Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex

Aurélie A. V. Albertini, *et al.*

Science **313**, 360 (2006);
DOI: 10.1126/science.1125280

points of notable rotation. We estimate that about 38.5 maneuvered N-RNA complex units are required to complete one round of the RNP structure with a radius of 245 Å.

The sequence of the vRNA in the RNP must be read by the polymerase complex during RNA synthesis. There are three possible mechanisms for copying the RNA sequence: (i) The vRNA is completely exposed in the RNP, so Watson-Crick base-pairing can occur without any change of the RNP structure. (ii) The vRNA is completely dissociated from the RNP, so it serves as a template like a naked RNA molecule. (iii) A pronounced structural change occurs in the RNP to allow the sequence of the vRNA to be read by the polymerase complex without disrupting the integrity of the RNP. The conformational arrangement of the RNA in the N protein as revealed by our structure suggests that Watson-Crick base-pairing could not occur when the N protein is closed on the vRNA. Bases in positions 5, 7, and 8 are completely shielded by the N protein such that their backbone conformation is held rigidly by the N protein, thus preventing the formation of an RNA duplex. The second possibility is also unlikely because the RNP remains intact after one round of RNA synthesis and may be used as the template again. Thus, it is likely that the vRNA is temporarily dissociated from the N protein within the active polymerase complex.

References and Notes

1. S. U. Emerson, R. R. Wagner, *J. Virol.* **10**, 297 (1972).
2. S. Naito, A. Ishihama, *J. Biol. Chem.* **251**, 4307 (1976).
3. A. S. Huang, R. R. Wagner, *J. Mol. Biol.* **22**, 381 (1966).
4. D. Thomas *et al.*, *J. Virol.* **54**, 598 (1985).
5. F. Iseni, F. Baudin, D. Blondel, R. W. Ruigrok, *RNA* **6**, 270 (2000).
6. S. U. Emerson, Y. Yu, *J. Virol.* **15**, 1348 (1975).
7. M. Howard, G. Wertz, *J. Gen. Virol.* **70**, 2683 (1989).
8. A. F. Howatson, G. F. Whitmore, *Virology* **16**, 466 (1962).
9. R. M. McCombs, M. B. Melnick, J. P. Brunschwig, *J. Bacteriol.* **91**, 803 (1966).
10. T. Nakai, A. F. Howatson, *Virology* **35**, 268 (1968).
11. G. Schoehn, F. Iseni, M. Mavrakis, D. Blondel, R. W. Ruigrok, *J. Virol.* **75**, 490 (2001).
12. J. Martin-Benito *et al.*, *EMBO Rep.* **2**, 313 (2001).
13. Z. Chen, T. J. Green, M. Luo, H. Li, *Structure* **12**, 227 (2004).
14. T. J. Green *et al.*, *J. Virol.* **74**, 9515 (2000).
15. T. J. Green, M. Luo, *Acta Crystallogr. D* **62**, 498 (2006).
16. J. D. Keene, B. J. Thornton, S. U. Emerson, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6191 (1981).
17. M. Carson, *Methods Enzymol.* **277**, 493 (1997).
18. W. L. Delano, *The PyMOL User's Manual* (Delano Scientific, San Carlos, CA, 2002).

19. We thank L. Andrew Ball for critical reading of the manuscript before submission. We thank the staff of the South East Regional Collaborative Access Team (SER-CAT) and Bio-CARS at the Advanced Photon Source, Argonne National Laboratory, for their assistance in data collection. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. SER-CAT supporting institutions may be found at www.ser-cat.org/members.html. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the U.S. Department of Energy, Office of Biological and Environmental Research, and by the NIH, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. This work is supported in part by NIH grants AI050066 (to M.L.) and R37AI012464 (to G.W.W.). The atomic coordinates and reflection file have been deposited with the Research Collaboratory for Structural Bioinformatics Protein Data Bank. The access code for the coordinates is 2GIC.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1126953/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Table S1
References

2 March 2006; accepted 5 June 2006
Published online 15 June 2006;
10.1126/science.1126953
Include this information when citing this paper.

Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex

Aurélie A. V. Albertini,[1]* Amy K. Wernimont,[2]* Tadeusz Muziol,[2] Raimond B. G. Ravelli,[2] Cedric R. Clapier,[2] Guy Schoehn,[1] Winfried Weissenhorn,[2]† Rob W. H. Ruigrok[1]

Negative-strand RNA viruses condense their genome into a helical nucleoprotein-RNA complex, the nucleocapsid, which is packed into virions and serves as a template for the RNA-dependent RNA polymerase complex. The crystal structure of a recombinant rabies virus nucleoprotein-RNA complex, organized in an undecameric ring, has been determined at 3.5 angstrom resolution. Polymerization of the nucleoprotein is achieved by domain exchange between protomers, with flexible hinges allowing nucleocapsid formation. The two core domains of the nucleoprotein clamp around the RNA at their interface and shield it from the environment. RNA sequestering by nucleoproteins is likely a common mechanism used by negative-strand RNA viruses to protect their genomes from the innate immune response directed against viral RNA in human host cells at certain stages of an infectious cycle.

Rabies virus, a member of the *Rhabdoviridae* family, is the causative agent of rabies, a fatal central nervous system disease (*1*), which constitutes a serious health problem in developing countries that lack effective vaccination programs (*2*). *Rhabdoviridae*, together with *Paramyxoviridae* (e.g., measles virus), *Filoviridae* (e.g., Ebola virus) and *Bornaviridae* (e.g., Borna disease virus), are RNA-containing enveloped viruses that use nonsegmented negative sense RNA as their genome. The RNA is condensed by the nucleoprotein (N) into a helical nucleocapsid (NC) (*3*) and this N-RNA complex constitutes the essential template for replication by the RNA-dependent RNA polymerase complex (*4*). The polymerase complex selects for either transcription or replication and is composed of the enzymatic active L protein and the phosphoprotein P (*5*–*7*). Replication produces a full-length (+) copy of the viral RNA (vRNA), which is the specific target for encapsidation by N and serves as a template for (–) RNA replication. The switch from transcription to replication is in part regulated by the abundance of free N in the cytoplasm (*8*). In order to understand the role of N-RNA complexes in the viral life cycle, we solved the crystal structure of a nucleoprotein oligomer from rabies virus complexed to a 99-nucleotide-long RNA segment.

Recombinant expression of most viral nucleoproteins from negative-strand RNA viruses leads to nonspecific host cell RNA encapsidation by N (*9*, *10*), resulting in either helical or ring-like structures depending on the length of the RNA (*10*, *11*). Rabies virus N expression in insect cells produced N-RNA rings containing 9, 10, 11, 12, or 13 copies of N as determined after purification (*12*). The undecameric ring produced the best diffracting crystals belonging to space group $P2_12_12$. The structure was solved by single anomalous dispersion (SAD) and refined to 3.5 Å resolution with an *R*-factor of 27.5 (R_{free} = 32.6) (table S1) (*13*). Two ~550-kDa large undecameric N-RNA rings pack head to head in the crystal asymmetric unit. Each ring has an outer diameter of 160 Å, an inner diameter of 60 Å, and a height of 74 Å (Fig. 1A). The N protomer consists of two main domains, which contact nine nucleotides of single-stranded RNA, as predicted (*14*), that are occluded in the center of the ring (Fig. 1, B, C, and D, and Fig. 2). The N-terminal core domain (NTD; residues 32 to 233) folds into a helical arrangement composed of 6 helices connected by large loops. The C-terminal core domain (CTD; residues 236 to 356 and 396 to 450) is composed of 11 helices joined by tighter loops (Figs. 1C and fig. S1). Two regions in N (NTD: 105 to 118 and CTD: 376 to 397) are

[1]Institut de Virologie Moléculaire et Structurale, FRE 2854 Université Joseph Fourier-CNRS, [2]European Molecular Biology Laboratory (EMBL), Boite Postale 181, 38042 Grenoble, France.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: weissen@embl-grenoble.fr

presumably flexible, because they are absent in the structure (Fig. 1C and fig. S1). In addition to the core domains, two smaller subdomains participate in domain exchange between protomers and stabilize polymerization (Figs. 1C and 3, A and B). Both the top (NTD) and bottom (CTD) domains act as "jaws" that clamp down onto the RNA strand and enclose it completely (Fig. 1D), an observation which is consistent with the fact that the RNA remains bound to N in CsCl gradients (*15*). The closest contact between the two jaws is between NTD residue N157 and CTD residue P435 (9 to 10 Å, depending on the protomer) (Fig. 1D). One likely reason for tight RNA sequestering is to prevent immune recognition. Tightly packaged RNA does not constitute a target for the innate immune system such as Toll-like receptors, which are present during transport of the nucleocapsid along the endosomal pathway into the cytoplasm to its replication site (*16*, *17*) and is protected from exonuclease activities triggered by the interferon antiviral defense system (*16*, *18*). In addition, complete genome protection could be crucial during transport of nucleocapsids to the site of virus assembly and budding.

The experimental electron density map showed continuous density representing the sugar-phosphate backbone and averaged densities for the bases indicating a 5′-3′ direction in clockwise orientation in the ring-structure (Figs. 1B and 2A). Although each N-RNA ring bound short random RNA from the expression host, differences between purines and pyrimidines could be discerned in the electron density map based on the angles and the sizes of the nucleotides, and they were modeled as either adenine or cytosine (Fig. 2A). The RNA strand is twisted clockwise in an irregular left-handed helix along the inner perimeter of the ring within the continuous cleft made up by the NTD (top) and CTD (bottom) interface in the ring structure (Fig. 1B and fig. S2). For each protomer, the RNA strand is roughly split into two halves, which wrap around the NTD jaw; this arrangement provides most of the interactions and forces the RNA to bulge out at the tip of the NTD jaw (Figs. 1D and 2A). Each nucleotide makes polar contacts either via its phosphate group (seven out of nine phosphates are recognized), as predicted (*19*), or its ribose moiety (two out of nine are recognized), which is modeled in the C3′ *endo* sugar pucker conformation (Fig. 2, A and B). The first three bases point toward the solvent and stack onto each other and onto the last two bases of the preceding N protomer (M–1). Nucleotide 5 is involved in a kink that allows the base of nucleotide 4 to stack onto the bases from nucleotides 6 and 7 pointing toward the protein moiety, while its own base points away from the protein. The last two bases then point again away from the protein (Fig. 2, A and B). Most of the basic residues involved in RNA coordination are conserved between rabies virus and vesicular stomatitis virus (genus *Vesiculovirus*), indicating a conserved RNA coordination network (fig. S1).

There are no significant interactions between the NTDs within the ring structure (Fig. 1B). Their position is mainly determined by the connection to the CTD, the bound RNA, and extensive crystal contacts between NTDs of the two rings in the crystal asymmetric unit. In contrast, the CTDs share a large interaction surface (2700 Å^2 total buried surface) (Fig. 1B) that is determined by hydrophobic and van der Waals contacts as well as multiple polar interactions (total of 15) involving many main chain contacts (total of nine contacts).

Two small subdomains emerge from NTD and CTD (Fig. 1C) and reach over to neighboring protomers, contacting them either clockwise (NTD; M+1) or counter clockwise (CTD; M–1) and so establish domain swapping (Fig. 3A). The extreme N terminus folds into a short β hairpin that nestles between the 3/10 helix η1 and helix 8 (fig. S1) of the CTD from M+1. The main contacts are made between the N11 and E266, as well as E20 and R254 in some monomers, in addition to hydrophobic interaction (V10, I22) (*22*). The β hairpin is followed by a stretch of residues (23 to 29) that are completely solvent exposed and constitute a potential hinge region between the subdomain and the core NTD, while Y28 and Y30 are used as anchor residues (Fig. 3B).

A second small subdomain emanates from the CTD as a coil region followed by helix 13 that reaches over to a neighboring protomer (M–1) in counterclockwise fashion (Fig. 3, A and B). The connection from helix 13 back to the core of the main CTD is disordered in all protomers (Fig. 1C). The CTD subdomain is much shorter and is firmly attached to the CTD by a hydrophobic core (F350, F349, F355), which is followed by a shorter potential hinge region (residues 351 to 356). The following loop region leading to helix 13 interacts with M–1 (R361 to E403; R357 to E403) and is in close contact with β strand 1 from M+1 (Fig. 3B).

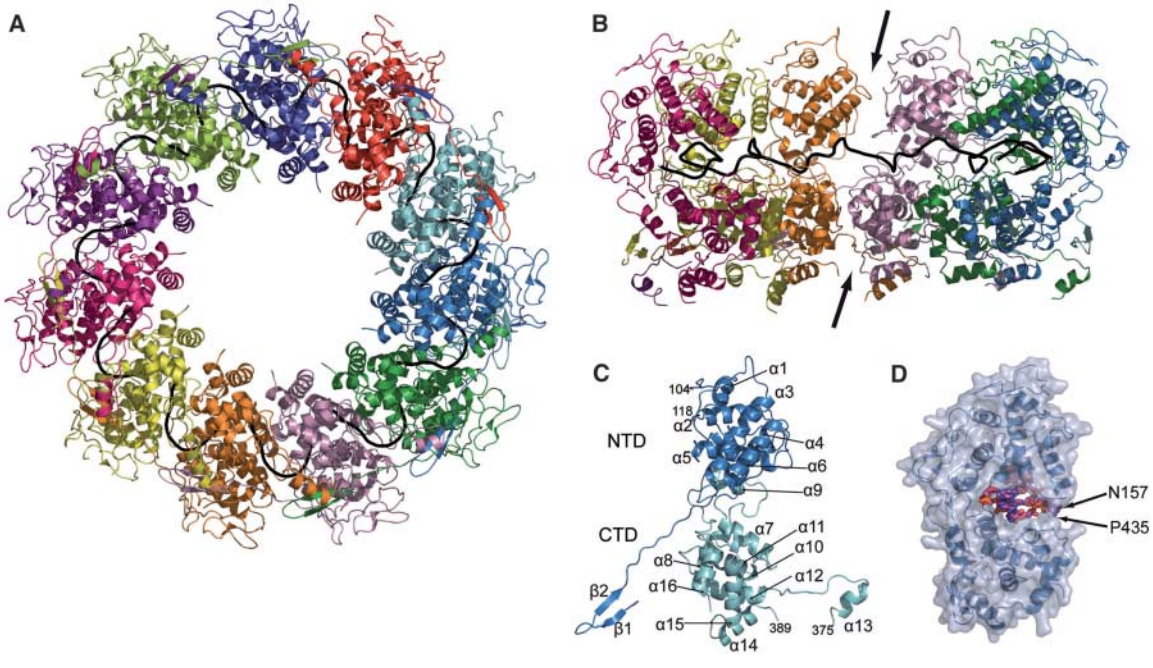

Fig. 1. Overall structure of the N-RNA complex. (**A**) Ribbon diagram of the 11-nucleotide oligomer N-RNA ring structure as viewed from the bottom. Each N protomer is colored differently. The RNA is shown as a black coil. (**B**) View of the inside of the ring structure; only six protomers are depicted in different colors with the NTD on the top and the CTD on the bottom. The path of the RNA is shown as a coil in a clockwise 5′ to 3′ orientation. The top arrow indicates that the NTDs do not interact with each other; the bottom arrow points toward the extensive interface between CTDs. (**C**) Ribbon diagram of the N protomer; the NTD is shown in dark blue and the CTD in light blue. The helical secondary structure elements are numbered consecutively. (**D**) Space-filling model of the N protomer reveals that the RNA is completely clamped at the interface of the NTD and the CTD and thus is not accessible as a template for the polymerase.

The structural overlay of the Calpha atoms of the 11 protomers reveals that the two swapped domains and an NTD region present the highest level of flexibility (fig. S3); this supports the notion that the two subdomains could act as potential hinges and lead to lateral opening of the ring. This flexible domain linkage thus permits not only the formation of differently sized ring-like structures but also the assembly of the nucleocapsid that adopts a diameter of ~75 nm, containing ~53 protomers per helical turn (*12*).

The tight sequestering of the RNA suggests that the rabies virus genomic RNA has to dissociate in order to become a template for the polymerase. However, simultaneously, the RNA must remain close to N, because the polymerase stays attached to N during its activity (*4*, *5*). Although each protomer spans a distance of 22 Å between nucleotide ends (fig. S2), the nine nucleotides could be stretched out to a length of ~49 Å. Thus, a local dissociation of the RNA from one or several N protomers could provide sufficient space for binding of the polymerase complex (*5*) (fig. S4).

Phosphoprotein P links the polymerase complex to N and binding requires that S389 be phosphorylated within a highly mobile region on the CTD of N (Fig. 1C and fig. S1) (*21*). This flexible region is disordered in the structure (Fig. 1C), but may become ordered upon P binding (*22*), as has been observed for the measles virus N-P interaction (*23*). P binding in the region of S389 (Fig. 3, A and B) may thus affect both subdomains including the hinge regions and could transfer a signal to the NTD, which would lead to the vertical opening of the NTD-CTD clamp, facilitating RNA exposure. The absence of any protein-protein contacts between NTDs (Fig. 1B) would facilitate a large NTD motion required for this process. In contrast, the CTD conformation is locked in by the large CTD-CTD interaction surface and by the tethering of the hinge loops. The latter interactions will thus maintain the contacts between the N protomers during the passage of the polymerase and the displacement of the RNA. Maintaining contact between N and RNA is required for rebinding of RNA and thus prevents the formation of double-stranded RNA during transcription and replication, which would otherwise require a helicase activity that is not encoded in the viral genome. Furthermore, an extensive production of double-stranded RNA would induce toll-like receptor 3 (TLR3) recognition and trigger immune reactions (*16*).

Even though N spontaneously associates with RNA in vivo in the absence of other viral proteins, a number of studies have shown that the 5′ region of the viral genomic sequence regulates RNA packaging into a functional nucleocapsid (*11*, *24*). Because the structure does not reveal any sequence specificity in RNA recognition, an alternative RNA-binding mechanism must exist. In the first protomer of the assembled nucleocapsid, the side extending the NTD hinge loop will be exposed and will not participate in polymerization (Fig. 1C); consequently, this region may provide a surface for sequence-specific RNA recognition. Conversely, at the 3′ end of the packaged genome, the CTD hinge loop (Fig. 1C) may adopt a different conformation that may facilitate access of the polymerase to the 3′ end,

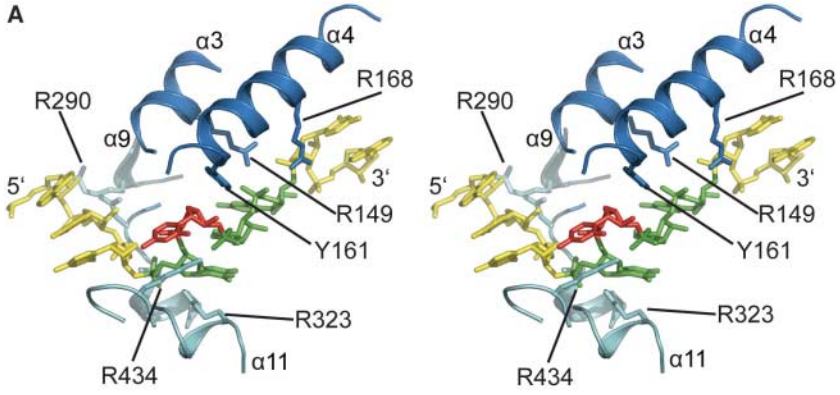

Fig. 2. Nucleoprotein-RNA interaction. (**A**) Stereo view of a close up of the interaction between N and the RNA. The RNA is shown as stick model and the protein moiety as a coil; the NTD is depicted in dark blue and the CTD in light blue. The nucleotides with bases facing the solvent are drawn in yellow and the bases facing the protein in green. The central nucleotide at the tip of the left-hand helical RNA segment is shown in red. Residues contacting the phosphates and ribose moieties are shown as sticks. (**B**) Schematic drawing highlighting the protein-RNA interactions with the phosphates and the ribose moieties. Color coding and orientations of the bases are the same as in (A). Most interactions are with the ribose-phosphate backbone.

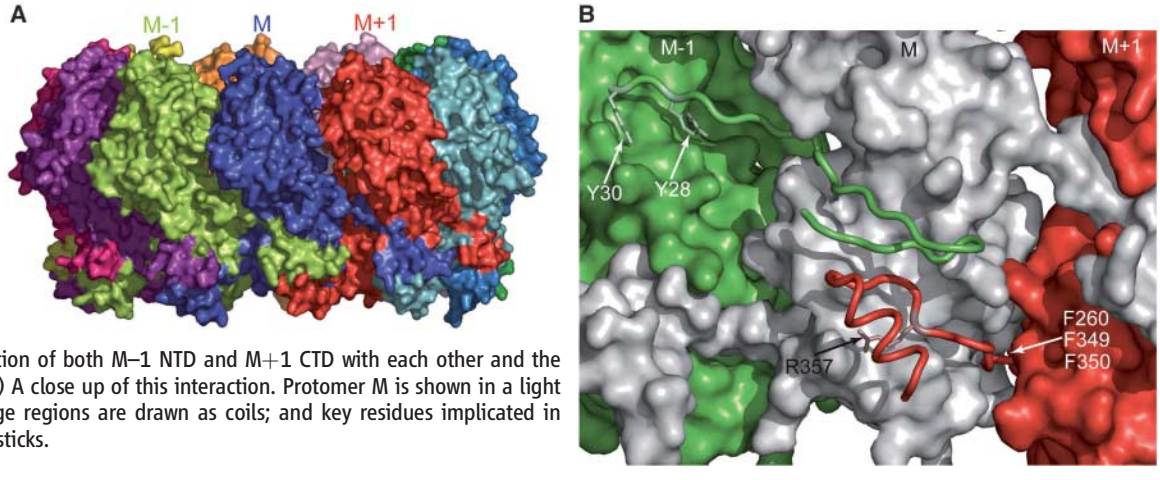

Fig. 3. Two hinge regions stabilize polymerization of N. (**A**) Space-filling model of the N-RNA ring as viewed from the side. Each protomer is shown in a different color. This indicates that the NTD hinge from protomer M reaches over to M+1 and its CTD hinge reaches over to M–1. This arrangement leads to the interaction of both M–1 NTD and M+1 CTD with each other and the surface of protomer M. (**B**) A close up of this interaction. Protomer M is shown in a light gray for clarity; both hinge regions are drawn as coils; and key residues implicated in interaction are shown as sticks.

which is the start site for replication and transcription (*25*).

Electron micrographs of nucleocapsids from measles virus (*Paramyxoviridae*) and Marburg virus (*Filoviridae*), as well as an RNA-free crystal structure from the Borna virus (*Bornaviridae*) nucleoprotein, suggest that these nucleoproteins also adopt a two-domain structure (*26*, *27*). This suggests that these enveloped viruses use an RNA-sequestering mechanism similar to that observed for the rabies virus N-RNA complex. The N-RNA polymer has thus evolved as the ideal template for the polymerase activity, which exposes the genomic RNA only temporarily to the host cell defense systems during replication. The tight sequestering of RNA observed in the crystal structure suggests further that the closed N-RNA conformation might be stabilized or frozen by small molecules, which could thus act as antiviral agents preventing rabies virus replication.

Note added in proof: This version of the manuscript is slightly changed relative to the version that was published *Science* Express on June 15. Improved refinement of the structure led to a better definition of the following regions, which have been changed accordingly: the N terminus, including the β hairpin leading to helix 1; Tyr^{161}, Arg^{168}, and Arg^{434} contacting the RNA; as well as helices 13 to 16 of the C-terminal domain. None of the changes made influences the overall conformation of the structure or any of the conclusions drawn from the structure.

References and Notes

1. B. Dietzschold, M. Schnell, H. Koprowski, *Curr. Top. Microbiol. Immunol.* **292**, 45 (2005).
2. D. L. Knobel *et al.*, *Bull. World Health Organ.* **83**, 360 (2005).
3. R. A. Lamb, D. Kolakofsky, in *Fields Virology*, D. M. Knipe and P. M. Howley, Eds. (Lippincott, Williams and Wilkins, Philadelphia, ed. 4, 2001), pp. 1305–1340.
4. H. Arnheiter, N. L. Davis, G. Wertz, M. Schubert, R. A. Lazzarini, *Cell* **41**, 259 (1985).
5. D. Kolakofsky, P. Le Mercier, F. Iseni, D. Garcin, *Virology* **318**, 463 (2004).
6. O. Poch, I. Sauvaget, M. Delarue, N. Tordo, *EMBO J.* **8**, 3867 (1989).
7. S. U. Emerson, M. Schubert, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5655 (1987).
8. D. Kolakofsky, *J. Virol.* **41**, 566 (1982).
9. D. Spehner, R. Drillien, P. M. Howley, *Virology* **232**, 260 (1997).
10. F. Iseni, A. Barge, F. Baudin, D. Blondel, R. W. Ruigrok, *J. Gen. Virol.* **79**, 2909 (1998).
11. B. M. Blumberg, C. Giorgi, D. Kolakofsky, *Cell* **32**, 559 (1983).
12. G. Schoehn, F. Iseni, M. Mavrakis, D. Blondel, R. W. Ruigrok, *J. Virol.* **75**, 490 (2001).
13. Supporting data are available on *Science* Online.
14. D. Thomas *et al.*, *J. Virol.* **54**, 598 (1985).
15. S. Lynch, D. Kolakofsky, *J. Virol.* **28**, 584 (1978).
16. S. Akira, S. Uematsu, O. Takeuchi, *Cell* **124**, 783 (2006).
17. I. Le Blanc *et al.*, *Nat. Cell Biol.* **7**, 653 (2005).
18. L. Espert *et al.*, *J. Biol. Chem.* **278**, 16151 (2003).
19. F. Iseni, F. Baudin, D. Blondel, R. W. Ruigrok, *RNA* **6**, 270 (2000).
20. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
21. B. Dietzschold *et al.*, *Virus Res.* **8**, 103 (1987).
22. H. Toriumi, A. Kawai, *Microbiol. Immunol.* **49**, 757 (2005).
23. S. Longhi *et al.*, *J. Biol. Chem.* **278**, 18638 (2003).
24. A. Kouznetzoff, M. Buckle, N. Tordo, *J. Gen. Virol.* **79**, 1005 (1998).
25. S. U. Emerson, *Cell* **31**, 635 (1982).
26. G. Schoehn *et al.*, *J. Mol. Biol.* **339**, 301 (2004).
27. M. G. Rudolph *et al.*, *Structure* **11**, 1219 (2003).

28. We thank D. Kolakofsky (Geneva), Y. Gaudin, and D. Blondel (Gif-sur-Yvette), for many discussions and ideas; F. Iseni and M. Mavrakis for early contributions to the project; S. Cusack for discussions and comments on the text; and all members of the EMBL European Synchrotron Radiation Facility (ESRF) Joint Structural Biology Group (JSBG) for access to the ESRF beam lines. This work was supported by the EMBL (W.W.), the Deutsche Forschungsgemeinschaft SFB 593 (W.W.), the Université Joseph Fourier (R.W.H.R.), the CNRS (G.S.), an EMBO fellowship (C.R.C.), and a Ph.D. fellowship from the French Ministry for Education, Research, and Technology (A.A.V.A.). Coordinates have been deposited in the Protein Data Bank (accession code no. 2GTT).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1125280/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References

23 January 2006; accepted 17 May 2006
Published online 15 June 2006;
10.1126/science.1125280
Include this information when citing this paper.

Characterization of the piRNA Complex from Rat Testes

Nelson C. Lau,[1]* Anita G. Seto,[1]* Jinkuk Kim,[2,3] Satomi Kuramochi-Miyagawa,[4] Toru Nakano,[4] David P. Bartel,[3,5] Robert E. Kingston[1]†

Small noncoding RNAs regulate processes essential for cell growth and development, including mRNA degradation, translational repression, and transcriptional gene silencing (TGS). During a search for candidate mammalian factors for TGS, we purified a complex that contains small RNAs and Riwi, the rat homolog to human Piwi. The RNAs, frequently 29 to 30 nucleotides in length, are called Piwi-interacting RNAs (piRNAs), 94% of which map to 100 defined (≤101 kb) genomic regions. Within these regions, the piRNAs generally distribute across only one genomic strand or distribute on two strands but in a divergent, nonoverlapping manner. Preparations of piRNA complex (piRC) contain rRecQ1, which is homologous to *qde-3* from *Neurospora*, a gene implicated in silencing pathways. Piwi has been genetically linked to TGS in flies, and slicer activity cofractionates with the purified complex. These results are consistent with a gene-silencing role for piRC in mammals.

Gene-silencing pathways guided by small RNAs, essential for maintaining proper cell growth and differentiation, operate at either the transcriptional or posttranscriptional level (*1*). Posttranscriptional gene silencing acts through mRNA destabilization or inhibition of mRNA translation (*1*), whereas TGS represses gene expression by altering chromatin conformation (*2*). Each pathway uses a core complex containing small RNA associated with a member of the Argonaute (Ago) protein family; however, the different mechanistic needs of each pathway require differences in complex composition. Although RNA-mediated TGS has been studied in fission yeast and other eukaryotes (*2*–*5*), the mechanism of this process in mammals remains elusive.

To identify candidate complexes for TGS in mammals, we exploited the previous observations that TGS might use small RNAs longer than the 21- to 23-nucleotide (nt) microRNAs (miRNAs). In *Arabidopsis*, *Tetrahymena*, *Drosophila*, and zebrafish, RNAs that are 24 nt and longer have been associated with TGS and/or genomic repeats, which are often silenced (*6*–*11*). In *Drosophila*, these repeat-associated small interfering RNAs (rasiRNAs) are enriched in the testis (*6*, *12*). Therefore, we prepared extract from rat testes and fractionated it on an ion-exchange Q column, monitoring the small RNAs. A peak of small RNAs longer than a 22-nt marker eluted in mild salt conditions, which suggested the presence of a novel ribonucleoprotein complex (Fig. 1A).

To characterize the small RNAs, we sequenced cDNA libraries made from flowthrough and eluate fractions, obtaining 61,581 reads from the eluate that matched perfectly to the *Rattus norvegicus* genome (*13*). In contrast to the flowthrough RNAs, which were mostly miRNAs (69%), the eluate RNAs derived primarily from regions of the genome not previously thought to be expressed (Fig. 1A). Some eluate reads matched expressed sequence tags (EST) (11%),

[1]Department of Molecular Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA. [2]Harvard-MIT Division of Health Sciences and Technology, E18-435, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. [3]Howard Hughes Medical Institute and Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [4]Department of Molecular Cell Biology, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamada-oka, Suita-shi, Osaka 565-0871, Japan. [5]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: kingston@molbio.mgh.harvard.edu