9 Progress in Molecular and Subcellular Biology

With Contributions by
M. Bachmann, C. Brunel, P. Jeanteur
R. Messer, W E G. Müller, R.E. Rhoads
H.C. Schröder, J. Sri-Widada

Edited by F. E. Hahn, D. J. Kopecko
W. E. G. Müller

Managing Editor: F. E. Hahn

9 Progress in Molecular and Subcellular Biology

With Contributions by
M. Bachmann, C. Brunel, P. Jeanteur
R. Messer, W. E. G. Müller, R. E. Rhoads
H. C. Schröder, J. Sri-Widada

Edited by F. E. Hahn, D. J. Kopecko
W. E. G. Müller

Managing Editor: F. E. Hahn

Springer-Verlag
Berlin Heidelberg New York Tokyo

Professor Fred E. Hahn, Ph. D.
Division of Communicable Diseases and Immunology
Walter Reed Army Institute of Research
Washington, D.C. 20307, USA

With 37 Figures

ISBN 3-540-15071-4 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-15071-4 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Catalog Card Number 75-79748

Printed in Germany.

Offsetprinting and Bookbinding: Konrad Triltsch, Graphischer Betrieb, 8700 Würzburg,
2131/3130-543210

Contributors

BACHMANN, M., Institut für Physiologische Chemie, Abt. "Angewandte Molekularbiologie", Universität, Duesbergweg, 6500 Mainz, FRG

BRUNEL, C., Laboratoire de Biochimie, Centre Paul Lamarque BP 5054, 34033 Montpellier and Laboratoire de Biologie Moléculaire, Université des Sciences et Techniques du Languedoc, 34060 Montpellier, France

JEANTEUR, P., Laboratoire de Biologie Moléculaire, Université des Sciences et Techniques du Languedoc, 34060 Montpellier, France

MESSER, R., Institut für Physiologische Chemie, Abt. "Angewandte Molekularbiologie", Universität, Duesbergweg, 6500 Mainz, FRG

MÜLLER, W.E.G., Institut für Physiologische Chemie, Abt. "Angewandte Molekularbiologie", Universität, Duesbergweg, 6500 Mainz, FRG

RHOADS, R.E., Department of Biochemistry, University of Kentucky Medical Center, Lexington, KY 40536-0084, USA

SCHRÖDER, H.C., Institut für Physiologische Chemie, Abt. "Angewandte Molekularbiologie", Universität, Duesbergweg, 6500 Mainz, FRG

SRI-WIDADA, J., Laboratoire de Biochimie, Centre Paul Lamarque BP 5054, 34033 Monpellier and Laboratoire de Biologie Moléculaire, Université des Sciences et Techniques du Languedoc, 34060 Montpellier, France

Contents

C. BRUNEL, J. SRI-WIDADA and P. JEANTEUR: snRNP's and scRNP's in Eukaryotic Cells (With 4 Figures)

A. Historical Background 1
B. The U snRNA's 3
C. Other Small RNA's Precipitable with Autoimmune Antibodies 26
D. Other Small RNP's of Organized Function 34
E. Concluding Remarks 36
References 38

C. SCHRÖDER, M. BACHMANN, R. MESSER and W.E.G. MÜLLER: Nucleotide-specific Ribonucleases from Eukaryotes. Their Possible Roles During Poly(A)(+)mRNA Maturation and Degradation (With 12 Figures)

A. Introduction 53
B. Posttranscriptional Events in mRNA Maturation 54
C. Eukaryotic Ribonucleases: a Brief Survey 59
D. Poly(A) Metabolism 62
E. Splicing 78
F. Role of Poly(A) Chain Length Controlled by Specific Nucleases and of snRNP's During Poly(A)(+)mRNA-nuclear Matrix Association and Nucleocytoplasmic Poly(A)(+)mRNA Transport 88
G. Conclusion and Future Directions 89
References 90

R.E. RHOADS: The Cap Structure of Eukaryotic Messenger RNA and its Interaction with Cap-binding Protein (With 12 Figures)

A. Introduction 104
B. 5'-Terminal Cap Structure of Messenger RNA 105
C. The Dependence of Translation on the Cap Structure 111
D. Identification of the Cap-binding Protein 120
E. Involvement of Cap-binding Proteins in Initiation of Protein Synthesis 127
F. Antibodies to Cap-binding Protein-related Polypeptides . 134
G. Effect of Virus Infection on Cap-binding Protein 136
H. Cap-binding Protein and the Cytoskeleton 139
I. Conclusion and Future Directions 141
K. 1984 Addendum 142
References 146

W.E.G. MÜLLER: Cell-Cell Recognition in Lower Multicellular Eukaryotes. Models Supporting th Modulation Theory of Adhesion (With 9 Figures)

A. Introduction 156
B. Cell-Cell Adhesion in Vertebrates 157
C. Cell Adhesion Molecules in Sponges 158
D. Cell Recognition Molecules in Corals 169
E. Conclusion 172
References 172

Subject Index 177

snRNP's and scRNP's in Eukaryotic Cells

C. Brunel, J. Sri-Widada and P. Jeanteur

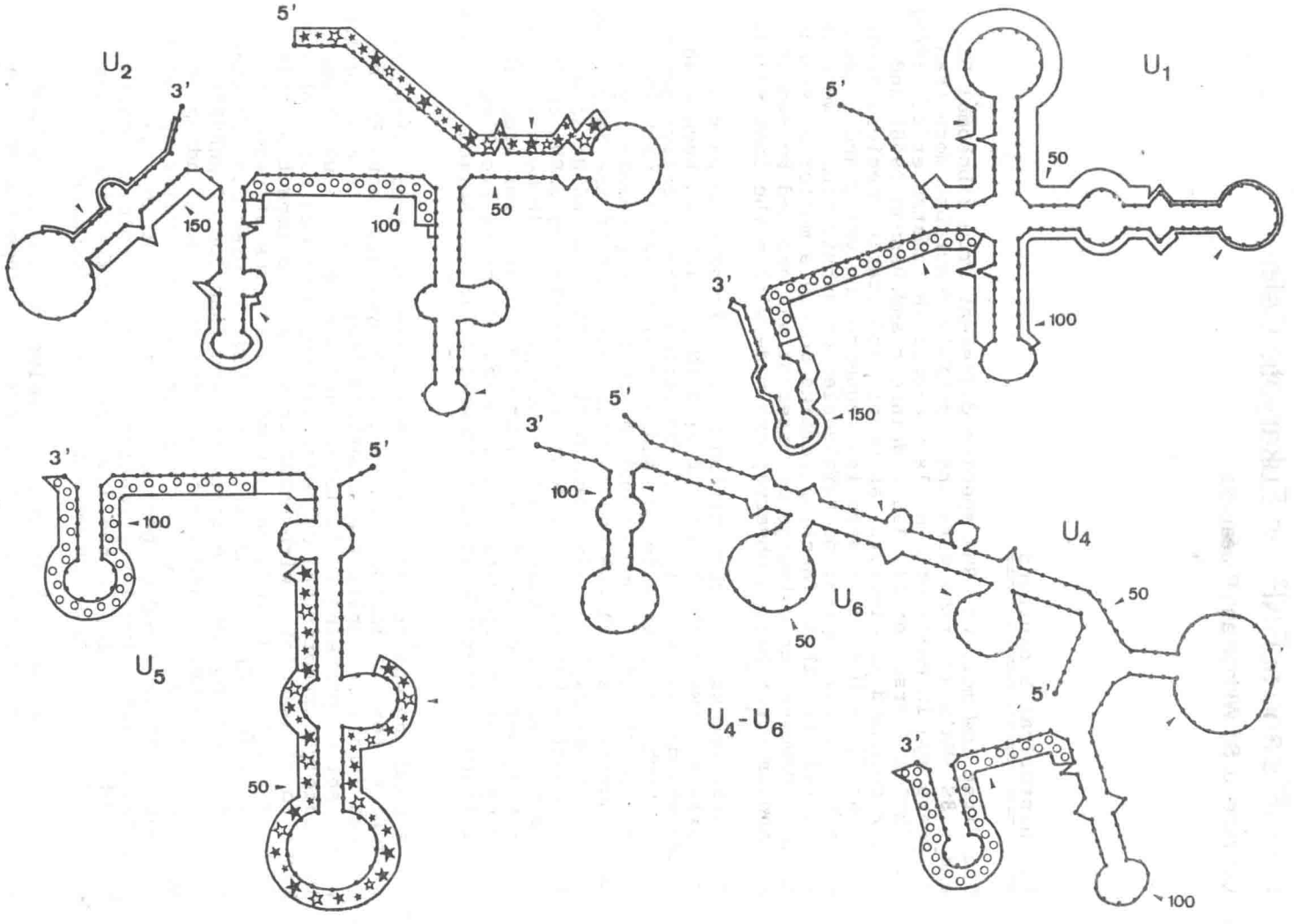
A. Historical Background

The first low mol. wt. RNA species different from ribosomal 5S and 5.8S RNA's and tRNA were the U snRNA's described more than 15 years ago in eukaryotic cells nucleoli (Muramatsu et al. 1966) and nuclei (Larsen et al. 1967; Weinberg and Penman 1968) and were recognized as discrete, abundant, and stable species. Their cellular distribution, synthesis, sequence analysis, and association with other nuclear macromolecules or substructures were investigated. Yet their function remains mostly a matter of speculation. However, an outsurge of interest was aroused by two simultaneous breakthrough observations made within the last years.

The first one was the recognition of a striking base pairing capability between the 5' end of U_1 snRNA, which has been highly conserved during evolution, and the intron side of splice junctions which also have a few consensus nucleotides at both their 5' (acceptor) and 3' (donor) end. A similar complementarity between adenovirus-2 VA RNA and a splice junction of the late adenovirus-2 transcript has also been noticed (Murray and Holliday 1979). On this basis, it has been proposed that U_1 snRNA could play a role in positioning intron-exon junctions (Rogers and Wall 1980; M.R. Lerner et al. 1980) in proper register for accurate splicing. Proteins specifically associated with U_1 are likely to participate in this recognition as well as in the splicing process itself.

The second one was the discovery that antibodies from patients suffering from autoimmune diseases, like Systemic Lupus Erythematosus (SLE), recognize the U snRNA's (except U_3) in association with a set of proteins (M.R. Lerner and Steitz 1979) which were immediately characterized as being neither histones nor proteins known to be associated with hnRNA in hnRNP. The impact of this finding has been enormous and opened new avenues of research along at least three major lines: firstly, it provided invaluable tools for probing both the structure and function of snRNP's; secondly, it provided important cues for the understanding of autoimmune diseases as well as a basis for a molecular approach of their pathology and pathogenesis; last but not least, it prompted an intensive investigation of other autoimmune antibodies which soon led to the discovery of new families of antigenic RNP's, many of which were found in the cytoplasm.

In this review, we will restrict ourselves to the low molecular RNA's which are found as RNP's. This will exclude the small nu-


U2
5'
3'
150
100
50
U1
5'
50
3'
100
150
U5
3'
5'
100
50
U4-U6
5'
3'
100
U6
50
U4
50
5'
3'
100

clear RNA's transcribed by RNA-polymerase I (Benecke and Penman 1977, 1979; Reichel et al. 1982) and referred to as snP1 RNA's. We will also overlook the entire field of translational effectors RNA's (Kennedy et al. 1978; Pluskal and Sarkar 1981).

B. The U snRNA's

In this chapter, our goal is to review recent results in the field of isolation, structure, and function of U snRNP's. Detailed accounts of the structural aspects of U snRNA's have already been published in recent reviews (Hellung-Larsen 1977; Reddy and Busch 1981; Zieve 1981; Busch et al. 1982), so that they will only be briefly summarized here.

I. Primary and Secondary Structures of U snRNA's

First recognized on the basis of their sedimentation and electrophoretic properties, the small nuclear RNA's were named U snRNA's on account of their relatively high content in uridylic acid residues. Several reviews (see above) have already been dedicated to this rather historical aspect which will intentionally not be covered here.

Actually six major species of these snRNA's are known and numbered U_1 to U_6. They are abundant (10^6 molecules by cell for U_1 and from 10^4 to 5×10^5 molecules by cell for the others) (Weinberg and Penman 1969; Marzluff et al. 1975), metabolically stable (half-lives of up to one cell cycle) (Busch et al. 1982) and were identified in all animal cells from insect to human. They are also present in dinoflagellates (Reddy et al. 1983a) and in plants (Krol et al. 1983). The 5' ends of U_1 to U_5 snRNA's are all capped by trimethylguanosine (Busch et al. 1982) while the nature of the modification of the of U_6 has not yet been completely identified

---

◀ Fig. 1. Localization of protected snRNA sequences within core snRNP's, hnRNP's and "native" snRNP's.
Lines with circles refer to core snRNP sequences protected against micrococcal nuclease (Liautard et al. 1982). *Star lines* refer to sequences protected in hnRNP's (Sri-Widada et al. 1983). *Blank lines* refer to sequences protected in "native" snRNP's (Reveillaud et al. 1984; Lelay-Taha et al. submitted). *Black points* represent nucleotides: The idealized schemes are based on, the following published primary and secondary structure data.

U_1: Branlant et al. 1980, 1981. Mount and Steitz 1981.
U_2: Branlant et al. 1982 and pers. Comm.; Reddy et al. 1981a.
U4 and U_5: Krol et al. 1981a,b.
U_6: Harada, Kato and Nishimura 1980.

The model showing the U_4 - U_6 interaction is from Hashimoto and Steitz (1984). For other published primary and secondary structures, see also: Epstein et al. 1980; Kato and Harada 1981a,b; Mount and Steitz 1982; Reddy, Henning and Busch 1980, 1981; Reddy et al. 1974a-c; Shibata et al. 1975; Wise and Weiner 1980; Myslinski et al. 1984

(Epstein et al. 1980). All contain modified nucleosides. The primary sequence of all U snRNA's are known in several species and secondary structures have been worked out as well for some of them (Fig. 1).

II. Subcellular Localization and Synthesis

The subcellular localization of U snRNA's (especially U_3, U_2, and U_1) has been a matter of controversy. snRNA's were originally found in the nucleus and claimed to be absent from the cytoplasm (Muramatsu et al. 1966; Larsen et al. 1967, 1969; Weinberg and Penman 1968; Dingman and Peacock 1968). Several investigators have then reported the presence of significant amounts of U_2 and U_1 in the cytoplasm (Frederiksen and Hellung-Larsen 1975; Zieve and Penman 1976). Conflicting results have been also obtained with nonaqueous methods of cell fractionation. When lyphilized cells were homogenized and centrifuged in anhydrous glycerol (Eliceiri and Gurney 1978; Gurney and Eliceiri 1980). snRNA species U_3, U_2, and U_1 were almost entirely (> 95%) nuclear. In contrast, Frederiksen et al. (1981) have found considerable amounts of U_2 (57%) and U_1 (43%) snRNA's in cytoplasmic fractions isolated by centrifugation in mixtures of cyclohexane and tetrachloromethane. It seems clear, however, that the presence of snRNA's in cytoplasmic fractions is either transient or merely artifactual and it is now widely accepted that their function, whatever it is, takes place in the nucleus where their mature, steady-state form belongs in vivo. As to U_3 snRNA, its stable localization is clearly nucleolar (Busch et al. 1982).

Notwithstanding their final cellular site of accumulation, at least some snRNA's or their precursors may transiently appear in the cytoplasm. U_3 is found in the cytoplasm within the first few minutes after its transcription, then equally distributed in the cytoplasm and the nucleolus 2 h later and finally almost entirely in the nucleolus after 16 h (Zieve and Penman 1976). Long precursor molecules of U_2, U_1, and U_4 have been detected in the cytoplasm (Eliceiri 1974; Frederiksen and Hellung-Larsen 1975; Zieve and Penman 1976; Eliceiri and Sayavedra 1976; Eliceiri and Gurney 1978; Eliceiri 1979, 1980). As they are recognized by both autoimmune sera and monoclonal Sm antibody (Chandrasekharappa et al. 1983; Madore et al. 1984a,c) these precursor molecules, between one and at least eight nucleotides longer than mature RNA in the case of U_1 (Madore et al. 1984a), assemble with proteins and then are rapidly processed at their 3' end prior to migrating into the nucleus (see Sect.B. VI.8. below). RNA's whose properties are compatible with being U_3 RNA precursors have been also detected (Chandrasekharappa et al. 1983).

III. In Vivo Transcription of U snRNA's

Controversial results have been reported about the type of RNA polymerase involved in U snRNA's synthesis. Using inhibitors like α-amanitin, actinomycin D1, camptothecin, and toyocamycin, Zieve et al. (1977) suggested that RNA polymerase I might be responsible

for the synthesis of U_3, U_2, U_1, U_4, and U_6. As of now, there has been no confirmation of this claim and the debate is more likely restricted to RNA polymerases II and III.

Strong but not quite final evidence suggests that RNA polymerase II, which is clearly established as responsible for the synthesi of hnRNA, also synthesizes U snRNA's (at least U_2, U_3, and U_1) in agreement with the first observation by Ro-Choi et al. (1976). These include: (a) The sensitivity of U_3, U_2, and U_1 snRNA's synthesis to low concentrations of α-amanitin in whole cell systems (Frederiksen et al. 1978; Ro-Choi et al. 1976, Gram Jensen et al. 1979; Eliceiri 1980; Tamm et al. 1980) as well as, in the case of U_1, in isolated nuclei (Roop et al. 1981) or in cell-free systems (Murphy et al. 1982); (b) The inhibition of the synthesis of U_3, U_2, and U_1 at high temperature in a mutant cell line of baby hamster kidney carrying a temperature-sensitive polymerase II (Hellung-Larsen et al. 1980); (c) The relative sensitivity of U_3, U_2, and U_1 snRNA's synthesis to DRB (5,6 dichloro-1-βD-ribofuranosyl benzimidazole) (Hellung-Larsen et al. 1981), a nucleoside analogue known to promote premature termination of hnRNA transcription (Tamm and Kikuchi 1979). No information pertaining to the synthesis of U_4, U_5, and U_6 species in particular is available yet.

Despite the above evidence for the involvement of RNA polymerase II there are at least three kinds of contradictory arguments: (1) Earlier experiments on isolated nuclei (Udvardy and Seifart 1976) and on nuclear preparations supplemented with exogenous RNA polymerase III (Sklar and Roeder 1977) suggested that this latter enzyme catalyzes the synthesis of U_3, U_2, and U_1 snRNA's and this was also concluded from experiments with mouse myeloma cells (Brown and Marzluff 1978); (2) The presence within U_1 of short sequences highly reminiscent of RNA polymerase III transcription initiation signals (Galli et al. 1981); (3) The existence of precursors which, if confirmed, would lead to the unprecedented situation of transcription initiation sites being clearly different from cap sites. In this respect, as well as in others which will be discussed later, it should be pointed out that U_6 seems to stand apart from the other U snRNA's.

IV. Genes and Pseudogenes for U snRNA's

Clones of genomic DNA from various species containing sequences at least partially complementary to U_1 (Roop et al. 1981; Denison et al. 1981; Van Arsdell et al. 1981; Westin et al. 1981; Murphy et al. 1982; Denison and Weiner 1982; Monstein et al. 1982; Manser and Gesteland 1982; Piechaczyk et al. 1982); Lund and Dahlberg 1984; Zeller et al. 1984), U_2 (Denison et al. 1981; Van Arsdell et al. 1981; Westin et al. 1981; Alonso et al. 1983; Van Arsdell and Weiner 1984a), U_3 (Wise and Weiner 1980; Denison et al. 1981; Van Arsdell et al. 1981; Berstein et al. 1983; Marzluff et al. 1983), U_4 (Hammarstrom et al. 1982) and U_6 (Ohshima et al. 1981b; Hayashi 1981) are now available. All earlier reports agree with the conclusion that these sequences, whether true genes or pseudogenes, are abundant, not clustered or in small tandem repeats but,

rather, scattered thoughout the genome. However, it appears from several recent reports that genes, especially true genes, are tandemly repeated. In sea urchin, genes coding for U_1 and U_2 snRNA's are in distinct tandemly repeated clusters of at least 30 kb with a repeat length of 1100 - 1400 base pairs (Card et al. 1982). In *Xenopus laevis*, the U_1 and U_2 genes are organized in tandemly repeated units of 830 bp and 1.8 kb respectively for U_2 and U_1 (Mattaj and Zeller 1983; Zeller et al. 1984). In *D. Melanogaster*, the number of U_2 snRNA genes is very small (four to five) and this seems also to be true for U_1 snRNA (Alonso et al. 1983). In man, genes for U_2 snRNA are organized as a nearly perfect tandem array of 10 to 20 copies per haploid genome, with a basic repeating unit of 6 kb (Van Arsdell and Weiner 1984a; Westin et al. 1984).

Attempts at localizing U snRNA's genes at the chromosomal level are just being started. Using cell fusion techniques, all true human U_1 genes, i.e., producing authentic U_1 snRNA, were found to be localized on chromosome 1 (Lund et al. 1983). By in situ hybridization to the chromosomes of *Drosophila* salivary glands, Saluz et al. (1983) were able to accurately localize genes for U_2, U_3, U_4, and U_6 on specific bands. The number of hybridizing sites (about three) correlates well with the number of genes estimated by Southern-blot analysis.

Clones containing sequences perfectly matching the sequence of the homologous U_1 snRNA were obtained from several species: chicken (Roop et al. 1981), rat (Watanabe-Nagasu et al. 1983), and man (Murphy et al. 1982; Monstein et al. 1982; Manser and Gesteland 1982; Lund and Dahlberg 1984). Sequences more or less closely related to the canonical TATA box were found upstream from the 5' end of the coding sequence in only two (Monstein et al. 1982; Lund and Dahlberg 1984) of the human U_1 clones studied so far, but also in clones for U_6 (Ohshima et al. 1981b), D_2 which is the *Dictyostelium* equivalent of U_3 (Wise and Weiner 1980) and U_1, U_2 of *D. Melanogaster* (Beck et al. 1984). In the case of one human U_1 clone (Murphy et al. 1982), transcription of a 592 bp fragment was analyzed both in vivo and in vitro. Deletions of 5' flanking sequences have demonstrated that the DNA sequence required for in vivo transcription is located at least 100 nucleotides upstream from the beginning of the coding sequence while in vitro transcription in a HeLa cell extract containing RNA polymerase II (Manley et al. 1980) yielded a molecule starting 183 nucleotides upstream of the capping site, thereby suggesting that U_1 snRNA might derive from a larger precursor. Coherent with this finding is recent work from Van Arsdell and Weiner (1984a), who identified two regions of strong homology in human U_1 and U_2 genes. Region II seems to correspond to that found by Murphy et al. 1982. However, it is located 174 bp further upstream from the 5' end of the snRNA sequence in U_1 gene that in U_2 gene, suggesting that its position relative to the initiation site can vary.

Numerous clones have been characterized as pseudogenes which, as pointed out by Denison and Weiner (1982) can include scattered base substitutions, short internal deletions and loss of sequences corresponding to the 3', but never the 5', end of the mature snRNA's. They are scattered throughout the genome in an approximate ratio of 10 to 1 as compared to the true genes.

According to Denison and Weiner (1982), three classes of U_1 pseudogenes have been distinguished. Class I pseudogenes have considerable flanking sequence homology with the authentic U_1 gene and were probably derived from it by gene duplication. Class II and III pseudogenes have flanking sequences unrelated to those of the true gene and short direct repeats flank class III pseudogenes. It was then proposed that class II and III pseudogenes were generated by a RNA-mediated mechanism involving the insertion of U_1 sequence information into a new chromosome locus either as a cDNA copy or as the RNA itself. Similar propositions were made recently for human U_2 (Van Arsdell and Weiner 1984b) and U_3 (Bernstein et al. 1983) pseudogenes which could be generated in vivo by integration of a self-primed cDNA copy of U_3 snRNA at new chromosomal sites.

V. The RNP Status of U snRNA's

All U snRNA's appear not to be free in the nucleus but rather directly associated with specific proteins.

1. Earlier Reports on the Existence of snRNP's

The first indication of a possible RNP structure came from an earlier report by Enger and Walters (1970) demonstrating that small methylated RNA molecules found in the nucleus sedimented more rapidly (10-30S) than the corresponding naked RNA's (5-10S). Rein (1971), studying the snRNA's in cells blocked in mitosis by colcemide, showed that they sedimented as 30 to 180S structures which were partially dissociated by 0.3 M NaCl. In neither case, however, was the nature of these structures further characterized. Raj et al. (1975) were the first to demonstrate that snRNP's represent a real association between snRNA's (in this case U_1 and U_2) and proteins in the form of an RNP structure sedimenting at about 20S. Similar results were obtained by Howard (1978), who isolated by sedimentation of a nucleoplasmic extract of mouse erythroleukemia cells a 10S particle containing snRNA's whose density (1.38 g cm^{-3}) in cesium chloride suggested an RNP nature.

2. Antigenicity of snRNP's versus Natural Autoimmune Antibodies

It has been known for a long time that autoimmune rheumatic diseases like Systemic Lupus Erythematosus (SLE) are characterized by antinuclear antibodies (ANA). Antibodies with specificities against DNA, histones, and other structures have been recognized in sera from these patients. The nature of these antibodies and their immunological as well as clinical significance have been extensively dealt with in two exhaustive reviews by Tan (1979, 1982) to which the reader is referred for a more detailed account. Among this wide class of autoimmune rheumatic disorders, a new distinct clinical entity called mixed connective tissue disease (MCTD) has been described by G.C. Sharp et al. (1972). High titers of an antibody against a nuclear antigen sensitive to both RNAse and trypsin (Northway and Tan 1972) and therefore

Table 1. Proteins associated with U snRNA's

Method		Cell	snRNA	Proteins		Reference
				M.W. < 20 k	M.W. > 20 k	
Immuno-precipitation with	Anti-Sm	Ehrlich ascites	$U_1 - U_6$	11 k, 12 k, 13 k, 16 k	22 k, 28 k, 33 k	(a)
	Anti-RNP		U_1			
	Anti-SM	F. eryth. leuk.	$U_1 - U_2$	3(13 k), 16 k	20 k, 25 k, 26 k, 31 k	(b)
	Anti-RNP	HeLa	U_1	10 k, (12-14 k)	26 k, 30 k	(c)
	Anti-Sm		$U_1 - U_6$			
	Anti-RNP		U_1			
	Anti-Sm	HeLa	$U_1 - U_6$	4(9-14 k)	20 k, 25 k, 28 k	(d)
	Anti-RNP		U_1			
	Anti-(U_2)RNP	HeLa	U_2	11 k, 12 k, 13 k, 16 k	28 k, 32 k	(o)
Immuno-affinity columns on	Anti-RNP	Rat liver	N.D.	13 k	30 k	(e)
		Calf thymus	N.D.	5(12-13 k)	2(30 k), 65 k	(f)
	Anti-Sm	Calf thymus	N.D.			
	Mixed	Rabbit thymus	(U_2-U_1) ?	10 k, 11 k, 13 k	25.5 k, 27.5 k	(g/h)
	Anti-Sm/ Anti-RNP				31 k, 35 k, 73 k	

Solid phase radioimmunoassay revealed with	Anti-Sm	Calf thymus		13 k		(f)
	Anti-RNP				30 k, 65 k	
Protein blots revealed with	Anti-RNP	Rabbit thymus			40 k, 2(70 k)	(i)
		Rat liver			30 k	(j)
		Rat liver		13 k		(j)
		Rabbit thymus				
	Anti-Sm	HeLa		13 k	2(16-27 k)	(h)
	Anti-RNP and Anti-RNP (monoclonal)	Rabbit thymus			73 k	(h)
		HeLa			69 k	
	Anti-(U_2)RNP	HeLa			32 k	(o)
Biochemical purification of snRNPs	CsCl+sarcosyl Centrifugation	HeLa	U_1	4/5(9-14 k)		(k)
			U_1 to U_6	4/5(9-14 k)		(l)
	Ion exchange chromatography	F.eryth.leuk.	U_1	11 k, 12 k, 13 k, 16 k	22 k, 28 k, 33 k, 68 k	(m)
		HeLa			22 k, 28 k, 29 k, 68 k	
		id	$U_2U_4U_5U_6$	11 k, 12 k, 13 k, 16 k	28 k, 32 k	(m)
		HeLa	U_1	8.5 k, 9.1 k, 10.2 k 12.3 k, 17.5 k	21.5 k, 23 k, 30 k 67 k	(n)
		id	U_2	8.5 k, 9.1 k, 10.2 k, 12.3 k	21.5 k, 23 k, 27 k	(n)

(a) Lerner and Steitz, 1979; (b) Matter et al., 1982; (c) Barque et al., 1981; (d) Assens et al., 1982a; (e) Douvas et al., 1979; (f) Takano et al., 1981; (g) White, Gardner and Hoch, 1981; (h) Billings and Hoch, 1983; (i) White and Hoch, 1981; (j) Douvas, 1982; (k) Sri-widada et al., 1982; (l) Brunel et al., 1981; (m) Hinterberger, Peterson and Steitz, 1983; (n) Kinlaw, Robberson and Berget, 1983; (o) Mimori et al., 1984.

designated RNP, are usually found in this disease, although they are not strictly diagnostic for it. This RNP antigen has been isolated from rat liver nuclei and characterized using affinity columns containing the antibody (Douvas et al. 1979). Two major polypeptides of mol. wt. 30 k and 13 k were identified in the protein moiety of this RNP. However, the RNA complement was not identified as snRNA's at this time. Another nuclear antigen, called Sm and characterized by its relative resistance to both trypsin and RNAse, had been identified even earlier and shown to be diagnostic for SLE (Tan and Kunkel 1966).

An important breakthrough came with the discovery by M.R. Lerner and Steitz (1979) that both anti-RNP and anti-Sm antibodies from SLE patients precipitated snRNP's from Ehrlich ascites tumors and HeLa cells. They definitively established that the anti-RNP antibody is directed against U_1 snRNP and accordingly proposed (M.R. Lerner and Steitz 1981) to rename it anti-U_1 RNP. Seven polypeptides with mol. wts. ranging from 12 to 35 k, designated A to G were found to be immunoprecipitated along with U_1 snRNA. Anti-Sm immunoprecipitated all the U_2, U_1, U_4, U_5, and U_6 snRNA's and the same seven polypeptides were found in the precipitate indicating that all these U snRNA's have a common set of associated proteins. This was later confirmed using a monoclonal antibody (E.A. Lerner et al. 1981). These proteins are clearly distinct from either histones or the major hnRNP polypeptides (M.R. Lerner and Steitz 1979). Two new sera have been described recently. The first one, discovered in a patient with scleroderma-polymyositis overlap syndrome, contains antibodies directed against U_2 snRNP. Along with the U_2 snRNA, six polypeptides are precipitated among which is a U_2 specific component (Mimori et al. 1984) (see Sect. V.5 and Table 1). The second one was found on scleroderma patients and recognizes, on protein blots, a 34 k component probably associated with U_3 snRNA (Reddy et al., pers. comm.). This seems to be the first demonstration that U_3 snRNA exist in a snRNP form.

3. Evidence for A core snRNP Structure

At the same time when immunological methods were introduced, a strictly biochemical approach to the purification of snRNP's was being developed in our laboratory. This procedure starts with purified hnRNP's to which at least a major fraction of snRNA's and, therefore, snRNP's are bound (see below). A ribonucleoprotein complex whose RNA complement consists exclusively of small nuclear RNA's has been purified from particles containing heterogenous nuclear RNA (hnRNP) from HeLa cells. This was accomplished by taking advantage of their ability to band at a density of about 1.43 g cm^{-3} in plain cesium chloride, as well as in cesium chloride gradients containing 0.5% sarcosyl without prior aldehyde fixation. After these two steps of equilibrium density centrifugation, these snRNP's were still largely contaminated by free proteins (and especially phosphoproteins). A final step of purification by velocity sedimentation in a sucrose gradient containing 0.5% M salt and 0.5% sarcosyl was efficient in completely eliminating all free proteins. The sedimentation coefficient measured in the presence of sarcosyl has been shown to be 11 to

12 S. All U snRNA's. U_2, U_1, U_4, U_5, and U_6, were found in these purified snRNP's, although the recovery was not as good for all of them, that of U_6 being noticeably lower. Electrophoresis of associated proteins revealed four of five bands with migration corresponding to mol. wt. ranging between 9 and 14 k (Brunel et al. 1981; Sri-Widada et al. 1981).

A further step was taken by the purification of the individual U_1 snRNP (Sri-Widada et al. 1982). This was obtained by combining the above procedure with the preferential release of this particle from hnRNP's under isotonic conditions as observed by other authors (Gallinaro and Jacob 1979; Zieve and Penman 1981). This U_1 snRNP contained no other snRNA and had a protein complement identical to that of the mixture of all snRNP's (Brunel et al. 1981). This finding did away with the possibility that each snRNA could be bound to only one of these small mol. wt. polypeptides and demonstrated that all individual snRNP's isolated by this procedure have the same protein composition and therefore that the specificity of their function depends only on their snRNA component and/or their cognate proteins which have been lost during the purification. Our rather drastic purification procedure has therefore yielded a "core particle" structure which is common to all nucleoplasmic snRNP's. Due to their simplified protein complement, these core particles can be expected to be more easily amenable to structural studies.

Although certainly far from being native, these core snRNP's retain an important part (46%) of their antigenicity towards anti-Sm antibodies (Assens et al. 1982b). Moreover, digestion of a great part of the RNA moiety does not abolish this antigenicity. This indicates that at least some of the Sm antigenic determinants are among the 4 - 5 small polypeptides with mol. wt. between 9 and 14 k, as confirmed by electroblotting techniques in which a 13 k protein is shown to react with anti-Sm (White and Hoch 1981; Takano et al. 1981). The antigenicity towards RNP antibody is nearly entirely lost in "core snRNP's" indicating that the RNP determinant must depend on proteins specifically interacting with U_1 and which are lost during the purification. The previously suggested (Takano et al. 1981) presence of RNP determinants among the high mol. wt. proteins was also confirmed by others (White and Hoch 1981; Billings et al. 1982) using electroblotting techniques (see below).

4. Attempts at Isolating Native snRNP's

Several recent reports were aimed at purifying snRNP's away from other cellular components on one hand and at separating individual particles on another hand. These procedures were based either on conventional fractionation methods or on the use of antibodies against the unusual cap ($m_3^{2,2,7}G$) which is found only in U snRNA's except U_6.

Along the first line, a procedure was developed by Hinterberger et al. (1983) to fractionate snRNP's from mouse Friend erythroleukemia and human HeLa cells. A nuclear extract obtained at 0.4 M salt was submitted to a series of gel filtration, ion ex-