MOLECULAR AND CELLULAR BIOLOGY, Aug. 1986, p. 2774–2783
0270-7306/86/082774-10$02.00/0

Vol. 6, No. 8

Primary Structure of Human Ribosomal Protein S14 and the Gene That Encodes It†

DOUGLAS D. RHOADS, APARNA DIXIT,‡ AND DONALD J. ROUFA*

Center for Basic Cancer Research, Division of Biology, Kansas State University, Manhattan, Kansas 66506

Received 21 February 1986/Accepted 21 April 1986

Chinese hamster ribosomal protein S14 cDNA was used to recognize homologous human cDNA and genomic clones. Human and Chinese hamster S14 protein sequences deduced from the cDNAs are identical. Two overlapping human genomic S14 DNA clones were isolated from a Charon 28 placental DNA library. A fragment of single-copy DNA derived from an intron region of one clone was mapped to the functional RPS14 locus on human chromosome 5q by using a panel of human × Chinese hamster hybrid cell DNAs. The human S14 gene consists of five exons and four introns spanning 5.9 kilobase pairs of DNA. Polyadenylated S14 transcripts purified from HeLa cell cytoplasm display heterogeneous 5′ ends that map within noncoding RPS14 exon 1. This precludes assignment of a unique 5′ boundary of RPS14 transcripts with respect to the cloned human genomic DNA. Apparently HeLa cells either initiate transcription at multiple sites within RPS14 exon 1, or capped 5′ oligonucleotides are removed from most S14 mRNAs posttranscription. In contrast to the few murine ribosomal protein and several other mammalian housekeeping genes whose structures are known, human RPS14 contains a TATA sequence (TATACTT) upstream from exon 1. Three related short sequence motifs, also observed in murine and yeast ribosomal protein genes, occur in this region of the RPS14 gene. RPS14 introns 3 and 4 both contain *Alu* sequences. Interestingly, the *Alu* sequence in intron 3 is located slightly downstream from a chromosome 5 deletion breakpoint in one human × hamster hybrid clone analyzed.

Ribosomal proteins (r-proteins) are encoded by complex multigene families in eucaryotes and procaryotes. Several bacterial and yeast r-protein genes have been investigated by both biochemical and genetic approaches. Substantial evidence indicates that bacterial and yeast r-protein genes are stringently coregulated by transcriptional, translational, and posttranslational mechanisms (16, 18, 22, 31, 38–40, 42, 47, 59, 60, 66–68). A few r-protein genes have been isolated from the DNAs of higher eucaryotes—*Drosophila melanogaster* (8, 14, 41, 57), *Xenopus laevis* (1, 6), mice (13, 43, 58, 63), rats (25, 36), and Chinese hamsters (34, 35, 45). For lack of appropriate genetic methods and markers, progress in elucidating mechanisms that regulate r-protein genes in these organisms has been slow. Despite this, recombinant DNA techniques permit us to distinguish active, intron-containing animal r-protein genes from processed pseudogenes (13, 23, 43, 58, 63) and indicate that they are regulated coordinately under a variety of experimental conditions (15, 17, 21, 24, 30, 33, 44, 51, 62).

The gene encoding Chinese hamster 40S ribosomal subunit protein S14 is the only mammalian r-protein gene currently amenable to somatic genetic analysis in tissue culture. This gene is the target of emetine resistance mutations, as *emtB* Chinese hamster ovary (CHO) cell mutants elaborate altered S14 mRNAs and polypeptides (4, 5, 27, 28, 35, 37, 45). We have isolated Chinese hamster S14 and other r-protein cDNAs (35) and have demonstrated that *emtB* S14 alleles differ from wild type by single base changes in two adjacent arginine codons at the 3′ ends of their coding sequences (45). S14 cDNA probes detect multiple bands in restriction digests of Chinese hamster genomic DNA (34, 35). However, because recessive *emtB* (S14) CHO cell clones are obtainable by single-step mutagenesis, it appears that they express only a single S14 gene (5, 27, 35). Others mapped the *emtB* locus to a region of chromosome 2 that is hemizygous in CHO cells (64, 65). Taken together, the data suggest that CHO cell DNA contains one functional RPS14 gene as well as several pseudogenes (35).

The distal portion of human chromosome 5 rescues emetine resistance in *emtB* Chinese hamster × normal human somatic cell hybrids (10, 11), and a DNA sequence complementary to S14 cDNAs is located on human chromosome 5 between bands q23 and q33 (9, 34). Apparently, interspecific mammalian hybrid cells assemble functional, hybrid ribosomes (53, 56, 61, 62), and the transcriptionally active human S14 gene is located in the region of chromosme 5q23-33.

We isolated complementary and genomic DNA clones of human r-protein S14 and determined their primary structures. Our data indicate that the human RPS14 gene is 5.9 kilobase pairs (kbp) long. It consists of five exons and four introns. Unlike a few mouse r-protein (13, 58, 63) and other characterized mammalian housekeeping genes, human RPS14 is preceded by a TATA sequence. Depending on its position relative to the transcriptional initiation site of the gene, the TATA sequence might constitute a conventional RNA polymerase I-binding site (7). Unfortunately, HeLa cell S14 transcripts display heterogeneous 5′ ends, and this precludes alignment of the TATA sequence with a unique transcript cap site by using DNA and mRNA sequence information alone.

MATERIALS AND METHODS

Materials. The Chinese hamster ovary cell line (CHO) used in these studies has been described before (4, 5, 27, 28). HeLa S3 cells were obtained from the American Type Culture Collection (Rockville, Md.). Methods for maintain-

* Corresponding author.

† Contribution no. 86-325-J from the Kansas Agricultural Experiment Station.

‡ Present address: Department of Pharmacology, Milton S. Hershey Medical Center, The Pennsylvania State University School of Medicine, Hershey, PA 17033.

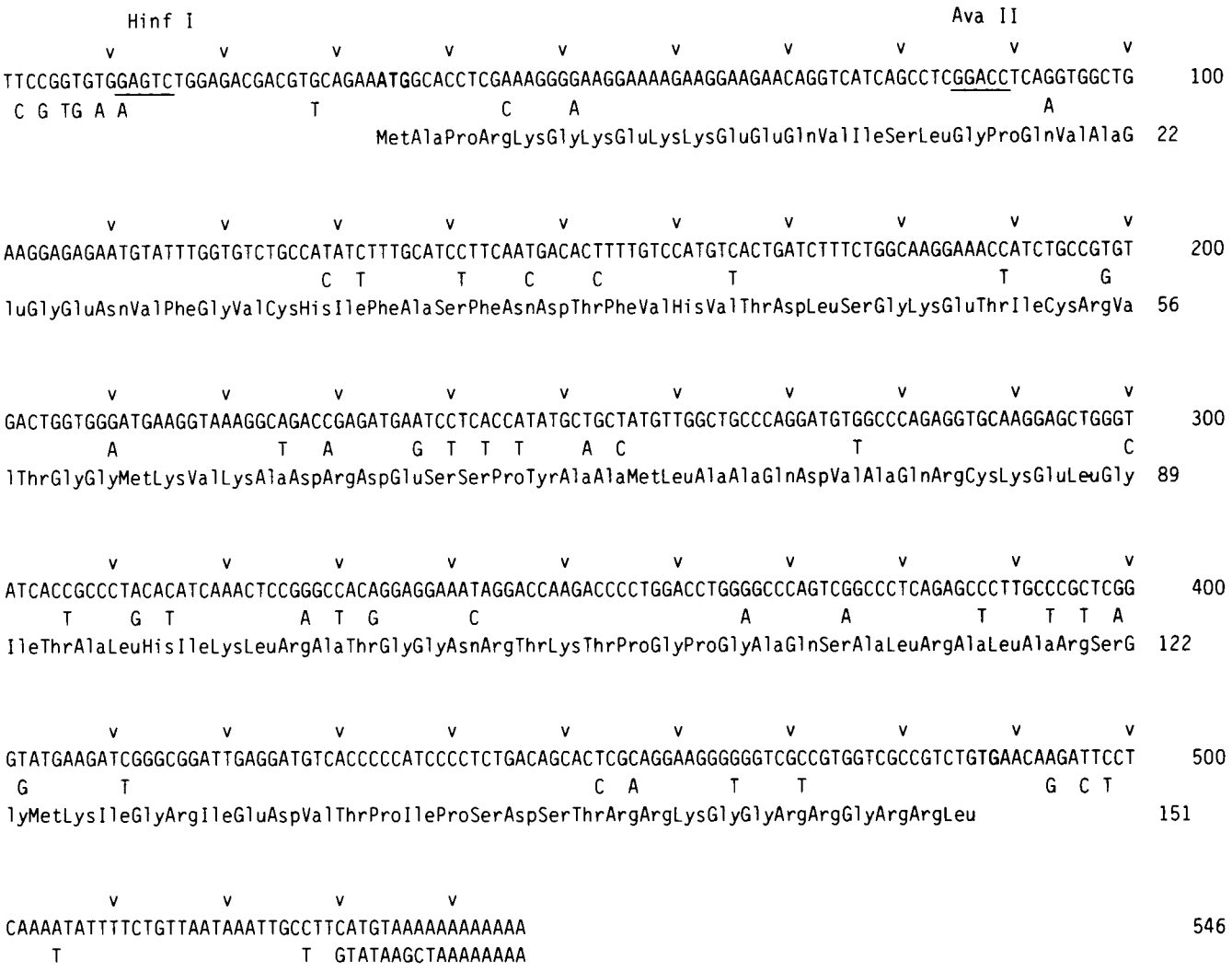


FIG. 1. Sequence of human r-protein S14 cDNA. Shown is the 546-bp sequence of pCS14-12, a cDNA clone prepared from HeLa mRNA. Single-base differences distinguishing human (top) from Chinese hamster (middle [45]) S14 cDNAs are indicated, as is the S14 polypeptide sequence (bottom). The *Hin*fI to *Ava*II restriction fragment noted was used as a primer in reverse transcriptase-catalyzed reactions to map 5′ ends of HeLa cell S14 mRNAs in the cloned RPS14 gene (see Fig. 5).

ing the cells in monolayer tissue culture also have been described previously (4, 5, 27, 28). A library of human placenta *Mbo*I DNA fragments constructed in the lambda vector Charon 28 was the generous gift of Philip Leder (Harvard University). [α-^{32}P]dCTP (800 to 900 Ci/mMol) was purchased from New England Nuclear Corp. (Boston, Mass.). Nonradioactive nucleotides and M13 17-mer were obtained from P-L Biochemicals, Inc. (Milwaukee, Wis.). Oligo(dT_{10-12})-cellulose was from Sigma Chemical Co. (St. Louis, Mo.). Restriction endonucleases; oligo(dG)-tailed, *Pst*I-cleaved pBR322 DNA; *Escherichia coli* DNA polymerase I (Klenow fragment); T4 DNA ligase; RNase T_1; and RNase H were products of Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). Avian myeloblastosis virus reverse transcriptase was obtained from Life Sciences, Inc. (St. Petersburg, Fla.). All enzymes were used according to procedures specified by the suppliers.

Isolation of human S14 cDNA and genomic clones. Polyadenylated mRNA was purified from exponentially growing HeLa cells and fractionated on the basis of size by sucrose gradient centrifugation. Double-stranded cDNAs were prepared from 10 to 12S mRNAs, and these were ligated into the oligo(dG)-tailed *Pst*I site of pBR322. Recombinant plasmids were introduced into $CaCl_2$-treated *E. coli* LE392. Resulting tetracycline-resistant, ampicillin-sensitive colonies were screened for human RPS14 cDNAs by colony hybridization with a purified CHO cell RPS14 cDNA fragment as the nucleic acid hybridization probe (19, 20, 35, 45). A 20-μg sample of size-fractionated mRNA yielded approximately 250 ng of double-stranded cDNA and, subsequently, six human RPS14 cDNAs.

A total of 9×10^5 PFU of bacteriophage lambda Charon 28 (46) carrying partial *Mbo*I fragments of human placental DNA were plated on *E. coli* LE392 and screened by plaque hybridization with CHO and human RPS14 cDNAs as nucleic acid probes (2). Two clones containing human S14 sequences, HGS14-1 and HGS14-2, were recognized.

Restriction endonuclease maps and purification of DNA fragments. Cloned DNAs were mapped by partial digestion with individual restriction endonucleases and by limit digestion with endonucleases used singly and in pairs. Individual fragments were purified by electrophoresis through agarose

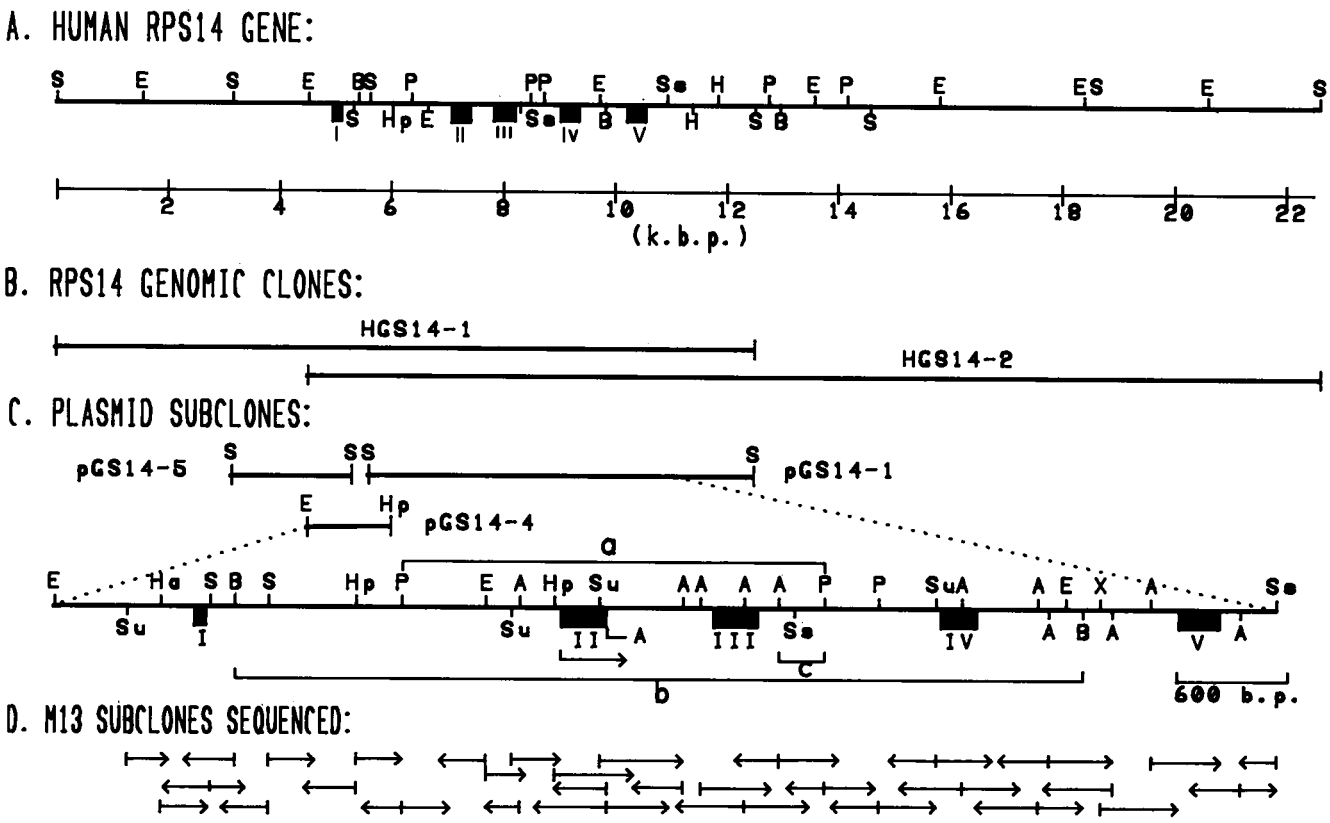


FIG. 2. Restriction endonuclease map and DNA sequencing strategy for human genomic S14 clones. (A) Restriction endonuclease map of two overlapping human placental S14 genomic clones isolated from a Charon 28 DNA library. Shaded boxes (labeled I through V) indicate homology to the cDNA probe (exons). (B) Boundaries of two human S14 genomic clones. HGS14-1 is approximately 12.5 kbp long, and HGS14-2 is 18 kbp. (C) Three genomic DNA fragments subcloned into pBR322. pGS14-1 and pGS14-4 were derived from HGS14-2, and pGS14-5 was from HGS14-1. A detailed map of the 5.9 kbp encoding the S14 sequence is indicated. a and b, Restriction fragments previously mapped to the functional human S14 locus at chromosome 5q23-q33 (34). c, Unique intervening sequence fragment also mapped to chromosome 5q23-q33 (see Fig. 3B). (D) M13mp10 and M13mp11 subclones whose DNA sequences were determined. The limits and directions of DNA sequence determinations are indicated by arrows. Restriction endonuclease cleavage sites: S, *Sma*I; E, *Eco*RI; B, *Bam*HI; Hp, *Hpa*I; P, *Pst*I; Ss, *Sst*I; H, *Hin*dIII; Ha, *Hae*III; Su, *Sau*3A; A, *Alu*I; X, *Xba*I.

or polyacrylamide gels. Bands were excised from gels, eluted by electrophoresis into 4 M NaCl, and precipitated by the addition of 2.5 volumes of ethanol. For subcloning, fragments were ligated into appropriate vectors (pBR322, M13mp10, or M13mp11) and reintroduced into $CaCl_2$-treated *E. coli* LE392 or JM103 (29).

DNA sequence and cap site determinations. DNA fragments subcloned into M13mp10 or M13mp11 were sequenced by the dideoxynucleotide chain termination method with M13 17-mer primer, *E. coli* DNA polymerase I (Klenow fragment), and denaturing polyacrylamide gradient gels (3, 29, 45, 50). Sequences were analyzed in a microcomputer by using the SEQAID package of software (45).

A 73-bp *Hin*fI-*Ava*II restriction endonuclease fragment derived from the 5′ end of a HeLa S14 cDNA clone (pCS14-12; see Fig. 1) was end labeled with Klenow fragment and [α-^{32}P]dCTP. Approximately 2 ng of the fragment was annealed wth 50 μg of poly(A)$^+$ HeLa mRNA in 80% formamide containing 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)], 1 mM EDTA, and 400 mM NaCl (pH 6.4) for 8 h at 55°C. Nucleic acids were recovered by ethanol precipitation and dissolved in 80 μl of 50 mM Tris hydrochloride–6 mM $MgCl_2$–40 mM KCl–1 mM dithiothreitol–100 ng of bovine serum albumin per ml (pH 8.3). The solution was adjusted to a 100 μM concentration of all four deoxyribonucleoside triphosphates (dATP, dGTP, dCTP, and dTTP), and aliquots (17 μl) were supplemented with 60 μM dideoxyATP (ddATP), 50 μM ddCTP, 25 μM ddGTP, or 100 μM ddTTP. Reverse transcriptase (30 U) was added to the four aliquots and to a control mixture that lacked dideoxynucleotides. After incubation for 30 min at 37°C, DNA synthesis was terminated by the addition of EDTA (25 mM), and reaction mixtures were extracted with phenol-chloroform-isoamyl alcohol (50:48:2). Primer extension products were resolved by electrophoresis through a denaturing polyacrylamide gel and detected by autoradiography.

RESULTS

Human S14 cDNA clone. Six nearly full-length cDNA clones of HeLa S14 mRNA were isolated, and the complete DNA sequence of one (pCS14-12) was determined (Fig. 1). The human S14 cDNA is identical to CHO S14 cDNA (45) in 476 of 540 residues. The two cDNAs differ most significantly at their 5′ and 3′ ends. Figure 1 illustrates the DNA sequence of pCS14-12, the residues that distinguish it from Chinese hamster S14 cDNA, and the amino acid sequence of the human S14 protein aligned with an initiator methionine codon at position 34. Human and CHO RPS14 polypeptide sequences are identical; the 41 single base differences that

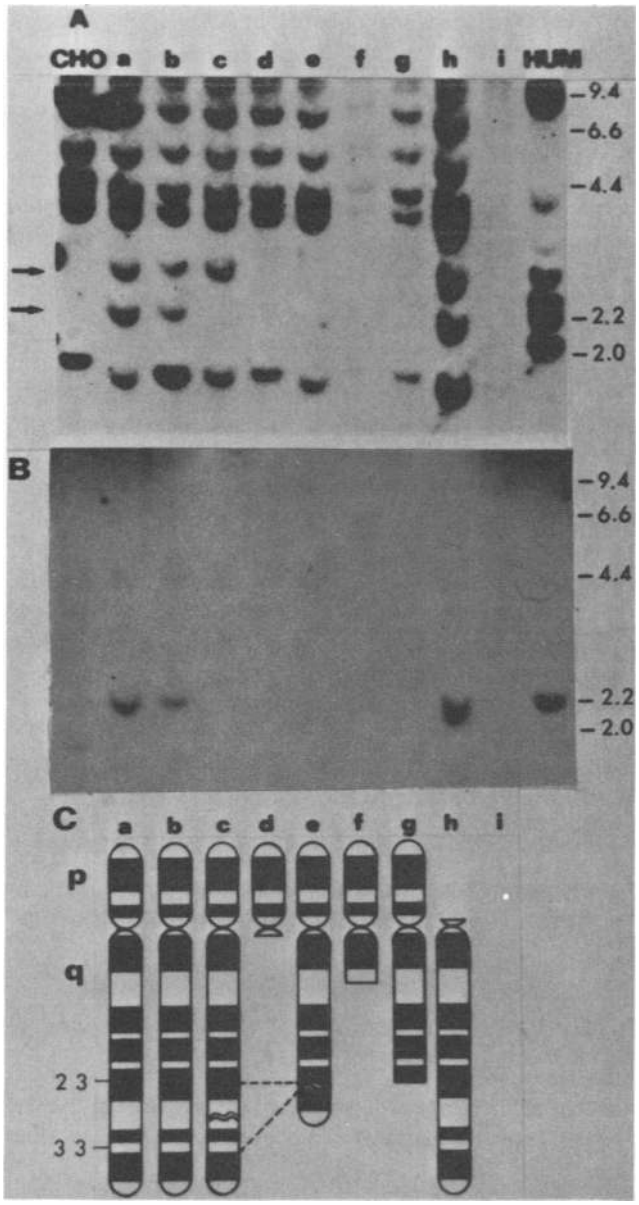


FIG. 3. Southern blot of DNAs from a human × Chinese hamster hybrid panel containing defined fragments of human chromosome 5. Genomic DNAs from Chinese hamster (CHO) and human (HUM) cells and nine CHO × human somatic cell hybrid clones, each possessing an intact human chromosome 5 or defined chromosome 5 fragment (34), were digested with *Pst*I, electrophoresed through agarose, and blotted to a GeneScreen Plus (New England Nuclear Corp.) membrane. The membrane was analyzed first with full-length human S14 cDNA (pCS14-12) as the hybridization probe (panel A) and then in panel B with a unique intervening sequence probe prepared from HGS14-2 (fragment c, Fig. 2C). Panel C depicts ideograms of the human chromosome 5 fragments in each hybrid clone. Molecular weight markers indicate migration of bacteriophage lambda *Hin*dIII fragments. Arrows to the left of panel A indicate the 2.2- and 2.5-kbp fragments of S14 DNA that map to human chromosome 5q.

distinguish human from Chinese hamster S14 coding sequences (30 transitions and 11 transversions) are silent with respect to the genetic code. Of the 41 nucleic acid differences, 40 affect third positions of their respective codons.

Transcriptionally active S14 locus on human chromosome 5. Purified Chinese hamster and human S14 cDNAs were used to survey a Charon 28 human placental DNA library. Two positive clones, HGS14-1 and HGS14-2, were detected. These clones contain 12.5 and 18 kbp of DNA that overlap for approximately 8 kbp (Fig. 2). Together the clones define 22.5 kbp of human chromosomal DNA whose restriction endonuclease map is depicted in Fig. 2A. DNA fragments containing S14 exon sequences (shaded boxes) were recognized by Southern blots (35, 52) with S14 cDNA as the hybridization probe.

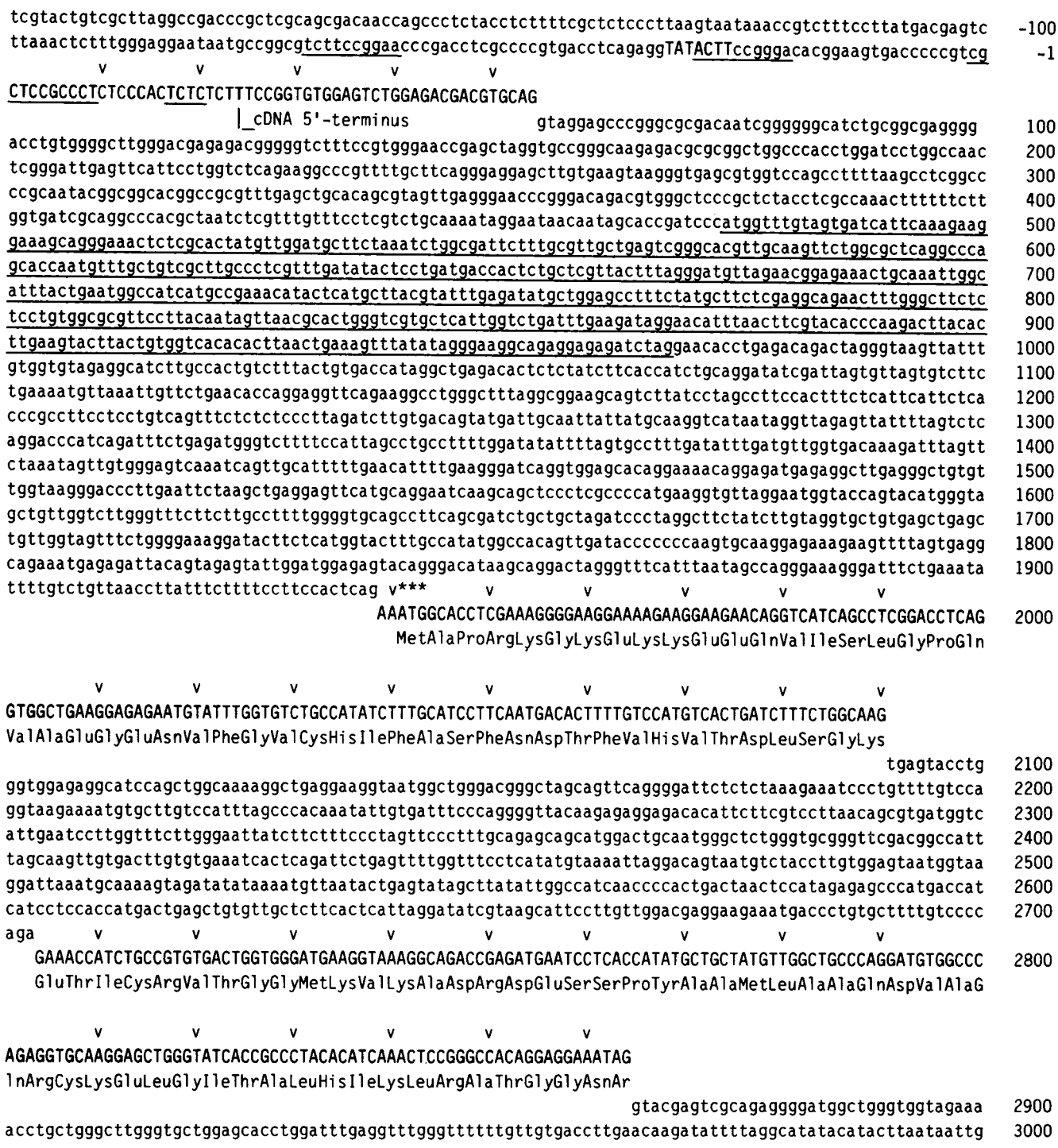
Two fragments of HGS14-2 and one of HGS14-1 that encompass the five S14 exons were subcloned into pBR322 for detailed restriction endonuclease and DNA sequence analyses. These are indicated in Fig. 2C together with a detailed restriction map of the DNA spanning the S14 exons. pGS14-5 is a *Sma*I fragment derived from HGS14-1. pGS14-1 (also a *Sma*I fragment) and pGS14-4 (an *Eco*RI-*Hpa*I fragment) were subcloned from HGS14-2.

Previously, when filter blots containing DNA restriction fragments prepared from panels of highly reduced human × Chinese hamster hybrid clones were analyzed with S14 cDNA probes, several bands were mapped to the functional human RPS14 locus on the long arm of chromosome 5 (34). Among these were a 4.5-kbp *Bam*HI fragment and a pair of *Pst*I fragments (2.2 and 2.5 kbp). The restriction map in Fig. 2C indicates three DNA fragments labeled a, b, and c. Fragments a and b are the same sizes as two of the RPS14 DNA bands mapped to the human locus on chromosome 5. Their presence in the genomic DNA clone strongly suggested that HGS14-1 and HGS14-2 derive from the active RPS14 locus. To confirm this, we purified a short restriction endonuclease fragment of single-copy DNA from intron 3 in HGS14-2 (labeled c in Fig. 2C). When used as a hybridization probe, this 260-bp fragment detected a single restriction fragment band in human DNA (lane HUM, Fig. 3B) and cross-reacted only slightly with Chinese hamser cell DNA (lane CHO, Fig. 3B).

A filter blot of *Pst*I-digested hybrid-cell DNAs was analyzed sequentially with full-length S14 cDNA (Fig. 3A) and fragment c (Fig. 3B) hybridization probes. The karyotypes of the hybrid clones with respect to human chromosome 5 and subchromosomal pieces derived from it are indicated by ideograms in Fig. 3C. As described before (34), S14 cDNA probe detected two chromosome 5q *Pst*I bands (2.2 and 2.5 kbp; Fig. 3A). On the other hand, the single-copy genomic DNA probe detected only one (the 2.2-kbp *Pst*I band; Fig. 3B). This confirmed that HGS14-1 and -2 derive from the transcriptionally active RPS14 gene on human chromosome 5q, the same chromosomal region which complements drug resistance in *emtB* CHO × human cell hybrids (9–11).

Human chromosome 5 in hybrid clone HHW 401 (lane c, Fig. 3) contains a submicroscopic interstitial deletion that includes S14 exons 1 to 3 (34). It lacks the 2.2-kbp fragment (labeled a in Fig. 2C) that is detected by the full-length cDNA probe (Fig. 3A). Because HHW 401 DNA includes the 2.5-kbp *Pst*I band encoding RPS14 exons 4 and 5 (Fig. 3A), the deletion breakpoint within the RPS14 gene must be to the left of the *Pst*I site 350 bp upstream from exon 4 (Fig. 2C). As fragment c probe detects no sequences in HHW 401 DNA, the deletion breakpoint must be very close to and likely downstream of the *Pst*I site at the right end of fragment c. Otherwise, fragment c hybridization probe would have detected a band in the DNA of the hybrid.

Primary structure of the human S14 gene. Portions of plasmids pGS14-1, -4, and -5 were subcloned into M13mp10



and M13mp11 for DNA sequence determination (Fig. 2D). DNA fragments with nonidentical 6-base restriction site termini (*Sst*I-*Pst*I or *Eco*RI-*Pst*I) were partially digested with 4-base endonucleases (*Alu*I or *Sau*3A) and ligated in defined orientations into the M13 polylinker. This strategy required that cloned DNAs contain a 6-base restriction site as one terminus and a 4-base cut site as the other. It thereby yielded nested sets of DNA fragments suitable for determining slightly more than 6 kbp of human S14 DNA sequence (Fig. 2D).

A total of 200 bp of nucleic acid sequence 5′ to the RPS14 exons were determined (Fig. 4). Similar to three mouse r-protein genes whose primary structures are known (13, 58, 63), the DNA immediately upstream of human RPS14 is extremely rich in pyrimidine nucleotides. In contrast to the murine r-protein and other mammalian housekeeping genes, human RPS14 contains a bona fide TATA sequence (residues −30 to −25; Fig. 4). In addition, the S14 5′-flanking sequence includes three copies of a 10-bp sequence motif at positions −69, −29, and +1 (underlined in Fig. 4). Se-

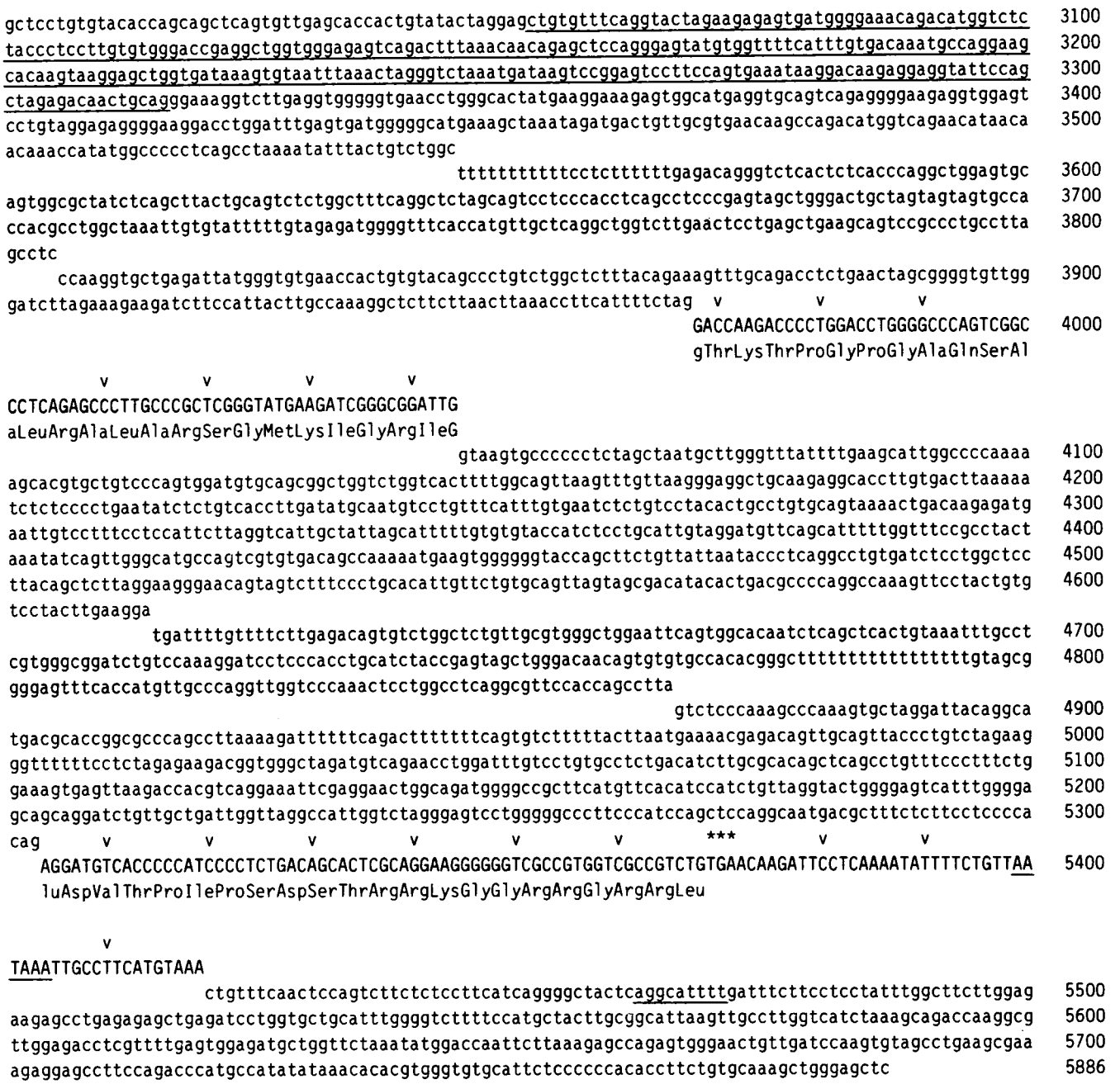


FIG. 4. DNA sequence of the human S14 locus. A DNA sequence including the five RPS14 exons (6,086 bp) was determined from both strands of the cloned S14 gene as indicated in Fig. 2D. Exons are displayed in uppercase type; flanking and intervening sequences are in lowercase type. Sequences underlined or otherwise highlighted are described in the text.

quences related to this motif also are observed in analogous 5′-flanking positions of several yeast and mouse r-protein genes (Table 1).

pCS14-12 cDNA aligns with the genomic DNA sequence beginning with residue 24 (Fig. 4). mRNA-directed primer extension (Fig. 5) and S1 protection experiments (not shown) indicated that HeLa S14 mRNAs exhibit heterogeneous 5′ termini. Less than 10% of the transcripts extend to pyrimidine residues at positions 1 and 11 (Fig. 4). The majority of S14 mRNA 5′ ends (>90%) mapped within the TCTCTC hexanucleotide between residues +17 and +22. Because the 5′-most dC residue indicated by primer extension is located 28 bp downstream from a TATA sequence and this conforms to expectations for a typical eucaryotic mRNA cap site (7), we indexed the sequence on that residue in Fig. 4. We emphasize, however, that no data are available (i) to implicate the TATA sequence at −33 (Fig. 4) in mechanisms that regulate the transcription of RPS14; (ii) to account for heterogeneity among the 5′ ends of S14 transcripts; and (iii) to document the presence (or absence) of methylated caps on all the S14 transcripts indicated in Fig. 5. This information awaits development of an RPS14 expression system and appropriate in vitro mutagenesis experiments.

The first RPS14 intron possesses an unusually high pyrimidine content that includes several oligo(dT) sequences. In

TABLE 1. Short sequence motif in the DNA upstream from S14[a]

Species	Gene	Position	Sequence
Human	S14	−69	TCTTCCGGAA
		−29	ACTTCCGGGA
		−2	CGCTCCGCCC
Mouse	L30	−77	CGGACCGGCT
		−65	TCTTCCGCTT
	L32	−69	GCTTCCCTTT
		−38	CCGTCCCATC
	S16	−211	TCTTCCGAGT
		−187	CCCTCCGGGC
		−157	CCCTCCGACA
		−142	CCCTCCCTTG
		−109	ACGTCCCCTC
Yeast	rp59	−155	CCATCCGCCT
		−149	GCCTCCTGGC
		−120	GCTTCCTCAG
	rp51	−149	GCTTCCTAAG
Consensus sequence[b]			SSYTCCBRNN

[a] Three variants of a 10-bp sequence motif from the chromosomal DNA preceding S14 exon 1 (underlined in Fig. 4) are aligned with each other and with related sequences from analogous positions in murine and yeast r-protein genes. Locations of the mammalian sequences are expressed according to the numbering scheme in Fig. 4. For lack of comparable information regarding the yeast genes, their positions are relative to translational initiator codons. DNA sequences of the mouse and yeast r-protein genes were taken from the following sources: mouse L30 (63), L32 (13), and S16 (58); yeast rp59 (54) and rp51 (55).

[b] S, G or C; Y, C or T; R, A or G; B, G, T, or C; N A, G, T, or C.

addition, it contains an open reading frame (underlined in Fig. 4) that could encode a 164-amino acid polypeptide. Because filter-bound HGS14-1 and -2 DNAs hybrid select only the 650-base transcript previously identified as S14 mRNA (35), we have no reason to believe that a stable human mRNA derives from this open reading frame. Further, computer-based surveys (26) of protein and nucleic acid sequence data bases failed to indicate known DNA or polypeptide sequences homologous to this open reading frame.

The coding sequence for r-protein S14 initiates with an ATG codon near the beginning of exon 2. An intron interrupts the exon after 50 amino acid codewords. The termini of this intron do not conform to the GT..AG rule (7, 32). It should be noted, however, that if the HeLa S14 gene (the source of cDNAs analyzed) were to differ from the cloned human placenta S14 allele by a single-base, silent polymorphism (G versus A at position 2090), the boundaries of intron 2 could be aligned to conform with the rule. That three independent HeLa S14 cDNAs and several Charon 28 and M13 clones of the placental S14 allele all displayed the intron 2 junctional DNA sequences depicted in Fig. 1 and 4 indicates that cloning and genetic artifacts in bacterial hosts do not account for this polymorphism. Because the placental tissue and DNA from which the Charon 28 library was constructed are no longer available, it is impossible to clarify this point.

Intron 3 was the source of the 260-bp single-copy DNA probe (fragment c, Fig. 2C) used to assign these clones to the transcriptionally active S14 locus at human chromosome 5q23-q33 (Fig. 4, underlined). Introns 3 and 4 both contain 250-bp repeated DNA sequences (Fig. 4, offset) that can be aligned with the BLUR8 clone of an *Alu* sequence (12, 48) to produce an approximately 80% sequence match (Fig. 6). Whether or not *Alu* sequences that flank S14 exon 4 are significant with respect to the evolutionary history of the gene is not clear. However, it is intriguing that the intron 3 *Alu* sequence resides close to one breakpoint of the interstitial deletion affecting the human chromosome 5 in hybrid cell clone HHW 401 described above.

Exon 5 encodes the carboxyl 22 amino acids of r-protein S14, including the two residues affected by *emtB* mutations in Chinese hamster cells, the Arg codons at positions 5360 and 5363 (45). The S14 translational reading frame terminates in a TGA codon at position 5369, followed by a polyadenylation signal (AATAAA) at residue 5399. S14 genomic and cDNA sequences diverge 16 bp downstream at the end of exon 5. Forty-two bases further 3′ (position 5462) is a 10-base sequence complementary to U1 small nuclear RNA (underlined in Fig. 4). This sequence, observed in the DNAs flanking 3′ ends of other eucaryotic viral and cellular genes, is thought to facilitate polyadenylation of mRNAs (49).

DISCUSSION

We describe full-length human RPS14 cDNA and genomic clones. The HeLa cell S14 cDNA encodes a 151-amino acid protein identical to the polypeptide specified by Chinese hamster S14 cDNA (45). The human and Chinese hamster

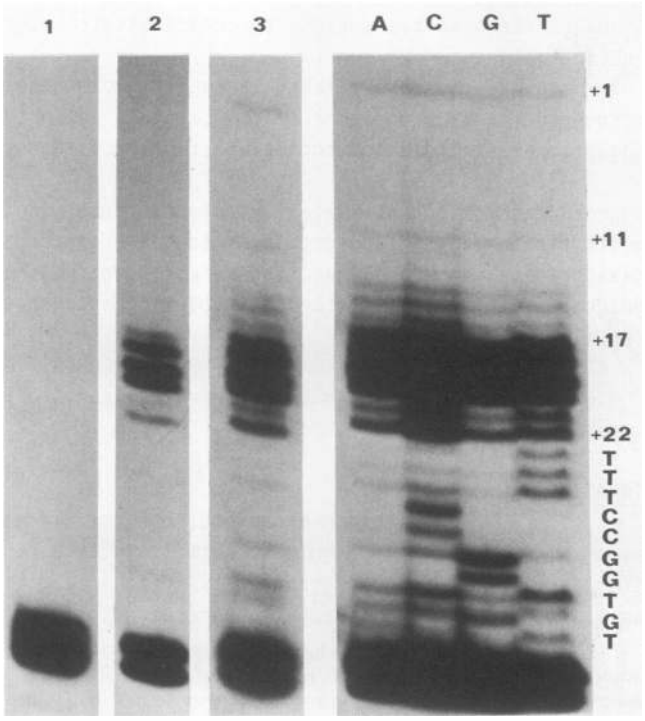


FIG. 5. Determination of the 5′ ends of HeLa S14 mRNAs. Polyadenylated HeLa mRNA was annealed to the 5′-end-labeled *Hin*fI-*Ava*II cDNA fragment indicated in Fig. 1, and the radioactive primer was extended by reverse transcription. The sample in lane 1 contained only the primer (18-day autoradiographic exposure). Lanes 2 (4-day exposure) and 3 (18-day exposure) contained a complete primer extension reaction including four deoxynucleoside triphosphates. Reactions in lanes A, C, G, and T (18-day exposure) also contained the dideoxynucleoside triphosphate required to read the coding-strand sequence indicated to the right of the figure. This sequence aligns perfectly with the genomic DNA sequence from the *Hin*fI site to the TCTCTC hexanucleotide (residues 34 to 23; Fig. 4) and verifies fidelity of the primer extension. The weak dC and dT bands at positions +1 and +11 as well as the TCTCTC bands between positions +17 and +22 are indicated at the right.

FIG. 6. *Alu* sequences in human S14 introns 3 and 4. Sequences in S14 introns 3 and 4 (IVSIII and IVSIV) (positions 3052 to 3316 and 3545 to 3805; Fig. 4) are aligned with the BLUR8 *Alu* sequence (12, 48) by using the computer program NUCALN (26). Alignments illustrated were based on "ktuples" of 2, gap penalties of 7, and scan windows of 20 bp. Matching base pairs are indicated by colons (:). Hyphens represent gaps suggested by NUCALN to optimize sequence alignment.

S14 protein-coding sequences (cDNAs) differ by 41 bp. All these differences are silent with respect to amino acid codewords. Of the 41 differences, 40 occur in codon third positions.

Three mouse r-protein genes whose structures are known were correlated with transcriptionally active loci by virtue of unique intervening sequences detected in high-molecular-weight nuclear transcripts (13, 58, 63). While these data indicated that the mouse genes are transcribed, they did not demonstrate that the transcripts are processed and translated into functional r-proteins. We used somatic cell genetics and cytogenetics and a single-copy DNA hybridization probe derived from intervening sequence to correlate human genomic clones with the functional RPS14 locus at chromosome 5q23-q33.

Human RPS14 exhibits several features similar to mouse r-protein genes already described (13, 58, 63). It is composed of multiple exons and introns, and its flanking and intervening sequences are unusually rich in pyrimidines. The human gene contains three related copies of a short sequence motif in its 5'-flanking region and first exon. Related versions of this motif also occur in analogous positions of several mouse and yeast r-protein genes (Table 1). Mouse r-protein genes contain a different repetitive motif (13, 58, 63) that is not a part of the human RPS14 gene.

Models describing cellular mechanisms that regulate multigene families such as r-protein genes can be investigated by two approaches. (i) One can compare the structures of individual genes to recognize common features that might interact with regulatory components; and (ii) one can modify specific regions of an active gene, such as the repeated sequences or the TATA motif preceding the RPS14 gene, to discern features whose alteration disrupts normal gene expression. RPS14 is the only mammalian r-protein gene suitable for both biochemical and genetic aspects of the second approach. It offers the possibility of modifying a cloned human or Chinese hamster r-protein gene carried in a mammalian expression vector and then assessing functional consequences by transformation of appropriate mutant (*emtB*) cells.

ACKNOWLEDGMENTS

We wish to thank Maggy Weston-Fina and Barbara Van Slyke for their excellent technical assistance and Susan Brown for her suggestions regarding the manuscript.

This research was supported by Public Health Service grant GM23013 to D.J.R. from the National Institutes of Health. D.D.R. is a predoctoral trainee supported by training grant CA09418 from the National Institutes of Health.

LITERATURE CITED

1. **Amaldi, F., E. Beccari, I. Bozzoni, Z. X. Luo, and P. Pierrei-Amaldi.** 1982. Nucleotide sequences of cloned cDNA fragments specific for six *Xenopus laevis* ribosomal proteins. Gene **17**:311–316.
2. **Benton, W., and R. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–182.
3. **Biggin, M., T. Gibson, and G. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.

4. **Boersma, D., S. McGill, J. Mollenkamp, and D. Roufa.** 1979. Emetine resistance in Chinese hamster cells. Analysis of ribosomal proteins prepared from mutant cells. J. Biol. Chem. **254**:559–567.
5. **Boersma, D., S. M. McGill, J. W. Mollenkamp, and D. J. Roufa.** 1979. Emetine resistance in Chinese hamster cells is linked genetically with an altered 40S ribosomal subunit protein S20. Proc. Natl. Acad. Sci. USA **76**:415–419.
6. **Bozzoni, I., E. Beccari, Z. X. Luo, and F. Amaldi.** 1981. *Xenopus laevis* ribosomal protein genes. Isolation of recombinant cDNA clones and study of their genomic organization. Nucleic Acids Res. **9**:1069–1086.
7. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
8. **Burns, D., B. Stark, M. Macklin, and W. Chooi.** 1984. Isolation and characterization of cloned DNA sequences containing ribosomal protein genes of *Drosophila melanogaster*. Mol. Cell. Biol. **4**:2643–2652.
9. **Dana, S., S. Chang, and J. Wasmuth.** 1985. Synthesis and incorporation of human ribosomal protein S14 into functional ribosomes in human-Chinese hamster cell hybrids containing human chromosome 5: human RPS14 gene is the structural gene for ribosomal protein S14. Somatic Cell Mol. Genet. **11**:625–632.
10. **Dana, S., and J. Wasmuth.** 1982. Selective linkage distribution in human-Chinese hamster cell hybrids: deletion mapping of the *leuS*, *hexB*, *emtB*, and *chr* genes on human chromosome 5. Mol. Cell. Biol. **2**:1220–1228.
11. **Dana, S., and J. J. Wasmuth.** 1982. Linkage of the *leu s*, *emt b* and *chr* genes on chromosome 5 in humans and expression of genes encoding protein synthetic components in human-Chinese hamster hybrids. Somatic Cell Genet. **8**:245–264.
12. **Deininger, P., D. Jolly, C. Rubin, T. Friedmann, and C. Schmid.** 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. **151**:17–33.
13. **Dudov, K., and R. Perry.** 1984. The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. Cell **37**:457–468.
14. **Fabijanski, S., and M. Pellegrini.** 1982. Isolation of a cloned DNA segment containing a ribosomal protein gene of *Drosophila melanogaster*. Gene **18**:267–276.
15. **Faliks, D., and O. Meyuhas.** 1982. Coordinate regulation of ribosomal protein mRNA level in regenerating rat liver. Study with the corresponding mouse cloned cDNAs. Nucleic Acids Res. **10**:789–801.
16. **Fallon, A. M., C. S. Jinks, G. D. Strycharz, and M. Nomura.** 1979. Regulation of ribosomal protein synthesis in *Escherichia coli* by selective mRNA inactivation. Proc. Natl. Acad. Sci. USA **76**:3411–3415.
17. **Geyer, P. K., O. Meyuhas, R. P. Perry, and L. F. Johnson.** 1982. Regulation of ribosomal protein mRNA content and translation in growth-stimulated mouse fibroblasts. Mol. Cell. Biol. **2**:685–693.
18. **Gorenstein, C., and J. Warner.** 1976. Coordinate regulation of the synthesis of eukaryotic ribosomal proteins. Proc. Natl. Acad. Sci. USA **73**:1547–1551.
19. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization, a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.
20. **Hanahan, D., and M. Meselson.** 1980. Plasmid screening at high colony density. Gene **10**:63–67.
21. **Ignotz, G. G., S. Hokari, R. M. Dephilip, K. Tsukada, and I. Lieberman.** 1981. Lodish model and regulation of ribosomal protein synthesis by insulin-deficient chick embryo fibroblasts. Biochemistry **20**:2550–2558.
22. **Kief, D. R., and J. R. Warner.** 1981. Coordinate control of syntheses of ribosomal ribonucleic acid and ribosomal proteins during nutritional shiftup in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **1**:1007–1015.
23. **Klein, A., and O. Meyuhas.** 1984. A multigene family of intron lacking and containing genes, encoding for mouse ribosomal protein L7. Nucleic Acids Res. **12**:3763–3776.
24. **Krauter, K. S., R. Soeiro, and B. Nadal-Ginard.** 1980. Uncoordinate regulation of ribosomal RNA and ribosomal protein synthesis during 16e9 myoblast differentiation. J. Mol. Biol. **142**:145–159.
25. **Kuwano, Y., O. Nakanishi, Y. O. Nabeshima, T. Tanaka, and K. Ogata.** 1985. Molecular cloning and nucleotide sequence of DNA complementary to rat ribosomal protein S26 messenger RNA. J. Biochem. **97**:983–992.
26. **Lipman, D., and W. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
27. **Madjar, J. J., M. Frahm, S. McGill, and D. J. Roufa.** 1983. Ribosomal protein S14 is altered by two-step emetine resistance genes in Chinese hamster cells. Mol. Cell. Biol. **3**:190–197.
28. **Madjar, J. J., K. Nielsen-Smith, M. Frahm, and D. J. Roufa.** 1982. Emetine resistance in Chinese hamster ovary cells is associated with an altered ribosomal protein S14 mRNA. Proc. Natl. Acad. Sci. USA **79**:1003–1007.
29. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
30. **Meyuhas, O.** 1984. Ribosomal protein gene expression in proliferating and nonproliferating cells, p. 243–271. *In* G. S. Stein and J. L. Stein (ed.), Recombinant DNA and cell proliferation. Academic Press, Inc., New York.
31. **Moreland, R., H. Nam, L. Hereford, and H. Fried.** 1985. Identification of a nuclear localization signal of a yeast ribosomal protein. Proc. Natl. Acad. Sci. USA **82**:6561–6565.
32. **Mount, S.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **19**:459–472.
33. **Nabeshima, Y. I., and K. Ogata.** 1980. Stimulation of the synthesis of ribosomal proteins in regenerating rat liver with special reference to the increase in the amounts of effective mRNAs for ribosomal proteins. Eur. J. Biochem. **107**:323–329.
34. **Nakamichi, N., F.-T. Kao, J. Wasmuth, and D. Roufa.** 1986. Ribosomal protein sequences map to human chromosomes 5, 8 and 17. Somatic Cell Mol. Genet. **12**:225–236.
35. **Nakamichi, N., D. D. Rhoads, and D. J. Roufa.** 1983. The Chinese hamster cell emetine resistance gene—analysis of cDNA and genomic sequences encoding ribosomal protein S14. J. Biol. Chem. **258**:13236–13242.
36. **Nakanishi, O., M. Oyanagi, Y. Kuwano, T. Tanaka, T. Nakayama, H. Mitsui, Y. I. Nabeshima, and K. Ogata.** 1985. Molecular cloning and nucleotide sequences of cDNAs specific for rat liver ribosomal proteins S17 and L30. Gene **35**:289–296.
37. **Nielsen-Smith, K., S. M. McGill, M. Frahm, and D. J. Roufa.** 1983. Biochemical evidence for multiple independent emetine resistance genes in Chinese hamster cells. Mol. Cell. Biol. **3**:198–202.
38. **Nomura, M., R. Gourse, and G. Baughman.** 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annu. Rev. Biochem. **53**:75–118.
39. **Nomura, M., and W. Held.** 1974. Reconstitution of ribosomes: studies of ribosome structure, function and assembly, p. 193–224. *In* M. Nomura, A. Tissieres, and P. Lengyel (ed.), Ribosomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Nomura, M., J. L. Yates, D. Dean, and L. E. Post.** 1980. Feedback regulation of ribosomal protein gene expression in *Escherichia coli*—structural homology of ribosomal RNA and ribosomal protein mRNA. Proc. Natl. Acad. Sci. USA **77**:7084–7088.
41. **O'Connell, P., and M. Rosbash.** 1984. Sequence, structure, codon preference of the *Drosophila* ribosomal protein 49 gene. Nucleic Acids Res. **12**:5495–5513.
42. **Pearson, N. J., H. M. Fried, and J. R. Warner.** 1982. Yeast use translational control to compensate for extra copies of a ribosomal protein gene. Cell **29**:347–355.
43. **Peled-Yalif, E., I. Cohen-Binder, and O. Meyuhas.** 1984. Isolation and characterization of four mouse ribosomal protein L18 genes that appear to be processed pseudogenes. Gene **29**: 157–166.

44. **Pierandrei-Amaldi, P., E. Beccari, I. Bozzoni, and F. Amaldi.** 1985. Ribosomal protein production in normal and anucleolate *Xenopus* embryos: regulation at the posttranscriptional and translational levels. Cell **42**:317–323.
45. **Rhoads, D. D., and D. J. Roufa.** 1985. Emetine resistance in Chinese hamster cells: structures of wild-type and mutant ribosomal protein S14 mRNAs. Mol. Cell. Biol. **5**:1655–1659.
46. **Rimm, D. L., D. Horness, J. Kucera, and F. R. Blattner.** 1980. Construction of coliphage lambda Charon vectors with Bam HI cloning sites. Gene **12**:301–309.
47. **Rosbash, M., P. K. W. Harris, J. L. Woolford, and J. L. Teem.** 1981. The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. Cell **24**:679–686.
48. **Rubin, C., C. Houck, P. Deininger, T. Friedmann, and C. Schmid.** 1980. Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature (London) **284**:372–374.
49. **Sadofsky, M., S. Connelly, J. Manley, and J. Alwine.** 1985. Identification of a sequence element on the 3′ side of AAUAAA which is necessary for simian virus 40 late mRNA 3′-end processing. Mol. Cell. Biol. **5**:2713–2719.
50. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94**:441–448.
51. **Santon, J. B., and M. Pellegrini.** 1980. Expression of ribosomal proteins during *Drosophilia* early development. Proc. Natl. Acad. Sci. USA **77**:5649–5653.
52. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by agarose gel electrophoresis. J. Mol. Biol. **98**:503–517.
53. **Stanners, C. P., G. L. Eliceiri, and H. Green.** 1971. Two types of ribosomes in mouse-hamster hybrid cells. Nature (London) New Biol. **230**:52–54.
54. **Teem, J., N. Abovich, N. Kaufer, W. Schwindinger, J. Warner, A. Levy, J. Woolford, R. Leer, M. Van Raamsdonk-Duin, W. Mager, R. Planta, L. Schultz, J. Friesen, H. Fried, and M. Rosbash.** 1984. A comparison of yeast ribosomal protein gene DNA sequences. Nucleic Acids Res. **12**:8295–8312.
55. **Teem, J., and M. Rosbash.** 1983. Expression of a β-galactosidasegene containing the ribosomal protein 51 intron is sensitive to the *rna2* mutation of yeast. Proc. Natl. Acad. Sci. USA **80**:4403–4407.
56. **Toniolo, D., and C. Basilico.** 1974. Complementation of a defect in the production of ribosomal RNA in somatic cell hybrids. Nature (London) **248**:411–413.
57. **Vaslet, C. A., P. O'Connell, M. Izquierdo, and M. Rosbash.** 1980. Isolation and mapping of a cloned ribosomal protein gene of *Drosophila melanogaster*. Nature (London) **285**:674–676.
58. **Wagner, M., and R. Perry.** 1985. Characterization of the multigene family encoding the mouse S16 ribosomal protein: strategy for distinguishing an expressed gene from its processed pseudogene counterparts by an analysis of total genomic DNA. Mol. Cell. Biol. **5**:3560–3576.
59. **Warner, J., G. Mitra, W. Schwindinger, M. Studeny, and H. Fried.** 1985. *Saccharomyces cerevisiae* coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover. Mol. Cell. Biol. **5**:1512–1521.
60. **Warner, J., R. Tushinski, and P. Wejksnora.** 1980. Coordination of RNA and proteins in eukaryotic ribosome production, p. 871–888. *In* G. Chambliss, G. R. Craven, J. Davies, K. Davis, L. Kahan, and M. Nomura (ed.), Ribosomes. Structure, function and genetics. University Park Press, Baltimore.
61. **Wejksnora, P. J., and J. R. Warner.** 1979. Hybrid mammalian cells assemble hybrid ribosomes. Proc. Natl. Acad. Sci. USA **76**:5554–5558.
62. **Wejksnora, P. J., and J. R. Warner.** 1981. Regulation of ribosomal RNA and proteins in mouse-hamster hybrid cells. J. Biol. Chem. **256**:9406–9413.
63. **Wiedemann, L., and R. Perry.** 1984. Characterization of the expressed gene and several processed pseudogenes for the mouse ribosomal protein L30 gene family. Mol. Cell. Biol. **4**:2518–2528.
64. **Worton, R. G., C. Duff, and C. E. Campbell.** 1980. Marker segregation without chromosome loss at the *emt* locus in Chinese hamster cell hybrids. Somatic Cell Genet. **6**:199–214.
65. **Worton, R. G., C. Duff, and W. Flintoff.** 1981. Microcell-mediated cotransfer of genes specifying methotrexate resistance, emetine sensitivity, and chromate sensitivity with Chinese hamster chromosome 2. Mol. Cell. Biol. **1**:330–335.
66. **Yates, J. L., A. E. Arfsten, and M. Nomura.** 1980. *In vitro* expression of *Escherichia coli* ribosomal protein genes—autogenous inhibition of translation. Proc. Natl. Acad. Sci. USA **77**:1837–1841.
67. **Yates, J. L., and M. Nomura.** 1980. *E. coli* ribosomal protein 14 is a feedback regulatory protein. Cell **21**:517–522.
68. **Zengel, J. M., D. Mueckl, and L. Lindahl.** 1980. Protein L4 of the *E. coli* ribosome regulates an eleven gene rprotein operon. Cell **21**:523–536.