3. H. D. Rundle, L. Nagel, J. W. Boughman, D. Schluter, *Science* **287**, 306 (2000).
4. M. White, *Modes of Speciation* (Freeman, San Francisco, 1978).
5. G. L. Bush, *Annu. Rev. Ecol. Syst.* **6**, 339 (1975).
6. E. Mayr, *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, MA, 1963).
7. S. P. Otto, J. Whitton, *Annu. Rev. Genet.* **34**, 401 (2000).
8. V. Grant, *Plant Speciation* (Columbia Univ. Press, New York, 1971).
9. A. R. Templeton, *Annu. Rev. Ecol. Syst.* **12**, 23 (1981).
10. L. H. Rieseberg, *Annu. Rev. Ecol. Syst.* **28**, 359 (1997).
11. D. Greig, E. J. Louis, R. H. Borts, M. Travisano, *Proc. R. Soc. London Ser. B* **269**, 1167 (2002).
12. G. Naumov, *Stud. Mycol.* **30**, 469 (1987).
13. N. Hunter, S. R. Chambers, E. J. Louis, R. H. Borts, *EMBO J.* **15**, 1726 (1996).
14. J. A. Coyne, H. A. Orr, *Philos. Trans. R. Soc. London Ser. B* **353**, 287 (1998).
15. Materials and methods are available as supporting material on *Science* Online.
16. D. Loidl, *Genetics* **139**, 1511 (1995).
17. D. R. Taylor, C. Zeyl, E. Cooke, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3690 (2002).
18. G. Marinoni *et al.*, *J. Bacteriol.* **181**, 6488 (1999).
19. F. Sebastiani, C. Barberio, E. Casalone, D. Cavalieri, M. Polsinelli, *Res. Microbiol.* **153**, 53 (2002).
20. I. Masneuf, J. Hansen, C. Groth, J. Piskur, D. Dubourdieu, *Appl. Environ. Microbiol.* **64**, 3887 (1998).
21. C. Groth, J. Hansen, J. Piskur, *Int. J. Syst. Bacteriol.* **49**, 1933 (1999).
22. R. K. Mortimer, *Genome Res.* **10**, 403 (2000).
23. L. H. Rieseberg, *Trends Ecol. Evol.* **16**, 351 (2001).
24. This work was supported by the Wellcome Trust.

Supporting Online Material
www.sciencemag.org/cgi/content/full/298/5599/1773/DC1
Materials and Methods
Table S1
Figs. S1 and S2
References and Notes

19 July 2002; accepted 24 September 2002

Essential Role for the SMN Complex in the Specificity of snRNP Assembly

Livio Pellizzoni,* Jeongsik Yong, Gideon Dreyfuss†

The Survival of Motor Neurons (SMN) protein, the product of the spinal muscular atrophy–determining gene, is part of a large macromolecular complex (SMN complex) that functions in the assembly of spliceosomal small nuclear ribonucleoproteins (snRNPs). Using cell extracts and purified components, we demonstrated that the SMN complex is necessary and sufficient to mediate the ATP-dependent assembly of the core of seven Sm proteins on uridine-rich, small nuclear ribonucleic acids (U snRNAs). In vitro experiments revealed strict requirements for ordered binding of the Sm proteins and the U snRNAs to the SMN complex. Importantly, the SMN complex is necessary to ensure that Sm cores assemble only on correct RNA targets and prevent their otherwise promiscuous association with other RNAs. Thus, the SMN complex functions as a specificity factor essential for the efficient assembly of Sm proteins on U snRNAs and likely protects cells from illicit, and potentially deleterious, nonspecific binding of Sm proteins to RNAs.

Nuclear pre-mRNA splicing, the process of removal of introns from pre-mRNAs, is an essential aspect of eukaryotic mRNA biogenesis that is carried out by the spliceosome. The snRNPs U1, U2, U4/U6, and U5 are essential and major components of the spliceosome. Each snRNP consists of one U snRNA molecule, a common core comprising a ring of seven Sm proteins, and several snRNP-specific proteins (*1*). The process of snRNP biogenesis, which occurs in the cytoplasm, requires the assembly of the Sm proteins on the Sm site, which is a uridine-rich sequence present in the U snRNAs, to form the Sm core (*2*). A properly assembled Sm core and the hypermethylated 5′ cap are both required to recruit the import receptors necessary for snRNP translocation into the nucleus (*1*, *3–6*). Once in the nucleus, snRNPs associate with specific proteins that are unique to each snRNA, and function in pre-mRNA splicing.

SnRNP assembly readily occurs in vitro from purified snRNP proteins and U snRNAs (*7–9*). This assembly reaction does not require ATP or non-snRNP factors. However, snRNP assembly in extracts of mammalian cells or *Xenopus laevis* eggs requires ATP hydrolysis (*10*, *11*). Considerable data reveal that a large macromolecular complex containing the Survival of Motor Neurons (SMN) protein is required for snRNP assembly (*11–14*). SMN is the product of the spinal muscular atrophy (SMA)–determining disease gene (*15*). Reduced levels of SMN protein result in SMA, a common neurodegenerative disease of the motor neurons (*16*, *17*). The SMN protein is associated with Gemin2, Gemin3, Gemin4, Gemin5, Gemin6, and Gemin7 in a large complex that localizes both in the cytoplasm and the nucleus (*18–24*). Although its role in snRNP assembly is better characterized, the SMN complex likely functions in the assembly and/or restructuring of several ribonucleoprotein particles including small nucleolar RNPs (snoRNPs) and the machineries that carry out transcription and pre-mRNA splicing (*25*). Despite advances in the characterization of the interactions and functions of the SMN complex, mechanistic insights into the molecular functions of the SMN complex in snRNP assembly have been lacking.

Using experiments in HeLa cell extracts, we showed that the SMN complex is necessary for snRNP assembly of all the major Sm site–containing U snRNAs and that this requires ATP hydrolysis (figs. S1 and S2). We therefore asked whether the SMN complex is not only necessary but also sufficient to mediate snRNP assembly. To do so, we tested the ability of purified SMN complexes to assemble Sm cores. Native SMN complexes were purified from cell lines expressing FLAG-tagged SMN or Gemin2 (*23*) by affinity chromatography on anti-FLAG beads and eluted with excess of the FLAG peptide (Fig. 1A) (*26*). SMN complexes purified from both FLAG-SMN and FLAG-Gemin2 cell lines are identical in composition and, at moderate salt concentrations, also contain Sm proteins (*23*). Purified SMN complexes were analyzed for their capacity to form Sm cores on the major Sm site U snRNAs. Through use of native gel electrophoresis under stringent conditions (*7*, *8*), we demonstrated that the SMN complex mediates snRNP assembly of Sm site–containing U snRNAs (Fig. 1B). The formation of Sm cores was confirmed by immunoprecipitation of the U snRNAs from assembly reactions with anti-Sm antibodies (*27*). ATP, but not ATP hydrolysis, is required for SMN complex–mediated Sm core assembly of U1 snRNA (Fig. 1C). These results demonstrate that a purified SMN complex containing the Sm proteins is necessary and sufficient to mediate the ATP-dependent assembly of snRNPs.

Sm core assembly can be reconstituted in vitro using purified total snRNP proteins (TPs) and in vitro–transcribed U snRNAs (*7–9*). The process of snRNP assembly with TPs is ATP-independent and does not require non-snRNP factors. We analyzed the effect of purified SMN complexes in the Sm core assembly with TPs. For these experiments, SMN complexes were purified under high salt conditions (500

Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104–6148, USA.

*Present address: Institute of Cell Biology, Consiglio Nazionale delle Ricerche, "Campus Adriano-Buzzati-Traverso," IBC 00016, Monterotondo Scalo, Rome, Italy.

†To whom correspondence should be addressed. E-mail: gdreyfuss@hhmi.upenn.edu

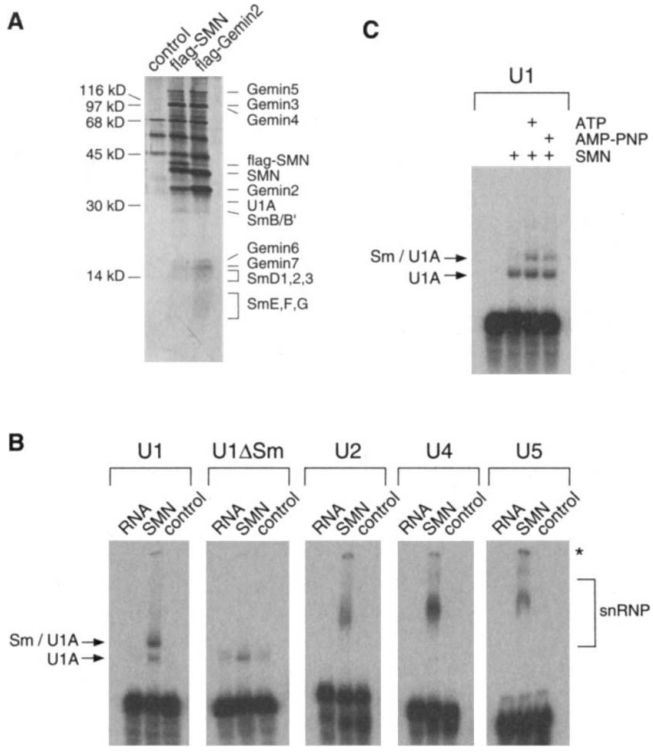

Fig. 1. Purified SMN complexes containing Sm proteins are sufficient to mediate the ATP-dependent snRNP assembly. (**A**) Purification of SMN complexes containing Sm proteins. SMN complexes were purified, as previously described, from HeLa cells that express FLAG-tagged SMN or Gemin2, but the high-salt wash procedure was omitted (*23*, *26*). Nonspecific proteins purified from the parental HeLa Tet-ON cells are also shown (control). Proteins were eluted with 3× FLAG peptide and analyzed by SDS-PAGE and silver staining. Molecular weight markers and known protein components are indicated. Substoichiometric amounts of a few additional unknown proteins co-purify with the SMN complex at low salt concentrations and may contribute to its function (*23*). (**B**) Purified SMN complexes containing Sm proteins mediate Sm core assembly of U snRNAs. Purified SMN complexes from the FLAG-Gemin2–expressing cells (SMN) or nonspecific proteins (control) were incubated with in vitro–transcribed [^{32}P]UTP-labeled U snRNAs for 45 min at 30°C (*26*). Reactions were analyzed by electrophoresis on native polyacrylamide gels. The identity of the RNP complexes is indicated. The asterisk marks complexes that migrate close to the origin of the gel and may contain both U snRNAs and the SMN complex. (**C**) Purified SMN complexes from the FLAG-Gemin2–expressing cells were incubated with in vitro–transcribed [^{32}P]UTP-labeled U1 snRNA for 45 min at 30°C in the absence or in the presence of 2.5 mM ATP or AMP-PNP as indicated. Reactions were analyzed by electrophoresis on native polyacrylamide gels. U1 snRNA complexes containing U1A or the Sm core and U1A are indicated.

mM NaCl) so that very little, if any, Sm proteins remained with the SMN complex (Fig. 2A, SMN) (*26*). TPs were purified according to standard procedures as described (Fig. 2B) (*7*, *26*). In vitro–transcribed [^{32}P]uridine triphosphate (UTP)–labeled U snRNAs were incubated with either purified SMN complexes or TPs, and formation of RNP complexes was analyzed by electrophoresis on native polyacrylamide gels. TPs efficiently assemble the Sm cores on the major Sm site–containing U snRNAs (U1, U2, U4, and U5), but not on U1ΔSm or U6 (Fig. 2C). After incubation of TPs with U1 snRNA, several complexes were detected that contain U1 snRNA bound to U1A, the Sm core and U1A, and the Sm core together with the U1-specific proteins U1A and U1 70K (Fig. 2C) (*7*, *11*). Only the U1A-containing complex forms on U1ΔSm. Under the stringent conditions of gel electrophoresis employed here, purified SMN complexes do not form stable complexes with U snRNAs, and only the U1A-containing complex on U1 is detected (Fig. 2C). We then analyzed the effect of adding purified SMN complexes together with TPs and U snRNAs. Purified SMN complexes inhibit snRNP assembly on U snRNAs, but not the interaction of U1A with U1 snRNA (Fig. 2C). These experiments demonstrate that, under these experimental conditions, the SMN complex inhibits the spontaneous assembly of TPs on U snRNAs.

In the above experiments (Fig. 2C), all the components were added at the same time. We performed order-of-addition experiments to test whether the SMN complex prevents the

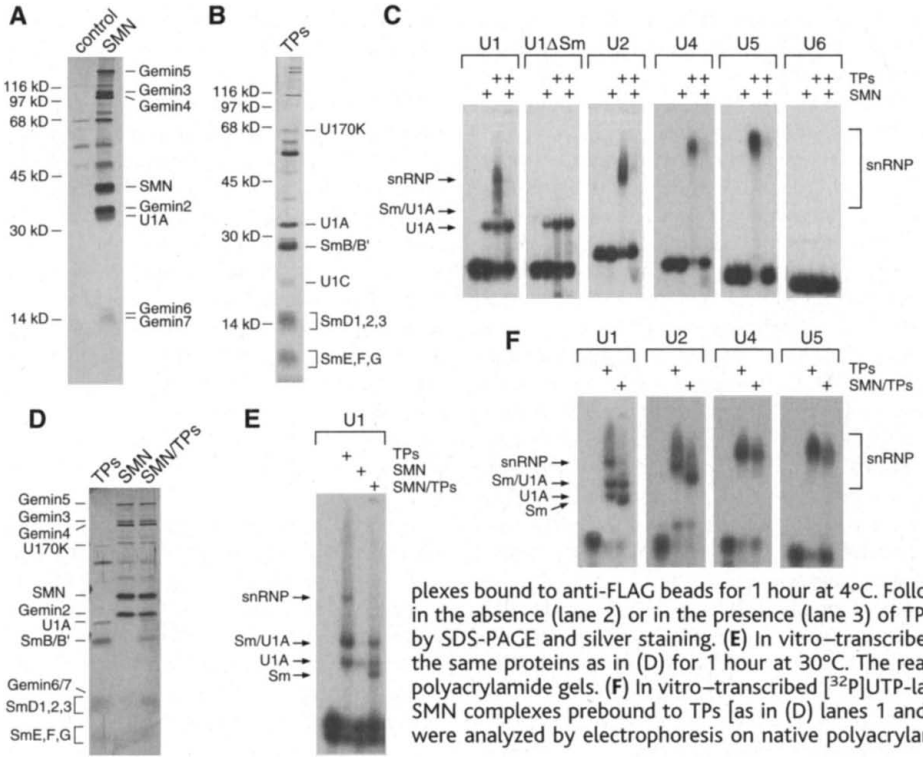

Fig. 2. The SMN complex imposes an ordered pathway to Sm core assembly. (**A**) SMN complexes (SMN) or nonspecific proteins (control) were affinity-purified from total extracts of a stable cell line expressing FLAG-Gemin2 or the parental cell line (Tet-ON), by using the high-salt wash procedure as described (*23*, *26*). Bound proteins were eluted with 3× FLAG peptide and analyzed by SDS-PAGE and silver staining. Molecular weight markers and known proteins are indicated. (**B**) Total snRNP proteins (TPs) were purified as described and analyzed by SDS-PAGE and silver staining (*7*, *26*). Molecular weight markers and known proteins are indicated. (**C**) The SMN complex inhibits TPs assembly on all U snRNAs. In vitro–transcribed [^{32}P]UTP-labeled U snRNAs were incubated with purified SMN complexes (SMN) and/or TPs as indicated for 1 hour at 30°C. The reactions were analyzed by electrophoresis on native polyacrylamide gels. Known RNP complexes are indicated. (**D**) The SMN complex prebound to Sm proteins mediates snRNP-core assembly. TPs (lane 1) or buffer were incubated with purified SMN complexes bound to anti-FLAG beads for 1 hour at 4°C. Following extensive washes, SMN complexes incubated in the absence (lane 2) or in the presence (lane 3) of TPs were eluted with 3× FLAG peptide and analyzed by SDS-PAGE and silver staining. (**E**) In vitro–transcribed [^{32}P]UTP-labeled U1 snRNA was incubated with the same proteins as in (D) for 1 hour at 30°C. The reactions were analyzed by electrophoresis on native polyacrylamide gels. (**F**) In vitro–transcribed [^{32}P]UTP-labeled U snRNAs were incubated with either TPs or SMN complexes prebound to TPs [as in (D) lanes 1 and 3, respectively] for 1 hour at 30°C. The reactions were analyzed by electrophoresis on native polyacrylamide gels. Known RNP complexes are indicated.

formation of the snRNP core or causes its disassembly. Once the Sm cores form by first incubating U snRNAs and TPs, addition of SMN complexes to the assembly reaction does not dissociate them (*27*). The Sm core also did not form when SMN complexes were preincubated with U snRNAs before TPs were added to the reaction (*27*). We then tested the possibility that Sm core assembly requires the interaction of the SMN complex with the Sm proteins before that with U snRNAs. Purified SMN complexes bound to anti-FLAG beads were preincubated with either buffer or an excess of TPs to allow Sm protein binding to the SMN complex. Unbound proteins were washed away, and the SMN complexes were eluted with FLAG peptides and analyzed by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). Immobilized SMN complexes bound efficiently all Sm proteins and some U1A but not U170K (Fig. 2D). None of these proteins bound to beads containing nonspecific proteins as a control (*27*). The proteins shown in Fig. 2D were tested for snRNP core assembly in vitro. SMN complexes prebound to Sm proteins were able to mediate Sm core formation on U snRNAs (Fig. 2, E and F). SMN complexes prebound to Sm proteins formed several U1A- and Sm core–containing complexes on U1 snRNA (Fig. 2E). Consistent with the lack of U1 70K association with the SMN complex, the corresponding complex was not formed. Therefore, in a purified in vitro system, such as in HeLa extracts (fig. S2), only Sm proteins bound to SMN complexes can mediate snRNP assembly on U snRNAs. These results indicate that the SMN complex imposes an ordered pathway of snRNP assembly that requires the association of Sm proteins with the SMN complex before recruitment of U snRNAs.

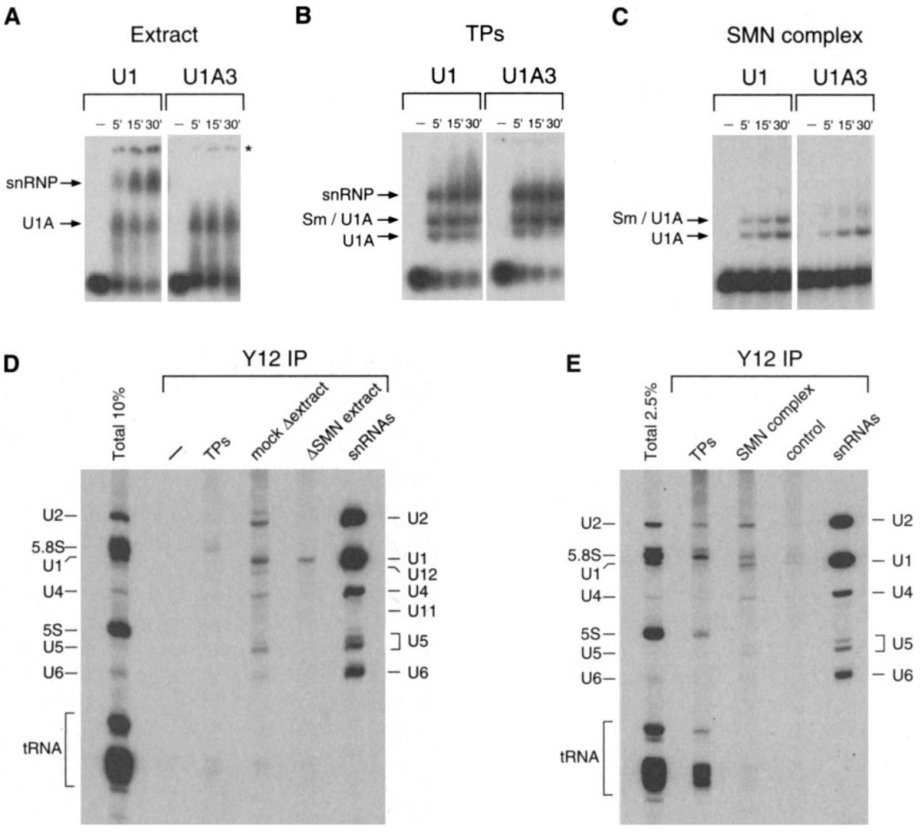

Fig. 3. The SMN complex is a specificity factor essential for snRNP assembly. (**A**) Sm core assembly of U1A3, an U1 snRNA mutant that does not interact with the SMN complex, is impaired in HeLa extracts. In vitro–transcribed [^{32}P]-labeled U1 or U1A3 snRNAs (*28*) were incubated with HeLa extracts at 30°C for the indicated times and analyzed by electrophoresis on native polyacrylamide gels. Known RNP complexes are indicated. The asterisk marks complexes that migrate close to the origin of the gel and may contain both U snRNAs and the SMN complex. (**B**) Purified snRNP proteins (TPs) do not discriminate between wild-type and mutant U1 snRNAs. In vitro–transcribed [^{32}P]UTP-labeled U1 or U1A3 snRNAs were incubated with TPs at 30°C for the indicated times and analyzed by electrophoresis on native polyacrylamide gels. (**C**) Impaired Sm core assembly of U1A3 by the SMN complex. In vitro–transcribed [^{32}P]UTP-labeled U1 or U1A3 snRNAs were incubated with SMN complexes purified as in Fig. 1A at 30°C for the indicated times and analyzed by electrophoresis on native polyacrylamide gels. [(D) and (E)] (**D**) The SMN complex provides the specificity for the efficient and correct assembly of U snRNAs. 10^6 cpm of [^{32}P]-labeled total RNA from HeLa cells (*26*) were incubated with buffer (–), TPs, and either mock- or SMN complex–depleted (Δ) HeLa extracts for 1 hour at 30°C. Reaction mixtures were immunoprecipitated with anti-Sm (Y12) monoclonal antibody, and bound RNAs were analyzed by electrophoresis on a polyacrylamide/urea gel (*26*). The total represents 10% of the input RNAs. U snRNAs isolated by immunoprecipitations with Y12 from [^{32}P]phosphoric acid–labeled HeLa extracts are also shown as a reference (snRNAs). Known RNAs are indicated. (**E**) We incubated 10^6 cpm of [^{32}P]-labeled total RNA from HeLa cells with TPs, SMN complexes (SMN), or nonspecific proteins (control) purified as in Fig. 1A for 1 hour at 30°C. Reaction mixtures were immunoprecipitated with anti-Sm (Y12) monoclonal antibody, and bound RNAs were analyzed by electrophoresis on a polyacrylamide/urea gel. Total represents 2.5% of the input RNAs. Panel (E) had a threefold longer exposure than that of (D).

To further investigate the difference between snRNP assembly mediated by the SMN complex and that of purified snRNP proteins, we analyzed the kinetics of Sm core assembly on U1 and U1A3, a stem-loop 1 mutant that is impaired in binding to the SMN complex (*28*). HeLa extracts efficiently assembled the Sm core on U1 but not on U1A3 (Fig. 3A). U1A, which binds stem-loop 2 independent of the SMN complex, associated with both U1 and U1A3. These results indicate an absolute requirement for SMN complex interaction with U1 for snRNP assembly. In strong contrast, TPs do not discriminate between U1 and U1A3 because they assemble on both RNAs efficiently and with much faster kinetics than that of HeLa extracts (Fig. 3B). To test directly whether the SMN complex is responsible for the discrimination between U1 and U1A3 observed in HeLa extracts, we carried out assembly reactions with purified SMN complexes containing Sm proteins. Similar to HeLa extracts, the SMN complex mediated Sm core assembly on U1 but not on U1A3 (Fig. 3C). These results point to a mechanistic difference in the assembly of snRNPs mediated by the SMN complex and by purified Sm proteins. The direct interaction of the SMN complex with U1 is required for U1 snRNP assembly and allows the discrimination between wild-type and mutant U1 snRNAs that do not bind SMN.

Several studies have demonstrated that Sm proteins are an evolutionarily conserved class of RNA binding proteins with high affinity for oligo-uridine stretches and form a thermodynamically stable heptameric core on the Sm site (*7–9*, *29–34*). To explain why cells need the SMN complex if Sm proteins can spontaneously assemble Sm cores, we reasoned that in the complex cellular environment the affinity of Sm proteins for the Sm site may not be sufficient to ensure their specific binding only to the correct target RNAs. To analyze Sm core formation on U snRNAs in a complex mixture of RNAs, we incubated [^{32}P]-labeled total RNA from HeLa cells with HeLa extracts, TPs, or purified SMN complexes under snRNP assembly conditions. Immunoprecipitations using anti-Sm antibodies were used to retrieve RNAs that associated with Sm proteins. As expected, HeLa extracts specifically assembled Sm cores

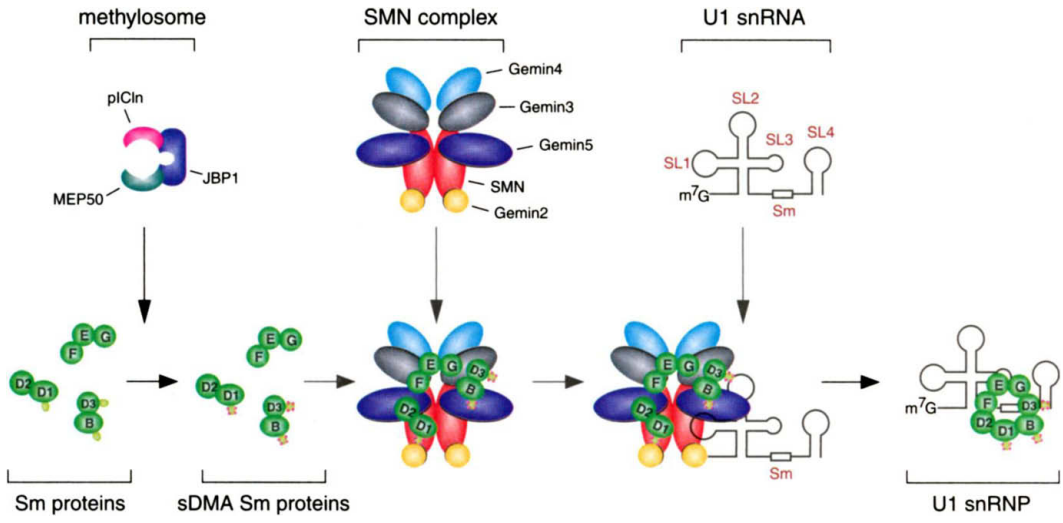

Fig. 4. A schematic model depicting the ordered pathway of snRNP assembly mediated by the SMN complex. Newly translated Sm proteins are captured by the 6*S* pICln complex (not shown) and by the 20*S* methylosome which modifies specific arginines in the RG domains of several of the Sm proteins to symmetrical di-methylarginines (sDMA) (*41*–*43*). This converts the modified proteins to high-affinity substrates for the SMN complex. Whereas SMN itself interacts directly with the sDMA-modified domains of the Sm proteins (*44*, *45*), the Sm-SMN complex interaction is further stabilized by the interaction of several of the other protein components of the SMN complex, including Gemin3, Gemin4, Gemin5, Gemin6, and Gemin7, with Sm proteins (*20*–*24*). For simplicity, Gemin6, Gemin7, and the U1 snRNP–specific proteins U1A, U170K, and U1C are not shown. The association of Sm proteins with the SMN complex precedes the direct binding of the SMN complex to specific domains of U snRNAs (*28*, *35*). As depicted, stem and loop 1 of U1 snRNA is necessary and sufficient for the association with the SMN complex (*28*). The SMN complex with bound Sm proteins is the active form for snRNP assembly, and this indeed accumulates in the cytoplasm poised to associate with newly exported U snRNAs and to mediate assembly of the Sm core.

on U1, U2, U4, and U5 as well as on the snRNA components of the minor spliceosome, U11 and U12 (Fig. 3D). The assembly requires the SMN complex because HeLa extracts depleted of SMN fail to assemble Sm cores on U snRNAs. Notably, purified Sm proteins (TPs) display neither efficient nor specific assembly with U snRNAs, because they associate also with other RNAs including 5*S* and 5.8*S* ribosomal RNAs and transfer RNAs (Fig. 3, D and E). In contrast, purified SMN complexes containing Sm proteins mediate the specific assembly of Sm cores on U1, U2, U4, and U5 snRNAs (Fig. 3E). Together, these results demonstrate that the SMN complex is essential for the efficient and specific assembly of Sm proteins on Sm site–containing U snRNAs.

In vitro assembly experiments have shown that purified Sm proteins can form Sm cores not only on full-length U snRNAs but also on short oligonucleotides that compose the Sm site sequence (*8*). In fact, the only requirement seems to be that the oligonucleotides be relatively rich in uridines (*34*). This indicates that the Sm proteins have a high propensity for self-association and this readily manifests itself without much specificity in respect to the RNA substrate. Thus, in physiological systems, there would be a need to control the Sm proteins and prevent their illicit assembly on RNAs other than the specific U snRNA targets. Adventitious binding of Sm proteins to RNAs would be deleterious to cells, because it would interfere with the functions of these RNAs. Our findings suggest that a key function of the SMN complex is to ensure the specificity of interaction of Sm proteins and restrict it exclusively to intended RNAs. To accomplish this, the SMN complex binds the Sm proteins so that there is no free pool of Sm proteins competent for assembly outside of the SMN complex, and it provides stringent control over RNA target selection by binding directly to specific sequences that are found in U snRNAs (*28*, *35*). This ensures the delivery of Sm proteins for assembly only on the appropriate RNAs, and the assembly can then occur on the adjacent Sm site through the intrinsic capacity of the Sm proteins for self-association. In addition to its critical role in ensuring the specificity of snRNP assembly, the SMN complex also likely enhances the efficiency of the Sm core assembly process by reducing the dimensionality of diffusion of the newly synthesized Sm proteins and the U snRNAs in the complex microenvironment of the cell. Our view of the snRNP assembly pathway is depicted in Fig. 4. This model also illustrates the ordered steps in the pathway suggested by the in vitro assembly experiments with purified SMN complexes.

The observations we report here on the function of the SMN complex as an ATP-dependent macromolecular machine (an "assemblyosome") responsible for the assembly of snRNPs provoke several thoughts on the parallels between this complex and other cellular machineries and on the molecular events that may underlie the pathology of SMA. A strong similarity can be drawn between the function of the SMN complex in the assembly of large RNP complexes and the role of chaperones in protein folding. Although the three-dimensional folding of proteins is specified by their amino acid sequence, this information is not sufficient to ensure that protein folding will occur in cells fast enough and with complete accuracy (*36*, *37*). In much the same way as protein folding can take place unaided in vitro, in cells this process uses molecular chaperones to assist proteins to rapidly attain their correct native conformations. Indeed, biological systems invest considerable energy to facilitate protein folding and prevent the accumulation of misfolded proteins, because these proteins could have undesirable properties and may tend to form aggregates that are harmful to cells. We envision that the SMN complex plays a similar role for large RNP complexes in cells. Protein misfolding, like the presence of expanded polyglutamine stretches in proteins, can result in formation of microaggregates and lead to neurodegenerative diseases (*38*–*40*). The molecular consequences of reduced SMN levels in motor neurons, as is the case in SMA patients, are not known. It is possible that deficiency in SMN results in less efficient assembly of snRNPs and other RNPs, but it is also possible that it leads to formation of aberrant RNPs as a result of nonspecific binding of Sm proteins to critical RNAs and that this contributes to the pathophysiology of SMA.

References and Notes

1. C. L. Will, R. Luhrmann, *Curr. Opin. Cell Biol.* **13**, 290 (2001).
2. I. W. Mattaj, E. M. De Robertis, *Cell* **40**, 111 (1985).
3. I. W. Mattaj, *Cell* **46**, 905 (1986).
4. U. Fischer, R. Luhrmann, *Science* **249**, 786 (1990).
5. J. Hamm, E. Darzynkiewicz, S. M. Tahara, I. W. Mattaj, *Cell* **62**, 569 (1990).
6. U. Fischer, V. Sumpter, M. Sekine, T. Satoh, R. Luhrmann, *EMBO J.* **12**, 573 (1993).
7. V. Sumpter, A. Kahrs, U. Fischer, U. Kornstadt, R. Luhrmann, *Mol. Biol. Rep.* **16**, 229 (1992).
8. V. A. Raker, K. Hartmuth, B. Kastner, R. Luhrmann, *Mol. Cell. Biol.* **19**, 6554 (1999).
9. V. A. Raker, G. Plessel, R. Luhrmann, *EMBO J.* **15**, 2256 (1996).
10. A. M. Kleinschmidt, J. R. Patton, T. Pederson, *Nucleic Acids Res.* **17**, 4817 (1989).
11. G. Meister, D. Buhler, R. Pillai, F. Lottspeich, U. Fischer, *Nature Cell Biol.* **3**, 945 (2001).
12. U. Fischer, Q. Liu, G. Dreyfuss, *Cell* **90**, 1023 (1997).

13. L. Pellizzoni, N. Kataoka, B. Charroux, G. Dreyfuss, *Cell* **95**, 615 (1998).
14. D. Buhler, V. Raker, R. Luhrmann, U. Fischer, *Hum. Mol. Genet.* **8**, 2351 (1999).
15. S. Lefebvre *et al.*, *Cell* **80**, 155 (1995).
16. A. H. Burghes, *Am. J. Hum. Genet.* **61**, 9 (1997).
17. J. Melki, *Curr. Opin. Neurol.* **10**, 381 (1997).
18. Q. Liu, U. Fischer, F. Wang, G. Dreyfuss, *Cell* **90**, 1013 (1997).
19. Q. Liu, G. Dreyfuss, *EMBO J.* **15**, 3555 (1996).
20. B. Charroux *et al.*, *J. Cell Biol.* **147**, 1181 (1999).
21. B. Charroux *et al.*, *J. Cell Biol.* **148**, 1177 (2000).
22. A. K. Gubitz *et al.*, *J. Biol. Chem.* **277**, 5631 (2002).
23. L. Pellizzoni, J. Baccon, J. Rappsilber, M. Mann, G. Dreyfuss, *J. Biol. Chem.* **277**, 7540 (2002).
24. J. Baccon, L. Pellizzoni, J. Rappsilber, M. Mann, G. Dreyfuss, *J. Biol. Chem.* **277**, 31957 (2002).
25. S. Paushkin, A. K. Gubitz, S. Massenet, G. Dreyfuss, *Curr. Opin. Cell Biol.* **14**, 305 (2002).
26. Materials and methods are available as supporting material on *Science* Online.
27. L. Pellizzoni *et al.*, data not shown.
28. J. Yong, L. Pellizzoni, G. Dreyfuss, *EMBO J.* **21**, 1188 (2002).
29. B. Seraphin, *EMBO J.* **14**, 2089 (1995).
30. I. Toro *et al.*, *EMBO J.* **20**, 2293 (2001).
31. J. Salgado-Garrido, E. Bragado-Nilsson, S. Kandels-Lewis, B. Seraphin, *EMBO J.* **18**, 3451 (1999).
32. Z. Palfi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8967 (2000).
33. H. Urlaub, V. A. Raker, S. Kostka, R. Luhrmann, *EMBO J.* **20**, 187 (2001).
34. T. Achsel, H. Stark, R. Luhrmann, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3685 (2001).
35. J. Yong, L. Pellizzoni, G. Dreyfuss, in preparation.
36. B. Bukau, A. L. Horwich, *Cell* **92**, 351 (1998).
37. F. U. Hartl, M. Hayer-Hartl, *Science* **295**, 1852 (2002).
38. A. L. Horwich, J. S. Weissman, *Cell* **89**, 499 (1997).
39. P. Kazemi-Esfarjani, S. Benzer, *Science* **287**, 1837 (2000).
40. C. J. Cummings *et al.*, *Hum. Mol. Genet.* **10**, 1511 (2001).
41. W. J. Friesen *et al.*, *Mol. Cell. Biol.* **21**, 8289 (2001).
42. G. Meister *et al.*, *Curr. Biol.* **11**, 1990 (2001).
43. W. J. Friesen *et al.*, *J. Biol. Chem.* **277**, 8243 (2002).
44. W. J. Friesen, S. Massenet, S. Paushkin, A. Wyce, G. Dreyfuss, *Mol. Cell* **7**, 1111 (2001).
45. H. Brahms, L. Meheus, V. de Brabandere, U. Fischer, R. Luhrmann, *RNA* **7**, 1531 (2001).
46. We thank members of our laboratory, in particular A. Gubitz, J. Baccon, and S. Paushkin, for stimulating discussions and for critical comments on this manuscript. We are grateful to J. Steitz for the anti-Sm Y12 monoclonal antibody. Supported by the Association Francaise Contre les Myopathies and by a grant from NIH. G.D. is an investigator of the Howard Hughes Medical Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/298/5599/1775/DC1
SOM Text
Figs. S1 and S2
References and Notes

11 June 2002; accepted 1 October 2002

Truncating Neurotrypsin Mutation in Autosomal Recessive Nonsyndromic Mental Retardation

Florence Molinari,[1] Marlène Rio,[1] Virginia Meskenaite,[2] Férechté Encha-Razavi,[1] Joelle Augé,[1] Delphine Bacq,[3] Sylvain Briault,[4] Michel Vekemans,[1] Arnold Munnich,[1] Tania Attié-Bitach,[1] Peter Sonderegger,[2] Laurence Colleaux[1]*

A 4–base pair deletion in the neuronal serine protease neurotrypsin gene was associated with autosomal recessive nonsyndromic mental retardation (MR). In situ hybridization experiments on human fetal brains showed that neurotrypsin was highly expressed in brain structures involved in learning and memory. Immunoelectron microscopy on adult human brain sections revealed that neurotrypsin is located in presynaptic nerve endings, particularly over the presynaptic membrane lining the synaptic cleft. These findings suggest that neurotrypsin-mediated proteolysis is required for normal synaptic function and suggest potential insights into the pathophysiological bases of mental retardation.

Moderate to severe mental retardation (IQ<50) affects 0.3 to 0.8% of the population and its prevalence increases up to 2% if mild MR is included (50<IQ<70) (*1*). The possible causes of MR are diverse and include environmental factors or teratogens, chromosomal anomalies and metabolic diseases impairing neuronal function or brain patterning during development (*2*).

By contrast, MR with apparently normal brain development and no other clinical features (i.e., nonsyndromic MR) represents the most common cognitive dysfunction and remains poorly understood. In fact, the broad genetic heterogeneity of MR and the scarcity of large pedigrees have hitherto hampered linkage analyses in nonsyndromic MR. Indeed, although 10 nonsyndromic X-linked MR genes have been found (*3–12*), none of the genes causing nonsyndromic autosomal recessive MR has been identified so far.

We have investigated a sibship of four mentally retarded (three girls and one boy) and four healthy children born to first cousin Algerian parents (family 1, fig. S1A). Cognitive impairment and a low IQ (below 50) were consistent features in the four affected sibs. A genome-wide screen that used 400 microsatellite markers identified a single region of shared homozygosity on chromosome 4q24–q25 in the four affected sibs. Further marker typing and haplotype analysis reduced the genetic interval to 18 cM (14 Mb) between markers D4S1570 and D4S1575 (Z_{max} = 3.33, at θ = 0 at D4S407 locus) (fig. S1A). This interval encompasses about 29 genes of known function including the *DKK2*, *RPL34*, *CASP6*, *ANK2*, *CAMK2D*, *TRPC3*, and *PRSS12* genes (*13–19*) (fig. S1B).

Neurotrypsin (PRSS12) is a secreted protein of 875 amino acids, belonging to the subfamily of trypsin-like serine proteases. This gene has been identified in a polymerase chain reaction (PCR)–based search for serine proteases that are expressed in the brain (*20*). Neurotrypsin was regarded as a strong candidate gene by both position and function, as expression studies in mouse embryos suggested its likely involvement in synapse maturation and neural plasticity (*21*). The genomic structure of the human neurotrypsin gene was identified by alignment of the cDNA sequence (GenBank accession number: NM_003619) with the sequence of PAC clone 265B3 (GenBank accession number: AC073025). Thirteen exons were identified, and primers were designed for direct sequencing of all coding exons and exon-intron junctions (*22*).

Homozygosity for a 4–base pair (bp) deletion in exon 7 of the *PRSS12* gene was detected in patient II.8 (fig. S2). This deletion disrupted an Aat II restriction site and resulted in a premature stop codon 147 nucleotides downstream of the deletion. Restriction analyses showed cosegregation of the mutation with the disease in all affected individuals, and that both parents were heterozygous (Fig. 1A). Reverse transcription PCR (RT-PCR) analysis of lymphoblastoid cell line RNAs from patient II.8 revealed a slightly reduced amount of mutant neurotrypsin transcript carrying the deletion (*23*). This mutation was not found in 200 unrelated control individuals (100 from Maghrebian and 100 from various ethnic origins).

To estimate the frequency of neurotrypsin mutations, the coding sequence of the gene was screened by DHPLC and direct DNA sequencing in 17 nonsyndromic inbred MR families and 23 mentally retarded children. The same 4-bp deletion was found in another child born to first cousin Algerian parents, apparently unrelated to family 1 but originating from the same area of Eastern Algeria. The course of the disease in this child was very similar to that of affected indi-

[1]Unité de Recherches sur les Handicaps Génétiques de l'Enfant, INSERM U-393, et Département de Génétique, Hôpital Necker-Enfants Malades, Paris, France. [2]Institute of Biochemistry, University of Zurich, Switzerland. [3]Centre National de Génotypage, Evry, France. [4]Service de Génétique, CHU de Tours, Hôpital Bretonneau, Tours, France.

*To whom correspondence should be addressed. E-mail: colleaux@necker.fr