Binding of the Human Prp31 Nop Domain to a Composite RNA-Protein Platform in U4 snRNP

Sunbin Liu, *et al.*

Science **316**, 115 (2007);
DOI: 10.1126/science.1137924

The ubiquitous occurrence of the *IGF1* B haplotype in a diverse panel of small breeds clearly does not support unorthodox explanations of phenotypic diversity in the dog such as elevated mutation or recombination rates. Rather, we show that a single *IGF1* allele is a major determinant of small size in dogs and that intense artificial selection has left a signature in the proximity of *IGF1* that can readily be found by genomic scans of breeds sharing a common phenotype. The ability to identify a gene contributing to morphology without doing a genetic cross, but instead by using centuries of dog breeding, highlights the contribution that the study of canine genetics can make to an understanding of mammalian morphogenesis. These results provide a precedent for future studies aimed at identifying the genetic basis for complex traits such as behavior and skeletal morphology in dogs and other species with small populations that have experienced strong artificial or natural selection.

References and Notes

1. R. K. Wayne, *Evolution* **40**, 243 (1986).
2. R. K. Wayne, *J. Morphol.* **187**, 301 (1986).
3. J. W. Fondon 3rd, H. R. Garner, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 18058 (2004).
4. C. Webber, C. P. Ponting, *Genome Res.* **15**, 1787 (2005).
5. W. Wang, E. F. Kirkness, *Genome Res.* **15**, 1798 (2005).
6. R. K. Wayne, *J. Zool.* **210**, 381 (1986).
7. P. Saetre *et al.*, *Brain Res. Mol. Brain Res.* **126**, 198 (2004).
8. P. Savolainen, Y. P. Zhang, J. Luo, J. Lundeberg, T. Leitner, *Science* **298**, 1610 (2002).
9. S. J. Olsen, *Origins of the Domestic Dog* (Univ. of Arizona Press, Tucson, AZ, 1985).
10. C. Vila *et al.*, *Science* **276**, 1687 (1997).
11. J. Sampson, M. M. Binns, in *The Dog and Its Genome*, E. A. Ostrander, K. Lindblad-Toh, U. Giger, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2006), vol. 44, pp. 19–30.
12. B. Van Valkenburgh, X. Wang, J. Damuth, *Science* **306**, 101 (2004).
13. K. Chase *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9930 (2002).
14. K. Chase, D. R. Carrier, F. R. Adler, E. A. Ostrander, K. G. Lark, *Genome Res.* **15**, 1820 (2005).
15. J. Baker, J. P. Liu, E. J. Robertson, A. Efstratiadis, *Cell* **75**, 73 (1993).
16. K. A. Woods, C. Camacho-Hubner, D. Barter, A. J. Clark, M. O. Savage, *Acta Paediatr. Suppl.* **423**, 39 (1997).
17. K. A. Woods, C. Camacho-Hubner, M. O. Savage, A. J. Clark, *N. Engl. J. Med.* **335**, 1363 (1996).
18. N. B. Sutter *et al.*, *Genome Res.* **14**, 2388 (2004).
19. K. Lindblad-Toh *et al.*, *Nature* **438**, 803 (2005).
20. J. P. Pollinger *et al.*, *Genome Res.* **15**, 1809 (2005).
21. R. Kooijman, *Cytokine Growth Factor Rev.* **17**, 305 (2006).
22. P. Cohen, *Horm. Res.* **65**, 3 (2006).
23. R. P. Favier, J. A. Mol, H. S. Kooistra, A. Rijnberk *J. Endocrinol.* **170**, 479 (2001).
24. J. E. Eigenmann, D. F. Patterson, E. R. Froesch, *Acta Endocrinol. (Copenh.)* **106**, 448 (1984).
25. H. G. Parker *et al.*, *Science* **304**, 1160 (2004).
26. H. Epstein, *The Origin of the Domestic Animals of Africa* (Africana Publishing, New York, 1971).
27. M. V. Sablin, G. A. Khlopachev, *Curr. Anthropol.* **43**, 795 (2002).
28. E. Tchernov, L. K. Horwitz, *J. Anthropol. Archaeol.* **10**, 54 (1991).

29. We thank the hundreds of dog owners who contributed samples; the AKC Canine Health Foundation; S. Hoogstraten-Miller and I. Ginty for assistance at dog shows; P. Cruz for assistance with automated PCR primer designs; S. Kim for analytical assistance, and R. Pelker for assistance with blood serum assays of IGF1. Funded by the National Human Genome Research Institute (E.A.O., N.B.S., E.K., S.D., P.Q., H.G.P., and D.S.M.), the NSF (R.K.W.), NIH grant no. 5 T32 HG002536 (M.M.G.), NSF grant 0516310 (C.D.B. and L.Z.), NSF grant DBI 0606461 (B.P.), NIH grant P50 HG002790 (K.Z. and M.N.), and the National Institute of General Medical Sciences 063056, the Judith Chiara Charitable Trust, and the Nestle Purina Company (K.G.L.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/316/5821/112/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S6
References

1 November 2006; accepted 8 March 2007
10.1126/science.1137045

Binding of the Human Prp31 Nop Domain to a Composite RNA-Protein Platform in U4 snRNP

Sunbin Liu,[1]* Ping Li,[1,2]* Olexandr Dybkov,[1] Stephanie Nottrott,[1] Klaus Hartmuth,[1] Reinhard Lührmann,[1]† Teresa Carlomagno,[2]† Markus C. Wahl[3]†

Although highly homologous, the spliceosomal hPrp31 and the nucleolar Nop56 and Nop58 (Nop56/58) proteins recognize different ribonucleoprotein (RNP) particles. hPrp31 interacts with complexes containing the 15.5K protein and U4 or U4atac small nuclear RNA (snRNA), whereas Nop56/58 associate with 15.5K–box C/D small nucleolar RNA complexes. We present structural and biochemical analyses of hPrp31-15.5K-U4 snRNA complexes that show how the conserved Nop domain in hPrp31 maintains high RNP binding selectivity despite relaxed RNA sequence requirements. The Nop domain is a genuine RNP binding module, exhibiting RNA and protein binding surfaces. Yeast two-hybrid analyses suggest a link between retinitis pigmentosa and an aberrant hPrp31-hPrp6 interaction that blocks U4/U6-U5 tri-snRNP formation.

Most eukaryotic pre-mRNAs contain introns that are removed before translation by a multi-megadalton ribonucleoprotein (RNP) enzyme, the spliceosome (*1*–*3*). A spliceosome is assembled anew on each intron from small nuclear (sn) RNPs and non-snRNP splice factors (*4*, *5*). The RNP network of the spliceosome is extensively restructured during its maturation (*2*, *6*, *7*), reflected by changing RNA interactions. The U6 snRNA is delivered to the pre-mRNA in a repressed state, in which catalytically important regions are base-paired to the U4 snRNA (*8*, *9*). During spliceosome activation, the U4-U6 interaction is disrupted, U4 snRNA is released, and U6 snRNA forms short duplexes with U2 snRNA and the pre-mRNA substrate (*6*). Understanding this catalytic activation of the spliceosome requires detailed structural information on the snRNPs.

As for other complex RNPs (*10*), the U4/U6 di-snRNP is built in a hierarchical manner. A U4 5′ stem loop (U4 5′-SL) between two base-paired stems of U4/U6 serves as a binding site for the highly conserved U4/U6-15.5K protein (*11*). 15.5K binds to and stabilizes a kink turn (K turn) in the U4 5′-SL (*12*) and is required for subsequent recruitment of the human (h) Prp31 protein to the U4/U6 di-snRNP (*13*). hPrp31 does not interact with either the 15.5K or the RNA alone (*13*, *14*), but it is not known whether 15.5K merely prestructures the RNA for subsequent binding of hPrp31 or whether 15.5K provides part of the hPrp31 binding site. hPrp31 is essential for pre-mRNA splicing (*15*) and is a component of both major and minor spliceosomes. In the latter, the U4 snRNA is replaced by the U4atac snRNA (Fig. 1A). Nevertheless, both snRNAs bind 15.5K, and both primary RNPs incorporate hPrp31 in a strictly hierarchical manner (*13*, *16*).

The 15.5K protein also binds to a K turn in box C/D small nucleolar (sno) RNAs (*17*, *18*), but subsequently Nop56 and Nop58 (Nop56/58; Nop5p in archaea) are recruited to the snoRNPs (Fig. 1A) (*17*, *19*). Stem II of the snRNAs and snoRNAs (Fig. 1A) encompasses crucial identity elements for secondary protein binding. In the box C/D snoRNAs, stem II is longer by one base

[1]Abteilung Zelluläre Biochemie, Max-Planck-Institut für Biophysikalische Chemie, Am Faßberg 11, D-37077 Göttingen, Germany. [2]AG Flüssig-NMR Spektroskopie, Max-Planck-Institut für Biophysikalische Chemie, Am Faßberg 11, D-37077 Göttingen, Germany. [3]AG Makromolekulare Röntgenkristallographie, Max-Planck-Institut für Biophysikalische Chemie, Am Faßberg 11, D-37077 Göttingen, Germany.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Reinhard.Luehrmann@mpi-bpc.mpg.de (R.L.); taco@nmr.mpibpc.mpg.de (T.C.); mwahl@gwdg.de (M.C.W.)

pair, and no sequence deviation is tolerated (*19*–*21*). Somewhat paradoxically, both hPrp31 and Nop56/58 contain a conserved, ~120-residue Nop domain (hPrp31$^{215\text{-}333}$) (*15*, *22*, *23*) (fig. S1), which seems to mediate binding to the different primary RNPs (*24*).

To delineate the structural basis for the ordered and selective binding of hPrp31, we first probed whether hPrp31 engages in direct contacts with 15.5K in the context of the U4 snRNP by using nuclear magnetic resonance (NMR) spectroscopy (*25*). [^{15}N]15.5K protein was bound to an RNA representing the entire U4 5′-SL [residues 20 to 52 of U4 snRNA (Fig. 1A)], and NMR chemical shift changes were monitored upon addition of hPrp31. Primarily residues in helices α2 and α3 of 15.5K were affected (Fig. 1B). Saturation transfer from the aliphatic protons of hPrp31 to the amide resonances of 15.5K in a ternary complex containing [^{15}N,^{2}D,^{1}H$_{N}$]15.5K confirmed direct contacts between hPrp31 and helices α2 and α3 of 15.5K (Fig. 1B).

By using limited proteolysis, we defined fragment hPrp31$^{78\text{-}333}$, whose binding activity resembled that of full-length hPrp31 [Supporting Online Material (SOM) text and fig. S2]. A reconstituted hPrp31$^{78\text{-}333}$-15.5K-U4 5′-SL complex yielded a 2.6 Å resolution crystal structure (table S1), in which residues 4 to 128 of 15.5K and 85 to 333 of hPrp31 (excluding residues 256 to 265 that form the tip of a flexible loop) and all RNA residues (nucleotides 20 to 52 of U4 snRNA) could be traced (fig. S3). The hPrp31$^{78\text{-}333}$–15.5K–U4 5′-SL complex is triangular, with one subunit at each vertex of the triangle and each subunit contacting the other two (Fig. 2A). The negatively charged RNA is sandwiched between positively charged areas on hPrp31$^{78\text{-}333}$ and 15.5K (fig. S4). The region of hPrp31$^{78\text{-}333}$ interacting with 15.5K exhibits alternating positively and negatively charged surface patches matched by a complementary set of patches on the 15.5K protein (fig. S4).

hPrp31$^{78\text{-}333}$ exhibits an all-helical fold with three domains (Fig. 2A and fig. S1). Residues 85 to 120 (helix α1) and 181 to 215 (helix α6) form two branches of an extended coiled coil, which is interrupted at the tip by a small globular module (residues 121 to 180; helices α2 to α5). An oval-shaped Nop domain (residues 215 to 333; helices α7 to α13) follows the coiled-coil motif at the C terminus. hPrp31$^{78\text{-}333}$ contacts the primary RNP exclusively via its Nop domain (Fig. 2A), suggesting that this element is the most crucial RNP interacting module in hPrp31.

The Nop domain of hPrp31$^{78\text{-}333}$ exhibits a flat surface formed by helices α9, α10, α12, and α13 (Fig. 2, A and B). The lower part of this surface (helix α9 and the C-terminal half of helix α12) interacts with the α2 and α3 region of 15.5K (contact regions a and b in Fig. 2B). Details of the interactions between hPrp31$^{78\text{-}333}$ and 15.5K (Fig. 3, A and B) agree well with full-length hPrp31-15.5K contacts mapped by NMR (Fig. 1B). The upper portion of the surface (helix α10 and the N-terminal half of helix α12) contacts the RNA on the side that is not associated with 15.5K (contact region 1 in Fig. 2B) and in the major groove of stem II (region 2). A loop following helix α10 interacts with the capping pentaloop (region 3). The surfaces of 15.5K and of the RNA buried upon binding of

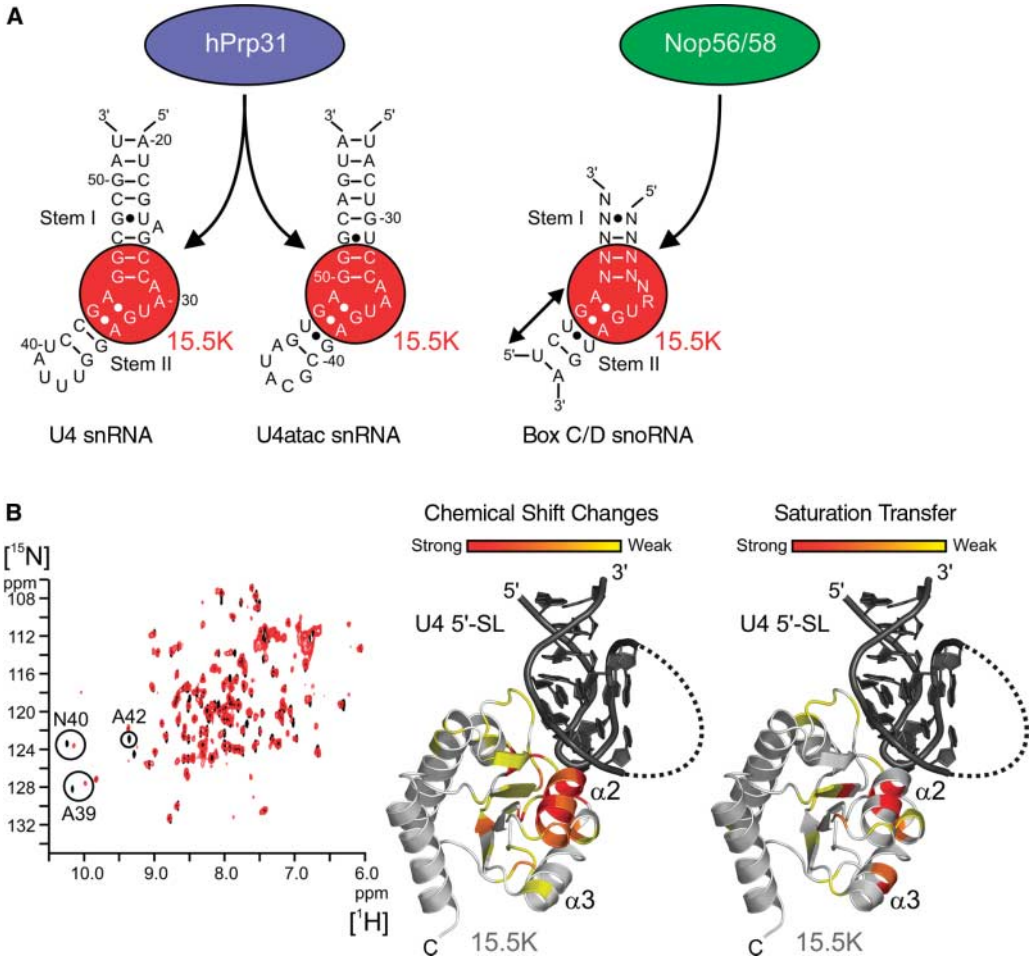

Fig. 1. (**A**) Schematics of the 5′-SLs of U4 snRNA (left), U4atac snRNA (middle), and the K-turn region of box C/D snoRNAs (right). N indicates any nucleotide; R, purine. Binding of 15.5K and the secondary binding proteins is indicated. Stem II of the K turn in the box C/D snoRNAs is longer by one base pair (*20*), and a single additional base pair in stem II is known to interfere with hPrp31 binding (*21*). Box C/D snoRNPs act as sequence-specific 2′-O methyltransferases, which posttranscriptionally modify several functional RNAs. (**B**) (Left) ^{1}H-^{15}N heteronuclear single-quantum coherence spectra of 15.5K in the binary 15.5K–U4 5′-SL complex (black) and in the ternary complex containing full-length hPrp31 (red). Assignments of selected resonances are indicated. ppm, parts per million. (Middle) Mapping of NMR chemical shift changes elicited by the addition of hPrp31 on the structure of the 15.5K-RNA complex [coordinates from (*12*); PDB ID 1E7K]. Dashed line is the disordered pentaloop of the RNA; 15.5K, light gray; and RNA, dark gray. All structure figures were prepared with PyMOL (*34*). (Right) Mapping of saturation transfer from hPrp31 to RNA-bound 15.5K, indicating residues of 15.5K that are directly contacted by hPrp31. Apparent contacts to the central β sheet of 15.5K arise from spin diffusion. Degrees of chemical shift changes and saturation transfer are color-coded: red, strong; orange, intermediate; and yellow, weak. A similar picture is obtained when mapping the contacts of hPrp31$^{78\text{-}333}$ on 15.5K in the framework of the 15.5K–U4 5′-SL complex (SOM text), confirming that the hPrp31$^{78\text{-}333}$ fragment interacts with 15.5K in the same way as the full-length protein.

the Nop domain are comparable (550 to 650 Å^2 each) and are confluent (fig. S5). Thus, the Nop domain presents an RNP recognition motif, as opposed to pure RNA interaction domains found in other proteins (*26*).

On the basis of this architecture, failure of hPrp31 to bind either 15.5K or the RNA alone (*13*, *14*) can be attributed to 15.5K and the RNA, each contributing about half of the hPrp31 interaction surface, so that neither of the components alone is able to supply a sufficiently large interface. In addition, 15.5K stabilizes the RNA K-turn region in a conformation favorable for hPrp31 binding and thus pays the entropic cost for immobilizing part of the RNA structure (SOM text).

The conformation of the core 15.5K-RNA complex is unaffected by the addition of hPrp31$^{78\text{-}333}$ (Fig. 2A). Furthermore, the Nop domain of hPrp31$^{78\text{-}333}$ closely resembles the corresponding domain of the archaeal Nop5p protein in the absence of a primary RNP (*24*) (table S2). Therefore, binding of the hPrp31 Nop domain to 15.5K and the K turn resembles a lock-and-key–type interaction. hPrp31$^{78\text{-}333}$ binds to one side of the bulged U31 of the K-turn via water-mediated interactions, van der Waals contacts, and hydrogen bonds to the backbone (Fig. 3C and contact region 1 in Fig. 2B). This situation is reminiscent of protein-DNA interactions, where the DNA-bound water structure provides important latching points for proteins (*27*). In addition, the short helix α10 of the Nop domain lines the major groove of stem II of the RNA where C247 (*28*) engages in hydrogen bonds to the bases of C41 and G43, representing the only sequence-specific contact of hPrp31$^{78\text{-}333}$ to the RNA (Fig. 3D; contact region 2 in Fig. 2B).

Major structural changes in the RNA upon formation of the ternary complex are confined to the RNA pentaloop (nucleotides 36 to 40), which becomes ordered on addition of hPrp31$^{78\text{-}333}$ (Fig. 2A). hPrp31$^{78\text{-}333}$ stabilizes the pentaloop by direct contacts and by reinforcing intramolecular interactions. H270 (*28*), from a flexible loop of hPrp31$^{78\text{-}333}$, stacks on the penultimate residue of the pentaloop, A39, which in turn stacks on the terminal base pair of stem II and forms a hydrogen bond across the loop (Fig. 3E and contact region 3 in Fig. 2B). The apparent malleability of the RNA pentaloop allows its remaining part to wrap around and engage in hydrogen bonds with H270 (Fig. 3E). The terminal U40 of the pentaloop stacks on a hydrophobic surface patch of the Nop domain (Fig. 3F). Because the above contacts ensue between flexible elements of the protein and the RNA, we conclude that hPrp31$^{78\text{-}333}$ recognizes and stabilizes the RNA pentaloop by induced fit interactions (SOM text).

Although the entire pentaloop of the U4 5′-SL can be removed without completely disrupting the binding of hPrp31, it becomes protected from hydroxyl radical cleavage upon binding of hPrp31 (*21*) or hPrp31$^{78\text{-}333}$ (Fig. 4A). Our structure reconciles this apparent discrepancy. Even though large portions of the pentaloop are exposed, the C4' atoms, which are the primary sites of hydroxyl radical attack (*29*), are partially or entirely buried for residues 36, 39, and 40 (Fig. 4B). The flexible loop of hPrp31$^{78\text{-}333}$ (residues 256 to 265) is suspended next to the C4' atoms of residues 37 and 38 (Fig. 4B), where it can scavenge radicals and protect the RNA even in the absence of direct contacts.

hPrp31 recognizes complexes of 15.5K with U4 or U4atac snRNA, whose sequences differ markedly in the hPrp31 contact regions defined by the crystal structure (Fig. 1A). In the U4 structure, hPrp31$^{78\text{-}333}$ avoids sequence-specific interactions with the RNA bases and instead maintains water-mediated interactions, contacts to the backbone, or stacking interactions with the bases (Fig. 3, C to F). With the exception of a single hydrogen bond from C247 of hPrp31$^{78\text{-}333}$ to C41 of the U4 snRNA, all contacts could be maintained in a complex containing the U4atac 5′-SL. Consistent with a similar interaction, H270 of hPrp31 has been ultraviolet light (UV)–cross-linked to U44 of the U4atac snRNA (*13*, *30*), implying an identical stacking arrangement as seen with the corresponding A39 of U4 snRNA (Fig. 3E).

The largely sequence-independent RNA contacts of hPrp31$^{78\text{-}333}$ suggest that structural rather than sequence differences preclude bind-

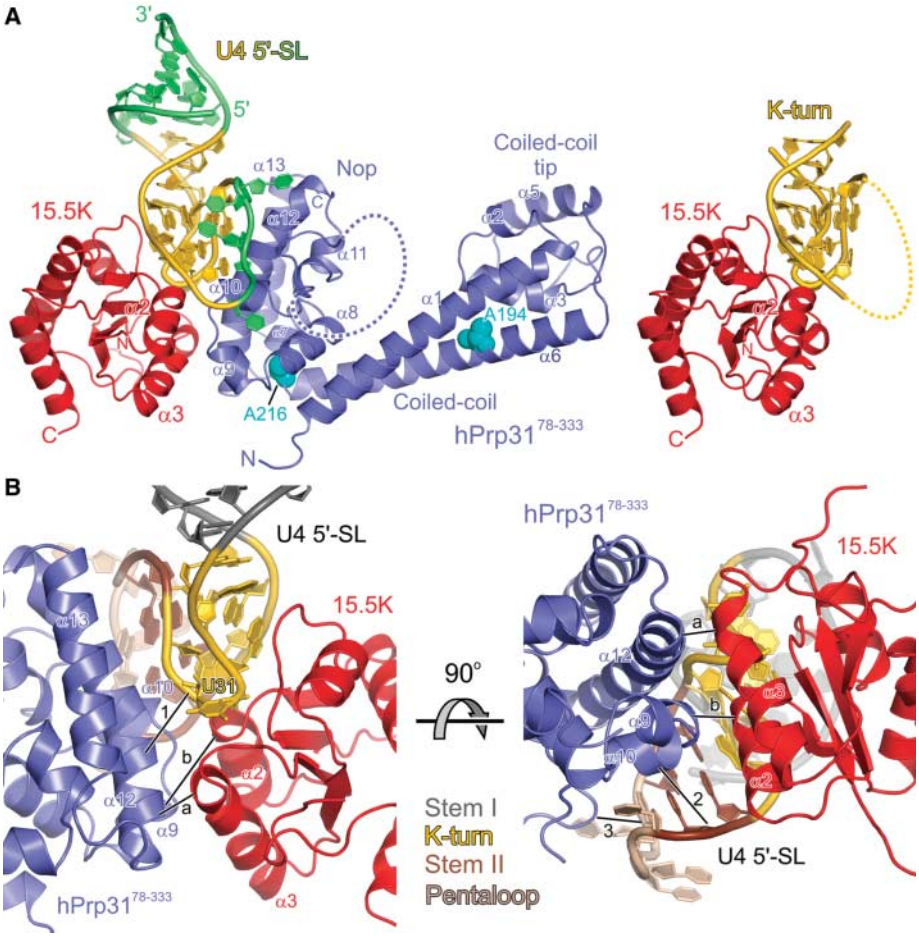

Fig. 2. (**A**) Overview of the hPrp31$^{78\text{-}333}$–15.5K–U4 5′-SL complex (left). hPrp31$^{78\text{-}333}$, blue; 15.5K, red; RNA, gold. RNA elements not seen in the binary 15.5K–22-mer RNA complex with a shorter stem I [right (*12*); PDB ID 1E7K] are in green. Positions A194 and A216, at which missense mutations have been linked to the RP11 form of retinitis pigmentosa, are shown as space-filling models and colored cyan. Dashed line in hPrp31$^{78\text{-}333}$, disordered loop. Dashed line in the binary complex, unstructured pentaloop. Although induced-fit interactions are the hallmark of most RNA-protein complexes (*32*), the structuring of the RNA pentaloop upon hPrp31$^{78\text{-}333}$ binding observed here is particularly pronounced. The crystal structure contains two crystallographically independent ternary complexes per asymmetric unit that are largely identical (table S2). (**B**) Close-up views of the complex from the back (left) and from the bottom (right). Main contact regions between hPrp31$^{78\text{-}333}$ and 15.5K and between hPrp31$^{78\text{-}333}$ and the RNA are indicated by connecting lines and are labeled by letters and numbers, respectively. Regions of the RNA are color-coded: distal portion of stem I, gray; K-turn region, gold; distal portion of stem II, brown; and pentaloop, beige. The bulged-out U31 denotes the tip of the K turn and is shown in sticks.

ing to box C/D-like RNAs. In our structure, all hPrp31 contacts with the RNA are mediated by the Nop domain, suggesting that this motif alone is able to discriminate against an extended stem II. To verify this hypothesis, we conducted gel mobility shift assays by using wild-type (WT) and mutant U4 5′-SLs and a Nop-domain fusion protein [maltose binding protein (MBP)–hPrp31$^{215-333}$]. Like full-length hPrp31, the Nop domain did not bind to RNPs, in which stem II of the RNA was extended by a noncanonical U-U (Fig. 4C, lanes 7 to 9) or by a canonical C-G base pair (lanes 10 to 12). These data confirm that the Nop domain is both required and sufficient for binding to the primary RNP and for decoding its structural specificity determinants.

An elongated stem II would reposition the pentaloop and thus the stacking platform for H270 (Fig. 3E). Loss of hPrp31 affinity to 15.5K complexes with elongated stem II RNAs could, therefore, arise due to the disruption of H270-A39 stacking. We tested this possibility by converting H270 to an alanine or to a lysine (as found in Nop56/58). Loss or alteration of the H270 side chain resulted in reduced affinity of hPrp31^{H270A} and hPrp31^{H270K} to the 15.5K-RNA 5′-SL complexes (Fig. 4D, lanes 1 to 5). However, the mutants retained significant binding activity and discriminated strongly against long stem II constructs (Fig. 4D, lanes 6 to 10), indicating that H270 is not required for measuring the length of stem II.

Next, we modeled an extended stem II A-form duplex into the present structure. The additional base pair would lie in the stacking level occupied by A39 in the WT complex. Although A39 fits snugly next to helix α10 of the Nop domain, the helical twist would lead to a severe clash between an additional Watson-Crick base pair and helix α10 (fig. S6). In contrast, in a tetraloop RNA, to which hPrp31 (*21*) and hPrp31$^{215-333}$ (Fig. 4C, lanes 4 to 6) still bind, the nucleotides would be turned away from helix α10. Thus, we conclude that helix α10 acts as a ruler for measuring the length of stem II by presenting a physical barrier to additional base pairs. A different recognition mechanism must be at work in Nop56/58, in which the Nop domain is compatible with elongated stem II RNAs.

The hPrp31$^{78-333}$ fragment encompasses both A194 and A216 (*28*) that have been linked to the RP11 form of retinitis pigmentosa (*23*). A194 maps to the second helix of the coiled coil, whereas A216 lies in a short loop connecting the coiled coil to the Nop domain (Fig. 2A and fig. S7). Thus, neither of the two residues interacts directly with 15.5K or the U4 5′-SL in the present structure (Fig. 2A). The Ala^{194}→Glu^{194} (A194E) substitution most likely disturbs the local structure and/or the surface properties of the coiled coil, because the A194 side chain is embedded in a hydrophobic environment (fig. S7A). The Ala^{216}→Pro^{216} (A216P) replacement (fig. S7A) potentially

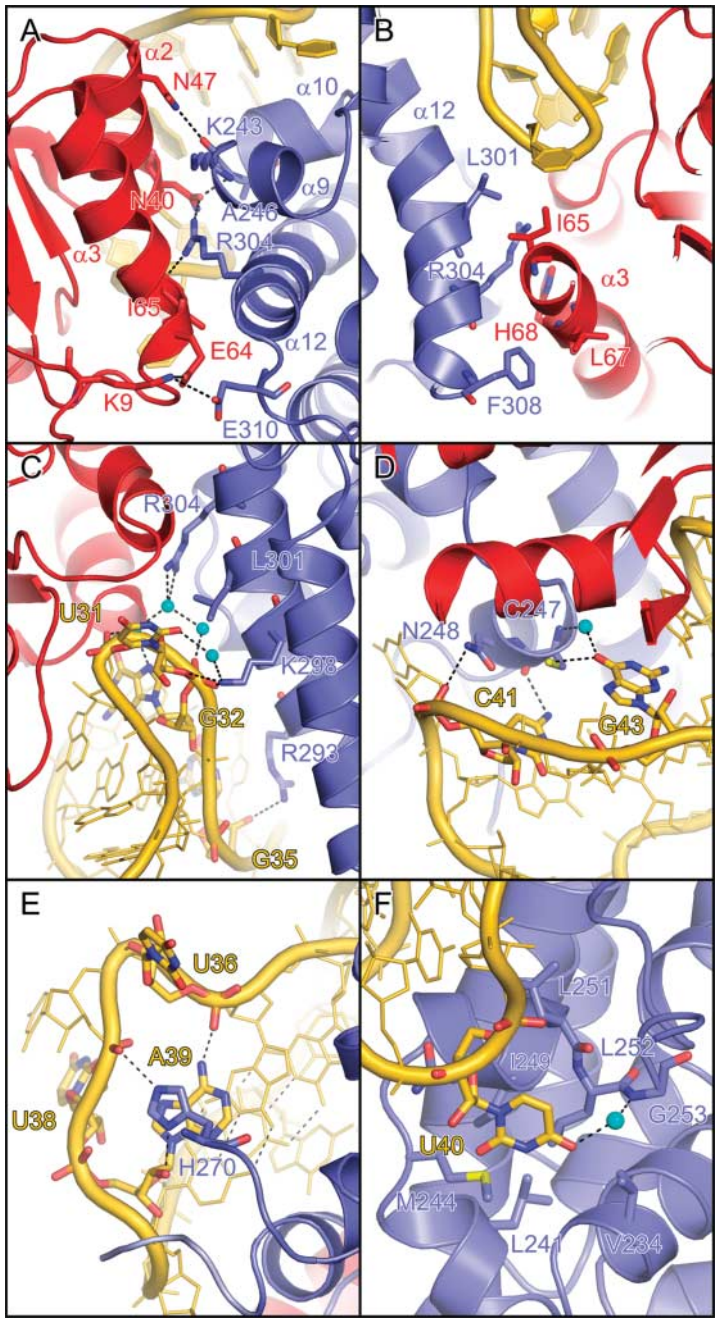

Fig. 3. Detailed interactions in the ternary complex. Selected interface residues (sticks) are labeled and color-coded by atom type (carbon and phosphorus, as the respective molecule; oxygen, red; nitrogen, blue; sulfur, yellow; bridging waters, cyan spheres). (**A**) Hydrogen bonds and salt bridges (dashed lines) involving helices α2 and α3 of 15.5K and the hPrp31$^{78-333}$ Nop domain (view from bottom of Fig. 2A). A network of alternating residues from 15.5K and hPrp31$^{78-333}$ extends from the backbone carbonyl of I65 (15.5K) over the side chain of R304 (hPrp31$^{78-333}$) and the side chain of N40 (15.5K) to the backbone nitrogen of A246 (hPrp31$^{78-333}$). (**B**) Hydrophobic contacts between the N terminus of helix α3 of 15.5K and the C terminus of helix α12 of hPrp31 (view as in Fig. 2A). Our crystal structure and NMR analysis show that residues 74 to 77 of 15.5K, which upon joint mutation inhibited the binding of hPrp31 (*33*), are not involved in contacts to hPrp31$^{78-333}$ or full-length hPrp31, respectively, and elicit their effect indirectly, for example, by influencing the structure of 15.5K. (**C** to **F**) Details of the interaction of hPrp31$^{78-333}$ with the K turn (C), with the major groove of stem II (D), with the 5′ portion of the pentaloop (E) and with the 3′ portion of the pentaloop (F). The central lock-and-key interaction region around the K turn [(C) and (D)] is of paramount importance for the stability of the ternary complex, because the RNA pentaloop can be removed without completely disrupting the binding of hPrp31 (*21*). The induction of a stable structure in the RNA pentaloop by hPrp31$^{78-333}$ [(E) and (F)] is reminiscent of the way some primary binding ribosomal proteins induce novel binding sites for secondary binding proteins (*10*). Here, three of the five pentaloop bases are turned outward (E) and conceivably provide a binding platform for another spliceosomal component.

alters the structure and flexibility of a protein loop, which may affect the relative orientation of the Nop and the coiled-coil domains. Either of the above mutations could influence other interactions of hPrp31, for example, with hPrp6 of the U5 snRNP. To test this hypothesis, we conducted a targeted yeast two-hybrid analysis using pGADT7-hPrp6 as prey and pGBKT7-hPrp31, pGBKT7-hPrp31^{A194E}, and pGBKT7-hPrp31^{A216P} as bait under stringent conditions. Whereas introduction of proline at position 216 had no effect on the interaction with hPrp6, a glutamate at position 194 significantly weakened the interaction (fig. S7B). This result supports the idea that the hPrp31 coiled coil is an interaction site for hPrp6 and links a retinitis pigmentosa mutation in a spliceosomal protein to an aberrant molecular communication.

By interacting concomitantly with both 15.5K and the RNA, the Nop domain reinforces the 15.5K-RNA interaction. The latter interaction is crucial for the transition from the spliceosomal B complex to the C complex (*11*), during which spliceosome activation occurs. Thus, on the one side hPrp31 may regulate the RNA-protein network at the U4 5′-SL and thereby facilitate disruption of the U4/U6 snRNA duplex. Our work also suggests that hPrp31 stabilizes the U4/U6-U5 tri-snRNP by concomitantly interacting with hPrp6 via a separate coiled-coil domain. Intriguingly, hPrp6 links hPrp31 to hBrr2 and hSnu114 (*14*), which have been shown to be the DEAD-box protein and regulatory guanosine triphosphatase (GTPase), respectively, that are crucial for both spliceosome activation and disassembly (*31*). Therefore, our structural data are in line with the previous hypothesis (*14*) that hPrp31 may represent one of the ultimate targets of the helicase and GTPase machinery of the U5 snRNP that acts during spliceosome activation.

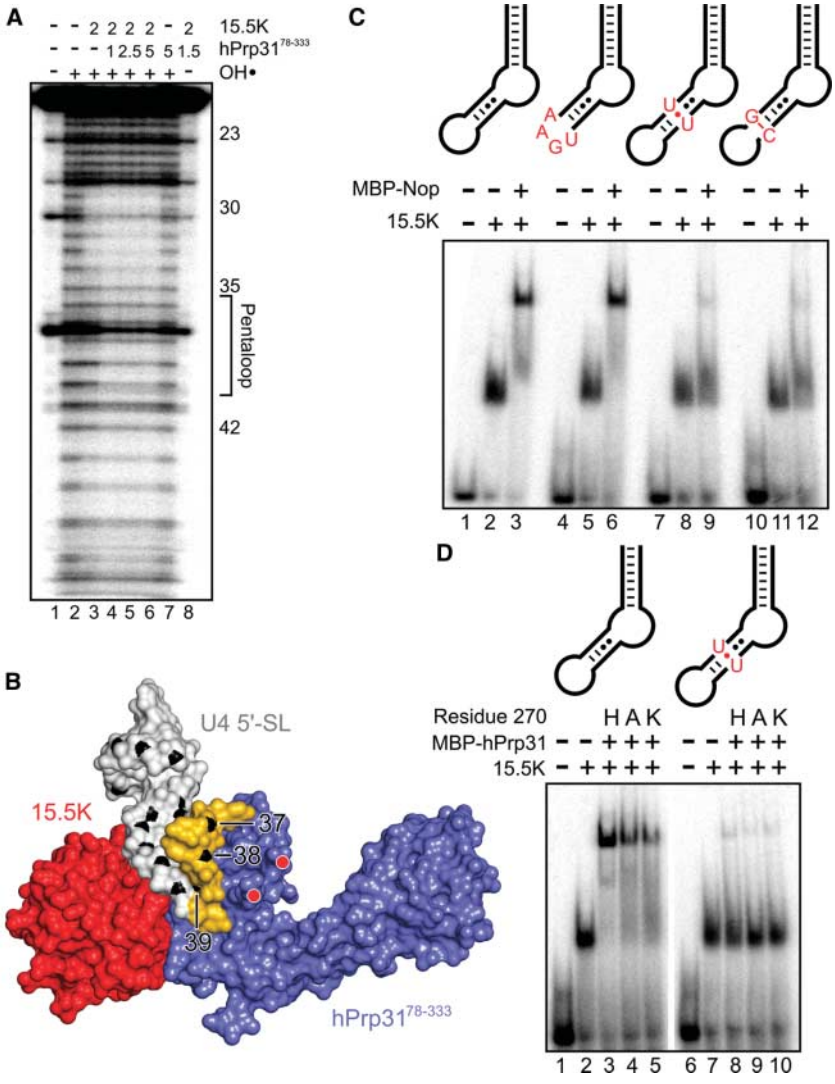

Fig. 4. (**A**) Hydroxyl radical footprinting of the U4 5′-SL in the absence of protein (lane 2), in the presence of only 15.5K (lane 3) and in the presence of 15.5K and increasing amounts of hPrp31$^{78-333}$ (lanes 4 to 6). Numbers indicate the protein concentration in μmol/l. Numbers on the right indicate positions in the U4 5′-SL. The location of the pentaloop is indicated. (**B**) Surface of the hPrp31$^{78-333}$–15.5K–U4 5′-SL complex (RNA, gray; pentaloop, gold). Sugar C4' atoms are highlighted in black and labeled for residues 37 to 39. Red dots, beginning (lower dot) and end (upper dot) of a flexible loop in the protein that is suspended next to the sugars of the pentaloop. (**C**) Gel mobility shift assays monitoring the binding of a Nop domain fusion protein (MBP-hPrp31$^{215-333}$) to U4 5′-SL constructs. Lanes 1 to 3, WT RNA sequence; lanes 4 to 6, replacement of the pentaloop by a UGAA tetraloop; lanes 7 to 9, addition of a U-U base pair to stem II following the sheared G-A pairs; and lanes 10 to 12, addition of a C-G base pair at the terminus of stem II. (**D**) Gel mobility shift assays monitoring the effects of converting H270 into an alanine (A) or a lysine (K). Mutant hPrp31 proteins bind less strongly to a WT U4 5′-SL (lanes 1 to 5) but still discriminate against RNAs with a longer stem II (lanes 6 to 10).

References and Notes

1. C. B. Burge, T. Tuschl, P. A. Sharp, in *The RNA World*, R. F. Gesteland, T. Cech, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999), pp. 525–560.
2. T. W. Nilsen, *Bioessays* **25**, 1147 (2003).
3. C. L. Will, R. Lührmann, in *The RNA World*, R. F. Gesteland, T. Cech, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2006), pp. 369–400.
4. R. Reed, L. Palandjian, in *Eukaryotic mRNA Processing*, A. R. Krainer, Ed. (IRL Press, Oxford, 1997), pp. 103–129.
5. J. Gornemann, K. M. Kotovic, K. Hujer, K. M. Neugebauer, *Mol. Cell* **19**, 53 (2005).
6. J. P. Staley, C. Guthrie, *Cell* **92**, 315 (1998).
7. D. A. Brow, *Annu. Rev. Genet.* **36**, 333 (2002).
8. P. Bringmann *et al.*, *EMBO J.* **3**, 1357 (1984).
9. C. Hashimoto, J. A. Steitz, *Nucleic Acids Res.* **12**, 3283 (1984).
10. S. C. Agalarov, G. S. Prasad, P. M. Funke, C. D. Stout, J. R. Williamson, *Science* **288**, 107 (2000).
11. S. Nottrott *et al.*, *EMBO J.* **18**, 6119 (1999).
12. I. Vidovic, S. Nottrott, K. Hartmuth, R. Lührmann, R. Ficner, *Mol. Cell* **6**, 1331 (2000).
13. S. Nottrott, H. Urlaub, R. Lührmann, *EMBO J.* **21**, 5527 (2002).
14. S. Liu, R. Rauhut, H. P. Vornlocher, R. Lührmann, *RNA* **12**, 1418 (2006).
15. O. V. Makarova, E. M. Makarov, S. Liu, H. P. Vornlocher, R. Lührmann, *EMBO J.* **21**, 1148 (2002).
16. C. Schneider, C. L. Will, O. V. Makarova, E. M. Makarov, R. Lührmann, *Mol. Cell. Biol.* **22**, 3219 (2002).
17. N. J. Watkins *et al.*, *Cell* **103**, 457 (2000).
18. L. B. Szewczak, J. S. Gabrielsen, S. J. Degregorio, S. A. Strobel, J. A. Steitz, *RNA* **11**, 1407 (2005).
19. N. J. Watkins, A. Dickmanns, R. Lührmann, *Mol. Cell. Biol.* **22**, 8342 (2002).
20. T. Moore, Y. Zhang, M. O. Fenley, H. Li, *Structure* **12**, 807 (2004).
21. A. Schultz, S. Nottrott, K. Hartmuth, R. Lührmann, *J. Biol. Chem.* **281**, 28278 (2006).
22. T. Gautier, T. Berges, D. Tollervey, E. Hurt, *Mol. Cell. Biol.* **17**, 7088 (1997).
23. E. N. Vithana *et al.*, *Mol. Cell* **8**, 375 (2001).
24. M. Aittaleb *et al.*, *Nat. Struct. Biol.* **10**, 256 (2003).
25. Materials and methods are available as supporting material on *Science* Online.
26. C. G. Burd, G. Dreyfuss, *Science* **265**, 615 (1994).
27. Z. Shakked *et al.*, *Nature* **368**, 469 (1994).
28. 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.

29. J. C. Wu, J. W. Kozarich, J. Stubbe, *J. Biol. Chem.* **258**, 4694 (1983).
30. E. Kühn-Hölsken, C. Lenz, B. Sander, R. Lührmann, H. Urlaub, *RNA* **11**, 1915 (2005).
31. E. C. Small, S. R. Leggett, A. A. Winans, J. P. Staley, *Mol. Cell* **23**, 389 (2006).
32. J. R. Williamson, *Nat. Struct. Biol.* **7**, 834 (2000).
33. A. Schultz, S. Nottrott, N. J. Watkins, R. Lührmann, *Mol. Cell. Biol.* **26**, 5146 (2006).
34. PyMOL, http://pymol.sourceforge.net/.
35. We thank S. Bhattacharya (University College London, UK) for supplying pTriEx-hPrp31, H. Urlaub for mass spectrometric analyses, G. Bourenkov (European Molecular Biology Laboratory, Hamburg, Germany) and the group of H. Bartunik (beamline BW6, Deutsches Elektronen Synchrotron, Hamburg, Germany) for beamline support, and C. Will for fruitful discussions and critical reading of the manuscript. This work was supported by the Max-Planck-Gesellschaft (to R.L., T.C., and M.C.W.), the Deutsche Forschungsgemeinschaft (LU294/12-3), the Fonds der Chemischen Industrie (to R.L.), and the Ernst-Jung-Stiftung (to R.L.). Coordinates and structure factors have been deposited with the RCSB Protein Data Bank (www.rcsb.org/pdb/) under accession code 2OZB and will be released upon publication. The authors declare that they have no competing financial interests.

Supporting Online Material

www.sciencemag.org/cgi/content/full/316/5821/115/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 and S2
References

22 November 2006; accepted 1 March 2007
10.1126/science.1137924

An ATP Gate Controls Tubulin Binding by the Tethered Head of Kinesin-1

Maria C. Alonso,[1] Douglas R. Drummond,[1] Susan Kain,[1] Julia Hoeng,[2] Linda Amos,[2] Robert A. Cross[1]*

Kinesin-1 is a two-headed molecular motor that walks along microtubules, with each step gated by adenosine triphosphate (ATP) binding. Existing models for the gating mechanism propose a role for the microtubule lattice. We show that unpolymerized tubulin binds to kinesin-1, causing tubulin-activated release of adenosine diphosphate (ADP). With no added nucleotide, each kinesin-1 dimer binds one tubulin heterodimer. In adenylyl-imidodiphosphate (AMP-PNP), a nonhydrolyzable ATP analog, each kinesin-1 dimer binds two tubulin heterodimers. The data reveal an ATP gate that operates independently of the microtubule lattice, by ATP-dependent release of a steric or allosteric block on the tubulin binding site of the tethered kinesin-ADP head.

Kinesin-1 molecular motors are adenosine triphosphate (ATP)–driven walking machines that move in 8-nm steps toward the plus ends of microtubules, turning over one ATP molecule per step under a range of loads (*1*–*5*). Even at very high backward loads, when the motor can be forced to step processively backward (*6*), stepping remains coupled to ATP binding (*6*, *7*). Between steps, the motor pauses stably in a dwell state. It is clear that ATP binding triggers exit from this dwell state, but the structural mechanism is controversial (*8*).

Two general types of model for this ATP gate have been proposed. In the first, kinesin dimers are proposed to dwell between steps with only one head attached to the microtubule, whereas the other diffuses to some extent on its tether but cannot access its next binding site along the microtubule because the site is too far away. ATP binding to the microtubule-attached head drives a conformational change that shifts the tethered head along the microtubule, biasing and focusing its diffusional search for its next binding site (*9*, *10*). In the second type of model, kinesin is proposed to dwell with both heads attached to the microtubule (*5*), and gating is ascribed to the effects of the resulting intramolecular strain, together with any external strain, on nucleotide exchange (*11*, *12*). These two types of model are not mutually exclusive; the influential Rice *et al*. model (*9*), for example, proposes that the first step in each run of steps uses the first type of gate and that subsequent steps use the second type. Both types of model require the microtubule lattice for their operation, either to set a prohibitive distance between binding sites or to apply strain to the kinesin heads. Here, we report an ATP gate that operates independently of the microtubule lattice.

We have found that kinesin-1 binds to free tubulin heterodimers in solution, causing tubulin-activated release of adenosine diphosphate (ADP). This shows that tubulin activation of the kinesin adenosine triphosphatase (ATPase) is not unique, as had previously been thought, to the depolymerizing kinesins (*13*). The degree of activation of the kinesin-1 ATPase by unpolymerized tubulin varies according to the source of kinesin-1 and tubulin, but clearly activation of the kinesin-1 ATPase does not require the interheterodimer interfaces that arise in the microtubule lattice. For a fungal kinesin-1 and a fungal tubulin, maximal activation by unpolymerized tubulin heterodimers is equivalent to that produced by assembled microtubules (Fig. 1A). For brain tubulin and brain kinesin, tubulin activation of the kinesin ATPase is modest compared with microtubule activation (Fig. 1, B to D).

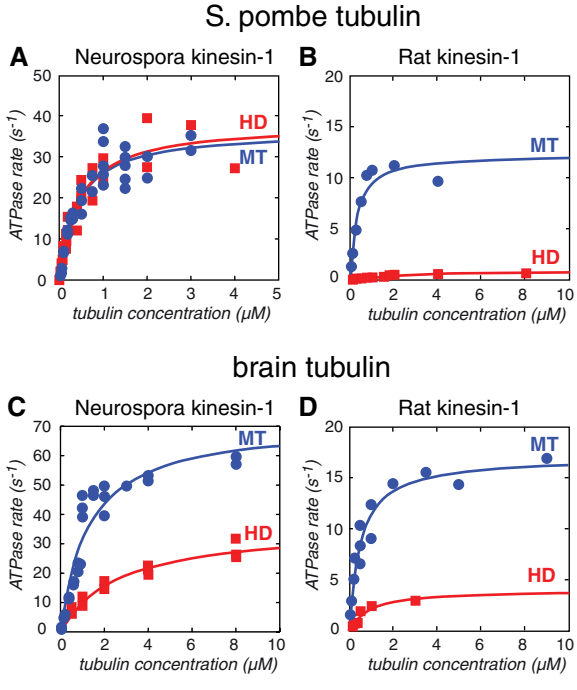

Fig. 1. Activation of kinesin dimers by tubulin and microtubules. The microtubule- or tubulin heterodimer–stimulated steady state ATPase activity of kinesin was measured at 25°C with an enzyme-linked assay in 20 mM Pipes, pH 6.9, 5 mM $MgCl_2$, 1 mM dithiothreitol (*22*). Values for V_{max} (the projected maximum ATPase rate) and K_m (the tubulin concentration giving half-maximal ATPase) were obtained by least-squares fitting to plots of ATPase versus tubulin heterodimer concentration, using Kaleidagraph 3.6.4 (Synergy Software, Reading, PA, USA). MT, microtubule; HD, heterodimer. (**A**) V_{max} 38.1 s^{-1}, K_m 0.44 μM for HD, V_{max} 36.4 s^{-1}, K_m 0.39 μM for MT. (**B**) V_{max} 0.9 s^{-1}, K_m 2.03 μM for HD, V_{max} 12.2 s^{-1}, K_m 0.28 μM for MT. (**C**) V_{max} 35.9 s^{-1}, K_m 1.29 μM for HD, V_{max} 71.1 s^{-1}, K_m 0.62 μM for MT. (**D**) V_{max} 4.0 s^{-1}, K_m 0.86 μM for HD, V_{max} 17.0 s^{-1}, K_m 0.49 μM for MT.

[1]Molecular Motors Group, Marie Curie Research Institute, The Chart, Oxted, Surrey RH8 0TL, UK. [2]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

*To whom correspondence should be addressed. E-mail: r.cross@mcri.ac.uk