

1975

## James Darnell, Jr., 1973

The Rockefeller University

Follow this and additional works at: <https://digitalcommons.rockefeller.edu/harvey-lectures>

---

### Recommended Citation

The Rockefeller University, "James Darnell, Jr., 1973" (1975). *Harvey Society Lectures*. 45.  
<https://digitalcommons.rockefeller.edu/harvey-lectures/45>

This Book is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Harvey Society Lectures by an authorized administrator of Digital Commons @ RU. For more information, please contact [nilovao@rockefeller.edu](mailto:nilovao@rockefeller.edu).

# THE ORIGIN OF mRNA AND THE STRUCTURE OF THE MAMMALIAN CHROMOSOME\*

JAMES E. DARNELL, JR.

*The Rockefeller University, New York, New York*

## I. INTRODUCTION: EVENTS IN THE DEVELOPMENT OF RNA BIOCHEMISTRY

IT was less than 25 years ago that Erwin Chargaff suggested, on the basis of variability in the average base composition (Chargaff, 1950), that DNA could not be a simple arrangement of tetranucleotides as had been proposed by P. A. Levene (Levene and Bass, 1931). These early chemical composition studies together with the still earlier demonstration of bacterial transformation from Avery's laboratory (Avery *et al.*, 1944) provided some legitimate confidence that DNA was involved in gene specificity even before the deluge of proof of the central biological role of DNA that followed the Watson-Crick discovery of the DNA structure (Watson and Crick, 1953). The respect accorded to DNA molecule as being profoundly complicated in information content yet simple in structure did not come as easily for RNA. In spite of the early discovery of RNA in plant viruses (Stanley, 1935; Bawden and Pirie, 1937) the role of RNA in virus transmission, and gene expression in general, remained obscure. The early studies mentioned above that showed the base composition of DNA to be variable, showed the total cell RNA, particularly animal cell RNA, to be constructed within narrower limits (Chargaff, 1955); in addition, there was also no evidence for base-pairing to suggest a regularity in structure.

Among the first to draw attention to a possible role for RNA in gene expression were the cytologists Brachet (1947, 1955) and Caspersson (1947), who observed that cells which contained a large amount of RNA produced a large amount of protein. The role carried out by RNA could not be accurately prophesied

\* Lecture delivered September 20, 1973.

at this early time because separation of RNA into classes had not been achieved. The first discrete function for cellular RNA molecules was proposed by Francis Crick (1958), who reasoned that an "adaptor" molecule, likely a small RNA molecule, would be required to decode information stored in DNA for use in the selection of the correct amino acid by the cellular protein-synthesizing machinery. This prediction was accompanied almost at the same time by the discovery in Zamecnik's laboratory of the first discrete class of cellular RNA molecules, "soluble" RNA (Hoagland *et al.*, 1958; Zamecnik, 1960), now known as transfer or tRNA, the approximately 70–80 nucleotide-long carriers of activated amino acids (Holley *et al.*, 1965). Perhaps the most important work in highlighting the crucial direct role of RNA in gene expression was that of Fraenkel-Conrat and Williams (1955), Fraenkel-Conrat *et al.* (1957), Gierer and Schramm (1956), and Mundry (1959) with tobacco mosaic virus. That work, of course, demonstrated the infectious nature of the whole TMV-RNA molecule and the determination by the RNA, not by the protein, of genetic specificity of the virus.

In addition to the intellectual legacy owed by all of modern biology to the early TMV workers, biochemical technology owes them an equally important debt for methods of preparation of intact high molecular weight RNA molecules. Especially important in this regard was the technique of phenol extraction which denatures proteins, leaving RNA free in aqueous solution (Gierer and Schramm, 1956).

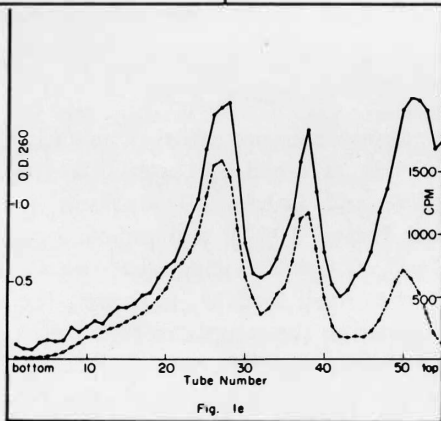
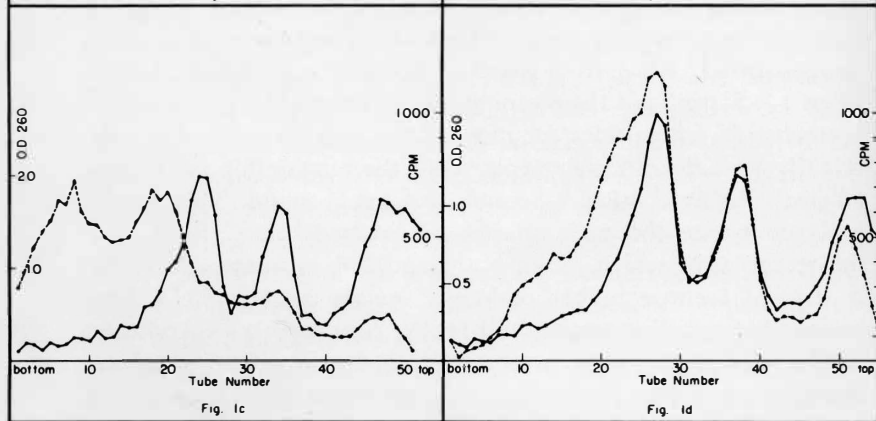
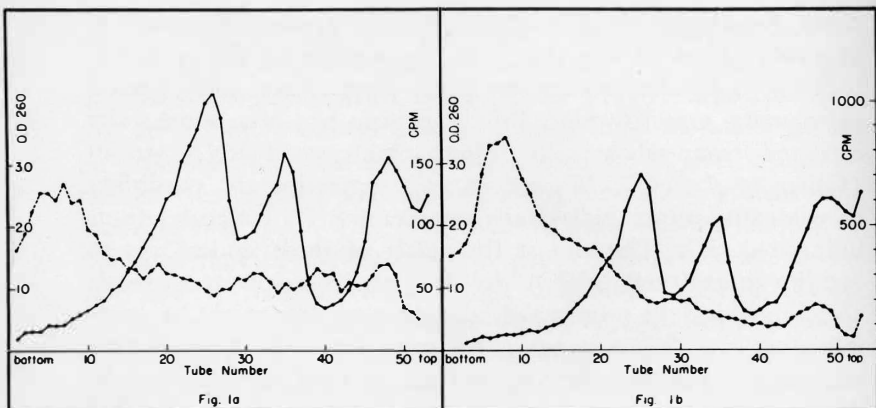
Thus, in the late 1950s when genetic studies with bacteria suggested the existence of an unstable intermediate acting between the genes and the protein-synthesizing system (Pardee *et al.*, 1959; Jacob and Monod, 1961), biochemical studies on extracted, protein-free RNA soon confirmed the theory of information transfer via the DNA  $\rightarrow$  RNA  $\rightarrow$  protein route. A small fraction of the total bacterial cell RNA, called messenger RNA (mRNA) was identified in association with ribosomes, the ubiquitous ribonucleoprotein particles which contained two regular-sized, presumably structural RNA molecules (McQuillen *et al.*, 1959; Zamecnik, 1960; Tissieres *et al.*, 1959; Kurland, 1960). The mRNA was hypothesized to provide instructions (i.e., the mes-

sage) for alignment of amino acids into proteins (Brenner *et al.*, 1961; Gros *et al.*, 1961) on the surface of the ribosome.

By 1961 detailed studies on RNA from animal cells seemed appropriate, especially since infectious virus had been successfully extracted from whole cells infected with small RNA viruses (Colter *et al.*, 1957; Wecker, 1958), suggesting the possibility of extracting intact, biologically important RNA molecules from uninfected cells. The aim at the outset of these studies was to identify animal cell mRNA, so that regulation of its synthesis and fate within the protein synthesizing machinery might be determined as was being done in bacteria. The first sedimentation analysis of "pulse-labeled" RNA from cultured mammalian cells (cells labeled for a small fraction of a generation time) revealed a much larger average size (faster sedimentation rate) than "pulse-labeled" RNA from bacteria (Scherrer and Darnell, 1962) (Fig. 1). Sorting out the meaning of this large RNA has occupied a number of laboratories for most of the past 10 years, and only recently have definite statements about the relationship of "pulse-labeled" RNA to mRNA become reasonably sound. This article will summarize the early studies on "pulse-labeled" RNA and the recent work which indicates that mRNA in mammalian cells is derived from a higher molecular weight nuclear RNA precursor. A concluding section will project how further experiments with nuclear mRNA precursor molecules might inform us about arrangement within chromosomes of structural gene regions and possibly how regulation of genetic expression is achieved.

## II. IDENTIFICATION AND SEPARATION OF CLASSES OF NUCLEAR RNA

Very soon after radioisotopic precursors of nucleic acids became available, evidence was obtained by both cell fraction (Marshak, 1948; Marshak and Calvet, 1949; Elson and Chargaff, 1952; Hurlbert and Potter, 1952) and radioautography (Goldstein and Plaut, 1955; Zalokar, 1959) that the initial labeling of RNA occurred in the cell nucleus. However, the majority of the cell RNA was located in the cytoplasm (Brachet, 1955) which only became the predominant location of labeled RNA after longer label times. Many of these early studies concluded that



nuclear turnover, without transport to the cytoplasm, was the fate of most, if not all, nuclear RNA because of the extremely rapid nuclear labeling and inability to observe nuclear radioactivity shift to the cytoplasm when exogenous label was removed, and the fact that the average base composition of nuclear and cytoplasmic RNA was found to be not exactly the same (Hurlbert and Potter, 1952; Smellie *et al.*, 1953; Barnum *et al.*, 1953; Moldave and Heidelberger, 1954). Particularly insistent on the viewpoint that nuclear RNA did not exit to the cytoplasm were Harris and his colleagues, whose experiments in 1959–1963 were carried out in cultured cells whereas many of the earlier studies were in whole animals (Harris, 1959; Harris and Watts, 1962; Harris *et al.*, 1963). Their emphasis on turnover was so strong that the suggestion was made that cytoplasmic synthesis might in fact be the source of cytoplasmic RNA (Harris and La Cour, 1963).

Since these early studies were performed on bulk RNA from the nucleus and cytoplasm, the possibility remained for differential stability of *different classes* of RNA molecules or for cytoplasmic transfer of a *part of a nuclear* molecule with nuclear turnover of the rest. The separation of RNA molecules of different classes was a necessary prerequisite for a solution of the relationship between nuclear and cytoplasmic RNA. By 1961 the techniques for extracting whole RNA molecules and separating them by sedimentation into different size classes encouraged us to examine the RNA from HeLa cells, labeled for various periods from a few minutes up to 24 hours, in search of an RNA fraction that might be mRNA (Fig. 1).

---

FIG. 1. Original sedimentation pattern of rapidly labeled HeLa cell RNA. Sucrose gradient analysis of RNA from HeLa cells growing with a 24-hour doubling time. Cells were labeled by exposure to uridine- $^{14}\text{C}$  0.03–0.07 mM, 0.54  $\mu\text{Ci}/\mu\text{mole}$ , according to the following schedule: (a) 10.0  $\mu\text{Ci}$ , 250 ml cells, 5 minutes; (b) 5.0  $\mu\text{Ci}$ , 150 ml cells, 30 minutes; (c) 2.5  $\mu\text{Ci}$ , 100 ml cells, 60 minutes; (d) 1.0  $\mu\text{Ci}$ , 75 ml cells, 4 hours; (e) 0.5  $\mu\text{Ci}$ , 100 ml cells, 24 hours; +2.0  $\mu\text{M}$  uridine (total specific activity, 0.17  $\mu\text{Ci}/\mu\text{mole}$ ). The 5–10% sucrose gradients contained NaCl, 0.05 M;  $\text{MgCl}_2$ ,  $10^{-4}$  M;  $\text{CH}_3\text{COONa}$ , 0.01 M; pH 5.1. ●—●,  $\text{OD}_{260}$ ; ●---●, counts per minute. From Scherrer and Darnell (1962).

Sucrose gradient zonal sedimentation analysis showed that briefly labeled HeLa cell RNA exhibited a much more rapid sedimentation than ribosomal RNA (Scherrer and Darnell, 1962). In addition most of the rRNA was in the cytoplasm whereas more than 90% of the pulse-labeled RNA was in the cell nucleus (Scherrer *et al.*, 1963) (Fig. 1). Pulse-labeled molecules sedimenting from 30–100 S were apparent with predominant peaks at 45 S and 32 S (at first termed 35 S) compared to the 28 and 18 S rRNA (Philipson, 1961). The use of "S" values in this work is for the purpose of naming molecules not for the purpose of calculating accurate molecular weights. The "S" values were based on comparison with the sedimentation of 23 S or 16 S rRNA from *E. coli* (Kurland, 1960).

#### A. Proof of the Precursor Role of 45 S Pre-rRNA

In an effort to determine whether the pulse-labeled nuclear RNA was related to rRNA or might be mRNA, the base composition of the 45 S and 32 S RNA labeled with  $^{32}\text{PO}_4^{3-}$  (Darnell, 1962; Scherrer *et al.*, 1963) was determined. These two species of nuclear RNA had a high guanine plus cytosine content similar to ribosomal RNA, and when briefly labeled cells were prevented from synthesizing further RNA by treatment with actinomycin D, the 45 S and 32 S peaks disappeared and labeled 28 S and 18 S ribosomal RNA appeared (Scherrer and Darnell, 1963; Perry, 1962). It was therefore concluded that the nuclear 45 and 32 S RNAs were ribosomal precursor RNA (pre-rRNA). Many experiments have confirmed the precursor-product relationship between pre-rRNA and rRNA (Girard *et al.*, 1964; Penman, 1966; Greenberg and Penman, 1966; Zimmerman and Holler, 1967; Jeanteur *et al.*, 1968; Salim *et al.*, 1970; Brown and Weber, 1968). These later experiments all take advantage of the fact that the molecules concerned, i.e., the 45 S, 32 S, and 20 S precursors and the 28 S and 18 S final products, are (1) discrete in size, and (2) make up substantial amounts of the total cellular RNA. Therefore purification and chemical relatedness between precursors and products could be firmly established. The 45 S pre-rRNA is about 14,000 nucleotides long and contains one 28 S

unit (5100 nucleotides long) and one 18 S unit (2000 nucleotides); the remainder of the 45 S is apparently turned over in the cell nucleus. Recent electron microscopic examination of the secondary structure of the precursor and the ribosomal molecules before and after 3' exonuclease action has in fact demonstrated the various regions pictorially: the 28 S is at (or close to) the 5' end followed by a region that is turned over, the 18 S, and finally other regions that are not conserved (Wellauer and Dawid, 1973) (Fig. 2). A thorough recent review of mRNA formation has been provided by Maden (1971).

### B. The "DNA-like" RNA of Heterogeneous Size

The existence of a nuclear RNA fraction separate from pre-rRNA was indicated in several early studies. Sibatani and co-workers (1959, 1962) reported that the base composition of pulse-labeled intestinal and later thymus cell RNA was similar to the DNA (uracil substituted for thymine). Georgiev and co-workers (Georgiev and Mantieva, 1962; Georgiev *et al.*, 1963) also reported that phenol extraction of rat liver and ascites tumor cells at increasing temperature yielded first a fraction with a high G + C base composition, which they called rRNA, and then a "DNA-like" fraction, which they called messenger RNA. This

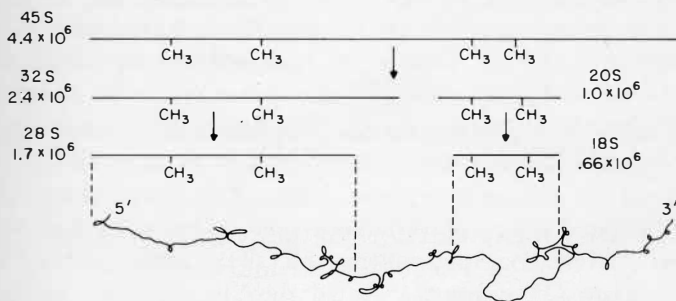


FIG. 2. The processing of 45 S pre-rRNA from HeLa cells. Top portion of diagram shows molecular weights of 45 S molecule, regions of addition of methyl groups, and cleavage products derived from it. Bottom portion is a tracing of an electron micrograph from Wellauer and Dawid (1973) of a 45 S molecule showing characteristic secondary structure pattern.



"DNA-like" base composition was a characteristic of bacterial messenger RNA, while bacterial rRNA differed widely from the average composition of the DNA. In the early experiments with HeLa cells in our laboratory, a fraction of briefly labeled nuclear RNA selected from the total by hybridization to cellular DNA was found to have a base composition different from pre-rRNA and like DNA (Scherrer *et al.*, 1963). Perhaps the first clear separations of pre-rRNA (high G + C) from undegraded "DNA-like" (low G + C) nuclear RNA was achieved by chromatography of the total labeled nuclear RNA on methylated albumin kieselguhr columns (Ellem and Sheridan, 1964; Yoshikawa-Fukada *et al.*, 1965; Kubinski and Koch, 1966). These studies called attention to the approximately equal labeling of this DNA-like RNA and pre-rRNA in "pulse" labels of growing cells. In addition duck erythroblasts which made no pre-rRNA did make a high molecular weight nuclear RNA with a DNA-like base composition (Scherrer and Marcaud, 1965). Reexamination of HeLa cell nuclear RNA by more careful sucrose gradient separation revealed that the RNA sedimenting faster than 45 S, which earlier had not been separately examined, did indeed have a "DNA-like" base composition (Soeiro *et al.*, 1966; Houssais and Attardi, 1966) (Fig. 3). The successful isolation of the nucleolus containing all the pre-rRNA (Penman *et al.*, 1966) provided, at the same time, a reliable means of obtaining the DNA-like RNA in radiochemically pure form in the extranucleolar fraction (Soeiro *et al.*, 1966). Because of its heterogeneous sedimentation pattern (20 S–100 S) and a lack of any clue as to what its function might be, this extranucleolar, DNA-like RNA was termed HnRNA or heterogeneous nuclear RNA.

### III. IDENTIFICATION OF mRNA: Is HnRNA A PRECURSOR TO mRNA?

Before the studies characterizing HnRNA were completed, the definition of mRNA had already been achieved. When polyribosomes were shown to be the site of protein synthesis in mammalian cells (Gierer, 1963; Noll *et al.*, 1963; Warner *et al.*, 1963), it was reasoned that they must contain the mRNA. Polyribosomes

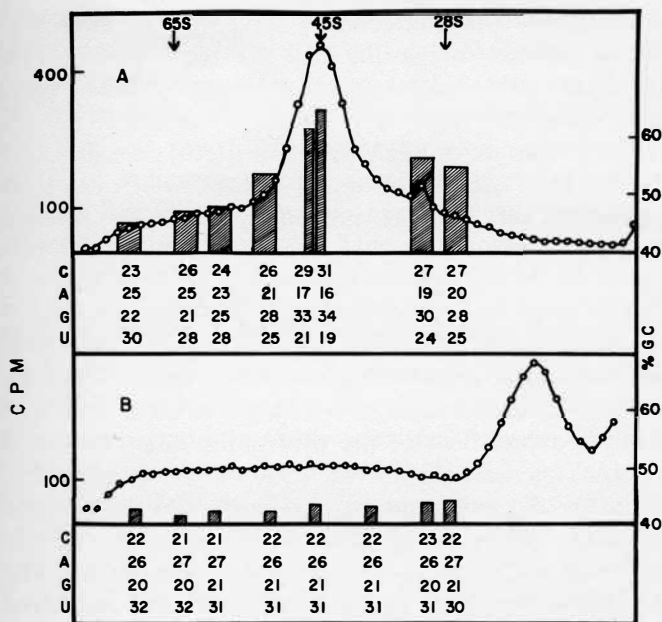


FIG. 3. Base composition of nuclear RNA fractions from HeLa cell.  $^{32}\text{P}$  labeled nuclear RNA was extracted from the nucleolar (A) and extranucleolar (B) fractions of HeLa cells prepared by the technique of Penman *et al.* (1966). After zonal sedimentation, samples of RNA of various sizes were analyzed for base composition (C, A, G, U). The total radioactivity and % G + C are given in the figures, and the total base analysis is printed beneath. From Soeiro *et al.* (1966).

from HeLa cells, labeled for 30 minutes or less, did indeed have a rapidly labeled DNA-like nonribosomal RNA fraction in the size ranges from 6 S to 30 S [i.e., much smaller than the HnRNA (Penman *et al.*, 1963)]. This RNA fraction could be discharged from polyribosomes by EDTA treatment of cell extracts or by puromycin treatment of cells before extraction (Darnell, 1968; Penman *et al.*, 1968). Most important, the polyribosomes of cells infected with various viruses were shown to contain virus-specific RNA when the only proteins being formed in the cells were virus-specific proteins (Penman *et al.*, 1968; Becker and Joklik, 1964). In addition, in the case of poliovirus, the entire viral

RNA molecule (7000 nucleotides) served as the mRNA in the virus polyribosomes, suggesting that the techniques of isolating mRNA from polyribosomes yielded whole mRNA molecules (Penman *et al.*, 1964).

The techniques of mRNA isolation from polyribosomes has now been widely used in obtaining mRNA which can stimulate the production of cell-free protein synthesis by heterologous ribosomes.

#### A. *Proof of the HnRNA Turnover*

With the characterization of two RNA species, both having a low G + C content and both rapidly labeled, and with the just-discovered precedent of pre-rRNA, the question was posed: Is HnRNA a precursor to mRNA (Scherrer *et al.*, 1963; Penman *et al.*, 1963)? A partial and negative answer to the question was soon available: all of the HnRNA could not possibly be destined to function as mRNA in the cell cytoplasm. Scherrer and Marcaud (1965) found that duck erythroblasts, nucleated red blood cell precursors, synthesized almost no pre-rRNA but did make HnRNA. When they removed these cells from medium containing radioactive precursor or added actinomycin D, about one-half of the labeled nuclear RNA disappeared to acid-soluble form without being detected in the cytoplasm. Erythroblasts, however, are highly differentiated cells devoted to the manufacture of one protein from a long-lived mRNA, and it was not certain that this "turnover" of HnRNA in the nucleus was also true for less specialized types of cells. For example it was established that a large portion ( $\sim 50\%$ ) of the pre-rRNA of HeLa cells turned over which might account for all turnover in growing cells. In addition, if in growing cells an mRNA with a rapid turnover existed, the rapidly labeled nuclear RNA might be a precursor to such a minor rapidly turning over cytoplasmic fraction in growing cells.

Therefore a study was undertaken on the relative labeling of HnRNA compared to pre-rRNA in HeLa cells. The polysomal mRNA ( $\sim 0.6 \times 10^6$  daltons) which had a half-life of at least 3-4 hours (Penman *et al.*, 1963) constituted less than 5% of

the total polysomal RNA (average of 8 ribosomes,  $2.4 \times 10^6$  daltons of RNA in each =  $20 \times 10^6$  daltons). However, it was found that HnRNA was labeled at least 4–5 times faster than pre-rRNA (Soeiro *et al.*, 1968) (Fig. 4). [The nucleotide pools for HnRNA and pre-rRNA appear to be the same (Wu and Soeiro, 1971) indicating that the faster labeling of HnRNA compared to pre-rRNA was not due to differential entry of label into separate acid-soluble pools.] Thus a turnover time of at most a few minutes for all the mRNA of the cell would have been required for all the HnRNA to be used as a precursor to mRNA.

Perhaps the most unambiguous demonstration of the rapid turnover of HnRNA was obtained by examining, after various label times, the base composition of the total labeled RNA sedimenting at 45 S, which could be easily identified because of the peak of UV absorbance of the 45 S pre-rRNA (Fig. 5). Completely labeled purified 45 S pre-rRNA has a guanine plus cytosine content of 68% (Maden, 1971) whereas the labeled nuclear RNA

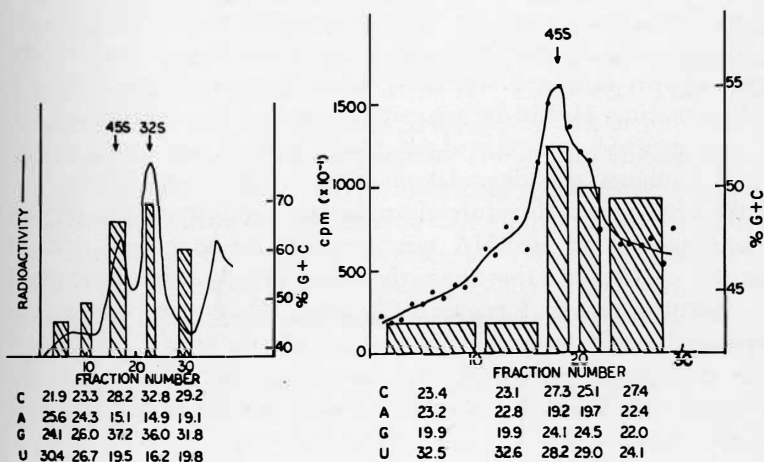


FIG. 4. Base composition of "pulse-labeled" nuclear RNA of HeLa cells. The total HeLa cell nuclear RNA from cells grown in  $^{32}\text{P}$  for 4 generations (left panel) or minutes (right panel) was fractionated by zonal sedimentation and base analysis was carried out on the indicated sections of the gradients. The 45 S peak was marked by the optical density tracing which followed exactly the radioactivity of the left panel. From Soeiro *et al.* (1968).

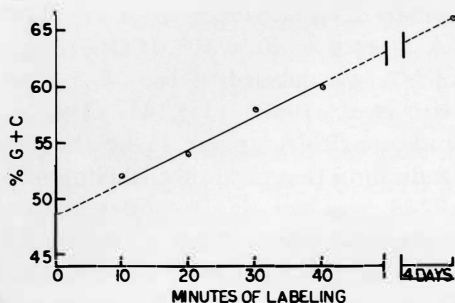


FIG. 5. Changing base composition of RNA in 45 S region.  $^{32}\text{P}$ -labeled nuclear RNA from cells labeled for various times was fractionated by zonal sedimentation as in Fig. 4. Samples equivalent to fractions 17–19 of Fig. 4 (right panel) were subjected to base composition analysis.

sedimenting in the 45 S region after a 10-minute label period was only 53% G + C (which calculates to be 6/10 HnRNA of 44% G + C and 4/10 pre-rRNA of 68%). After longer label times, the average G + C composition in the 45 S region rose gradually to finally attain 65% after several generations. Even in this restricted size class, 45 S, there was more label initially entering HnRNA than pre-rRNA and the lifetime of the 45 S pre-rRNA was only about 10–15 minutes indicating a very rapid synthesis and degradation of most of the HnRNA.

By 1968 it was therefore clear that the majority of the nucleotide sequences in HnRNA were destined to be turned over in the cell nucleus and that the early studies which claimed complete or almost complete turnover of nuclear RNA were due to the combined behavior of HnRNA (e.g., Harris *et al.*, 1963) and that fraction of pre-rRNA that turns over. None of this work, however, precluded the possibility that a portion of the HnRNA might become mRNA.

### B. Failure of RNA:DNA Hybridization to Prove HnRNA Conversion to mRNA

The problem of the possible origin of mRNA seemed susceptible to test by RNA:DNA hybridization studies. A number of

workers had demonstrated that a portion of the nuclear RNA or total cellular "pulse"-labeled RNA would indeed bind to homologous cell DNA as a ribonuclease-resistant hybrid (Scherrer *et al.*, 1963; Hoyer *et al.*, 1963; Birnboim *et al.*, 1967; Whiteley *et al.*, 1966), and if cytoplasmic mRNA were derived from HnRNA it should block ("compete with") such hybrid formation. Many such competition hybridization studies were reported where labeled nuclear RNA or labeled mRNA were hybridized alone or with unlabeled competing RNA. The general result of all these studies was that nuclear RNA could block hybrid formation by cytoplasmic or mRNA but that cytoplasmic RNA either failed to block all nuclear RNA hybrids or did so much less efficiently than did nuclear RNA itself (Birnboim *et al.*, 1967; Shearer and McCarthy, 1967, 1970; Soeiro and Darnell, 1970).

Unfortunately for this line of experimentation, the specificity of the hybridization and competition reactions were not sufficient to allow firm conclusions about the relatedness of HnRNA and mRNA (Fig. 6). Britten and Kohne recognized in 1968 (Britten and Kohne, 1968) that some DNA regions were repeated many times in the genome because a portion of denatured mammalian DNA reannealed much faster than the remainder. This rapid annealing was first studied for DNA:DNA interactions but is also true for RNA:DNA interactions (Melli and Bishop, 1969; Pagoulatos and Darnell, 1970). Many of the repeated regions were recognized to be nonidentical but similar enough to allow cross-hybridization (Melli and Bishop, 1969; Pagoulatos and Darnell, 1970; McCarthy and Duerksen, 1970). Thus, the identity of rapidly hybridizing sequences in mRNA and HnRNA could

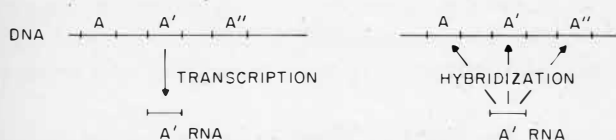


FIG. 6. Lack of specificity of hybridization of RNA transcribed from repeated DNA. Mammalian cell DNA contains sites which are similar but nonidentical (A, A', A'', Britten and Kohne, 1968). When these sites are transcribed ( $A_{RNA}'$ ) the resulting RNA may hybridize equally well to any similar site thus making it impossible to conclude from which site the RNA originated.

not be proved from RNA:DNA hybridization studies (Fig. 6). All that could be legitimately concluded from such work was that both HnRNA and mRNA had regions which were transcribed from these quasi-repeated sites in DNA, and that the HnRNA apparently contained more such sites, including highly repeated regions not represented at all in the mRNA (Shearer and McCarthy, 1967, 1970; Darnell and Balint, 1970). Finally, both HnRNA and mRNA contained slowly hybridizing regions, perhaps but not proved to be, transcribed from regions of DNA that specified proteins (Perry *et al.*, 1970; Scherrer *et al.*, 1970; Pagoulatos and Darnell, 1970; Darnell and Balint, 1970; Georgiev *et al.*, 1972). Whether these slowly hybridizing regions in the HnRNA and mRNA were the same remained unknown.

#### IV. THE DERIVATION OF mRNA FROM HnRNA

With the failure of the hybridization experiments to resolve the issue of mRNA origin, it remained an important unproved possibility that mRNA was derived in some selective way from HnRNA. Such specific mechanisms drew the interest of many workers and detailed models were described as to how mRNA derivation might be accomplished (Scherrer and Marcaud, 1968; Georgiev, 1969).

What was missing in the attempt to link HnRNA with mRNA biogenesis was, however, a distinct, reliable, sequence identity between the two. Since 1970 our laboratory, among others, has been concerned with two types of sequences that have been found in both HnRNA and mRNA. These recent results strongly indicate that mRNA does arise from posttranscriptional modification of HnRNA molecules. These two shared sequences are (a) viral specific RNA in cells transformed by DNA viruses and (b) polyadenylic acid. A summary of recent work in each area follows.

##### A. *The RNA of Virus Transformed Cells*

It had been found in the 1960s that cells transformed by small DNA tumor viruses produced virus-specific antigens (Habel, 1962; Huebner *et al.*, 1963) and that the total cell RNA contained virus-specific sequences (Benjamin, 1966). When it was

demonstrated that the DNA of SV40 virus was covalently integrated into cellular DNA (Sambrook *et al.*, 1968), it became reasonable to test whether HnRNA of virus transformed cells contained virus-specific sequences and whether the size of any HnRNA molecules bearing virus-specific sequences was larger than the RNA molecules from polyribosomes containing the same sequences (Fig. 7). The first results with SV40 transformed cells showed that all of the polysomal, virus-specific, presumed mRNA was smaller than 28 S while all of the HnRNA containing virus-specific sequences was larger than 32 S (7000 nucleotides) and a sizable fraction (30%) was larger than 45 S (14,000 nucleotides) (Lindberg and Darnell, 1970). Since the virus DNA contained only 4000–5000 base pairs, and it is now known to exist in only 1 or 2 copies per cell (Gelb *et al.*, 1971), it seemed likely that the virus RNA regions of HnRNA molecules were transcribed as part of a molecule covalently linked with cell RNA regions. HnRNA molecules from cells transformed with either SV40 virus or adenovirus type 2 (Ad-2) were selected by hybridization to virus DNA and found to contain both cell and virus sequences; but the polysomal virus-specific mRNA did not contain

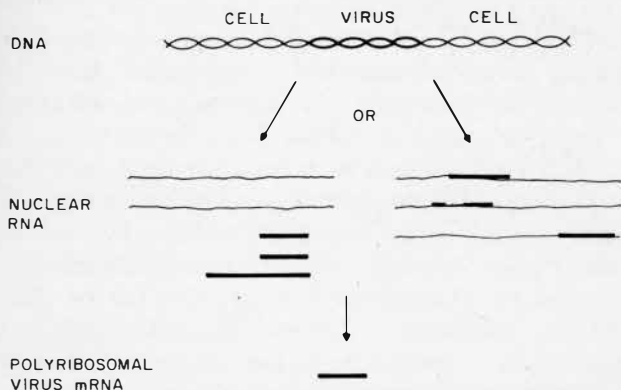


FIG. 7. Transcription of integrated viral DNA. Cells transformed by DNA viruses (Sambrook *et al.*, 1969) offer a test of whether large HnRNA contains sequences also present in polysomal mRNA. Experiments showed the model on the right was likely (Lindberg and Darnell, 1970; Wall and Darnell, 1971; Wall *et al.*, 1973).



cell sequences (Wall and Darnell, 1971; Wall *et al.*, 1973). Also with AD-2 transformed cells it has been shown that (a) the majority of the nuclear and cytoplasmic sequences competed for the same portion of Ad-2 DNA (Shimada *et al.*, 1972) and (b) while the nuclear virus-specific regions were heterogeneous in sedimentation, the polyribosomal mRNA molecules consisted of several discrete size classes (1 major and 1 or 2 minor peaks) (Wall *et al.*, 1973). Thus, work with transformed cells suggested that regions of the cell genome containing the integrated virus DNA were transcribed into long HnRNA molecules which were subsequently cleaved to form virus-specific mRNA molecules (Fig. 7, right side).

Results obtained during cytolytic infection of cells with both adenovirus and herpes virus also demonstrated that processing of large, probably entirely virus-specific, nuclear molecules probably occurred during the manufacture of virus mRNA (Roizman *et al.*, 1970; Parsons *et al.*, 1971; Wall *et al.*, 1972).

### B. Poly(A) in mRNA Biogenesis

#### 1. Presence in HnRNA and mRNA and Characterization

The second sequence discovered to be present in both mRNA and HnRNA is polyadenylic acid. This unusual homopolymeric RNA segment was discovered by Edmonds and Abrams (1959, 1962) while they were searching for enzymes responsible for general RNA synthesis in thymus nuclei. Instead, they discovered an enzyme that added polyadenylic acid either to synthetic polyadenylic acid or an endogenous primer, which they subsequently demonstrated was adenylate rich. Later, similar enzymes were found in bacteria (Gottesman *et al.*, 1962), but no physiologic role for these activities was suggested. Perhaps, decline in interest in poly(A) polymerase can be blamed on the discovery that this enzyme activity disappeared in *E. coli* cells infected with bacteriophage T4, demonstrating its dispensability at least in that genetic system (Ortiz *et al.*, 1965).

Improved evidence of the nature of cellular poly(A) came from Hadjivasilou and Brawerman (1966), who clearly identified

a "4 S" polyadenylic acid fraction in rat liver microsomes. In 1969 Edmonds and Caramela (1969) again identified a segment of T1 RNase resistant, adenylate-rich RNA which migrated during gel electrophoresis at about 9 S and was thought to originate in the nucleus of tumor cells. In none of this early work was it appreciated that poly(A) was part of other cellular RNA molecules.

It was not until 1970 that a more general role of poly(A) in mRNA metabolism was suggested. Lim and Canellakis (1970) reported that hemoglobin mRNA contained an approximately 70-unit adenylate-rich fragment and Kates and Beeson (1970) demonstrated that vaccinia mRNA, but not cellular mRNA, contained poly(A) about 150 nucleotides long; later Kates (1970) suggested that HeLa cell mRNA also contained poly(A). Work then centered on whether the cellular poly(A) was covalently attached to larger RNA molecules, and whether poly(A) was involved in HnRNA and general mRNA metabolism. It was found that some HnRNA molecules and most of the rapidly labeled polyosomal mRNA, in fact, did contain poly(A) covalently associated with the RNA chains (Edmonds *et al.*, 1971; Lee *et al.*, 1971; Darnell *et al.*, 1971a,b; Sheldon *et al.*, 1972). Moreover the size distribution of the poly(A) segments in HnRNA and that newly arrived in mRNA were initially the same (Fig. 8; Sheiness and Darnell, 1973). The poly(A) segment derived by pancreatic RNase treatment was shown to contain only adenylate residues (Molloy and Darnell, 1973; Mendecki *et al.*, 1972; Molloy *et al.*, 1972a; Sheldon *et al.*, 1972; Nakazato *et al.*, 1973) (Table I) and to be exclusively at the 3' terminus in both HnRNA and mRNA. Also, T1 RNase, which cleaves on the 3' side of all G residues, produced from both HnRNA and mRNA a fragment containing no G, 1 C, and 2 U's suggesting a similar terminal structure—G(C<sub>1</sub>, U<sub>2</sub>) A<sub>200</sub> (Molloy and Darnell, 1973) in HnRNA and mRNA (Table I).

## 2. Nuclear Origin and Posttranscriptional Addition of Poly(A)

Where in the cell, and how, does a homopolymeric segment in larger molecules originate? As mentioned previously Edmonds

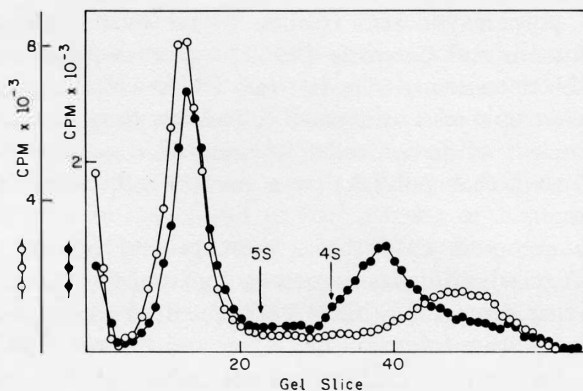


FIG. 8. Gel electrophoresis of HeLa cell poly(A). Both HnRNA and mRNA were prepared from cells labeled for 12 minutes with adenosine- $^3\text{H}$  purified and digested with T1 ribonuclease. The digests were passed through poly(U) Sepharose and the specifically bound portion of each sample was subjected to electrophoresis with  $^{32}\text{P}$  4 S and 5 S markers which were in the same gel slice in both samples; therefore the  $^3\text{H}$  nuclear and cytoplasmic poly(A) are plotted together to show they are indistinguishable in size (●—●, cytoplasmic digest; ○—○, nuclear digest; from Sheiness and Darnell, 1973).

and Abrams (1959, 1962) had found a nuclear enzyme capable of adding poly(A) to a primer without a template. When the distribution of poly(A) was examined in cells which had been

TABLE I

BASE COMPOSITION OF 3' TERMINAL POLY(A) FRAGMENTS<sup>a</sup>

|                        | T1 RNase          | Pancreatic RNase |
|------------------------|-------------------|------------------|
| 5'—G-A <sub>200</sub>  | Only A            | 1 or more G/200A |
| 5'—Py-A <sub>200</sub> | 1 or more Py/200A | Only A           |
| Actual results         | A 200             | A 200            |
|                        | G <0.1            | G <0.1           |
|                        | C 1               | C <0.1           |
|                        | U 2               | U <0.1           |

<sup>a</sup>  $^{32}\text{P}$ -labeled HnRNA and mRNA were digested with either RNase T1 or pancreatic RNase. The nucleotide ratios of the resulting purified poly(A) fragments are given in the table (see Molloy and Darnell, 1973).

TABLE II  
DISTRIBUTION OF POLY(A) IN  
HELA CELLS<sup>a</sup>

| Label time | Percent in |           |
|------------|------------|-----------|
|            | Nucleus    | Cytoplasm |
| 45 sec     | 92         | 8         |
| 1.5 min    | 97         | <3        |
| 7.5 min    | 82         | 18        |
| 20 min     | 63         | 37        |

<sup>a</sup> See Jelinek *et al.* (1973a).

labeled with adenosine-<sup>3</sup>H for very short periods (less than 5 minutes) all the poly(A) was found in the nucleus attached to HnRNA (Darnell *et al.*, 1971b; Jelinek *et al.*, 1973a) (Table II). It appeared therefore to originate as a part of nuclear molecules. Further information on the enzymatic nature of poly(A) synthesis was obtained from studies with inhibitors of nucleic acid synthesis. Actinomycin D, which almost immediately stops DNA-dependent RNA synthesis (Reich *et al.*, 1962), did not block poly(A) addition to HnRNA during the first few minutes of treatment (Darnell *et al.*, 1971b; Jelinek *et al.*, 1973a). Thus progressive movement of the RNA polymerase, which presumably makes HnRNA, was not necessary for poly(A) synthesis; rather the poly(A) appeared to be a posttranscriptional addition product. In addition, during the productive infection of HeLa cells with AD-2 virus, both virus-specific large nuclear RNA and mRNA which contained poly(A) were produced, but the AD-2 genome contained no poly(dT) regions to which poly(A) will hybridize (Philipson *et al.*, 1971). HeLa cell DNA has also been searched for polydeoxypyrimidines, but no poly(T) long enough to encode the approximately 200 nucleotide long poly(A) segment was found (Birnboim *et al.*, 1973). These results strongly support the suggestion of posttranscriptional addition of poly(A) (Fig. 9).

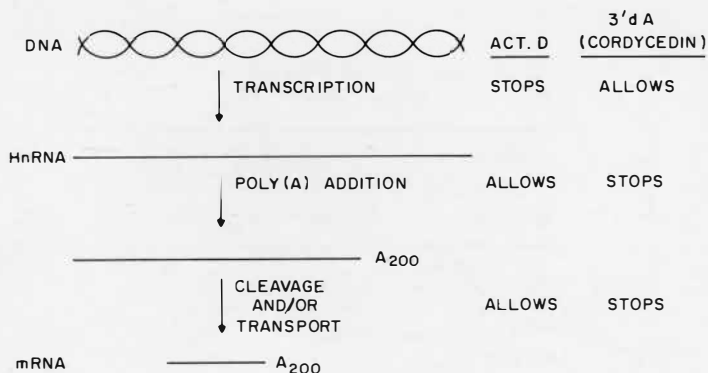


FIG. 9. mRNA biogenesis and effects of drugs. Act. D = actinomycin D.

### 3. Nuclear Role of Poly(A) in mRNA Biogenesis

Studies with another inhibitor of nucleic acid synthesis have added additional weight to the conclusion of posttranscriptional addition of poly(A) and provided the only presently available information about the role of poly(A) in mRNA metabolism. Cordycepin, 3'-deoxyadenosine (3'dA), inhibits RNA chain elongation by bacterial enzymes (Shigeura and Boxer, 1964; Klenow and Frederikson, 1964). This drug inhibits the incorporation of radioactive precursors in pre-rRNA, but not into HnRNA in HeLa cells (Siev *et al.*, 1969; Penman *et al.*, 1970); it does, however, prevent accumulation of radioactive mRNA (Penman *et al.*, 1970). In light of the presence of poly(A) in mRNA and HnRNA, it seemed possible that 3'dA might effect a block in mRNA biosynthesis by preventing poly(A) addition to HnRNA and prevent the successful processing of HnRNA. This proved to be the case (Darnell *et al.*, 1971b; Adesnik *et al.*, 1972; Mendecki *et al.*, 1972). 3'dA quickly blocks synthesis of the large 200-plus nucleotide unit of poly(A) in the nucleus and less than 10% of the normal amount of labeled mRNA reaches the polyribosomes. The locus of action of the 3'dA was also shown to be posttranscriptional (Adesnik *et al.*, 1972; Darnell *et al.*, 1973). Cells labeled for only 5 minutes have almost no radioactivity in polyribosomal mRNA. Treatment with actinomycin stops further RNA synthesis, but prelabeled RNA reaches

the polyribosomes as mRNA. This "actinomycin chase" of mRNA into the polyribosomes is greatly reduced by 3'dA. Thus it appears that poly(A) addition is necessary for mRNA appearance, and again it seems that the addition of poly(A) occurs after transcription is complete (Fig. 9).

The exact locus of action of 3'dA remains unknown. By blocking poly(A) addition, the drug could stop specific nucleolytic cleavage of HnRNA, movement of HnRNA within the nucleus or of mRNA from nucleus to cytoplasm, protect mRNA from otherwise immediate destruction in the cytoplasm, or probably other imaginable possibilities (Darnell *et al.*, 1973).

#### 4. Does All Poly(A) Reach the Cytoplasm?

In addition to indicating some role of poly(A) in proper biogenesis of mRNA, the experiments with 3'dA also contribute to a still unsettled question of considerable importance: Does every poly(A) unit (and presumably the associated mRNA) exit to the cytoplasm? When 3'dA was added to cell cultures before <sup>3</sup>H-labeled adenosine, no labeled 200 nucleotide-long poly(A) appeared in either the nucleus or the cytoplasm (Darnell *et al.*, 1971b). When cells were labeled for 5–10 minutes, so that 70–90% of the labeled cell poly(A) was in the nucleus, before 3'dA was added, further synthesis of the poly(A) segment was stopped by 3'dA; a fall in nuclear poly(A) and a rise in cytoplasmic poly(A) was observed (Jelinek *et al.*, 1973) (Table III). The labeled 200 plus nucleotide-long poly(A) appearing in the cytoplasm after 3'dA must have come from the nucleus. However, within about 60–90 minutes a fall in the total cytoplasmic poly(A) was observed, and only about 30–40% of the poly(A) which existed in the nucleus at the time of 3'dA addition could be accounted for in the cytoplasm at any one time after 3'dA. This could result either because (1) some poly(A) terminated nuclear molecules were destroyed in the nucleus and therefore the poly(A) did not exit to the cytoplasm or (2) during the 3'dA treatment some of the early-arriving, cytoplasmic poly(A)-containing molecules decayed before other poly(A)-containing molecules arrived from the nucleus.

TABLE III

APPEARANCE OF POLY(A) IN CYTOPLASM AFTER ADDITION OF  
3'-DEOXYADENOSINE<sup>a</sup>

| Time after 3'dA addition<br>(min) | Cytoplasmic poly(A) (cpm $\times 10^{-3}$ ) |     |     |    |
|-----------------------------------|---|-----|-----|----|
| 0                                 | 11  | 99  | 7.5 | 4  |
| 5                                 | —   | —   | 14  | 20 |
| 7.5                               | 28  | —   | 18  | —  |
| 10                                | —   | —   | —   | —  |
| 12.5                              | —   | —   | —   | 38 |
| 15                                | 36  | —   | —   | —  |
| 20                                | —   | 180 | 29  | —  |
| 30                                | 31  | —   | —   | 33 |
| 90                                | 16  | —   | —   | —  |
| 100                               | —   | 40  | —   | —  |

<sup>a</sup> The poly(A) in the cytoplasmic RNA of HeLa cells labeled for 5 or 7.5 minutes was determined as 0 time; 3'dA, 100  $\mu\text{g}/\text{ml}$  was then added, and samples were taken at intervals for poly(A) determination (see Jelinek *et al.*, 1973a). The four columns of figures are 4 separate experiments.

Another type of evidence about the conservation of nuclear poly(A) came from the measurement of the accumulation of adenosine-<sup>3</sup>H-labeled poly(A) in the nucleus compared to the cytoplasm. Such experiments were performed in growing HeLa cells and the accumulation of both poly(A) and total radioactive RNA was measured in cells treated with low doses of actinomycin so that rRNA synthesis was suppressed. In both situations, poly(A) accumulated in the nucleus to a fairly constant amount within about an hour while accumulation of labeled poly(A) in the cytoplasm continued and reached a larger total amount. In contrast, the total radioactivity in HnRNA remained much higher than the total radioactivity in mRNA. Thus it appeared that the poly(A) was conserved much more completely than the total HnRNA (Jelinek *et al.*, 1973a). But the question: "Is all the poly(A) conserved?" remained unanswered.

Perry *et al.* (1970) have investigated the labeling of

poly(A) in L cells using much higher absolute concentrations of adenosine than were used in HeLa cells in an attempt to quickly achieve and maintain a constant internal ATP specific activity. The total nuclear poly(A) (measured as total RNase-resistant poly(U)-adsorbable material) increased for many hours whereas the rate of cytoplasmic poly(A) accumulation reached a maximum soon after labeling. If poly(A) truly accumulates in the nucleus for almost half a generation, then probably all the nuclear poly(A) cannot be precursor to cytoplasmic poly(A).

However, it was recognized a number of years ago that cells must be grown for several generations in purine precursors to completely stop *de novo* purine synthesis. Therefore immediately after exposure to any purine precursor it is very likely not possible to quickly achieve a stable maximum nucleotide pool specific activity, and the results in the L cells could arise from a continuing increase in pool ATP-specific activity (Salzman and Sebring, 1959; McFall and Magasanik, 1960). Therefore, we have recently reexamined the kinetics of adenine-<sup>3</sup>H-labeling of nuclear and cytoplasmic poly(A) in HeLa cells previously grown for several generations in medium containing unlabeled adenine (Puckett and Darnell, unpublished observations). The results showed that the ATP pool rose in specific activity for several hours and that the accumulation of nuclear poly(A) closely followed the rising curve of pool specific activity. Nuclear poly(A) did not continue to accumulate after the nucleotide pool specific activity became constant, but reached saturation within about an hour. These results, coupled with the fact that cytoplasmic poly(A) far exceeds nuclear poly(A) in total amount, are consonant with, but do not prove, total conservation of nuclear poly(A) in the transport to the cytoplasm. To prove or disprove nuclear poly(A) turnover it would be necessary to show that more poly(A) was synthesized than ever appeared in the cytoplasm and this would require a perfect chase experiment of some kind.

The safe conclusion would seem to be that at least 30–40% of nuclear poly(A) is transported to the cytoplasm, and, while nuclear poly(A) turnover may occur, no acceptable evidence in favor of such turnover exists. The issue is clearly important since if no nuclear poly(A) turnover occurs, then all the HnRNA



that is terminated in poly(A) is probably mRNA precursor; if nuclear poly(A) turnover occurs, then some selection of poly(A) terminated HnRNA molecules appears possible.

### C. Generality of Pathway of mRNA Formation

The studies on viral RNA in transformed cells and the apparent role of poly(A) in mRNA metabolism in many different vertebrate cells (Kates, 1970; Lim and Canellakis, 1970; Edmonds *et al.*, 1971; Lee *et al.*, 1971; Darnell *et al.*, 1971a; Pemberton and Baglioni, 1972; Swan *et al.*, 1972) have provided the first strong evidence that some HnRNA is a precursor to mRNA. Furthermore it appears that perhaps all eukaryotic cells may utilize this pathway in manufacturing mRNA. For example, slime molds, eukaryotic cells of a lower order, contain a 3' poly(A) in both nuclear RNA and mRNA, but in this case the nuclear RNA is only slightly larger than mRNA (Firtel *et al.*, 1972). When these cells are labeled and then exposed to inhibitors of RNA synthesis, the nuclear poly(A) containing RNA disappears and appears as slightly lower molecular weight mRNA in the cytoplasm.

Another approach that demonstrates mRNA sequences in HnRNA is the identification of specific mRNA sequences for hemoglobin in large HnRNA molecules. Hemoglobin mRNA can be purified from reticulocytes and a radioactively labeled complementary copy of the mRNA prepared either as an RNA copy with *B. lysodeikticus* RNA polymerase (Melli and Pemberton, 1972) or as a DNA copy with the enzyme reverse transcriptase, available from RNA tumor viruses (Imaizumi *et al.*, 1973). Both types of mRNA copy have been shown to hybridize to high molecular weight RNA from the nuclei of duck erythroblasts. In addition, high molecular weight RNA from erythroblasts was injected into frog oocytes, which will translate exogenous mRNA, and the injected oocytes made hemoglobin (Williamson *et al.*, 1973). All these experiments have been criticized because of the possible contamination of high molecular weight fraction with the smaller hemoglobin mRNA which is very abundant in reticulocytes. In

the experiments of Imaizumi *et al.* (1973), however, an attempt was made to overcome this difficulty by treating the HnRNA with DMSO; there still remained a small amount of very high molecular weight RNA complementary to a hemoglobin mRNA copy.

Recently, Ruiz-Carillo *et al.* (1973) have reported a most convincing demonstration that high-molecular-weight RNA does contain hemoglobin sequences. They purified HnRNA from nucleated chicken erythroblasts to which rabbit hemoglobin mRNA had been purposely added and found that the hemoglobin synthesis dictated by the large HnRNA in a cell-free protein synthesizing system was chicken hemoglobin.

It seems safe to conclude at this point that most if not all eukaryotic cells make at least a large proportion of their mRNA via the pathway of poly(A) addition to HnRNA followed by cleavage of the HnRNA to yield mRNA. It should be mentioned here that the poly(A) pathway is not obligatory for the production of mRNA by eukaryotic cells since histone mRNA appears to lack poly(A) entirely (Adesnik and Darnell, 1972; Greenberg and Perry, 1972) and to appear in polysomes almost instantly after synthesis.

## V. REGULATION AND CHROMOSOME STRUCTURE

### A. Types of Regulation

On the assumption that the major pathway of mRNA biogenesis is at least partly understood, the question can be posed; How is regulation of protein synthesis accomplished in mammalian cells? The only organisms in which gene regulation is presently understood are bacteria which utilize transcriptional control to the virtual exclusion of other possibilities. A bacterium regulates protein synthesis by either preventing a given mRNA from being transcribed from DNA [e.g., *lac* repressor prevents  $\beta$ -galactosidase mRNA formation (Jacob and Monod, 1961; Zubay *et al.*, 1970)] or greatly enhancing the transcription of a given mRNA [e.g., the *ara C* gene product enhances the synthesis of mRNA

for arabinose-metabolizing enzymes (Zubay *et al.*, 1971; Greenblatt and Schleif, 1971)]. Protein synthesis automatically follows mRNA production because ribosomes engage the mRNA even before transcription has finished (Miller *et al.*, 1970); therefore regulation in bacteria occurs by the transcription of a given mRNA.

The information that proved transcriptional control of mRNA manufacture in bacteria was obtained first by genetic analysis: mutant cells were obtained in which aberrations in control existed (Jacob and Monod, 1961); the mutations were inserted next to and far from genes of interest. These experiments identified the aforementioned regulatory genes which produce diffusible regulator proteins (repressor and activators) and binding sites on the chromosome where such regulators interact (operators, promoters) to control structural genes. Finally, biochemical experiments demonstrated changes in levels of specific mRNA in accord with the fluctuations in the rate of specific protein synthesis (Hayashi *et al.*, 1963). Thus through a combination of genetic exploration of chromosome structure and a biochemical analysis of the chromosome transcripts, bacterial gene regulation was understood.

In eukaryotic cells, transcriptional regulation of mRNA production most probably also exists. For example, radioautographic examination of insect chromosomes from cells exposed to  $^3\text{H}$ -labeled nucleosides demonstrates changing patterns of RNA synthesis in different morphologically identifiable regions of the chromosomes during development (Pelling, 1970; Daneholt *et al.*, 1970). Also, cells which make ovalbumin and avidin (Means *et al.*, 1972; Palacios *et al.*, 1973) or hemoglobin (Ross *et al.*, 1972; Terada *et al.*, 1972) have a high concentration of specific mRNA after the rate of synthesis of the specific protein has been increased. Since none of these proteins nor mRNA can be detected in the presumed precursor cells, it seems likely that transcriptional regulation is important in these cases.

It by no means follows, however, that simple transcriptional regulation of mRNA production is the only possible mode of regulation of rates of protein synthesis in mammalian cells (see Table IV). First of all, the half-life of the total mRNA in cultured mammalian cells is generally many hours, not a few minutes

TABLE IV

REGULATION OF RATES OF PROTEIN SYNTHESIS IN MAMMALIAN CELLS

| Location          | Life history of mRNA        | Possible levels of regulation |
|-------------------|-----------------------------|-------------------------------|
| Nucleus           | 1. Transcription of HnRNA   | 1. Transcriptional            |
| Nucleus           | 2. Poly(A) addition         | 2. Posttranscriptional        |
| Nucleus           | 3. Cleavage                 |                               |
| Nucleus/cytoplasm | 4. Transport of mRNA        |                               |
| Cytoplasm         | 5. Translation into protein | 3. Translational modulation   |

(Penman *et al.*, 1963; Singer and Penman, 1973; Perry and Kelley, 1973). Nevertheless, cultured cells can be shown to vary the synthesis of both specific and total proteins over very short periods of time (Tompkins *et al.*, 1969; Fan and Penman, 1970; McCormick and Penman, 1968). Thus, without fluctuations in the amount of mRNA, *translational modulation*—the greater or lesser use of existing mRNA (Darnell *et al.*, 1973)—can change the rate of protein output.

Further, even in those cases where increased accumulation of specific mRNA can be documented (Means *et al.*, 1972; Palacios, 1973; Ross *et al.*, 1972; Terada *et al.*, 1972), it is not clear that increased transcription per se is solely responsible. Because the mechanisms of mRNA formation in animal cells that has been described earlier in this paper involves several steps not carried out by bacteria (Table IV, life history steps 2–4), it seems reasonable to consider the possibility of *posttranscriptional regulation* of mRNA formation. Such regulation would involve a decision(s) to use some but not all potentially usable mRNA sequences available in HnRNA. Such decisions could involve selected destruction (or preservation) of parts of some molecules or destruction of whole unneeded or uncalled-for molecules (Table IV). Some implications of posttranscriptional regulation have been considered in a previous paper (Darnell *et al.*, 1973) and will be briefly discussed again in the conclusion of this paper. To call for consideration of such models is not to suggest that they truly operate in cells.

Unlike the bacterial physiologist, the animal cell biologist interested in molecular events of protein synthesis regulation can expect, at best, feeble and distant assistance from detailed genetics of regulatory "genes." However, the logic of bacterial physiology that may hold promise for animal cell biology is the study of chromosome structure through nucleic acid biochemistry with a view to discovering how the structural gene regions are arranged among regions that might have a regulatory function.

Such information might be obtained by studying the arrangement of sequences in DNA, but it should also be very important to study the arrangement of sequences transcribed into HnRNA, which is a copy of a long region of DNA containing an mRNA region or regions. If posttranscriptional regulation of mRNA manufacture does occur, sites in HnRNA molecules should exist, some of which might be common among many molecules, to provide recognition points for agents (? proteins) involved in regulation. Even if posttranscriptional regulation does not occur, there are many posttranscriptional modifications necessary to generate mRNA from HnRNA molecules. Some common recognition points would appear to be needed. Accordingly, our laboratory has recently been engaged in studying arrangement of identifiable, repeated sequences in HnRNA molecules that are not found in cytoplasmic mRNA to determine whether some regularity of arrangement exists which might ultimately point to specific functions.

## *B. A Method for the Study of Chromosomal Order*

### *1. The Structure of HnRNA*

To determine whether there is a regularity in arrangement of particular segments within a series of different long molecules it is necessary (a) to have a marker for one particular spot in every molecule, preferably one of the two ends; (b) to be able to recognize and measure sequences common to some if not all the molecules, (c) to have a means of partial degradation followed by isolation of the marked region(s) of the starting molecule. In the analysis of the structure of HnRNA, the first of

these requirements is met by selecting poly(A)-containing HnRNA which therefore locates a 3' marker. Several regions have been found to be common to many HnRNA molecules, which satisfies the second requirement for studying sequence arrangements: (1) These include a T1 RNase-resistant uridylylate-rich region called oligo(U) (Molloy *et al.*, 1972b) and (2) portions of the HnRNA which arise from transcription of repeated regions of the cell DNA (Pagoulatos and Darnell, 1970; Jelinek and Darnell, 1972; Jelinek *et al.*, 1973b). To perform the sequence ordering, partial degradation of poly(A) terminated molecules can be achieved by limited alkali treatment, reselection of the 3' portion of the molecules, and assay for the repeated oligonucleotides (Molloy *et al.*, 1974).

In addition to general information about HnRNA structure provided by studying common regions that do not exit to the cytoplasm it would, of course, be desirable to examine the distribution of some sequence(s) which is destined to become mRNA. With this goal in mind, we have also begun an examination of the position of the virus-specific sequences in HnRNA from cells transformed by adenovirus (Wall *et al.*, 1973). The results of these various studies on HnRNA structure will be separately summarized.

## 2. *Oligo(U)*

HnRNA from HeLa cells contains uridylylate-rich oligonucleotides which are released by T1 RNase digestion and can be assayed by affinity chromatography on poly(A) Sepharose followed by gel electrophoresis (Fig. 10) (Molloy *et al.*, 1972b). The average size of these regions is about 30 nucleotides, estimated both from electrophoretic migration and the presence of one 3' terminal GMP residue per 30 total residues ( $U_{26}C_2A_2G$ ). The occurrence in HnRNA at a fairly high content [20–30 nucleotides in oligo(U) per 20,000 total nucleotides in HnRNA] of a fragment of this sort suggested repetition in many HnRNA molecules. Since the oligo(U)-segments are internal (i.e., not immediately at the 3' terminus), they were thought likely to arise by transcription of a repeated complementary DNA site. This was confirmed by finding rapid hybridization of oligo(U) to cell DNA as well as by

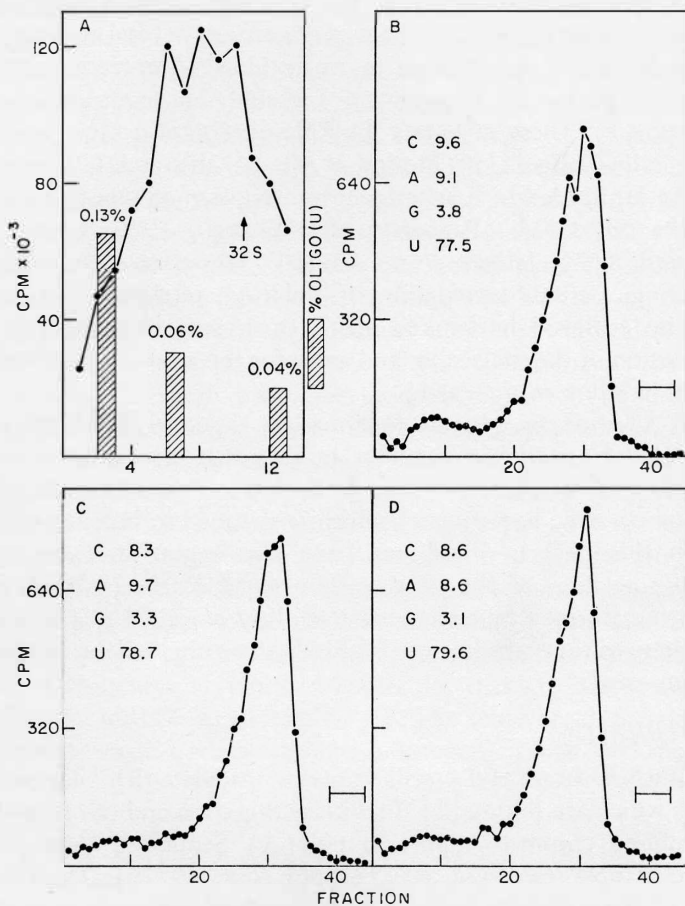


FIG. 10. The presence of uridylate-rich sections in HnRNA.  $^{32}\text{P}$ -labeled HnRNA from HeLa cells, suppressed for pre-rRNA formation by exposure to a low level of actinomycin D, was fractionated by zonal sedimentation. Samples of three sizes were digested with RNase T1 and exposed to poly(A) Sepharose; specifically bound material was collected and analyzed by gel electrophoresis revealing the uridylate-rich fragments about 20–40 nucleotides long. The base composition of the fragments (C, A, G, U, insets) was determined after elution from the acrylamide gel. From Molloy *et al.* (1972b).

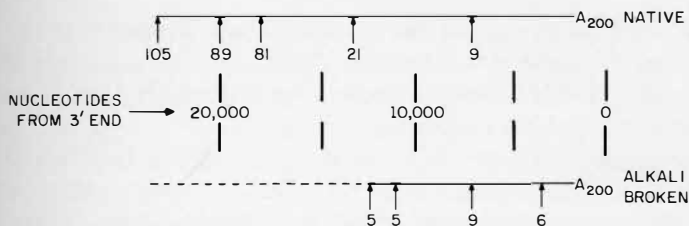


FIG. 11. Distribution of oligo(U) in HnRNA.  $^{32}\text{P}$ -labeled HnRNA terminated with the poly(A) was selected by poly(U) Sepharose chromatography from several size classes. The poly(A) content was measured to determine the average length of a sample

| 200                      | cpm in poly(A)      |
|--------------------------|---------------------|
| Avg. length of molecules | total cpm in sample |

and the oligo(U) was measured to determine the amount of oligo(U) in molecules of different lengths. Samples 20,000 nucleotides in length were alkali treated; the resulting 3' fragments were recollected by poly(U) Sepharose chromatography and subjected to zonal sedimentation, and poly(A) and oligo(U) were measured in these fragments. The results show the average number of nucleotides in oligo(U) as a function of the length of poly(A)-terminated HnRNA. Neither the alkali broken nor "native" molecules shorter than 12,000–15,000 nucleotides have much oligo(C). From Molloy *et al.* (1974).

the presence of dT:dA rich regions in DNA (Shenkin and Burdon, 1972), transcription of which might lead to oligo(U).

The first experiments on the distribution of these oligo(U) segments in HnRNA showed a higher absolute percentage in larger molecules than in smaller ones indicating a considerably greater molar percentage in long HnRNA molecules (Fig. 10).

Poly(A) terminated HnRNA has now been found (Molloy *et al.*, 1974) to contain approximately 100 residues in oligo(U) regions out of 20,000 residues in the longest HnRNA; shorter poly(A) terminated HnRNA molecules (less than 10,000 nucleotides) from the cell have less or no oligo(U), and mRNA has none (Molloy *et al.*, 1972b) (Fig. 11). In addition very long HnRNA molecules containing oligo(U) were subjected to alkali breakage, and all of the 3' terminal poly(A) containing fragments, even those as long as 10,000–15,000 nucleotides, lacked oligo(U).



Several conclusions can be drawn from these experiments:

1. Oligo(U) does not mark the 5' boundary between mRNA regions and HnRNA not destined to become mRNA; it is too far from the 3' poly(A) terminus.

2. Since oligo(U) may be close to the 5' end of HnRNA (it is at least a great distance from the 3' end), some conclusions about the nature of shorter HnRNA molecules seem possible. If the shorter molecules were nascent (partially synthesized) they should contain a higher proportion of 5' regions than the longer molecules, but the opposite is found to be true by oligo(U) analysis. Therefore the majority of the smaller molecules which are present in cell nuclei labeled for several hours must either be derived from longer molecules or be unrelated to longer molecules.

### 3. *Repeated Sequences in HnRNA*

As mentioned previously, Britten and Kohne (1968) realized that base-pairing association ("hybridization") of a fraction of mammalian cell nucleic acids at rates as fast, or in some cases faster, than viral or bacterial nucleic acids must mean that some sequences are repeated many times in mammalian cell DNA. Since then it has been shown that some repeated DNA sequences (particularly satellite DNAs) are not transcribed (Flamm *et al.*, 1969) whereas other repeated sequences apparently are (Melli and Bishop, 1969; Shearer and McCarthy, 1970; McCarthy and Duerksen, 1970; Pagoulatos and Darnell, 1970; Darnell and Balint, 1970; Perry *et al.*, 1970; Scherrer *et al.*, 1970). Transcription of these repeated regions accounts for the most rapidly hybridizing portions of HnRNA. Several points about repeated sequences in HnRNA and mRNA have been established for some time: (a) most HnRNA molecules contain repeated sequences (Darnell and Balint, 1970) and (b) the total fraction of HnRNA contained in repeated sequences is 10–30% with the remainder hybridizing sufficiently slowly so that it might come from "unique" DNA (Pagoulatos and Darnell, 1970; Perry *et al.*, 1970; Scherrer *et al.*, 1970); (c) some, if not most, repeated sequences in HnRNA were not present or were present in much lower quantities in mRNA (Shearer and McCarthy, 1967, 1970; Darnell and Balint, 1970).

Several new aspects of the repeated sequences in HnRNA have been recognized in the past year or two. If HnRNA is digested with RNase A, about 2–3% resists digestion. The RNase-resistant portion appears to be double-stranded based on its density in  $\text{Cs}_2\text{SO}_4$  and its base composition (Harel and Montagnier, 1970; Kronenberg and Humphries, 1972; Jelinek and Darnell, 1972; Ryskov *et al.*, 1972). The double-stranded regions arise from intramolecular base-pairing ("loops") because exposure to DMSO or boiling and quenching (quick return to  $0^\circ$ ) of the native HnRNA molecules does not decrease the amount of RNase-resistant "double-stranded" portion (Jelinek *et al.*, 1973a). When the RNase-resistant double-stranded fraction is prepared and exposed to DNA, no DNA–RNA hybridization occurs *unless it is boiled* to destroy secondary structure, after which the RNA rapidly hybridizes to DNA (Harel and Montagnier, 1971; Jelinek *et al.*, 1973b). Thus the "double-stranded" regions are derived from some repeated sites in DNA.

Further study of the effect of incomplete breakdown on the hybridizing capacity of HnRNA has shown that alkali digestion to segments about 500 nucleotides long or partial nuclease digestion (about 30–40% conversion to acid solubility) plus boiling releases about 5 times as many rapidly hybridizable (i.e., repeated) sites as exist in native HnRNA molecules (Jelinek *et al.*, 1973b). The results would seem to favor a model for HnRNA containing loops and stems. The stems and perhaps part of the loops must originate from repeated sites in DNA (Fig. 12).

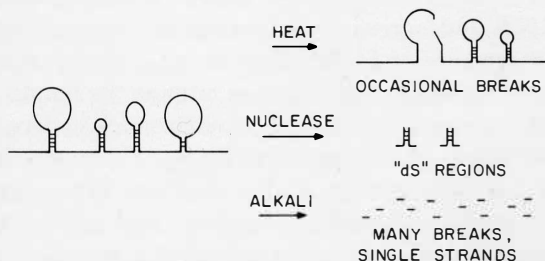


FIG. 12. Model of repeated sequences in HnRNA. Effect of simple denaturation, nuclease treatment and denaturation, and extensive alkali treatment in releasing internally base-paired regions of HnRNA.

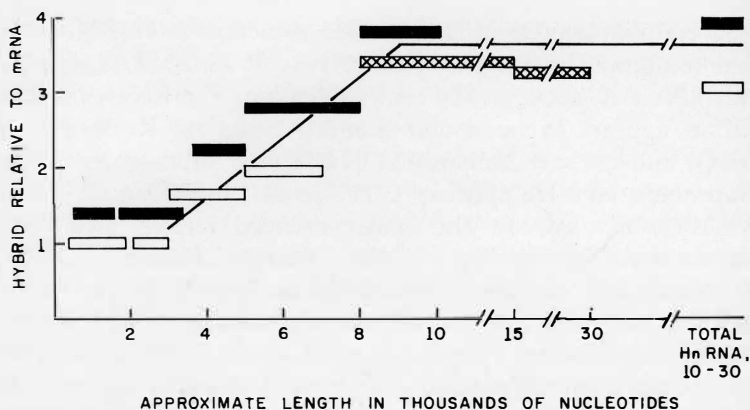


FIG. 13. Hybridization rate of HnRNA fragments.  $^{32}\text{P}$ -labeled large molecules of HnRNA terminated with poly(A) were prepared by poly(U) Sepharose chromatography. A portion of these molecules was alkali-treated, and the 3'-terminal fragments were recollected and separated into various size classes. The hybrids formed with a vast excess of cellular DNA at 6 and 24 hours were determined for each sample and compared to the hybridization rate of mRNA. The results show the 3'-terminal 4000 nucleotides of HnRNA hybridize at the same rate as mRNA while the longer HnRNA molecules hybridize at a faster rate (Molloy *et al.*, 1974).

With the availability of techniques for distinguishing the location of regions of HnRNA relative to the 3' terminal poly(A), experiments to locate rapidly hybridizable (repeated) regions within the HnRNA were carried out (Fig. 13). The shortest 3' terminal fragments (3000 nucleotides or less) obtained from HnRNA showed a substantially slower hybridization than the total HnRNA and agreed in hybridization rate with mRNA. 3'-Terminal fragments of HnRNA between 3000 and 8000 nucleotides showed increasing hybridization rates as the length increased, but fragments from 8000 to 30,000 nucleotides in length hybridized at the same rate as total HnRNA. These results indicate that the repeated sequences in HnRNA are not present in the 3'-terminal, presumably mRNA region, but are approximately evenly interspersed with nonrepeated regions beginning 3000-4000 nucleotides from the 3' end, continuing all the way up to 20,000-30,000 nucleotides (Fig. 11). It should be empha-

sized that this is an average distribution for all HnRNA molecules, not a detailed picture of any one.

#### 4. *Adenovirus-Specific RNA*

As mentioned previously, it is possible to study in the HnRNA of DNA virus-transformed cells, e.g., Ad-2 (adenovirus, type 2) transformed cells least some sequences that will eventually become mRNA. As tested by RNA:DNA hybridization, the majority of the Ad-2 sequences in nuclear RNA labeled for several hours appear to be the same as the sequences found in the cytoplasm. The HnRNA molecules with Ad-2 regions (i.e., hybridizable to Ad-2 DNA) appear to be distributed throughout the total size range of HnRNA with about 30% larger and 70% smaller than 45 S (Wall *et al.*, 1973). The larger Ad-2 containing HnRNA molecules contain host-specific sequences (Wall *et al.*, 1973). Also, it has been shown that the Ad-2 mRNA molecules are terminated with poly(A) like the cell mRNA. We have recently examined the distribution of Ad-2 sequences in the largest poly(A) terminated HnRNA. The Ad-2-specific RNA per molecule, i.e., per poly(A) terminus, increases until average length of about 15,000 nucleotides is reached (Table V) (M. Salditt and Darnell, 1974, in press). When long poly(A) terminated HnRNA molecules of average length 30,000 nucleotides are exposed to alkali and the 3' portion reselected, about two-thirds of the adenovirus-specific RNA remains associated with poly(A)-terminated fragments which have an average size of 6000–8000 nucleotides (Table VI). These results indicate then that (a) only HnRNA molecules greater than 10,000 nucleotides contain the maximum amount of Ad-2 specific RNA per molecule and (b) that over 60% of such virus specific sequences occur in the 6000–8000 nucleotides at the 3' terminus. Since the fragment of adenovirus DNA integrated in these cells is about 12,000 base pairs (Sharp *et al.*, 1974), it may be that the entire region is transcribed toward the 3' end of an HnRNA molecule. Modification by poly(A) addition at the 3' terminus then occurs, followed by cleavage to yield the virus-specific mRNA. However, even if all the adeno-2 sequences in poly(A)-terminated molecules

TABLE V  
DISTRIBUTION OF ADENOSINE TYPE 2 (Ad-2) SEQUENCE<sup>a</sup>

| Starting material        | Approximate size<br>Poly(U)<br>selected molecules | % Ad-2 | Ad-2 per molecule<br>(% Ad-2/% Poly(A)) |
|--------------------------|---|--------|---|
| Large HnRNA<br>(>40 S)   | 50-60 S   | 0.0092 | 0.0200                                  |
|                          | 45-50 S   | 0.0128 | 0.0204                                  |
|                          | 40-45 S   | 0.0196 | 0.0190                                  |
|                          | 30-40 S   | 0.0217 | 0.0161                                  |
|                          | 10-30 S   | 0.100  | 0.0470                                  |
| Small HnRNA<br>(10-40 S) | 35-40 S   | 0.031  | 0.012                                   |
|                          | 30-35 S   | 0.031  | 0.010                                   |
|                          | 28 S  | 0.037  | 0.008                                   |
|                          | 20-28 S   | 0.145  | 0.027                                   |
|                          | 15-20 S   | 0.250  | 0.041                                   |
|                          | 5-15 S  | 0.137  | 0.0137                                  |

<sup>a</sup> Adenosine-<sup>3</sup>H and uridine-<sup>3</sup>H-labeled HnRNA from Ad-2-transformed rat cells was separated by zonal sedimentation into > and < then 40 S and poly(A)-terminated molecules selected by poly(U) chromatography and again subjected to zonal sedimentation. The fraction of Ad-2 specific RNA was measured by hybridization to Ad-2 DNA, and poly(A) was measured as in Fig. 8.

are 3' terminal, the location of virus-specific regions in HnRNA *not* containing poly(A) is unknown.

## VI. CONCLUSIONS AND PROJECTIONS

The experiments described in this review provide evidence that a general structure may exist for poly(A) terminated HnRNA molecules and that mRNA is derived from such molecules (Fig. 14). HeLa cell HnRNA contains very long polyribonucleotides (>20,000 nucleotides) to which poly(A) is added at the 3' terminus. These regions contain repeated sequences, probably largely in internally base-paired loops, beginning some 3000 to 4000 nucleotides from the 3' end and continuing interspersed with non-repeated regions toward the 5' end. At least 12,000-15,000 nucleotides from the 3' end there occur uridylate-rich regions (or a region) averaging about 100 nucleotides in length. In poly(A)

TABLE VI  
ASSOCIATION OF ADENOSINE TYPE 2 (AD-2)-SPECIFIC RNA\*

| Sample                          | Total cpm         | AD-2 | Poly(A) |                 |
|---------------------------------|-------------------|------|---------|-----------------|
|                                 |                   |      | %       | Calculated size |
| Alkali treated, poly(U) bound   | $1.7 \times 10^6$ | 358  | 2.4     | 8,000           |
| Alkali treated, flow through    | $6.9 \times 10^6$ | 222  | 0       | —               |
| Original poly(U) selected HnRNA | —                 | —    | 0.6     | 33,000          |

\*  $^{32}\text{P}$  labeled HnRNA from Ad-2 transformed rat cells was selected by poly(U) Sepharose chromatography, and its size was estimated from poly(A) content. The sample was then exposed to alkali and the 3' portions were recollected by poly(U) selection. The bound and flow-through fractions after alkali treatment were both assayed for Ad-2 sequences and poly(A).

terminated HnRNA from rat cells transformed by Ad-2, virus-specific RNA sequences at least some of which are also found in mRNA, are mostly, if not entirely, 3' terminal. This result is consistent with the model of mRNA derivation from the 3' end of HnRNA drawn from the study of HeLa cell HnRNA (Fig. 14).

### *The Remaining Puzzle*

Unfortunately, a very major question bearing on genetic regulation appears to be unresolved: What is the form in which

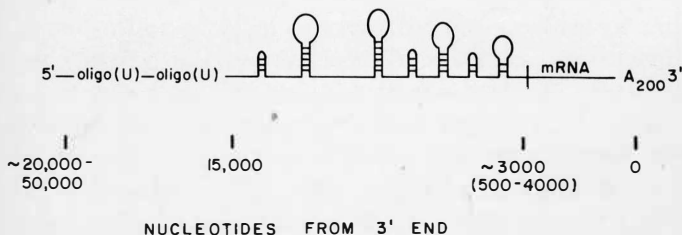
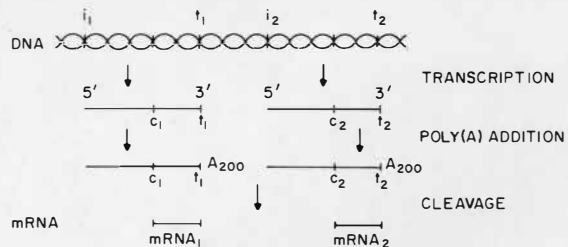


FIG. 14. Model of HnRNA from HeLa cells. The model is not drawn to constant scale, and the configuration or size of the loops and stems is not supposed to be the exact.



OR

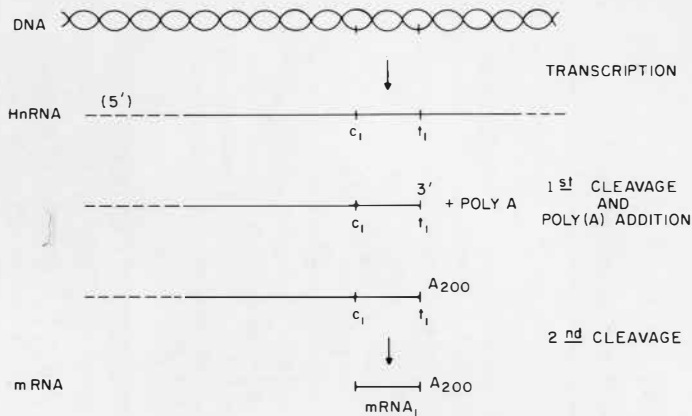
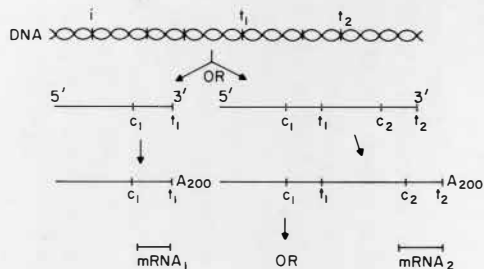


FIG. 15. Possible modes of HnRNA transcription and mRNA production. Left side: Definite sites for initiation and termination. Fixed initiation and termination points means that the starting or finishing site for a polymerase copying a given DNA region is definite;  $i$  is an initiation site;  $t$ , a termination site. For a polymerase to copy through  $t_1$  and proceed to  $t_2$  requires that a proper termination cofactor (? termination protein) is not available to stop a polymerase at  $t_1$ . HnRNA conversion to mRNA proceeds through poly(A) addition at  $t_1$  or  $t_2$ , and cleavage occurs at  $c$ . Right side: Indefinite sites for initiation and termination. Without concern as to possible specific initiation and termination sites, HnRNA is cleaved once at  $c$  to reveal a 3' terminus for poly(A) addition followed by a second cleavage at  $t$  to produce the mRNA.

HnRNA is synthesized? If processing of HnRNA (nucleolytic attack and posttranscriptional additions) is rapid, then the nature of the initial transcription product from mammalian chromosomes is still unclear. The pattern of HnRNA synthesis and mRNA formation can be envisioned in two extreme forms (Fig. 15). (1) For each transcription unit there is a fixed initiation and one or more fixed termination points for an RNA polymerase, e.g., the DNA of an entire chromomere as suggested by Judd *et al.* (1971). (2) There may or may not be fixed initiation points and there may or may not be fixed termination points, but frequently a 3' terminus which can act as a poly(A) receptor is derived by endonucleolytic attack followed by poly(A) addition and mRNA derivation (Fig. 15). To settle these questions some specific (? mRNA) nucleotide regions must be located with respect to both 5' and 3' ends of unprocessed as well as processed molecules to discover if they are always in a fixed position. If such ordering experiments were accomplished for several specific regions and a general pattern emerged, a satisfactory conclusion could be made about a fixed or flexible pattern of HnRNA synthesis. Such a task is at the moment extraordinarily difficult because the only available technique for ordering sequences within HnRNA involves selecting molecules on the basis of poly(A) content, and these molecules apparently have already entered the processing pathway. This later difficulty might be surmounted by using purified poly(A) polymerase (Winters and Edmonds, 1973) to add poly(A) to HnRNA molecules lacking it followed by positioning experiments of the type described here for poly(A)-terminated molecules.

A second major problem concerns the measurement of mRNA regions in *newly synthesized* HnRNA molecules. At present this is possible only in labeled HnRNA from whole cells. To illustrate the difficulty of measuring unprocessed HnRNA let us examine one of the current methods of measuring mRNA and mRNA sequences within HnRNA. The technique involves the use of reverse transcriptase to prepare a labeled DNA copy of a purified mRNA, followed by hybridization of the copy to unlabeled HnRNA (Imaizumi *et al.*, 1973; Axel *et al.*, 1973). This technique does not allow observation of *newly* formed HnRNA, but



only the accumulated products of synthesis and processing. If isolated nuclei (Zylber and Penman, 1971) or enzyme-chromatin preparations (Axel *et al.*, 1973) can be shown to properly start and complete very high molecular weight HnRNA, containing measurable mRNA sequences then perhaps the task of studying just completed molecules can be simplified.

A possibly promising line of attack on the problem of regularity of transcription units involves the RNA transcribed from certain sites in insect chromosomes. In the salivary glands of *C. tentans* there is the production of an RNA molecule from a single expanded region of the chromosome (a "puff") which appears to nearly match in length the entire DNA component of this region of the chromosome (Daneholt *et al.*, 1970; Daneholt and Hosick, 1973). Moreover, this giant "75 S" RNA moves apparently intact to the cell cytoplasm, where it may be a giant mRNA for a very high molecular weight secreted protein. Perhaps this interesting case and possibly others in insects (Suzuki and Brown, 1972; Suzuki *et al.*, 1972) are examples of transcriptional units which can be studied, but if they represent specialized instances in which no or minor processing occurs before the entire unit appears in the cytoplasm, then they may not be useful models for understanding general HnRNA transcription or processing. Possibly, labeled RNA from virus-transformed cells may still offer the best material for study of these problems.

Understanding the regulation of mRNA production in mammalian cells will be greatly aided by knowing whether the transcription unit is a relatively free or rigidly prescribed portion of the genome. The task of understanding regulation is not finished with this knowledge however. A fixed transcription unit would seem to favor *transcriptional* compared to *posttranscriptional* regulation of mRNA formation. But it is still possible that as cells are exposed to changing circumstances a variable amount of HnRNA from a given transcriptional unit might be processed into mRNA. An irregular sized transcription unit containing a potential mRNA in a variable position would seem to make *posttranscriptional* regulation more plausible. However, it may be that even from such an irregular transcriptional unit (or units) there is a constant probability of posttranscriptional success in

making an mRNA, and thus the important decision in regulation could still be at the level of transcription.

When a transcriptional unit is identified for a particular mRNA (Fig. 15), the final question must then be posed: During a period of regulation, when mRNA can be demonstrated to accumulate, is more of a given transcriptional unit containing that mRNA manufactured or is the same amount manufactured but a variable proportion processed into mRNA? Now that this difficult experiment can be accurately phrased we can hope for an answer soon.

#### REFERENCES

- Adesnik, M., and Darnell, J. E. (1972). *J. Mol. Biol.* **67**, 397.
- Adesnik, M., Salditt, M., Thomas, W., and Darnell, J. E. (1972). *J. Mol. Biol.* **71**, 21.
- Avery, O. T., Macleod, C. M., and McCarty, M. (1944). *J. Exp. Med.* **79**, 137.
- Axel, R., Cedar, H., and Felsenfeld, G. (1973). *Proc. Nat. Acad. Sci. U.S.* **70**, 2029.
- Barnum, C. P., Huseby, R. A., and Vermind, H. (1953). *Cancer Res.* **13**, 880.
- Bawden, F. C., and Pirie, N. W. (1937). *Proc. Roy. Soc. Ser. B* **123**, 274.
- Becker, Y., and Joklik, W. K. (1964). *Proc. Nat. Acad. Sci. U.S.* **51**, 577.
- Benjamin, T. (1966). *J. Mol. Biol.* **16**, 359.
- Birnboim, H. C., Pene, J. J., and Darnell, J. E. (1967). *Proc. Nat. Acad. Sci. U.S.* **58**, 320.
- Birnboim, H. C., Mitchel, R. E. J., and Strauss, N. A. (1973). *Proc. Nat. Acad. Sci.* **70**, 2189.
- Brachet, J. (1947). *Symp. Soc. Exp. Biol.* **1**, 207.
- Brachet, J. (1955). In "The Nucleic Acids" (E. Chargaff, and J. N. Davidson, eds.), Vol. II, p. 475. Academic Press, New York.
- Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature (London)* **190**, 576.
- Britten, R. J., and Kohne, D. E. (1968). *Science* **161**, 529.
- Brown, D. D., and Weber, C. S. (1968). *J. Mol. Biol.* **34**, 681.
- Caspersson, T. (1947). *Symp. Soc. Exp. Biol.* **1**, 127.
- Chargaff, E. (1950). *Experientia* **6**, 201.
- Chargaff, E. (1955). *Rend. 1st. Lomb. Sci. Lett. Cl. Sci. Mat. Nat. Pt. 1*, **89**, 101.
- Cold Spring Harbor Symp. Quant. Biol.* (1961) **26**.
- Colter, J. S., Bird, H. H., Moyer, A. W., and Brown, R. A. (1957). *Virology* **4**, 522.
- Crick, F. H. C. (1958). *Symp. Soc. Exp. Biol.* **12**, 138.
- Daneholt, B., and Hosick, H. (1973). *Proc. Nat. Acad. Sci. U.S.* **70**, 442.

- Daneholt, R., Edstrom, J. E., Egyhavi, E., Lambert, B., and Ringborg, U. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 513.
- Darnell, J. E. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 149.
- Darnell, J. E. (1968). *Bacteriol. Rev.* **32**, 262.
- Darnell, J. E., and Balint, R. (1970). *J. Cell. Physiol.* **76**, 349.
- Darnell, J. E., Philipson, L., Wall, R., and Adesnik, M. (1971a). *Science*, **174**, 507.
- Darnell, J. E., Wall, R., and Tushinski, R. J. (1971b). *Proc. Nat. Acad. Sci. U.S.* **68**, 1321.
- Darnell, J. E., Jelinek, W. R., and Molloy, G. R. (1973). *Science* **181**, 1215.
- Edmonds, M., and Abrams, R. (1959). *J. Biol. Chem.* **235**, 1142.
- Edmonds, M., and Abrams, R. (1962). *J. Biol. Chem.* **238**, 1186.
- Edmonds, M., and Caramela, M. G. (1969). *J. Biol. Chem.* **244**, 1314.
- Edmonds, M. P., Vaughan, M. H., and Nakazato, H. (1971). *Proc. Nat. Acad. Sci. U.S.* **68**, 1136.
- Ellem, K. A. O., and Sheridan, J. W. (1964). *Biochem. Biophys. Res. Commun.* **16**, 505.
- Elson, D., and Chargaff, E. (1952). In "Phosphorus Metabolism," Vol. 2, pp. 329. Johns Hopkins Press, Baltimore, Maryland.
- Fan, H., and Penman, S. (1970). *J. Mol. Biol.* **50**, 655.
- Firtel, R. A., Jacobson, A., and Lodish, H. F. (1972). *Nature (London) New Biol.* **239**, 235.
- Flamm, W. G., Walker, P. M. B., and McCallum, M. (1969). *J. Mol. Biol.* **40**, 423.
- Fraenkel-Conrat, H., and Williams, R. C. (1955). *Proc. Nat. Acad. Sci.* **41**, 690.
- Fraenkel-Conrat, H., Singer, B., and Williams, R. C. (1957). *Biochem. Biophys. Acta* **25**, 87.
- Gelb, L. D., Kohne, D. E., and Martin, M. A. (1971). *J. Mol. Biol.* **57**, 129.
- Georgiev, G., and Mantieva, V. L. (1962). *Biochim. Biophys. Acta* **61**, 153.
- Georgiev, G. P., Samarina, O. P., Lerman, M. I., and Smirnov, M. N. (1963). *Nature (London)* **200**, 1291.
- Georgiev, G. P. (1969). *J. Theoret. Biol.* **25**, 473.
- Georgiev, G. P., Ryskov, A. P., Coutelle, C., Mantieva, V. L., and Avaryan, E. R. (1972). *Biochim. Biophys. Acta* **259**, 259.
- Gierer, A. (1963). *J. Mol. Biol.* **6**, 148.
- Gierer, A., and Schramm, G. (1956). *Nature (London)* **177**, 702.
- Girard, M., Penman, S., and Darnell, J. E. (1964). *Proc. Nat. Acad. Sci. U.S.* **51**, 505.
- Goldstein, L., and Plaut, W. (1955). *Proc. Nat. Acad. Sci. U.S.* **41**, 874.
- Gottesman, M. E., Canellakis, Z. N., and Canellakis, E. S. (1962). *Biochim. Biophys. Acta* **61**, 34.
- Greenberg, H., and Penman, S. (1966). *J. Mol. Biol.* **21**, 527.
- Greenberg, J. R., and Perry, R. P. (1972). *J. Mol. Biol.* **72**, 91.

- Greenblatt, J., and Schleif, R. (1971). *Nature (London) New Biol.* **223**, 166.
- Gros, F., Gilbert, W., Hiatt, H. H., Attardi, G., Spohr, P. R., and Watson J. D. (1961). *Cold Spring Harbor Symp. Quant. Biol.* **26**, 111.
- Habel, K. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 433.
- Hadjivasiliou, A., and Brawerman, G. (1966). *J. Mol. Biol.* **20**, 1.
- Harel, L., and Montagnier, L. (1971). *Nature (London) New Biol.* **229**, 106.
- Harris, H. (1959). *Biochem. J.* **13**, 362.
- Harris, H., and La Cour, L. R. (1963). *Nature (London)* **200**, 227.
- Harris, H., and Watts, J. W. (1962). *Proc. Roy. Soc. Ser. B* **156**, 109.
- Harris, H., Fisher, H. W., Rodgers, A., Spencer, T., and Watts, J. W. (1963). *Proc. Roy. Soc. Ser. B* **151**, 177.
- Hayashi, M., Spiegelman, S., Franklin, N. C., and Luria, S. E. (1963). *Proc. Nat. Acad. Sci. U.S.* **49**, 729.
- Hoagland, M. G., Stephenson, M. L., Scott, J. F., Hecht, L. I., and Zamecnik, P. C. (1958). *J. Biol. Chem.* **231**, 241.
- Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A. (1965). *Science* **147**, 1462.
- Houssais, J. F., and Attardi, G. (1966). *Proc. Nat. Acad. Sci. U.S.* **56**, 616.
- Hoyer, B. A., McCarthy, B. J., and Bolton, E. T. (1963). *Science* **140**, 1408.
- Huebner, R. J., Rowe, W. P., Turner, H. C., and Lane, W. T. (1963). *Proc. Nat. Acad. Sci. U.S.* **50**, 379.
- Hurlbert, R. B., and Potter, V. R. (1952). *J. Biol. Chem.* **195**, 257.
- Imaizumi, T., Diggelman, H., and Scherrer, K. (1973). *Proc. Nat. Acad. Sci. U.S.* **70**, 1122.
- Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
- Jeanteur, P., Amaldi, F., and Attardi, G. (1968). *J. Mol. Biol.* **33**, 757.
- Jelinek, W., and Darnell, J. E. (1972). *Proc. Nat. Acad. Sci. U.S.* **69**, 2537.
- Jelinek, W., Adesnik, M., Salditt, M., Sheiness, D., Wall, R., Molloy, G., Philipson, L., and Darnell, J. E. (1973a). *J. Mol. Biol.* **75**, 515.
- Jelinek, W., Molloy, G., Fernandez-Munoz, R., Salditt, M., and Darnell, J. E. (1973b). *J. Mol. Biol.* **81**, in press.
- Judd, B. H., Shen, M. W., and Kaufman, T. C. (1971). *Genetics* **71**, 139.
- Kates, J. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 743.
- Kates, J., and Beeson, J. (1970). *J. Mol. Biol.* **50**, 19.
- Klenow, H., and Frederikson, S. (1964). *Biochim. Biophys. Acta* **87**, 495.
- Kronenberg, L. H., and Humphries, T. (1972). *Biochemistry* **11**, 2020.
- Kubinski, H., and Koch, G. (1966). *Biochem. Biophys. Res. Commun.* **22**, 346.
- Kurland, C. G. (1960). *J. Mol. Biol.* **2**, 83.
- Lee, Y., Mendecki, J., and Brawerman, G. (1971). *Proc. Nat. Acad. Sci. U.S.* **68**, 1331.
- Levene, P. A. T., and Bass, L. W. "Nucleic Acids," Chemical Catalog Co., New York, 1931.

- Lim, L., and Canellakis, E. S. (1970). *Nature (London)* **227**, 710.
- Lindberg, U., and Darnell, J. E. (1970). *Proc. Nat. Acad. Sci. U.S.* **65**, 1089.
- McCarthy, B. J., and Duerksen, J. D. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 621.
- McCormick, W., and Penman, S. (1968). *J. Mol. Biol.* **39**, 315.
- McFall, E., and Magasanik, B. (1960). *J. Biol. Chem.* **235**, 2103.
- McQuillen, K., Roberts, R. B., and Britten, R. J. (1959). *Proc. Nat. Acad. Sci. U.S.* **45**, 1437.
- Maden, B. E. H. (1971). *Progr. Biophys. Mol. Biol.* **22**, 129.
- Marshak, A. (1948). *J. Cell. Comp. Physiol.* **32**, 381.
- Marshak, A., and Calvet, F. (1969). *J. Cell. Comp. Physiol.* **34**, 451.
- Means, A. R., Comstock, J. P., Rosenfield, G. C., O'Malley, B. W. (1972). *Proc. Nat. Acad. U.S.* **69**, 1146.
- Melli, M., and Bishop, J. (1969). *J. Mol. Biol.* **40**, 117.
- Melli, M., and Pemberton, R. (1972). *Nature (London) New Biol.* **236**, 172.
- Mendecki, J., Lee, Y., and Brawerman, G. (1972). *Biochemistry* **11**, 1792.
- Miller, O. L., Hamkalo, B. A., and Thomas, C. A. (1970). *Science* **169**, 392.
- Moldave, K., and Heidelberger, C. (1954). *J. Amer. Chem. Soc.* **76**, 679.
- Molloy, G., and Darnell, J. E. (1973). *Biochemistry* **12**, 2324.
- Molloy, G. R., Sporn, M., Kelley, D. E., and Perry, R. P. (1972a). *Biochemistry* **11**, 3256.
- Molloy, G. R., Thomas, W. L., and Darnell, J. E. (1972b). *Proc. Nat. Acad. Sci. U.S.* **69**, 3684.
- Molloy, G. R., Jelinek, W., Salditt, M., and Darnell, J. E. (1974). *Cell* **1**, 43.
- Mundry, K. (1959). *Virology* **9**, 722.
- Nakazato, H., Kupp, D., and Edmonds, M. (1973). *J. Biol. Chem.* **248**, 1472.
- Noll, H., Staehelin, T., and Wettstein, F. O. (1963). *Nature (London)* **198**, 632.
- Ortiz, P., August, J. T., Watanabe, M., Kaye, A. M., and Hurwitz, J. (1965). *J. Biol. Chem.* **240**, 432.
- Pagoulatos, G. N., and Darnell, J. E. (1970). *J. Mol. Biol.* **54**, 517.
- Palacios, R., Sullivan, D., Summers, N. M., Kiely, M. L., and Schimke, R. T. (1973). *J. Biol. Chem.* **248**, 540.
- Pardee, A., Jacob, F., and Monod, J. (1959). *J. Mol. Biol.* **1**, 165.
- Parsons, J. T., Gardner, J., and Green, M. (1971). *Proc. Nat. Acad. Sci. U.S.* **68**, 557.
- Pelling, C. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 521.
- Pemberton, R., and Baglioni, C. (1972). *J. Mol. Biol.* **65**, 531.

- Penman, S. (1966). *J. Mol. Biol.* **17**, 177.
- Penman, S. (1968). *J. Mol. Biol.* **39**, 315.
- Penman, S., Scherrer, K., Becker, Y., and Darnell, J. E. (1963). *Proc. Nat. Acad. Sci. U.S.* **49**, 654.
- Penman, S., Becker, Y., and Darnell, J. E. (1964). *J. Mol. Biol.* **8**, 541.
- Penman, S., Smith, I., and Holtzman, E. (1966). *Science* **154**, 786.
- Penman, S., Rosbash, M., and Penman, M. (1970). *Proc. Nat. Acad. Sci. U.S.* **67**, 1878.
- Perry, R. P. (1962). *Proc. Nat. Acad. Sci. U.S.* **48**, 2179.
- Perry, R. P., and Kelley, D. E. (1973). *J. Mol. Biol.* **79**, 681.
- Perry, R. P., Kelley, D., and Latorre, J. (1974). *J. Mol. Biol.* **82**, 315.
- Perry, R. P., Greenberg, J. R., and Tartof, K. D. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 577.
- Philipson, L. (1961). *J. Gen. Physiol.* **44**, 899.
- Philipson, L., and Darnell, J. E. (1973). *J. Mol. Biol.* **72**, 91.
- Philipson, L., Wall, R., Glickman, R., and Darnell, J. E. (1971). *Proc. Natl. Acad. Sci. U.S.* **68**, 2806.
- Reich, E., Franklin, R. M., Shatkin, A. J., and Tatum, E. (1962). *Science* **134**, 556.
- Roizman, B., Bachenheimer, S., Wagner, E. K., and Savage, T. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 753.
- Ross, T., Ikawa, Y., and Leder, P. (1972). *Proc. Nat. Acad. Sci. U.S.* **69**, 3260.
- Ruiz-Carillo, A., Beato, M., Schurtz, G., Feigelson, P., and Allfrey, V. G. (1973). *Proc. Natl. Acad. Sci. U.S.* **70**, 3641.
- Ryskov, A. P., Saunders, G. F., Farashyam, V. R., and Georgiev, G. P. (1973). *Biochim. Biophys. Acta* **312**, 152.
- Salim, M., Maden, B. E. H., and Williamson, R. (1970). *FEBS Lett.* **12**, 109.
- Salzman, N. P., and Sebring, E. D. (1959). *Arch. Biochem. Biophys.* **84**, 143.
- Sambrook, J., Westphal, H., Srinivasan, P. R., and Dulbecco, R. (1968). *Proc. Nat. Acad. Sci. U.S.* **60**, 1288.
- Scherrer, K., and Darnell, J. E. (1962). *Biochem. Biophys. Res. Commun.* **7**, 486.
- Scherrer, K., and Marcaud, L. (1965). *Bull. Soc. Chim. Biol.* **47**, 1697.
- Scherrer, K., and Marcaud, L. (1968). *J. Cell Physiol.* **72**, Suppl. 1, 181.
- Scherrer, K., Latham, H., and Darnell, J. E. (1963). *Proc. Nat. Acad. Sci. U.S.* **49**, 240.
- Scherrer, K., Spohr, G., Granbauran, N., Morel, C., Grosclaude, J., and Chezzi, C. (1970). *Cold Spring Harbor Symp. Quant. Biol.* **35**, 539.
- Sharp, P., Petersson, U., and Sambrook, J. (1974). *J. Mol. Biol.* in press.
- Shearer, R. N., and McCarthy, B. J. (1967). *Biochemistry* **6**, 283.
- Shearer, R. W., and McCarthy, B. J. (1970). *J. Cell Phys.* **75**, 97.

- Sheiness, D., and Darnell, J. E. (1973). *Nature (London) New Biol.* **241**, 265.
- Sheldon, R., Kates, J., Kelley, P., and Perry, R. P. (1972). *Biochemistry* **11**, 3829.
- Shenkin, A., and Burdon, R. H. (1972). *FEBS Lett.* **22**, 157.
- Shigeura, H. T., and Boxer, G. E. (1964). *Biochem. Biophys. Res. Commun.* **17**, 758.
- Shimada, K., Fujinaga, K., Hama, S., Sekikawa, K., and Ito, Y. (1972). *J. Virol.* **10**, 648.
- Sibatani, A., Yamana, K., Kimura, K., and Okagaki, H. (1959). *Biochim. Biophys. Acta* **33**, 590.
- Sibatani, A., deKloet, S. R., Allfrey, V. G., and Mirsky, A. E. (1962). *Proc. Nat. Acad. Sci. U.S.* **48**, 471.
- Siev, M., Weinberg, R., and Penman, S. (1969). *J. Cell Biol.* **41**, 510.
- Singer, R. H., and Penman, S. (1973). *J. Mol. Biol.* **78**, 321.
- Smellie, R. M. S., McIndoe, W. M., Logan, R., and Davidson, J. N. (1953). *Biochem. J.* **54**, 280.
- Soeiro, R., Birnboim, H. C., and Darnell, J. E. (1966). *J. Mol. Biol.* **19**, 362.
- Soeiro, R., Vaughan, M. H., Warner, J. R., and Darnell, J. E. (1968). *J. Cell Biol.* **39**, 112.
- Soeiro, R., and Darnell, J. E. (1970). *J. Cell Biol.* **44**, 467.
- Stanley, W. M. (1935). *Science* **81**, 644.
- Suzuki, Y., and Brown, D. D. (1972). *J. Mol. Biol.* **63**, 409.
- Suzuki, Y., Gage, P., and Brown, D. D. (1972). *J. Mol. Biol.* **70**, 637.
- Swan, D., Aviv, H., and Leder, P. (1972). *Proc. Nat. Acad. Sci. U.S.* **69**, 1967.
- Terada, M., Cantor, L., Metafora, S., Rifkind, R. A., Banks, A., and Marks, P. A. (1972). *Proc. Nat. Acad. Sci.* **69**, 3578.
- Tissieres, A., Watson, J. D., Schlessinger, D., and Hollingworth, B. R. (1959). *J. Mol. Biol.* **1**, 221.
- Tompkins, G. M., Gelehrter, T., Granner, D., Martin, D. W., Samuels, H., and Thompson, E. B. (1969). *Science* **166**, 1474.
- Wall, R., and Darnell, J. E. (1971). *Nature (London) New Biol.* **232**, 73.
- Wall, R., Philipson, L., and Darnell, J. E. (1972). *Virology* **50**, 27.
- Wall, R., Weber, J., Gage, Z., and Darnell, J. E. (1973). *J. Virol.* **11**, 953.
- Warner, J. R., Knopf, P., and Rich, A. (1963). *Proc. Nat. Acad. Sci. U.S.* **49**, 122.
- Watson, J. D., and Crick, F. H. C. (1953). *Nature (London)* **171**, 964.
- Wecker, E. (1958). *Virology* **7**, 241.
- Wellauer, P. K., and Dawid, I. B. (1973). *Proc. Nat. Acad. Sci. U.S.* **70**, 2827.
- Whiteley, A. H., McCarthy, B. J., and Whiteley, H. R. (1966). *Proc. Nat. Acad. Sci. U.S.* **55**, 519.
- Williamson, R., Drewienkiewicz, C., and Paul, J. (1973). *Nature (London) New Biol.* **241**, 66.

- Winters, M. A., and Edmonds, M. (1973). *J. Biol. Chem.* **248**, 4756-4763.
- Wu, R., and Soeiro, R. (1971). *J. Mol. Biol.* **58**, 481.
- Yoshikawa-Fukada, M., Fukada, T., and Kawade, Y. (1965). *Biochim. Biophys. Acta* **103**, 383.
- Zalokar, M. (1959). *Nature (London)* **183**, 330.
- Zamecnik, P. C. (1960). *Harvey Lect.* **54**, 256.
- Zimmerman, E., and Holler, R. W. (1967). *J. Mol. Biol.* **20**, 1.
- Zubay, G., Schwartz, D., and Beckwith, J. (1970). *Proc. Nat. Acad. Sci. U.S.* **66**, 104.
- Zubay, G., Gielow, L., and Engelsberg, E. (1971). *Nature (London) New Biol.* **233**, 164.
- Zylber, E., and Penman, S. (1971). *Proc. Nat. Acad. Sci. U.S.* **68**, 2861.