

In vivo assembly of functional U7 snRNP requires RNA backbone flexibility within the Sm-binding site

Nikolay G Kolev & Joan A Steitz

Most histone precursor mRNAs (pre-mRNAs) in metazoans are matured by 3'-end cleavage directed by the U7 small nuclear ribonucleoprotein (snRNP). RNA functional groups necessary for *in vivo* assembly and activity of the U7 snRNP were examined by nucleotide-analog interference mapping and mutagenesis using a chimeric mouse histone H4 pre-mRNA–U7 snRNA construct that is cleaved *in cis* in *Xenopus laevis* oocytes. Assembly of the unique U7 Sm protein core is rate limiting for processing *in vivo* and requires four conserved nucleotides within the U7 Sm-binding site, as well as the correct positioning and size of the U7 terminal stem-loop structure. To our surprise, pseudouridine substitution revealed a requirement for backbone flexibility at a particular position within the U7 Sm site, providing *in vivo* biochemical evidence that an unusual C2'-endo sugar conformation is necessary for assembly of the Sm ring.

U7 is a metazoan small nuclear RNA (snRNA) necessary for the 3'-end formation of replication-dependent histone mRNAs^{1–3}. It ranges in length from 57 nucleotides (nt) in sea urchins¹ to 74 nt in *Drosophila melanogaster*³ and has a single secondary structure feature, a hairpin at the 3' end. U7 follows the same maturation pathway as U1, U2, U4 and U5 spliceosomal snRNAs. After transcription, it is exported to the cytoplasm and assembles with seven Sm and Sm-like core proteins in an ATP-dependent process involving the survival of motor neuron (SMN) and the protein arginine methyltransferase-5 (PRMT5) complexes^{4–6}. The U7 RNA then undergoes cap hypermethylation, trimming of its 3' end and reimport into the nucleus⁷.

The noncanonical Sm-binding site of U7 snRNA determines the assembly of a unique core of Sm proteins: SmB/B', SmD3, SmE, SmF and SmG are shared with spliceosomal snRNPs, whereas SmD1 and SmD2 are replaced by the Sm-like proteins Lsm10 and Lsm11 (refs. 4,5,8). The unique Sm core determines the low abundance of the U7 snRNP⁹, its cytoplasmic assembly and nuclear import at rates slower than those of spliceosomal snRNPs^{7,9} and its localization to Cajal bodies¹⁰.

The pre-mRNAs transcribed from the replication-dependent histone genes of metazoans are intronless. They undergo endonucleolytic cleavage in the 3' untranslated region^{5,11} without polyadenylation. The signals for this maturation process are a conserved stem-loop upstream and a histone downstream element (HDE) 10 or 11 nucleotides downstream of the cleavage site. The HDE sequence is complementary to the 5' end of U7 snRNA, and a requirement for intermolecular base pairing has been demonstrated by complementation of HDE mutants that do not support processing with compensatory changes in the 5' end of U7 (refs. 12,13). A stem-loop-binding protein (SLBP) coordinates the high expression of histone messages

during the S phase of the cell cycle and enhances recruitment of U7 snRNPs to the pre-mRNAs¹⁴. Recent evidence implicates components of the cleavage and polyadenylation machinery, namely symplekin¹⁵ and the cleavage and polyadenylation specificity factor 73-kDa subunit (CPSF-73)^{15,16}, as intimately involved in the processing event.

We set out to identify RNA determinants crucial for the *in vivo* assembly and activity of the U7 snRNP using a chimeric RNA construct that contains a processing substrate covalently attached to the mouse U7 snRNA sequence. A similar construct, when injected into the cytoplasm of *Xenopus* oocytes, has been shown to assemble into a processing-competent RNP that directs cleavage *in cis* of the attached pre-mRNA sequence¹⁷. U7 is therefore the only Sm snRNP in metazoan cells that can easily be subjected to an analysis that couples assembly to function. Pseudouridine substitution at a particular position within the U7 Sm-binding site produced unexpected inhibition of RNP assembly, revealing a requirement for flexibility of the sugar-phosphate backbone. Our findings provide *in vivo* evidence for the importance of a C2'-endo sugar conformation to achieve stable assembly with the Sm core proteins.

RESULTS

Pseudouridine at position 24 in U7 impairs RNP assembly

Nucleotide-analog interference mapping (NAIM) is a powerful technique that identifies individual functional groups that are crucial for the activity of an RNA molecule^{18,19}. The method relies on selection of functional species within a pool of RNA molecules containing chemical substitutions at single, random positions. A chimera consisting of the mouse histone H4-12 pre-mRNA and the mouse U7 snRNA (Fig. 1) is perfectly suited for *in vivo* NAIM²⁰, as it undergoes *in cis* cleavage in the cytoplasm of *Xenopus* oocytes¹⁷, generating shorter processed RNAs.

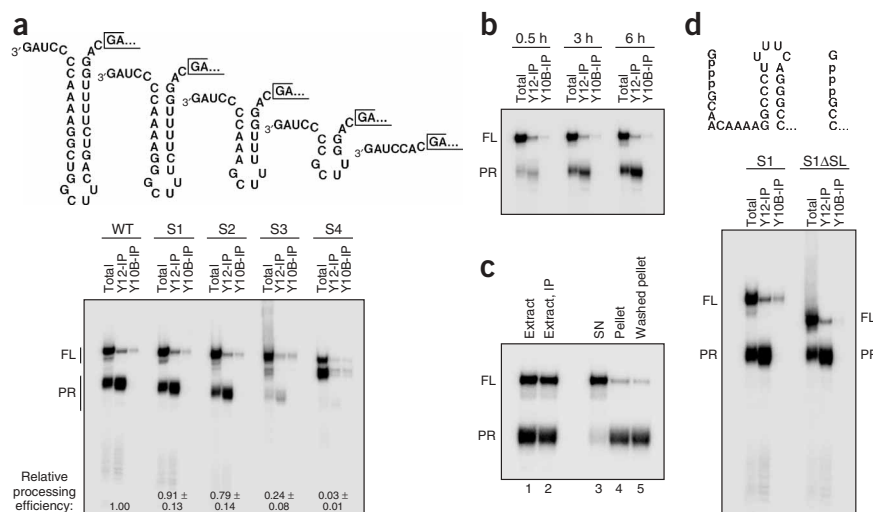
Howard Hughes Medical Institute, Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06536, USA. Correspondence should be addressed to J.A.S. (joan.steitz@yale.edu).

Received 6 July 2005; accepted 17 February 2006; published online 19 March 2006; doi:10.1038/nsmb1075



Figure 3 The Y12 immunoprecipitation (IP) assay simultaneously assesses U7 Sm core assembly and processing efficiency of the chimera.

(a) Shortening the U7 hairpin to less than 5 bp impairs snRNP assembly and processing. Total (33%) and immunoprecipitated (100%) chimeras containing the U7 hairpins shown were separated on a denaturing gel. Processing efficiencies (data from two independent experiments, with \pm values indicating the range) are relative to the wild type. FL, full-length chimera; PR, processed U7-containing product. Some nonspecific shortening of the injected RNA was observed, especially in the case of S4, most probably because of exonucleolytic 3'-end trimming. The bands appearing as a ladder near the bottom of the gel (for WT, S1 and S2) correspond to processed chimeric RNAs that are 5'-exonucleolytically trimmed to the 5' end of the Sm-binding site of the chimera (data not shown). (b) S1 chimera was incubated in oocytes for shorter time periods and subjected to the Y12 IP assay. (c) S1 chimera was incubated overnight in oocytes and extract was prepared and treated as follows: lane 1, total extract treated with phenol to isolate RNA; lane 2, total extract subjected to Y12 IP before extraction with phenol; lane 3, supernatant after Y12 IP; lane 4, pellet after Y12 IP; lane 5, pellet after Y12 IP, washed four times with TBS. (d) Efficient processing of the chimera does not require the histone stem-loop. S1 and S1 Δ SL RNAs were assayed by Y12 IP. Shown above gel are the 5' ends of the two constructs.



series of immunoprecipitations using the Y12 antibody. Capped internally labeled mutant chimeric RNAs were microinjected into the cytoplasm of oocytes, and the ability of full-length and processed RNAs from each pool of oocytes to be immunoprecipitated with Y12 versus the control Y10B antibody to ribosomal RNA²¹ was compared.

We established (Fig. 3a) that shortening the U7 hairpin, previously shown to be a structural feature of the RNA that tolerates sequence changes²³, to 8 or 5 base pairs (bp) (S1 and S2, respectively) has only a modest inhibitory effect on assembly and subsequent processing (resulting in efficiencies of 0.91 and 0.79, respectively, relative to the wild-type chimera). Truncation of an additional 3 bp (S3), leaving only an unstable 2-bp hairpin, reduced processing efficiency to \sim 25%, and complete removal of the U7 hairpin (S4) yielded only trace levels of cleaved product.

Notably, the amount of RNA immunoprecipitated with Y12 was always directly proportional to the efficiency of processing, and the vast majority of precipitated RNA was in its processed form. This guarantees that the immunoprecipitation specifically assessed the assembly of functional snRNPs (rather than detecting randomly bound Sm proteins) and suggests that snRNP assembly is rate limiting for *in vivo* processing of the chimera. To validate this conclusion, we checked whether the overnight incubation (\sim 16 h) in oocytes might be too long to capture snRNP particles that were newly assembled but had not undergone processing; incubating the chimera in oocytes for much shorter time periods before the Y12 immunoprecipitation yielded no increase in the amount of precipitated full-length RNA (Fig. 3b). The cleavage of full-length RNAs during the immunoprecipitation can also be ruled out, as extracts subjected to Y12 immunoprecipitation conditions (incubation with the resin-attached antibody without separation of pellet and supernatant) resulted in no increase in the fraction of processed RNA compared to extracts that were immediately deproteinized (Fig. 3c, lanes 1 and 2). Together, these observations demonstrate the lack of an appreciable lag between Sm core assembly and processing of the chimeric RNAs *in vivo*.

As the highly conserved histone stem-loop is the binding site for SLBP, which facilitates recruitment of the U7 snRNP¹⁴, we asked

whether this stem-loop is needed for efficient cleavage *in cis* of the histone pre-mRNA–U7 snRNA chimera. In the Y12 immunoprecipitation (Fig. 3d), deletion of the stem-loop in S1 Δ SL resulted in only a small decrease in the efficiency of processing (0.87 relative to S1). This suggests that the histone stem-loop and SLBP do not play an essential role in the *in vivo* cleavage of our chimera in *Xenopus* oocytes.

Using the S2 chimera, we created a series of mutants with deletions in the sequence that separates the U7 hairpin (here, 5 bp) from the Sm site (Fig. 4) and found that spacing is also important for assembly of functional RNP. Whereas removal of the 3' single-stranded tail of the chimera (Fig. 4b) had no adverse effect on assembly and processing,

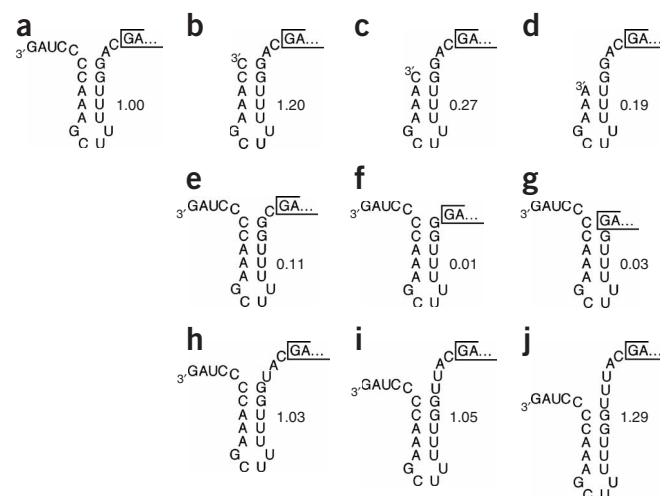


Figure 4 Spacing between the U7 terminal stem and the Sm site affects U7 snRNP assembly and processing. (a–j) The S2 chimera (a) and mutants whose 3' portions are shown (b–j) were used to determine relative processing efficiencies indicated (fraction of RNA processed compared to the S2 chimera). Y12 immunoprecipitability was directly proportional to processing efficiency.

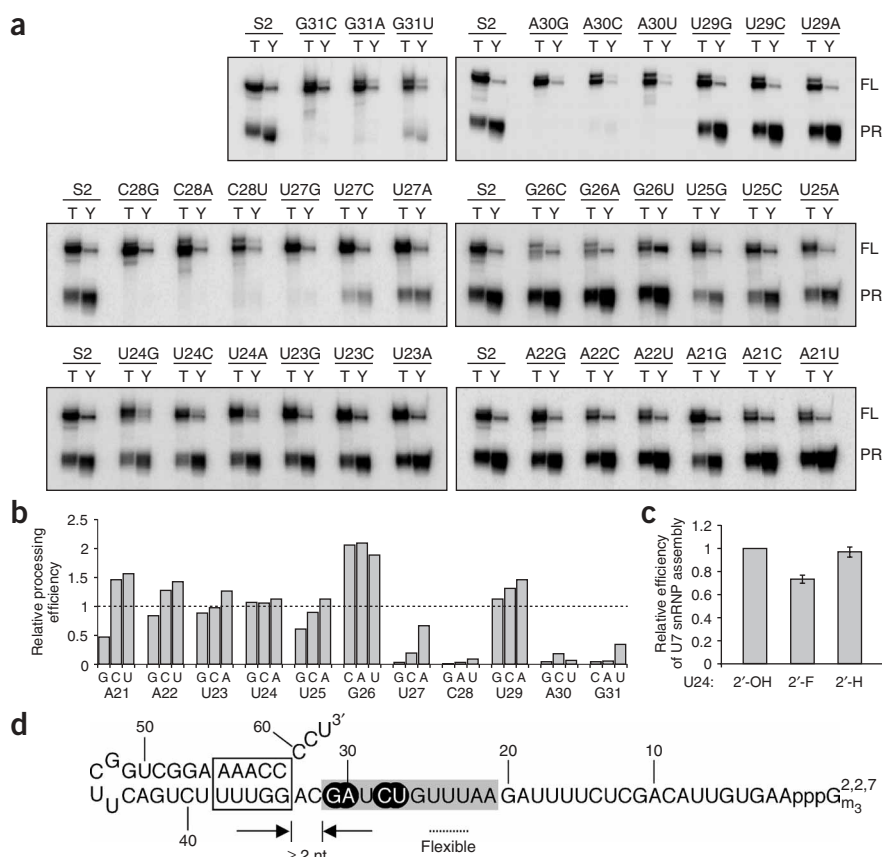


Figure 5 Effects of single-base substitutions within the Sm-binding site on assembly and processing. **(a)** Y12 IP assays of S2 chimeric RNAs with the indicated base substitutions. T, 33% of total RNA isolated from the oocytes; Y, RNA immunoprecipitated by Y12; FL, full-length chimera; PR, processed chimera. **(b)** Processing efficiency of each mutant, relative to construct S2. **(c)** Assembly of singly substituted U7 snRNAs, assayed by Y12 IP. S2 U7 RNAs were prepared by two-piece ligation using ^{32}P -phosphorylated synthetic oligoribonucleotides having a 2'-hydroxyl (2'-OH), 2'-fluoro (2'-F) or 2'-deoxy (2'-H) substituent at nucleotide U24 (see **Supplementary Methods**). Their relative assembly efficiencies (from four experiments, with s.e. shown) were calculated after PhosphorImager quantification of band intensities for total and Y12-bound RNAs compared to the U24 2'-OH. **(d)** Crucial determinants for *in vivo* assembly of the U7 snRNP. Shown is the sequence and secondary structure of the mouse U7 snRNA⁴³. Gray box, Sm-binding site; black circles, crucial nucleotides within the Sm site; open box, required 3' hairpin structure. The minimal distance of 2 nt between the Sm site and the hairpin is indicated, as is the proposed requirement for structural flexibility of the uridines in the 5' portion of the Sm site.

deletion of one or two additional nucleotides (**Fig. 4c,d**) severely impaired function, most probably by destabilizing the U7 hairpin. Deletion of A33, C32 A33 or G31 C32 A33 (**Fig. 4e-g**) resulted in a large decrease in the efficiency of assembly and processing of the chimera, whereas insertion of 1, 2 or 3 uridines between A33 and G34 (**Fig. 4h-j**) had no adverse effect. We conclude that a minimal spacing of 2 nt is needed between the U7 Sm-binding site and the 3' hairpin.

Four bases within the U7 Sm-binding site are crucial

Changing the unique U7 Sm-binding site, 5'-AAUUUGUCUAG-3', into a canonical spliceosomal Sm site, 5'-AAUUUUUGGAG-3' (U7 OPT), results in the assembly of an snRNP that is nonfunctional in 3'-end processing; a mutant with multiple substitutions within the Sm site (5'-AACGCGUCAUG-3') cannot assemble an snRNP at all^{7,17}. We analyzed the contribution of each nucleotide in the U7 Sm site to the assembly and processing of the RNA chimera by mutating every position to the other three possible bases.

The identities of only four bases, which reside in the 3' half of the Sm site, are crucial for assembly of an active U7 snRNP (**Fig. 5a,b**). Any substitution of C28, A30 or G31 resulted in pronounced inhibitory effects, whereas U27A alteration was partially deleterious. Notably, the loss of immunoprecipitation by the Y12 antibody indicates a failure of Sm proteins to assemble at all, rather than a simple exchange of the U7-specific proteins Lsm10 and Lsm11 for the canonical SmD1 and SmD2.

To our surprise, U24 could be replaced individually with any other unmodified base with little effect on assembly or processing of the chimera (**Fig. 5b**). This contrasts with the strong inhibitory effect of the U24 Ψ substitution revealed by NAIM (**Fig. 2**). Substitutions with G at A21, U25 and other positions within the 5' end of the U7 Sm site

were somewhat inhibitory. All mutations of G26 (which is not conserved between different species; **Supplementary Fig. 1** online) resulted in a two-fold increase in assembly and processing, possibly reflecting a minimized potential for base pairing between the Sm site and the region immediately downstream of the cleavage site in the chimera, which would otherwise interfere with RNP assembly on the single-stranded Sm site. Contrary to our expectation, the A30U mutation, even though it creates a *Drosophila* U7 Sm-binding site sequence³, yielded a chimeric RNA that was inactive in *Xenopus* oocytes (**Fig. 5a,b**).

A required backbone conformation at U24

We further investigated the seemingly contradictory finding that pseudouridine at position 24 in the U7 Sm site is deleterious to active U7 snRNP assembly, whereas substitution with any canonical base is not. Pseudouridine preserves the Watson-Crick base pairing potential of U, but it confers structural rigidity on the sugar-phosphate backbone. This has been attributed to its ability to coordinate a water molecule between the N1-H, the 5'-phosphate, and the phosphate of the preceding nucleotide²⁴⁻²⁶, which favors the C3'-endo conformation of the ribose (see **Supplementary Fig. 2** online), as well as to increased stacking involving the Ψ base²⁷. We tested modifications of the ribose ring at position U24 and found that introduction of 2'-F, which likewise favors the C3'-endo sugar conformation²⁸, is also deleterious for snRNP assembly (**Fig. 5c**). The smaller inhibitory effect of 2'-F (**Fig. 5c**) compared to the greater than two-fold effect exerted by the Ψ substitution (**Fig. 2b-d**) is expected, as pseudouridine has an additional backbone-rigidifying effect (because of increased base stacking²⁷). In contrast, 2'-deoxy modification, which favors the C2'- over the C3'-endo conformation^{28,29}, has no negative

effect (Fig. 5c). These results provide excellent confirmation that the interference observed with U24 Ψ is due at least in part to a negative effect on Sm protein assembly conferred by the wrong sugar pucker at this position within the Sm-binding site.

DISCUSSION

The RNA determinants for *in vivo* assembly of functional U7 snRNP (Fig. 5d) include a stable hairpin downstream of the Sm-binding site, a spacing of 2 or more nucleotides between the hairpin and the Sm-binding site, correct identities of four bases in the 3' half of the Sm site and flexibility of the sugar-phosphate backbone at U24 in the 5' portion of the Sm-binding site. These features are conserved in putative and known U7 RNA sequences from different species (Supplementary Fig. 1), except *Drosophila*, where the penultimate Sm-site nucleotide is a U rather than an A.

The four nucleotides that are crucial for active U7 snRNP assembly are all positioned within the 3' half of the unique U7 Sm-binding site, which has been shown to UV cross-link to a 40-kDa U7-specific protein in oocyte extracts⁷, most probably Lsm11 (ref. 4). Changing these nucleotides individually abrogated RNP formation rather than inducing assembly of a canonical Sm core of proteins (Fig. 5a,b). Moreover, none of the mutations tested uncoupled processing of the RNA chimera from its assembly with the U7 Sm core. It seems that processing immediately follows formation of the Sm protein complex, as only trace amounts of full-length chimera could be immunoprecipitated with Y12 antibody, even after short incubation times (30 min) in the oocyte (Fig. 3b). Finally, consistent with mutational analysis of the mouse histone H4-12 pre-mRNA³⁰, even complete removal of the binding site for SLBP does not affect processing of the chimera, and additional factors beyond the assembled U7 snRNP (for example, symplekin¹⁵ and CPSF-73 (refs. 15,16)) do not seem to be limiting in *Xenopus* oocytes.

Our results extend previous studies of Sm snRNP assembly. An *in vivo* saturation-mutagenesis study of the Sm-binding site of yeast U5 snRNA³¹ found no major phenotypic effect for any single-nucleotide substitution in the 5' half of the Sm site. Instead, substitutions only in the 3' half of the U5 Sm site were deleterious, similar to our observations for U7. Previous *in vitro* studies with U1 snRNA using NAIM³² have revealed the importance of phosphates within the terminal stem-loop for formation of a stable RNP, suggesting a 'clamp' function for the U1 stem-loop in holding the Sm core of proteins in place. Although we did not detect phosphorothioate effects mapping to the U7 hairpin, deletion analyses revealed that a stable stem-loop downstream of the Sm site is essential (Figs. 3 and 4). We cannot exclude participation of another assembly factor, but it is unlikely that we have monitored interactions with the SMN complex of proteins. A recent report³³ has concluded that the SMN complex recognizes the Sm site plus the adjacent 3' stem-loop; either changing the base or inserting phosphorothioate at positions in viral U snRNAs comparable to U23 and U25 in U7 RNA was most deleterious. The fact that we did not see appreciable interferences at these positions (Figs. 2 and 5) supports the argument that our *in vivo* assays with the U7 chimera assess Sm rather than SMN interactions.

Most notably, we found that individual base alterations in the uridines (U23, U24 and U25) considered the 'signature' of the Sm-binding site³⁴⁻³⁷ do not substantially inhibit assembly (Fig. 5). Instead, the *in vivo* NAIM combined with nuclear import and Y12 immunoprecipitation assays of singly substituted U7 molecules showed that pseudouridine at position 24 substantially impairs RNP formation (Fig. 2). As Ψ does not alter the Watson-Crick face of the base, which is recognized in the nucleotide-binding pocket of Sm

proteins^{38,39}, and as substitution of any other base at that position leads to efficient assembly, it is evident that the Ψ must influence not base recognition, but some other aspect of Sm core assembly.

We hypothesize that pseudouridine within the U7 Sm site impairs assembly of the Sm ring because of both its backbone rigidity and its preference for assuming the C3'-endo sugar conformation. Inhibition of assembly by 2'-fluorodeoxyribose, which also prefers the C3'-endo sugar pucker²⁸, was likewise observed at position 24, whereas 2'-deoxyribose, as predicted, had no effect (Fig. 5c). Although the C3'-endo sugar pucker is a general feature of RNA molecules, in this particular position it seems to be incompatible with proper assembly of the Sm core proteins.

We predict that when crystal structures of the heteroheptameric Sm ring of the U7 snRNP become available, the C2'-endo sugar conformation will be seen at position U24 and perhaps at the neighboring U positions. Indeed, structural analyses of RNA bound to an archaeal homoheptameric Sm protein complex³⁸ and to homohexameric Hfq⁴⁰, the bacterial Sm homolog, have led the authors to suggest that the bound nucleotides assume a C2'-endo sugar conformation. This configuration promotes a more extended structure for the backbone and could enable the binding of consecutive bases in neighboring binding pockets^{38,40}. Our observation that pseudouridine is deleterious for *in vivo* assembly of the U7 RNP (Fig. 2) provides evidence that the ability of the RNA backbone within the Sm site to adopt the C2'-endo configuration is not only a structural feature but functionally important.

We asked whether similar backbone constraints apply to assembly of the canonical Sm ring, as this portion of the U7 Sm site cross-links to the common SmG protein⁷. We examined U5 RNA containing Ψ at the position equivalent to U24 but found no inhibitory effect on assembly with Sm proteins in *Xenopus* oocytes (data not shown); similar results were obtained for U7 RNA with a canonical spliceosomal Sm site (U7 OPT, Fig. 2d). Both of these RNAs assemble with the Sm core more efficiently than does the wild-type U7 RNA (Fig. 2d), and they have Sm sites containing six and five consecutive uridines, respectively, compared to three in U7. It may be that multiple uridines contribute sufficient flexibility to override the effect of a single pseudouridine within a consensus Sm site and that the relatively inefficient assembly of the U7 snRNP has allowed us to detect the deleterious effect of Ψ . Regardless, the unstructured nature of oligoU⁴¹, and its expected facilitation of assembly with Sm proteins, has probably contributed to the conservation of U residues within the Sm-binding sites of all U RNAs.

To our knowledge, this is the first demonstration that the backbone conformation of a particular nucleotide is crucial for the *in vivo* assembly of an RNA-protein complex. More generally, our results provide support for the idea that the evolutionary pressure that maintains pseudouridine modifications at conserved positions within many important cellular RNAs²⁶ (including ribosomal RNAs and snRNAs⁴² as well as transfer RNAs) reflects a requirement—either during assembly or functioning—for rigidity of the sugar-phosphate backbone at certain sites.

METHODS

DNA templates for *in vitro* transcription. See Supplementary Methods online.

***In vivo* NAIM.** Capped wild-type histone pre-mRNA-U7 snRNA chimeras containing a single, randomly positioned α -thioated nucleotide analog or α -thioated parent nucleotide were synthesized^{18,19}. The RNA was labeled at the 3' end as described in ref. 20, with modifications (see Supplementary Methods) and microinjected into oocytes (~10 fmol chimera per oocyte, ~100 oocytes for each sample). After overnight incubation at 18 °C, total

oocyte RNA was prepared by proteinase K digestion, extraction with phenol and with a ratio of chloroform to isoamyl alcohol of 24:1, and precipitation with ethanol. Processed RNA was eluted from 1.5-mm-thick denaturing 12% (w/v) polyacrylamide gels. Treatment of the processed RNAs and uninjected full-length RNA controls with iodine and analysis of the data were performed essentially as described previously^{18–20}.

Oocyte microinjections. Stage VI *Xenopus laevis* oocytes were cytoplasmically injected with 13.8 nl of RNA solution (corresponding to 3.5×10^3 c.p.m. or ~ 5 fmol, unless stated otherwise) containing trace amounts of bromophenol blue and $1 \mu\text{l}^{-1}$ RNasin. Oocytes were kept in OR2 buffer (5 mM HEPES, 82.5 mM NaCl, 2.5 mM KCl, 1 mM Na_2HPO_4 , 1 mM MgCl_2 and 1 mM CaCl_2 (pH 7.8)). Injected oocytes were incubated overnight (~ 16 h) at 18 °C.

Nuclear import assay. Preparation of singly substituted U7 RNAs by DNA-mediated RNA ligation is described in **Supplementary Methods**. Gel-purified RNAs (20 fmol) were microinjected into the oocyte cytoplasm and incubated overnight at 18 °C in OR2. The cytoplasm and nucleus of each cell were manually separated, with 5 oocytes per sample in J buffer (10 mM HEPES (pH 7.4), 70 mM NH_4Cl , 7 mM MgCl_2 , 0.1 mM EDTA and 2.5 mM DTT), and immediately digested with proteinase K, extracted with phenol and with a ratio of chloroform to isoamyl alcohol of 24:1, and precipitated with ethanol. After electrophoresis under denaturing conditions, the percent RNA imported was calculated on the basis of PhosphorImager quantification of the individual band intensities.

Immunoprecipitation. Protein A Sepharose CL-4B beads (Amersham Biosciences) were preswollen in TBS (50 mM Tris-HCl (pH 7.4) and 150 mM NaCl). For each sample, 70 μl swollen beads were preincubated with the indicated antibodies in 0.5 ml TBS for 1 h at room temperature, washed three times with 1 ml TBS and kept in 0.5 ml TBS on ice. For each sample, ten oocytes were homogenized in 0.2 ml ice-cold 130 mM NaCl and 34 mM sodium phosphate (pH 7.5), then centrifuged at 13,400g for 1.5 min at room temperature; the clear extract between the insoluble pellet and the lipid layer was transferred to a tube containing antibody-bound protein A Sepharose and incubated with rotation for 1 h at 4 °C. The beads were washed four times with 1 ml TBS and bound RNA was eluted by extraction with phenol. Total oocyte RNA was prepared by homogenizing ten oocytes with 200 μg proteinase K in 0.2 ml of solution containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 12.5 mM EDTA and 1% (w/v) SDS, then incubating for 30 min at 45 °C, extracting with phenol and with a ratio of chloroform to isoamyl alcohol of 24:1, and precipitating with ethanol. After electrophoresis under denaturing conditions, assembly efficiencies were calculated by PhosphorImager quantification of the percentage of precipitated RNA from band intensities for total and Y12-bound RNAs. Processing efficiencies were calculated as the fraction of processed RNA in the lanes containing total oocyte RNA.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank S. Strobel (Yale University) for generous gifts of 5'-O-(1-thio)-nucleoside analog triphosphates; R. Breaker, K. Tycowski and L. Weinstein-Szewczak for stimulating discussions; T. Steitz, K. Tycowski, S. Vasudevan and A. Alexandrov for critical reading of the manuscript; and A. Miccinello for secretarial help. This work was supported by US National Institutes of Health grant GM26154 to J.A.S. J.A.S. is an investigator of the Howard Hughes Medical Institute.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Strub, K., Galli, G., Busslinger, M. & Birnstiel, M.L. The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. *EMBO J.* **3**, 2801–2807 (1984).
2. Mowry, K.L. & Steitz, J.A. Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. *Science* **238**, 1682–1687 (1987).

3. Dominski, Z., Yang, X.C., Purdy, M. & Marzluff, W.F. Cloning and characterization of the *Drosophila* U7 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* **100**, 9422–9427 (2003).
4. Pillai, R.S. *et al.* Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* **17**, 2321–2333 (2003).
5. Schümperli, D. & Pillai, R.S. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell. Mol. Life Sci.* **61**, 2560–2570 (2004).
6. Azzouz, T.N. *et al.* Toward an assembly line for U7 snRNPs: interactions of U7-specific Lsm proteins with PRMT5 and SMN complexes. *J. Biol. Chem.* **280**, 34435–34440 (2005).
7. Stefanovic, B., Hackl, W., Lührmann, R. & Schümperli, D. Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into *Xenopus* oocytes. *Nucleic Acids Res.* **23**, 3141–3151 (1995).
8. Pillai, R.S., Will, C.L., Lührmann, R., Schümperli, D. & Müller, B. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J.* **20**, 5470–5479 (2001).
9. Grimm, C., Stefanovic, B. & Schümperli, D. The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J.* **12**, 1229–1238 (1993).
10. Wu, C.H., Murphy, C. & Gall, J.G. The Sm binding site targets U7 snRNA to coiled bodies (spheres) of amphibian oocytes. *RNA* **2**, 811–823 (1996).
11. Marzluff, W.F. & Duronio, R.J. Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. *Curr. Opin. Cell Biol.* **14**, 692–699 (2002).
12. Schaufele, F., Gilmartin, G.M., Bannwarth, W. & Birnstiel, M.L. Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of H3 messenger RNA. *Nature* **323**, 777–781 (1986).
13. Bond, U.M., Yario, T.A. & Steitz, J.A. Multiple processing-defective mutations in a mammalian histone pre-mRNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. *Genes Dev.* **5**, 1709–1722 (1991).
14. Dominski, Z., Zheng, L.X., Sanchez, R. & Marzluff, W.F. Stem-loop binding protein facilitates 3'-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. *Mol. Cell. Biol.* **19**, 3561–3570 (1999).
15. Kolev, N.G. & Steitz, J.A. Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes Dev.* **19**, 2583–2592 (2005).
16. Dominski, Z., Yang, X.C. & Marzluff, W.F. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* **123**, 37–48 (2005).
17. Stefanovic, B., Wittop Koning, T.H. & Schümperli, D. A synthetic histone pre-mRNA-U7 small nuclear RNA chimera undergoing *cis* cleavage in the cytoplasm of *Xenopus* oocytes. *Nucleic Acids Res.* **23**, 3152–3160 (1995).
18. Ryder, S.P. & Strobel, S.A. Nucleotide analog interference mapping. *Methods* **18**, 38–50 (1999).
19. Ryder, S.P., Ortoleva-Donnelly, L., Kosek, A.B. & Strobel, S.A. Chemical probing of RNA by nucleotide analog interference mapping. *Methods Enzymol.* **317**, 92–109 (2000).
20. Szewczak, L.B., DeGregorio, S.J., Strobel, S.A. & Steitz, J.A. Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem. Biol.* **9**, 1095–1107 (2002).
21. Lerner, E.A., Lerner, M.R., Janeway, C.A., Jr. & Steitz, J.A. Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. *Proc. Natl. Acad. Sci. USA* **78**, 2737–2741 (1981).
22. Fischer, U., Sumpter, V., Sekine, M., Satoh, T. & Lührmann, R. Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J.* **12**, 573–583 (1993).
23. Gilmartin, G.M., Schaufele, F., Schaffner, G. & Birnstiel, M.L. Functional analysis of the sea urchin U7 small nuclear RNA. *Mol. Cell. Biol.* **8**, 1076–1084 (1988).
24. Arnez, J.G. & Steitz, T.A. Crystal structure of unmodified tRNA(Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* **33**, 7560–7567 (1994).
25. Newby, M.J. & Greenbaum, N.L. Investigation of Overhauser effects between pseudouridine and water protons in RNA helices. *Proc. Natl. Acad. Sci. USA* **99**, 12697–12702 (2002).
26. Charette, M. & Gray, M.W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351 (2000).
27. Davis, D.R. Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res.* **23**, 5020–5026 (1995).
28. Uesugi, S., Miki, H., Ikehara, M., Iwahashi, H. & Kyogoku, Y. A linear relationship between electronegativity of 2'-substituents and conformation of adenine nucleosides. *Tetrahedr. Lett.* **20**, 4073–4076 (1979).
29. Altona, C. & Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. Improved method for the interpretation of proton magnetic resonance coupling constants. *J. Am. Chem. Soc.* **95**, 2333–2344 (1973).
30. Streit, A., Koning, T.W., Soldati, D., Melin, L. & Schümperli, D. Variable effects of the conserved RNA hairpin element upon 3' end processing of histone pre-mRNA *in vitro*. *Nucleic Acids Res.* **21**, 1569–1575 (1993).
31. Jones, M.H. & Guthrie, C. Unexpected flexibility in an evolutionarily conserved protein-RNA interaction: genetic analysis of the Sm binding site. *EMBO J.* **9**, 2555–2561 (1990).
32. McConnell, T.S., Lokken, R.P. & Steitz, J.A. Assembly of the U1 snRNP involves interactions with the backbone of the terminal stem of U1 snRNA. *RNA* **9**, 193–201 (2003).



33. Golembe, T.J., Yong, J. & Dreyfuss, G. Specific sequence features, recognized by the SMN complex, identify snRNAs and determine their fate as snRNPs. *Mol. Cell. Biol.* **25**, 10989–11004 (2005).
34. Jarmolowski, A. & Mattaj, I.W. The determinants for Sm protein binding to *Xenopus* U1 and U5 snRNAs are complex and non-identical. *EMBO J.* **12**, 223–232 (1993).
35. Hartmuth, K., Raker, V.A., Huber, J., Branlant, C. & Lührmann, R. An unusual chemical reactivity of Sm site adenosines strongly correlates with proper assembly of core U snRNP particles. *J. Mol. Biol.* **285**, 133–147 (1999).
36. Raker, V.A., Hartmuth, K., Kastner, B. & Lührmann, R. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol. Cell. Biol.* **19**, 6554–6565 (1999).
37. Urlaub, H., Raker, V.A., Kostka, S. & Lührmann, R. Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J.* **20**, 187–196 (2001).
38. Törö, I. *et al.* RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.* **20**, 2293–2303 (2001).
39. Thore, S., Mayer, C., Sauter, C., Weeks, S. & Suck, D. Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya. *J. Biol. Chem.* **278**, 1239–1247 (2003).
40. Schumacher, M.A., Pearson, R.F., Møller, T., Valentin-Hansen, P. & Brennan, R.G. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.* **21**, 3546–3556 (2002).
41. Inners, L.D. & Felsenfeld, G. Conformation of polyribouridylic acid in solution. *J. Mol. Biol.* **50**, 373–389 (1970).
42. Newby, M.I. & Greenbaum, N.L. Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat. Struct. Biol.* **9**, 958–965 (2002).
43. Cotten, M., Gick, O., Vasserot, A., Schaffner, G. & Birnstiel, M.L. Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the *in vitro* 3' RNA processing reaction. *EMBO J.* **7**, 801–808 (1988).