PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Genetic Transcription

G. S. Stent

Proc. R. Soc. Lond. B 1966 **164**, 181-197
doi: 10.1098/rspb.1966.0022

Genetic transcription*

By G. S. Stent

Department of Molecular Biology and Virus Laboratory, University of California, Berkeley

The purine and pyrimidine base sequence of *DNA*, permanent repository of the genetic information, must be transcribed on to ribopolynucleotides before genotype can be translated into phenotype. This transcription gives rise to three recognizably different classes of *RNA* molecules: (1) two species of *ribosomal RNA*, one about 1500 and the other 3000 nucleotides in length, that form part of the structural members of the engine for cellular protein synthesis; (2) several dozen species of *transfer RNA*, each about 100 nucleotides in length, that provide adaptors in protein synthesis for the twenty 'standard' amino acids; and (3) hundreds, or thousands, of species of *messenger RNA*, probably of variable length, but reaching into the tens of thousands of nucleotides, that furnish the templates for orderly copolymerization of amino acids into specific polypeptides. To the particular purine and pyrimidine base sequence of every one of these multifarious *RNA* species there corresponds some homologous sector of the *DNA* where this sequence recurs in one of the complementary deoxypolynucleotide strands.

It is generally assumed that the *DNA* strand of base sequence complementary to that of the *RNA* transcript acts as the template in transcription, because, in analogy with the mechanism of *DNA* replication, one imagines that formation of the complementary base pairs (Watson & Crick 1953) is responsible for specific alinement of the ribonucleotide monomers. But this point has not yet been firmly established, at least for the actual *in vivo* transcription. It would be prudent, therefore, to keep in mind for the time being that for transcription some other base-pairing mechanism might conceivably obtain. For instance, pairing might also involve identical, rather than complementary, bases (Donohue & Stent 1956), in which case the *DNA* strand of identical base sequence could act as transcription template, or pairing might involve base triplets, rather than pairs (Stent 1958), in which case *both DNA* strands of the homologous sector could act jointly as transcription template.

The RNA–DNA *complex*

During *RNA* synthesis *DNA* template and nascent *RNA* molecule must remain in contact long enough to allow completion of the transcribed unit. What is the nature of the transient complex in which template and nascent transcript are held together? When it was discovered that *DNA* and *RNA* molecules of complementary base sequence can be induced *in vitro* to form deoxy-ribopolynucleotide hybrid molecules (Hall & Spiegelman 1961), it seemed not too far-fatched to suppose that also the purine and pyrimidine bases of nascent *RNA* molecules

* This research was supported by Research Grant CA 02129 from the National Cancer Institute, Public Health Service.

remain hydrogen-bonded to their *DNA* template partners to produce a transient *in vivo* *DNA–RNA* hybrid. The isolation of just such a 'natural' *DNA–RNA* hybrid was reported in 1961 by Spiegelman, Hall & Storck (1961). Here, bacteria infected with ^{32}P-labelled bacteriophages were exposed briefly to ^{3}H-labelled uridine at early stages of intracellular phage growth and then lysed by detergent treatment. The nucleic acids of the lysate, after having been freed from protein, were banded in caesium chloride density gradients. A minor ^{3}H- and ^{32}P-labelled band was observed in the density gradient at a position slightly denser than that of the ^{32}P-labelled band of phage *DNA* and was taken to demonstrate the presence of the *DNA–RNA* 'hybrid'.

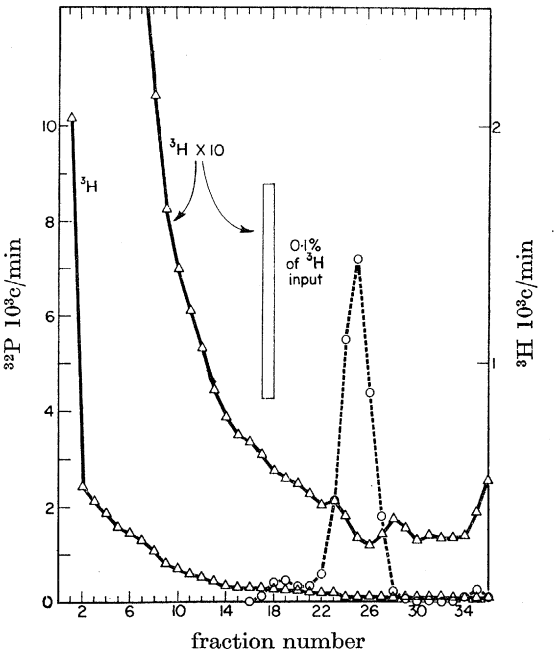

FIGURE 1. Caesium chloride density gradient analysis of deproteinized phage messenger *RNA* (^{3}H label) (△), and parental T4 phage *DNA* (^{32}P label) (○) extracted from phage-infected *E. coli* by detergent lysis 6 min after the onset of intracellular phage growth (incubation temperature: 37 °C). From Konrad & Stent (1964).

An attempt to repeat this experiment (Konrad & Stent 1964) led to the result shown in figure 1. The major band of ^{32}P activity represents the free parental phage *DNA* present in the extract, which bands at its buoyant density of 1·7 g cm^{-3}. Just as in the earlier experiment, a minor ^{32}P-labelled band is found near the slightly higher density of the unlabelled bacterial host *DNA* also present in this extract. It was this minor peak which Spiegelman *et al.* had identified as the *DNA* moiety of the 'natural' hybrid. In contrast to their finding, however, here no distinct class of ^{3}H-labelled *RNA* molecules appears to be associated with this minor *DNA* band. Indeed, the data of figure 1 allow one to estimate that less than 0·05 % (and probably less than 0·01 %) of the *RNA* synthesized during the short 1·5 min exposure of the infected culture to ^{3}H-uridine can be associated with the minor

DNA band. Further experiments then revealed that this minor *DNA* band is only an artifact of density gradient analysis, its ^{32}P activity representing phage *DNA* that is merely *trapped* in the band of unlabelled bacterial host *DNA* (Konrad & Stent 1964). Two main possibilities were then considered for the failure to find the putative *RNA–DNA* complex in this experiment: (*a*) the time required to synthesize individual *RNA* molecules is so short (less than 1 s) that even under brief labelling programmes less than 0·1 % of the ^{3}H-labelled *RNA* is nascent; or (*b*) appreciable quantities of the *RNA* molecules are nascent, but the complex is destroyed during the isolation procedures employed here. The first possibility seemed intrinsically implausible and the following experiments showed that the second possibility is probably correct.

The enzyme '*RNA*-polymerase', responsible for catalysing the ordered copolymerization of ribonucleotides on a *DNA* template, can be isolated from a variety of biological sources, including bacteria (Weiss 1960; Hurwitz, Bresler & Diringer 1960; Stevens 1960; Chamberlin & Berg 1962). If a reaction mixture containing the polymerase, *DNA* molecules as templates and the ribonucleoside 5′-triphosphates of adenine, guanine, uracil and cytosine (*ATP*, *GTP*, *UTP* and *CTP*) is incubated, the nucleotides are polymerized into *RNA*. Since the *RNA* synthesized in this way has the same nearest-neighbour base relations as the template *DNA* (Weiss & Nakamoto 1961; Chamberlin, Baldwin & Berg 1963), and, if the *DNA* is properly double-stranded, the product stimulates cell-free protein synthesis (Ning & Stevens 1962; Wood & Berg 1962), it may be assumed that such *in vitro* synthesis of *RNA* on a *DNA* template is a reasonable model for *in vivo* transcription. Bremer & Konrad (1964) set out to search for the hitherto elusive *RNA–DNA* complex in this relatively simple model system. They allowed *RNA* synthesis to proceed in a reaction mixture containing *RNA*-polymerase, ^{32}P-labelled T-even bacteriophage *DNA*, *ATP*, *GTP*, *UTP* and ^{3}H-labelled *CTP*, and then subjected one sample of the whole reaction mixture to zone sedimentation analysis in a sucrose density gradient. The result of this experiment is presented in figure 2*a*, where it can be seen that the ^{32}P-labelled *DNA* sediments with a velocity of about 40S and that the ^{3}H-labelled *RNA* has a sedimentation distribution whose maximum coincides with that of the *DNA*. (The high ^{3}H activity at the top of the gradient derives from contamination by free ^{3}H-*CTP*, and is not *RNA*). The joint sedimentation of *DNA* and *RNA* in this gradient thus suggests that *the nascent* RNA *is bound to its* DNA *template*. A second sample of the reaction mixture was treated with a protein-denaturing detergent before zone sedimentation analysis, with the result shown in figure 2*b*. It can be seen that detergent treatment produces an important change in the sedimentation profile: whereas the template *DNA* still sediments with a velocity of about 40*S*, the ^{3}H-labelled *RNA* now sediments much more slowly, with an average velocity of only 12*S*. Thus, *denaturation of the protein has freed nascent* RNA *from* DNA *template*, since the two polynucleotide species now sediment independently.

This finding suggests that a protein, namely the *RNA*-polymerase, holds the nascent *RNA* to its *DNA* template. Such an *RNA–DNA*–protein complex could not, of course, have survived the detergent lysis and protein extraction employed

in the experiment of figure 1 in the search for the *in vivo* complex in the phage-infected bacterium, and hence it is small wonder that no complex can be found by such methods. Bremer & Konrad proposed a structure for the complex in which, in contrast to the extensive base pairing of the imagined *RNA–DNA* hybrid, only the few most recently added ribonucleotides constitute the region of contact between *DNA* template and growing ribopolynucleotide chain, as indicated schematically in figure 3.

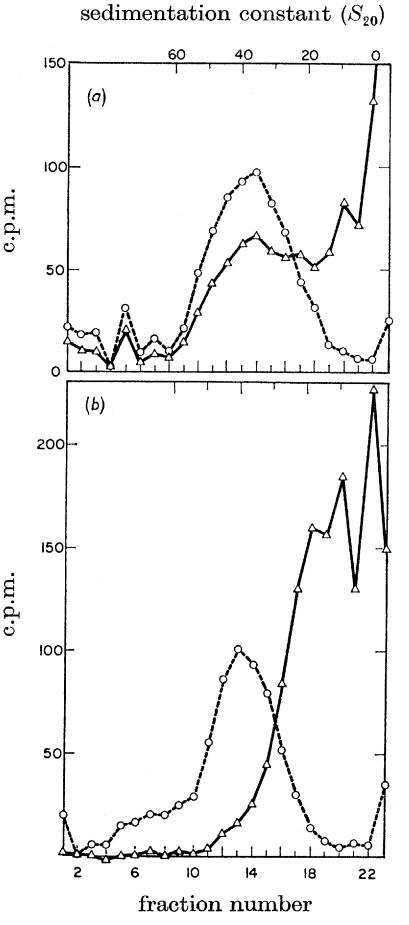

FIGURE 2. Sucrose gradient zone sedimentation analysis 10 min after the onset of *in vitro RNA* synthesis of an *RNA*-polymerase reaction mixture containing ^{32}P-labelled T 4 phage *DNA* as template (○) and ^{3}H-labelled *CTP* as one of the substrates (△). (*a*) The untreated reaction mixture. (*b*) The reaction mixture after detergent treatment. From Bremer & Konrad (1964).

Molecular kinetics of RNA *synthesis*

Figure 4 presents the kinetics of *RNA* synthesis in two similar reaction mixtures, of which one has twice the *RNA*-polymerase concentration of the other. It can be seen that the reaction begins at a constant rate proportional to the enzyme concentration, then slows down after about 20 min and finally comes to a complete halt by about 100 min, the ultimate amount of *RNA* synthesized being proportional to the enzyme concentration. Bremer & Konrad were able to show that these

kinetics reflect the ability of each enzyme molecule to catalyse synthesis of just one *RNA* molecule, which, at the end of the reaction, still remains bound to its *DNA* template. Measurement of the average molecular weight of the *RNA* chains produced after various reaction times allows an estimate of their molecular growth rate. This growth rate turns out to have the unexpectedly low value of about 2 to 3 nucleotides per second.

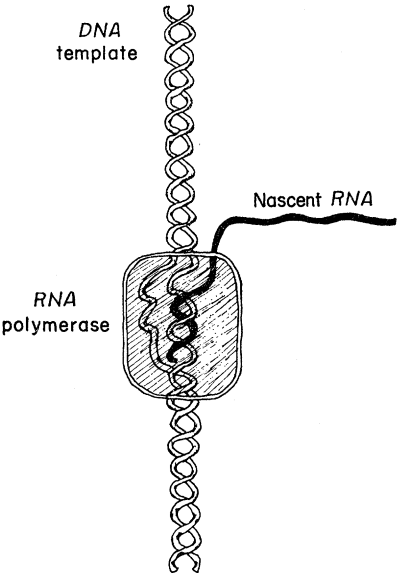

FIGURE 3. Schematic representation of the complex between double-stranded template *DNA*, polymerase enzyme and nascent *RNA*, according to Bremer & Konrad.

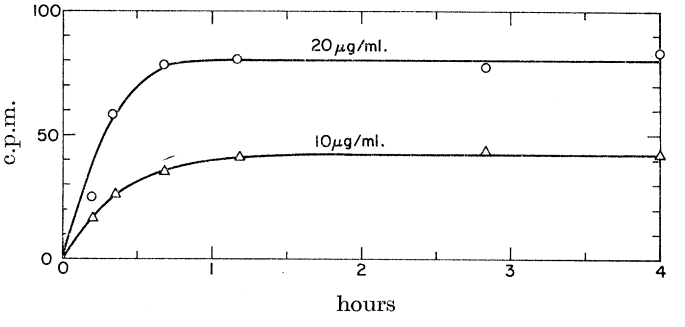

FIGURE 4. Kinetics of *RNA* synthesis at 37 °C in two *RNA*-polymerase reaction mixtures containing ^{14}C-labelled *CTP* as one of the substrates and one of the two concentrations, of polymerase protein indicated on each curve. The ordinate indicates the amount of cold-acid-insoluble radioactivity produced after the reaction time indicated on the abscissa. From Bremer & Konrad (1964).

These observations indicate that the *in vitro* model system for *RNA* synthesis departs from the real *in vivo* situation in three important respects: (1) *in vitro*, each polymerase molecule synthesizes only one *RNA* molecule, whereas *in vivo* it certainly synthesizes many; (2) *in vitro*, the finished *RNA* molecule remains attached to the template, whereas *in vivo* it is certainly quickly liberated from the

RNA–*DNA*–protein complex; and, (3) *in vitro* the rate of growth of the *RNA* molecules is lower, probably by one, or even two, orders of magnitude than what can be guessed to be its *in vivo* rate. It seems not unlikely that all three of these departures from the *in vivo* situation reflect the absence of an essential component from the *in vitro* model system: a process that actively removes the nascent *RNA* transcript from the *DNA*–polymerase complex.

What kind of system might assure *in vivo* removal of the nascent *RNA* from its template? The ribosomes offer themselves as likely agents for this removal

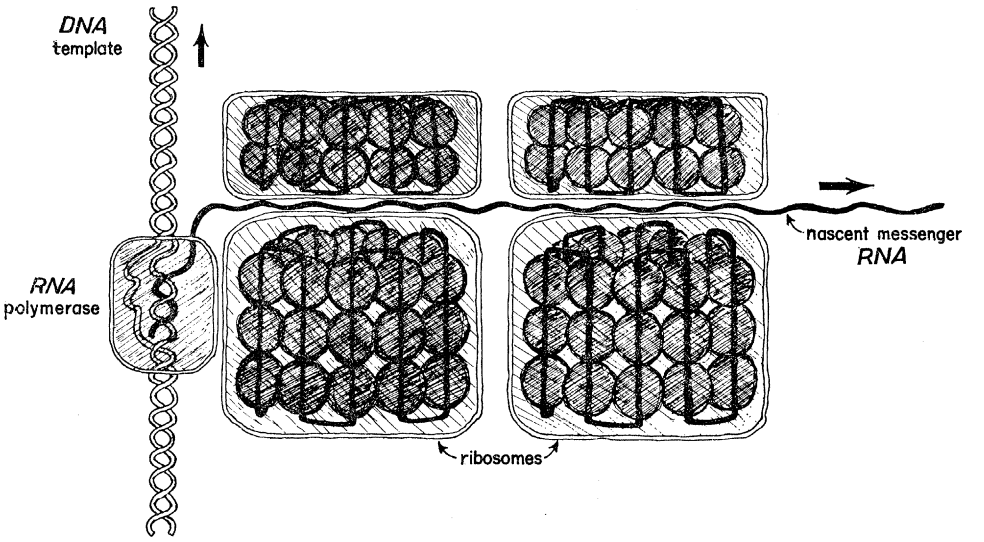

FIGURE 5. Schematic representation of the formation of polyribosomes on nascent messenger *RNA*.

process. In particular, in the case of messenger *RNA*, one can imagine that ribosomes engage the nascent polynucleotide as soon as it emerges from the *RNA*-polymerase and begin forming a polyribosome aggregate before formation of the whole messenger molecule is even complete, as shown schematically in figure 5. Protein synthesis could then commence at once upon the nascent messenger, and the relative motion of messenger and ribosome in protein synthesis, for which the coupled hydrolysis of *GTP* is though to provide the necessary energy (Noll, Staehelin & Wettstein 1963) would then pull nascent messenger away from *DNA*–polymerase complex. Though ribosomes might pull on nascent messenger *RNA* in this way, what mechanisms is one to imagine for removal of the two species of nascent ribosomal *RNA*? Here, one would propose that nascent ribosomal *RNA* also serves as a messenger, in particular as template for synthesis of ribosomal protein. The following considerations can be thought to support the notion that ribosomal protein structure is encoded in ribosomal *RNA*. (1) Although ribosomal *RNA* extracted from mature ribosomes does not appear to stimulate protein synthesis in *in vitro* systems, nascent ribosomal *RNA* does (Otaka, Osawa & Sibatani 1964). (2) In bacteria, ribosomal *RNA* is relatively rich in guanine and ribosomal

protein relatively rich in arginine, glutamic acid, glycine, valine and alanine, i.e. in amino acids whose codon triplets contain two or more guanylic acid residues (Nirenberg *et al.* 1965). (3) Since the triplet code engenders a nucleotide–protein coding ratio of about ten (by weight) and since ribosomes contain equal weights of *RNA* and protein, it would suffice for the production of the ribosomal protein of the cell that the average molecule of nascent ribosomal *RNA* is read about ten times, i.e. has a functional life similar to that of the average bacterial messenger. Thus the nascent ribosomal *RNA* would be removed from its template also by ribosomes, just like any other species of messenger *RNA*. However, unlike ordinary messenger *RNA* species that are eventually destroyed, nascent ribosomal

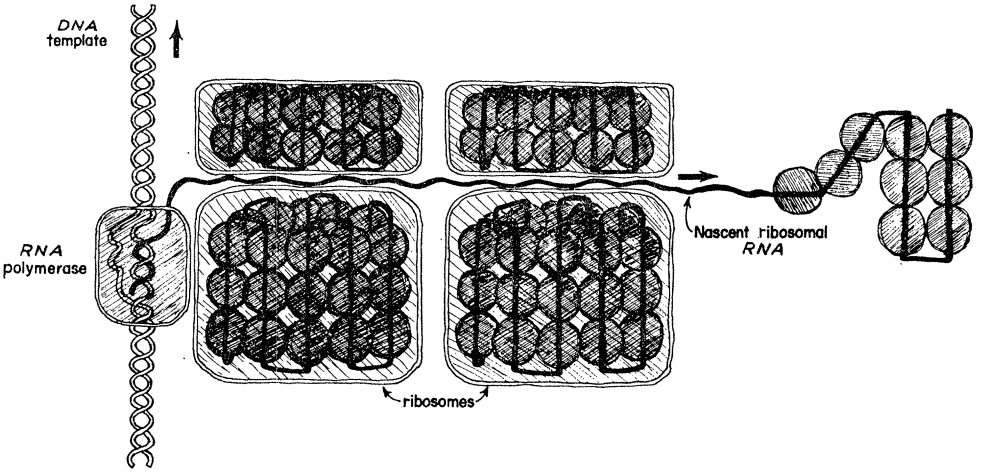

FIGURE 6. Schematic representation of the condensation of skeletal ribosomal protein subunits described by Spirin (1964) with nascent ribosomal *RNA*.

RNA would cease its template function upon condensation with the ribosomal protein subunits (Spirin 1964) to which it has given rise, to form the intact nascent ribosome as shown in figure 6. How, under this picture, transfer *RNA* molecules would be removed from the complex is rather less obvious. It is formally admissible, of course, that, before the characteristic secondary chemical modification of their nucleotides, transfer *RNA* molecules also function as messengers for some, albeit rather short, polypeptides, but this does not seem very plausible. Instead, one could conceive that transfer *RNA* molecules are first synthesized as non-translatable tail ends of long messenger chains, from which they are later cut away by nuclease action. In this way, the pull exerted on the nascent messenger front by its incipient translation would also serve to release the nascent transfer *RNA* segments at the tail.

The view espoused here that ribosomes are involved not only in protein synthesis but also in *RNA* synthesis can draw support from the finding that, in bacteria recovering from severe depletion of their ribosomes, the rate of synthesis of both messenger and ribosomal *RNA* is proportional to the degree to which the intracellular ribosome content has been restored (Kennel & Magasanik 1962; McCarthy 1962; Naono *et al.* 1965).

An RNA–DNA–*ribosome complex*

The model shown in figure 5 thus envisages that there should arise an *in vivo* complex between template *DNA*, *RNA*-polymerase, nascent *RNA* and ribosomes. That such complexes probably exist was shown in an experiment presented in figure 7. Here, bacteria were infected with a T-even bacteriophage and the newly synthesized phage *DNA* was labelled by exposure of the infected culture to ^{32}P-phosphate. After intracellular progeny phages were already present, a large excess of non-radioactive phosphate was added to stop further uptake of ^{32}P and to eliminate as much ^{32}P as possible from the rapidly metabolized phage

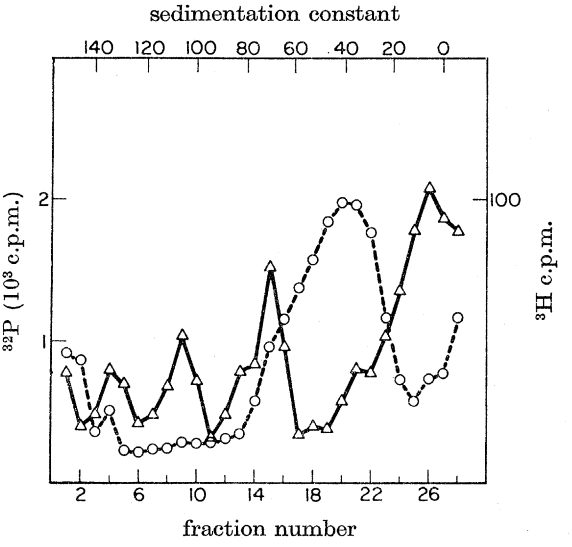

FIGURE 7. Sucrose gradient zone sedimentation analysis of (△) messenger *RNA* (^{3}H label) and (○) T4 phage *DNA* (^{32}P label) recovered from *E. coli* infected at a multiplicity of 4·5 phage/cell and lysed by freezing and thawing 50 min after infection (incubation temperature: 30 °C). $^{32}PO_4$ was added to the culture 14 min after infection and the specific $^{32}PO_4$ activity of the medium was reduced 20-fold by addition of excess $^{32}PO_4$ 44 min after infection. From an unpublished experiment of Bremer, Konrad & Stent.

messenger *RNA*. After a few more minutes, ^{3}H-uridine was added to the culture medium as a specific messenger *RNA* label for 5 s before the culture was lysed by freezing and thawing. The resulting lysate was subjected *directly* to zone sedimentation in a sucrose density gradient, that is without detergent treatment or protein extraction. As can be seen in figure 7, the resulting sedimentation profile contains three major ^{3}H peaks corresponding to sedimentation constants of 70*S*, 105*S* and 130*S*, representing, no doubt, messenger *RNA* attached to ribosomes and polyribosomes. The ^{32}P sedimentation profile shows one major peak, corresponding to the rate of sedimentation of free phage *DNA* molecules. But at least 10 % of the ^{32}P-labelled *DNA* is seen to sediment at velocities much greater than that of the major peak of free phage *DNA* and moves into the zone of ribosomes and polyribosomes. This minor fraction of rapidly sedimenting phage *DNA*, which, since it is resistant to alkaline hydrolysis, cannot represent residual ^{32}P-labelled

messenger-*RNA* molecules, is, most likely, the *DNA* moiety of an intracellular *DNA*–polymerase–*RNA*–ribosome complex.

An analogous complex was identified also by Byrne, Levin, Bladen & Nirenberg (1964) in an *in vitro* system, containing, as the reaction mixture of the experiment of figure 2, radioactively labelled T-even phage *DNA*, *RNA*-polymerase, *ATP*, *GTP*, *UTP*, and *CTP*, and, in addition, ribosomes, a 'soluble' bacterial extract, and the substrates that allow protein synthesis by use of the *RNA* produced in the reaction. Upon subjecting this reaction mixture to sucrose density gradient zone sedimentation, it was found that the template *DNA* of the mixture sediments at a much greater velocity than the free phage *DNA*, suggesting that it exists in a complex with ribosomes and polyribosomes. Since formation of this complex is dependent on the synthesis of *RNA* and on the presence of the 'soluble' extract (containing the transfer-*RNA* species necessary for protein synthesis), and since treatment with ribonuclease destroys the complex, it can be inferred that the complex is not a simple direct aggregate between *DNA* and ribosomes. On the basis of their findings, Byrne *et al.* suggested a model very similar to that depicted in figure 6, namely that there exists 'an intermediate stage in protein synthesis in which *DNA* is joined to functionally competent ribosomes by means of [messenger] *RNA*.'

Asymmetry and starting points of transcription

If double-stranded *DNA* serves as template for *RNA* synthesis through formation of base pairs between single template polynucleotide strands and ribonucleotide monomers, then it seems most probable that in the domain of any operon, the supposed genetic unit incorporated into a single messenger molecule (Martin 1963), only one of the two *DNA* strands is actually transcribed. For it seems well-nigh inconceivable that the cistronic codon sequence forms a sophisticated palindrome that spells out the same polypeptide amino acid sequence whether one *DNA* strand is read forwards or the other *DNA* strand is read backwards. Indeed, a variety of experiments bears out this expectation. The more convincing of these (Greenspan & Marmur 1963; Hayashi, Hayashi & Spiegelman 1963*a*; Tocchini-Valentini *et al.* 1963) involve the intracellular synthesis of radioactively labelled messenger *RNA* by bacteriophages which either carry only a single-stranded *DNA* in their extracellular, infective form and produce a double-stranded replicative *DNA* in their vegetative form, or carry in both infective and vegetative forms a double-stranded *DNA* composed of two complementary but readily resolvable strands of unequal buoyant density. The results of these experiments agree in showing that the viral messenger *RNA* extracted from the infected bacterium can be induced to form molecular deoxy-ribopolynucleotide hybrids with—i.e. is complementary in base sequence to—only one of the two viral *DNA* strands. One may now inquire into the mechanism by which such asymmetric transcription proceeds: how does the *RNA*-polymerase select the *DNA* strand on which it is to work?

This problem seemed all the more puzzling at first when it was found that in its *in vitro* action the *RNA* polymerase appears to be symmetric—i.e. does not seem to discriminate between the two complementary strands of the viral *DNA* and produces an *RNA* product complementary in base sequence to *both* template strands

(Chamberlin & Berg 1962; Geiduschek, Nakamoto & Weiss 1961; Hayashi, Hayashi & Spiegelman 1963*b*). More recent observations make it appear, however, that asymmetric transcription does proceed also *in vitro*, provided that the physical integrity of the double-stranded *DNA* template has been conserved. The first such study was made by Hayashi, Hayashi & Spiegelman (1964), who found that use of the circular, double-stranded replicative *DNA* of bacteriophage ØX 174 as template for *in vitro* action of *RNA*-polymerase produces asymmetric transcription, the asymmetry being of the same nature as that which occurs *in vivo*, in the infected bacterium. Hayashi *et al.* showed, furthermore, that symmetric *in vitro* transcription proceeds after sonic degradation of the template *DNA*, and concluded that *circularity* of the template is the factor responsible for asymmetric transcription. Later experiments (Geiduschek, Tochini-Valentini & Sarnat 1964; Green 1964) showed, however, that asymmetric *in vitro* transcription can be achieved also by non-circular viral *DNA*, even after some fragmentation of the template molecules. But sonic disruption, or heat denaturation, of the template *DNA* engenders symmetric transcription. Thus it would follow that it is not the circularity of the replicative forms of viral *DNA* that assures selection of the proper template strand to be transcribed but rather some feature of the native *DNA* double helix.

A problem closely related to strand selection of transcription is that of the nature of its starting signals. For, unless the whole genome is first transcribed as one giant continuous *RNA* molecule that is only later cut into appropriate operon segments, there must exist starting points on the *DNA* template at which segmental transcription is to commence. Bremer, Konrad & Bruner (1966) have attempted to find evidence for such specific starting points in the *in vitro* action of *E. coli RNA*-polymerase on T-even phage *DNA* as template. For this purpose, they estimated the upper limit of the number of *RNA* molecules that can grow simultaneously on a single T-even phage *DNA* complement by measuring the dependence of the initial rate of *in vitro RNA* synthesis at a fixed polymerase concentration on the number of *DNA* molecules added to the reaction mixture. As long as the number of polymerase molecules in the mixture is in great excess, maximal use should be made of all the *DNA* template molecules. And, knowing the chain growth rate from their studies already described, the number of chains growing per *DNA* molecules can be readily calculated from the observed overall rate of polymerization. Such calculations lead to a limit of about 180 *RNA* molecules capable of being initiated per T-even phage *DNA* complement. That no more than 180 *RNA* chains can be initiated suggests that the number of possible polymerase starting points on the T-even phage *DNA* is limited, since, if the only limitation on the number of growing *RNA* molecules were the close-packing of the polymerase protein on the *DNA* template, it should be possible to achieve simultaneous growth of about 5000 *RNA* molecules per phage *DNA* complement. (The figure of 5000 molecules is reckoned on the basis of a total length of 50 μm for the T-even phage *DNA* molecule and an approximate diameter of 0·01 μm for the polymerase molecule.) One possible explanation for the limited number of starting points would be that the polymerase can start only at the two physical ends of the *DNA* template, where its attachment can actually be demonstrated (Berg *et al.*

1965), so that at the onset of the reaction hundreds of polymerase molecules would begin working in tandem at each template end. This explanation is quite unlikely, however, since fragmentation of the phage *DNA*, and thus creation of additional ends, does not generate additional starting points.

Similar studies with heat-denatured, i.e. largely single-stranded, T-even phage *DNA* lend further support to the conclusion that there exists only a limited number of starting points on native, double-stranded *DNA*. First of all, the number of binding sites for *RNA*-polymerase per *DNA* nucleotide unit is greater on single-stranded *DNA* than on native, double-stranded *DNA* (Hurwitz *et al.* 1962; Fox & Weiss 1964; Wood & Berg 1964). Secondly, Bremer, Konrad & Bruner (1966) found that the maximum number of growing *RNA* molecules per T-even phage complement is about 6 times greater for heat-denatured than for native phage *DNA* (though the rate of chain growth on denatured is much smaller than on native *DNA*) and begins to approach the theoretical upper limit to be expected from close packing of polymerase protein on the *DNA* template. It can be supposed, therefore, that the polymerase recognizes specific starting points only on native double-stranded but not on single-stranded *DNA*.

Just as for strand selection, so also for starting points, one reaches the conclusion that the specificity of attack by the polymerase depends on the double-stranded character of the template *DNA* (Berg *et al.* 1965). One obvious possible explanation of both specificities is that the *RNA*-polymerase attacks *only* single-stranded or denatured regions of the double-stranded template (Chamberlin & Berg 1964), in which one of the strands has been rendered accessible to the polymerase by some molecular singularity. This starting-point singularity could represent interruptions in the phosphate diester bonds of one of the polynucleotide chains, or particular base sequences or chemical modifications of the purine or pyrimidine rings that favour spontaneous 'melting out' of the inter-strand hydrogen bonds.

Chemical direction of RNA *synthesis*

The model of figure 5 suggests the following question: can synthesis of polypeptides encoded in the 'head' of the messenger transcribed first begin before the messenger 'tail' has even been formed, so that, by the time messenger synthesis is complete, one molecule of every polypeptide inscribed into the unit of transcription has already been formed? In order that the anwer to this question be 'yes', it is necessary that the direction of *RNA* synthesis be the same as the direction in which the messenger *RNA* is translated by transfer *RNA* into an ordered polypeptide assembly of amino acids. For if transfer *RNA* were to translate the purine–pyrimidine base sequence of the messenger *RNA* in a direction opposite from that in which the nucleotides are themselves assembled on the *DNA* template, synthesis of the nascent molecule would have to be complete before it could begin to direct protein synthesis.

The following experiments are relevant to the first of these processes, the chemical direction of *RNA* synthesis (Bremer, Konrad, Gaines & Stent 1965). In the *RNA*-polymerase-catalysed polymerization of the ribonucleoside 5′-triphosphates *ATP*, *GTP*, *UTP* and *CTP* into *RNA*, the 5′-α-phosphate of one ribonucleotide is

esterified to the 3′-hydroxyl group of the ribose moiety of another, with attendant liberation of the 5′-β-γ-phosphates as inorganic pyrophosphate, for formation of the phosphate diester bonds responsible for the continuity of the ribopolynucleotide chain. One may now ask whether in this growth process the free 3′-hydroxyl of a monomeric nucleotide attacks the terminal 5′-triphosphate of the last nucleotide to have been incorporated into the growing chain (in which case the chain is said to grow from its 3′ end), or whether the 5′-triphosphate of a monomeric nucleotide attacks the terminal free 3′-hydroxyl of the last nucleotide to have been incorporated into the growing chain (in which case the chain is said to grow from its 5′ end). These two possibilities are illustrated diagrammatically in figure 8.

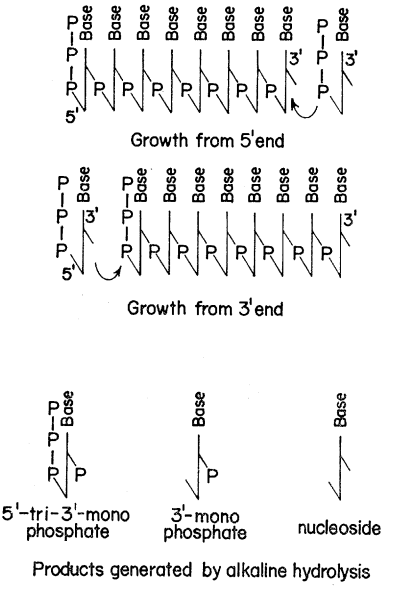

FIGURE 8. Two possibilities of the direction of chain growth in *RNA* synthesis and the expected three products of alkaline hydrolysis of the ribopolynucleotide product.

Under either of these two alternatives, alkaline hydrolysis of nascent *RNA*, which splits the internucleotide phosphate diester bonds between the phosphate and the 5′-hydroxyl group, should generate the three products shown in figure 8: a 3′- or 2′-monophosphate, representing all the nucleotides from the body of the nascent molecules, as well as a 5′-tri-3′-monophosphate (Maitra *et al.* 1965) and a nucleoside, representing the two ends of the nascent molecule. The problem of the direction of growth can then be resolved by ascertaining which of the two kinds of terminal residues comes from the starting, and which from the growing, end.

Figure 9 presents the results of an experiment in which 8 min after the start of reaction a sample was removed from an *RNA*-polymerizing mixture containing ^{3}H-labelled *ATP* from the beginning. The polymeric product of the sample was separated from unincorporated ^{3}H-*ATP* substrate by repeated acid precipitations and subjected to alkaline hydrolysis. The hydrolysis products were then resolved by paper electrophoresis. The ordinate of figure 9 indicates (on a logarithmic scale)

the ^3H activity of successive segments of the electrophoresis paper, whose number is registered on the abscissa. The electrophoretic distribution shows one major peak that contains more than 99 % of the total radioactivity; this peak represents the adenosine 3′- (and 2′-) monophosphate (*Ap*) from the body of the polymer. In addition, three minor peaks are manifest. The left-most of these peaks, on the negative pole side of the origin at which the hydrolysate had been applied to the paper, represents adenosine (*A*), and the right-most peak that migrated fastest to the positive pole appears to be adenosine 5′-tri-3′-monophosphate (*pppAp*). The third minor peak is *ATP* (*pppA*) and represents unincorporated substrate carried over as contaminant from the reaction mixture.

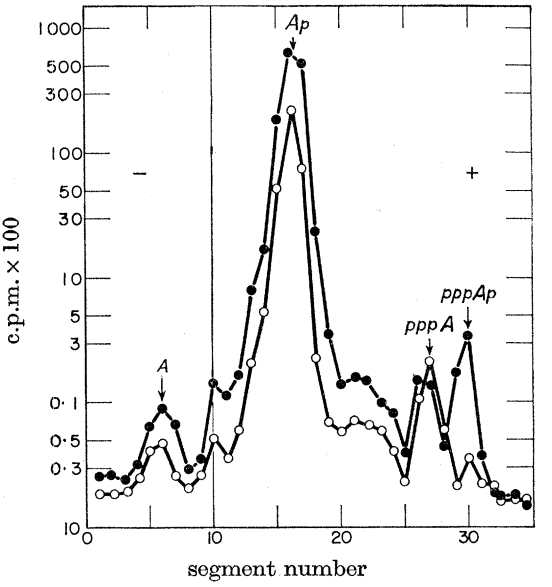

FIGURE 9. Paper electrophoresis of the products of alkaline hydrolysis of two *RNA*–polymerase reaction mixtures containing T4 phage *DNA* as template. One mixture contained ^3H-labelled *ATP* substrate from the beginning (●) and was incubated for 8 min. To the other mixture (○), which initially contained unlabelled *ATP*, ^3H-labelled *ATP* substrate was added after the 8th minute of reaction and the reaction was then allowed to progress for 2 more minutes. The positions of radioactivity peaks corresponding to adenosine, *Ap* and *pppA* were identified by u.v. absorption of the corresponding unlabelled carrier compounds added to the hydrolysates before electrophoresis. The identification of the right-most peak as *pppAp* is inferential. From Bremer, Konrad, Gaines & Stent (1965).

In order to ascertain whether adenosine or *pppAp* represents the growing end of the *RNA* molecule, a second *RNA*-polymerizing reaction mixture, initially containing no labelled *ATP*, was incubated for 8 min and only after that time ^3H-labelled *ATP* was introduced. The reaction was then allowed to proceed for two more minutes, after which a sample of the mixture was purified, hydrolysed, and subjected to electrophoresis, exactly as the sample of the first mixture. The result is also presented in figure 9. It can be seen, first of all, that the major, or *Ap*, peak of the second sample contains about one-fourth as much activity as that

of the first sample, in agreement with the relative amounts of ^{3}H-labelled *RNA* expected to be produced in the two reaction mixtures. Secondly, the activity in the minor *pppA* peak is evidently the same in both samples, in agreement with the assertion that this peak is derived from unincorporated ^{3}H-*ATP* substrate contaminating the hydrolysate. Thirdly, the activity in the left-most, or adenosine, peak of the second sample is about half that of the first sample. Fourthly, and most important, the activity in the right-most, or *pppAp*, peak of the second sample is less than one-tenth that of the first sample. The much smaller activity of this peak in the second sample can mean only that the *pppAp* nucleotide represents the starting ends of the *RNA* molecule. For in the second reaction mixture all *RNA* molecules were started with non-labelled *ATP* during the first 8 min of the reaction and had thus little opportunity to incorporate any label into their starting end during the last 2 min, after addition of ^{3}H-*ATP*. Conversely, the significant labelling of the adenosine peak in the second sample must mean that adenosine represents the gowing end of the *RNA* molecule. That only half as much label was incorporated into the growing adenosine ends in the second reaction mixture implies, however, that about half of the *RNA* molecules originally started have already stopped growing by the 8th minute of the reaction, and hence can no longer be labelled when ^{3}H-*ATP* is added to the substrates after that point. *It can be concluded, therefore, that the* RNA *chain is growing from its 5′ end.*

It is apparent from the data of figure 9 that, in the sample of the first reaction mixture that contained ^{3}H-*ATP* from the beginning, the chain-initial *pppAp* peak contains three to four times as much label as the chain-terminal adenosine peak. This finding leads to an unexpected conclusion, supported also by other data not shown here: the *in vitro* growth of *RNA* chains is initiated predominantly by adenosine nucleotides, whereas the terminal nucleotide of the chains at any stage of molecular growth may be any one of the four nucleotides. This preference for *RNA* chain initiation by *ATP* could be an artifact of the *in vitro* system; but it might also reflect something fundamental about the nature of the specific starting points for *RNA*-polymerase attack on the *DNA* template. Finally, the ratio of ^{3}H activities in the major *Ap* peak and minor *pppAp* peak allows an estimate of the average chain length of the polymer produced (bearing in mind that most chains originate with *pppAp* but that, on the basis of the overall nucleotide composition of the product *RNA*, only every third nucleotide in the body of the polymer is an *Ap*). This estimate leads to an approximate length of 1000 nucleotides per chain after 8 min of reaction time, or to a growth rate of about 2 nucleotides per second, in good agreement with the previous estimate of the molecular growth rate.

At present, the direction in which the messenger *RNA* is translated is still not known with certainty, although some more or less indirect experiments have been reported that seem to favour a direction of translation from the 3′ end. In one type of such experiments (Cramer, Küntzel & Matthaei 1964), a synthetic polyribouridylic–guanylic copolymer rich in uridylic acid was hydrolysed by means of specific nucleases that generate oligonucleotides of the type ... *pUpUpUpG*. It was reported that such guanylic-terminal oligonucleotides stimulate specific binding of leucyl but not phenylalanyl transfer *RNA* to ribosomes. Since one of the leucyl

transfer-*RNA* codons is *UUG*, and the phenylalanyl transfer-*RNA* codon is *UUU*, it was inferred that translation proceeds from the 3′ rather than the 5′ terminal end of the 'messenger' oligonucleotide.

A second, more directly relevant type of experiment (Eikenberry & Rich 1965; Williamson & Schweet 1965) that leads to the same conclusion on the direction of messenger translation involves the *in vitro* synthesis of haemoglobin by polyribosomes isolated from reticulocytes. Here digestion of the messenger *RNA* attached to the polyribosomes by two different specific exonucleases, one working from the 3′ end and the other from the 5′ end of the polyribonucleotide, produces different effects: destruction of the 5′ end of the messenger *RNA* appears to prevent further synthesis of those parts of the haemoglobin molecule that are near the carboxy-terminal end of the polypeptide known to be synthesized last, whereas destruction of the 3′ end of the messenger *RNA* prevents further initiation of new polypeptide chains at the amino-terminal end of the polypeptide known to be synthesized first, while still allowing completion of the carboxy-terminal end.

Thus taken together, *in vitro* studies on messenger synthesis and translation lead to the conclusion that these two processes proceed in *opposite* directions. Hence on the basis of these results, the answer to the question posed earlier seems to be 'no', since synthesis of polypeptides encoded in a unit of transcription could not begin before its entire messenger molecule has been formed. This answer is rather disappointing; if operons *were* transcribed and translated in the *same* direction, it would be much easier to understand the apparent regulatory connexion that exists between messenger formation and messenger translation, as the following considerations show.

Control of transcription and translation

In their model of the genetic regulation of protein synthesis, Jacob & Monod (1961) favoured the view that primary control of formation of specific enzymes occurs at the level of messenger synthesis. They proposed that interaction of a repressor molecule with an operator region of the *DNA* inhibits formation of messenger *RNA*, and hence transcription and, *a fortiori*, translation of the operon. Subsequent experiments showed, in agreement with this notion, that growth of bacteria in the presence of specific inducing effectors greatly increases the quantity of messenger *RNA* capable of forming hybrids with a homologous *DNA* fraction carrying the relevant operon (Attardi *et al.* 1963). This theory suggests, and later studies, when interpreted on this basis, support, that the transcription of operons proceeds from their operator end (Jacob, Ullman & Monod 1964), whereas a variety of physiologico-genetic data, particularly a recent study (Alpers & Tomkins 1965), strongly favour the view that operons are also translated from their operator end. Thus, here transcription and translation would proceed in the *same*, rather than in opposite, directions.

Opposite directions of transcription and translation present even greater difficulties if, in contrast to this model, primary control of synthesis of specific proteins occurs at the level of translation rather than transcription, a possibility that is by no means ruled out (Stent 1964). Here, cause and effect of variations in

intracellular concentration of specific messengers and enzymes are reversed, in that a feedback loop would exist between messenger synthesis and function: the faster the messenger can work in protein synthesis, the faster it is synthesized. Now if transcription and translation were to proceed in the *same* direction, the model of figure 5 would provide a ready explanation of just such a feedback loop. Function of the polymerase would depend on continuous removal of the nascent *RNA* by relative motion of messenger and ribosomes in protein synthesis. Synthesis of a particular messenger species would decelerate as soon as a codon sequence appears in the nascent messenger for which no corresponding transfer *RNA* is available. Thus, if as outlined in a recent proposal (Stent 1964), the regulation of protein synthesis proceeds by the control of availability of special, or 'modulating' transfer-*RNA* species corresponding to special modulating codons that concern degenerate representations of certain amino acids, synthesis of any messenger *RNA* whose cistrons are encumbered with one or more modulating codons could proceed only as long as the corresponding modulating transfer-*RNA* species are available.

It might be prudent, therefore, to reserve final judgement on the *in vivo* significance of the depressing results of *in vitro* studies that indicate opposite directions of transcription and translation, and meanwhile take courage in Eddington's dictum that, 10 years ago, stood another contributor (Wollman, Jacob & Hayes 1956) to this Symposium in good stead: '...it is also a good rule not to put overmuch confidence in the observational results that are put forward until they have been confirmed by theory.'

I am indebted to my colleagues, H. Bremer, J. C. Gerhardt and F. O. Wettstein, for stimulating discussions during the preparation of this essay.

Postscript: Within a month of the presentation of this paper, two independent sets of studies were completed that seem to leave little doubt that translation proceeds from the 5′ end of the messenger polynucleotide, and not from its 3′ end as the experiments on transfer-*RNA* binding by oligonucleotides and on haemoglobin synthesis after messenger degradation had indicated. One of these studies concerns the amino acid sequence determination of oligopeptides whose *in vitro* synthesis is directed by synthetic oligonucleotides of known nucleotide sequence (Salas *et al.* 1965). The other study represents the analysis of the effects of mutational reading frame shifts in a phage *DNA* cistron on the amino acid sequence of the phage protein coded by that cistron (Terzaghi *et al.* 1965). Hence, it now appears that synthesis and translation of *RNA* proceed in the *same* chemical direction and that a necessary condition for the model of figure 5 is actually met.

References (Stent)

Alpers, D. H. & Tomkins, G. M. 1965 *Proc. Nat. Acad. Sci., Wash.* **53**, 557.

Attardi, G., Naono, S., Rouviere, J., Jacob, F. & Gros, F. 1963 *Cold Spr. Harb. Symp. Quant. Biol.* **28**, 363.

Berg, P., Kornberg, R. D., Fancher, H. & Dieckmann, M. 1965 *Biochem. Biophys. Res. Commun.* **18**, 932.

Bremer, H. & Konrad, M. W. 1964 *Proc. Nat. Acad. Sci., Wash.* **51**, 801.
Bremer, H., Konrad, M. W. & Bruner, R. 1966 *J. Mol. Biol.* **14** (in the Press).
Bremer, H., Konrad, M. W., Gaines, K. & Stent, G. S. 1965 *J. Mol. Biol.* **13**, 540.
Byrne, R., Levin, J. G., Bladen, H. A. & Nirenberg, M. W. 1964 *Proc. Nat. Acad. Sci., Wash.* **52**, 140.
Chamberlin, M., Baldwin, R. L. & Berg, P. 1963 *J. Mol. Biol.* **7**, 334.
Chamberlin, M. & Berg, P. 1962 *Proc. Nat. Acad. Sci., Wash.* **48**, 81.
Chamberlin, M. & Berg, P. 1964 *J. Mol. Biol.* **8**, 708.
Cramer, F., Küntzel, H. & Matthaei, J. H. 1964 *Angew. Chem.* (int. ed.), **3**, 589.
Donohue, J. & Stent, G. S. 1956 *Proc. Nat. Acad. Sci., Wash.* **42**, 734.
Eikenberry, E. F. & Rich, A. 1965 *Proc. Nat. Acad. Sci., Wash.* **53**, 668.
Fox, C. F. & Weiss, S. B. 1964 *J. Biol. Chem.* **239**, 175.
Geiduschek, E. P., Nakamoto, T. & Weiss, S. B. 1961 *Proc. Nat. Acad. Sci., Wash.* **47**, 1405.
Geiduschek, E. P., Tocchini-Valentini, G. P. & Sarnat, M. T. 1964 *Prov. Nat. Acad. Sci., Wash.* **52**, 486.
Green, M. H. 1964 *Proc. Nat. Acad. Sci., Wash.* **52**, 1388.
Greenspan, C. & Marmur, J. 1963 *Science*, **142**, 387.
Hall, B. D. & Spiegelman, S. 1961 *Proc. Nat. Acad. Sci., Wash.* **47**, 137.
Hayashi, M., Hayashi, M. N. & Spiegelman, S. 1963*a* *Proc. Nat. Acad. Sci., Wash.* **50**, 664.
Hayashi, M., Hayashi, M. N. & Spiegelman, S. 1963*b* *Science*, **140**, 1313.
Hayashi, M., Hayashi, M. N. & Spiegelman, S. 1964 *Proc. Nat. Acad. Sci., Wash.* **51**, 351.
Hurwitz, J., Bresler, A. & Diringer, R. 1960 *Biochem. Biophys. Res. Commun.* **3**, 15.
Hurwitz, J., Furth, J. J., Anders, M. & Evans, A. 1962 *J. Biol. Chem.* **237**, 3752.
Jacob, F. & Monod, J. 1961 *J. Mol. Biol.* **3**, 318.
Jacob, F., Ullman, A. & Monod, J. 1964 *C.R. Acad. Sci., Paris*, **258**, 3125.
Kennel, D. E. & Magasanik, B. 1962 *Biochim. Biophys. Acta*, **55**, 139.
Konrad, M. W. & Stent, G. S. 1964 *Proc. Nat. Acad. Sci., Wash.* **51**, 647.
McCarthy, B. J. 1962 *Biochim. Biophys. Acta*, **55**, 880.
Maitra, U., Novogrodsky, A., Baltimore, D. & Hurwitz, J. 1965 *Biochem. Biophys. Res. Commun.* **18**, 801.
Martin, R. G. 1963 *Cold Spr. Harb. Symp. Quant. Biol.* **28**, 357.
Naono, S., Rouvière, J., Schedlovsky, A. & Gros, F. 1965 (In the Press.)
Ning, C. & Stevens, A. 1962 *J. Mol. Biol.* **5**, 650.
Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F. & O'Neal, C. 1965 *Proc. Natl. Acad. Sci. Wash.* **53**, 1161.
Noll, H., Staehelin, T. & Wettstein, F. O. 1963 *Nature, Lond.* **198**, 632.
Otaka, E., Osawa, S. & Sibatani, A. 1964 *Biochem. Biophys. Res. Commun.* **15**, 568.
Salas, M., Smith, M., Stanley, W. M. Jr., Wahba, A. & Ochoa, S. 1965 *J. Biol. Chem.* (In the Press.)
Spiegelman, S., Hall, D. & Storck, R. 1961 *Proc. Nat. Acad. Sci., Wash.* **47**, 1135.
Spirin, A. S. 1964 *Macromolecular structure of ribonucleic acids*, p. 190. New York: Reinhold Publishing Corp.
Stent, G. S. 1958 *Advances in virus research*, **5**, 95.
Stent, G. S. 1964 *Science*, **144**, 816.
Stevens, A. 1960 *Biochem. Biophys. Res. Commun.* **3**, 92.
Terzaghi, E., Okada, Y., Streisinger, G., Inoue, H. & Tsugita, A. 1965 (In the Press.)
Tocchini-Valentini, G. P., Stodolsky, M., Aurisicchio, A., Sarnat, M., Graziosi, F., Weiss, S. B. & Geiduschek, E. P. 1963 *Proc. Nat. Acad. Sci., Wash.* **50**, 935.
Watson, J. D. & Crick, F. H. C. 1953 *Cold Spr. Harb. Symp. Quant. Biol.* **18**, 123.
Weiss, S. B. 1960 *Proc. Nat. Acad. Sci., Wash.* **46**, 1020.
Weiss, S. B. & Nakamoto, T. 1961 *Proc. Nat. Acad. Sci., Wash.* **47**, 1400.
Williamson, A. R. & Schweet, R. 1965 *Nature, Lond.* **206**, 29.
Wood, W. B. & Berg, P. 1962 *Proc. Nat. Acad. Sci., Wash.* **48**, 94.
Wood, W. B. & Berg, P. 1964 *J. Mol. Biol.* **9**, 452.
Wollman, E. L., Jacob, F. & Hayes, W. 1956 *Cold Spr. Harb. Symp. Quant. Biol.* **21**, 141.