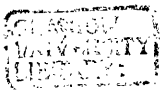
THE GENES FOR MOUSE GLOBIN AND THE POST-TRANSCRIPTIONAL CONTROL OF THEIR EXPRESSION

Stephen E. Humphries

A Thesis submitted for the Degree of Doctor of Philosophy in the Faculty of Medicine, The University of Glasgow.

Beatson Institute for Cancer Research,
Glasgow. October, 1974.



CONTENTS

ACKNOWLEDGMENTS

During the course of this work, my supervisor Dr. Bob Williamson has given me a great deal of encouragement and advice, and I thank him for this, as well as for his example.

Also over the past three years I have received help with many aspects of this work from all members of the Beatson Institute. For this I thank them sincerely, especially the Director, Dr. J. Paul, and my collaborators, Drs. G. D. Birnie, R. S. Gilmour, P. R. Harrison, A. Hell and B. D. Young. I would specifically like to thank Mrs. J. Crossley for help in running the fingerprints, Dr. I. B. Pragnell for assistance with the Krebs ascites cell-free system, my collaborator Dr. M. Doel for his help with the rabbit cell-free system, and Dr. M. Getz for advice in using polynucleotide phosphorylase, and for many helpful discussions. I also acknowledge and thank the assistance and cheerful co-operation of Mrs. Alice Walker and her animal house staff, Mrs. June Peffer, who typed this manuscript, and Susie, my wife, whose support and understanding has been invaluable to me.

SUMMARY

In this work I have investigated the possible role of gene reiteration, amplification or deletion, as a means of regulating the expression of the globin gene during tissue differentiation in the mouse. This has employed a probe of complementary DNA (cDNA) prepared by copying reticulocyte 9S RNA using avian myeloblastosis virus reverse transcriptase. Since this RNA contains the messenger RNAs for α and β globin, the cDNA produced is complementary to the α and β globin genes. The fidelity of this cDNA and its use as a probe for globin sequences in DNA and RNA are rigorously justified. The rates of hybridisation of this probe to whole mouse embryo DNA, mouse erythroid DNA from 14 day foetal livers, and mouse germ line DNA from sperm are compared. The small corrections required for the size differences of the hybridising molecules, and the slight mismatching in the cDNA are estimated.

The study demonstrates that there are 1 – 2 copies of the globin α and β genes in all tissues studied, and thus mechanisms other than gene amplification or deletion are required to explain the control of globin gene expression.

The existence of post–transcriptional mechanisms by which this regulation may occur was then investigated. The first of these is the translational role of the poly(A) segment found on the 3' end of eukaryotic mRNAs. This segment was removed from reticulocyte 9S RNA, using polynucleotide phosphorylase. Several experiments were performed to demonstrate the absence of the poly(A) sequence from the mRNA. The ability of this mRNA to code for α and β globin chain synthesis in several cell–free protein synthesising systems

was compared with intact mRNA. The results demonstrate that the poly(A) segment is not required for efficient initiation, elongation, termination or reinitiation of protein synthesis in a cell-free system.

The role of the poly(A) sequence in determining the nuclear metabolism of globin RNA was then investigated. Low levels of globin RNA sequences were detected using cDNA in the nucleus of several non-erythroid tissues, adult mouse brain and liver, and a cultured lymphoma cell line, L5178Y. Very much lower levels of globin RNA sequences were detected in the cytoplasm of these cells. In contrast, in the erythroid 14 day foetal mouse liver, the amounts of globin RNA sequences found in the cytoplasm were very much higher than that in the nucleus. This indicates the operation of a post-transcriptional regulatory mechanism, whereby the nuclear and cytoplasmic levels of globin RNA sequences are regulated independently. However, this mechanism appears not to be mediated through association of nuclear globin RNA sequences with poly(A) since in both erythroid and non-erythroid tissues, where the fraction of total cell globin RNA sequences in the cytoplasm varies considerably, the proportion of polyadenylated nuclear globin RNA sequences is the same.

At the present time, the possible functions of this low-level transcription of the globin gene in non-erythroid tissues are unclear.

INTRODUCTION

INTRODUCTION

1. In recent years, the isolation of mRNAs coding for specific proteins has enabled considerable advances to be made in the understanding of the control of eukaryotic gene expression. Using these probes, it has been possible to examine the arrangement of specific gene sequences in the DNA, and to study the control of nuclear and cytoplasmic RNA metabolism, during cellular differentiation. In these ways, it is hoped to gain an insight into the molecular events in pathological disorders such as cancer or viral infections, as well as inherited defects such as mongolism or thalassaemia.

2. Control of gene expression at the DNA level

One possible mechanism by which the expression of the information encoded in the DNA may be regulated, is by changing the relative amounts of specific DNA sequences in the cell genome. Thus, during differentiation, some of the sequences that are not required in the new cell type may be deleted from the DNA, or other required sequences may be specifically amplified.

a) Deletion of gene sequences

It is now accepted that in all but a few situations, the informational content of the DNA is the same in all the cell types of a particular species. Evidence has been accumulating for this since 1892 when Driesch showed that the nuclei from the endoderm region of developing sea urchin embryos, could be successfully manipulated to direct the synthesis of normal mesoderm development. Recent work by Gurdon (1962) and Gurdon *et al*. (1966) has extended this. Nuclei from the fully differentiated intestinal cells of adult *Xenopus* toads can

support the complete development of the fertile adult when stimulated by the correct environment – in this case the cytoplasm of an enucleate Xenopus egg.

McCarthy and Hoyer (1964) have obtained evidence for the genomic equivalence of differentiated tissues, by the application of DNA hybridisation techniques. Their experiments involved the immobilisation of denatured mouse embryo DNA in agar. It was found that a significant amount of labelled mouse L cell DNA was able to form a stable hybrid with this DNA, and in subsequent experiments, the hybridisation of the labelled DNA could be equally well 'competed out', by the addition of increasing amounts of DNA from a wide variety of tissues such as brain, thymus or liver.

These results demonstrate that eukaryotic cellular differentiation does not involve the obligatory loss of informational content of the DNA. Some exceptions to this are known, such as Ascaris, or the Gall midge, where most of the chromosomes are lost during somatic differentiation, and only the cells of the gonads in the adult have a full complement of DNA (Wilson, 1925).

b) Amplification and reiteration

As well as the possibility of specific gene sequences becoming amplified during the differentiation of certain tissues, it has been suggested that some genes may be present as multiple copies in the genome of all cells. Reiterated genes may be necessary where cellular functions demand the rapid production of large amounts of specific protein or RNA molecules. There are, however, several genetic arguments against the existence of reiterated genes in the genome, one of which is the requirement for a mechanism to maintain uniformity between these many copies.

Several alternative evolutionary schemes for groups of reiterated genes have been proposed, including the master–slave hypothesis (Callan 1967), and the saltatory replication theories (Britten and Kohne, 1968; Amaldi et al., 1973). A more likely mechanism of crossover fixation has recently been suggested by Smith (1974).

The reiteration frequency, and the possibility of specific amplification have now been investigated for several gene sequences. The most widely used technique for measurement of this value first requires the isolation of the RNA transcribed from the gene. This probe is either 1) isolated from cells grown in the presence of a radioactive label, 2) chemically labelled in vitro, usually by methylation with tritiated dimethyl sulphonate, or iodinated with I^{125}, 3) copied enzmatically to produce a highly labelled 'anti–strand' probe of RNA or DNA.

It is possible to estimate the number of copies of a gene either from the amount of DNA that is complementary to the probe (saturation analysis), or from the rate at which the probe hybridises with its complement (kinetic analysis).

In the first of these, the probe, usually RNA, is in excess and 'drives' the reaction. In the second only trace amounts of the probe are added, with the DNA being in vast excess (at least 100 times more gene sequences than the probe).

(1) Ribosomal RNA genes

Probably the best example of a reiterated gene is that of the sequences coding for the ribosomal RNA (rRNA). The redundancy of information, which is the same for both the 18S and 28S rRNA species appears to be 5 – 10 copies in the bacteria studied (Yankofsky and Spiegelman, 1962) and between

one and several hundred copies per haploid genome in eukaryotes (Brown and Weber, 1968a; Jeanteur and Attardi, 1969; Ritossa et al., 1966b). Several studies have demonstrated that these gene sequences found in any one species are identical (Dawid, Brown and Reeder, 1970; Birnstiel et al., 1969, Brown, Wensink and Jordan, 1972).

The genes coding for both the 5S and 4S RNA species are also reiterated (Brown et al., 1971; Ritossa et al., 1966b), but some heterogeneity appears to exist in the 5S genes of somatic cells (Wegnez et al., 1972; Ford and Southern, 1973).

Calculations based on known transcription rates of RNA, have demonstrated that the observed redundancy of ribosomal genes is necessary to produce the amount of rRNA detected in normal cells (Attardi and Amaldi, 1970). In agreement with this is the fact that some Drosophila mutants(bobbed) with a large reduction in the number of ribosomal genes, show a characteristically slow growth rate (Ritossa et al., 1966a).

It is known that the gene copies for the rRNA are clustered together in the genome (Brown and Weber, 1968b; Birnstiel et al., 1968). They are arranged in a repeating pattern of gene and spacer (Miller and Beatty, 1969a and b; Wensink and Brown, 1971) and because of their high G + C content, can be isolated by centrifugation on buoyant density gradients (Birnstiel et al., 1966; Brown and Dawid, 1968). Cytological evidence from both anucleolate Xenopus (Brown and Gurdon) and bobbed Drosophila (Rittossa et al., 1966a) also indicates that the ribosomal genes are clustered, and are located in the DNA associated with the nucleolus.

The ribosomal genes are also the only known genes that undergo amplification. This has been found to occur at the pachytene stage of oogenesis, and has been studied in amphibia (Gall 1968; MacGregor, 1968) and beetles (Gall and Rochaix, 1974). This amplification results in the formation of about 4,000 extra chromosomal nucleoli in the nucleus of Xenopus oocytes (Brown and David, 1968; Perkowska et al.,1968). These are similar to somatic cell nucleoli in ultra-structure and are active in rRNA production until the first meiotic division. It is known that very high numbers of ribosomes are required to support the early stages of Xenopus embryogenesis, and it appears that amplification of the ribosomal genes is necessary for the synthesis of this large amount of ribosomal RNA. In all other tissues the number of ribosomal genes is found to be constant (Brown and Weber, 1968a).

(2) Histone genes

Reasonably pure histone mRNA was first isolated from developing sea urchins (Kedes and Gross, 1960 and 1969; Kedes et al., 1969). This 9 - 10S RNA becomes rapidly labelled in the dividing cells of the embryo, and can be resolved by polyacrylamide gel electrophoresis into 5 major species, of the expected sizes to code for sea urchin histones (Weinberg et al., 1972). Recently this RNA has been translated into histones in a cell-free protein synthesising system (Gross et al., 1973).

Trace amounts of this labelled histone mRNA were used to probe reannealing sea urchin sperm DNA. The label becomes RNase resistant at a lower C_0t than either the bulk of the DNA sequences, or total sea urchin mRNA. Kedes and Birnstiel (1971) interpret this data as showing that the histone genes are reiterated about 400 times per haploid genome. This

labelled probe hybridises in a similar manner with Xenopus and Rhynchosciara DNA, indicating that the genes in these species are also reiterated.

More recently, histone mRNA has been isolated from dividing HeLa cells and translated in a cell-free system (Jackobs-Lorena, et al., 1972; Gallwitz and Breindl, 1972). This RNA has been copied and used to probe reannealing human placental DNA. Preliminary evidence suggests that the histone genes in the human genome are reiterated 50 – 100 times (Getz, Borun and Birnie, unpublished).

Because of their high G + C content, it has proved possible to isolate the histone genes from the bulk of the DNA (Kedes and Birnstiel, 1971; Birnstiel et al., 1974). The gene sequences are clustered in the DNA, in a similar manner to the ribosomal genes, and there are several studies that indicate that the histone gene copies are identical within a given species (Weinberg et al., 1972; DeLange et al., 1968a, b; Grunstein et al., 1973).

(3) Globin genes

One of the first mammalian mRNAs to be isolated in a pure form was the message coding for globin. Globin is almost the only protein being synthesised in immature blood cells (Kruh and Borsook, 1956; Dintzis et al., 1958). These cells are easily isolated from anaemic animals, and a 9 – 10S RNA component purified from the reticulocyte polysomes (Burny and Marbaix, 1965; Marbaix et al., 1966; Huez et al., 1967; Williamson et al., 1971). This component has been shown unequivocally to be the messenger RNA for globin, coding for both α and β globin chains in cell-free systems (Lingrel et al., Lanyon et al., 1972) and in the Xenopus oocyte (Lane et al., 1971).

The first attempt to measure the reiteration frequency of the globin gene was made by Williamson *et al*. (1970). Mouse globin mRNA methylated *in vitro* was used in saturation experiments with DNA immobilised on filters. Large amounts of RNA rapidly hybridised to the DNA, giving a calculated gene reiteration frequency of 50,000. However, the low T_m of this hybrid, and the ability to compete out the hybridisation with poly(A) (Morrison *et al*., 1972), suggested that the poly(A) segment of the mRNA was hybridising to the poly(dT) tracts detected in mouse DNA by Shenkin and Burdon (1972). Similar values were obtained by de Jiminez *et al*. (1971) using *in vitro* labelled chicken globin mRNA.

In vivo labelled globin mRNA can be easily isolated using avian nucleated reticulocytes and Bishop *et al*. (1972) have used this RNA to estimate the number of globin genes in the duck genome. Their kinetic method of analysis (Melli *et al*., 1971) gave a reiteration frequency of 10 copies, by comparison with the rates of hybridisation of DNA and RNA standards. Since corrections are required for both the G + C content of the mRNA, and the different rates of reaction of RNA–DNA and DNA–DNA hybridisation, the best interpretation of their data is that there are no more than 5 copies of the globin gene per haploid duck genome.

The isolation of an RNA dependent DNA polymerase (reverse transcriptase) from avian myeloblastosis virus (Packman *et al*., 1972; Verma *et al*., 1972; Kacian *et al*., 1972) has enabled highly labelled DNA copies of mRNAs to be synthesised. Trace amounts of these complementary DNAs (cDNA) can be followed in hybridisation reactions.

Harrison et al. (1972a) used cDNA copied from mouse globin mRNA in a kinetic analysis to estimate the number of mouse globin genes in mouse embryo DNA, and obtained values of 1 – 2 copies per haploid genome. Similar low figures were also obtained by Bishop and Rosbash (1973) and Packman et al., (1972) for the number of globin genes in the duck using a cDNA technique. These last authors also compared the number of globin genes in DNA from adult duck liver and from reticulocytes. Their preliminary data indicated no amplification of the globin genes in the reticulocyte DNA.

In the work reported here, the observations of Harrison et al.(1972) have been confirmed and extended. I have compared the number of globin genes in the mouse germ line (sperm) DNA, and DNA from an erythropoietic tissue (foetal liver) with total mouse embryo DNA (Harrison et al., 1974a). The corrections required for the effects of mismatching of bases in the hybrid, and the base composition and size of the hybrid have been determined. The results show that the globin genes are present in 1 – 2 copies per haploid genome and are not amplified or lost during differentiation.

(4) Immunoglobulin genes

Mammals have a highly developed immune system, and respond to a wide diversity of antigens by the production in specialised cells of large amounts of specific immunoglobulin (Ig) proteins. These proteins differ by only a few amino acids and are made up of heavy (H) chain and light (L) chain subunits. Each of these chains is encoded by one variable (V) and one constant (C) gene (Dreyer and Bennet, 1965). Heavy chains of all classes share the same V_H-gene pool and different C_H-genes can be linked to the same V_H gene. It is possible that every cell contains all immunoglobin (Ig) genes (germ line theory, Dreyer et al., 1967), or that variability is created somatically, so different antibody producing cells have different (and only a few) Ig genes (Brenner and Milstein, 1969; Gally and Edelman, 1970).

These hypotheses have been tested by several workers. The mRNA for a single specific Ig can be labelled in vivo, and isolated from myeloma cell lines. Using this probe, Premkumar et al., 1974 have shown that the mRNA coding for a heavy chain protein hybridises to mouse DNA in a biphasic manner. The first component (assumed to be the V_H gene pool), appears to be reiterated in the genome about 5,000 times; the second (C_H gene pool) component about 4 times. Since widely divergent V_H sequences would not be expected to cross-hybridise, this represents a minimum estimate of gene number. If the L chain variable region gene pool is of the same size, the information present in the germ line would be enough for 2.5×10^7 antibodies. These hybridisation values are the same for DNA from mouse embryos, lymphoid and myeloma cells, and indicate that specific loss or amplification of the Ig genes does not occur.

(5) Silk fibroin genes

The posterior silk gland of the larvae of the silk moth Bombyx mori is involved in the synthesis of large amounts of the silk fibroin protein (Tashiro et al., 1968). The 45S mRNA for silk fibroin is very stable in these cells (Suzuki and Brown, 1972) and has been isolated in a highly labelled form. This probe hybridises to whole larvae DNA and silk gland DNA, at a saturation level indicating that in both there are 1 – 3 silk genes per haploid complement of DNA (Suzuki et al., 1972).

However, in development of the silk gland, as in many other terminally differentiated tissues of dipteran larvae, extensive DNA synthesis occurs, without cellular division (Tashiro et al., 1968; Gall et al., 1971; Dickson, et al., 1971). The chromatids become paired and the euchromatic portion of each is replicated several hundred times (Bridges, 1935; Painter and Griffin, 1937; Beerman, 1956). This results in large banded structures visible under the light microscope – polytenised giant chromosomes. Thus, although the reiteration frequency of the silk gene remains the same in whole larvae and silk gland DNA, there is a very high number of silk genes in each cell. Suzuki et al. (1972) calculate that transcription of this large number of genes is necessary to explain the observed number of 10^{15} fibroin molecules present in the cells of the silk gland.

These conclusions for the low reiteration frequencies of specific genes have also been substantiated for total cell mRNAs in several studies. The rates of hybridisation of in vivo labelled mRNA from mouse L cells (Greenberg and Perry, 1971), sea urchins (Goldberg et al., 1973) or HeLa cells (Klein et al., 1974) are as predicted for non-repetitive transcripts.

A similar conclusion has been reached from the observed rate of annealing of a cDNA copied from the poly(A) containing RNA of mouse M2 cell polysomes (Birnie et al., 1974), or from HeLa cell mRNA (Bishop et al., 1974).

However, most of these studies cannot exclude the presence of short repetitive sequences in the mRNAs. These sequences could be between 30 and several hundred nucleotides long, and may be located in the untranslated region of the mRNA, where they may fulfil a regulatory function. Using techniques whereby partially hybridised molecules can be isolated on hydroxyapatite, Goldberg et al. (1973) have demonstrated that sea urchin gastrula mRNA is transcribed almost exclusively from single copy DNA, and contains no detectable, rapidly hybridising tags. Similarly for HeLa mRNA, Klein et al. (1974) have shown that more than 90% of the RNA contains no repetitive tags. However, about 6% of this mRNA appears to be transcribed entirely from repetitive DNA sequences. Controls demonstrate that this cannot be explained by HnRNA, or rRNA contamination, and since it contains poly(A), is unlikely to be histone mRNA. This indicates the presence of low levels of repetitive sequences coding for mRNA in the HeLa DNA.

It thus appears that for most structural genes, germ line reiteration, or amplification during differentiation does not occur. In addition a quantitative consideration (Kafatos 1972a,b) has demonstrated that even the relatively large accumulations of specific proteins that occur in some terminally differentiated tissues, such as the silk gland, or red blood cell, can be explained by the known rates of transcription and translation, assuming a fairly high stability of the mRNAs. Only the histone genes are known to be reiterated. Histone

mRNAs are only synthesised during the short DNA replication phase of the cell cycle, after which they are rapidly broken down (Robbins and Borun, 1967). This compares with other mRNAs, which are quite stable, and may accumulate over several cell generation times (Greenberg 1972; Perry and Kelly, 1973; Murphy and Attardi, 1973). Gene reiteration may also be necessary for production of the large amounts of histones required during the stages of rapid cell division in early embryogenesis.

The only other known informational sequences that are reiterated are the genes coding for the ribosomal RNAs. Here too the large numbers of ribosomes necessary for cell growth requires the transcription of many gene copies. Also, there is no translational "amplification" of the gene product, since the RNA is used directly by the cell.

It is of interest that both these classes of reiterated genes are known to be grouped together in the genome. Since the products of these gene clusters are homogeneous in any one species, their similar arrangement in the DNA may reflect the mechanism by which the uniformity of the reiterated sequences is maintained.

Since it appears that gene expression is not regulated at this level of DNA arrangement, other mechanisms of gene control must be investigated.

3. Transcriptional control of gene expression

There is considerable evidence that the transcriptional availability of sequences is the major site of the control of gene expression in eukaryotes. The theoretical basis of these controls was worked out in prokaryotic systems, where the regulation of the lac operon by the negative control of gene activity is best understood (Jacob and Monod, 1961; Gilbert and Muller-Hill, 1967; Beckworth and Zipser, 1970).

Transcription of the lac genes will occur unless the operator site of about 30 base pairs (Gilbert and Maxam, 1973) is occupied by the repressor molecule. This molecule has recently been shown to be a tetrameric protein of about 150,000 molecular weight (Adler *et al.*, 1972; Steitz *et al.*, 1974). The operator site is situated 5' to the coding sequence. Inducer molecules bind with the repressor protein and prevent its binding to the DNA, thus allowing transcription of the genes (Jacob and Monod, 1961). A polycistronic mRNA transcript of the 3 genes in the operon ensures co-ordinate expression of the information required.

In eukaryotic cells, the first line of evidence for the regulation of transcription comes from a comparison of the RNA sequences found in the nuclei of different tissues. It is assumed that this will approximate to the population of RNA molecules being transcribed. Examples of results from these studies are that there is less than 10% homology between the RNA sequences stored in *Xenopus* oocytes and synthesised in the blastulae (Davidson *et al.*, 1968), while there is up to 30% homology between the RNA sequences of mouse brain liver and spleen (Brown and Church, 1972). Precise interpretations of these experiments are doubtful because of several technical difficulties, but the conclusion that the cell can regulate

the information it transcribes is clear.

Difficulties also exist in the second line of evidence, which comes from the *in vitro* transcription of cell chromatin. In eukaryotic cells, the DNA is complexed with both histone and nonhistone proteins. This material stains with basic dyes *in situ* and is termed chromatin. During cell division the chromatin can be seen as condensed super coiled chromosomes, but in this form no transcription can occur (Johns, 1969). Interphase chromosomes are only visible as distinct entities in a few examples, such as the polytenised chromosomes of dipterans.

Several workers have shown that the RNA sequences transcribed from chromatin *in vitro* by *E. coli* or homologous RNA polymerase, reflects the differences found in the *in vivo* nuclear RNA (Paul and Gilmour, 1968; Gilmour and Paul, 1969; Flickinger *et al*., 1965; Marushige and Bonner, 1966; Marushige and Ozaki, 1967). In experiments where the chromatin from several tissues was dissociated into its constituents and then reassociated, Paul and Gilmour pointed to the involvement of non-histone proteins in regulating this differential gene activity. However, in much of this work only the hybridisation of repetitive sequences was followed and the results may be explained by contamination with endogenous RNA sequences. It is only recently with the use of specific gene probes, such as cDNA, that the general validity of these ideas has been substantiated.

Using these probes, it has been shown that globin sequences are transcribed by *E. coli* polymerase from reticulocyte chromatin and not liver chromatin (Axel *et al*., 1973) and from mouse foetal liver chromatin and not brain chromatin (Gilmour and Paul, 1973). Others have used mammalian

polymerase to the same effect (Steggles et al., 1974).

Although these data show that most of the genome is inactive in any particular tissue, the number of regulatory events involved in this control is very large. For example, Davidson and Britten (1973) calculate that if the complexity of the RNA transcribed in mouse brain is approximately 2×10^8 nucleotides (Hahn and Laird, 1971), this may represent the regulation of between 10^3 to 10^4 separate and diverse loci in the mouse genome (assuming the size of the transcripts to be between 10^4 to 10^5 nucleotides). Paul (1973) has suggested that the number of molecules involved in this regulation may be even higher. A single regulator protein of 20,000 – 50,000 molecular weight, would probably bind to a length of DNA of between 50 – 100 Å. This constitutes less than one third of the DNA–histone supercoil (Ris 1961 and 1962), and indicates that more than one sequence length may be required to destabilise the nucleohistone sufficiently to allow transcription to start. Crick (1971) has proposed that these recognition sites may exist as single stranded regions of DNA, held open by a hairpin structure in the compact nucleohistone.

4. Post-transcriptional control mechanisms of gene expression

Scherrer and Marcaud, (1968) in their cascade hypothesis, suggested that, superimposed on transcriptional regulation, were other nuclear and cytoplasmic mechanisms to ensure the stability of the differentiation processes. These mechanisms may impose both quantitative and qualitative control on the expression of the transcribed information. Darnell et al. (1973) have pointed out that post-transcriptional regulation requires the overproduction of potential mRNA sequences. Different cellular conditions will then alter the probability of a potential mRNA exiting from the pool of transcribed sequences, and functioning in the cytoplasm. Thus, the specificity of post-transcriptional control must lie in the sequences of the RNA.

Some insight into the mechanisms by which this control is exerted in the cytoplasm can be gained by a study of the structure and functioning of the messenger RNAs.

4(i) Cytoplasmic control processes

a) General characteristics of Messenger RNAs

Messenger RNA was first identified in bacterial systems, by its rapid labelling with radioisotopes, a rapid turnover relative to rRNA, and its 'DNA like' base composition (Gross et al., 1961; Jacob and Monod, 1961). In eukaryotes the stability of many messenger RNAs is much greater than in bacteria, where decay constants of less than 5 minutes have been recorded (Levinthal et al., 1962). Functional mRNA is located on polysomes during protein synthesis (Rich et al., 1963) when it is very sensitive to RNase degradation. It is dissociable from the polysomes by chelating agents such as EDTA, and is usually released as an RNA-protein particle (mRNP) (Chantrenne et al., 1967).

The definitive proof that an RNA can function as a message is its ability to direct the synthesis of a specific protein in a cell-free protein synthesising system. In several instances, other criteria have been used to identify mRNAs, for example, the large size (30S) of the myosin mRNA in muscle cells (Heywood et al., 1967), or the distinctive base composition of the silk fibroin mRNA (Suzuki and Brown, 1972).

b) Poly(A) sequences

One major distinguishing feature of most eukaryotic mRNAs is the presence on the 3' end of the molecule of a poly adenylic acid sequence (poly(A)). The ability of this poly(A) sequence to hybridise to poly deoxythymidylate oligomers (poly(dT)) covalently bound to cellulose, enabled Edmonds et al. (1971) to isolate messenger like RNA from HeLa cells. Several other similar techniques have been employed to purify mRNAs. It has been found that the poly(A) containing RNA will stick to millipore (nitrocellulose) filters (Brawerman et al., 1972; Rosenfeld et al., 1972) and will hybridise to polyuridilic acid (poly(U)) that has been immobilised on sepharose (Wagner et al., 1972), or glass fibre discs (Sheldon et al., 1972b).

At the present time the only known eukaryotic mRNA that does not contain these sequences are the histone messenger RNAs (Adesnik et al., 1972; Adesnik and Darnell, 1972; Schochetman and Perry, 1972). Thus, when labelled polysomal preparations are fractionated by poly(A) selection techniques, between 60 – 85% of the mRNA behaves like poly(A) containing material (Lindberg and Person, 1972; Lee et al., 1971; Adesnik et al., 1972). This suggests that the majority of mRNAs contain a poly(A) sequence, but the occurrence of some mRNA species, other than the histone mRNA

that lack poly(A), cannot be excluded.

The poly(A) segment consists of a pure poly(A) sequence, as indicated by its resistance in high salt conditions to pancreatic and T_1 RNase digestion. Poly(A) isolated in this way from HeLa cells contains 1 adenosine per 200 AMP residues, and no other nucleotides. This indicates that the poly(A) segment is 200 nucleotides long and is terminated by a free 3' OH group (Molloy and Darnell, 1973). Similar results have been obtained with mouse sarcoma mRNA (Mendecki et al., 1972).

A segment of this configuration could only originate from the 3' terminus of the mRNA, and this location of the sequence has been recently confirmed by digestion of mRNA with an exonuclease that requires a free 3' OH. This rapidly reduces the ability of the RNA to bind to poly(U) sepharose (Molloy et al., 1972), or nitrocellulose filters (Sheldon et al., 1972b). Also a considerable fraction of the poly(A) sequence becomes labelled when the 3' terminus of the mRNA is oxidised with periodate and reduced with ^{3}H borohydride (Nakazato et al., 1973).

The poly(A) isolated from newly synthesised mRNA migrates as a relatively homogeneous component in polyacrylamide gels, and with the mobility expected of a molecule of about 200 nucleotides in length (Edmonds et al., 1971; Adesnik et al., 1972; Sheiness and Darnell, 1973; Mendecki et al., 1972). mRNA from lower eukaryotes has been demonstrated to contain a shorter 3' poly(A) segment of about 100 nucleotides in Dictyostelium (Firtel et al., 1972) and about 50 nucleotides in yeast (McLaughlin et al., 1973).

Several studies with cells grown in culture have indicated that the size of the messenger poly(A) sequence, decreases in the cell cytoplasm.

Thus, the poly(A) sequence isolated from total HeLa cell mRNA has been shown to migrate as a heterogeneous band of RNA on acrylamide gels (Darnell et al., 1971a, Mendecki et al., 1972). This heterogeneity in size cannot be explained by aggregation of the molecules (Sheiness and Darnell, 1973) and these authors suggest it might result from the metabolic processing of an originally uniform sequence length.

In several pulse chase experiments it has been shown that the poly(A) sequence of labelled HeLa mRNA becomes progressively shorter and more heterogeneous over a period of several hours. After 6 hours the size of the poly(A) has decreased from 190 $\pm$ 20 nucleotides to 100 $\pm$ 20, with some molecules of only 50 nucleotides occurring after longer chase periods. Similar results have been found using mouse L cells (Brawerman 1973).

These data suggest that the shortening of the poly(A) in the cytoplasm may be related to the utilisation of mRNA during protein synthesis. A ticketing hypothesis to regulate mRNA half life has been suggested by Sussman (1970). However, it has been found that the shortening of poly(A) is not prevented when protein synthesis is inhibited (Sheiness and Darnell, 1973; Brawerman, 1973). Also, in a 6 hour period, a mRNA is probably translated at least 3,000 times. Clearly the removal of one nucoeltide per ribosomal transit is not tenable.

These ideas have been substantiated by a study of the

poly(A) sequence of several specific mRNAs. Rapidly labelled duck reticulocyte 9S globin mRNA has been found to contain a poly(A) sequence of about 190 nucleotides, by a comparison of the migration on acrylamide gels of the RNase resistant fragment, with 4S and 5S (Pemberton and Baglioni, 1972). Similarly, calf lens has been shown to have a poly(A) sequence of about 150 nucleotides (Favre et al., 1974) as does chick myosin (Mondel et al., 1974) and collagen (Harwood et al., 1974).

However, a shorter poly(A) sequence has been described for the long lived reticulocyte globin mRNA. Thus, Lim and Cannelekis (1970) described a poly(A) sequence of between 50 – 70 nucleotides in rabbit globin mRNA, but Burr and Lingrel (1971) found a poly(A) sequence no longer than 10 nucleotides at the 3' end of this mRNA, and Hunt (1973) of between 30 – 40 long. The poly(A) segment of mouse globin mRNA has been shown to be 50 – 70 nucleotides long, and by ^{3}H borohydride studies, to be on the 3' end of the RNA (Morrison et al., 1973; Mansbridge et al., 1974). Sequencing studies have shown that there are no other bases present in this fragment (Mansbridge et al., 1974).

The only well characterised eukaryotic message to lack poly(A) is the histone mRNA. Several differences in the metabolism of this mRNA may be related to the absence of poly(A). Histone mRNAs are only made during the DNA synthesis phase of the cell cycle, when other mRNA production is minimal (Borun et al., 1967; Gallwitz and Mueller, 1969). After being synthesised, histone mRNAs exit rapidly from the nucleus (Schochetman and Perry, 1972), in contrast to the 10 – 20 minute lag observed for other mRNAs (Penman et al., 1968; Jelenik et al., 1973). This suggests that bypassing

the normal poly(A) pathway may be related to a cell requirement for rapid histone synthesis. The histone mRNA has a similar short half-life to the non-adenylated bacterial mRNAs, implying a role for poly(A) in mRNA stability. However, both this stability, and the rapid appearance of the mRNA in the cytoplasm may also be explicable by general changes in the cell metabolism during this stage of the cell cycle.

The possible translational role of poly(A) has been investigated in this work. The poly(A) sequences of mouse globin mRNA was removed enzymatically, and the ability of this RNA to direct the synthesis of globin protein, studied in various cell-free systems. By comparison with intact mRNA, initiation, elongation, termination and reinitiation on this mRNA is not impaired. This has also been demonstrated for total mouse L cell mRNA by Bard et al., (1974).

However, recent work by Huez et al. (1974), has indicated a role for poly(A) in the stability of the mRNA. In these experiments the poly(A) sequence was enzymatically removed from rabbit globin, and the translational ability of this message followed over several hours by injection into Xenopus oocytes. The amount of globin synthesised on the treated mRNA as compared with native mRNA was found to be the same for the first hour of translation, but to decrease after this. The mechanism by which poly(A) may confer this stability is unknown.

c) mRNA size

A feature of all mRNAs so far studied is that their sequence length is greater than required for coding of the polypeptide, though not long enough to allow the possibility of polycistronic mRNAs (Paul et al., 1972; Murphy and Attardi,

1973). Thus, mouse or rabbit globin mRNA is about 220,000 $\pm$ 10,000 molecular weight, approximately equal to 660 bases (Williamson et al., 1971; Gaskill and Kabat, 1971; Labrie, 1969). Since in the mouse, 423 – 435 bases are required to code for α or β globin, and the size of the poly(A) segment is 50 – 70 bases (Mansbridge et al., 1974), there may be up to 180 untranslated bases in the globin message.

Some of these sequences are known to be located at the 3' end of human globin mRNA. An α globin chain variant is known (Haemoglobin Constant Spring) which exceeds the normal polypeptide length by 31 residues (Clegg et al., 1971). From sequence studies of these amino acids, it appears that a mutation in the normal termination codon has resulted in read–through of these usually untranslated sequences, till a further termination codon is reached 93 bases later. This interpretation is substantiated by a second 'frame shift' variation (Haemoglobin Wayne), which has 5 additional carboxyl amino acids (Flatz et al., 1971).

Computer studies (Hunt and Dayhoff, 1972) have suggested that these extra sequences may have evolved by the duplication, and subsequent mutation of α protein sequences.

Recent sequencing work by Proudfoot and Brownlee (1974) on a mixture of rabbit α and β globin mRNA, have revealed the same sequence of AUUGC poly(A) at the 3' end of both messages. Since the poly(A) sequence is not transcribed by RNA polymerase but added by a different enzyme, this sequence may represent the recognition site for the enzyme.

Little is known about possible untranslated sequences at the 5' end of the mRNA. From studies on the events leading to the initiation of protein synthesis in prokaryotes, it is

thought that the secondary structure of a messenger RNA molecule must signal that a particular AUG triplet is an initiation codon. This triplet is recognised by a charged f-met-tRNA, and an mRNA - ribosome complex is formed. The segment of mRNA in this complex that is protected from RNase digestion, has been isolated and sequenced by several workers, (Steitz 1969; Staples et al., 1971; Robertson et al., 1973) using mRNAs from viral and prokaryotic sources. Between 9 and 25 bases 5' to the AUG codon, are found to be protected (Maizels, 1974).

In bacteria, ribosome associated factors have been described, that act at the level of initiation, by affecting ribosomal recognition of individual mRNAs (Revel et al., 1970; Lee-Huang and Ochoa, 1971). The presence of similar factors in differentiated mammalian cells was first described by Heywood (1970). These factors have now been clearly demonstrated for myosin and myoglobin translation in a rabbit reticulocyte cell-free system (Heywood et al., 1974) and for the efficient translation of several mRNAs in the Krebs cell-free system (Nudel et al., 1973; Wigle and Smith, 1973). See ref These factors may recognise a specific sequence at the 5' end of the mRNA molecule, which may not be translated.

Another possibility is that these factors may recognise a secondary structure feature of the RNA. Several of the well characterised mRNAs such as globin or ovalbumin are known to exhibit considerable secondary structure (Williamson et al., 1971; Lingrel et al., 1971; Palacios et al., 1973). RNA secondary structure has been demonstrated in tRNAs (Holley et al., 1965) and in viral RNAs (Min Jow et al., 1972). These structural features may increase the RNase resistance of the

molecule or may be essential for the correct functioning of the molecule.

d) Messenger RNA–protein particles

When messenger RNA is released from polysomes by EDTA treatment (Burny et al., 1969), or by reaction with puromycin (Bloebel, 1972), it is found associated with protein as an mRNP particle.

Globin mRNA released from rabbit reticulocyte polysomes by the puromycin treatment contains 2 types of proteins with molecular weights of 78,000 and 52,000 (Bloebel, 1972). Other minor components may also be present. These proteins appear to be tightly associated with the RNA, and remain bound in conditions which dissociate many of the ribosomal proteins (Bloebel, 1972). This suggests a specific and functional interaction of these proteins. Similar sized proteins have been found on the mRNA of duck and mouse reticulocytes, as well as total mRNAs from mouse liver or L cells (Morel et al., 1971; Sampson et al., 1972; Perry and Kelly, 1968). Electron microscope studies indicate regular location of these proteins along the mRNA (Dubochet et al., 1973), with an RNA: protein ratio of about 1:3 (Lissitzky et al., 1970). Considerable evidence now exists for the association of the 78,000 MW proteins with the poly(A) sequence of mRNAs (Kwan and Brawerman, 1972; Bloebel, 1973). The other major protein component may be located at the 5' end of the mRNA. These proteins could be involved in protecting the untranslated ends of the message from nuclease digestion – the rest of the mRNA being protected by ribosomes. It is known that mRNA and mRNP are equally well translated in cell-free systems (Lingrell et al., 1971; Sampson et al., 1972), indicating that they are not involved in the translation

process itself.

The post-transcriptional functioning of messenger-protein particles has been clearly implicated in several systems such as the eggs of sea urchins (Denny and Tyler, 1963), loach embryos (Spirin 1969), and Xenopus oocytes (Rosbash and Ford, 1974). In these cells, maternally derived mRNA (Davidson 1968; Rosbash and Ford, 1974) is stored as an mRNP, probably both to prevent accidental translation, and to protect the mRNA against degradation (Spirin, 1966). It has been suggested that in the sea urchin, the activation of this stored message may be triggered by polyadenylation of the mRNA (Slater et al., 1974).

In differentiating cells, mRNAs may also accumulate in a stored form, to subsequently allow the rapid, co-ordinated synthesis of large amounts of particular proteins such as haemoglobin (Wilt 1965) or pancreas enzymes (Wessels and Wilt, 1965). Similarly, proteins may also be involved in a rapid and co-ordinated switching-off of a set of translating mRNAs.

Clearly the association of protein with mRNA in these ways, represents a powerful method by which the cytoplasmic regulation of gene expression may occur.

4ii) Nuclear post-transcriptional control mechanisms

Several features of nuclear RNA metabolism may be involved in regulating the appearance of transcribed sequences in the cytoplasm.

a) The size of nuclear RNA

Heterogeneous RNA of a high molecular weight (HnRNA) has been found in the nuclei of many eukaryotic cells. The original estimates of the size of this RNA, using sucrose gradients or acrylamide gel electrophoresis exceeded 10^6 nucleotides in length (Attardi et al., 1966; Scherrer et al., 1966; Warner et al., 1966). However, aggregation, or inter-strand partial duplex formation of RNA has been demonstrated to occur under non-denaturing conditions (Mayo and deKloet 1971; Bramwell, 1972), and recent experiments using denaturing agents such as formamide or dimethyl sulphoxide give size ranges for nuclear RNA of mouse kidney cells of $0.5 - 1.4 \times 10^4$ nucleotides (Acheson et al., 1971), and $2 - 4 \times 10^4$ nucleotides in rat ascites cells (Holmer and Bonner, 1973). These molecules are still considerably larger than the average size of mRNA molecules, $1 - 3 \times 10^3$ nucleotides (Murphy and Attardi, 1973; Fan and Penman, 1970).

Electron microscope studies substantiate these values (Holmes and Bonner, 1973; Granboulan and Scherrer, 1969) for HeLa and duck erythroblast HnRNA. RNA transcripts larger than this have been observed in Triturus lampbrush loops (up to several hundred thousand nucleotides (Miller and Bakken, 1972)), and in the transcript of a puffed band of a Chironomus polytene chromosome (approximately 5×10^4 nucleotides (Beerman, 1967; Lambert, 1973).

It has long been thought that HnRNA may represent the primary gene transcript, and that these large molecules are the precursors to the mRNA found in the cytoplasm (Scherrer et al., 1963; Penman et al., 1963). A precedent for this exists in the cleavage of 18S and 28S ribosomal RNA from a large precursor molecule. This precursor forms a large part of the nuclear RNA synthesised under certain growth conditions, can be isolated as a discrete 45S molecule (Scherrer et al., 1963; Penman, 1966; Jeanteur et al., 1968), and is processed in discrete stages to the rRNA found in the cytoplasm. The base composition of rRNA is distinctive and contains modified bases, that have been utilised to prove detailed sequence similarities (Scherrer et al., 1963; Maden 1971; Jeanteur et al., 1968). The enzymes involved in this processing have not been identified, but both the conserved and discarded sequences at all stages, have now been isolated and visualised using the electron microscope (Wellauer and Dawid, 1973).

The proof of a similar precursor–product relationship for HnRNA and mRNA is complicated by the high number of genes normally expressed in the cell. Several early studies on general sequence similarities of HnRNA and mRNA were compatible with a precursor–product relationship (Georgiev et al., 1963, 1972 and Scherrer et al., 1970). The base sequence homology between the two populations is mainly at the 3' end of the HnRNA (Georgiev et al., 1963; Ryskov et al., 1972).

However, experiments utilising drugs such as actinomycin D to repress rRNA synthesis, gave values for HnRNA half life of approximately 3 minutes (Soeiro et al., 1968).

Since the time for transcription of an average HnRNA

molecule approaches 3 minutes, and no simultaneous transcription–translation complexes (such as are found in prokaryotes) have been detected in eukaryotic chromatin (Miller and Bakken, 1972), this would indicate that most of the HnRNA is degraded as it is made, and is never exported to the cytoplasm (Harris, 1962).

Recently Brandhorst and McConkey have circumvented the use of drug inhibitors, and estimated a much longer first-order decay rate for L cell HnRNA of 23 minutes. This compares with an mRNA half life in these cells of 600 minutes (Greenberg, 1972). Similar estimates of HnRNA half life have been made in HeLa cells (Penman et al., 1968) and duck reticulocytes (Attardi et al., 1966). These values would allow processing of a portion of the HnRNA to the cytoplasm, and Brandhorst and McConkey (1974) estimate that about 2% of labelled HnRNA may exit to the cytoplasm during long chase periods.

b) The poly(A) sequence of nuclear RNA

One of the best indirect pieces of evidence for the precursor–product relationship of HnRNA and mRNA, has been the detection of a poly(A) segment in both molecules. The HnRNA poly(A) sequence has been found to be the same size as the cytoplasmic sequence and also on the 3' end of the RNA (Edmonds et al., 1971; Darnell et al., 1971b; Jelenik et al., 1973).

Poly(A) was first discovered during the studies by Edmonds and Abrams (1960, 1963) of a nuclear enzyme that specifically incorporates adenylic acid into RNA, without a DNA template. Three findings suggest that, rather than having a transcriptional origin, a similar enzyme may be

responsible for the polyadenylation of HnRNA molecules.

1. When added to cultured HeLa cells, the drug actinomycin D almost completely prevents DNA dependent RNA synthesis in less than 2 minutes. However, it has no effect on poly(A) synthesis for several minutes after this, indicating that the poly(A) sequence is not added by RNA polymerase (Jelenik et al., 1973).

2. Adenovirus 2 DNA sequences, which are found integrated into host cell DNA, are transcribed into HnRNA and processed into mRNA. Both these molecules contain a poly(A) sequence of 200 nucleotides at their 3' end. However, adenovirus DNA has been shown not to contain thymidine stretches from which these poly(A) sequences could have been transcribed (Philipson et al., 1971).

3. It is possible to isolate deoxypyrimidine nucleotide stretches from DNA because of their resistant to acid hydrolysis (Burton and Peterson, 1957). Oligo(dT) stretches can be isolated from HeLa DNA (Birnboim et al., 1972), or from mouse DNA (Mansbridge, unpublished observations), but these tracts are too small, and not present in large enough amounts to account for the total number of poly(A) sequences found on different mRNAs. However, in the slime mould Dictyostelium, Firtel et al. (1972) and Jacobson et al. (1974) have demonstrated that a short oligo(A) sequence is transcribed at the 3' end of the HnRNA, and this sequence length is then increased by a different enzyme. The (dT) tracts found in eukaryotic DNA are large enough to fit this hypothesis but no direct proof for this has yet been obtained.

The poly(A) segments are rapidly synthesised, and kinetic evidence, and the inability to find unattached nuclear poly(A)

molecules, indicate that they are formed by the stepwise addition of adenylate residues to the end of the HnRNA (Jelenik et al., 1973).

c) The role of poly(A) in nuclear post-transcriptional regulation

In HeLa cells, after the addition of ^{3}H adenosine, labelled poly(A) is confined to the nucleus for the first 1 – 2 minutes (Jelenik et al., 1973), showing that the nucleus is the major site of poly(A) synthesis. After this, labelled poly(A) begins to accumulate in the cytoplasm, while the majority of labelled HnRNA remains in the nucleus. This implies that during nuclear processing, a greater proportion of the poly(A) sequences are conserved than the HnRNA sequences.

Information about the role of poly(A) has been obtained with the drug cordycepin (3' deoxyadenosine). In HeLa cells, this drug terminates RNA chains prematurely, and rRNA synthesis rapidly ceases (Penman et al., 1970). Although HnRNA synthesis is not affected, the appearance of newly synthesised mRNA in polysomes is reduced by more than 80%. This is apparently due to an inhibition of the enzyme that adds the poly(A) sequence, and those mRNAs that do appear in the cytoplasm have either no poly(A) sequences and are probably histone mRNAs, or have a considerably shorter poly(A) sequence (Jelenik et al., 1973).

These authors have found labelled poly(A) attached to all size classes of HeLa HnRNA, 45 seconds after the addition of ^{3}H adenosine. If RNA synthesis is then blocked by actinomycin D, this addition of poly(A) to all size classes of RNA still occurs, and even if cordycepin is added, previously synthesised poly(A) appears in the cytoplasm in

mRNA after 10 to 15 minutes. After another 15 minutes about 30 – 40% of the previously labelled poly(A) can be found in the cytoplasm.

It has been estimated that between 30 to 40% of the largest HeLa HnRNA molecules contain poly(A) (Jelenik et al., 1973). Several possibilities that cannot be distinguished may give rise to the molecules that do not contain poly(A). Some HnRNA molecules may not contain RNA sequences to be expressed in the cytoplasm, are thus not adenylated, and turn over in the nucleus. 5' RNA sequences produced by the processing of mRNAs from HnRNAs as well as sequences from molecules artifactually cleaved during preparation, and incomplete HnRNA molecules, will not contain poly(A). It is also possible, that precursor molecules are first cleaved and the revealed site then polyadenylated, although this is not consistent with a short lag period between the termination of transcription and polyadenylation (Jelenik et al., 1973).

From this data it has been suggested that the addition of poly(A) to an HnRNA sequence is a post-transcriptional control mechanism, poly(A) being a 'tag' for the conservation and subsequent processing of the adjacent sequence to cytoplasmic mRNA (Jelenik et al., 1973; Darnell et al., 1973).

These views have been challenged by Perry and his colleagues, based on their work in mouse L cells (Latore and Perry, 1973; Perry et al., 1974). They find that only a small fraction of nuclear poly(A) exits to the cytoplasm, indicating a significant amount of nuclear poly(A) turnover. The appearance of poly(A) in the cytoplasm is very rapid, which is inconsistent with polyadenylation being an early step in mRNA processing. They further suggest from experiments using cordycepin, that a large pool of polyadenylated HnRNA

molecules may turn over in the nucleus, while a separate, smaller pool of polyadenylated HnRNA serves as a precursor to cytoplasmic mRNA. The conclusion from these results is that poly(A)does not have an early regulatory role in post-transcriptional processing,as suggested by Darnell and his colleagues.

In this work I have investigated the association of poly(A) with nuclear globin RNA sequences in erythroid and non-erythroid tissues. The results are consistent with the views of Perry et al. (1974), and indicate that polyadenylation of a nuclear RNA molecule is not in itself sufficient to ensure that the RNA will be processed and transported to the cytoplasm.

At the present time, the nuclear function of the poly(A) segment is unclear. However, the polyadenylation pathway appears to be obligatory, since when it is blocked, most mRNAs are not processed to the cytoplasm (Jelenik : et al., 1973).

The fact that histone mRNA is not adenylated may be related to the requirement for rapid histone synthesis during division, with the mRNA bypassing the slower poly(A) pathway (Perry and Kelly, 1968; Schochetman and Perry, 1972). The mechanism by which this occurs may be related to the 3' nucleotide sequence, or secondary structure of the histone mRNA, or be a function of its production at a particular time in the cell cycle (Robbins and Borun, 1967).

The fact that some viruses which replicate in the cytoplasm also contain poly(A) (Yogo and Wimmer, 1972; Johnston and Bose, 1972), does not conflict with a nuclear role for these sequences – virus mRNAs may encounter similar problems as that of the cell mRNA in getting from the site of manufacture to the site of translation. For example, the replication of polio virus occurs on smooth membranes, and the translation of viral poly(A) containing mRNA, on the rough

membranes of the cell (Caliguiri and Tamm, 1970).

d) Direct evidence for the presence in HnRNA of mRNA sequences

This requires identification of a specific mRNA, either by its sequence, or by its ability to code for a specific protein. A translational approach has been followed by several workers. R. Williamson et al. (1973) have micro-injected HnRNA from mouse foetal liver cells into Xenopus oocytes and obtained globin synthesis. The RNA used for these experiments sediments faster than 35S in sucrose gradients, but aggregation of the RNA could not be excluded. However, brain HnRNA prepared in the presence of reticulocyte polysomes did not cause globin synthesis in oocytes.

It should be pointed out that the detection of mRNA sequences in large nuclear RNA does not prove that these sequences are a precursor to cytoplasmic RNA. The best known evidence for such a precursor-product relationship for a specific gene comes from the work of Stevens and A. Williamson (1973) who have purified from a mouse myeloma cell line, the HnRNA and the mRNA coding for an antibody heavy chain. This antibody protein binds tightly to its mRNA, and can be then immunologically precipitated. Using this method, a discrete nuclear RNA of about 2×10^6 daltons has been isolated, and found to direct the synthesis of heavy chain proteins when injected into oocytes. Pulse-chase experiments show that this precursor, which is polyadenylated, is processed to the mRNA size of about 6×10^5 daltons, and then transported to the polysomes. After extended chase periods more than 96% of the labelled, antibody precipitable RNA, has left the nucleus and is present in the cytoplasm, confirming a precursor-product relationship. However, when the cells are not growing, more

than 80% of the precipitable RNA remains in the nucleus, most of which is mRNA sized, indicating a block at, or just before, the transport of the RNA from the nucleus. The authors calculate that in a rapidly growing cell, about 90 pre-mRNA molecules exist in the nucleus, and about 5,000 mRNA molecules in the cytoplasm (A. Williamson, 1974), and point out that a larger transcript may have escaped detection if it is processed rapidly, or if some processing of the transcript is required to make the protein binding site available.

A rigorous proof of the presence of mRNA sequences in HnRNA has recently been published by Ruiz-Carillo et al. (1973). Total RNA prepared from duck erythroblasts was sedimented through dimethyl sulphoxide - a solvent known to break hydrogen bonds (Strauss, et al., 1968), and the RNA sedimenting faster than 45S added to a Krebs cell-free system. This RNA codes quite efficiently for duck globin, and was shown not to code for rabbit globin when rabbit 9S RNA was cosedimented with the sample. This makes contamination of the large RNA with mature 9S mRNA very unlikely.

This data is in agreement with the earlier observations of Imaizami et al. (1973) who used cDNA prepared from duck 9S RNA, to demonstrate the presence of globin sequences in HnRNA of up to 10×10^6 daltons. This material was recovered from denaturing DMSO gradients, which were found to break down a high proportion of the large HnRNA isolated from sucrose gradients. Newly synthesised HnRNA did not appear to be as labile to DMSO denaturation as the total HnRNA. The authors suggest that this is indicative of the presence in the RNA of hidden processing nicks.

In contrast with these results, MacNaughton et al. (1974), using duck globin cDNA and nuclear RNA from duck erythroblasts, cannot find nuclear globin RNA sequences sedimenting faster than 14S in denaturing formamide gradients. They conclude that the molecular weight of the precursor is $6 - 7 \times 10^5$ daltons – 3 times larger than the mRNA.

Although there is evidence for the existence of precursor molecules for globin and immunoglobin, searches for the precursors to several larger mRNAs have proved negative. McKnight and Shimke (1974), have not been able to find evidence for a precursor to ovalbumin 18S mRNA in oestrogen stimulated chick oviduct. Using cDNA, these workers calculate they could have detected less than one molecule of precursor per cell.

Similarly Suzuki and Brown (unpublished) have not been able to find a precursor to the giant 45S silk fibroin mRNA using sequencing techniques.

Although these studies clearly demonstrate that HnRNA molecules contain mRNA sequences, most of the HnRNA synthesised never leaves the nucleus, and thus cannot be a precursor of cytoplasmic mRNA (Penman et al., 1968; Soiero et al., 1968; Brandhorst and McConkey, 1974). This may be because only a small number of the nuclear RNA molecules are mRNA precursors, or because only a small fraction of each HnRNA molecule is destined to appear as cytoplasmic mRNA.

Many of the early experiments to distinguish between these possibilities attempted to compare the complexity of nuclear and cytoplasmic RNA using filter hybridisation techniques. Since mainly only repetitive sequences hybridise in these experiments, the results, though indicating a high nuclear RNA

complexity, are not definitive (Scherrer et al., 1970; Hahn and Laird, 1971; Church and Brown, 1972). In a more recent analysis, Getz et al.(1974) have estimated the complexity of the poly(A) associated nuclear RNA of mouse M2 cells. This nuclear RNA can be copied by reverse transcriptase into a cDNA, and the complexity of the template RNA measured by a kinetic analysis of the hybridisation of the RNA and cDNA. A similar analysis has been carried out on the polysomal mRNA of the cells (Birnie et al., 1974), and from these studies, several points emerge.

1. Both the nuclear and cytoplasmic RNA sequences adjacent to the poly(A) segment are transcribed from non-repetitive DNA in the mouse genome.

2. The complexity of the nuclear RNA sequences is about 5 times higher than that of the mRNA sequences, being complementary to about 3% (as compared with 0.6%) of unique mouse DNA. Within this nuclear RNA, 2 frequency classes of sequences can be distinguished, with the majority of the RNA being represented by only a few copies per nucleus.

3. A large proportion of the sequences in nuclear RNA that are next to the poly(A) segment, are also found in polysomal mRNA. However, the relative proportions of the two classes of nuclear RNA, appear to be different in the cytoplasm.

These data are thus consistent with the hypothesis that only a small fraction of each HnRNA molecule – the sequence adjacent to the 3' poly(A) segment – are transcribed from structural gene sequences, and are conserved and processed to the cytoplasm where they function as mRNAs. It would also appear that superimposed upon this process, a post-transcriptional mechanism is operating that controls the cytoplasmic expression

of the nuclear RNA.

The existence of similar post-transcriptional regulation has been recently indicated by work on the nuclear and cytoplasmic RNA of developing sea urchin embryos (Galau et al., 1974; Hough, Smith and Davidson, unpublished).

e) Sequence differences between HnRNA and mRNA.

It has been demonstrated by several workers that many or all HnRNA molecules contain sequences that are transcribed from reiterated DNA. These sequences become RNase resistant after very short incubations with homologous DNA, while the rest of the RNA hybridises to unique DNA sequences (Darnell and Balint, 1970; Melů et al., 1971; Church and McCarthy, 1967). RNA can be fragmented to different lengths by sonication or brief alkali treatment, and the rate of hybridisation of these fragments, and their association with poly(A) studied.

From several studies it appears that sequences transcribed from reiterated DNA are present in all HnRNA molecules longer than the expected mRNA size ($1 - 3 \times 10^3$ nucleotides)(Holmes and Bonner, 1974; Jelenik et al., 1974; Molloy et al., 1974; Smith et al., 1974). Such reiterated sequences have not been found in either HeLa mRNA (Klein et al., 1974) or sea urchin mRNA (Goldberg et al., 1973). Reiterated sequences, which may be 200 – 300 nucleotides long (Holmes and Bonner, 1974; Jelenik et al., 1974), appear to be finely interspersed with sequences of 1000 – 5000 nucleotides, transcribed from unique DNA (Holmes and Bonner, 1974; Molloy et al., 1974). Sequences next to the 3' poly(A) segment, appear to be entirely non-repetitive (Molloy et al., 1974; Getz et al., 1974). The reiterated RNA may represent transcription regulatory molecules (Britten and Davidson (1969) or may reflect

transcription of sequences involved in DNA synthesis, or recombination (Holmes and Bonner, 1974).

Some of these reiterated sequences are present in HnRNA as double stranded regions (Jelenik and Darnell, 1972). When HnRNA is denatured, these self complementary sequences reform very rapidly, and, following considerable RNA fragmentation, are available to hybridise to reiterated DNA sequences. The double stranded regions isolated by their RNase resistance, appear to be about 300 nucleotides long but are very heterogeneous in size (Jelenik and Darnell, 1972). The distances between these self complementary regions varies widely between a few hundred and several thousand nucleotides (Jelenik et al., 1974; Molloy et al., 1974).

Molloy et al. (1972b) have discovered an oligo(U) rich sequence in HeLa cell HnRNA that is absent from the mRNA. This fragment is about 30 nucleotides long, contains about 80% uridylic acid and hybridises rapidly to HeLa DNA. Mammalian DNA has been shown to contain poly(A) rich stretches that when copied, may give rise to oligo(U) (Burdon and Shenkin, 1972; Shenkin and Burdon, 1972). Recent data demonstrates that every completed HnRNA may contain 2 – 3 such oligo(U) sequences located in the 5' region of the molecule (Molloy et al., 1974).

Nakazato et al. (1974) have recently described an oligo(A) sequence of similar length to this in the 5' region of HeLa HnRNA. The distribution of this sequence among different HnRNA molecules, is not known.

It has been suggested that the reiterated sequences, and the other distinctive regions discovered, may be involved in the post-transcriptional processing of the RNA (Jelenik et al., 1974). One example of this is that the enzyme that rapidly

processes ribosomal RNA in E. coli has been shown to act specifically at double stranded sites (Dunn and Studier, 1973). However, it has been suggested that in the cell, these regions may normally be complexed with proteins (Jelenik et al., 1974).

Although RNA can form artificial complexes with cellular proteins in high salt (Girard and Baltimore, 1966; Baltimore and Huang, 1970), recent experiments by Pederson (1974a) have employed stringent controls to establish the identity of HnRNA protein (nRNP) particles in HeLa cells. These particles have been observed in cytological studies (Gall, 1956; Callan and Loyd, 1960; Stevens and Swift, 1966; Miller and Bakken, 1973), and can be isolated from different cell nuclei (Niessings and Sekeris, 1971b; Spirin 1969). Although the complexity of nRNP proteins is still in doubt (Kriekevskaya and Georgiev 1969; Faiferman et al., 1971; Albrecht and Van Zyl, 1973), both poly(A) synthetase activity (Niessing and Sekeris, 1973) and endonucleolytic activity (Niessing and Sekeris, 1970) have been found associated with the nRNP particle.

It is thus possible that either certain regions of the HnRNA, or proteins associated with these regions, may serve as site-specific signals, and direct the action of processing endonucleases, or be involved in the selection of certain RNA sequences for transport to the cytoplasm (Pederson, 1974b).

f) The organisation of unique and repetitive sequences in the genome

The information on the interspersion of unique and repeated sequences, is in close agreement with recent data on sequence organisation in the eukaryotic genome (Davidson et al., 1973; Graham et al., 1974). Partially hybridised DNA fragments of differing lengths can be bound to hydroxyapatite

after low C_0t incubations, and the melting behaviour, and nuclease sensitivity of these molecules studied. It appears that in both the sea urchin and Xenopus, just over half of the genome consists of alternating repetitive sequences of about 300 nucleotides long and non-repetitive sequences of 700 - 1100 nucleotides long. These non-repetitive sequences are long enough to code for many of the polypeptides that occur in the cell, and the repetitive sequences may be represented in the HnRNA, and subsequently degraded during nuclear processing.

5 to 15% of the genome consists of clustered, precisely repeated, sequences, that are probably located in the chromosome centromeres (Jones, 1970; Pardue and Gall, 1970). These sequences have been observed in many eukaryotic DNAs, and are sufficiently different in base composition from the bulk of the cell DNA, to appear as a separate 'satellite' band in a caesium chloride density gradient (Flamm et al., 1967). Satellite DNA appears to be of recent evolutionary origin (Walker and McLaren, 1965) and is not transcribed into RNA (Flamm et al., 1969).

The remaining 25-35% of the genome consists of long unique sequences, with very few interspersed repetitive sequence elements. These regions may be the location of other structural genes, and will also be transcribed.

These data are in substantial agreement with portions of the model for the regulation of gene activation proposed by Britten and Davidson (1969) and later expanded by them (Davidson and Britten, 1971, 1973). A major feature of the model is that the reiterated 'receptor' sequences will be adjacent in the DNA to gene sequences. Co-ordinated activation of non-contiguous genes may then occur by a single

activator molecule that can bind to these many receptor sequences.

This model does not require the transcription of large, polycistronic HnRNA, and a complete explanation for the size of the transcribed RNA has not yet been put forward. Recently Paul (1972) has suggested that the first part of the transcribed RNA may act as a polyanion, and assist in the unwinding of the nucleohistone supercoil as the RNA polymerase proceeds.

MATERIALS AND METHODS

MATERIALS

All chemicals used were supplied by British Drug Houses (AR grade) unless otherwise stated.

Isotopes

All isotopes were obtained from the Radiochemical Centre, Amersham, England.

Chromatography

Sephadex and Sepharose gel filtration media were obtained from Pharmacia, Upsalla, Sweden. Carboxymethyl cellulose, CM52, Carboxymethyl Sephadex C50 and diethyl-aminoethyl cellulose, DE52, and DE23 were supplied by Whatman, Maidstone, England.

Buffers

Trizma-HCl and Trizma base, HEPES (N-2-Hydroxyethyl-piperazine-N'-2-ethanesulphonic acid) and PIPES (Piperazine-N-N'-bis [2-ethanesulphonic acid]) were obtained from Sigma Chemical Co., St. Louis, Missouri.

Nucleic Acids

E. coli DNA, calf thymus DNA and poly(U) were obtained from Sigma, and A_7 and A_{10} from Bohringer, Germany. Poly(rA) and poly(rA)-(dT) were obtained from Miles, Illinois, U.S.A., and oligo$(dT)_{12-17}$ from P.L. Biochemicals, Milwaukee, Wisconsin.

Enzymes

RNase free DNase, Bacterial alkaline phosphatase and polynucleotide phosphorylase were obtained from Worthington Biochem. Co., Freehold, New Jersey, U.S.A. Ribonuclease T1 was

obtained from Sankyo Co. Ltd., Tokyo, Japan. Polynucleotide kinase was a kind gift from Dr. K. Murray (University of Edinburgh). Bovine serum albumin, catalase and pancreatic RNase A were obtained from Sigma.

Electrophoresis

Acrylamide and N.N'Methylenebisacrylamide were purchased from Serva, Heidleberg, N.N.N'N'-tetramethylethylenediamine (TMED) from Kodak Ltd., London, and ammonium persulphate from May & Baker, Dagenham, England.

Sucrose was purchased from Hopkins & Williams, Essex, England.

Diethyl pyrocarbonate (DEP) 'Baycovin' was a gift from Bayer Co. Ltd., London.

2.1 Determination of radioactivity

Radioactivity was measured in a Beckman LS-100 liquid scintillation counter, a Nuclear Chicago Mark I scintillation counter or a Packard Tri-Carb liquid scintillation counter. ^{32}P was measured by Cerenkov radiation. The efficiency of counting was almost 30%.

Radioactive samples on filters were dried, then counted in 10 mls toluene based scintillator (TBS) containing 0.5%, 2,5-diphenyloxazolyl (PPO, scintillation grade, Nuclear Enterprises, Edinburgh) and 0.03% 1,4 bis (2-(5-(phenyloxazolyl) benzene (POPOP, scintillation grade, Nuclear Enterprises, Edinburgh) in toluene (sulphur free from May & Baker, Dagenham, England).

Samples in aqueous solution were made up to 1 ml with water. 10 mls TBS/Triton X-100 (Lennig Chemicals, Croydon) in the ratio 2:1 were added (TXBS). The mixtures were shaken then counted. Corrections were made for counting efficiency using standard quench curves. The efficiency of counting of ^{3}H was 25 - 30% for the Beckman and Nuclear Chicago and 63% in the Packard counter. Aqueous samples of greater than 6 mls were counted in Instagel (Packard, Zurich, Switzerland) at an efficiency of 15 - 20%.

2.2 Estimation of DNA and RNA

A modified version of the Schmidt & Tannhauser method was used (Hutchinson & Munro, 1961). Tissue was extracted 3 times with 0.6 M PCA, and lipids removed with ethanol, ethanol/chloroform (3 + 1), ethanol/ether (3 + 1) and ether. The pellet thus obtained was hydrolysed in 0.3N NaOH at 37° for 90 minutes

and the DNA and denatured protein precipitated by the addition of concentrated PCA. RNA was estimated by measuring the A_{260} of the supernatant. The pellet was extracted in 0.6 M PCA at 90° for 20 minutes, spun, and the DNA estimated by measuring the A_{260} of this supernatant.

2.3 Preparation of mouse reticulocytes

In order to promote production of reticulocytes, anaemia was induced in Porton strain mice by one subcutaneous injection of 0.1 ml of 2.5% phenylhydrazine hydrochloride (Sigma), 0.16 M sodium acetate (pH 7.0) (Wintrobe, 1961). After 5 days the blood was collected into sterile balanced salt solutions (BSS) at 0° (Paul, 1965) containing 0.1 ml heparin/200 ml (Pularin heparin 5000 I.U./ml; Evans Medical Ltd., Liverpool), filtered through glass wool and centrifuged at 800 g for 10 mins. The plasma and buffy coat were removed by aspiration. The cells were resuspended in BSS and these steps repeated twice more, before storing the cells at -70°C.

A sample of the blood was taken for determination of the reticulocyte count. 10 drops of 1% brilliant cresyl blue in normal saline were incubated with one drop of blood for 15 mins. at 37°. A slide was prepared and the blue staining reticulocytes counted under oil immersion (Mag.x 1000) (Dacie, 1956). The count was usually between 50 - 70%.

2.4 Preparation of RNA from reticulocytes

During RNA preparation, care must be taken to prevent the action of contaminating ribonuclease (RNase).

Mammalian reticulocytes contain little endogenous RNase (Mathias et al., 1964). RNase contamination of the solutions on glassware used was inactivated by pretreatment with 0.05 %

b) Polysomal RNA

The post 20,000 *g* supernatant from the lysed reticulocytes was centrifuged at 150,000 *g* for 1 hour. The polysome pellet was resuspended in ANE, 0.5% SLS, and extracted as above. Alternatively, the polysomes were suspended in 10 mM Tris (pH 7.0), 10 mM KCl and dissociated by the addition of a half volume of 0.1 M EDTA, (pH 7.0).

2.5 Preparation of RNA from Mouse Foetal Livers

14 day pregnant mice were killed by dislocation of the neck. The embryos were removed into Hanks BSS at 0°, and the livers excised (Paul, 1965).

a) Total RNA

The livers were homogenised in ANE/SLS buffer and extracted with phenol/chloroform as above.

b) Cytoplasmic and nuclear RNA

Two methods for this were used:

1. Using a non-ionic detergent

The livers were disaggregated and the cells lysed by gentle homogenisation in DEP treated 0.14 M NaCl; 10 mM Tris (pH 7.4); 1.5 mM magnesium acetate; 0.5% Nonidet NP-40 (Borun *et al.*, 1967). Nuclei and membranes were pelleted at 20,000 *g* for 10 minutes. Microscopic examination of the pellet showed that most of the cells had been lysed by this procedure. The supernatant was removed, and total cytoplasmic RNA extracted from this with phenol/chloroform. The nuclear pellet was homogenised in ANE/SLS buffer and nuclear nucleic acid extracted with phenol/chloroform.

Diethyl pyrocarbonate (DEP). DEP was removed by boiling for 10 minutes or leaving for 16 hours at room temperature (Fedorisac et al., 1969).

The reticulocytes were lysed by the addition of 2 volumes of DEP treated 1 mM $MgCl_2$ at 0°. Cell debris and remaining white cells were pelleted at 20,000 g for 10 minutes and discarded.

a) Total RNA

The supernatant was made 0.5% with respect to sodium lauryl sulphate (SLS) and shaken with an equal volume of a phenol/chloroform mixture. (Equal volumes of 90% phenol, 0.1% 8-hydroxyquinoline, and chloroform, with the mixture saturated with 0.1 M sodium acetate; 0.1 M NaCl; 10 mM EDTA, (pH 6.0) (ANE buffer, Perry et al., 1972). After centrifugation at 20,000 g for 5 minutes, the aqueous phase was removed, and re-extracted with phenol/chloroform until there was no visible protein at the interface. To the final aqueous phase was added 1/10 volume of 30%, NaCl and 2 volumes of ethanol, and the RNA precipitated by standing for 16 hours at -20°.

The RNA was pelleted by centrifugation at 20,000 g for 15 minutes, washed twice with 95% ethanol, and recentrifuged. This pellet was dried with 95% nitrogen/5% oxygen and dissolved in 0.1 M NaCl; 10 mM-Tris HCl, (pH 7.4); 1 mM EDTA; 0.5% SLS, (NETS) and a sample scanned in the Unicam SP 800 spectrophotometer, to determine the concentration of RNA. It was assumed that a solution of 35 μg RNA/ml gives an absorbance of 1.0 at 260 nm. The RNA was stored at -20°.

2. Using citric acid

The livers were homogenised in 5.0 mM Tris (pH 7.5); 50 mM KCl; 1.5 mM $MgCl_2$; 0.25 M Sucrose (TKM - sucrose) and nuclei; membranes and unbroken cells pelleted at 20,000 *g* for 10 minutes. After addition of SLS to the supernatant, cytoplasmic RNA was extracted. The pellet was resuspended by homogenisation in 5 mls of 1.5% citric acid, 0.25 M sucrose and layered over 5 mls of 1.5% citric acid, 0.88 M sucrose (Busch and Smetana, 1970). This was centrifuged at 400 *g* for 10 minutes to pellet the dense nuclei, and unbroken cells and membranes discarded in the supernatant. This procedure was repeated twice or until nuclei were free of cytoplasmic tags and cell debris, as visualised by phase-contrast microscopy. The pellet was then homogenised in ANE/SLS buffer, and nuclear nucleic acid extracted with phenol/chloroform.

The nuclear nucleic acid extracted by either method was resuspended by homogenisation in 10 mM magnesium acetate; 25 mM NaCl; 2 mM $CaCl_2$; 0.1 M HEPES (pH 7.0) (HMC) (Birnie et al., 1973). 10 µg/ml DNase was added and incubated for 7 minutes at 4°. The solution was made 0.5% with respect to SLS, extracted once with phenol/chloroform, and the RNA precipitated. Fragments and nucleotides of DNA were removed by gel filtration through Sephadex G50 equilibrated in NETS. The excluded material was reprecipitated with ethanol and NaCl.

c) Polysomal RNA

Because of the high levels of RNase present in livers it was found necessary to modify the techniques for the preparation of polysomes. The cells were lysed in the Nonidet NP-40 buffer as before, but with the addition of 2 mg/ml yeast RNA (K. Light & Co. Ltd., Colnbrook, England). The 20,000 *g* supernatant was layered over a cushion of this buffer containing

2.5% sucrose and then spun at 150,000 g for 90 minutes. RNA was extracted from the pellets as before.

2.6 Preparation of RNA from adult mouse livers

Adult male mice were killed by cervical dislocation. The livers were perfused *in vivo* with 10 mls of TKM - sucrose at 37°, by injection from the portal vein. Any lobes that were not completely perfused were discarded. The livers were excised, cut into small pieces and washed 3 times with TKM - sucrose at 0°.

The pieces were then homogenised in TKM - sucrose, and cytoplasmic and nuclear RNA prepared by the citric acid method as before.

2.7 Preparation of RNA from adult mouse brain

Adult male mice were killed by cervical dislocation. The brains were excised into Hanks BSS at 0°, cut into pieces and washed 3 times. The pieces were gently homogenised in 1 mM Tris (pH 8.0); 5 mM $MgCl_2$; 0.5 M Dithiothreitol (DTT) (Sigma, U.S.A.); 3 mM $CaCl_2$ and left at 0° for 5 minutes. The solution was then made isotonic by the addition of 1/4 volume of 1.0 M sucrose, and spun at 1,000 g for 10 minutes (Gilmour and Paul, 1974). SLS was added to the supernatant and cytoplasmic RNA extracted.

The pellet was homogenised in 2.2 M sucrose and spun in a swing out rotor at 20,000 g for 1 hr. The pellet was homogenised in ANE/SLS buffer and nuclear RNA extracted.

2.8 Removal of salt from RNA samples

A column of G_{25} Sephadex, swollen in DEP treated water, was poured in a 10 ml syringe barrel (Gillette Scimitar, Middlesex, England). A GF/B and GF/C filter (Whatman) cut to

the correct size, were used to hold the column. Water was removed by spinning at 1,000 *g* for 15 minutes. The RNA, in a maximum volume of 1 ml and maximum salt concentration of 0.1*4* M, was layered on, and the column respun at 0^{o} at 1,000 *g* for 15 minutes using a DEP treated tube. The salt free sample was scanned in the SP800 and lyophilised if required. Between 95 - 98% of the A_{260} was recovered.

2.9 Purification and analysis of globin mRNA

a) Affinity chromatography on Poly(U)-Sepharose

Sepharose 6B was activated by stirring at 0^{o} with cyanogen bromide (Koch-Light Ltd., Colnbrook, England). The pH was kept constant during this step by the addition of NaOH. The sepharose was washed free of unreacted CnBr and then stirred overnight at 4^{o} with poly(U) (ave. Mol Wt. > 100,000) at 10 mg/ml in NETS and an equal volume of 0.1 M PIPES pH 6.0. After washing in NETS the poly(U) sepharose was stored in 50% glycerol and NETSAR (as NETS, but with sodium lauryl sarcosine replacing the SLS) at -20^{o} (Wagner et al., 1971; Adesnik et al., 1972). This material has a binding capacity of approximately 175 µg mRNA/ml sepharose.

The sepharose was poured as a slurry into a column packed with glass wool and washed extensively with NETS. RNA was applied to the column in this buffer, and unretained material washed off. Bound material was eluted in 90% formamide; 10 mM Tris HCl (pH 7.4); 0.2% SLS and precipitated with NaCl and ethanol. The column was regenerated by washing in NETS. Later work was carried out with commercially available poly(U) sepharose, using the same technique.

b) Affinity chromatography on oligo dT cellulose

Oligo dT cellulose (Collaborative Research, Waltham, Mass.) columns were poured as above and washed with 0.5 M NaCl; 10 mM Tris (pH 7.5); 0.2% sodium lauryl sarcosine. Samples were applied in this buffer and unretained material washed off. Bound RNA was eluted using 10 mM Tris (pH 7.5); 0.2% sarcosine. The column was washed with 0.1 N NaOH and re-equilibrated in binding buffer (Aviv and Leader, 1972).

2.10 Separation of RNA species by sucrose density gradient centrifugation

a) Centrifugation in swing-out rotors

10 - 30% gradients were prepared by layering 10, 17, 24 and 30% sucroses in TKM in the centrifuge tube and leaving overnight at 4° before use. 15 - 30% gradients in NETS were layered as 15, 20, 25 and 30% sucroses and left for 16 hrs at room temperature. 6.5 ml tubes were layered with a sample volume of 50 µl and 25 ml tubes with a sample volume of 300 µl.

Gradients were analysed by upward displacement with 40%, or 60% sucrose containing phenol red. Absorbance was continuously monitored at 254 nm using a Uvicord II absorptiometer (L.K.B. Instruments Ltd., London) and fractions collected with an Ultrorac (L.K.B).

b) Centrifugation in the zonal rotor

For handling large amounts of material, a titanium B-XIV zonal rotor was used in the MSE 65 superspeed ultracentrifuge. It was loaded as described by Anderson (1966) and the MSE publication, No.49. Convex and linear sucrose gradients were formed by an LKB ultrograd, and pumped into the rotor, light end first. A 40% sucrose 'cushion' was then introduced until the rotor was full. The sample, in 2.5% sucrose, was pumped into the centre of the rotor and positioned by an 'overlay'

of buffer. The rotor was sealed and accelerated from its loading speed of 2,500 rpm. Unloading was achieved by pumping 40% sucrose in from the edge of the rotor. The effluent from the centre pipe was monitored by passing through a variable path length flow cell (Perkin-Elmer Ltd., London) in a Unicam SP-800 spectrophotometer, with an attached servoscribe chart recorder.

B-XIV rotor content volumes (Capacity: 650 ml)

Gradient	560 ml
Sucrose cushion	30 ml
Sample	8 ml
Overlay	52 ml

Ribonucleoprotein particles from EDTA-treated polysomes were separated in a 15 - 30% (w/v) sucrose gradient in 10 mM Tris; 10 mM KCl, (pH 7.0).

2.11 Recovery of RNA

mRNP fractions were made 3% with respect to NaCl, and then precipitated by the addition of 2 volumes of ethanol. After standing at -20° for 16 hrs., the mRNP was then pelleted at 20,000 g for 15 minutes and RNA extracted as before.

2.12 Separation of RNA species by polyacrylamide gel electrophoresis

Electrophoresis of RNA was carried out according to the method of Loening (1967). Gels were run in 0.7 x 7 cm or 0.7 x 12 cm perspex tubes, with a buffer of 40 mM Tris (pH 7.8); 36 mM sodium dihydrogen phosphate; 1 mM M EDTA; 0.2% SLS. For 2.6% gels the methylenebisacrylamide was 5% that of the acrylamide. For higher concentration gels this was reduced to 2.5%. 2.6% gels contained 0.26 gm acrylamide; 0.013 gm bisacrylamide; 0.008 ml TMED and 0.08 ml 10% ammonium persulphate in 10 mls buffer. 6% gels contained 0.6 gm

acrylamide, 0.015 gm bisacrylamide, 0.016 ml TMED and 0.07 ml 10% ammonium persulphate in 10 mls buffer.

Samples were made up in 2.5% sucrose in buffer. Gels were pre-run for 1 hour at 17 v/cm. After sample application, electrophoresis was carried out for 20 minutes at 3 v/cm then for the required time at 17 v/cm. Electrophoresis was performed at room temperature. The gels were scanned at 260 nm with the Joyce-Loebel UV scanner using a liquid filter (0.001% p-dimethylamino-benzaldehyde, (Sigma) in methanol). RNA in gels was stained in 0.2% toluidine blue (Williamson et al., 1971) in water for 1 hour, and gels were destained overnight.

Estimation of radioactivity in gels

Gels were sliced directly into vials, using a Gibson automatic gel slicer (Villiers-le-Bel, France).

For ^{3}H counting, 1 ml of 10% piperidine in water (v/v) was added to each vial which was incubated at 60°C to evaporate the piperidine. 1 ml of distilled water was added to swell the gels, followed by 10 ml of TXBS.

2.13 Detection of mRNA activity

(a) by means of a duck reticulocyte cell-free system

The duck reticulocyte lysate system described by Lingrel (1971) was used to measure the incorporation of (^{14}C) L-leucine and (^{3}H) L-leucine into mouse α and β-globin chains. Two lysate incubation mixtures, one containing (^{3}H) leucine and RNA under test, the other containing (^{14}C) leucine but without added RNA, were incubated for 2 h at 25°C. Between 1 – 10 μg of mRNA were added, which is within the linear response range to mRNA for this system (Lingrel 1971). The mixtures were then cooled in ice and combined; 40 mg of mouse carrier globin was added and the protein extracted with acid-acetone at 20°, (0.15 N HCl),

pelleted, and washed twice with acetone. The pellet was dried under vacuum, and dissolved in a buffer containing 8 M urea; 50 mM mercaptoethanol to dissociate the mouse α and β-globin subunits, which were then separated on a carboxymethyl cellulose column (Whatman, CM 52 microgranular) using a gradient of 0.01-0.1 M Na_2HPO_4, (pH 6.9.) The effluent (30 ml/hour) was monitored with an LKB Uvicord II absorptiometer, at 280 nm, and 10 ml fractions collected. One drop of a 2% solution of bovine serum albumin was added to each fraction to aid precipitation of the proteins after the addition of an equal volume of 20% trichloroacetic acid (TCA). The precipitates were trapped on glass fibre discs (Whatman, GF/A), washed with 1% TCA then dried at 80°C for 1 hour. The proteins on each disc were solubilised by the addition of 0.05 ml of water followed by 0.5 ml of NCS reagent (Amersham/Searle Corporation), and incubated overnight at 37°C. To each vial was added 10 ml of TBS and the filter counted in a Nuclear Chicago Mark I liquid scintillation counter, adjusted for minimum overlap of ^{14}C into the ^{3}H channel.

(b) <u>by means of a rabbit reticulocyte cell-free system</u>

The rabbit lysate was prepared as described by Palmiter (1973). Assays were carried out in a reaction mixture of 250 microlitres containing 100 µl of a mixture of essential amino acids, ATP and an ATP generating system and 2.5 microcuries of ^{3}H isoleucine, 100 µl of lysate and 50 µl of the RNA solution under test. Between 2 - 6 picomoles of globin mRNA were added to each assay mixture, which is known to be in the linear response range of the system for added mRNA.

After incubation for 6 or 90 minutes at 26°, non-radioactive mouse globin was added as carrier, and total globin extracted and chromatographed on carboxymethylcellulose with a 10 mM - 50 mM phosphate gradient. Fractions were analysed as before.

(c) by means of a Krebs ascites cell-free system

A preincubated ascites cell free system was prepared as described by Mathews and Korner (1970); Mathews et al., 1972; Mathews 1972. Routine assays of 50 µl contained 15 µl preincubated ascites S-30, 10 µl of a cocktail, containing ATP, an ATP generating system, amino acids, and 5 µCi of ^{3}H leucine (30-60 Ci/mMole), 10 µl of 12.5 mM $MgCl_2$ and 0.5 M KCl, and 15 µl of the RNA solution being treated.

The dependence of incorporation on added globin mRNA was determined by incubation for 60 min at 37° followed by 10 min digestion with alkali. Protein was precipitated with 10% CCl_3COOH after addition of 50 µg of bovine serum albumin and the precipitates were collected onto Whatman GF/C filters, washed, dried, and counted in TBS.

Product analysis was performed on 100 µl incubation mixtures each containing 5 µg/ml of mRNA, incubated for 60 min at 37°. Carrier mouse globin was added, the protein extracted with acid acetone at -20°, and the α and β chains separated on CM cellulose as before.

2.14 Use of polynucleotide phosphorylase

(a) Preparation of ^{3}H Poly(A)

^{3}H ADP (1 mC/mMole) was incubated with 1.25 mg/ml polynucleotide phosphorylase (from Micrococcus lysodiekticus, EC 2.7.7.8) in 14 mM $MgCl_2$; 0.15 M Tris (pH 9.0); 28 mM ADP for 2 h at 37° (Grunberg-Manago, 1963). The reaction was stopped by the addition of SLS, the RNA deproteinised with phenol/chloroform, and precipitated with ethanol. ^{3}H poly(A) of a specific activity of 3.3×10^{3} cpm/µg was obtained.

(b) Assay of polynucleotide phosphorylase degredative activity

Assays were carried out in 200 µl, containing 3 µg of ^{3}H poly(A); 50 mM Tris (pH 7.5); 15 mM $MgCl_2$, usually in 15 mM potassium phosphate. After incubation at 37°, undegraded poly(A) was precipitated by the addition of 50 µl of bovine serum albumin (10 µg/ml) and 2 mls of 5% perchloric acid (PCA) at 0°. The assays were kept at 0° for 20 minutes, and then filtered through Whatman GF/A discs. Each filter was washed with 20 mls of 5% PCA, dried, and counted in TBS.

2.15 Digestion of globin mRNA with polynucleotide phosphorylase

Mouse globin mRNA (100 µg/ml) was incubated with 500 µg/ml of polynucleotide phosphorylase (Micrococcus lysodiekticus, (EC 2.7.7.8) in 50 mM Tris (pH 7.5); 15 mM $MgCl_2$; 15 mM potassium phosphate at 37° for 7 min. The reaction was stopped by adding one-tenth volume of 10% SLS and the reaction mixture was deproteinised by shaking with a half-volume of phenol/chloroform. The RNA was precipitated with ethanol.

2.16 Fingerprint analysis of RNA

Two-dimensional fingerprint analysis of RNA was performed after digestion for 45 min at 37° with pancreatic and T1 ribonucleases each at 0.05 mg/ml in 10 µl of 10 mM Tris (pH 7.4). The digestion was terminated by addition of an equal volume of water-saturated phenol-chloroform (1:1) and the aqueous phase was dried down after deproteinisation. The oligonucleotide mixture was dissolved in 1% sodium dodecyl sulphate containing 0.1 mg/ml of bacterial alkaline phosphatase and incubated for 30 min at 37°. The reaction mixture was deproteinised, dried down, and then taken up in 1 µl of 0.1 M Tris, 11 mM mercaptoethanol, 11 mM $MgCl_2$; 0.27 mM ^{32}P ATP, and 2 units of polynucleotide kinase. Incubation was carried

out at 37° for 45 min. The mixture was dried, and the ^{32}P-labelled oligonucleotides were taken up in 0.7 µl of dye mixture and separated by electrophoresis on cellulose acetate strips in the first dimension and by chromatography on PEI-cellulose thin-layer plates in the second dimension (Southern and Mitchell, 1971; Crossley et al., 1974).

After the plates had run they were auto-radiographed, using Kodirex X-ray film (Kodak). Individual oligonucleotide spots were cut out and counted.

2.17 Preparation of DNA

DNA was prepared using a method adapted by Hell et al., (1972) from the earlier work of Britten et al. (1969).

a) From foetal livers

Foetal livers were suspended in 20 vol 8 M urea; 0.24 M sodium phosphate buffer, (pH 6.8) (MUP), containing 1% (w/v) SLS and 10 mM EDTA. This was blended in a sealed, filled container with an MSE homogenizer run at 14,000 rpm for 6 periods of 14 s each, alternating with 30 s cooling in ice. The homogenate was poured into a thick slurry of hydroxyapatite (HAP)(Bio-Rad Bio-Gel HTP, Richmond, California) in MUP, and the mixture was stirred gently and allowed to stand at room temperature for about 1 hr. (The amount of HAP used varied according to the capacity of the particular batch of HAP available). The slurry was poured into Pyrex No.3 sintered-glass funnels (approximately 30 gm HAP per funnel of diameter 9.5 cm) and the liquid was drawn off under a vacuum of 5 psi. The HAP was washed free of RNA and protein with MUP until no further material was eluted as judged by A_{260} and A_{280} measurements on the effluent (approximately 1,500 ml/30 gm HAP),

then with a further 500 ml of MUP. Urea was removed by washing with 14 mM sodium phosphate, (pH 6.8)(approximately 500 ml required/30 gm HAP). The removal of urea was monitored by refractive index measurements on the effluent. The DNA was eluted with 0.4 M sodium phosphate, (pH 6.8).

The DNA solution was dialysed for 3 x 16 hours against 10 volumes of 0.01 M NaCl at 4° and then decalcified by passage through Dowex 50 Na^+ (Bio-Rad, analytical-grade cation exchange resin AG 50W - X8, 1 gm resin/50 mg DNA). The preparation was freeze dried and taken up in water to give a DNA concentration of about 2.5 mg/ml and NaCl concentration of 0.1 M. Aliquots of 5 mls were sonicated in a Dawe sonicator using the microprobe, (6, 10-s treatments punctuated by 30-s cooling periods at 0°C). After precipitation with alcohol, the DNA fragments were fractionated on an alkaline surcose gradient in an MSE BXIV zonal rotor or a 3 x 25 ml rotor, and fragments of the appropriate size were obtained and purified.

b) DNA from mouse sperm

Mouse vasa deferentia were excised into balanced salts solution (Paul, 1965). The sperm were freed by gently stroking the vasa deferentia with a glass rod. The sperm were counted in a counting chamber (x 250 magnification).

Initial isolation of the DNA requires reduction of the protein-sulphur linkages of the keratin-like membrane of the sperm head (Borenfreund et al., 1961). The sperm were centrifuged at 8,000 g for twenty minutes and washed twice with 0.15 M NaCl; 15 mM sodium citrate, once with ethanol and once with a 1:1 ethanol, ether mixture. After each wash the sperm were pelleted by centrifugation at 8,000 g for thirty minutes. To each 100 mg sample was added 20 mls 0.15 M NaCl; 15 mM sodium citrate; 0.25 M mercaptoethanol. The sample was incubated for 2 hours at 4°C with gentle stirring.

10 mg trypsin were added and the incubation continued for 1 hour at room temperature.

The pellet obtained after the sample was centrifuged at 12,000 g for 1 hour was re-extracted with 10 mls buffer and 5 mg trypsin as before. DNA and protein were precipitated from the combined supernatants by the addition of two volumes of ethanol.. The pellet was dissolved in MUP containing SLS and EDTA, and DNA purified as for the foetal liver.

2.18 Sucrose density gradient fractionation of DNA

a) In swing-out rotors

Alkaline gradients for denatured DNA contained 0.9 M NaCl, 0.1 M NaOH. To determine the molecular weight of small DNA fragments, 5 - 10% (w/w) sucrose gradients were used. Samples were sedimented through either 14 ml gradients in the MSE 6 x 15 swingout rotor, or through 25 ml gradients in the MSE 3 x 25 swingout rotor, in MSE superspeed 65 or 50 ultracentrifuges. Fractions (0.75 or 1.0 ml, respectively) were collected by upward displacement by 20% (w/w) sucrose in 1 M NaCl. The sedimentation coefficients were determined using tables of McEwen (1967) and, from them, molecular weights were calculated using equations given by Studier, (1965). Fractions were neutralised, after the addition of 1/40th volume of 1.0 M HEPES (pH 7.0), and one drop of phenol red, by the dropwise addition of 1N HCl. DNA was precipitated with ethanol and desalted through Sephadex G25.

b) Large-scale fractionation in a zonal rotor

A titanium BXIV zonal rotor was loaded with a 500 ml 5-20% (w/w) sucrose gradient, linear with rotor radius. The gradient was prepared in apparatus similar to that described by Birnie

and Harvey (1968). The mixing vessel contained 600 ml 5% (w/w) sucrose to which was added 500 ml (w/w) sucrose. The sucrose solutions contained 0.9 M NaCl; 0.1 M NaOH. The sample (10 ml) of fragmented DNA in 0.9 M NaCl; 0.1 M NaOH was overlaid with 100 ml 0.9 M NaCl; 0.1 M NaOH. The rotor was spun at 25,000 rpm for 17 hr at 20°. 20 ml fractions were collected, neutralised and treated as above.

2.19 Purification of reverse transcriptase from avian myeloblastosis virus

The method is essentially as described by Kacian et al., (1971).

a) Purification of virus

The virus was purified from chicken plasma by spinning the particles onto a glycerol pad. These were suspended in 0.15 M NaCl, 10 mM Tris HCl (pH 8.8) and the procedure repeated twice, discarding the supernatant each time. The particles were then taken up in a minimum volume of 0.15 M NaCl, 10 mM Tris–HCl (pH 8.5), 1 mM EDTA and the volume adjusted to give a protein concentration of 4–5 mg/ml.

b) Isolation and purification of enzyme

The virus particles were lysed by the addition of 1/10th volume Nonidet NP–40, 1/10th volume of 10% deoxycholate in water and 3/10th volume of 4 M KCl. The debris was removed by centrifuation and discarded. The supernatant was diluted 10–fold with 10 mM potassium phosphate, (pH 7.2), 2 mM dithiothreitol (DTT), 10% glycerol, and loaded onto a DE52 column equilibrated with the same buffer. The column was washed with three volumes of the loading buffer and then with four column volumes of 50 mM phosphate (pH 7.2), DTT, glycerol, before being eluted with two volumes of 0.3 M

phosphate (pH 7.2), DTT, glycerol. The eluate was monitored with an LKB Uvicord absorptiometer at 280 nm and fractions containing protein were pooled. The protein washed off with 50 mM phosphate was discarded. The protein eluted with 0.3 M phosphate was diluted with two volumes of 10 mM phosphate (pH 8.0), DTT, glycerol, before being loaded on a carboxymethyl Sephadex column, C50, equilibrated in this buffer.

This column was washed with 0.1 M phosphate (pH 8.0), DTT, glycerol. The column was then eluted with 0.3 M phosphate, (pH 8.0), DTT, glycerol. Fractions were collected, made 50% with respect to glycerol and assayed for enzyme activity. Active fractions were pooled, and stored at -20°.

c) Assay of enzyme activity

Assays were performed in a total volume of 120 µl containing: (1) 1 µg poly(rA)-poly(dT) being used as a template for 10 µl of enzyme solution; (2) 5 µCi of ^{3}H dTTP; (3) 200 µM dTTP, dATP, dCTP and dATP (Bohringer, London); (4) 10 µg catalase; (5) 50 mM KCl, 50 mM Tris-HCl, (pH 8.2).

After incubation at 37° for 30 minutes the mixture, to which EDTA and E. coli DNA were added to give final concentrations of 2 mM and 50 µg/ml respectively, was chromatographed on a Sephadex G50 column above a pad of Chelex 100 (Bio-Rad Laboratories), equilibrated with 0.1 M NaCl. The excluded peak was pooled and counted.

The activity of the enzyme solution was calculated as follows:

$$\text{Activity} = \text{total } \frac{\text{d.p.m. in product}}{4.4 \times 10^{5}} \times 100 \text{ units/ml.}$$

(Kacian et al., 1972).

2.20 Preparation of cDNA complementary to globin mRNA

cDNA was prepared as described by Harrison et al., (1972a, b) by incubating for 90 min at 37°C the following mixture: 40 μM–d ^{3}H TPP (8 to 10 Ci/mmol; 40 μM–d ^{3}H –CTP (8 to 10 Ci/mmol); 500 μM–dATP; 500 μM–dGTP; 20 μg actinomycin D/ml (Sigma); 1 to 2 μg oligo$(dT)_{12-18}$/ml (PL Biochemicals); 5 to 10 μg 9S RNA/ml; 100 μg catalase/ml; 50 mM–Tris (pH 8.2); 50 mM–KCl; 10 mM dithiothreitol; 5 mM–magnesium acetate; 0.2 vol. reverse transcriptase in 50% glycerol; 0.15 M potassium phosphate (pH 8.0). The specific activities of d ^{3}H TTP and d ^{3}H CTP were always adjusted to be equal.

Tritium–labelled deoxyribonucleoside triphosphates were supplied in 50% ethanol containing 1% ammonium bicarbonate and stored at –20°. Bicarbonate was removed from these radioactive solutions, after adding the requisite unlabelled deoxyribonucleoside triphosphates, by mixing them with an excess of Dowex resin (hydrogen–form, AG 50W x 8, which has been washed extensively with 50% ethanol) at 0°, followed by removal of the resin by centrifugation (Harrison et al., 1972b). The resultant supernatant solution was evaporated to dryness under vacuum. All these manipulations, and incubation of enzyme with substrates, were performed in glassware coated twice with Repelcote (Hopkins & Williams, Essex, England) and then rinsed with distilled water.

After incubation the mixture was chromatographed on Sephadex G50 as in a normal reverse transcriptase assay.

The excluded fraction was adjusted to contain 100 μg/ml E. coli DNA and precipitated with ethanol. The precipitate was then dissolved in 1 ml of 0.9 M NaCl; 0.1 M NaOH and layered on a linear 25 ml 5–10% (w/w) sucrose gradient containing 0.9 M NaCl and 0.1 M NaOH.

The gradient was centrifuged at 29,000 rpm in an MES 3 x 25

swing-out rotor at 20°C for 24 h; 1 ml fractions were collected in tubes containing 10 μg of E. coli DNA. 5 microlitres of each fraction were counted in TXBS, and the molecular weight of cDNA in each fraction was determined from S values calculated according to McEwen. The higher molecular weight fractions were pooled, neutralised and passed through Sephadex G-50 - Chelex 100 equilibrated in distilled water.

Amounts of cDNA were calculated on the basis of the quoted specific activities of the triphosphates, assuming that cDNA consists of 50% (C + T).

In later experiments only ^{3}H dCTP was used and the specific activity calculations adjusted accordingly.

2.21 RNA - DNA hybridisation

cDNA was mixed with the appropriate amount of salt free RNA and incubated at 43° in hybridisation buffer - 0.5 M NaCl; 25 mM HEPES; 0.5 mM EDTA; (pH 6.8); 50% formamide (Fluka, Switzerland) (Harrison et al., 1974a).

The salt solutions, before addition of formamide, were passed through Chelex 100 and then DEP treated. Usually the hybridisation solution contained 500 μg/ml E. coli RNA.

The mixture was incubated for the appropriate period at 43° in a sealed repelcoted capillary, usually in a volume of 10 μl.

2.22 Analysis of hybridisation using a single-strand specific nuclease

a) Preparation of nuclease (S_1)

S_1 nuclease was obtained as described by Sutton (1971), 1 g of Takadiastase powder (Park Davis Ltd., London) was dissolved in 5 ml of distilled water and dialysed overnight at

4° against two hundred of 10 mM phosphate buffer, (pH 7.0). The sticky precipitate was discarded, while the dialysed Takadiastase solution was applied to a 2 cm x 1 cm diameter column of DEAE–cellulose (Whatman DE–23) previously equilibrated at 4° with 10 mM sodium phosphate buffer, (pH 7.0) (starting buffer). This column was immediately rinsed at 4° with successive 50–ml portions of starting buffer supplemented with (a) no addition; (b) 50 mM NaCl; 10 mM KH_2PO_4; (c) 0.1 M NaCl; 20 mM KH_2PO_4; (d) repeat of (c), and (e) 0.2 M NaCl; 40 mM KH_2PO_4.

4 ml fractions were collected from the final rinse, each being immediately supplemented with an equal volume of chilled 50% (v/v) glycerol and cooled to 0°. Each fraction was rapidly assayed for protein concentration ($A_{280\ nm}$) and nuclease S_1 activity, and then stored at −20°.

b) Assay of enzyme activity

Aliquots of denatured ^3H DNA from mouse LS cells (a gift from Dr. G. D. Birnie) were incubated at 37° with dilutions of each enzyme fraction, in nuclease assay buffer; 50 mM – sodium acetate (pH 4.5); 2 mM $ZnSO_4$; 0.1 M NaCl and 10 µg denatured calf thymus DNA/ml. After 30 minutes a portion of the incubation mixture was taken to determine the total radioactivity; a further portion was acidified (1 N–$HClO_4$) at 4°C after addition of mouse DNA and bovine serum albumin (5 µg and 50 µg, respectively) and then centrifuged for 15 min at 2500 revs/min. A portion of the supernatant fluid was then taken to determine the acid–soluble radioactivy.

c) Assay of hybridisation of cDNA

Before use in the hybridisation assay, each fraction of enzyme was diluted with 0.2 M NaCl, 40 mM KH_2PO_4, 50% glycerol, so that 20 µl would render 50% of the ^3H DNA acid soluble, in 30 minutes (G. Birnie, personal communication).

The hybridisation reaction was stopped by flushing out the capillaries with 250 µl of nuclease assay buffer. This was incubated with 100 µl of S_1-nuclease for 2 h at 37°. 100 µl of the incubation mixture was counted to determine total radioactivity and 200 µl was precipitated by the addition of 50 µl of 6.0 N PCA, and 50 µl of carrier bovine serum albumin (2 mg/ml). This was centrifuged for 15 min at 2500 revs/min. and 200 µl counted, to determine the proportion of acid soluble counts.

For hybridisation reactions larger than 10 µl, the volumes used were adjusted, reducing the final formamide concentration to below 5%, to prevent inactivation of the S_1 enzyme.

2.23 Annealing of cDNA to DNA, and analysis on hydroxyapatite

The appropriate amounts of cDNA and mouse or E. coli DNA (at about 2 mg/ml) were mixed in water, lyophilized and redissolved at 10 mg/ml before adding 0.2 vol. 0.72 M-sodium phosphate (pH 6.8). DNA concentrations were calculated on the basis of absorption measurements on acid-hydrolysed portions of the dilute DNA solutions. The cDNA/DNA mixtures, either undiluted or 100 fold diluted, were denatured in a boiling water bath for 10 min, after which they were annealed at 60°. The reaction was stopped by the addition of 10 vols of 0.1 M KCl and the samples stored at -20°.

Analysis was carried out on hydroxyapatite (Bio-Rad, Biogel HTP batch 10936) using a ratio of 1 mg DNA to 3 ml packed hydroxyapatite, as described by Harrison et al., 1972a. Colums were eluted with 8 ml of 30 mM-phosphate, twice with 8 ml of 0.16 M-phosphate, and finally twice with 4 ml of 0.4 M-phosphate. Absorbance measurements of the fractions of eluate were made, before counting in Instagel.

2.24 Melting behaviour of hybrids

a) Using hydroxyapatite

After loading the hydroxyapatite with the reannealed duplexes and washing with 30 mM-phosphate, the column was eluted 3 times with 8 ml of 0.16 M-phosphate, at 5 degree intervals from 60 to 95°C. Finally, the column was eluted with 0.4 M-phosphate; this removed a negligible amount of DNA and radioactivity. The absorbance at 260 nm was measured, and the samples then counted in Instagel.

b) In solution

After a suitable period of annealing, in a sealed capillary, the solution containing hybrid was diluted several fold with formamide/hybridisation buffer, and aliquots resealed in capillaries. These were heated for 5 minutes at the appropriate temperature, before being rapidly cooled and diluted with 25 volumes of nuclease assay buffer. S_1 analysis of these samples was then as previously described.

RESULTS

3.1 Isolation and characterisation of 9S RNA from mouse reticulocytes

a) Preparation of total RNA

500 mice routinely gave a yield of 90 ml of packed cells, after washing 3 times in BSS. This also removed the majority of white blood cells. The 1 mM $MgCl_2$ used to lyse the reticulocytes leaves most of the remaining white cells intact, and RNA prepared from the 20,000 g supernatant was essentially free of DNA contamination. 30 - 40 mg of total RNA was isolated from 40 mls of packed red blood cells.

Figure 1(a) shows the banding pattern of total RNA when run on 2.6% polyacrylamide gels; the major bands correspond to the 28S, 18S and 4S and 5S RNA fractions. Between the 18S and 5S two minor bands of 12S and 9S RNA can be distinguished. A third minor band running at 7S can sometimes also be seen. 9S RNA accounts for between 1 - 2% of the total RNA of the reticulocytes.

b) Isolation of 9S RNA using affinity chromatography

Both affinity chromatography methods used during this work gave essentially similar results. The RNA that was retained by the poly(U) sepharose or oligo(dT) cellulose was highly enriched in 9S RNA, as shown in Figure 1(c). Usually the percentage of 9S RNA was increased to 40 - 50% by one passage through a column, and to $>$ 90% by a second passage. A sample of this purified 9S RNA when run on 2.6% acrylamide gels migrates as a single diffuse peak (Figure 1(d)). The molecular weight of this RNA is 220,000, (about 650 nucleotides). This value was obtained by comparison with standard RNAs on gel electrophoresis in denaturing conditions, or by analytical

2.6% polyacrylamide gel electrophoresis of:

a) 90 µg of total reticulocyte polysomal RNA.

b) 90 µg of polysomal RNA not-retained by poly(U) sepharose (first cycle).

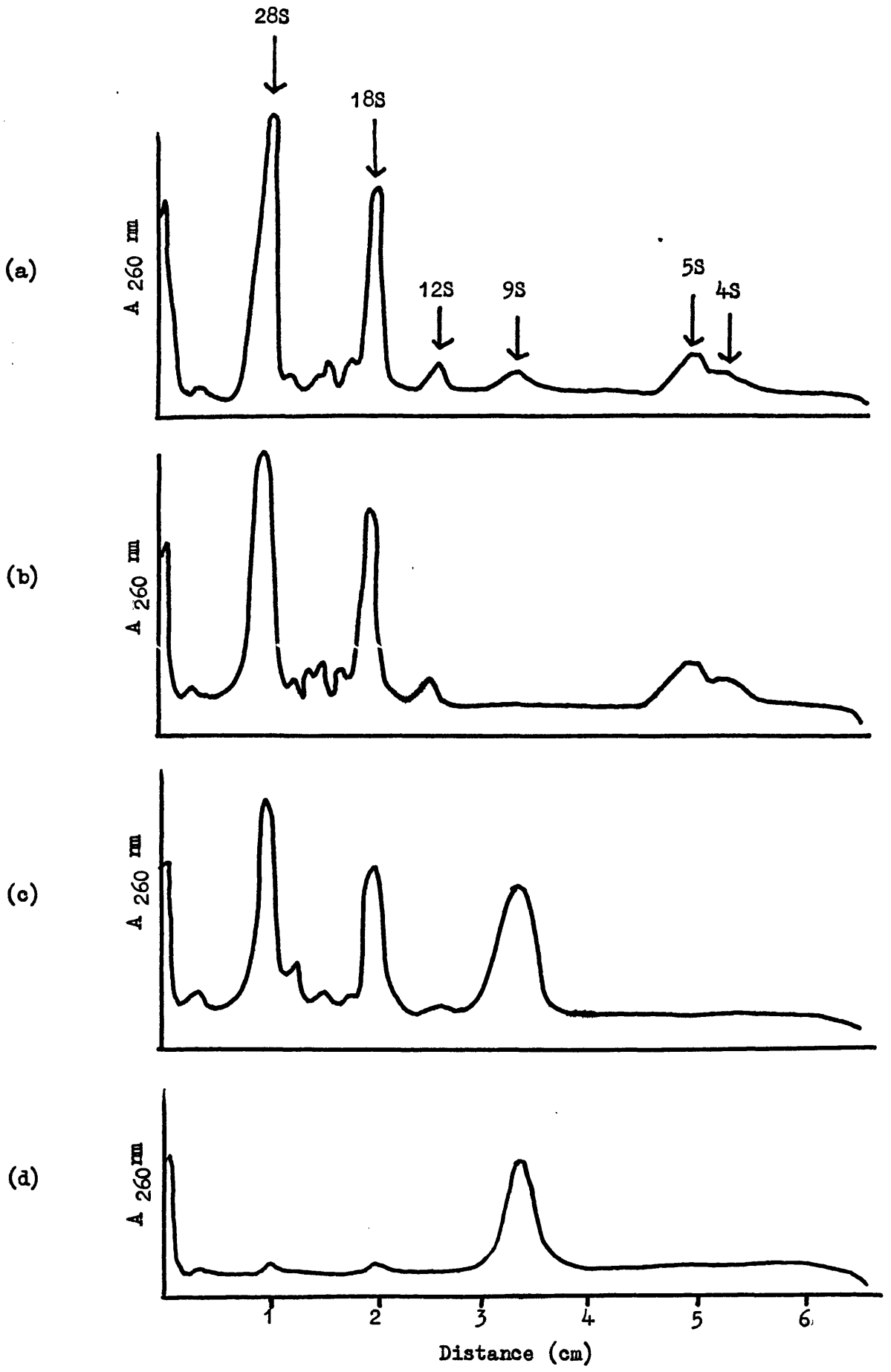
c) 30 µg of retained RNA (first cycle).

d) 5 µg of retained RNA (second cycle).

Gels were 0.7 cm diameter, and 7 cm long.

Electrophoresis was for 15 minutes at 2 mA/gel, and then 70 minutes at 6 mA/gel, 17 v/cm. Gels (a) and (b) were scanned at one half times magnification, and gels (c) and (d) at two times magnification on a Joyce-Loebel UV scanner.

Figure 1


28S
18S
12S
9S
5S
4S
(a)
(b)
(c)
(d)
A 260 nm
A 260 nm
A 260 nm
A 260 nm
1
2
3
4
5
6
Distance (cm)

centrifugation (Williamson et al., 1971). About 80% of the theoretical yield of 9S RNA was recoverable from polysomal RNA using these techniques.

Elution of some batches of poly(U) sepharose with formamide resulted in the release of small amounts of poly(U). It is known that the addition of poly(U) to polysome systems results in the synthesis of polyphenylalanine (Nirenberg and Matthaei 1961; Arnstein et al., 1962) and the reduction in the synthesis of endogenous proteins (Weinstein et al., 1963; Marcus et al., 1963). Thus for some experiments oligo(dT) cellulose was used for the purification of 9S RNA.

c) Isolation of the 14S mRNP containing 9S RNA using zonal ultracentrifugation

EDTA treatment of polysomes results in their dissociation into ribosomal subunits with release of a 14S mRNA-protein (mRNP) complex. The sedimentation profile of EDTA treated mouse polysomes in the B-XIV zonal rotor is shown in Figure 2. Three main peaks were resolved. These were haemoglobin, the 40S ribosomal subunit, and the 60S ribosomal subunit. The mRNP can be isolated as a small peak sedimenting between haemoglobin and the 40S subunit, at about 14S. Between 0.25 - 0.4 mg 9S RNA was recovered from 30 - 40 mg polysomes centrifuged in the B-XIV rotor.

d) Assay for messenger activity

The ability of 9S RNA to direct the synthesis of mouse α and β globin was tested in a duck reticulocyte cell-free system. Figure 3 shows the 280 nm absorbance profile from a CMC column with mouse α and β globin resolving clearly from the two duck globins.

FIGURE 2.

Zonal ultracentrifugation of EDTA-treated mouse reticulocyte polysomes.

40 mg of polysomes in 10 mM KCl; 10 mM Tris (pH 7.4); 0.1 M EDTA, layered on to a 10 - 30% sucrose gradient in TKM buffer, in an M.S.E. B-XIV titanium rotor, spun at 46,000 rpm for 3 hrs at 7°.

Figure 2

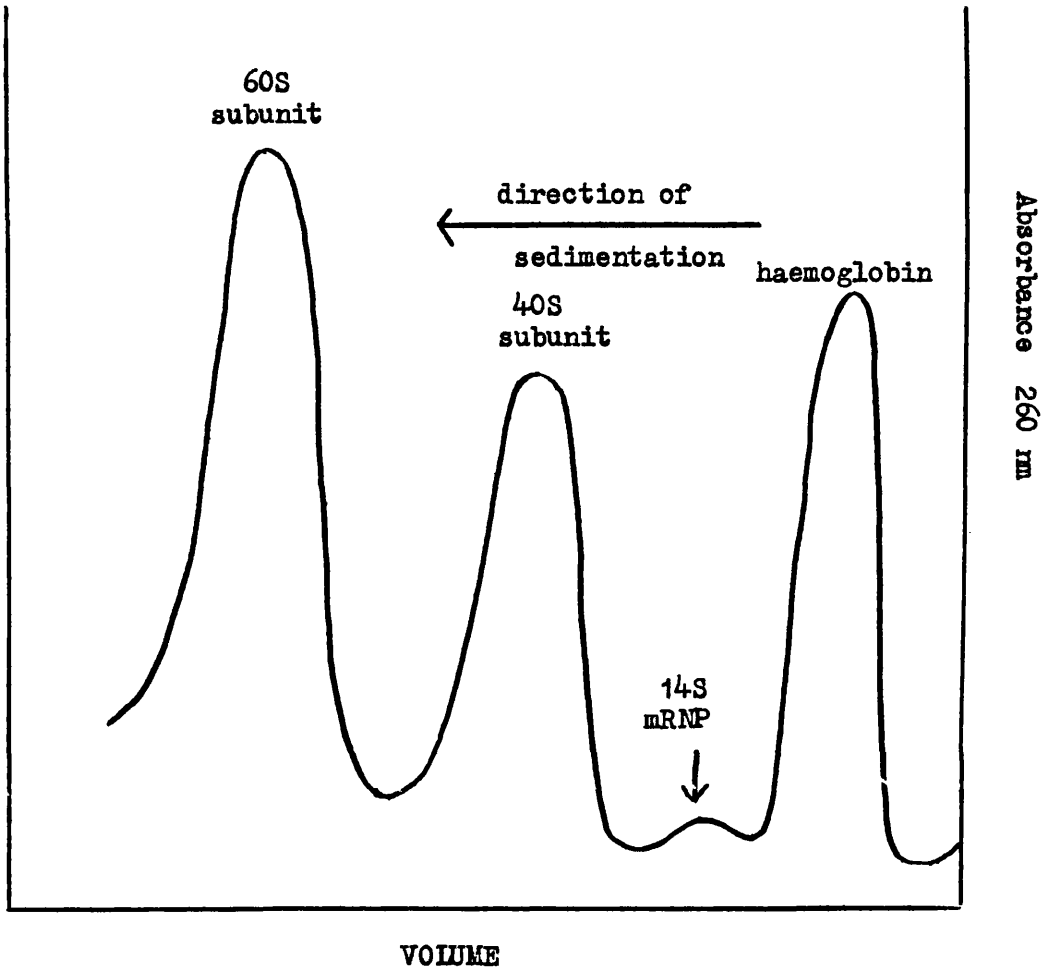

60S
subunit
direction of
sedimentation
haemoglobin
40S
subunit
14S
mRNP
Absorbance 260 nm
VOLUME

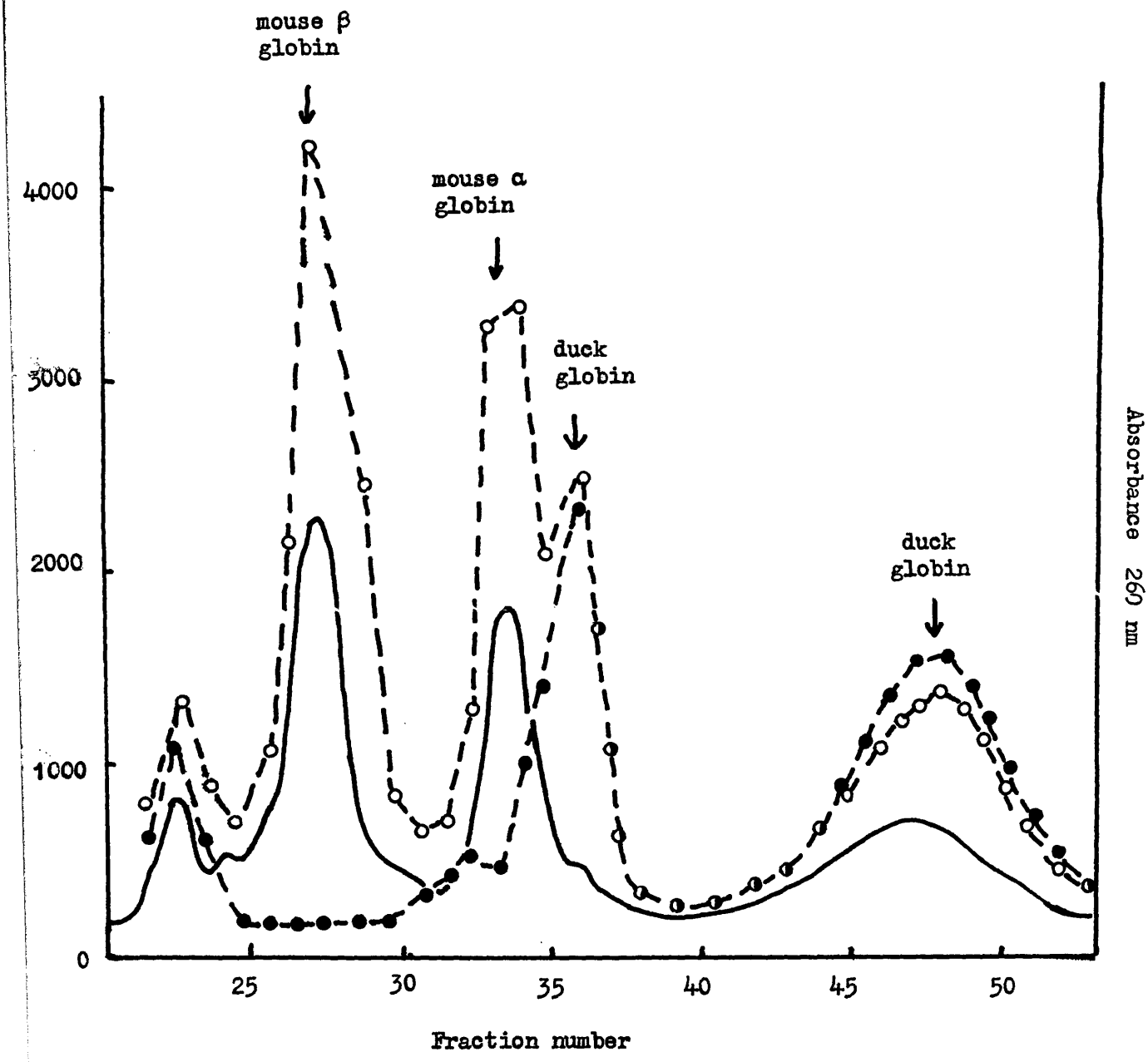
FIGURE 3.

Product analysis of the globin chains synthesised in a duck reticulocyte cell-free system, with added mouse globin mRNA.

7 µg of mouse reticulocyte 9S RNA were incubated in a 250 µl assay, with ^{3}H leucine. A parallel 250 µl assay was incubated with ^{14}C leucine, without added mouse message. The incubations were combined, the globin extracted, and the sample chromatographed on carboxymethyl cellulose, as described. The absorbance of the column eluate was monitored at at 280 nm (———). The fractions were prepared for counting as described:

—o—o—, ^{3}H leucine cpm; —●—●—, ^{14}C leucine cpm.

Figure 3


mouse β globin
mouse α globin
duck globin
duck globin
4000
3000
2000
1000
0
25
30
35
40
45
50
Fraction number
Absorbance 260 nm

A stimulation of leucine incorporation into mouse α and β globin can be seen on addition of 9S RNA to the duck lysate.

3.2 Synthesis and characterisation of complementary DNA

a) In the presence of oligo(dT) primer and actinomycin D, the rate of synthesis of cDNA by reverse transcriptase was proportional to the concentration of 9S RNA in the range of 1 - 10 μg per ml. Under the conditions routinely used, the yield of high specific activity cDNA (25×10^6 to 40×10^6 disints/min per μg) was about 20% of the amount of 9S RNA used. No labelled DNA was synthesised when reverse transcriptase was incubated with the oligo(dT) primer alone, using all four labelled deoxyribonucleoside triphosphates as precursors. This indicates that the dependence of reverse transcriptase on added template is absolute. When sedimented on a 5 - 10% alkaline sucrose gradient, the mean molecular weight of cDNA was found to be 110,000 (330 bases) which is about 50 - 60% of the length of 9S RNA (560 - 660 bases; Williamson et al., 1971)(Figure 4).

For most experiments, only the higher molecular weight cDNA fragments (greater than about 240 bases) were isolated and characterised further. The size distribution of the cDNA used in most of the experiments is shown in Figure 10.

Purified cDNA was incubated at 43° for two weeks in hybridisation buffer. The product was virtually all degraded (95%) by S_1 nuclease. Therefore, cDNA contained a negligible fraction of self-complementary sequences.

b) Hybridisation of complementary DNA to 9S RNA

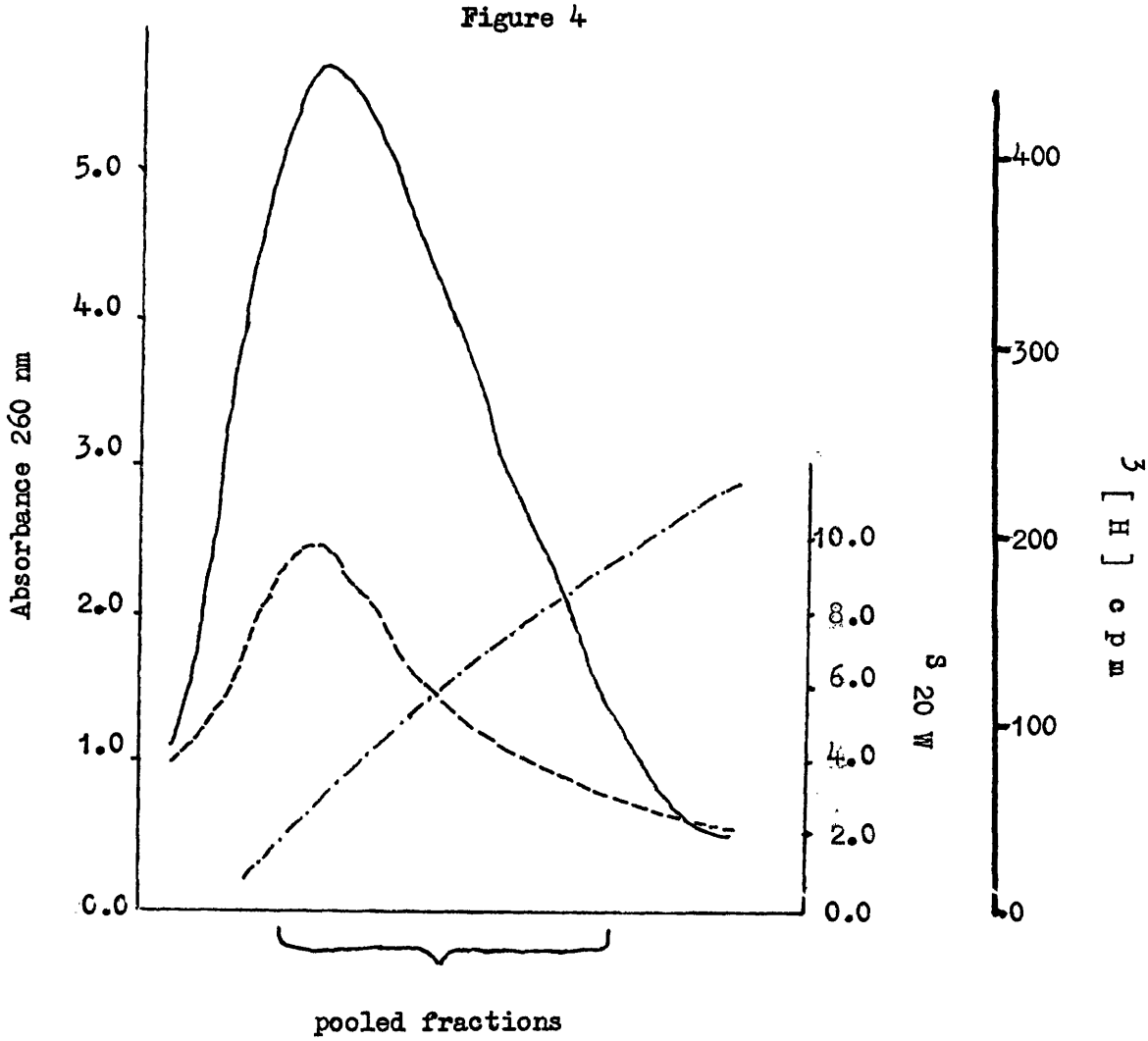
In order to minimise degradation of RNA during hybridisation, cDNA was annealed (usually at 43°) with 9S

FIGURE 4.

Sedimentation profile of crude cDNA on 5 – 10% alkaline sucrose gradients.

Samples were centrifuged at 29,000 rpm in an M.S.E. 3 x 25 swing-out rotor, for 18.5 hours at 20°. 1 ml fractions were collected, and 5 µl of each counted ——— . The A_{260} of marker DNA run in a parallel gradient was read– – – –. The molecular weight was calculated from the S value —·—·— ard fractions 5–15 pooled.

Figure 4


Absorbance 260 nm
5.0
4.0
3.0
2.0
1.0
0.0
10.0
8.0
6.0
4.0
2.0
0.0
s 20 w
400
300
200
100
0
η [H] o p m
pooled fractions

RNA in hybridisation buffer containing formamide and excess E. coli RNA. Under these conditions, the T_m of a 40% (G + C) RNA-DNA hybrid is 58° (Bishop, 1972). No hybridisation of cDNA to E. coli RNA was detected (Figure 5). However, about 80 to 85% of the cDNA was converted into hybrid by incubation with excess 9S RNA (Figure 5). Neither the amount of cDNA hybridised nor the rate of ribonuclease-resistant hybrid formation was greatly affected by annealing at different temperatures (37, 43 or 48°). Furthermore, the sensitivity of the hybrids to ribonuclease was not affected by annealing at different temperatures. Thus, the annealing conditions chosen appear to be within the optimum range and are sufficiently stringent.

In order to test the quality of the 9S RNA-cDNA hybrid directly, the size of the cDNA in such hybrids was estimated before and after treatment of the hybrids with S_1 nuclease. The size distribution of cDNA was not affected by the annealing conditions themselves (Figure 6). Under conditions in which cDNA alone was completely degraded, cDNA in the 9S RNA-cDNA hybrid was reduced to about half its initial size by such treatment, indicating that on average about one nick per hybrid molecule was introduced by the treatment with S_1 nuclease (Figure 6). This indicates that there are no extensive regions of S_1 nuclease-sensitive cDNA distributed throughout the cDNA-9S RNA hybrid. However, this hybrid had a T_m of 69° in 0.16 M phosphate (Figure 7), somewhat lower (3°) than predicted for a perfect hybrid.

FIGURE 5.

Hybridisation of cDNA to reticulocyte 9S RNA, ●—● , reticulocyte polysomal RNA + — + , and E. coli ribosomal RNA O—O.

0.001 µg of cDNA was incubated with increasing amounts of RNA, in 10 µl of formamide/hybridisation buffer to a $D_o t$ of 0.2 mol l^{-1} sec. (7 days) at 43°. Analysis of the hybrids was as described.

Figure 5

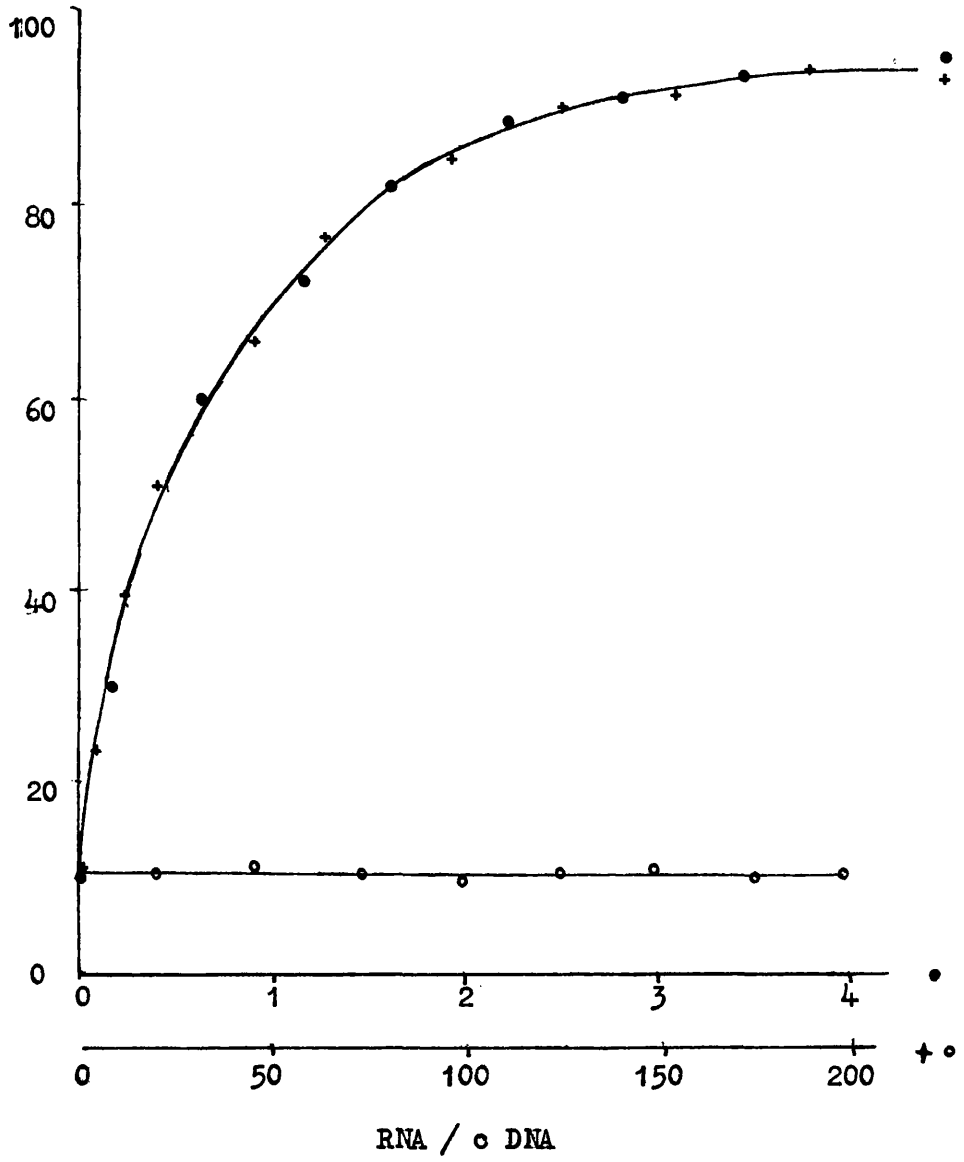

100
80
60
40
20
0
% o DNA HYBRIDIZED
0
1
2
3
4
0
50
100
150
200
RNA / o DNA

FIGURE 6.

Sedimentation of cDNA before and after hybridisation, on a 5 – 10% alkaline sucrose gradient.

Samples were centrifuged at 29,000 in an M.S.E. 3 x 25 swing-out rotor for 24 hours at 20°. 1 ml fractions were collected and 0.5 ml counted; –Δ–Δ–, cDNA from cDNA–9S RNA hybrid; –O–O–, cDNA from cDNA–9S RNA hybrid after treatment of hybrid with S_1 nuclease; –▲–▲–, cDNA before hybridisation.
A marker DNA of 4.35 (about 200 nucleotides) was included in one of the gradients. The molecular weight was calculated as described in the text –●–●–

Figure 6

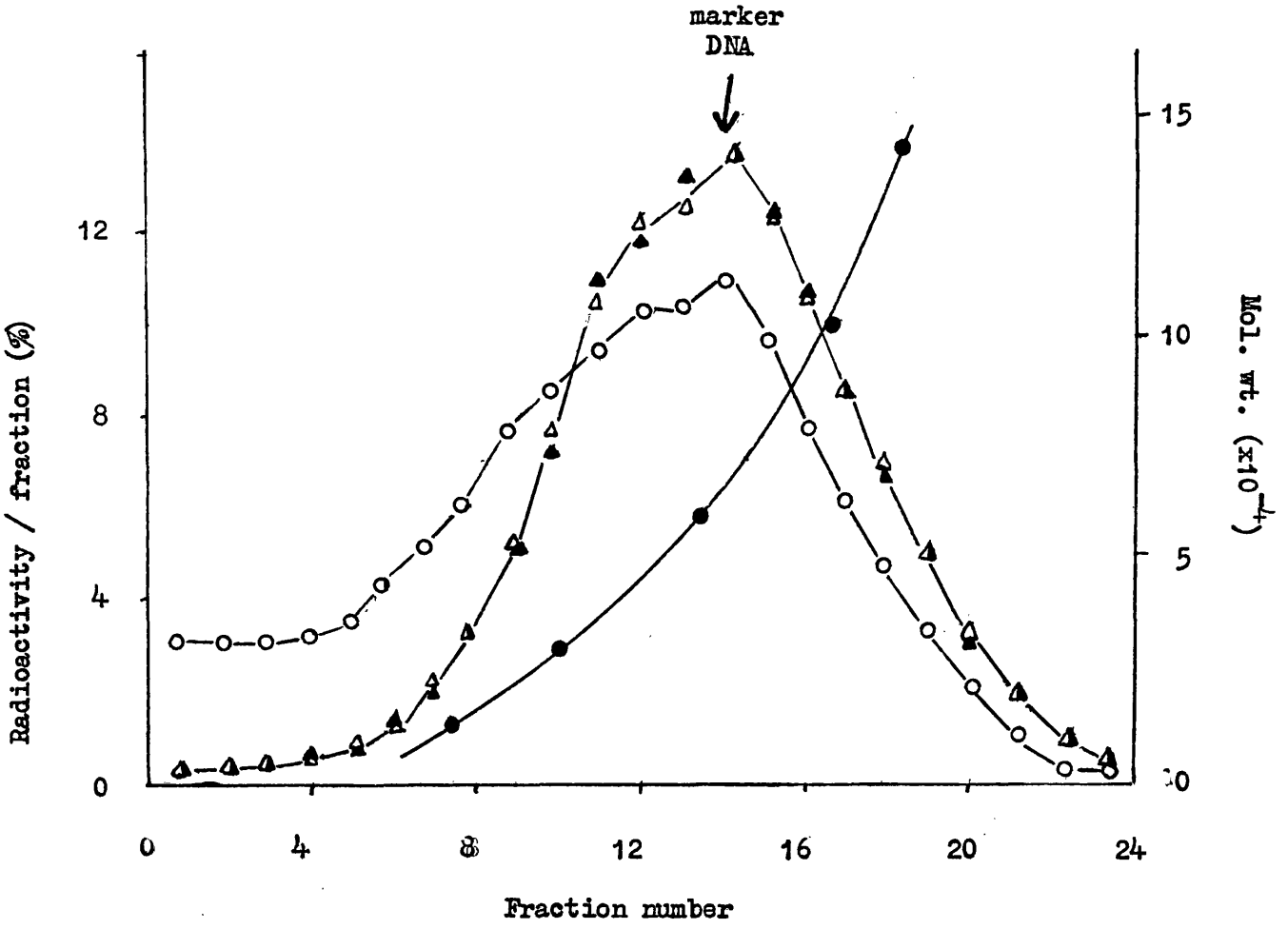
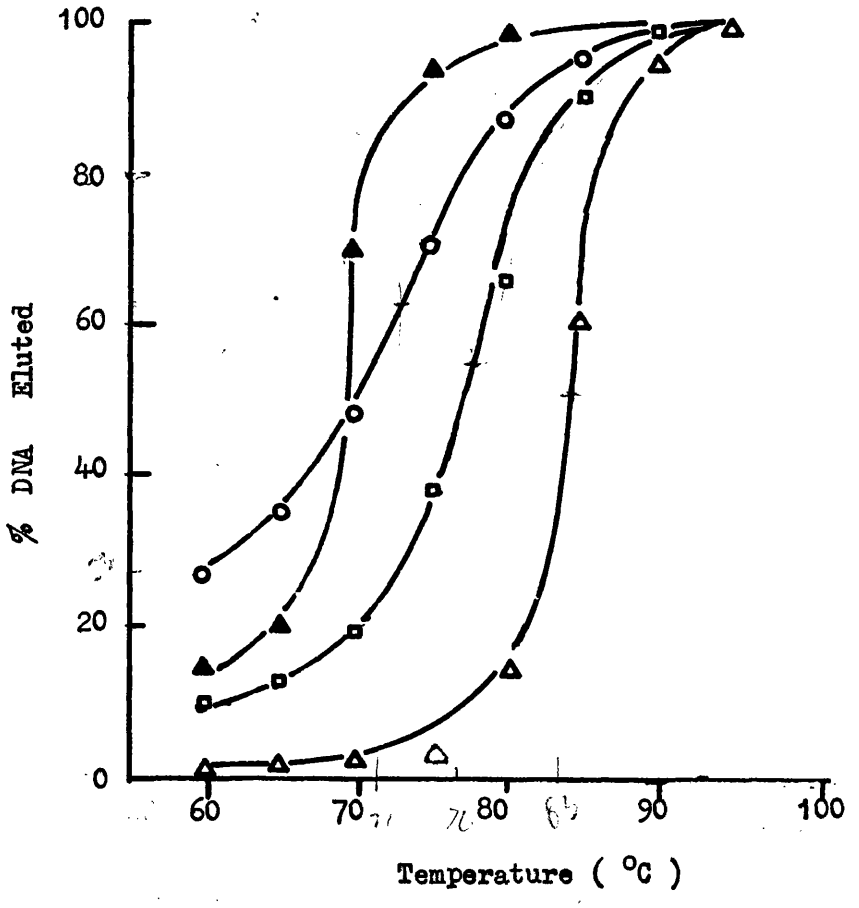

marker
DNA
Radioactivity / fraction (%)
Mol. wt. (x10⁻⁴)
Fraction number
0
4
8
12
16
20
24
15
10
5

FIGURE 7.

Melting behaviour of reannealed duplexes.

Melts were carried out on HAP and unretained material eluted with 0.16 M phosphate; —▲—▲—, cDNA—9S RNA hybrid; —Δ—Δ—, native foetal liver DNA (330 nucleotides long); —□—□— reannealed foetal liver DNA; —o-o— cDNA annealed to foetal liver DNA.
Reannealed DNA samples were incubated in 0.12 M phosphate to a $C_o t$ value of 20,000.

Figure 7



c) Analysis of complementary DNA-RNA titration curves

Hybridisation of cDNA to RNA may be used as a general technique to estimate the fraction of an RNA preparation that is complementary to cDNA. The simplest approach is to titrate cDNA with increasing amounts of RNA, under conditions that allow completion of hybridisation reactions. If cDNA were of homogeneous size, the titration curve, namely the plot of the fraction of cDNA hybridised versus the RNA/cDNA ratio, would consist of a line through the origin for cDNA excess and a horizontal line for RNA excess. The proportion of the RNA preparation that is complementary to cDNA could be found from the RNA/cDNA ratio at the point of intersection, because at that value there would be an equal number of cDNA and complementary RNA molecules present. In practice, however, cDNA is heterogeneous in size, so that in cDNA excess there is a greater probability of hybridisation to larger cDNA molecules.

Young et al. (1974) have shown that the kinetics of the hybridisation of cDNA to 9S RNA are in agreement with second-order reaction theory. However, the conclusions drawn from such titration experiments are valid only if all reactions used to establish the plots are complete. It is, therefore, an important prediction of second-order kinetics that the slowest reaction occurs when equal amounts of the two components are present. Reactions are characterised in terms of $D_o t$, where D_o is the initial cDNA concentration, and t is the time of incubation in seconds.

Data presented by Young et al., indicates that the slowest cDNA-RNA reaction has $D_o t_{\frac{1}{2}} = 5.5 \times 10^{-3}$ mol l^{-1}.

This is the value of $D_o t$ when half of the cDNA is converted to duplex, and, hence, incubation to 40 times this value (0.22 mol l^{-1}s) should ensure completion of all hybridisation reactions.

In order to establish that sufficiently large incubation periods were used, two titration curves using reticulocyte 9S RNA were obtained by annealing to $D_o t = 0.22$ and $D_o t = 1.7$. The titration curves were identical to that obtained in Figure 5, and it is apparent that incubation to $D_o t = 0.22$ is sufficient to ensure completion of the hybridisation reactions. This result also demonstrates that RNA degradation is not a problem during hybridisation up to $D_o t = 1.7$. This is important, because it is assumed for the theoretical analysis that all RNA molecules remain larger than the largest cDNA molecules. No appreciable degradation of cDNA occurs during hybridisation (Figure 6).

Figure 5 shows titration curves in which constant amounts of cDNA (0.001 µg) were hybridised to increasing amounts of reticulocyte 9S RNA and reticulocyte polysomal RNA. correcting for cDNA heterogeneity (Young et al 1974)
With 9S RNA, maximum levels of hybridisation were obtained at input RNA/cDNA ratios of 1-1.4, consistent with the finding that about 60% of the 9S sequence is represented in cDNA. With reticulocyte polysomal RNA maximum hybridisation is obtained at input ratios of about 50 which is consistent with the known 9S content of reticulocyte polysomal RNA (2%).

A small fraction of cDNA (about 10%) was apparently incapable of forming a hybrid with 9S RNA (Figure 5). Data from Harrison et al. (1974a) demonstrates that this fraction also does not hybridise if isolated and rehybridised. It is thought that transcriptional errors by the reverse transcriptase, results in the formation of these sequences, which are not found in 9S RNA (Harrison et al., 1974a).

3.3 Annealing of complementary DNA to DNA

a) Preparation of DNA

In order to simplify the interpretation of the results of these experiments, the various mouse and *E. coli* DNAs were each sonicated and fractionated to give a size distribution of fragments similar to that of cDNA itself (Figure 10). 90 mg of foetal liver DNA were fractionated using a 10 – 40% alkaline sucrose gradient in the BXIV zonal rotor, as shown in Figure 8. The size of DNA from the fractions indicated were checked in analytical gradients. From these data, fractions were pooled and 50 mg of DNA were obtained, with a size range between 4.9 – 7.7S and a peak size of 6.6S (Figure 10).

10 mg of DNA were obtained from approximately 0.5 gm (dry weight) of mouse sperm. This was fractionated on 5 – 10% alkaline sucrose gradients in a 3 x 25 swing-out rotor (Figure 9). 5 mg of DNA, with a size range between 4.9 – 7.7S and a peak size of 6.6S (Figure 10), were pooled from the gradient fractions.

b) Annealing of cDNA to DNA

Samples of these DNAs were mixed with cDNA in a ratio (cDNA:DNA) of $1:8 \times 10^6$ or $1:16 \times 10^6$ (w/w), denatured and allowed to anneal to various $C_o t$ values.

These experiments used DNA fragments with an average length of 330 bases (Figure 10). Under these conditions, *E. coli* DNA reassociated with a $C_o t_{\frac{1}{2}}$ value of about 6 (Figure 11a). Using mouse (foetal liver and sperm) DNA fragments of the same size, a small fraction (about 10%) reassociated at a low $C_o t$ value, whereas the bulk of the foetal liver and sperm DNA fragments reassociated with a $C_o t_{\frac{1}{2}}$ value of about 800 (Figure 11, a & b).

FIGURE 8.

Zonal ultracentrifugation of foetal liver DNA.

90 mg of sonicated foetal liver DNA were spun at 25,000 rpm for 17 hours at 20° on a 10 – 40% alkaline sucrose gradient in an M.S.E. B–X1V titanium rotor. 20 ml fractions were collected, the A_{260} of each measured. –●–●–, and neutralised as described.

Aliquots from fractions 9, 12, 14,17, 20 and 24 were taken for accurate size analysis on 14 ml, 5 – 10% alkaline sucrose gradients, –▲–▲–. From this information fractions of between 4.9 – 7.8S (12–16) were pooled from the zonal gradient.

Figure 8

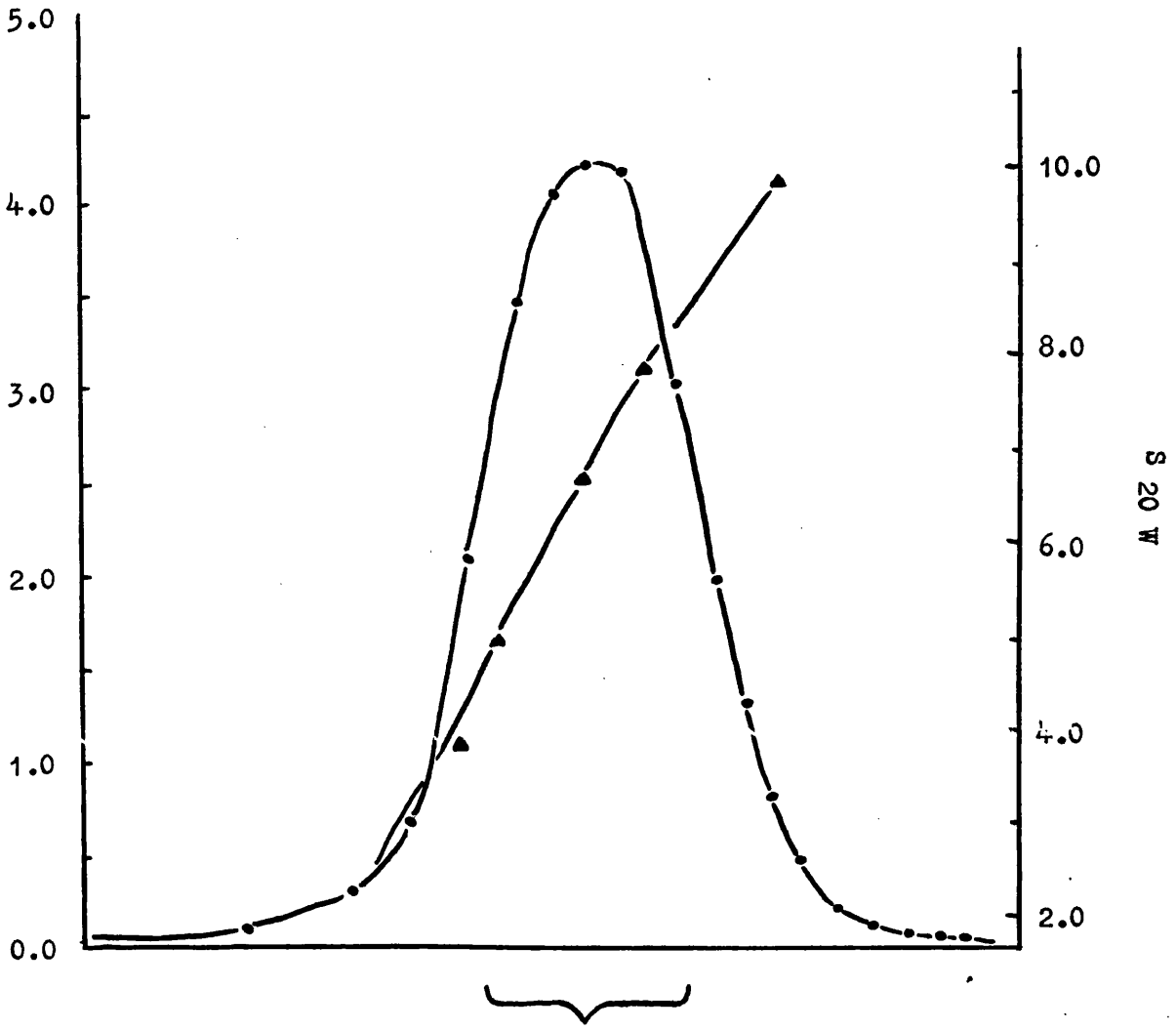


FIGURE 9.

Sedimentation of sonicated sperm DNA on a 5 – 10% alkaline sucrose gradient.

5 mg of DNA were centrifuged at 29,000 rpm in an M.S.E. 3 x 25 swing-out rotor for 17 hours at 20°. 1 ml fractions were collected, and the A_{260} measured. -●-●-. The S_{20_W} value was calculated -▲-▲-, and fractions 7 – 11 were pooled.
The peak sedimentation of a 4.5S marker DNA run in a parallel gradient is shown.

Figure 9

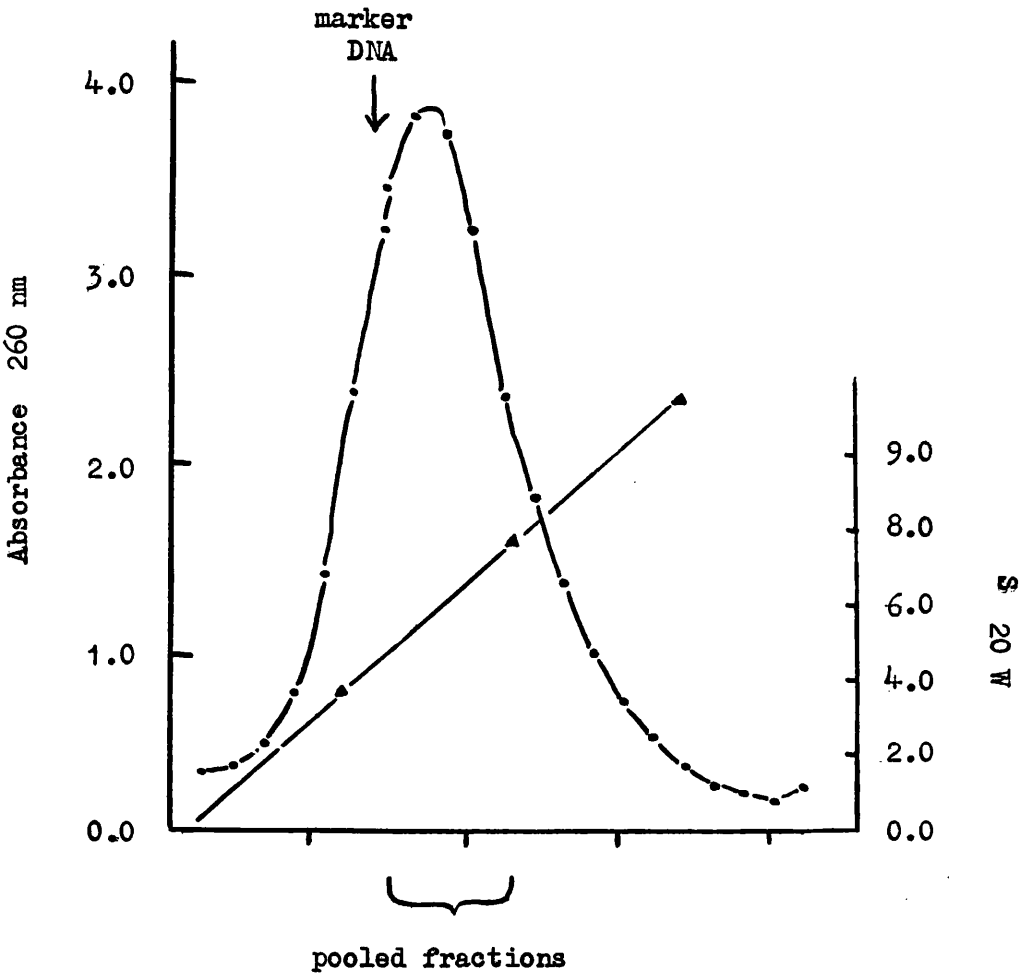

marker
DNA
4.0
3.0
2.0
1.0
0.0
Absorbance 260 nm
9.0
8.0
6.0
4.0
2.0
0.0
S 20 W
pooled fractions

FIGURE 10.

Sedimentation of cDNA and mouse DNA fragments in 5 – 10% alkaline sucrose gradients.

Samples were centrifuged at 29,000 in an M.S.E. 3 x 25 ml swing-out rotor for 19 hours at 20°; -·-·-· , 70 µg of sperm DNA; ----- 60 µg of foetal liver DNA; ····· 1500 cpm of cDNA; molecular weight -▲-▲-

Figure 10

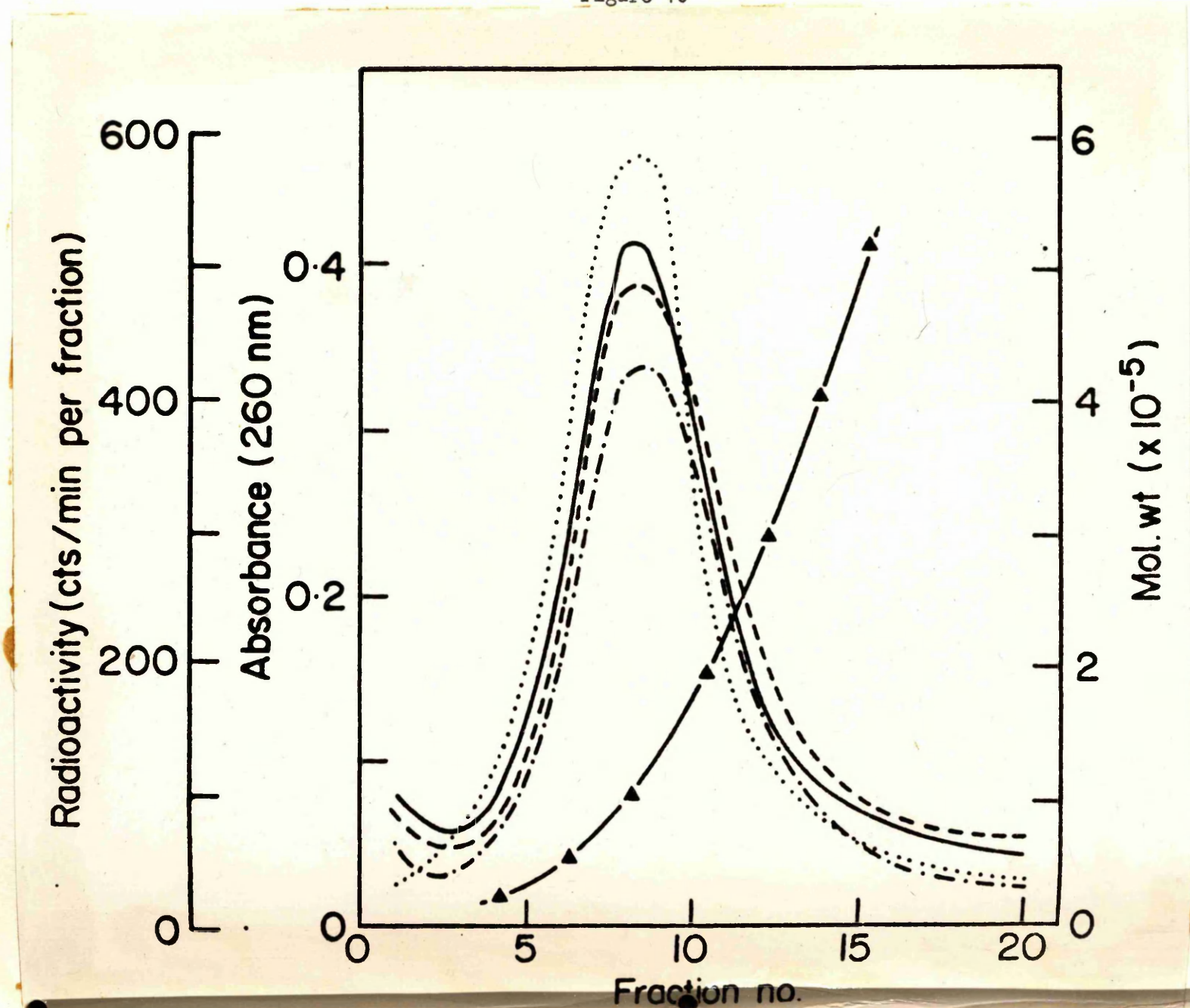

Radioactivity (cts/min per fraction)
600
400
200
0
Absorbance (260 nm)
0·4
0·2
0
0
5
10
15
20
Fraction no.
6
4
2
0
Mol. wt (x 10⁻⁵)

FIGURE 11.

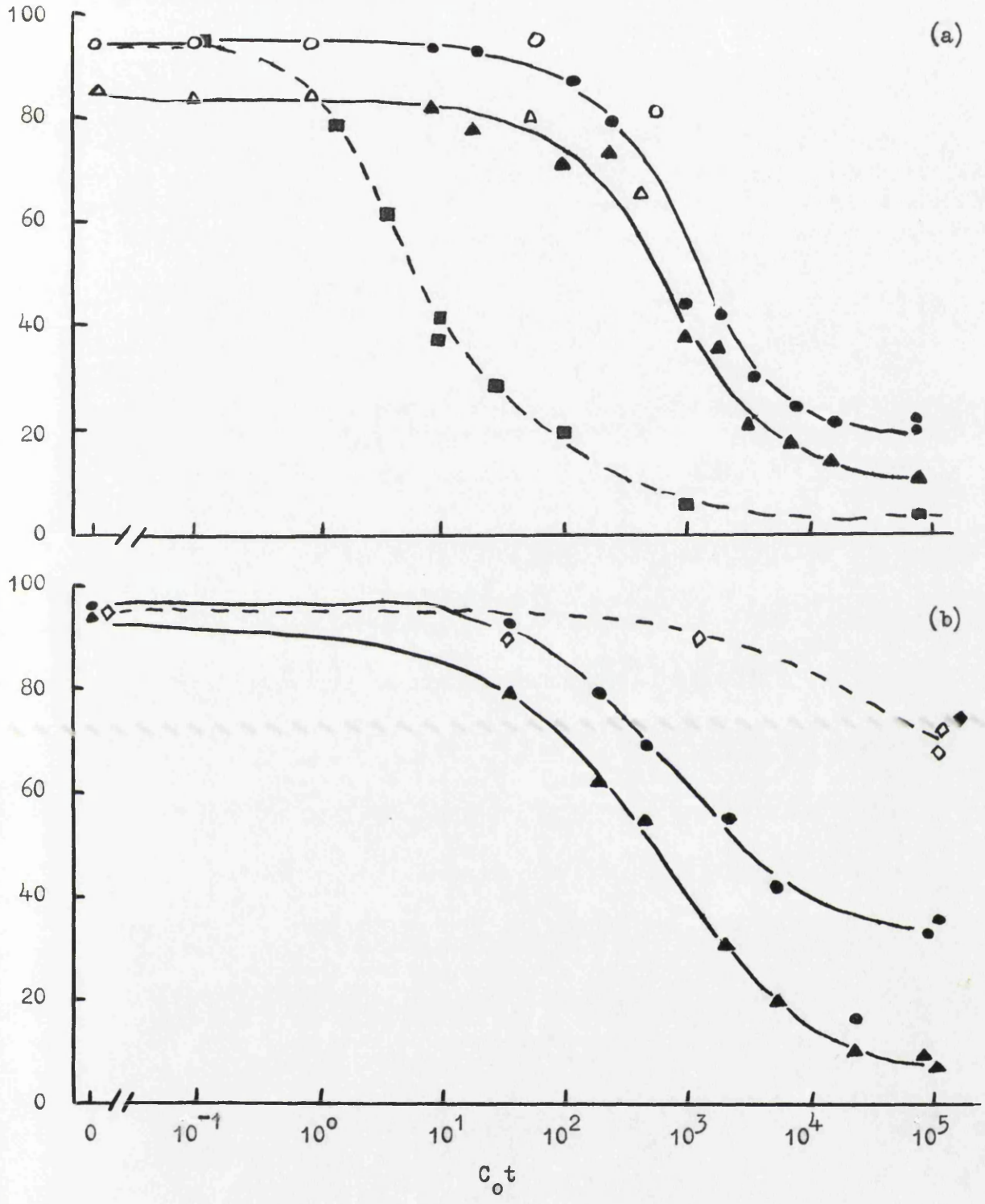
Incubation of cDNA with E. coli and mouse DNA, and reannealing of E. coli and mouse DNA fragments, measured by chromatography on hydroxyapatite.

a) □ reannealing of E. coli DNA/fragments; ◇ incubation of cDNA with E. coli DNA; ◆ incubation of poly (dT) with mouse DNA fragments; Δ ▲ reannealing of foetal liver DNA fragments; o ● annealing of cDNA to foetal liver DNA fragments.

b) ▲ reannealing of sperm DNA fragments; ● annealing of cDNA to sperm DNA fragments.

In a) the ratio of mouse DNA to cDNA was 1.6×10^7:1 (w/w), and in b) 8×10^6:1 (w/w). Open symbols in (a) refer to experiments in which the cDNA concentrations were diluted 100-fold.

Figure 11


(a)
(b)
100
80
60
40
20
0
DNA eluted by 0.16M phosphate (%)
0
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
$C_o t$

From the data in Figure 11, and knowing the complexity of the cDNA, the complexity of the sequences in mouse DNA that anneal to cDNA can be calculated. This will give an independent estimate of the number of globin genes in each tissue studied. The approach assumes that the cDNA that remains unannealed at high C_0t values represents excess cDNA sequences, above the number of complementary sequences present in mouse DNA.

The ratios of cDNA : foetal liver DNA, or sperm DNA that were used, were, $1 : 1.6 \times 10^7$ and $1 : 8 \times 10^6$ respectively. However, since 5% of the cDNA is S_1 resistant after low C_0t incubations, and 10% of the cDNA never forms a hybrid, 85% of the cDNA added can take part in the reaction. From this value, and the proportion of cDNA unannealed at high C_0t values (Figure 11), the proportion of the genome complementary to cDNA can be calculated to be;

$$\frac{0.85}{1.6 \times 10^7} \times \frac{85-20}{20} = 1.73 \times 10^{-7} \text{ for foetal liver}$$

$$\frac{0.85}{0.8 \times 10^7} \times \frac{85-35}{35} = 1.54 \times 10^{-7} \text{ for sperm DNA.}$$

Since the double-stranded molecular weight of the haploid mouse genome is 1.8×10^{12}, these values indicate that cDNA is complementary to 310,800 daltons of foetal liver DNA, and 271,800 daltons of sperm DNA.

The cDNA has a complexity of between 700-800 bases (210,000-240,000 daltons) (Young et al., 1974). Thus, the cDNA could be complementary to between 1.5 - 1.3 copies of the globin gene in the foetal liver DNA, and 1.3 - 1.1 copies in the sperm DNA. The cDNA is thus hybridising to a similar, low number of complementary sequences in both DNAs studied.

As a control experiment, cDNA was incubated with E. coli DNA. Only about 10% of the cDNA annealed by a $C_o t$ of 1000, during the period when all the E. coli DNA reassociated (Figure 11b). However, after incubation to a $C_o t$ of 80,000 a further 20% of the cDNA was not eluted from hydroxyapatite by 0.16 M phosphate. Evidence that this is not due to self-annealing of the cDNA sequences has been given in a previous section. Caution must be exercised in interpreting this result, since when high specific activity poly (dT) was incubated with mouse DNA to a $C_o t$ value of 80,000, about 30% of the poly(dT) behaved similarly.

With DNA from both mouse tissues studied, cDNA annealed at a $C_o t_{\frac{1}{2}}$ value identical to that for the reannealing of the bulk of the mouse DNA sequences (Figure 11, a & b). These relative rates of annealing of cDNA to mouse DNA, and reannealing of mouse DNA fragments would be distorted if:- (1) there were preferential degradation of cDNA during annealing, relative to the mouse DNAs, or (2) if mismatching of cDNA sequences occurred with respect to the globin genes in the mouse DNA.

(1) Samples of cDNA and mouse DNA fragments (initial size 330 bases) were taken for size determinations after incubation to a $C_o t$ value of 8000, when the annealing reaction was essentially complete. Negligible degradation of either cDNA or mouse DNA fragments was detected at this time (Figure 12), although after incubation to a $C_o t$ value of 80,000, some degradation of cDNA and mouse DNA fragments occurred. Thus, neither the observed rate of annealing of cDNA to mouse DNA nor rate of annealing of mouse DNA fragments requires correction for changes in size of fragments during annealing.

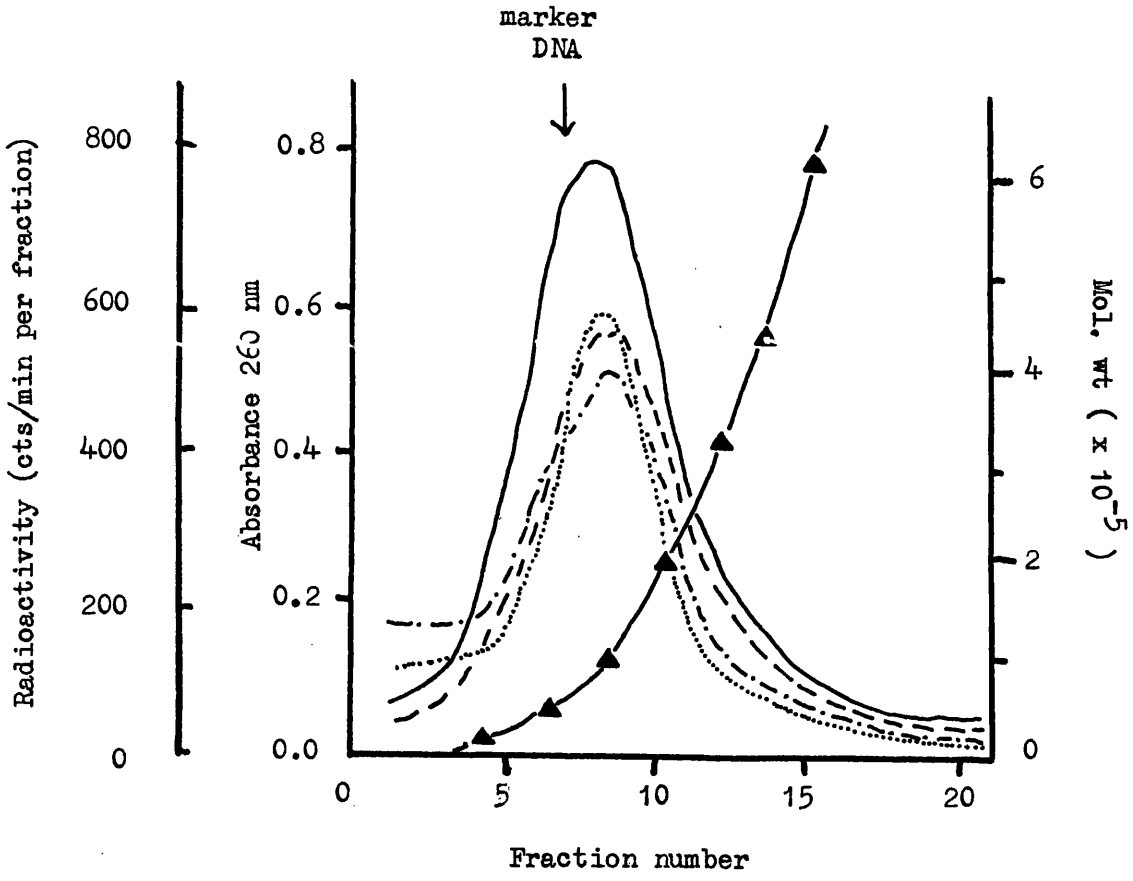
FIGURE 12.

Sedimentation of DNA and cDNA before and after hybridisation.

Hybridisation was carried out for 7 days at 60° in 0.12M phosphate (to a $C_o t$ of 8000). Samples were centrifuged at 29,000 in an M.S.E. 3 x 25 swing-out rotor for 16 hours at 20°. 1 ml fractions were collected, an aliquot counted, and the A_{260} measured; ---- foetal liver DNA before hybridisation; -·-·-· cDNA before hybridisation; ——— foetal liver DNA after hybridisation; ····· cDNA after hybridisation; —▲—▲— molecular weight.

A marker DNA of 66,000 molecular weight was centrifuged in a parallel gradient.

Figure 12


marker
DNA
Radioactivity (cts/min per fraction)
800
600
400
200
0
Absorbance 260 nm
0.8
0.6
0.4
0.2
0.0
Mol. wt (x 10⁻⁵)
6
4
2
0
0
5
10
15
20
Fraction number

(2) For foetal liver DNA, the T_m of the cDNA-mouse DNA annealed duplexes was about 3° lower than that of the reannealed mouse DNA duplexes themselves (Figure 7). After correcting for the slight differences in (G + C) content of cDNA and mouse DNA (Gruenwedel et al., 1971), this implies no more than about 2% mismatching between cDNA and the mouse DNA sequences to which cDNA annealed in these experiments (Laird et al., 1969; Ullman and McCarthy, 1973a,b). This mismatching would reduce somewhat the rate of annealing of cDNA to mouse DNA. The magnitude of this effect is difficult to estimate precisely. On the basis of data for the effect of mismatching on the rate of reassociation of satellite DNA, the rate of annealing of cDNA to mouse DNA might be reduced to 60% of the rate if there was no mismatching (Southern, 1971; Sutton and McCallum, 1971). However, similar studies on the effect of mismatching on the rate of reassociation of bacterial DNAs (Britten and Bonner, 1971) suggest that this correction to the observed rate of annealing of cDNA to mouse DNA might be much smaller.

This means that the sequences in mouse DNA to which cDNA anneals are as frequent (certainly not more than twice as frequent) as the bulk of the mouse sequences.

3.4 Treatment of mRNA with polynucleotide phosphorylase

Conditions for polynucleotide phosphorylase digestion of poly(A) were optimised using ^{3}H poly(A) as substrate.

a) Phosphate dependence

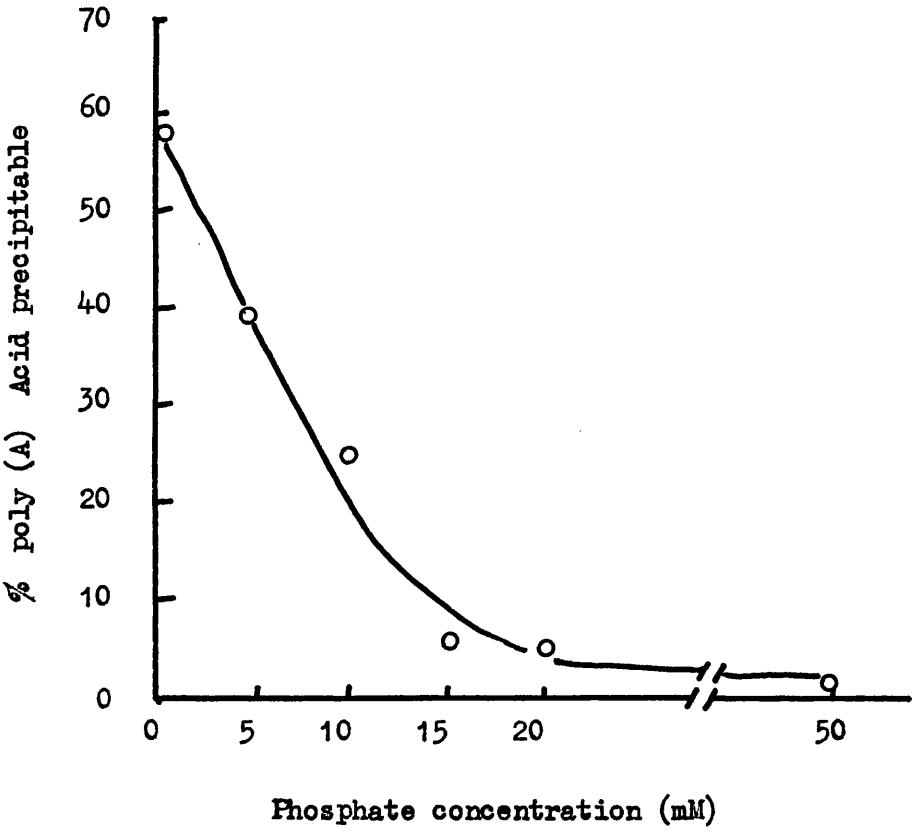
Figure 13 shows the effect of different concentrations of potassium phosphate on the degradation of ^{3}H poly(A). The conditions were 10 µg/ml ^{3}H poly(A); 1.0 µg/ml enzyme; 50 mM

FIGURE 13.

Phosphate dependance of polynucleotide phosphorylase digestion of ^{3}H poly(A).

Each 200 µl incubation contained 2 µg ^{3}H poly(A); 200 µg polynucleotide phosphorylase; 50 mM Tris (pH 7.5); 15 mM $MgCl_2$, and varying concentrations of potassium phosphate. Digestion was carried out for 1 hour at 37°.

Figure 13


% poly (A) Acid precipitable
70
60
50
40
30
20
10
0
0
5
10
15
20
50
Phosphate concentration (mM)

Tris (pH 7.5); 15 mM $MgCl_2$, for 60 minutes at 37°. (Grunberg-Manago, 1963). It can be seen that 15 mM phosphate is required for complete degradation during this period.

b) Enzyme concentration

Figure 14 shows the rate of digestion with enzyme: substrate ratios of 60:1, 30:1 and 6:1. It was known from Mansbridge et al., 1974 that the poly(A) sequence in 9S RNA represents between 10 - 15% of the molecule length. Thus, enzyme : 9S RNA ratios of 30:6 were used for the digestion of mRNA.

c) Time of incubation

Aliquots of 9S mRNA were incubated with polynucleotide phosphorylase for increasing lengths of time, using the conditions described above. The reaction was stopped by the addition of SLS, and a sample taken for size analysis on polyacrylamide gels. The remainder was deproteinised with phenol/chloroform, and RNA precipitated with ethanol. The pellet was taken up in NETS, and chromatographed on poly(U) sepharose. The RNA that was retained on the column was eluted.

The percentage of RNA that bound to the column decreased rapidly during the first 7 minutes of incubation, and only slowly after this (Figure 15).

The size of the retained and unretained RNAs were compared with that of the original mRNA on 6% polyacrylamide gels, with added marker 4S and 5S RNA from mouse reticulocytes (Loening, 1969). (Figure 16, a-e, and Figure 17).

FIGURE 14.

Rate of digestion of ^{3}H poly(A) with different concentrations of polynucleotide phosphorylase.

Each 200 µl assay contained 1.5 µg ^{3}H poly(A), with; -O-O-, 0.1 mgs enzyme/ml; -●-●-, 0.5 mgs/ml; -▲-▲-, 1.0 mgs/ml. Digestion was for 1 hour at 37°.

Figure 14

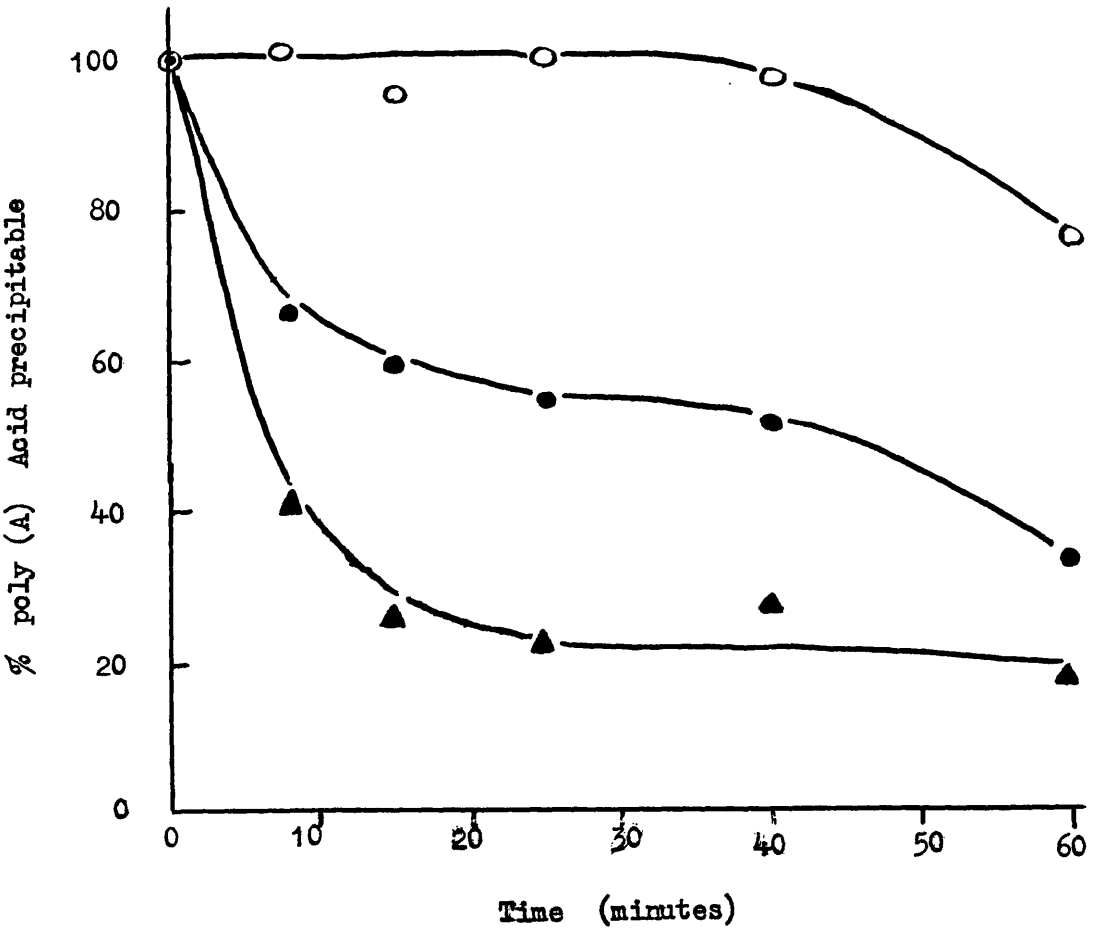
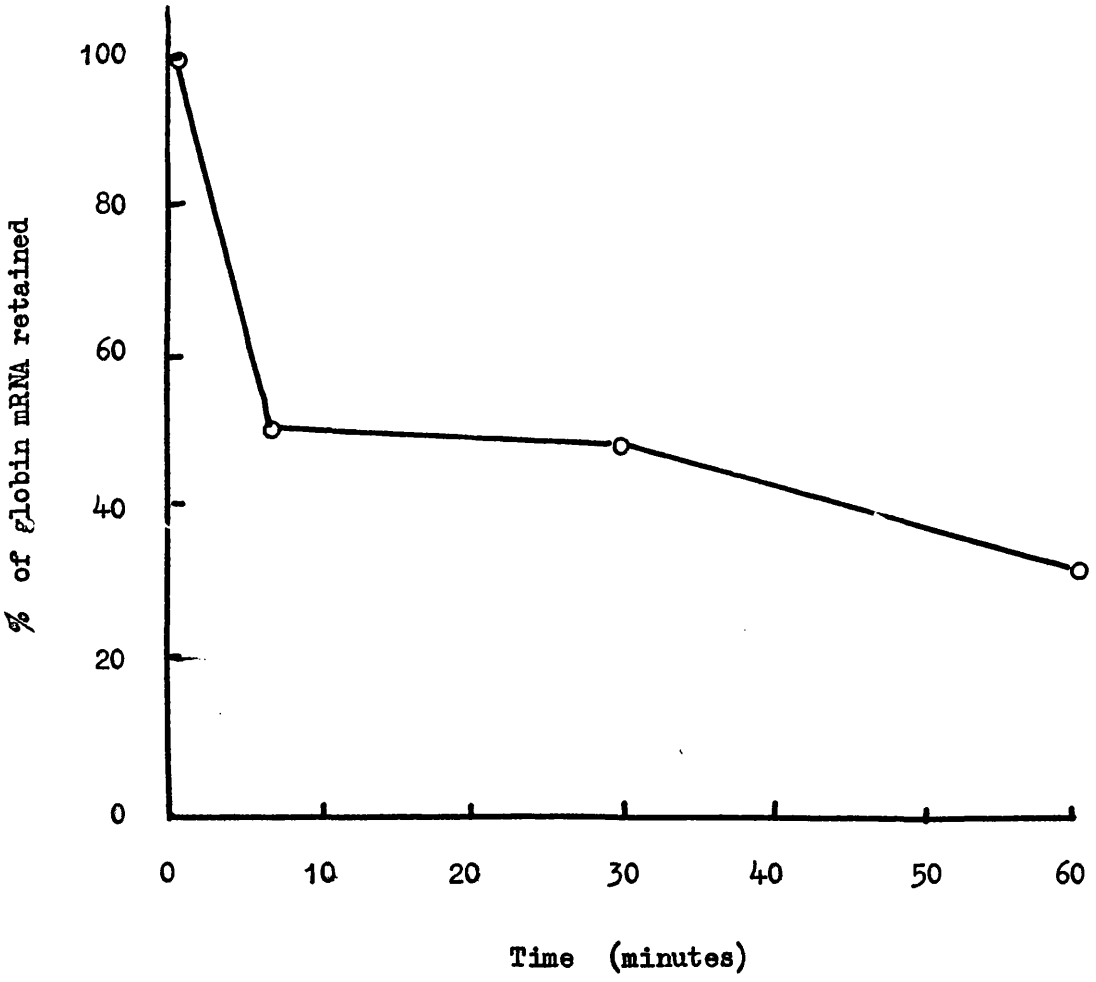

% poly (A) Acid precipitable
100
80
60
40
20
0
0
10
20
30
40
50
60
Time (minutes)

FIGURE 15.

Effect of increasing periods of polynucleotide phosphorylase digestion on the retention of 9S RNA by poly(U) sepharose.

20 µg 9S RNA were incubated at 37° with 100 µg of enzyme in each 200 µl assay. Other conditions were as described.

Figure 15


100
80
60
40
20
0
% of globin mRNA retained
0
10
20
30
40
50
60
Time (minutes)

6% polyacrylamide gel electrophoresis of globin mRNA, following incubation with polynucleotide phosphorylase.

a) 10 µg retained RNA, 0 minutes incubation.
b) 12 µg nonretained RNA, 7 minutes incubation.
c) 11 µg retained RNA, 7 minutes incubation.
d) 15 µg non-retained RNA, 30 minutes incubation.
e) 13 µg non-retained RNA, 60 minutes incubation.

Migration from left to right; gels 12 cm long, and 0.7 cm diameter; electrophoresis at 2.5 mA/gel for 30 minutes, and then 10 mA/gel for 4 hours. Gels were scanned at two times magnification, using a Joyce-Loebel UV scanner.

Figure 16

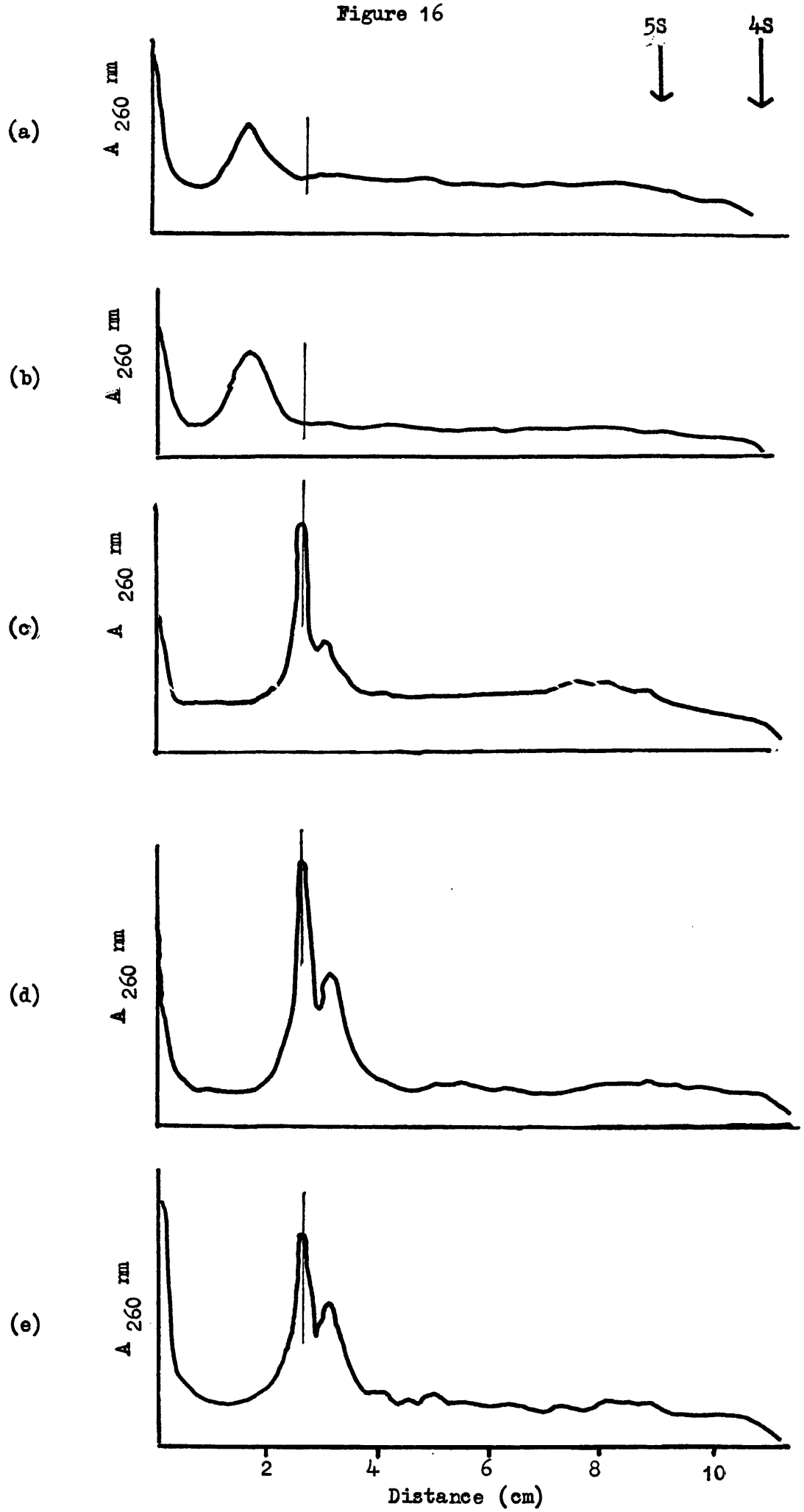


FIGURE 17.

6% polyacrylamide gel electrophoresis of globin mRNAs following incubation with polynucleotide phosphorylase.

a) 15 µg 9S mRNA.

b) 6 µg retained RNA, 7 minutes incubation.

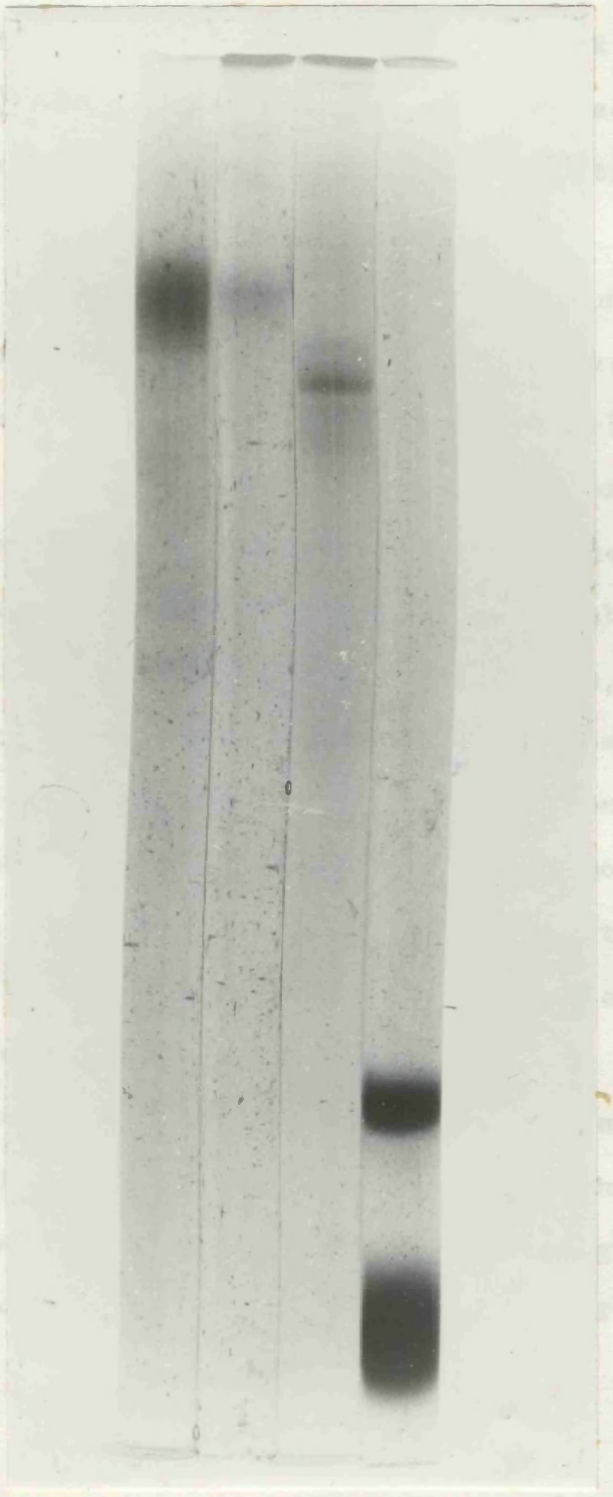
c) 12 µg non-retained RNA, 7 minutes incubation.

d) 35 µg mouse 5S and 4S RNA.

Conditions of electrophoresis were as for Figure 16. Gels were stained with toluidine blue as described by Williamson, et al. (1971).

Figure 17

(a) (b) (c) (d)



The molecular weight of the peak of unbound phosphorylase-treated mRNA varied in five preparations between 185,000 and 200,000 (Ave. mol. wt. 190,000) taking the molecular weight of the untreated mRNA as 220,000 (Figure 18). The treated bound mRNA migrates to the same position as untreated mRNA. A second smaller peak (of ave. mol. wt. 175,000) was usually seen in the treated unbound preparations, as was small amounts of material migrating in the 4S-5S region.

9S mRNA, incubated for 7 minutes under phosphorolysis conditions but in the absence of polynucleotide phosphorylase, was found to be 100% retained by a poly(U) sepharose column. This material behaves in an exactly similar manner to the original 9S mRNA in gel and sequence analyses.

Thus, in the conditions used, incubation of the mRNA with polynucleotide phosphorylase for 7 mins, results in removal of enough of the adenosine residues from a proportion of the molecules to prevent their retention by poly(U) sepharose. Neither these unretained molecules, nor the retained RNA suffer significant degradation by exonucleases (Figure 16, b & c). These incubation conditions were used in subsequent work.

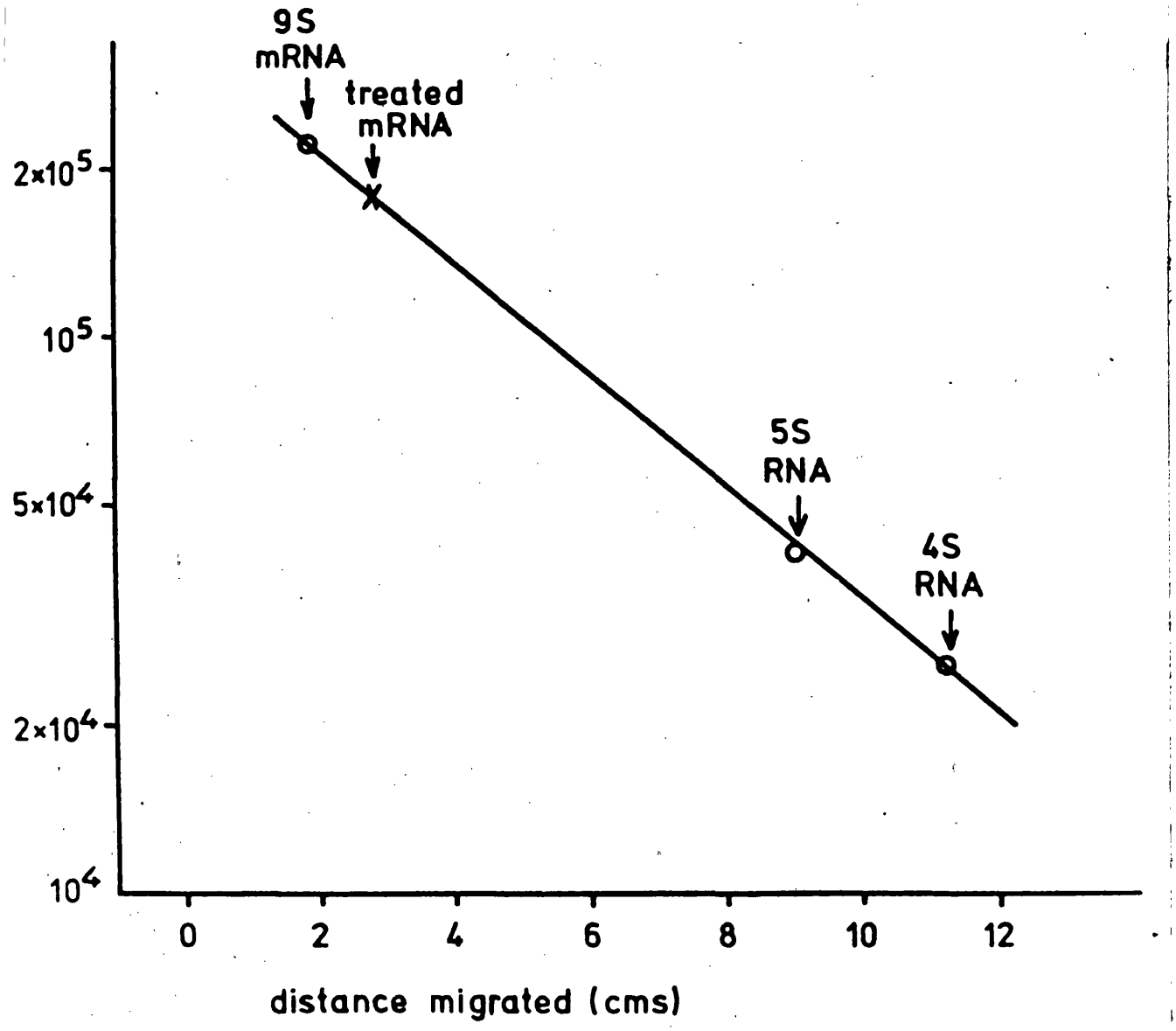
Because of the possibility of contamination of the RNAs with small fragments of poly(U), the preparation of treated mRNA for the analysis of messenger activity, was carried out using oligo(dT) cellulose affinity chromatography.

FIGURE 18.

Determination of the molecular weight of treated globin mRNA.

The treated non-retained mRNA was run on 6% polyacrylamide gels with markers of untreated 9S globin mRNA (mol. wt. 2.20×10^5, Williamson et al., 1971), mouse 5S RNA (mol. wt. 4.1×10^4, Williamson and Brownlee, 1969) and mouse 4S RNA (mol. wt. 2.6×10^4, Staehelin et al., 1968).

Figure I8


9S
mRNA
treated
mRNA
5S
RNA
4S
RNA
2×10^5
10^5
5×10^4
2×10^4
10^4
0
2
4
6
8
10
12
distance migrated (cms)

3.5 Characterisation of treated, non-retained RNA

a) Fingerprint analysis

Pancreatic RNase hydrolyses internucleotide bonds in RNA specifically after cytidine and uridine residues, giving first a 2':3'-cyclic phosphate and then in turn hydrolysing this to the 3'-phosphate. Poly(A) sequences are also hydrolysed slowly to give oligo(A) fragments by pancreatic RNase, or by a minor contaminating enzyme activity, particularly in solutions of low ionic strength (Beers, 1960). T_1 RNase hydrolyses internucleotide bonds specifically after guanosine residues, giving oligonucleotides terminating in the 3'-phosphate. Bacterial alkaline phosphatase will hydrolyse 3'-phosphate groups (as well as 5'-phosphate groups), but will not attack 2':3'-cyclic phosphates or phosphodiester bonds. Bacteriophage phosphokinase transfers a phosphate group from the γ-position of ATP to the 5'-end of an oligonucleotide or a mononucleotide, but not to a mononucleoside.

Therefore, depending on whether the terminal phosphate is cyclic or not, and whether phosphatase is used, series of oligonucleotides of the form $(pA)_n pN\text{-}OH$ and $(pA)_n pNp$ will be generated after treatment of RNA with T_1 plus pancreatic RNases where N can be any of the four ribonucleosides.

Two-dimensional fingerprints of T_1 plus pancreatic RNase-treated 9S RNA, phosphorylase-treated non-retained mRNA and poly(A) are shown in Figure 19. Since pancreatic RNase hydrolyses cyclic purine nucleotides to 3' phosphate only slowly, if at all (Markham, 1957), two series of oligonucleotide spots are seen for poly(A), corresponding to $(pA)_n\text{-}p$ and $(pA)_n\text{-}OH$. (Figure 19a). For 9S mRNA, the proportion of the total

FIGURE 19.

Fingerprints of commercial poly(A), untreated 9S mRNA and phosphorylase-treated non-retained mRNA:

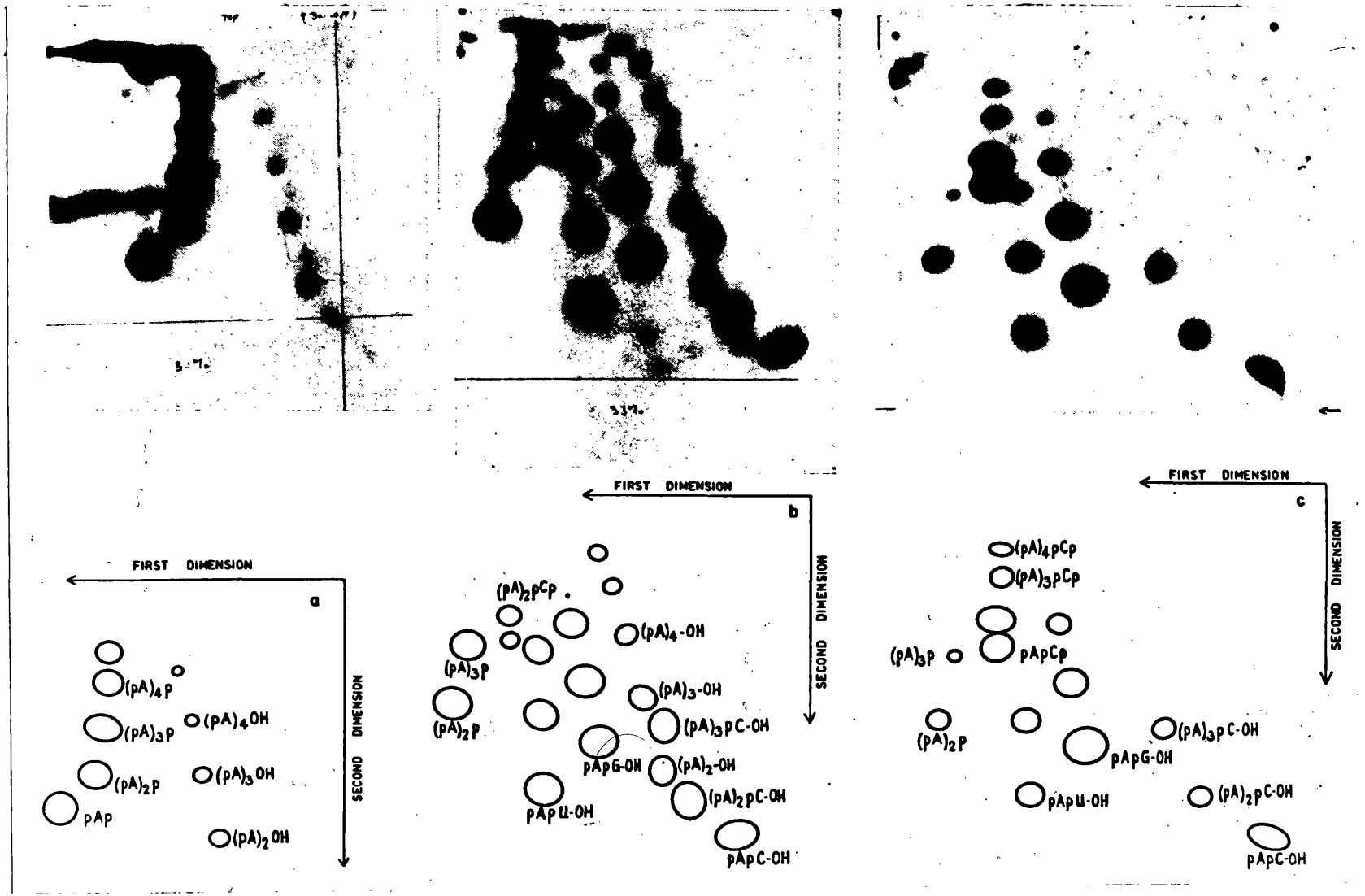
a) Commercial poly(A) treated with pancreatic RNase and labelled with phosphokinase.

b) 9S globin mRNA, treated with T_1 plus pancreatic ribonucleases followed by phosphatase and then labelled at the 5' end with phosphokinase.

c) Phosphorylase-treated non-retained mRNA, treated as in (b).

3 - 5 µg of each RNA was used.

Figure 19


FIRST DIMENSION
SECOND DIMENSION
a
(pA)4p
(pA)3p
(pA)2p
pAp
(pA)4OH
(pA)3OH
(pA)2OH
b
(pA)2pCp
(pA)4-OH
(pA)3p
(pA)3-OH
(pA)2p
(pA)3pC-OH
pApG-OH
(pA)2-OH
pApU-OH
(pA)2pC-OH
pApC-OH
c
(pA)4pCp
(pA)3pCp
(pA)3p
pApCp
(pA)2p
(pA)3pC-OH
pApG-OH
pApU-OH
(pA)2pC-OH
pApC-OH

radioactivity found in these two oligonucleotide isopliths derived from poly(A) (but not including pAp, which is also found in fingerprints of rRNA) is approximately 20 ± 3%. (Figure 19b). The proportion of radioactivity in the oligo(A) tracts for the treated non-bound mRNA is 1.4 ± 0.2% and the great majority is found in the smallest oligonucleotides. (Figure 19c).

This demonstrates the absence of at least 90% of the oligo(A) sequences derived from the poly(A) tract, in the treated non-retained RNA.

b) Template activity with reverse transcriptase

It is known that reverse transcriptase requires a double stranded primer to commence transcription on the template RNA. In an experiment in which both untreated and treated non-retained globin mRNA were copied by reverse transcriptase, using oligo$(dT)_{12-17}$ as primer, the incorporation obtained is shown in Table I.

The low level of transcription of the treated non-retained mRNA - less than 1% of the untreated mRNA - indicates that the preparation contains few, if any, molecules with a poly(A) sequence long enough to form a stable hybrid with oligo$(dT)_{12-17}$ under these conditions.

3.6 Translational ability of treated non-retained RNA

a) Using the Krebs II ascites cell-free system

The dependence of amino acid incorporation on added treated bound and treated non-bound globin mRNA is shown in Figure 20. For both RNAs the stimulation of incorporation is linear and similar up to 4 picomoles of mRNA added per 50 µl of incubation mixture. The amount of added mRNA was calculated using molecular weights of 220,000 (control) and

TABLE I

Transcription of untreated and treated non-bound globin mRNA with reverse transcriptase in the presence of oligo(T) primer.

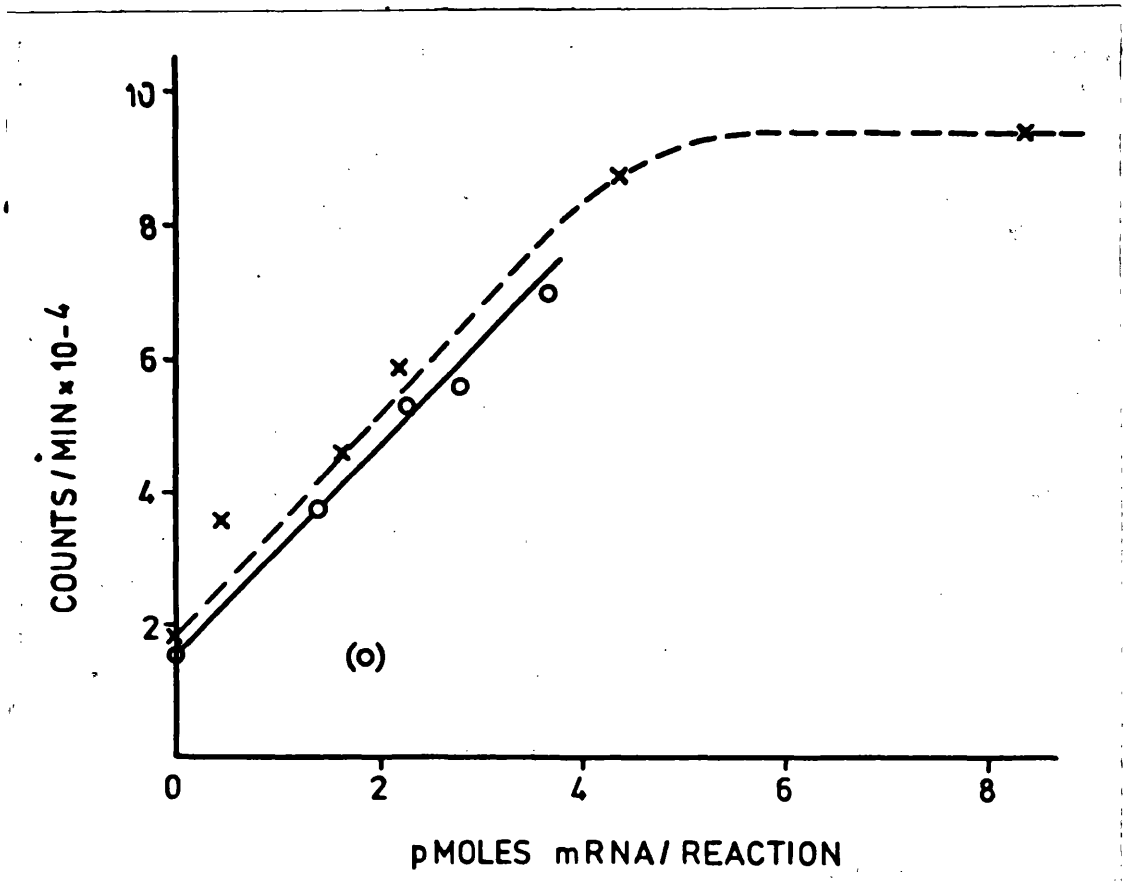
	Cpm/μg of Template
Control globin mRNA	7.4×10^6
Treated globin mRNA (non-bound)	4.9×10^4

FIGURE 20.

Dependence of incorporation upon concentration of added globin mRNA, in the Krebs ascites cell-free system.

Assays were carried out as described in the text. -O-O-, retained mRNA; -X-X-, non-retained mRNA. Regression analysis gives slopes of 14,900 (non-retained) and 14,700 (retained, omitting one low point).

Figure 20



185,000 (treated, non-bound).

The elution profile of globin chains synthesised after addition of treated bound and non-bound mRNAs in this linear range to the Krebs ascites system is shown in Figure 21. The counts recovered in each globin peak and in the breakthrough peak are given in Table II. The non-bound mRNA is approximately half as active as the control mRNA judged by product analysis, and there is a greater depression of α-chain synthesis than β-chain synthesis.

b) Using the rabbit reticulocyte lysate cell-free system

Figure 22 shows the separation of the globin chains synthesised in the presence of added mouse globin mRNA after 6 minutes incubation. Only the mouse β-globin chain is resolved from the rabbit globins (Lingrel et al., 1971).

It has been shown (Palmiter, 1973) that the rabbit lysate contributes 5 nM of isoleucine to each 250 µl assay, and thus 1.6% of the isoleucine incorporated will be labelled. Mouse β-globin contains 4 isoleucines/chain and rabbit globin an average of 2/chain (Dayhoff, 1972). It can be calculated that an incorporation of approximately 2,000 dpm represents one picomole of mouse β-globin. Analysis of the translation of mRNA in the Krebs cell-free system suggested that the mouse α-globin mRNA was translated at about 50% of the efficiency of the β mRNA, and a small correction to the counts in the rabbit globin peak has been made for this (Table III).

The amount of mouse message added was known, and Palmiter has estimated that the rabbit lysate contains 10.9 picomoles of rabbit mRNA/assay. The number of times each message was translated was calculated, and the data normalised for the different amounts of rabbit globin made in each assay.

FIGURE 21.

Product analysis of the globin chains synthesised in the Krebs ascites cell-free system, with added mouse globin mRNAs.

5 picomoles of mRNA in 100 µl assays were incubated for 60 minutes at 37°. Two urea-phosphate, CM cellulose columns were run, and the results plotted together:-O-O- retained mRNA;-X-X-, non retained mRNA.

Figure 2I

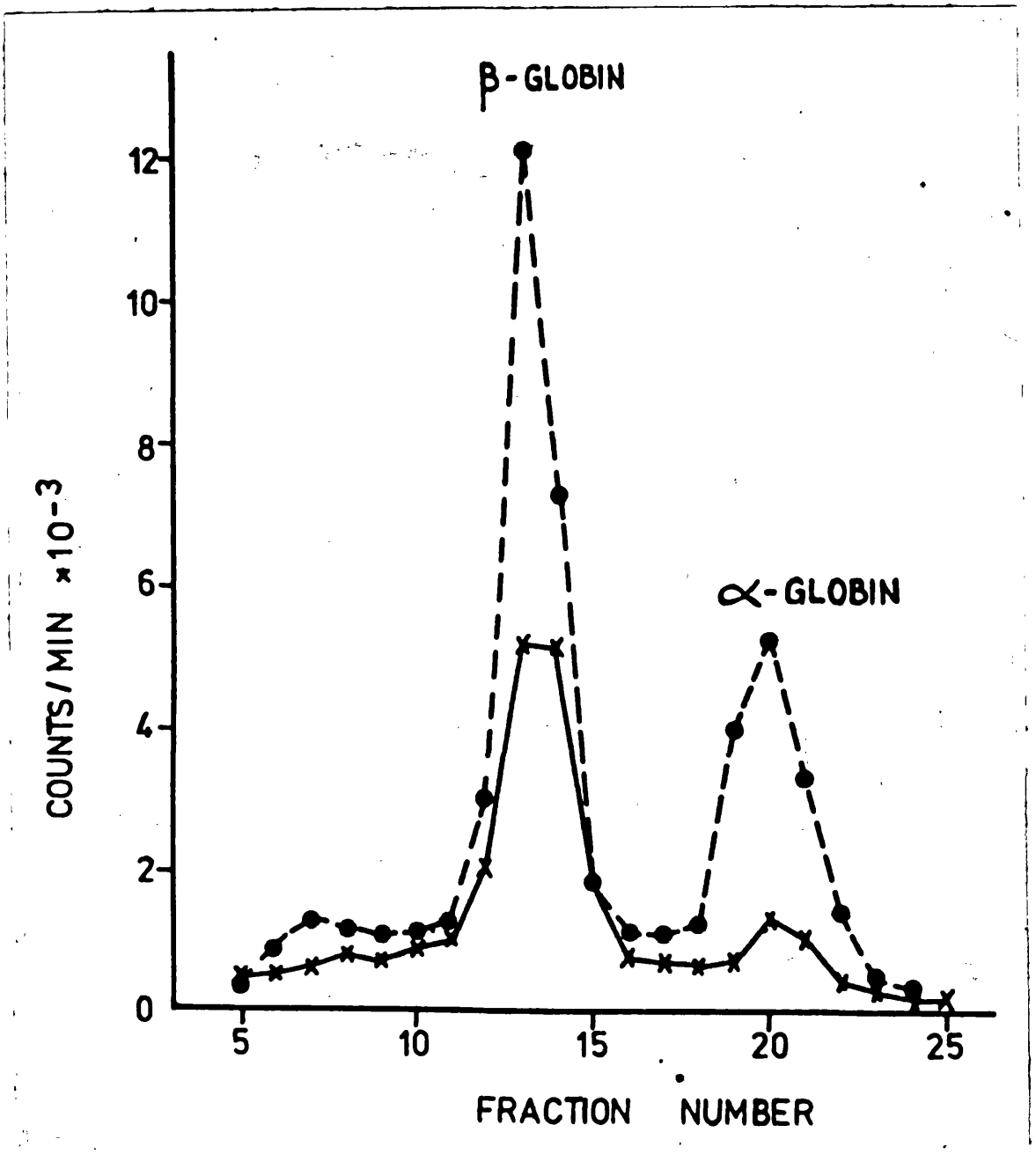

β-GLOBIN
α-GLOBIN
COUNTS/MIN ×10⁻³
0
2
4
6
8
10
12
5
10
15
20
25
FRACTION NUMBER

TABLE II

Amino acid incorporation into separated globin chains in the Krebs ascites system

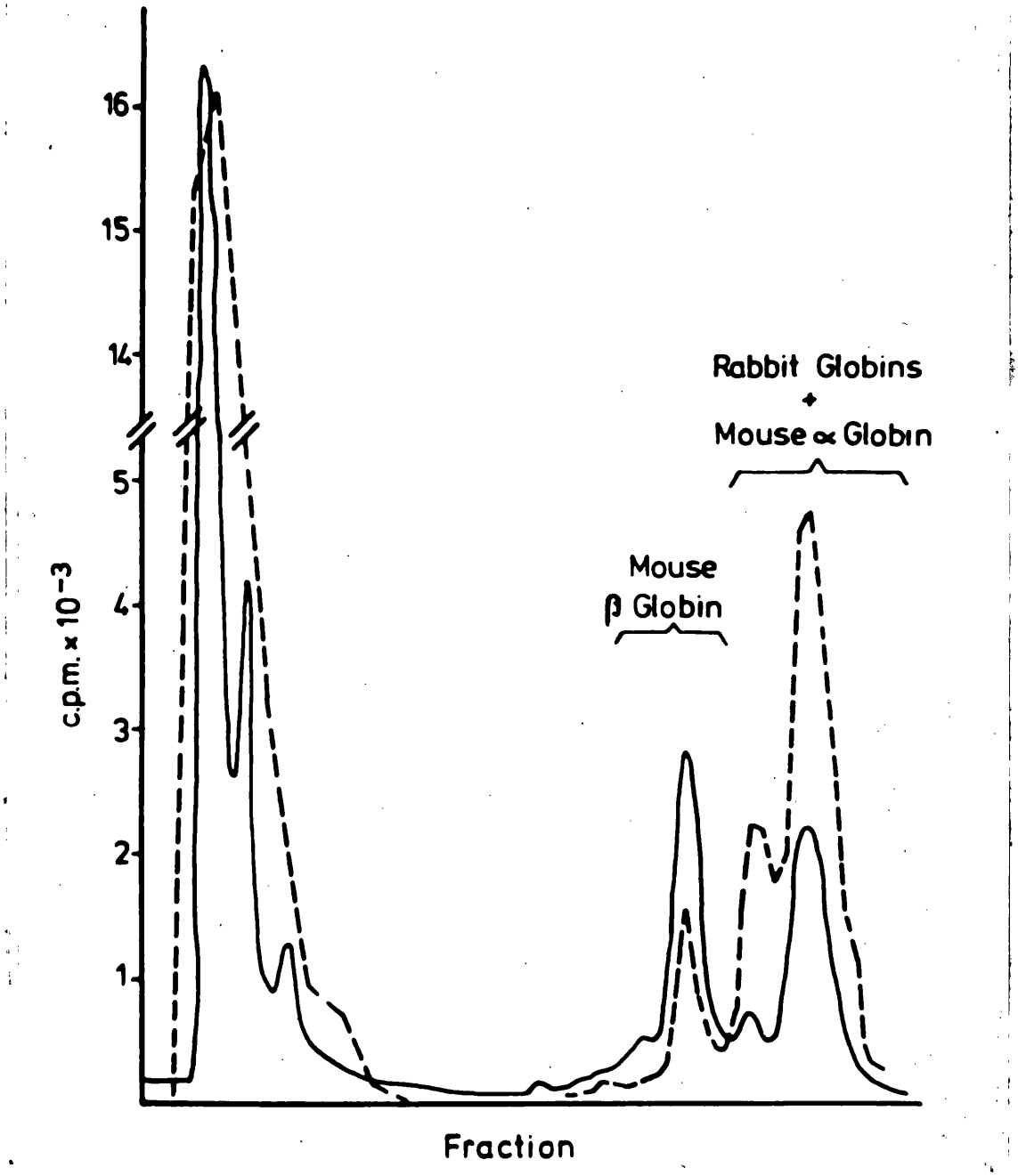
CCl_3COOH-Precipitable Counts in Peak	Bound 9S	Non-bound 9S	Non-bound: Bound (%)
α-Globin	15,370	3,890	25
β-Globin	24,435	15,513	63
Breakthrough	60,000	70,756	
α + β	39,805	19,203	48
Total	100,552	89,959	89.5

FIGURE 22.

Product analysis of the globin chains synthesised in a rabbit reticulocyte cell-free system, with added mouse globin mRNA.

2.3 picomoles of control mouse globin mRNA were incubated in a 250 µl assay, and extracted globin chromatographed on carboxymethyl cellulose as described. The absorbance was monitored at 280 nm (———) and the fractions counted at 48% efficiency (- - - - -).

Figure 22


16
15
14
5
4
3
2
1
c.p.m. × 10⁻³
Rabbit Globins
+
Mouse α Globin
Mouse
β Globin
Fraction

Amino acid incorporation into separated globin chains made in the rabbit reticulocyte cell-free system.

		dpm in each fraction	picomoles globin chain made	picomoles of mRNA added[1]	Number of globin chains made per mRNA added	corrected number of chains made[2].
6 minute incubation with added deadenylated mRNA	mouse β globin	16,080	3.81	2.6	1.46	1.9
	total rabbit globins	209,400	99.17	10.9	9.1	10.5
6 minute incubation with added control mRNA	mouse β globin	24,803	5.87	2.3	2.55	2.8
	total rabbit globins	241,763	114.47	10.9	10.5	10.5
90 minute incubation with added deadenylated mRNA	mouse β globin	59,280	14.01	1.3	10.78	11.2
	total rabbit globins	918,130	434.72	10.9	39.88	41.53
90 minute incubation with added control mRNA	mouse β globin	85,160	20.16	1.15	17.53	17.53
	total rabbit globins	956,140	452.72	10.9	41.53	41.53

Thus the amount of globin chains made by the deadenylated mRNA is:

$$\frac{1.9}{2.8} = 68\% \text{ after 6 minutes}$$

$$\text{and } \frac{11.2}{17.53} = 64\% \text{ after 90 minutes.}$$

[1] The amount of β globin mRNA in both the control and deadenylated mRNA was assumed to be 50% of the total RNA.

[2] The figures in the last column have been corrected for small differences (± 10%) in the amount of rabbit globin made in each assay.

After six minutes the deadenylated mouse mRNA had been translated twice per molecule on average and the treated bound mRNA 2.8 times per molecule on average; after 90 minutes the deadenylated mRNA had been translated 11.2 times on average and the treated bound mRNA 19.2 times on average. The deadenylated mRNA consistently showed only 60 - 70% the activity of the control mRNA. However, this figure was similar at 6 and 90 minutes.

3.7 Quantitation of poly(U) sepharose

The ability of poly(U) sepharose to retain poly(A) containing sequences quantitatively and reproducibly, is important for the interpretation of the data discussed in later sections. Several experiments were devised to test this.

a) With reticulocyte polysomal RNA

1½ mg of reticulocyte polysomal RNA were chromatographed on 5 mls of poly(U)-sepharose, and the RNA in both retained and non-retained fractions was precipitated. The non-retained material was recycled through poly(U)-sepharose 3 times. Small amounts of RNA were retained on the column at each cycle. Samples of these RNAs were titrated with cDNA, and the percentage of each that was complementary to cDNA was calculated, by comparison with a titration curve of cDNA and 9S RNA.

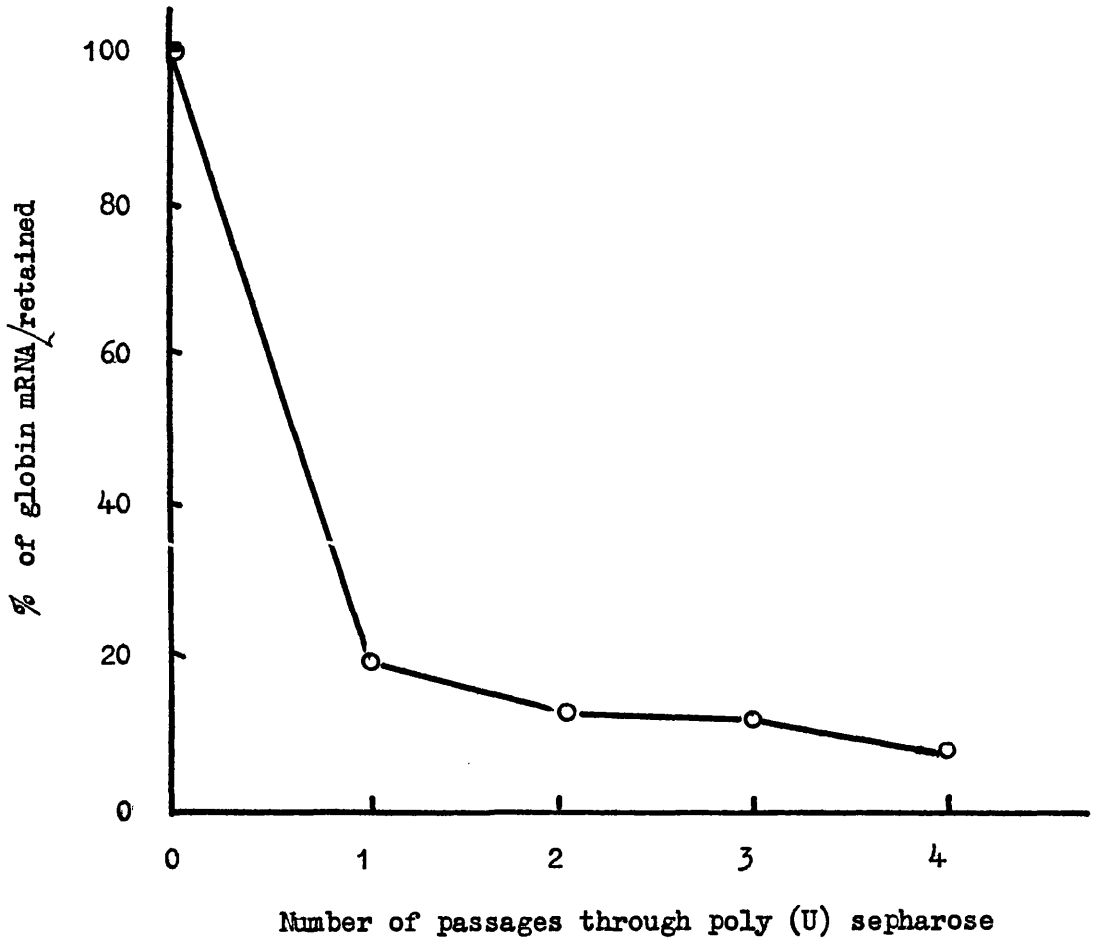
80% of the globin sequences present in the polysomal RNA are retained by poly(U) sepharose, on the first cycle (Figure 23 and Table IV). Some of the globin RNA sequences that are not retained by the first passage, are retained in second or subsequent cycles. After 4 cycles through poly(U) sepharose more than 90% of the globin mRNA sequences have been retained by the column.

FIGURE 23.

Retention of globin mRNA sequences by poly(A) sepharose.

The RNA that was not retained by the column was recycled, and the amounts of globin RNA sequences estimated by cDNA titration. The percentage of the total globin RNA sequences that was non-retained by the column at each stage is plotted.

Figure 23


100
80
60
40
20
0
% of globin mRNA retained
0
1
2
3
4
Number of passages through poly (U) sepharose

b) With retained RNA

210 µg of the material that was retained by the column on the first cycle, was rechromatographed over 2 mls of poly(U) sepharose. The RNA in the retained and non-retained fractions was titrated with cDNA. The percentage of RNA sequences complementary to cDNA was calculated by comparison with a 9S - cDNA titration curve.

More than 97% of the globin RNA sequences were retained by the column, in this second cycle (Table V).

c) With mRNP RNA

Globin mRNP was prepared from anaemic mouse reticulocyte polysomes, by taking a crude 11S cut from a BXIV zonal sucrose gradient (Figure 2). RNA was extracted from this, and a proportion analysed on 2.6% acrylamide gels (Figure 2b). About 20% of the RNA ran in the 9S region, by comparison with standard RNA markers. 15 µg of this RNA was chromatographed on 1 ml of poly(U) sepharose, and the retained and non-retained fractions titrated with cDNA. The proportion of sequences complementary to globin cDNA, was calculated by comparison with a 9S RNA-cDNA titration curve (Table VI).

More than 98% of the globin RNA sequences prepared from 11S mRNP were retained by poly(U) sepharose.

d) Using oligo(A) of different lengths

Known amounts of poly(A) (mol. wt. greater than 100,000), oligo $(A)_{10}$ and oligo$(A)_7$ were applied to 2 ml columns of poly(U) sepharose. Non-retained material was washed off with a known volume of NETS, and the absorbance at 260 nm was measured. The amounts of each RNA that were not retained by the column, were expressed as a percentage of the loading sample (Table VII).

Under these conditions, only oligo$(A)_7$ was not appreciably retained by the column.

Retention of retained globin mRNA.

	Amount of RNA in each fraction (µg).	Percentage of RNA complementary to globin cDNA	Amount of globin mRNA in each fraction (µg).	Percentage of Total globin mRNA sequences
Retained RNA First cycle	217	62	134.5	-
Retained RNA second cycle	153	91	139.2	97.3
Non-retained RNA second cycle	64	5.9	3.8	2.7

FIGURE 24.

2.6% polyacrylamide gel analysis of RNA extracted from 14S mRNP.

6 µg of RNA were layered on to gels 0.7 cm in diameter, and 7 cm long. Electrophoresis was for 20 minutes at 2 mA/gel, and then 70 minutes at 6 mA/gel.
The gel was scanned at 2 times magnification on a Joyce–Loebel gel scanner. The positions of 18S, 9S and 5S RNA bands that were run on a parallel gel, are shown by arrows.

Figure 24

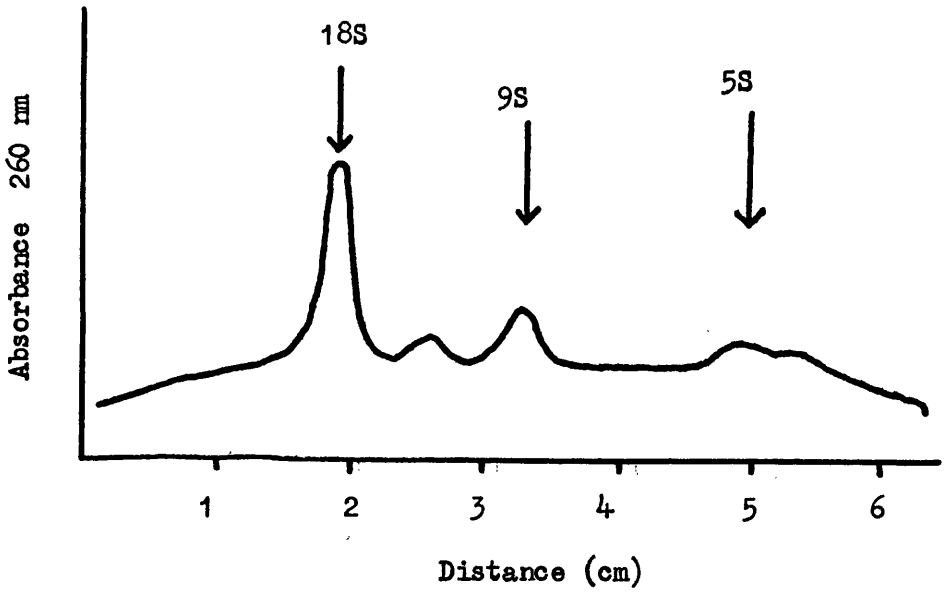


TABLE VI

Retention of mRNP globin mRNA by poly(U) sepharose

	Amount of RNA in each fraction (μg)	Percentage of RNA complementary to globin cDNA	Amount of globin mRNA in each fraction (μg)	Percentage of total globin mRNA sequences
Retained RNA	3.3	48.2	1.591	98.4
Non-retained RNA	3.7	0.7	0.026	1.6

TABLE VII

Retention of Adenosine oligomers by poly(U) sepharose

	Amount of RNA (μg)	Percentage of RNA not retained
poly(rA)	133	15.8
$oligo(rA)_{10}$	102	< 3
$oligo(rA)_7$	112	> 98

Samples applied to a 1 ml poly(U) sepharose column.

c) With nuclear RNA

^{3}H adenosine labelled nuclear RNA prepared from cells of the Friend erythroleukaemic line (clone M2) was kindly donated by Dr. M. Getz. Cells were harvested in mid log-phase growth, after labelling for 16 hrs (1 mCi, 1 µmole of ^{3}H adenosine per litre) (Gilmour et al., 1974). Nuclei were prepared using the citric acid procedure, and nuclear RNA isolated as previously described. This RNA was fractionated on 2 mls of poly(U) sepharose. 22% of the labelled RNA was retained by the column (Table VIII).

The retained and non-retained RNAs were digested in high salt with a mixture of pancreatic and T_1 RNases, as described before, and the reaction stopped by the addition of SLS. The long poly(A) sequences were separated from nucleotides and other oligo nucleotide sequences, by chromaography on poly(U) sepharose. The small amounts of ^{3}H poly(A) were eluted from the column with buffer containing oligo$(rA)_{10}$ and mouse 4S and 5S RNA. The RNA was precipitated with ethanol, taken up in NETS, and a proportion counted.

The amounts of RNA binding to poly(U) sepharose at this stage are shown in Table VIII. The remainder of the sample was run on 12% polyacrylamide gels (Figure 25). About 98% of the nuclear poly(A) sequences are found in the RNA that binds to poly(U) sepharose before digestion. This poly(A) is heterogeneous in size, with material running between 4S and 5S RNA markers. Digestion releases an additional 2% of poly(A) that can be retained by poly(U) sepharose. This material consists entirely of small sequences of 10 - 20 residues in length.

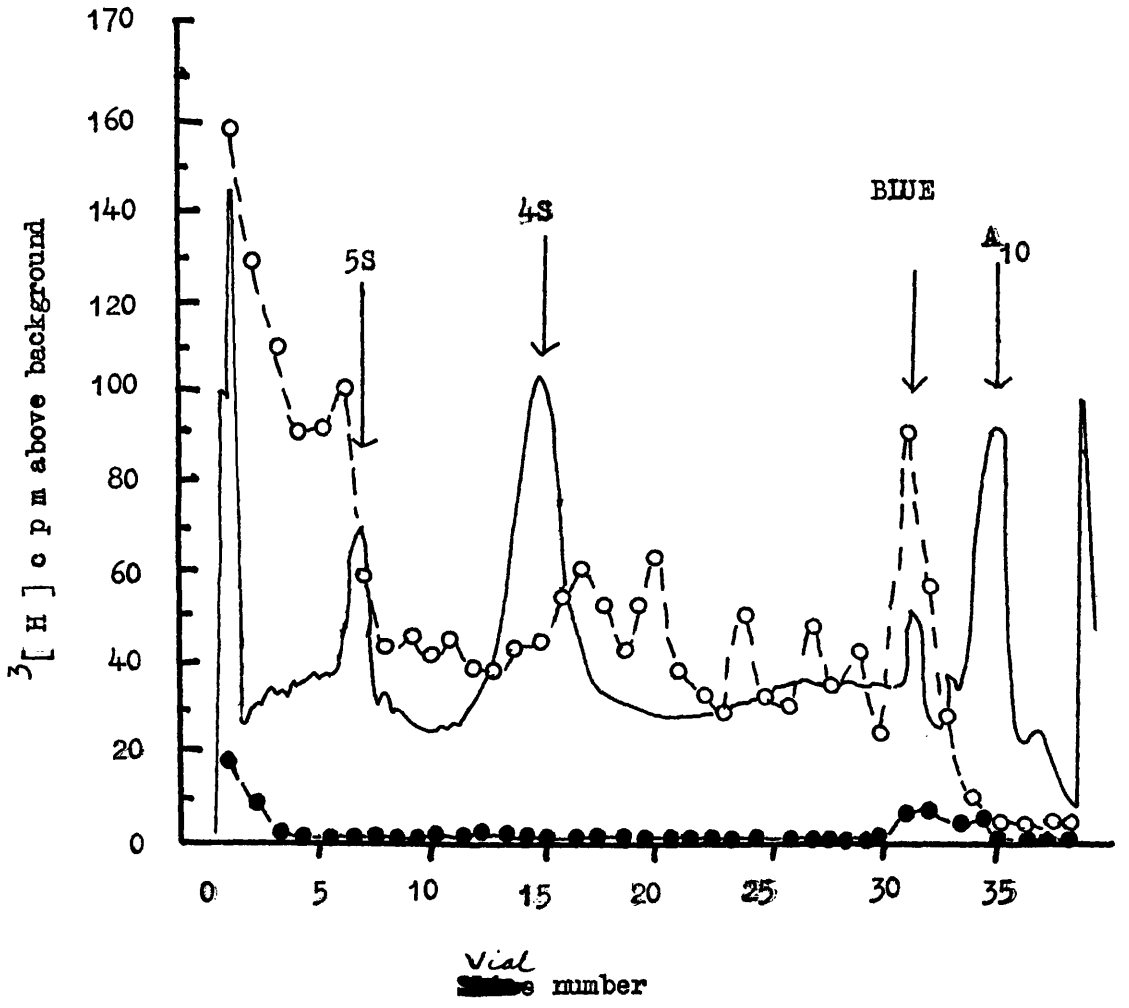
Retention of nuclear RNA by poly(U) sepharose before and after RNase digestion.

	Percentage of RNA in each fraction	Amount of RNA retained after digestion (cpm)	Amount of RNA non-retained after digestion (cpm)	Percentage of RNA retained after digestion
Retained nuclear RNA	22.2	3,672	447,645	0.82
Non-retained nuclear RNA.	77.8	360	1,564,797	0.02

12% polyacrylamide gel electrophoresis of nuclear poly(A).

After digestion with a mixture of T_1 and pancreatic RNases, the resistant material that was retained by poly(U) sepharose, was subjected to electrophoresis along with markers of 5S and 4S RNA, (A_{10}) and bromophenol blue. Electrophoresis was at 1.5 mA/gel for 30 minutes, and then at 8 mA/gel for 4 hours, on 0.7 cm by 12 cm gels. The gels were scanned at one half magnification, and then sliced and counted. The positions of the marker RNAs were used to align the two experimental samples: ———, absorbance at 260 nm; —o—o—, cpm per 3 slices for the retained nuclear RNA; —●—●—, cpm per 3 slices for the non-retained RNA.

Figure 25


170
160
140
120
100
80
60
40
20
0
0
5
10
15
20
25
30
35
5S
4S
BLUE
A10
^{3}[H] c p m above background
Vial number

3.8 Identification of globin mRNA in mouse foetal livers

During the 11th to the 16th day of foetal mouse development, the liver becomes the main erythropoietic organ, and the synthesis of adult globin chains occurs in the cytoplasm of the cells (Kovach et al., 1967; Cole et al., 1968). RNA sequences coding for globin have been previously detected in the nuclear RNA from 14 day foetal livers (Williamson et al., 1973).

a) Presence of globin mRNA in 14 day foetal liver polysomes

220 A_{260} units of polysomes were prepared from 68 14 day foetal livers. 3 A_{260} units of polysomes were analysed on a 10 - 30% sucrose gradient in TKM (Figure 26). The majority of the material consists of polysomes of greater than 4 ribosomes/message.

RNA was prepared from the total polysomes, and chromatographed on poly(U) sepharose. Approximately 1% of the RNA was retained by the column. The RNA was analysed on 2.6% acrylamide gels (Figure 27). The foetal liver polysomal RNA shows 3 minor peaks running between 18S and 5S (Figure 27a), as does the RNA that is not retained by the column (Figure 26b). The retained material shows a considerable enrichment of RNA in the 9S region (Figure 27c), as compared with a reticulocyte 9S RNA (Figure 27d). About 15% of the retained RNA is in the 9S region, with most of the other RNA running at 28S and 18S.

b) Messenger activity of foetal liver RNA

30 µg of retained RNA was tested for globin messenger activity in a duck cell-free system. Figure 28 shows the elution profile of globin chains from a CM cellulose column, and the stimulation of mouse α and β globin chain synthesis by the RNA sample. This stimulation is approximately one quarter

Sedimentation of 14 day foetal liver polysomes.

0.75 A_{260} units of foetal liver polysomes were layered onto a 10 – 30% sucrose gradient in TKM and spun in the M.S.E. 3 x 6.5 ml titanium swing-out rotor at 60,000 rpm for 20 minutes at 4°. The absorbance at 260 nm was monitored (——) and the percentage sucrose (- - -) estimated from the refractive index of each 0.5 ml fraction.

Figure 26

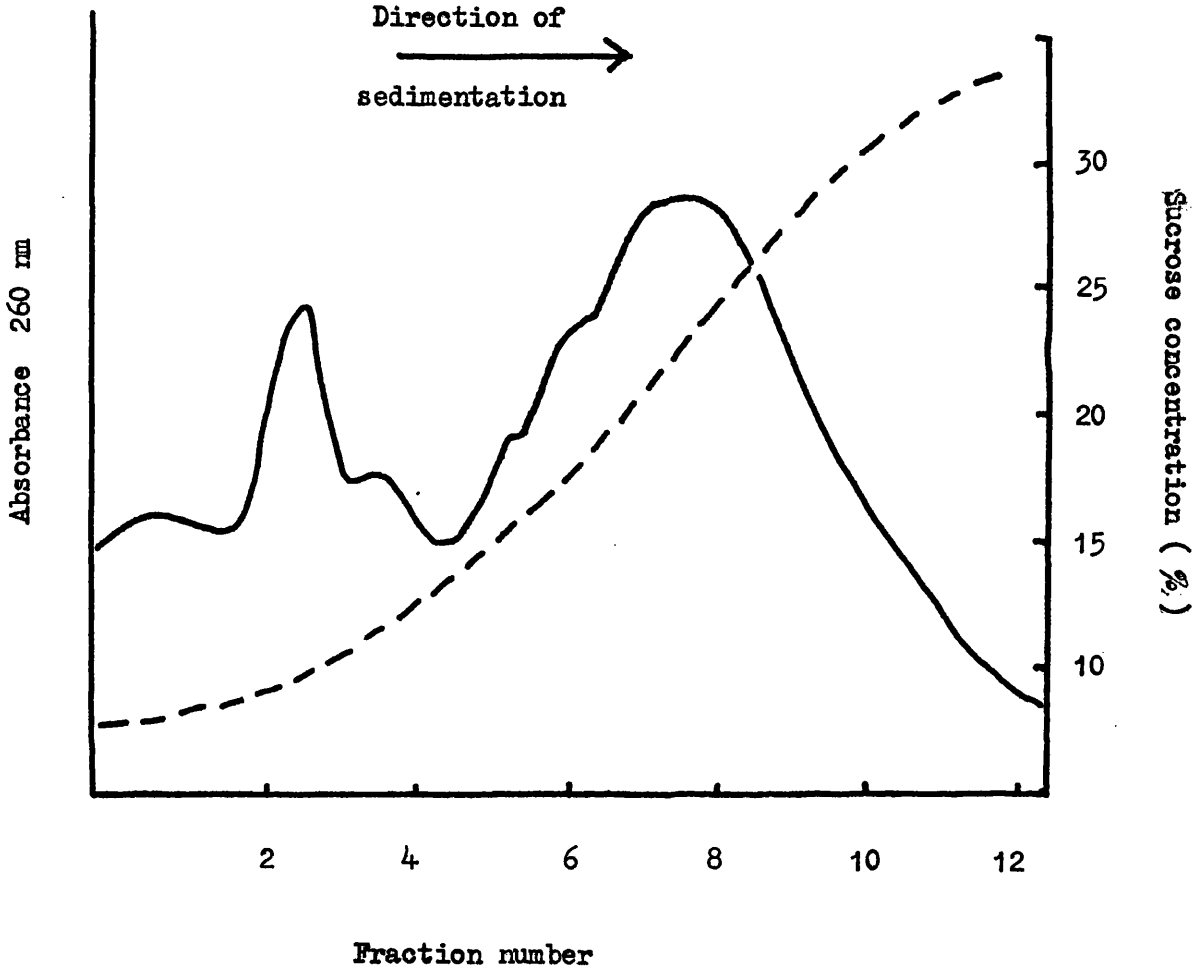

Direction of
sedimentation
Absorbance 260 nm
Sucrose concentration (%)
30
25
20
15
10
2
4
6
8
10
12
Fraction number

FIGURE 27.

2.6% polyacrylamide gel analysis of foetal liver polysomal RNA.

Electrophoresis was in 0.7 cm by 7 cm gels, for 15 minutes at 2 mA/gel, and then 70 minutes at 6 mA/gel. Gels were scanned at two times magnification on a Joyce-Loebel UV scanner:

a) 60 µg foetal liver polysomal RNA.

b) 70 µg foetal liver non-retained polysomal RNA.

c) 17 µg retained polysomal RNA.

d) 4 µg reticulocyte 9S globin mRNA.

Figure 27

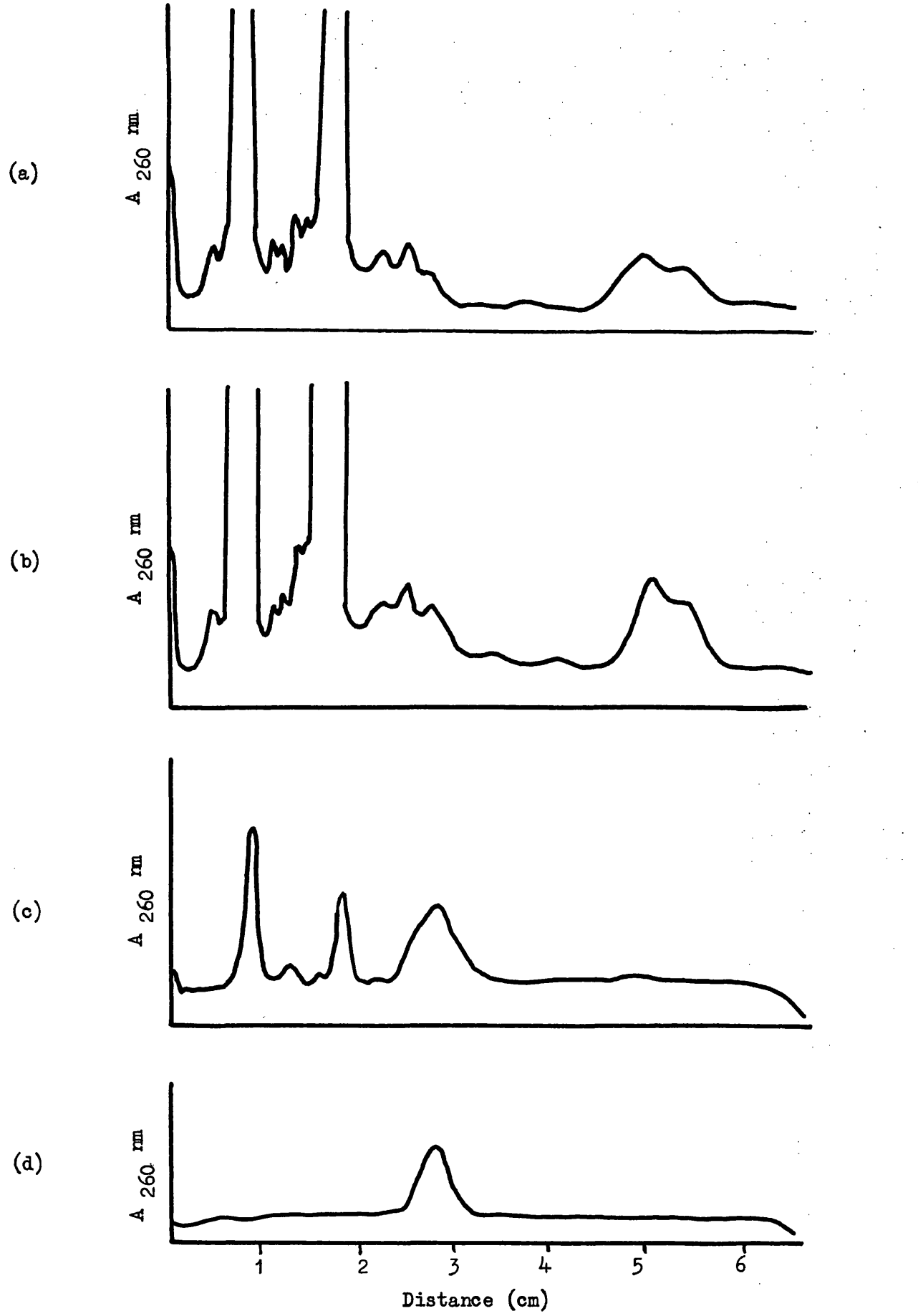
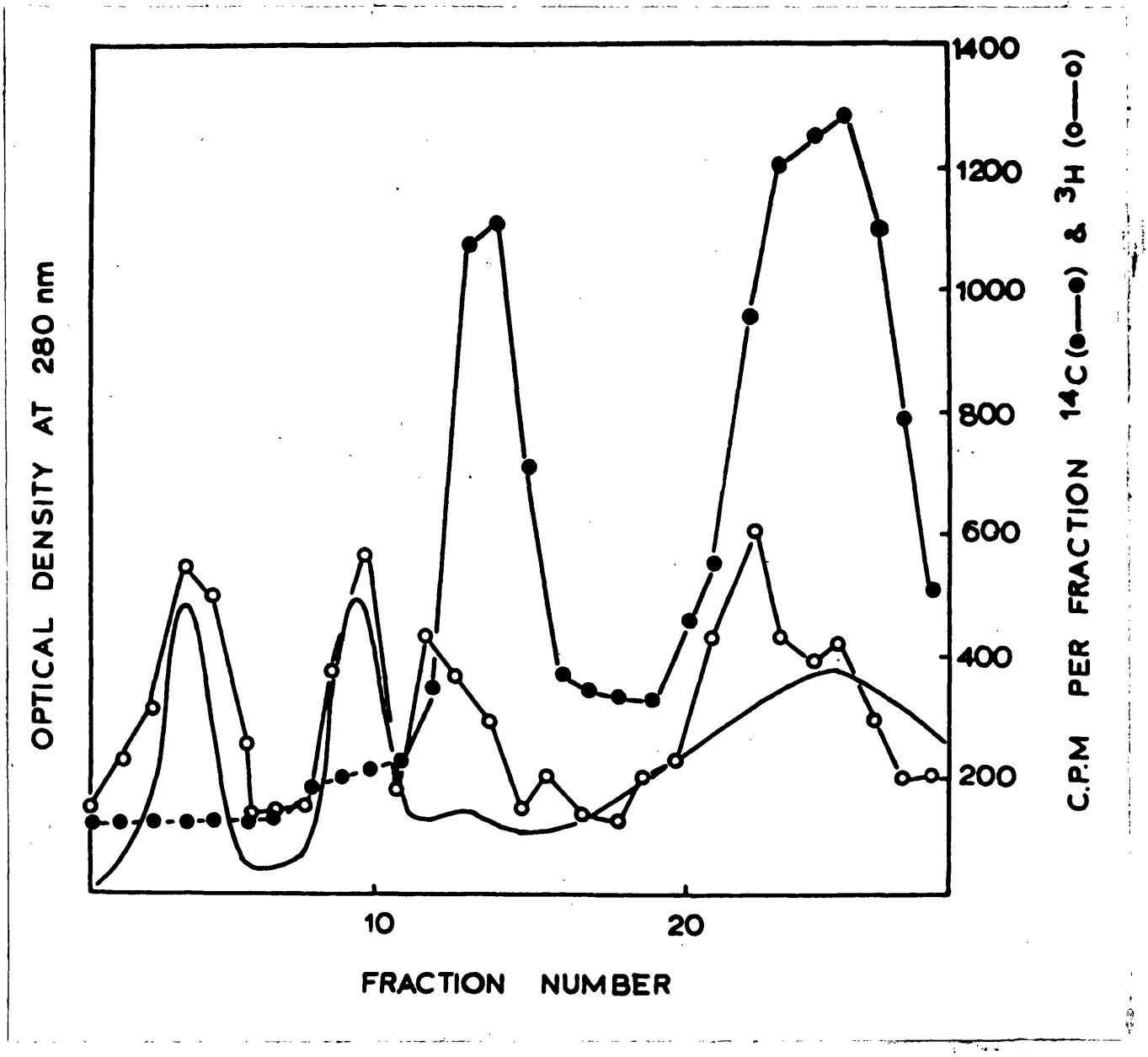


FIGURE 28.

CM cellulose chromatogram of globins from a duck reticulocyte cell-free system, to which was added poly(U) sepharose retained polysomal RNA from 14 day mouse foetal livers.

The absorbance a 280 nm was monitored ——— ;
—O—O—, endogenous incorporation of ^{14}C-leucine;
—O—O—, incorporation of ^{3}H-leucine in the presence of 30 µg of retained foetal liver polysomal RNA.

Figure 28


OPTICAL DENSITY AT 280 nm
C.P.M PER FRACTION ^{14}C(●—●) & ^{3}H (o—o)
1400
1200
1000
800
600
400
200
10
20
FRACTION NUMBER

to one fifth that of pure 9S RNA prepared from reticulocytes.

This value is in agreement with the purity of the RNA as estimated from 2.6% acrylamide gel analysis. Translational studies thus indicate that 0.20 - 0.25% of the foetal liver polysomal RNA is globin mRNA. This is approximately one tenth of the amount of globin mRNA found in anaemic mouse reticulocytes.

3.9 Retention of foetal liver globin sequences by poly(U) sepharose

a) Polysomal RNA

Samples of the polysomal RNA, and the retained and non-retained RNA were titrated with cDNA. 0.48% of the sequences in polysomal RNA are complementary to globin cDNA, by comparison with a cDNA - 9S RNA titration (Table IX). This value is approximately two fold higher than indicated by translational work. The RNA that is retained by the poly(U) sepharose, is considerably enriched in globin sequences. About 38% of the RNA is complementary to globin cDNA. Since 1% of the total RNA was retained by the column, it can be calculated that 80% of the polysome associated globin sequences, are retained by poly(U) sepharose (Table IX). This figure is in good agreement with the value obtained for the retention of reticulocyte polysome associated globin sequences.

b) Cytoplasmic RNA

Total cytoplasmic RNA was prepared using Nonidet NP-40. Cytoplasmic RNA contains a slightly greater proportion of 4S and 5S RNA than foetal liver polysomal RNA, but appears otherwise identical on 2.6% acrylamide gel analysis (Figure 29a). This RNA was chromatographed on poly(U) sepharose, and the starting RNA, and the retained and non-retained RNA titrated with cDNA. By comparison with a cDNA-9S RNA titration curve,

cDNA estimation of globin mRNA sequences in the RNA of 14 day foetal liver polysomes, and the retention of the globin RNA sequences by poly(U) sepharose.

	Proportion of Total RNA in each fraction	Percentage of RNA complementary to globin cDNA	Column 1 x Column 2	Percentage of total globin sequences
Total foetal liver polysomal RNA	1.0	0.48	0.48	-
Retained RNA	0.01	38.0	0.38	78
Non-retained RNA	0.99	0.11	0.109	22
			0.489	

2.6% polyacrylamide electrophoresis of nuclear and cytoplasmic RNAs from erythroid and non-erythroid tissues.

Electrophoresis was in 0.7 cm by 7 cm gels for 15 minutes at 2 mA/gel and then 70 minutes at 6 mA/gel. Gels were scanned at two times magnification, on a Joyce Loebel UV scanner.

a) 70 µg of 14 day foetal liver cytoplasmic RNA.

b) 15 µg of 14 day foetal liver nuclear RNA.

c) 60 µg of adult brain cytoplasmic RNA.

d) 18 µg of adult brain nuclear RNA.

e) 70 µg of adult liver cytoplasmic RNA.

f) 15 µg of adult liver nuclear RNA.

g) 80 µg of L Y cell cytoplasmic RNA.

h) 50 µg of L Y cell nuclear RNA.

Figure 29

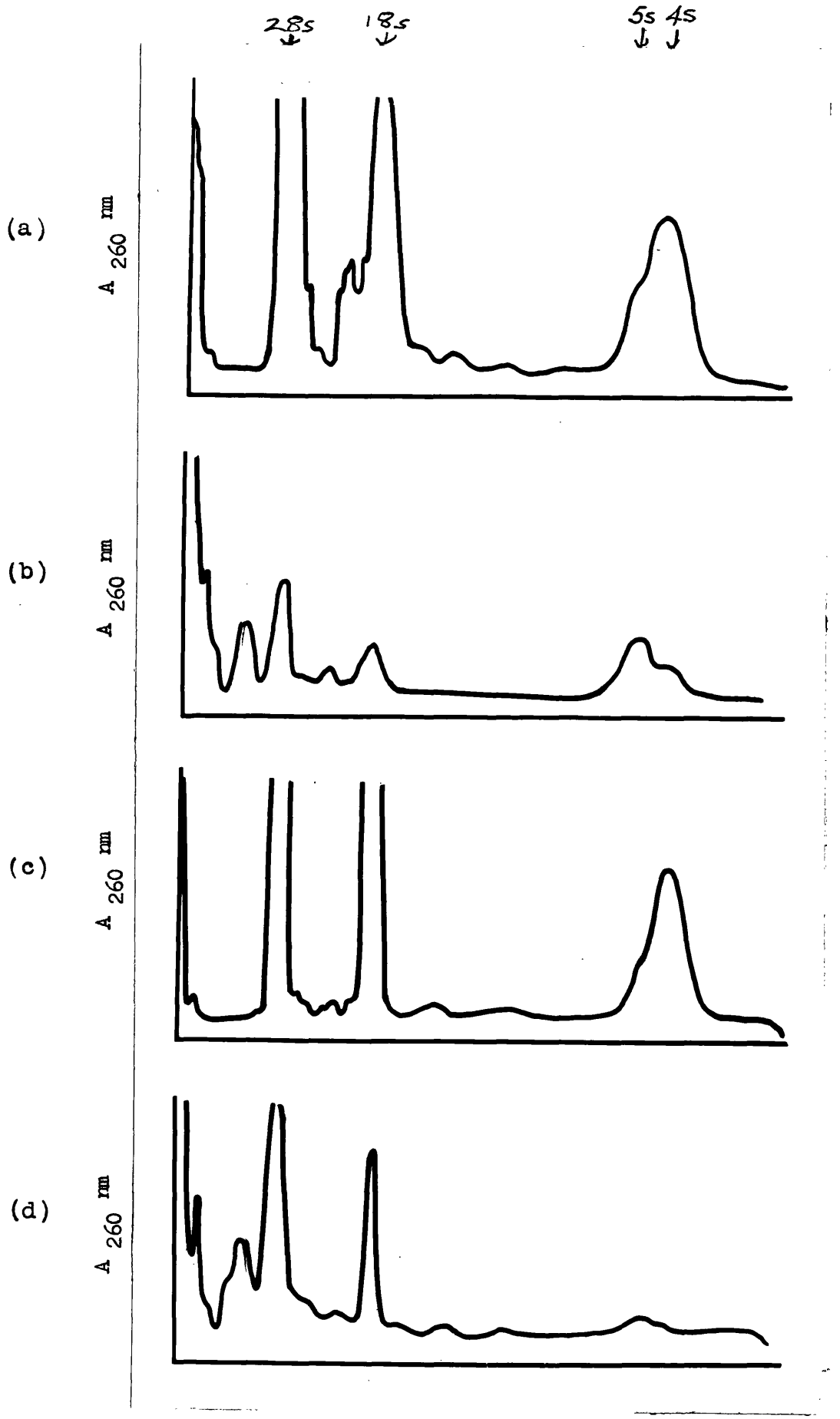
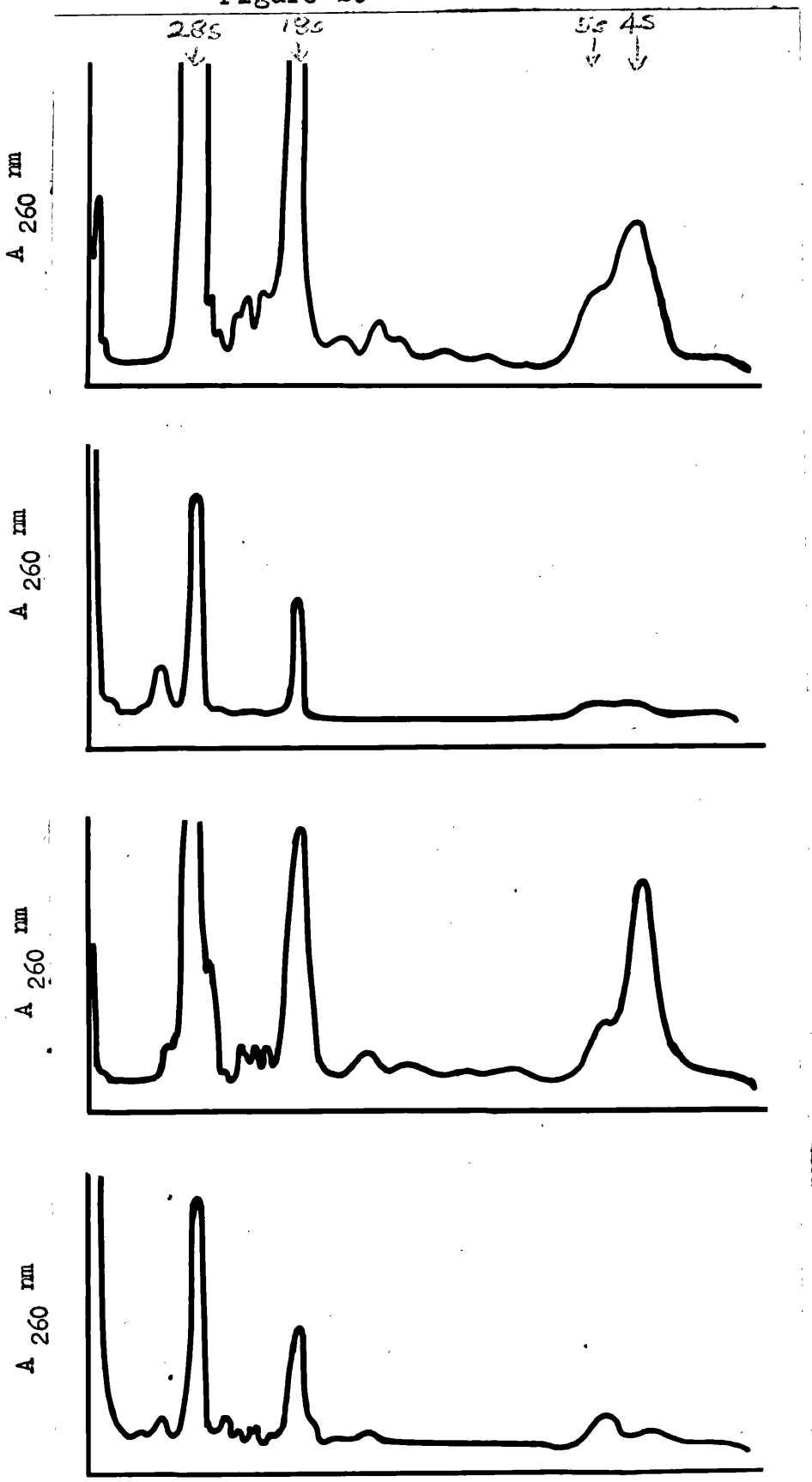


Figure 29



0.3% of the starting RNA is complementary to globin cDNA (Figure 30a). About 0.7% of the RNA was retained by the column. This RNA was considerably enriched in globin RNA sequences, and contained 69% of the total globin sequences present in cytoplasmic RNA (Table X). This figure is about 10% lower than the amount of retained globin RNA sequences in foetal liver polysomal RNA.

In other preparations, between 67 - 72% of the cytoplasmic globin RNA sequences are retained by poly(U) sepharose, whether the RNA is made by Nonidet NP-40, or by the citric acid method (lysis in TKM-sucrose). However, the absolute levels of globin RNA sequences in RNA prepared by NP-40 are approximately twice that of RNA prepared by lysis in TKM-sucrose.

c) In total nuclear RNA

Nuclei were isolated from 14 day foetal livers using either citric acid or Nonidet NP-40. The yield of RNA using the citric acid procedure was about one half that of the Nonidet NP-40 technique. RNA prepared by either method exhibits sharp ribosomal and ribosomal precursor peaks on 2.6% acrylamide gels (Figure 29b), and shows no detectable signs of degradation. Material migrating slower than 32S can be seen on the gel scans. Approximately 10% of the total nuclear RNA prepared in this way sediments at greater than 35S in NETS/sucrose gradients (Dr. M. Getz, personal communication).

Samples of both RNAs were titrated with cDNA. By comparison with a cDNA-9S RNA titration (Figure 30a), 0.15% (± 0.02%) of the RNA is complementary to globin cDNA (average of four different preparations). The RNA was chromatographed on poly(U) sepharose, and the retained and non-retained material titrated with cDNA. Between 1.6 - 2.0% of the RNA was retained

cDNA titration of RNAs from erythroid and non-erythroid tissues.

0.3 ng of cDNA were incubated with increasing amounts of test RNA to a $D_o t$ of 0.12 mol. l^{-1} s (for 1½ days), and hybrids analysed as described. The titration curve for the 9S RNA-cDNA was calculated, and the scale for the other RNAs adjusted for the values to fit this curve.

a) ● 9S RNA-cDNA.

■ foetal liver cytoplasmic RNA-cDNA.

△ foetal liver nuclear RNA-cDNA.

b) ● 9S RNA-cDNA.

▲ adult liver nuclear RNA-cDNA.

□ adult brain nuclear RNA-cDNA.

× L Y cell nuclear RNA-cDNA.

○ E. Coli ribosomal RNA-cDNA.

Figure 30 (a)

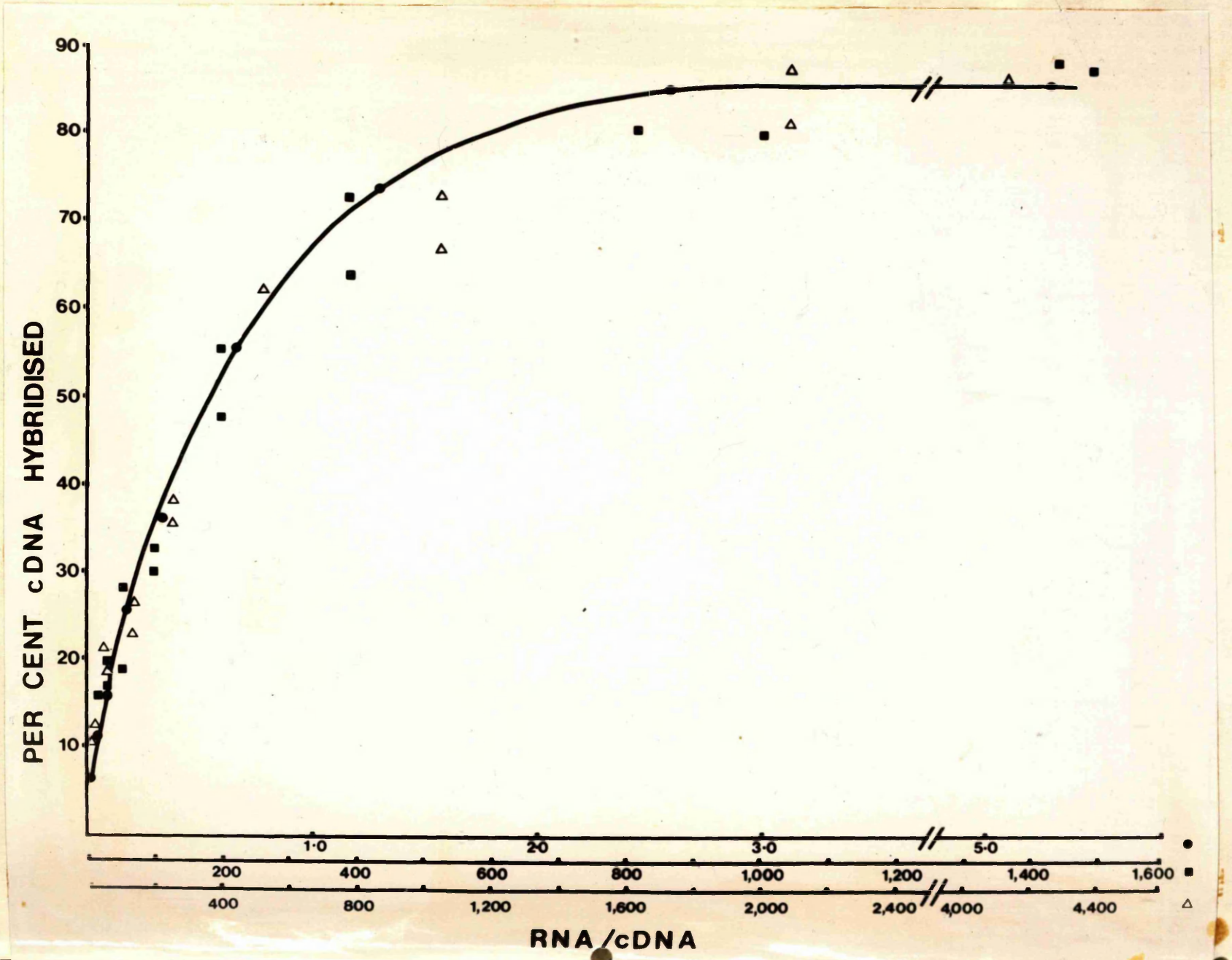

PER CENT cDNA HYBRIDISED
90
80
70
60
50
40
30
20
10
1·0
2·0
3·0
5·0
200
400
600
800
1,000
1,200
1,400
1,600
400
800
1,200
1,600
2,000
2,400
4,000
4,400
RNA /cDNA

Figure 30 (b)

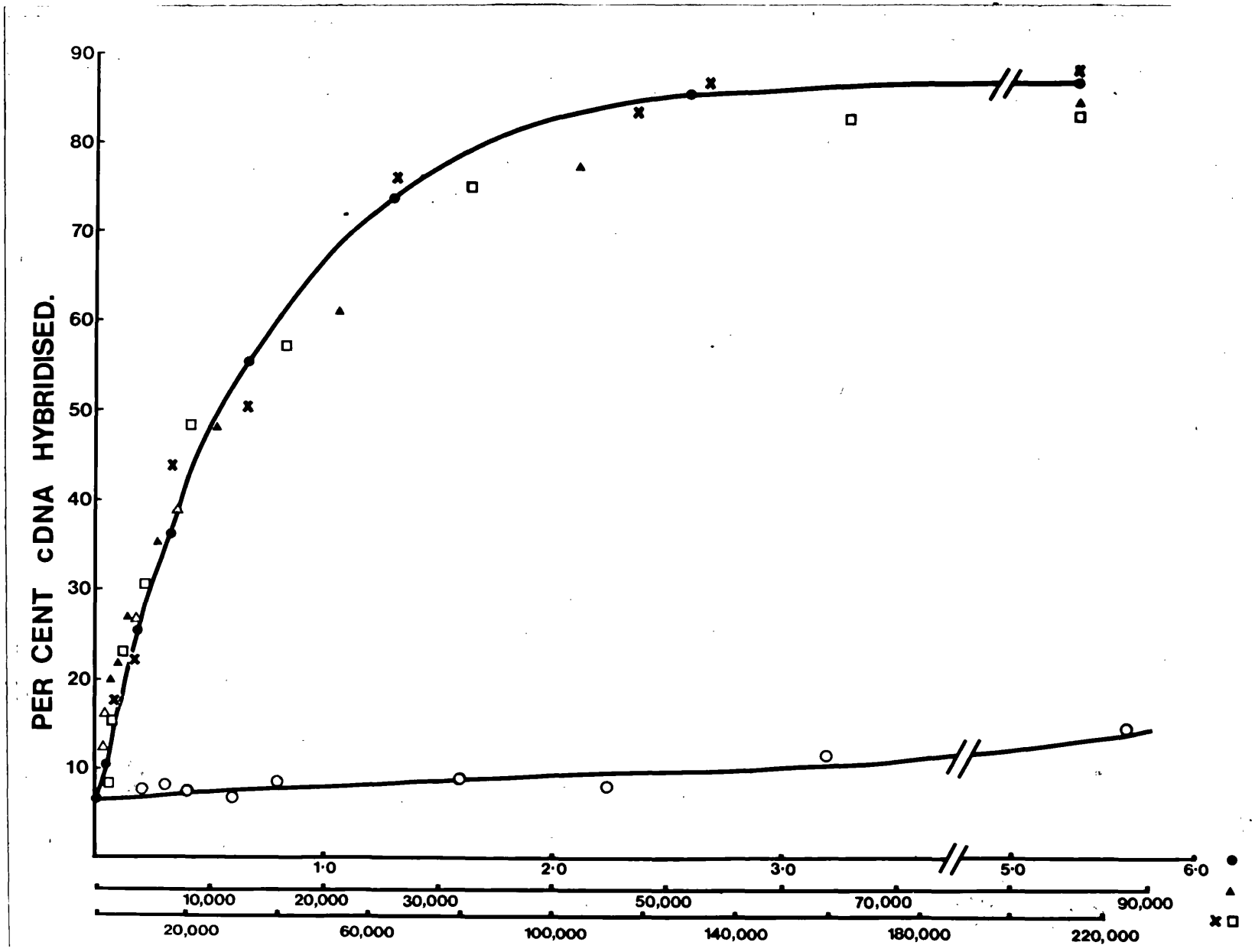

PER CENT cDNA HYBRIDISED.
90
80
70
60
50
40
30
20
10
1·0
2·0
3·0
5·0
6·0
10,000
20,000
30,000
50,000
70,000
90,000
20,000
60,000
100,000
140,000
180,000
220,000

TABLE X.

cDNA estimation of the globin mRNA sequences present in 14 day foetal liver total cytoplasmic RNA, and the retention of these sequences by poly(U) sepharose.

	Proportion of RNA in each fraction	Percentage of RNA complementary to globin cDNA	Column 1 x Column 2.	Percentage of total globin sequences
Total cytoplasmic RNA	1.0	0.30	0.30	-
Retained cytoplasmic RNA	0.7	35	0.245	69.2
Non-retained cytoplasmic RNA	0.993	0.105	0.104	29.8
			0.349	

by the column. This RNA was enriched in globin RNA sequences, and contained about 50% of the total nuclear globin sequences (Table XI).

2.10 Presence of globin RNA sequences in other tissues

a) Preparation of nuclear and cytoplasmic RNA

The method of preparation of RNA was modified for different tissues.

1. Adult mouse livers

Nuclear and cytoplasmic RNA was made from perfused livers of adult male mice using either the citric acid, or the Nonidet NP-40 method. In one preparation, using the citric acid method, 8 mg of cytoplasmic RNA, and 100 μg of G50 - excluded nuclear RNA were obtained from 3 perfused livers. In similar preparations using Nonidet NP-40, 5 - 10 times more nuclear RNA was obtained. Adult liver nuclei are known to be fragile during preparation (Dr. Gilmour, personal communication).

2. Adult mouse brains

20 brains were used to prepare 3 mg of cytoplasmic RNA and 0.6 mg of G50 excluded nuclear RNA using an osmotic shock technique to lyse the cells. Nuclei were pelleted from dense sucrose to reduce lipid contamination.

3. From mouse lymphoma L5178Y cells

4×10^9 cells (a kind gift of G. Lanyon) were grown as described in Gilmour et al., 1974, and used to prepare 9 mg of cytoplasmic RNA and 2 mg of nuclear RNA by the citric acid method.

TABLE XI.

cDNA estimation of the globin RNA sequences in the nuclear RNA of 14 day foetal livers, and the retention of the globin RNA sequences by poly(U) sepharose.

	Proportion of Total RNA in each fraction	Percentage of RNA complementary to globin cDNA	Column 1 x Column 2.	Percentage of total globin sequences.
Total nuclear RNA.	1.0	0.15	0.15	
Retained nuclear RNA	0.017	3.0	0.051	49
Non-retained nuclear RNA.	0.983	0.055	0.054	51
			0.105	

These results are the average of 3 different preparations.

b) Gel analysis of RNA

These cytoplasmic and nuclear RNAs were run on 2.6% acrylamide gels (Figure 29, c–h). The cytoplasmic RNAs contain distinct 28S, 18S and 4S and 5S RNA, and show little evidence of breakdown. The nuclear RNAs are slightly contaminated with material migrating in the 4S region of the gel. This may be fragments of DNA, or represent some breakdown of nuclear RNA. Small amounts of 28S and 18S RNA are present which may reflect a degree of contamination of the nuclei with cytoplasmic RNA. However, distinct 32S ribosomal precursor RNA can be seen, as well as a considerable amount of RNA migrating slower than 32S, or not entering the gel.

c) Analysis of globin RNA sequence content of RNA

These RNAs were titrated with cDNA. The levels of globin RNA sequences in each were calculated by comparison with a 9S RNA–cDNA titration curve (Figure 30b and Table XII).

Hybridisation of cDNA to RNA sequences can be detected in all cytoplasmic and nuclear RNAs. However, cDNA does not become S_1 resistant if incubated to $D_o t$ of 0.033 with large amounts of *E. coli* RNA (Figure 30b). The levels of globin RNA sequences in the nuclear RNAs of brain, adult liver and L Y cells, are 30 – 70 times lower than those found in the nuclear RNA of foetal liver. The amounts found in the cytoplasmic RNA of these tissues are lower than that found in the nuclei by a factor of 2 – 10, and lower than that found in the foetal liver cytoplasm by greater than a factor of 300. Very low levels of globin RNA were detected in the cytoplasmic RNA of L Y cells, and an accurate titration curve was not obtained.

cDNA estimation of the levels of globin RNA sequences in the RNA of adult mouse tissues, and a mouse cell line.

	Percentage of RNA complementary to globin cDNA
Adult liver nuclear RNA	0.005
Adult liver cytoplasmic RNA	0.0009
Adult brain nuclear RNA	0.002
Adult brain cytoplasmic RNA	0.0015
L5178Y cell nuclear RNA	0.002
L5178Y cell cytoplasmic RNA	not greater than 0.00005

d) Retention of the globin RNA sequences from brain and liver RNA by poly(U) sepharose

A sample of the nuclear and cytoplasmic RNAs from these tissues was chromatographed on poly(U) sepharose, and the retained and non-retained RNAs titrated with cDNA. The percentage of sequences complementary to cDNA was calculated by comparison with a cDNA-9S RNA titration curve. The retained RNAs are enriched in globin sequences. 45% of the liver nuclear globin RNA, and 41% of the brain nuclear globin RNA sequences are retained by the column (Table XIII). These values are slightly lower than that found in foetal liver nuclear RNA. Less than 40% of the globin sequences found in liver cytoplasmic RNA are retained by poly(U) sepharose, but more than 80% of those found in brain cytoplasmic RNA are retained.

3.11 Estimation of the reticulocyte contamination of the nuclei preparations

The levels of globin sequences found in the cytoplasm of the adult tissues can be partly explained by contamination with circulating reticulocytes. In normal adult mice, about 2.0% of the circulating blood cells are reticulocytes (Wintrobe, 1961). However, contamination of nuclear material with intact reticulocytes seems unlikely.

This was tested by homogenising the blood cells from anaemic mice in TKM-sucrose, followed by the normal citric-acid procedure. The amount of nucleic acid in each fraction was estimated from the absorbance at 260 nm of the sample, after phenol extraction and precipitation with ethanol (Table XIV). The fraction normally used to prepare cytoplasmic RNA contains 5% of the possible reticulocyte RNA, and the nuclear

cDNA estimation of the retention of the globin RNA sequences present in the RNA of adult mouse brain, and liver

	Proportion of total RNA in each fraction	Percentage of RNA complementary to globin cDNA	Column 1 x Column 2	Percentage of total globin sequences
Retained adult liver nuclear RNA	0.017	0.12	0.00204	45.4
Non-retained adult liver nuclear RNA	0.983	0.0025	0.00246	54.6
Retained adult liver cytoplasmic RNA	0.005	0.05	0.00025	36.0
Non-retained adult cytoplasmic RNA	0.995	0.00045	0.00045	64.0
Retained adult brain nuclear RNA	0.023	0.045	0.00103	41.4
Non-retained adult brain nuclear RNA	0.977	0.0015	0.00147	58.6
Retained adult brain cytoplasmic RNA	0.016	0.155	0.00248	84.8
Non-retained adult brain cytoplasmic RNA	0.984	0.00045	0.00044	15.2

TABLE XIV

Distribution of globin mRNA sequences from anaemic mouse reticulocytes following extraction by citric acid procedure.

	Amount of nucleic acid in each fraction(μg)	Percentage of total nucleic acid in each fraction
TKM/sucrose homogenisation	298	5.1
Combined supernatants of citric acid washes	5,600	94.7
Pellet after 3 X citric acid washes	12.6	0.2

fraction 0.2%. Gel analysis of these fractions showed that the nucleic acid in the nuclear pellet was DNA, presumably from white cells in the blood. The material in all other fractions was indistinguishable from normal reticulocyte cytoplasmic RNA.

3.12 Analysis of hybrids

The fidelity of the hybrids formed between globin cDNA, and the sequences present in brain, adult liver and LY cell nuclear RNA, and foetal liver nuclear and cytoplasmic RNA, were compared with a globin RNA-cDNA hybrid.

a) Melting analysis of hybrids

An estimate of mismatching in the hybrid formed between cDNA, and sequences in the RNA of different sources, can be obtained by comparing the T_m of the hybrid with that of a cDNA-9S RNA hybrid. Hybrids were melted in 50% formamide/ hybridisation buffer. The T_m of cDNA-9S RNA under these conditions is 69° (Figure 31). The T_m of the other hybrids is slightly lower than this (Figure 31, Table XV). This difference represents at most 3% mismatching of sequences in the hybridising RNA (Laird et al., 1969; Ullman and McCarthy, 1973a,b) as compared with 9S RNA-cDNA. This is not significantly greater than the expected experimental error.

b) Size analysis of hybrid molecules

A sample of each hybrid was treated with S_1 nuclease. Any short single stranded regions in the RNA-cDNA hybrid will be nicked. Recent data suggests that S_1 nuclease may nick a single mismatched base in a hybrid molecule (Dr. P. Harrison, personal communication). Carrier RNA was added, and the sample desalted through G_{25} sephadex. Each sample was then mixed

Melting behaviour of cDNA–RNA hybrids from erythroid and non–erythroid cells.

20,000 cpm of globin cDNA were incubated in 30 µl of formamide–hybridisation buffer with test RNA to 14 times $D_o t_{\frac{1}{2}}$ (5 days). This solution was diluted 10 fold with formamide–hybridisation buffer, and incubated in sealed capillaries containing 20 µl for 5 minutes, at the appropriate temperature. Analysis of hybrids was then as described in the text. The ratios ofRNA/cDNA during annealing were the lowest necessary to convert 80% of cDNA into hybrid form. The data were normalised to correct for the amount of cDNA which failed to form S_1 nuclease resistant hybrid (10 – 15%) and for the S_1 resistant fraction of cDNA at 96° (about 5%).

- O ● 9S RNA–cDNA.
- ▲ adult liver nuclear RNA–cDNA.
- ■ foetal liver nuclear RNA–cDNA.
- □ adult brain nuclear RNA–cDNA.
- × L Y cell nuclear RNA–cDNA.
- △ foetal liver cytoplasmic RNA–cDNA.

Figure 3I

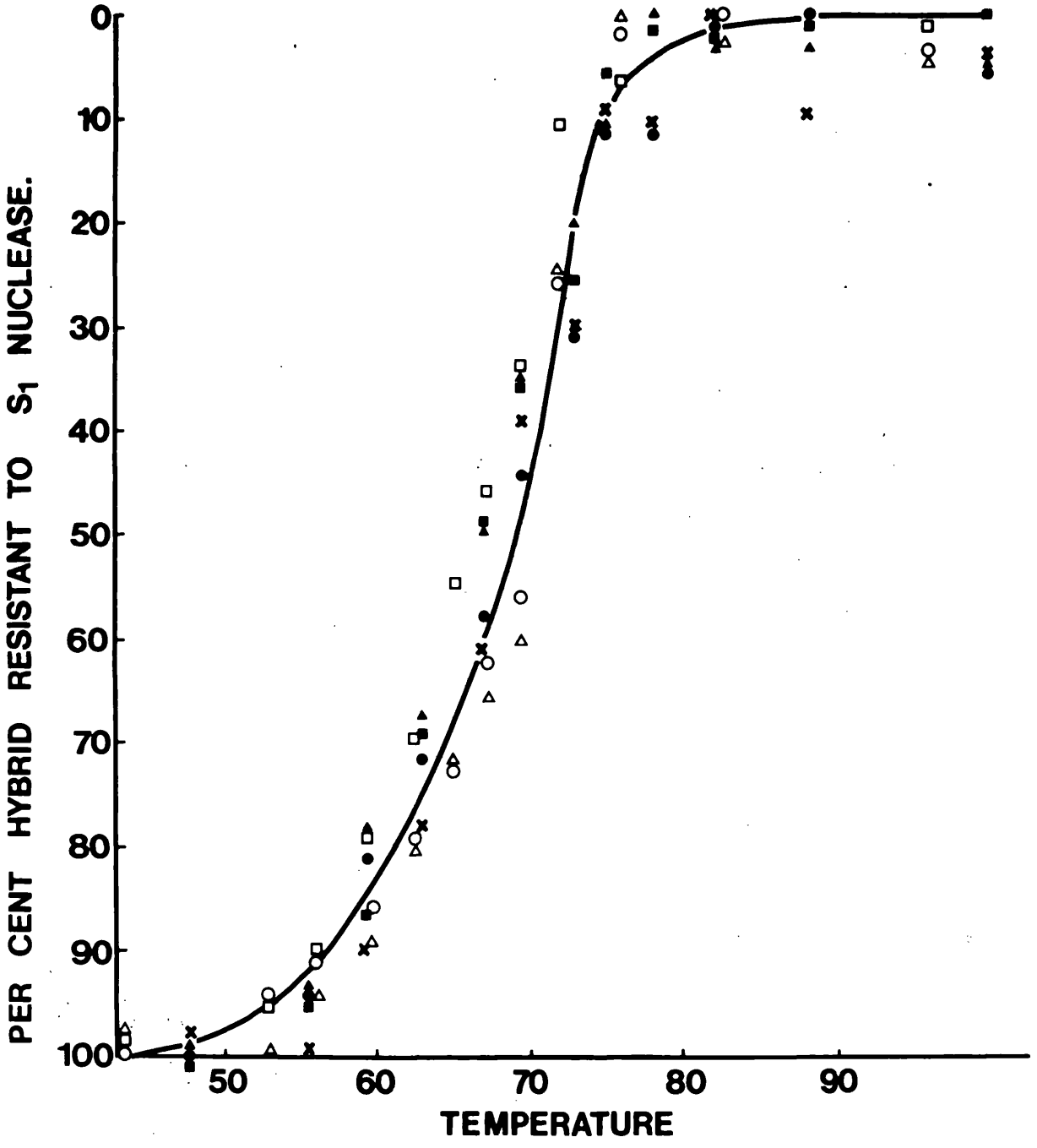

PER CENT HYBRID RESISTANT TO S_1 NUCLEASE.
0
10
20
30
40
50
60
70
80
90
100
50
60
70
80
90
TEMPERATURE

TABLE XV

Analysis of hybrids of cDNA, and RNA from erythroid and non-erythroid tissues

	Melting Temperature °C	Maximum Percentage Mismatching	Mo.wt of cDNA after S_1 treatment Daltons	Possible Number of Nicks/hybrid
9S RNA	69.2	-	49,000	1
Brain nuclear RNA	66.6	3%	33,000	2
Liver nuclear RNA	66.5	3%	34,000	2
Foetal liver nuclear RNA	68.2	1%	49,000	1
Foetal liver cytoplasmic RNA	67.0	2%	29,000	2
L Y cell nuclear RNA	69.0	0%	-	-

The original size of the cDNA used in this experiment was 100,000 daltons.

with a marker DNA of known size, and layered onto a 5 - 10% alkaline sucrose gradient.

After centrifugation, the DNA was fractionated by upward displacement with 20% sucrose, the refractive index and absorbance at 260 nm measured, and an aliquot of each fraction counted (Figure 32). The sedimentation profile of the marker DNA was the same in each tube.

From the calculated peak molecular weight of each cDNA (Table XV) it can be seen that after S_1 nuclease treatment, the cDNA hybridised to 9S RNA and foetal liver nuclear RNA has been nicked on average once, and the hybrid from foetal liver cytoplasmic RNA, and brain and adult liver nuclear RNA, has been nicked twice.

3.13 Absolute amounts of globin RNA sequences in different tissues

Using a modified Schmidt-Tannhauser method, the relative amounts of DNA and RNA were estimated in the erythroid and non-erythroid tissues, and in nuclei isolated from these tissues. From this the relative amounts of DNA in the nucleus and cytoplasm can be calculated (Table XVI). Combining this with data in Table XII, the absolute concentration of globin sequences in the tissue studied was calculated (Table XVII).

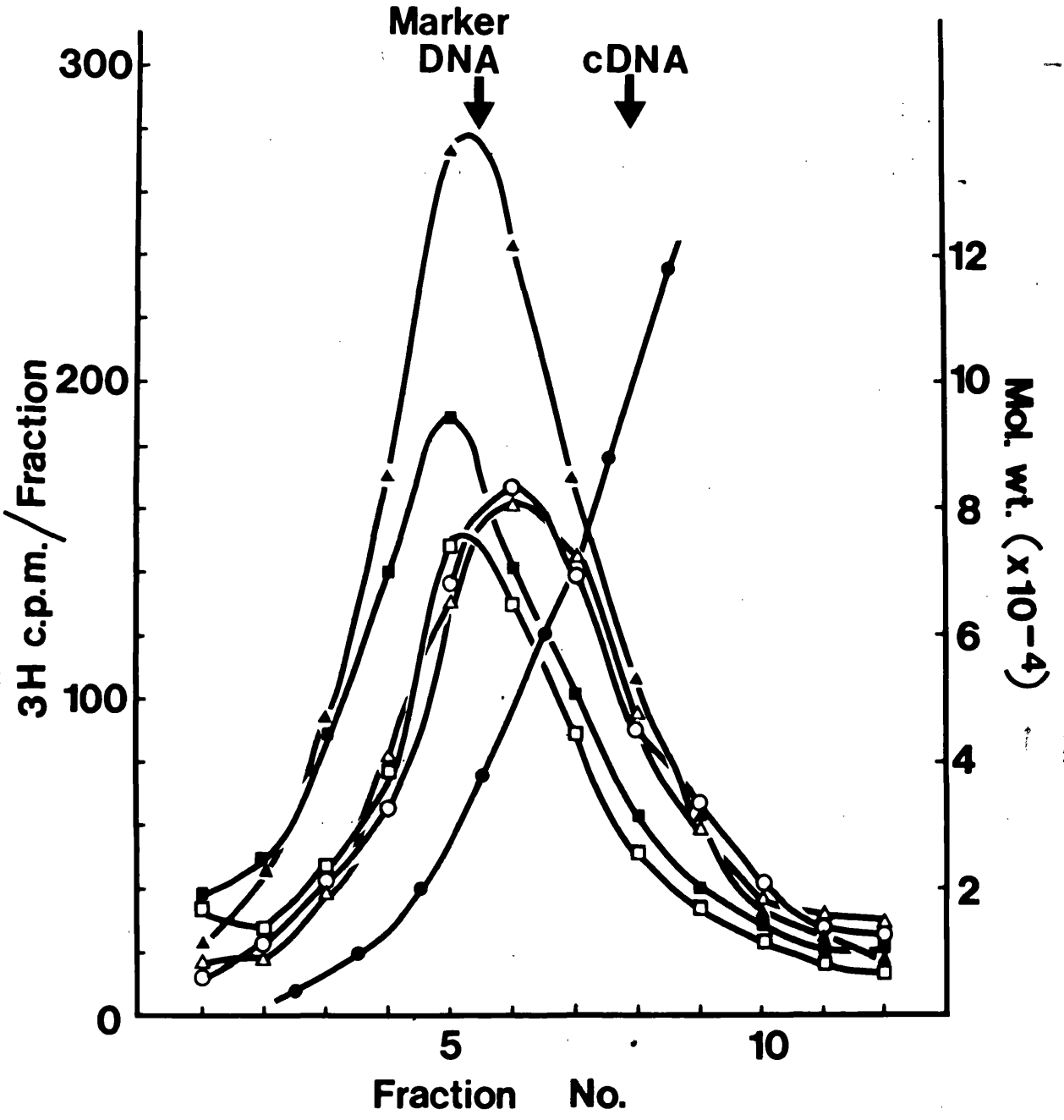
There is a 25 - 75 fold difference in globin sequences in the nucleus of the erythropoietic tissue, when compared with the other tissues, but a much greater difference of 200-4,000 fold in the cytoplasm. A higher percentage of the cellular globin sequences are thus found in the cytoplasm of foetal livers, than adult livers or brains.

FIGURE 32.

Size analysis of the S_1 nuclease resistant cDNA hybridised to RNA from erythroid and non-erythroid cells.

Samples were centrifuged on 5 – 10% alkaline sucrose gradients in an M.S.E. 6 x 15 ml swing-out rotor, at 24,000 rpm for 38 hours at 20°. 0.75 ml fractions were collected, an aliquot counted and the A_{260} measured. The peak absorbance of the 4S marker DNA was the same in each gradient. The refractive index of each fraction was measured, and the molecular weight calculated —●—●—; The size of cDNA from a 9S RNA hybrid —○—○—, adult liver nuclear RNA—▲—▲—, foetal liver nuclear RNA —■—■—, adult brain nuclear RNA—□—□—, and foetal liver cytoplasmic RNA —Δ—Δ—, is shown. The peak fraction of unhybridised cDNA, centrifuged in a parallel gradient is shown by an arrow.

Figure 32


Marker DNA
cDNA
300
200
100
0
3H c.p.m./Fraction
12
10
8
6
4
2
Mol. wt. (x10-4)
5
10
Fraction No.

Estimation of the relative amounts of RNA in the nucleus and cytoplasm of mouse brains, and adult and foetal livers.

	RNA : DNA ratio of total material	RNA : DNA ratio of nuclear material	Ratio of cytoplasmic RNA: nuclear RNA	Percentage of Total cell RNA in cytoplasmic and nuclear RNA
Adult mouse brain	1.065 : 1	0.155 : 1	0.91 : 0.155	85.4 : 14.67
Adult mouse liver	3.82 : 1	0.187 : 1	3.633 : 0.187	95.1 : 4.9
Mouse foetal liver	1.04 : 1	0.148 : 1	0.892 : 0.148	85.6 : 14.4

TABLE XVII

Estimation of the number of molecules of globin RNA present in the nucleus and cytoplasm of erythroid and non-erythroid tissues.

	RNA content in pg/cell	Fraction of RNA complementary to cDNA x 10^6	Number of molecules of 9S mRNA per cell	Percentage of total globin RNA
Adult Liver nuclear RNA	1.12	50	150	38
Adult Liver cytoplasmic RNA	21.8	4	240	62
Adult Brain nuclear RNA	0.93	20	50	17
Adult Brain cytoplasmic RNA	5.46	16	240	83
L51787 cell nuclear RNA	0.9	20	50	83
L51787 cell cytoplasmic RNA	5.1	0.5	10	17
14 day foetal liver nuclear RNA	0.89	1,500	3,700	8
14 day foetal liver cytoplasmic RNA	5.35	3,000	44,000	92
Reticulocyte cytoplasmic RNA	2.6	20,000	140,000	-

$$\text{Number of molecules 9S RNA} = \frac{\text{Avogadros Number}}{\text{Mol. wt. 9S mRNA}} \times \text{weight RNA complementary to cDNA}$$

DISCUSSION

DISCUSSION

4.1 Before attempting to draw conclusions from the work presented here, two fundamental assumptions must be critically examined. They are:-

a) That mouse reticulocyte 9S RNA comprises mainly the messenger RNAs for α and β globins.

b) That cDNA is a faithful transcript of most of the sequences present in 9S RNA.

a) Evidence that 9S RNA comprises α and β globin mRNA

1) Since about 90% of protein synthesis in reticulocytes is represented by globin, it is assumed that the purity of the mRNA prepared from reticulocytes must approach this value.

2) The 9S RNA is selected by its retention by poly(U) sepharose, and must therefore contain poly(A). This excludes contamination with non-messenger RNA molecules, such as fragments of ribosomal RNA. This retained RNA runs as a diffuse 9S peak in 2.6% (Figure 1) and 6% (Figure 16a) polyacrylamide gels, with little other contaminating RNAs.

The peak molecular weight of this RNA has been calculated to be 220,000, with a molecular weight range in the band of 40,000. This represents 660 bases $\pm$ 120 (Williamson et al., 1971). When run on denaturing formamide gels, only two peaks of RNA can be resolved (Morrison et al., 1974).

Thus, by these physical criteria, 9S RNA contains enough information to code for two proteins of about 16,000 molecular weight.

3) 9S RNA directs the synthesis of mouse α and β globin in the duck and rabbit reticulocyte cell-free system, and the Krebs II ascites cell-free system (Figure 3, 22 and 21 respectively). Lane et al. (1971) have studied the tryptic

peptides released from non-oocyte proteins synthesised after injection of rabbit reticulocyte 9S RNA into oocytes. No peptide was observed that did not correspond to peptides obtained by tryptic digests of carrier haemoglobin.

These three points are strong evidence that 9S RNA comprises mainly the messages for α and b globin.

b) Evidence that complementary DNA is a faithful transcript of 9S RNA

1) cDNA is about 60% of the size of the 9S RNA molecule and contains about 40 – 45% (G + C) (Harrison et al., 1974a). Moreover, after depurination, only about 4% of the nucleotides in cDNA bind to poly(A) sepharose under conditions in which poly(dT) binds completely (J. Mansbridge, personal communication). cDNA contains, therefore, a very small proportion of poly(dT) sequences. About 85% of cDNA sequences hybridise to excess 9S RNA at a temperature 15° below the T_m of a 40% (G + C) RNA–DNA hybrid under these conditions (Bishop, 1972).

2) The bulk of the cDNA sequences represent reasonably faithful transcripts of the 9S RNA sequences. This is shown by the fact that only about one nick per 9S RNA–cDNA hybrid molecule is introduced by treatment with S_1 nuclease (Figure 6). Any regions of mismatching, therefore, must be either too small for S_1 nuclease to recognise or be clustered. An independent estimate of mismatching in the cDNA–9S RNA hybrid can be obtained by comparing its T_m (69°) with that of reannealed mouse DNA fragments (77 to 78°) (Figure 7). After correcting for slight differences in (G + C) content, this means that the T_m of cDNA–9S RNA hybrid is 9 – 10° below that of the reannealed DNA duplex of the same (G + C) content (Gruenwedel et al., 1971). This difference in T_m is 3 – 4° more than that reported for

perfect RNA–DNA hybrids (Kohne, 1969; Gelderman et al., 1971), which implies at most 2% mismatching of sequences in cDNA with respect of 9S RNA (Laird et al., 1969; Ullman and McCarthy, 1973a, b). One explanation of this mismatching is that cross–hybridisation of α and β globin cDNAs to β and α globin mRNAs in 9S RNA might occur. However, this appears most unlikely from consideration of the amino acid sequences of α and β mouse globins (Dayhoff, 1972). The most likely explanation of the mismatching is that reverse transcriptase introduces a low frequency of mistakes during transcription of 9S RNA in vitro.

The results show that most of the cDNA represents a faithful, but partial, transcript of sequences present in the purest preparations of 9S RNA available, which contain mainly globin mRNAs. Nevertheless, it is conceivable that cDNA might represent transcripts of contaminant RNA sequences in 9S RNA that do not specify globins.

3) The titration of a fixed amount of cDNA with increasing amounts of 9S RNA can be used to determine whether cDNA consists of copies of all or some of the classes of molecules present in 9S RNA. The result of such a titration experiment is shown in Figure 5.

The concentration of cDNA used for each point on the titration curve was determined from the specific activity of cDNA, which in turn was estimated from the specific activity of the nucleotide precursor, assuming a 50% (C + T) content in cDNA. With such small amounts of cDNA it is not possible to confirm independently the assumed specific activity, and hence any errors in the estimate will affect the RNA/cDNA ratios. This ratio is similarly affected by

errors in the RNA concentrations, which were obtained by serial dilution.

Maximum levels of hybridisation were obtained at RNA/cDNA ratios of 1 – 1.4. This indicates that the cDNA is complementary to 60% of the RNA sequences, a finding that is consistent with the size of the cDNA.

In a more rigorous theoretical analysis, Young et al. (1974) include the assumption that in cDNA excess, there is a greater probability of hybridisation of the larger fragments of cDNA. Predictions from this are in close agreement with the observed data in Figure 5, and it can be concluded that experimental errors in the RNA/cDNA ratios are minimal, and also that virtually all classes of molecules in 9S RNA are transcribed.

4) Young et al. (1974) have measured the complexity of this cDNA. The kinetics of the hybridisation of cDNA to 9S RNA are in agreement with second-order reaction theory. Hence, the time-course of the hybridisation of cDNA to an equal amount of 9S RNA can be converted, using the appropriate factor, to the DNA-DNA reaction of the same base composition and complexity. By comparison with the renaturation of E. coli DNA, it was estimated that cDNA was transcribed from a total RNA sequence of between 700 and 800 bases, i.e. longer than a single mRNA molecule. This suggests that cDNA consists of partial copies of two different mRNAs, giving a total complexity of 700 – 800 bases, with a mean fragment size of 330 bases. Since the titration experiment shows that cDNA consists of copies of virtually all classes of molecules present in 9S RNA, it is probable that cDNA includes transcripts of both α and β mRNAs. If cDNA transcripts of α and β mRNA are sufficiently different to prevent cross-reaction (Dayhoff, 1972), the most likely

interpretation of these results is that cDNA comprises partial transcripts of both α and β globin mRNA. Furthermore, the requirement for poly(dT) primer for 9S RNA transcription and the presence of adenine-rich sequences at the 3' terminal (Mansbridge et al., 1974; Burr and Lingrel, 1971; Lim and Canellakis, 1970) strongly suggests that such partial transcripts are initiated at the 3' end of the mRNA.

The results from these experiments thus support the conclusion that cDNA represents transcripts of both α and β globin mRNAs, each transcript being a partial, non-random copy of its template 9S RNA molecule.

Although there is as yet no proof, it is assumed that any slight imbalance of α and b globin mRNAs, as noted by Morrison et al. (1974), will be reflected by a similar imbalance in the cDNA transcript.

4.2 Factors affecting hybridisation reactions

The use of cDNA as a probe for globin sequences in DNA and RNA first requires a knowledge of the factors affecting (1) the specificity of the reaction between cDNA and sequences complementary to it, and (2) the rate at which this reaction will occur.

1) The specificity of the hybridisation reactions

The related problems here are the complementarity required of 2 reacting molecules before a stable duplex between them can form, and the proportion of non-complementary base pairs that can be accommodated before the partial hybrid will be detected as single stranded. With respect to this latter point, the two methods of detection used in these studies differ considerably in their stringency. The first, the binding of hybridised molecules on hydroxyapatite is relatively non-stringent, and probably a region of about 17 base pairs may allow the retention

of a large, partially-hybridised molecule (Wilson and Thomas, 1973). The second method, the recognition of unhybridised sequences by S_1 nuclease, is highly specific, with both single-stranded ends, and internal regions of mis-matching (possibly as little as one mis-matched base pair (Dr. P. Harrison, personal communication). being degraded by the enzyme.

The properties of a hybrid molecule are most easily defined by its mean thermal dissociation (Tm), and using this parameter, the effect of several factors on the stringency of reaction can be determined.

The temperature of incubation and the cation concentration considerably affect the rate and specificity of reaction. The reaction proceeds most rapidly at high cation concentrations, and at elevated temperatures up to 20 - 25° below the Tm of the hybrid (Marmur and Doty, 1962; Wetmur and Davidson, 1968). Short, or highly mis-matched duplexes have a low Tm, and thus will not form at high temperatures or low salt concentrations. Such conditions are thus used for experiments with cDNA, where the formation of mis-paired hybrids is not desired. For studies using heat-labile RNA, similar stringent conditions at lower temperatures have been worked out with organic solvents such as formamide (McConaughy et al., 1969).

From several studies it appears that the minimum stable length of a hybrid may be about 10 - 20 base pairs, with a G - C rich fragment requiring a shorter minimum length, than an A - T rich fragment. Above this length, the stability of the oligomer approaches that of the polymer of similar base composition (Lipsett, 1964; Niyogi and Thomas, 1967; Rüger and Bautz, 1968). The effect of mis-matching on the Tm of a hybrid has been determined from studies using synthetic polymers and

modified bacterial polynucleotides, and indicates that the Tm of the resulting duplexes is reduced by about 0.7°C per 1% base modification (Bautz and Bautz, 1964; Laird et al., 1969; Ullman and McCarthy, 1973a and b).

Thus even under stringent conditions of incubation, a considerable degree of mismatched bases may not significantly reduce the ability of two molecules to form a hybrid. If the minimum stable length of hybridisation is 20 nucleotides, it can be estimated (Walker, 1969) that up to 14% random base change in a molecule of 500 nucleotides, will not greatly reduce the proportion of hybridising molecules which have less than one of these 20 nucleotide regions. All these molecules will thus be detected as hybrids by hydroxyapatite. The proportion of molecules that can form a hybrid rapidly decreases as further base changes accumulate, and about 40% base changes will prevent hybrid formation completely.

The effect of unhybridised free ends on the Tm appears to be smaller than for internally unhybridised regions (Subirana, 1966; Marmur et al., 1963).

Finally, the sugar backbone of the reacting molecules will affect the stability of the hybrid. The greater range of stabilities of RNA–DNA hybrids is reflected by a greater breadth of thermal transition (Chamberlin and Patterson, 1965), but for natural polynucleotides, the RNA–DNA hybrids are less stable than DNA–DNA hybrids (Bolton and McCarthy, 1964).

Factors affecting the rate of hybridisation

It has now been well established that the renaturation of DNA follows second order kinetics (Nygaard and Hall, 1969; Britten and Kohne, 1968; Wetmur and Davidson, 1969). The rate limiting step in this reaction appears to be the formation

of one, or a few, correct base pairs at some in-register point along the two strands, and is followed by a rapid zippering reaction of the rest of the base pairs. Although the rate of reaction will thus increase with fragment size, a steric effect of large molecules results in the observation that the rate is proportional to the square root of the length (Wetmur and Davidson, 1968). The rate is a maximum at a temperature 25° below the hybrid Tm (Marmur and Doty, 1962), and is reduced by solvent viscosity (Subirana, 1966). High G – C content and high ionic conditions will also increase the rate of reaction (Wetmur and Davidson, 1968; Schildkraut et al., 1961).

Thus for any unique DNA (such as a viral or prokaryotic DNA) the rate of reassociation is inversely proportional to the sequence complexity (Marmur and Doty, 1961). In higher organisms, both fast, intermediate and slowly annealing sequences have been observed (Britten and Kohne, 1968). However, interpretation of these annealing classes depends on the effect of mismatching on the rate of annealing of these sequences. During the formation of a duplex, the instability introduced by a mismatched base appears to have a large effect on the rate of hybrid formation (Southern 1971; Sutton and McCallum, 1971; Britten and Bonner, 1971). Thus if base sequences differ only by 10%, the rate of the heterologous hybridisation reaction between them is so reduced that the reaction with the exactly complementary partner will occur.

Clearly for experiments where the rate of DNA annealing is being measured, the presence of mismatched sequences must be taken into account.

4.3 Estimation of the number of globin genes in the mouse genome

As can be seen from Figure 11, the rate of annealing of cDNA to mouse DNA is the same as that for the reassociation of mouse DNA, whether mouse DNA is isolated from erythroid cells (foetal liver) or sperm. The sizes of cDNA and the mouse DNA fragments are not reduced significantly during the period of annealing (Figure 12), and little correction is required for mismatching in the cDNA–mouse DNA duplexes (Figure 7). Since the slight extent of mismatching in cDNA–9S RNA hybrids is the same as that in cDNA–mouse DNA duplexes, cDNA must anneal only to those globin genes expressed in mouse reticulocyte 9S RNA, or to very closely related variants. In the mouse, the only globin genes that might possibly be sufficiently similar in this respect are the β and z chains (14% amino acid differences). Mouse α and x chains are quite dissimilar (36% amino acid differences)(Steinheider *et al.*, 1971).

These facts taken together show that in all the mouse tissues studied, the sequences in their DNA complementary to cDNA – the α and β globin genes – are about as frequent as the most slowly reassociating DNA sequences.

However, by comparison with *E. coli* DNA the observed rate of hybridisation of mouse DNA is about five times as high as predicted, suggesting that the "unique sequences" are present as five copies in the mouse genome. Harrison *et al.* (1974a) have compared the rates of reassociation of mouse DNA fragments of various sizes. They observed that with fragments of larger or smaller size, the bulk of the DNA reassociates more quickly or slowly respectively, as predicted by Wetmur and Davidson (1968) and Sutton and McCallum (1971). The discrepancy between observed and predicted values decreased with decreasing fragment

size, and with the smallest fragment size (120 bases) the bulk of the mouse DNA sequences appear to be unique. This is almost certainly explained in terms of the greater separation of unique and repetitite sequences in mouse DNA with very small fragments.

Other workers using DNA fragments of about 500 nucleotides have also observed this faster rate of reassociation. Thus, Melli et al. (1971) observed a fourfold faster reassociation of sonicated rat DNA fragments than predicted for single-copy sequences. Bishop et al. (1972) and Bishop and Rosbash (1973) have reported a similar, though slightly smaller, discrepancy for duck RNA fragmented by sonication, whether duplex formation was assayed optically, by hydroxyapatite chromatography or resistance to single-strand-specific nuclease. This eliminates the possibility that the method for assay of duplex is the explanation of the discrepancy.

From the data in Figure 11, and the complexity of the cDNA (Young et al., 1974) an estimate can be made of the amount of DNA per haploid mouse genome, that is complementary to the cDNA. These calculations substantiate the conclusion that the cDNA is hybridising to only a single copy of the α and β genes in the DNA from the mouse tissues studied (Page 77).

In a study of DNA made from whole 14 day mouse embryos, similar values for the rates of hybridisation of cDNA to the DNA, and for the number of copies of globin genes per genome, have been observed (Harrison et al., 1974a).

It can thus be concluded that in the mouse: (a) there are probably only single, but not more than four to five copies each of the α and β globin genes in the mouse germ line; (b) that there is no globin gene amplification in erythroid cells, since

foetal liver in this stage of development contains 80% erythroid cells, of which all but the most immature contain detectable amounts of globin mRNA (Harrison *et al.*, 1973); and (c) that there is no widespread elimination of globin genes during non-erythroid somatic differentiation, since the globin gene dosage in sperm, total embryos and erythroid cells is the same. This conclusion concerning the numbers of genes in various mouse tissues would be valid even if episomal copies of the globin gene existed. Since we prepared all DNAs by the hydroxyapatite method (which would retain even gene-size fragments of DNA), the DNAs used in these experiments should include any nuclear DNA sequences, whether of chromosomal origin or otherwise.

These findings are in agreement with other recent reports, which compare the rate of annealing of cDNA or *in vivo* labelled duck reticulocyte 10S RNA to excess duck DNA, with that expected theoretically for reassociation of unique duck DNA sequences. These reports show that the globin genes in the duck may be slightly reiterated (two to three copies; Bishop *et al.*, 1972; Packman *et al.*, 1972; Bishop and Rosbash, 1973), and that there is no globin gene amplification in duck erythroid cells (Packman *et al.*, 1972). These studies did not investigate whether degradation of cDNA or mismatching of cDNA with respect to duck DNA introduced any correction to these quoted values.

These results imply that transcriptional and post-transcriptional mechanisms must be entirely capable of regulating the levels of globin gene expression in mouse tissues. Different transcriptional availability of the globin genes has been clearly shown by several workers (Axel *et al.*, 1973;

Gilmour and Paul, 1973; Steggles et al., 1974). Post-transcriptional mechanisms for the regulation of gene expression are less clearly documented, although control processes at the level of mRNA translation have been demonstrated (Nudel et al., 1973; Wiggle and Smith, 1973; Heywood et al., 1974).

The poly(A) sequence of eukaryotic nuclear RNA, and cytoplasmic mRNA has been implicated in both nuclear and cytoplasmic mechanisms for the control of gene expression at the RNA level (Jelenik et al., 1972; Adesnik et al., 1973; Darnell et al., 1973; Perry et al., 1974; Sheiness and Darnell, 1973; Brawerman, 1973). The possible cytoplasmic role of the globin mRNA poly(A) sequence has first been investigated.

4.4 Translation of mouse reticulocyte globin mRNA from which the poly(A) sequence has been removed

a) Demonstration of removal of poly(A)

Optimisation of the conditions of polynucleotide phosphorylase digestion was important for several reasons:- (1) In order to minimise endonucleitic attack of 9S mRNA, it was desirable for the incubation time to be short, and the amount of protein added to be small. (2) It was also thought possible that a high salt concentration might result in artefacts of degradation due to secondary structure of the RNA.

The isolation of the treated mRNA used in these experiments depends upon its nonbinding to oligo(dT)-cellulose. This material migrates at a maximum molecular weight of 200,000 and an average molecular weight of 185,000 (Figure 16b), at least 20,000 below the molecular weight of intact mouse globin mRNA.

The uncertainty as to the exact molecular weight of mouse globin mRNA does not affect the calculation of the difference in molecular weights, which depends upon the relative mobility of the intact and treated 9S RNA.

Only small amounts of absorbing material migrate other than in the major component of mol. wt. 185,000 after phosphorylase treatment (Figure 16b). At least 80% of the RNA is in this peak, and there is little if any material migrating in the position of intact 9S mRNA.

It can be seen from the stained gels (Figure 17) that removal of the poly(A) sequence modifies the behaviour of 9S RNA, with a considerable reduction in the heterogeneity of migration of the RNA. With this sharp peak of molecular weight 185,000, a faster moving species of average molecular weight 175,000 is frequently seen (Figures 16b and 17). The relative proportion of this smaller molecule increases with the length of the digestion period. Since neither molecule is retained by oligo(dT) cellulose, this smaller RNA must represent a second distinct digestion product.

The poly(A) sequence of mouse globin mRNA is approximately 50–70 residues in length (Lim and Canellakis, 1970; Morrison et al., 1973; Mansbridge et al., 1974) and therefore the molecular weight reduction is calculated to be sufficient to remove all of the poly(A) sequence. It is of interest that after fractionation of the treated mRNA the bound material still migrates at a position identical with untreated mRNA. This may be due to a total lack of degradation by the enzyme, or to a secondary structure which is dependent for its maintenance on the existence of at least a short poly(A) sequence. When the entire poly(A) sequence is removed, the secondary structure and migration properties may change considerably. Further

analysis of the poly(A) sequence and its interaction with the rest of the mRNA molecule will be necessary to resolve this point.

The phosphokinase fingerprint of the pancreatic RNase T_1 digest (Figure 19), demonstrates the absence of at least 90% of the oligo(A) sequences derived from the poly(A) tract in intact globin mRNA. Therefore it can be concluded from the lack of binding to oligo(dT)–cellulose, the reduction in molecular weight, the lack of template activity with reverse transcriptase and the absence of oligo(A) isopliths, that the treated preparation is essentially free of poly(A) sequences on the 3' end.

Since human globin mRNA, which migrates to the same position as mouse globin mRNA on gels, is known to contain long untranslated nucloetide sequences between the normal termination codon and the poly(A) sequence of both α– and β–globin mRNAs (Clegg et al., 1971; Flatz et al., 1971), similar sequence may occur in other mammals. The reduction in size obtained after enzyme treatment would not be sufficient to affect the globin coding sequence of minimum mol. wt. 143,000 (α chain) and 148,000 (β chain).

Polynucleotide phosphorylase attacks RNA from the 3' end and is particularly active in digesting homopolymers, including poly(A) (Grunberg–Manago, 1963).

With 9S mRNA, digestion by the enzyme does not proceed rapidly beyond the first hundred nucleotides or so. This may be because of the rapid rate of digestion of homopolymer as compared with RNA, or due to secondary structure of the RNA molecule. Globin mRNA is known to have a high hyper–chromicity, indicating considerable secondary structure (Williamson et al., 1971; Lingrel et al., 1971).

b) Translation of deadenylated mRNA in cell-free systems

It can be seen from Table II that in the Krebs ascites cell-free system, the incorporation into mouse α and β globin directed by the treated, non-bound mRNA is about 50% that of the control RNA. A similar observation is made for β globin chain synthesis in the rabbit reticulocyte system (Table III). This value is similar after 6 minutes and 90 minutes incubation. The amount of mouse globin synthesis obtained with added mouse mRNA is that expected for the rabbit lysate system, as Palmiter has previously shown a transit time of approximately three minutes for translation of a globin polypeptide chain.

The somewhat lower globin synthesis directed by the deadenylated as compared to the control mRNA may reflect an inherent effect due to the removal of the poly(A) sequence, but it is more likely to reflect contamination of the deadenylated mRNA by small fragments of RNA broken by contaminating endonucleases, which contribute to the optical density of the sample but cannot contribute to polypeptide synthesis.

This incorporation seen is still very much more than could be accounted for by contamination with intact globin mRNA. The fingerprint data demonstrate that the treated mRNA contains less than 10% of the oligo(A) fragments derived from poly(A) in intact globin mRNA. This small amount of material which could be derived from poly(A) sequences in the mRNA is very short in fragment length and shows a different distribution from that obtained with untreated mRNA, making it unlikely that it is obtained from a proportion of the intact mRNA molecules.

The incorporation experiments were carried out in a response range where amino acid incorporation is linear with added mRNA. Even 10% contamination with intact globin mRNA would at most account for 20% of the observed incorporation into globin chains.

The Krebs cell-free protein synthesis system used to assay for globin synthesis in these experiments is much less efficient than *in vivo* protein synthesis and it is unlikely that each exogenous mRNA molecule is translated more than once (Mathews *et al.*, 1972).

In Figure 20, the dependence of incorporation upon the concentration of added globin mRNA is shown for the Krebs system. The initial slopes of the lines are the same, demonstrating that the messenger activity of the treated and untreated mRNA is the same. If removal of the poly(A) sequence had reduced the ability of the RNA preparation to initiate protein synthesis, the slope of this line for the treated RNA would be lower.

There is, however, a puzzling discrepancy between the fact that CCl_3COOH-precipitable counts for treated and control mRNAs are equal, but globin chain synthesis is less for the treated messenger (Figure 21). This may be due to an increased number of incomplete polypeptide chains; analysis of the "break-through" peak of the CM-cellulose column showed more counts for the treated than control mRNA, which increase the measured incorporation to an approximately equal level (Table II). However, although this peak contains polypeptide material (Mathews and Osborn, 1971), which may represent incomplete globin chains, it is not clear if it is globin mRNA directed, and further work is required to elucidate the reason for this difference.

The difference in ratio of α- to β-globin chains for the treated mRNA may be due to the characteristics of the cell-free system, in which a factor required for the specific synthesis of α-globin has been demonstrated (Wiggleand Smith, 1972).

phosphorylase digestion, possibly because of its secondary structure. Gorski et al. (1974) have demonstrated that although the poly(A) sequence of mouse globin mRNA is heterogeneous, the size range of the sequence is similar on both α and β messages. The translational ability of the 175,000 daltons band was not tested separately, but its molecular weight makes it unlikely that it lacks the codons for the terminal region of the globin protein.

These experiments thus demonstrate that globin chain synthesis is primed in a heterologous cell-free system by deadenylated globin mRNA. They also show that the efficiency of reinitiation of this mRNA is not significantly lower than control mRNA. This indicates that the poly(A) sequence, is not involved directly in the processes of translation, namely, initiation, elongation, termination or reinitiation. These results are in agreement with those of Bard et al. (1974).

Recently however, Huez et al. (1974) have studied the fate of deadenylated rabbit globin mRNA during translation in Xenopus oocytes. Although initially this mRNA is translated efficiently, after a few hours the rate of haemoglobin synthesis decreases considerably, as compared with the synthesis by control, adenylated globin mRNA. It thus appears that although the poly(A) sequence is not necessary for mRNA translation, it is required to sustain high functional stability and repeated translation.

It is not yet known how the poly(A) sequence may function to increase mRNA stability. It has been suggested that poly(A) secondary structure may reduce the effect of exonucleitic degradation of the coding sequences. It may also serve as a site of attachment for a protecting protein (Kwan and Brawerman, 1972; Bloebel, 1973). The very low stability of the only known eukaryotic mRNAs to lack poly(A) — the histone messengers — suggests such a role for the poly(A) sequence.

4.5 Quantitation of poly(U) sepharose

For interpretation of the results discussed in later sections, the ability of the poly(U) sepharose colums to retain all poly(A) containing RNA was crucial. The experiments carried out to test this can be summarised as follows.

a) About 80% of the globin RNA sequences present in reticulocyte polysomal RNA are retained by the column on the first cycle. The length of the poly(A) sequence on retained molecules in 50 – 70 nucleotides (Mansbridge et al., 1974). In subsequent passages through poly(U) sepharose, small amounts of globin RNA are additionally retained by the column (Table IV). After four passages through poly(U) sepharose 90% of the globin RNA sequences have been retained.

b) More than 97% of globin RNA sequences that have been once retained on poly(U) sepharose, are retained when rechromatographed on the column (Table V).

c) More than 98% of globin RNA sequences made from 14S mRNP are retained on poly(U) sepharose (Table VI).

d) 22% of labelled nuclear RNA from growing cells is retained by poly(U) sepharose (Table VIII). This agrees well with published figures of Jelenik et al.(1972) and others. 0.82% of this RNA is resistant to digestion by RNase, and migrates as several heterogeneous peaks in polyacrylamide gels. 0.02% of the RNA that is not retained by poly(U) sepharose is RNase resistant (Table VIII). This material is approximately 10 nucleotides long (Figure 25). It thus appears that in the loading buffer used, the secondary structure of the nuclear RNA does not prevent the retention of poly(A) containing RNA by poly(U) sepharose.

e) In conditions where more than 97% of oligo $A_{(10)}$ is retained by poly(U) sepharose, less than 2% of oligo $A_{(7)}$ is retained (Table VII). This shows that the minimum stable length of hybrid

between poly(U) and poly r(A), in 0.1M NaCl at 21°C is about 10 nucleotides. This is in agreement with published figures (Lipsett, 1964).

From the values given in Tables IV and V it can be calculated that a 1 ml column of poly(U) sepharose will retain at least 50 – 70 μg of globin mRNA. This value is in agreement with published data (Lindberg and Persson, 1972).

The results summarised above demonstate the validity of using poly(U) sepharose as a means of isolating the vast majority of the poly(A) containing RNA molecules from both nuclear and cytoplasmic RNAs.

4.6 Globin RNA metabolism in the mouse foetal liver

The 14 day mouse foetal liver was chosen to study other possible mechanisms of post-transcriptional control at the RNA level. A large proportion of the cells of this tissue are actively synthesising α and β globin in the cytoplasm (Kovach et al., 1967; Cole et al., 1968). It has also been demonstrated that the HnRNA of these cells contain globin RNA sequences (Williamson et al., 1973).

a) Detection of globin mRNA in the 14 day foetal liver cytoplasm by translational studies

In several of the early attempts to make cytoplasmic RNA from foetal livers, considerable degradation of the RNA occurred. The modified method described circumvents these problems, as can be seen by the polysome profile obtained (Figure 26) and the bands observed in gel analysis (Figure 27).

A large proportion of the foetal liver RNA that is retained by poly(U) sepharose runs as a heterogeneous peak between 18S and 5S, with a peak at about 9S, by comparison with a standard RNA preparation (Figure 27). This retained material contains the messenger RNA for mouse α and β globin,

as seen by its ability to prime for α and β globin chain synthesis in the duck reticulocyte cell-free system (Figure 28). These translational studies indicate that the amount of globin mRNA in the total polysomes of 14 day foetal livers, is about one tenth of the amount found in anaemic mouse reticulocytes.

b) Detection of globin RNA sequences in the cytoplasm of 14 day foetal livers using cDNA

The amount of globin RNA sequences in foetal livers can most easily be quantitated using a cDNA probe. The cDNA forms a stable hybrid with the sequences present in foetal liver cytoplasmic RNA, and the hybrid melts 2.2° below the T_m of a cDNA-9S RNA hybrid (Figure 31). A similar decrease in the T_m of a foetal liver cytoplasmic RNA-cDNA hybrid has been noted by other workers in this laboratory (Harrison et al., 1974c). A possible explanation for this decrease would be the hybridisation of cDNA to RNA sequences coding for other globin chains. However, the synthesis of non-adult haemoglobins has not been observed in foetal livers (Kovach et al., 1967; Cole et al., 1968). After treatment of the cDNA-foetal liver RNA hybrid with S_1 nuclease, the size of the cDNA is only reduced 3 fold (Figure 32), indicating a maximum of 2 regions of mismatching.

These observations make it highly unlikely that hybridisation of cDNA to sequences other than α and β globin mRNAs is occurring.

Titration of 14 day foetal liver polysomal RNA with globin cDNA demonstrates that about 0.48% of the RNA is complementary to globin cDNA (Table IX). This is about one quarter of the value observed for reticulocyte polysomes.

80% of the foetal liver polysome-associated globin mRNA sequences are retained by poly(U) sepharose. This value is similar to those obtained by Adesnik et al. (1972) in HeLa cells, and Greenberg and Perry (1972) in L cells. Since most mRNAs are thought to be polyadenylated, these low values may result from the action of nucleases during RNA extraction, as has been observed in the removal of the poly(A) sequence from globin mRNA extracted from anaemic mouse spleen (Cheng et al., 1974). However, a similar value of 80% retention has been obtained using the polysomes of anaemic mouse reticulocytes (Table IV) which are known to be low in nucleases (Mathias et al., 1964). This makes this explanation unlikely.

It has been suggested that a decrease in the length of the poly(A) sequence of cytoplasmic mRNAs may be part of the messenger ageing processes of the cell (Sheiness and Darnell, 1973). This would result in a considerable heterogeneity in the length of the messenger poly(A) sequences especially in long-lived messengers such as globin mRNA. Since oligo(A) sequences shorter than $(A)_{10}$ will not be retained by poly(U) sepharose (Table VII) this ageing process would result in the non-retention of a proportion of the globin mRNA present in both reticulocyte and foetal liver polysomes.

The message that is no longer being translated will be present in the cytoplasm of the cell until it is degraded. These sequences will represent a class of cytosol mRNAs that will not be retained by poly(U) sepharose. Other mRNAs free in the cytoplasm will be newly synthesised molecules that have just exited from the nucleus. These messages will all contain a long poly(A) sequence, and will be retained by poly(U) sepharose.

This scheme of mRNA turnover in the cytosol and polysomes is in agreement with the observation that a lower proportion (70%) of the globin RNA sequence in the total cytoplasmic RNA than in polysomal RNA, are retained by poly(U) sepharose (Table X).

The level of globin RNA sequences in the cytoplasm of 14 day foetal livers in about 0.3% (Figure 30a). However, since erythroid cells contain more RNA than reticulocytes (Table XVII) this value can be calculated to represent about 30% of the number of globin mRNA molecules found in reticulocytes (Table XVII).

At this stage of foetal development the liver comprises of up to 70% erythroid cells, made up of both stem cells, and recognisable erythroid cell types, from immature proerythroblasts to mature erythrocytes (Paul et al., 1969). In situ hybridisation of globin cDNA to these cells has demonstrated the presence of globin RNA sequences in the cytoplasm of all erythroid cells, with the exception of the proerythroblasts (Harrison et al., 1974c). The total percentage of cells in the foetal liver that do not contain cytoplasmic globin mRNA can be estimated to be about 40% (data from Harrison et al., 1974c). This will increase the relative proportion of globin mRNA in the cytoplasm of the erythropoietic cells from 0.3% to 0.5%. From this it can be estimated that the number of globin mRNA molecules in a foetal liver erythroid cell is about 75,000. This value is about half the calculated number of messages present in the reticulocyte (Table XVII).

c) cDNA estimation of globin RNA sequences in the nucleus of foetal liver

Globin RNA sequences have been identified in the nuclear RNA of 14 day foetal livers by hybridisation with globin cDNA. By a comparison of the melting behaviour of the hybrids, and the size of the cDNA isolated from the S_1 nuclease-treated hybrids, these RNA sequences are indistinguishable from the α and β globin mRNAs found in the adult reticulocyte (Table XV). About 0.15% of 14 day foetal liver nuclear RNA is complementary to cDNA, by comparison with a titration curve of cDNA-9S RNA (Figure 30a). This value is about 10 times higher than is found in the nuclear RNA of DMSO induced Friend erythro-leukaemic cells (Harrison et al., 1974b; Gilmour et al., 1974), but is similar to the estimate obtained by cDNA hybridisation to adult human bone marrow nuclear RNA (Lanyon et al., 1974).

When mouse foetal liver nuclear RNA is chromatographed on poly(U) sepharose, about 50% of the globin RNA sequences are retained by the column (Table XI). There are several ways in which nuclear processes might combine to generate this value.

1) Cleavage of RNA during preparation

It is difficult to prevent completely nuclease action during the preparation of nuclear RNA. This may cause random RNase breaks, following distruption of the cells by homo-genisation. If these nicks occur between the globin RNA sequences, and the poly(A) sequence, the amount of globin RNA retained by poly(U) sepharose will be reduced. However, it is thought that the protein coding sequences are usually found on the 3' end of the HnRNA molecule, next to the poly(A) sequence (Edmonds et al., 1971; Darnell et al., 1971; Jelenik et al., 1973). Because of this proximity, the

probability of a random cleavage occurring between these sequences is low.

However, it is possible that there is more than one globin-coding sequence in each polyadenylated HnRNA molecule. Cleavage of this molecule during processing will release 5' globin RNA sequences from association with poly(A). Although these sequences may be rapidly readenylated or degraded, the occurrence of either processing, or random nuclease breaks during preparation of the nuclear RNA, will increase the proportion of non-retained globin RNA sequences. However, recent data on the transcription of the linked human beta and delta globin genes (Lanyon et al., 1974) and the low reiteration frequency of the mouse globin genes makes this explanation unlikely.

Apart from these possibilities of artifactual RNA cleavage during preparation, there are two major alternatives of nuclear RNA metabolism that will effect the value for the retention of globin RNA sequences by poly(U) sepharose.

2) Time of polyadenylation

The addition of poly(A) to nuclear RNA molecules is a post-transcriptional event (Philipson et al., 1971; Jelenik et al., 1973). Estimates of the time that elapses between the end of transcription and the addition of poly(A) vary between a few minutes (Jelenik et al., 1973) in HeLa cells, to 13 minutes (Perry et al., 1974) in L cells. This last figure would indicate that polyadenylation is a relatively late post-transcriptional event, occurring just prior to the appearance of the mRNA in the cytoplasm, and would be consistent with the presence in the foetal liver nucleus of a large proportion of non-adenylated globin RNA sequences.

However, the value obtained from L cells depends critically on the cell pool sizes for poly(A) precursors, and some doubt has recently been expressed about the accuracy of this approach (Darnell 1974). Although some of the non-retained globin RNA sequences will certainly represent HnRNA molecules that have not yet been adenylated, other explanations for the non-retention of these sequences must be considered.

3) Other mechanisms of poly(A) metabolism.

It has long been known that HnRNA has a short half life, with the majority of the RNA being broken down in the nucleus. However, a much higher proportion of the nuclear poly(A) appears to be conserved (Jelenik et al., 1973). These authors suggest that the addition of poly(A) to nuclear RNA sequences may serve as a 'tag' for their conservation, and processing to cytoplasmic mRNAs. This pathway is thought to be obligatory for the majority of mRNAs, since most mRNAs contain poly(A). Also, the prevention of poly(A) addition by cordycepin blocks the appearance of mRNAs in the cytoplasm (Adesnik et al., 1972; Jelenik et al., 1973).

There is however considerable doubt as to whether all the nuclear poly(A) is conserved during processing. Values for the fraction of nuclear poly(A) that enters the cytoplasm over long chase periods, varies between 80% (Jelenik et al., 1973) for HeLa cells, to 25% (Perry et al., 1974) for L cells. Turnover of poly(A) sequences on nuclear globin RNA would reduce the amount of globin RNA retained by poly(U) sepharose. Also a pool of globin RNA sequences may exist which are destined never to become associated with poly(A) and are turned over in the nucleus.

Both these possibilities would indicate that a distinct regulatory mechanism exists in the processing of HnRNA molecules, independent of poly(A) addition. In this, only a proportion of the nuclear globin RNA sequences are selected for transport to the cytoplasm and the rest would turn over in the nucleus. Over production of the required RNA sequences is a necessary prerequisite for post-transcriptional modulation of the levels of these sequences in the nucleus or cytoplasm.

At the present time it is not possible to distinguish between these explanations. However, some clarification can be obtained by a study of the fate of globin RNA sequences in the nucleus and cytoplasm of other, non-erythropoietic tissues.

4.7 Presence of globin RNA sequences in the RNA of non-erythropoietic tissues

Globin RNA sequences have been detected in the nuclear and cytoplasmic RNA of adult mouse brain and liver, and a mouse lymphoma cell line L5178Y (LY). The fidelity of the hybrids between globin cDNA, and nuclear RNA from these tissues has been demonstrated by their melting behaviour (Figure 31) and the size of the cDNA after S_1 treatment of the hybrid (Figure 32). By comparison with a cDNA-9S RNA hybrid, these RNAs contain sequences coding for mouse α and β globin.

The levels of globin RNA sequences detected in the nuclear RNA of the non-erythropoietic cells is between 20-90 fold less than that found in the foetal liver, when the calculated number of molecules of globin RNA in each is compared (Table XVII). These values will reflect both the relative availability of the globin genes to transcription, and the rate at which these transcribed sequences are processed to mRNA and exported to the cytoplasm, or are degraded in the nucleus.

Gilmour and Paul (1973) and Gilmour *et al*. (1974) have studied the *in vitro* transcription of chromatin prepared from adult mouse brain, foetal livers, and L Y cells. They have demonstrated a clear difference in the transcription of the globin genes in these tissues, with high levels of globin sequences being only detected in the chromatin primed RNA from foetal livers. Differences in the stability of these nuclear RNA sequences are difficult to assess. However, the data in Table XIII demonstrate that these differences are not related to the association of the sequences with poly(A), since the proportion of the nuclear globin RNA sequences that is retained by poly(U) sepharose is similar in brain, and adult and foetal liver.

The amounts of globin RNA detected in the cytoplasm of these tissues is 200 fold lower for brain and adult liver, and 4000 fold lower for L Y cells than is found in the foetal liver cytoplasm (Table XII). Although the majority of these sequences from brain and foetal liver cytoplasm are retained by poly(U) sepharose, the value is about 2 fold lower for adult liver (Table XIII). The reason for this difference is unknown.

An estimate of the effect of reticulocyte contamination, on the levels of globin RNA sequences detected in the nuclear and cytoplasmic preparations of adult brain and liver can be made. It was assumed that blood cells may represent a maximum of 10% of the cells in a tissue. If 2% of these cells are reticulocytes, with each reticulocyte containing 140,000 globin messenger RNA molecules (Mathias *et al*., 1964 and Table XVII), it can be calculated that the cytoplasmic preparation of an adult mouse tissue could be contaminated with 280 reticulocyte mRNA molecules per cell, and the nuclear preparation with about 0.2% of this = 0.5 molecules per nucleus.

Thus, although the levels of globin RNA detected in the cytoplasm of adult tissues may be explained by reticulocyte contamination, the levels detected in the nuclear RNA could not come from this source. In agreement with this, the amount of globin RNA sequences detected in the cytoplasm of L Y cells is very low, but the nuclear levels are similar to those observed in the adult brain and liver.

4.8 The nuclear post-transcriptional regulation of globin gene expression

The operation of several distinct post-transcriptional control mechanisms can be inferred from the data in Tables XIII and XVII. It can be seen that the amount of nuclear globin RNA does not in itself determine the amount of globin messenger RNA in the cytoplasm (Table XVII). This implies that the levels of globin RNA sequences in the two compartments, though related in a precursor - product manner, are under independent regulatory mechanisms. These mechanisms may operate through the nuclear processing of the sequences, or at the level of cytoplasmic stability, or both.

Messenger stability might be altered through factors that are required for the translation of the specific mRNA, with the free message being rapidly degraded. Harrison *et al.* (1974b) and Gilmour *et al.*, (1974) have recently suggested that such a mechanism may explain the observed levels of globin mRNA in a non-erythroid cell line.

The way in which nuclear control processes may act is unknown but the data (Table XIII) demonstrate that it is unlikely to be mediated through the post-transcriptional association of globin sequences with poly(A). Thus, although the relative amounts of nuclear globin RNA that appear in the cytoplasm are very different in erythroid and non-erythroid

cells, the proportion of nuclear globin sequences associated with poly(A) in these tissues is the same (Table XIII). This strongly indicates that although the addition of poly(A) to a potential mRNA sequence in the nucleus may be obligatory for its processing to the cytoplasm, polyadenylation is not a sufficient event for this processing to occur.

It is difficult to explain the high levels of globin RNA sequences detected in brain, adult liver and L Y nuclei, since these are not erythropoietic tissues. If this amount of background transcription of all genes occurred, the sequences found in nuclear RNA would be similar in all tissues, and the complexity of the RNA would approach that of the mouse genome. Although the complexity of nuclear RNA from any tissue is known to be high ($10^7 - 10^8$ nucleotides) (Hahn and Laird, 1971; Holmes and Bonner, 1973; Getz et al., 1974), this ~~is transcribed~~ represents transcription of only 1–5% of the tissues' unique DNA. Also, several studies have demonstrated clear differences between the sequences present in the nuclear RNA of different tissues (Brown and Church, 1972; Church and Brown, 1972; Liarakos, et al., 1973).

In mouse M2 cells, Getz et al. (1974) have described 2 classes of poly(A) containing nuclear RNA. About 25% of nuclear RNA appears to be transcribed from about 0.06% of the genome, with about 300 RNA copies of each gene sequence being present per nucleus. The majority of the nuclear RNA has a much higher complexity (being transcribed from about 3% of the genome), but is only present at a level of about 5 copies per nucleus. The levels of globin RNA sequences detectable in non–erythroid cells are thus similar to that of the high frequency class of nuclear RNAs. Both classes of nuclear RNAs can be detected in M2 cell polysome associated mRNA (Birnie et al., 1974) though in a different ratio from their nuclear levels (Getz et al., 1974).

The possibility must therefore be considered that either the production of globin RNA, or the production of globin protein, may be necessary for the functioning of the cell. Perhaps low levels of haemoglobin are required in the cell cytoplasm. Several examples are known where proteins that occur in high amounts in certain differentiated cell types, such as actin, myosin and tropomyosin, have been detectable in many other cell types (Ishikawa et al., 1969; Adelstein et al., 1972; Cohen and Cohen, 1972).

However, the possibility cannot be discounted that the globin RNA sequences themselves are involved in a nuclear function. Another less likely possibility is that the globin genes are linked to genes that are expressed in all tissues, with the globin gene sequence being cotranscribed, or having a high probability of being transcribed.

Although at the present time, the reasons for the expression of the globin genes in non-erythroid tissues are unclear, the system lends itself well to further study. Denaturing formamide gradients can be used to isolate RNA molecules of different sizes, and coupled with the extreme sensitivity of the cDNA assay for globin RNA sequences, it should be possible to follow the nuclear processing of globin RNA sequences in both erythroid and non-erythroid cells. A comparison of these processes in the two classes of tissues should help to elucidate the function of the non-erythroid globin RNA sequences. It may also indicate the location of the metabolic blocks that appear to occur in some human diseases such as β^+ thalassaemia, where the β genes are present (Ottolenghi, Williamson and Weatherall, unpublished) and may be transcribed, but where β globin mRNAs do not appear in the cytoplasm.

ABBREVIATIONS

RNA	Ribonucleic acid
DNA	Deoxyribonucleic acid
mRNA	Messenger RNA
tRNA	Transfer RNA
rRNA	Ribosomal RNA
mRNP	Messenger ribonucleoprotein particle
cDNA	Complementary DNA
A, C, G, U, T	The bases adenine, cytosine, guanine, uracil and thyamine respectively.
EDTA	Ethylenediamine tetra-acetic acid
Tris	2-amino-2-(Hydroxymethyl)-propane-1:3-diol.
ANE	0.1 M sodium acetate; 0.1 M NaCl; 10 mM EDTA
TKM	5 mM Tris (pH 7.5); 50 mM KCl; 1.5 mM $MgCl_2$
NETS	0.1 M NaCl; 10 mM Tris (pH 7.5); 1 mM EDTA; 0.5% SLS
H.M.C.	0.1 M HEPES (pH 7.0); 10 mM magnesium acetate; 25 mM NaCl; 2 mM $CaCl_2$
HEPES	N-2-Hydroxyethyl-piperazine-N'-2-ethanesulphonic acid
PIPES	Piperazine-N-N'-bis (2-ethanesulphonic acid)
C.M.C.	Carboxymethyl cellulose
DEC	Diethyl-aminoethyl cellulose
BSS	Hank's balanced salts solution (Paul 1965)
PCA	Perchloric acid
TCA	Trichloroacetic acid
TBS	Toluene-based scintillant
TXBS	Toluene-based scintillant, plus Triton X 100.
DMSO	Dimethyl sulphoxide
SLS	Sodium lauryl sulphate
DTT	Dithiothreitol

DNase	Deoxyribonuclease
RNase	Ribonuclease
A_{260}	Absorbance 260 nm
DEP	Diethyl pyrocarbonate
Tm	Mean thermal denaturation
Co	Concentration of DNA in moles nucleotides/l, in reannealing experiments
Do	Concentration of cDNA in moles nucleotides/l in cDNA–RNA hybridisation experiments
$t_{\frac{1}{2}}$	time (seconds) for 50% of the reaction to take place

BIBLIOGRAPHY

Acheson, N.H., Buetti, E., Scherrer, K. and Weil, R. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 2231.

Adelstein, R.S., Conti, M.A., Johnstone, G.S., Pastan, I. and Pollard, T.D. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 3693.

Adesnik, M. and Darnell, J.C. (1972). J. Mol. Biol. 67, 397.

Adesnik, M., Saldilt, M., Thomas, W. and Darnell, J. (1972). J. Mol. Biol. 71, 21.

Adler, K., Beyruther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Muller-Hill, B., Pfahl, M. and Schmitz, A. (1972). Nature 237, 322.

Albrecht, C. and Zyl, I.M. van. (1973). Exptl. Cell Res. 76, 8.

Amaldi, F., Lava-Sanchez, P.A. and Buongiorno-Nardelli, M. (1973). Nature 242, 615.

Anderson, N.G. (1966). Science, 154, 103.

Arnstein, H.R.V., Cox, R.A and Hunt, J.A. (1962). Nature, 194, 1042.

Attardi, G. and Amaldi, F. (1970). Ann. Rev. Biochem. 39, 216.

Attardi, G., Parnas, H., Hwang, M.-I.H. and Attardi, B. (1966). J. Mol. Biol. 20, 145.

Aviv, H. and Leder, P. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 1408.

Axel, R., Cedar, H. and Felsenfeld, G. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2029.

Baltimore, D. and Huang, A.S. (1970). J. Mol. Biol. 47, 263.

Bard, G., Efron, D., Marcus, A. and Perry, R.P. (1974). Cell 1, 101.

Bautz, E.F.K. and Bautz, F.A. (1964). Proc. Nat. Acad. Sci. U.S.A. 52, 1476.

Beckworth, J.R. and Zipser, D. (eds.) (1970). "The Lactose Operon". (Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.).

Beerman, W. (1956). Cold Spring Harbour Symposium 21, 217.

Beerman, W. (1967). In R. A. Brink (ed.) "Heritage from Mondel". 179. (Univ. of Wisconsin Press, Madison).

Beers, R.F. (1960). J. Biol. Chem. 235, 2393.

Birnboim, H.C., Mitchel, R.E.J. and Strauss, N.A. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2189.

Birnie, G.D., Delcour, J., Angus, D., Threlfall, G. and Paul, J. (1973). In Cold Spring Harbour Monograph "Control of Proliferation in Animal cells".

Birnie, G.D. and Harvey, D.R. (1968). Anal. Biochem. 22, 171.

Birnie, G.D., MacPhail, E.,Young, B.D., Getz, M.J. and Paul, J. (1974). Cell Different. 3, in press.

Birnstiel, M.L., Grunstein, M., Speirs, J. and Hennig, W. (1969). Nature, 223, 1265.

Birnstiel, M., Speirs, J., Purdom, I., Jones, K. and Loening, U.E. (1968). Nature, 219, 454.

Birnstiel, M.L., Telford, J., Weinberg, E. and Stafford, D. (1974). Proc. Nat. Acad. Sci. U.S.A. (in press).

Birnsteil, M.L., Wallace, H., Sirlin, J.L. and Fischberg, M. (1966). Nat. Cancer Inst. Monog. 23, 431.

Bishop, J.O. (1972). Biochem. J. 126,171.

Bishop, J.O., Morton, J.G., Rosbash, M. and Richardson, M. (1974). Nature, 250, 199.

Bishop, J.O., Pemberton, R. and Baglioni, C. (1972). Nature New Biol. 235, 231.

Bishop, J.O. and Rosbash, M. (1973). Nature New Biol. 241, 204.

Blobel, G. (1972). Biochem. Biophys. Res. Commun. 47, 88.

Blobel, G. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 924.

Boedtker, H. (1968). J. Mol. Biol. 35, 61.

Bolton, E.T. and McCarthy, B.J. (1964). J. Mol. Biol. 8, 201.

Borenfreund, E., Fitt, E. and Bendich, A. (1961). Nature, 191, 1375.

Borun, T.W., Scharff, M.D. and Robbins, E. (1967). Biochim. Biophys. Acta, 302.

Bramwell, M.E. (1972). Biochim. Biophys. Acta. 281, 329.

Brandhorst, B.P. and McConkey, E.M. (1974). J. Mol. Biol. 85, 451.

Brawerman, G. (1973). Mol. Biol. Rep. 1, 7.

Brawerman, G., Mendecki, J. and Lee, S.Y. (1972). Biochemistry, 11, 637.

Brenner, S. and Milstein, C. (1969). Nature, 211, 242.

Bridges, C.B. (1935). J. Hered. 26, 60.

Britten, R.J. and Bonner, T. (1971). "Yearb. Carneg. Instn." 70, 373.

Britten, R.J. and Davidson, E.H. (1969). Science, 165, 349.

Britten, R.J. and Davidson, E.H. (1971). Quart. Rev. Biol. 46, 111.

Britten, R.J. and Kohne, D.E. (1968). Science, 161, 529.

Britten, R.J., Pavich, M. and Smith, J. (1969). "Carnegie Inst. of Washington Yearb. 68, 400.

Brown, D.D. and Dawid, I.B. (1968). Science, 160, 272.

Brown, D.D. and Weber, C.S. (1968a). J. Mol. Biol. 34, 661.

Brown, D.D. and Weber, C.S. (1968b). J. Mol. Biol. 34, 681.

Brown, D.D., Wensink, P.C. and Jordan, E. (1971). Proc. Nat. Acad. Sci. 68, 3175.

Brown, D.D., Wensink, P.C. and Jordan, E. (1972). J. Mol. Biol. 63, 57.

Brown, I.R. and Church, R.B. (1972), Develop. Biol. 29, 73.

Burdon, R.H. and Shenkin, A. (1972). FEBS Letters, 24, 11.

Burny, A., Huez. G., Marbaix, G. and Chantrenne, H. (1969). Biochim. Biophys. Acta. 190, 228.

Burny, A. and Marbaix, G. (1965). Biochim. Biophys. Acta. 103, 409.

Burr, M. and Lingrell, J.B. (1971). Nature New Biol. 233, 41.

Burton, K. and Peterson, G.B. (1957). Biochim. Biophys. Acta. 26, 667.

Busch, H. and Smetana, K. (1970). "The Nucleolus". (Academic Press, N.Y.).

Caliguiri, L.A. and Tamm, I. (1970). Virology 42, 112.

Callen, H.G. (1967). J. Cell Sci. 2, 1.

Callen, H.G. and Lloyd, L. (1960). Phil. Trans. Rog. Soc. ser. B. 243, 135.

Chamberlin, M.J. and Patterson, D.L. (1965). J. Mol. Biol. 12, 410.

Chantrenne,H., Burny, A. and Marbaix, G. (1967). Prog. Nuc. Acid Res. and Mol. Biol, 7, 173.

Cheng, T., Polmar, S.K. and Kazajian, H.H. (1974). J. Biol. Chem. 249, 1781.

Church, R.B. and Brown, I.R. (1972). In Results and Problems of Differentiation (ed. H. Ursprung) (Springer-Verley, New York), 3, 11.

Church, R.B. and McCarthy, B.J. (1967). J. Mol. Biol. 23, 477.

Clegg, J.B., Weatherall, D.J. and Milner, P.F. (1971). Nature,234, 337.

Cohen, I. and Cohen, C. (1972). J. Mol. Biol. 68, 383.

Cole, R.J., Hunter, J. and Paul, J. (1968). Brit. J. Haemat. 14, 477.

Crick, F. (1971). Nature 234, 25.

Crossley, J., Mansbridge, J. and Williamson, R. (1974). Biochem. J. 137, 299.

Dacie, J.V. (1956). In "Practical Haematology". (J. and A. Churchill, Ltd., London).

Darnell, J.E. (1974). reported at Biochem. Soc. Meeting (Glasgow).

Darnell, J.E. and Balint, R. (1970). J. Cell. Physiol. 76, 349.

Darnell, J.E., Jelinek, W.R. and Molloy, G.R. (1973). Science 181, 1215.

Darnell, J.E., Philipson, L., Wall, R. and Adesnick, M. (1971b). Science 174, 507.

Darnell, J.E., Wall, R. and Tushinski, R.J. (1971a). Proc. Nat. Acad. Sci. U.S.A. 68, 1321.

Davidson, E.H. (1968). Gene Activity in Early Development, (Academic Press, New York).

Davidson, E.H. and Britten, R.J. (1971). J. Theoret. Biol. 32, 123.

Davidson, E.H. and Britten, R.J. (1973). Quart. Rev. Biol. 48, 565.

Davidson, E.H., Crippa, M. and Mivsky, A.E. (1968). Proc. Nat. Acad. Sci. 60, 152.

Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973). J. Mol. Biol. 77, 1.

Dayhoff, M.O. (1972). (ed.) "Atlas of Protein Sequence and Structure". 5 (National Biomedical Research Foundation).

Dawid, I.B., Brown, D.D. and Reeder, R.H. (1970). J. Mol. Biol. 51, 341.

Denny, P.C. and Tyler, A. (1964). Biochem. Biophys. Res. Commun. 14, 245.

DeLange, R.J., Farnbrough, D.M., Smith, E.L. and Bonner, J. (1968). J. Biol. Chem. 243, 5906.

DeLange, R.J., Farnbrough, D.M., Smith, E.L. and Bonner, J. (1969). J. Biol. Chem. 244, 5669.

Dickson, E., Boyd, J.B. and Laird, C.D. (1971). J. Mol. Biol. 61, 615.

Dintzis, H.M., Borsook, H. and Vinograd, J. (1958). in "Microsomal particles and Protein Synthesis". (Roberts, R.B. Ed.). 95. Pergamon Press, London.

Dreyer, W.J. and Bennet, J.C. (1965). Proc. Nat. Acad. Sci. U.S.A. 54, 864.

Dreyer, W.J., Gray, W.R. and Hood, L. (1967). Cold Spring Harbour Symposium Quant. Biol. 32, 353.

Driesch, H. (1892) Entwicklungsmechanik Anat. Anz. 7, 584.

Dubochet, J., Morel, C., Leblieu, B. and Henzberg, M. (1973). Eur. J. Biochem. 36, 465.

Dunn, J.J. and Studier, F.W. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 3296.

Edmonds, M. and Abrams, R. (1960). J. Biol. Chem. 235, 1142.

Edmonds, M. and Abrams, R. (1963). J. Biol. Chem. 238, 1186.

Edmonds, M., Vaughan, M.M. and Nakazato, H. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 1336.

Faiferman, I., Hamilton, M.G. and Pogo, A.O. (1971). Biochim. Biophys. Acta. 232, 685.

Fan, H. and Penman, S. (1970). J. Mol. Biol. 50, 655.

Favre, A., Bertazzoni, V., Berns, A.J.M. and Bloemendal, H. (1974). Biochem. Biophys. Res. Commun. 56, 273.

Faust, G.H., Diggelmann, H. and Mach, B. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 2491.

Fedorisak, I., Natarajan, A.T. and Ehrenberg, L. (1969). Eur. J. Biochem. 10, 450.

Firtel, R.A., Jacobson, A. and Lodish, H.F. (1972). Nature New Biol. 239, 225.

Flamm, W.G., McCallum, M. and Walker, P.M.B. (1967). Proc. Nat. Acad. Sci. U.S.A. 57, 1729.

Flamm, W.G., Walker, P.M.B. and McCallum, M. (1969). J. Mol. Biol. 40, 423.

Flatz, G., Kinderlever, J., Kilmartin, J.V. and Lehmann, H. (1971). Lancet, 732.

Flickinger, R.A., Coward, S.J., Miyagi, M., Moser, C. and Rollins, F. (1965). Proc. Nat. Acad. Sci. U.S.A. 53, 789.

Ford, P.J. and Southern, E.M. (1973). Nature New Biol. 241, 7.

Gall, J.G. (1956). J. Biochem. Biophys. Cytol. 2 (suppl.) 393.

Gall, J.G. (1968). Proc. Nat. Acad. Sci. U.S.A. 60, 553.

Gall, J.G., Cohen, E.H. and Polam, M.L. (1971). Chromosoma, 33, 319.

Gall, J.G. and Rochaix, J.D. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1819.

Gallau, G., Britten, R. J. and Davidson, E.H. (1974). Cell 2, 9.

Gallwitz, D. and Breindl, M. (1972). Biochem. Biophys. Res. Commun. 47, 1106.

Gallwitz, D. and Meuller, C. (1969). J. Biol. Chem. 244, 5947.

Gally, J.A. and Edelman, G.M. (1970). Nature 227, 341.

Gaskill, P. and Kabot, D. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 72.

Gelderman, A.H., Rake, A.V. and Britten, R.J. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 1.72.

Georgiev, G.P. (1969). J. Theoret. Biol. 25, 473.

Georgiev, G.P. (1972). Curr. Topics Develop. Biol. 7, 1.

Georgiev, G.P., Ryskov, A.P., Coutelle, C., Mantieva, U. L. and Avakyan, E.R. (1972). Biochem. Biophys. Acta. 259, 259.

Georgiev, G.P., Samarina, O.P., Leman, M.I., Smirnov, M.N., and Severitzov, A.N. (1963). Nature, 200, 192.

Getz, M.J., Birnie, G.D., Young, B.D., MacPhail, E. and Paul, J. (1974) Cell, in press.

Gilbert, W. and Maxam, A. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 3581.

Gilbert, W. and Muller-Hill, B. (1967). Proc. Nat. Acad. Sci. U.S.A. 58, 2445.

Gilmour, R.S., Harrison, P.R., Windass, J.D., Affara, N.A. and Paul, J. (1974). Cell Diff. 3, 9.

Gilmour, R.S. and Paul, J. (1969). J. Mol. Biol. 40, 137.

Gilmour, R.S. and Paul, J. (1973). Proc. Nat. Acad. Sci. 70, 3440.

Girard, M. and Baltimore, D. (1966). Proc. Nat. Acad. Sci. U.S.A. 56, 999.

Goldberg, R.B., Galou, G.A., Britten, R.J. and Davidson, E.H. (1973). Proc.Nat. Acad. Sci. U.S.A. 70, 3517.

Gorski, J., Morrison, M.R., Merkel, C.G. and Lingrel, J.B. (1974). J. Mol. Biol.

Graham, D.E., Neufeld, B.R., Davidson, E.H. and Britten, R.J. (1974). Cell. 1, 127.

Granboulan, N. and Scherrer, K. (1969). Europ. J. Biochem. 9, 1.

Greenberg, J. (1972). Nature, 240, 102.

Greenberg, J.R. and Perry, R.P. (1971). J. Cell Biol. 50, 774.

Greenberg, J.R. and Perry, R.P. (1972). J. Mol. Biol. 72, 3.

Gross, F., Hiatt, H., Gilbert, W., Kurland, C.G., Risebrough, R.W. and Watson, J.O. (1961). Nature, 190, 581.

Gross, K., Ruderman, J., Jacobs-Lorena, M., Baglioni, L. and Gross, P.R. (1973). Nature New Biol. 241, 272.

Gruenwedel, D.W., Hsu, C.H. and Lu, D.S. (1971). Biopolymers, 10, 47.

Grunberg Manago, M. (1963). Progr. Nucl. Acid. Res. 1, 93.

Grunstein, M., Levy, S., Schedl, P. and Kedes, L. (1973). Cold Spring Harb. Symp. Quant. Biol. 38,

Gurdon, J.B. (1962). Develop. Biol. 4, 256.

Gurdon, J.B. and Vehlinger, U. (1966). Nature, 210, 1240.

Hahn, W.E. and Laird, C.D. (1971). Science, 173, 158.

Haldane, J.B.S. (1957). Genet. 55,511.

Harris, H. (1962). Biochem. J. 84, 60.

Harrison, P.R., Birnie, G.D., Hell, A., Humphries, S., Young, B.D. and Paul, J. (1974a) J.M.B. 84, 539.

Harrison, P.R., Conkie, D., Affara, N. and Paul, J. (1974c). J. Cell. Biol. in press.

Harrison, P.R., Conkie, D., Paul, J. and Jones, K. (1973). FEBS. Letters, 32, 109.

Harrison, P.R., Gilmour, R.S., Affara, N., Conkie, D. and Paul, J. (1974b). Cell. Diff. 3, 23.

Harrison, P.R., Hell, A., Birnie, G.D. and Paul, J. (1972a). Nature, 239, 219.

Harrison, P.R., Hell, A. and Paul, J. (1972b). FEBS Letters, 24, 73.

Harwood, R., Connelly, A.D., Grant, M. and Jackson, D.S. (1974). FEBS Letters, 41, 85.

Hell, A., Birnie, G.D., Slimming, T.K. and Paul, J. (1972). Anal. Biochem. 48, 369.

Heywood, S.M. (1970). Proc. Nat. Acad. Sci. U.S.A. 67, 1782.

Heywood, S.M., Dowben, R.M. and Rich, A. (1967). Proc. Nat. Acad. Sci. U.S.A. 57, 1002.

Heywood, S.M., Kennedy, D.S. and Bester, A.J. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 2428.

Holley, R.N., Apgar, J., Everett, G.A., Madison, J.T., Marquisee, M., Merrill, S.M., Penswick, R. and Zamir, A. (1965). Science, 147, 1462.

Holmes, D.S. and Bonner, J. (1973). Biochemistry 12, 2330.

Holmes, D.S. and Bonner, J. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1108.

Huang, R.C. and Bonner, J. (1962). Proc. Nat. Acad. Sci. U.S.A. 48, 1216.

Huez, G., Burny, A., Marbaix, G. and Schram, E. (1967). Europ. J. Biochem. 1, 179.

Huez, G., Marbaix, G., Hubert, E., Lecherq, M., Nudel, U., Soreq, H., Salomen, R., Lebleu, B., Revel, M. and Lillauer, U.Z. (1974). Proc. Nat. Acad. Sci. U.S.A. 71

Hunt, J.A. (1973). Biochem. J. 131, 327.

Hunt, L.T. and Dayhoff, M.O. (1972). Biochem. Biophys. Res. Commun. 47, 699.

Hutchison, W.C. and Munro, H.N. (1961). Analyst. 86, 768.

Imaizumi, T., Diggelmann, H. and Scherrer, K. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 1122.

Ishikawa, H., Bischoff, R. and Holtzer, H. (1969). J. Cell. Biol. 43, 312.

Jacob, F. and Monod, J. (1961). J. Mol. Biol. 3. 318.

Jacobs-Lorena, M., Baglioni, C. and Borun, T.W. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 2095.

Jacobson, A., Firtel, R.A. and Lodish, H.F. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1607.

Jeanteur, P., Amaldi, F. and Attardi, G. (1968). J. Mol. Biol. 33, 757.

Jeanteur, P. and Attardi, G. (1969). J. Mol. Biol. 45, 305.

Jelinek, W., Adesnik, M., Salditt, M., Sheiness, D., Wall, R., Molloy, G., Phillipson, L. and Darnell, J.E. (1973). J. Mol. Biol. 75, 515.

Jelinek, W. and Darnell, J.E. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 2537.

Jelinek, W., Molloy, G., Fernandez-Munoz, R., Salditt, M. and Darnell, J.E. (1974). J. Mol. Biol. 82, 361.

deJiminez, E.S., Dominguez, J.L., Webb, F.H. and Bock, R.M. (1971). J. Mol. Biol. 61, 59.

Jones, K.W. (1970). Nature. 225, 912.

Johns, E.W. (1969). Ciba Symposium on Homeostatic Regulators, (edit. by Wolstenholme, G.E.W. and Knight, I.J.) 128. (Churchill, London).

Johnston, R.E. and Bose, H.R. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 1514.

Kacian, D.L., Spiegelman, S., Bank, A., Terada, M., Melafora, S., Dow, L. and Marks, P.A. (1972). Nature New Biol. 235, 167.

Kacian, D.L., Watson, K.F., Burny, A. and Spiegelman, S. (1971). Biochem. Biophys. Acta. 246, 365.

Kafatos, F.C. (1972a). Acta. Endocrinologica. Suppl. 168, 319.

Kafatos, F.C. (1972b). Curr. Topics in Develop. Biol. 7, 125.

Kedes, L.H. and Birnstiel, M.L. (1971). Nature New Biol. 230, 165.

Kedes, L.H. and Gross, P.R. (1960). J. Mol.Biol. 42, 559.

Kedes, L.H. and Gross, P.R. (1969). Nature 223, 1335.

Kedes, L.H., Gross, P.R., Cognetti, G. and Hunter, A.L. (1969). J. Mol. Biol. 45, 337.

Klein, W.H., Murphy, W., Attardi, G., Britten, R.J. and Davidson, E.H. (1974). Proc. Nat. Acad.Sci. U.S.A. 71, 1785.

Kohne, D.E. (1969). Biophys. J. 8, 1104.

Kovach, J.S., Marks, P.A., Russel, E.S. and Elper, H. (1967). J. Mol. Biol. 25, 131.

Kruk, J. and Borsook, H. (1956). J. Biol. Chem. 220, 905.

Kwan, S. and Brawerman. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 3247.

Labrie, F. (1969). Nature, 221, 1217.

Laird, C.D.,McConaughty, B.L. and McCarthy, B.J. (1969). Nature, 224, 149.

Lambert, B. (1973). Nature, 242, 51.

Lane, C.D., Marbaix, G. and Gurdon, J.B. (1971). J. Mol. Biol. 61, 73.

Lanyon, G., Paul, J. and Williamson, R. (1972). Eur. J. Biochem. 31, 38.

Lanyon, G., Ottolenghi, S. and Williamson, W. (1974). Proc. Nat. Acad. Sci. U.S.A. in press.

La Torre, J. and Perry, R.P. (1973). Biochim. Biophys. Acta. 335, 93.

Lee, S.Y., Mendecki, J. and Brawerman, G. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 1331.

Lee-Huang, S. and Ochoa,S. (1971). Nature New Biol. 234, 236.

Levinthal, C., Keynan, A. and Higa, A. (1962). Proc. Nat. Acad. Sci. U.S.A. 48, 1631.

Liarakos, G.D., Rosen, J.M. and O'Malley, B.W. (1973). Biochemistry, 12, 2809.

Lim, L. and Canellakis, E.S. (1970). Nature, 227, 710.

Lindberg, U. and Persson, T. (1972). Eur. J. Biochem. 31, 246.

Lingrell, J.B., Lockard, R.E., Jones, R.F., Burr, H.E. and Holder, J. W. (1971). in "Series Haematologica, Vol. 4. (Jenson, K.G., and Killman, S., Ed., Munksgaard, Copenhagen). 37.

Lipsett, M.N. (1964). J. Biol.Chem. 239, 1256.

Lissitzky, S., Poiree, J. C., Carttouzou, G. and Gregoire, J. (1970).Eur. J. Biochem.12, 104.

Loening, U.E. (1967). Biochem. J. 102, 251.

Loening, U.E. (1969). Biochem. J. 113, 131.

London, J.M., Shemin, D. and Rittœnberg, D. (1950). J. Biol. Chem. 183, 749.

MacGregor, H.C. (1968). J. Cell. Sci. 3, 437.

MacNaughton, M., Freeman, K.B. and Bishop, J.O. (1974). Cell, 1, 117.

Maden, B.E.H. (1971). Progr. Biophys. Mol. Biol. 22, 129.

Maizles, N. (1974). Nature, 249, 647.

Mansbridge, J.N., Crossley, J.A., Lanyon, W.G. and Williamson, R. (1974). Eur. J. Biochem. 44, 261.

Marbaix, G., Burny, A., Huez, G. and Chantrenne, H. (1966). Biochem. Biophys. Acta. 114, 404.

Marcus, L., Bretthauer, R.K., Bock, R.M. and Halvorson, H.O. (1963). Proc. Nat. Acad. Sci.. 50, 782.

Markham, R. (1957). Methods Enzymol. 3, 805.

Marmur, J. and Doty, P. (1962). J.. Mol. Biol. 5, 109.

Marmur, J. and Dpty, P. (1961). J.. Mol. Biol. 3, 585.

Marmur, J., Rownd, R. and Schildkraut, C.L. (1963). Progr. Nucleic Acid Res. 1, 231.

Marushige, K. and Bonner, J. (1966). J. Mol. Biol. 15, 160.

Marushige, K. and Ozaki, M. (1967)). Develop. Biol. 16, 474.

Mathews, M.B. (1972). Biochem. Biophys. Acta. 272, 108.

Mathews, M.B. and Korner,A. (1970)). Eur. J. Biochem. 17, 328.

Mathews, M.B. and Osborn, M. (1971). Nature New Biol. 233, 206.

Mathews, M.B., Pragnell, I.B., Osborn, M. and Arnstein, H.R.V. (1972). Biochim. Biophys. Acta. 287, 113.

Mathias, A.P., Williamson, R., Huxley, M.E. and Page, S. (1964). J. Mol. Biol. 9, 154.

Mayo, U.S. and DeKloet, S.R. (1971). Biochem. Biophys. Acta. 247, 74.

McCarthy, B.J. and Hoyer, B.H. (1964). Proc. Nat. Acad. Sci. U.S.A. 52, 915.

McConaughy, B.L., Laird, C.D. and McCarthy, B.J. (1969). Biochemistry 8, 3289.

McEwen, C.R. (1967). Anal. Biochem. 20, 114.

McKnight, G.S. and Schimke, R.T. (1974). Proc. Nat. Acad. Sci. U.S.A.

McLaughlin, C.S., Warner, J.R., Edmonds, M. Nakazato, M. and Vaughn, M.H. (1973). J. Biol. Chem. 248, 1466.

Melli, M., Whitfield, C., Rao, K.V., Richardson, M. and Bishop, J.O. (1971). Nature New Biol. 231, 8.

Mendecki, J., Lee, Y. and Brawerman, G. (1972). Biochemistry, 11, 792.

Miller, O.L. and Bakken, A.M. (1972). Acta. Endocrinologica, Suppl. 168, 155.

Miller, O.L. and Bakken, A.M. (1973). Fifth Karolinska Symposium on Res. Methods im Reproductive Endocrinology (Diczfalusy, E. Ed.).155. ((Bogtrykkerit Forum, Copenhagen).

Miller, O.L. and Beatty, B.R. (1969a). Science, 164, 455.

Miller, O.L. and Beatty, B.R. (1969b). Genetics, Supply. 61, 1, 133.

Min-Jou, W., Haegerman, G., Ysebaerl, M. and Friers, W. (1972). Nature, 237, 82.

Molloy, G.R. and Darnell, J.E. (1973). Biochemistry, 12, 2324.

Molloy, G.R., Jelenick, W., Saldiitt, M. and Darnell, J.E. (1974). Cell 1, 43.

Molloy, G.R., Sporn, M.B., 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.A. 69, 3684.

Mondel, M., Sutton, A., Chen. V., and Sarkou, S. (1974). Biochim. Biophys. Res. Commun. 56, 988.

Morel, L., Kayibandu, B. and Scherrer, K. (1971). FEBS Letters 18, 84.

Morris, G.E., Buzash, E.A., Rourke, A.W., Tepperman, K., Thompson, W.C. and Heywood, S.M. Cold Spring Harb. Quant. Biol. 37, 535.

Morrison, M.R., Brinkley, S.A., Gorshi, J. and Lingrel, J.B., (1974). J. Biol.Chem.

Morrison, M.R., Merkel, C.G. and Lingrel, J.B. (1973). Mol. Biol. Rept. 1, 55.

Morrison, M.R., Paul, J. and Williamson, R. (1972). Eur. J. Biochem. 27, 1.

M.S.E. Technical publications. (1966). No.49.

Murphy, W. and Attardi, G. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 115.

Nakazato, H., Edmonds, M. and Kopp, D.W. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 200.

Nakazato, H., Kopp, D.W. and Edmonds, M. (1973). J. Biol. Chem. 248, 1472.

Niessing, J. and Sekeris, C.E. (1970). Biochim. Biophys. Acta. 209, 484.

Niessing, J. and Sekeris, C.E. (1971b). Biochim. Biophys. Acta. 247, 391.

Niessing, J. and Sekeris, C.E. (1973). Nature New Biol. 243, 9.

Niyogi, S.K. and Thomas, C.A. (1967). Biochem. Biophys. Res. Commun. 26, 51.

Nirenberg, M.W. and Matthaei, J.H. (1961). Proc. Nat. Acad. Sci. 47, 1588.

Nudel, U., Lebleu, B. and Revel, M. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2139.

Nygaard, A.P. and Hall, B.D. (1964). J. Mol. Biol. 9, 125.

Packman, S., Aviv, H., Ross, J. and Leder, P. (1972). Biochem. Biophys. Res. Commun. 49, 813.

Painter, T.S. and Griffen, A.B. (1937). Genetics, 22, 612.

Palacios, R., Sullivan, D., Summers, N.M., Kiely, M.L. and Schimke, R.T. (1973). J. Biol. Chem. 248, 540.

Palmiter, D. (1973). J. Biol.Chem. 248, 2095.

Pardue,M.L. and Gall, J.G. (1970). Science, 168, 1356.

Paul, J. (1965). "Cell and Tissue Culture" 83. (E. and S. Livingstone Ltd., Edinburgh and London).

Paul, J. (1972). Nature, 238, 444.

Paul, J., Conkie, D. and Freshney, R.I. (1969). Cell Tissue Kinet. 2, 283.

Paul, J. and Gilmour, R.S. (1968). J. Mol. Biol. 34, 305.

Paul, M., Goldsmith, M.R., Hunsley, J.R. and Kafatos, F.C. (1972). J. Cell. Biol. 55, 653.

Pederson, T. (1974a). J. Mol. Biol. 83, 164.

Pederson, T. (1974b). Proc. Nat. Acad. Sci. 71,

Pemberton, R.E. and Baglioni, C. (1972). J. Mol. Biol. 235, 99.

Penman, S. (1966). J. Mol. Biol. 17, 177.

Penman, S., Rosbash, M. and Penman, M. (1970). Proc. Nat. Acad. Sci. U.S.A. 67, 1878.

Penman, S., Sherrer, K., Becker, Y. and Darnell, J.E. (1963). Proc. Nat. Acad. Sci. U.S.A. 49, 654.

Penman, S., Vesco, C. and Penman, C.(1968) J. Mol. Biol. 34, 49.

Perkowska, E., MacGregor, H.L. and Birnstiel, M.L. (1968). Nature, 217, 649.

Perry, R.P. and Kelly, D.E. (1968). J. Mol. Biol. 35, 37.

Perry, R.P. and Kelly, D.E. (1973). J. Mol. Biol. 79, 681.

Perry, R.P., Kelly, D.E. and LaTorre, J. (1974).J. Mol. Biol. 82, 315.

Perry, R.P.,LaTorre, J., Kelly, D.E. and Greenberg, J.R.(1972). Biochim. Biophys. Acta. 262, 220.

Philipson, L., Wall, R., Glickman, G. and Darnell, J.E. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 2806.

Premhumer, E., Shoyab, M. and Williamson, A.R. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 99.

Proudfoot, N.J. and Brownlee, G.G. (1974). FEBS Letters 38, 179.

Revel, M., Aviv, H., Groner, Y. and Pollack, V. (1970). FEBS Letters, 9, 213.

Rich, A., Warner, J.R. and Goodman, H.M. (1963). Cold Spring Harbor Symp. Mol. Biol. 28, 269.

Ris, H. (1960). 4th Int. Conf. Electron Micros. 2, 199. (Springer-Verlay, Berlin).

Ris, H. (1961). Can. J. Genet. Cytol. 3, 95.

Ritossa, F.M., Atwood, K.C., Kindsley, D.L. and Spiegelman, S. (1966b). Natl. Cancer Inst. Monograph, 23, 449.

Ritossa, F.M., Atwood, K.C. and Spiegelman, S. (1966a). Genetics, 54,663.

Robertson, H.D., Barrell, B.G., Weith, H.L. and Donelson, J.E. (1973). Nature New Biol. 241, 38.

Robbins, E. and Borun, T.W. (1967). Proc. Nat. Acad. Sci. U.S.A. 57, 409.

Rosbash, M. and Ford, P.J. (1974). J. Mol. Biol. 85, 87.

Rosenfeld, G.C., Conistock,J.P., Mears, A.R. and O'Malley, B.W. (1972). Biochem. Biophys. Res. Commun. 47, 387.

Ross, J., Aviv, H., Scolnick, E. and Leder, P. (1972b). Proc. Nat.Acad. Sci. U.S.A. 69, 264.

Ruger, W. and Bautz, E.K.F. (1968). J. Mol. Biol. 31, 83.

Ruiz-Carrillo, A., Beato, M., Schultz, G., Feigelson, P. and Allfrey, U.G. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 3641.

Ryskov, A.P., Farashyan, V.R. and Georgiev, G.P. (1972). Biochim. Biophys. Acta. 262, 568.

Sampson, J., Mathews, M.B., Osborn, M. and Borghetti, A.F. (1972). Biochemistry, 11, 3636.

Scherrer, K., Latham, H. and Darnell, J.E. (1963). Proc. Nat. Acad. Sci. U.S.A. 49, 240.

Scherrer, K. and Marcaud, L. (1968). J. Cell. Physiol. 72, Suppl,1, 181.

Scherrer, K., Marcaud, L., Zajdela, F., London, I.M. and Gros, F. (1966). Proc. Nat. Acad. Sci. U.S.A. 56, 1571.

Scherrer, K., Spokr, G., Granboulan, N., Morel, C., Groselaude, J. and Chezzi, C. (1970). Cold Spring Harbor Symp. Quant. Biol. 35, 539.

Schochetman, G. and Perry, R.P. (1972). J. Mol. Biol. 63. 591.

Sheiness, D. and Darnell, J.E. (1973). Nature New Biol. 241, 265.

Sheldon, R., Jarale, C. and Kates, J. (1972a) Proc. Nat. Acad. Sci. U.S.A. 69, 417.

Sheldon, R., Kates, J., Kelly, D.E. and Perry, R.P. (1972b). Biochemistry, 11, 3829.

Shenkin, A. and Burdon, R.H. (1972). FEBS Letters, 22, 157.

Slater, I. and Slater, D.W. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1103.

Smith, G.P. (1973). Cold Spring Harb. Symp. Quant. Biol. 38, 507.

Smith, M.J., Hough, B.R., Chamberlin, M.E. and Davidson, E.H. (1974). J. Mol. Biol. 84,

Soeiro, R., Vaughan, M.H., Warner, J.R. and Darnell, J.E. (1968). J. Cell. Biol. 39, 112.

Southern, E.M. (1971). Nature New Biol. 232, 82.

Southern, E.M. and Mitchell, A.R. (1971). Biochem. J. 123, 613.

Spirin, A.S. (1966). In Current Topics in Developmental Biology. (Eds. A.A. Moscona and A. Monroy, Academic Press, New York) 1, 1.

Spirin, A.S. (1969). Eur. J. Biochem. 10, 20.

Staekelin, M., Rogg, H., Bagaley, B.C., Grunsberg, T. and Weholi, W. (1968). Nature, 219, 1363.

Staples, D.H., Hindley, J., Billeter, M.A. and Weissmann, C. (1971). Nature New Biol. 234, 202.

Steggles, A.W., Wilson, G.N., Kantor, J.A., Picciano, D.J., Falvey, A.K. and Anderson, W.F. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1219.

Steinheider, G., Medleris, H. and Ostertag, W. (1971). in Syntheses, Struktur and Funktion des Hamoglobins (Martinand Nowichi, Eds.) 225. J. F. Lehmans, Verlag, Munchen.

Steitz, J.A. (1969). Nature, 224, 957.

Stevens, B.J. and Swift, H. (1966). J. Cell. Biol. 31, 55.

Stevens, R.H. and Williamson, A.R. (1973). Nature New Biol. 245, 101.

Storb, U. (1974). Biochem. Biophys. Res. Commun. 57, 31.

Strauss, J.H., Kelly, R.B. and Sinsheimer, R.L. (1968). Biopolymers, 6, 793.

Studier, F.W. (1965). J. Mol. Biol. 11, 373.

Subirana, J.A. (1966). Biopolymers, 4, 171.

Sussman, M. (1970). Nature, 225, 1245.

Sutton, W.D. (1971). Biochim. Biophys. Acta, 240, 522.

Sutton, W.D. and McCallum, M. (1971). Nature New Biol. 232, 83.

Suzuki, Y. and Brown, D.D. (1972). J. Mol. Biol. 63, 409.

Suzuki, Y., Gage, L.P. and Brown, D.D. (1972). J. Mol. Biol. 70, 637.

Tashiro, Y., Morimoto, T., Matsuura, S. and Nagata, S. (1968). (1968). J. Cell Biol. 38, 574.

Ullman, J.S. and McCarthy, B.J. (1973a). Biochim. Biophys. Acta. 294, 405.

Ullman, J.S. and McCarthy, B.J. (1973b) Biochim. Biophys. Acta. 294, 416.

Verma, I.M., Temple, G.F., Fan, H. and Baltimore, D. (1972). Nature New Biol. 235, 163.

Wagner, A.F., Bugianesi, R.L. and Shen, T.Y. (1971). Biochem. Biophys. Res. Commun. 45, 1184.

Walker, P.M.B. (1969). Prog. Nucl. Acad. Res. 9, 301.

Walker, P.M.B. and McLaren, A. (1965). J. Mol. Biol. 12, 394.

Warner, J.R., Soeiro, R., Birnboim, H.C., Girard, M. and Darnell, J.E. (1966). J. Mol. Biol. 19, 349.

Wegnez, M., Monier, R. and Denis, H. (1972). FEBS Letters, 25, 13.

Weinberg, E.S., Birnstiel, M.L., Purdom, D.R. and Williamson, R. (1972). Nature 240, 225.

Weinstein, B.I., Schechter, A.N., Burka, E.R. and Marks, P. (1963). Science, 140, 314.

Wellauer, P.K. and Dawid, I.B. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2827.

Wensink, P.C. and Brown, D.D. (1971). J. Mol.Biol. 60, 253.

Wessels, N. and Wilt, F.H. (1965). J. Mol. Biol. 13, 767.

Wetmur, J.G. and Davidson, W. (1968). J. Mol. Biol. 31, 349.

Wiggle, D.T. and Smith, A.E. (1973). Nature New Biol. 242, 136.

Williamson, A.R. (1974). Reported at Biochem. Soc. Meeting, Glasgow.

Williamson, R.and Brownlee, G.G. (1969). FEBS Letters, 3, 306.

Williamson, R., Drewienkiewicz, C.E. and Paul, J. (1973). Nature New Biol. 241, 66.

Williamson, R., Morrison, M., Lanyon, W.G., Eason, R. and Paul, J. (1971). Biochemistry, 10, 3014.

Williamson, R., Morrison, M. and Paul, J. (1970). Biochem. Biophys. Res. Commun. 40, 740.

Wilson, D.A. and Thomas, C.A. (1973). Biochim. Biophys. Acta. 331, 333.

Wilson, E.B.(1965). The Cell in Development and Heredity, 3rd ed. Macmillan, New York.

Wilt, F.H. (1965). J. Mol. Biol. 12, 331.

Wintrobe, M.M. (1961). in 'Clinical Haematology' 5th ed. (Kimpton, H. Ed.) 116, London.

Yankofsky, S.A. and Spiegelman, S. (1962). Proc. Nat. Acad. Sci. U.S.A. 48, 1466.

Yogo, Y. and Wimmer, E. (1972). Proc. Nat. Acad. Sci. U.S.A. 69, 1877.

Young, B.D., Harrison, P.R., Gilmour, R.S., Birnie, G.D., Hell, A., Humphries, S. and Paul, J. (1974). J. Mol. Biol. 84, 555.

TRANSCRIPTION OF THE GLOBIN GENE

R.S. GILMOUR, S.E. HUMPHRIES, E.C. HALE and J. PAUL
Beatson Institute for Cancer Research
132, Hill Street, Glasgow G 3 6 UD

INTRODUCTION

Previous reports from this laboratory (1-3) have provided evidence that there exists in animal cells a control mechanism which provides for the transcription of tissue specific RNA sequences. In these studies the overall spectrum of RNA sequences transcribed in vitro from chromatin by a bacterial polymerase was analysed by a low C_0t, RNA driven hybridisation reaction, where only annealing to the relatively reiterated DNA sequences takes place. It has also been reported recently (4,5) that globin mRNA, a tissue specific messenger, is transcribed from a unique DNA sequence, and therefore would not have been detected by the previous type of analysis. The recent use of the RNA dependent polymerase (reverse transcriptase) from avian myeloblastosis virus to obtain DNA copies (cDNA) of globin mRNA from reticulocytes (6-8) offers a hybridisation method for analysing RNA transcribed in vitro from chromatin for a specific type of RNA sequence derived from the unique fraction of the genome. In the present report chromatin of foetal mouse liver, a haemopoietic tissue, is examined using this new approach.

RESULTS

(a) The detection of globin mRNA in embryonic mouse livers.

Total polysomal RNA was isolated from 14 day embryonic mouse livers, a tissue that contains greater than 70% nucleated erythropoietic cells. Cytoplasmic extracts were prepared with NP-40 (9) and polysomes sedimented at 60,000 g x 90 mins. Messenger RNA was fractionated from the total RNA on Sepharose-bound polyU columns as described by Darnell et al. (10). RNA bound to the column was analysed on polyacrylamide gels and tested for the ability to direct globin synthesis in a duck lysate cell free system as described by Lanyon et al. (11) and Lockard and Lingrel (12).

Fig. 1 shows the analysis of polyU bound polysomal RNA on a 2.6% polyacrylamide gel. Small amounts of 28S and 18S ribosomal RNA were present in addition to a substantial peak of material running in the 9S region as judged by co-chromatography with known globin mRNA isolated from mouse reticulocytes. When tested in a duck reticulocyte lysate cell free system it could be demonstrated that the 9S RNA contained mRNA for mouse globin (Fig. 2). Clearly, the appearance of haemoglobin in the developing embryonic mouse liver can be correlated with presence of 9S globin mRNA in the cytoplasm.

(b) The detection of globin mRNA sequences in RNA synthesised from chromatin.

It is of considerable interest to ask whether the globin mRNA of embryonic liver exists as presynthesised RNA or whether it is actively transcribed from the chromatin. Because of the relative longevity of mRNA in the cytoplasm amounts large enough to be detected by a cell free system will accumulate. This favoured situation will not obtain to RNA synthesised in vitro from chromatin. Here a much more sensitive analysis has been adopted.

9S RNA was purified from the reticulocytes of phenylhydrazine anaemic mice by passing polysomal RNA twice through a polyU-Sepharose column. The bound RNA ran as a single 9S component on polyacrylamide gels and directed the synthesis of globin in the duck lysate cell free system. A complementary DNA copy (cDNA) of 9S RNA was prepared using the reverse transcriptase as described by Harrison et al. (13). [^{3}H] -dTTP and [^{3}H] -dCTP were incorporated into

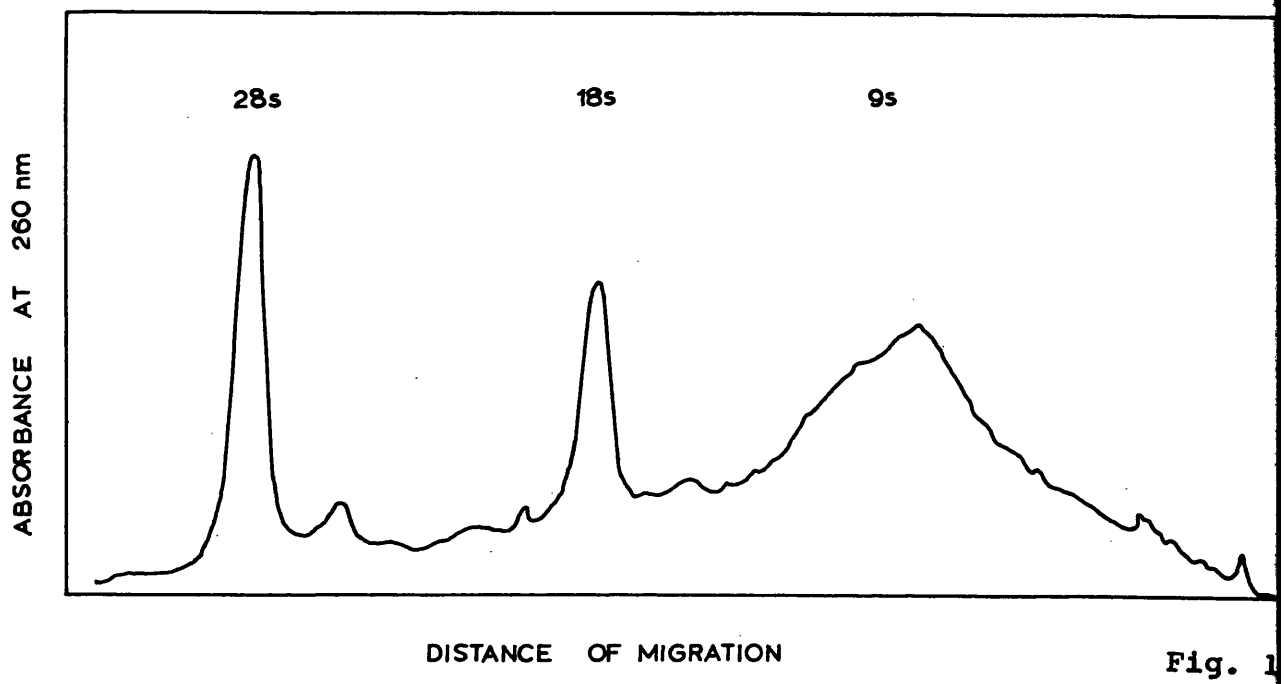


Fig. 1

FIG. 1. Electrophoresis of polyU-Sepharose bound embryonic mouse liver polysomal RNA on a 2.6% polyacrylamide gel.

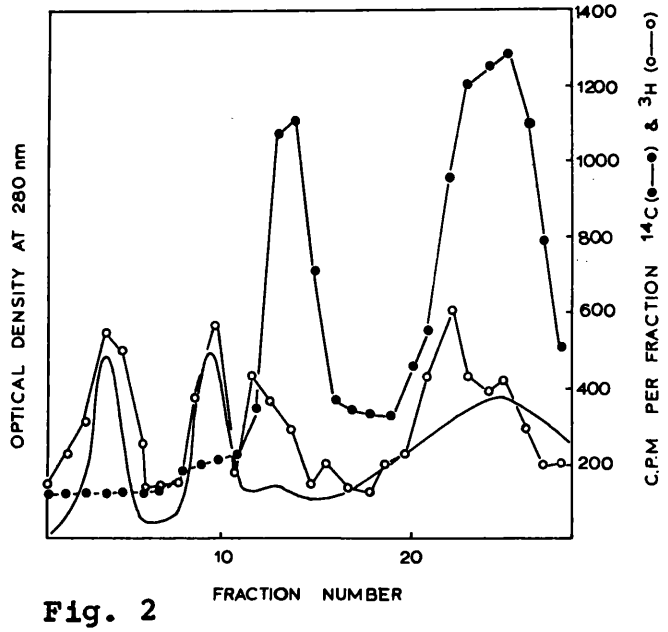


Fig. 2

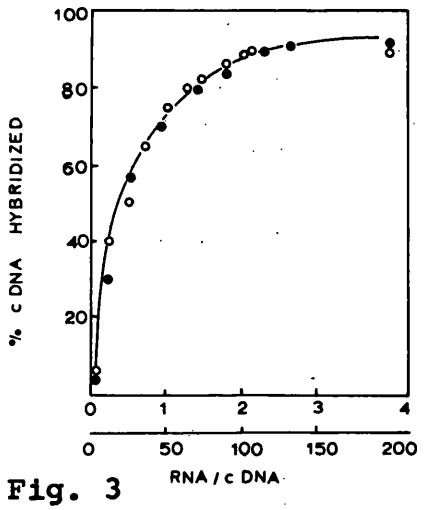


Fig. 3

FIG. 2. CMC chromatogram of globins from a duck reticulocyte cell free system to which was added polyU-Sepharose bound polysomal RNA from embryonic mouse liver.

Solid line is the OD profile. β and α carrier mouse globin peaks elute first followed by minor and major duck peaks.

(●—●) Endogenous incorporation of ^{14}C-leucine.

(○—○) Incorporation of ^{3}H-leucine in the presence of polyU-Sepharose bound RNA.

FIG. 3. Hybridization to cDNA of 9S reticulocyte globin mRNA (○——○) and reticulocyte polysomal RNA (●——●).

cDNA to give a final specific activity of 20 x 10^6 d.p.m./μg. Titration curves were carried out with varying amounts of test RNA by hybridising in 10 μl 50% formamide, 0.5 M NaCl; 0.025 M HEPES pH 7; 0.01 M EDTA at 43°C for 3.5 days. The hybrid formed was assayed using the single stranded nuclease of Sutton (14) prepared from Takadiastase.

Fig. 3 shows titration curves in which constant amounts of cDNA (0.001 μg) were hybridised to increasing amounts of reticulocyte 9S RNA and reticulocyte polysomal RNA. With 9S RNA, maximum levels of hybridisation were obtained at input RNA/cDNA ratios of 1-1.4, consistent with the finding that about 60% of the 9S sequence is represented in cDNA. With reticulocyte polysomal RNA maximum hybridisation is obtained at input ratios of about 50 which is consistent with the known 9S content of reticulocyte polysomal RNA (2%).

RNA was transcribed *in vitro* from embryonic mouse liver and mouse brain chromatin using *E.coli* RNA polymerase prepared according to the method of Burgess (15).

Hybridisation to cDNA of RNA obtained in this way is shown in Fig. 4. RNA from brain chromatin did not show significant hybridisation while that from embryonic liver chromatin hybridised to about 40% of the cDNA at input ratios above 25,000. It is important to eliminate the possibility that the hybridising RNA from liver chromatin was derived from endogenous RNA and not from RNA transcribed *de novo* by the bacterial polymerase. This was confirmed by preparing RNA from liver chromatin as before with polymerase and nucleoside triphosphates one of which was highly labelled with P. After hybridisation of the RNA to cDNA the reaction mixture was treated with ribonuclease for 1 h at 37°C, passed through Sephadex G.75 and the excluded material centrifuged to equilibrium in a CsCl gradient. Hybridised and unhybridised cDNA can be separated by this technique while remaining RNA fragments are pelleted. Fig. 5 shows the distribution of ^3H and ^{32}P label obtained. The majority of the ^3H label appeared in the lower half of the gradient normally expected for hybridised material while unhybridised cDNA formed a smaller less dense band. ^{32}P labelled RNA was found to be associated with the hybrid material such, that it comprised at least 80% of hybridised RNA. These experiments show that the ability of embryonic liver chromatin to synthesis globin mRNA sequences remains intact in the isolated chromatin.

(c) Reconstitution studies with chromatin

Previous studies (16) suggest that the specificity of transcription in chromatin is an inherent property in that if the chromatin components are dissociated in high salt/urea solutions and then gradually reconstituted by gradient dialysis to low salt concentrations, the resulting chromatin still possesses its original specificity. This was re-investigated by dissolving embryonic liver chromatin at 0.5 mg/ml in 2 M NaCl; 5 M urea; 0.01 M Tris, pH 8.3 and dialysing against 5 M urea, 0.01 M Tris containing 0.6 M, 0.4 M and 0.2 M NaCl for 16 h, 2 h and 2 h respectively. RNA was synthesised *in vitro* from the reconstituted chromatin and compared by hybridisation to cDNA with RNA from untreated embryonic liver chromatin and reconstituted brain chromatin (Fig. 6). Globin sequences were transcribed from both control and reconstituted liver chromatins while background levels were obtained with RNA from brain chromatin. It was also suggested from previous work that the non-histone protein fraction of the chromatin was responsible for conferring specificity of transcription in chromatin. A non-histone protein fraction was obtained from embryonic liver chromatin using the hydroxylapatite method of MacGillivray, *et al.* (17). This fraction was added to brain chromatin in high salt/urea and the chromatin reconstituted as before. The RNA transcribed from this chromatin

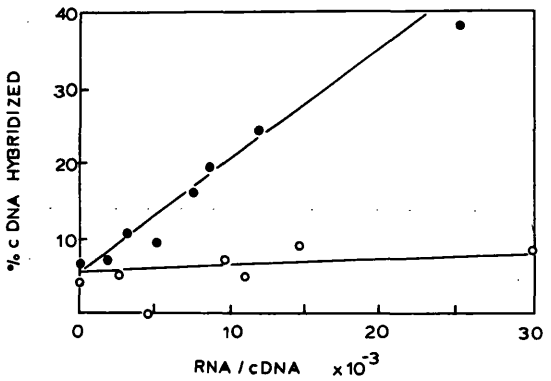


Fig. 4

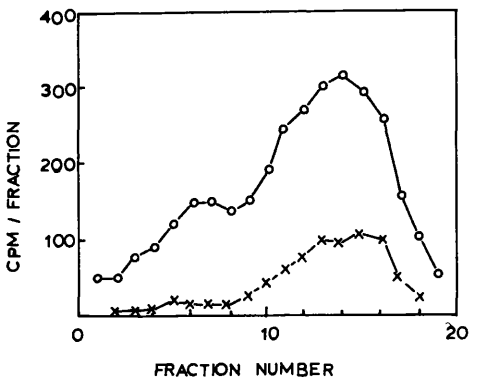


Fig. 5

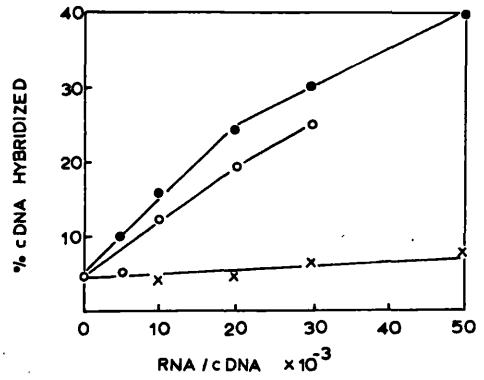


Fig. 6

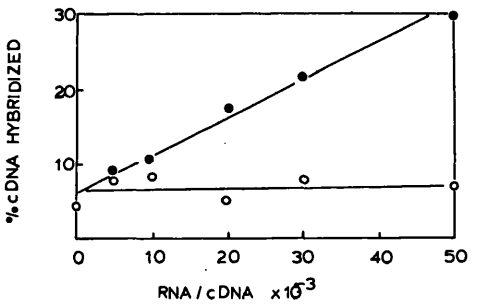


Fig. 7

FIG. 4. Hybridization to cDNA of RNA transcribed *in vitro* from embryonic mouse liver chromatin (●——●) and mouse brain chromatin (o——o).

FIG. 5. Caesium chloride gradient analysis of hybridization of ^{32}P RNA (x——x) transcribed *in vitro* from embryonic mouse liver chromatin and ^{3}H -cDNA (o——o). The larger, more dense peak represents hybridized cDNA.

FIG. 6. Hybridization to cDNA of RNA transcribed *in vitro* from native embryonic liver chromatin (●——●)
reconstituted embryonic liver chromatin (o——o)
reconstituted brain chromatin (x——x)

FIG. 7. Hybridization to cDNA of RNA transcribed *in vitro* from reconstituted brain chromatin (o——o)
and brain chromatin reconstituted with liver non histone proteins, (●——●).

now showed a significant ability to hybridise with cDNA when compared with brain chromatin which had been reconstituted in the absence of liver non-histone material.

REFERENCES

1. Paul, J. and Gilmour, R.S., J. Mol. Biol. (1966) 16, 242

2. Paul, J. and Gilmour, R.S., J. Mol. Biol. (1968) 134, 305

3. Gilmour, R.S. and Paul, J., FEBS Letts.,(1970) 9, 242

4. Harrison, P.R., Hell, A., Birnie, G.D. and Paul, J., Nature (1972) 239, 219

5. Bishop, J.O., Pemberton, R. and Baglioni, G., Nature (1972) 235, 231

6. Verma, I.M., Temple, G.F., Fan, H.and Baltimore, D., Nature (1972) 235, 163

7. Ross, J., Aviv, H., Scolnick, E. and Leder, P., Proc. Natl. Acad. Sci. U.S. (1972) 69, 264

8. Kacian, D.L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L. and Marks, P.A., Nature (1972) 235, 167

9. Borun, T.W., Robbins, E. and Scharff, M.D., Proc. Natl. Acad. Sci. U.S. (1967) 58, 1977

10. Darnell, J.E., Phillipson, L., Wall, R. and Adensik, M., Science (1971) 174, 507

11. Lanyon, W.G., Paul, J. and Williamson, R., Eur. J. Biochem. (1972) 31, 38

12. Lockard, R.E. and Lingrel, J.B., J. biol. Chem. (1972) 247, 4174

13. Harrison, P.R., Hell, A. and Paul, J., FEBS Letts. (1972) 24, 73

14. Sutton, W.D., Biochim. biophys. Acta (1971) 240, 522

15. Burgess, R.R., J. biol. Chem. (1969) 244, 6160

16. Gilmour, R.S. and Paul, J., J. Mol. Biol. (1969) 40, 137

17. MacGillivray, A.J., Cameron, A., Krauze, R.J., Rickwood, D. and Paul, J., Biochim. biophys. Acta (1972) 227, 384

Summary

This report provides a more rigorous proof of previous findings that the RNA transcribed *in vitro* from the chromatins of different organs shows different sequence specificities. Here the particular case of the globin gene is considered for a comparison of embryonic mouse liver chromatin and mouse brain chromatin using the reverse transcriptase cDNA copy of globin 9S mRNA as a definitive probe. It can be shown that globin sequences are transcribed *in vitro* from embryonic liver chromatin and not brain chromatin. This specificity in liver chromatin can be reconstituted after dissociation of the structural elements of the chromatin. It can be shown that the non-histone protein fraction of liver chromatin can confer specificity for the transcription of globin sequences from brain chromatin which otherwise lacks this ability.

Résumé

Transcription du gène de la globine

Ce rapport fournit une preuve plus rigoureuse des résultats précédents, que les RNAs transcrits *in vitro* des chromatines d'organes divers montrent des différentes spécificités de séquences. Ici, pour la comparaison des chromatines du foie d'embryon de souris et du cerveau de souris, on considère le cas spécial du gène pour l'hémoglobine, en utilisant comme source définitive le cDNA, copie synthétisée à l'aide de transcriptase-reverse, du mRNA 9S pour la globine. On peut montrer que des séquences de la globine sont transcrites *in vitro* à partir de la chromatine du foie embryonnaire, mais pas à partir de la chromatine de cerveau. Cette spécificité de la chromatine de foie peut être reconstituée après la dissociation des éléments de structure de la chromatine. On peut montrer que la fraction de protéine non-histone de la chromatine de foie transmet la spécificité de transcription de séquences de globine à la chromatine de cerveau, un tissu qui en est autrement incapable.

[Reprinted from Biochemistry, (1974) 13, 703.]

Translation of Mouse Globin Messenger Ribonucleic Acid from Which the Poly(adenylic acid) Sequence Has Been Removed†

Robert Williamson,* Jennifer Crossley, and Stephen Humphries

ABSTRACT: Mouse globin mRNA was isolated using oligo(dT)-cellulose and treated with *Micrococcus* polynucleotide phosphorylase to remove nucleotides from the 3′ end. The material which no longer bound to oligo(dT)-cellulose was then separated from that which still bound. The molecular weight of the major peak of the nonbinding mRNA was reduced from 220,000 to 185,000, as calculated from mobility in polyacrylamide gel electrophoresis. The treated mRNA which still bound to oligo(dT)-cellulose migrated at the same mobility as untreated globin mRNA. The treated mRNA did not function as a template for reverse transcriptase. Two-dimensional fingerprint analysis of phosphokinase-labeled fragments from a combined pancreatic T1 ribonuclease digest demonstrated the absence of any large oligo(A) fragments from the nonbound mRNA; at least 90% of the poly(A) region had been removed. The treated nonbound mRNA stimulated globin synthesis in the Krebs ascites cell system, but the nonbound mRNA was only approximately 50% as efficient as treated bound mRNA as judged by product analysis on carboxymethylcellulose columns. In the case of mouse globin mRNA the presence of a long poly(A) sequence is not required for successful translation.

Poly(adenylic acid) sequences have been demonstrated to occur at the 3′ termini of all mRNAs from higher cells (Edmonds *et al.*, 1971; Darnell *et al.*, 1971; Lee *et al.*, 1971) with the exception of histone mRNAs (Adesnik *et al.*, 1972; Adesnik and Darnell, 1972; Schochetman and Perry, 1972). Various roles may be suggested for these sequences: (a) markers for the processing of heterogeneous nuclear RNA (HnRNA) to messenger RNA (mRNA) (Edmonds *et al.*, 1971; Philipson *et al.*, 1971; Jelinek *et al.*, 1973; Darnell *et al.*, 1972); (b) transport from nucleus to cytoplasm (Philipson *et al.*, 1971); (c) regulation of mRNA half-life (Sheiness and Darnell, 1973; Sussman, 1970); and (d) binding of poly(A)-specific proteins (Kwan and Brawerman, 1972; Blobel, 1973).

Histone mRNAs contain no poly(A) sequences yet migrate to the cytoplasm and are translated (Adesnik *et al.*, 1972; Adesnik and Darnell, 1972; Schochetman and Perry, 1972; Jacobs-Lorena *et al.*, 1972). On the other hand, viral mRNAs may contain poly(A) sequences although synthesized in the cytoplasm (Yogo and Wimmer, 1972; Johnston and Bose, 1972). We have used polynucleotide phosphorylase to remove the poly(A) sequence from mouse globin mRNA and tested the product in the Krebs ascites protein synthesis system for its ability to direct mouse globin synthesis.

Materials and Methods

Mouse reticulocyte polysomal RNA was prepared as previously described (Williamson *et al.*, 1971) and the globin mRNA was isolated using oligo(dT)-cellulose (Aviv and Leder, 1972). Oligo(dT)-cellulose was purchased from Collaborative Research, Waltham, Mass. The purity of the mRNA preparation was confirmed by 2.6% polyacrylamide gel electrophoresis and by its ability to direct mouse globin synthesis in a duck reticulocyte lysate cell-free system, both as previously described (Williamson *et al.*, 1971; Lingrel *et al.*, 1971; Lanyon *et al.*, 1972). [^{3}H]- and [^{14}C]leucine and [γ-^{32}P]-ATP were purchased from the Radiochemical Centre, Amersham.

Micrococcus lysodiekticus polynucleotide phosphorylase (EC 2.7.7.8) was purchased from Worthington Biochemicals. Mouse globin mRNA (100 μg/ml) was incubated with 500 μg/ml of polynucleotide phosphorylase in 50 mM Tris (pH 7.5)–15 mM $MgCl_2$–15 mM potassium phosphate at 37° for 7 min (Grunberg-Manago, 1963). The reaction was stopped by adding one-tenth volume of 10% sodium dodecyl sulfate and the reaction mixture was deproteinized by shaking with a half-volume of phenol–chloroform. The RNA was precipitated with ethanol, dissolved in 0.5 M NaCl–0.01 M Tris (pH 7.5)–0.1% sodium laurylsarcosine, and passed through a dT-cellulose column. The material which was not retained was collected and ethanol precipitated; the proportion of unretained RNA varied from preparation to preparation but was of the order of 80%. The size of the treated mRNA was compared with that of the original mRNA on 6% polyacrylamide gels, with added marker 4S and 5S RNA from mouse reticulocytes.

The ability of the treated nonbound mRNA to be copied by reverse transcriptase with an oligo(dT) primer was compared with that of untreated mouse globin mRNA. The enzyme was prepared and the assay carried out as described by Harrison *et al.* (1972).

Two-dimensional fingerprint analysis of treated and untreated mRNA was performed after digestion for 45 min at 37° with pancreatic and T1 ribonucleases each at 0.05 mg/ml in 10 μl of 0.01 M Tris (pH 7.4). The digestion was terminated by addition of an equal volume of water-saturated phenol–chloroform (1:1) and the aqueous phase was dried down after deproteinization. The oligonucleotide mixture was dissolved in 1% sodium dodecyl sulfate containing 0.1 mg/ml of bacterial alkaline phosphatase and incubated for 30 min at 37°. The reaction mixture was deproteinized, dried down, and then taken up in 1 μl of 0.1 M Tris, 11 mM mercaptoethanol, 11 mM $MgCl_2$, 0.27 mM [γ-^{32}P]ATP, and 2 units of polynucleotide kinase (a gift from Dr. Ken Murray, University of Edinburgh). Incubation was carried out at 37° for 45 min,

† The Beatson Institute for Cancer Research, Glasgow, G3 6UD, Scotland. *Received July 30, 1973.* This work was supported by grants to the Beatson Institute for Cancer Research from the Medical Research Council and the Cancer Research Campaign.

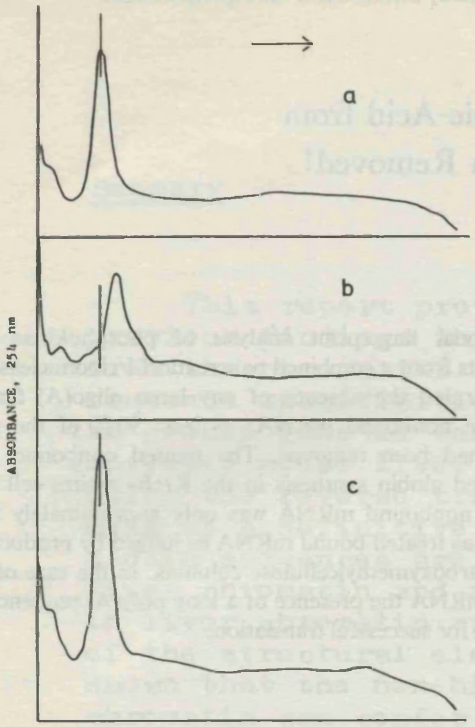


FIGURE 1: 6% polyacrylamide gel electrophoresis of globin mRNAs following incubation with polynucleotide phosphorylase. Migration from left to right; gels 12 cm long and 0.7 cm diameter; electrophoresis at 2.5 mA/gel for 30 min and then 10 mA/gel for 4 hr (Williamson *et al.*, 1971): (a) 13 μg of untreated 9S globin mRNA; (b) 13 μg of phosphorylase-treated nonbound mRNA; (c) 15 μg of phosphorylase-treated bound mRNA.

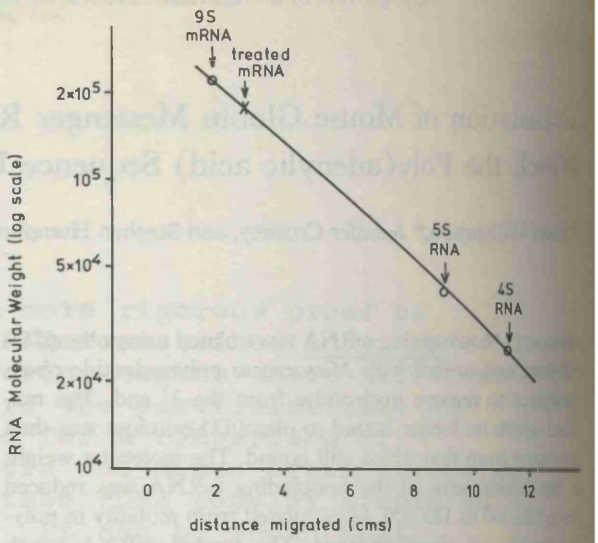


FIGURE 2: Determination of molecular weight of treated globin mRNA. The treated nonbound mRNA was run on 6% polyacrylamide gels with markers of untreated 9S globin mRNA (mol wt 2.20×10^5, Williamson *et al.*, 1971), mouse 5S RNA (mol wt 4.1×10^4, Williamson and Brownlee, 1969), and mouse 4S RNA (mol wt 2.6×10^4, Staehelin *et al.*, 1968).

the mixture was dried, and the ^{32}P-labeled oligonucleotides were taken up in 0.7 μl of dye mixture and separated by electrophoresis on cellulose acetate strips in the first dimension and by chromatography on PEI-cellulose thin-layer plates in the second dimension, as described elsewhere (Southern and Mitchell, 1971; Crossley *et al.*, 1974).

The ability of the treated mRNA to direct protein synthesis was assayed in a preincubated Krebs ascites cell-free system (Mathews and Korner, 1970; Mathews *et al.*, 1972; Mathews, 1972). The dependence of incorporation on added globin mRNA was determined by incubation for 60 min at 37° followed by 10-min digestion with alkali. Protein was precipitated with 10% CCl_3COOH after addition of 50 μg of bovine serum albumin and the precipitates were collected onto Whatman GFC glass fiber filters, washed, dried, and counted in toluene-based scintillator. Product analysis was performed on 100-μl incubation mixtures each containing 5 μg/ml of mRNA, incubated for 60 min at 37°. Carrier mouse globin was added and the protein was extracted with acid acetone at −20°, pelleted, and washed twice with acetone. The globin was then chromatographed on carboxymethylcellulose (Whatman, CM-52) using a gradient of 0.01–0.1 M disodium phosphate (pH 6.8) in 8 M urea–0.05 M mercaptoethanol. Fractions of 5 ml were collected, precipitated onto glass fiber disks, solubilized with NCS (Nuclear-Chicago Corp.), and counted in toluene-based scintillator.

Results

The profiles obtained after electrophoresis on 6% acrylamide gels for untreated mRNA, treated mRNA which does not bind to oligo(dT)-cellulose, and treated mRNA which still binds to oligo(dT)-cellulose are shown in Figure 1. The molecular weight determination for the treated unbound mRNA is shown in Figure 2.

The molecular weight of the peak of unbound phosphorylase-treated mRNA varied in five preparations between 185,000 and 200,000 (av mol wt 190,000) taking the molecular weight of the untreated mRNA as 220,000. The treated bound mRNA migrates to the same position as untreated mRNA. A second smaller peak (of av mol wt 175,000) was usually seen in the treated unbound preparations, as was small amounts of material migrating in the 4S–5S region.

9S mRNA, incubated for 7 min under phosphorolysis conditions but in the absence of polynuceotide phosphorylase, was found to be 100% retained by a dT-cellulose column. This material behaves in an exactly similar manner to the original 9S mRNA in gel and sequence analyses.

Two-dimensional fingerprints of T1 plus pancreatic RNase-treated 9S RNA, phosphorylase-treated nonbound mRNA and poly(A) are shown in Figure 3. Since pancreatic RNase hydrolyzes cyclic purine nucleotides to 3′-phosphate only slowly if at all (Markham, 1957), two series of oligonucleotide spots are seen for poly(A), corresponding to $(pA)_n$-p and $(pA)_n$-OH. For 9S globin mRNA, the proportion of the total radioactivity found in these two oligonucleotide isopliths derived from poly(A) (but not including pAp, which is also found in fingerprints of rRNA) is approximately 20 ± 3%. (A full quantitative analysis of these data for untreated 9S mRNA is in preparation and will appear elsewhere.) The proportion of radioactivity in the oligo(A) tracts for the treated nonbound mRNA is 1.4 ± 0.2%, and the great majority is found in the smallest oligonucleotides.

In an experiment in which both untreated and treated nonstuck globin mRNA were copied by reverse transcriptase, the incorporation obtained is shown in Table I.

The dependence of amino acid incorporation on added treated bound and treated nonbound globin mRNA is shown in Figure 4. In both cases the stimulation of incorporation is linear and similar up to 4 pmol of mRNA added per 50 μl of incubation mixture. The amount of added mRNA was cal-

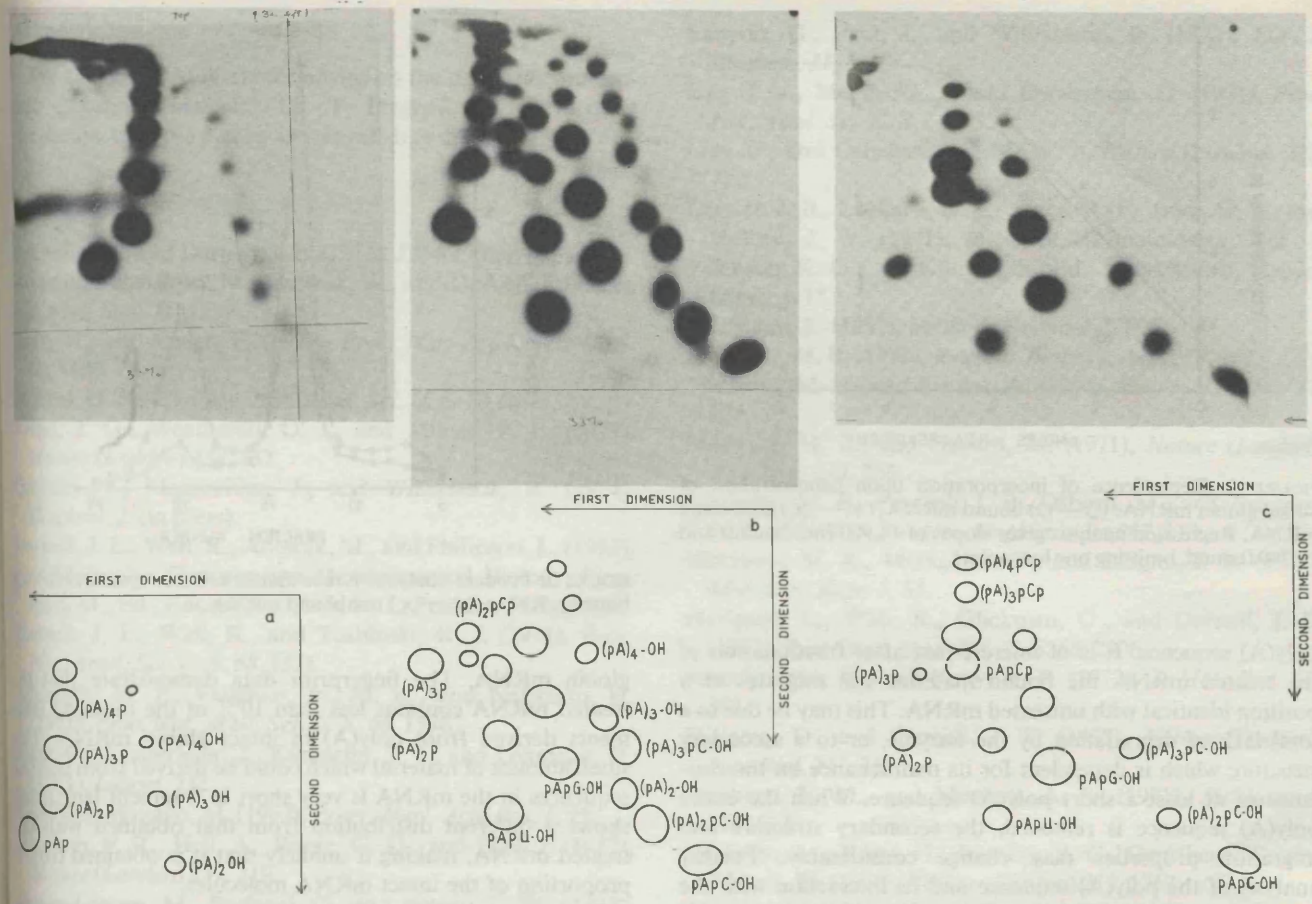


FIGURE 3: Fingerprints of commercial poly(A), untreated 9S mRNA, and phosphorylase-treated nonbound mRNA: (a) commercial poly(A) treated with pancreatic RNase and labeled with phosphokinase; (b) 9S globin mRNA; (c) phosphorylase-treated nonbound mRNA–treated with T1 plus pancreatic ribonucleases followed by phosphatase and then labeled at the 5′ end with phosphokinase. Note the complete absence in (c) of the isoplith $(pA)_n$-OH and the absence of any spots larger than $(pA)_3$-p in the isoplith $(pA)_n$-p.

TABLE I: Transcription of Untreated and Treated Nonbound Globin mRNA with Reverse Transcriptase in the Presence of Oligo(T) Primer.

	Cpm/μg of Template
Control globin mRNA	7.4×10^6
Treated globin mRNA (nonbound)	4.9×10^4

culated using molecular weights of 220,000 (control) and 185,000 (treated, nonbound).

The elution profile of globin chains synthesized after addition of treated bound and nonbound mRNAs in this linear range to the Krebs ascites system is shown in Figure 5. The counts recovered in each globin peak and in the breakthrough peak are given in Table II. The nonbound mRNA is approximately half as active as the control mRNA judged by product analysis, and there is a greater depression of α-chain synthesis than β-chain synthesis.

Discussion

The isolation of the treated mRNA used in these experiments depends upon its nonbinding to oligo(dT)-cellulose. This material migrates at a maximum molecular weight of 200,000 and an average molecular weight of 185,000, at least

TABLE II: Amino Acid Incorporation into Separated Globin Chains.

CCl_3COOH-Precipitable Counts in Peak	Bound 9 S	Nonbound 9 S	Nonbound: Bound (%)
α-Globin	15,370	3,890	25
β-Globin	24,435	15,513	63
Breakthrough	60,000	70,756	
$\alpha + \beta$	39,805	19,203	48
Total	100,552	89,959	89.5

20,000 below the molecular weight of intact mouse globin mRNA. The uncertainty as to the exact molecular weight of mouse globin mRNA does not affect the calculation of the difference in molecular weights, which depends upon the relative mobility of the intact and treated 9S RNA.

Only small amounts of absorbing material migrate other than in the major component of mol wt 185,000 after phosphorylase treatment. We estimate that at least 80% of the RNA is in this peak, and there is little if any material migrating in the position of intact 9S mRNA,

The poly(A) sequence of mouse globin mRNA is approximately 50–70 residues in length (Lim and Canellakis, 1970; Morrison *et al.*, 1973), and therefore the molecular weight reduction is calculated to be sufficient to remove all of the

Acknowledgments

We thank Dr. M. Getz for advice on the use of polynucleotide phosphorylase and Dr. I. Pragnell for performing the assays with the Krebs ascites cell-free system.

References

Adesnik, M., and Darnell, J. E. (1972), *J. Mol. Biol. 67*, 397.
Adesnik, M., Salditt, M., Thomas, W., and Darnell, J. (1972), *J. Mol. Biol. 71*, 21.
Aviv, H., and Leder, P. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 1408.
Bloebel, G. (1973), *Proc. Nat. Acad. Sci. U. S. 70*, 924.
Clegg, J. G., Weatherall, O. J., and Milner, P. F. (1971), *Nature (London) 234*, 337.
Crossley, J., Mansbridge, J., and Williamson, R. (1974), *Biochem. J.* (in press).
Darnell, J. E., Wall, R., Adesnik, M., and Philipson, L. (1972), *in* Molecular Genetics and Developmental Biology, Sussman, M., Ed., Engelwood Cliffs, N. J., Prentice-Hall, p 201.
Darnell, J. E., Wall, R., and Tushinski, R. J. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 1321.
Edmonds, M. P., Vaughan, Jr., M. H., and Nakazato, H. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 1336.
Flatz, G., Kinderlerer, J., Kilmartin, J. V., and Lehmann, H. (1971), *Lancet*, 732.
Grunberg-Manago, M. (1963), *Progr. Nucl. Acid Res. 1*, 93.
Harrison, P. R., Hell, A., Birnie, G. D., and Paul, J. (1972), *Nature (London) 239*, 219.
Jacobs-Lorena, M., Baglioni, C., and Borun, T. W. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 2095.
Jelinek, W., Adesnik, M., Salditt, M., Sheiners, D., Wall, R., Molloy, G., Philipson, L., and Darnell, J. E. (1973), *J. Mol. Biol. 75*, 515.
Johnston, R. E., and Bose, H. R. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 1514.
Kwan, S., and Brawerman, G. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 3247.
Lanyon, G., Paul, J., and Williamson, R. (1972), *Eur. J. Biochem. 31*, 38.
Lee, S. Y., Mendecki, J., and Brawerman, G. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 1331.
Lim, L., and Canellakis, E. S. (1970), *Nature (London) 227*, 710.
Lingrel, J. B., Lockard, R. E., Jones, R. F., Burr, H. E., and Holder, J. W. (1971), *in* Series Haematologica, Vol. 4, Jenson, K. G., and Killman, S., Ed., Munksgaard, Copenhagen, p 37.
Markham, R. (1957), *Methods Enzymol. 3*, 805.
Mathews, M. B. (1972), *Biochim. Biophys. Acta 272*, 108.
Mathews, M. B., and Korner, A. (1970), *Eur. J. Biochem. 17*, 328.
Mathews, M. B., and Osborn, M. (1971), *Nature (London), New Biol. 233*, 206.
Mathews, M. B., Pragnell, I. B., Osborn, M., and Arnstein, H. R. V. (1972), *Biochim. Biophys. Acta 287*, 113.
Morrison, M. R., Merkel, C. G., and Lingrel, J. B. (1973), *Mol. Biol. Rept. 1*, 55.
Philipson, L., Wall, R., Glickman, G., and Darnell, J. E. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2806.
Schochetman, G., and Perry, R. P. (1972), *J. Mol. Biol. 63*, 591.
Sheiness, D., and Darnell, J. E. (1973), *Nature (London), New Biol. 241*, 265.
Southern, E. M., and Mitchell, A. R. (1971), *Biochem. J. 123*, 613.
Staehelin, M., Rogg, H., Bagaley, B. C., Grinsberg, T., and Weholi, W. (1968), *Nature (London) 219*, 1363.
Sussman, M. (1970), *Nature (London) 225*, 1245.
Wigle, D. T., and Smith, A. E. (1973), *Nature (London), New Biol. 242*, 136.
Williamson, R., and Brownlee, G. G. (1969), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 3*, 306.
Williamson, R., Morrison, M., Lanyon, G., Eason, R., and Paul, J. (1971), *Biochemistry 10*, 3014.
Yogo, Y., and Wimmer, E. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 1877.

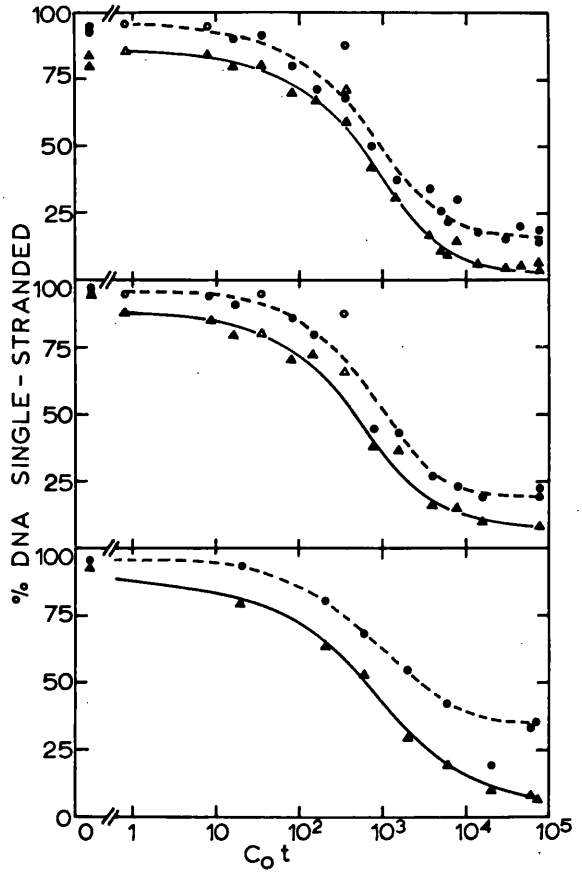


Figure 1. Annealing of mouse globin cDNA to mouse DNA in the presence of a great excess of the latter. (*Top*) Total embryo DNA/cDNA = 10^7 (w/w); (*middle*) fetal liver DNA/cDNA = 10^7 (w/w); (*bottom*) sperm DNA/cDNA = 5×10^6 (w/w). Total DNA shown by continuous lines and triangles; cDNA shown by interrupted lines and circles. General conditions for preparation of DNA and cDNA and estimation of single- and double-stranded DNA were as described in Harrison et al. (1972).

incubating [^{3}H]deoxynucleoside triphosphates with globin mRNA in the presence of AMV reverse transcriptase. The titrations were performed by incubating a standard amount of cDNA with increasing amounts of RNA in conditions that should lead to complete hybridization; the amount of cDNA hybridized was then estimated by measuring the fraction that resisted digestion by S1 nuclease, which attacks single-stranded but not double-stranded DNA (Sutton, 1971). The results of typical titrations are shown in Figure 2, from which it can be seen that the RNA isolated from the incubation mixture containing chromatin from erythropoietic tissue included sequences complementary to the cDNA, whereas the RNA isolated from the mixture containing brain chromatin did not.

Chromatin always contains some RNA. Therefore, the possibility always exists, especially when minute amounts of transcripts from the globin genes are being detected, that these could arise entirely from RNA originally present in the chromatin. To investigate this possibility we performed the following experiments. First, RNA was isolated from fetal liver chromatin and hybridized with cDNA; this was not found to contain detectable amounts of globin mRNA in concentrations equivalent to those used in the experiments illustrated in Figure 2. Second, control incubations, complete except for the omission of GTP, did not yield RNA containing measurable amounts of globin mRNA in similar concentrations. Finally, to demonstrate that the RNA which formed a hybrid with globin cDNA was actually made in the reaction, an incubation mixture was prepared with chromatin from fetal liver as before, but ^{32}P-labeled ATP was included. After hybridization of the purified RNA with cDNA, the hybrid was isolated and banded in a cesium gradient (Fig. 3). This revealed cDNA with the same buoyant density as an RNA-DNA hybrid between globin mRNA and its cDNA (1.78); associated with this were ^{32}P counts. The most likely explanation of this finding was the formation of a hybrid between globin cDNA and the RNA made in the reaction. An alternative, though exceedingly unlikely, explanation was the formation of a complex between poly(A) and cDNA. We excluded this possibility by hybridizing poly(A) (mol. wt. 100,000) to cDNA and isolating and banding the nucleic acids from the reaction mixture as described in Figure 3. Most of the cDNA then banded at 1.71; there was no evidence for the formation of a complex of significantly higher buoyant density. It can therefore be concluded that cDNA was indeed hybridized to RNA that had been made during the incubation. The relative amounts of labeled and unlabeled RNA in the hybrid could be calculated at approximately 4:1; hence, about 80% of the globin mRNA sequences isolated from the incubation mixture appear to be present in newly synthesized RNA. These experiments show that RNA is transcribed from globin genes in fetal liver chromatin but not in brain chromatin, thereby substantiating the conclusion we had previously drawn that there is an organ-specific "masking" of DNA in chromatin.

In previous work we showed that the specificity of chromatin templates could be altered by transferring the nonhistone components (Gilmour and Paul, 1970). We therefore explored this phenomenon in the present series of experiments. As shown in Figure 4, when brain chromatin was dissociated and reassociated in the presence of nonhistone proteins from fetal liver, evidence was obtained that this gave a nucleoprotein from which globin mRNA sequences could be transcribed, although

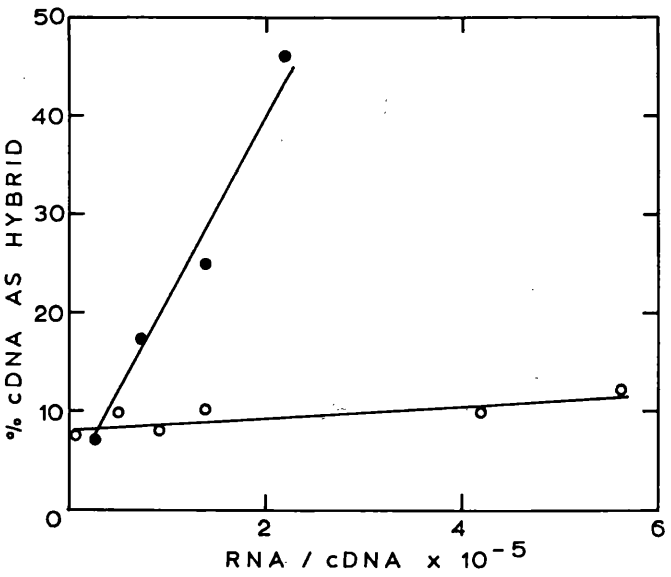


Figure 2. Titrations of chromatin-primed RNA against globin cDNA. (●——●) Mouse fetal liver chromatin; (○——○) mouse brain chromatin. Globin mRNAwas purified from mouse reticulocytes by passing polysomal RNA twice through a poly(U)-Sepharose column; its activity was measured in a duck reticulocyte lysate system. cDNA was prepared as described by Harrison et al. (1972), with a specific activity of 20 × 10^6 dpm/μg. Chromatin was prepared from nuclei isolated in 0.25 M sucrose by successive washing with 10 vol 0.1 M Tris pH 7.4 and 10 vol 0.14 M NaCl. This was incubated with *E. coli* RNA polymerase and nucleoside triphosphates using incubation conditions previously described (Paul and More, 1972), and the RNA was isolated by the method of Penman (1966). Different amounts of this RNA were mixed with 0.001 μg cDNA in 10 μl 50% formamide, 0.5 M NaCl, 0.025 M HEPES pH 7, 0.01 M EDTA at 43°C for 3–4 days. The hybrid was assayed by measuring the amount of cDNA resistant to the S1 nuclease described by Sutton (1971).

reconstituted brain chromatin did not by itself exhibit this property. This provides an indication that the nonhistone proteins of fetal liver can "unmask" the globin genes in brain chromatin. Moreover, by experiments similar to those described above evidence was obtained that the hybrid was composed of cDNA and newly made RNA. Hence, an element of the transcriptional control in erythroid cells would seem to be of the "gene-masking" type. We have previously presented evidence that the specific regulation of RNA synthesis in chromatin is due to nonhistone proteins (Paul and

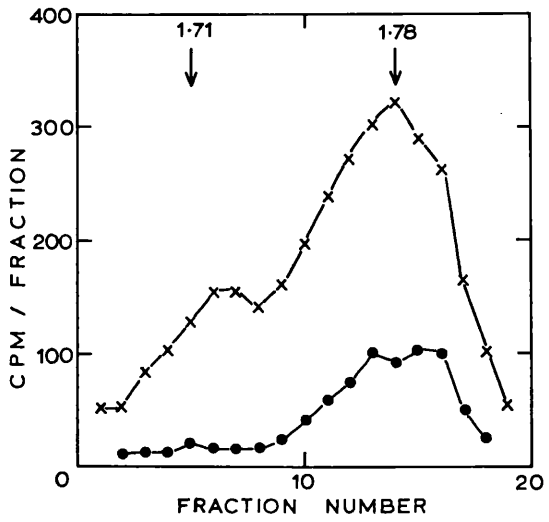


Figure 3. Isopycnic banding in CsCl of a hybrid formed between [^{3}H]cDNA and [^{32}P]RNA transcribed from mouse fetal chromatin (x——x) ^{3}H counts; (●——●) ^{32}P counts. Chromatin was prepared and RNA transcribed from it as described in the legend to Figure 2, except that [^{32}P]ATP (1 Ci/mmole) was included in the incubation mixture. After hybridization with cDNA, the reaction mixture was diluted to 1 ml with 0.1 M NaCl and treated with 20 μg/ml ribonuclease for 1 hr at 37°C. It was then passed through a Sephadex G-75 column in 0.1 M NaCl and the excluded material centrifuged to equilibrium in a CsCl gradient. The fractions were collected and analyzed.

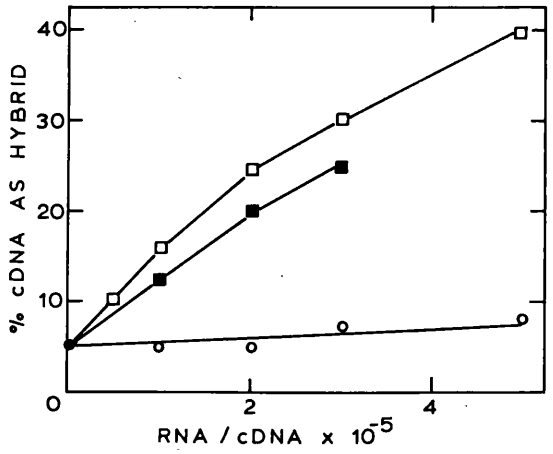


Figure 4. Titration of globin cDNA against RNA transcribed from reconstituted chromatin. (○——○) Reconstituted brain chromatin; (□——□) reconstituted liver chromatin; (■——■) brain chromatin reconstituted in the presence of nonhistone proteins from fetal liver. Reconstitution was performed by taking up the chromatin or other components in 5 M urea–2 M NaCl (which produces dissociation) and then progressively lowering the concentrations of urea and salt (Gilmour and Paul, 1970). Nonhistone proteins from fetal liver were obtained by fractionation of chromatin on hydroxyapatite (MacGillivray et al., 1971). These were mixed with brain chromatin in the ratio 1 NHP:2 DNA (w/w). RNA was transcribed, isolated, and hybridized as in the legend to Figure 2.

Table 1. Concentration of Globin mRNA Sequences in RNA Pools in Friend Erythroleukemic Cells Induced or Not Induced to Synthesize Hemoglobin

	Fraction of RNA as Globin mRNA × 10^5			Fraction of Cells Hemoglobinized (%)
	Polysomal RNA	Cytosol RNA	Total Nuclear RNA	
Uninduced	3–5	4	18	<1
Induced	50–70	15–18	20	50–70

Friend erythroleukemic cells, clone 707, were grown in suspension in Ham's F12 medium and MEM nonessential amino acids + 16.5% horse serum. DMSO was added to the induced set of cultures to 1.5% v/v. Samples were harvested at intervals. The table shows findings from cells grown for 4 days in these conditions. The fraction of cells hemoglobinized was determined by benzidine staining. The cells were fractionated and RNA extracted from the fractions. This was titrated against cDNA in the manner outlined in the legend to Figure 2, in comparison with an authentic sample of purified mouse globin mRNA.

Gilmour, 1968; Gilmour and Paul, 1970) and a considerable amount of evidence has now accumulated in support of this view (MacGillivray et al., 1972). These preliminary experiments not only reinforce earlier work but offer a technical approach of considerable power.

The Transcription Product in the Cell

The speculation that high molecular weight RNA might be a precursor of mRNA has now been tested. Williamson et al. (1973) injected high molecular weight RNA (>30 S) from mouse fetal liver into Xenopus oocytes and demonstrated the synthesis of mouse α- and β-globin chains. Imaizumi et al. (1973) have more recently shown that nuclear HnRNA contains globin mRNA sequences by hybridization with cDNA, and we have made similar observations (Table 1). This result has been anticipated for some time, but an important piece of information likely to be obtained from current studies is the size of the largest molecule containing globin mRNA sequences, since this may set a lower limit on the size of the transcriptional unit. This is probably at least 50 S, equivalent to a molecular weight of several million. There is a possibility that some of these very large molecules may be aggregates. They could even be formed by the ligation of smaller molecules. However, present evidence suggests, on the contrary, that the larger molecules are precursors of the smaller. It seems very likely, therefore, that the transcriptional unit is very much larger than the sequences coding for globin chains.

Globin mRNA Synthesis and Processing during Induction of Hemoglobin Synthesis

Friend and her colleagues have demonstrated that lines of cells isolated from DBA/2 mice, rendered leukemic by inoculation with Friend virus, are capable of hemoglobin synthesis (Scher et al., 1971). Moreover, when these cells are treated with 1.5–2% dimethyl sulfoxide (DMSO), they can be induced to form hemoglobin at a much higher rate (Friend et al., 1971). Ostertag et al. (1972) and Ross et al. (1972a) have made similar observations, and the latter group has demonstrated that the increased hemoglobin synthesis is accompanied by accumulation of globin mRNA in the treated cells.

In studies with this system, we have measured separately the concentrations of mRNA sequences in nuclear, cytosol, and polysomal RNA. As summarized in Table 1, we found that the concentration of mRNA sequences was not very different in the nucleus and only moderately greater in the cytosol in induced cells in comparison with uninduced ones, but that there was a dramatic increase in polysomes. These findings strongly suggest the intervention of a factor (a translation initiation factor?) which results in incorporation of globin mRNA into polysomes.

Chromatin from cells induced to synthesize hemoglobin was examined for its capacity to act as a template for synthesis of globin mRNA. As shown in Figure 5, evidence could be obtained for new synthesis of RNA, which formed a hybrid molecule with cDNA. (This hybrid had a slightly lower buoyant density than might have been expected of a pure cDNA/mRNA hybrid.) At the time of writing, experiments to compare chromatin from induced and uninduced cells are incomplete. Hence, although evidence for translational control of hemoglobin synthesis in these cells is good, it is not yet possible to state whether there is also some control at the level of transcription. However, taken together with the evidence outlined earlier,

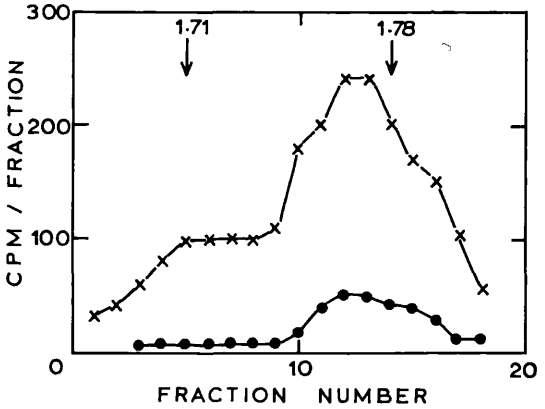


Figure 5. Isopycnic banding in CsCl of hybrid formed between [^{3}H]cDNA and [^{32}P]RNA transcribed from chromatin from induced Friend cells. (x——x) ^{3}H counts; (●——●) ^{32}P counts. Cell cultures were prepared as described in Table 1 and [^{32}P]RNA prepared from chromatin, hybridized to cDNA, and banded as described in the legend to Figure 3.

that the globin gene can be transcribed from fetal liver chromatin but not brain chromatin, it seems likely that both transcriptional and translational controls are involved in the differentiation and maturation of erythroid cells.

From our own findings and those of other investigators the following points are reasonably well established. First, in the mouse the genes for α- and β-globin chains are represented a very small number of times, probably once, in the genome. Second, the initial transcriptional product is of much greater molecular weight (possibly $>10^7$) than globin mRNA (220,000); hence, the transcriptional units are presumably much bigger than the regions coding for globin chains. Third, the globin genes are accessible to bacterial RNA polymerase in chromatin from erythropoietic tissue but not in chromatin from brain. Finally, nonhistone factors, probably proteins, seem to be instrumental in determining this accessibility.

These findings are generally compatible with Paul's (1972) model of the transcriptional unit. This proposes that, possibly for physicochemical reasons, there is a constraint on the minimum size of the transcriptional unit in chromatin, and that where a very small number of functional gene copies is demanded, most of the DNA in a transcriptional unit may be "nonsense DNA," in that it conveys no information in its nucleotide sequences. It also proposes that nonhistone proteins cause a partial unwinding of nucleohistone which, in their absence, probably forms a highly compact, supercoiled structure. Special DNA configurations, termed "address sites," that serve as binding sites for nonhistone proteins are envisaged as being adjacent to or a component of polymerase binding sites; they correspond roughly to promoter loci. In the original model no proposal was made about specificity of address sites and the corresponding proteins. However, this is clearly called for by the findings in these and other experiments. If the model has any validity, it also may be necessary to postulate a secondary modification of the nucleoprotein in the address sites. This could occur by an allosteric modification of the protein bound there, which in turn could be achieved in a number of ways, including interaction with other proteins such as effector binding proteins.

If specificity were to be ascribed to "address sites," it would seem likely that the same address sites might be common to several genes, for example, the α- and β-globin genes, to achieve coordination. Moreover, each gene might obviously have several different address sites, so that different sets of genes with common components (e.g., α- and β-globin in the adult and α- and γ-globin in the fetus) could be operated independently.

Some of the evidence presented by other speakers in this symposium strongly suggests that unique and moderately repetitive DNA sequences alternate within chromomeres and that the moderately repetitive sequences represent clusters of different sequences. It is perhaps worth pointing out that these findings, too, are clearly compatible with Paul's model, although, of course, both they and the findings which we have described above fall far short of providing strong evidence for it.

Acknowledgments

We wish to acknowledge the help of our colleagues, especially Drs. R. Williamson, A. J. MacGillivray, and D. Rickwood for collaboration and discussions and Mr. A. C. McKirdy, Miss Elizabeth McPhail, and Miss Claire Hale for skilled technical assistance. We are very grateful to Dr. Charlotte Friend of the Mount Sinai Hospital, N.Y. for providing us with inocula of the cells used in the experiments and to Dr. J. W. Beard of Duke Univeristy, N.C. for supplies of avian myeloblastosis virus. The work was supported by grants from the Medical Council and Cancer Research Campaign.

References

Benz, E. J. and B. G. Forget. 1971. Defect in messenger RNA for human haemoglobin synthesis in beta-thalassemia. *J. Clin. Invest.* **50:** 2755.

Bishop, J. and M. Rosbash. 1973. Reiteration frequency of duck haemoglobin genes. *Nature New Biol.* **241:** 204.

Bishop, J. O., R. Pemberton, and C. Baglioni. 1972. Reiteration frequency of haemoglobin genes in the duck. *Nature New Biol.* **235:** 231.

Chantrenne, H., A. Burny, and G. Marbaix. 1967. The search for the messenger RNA of haemoglobin. *Progr. Nucleic Acid Res. Mol. Biol.* **7:** 173.

Friend, C., W. Scher, J. G. Holland, and T. Sato. 1971. Hemoglobin synthesis in murine virus-induced leukemic cells *in vitro*: Stimulation of erythroid differentiation by dimethyl sulfoxide. *Proc. Nat. Acad. Sci.* **68:** 378.

Gilmour, R. S. and J. Paul. 1969. RNA transcribed from reconstituted nucleoprotein is similar to natural RNA. *J. Mol. Biol.* **40:** 137.

———. 1970. Role of non-histone components in determining organ specificity of rabbit chromatin. *FEBS Letters* **9:** 242.

Harrison, P. R., A. Hell, G. D. Birnie, and J. Paul. 1972. Evidence for single copies of globin genes in the mouse genome. *Nature* **239:** 219.

Imaizumi, T., H. Diggelman, and K. Scherrer. 1973. Demonstration of globin messenger sequences in giant nuclear precursors of messenger RNA of avian erythroblasts. *Proc. Nat. Acad. Sci.* **70:** 1122.

Kacian, D. L. and S. Spiegelman. 1972. *In vitro* synthesis of DNA components of human genes for globins. *Nature New Biol.* **235:** 167.

Lingrel, J. B., R. E. Lockard, R. F. Jones, H. E. Burr, and J. W. Holder. 1971. Biologically active messenger-RNA for haemoglobin. *Series Haemat.* **4:** 337.

MacGillivray, A. J., D. Carroll, and J. Paul. 1971. The heterogeneity of the non-histone chromatin proteins from mouse tissues. *FEBS Letters* **13**: 204.

MacGillivray, A. J., J. Paul, and G. Threlfall. 1972. Transcriptional regulation in eukaryotic cells. *Advanc. Cancer Res.* **15**: 93.

Morrison, M. R., J. Paul, and R. Williamson. 1972. The DNA-RNA hybridization of ^{3}H-methylated mouse-globin mRNA in conditions of RNA excess. *Eur. J. Biochem.* **27**: 1.

Ostertag, W., H. Melderis, G. Steinheider, N. Kluge, and S. Dube. 1972. Synthesis of mouse haemoglobin and globin mRNA in leukaemic cell cultures. *Nature New Biol.* **239**: 231.

Paul, J. 1972. A general theory of chromosome structure and gene activation in eukaryotes. *Nature* **238**: 444.

Paul, J. and R. S. Gilmour. 1968. Organ-specific restriction of transcription in mammalian chromatin. *J. Mol. Biol.* **34**: 305.

Paul, J. and I. A. R. More. 1972. Properties of reconstituted chromatin and nucleohistone complexes. *Nature New Biol.* **239**: 134.

Penman, S. 1966. RNA metabolism in the HeLa cell nucleus. *J. Mol. Biol.* **17**: 117.

Ross, J., Y. Ikawa, and P. Leder. 1972a. Globin messenger-RNA induction during erythroid differentiation of cultured leukemia cells. *Proc. Nat. Acad. Sci.* **69**: 3620.

Ross, J., H. Aviv, E. Scolnick, and P. Leder. 1972b. *In vitro* synthesis of DNA complementary to purified rabbit globin mRNA. *Proc. Nat. Acad. Sci.* **69**: 264.

Schapira, G., J. C. Dreyfus, and N. Maleknia. 1968. The ambiguities in rabbit haemoglobin; evidence for a messenger RNA translated specifically into haemoglobin. *Biochem. Biophys. Res. Commun.* **32**: 558.

Scher, W., J. G. Holland, and C. Friend. 1971. Haemoglobin synthesis in murine virus-induced leukaemic cells *in vitro*. I. Partial purification and identification of haemoglobin. *Blood* **37**: 428.

Sutton, W. D. 1971. A crude nuclease preparation suitable for use in DNA reassociation experiments. *Biochim. Biophys. Acta* **240**: 552.

Suzuki, Y., L. P. Gage, and D. D. Brown. 1972. Genes for silk fibroin in *Bombyx mori*. *J. Mol. Biol.* **70**: 637.

Verma, I. M., G. F. Temple, H. Fan, and D. Baltimore. 1972. *In vitro* synthesis of DNA complementary to rabbit reticulocyte 10S RNA. *Nature New Biol.* **235**: 163.

Williamson, R., C. E. Drewienkiewicz, and J. Paul. 1973. Globin messenger sequences in high molecular weight RNA from embryonic mouse liver. *Nature New Biol.* **241**: 66.

Williamson, R., M. Morrison, G. Lanyon, R. Eason, and J. Paul. 1971. Properties of mouse globin messenger RNA and its preparation in milligram quantities. *Biochemistry* **10**: 3014.

THE TRANSLATION OF MOUSE GLOBIN mRNA FROM WHICH THE POLYADENYLIC ACID SEQUENCE HAS BEEN REMOVED IN A REINITIATING PROTEIN SYNTHESIS SYSTEM

by

Stephen Humphries [1], Michael Doel [2] and Robert Williamson [1]

1. The Beatson Institute for Cancer Research, 132 Hill Street, Glasgow G3 6UD, Scotland.
2. Searle Research Laboratories, High Wycombe, Bucks, England.

Received April 9,1974

SUMMARY: The translation of mouse globin messenger RNA from which the polyA sequence had been removed enzymatically was compared with the translation of control globin mRNA in a rabbit reticulocyte lysate cell-free protein synthesis system. Re-initiation of synthesis occurs on average eleven times for mouse β-globin, demonstrating that the polyA sequence is not required for repeated initiation on the same mRNA molecule.

Recent work from this laboratory with mouse globin mRNA (1) and by others using total L cell mRNA (2) has demonstrated that the polyA sequence found on the 3'-terminus of most eukaryotic mRNAs is not required for the initiation of at least a single round of protein synthesis in a heterologous cell-free system. However, neither of these results definitively demonstrated reinitiation of protein synthesis on deadenylated mRNA. The experiment described below shows that reinitiation is possible on an mRNA from which the polyA sequence has been removed. This was done using a heterologous system derived from rabbit reticulocytes, which is known to translate added mRNA many times during a 90 minute period of incubation (3).

MATERIALS AND METHODS

Polysomal mouse globin messenger RNA was prepared using oligo-dT cellulose (Searle Ltd., High Wycombe, Bucks.) and incubated with polynucleotide phosphorylase to remove the polyA sequence as described previously (1). After this treat-

ment, between 60-80% of the mRNA was no longer retained by oligo-dT cellulose. This deadenylated mRNA is free of 3'-terminal polyA sequences as judged by non-retention by oligo-dT cellulose, fingerprint analysis, template ability with oligo-dT primer and reverse transcriptase, and molecular weight reduction on polyacrylamide gel electrophoresis (1). The RNA retained by the column is identical in sequence and behaviour to untreated mRNA and was used as a control in the protein synthesis experiments.

The activity of the mRNA in directing mouse globin synthesis was measured using a rabbit lysate system as described by Palmiter (3). Assays were carried out in a reaction mixture of 250 microlitres containing 2.5 microcuries of ^{3}H-isoleucine (30 Ci/mM, Radiochemical Centre, Amersham). Between 2-6 picomoles of globin mRNA were added to an assay mixture, which is known to be in the linear response range of the system for added mRNA.

After incubation for 6 or 90 minutes at 26°, non-radioactive mouse globin was added as carrier, and total globin extracted and chromatographed on carboxymethyl-cellulose with a 0.01M - 0.05M phosphate gradient as described by Lingrel *et al.* (4).

RESULTS

Figure 1 shows a typical separation of mouse β-globin from mouse α-globin and rabbit globin after 6 minutes incubation. Palmiter has shown that for this system 1.6% of the isoleucine incorporated is radioactive under these conditions (3). Mouse β-globin contains twice as much isoleucine as rabbit globin and it can be calculated that an incorporation of approximately 2,000 dpm represents one picomole of rabbit globin and 0.5 picomole of mouse β-globin. The amounts of mouse β-globin made in each assay is shown in Table 1.

After six minutes the deadenylated mouse mRNA has been translated twice per molecule on average and the control mRNA 2.8 times per molecule on average; after 90 minutes the deadenyl-

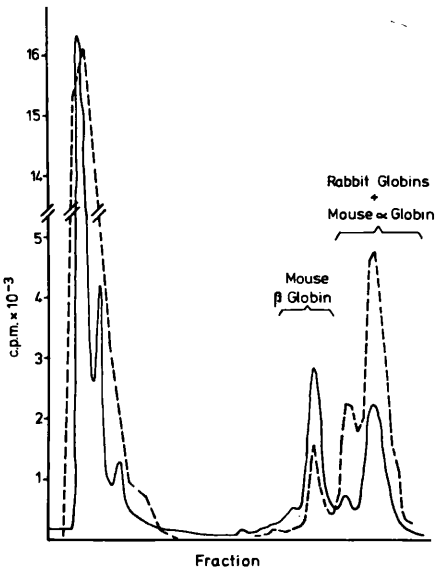


Figure 1. Product analysis of globin chains synthesised in a rabbit reticulocyte lysate cell-free system with added mouse globin mRNA. 2.3 picomoles control mouse globin mRNA were incubated in the system for 6 min at 26°. The mix was cooled to 0° and 40mg unlabelled mouse globin was added. Globin was extracted with acid-acetone and chromatographed on carboxymethylcellulose. The absorbance was monitored at 280 nm (——). Five ml fractions were precipitated onto glass fibre discs, solubilised with NCS (Nuclear Chicago Ltd.) and counted in toluene-based scintillator in a Packard scintillation counter at an efficiency of 48% (– – – –).

ated mRNA has been translated 11.2 times on average and the control mRNA 19.2 times on average. As previously reported, the deadenylated mRNA consistently shows only 60-70% the activity of the control mRNA. However, this figure is similar at 6 and 90 minutes.

DISCUSSION

The amount of mouse globin synthesis obtained with added mouse globin mRNA is that expected for this system, as Palmiter has previously shown a 'transit time' of approximately three minutes for translation of a globin polypeptide chain (3). The somewhat lower globin synthesis

Table 1. The amount of mouse β-globin synthesised at 6 and 90 minutes in a rabbit reticulocyte cell free system.

Added mRNA	Time	pMoles mouse β-globin made	pMoles mRNA added	No. of times each message translated
deadenylated	6 min.	3.81	5.2	1.9
control	6 min.	5.87	4.6	2.8
deadenylated	90 min.	14.01	2.6	11.2
control	90 min.	20.16	2.3	19.2

The picomoles of mRNA added was calculated from the molecular weight of the control (220,000) and deadenylated (185,000) RNAs as determined by gel electrophoresis (1). It was assumed that in both cases 50% of the added mRNA was β-globin message. After 6 minutes each rabbit message had been translated about 10 times, and after 90 minutes about 40 times. The figures in the last column have been corrected for small differences (±10%) in the amount of rabbit globin made in each assay.

directed by the deadenylated as compared to the control mRNA may reflect an inherent effect due to the removal of the polyA sequence, but we feel it is more likely to reflect contamination of the deadenylated mRNA by small fragments of RNA broken by contaminating endonucleases, which contribute to the optical density of the sample but cannot contribute to polypeptide synthesis.

These experiments and others (1,2) demonstrate that the polyA sequence is not obligatory for initiation, reinitiation, translation or termination. However, it is known that the length of the polyA sequence of total HeLa cell cytoplasmic mRNA decreases with time during *in vivo* translation (5). It is therefore likely that polyA is involved in the determination of the lifetime of a messenger RNA in the cytoplasm.

Acknowledgments: This work was carried out at Searle Research Laboratories, High Wycombe, Bucks. The authors would like to thank Dr. Norman Carey for his co-operation in arranging this collaboration and Dr. Mike Getz and Mrs. Jenny Crossley for helpful discussions. The Beatson Institute is supported by the Medical Research Council and the Cancer Research Campaign.

REFERENCES:

1. Williamson, R., Crossley, J. and Humphries, S. (1974) Biochemistry 13, 703-707.
2. Bard, E., Efron, O., Marcus, A. and Perry, R.P. (1974) Cell 1, 101-106.
3. Palmiter, R.D. (1973). J. Biol. Chem. 248, 2095-2106.
4. Lingrel, J.B., Lockard, R.E., Jones, R.F., Burr, H.E. and Holder, J.W. (1971) in: Series Haematologica, Vol. 4, Hemoglobin Synthesis, K.G. Jenson and S. Killman, Eds., Munksgaard, Copenhagen, p.p. 37-69.
5. Sheiness, D. and Darnell, J.E. (1973) Nature New. Biol. 241, 265-268.

Reprinted from *J. Mol. Biol.* (1974) **84**, 539–554

Kinetic Studies of Gene Frequency

I. Use of a DNA Copy of Reticulocyte 9 S RNA to Estimate Globin Gene Dosage in Mouse Tissues

P. R. Harrison, G. D. Birnie, Anna Hell, S. Humphries
B. D. Young and J. Paul

J. Mol. Biol. (1974) 84, 539–554

Kinetic Studies of Gene Frequency

I. Use of a DNA Copy of Reticulocyte 9 S RNA to Estimate Globin Gene Dosage in Mouse Tissues

P. R. Harrison, G. D. Birnie, Anna Hell, S. Humphries
B. D. Young and J. Paul

The Beatson Institute for Cancer Research
132 Hill Street, Glasgow, G3 6UD, Scotland

(*Received 4 May 1973, and in revised form 10 January 1974*)

This and the accompanying paper justify rigorously the use of a reverse transcriptase-prepared complementary DNA copy of mouse reticulocyte 9 S RNA to estimate globin gene dosage in mouse tissues. It is demonstrated that globin cDNA† is a faithful, but partial, transcript of reticulocyte 9 S RNA, which contains both α and β globin messenger RNAs. Detailed calculations suggest that, at most, 2% base errors are introduced into globin cDNA during transcription of 9 S RNA *in vitro*.

Globin cDNA is used to estimate the reiteration frequency of the globin genes by annealing to germ-line or somatic mouse DNA obtained under conditions that extract total DNA, including any small gene-size fragments (e.g. episomal gene copies). After fragmentation to about 120 bases, the bulk of the mouse DNA reassociates at the rate predicted for unique sequences, by comparison with the rate of reassociation of *Escherichia coli* DNA under the same conditions. With somewhat larger fragments (180 to 330 bases), the bulk of the mouse DNA sequences reassociate as if they were repeated four to five times. We interpret these data in terms of intimate interspersion of repetitive and unique sequences.

The rate of annealing of globin cDNA to mouse DNA is the same as that of reannealing of the bulk of the mouse DNA sequences, with fragments with a mean size of 180 or 330 bases. This is true whether germ-line DNA (sperm), predominantly erythroid DNA (14-day foetal liver) or total embryo DNA is studied. Neither cDNA nor the mouse DNA fragments are degraded during annealing; and only a slight correction for the effect of mismatching on the rate of annealing of cDNA to mouse DNA is necessary. Three conclusions therefore follow: (1) there are probably only single, but at most four to five, copies of the globin genes in the mouse germ line; (2) there is no globin gene amplification in mouse erythroid cells; and (3) there is no widespread elimination of globin genes in non-erythroid somatic mouse tissues.

1. Introduction

It is of crucial importance in biology to determine the extent of reiteration of specific genes in eukaryotes. Whilst genetic evidence in general argues for a single copy of each gene, the size of the haploid genome in eukaryotes is very much larger than could reasonably be expected if it consisted entirely of single copies of genes (e.g. King & Jukes, 1969; Ohno, 1971; Ohta & Kimura, 1971; O'Brien, 1973; Beermann,

† Abbreviation used: cDNA, complementary DNA.

1972). In fact, the large DNA complement of eukaryotes consists of a diversity of sequences, some of which are reiterated hundreds or thousands of times, but the majority of which are unique. By annealing very highly labelled RNA molecules of specific types with excess eukaryotic DNA (Melli *et al.*, 1971), estimates of the reiteration frequencies of specific genes have been possible. Some genes, for example those specifying the immunoglobulin light chain (Delovitch & Baglioni, 1973), histones (Kedes & Birnstiel, 1971; Weinberg *et al.* 1972) and ribosomal RNA (Birnstiel *et al.*, 1969) are reiterated tens, hundreds or thousands of times. However, fibroin genes (Suzuki *et al.*, 1972) and globin genes (Bishop *et al.*, 1972) are reiterated infrequently, if at all. Similar studies have shown that in amphibians, amplification of ribosomal genes occurs specifically in the oocyte (Attardi & Amaldi, 1970; Birnstiel *et al.*, 1971). On the other hand, in the silkworm not only the fibroin genes but the whole of the genome is amplified during polytenization (Suzuki *et al.*, 1972). However, no amplification of immunoglobulin light chain genes has been observed in antibody-producing cells (Delovitch & Baglioni, 1973).

It has been well understood for some time that hybridization of labelled RNA molecules in DNA-excess experiments of this type involves certain theoretical and technical problems (Bishop, 1972*a*). Therefore, in a previous report (Harrison *et al.*, 1972*a*) we extended this approach by using reverse transcriptase to prepare a single-stranded DNA copy of mouse reticulocyte 9 S RNA (which contains α and β globin mRNAs) and then determined the reiteration frequency of the globin genes by annealing cDNA† to excess mouse DNA in the usual way. Other workers have subsequently used the same approach to determine the reiteration frequency of the globin genes in the duck (Packman *et al.*, 1972; Bishop & Rosbash, 1973). The present papers justify rigorously the general validity of conclusions drawn from the use of cDNA to determine gene frequency, with specific reference to globin genes in the mouse. The biological conclusions that emerge from this work are, that there are very few (probably single but not more than about four) copies of the globin gene either in the mouse germ line, or in erythroid cells; and that the globin genes are not eliminated during differentiation of non-erythroid somatic tissues.

2. Materials and Methods

(a) *Reverse transcriptase*

Reverse transcriptase was isolated from avian myeloblastosis virus as described by Kacian *et al.* (1971). The enzyme solution obtained had an activity of 100 units/ml (Kacian *et al.*, 1972).

(b) *Preparation of DNA*

Mouse embryo, foetal mouse liver and *Escherichia coli* DNAs were isolated by the hydroxyapatite method described previously (Hell *et al.*, 1972). Mouse sperm were obtained as described previously (Borenfreund *et al.*, 1961), and sperm DNA was prepared similarly. Isolated DNAs (dissolved in 0·1 M-NaCl at about 2·5 mg/ml) were sonicated in 5-ml portions in a Dawe sonicator using the microprobe (6, 10-s treatments punctuated by 30-s cooling periods at 0°C) or fragmented by treatment with deoxyribonuclease I (Hell *et al.*, 1972). After precipitation with alcohol, the DNA fragments were fractionated on an alkaline sucrose gradient in an MSE BXIV zonal rotor or a 3×25 ml rotor, and fragments of the appropriate size were obtained and purified as described previously (Hell *et al.*, 1972). Sizes of DNA fragments were estimated by sedimentation in 5% to 10% (w/w) alkaline sucrose gradients (Hell, 1972). The $s_{20,w}$ values (McEwen, 1967) and

† Abbreviation used: cDNA, complementary DNA.

molecular weights (Studier, 1965) were calculated. A marker DNA ($s_{20,w} = 4{\cdot}3$) was included in each rotor load. The sedimentation coefficient of this marker DNA has also been determined independently by sedimentation in the analytical ultracentrifuge by Dr Ailsa Campbell of the Department of Biochemistry, University of Glasgow. Both estimates of the size of the marker agreed.

E. coli ^{14}C-labelled DNA was isolated from a thymine-less strain of *E. coli* (NC1B 8583) grown in [^{14}C]thymine (56 μCi/μmol).

(c) *Preparation of 9 S RNA*

9 S RNA was isolated from mouse reticulocytes, either from EDTA-treated polysomes (Williamson *et al.*, 1971) or from total polysomal RNA by the poly(U)- Sepharose method (Adesnik *et al.*, 1972); it was tested for biological activity in the duck lysate (Lingrel, 1972) or oocyte (Lane *et al.*, 1971) systems. Both preparations of 9 S RNA ran as single bands on either 2·6% or 6% polyacrylamide gels (G. Lanyon, personal communication).

(d) *Preparation of ^{3}H-labelled cDNA and ^{3}H-labelled poly(dT)*

cDNA was prepared as described previously (Harrison *et al.*, 1972*a*,*b*) by incubating for 90 min at 37°C the following mixture: 40 μM-d[^{3}H]TTP (8 to 10 Ci/mmol); 40 μM-d[^{3}H]-CTP (8 to 10 Ci/mmol); 500 μM-dATP; 500 μM-dGTP; 20 μg actinomycin D/ml; 1 to 2 μg oligo(dT)$_{12-18}$/ml (PL Biochemicals); 5 to 10 μg 9 S RNA/ml; 100 μg catalase/ml; 50 mM-Tris (pH 8·2); 50 mM-KCl; 10 mM-dithiothreitol; 5 mM-magnesium acetate; 0·2 vol. reverse transcriptase in 50% glycerol; 0·15 M-potassium phosphate (pH 8·0). The specific activities of d[^{3}H]TTP and d[^{3}H]CTP were always adjusted to be equal. cDNA was fractionated and purified as described previously (Harrison *et al.*, 1972*a*), and fractions in the desired range of molecular weight were isolated. Amounts of cDNA were calculated on the basis of the quoted specific activities of the triphosphates, assuming that cDNA consists of 50% (C + T).

^{3}H-labelled poly(dT) was prepared in the same way as ^{3}H-labelled cDNA, except that poly(A) (Miles-Seravac) replaced 9 S RNA, and 500 μM-dCTP (unlabelled) replaced the ^{3}H-labelled dCTP. The resultant poly(dT) had a mean mol. wt of about 10,000 to 20,000.

(e) *Preparation of RNA–DNA hybrids of known (G + C) contents*

These were prepared by transcribing heat-denatured *E. coli* or mouse DNA with RNA polymerase from *E. coli*, followed by treatment with single-strand-specific nuclease from *Neurospora crassa*, as described previously (Birnie, 1972).

(f) *RNA–DNA hybridization*

cDNA was mixed with the appropriate amount of RNA and incubated, usually at 43°C, in hybridization buffer, either (a) 0·3 M-sodium phosphate (pH 6·8), 50% formamide (Fluka); or (b) 0·5 M-NaCl, 25 mM-HEPES, 0·5 mM-EDTA (pH 6·8), 50% formamide. The salt solutions (before addition of formamide) were all passed through Chelex-100 resin (Bio-Rad Laboratories), then treated with diethylpyrocarbonate (Baycovin, a gift from Bayer Chemicals, Ltd, Glasgow) and autoclaved. Usually the hybridization solutions also contained 500 μg *E. coli* RNA/ml, which had been passed through Chelex. Where appropriate, the hybridization mixtures were treated with pancreatic ribonuclease (20 μg/ml; Worthington) for 1 h at 20°C, after dilution about 20 times with 0·15 or 0·3 M-NaCl.

(g) *Annealing of complementary DNA or poly(dT) to DNA and analysis on hydroxyapatite*

The appropriate amounts of cDNA or poly(dT) and mouse or *E. coli* DNA (at about 2 mg/ml) were mixed in water, lyophilized and redissolved at 10 mg/ml before adding 0·2 vol. 0·72 M-sodium phosphate (pH 6·8). DNA concentrations were calculated on the basis of absorption measurements on acid-hydrolysed portions of the dilute DNA solutions. The cDNA/DNA mixtures were annealed for the appropriate times at 60°C (either undiluted or 100-fold diluted) and analysed on hydroxyapatite (Bio-Rad, Biogel HTP batch 10936, using a ratio of 1 mg DNA to 3 ml packed hydroxyapatite) essentially as described previously (Harrison *et al.*, 1972*a*). With this batch of hydroxyapatite, the best resolution

of native and denatured DNA at 60°C was obtained by eluting the single-stranded material with 0·16 M-phosphate. At this phosphate concentration, about 85% of heat-denatured mouse DNA (size 330 bases), but only 5% of native mouse DNA of the same size, was eluted. Therefore, the 3-ml hydroxyapatite columns were eluted with 8 ml of 0·03 M-phosphate, twice with 8 ml of 0·16 M-phosphate, and finally twice with 4 ml of 0·4 M-phosphate. Absorbance measurements of the fractions of eluate were made, before counting in Instagel (Packard, Ltd).

Measurements of the melting temperature, T_m, were made similarly; after loading the hydroxyapatite with the reannealed duplexes and washing with 0·03 M-phosphate, the column was eluted 3 times with 8 ml of 0·16 M-phosphate at 5 deg. intervals from 60 to 95°C. Finally, the column was eluted with 0·4 M-phosphate; this removed a negligible amount of DNA and radioactivity.

(h) *Equilibrium centrifugation in caesium chloride and sodium iodide*

Samples were dissolved in 4·6 ml caesium chloride ($\rho_{20} = 1{\cdot}751$ g/cm^3) or sodium iodide ($\rho_{20} = 1{\cdot}546$ g/cm^3) containing 0·1 M-Tris, 0·01 M-EDTA (pH 8·0) and 2 to 10 μg native and denatured mouse DNA as carrier, and centrifuged to equilibrium for about 66 h at 20°C at 45,000 revs/min in the MSE titanium 10×10 ml rotor; fractions of 0·2 ml were unloaded, refractive indices measured and samples counted as described previously (Birnie, 1972). Labelled, heat-denatured LS cell (a substrain of mouse NCTC strain L929, which grows in suspension) DNA was also centrifuged as a marker in each rotor load.

(i) *Nuclease assay*

S1 nuclease was obtained as described by Sutton (1971). Samples were incubated for 2 h at 37°C in 0·05 M-sodium acetate (pH 4·5), 2 mM-$ZnSO_4$, 0·1 M-NaCl, 10 μg denatured mouse DNA/ml. Sufficient enzyme was added to ensure complete degradation of cDNA alone. A portion of the incubation mixture was taken to determine the total radioactivity; a further portion was acidified (1 N-$HClO_4$) at 4°C after addition of mouse DNA and bovine serum albumin (5 μg and 50 μg, respectively) and then centrifuged for 15 min at 2500 revs/min. A portion of the supernatant fluid was then taken to determine the acid-soluble radioactivity.

3. Results

(a) *Characteristics of complementary DNA*

In the presence of oligo(dT) primer and actinomycin D, the rate of synthesis of cDNA by reverse transcriptase was proportional to the concentration of 9 S RNA in the range of one to ten μg per ml. Under the conditions routinely used, the yield of high specific activity cDNA (25×10^6 to 40×10^6 disints/min per μg) was about 20% of the amount of 9 S RNA used. No labelled DNA was synthesized when reverse transcriptase was incubated with the oligo(dT) primer alone, using all four labelled deoxyribonucleoside triphosphates as precursors. This indicates that the dependence of our preparations of reverse transcriptase on added template is absolute (compare Kacian *et al.*, 1971). The mean molecular weight of cDNA was about 110,000 (330 bases), about 60% of the length of 9 S RNA (600 bases; Williamson *et al.*, 1971). For most experiments, only the higher molecular weight cDNA fragments (greater than about 240 bases) were isolated and characterized further. The size distribution of the cDNA used in most of the experiments is shown in Figure 5.

Purified cDNA was incubated at 43°C for two weeks in hybridization buffer (to a D_0t† of 0·3). The product banded in sodium iodide gradients as denatured DNA and was virtually all degraded (95%) by S1 nuclease. Therefore, cDNA contained a

† D_0t, Product of concentration of cDNA and time of incubation (mol s l^{-1}).
$D_0t_{\frac{1}{2}}$, Value of D_0t when half of hybridizable cDNA is converted to hybrid.

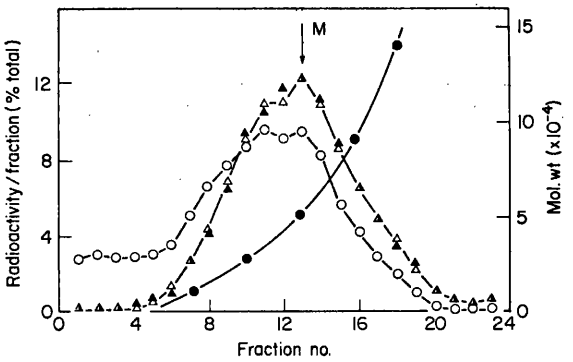


FIG. 2. Sedimentation of cDNA from 9 S RNA–cDNA hybrid in 5-ml 5% to 20% (w/w) alkaline sucrose gradients. M, marker DNA, $s_{20,w} = 4{\cdot}3$ (200 nucleotides); —△—△—, cDNA from cDNA–9 S RNA hybrid; —○—○—, cDNA from cDNA–9 S RNA hybrid after treatment of hybrid with S1 nuclease; —▲—▲—, cDNA before hybridization. The cDNA used in this experiment was of somewhat smaller size than in the other experiments described in this paper. Direction of sedimentation is from left to right.

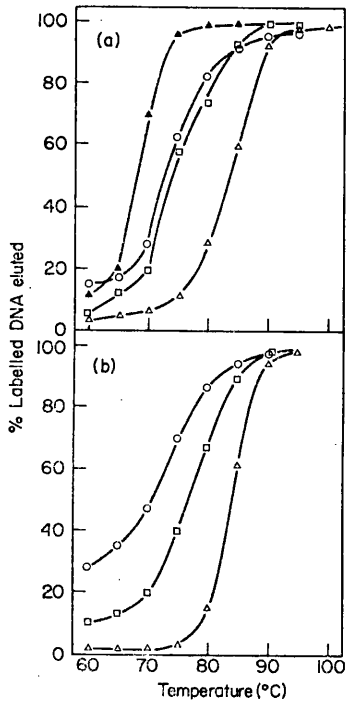


FIG. 3. T_m values of reannealed duplexes in 0·16 M-phosphate.

(a): —△—△—, Native mouse embryo DNA (330 base pairs); —□—□—, reannealed mouse embryo DNA; —○—○—, cDNA annealed to mouse embryo DNA; —▲—▲—, cDNA–9 S RNA hybrid obtained as in Fig. 1.

(b): —△—△—, Native foetal liver DNA (330 base pairs); —□—□—, reannealed foetal liver DNA; —○—○—, cDNA annealed to foetal liver DNA. Reannealed DNA samples were incubated in 0·12 M-phosphate to a C_0t value of 20,000. The T_m of reannealed mouse DNA sequences is about 6 deg. C lower than that of native mouse DNA. This difference may be partially due to the fact that the T_m values were determined on hydroxyapatite in 0·16 M-phosphate, rather than in 0·12 M-phosphate, for the reasons given in Materials and Methods.

In order to test the quality of the 9 S RNA–cDNA hybrid directly, the size of the cDNA in such hybrids was estimated before and after treatment of the hybrids with S1 nuclease. The size distribution of cDNA was not affected by the annealing conditions themselves (Fig. 2). Under conditions in which cDNA alone was completely degraded, cDNA in the 9 S RNA–cDNA hybrid was reduced to about half its initial size by such treatment, i.e. on average about one nick per hybrid molecule was introduced by the treatment with S1 nuclease (Fig. 2). This indicates that there are no extensive regions of S1 nuclease-sensitive cDNA distributed throughout the cDNA–9 S RNA hybrid. However, this hybrid had a T_m of 69°C in 0·16 M-phosphate (Fig. 3(a)), somewhat lower (3 deg. C) than predicted for a perfect hybrid (see Discussion).

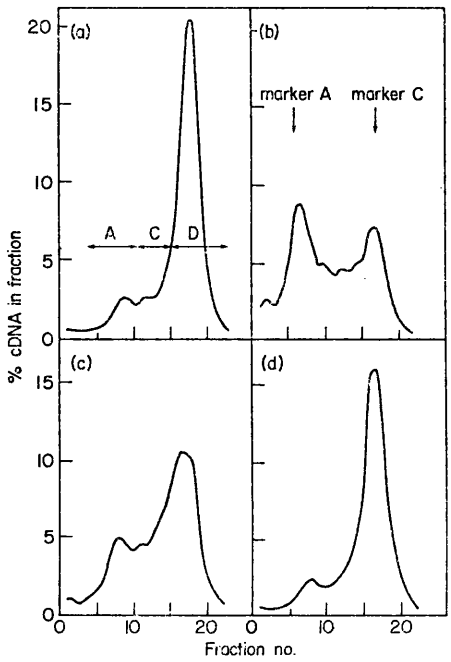


FIG. 4. Hybridization of cDNA to 9 S RNA.

(a) cDNA annealed to 14× excess poly(U)-Sepharose-prepared 9 S RNA as described in Materials and Methods (to $D_0t = 0{\cdot}3$) and treated with ribonuclease was banded in CsCl. cDNA banding in 3 regions of the gradient (arrows) was recovered free from 9 S RNA, reannealed to excess 9 S RNA, treated with ribonuclease and banded again in CsCl.

(b) Annealing to excess 9 S RNA of cDNA from region A in (a); (c) annealing to excess 9 S RNA of cDNA from region C in (a); (d) annealing to excess 9 S RNA of cDNA from region D in (a). Characteristics of CsCl gradient are as in Fig. 3. Markers are those described in Fig. 1.

A small fraction of cDNA was apparently incapable of forming a ribonuclease-resistant hybrid with 9 S RNA. To investigate this, cDNA was incubated with excess 9 S RNA, treated with ribonuclease and banded in CsCl. About 8% of the cDNA banded as DNA, 17% as hybrid of low density, and 75% as hybrid of high density (Fig. 4(a)). cDNA from each of these fractions of the gradient was recovered separately, reincubated with excess 9 S RNA and treated with ribonuclease. The results obtained after banding each of these fractions of cDNA in CsCl are shown in Fig. 4

((b) to (d)). Almost all the cDNA isolated from the ribonuclease-resistant hybrid formed in the first annealing formed ribonuclease-resistant hybrid in a second annealing to 9 S RNA; whereas about half of the cDNA that failed to hybridize to 9 S RNA in the first annealing did not hybridize at the second stage. About one-third of the cDNA isolated from the intermediate part of the gradient failed to re-hybridize to 9 S RNA. These results indicate, therefore, that about 10% of the total cDNA is not capable of forming ribonuclease-resistant hybrid during annealing to 9 S RNA.

(c) *Annealing of complementary DNA to DNA*

(i) *Reannealing of* E. coli *and mouse DNA fragments*

Most experiments used DNA fragments with an average length of 330 bases (Fig. 5). Under these conditions, *E. coli* DNA reassociated with a $C_0t_{\frac{1}{2}}$† value of about 6 (Fig. 6(a)). Using mouse DNA fragments of the same size, a small fraction (about 10%) reassociated at a low C_0t value, whereas the bulk of the mouse embryo, foetal liver and sperm DNA fragments reassociated with a $C_0t_{\frac{1}{2}}$ value of about 800 (Fig. 6(a) to (c)). With mouse embryo DNA fragments of larger or smaller size, the bulk of the

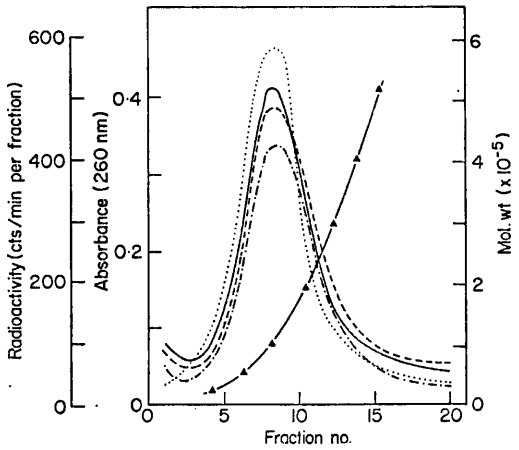


FIG. 5. Sedimentation of cDNA and mouse DNA fragments in 20-ml 5% to 10% alkaline sucrose gradient. (·····) cDNA; (–––––) foetal liver DNA fragments; (————) mouse embryo DNA fragments; (–·–·–·–·–) sperm DNA fragments; —▲—▲—, molecular weight.

DNA reassociated more quickly or more slowly, respectively (Table 1). With the exception of the smallest fragment, the rate of reassociation of mouse DNA fragments is roughly proportional to the square root of the fragment size, as observed also by Wetmur & Davidson (1968) and by Sutton & McCallum (1971). By comparison with *E. coli* DNA, the predicted rates of reassociation of unique sequences in mouse DNA fragments of different sizes are calculated (Table 1). These calculations show that the discrepancy between observed and predicted values decreased with decreasing fragment size. With the smallest fragment size (120 bases), the bulk of the mouse DNA sequences appear to be unique. This is almost certainly explained in terms of the greater separation of unique and repetitive sequences in mouse DNA with very small fragments (see Discussion).

† C_0t, Product of mouse or *E. coli* DNA concentration and incubation time (mol s l^{-1}).
$C_0t_{\frac{1}{2}}$, Value of C_0t when half of DNA (or cDNA) is converted to duplex.

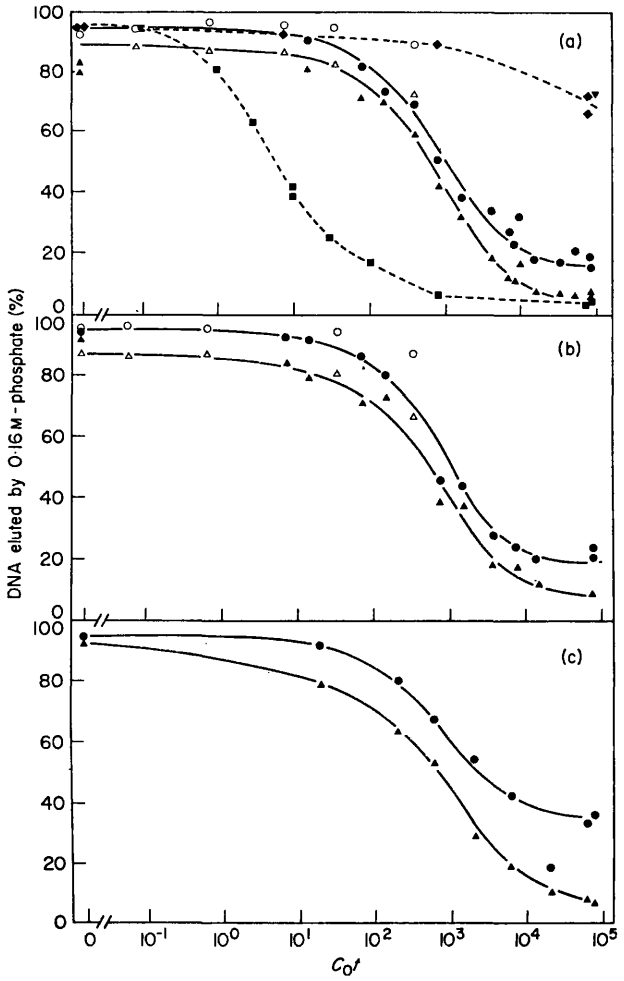


FIG. 6. Incubation of cDNA with *E. coli* and mouse DNA and reannealing of *E. coli* and mouse DNA fragments, measured by chromatography on hydroxyapatite, as described in Materials and Methods.

(a) (▲), (△) Reannealing of mouse embryo DNA fragments; (●), (○) annealing of cDNA to mouse embryo DNA fragments; (■) reannealing of *E. coli* DNA fragments; (◆) incubation of cDNA with *E. coli* DNA; (▼) incubation of poly(dT) with mouse DNA fragments.

(b) (▲), (△) Reannealing of foetal liver DNA fragments; (●), (○) annealing of cDNA to foetal liver DNA fragments.

(c) (▲) Reannealing of sperm DNA fragments; (●) annealing of cDNA to sperm DNA fragments.

In (a) and (b) the ratio of mouse DNA to cDNA was $1{\cdot}6\times10^7$ (w/w), but in (c), 8×10^6 (w/w). Open symbols refer to experiments in which the DNA concentrations were diluted 100-fold.

(ii) *Annealing of complementary DNA and poly(dT) to* E. coli *and mouse DNAs*

In order to simplify the interpretation of the results of these experiments, the various mouse and *E. coli* DNAs were each sonicated and fractionated to give a size distribution of fragments similar to that of cDNA itself (Fig. 5). Samples of these DNAs were mixed with cDNA in a ratio of 8×10^6 to 16×10^6 (w/w), denatured and allowed to anneal to various C_0t values. As a control experiment, cDNA was incubated

with *E. coli* DNA. Only about 10% of the cDNA annealed by a C_0t of 1000, during the period when all the *E. coli* DNA reassociated (Fig. 6(a)). However, after incubation to a C_0t of 80,000, a further 20% of the cDNA was not eluted from hydroxyapatite by 0·16 M-phosphate. Evidence that this is not due to self-annealing of the cDNA sequences has been given in a previous section. Caution must be exercised in interpreting this result, since when high specific activity poly(dT) was incubated with mouse DNA to a C_0t value of 80,000, about 30% of the poly(dT) behaved similarly.

With DNA from each mouse tissue studied, cDNA annealed at a $C_0t_{\frac{1}{2}}$ value identical to that for the reannealing of the bulk of the mouse DNA sequences (Fig. 6(a) to (c)). This was observed with cDNA and DNA fragments with lengths of 180 and 330 bases (Table 1). These relative rates of annealing of cDNA to mouse DNA and reannealing of mouse DNA fragments would be distorted if there were preferential degradation of cDNA during annealing relative to the mouse DNAs, or if mismatching of cDNA sequences occurred with respect to the globin genes in the mouse DNA. (1) Samples of cDNA and mouse DNA fragments (initial size 330 bases) were taken for size determinations after incubation to a C_0t value of 8000, when the annealing reaction was essentially complete. Negligible degradation of either cDNA or mouse DNA fragments was detected at this time, although after incubation to a C_0t value of 80,000, some degradation of cDNA and mouse DNA fragments occurred. Thus, neither the observed rate of annealing of cDNA to mouse DNA nor rate of annealing of mouse DNA fragments requires correction for changes in size of fragments during annealing. (2) For both total embryo and foetal liver DNA, the T_m of the cDNA–mouse DNA annealed duplexes was about three degrees lower than that of the reannealed mouse DNA duplexes themselves (Fig. 3). After correcting for the slight differences in (G + C) content of cDNA and mouse DNA (Gruenwedel *et al.*, 1971), this implies no more than about 2% mismatching between cDNA and the mouse DNA sequences to which cDNA annealed in these experiments (Laird *et al.*, 1969; Ullman & McCarthy, 1973*a*,*b*). This mismatching would reduce somewhat the rate of annealing of cDNA to mouse DNA. The magnitude of this effect is difficult to estimate precisely. On the basis of data for the effect of mismatching on the rate of reassociation of satellite DNA, the rate of annealing of cDNA to mouse DNA might be reduced to 60% of the rate if there was no mismatching (Southern, 1971; Sutton & McCallum, 1971). However, similar studies on the effect of mismatching on the rate of reassociation of bacterial DNAs (Britten & Bonner, 1971) suggest that this correction to the observed rate of annealing of cDNA to mouse DNA might be much smaller. This means that the sequences in mouse DNA to which cDNA anneals are as frequent (certainly not more than twice as frequent) as the bulk of the mouse sequences.

4. Discussion

(a) *Evidence that complementary DNA represents a replica of globin genes*

All conclusions from the present work concerning the number of globin genes are based on the belief that most of the cDNA is a faithful replica of part of the α and β globin genes. This stems from the evidence (1) that mouse reticulocyte 9 S RNA comprises mainly the messenger RNAs for α and β globins and (2) that cDNA is a faithful transcript of most of the sequences present in 9 S RNA.

(b) *Evidence that 9 S RNA comprises α and β globin messenger RNAs*

Using the duck lysate and oocyte systems devised by Lingrel (1972) and by Gurdon and co-workers (Lane *et al.*, 1971), our colleagues have demonstrated that the preparations of polysome-derived 9 S RNA used in the present experiments direct the synthesis of both α and β globin chains (Lanyon *et al.*, 1972; Williamson *et al.*, 1973). Since about 90% of protein synthesis in reticulocytes is represented by globin, it is generally assumed that the purest preparations of 9 S RNA available contain mainly globin mRNAs. Since these preparations of 9 S RNA are selected by their ability to bind poly(U)-Sepharose and by size, any mRNA contaminants must contain poly(A) regions and specify proteins about the same size as globin. Lane *et al.* (1971) have studied the tryptic peptides released from non-oocyte proteins synthesized after injection of rabbit 9 S RNA into oocytes. No peptide was observed that did not correspond to peptides obtained by tryptic digestion of carrier haemoglobin. This argues strongly that 9 S RNA contains mainly globin mRNAs.

These experiments do not indicate the ratio of mRNAs specifying α and β globin. Since mouse reticulocyte 9 S RNA can be only partially resolved into the components specifying α and β globin (Lanyon *et al.*, 1972), it has not yet proved possible to answer this question directly.

(c) *Evidence that complementary DNA is a faithful transcript of 9 S RNA*

cDNA is about 60% of the size of the 9 S RNA molecule and contains about 40 to 45% (G + C). Moreover, after depurination, only about 4% of the nucleotides in cDNA bind to poly(A)-Sepharose under conditions in which poly(dT) binds completely (J. Mansbridge, personal communication). cDNA contains, therefore, a very small proportion of poly(dT) sequences. About 85% of cDNA sequences hybridize to excess 9 S RNA at a temperature 15 degrees C below the T_m of a 40% (G + C) RNA–DNA hybrid under these conditions (Bishop, 1972*b*).

The bulk of the cDNA sequences represent reasonably faithful transcripts of the 9 S RNA sequences. This is shown by the fact that only about one nick per 9 S RNA–cDNA hybrid molecule is introduced by treatment with S1 nuclease. Any regions of mismatching, therefore, must be either too small for S1 nuclease to recognize or be clustered. An independent estimate of mismatching in the cDNA–9 S RNA hybrid can be obtained by comparing its T_m (69°C) with that of reannealed mouse DNA fragments (77 to 78°C). After correcting for slight differences in (G + C) content, this means that the T_m of cDNA–9 S RNA hybrid is nine to ten degrees C below that of a reannealed DNA duplex of the same (G + C) content (Gruenwedel *et al.*, 1971). This difference in T_m is three to four degrees C more than that reported for perfect RNA–DNA hybrids (Kohne, 1969; Gelderman *et al.*, 1971), which implies at most 2% mismatching of sequences in cDNA with respect to 9 S RNA (Laird *et al.*, 1969; Ullman & McCarthy, 1973*a*,*b*). One explanation of this mismatching is that cross-hybridization of α and β globin cDNAs to β and α globin mRNAs in 9 S RNA might occur. However, this appears most unlikely from consideration of the amino acid sequences of α and β mouse globins (Dayhoff, 1972). The most likely explanation of the mismatching is that reverse transcriptase introduces a low frequency of mistakes during transcription of 9 S RNA *in vitro*.

The results discussed above show that most of the cDNA represents a faithful, but partial, transcript of sequences present in the purest preparations of 9 S RNA

available, which contain mainly globin mRNAs. Nevertheless, it is conceivable that cDNA might represent transcripts of contaminant RNA sequences in 9 S RNA that do not specify globins. For this reason, we have devised and proved the method of titrating cDNA with increasing amounts of 9 S RNA (see accompanying paper, Young *et al.*, 1974). In the same paper, we also determine the sequence complexity of cDNA. The results from both experimental approaches support the conclusion that cDNA represents transcripts of both α and β globin mRNAs, each transcript being a partial, non-random copy of its template 9 S RNA molecule.

(d) *Estimation of the number of globin genes*

(i) *Reassociation of mouse DNA*

The bulk of the mouse DNA sequences reassociate at a rate that depends on the fragment size (Table 1). A comparison of experimental $C_0t_{\frac{1}{2}}$ values with those predicted on the basis of the observed $C_0t_{\frac{1}{2}}$ value for reassociation of *E. coli* DNA, the

TABLE 1

Rates of annealing of mouse embryo DNA, E. coli *DNA and complementary DNA fragments*

Initial mean DNA fragment size (bases)			$C_0t_{\frac{1}{2}}$		Apparent
Mouse	cDNA	Mouse DNA	cDNA–mouse DNA	Mouse DNA predicted	reiteration frequency
120		6000	—	6100	1
180†	240†	1400†	1600†	5000	4
330	330	800	800	3700	5
900	—	500	—	2200	4

The experimental $C_0t_{\frac{1}{2}}$ value for *E. coli* DNA reassociation (300 bases) was 6. The predicted $C_0t_{\frac{1}{2}}$ value is calculated from the relation between the rate of reassociation and the genetic complexity established by Britten & Kohne (1968), assuming the genome sizes of *E. coli* and the mouse to be $2{\cdot}8\times10^9$ and $1{\cdot}8\times10^{12}$, respectively, and assuming that the rate of reassociation of *E. coli* DNA is proportional to the square root of mean fragment size (Wetmur & Davidson, 1968). The apparent reiteration frequency is obtained by dividing the predicted $C_0t_{\frac{1}{2}}$ value for mouse DNA sequences by the observed $C_0t_{\frac{1}{2}}$ value. Fragments of 120 or 180 nucleotides were obtained by degradation with deoxyribonuclease I, whereas those of 330 or 900 nucleotides were obtained by sonication. Sizes of DNA fragments were checked as described in Materials and Methods.

† Data from Harrison *et al.* (1972*a*).

observed dependence of rate of reassociation on fragment size and the genetic complexities of the mouse and *E. coli* show good agreement only for the smallest fragments (120 nucleotides). This may mean that it is necessary (using sonication or enzymic methods of fragmentation) to reduce the size of mouse DNA fragments to about 120 nucleotides in order to obtain nearly complete separation of unique and repetitive sequences. Alternatively, the result for the smallest DNA fragments could be peculiar, for example, due to a marked change in the stringency of binding of very small duplexes by hydroxyapatite (Martinson, 1973*a*,*b*). However, this interpretation would imply that most DNA sequences are reiterated four to five times in the mouse genome.

Other workers, using sonication to fragment DNA to about 500 nucleotides in length, have obtained results similar to our own: namely, that the most slowly reassociating sequences reassociate more quickly than predicted for single-copy sequences. Thus Melli *et al.* (1971) observed a fourfold faster reassociation of sonicated rat DNA fragments than predicted for single-copy sequences. Bishop *et al.* (1972) and Bishop & Rosbash (1973) have reported a similar, though slightly smaller, discrepancy for duck RNA fragmented by sonication, whether duplex formation was assayed optically, by hydroxyapatite chromatography or resistance to single-strand-specific nuclease. This eliminates the possibility that the method for assay of duplex is the explanation of the discrepancy. However, the method of fragmenting DNA may be important: Mori *et al.* (1972) have found that the rate of reassociation of rat DNA fragmented by sonication is markedly faster than that of similar sized fragments produced by digestion with deoxyribonuclease II.

Britten & Smith (1969) and Grouse *et al.* (1972) found that (using hydrodynamic shearing) it was necessary to fragment eukaryotic DNAs at to least 400 to 500 nucleotides in order to prepare unique sequences in reasonable yield. However, such fragments of total DNA are reported still to contain interspersed some repetitive sequences with the unique sequences (Hoyer & Van de Velde, 1971; Davidson *et al.*, 1973). Nevertheless, the rates of reassociation of sequences in total DNA fragmented in this way correspond quite closely with those predicted for unique sequences in the calf (Britten & Kohne, 1968; Britten & Smith, 1969), in the mouse (Hahn & Laird, 1971), in the *Zenopus* oocyte (Davidson & Hough, 1971) and in *Chironomus tentans* (Sachs & Clever, 1972). Britten & Smith (1969) have argued that in calf DNA at least, it is unlikely that sequences exist that are repeated only a few times. However, in a recent analysis of the interspersion of repetitive and unique DNA sequences in *Xenopus*, Davidson *et al.* (1973) conclude that about 10% of the total DNA comprises sequences repeated about 20 times. Similar intimate interspersion of unique and slightly repetitive sequences in mouse DNA would explain our own data.

(ii) *Annealing of complementary DNA to mouse DNAs*

Using fragments of 180 or 330 nucleotides, cDNA anneals to excess mouse DNA at the same rate as reassociation of mouse DNA itself, which by comparison with *E. coli* DNA is that expected if most of the sequences in mouse DNA are reiterated four to five times. We have not attempted to determine whether, using very small fragments, the rate of annealing of cDNA to mouse DNA corresponds to that expected for unique sequences (as is the case for the bulk of the mouse sequences themselves), owing to the complications introduced by the significant proportion of poly(dT) sequences in cDNA of such small size.

In principle, as we indicated in a previous report (Harrison *et al.*, 1972*a*), it should be possible to clarify this issue by determining the complexity of the sequences in mouse DNA that anneal to cDNA, from the fraction of cDNA that anneals to mouse DNA at high C_0t values. In the experiments described in Figure 6(a) and (b), or (c), it can be calculated (assuming the value for the complexity of cDNA given in the accompanying paper) that 28 or 43%, respectively, of the cDNA should remain unannealed at high C_0t values if there were single copies of the α and β globin genes. In fact, the experimental values are 20% and 35%, respectively. We have attempted to investigate systematically how the proportion of cDNA annealed after "high C_0t" incubation varies with the ratio of mouse DNA/cDNA, and have attempted to

rationalise the result in terms of a titration theory that could be used to determine the complexity of the mouse DNA sequences that anneal to cDNA. However, the results of these experiments have not been consistent and we have not pursued this method further.

With fragments of 330 nucleotides, the rate of annealing of cDNA to mouse DNA is the same as that for the reassociation of mouse DNA, whether mouse DNA is isolated from total embryos, erythroid cells (foetal liver) or sperm. The sizes of cDNA and the mouse DNA fragments are not reduced significantly during the period of annealing, and little correction for mismatching in cDNA–mouse DNA duplexes is required. These facts taken together show that in all the mouse tissues studied, the sequences in their DNA complementary to cDNA are about as frequent as the most slowly reassociating DNA sequences. Since the slight extent of mismatching in cDNA–9 S RNA hybrids is the same as that in cDNA–mouse DNA duplexes, cDNA must anneal only to those globin genes expressed in mouse reticulocyte 9 S RNA, or to very closely related variants. In the mouse, the only globin genes that might possibly be sufficiently similar in this respect are the β and z chains (14% amino acid differences). Mouse α and x chains are quite dissimilar (36% amino acid differences) (Steinheider *et al.*, 1971).

In view of the arguments developed so far, it can be concluded that in the mouse: (a) there are probably only single, but not more than four to five copies each of the α and β globin genes in the mouse germ line; (b) that there is no globin gene amplification in erythroid cells, since foetal liver in this stage of development contains 80% erythroid cells, of which all but the most immature contain detectable amounts of globin mRNA (Harrison *et al.*, 1973); and (c) that there is no widespread elimination of globin genes during non-erythroid somatic differentiation, since the globin gene dosage in sperm, total embryos and erythroid cells is the same. This conclusion concerning the numbers of genes in various mouse tissues would be valid even if episomal copies of the globin gene existed. Since we prepared all DNAs by the hydoxyapatite method (which would retain even gene-size fragments of DNA), the DNAs used in these experiments should include any nuclear DNA sequences, whether of chromosomal origin or otherwise.

These conclusions extend the findings reported in less detail in a previous paper (Harrison *et al.*, 1972*a*) and are broadly in agreement with other recent reports, which compare the rate of annealing of cDNA or *in vivo* labelled duck reticulocyte 10 S RNA to excess duck DNA with that expected theoretically for reassociation of unique duck DNA sequences. These reports show that the globin genes in the duck may be slightly reiterated (two to three copies; Bishop *et al.*, 1972; Packman *et al.*, 1972; Bishop & Rosbash, 1973), and that there is no globin gene amplification in duck erythroid cells (Packman *et al.*, 1972). These studies did not investigate whether degradation of cDNA or mismatching of cDNA with respect to duck DNA introduced any correction to these quoted values.

Both our own and these other estimates of the reiteration frequency of the globin genes in mouse and duck cells are markedly different to those obtained by hybridizing excess chemically labelled mouse or chicken reticulocyte 9 S RNA to mouse or chicken DNA (about 50,000: Williamson *et al.*, 1970; Morrison *et al.*, 1972; de Jiminez *et al.*, 1971). These results have been interpreted in terms of the hybridization of labelled poly(A) regions in the 9 S RNA to repetitive DNA sequences (Morrison *et al.*, 1972), which implies that about 0·4% of mouse DNA sequences (presumably

poly(dT)) are complementary to the poly(A) regions in 9 S RNA. This value is approximately the same as that obtained for the proportion of poly(dA) regions in mouse DNA by Shenkin & Burdon (1972).

The authors are indebted to the following people: Dr J. W. Beard, Duke University, U.S.A., for continuing supplies of avian myeloblastosis virus; Dr Ailsa Campbell of the Department of Biochemistry, Glasgow University, for checking the size of DNA fragments on the analytical ultracentrifuge; Dr G. Threlfall for preparating synthetic RNA–DNA hybrids; Dr R. Williamson, Mr G. Lanyon and Miss S. Laird for supplies of 9 S RNA; and Miss E. MacPhail, Mr A. McPhee and Mr A. McNab for technical assistance. We are grateful to other members of staff for useful discussion and criticisms, especially to Drs R. Williamson and R. S. Gilmour. This work was supported by grants from the Medical Research Council and the Cancer Research Campaign.

REFERENCES

Adesnik, M., Salditt, M., Thomas, W. & Darnell, J. E. (1972). *J. Mol. Biol.* **71**, 21–30.
Attardi, G. & Amaldi, F. (1970). *Annu. Rev. Biochemistry*, **39**, 183–226.
Beermann, W. (1972). In *Results and Problems in Cell Differentiation* (Beermann, W., ed.), vol. 4, pp. 1–33, Springer-Verlag, Berlin, Heidelberg and New York.
Birnie, G. D. (1972). *FEBS Letters*, **27**, 19–22.
Birnstiel, M. L., Grunstein, M., Speirs, J. & Hennig, W. (1969). *Nature (London)*, **223**, 1265–1267.
Birnstiel, M. L., Chipcase, M. & Speirs, J. (1971). *Prog. Nucleic Acid Res. Mol. Biol.* **11**, 351–389.
Bishop, J. O. (1972*a*). *Karolinska Symp. on Res. Methods in Reproductive Endocrinology* (Diczfalusy, E., ed.), pp. 247–276, Karolinska Institutet, Stockholm.
Bishop, J. O. (1972*b*). *Biochem. J.* **126**, 171–185.
Bishop, J. O. & Rosbash, M. (1973). *Nature New Biol.* **241**, 204–207.
Bishop, J. O., Pemberton, R. & Baglioni, C. (1972). *Nature New Biol.* **235**, 231–234.
Borenfreund, E., Fitt, E. & Bendich, A. (1961). *Nature (London)*, **191**, 1375–1377.
Britten, R. J. & Bonner, T. (1971). *Yearb. Carneg. Instn*, **70**, 373–374.
Britten, R. J. & Kohne, D. E. (1968). *Science*, **161**, 529–540.
Britten, R. J. & Smith, J. (1969). *Yearb. Carneg. Instn*, **68**, 378–386.
Davidson, E. H. & Hough, B. R. (1971). *J. Mol. Biol.* **56**, 491–506.
Davidson, E. H., Hough, B. R., Amenson, C. S. & Britten, R. J. (1973). *J. Mol. Biol.* **77**, 1–23.
Dayhoff, M. O. (1972). Editor of *Atlas of Protein Sequence and Structure*, vol. 5, National Biomedical Research Foundation.
de Jiminez, E. S., Dominguez, J. L., Webb, F. H. & Bock, R. M. (1971). *J. Mol. Biol.* **61**, 59–71.
Delovitch, T. L. & Baglioni, C. (1973). *Proc. Nat. Acad. Sci., U.S.A.* **70**, 173–178.
Gelderman, A. H., Rake, A. V. & Britten, R. J. (1971). *Proc. Nat. Acad., U.S.A.* **68**, 172–176.
Grouse, L., Chilton, N. D. & McCarthy, B. J. (1972). *Biochemistry*, **11**, 798–805.
Gruenwedel, D. W., Hsu, C. H. & Lu, D. S. (1971). *Biopolymers*, **10**, 47–68.
Hahn, W. E. & Laird, C. D. (1971). *Science*, **173**, 158–161.
Harrison, P. R., Hell, A., Birnie, G. D. & Paul, J. (1972*a*). *Nature (London)*, **239**, 219–221.
Harrison, P. R., Hell, A. & Paul, J. (1972*b*). *FEBS Letters*, **24**, 73–76.
Harrison, P. R., Conkie, D., Paul, J. & Jones, K. (1973). *FEBS Letters*, **32**, 109–112.
Hell, A. (1972). *MSE Application Information Sheet* A6/6/72.
Hell, A., Birnie, G. D., Slimming, T. K. & Paul, J. (1972). *Anal. Biochem.* **48**, 369–377.
Hoyer, B. H. & van de Velde, N. W. (1971). *Yearb. Carneg. Instn*, **70**, 375–376.
Kacian, D. L., Watson, K. F., Burny, A. & Spiegelman, S. (1971). *Biochim. Biophys. Acta*, **246**, 365–383.
Kacian, D. L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L. & Marks, P. A. (1972). *Nature New Biol.* **235**, 167–169.

Kedes, L. H. & Birnstiel, M. L. (1971). *Nature New Biol.* **230**, 165–169.
King, J. L. & Jukes, T. H. (1969). *Science*, **164**, 788–798.
Kohne, D. E. (1969). *Biophys. J.* **8**, 1104–1118.
Laird, C. D., McConaughty, B. L. & McCarthy, B. J. (1969). *Nature* (*London*), **224**, 149–154.
Lane, C. D., Marbaix, G. & Gurdon, J. B. (1971). *J. Mol. Biol.* **61**, 73–91.
Lanyon, W. G., Paul, J. & Williamson, R. (1972). *Eur. J. Biochem.* **31**, 38–43.
Lingrel, J. B. (1972). *Methods in Protein Biosynthesis*, Methods in Molecular Biology Series, vol. 2 (Laskin, A. E. & Last, J. A., eds), Marcel Dekker, New York.
Martinson, H. G. (1973*a*). *Biochemistry*, **12**, 139–145.
Martinson, H. G. (1973*b*). *Biochemistry*, **12**, 145–150.
McEwen, C. R. (1967). *Anal. Biochem.* **20**, 114–149.
Melli, M., Whitfield, C., Rao, K. V., Richardson, M. & Bishop, J. O. (1971). *Nature New Biol.* **231**, 8–12.
Mori, K., Winterith, M. & Mandel, P. (1972). *Biochimie*, **54**, 1427–1434.
Morrison, M. R., Paul, J. & Williamson, R. (1972). *Eur. J. Biochem.* **27**, 1–9.
O'Brien, S. J. (1973). *Nature New Biol.* **242**, 52–54.
Ohno, S. (1971). *Nature* (*London*), **234**, 134–137.
Ohta, T. & Kimura, M. (1971). *Nature* (*London*), **233**, 118–119.
Packman, S., Aviv, H., Ross, J. & Leder, P. (1972). *Biochem. Biophys. Res. Commun.* **49**, 813–819.
Sachs, R. I. & Clever, U. (1972). *Expt. Cell Res.* **74**, 587–591.
Shenkin, A. & Burdon, R. H. (1972). *FEBS Letters*, **22**, 157–160.
Southern, E. M. (1971). *Nature New Biol.* **232**, 82–83.
Steinheider, G., Medleris, H. & Osterlag, W. (1971). In *Syntheses, Struktur und Funktion des Hamoglobins* (Martin & Nowicki, eds), p. 225, J. F. Lehmans, Verlag, Munchen.
Studier, F. W. (1965). *J. Mol. Biol.*, **11**, 373–390.
Sutton W. D. (1971). *Biochim. Biophys. Acta*, **240**, 522–531.
Sutton, W. D. & McCallum, M. (1971). *Nature New Biol.* **232**, 83–85.
Suzuki, Y., Gage, L. P. & Brown, D. D. (1972). *J. Mol. Biol.*, **70**, 637–649.
Ullman, J. S. & McCarthy, B. J. (1973*a*). *Biochim. Biophys. Acta*, **294**, 405–415.
Ullman, J. S. & McCarthy, B. J. (1973*b*). *Biochim. Biophys. Acta*, **294**, 416–424.
Weinberg, E. S., Birnstiel, M. L., Purdom, I. F. & Williamson, R. (1972). *Nature* (*London*), **240**, 225–228.
Wetmur, J. G. & Davidson, W. (1968). *J. Mol. Biol.*, **31**, 349–370.
Williamson, R., Morrison, M. & Paul, J. (1970). *Biochem. Biophys. Res. Commun.* **40**, 740–745.
Williamson, R., Morrison, M., Lanyon, G., Eason, R. & Paul, J. (1971). *Biochemistry*, **10**, 3014–3021.
Williamson, R., Drewienkiewicz, C. E. & Paul, J. (1973). *Nature New Biol.*, **241**, 66–68.
Young, B. D., Harrison, P. R., Gilmour, R. S., Birnie, G. D., Hell, A., Humphries, S. & Paul, J. (1974). *J. Mol. Biol.* **84**, 555–568.

Reprinted from *J. Mol. Biol.* (1974) **84**, 555–568

Kinetic Studies of Gene Frequency

II. Complexity of Globin Complementary DNA and its Hybridization Characteristics

B. D. Young, P. R. Harrison, R. S. Gilmour, G. D. Birnie
Anna Hell, S. Humphries and J. Paul

J. Mol. Biol. (1974) 84, 555–568

Kinetic Studies of Gene Frequency

II. Complexity of Globin Complementary DNA and its Hybridization Characteristics

B. D. Young, P. R. Harrison, R. S. Gilmour, G. D. Birnie
Anna Hell, S. Humphries and J. Paul

The Beatson Institute for Cancer Research
132 Hill Street, Glasgow G3 6UD, Scotland

(*Received 4 May 1973, and in revised form 10 January 1974*)

Single-stranded complementary DNA and double-stranded complementary DNA were synthesized on a template of mouse reticulocyte 9 S RNA using reverse transcriptase from avian myeloblastosis virus. cDNA^s† (mean size 330 nucleotides) was 95% sensitive to single-strand-specific nuclease (S1) after extensive incubation, indicating that there were very few complementary sequences in cDNA^s. On the other hand, cDNA^d (mean denatured size 105 nucleotides) was only 35% sensitive to S1 nuclease after being denatured and incubated for a similar period, indicating that there were many complementary sequences in cDNA^d.

The kinetics of the hybridization of cDNA^s to 9 S RNA are in agreement with second-order reaction theory. Hence, the time-course of the hybridization of cDNA^s to an equal amount of 9 S RNA can be converted, using the appropriate factor, to the DNA–DNA reaction of the same base composition and complexity. By comparison with the renaturation of *Escherichia coli* DNA, it was estimated that cDNA^s had been transcribed from RNA sequences totalling 700 to 800 bases. After correction for fragment size and base composition, the reassociation of denatured cDNA^d was also compared with the *E. coli* DNA renaturation curve and it was estimated that cDNA^d was transcribed from RNA sequences totalling 330 bases.

The titration of cDNA^s with increasing amounts of 9 S RNA has been analysed by taking into account the cDNA^s size distribution. This analysis indicates that all classes of molecules present in 9 S RNA are partially transcribed into cDNA^s. The most likely interpretation of all our results is that cDNA^s and cDNA^d both contain sequences transcribed from a limited region of two such classes, probably those messenger RNA molecules coding for α and β globin chains.

1. Introduction

Reverse transcriptase has been used to synthesize radioactive complementary DNA on a template of reticulocyte 9 S RNA (Ross *et al.*, 1972; Verma *et al.*, 1972; Kacian *et al.*, 1972). Since there is good evidence (Lockard & Lingrel, 1969; Mathews *et al.*, 1971; Housman *et al.*, 1971) that reticulocyte 9 S RNA consists largely of messenger RNA coding for globin, it has been implied that cDNA† comprises at least partial copies of globin genes. Harrison *et al.* (1972*a*,1974) have established that cDNA transcribed from mouse reticulocyte 9 S RNA anneals with the slowest renaturing

† Abbreviations used: cDNA^s, single-stranded complementary DNA; cDNA^d, double-stranded complementary DNA.

fraction of mouse DNA, implying that globin genes are reiterated infrequently, if at all, in the mouse genome, even in erythroid tissue. Similarly, by comparing the rate of annealing of cDNA transcribed from duck reticulocyte 9 S RNA with the expected renaturation kinetics of unique duck DNA, Packman *et al.* (1972) and Bishop & Rosbash (1973) have concluded that globin genes may be reiterated no more than two or three times in the duck genome.

Mouse reticulocyte 9 S RNA is known to contain both α and β globin mRNA (Lanyon *et al.*, 1972; Williamson *et al.*, 1973) and, therefore, it is important to establish conclusively that cDNA has been transcribed from these mRNAs and is not merely a copy of a minor contaminant RNA. Since this information cannot be obtained from the kinetics of annealing cDNA to mouse DNA, it has been necessary to examine in detail the hybridization characteristics of cDNA. Our strategy has been, first, to determine the base sequence complexity of both cDNA^s and cDNA^d, and second to demonstrate that cDNA^s is transcribed from all classes of molecules present in 9 S RNA.

2. Materials and Methods

Details of all methods, including the preparation of cDNA^s, are described fully in the accompanying paper (Harrison *et al.*, 1974), except as indicated below.

(a) *Preparation of double-stranded complementary DNA*

cDNA^d was prepared by incubating mouse reticulocyte 9 S RNA with reverse transcriptase exactly as described by Harrison *et al.* (1974), except that actinomycin D was omitted. The reaction mixture was chromatographed on Sephadex G50 in 0·1 M-NaCl and the excluded fraction collected. The size of the labelled DNA so obtained was determined by sedimentation of a sample in alkaline sucrose (Fig. 1). The remainder of the excluded fraction was adjusted to contain 0·05 M-sodium acetate (pH 4·5), 2 mM-$ZnSO_4$, 0·1 M-NaCl, 10 μg denatured mouse DNA/ml, and 60 μg pancreatic ribonuclease/ml, incubated for 2 h at 37°C with S1 nuclease (sufficient to degrade denatured DNA completely within 30 min), and chromatographed on Sephadex G50 in 0·1 M-NaCl. The excluded fraction was sedimented in alkaline sucrose (Fig. 1); both strands of the cDNA^d were obtained from gradient fractions 6 to 14 by neutralization, chromatography on Sephadex G50 equilibrated with distilled water and lyophilization.

(b) *Treatment with* Neurospora crassa *nuclease*

N. crassa nuclease was obtained as described by Linn & Lehman (1965) omitting the phosphocellulose and hydroxyapatite steps. Samples were incubated for 2 h at 37°C in 0·1 M-Tris (pH 7·5), 0·01 M-$MgCl_2$, 10 μg native DNA/ml after addition of sufficient enzyme to ensure complete degradation of denatured DNA within 30 min. The extent of nuclease degradation was then determined as for S1 nuclease (Harrison *et al.*, 1974).

(c) *Preparation of [^{14}C]DNA from LS cells*

[^{14}C]DNA was prepared (Hell *et al.*, 1972) from the nuclei of LS cells (a substrain of mouse NCTC strain L929, which grows in suspension) that had been grown in the presence of [α-^{14}C]thymidine (59 μCi/μmol; 2 μCi/ml) for 5 days.

3. Theory

(a) *Analysis of complementary DNA–9 S RNA hybridization kinetics*

It is well established that the renaturation of DNA follows second-order kinetics (Britten & Kohne, 1968; Wetmur & Davidson, 1968), the rate-constant being dependent on temperature, fragment size, viscosity and base composition. For DNA of

uniform complexity, the parameter $C_0t_{\frac{1}{2}}$† has been shown to be proportional to the unique base sequence complexity. This relation may be deduced from the facts that (a) for a fixed amount of DNA the concentration of unique sequences is inversely proportional to genome size, and (b) for an ideal second-order reaction with equal initial concentrations of the two reactants, $C_0t_{\frac{1}{2}}$ is constant.

The dependence of the initial rate of hybridization on the concentrations of RNA and DNA observed by Nygaard & Hall (1964) indicates that the initial part of the hybridization time-course can be explained in terms of second-order reaction theory. There is also evidence from the experiments of Laird & McCarthy (1968), Bishop (1969) and Birnstiel *et al.* (1972) that the rate of hybrid formation is approximately inversely proportional to the complexity of the complementary RNA and DNA sequences that take part in the reaction. We consider, therefore, cDNA–9 S RNA hybridization to be represented by an irreversible bimolecular reaction, where D_0 and R_0 are the initial concentrations of cDNA and RNA, respectively, H is the concentration of hybrid and k is the rate-constant.

The general rate equation for such a system may be written in the form:

$$\frac{dH}{dt} = k(D_0 - H)(R_0 - H). \tag{1}$$

If the initial concentrations R_0 and D_0 are unequal, by integration, the time-course of this reaction is given by:

$$H = \frac{R_0D_0(1 - e^{(R_0 - D_0)kt})}{D_0 - R_0\, e^{(R_0 - D_0)kt}}. \tag{2}$$

If, however, the initial concentrations are equal, a condition always fulfilled by native DNA, the time-course is described by:

$$H = \frac{D_0{}^2kt}{1 + D_0kt}. \tag{3}$$

This equation describes the "C_0t" curves used by Britten & Kohne (1968) to interpret DNA renaturation. By analogy with the parameter $C_0t_{\frac{1}{2}}$, it is convenient to characterize the hybridization reaction using the parameter $D_0t_{\frac{1}{2}}$. The dependence of this term on the concentration ratio of the two species (R_0/D_0) can be demonstrated by considering three cases, namely $R_0 > D_0$, $R_0 = D_0$ and $R_0 < D_0$. It can be shown from equations (2) and (3) that for

$$R_0 > D_0, \quad D_0t_{\frac{1}{2}} = \frac{\log_e\left(2 - \frac{D_0}{R_0}\right)}{\left(\frac{R_0}{D_0} - 1\right)k}; \tag{4}$$

$$R_0 = D_0, \quad D_0t_{\frac{1}{2}} = \frac{1}{k}; \tag{5}$$

† C_0 is used here to denote concentration of DNA in moles nucleotides/l in DNA reannealing experiments. D_0 and R_0 are used to denote concentrations of DNA and RNA, respectively, in moles nucleotides/l in DNA–RNA hybridization experiments. $t_{\frac{1}{2}}$ is the time (s) for 50% of the reaction to take place.

$$R_0 < D_0, \quad D_0 t_{\frac{1}{2}} = \frac{\log_e\left(2 - \frac{R_0}{D_0}\right)}{\left(1 - \frac{R_0}{D_0}\right)k}. \tag{6}$$

The terms $D_0 t_{\frac{1}{2}}$ and R_0/D_0 can be experimentally determined and thus an estimate of the rate-constant for cDNA–9 S RNA hybridization can be obtained. With appropriate correction factors it is possible, therefore, to estimate the base sequence complexity of cDNA^s. Another feature of equations (4), (5) and (6) is that it can be shown (see Fig. 3) that the slowest reaction, measured in terms of $D_0 t_{\frac{1}{2}}$, occurs for $R_0 = D_0$. The importance of this result is that it allows conditions to be chosen such that all hybridization reactions are complete.

(b) *Analysis of single-stranded complementary DNA–RNA titration curves*

Hybridization of cDNA^s to RNA may be used as a general technique to estimate the fraction of an RNA preparation that is complementary to cDNA^s. The simplest approach is to titrate cDNA^s with increasing amounts of RNA, under conditions that allow completion of hybridization reactions. If cDNA^s were of homogeneous size, the titration curve, namely the plot of the fraction of cDNA^s hybridized *versus* the RNA/cDNA^s ratio, would consist of a line through the origin for cDNA^s excess and a horizontal line for RNA excess. The proportion of the RNA preparation that is complementary to cDNA^s could be found from the RNA/cDNA^s ratio at the point of intersection, because at that value there would be an equal number of cDNA^s and complementary RNA molecules present. In practice, however, cDNA^s is heterogeneous in size, so that in cDNA^s excess there is a greater probability of hybridization to larger cDNA^s molecules.

In order to analyse the effect on titration curves of the size of cDNA^s molecules, it is necessary to consider cDNA^s as being composed of a number of discrete size classes (m). We define the ith class as the total number n_i of molecules of length l_i nucleotides, where $i = 1, 2, \ldots m$. It can be assumed that the hybridizing RNA molecules have a constant length r nucleotides, which is larger than the largest cDNA molecules. We also define f_i as the number of hybridized cDNA molecules of the ith class. Hence $0 \leq f_i \leq n_i$ is true for $i = 1, 2, \ldots m$. Provided that only one cDNA^s molecule can hybridize to each RNA molecule, the number of RNA molecules (when in cDNA excess) is given by $\sum_i f_i$. Hence the RNA/DNA ratio (R_0/D_0) is given by

$$R_0/D_0 = \frac{r\sum_i f_i}{\sum_i n_i l_i}. \tag{7}$$

Similarly, the ratio of hybridized cDNA to the initial amount is given by

$$H/D_0 = \frac{\sum_i f_i l_i}{\sum_i n_i l_i}. \tag{8}$$

It may be assumed that the probability that an RNA molecule will hybridize to a cDNA molecule of the ith class is proportional to the number of possible nucleation

sites, l_i, and to the number of unhybridized cDNA molecules present, $n_i - f_i$. Hence the probability of hybridization to the ith class is proportional to $(n_i - f_i)l_i$. We consider the hybridization of a small amount of RNA which produces small changes, Δf_i, in f_i. It follows that:

$$\frac{\Delta f_i}{\sum_j \Delta f_j} = \frac{(n_i - f_i)l_i}{\sum_j (n_j - f_j)l_j}, \qquad \text{for } i = 1, \ldots m. \tag{9}$$

Hence we define $\Delta\alpha$ by

$$-\Delta\alpha = \frac{\Delta f_1}{(n_1 - f_1)l_1} = \frac{\Delta f_2}{(n_2 - f_2)l_2} = \ldots = \frac{\sum_j \Delta f_j}{\sum_j (n_j - f_j)l_j}. \tag{10}$$

We let $\Delta\alpha \to \mathrm{d}\alpha$ and $\Delta f_1 \to \mathrm{d}f_1$, etc. and then we have

$$-\mathrm{d}\alpha = \frac{\mathrm{d}f_1}{(n_1 - f_1)l_1} = \frac{\mathrm{d}f_2}{(n_2 - f_2)l_2} = \ldots . \tag{11}$$

It can be shown by integration that the above equations are satisfied if

$$f_i = n_i(1 - \mathrm{e}^{l_i\alpha}). \tag{12}$$

Thus we have from equations (7) and (8):

$$R_0/D_0 = \frac{r\sum_i n_i(1 - \mathrm{e}^{l_i\alpha})}{\sum_i n_i l_i}, \tag{13}$$

$$H/D_0 = \frac{\sum_i n_i l_i(1 - \mathrm{e}^{l_i\alpha})}{\sum_i n_i l_i}. \tag{14}$$

Both R_0/D_0 and H/D_0 are expressed above as functions of the independent variable α, which ranges from $0 (R_0/D_0 = H/D_0 = 0)$ to $-\infty$ $(H/D_0 = 1)$. The titration curve expected theoretically for any cDNA size distribution can be obtained by selecting appropriate values of α in the range 0 to $-\infty$, and calculating the corresponding values of R_0/D_0 and H/D_0 from the above equations. In such calculations the experimental size distributions are equivalent to plots of $n_i l_i$ *versus* l_i, and the difference between the ith and the $(i + 1)$th class has the minimum value of one nucleotide. Thus the number of terms (m) in the above summations corresponds to the difference in nucleotides between the largest and smallest cDNA[s] molecules.

4. Results

(a) *Characteristics of complementary DNA*

The characteristics of cDNA[s] have been fully described elsewhere (Harrison *et al.*, 1974). It has been shown (McDonnell *et al.*, 1970) that without actinomycin D in the incubation mixture, reverse transcriptase synthesizes both RNA–DNA hybrids and single and double-stranded DNA. In order to prepare cDNA[d], it is necessary to

TABLE 1

Nuclease sensitivity of mouse DNA and reverse transcriptase product after 120 minutes

Nuclease	Denatured LS DNA (%)	Reverse transcriptase product (%)
S1	93	17
S1 + RNAase	—	32
N. crassa	90	20
N. crassa + RNAase	—	40

remove hybrid molecules by simultaneous treatment with ribonuclease and nuclease specific for single-stranded DNA molecules. The action of S1 and *N. crassa* nucleases on the reverse transcriptase product and on denatured LS cell DNA is compared in Table 1. It can be seen that the addition of ribonuclease to the reverse transcriptase reaction product increases the degradation of cDNA by releasing cDNA[s] from hybrid molecules, but that there is little difference in the proportion degraded in 120 minutes by S1 and *N. crassa* nucleases. S1 nuclease was used for the analysis of the annealing reactions because, unlike *N. crassa* nuclease, it is not inhibited by low concentrations of formamide. Sedimentation in an alkaline sucrose gradient (Fig. 1) shows that

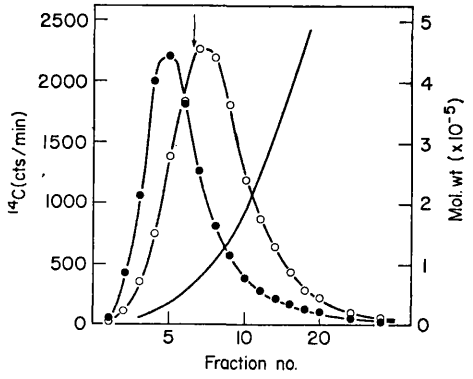


FIG. 1. Alkaline sucrose sedimentation of labelled DNA produced by reverse transcriptase without actinomycin D, as described in Materials and Methods. —○—○—, Labelled product excluded from Sephadex G50; —●—●—, labelled product treated with ribonuclease and S1 nuclease and excluded from Sephadex G50. The arrow indicates the position of a 4·3 S marker. (——) Molecular weight.

treatment with ribonuclease and S1 nuclease reduced the mean fragment size of cDNA from 80,000 (240 nucleotides) to 35,000 (105 nucleotides). The buoyant density in CsCl of cDNA[d] prepared in this way was 1·705 g/cm^3 ,which, if it is double-stranded DNA (see below), corresponds to a base composition of about 45% (G + C). Furthermore, sedimentation in alkaline CsCl did not reveal any inter-strand bias in base composition.

The thermal denaturation of cDNA^d was studied by elution from hydroxyapatite and a sharp melting temperature (T_m) of 76°C was obtained. This value may have been depressed because of the very small size of the duplex resulting in premature elution of partially melted cDNA^d (Martinson, 1973*a*,*b*). However, the T_m of cDNA^d that had been denatured and allowed to reanneal was 75°C. If these are true melting temperatures, the small change in T_m suggests that the increase in base mispairing produced by reannealing thermally denatured cDNA^d is not more than 1·5% (Laird *et al.*, 1969; Ullmann & McCarthy, 1973*a*,*b*).

(b) *Hybridization time-courses*

The S1 nuclease assay was used to determine hybridization time-courses with 9 S RNA/cDNA^s concentration ratios in the range of 100 to 0·1. The normalized time-courses for the ratios of 100, 10, 1 and 0·1 are plotted in Figure 2(a) as functions of D_0t. At the start of each reaction, 5 to 10% of the cDNA^s is nuclease resistant, and under conditions of 9 S RNA excess, the final amount of cDNA^s incorporated into nuclease-resistant hybrids is not more than 80%. About half of the final 20% nuclease sensitivity can be explained by cDNA sequences that cannot hybridize

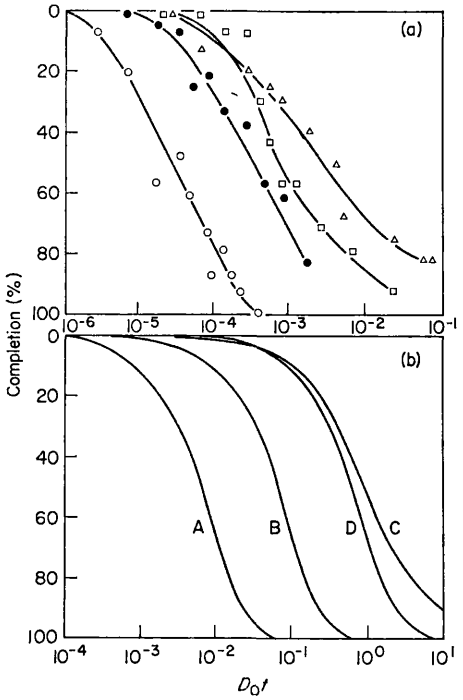


FIG. 2. (a) Hybridization of cDNA^s to reticulocyte 9 S RNA determined using the S1 nuclease assay. Four initial concentration ratios (R_0/D_0) were used: (○) $R_0/D_0 = 100$; (●) $R_0/D_0 = 10$; (△) $R_0/D_0 = 1$; (□) $R_0/D_0 = 0{\cdot}1$. The results have been normalised to remove the initial 5 to 10% nuclease resistance and the final 20 to 30% nuclease sensitivity.

(b) Theoretical time-courses for hydridization of cDNA to its template RNA, assuming that the formation of hybrid molecules is a second-order reaction with a rate-constant of 1. The 4 time-courses, A, B, C and D have initial RNA to DNA concentration ratios of 100, 10, 1 and 0·1, respectively.

(Harrison *et al.*, 1974), and the remainder may be due to digestion of cDNA "tails". For cDNA^s excess, the amount remaining unannealed is dependent on the 9 S RNA concentration. There is good general agreement between these experimental time-courses and the theoretical curves (Fig. 2(b)) that were computed for the same values of R_0/D_0 using the second-order equations (2) and (3) with the rate constant $k = 1$. In particular, it is evident that, as predicted by equations (4), (5) and (6), the slowest reaction is that for which $R_0/D_0 = 1$.

The effect of the initial concentration ratio, R_0/D_0, on the rate of hybridization is shown in Figure 3, where it can be seen that equal amounts of 9 S RNA and cDNA^s react with a $D_0t_{\frac{1}{2}}$ of approximately $5{\cdot}5\times10^{-3}$ mol l^{-1} s. From equation (5) we have the rate constant $k = 1/D_0t_{\frac{1}{2}} = 182$ l mol^{-1} s^{-1}. Furthermore, the experimental data fit the theoretical curve, also shown in Figure 3, which was derived from equations (4), (5) and (6) using $k = 182$ l mol^{-1} s^{-1}. Hence these results confirm our assumption that cDNA–9 S RNA hybridization can be described adequately by second-order kinetics.

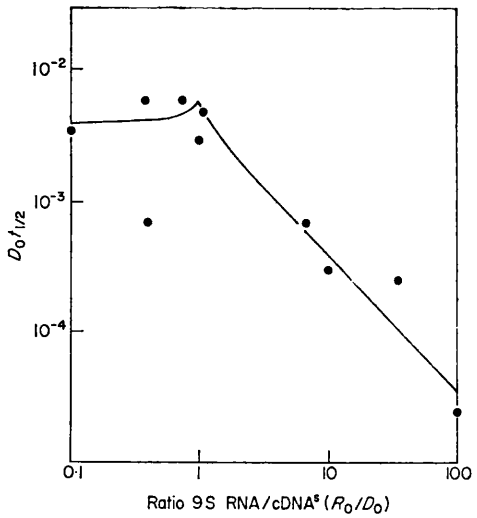


FIG. 3. Experimental variation of $D_0t_{\frac{1}{2}}$ with the 9 S RNA/cDNA^s ratio. The superimposed curve is theoretically derived, assuming that cDNA^s–9 S RNA hybridization is an ideal second-order reaction with a rate-constant of 182 l mol^{-1} s^{-1}.

(c) *Renaturation of complementary DNA*

When cDNA^s was incubated to 0·1 mol l^{-1} s, (Fig. 4) only 5% was found to be resistant to S1 nuclease, thus confirming that cDNA^s contained few complementary sequences and little secondary structure. However, when denatured cDNA^d was incubated to 0·05 mol l^{-1} s, 65% was resistant to S1 nuclease, and hence this preparation of cDNA included a large proportion of complementary sequences. As shown in Figure 4, renaturation of the cDNA^d followed a typical DNA C_0t curve, with a $C_0t_{\frac{1}{2}}$ of 10^{-3} mol l^{-1} s.

It cannot necessarily be assumed that cDNA^s and cDNA^d are transcribed from the same RNA sequences. The addition of cDNA^s to an equal amount of denatured cDNA^d will only increase the rate of renaturation if the cDNA^d sequences are represented in cDNA^s. It can be shown from equation (4) that, if cDNA^d consists solely

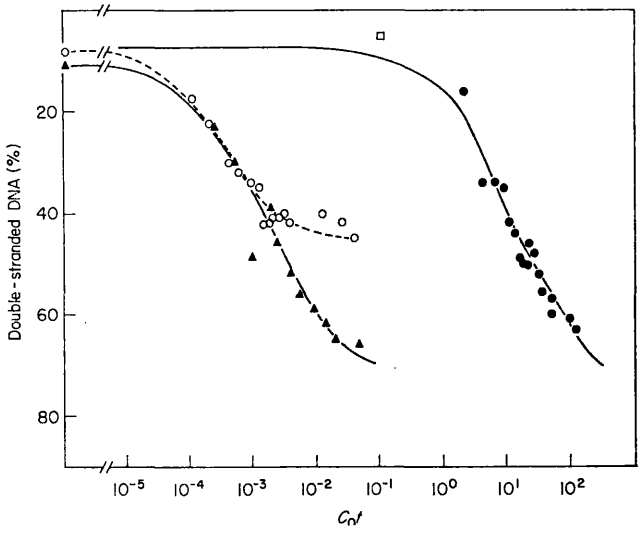


FIG. 4. The rates of DNA reassociation at 43°C were determined using the S1 nuclease assay. Labelled *E. coli* DNA (●) and cDNA^d (▲) were denatured and allowed to reanneal. cDNA^s (□) that had been denatured showed little double-stranded structure by a C_0t of 0·1 mol l^{-1} s. A mixture of equal proportions of cDNA^s and cDNA ^{d} was denatured and allowed to reanneal (○).

of equal proportions of cDNA^s sequences and their anti-strand sequences, the $C_0t_{\frac{1}{2}}$ (C_0 = concentration of cDNA^d) will be reduced to 0·25 of its former value. The renaturation curve shown in Figure 4 indicates that the addition of an equal amount of cDNA^s to cDNA^d reduces the $C_0t_{\frac{1}{2}}$ from 10^{-3} mol l^{-1} s to $0{\cdot}3\times10^{-3}$ mol l^{-1} s. This result suggests that a large proportion of the sequences transcribed into cDNA^d are also present in cDNA^s. A more precise interpretation is not possible because of the difference in base sequence complexity observed for cDNA^s and cDNA^d (see below).

The renaturation of *E. coli* DNA (mean denatured size 330 nucleotides) was followed using the S1 nuclease assay (Fig. 4) in order to provide a standard against which the reassociation of cDNA could be compared. The $C_0t_{\frac{1}{2}}$ was estimated to be 7 mol l^{-1} s.

(d) *Single-stranded complementary DNA–9 S RNA titration curves*

The titration of a fixed amount of cDNA^s with increasing amounts of 9 S RNA can be used to determine whether cDNA^s consists of copies of all or some of the classes of molecules present in 9 S RNA. The results of several titration experiments are shown in normalized form in Figure 5, where it can be seen that the titration curves are similar whether the hybrid is assayed by resistance to S1 nuclease or by banding in CsCl. Normalization corrected for about 5% of self-annealing in cDNA^s with no RNA present and about 15 to 20% of the cDNA^s that does not hybridize, even in large 9 S RNA excess. The concentration of cDNA^s used for each point on the titration curve was determined from the specific activity of cDNA^s, which in turn was estimated from the specific activity of the nucleotide precursor, assuming a 50% (C + T) content in cDNA^s. With such small amounts of cDNA^s it is not possible to confirm independently the assumed specific activity, and hence any errors in our estimate will affect the RNA/cDNA^s ratios. This ratio is similarly affected by errors in the RNA concentrations, which were obtained by serial dilution. A theoretical

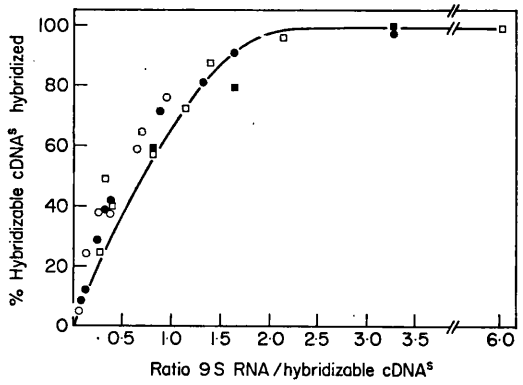


FIG. 5. Titration of cDNA^s with 9 S RNA prepared by the poly(U)-Sepharose method. The hybrid was assayed by resistance to S1 nuclease (open symbols) or by banding in CsCl (closed symbols) and the results have been normalized to include only hybridizable cDNA^s (about 80% of total cDNA^s). Each set of symbols represents a separate experiment. The solid line indicates the theoretical titration curve calculated from the size distribution of cDNA^s.

titration curve was computed by applying equations (13) and (14) to the size distribution of cDNA^s (Harrison *et al.*, 1974) and assuming (a) that all the 9 S RNA molecules (approximate size 570 nucleotides) remain larger than the cDNA molecules, and (b) that all classes of 9 S RNA molecules are transcribed into cDNA^s. The close agreement between experiment and theory (Fig. 5) indicates, first, that our experimental errors in the RNA/cDNA^s ratios are minimal and, second, that virtually all classes of molecules in 9 S RNA are transcribed. We discuss evidence below (see Discussion) that transcription is partial and non-random in the sense that only a particular region of each class is transcribed.

The conclusions drawn from such titration experiments are valid only if all reactions used to establish the plots are complete. It is, therefore, an important prediction of second-order kinetics that the slowest reaction, measured in terms of D_0t, occurs when equal amounts of the two components are present. The data in Figure 3 indicate that the slowest cDNA–RNA reaction has $D_0t_{\frac{1}{2}} = 5{\cdot}5\times10^{-3}$ mol l^{-1} s and, hence, incubation to 40 times this value (0·22 mol l^{-1} s) should ensure completion of all hybridization reactions. In order to establish that sufficiently large incubation periods were used, two titration curves using reticulocyte 9 S RNA were obtained by annealing to $D_0t = 0{\cdot}22$ and $D_0t = 1{\cdot}7$. The titration curves were identical, and it is apparent that incubation to $D_0t = 0{\cdot}22$ (used for results in Fig. 5) is sufficient to ensure completion of the hybridization reactions. This result also implies that RNA degradation is not a problem during hybridization up to $D_0t = 1{\cdot}7$. This is important, because it is assumed for the theoretical analysis that all RNA molecules remain larger than the largest cDNA molecules. It has already been demonstrated (Harrison *et al.*, 1974) that no appreciable degradation of cDNA^s occurs during hybridization.

The theoretical analysis is based on the assumption that, in cDNA^s excess, there is a greater probability of hybridization to the larger fragments of cDNA^s. Direct evidence for this assumption has been obtained by determining the mean sizes of cDNA^s hybridized at 9 S RNA/cDNA^s ratios of 0·3 and 3. It was found that the mean size of cDNA^s hybridized at a ratio of 0·3 was 60% larger than that for a ratio

of 3. Finally, in order to demonstrate the effect of cDNA^s size, two different size ranges of cDNA^s were used to establish the titration curves shown in normalized form in Figure 6. As would be expected theoretically, the cDNA^s with the larger mean size of 210 nucleotides produced a "steeper" titration curve than that for the smaller cDNA^s (150 nucleotides).

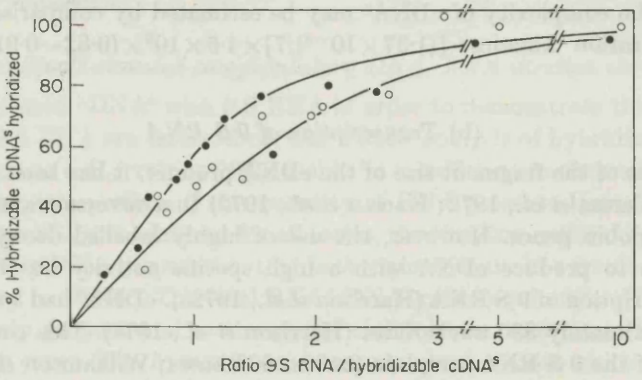


FIG. 6. Titration of cDNA^s of mean size 210 nucleotides (●) and of mean size 150 nucleotides (○) with 9 S RNA.

5. Discussion

(a) *Estimation of complexity of single-stranded and double-stranded complementary DNA*

With appropiate correction factors for fragment size and base composition, the rates of renaturation and hybridization may be used to estimate the base sequence complexity of cDNA^d and cDNA^s. Under identical conditions, the $C_0t_{\frac{1}{2}}$ of *E. coli* DNA (mean size 330 base pairs) was 7 mol l^{-1} s, and the $C_0t_{\frac{1}{2}}$ of cDNA^d (mean size 105 base pairs) was 10^{-3} mol l^{-1} s. According to the relation between DNA chain-length and rate of reassociation determined by Wetmur & Davidson (1968), the $C_0t_{\frac{1}{2}}$ value of a DNA with a mean size of 330 nucleotides is 0·56 times the $C_0t_{\frac{1}{2}}$ value of DNA of mean size 105 nucleotides. Although there is some disagreement in the literature on the effect of base composition on the rate of renaturation (Bak *et al.*, 1969; Gillis *et al.*, 1970), we have used a correction factor of 1·8% of the observed $C_0t_{\frac{1}{2}}$ for every mole per cent (G + C) difference from 51% (G + C) (Seidler & Mandel, 1971; Wetmur & Davidson, 1968). On this basis, the $C_0t_{\frac{1}{2}}$ value of a DNA with a base composition of 50% (G + C) is 0·91 times the $C_0t_{\frac{1}{2}}$ value of a DNA with 45% (G + C), such as cDNA^d. Hence, taking the complexity of *E. coli* DNA as $4{\cdot}5\times10^6$ base pairs (Cairns, 1963) and applying both correction factors, the complexity of cDNA^d may be estimated as: $(10^{-3}/7)\times4{\cdot}5\times10^6\times0{\cdot}56\times0{\cdot}91=330$ nucleotide pairs.

The kinetics of the hybridization reaction between cDNA^s and 9 S RNA fit second-order reaction theory. However, in order to estimate the complexity of cDNA^s, it is necessary to convert the DNA–RNA reaction to the equivalent DNA–DNA reaction of the same complexity. By hybridizing RNA synthesized *in vitro* to excess template DNA, Bishop (1972) has estimated the rate-constants for both hybridization and renaturation under the same conditions. Interpolation of the data provided suggests that the ratio of hybridization rate-constant to renaturation rate-constant

at 43°C in 3 × SSC/50% formamide is approximately 0·25 (SSC is 0·15 M-NaCl, 0·015 M-sodium citrate). If this value is applicable to 9 S RNA–cDNA^s hybridization, the $C_0t_{\frac{1}{2}}$ value of the equivalent DNA–DNA reaction of the same complexity and base composition may be estimated as $0{\cdot}25/182 = 1{\cdot}37 \times 10^{-3}$ mol l^{-1} s. cDNA^s had a mean fragment size of 330 nucleotides and a base composition of 40 to 45% (G + C) (Harrison *et al.*, 1974). On the assumption that no correction factor for size is required, the complexity of cDNA^s may be estimated by comparison with *E. coli* DNA renaturation kinetics: $[(1{\cdot}37 \times 10^{-3})/7] \times 4{\cdot}5 \times 10^6 \times (0{\cdot}82 - 0{\cdot}91) = 700$ to 800 nucleotides.

(b) *Transcription of 9 S RNA*

On the basis of the fragment size of the cDNA product, it has been reported (Ross *et al.*, 1972; Verma *et al.*, 1972; Kacian *et al.*, 1972) that reverse transcriptase copies all or most globin genes. However, the use of highly labelled deoxyribonucleoside triphosphates to produce cDNA with a high specific activity has been shown to inhibit transcription of 9 S RNA (Harrison *et al.*, 1972*b*). cDNA^s had a mean fragment size of approximately 330 nucleotides (Harrison *et al.*, 1974). This could imply that about 60% of the 9 S RNA template (540 to 600 bases; Williamson *et al.*, 1971) was transcribed or that fragmentation of cDNA to a mean size of 330 nucleotides occurred after complete transcription. However, our results indicate that cDNA^s was transcribed from a total RNA sequence of between 700 and 800 bases, i.e. longer than a single mRNA molecule, suggesting that cDNA^s consists of partial copies of two different mRNAs, giving a total complexity of 700 to 800 bases, with a mean fragment size of 330 bases. Since we have also shown, by titration experiments, that cDNA^s consists of copies of virtually all classes of molecules present in 9 S RNA, it is probable that cDNA^s includes transcripts of both α and β mRNAs. If cDNA transcripts of α and β mRNA are sufficiently different to prevent cross-reaction (Dayhoff, 1972), the most likely interpretation of our results is that cDNA^s comprises partial transcripts of both α and β globin mRNA. Furthermore, the requirement for poly(dT) primer for 9 S RNA transcription and the presence of adenine-rich sequences at the 3′ terminal (Burr & Lingrel, 1971; Lim & Canellakis, 1970) strongly suggests that such partial transcripts are initiated at the 3′ end of the mRNA.

The evidence concerning the transcription of cDNA^d from 9 S RNA is less conclusive. Because of the relatively low complexity of cDNA^d (about 330 base pairs), it is possible that cDNA^d consists of sequences (mean size 105 nucleotides) transcribed randomly from a restricted region (330 bases) of a single mRNA. However, the increase in the rate of renaturation of cDNA^d that is caused by the addition of cDNA^s, suggests that most of the cDNA^d sequences are represented in the same ratio as in cDNA^s. It is more likely, therefore, that cDNA^d consists of partial copies of more than one mRNA.

When reverse transcriptase product is treated with ribonuclease and S1 nuclease, the mean size of denatured cDNA fragments is reduced from 240 to 105 nucleotides. This suggests that the nuclease-resistant parts of cDNA duplexes are shorter than the nuclease-sensitive cDNA molecules, and that the former also contain single-stranded cDNA tails. This result, together with our estimates of complexity and fragment size, is consistent with the conclusion of McDonnell *et al.* (1970), that in the absence of actinomycin D, transcription of RNA into DNA is a two-stage process. In our conditions, the first stage would appear to involve the synthesis of a cDNA

strand complementary to about 60% of the mRNA template with which it forms a hybrid duplex. The second stage involves the synthesis, on this hybrid molecule, of a cDNA strand complementary to about 30% of the first cDNA transcript. Treatment with ribonuclease and S1 nuclease would remove hybrid molecules and single-stranded cDNA tails, leaving only the relatively short $cDNA^d$ duplexes. The presence of actinomycin D appears to prevent the second stage of transcription, resulting in larger $cDNA^s$ molecules.

(c) *Single-stranded complementary DNA–RNA titration curves*

We have titrated $cDNA^s$ with 9 S RNA in order to demonstrate that all classes of molecules in 9 S RNA are transcribed. Our kinetic analysis of hybridization indicates that incubation to a D_0t value of 0·22 mol l^{-1} s is sufficient to ensure the completion of all reactions, irrespective of the quantity of RNA present. Hence, provided this minimum value of D_0t is achieved for all ratios, the titration method can equally well be applied to any RNA preparation that is thought to contain a fraction of sequences complementary to $cDNA^s$. The total RNA/$cDNA^s$ ratio required for 100% hybridization is inversely proportional to the size of this fraction, and hence the $cDNA^s$–9 S RNA titration curve (Fig. 5) can be used as a standard against which other titration curves may be compared. Thus, it is merely necessary to estimate the abscissa scale factor required to superimpose the titration curve on the standard $cDNA^s$–9 S RNA titration curve. This technique provides a sensitive assay for the presence of globin mRNA in any RNA preparation.

The authors are indebted to Dr J. W. Beard, Duke University, U.S.A., for supplies of avian myeloblastosis virus and to Miss E. MacPhail and Mr A. McPhee for technical assistance. This work was supported by grants from the Medical Research Council and the Cancer Research Campaign.

REFERENCES

Bak, A. L., Black, F. T., Christiansen, C. & Freundt, E. A. (1969). *Nature (London)*, **224**, 1209–1210.
Birnstiel, M. L., Sells, B. M. & Purdom, I. F. (1972). *J. Mol. Biol.* **63**, 21–39.
Bishop, J. O. (1969). *Biochem. J.* **113**, 805–811.
Bishop, J. O. (1972). *Biochem. J.* **126**, 117–185.
Bishop, J. O. & Rosbash, M. (1973). *Nature New Biol.* **241**, 204–207.
Britten, R. J. & Kohne, D. E. (1968). *Science*, **161**, 529–540.
Burr, M. & Lingrel, J. B. (1971). *Nature New Biol.* **233**, 41–43.
Cairns, J. (1963). *Cold Spring Harbor Symp. Quant. Biol.* **28**, 43–46.
Dayhoff, M. O. (1972). Editor of *Atlas of Protein Sequence and Structure*, vol. 5, National Biomedical Foundation, Washington, D.C.
Gillis, M., DeLey, J. & DeCleen, M. (1970). *Eur. J. Biochem.* **12**, 143–153.
Harrison, P. R., Hell, A., Birnie, G. D. & Paul, J. (1972*a*). *Nature (London)*, **239**, 219–221.
Harrison, P. R., Hell, A. & Paul, J. (1972*b*). *FEBS Letters*, **24**, 73–76.
Harrison, P. R., Birnie, G. D., Hell, A., Humphries, S., Young, B. D. & Paul, J. (1974). *J. Mol. Biol.* **84**, 539–554.
Hell, A., Birnie, G. D., Slimming, T. K. & Paul, J. (1972). *Anal. Biochem.* **48**, 369–377.
Housman, D., Pemberton, R. & Taber, R. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 2716–2719.
Kacian, D. L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L. & Marks, P. A. (1972). *Nature New Biol.* **235**, 167–169.
Laird, C. D. & McCarthy, B. J. (1968). *Genetics*, **60**, 303–322.
Laird, C. D., McConaughy, B. L. & McCarthy, B. J. (1969). *Nature (London)*, **224**, 149–154.

Lanyon, G., Paul, J. & Williamson, R. (1972). *Eur. J. Biochem.* **31**, 38–43.
Lim, L. & Canellakis, E. S. (1970). *Nature (London)*, **227**, 710–712.
Linn, S. & Lehman, I. R. (1965). *J. Biol. Chem.* **240**, 1287–1293.
Lockard, R. E. & Lingrel, J. B. (1969). *Biochem. Biophys. Res. Commun.* **37**, 204–212.
Martinson, H. G. (1973*a*). *Biochemistry*, **12**, 139–145.
Martinson, H. G. (1973*b*). *Biochemistry*, **12**, 145–150.
Mathews, M., Osborn, M. & Lingrel, J. (1971). *Nature New Biol.* **233**, 206–209.
McDonnell, J. P., Garapin, A., Levinson, W. E., Quintrell, N., Fanshier, L., & Bishop, J. M. (1970). *Nature (London)*, **228**, 433–435.
Nygaard, A. P. & Hall, B. D. (1964). *J. Mol. Biol.* **9**, 125–142.
Packman, S., Aviv, H., Ross, J. & Leder, P. (1972). *Biochem. Biophys. Res. Commun.* **49**, 813–819.
Ross, J., Aviv, H., Scolnick, E. & Leder, P. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 264–268.
Seidler, R. J. & Mandel, M. (1971). *J. Bacteriol.* **106**, 608–614.
Ullmann, J. S. & McCarthy, B. J. (1973*a*). *Biochim. Biophys. Acta*, **294**, 405–415.
Ullmann, J. S. & McCarthy, B. J. (1973*b*). *Biochim. Biophys. Acta*, **294**, 416–424.
Verma, I. M., Temple, G. F., Fan, H. & Baltimore, D. (1972). *Nature New Biol.* **235**, 163–167.
Wetmur, J. G. & Davidson, W. (1968). *J. Mol. Biol.* **31**, 349–370.
Williamson, R., Morrison, M., Lanyon, G., Eason, R. & Paul, J. (1971). *Biochemistry*, **10**, 3014–3021.
Williamson, R., Drewienkiewicz, C. E. & Paul, J. (1973). *Nature New Biol.* **241**, 66–68.