

# Molecular Cloning, Expression, and Chromosome 19 Localization of a Human Ro/SS-A Autoantigen

Daniel P. McCauliffe, F. Anthony Lux, Tsu-San Lieu, Ignacio Sanz, Jeffrey Hanke, Marianna M. Newkirk, Linda L. Bachinski,\* Yasuhiko Itoh,† Michael J. Siciliano,\* Morris Reichlin,‡ Richard D. Sontheimer, and J. Donald Capra  
Departments of Dermatology, Internal Medicine, and Microbiology, University of Texas, Southwestern Medical Center, Dallas, Texas 75238; \*M. D. Anderson Hospital and Tumor Institute, Houston, Texas 77030; and †Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104

## Abstract

Ro/SS-A antibodies are found in a number of human autoimmune disorders including Sjogren's syndrome and several systemic lupus erythematosus-related disorders. These heterogeneous autoantibodies are known to recognize several distinct cellular antigens. With synthetic oligonucleotides corresponding to amino acid sequence information we have isolated a full-length cDNA clone which encodes a human Ro ribonucleoprotein autoantigen. The 1,890-base pair clone contains an open reading frame that encodes a 417-amino acid, 48-kD polypeptide that migrates aberrantly at 60 kD by SDS-PAGE. Rabbit antibodies raised against this protein's recently described amino-terminal epitope react with a previously identified 52-kD human Ro protein and immunoprecipitate the human cytoplasmic RNAs. Ultraviolet light cross-linking studies suggest that this Ro protein binds each of the four major human cytoplasmic RNAs. The deduced amino acid sequence is 63% homologous to an *Onchocerca volvulus* antigen. Southern filter hybridization analysis indicates that this gene is not highly polymorphic and exists as a single copy in the human genome. Chromosomal localization studies place this gene on the short arm of chromosome 19 near the gene encoding the low density lipoprotein receptor. (*J. Clin. Invest.* 1990; 85:1379-1391.) autoantigen • human cytoplasmic RNA • onchocerciasis • ribonucleoprotein • Sjogren's syndrome • systemic lupus erythematosus

## Introduction

The Ro autoantigens are of clinical interest in that antibodies directed against them are found in the majority of patients with primary Sjogren's syndrome, subacute cutaneous lupus erythematosus, neonatal lupus erythematosus, anti-nuclear antibody-negative lupus erythematosus and systemic lupus erythematosus-like disease secondary to homozygous C2 or C4

Dr. Newkirk's present address is Department of Medicine, McGill University, Montreal, Quebec H3G 1A4, Canada. Dr. Sanz's present address is Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX 78284. Dr. Hanke's present address is Department of Molecular Genetics, Pfizer, Inc., Groton, CT 06340.

Address reprint requests to Dr. Capra, Department of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75238-9048.

Received for publication 22 February 1989 and in revised form 14 November 1989.

*J. Clin. Invest.*

© The American Society for Clinical Investigation, Inc.

0021-9738/90/05/1379/13 \$2.00

Volume 85, May 1990, 1379-1391

complement deficiency (1-6). The significance of these antibodies is uncertain, but there is substantial evidence that they have a major role in the pathogenesis of disease (7).

In 1969 Clark et al. (8) first demonstrated the presence of these antibodies in sera from patients with systemic lupus erythematosus and Sjogren's syndrome. These sera precipitated antigens from cell and tissue extracts by immunodiffusion. They designated these the Ro antibody and Ro antigen, respectively. In 1975 Alspaugh and Tan (9) similarly described the presence of three types of precipitating antibodies in Sjogren's syndrome patient sera; these authors referred to the three types as SS-A, SS-B, and SS-C. SS-A antibodies were later shown to be immunologically equivalent to the Ro antibodies (10).

In 1981 Lerner et al. (11) demonstrated that human Ro antisera precipitated a novel class of small RNAs designated the human (h)<sup>1</sup> cytoplasmic (Y) RNAs, although there is now evidence that the hY RNAs may also be intranuclear (12). The immunoprecipitation of the hY RNAs required the presence of protein; thus it was deduced that the hY RNAs are bound to a protein antigen which in 1984 was shown to be a 60-kD protein (13). We subsequently isolated a 60-kD protein from a human B-cell line which reacts with human Ro antisera. The amino terminus of this protein was sequenced and synthetic peptides corresponding to this sequence are reactive to Ro antisera (14). These data as well as a significant body of other experimental data at that time suggested that Ro antibodies were directed at a single 60-kD ribonucleoprotein (RNP) that was designated the Ro protein, Ro RNP, or Ro (auto)antigen. Recently, however, it has been shown that Ro antisera react with at least four immunologically distinct "Ro" proteins of three different molecular masses: a 52- and 54-kD protein and two 60-kD proteins (15). It is not known whether all of these proteins bind the hY RNAs or if they are structurally or functionally related. Further molecular characterization of these proteins will help address these questions and should further clarify the cellular function(s) and the pathogenic role(s) of these protein autoantigens. Here we report the molecular characterization of a human Ro RNP through the cloning and analysis of its cDNA.

## Methods

### *Protein purification and sequence analysis*

The Ro protein was purified from the human Wil-2 cell line (an Epstein-Barr virus-transformed lymphoblastoid B-cell line) as previously described (14). *Staphylococcus aureus* V8 (Boehringer Mannheim Biochemicals, Indianapolis, IN) and cyanogen bromide (Sigma Chemical

1. *Abbreviations used in this paper:* CHO, Chinese hamster ovary; h, human; LDLR, LDL receptor; poly-A, polyadenylated; RNP, ribonucleoprotein; UVC, short wavelength ultraviolet light; Y, cytoplasmic.

Co., St. Louis, MO) cleavage fragments were generated according to established protocols (16) and sequenced using a model 470A protein sequencer/model 120A PTH analyzer (Applied Biosystems, Inc., Foster City, CA), as previously described (14).

#### *Synthetic oligonucleotide construction*

A codon utilization table was employed to convert the amino acid sequence into its most probable nucleic acid sequence (17). The oligonucleotides were synthesized using a model 380B DNA synthesizer (Applied Biosystems).

#### *cDNA library construction*

Total RNA was isolated from the Wil-2 cell line by the guanidinium method and enriched for the polyadenylated (poly-A) fraction with an oligo(dT)-cellulose column (18). cDNA was made from the poly-A-enriched fraction with the cDNA synthesis system (Bethesda Research Laboratories, Gaithersburg, MD). The cDNA was dG-tailed with dGTP and terminal transferase and ligated into similarly dC-tailed pGEM plasmid DNA with T4 DNA ligase (19). DH5 *Escherichia coli* competent cells (Bethesda Research Laboratories) were transformed with the cDNA-pGEM ligation mixture and a cDNA library was constructed (18). A human hybridoma cDNA library was similarly constructed.

The enzymes used in the various recombinant nucleic acid techniques were obtained from Promega Biotec, Madison WI, or Pharmacia, Inc., Piscataway, NJ, unless stated otherwise.

#### *cDNA isolation*

The synthetic oligonucleotides were radiolabeled and hybridized with nitrocellulose filters to which the cDNA-containing bacterial colonies had been fixed (18). A single colony containing a 1.2-kb cDNA insert was isolated. Later this 1.2-kb cDNA was radiolabeled and used to screen a human hybridoma cell cDNA library, and a single 1.9-kb cDNA insert was isolated.

#### *cDNA characterization*

**Restriction enzyme analysis.** The 1.9-kb cDNA was digested with various restriction enzymes and the restriction fragments were analyzed by Southern filter hybridization with radiolabeled synthetic oligonucleotides (19).

**Sequencing.** Several of the cDNA restriction fragments were electroeluted from a 1% agarose gel and subcloned into M13mp18 and M13mp19 plasmid vectors (Boehringer Mannheim Biochemicals) and single-stranded DNA complementary to both strands of cDNA was generated (20). This DNA was sequenced by the Sanger dideoxy method with [ $\alpha$ -<sup>35</sup>S]dATP and modified T7 DNA polymerase (Sequenase) according to the manufacturer's recommendations (United States Biochemical Corp., Cleveland, OH).

**Northern filter hybridization.** Total RNA and poly-A-enriched RNA from several human white blood cell lines (obtained as outlined above) were electrophoresed in a 1% agarose-formaldehyde gel, electrophoretically transferred to Zeta-Probe nylon-reinforced support membrane according to the manufacturer's guidelines (Bio-Rad Laboratories, Richmond, CA), hybridized with radiolabeled cDNA, and then washed at 65°C in 0.25× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate, pH 7.0) and 0.1% sodium dodecyl sulfate (SDS).

**Southern filter hybridization.** 15 µg of human genomic DNA was digested with various restriction enzymes, separated by 0.6% agarose gel electrophoresis, and transferred to a nitrocellulose support membrane where it was hybridized with radiolabeled full- or partial-length cDNA. Membranes were washed in a 0.5× SSC, 0.1% SDS solution at 65°C (19).

#### *cDNA expression analysis*

The bacteriophage T7 RNA polymerase/promoter system developed by Tabor and Richardson (21) was utilized for cDNA expression. The 1.9-kb cDNA was amplified with the polymerase chain reaction with two degenerate 24-base oligonucleotides in order to incorporate an

Nde I restriction enzyme site at the ATG translation start site and 64 base pairs beyond the stop codon (22). The resultant polymerase chain reaction-derived fragment was then ligated into the Nde I site of the pT7-7 expression vector which was obtained through Dr. S. Tabor, Department of Biological Chemistry, Harvard Medical School, Boston, MA. TG1 *E. coli* cells were transformed with the resultant ligation mixture and a colony which had the cDNA insert in the correct orientation with preservation of the Nde I sites, as determined by restriction enzyme analysis, was chosen for further study. Cells from this colony were subsequently transformed with the pGP1-2 vector. The cDNA was translated in the presence of rifampicin and [<sup>35</sup>S]methionine, the cells were harvested and samples of bacterial lysate were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (21).

#### *Radiolabeling and autoradiography*

Synthetic oligonucleotides were end labeled with [ $\gamma$ -<sup>32</sup>P]ATP using T4 polynucleotide kinase (19). cDNA was radiolabeled using the heximer extension method with heximer primers (Pharmacia, Inc.), [ $\alpha$ -<sup>32</sup>P]-dCTP and *E. coli* DNA polymerase I (Klenow fragment) (19). Radionucleides were obtained from New England Nuclear Corp., Boston, MA.

Filters and dried gels were exposed to X-OMAT-AR film (Eastman Kodak Co., Rochester, NY) between intensifying screens for an optimal period of time, and the film was then developed on a QX-60A film processor (Konica Medical Corp., Wayne, NJ).

#### *Chromosomal localization*

Somatic cell hybrid clone panels were formed by polyethylene glycol-mediated fusion of human lymphocytes to Chinese hamster ovary (CHO) cell lines defective for various DNA repair capabilities. Cytogenetic analysis was used to determine the presence or absence of human chromosomes in each of the hybrid clones. Due to frequent human chromosomal alterations in these clones, the human chromosomes were more definitively detected by analysis of isoenzyme and DNA markers (23, 24). Probes for complement component 3 (C3) and low density lipoprotein receptor (LDLR) were used to identify the short arm of chromosome 19.

#### *Computer-based sequence analysis*

The 1.9-kb cDNA nucleic acid sequence and its deduced amino acid sequence were analyzed for homologies to other published sequences. This was done with the University of Wisconsin Computer Genetics Group's Genetics Analysis software and the FASTA/FASTP programs. The nucleic acid sequence was compared to the European Molecular Biology Lab database-Version 13 (April 1988) and the Genbank database-Version 56 (July 88). The protein sequence was compared to the National Biomedical Research Foundation database-Version 13 (March 1988) (25).

#### *Deglycosylation analysis*

The purified Ro protein was digested with neuraminidase, endo- $\alpha$ -N-acetylgalactosaminidase, and glycopeptidase F according to the manufacturer's recommendations (Boehringer Mannheim Biochemicals), and then subjected to SDS-PAGE.

#### *Rabbit antiserum immunopurification and immunoblotting*

A female New Zealand White rabbit was immunized with a synthetic peptide corresponding to amino acids 6–19 of the mature Ro protein as previously described (14). 10% polyacrylamide gels with 4.5% polyacrylamide stacking gels were run using the discontinuous buffer method previously described (26). Samples were denatured by boiling for 10 min in the presence of 5% 2-mercaptoethanol (Sigma Chemical Co.). The 1.5-mm gels were run at 20 mA constant current on a mini-gel apparatus model SE-200 (Hoefer Scientific Instruments, San Francisco, CA). Protein that had been electrophoresed was blotted onto nitrocellulose membranes and incubated with antibodies in the Western blot procedure as previously described (27). For elution of

affinity-purified antibodies, a vertical electrophoresis apparatus (14 × 16 cm) (model 2001, Pharmacia LKB, Piscataway, NJ) was used. The method for the elution of antibodies has been described (27). Briefly, the strips were washed thoroughly and the antibody eluted with 3 M NaSCN. The eluate was dialyzed to PBS (Tris-buffered saline with Tween 20 for Western immunoblotting or IPP buffer [10 mM Tris-HCl, 500 mM NaCl, 0.1% NP-40, pH 8.0] for RNA immunoprecipitation) by repeated dilution and concentration using a Centriprep concentration device as recommended by the manufacturer (Amicon Corp., Lexington, MA).

### RNA binding studies

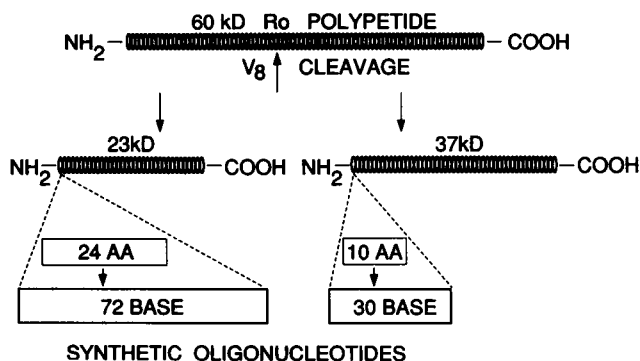
**Immunoprecipitation.** Immunoprecipitation, electrophoresis, and silver staining of the hY RNAs was performed according to a well-established protocol (28).

**Short wavelength ultraviolet light (UVC) cross-linking.** The purified Ro protein was irradiated with 25–125 mJ/cm<sup>2</sup> of UVC using four Phillips TUV-15W germicidal lamps (Gulf Coast Electric, Houston, TX). The Ro protein was analyzed by Western immunoblotting as previously described (14). A portion of the cross-linked purified extract was also incubated with ribonuclease A (1 mg/ml overnight at room temperature) before immunoblot analysis.

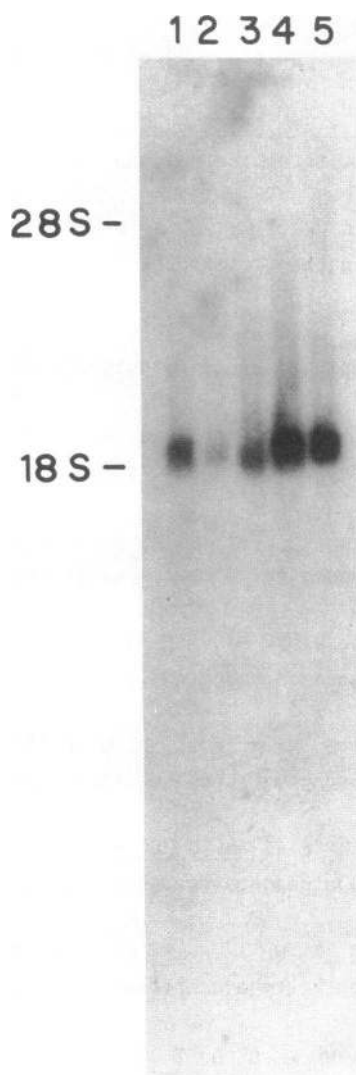
## Results

**Amino acid sequencing and synthetic oligonucleotide construction.** We isolated a 60-kD protein with Ro antigenic activity from the Epstein-Barr virus-transformed human Wil-2 B-cell line and subjected it to a limited *S. aureus* V8 protease digestion. This produced 23- and 37-kD fragments which were identified by SDS-PAGE. The amino-terminal end of the 60-kD protein and its 23- and 37-kD fragments were sequenced and this information was used to construct two non-degenerate synthetic oligonucleotides (Fig. 1). The amino terminus of the 23-kD fragment was identical to that of the 60-kD protein.

**cDNA isolation, bacterial expression, and sequence analysis.** A single 1.2-kb cDNA clone was isolated from the Wil-2 cell cDNA library with the two synthetic oligonucleotides. This clone was characterized by restriction enzyme analysis and sequenced. The cDNA encoded the previously determined amino acid sequences, however, this reading frame contained no termination codon, indicating that the cDNA



**Figure 1.** Synthetic oligonucleotide construction. The native 60-kD Ro polypeptide was purified from a Wil-2 cell extract and subjected to a limited *S. aureus* V8 protease digestion which cleaved the 60-kD polypeptide into 23- and 37-kD fragments. The amino terminus of each fragment was sequenced and this amino acid (AA) sequence information was converted into the most probable nucleic acid sequence for the construction of two nondegenerate synthetic oligonucleotides.



**Figure 2.** Northern filter hybridization. The 1.9-kb Ro cDNA was radiolabeled and hybridized with RNA from several different human white blood cell lines. Lanes 1–4 were each loaded with 30 μg of total RNA isolated from a human hybridoma cell line (lane 1), normal peripheral blood leukocytes (lanes 2 and 3), and the Wil-2 B-cell line (lane 4). Lane 5 was loaded with 5 μg of poly-A selected RNA from the Jurkat T-cell line. A 2-kb RNA species is identified. The positions of the 28-S and 18-S ribosomal RNA bands are indicated as references for RNA size.

was truncated at the 3' end. Northern filter hybridization with this cDNA (Fig. 2) identified a 2-kb RNA species but no 1.2-kb species, confirming that our cDNA was abbreviated. A human B-cell hybridoma cDNA library was subsequently screened with the 1.2-kb cDNA, and a single 1.9-kb cDNA clone was isolated and sequenced. The first 1,238 basepairs of this clone are identical to the entire sequence of the 1.2-kb clone. The 1.9-kb clone contains 1,890 basepairs which include a single 1,251-base open reading frame beginning with an AUG start site at position 67 as part of a putative Kozak ribosomal translation initiation site and ending with the termination codon UAG (Fig. 3) (29). The sequence AUUAAA (Fig. 3) is a putative polyadenylation signal (30), but there is not a typical poly-A sequence between this signal and the end of the cDNA sequence, suggesting that this 1.9-kb cDNA may be minimally truncated.

The deduced polypeptide has a molecular mass of 48 kD which includes a 17-amino acid hydrophobic leader segment that is not present in the purified mature protein. The *M<sub>r</sub>* of the deduced polypeptide without the leader segment is ~ 14 kD less than that of the "60"-kD protein as measured by SDS-PAGE. The amino acid sequence contains one potential site for NH<sub>2</sub>-linked glycosylation (amino acid position 344 in Fig.

ccgtactgcagagccgctgccggagggctgttttaaggccgctgtgccgccccctcgccccgcatgctgctatccgtgccgctgctgctcgccctcc 12 100  
G L A V A E P A V Y F K E Q F L D G D G W T S R W I E S K H K S D 45  
 tcggcctggcctgcccagcccgcgttacttcaaggagcagtttctggacggagagcgggtggacttcccgtggatcgaatccaacacaagtcaga 200  
 F G K F V L S S G K F Y G D E E K D K G L Q T S Q D A R F Y A L S 78  
 ttttgcaaatcgttctcagttccggcaagttctacggtgacgaggagaaagataaaggtttgacagacaagccaggatgcacgcttttatgctctgtcg 300  
 A S F E P F S N K G Q T L V V Q F T V K H E Q N I D C G G G Y V K L 112  
 gccagtttcgagcctttcagcaacaaaggccagacgctgggtggcagttcacggtgaaacatgacgagaacatcgactgtggggggcggctatgtgaagc 400  
 F P N S L D Q T D M H G D S E Y N I M F G P D I C G P G T K K V H 145  
 tgtttctaatagtttggaccagacagacatgcacggagactcagaatacaacatcatgtttgggtcccacatctgtggcctggcaccagaaggttca 500  
V I F N Y K G K N V L I N K D I R C K D D E F T H L Y T L I V R P 178  
 tgtcatcttcaactacaagggcaagaacgtgctgatcaacaaggacatccgttgcaaggatgatgagtttacacacctgtacacactgattgtgcgcca 600  
 D N T Y E V K I D N S Q V E S G S L E D D W D F L P P K K I K D P D 212  
 gacaacacctatgaggtgaagattgacaacagccaggtggagtcggctccttggaagacgattgggacttctgccaccaagaagataaaggatcctg 700  
 A S K P E D W D E R A K I D D P T D S K P E D W D K P E H I P D P 245  
 atgcttcaaaaccggaagactgggatgagcgggcaagatcgatgatccacagactccaagcctgaggactgggacaagcccagcatatccctgacc 800  
 D A K K P E D W D E E M D G E W E P P V I Q N P E Y K G E W K P R 278  
 tgatgctaagaagcccaggactgggatgaagagatggacggagagtggaacccccagtgattcagaaccctgagtacaagggtgagtggaagcccgg 900  
Q I D N P D Y K G T W I H P E I D N P E Y S P D P S I Y A Y D N F G 312  
 cagatgcacaaccagattacaagggcacttggatccaccagaattgacaacccccagtgatttctcccgatcccagtatctatgcctatgataactttg 1000  
 V L G L D L W Q V K S G T I F D N F L I T N D E A Y A E E F G N E 345  
 gcgtgctggcctggacctctggcaggtaagttctggcaccatcttgacaacttctcatcaccaacgatgaggcatacgtgaggagtttgcaacga 1100  
 T W G Y T K A A E K Q M K D K Q D E E Q R L K E E E E D K K R K E 378  
 gacgtggggcgtacaagaaggcagcagagaacaatgaaggacaacaggacgaggagcagaggcttaaggaggaggaagaagacaagaacgcaagag 1200  
 E E E A E D K E D D E D K D E D E E D E E D K E E D E E E D V P G Q 412  
 gaggaggaggcagaggacaaggaggatgatgaggacaagatgaggatgaggaggatgaggaggacaaggaggaagatgaggaggaagatgtccccggc 1300  
 A K D E L 417  
 aggccaaggacgagctgtagagaggcctgcctccagggtggactgaggcctgagcgtcctgccgagagcttgccgcgcaataatgtctctgtgag 1400  
 actcgagaactttcattttttccaggctggttcggatttgggtggatttgggtttgttcccctcctccacttccccccccctccccgcccttt 1500  
 ttttttttttttaactggatatttatcctttgattctccttcagccctcaccctgggttctcatctttcttgatcaacatctttcttgctctgtgc 1600  
 cccttctctcatctcttagctccccccaacctggggggcagtggtgtggagaagccacaggcctgagatttcattctgctctccttcttgagcccagag 1700  
 gagggcagcagaagggggtggtgtctccaacccccagcactgaggaagaacggggcttctctatttccccctcccttctccccctgccccaggact 1800  
 gggccacttctgggtggggcagtggttcccagattggctcacactgagaatgtaagaactacaacaaaatttctattaataattttt 1890

3) but deglycosylation analysis of the purified 60-kD protein shows no evidence of NH<sub>2</sub>- or COOH-linked glycosylation (data not shown). Bacterial expression of the cDNA produced a protein which migrates at ~ 60 kD by SDS-PAGE, which is inclusive of the ~ 2-kD hydrophobic leader segment (data not shown). A similar *M<sub>r</sub>* discrepancy between that encoded by the cDNA and that measured by SDS-PAGE, has been reported in the cloning of several other proteins as shown in Table I. These proteins have a highly charged region in common which may cause retarded gel migration and thus an overestimation of the *M<sub>r</sub>*'s by SDS-PAGE. Our protein has a highly charged region between residues 358 and 408 where 47 of 51 residues are strongly charged (36 are negatively charged and 11 are positively charged). The calculated isoelectric point of this polypeptide is 4.14 which closely approximates our value of 4.67 measured from the native purified protein (36).

This protein contains three different sets of repeating sequences (Fig. 4) which may have arisen from internal replications and may be of functional importance. The first set has 82% of its nucleic acid sequence and 82% of its amino acid sequence conserved. The second set has 78% nucleic acid and 71% amino acid sequence conservation, and the third set has 83% nucleic acid and 80% amino acid sequence conservation. This protein also contains several PEST regions (Fig. 4, *upper panel*) as proposed by Rogers et al. (37): these so-called PEST regions are rich in the amino acids proline (P), glutamic acid (E), serine (S), and/or threonine (T), and to a lesser extent aspartic acid (D). These regions are thought to make a protein susceptible to rapid intracellular degradation.

The primary structure of this protein is unique but computer-based analysis of the nucleic and amino acid sequence shows striking 63% amino acid sequence homology to an antigen recently partially cloned from *Onchocerca volvulus* (Fig. 5) (38). 14 of the 15 amino acids at the amino terminus of the mature protein are identical with the 15-amino terminal residues recently deduced from the purified rabbit calregulin protein (39). The negatively charged carboxy-terminal region has some minor amino acid sequence homology with a number of other proteins of diverse origin and function. The most striking of these homologies is with residues 44–80 of the 17-kD subunit of yeast ubiquinone cytochrome *c* reductase (40), where 18 of 36 residues are identical and 9 of the remaining 18 residues are a Asp for Glu or a Glu for Asp substitution with Ro residues 379–415. Within this negatively charged region is a sequence (residues 384–395) which has some homology to previously described calcium binding domains (41), although this sequence is not flanked by hydrophobic residues as in the classic "EF hand" calcium-binding domain (42). The carboxy-terminal sequence Lys-Asp-Glu-Leu (KDEL) follows the negatively charged region (Fig. 4, *upper panel*) and is identical to the carboxy signal sequence which has been shown to be crucial for the retention of several proteins in the endoplasmic reticulum (43). These other proteins likewise have a highly negatively charged region just proximal to the KDEL sequence. The 17-amino acid hydrophobic leader sequence (Fig. 4, *upper panel*) is similar to that of a number of other precu-

Table I. Proteins with Highly Charged Regions and Retarded SDS-PAGE Gel Migration

Protein	<i>M<sub>r</sub></i>		Reference
	cDNA encoded	SDS-PAGE determined*	
	<i>kD</i>		
Amphibian nucleoplasmin	22	33	32
Amphibian N1/N2 histone-binding protein	65	110	34
Bovine chromogranin A	53	75	33
Yeast GCN4 transcription activator	31	45	31
Human nuclear RNP particle C2	32	40–44	57
Human U1–70-kD small nuclear RNP	52	70	35

\* Without any carbohydrate moiety.

ror proteins and indicates that this protein is transported into the endoplasmic reticulum (44).

There is no striking sequence similarity to other RNA-binding proteins, including another recently sequenced Ro cDNA (45). There is no major homology to the RNP consensus sequence (46) and there are no zinc finger (47) or leucine zipper (48) nucleic acid-binding motifs.

Chou-Fasman computer-based secondary structure analysis (49) predicts a complex secondary structure (Fig. 6), which includes several helix-turn-helix units centered around residues 57, 70, 210, 233, and 246. Three of these units are found within the sequence triplications between residues 207 and 300 (Fig. 4, *upper panel*), one unit per each sequence repeat. There are also several  $\beta$ -sheet-rich areas between residues 1 and 17, 144 and 186, and 285 and 333. The carboxy-terminal residues 349–417 are predicted to have an  $\alpha$ -helical array.

Kyte-Doolittle hydrophobic analysis (50) predicts a strongly hydrophobic leader segment and several smaller regions of hydrophobicity, including an area just proximal to the negatively charged carboxy-terminal residues which could be a membrane-spanning region. This analysis also predicts several strongly hydrophilic domains particularly between amino acids 210–300 and 350–417. Residues 210–300 include the first two sets of sequence triplications and residues 350–417 span the negatively charged carboxy end which includes the set of sequence duplications.

Jameson-Wolf antigenicity analysis (51) (Fig. 6) predicts the location of several potential epitopes including the previously characterized epitope at the amino terminus of this polypeptide, (synthetic peptide 6–19) (14), and a recently characterized epitope (synthetic peptide 171–194) corresponding to residues 171–194 (52).

*Southern filter hybridization analysis and chromosomal localization.* Southern filter hybridization of Eco RI digested genomic DNA from 10 normal individuals shows a single

Figure 3. The 1.9-kb Ro cDNA nucleic acid and encoded amino acid sequence. The 1,890-base coding strand encodes a 417-amino acid polypeptide which includes four previously determined amino acid sequences (*underlined*) from sequencing the native protein and cyanogen bromide and *S. aureus* V8 cleavage products. The eukaryotic ribosomal consensus sequence for the initiation of translation is boxed and the putative polyadenylation signal is overlined. These sequence data are available from EMBL/GenBank/DDBJ under accession number M32294.

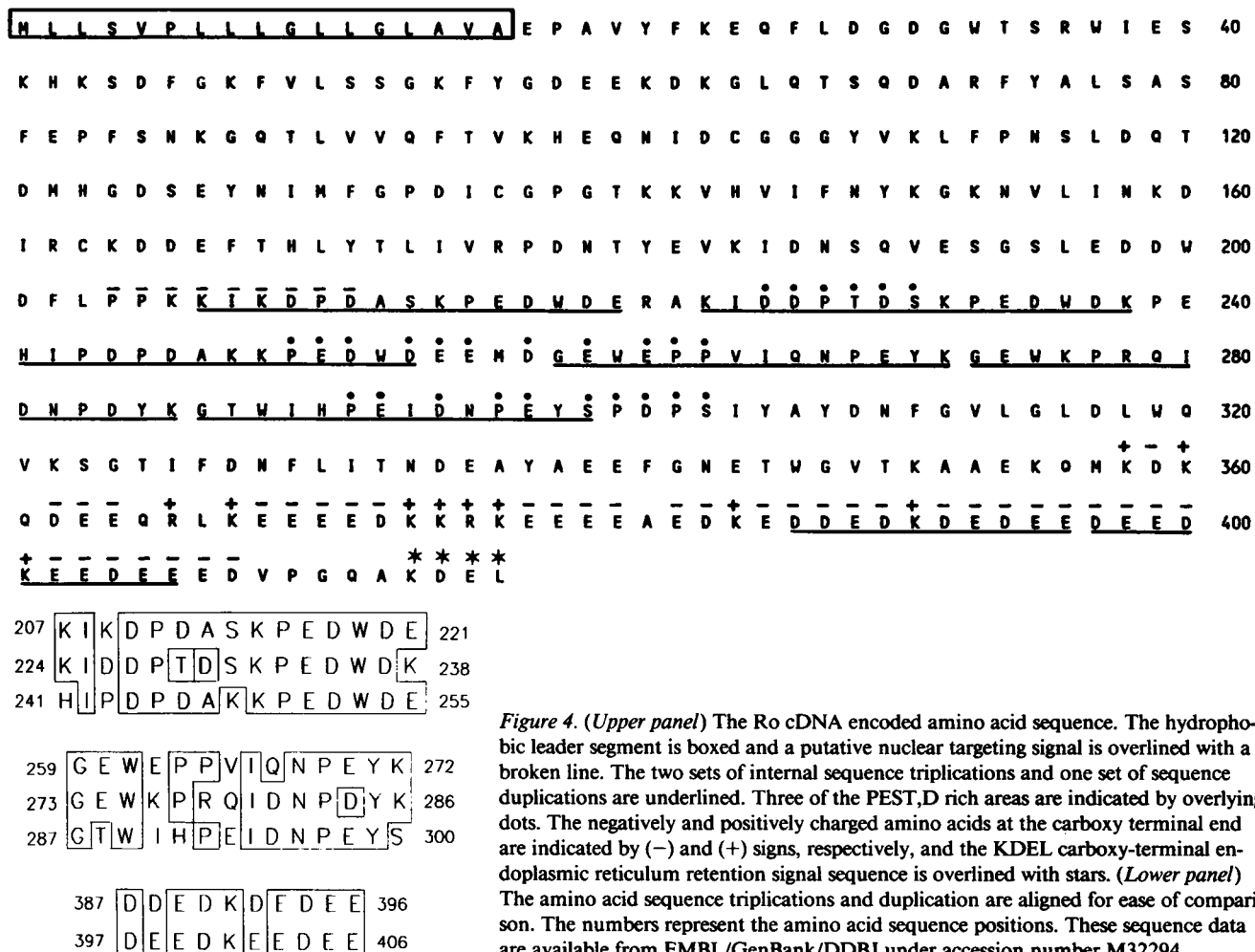


Figure 4. (Upper panel) The Ro cDNA encoded amino acid sequence. The hydrophobic leader segment is boxed and a putative nuclear targeting signal is overlined with a broken line. The two sets of internal sequence triplications and one set of sequence duplications are underlined. Three of the PEST,D rich areas are indicated by overlying dots. The negatively and positively charged amino acids at the carboxy terminal end are indicated by (-) and (+) signs, respectively, and the KDEL carboxy-terminal endoplasmic reticulum retention signal sequence is overlined with stars. (Lower panel) The amino acid sequence triplications and duplication are aligned for ease of comparison. The numbers represent the amino acid sequence positions. These sequence data are available from EMBL/GenBank/DDBJ under accession number M32294.

13.5-kb hybridizing fragment. Other restriction enzyme digests were similarly analyzed, with no difference in the pattern of bands between individuals, suggesting that the gene is not highly polymorphic and that it exists as a single copy (data not shown). A similar analysis using several different radiolabeled portions of the 1.9-kb cDNA allowed the construction of a genomic restriction map as shown in Fig. 7. This Ro gene occupies ~ 6 kb of genomic DNA indicating that introns may contribute 4 kb to this gene.

The Ro cDNA was used to determine chromosomal location by Southern filter hybridization analysis of Hind III digested DNA extracted from 38 independently derived human × CHO somatic cell hybrids that had randomly segregated human chromosomes. As seen in Fig. 8, Ro cDNA hybridized to both human and CHO DNA fragments and the resolvable difference in fragment size (human at 19–20 kb and CHO at 5.7 kb) made it easy to determine the presence or absence of human genomic DNA among the hybrid clones. The data are scored in Table II along with the discordancy analysis of Ro with respect to each human chromosome. The low discordancy between Ro-hybridizing human sequences and human chromosome 19 (8%) and the apparent random association between Ro-hybridizing human sequences and every other human chromosome (34–67% discordancy) suggest the chromosome 19 location of this gene.

The markers used to determine the presence or absence of

human chromosome 19 in the hybrids (PEPD and GPI) are both located on the long arm of this small, slightly submetacentric chromosome. Therefore, the three clones (1HL14, 9HL9, and 24HL8) discordant for the chromosome 19 markers and Ro suggest that Ro might be on the short arm of the chromosome. This hypothesis was tested by examining those three hybrids for the presence of known chromosome 19 short-arm markers, C3 and LDLR. The results are summarized in Table III. Ro is perfectly concordant with LDLR in this set, clearly placing the gene on the short arm of chromosome 19. The discordancy of LDLR with C3-PEPD-GPI in clone 24HL8 is consistent with the linkage data placing LDLR distal to C3 (53).

**Immunoblotting.** Rabbit anti-synthetic peptide 6–19 (SP 6–19) antiserum bound to the 60-kD protein and a previously identified 52-kD Ro protein on the same immunoblot (data not shown). When the rabbit antibodies are eluted from the 52-kD protein they react with the 60-kD protein.

**RNA binding studies.** Previous studies have demonstrated that human Ro antisera immunoprecipitate the hY RNAs which are noncovalently bound to a 60-kD protein to which the antisera reacts (13). Recently the hY RNAs were similarly immunoprecipitated with immunoaffinity-purified Ro antisera directed against either a 52- or a 60-kD protein (54). With immunopurified rabbit anti-SP 6–19 antibodies we were able to immunoprecipitate the hY RNAs from HeLa cells (Fig. 9,

Hu - M L L S V P L L L G L L G L A V A E P A V Y F K E Q F L D G D G W T S R 36

Hu - W I E S K H K S D F G K F V L S S G K F Y G D E E K D K G L Q T S Q D A 72  
 On - - - - - incomplete NH<sub>2</sub> end - - F Y G D A V K D K G L K T T Q D A

Hu - R F Y A L S A S F E - P F S N K G Q T L V V Q F T V K H E Q N I D C G G 107  
 : F Y : : : A . F : : F S N K G : : L V : Q F : V K H E Q : I D C G G  
 On - K F Y S I G A K F D K S F S N K G K S L V I Q F S V K H E Q D I D C G G

Hu - G Y V K L F P N S L D Q T D M H G D S E Y N I M F G P D I C G P G T K K 143  
 G Y V K L . : . : . : . D H G : : Y : I M F G P D I C G P G T K K  
 On - G Y V K L M A S D V N L E D S H G E T P Y H I M F G P D I C G P G T K K

Hu - V H V I F N Y K G K N V L I N K D I R C K D D E F T H L Y T L I V R P D 179  
 V H V I F : Y K : : N : I : K D I R C K D D F T H L Y T L I V . : D  
 On - V H V I F H Y K D R N H M I K K D I R C K D D V F T H L Y T L I V N S D

Hu - N T Y E V K I D N S Q V E S G S L E D D W D F L P P K K I K D P D A S K 215  
 N T Y E V : I D . . . E S G . L E . D W D F L P P K K I K D P D A . K  
 On - N T Y E V Q I D G F K A E S G E L E A D W D F L P P K K I K D P D A K K

Hu - P E D W D E R A K I D D P T D S K P E D W D K P E H I P D P D A K K P E 251  
 P E D W D E R . I D D . D . K P E D W D K P E H I P D P D A K K P E  
 On - P E D W D E R E F I D D E D D K K P E D W D K P E H I P D P D A K K P E

Hu - D W D E E M D G E W E P P V I Q N P E Y K G E W K P R Q I D N P D Y K G 287  
 D W D : E M D G E W E P P : : : N P E Y K G E W K P : Q . N P . Y K G  
 On - D W D E E M D G E W E P P M V D N P E Y K G E W K P K Q K K N P A Y K G

Hu - T W I H P E I D N P E Y S P D P S I Y A Y D N F G V L G L D L W Q V K S 323  
 . W I H P E I : P : Y : P D : : Y . Y D : : G . : G : D L W Q V K S  
 On - K W I H P E I E I P D Y T P D D N L Y V Y D D I G A I G F D L W Q V K S

Hu - G T I F D N F L I T N D E A Y A E E F G N E T W G V T K A A E K Q M K D 359  
 G T I F D : : : T : . : A . . F G : . T : : T : . : E K : K :  
 On - G T I F D D V I V T D S V E E A K K F G E K T L K I T R E G E K K - K G

Hu - K Q D E E Q R L K E E E E D K K R K E E E E A E D K E D E D K D E D E 395  
 K : . . . Q : K . E . : : K : K E . . . : : . . K .  
 On - K K T K K Q K - K K E K N E K I K K E K M K R K R A N R K K K K \* end

Hu - E D E E D K E E D E E E D V P G Q A K D E L 417

Figure 5. Amino acid sequence homology between the human Ro protein (Hu) and an *Onchocerca volvulus* antigen (On). The On cDNA sequence was truncated at the 5' end and thus the encoded amino terminus of this antigen is incomplete (incomplete NH<sub>2</sub> end) for full homology comparison. The numbers correspond to the Hu amino acid sequence position. Symbols: (·) represents a conservative amino acid substitution; (•) represents a semiconservative substitution; (–) indicates a gap between amino acids; (\*) denotes the end of the On protein. These sequence data are available from EMBL/GenBank/DDBJ under accession number M32294.

left panel). We have previously demonstrated the presence of RNA in our purified Ro protein product by ultraviolet absorbance analysis (14). To determine whether this was hY RNA, the purified protein was irradiated with UVC to cross-link any associated RNA to the Ro protein so that a shift in molecular mass could be detected by SDS-PAGE analysis. This method of cross-linking protein to intimately associated RNA molecules has been well established (55). In this manner we demonstrated that the purified Ro protein which normally migrates at 60 kD by SDS-PAGE, migrates at four different higher molecular masses between ~ 86 and 96 kD after UVC cross-linking (Fig. 9, right panel). The UVC-induced increase in molecular mass could be attenuated by digesting the cross-linked sample with ribonuclease before SDS-PAGE (data not shown). Rabbit anti-SP 6-19 serum and a human Ro antiserum both react to the 60-kD protein and each of the four higher molecular mass cross-linked species by immunoblot analysis (data not shown).

**Discussion**

We have isolated, characterized, and expressed a cDNA clone which encodes a 46-kD Ro RNP autoantigen which migrates

at 60 kD by SDS-PAGE. The deduced amino acid sequence although novel is 63% homologous with an *Onchocerca volvulus* antigen. The amino acid sequence includes a hydrophobic leader segment, an endoplasmic reticulum retention sequence, a strongly charged hydrophilic carboxy terminus, three sets of repeating sequences, a putative nuclear translocation signal, and a putative calcium-binding domain. The amino acid sequence contains a region recognized by antibodies, which when immunopurified bind both our purified "60"-kD protein and a previously identified 52-kD Ro protein. These antibodies also immunoprecipitate the hY RNAs. The gene encoding the 60-kD protein exists in a single copy on the short arm of chromosome 19.

The *M<sub>r</sub>* disparity between that measured by SDS-PAGE and that calculated from the encoded polypeptide is not difficult to reconcile in light of similar discrepancies reported with a number of other proteins which have a highly charged region similar to this Ro polypeptide. This highly charged region apparently is responsible for the retarded gel migration observed with these proteins. Our cDNA produces a 60-kD protein when expressed in bacteria which supports aberrant gel migration as the likely cause of the *M<sub>r</sub>* disparity.

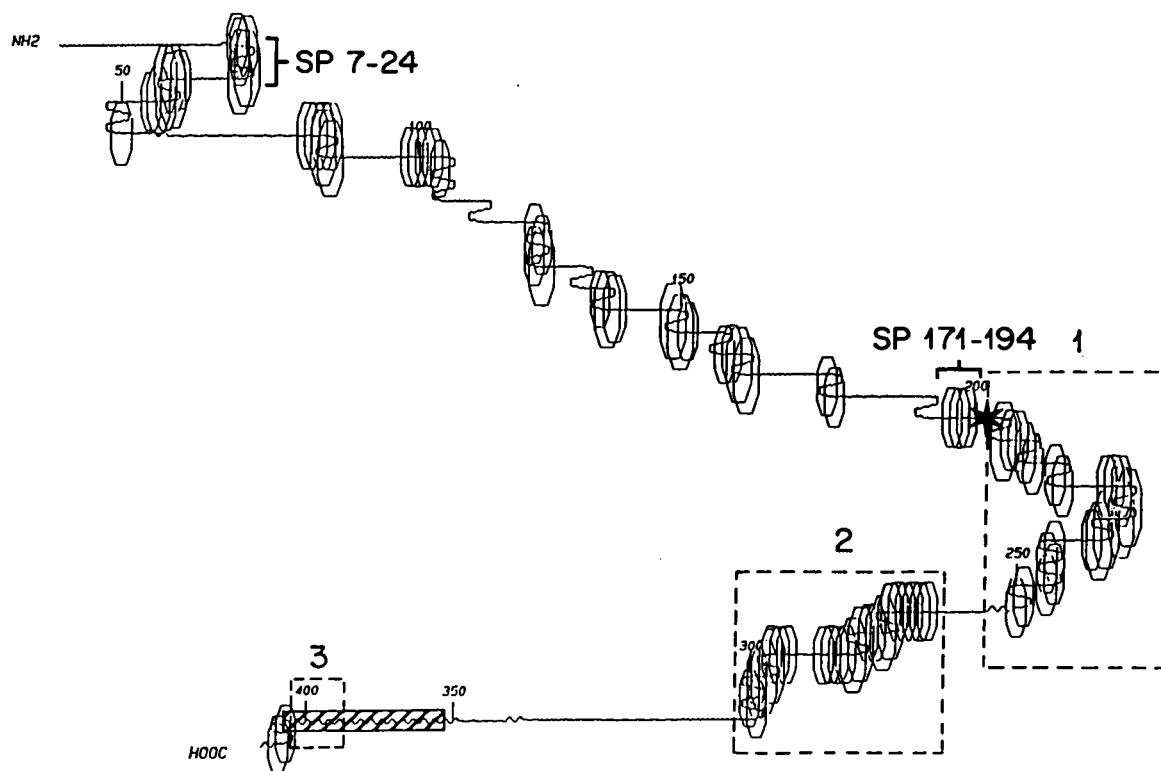


Figure 6. Chou-Fasman structural and Jameson-Wolf antigenicity predictions of the Ro polypeptide. The small numbers represent amino acid sequence positions, pleated lines represent  $\beta$  sheets, wavy lines represent  $\alpha$  helices, and directional changes represent turns. The octagons represent potential antigenic sites. SP 7-24 and SP 171-194 represent the amino acid sequence positions for which corresponding synthetic peptides were previously shown to react with human Ro antisera. The region of a putative nuclear targeting signal is indicated by a star, and the strongly charged carboxy terminus is shaded. The two sets of sequence duplications are boxed (1 and 2) as is the set of sequence duplications (3).

The fact that Ro antibodies specific for a sequence encoded by our cDNA recognize our 60-kD protein from Wil-2 cells and a 52-kD protein from chronic lymphocytic leukemia cells, on the same immunoblot, suggests that both proteins have one or more shared epitopes. The two proteins might be separate gene products, or the 52-kD protein might be a degradation product of the 60-kD protein. Similarly the 60-kD protein might result from a posttranslational modification of the 52-kD protein. It is unlikely that these two proteins result from different transcripts from the same gene as only one molecular weight species of RNA was detected on Northern filter hybridization. We found no evidence that the 60-kD protein is glycosylated or ubiquitinated (personal observation). Differences

in purification techniques or cell/tissue source might explain why two different  $M_r$  proteins were isolated.

The absence of a typical poly-A tail at the 3' end of our 1.9-kb cDNA suggests that it is truncated as was the 1.2-kb clone. This may have arisen from aberrant cDNA synthesis or from subsequent deletion of the poly-A tail after cDNA synthesis. Another explanation would be that the Ro mRNA is not poly-A tailed, like histone mRNA. However, there is no comparable 3' end processing signal sequence as found in the histones (56) and the 1.9-kb clone does have a poly-A signal sequence.

No major similarities were found between the RNP consensus sequence and the deduced amino acid sequence of our

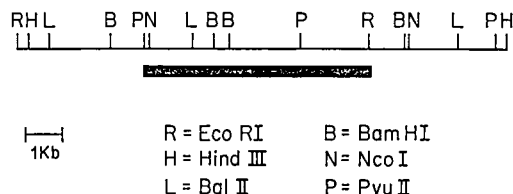


Figure 7. Genomic restriction map. Various portions of the Ro cDNA were radiolabeled and hybridized to multiple restriction enzyme digests of human genomic DNA by the Southern technique. The length of each labeled fragment was determined and a composite restriction map was thus constructed. The map indicates that this Ro gene resides within a 6-kb stretch of chromosomal DNA. Restriction enzyme sites are indicated.

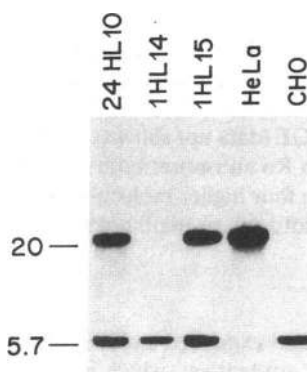


Figure 8. Southern filter hybridization with radiolabeled Ro cDNA reveals a 19-20-kb hybridizing Hind III fragment in the human (HeLa) cell line in comparison to a 5.7 hybridizing fragment in the hamster (CHO) cell line. Human  $\times$  hamster hybrid clones 24HL10 and 1HL15 contain that portion of human chromosomal DNA where this Ro gene resides whereas clone 1HL14 does not.



Table II. Human Chromosomes and DNA Sequences Present or Absent in Hybrids

Hybrid clone	Chromosomes																						Ro probe	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22		X
1HL14	-	+	-	-	-	+	+	-	-	-	+	+	-	-	+	+	-	+	+	+	+	+	-	-
1HL15	+	+	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL3	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL4	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL10	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL11	-	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL14	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5HL15	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL16	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL17	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5HL20	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL3	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL7	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL9	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL10	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL11	+	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9HL12	+	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
20HL1	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20HL4	+	+	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20HL6	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20HL7	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20HL21-4	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24HL5	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24HL7-5	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24HL8	-	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24HL10	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
41HL7	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
41HL10	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL7	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL9	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL11	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL12	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135HL15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3B2	-	-	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5A3	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Percentage discordant with Ro (%)	42	61	49	53	45	53	57	34	47	54	50	42	57	37	42	65	61	49	8	50	50	47	45	45

Symbols: +, present; -, absent. Blank spaces were not read.

Table III. Chromosome 19 Markers in Hybrid Clones with Discordancies

Hybrid clone	Markers				
	Ro	LDLR	C3	PEPD	GPI
24HL8	+	+	-	-	-
9HL9	-	-	-	+	+
1HL14	-	-	-	+	+

Symbols: +/-, presence/absence of marker hybridization.

cDNA (46). However, the RNP consensus sequence is not necessarily a requirement nor a universal property of RNA binding proteins for it is absent in ribosomal proteins, in many viral RNA-binding nucleocapsid proteins and in the Sm-D RNP autoantigen (57, 58). The three duplications between residues 207 and 255 each have a helix-turn-helix configuration and may be a site of RNA binding. However, these helix-turn-helix units have a larger turn component than those described in other nucleic acid binding proteins (59). It is of interest to note that another RNA-binding protein, the human 70-kD small nuclear RNP, also resides on chromosome 19 and has a similar  $M_r$  disparity, though it has no significant sequence homology with Ro protein (35).

The role of this protein in cellular function and its precise cellular location are unknown. Analysis of the amino acid sequence gives some insight into these issues. The hydrophobic

leader segment of the polypeptide suggests that this protein undergoes transmembrane transport. This sequence may serve to transport this protein across the endoplasmic reticulum for modification, however there is no evidence of glycosylation. The KDEL carboxy-signal sequence suggests that this protein may reside in the endoplasmic reticulum.

The amino terminal amino acid sequence similarity with rabbit calregulin, a calcium-binding protein with an  $M_r$  and pI similar to this 60-kD Ro protein (39), and a region of putative calcium binding in our protein suggest that this protein may have a similar function or may even be the human equivalent to rabbit calregulin. This is even more interesting in light of studies which have shown that calregulin is sequestered in the endoplasmic reticulum or another membrane-bound cytoplasmic organelle (60). Thus the 60-kD Ro protein might have a hydrophobic leader to allow transport into a membrane bound organelle where it might bind calcium, and the KDEL sequence would insure retention of the protein within that organelle.

Indirect immunofluorescence microscopy on cultured mouse L cells (fibroblasts) and human Hep-2 cells (epithelioid cells) with rabbit antisera raised against SP 6-19, reveals predominantly perinuclear and cytoplasmic staining (data not shown), similar to the pattern reported by Hendrick and co-workers with Ro antisera (61), and also similar to the pattern seen with antibodies directed against calregulin (60). Other investigators have reported predominantly intranuclear localization of the Ro antigens by indirect immunofluorescence (12). Whether this discrepancy in subcellular localization is related to the method of cell fixation, cell substrate, or the type of Ro antisera used, needs to be further investigated. Isolation of antibodies specific for each Ro antigen should allow a more precise subcellular localization of each antigen and these results may help explain the discrepancies encountered thus far with immunofluorescence staining.

The Ro polypeptide does contain the sequence PPKKIKDPD (residues 203-212 in Fig. 4, upper panel) which is very similar to nuclear targeting signals of other nuclear proteins (62). This sequence might facilitate transport of this protein into the nucleus. Histone-binding proteins similarly have nuclear targeting signals and highly negatively charged regions (32, 34).

There has been mounting evidence that foreign microbial antigens may trigger an inappropriate immune response against self-antigens through molecular mimicry (63). Initial computer search for sequence homology to microbial agents has not been fruitful. As the Ro epitopes become better defined it may become more apparent whether microbial agents play a role in the pathogenesis of this autoimmune response. The sequence homology with the *Onchocerca volvulus* antigen does suggest the possibility that a foreign protein homologous to a self-protein might trigger an immune response which reacts with the self protein. *Onchocerca volvulus* is a filarial nematode which causes river blindness, sclerosing lymphadenitis, and dermatologic disease in humans residing in parts of Africa and Central America (64). Studies are underway to determine whether sera from patients with this disease contain Ro antibodies.

The relationship between this Ro protein and the others is unknown. They may be structurally and (or) functionally related as are the antigenic U series of RNPs (65). Taken to-

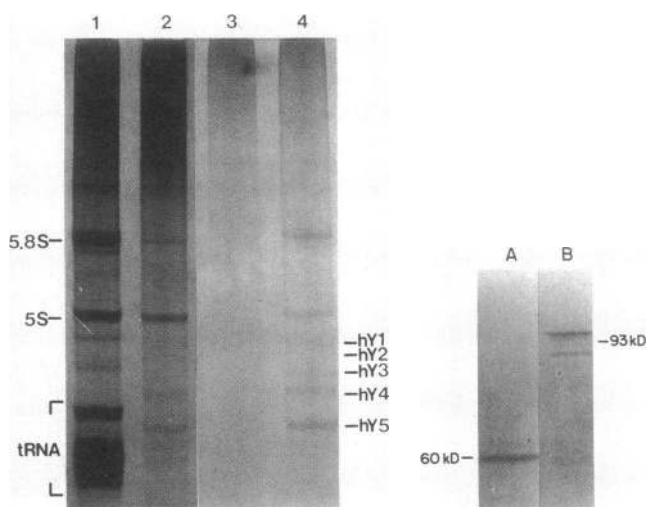


Figure 9. (Left) Silver-stained RNA fractions in a polyacrylamide gel. Extracted total RNA from HeLa cells was either loaded directly on the gel (lane 1) or immunoprecipitated first with: rabbit anti-SP 6-19 antibodies eluted from the 52-kD Ro (lane 2), normal rabbit serum antibodies eluted from the 52-kD Ro (lane 3), or a patient Ro antiserum (lane 4). The 5.8-S and 5-S ribosomal fractions and tRNA fractions are indicated in lane 1. The hY RNA species are demonstrated in both lanes 2 and 4. (Right) Ultraviolet light cross-linking of the 60 kD Ro to the hY RNAs. This is a Coomassie Blue-stained SDS-PAGE gel with non-UVC-exposed purified Ro protein in lane A and UVC cross-linked purified Ro protein in lane B. Four separate higher molecular mass bands are seen in the cross-linked sample, though the two middle bands are faint. The positions of molecular mass markers are indicated.

gether, our data and the hY RNA-binding data from other investigators suggest that a 52-kD and two 60-kD Ro proteins bind hY RNA. Ro antisera specific for a 52- or a 60-kD protein have been shown to immunoprecipitate the hY RNAs from cellular extracts. However, whenever the 52-kD specific antibodies were used in this study, the hY RNAs and 52-kD protein were precipitated along with a small amount of 60-kD protein (13, 54). Thus it is not certain whether the 52-kD species binds the hY RNA directly or indirectly through its association with a 60-kD hY RNA-binding protein. Binding of a hY RNA has also been demonstrated in reconstitution studies with another recently characterized 60-kD Ro RNP, though the efficiency of binding was reportedly quite low (45). Immunopurified rabbit antibodies directed at our 60-kD protein's amino-terminal amino acid sequence immunoprecipitate the hY RNAs and also recognize a recently characterized 52-kD protein, making it uncertain whether the immunoprecipitated hY RNA was bound to the 52-kD protein, the 60-kD protein, or both. The UVC cross-linking studies give more direct evidence that the "60"-kD protein binds each of the four major hY RNAs. Each of the four higher molecular mass species is consistent with the addition of one of the four major hY RNAs (13). These findings support the concept of only one hY RNA molecule bound per Ro molecule as suggested by Wolin and Steitz (13), who demonstrated that the hY RNA-protein complexes sediment at  $\sim 7S$  (equivalent to 93 kD) in sucrose gradients (13). This is consistent with the  $M_r$  of one 60-kD protein molecule complexed with one of the hY RNA molecules which range from 28 to 38 kD in size (an average of 33 kD).

Ro was a term first used to define a soluble cytoplasmic antigen which formed a unique line of precipitation in double-immunodiffusion studies with sera from patients with systemic lupus erythematosus and Sjogren's syndrome (8). From the work of Steitz and co-workers, Ro has been further defined as a cytoplasmic hY RNA binding protein which migrates at 60 kD by SDS-PAGE (13). Like the originally described Ro protein, the protein we describe measures 60 kD by SDS-PAGE and is reactive by Western immunoblotting to prototypal monospecific human Ro antisera from several different laboratories including the laboratory of the Center for Disease Control (AF/CDC7). Rabbit antiserum raised against the amino-terminal portion of this protein (anti-SP 6-19) demonstrates a cytoplasmic pattern with indirect immunofluorescence staining, and contains antibodies that immunoprecipitate the hY RNAs. These data suggest that this Ro autoantigen may be the originally described 60-kD, cytoplasmic, hY RNA-binding protein.

Now that several proteins with Ro antigenicity have been identified, including two different 60-kD proteins, a system of classification needs to be developed so that each Ro antigen gets a more unique designation. This should be accomplished as the antigens become better characterized.

Whether or not patients with Ro antibodies can be clinically categorized by which of the Ro proteins or epitopes their sera recognize and whether or not this is related to their HLA type has not yet been determined. The ability to categorize patients based on Ro epitope recognition could have great clinical utility if the patient's clinical course and/or response to therapy could be predicted by these results. The characterization of the various Ro cDNAs and their encoded epitopes

should be helpful in this regard and should also provide a means to further clarify the functional and pathologic roles of these protein autoantigens.

## Acknowledgments

We are grateful to the assistance from Peter Roome and Dr. Robert Fisher in the computer based sequence analysis and to Jeff Wilson and Carol Williams for their excellent technical assistance.

These studies were in part supported by grants AI-12127, GM-31689, AR-19101, and AR-07341 from the National Institutes of Health and by support from the Dallas Biomedical Corporation. Dr. Sontheimer is the recipient of National Institutes of Health Research Career Development Award AR01784.

## References

1. Martinez-Lavin, M., J. H. Vaughan, and E. M. Tan. 1979. Auto-antibodies and the spectrum of Sjogren's syndrome. *Ann. Intern. Med.* 91:185-190.
2. Sontheimer, R. D., P. J. Maddison, M. Reichlin, R. E. Jordon, P. Stastny, and J. N. Gilliam. 1982. Serologic and HLA associations in subacute cutaneous lupus erythematosus, a clinical subset of lupus erythematosus. *Ann. Intern. Med.* 97:664-671.
3. Maddison, P. J., T. T. Provost, and M. Reichlin. 1981. ANA negative systemic lupus erythematosus: serological analysis. *Medicine (Baltimore)*. 60:87-94.
4. Kephart, D., A. F. Hood, and T. T. Provost. 1981. Neonatal lupus: serologic findings. *J. Invest. Dermatol.* 77:331-333.
5. Provost, T. T., F. C. Arnett, and M. Reichlin. 1983. Homozygous C2 deficiency, lupus erythematosus and anti-Ro(SSA) antibodies. *Arthritis Rheum.* 26:1279-1282.
6. Meyer, O., G. Hauptmann, G. Tuppeiner, H. D. Ochs, and F. Mascart-Lemone. 1985. Genetic deficiency of C4, C2 or C1q and lupus syndromes: association with anti-Ro(SS-A) antibodies. *Clin. Exp. Immunol.* 62:678-684.
7. McCauliffe, D. P., F. Lux, T. S. Lieu, I. Sanz, J. Hanke, M. Newkirk, M. J. Siciliano, R. D. Sontheimer, and J. D. Capra. 1989. Ro/SS-A and the pathogenic significance of its antibodies. *J. Autoimmun.* 2:375-381.
8. Clark G., M. Reichlin, and T. B. Tomasi. 1969. Characterization of a soluble cytoplasmic antigen reactive with sera from patients with systemic lupus erythematosus. *J. Immunol.* 102:117-122.
9. Alspaugh, M. A., and E. M. Tan. 1975. Antibodies to cellular antigens in Sjogren's syndrome. *J. Clin. Invest.* 55:1067-1073.
10. Alspaugh, M. A., and P. Maddison. 1979. Relation of the identity of certain antigen-antibody systems in systemic lupus erythematosus and Sjogren's syndrome: an interlaboratory collaboration. *Arthritis Rheum.* 22:796-798.
11. Lerner, M. R., J. Boyle, J. A. Hardin, and J. A. Steitz. 1981. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science (Wash. DC)*. 211:400-402.
12. Harmon, C. E., J. S. Deng, C. L. Peebles, and E. M. Tan. 1984. The importance of tissue substrate in the SS-A/Ro antigen-antibody system. *Arthritis Rheum.* 27:166-173.
13. Wolin, S. L., and J. A. Steitz. 1984. The Ro small cytoplasmic ribonucleoproteins: identification of the antigenic protein and its binding site on the Ro RNAs. *Proc. Natl. Acad. Sci. USA*. 81:1996-2000.
14. Lieu, T. S., M. M. Newkirk, J. D. Capra, and R. D. Sontheimer. 1988. Molecular characterization of human Ro/SS-A antigen. *J. Clin. Invest.* 82:96-101.
15. Rader, M. D., C. O'Brien, Y. Liu, J. B. Harley, and M. Reichlin. 1989. The heterogeneity of the Ro/SSA antigen: different molecular forms in lymphocytes and red blood cells. *J. Clin. Invest.* 83:1293-1298.

16. Darbre, A. 1986. *Practical Protein Chemistry—A Handbook*. John Wiley & Sons, Inc., New York.
17. Lathe, R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. *J. Mol. Biol.* 183:1–12.
18. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. *Molecular Cloning—A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
19. Ausubel, F. M. 1987. *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., New York.
20. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13mp phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. 33:103–119.
21. Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled expression of specific genes. *Proc. Natl. Acad. Sci. USA*. 82:1074–1078.
22. Erlich, H. H., D. H. Gelfand, and R. K. Saiki. 1988. Specific DNA amplification. *Nature (Lond.)*. 33:461–462.
23. Stallings, R. L., E. Olson, A. W. Strauss, L. H. Thompson, L. L. Bachinski, and M. J. Siciliano. 1988. Human creatine kinase genes on chromosomes 15 and 19, and proximity of the gene for the muscle form to the genes for apolipoprotein C2 and excision repair. *Am. J. Hum. Genet.* 43:144–151.
24. McBride, O. W., B. Z. Zmudzka, and S. H. Wilson. 1987. Chromosomal location of the human gene for DNA polymerase B. *Proc. Natl. Acad. Sci. USA*. 84:503–507.
25. Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387–395.
26. Laemeli, U. K. 1970. Cleavage of structural proteins during the assembly of the bacteriophage T4. *Nature (Lond.)*. 227:680.
27. Olmstead, J. B. 1981. Affinity purification of antibodies from diazotized paper blots of heterogeneous protein samples. *J. Biol. Chem.* 256:11955–11957.
28. Forman, M. S., M. Nakamura, T. Mimori, C. Gelpi, and J. A. Hardin. 1985. Detection of antibodies to small nuclear ribonucleoproteins and small cytoplasmic ribonucleoproteins using unlabeled cell extracts. *Arthritis Rheum.* 28:1356–1361.
29. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*. 44:283–292.
30. Nevins, J. R. 1983. The pathway of eukaryotic mRNA formation. *Ann. Rev. Biochem.* 52:441–66.
31. Hope, I. A., and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell*. 46:885–894.
32. Dingwall, C., S. M. Dilworth, S. J. Black, S. E. Kearsey, L. S. Cox, and R. A. Laskey. 1987. Nucleoplasmic cDNA sequence reveals polyglutamic acid tracts and a cluster of sequences homologous to putative nuclear localization signals. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:69–74.
33. Benedum, U. M., P. A. Baeuerle, D. S. Konecki, R. Frank, J. Powell, J. Mallet, and W. B. Huttner. 1986. The primary structure of bovine chromogranin A: a representative of a class of acidic secretory proteins common to a variety of peptidergic cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:1495–1502.
34. Kleinschmidt, J. A., C. Dingwall, G. Maier, and W. Franke. 1986. Molecular characterization of a karyophilic, histone-binding protein: cDNA cloning, amino acid sequence and expression of nuclear protein N1/N2 of *Xenopus laevis*. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:3547–3552.
35. Spritz, R. A., K. Strunk, C. S. Surowy, S. O. Hoch, D. E. Barton, and U. Francke. 1987. The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing and RNA-binding. *Nucleic Acids Res.* 15:10373–10391.
36. Lieu, T.-S., M. Jiang, J. C. Steigerwald, and E. M. Tan. 1984. Identification of the SS-A/Ro intracellular antigen with autoimmune sera. *J. Immunol. Methods*. 71:217–228.
37. Rogers, S., R. Wells, and M. Rechsteiner. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science (Wash. DC)*. 234:364–368.
38. Unnasch, T. R., M. Y. Gallin, P. T. Soboslay, K. D. Erttmann, and B. M. Greene. 1988. Isolation and characterization of expression cDNA clones encoding antigens of *Onchocerca volvulus* infective larvae. *J. Clin. Invest.* 82:262–269.
39. Khanna, N. C., M. Tokuda, and D. M. Waisman. 1987. Comparison of calregulins from vertebrate livers. *Biochem. J.* 242:245–251.
40. Van Loon, A. P., R. J. De Groot, M. De Haan, A. Dekker, and L. A. Grivell. 1984. The DNA sequence of the nuclear gene coding for the 17-kd subunit VI of the yeast ubiquinol-cytochrome c reductase: a protein with an extremely high content of acidic amino acids. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1039–1043.
41. Jongstra, J., G. F. Tidmarsh, J. Jongstra-Bilen, and M. M. Davis. 1988. A new lymphocyte-specific gene which encodes a putative Ca<sup>2+</sup>-binding protein is not expressed in transformed T lymphocyte lines. *J. Immunol.* 141:3999–4004.
42. Baudier, J., and D. Gerard. 1983. Ions binding to S100 proteins: structural changes induced by calcium and zinc on S100a and S100b proteins. *Biochemistry*. 22:3360–3369.
43. Munro S., and H. R. Pelham. 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell*. 48:899–907.
44. Walter, P., R. Gilmore, and G. Blobel. 1984. Protein translocation across the endoplasmic reticulum. *Cell*. 38:5–8.
45. Deutscher, S. L., J. B. Harley, and J. D. Keene. 1988. Molecular analysis of the 60 kD human Ro ribonucleoprotein. *Proc. Natl. Acad. Sci. USA*. 85:9479–9483.
46. Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss. 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol. Cell. Biol.* 6:2932–2943.
47. Evans, R. M., and S. M. Hollenberg. 1988. Zinc fingers: guilt by association. *Cell*. 52:1–3.
48. Landschulz, W. H., P. F. Johnson, and S. L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science (Wash. DC)*. 240:1759–1764.
49. Chou, P. Y., and G. D. Fassman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 47:45.
50. Kyte, J., and R. R. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105.
51. Jameson, B. A., and H. Wolfe. 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. *Comput. Appl. Biosci.* 4:181–6.
52. Lieu, T. S., M. M. Newkirk, F. C. Arnett, L. A. Lee, J. S. Deng, J. D. Capra, and R. D. Sontheimer. 1989. A major autoepitope is present on the amino terminus of the human Ro/SS-A polypeptide. *J. Autoimmun.* 2:367–374.
53. Lusic, A. J., C. Heinzmann, R. S. Sparkes, R. Geller, M. C. Sparkes, and T. Mohandas. 1985. Regional mapping on human chromosome 19: apolipoprotein E, apolipoprotein C11, low density lipoprotein (LDL) receptor, peptidase D, glucose phosphate isomerase. *Cytogenet. Cell Genet.* 40:683.
54. Ben-Chetrit, E., E. K. Chan, K. F. Sullivan, and E. M. Tan. 1988. A 52-kD protein is a novel component of the SS-A/Ro antigenic particle. *J. Exp. Med.* 167:1560–1571.
55. Setyono, B., and J. R. Greenberg. 1981. Proteins associated with Poly(A) and other regions of mRNA and hnRNA molecules as investigated by crosslinking. *Cell*. 24:775–783.
56. Mowry, K. L., and J. A. Steitz. 1987. Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. *Science (Wash. DC)*. 238:1682–1687.
57. Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss. 1987. Primary structure of human nuclear ribonucleic particle C pro-

teins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA and pre-rRNA-binding proteins. *Mol. Cell. Biol.* 7:1731-1739.

58. Rokeach, L. A., J. A. Haselby, and S. O. Hoch. 1988. Molecular cloning of a cDNA encoding the human Sm-D autoantigen. *Proc. Natl. Acad. Sci. USA.* 85:4832-4836.

59. Pabo, C. O. 1984. Protein-DNA recognition. *Annu. Rev. Biochem.* 53:293-321.

60. Khanna, N. C., M. Tokuda, and D. M. Waisman. 1987. Calregulin: purification, cellular localization, and tissue distribution. *Methods Enzymol.* 139:36-50.

61. Hendrick, J. P., S. L. Wolin, J. Rinke, M. R. Lerner, and J. A. Steitz. 1981. Ro small cytoplasmic ribonucleoproteins are a subclass of La ribonucleoproteins: further characterization of the Ro and La small

ribonucleoproteins from uninfected cells. *Mol. Cell. Biol.* 1:1138-1149.

62. Burglin, T., and E. M. De Robertis. 1987. The nuclear migration signal of *Xenopus laevis* nucleoplasmin. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2617-2625.

63. Oldstone, M. B. 1987. Molecular mimicry and autoimmune disease. *Cell.* 50:819-820.

64. Connor, D. H., G. H. George, and D. W. Gibson. 1985. Pathologic changes of human onchocerciasis: implications for future research. *Reviews of Infect. Dis.* 7:809-819.

65. Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, A. Kramer, D. Frendewey, and W. Keller. 1987. Functions of the abundant U-snRNAs. In *Small Nuclear Ribonucleoprotein Particles: Structure and Functions*. M. L. Birnstein, editor. Springer-Verlag New York, Inc., New York. 115-154.