
Poliovirus genome RNA hybridizes specifically to higher eukaryotic rRNAs

Marcella A. McClure*⁺ and Jacques Perrault⁺

Department of Microbiology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, and Molecular Biology Institute and Biology Department, San Diego State University, San Diego, CA 92182, USA

Received 28 August 1985; Accepted 12 September 1985

ABSTRACT

The RNA genome of poliovirus hybridizes to 28S and 18S rRNAs of higher eukaryotes under stringent conditions. The hybridization detected by Northern blot analyses is specific since little or no signal was detected for yeast or prokaryotic rRNAs or other major cellular RNAs. Southern blot analysis of DNA clones of mouse rRNA genes leads us to conclude that several regions of 28S rRNA, and at least one region in 18S rRNA, are involved in the hybridization to polio RNA, and that G/C regions are not responsible for this phenomenon. We have precisely mapped one of these hybridizing regions in both molecules. Computer analysis confirms that extensive intermolecular base-pairing (81 out of 104 contiguous bases in the rRNA strand) could be responsible for this one particular site of interaction (polio genome, bases 5075-5250; 28S rRNA, bases 1097-1200). We discuss the possible functional and/or evolutionary significance of this novel type of interaction.

INTRODUCTION

Picornaviruses are cytoplasmic RNA viruses of positive genome polarity, and include human pathogens such as polio, Coxsackie, and rhinoviruses, as well as the agriculturally important foot-and-mouth disease virus (FMDV). In contrast to the usual eukaryotic mRNAs, these viruses do not contain a cap structure at the 5' end of their messenger species although the genomes of the subgroup cardioviruses are very efficiently translated *in vitro* (1). The packaged genome strand contains a 5' end covalently-linked, virus-coded protein (VPg) which is apparently removed by cellular enzymes before translation on cellular polyribosomes (2). Classically, picornaviruses are thought to translate their monocistronic message from a single initiation site (741 bases from the 5' end in the case of polio) into a large polyprotein that is subsequently cleaved by viral, and possibly host protease(s), to give rise to all protein products (reviewed in ref. 3).

Ribosomal RNAs have long been known to play more than a passive structural role in the translation process. For prokaryotes, it is generally accepted that a short purine rich sequence, near the 3' end terminus of the small ribosomal subunit RNA, is involved through base-pairing in recognition of the

5' terminal translation initiation site of most, if not all, mRNAs. In contrast, there is little, if any, evidence for such a "Shine-Dalgarno" base-pairing interaction in initiation of mRNA translation in eukaryotic organisms. The 5' terminal cap structure, however, is thought to be involved in initial binding of the 40S ribosomal subunit at this end of the molecule. In addition, a "scanning function" for positioning the 40S subunit at the correct AUG start codon has been proposed (for a recent review of translation see ref. 4). For picornaviruses this would entail traversing several hundred nucleotides, including 7 AUG codons in the case of poliovirus, before arriving at the correct site for translation of the polyprotein (5,6).

We describe here, for the first time to our knowledge, that genomic RNA from polio, as well as Mengo and Coxsackie viruses, contain regions of complementarity to the higher eukaryotic-specific domains of both 18S and 28S rRNAs. We have mapped one of these regions of hybridization between mouse 28S rRNA and poliovirus RNA and discuss the possibility that this type of interaction may be involved in the translation of picornavirus genomes. We also consider common evolutionary ancestry for picornaviruses and rRNAs, and recombination/adaptive events as other possibilities underlying this phenomenon.

MATERIALS AND METHODS Cells, Viruses and RNA

Cells and picornaviruses were grown and viral RNA extracted as previously described (7), except that prior to sonication of purified virus, the 0.1 ml suspension was made 5 mM iodoacetic acid and the soluble carrier RNA was either omitted or replaced by addition of glycogen (20 µg/ml). All RNAs transferred to nitrocellulose paper, including those fractionated by agarose gel electrophoresis, were first glyoxal-denatured as described previously (7). Vesicular stomatitis virus (VSV) RNA from the Mudd-Summers Indiana strain was purified from virions as described before (8). RNA from cytoplasmic extracts and wheat germ S-30 fraction was purified by phenol extraction and Sarkosyl-proteinase K digestion (8). Purified cytoplasmic RNA from primary chicken embryo fibroblast was a gift from Dr. M. Schlesinger, purified rRNAs of yeast (Saccharomyces cerevisiae) from Dr. M. Olson, and gel purified B. stearothermophilus rRNAs and purified RNA from E. coli from Dr. D. Schlessinger (Washington University Medical School). Purified rRNAs from calf-liver and Saccharomyces cerevisiae were also obtained from P-L Biochemicals.

Growth, Purification and Restriction of Plasmids and Lambda Vector

The four pBR322-derived plasmids containing the inserts described in the text, and the λ gtWES/mrl100 vector were obtained from Dr. Lewis Bowman (9). All plasmids were grown and purified using a scaled up version of the method of Birnboim and Doly (10) followed by ethidium bromide-CsCl density centrifugation (11). The λ vector was essentially grown and purified as outlined by Tiemeier et al. (12) and its DNA extracted with proteinase K and phenol-chloroform as above. The infectious DNA clone of the poliovirus genome (pRV16) was obtained from Dr. B. Semler and grown as described (13).

Restriction enzymes were purchased from either Boehringer Mannheim (BamHI, EcoRI, Sall, HinfI) or New England Biolabs (AluI, HaeIII, HpaII, RsaI), and used as recommended by the suppliers, except for overnight incubations.

Gels, Electrophoretic Transfer, Probes and Hybridization

Electrophoresis of glyoxal-denatured RNAs on sodium phosphate-buffered agarose gels was carried out as before (7). Analysis of DNA on Tris-acetate agarose gels or acrylamide gels, and recovery of purified gel bands by electroelution into dialysis bags or overnight diffusion from slices of acrylamide were carried out according to Maniatis et al. (11). Restriction fragments were treated with calf alkaline phosphatase, and labeled with polynucleotide kinase and γ -³²P-ATP as previously described (14). Sequencing was carried out using the procedure of Maxam and Gilbert (15).

Electrophoretic transfer of DNA to nitrocellulose paper (Schleicher & Schuell, BA85) was carried out exactly as described by Smith et al. (16). It should be noted that these authors have shown equal efficiency of transfer and retention of DNA fragments onto nitrocellulose paper ranging in size from 2 kb-23 kb (from agarose gels), and 40 bases to 800 bases (from acrylamide gels) using this technique. For RNA transfers, alkaline denaturation was omitted and the blots were equilibrated with 20X SSC (3 M NaCl/0.3 M Na citrate) by diffusion through Whatman 3 MM paper (as in Southern transfer) for 30 min prior to baking in vacuo at 80° for 2 hr (this step reverses glyoxalation). The efficiency of transfer and retention throughout hybridization using this method was ~85-90% for ³²P-labeled 28S, 18S, and 4-5.8S RNAs (M.A. McClure, Ph.D. thesis, Washington University, St. Louis, 1984)

All RNA probes used for either Northern or Southern blot analyses were prepared by boiling the RNA in 100% formamide (deionized) for 40 min (modal size distribution after this treatment was ~75 nucleotides), before end-labeling by the polynucleotide kinase reaction. The average specific activity of RNA labeled under these conditions was 5×10^7 cpm/ μ g, and $5-20 \times 10^6$

cpm were used to probe each blot. The gel purified Sall/BamHI insert of clone 7 was "nick-translated" using the procedure of Meinkoth and Wahl (17).

All RNA:RNA blots, and the DNA:RNA blots of figures 3 and 6 were pre-hybridized at 42°C for 4-16 hrs, in 5X SSC, 0.05 M Na-phosphate (pH 6.5), soluble yeast RNA at 150 µg/ml, 0.2% SDS, and 50% formamide. Hybridization was then carried out at 42°C for 48 hrs in 4X SSC, 0.04 M Na phosphate (pH 6.5), soluble yeast RNA at 150 µg/ml, 0.16% SDS, 48% formamide and 10% dextran sulfate (Pharmacia). Following hybridization, the blots were washed at room temperature 4 times, 5 min each, in 2X SSC, 0.1% SDS, and then twice for 15 min in 0.1X SSC, 0.1% SDS at 60°C. These washing conditions are more stringent than the hybridization reaction, and are equivalent to ~10°C nearer the T_m value of the hybrids (18,19). All other DNA:RNA blot hybridizations of restriction fragments were carried out as above except for a hybridization temperature of 37°C and a formamide concentration of 40%, while the final set of washes were done at 52°C in 1X SSC. The stringency of this second set of hybridization conditions is lower than that of the first set by ~13°C for equivalent salt and formamide concentrations. The latter washes, however, are roughly equivalent to the hybridization conditions, and are less stringent than the washes from the previous set of blots by an equivalent of -23-25°C (18,19). Following the final washes the blots were dried briefly, wrapped in Saran Wrap, and exposed to Kodak XAR-5 film for 1-8 days. Quantitation of the labeled species in the autoradiograms was carried out by densitometry (EC910 Transmission Densitometer).

Computer Analysis

Analysis was carried out on a VAX 11/780 computer with a VMS operating system. To generate the random set of sequences, polio RNA was arbitrarily divided into four equal segments, while the rat 28S rRNA was subdivided into 16 segments beginning at positions 1, 113, 276, 329, 1228, 1280, 1301, 1964, 2119, 2291, 2620, 2720, 3329, 3801, 3889 and 4450.

The SEQH program of Goad and Kanehisa (20) was used with default criteria of 3 matches/mismatch or deletion-insertion. The RNAFOLD program employed was the original version described by Zuker and Stiegler (21), modified to include the base-pairing energy rules which have been recently published by Jacobson *et al.* (22). An additional modification of this program was included to allow the two blocks of sequence to fold as if part of the same molecule. This was accomplished by linking the two blocks with a 13 base-long stretch of residues (GGGGAAAAACCCC) which is forced to base-pair its G and C residues regardless of the overall secondary structure. The use of these modified programs, and statistical analysis of the randomized and actual polio RNA and 28S rRNA

sequences constitute the computer-simulated hybridization methodology we have developed.

RESULTS

Poliovirus Genome RNA Hybridizes Specifically to Higher Eukaryotic rRNAs

Since we initially discovered this phenomenon with RNAs extracted from HeLa cell cytoplasm, we explored whether cytoplasmic RNAs from a variety of sources could also hybridize with our polio RNA probe using stringent hybridization conditions (see Materials and Methods). Fragmented, end-labeled poliovirus genomic RNA was used as a probe for hybridizations to equal amounts of glyoxal-denatured RNAs fractionated on agarose gels, and electrophoretically transferred to nitrocellulose paper (see Materials and Methods). Figure 1A demonstrates that this hybridization was much stronger for 28S and 18S rRNA of higher eukaryotes, represented by mouse L cell cytoplasmic extracts (lane a), as compared to purified yeast rRNAs (lane b), yeast cytoplasmic extracts (lane c), gel purified B. stearothermophilus rRNAs (lane d) or total cell RNAs of E. coli (not shown), or VSV genomic RNA (lane e). All the latter signals were ~50-100 fold lower (estimated by densitometric scanning), and we assume for the purposes of this study that this low level of hybridization represents background. In contrast to the above, when similarly prepared HeLa cytoplasmic RNA was used as a probe, roughly equal hybridization signals were observed with all sources of rRNA (not shown). This self-hybridization was expected since these molecules display a high degree of secondary structure and extensive conservation of sequences between prokaryotes and eukaryotes. These results indicate that contamination of our polio RNA probe with HeLa rRNA did not account for the species specificity observed.

We next tested whether poliovirus RNA hybridized to a number of different sources of higher eukaryotic rRNAs. Cytoplasmic extracts of human (HeLa), hamster (BHK), mouse (L), chicken (embryo fibroblasts), and wheat germ (S30 fraction), containing an estimated 80-90% rRNAs were assayed as outlined above. As shown in figure 1B, the signal for the rRNA of the larger subunit was ~3-4 fold higher than the 18S signal for all higher eukaryotic species examined, except HeLa extract where the 28S signal was ~10 fold higher than 18S, and wheat germ where no significant 18S signal was detectable. No hybridization signal was observed in any other region of these gels including the 4-5.8S size range (the presence or absence of soluble yeast RNA in the hybridization buffer did not affect this result), even though electrophoretic transfer and retention of these molecules on nitrocellulose under our

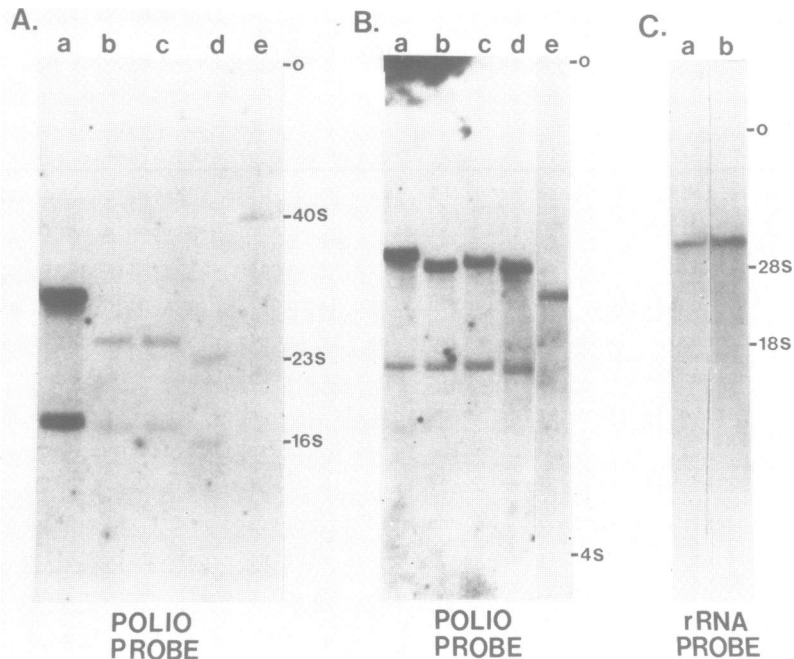


FIGURE 1. Hybridization between picornavirus genomes and higher eukaryotic rRNAs. All RNAs (~200 ng) were glyoxal-denatured, electrophoresed on a 1.1% agarose gel, electrophoretically transferred to nitrocellulose paper, and probed with poliovirus RNA (parts A and B), or HeLa cytoplasmic RNAs (part C) as described in Materials and Methods. Part A: lane a, mouse L cell cytoplasmic RNA; lane b, purified yeast 26S and 18S rRNA; lane c, yeast cytoplasmic RNAs; lane d, purified *B. stearothermophilus* 23S and 16S rRNA; lane e, VSV genomic RNA (40S). Part B: cytoplasmic RNAs of HeLa (lane a), BHK (lane b), L cells (lane c), chicken embryo fibroblasts (lane d), and wheat germ S-30 fraction (lane e). The position of 4S size tRNA is indicated. Part C: lane a, poliovirus RNA; lane b, Mengovirus RNA. Positions of HeLa rRNA markers run on a parallel lane are indicated.

conditions was as efficient as that of the two larger rRNAs (see Materials and Methods). Commercially available preparations of purified calf liver 28S and 18S rRNAs were also positive when probed as in figure 1, and identical results were also obtained using uniformly labeled ^{32}P -poliovirus RNA as the probe (not shown).

Figure 1C, shows that, using HeLa cytoplasmic RNA as the probe, we also obtained specific hybridization to full-size genome RNA of poliovirus (lane a), Mengovirus (lane b), and Coxsackie virus (not shown). This phenomenon therefore extends to other picornavirus genomes. Furthermore, contaminating rRNAs are clearly not found in these viral RNA preparations.

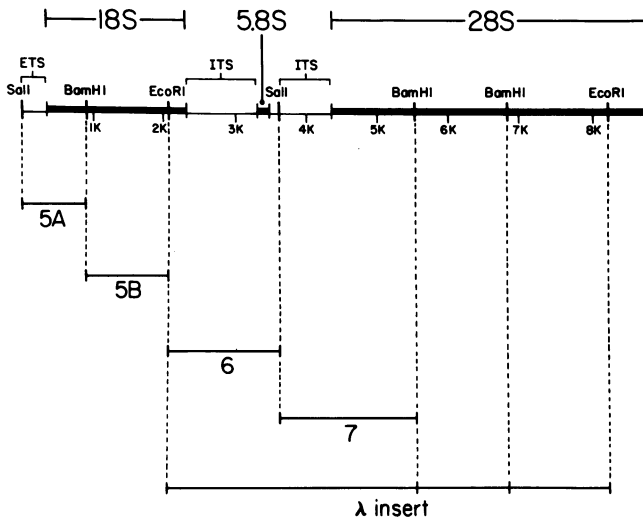


FIGURE 2. Map of the transcribed region from the mouse rDNA genes and cloned inserts. This map shows the positions of the clones of the mouse rRNA genes we have used for probing with poliovirus RNA in figure 4. The first four clones are pBR322-derived plasmids, while the fifth clone is a λ vector.

Mapping of rRNA Regions Hybridizing to Poliovirus RNA by Southern Blot Analysis of Cloned Mouse rDNA Genes

We initially examined four plasmid clones of mouse rDNAs, three of which together contained all of 18S, 5.8S and accompanying external transcribed spacer (ETS) and internal transcribed spacer (ITS) region, while the fourth contained 1.24 kb of the 5' end of 28S rRNA and an upstream accompanying ITS (see Fig. 2). Southern blot analysis of the purified plasmids under stringent hybridization and washing conditions (see Materials and Methods), showed that clone 7 (5' end region of 28S rRNA) and clone 5B (internal portion of 18S rRNA) gave the strongest signals when probed with polio RNA (Fig. 3). The signal in clone 5B was ~4 fold lower than that of clone 7 for roughly equivalent amounts of insert DNA on a molar basis (Fig. 3A, lane a vs b) while clones 5A and 6 (Fig. 3B) resulted in lower signals than 5B. Quantitation of these signals (correcting for the amounts of inserts) as well as those from a separate analysis in which all four clone inserts were present in approximately equal amounts (not shown) revealed that both clones 5A and 6 gave ~10 fold lower strength signals than clone 5B. The multiple bands seen in each lane of figure 3 represent the various conformations of the purified plasmids which were not linearized before analysis. We conclude from this

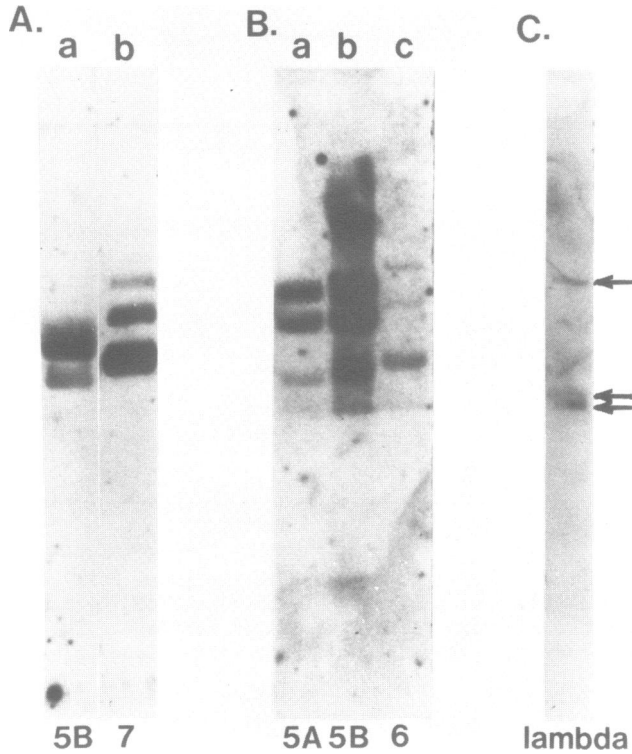


FIGURE 3: Southern blot analysis of mouse rDNA gene sequences probed with poliovirus RNA. Plasmid preparations, as well as the DNA (see Fig. 2), were electrophoresed on 1% agarose gels, electrophoretically transferred to nitrocellulose paper, and probed with labeled poliovirus RNA as in figure 1. Part A: lane a, clone 5B (~500 ng); lane b, clone 7 (~500 ng). Part B: lane a, clone 5A (~420 ng); lane b, clone 5B (~880 ng); lane c, clone 6 (~200 ng). The fastest migrating band in each plasmid preparation represents the covalently closed circular forms. Part C: EcoRI/BamHI digested λ gtWES/Mr100 DNA (~2 μ g) probed with poliovirus RNA as above. The positions of the three BamHI fragments within the EcoRI fragment insert are indicated by arrows.

experiment that there is at least one region of relatively weak complementarity to polio RNA present in the body of 18S rRNA while one or more stronger signal(s) is situated within the first 1.24 kb of 28S rRNA and/or upstream in the ITS. Although clones 5A and 6 also gave some very weak signals, we considered these to be too close to background level for further analysis.

To examine whether other regions in 28S rRNA might be involved, we also probed a λ vector containing an insert beginning at the EcoRI site at position 1644 from the 5' end of 18S and ending at the EcoRI site 583 bases from the 3'

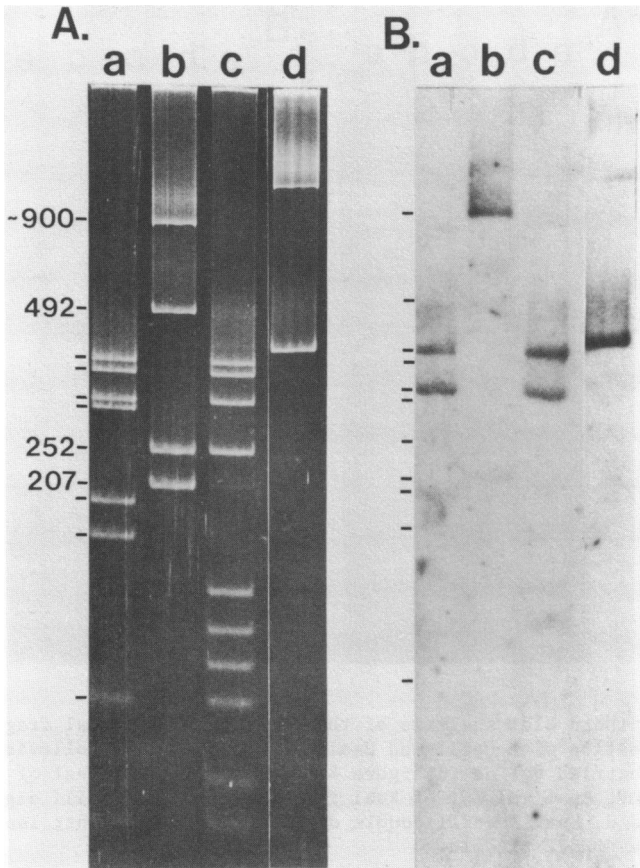


FIGURE 4: Southern blot analysis of the purified rDNA clone 7 insert following restriction enzyme digestion and probed with poliovirus RNA. Part A: lanes a to d show the ethidium bromide stained gel of the clone 7 insert (~2 μ g) after digestion with *Hinfl*, *RsaI*, *RsaI/Hinfl* and *AluI*, respectively. Part B: Southern blot analysis of the same gel probed with poliovirus RNA as described in the text. The size of the various fragments (bp) was estimated from other gels containing parallel lanes of *Hinfl* and *HaeIII* digested pBR322 as markers.

end of 28S (see Fig. 2). The gel purified insert was cut into three pieces with *BamHI* (12). Southern blot analysis of this digest revealed that all three insert fragments could be specifically probed with poliovirus RNA while none of the λ genome fragments reacted (Fig. 3C). Additional experiments showed that the signal from the smallest *BamHI* fragment was at least equal to that of the largest fragment which includes the sequences present in the clone 7 plasmid. The signal in the remaining *BamHI* fragment was at least 10 fold lower than that in the smallest band. Since we could not obtain a plasmid

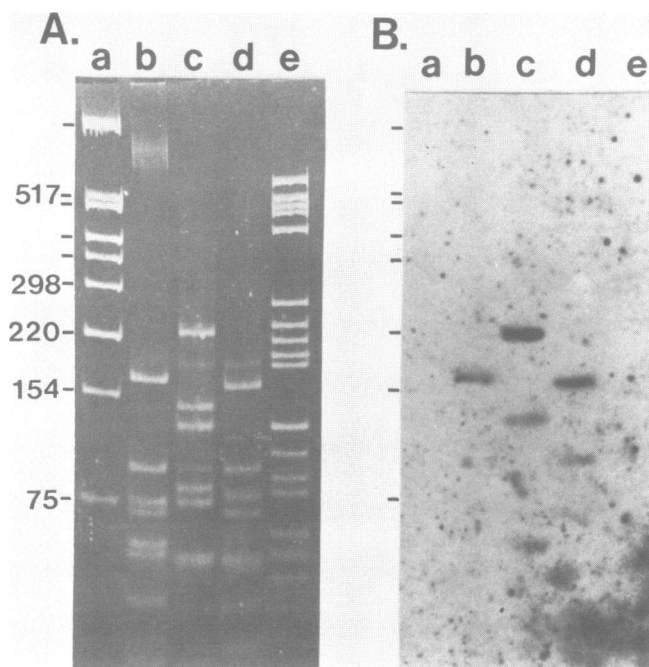


FIGURE 5: Southern blot analysis of the purified 875 bp RsaI fragment following digestion with HpaII and HaeIII and probed with poliovirus RNA. Analysis was carried out as in figure 4. Lane a, HindIII digest of pure pBR322 plasmid; lane b, HpaII digest of RsaI fragment; lane c, HaeIII digest of RsaI fragment; lane d, HpaII/HaeIII double digest of RsaI fragment; lane e, HaeIII digest of pure pBR322 plasmid.

clone containing the remaining terminal 3' end sequence of mouse 28S rRNA, we examined a clone containing the analogous human 28S rRNA region (~580 bp). This also gave a signal at least 10 fold lower than clone 7 (data not shown). We conclude from this that at least two or more sites within the λ insert can base-pair with poliovirus RNA. In addition at least one more site, albeit weaker, is also found in the 3' terminal region of 28S rRNA.

To enable us to assess the significance of these hybridization interactions we chose to precisely map at least one of the regions in the 28S rRNA which resulted in a strong signal. Accordingly, the insert from plasmid clone 7 was gel purified and used for subsequent restriction analysis. The restriction patterns of this fragment after digestion with HindIII, RsaI, HindIII/RsaI and AluI (Fig. 4A, lanes a-d), and the Southern blot analysis with the polio RNA probe (Fig. 4B, lanes a-d) show that the signal was entirely contained within the single 875 bp RsaI fragment (lane b) or ~450 bp AluI

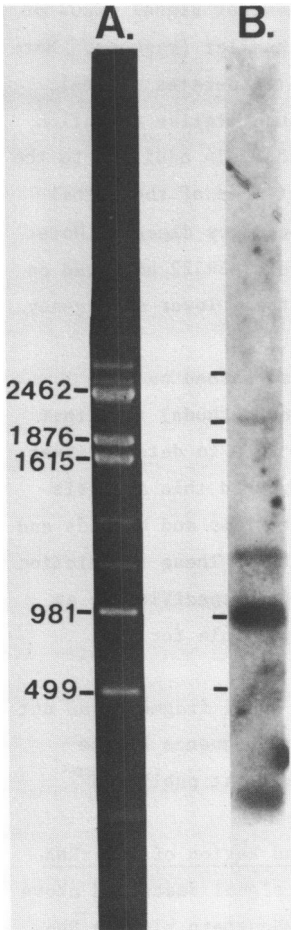


FIGURE 6: Southern blot analysis of the infectious cDNA of the poliovirus genome following restriction enzyme digestion and probed with the purified Sall/BamHI fragment from clone 7. Part A: ethidium bromide stained gel of the polio cDNA insert (~1.5 μ g) after digestion with HaeII. The polio specific fragments and sizes are indicated. The pale bands are due to contaminating pBR322. Part B: Southern blot analysis of the same gel probed with the Sall/BamHI fragment as described in the text. The minor signals present in part B are due to very small amounts of contaminating pBR322 sequences present in probe.

fragment (lane d) while HinfI split the signal into two fragments (375 bp and 312 bp) of roughly equal intensity (lanes a and c). The published mouse 28S rRNA sequence (23) places the RsaI fragment at position 363-1237 and the two HinfI fragments at positions 464-838 and 839-1149. Although no AluI site is predicted from this published sequence our restriction analysis using the end-labeled insert maps the AluI site ~450 bp from the BamHI site (position ~787-1237, data not shown).

To further delineate the region(s) within the 875 bp RsaI fragment from clone 7 which reacted with polio, we gel purified this fragment, restricted it with HpaII, HaeIII and HaeIII/HpaII, and again probed with polio RNA. The results of this analysis (Fig. 5A and B) show that most of the signal was found in a 174 bp HpaII fragment (position 1048-1221, lanes b and d) and a 230

bp HaeIII fragment (position 984-1213, lane c), while a minor signal (<20% of the total signal in each case), was also observed in a smaller fragment. Note that HpaII and/or HaeIII digestion of this RsaI fragment generates several small fragments < 40 bp, which is the lower limit for quantitative retention of fragments to the nitrocellulose paper using this method, in addition to the several larger fragments. It is therefore possible that some of the signal contained in the RsaI fragment is destroyed by these secondary digests. Note also in figure 5 that the HinfI and HaeIII digests of pure pBR322 analyzed on the same gel did not hybridize to polio RNA even under these lower stringency conditions (lanes a and e).

The Southern blots of figures 4 and 5 were probed and washed under considerably less stringent conditions (see Materials and Methods) than that of figure 3 because we reasoned that we might be better able to detect weaker signals spread over more than one fragment. We have repeated this analysis under the more stringent conditions described in the Materials and Methods and detected no change in the signals. Clearly, the majority of these restriction fragments do not react with the polio RNA probe, indicating specificity, as well as ruling out any contaminating rRNAs as being responsible for the phenomenon.

Since the sequence of 28S rRNA corresponding to the HpaII fragment was not known at the time of these studies we also determined its sequence by the Maxam-Gilbert method. Our mouse sequence was identical to that published subsequently by Hassouna *et al.* (23).

Mapping of Poliovirus Genome Region Hybridizing to 5' End Region of 28S rRNA

To demonstrate the specificity of the hybridization signal described above with respect to the poliovirus sequence, we analysed by Southern blots a Hae II restriction digest of a gel-purified DNA insert representing the complete poliovirus genome (see Materials and Methods). Figure 6 clearly indicates that only one of the five Hae II fragments representing the polio RNA genome can be probed by the nick-translated Sall/Bam HI insert fragment from the clone 7 plasmid described above. The 982bp fragment which gives the signal, spans position 4837-5818 of the polio genome.

Ruling Out Non-Specific G/C Rich Interactions

The species specificity of rRNAs reacting with poliovirus RNA, as well as the specificity of restriction fragment probing shown above, suggested to us that the phenomenon was not simply due to an artifact. It is well known, however, that the guanine and cytosine (G/C) content of rRNAs, especially 28S (24,25), increases as a function of evolutionary complexity. We therefore

considered the possibility that the signal could reflect non-specific base-pairing with these high G/C content regions. Such a phenomenon has recently been well documented by Rasmussen *et al.* (26) who have demonstrated that the human cytomegalovirus genome can hybridize to the avian retrovirus oncogene V-MYC sequence in regions which are >90% G/C, ranging in size from 15 bp to 40 bp.

Several lines of evidence, however, argue very strongly against this possibility for the phenomenon we report here. First, a closer look at the experiment of figure 3 reveals that clone 6 plasmid which gave the lowest or background signal with polio RNA, contains 998 bp of an ITS (3' side of 18S sequence) which has an average G/C content of 70%, including two stretches of 48 and 49 bases with a 90% G/C content (27). Second, the sequence represented by the three smallest RsaI fragments from the clone 7 plasmid insert (Fig. 4), which also includes a 70% G/C rich ITS (5' side of 28S sequence), contains several stretches >27 bp long with G/C content >90% (27,28). Even under the lower stringency hybridization and washing conditions of figure 4 (see Materials and Methods) none of these G/C rich regions reacted with poliovirus RNA. Third, we attempted hybridization of the poliovirus probe to homopolymers of G and C (6-13S in size) dotted onto nitrocellulose. Results showed that ~50 ng of rRNA gave a 10-fold stronger signal than ~10 ng of either homopolymer (M.A. McClure, Ph.D. thesis, Washington University, St. Louis, 1984). Lastly, using hybridization conditions similar to those of Jones and Hyman (29) who were able to compete out a presumed nonspecific G/C rich interaction between the herpesvirus genome and cellular DNA with poly G and poly (G,U), we were not able to reduce the signal between the polio RNA probe and the largest RsaI fragment derived from the Sall/BamHI insert even in the presence of 1,000 fold excess of these unlabelled polynucleotides. Furthermore, several different restriction digests (HinfI, HaeIII, AvaII and XmaIII) of the Sall/Bam HI insert fragment from clone 7 were also tested with a 100-100,000 fold excess of poly G and poly (G,U) and we were still unable to detect any loss of the polio signal (data not shown).

Confirmation of Base-Pairing Potential Between the Poliovirus Genome and Rat 28S rRNA by Computer Simulated Hybridization Analysis

We compared the complement of the entire sequence of the poliovirus genome to that of rat 28S rRNA (mouse sequences were not available at the time of this analysis) by the SEQH homology program of Goad and Kanehisa (20). From the 662 possible alignments generated ranging from 24 bp to 9 bp, we next considered only those involving at least 16 bp (including G:U pairing). These

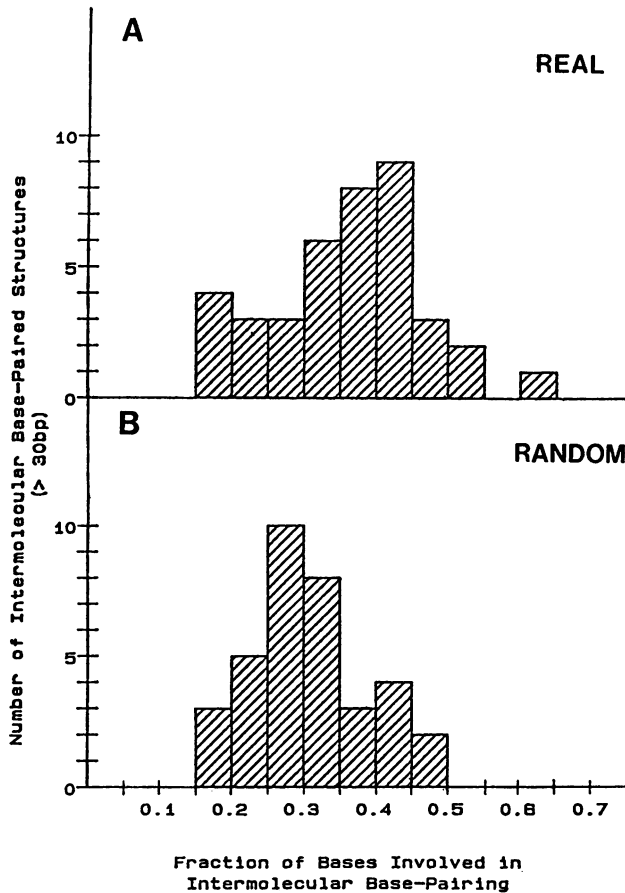


FIGURE 7: Distribution of all possible hybrid structures as a function of the fraction of bases involved in intermolecular base-pairing (minimum of 30 bp). Each possible hybrid structure is represented as the ratio of bases involved in intermolecular base-pairing (minimum of 30 bp)/total number of bases in the structure whose borders are defined by the terminal intermolecular base pairs on each side. The upper histogram represents the possibilities generated by comparing the actual or real poliovirus sequence vs rat 28S rRNA, while the lower histogram depicts the hybrids generated by comparing randomly shuffled sequences analogous to polio RNA and rat 28S rRNA (see text).

64 sites, with an additional flanking 100 bases on either side on both strands, were then examined for their ability to base-pair intra and intermolecularly using a modified version of the Zuker and Stiegler RNAFOLD program (21) (see Methods). The folded structures generated thus reveal only those potential intermolecular base-pairing interactions which are predicted to be more stable than intramolecular folding within the same local regions.

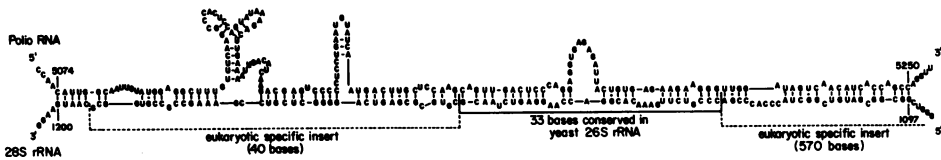


FIGURE 8: Predicted complementary alignment between bases 1097-1200 of 28S rRNA and 5074-5250 of poliovirus RNA. The structure shown is that predicted by the computer simulated hybridization methodology outlined in the text. The regions of the 28S rRNA sequence which correspond to higher eukaryotic specific inserts, as well as the 33 bases conserved in yeast 26S rRNA, are indicated. Intermolecular base-pairing (counting G:U base pairs) includes 81 out of 104 contiguous bases in the rRNA strand, and 81 out of 176 contiguous bases in the polio strand within the region whose boundaries are defined by the last intermolecular base pair of either side.

Of these possibilities, we arbitrarily chose 30 intermolecular base pairs as a minimum stability criterion. The frequency distribution of the resulting 39 structures plotted as a function of the fraction of base-pairs involved in intermolecular base-pairing, is shown in figure 7A.

To obtain a relative measure of potential base-pairing between poliovirus RNA and rat 28S rRNA expected on the basis of chance alone, we carried out the same analysis using randomly shuffled versions of both sequences. Note that both molecules were first subdivided into segments before shuffling as described in Methods. Our rationale for subdividing in this fashion was that, in contrast to poliovirus RNA, rat 28S rRNA shows wide variation of base composition as a function of domains, particularly with respect to the G/C rich eukaryotic specific inserts. Preserving local sequence bias in this fashion has been shown previously to provide a more stringent test of statistical significance than either preserving total base composition or nearest neighbor frequency (30). The resulting distribution is shown in figure 7B. Although the average intermolecular base-pairing potential appears to be significantly higher in the real sequence comparison vs the random (means = 0.367 ± 0.108 and 0.318 ± 0.083 , $p = 0.015$) the difference is not striking except for the one real structure with an intermolecular base-pairing ratio >0.6 . Most significantly, the rRNA and poliovirus sequences predicted to be involved in this structure map within the fragments which we identified above in our Southern blot analysis, i.e., the HpaII fragment from rRNA (position 1048-1221) and the poliovirus HaeII fragment (position 4837-5818). Whether the remaining sites of hybridization between 28S rRNA and polio RNA, which have not been precisely mapped, are also represented within the

structures with relatively high base-pairing ratios remains to be determined.

Attempts to extend the predicted level of base-pairing interaction for the one structure with a base-pairing ratio >0.6 by considering additional stretches of sequence on either side resulted in the structure shown in figure 8 ($\Delta G = -139.8$ kcal). Although the limitations inherent in any RNA folding program available to date do not allow us to firmly conclude that the structure shown does in fact reflect one of the base-pairing interactions seen between these two RNA molecules, the extent of intermolecular base-pairing in the rRNA strand (81 out of 104 contiguous bases) is nonetheless impressive. Note that the polio strand in this structure is only 46% GC while that of rRNA is 66%.

This computer methodology also suggests that an additional base-pairing region between position 780-895 of 28S rRNA and position 389-525 of polio RNA, and/or 744-878 of 28S rRNA and 3283-3363 of polio RNA (53 and 63 intermolecular bp respectively) might be responsible for the positive signals obtained with the two *Hin*I fragments in figure 4, lanes a and c. Although the ΔG 's for these structures (not shown) are very similar to that of the structure shown in figure 8, we have nonetheless been unable to detect hybridization of these regions to the *Hae*II digest of the polio insert (contained within the two largest polio fragments) in figure 6. The lack of hybridization to these polio regions may be due to preferential long distance intramolecular interactions. This is consistent with preliminary computer analysis using a newer version of the folding program which allows 1000 bases as input and which results in loss of the intermolecular structures around the *Hin*I sites on the rRNA molecule but essentially retains the structure of figure 8.

DISCUSSION

We have shown here that poliovirus genomic RNA can hybridize specifically to all 28S and most 18S rRNAs of higher eukaryotes examined but not to that of yeast or the two eubacterial species tested. The hybridization cannot be due merely to non-specific interactions with regions of high G/C content since such regions present in plasmid clones of rRNA genes or pure G and C homopolymers did not react with the polio probe. In addition, we have shown that labeled polio RNA does not hybridize significantly to other abundant cytoplasmic RNAs, such as tRNA or 5-5.8S RNA, or VSV genomic RNA, or with the DNA of lambda phage or pBR322 or, as shown elsewhere (M.A. McClure, Ph.D. thesis, Washington University, 1984) that of adenovirus type 2. These results clearly attest to the specificity of this novel hybridization phenomenon.

Based on our Southern blot analysis of DNA clones corresponding to all of the rRNA transcription unit, we conclude that there are several regions of hybridization between polio RNA and 28S rRNA, and at least one in 18S rRNA. We have identified one of these regions in 28S rRNA and poliovirus RNA by restriction analysis of the clones and find that a fragment corresponding to position 1097-1200 in the ribosomal gene hybridizes to a fragment corresponding to position 5075-5220 in polio. The computer-simulated hybridization methodology (see Methods) used to compare the complete sequences of the poliovirus genome and 28S rRNA revealed that the most significant potential base-pairing interaction between these molecules did indeed correspond to the rRNA region and polio region mapped by Southern blotting. The predicted structure (Fig. 8) is best described as a site of "patchy complementarity" with 81 out of 104 contiguous bases of the rRNA strand base-paired to polio, and is consistent with the lack of hybridization with yeast 26S rRNA or prokaryotic 23S rRNA. The rRNA sequence of the predicted hybrid begins in one, and spans another, of the higher eukaryote specific insert regions, while only the bases at position 1123-1155 are found in the yeast 26S rRNA (position 634-666 (31)). Furthermore, by substituting the yeast 26S rRNA sequence for that of rat in our computer analysis, we were unable to form any significant hybrid with position 5075-5250 of polio RNA.

The biological significance of base-pairing interactions between poliovirus RNA and rRNAs is difficult to assess and will require further study, but this unexpected phenomenon raises interesting questions. We have found that a number of other RNA viruses are also able to hybridize with rRNA (manuscript submitted) but so far, this does not seem to be the case for DNA viruses (see above). Therefore, this phenomenon is not limited to picornaviruses but may well be specific to RNA viruses. Whether some cellular mRNAs might also be capable of this type of interaction is also an open question since their low abundance in the cytoplasmic extracts examined in our study here may have precluded our detecting them. However, a large number of cellular gene sequences have been used as probes in numerous studies for several years and no such phenomenon has yet been reported. If we assume for now that the phenomenon is more or less specific to RNA viruses, one possible explanation is that it represents a remnant of evolutionary relatedness between higher eukaryotic rRNAs and RNA virus genomes. It is well known that rRNAs are considered to be among the earliest nucleic acids in evolution (32,33) and that the origin of RNA viruses, while a subject of past speculation (34-36), is unknown. The data presented here would suggest that such a relationship, at least as far as polio is concerned, would have to be

with those elements that represent "higher" eukaryotic specific domains in rRNAs. Interestingly, the genomes of the RNA bacteriophages, Q β and MS2, have been shown to contain short sequences homologous to 16S rRNA (37). The authors of that study have also suggested a possible common evolutionary ancestry, perhaps reflecting conserved sequence elements which can bind ribosomal proteins. The involvement of the ribosomal S1 protein, and the translation elongation factors EF-Tu and EF-Ts as essential subunits of the Q β replicase is consistent with this possibility, and raises the question as to whether a similar situation may hold for eukaryotic RNA viruses.

Whether common ancestry is involved or not, RNA virus genomes are known to evolve rapidly as compared to DNA genomes (38). Hybridization to rRNA may thus reflect an important conserved function. The likely candidate function would be a role in translation where the viral mRNAs may have the potential to base-pair with exposed regions of rRNA on the ribosomes. The function of the eukaryotic specific inserts in rRNAs or their accessibility on the ribosomal complex is not yet known (Noller, personal communication). It is worth noting, however, that it is the 28S rRNA, and not the 18S species, which hybridizes more strongly to polio RNA. Most eukaryotic studies to date have indicated that it is the 40S ribosomal subunit which binds to cellular mRNAs and recognizes the AUG codon (4). We find it intriguing, however, that the region of complementarity to 28S rRNA in the polio genome shown in figure 8 (a structure utilizing at least some of the bases between positions 5075-5250), not only includes within it an AUG codon which could code for the P3/1b precursor, but is also positioned just upstream from the four possible AUG start codons (positions 5290, 5346, 5466 and 5517) that could be responsible for the internally initiated proteins described by Dorner *et al.* (39). Although internal initiation on polio mRNA could only be detected in reticulocyte lysates in this study and not in HeLa cell extracts, *in vivo* studies of Koch *et al.* (40) are consistent with such internal initiation early in infection or following hypertonic shock in HeLa cells. This putative internal initiation region also corresponds to one of the three poliovirus genome ribosome binding sites (position 5300) in the studies of McClain *et al.* (41). The foregoing suggests that if internal initiation does occur in poliovirus, albeit less frequently than external 5' end initiation, the mechanism may be quite different and perhaps utilize features of 28S rRNA not involved in 5' end initiations. It has also been proposed from studies with prokaryotes, which differ from eukaryotes in that a large fraction of translation initiation occur internally, that the 70S ribosome complex is responsible for recognition of internal initiation sites as opposed to the 30S

complex which binds to the external (5' end) site (42).

Another explanation for the hybridization phenomenon described in this study might be that poliovirus and perhaps other RNA viruses have "adapted" to the eukaryotic rRNAs through some form of recombination. RNA recombination in picornaviruses has been convincingly demonstrated (43). We should also consider the possibility that, regardless of exact primary sequence or biological function, some large RNA molecules may be capable of folding into "complementary" three-dimensional domains which favor intermolecular interactions just as in the case of protein-protein interactions. Formation of these complexes may be precluded in the cellular environment because of protein-RNA associations or unfavorable conditions for base-pairing.

Lastly, but importantly, a more immediate practical concern raised by our studies is that the use of nucleic acid probes for detecting the presence of RNA virus genomes in diseased tissues could be misleading, unless care is taken to use only those sequences which are non-crossreactive with rRNAs.

ACKNOWLEDGEMENTS

We thank Drs. Ira Wool, Harry Noller, and Michael Lieberman, for kindly providing us with preprints of their work and Dr. Bert Semler for a gift of the poliovirus infectious clone. We also thank Drs. Harry Noller, David Kohne, John Holland and Elvera Ehrenfeld for helpful discussions, and Dr. M. Zuker and the National Research Council of Canada for providing us with the RNAFOLD program. This work was supported by grants AI-14365 and AI-21572 from the National Institutes of Health and PCM-8303955 from the National Science Foundation to J.P. who is also a recipient of a Research Career Development Award from the NIH.

*To whom reprint requests should be addressed.

[†]Present address: Biology Department, San Diego State University, San Diego, CA 92182.

REFERENCES

1. Perez-Bercoff, R. (1982) in *Protein Biosynthesis in Eukaryotes*, Perez-Bercoff, R. Ed., pp. 245-251, Plenum Press, New York.
2. Ambros, V. Pettersson, R.F., and Baltimore, D. (1978) *Cell* 15, 1439-1446.
3. Pallansch, M.A., Kew, O.H., Semler, B.L., Omilianowski, D.R., Anderson, C. W. Wimmer, E., and Rueckert, R.R. (1984) *J. Virol.* 49, 873-880.
4. Kozak, M. (1983) *Microbiol. Rev.* 47, 1-45.
5. Kitamura, N., Semler, B.L., Rotherberg, P.G., Larsen, G.R., Adler, C.J., Dorner, A.J., Emini, E.A., Hanecak, R., Lee, J.J., van der Werf, S., Anderson, C.W., and Wimmer, E. (1981) *Nature* 291, 547-553.
6. Racaniello, V.R., and Baltimore, D. (1981) *Proc. Natl. Acad. Sci. USA* 78, 4887-4891.
7. McClure, M.A., Holland, J.J., and Perrault, J. (1980) *Virology* 100, 408-418.

8. Perrault, J., Clinton, G.M., and McClure, M.A. (1983) *Cell* 35, 175-185.
9. Bowman, L.H., Rabin, B., and Schlessinger, D. (1981) *Nucleic Acids Res.* 9, 4951-4966.
10. Birnboim, H.C., and Doly, J. (1979) *Nucleic Acids Res.* 7, 1513-1522.
11. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning a Laboratory Manual*, Cold Spring Harbor Laboratory, New York.
12. Tiemeier, D.C., Tilghman, S.M., and Leder, P. (1977) *Gene* 2, 131-191.
13. Semler, B.L., Dorner, A.J., and Wimmer, E. (1984) *Nucleic Acids Res.* 12, 5123-5141.
14. Perrault, J., and Semler, B.L. (1979) *Proc. Natl. Acad. Sci. USA* 76, 6191-6195.
15. Maxam, A.J., and Gilbert, W. (1980) *Methods in Enzymol.* 65, 499-560.
16. Smith, M.R., Devine, C.S., Cohn, S.M., and Lieberman, M.W. (1984) *Anal. Biochem.* 137, 120-124.
17. Meinkoth, J., and Wahl, G. (1984) *Anal. Biochem.* 138, 267-284.
18. Casey, J., and Davison, N. (1977) *Nucleic Acids Res* 4, 1539-1552.
19. Schildkraut, C., and Lifson, S. (1965) *Biopolymers* 3, 195-208.
20. Goad, W.B., and Kanehisa, M.I. (1982) *Nucleic Acid Res.* 10, 247-263.
21. Zuker, M., and Stiegler, P. (1981) *Nucleic Acid Res.* 9, 133-148.
22. Jacobson, A., Good, L., Simonetti, J., and Zuker, M. (1984) *Nucleic Acid Res.* 12, 45-52.
23. Hassouna, N., Michot, B. and Bachellerie, J-P., (1984) *Nucleic Acid Res.* 12, 3563-3583.
24. Choi, U.C., and Ro-Choi, T. (1980) in *Cell Biology: A Comprehensive Treatise*, Goldstein, L., and Prescott, D.M. Eds., Vol 3, pp. 609-667, Academic Press, New York.
25. Chan, Y., Olvera, J., and Wool, I.G. (1983) *Nucleic Acids Res.* 11, 7819-7831.
26. Rasmussen, R.D., Staprans, S.I., Shaw, S.B., and Spector, D.H. (1985) *Mol. Cell Biol.* 5, 1525-1530.
27. Goldman, W.E., Goldberg, G. Bowman, L.H., Steinmetz, D., and Schlessinger, D. (1983) *Mol. Cell Biol.* 3, 1488-1500.
28. Michot, B., Bachellerie, J., and Raynal, F. (1982) *Nucleic Acids Res.* 10, 5273-5283.
29. Jones, T.R. and Hyman, R.W. (1983) *Virology* 131, 555-560.
30. Lipman, D.J., Wilbur, W.J., Smith, T.F., and Waterman, M.S. (1984) *Nucleic Acids Res.* 12, 215-226.
31. Rabtsov, P.M., Musakhanov, M.M., Zakharyev, V.M., Krayev, A. S., Skryabin, K.G., and Bayev, A.A. (1980) *Nucleic Acids Res.* 8, 5779-5794.
32. Woese, C.R. (1981) *Sci. Amer.* 244, 98-122.
33. Crothers, D.M. (1982) *J. Mol. Biol.* 162, 379-391.
34. Baltimore, D. (1980) *Ann. N.Y. Acad. Sci.* 354, 491-497.
35. Reanney, D.C. (1982) *Ann. Rev. Microbiol.* 36, 47-73.
36. Zimmern, D. (1982) *Trends Biochem. Sci.* 7, 25-207.
37. Engelberg, H., and Schoulaker, R. (1976) *J. Mol. Biol* 106, 709-733.
38. Holland, J., Spendler, K., Horodyski, F., Grabau, E., Nichol, S., and Van de Pol, S. (1982) *Science* 215, 1577-1585.
39. Dorner, A.J., Semler, B.L., Jackson, R.J., Hanecak, R., Duprey, E., and Wimmer, E. (1984) *J. Virol.* 52, 719-721.
40. Koch, G., Koch, F., Billelo, J.A., Hiller, E., Scharli, C. Warnecke, G. and Weber, C. (1982) in *Protein Biosynthesis in Eukaryotes*, Perez-Bercoff, R. Ed., pp. 275-309, Plenum Press, New York.
41. McClain, K., Stewart, M., Sullivan, M., and Maizel, J.V. (1981) *Virology* 113, 150-167.
42. Peterson, H.U., Joseph, E., Ullmann, A., and Donchin, A. (1978) *J. Bact.* 135, 453-459.
43. King, A.M.Q., McCahon, D., Slade, W.R., and Newman, J.V.I. (1982) *Cell* 29, 921-928.