins can be considered as processivity factors. The formation of the core boxB/N/NusA complex has been shown to occur first by the binding of the amino terminal region of N as a bent α-helix with the 5′-strand of the boxB stem and the first three residues of the loop (Legault *et al.*, 1998). This is accompanied by the extrusion of the fourth nucleotide in the loop so that the remaining nucleotides form a GNRA fold. NusA interacts with this newly created boxB/N binding surface, while also forming a protein–protein contact with amino acids 34–47 of N (Mogridge *et al.*, 1998b), and by doing so further strengthening the interaction of N with RNAP (Fig. 1). Although NusA has been shown to bind fairly tightly to the free N protein with Kds as high as 70 nM (Van Gilst *et al.*, 1997), mutations in the boxB loop such as the deletion of the extruding fourth nucleotide of boxB lead to a loss of NusA-

binding *in vitro*, while still maintaining the boxB/N interaction (Chattopadhyay *et al.*, 1995; Legault *et al.*, 1998).

One possible explanation for the dramatic decrease in anti-termination activity upon lengthening the boxB stem and the N peptide linker (Figs 2B and 5B and C) may be the loss of cooperativity in the formation of the boxB/N/NusA complex, indicating that the spatial requirement of the individual components is important for stable ternary complex formation. However, considering that the stability of the boxB/N and RRE/Rev peptide have been shown to be fairly similar by *in vitro* gel shift experiments under similar conditions (20 nM and 40 nM respectively) (Tan and Frankel, 1995; Harada *et al.*, 1996), loss of the NusA interaction, for example by peptide linker lengthening, would be expected to result in residual activity similar to that of the RRE/Rev interaction (3+). Therefore, a more likely reason for the dramatic loss of activity in the wild-type context may be that the orientation of the boxB/N/NusA complex within the anti-termination complex is restricted, and that slight changes in the orientation of this ternary complex may cause strain to other interactions such as those involving RNAP, boxA, and other host factors, and destabilize the anti-termination complex. The partial suppression of the complete loss of activity upon lengthening the boxB stem by 5 bp (Figs 2B and 4, boxB/N, st+5) by the insertion of a 16 base linker to 4+ (Fig. 4, boxB/N, st+5/sp+16) may be regarded as a partial restoration of the cooperativity of the boxB/N/NusA ternary complex within the anti-termination complex. On the other hand, the introduction of mutations in the boxB loop that disrupt NusA-binding, while leading to an overall decrease in anti-termination activity, may be expected to result in a relaxation of the strict spatial requirement for the boxB/N complex, as in the case of the heterologous RNA/peptide interactions. In order to test this hypothesis, mutant boxB reporter plasmids where the fourth base (A) in the loop was substituted to an G, C and U, or deleted were constructed, and anti-termination activities in the presence of wild-type N and mutant N's with peptide linkers were analysed (Fig. S5 and Table S4). However, the anticipated relaxation of the stringent spatial requirement was not observed for the single nucleotide mutations in the loop, presumably because these mutations were not sufficient to abolish NusA binding *in vivo*. The deletion of the fourth nucleotide in the loop presumably led to disruption of both NusA- and N-binding, and diminished anti-termination activity.

The lengthening of the RNA spacer connecting boxB and the boxA element, which is mainly involved in positioning the processivity factors NusB, NusG and S10, resulted in only a gradual decrease in activity (Fig. 2C). As we also showed that relatively large stem-loops could be inserted into the *nut* RNA linker (Fig. 2C), the gradual decrease in activity for the *nut* RNA linker is most likely due to the increased entropic costs involved with an increase in flexibility as observed in the case of artificially engineered DNA-binding or RNA-binding proteins (Ribas de Pouplana *et al.*, 1996; Kim and Pabo, 1998; Campisi *et al.*, 2001; Moore *et al.*, 2001). In addition, for the boxB/N complex, optimal anti-termination activities were observed in the wild-type context, indicating that the spatial context of this interaction has been evolutionarily optimized.

Relaxed spatial requirements for the heterologous RNA–peptide modified anti-termination complex and implications for the engineering of functional ribonucleoprotein complexes

Replacement of the boxB/N complex with heterologous RNA–peptide interactions was shown to lead to a considerable relaxation of the strict spatial requirements for the RNA stem and N peptide linker observed in the formation of the wild-type anti-termination complex (Figs 2B and 5B and C), presumably due to the absence of the interaction of boxB/N with NusA in the wild-type complex. However, as a consequence of the replacement of the boxB/N interaction by heterologous RNA–peptide interactions, a considerable decrease in anti-termination activity was observed. For example, replacement of the boxB/N interaction with RRE/Rev resulted in a decrease in colony colour of 8+ to 3+, and a 40-fold reduction in β-galactosidase units. As the Kd of the boxB/N interaction (20 nM) has been shown to be similar to that of the RRE/Rev interaction (40 nM) (Tan and Frankel, 1995; Harada *et al.*, 1996), the difference in anti-termination activity of 8+ and 3+, respectively, may account for the contribution of the interaction of boxB/N and NusA on the stability of the anti-termination complex. On the other hand, replacement of the RRE/Rev interaction with those of RRE/RSG-1.2 and RRE/DLA with Kds of 6 nM and 0.5 nM, resulted in increases in anti-termination activities to 5+ and 6+ respectively (Harada *et al.*, 1997; Sugaya *et al.*, 2008a). This shows that the decrease in anti-termination activity due to the loss of the NusA interaction can be partially recovered to about 15% that of the boxB/N interaction in β-galactosidase units by increasing the affinity of the heterologous RNA–peptide interaction, even though the RNA–protein interaction network within the anti-termination complex has been considerably modified.

While it is not clear whether full anti-termination activity as seen in the wild-type context can be achieved for the heterologous RNA–peptide interactions (Xia *et al.*, 2003), a number of strategies are conceivable. First, optimization of the interaction of the individual components may further increase the stability of the anti-termination complex. In fact, we have found that optimization of the NusA-binding

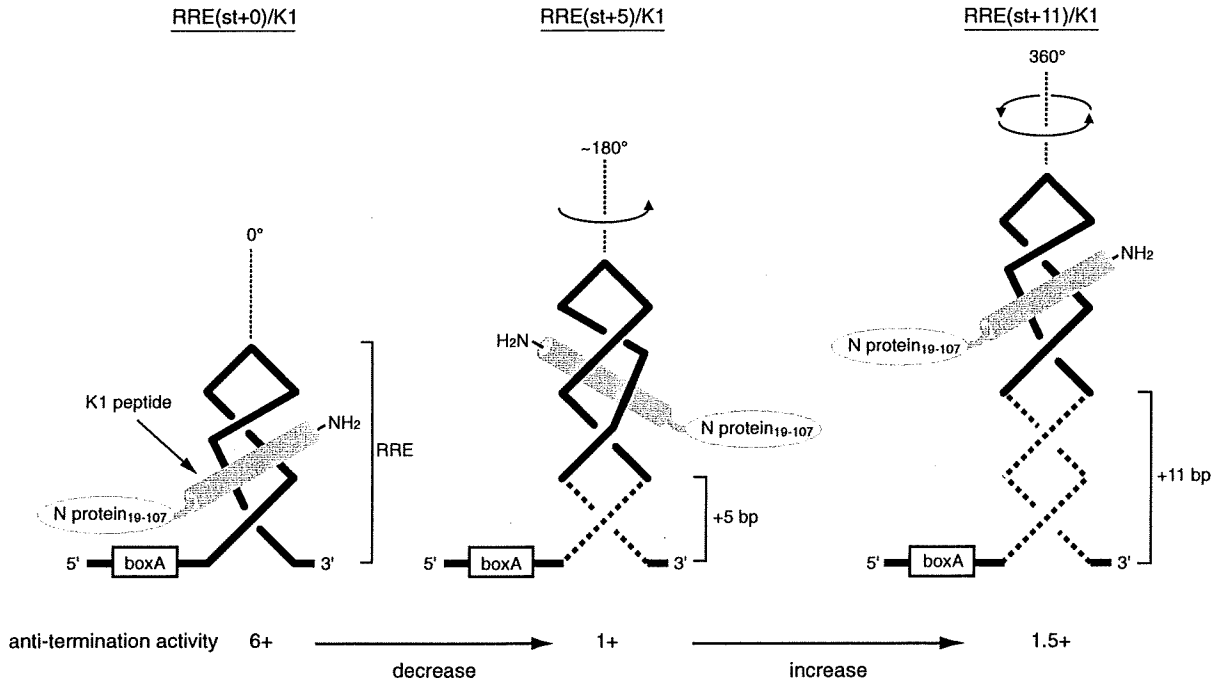

Fig. 6. A schematic illustration of the topological change of the boxB RNA stem-loop and the N protein upon elongation of the bottom stem in RRE/K1.

domain of N (residues 34–47) by randomization and selection leads to N variants with increased anti-termination activity (H. Suzuki and K. Harada, unpublished), demonstrating the potential of such an approach. A second strategy for the improvement of anti-termination activity is the simultaneous manipulation of the orientation of multiple components of the anti-termination complex as attempted in this study. While dramatic increases in activity were not observed in this study, a further understanding of the structural organization of the anti-termination complex may lead to the identification of alternative parameters for optimizing the spatial orientation of individual domains.

This study has also highlighted the importance of the length of RNA stem regions in the structural manipulation of the anti-termination complex, and possibly also for the engineering of any functional ribonucleoprotein complex. This is because even small changes in the length of RNA stem regions may result in dramatic changes in the orientation of associated factors. As illustrated in Fig. 6, insertion of a 5 bp stem would result in a completely different orientation for the remainder of the N protein (19–107) that contains the NusA and RNA polymerase binding domain, while insertion of an 11 bp stem would result in an orientation similar to the RNA site without an insertion (st+0). This may explain why a 5 bp insertion leads to a dramatic decrease in activity, while further decrease in activity is suppressed, or in the case of RRE/K1 a slight recovery of activity is observed, compared with st+5 upon further lengthening to st+11 (Fig. 2B and Table S1). Similarly, this may also explain why in the boxB/N interaction, the loss of activity in the st+5 stem was optimally suppressed by the sp+16 linker, while the shorter sp+8 linker was optimum for the st+11 stem (Fig. 4).

Taken together, the above results lead us to believe that a further understanding of the spatial and structural requirements for anti-termination complex formation may enable the manipulation of the anti-termination complex stability in a predictable manner, and as a result, allow the accommodation of a wider range of RNA–peptide interactions. Such studies are also expected to provide useful insights into general principles for the engineering of functional ribonucleoprotein complexes.

Experimental procedures

Plasmids

Oligonucleotides used to construct the pAC and pBR constructs described in this paper were synthesized at Hokkaido System Science or Espec Oligo Service. Reporter pAC plasmids containing a wild-type spacer (sp+0) or shortened spacer (sp-4 or sp-2) at the *nut* site, and a part of reporters containing spacer with XhoI site were constructed by cloning

synthetic oligonucleotides into the PstI and BamHI sites of the previously described pAC plasmid (referred to as pAC-TAT13 in Franklin, 1993). Based on the generated pAC plasmids containing the XhoI site, synthetic oligonucleotide cassettes were cloned into the PstI and XhoI site to construct spacer elongated constructs, and XhoI and BamHI for the replacement of the boxB site with appropriate sequences. These newly generated plasmids were further digested with the combination of PstI and XhoI, or XhoI and BamHI, and replaced with synthetic oligonucleotide cassettes to yield appropriate constructs. N or N-fusion protein expressors with peptide linkers were generated by insertion of synthetic oligonucleotides cassettes into the BsmI sites of the previously described pBR-based plasmids encoding N (Franklin, 1993), HIV Rev-N and BIV Tat-N (Harada *et al.*, 1996), RSG-1.2–N (Harada *et al.*, 1997) or U1A-N and K1-N (Peled-Zehavi *et al.*, 2003). The sequence of the sense strands of inserted linkers is as follows, A5: 5′-GCCGCTGCGGCCGCGGCA-3′; A10: 5′-GCCGCTGCGGCCGCGGCAGCCGCCGCGGCTG CA-3′; A20: 5′-GCCGCTGCGGCCGCGGCAGCCGCCGC GGCTGCAGCTGCGGCGGCGGCCGCTGCGGCTGCCGC AGCA-3′; GS5: 5′-GGCAGCGGTAGCGGCGCA-3′; GS10: 5′-GGCAGCGGTAGCGGCGGAAGTGGCAGCGGTGCA-3′; GS20: 5′-GGCAGCGGTAGCGGCGGAAGTGGCAGCGGT GGAAGCGGTAGTGGCGGCAGCGGCAGCGGTGCA-3′.
The pBR N^- plasmid which does not express N protein was generated by insertion of synthetic oligonucleotides cassettes, whose sense strand is 5′-CATGGCCTGACTG ACTGACTGACGAATGCA-3′, into the NcoI and BsmI sites of the pBR plasmid described above.

Monitoring β-galactosidase expression by colony colour and solution assays

Competent *E. coli* N567 cells were prepared by standard $CaCl_2$ methods and transformed with the appropriate plasmids by heat shock as described previously (Harada and Frankel, 1999). Each transformant was spread onto tryptone plates containing ampicillin (200 mg l^{-1}), chloramphenicol (20 mg l^{-1}), isopropyl β-D-thiogalactoside (IPTG) (0.05 mM) and Xgal (80 mg l^{-1}). When plates were incubated at 37°C for 18–24 h and further at 24°C for 16–24 h, blue intensity of these colonies was monitored and represented in the number of plusses from 0 to 8+ (see Fig. 3 for representative colours). Solution assays using ONPG were carried out as described previously (Cocozaki *et al.*, 2008) except for differences in antibiotic concentration for overnight culturing (100 mg l^{-1} ampicillin and 20 mg l^{-1} chloramphenicol).

Comparison of the copy number of reporter plasmids

Escherichia coli N567 cells transformed with reporter pAC plasmids were spread onto tryptone plates containing chloramphenicol (20 mg l^{-1}) and tetracycline (6 mg l^{-1}) as antibiotics and incubated at 37°C overnight. Three independent colonies were picked and cultured in tryptone medium containing both antibiotics with aeration at 37°C overnight. DNA was prepared from each overnight culture (1 ml) using QIAprep® Spin Miniprep Kit (Qiagen). Each DNA was eluted from the QIAprep spin column with 150 μl of Buffer EB (10 mM Tris-Cl, pH 8.5).

RNA probes for North Western analysis

RNAs (boxB and RRE) for the preparation of biotinylated probes were transcribed *in vitro* using T7 RNA polymerase. The RNAs were transcribed at 37°C for 2 h in the reaction mixture containing 80 mM HEPES-KOH (pH 8.1), 5 mM DTT, 1 mM spermidine, 0.001% Triton X-100, 80 mg ml^{-1} polyethylene glycol-8000, 42 mM $MgCl_2$, 8 mM GTP, 8 mM UTP, 4 mM ATP, 4 mM CTP, 0.4 U μl^{-1} RNasin® Plus RNase Inhibitor (Promega), 1.6 U μl^{-1} T7 RNA polymerase (Promega) and 500 nM synthetic DNA template (boxB, 5′-GGCCCTTTTTC AGGGCCTTCCTATAGTGAGTCGTATTAC-3′; RRE, 5′-GGC CTGTACCGTCAGCTTGCGCTGCGCCCAGACCTATAGTG AGTCGTATTAC-3′) annealed with T7 promoter sequence (5′-GTAATACGACTCACTATA-3′). RNAs were treated with Turbo™ DNase (0.02 U μl^{-1}) (Ambion) at 37°C for 15 min, purified on 12% polyacrylamide/8 M urea gels, eluted from the gels in 0.4 M sodium acetate (pH 5.4), isopropyl alcohol precipitated, and washed with 70% ethanol. Purified RNAs were labelled with biotin at their 5′-ends using 5′-EndTag™ Nucleic Acid Labeling System (Vector Laboratories) and treatment with biotin maleimide (Vector Laboratories). The RNAs were dissolved to 10 μM in renaturing buffer containing 200 mM Tris-HCl, pH 7.5 and 1 M NaCl, then heated at 70°C for 5 min and slowly cooled to room temperature for annealing.

Western blotting and North Western blotting

Escherichia coli N567 cells transformed with pBR plasmids were spread onto tryptone plates containing ampicillin (100 mg l^{-1}) and tetracycline (6 mg l^{-1}) as antibiotics and incubated at 37°C overnight. Generated colonies were inoculated to tryptone medium containing both antibiotics and incubated with aeration at 37°C overnight. The overnight cultures were diluted 1:50 in A medium (4 or 5 ml) containing 10.5 g l^{-1} K_2HPO_4, 4.5 g l^{-1} KH_2PO_4, 1.0 g l^{-1} $(NH_4)_2SO_4$ and 0.5 g l^{-1} trisodium citrate dihydrate supplemented with 0.4% (w/v) glucose, 1 μg ml^{-1} vitamin B1, 1 mM $MgSO_4$ and both antibiotics, and incubated with aeration for 6 h at 37°C. Then, each culture was divided into two aliquots and further incubated with or without addition of 0.5 mM IPTG for 30 min. The cells were harvested from each culture (1 or 2 ml) and washed with ice-cold acetone (1 ml). Each pellet was resuspended in 100 or 150 μl of 1× sodium dodecyl sulphate (SDS) sample buffer containing 62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 0.002% (w/v) bromophenol blue, 10% (v/v) glycerol and 5% (v/v) 2-mercaptoethanol, and heated at 100°C for 5 min. These solutions were separated by 15% SDS-PAGE and then transferred to polyvinylidene difluoride membrane (Millipore) in transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol). The amounts of the solutions loaded on the gels are otherwise noted. For Western blotting, the membrane was blocked with 2% (w/v) Membrane Blocking Agent (GE Healthcare) in TBST (20 mM Tris-HCl, pH 7.5; 154 mM NaCl; 0.05% Tween® 20). The blocked membrane was treated with antisera against N protein (a gift from Franklin, N. C.) and ECL Rabbit IgG, HRP (horse radish peroxidase)-Linked Whole Ab (from donkey) (GE Healthcare), and visualized with Amersham ECL Plus Western blotting Detection Reagents (GE Healthcare). Chemiluminescent signals were detected with Cooled CCD Camera System Light-Capture II (ATTO). For North Western blotting, the membranes after transferring were

blocked in wash buffer (10 mM Tris-HCl, pH 7.5; 100 mM NaCl; 1 mM EDTA, pH 8.0; 0.01% (v/v) NP-40) containing 2% (w/v) Membrane Blocking Agent (GE Healthcare) and 5 mg ml^{-1} ribonucleic acid from torula yeast type VI (Sigma) for 1–2 h at room temperature. The blocked membrane was washed by several changes of wash buffer. Then, the membrane was incubated with 50 nM biotinylated RNA probes in wash buffer for 1 h at 4°C. After several washings, the membrane was incubated in streptavidin-HRP (Invitrogen) diluted to 1:5000 with wash buffer for 1–5 h at room temperature or 4°C, washed by several changes of wash buffer, visualized and detected as for Western blotting.

Acknowledgements

The authors thank Naomi Franklin for the kind gift of N antisera, Akihiro Oguro and Koichi Ito for advice on experimental techniques, and Colin Smith for comments on the manuscript. This work was supported by a Grant-in-Aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (K.H. and S.M.).

References

Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. (2003) Differential modes of recognition in N peptide-boxB complexes. *Biochemistry* **42:** 14957–14967.

Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., *et al.* (1996) α helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273:** 1547–1551.

Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., *et al.* (1998) Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. *Nat Struct Biol* **5:** 203–212.

Campisi, D.M., Calabro, V., and Frankel, A.D. (2001) Structure-based design of a dimeric RNA-peptide complex. *EMBO J* **20:** 178–186.

Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K., and Das, A. (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage λ. *Proc Natl Acad Sci USA* **92:** 4061–4065.

Cilley, C.D., and Williamson, J.R. (1997) Analysis of bacteriophage N protein and peptide binding to *boxB* RNA using polyacrylamide gel coelectrophoresis (PACE). *RNA* **3:** 57–67.

Cilley, C.D., and Williamson, J.R. (2003) Structural mimicry in the phage φ21 N peptide-boxB RNA complex. *RNA* **9:** 663–676.

Cocozaki, A.I., Ghattas, I.R., and Smith, C.A. (2008) Bacteriophage P22 antitermination boxB sequence requirements are complex and overlap with those of λ. *J Bacteriol* **190:** 4623–4271.

Das, A. (1993) Control of transcription termination by RNA-binding proteins. *Annu Rev Biochem* **62:** 893–930.

Doelling, J.H., and Franklin, N.C. (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.

Franklin, N.C. (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.

Franklin, N.C. (2004) Morphing molecular specificities between Arm-peptide and NUT-RNA in the antitermination complexes of bacteriophages λ and P22. *Mol Microbiol* **52:** 815–822.

Gosser, Y., Hermann, T., Majumdar, A., Hu, W., Frederick, R., Jiang, F., *et al.* (2001) Peptide-triggered conformational switch in HIV-1 RRE RNA complexes. *Nat Struct Biol* **8:** 146–150.

Greenblatt, J., Nodwell, J.R., and Mason, S.W. (1993) Transcriptional antitermination. *Nature* **364:** 401–406.

Gusarov, I., and Nudler, E. (2001) Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell* **107:** 437–449.

Harada, K., and Frankel, A.D. (1998) *In vivo* selection of specific RNA binding polypeptides using a transcription anti-termination reporter assay. In *RNA:Protein Interactions: A Practical Approach.* Smith, C.W.J. (ed.). Oxford: Oxford University Press, pp. 217–236.

Harada, K., and Frankel, A.D. (1999) Screening RNA-binding libraries using a bacterial transcription antitermination assay. *Methods Mol Biol* **118:** 177–187.

Harada, K., Martin, S.S., and Frankel, A.D. (1996) Selection of RNA-binding peptides *in vivo*. *Nature* **380:** 175–179.

Harada, K., Martin, S.S., Tan, R., and Frankel, A.D. (1997) Molding a peptide into an RNA site by *in vivo* peptide evolution. *Proc Natl Acad Sci USA* **94:** 11887–11892.

Iwazaki, T., Li, X., and Harada, K. (2005) Evolvability of the mode of peptide binding by an RNA. *RNA* **11:** 1364–1373.

Jaeger, J.A., Turner, D.H., and Zuker, M. (1989) Improved predictions of secondary structures for RNA. *Proc Natl Acad Sci USA* **86:** 7706–7710.

Jaeger, J.A., Turner, D.H., and Zuker, M. (1990) Predicting optimal and suboptimal secondary structure for RNA. *Methods Enzymol* **183:** 281–306.

Kim, J.-S., and Pabo, C.O. (1998) Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc Natl Acad Sci USA* **95:** 2812–2817.

Lazinski, D., Grzadzielska, E., and Das, A. (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59:** 207–218.

Legault, P., Li, J., Mogridge, J., Kay, L.E., and 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.

McColl, D.J., Honchell, C.D., and Frankel, A.D. (1999) Structure-based design of an RNA-binding zinc finger. *Proc Natl Acad Sci USA* **96:** 9521–9526.

Mason, S.W., Li, J., and Greenblatt, J. (1992) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage λ. *J Biol Chem* **267:** 19418–19426.

Mogridge, J., Mah, T.-F., and Greenblatt, J. (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.

Mogridge, J., Mah, T.-F., and Greenblatt, J. (1998a) Involve-

ment of *boxA* nucleotides in the formation of a stable ribonucleoprotein complex containing the bacteriophage λ N protein. *J Biol Chem* **273:** 4143–4148.

Mogridge, J., Legault, P., Li, J., Van Oene, M.D., Kay, L.E., and Greenblatt, J. (1998b) 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.

Moore, M., Choo, Y., and Klug, A. (2001) Design of polyzinc finger peptides with structured linkers. *Proc Natl Acad Sci USA* **98:** 1432–1436.

Neely, M.N., and Friedman, D.I. (1998) Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol Microbiol* **28:** 1255–1267.

Neely, M.N., and Friedman, D.I. (2000) N-mediated transcription antitermination in lambdoid phage H-19B is characterized by alternative NUT RNA structures and a reduced requirement for host factors. *Mol Microbiol* **38:** 1074–1085.

Nudler, E., and Gottesman, M.E. (2002) Transcription termination and anti-termination in *E. coli. Genes Cells* **7:** 755–768.

Olson, E.R., Flamm, E.L., and Friedman, D.I. (1982) Analysis of *nutR*: a region of phage lambda required for antitermination of transcription. *Cell* **31:** 61–70.

Oubridge, C., Ito, N., Evans, P.R., Teo, C.-H., and Nagai, K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372:** 432–438.

Peled-Zehavi, H., Horiya, S., Das, C., Harada, K., and Frankel, A.D. (2003) Selection of RRE RNA binding peptides using a kanamycin antitermination assay. *RNA* **9:** 252–261.

Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **270:** 1200–1203.

Rees, W.A., Weitzel, S.E., Yager, T.D., Das, A., and von Hippel, P.H. (1996) Bacteriophage λ N protein alone can induce transcription antitermination *in vitro*. *Proc Natl Acad Sci USA* **93:** 342–346.

Ribas de Pouplana, L., Auld, D.S., Kim, S., and Schimmel, P. (1996) A mechanism for reducing entropic cost of induced fit in protein-RNA recognition. *Biochemistry* **35:** 8095–8102.

Rosenberg, M., Court, D., Shimatake, H., Brady, C., and Wulff, D.L. (1978) The relationship between function and DNA sequence in an intercistronic regulatory region in phage λ. *Nature* **272:** 414–423.

Salstrom, J.S., and Szybalski, W. (1978) Coliphage *λnutL*-: a unique class of mutants defective in the site of gene *N* product utilization for antitermination of leftward transcription. *J Mol Biol* **124:** 195–221.

Schärpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S., and Rösch, P. (2000) Antitermination in bacteriophage λ. The structure of the N36 peptide-*boxB* RNA complex. *Eur J Biochem* **267:** 2397–2408.

Somasekhar, G., Drahos, D., Salstrom, J.S., and Szybalski, W. (1982) Sequence changes in coliphage lambda mutants affecting the *nutL* antitermination site and termination by t_{L1} and t_{L2}. *Gene* **20:** 477–480.

Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L.A., and Weiss, M.A. (1997a) RNA recognition by a bent α-helix regulates transcriptional antitermination in phage λ. *Biochemistry* **36:** 12722–12732.

Su, L., Radek, J.T., Labeots, L.A., Hallenga, K., Hermanto, P., Chen, H., *et al.* (1997b) An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. *Genes Dev* **11:** 2214–2226.

Sugaya, M., Nishino, N., Katoh, A., and Harada, K. (2008a) Amino acid requirement for the high affinity binding of a selected arginine-rich peptide with the HIV Rev-response element RNA. *J Pept Sci* **14:** 924–935.

Sugaya, M., Nishimura, F., Katoh, A., and Harada, K. (2008b) Tailoring the peptide-binding specificity of an RNA by combinations of specificity-altering mutations. *Nucleosides Nucleotides Nucleic Acids* **27:** 534–545.

Tan, R., and Frankel, A.D. (1995) Structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci USA* **92:** 5282–5286.

Tan, R., Chen, L., Buettner, J.A., Hudson, D., and Frankel, A.D. (1993) RNA recognition by an isolated α helix. *Cell* **73:** 1031–1040.

Van Gilst, M.R., Rees, W.A., Das, A., and von Hippel, P.H. (1997) Complexes of N antitermination protein of phage λ with specific and nonspecific RNA target sites on the nascent transcript. *Biochemistry* **36:** 1514–1524.

Vieu, E., and Rahmouni, A.R. (2004) Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed *in vivo*. *J Mol Biol* **339:** 1077–1087.

Whalen, W., and Das, A. (1990) Action of an RNA site at a distance: role of the *nut* genetic signal in transcription antitermination by phage-λ N gene product. *New Biol* **2:** 975–991.

Whalen, W., Ghosh, B., and Das, A. (1988) NusA protein is necessary and sufficient *in vitro* for phage λ *N* gene product to suppress a ρ-independent terminator placed downstream of *nutL*. *Proc Natl Acad Sci USA* **85:** 2494–2498.

Wilhelm, J.E., and Vale, R.D. (1996) A one-hybrid system for detecting RNA–protein interactions. *Genes Cells* **1:** 317–323.

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

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

Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science* **244:** 48–52.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Nucleic Acids Research Advance Access published February 15, 2010
Nucleic Acids Research, 2010, 1–13
doi:10.1093/nar/gkq027

Identification of antisense RNA stem–loops that inhibit RNA–protein interactions using a bacterial reporter system

Akiko Yano[1], Satoru Horiya[2], Takako Minami[1], Eri Haneda[1], Makiko Ikeda[1] and Kazuo Harada[1,*]

[1]Department of Life Sciences, Tokyo Gakugei University, Koganei, Tokyo 184-8501 and [2]Department of Molecular Biology, The Jikei University School of Medicine, Minato-ku, Tokyo 105-8461, Japan

Received May 26, 2009; Revised and Accepted January 11, 2010

ABSTRACT

Many well-characterized examples of antisense RNAs from prokaryotic systems involve hybridization of the looped regions of stem–loop RNAs, presumably due to the high thermodynamic stability of the resulting loop–loop and loop–linear interactions. In this study, the identification of RNA stem–loops that inhibit U1A protein binding to the hpII RNA through RNA–RNA interactions was attempted using a bacterial reporter system based on phage λ N-mediated antitermination. As a result, loop sequences possessing 7–8 base complementarity to the 5′ region of the boxA element important for functional antitermination complex formation, but not the U1 hpII loop, were identified. *In vitro* and *in vivo* mutational analysis strongly suggested that the selected loop sequences were binding to the boxA region, and that the structure of the antisense stem–loop was important for optimal inhibitory activity. Next, in an attempt to demonstrate the ability to inhibit the interaction between the U1A protein and the hpII RNA, the rational design of an RNA stem–loop that inhibits U1A-binding to a modified hpII was carried out. Moderate inhibitory activity was observed, showing that it is possible to design and select antisense RNA stem–loops that disrupt various types of RNA–protein interactions.

INTRODUCTION

RNA–protein interactions play important roles in gene regulation, in the assembly of functional RNA–protein complexes such as the ribosome, and in viral replication. Therefore, molecules that regulate specific RNA–protein interactions provide an attractive means to dissect molecular steps of various biological processes, and to establish the validity of targeting an RNA–protein interaction for future drug design.

Various strategies have been developed for the inhibition of RNA–protein interactions, and can be classified into two groups depending on whether the protein or the RNA is targeted. Methods for targeting the protein include the use of RNA decoys or *in vitro* selected DNA or RNA aptamers. In the case of the human immunodeficiency virus (HIV) regulatory proteins Tat and Rev, RNA decoys corresponding to the respective RNA sites, the trans-activating response region (TAR) and the Rev-responsive element (RRE), as well as aptamers have been shown to inhibit viral replication (1). In particular, several Rev aptamers with affinities significantly higher than the wild-type RRE that compete with the RRE for Rev-binding have been generated (2).

Approaches for targeting RNA range from the use of small molecules (3) and peptides (4) to nucleic acid-based agents such as antisense RNA/DNA(5), siRNA(6) and aptamers (7,8). Targeting RNA using small molecules is a particularly attractive approach because such molecules may directly lead to the development of therapeutic agents; however, the desired specificity has been difficult to achieve by such compounds (3). On the other hand, nucleic-acid-based agents, such as antisense RNA/DNA and siRNA, have been shown to be effective in regulating gene expression, and a useful tool in elucidating molecular mechanisms (9,10). However, stable RNA secondary structure formation has been known to be an obstacle for both antisense oligonucleotides (11) and siRNA (12).

In many prokaryotic antisense control systems, RNA stem–loops are used for initial recognition, resulting in hairpin loop–loop ('kissing') and loop–linear interactions (13–15). Loop–loop interactions are also observed in

*To whom correspondence should be addressed. Tel: +81 42 329 7550; Fax: +81 42 329 7550; Email: harada@u-gakugei.ac.jp

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

RNA folding (16–18) and in the dimerization of retroviral genomic RNAs (19–21). These interactions appear to have been optimized for rapid and stable intermolecular interactions which are essential for their function (22). While loop–loop interactions generally use only five to seven complementary base pairs to join the two hairpin loops, this short complementary region may be an advantage since increasing affinity by increasing complementarity may be a source of decreased specificity (23,24).

However, the rational design of novel loop–loop interactions is not straightforward because the factors governing stable loop–loop complex formation appear to be complex and diverse, and the stability of loop–loop interactions are difficult to predict (25). For example, the stability of the extensively studied loop–loop interaction derived from RNA I and RNA II from plasmid ColE1, which consists of seven bases in the loop, of which all seven form base pairs, has been shown to increase 350-fold by simply inverting the loop sequences of the hairpins 5′ to 3′ (26). In this case, the major determinant of complex stability was found to be the identity of the base at the first and seventh position in the loop (27). An *in vitro* selected antisense stem–loop targeting the HIV TAR with a six base-pair loop, has an eight base loop with a closing G-A base-pair that has been shown to be crucial for stable complex formation (28). In the case of the dimerization initiation site (DIS) of HIV, six of the nine loop bases participate in base-pair formation, while the remaining three purine bases are important for stacking interactions (29–32). Surprisingly, stable loop–loop complexes with only two intermolecular G-C base-pairs have also been found (33).

In this study, we have attempted to identify RNA stem–loops that inhibit RNA–protein interactions through the formation of loop–loop interactions between the antisense RNA stem–loop and the target RNA structure. The complex formed between hairpin II of U1 snRNA (U1 hpII) and U1A protein, which is a component of the U1 snRNP, was chosen as a target (34). U1 hpII RNA contains a 10-nt apical loop, which is recognized by the N-terminal RRM of U1A protein with high specificity and affinity (35), and was expected to be a potential target for kissing complex formation. As it is difficult in general to predict the stability of loop–loop interactions as described above, an RNA stem–loop library was screened for sequences that bind to the target U1 hpII loop and inhibit U1A protein binding. A bacterial two-plasmid system for detecting RNA–polypeptide interactions based on bacteriophage λ N protein-mediated antitermination was used (Figure 1A) (36). In this system, N protein is expressed from a pBR322-based N expressor plasmid and LacZ is expressed from a pACYC184-based reporter plasmid containing the nut site (boxA–boxB) and four terminators upstream of *LacZ*. Binding of the N-terminal RNA-binding domain of N (N peptide) to the nut site boxB stem–loop of the nascent RNA transcript nucleates the formation of an antitermination complex, which includes the bacterial host factors NusA, NusB, NusG and S10, thereby causing transcription antitermination by RNA polymerase and expression of LacZ. This system can be modified to study heterologous RNA–polypeptide interactions by replacing the pBR and pACYC plasmid DNA regions corresponding to the N peptide and boxB RNA with those of the peptide and RNA of interest (37).

In this study, the N/boxB interaction was replaced by that of the U1A protein and the U1 hpII RNA, and an RNA stem–loop library with a completely randomized 10-nt loop was placed 38 nt downstream of U1 hpII (Figure 1B). Individual clones showing repressed reporter gene expression were isolated and the sequences were analyzed for intramolecular RNA–RNA binding. Contrary to our expectations, it was found that the selected RNA stem–loop was most likely binding to the boxA region, which is an essential element in antitermination complex formation (Figure 1C, left) (38,39). We therefore used a rational approach to design RNA stem–loops that inhibit the binding of the U1A protein to a U1 hpII variant (Figure 1C, right). The result shows that it is possible to inhibit RNA–protein interaction in an efficient way using antisense RNA stem–loops.

MATERIALS AND METHODS

Selection of RNA stem–loops that inhibit U1A-mediated antitermination complex formation using a bacterial reporter assay

The pAC hpII reporter plasmid containing a randomized stem–loop library downstream of the U1 hpII site was constructed in the following manner. A synthetic oligonucleotide cassette containing the boxA of nut, U1 hpII with a 9-bp stem, and a BsrGI site (denoted boxA–hpII; Tables S1 and S2) was cloned into the unique PstI and BamHI sites of pAC nut (40), to give pAC hpII. RNA stem–loop library dsDNA was prepared by annealing the two synthetic oligonucleotides, BsrGI-linker-1 and random-stem–loop (Tables S1 and S2) complementary at the 3′-end, and second-strand synthesis with *Taq* polymerase. The resulting dsDNA was introduced into the BsrGI and BamHI sites of the pAC hpII plasmid.

The procedure for the *in vivo* selection of library sequences that resulted in reduced antitermination activity was based on a previously described method for peptide selection of RNA binders (4,37). For the primary screen, the RNA stem–loop library plasmid was prepared by ligation of the library insert (7.5 ng) into pAC hpII reporter plasmid (250 ng), followed by phenol extraction and concentration to 10 μl using a filter unit (Montage PCR, Millipore). Plasmids were electroporated into N567/pBR U1A-N cells (80 μl) in 1-mm cuvettes at 2.0 kV using 1 μl of the above solution, Super Optimal Broth (SOC) medium (5 ml) was added immediately after electroporation, and cells were allowed to recover by incubating at 37°C; for 1 h. Transformants were spread onto tryptone plates (ϕ150 mm) containing ampicillin (100 μg/ml), chloramphenicol (20 μg/ml), isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.05 mM) and X-gal (80 μg/ml) and incubated at 37°C; for 28 h. A

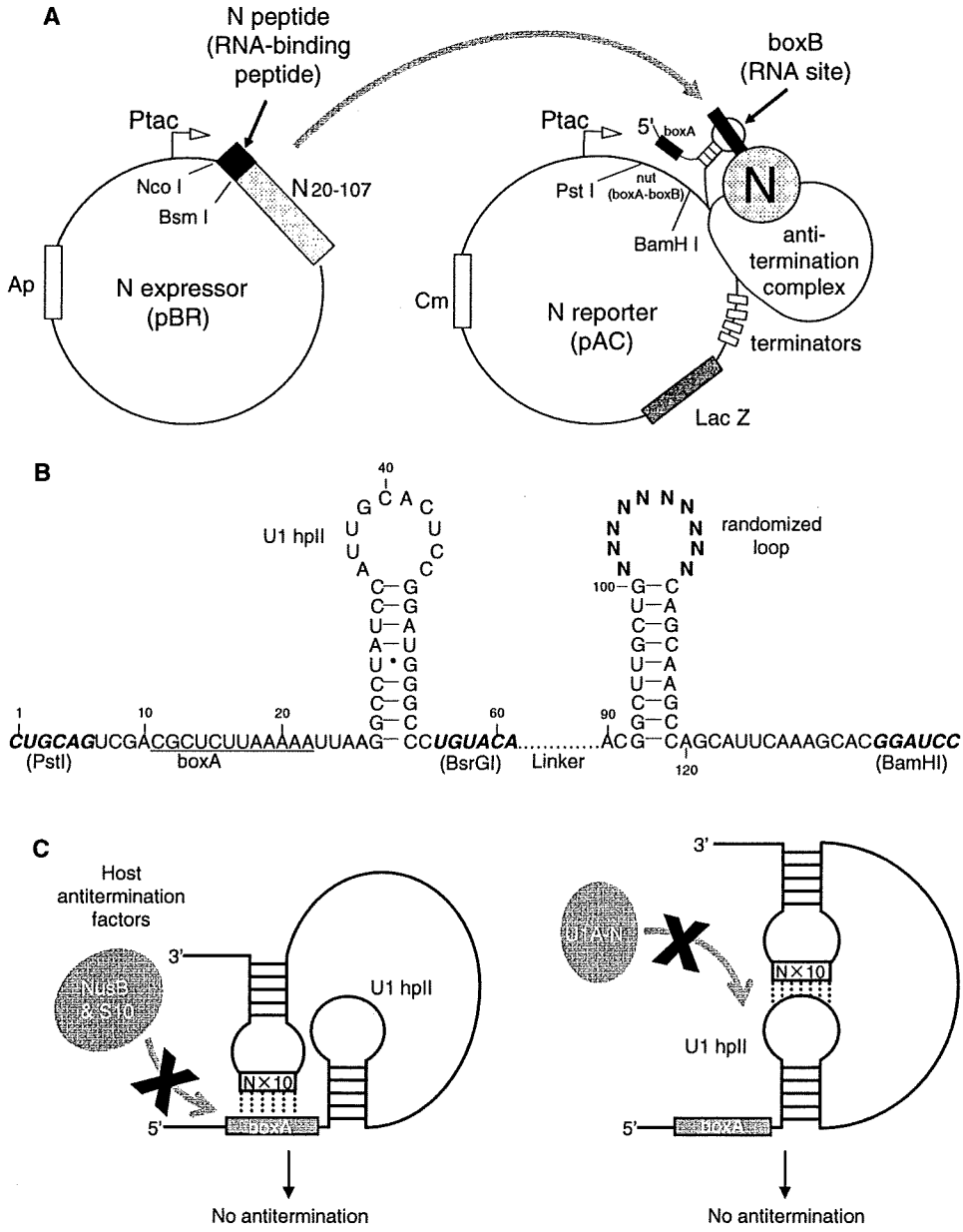

Figure 1. The strategy and design of a bacterial assay for detecting RNA loop–loop interactions that inhibit U1A–hpII-mediated antitermination complex formation. (**A**) A bacterial two-plasmid system based on phage λ N-mediated antitermination for detection of RNA–polypeptide interactions. (**B**) The secondary structure of the region of the RNA transcript containing the target U1 hpII and downstream RNA stem–loop library. The randomized nucleotides are indicated by N, and the regions corresponding to the PstI, BsrGI, and BamHI sites are shown in italic. (**C**) Possible mechanism for the disruption of antitermination complex formation by antisense RNA stem–loops.

total of 1.5×10^5 colonies were obtained, and the degree of colony color was visually scored by comparison with the blue intensity of a standardized set of controls that included the U1A–U1 hpII interaction (5+), as well that of the RSG-1.2 peptide and the HIV Rev-response element (RRE) (2+), the HIV Rev peptide and the RRE (1+), U1A and the RRE (0.5+) and the Rev peptide and U1 hpII (0). Individual light blue colonies (white to 3+; 1056) were then grown to saturation in 96-well plates containing tryptone and antibiotics, cultures were pooled and plasmid DNA was isolated. The library region of the selected pAC plasmid was amplified by polymerase chain reaction (PCR) using pAC forward primer (5′-GGCTTATCGAAATTAATA CG-3′) and reverse primer (5′-ACGGTAAGAGTGCCA GTG-3′). The amplified fragments were then digested with BsrGI and BamHI, phenol-extracted and purified on a native 8% polyacrylamide gel.

In the secondary screen, library inserts from the primary screen (white to 3+ and >4+) were reintroduced into the pAC hpII reporter plasmid. Ligation mixtures were phenol-extracted and concentrated using Montage PCR (Millipore), and individually electroporated into N567/pBR U1A-N cells as described above. Transformants were spread onto X-gal plates, and plasmid DNA from individual colonies with various intensity of blue color were isolated and used to transform N567/pBR U1A-N cells by heat shock to confirm activity. The sequences of 25 clones with varying colony colors were determined.

Mutational analysis of intramolecular RNA stem–loop binding to boxA

Reporter plasmids with base substitutions in the boxA and antisense stem–loop region were prepared by ligating synthetic double-stranded inserts corresponding to the boxA–hpII (PstI–BsrGI fragments) and the antisense stem–loop (BsrGI–BamHI fragments) region into the PstI and BamHI sites of the pAC plasmid. The boxA–hpII inserts were prepared by annealing two complementary oligonucleotides listed in Supplementary Table S1 in the combinations indicated in Supplementary Table S2. BoxA–hpII and boxA–hpII A_9 dsDNAs were digested with BsrG I, phenol-extracted and ethanol-precipitated. Inserts corresponding to the antisense stem–loops of clones **1-4**, **1-2** and **1-8** were amplified by PCR using a pAC forward primer and reverse primer from selected plasmids, then digested with BsrGI and BamHI, phenol-extracted, and purified on a 10% polyacrylamide gel. antisense stem–loops with base substitutions were prepared by the same method as the preparation of library insert using the oligonucleotides listed in Supplementary Table S1 in the combinations indicated in Supplementary Table S2. The activities of these constructs were assessed using the LacZ colony color assay with N567/pBR U1A-N cells, and quantitated by using a β-galactosidase solution assay (41).

Native PAGE analysis of intra- and intermolecular RNA–RNA interactions between the boxA region and the selected antisense RNA stem–loops

RNA substrates (114-mer) for intramolecular RNA binding experiments and 18- or 34-mer RNAs for intermolecular RNA-binding experiments were transcribed *in vitro* by T7 RNA polymerase. DNA templates for 114-mer RNAs were PCR-amplified from the pAC constructs used in the *in vivo* antitermination assay using a T7 boxA primer and a reverse primer (Supplementary Table S3). 114-mer RNA substrates (a total of 48 pmol per experiment) in H_2O (typically 24 μl) were heated at 95°C for 5 min and immediately cooled on ice for 5 min. To this solution, 4X PN buffer was added to give a final concentration of 10 mM sodium phosphate (pH 7.0) and 50 mM NaCl (1X PN buffer) and the mixture was incubated at 37°C; for 30 min, cooled on ice for 20 min and one-fourth of the volume of loading buffer containing 50% glycerol was added. One-fourth of this solution was analyzed separately by gel electrophoresis at 4°C; on 10% native polyacrylamide gels (acrylamide:methylenebisacrylamide = 20:1) in TBM (89 mM Tris, 89 mM borate, and 0.1 and 0.5 mM $MgCl_2$). RNA bands were visualized by staining with ethidium bromide and irradiation with a UV transilluminator.

Gel mobility shift assays using ^{32}P-labeled RNAs were carried out as follows. DNA templates for boxA (18mer) and antisense stem–loop (34mer) RNAs were prepared by annealing synthetic oligonucleotides (Supplementary Table S3) and T7 promoter DNA (5′-GTAATACGACT CACTATA-3′). Internally ^{32}P-labeled RNAs (SL1-4 and SL1-4 GC) redissolved to 20 nM in 1X PN buffer were heated at 95°C and quickly cooled on ice. The RNAs were diluted to 5 nM with 1X PN buffer containing 50% glycerol. Unlabeled RNAs (boxA and boxA GC) diluted to 64 μM in 1X PN buffer were heated at 95°C and quickly cooled on ice, then diluted to 0.25–32 μM in 1X PN buffer. The 5 nM labeled RNAs (0.5 μl) were mixed with 4 vol of 1X PN buffer or the unlabeled RNAs (0.25–64 μM) and incubated at 37°C for 30 min, then chilled on ice for 20 min and subjected to 12% PAGE in TBM (0.1 mM $MgCl_2$) at 4°C. The gels were dried up on 3MM Chromatography paper (Whatman) and analyzed with a fluorescence/radioisotope image analyzer (Fujifilm FLA-2000).

Inhibition of U1A protein binding to a mutant U1 hpII RNA hairpin by designed antisense stem–loop RNAs

pAC reporter plasmids containing mutant U1 hpIIs were constructed by insertion of synthetic oligonucleotide cassettes prepared using the oligonucleotides in Supplementary Table S4 in the combinations shown in Supplementary Table S5 into the PstI and BamHI site of pAC nut$^-$ as described above. The U1A protein-binding activities of the resulting hpII RNA mutants were determined by transformation of N567/pBR U1A cells using the resulting pAC hpII mutant plasmids, and scoring antitermination activity as described above.

pAC reporter plasmids containing the U1 hpII Δ16/18A/10G (hpIIΔ^{DIS}) mutant and a downstream BsrGI site for subsequent cloning were prepared by annealing oligonucleotides shown in Supplementary Table S4, and inserting into the PstI and BamHI site of pAC nut$^-$ to yield pAC hpIIΔ^{DIS}. Oligonucleotide cassettes encoding the downstream antisense stem–loop structures, prepared by annealing oligonucleotide shown in Supplementary Table S4 in the combinations shown in Supplementary Table S5, second-strand synthesis, and digestion by BsrGI and BamHI, were inserted into the BsrGI and BamHI site of pAC hpIIΔ^{DIS} to give pAC hpIIΔ^{DIS}–aSL. Inhibition of the U1A–hpIIΔ^{DIS} interaction by the antisense stem–loops was determined by transformation of N567/pBR U1A cells using the resulting pAC hpIIΔ^{DIS}–aSL plasmids, and scoring antitermination activity as described above.

Native PAGE analysis of inhibition of the U1A-hpIIΔ^{DIS} interaction by the antisense stem–loops

RNAs were transcribed *in vitro* by T7 RNA polymerase from synthetic template DNAs (Supplementary Table S6)

annealed with T7 promoter DNA. The plasmid expressing U1A with a C-terminal $(His)_6$-tag was constructed as follows. The sequence of $U1A_{1-102}$ followed by a glycine spacer was PCR-amplified from pBR U1A-N (40) and cloned into pBAD/*Myc*-His A (Invitrogen) at the NcoI and EcoRI sites. The generated plasmid was transformed into TOP 10 *E. coli* cells, and U1A-Myc-His was expressed by the induction of L-arabinose and purified using a Ni-NTA column (QIAGEN). The purified U1A-Myc-His protein was dialyzed against dialysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl and 20% glycerol and the concentration of the protein was determined by UV absorption at 280 nm using $\varepsilon = 5120\,M^{-1}\,cm^{-1}$(42). The protein was diluted to 163.84 μM with dialysis buffer and further diluted with 4 vol of TBS containing Triton X-100 [20 mM Tris–HCl (pH 7.6), 50 mM NaCl and 0.5 % Triton X-100]. The diluted protein was serially diluted to 0.125–16384 nM with the 1:4 mixture of dialysis buffer and TBS containing Triton X-100. Unlabeled RNAs (antisense stem-loop SL and antisense stem-loop SL 15C) diluted to 8192 nM in 1X PN buffer were heated at 95°C and quickly cooled on ice, then diluted to 8-4096 nM in 1X PN buffer. ^{32}P-labeled RNAs (hpII and hpIIΔ^{DIS}) diluted to 2 nM in 1X PN buffer were heated at 95°C, quickly cooled on ice and diluted to 0.1 nM for hpII or 0.8 nM for hpIIΔ^{DIS} with 1X PN buffer. The labeled RNAs (2.5 μl) were mixed with the same volume of 4X binding buffer [40 mM Tris–HCl (pH 7.6) 850 mM NaCl, 1.2 mM $MgCl_2$, 1.0 mg/ml tRNA and 40% glycerol]. To this mixture, 2.5 μl of the U1A protein solution (0 or 0.125–32768 nM) and 2.5 μl of the unlabeled RNA solution (0 or 8-8192 nM) were added and incubated at room temperature (22°C) for 20 min and further on ice for 20 min. Then, the mixtures (10 μl) were subjected to 12% PAGE in TBM (0.4 mM $MgCl_2$) at 4°C. The gels were analyzed as described above.

RESULTS

Screening for RNA stem–loops that inhibit RNA–protein interactions using a bacterial reporter assay

The bacterial reporter system for detecting RNA–protein interactions described above was modified so that antisense RNA stem–loops that inhibit the U1A-hpII interaction through RNA–RNA interactions may be identified (Figure 1A and B). First, a DNA insert containing the boxA site followed by the wild-type U1 hpII stem–loop with an additional 4 bp in the stem to stabilize the structure and a BsrGI site was inserted into the PstI and BamHI sites of the pAC reporter plasmid, thereby yielding the pAC hpII reporter plasmid. Next, an RNA stem–loop library, consisting of a 9-bp stem with a 10-nt randomized loop encoding $4^{10} = 1.0 \times 10^6$ sequences, was introduced 38 nt downstream of the hpII stem–loop using the BsrGI and BamHI sites (Figure 1B). N567/pBR U1A-N cells that express the U1A-N fusion protein were transformed using the pAC library plasmid and spread onto tryptone plates containing X-gal. Colonies exhibiting lighter blue intensities compared with that of the U1 hpII–U1A interaction, scored as 5+, were selected. Antitermination activities were visually scored by colony color intensities as the number of plusses, with more plusses indicating stronger intensities, by comparison with a set of standardized controls. We have previously shown that antitermination activities visually scored as the intensity of colony color correlate well with the stability of the RNA–peptide interaction replacing the boxB RNA and N peptide, and provide an accurate measure for the stability of the antitermination complex (37,43).

In the primary screen, $\sim 1.5 \times 10^5$ colonies were screened, of which ~1.8% were white or light blue (white, 0.49%; 1+, 0.09%; 2+, 0.54%; and 3+, 0.65%) compared to the U1 hpII–U1A interaction (5+). A total of 1056 light blue colonies with varying blue intensity (white, 96; 1+, 125; and 2+~3+, 835) and 96 blue colonies (>4+) as a control were individually pooled and plasmid DNA was isolated. In the secondary screen, to eliminate pAC reporter plasmid-related false-negative clones (37), the library region was PCR-amplified, reintroduced into pAC U1 hpII reporter plasmid and transformed into N567/pBR U1A cells yielding $\sim 9.7 \times 10^3$ colonies, of which ~29% were white or light blue (white, 2.0%; +, 0.8%; 2+, 1.3%; and 3+~4+, 24.9%) on X-gal plates. The library region of the control plasmid DNA isolated from the pool of blue colonies (>4+) was also reintroduced into pAC U1 hpII reporter plasmid, and transformed into N567/pBR U1A cells yielding $\sim 1.3 \times 10^3$ colonies with varying blue intensity (white, 1.8%; 1+, 0.2%; 2+, 1.3%; 3+~4+, 55.4%; >5+, 27.9%; and 13.4% that could not be determined due to small colony size) on X-gal plates.

Plasmids were isolated from clones of varying blue intensity and reintroduced into N567/pBR U1A cells to confirm colony color. As a result, 13 unique sequences with colony colors 1 + ~4+ were found, and the six clones exhibiting weak colony color (1 + ~3+) possessed 7–8 base complementarity to the 5′ region of the boxA element, and not to U1 hpII loop (Table 1). This suggested that these stem–loops bind to the boxA region and inhibit binding of host factors (NusB and S10), which are important for antitermination (38,39,44). Interestingly, three clones exhibiting particularly low antitermination activity, clone **1-2**, **1-4** and **1-8**, all possessed the same 8 nt sequence complementary to the 5′-region of boxA at the 3′-side of 10-nt loop. The remaining seven clones showed reduced antitermination activity (3 + ~4+), but little complementarity to both the hpII loop and the boxA region (Table 1). The RNA loop of **2-11** had 2-nt deletion in the linker region. Using an RNA folding algorithm, Mulfold, it was predicted that this nucleotide deletion caused incorrect folding, resulting in the linker region binding to boxA. The RNA loops of a number of the remaining six clones showed 3–4 nt complementary to the U1 hpII loop (Table 1). Although it is possible that these antisense RNA stem–loops bound to U1 hpII and destabilized the U1 hpII–U1A interaction, further analysis was not carried out since the inhibitory effect was small.

Table 1. The nucleotide sequences of selected clones that exhibited low antitermination activity

	boxA	hpII loop	random loop		
Clone#	GUCGA**CGCUCUUAAAAA**- -	-AUUGCACUCC- -	-NNNNNNNNNN	clones	β-gal (X-gal)
1-1	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CUACGUCGAC	2	3+
1-2	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-UUGCGUCGAC	5	1+
1-4	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-ACGCGUCGAC	3	1+
1-7	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-AAGGCGUCGA	1	3+
1-8	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CUGCGUCGAC	5	1.5+
1-13	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CUAGCGUCGA	1	3+
1-16	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CAAUGUCUCA	2	4+
2-3	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-AAUAAGCCAC	1	4+
2-4	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CCCUUCCAUU	1	4+
2-11	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-AUCCAUCAAA	1	3+
2-14	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-CACCUGCAUA	1	4+
2-20	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-UAAAUACCCG	1	4+
4-3	GUCGACGCUCUUAAAAA- -	-AUUGCACUCC- -	-AAAACUCUGC	1	4+

Bases in the boxA 5′-region and selected loop that were complementary to each other are underlined. Bases in the hpII loop and selected loop that were complementary to each other are shaded. The number of plusses indicates blue colony color scored by the colony color (X-gal) assay using the following controls: U1A/hpII, 5+; RSG-1.2/RRE, 2+; Rev/RRE, 1+; U1A/RRE, 0.5+; Rev/hpII, 0+.

In vivo mutational analysis of intramolecular RNA stem–loop binding to boxA

In order to confirm the presence of intramolecular binding of the selected loop sequences to the 5′-region of boxA, boxA and loop mutants based on **1-2**, **1-4** and **1-8** that showed low antitermination activity (1+∼1.5+) were further examined *in vivo* using the antitermination assay (Table 2). LacZ expression scored by colony color correlated well with β-galactosidase activities quantitated using o-nitrophenyl β-D-galactopyranoside (ONPG). Clone **1-4** U_{107}, having a C to U substitution in the selected RNA loop, showed high antitermination activity (Table 2, 4+), presumably due to the disruption of the RNA–RNA interaction; however, the activity was not as high as wild-type U1 hpII-U1A (5+). This weak disruption of inhibitory activity for the **1-4** U_{107} mutant could be the result of G-U base pair formation, and residual binding of the loop sequence and the boxA region. Conversely, clone **1-4** A_9 mutant, having a G to A substitution upstream of boxA, showed antitermination activity (5+) similar to that of the wild-type U1 hpII-U1A (Table 2), suggesting that the intramolecular base pairing had been completely disrupted. Clone **1-4** A_9/U_{107} mutant containing both the G to A substitution upstream of boxA and the C to U substitution in the original **1-4** RNA loop (Table 2), which restores complementarity, showed low antitermination activity (3+); however, the antitermination activity was not as low as the original clone **1-4** (Table 2, 1+). This weak restoration of inhibitory activity for clone **1-4** A_9/U_{107} mutant could be the result of the substitution of a G-C pair in clone **1-4** to an A-U base pair in the double mutant. In an attempt to bring about more dynamic change in disruption and restoration of inhibitory activity, mutants containing double nucleotide substitutions were constructed. The mutants containing CG to GC substitutions either upstream of boxA or in the original **1-2**, **1-4** and **1-8** RNA loop, denoted **1-2**, **1-4** and **1-8** G_8C_9 or **1-4** $G_{107}C_{108}$, showed high antitermination activity (5+) comparable to wild-type U1 hpII-U1A (Table 2), suggesting that double nucleotide substitutions abolished the intramolecular RNA–RNA interactions. Conversely, CG to GC double nucleotide substitutions both upstream of boxA and in the original **1-4** RNA loop, denoted **1-4** $G_8C_9/G_{107}C_{108}$, restored complementarity and low antitermination activity (2+) comparable to the original clone **1-4** (Table 2). To confirm that the U1 hpII stem–loop was not involved in the inhibition of antitermination activity, the U1 hpII region was replaced by the boxB element (data not shown). As a result, the clone **1-4** SL repressed the antitermination activity, while the U_{107} and $G_{107}C_{108}$ mutants had no effect. Taken together, the above results strongly suggested that the selected loop sequences were binding to the boxA region and disrupting antitermination complex function.

Table 2. Mutational analysis of intramolecular binding of the selected RNA loops to the boxA 5′-region

			β-gal	
Clone#	boxA	random loop	Xgal	ONPG
1-4	GUCGA**CGCUCUUAAAAA**- -	-ACGCGUCGAC	1+	31.7
1-4 U_{107}	GUCGA**CGCUCUUAAAAA**- -	-ACGCGUUGAC	4+	74.6
1-4 A_9	GUCAA**CGCUCUUAAAAA**- -	-ACGCGUCGAC	5+	168
1-4 A_9/U_{107}	GUCAA**CGCUCUUAAAAA**- -	-ACGCGUUGAC	3+	62.6
1-4 $G_{107}C_{108}$	GUCGA**CGCUCUUAAAAA**- -	-ACGCGUGCAC	5+	174
1-4 G_8C_9	GUGCA**CGCUCUUAAAAA**- -	-ACGCGUCGAC	5+	116
1-4 $G_8C_9/G_{107}C_{108}$	GUGCA**CGCUCUUAAAAA**- -	-ACGCGUGCAC	2+	43.2
1-2	GUCGA**CGCUCUUAAAAA**- -	-UUGCGUCGAC	1+	39.1
1-2 G_8C_9	GUGCA**CGCUCUUAAAAA**- -	-UUGCGUCGAC	5+	123
1-8	GUCGA**CGCUCUUAAAAA**- -	-CUGCGUCGAC	1.5+	40.8
1-8 G_8C_9	GUGCA**CGCUCUUAAAAA**- -	-CUGCGUCGAC	5+	85.2

Bases in the boxA 5′-region and selected loop that were complementary to each other are underlined. The boxA nucleotides are indicated in bold. The number of plusses indicates blue colony color scored by the colony color (X-gal) assay using the following controls: U1A/hpII, 5+; RSG-1.2/RRE, 2+; Rev/RRE, 1+; U1A/RRE, 0.5+; Rev/hpII, 0+.

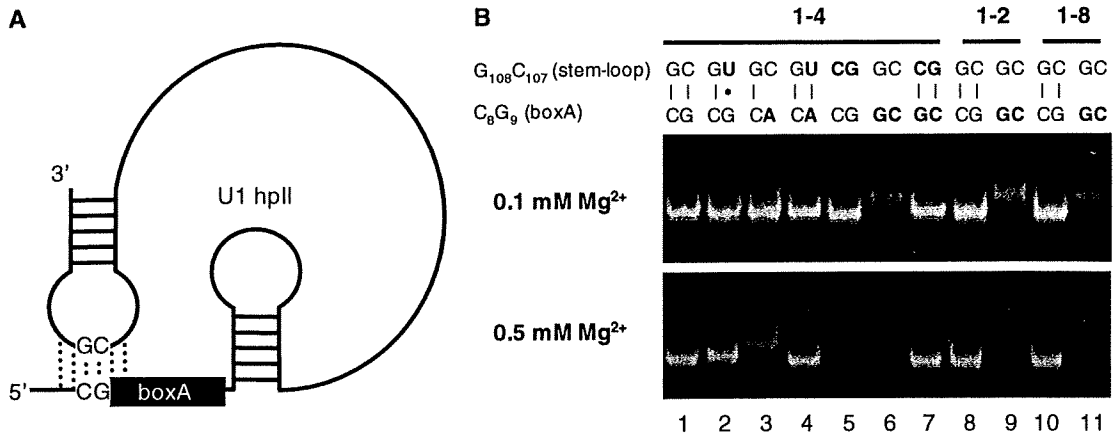

Figure 2. Native gel electrophoretic analysis of the intramolecular binding of the selected RNA stem–loop to boxA. (**A**) Schematic representation of the pseudoknot-like RNA folding by intramolecular RNA–RNA binding. (**B**) TBM gel (0.1 mM and 0.5 mM Mg^{2+}) analysis of RNA substrates. All gels were run at 4°C;. Lane 1, **1-4**; lane 2, **1-4** U_{107}; lane 3, **1-4** A_9; lane 4, **1-4** A_9/U_{107}; lane 5, **1-4** $G_{107}C_{108}$; lane 6, **1-4** G_8C_9; lane 7, **1-4** $G_8C_9/G_{107}C_{108}$; lane 8, **1-2**; lane 9, **1-2** G_8C_9; lane 10, **1-8**; lane 11, **1-8** G_8C_9.

In vitro mutational analysis of intramolecular RNA stem–loop binding to boxA

In order to further confirm that the selected loop sequences bound to the boxA region, the structures of RNA substrates derived from the mutants used in the *in vivo* assay described above were analyzed by native polyacrylamide gel electrophoresis. It was predicted that RNA substrates that form pseudoknot-like structures as a result of intramolecular RNA–RNA binding as shown in Figure 2A would be compactly folded and migrate faster on native gels compared to RNAs that cannot form pseudoknots.

When the 11 RNA substrates shown in Table 2 were run on a Tris-boric acid (TB) gel containing 0.1 mM Mg^{2+} (TBM gel), most of the RNAs showed similar mobilities with the exception of the G_8C_9 mutants of **1-4**, **1-2** and **1-8**, which showed lower mobilities (Figure 2B). Since pseudoknot formation has been shown to be stabilized by divalent metal ions (25), the mobility of the pseudoknot-forming RNAs was expected to increase relative to the non-pseudoknot forming RNAs with increasing Mg^{2+} in the gel. Indeed, the original **1-4** RNA showed a significant increase in mobility relative to the **1-4** A_9 mutant RNA at high Mg^{2+} concentrations, suggesting the formation of a compact pseudoknot conformation for **1-4** RNA as opposed to an open conformation for **1-4** A_9 (Figure 2B, lanes 1 and 3). In contrast, the **1-4** U_{107} mutant only showed a slight difference in mobility compared to **1-4** (Figure 2B, lanes 1 and 2), presumably due to the retention of a G-U wobble base-pair, agreeing well with residual inhibitory activity observed *in vivo* (Table 2). On the other hand, the **1-4** A_9/U_{107} double mutant, which was expected to restore pseudoknot formation, showed a high mobility comparable to that of **1-4** (Figure 2B, lanes 4). Similarly, the two mutants, **1-4** $G_{107}C_{108}$ and **1-4** G_8C_9, showed lower gel mobilities compared to the original **1-4**, indicating that pseudoknot formation had been disrupted (Figure 2B, lanes 5 and 6). Conversely, the double mutant **1-4** $G_8C_9/G_{107}C_{108}$, in which pseudoknot formation was expected to be restored, a high gel mobility comparable to **1-4** was observed (Figure 2B, lane 7). The gel mobility of the G_8C_9 mutants of the RNAs corresponding to clones **1-2** and **1-8** also lead to a decrease in gel mobility relative to the original molecules (Figure 2B, lanes 8–11), indicative of destabilization of pseudoknot formation. These results correlate with the *in vivo* mutational analysis, strongly supporting the notion that the selected loop sequences are binding to the boxA region of the RNA transcript and disrupting antitermination complex formation.

In vitro mutational analysis of intermolecular RNA stem–loop binding to boxA

The binding affinity and specificity of the selected stem–loop RNAs toward the boxA RNA region was examined by a gel mobility shift assay (Figure 3). The boxA RNA and the mutant boxA containing the double nucleotide substitution, denoted as boxA and boxA GC, respectively, was mixed in different concentrations with internally ^{32}P-labeled antisense stem–loop from clone **1-4** and the mutant with the double nucleotide substitution, denoted SL1-4 and SL1-4 GC, respectively, and analyzed on TBM gels containing 0.1 mM Mg^{2+}. SL1-4 was found to bind to the boxA RNA with an apparent K_d of 0.50 μM, while no interaction with boxA GC could be detected up to a concentration of 50 μM. Similarly, only a very weak interaction ($K_d > 50$ μM) could be observed between the mismatched SL1-4 GC and wild-type boxA. On the other hand, restoration of complementarity in the case of SL1-4 GC and boxA GC resulted in a considerable recovery of binding affinity to 3.8 μM, which is somewhat weaker than that of SL1-4 and boxA, but correlates well with the slightly weaker inhibitory effect of the SL1-4 GC/boxA GC interaction in vivo (2+) (Table 2), compared to the wild-type SL1-4/boxA interaction (1+). Taken together, these results support the notion that the

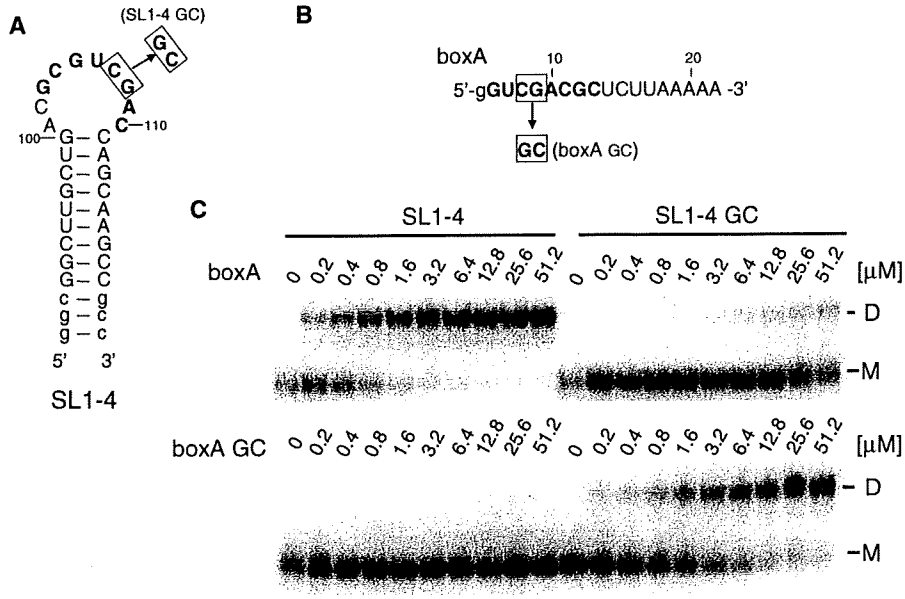

Figure 3. Gel mobility shift analysis of the intermolecular binding of the antisense stem–loop and boxA. (**A**) The secondary structure of the 'inhibitor' RNA stem–loop substrate. (**B**) The sequence of the boxA RNA substrates. (**C**) Gel shift analysis of the binding of ^{32}P-labeled wild-type and mutant SL1-4 RNAs with boxA and mutant RNAs on a TBM gel (0.1 mM Mg^{2+}) at 4°C.

antisense stem–loop bound to the 5′-region of boxA and prevented the association of host factors important for antitermination complex formation.

In vivo analysis of the effect of the disruption of the stem region of antisense RNA stem–loops

The importance of the stem–loop structure for efficient inhibition of RNA–protein interaction by the antisense RNA stem–loops was investigated by disrupting the upper stem region and measuring antitermination activity using the *in vivo* reporter system (Figure 4). When the single and double mismatches (1 bp and 2 bp, respectively) were introduced into the upper stem of clones **1-4** and **1-4** $G_8C_9/G_{107}C_{108}$, a gradual increase in antitermination from 1+ to 3.5+ and from 1.5+ to 3+ was observed, respectively. On the other hand, constructs with mismatches in the loop region, clones **1-4** $G_{107}C_{108}$ and **1-4** G_8C_9, which do not inhibit RNA–protein interaction, showed no change. This suggested that the structure of the antisense RNA stem–loop, and not just the loop region is important for efficient inhibition of RNA–protein interaction.

Rational design of a modified U1 hpII RNA and cognate antisense RNA stem–loop

In order to show that an RNA–protein interaction resembling that of U1A–hpII could be disrupted by RNA loop–loop complex formation, rational design of a modified U1 hpII RNA and cognate antisense RNA stem–loop was carried out. The U1 hpII was modified so that it resembles the HIV dimerization initiation site (DIS) loop, which is known to dimerize through a loop–loop interaction, while still retaining significant affinity toward the U1A protein. Since the requirements for stable loop–loop complex formation by HIV DIS-like RNA stem–loops are fairly well understood both *in vitro* and *in vivo* (45–49), this would allow the rational design of an antisense RNA stem–loop that binds to the modified U1 hpII loop, which may inhibit binding to the U1A protein. Modification of the U1 hpII should also make possible the adjustment of the affinity toward U1A protein, since the wild-type interaction may be too strong for inhibition by RNA loop–loop complex formation.

The HIV DIS loop contains nine bases, of which positions 3 through 8 are involved in Watson–Crick base-pair formation, and the purine residues at positions 1, 2 and 9 are important for base stacking interactions (Figure 5A) (45,46). A sheared A–A base-pair was proposed to be important for positions 1 and 9 (46). The relative stability of the loop–loop complexes was found to correlate with the stability of the duplex region as estimated using nearest-neighbor parameters determined by Turner and co-workers (50), so that a higher G-C content leads to stable complex formation (47,48). On the other hand, the U1 hpII loop has a 10-nt loop (Figure 5B), which is one base larger than that of the DIS loop, of which the first 7 nt contact the U1A protein (34), and the remaining three nucleotides can be replaced by unnatural ethylene glycol linkers (51).

In order to modify the U1 hpII loop so that it resembles the DIS loop, first, the eighth nucleotide in the U1 hpII loop, U16, was deleted, and C18 was substituted by an A to yield Δ16/18A, which resulting in a drop in antitermination activity of 5+ to 4+ (Figure 5E). Next,

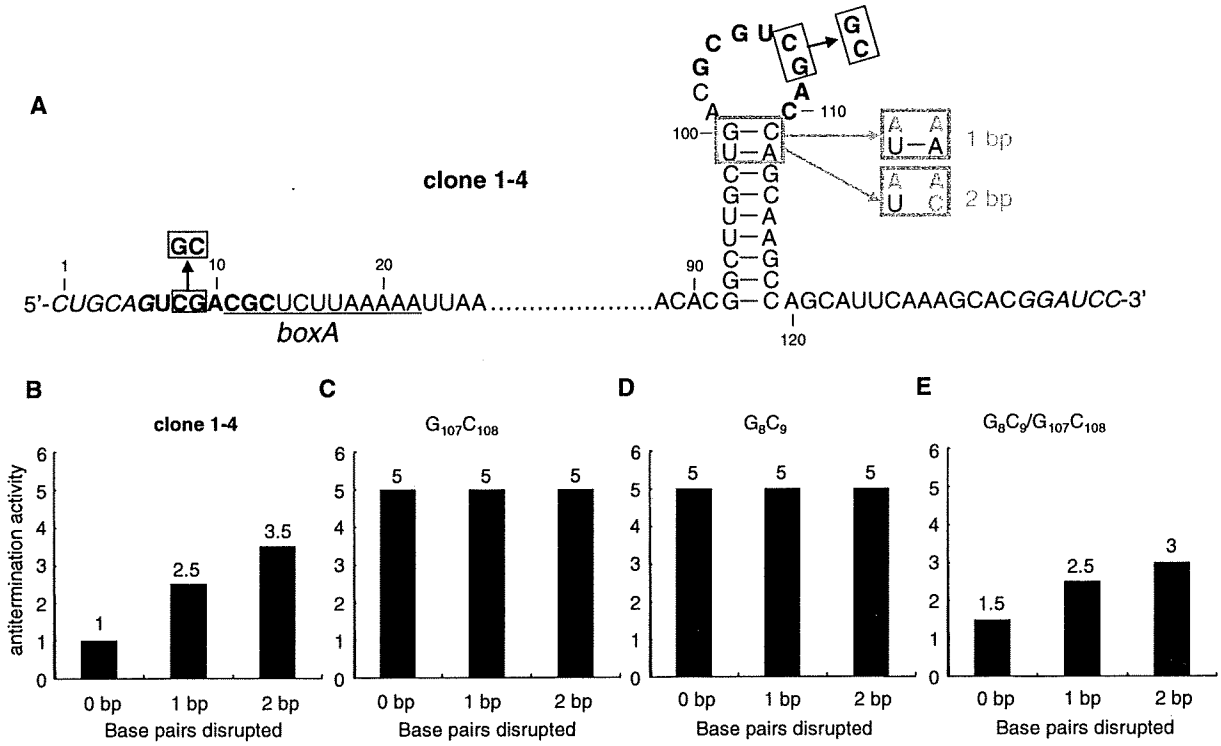

Figure 4. The effect of the disruption of 1 and 2 bp in the upper stem region of the antisense RNA stem–loop of clone **1-4** and mutant constructs. (**A**) The secondary structure of the boxA and antisense stem–loop region of clone **1-4** and base substitutions investigated. The effect of the disruption of 1 and 2 bp in the upper stem of the antisense RNA on the antitermination activity of (**B**) clone **1-4**, (**C**) clone **1-4** $G_{107}C_{108}$, (**D**) clone **1-4** G_8C_9 and (**E**) clone **1-4** $G_8C_9/G_{107}C_{108}$.

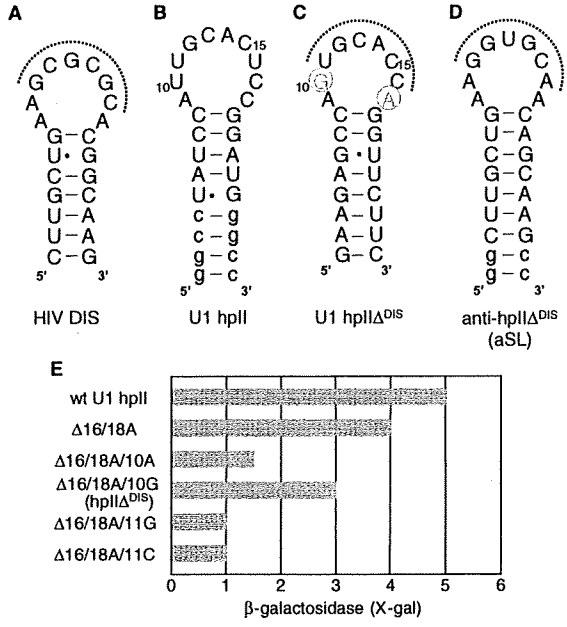

Figure 5. Design of a modified U1 hpII RNA and cognate RNA stem–loop. The secondary structure of (**A**) HIV DIS, (**B**) U1 hpII, (**C**) U1 hpIIΔ^{DIS} and (**D**) the designed antisense stem–loop (aSL). (**E**) Antitermination activities of modified U1 hpII loop sequences.

the second nucleotide in the loop, U10, was mutated to an A or G to improve base stacking interactions, and the third nucleotide, U11, was mutated to a C or G to increase the stability of complementary base pairing. However, a large decrease in antitermination activity was observed, except for the 10G mutation, which showed an activity of 3+ (Figure 5E). Therefore, the hpII triple mutant Δ16/18A/10G, designated hpIIΔ^{DIS} (Figure 5C), was chosen as a target for the design of an antisense stem–loop (Figure 5D).

Inhibition of U1A protein binding to a modified hpII loop by a designed antisense stem–loop

First, the antitermination system was used to determine whether the designed antisense stem–loop can indeed inhibit the interaction between the modified hpII (hpIIΔ^{DIS}) and the U1A protein (Figure 6). A synthetic DNA cassette encoding the antisense stem–loop was inserted into the parent reporter plasmid containing the hpIIΔ^{DIS}, so that the two stem–loops were separated by 37 nt. As a result, a considerable decrease in antitermination activity from 4+ for the parent plasmid containing the BsrGI site, which was somewhat higher than the activity of the hpIIΔ^{DIS} construct in Figure 5E (3+), to 1.5+ was observed, suggesting that antisense stem–loop was indeed inhibiting the interaction between

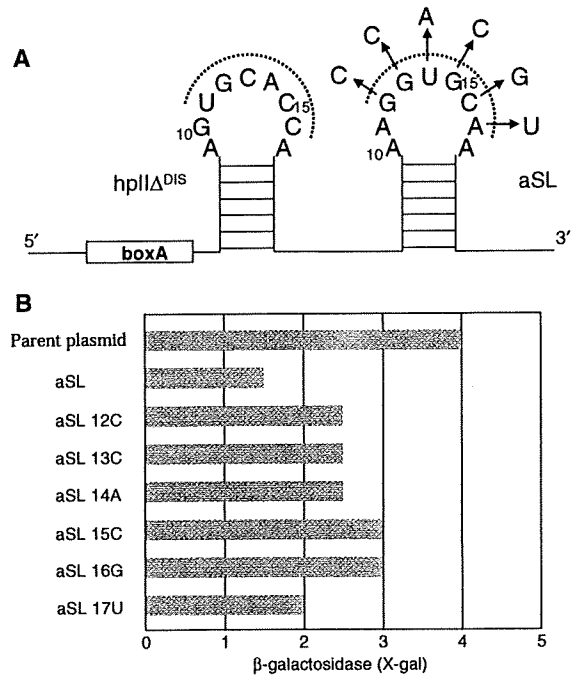

Figure 6. *In vivo* analysis of the inhibition of U1A protein binding to a modified hpII loop by the designed antisense stem-loop (aSL) and the effect of loop mutations. (**A**) The secondary structure of modified hpII-aSL region. (**B**) Antitermination activities of the designed constructs consisting of the the designed hpII–aSL, and aSL mutants.

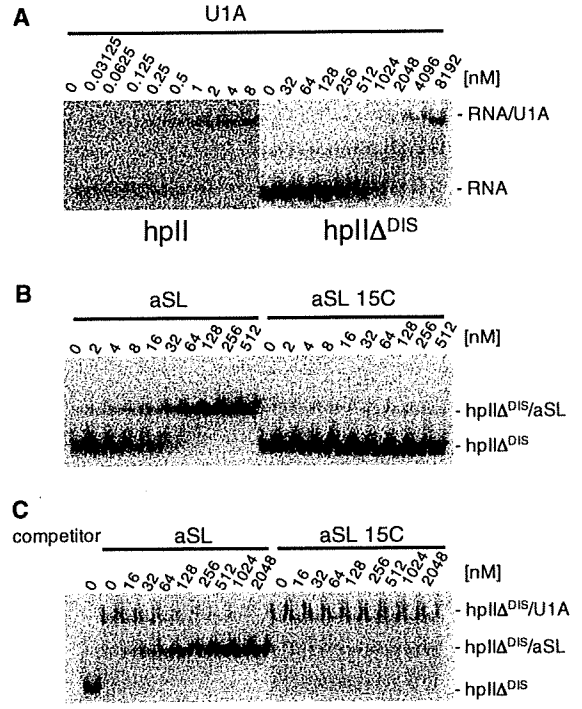

Figure 7. Gel mobility shift analysis of the inhibition of U1A protein binding to the modified hpII RNA (hpIIΔ^{DIS}) by the designed antisense RNA stem-loop (aSL). (**A**) U1A protein binding to hpII and hpIIΔ^{DIS}. (**B**) Binding of aSL and a mutant aSL (aSL 15C) to hpIIΔ^{DIS}. (**C**) Inhibition of U1A–hpIIΔ^{DIS} binding by aSL and aSL 15C.

hpIIΔ^{DIS} and the U1A protein. In order to show that the decrease in antitermination activity was due the specific interaction between the hpIIΔ^{DIS} and antisense stem–loop loop regions, the six bases within the antisense stem–loop loop (positions 12–17), which were anticipated to bind to the hpIIΔ^{DIS}, were individually substituted. As expected, an increase in antitermination activity to 2–3+ was observed, showing that a mismatch in this region leads to a decrease in the inhibitory activity of antisense stem–loop. Mutations at positions 15 and 16 lead to the largest decrease in inhibitory activity, and appeared to be the most important for the loop–loop interaction.

Next, the affinity of the modified hpII toward the U1A protein and the designed antisense stem–loop was determined by a gel shift assay (Figure 7) In the presence of 0.4 mM Mg^{2+}, the apparent K_d of the interaction of the modified hpII and the U1A protein was between 1 and 5 μM, which was more than three orders of magnitude weaker than that of the wild-type hpII with the U1A protein (0.60 nM) (Figure 7A). The K_d value for the wild-type hpII/U1A interaction was similar to that determined under similar conditions (0.35 nM) (52). On the other hand, the interaction between the modified hpII and the designed antisense stem–loop was found to be 35 nM in the presence of 0.4 mM Mg^{2+} (Figure 7B). This indicated that it should be possible to inhibit the U1 hpIIΔ^{DIS}–U1A interaction by the antisense stem–loop RNA. Indeed, when a competition experiment was carried out, antisense stem–loop was found to effectively compete with the modified hpII-U1A complex with an apparent K_i of ~160 nM, while antisense stem–loop 15C showed no inhibitory effect (Figure 7C). It should be noted that changes in the Mg^{2+} concentration showed opposite effects on the stability of the RNA loop–loop interaction and that of the RNA–protein interaction, so that stronger inhibitory effect by the antisense stem–loop would be expected to be observed at higher Mg^{2+} concentrations.

DISCUSSION

Inhibition of RNA–protein interactions by an antisense RNA stem–loop through long-range pseudoknot formation

In this study, a novel method for the identification of RNA stem–loops that disrupt the formation of RNA–protein interactions through RNA–RNA interactions is described. A bacterial reporter system for the detection of RNA–protein interactions, based on N-mediated antitermination (36), was reconstructed so that an stem–loop RNA library with 10 randomized loop nucleotides was positioned downstream of the target RNA, therefore enabling the selection of stem–loops that bind intramolecularly to the upstream target RNA and disrupt protein binding. While at least two mechanisms for the reduction of antitermination by the antisense