10. U. Gubler and B. J. Hoffman, *Gene* **25**, 263 (1983).
11. The initial screen of the library yielded 12 clones that hybridized to probe UM1 (Table 1). Results of DNA sequence analysis showed that the longest clone [4.3; 1800-bp insert] and an overlapping clone (4.4; 950-bp insert) contained an open reading frame but lacked an initiation codon and leader sequence. To obtain the additional missing 5′ sequence, the library was screened again with an antisense 30-nucleotide probe that corresponded to the 5′ end of clone 4.4. This yielded additional clones, including clone 6.7 that contains 105 nucleotides of 5′ untranslated region and encodes the putative initiating methionine, the leader sequence, and the mature amino terminus of the polypeptide.
12. A possible exception is rat heart, where a faint 2.4-kb band could be seen after a 3-week exposure of the blot.
13. S. Kumar *et al.*, *Histochem. J.* **17**, 1251 (1985).
14. I. Tamm and F. Horsfall, *J. Exp. Med.* **95**, 71 (1952).
15. This clone contains a portion of the 5′ untranslated region (32 bp), an open reading frame that potentially encodes a 24–amino acid leader, a 615–amino acid mature protein, and a portion of the 3′ untranslated region (26 bp). For transient expression in murine L cells, the gene was excised from pUC19 and inserted into the polylinker region of pBG312MO, an expression vector that contains the Moloney long terminal repeat in place of the SV40 early and adenovirus major late promoter region of pBG312 (*8*).
16. Both the SEQHP and SEQDP programs were developed by M. Kaneshia, Laboratory of Mathematical Biology, National Cancer Institute.
17. T. C. Sudhoff *et al.*, *Science* **228**, 893 (1985); J. Scott *et al.*, *ibid.* **221**, 236 (1983).
18. E. Ruoslahti and M. D. Pierschbacher, *Cell* **44**, 517 (1986).
19. A. Sherblom, J. Decker, A. Muchmore, in preparation.
20. E. Schenk, R. Schwartz, R. Lewis, *Lab. Invest.* **25**, 92 (1971); J. R. Hoyer *et al.*, *ibid.* **30**, 757 (1974).
21. A. M. Grant, L. R. I. Baker, A. Neuberger, *Clin. Sci.* **44**, 377 (1964).
22. J. S. Hunt and A. R. McGiven, *Immunology* **35**, 391 (1978).
23. A. Abbondanza *et al.*, *Biochem. J.* **187**, 525 (1980).
24. Half-maximal inhibition of antigen-specific T-cell proliferation was determined for the eight batches of uromodulin representing a total of 21 individual assays; similarly, the eight batches of Tamm-Horsfall glycoprotein were assayed a total of 14 times. If a preparation did not inhibit proliferation, it was arbitrarily assigned the highest value of uromodulin or Tamm-Horsfall protein tested (this ranged from 500 ng to 2 μg per culture). Since four of the eight Tamm-Horsfall preparations did not inhibit at any dose, this method yields a minimum estimate.
25. J. Hoyer and M. Seiler, *Kidney Int.* **16**, 279 (1979).
26. B. A. Beutler, I. W. Milsark, A. Cerami, *J. Immunol.* **135**, 3972 (1985).
27. H. R. MacDonald *et al.*, *FEBS Lett.* **209**, 295 (1986).
28. A. V. Muchmore *et al.*, in preparation.
29. D. Pennica *et al.*, *Science* **236**, 83 (1987).
30. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
31. G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984); R. Tizard and H. Nick, unpublished information.
32. B. Wallner *et al.*, *Nature (London)* **320**, 77 (1986).
33. A. P. Feinberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983).
34. L. M. Sompayrac and K. J. Danna, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7575 (1981).

17 February 1987; accepted 15 May 1987

Saccharomyces cerevisiae Has a U1-Like Small Nuclear RNA with Unexpected Properties

PAUL G. SILICIANO, MICHELE HALTINER JONES, CHRISTINE GUTHRIE

Previous experiments indicated that only a small subset of the ≃24 small nuclear RNAs (snRNAs) in *Saccharomyces cerevisiae* have binding sites for the Sm antigen, a hallmark of metazoan small nuclear ribonucleoproteins (snRNPs) involved in pre-messenger RNA splicing. Antibodies from human serum to Sm proteins were used to show that four snRNAs (snR7, snR14, snR19, and snR20) can be immunoprecipitated from yeast extracts. Three of these four, snR7, snR14, and snR20, have been shown to be analogs of mammalian U5, U4, and U2, respectively. Several regions of significant homology to U1 (164 nucleotides) have now been found in cloned and sequenced snR19 (568 nucleotides). These include ten out of ten matches to the 5′ end of U1, the site known to interact with the 5′ splice site of mammalian introns. Surprisingly, the precise conservation of this sequence precludes perfect complementarity between snR19 and the invariant yeast 5′ junction (GTATGT), which differs from the mammalian consensus at the fourth position (GTPuAGT).

NUCLEAR PRE-MESSENGER RNA (mRNA) splicing is mediated by a set of small nuclear RNAs (snRNAs) (U1, U2, and U4 to U6) complexed with proteins that react with human autoimmune sera of the Sm specificity (*1*). The splicing reaction appears to be guided by specific small nuclear ribonucleoprotein (snRNP): intron and snRNP:snRNP interactions, the best characterized of which is that between U1 and the pre-mRNA. This interaction was first postulated 7 years ago based on the striking complementarity between the 5′ ten nucleotides of U1 and the conserved sequence found at metazoan 5′ splice junctions (*2*). Subsequent data from in vitro splicing systems have provided persuasive support for this model (reviewed in *3*). Furthermore, Zhuang and Weiner (*3*) recently demonstrated base pairing at two of the ten positions by constructing base changes in U1 to compensate for point mutations in the 5′ junction.

We have previously identified snRNAs in *Saccharomyces cerevisiae* by criteria that include possession of a characteristic trimethylguanosine (TMG) cap (*4*). In contrast to our expectations from work in higher eukaryotes (i) yeast snRNAs are present in relatively low abundance (10 to 10^3 copies per cell in yeast versus 10^5 to 10^6 in mammalian cells) (*5*); (ii) the yeast snRNA family contains at least two dozen species, some of which are quite large (up to six times the largest metazoan snRNA) (*4*); and (iii) at least six of these are dispensable for growth (*6*).

In order to identify yeast snRNAs likely to function in mRNA splicing, we looked for those RNAs that were associated with Sm proteins. Because reported attempts to immunoprecipitate snRNAs from whole cell sonicates with antibodies to the Sm proteins (anti-Sm) were unsuccessful (*7*), we used the alternative approach of microinjecting ^{32}P-labeled yeast snRNAs into *Xenopus* oocytes (*8*), which stockpile the Sm protein (*9*). We reasoned that if any yeast snRNAs contained Sm binding sites (*10*), they would assemble into Sm-snRNPs and become immunoprecipitable. By this assay, two snRNA species, snR7 and snR14, were detected after immunoprecipitation with human antibodies (*8, 11*). We have since demonstrated that snR7 (179 nt) has limited sequence-specific but strong structural homology to U5 (116 nt) (*12*), while snR14 (160 nt) has several blocks of sequence homology to U4 (144 nt) (*13*). By sequence analysis, snR7 and snR14 each contain a consensus Sm binding site ($AU_{5\text{-}6}$ GPu). Moreover, both RNAs are essential for viability (*12, 13*) and can be found in association with the spliceosome (*14*). This is also true for a third snRNA, snR20 (1175 nt; also called LSR1) (*14*), which has a 5′ domain of striking sequence and structural homology to U2 (187 nt), including a presumptive Sm binding site (*15*).

These results suggested that, despite differences in size and structure, yeast might contain a full complement of functional analogs to the Sm snRNPs of higher eukaryotes, including the U1 counterpart. Trivial explanations, such as degradation, could account for the failure of species other than snR7 and snR14, including snR20, to be efficiently immunoprecipitated after microinjection. We now have turned to direct analysis of yeast snRNPs. Knowing that yeast snRNAs are much less abundant than their mammalian counterparts (*4, 5*), we have used active splicing extracts as starting material because these are likely to be enriched for the snRNPs of interest.

Extracts were prepared and incubated in the presence of protein A–Sepharose that

Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143. The order of the first two authors is arbitrary.

had been bound with either anti-Sm or antibodies to the snRNA-specific trimethylguanosine cap (anti-TMG), or nonimmune human serum as a control. Following immunoprecipitation, the pellet and supernatant fractions were deproteinized and the RNAs were 3′ end labeled with pCp and analyzed by electrophoresis (Fig. 1). Immunoprecipitation with anti-TMG (Fig. 1, lane 2) reveals the full complement of snRNAs present in the splicing extract; these approximately ten species are only a portion of the snRNA family in yeast (*4*). Anti-Sm from two different patients (Sm1, lane 3; Sm2, lane 4) immunoprecipitate a subset of these species; snR7(L and S), snR14, snR19, and snR20. The specificity of the reaction is demonstrated by the absence of RNA in the nonimmune pellet (NHS, lane 5) and by the failure of snR13, snR17, and snR18 to be immunoprecipitated despite their abundance in both the anti-TMG pellets and the anti-Sm supernatants. We conclude that four major yeast snRNAs in splicing extracts associate with an Sm antigen, namely snR7, snR14, snR19, and snR20 (*16*).

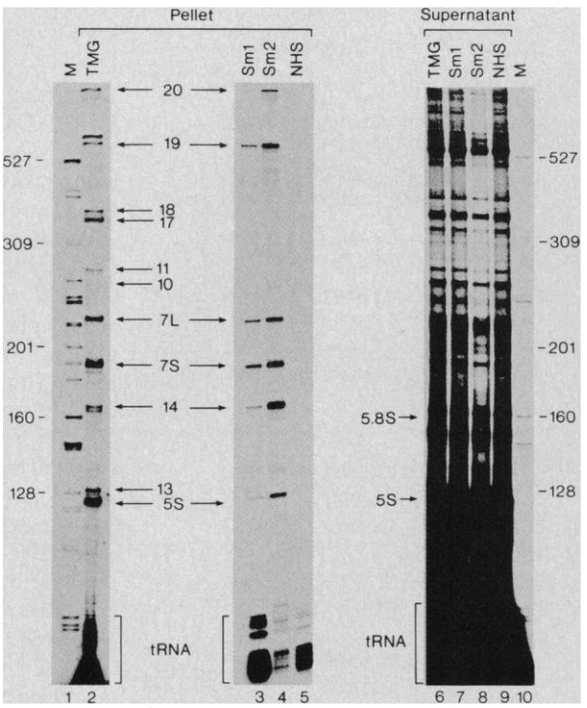

Fig. 1. Immunoprecipitation of yeast snRNAs with anti-Sm. A modified procedure based on that of Wolin and Steitz (*24*) was used. Protein A coupled to Sepharose beads (Pharmacia) was incubated for 1 hour at room temperature with either anti-TMG, or one of the two anti-Sm or nonimmune human serum in NET-2 buffer (50 m*M* tris, *p*H 7.6, 150 m*M* NaCl, 0.05% Nonidet P-40). Splicing extract (*25*) was prepared and incubated for 1 hour at 4°C with the washed, bound antibodies in NET-2 buffer with protease inhibitors and then centrifuged. The pellets were washed extensively, and both the pellet and the supernatant fraction were extracted with a phenol-chloroform mixture and precipitated with ethanol. The RNA (one-tenth of the supernatant) was 3′ end labeled with 3′,5′-cytidine [^{32}P]bisphosphate (3000 Ci/mmol, Amersham) (*26*). The pellet (lanes 2 to 5) and supernatant (lanes 6 to 9) were analyzed by electrophoresis on 6% denaturing acrylamide gels and then autoradiography. (Lanes 1 and 10) DNA size markers (M), end labeled pBR325 Hpa II fragments; (lanes 2 to 9) pellet and supernatant, respectively, of immunoprecipitation reactions as follows: (lanes 2 and 6) anti-TMG; (lanes 3 and 7) first anti-Sm (Sm1) (snR20 is visible in lane 3 after longer exposure); (lanes 4 and 8) second anti-Sm (Sm2); (lanes 5 and 9) nonimmune human serum (NHS). This experiment is representative of at least ten independent immunoprecipitations with three different yeast extracts and a third antiserum in addition to the two shown.

The cloned sequences hybridize with an RNA of the size expected for snR19 in both total yeast RNA (Fig. 2A, lane 1) and fraction I RNA (Fig. 2A, lane 2). Furthermore, the clone hybridizes with an RNA of the same size in the pellet fraction of an Sm immunoprecipitation but not of a nonimmune control immunoprecipitation. Southern analysis of yeast genomic DNA (Fig. 2B) demonstrates that *SNR19* is a single copy gene. The location of the 5′ end of snR19 was determined by primer-directed RNA sequencing (Fig. 2C). The 3′ end was located by S1 nuclease protection analysis and enzymatic RNA sequencing. The length of the RNA determined by these methods is 568 nt.

The entire sequence of *SNR19*, including 5′ and 3′ flanking regions, is shown in Fig. 3. A presumptive TATA box, TATAAATA, is located 98 nt upstream of the transcription start site, a position in accordance with that for other yeast snRNA genes (*6*). Three potential Sm binding sites (Fig. 3) agree

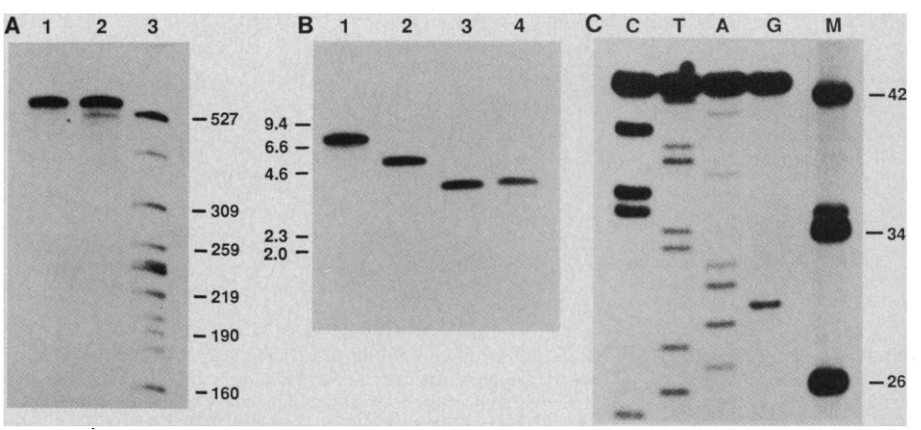

Fig. 2. (**A**) Northern analysis of *SNR19*. RNA was separated on a 6% polyacrylamide 7*M* urea gel and transferred electrophoretically to a nylon membrane (Hybond, Amersham). The filter was probed with the original genomic clone (10 kb) which had been labeled by nick translation. Lane 1 contained total yeast RNA (25 μg) while lane 2 contained RNA (5 μg) isolated from fraction I of the in vitro splicing extract (*27*). Lane 3 contained end labeled pBR325 Hpa II fragments as size markers. Hybridization conditions were as described in Wise *et al.* (*5*). (**B**) Southern analysis of *SNR19*. A Southern filter, prepared as described, was probed with a 900-nt restriction fragment (isolated from a Bal-31 deletion clone) containing the entire *SNR19* coding region as well as 150 nt of 3′ flanking sequence and 80 nt of 5′ flanking sequence. Yeast genomic DNA, 2 μg per lane, was digested with Bam HI (lane 1), Bgl II (lane 2), Hind III (lane 3), and Eco RI (lane 4). Size markers are lambda phage Hind III fragments visualized by ethidium bromide staining. Hybridization conditions were as described in Wise *et al.* (*5*). (**C**) RNA sequence analysis of snR19. A ^{32}P-labeled oligonucleotide (19G), complementary to positions 43 to 26 of snR19, was annealed to total yeast RNA (5 μg) and extended with avian myoblastosis virus reverse transcriptase (Life Sciences) in the presence of dideoxynucleotide triphosphates (*28*). The reactions were analyzed on a 12% polyacrylamide 7*M* urea gel. Markers are end labeled pBR325 Hpa II fragments.

with the consensus derived from metazoan and yeast snRNAs ($AU_{5\text{-}6}GPu$).

Comparison of snR19 (568 nt) to U1 (164 nt) reveals two stretches of significant homology within the first 50 nt (Fig. 4A). There is a perfect match between the 10 nt at the 5′ end of snR19 and the 5′ end of all known metazoan U1 RNAs. This is followed by nine out of ten matches at positions 26 to 35 of snR19 to positions 28 to 37 of U1. The latter sequence is found on top of hairpin I in the single-stranded loop designated "a" by Branlant *et al.* (*17*). The corresponding nucleotides in snR19 can be drawn in a comparable structure (Fig. 4A). The sizes of the loop (10 nt in each) and the underlying stem (9 bp in snR19 and 10 bp in U1, with one unpaired residue in each) are in good agreement. Moreover, the position of this domain in the two molecules is virtually identical with respect to the 5′ terminus. Finally, a change from C to A at position 33 in loop a of human U1, which produces ten out of ten matches to the snR19 loop, does not measurably inhibit the function of U1 in an in vivo suppression assay (*18*).

Additional regions of homology between U1 and snR19 can be identified. A potential counterpart in snR19 to the stem II-loop b region of U1, including seven out of eight matches in the loop, is shown in Fig. 4B. However, in snR19 the stem is less strongly base-paired, and the homology is located 24 nt farther from the 5′ end. A weak homology to the loop of the third hairpin of U1 was also observed as five out of seven matches (285 to 289 nt) at the top of a 7-bp stem. However, the statistical significance of this short match is questionable. Moreover, data for the U1 analog from the fission yeast, *Schizosaccharomyces pombe*, suggest that this loop (as well as that of the fourth hairpin in U1) is not phylogenetically conserved (*19*).

We conclude that snR19 is the *S. cerevisiae* analog of U1. Despite substantial differences in length and overall organization, regions of strong homology were identified in putative single-stranded regions, as is the case for snR20 and U2 (*15*) and for snR7 and U5 (*12*). As would be expected from its proposed role in splicing, snR19 has been observed to be a component of the spliceosome (*20*).

The precise conservation of the first 10 nt of snR10 and U1 was unexpected since this precludes Watson-Crick pairing at the highly conserved fourth position (and the less conserved −2 position) of the yeast intron (Fig. 4C). That is, based on the strict sequence conservation of the 5′ junction in *S. cerevisiae* (*21*), it was anticipated that the 5′ splice site and the yeast U1 analog would be perfectly complementary. The functional

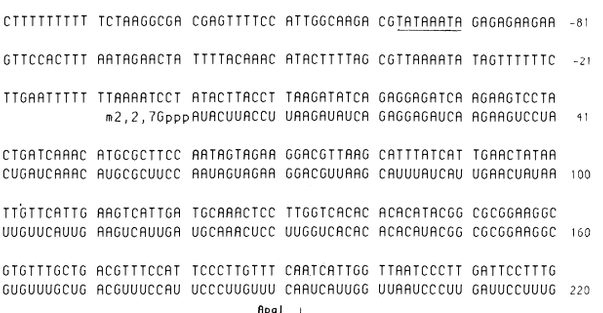

Fig. 3. Sequence of *SNR19*. The sequence of SNR19 and its flanking regions is shown. In the coding region, the RNA is drawn below the DNA sequence. A presumptive TATA box at −98 is underlined, three consensus Sm binding sites are boxed, and the positions of relevant restriction sites are indicated by arrows. The original clone isolated from the YEP13 library contained a 10-kbp insert. A 1.1-kb Hind III to Acc I restriction fragment carrying *SNR19* was identified by Southern hybridization with the oligonucleotide, subcloned into Bluescript M13 vectors (Stratagene Cloning Systems), and sequenced by the dideoxy method with the use of primers and protocols as recommended by the manufacturer. Sequence was also determined from the Apa I and Asu II sites as indicated, and from four synthetic oligonucleotides within the coding region: 19G, complementary to positions 43 to 26; 19C, complementary to positions 391 to 372; 19E, identical to positions 486 to 503; and 19A, complementary to positions 548 to 530. Additional sequence was obtained using a Bal-31 deletion which removes sequences between the Acc I site and the coding region.

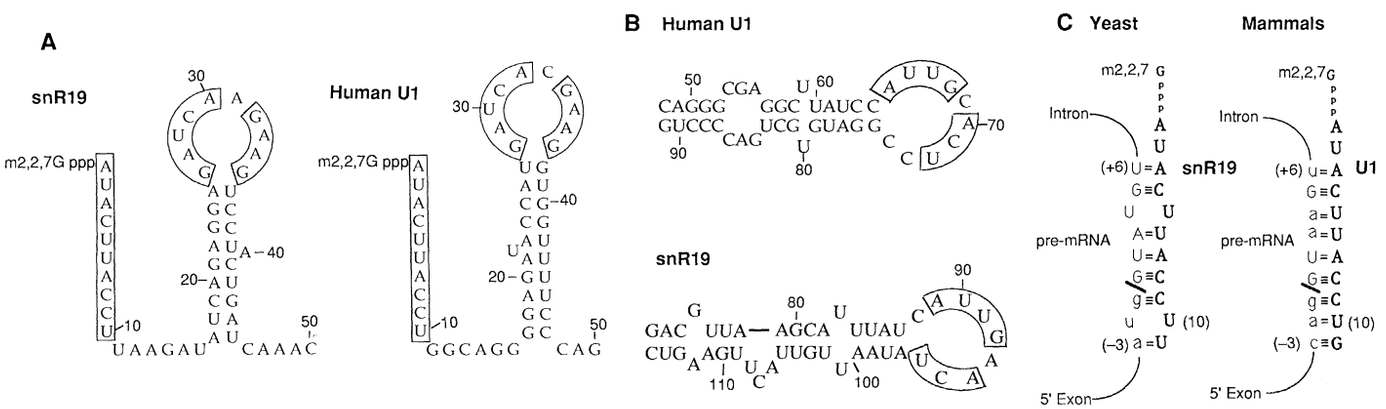

Fig. 4. (**A**) Homologies between snR19 and the 5′ end of human U1. The first 50 nt of snR19 and human U1 are drawn in the structure of Mount and Steitz (*29*). The G and U residues at the bottom of the loop are shown as base-paired. In the mammalian U1 RNA, U_5 and U_6 are actually pseudo U; it is not known whether this is the case in yeast snR19. (**B**) Loop b homologies. Stem II and loop b of human U1 (*29*) are shown along with the homologous region of snR19. While other homologies to loop b can be found in snR19, the optimal alignment is shown. (**C**) Potential base pairing with the 5′ splice junction. Potential base pairings between the first 10 nt of snR19 and the yeast consensus 5′ splice junction sequence are drawn, along with the idealized pairing between mammalian U1 and the mammalian consensus sequence (*23*). Bold letters indicate snR19 and U1 and pre-mRNA is shown in light type. For each pre-mRNA, nucleotides showing more than 80% conservation (*21, 23*) are shown as upper case letters, while those nucleotides conserved to less than 80% are shown as lower case letters. The diagonal line indicates the start of the intron (position +1). The number of potential hydrogen bonds is shown for each base pair.

importance of this conserved intron sequence has been demonstrated genetically; mutations at positions 1, 2, and 5 significantly impair splicing (*22*) (mutations at position 4 have not been tested). Furthermore, while few mammalian 5′ junctions fit their consensus perfectly (*23*), Zhuang and Weiner demonstrated that increasing the degree of complementarity from 5 to 6 bp (of 9 bp possible) between U1 and a mammalian 5′ splice site resulted in more efficient use of that junction (*3*).

On the other hand, perfect (nine out of nine) complementarity may be detrimental to optimal splicing efficiency. It is also possible that position 4 of the intron is not needed to form a canonical base pair with snR19, but may be involved in a separate recognition event. In addition, there may be a structural requirement for a U at position 5 (and 10) of snR19, perhaps for tertiary interactions. Construction of mutations in conserved sequences of snR19 should provide a genetic test of their function. Moreover, making the complementary changes in the intron will allow us to assess the specific contribution of Watson-Crick complementarity to 5′ splice junction recognition by snR19.

REFERENCES AND NOTES

1. T. Maniatis and R. Reed, *Nature (London)* **325**, 673 (1987).
2. M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wolin, J. A. Steitz, *ibid.* **283**, 220 (1980); J. Rogers and R. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1877 (1980). The original model included a more extensive interaction between U1 and the intron, which also included base pairing to the 3′ acceptor; the latter interaction has been ruled out on the basis of phylogenetic comparison (*25*).
3. Y. Zhuang and A. Weiner, *Cell* **46**, 827 (1986).
4. N. Riedel, J. A. Wise, H. Swerdlow, A. Mak, C. Guthrie, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8097 (1986).
5. J. A. Wise *et al.*, *Cell* **35**, 743 (1983).
6. D. Tollervey, J. A. Wise, C. Guthrie, *ibid.* p. 753; D. Tollervey and C. Guthrie, *EMBO J.* **4**, 3873 (1985); R. Parker *et al.*, in preparation.
7. M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wolin, J. A. Steitz, *Nature (London)* **283**, 220 (1980).
8. N. Riedel, S. Wolin, C. Guthrie, *Science* **235**, 328 (1987).
9. E. M. De Robertis, S. Lienhard, R. F. Parisot, *Nature (London)* **295**, 572 (1982); D. J. Forbes, T. B. Kornberg, M. W. Kirschner, *J. Cell Biol.* **97**, 62 (1983).
10. C. Branlant *et al.*, *EMBO J.* **1**, 1259 (1982); J.-P. Liautard, J. Sri-Wadada, C. Brunel, P. Jeanteur, *J. Mol. Biol.* **162**, 623 (1982); I. Mattaj and E. De Robertis, *Cell* **40**, 111 (1985).
11. D. Tollervey and I. Mattaj, *EMBO J.* **6**, 469 (1987).
12. B. Patterson and C. Guthrie, *Cell* **49**, 613 (1987). The single copy gene for snR7 encodes two forms of the RNA, designated L (214 nt) and S (179 nt), which differ at their 3′ ends.
13. P. Siliciano, D. Brow, H. Roiha, C. Guthrie, *ibid.*, in press.
14. C. W. Pikielny, B. C. Rymond, M. Rosbash, *Nature (London)* **324**, 341 (1986); N. Riedel *et al.*, in preparation.
15. M. Ares, Jr., *Cell* **47**, 49 (1986).
16. D. Tollervey and I. Mattaj (*11*) have reported anti-Sm precipitation of snR7 and snR14. Precipitation of other, higher molecular weight RNAs was not observed.
17. C. Branlant, A. Krol, J.-P. Ebel, *Nucleic Acids Res.* **9**, 841 (1981).
18. A. Weiner, personal communication.
19. J. A. Wise, personal communication.
20. B. Rymond, D. Torrey, M. Rosbash, *Genes Dev.* **1**, 238 (1987); N. Riedel *et al.*, in preparation.
21. J. L. Teem *et al.*, *Nucleic Acids Res.* **12**, 8295 (1984).
22. R. Parker and C. Guthrie, *Cell* **41**, 107 (1985); A. Jacquier, J. R. Rodriguez, M. Rosbash, *ibid.* **43**, 423 (1985); L. A. Fouser and J. D. Friesen, *ibid.* **45**, 81 (1986); U. Vijayraghavan *et al.*, *EMBO J.* **5**, 1683 (1986).
23. S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982).
24. S. L. Wolin and J. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1996 (1984).
25. R. J. Lin, A. J. Newman, S. Cheng, J. Abelson, *J. Biol. Chem.* **260**, 14780 (1985).
26. T. E. England and O. C. Uhlenbeck, *Nature (London)* **275**, 560 (1978).
27. S.-C. Cheng and J. Abelson, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2387 (1986).
28. D. McPheeters, A. Christensen, E. Young, G. Stormo, L. Gold, *Nucleic Acids Res.* **14**, 5813 (1986).
29. S. M. Mount and J. A. Steitz, *Nucleic Acids Res.* **9**, 6351 (1981).
30. The *SNR19* gene has also been cloned and sequenced by L. Kretzner, B. Rymond and M. Rosbash (*Cell*, in press).

31. We thank J. Steitz for giving us antibodies; A. Weiner and J. A. Wise for communication of results prior to publication; J. Tamm for help with the RNA sequencing; N. Riedel and C. Desplan for discussion and advice; and T. Simmons, B. Patterson, and K. Yamamoto for comments on the manuscript. Supported by NIH grant GM21119 and NSF grant DCB8603926 (C.G.), by a postdoctoral fellowship from the Leukemia Society (M.H.J.), and by a Damon Runyon–Walter Winchell Cancer Fund Fellowship, DRG-872 (P.G.S.)

8 July 1987; accepted 14 August 1987

A Sea Urchin Gene Encodes a Polypeptide Homologous to Epidermal Growth Factor

DEBORAH A. HURSH, MARY E. ANDREWS, RUDOLF A. RAFF

A sea urchin DNA clone complementary to an embryonic messenger RNA whose protein product bears striking homology to the epidermal growth factor family of proteins has been identified and characterized. The structure of the protein is similar to that of previously identified regulatory genes in *Drosophila* and *Caenorhabditis*. RNA gel blot hybridization showed a unique temporal pattern of expression of this gene during embryogenesis and transcript enrichment in the embryonic ectoderm. These results suggest that this member of the epidermal growth factor gene family plays a role in early development decisions in sea urchin embryos.

THE EPIDERMAL GROWTH FACTOR (EGF) family represents a rapidly expanding collection of proteins that have a variety of cellular functions, including differentiation, cell proliferation, and neoplastic transformation (*1*). Recently, the products of two developmental switch genes, the *notch* gene of *Drosophila* (*2, 3*) and the *lin 12* gene of *Caenorhabditis elegans* (*4*), were shown to be members of this gene family, heretofore thought to be exclusively mammalian. Our isolation of a complementary DNA (cDNA) clone encoding a protein resembling the products of these two genes suggests that the EGF peptide domain is widely used in metazoan development.

We isolated a 1.5-kb cDNA clone, uEGF-1, during a screen for cell lineage–specific genes in the sea urchin *Strongylocentrotus purpuratus*. This cDNA clone recognizes transcripts of 3.0 kb and 4.0 kb on RNA gel blots and is preferentially expressed in embryonic ectoderm, as determined by cell fractionation techniques (*5, 6*) (Fig. 1).

The sequence of this cDNA revealed a long open reading frame (ORF) of 1447 nucleotides ending with three in-frame stop codons (Fig. 2). Codon bias for this reading frame agrees with that of other sea urchin ORFs (*7*). A search of the Protein Information Resource (PIR) database by means of

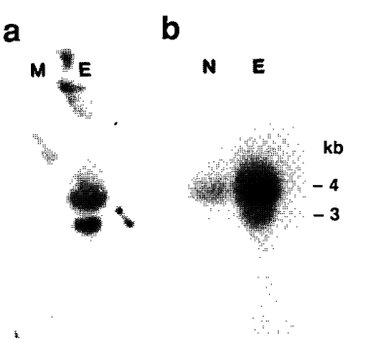

Fig. 1. RNA gel blot analysis of embryonic tissues. (**a**) Mesenchyme blastula embryos (27 hours) were fractionated into mesenchyme (M) and ectoderm (E) (*5*). (**b**) Early pluteus-stage embryos (68 hours) were fractioned into a mesenchyme-endoderm fraction (N) and ectoderm (E) (*6*). Total RNA was extracted from each fraction, electrophoresed in formaldehyde gels, and transferred to nitrocellulose. Ten micrograms of RNA was loaded from each fraction. Gels were stained with ethidium to monitor concentration. RNA extractions, blots, and hybridizations were as described (*26*).

Institute for Molecular and Cellular Biology, Department of Biology, Indiana University, Bloomington, Indiana 47405.