

# Short-Range DNA Looping by the *Xenopus* HMG-Box Transcription Factor, xUBF

David P. Bazett-Jones, Benoît Leblanc, Manfred Herfort, Tom Moss\*

*Xenopus* UBF (xUBF) interacts with DNA by way of multiple HMG-box domains. When xUBF binds to the ribosomal promoter, the carboxyl-terminal acidic tail and amino-terminal HMG-box interact. Binding also leads to negative DNA supercoiling and the formation of a disk-like structure, the enhancesome. Within the enhancesome, an xUBF dimer makes a low-density protein core around which DNA is looped into a single 180-base pair turn, probably by in-phase bending. The enhancesome structure suggests a mechanism for xUBF-dependent recruitment of the TATA box-binding protein complex without direct interaction between the two factors.

Ribosomal transcription in many eukaryotes depends on a complex array of repeated enhancers, duplicated promoters, and terminators (1–3). Most of these DNA elements bind to the HMG-box factor, UBF (4, 5), a factor required for efficient promoter activation (6–10). The ribosomal promoter contains two critically spaced sequences, the upstream control element (UCE) and Core (10–13). UBF binds within both these promoter sequences and in so doing greatly enhances recruitment of the RNA polymerase I-specific TATA-binding protein (TBP) complex (7, 9, 14, 15).

The repeated HMG-boxes of *Xenopus* UBF (xUBF) interact with the *Xenopus* Core promoter in a colinear manner (4) (Fig. 1A). DNase I accessibility of important Core promoter sequences (–11 to +15) requires DNA sequences downstream of +20 as well as sequences in the COOH-

terminal half of xUBF. Figure 1B shows that deletion of the COOH-terminal acidic tail domain (Nbox15) also suppresses Core promoter accessibility without otherwise modifying the xUBF footprint (16). Hence, the acidic tail of xUBF must fold back in order to specifically modify NH<sub>2</sub>-terminal HMG-box1 binding. This interaction between distal acidic and basic domains may be relevant to the role of the acidic tail in activation (8, 17, 18). The data also suggest that the promoter DNA may need to bend to allow the interaction to occur (Fig. 1A). Consistent with such bending, xUBF stabilizes negative DNA supercoiling in a dose-dependent manner (19) (Fig. 1C); supercoiling is apparent at near equimolar levels, for example, 30-ng track. When the plasmid was relaxed before incubation with xUBF, supercoiling was induced but required a higher xUBF concentration and was limited to a small change in linking number.

The enhancers of the *Xenopus* intergenic spacer (IGS) form a long, naturally repeated xUBF binding site (20) that is suitable for electron spectroscopic imaging (ESI). Conditions necessary for the assembly of single xUBF complexes onto the enhancer

DNA were sought and the contour lengths of naked and complexed DNA fragments were compared (21) (Fig. 2). It was estimated that a single xUBF complex shortened the DNA contour length ( $\Delta$ contour length) by about 190 base pairs (bp), confirming that xUBF induces appreciable DNA folding. The masses of the DNA and protein components of 20 individual complexes were also estimated directly from total mass and net phosphorous ESI images (22) (Fig. 3, A to F). The average phosphorous content corresponds to  $185 \pm 30$  bp, a value in close agreement with the  $\Delta$ contour length of  $173 \pm 40$  bp for this group of complexes (Fig. 3G). Fourteen of the 20 complexes analyzed gave a mean protein component of  $217 \pm 33$  kD, which indicates that a dimer of the GST-xUBF fusion protein (105 kD) was present. Of the other six complexes analyzed, two appeared to contain an xUBF tetramer (406 and 444 kD) and four contained a monomer ( $119 \pm 19$  kD) (Fig. 3G).

The net phosphorous images of the xUBF-enhancer complex indicate that the DNA component is concentrated toward its periphery (Fig. 3, A and B). In many cases, an approximately 360° loop of DNA is also evident (Fig. 3, C to G). Superimposition of the net phosphorous and total mass images (Fig. 4A) further confirms that the DNA lies at the exterior of the complex. The DNA loop diameters of about 19 nm, the average DNA content of 180 bp, and the  $\Delta$ contour length of about 170 bp together indicate that each complex contains a single turn of DNA. Hence, an xUBF dimer loops 180 bp of enhancer DNA into a single, approximately 360° turn, in a structure we call an enhancesome.

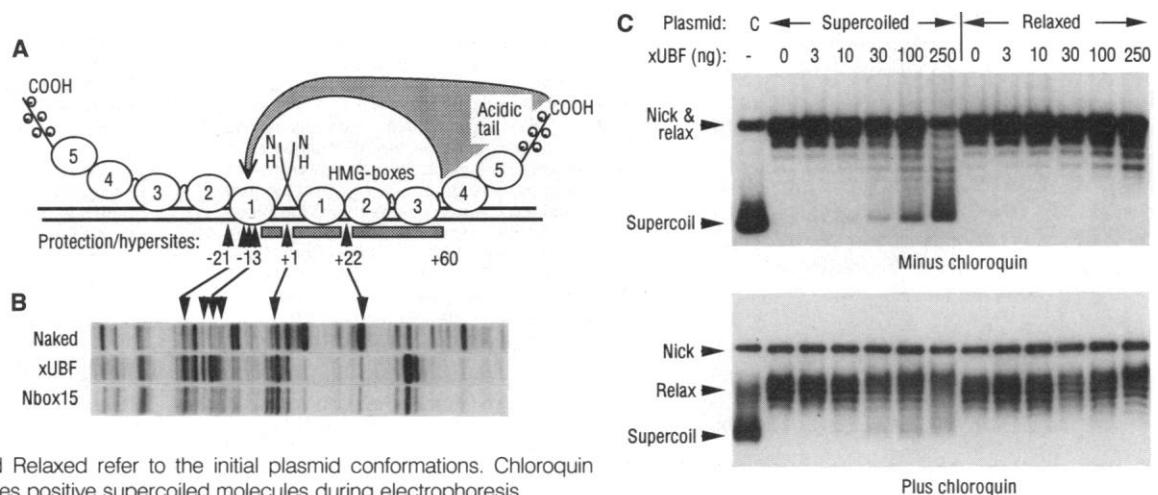
The 120 to 130 kD of enhancesome DNA lies at the periphery of the complex, yet in total mass images the enhancesome appears to be a disk of about even density. Thus, most of the 160 kD of xUBF protein component occupies the center of the DNA

D. P. Bazett-Jones and M. Herfort, Department of Medical Biochemistry and Anatomy, Faculty of Medicine, Health Sciences Center, University of Calgary, Alberta T2N 4N1, Canada.

B. Leblanc and T. Moss, Département de Biochimie et Centre de Recherche en Cancérologie de l'Université Laval, Hôtel-Dieu de Québec, 11 Côte du Palais, Québec G1R 2J6, Canada.

\*To whom correspondence should be addressed.

**Fig. 1.** (A) xUBF interacts colinearly with the promoter DNA (4). The boundaries of the observed HMG-box contacts are indicated. Interaction between the COOH-terminal domain and HMG-box 1 (4) is indicated by the large shaded arrow. (B) Deoxyribonuclease I footprinting of recombinant xUBF and Nbox15, on the Core promoter (16). Naked indicates unprotected DNA. (C) DNA supercoiling by xUBF (19). Supercoiled and Relaxed refer to the initial plasmid conformations. Chloroquin retards negative and advances positive supercoiled molecules during electrophoresis.



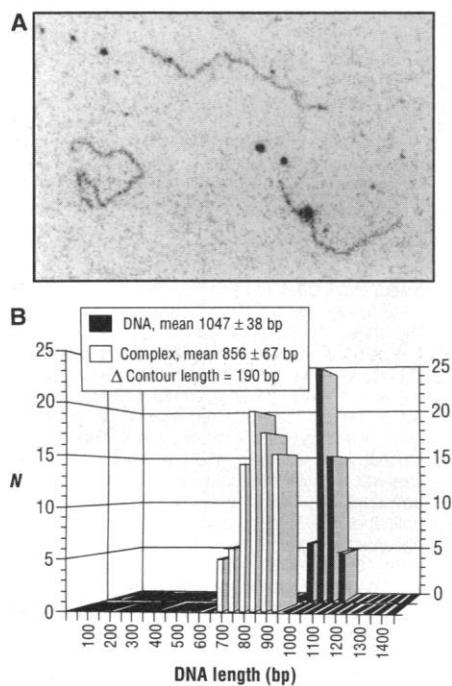
loop. The HMG-box domains represent over 45 kD of the 80-kD xUBF (23). We have shown that each HMG-box binds to the same face of the DNA (4). Together, these data suggest that the HMG-boxes lie on the inside of the enhesome DNA loop. The length of enhesome DNA ( $185 \pm 30$  bp) corresponds closely with the expected binding site for 8, if not all 10, HMG-boxes of an xUBF dimer. By combining these observations we have constructed a low-resolution model of the enhesome in which an xUBF-dimer lies inside an approximately 200-bp DNA loop, the tandemly arranged HMG-boxes each binding to  $\sim 20$  bp of DNA (Fig. 4B) (4, 8). The very low density of the enhesome protein core is apparent in the model and indicates that few protein-protein interactions could occur across the inside of the loop. The DNA loop cannot, therefore, be stabilized by a protein core, as is the case in the chromatin core particle (24). Instead, the model predicts that DNA looping by xUBF is predominantly the result of a series of in-phase bends induced by the repeated binding of the HMG-boxes. The bend angle per HMG-box can be estimated from our data to be  $\sim 36^\circ$ , if all five boxes of xUBF (23) bend DNA, or  $\sim 60^\circ$  if only the tight DNA binding boxes 1 to 3 do so. HMG-boxes are known to bind four-way junctions (25) and hence might interact

with the DNA crossover that occurs at the exit of the DNA duplex from the enhesome (Fig. 4B). Our previous mapping of the xUBF-DNA interaction (4) and that of hUBF (8) exclude HMG-boxes 1, 2, and 3 from such an interaction. We believe that the repeated arrangement of xUBF binding sites in the ribosomal DNA also excludes HMG-boxes 4 and 5.

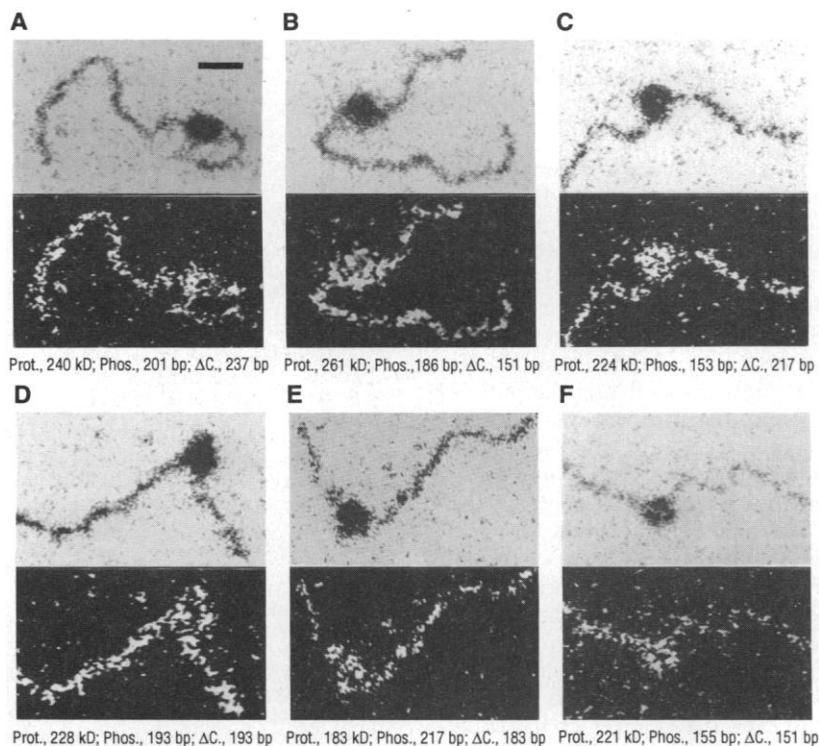
xUBF positions itself equivalently on pairs of contiguous 60- and 81-bp enhancers (20). The enhesome must therefore be able to accommodate various lengths of DNA between 120 and 162 bp. We have shown that HMG-boxes 1 to 3 interact with 60 bp of DNA and have higher affinity for DNA than boxes 4 and 5 (4). DNA contact by HMG-boxes 4 and 5 might therefore be facultative (Fig. 4C). On the ribosomal enhancers, the resulting repeated enhesome structure would resemble an unbroken DNA superhelix of about 180 bp

per turn (Fig. 4C). This suggestion and the observation that the apparently unrelated *Xenopus* and mouse enhancers are functionally interchangeable (26) indicates that, as in *Escherichia coli* (27), DNA bending may be the key to ribosomal enhancer function. DNA bending and constrained DNA loops have also been implicated in estrogen-dependent and in lymphoid-specific mRNA promotion (28, 29).

The ribosomal promoter in mammals and amphibians consists of two DNA elements, UCE and Core, which function cooperatively and bind the RNA polymerase I TBP complex. Previous work (4, 8, 30) has shown that two UBF complexes (which the present study suggests are xUBF dimers) bind within the ribosomal promoter, one centered around +1 and the other within the UCE at around  $-90$  to  $-100$  bp. Hence, we predict that two enhesomes form on the promoter.

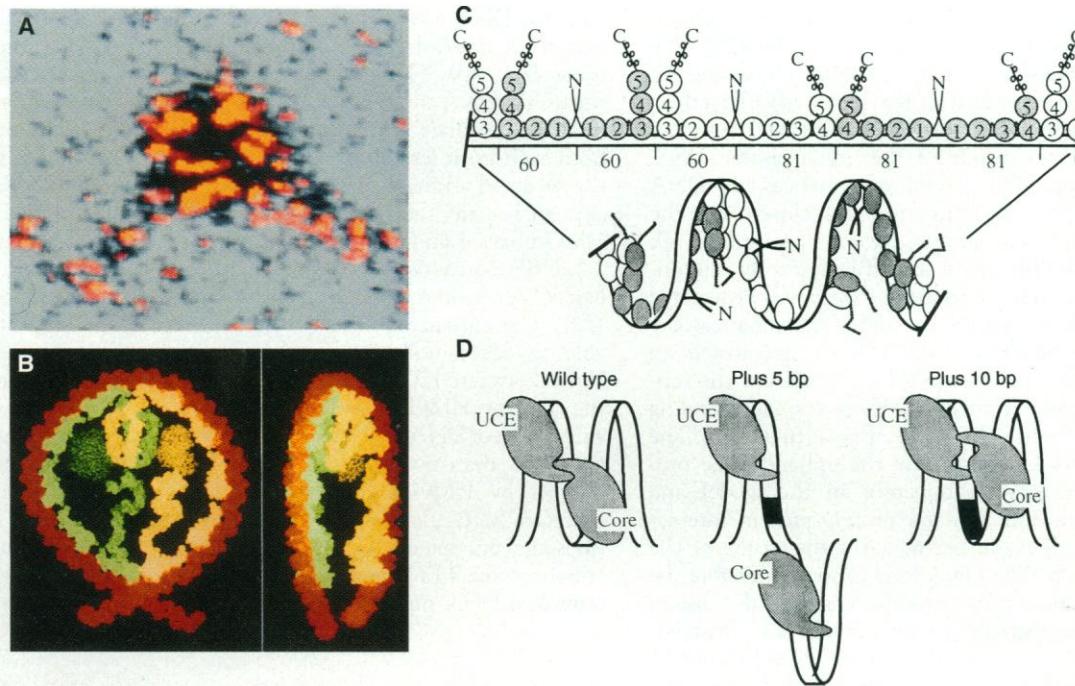


**Fig. 2.** (A) A typical ESI field showing a unique xUBF complex and two uncomplexed DNA fragments (21). (B) Histograms of DNA lengths obtained from enhancer fragments carrying a unique xUBF complex (Complex) and from uncomplexed DNA (DNA).



**Fig. 3.** (A to F) Analysis of individual xUBF-DNA images (22). The upper (positive) images show both DNA and protein and the lower (negative) images phosphorous. The protein (Prot.) and DNA contents, the latter from phosphorous (Phos.), or contour length ( $\Delta$ C.) are given below each pair of images. The scale bar in (A) represents 35 nm. (G) Histograms of the estimated protein and DNA (net phosphorous, Phosphorous or contour length,  $\Delta$ Contour) contents of individual complexes.

**Fig. 4.** (A) Example of a false color net phosphorous image superimposed on the corresponding total mass image (gray tone). (B) Low-resolution model of the enhancesome. Five HMG-box folds (39) have been arranged within a loop of about 200 bp of DNA. The NH<sub>2</sub>-terminal dimerization and COOH-terminal acidic domains of xUBF have been modeled as broken 3<sub>10</sub> helices to complete the space filling model (no structural prediction for these domains is intended). Two views of the water-accessible surface of the structure are shown, one on the face of the DNA loop (left) and the other from one edge of the loop (right). (C) Arrangement of multiple xUBF-dimers on the enhancer repeats. The colinear arrangement of xUBF dimers (upper diagram) would lead to a continuous superhelix of interpenetrating enhancesomes (lower diagram). (D) An explanation for the synergy between UCE and Core promoter elements. Two TBP-complexes (shaded) are shown binding on the surface of an extended enhancesome



superhelix generated by xUBF binding within UCE and Core promoter elements.

[Their proximity to each other and the previous footprinting data (4) suggest DNA contact by HMG-box 3 may be facultative within the promoter (Fig. 1A)]. The TBP complex extends the UBF footprint on the UCE from -115 to beyond -160 bp and also protects the Core element near the initiation site [see, for example, (30)]. In modeling the promoter, we found that the two adjacent UBF-DNA complexes would, if folded as enhancesomes, present the TBP-complex binding sites on the surface of a superhelix. This might then facilitate cooperative binding of the TBP complex to both sites (Fig. 4D), either by allowing interaction between Core and UCE-bound TBP complexes [TBP has been shown to dimerize (31)] or by allowing a single TBP complex to interact with both promoter elements. Spacing changes of half a duplex turn between the UCE and Core elements diminish promoter activity, whereas changes involving a full turn only mildly affect promoter activity (11, 12). Figure 4D shows how the corresponding predicted changes in enhancesome topology may explain these observations. A similar model might also explain the observed coupling of terminator and promoter (32-34).

## REFERENCES AND NOTES

1. T. Moss, K. Mitchelson, R. F. J. De Winter, *Oxf. Surv. Eukaryotic Genes* 2, 207 (1985).
2. B. Sollner-Webb and E. B. Mougey, *Trends Biochem. Sci.* 16, 58 (1991).
3. R. F. J. De Winter and T. Moss, *J. Mol. Biol.* 196, 813 (1987).
4. B. Leblanc, C. Read, T. Moss, *EMBO J.* 12, 513 (1993).
5. C. S. Pikaard, B. McStay, M. C. Schultz, S. P. Bell, R. H. Reeder, *Genes Dev.* 3, 1779 (1989).
6. A. Kuhn and I. Grummt, *Proc. Natl. Acad. Sci. U.S.A.* 89, 7340 (1992).
7. S. D. Smith *et al.*, *Mol. Cell. Biol.* 10, 3105 (1990).
8. H. M. Jantzen, A. M. Chow, D. S. King, R. Tjian, *Genes Dev.* 6, 1950 (1992).
9. B. McStay, C. H. Hu, C. S. Pikaard, R. H. Reeder, *EMBO J.* 10, 2297 (1991).
10. C. Read *et al.*, *J. Biol. Chem.* 267, 10961 (1992).
11. W. Q. Xie and L. I. Rothblum, *Mol. Cell. Biol.* 12, 1266 (1992).
12. J. J. Windle and B. Sollner-Webb, *ibid.* 6, 4585 (1986).
13. M. M. Haltiner, S. T. Smale, R. Tjian, *ibid.*, p. 227.
14. S. P. Bell, R. M. Learned, H.-M. Jantzen, R. Tjian, *Science* 241, 1192 (1988).
15. A. Schnapp and I. Grummt, *J. Biol. Chem.* 266, 24588 (1991).
16. The COOH-terminal deletion mutant Nbox15 corresponds to amino acids 16 to 567 of xUBF2 (23, 35). It was expressed and purified as a GST fusion protein as described (4). Footprinting was performed on the wild-type *X. laevis* ribosomal promoter (10).
17. R. Voit *et al.*, *EMBO J.* 11, 2211 (1992).
18. B. McStay, M. W. Frazier, R. H. Reeder, *Genes Dev.* 5, 1957 (1991).
19. Recombinant GST-xUBF fusion protein (4) was incubated in 10 mM Hepes (pH 7.9), 80 mM KCl, 3 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 0.5 mM dithiothreitol (DTT), with 20 ng of pT7AB5'Δ-9, which contained -9 to +92 bp of the *X. laevis* ribosomal promoter (10). After 20 min on ice, 0.7 U of topoisomerase (Promega) was added for 10 min at 37°C. The plasmid topoisomers were separated on a 1% agarose gel in 45 mM tris-borate, 1.25 mM EDTA with or without chloroquin. The DNA was visualized by hybridization after transfer to Hybond N (Amersham).
20. C. D. Putnam and C. S. Pikaard, *Mol. Cell. Biol.* 12, 4970 (1992).
21. GST-xUBF (1.2 μg) (4) was incubated in 25 μl of 50 mM Hepes (pH 7.6), 5 mM MgCl<sub>2</sub>, 80 mM KCl, 1 mM DTT with 200 ng of the *X. laevis* 1.1-kb Bam HI enhancer DNA fragment (3, 36). After 15 min at room temperature, the mixture was chromatographed on a 0.5-ml column of Sepharose CL-2B to separate DNA-bound xUBF from free protein. The column buffer contained 10 mM Hepes (pH 7.2), 5 mM MgCl<sub>2</sub>, 1% formaldehyde, and 0.5% glutaraldehyde. Five microliters of the peak DNA fraction were placed on a 1000-mesh copper electron microscope grid, which had been coated with a 3-nm carbon film and glow-discharged immediately before use (37). After 30 s, excess sample was washed from the grid with H<sub>2</sub>O and the grid was air-dried after all but a thin layer of the H<sub>2</sub>O had been removed. [For details of ESI analysis, see (37, 38)]. A control reaction was performed with the use of an equivalent volume of a mock protein preparation from bacteria expressing wild-type GST. Putative complexes were observed about 800 times less frequently than with GST-xUBF. Five such complexes were analyzed. Three contained no protein component and were formed by a DNA fold-back, and the other two showed no DNA length contraction.
22. ESI analysis of DNA-protein complexes has been previously described (37, 38). Estimation of the masses of the xUBF-DNA complexes was carried out on a reference image recorded at 120 eV in the electron energy loss spectrum. DNA was used as an internal mass standard, and the mass of the complex was estimated by comparison of the integrated optical density of the complex with the integrated optical density over a defined length of DNA. Net phosphorous images were obtained by subtraction of the 120-eV reference image from a 155-eV energy loss image recorded at the peak of the phosphorous L<sub>2,3</sub> ionization edge, after alignment and normalization. Results were compared quantitatively with a multiple parameter background correction by use of two pre-edge images recorded at 105 and 120 eV (37). The phosphorous content of the complex was estimated by comparison of the integrated phosphorous signal in the complex to that of a defined length of DNA. This value was used to calculate the DNA content in the complex and subtracted from the total mass estimate to obtain the protein content.
23. D. Bachvarov and T. Moss, *Nucleic Acids Res.* 19, 2331 (1991).

24. M. M. Struck, A. Klug, T. J. Richmond, *J. Mol. Biol.* **224**, 253 (1992).
25. D. M. J. Lilley, *Nature* **357**, 282 (1992).
26. C. S. Pikaard *et al.*, *Mol. Cell. Biol.* **10**, 4816 (1990).
27. N. Nachaliel, J. Melnick, R. Gafny, G. Glaser, *Nucleic Acids Res.* **17**, 9811 (1989).
28. C. Schild, F. X. Claret, W. Wahli, A. P. Wolffe, *EMBO J.* **12**, 423 (1993).
29. K. Giese, J. Cox, R. Grosschedl, *Cell* **69**, 185 (1992).
30. S. P. Bell, H.-M. Jantzen, R. Tjian, *Genes Dev.* **4**, 943 (1990).
31. C. Icard-Liepkalns, *Biochem. Biophys. Res. Commun.* **193**, 453 (1993).
32. B. McStay and R. H. Reeder, *Genes Dev.* **4**, 1240 (1990).
33. S. Firek, C. Read, D. R. Smith, T. Moss, *Mol. Cell Biol.* **9**, 3777 (1989).
34. T. Moss, A.-M. Larose, K. Mitchelson, B. Leblanc, *Biochem. Cell Biol.* **70**, 324 (1992).
35. A. Guimond and T. Moss, *Nucleic Acids Res.* **20**, 3361 (1992).
36. T. Moss, *Nature* **304**, 562 (1983).
37. D. P. Bazett-Jones, *Microbeam Anal.* **2**, 69 (1993).
38. ——— and M. L. Brown, *Mol. Cell. Biol.* **9**, 336 (1989).
39. C. M. Read, P. D. Cary, C. Crane-Robinson, P. C. Driscoll, D. G. Norman, *Nucleic Acids Res.* **21**, 3427 (1993).

40. We thank M. Boissinot for help with molecular modeling and C. Crane-Robinson and V. Stefanovsky for discussions. Supported by the Medical Research Council of Canada (T.M. and D.P.B.-J.) and the National Cancer Institute with funds from the Canadian Cancer Society (D.P.B.-J.). B.L. was FCAR Fellow, and T.M. is FRSQ Senior Researcher. The Centre de Recherche en Cancérologie de l'Université Laval is supported by the Fonds de la Recherche en Santé du Québec and the Fonds Pour la Formation de Chercheurs et l'Aide à la Recherche de Québec.

24 January 1994; accepted 11 April 1994

## RNAs with Dual Specificity and Dual RNAs with Similar Specificity

Gregory J. Connell and Michael Yarus\*

The biological role of RNA is delimited by its possible reactions, which can be explored by selection. A comparison of selected RNAs that bind one ligand with those that bind two related ligands suggests that a single nucleotide substitution can expand binding specificity. An RNA site with dual (joint) specificity has adenine and cytosine bases whose  $pK_a$ 's appear shifted upward, thereby mimicking an efficient general acid-base catalyst. The joint site also contains two conserved, looped arginine-coding triplets implicated in arginine site formation. Two selected joint RNAs are identical in some regions and distinct in others. The distinct regions, like some peptides, seem to function similarly without being similar in primary structure.

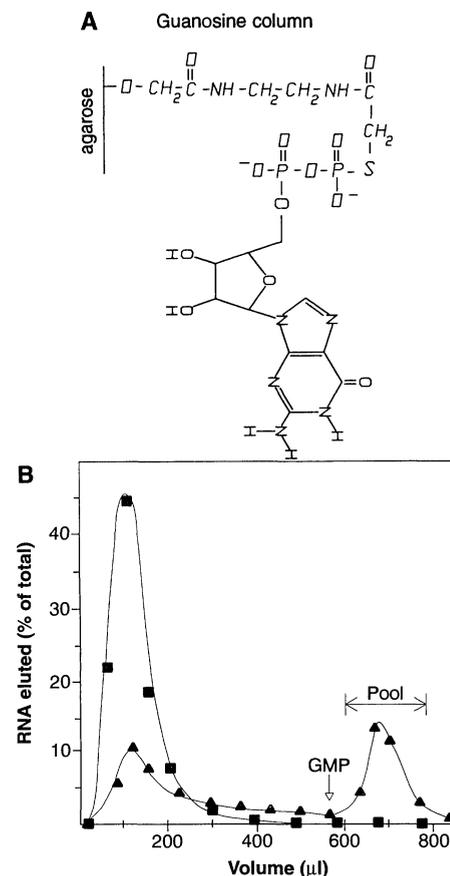
A bulged helix within the *Tetrahymena* group I intron binds guanosine during the intron excision reaction (1). It also specifically binds L-arginine, but not other normal amino acids (2). This dual affinity suggested that comparison of RNA sites binding related ligands might show how an RNA site changes specificity. Here we used selection-amplification to isolate sites that bind both guanosine and arginine and sites that bind guanosine alone. Selection-amplification has been used to obtain RNA molecules capable of binding various proteins (3), dye affinity ligands (4), free arginine (5), and adenosine triphosphate (ATP) (6).

Joint arginine-guanosine sites were isolated by requiring two related affinities from a limited number of randomized nucleotides. We first selected by arginine affinity chromatography on RNA with 25 randomized nucleotides (5, 7) and elution with L-arginine. This column was alternated with a guanosine affinity column (7) eluted with guanosine monophosphate (GMP). After the seventh selection, pooled RNA was rerun on the guanosine column. The GMP-eluted RNA was reverse transcribed, amplified, and cloned in pUC19 for sequencing.

For the selection of guanosine-binding RNAs by affinity chromatography we used the same guanosine diphosphate (GDP)-agarose and elution (Fig. 1) (8). Ligand-specific elution of RNA in both selections virtually ensures that selected RNAs bind the free ligand and its immobilized affinity derivative at the same site.

The predominant RNA sequences from guanosine and joint arginine-guanosine selections are identical, except at two locations (Fig. 2A). The joint site consensus has a 5' extension and also conserves C10 and A11. The C10 (Fig. 2B) is not the only nucleotide that occurs at this position; one to two nucleotides of any kind except U appear here in the guanosine RNA. In contrast, only A occurs at position 11 of the joint site and the guanosine-binding RNA never contains a purine here.

Binding specificity is determined by this internal-loop CA (Fig. 2A). A truncated joint motif transcript spanning positions W and Z (Fig. 2C) binds to both guanosine and arginine columns with dissociation constant ( $K_d$ ) values 1.4 times that of the full RNA. A transcript spanning positions X and Z binds to the guanosine and arginine columns with, respectively, 1.8 and 2.3 times the  $K_d$  of the full-length RNA. A truncate spanning positions Y and Z showed no measurable affinity for the arginine column and minimal affinity for the guanosine column (20 times the  $K_d$  of the full-length



**Fig. 1.** Guanosine affinity chromatography. (A) Structure of the guanosine affinity resin. (B) Elution profiles (7) for cycle 1 (■) and cycle 5 (▲) of the guanosine selection are shown.

transcript). These boundary determinations indicate that the first nucleotide of the joint consensus may contribute slightly to binding, but not significantly to specificity. The A5 of the joint consensus is present in some guanosine motif representatives (Fig. 2B) and therefore is not specific either. The remaining characteristic structure found in the joint site, the internal-loop CA, previously appeared in three independent arginine-binding RNAs (5). Taken together, these observations hint that DNA sites may be assembled from fixed smaller units.

Elution with nucleoside analogs (9, 10) suggests that the binding site in both RNAs

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309-0347, USA.

\*To whom correspondence should be addressed.