Solution Structure of a Bovine Immunodeficiency Virus Tat-TAR Peptide-RNA Complex

Joseph D. Puglisi,* Lily Chen, Scott Blanchard, Alan D. Frankel

The Tat protein of bovine immunodeficiency virus (BIV) binds to its target RNA, TAR, and activates transcription. A 14–amino acid arginine-rich peptide corresponding to the RNA-binding domain of BIV Tat binds specifically to BIV TAR, and biochemical and in vivo experiments have identified the amino acids and nucleotides required for binding. The solution structure of the RNA-peptide complex has now been determined by nuclear magnetic resonance spectroscopy. TAR forms a virtually continuous A-form helix with two unstacked bulged nucleotides. The peptide adopts a β-turn conformation and sits in the major groove of the RNA. Specific contacts are apparent between critical amino acids in the peptide and bases and phosphates in the RNA. The structure is consistent with all biochemical data and demonstrates ways in which proteins can recognize the major groove of RNA.

RNA-protein interactions are important in many cellular processes. Recent cocrystal structures of RNA-protein complexes (*1*–*6*) have provided the first detailed views of RNA-protein recognition. In these complexes, many specific contacts are made to bases, sugars, and phosphates located in or adjacent to loop regions of RNA, within which chemical information is readily accessible. Conformational flexibility of RNA can play an important role in recognition; for example, nucleotides in transfer RNA anticodon loops and in the U1 hairpin become splayed on binding (*2*, *4*, *6*). Because the major groove of RNA is narrow, recognition of helical regions appears to be restricted to the ends of helices (*2*) or to the minor groove, unless the helices are distorted. We now present the nuclear magnetic resonance (NMR) structure of an RNA-peptide complex that shows how the major groove of the BIV TAR site is recognized by the BIV Tat protein.

BIV Tat is a transcriptional activator that binds to the BIV TAR hairpin located at the 5′ end of the viral mRNAs (*7*). BIV Tat is closely related to human immunodeficiency virus (HIV) Tat, but the sequences of the arginine-rich RNA-binding domains differ (*8*). The BIV and HIV TAR hairpins show sequence similarities in the stem regions, but the structures and sequences of bulge and loop regions differ (*7*). Unlike HIV TAR, the loop sequence of BIV TAR is not essential for BIV Tat function in vivo (*7*, *9*). A peptide corresponding to the RNA-binding domain of BIV Tat binds to BIV TAR with high affinity and specificity, recognizing the bulge and upper stem regions (*7*). In vitro and in vivo experiments show that base pairs G11 · C25, G14 · C23, and C15 · G22 as well as a bulged nucleotide at position 10 of BIV TAR are essential for binding (Fig. 1A), and that amino acids Arg^{70}, Gly^{71}, Thr^{72}, Arg^{73}, Gly^{74}, Gly^{76}, Arg^{77}, and Ile^{79} in BIV Tat are critical (*10*) (Fig. 1B).

The conformations of the unbound 28-nucleotide RNA oligomer and 14–amino acid peptide were characterized qualitatively by NMR. The free RNA adopts the secondary structure shown in Fig. 1A and suggested by ribonuclease mapping (*7*). The two stem regions form A-form helices, and the G11 · C25 pair is formed. The bulged nucleotide U10 is stacked between G9 · C26 and G11 · C25, whereas U12 is flipped out in solution. The unbound peptide (Fig. 1B) is unstructured in aqueous solution, as suggested by circular dichroism spectroscopy (*10*).

A 1:1 peptide-RNA complex was formed (Fig. 1C), and the solution structure was determined by NMR with the use of ^{13}C-labeled RNA (*11*–*13*) and ^{15}N-labeled peptides to facilitate the analyses (*14*). Structural statistics for the 20 final simulated structures are listed in Table 1, and the superposition of structures is shown in Fig. 2A (*15*). Nucleotides 5 to 11, 13 to 16, and 21 to 30 and amino acids 70 to 79 are particularly well defined in the ensemble of structures [heavy-atom root-mean-square deviation, 1.49 Å (Fig. 2)].

The stem regions of BIV TAR in the complex form a continuous A-form helix. The two bulge nucleotides (U10 and U12) are not intercalated in the helix. U12 is disordered in the NMR structure, but is found on the minor groove side of the helices and does not interact with the peptide. U10 is well positioned by the NMR data in the major groove and interacts with the peptide. U12 can be deleted from BIV TAR with little effect on peptide binding, whereas deletion of U10 results in a 17-fold decrease in binding affinity (*7*). NMR experiments with the U12 deletion mutant (*16*) demonstrate that the peptide forms the same complex as that formed with wild-type BIV TAR, consistent with the nonessential nature of U12. The unstacking of U10 is the only conformational change in the RNA that occurs on peptide binding. There is a minor distortion in RNA helical structure at the junction of the G11 · C25 pair and the upper and lower helical stems, resulting in a widening of the major groove relative to an A-form helix (15.5 versus 10.5 Å) (*17*). The loop region of BIV TAR (C17 to U20) does not participate in peptide binding and is poorly defined in the ensemble of structures.

The peptide adopts a β-hairpin conformation in the complex. Residues Arg^{70}

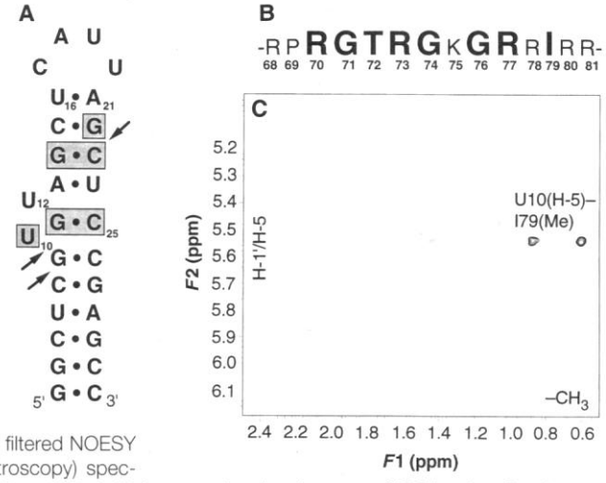

Fig. 1. (**A**) Sequence and secondary structure of BIV TAR RNA. Numbering corresponds to nucleotide positions in BIV mRNA. Nucleotides that are important for protein binding and function are boxed, and phosphates whose ethylation markedly interferes with binding are indicated by arrows (*7*). (**B**) Sequence of the BIV Tat RNA-binding peptide. Numbering corresponds to positions in the intact protein. Large letters indicate amino acids whose mutation results in a decrease in binding affinity or function (*10*). R, Arg; P, Pro; G, Gly; T, Thr; K, Lys; and I, Ile. (**C**) An *F*2-^{13}C-half filtered NOESY (nuclear Overhauser effect spectroscopy) spectrum of the BIV TAR-peptide complex at 1:1 stoichiometry, showing the strong NOE (nuclear Overhauser effect) interactions between the U10 H-5 proton and the methyl (Me) protons of Ile^{79}. Data were acquired at 25°C with a mixing time of 50 ms; ppm, parts per million.

J. D. Puglisi and S. Blanchard, Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA.
L. Chen and A. D. Frankel, Department of Biochemistry and Biophysics, and Gladstone Institute of Virology and Immunology, University of California, San Francisco, CA 94143, USA.
*To whom correspondence should be addressed.

through Ile[79] are well defined by the NMR data (Table 1) and form a β ribbon–like conformation that fits in the major groove of the RNA. The β turn is defined by Gly[74] to Arg[77] and is type I′ (*18*). Mutagenesis has highlighted the importance of Gly[76] (*10*), which assists in turn formation. The turn does not play a passive role in RNA recognition; it is buried deep in the major groove, with Gly[74] in close contact with the RNA and Lys[75] and Gly[76] exposed to solvent. The structural data agree with mutagenesis data: Mutation of Lys[75] to Ala has no effect on binding, whereas mutation of Gly[74] or Gly[76] to Ala markedly decreases binding affinity (*10*). Larger side chains at position 74 would not allow deep penetration of the peptide in the major groove.

Side chain and main chain atoms from both strands of the β hairpin make specific contacts with the RNA (Fig. 3) (*19*). One region of contact involves nucleotides in the upper helix of BIV TAR and three residues near the NH_2-terminus of the peptide. The guanidinium group of Arg[70] is within hydrogen-bonding distance of G14; this contact is well defined in the NMR structure and may involve bifurcated hydrogen bonds from Arg[70] to O-6 and N-7 of G14. Corresponding mutagenesis and modification data show that (i) Arg[70] cannot be changed to Lys without a substantial loss in affinity, (ii) mutation of the G14 • C23 pair reduces binding affinity, and (iii) methylation of N-7 of G14 interferes with peptide binding (*7*, *10*). The main chain of Gly[71] is within hydrogen-bonding distance of G22 (N-7). Mutation of Gly[71] to Ala reduces binding affinity, the C15 • G22 pair can be changed to U15 • G22 but not to other pairs, and methylation of N-7 of G22 interferes with peptide binding (*7*, *10*). There is probably a steric requirement for glycine, as addition of a side chain would prevent the close contact necessary for main chain hydrogen bonding. Thr[72] is in close contact with both G22 and C23. The methyl group of Thr[72] appears to make hydrophobic contacts with the ribose ring of G22; the major groove face of the ribose ring is the most hydrophobic. Ethylation of the phosphate between G22 and C23 interferes with peptide binding, and Thr[72] likely also contacts this phosphate because its hydroxyl group points toward it in the average NMR structure.

A second region of contact involves the bulge region of the RNA. The position of the Arg[73] side chain is well defined by the NMR data, and the guanidinium group is within hydrogen-bonding distance of N-7 of G11 and close to O-6 of G11 and the phosphate between U10 and G11. Corresponding mutagenesis and modification data show that (i) mutation of Arg[73] to Lys decreases binding affinity, (ii) changes to G11 • C25 markedly reduce binding, and (iii) ethylation of the phosphate between U10 and G11 interferes with peptide binding (*7*, *10*). Arg[77] is in close proximity to the RNA near the G9 • C26 base pair below U10, but the position of the Arg[77] side chain is not well defined by the NMR data. Mutagenesis has demonstrated the importance of Arg[77] for peptide binding, whereas changing G9 • C26 has only a small effect (*7*, *10*). These data are consistent with hydrogen bonding or electrostatic contacts between Arg[77] and a phosphate in the lower stem, perhaps between C8 and G9.

The side chain of Ile[79] makes van der Waals contacts with the aromatic ring of the bulge nucleotide U10, and this hydrophobic interaction apparently stabilizes U10 in the major groove of the RNA. Similar interactions with Ile[79] are observed in the U12 deletion mutant–peptide complex (*16*). Ile[79] may also help buttress the aliphatic portions of the surrounding Arg[73] and Arg[77] side chains. The hydrophobic nature of the interaction is emphasized by the mutagenesis data: U10 can be changed to other nucleotides but cannot be deleted without disrupting binding, and peptides in which Ile[79] is replaced with other hydrophobic amino acids retain significant activity, although isoleucine forms the highest affinity interaction (*10*).

The structure of the BIV RNA-peptide complex explains the complete set of available in vivo and in vitro mutagenesis results. The three G • C base pairs and the U10 bulge that are required for specific peptide binding are directly contacted by peptide. Phosphates whose ethylation interferes with peptide binding are also likely in contact with amino acid side chains. The

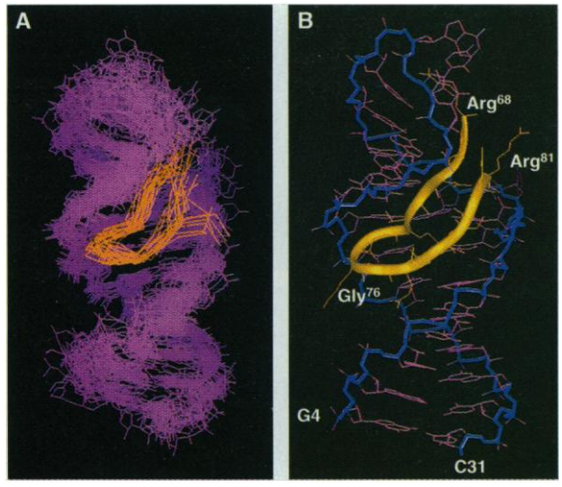

Fig. 2. (**A**) Best-fit superposition of the 20 final simulated annealing structures of the BIV TAR-peptide complex. The RNA is pink, and the peptide backbone (N, Cα, and C) is yellow. (**B**) Energy-minimized, average structure of the BIV TAR-peptide complex. All heavy atoms are shown. The RNA is pink, and the peptide is yellow. The RNA backbone is highlighted in blue.

Table 1. Structural statistics and atomic root-mean-square (rms) deviations.

	⟨SA⟩*	$(SA)_r$	⟨SA⟩ versus SA	⟨SA⟩ versus $(SA)_r$
	Structural statistics			
rms deviation from experimental distance restraints (Å)†				
All (384)	0.0391 ± 0.0037	0.0330		
RNA (301)	0.0383 ± 0.0040	0.0332		
Peptide (57)	0.0490 ± 0.0130	0.0300		
RNA-peptide (26)	0.0426 ± 0.0140	0.0360		
rms deviation from experimental dihedral restraints (degrees) (76)	0.084 ± 0.005	0.017		
Deviations from idealized geometry				
Bonds (Å)	0.0261 ± 0.0001	0.0258		
Angle (degrees)	0.0609 ± 0.0008	0.0593		
Impropers (degrees)	0.3537 ± 0.0141	0.3635		
	Atomic rms deviations			
Heavy-atom rms deviation				
All RNA + peptide			2.48	3.02
Ordered RNA + peptide‡			1.49	1.92

*⟨SA⟩ refers to the final 20 simulated annealing structures, SA to the average structure, and $(SA)_r$ to the structure obtained by taking the average coordinates of the 20 SA structures best-fitted to one another and applying restrained energy minimization. †The 20 final structures did not contain distance violations of >0.35 Å or dihedral violations of >5°. Numbers in parentheses refer to number of restraints. ‡RNA residues 5 to 11, 13 to 16, 21 to 31, and peptide residues 70 to 79.

important protein residues can be divided into two classes: those that make direct contact with the RNA (Arg^{70}, Gly^{71}, Thr^{72}, Arg^{73}, Arg^{77}, and Ile^{79}) and those that play an important structural role (Gly^{74} and Gly^{76}). Amino acids that can be changed in mutagenesis experiments with little effect on binding (Arg^{68}, Pro^{69}, Lys^{75}, Arg^{78}, Arg^{80}, and Arg^{81}) do not make RNA contacts or contribute specifically to the peptide structure.

Recognition of the BIV TAR major groove involves a widening of the groove and deep penetration by the peptide. Nevertheless, the groove is still narrow compared to the major groove of DNA, and Gly^{71} and Gly^{74} are necessary to allow peptide binding deep within the groove, apparently resulting in a tight fit between the peptide and major groove. This type of shape-selective recognition resembles that associated with interactions of proteins and drugs in the DNA minor groove (*10*). Widened RNA major grooves may generally be found near bulge or loop regions (*20*). Argininamide, an analog of arginine, binds to HIV TAR RNA in a helical region adjacent to a three-nucleotide bulge (*21*), whereas an α-helical peptide from the HIV Rev protein binds to an RNA major groove containing two purine-purine base pairs (*22*), which may induce greater widening of the groove than the single bulge nucleotide in BIV TAR (*23*).

The structure of the BIV Tat peptide–TAR complex has features seen in both RNA-protein and DNA-protein complexes. The β-sheet motif is used by the U1A protein and aspartyl–transfer RNA synthetase to contact nucleotides in hairpin loops, with amino acids on the surface of the sheet recognizing flipped-out bases (*1*, 6). In contrast, the BIV Tat β hairpin recognizes bases in the relatively unexposed major groove of BIV TAR. In DNA recognition, a β-ribbon motif is used by the MetJ and Arc repressors to bind to the major groove of the DNA helix (*24*, *25*), and a β-turn motif interacts with the DNA minor groove in a paired domain–DNA complex (*26*). Hydrophobic contacts play important roles in nucleic acid recognition. Isoleucine and leucine have been shown to intercalate between bases in DNA-protein complexes (*27*, *28*), and stacking of aromatic residues on or between bases has been observed in both DNA-protein and RNA-protein complexes (6, *29*, *30*). In the BIV Tat-TAR complex, the isoleucine interaction apparently stabilizes the position of U10 in the major groove.

BIV Tat-TAR recognition shows features common to many RNA-protein interactions, including structural distortions that open the major groove of RNA helices and the use of unpaired nucleotides for protein binding. Our data further emphasize the structural versatility of RNA and the proteins that recognize it.

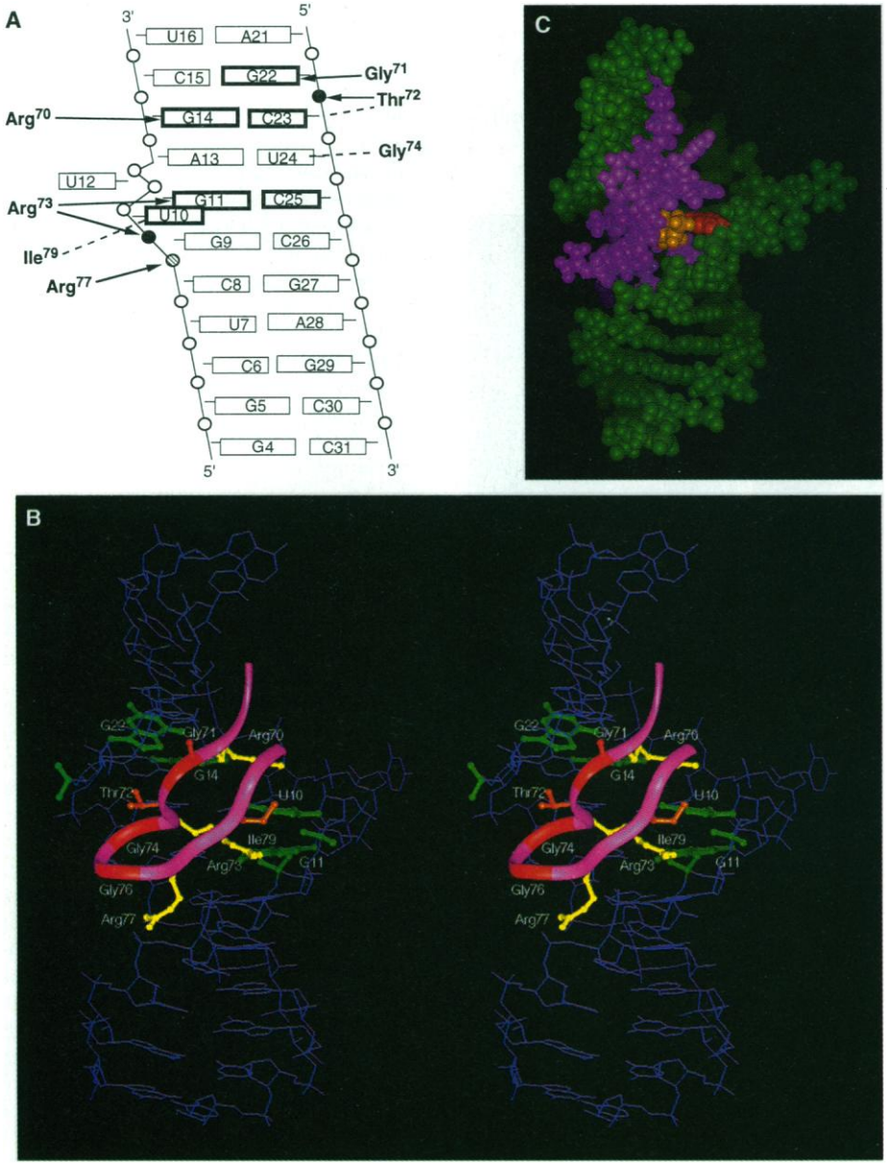

Fig. 3. (**A**) Schematic diagram of protein-RNA contacts observed in the BIV TAR-peptide complex. Hydrogen bonding and van der Waals contacts are indicated by arrows and dashed lines, respectively. Nucleotides required for specific binding are highlighted; phosphates whose ethylation strongly or weakly interferes with peptide binding are indicated by solid and hatched circles, respectively. (**B**) Stereo view of the energy-minimized, average structure of the BIV TAR-peptide complex, highlighting the agreement between biochemical and structural studies. Important nucleotides and phosphates are in green. Critical glycine residues are indicated in red, arginines in yellow, and threonine and isoleucine in orange. (**C**) Space-filling model of the BIV TAR-peptide complex; RNA is green and peptide is violet. The interaction of Ile^{79} (yellow) with the bulge nucleotide U10 (orange) is highlighted.

REFERENCES AND NOTES

1. J. Cavarelli, B. Rees, M. Ruff, J.-C. Thierry, D. Moras, *Nature* **362**, 181 (1993).
2. M. Ruff *et al.*, *Science* **252**, 1682 (1991).
3. M. A. Rould, J. J. Perona, D. Söll, T. A. Steitz, *ibid.* **246**, 1135 (1989).
4. M. A. Rould, J. J. Perona, T. A. Steitz, *Nature* **352**, 213 (1991).
5. K. Valegård, J. B. Murray, P. G. Stockley, N. J. Stonehouse, L. Liljas, *ibid.* **371**, 623 (1994).
6. C. Oubridge, N. Ito, P. R. Evans, C.-H. Teo, K. Nagai, *ibid.* **372**, 432 (1994).
7. L. Chen and A. D. Frankel, *Biochemistry* **33**, 2708 (1994).
8. L. A. Pallansch, C. S. Lackman-Smith, M. A. Gonda, *J. Virol.* **66**, 2647 (1992).
9. S. Carpenter, S. A. Nadin-Davis, Y. Wannemuehler, J. A. Roth, *ibid.* **67**, 4399 (1993).
10. L. Chen and A. D. Frankel, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5077 (1995).
11. R. T. Batey, M. Inada, E. Kujawinski, J. D. Puglisi, J. R. Williamson, *Nucleic Acids Res.* **20**, 4515 (1992).
12. E. P. Nikonowicz *et al.*, *ibid.*, p. 4507.
13. E. P. Nikonowicz and A. Pardi, *J. Mol. Biol.* **232**, 1141 (1993).
14. RNA and peptide were synthesized and purified by standard methods (*7*) [J. D. Puglisi and J. R. Wyatt, *Methods Enzymol.* **261**, 323 (1995)]. NMR spectra were obtained with a Varian Unity+ 500-MHz NMR

spectrometer at 25°C in 50 mM NaCl and 10 mM sodium phosphate (pH 6.5). Data were obtained on complexes of unlabeled RNA and unlabeled peptide, unlabeled RNA and specific [^{15}N]glycine-containing peptide, and uniformly ^{13}C-labeled RNA and unlabeled peptide at concentrations of 1 to 2 mM. Initially, NMR spectra of the unlabeled complex were partially assigned from standard homonuclear and heteronuclear experiments. Complete assignment of all RNA nonexchangeable protons, and imino and amino protons in base-paired nucleotides, was achieved with a combination of SSNOESY [S. H. Smallcombe, *J. Am. Chem. Soc.* **115**, 4776 (1993)], constant-time HSQC [J. Santoro and G. King, *J. Magn. Reson.* **97**, 202 (1992)], 3D-HCCH-TOCSY and 3D-NOESY-HMQC [G. M. Clore *et al.*, *Biochemistry* **29**, 8172 (1990)], and ^{13}C-filtered NMR [G. Otting and K. Wüthrich, *Q. Rev. Biophys.* **23**, 39 (1990)] experiments (where HSQC is heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; and 3D, three-dimensional). Peptide assignments in the complex (all amide, α, and most side chain protons) were confirmed with peptides labeled with [^{15}N]glycine at positions 74 or 76. Assignments of the bulge nucleotides U10 and U12 in BIV TAR were confirmed by studies of the U12 deletion mutant. Distance restraints were obtained from NOESY with mixing times of 50 and 100 ms, with certain intermolecular restraints derived from 3D-NOESY-HMQC with a mixing time of 200 ms. Distance restraints were classified as strong (<2.5 Å), medium (<3.5 Å), or weak (<5.0 Å); appropriate pseudoatom corrections were applied for all nonstereospecifically assigned methylene groups. Dihedral torsion angle restraints were obtained from double-quantum–filtered correlation spectroscopy (DQF-COSY) experiments. Ribose conformations were restrained to be either C-3′-endo or C-2′-endo. Ribose sugars with mixed sugar conformations were not restrained. The phosphodiester torsion angles α and ζ were constrained loosely to adopt non-trans conformations, because of the absence of upfield shifted ^{31}P resonances. Other torsion angles in the RNA were not restrained.

15. Structures were calculated with a simulated annealing protocol within the InsightII NMRArchitect package (Biosym Technologies, San Diego, CA). A randomized array of atoms corresponding to peptide and RNA was heated to 1000 K, and bonding, distance, and dihedral restraints and a repulsive quartic potential were gradually increased to full value over 30 ps of molecular dynamics. The molecules were then cooled during 10 ps to 273 K and subjected to a final energy-minimization step that included an attractive Lennard-Jones potential. No electrostatic term was included in the target function. A total of 384 distance restraints were used, including 108 intranucleotide RNA restraints, 193 internucleotide RNA restraints, 30 hydrogen-bonding restraints in RNA, 27 intraresidue peptide restraints, 30 interresidue peptide restraints, and 26 RNA-peptide restraints; only conformationally significant restraints were included. A total of 76 experimental dihedral restraints were used, comprising eight peptide and 68 RNA restraints. Additional restraints were included to maintain chirality and peptide-bond planarity. The final force constants for distance, dihedral, and peptide-bond planarity restraints were 40, 60, and 200 kcal mol^{-1}, respectively. All color figures were generated with the program InsightII (Biosym Technologies). Experimental restraints as well as coordinates for the 20 final simulated annealing structures and the average energy-minimized structure will be deposited in the Brookhaven Protein Data Base.
16. J. D. Puglisi, data not shown. The U12 deletion mutant RNA–peptide complex yielded the same intramolecular and intermolecular NOEs as the wild-type complex.
17. Major groove width is defined as the phosphate-phosphate distance between G22 and G9. A canonical A-form helix was compared to the minimized average structure of the BIV TAR-peptide complex. The distance was measured from the center of mass of the phosphates, not considering van der Waals radii.
18. T. E. Creighton, *Proteins: Structures and Molecular Properties* (Freeman, New York, 1993). The turn is defined by a hydrogen bond between the Gly^{74} carbonyl group and the Arg^{77} amide nitrogen.
19. Peptide-RNA NOE restraints used in structure calculations included G14(H-1)–Arg^{70} (δ), C23(NH$_2$)–Arg^{70} (α, γ, δ), G22(H-8)–Gly^{71}(NH, α), G22(H-8)–Thr^{72}(β, Me), G22(H-3′)–Thr^{72}(Me), G22(H-5′/5″)–Thr^{72}(Me), C23(H-6)–Thr^{72}(Me), C23(H-5)–Thr^{72}(Me), C25(H-5)–Arg^{73}(γ, δ), C25(NH$_2$)–Arg^{73}(γ, δ), C23(H-5)–Gly^{74}(α), U24(H-5)–Gly^{74}(α), U24(H-3′)–Gly^{74}(α), U10(H-5)–Ile^{79}(γ, δ, β), G11(H-1)–Arg^{73}(δ), C26(NH$_2$)–Arg^{77}(α, γ, δ), C26(NH$_2$)–Arg^{77}(α, γ, δ), and G9(H-1)–Arg^{77}(δ). All contacts discussed in the text are constrained by direct NOE restraints.
20. K. M. Weeks and D. M. Crothers, *Biochemistry* **31**, 10281 (1992); *Science* **261**, 1574 (1993).
21. J. D. Puglisi, R. Tan, B. J. Calnan, A. D. Frankel, J. R. Williamson, *Science* **257**, 76 (1992); J. Tao and A. D. Frankel, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2723 (1992).
22. R. Tan, L. Chen, J. A. Buettner, D. Hudson, A. D. Frankel, *Cell* **73**, 1031 (1993); J. L. Battiste, R. Tan, A. D. Frankel, J. R. Williamson, *Biochemistry* **33**, 2741 (1994); R. D. Peterson, D. P. Bartel, J. W. Szostak, S. J. Horvath, J. Feigon, *ibid.*, p. 5357.
23. R. Tan and A. D. Frankel, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5282 (1995).
24. W. S. Somers and S. E. V. Phillips, *Nature* **359**, 387 (1992).
25. J. N. Breg, J. H. J. van Opheusden, M. J. M. Burgering, R. Boelens, R. Kaptein, *ibid.* **346**, 586 (1990); B. E. Raumann, M. A. Rould, C. O. Pabo, R. T. Sauer, *ibid.* **367**, 754 (1994).
26. W. Xu, M. A. Rould, S. Jun, C. Desplan, C. O. Pabo, *Cell* **80**, 639 (1995).
27. C. Y. King and M. A. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11990 (1993); M. H. Werner, J. R. Huth, A. M. Gronenborn, G. M. Clore, *Cell* **81**, 705 (1995).
28. M. A. Schumacher, K. Y. Choi, H. Zalkin, R. G. Brennan, *Science* **266**, 763 (1994).
29. J. L. Kim, D. B. Nikolov, S. K. Burley, *Nature* **365**, 520 (1993).
30. Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *ibid.*, p. 512.
31. We thank E. V. Puglisi for helpful discussions and suggestions on the manuscript, R. Batey for helpful discussions, I. Tinoco Jr. for reading the manuscript, and H. Noller for encouragement. Supported by grants from the Dreyfus Foundation, the Packard Foundation (J.D.P.), and NIH (AI29135 to A.D.F.) as well as by NIH postdoctoral fellowship AI08591 (L.C.). The NMR facility is supported by a grant from the Markey Foundation to the Center for Molecular Biology of RNA at the University of California, Santa Cruz.

3 July 1995; accepted 22 September 1995

Inflammatory Bowel Disease and Adenomas in Mice Expressing a Dominant Negative N-Cadherin

Michelle L. Hermiston and Jeffrey I. Gordon*

Cadherins mediate cell adhesion and are essential for normal development. Embryonic stem cells were transfected with a dominant negative N-cadherin mutant (NCADΔ) under the control of promoters active in small intestinal epithelial cells and then introduced into C57BL/6 mouse blastocysts. Analysis of adult chimeric mice revealed that expression of NCADΔ along the entire crypt-villus axis, but not in the villus epithelium alone, produced an inflammatory bowel disease resembling Crohn's disease. NCADΔ perturbed proliferation, migration, and death programs in crypts, which lead to adenomas. This model provides insights about cadherin function in an adult organ and the factors underlying inflammatory bowel disease and intestinal neoplasia.

Precise control of cell adhesion is necessary during embryogenesis. In adult organisms, perturbations of cell adhesion are associated with tumor invasion and metastasis. Cadherins are transmembrane glycoproteins that mediate homophilic adhesive interactions between cells (*1*). Their conserved cytoplasmic domains interact with β-catenin or plakoglobin, which bind α-catenin (*2*). These interactions are essential for linkage to the actin cytoskeleton and for productive adhesion (*3*). Results from cell culture studies indicate that cadherin-catenin complexes regulate cell polarity, formation of junctional complexes, migration, and proliferation (*1*, *3*). Disruption of endogenous cadherin production results in embryonic lethality in mice and *Xenopus*, establishing the fundamental role of cadherins in development (*4*). We describe here the consequences of disrupting cadherin function in the crypt and villus epithelium of the adult mouse small intestine.

The mouse intestinal epithelium expresses a sequence of "developmental events"—proliferation, lineage allocation, migration, differentiation, and death—throughout life (*5*). Proliferation is confined to the crypts of Lieberkühn. The crypt's multipotent stem cell gives rise to enterocytes, mucus-producing goblet cells, enteroendocrine cells, and Paneth cells. Member cells of these four lineages differentiate during an orderly migration and are subsequently eliminated by apoptosis and exfoliation or phagocytosis. Renewal is rapid (3 to 20 days).

The organization of the small intestinal

Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO 63110, USA.

*To whom correspondence should be addressed.