STUDIES OF POST-TRANSCRIPTIONAL GENE REGULATION IN RAT LIVER AND HEPATOMA CELLS

by

HOWARD TREVOR JACOBS

Thesis submitted for the degree of Doctorate of Philosophy, in the University of Glasgow, being an account of research conducted at the Beatson Institute for Cancer Research, Glasgow.

November 1980

TABLE OF CONTENTS

Page No.

Page No.

Page No.

Page No.

Page No.

Page No.

NOTE

Some of the work described in this thesis has already been published, and reprints are enclosed with the bound copies.

ABBREVIATIONS

HTC - Hepatoma tissue-culture cells

C,A,G,T,U - Cytosine, Adenine, Guanine, Thymidine and Uracil of their equivalent nucleotides incorporated into DNA or RNA

poly(A), poly(U) - polyadenylic acid, polyuridylic acid

oligo(A) - oligoadenylic acid

oligo(dT) - oligodeoxythymidylic acid

mRNA - messenger RNA

hnRNA - heterogeneous nuclear RNA

rRNA, tRNA - ribosomal RNA, transfer RNA

RNP, mRNP - ribonucleoprotein, messenger ribonucleoprotein

DNase, RNase - deoxyribonuclease, ribonuclease

cDNA - complementary DNA

tDNA - DNA encoding tRNA

u.v. - ultra-violet light

DMSO - dimethyl sulphoxide

DRB - 5,6 dichloro-1-β-ribofuranosylbenzimidazole

EDTA - Ethylenediaminetetraacetic acid

DTT - Dithiothreitol

HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulphonate

Tris - 2-amino-2-hydroxymethyl-1,3-propanediol

SDS - Sodium dodecyl sulphate

HAP - Hydroxylapatite

TCA - Trichloroacetic acid

ER - Endoplasmic reticulum

TEMED - N,N,N',N'-tetramethylethylenediamine

$poly(A)^+$ - polyadenylated (poly(A)-containing)

$poly(A)^-$ - non-polyadenylated

$C_o t, R_o t$ - product of DNA, RNA concentration (in moles of nucleotide per litre) and time (in seconds)

$C_o t_{1/2}$,
$R_o t_{1/2}$ - $C_o t$, $R_o t$ values at which half-maximal extent of hybridisation is reached

T_m - melting temperature (mid-point of the temperature/absorbance profile of a nucleic acid)

OD_{260},
OD_{280} etc. - optical density at 260, 280 nm etc.

ATP - adenosine triphosphate

AMP.P(NH)P- β,γ -imido-substituted analogue of ATP

cAMP - 3',5' cyclic adenosine monophosphate

SSC - 0.15 M NaCl, 0.015 M Sodium citrate

LIST OF TABLES

LIST OF FIGURES

LIST OF CHARTS

ACKNOWLEDGEMENTS

I would like to thank Dr John Paul and the staff of the Beatson Institute for the provision of research facilities, and the Lady Tata Memorial Trust for financial support in the form of a scholarship. The Beatson Institute has been supported by M.R.C. and C.R.C.

I am particularly indebted to my supervisor, Dr George Birnie, for his invaluable advice and encouragement, to Anne Sproul, Lesley Frew and Diana Morgan for my initial training, to Rosemary Shott for permission to use her cDNA clones prior to publication, and to my fellow students in numerous ways. I am also very grateful to Mrs Lorna Peedle for her expeditious typing of the manuscript, and to David Tallach for the artwork.

SUMMARY

Studies of the composition of cellular RNA populations have indicated that very considerable shifts in phenotype, such as maturation along a given differentiation pathway, the induction or cessation of growth, the expression of specialised physiological functions, or carcinogenic transformation, are marked by essentially quantitative rather than qualitative changes in gene expression. Furthermore, these phenotypic alterations seem to be brought about by a complex interaction of regulatory events, operating at multiple sites in the pathway of mRNA synthesis. The rates of transcription, post-transcriptional processing, nucleocytoplasmic transport, cytoplasmic mRNA turnover and recruitment into translationally active polysomes all appear to be subject to sequence-selective controls whose specificity varies according to the physiological context.

In this study, the poly(A)$^+$ mRNA and hnRNA populations of rat liver and hepatoma (HTC) cells have been compared by various hybridisation methodologies (single-copy DNA saturation, kinetics of cDNA cross-hybridisation, titration against cloned cDNAs) in order to characterise the extent to which post-transcriptional controls determine the overall pattern of relative mRNA abundances. The results indicate that the two cell-types transcribe essentially the same sequences, and express the same set of genes on their polysomes. However, their patterns of relative mRNA abundances are radically different. Liver expresses a set of superabundant, presumably differentiation-specific mRNAs, encoding both secretory and intracellular products, which are orders of magnitude rarer in HTC cells. No such dramatic difference in mRNA frequencies is evident in the opposing direction; however, a considerable number of mRNAs at intermediate abundance in liver are modestly increased in concentration in the hepatoma. The superabundant liver mRNAs which are depleted in HTC cells appear to be regulated mainly at a post-transcriptional level,

since the disparity in their concentrations in hnRNA in the two cell-types is much less striking (3-fold, on average, as compared with 100-fold at the level of polysomal mRNA). Two other aspects of post-transcriptional selection in rat liver emerge from the results. A large proportion of the poly(A)-adjacent sequence complexity of hnRNA is unrepresented on the polysomes. This may represent a qualitative control mechanism, defining the characteristic expressed "gene set" of a committed cell, a control which seems to function unaltered in the hepatoma. Liver hnRNA also contains a population of relatively abundant poly(A)-adjacent sequences whose levels on the polysomes appear to be post-transcriptionally suppressed.

In order to obtain more detailed information on the levels at which mRNA abundances are generated and modulated between cell-types and to gain insight into the mechanisms responsible, it is necessary to develop in vitro systems which maintain the physiological specificity of gene regulation in vivo. The clear-cut, post-transcriptionally determined differences in mRNA abundance, both within and between the cell-types considered here, provide a set of convenient and sensitive assays for the degree to which post-transcriptional selectivity is preserved in isolated nuclei.

A system was therefore developed, based on detergent-treated HTC cell nuclei, incubated in the presence of cytosol protein. The medium used for the incubations supported ATP- and cytosol-dependent transport of prelabelled mRNA and rRNA and maintained nuclear integrity. The composition of the population of polyadenylated molecules transported from (unlabelled) nuclei was investigated by various hybridisation methodologies. Despite low levels of both non-specific leakage, and possible release of residual adherent RNA of cytoplasmic origin, the bulk of in vitro-transported

poly(A)$^+$ RNA could not be explained as being due to either of these artifacts. Its pattern of relative abundances indicated that highly sequence-selective processing and/or transport mechanisms are able to operate *in vitro*. Liver cytosol was found to have no effect on the specificity of mRNA transport from isolated HTC cell nuclei, even when highly sensitive assays (both heterogeneous and cloned cDNAs) for post-transcriptionally regulated messengers were used. This cannot be formally interpreted until the precise degree of physiological equivalence of the system is established which must await a more precise characterisation of the relative contributions of intranuclear and cytoplasmic events to polysomal abundances *in vivo*, using cloned cDNAs.

The results are discussed in terms of their implications for post-transcriptional gene control mechanisms, in relation to development, growth, differentiation and carcinogenesis.

1. PREFACE

A major goal of molecular biology is the elucidation of the mechanisms of selective gene expression. In eukaryotes, where structural genes are separated from the site of their expression (polypeptide synthesis) by a cascade of physical and functional barriers, there exist many loci at which gene expression could be, and is, regulated. Such an arrangement allows for considerable biological sophistication in the response of individual cells to changes in their immediate environment, and in their ability to participate effectively in the integrated responses of the whole organism. It also facilitates a developmental programme which is both orderly and flexible. This complexity, however, greatly complicates the task of molecular biology in identifying those levels in the pathway of cellular information flow at which controls actually operate, how they are inter-related, and by what mechanisms they are brought about.

These questions, and the contribution to their resolution of molecular hybridisation methodology, are the broad subject matter of this thesis. More particularly, attention is focussed on gene regulation at the post-transcriptional level; its nature and extent and the mechanisms by which it may be brought about. In the first instance this involves studies of the informational content of nucleic acid sequences found in various cellular compartments *in vivo*. Not only does this directly provide data relating to the levels at which gene controls are exerted. It also supplies assays for the degree to which the specificity of gene control mechanisms can be accurately reproduced *in vitro*. Adequately characterised *in vitro* systems are an essential tool for the detailed investigation of these mechanisms. The experimental work which will be described consists of an analysis of cellular RNA populations in a biological system considered potentially fruitful for studies of post-transcriptional control, namely

rat liver and hepatoma cells, and the application of the findings to measuring the physiological equivalence of RNA processing and transport in isolated hepatoma nuclei, with a view to defining the influence of tissue-specific factors on post-transcriptional regulation.

The choice of this particular experimental system was to some extent arbitrary, being governed largely by technical features which make it particularly suitable for this type of investigation. The results are discussed, therefore, against a wider backcloth of findings in disparate biological systems, emphasising their relevance to gene control mechanisms in general. The justification for this derives from the observation that the principles underlying gene organisation and expression amongst metazoans and metaphytans appear to be quite general.

In order to place the experimental work in its proper context, the Introduction is set out as follows. Firstly, the various steps in the pathway of the expression of a eukaryotic gene will be briefly surveyed, together with the evidence from genetic, biochemical, cell biological and ultrastructural studies that any of them might function to control relative gene activity. The results of molecular hybridisation studies of cellular RNA populations will then be presented, following a concise description of the technologies used. The significance of these findings will be discussed in terms of the relative importance of transcriptional and post-transcriptional controls in determining the qualitative and quantitative composition of polysomal mRNA, and in terms of how altered specificity at various levels may correlate which particular changes in phenotype. The next section sets out to evaluate the various strategies which have been adopted, in the construction of <u>in vitro</u>

systems capable of reproducing the specificity of the gene control mechanisms inferred from in vivo studies. Particular emphasis will be placed on the use of hybridisation assays, and their suitability for this characterisation. Finally, the rationale for the choice of this particular experimental system will be presented in somewhat greater detail.

2. INTRODUCTION

2.1 Levels of control of eukaryotic gene expression

The steps in the synthesis of a polypeptide encoded in the nuclear DNA of a eukaryote may be traced as follows:- firstly a specific region of DNA is transcribed, the efficiency of which depends on the primary sequence of the DNA as well as upon its structural configuration, and the specificity of the transcription machinery; secondly, the RNA transcript undergoes a series of processing reactions which alter its covalent structure, and which result in its interaction with proteins and perhaps other RNA species; thirdly, the processed transcript is translocated to the cytoplasm where, finally, it is translated on the polysomes until it is degraded by the action of nucleases. Since the experimental work which will be described is principally in the area of post-transcriptional control, transcription and post-transcriptional events will here be considered separately.

2.1.1. Transcriptional control

Control at the level of the DNA sequences available for transcription can be envisaged as taking various forms. An extreme model would postulate somatic re-organisation, including gene deletions, amplifications and re-arrangements as being responsible for differential gene activity during development. A number of classical experiments would seem to rule out such somatic plasticity of the genome as a general mechanism. The regeneration of entire plants from single cells derived from differentiated tissues (1), or the re-programming of somatic nuclei from adult frogs by serial transplantation into enucleated eggs (2), indicate that the germ-line state of the DNA is preserved in many cell-types during development. The low frequency of such re-programming

events, however, implies that differentiation does involve profound alterations in the structural organisation of the genome, which are not easily reversed. These may include, for example, the pattern of base modifications, the packing of the DNA, or its interaction with regulatory macromolecules.

Hybridisation studies, of the frequency and organisation in different tissues of particular structural genes, have confirmed that gene re-arrangements are used only infrequently as a mechanism of selective gene control. For example, the restriction patterns of rat albumin and α-fetoprotein genes in liver and hepatoma are indistinguishable (3). However, isolated examples of gene amplification may be cited, both in normal development, for example ribosomal genes during amphibian oogenesis (4) or chorion genes in Drosophila (5), and in response to a selective pressure, such as applies in the case of the dihydrofolate reductase gene in mouse cells cultured in the presence of methotrexate (6). A number of instances of gene re-arrangement are also known. Pathogenic trypanosomes are able to express novel surface antigens by duplicating the relevant gene and transposing it into an expression site (7). A similar mechanism is responsible for mating type selection in yeast (8). Deletions during lymphocyte differentiation are responsible for the clonal expression of a given pair of immunoglobulin variable region genes, by means of somatic recombinations which transpose them adjacent to constant region genes of the appropriate class, forming single transcriptional units (9). Such mechanisms are also associated with the progressive switching in immunoglobulin heavy chain class (10, 11), and in part with the generation of antibody diversity, by virtue of the 'stutter' in the joining of V_H, D, and J_H gene segments (12, 13). Translocations involving palindromic elements during sea urchin embryogenesis (14) and gene amplifications during chick

development (15) have been postulated on the basis of DNA re-association studies.

In interphase nuclei, DNA is compacted into an orderly structure, chromatin (16), by virtue of its interaction with a protein matrix (17), and at a lower order of structure is wound around histone core-particles called nucleosomes (18). Both biochemical and ultrastructural studies indicate that gross changes in this orderly structure accompany alterations in the pattern of transcription. Pulse-labelling and electron microscopy have shown that highly condensed regions (hetero-chromatin) are transcriptionally inert (19), whilst actively transcribed sequences differ in their arrangement of nucleosomes from untranscribed sequences, both in amphibia (20) and insects (21). Structural changes in chromatin are correlated with hormonally-induced developmental changes in gene expression in insects (22): the induction of new transcriptional activity is accompanied by the formation of new puffs at specific loci in polytene chromosomes (23), where RNA polymerase, as detected by immunofluorescence, rapidly becomes re-located (24). Puffs induced by heat-shock in Drosophila are the sites of specific transcription of heat-shock mRNAs, as shown by <u>in situ</u> hybridisation (25-27).

The differential nuclease-sensitivity of expressed and un-expressed genes supports the idea of gross structural controls on transcription. For example, hybridisation studies have shown that DNAse I preferentially digests globin genes in hen erythrocyte nuclei but not oviduct nuclei (28), and ovalbumin genes in oviduct but not erythrocytes (29). The assay may be made semi-quantitative by examining the rate of digestion of specific restriction bands in a blot hybridisation (30, 31), which has indicated that ovalbumin in oviduct is more susceptible to digestion that globin in erythrocytes (30). Genes in the expressible state also contain specific

sites hypersensitive to nucleases (32-35). These sites have been postulated as having a directive role in nucleosome phasing, allowing recognition of promotor sites (35-37). The property of preferential digestion applies to the entire set of sequences transcribed into nuclear RNA (28, 38), including the very rarest mRNA sequences (39, 40), which in chick oviduct are similar in sensitivity to ovalbumin (40). The switch from embryonic to adult β-globin expression in chickens is accompanied by a decrease in the nuclease-sensitivity of the embryonic gene, whilst the adult β-globin gene is nuclease-sensitive throughout development (31). The local configuration of chromatin may therefore earmark genes as 'transcribable', perhaps during cell commitment, with the activation of a particular transcriptional promotor during subsequent cellular maturation. Hence, the globin genes are equally DNase I-sensitive before or after DMSO-induction of globin mRNA synthesis in Friend erythroleukaemia cells (41), or indeed, in non-inducible, and even non-erythroid hematopoietic cells (41). The use of DNase II as a probe, however, has given a contrary finding (42). DNAse I-sensitivity appears to be conferred by a class of non-histone chromosomal proteins, specifically high mobility group proteins 14 and 17 (43), which are capable of specific interaction with 'active' nucleosomes, in chromatin reconstruction (44). Transcriptionally active genes are also associated with a more highly acetylated histone fraction (45, 46), which may facilitate their recognition by the high mobility group proteins. The length of nucleosome-associated DNA however is invariant within a given cell-type, between transcriptionally active and inactive regions of the genome (47). DNAse I preferentially releases at least one protein other than RNA polymerase from Drosophila puffs (48). Specific protein-DNA interactions involved in this coarse level gene control are also implied by the finding that unique, messenger-complementary sequences are highly enriched

in the fraction of sheared rat liver DNA which associates *in vitro* with non-histone proteins from liver (49). Many of the effects of steroid hormones on avian and mammalian gene activity appear to involve local modifications in chromatin structure (50), leading to altered transcription. For example, there is an increase in the number of sites available for initiation by exogenous RNA polymerase following estrogen treatment of chick oviduct (51). The enhancement of chromatin properties indicative of transcriptional activity (46,52,53,54), by *in vitro* histone acetylation, is further evidence for the dependence of transcription on chromatin structure.

Sensitivity to nucleases extends well beyond the boundaries of individual transcriptional units (55), supporting the idea that chromatin is organised into large transcriptional 'domains' subject to coarse regulation. This hypothesis is also favoured by the observation that some groups of structurally and/or functionally related genes are clustered in the genome often in the order of their developmental expression such as the human β-like globins (56), three chick genes (ovalbumin, 'X' and 'Y') co-ordinately expressed under hormonal control in stimulated oviduct (57), and at least one group of Drosophila heat-shock genes (58). A domain theory is an attractive explanation for the phenotypic manifestations of some of the thalassemia lesions; for example, in β thalassemia, a deletion in the human β-like globin locus far from the β-globin gene suppresses all transcription of the β gene, as well as the γ and δ genes. Co-ordinately regulated genes in Aspergillus (59), and co-expressed families of silkmoth chorion genes (60) are also clustered. However, these chromosomal clusterings may represent the evolutionary histories, rather than the current modes of expression, of co-ordinately functioning genes. It

is noteworthy that other sets of simultaneously expressed genes, such as the α and β globins, can be totally unlinked (61).

Another possible gross mechanism of transcriptional control is selective methylation of cytosine residues in DNA. The extent of methylation of the chick ovalbumin and human β-like globin loci has been studied by blot hybridisation to DNA digested with methylation-sensitive restriction endonucleases, which has shown a consistently low level of methylation in tissues expressing a given gene, with a variable level in other tissues (62, 63). The sequences found in the methylated and unmethylated compartments of the sea urchin genome, however, do not alter dramatically during development, with transcribed genes found always in the unmethylated DNA (64). Beyond such a gross control exerted at the level of the transcribability of chromatin, transcription may be modulated at the initiation, elongation of termination steps of RNA synthesis, by the interaction of specific effector molecules with the DNA. Such mechanisms are well documented in prokaryotes (65-67). The variable packing of transcriptional complexes on chromatin, as visualised by electron microscopy (68), suggests that different genes are transcribed at different rates. Numerous pieces of indirect evidence support the view that switches in the pattern of gene expression may be effected by changes in relative transcription rates. For example, the RNA polymerase inhibitor α-amanitin blocks the alteration in mRNA translational template activities which occurs in imbibing wheat embryos (69), and actinomycin D prevents the induction by interferon of translation-regulatory enzymes in mouse L cells (70). Studies of the composition of cellular RNA populations (discussed in Section 2.3) also contribute powerfully to the evidence for qualitative and quantitative control of transcription.

2.1.2. Post-transcriptional control

Since almost all the evidence for post-transcriptional regulation of relative gene expression derives from hybridisation studies of RNA populations, this section will merely delineate the possible loci at which control may be exerted, and indicate a number of observations which indirectly support the notion that post-transcriptional processes play a major role in controlling gene activity.

The superinduction of glucocorticoid-inducible hepatic enzymes (71), or α2u globulin (72) by actinomycin D, or of interferon mRNA by DRB (73), both agents which block transcription, has been taken to imply that these genes are subjected normally to post-transcriptional 'repression'. This control could be exerted at any of the many processing steps on the pathway from primary transcript to mature mRNA. These include the excision of non-coding regions by trimming and/or splicing (74), the capping of the 5' end (75), poly(A) addition to the 3' end (76), association with nuclear proteins (77) and base modifications (107). The fact that a large proportion of hnRNA turns over entirely within the nucleus (more than 90 percent in the case of resting human lymphocytes (78)) strongly suggests that some transcripts are selectively processed whereas others are preferentially degraded. However, since many mRNAs are transcribed into much larger precursors, as evidenced by kinetic labelling studies (79), electron microscopic visualisation of transcription complexes (80) or u.v. target size determination for the inhibition of mRNA synthesis (81), the extent of sequence selection during processing remains uncertain.

The addition of a cap at a particular site, which is an early processing event (82) may indirectly influence the relative levels of different mRNAs, since the extent of secondary methylation at the cap-adjacent nucleotide, itself a probable

determinant of relative mRNA stability (83), varies according to the precise sequence selected for cap addition (83). The blocking of poly(A)$^+$ mRNA formation by the polyadenylation inhibitor 3' deoxyadenosine (cordycepin) (76) has been taken to indicate a selective role for poly(A) addition, although the drug may have other effects (84). The extent of conservation of poly(A), and by implication poly(A)-adjacent sequences, remains disputed (85-88), although in the case of the 3' co-terminal sets of late adenoviral mRNAs there is good evidence for poly(A) conservation between nucleus and cytoplasm (89) which, moreover, determines the relative abundances of these messenger families. It remains unclear whether poly(A)$^-$ mRNAs are processed via an alternative pathway, or whether, as is apparent in at least some cases (90) they are de-adenylated cytoplasmically. Poly(A) may also function as an indirect determinant of messenger abundance by stabilising mRNAs to which it is attached, as is demonstrable for globin mRNA and histone mRNAs (91,92) in Xenopus oocytes, although not for human interferon mRNA (93). The rate of poly(A)-shortening is correlated with the lifetime of late adenoviral mRNAs (94).

The possible importance of splicing as a control point in processing has been deduced from hybridisations and gene transfer studies, and will not be discussed in detail in this section. Genetic studies on yeast mitochondrial genes suggest that splicing efficiencies are dependent on trans-acting factors, encoded in this case within the genes affected (95). Early deletion mutants in SV40, entirely within introns, alter the relative levels of products encoded in alternatively spliced mRNAs, suggesting that the frequency of splicing depends on the structure of the transcript as a whole (96). Since the way in which an RNA chain folds depends critically upon its primary sequence (97), the exact sites of transcriptional initiation and termination and the

sites of capping, polyadenylation, and particular splicing events, may profoundly alter its stability, by exposing or shielding particular sequences from nucleolytic attack, both during processing and after export to the cytoplasm.

An additional potential control level may be the interaction of mRNA precursors with nuclear proteins, to form RNP complexes containing ubiquitous and tissue-specific proteins (77,98). The non-random organisation of RNP complexes along nascent RNP fibrils (99), which by implication is sequence-specific, suggests that the rate of formation of these complexes may govern the relative processing efficiencies of different transcripts. Most information relating to nucleo-cytoplasmic transport has come from in vitro studies (see Section 2.4.4). Once in the cytoplasm, a messenger may remain untranslated in the form of an mRNP, distinguishable from actively translated mRNPs by its protein composition (100), or it may be recruited into polysomes, or it may be degraded by nucleases. Selective translation can be envisaged as the result either of preferential stabilisation or of the action of mRNA-specific translation factors, or of variation in the availability of a common factor required in the rate-limiting step of translational initiation, as proposed by Lodish (101). Such a mechanism would exploit differences in the affinity for this factor of different mRNAs. The existence of mRNA-specific factors has been documented only in the case of myosin mRNA in chick embryonic muscles (102). Shifts in the relative efficiencies of translation of different mRNAs have been reported in the early development of the clam (103), the mouse (104) and Dictyostelium (105), and in response to the plant hormone abscissic acid in aleurone layers of germinating barley seeds (106). Other than those already mentioned, factors which may influence relative mRNA stabilities, include internal base modifications (107), interactions with membranes (109), translation efficiency (108) and sequence-specific endonucleases.

Tobin has proposed (110) that the stochastic nucleolytic degradation of a given mRNA or its precursor should be regarded as being in competition, at every step in the pathway of its expression, with further useful processing. Viewed in these terms, the rate of synthesis of any polypeptide is a complex function relating to the efficiency of each processing step, coupled with its effect on both the stability and subsequent processing efficiency of the product. A simple mathematical treatment, based on available data on the synthesis and decay rates of well-characterised mRNAs (110), leads to the conclusion that the observed differential rates of polypeptide synthesis can only be explained as the result of a cascade of selective mechanisms, operating at multiple levels in the pathway of gene expression. The detailed elucidation of exactly which steps are regulated, and their relative contributions, as well as by what mechanisms regulation is brought about, have required the development of techniques for analysing the informational composition of nucleic acid populations in the cell. The successful approach has been that of nucleic acid hybridisation.

2.2. Nucleic acid hybridisation methodology

2.2.1. The theoretical basis of hybridisation reactions

The property of single-stranded nucleic acid molecules of complementary sequence to base-pair (hybridise) with one another in solution has made possible a detailed analysis of eukaryotic gene structure and genome organisation, and of the expression of genetic information in RNA at various levels in the cell, leading to important conclusions regarding eukaryotic gene control mechanisms. Before evaluating the results of these studies it is useful to consider the theoretical basis of molecular hybridisation, the types of information which can be derived from different experimental approaches, and the limitations in interpretation of

experimental data, which are not always scrupulously respected.

The reassociation of a perfectly homogeneous preparation of denatured DNA is theoretically a second-order reaction, whose rate-limiting step is the formation of 'nucleation centres', short regions of perfectly matched duplex from which duplex formation proceeds by a rapid 'zippering-up' mechanism (111). Second-order kinetics are obeyed experimentally by dilute solutions of purified bacterial or viral DNA (111-114). The rate constant for such a reaction is a function of temperature, ionic strength, base composition, DNA fragment size, and the presence of denaturants such as formamide. Under a variety of ionic conditions optimal reaction rate is obtained at a temperature 25° below the melting temperature (T_m) of the given DNA (115), defined as the mid-point of the temperature-absorbance transition. Under these conditions, duplex formation is minimally impeded by the formation of intrastrand secondary structure. High ionic strength favours duplex formation and increases T_m by screening the phosphate groups of the polynucleotide. T_m is approximately a function of the logarithm of ionic strength (116), and is inversely related to the ionic radius of the cation used (117). At very high ionic strengths, the stabilising effect of increasing cation concentration tends to drop off. The base composition affects T_m's, and hence relative reassociation rates at a given temperature, by virtue of the enhanced stability of G-C pairs: in SSC, for example, T_m increases by 0.41° per mole per cent G-C (118). T_m is also affected by the distribution of G-C-rich tracts along the molecule, which are potential nucleation centres (111). The dependence of reassociation rate on fragment size, is due to the fact that in longer molecules each nucleation event leads to the 'zippering-up' of a larger mass of DNA (115,119). The duplex-destabilising effects of denaturants such as formamide, which lower the effective T_m, allow the rate-enhancing effects

of high ionic strengths to be exploited at much lower temperatures, significantly reducing thermal degradation of nucleic acids due to strand scission or depurination (120). Such denaturants stabilise single-stranded nucleic acids by increasing the hydrophobic character of the solvent (121). Similar arguments apply to the hybridisation of complementary RNA and DNA sequences. Since formamide stabilises single-stranded DNA more than single-stranded RNA, high formamide conditions may be used to favour RNA-DNA hybridisation over DNA-DNA re-association (122).

One of the main causes of deviation from ideal second-order kinetics is aggregation which occurs at high DNA or RNA concentrations, and at high ionic strengths (123). The apparent decrease in the rate of reaction in the latter stages of hybridisation can be traced to an additional second-order component, operating at high nucleic acid concentrations, which represents the formation of true duplexes from previously aggregated molecules (112). This is of particular importance in the use of hybridisation methodology to study the composition of extremely diverse mixtures of nucleic acids. Although the concentration of each individual reacting species may be very small in such a mixture, the total DNA or RNA concentration required to achieve complete reaction is such that aggregation effects are unavoidable. There is considerable uncertainty attached, therefore, to the meaning of the kinetic parameters relating to the final stages of such reactions.

The number of different species present in a heterogeneous mixture of DNA sequences can be deduced from the reaction kinetics of DNA re-association. Since the reaction rate for any given pair of complementary sequences is dependent on their concentration, the number of different species present in a complex mixture is inversely proportional to the reaction half-time (124). Since the range of sequence diversities encountered far exceeds the practical range of

reaction times, a more useful parameter is $C_o t$, being the product of the starting DNA concentration (C_o, expressed in moles per litre of nucleotide) and time (t, in seconds) (124). The value of this parameter at which the reaction is half complete ($C_o t_{1/2}$) is a direct measure of the diversity of sequences present in a given DNA sample, usually denoted as its base sequence complexity and expressed in nucleotides or in nucleotide-pairs, or in Dalton equivalents by reference to a kinetic standard of known complexity, such as E. coli DNA (125). This assumes, however, that all sequences are equally represented in the DNA sample, which is not the case for most eukaryotic DNAs. Here reactions are found to span far more than the 2 log units of $C_o t$ predicted for an ideal second-order reaction, indicating the presence of sequences in the population reannealing at different rates due to their different relative frequencies in the genome. By resolving such reactions into ideal components, the existence of highly and moderately repetitive genome elements has been deduced (124). The slowest reassociating component generally represents sequences present once per haploid genome, which are usually called unique or single-copy, although this class does include, in practice, poorly cross-reacting members of ancestrally related gene families, such as the vertebrate hemoglobins (126). Gene reiteration frequencies may be deduced from the sequence complexities of the various components, the proportion of the total DNA for which they account, and the overall haploid DNA-content. Repeated DNA sequences are also distinguishable from single-copy DNA by virtue of the thermal instability of the duplexes they form in hybridisation reactions, which is the result of base-mismatching (1 percent mismatch lowering T_m by 1^o (127)) between diverged members of given families of repeats (124), despite the presence of some truly re-iterated sequences.

2.2.2. Saturation hybridisation using single-copy DNA probes

In order to determine the proportion of unique genomic sequences transcribed into a given RNA population it is necessary to isolate this unique fraction of the genome. (Early studies of transcriptional activity based on competition hybridisation with DNA immobilised on filters or in agar (128,129) were inadequate for this purpose, since the effective $C_o t$ values reached are much less than in free solution (130), and almost all hybridisation is of transcripts of repetitive genes). The isolation of single-copy DNA may be accomplished by partial re-association followed by hydroxy-apatite chromatography (131-134), such that the unhybridised material is free of reiterated sequences by criteria of thermal stability of the duplexes it forms (131,133) and its hybridisation kinetics in vast total DNA excess (133,134). Labelled single-copy DNA probes have been prepared in vivo or by 'gap' translation using E. coli DNA polymerase I (140,145). Such probes, if of sufficiently high specific radioactivity, may be hybridised in vast RNA excess, such that self-annealing of the probe does not significantly remove sequences from the reaction. The proportion of single-copy genomic sequences represented in the unknown RNA may be simply computed from the saturation hybridisation level, corrected for the probe reactivity (132,134,137,138) and for sequences in duplex due to DNA-DNA reaction. This can be accomplished using an assay method capable of distinguishing directly between RNA-DNA and DNA-DNA hybrids, such as thiol-sepharose chromatography, using mercurated RNA (139,140), or alternatively, by prior treatment of aliquots of the hybrid-containing solution with RNase (138).

The reliability of saturation estimates of sequence complexity (see Table 1) depends critically on the purity of the RNA sample, since only modest levels of contamination by weight

TABLE 1. Survey of saturation and kinetic estimates of hnRNA and mRNA sequence complexity and abundance distributions.

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
MAN	Liver	A+	P	K	24.4	-	63	(197)
	Acute leukaemic cells	A+	P	K	18.6	-	57	(197)
	HeLa	A+	C	K	18	-	33	(141)
		A+	C	K	64	22	50	(142)
		A+	P	S	41			(143)
		T	P	S	45.9			(143)
		T	N	S	>410			(143)
RAT	Liver	A+	P	K	48.4	18	37	(146)
		A+	P	S	56			(144)
		A+	P	K	44	18	45	(144)
		A+	P	S	92			(145)
		A+	C	K	16.5	29	31	(372)
		T	C	S	86			(147)
		A+	N	S	260			(145)
		A+	N	S	360			(307)
		T	N	S	410			(136)
		T	N	S	260			(148)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
RAT	Regenerating Liver	A+	C	K	22	-	41	(150)
		A+	P	S	92			(145)
		A+	N	S	260			(145)
	Kidney	T	C	S	58			(147)
		T	N	S	230			(136)
	Brain	A+	C	S	170			(147)
		A-	C	S	170			(147)
		T	P	S	290			(147)
		T	C	S	360			(147)
		A+	N	S	650			(307)
		T	N	S	590			(136)
	Pancreas	A+	T	K	>0.3	90	1.5	(151)
	Adult uterus	A+	P	K	97	19	39	(152)
		A+	P	S	106			(152)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
RAT	Immature stimulated uterus	A+	P	K	14.2	28	38	(152)
		A+	P	S	22.9			(152)
	Spleen	T	N	S	180			(136)
	Thymus	T	N	S	170			(136)
	Seminal Vesicle	A+	T	K	106	50	20	(153)
	Prostate	A+	T	K	14.3	40	29	(154)
	Myoblast	A+	C	K	18	-	47	(155)
	Myotubes	A+	C	K	18.9	36	44	(156)
	Neuroblastoma	A+	N	S	490			(307)
	C6 Glioma	A+	N	S	520			(307)
	Novikoff hepatoma	A+	P	K	41.8	-	36	(146)
		A+	C	K	54.5	-	51	(150)
		T	N	S	220			(157)

TABLE 1 (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	Percentage (5)	Ref.
MOUSE	Liver	A+	P	S	22.3			(158)
		A-	P	S	16.9			(158)
		T	P	S	39.9			(158)
		A+	C	K	25.4	22	37	(159)
		A+	C	S	76			(201)
		T	N	S	181			(158)
		T	N	S	325			(201)
		T	T	S	115			(134)
	Brain	A+	P	S	140			(161)
		A+	P	S	130			(162)
		A+	P	S	120			(163)
		A-	P	S	110			(163)
		A-	P	K	90			(163)
		T	P	S	260			(163)
		A+	C	K	34.5			(171)
		A+	C	K	22.9	7	49	(159)
		A+	C	S	145			(201)
		A+	N	S	500			(161)
		T	N	S	510			(201)
		T	N	S	860			(161)
		T	N	S	616			(162)
	Kidney	A+	C	K	22.7	10	45	(159)
		T	T	S	115			(134)

TABLE 1 (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
MOUSE	Whole neonate	T	T	S	460			(132)
	Embryonal carcinoma	A+	P	K	15.5	-	36	(164)
		A+	P	S	21.6			(164)
		A+	N	K	102	-	89	(164)
		A+	N	S	108			(164)
		T	N	S	220			(360)
	Embryoid bodies	A+	T (C)	K	3.4	10	25	(149)
		A+	N	K	34	-	48	(149)
	Friend erythro-leukaemia cells	A+	P	S	48.8			(165)
		A+	P	K	19	-	10	(166)
		A+	N	K	103	-	70	(167)
		A+	N	S	206			(165)
		T	N	S	224			(165)
	Ehrlich carcinoma	A+	P	K	40	16	38	(168)
	AKR-2B cells	A+	P	K	22.5			(169)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
MOUSE	3T6 (growing)	A+	C	K	15	-	22	(141)
	3T6 (resting)	A+	C	K	17	-	28	(141)
	Mammary gland	A+	T	K	12.5	32	35	(206)
		A+	T	S	315			(206)
	Mammary carcinoma	A+	T	K	20.1	-	30	(206)
		A+	T	S	315			(206)
	EC-derived myoblasts	A+	P	K	26	-	36	(170)
	L cells	A+	P	S	50			(155)
		A+	C	K	14.4	5	50	(171)
	Neuroblastoma cells	A+	P	K	12.3			(384)
		A+	C	S	46			(172)
		T	N	S	120			(172)
	PYAL/N cells (Polyoma-transformed)	A+	P	S	42.4			(158)
		A-	P	S	33.9			(158)
		T	P	S	67.9			(158)
		T	N	S	636			(158)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
MOUSE	Taper	A+	P	K	15.3	20	42	(214)
	hepatoma	A+	C	K	14.8	28	36	(214)
HAMSTER	13-day	A+	P	S	58			(173)
	embryo	T	N	S	410			(173)
	SHE(embryonic)	A+	P	S	27			(173)
	cells	T	N	S	116			(173)
	BP6T cells	A+	P	S	27			(173)
		T	N	S	116			(173)
	CHO cells	A+	P	K	15.1			(433)
CHICK	Myoblasts	A+	C	K	32.4			(309)
	Liver	A+	P	K	22.2	65	35	(174)
		A+	P	S	34			(174)
	Oviduct	A+	P	K	15	25	43	(381)
	(hormone	A+	P	K	26.6	16	44	(174)
	stimulated)	A+	P	S	30			(174)
		T	N	S				(175)

TABLE 1 (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
CHICK	Immature erythroid cells	A+	P	K	0.11	90	-	(220)
		A+	N	K	5.1	-	50	(220)
XENOPUS	Kidney	A+	C	K	30.9	-	40	(223)
	Liver	A+	C	K	24			(176)
		A+	N	K				(176)
	Oocyte	A+	C	K	30.1	-	53	(223)
		A+	C	S	35.2			(223)
		T	T	S	20			(137)
	Tadpole	A+	C	K	35.5	-	25	(223)
RANA	Embryos (neurula stage)	A+	P	S	174			(177)
		T	N	S	420			(177)
	Embryos (larval stage)	A+	P	S	322			(177)
		T	N	S	450			(177)

TABLE 1. (Contd.

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
TROUT	Testis	A+	C	K	11	15	40	(39)
SEA URCHIN	Adult	T	P	S	6			(179)
	Intestine	T	N	S	230			(178)
	Adult coelomocytes	T	P	S	4.3			(179)
	Adult tubefoot	T	P	S	2.7			(179)
	Adult ovary	T	P	S	20			(179)
	Oocyte (mature)	T	C	S	38			(179)
	Oocyte (previtellogenic)	T	C	S	16			(180)
		T	N	S	160			(180)
	Blastula	T	P	S	26	-	10	
		T	P	S	27			(179)
	Gastrula	T	P	S	17			(179)
		T	N	S	174			(181)
	Exogastrula	T	P	S	14			(179)
	Pluteus	T	P	S	14			(179)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
DROSOPHILA	Egg	A+	C	K	29			(163)
		T	C	S	12			(182)
	Tissue-culture cells	A+	C	K	8.48			(183)
		T	C	S	6.8			(183)
		A+	N	K	42			(160)
		T	N	S	52			(160)
MUSCA	Egg	T	C	S	24			(182)
TOBACCO	Leaves	A+	P	S	30			(184)
		A+	P	K	15	10	39	(185)
		T	P	S	33.3			(185)
		T	N	S	119			(185)
	Stem	A+	P	S	32.4			(186)
	Root	A+	P	S	30			(186)
	Petal	A+	P	S	33			(186)
	Anther	A+	P	S	32.3			(186)
	Ovary	A+	P	S	31.1			(186)

TABLE 1. (Contd.)

Organism	Cell-type	RNA (1)	studied (2)	Method (3)	Complexity (nucleotides x 10^6)	Percentage (4)	(5)	Ref.
BARLEY	Leaves	A+	P	S	45			(187)
PARSLEY	Leaves	A+	P	K	12.9			(188)
		A+	P	S	35			(188)
	Callus	A+	P	K	16.1			(188)
		A+	P	S	49			(188)
SOYBEAN	Suspension culture cells	A+	T	K	45			(189)
		A+	T	S	33			(189)
		T	T	S	64			(189)
	Rhizobium-infected root	A+	P	K	27	24	51	(190)
EUGLENA	Dark-grown cells	T	T	S	80			(191)
	Light-grown cells	T	T	S	65			(191)
Schizophyllum commune (FUNGUS)		A+		S	11			(192)
Achyla ambisexualis		A+	P	K	3.7	-	36	(193/194)
		A+	P	S	3.7			(193/194)
YEAST		A+	T/P	K	4.35	23	26	(195)
		A+	T	S	5.45			(195)

Notes:

(1) Total, denoted T, poly(A)$^+$, denoted A+, or poly(A)$^-$, denoted A-.

(2) Total (T), Polysomal (P), Cytoplasmic (C) or Nuclear (N).

(3) Saturation (S) or Kinetic (K).

Where saturation estimates are not quoted in nucleotides in the original reference, calculation is based on consensus (or most recently published) estimates of unique sequence complexity of the organism concerned.

(4) Percentage of sequences at mass fraction > 1 percent in the population (abundant sequences).

(5) Percentage of sequences at mass fraction < 0.02 percent in the population (rare sequences).

(for example of polysomal by the much more complex nuclear RNA) can lead to serious overestimates. The effect of cellular heterogeneity in the biological source material must also be considered. Since the vast bulk of both nuclear and cytoplasmic sequence complexity in RNA is at fairly uniform relative abundance, such contamination effects are easily detected by the failure to conform to the expected single-component first order kinetics (in vast RNA-excess), or the failure to achieve a satisfactory plateau level of hybridisation. In many instances, the deviation from ideal first-order kinetics is slight, and a reasonably accurate estimate may be derived of the proportion of the RNA mass which represents the complex class of sequences driving the reaction. For example this is approximately 8 percent of the total in the case of sea urchin blastula polysomal mRNA (178). In other cases, however, it is evident that the complex class of sequences is so rare in the driver that the reaction never reaches completion, or worse, an insufficient RNA excess is achieved. The development of techniques for purifying complex populations away from the bulk of simpler sequences (especially rRNA and its precursors) has greatly decreased the technical problems associated with this method. These include puromycin or EDTA-induced release of mRNPs from polysomes (137), and oligo(dT)-cellulose chromatographic purification (196) of polyadenylated nuclear and polysomal RNAs (161,165). These purified fractions contain a much higher content of the heterogeneous sequences driving the single-copy DNA reactions, so that the $R_o t$ values (where $R_o t$ is the counterpart of $C_o t$ for an RNA-driven reaction) at which saturation is achieved are much more easily obtainable in solution. Saturation values obtained without using such purification procedures (for example in Refs. 132, 134 or 165) must therefore be treated as minimum estimates, to which is attached considerable uncertainty. Much of the data in the literature, relating to comparisons of single-copy sequences expressed in different RNA populations, have failed to take account of these uncertainties. Although

technically more reliable, complexity estimates using poly(A)$^+$ or other enriched mRNA preparations do suffer the uncertainties resulting from the possible exclusion of a significant set of sequences by the purification method itself. This is particularly important, since it is clear that poly(A)$^-$ mRNAs may constitute a distinct and significant set of sequences (see Section 2.3.3). Once again, these difficulties have often been overlooked by investigators.

The reproducibility of complexity estimates by this method, especially where care has been taken to account for the possible contribution of very rare sequences, suggests that it has general validity. Particularly striking are the relative complexities of nuclear and cytoplasmic or nuclear and polysomal RNAs which show a similar relationship across the phylogenetic spectrum (see Table 1). The biological significance of this and other findings will be discussed subsequently. A number of points may be made here in relation to the methodology. Firstly, as indicated by Kiper (188), the interspersion of unique and repetitive elements in eukaryotic genomes (198,199) means that randomly sheared single-copy DNA prepared by partial re-association, will be enriched in those unique sequences which occur in long tracts uninterrupted by repetitive elements. This may enrich for or deplete mRNA sequences in the probe, depending on whether coding sequences are or are not preferentially associated with repeats. If the single-copy probe is depleted of coding or non-coding sequences, the saturation values of single-copy DNA hybridisation will not represent the true sequence complexities of the RNA samples under investigation. This may be particularly serious for estimates of RNA complexities in plants, where repetitive DNA makes up a major portion of the genome (188), or in those species where coding sequences are generally adjacent to repetitive tracts as in sea urchins (200) and in the mouse (201). A second problem cited by Kiper (188) arises from the contiguity of segments of mRNA

sequences with non-coding regions (introns) in the genome, which means that estimates of mRNA complexity based on saturation hybridisation to unique genomic fragments will tend to be too high. These sources of error may account for the generally higher complexity estimates derived from single-copy DNA saturation methodology than from cDNA hybridisation kinetics (further discussed below in Section 2.2.3). Investigators other than Kiper have maintained, however, that saturation and kinetic estimates of complexity concur, even in the case of plant RNAs, if the measurements are carried out in parallel, under favourable conditions (202, 144).

Other more technical problems arise in saturation estimates of complexity. The very long incubation times needed to achieve saturation, often at elevated temperatures, could result in RNA or DNA degradation, generating spurious plateau hybridisation levels. The integrity of probe and driver have not generally been checked after reaction, except in the earliest studies (131). Residual contamination of the probe with repetitive sequences will result in falsely high complexity values if these sequences are represented in RNA. This effect can be minimised by checking the 'uniqueness' of the single-copy probe and of the fraction of it hybridised by RNA, both by thermal stability and re-association kinetic criteria (132,134).

2.2.3. Hybridisation kinetics using complementary DNA probes

The results of saturation hybridisation contribute almost no information on the relative abundances of different sequences in an RNA population. For this, a method analogous to the kinetic analysis of DNA re-association is required. This is made possible by the reverse transcription of RNA into complementary DNA (cDNA) using retroviral reverse transcriptase. Although *in vivo* this enzyme copies only viral RNA, using a

cellular RNA as a primer, in vitro it will transcribe any RNA given a suitable primer (203-205, 142, 207). Transcription of poly(A)$^+$ RNAs is achieved by adding a large excess of oligo(dT) fragments, such that DNA synthesis is primed by the fragment immediately adjacent to the mRNA sequence, avoiding the synthesis of large poly(dT) tracts which would cross-hybridise with any poly(A)$^+$ RNA. The rate of RNA-cDNA hybridisation is a measure of the complexity of the population, since in a second-order reaction, the rate depends on the concentration of each individual reacting species (208-212). The mathematics may be simplified if the reaction is performed in vast RNA excess (208), such that pseudo first-order kinetics according to the RNA concentration is observed (208,213,159). As in the case of DNA-DNA re-association, complexities are calculable with reference to a suitable kinetic standard, such as the hybridisation rate of mouse globin mRNA with its cDNA (211), using the kinetic parameter $R_o t_{1/2}$ (208,213), formally the counterpart of $C_o t_{1/2}$ in DNA renaturation. Under these conditions, reaction proceeds to completion within a narrower range of $R_o t$ values than is the case for a strictly second-order reaction with the starting concentrations of the reactants equal (208). The numerical relationship of complexity and $R_o t_{1/2}$ is therefore not the same as between complexity and $C_o t_{1/2}$ in DNA renaturation.

The overall kinetics of cDNA hybridisation in RNA excess can be used to characterise the composition of a population of sequences at heterogeneous relative abundances (166,142,155), by resolving the kinetic curve into pseudo first-order components by a least-squares curve-fitting method. The total complexity of the population is then the sum of that of the hypothetical components. If values are measured or estimated for the average mRNA length, and the total RNA content of the cell, the number of different mRNA species and the number of copies of each per cell may be computed. The degree to

which this division into 'abundance classes' reflects the true mRNA (or hnRNA) composition of the cell is disputed. HeLa cells (206) and yeast (195) do seem to contain fairly discrete classes of mRNA at homogeneous abundance, but for most other cell-types clear-cut kinetic transitions cannot easily be identified, and the data may be equally well resolved into any number of hypothetical components larger than two or three, implying a continuum of relative abundances in the cell. This is the case, for example, in mouse L cells (171).

The use of this method to study nuclear RNA populations carries additional uncertainty in that the probe is no larger than that obtained from polysomal mRNA, whereas the template RNA is very considerably longer on average (if undegraded), hence the cDNA must be regarded as a probe for the 3' proximal region only. This would only seriously affect complexity estimates, however, if different nuclear RNA species carried common 3' ends (or vice versa). The RNA driver length should not affect the reaction rate directly (166), provided it is longer than the probe. The picture is also complicated by the possible self-priming activity of nuclear RNA, reported by some investigators (215).

The interpretation of cDNA hybridisation kinetics in terms of relative abundances and sequence complexities depends upon the number and accuracy of data points, and also upon the validity of certain assumptions regarding the method. Firstly, it is assumed that the cDNA is a faithful copy of the RNA population (i.e. that all sequences are equally efficiently reverse transcribed). Consideration of the relative mRNA and cDNA sizes, the overall transcriptional yield, and the mechanism of reverse transcription (216-218) led Birnie *et al* (166) to conclude that almost all mRNA sequences were reverse transcribed, hence that cDNAs were unlikely to deviate grossly from the composition of their templates. By constructing

mixtures of globin or mouse liver mRNA with different stoichiometries, and analysing the hybridisation kinetics of the resultant cDNAs, Hastie and Bishop (159) showed that the frequency of a given mRNA in the population was not a determinant of the efficiency of its transcription. However, the existence of specific 'pause' sites in the reverse transcription of particular RNAs (210,219) leaves open the possibility that cDNA populations are selectively depleted of some sequences. Transcriptional efficiencies with nuclear RNA are considerably less than can be accounted for by the increased template size, although this is apparently due to the action of a non-specific inhibitor (215) rather than sequence selective transcription.

A second assumption is that the unreactable cDNA sequences are drawn equally from all frequency classes (141). If their unreactivity is due to degradation during the long incubation times required to reach completion, their composition may in fact be biassed towards the rarest, most complex sequences which participate only in the final stages of the reaction. Complexities would thus be seriously underestimated. Aggregation effects may also be responsible for biassing the composition of the unreacted sequences towards the rarest species, since abundant sequences which become non-specifically aggregated with RNA or DNA are more likely to proceed into stable duplex during transient release from the aggregated state. The complexity of the rarest component is often difficult to estimate if no clear plateau of hybridisation is reached for these reasons, and this is especially a problem in measurements of hnRNA complexity (221). In the extreme, ultra-rare components of high complexity may go undetected, especially if, as implied by single-copy DNA saturation experiments, they represent only a very tiny fraction by weight of the total (136).

A third problem arises from the limitation imposed by the use of poly(A)$^+$ sequences only as templates, which is the case for the vast bulk of published experiments using this method. It is now possible, however, to prepare a cDNA to poly(A)$^-$ mRNA or hnRNA populations by the use of a random primer (207), although ribosomal sequences must be removed by saturation with rRNA.

Investigations of the specificity of mRNA or hnRNA populations between cell-types (or between cellular compartments), are made possible by cross-hybridising cDNA probes with heterologous RNAs (141,171,220,221,222,176). The difference in the final plateau level of hybridisation is a measure of the extent of qualitative differences (by weight) between the two populations, although care must be exercised to ensure a sufficient ratio of driver heterologous RNA to cDNA, in order that large quantitative differences are not erroneously scored as qualitative ones. This can usually be accomplished by titration of the cDNA with increasing ratios of the heterologous RNA, taking the reaction to a $R_o t$ value at or beyond the commencement of the apparent plateau. This control has not universally been applied, hence apparent instances of qualitative differences between populations (220) must be treated with a certain amount of caution. A shift in the reaction kinetics between homologous and heterologous hybridisations, carried out at high RNA:cDNA ratios, indicates differences in the relative abundances of sequences held in common between the populations (222). Because heterologous reactions cannot be analysed mathematically in the same way as homologous reactions, the proportion of the RNA driving the hybridisation of each component of the cDNA being uncertain, their interpretation in terms of numbers of sequences involved in abundance shifts, and the actual extent of these shifts, cannot be precise.

A more detailed analysis of abundance differences can be achieved by the use of cDNA probes fractionated, by partial hybridisation, into abundant and rare components (145,159, 176,224). Probe fractionation has also been used to amplify supposed qualitative differences between RNA populations (141), although once again, the interpretation of the results depends critically on the RNA:cDNA ratio used in the preparative hybridisation, which must be sufficient to exclude sequences which are only quantitatively regulated.

2.2.4. Molecular cloning: the use of gene-specific probes

The study of eukaryotic gene expression has been revolutionised by the development of the technology for the isolation and bulk preparation of DNA segments corresponding to individual genes, both at the level of the genome, and of mRNA (225,226). Originally designed for the study of genes coding for specific highly abundant proteins, such as globin, ovalbumin or silk fibroin (227-229), these methods have now been adapted to permit the construction of 'libraries' - plasmid or phage banks incorporating an entire repertoire of messenger sequences or genomic information from a given cell-type (230-232). By the appropriate screening technique (233,234) clones containing a particular sequence may be selected from these libraries. Alternatively, a random set of clones may be selected in order to test the generality of any hypothesis concerning gene control mechanisms (231,235), or gain general information about the processes of gene expression. This latter application is particularly appropriate for verifying, extending and refining the results of heterogeneous cDNA hybridisation studies as described above (Section 2.2.3): gene-specific probes may be used to determine directly the relative abundance of a given sequence in many different RNA populations, and permit even quite subtle quantitative modulations to be analysed in great detail.

A number of methods are available for determining the level of a given mRNA sequence, using a cloned hybridisation probe. These include measurement of the kinetics of RNA-excess hybridisation (236), where the $R_o t_{1/2}$ for the hybridisation of a sequence of known complexity is a measure of the proportion of the RNA driving the reaction; and titration in cloned DNA excess, with increasing RNA:DNA ratios, to fix the amount of RNA needed to saturate the probe (236). Both these methods require the preparation of strand-separated plasmid probes, which is particularly tedious and expensive on material. An alternative method is titration of labelled RNA, or its cDNA, with an excess of filter-bound plasmid recombinant DNA. If taken to completion, (231,237) the proportion of ribonuclease-resistant or S1 nuclease-resistant counts bound to the filters indicates the proportion of the RNA or cDNA which is complementary to the given recombinant sequence (provided appropriate corrections are made for plasmid insert-size, mRNA size, and probe reactivity). Since the reaction is carried out under considerable filter-bound DNA-excess, it should exhibit pseudo first-order kinetics such that at any given time, the proportion of counts bound will be specified by the equation

$$\frac{c}{c_o} = 1 - e^{-ft} \qquad \text{(Eqn.1)}$$

c = counts bound
c_o = total counts in solution at time $t = 0$
t = time

where f, being a function of the mass per unit area of reacting DNA on the filters, is independent of the starting concentration of that sequence in the RNA or cDNA (238). For different sequences hybridised under identical conditions, this proportion of bound counts at any given time should be the same. Hence the use of an appropriate standard should permit the use of this method to determine the mass fraction of many different mRNAs in many different RNA populations, using only a tiny quantity of plasmid DNA (since highly

labelled cDNA probes are sufficiently sensitive to detect mRNA sequences at the observed limits of cellular abundance).

As well as enabling a direct measurement of relative mRNA abundances, gene-specific probes have also contributed indirect evidence on gene control mechanisms, by virtue of facilitating the analysis of gene structure. This has involved genomic blotting (239) in conjunction with restriction mapping (240), in order to determine the arrangement of coding and non-coding sequences, and RNA blotting (241), heteroduplex analysis (242) and S1 mapping (243) to characterise individual transcriptional units, and processing intermediates on the pathway to mature mRNAs. Because their expression pathway may be traced in such detail, gene-specific probes are particularly useful for testing the *in vivo* equivalence of *in vitro* systems for the study of gene expression (see Section 2.4).

2.3. Sequences in cellular RNA populations

2.3.1. Transcripts of unique and repetitive DNA sequences

Kinetic studies of DNA re-association (244-249) have shown great diversity amongst eukaryotes in the relative proportions of unique and repetitive sequences, in sequence re-iteration frequencies, in total unique sequence complexities and in the organisation of unique and repetitive elements. This diversity shows no obvious phylogenetic rationale. Most mammals (244), amphibia (244) and higher plants (245), many invertebrates (244,246) and some eukaryotic protists (250, 251) show interspersion of short regions (typically 300 base-pairs) of repetitive sequence of intermediate re-iteration frequency ('middle-repetitive DNA') with unique regions about 1-2 kilobase-pairs long. Some insects, such as Drosophila (247), other invertebrates (248) and protists (252), however, display a much longer interspersion periodicity, with unique segments averaging more than 13 kilobase-pairs in length.

Avian genomes that have been examined seem to have an intermediate interspersion pattern (249), and some of the lower eukaryotes do not show interspersion at all (253,254). Total unique sequence complexity bears no simple relationship with biological complexity, even for taxonomically close species (255,256).

From the pattern of transcripts of unique and repetitive sequences found in cellular RNA populations, tentative conclusions emerge relating to the role of these different sequences in gene expression, and to the way in which genome organisation may influence its control at one or several levels. The first point to be established is the identify of the different classes of sequence, in respect of structural genes. The vast majority of structural genes are encoded in one or a very few copies per genome. This applies both to specific individual genes whose frequencies have been determined, such as chick ovalbumin (257), rat serum albumin (3) or silkmoth fibroin (258), and to the bulk of mRNA sequences, as shown by the 'unique' re-association kinetics of messenger cDNAs with excess total DNAs, from a wide variety of cell-types and organisms (135,159,183,185,189,259,260). Even highly abundant mRNAs are transcribed, in general, from single-copy genes (135), the most extreme example being that of the vertebrate major hemoglobins, which in reticulocytes account for more than 90 percent of the total mRNA (261). A number of structural genes are encoded in related multigene families, such as the silkmoth chorion genes (262), the Drosophila actins (263) or the Xenopus vitellogenins (264), but with the possible exception of the histones, where gene dosage may be related to the genome size and the duration of S-phase (265-267), there does not appear to be any general correlation between gene frequency and either mRNA abundance or the demands of protein synthesis. Unique sequence complexities in general far exceed those needed to encode the number of structural genes postulated on the basis of

genetic considerations (268-270) or the complexities of mRNA and hnRNA populations (136,179,186 and Table 1). Various 'functions' for the extra DNA have been proposed (255, 271-273), including the suggestion that its persistence is due to the acquisition of 'selfish' properties merely favouring its own survival (272,273).

Middle-repetitive DNA includes, in effect, all the re-iterated sequences except those simple sequences banding as distinct satellites in density-gradients (274). Middle-repetitive DNA sequence families range from as few as tens of copies per genome (275) to those constituting as much as 3 percent of the total DNA (276), and include a number of structural gene families, such as the histone genes (265-267). Highly re-iterated sequences may have important general functions in the replication or expression of the genome (277-279). Sequences related, for example, to the human 'Alu I' family are expressed in vivo in double-stranded hnRNA, where they may constitute generalised processing signals. They are also found in short, in vitro RNA polymerase transcripts, and are uniquitously represented in mammalian repetitive DNAs, in viral DNA replication origins, and in the short RNAs found hydrogen-bonded to polyadenylated messengers (278). Shorter, highly conserved sequences have also been observed in or adjacent to many structural genes, and constitute putative signals for transcription, capping, splicing, polyadenylation, translational initiation and somatic recombination (280-283), in species ranging from yeast (283) to man (284).

The possible function of middle-repetitive DNA as a regulatory element in the control of structural gene expression has been discussed extensively by Davidson and Britten (199,285,286). They propose that sets of structural genes may be co-ordinately expressed, by virtue of their association in the genome with specific families of repeat elements, which act as recognition sites for complementary transcripts distributed in a tissue-

specific manner. The heterogeneity of genomic interspersion patterns suggests that if such mechanisms operate they must be highly adaptable. Repeated sequences which might function as control elements have been identified within co-ordinately expressed gene families in Dictyostelium (287) and adjacent to isocoding sea urchin histone gene variants (283). The human and rabbit β-like globin gene loci show a complex pattern of interspersed repetitive elements, including palindromic sequences (56, 288,289), which may control their expression. In the mouse, the 3' ends of both α and β-like globin gene sets are bordered by a common repeat, located 1.5 kilobase-pairs from the genes (290). The location of a large proportion of structural genes in both the sea urchin (200) and mouse (201) genomes, adjacent to repetitive elements, as has been deduced from the hybridisation of messenger cDNAs with 'repeat-continguous' fragments of single-copy DNA, favours the idea that such elements may be important in their expression, although by what mechanism, or at what level, remains unknown. The partially clustered arrangement of repeat elements within one part of the genome, with their 'scrambled' representation elsewhere, interspersed with unique sequences, as predicted by the Britten-Davidson model (285), is evidenced by the re-association properties of sheared repetitive DNA of sea urchins (291) and amphibia (292), and by the chromosomal organisation of individual cloned repeats of Drosophila (293) and chicken (294) DNA.

The patterns of repeat sequence expression in nuclear RNA provide considerable support for such a model. Both in sea urchins (181) and mammalian cells (295), nuclear RNA contains a significantly greater proportion of repetitive sequence transcripts than does polysomal RNA. Short 'chromosomal RNA' transcripts in rat Novikoff hepatoma cells consist more than 80 percent by weight of repetitive sequences (133), and repetitive sequences are also found in high molecular weight

hnRNA, where about 8 percent of the different repeat families are represented (157). Repeat sequence expression in chromosomal RNA follows a definite tissue-specific pattern (296). The use of cloned repeat elements has shown a tissue-specific distribution of repeats in sea urchin developmental stages (236), both strands of a given repeat family being represented in hnRNA, but remaining nucleus-confined. Repeat sequences are also represented symmetrically in sea urchin oocytes (235) where, despite differences in relative genomic frequency, the pattern of expression is highly conserved between species (297). Interspersed unique and repetitive elements are covalently linked in echinoderm (298) and mammalian (279) hnRNA, and in sea urchin egg poly(A)$^+$ RNA (299), where at least half the mass of maternal mRNA, and most of the sequence complexity is accounted for by such transcripts (299). Such structures are expected if post-transcriptional processing is regulated by the presence of particular families of repeat sequences in primary transcripts as predicted by Davidson and Britten (285). If RNA duplex formation in nuclear RNA is associated with processing selectivity (either of a qualitative nature, as originally proposed by Britten and Davidson, or equally likely of a quantitative nature) as specified by repeat sequence transcripts, then repetitive sequences should comprise most if not all dsRNA tracts (and vice versa). HeLa cell hnRNA contains transcripts of middle repetitive DNA which are able to form intermolecular RNA duplexes, as well as intramolecular snapback structures (295), and internal palindromes are found in CHO cell nuclear RNA (300). Snapback regions in mouse Ehrlich carcinoma cell hnRNA (277) consist of a few highly re-iterated sequences, whereas the longer sequences found in intermolecular duplexes are relatively complex, and have some homology with sequences found in mature mRNA (301). A role in RNA processing has been proposed for the small nuclear RNAs (302) which are found associated with hnRNPs as small RNPs (303,304), on the basis of the observed complementarity between their

sequences and those of splice-junctional sequences in hnRNA. Conceivably, the relative availability of different variants of these splice-directing RNAs might regulate the differential processing rates of transcripts with different versions of the consensus splice-junctional sequence. In Xenopus, for example, the small nuclear RNA moieties which are at highest abundance alter during development (305). These ideas remain necessarily speculative. A 'splice-director' function is also proposed for the action of viral-coded VA-RNA, in the synthesis of adenoviral mRNAs (306). This RNA will hybridise with spliced late mRNA, but not with genomic DNA, and is found complexed with high molecular weight nuclear RNA in infected cells.

These observations constitute strong circumstantial evidence in support of the post-transcriptional regulation of the relative levels of expression of structural gene sets, by tissue-specific transcription of repetitive DNA sequences.

2.3.2. Complexities and abundance components of nuclear and cytoplasmic RNA

Although hybridisation studies of genome organisation can yield suggestive evidence of gene control mechanisms, the use of the methodology to investigate directly the qualitative and quantitative composition of the population of sequences transcribed into RNA and processed and transported onto the polysomes, offers a more detailed analysis of the selective processes. Table 1 summarises the results of saturation and kinetic estimates of nuclear and cytoplasmic RNA complexities, and where appropriate, abundance distributions, for a great variety of organisms and cell-types. These data are not exhaustive, although they are representative of findings in this field, and illustrate a number of important points. Firstly, as will be obvious, the discrepancies are far more than can be accounted for by the inherent uncertainties of the different techniques (see Sections 2.2.2 and 2.2.3),

with different investigators often obtaining wildly divergent estimates of essentially the same parameter. This does not necessarily imply that some investigators are incompetent, or have failed to observe or record important methodological precautions. What it does mean, however, is that the numerical results obtained from such studies are often very sensitive to minor differences in technique, in particular of tissue preparation, which are not always easily controlled. Therefore, it is unwise to attach too much significance to individual reports of complexity differences (or similarities), especially when these are against a general trend, except where very rigorous controls have been carried out. The data presented in Table 1, together with other findings, will be considered subsequently from the viewpoint of differential gene expression in related but distinct cell-types, and its implications for gene control mechanisms (Sections 2.3.5 and 2.3.6). Here, the data will be considered 'vertically' - in other words from the viewpoint of their implications for RNA metabolism within individual cell-types, as the pathway of information flow, from genome to polypeptide synthesis, is traversed.

A second feature which emerges from the data of Table 1 is the fact that in all cases only a fraction of the total unique sequence complexity of eukaryotic genomes is ever transcribed in a given cell-type into nuclear RNA, although the absolute sequence complexity of hnRNA, both in nucleotides and expressed as a percentage of the genetic information available in a given genome, is extremely variable even between the tissues of a single organism such as the rat (136,307). Mammalian brain nuclear RNAs, for example, are more than one hundred times as complex as those of a lower eukaryote, such as yeast (195). Similarly wide spans are evident at the level of polysomal RNA, where sea urchin coelomocyte polysomes, for example, contain less than 1 percent of the sequence complexity of mouse brain (161), and only 10 percent that of the polysomes in a cloned mouse neural cell-line (172),

but may be of comparable RNA complexity to mouse embryoid body polysomes (149).

Thirdly, nuclear RNA complexities, however measured, and whether or not consideration is restricted to the poly(A)$^+$ sequences, are at least several-fold (177) and sometimes as much as thirty-fold (178,179) greater than the complexities of cytoplasmic or polysomal RNA, with the exception of the lower eukaryotes (193-195). The ubiquitousness of this finding, despite all the numerical uncertainties, is strongly suggestive of its biological importance. It can be interpreted either as the qualitative selection of a limited set of transcripts for further processing, and ultimate expression, or as the elimination during processing of a large portion of the unique sequences associated with each species of primary transcript. Whilst the processing of large primary transcripts to less complex mRNAs is undoubtedly general, the question of whether post-transcriptional qualitative selection is superimposed upon it, contributing to the complexity differential of nuclear and polysomal RNA, remains an unresolved question which will be further discussed below (Section 2.3.3).

Fourthly, the pattern of relative abundances, both in nuclear and polysomal RNAs, shows equally great diversity. In rat pancreas (151) as in mouse reticulocytes (261) more than 90 percent of the mass of mRNA is accounted for by a few superabundant species and moreover, in mouse brain, just 1 percent of the poly(A)$^+$ cytoplasmic RNA represents more than 75 percent of the sequence complexity (201). Conversely, in poly(A)$^+$ nuclear RNA from embryonal mouse carcinoma cells, the complex sequences are in a homogeneous class, comprising about 90 percent of the mass (164). Again, it is difficult to extract any meaningful general pattern from such diverse data, although it is evident that highly abundant mRNAs are a feature of individual specialised cell-types, particularly those in terminal differentiation, whereas ultra-rare sequences

are typical of highly heterogeneous tissues such as brain (or whole embryos). Rare mRNAs, comprising a variable proportion of the total mass, are nevertheless found in all cells, and are thought to encode many basic and ubiquitous metabolic ('housekeeping') functions.

Relative mRNA abundances, as predicted by cDNA hybridisation kinetics, are well correlated with translational template activities in a heterologous cell-free system, as shown by the kinetics of hybridisation-mediated translational arrest (308,309). The diversity of mRNA abundances has been confirmed from studies using polysomal cDNA recombinant libraries, in CHO cells (231), silkmoth larvae (232), developing Xenopus tadpoles (310), Dictyostelium (311), sea urchins (312) and Drosophila (313).

A final point which may be extracted from the data of Table 1, despite uncertainties about the accuracy of numerical estimates, is that cell-types of phylogenetically disparate organisms, where they are biologically comparable, show not dissimilar mRNA (or hnRNA) complexities. For example, undifferentiated human (141), mouse (171), plant (189) and even fungal (192) cells in tissue-culture, all express about 10,000 different poly(A)$^+$ messengers, as determined by kinetic measurements of complexity. More specialised somatic tissues of mammals (excluding neural tissue) (144,159), plants (186) and avian species (174), all express 15-25,000 different mRNAs. Identical oocyte mRNA complexities have been observed in amphibia of vastly different genome sizes (259). These correlations are once again strongly suggestive of the link between mRNA expression and phenotype, even for rare sequences, which as Galau et al (314) have pointed out, are capable of supporting a sufficient rate of polypeptide synthesis to maintain observed levels of representative cellular enzymes.

In many studies polysomal and total cytoplasmic sequence sets have not been distinguished. This may be significant in comparing RNA populations in distinct cellular compartments, since cytoplasmic RNA will obviously contain a variable, but in some cases important proportion of contaminating RNPs of nuclear origin, as well as genuinely untranslated mRNPs, whose identity as a distinct sequence set may be relevant to control at the level of translation. This latter point will be discussed subsequently (Section 2.3.6). However, its possible influence on the apparent differences between nuclear and cytoplasmic RNA populations should not be discounted, nor should the possible effects of nuclear contamination of cytoplasmic RNA.

2.3.3. Qualitative aspects of post-transcriptional selection

The much greater complexity of nuclear than cytoplasmic RNA can, as has already been indicated, be interpreted in two ways: either as qualitative post-transcriptional selection or as generalised processing of large primary transcripts. In order to distinguish these it is necessary to digress briefly into what is known of gene structure, and the definition of the transcriptional unit.

As shown by restriction mapping, heteroduplexing, and sequencing studies of cloned DNAs, structural genes are in general encoded discontinuously in DNA; that is, regions which appear in mature mRNA (exons in Gilbert's terminology (315)) are interrupted by non-coding regions (introns). The total length of intron DNA is generally at least several times that of the mRNA; three times as large, for example in the case of chick ovalbumin (287), although this ratio is extremely variable: the entire mouse dihydrofolate reductase gene being about 25 times the size of its mRNA (316), whereas the silk fibroin intron is trivial in size compared with the gene in which it is located (317). A number of structural genes, however,

contain no introns, such as yeast cytochrome c (318), Drosophila heat shock genes (319), the sea urchin histones (265) and adenovirus polypeptide IX (412) . The intron-exon type of organisation spans the evolutionary hierarchy, applies both to germ-line DNA and the DNA in a cell-type in which a given gene is expressed (257,321), and to genes coding for mRNAs in all frequency classes, including super-abundant, differentiation-related mRNAs such as hemoglobin (240), ubiquitous abundant mRNAs such as actin (282), and rare mRNAs encoding housekeeping functions, such as dihydro-folate reductase (316). In addition, eukaryotic viral genes show a similar organisation, despite the pressure for genetic economy (322-325). The origins of split gene organisation, and the possible functions it may serve are unclear. Gilbert (315) has proposed that exons correspond to structural and/or functional domains of globular proteins, for which there is considerable evidence from structural studies of the genes and proteins of hemoglobin (320), chick lysozyme (326), chick ovomucoid (327) and the immunoglobulins (328). This arrangement may facilitate evolution by allowing new proteins to be 'sewn together' in novel ways, from pieces of old proteins. An alternative proposal is that divergent intron sequences diminish the chance of undesirable recombinations taking place between closely related gene sequences situated adjacent to one another in the genome (329). Regardless of whether it is their main 'function' introns may play a major role in the control of gene expression, either as recognition elements for regulatory factors, or simply by providing an additional level of mRNA processing (splicing) at which sequence-selectivity may operate.

The mosaic organisation of eukaryotic genes offers at least a partial explanation for the much greater complexity of nuclear than cytoplasmic RNA, since intron sequences are transcribed in all cases studied, alongside the exons of the genes which they interrupt (330-335), and are subsequently

removed by RNA splicing (74) which takes place wholly within the nucleus (336). Intron sequences removed by splicing are degraded in the nucleus, being undetectable in cytoplasm above the level of nuclear contamination (337). Transcriptional units do not appear to extend significantly beyond the 5' and 3' ends of the terminal exons (338), except in the case of the papovaviruses (339), where multi-genome transcripts appear to be processed into single mRNAs. This means that the ratio of mRNA complexity to 'nucleus-confined' complexity should be the same as the ratio of total intron to exon lengths, averaged over all genes, in the absence of further qualitative post-transcriptional selection. Unfortunately, the number of individual genes whose structure is known in detail is still insufficient to permit any definite conclusion on this matter, especially in view of the uncertainties attached to complexity estimates of nuclear and polysomal RNA. Other attempts to fix the ratios of primary transcript size to mRNA length suggest that in general, genes are several fold larger than their corresponding mRNAs, though leaving considerable doubt that this can explain away all the extra complexity of nuclear RNA in a given cell-type. These studies include determination of the u.v. target size for inactivation of mRNA synthesis (81), and hybridisation studies of the species present in nuclear RNA corresponding either to a heterogeneous group of mRNAs, such as the short mRNAs of rat liver, 30 percent of whose nuclear counterparts are more than twice as long (340), or to a randomly selected set of cloned mRNAs from a cDNA library in CHO cells (231). The generality of splicing, preserving the 5' and 3' terminal sequences of primary transcripts, has also been inferred from the finding that a complex cDNA, transcribed from 5' triphosphorylated fragments of Ehrlich carcinoma hnRNA polyadenylated <u>in vitro</u>, is substantially hybridised by polysomal RNA (341), as is the bulk of the (3') poly(A) adjacent sequences (by weight).

The observed evolutionary divergence of the bulk of nucleus-confined sequences in poly(A)$^+$ RNA, including those sequences adjacent to poly(A) (375), suggests that in mammalian cells, the majority of nucleus-confined transcripts are not potential mRNAs: intron sequences show far greater evolutionary divergence than the exons of the genes they interrupt (329, 342). However, in sea urchins (178,343), higher plants (186) and frogs (177) there is convincing evidence that single-copy sequences which are expressed in some tissues or developmental stages but not others, are nevertheless uniquitously represented in nuclear RNA. This has been inferred by the use of heterogeneous single-copy DNA probes, and has recently been confirmed using cloned cDNAs in the sea urchin (312). Qualitative post-transcriptional selection therefore regulates tissue-specific gene expression in these organisms, with transcripts of expressed and unexpressed rare class mRNAs in a given cell-type being at very similar concentrations in the nuclear RNA (178). The question as to whether a similar phenomenon may have gone undetected in mammals remains open. Whether there is ample 'room' for such transcripts within the observed complexities of nuclear RNA (161,136,143) is debatable but even so they may be swamped by the much greater complexity due to the introns of both expressed and unexpressed genes, and so not detectable by the criterion of evolutionary conservation (375). A further uncertainty arises in the exact amount of the single-copy mammalian genome which is ever expressed on the polysomes, as no reliable study has been undertaken to determine the extent of tissue-specific gene expression, in terms of unique sequence complexities. Alternatively, qualitative post-transcriptional selection may only operate within a given committed cell lineage in mammals, such that the nuclear RNAs in a given lineage are lacking in the precursors to those mRNAs never expressed on the polysomes of that particular lineage.

In HeLa cell nuclear RNA two fairly discrete abundance components of complex class sequences have been observed by Holland et al (143), corresponding to transcripts present at 0.4 and 0.02 copies per cell respectively. Messenger cDNA cross-hybridisation has identified a substantial proportion of the former as mRNA sequences, in which case the rarer component might consist largely of decaying intron transcripts. The relative nuclear concentrations of intron and exon sequences of the ovalbumin gene, for example, are comparable with these figures (337). If this were the case, simple calculation shows that the complexity of the more abundant class of complex sequences in HeLa hnRNA is still far larger than that of polysomal mRNA, suggesting that it might contain mRNA sequences whose expression on the polysomes is blocked post-transcriptionally. Of possible relevance to this question is the observation that the disparity in nuclear and cytoplasmic complexities is much less evident in highly heterogeneous tissues (161,177,201) such as brain, where sequences post-transcriptionally repressed in one cell-type may be expressed in other cell-types which are also constituents of the tissue as a whole.

Nucleus-confined sequences adjacent to poly(A) have been detected by cross-hybridisation of nuclear cDNAs with poly(A)$^+$ or total polysomal RNAs. These reactions do not reach completion, but plateau at about 70 percent of the hybridisability of the cDNA, implying that some 30 percent by weight of the sequences adjacent to poly(A) in a given cell-type are unrepresented, or else exceedingly rare, on the polysomes, in mouse Friend cells,(346), Xenopus liver (176), HeLa cells (215) and mouse embryoid bodies (149). Hybridisations using cDNA probes enriched for these nucleus-confined poly(A)-adjacent sequences show that they are present in hnRNA at very low abundance, and account for the bulk of the poly(A)-adjacent sequence complexity (176,346). Attempts to demonstrate

their expression in the mRNA of heterologous cells have been inconclusive (345,346). The interpretation of this is obscured by the possibility that some, at least, of these sequences may represent oligo(A)-adjacent regions in mRNA-containing transcripts (279).

A possible example of a set of genes under qualitative post-transcriptional control in mammalian cells are the histone mRNAs during the cell cycle. Cytoplasmic histone mRNA was found to be detectable in HeLa cells during S phase, but not G1 phase (347,348), decaying rapidly after the cessation of DNA synthesis, whether at the natural end of S phase, or as a result of the introduction of inhibitors (349). Histone mRNA sequences have been reported in nuclear RNA, however, throughout the cell cycle (350,351), although (352) there is a contrary report. Since these studies all used heterologous (350,351) or uncloned (352) hybridisation probes, the question remains unresolved, although a recent report has questioned the cell-cycle dependence even of cytoplasmic histone mRNA (353).

A further set of observations supporting qualitative post-transcriptional selection as a mechanism of gene control is the measurement of the levels of differentiation specific mRNAs in nuclear and polysomal RNAs of tissues in which they are not expressed. Globin mRNA, for example, is detectable in liver nuclei, and also at a low level in liver polysomes, possibly due to blood cell contamination of liver. (354,355). Globin mRNA is also expressed at a high level in the nuclear RNA of at least one non-inducible variant of mouse Friend cells (356) and in a chick erythroblast line, transformed with avian erythroblastosis virus (357). In both cases, an altered selectivity of post-transcriptional processing appears to be responsible for the failure to express globin mRNA cytoplasmically. During adenovirus infection, cellular mRNAs disappear from HeLa cell polysomes, but are still detectable in hnRNA (358). All these 'qualitative' controls may, however, on closer examination turn out to be extreme examples of quantitative regulation.

2.3.4. Quantitative aspects of post-transcriptional selection

The operation of quantitative post-transcriptional controls has been inferred from the heterologous hybridisations of nuclear and polysomal cDNAs. Polysomal cDNAs are hybridised typically 1-2 orders of magnitude more slowly by poly(A)$^+$ nuclear than by poly(A)$^+$ polysomal RNA, for example in mouse Ehrlich carcinoma cells (168), mouse teratoma (164) or HeLa cells (215), and are present in a narrower range of concentrations in hnRNA (359,164). Nuclear cDNAs, conversely, are hybridised by polysomal RNA with similar or slightly accelerated kinetics (164,176,344), which suggests that while the bulk of nuclear transcripts do not alter profoundly in relative abundance post-transcriptionally, a small proportion of them, which are destined to constitute the high abundance polysomal mRNAs, are selected for very considerable post-transcriptional abundance modulation. This is supported by the use of fractionated abundant and rare polysomal cDNA probes. The disparity in their relative abundance is greatly diminished at the level of poly(A)$^+$ nuclear RNA, in mouse Friend cells (344), embryonal carcinoma cells (360), HeLa cells (215) and rat liver (224). Studies with individual, cloned cDNAs have yielded similar conclusions: abundant CHO cell mRNAs are an order of magnitude rarer in hnRNA (231); rat liver albumin mRNA is 100-fold less abundant in nuclear RNA (361), and IgG light-chain mRNAs are 200-fold concentrated in polysomal poly(A)$^+$ over nuclear poly(A)$^+$ RNA in mouse myeloma cells (333). Sequences which shift in polysomal abundance between cell-types can be at similar nuclear abundances, such as is the case for at least one cloned CHO cell mRNA, raised in abundance in hamster liver (362).

Since most studies comparing nuclear and polysomal abundances have used steady-state nuclear RNA, the contribution of post-transcriptional events to the pattern of relative

abundances may have been underestimated. However, in HeLa cells, the composition of pulse-labelled poly(A)-adjacent fragments of hnRNA is very similar to that of steady-state hnRNA, as judged by the kinetics of their hybridisation with excess polysomal cDNA (363). The qualitative and quantitative aspects of post-transcriptional selection in these cells are not radically affected by the use of steady-state hnRNA. A similar result has been inferred from the hybridisation characteristics of pulse-labelled hnRNA in sea urchin gastrulae (181), which indicate that the rare sequences are turning over at a similar rate to the more abundant sequences.

A major uncertainty associated with post-transcriptional sequence selection in many cell-types concerns the relative homologies, complexities and abundance patterns of polyadenylated and non-polyadenylated fractions, and whether polyadenylation, as such, has any bearing on post-transcriptional stabilisation of particular sequences. In sea urchin embryos, only 8 percent of the poly(A)$^+$ derived polysomal cDNA is hybridised by poly(A)$^-$ RNA (364). In HeLa cells, translation data have suggested that some abundant poly(A)$^+$ mRNAs are also present in the poly(A)$^-$ fraction, but that few if any of the poly(A)$^-$ mRNAs which are at relatively high abundance, are unique to this compartment (365), despite drastic differences in abundance. The extent and kinetics of cDNA hybridisations suggest a minimal overlap between the two populations of sequences (365,366), with only a small fraction (about 10 percent by weight) of the abundant poly(A)$^+$ mRNAs behaving bimorphically. However, single-copy DNA saturation has shown that the total complexity of the poly(A)$^-$ mRNA fraction is only a small proportion (about 10 percent) of that of the poly(A)$^+$ compartment in HeLa cells (143). Similar findings have been reported for embryonic rat myoblasts (367), mouse brain and liver (201), Drosophila tissue-culture cells (183), yeast (195) and tobacco leaves (185,186). Contrary reports exist, however, for mouse brain (163) and liver (158),

polyoma-transformed mouse tissue-culture cells (158) and rat brain (147), where the compartments have been reported as being essentially non-overlapping sets of similar, high complexity. At the level of nuclear RNA, the literature is equally confusing, the poly(A)$^-$ sequences having been reported as being distinct and of high complexity in mouse brain (161) and soyabean tissue-culture cells (189), but of relatively low specific complexity in Drosophila tissue-culture cells (183) and mouse Friend cells (165). Relative abundances between the two compartments have not been reported in detail. Evidence that polyadenylation may be a level at which post-transcriptional abundance modulation is brought about will be discussed subsequently (see Section 2.3.6).

2.3.5. Differential gene expression between cell-types: qualitative and quantitative aspects

Studies of differential gene expression in development, differentiation, growth and carcinogenesis, by comparing the composition of nuclear and polysomal RNA populations, give further evidence of the complexity and diversity of gene control mechanisms. Comparisons of the sequence sets expressed in nuclear RNA of different mammalian echinoderm, and plant tissues has shown that despite a proportion of tissue-specific transcripts, many transcribed sequences are held in common. For example, in the rat, the nuclear RNAs of brain, kidney and liver constitute a nested set (136), as do the poly(A)$^+$ nuclear RNAs expressed in whole rat brain, functionally distinct regions of it, and cultured neural cell-lines (307). Nuclear RNAs of different sea urchin developmental stages also share a large block of common sequences (368), the tissue-specific component being very variable in size, ranging from the undetectable, in the case of gastrula and pluteus stages (343), to an amount greater than the polysomal mRNA complexity,in adult intestine nuclear RNA compared with total gastrula nuclear RNA (368).

The nuclear RNA complexity in a teratoma-derived myogenic line is increased by about 30 percent relative to the undifferentiated embryonal carcinoma (360), which may imply an increase in transcriptional repertoire during early development. The data on the composition of nuclear and polysomal sequence sets during avian and mammalian development are insufficient to permit any conclusion as to whether changes in the latter are brought about transcriptionally or post-transcriptionally (286,345,174,159).

The size of the tissue-specific component of polysomal RNA sequence sets is remarkably small, both in adult mammals and birds, and in cell-lines derived from them. For example, more than 80 percent of the sequences expressed in chick liver and oviduct are held in common (174), and the situation is similar in mouse (159,222,286) and human (197) tissues except brain, where most of the sequence complexity is specific for the one tissue, and is a significant fraction of the total mRNA by weight (about 30 percent) (171). Mouse Friend cells, a committed cell-line, share most of their polysomal sequences with transformed fibroblasts (369), or embryonal carcinoma cells (170,370). Closely related tissues may express sequence sets indistinguishable by single-copy DNA saturation technology, for example rat prostate and seminal vesicle (153). Cell commitment and development are characterised at the mRNA level (as well, as has already been suggested, at the level of hnRNA) by an expansion of the set of expressed sequences. For example, rat uterine development is marked by the appearance of some 40,000 new mRNA species (152),which may represent the greatly increased cellular heterogeneity of the specialist tissue. Mouse embryonal carcinoma cells share their 7500 mRNA sequences with myoblasts derived from them, but the latter also express a further 4500 mRNAs not found in the undifferentiated cells (170). As committed cells pass through the later stages along their differentiation pathway, however,

changes become more subtle. Some studies have suggested that the set of expressed sequences may narrow again in response to the demands of specialisation, for example as in myoblasts whose kinetic mRNA complexity declines from about 28,000 to 10,000 mRNAs during fusion to form myotubes (156). Paterson and Bishop (309) found no such complexity change in primary myoblast cultures, using a single-copy DNA probe, and it is therefore possible that the findings of Leibovitch et al (156) are due to the increasing difficulty of making kinetic measurements of the complexity of very rare mRNAs when superabundant mRNAs are being induced to a high level, as here in myogenesis, or indeed, in other differentiating systems. A similar finding of unchanged polysomal complexity has been reported for mouse neuroblastoma cells induced to 'differentiate' with bromodeoxyuridine (172), although some new sequences, of undetermined complexity, were detected. The ultimate steps in mammalian differentiation may occur without changes in the overall sequence complexity of the polysomal RNA, as in Friend cells induced to hemoglobinise in culture (369) (where the use of an enriched 'rare' sequence cDNA probe failed to detect any differences (170)), the rhizobium-induced differentiation of root nodules in leguminous plants (190), and the many cases of hormonal activation of specialised cell functions. These include chick oviduct stimulation by oestrogen (381), which is also characterised by an unchanged set of sequences detectable in nuclear RNA (175), the androgen-induced specialisation of the rat prostate (154) or seminal vesicle (153) and thyroid-hormone stimulation of liver function in rats (372). Comparisons between quiescent and growing cells also shows no change in mRNA sequence sets. These studies include mouse kidney (373) or rat liver (145, 374) regeneration, and cultured mouse fibroblasts in the growing or non-growing state (141).

In plants and in echinoderms, the size of stage- or tissue-specific polysomal sequence sets is somewhat larger, constituting about a third of the total mRNA complexity in various tobacco organ systems (186), or in sea urchin gastrulae compared with blastulae (179). Moreover, the developmental programme in sea urchins appears to be quite distinct from that in the higher eukaryotes which was hypothesised above. The fertilised egg contains a high complexity mixture of maternal mRNAs, most of which accumulate during the latter stages of oogenesis (180), and which declines in complexity during development (179). Early embryos synthesise a set of mRNAs very similar to that of the oocyte (376). However, the process of cell commitment which begins very early is marked by the asymmetric repartition of these mRNAs to the daughter cells, such that each receives (and then proceeds to synthesise) a distinct subset of single-copy mRNAs (377). Isolated micromeres, for example, express only 70-80 percent of the total 16-cell stage sequence complexity (377). Later during development, an increasing number of stage-specific mRNAs are newly expressed. In individual adult tissues, the number of mRNAs expressed is small compared with the number expressed in whole embryos at any given stage of development (179), a situation paralleled to some extent in the mouse, where whole embryo polysomal RNA is much more complex than that of individual, isolated embryonic cell-lines (173).

In Drosophila, the developmental programme is possibly different again, with a rather low complexity of egg RNA (182), which remains essentially unchanged during the first 19 hours of development, losing only about 40 mRNA species (378). Thereafter, new sequences begin to accumulate. In all these cases, however, many more diverse genes seem to be expressed during development than are required to encode the house-keeping functions common to all cells, or the specialist functions of adult tissues. A similar situation may even

pertain in Euglena, where the number of mRNA sequences expressed in mature light-grown cells is considerably less than that expressed in dark-grown cells newly exposed to the light, and undergoing chloroplast development (191).

With the exception of early development, it therefore seems that the vast differences in phenotype between the cell-types of a eukaryote are not principally the result of qualitative differences in the set of mRNAs expressed, and it is in the direction of quantitative differences, at the level of mRNA abundance, that control mechanisms must be sought.

The assumption that phenotypic differences result principally from differences in relative mRNA abundance is supported by the results of numerous cDNA cross hybridisation experiments. Sequences abundant in one tissue are present, but rare, in others (153,159,222), and even super-abundant mRNAs characteristic of specialised cells may be expressed at a low level in unrelated, heterologous tissues. Such is the case for chick ovalbumin (174,335) or mouse globin (354,355), although there are contrary findings (379), and the presence of a sub-population of undifferentiated cells which remain multipotent cannot be entirely ruled out as an explanation. The accumulation of these differentiation-specific mRNAs has been followed both <u>in vivo</u> and in cell culture. The high abundance mRNAs of adult rat pancreas accumulate 500-fold during fetal development (from day 14 to adulthood), during which time the relative concentration of those sequences held in common with hepatoma cells declines 10-fold (151). Similarly, a high abundance class of mRNAs appears during myoblast differentiation in culture (309), following oestrogen treatment of chick oviduct (380) or androgen stimulation of rat prostate (154), in Friend cells induced to differentiate <u>in vitro</u> (369), or rhizobium-induced root nodule formation in soyabeans (190). The high abundance component disappears from oviduct polysomes during hormone withdrawal (381).

More subtle abundance changes have also been recorded during differentiation and changes in proliferative state. In Friend cells, some moderately abundant sequences decline in abundance following DMSO induction (370), and some rarer, non-globin mRNAs may become more abundant (382). Changes in abundance also accompany the earliest stages of uterine growth (383), neuroblastoma differentiation in culture (324) and Xenopus (223) and sea urchin (624) development. Specialised quiescent cells, when stimulated to re-enter the cell cycle and ultimately divide, such as regenerating mouse kidney (373) or rat liver (145), show a modest depletion of abundant mRNAs. The transition from quiescence to growth may be associated with an increased relative concentration of a set of mRNAs which enter the moderate abundance class. This has been inferred from hybridisation studies of rat liver regeneration (374), and of a yeast mutant temperature-sensitive for growth (371) where the abundance changes seem to be confined to about 100-150 mRNA sequences. Serum stimulation of resting mouse fibroblasts is also marked by only a very modest shift in mRNA levels, with at most 3 percent of the sequences altering radically in concentration (141). Similar conclusions have been drawn for viral (385,386) or chemical (173,387) transformation of cells in culture where, despite the dramatic change in phenotype, only subtle quantitative changes appear to take place in the mRNA population. The only qualitative changes which do seem to occur in proliferating or neoplastic cells, as compared with quiescent homologues, affect the subcellular distribution of repetitive sequence transcripts, which may not have any direct bearing on mRNA expression. Sequences of this type, confined to the nucleus in normal liver, have been detected in the cytoplasm of regenerating liver (388) and hepatoma cells (389), although repeat sequence expression in cellular RNA is unaffected by polyoma transformation of mouse cells (390).

Differentiation, growth and carcinogenesis are also marked by subtle changes in poly(A)$^+$ nuclear RNA populations (369,206, 175, 374). Small qualitative changes have been reported in nuclear RNA in 16-hour regenerating rat liver (148,374) even when the polysomal mRNA population is unchanged. These may represent the induction of new nucleus-confined transcripts of unknown function, or of sequences which are to be expressed on the polysomes at a later stage of the regenerative response.

Various biochemical stresses, such as heat-shock, anoxia, or metabolic poisons (391,392) induce a dramatic change in the mRNA population in Drosophila cells, both in vivo and in cell culture, with a relatively small group of previously rare mRNAs being transcribed at a high level, and displacing all other sequences rapidly from the polysomes. A similar effect is seen in cells of higher eukaryotes, for example chick embryo fibroblasts in culture (393).

Studies using cloned, or otherwise highly purified cDNAs, have confirmed and added detail to the information deduced from heterogeneous cDNA hybridisations. Probes specific for α and β globins have confirmed the accumulation of these mRNAs in induced Friend cells (394). Hormone-mediated increases in the abundance of differentiation-linked mRNAs are seen in the glucocorticoid or prolactin induction of casein mRNA in mouse mammary gland in vivo (395), or in rat mammary organ culture (396), oestrogen-induced accumulation of ovomucoid (397) and conalbumin (398) mRNAs in chick ovidcut, and of preprolactin mRNA in rat pituitary (399), androgen induction of a mouse kidney-specific abundant mRNA (400) and of major urinary proteins in mouse liver (401), and progesterone induction of uteroglobin mRNA in rabbit endometrium (402). Ovalbumin (403), uteroglobin (402) and conalbumin (transferrin) (398) have been shown to be under differential hormonal control in different tissues. A growth-related

promotion of a rare mRNA to moderate abundance is seen for actin mRNA in Dictyostelium, following spore germination (105). Developmentally regulated changes in mRNA abundance have been reported for Xenopus (404), Dictyostelium (311), the silkmoth (232) and Drosophila (313) using cloned cDNA libraries. Gene switching during development has been shown in the case of the sea urchin histone gene variants (406). Changes have also been reported in the level of differentiation-linked mRNAs during development, such as for serum albumin and α-fetoprotein in rat liver (407).

2.3.6. Inferences concerning gene control mechanisms

The preceding sections illustrate the degree to which, in any eukaryotic cell, qualitative and quantitative controls determine the population of sequences which are transcribed and subsequently transmitted to the polysomes. They further show the effects of altered qualitative and quantitative controls on the cell's phenotype. This section will attempt to relate these two phenomena, by examining the evidence that particular regulatory events affecting phenotype are brought about at particular levels in the pathway of cellular information flow.

The best evidence for transcriptional regulation of gene expression, from hybridisation studies of *in vivo* RNA populations, is provided by comparing the proportions of particular sequences in the population of newly synthesised molecules, after a brief pulse-label. Thus, globin mRNA is increased 10-fold in pulse-labelled RNA after DMSO induction of Friend cells (408). If the labelling time is increased from 50 seconds to 15 minutes, the degree to which globin is specifically induced is greatly enhanced, reaching some 280-fold, indicating the co-ordinate operation of transcriptional and post-transcriptional selection. However, even the use of very short pulses cannot distinguish with certainty between

transcriptional and immediate post-transcriptional regulation. The relative contribution of post-transcriptional controls to overall mRNA abundances can similarly be deduced by measuring the concentrations of particular mRNA sequences in steady-state nuclear and polysomal RNA, before and after changes known to induce altered abundances. Using Friend cell induction once more as an example, clone M2 shows a 9-fold increase in globin mRNA levels in poly(A)$^+$ nuclear RNA, but a 50-fold enhancement at the level of poly(A)$^+$ polysomal RNA several days after induction (369). Similar measurements on the rate of pulse-labelling, and steady-state abundances of mRNAs corresponding to clones selected from a CHO cell cDNA library (362), suggest the operation of both transcriptional and post-transcriptional controls in determining altered abundances between tissues. By contrast, such methods indicate that the glucose-repression of yeast iso-1-cytochrome c is brought about wholly at the level of transcription (409).

Hybridisation studies <u>in vivo</u> have not succeeded in contributing any more precise information on the particular step(s) in transcription which might be subject to regulation. Studies of adenoviral transcription have suggested that premature transcriptional termination occurs frequently, and may be responsible for determining the relative levels of different mRNA products sharing a transcriptional start (410), although there is no evidence that such a mechanism may operate for cellular genes. The drug DRB, however, appears to interfere with transcription by promoting premature termination events (411), possibly at splice-donor sites (412). Genes with no introns such as adenovirus polypeptide IX (413) appear to be resistant to its effects, suggesting that altered premature termination frequency <u>in vivo</u> may generate a novel pattern of relative transcript frequencies. Alternatively, if DRB is in fact changing the specificity of the splicing

system, this may imply that relative levels of spliced and unspliced mRNAs could be regulated in vivo by a similar modulation.

The disparity in the concentration of abundant and rare Friend cell mRNAs is much less at the level of poly(A)$^-$ than poly(A)$^+$ nuclear RNA (344), suggesting that polyadenylation may be a selective mechanism responsible for generating differential abundances. There are few examples, however, of sequences shifting between poly(A)$^+$ and poly(A)$^-$ compartments, where this may be correlated with alterations in relative abundance, but myosin heavy-chain mRNA accumulation in fusing myoblasts is accompanied by a radical increase in the proportion of non-polyadenylated molecules synthesised (414) (defined as those with 13 or fewer A residues).

Splicing as a potential control point in the determination of mRNA abundances has been inferred indirectly from a number of experimental results. A sequence-specific lesion in splicing appears to be responsible for the non-infectivity of SV40 in undifferentiated mouse teratocarcinoma cells, where unspliced early region transcripts are detectable in nuclear RNA, but not in the cytoplasm (415). Following retinoic acid-induced differentiation, correctly spliced viral mRNAs accumulate in the cytoplasm, and infectivity is restored (416). Differential splicing of a given primary transcript generates a family of mRNAs in early and late papovavirus gene expression (323,324) and may also occur in adenovirus (417), although there is no clear evidence that relative splicing efficiencies are modulated in vivo to bring about changes in mRNA levels. Differential splicing has also been postulated to occur in lymphocytes, facilitating the co-expression of immunoglobulin μ and δ heavy chains carrying the same V region (11,418), or receptor and secretory forms of the same IgM μ-heavy chain, differing at their

carboxy termini (419). The latter may represent an example of indirect abundance modulation by the choice of polyadenylation site, since this may determine which of the two alternative splicing patterns is adopted. Differential splicing may also explain the detection of multiple mRNA species in sea urchin embryos, sharing a region of single-copy sequence (420), and of a variant of human growth hormone (421,422).

An important question, which may at least partially be resolved by hybridisation studies in vivo, concerns the relative contribution of intranuclear and post-nuclear events to the post-transcriptional determination of mRNA abundances. The hybridisation kinetics of newly released HeLa cell (pulse-labelled) cytoplasmic RNA, with excess 'steady-state' polysomal cDNA, is very similar to the kinetics of the reaction between steady-state polysomal cDNA and its template, in excess (363). This implies that in HeLa cells, post-nuclear events contribute little to the overall pattern of relative mRNA abundances. A more exhaustive study of the possible influence of relative cytoplasmic stabilities on abundance, was carried out by Lenk et al (423), who followed changes in the kinetics of homologous and heterologous hybridisations, after actinomycin D inhibition of new mRNA synthesis, and also compared the kinetics of hybridisation of labelled, newly released RNA with excess steady-state cytoplasmic cDNA. These studies confirmed the unimportance of stability as a determinant of relative abundance in HeLa cells but, conversely, found that stability and abundance were closely related in Drosophila tissue-culture cells. A similar finding in mouse L cells (424) is accompanied by the observation that shorter RNAs are also strikingly enriched in abundant sequences, comprising a high proportion of stable species. An inverse relationship has also been reported in cultured insect cells (425,426) and resting human lymphocytes (427), between mRNA size and stability, suggesting that mRNA lifetimes may be determined, at least in some cases, by stochastic (one-hit) decay, itself

related to messenger length. Individual high abundance mRNAs, such as those for globin (428), ovalbumin (429), myosin (430) and silk fibroin (431) have also been shown to be of significantly higher stability than the bulk of mRNAs in the cells in which they are expressed. As Tobin (110) has pointed out, however, the range of observed stabilities is insufficient to account for the range of mRNA abundances documented by hybridisation studies, and additional transcriptional and post-transcriptional mechanisms must be operating concomitantly to produce the observed patterns of abundance.

Evidence for altered mRNA abundances arising directly from changes in relative mRNA stability is provided in a number of cases of hormonal induction of specific mRNA function. Prolactin, for example, specifically increases the half-life of casein mRNA in rat mammary organ culture by 17-25-fold, which together with an effect on the rate of its synthesis, raises the steady-state level of the messenger 100-fold (396). During hormone withdrawal, the oestrogen-inducible mRNAs of the chick oviduct are selectively destabilised (381,432). Globin mRNA, however, is found to be equally long-lived in both induced and uninduced Friend cells (428), although its half-life is specifically decreased late in induction.

The possible contribution of translational control to polysomal abundance patterns can be studied by comparing the sequences in polysomal and non-polysomal poly(A)$^+$ cytoplasmic RNA. The only thorough investigation of this, using cDNA cross-hybridisation methodology, has been carried out in CHO cells, where the polysomal and informosomal populations were found to be of similar kinetic complexity, to exhibit complete homology, and to show only minor abundance differences (less than 2-fold, on average, in either direction) (433), which were confined to the most abundant sequences. Similar results

were found for mouse Taper hepatoma (214) and Drosophila oocytes and embryos (434), although this contradicted earlier findings (435). Studies using cloned, or otherwise highly purified cDNA hybridisation probes, have revealed a more complex pattern. Myosin heavy-chain mRNA is preferentially located in untranslated mRNPs in pre-fusion myoblasts, whereas it is actively translated in fused myotubes (405). Conversely, the modulation of albumin synthesis in rat liver, also directed by a highly abundant mRNA, is independent of translational controls (436-438). The same is true for the differential expression of rat α-fetoprotein in liver and hepatomas (439), the synthesis of rat pancreatic amylase during development (440) and the expression of stage-specific histone mRNAs in sea urchin embryos (441). Using clones from a Xenopus embryo cDNA library, Dworkin and Dawid observed that mRNAs that were predominantly informosomal at any given stage remained so throughout development (404). However, in heat-shocked Drosophila cells, non-heat-shock mRNAs go into a reserve pool of untranslated mRNPs (392), as newly synthesised heat-shock mRNAs sequester the translation apparatus. More subtle forms of translational control may operate in relation to mRNA stability and translatability on free as opposed to membrane-bound polysomes, as appears to be the case for ferritin mRNA in liver (442). Although untranslated mRNPs do not in general appear to constitute a distinct set of sequences of high complexity, the evidence that subtle translational modulation may sometimes be a determinant of polysomal abundance, should be taken into account wherever hybridisation studies show altered polysomal abundances as responsible for a change in phenotype.

2.4. In vitro systems for studying eukaryotic gene expression

2.4.1. Classes of in vitro system

In vitro systems which have proved useful in the study of eukaryotic gene expression, using nucleic acid hybridisation methodology, may be divided into three categories. Firstly, are those in which an exogenous gene or genes, usually in the form of a foreign nucleus or a cloned DNA sequence, is introduced into what is essentially a viable living cell (see sub-section 2.4.2). By analysing the RNAs (and polypeptides) synthesised in the acceptor cell considerable information may be derived concerning how the transferred gene is expressed in that cell, or what features (for example, particular regions of its primary DNA sequence, or its structural configuration) are involved in its correct expression or regulation. A further development of some systems of this type is the introduction (or co-introduction) into the recipient cell of putative regulatory macromolecules governing the expression of the transferred gene. The only serious drawback to the use of such systems is that they tend to treat the whole gene expression machinery as a 'black box', in which it is often difficult to detect anything other than the end product of transcription and processing. Their great advantage, however, is that they involve the minimum of perturbation to the cell's genetic apparatus, and are therefore free of many of the artifacts which have plagued cell-free systems. A number of assumptions must still be made, however, in the interpretation of results using gene transfer systems. Firstly, it is assumed that all components necessary for the correct expression of the foreign gene are present in the recipient cell. Secondly, the fact that an abnormal gene dosage has been set up is generally not considered, and the possible effects involving limiting components of the gene expression machinery are ignored. Thirdly, it is assumed that the introduction of a gene in a state very different from its native configuration (e.g. as naked supercoiled DNA in association with a prokaryotic plasmid) does not influence its expression. Since all three

of these assumptions have not proved universally valid, experiments should be interpreted with some caution, although these findings in themselves contribute considerable information on how genes are selectively expressed. Gene transfer systems are the method of choice for investigations of the role of cytoplasmic factors on the control of gene expression.

A second type of *in vitro* system is the use of isolated nuclei, often in association with a surrogate cytoplasm fraction, in which mRNA transcription, processing and transport may be followed, after particular pre-treatments of the cells from which they are derived, or in response to specific manipulations *in vitro* (see sub-sections 2.4.3 and 2.4.4). Analysis may be at the level of an individual gene, or of whole populations of sequences, depending on the type of hybridisation probe used. The main problem of such systems is that the micro-environment of the nuclei *in vitro* may be very different from that which they experience in the cell *in vivo*, in ways that may not easily be measured or controlled. The properties of isolated nuclei may therefore be due to spurious processes unrelated to normal gene controls. Great attention must therefore be paid to the design of stringent assays for *in vivo* equivalence, which is particularly difficult for those processes which are not easily accessible to study *in vivo*, such as nucleocytoplasmic transport. Despite this ever-present danger of artifacts, isolated nuclei systems possess the great advantage that they permit experimental manipulation of the expression of a natural gene, in what approximates to its native state. They are in many ways, therefore, a compromise between living-cell gene transfer systems, and soluble cell-free systems. Like gene transfer systems based on whole cells, there is a tendency to treat the gene expression machinery as a 'black box', although there is at least a clear distinction here between intra-nuclear and post-nuclear events, and because *in vitro* manipulations are feasible, further dissection of the level

at which regulation is occurring can be achieved. They are also useful for examining the effect of cytoplasmic factors on nuclear processes, although there is uncertainty as to the true cellular origin of factors in 'cytosol' which may influence nuclei, a problem which does not arise in whole unruptured cell systems.

The third class of *in vitro* system comprises those in which soluble extracts of cells or nuclei are used to transcribe or process substrate DNAs or RNAs, which are either present already in the extracts or, more usually, added exogenously (see sub-section 2.4.5). Because such systems are the result of the greatest disruption of the cellular machinery, they are the most prone to artifacts, and the most difficult to characterise from the viewpoint of *in vivo* equivalence. Nevertheless, they remain the only feasible way of examining individual levels of regulation in isolation from other processes (other than the equally undesirable use of supposedly specific inhibitors). However, this assumes that this type of separation of the processes involved in gene expression, itself has a physiological basis.

2.4.2. Gene transfer into living cells

Early cell fusion experiments (443,444) led to the general conclusion that cytoplasm contains regulatory elements governing nuclear events in gene expression. For example, otherwise inert chick erythrocyte nuclei can be induced to synthesise RNA and chick specific proteins, after fusion with HeLa cells (443), this being accompanied by the migration of HeLa cell cytoplasmic proteins into the erythrocyte nucleus. The activation of differentiation-specific markers in hybrid cells, such as of mouse albumin in mouse lymphoma x rat hepatoma hybrids (445), or the extinction of such functions, for example of globin synthesis in Friend cell x mouse fibroblast hybrids (446) have been taken to indicate the presence of trans-acting cytoplasmic activators or repressors specific for given genes

or sets of genes. The interpretation of such experiments is in some doubt, however, due to the many instances of chromosome segregation, the uncertain developmental competence of the parent cells, and the possible artifacts induced by fusion techniques, and more recent findings have cast doubt on some of the earlier results. The existence of trans-acting repressors is disputed, for example, following the observed co-expression of alternative differentiation markers in hybrids between different murine hematopoietic cell-lines (447). Appropriate regulation of the human β-globin gene has been reported in hybrids between human non-erythroid (lymphoblastoid) cells, and Friend cells (448), in which human globin expression is DMSO-inducible, and dependent on the retention of a particular human chromosome. Trans-acting activators are therefore apparently able to induce the expression of a gene in response to an external signal.

A rather more defined environment, for the investigation of the possible role of cytoplasmic factors in gene activation, is provided by the fusion of karyoplasts with enucleated cytoplasts (499), both obtained by centrifugation of cyto-chalasin B-treated cells. Mouse fibroblast nuclei, transplanted into rat hepatoma (HTC Cell) cytoplasts are able to synthesise mouse tyrosine aminotransferase, a liver-specific enzyme, which is, moreover, steroid-inducible in the hybrid cells (450). This may prove a useful system for studying the identity of the regulatory element involved in the activation, and the mechanism by which it acts. Nuclear transplantation into amphibian eggs or oocytes (2,451) also results in specific gene activation or repression. After injection into Xenopus oocytes, HeLa cell nuclei synthesise only some of the proteins which they manufacture <u>in vivo</u>, and the regulation appears to be at the mRNA (pre-translation) level (451). In addition, some new, non-HeLa, non-Xenopus polypeptides are synthesised. Since this reprogramming appears to be once again associated with the migration of oocyte cytoplasmic proteins (including newly synthesised species) into the donor

nuclei, the system is potentially useful for assaying the regulatory potential of macromolecular fractions co-injected with the nuclei. Nuclear or cytoplasmic proteins injected into oocytes rapidly partition themselves between oocyte nucleus and cytoplasm, in accord with their normal cellular location (452), suggesting that the oocyte is capable of identifying potential gene regulatory molecules and directing them to the sites of their action. Although specific gene activation and repression clearly occur in transplanted HeLa cell nuclei, it could be argued that the mechanism(s) by which they are brought about may be unphysiological since mammalian nuclei are responding to factors in amphibian cytoplasm. A more easily interpreted example of the use of the oocyte injection technique is the introduction of somatic Xenopus (kidney cell) nuclei into the oocytes of another amphibian species, Plurodeles (453). This results in the specific activation of a set of Xenopus oocyte proteins, and the cessation of all protein synthesis characteristic of kidney cells. Similar reprogramming has been reported for liver nuclei transplanted into oocytes, using two distinct species of axolotl (454), in which liver-specific alcohol dehydrogenase expression was extinguished, but both 'donor' and 'host' forms of lactic dehydrogenase continued to be expressed.

Another potent method for introducing foreign genetic or gene regulatory material into cells is the use of erythrocyte ghosts, containing the test substance(s), which are fused with recipient cells (455). This method has confirmed the ability of exogenous non-histone chromosomal proteins to be incorporated into chromatin, where they persist (456).

Micro-injection into oocytes is proving a powerful technique for studying the regulation of expression of purified DNA preparations, particularly of cloned genes. Following micro-

injection, efficient and correct transcription and processing has been reported (where appropriate, including RNA splicing) for yeast tRNA genes (457,458), at least 2 early and 2 late genes encoded in SV40 DNA (459,460), sea urchin histone genes H2A and H2B (although not the remainder) (461) and Xenopus 5S DNA genes (462). Circularised DNA, such as that of SV40, micro-injected into Xenopus oocytes, adopts a supercoiled configuration, and associates into a nucleoprotein complex which resembles the native minichromosome (464). The formation of this complex can be reproduced in vitro in a cell-free system derived from components of the Xenopus egg (464). The processing events which occur during the maturation of tRNA, transcribed from micro-injected cloned tDNA, all take place in the oocyte nucleus, and follow a fixed order (458). For micro-injected sea urchin histone genes, the TATAAA box apparently functions as a specificity element in transcription, since its deletion generates a novel set of 5' ends (465). Furthermore, deletion of the normal cap site induces transcription to start at a new unique site, 24 nucleotides downstream of the TATAAA box, and deletions further upstream cause either an increase or a decrease in the amount of transcription suggesting that sequences 5' to the gene may interact with transcriptional regulators in vivo. The use of oocyte micro-injection, combined with specific modifications to the substrate DNA (so-called surrogate genetics) is clearly a powerful technique for identifying regulatory signals, and elucidating how these are utilised in the cell.

Chick ovalbumin DNA is also transcribed and processed in Xenopus oocytes, generating functional ovalbumin mRNA (466) albeit at a low level compared with SV40 DNA. This is perhaps not surprising, in view of its much more complex maturation pathway. The level of ovalbumin mRNA synthesis in oocytes does not, however, depend upon cap site and TATAAA box sequences,

and hybridisation analysis of ovalbumin RNA-containing transcripts shows that these are mostly read-through products which are not correctly initiated or terminated. Cloned ovalbumin cDNA is also transcribed to a similar extent, but gives rise to no functional ovalbumin mRNA, suggesting that sequences within the introns may be involved in specifying the correct processing pathway, or that splicing and nucleo-cytoplasmic transport may be intimately connected. Since correct transcriptional initiation and termination do not appear to be required for correct processing, it would appear that the requirements for accurate splicing do not include the whole of the structure of the transcript. These findings illustrate that the complete information for the expression of the ovalbumin gene is not provided by the cloned chromosomal fragment. The missing information may be contained in DNA sequences more distant from the structural gene, or in proteins with which the DNA is complexed in its native state in chicken cells. Alternatively, they may indicate that the oocyte is an inadequate system for ovalbumin DNA expression and studies of how it is controlled.

Amphibian oocytes have also proved a useful system for studying the structural features of mRNAs which are important as determinants of translational efficiency and stability (91, 92).

An additional method for introducing purified DNAs into eukaryotic cells is by the use of a eukaryotic virus as a cloning vector. Mouse β-globin major has been successfully synthesised in mouse cells in culture, infected with SV40/ globin DNA recombinants (467), where transcription is from a viral promotor. Using such recombinants with varying organisations of viral and chromosomal sequences, it has been shown that splicing is dependent on a correctly oriented splice-donor and splice-acceptor site in the primary transcript (468) and that splicing is required for the formation of a

stable RNA product, at least in this case. This supports the idea that splicing may be a major control point in mRNA synthesis.

Direct DNA transfer into cells, using co-transformation with a selective marker, also provides a promising system for studying the factors controlling gene expression. For example, mouse L cells, stably transformed with cloned rabbit β-globin DNA, express rabbit, but not mouse β-globin (469,470). This implies one or more of the following: either the normal repression of β-globin synthesis in these cells depends on DNA sequences far from the gene, acting in cis, or upon structural features of the gene, such as methylation or association with chromosomal proteins, not shared with the transforming DNA, or that the control sequences associated with the gene, which interact with a trans-acting repressor, have diverged significantly between rabbit and mouse. All the transcription and processing signals necessary for rabbit β-globin mRNA synthesis, however, are correctly and efficiently utilised by the heterologous cell. In other cases this is not so, however. In yeast cells, transforming rabbit β-globin DNA is aberrantly transcribed, the transcript remaining unspliced (471), despite the presence in the yeast genome of quite recognisable signals governing transcriptional initiation and splicing (472). Cloned ovalbumin DNA is also transcribed aberrantly in mouse L cells when introduced by DNA-mediated gene transfer, with initiation occuring considerably upstream of the normal cap site (473), although the transcript is correctly spliced and polyadenylated to yield ovalbumin mRNA, albeit at a low level. Other investigators, using similar technology, have detected ovalbumin synthesis in L cells transformed with the chick gene (474). Thus far, cloned eukaryotic genes used to transform cells concomitantly with a selective marker such as the herpesvirus thymidine kinase gene, seem in general to be constitutively expressed, hence such systems may be of limited

use in studying their normal mode of control. Chromosome-mediated gene transfer where, in stable transformants, small fragments of donor chromosomes are integrated into recipient cell chromosomes (476) may prove a more useful model system for normal regulation. This would obviously be so if this involves gene positioning in an overall chromosome structure, or control sequences relatively distant from the structural gene. The inappropriate regulation of genes introduced by DNA-mediated gene transfer may be contrasted with correct regulation following the introduction of whole chromosomes, using cell fusion technology (448).

Highly efficient gene transfer has now been achieved by direct fusion of recombinant plasmid-bearing bacterial cells with eukaryotic recipient cells (477). This should, at the very least, provide a powerful method for isolating genes encoding particular functions expressed normally in rare mRNAs, and studying the factors governing their expression.

2.4.3. In vitro transcription studies using isolated nuclei

Extended RNA synthesis in isolated nuclei, sensitive to low levels of α-amanitin (i.e. resulting from transcription by RNA polymerase II), has been reported for a wide variety of cell-types, including mouse myeloma (478) or L cells (479), HeLa cells (480), adenovirus-transformed (481) or infected cells (482), SV40-infected monkey cells (483), chick oviduct (397,484,485) and liver (486), rat liver and hepatoma cells (487,488), Xenopus tissue-culture cells (489) and embryos (490), Drosophila tissue-culture cells (491) and embryos (492) and slime-moulds (493). Only in a minority of cases, however, has chain re-initiation been conclusively demonstrated, for example by the incorporation of γ^{32}P-labelled nucleotides into polyphosphorylated 5' termini in L cells (479) or chick liver (486), the incorporation of label into cap-adjacent nucleotides in adenovirus-infected HeLa nuclei (494), or by

the retention on mercury-sepharose columns of *in vitro* transcripts made by Drosophila cell nuclei in the presence of 5' γ-sulphurated purine nucleotides (491). Re-initiation *in vitro* has also been reported for mouse myeloma nuclei (495), on the basis of buoyant density measurements of transcripts synthesised in the presence of mercurated nucleotides. Reported conditions for efficient *in vitro* transcription vary widely, and most systems have been optimised for elongation rather than re-initiation efficiency. Both the length and poly(A) content of *in vitro* transcripts appear to depend on the nature of the divalent cation, and the KCl and ribonucleotide concentrations (496). Non-ionic detergent treatment of rat pituitary tumour nuclei does not impair their capacity for synthesising hnRNA-like molecules (497).

Attempts to demonstrate an effect on transcription by cytosol have generated contradictory findings. Cytosol seems to be required in many instances for extended RNA synthesis (498, 499). One cytosol-independent system, however, based on rat pituitary tumour nuclei (497) does require the presence of rat liver nuclease-inhibitor, which is derived from a cytosol fraction. Heterologous (regenerating liver) cytosol has been reported to stimulate transcription of both unique and moderately re-iterated sequences in isolated rat liver (487) or hepatoma (488) nuclei, and cytosol is able to stimulate RNA synthesis in SV40-infected monkey cell nuclei (483). Promotion of transcriptional activity in Xenopus tissue-culture cell nuclei by Xenopus oocyte cytosol (489) is an effect confined to RNA polymerases I and III. A specific transcriptional effect has been demonstrated, however, in response to a factor present in the cytoplasm of heat-shocked Drosophila tissue-culture cells, which induces heat-shock chromosomal puffing, and transcription of heat-shock mRNAs, in isolated salivary gland nuclei not directly exposed to the heat-shock stimulus (500). Purified oestrogen/receptor complex is able to stimulate RNA synthesis in isolated

chick oviduct nuclei (501), and SV40 T-antigen acts similarly on nuclei from rat liver or quiescent hamster cells (502), although in neither of these cases have gene-specific transcriptional effects been demonstrated. Cytosol fractions have no detectable effect on at least one efficiently re-initiating system (derived from adenovirally infected HeLa cells (494). Attempts to overcome the generally poor level of endogenous initiation include the use of added bacterial RNA polymerase (503), which in Friend cell nuclei did not alter the proportion, only the absolute amount of viral transcripts, or the use of cells made permeable to small molecules by lysolecithin treatment (504), which so far has been developed only for studying transcription by RNA polymerase I in HeLa cells. Neither of these techniques has yet gained wide acceptance. Small nuclear RNA addition has been shown not to affect transcription in isolated rat liver nuclei (505).

The fidelity of transcription in isolated nuclei has been demonstrated for Friend cells, by the absence of anti-sense RNA, and the correct ratio of α : β globin mRNA in the de novo transcript (506), and for adenovirally transformed (481) or infected (482) cells, by the close correlation with the pattern of relative transcription which pertains in vivo, including transcript sizes, map positions, relative abundances, and accuracy of initiation (494). The transcriptional unit for chick ovalbumin in isolated oviduct nuclei, is the same as that observed in vivo by pulse-labelling, with the immediately 5' and 3'-adjacent sequences remaining untranscribed (507). One problem which has arisen in the interpretation of such experiments is that the use of mercurated nucleotides in the isolation of de novo transcripts may cause premature chain termination and processing defects (508), although this is disputed (509).

Isolated nuclei systems have been widely used to help dissect out the levels at which particular genes are regulated, in particular, to test hypotheses of transcriptional control.

For these studies, re-initiation is not strictly necessary, since the relative rates of specific chain extension in a pulse-labelled nuclear suspension will depend upon the frequency with which particular genes have been initiated in the population of nuclei prior to their isolation, provided a brief pulse label is used. In practice, however, control at the levels of transcriptional initiation, transcriptional elongation, and immediate post-transcriptional processing and/or turnover of the newly-synthesised transcript, are impossible to distinguish with present techniques. The most widely studied systems are those under steroid hormone control, such as the chick oviduct. Synthesis of both ovalbumin (485) and ovomucoid (397) RNA can be detected in oviduct nuclei only after the tissue has been stimulated with oestrogen, and not from non-target tissue or untreated oviduct nuclei, suggesting that the expression of these genes is controlled transcriptionally by the hormone. In the case of ovalbumin, the synthesis of mRNA and intron sequences is induced by the same factor, and follows similar kinetics following hormone re-administration, when the products of pulse-labelled isolated nuclei transcription are analysed, once again strongly suggesting transcription level regulation (485). However, it should be noted that the proportion of the in vitro transcript which is complementary to ovalbumin or ovomucoid DNA (397) is much less than that found in steady-state hnRNA in vivo, suggesting that post-transcriptional selection may also be operating. Comparisons of in vitro transcription and in vivo accumulation rates for ovalbumin and conalbumin mRNAs (484) suggests differential hormonal effects on the transcription and cytoplasmic stabilities of these messengers.

The isolated nuclei technique has also demonstrated transcriptional modulation of casein mRNA synthesis in mouse mammary gland, by glucocorticoids (395), of transferrin mRNA synthesis in chick liver in response to oestrogen or iron deficiency (510), of an abundant Dictyostelium mRNA during

development (493), and of globin mRNA production after Friend cell induction (506). The absence of $\alpha 2_u$ globulin mRNA in female rat liver nuclei is parallelled by its absence from the in vitro transcript from female rat liver nuclei (511), although its relative transcription rate in male liver nuclei is grossly insufficient to account alone for its high level on the polysomes, suggesting a major role for post-transcriptional selection in the determination of its abundance. The cell-cycle-dependent transcriptional control of histone synthesis in HeLa cells has been inferred from the failure to detect histone gene transcription in isolated G1 phase (as opposed to S phase) nuclei (480). Doubts persist, however, due to the inadequacy of uncloned hybridisation probes in this system.

2.4.4. RNA processing and transport studies using isolated nuclei

RNA processing and transport in isolated nuclei are far less well characterised than transcription, in part because there are few reliable ways of testing for in vivo equivalence, and also because of the multiplicity of processing steps. Only in the case of individual genes, for which purified hybridisation probes are available, can the mRNA maturation pathway be traced in any detail, in such a system. Elsewhere, there is a tendency to consider all nuclear post-transcriptional events as a single process, culminating in transport of a mature mRNA to the medium, which may be grossly misleading. Most attention has rightly been paid to the conditions under which accurate processing and transport take place, although the lack of appropriate assays has meant that in vivo equivalence has not been satisfactorily demonstrated for the bulk of genes in any system.

A number of systems which are capable of supporting extended transcription are also able to carry out several or all of the subsequent processing steps, often leading to the release of apparently fully matured mRNA into the medium. Such is the case for rat pituitary tumour nuclei, in the presence of ribonuclease-inhibitor (497), which package the *in vitro*-synthesised and processed mRNA into mRNP-like particles (512,513). Rat liver (514) and mouse myeloma nuclei (515) also transport fully processed mRNPs synthesised *in vitro*, in the case of the latter being able also to bind ribosomes. Drosophila embryo nuclei (492) will polyadenylate RNA synthesised *in vitro*, and transport of processed mRNAs synthesised *in vitro* in immobilised rat liver nuclei, subjected to continuous perfusion, has also been reported (516). Transcription and polyadenylation are temporally separable in mouse myeloma nuclei (517) and may also be distinguished by virtue of the dependence of polyadenylation (but not transcription) on the addition of a crude nuclear extract (517), although other investigators have reported polyadenylation to occur in this system independently of added factors (470). RNA transport from rat pituitary tumour nuclei is dependent largely upon polyadenylation, since the polyadenylation inhibitor cordecypin triphosphate blocks it (512). The *in vivo* equivalence of isolated nuclei systems in which transcription and processing occur successively is extremely difficult to test, since correct specificity may be reproduced at some but not all levels, whereas most assays can test only the overall specificity. In general, sequence-specific assays of *in vivo* equivalence have not been applied to this type of system.

Other investigators have concentrated on the development of systems which merely complete the processing and transport of transcripts synthesised *in vivo*. Such systems should ideally be inactive in *de novo* transcription, and this should be

verified, if they are to offer a specific environment in which to test strictly for post-transcriptional control. However, the proportion of the RNA released from isolated nuclei which is transcribed de novo is very small (as can be verified by comparing the recorded amount of release in such systems, with the amount of in vitro transcription, even in relatively efficient isolated nuclei systems), and can, in practice, be ignored. Furthermore, the incubation medium used in many systems contains factors, such as calcium ions, which are known to inhibit transcription (606).

The release from incubated nuclei, of messenger-like RNA, previously pulse-labelled in vivo, has been reported for rat liver (518,519), rat brain (520) and mouse tissue-culture cells (521). The fierce debate in the literature, as to which are the appropriate conditions for this, relating in particular to the need for and effects of cytosol and ATP, illustrate the essential problem of such systems, namely the absence of reliable and specific assays for physiological equivalence. Under many conditions used in the earlier studies nuclear integrity in general, and DNA retention in particular, were not controlled (or maintained),these being dependent, as shown subsequently, on the presence of spermidine (518) and calcium and manganese ions (522). A need for cytosol has been reported for liver nuclei by one group of investigators (518,519) but not by others (522, 523), although ribonuclease-inhibitor was included in one case (523). Further groups of investigators found that cytosol was necessary to prevent massive spurious RNA release (524), a result reported elsewhere for brain nuclei (520), or else reported that cytosol promoted the packaging of the RNA into dense RNP particles (525).

The requirement for ATP has not been found to be universal, even when similar incubation media have been employed. Under conditions in which ATP is required for mRNA transport from liver nuclei, transport from hepatoma nuclei, or from liver

nuclei derived from carcinogen-treated rats, is only partially dependent on ATP (526-528). Carcinogenesis is not, as was originally proposed, the cause of this discrepancy, since nuclei from rat brain, carcinogen-treated rat-brain, and glioma all show equivalent ATP-dependence in the same medium (520). These discrepancies have not been satisfactorily explained. A requirement for ATP hydrolysis in the energisation of RNA transport, reported by a number of groups (518,519,523, 529) is disputed by others (524), who found that non-hydrolysable analogues of ATP, or even a mixture of AMP and pyrophosphate could substitute, on the basis of which they advanced a 'conformational hypothesis' of the role of ATP in promoting RNA transport from isolated nuclei. In this latter case, the integrity of the nuclei may have been impaired by the absence of calcium and manganese ions from the medium, leaving the significance of the finding in some doubt.

Transport rates have been found to correlate with the total nucleotide energy charge in the medium (523), and inhibitor and kinetic studies have indicated a reciprocity between transport activity and nuclear membrane pore NTPase activity, both in rat liver (529) and in transformed mouse fibroblasts (530,531), although the existence of this NTPase is disputed elsewhere (524). Cyclic nucleotides have been reported as stimulating (532) or inhibiting (533) pulse-labelled RNA transport from liver nuclei under identical conditions, and the stimulatory activity could be achieved by prior exposure of the tissue to cAMP <u>in vivo</u> (532). Hydrocortisone, which also inhibits transport (533), can be mimicked by the use of cytosol from hormone-treated rats in the incubation.

Inhibitor studies have indicated that mRNA transport in these systems is dependent upon polyadenylation (by the use of cordecypin) (521,534), and upon processes mediated by RNA-RNA duplex formation (by the use of proflavine) (535). A role for the nuclear membrane phospholipid in RNA transport, both from transformed mouse cell (531) and tetrahymena (536) nuclei,

has been inferred from the temperature-profile, which resembles that of membrane fluidity. Incorporation of mRNAs into mRNP-like particles prior to release is also widely reported (519,537) and ribosomal RNA release which is promoted under similar conditions (538, 539, 524) occurs in the form of mature ribosomal subunits. The in vivo equivalence of such systems for the processing and transport of ribosomal RNA has been inferred from the kinetics of transfer of label to smaller RNA species, during the course of the incubation, together with the observation that the 18S : 28S ratio of the released RNA is characteristic of their nuclear precursor pool sizes (538). Since such properties as RNA size, poly(A)-content, or association with proteins are very crude tests for in vivo-like specificity, various attempts have been made, using hybridisation methodology, to characterise the informational content of the released RNA. Early hybridisation-assays, based on competition hybridisation with filter-bound genomic DNA, were used to test where in vivo sequence selectivity was maintained in vitro. In the presence of homologous cytosol, rat liver nuclei were found to transport only those sequences found normally in the cytoplasm (537), although some release of nucleus-confined transcripts has been detected by other investigators using this system (G.A. Clawson, personal commun.). Since the assay applies only to repetitive sequence transcripts, it may not apply to the bulk of mRNA sequences, whose nucleus-restriction may or may not be maintained in vitro. The operation of cytosol-induced specificity was similarly inferred on the basis of the transport of otherwise nucleus-restricted sequences in the presence of cytosol from hepatoma (537), regenerating liver (537), or liver from carcinogen-treated rats (541), in all of which cell-types, a similarly altered pattern of repeat sequence transcript expression can be detected in vivo (see Section 2.3.5). The possibility that the phenomenon was a non-specific effect of cytosols from these cell-types, causing a spurious release of nuclear sequences, was not thoroughly investigated. Release of nucleus-restricted repeat sequence transcripts in response to

serum fractions from tumour-bearing rats has also been reported (541,542). This apparent maintenance, in the *in vitro* system, of the specificity of repeat sequence transport inferred *in vivo* prompts the question whether the qualitative or quantitative post-transcriptional controls deduced to be operating on the basis of *in vivo* studies of mRNA populations may also be reproduced in such an *in vitro* system. This is a question which will be addressed directly in the experimental work to be reported here.

Hybridisation studies using probes for the processing of specific mRNA sequences in isolated nuclei are still in the earliest stages of development. Correct *in vitro* splicing has been reported for SV40 (540) or adenovirus (543,544) mRNAs in infected cell nuclei. The latter either depends upon (543) or else it is greatly stimulated (544) by the addition of cytosol extracts. Since cytosol is active only if cells are lysed hypotonically, it has been suggested that its action is to restore some factor or factors which leach out of nuclei during cell lysis (544). Nuclear extract can also restore splicing activity to hypotonically lysed SV40-infected cell nuclei (540).

2.4.5. Transcription, processing and translation in soluble cell-free systems

The interpretation of *in vitro* chromatin transcription experiments, using exogenous bacterial RNA polymerase, is seriously in doubt, due to the copying of endogenous RNA by the enzyme (545,546), the contamination of *de novo* transcripts with endogenous RNA (547) and the generally poor specificity of the enzyme when these artifacts have been properly controlled (548,549). It is therefore proposed not to discuss these experiments in detail. Chromatin transcription based on endogenous RNA polymerase II activity is largely the extension of pre-initiated chains (550), and re-initiation is inefficient and short-lived even when supplementary exogenous polymerase

is provided (551). The reproduction of *in vivo*-like specificity in such systems has not been demonstrated, nor in endogenous chain elongation systems based on nucleoplasm extracts (552).

Soluble transcription systems based on nuclear extracts or purified RNA polymerases, often supplemented with cytoplasm preparations, have been shown to carry out accurate *de novo* transcription on exogenous, naked DNA templates. Coupled with a surrogate genetics approach, these systems have contributed considerable information on the mechanisms controlling transcription, albeit of a preliminary nature. For example, accurate RNA polymerase III transcription of Xenopus 5S DNA in a Xenopus oocyte nuclear extract (553) is dependent on a control region in the centre of the structural gene itself (554,555) which binds a specific transcription regulatory factor (556). This factor becomes associated with 5S RNA in developing oocytes *in vivo*, providing a mechanism whereby the accumulation of a gene product is able to switch off its own synthesis (557). Xenopus oocyte nuclear extracts also contain (separable) activities for the transcription and processing of tRNA genes (558). Correct transcriptional initiation here depends on sequences in both the 3' and 5' parts of the structural gene (559). A polymerase III-system which faithfully transcribes adenovirus 'VA' genes also gives rise to specific short transcripts from the human β-like globin locus, derived from repetitive sequences bordering the structural genes (560) which are suggested as possible elements in globin gene regulation.

Accurate RNA polymerase II transcription *in vitro* has been reported for cloned chick ovalbumin (561) and conalbumin (561) genes, adenovirus (562,563) and mouse β-globin (564). The relative efficiencies with which these DNA templates are transcribed *in vitro* is not simply related to their relative transcription rates *in vivo* (561,564). This is shown most clearly for ovalbumin and conalbumin, genes

which are transcribed at similar rates *in vivo*, where the former is a very poor template for cell-free transcription (561), even though there is no direct competition between templates for a limiting factor. It may therefore be inferred that transcription rates *in vivo* depend upon elements other than the DNA sequences of the transcriptional unit itself. These may be DNA sequences distant from the structural gene, gross chromosomal structure, or factors inactive in, or absent from, the transcription systems which are used. Accurate transcriptional initiation on adenoviral DNA, generating long transcripts, has now been reported in a crude whole-cell extract of HeLa cells (562), which may help elucidate this point.

Surrogate genetics has shown that the 5' flanking regions of polymerase II-transcribed genes contain sequences which are essential for accurate and efficient transcriptional initiation *in vitro*, including the TATAAA-box sequence at approximately position -25 with respect to the gene, and also sequences closer to the start site (561). Such an approach may also prove useful in studying the sequences involved in specific gene regulation, for example by steroid hormones, provided the sequences required are sufficiently close to the gene to be included on the same template molecule.

Various post-transcriptional processing reactions have been successfully carried out in soluble cell-free systems, generally derived from nuclear extracts, although the reconstitution of *in vivo* specificity (if any) is generally unproven. Polyadenylation of primer RNA derived from homologous nuclei, by partially purified rat liver poly(A)-polymerase (565), is sensitive to inhibition by 2' or 3' dATP, and is modulated by *in vitro* phosphorylation which enhances the initiation rate of poly(A) synthesis (566). Poly(A)-polymerase from hepatoma is more highly phosphorylated (567), which may imply control of gene activity between liver and hepatoma by differential rates of polyadenylation (coupled, for example, to different

affinities of different transcripts for the enzyme). The poly(A) polymerase is antigenically related to (and possibly identical with) the poly(A)-binding protein found associated with nuclear or cytoplasmic mRNPs (568) from liver.

Capping (transguanylylation) and cap-methylation of RNA molecules newly synthesised by RNA polymerase II, have been reported in soluble extracts of HeLa (569,570) and rat liver (571) nuclei. Only polyphosphate 5' termini can act as substrates (569-571), and capping and methylation activities are physically separable (571,572).

Soluble splicing systems have so far been reported only for yeast tRNAs (573,574). The yeast enzyme(s) catalyse a two-step excision-ligation reaction (575) which generates unusual 3' phosphate and 5' hydroxyl-terminated intermediates (576). Accurate splicing of yeast tRNA precursors is also supported in a Xenopus oocyte nuclear extract (574).

Efficient translation of exogenous mRNAs has been demonstrated in wheat-germ extracts (577) and nuclease-treated reticulocyte lysates (578). No convincing evidence exists for translational modulation in these systems by mRNA-specific factors. A number of treatments, such as heme deprivation (579), exposure to double-stranded RNA (580) or oxidised glutathione (581), cause a radical decrease in the generalised rate of polypeptide chain initiation. Such mechanisms have been proposed by Lodish (101), as ways of altering the relative translation rates of high and low affinity mRNAs. Studies of these regulatory phenomena have led to a detailed understanding of the mechanisms responsible, although they are complex, and this understanding remains incomplete. The rate-limiting step of chain initiation appears to be modulated by the reversible phosphorylation of factors involved in its catalysis (581,582,583). This may provide a model for regulatory mechanisms operating at other levels of gene expression.

Modulation of mRNA stabilities *in vitro* has been demonstrated in response to interferon treatment. An endonuclease is induced in interferon-treated cells, in the presence of double-stranded RNA (584), which is active in soluble extracts, and which degrades different mRNAs at different rates when added to reticulocyte lysates (585).

2.5. Rationale for choice of experimental system

The particular aim of this project was the development of an *in vitro* system in which the *in vivo* specificity of post-transcriptional selection was maintained, and which might be useful for investigating the role of cytoplasmic factors in the modulation of gene expression in nuclei of different cell-types. Clearly, a number of biological systems could have been chosen for this study - the aforegoing discussion illustrates the enormous range of organisms and cell-types whose investigation has contributed to our understanding of selective gene controls. The choice of system was therefore based on consideration of the type of assay which needed to be developed, to test critically for *in vivo* equivalence of processing and transport *in vitro*. Nevertheless, the choice of system is to some degree arbitrary, and unavoidably biasses the type of selective mechanism which can be studied.

As indicated previously, the development of a cell-free system for studying post-transcriptional controls depends on the provision of assays for sequences which are regulated at this level, either qualitatively or quantitatively. These assays, which must be based upon nucleic acid hybridisation methodology as the only way of ensuring sufficient specificity, are of two types: firstly, for sets of sequences under a defined post-transcriptional modulation within a single cell-type; secondly, for sets of sequences whose polysomal concentrations are modulated between cell-types at a post-transcriptional level. Since any attempt to reproduce the latter phenomenon *in vitro* is likely to involve the use of heterologous cell

extracts, it is desirable that the biological system chosen should involve cell-types which are as closely related as possible, but where the differences are sufficiently clear-cut to provide useful hybridisation assays. The first type of assay presents no particular problems, in that large qualitative differences between the sequences present in nuclear and polysomal RNA are evident (see Table 1) for a wide range of cell-types. Quantitative differences between nuclear and polysomal mRNA populations are also widespread, and together these should provide a sufficient battery of assays for judging whether the processing and transport of RNA sequences in vitro follows a specificity which resembles that seen in vivo, using any one of a number of sources of biological material. The second type of assay is far more problematic, however. As already discussed (Section 2.3) the sequence sets present in nuclear and polysomal RNA of even quite different cell-types are often qualitatively very similar, and when closely related cell-types are considered, such as those in different states of differentiation or proliferative capacity, selective gene expression appears to be almost entirely at the level of relative mRNA abundances. These may be, moreover, very subtle, even when phenotypic differences are dramatic. The type of clear-cut qualitative post-transcriptional modulation between related cell-types, of the kind observed for repeat sequence transcripts (388, 389), is almost certainly not applicable to mRNA sequences.

The biological system which was, in fact, selected, that of the rat liver in comparison with a line of cultured rat hepatoma cells (HTC), was considered particularly advantageous from the viewpoint of the possible operation of post-transcriptional controls which might give rise to sufficiently clear-cut hybridisation assays for probing an in vitro system. Although deriving from the same cell lineage, these two cell-types differ phenotypically along a number of parameters which might be expected to be reflected in differences at the mRNA level. The use of the liver and hepatoma also

provided a link with earlier studies, in which qualitative differences in repeat sequence expression could be reproduced in vitro (537). Furthermore, the observed phenomenology of repeat sequence expression in these cells offered some hope that mRNA expression might also be regulated post-transcriptionally, whether qualitatively or quantitatively.

The phenotypic difference between liver cells and HTC cells may be summarised as follows. Firstly, hepatocytes are highly specialised, in the sense of exhibiting differentiation-linked markers, such as albumin synthesis, not detected in HTC cells (586,612), or constitutively expressing specialised enzymes, such as aminotransferases, whose expression in HTC cells is at a low level, but may be induced by glucocorticoid treatment (587). Secondly, whereas HTC cells are actively growing, and proceeding through the cell cycle, liver is composed of quiescent cells in GO phase. Thirdly, liver cells are subject to normal growth controls, are contact inhibited in primary (590) cell culture, will not grow in soft agar, and are non-tumorigenic, whereas HTC cells are neoplastic, and are not subject to any of these restrictions. Fourthly, liver cells are adapted to functioning as part of a multicellular array in vivo, whereas HTC cells are adapted to independent growth in suspension culture. These differences are sufficiently multifarious to hold out some hope that at least a clear-cut quantitative difference in their mRNA populations could be detected by hybridisation methodology, and that some, if not all of their differences, could be traced to altered selectivity at the post-transcriptional level. The potential comparison with the regenerating liver is also of some interest.

Other aspects of the liver/hepatoma comparison make it particularly suitable for the type of study envisaged here. Firstly, the properties of liver and hepatoma nuclei incubated in vitro, in particular their transcriptional and RNA processing

and transporting activities, are well characterised, probably more so than any other cell-types (see Sections 2.4.3 and 2.4.4). This means that the development of an *in vitro* system is far less susceptible to simple artifacts, arising from non-specific leakage of nuclear components, or the action of non-specific nucleases or proteases, since the conditions under which they are likely to be minimised have already largely been determined. Secondly, rat liver cytosol is known to be low in nuclease activity, since it is the source of an active ribonuclease inhibitor (588), which enhances the prospect of recovering intact mRNA-like sequences transported *in vitro*. Thirdly, large quantities of material, both active nuclei and RNA, may be prepared under well controlled and easily reproduced conditions. Finally, the well characterised physiology of the liver makes possible the relationship of experimental findings to phenotype, for example, in relation to the liver's secretory functions.

3. MATERIALS AND METHODS

Materials used generally were as follows. Except where indicated, all reagents were AnalaR Grade British Drug Houses. All radioisotopes were supplied by The Radiochemical Centre, Amersham. Tissue-culture media, including supplements and serum, were supplied by Flow Laboratories or by Gibco-Europe, except where stated. Ultra-pure (density-gradient grade) sucrose was obtained from Schwarz-Mann. Tris was supplied by Sigma.

All glassware for handling RNA or DNA was siliconised by treatment with Repelcote (Hopkins and Williams), and baking for at least 2 hours at 200°. Glassware was sterilised by boiling for 1 hour in a 0.05 percent (v/v) solution of diethyl pyrocarbonate (Sigma). Solutions were sterilised (except where stated) by autoclaving, preceded by the addition of diethyl pyrocarbonate to 0.01 percent where the solution was to be used for the handling of RNA. Solutions containing sucrose were sterilised by treatment with diethyl pyrocarbonate (0.01 percent) which was left to hydrolyse overnight at room temperature.

Formamide (Fluka), DMSO and acrylamide solutions were de-ionised for several hours by stirring with AG501-X8(D) ion-exchange resin beads (Bio-Rad). Phosphate buffer was prepared by mixing suitable quantities of Sodium dihydrogen orthophosphate and di-Sodium hydrogen orthophosphate to obtain a pH of 6.8. All radioactivity determinations were by scintillation counting, using MI-96 (Packard) scintillant, except where stated.

Where detailed methods of individual experiments are not cited here, they are to be found in the appropriate figure legends, or in the text.

3.1. Animals and cell-culture

Livers were excised from slaughtered male Wistar rats, aged 3-6 months immediately following perfusion via the portal vein with STKM buffer (0.25 M sucrose, 50 mM Tris/HCl, 3 mM $MgCl_2$, 25 mM KCl, pH 7.4), warmed to 37°. Excised livers were transferred to the same buffer on ice, chopped and washed several times.

Rat hepatoma tissue-culture (HTC) cells, checked previously in this laboratory for glucocorticoid inducibility of tyrosine aminotransferase, a hepatic marker enzyme, were grown and passaged in monolayer culture in plastic bottles (Falcon). The medium used was Ham's SF12/20 percent fetal calf serum, buffered with 1 percent w/v Sodium bicarbonate/5 percent $_v/_v$ CO_2 (gas phase), containing 50 units/ml 'Crystapen' benzylpenicillin (Glaxo). Cells were passaged by detachment in phosphate-buffered isotonic saline (Oxoid)/1 mM EDTA, followed by re-innoculation at a cell density of 1-2 x 10^4 cells/ml. Cells generally reached confluence in 6-7 days. Bulk cultures were grown in continuously stirred suspension, either in 1 litre glass blood-bottles or 5 litre aspirators, in the same medium. These were seeded at 10^5 cells/ml and harvested at about 10^6 cells/ml, during exponential growth. After a lag phase, cells were found to grow exponentially with a generation time of about 22 hours. In some batches of serum, cell clumping occurred at high cell densities, but where this was serious, cultures were discarded.

Suspension-culture cells were harvested by centrifugation in 1 litre plastic bottles (MSE) for 20 minutes at 1500 rev/min (700 g_{max}) at 4°. Cell pellets were washed twice with ice-cold phosphate-buffered saline.

A fresh monolayer culture was started every 3-4 months, by thawing at 37° an ampoule of cells stored frozen in liquid N_2 in their usual medium plus 10 percent DMSO, recovering cells by centrifugation, washing with warmed medium, and innoculation at high density (10^6 cells/ml).

3.2. Cell homogenisation and fractionation

3.2.1. Preparation of liver polysomes

Polysomes were prepared from rat liver essentially by the method of Wilkes et al (145). Chopped livers were homogenised on ice in STKM buffer (approximately 5 ml per g wet weight), in a teflon/glass homogeniser (6 strokes, hand-held, loosely fitting pestle). Nuclei and debris were pelleted by centrifugation at 2000 rev/min (1000 g_{max}) for 15 minutes at 4°, after which the supernatant was recovered and re-centrifuged at 8000 rev/min (10,000 g_{max}) for 10 minutes to remove mitochondria and other organelles. The post-mitochondrial supernatant was treated for 15 minutes on ice with 40 units/ml α-amylase (Worthington) in order to remove glycogen, the presence of which was observed, in accordance with the findings of Gamulin et al (589) to hinder resuspension of the final polysomal pellets, so reducing yields, and also to cause problems in oligo(dT)-cellulose chromatography. Treatment with α-amylase was found to cause no detectable changes to the polysome profile (see Fig. 1) nor to the proportion of material retained by oligo(dT)-cellulose, hence the enzyme was judged to be RNase-free. Millipore-filtered Triton X-100 (Sigma) was then added to a final concentration of 1 percent v/v, and the suspension was layered over 5 ml pads of STKM buffer containing 2 M sucrose, in 25 ml polycarbonate centrifuge tubes. Polysomes were pelleted by centrifugation at 55,000 rev/min (350,000 g_{max}) for 3.5 hours at 4°, in an MSE 8 x 25 Titanium fixed-angle rotor. The supernatants were carefully removed, so as to avoid contaminating the pellets,

and the tubes wiped with sterile gauze. Polysome pellets were either stored at -20°, or re-suspended immediately for rate-zonal centrifugation (Section 3.2.5).

3.2.2. Preparation of liver nuclei by sucrose/citric acid method

The method was adapted from Getz et al (167). All manipulations were carried out at 4°. Chopped excised liver, or crude nuclear pellets from polysome preparations, were resuspended in 5 percent citric acid (approximately 5 ml per g wet weight of tissue) and homogenised in a teflon/glass homogeniser (motor-driven, tight fitting pestle, 15 strokes). After filtration through four layers of sterile gauze, the homogenate was centrifuged for 10 minutes at 1000 rev/min (250 g_{max}). Pellets were resuspended in a similar volume of 1.5 percent citric acid, 0.25 M sucrose, re-homogenised (5 strokes), and layered over an equal volume of 1.5 percent citric acid, 0.88 M sucrose. After centrifugation for 5 minutes at 2000 rev/min (1000 g_{max}) nuclei pellets were resuspended in 1.5 percent citric acid, 0.25 M sucrose, re-homogenised (3 strokes) and recovered once more by pelleting through 1.5 percent citric acid, 0.88 M sucrose. Microscopic examination (under phase-contrast) showed that by this stage nuclei were free of contamination with cytoplasmic 'tags', and they were therefore washed finally with STKM buffer, and recovered by centrifugation at 1000 rev/min (250 g_{max}) for 5 minutes. Nuclear pellets were stored at -20°, or subjected immediately to procedures for extraction of RNA or DNA (Sections 3.3 and 3.4.2). Nuclei prepared by this method were considerably free of cytoplasmic contamination, as evidenced by the kinetics of nuclear RNA-driven hybridisation of polysomal cDNAs (Section 4.3).

3.2.3. Preparation of free and membrane-bound liver polysomal fractions

Livers were homogenised in MHB (0.25 M sucrose, 75 mM KCl, 5 mM $MgCl_2$, 50 mM Tris/HCl, 3 mM DTT) under similar conditions to those used for total polysome preparations. After nuclei and debris had been removed by centrifugation at 2000 rev/min (1000 g_{max}) for 15 minutes at 4^o, the supernatant was made 250 mM in KCl, and mitochondria were pelleted at 8000 rev/min (10,000 g_{max}) for 10 minutes. After α-amylase treatment (see Section 3.2.1), the supernatant was layered in 25 ml polycarbonate centrifuge tubes, over a 3 ml pad of MHB containing 1.8 M sucrose and 250 mM KCl, which was layered over a 5 ml pad of MHB containing 2.0 M sucrose and 250 mM KCl. The samples were centrifuged at 4^oC for 4.5 hours at 55,000 rev/min (350,000 g_{max}). Membrane-bound polysomes were removed from the band of material sedimenting at the interface between the two dense sucrose pads: an aliquot of this material was removed for protein analysis by gel electrophoresis (see Section 3.10) and the remainder was centrifuged for 9 hours at 40,000 rev/min (170,000 g_{max}) in order to pellet polysomes. Free polysomes were recovered as pellets from the initial step-gradient.

3.2.4. Preparation of nuclei and polysomes from HTC cells

The methods previously used for mouse Friend cells (167) were employed. HTC cell pellets were resuspended in TNM buffer (140 mM NaCl, 10 mM Tris/HCl, 1.5 mM $MgCl_2$, pH 7.6) on ice (4-7 x 10^7 cells/ml), and NP40 was added to a final concentration of 0.5 percent (w/v). Cells were homogenised after 10 minutes in a teflon/glass homogeniser (6 strokes, hand-held, loosely fitting pestle). TNM buffer containing 2 M sucrose was added, such that the final concentration of sucrose was 0.25 M, and nuclei and debris were spun out at 2000 rev/min (1000 g_{max}) for 15 minutes at 4^o. Nuclei were purified from the pellets by the sucrose/citric acid method, exactly as described for liver (Section 3.2.2). The supernatant was centrifuged at 8000 rev/min (10,000 g_{max})

for 10 minutes, and Triton X-100 added to the post-mitochondrial supernatant to a final concentration of 0.5 percent (v/v). Polysomes were pelleted as already described for liver (Section 3.2.1) except that the dense sucrose pad consisted of 2 M sucrose in TNM, not STKM buffer.

3.2.5. Rate-zonal sedimentation of polysomes

Liver and HTC cell polysome pellets were resuspended in a minimal volume, usually 1/2 - 1 ml of TNM buffer, on ice, by gentle swirling for up to 1 hour. After centrifugation at 1000 rev/min (250 g_{max}) for 5 minutes to remove aggregated material, polysomal suspensions were layered carefully over 32 ml 15-40 percent sucrose gradients in TNM buffer (<40 OD_{260} units per gradient) and centrifuged at 4^{o} for 2 hours at 25000 rev/min (110,000 g_{max}) in an IEC SB-110 swing-out rotor. Gradients were analysed and fractionated by upward displacement using fluorochemical FC43 (supplied by 3M (UK) Ltd.), passed through a u.v. absorbance scanner, and collected manually on ice. Material sedimenting at > 100S was pooled and precipitated overnight at -20^{o} by adding 1/9 volume of 5 M NaCl, and 2 volumes of ethanol. Typical polysome profiles are shown in Fig. 1, where it can be seen that liver and HTC cell polysomal preparations are similarly undegraded. The α-amylase treatment of liver polysomes results in no detectable degradation. In the few cases where polysome profiles were clearly degraded, the preparation was abandoned at this stage. This procedure is essentially the same as that used by Young et al. (222), and removes almost all contaminating hnRNPs from the polysome preparation. All polysomal RNA used in the experiments was purified by this method, except for free and membrane-bound polysomal fractions from liver.

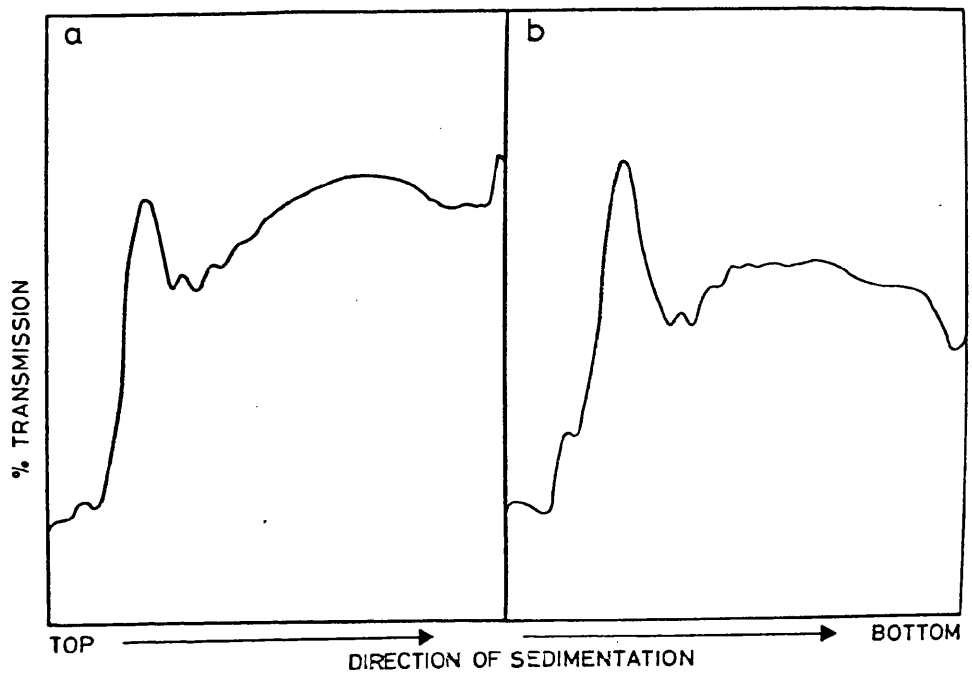


FIGURE 1. Polysome profiles from (a) rat liver and (b) HTC cells analysed on 15-40 percent TNM-sucrose gradients. Percent transmission of uv light, shown on the y-axis, is an arbitrary scale obtained empirically to produce an appropriate range of deflection for sample analysis. Centrifugation conditions are as stated in the text. The major peak, sedimenting near the top of each gradient, is estimated at 75-80S by computer analysis. (222), hence is identifiable as monoribosomes.

3.3. Preparation of rat liver DNA

Rat liver nuclei, prepared by the sucrose/citric acid method, were resuspended in 100 mM Tris/HCl pH 8.0 (10 ml per g wet weight starting tissue) and sodium dodecyl sarcosinate (Sigma) and proteinase K (Boehringer) were added to concentrations of 4 percent (w/v) and 50 µg/ml respectively. The mixture was hand homogenised gently and incubated for 45 minutes at 37°. After chilling 2-3 minutes on ice, DNA was ethanol precipitated at -20° overnight, by the addition of 1/9 volume 5 M NaCl and 2 volumes ethanol, spooled off and redissolved in 50 mM NaCl, 10 mM HEPES (Sigma) pH 7.0 by gentle hand homogenisation. Pre-boiled RNase A (Sigma) was added to a final concentration of 50 µg/ml, and following a 2 hour incubation at 37°, Sodium dodecyl sarcosinate and proteinase K were added to 4 percent (w/v) and 10 µg/ml respectively and incubation was continued for a further 30 minutes at 37°. The mixture was phenol/chloroform extracted by the addition of an equal volume of phenol/chloroform/iso-amyl alcohol (100:100:1) to which had been added 0.01 percent (w/w) 8-hydroxyquinoline as an antoxidant and which had been saturated with NETS buffer (100 mM NaCl, 1 mM EDTA, 10 mM Tris/HCl, 0.5 percent SDS (Sigma), pH 7.4). After vigorous agitation, the phases were separated by centrifugation at 10,000 rev/min (16,000 g_{max}) and the whole procedure was repeated until the interface was clear. DNA was recovered from the aqueous phase by ethanol precipitation, and the precipitate, isolated by centrifugation at 10,000 rev/min (16,000 g_{max}) for 15 minutes, was redissolved in 100 mM NaCl to a concentration of 1-2 mg/ml and sonicated on ice using a Branson sonifier with 64 mm tip, in five 20 second bursts separated by 30 second cooling pauses. The sonicated DNA was lysophilised and desalted in a minimal volume by passage through a Sephadex G-50 (Pharmacia) column, with a 2 cm pad of dowex chelating resin (Sigma) at the top.

Sonicated DNA was stored in water at -20°. The procedure described has been used previously by Paul et al. (590).

3.4. Preparation of liver and HTC cell poly(A)$^+$ polysomal and nuclear RNA

3.4.1. CsCl extraction of polysomal RNA

Polysomal RNA was prepared from ethanol precipitates of density-gradient-purified polysomes, recovered by centrifugation, or from pellets of free and membrane-bound liver polysomes, by CsCl extraction (222). Pellets were resuspended by gentle swirling at room temperature in 150 mM NaCl, 50 mM HEPES, 10 mM EDTA, 1 percent SDS, pH 7.5 to a concentration of 1-2 mg/ml (final OD_{260} of 10-20), and dowexed CsCl was added (1.4 g per ml of solution) and dissolved by vortexing. The protein precipitate was packed into a hard surface pellicle by centrifugation at 10,000 rev/min (16,000 g_{max}) for 25 minutes at 25°, and the RNA-containing solution was removed by means of a syringe fitted with a long canula, millipore filtered, and ethanol precipitated overnight at -20° by the addition of 3 volumes of distilled water and 8 volumes of ethanol RNA purity at this stage was checked spectrophotometrically.

3.4.2. Isolation of nuclear RNA

Nuclear RNA was extracted using either the guanidinium thiocyanate method, modified from that of Chirgwin et al (591), or by successive digestion with Proteinase K, DNase I and phenol/chloroform extraction (167). The guanidinium thiocyanate method was used to prepare all samples for use in single-copy DNA saturation hybridisation experiments. Nuclei (approximately 10^8 per ml) were extensively homogenised on ice, (teflon/glass homogeniser, tight-fitting pestle, motor driven) in millipore-filtered 5 M guanidinium thiocyanate

(Merck), 50 mM Tris/HCl, 50 mM EDTA, 5 percent 2-mercapto-ethanol (Sigma) pH 7.0. The homogenate was made 2 percent in Sarcosyl NL97 (Ciba-Geigy) after heating for 2 minutes at 50°, and layered over 1.2 ml pads of millipore-filtered 5.7 M CsCl, 50 mM EDTA, pH 7.0 in 4 ml polypropylene centrifuge tubes. After spinning for 20 hours at 35,000 rev/min (150,000 g_{max}) in an IEC SB-405 swing-out rotor at 15° the supernatants were carefully decanted and RNA pellets were redissolved in distilled water, and ethanol precipitated (see Section 3.3).

Alternatively, nuclei were resuspended in 100 mM Tris/HCl pH 8.0 and protease K digested as described previously (Section 3.3). Following this, the ethanol-precipitated nucleic acids were divided into 2 fractions, a DNA-rich fraction which was spooled off, and an RNA-rich fraction recovered by centrifugation. Each fraction was washed in 70 percent ethanol and dried thoroughly by vacuum dessication. The DNA-rich fraction was dissolved in 100 mM HEPES, 10 mM Magnesium acetate, 25 mM NaCl, 2 mM $CaCl_2$, pH 7.0 by gentle hand homogenisation, and DNase I (Worthington) added to a final concentration of 50 µg/ml. Following incubation at 37° for 12 minutes, the RNA-rich fraction dissolved in the same buffer was added for a further one minute of incubation, after which the solution was made 0.5 percent in SDS and 0.5 M in NaCl, phenol/chloroform extracted repeatedly until the interface was clear (see Section 3.3) and ethanol precipitated. Oligo-deoxyribonucleotides were removed by passing the ethanol-precipitate, redissolved in NETS buffer, over a Sephadex G-100 gel filtration column, incorporating a pad of dowex chelating resin. Excluded material was pooled and ethanol precipitated, and its purity checked spectro-photometrically.

3.4.3. Oligo(dT)-cellulose chromatograph

Poly(A)$^+$ polysomal and nuclear RNAs were isolated by two cycles of oligo(dT)-cellulose chromatography, as described by Birnie et al (166), modified from Aviv and Leder (196). Pellets of ethanol-precipitated RNA were washed in 70 percent ethanol, dried by vacuum dessication, and redissolved in 0.5 M NaCl, 10 mM Tris/HCl, 1 mM EDTA, 0.1 percent Sodium dodecyl sarcosinate, pH 7.5 at < 100 µg/ml in RNA ($OD_{260} < 2.5$). The sample was applied to an appropriately sized column of oligo(dT) cellulose (Type 3, Collaborative Research, binding capacity approximately 500 µg/ml poly(A)$^+$ RNA per ml of column) pre-equilibrated with the same buffer. Poly(A)$^+$ RNA was eluted in sterile water, lyophilised and desalted on a Sephadex-G50/dowex column. RNA samples were stored at -20^o for no more than 3 months. In general, poly(A)$^+$ RNA yields were 1.5 - 2 percent of the total for polysomal RNA, 8-10 percent for guanidinium thiocyanate-prepared nuclear RNA, and 2-4 percent for DNase I-treated nuclear RNA. Wherever yields were substantially lower than this, samples were discarded, except in the case of membrane-bound liver polysomal RNA (see Section 4.4.5). Purity was checked spectrophotometrically, and OD_{260}/OD_{260} ratios were always very close to 2.0. Samples of poly(A)$^+$ nuclear and polysomal RNAs from 16-hour regenerating rat liver (145), prepared using the same methods, were kindly donated by Dr Peter Wilkes for use in the experiment described in Section 4.5.3.

3.5. Isolation and labelling of single-copy rat DNA

3.5.1. Isolation of single-copy DNA by partial re-association

Sonicated rat DNA was lyophilised and dissolved in HB/F (0.5 M NaCl, 0.5 mM EDTA, 50 mM HEPES, pH 6.8, 50 percent formamide, dowex-treated and diethyl pyrocarbonate-sterilised prior to addition of the formamide) to a concentration of

1-2 mg/ml. After heat denaturation for 5 minutes at 70°, the sample was re-associated to $C_o t = 250$ moles.s.l^{-1} at 43° ($C_o t = 1$ being reached in 1 hour at a DNA concentration of 83 µg/ml (124)), flushed into a large excess of 30 mM phosphate buffer, 150 mM NaCl, and passed at 60° over a column of hydroxy-apatite (Bio-Gel HTP, Bio Rad), previously de-fined and boiled in 1.0 M phosphate buffer, and pre-equilibrated at 60° with 30 mM phosphate buffer. The column was washed several times with 30 mM phosphate buffer, after which single-stranded DNA was eluted with 140 mM phosphate buffer at 60°. Double-stranded DNA was eluted at 60° with 0.4 M phosphate buffer. These conditions had been determined as optimal, for the use of the given batch of HAP, for fractionating single and double-stranded DNAs. Single-stranded DNA was lyophilised, desalted on a Sephadex G-50 column incorporating a pad of the Na^+ form of the ion-exchange resin AG 50W-X8 (Bio Rad) and lyophilised once again. The whole procedure was repeated. After the second cycle, single-stranded DNA, consisting almost exclusively of single-copy sequences (see Section 4.1.1) was lyophilised, redissolved in HB/F to a concentration of 20 mg/ml, denatured in a sealed siliconised sterile glass capillary at 70°, and reassociated at 43° to $C_o t = 25{,}000$. Reannealed single-copy DNA was flushed out with 150 mM NaCl and stored at -20°at a concentration of 0.5 mg/ml.

3.5.2. Labelling of single-copy DNA by 'gap'-translation

Reannealed single-copy DNA was radio-labelled by a variant of the nick-translation reaction, in which single-stranded tails and gaps are made double-stranded by E.coli DNA polymerase I (591). The reaction mix contained ^{3}H-dCTP (19 Ci/mmol) or ^{32}P-dCTP (400 Ci/mmol) at 40 µM, dATP, dGTP and dTTP each at 100 µM, reannealed single copy DNA at 20 µg/ml, E.coli DNA polymerase I (Boehringer) at 50 units/ml, 50 mM Tris/HCl (pH 7.9), 10 mM $MgCl_2$, 1 mM DTT and 50 µg/ml

BSA (Sigma). After incubation at 12° for 20 hours, EDTA, E.coli DNA carrier (Sigma), SDS and NaCl were added to 10 mM, 20 µg/ml, 0.5 percent and 100 mM respectively and the solution was phenol/chloroform extracted (see Section 3.3). The aqueous phase was made 120 mM in phosphate buffer by considerable dilution, denatured by boiling for 10 minutes, and cooled to 60° for 30 seconds. Hairpin-like structures were removed by dilution with 3 volumes of 200 mM NaCl at 60°, and passage over a HAP column as already described (Section 3.5.1) care being taken to ensure that all un-incorporated radio-isotope was eluted by washing with 30 mM phosphate buffer, prior to recovery of the single-stranded and double-stranded DNA fractions. The counts eluted in these fractions were used to compute specific radioactivity ($2-5 \times 10^6$ cpm/µg for ^{3}H-labelling, 10^7 cpm/µg for ^{32}P-labelling) and percentage incorporation into hairpin-like structures (generally 30-40 percent). The radio-labelled single-copy DNA was sized by sedimentation of an aliquot on a 12 ml 4-11 percent alkaline sucrose (0.9 M NaCl, 0.1 M NaOH) gradient (167), in the IEC SB-283 swing-out rotor, for 16 hours at 40,000 rev/min (270,000 g_{max}) at 20°C. Weight-average single-strand length, determined by use of a computer programme for calculating S values (593) was 200 nucleotides. Radio-labelled single-copy DNA was lyophilised, desalted on a Sephadex G50/AG 50W-X8 (Na^+ form) column, and stored at -20°.

3.6. Preparation of cDNA probes by reverse transcription

3.6.1. Reverse transcription of poly(A)$^+$ RNAs

Nuclear, polysomal and *in vitro*-transported poly(A)$^+$ RNAs were reverse transcribed essentially as described by Getz *et al* (167). Avian myeloblastosis virus reverse transcriptase (obtained from the Division of Cancer Cause and Prevention, National Cancer Institute) was added (30-60 units/µg RNA depending on enzyme batch) to reaction mixtures containing:

50 mM Tris/HCl (pH 8.2), 10 mM Magnesium acetate, 2 mM DTT, 40 mM NaCl; poly(A)$^+$ RNA at 10-20 µg/ml, oligo(dT)$_{17}$ (PL Biochemicals) at the same concentration (w/w) as RNA: 100 µg/ml actinomycin D (Sigma), dATP, dGTP and dTTP (Boehringer) at 2 mM, ^{3}H-dCTP (19 Ci/mmol) at 50 µM, giving a final specific activity of 9.5 x 10^6 cpm/µg cDNA, or 2.5 µM ^{32}P-dCTP (400 Ci/mmol), giving a final specific activity of 4-6 x 10^8 cpm/µg, depending on the freshness of the isotope; and RNase inhibitor (Searle) at 10 units/ml.

Following incubation at 37° for 2 hours, reactions were quenched on ice for 2-3 minutes with the addition of EDTA to 10 mM, and E.coli DNA, SDS and NaCl were then added to 20-100 µg/ml, 0.5 percent and 100 mM respectively, after which the mixtures were phenol/chloroform extracted (Section 3.3). ^{32}P-labelled cDNAs were alkali-treated by making the aqueous phase 0.3 N in NaOH, and incubating for 1 hour at 37°, after which the solution was neutralised by dropwise addition of 1 M Sodium acetate pH 4.5, with checking of pH by indicator papers, gel filtered into 1/2 x SSC on a Sephadex G-50/dowex column, and lyophilised. The cDNAs were dissolved in the appropriate solution for use in filter-bound DNA-driven hybridisation reactions (see Section 3.12.2). ^{3}H-labelled cDNAs were gel-filtered, after phenol extraction, into 50 mM NaCl, 10 mM HEPES, pH 7.0 on Sephadex G50, and lyophilised or ethanol precipitated.

3.6.2. Size characterisation and fractionation of cDNA probes

^{3}H-labelled cDNAs were sized on 4-11 percent alkaline sucrose gradients (see Section 3.5.2) and fractionated by upward displacement using fluorochemical FC43. Where necessary (only in the case of nuclear cDNA preparations) material sedimenting at < 3S was excluded, the remainder being pooled and ethanol precipitated. Size comparisons of polysomal cDNAs to rat liver and HTC cell mRNA preparations are shown in Figs. 2a-b. They are almost identical, showing approximate mean single-strand lengths of 350 nucleotides on the basis of computer

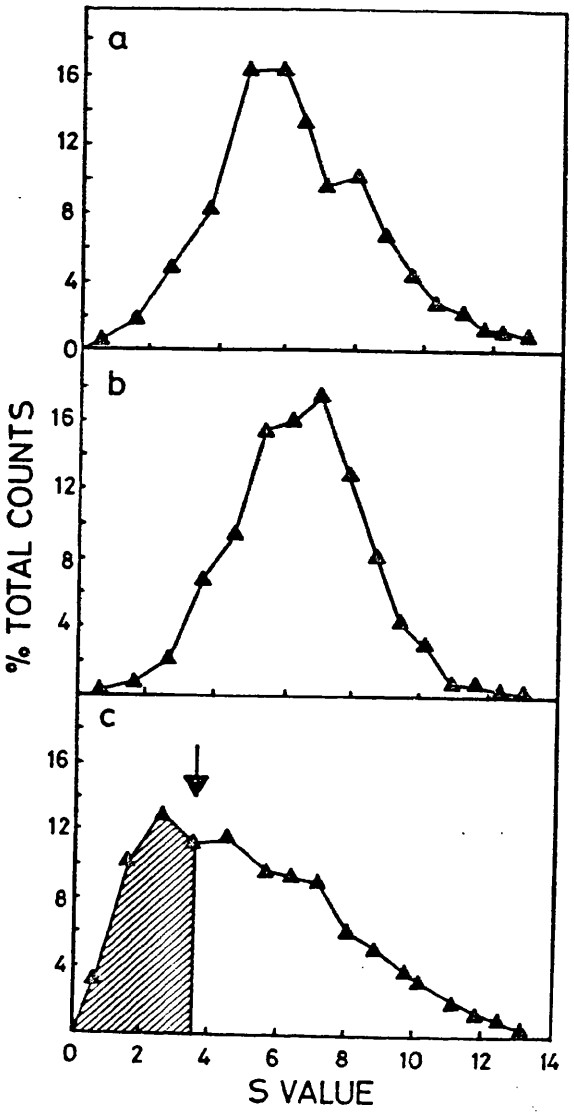


FIGURE 2. Sizing of complementary DNA probes on 4-11 percent alkaline sucrose gradient , centrifugation conditions are as stated in the text.

(a) Rat liver polysomal cDNA.
(b) HTC cell polysomal cDNA.
(c) Rat liver nuclear cDNA.

The proportion of counts in each gradient fraction is plotted against its S value as determined by computer analysis (211). Molecular weight is given by the equation $S_{20,w} = 0.0528\ M^{0.4}$ (ref. 593). Shading indicates gradient fractions discarded in order to eliminate short fragments (< 50 nucleotides), and generate a probe of reasonably homogeneous size.

analysis (593). Nuclear cDNA preparations were invariably shorter (Fig. 2c). cDNA probes were recovered from ethanol precipitates by centrifugation, redissolved in sterile water, desalted on Sephadex G50/dowex, and stored at -20°.

3.7. Fractionation of cDNA probes to enrich for 'abundant' sequences

3.7.1. Mercuration of RNA

Poly(A)$^+$ RNA preparations and E.coli RNA (Sigma) were mercurated as described by Wilkes et al (145) by incubation at 100 µg/ml, for 90 minutes at 50° in a freshly prepared solution, containing 1 mg/ml Mercuric acetate, 5 mM Sodium acetate, pH 6.0. The reaction was quenched by addition of EDTA to 10 mM, and mercurated RNA was isolated by gel filtration into sterile water on a Sephadex G50 column (without dowex), lyophilised, and stored at -20°C.

3.7.2. Partial hybridisation reactions

^{3}H-labelled cDNA probes were lyophilised with at least a ten-fold mass excess of their template RNAs following mercuration, and dissolved in HB/F supplemented with 5 mM 2-mercaptoethanol (Sigma), to a final concentration of 1-2 mg/ml in RNA. After heat denaturation at 70° for 5 minutes, the reaction mixtures in sealed, siliconised sterile capillaries were hybridised at 43° to an appropriate $R_o t$ value, selected on the basis of the kinetics of homologous hybridisation (indicated in the text in the results section). The partially hybridised cDNAs were flushed out from the capillaries by a large excess of NETS buffer and fractionated by thiol-sepharose chromatography.

3.7.3. Thiol-sepharose chromatography

1 ml columns of thiol-sepharose (prepared in the Beatson Institute by Drs Allan Balmain and Peter Wilkes, following the procedure of Dale and Ward (594), from cross-linked sepharose 4B supplied by Pharmacia) were activated for 30 minutes in 50 mM DTT, 0.5 M Tris/HCl, pH 8.0 at room temperature, and washed extensively at 60° with NETS buffer (about 20 bed volumes). Samples were applied dropwise at a rate of less than 0.1 ml/min, in NETS buffer at 60°, together with 10 µg of E.coli DNA carrier and where the amount of mercurated RNA present was less than 1 µg, 1 µg of mercurated E.coli RNA was also added to the sample. The unbound (unhybridised) material was eluted by sequential washes with NETS buffer at 60°. Following further extensive washing at 20° with NETS buffer, bound material was eluted with NETS containing 0.1 M β-mercaptoethanol. The use of thiol-sepharose chromatography in RNA-driven hybridisation (both preparatively and analytically) is discussed by Brown and Balmain (595). Hybridised or unhybridised DNA from preparative fractionations was recovered by lyophilisation of the appropriate eluate, alkaline hydrolysis in 0.5 N NaOH at 37° for 1 hour in a minimal volume, neutralisation by dropwise addition of 1 M sodium acetate, pH 4.5 using pH indicator paper and desalting on a Sephadex G50/dowex column. The hybridised material in a partial hybridisation reaction corresponds with the fraction enriched for abundant sequences. Thiol-sepharose was re-usable, and was regenerated by extensive washing with NETS, in which it was stored at 4°, and re-activation before use.

3.8. Solution hybridisation reactions

3.8.1. Reaction conditions

Samples for RNA-driven hybridisation, consisting of mixtures of cDNA and an appropriate mass excess (detailed in Figure

legends) of RNA, were lyophilised and redissolved either in HB/F or in HB/S (0.6 M NaCl, 50 mM HEPES, 0.1 percent SDS, pH 7.5, chelexed and diethyl-pyrocarbonate-sterilised). Aliquots of 0.5 - 10 µl, as appropriate, were loaded into siliconised, sterile capillaries and sealed in a Bunsen flame. Heat denaturation was at 70° for reactions in HB/F, or at 100° for reactions in HB/S, and hybridisations were carried out at 43° in HB/F or 70° in HB/S. Reactions were quenched at the appropriate time-points by plunging the capillaries into ice-cold ethanol, wiping dry, and immediately flushing out with a solution appropriate to the assay which was to be employed (see following section 3.8.2).

3.8.2. Assays of RNA-driven hybridisation

Analytical RNA-driven hybridisation reactions were assayed either by S1 nuclease digestion or by thiol-sepharose chromatography. Where the S1 nuclease assay (167) was to be employed, reaction capillaries were flushed out with 0.25 ml S1 buffer (140 mM NaCl, 70 mM Sodium acetate, 2.8 mM $ZnSO_4$, pH 4.5) containing 14 µg/ml calf-thymus DNA, and stored at -20°. All samples from a given reaction, including zero time controls, were assayed in parallel by the addition of 0.1 ml S1 buffer containing 10 units S1 nuclease (Sigma) per sample, incubation for 1 hour at 37°, and comparison of total and acid-soluble counts (the latter being determined by precipitation for 15 minutes on ice by 0.5 N perchloric acid, in the presence of 5 µg and 50 µg respectively of carrier (calf-thymus) DNA and BSA.

Where assay was by thiol-sepharose chromatography, 5 mM 2-mercaptoethanol was included in the hybridisation buffer (only HB/F was employed in this case). Reactions were flushed out in 0.5 ml of NETS, and chromatographed on thiol-sepharose as described previously (see Section 3.7.3). The extent of reaction was determined by comparing the counts

in the bound and unbound fractions.

For reactions in HB/S at 70°, $R_o t$ values are corrected for the (five-fold) accelerated rate of hybridisation, and all values quoted are 'equivalent $R_o t$'.

3.8.3. Assay of DNA-reassociation

DNA reassociation reactions were set up and performed exactly as described for RNA-driven hybridisations. Capillaries were flushed out with 30 mM phosphate buffer, 150 mM NaCl, and samples were analysed by HAP chromatography as already described (Section 3.5.1).

3.9. Cell-free translation

Poly(A)$^+$ RNAs were translated in the mRNA-dependent reticulocyte lysate, prepared and used exactly as described by Pelham and Jackson (578). Materials supplied were as follows: rabbit reticulocyte lysate (prepared and kindly donated by Ms. Carol Finnegan), micrococcal nuclease (Worthington), creatine phosphate (Sigma) creatine phosphokinase (Sigma). The labelled amino acid used was ^{35}S-Methionine. Aliquots of nuclease-treated lysate were stored without label at -70°. Reactions were quenched after 60 minutes by the addition of 4 volumes of gel sample buffer (0.1 M DTT, 25 mM Tris/HCl, pH 6.8, 2 percent SDS, 10 percent (v/v) Glycerol, 0.0008 percent bromophenol blue).

3.10. Polypeptide analysis by SDS-polyacrylamide gel electrophoresis

SDS-15 percent polyacrylamide gels were used to compare the sets of polypeptides synthesised in cell-free translation, and in cellular fractions *in vivo*. The method used was essentially that of Laemmli (596) as modified by Anderson

et al (597), using a 5 percent polyacrylamide stacking gel. Samples were prepared from cell-free translation as already described, and from aliquots of post-mitochondrial supernatant, obtained during rat liver and HTC cell polysome preparation (Sections 3.2.1 and 3.2.4) by the addition of 4 volumes of gel sample buffer (Section 3.9). In addition, a sample was similarly prepared from the membrane-bound polysome-enriched fraction (rough ER) described in Section 3.2.3. For comparative purposes, samples of cell-free translation products containing equal numbers of counts precipitable in 8 percent TCA at room temperature (578) were electrophoresed. The protein-content of in vivo-polypeptide samples was determined, using Bradford's Coomassie blue assay (598), using unused disposable plastic curettes and glass test-tubes, and once again equal amounts were loaded on gels for comparative purposes. Samples were heated to 100° for 2 minutes prior to loading. Marker polypeptides (Pharmacia) were electrophoresed in parallel with experimental samples. Gels were run overnight at 8 mA, in running buffer containing glycine (28.8 g/1), Tris base (6 g/1) and 0.1 percent SDS, pH 8.5.

Gels were fixed in 10 percent acetic acid, 50 percent methanol for 30 minutes, stained in 1.25 percent coomassie blue (Sigma) in the same solvent for 30 minutes, and destained in 10 percent methanol, 7 percent acetic acid at 60° with several changes. Destained gels were photographed, dried down under vacuum, and, if appropriate, autoradiographed using Fuji R-X X-ray film.

3.11. Isolation and characterisation of plasmid DNAs

3.11.1. Growth of bacteria and plasmid DNA extraction

Plasmid pAT 153, a derivative of pBR 322 (599) was grown in E.coli strain C600. Large-scale cultures (e.g. 1/2 1 in a 2l flask) were seeded in L-broth (yeast extract and bacto-

tryptone were from Difco) at an OD_{595} of 0.05, and incubated at 37° with continuous shaking. During exponential growth, when the OD_{595} had reached 0.6, chloramphenicol (Sigma) was added to a final concentration of 150 µg/ml and the cultures amplified for a further 18 hours at 37° with shaking. Cells were harvested by centrifugation at 5000 rev/min (4000 g_{max}) for 5 minutes at 4°, washed in SSC, and resuspended in 50 mM Tris/HCl, pH 8.0, 25 percent sucrose (approximately 10 ml per 1 of culture). Lysosyme (Sigma) freshly dissolved at 10 mg/ml in 250 mM Tris/HCl, pH 8.0, was added (320 µl per ml of suspension), and after incubation at 4° for 15 minutes with gentle shaking, 250 mM EDTA, pH 8.0, was added (440 µl per ml of suspension) and the mixture incubated a further 5 minutes at room temperature. Following dropwise addition of 2.5 percent SDS (16 ml per litre of original culture) with continuous swirling, and swirling for a further 5 minutes, 2 ml (per litre of original culture) of self-digested pronase (Calbiochem) solution (20 mg/ml), were added, and the lysate incubated 1.5 hours at 37°. The lysate was then spun for 2 hours at 21,000 rev/min (46,000 g_{max}) in 35 ml tubes in the MSE 8 x 35 Titanium fixed-angle rotor at 20°, and the supernatant carefully decanted from the loose pellet of chromosomal DNA. CsCl (1 g per ml of solution, plus 1 g per ml of additional volume to be introduced subsequently with Ethidium bromide addition) was added, dissolved by repeatedly inverting the vessel and warming to 37°, and the solution centrifuged at 10,000 rev/min (16,000 g_{max}) for 10 minutes at 20° to pack the protein precipitate into a hard pellicle, thereby allowing the DNA-containing solution to be removed easily using a syringe fitted with a canula. Ethidium bromide (Sigma) dissolved in 50 mM Tris/HCl, 20 mM EDTA, pH 8.0, was added to a final concentration of 500 µg/ml, and mixed by magnetic stirring in the dark. The solution was transferred to new 25 ml polycarbonate centrifuge tubes which were filled up and balanced using mineral oil, and spun at 36,000 rev/min (140,000 g_{max}) for 60 hours at 20°C

in an MSE 8 x 25 Titanium rotor. The lower of the two visible bands, containing the plasmid DNA, was carefully removed by downward displacement with water, after the upper band had been removed and discarded. The solution was extracted several times, until colourless, with an equal volume of isopropanol, saturated with CsCl, and dialysed extensively against TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 8.0). This plasmid DNA isolation method is essentially that of Birnie (600).

Recombinant plasmid DNAs from a rat liver polysomal cDNA library (constructed by Ms. Rosemary Shott in this laboratory, by blunt-end ligation of double-stranded cDNA into the Bam Hl site of pAT 153) were isolated similarly from bulk cultures grown in E.coli strain HB 101, amplified and lysed by Ms. Rosemary Shott under appropriate containment conditions.

3.11.2. Characterisation of plasmid DNAs by agarose gel electrophoresis and restriction analysis

Plasmid DNAs were electrophoresed on 1 percent agarose gels (1 percent agarose (Sigma), 50 mM Tris/HCl, 20 mM sodium acetate, 1 mM EDTA, pH 8.0). Plasmid DNA samples were ethanol-precipitated and redissolved in TE buffer at 100 µg/ml, and 2.5 µl of each sample was added to 12.5 µl distilled water and 5 µl DNA-gel sample buffer (30 percent sucrose, 5 percent SDS, 0.1 percent bromophenol blue, 10 mM Tris/HCl, 1 mM EDTA, pH 8.0) for loading onto the gel (run horizontally). Samples were run into the gel at 60V, then run overnight at 25V, stained with ethidium bromide, and photographed under u.v. Size markers of Eco RI-digested phage λ DNA (donated by Ms. Mary Branwood) were run in parallel.

Further samples of the plasmid DNAs precipitated and redissolved in TE, were digested at 37° for 1 hour with Eco RI (Miles) and Hinf I (BRL) restriction endonucleases, in a reaction

mixture containing 100 mM Tris/HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 12.5 µg/ml DNA, 1 unit/µg DNA Eco RI and 9 units/µg DNA Hinf I. Reactions were quenched by the addition of a tenth volume (4 µl) of DNA-gel sample buffer, and loaded onto a 6 percent polyacrylamide/0.15 percent bis-acrylamide gel (containing 40 mM Tris/HCl, 20 mM sodium acetate, 2 mM EDTA, pH 7.8, polymerised with 150 µl TEMED and 670 µl 10 percent Ammonium persulphate per 100 ml gel mix). The gel was run overnight at 30V in a circulating buffer of the same composition as that in the gel mix, stained with 1 µg/ml ethidium bromide, and photographed under u.v.

3.12. Hybridisations to filter-bound plasmid DNA

3.12.1. Binding of DNA to filters

Plasmid DNAs were nicked by boiling for 5 minutes in TE buffer. Aliquots of 1 µg of plasmid DNA in 0.4 ml SSC were denatured by the addition of 0.4 ml 1 N NaOH at room temperature for 30-60 minutes. The solution was neutralised by dropwise addition of 0.2 N HCl, 9 x SSC containing phenol red indicator, and immediately passed slowly (in 30 seconds) through a 1 cm diameter nitrocellulose filter (Millipore, catalogue number HAWP 01300), pre-wetted with 6 x SSC. After washing with 5 ml 6 x SSC, passed at the same flow rate, each such filter was immersed in 70 percent ethanol and baked for 2 hours at 80^{o}.

3.12.2. Hybridisation conditions and assay

Filters were prehybridised at 43^{o} in 1X millipore-filtered Denhardt's solution (601),5 x SSC, 50 percent formamide, 0.1 percent SDS, 100 µg/ml salmon sperm DNA (Sigma, phenol/ chloroform extracted and re-precipitated) for 3-4 hours. Lyophilised ^{32}P-labelled cDNA probes were each redissolved in an appropriate volume, such that final concentrations in the probe-containing hybridisation solutions were 1 x

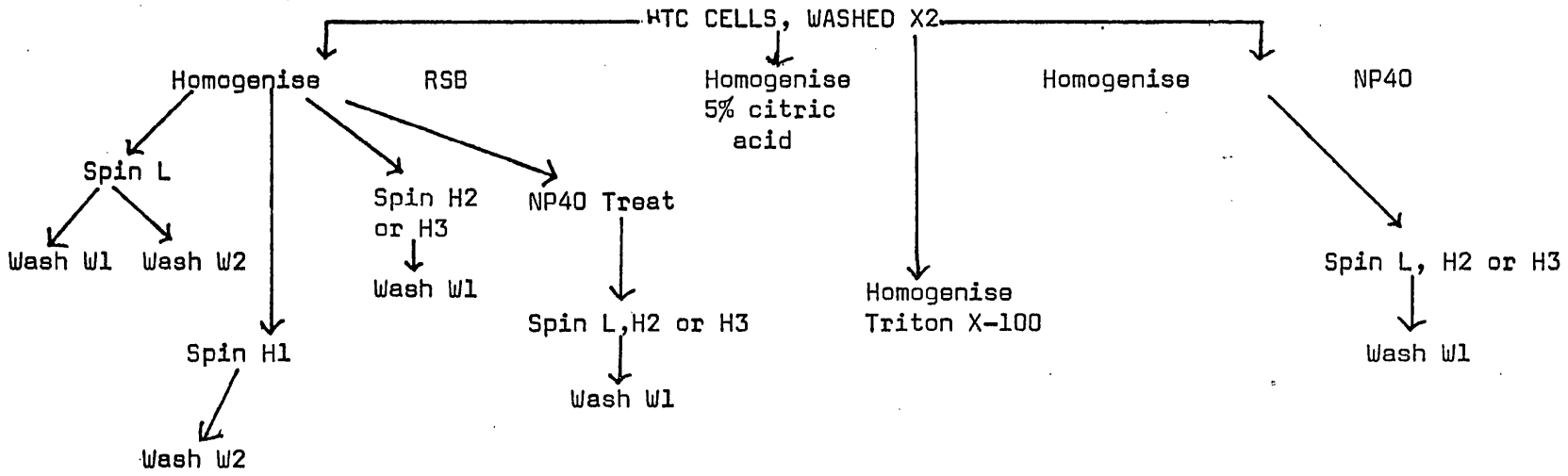
millipore-filtered Denhardt's solution, 5 x SSC, 0.1 percent SDS, 50 percent formamide, 100 μg/ml salmon sperm DNA. Replicate sets of filters, bearing different plasmid DNAs, were each hybridised for 55 hours at 43° with 1 ml of the probe-containing solutions, previously denatured for 5 minutes at 70°. Aliquots of the probe-containing solutions were retained for counting. After hybridisation, filters were washed 4 times, for half an hour, in a large volume of 0.1 percent SDS, 0.1 x SSC at 60°, dried, and counted. Extent of hybridisation, as measured by the proportion of counts bound to the filters, was compared with that of a control reaction, carried out under identical conditions, between mouse reticulocyte cDNA (donated by Mr Colin Casimir) and a mouse α-globin-cDNA-containing plasmid, derived from pCR1 (601) (provided by Mr Alan Lyons, originally donated by the laboratory of Dr Charles Weissmann).

3.13. Preparation of nuclei and cytosol for studies of RNA transport in vitro

3.13.1. Variants of methods for preparation of nuclei

As will be described in Section 4.7.1 various methods were investigated in order to determine the most appropriate procedure for preparing HTC cell nuclei for use in in vitro studies of mRNA transport. These are summarised in Chart 1, which takes the form of a flow-chart showing the various alternatives employed. Aliquots of HTC cells were resuspended in RSB (10 mM Tris/HCl, 10 mM NaCl, 1.5 mM $MgCl_2$, pH 7.5) plus or minus NP40 (0.5 percent w/v). The various options used after homogenisation and recovery of the crude nuclear pellet (by centrifugation at 2000 rev/min for 5 minutes), included: low speed spinning (2000 rev/min) from 0.3 M sucrose, 3 mM $CaCl_2$ through 0.88 M sucrose, 3 mM $CaCl_2$; high speed spinning (25000 rev/min in IEC SB-110 swing out rotor for 2 hours) from 1.6 M sucrose, 3 mM $CaCl_2$ through either 1.9 M

CHART 1: Flow chart showing methods used in HTC nuclei preparation.



Homogenisation conditions: RSB: resuspend in RSB, swell osmotically, homogenise, spin out crude nuclei, 500 g_{max}.

NP40: resuspend in RSB/0.5% NP40 and treat similarly.

TRITON X-100: resuspend in 0.1% Triton X-100 containing buffer and proceed as Roy et al. (ref. 514).

NP40-Treat.: Nuclei resuspended in 0.3 M sucrose, 0.5% NP40, 3 mM $CaCl_2$ and hand homogenised.

Spins L,H1,H2,H3: Spins through 3 mM $CaCl_2$/0.88 M sucrose at 1000 g_{max}, or 1.6, 1.9 or 2.2 M sucrose/3 mM $CaCl_2$ at 110,000 g_{max}.

Wash W1: wash in 1 mM $CaCl_2$.

Wash W2: wash in 0.25 M sucrose, 0.1% Triton X-100, 3 m $CaCl_2$ and repeat without Triton.

or 2.2 M sucrose, 3 mM $CaCl_2$, or direct pelleting; these same options but including treatment with 0.5 percent NP40 in 0.3 M sucrose, 3 mM $CaCl_2$ prior to dilution of the detergent with a large volume of denser sucrose; final wash either in 1 mM $CaCl_2$ or in 0.25 M sucrose, 3 mM $CaCl_2$ following treatment with Triton X-100 (0.1 percent v/v) in the same solution. A method employing initial resuspension in a buffer containing 0.1 percent Triton X-100, similar to that of Roy et al (514), and the sucrose/citric acid method (see Section 3.2.2) were also used.

3.13.2. Method of choice for preparation of nuclei and cytosol

HTC cell pellets were resuspended on ice in RSB (2.5×10^7 cells/ml), swollen for 10 minutes, and motor-homogenised in a teflon/glass homogeniser (3-4 strokes, lysis checked microscopically). The suspension was made 0.3 M in sucrose, and centrifuged for 5 minutes at 4^o at 2000 rev/min (1000 g_{max}). The crude nuclei were resuspended in ice-cold 0.3 M sucrose, 3 mM $CaCl_2$ by hand homogenisation (about 8×10^7 nuclei/ml). NP40 was added to 0.5 percent (v/v) and the suspension briefly vortexed and hand-homogenised in a single stroke. 2 M sucrose, 3 mM $CaCl_2$ was added to a final sucrose concentration of 1.6 M, and the suspension layered over a 10 ml pad of 2 M sucrose, 3 mM $CaCl_2$ in 35 ml polypropylene centrifuge tubes, and spun at 25,000 rev/min (110,000 g_{max}) for 2 hours in the IEC SB-110 rotor at 4^o. Nuclei were resuspended by hand homogenisation in 1 mM $CaCl_2$ and recovered by centrifugation at 4^o for 5 minutes, at 1000 rev/min (250 g_{max}).

Cytosol was prepared from HTC cells swollen and homogenised using the same procedure. After centrifugation at 10,000 rev/min (16,000 g_{max}) to remove large organelles, the sample was centrifuged at 55,000 rev/min (350,000 g_{max}) for 4 hours in an MSE 8 x 25 Titanium swing-out rotor, at 4^o. The post-microsomal supernatant was carefully removed and vacuum

dialysed against distilled water at 4^{o} using a Sartorius membrane filter apparatus. For bulk preparations the supernatant was sealed in dialysis tubing and concentrated by treatment with polyethylene glycol (Koch-Light) at 4^{o}, followed by extensive dialysis against distilled water at 4^{o}. Protein concentrations were monitored using Bradford's assay (598), and cytosol preparations after removal of insoluble material by low speed centrifugation, were stored at 5-10 mg/ml of protein in aliquots at -20^{o}. Rat liver cytosol was prepared similarly.

3.14. Assay of RNA transport from isolated HTC cell nuclei

3.14.1. Cell labelling regimes

HTC cells undergoing exponential growth in suspension culture at cell densities of 0.5 - 0.8 x 10^{6} cells/ml were steady-state labelled (in nuclear RNA) by incubation for 18 hours in their usual medium, supplemented with 1 mCi/l ^{3}H-uridine (50 Ci/mmol). Pulse-labelling was accomplished by centrifuging cells at 25^{o} for 15 minutes at 1500 rev/min (700 g_{max}), and resuspending at 3 x 10^{7} cells/ml in fresh warmed medium. After incubation for 15 minutes at 37^{o} with continuous swirling, 100 µCi/ml ^{3}H-uridine was added for a further 15 minutes. Following this, cells were cooled on ice and centrifuged for 5 minutes at 4^{o} at 1500 rev/min (600 g_{max}) and washed as normally.

3.14.2. Incubation conditions for in vitro RNA transport

The procedure followed was essentially similar to that of Schumm and Webb (519). HTC nuclei were incubated at 30^{o} at a density of approximately 0.5-1 x 10^{8} nuclei/ml in a medium containing 50 mM Tris/HCl, pH 7.5, 25 mM KCl, 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.3 mM $MnCl_2$, 5 mM NaCl, 5 mM Spermidine (Sigma), 2 mM DTT, 2.5 mM Na_2HPO_4, 10 mM creatine phosphate, 50 µg/ml creatine phosphoteinase, 5 mM ATP (Sigma), 350 µg/ml E.coli

(Sigma) RNA as carrier and HTC cell or liver cytosol at 2.4 mg/ml of protein, except where stated in the text. In the course of experimentation, various parameters of the incubation medium were varied, and this is indicated where appropriate in the text or in Figure legends.

3.14.3. Assay of RNA release to the medium

At various time-points, sample incubations were cooled on ice for 30 seconds, and the amount of labelled RNA released to the medium was measured by centrifugation for 2 minutes at 4^{o} in a micro-centrifuge to remove nuclei. Aliquots were counted, of the post-nuclear supernatant and of the total suspension, in Dimilune-30 scintillant (Packard) following solubilisation in Soluene^{-350} (Packard). Acid-precipitable counts released to the medium were determined by pre-initiation of total and released RNA on ice, in 10 percent TCA, followed by filtration GF/C filter circles, and extensive washing with 10 percent TCA, 1 percent pyrophosphate.

3.15. Sizing of in vitro-transported RNA and RNP by rate-zonal sedimentation

RNA transported in vitro under various conditions was recovered from post-nuclear supernatants made 0.5 percent in SDS and 0.1 M in NaCl by phenol/chloroform extraction (Section 3.3) and ethanol precipitation. RNA was resuspended in a minimal volume of NETS buffer and centrifuged through 12 ml 15-30 percent sucrose/NETS gradients in an IEC SB-283 swing-out rotor, at 20^{o} at 40000 rev/min (270,000 g_{max}) for 6 hours. Gradients were analysed and fractionated by upward displacement with fluorochemical FC43, followed by TCA precipitation and counting (Section 3.14.3).

RNPs were analysed by centrifuging aliquots of post-nuclear supernatant on 12 ml 15-40 percent sucrose gradients either in TNM buffer (Section 3.2.4) or in TNM containing 10 mM EDTA instead of 1.5 mM $MgCl_2$). Centrifugation was at 4^o in an IEC SB-283 swing-out rotor, for 9 hours at 20,000 rev/min (68,000 g_{max}). Gradients were analysed as described above.

4. RESULTS

4.1 Complexity and diversity of rat liver and HTC cell nuclear and polysomal poly(A)$^+$ RNAs

4.1.1. Preparation and characterisation of single-copy rat DNA probe

As already indicated, the most reliable estimates of base sequence complexity, and of the qualitative differences between RNA populations, are by saturation hybridisation of single-copy DNA in vast RNA excess. Therefore, single-copy rat DNA was isolated and radiolabelled with ^3H or ^{32}P, by gap translation (Section 3.5.2). Since the accuracy of complexity measurements depends upon the 'uniqueness' of such a probe, its re-annealing characteristics, driven by excess total sonicated rat DNA, were followed by hydroxylapatite chromatography. The data (Fig. 3) are fitted to a theoretical second-order rate equation by least-squares analysis by computer (602). The close fit indicates that the great majority (probably more than 98 percent) of sequences in the probe are at homogeneous frequency in genomic DNA, the increased $C_o t_{1/2}$ compared with that of the most slowly re-annealing component of total rat DNA, being due to the smaller size of the probe (200 nucleotides as judged by sedimentation on 4-11 percent alkaline sucrose gradients) than that of extensively sonicated DNA (about 600 nucleotides). The hybridisability of freshly labelled probes, which were used wherever possible, was approximately 75 percent, regardless of the label used. The reactivity of ^3H-labelled probes declined by approximately 5 percent per month when stored in aqueous solution at -20^o. In all saturation hybridisations the extent of reaction was corrected for probe hybridisability, which was measured at the time of reaction by annealing a portion of the probe to $C_o t = 30{,}000$ moles.s.l^{-1} in a large excess of sonicated total rat DNA.

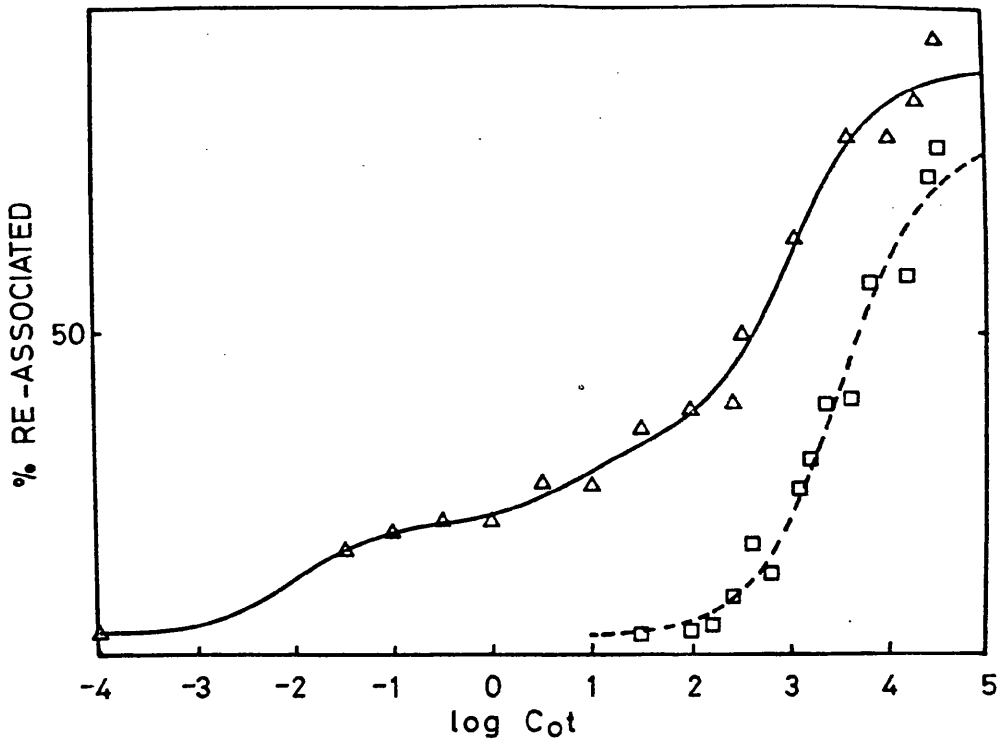


FIGURE 3. Re-association of ^{3}H labelled single-copy rat DNA, driven by excess sonicated total rat DNA (□ , ---- curve) and of driver DNA (Δ , ——— curve), determined by OD_{260} of single and double-stranded fractions, assuming $\varepsilon = 40\ g^{-1}1.cm^{-1}$ for single-stranded DNA in salt, and $\varepsilon = 50\ g^{-1}.1.cm^{-1}$. for double-stranded DNA. Curves fitted by computer analysis (602).

4.1.2. Saturation hybridisations using single-copy DNA

Single-copy DNA saturation hybridisations were carried out in order to quantify the complexities of mRNA and hnRNA populations of liver and HTC cells, to characterise the extent of qualitative post-transcriptional regulation between nucleus and polysomes in these cell-types, and to determine the homology in sequence expression between them. Mercurated poly(A)$^+$ nuclear and polysomal RNA preparations were hybridised in 100-200-fold excess with aliquots of radiolabelled single-copy rat DNA. Under these conditions of excess, sequences which represent as little as 10^{-8} of the RNA should be completely hybridised at sufficiently high $R_o t$ values. Reactions were taken to a $R_o t$ of 10-30,000 moles.s.l^{-1}, and hybridisation levels were assayed by thiol-sepharose chromatography. In addition, mixtures of nuclear or polysomal RNAs from the two cell-types were hybridised with the probe under identical conditions, in order to determine the extent to which they overlap. The saturation level should be the sum of those obtained with the RNAs individually if the two populations are entirely distinct, or less than this, according to the degree of homology between them. This method was used by Axel et al (174) to demonstrate tissue-specific mRNA complexity in the chick. In all cases the background hybridisation with mercurated poly(U) or E.coli RNA (shown in Fig. 4a) was deducted from the hybridisation levels obtained with the experimental samples. The corrected data for the hybridisations with polysomal poly(A)$^+$ RNAs are shown in Figs. 4b-d, and with poly(A)$^+$ nuclear RNAs in Figs. 5a-c. Saturation hybridisation levels have been deduced by inspection (strictly mathematical curve-fitting using a pseudo-first-order single-component rate equation being considered no more appropriate, in view of the diverse relative abundance patterns revealed by kinetic cDNA hybridisation - see Section 4.2.

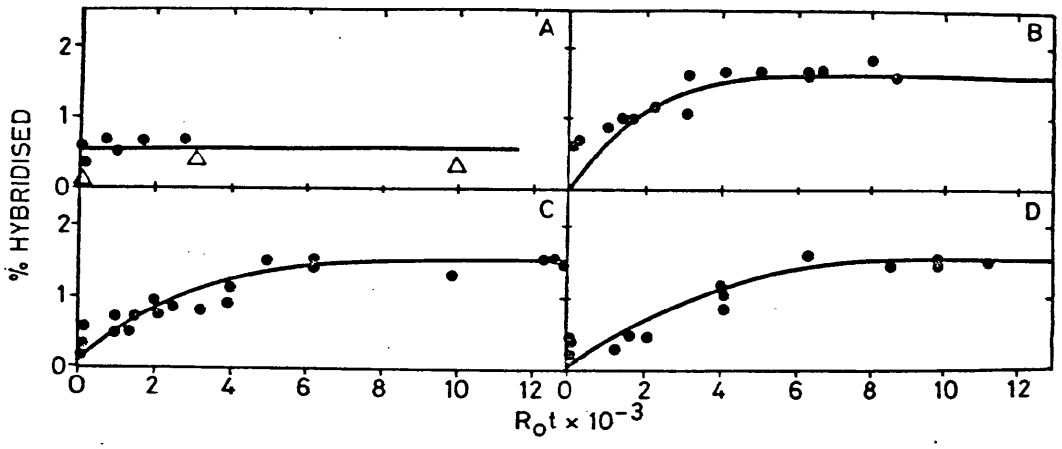


FIGURE 4. Hybridisation of ^3H labelled single-copy rat DNA with mercurated RNA samples. (a) Background hybridisation with mercurated poly U (△) or E.Coli RNA (●) at 5 mg/ml, which is subtracted from experimental values plotted in (b), (c) and (d). Experimental hybridisations were with (b) mercurated rat liver polysomal poly(A)$^+$ RNA at 5 mg/ml (100 fold RNA excess, (c) mercurated HTC cell polysomal poly(A)$^+$ RNA at 5 mg/ml (100 fold RNA excess), and (d) a mixture of mercurated rat liver and HTC cell polysomal poly(A)$^+$ RNAs, each at 5 mg/ml (200 fold RNA excess). RNA concentrations determined spectrophotometrically (assuming ε = 40 $g^{-1}.l.cm^{-1}$). Hybridisation was at 43°, analysed by thiol-sepharose chromatography. All values are normalised for probe hybridisability (69 percent).

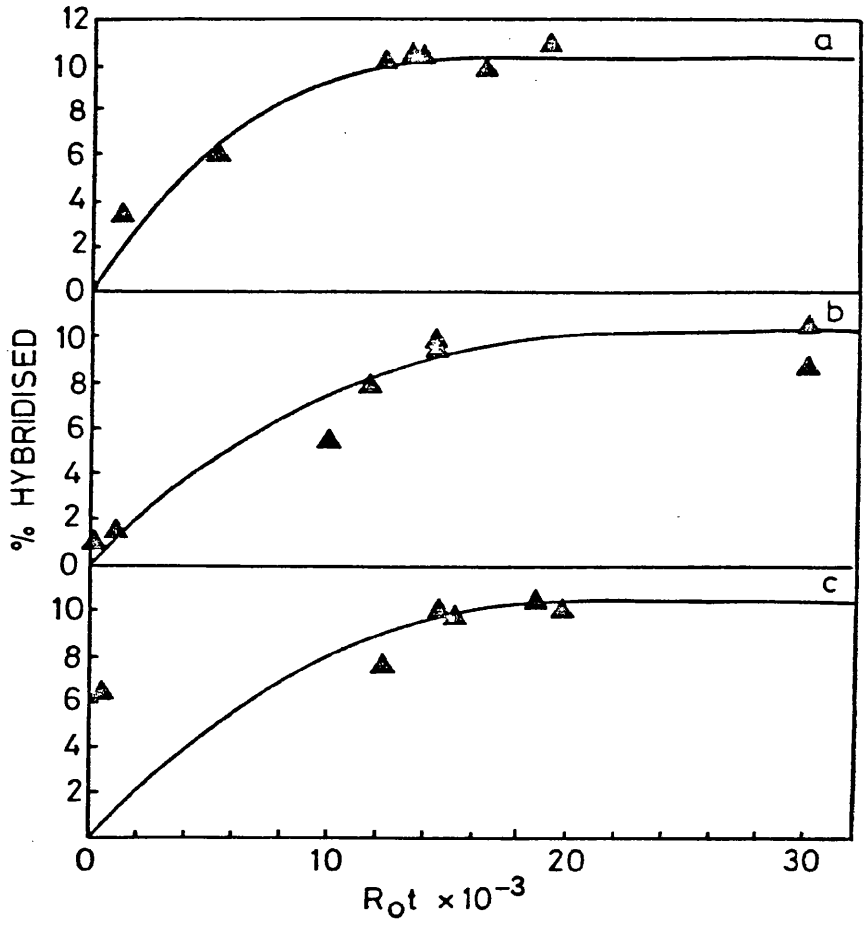


FIGURE 5. Hybridisation of ^{32}P-labelled single-copy rat DNA with (a) mercurated rat liver poly(A)$^+$ nuclear RNA at 20 mg/ml (1330-fold RNA excess); (b) mercurated HTC cell poly(A)$^+$ nuclear RNA at 20 mg/ml (1330-fold RNA excess); (c) mercurated rat liver and HTC cell poly(A)$^+$ nuclear RNAs each at 10 mg/ml (1330-fold RNA excess). All values are normalised for probe hybridisability (75 percent) with background hybridisation (Fig. 4a) subtracted. Hybridisation was at 43°, analysed by thiol-sepharose chromatography.

The proportions of unique sequence complexity expressed in these RNA populations, given both as a percentage and in nucleotides of transcribed DNA, are summarised in Table 2. In terms of absolute complexity, these values are subject to a number of potential errors concerning the probe and assay method, which have already been discussed (Section 2.2.2). Clearly, if any repetitive sequences which remain in the probe are also present in the driver RNA, the saturation levels will tend to overestimate complexity, which will also be the case if there is insufficient RNA excess to prevent DNA re-annealing with single-stranded DNA tails of RNA-DNA hybrids. The reasonably good plateau levels of hybridisation obtained here suggest that this is not a problem. However, since demercuration of RNA incubated for prolonged reaction times can lead to spurious underestimates of hybridisation (595), stable plateau levels could conceivably be generated by opposing artifacts. Despite these uncertainties, the validity of the comparative complexity estimates of liver and hepatoma RNAs depends only on the reproducibility of the plateau hybridisation levels. These are estimated, from the data obtained, as being accurate to approximately $\pm$ 7 percent of the observed saturation levels. No significant difference is therefore detectable between liver and hepatoma RNA complexities, neither at nuclear, nor polysomal levels. Furthermore, the mixed RNA-driven reactions reached the same plateau levels. This indicates that the two cell-types contain essentially indistinguishable sets of sequences, although as already stated, this technique is insensitive to small differences in complexity which may be of great biological significance. Notwithstanding, these results indicate that the number of polysomal poly(A)$^+$ mRNA sequences specific to either cell-type is unlikely to exceed about 1000, out of a total of some 27000 different species of average length 2000 nucleotides. At the level of poly(A)$^+$ nuclear RNA, sequences specific to either cell-type are no more than three times this number. Conversely, the data reveal highly

TABLE 2. Saturation estimates of poly(A)$^+$ hnRNA and mRNA complexity from rat liver and HTC cells: derived from the data of Figs. 4,5.

RNA Fraction	Units	Rat liver complexity	HTC cell complexity
Poly(A)$^+$ hnRNA	percentage single-copy sequences	10.2 ± 0.7 (b)	10.5 ± 0.7
	nucleotides of transcribed unique sequence (a)	3.88×10^8	3.99×10^8
Poly(A)$^+$ mRNA	percentage single-copy sequences	1.57 ± 0.1	1.52 ± 0.1
	nucleotides of expressed unique sequence (a)	5.50×10^7	5.32×10^7

(a) Assuming approximztely 3.5×10^9 nucleotides are represented in the unique rat genome (ref. 157).

(b) All values quoted to 3 sig. fig.; percentages given with mean SE computed from scatter of plateau hybridisation values.

significant differences within each cell-type, in the complexities of nuclear as opposed to polysomal poly(A)$^+$ RNA. These imply that post-transcriptional processing is responsible for the elimination of about 85 percent of the sequence complexity from the precursors of polyadenylated mRNA, although the extent to which this represents qualitative selection amongst different potential mRNAs, rather than the mere removal of non-coding sequences from primary transcripts, is left unanswered. The relative complexities of nuclear and polysomal RNA reported here are similar to those found generally in eukaryotic cells (see Table 1).

The precise values computed here for the complexities of liver and hepatoma hnRNA populations are of roughly the same order of magnitude as those reported by other investigators for nuclear RNA in eukaryotic cells (136,143,161,201,162,165,198, 173,178) generally. In the particular case of rat liver hnRNA complexity, estimates in the literature vary considerably, and the data presented here accord with those of Chikaraishi et al (136) rather than with those of some other investigators (145,148) who reported significantly lower values. One explanation for this may be in the use of guanidinium thiocyanate to prepare undegraded samples of nuclear RNA, enabling a significantly higher proportion of the total complexity to be recovered in the polyadenylated fraction. The use of earlier methods to prepare nuclear RNA, particularly those involving the use of DNase I, which in this laboratory has been found to be variably but unavoidably contaminated with ribonuclease activity, may result in spuriously low estimates of complexity. When results were compared using poly(A)$^+$ nuclear RNA prepared by both methods, significantly lower saturation hybridisation levels were observed for single-copy rat DNA hybridised by DNase I-prepared nuclear RNA than by guanidinium thiocyanate-prepared RNA (approximately 7 percent, as opposed to 10 percent). This was also reflected in the proportion of the RNA preparation which bound to oligo-(dT)-cellulose (2-4 percent as compared with 8-10 percent). Saturation estimates of hnRNA complexity

obtained with total rather than poly(A)$^+$ preparations must also be regarded as unreliable, since the attainable effective $R_o t$ values and RNA/DNA excesses are an order of magnitude less. The use of oligo(dT)-cellulose fractionation does of course exclude sequences in the strictly non-polyadenylated compartment. This may contribute materially to the complexity estimates if, indeed, this compartment does comprise a set of distinct sequences of high complexity, as suggested by results in other systems, particularly at the level of polysomal RNA. Where poly(A)$^+$ and total nuclear RNA complexities have been compared simultaneously (165,183) or even, in some cases, by different investigators (compare Ref.161 with 160,162) they have generally not been very significantly different, however. The similarity of the numerical estimate obtained here with that of Chikaraishi (136) suggests that the non-polyadenylated compartment is not of significant complexity in liver hnRNA.

Estimates of the complexity of liver and HTC cell polysomal poly(A)$^+$ RNA derived here are again similar to those obtained generally for mammalian somatic tissues or tissue-culture lines derived from them (Table 1, refs. 144,145,146,152,153,158, 143, 201,135,165,172) in particular) using a similar method. They accord particularly with other estimates of the complexity of rat liver poly(A)$^+$ mRNA (144,145) or total mRNA (147), again indicating that the uniquely poly(A)$^-$ mRNAs are of low complexity in this tissue. The clear-cut difference in nuclear and polysomal complexity, and the satisfactory plateau levels observed, indicate that cross contamination of polysomal with nuclear RNA is below the limits of detection. Since plateau levels of polysomal RNA-driven hybridisation are maintained at $R_o t$ values considerably beyond the main transition in the nuclear RNA driven reaction, the contamination level must be very low, probably less than 3 percent.

4.1.3. Saturation hybridisation using fractionated single-copy DNA

Direct confirmation of the relative complexities of nuclear

and polysomal poly(A)$^{+}$ RNA in rat liver was obtained by the preparation of a probe enriched for those unique sequences expressed in nuclear poly(A)$^{+}$ RNA. ^{3}H-labelled single-copy rat DNA was hybridised with a 100-fold excess of mercurated rat liver poly(A)$^{+}$ nuclear RNA to a $R_o t$ of 30,000 moles.s.l^{-1}, and the hybridising sequences were recovered by thiol-sepharose chromatography. This 'transcribed single-copy sequence' (N-DNA) probe was then hybridised to saturation with mercurated nuclear and polysomal RNA from rat liver, in large RNA-excess, the extent of reaction being assayed by thiol-sepharose chromatography. Approximately 15 percent of the sequences in the N-DNA probe which were capable of reaction with nuclear RNA, were also hybridised by polysomal RNA (Fig. 6). The poor hybridisability of the probe with nuclear RNA is attributable partially to incomplete fractionation in its preparation, and partially to the fact that the ^{3}H- DNA used was already three months old, and detectably shorter than immediately after gap-translation (about 150 nucleotides single-strand length as judged by alkaline sucrose gradient sizing). The result nevertheless confirms the implications of saturation hybridisation using unfractionated single-copy DNA, namely that nuclear poly(A)$^{+}$ RNA is some 6 times more complex than polysomal poly(A)$^{+}$ RNA in these cell-types.

4.2. Sequence components of rat liver and HTC cell nuclear and polysomal poly(A)$^{+}$ RNAs

RNA populations of similar or indistinguishable complexity may contain sequences at radically divergent relative abundances. Sequences abundant in one population may even be absent in the other. Conversely, populations of distinct complexity may show very considerable homology, by weight. In order to extend the analytical comparison of liver and hepatoma nuclear and polysomal RNA beyond the measurement of purely qualitative differences, the kinetics of homologous and heterologous cDNA hybridisations were studied. The results of these studies

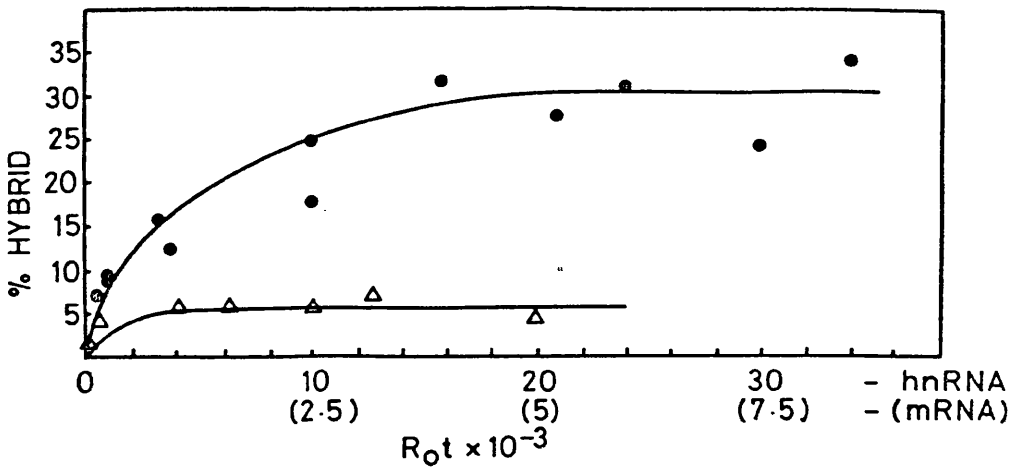


FIGURE 6. Hybridisation of ^{3}H-labelled single-copy rat DNA enriched for transcribed single-copy liver sequences (N-DNA) with mercurated poly(A)$^+$ nuclear (●) and polysomal (△) rat liver RNAs, at 10 mg/ml and 2 mg/ml respectively (RNA excess of 1600-fold and 320-fold respectively). Hybridisations were at 43°, analysed by thiol-sepharose chromatography.

are considered in three separate sections. Firstly, here, the kinetics of homologous reactions are presented, and analysed in terms of their implications for the pattern of relative sequence abundances within each population. Their possible significance regarding quantitative regulation of gene expression both within and between cell-types will be only briefly referred to, since it is discussed in more detail in the following sections. In the next section (4.3), the results of heterologous hybridisations between nuclear and polysomal RNAs and their cDNAs will be considered in terms of quantitative post-transcriptional sequence selection in rat liver. The third set of results consists of heterologous hybridisation kinetics of liver and hepatoma cDNAs with RNAs from the other cell-type, which illustrates (Section 4.4) the nature and extent of quantitative mRNA regulation between the cell-types considered.

4.2.1. Hybridisation kinetics of nuclear cDNAs

Nuclear cDNAs were prepared from liver and HTC cell poly(A)$^+$ nuclear RNA templates, and were hybridised in large template RNA excess, the extent of reaction being assayed at various values of $R_o t$ by S1 nuclease digestability. The results, plotted in Fig. 7, were analysed by least-squares fitting to a rate-equation comprising a minimum number of pseudo-first order components (602). For convenience of comparison the curves have been normalised with respect to the hybridisabilities of the various probes (73 percent and 74 percent for liver nuclear cDNA, 55 percent for hepatoma nuclear cDNA). Table 3 summarises the results of the computer analysis in terms of the hypothetical abundance components of the RNA populations.

The errors in these estimates are a function of the number and accuracy of data points, and it must be stated that the error in each statistic is generally not lower than $\pm$ 20 percent.

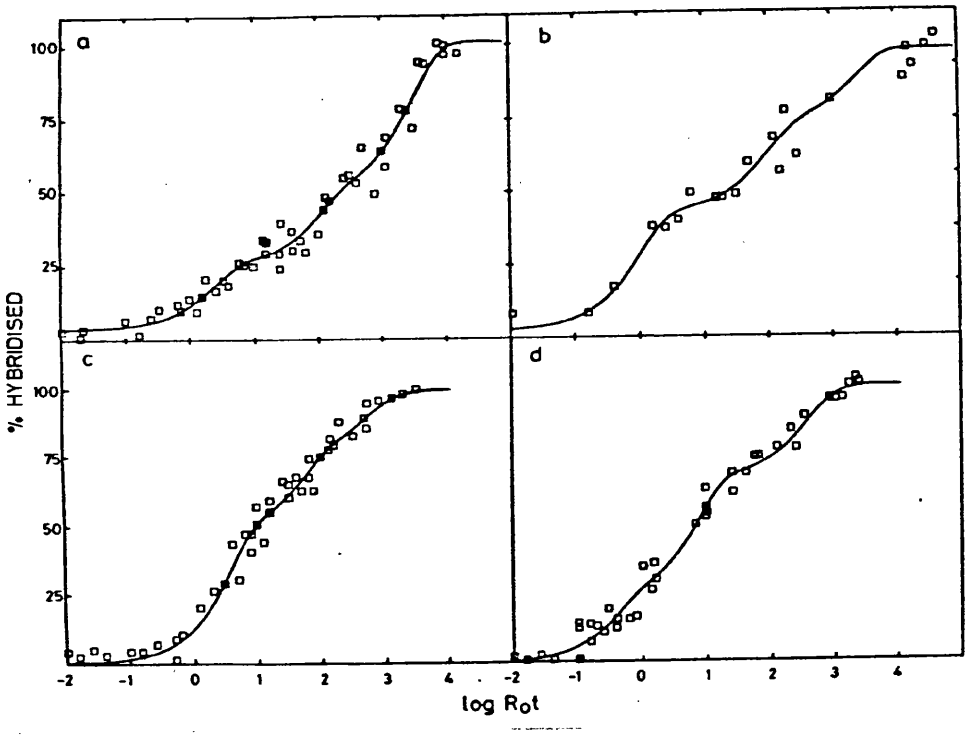


FIGURE 7. Hybridisation kinetics of rat liver and HTC cell nuclear and polysomal cDNAs with their template RNAs. All values normalised for probe hybridisability, with S1 background (2-5 percent) subtracted. Poly(U) (1/2 mg/ml) was added to all reactions where RNA concentration was less than this.

(a) Liver nuclear cDNA hybridised with liver nuclear poly(A)$^+$ RNA at 10 µg/ml, 100 µg/ml and 4 mg/ml, in HB/S at 70° or HB/F at 43°. Analysis was by S1 nuclease. Probe hybridisability = 73 percent or 74 percent.

(b) HTC nuclear cDNA hybridised with HTC nuclear poly(A)$^+$ RNA at 50 µg/ml, or 8 mg/ml in HB/S at 70°. Analysis by S1 nuclease. Probe hybridisability = 55 percent.

(c) Liver polysomal cDNA hybridised with liver polysomal poly(A)$^+$ RNA at 10 µg/ml, 500 µg/ml or 5 mg/ml in HB/F at 43°. Analysis was by S1 nuclease. Probe hybridisability = 85 percent.

(d) HTC polysomal cDNA hybridised with HTC polysomal poly(A)$^+$ RNA at 10 µg/ml, 500 µg/ml or 5 mg/ml in HB/F at 43°. Analysis was by S1 nuclease. Probe hybridisability = 83 percent.

Curve fitting is by computer analysis (602).

TABLE 3. Kinetic estimates of complexity and abundance distributions of poly(A)$^+$ hnRNA and mRNA from rat liver and HTC cells: derived from data of Fig. 7.

RNA	Cell-Type	Kinetic Component (a)	$R_o t_{1/2}$ (b)	Percentage total cDNA (c)	Complexity (d)	No. of sequences (e)	Mass fracti per sequence (f)
hnRNA	Rat liver	I. (Abundant)	1.7	22	1.2×10^5	34	6.5×10^{-3}
		II. (Intermediate)	75	25	6.1×10^6	1700	1.5×10^{-4}
		III. (Rare)	2200	53	3.7×10^8	110,000	4.8×10^{-6}
	HTC Cells	I. (Abundant)	0.71	41	9.5×10^4	27	1.5×10^{-2}
		II. (Intermediate)	67	31	6.8×10^6	1,900	1.6×10^{-4}
		III. (Rare)	1630(g)	28	1.5×10^8	43,000(g)	6.5×10^{-6}
mRNA	Rat liver	I. (Abundant)	0.26	21	1.8×10^4	9	2.3×10^{-2}
		II. (Intermediate)	5.0	45	7.5×10^5	370	1.2×10^{-3}
		III. (Rare)	270	34	3.0×10^7	15,000	2.3×10^{-5}
	HTC Cells	I. (Abundant)	2.3	49	3.6×10^5	180	2.7×10^{-3}
		II. (Intermediate)	39	28	3.6×10^6	1,800	1.6×10^{-4}
		III. (Rare)	390	23	2.9×10^7	14,000	1.6×10^{-5}

./.

All values are quoted to 2 significant figures.

(a) Chosen arbitrarily by curve-fitting (602).

(b) Observed $R_o t_{1/2}$ in moles.s.l^{-1}.

(c) Proportion of reactable probe hybridised in each hypothetical transition.

(d) Computed by calculating corrected $R_o t_{1/2}$ (product of (b) and (c)) and taking as kinetic standard the hybridisation of globin mRNA (1300 nucleotides) with its cDNA (5-6S) which under these conditions was found to exhibit a $R_o t_{1/2}$ = 4×10^{-3}. Because of the many uncertainties regarding the kinetic complexity estimates of hnRNA (outlined in the text), no attempt has been made to correct the values derived for the effects of a short probe and long driver, neither theoretically nor empirically. Such a correction is probably unnecessary for the mRNA-cDNA analysis since the size of reacting species in the globin standard reaction is very similar to the experimental case.

(e) Number of sequences is stated for mRNA, as the number of average mRNA-sized sequences (i.e. of 2000 nucleotides length). For poly(A)$^+$ nuclear RNA the experimental weight-average size (generally about 3500 nucleotides on sucrose gradients) was used in the computation.

(f) Proportion of the total population represented by each member of the given class.

(g) Very unreliable estimate due to insufficient number of data points.

Furthermore, the number of components generated is somewhat arbitrary, since any larger number of components than three will satisfy the data equally well, without a dramatic effect on the total estimated kinetic complexities. Since the analysis for HTC cell nuclear cDNA is based on fewer data points, errors are correspondingly greater, especially in terms of the size and complexity of the rarest component. The division into kinetic components, it must be stressed, is to be regarded as a mathematical device for conceptualising and analysing the relative abundance patterns of the RNA populations. It is not intended to support or refute the hypothesis of discrete abundance classes (142,159,171,195) question which cannot be settled by consideration of heterogeneous cDNA hybridisation kinetics alone. These arguments apply equally to the analysis of polysomal cDNA hybridisation (next section 4.2.2).

The data analysis shown in Table 3 indicates that the poly(A)-adjacent sequences in hnRNA of both liver and hepatoma are at diverse relative abundances, the frequencies of the rarest species being significantly more than 3 orders of magnitude lower than those of the most abundant ones. Although the existence of a further component comprising 'ultra-rare' sequences cannot be ruled out, it is notable that the total kinetic complexities of both populations at least in order of magnitude are comparable with saturation estimates of complexity (Table 2). Provided all poly(A)-adjacent sequences are represented equally in the cDNA,kinetic estimates of RNA complexity should be unaffected by the fact that cDNAs, in general, comprise only a partial transcript of each RNA sequence. This is because the rate of hybridisation of each cDNA species is dependent upon the mass fraction of its complementary messenger in the RNA population. Therefore, the short size of nuclear cDNAs (190 nucleotides weight-average single-strand length, judged on alkaline sucrose gradients, Fig. 2) should not result in an underestimate of

hnRNA complexity. However, Chamberlin et al (603) have shown empirically that hybridisation rate is affected by the ratio of driver and tracer lengths, which means that the comparison with a globin cDNA-mRNA kinetic standard is not strictly justified in this case, and will lead to an erroneous estimate of complexity. Moreover, a considerable proportion of the mass of poly(A)$^+$ nuclear RNA may consist of non-polyadenylated species co-isolated with the true poly (A)$^+$ hnRNA (A. Balmain, personal communication) due to the formation of regions of intermolecular duplex of relatively low complexity, of the type proposed by Davidson and Britten (286) or observed by Kramerov et al (301) and Jelinek et al (295) in mouse and human cells respectively. This would tend to exaggerate the estimates of nuclear RNA complexity, since measured $R_o t$ values would be spuriously inflated by the contribution from non-adenylated sequences unrepresented in the probe.

The possibility of internal initiation of reverse transcription on oligo(A) sequences must also be considered. If the distribution of such sequences, between different RNA species showed considerable variation, the relative composition of the probe would be different from that of the template, and the kinetic analysis would no longer be strictly applicable.

The short size of nuclear cDNAs may be due, in part, to the formation of intra- or intermolecular RNA duplexes which block the progress of the polymerase. Once more, if the distribution of such tracts with respect to poly(A) is not random (i.e. if duplex regions are found very close to poly(A) in some molecules but not in others), the composition of the probe would again be biassed.

For all these reasons, the kinetic estimates of hnRNA complexity must be regarded as very approximate, and their apparent similarity with saturation estimates may not therefore be construed as necessarily supporting Kiper's explanation

(188) for the discrepancy observed at the level of polysomal complexity measurements by the two methods. Despite uncertainties in the calculated values shown in Table 3, it is evident that the relative abundance patterns of poly(A)-adjacent sequences in hnRNA of the two cell-types considered are not radically different, although the high complexity class in HTC cell hnRNA appears to account for a somewhat lower proportion of the total mass than in rat liver. Put in this way, however, this observation is somewhat misleading, since the 'lower' complexity classes have a combined complexity which is a substantial fraction of the total complexity expressed in polysomal mRNA (see Table 2 and Section 4.2.2). If significant, this difference in hnRNA abundance patterns would be compatible with the idea that some sequences might be at increased levels in HTC cells due to regulation at the level of transcription or early post-transcriptional processing. There is no evidence, however, from these data that such sequences are represented in mRNA, a question which can only be resolved by directly studying the composition of hnRNA using polysomal cDNA probes (Sections 4.3 and 4.4.4).

4.2.2. Hybridisation kinetics of polysomal cDNAs

The kinetics of hybridisation of rat liver and HTC cell polysomal cDNAs with their templates, carried out and assayed as for nuclear cDNA hybridisations, are shown in Fig. 7c-d, with a similar computer analysis presented in Table 3. The distributions of mRNA abundances are also very wide, spanning some three orders of magnitude in liver, somewhat less in HTC cells. Noteworthy is the absence of a detectable class of ultra-high abundance mRNAs in HTC cells, where a three component solution was found to provide only a marginally better fit to the data than a two component solution. Total kinetic complexities of the two cell-types messenger populations are once again very similar, although rather lower than those estimated from saturation hybridisation of single-copy DNA (Table 2). This may reflect inherent

errors in either of the methods, which have already been discussed (Sections 2.2.2, 2.2.3 and 4.1.2), and tends to support the arguments of Kiper (1^{88}). The relative abundance pattern of liver polysomal mRNAs deduced here (Table 3) is in general agreement with that reported by other authors (144,145), and resembles the pattern observed in other specialised mammalian tissues (152,153,154,159). The pattern of relative mRNA abundances in HTC cells is correspondingly similar to those of some other mammalian tissue-culture lines, whether 'normal' or neoplastic, including HeLa cells (141), 3T6 cells (141), embryonal carcinoma-derived mouse myoblasts (170), L cells (171) and Novikoff hepatoma (146). Some cell-lines do appear to contain super-abundant mRNAs, however, as judged by some investigators, including HeLa cells (142) and Ehrlich carcinoma (168), so it cannot be concluded that the absence of such a class is a necessary concomitant of growth in tissue-culture. Once again, the arbitrary nature of the division into kinetic components is stressed. The significance of the apparent difference between liver and HTC cell mRNA populations can only be adequately assessed by recourse to heterologous hybridisations (see Section 4.4).

4.3. Homologies between nuclear and polysomal poly(A)$^+$ RNA in rat liver

4.3.1. Heterologous hybridisation of nuclear cDNA: evidence for qualitative post-transcriptional selection

The results of single-copy DNA saturation hybridisation (Sections 4.1.2 and 4.1.3) and homologous cDNA hybridisation kinetics (Section 4.2, Table 3) demonstrate that a large proportion of the sequence complexity expressed in poly(A)$^+$ nuclear RNA in both liver and HTC cells is qualitatively absent from polysomal poly(A)$^+$ RNA. This does not, however, indicate what proportion of the mass of steady-state hnRNA is unrepresented in polysomal mRNA, nor does it distinguish post-transcriptional selection between different potential

mRNAs from generalised processing. Since the primary transcripts of most mammalian genes so far studied are substantially more complex than their corresponding mature mRNAs, the only doubt which remains is whether such processing can account for the entire additional complexity of hnRNA over mRNA. Since the regions adjacent to poly(A) in presumptive primary transcripts of mammalian genes are conserved into mature mRNA (see Section 2.3.3), a necessary, though not sufficient condition for qualitative selection between different potential mRNAs, would be the detection of a class of poly(A)-adjacent nucleus-restricted sequences of high complexity. Such sequences have been reported in other tissues (346,176,375,149,363), therefore it was of interest to investigate this phenomenon in the cell-types studied here. Rat liver nuclear cDNA was therefore hybridised with a large excess of liver polysomal poly(A)$^+$ RNA, and the extent of reaction at various values of $R_o t$ determined by resistance to S1 nuclease (Fig. 8a). The level of hybridisation was observed to reach a plateau by a $R_o t$ of 2,000 moles.s.l^{-1}, as does the hybridisation of liver polysomal cDNA with its template (see Fig. 7c), such that some 30 percent of the sequences represented in the probe (by weight) were not hybridised by polysomal mRNA. This indicates that qualitative post-transcriptional selection in rat liver applies to poly(A)-adjacent sequences. This also confirms the low level of contamination of polysomal RNA with nuclear RNA, which must be below the limits of detection, supporting the conclusion reached from inspection of single copy DNA saturation levels (Section 4.1.2).

The complexity of the nucleus-restricted sequences cannot be simply determined, since their preparative fractionation, even when carried out at substantial RNA/cDNA ratios (e.g. 100-fold), results in co-isolation of sequences whose post-transcriptional regulation appears to be only quantitative (see Section 4.3.2). Despite the detection in a probe thus fractionated, of sequences hybridising at relatively low $R_o t$

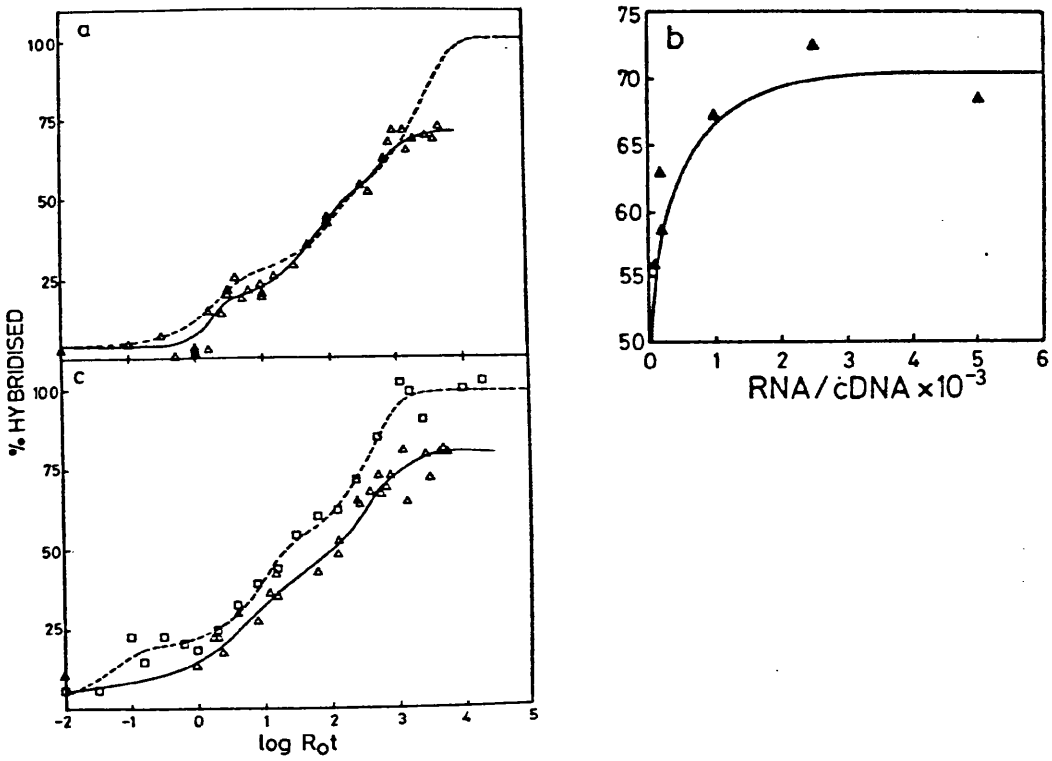


FIGURE 8. Cross-hybridisation of total and fractionated liver nuclear cDNAs with liver poly(A)$^+$ polysomal RNA. Poly(U) added to reactions as previously.

(a) Hybridisation of total nuclear cDNA with polysomal RNA at 25 µg/ml, 330 µg/ml or 5 mg/ml (20,000-fold RNA excess at highest $R_o t$ values) in HB/S at 70° or HB/F at 43°. Normalised for probe hybridisability (73 percent) with S1 background subtracted. Homologous reaction (Fig. 7a) shown as dashed curve.

(b) Hybridisation of total nuclear cDNA with polysomal RNA at 5 mg/ml to $R_o t$ = 2000 moles.s.l^{-1}, at various RNA:cDNA ratios. Each point plotted is the mean of at least two replicate hybridisations, analysed by S1 nuclease.

(c) Hybridisation of abundant nuclear cDNA with nuclear poly(A)$^+$ RNA (□ , dashed curve) and polysomal poly(A)$^+$ RNA (△ , solid curve) at 50 µg/ml and 5 mg/ml (10,000 fold RNA excess at highest $R_o t$ values in HB/F at 43°. Normalised for probe hybridisability (60 percent) with S1 background subtracted.

values, when re-hybridised with poly(A)$^+$ nuclear RNA, its overall composition does show some enrichment for rarer sequences (data not shown) which is consistent with the nucleus-restricted sequences adjacent to poly(A) being of high complexity. Even this does not, however, prove that they correspond with potential mRNA sequences, since the possibility that they consist largely of oligo(A) rather than poly(A)-adjacent sequences cannot be excluded; nor can the existence of intron-like polyadenylated sequences, 3' to the poly(A)-addition site of mature mRNAs, or even non-specifically polyadenylated hnRNA species, be entirely ruled out. A similar group of sequences is found in HTC cell nuclear RNA, as shown by the saturation level of the corresponding heterologous hybridisation (shown in Fig. 26).

4.3.2. Heterologous hybridisation kinetics using total and fractionated nuclear and polysomal cDNAs: evidence for quantitative post-transcriptional selection

The kinetics of the cross-hybridisation shown in Fig. 8a indicate that, on average, the frequency of the sequences adjacent to poly(A) in rat liver hnRNA, which are shared with polysomal mRNA, are at similar or slightly elevated levels on the polysomes. Since in poly(A)$^+$ nuclear RNA they are diluted out by nucleus-confined sequences both proximal and distal to poly(A), they are likely to be diluted out by other sequences in poly(A)$^+$ polysomal RNA. Unless these were to consist of poly(A)-adjacent sequences in mature mRNA which were not detectable next to poly(A) in hnRNA (arising, for example, by cytoplasmic splicing and/or polyadenylation for which there is no evidence in any studies of individual genes or populations), the implication would be that they were sequences at high abundance on the polysomes, but much rarer in nuclear RNA. The curve shown in Fig. 8a is indeed strikingly similar to the curve of Fig. 7c, but without the super-abundant component. The implication that the bulk of mRNA sequences occupy a more restricted frequency range

in hnRNA than on the polysomes is consistent with the findings of other investigators in various systems (344, 359,360,224,231,333,361).

A far more direct demonstration of this is provided by the converse heterologous hybridisation, namely of polysomal cDNA by poly(A)$^+$ nuclear RNA (Fig. 9a). This shows a marked rate decrease compared with the homologous hybridisation with template RNA (Fig. 7c) of, on average, between one and two orders of magnitude. As represented crudely by the parameter $R_o t_{1/2}$, the rate of the heterologous reaction is more than a 100-fold slower. Most importantly,the reaction takes place over a much more restricted range of $R_o t$ values (about 3 rather than 4-5 decades). When the homologous and heterologous reactions of a fractionated cDNA probe, enriched for abundant polysomal sequences, are compared (Fig. 9b) it is evident that the disparity in polysomal and nuclear abundance is even further enhanced (some 300-fold as judged by a comparison of $R_o t_{1/2}$ values) and the curves no longer approach one another at high $R_o t$ values. Even allowing for the presence in poly(A)$^+$ nuclear RNA of a considerable mass of nucleus-restricted sequences distal to poly(A) and of non-polyadenylated RNA co-isolated with the polyadenylated fractions, which must account for some of this disparity, it is still evident that the more abundant polysomal sequences are at much lower relative abundance in hnRNA, compared with the bulk of mRNA sequences. These findings are essentially similar to those reported by Sippel et al (224). The high relative abundance of some liver polysomal mRNAs would therefore appear to be the result of positive post-transcriptional selection, at levels beyond those in the pathway of cellular information flow at which the sequence distribution of steady-state hnRNA is determined.

In probing for evidence that relative mRNA abundances are influenced post-transcriptionally, the converse case, namely of sequences relatively abundant in hnRNA but significantly depleted in polysomal mRNA, must also be considered. The

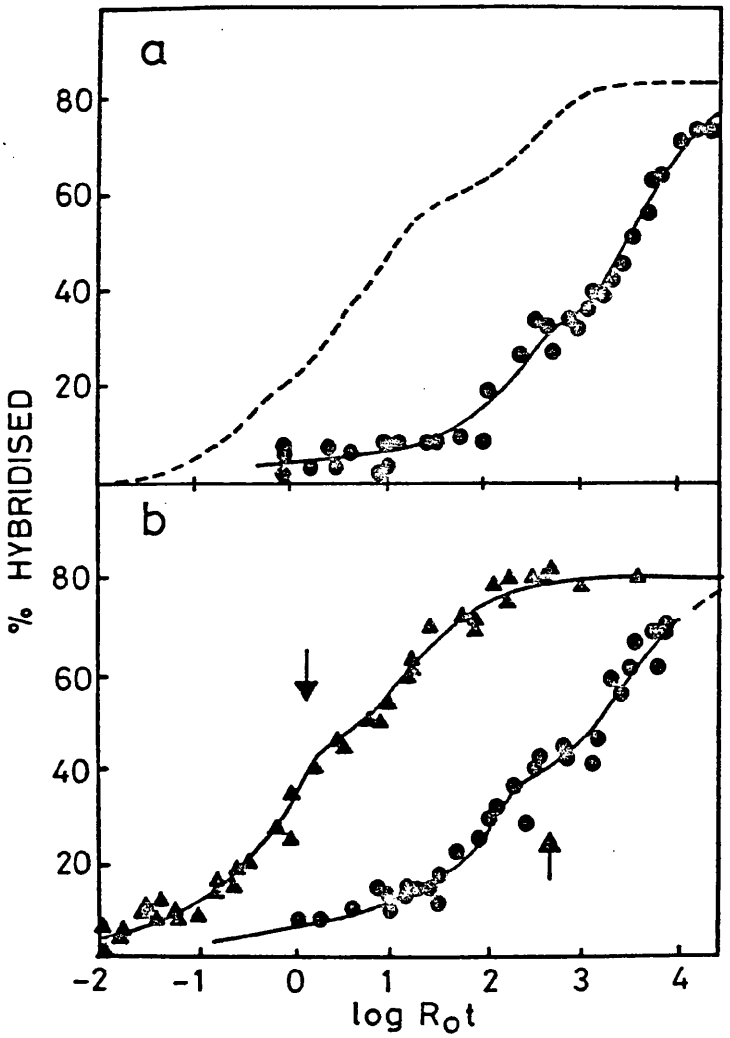


FIGURE 9. Cross-hybridisation of total and fractionated liver polysomal cDNAs with liver poly(A)$^+$ nuclear RNA. Poly(U) added to reactions as previously.

(a) Hybridisation of total polysomal cDNA with nuclear RNA at 1 mg/ml and 10 mg/ml (18,000-fold RNA excess at highest R$_o$t values) in HB/S at 70°. Homologous reaction (Fig. 7c) shown as dashed curve.

(b) Hybridisation of abundant polysomal cDNA with polysomal poly(A)$^+$ RNA (▲) and nuclear poly(A)$^+$ RNA (●) at 10 µg/ml (polysomal RNA only), 100 µg/ml and 5 mg/ml (7,000-fold RNA excess at highest R$_o$t values), in HB/S at 70° or HB/F at 43°.

data of Fig. 8a indicate that such sequences cannot account for a very large proportion of poly(A)-adjacent hnRNA. A crude estimate of the proportion of poly(A)-adjacent sequences in nuclear RNA which are orders of magnitude rarer in polysomal poly(A)$^+$ RNA may be derived from partial titration of nuclear cDNA with polysomal RNA. Under the high RNA-excess conditions of the reaction shown in Fig. 8a, effectively all cDNA species represented in the heterologous RNA will react. If the RNA/cDNA ratio is progressively decreased, however, sequences which are much rarer in the heterologous RNA driver than in the probe will no longer be in RNA-excess and will no longer react to completion (nor will the kinetics be first-order). The proportion of counts no longer hybridised at lower and lower RNA/cDNA ratios is therefore a measure of the proportion of sequences no longer in substantial RNA excess at any given ratio. The results of such a partial titration are shown in Fig. 8b. The reactions were taken to a $R_o t$ value of 2,000 moles.s.l^{-1}, at which value the reaction in large RNA excess (Fig. 8a) had reached a plateau. When the RNA/cDNA ratio is lowered to 100, some 15 percent of the reactable sequences in the probe no longer hybridise by this $R_o t$ value. This indicates that approximately this proportion of the sequences adjacent to poly(A) in hnRNA are at least one order of magnitude rarer in polysomal mRNA, although they are still detectably present.

As shown above for polysomal cDNA, fractionation into abundant and rare components can amplify differences in relative abundance which are detectable by heterologous hybridisation, and this technique is particularly applicable where the shifts in abundance affect only a modest proportion of the total sequences. An abundant nuclear cDNA probe was therefore prepared, by partial hybridisation with template RNA to $R_o t$ = 100 moles.s.l^{-1}, under which conditions some 35-40 percent of the probe consistently hybridised. The fractionated probe was re-hybridised with nuclear and polysomal poly(A)$^+$ rat liver RNAs, in large RNA excess. The reaction kinetics

(Fig. 8c) show that enrichment for sequences abundant in liver hnRNA (represented crudely by the shift in $R_o t_{1/2}$ for the normalised reaction from 210 to 20 moles.s.l^{-1} is not accompanied by a comparable enrichment for sequences abundant on the polysomes (the $R_o t_{1/2}$ for the polysomal RNA-reactable sequences shifting only from 33 to 23).

This implies that a proportion of relatively abundant nuclear sequences is post-transcriptionally depleted in abundance with respect to the bulk of the mRNA population. In order to establish that this is not the result of a hybridisation artifact, it should be possible to prepare, by successive cDNA probe fractionation, a probe specifically enriched for such sequences. Nuclear cDNA was therefore hybridised with a 100-fold excess of mercurated polysomal poly(A)$^+$ RNA to a $R_o t$ of 1,000 moles.s.l^{-1}, after which the hybridised and unhybridised material (the latter including both nucleus-restricted sequences and those significantly depleted on the polysomes), were separated by thiol-sepharose chromatography. Each fraction was re-hybridised with template RNA to $R_o t$ = 100 moles.s.l^{-1}, in order to recover sequences at high abundance in hnRNA. The two probes thus fractionated should correspond, in theory, with abundant nuclear sequences which are, on the one hand, very rare in (or absent from) polysomal RNA, and on the other hand, at comparable (or elevated) abundance on the polysomes. The back hybridisation of each of these fractionated probes, with poly(A)$^+$ nuclear and polysomal RNA (Fig. 10) confirmed the fractionation. The curve shown in Fig. 10b is an important control for those shown in Fig. 10a, since it indicates that the poor hybridisability of the probe for 'post-transcriptionally depleted sequences' with polysomal RNA is not a simple artifact caused by the successive fractionations.

The data of Figs. 8-10 imply that in rat liver, the post-transcriptional determination of relative mRNA abundance is highly complex, involving not merely the operation of special mechanisms to boost the concentration of the high abundance

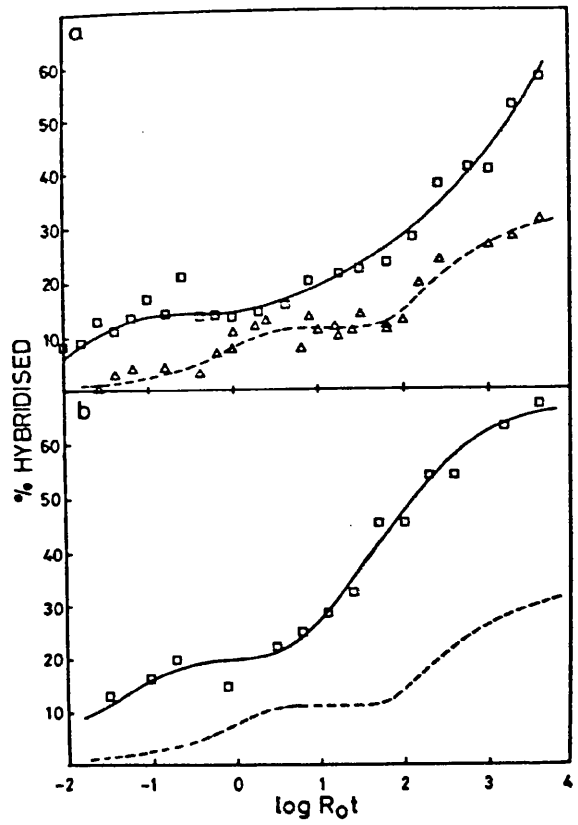


FIGURE 10. Hybridisation of fractionated abundant liver cDNAs, complementary to sequences abundant and rare in polysomal RNA, with poly(A)$^+$ nuclear and polysomal RNAs from rat liver.

(a) Probe for sequences depleted in polysomes, hybridised with poly(A)$^+$ nuclear RNA (□ , solid curve) and poly(A)$^+$ polysomal RNA (△, dashed curve) at 500 µg/ml and 5 mg/ml (6000-fold RNA excess at highest $R_o t$ values) in HB/F at 43°.

(b) Probe for sequences not depleted in polysomes, hybridised with polysomal poly(A)$^+$ RNA at 100 µg/ml and 2 mg/ml (4000 fold excess at highest $R_o t$ values) in HB/F at 43°. Reaction of 'depleted-sequence' probe with the same RNA, from Fig. 10a, shown as dashed curve.

messengers, but also negative selection, applied against some sequences already at a high frequency in steady-state nuclear RNA (whether as a result of high gene dosage, preferential transcription, or some other mechanism). A possible biological rationale for this is discussed in Section 5.2.

4.4. Modulation of mRNA abundances between rat liver and HTC cells

The data already presented indicate that in both liver and hepatoma cells, there is qualitative and quantitative regulation of gene expression, both at the level of transcription (in that only a limited part of the genome appears to be transcribed into nuclear RNA, and that there is a wide disparity in the relative concentrations of different poly(A)-adjacent sequences in steady-state nuclear RNA) and post-transcriptionally (in that only a small proportion of the total hnRNA sequence complexity, including the poly(A)-adjacent complexity, ever reaches the polysomes, and that there are alterations in relative sequence abundance between nuclear and polysomal poly(A)$^+$ RNA, at least in rat liver). The data also show that there are no large qualitative differences between rat liver and hepatoma in transcriptional or post-transcriptional selection mechanisms (i.e. ones that could be detected using heterogeneous single-copy DNA or cDNA technology). The phenotypic differences between the cell-types are therefore likely to be represented by differences in the pattern of relative mRNA abundances. An indication that this might be so has already arisen, from the analysis of homologous cDNA hybridisation kinetics shown in Table 3, indicating a quantitatively altered messenger population in HTC cells, with an absence of super-abundant mRNAs. A more thorough approach to this question is provided by the use of cDNA cross-hybridisation, which can also contribute information on whether the differences in mRNA abundance between the two cell-types is the result of altered transcriptional or post-transcriptional selectivity. As already indicated (Section 2.5), the provision of a satisfactory hybridisation

assay for post-transcriptionally regulated changes in gene expression is necessary for the development of an *in vitro* system for studying the mechanisms by which they are brought about, which is the main purpose of this project.

4.4.1. Hybridisation kinetics of polysomal cDNAs with heterologous cell RNAs

Rat liver and HTC cell polysomal cDNAs were cross-hybridised with the poly(A)$^+$ polysomal RNAs from the heterologous cell-types, under conditions exactly analogous to those of homologous hybridisations described previously (Section 4.2.2), the two sets of hybridisation and assays being conducted, in fact, in parallel. The kinetics of the heterologous reactions are plotted in Figs. 11a-b. The hybridisation of HTC cell polysomal cDNA by liver RNA (Fig. 11b) proceeds to completion, but at a slightly slower rate than the homologous reaction, indicating a modest depletion in liver cells of some abundant HTC cell mRNAs. This must either involve only a small proportion of abundant hepatoma mRNAs, or, if more generalised, is only very small in magnitude. Conversely, the hybridisation of liver polysomal cDNA by HTC cell RNA (Fig. 11a) reveals a radical shift in abundance. By the highest $R_o t$ values reached, the reaction is incomplete, having attained no clear plateau level. The shift in kinetics, of rather more than one decade of $R_o t$, implies that a significant proportion of liver mRNAs is order(s) of magnitude rarer in HTC cells; indeed, some may be absent altogether. A precise quantification of the magnitude and generality of this shift in abundance is not possible, however, for reactions involving whole cDNA populations, since, as has already been implied, it is not possible to distinguish large changes involving few sequences from smaller changes involving many sequences. A clearer picture emerges from the use of fractionated abundant (or rare) cDNAs (see Section 4.4.2) but, ultimately, precise quantification requires the purification, by recombinant DNA technology, of DNAs corresponding to individual messengers (see Section 4.5).

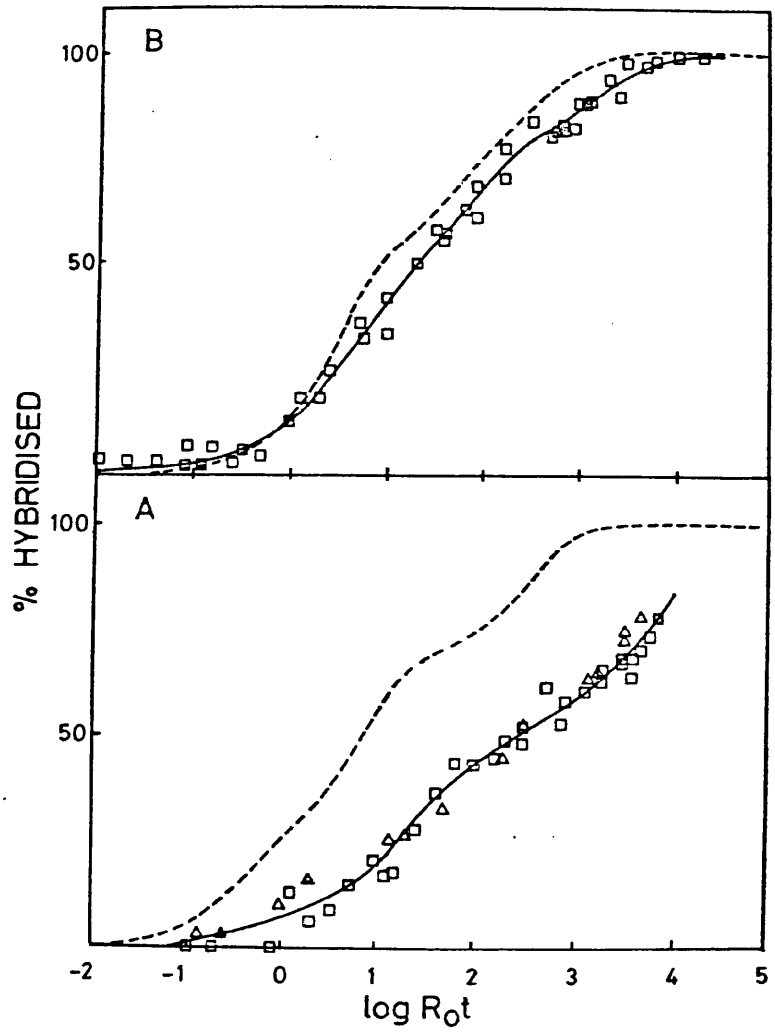


FIGURE 11. Cross-hybridisation of rat liver and HTC cell polysomal cDNAs with heterologous polysomal poly(A)$^+$ RNAs. Poly(U) added to reactions as previously.

(a) Liver cDNA hybridised with two different samples (□ , △) of HTC cell RNA at 20 µg/ml, 500 µg/ml and 5 mg/ml.

(b) HTC cell cDNA hybridised with liver RNA at 20 µg/ml, 500 µg/ml, 5 mg/ml.

Reactions were in 3300-fold RNA excess at highest $R_o t$ values and carried out in HB/F at 43°. Homologous reactions (from Fig. 7c and 7d respectively) shown as dashed curves; all reactions normalised for probe hybridisability, with S1 background subtracted.

The results shown in Figs. 11a-b are not due to trivial differences in the probes or drivers used. As already indicated, liver and HTC cell polysome preparations showed similar, undegraded and reproducible profiles (Fig. 1) and similar percentage binding to oligo(dT)-cellulose columns. Reverse transcription efficiencies were also comparable, and the sizes of the cDNAs, on alkaline sucrose gradients were very similar (Fig. 2) as were their hybridisabilities with template RNAs (Figs. 7c-d). As can be seen from the data plot of Fig. 11a, the same result was obtained, using a second preparation of RNA, and in addition, the use of several different batches of cDNA in both heterologous hybridisation reactions gave very similar results to those shown in Fig. 11.

4.4.2. Hybridisation kinetics of fractionated (abundant) polysomal cDNAs

Polysomal cDNA from each cell-type was fractionated by partial hybridisation with template RNA to $R_o t$ = 10 moles.s.l^{-1}, followed by thiol-sepharose chromatography (see Materials and Methods). Each of the abundant cDNA fractions was re-hybridised with a large excess of the original template RNAs. The kinetics of these reactions, shown in Fig. 12 (the data in Fig. 12a having already been presented in Fig. 9b), indicate the degree of enrichment for abundant sequences in each probe. Also shown are the kinetics of hybridisation with the heterologous cell polysomal poly(A)$^+$ RNAs, which indicate the extent to which abundant messengers are depleted in the polysomes of the heterologous cell. By reference to Fig. 11, they also serve to clarify the overall nature of shifts in relative mRNA abundance between liver and hepatoma.

Enrichment for abundant sequences is demonstrated in each case by the decreased kinetic span of the homologous reactions (effectively complete within 3 rather than 4 or 5 decades of

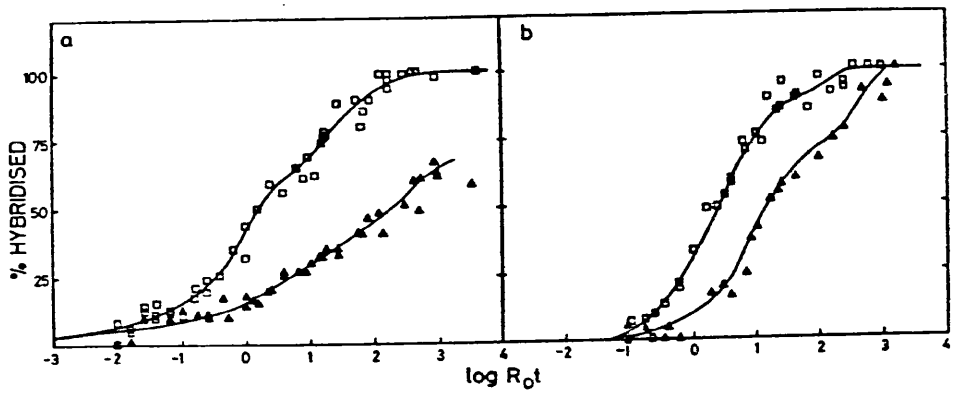


FIGURE 12. Hybridisation of fractionated abundant rat liver and HTC cell polysomal cDNAs with homologous (□) and heterologous (▲) polysomal poly(A)$^+$ RNAs.

(a) Abundant liver cDNA, (b) abundant HTC cell cDNA, hybridised with 20 µg/ml, 500 µg/ml and 5 mg/ml of each RNA (3300-fold RNA excess at highest $R_o t$ values) in HB/F at 43°. All reactions have been normalised for probe hybridisability: (a) 81 percent and 67 percent (2 samples) (b) 60 percent. Poly(U) was included in reactions as previously. S1 backgrounds were subtracted as previously.

$R_o t$, comparing Fig. 12 with Figs. 7c-d), and the lowered $R_o t_{1/2}$ values, indicating greater than 6-fold enrichment in either case. The kinetics of hybridisation of abundant liver cDNA with homologous and heterologous cell mRNAs (Fig. 12a) show a kinetic shift of about two orders of magnitude. The heterologous reaction again shows no clear plateau level, and by the highest $R_o t$ values reached, has only proceeded to about 60 percent of the level seen with liver RNA. The disparity between the kinetics of homologous and heterologous reactions observed using total cDNA (Fig. 11a) has been considerably amplified by the use of a probe enriched for abundant sequences (Fig. 12a). It would therefore appear that the bulk of this disparity noted in Fig. 11a is due to the loss (or massive depletion) of high abundance liver messengers in HTC cells, rather than the loss of a larger number of liver mRNAs at lower abundance. Similarly, the use of a probe enriched for abundant HTC cell mRNA sequences, in the converse set of hybridisations (Fig. 12b), showing, on average, a five-fold depletion of these in liver polysomes as judged by $R_o t_{1/2}$ values, implies that the quite similar kinetics observed between homologous and heterologous reactions of unfractionated HTC cell polysomal cDNA (Fig. 11b) are due to subtle shifts of abundance in both directions.

As already stated, even the use of fractionated probes cannot permit a precise quantification of abundance shifts. However, it is possible to calculate the extreme 'real' situations with which the data are compatible. This is necessarily an approximate calculation, involving several simplifying assumptions. The results of such an exercise are shown in Table 4. The resolution of the homologous curves of Fig. 12 into hypothetical components, upon which the calculation is based, was carried out assuming that the hypothetical computer analysis shown in Table 3 corresponds to the actual pattern of mRNA abundances. The extreme situations presented are those with shifts in abundance involving, on the one hand,

TABLE 4. Extremes of magnitude and generality of abundance shifts between liver and hepatoma inferred from cDNA cross hybridisation. Values presented are hypothetical extreme estimates compatible with the hybridisation data shown in Fig. 12, derived by splitting the kinetic curves into arbitrary components according to the analysis shown in Table 3, and considering separately the different possible fates of the different sequence sets, which together could generate data similar to those observed.

mRNA set considered	Type of Abundance Shift	Number of mRNAs affected	Mass fraction per mRNA in rat liver	Mass fraction per mRNA in HTC cells	Relative Abundance Ratio (c)
Sequences abundant in rat liver but rarer in HTC cells (a)	Most Specific	10	1×10^{-2}	$< 5 \times 10^{-6}$	> 2000
	Most General	400	1×10^{-3}	1×10^{-5}	100
Sequences abundant in HTC cells but rarer in liver (b)	Most Specific	20	3×10^{-5}	7×10^{-3}	200
	Most General	400	2×10^{-4}	1×10^{-3}	5

(a) Relating to data shown in Fig. 12a.

(b) Relating to data shown in Fig. 12b.

(c) Ratio of mass fractions in the two cell types (homologous/heterologous).

the greatest number of different mRNA species (most general solutions), and on the other hand, the greatest disparity in individual mRNA concentrations between the cell-types (most specific solutions). Stated verbally, it is concluded that the data of Fig. 12a could represent anything from the effective absence in HTC cell RNA of a handful of superabundant liver mRNAs, to the uniform depletion of several hundred messengers, by a factor of about 100. Similarly, the data of Fig. 12b could be interpreted as 100-fold drop in concentration of just a dozen or so mRNAs, rather than a generalised five-fold depletion of many hundreds of messengers.

4.4.3. Hybridisation kinetics of nuclear cDNAs with heterologous cell hnRNAs

The foregoing data, which indicate the modulation of mRNA abundances between liver and hepatoma, do not imply at which level in the cell the changes are brought about. One approach to this question is the use of similar technology to compare the relative compositions of the populations of poly(A)-adjacent sequences in steady-state hnRNA in the two cell-types, which offers insight into the likely relative contributions of transcriptional and post-transcriptional regulation.

Fig. 13 shows the kinetics of hybridisation of rat liver and HTC cell nuclear cDNAs with excess poly(A)$^+$ nuclear RNA from the heterologous cell-type. By contrast with the equivalent heterologous hybridisations using polysomal cDNAs and RNAs (Fig. 11), both reactions proceed effectively to completion, and only a modest kinetic shift is seen, in either case, with respect to the rate of homologous hybridisation. The largely unidirectional, and radical shift in relative mRNA abundances between the cell-types at the level of polysomal RNA is not, therefore, reflected in an equivalent shift at the level of nuclear RNA. This suggests that the abundance modulation may have been brought about post-transcriptionally.

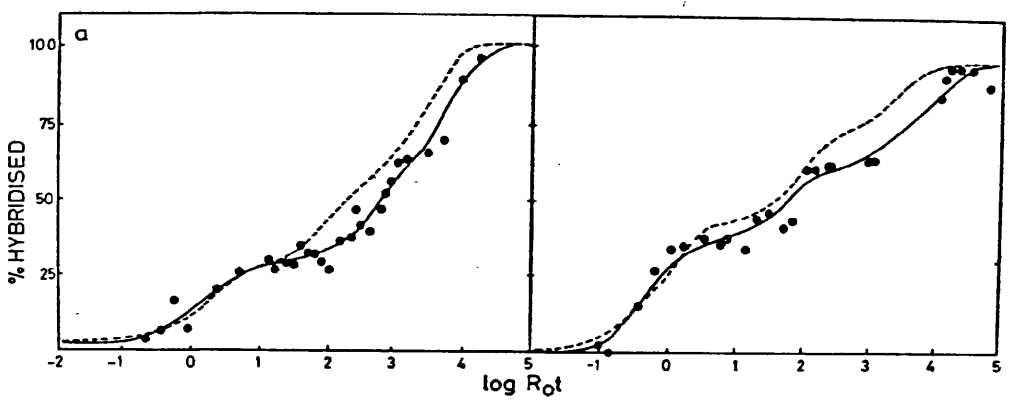


FIGURE 13. Cross hybridisation of rat liver and HTC nuclear cDNAs with heterologous nuclear poly(A)$^+$ RNAs. Poly(U) was included in reactions as previously.

(a) Liver nuclear cDNA hybridised with HTC nuclear RNA at 50 µg/ml or 8 mg/ml.

(b) HTC nuclear cDNA hybridised with liver nuclear RNA at 50 µg/ml or 8 mg/ml.

Reactions were in 8000-fold RNA excess at highest $R_o t$ values, and were carried out in HB/S at 70°. Homologous reactions (from Figs. 7a and 7b respectively) are shown as dashed curves; all reactions normalised for probe hybridisability with S1 background subtracted.

Since, however, a considerable proportion of nuclear poly(A)-adjacent sequences are either absent from, or much rarer on the polysomes (Section 4.3), and since abundant polysomal sequences are not represented at a high level in nuclear RNA (Fig. 9), the results shown in Fig. 13 do not necessarily reflect the relative levels in nuclear RNA of those sequences which are modulated at the level of the polysomal mRNA, between the cell-types, differences which could be swamped by changes (or similarities) in other sequences.

4.4.4. Hybridisation kinetics of polysomal cDNAs with homologous and heterologous cell nuclear RNAs: evidence for post-transcriptional abundance modulation

In order to overcome these uncertainties, total and fractionated (abundant) polysomal cDNAs were hybridised with poly(A)$^+$ nuclear RNA from both cell-types (Fig. 14). A striking observation is that sequences which are at dramatically different concentrations on the polysomes of liver and hepatoma cells are at very similar concentrations in their nuclei. Total polysomal cDNA from rat liver was hybridised with very similar kinetics by nuclear RNA from either source (Fig. 14a). Hepatoma polysomal cDNA was hybridised by liver nuclear RNA with these same kinetics, despite the absence of a superabundant mRNA component (Fig. 14a), a further indication that the abundance of these liver mRNAs may have been generated post-transcriptionally. Most notable of all is the fact that the fractionated (abundant) liver polysomal cDNA probe, which was previously shown to hybridise 100 times more slowly with HTC cell than with liver polysomal RNA (Fig. 12a), was hybridised only about 3-fold slower by HTC cell than by liver nuclear RNA (Fig. 14b). Moreover, comparing Figs. 14a and 14b, it is evident that the high abundance liver mRNAs, which were so depleted on hepatoma polysomes, are at a concentration in hepatoma hnRNA comparable with that of the bulk of polysomal sequences in hnRNA generally.

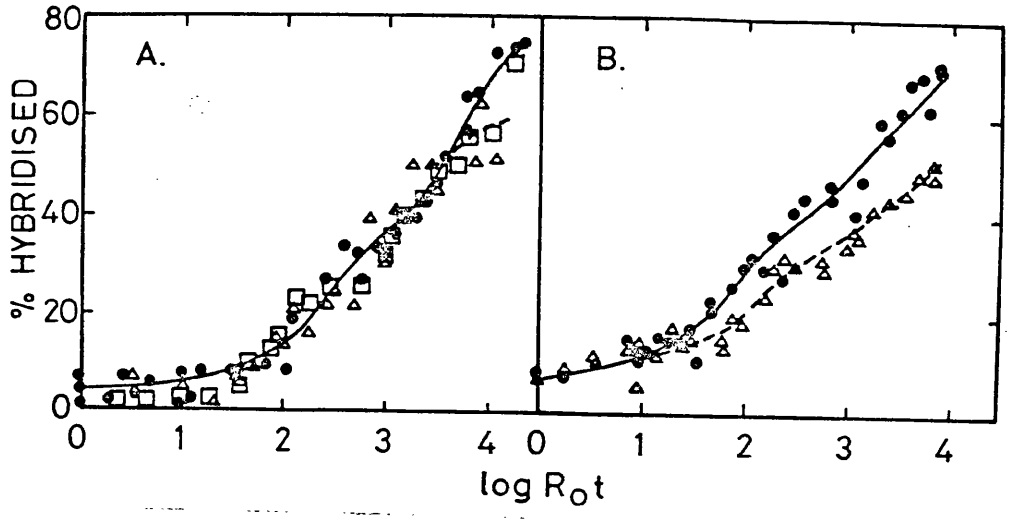


FIGURE 14. Hybridisation of rat liver and HTC cell polysomal cDNAs with homologous and heterologous nuclear poly(A)$^+$ RNAs, (all S1 backgrounds subtracted).

(a) Reactions of total liver polysomal cDNA with liver nuclear RNA (●) and HTC nuclear RNA (△), and of total HTC polysomal cDNA with liver nuclear RNA (□), carried out in HB/S at 70° at RNA concentrations of 1 mg/ml or 10 mg/ml (18,000-fold RNA excess at highest R t values. The former data are replotted from Fig. 9a.

(b) Reactions of abundant liver polysomal cDNA with liver nuclear RNA (●) and HTC nuclear RNA (△) carried out in HB/S at 70°, at RNA concentrations of 100 µg/ml and 5 mg/ml (7000 fold RNA-excess at highest $R_o t$ values. Poly (U) was included as previously. The former data are replotted from Fig. 9b.

A definitive interpretation of these experiments is not possible, due to the difficulty in driving the reactions to completion (such an appreciable proportion of the probe being very rare in the driver). It cannot therefore be excluded that some of the liver sequences depleted in HTC cells are regulated transcriptionally, although it should be borne in mind that since the composition of steady-state poly(A)$^+$ nuclear RNA may itself be partly determined by the selectivity of post-transcriptional processing, the experimental strategy used here may well underestimate the importance of post-transcriptional regulation. A minimum estimate of this contribution may be derived, once again, by considering extreme situations with which the data are compatible. Thus, if the highest hybridisation level reached in the reaction of abundant liver polysomal cDNA with HTC cell nuclear RNA (dashed curve in Fig. 14b), namely about 65 percent of the total reactable sequences, were in fact a plateau hybridisation level, and if the sequences unrepresented in HTC cell hnRNA were in fact the most abundant ones in liver mRNA (i.e. were regulated transcriptionally), then some two-thirds of the total abundant liver mRNAs would still be at comparable abundances in hnRNA in the two cell-types. Juxtaposing this extreme postulate, with one extreme case hypothesised in Table 4 (the 'most specific' solution), then it is just conceivable that a purely transcriptional regulation is responsible for the observed kinetics of hybridisation. However, this extreme view is incompatible with translation data (considered in Section 4.4.6) and with the results of preliminary studies using cloned cDNAs (see Section 4.5), and can therefore be all but eliminated.

4.4.5. Hybridisation kinetics of polysomal cDNAs with liver mRNA sub-fractions

Amongst the most highly abundant mRNAs in rat liver are those which encode the major secretory proteins, whose synthesis is

one of the major physiological functions of the mammalian liver. These include the mRNA for serum albumin, which has been reported as constituting up to 10 percent of the total liver mRNA population (361), and is located almost exclusively on membrane-bound polysomes (442). A pertinent question is therefore whether the observed depletion of high abundance liver mRNAs in HTC cell polysomes could be the indirect result of a lesion in recruitment into membrane-bound polysomes. The mRNA for at least one serum protein, $\alpha 2_u$ globulin, has been shown to be preferentially recruited into membrane-bound polysomes (relative to free polysomes and untranslated mRNPs) following growth hormone administration (604). If the mechanism of this recruitment were permanently switched off in HTC cells, surplus untranslated messengers for serum proteins might be rapidly degraded, causing a large drop in their abundance. Their low frequency on the polysomes might even be a primary consequence of translational control. If either were the case, the liver mRNAs depleted in HTC cells would be expected to be confined to those predominantly in the membrane-bound compartment. In order to investigate this possibility, cellular fractions of liver mRNA were prepared, and were hybridised in excess with liver and HTC cell polysomal cDNAs.

A number of technical difficulties were encountered in the preparation, with some bearing on the results. Undegraded free polysomal mRNA was routinely prepared (as evidenced by polysome profiles and oligo(dT)-cellulose-binding), although it may have been contaminated, to an undetermined extent, with membrane-bound material detached during isolation. The high salt conditions used should have minimised bound mRNA degradation (605). However, membrane-bound poly$(A)^+$ RNA was obtained in much lower than expected yields, was considerably degraded by all the above criteria, and was a very poor template for reverse transcription. Although simple calculation of the likely size of this RNA fraction, from oligo(dT)-cellulose-binding levels observed (it was not determined directly), suggest that it is still longer than the cDNA

probes with which it was hybridised, it may not contain a full (or representative) complement of membrane-bound mRNA sequences. The interpretation of hybridisation reactions driven by it is therefore in some doubt, and only very broad, general inferences will be drawn.

Figs. 15b,a show the reactions of HTC cell total polysomal cDNA, and abundant rat liver polysomal cDNA, respectively, hybridised with liver free polysomal poly(A)$^+$ RNA. Comparing the kinetics of these reactions with those of the hybridisations of the same probes driven by total liver polysomal poly(A)$^+$ RNA (shown as dashed curves) it is clear that the sequences common to liver and hepatoma, as well as those massively depleted in the hepatoma, are all represented in the free polysomal fraction (since complete hybridisation is seen in both cases). Whereas, however, the sequences present in the two cell-types at similar abundance, as represented by HTC cell polysomal cDNA (Fig. 15b), are very slightly enriched in the free polysomal fraction of liver mRNA, the sequences which are most dramatically regulated in abundance between the cell-types, as represented by the abundant liver polysomal cDNA (Fig. 15a) are somewhat depleted, about 3-fold on average, in the free mRNA. Together, these observations indicate that some, but not all of the mass of sequences depleted in HTC cells, is to be found preferentially on membrane-bound polysomes, and hence encodes secretory proteins. Unless, however, the free polysomal mRNA fraction is massively contaminated with membrane-bound material, a significant proportion of the abundant liver mRNAs under post-transcriptional regulation are likely to code for proteins synthesised on free polysomes, which perform intracellular functions.

The hybridisation of abundant liver polysomal cDNA, driven by liver membrane-bound poly(A)$^+$ RNA (Fig. 15c) does not proceed to completion, but reaches an ill-defined plateau representing about 35-40 percent of the reactable sequences,

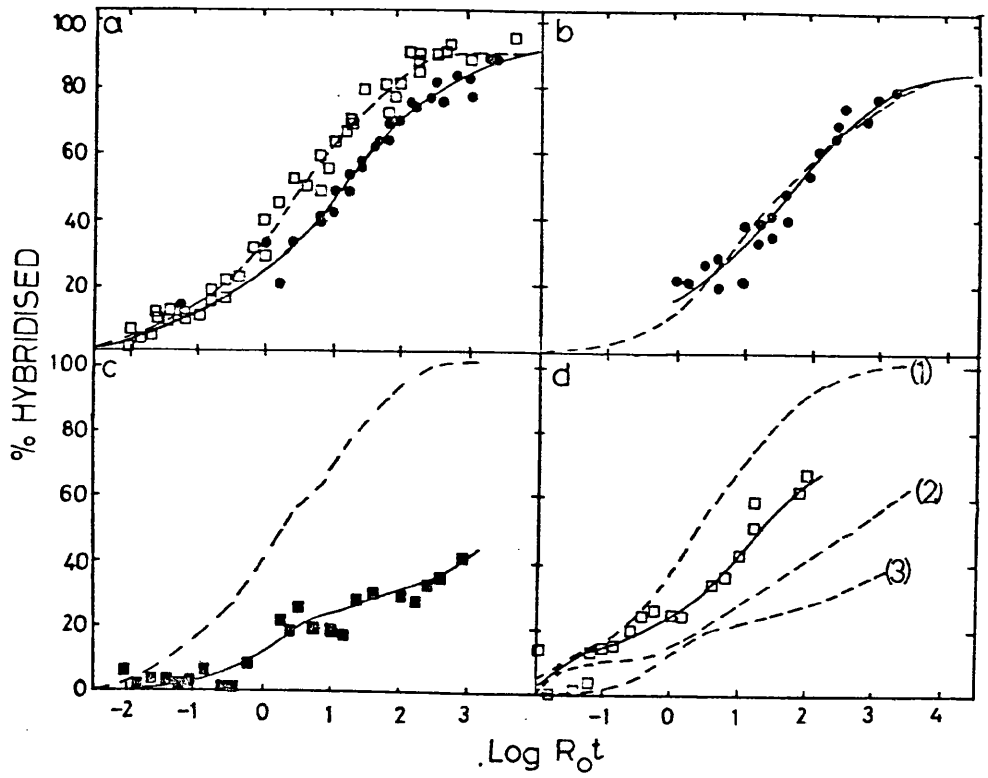


FIGURE 15. Hybridisation of rat liver and HTC cell polysomal cDNAs with free and membrane-bound polysomal poly(A)$^+$ RNAs from rat liver. Poly (U) was included in all reactions, as previously, and curves in (c) and (d) are normalised for probe hybridisability, with S1 background subtracted.

(a) Abundant liver polysomal cDNA hybridised with (●) free liver polysomal RNA at 500 µg/ml and 10 mg/ml in HB/S at 70°. RNA-excess at highest $R_o t$ values was 5000-fold. Homologous reaction with total liver polysomal poly (A)$^+$ RNA (shown as □, dashed curve) reproduced from Fig. 9b.

(b) HTC cell polysomal cDNA hybridised with free liver polysomal RNA at 500 µg/ml and 10 mg/ml in HB/S at 70°. Heterologous reaction of the same probe with total liver polysomal poly(A)$^+$ RNA shown as dashed curve (reproduced from Fig. 11b). RNA excess at highest $R_o t$ values was 5000-fold.

(c) Abundant liver polysomal cDNA hybridised with membrane-bound liver polysomal RNA at 100 µg/ml and 2 mg/ml in HB/F at 43° (and with total liver polysomal RNA(shown as dashed curve), reproduced from Fig. 9b). RNA-excess was 1000-fold at high $R_o t$ values.

(d) Abundant liver polysomal cDNA hybridised with 50:50 mixture of HTC cell polysomal RNA and membrane-bound liver polysomal RNA, each at 100 µg/ml or 1 mg/ml, in HB/F at 43° RNA excess at highest $R_o t$ values was 2400-fold. Reactions of the same probe with (1) total liver polysomal RNA, (2) HTC cell polysomal RNA alone, and (3) membrane-bound liver polysomal RNA alone, are shown as dashed curves.

suggesting that the driver RNA contains only a subset of the poly(A)-adjacent sequences of abundant liver mRNA. Since it was shown previously that the same probe was also hybridised incompletely, and with slowed kinetics, by HTC cell total polysomal poly(A)$^+$ RNA (Fig. 12a) the question arises whether HTC cell polysomes and liver membrane-bound polysomes might contain mutually exclusive subsets of the abundant liver mRNAs (in other words, whether HTC cell mRNA can be crudely depicted as liver mRNA, minus the membrane-bound messengers, effectively the same hypothesis which the data of Figs. 15a-b seemed to disfavour). In order to answer this, the same probe was reacted with a 50/50 mixture of membrane-bound liver mRNA and HTC cell total mRNA (Fig. 15d). The kinetics of this reaction are still an order of magnitude slower than those of the strictly homologous reaction, driven by total liver polysomal RNA, despite the degree of syngergism observed. It can be concluded that although some of the mRNAs depleted in HTC cell polysomes are preferentially located in the membrane-bound compartment in liver, others of them cannot be detected in this fraction. Despite the uncertainties associated with the use of the membrane-bound poly(A)$^+$ RNA as driver, the overall implications are consistent with the earlier findings, namely that mRNAs coding for major secretory and intracellular polypeptide products of the rat liver are both depleted in HTC cell polysomes. The depletion cannot, therefore, be regarded simply as a consequence of a lesion in recruitment of mRNAs into membrane-bound polysomes.

The difficulty encountered in the preparation of membrane-bound polysomal RNA further suggests that the method employed for tissue homogenisation and fractionation may substantially deplete the liver polysomal RNA of membrane-bound sequences. Indeed, the hybridisation kinetics of total liver polysomal cDNA with its template (Fig. 7c) are not wholly consistent with other reports of a single major (secretory) species, albumin, accounting for up to 10 percent of the total mRNA (361,43

(611). Although preferential template specificities in reverse transcription may account for some of this discrepancy, it would also appear that the albumin mRNA-containing (i.e. membrane-bound) mRNA fraction has been selected against. Translation data (Section 4.4.6) and the measured relative abundances of the most prevalent cloned transcripts (Section 4.5) support this view. Gentle homogenisation, which is necessary in order to prepare undegraded total polysomes, followed by low speed centrifugation, probably results in significant preferential loss of membrane-associated mRNA, in the crude pellet. Therefore, the use of cDNA cross-hybridisation (Fig. 11) may have detected, primarily, the depletion in HTC cells of abundant liver mRNAs coding for intracellular proteins. The loss of secretory function, as evidenced elsewhere by the loss of albumin translational template activity in HTC cells (612), may be super-imposed upon this.

4.4.6. Translational template activities of rat liver and HTC cell polysomal poly(A)$^+$ RNA

The use of heterogeneous cDNA hybridisation kinetics to study the composition of mRNA populations is open to the criticism that it may be influenced by the possible non-randomness of reverse transcription efficiencies, or the presence, conceivably, of non-coding polyadenylated RNA species. Data so obtained may therefore be usefully compared and complemented with the results of translational template assays applied to the same mRNA populations. Polysomal poly(A)$^+$ RNAs from both rat liver and HTC cells were therefore translated in a messenger-dependent (heterologous) translation system, the nuclease-treated rabbit reticulocyte lysate. Translation was carried out in the presence of ^{35}S-methionine, at an mRNA concentration which was previously determined to be optimal for both preparations (50 µg/ml), that is, the concentration giving the highest amount of incorporation of label, but which was still in the linear part of the mRNA

dose-response curve (not shown). The products were analysed by one-dimensional (SDS-polyacrylamide) gel electrophoresis, alongside samples of the major cytoplasmic polypeptides of the two cell-types, and a sample of rough ER material from the liver (sample preparation, electrophoresis and visualisation are described under Materials and Methods). The use of a simple one-dimensional gel system for comparing *in vivo* polypeptides and *in vitro* translation products is considered more informative, in this case, than more sophisticated two-dimensional systems since it is unaffected by differences in post-translational modification between the cell-types themselves, or between *in vivo* and *in vitro* environments. Although it suffers the disadvantage of providing poor resolution of distinct polypeptides of similar molecular weights, it nevertheless provides an easily interpreted record of the pattern of differences between liver and hepatoma mRNAs, and their relationship with phenotype.

The gels, stained or autoradiographed as appropriate, are shown in Fig. 16. By scoring for the relative intensities of the various bands, Table 5 has been constructed showing the pattern of abundance shifts between the cell-types, which is broadly in conformity with the results of the hybridisation studies. The translational assay, furthermore,despite the poor resolution,appears to support the view that the depletion of abundant liver mRNAs in HTC cells, whilst not completely generalised, does involve most of the more abundant sequences and would therefore eliminate the 'most specific' solution of Table 4, and in turn, the extreme transcriptional model of regulation hypothesised in Section 4.4.4.

The close correlation between the patterns of *in vivo* polypeptides and *in vitro* translation products suggests a close relationship between mRNA abundances and phenotype, and also supports the view that the depleted messengers include some coding for major intracellular proteins. A further indication of the relative contribution to the abundance shift of mRNAs coding for secretory proteins comes from the observation that just three of the seventeen abundant liver polypeptides

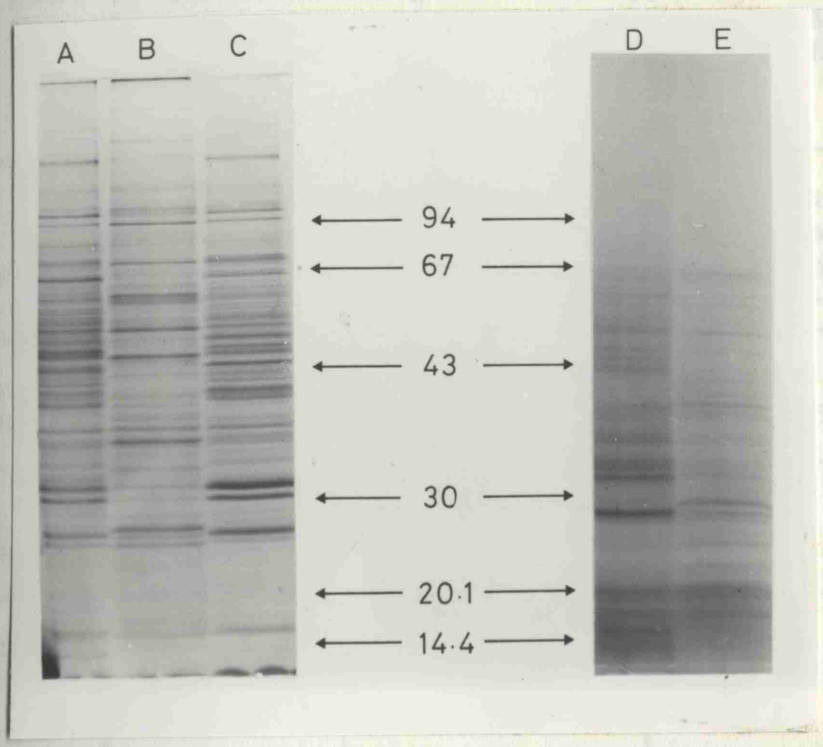


FIGURE 16. Gel electrophoretic comparison of polypeptides and *in vitro* mRNA translation products from rat liver and HTC cells, using SDS-15 percent polyacrylamide gels, visualised by coomassie blue staining (Tracks A, B and C) or autoradiography (D and E). (A) Sample of rat liver proteins enriched in rough ER material; (B) HTC cell post-mitochondrial supernatant; (C) rat liver post-mitochondrial supernatant; (D) ^{35}S-labelled products of *in vitro* translation of rat liver poly(A)$^+$ polysomal RNA, (E) ^{35}S-labelled products of *in vitro* translation of HTC cell poly(A)$^+$ polysomal RNA. Marker Molecular weights in kilodaltons, as shown.

TABLE 5. Relative abundances in the equivalent polypeptide pool of the heterologous cell-type of the major cytoplasmic polypeptides and *in vitro* mRNA translation products from rat liver and HTC cells

Source of polypeptides	No. scored as major bands	No. highly depleted in heterologous cell-type	No. at similar abundance in heterologous cell-type
Rat liver cytoplasm	17	12	5
HTC cell cytoplasm	16	3	13
Liver mRNA translation	18	12	6
HTC mRNA translation	16	5	11

which are severely depleted in HTC cells are enriched in the sample of rough ER material (i.e. are either secretory products, or protein constituents of the secretory machinery). One of these at 66 kilodaltons is at an apparent molecular weight close to that predicted for serum albumin.

Translational assays of the composition of mRNA populations suffer their own special objections. The first of these is that relative template activities in the heterologous system may be significantly different from those which pertain in vivo in the homologous cells, and even if this is not so, they may give a distorted picture of mRNA abundances, since translational initiation rates on different mRNAs may vary considerably (although it could be argued that translation rates are a more valid measurement of 'gene expression' than mRNA abundances). Secondly, the isolation of mRNA by oligo(dT)-cellulose chromatography, if there is even the slightest degree of degradation, will result in selection against translationally functional longer mRNAs. Thirdly, the translational 'elongation-lifetime' of the cell-free system will influence the number of completed chains, which will be unrepresentatively small for the longest mRNAs, or those with translational pause sites. Fourthly, if initiation is significantly more labile than elongation, under given conditions, the proportion of completed chains of very small polypeptides will also be depleted. Despite these reservations, the similar inferences drawn from hybridisation and translation studies, regarding the relative compositions of liver and hepatoma poly(A)$^+$ mRNA populations, strongly reinforce each other.

4.5. Quantitation of abundance differences using recombinant plasmids from a rat liver polysomal cDNA library

4.5.1. Methodological considerations: verification of theoretical predictions concerning plasmid-DNA-driven hybridisation

The most accurate hybridisation methods for quantifying the concentration of a given individual messenger sequence in a

heterogeneous RNA population are by DNA-excess titration, or RNA-excess hybridisation kinetics, using strand-separated DNA (see Sections 2.2.4). Hybridisation of labelled cDNA by excess filter-bound recombinant plasmid DNA is a convenient alternative method, however, which may be made semi-quantitative with relatively little expenditure of material or effort. This depends primarily upon the verification of equation (1) (see Section 2.2.4) and the provision of a suitable kinetic standard. In order to verify the applicability of equation (1);

$$\frac{c}{co} = 1 - e^{-ft}$$

for reactions carried out under the conditions described in Materials and Methods, a pilot study was carried out, using DNA from a recombinant plasmid (pCR1-MαG) containing a mouse α-globin cDNA sequence, titrated against a cDNA probe containing various amounts of α-globin sequences. Equal amounts of the plasmid DNA were bound to a series of filters, and reacted with rat liver polysomal cDNA to which had been added various proportions of mouse reticulocyte cDNA of the same specific activity. High specific activity probes, labelled with ^{32}P-dCTP, were used for the hybridisation, in order to ensure a vast excess of filter-bound reacting DNA. The prediction of equation (1), namely that the proportion of globin-complementary counts bound to the filters, at any given time, should be independent of the absolute concentration of globin sequences in the probe was obeyed, as shown in Fig. 17. Allowing for the 70 percent overall hybridisability of the probe, about 12 percent of the α-globin sequences in the probe annealed to the filters after 55 hours of hybridisation, irrespective of how diluted out they were by non-reacting liver cDNA sequences. (This also assumes that α-globin cDNA represents approximately 50 percent of reticulocyte cDNA.) Although this does not prove that equation (1) holds for all sequences, there is no theoretical reason for assuming that the pseudo-first order rate constant for the reaction should vary dramatically for different sequences hybridised under identical conditions (although smallvariations

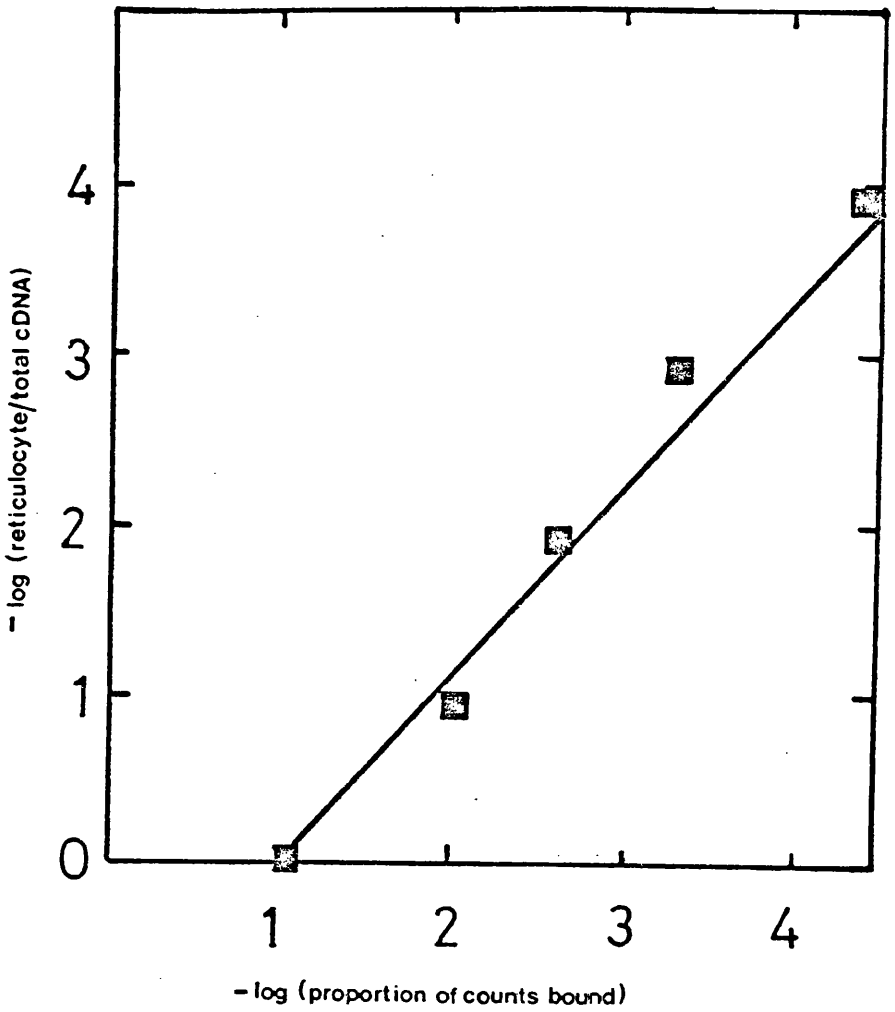


FIGURE 17. Hybridisation kinetics of filter-bound α-globin-complementary plasmid DNA with mouse reticulocyte cDNA. The proportion of counts bound to filters after 55 hours of hybridisation is shown for various mixtures of rat liver polysomal cDNA and mouse reticulocyte cDNA. A log/log plot is used, the slope of which, for a linear relationship, should be +1. The observed linear regression coefficient of x on y, of 0.85, is not significantly different from 1, using a t-test.

due to differences in G-C content might be expected). Moreover, if the amount of reacting DNA per unit area on the filters is held constant for different messenger sequences, the proportion of reacting sequences bound at a particular time during the reaction should be the same (i.e. if 12 percent of the α-globin sequences are hybridised by an α-globin DNA-bearing filter, 12 percent of an unknown mRNA sequence in a given probe should bind to filters bearing the appropriate recombinant plasmid DNA). By this use of α-globin as a kinetic standard, the proportion of total cDNA counts bound by any given recombinant-bearing filter can be converted into a direct measurement of the mass fraction of that particular cDNA sequence in the probe. If, in practice, slightly different amounts of reacting DNA are bound to different filters, varying the parameter 'f' in equation (1) (in effect, the product of mass per unit area and first-order rate constant) a simple correction may be applied, assuming direct proportionality between the mass per unit area of reacting DNA, and the parameter 'f', i.e. that the rate-constant is unaffected by altering the DNA per unit area. Even if this correction is an over-simplification, the data may be regarded as supplying an estimate of abundance, correct to within considerably better than order-of-magnitude accuracy, provided the amount of reacting DNA on the filters varies no more than about 5-fold, and depending, of course, on experimental reproducibility.

4.5.2. Selection and characterisation of recombinant plasmids

Provided hybridisation conditions are the same, the result of the pilot experiment indicates that the method can be used to estimate the mass fraction of individual messenger sequences in a cDNA probe, and by implication in its template RNA. In order to exploit this in the investigation of post-transcriptional controls on mRNA abundance in liver and hepatoma cells, suitable plasmids were selected from a rat

liver polysomal cDNA recombinant library, constructed in this laboratory by Ms. Rosemary Shott. The basis for the selection was an initial characterisation of the library, carried out by Ms. Shott, using the Grunstein-Hogness colony hybridisation procedure (234). This showed that a large proportion of the recombinants which hybridised strongly with rat liver polysomal cDNA, reacted weakly or not at all with HTC cell polysomal cDNA, as predicted on the basis of heterogeneous cDNA hybridisation experiments described here (Section 4.4). A minority of strongly positive clones was also highly abundant, and in a few cases more abundant, in the hepatoma probe.

Four recombinants were selected for further study, showing distinct relative abundance patterns in this preliminary screening: these are designated pRR 83 and pRR 5B, both of which scored as highly positive with liver cDNA, but negative with hepatoma cDNA, pRR 133, which was at similar abundance in the two probes, and pRR 117, which was slightly stronger on screening with hepatoma cDNA. Plasmids pRR 83 and pRR 5B were assumed to contain different cDNA inserts, due to the consistently stronger signal provided by the latter in preliminary screening. The selected clones are not intended to represent an unbiassed sampling of the liver cDNA library (indeed they are not, since there were far more recombinants behaving like pRR 83 and pRR 5B than like pRR 117). However, the four selected plasmids do represent the distinct patterns of modulation of mRNA abundances postulated on the basis of heterogeneous cDNA hybridisations to be occurring between rat liver and HTC cells, namely massive depletion, no change and modest increase. They were therefore considered useful for a preliminary study attempting to quantify the extent of abundance shifts, and the contribution to them of post-transcriptional control(s), in these cell-types.

Plasmid DNAs were isolated from amplified cultures of recombinant and wild-type (pAT 153) plasmids in E.coli

strains HB101 and C600 respectively. Plasmid DNA yields are shown in Table 6. The five plasmids were characterised by electrophoresis on a 1 percent agarose gel (Fig. 18a) which gives an indication of their purity, intactness and size, and also by double digestion with restriction endonucleases Hinf and Eco RI, followed by electrophoresis on a 6 percent polyacrylamide gel (Fig. 18b), which indicates the size and position of the inserted cDNA. Insert sizes were confirmed independently by heteroduplexing with plasmid pBR 322, carried out by Ms. Alison Slater. The information is schematised in Fig. 19, and insert sizes are indicated in Table 6. It will be noted that plasmid pBR 117 contains a large deletion at the insertion site, and the precise size of the insert can only be deduced from heteroduplex analysis. As can be seen from the gel shown in Fig. 18a, the four recombinants were isolated predominantly as supercoiled monomers, whereas the wild-type plasmid was mainly in the relaxed circular and concatameric forms. The reason for this is unclear: it may reflect different amplification properties of the host strains used. Differences in the plasmid copy numbers attained during amplification are the likely cause of most of the discrepancies in yields obtained with the different plasmids.

4.5.3. Filter hybridisations using recombinant plasmids: relative abundances in rat liver and HTC cell nuclear and polysomal poly(A)$^+$ RNAs

Recombinant and wild-type plasmid DNAs were bound to nitrocellulose filters as described under Materials and Methods. In order to standardise the conditions of filter-binding, the same mass of plasmid DNA was bound to each filter (1 µg) requiring the use of the correction factor already discussed (Section 4.5.1) to compensate the amount of hybridisation for the slight differences in mass of reacting DNA per unit area on the different filters. The additional uncertainty introduced into the estimates of relative abundance using this procedure

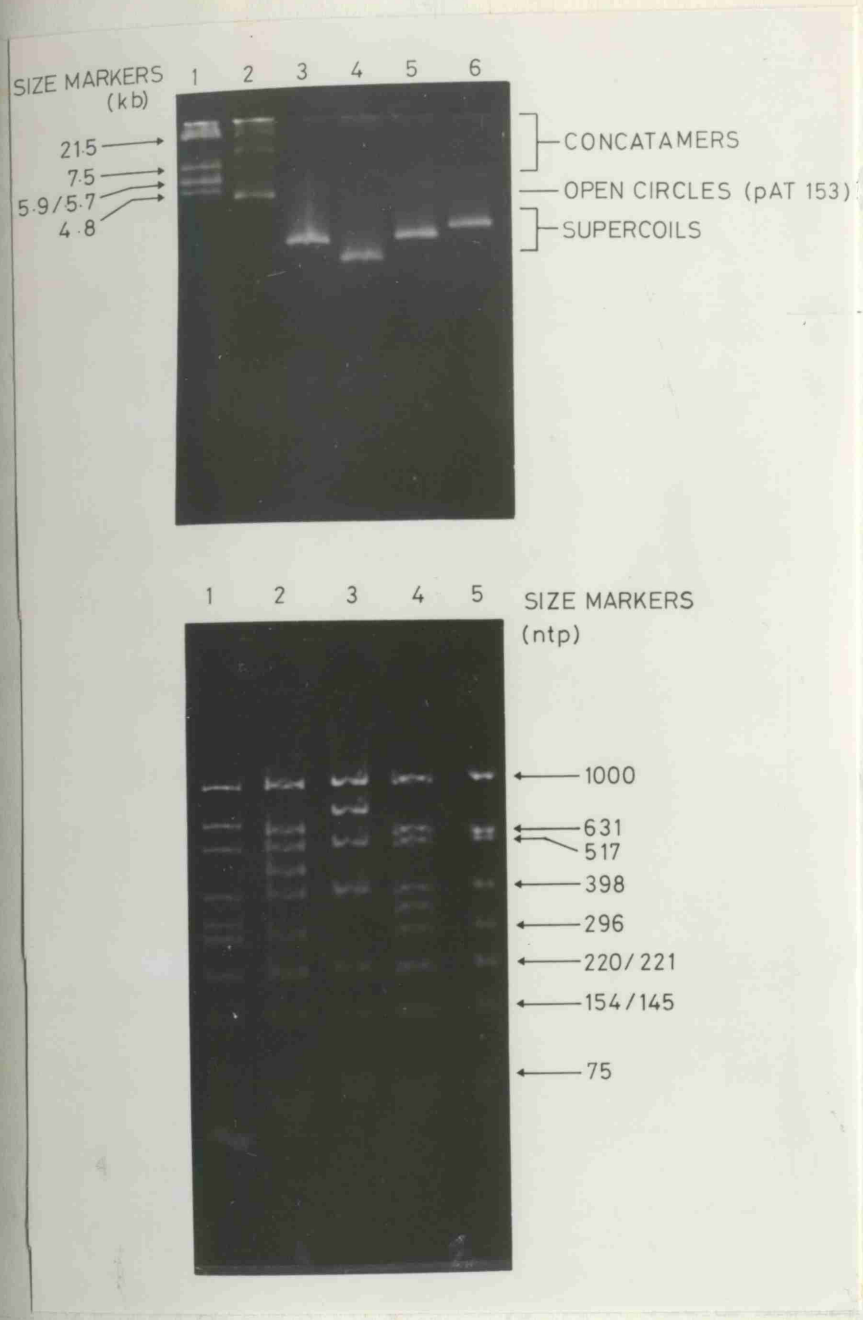


FIGURE 18. Gel electrophoretic characterisation of recombinant plasmids from a rat liver polysomal cDNA library.

(a) 1 percent agarose gel. Track 1: Eco R1 digested λ phage DNA markers; track 2: pAT 153 DNA (wild-type); track 3: pRR 83; track 4: pRR 117 track 5: pRR 133; track 6: pRR 5B. 1/2 µg DNA was run in each track.

(b) 6 percent polyacrylamide gel of Eco RI/HinfI digested plasmid DNAs: Track 1: pRR 5B, track 2: pRR 133; track 3: pRR 117; track 4: pRR 83; track 5: pAT 153.

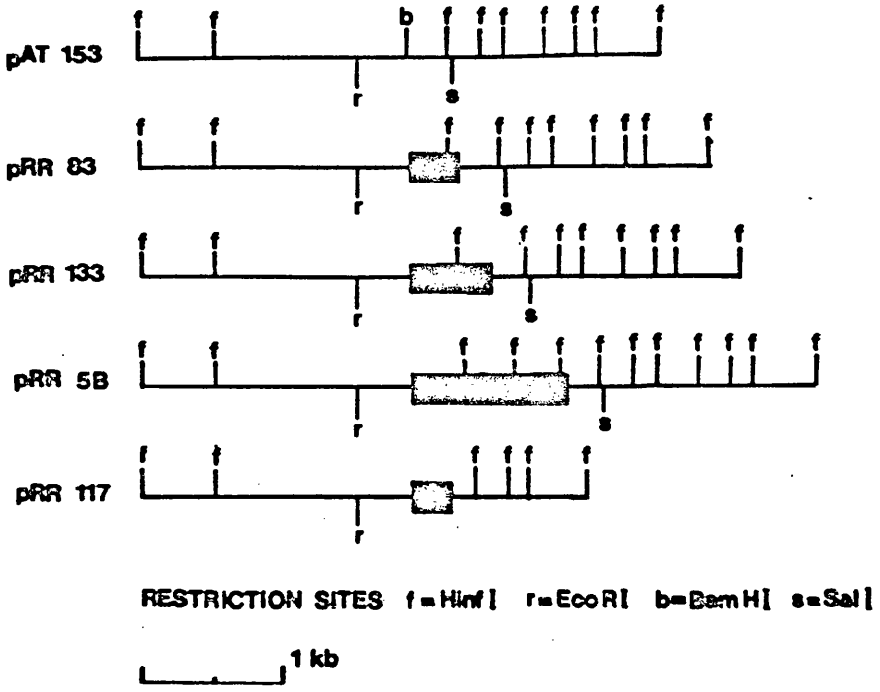


FIGURE 19. Simplified restriction maps of recombinant plasmids selected from a rat liver polysomal cDNA library. Data are from Figs. 18a and 18b, heteroduplex analysis (A. Slater), and Sal I restriction (R. Shott). Inserts are shown as bold lines.

TABLE 6. Studies of relative mRNA abundances in nuclear and polysomal poly(A)$^+$ RNAs from rat liver and HTC cells, using selected recombinants from a rat liver polysomal cDNA library (based on Fig. 16).

DESIGNATION	pRR 83	pRR 5B	pRR 117	pRR 133
DNA yield (a)	130	370	100	150
Insert cDNA size (b)	370	1100	260	550
Mass fraction: liver mRNA (c)	0.44	1.5	0.067(e)	0.44
Mass fraction: liver hnRNA	0.037	0.067(f)	0.0079	0.0094
Mass fraction: HTC mRNA	0.0072(g)	0.018	0.27	0.63
Mass fraction: HTC hnRNA	0.014	0.034	0.042	0.033
Mass fraction: regenerating liver mRNA	0.23	0.78	0.13	0.32
Mass fraction: regenerating liver hnRNA	0.018	0.15	0.14	0.012
Abundance ratio(d) liver mRNA/hnRNA	12	22	8.5	47
Abundance ratio liver/HTC mRNA	62	84	0.22	0.69
Abundance ratio Liver/HTC hnRNA	2.6	2.0	0.19	0.28

./.

All values quoted to 2 significant figures. SE = $\pm$ 18 percent, except where stated, computed as a percentage of the counts above background of sum of absolute SE's of experimental and background samples. Significance of each statistic above background ($p < 0.05$ in all cases) is computed using a t-test, which involves calculation of a combined SE from background and experimental samples, whose value is different and generally less.

(a) µg DNA per litre of culture (Wild-type yield = 800 µg/l)

(b) Nucleotide pairs.

(c) As a percentage of the cDNA to liver poly(A)$^+$ polysomal RNA: other mass fractions computed similarly for the corresponding cDNAs.

(d) Ratio of mass fractions quoted in the table

(e) Sample SE $\pm$ 41 percent, $p < 0.05$ (t test)

(f) Sample SE $\pm$ 2 percent.

(g) Sample SE $\pm$ 57 percent, $p < 0.05$ (t test).

is probably insignificant compared with other, more serious problems, which mean that the method can only be regarded as semi-quantitative. These will be discussed below, in this section.

Filters were hybridised in replicate sets with high specific activity (^{32}P-labelled)cDNA probes to rat liver, regenerating rat liver, and HTC cell nuclear and polysomal poly(A)$^+$ RNAs, under identical conditions to those used in the pilot study. On the basis of the counts hybridised in each case above background (the level seen with wild-type plasmid DNA), and applying the necessary corrections for probe hybridisability, mass of reacting DNA per unit filter area, and extent of reaction of the α-globin control, the mass fractions of each recombinant in each of the cDNA probes was computed, the data being summarised in Table 6. Values quoted are generally the means of duplicate, or where possible, triplicate filters, and where hybridisation levels above background were comparable with background itself, the entire experiment was repeated, using more counts if possible, and the overall mean computed by weighting the two sample means in proportion to the number of input counts in each case. Standard deviations were calculated for each set of replicate observations, and an average standard error of $\pm$ 18 percent was deduced. This is probably the most meaningful estimate of the error in each statistic in Table 6, with the exception of those where additional replicates were used, or where the sample variance was itself judged to be significantly above average. In these cases, the observed sample standard errors are quoted. All the values quoted in Table 6 are statistically significant above background, at a 95 percent confidence level. As well as showing the computed estimates of the mass fractions of each recombinant in each cDNA probe, Table 6 shows calculated ratios of abundance in liver polysomal/nuclear cDNA, in liver polysomal/hepatoma polysomal cDNA, and in liver nuclear/hepatoma nuclear cDNA.

Clearly, the interpretation of this experiment rests upon a number of assumptions. In common with the heterogeneous cDNA hybridisations, it is assumed that the relative frequencies of sequences in the cDNA probes accurately reflects the levels in the RNA population of the messengers from which they were transcribed. It is also the case that the filter hybridisation assay measures relative abundance by number rather than by mass, assuming that cDNA sizes are unrelated to template length under conditions where the bulk of cDNA molecules is considerably shorter than the template (as here). Although the length of insert cDNA sequences in the plasmids is not a direct determinant of the extent of reaction, it is assumed in comparing the hybridisation of different sequences, that in all cases, the length of messenger sequence represented in the plasmid driver exceeds and includes that in the cDNA probe. This would not be the case if the insert were too short, if it represented a 5' rather than 3'-proximal region of the mRNA sequence, or if the mRNA gave rise with a high frequency to full-length or near full-length reverse transcripts. If any of these conditions is not fulfilled, the percentage of the probe that is retained in hybrid on the filters will under-represent the abundance of the given mRNA sequence in the population. The cDNA inserts in the plasmids used here are as long as, or longer than the vast majority of cDNA transcripts from nuclear and polysomal RNA hence these problems are unlikely to arise (except possibly for clone pRR 117 if this is not an extreme 3' end clone of the mRNA sequence). In addition, it is conceivable that some clones could cross react, even under conditions of high stringency, with related gene sequences carrying limited homology regions, as would be the case for some members of the globin family in most vertebrates. The number of hybridised counts may, in the extreme, represent all the more abundantly expressed members of a family, rather than an individual mRNA species which may be expressed at a low level.

The data of Table 6 must therefore be regarded as preliminary, and at best, semi-quantitative. Their true significance can only be assessed when the particular recombinants selected have been much more thoroughly characterised, and preferably complemented by similar studies using many other clones. The type of characterisation that is necessary involves thorough restriction mapping, genomic blotting, RNA blotting and translational identification studies, in order to establish that the recombinants are all of distinct, non cross-reacting genes; measurements of the melting temperature of hybrids formed with the plasmid DNA to check that cross-reaction is not occurring with other mRNA species; and determination of the location, within the messenger, of the insert cDNA sequence.

Despite the uncertainties, and the preliminary nature of such a study involving just four recombinants, a number of general conclusions may be drawn from the data of Table 6. Firstly, all four sequences show distinct relative abundance patterns in the different RNA populations tested, strengthening the view that they represent different mRNAs. Secondly, the measured abundances of the clones correspond with those expected from heterogeneous hybridisation studies. Thus, the frequencies of mRNAs corresponding to clones pRR 83, pRR 133, and pRR 5B, which on the basis of Grunstein-Hogness screening are typical of the more abundant species, are close to those postulated in Table 3 for the most abundant liver messengers. Clone pRR 5B would be regarded as representing a super-abundant messenger sequence, whilst pRR 83 and pRR 133 are on the fringes of so being. Clone pRR 117 is typical of a group of mRNA sequences at moderate abundance. There is no evidence here for sequences at even higher relative abundance in liver mRNA (such as determined elsewhere for albumin messenger), supporting the view that liver polysomes, as prepared here, may be selectively depleted of sequences coding for secretory proteins. (Indeed, one of pRR 83 or pRR 5B may correspond to albumin cDNA).

Thirdly, statistically significant differences in the polysomal/ nuclear abundance ratio, not simply related to abundance,

imply that polysomal frequencies are determined by sequence-specific post-transcriptional selection. For example, clones pRR 83 and pRR 133, which are at very similar polysomal abundances, are at quite different levels in steady-state liver hnRNA. It will be noted that for abundant sequences in general, relative nuclear abundances are at least an order of magnitude lower than those on the polysomes, as found in CHO cells by Harpold et al (231). Conversely, sequences which in a given context are at 'rare-class' levels on the polysomes are at comparable, or slightly higher relative levels in hnRNA, such as is the case for pRR 83 or pRR 5B mRNAs in HTC cells. This confirms the inferences drawn from cross-hybridisation studies using heterogeneous nuclear and polysomal cDNAs (Section 4.3.2). Fourthly, the drastic depletion of high abundance liver mRNAs in HTC cells, deduced from the data presented in Section 4.4, is shown for two typical representatives of the class, pRR 83 and pRR 5B. Their detection above background in the hepatoma, demonstrates that their modulation is quantitative, rather than qualitative. Other investigators have concluded, on the basis of hybridisation data similar to that presented earlier, that qualitative changes occur in the mRNA population in hepatoma cells (146,150). At least in the case of HTC cells, such a conclusion is certainly not supported by the data of Table 6. The quantitative extent of the depletion is similar to that postulated on the basis of heterogeneous hybridisation kinetics (Section 4.4.2), namely approaching two orders of magnitude. If the clones are typical in this respect, the data would appear to support the 'most general' solution of Table 4. In addition, measurement of their levels in poly(A)-adjacent hnRNA confirm that the predominant mechanism regulating their abundance is post-transcriptional, since only a 2-3 fold depletion is evident in steady-state hepatoma nuclear RNA. Similar filter hybridisations, using pulse-labelled HTC cell nuclear RNA (data not shown) suggest that they are transcribed at similar rates to those sequences which remain at high abundance, such as represented by pRR 133 or pRR 117.

A fifth deduction from the data of Table 6 is that the abundant liver mRNAs massively depleted in hepatoma polysomes continue to be expressed at a high level in 16-hour regenerating rat liver. This may simply reflect their long half-lives. Alternatively, it could indicate that differentiation-linked functions, which are post-transcriptionally 'repressed' in HTC cells, are not substantially modulated during liver regeneration. The very modest decline in their abundance in regenerating liver (145 and Table 6) can be attributed to the increased steady-state levels of growth-related mRNAs. The regenerating liver may therefore differ from quiescent liver only in terms of its proliferative state, not the degree of its physiological specialisation. A sixth and final point, arising from Table 6, is the confirmation that at least one moderately abundant liver mRNA (that represented by clone pRR 117) is at elevated levels in both HTC cells and regenerating liver, and is therefore a strong candidate for a growth-related or growth-inducible gene. The use of simple screening techniques may permit the isolation of a large number of such sequences, and investigation of the degree to which they are co-ordinately controlled. The data suggest that the level of the mRNA corresponding to clone pRR 117 may be raised by accelerated transcription (or early post-transcriptional processing), since it is induced to the same extent in nuclear and polysomal HTC cell RNAs. Its even higher level in regenerating liver nuclear RNA may reflect maximal transcription whilst it is still accumulating on the polysomes.

Although the preliminary nature of this investigation using cloned cDNAs must again be stressed, its principal significance lies in its confirmation of the inferences from heterogeneous cDNA hybridisation studies. The patterns of expression and regulation of mRNA sequences, hypothesised on the basis of the data presented in Sections 4.2, 4.3 and 4.4, are matched by those of real mRNA species, and are not, therefore the result of methodological artifacts.

4.6. Summary of conclusions from hybridisation studies of in vivo RNA populations

It is useful at this point to consider the implications of the results reported thus far, for the development and characterisation of a cell-free system in which the physiological specificity of post-transcriptional controls is maintained. Taking an overview of the results quoted in Sections 4.1 - 4.5, it would appear that the mRNA and hnRNA populations of rat liver and HTC cells are qualitatively very similar. Their different phenotypes result from differences in relative mRNA levels which are determined and modulated transcriptionally and post-transcriptionally. In particular, rat liver expresses a group of high abundance messengers which are post-transcriptionally depleted in HTC cells, and whose place in the mRNA population is taken by a much larger number of rarer sequences which are more subtly elevated in the tumour.

Sensitive hybridisation assays for post-transcriptionally regulated sequences are essential for judging the physiological equivalence of in vitro RNA processing and transport, and the above in vivo study of RNA populations suggests several types of probe useful for such a characterisation. The maintenance of qualitative post-transcriptional selectivity can be judged by the use of single-copy DNA probes: single-copy DNA, enriched for sequences present in hnRNA, but absent from the polysomes, provides the most sensitive qualitative test for the presence of nucleus-restricted sequences. The degree to which the quantitative composition of in vitro-transported RNA resembles that of in vivo RNA populations may be determined by the use of probes for sequences which change radically in relative abundance between nucleus and cytoplasm. These include total (or abundant) polysomal cDNA, the bulk of which is much less abundant in hnRNA than on the polysomes, cloned cDNA probes for

sequences modulated to different degrees by post-transcriptional selection, and a heterogeneous cDNA probe for abundant nuclear poly(A)-adjacent sequences which are substantially rarer in polysomes, such as the one described in Section 4.3.2. An alternative approach is to prepare a heterogeneous cDNA for in vitro-transported poly(A)-adjacent sequences, and measure its relative content of sequences abundant in nuclei and polysomes, by heterologous hybridisation with RNAs derived from these sources.

The effect of cytosol on the specificity of mRNA processing and transport in vitro, in terms of the cell-type of origin of the cytosol, can be judged by using hybridisation assays for sequences post-transcriptionally regulated between cell-types (see Section 2.5). Results indicate several hybridisation probes which are useful for this in the context of the liver/hepatoma system. These include heterogeneous abundant liver polysomal cDNA, and cloned cDNAs for individual abundant messengers showing post-transcriptional depletion in HTC cells, as well as other cloned cDNAs regulated more subtly at a post-transcriptional level, as revealed by further screening of liver and HTC cell cDNA libraries.

The results of in vivo studies therefore provide a springboard for developing an in vitro system in which regulatory mechanisms may be investigated in greater detail. They also lead to certain important conclusions regarding the way in which mRNA abundances are related to phenotype, and at what levels in the cell they are modulated. The nature of these conclusions, viewed in the light of the many related studies referred to in the introduction (Chapter 2), will be discussed subsequently (Section 5.7).

4.7. Transport of RNA from isolated HTC cell nuclei: preliminary characterisation of an in vitro system for studying post-transcriptional control

In order to establish the conditions under which isolated nuclei will process and transport RNA sequences with in vivo-like specificity, as envisaged in the foregoing section, it is necessary to analyse the composition of the transported RNA using hybridisation methodology. Before this can be attempted, however, it is essential to ensure that the crude properties of the system, and of the transported RNA, correspond with those expected on the basis of in vivo observations.

The starting point for the development of such a system is provided by the conditions used by previous investigators (519), in which nuclear lysis is minimal, the release of pre-labelled RNA is dependent on both ATP and cytosol, continues linearly for up to an hour, and is messenger-like by criteria of size, poly(A)-content, association with proteins in mRNP-like particles and the absence of nucleus-restricted repetitive sequence transcripts (537). These conditions were originally used for rat liver nuclei, although similar conditions have been used for other cell-types (521,530) including tissue-culture cells. For the purposes of this study, it was decided to adapt the conditions to the incubation of HTC cell nuclei. This was principally due to the ease with which nuclear RNA may be pulse-labelled to high specific radioactivity in cell-culture, facilitating both the confirmation of the crude properties of the system and of the transported RNA and (possibly) hybridisation analysis of its composition. In addition, the use of hepatoma nuclei may favour the detection of cytosol-induced specificity, assuming that the factors which determine it act in a 'positive' manner, permitting the processing and transport of particular mRNAs or sets of mRNAs to be more efficient than that of the bulk of mRNAs.

The initial characterisation of the system therefore required confirmation of these crude properties for HTC cell nuclei. Firstly, this involved a thorough investigation of the methods of preparation of the nuclei, giving the most acceptable yields, with full retention of metabolic activity and structural integrity during subsequent incubation, but minimal contamination with RNA of cytoplasmic origin (Section 4.7.1). Secondly, the extent, kinetics, and energy- and cytosol-dependence of *in vitro* transport of steady-state labelled (mainly ribosomal) and pulse-labelled (mainly heterogeneous) nuclear RNA were measured, and the effects upon them of alteration in the conditions of incubation were determined (Section 4.7.2). Thirdly, the transported RNA (ribosomal and messenger-like) was sized, and characterised with respect to poly(A)-content and association into mRNPs (Section 4.7.3).

4.7.1. Preparation of nuclei and cytosol for in vitro RNA transport

The main problems which have been encountered in the preparation of metabolically active nuclei from tissue-culture cells are leakage or degradation of nuclear components, and contamination with cytoplasmic material. In this instance, it is also necessary to ensure that the nuclear preparation is free of whole cells and of lysed nuclear fragments. Leakage or degradation of nuclear components may lead to a (not necessarily random) loss of RNA species which are the substrates for processing and transport, or of the factors responsible for catalysing these reactions or determining their specificity. The presence of divalent cations, such as Ca^{2+} or Mg^{2+}, is known to promote the retention of DNA and RNA (606), although Ca^{2+} may impair enzymic functions such as RNA polymerase (606). The use of polyamines also favours nuclear integrity (606). Isolation procedures based on organic solvents appear to maximise the conservation of nuclear enzymic

activities (607), but are not recommended due to the danger of disturbing the integrity of other cellular structures, with uncertain effects on the properties of nuclei. Nuclei can be seriously disrupted by the release of degradative lysosomal enzymes during cell fractionation, and this has generally been minimised by the inclusion of osmotic stabilisers, such as sucrose, and the avoidance of detergents in media used for cell homogenisation. Cytoplasmic contamination arises principally from the contiguity of the nuclear envelope with the membrane system of the ER. Detergent treatments (606) and sedimentation through a dense sucrose pad (which eliminates material of lower buoyant density than pure nuclei) (608), have been used to overcome this problem. The citric-acid method, used to prepare ultra-pure nuclei for studies of the composition of in vivo steady-state hnRNA (see Section 3.2.2) is inappropriate here, since the acidity (pH 3-4) inactivates enzymic functions (606).

The approach used here has been to investigate, from the points of view of yield and purity, variants of a basic method which fulfils most of the criteria discussed above. Technical details of the methods are described in Section 3.13 and Chart 1, but a brief summary of their main features is included in Table 7. HTC cells were pre-labelled in culture for three days with 1 mCi/l ^{3}H-uridine, and for the final 18 hours, in fresh medium, with 50 µCi/l ^{14}C-thymidine. Nuclear samples were prepared by the various different procedures, according to an experimental flow programme, and were scored for the percentage recovery of ^{14}C counts (being an indication of nuclear yield), and for the ^{3}H/^{14}C ratio. This ratio, compared with that obtained using the sucrose/citric acid method, gives an indication of the proportion of cytoplasmic RNA which remains associated with the nuclei (assuming that citric acid nuclei are effectively free of cytoplasmic contamination, which is reasonable in view of the absence of cytoplasmic 'tags' visible in the light microscope

TABLE 7. Yield and purity of HTC cell nuclei prepared by various methods.

Lysis Method	Crude nuclei NP40 treated?	Spin	Sample density (sucrose)	Pad density (sucrose)	Percentage contaminating RNA retained	Percentage Yield (a)
Citric acid	NO	Low-speed	0.25 M	0.88 M	0 (baseline)	23 ± 2
Osmotic	NO	Low-speed	0.3 M	0.88 M	22.8 ± 2.2	63 ± 6
Osmotic	NO	High-speed	1.6 M	1.9 M	13.3 ± 1.6	23 ± 2
Osmotic	NO	High-speed	1.6 M	2.2 M	?	0
Osmotic	NO	High-speed	1.975M	None	10.9 ± 1.5	16 ± 2
Osmotic	YES	Low-speed	0.3 M	0.88 M	24.0 ± 2.2	88 ± 9
Osmotic	YES	High-speed	1.6 M	1.9 M	12.9 ± 1.6	76 ± 8
Osmotic	YES	High-speed	1.6 M	2.2 M	9.2 ± 1.4	37 ± 4
NP40	-	Low-speed	0.3 M	0.88 M	23.8 ± 2.2	18 ± 2
NP40	-	High-speed	1.6 M	1.9 M	24.6 ± 2.3	12 ± 1
NP40	-	High-speed	1.6 M	2.2 M	19.0 ± 1.9	6 ± 1
NP40	-	High-speed	1.62 M	None	46.6 ± 3.5	52 ± 5
Triton-X 0.1 percent	-	High-speed	1.6 M	2.2 M	53.2 ± 3.8	38 ± 4 (c)

(a) Percentage recovery of ^{14}C counts.

(b) Percentage of "cytoplasmic" ^{3}H counts remaining associated with the nuclei above baseline, computed using the equation

$$C = \frac{(R_x - R_c)}{R_T - R_c} \times 100 \text{ percent}$$

where $R_x = {}^{3}H/{}^{14}C$ ratio of sample
$R_c = {}^{3}H/{}^{14}C$ ratio of citric acid nuclei
$R_T = {}^{3}H/{}^{14}C$ ratio of whole cell suspension.

(c) Method used by Roy et al. (514) for preparing rat liver nuclei for in vitro RNA synthesis and release.

Errors calculated as SE's from replicate determinations.

and that there is no leakage from them of true nuclear RNA. Table 7 summarises the yield and purity of nuclei prepared by each variant method, as determined by this procedure. The overall conclusions from this study are as follows. Firstly, both detergent treatment and sedimentation through a dense sucrose pad are essential for obtaining satisfactory yields of acceptably pure nuclei. Secondly, detergent treatment should only be introduced after initial fractionating by low speed centrifugation, in order to avoid catastrophic losses. Thirdly, centrifugation through sucrose pads denser than 2.0 M has only a marginal effect on nuclear purity, but drastically reduces yield. Fourthly, final washing of the nuclei in 1 mM $CaCl_2$, as used previously for the preparation of nuclei for <u>in vitro</u> RNA transport (538) does not result in detectable losses, and is therefore an acceptable way of eliminating residual sucrose and detergent. Finally, it must be noted that all the methods used yielded nuclei substantially more contaminated than by the sucrose/citric acid method (which may, however cause some losses of true nuclear RNA). Accurate estimation of the true extent of this contamination, relative to the amount of nuclear RNA, is not possible since the precise relationship of nuclear and cytoplasmic specific radioactivities in RNA is unknown. Although the specific activity of nuclear RNA is probably lower than that of cytoplasmic RNA under the labelling regime used, a maximum estimate of the contamination may be derived, assuming that nuclear and cytoplasmic specific activities are similar. In this case, contaminating RNA of cytoplasmic origin could represent up to 15 percent of the RNA associated with nuclei prepared by the cleanest methods.

The method of choice of nuclear preparation, subject to clear demonstration that it generates nuclei capable of sustained RNA-transporting activity similar to that previously demonstrated for liver nuclei, is therefore as follows: cells are osmotically swelled, homogenised, and centrifuged

in 0.3 M sucrose at low speed; the crude nuclear pellet is resuspended in calcium-containing sucrose, treated with NP40, and spun at high speed through a 2.0 M sucrose pad, after which nuclei are washed in 1 mM $CaCl_2$. The method is stated in full detail in Section 3.13.2.

The adherent RNA of cytoplasmic origin seems to have associated non-specifically with nuclei, and is not therefore accounted for by polysomes bound to the nuclear envelope. This can be seen from the results of a simple mixing experiment, in which the crude post-nuclear supernatant from cells pre-labelled for 18 hours in the presence of 1 mCi/l ^{3}H-uridine was mixed with a suspension of unlabelled cells immediately prior to homogenisation. About 8 percent of the label remained associated with the nuclei after preparation by the method of choice. This contamination would be unimportant if the contaminating adherent RNA were not released from the nuclei during incubation <u>in vitro</u>. Unfortunately, this is not the case, as will be shown subsequently. Therefore it must be taken account of in the design of experiments to study the composition of <u>in vitro</u>-transported RNA.

The method used to prepare cytosol (see Section 3.13.2) follows closely that used by previous investigators (538,487), despite its rather crude nature. In order to minimise contamination with material originating in other sub-cellular organelles (which includes nucleases), cells were osmotically swelled, and exposed at no time to detergent. Particulate material was removed by high speed centrifugation, and the supernatant was concentrated by vacuum dialysis, or on one occasion, by polyethylene glycol treatment in a bulk preparation, after which it was dialysed extensively against distilled water, again following the procedure adopted by other investigators. This crude preparation may contribute to the selective loss of activities in the cytosol, which may influence its properties in relation to <u>in vitro</u> RNA transport. A proportion of the protein inevitably precipitates on dialysis

against distilled water, and this needs to be removed by centrifugation. Aliquots of cytosol were stored at -20°, were not inactivated by one cycle of freeze-thawing, and were stable for several months. The protein concentration of cytosol preparations was determined by Bradford's assay (598). Liver cytosol was prepared by essentially the same procedure. Perfusion of livers prior to excision, and very gentle homogenisation conditions, should have minimised contamination with blood cells, serum, and liver cell-types other than hepatocytes.

4.7.2. Kinetics, energy and cytosol-dependence of in vitro RNA transport

Nuclei were prepared from both steady-state and pulse-labelled HTC cells (see Section 3.14.1 for cell labelling regimes). The pulse-labelling regime is similar to that employed by Harpold et al (231) and results in the incorporation of about 15,000 dpm per 10^6 cells into acid-precipitable material, using ^{3}H-uridine. Labelled nuclei were incubated at 30° in essentially the same medium as that of Schumm and Webb (519), but with the following modifications: E.coli RNA, rather than yeast RNA was used as carrier, at a concentration of 350 µg/ml; the ATP regenerating system consisted of creatine phosphate and creatine phosphokinase, rather than phosphoenol-pyruvate and pyruvate kinase; nuclear suspension densities were generally 10-fold greater than those used previously, in order to conserve cytosol, and achieve a more realistic nucleus: cytosol protein ratio. The medium contained spermidine, Mn^{2+}, and Ca^{2+} ions, which have been reported as assisting in the maintenance of nuclear integrity (519,522), as well as dithiothreitol and other salts. Counts released to the post-nuclear supernatant were determined at different time-points, under various conditions. Kinetics of release of steady-state labelled and pulse-labelled nuclear RNA are shown in Fig. 20. 99.5 percent of the counts associated with steady-state labelled nuclei were found to be acid-

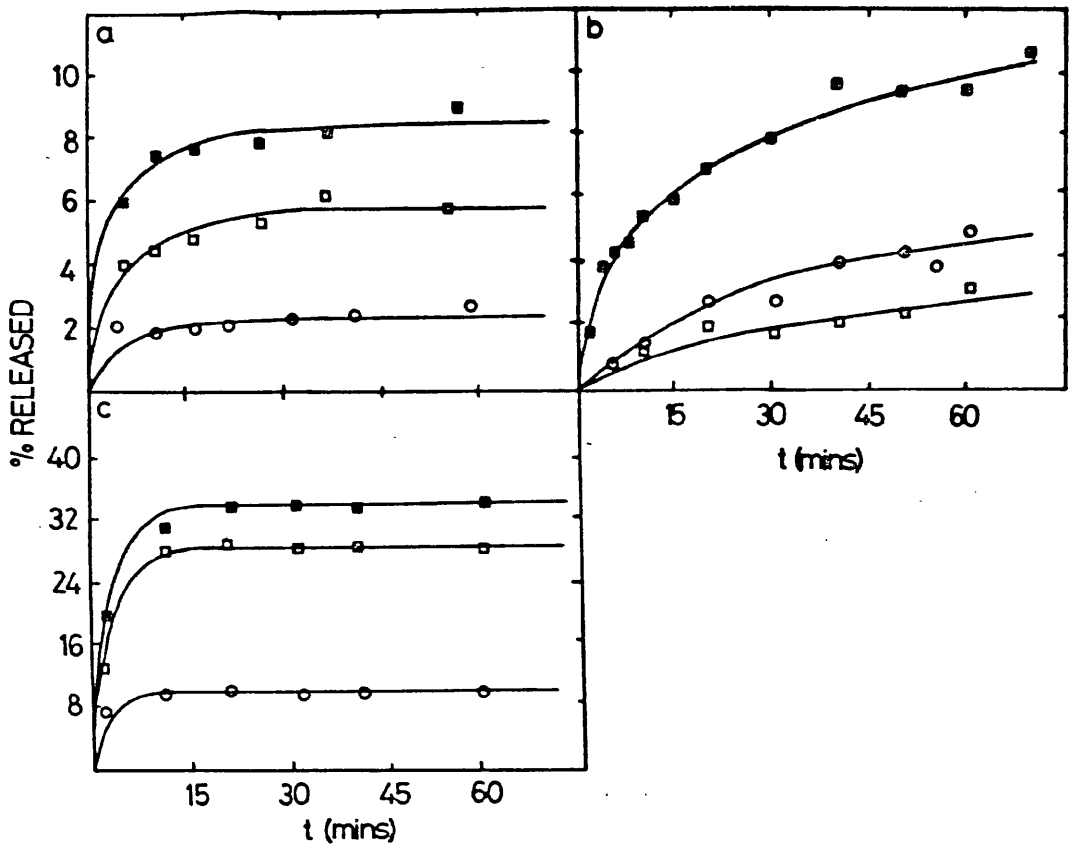


FIGURE 20. Kinetics of release of steady-state labelled and pulse-labelled RNA from isolated HTC cell nuclei.

(a) release of steady-state labelled RNA: +ATP (5 mM) (■); + AMP.P (NH)P (5 mM) (□); -ATP (○)

(b) release of pulse-labelled RNA: +ATP (5 mM) (■); +AMP.P(NH)P (5 mM)(□); -ATP (○).

(c) release of adherent RNA of cytoplasmic origin (labelled steady-state RNA derived from cytoplasm with which unlabelled cells were homogenised): +ATP(5 mM) (■); +AMP.P(NH)P (5 mM) (□); -ATP (○).

precipitable, therefore only total released counts were determined in this case, whereas released counts from pulse-labelled nuclei are acid-precipitable counts. Fig. 20 also shows the release of 'adherent nuclear RNA' - labelled RNA which remained associated with unlabelled nuclei in the mixing experiment described in the previous section. In addition to the kinetics of transport in the complete medium, Fig. 20 also shows the kinetics observed when ATP and its regenerating system were omitted, or substituted by a non-hydrolysable analogue.

A number of general deductions are possible from the data of Fig. 20. Firstly, the nuclei as prepared, despite detergent treatment, are active in the transport of both steady-state (mainly ribosomal) and pulse-labelled (mainly heterogeneous) RNA, the kinetics being similar to those reported for RNA transport from rat liver nuclei (519,538). Secondly, the release of steady-state labelled RNA is distinguishable from that of pulse-labelled RNA, by criteria of kinetics and ATP-dependence. Pulse-labelled RNA transport shows a much less dramatic initial burst, continues almost linearly for more than an hour, is decreased some 60 percent if ATP is omitted, and is further inhibited by the non-hydrolysable analogue AMP.P(NH)P. It would therefore appear to be an active process dependent upon ATP hydrolysis. By contrast, steady-state labelled RNA release, although also ATP-dependent, can be promoted by a non-hydrolysable analogue in its place. This suggests that the transport of rRNA to the medium in this system is either passive, or artifactual. The effect of ATP may be either through cation chelation, or by allosteric interaction with a site in the nuclear envelope, or just possibly, involving pyrophosphate cleavage of ATP (between α and β phosphate groups). The energy-dependence of RNA transport <u>in vitro</u> is further discussed in section 5.4.

A third deduction from Fig. 20 is that nuclear-adherent RNA is released with similar kinetics and energy-dependence to steady-state labelled nuclear RNA, and moreover, a larger proportion of it is released than of true nuclear RNA. Even taking account of the likely higher specific activity of the adherent RNA, it is likely that a significant proportion of total released RNA is cytoplasmic in origin. This problem, can, however, be largely overcome by using a brief pre-incubation step. Within a five minute pre-incubation, some 80-90 percent of the removable adherent RNA is released, whereas relatively little of the heterogeneous RNA available for transport is released. In a further 75 minutes of incubation, the kinetic data of Fig. 20 imply that more than 95 percent of the heterogeneous RNA transported from nuclei is of true nuclear origin (assuming that the proportion of adherent RNA which is polyadenylated mRNA is similar to that found in total cytoplasm (< 2 percent). One problem is that the population of sequences transported during the pre-incubation may not be the same as that transported during the following 75 minutes. This is unlikely to arise if the pre-incubation time is short compared even with the processing time for the most efficiently processed messengers, which is a reasonable assumption. Indeed, consideration of observed processing times for individual mRNAs, coupled with estimates of processing efficiency (relative sequence conservation) deduced from <u>in vivo</u> studies (609) suggest that the extent and kinetics of <u>in vitro</u> pulse-labelled RNA transport could be due to physiological processes.

If RNA transport <u>in vitro</u> is the result of a physiological mechanism, rather than, say, non-specific leakage of nuclear components, there should be only minimal release of such components to the medium. Nuclei were therefore double-labelled by culturing HTC cells for 18 hours with ^{14}C-thymidine, and adding ^{3}H-uridine for a final 15 minute pulse. RNA and DNA release were followed in relation to variables of the medium

composition in order to test, in particular, the effect of spermidine and carrier RNA on nuclear integrity and RNA transport. The results, shown in Fig. 21 indicate a stringent maintenance of nuclear integrity under the conditions recommended by Schumm and Webb (519), where DNA release over 60 mins. is undetectable, and RNA release is optimal. Nuclear integrity appears to be well maintained above a level of 2 mM spermidine, but drops off dramatically below this concentration. Pulse-labelled RNA transport is also inhibited by high levels of spermidine, but is still substantial at spermidine concentrations which prevent DNA release. The apparent effect of carrier RNA is slightly to inhibit the release of RNA, but this finding should be treated with some caution, since other conditions (see below), which favour enhanced RNA release, are accompanied by increased activity of nonspecific ribonucleases.

The integrity of the post-incubation nuclei was confirmed by light microscopic examination, although at high suspension densities considerable clumping occurred, resulting in the formation of macroscopic aggregates. These aggregates were efficiently dispersed by vortexing or gentle hand homogenisation, and their formation did not impair RNA transport. The cause of this aggregation is not known, but it may be due to a very limited degree of nuclear lysis, releasing chromatin fibres into the medium which enmesh the nuclei.

The cytosol dependence of <u>in vitro</u> RNA transport, which is the subject of considerable controversy in the literature, was studied by measuring the dose-response to homologous (HTC) and heterologous (rat liver) cytosol, of pulse-labelled HTC nuclear RNA release. The results (Fig. 22) show an unexpected phenomenon. In the complete absence of cytosol, there is a massive release of pulse-labelled RNA (and of steady-state labelled RNA) whereas at all tested cytosol concentrations, the extent of RNA transport shows only a very slight dose-dependence. In a double-label experiment, the release of DNA,

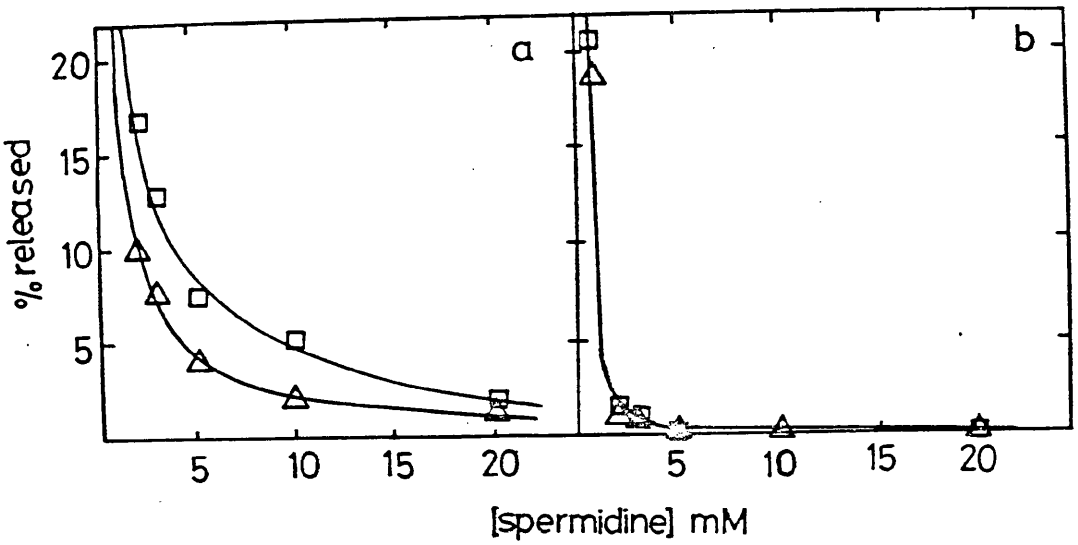


FIGURE 21. Effect of spermidine and exogenous RNA on in-vitro RNA transport and nuclear integrity. Cells were double labelled in RNA (^{3}H) and DNA (^{14}C) (see text) and isolated nuclei incubated for 60 minutes in various media. Release of (a) pulse labelled RNA (^{3}H counts) and (b) DNA (^{14}C counts) were determined in the presence (△) and absence (□) of 350 µg/ml E.coli RNA, at various concentrations of spermidine.

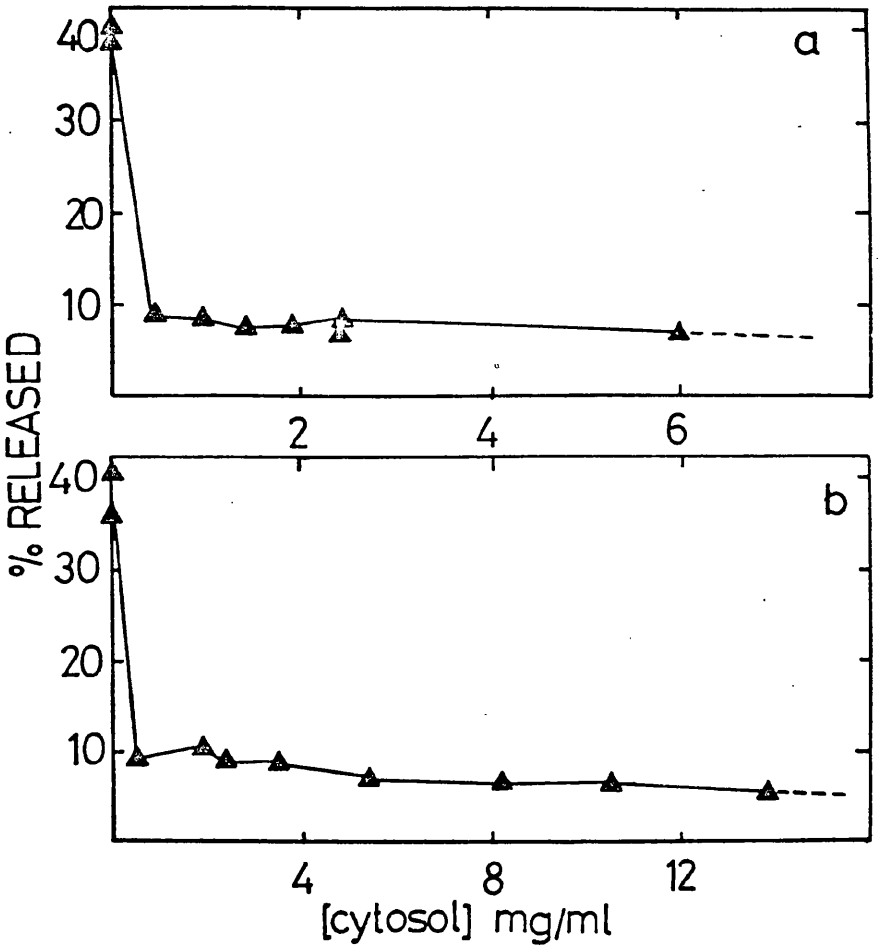


FIGURE 22. Dose response of in vitro RNA transport to (a) homologous (HTC cell) and (b) heterologous (liver) cytosol. Incubations were carried out for 60 minutes under standard conditions (Section 3.14.2).

in the absence of cytosol, was at the same low level as observed previously. Nuclear lysis would not therefore seem to be the cause of this phenomenon, as has been suggested elsewhere (530). Size characterisation of the transported RNA (see following section, 4.7.3) suggests that it may be due to uncontrolled, nonspecific ribonuclease action, in the absence of a nuclease inhibitor present in cytosol. The poor dose-dependence of RNA transport upon cytosol is hard to reconcile with ideas that cytosol may be restoring vital physiological functions to the nuclei, and suggests that the specificity of *in vitro* RNA transport may not be an exact model for that which pertains *in vivo*.

4.7.3. Crude characterisation of RNA transported from HTC cell nuclei in vitro

RNA transported over 75 minutes of incubation from both steady-state labelled and pulse-labelled HTC cell nuclei, in the presence or absence of homologous cytosol, was isolated from post-nuclear supernatants by phenol-chloroform extraction (Section 3.3) and ethanol precipitation, and sized by rate-zonal centrifugation on sucrose density-gradients, alongside ^{3}H-labelled cytoplasmic RNA markers (Section 3.15, Fig. 23). Also shown is the size profile (together with size markers) of pulse-labelled nuclear RNA on a denaturing gradient (99 percent DMSO). These profiles demonstrate graphically the undegraded nature of steady-state RNA transported in the presence of homologous cytosol, consisting primarily of mature ribosomal (and possibly transfer) RNA species sedimenting at 4/5S, 18S and 28S, but lacking the 45S species which is found exclusively in total untransported nuclear steady-state RNA. In the absence of cytosol, however, the RNA is seriously degraded, with only a small proportion sedimenting at 18S, none detectable at 28S, and most consisting of very small fragments. This is compatible with the suggestion that in the absence of cytosol, nonspecific ribonuclease action promotes massive RNA release from isolated nuclei. Pulse-labelled

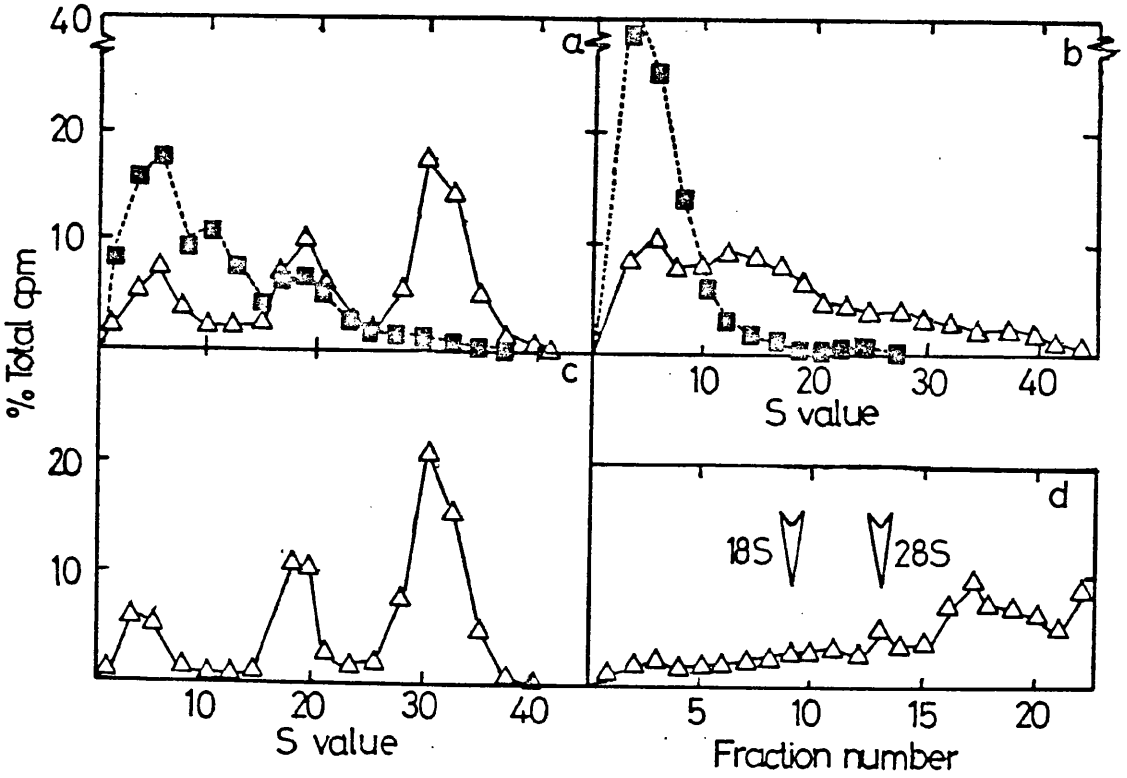


FIGURE 23. Sizing of in vitro-transported RNA by sedimentation on 15-30 percent NETS-sucrose gradients. Centrifugation conditions were as stated in the text. The proportion of acid-precipitable counts in each gradient fraction is plotted against S value (derived from computer analysis).

(a) Steady state labelled RNA transported in presence of cytosol (—△—) or in absence of cytosol (--□--).

(b) Pulse-labelled RNA transported in presence of cytosol (—△—) or in absence of cytosol (--□--).

(c) Steady-state labelled HTC cell cytoplasmic RNA (size markers)

(d) Pulse-labelled HTC cell hnRNA analysed on a 99 percent DMSO, 0-10 percent sucrose (denaturing) gradient as described by N.A. Affara (MSE Application Information Sheet No. A14/9/76), with size markers run in parallel as in (c) above.

RNA follows a similar pattern, with the bulk of the material in the unincubated nuclei larger than 28S, and transported RNA being heterodispersed between 7-20S, when cytosol was present (Fig. 23), but being seriously degraded when cytosol was absent.

A crude estimate of the proportion of poly(A)$^+$ sequences in transported RNA was obtained by oligo-(dT)-cellulose chromatography. Transported pulse-labelled RNA was reproducibly about 25 percent poly(A)$^+$, compared with about 50 percent for the total pulse-labelled RNA in unincubated nuclei, on a single passage. Taken together with the sizing described above, this suggests that, allowing for some degradation, transported pulse-labelled RNA is messenger-like, whereas transported steady-state labelled RNA was predominantly ribosomal (exhibiting oligo-(dT)-cellulose binding of less than 1/2 percent). A further crude test of the messenger-like properties of transported RNA is formation of mRNP-like particles. Pulse-labelled HTC cell nuclei were incubated *in vitro* for 75 minutes, after which the post-nuclear supernatant was divided into aliquots and layered onto sucrose density-gradients in different buffers (Section 3.15). A preparation of post-mitochondrial supernatant from steady-state labelled cells was centrifuged in parallel, in order to provide markers for ribosomal subunit sizes. As shown in Fig. 24, the transported pulse-labelled RNA was incorporated in particles of heterogeneous size, sedimenting at 10-70S. Although a proportion of this undoubtedly consists of ribosomal subunits, it is also clear that a major fraction peaks in the range 20-40S in 1.5 mM $MgCl_2$, or slightly smaller in the presence of EDTA, in other words has the sedimentation properties expected for mRNPs.

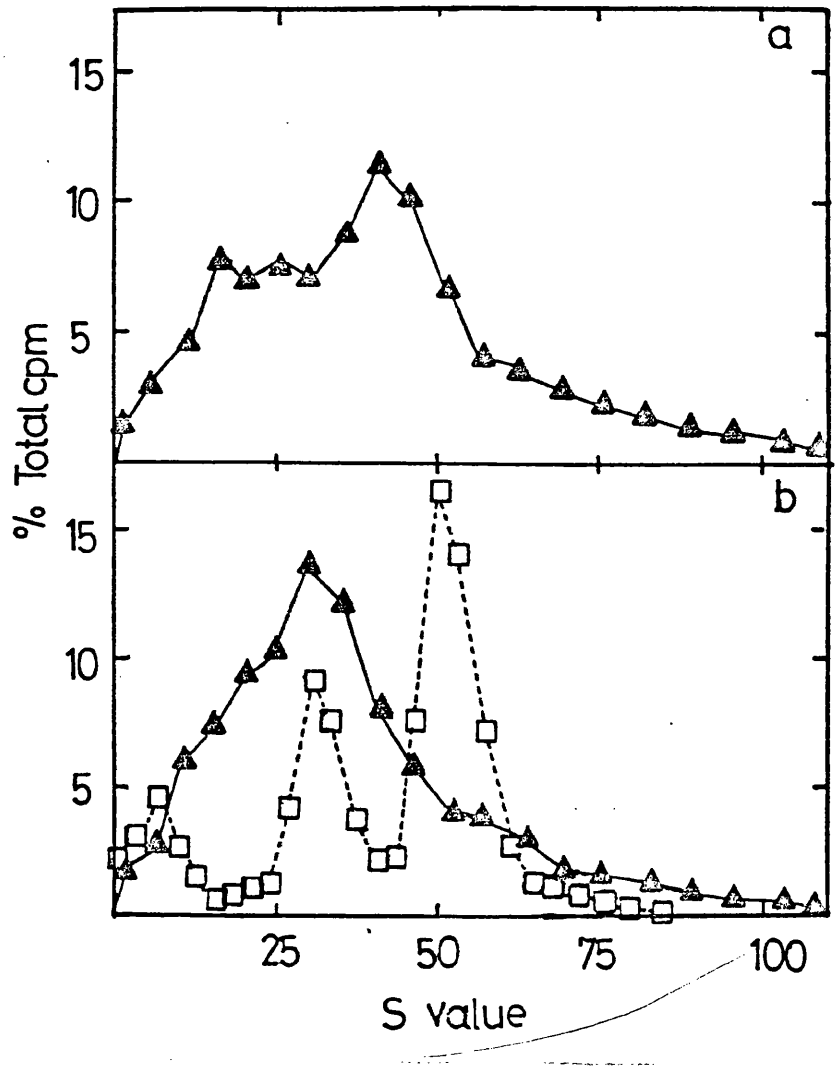


FIGURE 24. Sizing of <u>in vitro</u>-transported RNP by sedimentation on 15-30 percent TNM and NET-sucrose gradients. Centrifugation conditions were as stated in the text. The proportion of acid-precipitable counts in each gradient fraction is plotted against S value (derived from computer analysis).

(a) <u>In vitro</u>-transported RNPs from pulse-labelled nuclei, sedimented on 15-30 percent TNM gradient.

(b) <u>In vitro</u>-transported RNPs from pulse-labelled nuclei, sedimented on 15-30 percent NET gradient (—▲—) and ^{3}H labelled RNPs from HTC cell post-mitochondrial supernatant, dissociated from polysomes by treatment with 10 mM EDTA, acting as size-markers for ribosomal subunits, sedimented on 15-30 percent NET gradient (-□--).

4.8. Hybridisation studies of RNA sequences transported in vitro from HTC cell nuclei

The results presented in the foregoing section indicate conditions under which HTC cell nuclei process and transport hnRNA sequences with the crude properties of mRNA. The present section describes a series of tests of the composition of the heterogeneous (polyadenylated) RNA transported *in vitro*, using some of the hybridisation assays envisaged for such a study in Section 4.6. Attention is focussed on three specific questions. Firstly, does the transported RNA have a qualitative and quantitative composition expected of a population of newly transported mRNAs? Secondly, could this composition be explained as being due to one or other of the principal artifacts suspected of occurring in this system (non-specific release of nuclear RNA by leakage, and release by a non-physiological mechanism of adherent RNA of cytoplasmic origin), or by a simple combination of these? Thirdly, if artifactual explanations can be eliminated, what implications would there be for the levels of gene control in HTC cells? The speculative nature of this last question must be stressed, since conclusive proof of *in vivo* equivalence for mRNA processing and transport *in vitro* requires firm information on the relative contributions of intranuclear and postnuclear events to post-transcriptional abundance modulation, which will require exhaustive *in vivo* studies using cloned cDNAs outside the scope of this thesis. All the present study can do is assess the importance of artifacts, and say whether, on the basis of hybridisation studies of steady-state nuclear and polysomal RNA populations, the composition of *in vitro*-transported RNA could be the result of physiological selectivity.

4.8.1. Isolation of poly(A)$^+$ RNA transported in vitro

Unlabelled HTC cell nuclei were incubated under the standardised conditions (Section 3.14) already referred to, following a pre-incubation as envisaged in Section 4.7.2. This pre-

incubation was carried out in the complete incubation medium, with HTC cytosol at 2.4 mg/ml in protein. After chilling on ice for 30 seconds, the nuclei were recovered by centrifugation for 2 minutes at 4°, at 2000 rev/min (1000 g_{max}) in the MSE 4L centrifuge, and resuspended in fresh medium of identical composition. Incubation was for a further 75 minutes at 30°. RNA was recovered from the post-nuclear supernatant by phenol-chloroform extraction (Section 3.3) and ethanol precipitation and passed twice through oligo-(dT)-cellulose (Section 3.4.3). In two separate preparations, 26 µg and 24 µg of poly(A)$^+$ RNA were recovered from 4 x 10^9 cells. As a control, the same amount of cytosol used in these incubations, when phenol-chloroform extracted in the presence of an equivalent amount of RNA carrier (E.coli RNA), yielded an undetectable amount of poly(A)$^+$ RNA (< 1 µg). The 'in vitro-transported' RNA recovered in the experiment must therefore have originated in the nuclei. The yield of RNA, as well as being reproducible, is in accord with that predicted on the basis of the kinetic experiments with pulse-labelled nuclear RNA (Section 4.7, Fig. 20, and Table 7).

4.8.2. Homologous cDNA hybridisation kinetics of poly(A)$^+$ RNA transported in vitro

Poly(A)$^+$ in vitro-transported RNA, prepared as above, was reverse transcribed in the presence of ^{32}P-labelled dCTP giving a yield of approximately 5 percent by weight. This compares with yields of 1.5-2 percent for poly(A)$^+$ nuclear RNA, and 10-20 percent for poly(A)$^+$ polysomal RNA. This 'in vitro-transported' (ivt) cDNA was hybridised with its template RNA, and the extent of hybridisation determined by resistance to S1 nuclease digestion (Fig. 25). Rot analysis (Table 8) indicates a somewhat more restricted range of relative frequencies than is evident in polysomal RNA, with a 2 component or 3 component solution being equally compatible with the data. The small number of data points

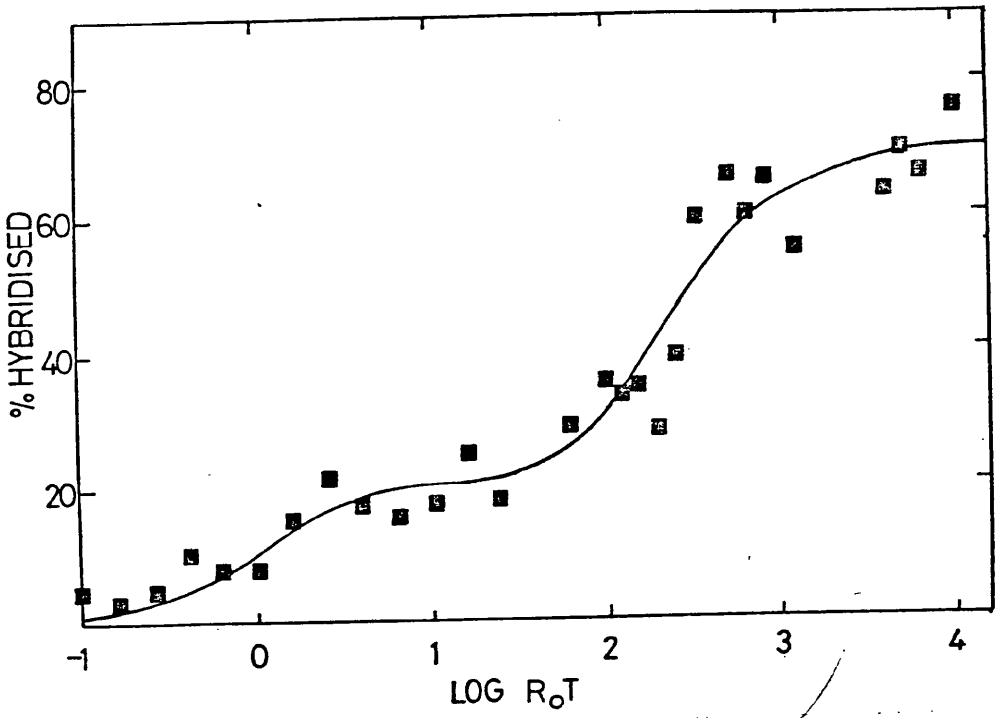


FIGURE 25. Homologous hybridisation kinetics of cDNA to *in vitro*-transported RNA, with its template RNA. Reactions were carried out at RNA concentrations of 10 µg/ml, 1 mg/ml and 10 mg/ml, in HB/S at 70° (cDNA was ^{32}P-labelled, and not purified on an alkaline sucrose gradient). Its hybridisability against sonicated rat DNA was 70 percent. Poly(U) was added to reactions as previously. S1 background (2 percent) was subtracted from all data points.

TABLE 8. Kinetic estimate of complexity and abundance components of *in vivo*-transported RNA.

Component	$R_o t_{1/2}$ (a)	Percentage cDNA (b)	Corrected $R_o t_{1/2}$ (c)	Complexity (nucleotides)
I (abundant)	1.1	29	0.32	1.0×10^5
II (intermediate)	180	57	100	3.3×10^7
III (rare)	1600	14	220	7.2×10^7 (d)

All data quoted to 2 significant figures.

(a) derived by resolving curve in Fig. 25 into kinetic components.

(b) proportion of total cDNA represented in each component.

(c) product of (a) and (b), taking account of the fact that only part of the mass is driving the hybridisation of each hypothetical component.

(d) roughly twice the kinetic complexity of polysomal poly $(A)^+$ RNA (Table 3). However, experimental scatter and small number of data points make error in this estimate very large. ($\pm$ 200 percent).

means that there is great uncertainty in the estimates of overall complexity. However, it is clear that the *in vitro*-transported RNA comprises a complex population of sequences comparable in diversity with that of polysomal RNA, and exhibiting a wide span of relative abundances.

4.8.3. Heterologous hybridisations of 'in vitro-transported' cDNA with HTC cell nuclear and polysomal poly(A)$^+$ RNAs

Since the amount of *in vitro*-transported poly(A)$^+$ RNA which is obtained from a large number of nuclei is relatively small, it is desirable to apply hybridisation assays of its composition which are as inexpensive as possible in the use of this RNA. One very convenient approach is to use the complementary DNA to this RNA (ivt cDNA) to probe the relative content in nuclear and polysomal poly(A)$^+$ RNAs of the sequences transported *in vitro*. Considering the data already presented (Section 4.3) on the cross-hybridisation of nuclear and polysomal cDNAs of rat liver, which are, as will be seen, applicable also to HTC cells, it is evident that the hybridisations of ivt cDNA can contribute two important pieces of information. Firstly, the extent of the reaction indicates whether the ivt cDNA contains poly(A)-adjacent nucleus-confined sequences: nuclear cDNA, which contains such sequences, is incompletely hybridised by polysomal RNA, and a clear plateau level is reached, whereas polysomal cDNA is hybridised to at least as great an extent by polysomal as by nuclear RNA. Secondly, the relative kinetics of the two reactions are a measure of how far the quantitative composition of the *in vitro*-transported RNA is determined by the operation of the post-transcriptional controls which are responsible for the altered composition of polysomal (as opposed to nuclear) steady-state RNA *in vivo*. Whereas nuclear cDNA is hybridised by nuclear or polysomal RNAs with comparable kinetics, polysomal cDNA is hybridised much more slowly (1-2 orders of magnitude, depending on the cell-type) by nuclear RNA than by polysomal RNA.

Hybridisation kinetics of ivt cDNA, driven by nuclear and polysomal poly(A)$^+$ RNAs, are shown in Fig. 26, alongside the equivalent curves for HTC cell nuclear and polysomal cDNAs (extracted largely from previous data). The data may be analysed in various ways. It is clear that all sequences in the probe are represented in polysomal RNA as well as in nuclear RNA. The concentration of nucleus-confined poly(A)-adjacent sequences is therefore below the limits of detection in the S1 nuclease assay. Since a difference in hybridisation level of about 3 percent is the smallest that could reliably be detected using this assay, and since the difference in hybridisation levels of nuclear cDNA with nuclear and polysomal RNAs is about 30 percent, it would appear that the qualitative composition of the poly(A)-adjacent sequences in <u>in vitro</u>-transported RNA resembles that of polysomal RNA, with at most, a 10 percent contamination with non-specifically leaking nuclear sequences.

The kinetic disparity in the nuclear and polysomal RNA-driven reactions is about one order of magnitude. The average concentration of ivt cDNA sequences in both nuclear and polysomal RNAs being intermediate between that of nuclear and polysomal cDNAs. Since detachment of adherent RNA of cytoplasmic origin may conceivably account for some (or all) of the transported sequences which are not leaking non-specifically from isolated nuclei, it is important to estimate the effect of this on the kinetics of cross-hybridisation as shown in Fig. 26. If it is assumed that non-specific leakage accounts for 10 percent of the <u>in vitro</u>-transported poly(A)$^+$ RNA, then if detachment of adherent RNA accounted for the remainder, the concentration in the resulting mixture, of an 'average' polysomal sequence represented in nuclear RNA at 1 percent of its polysomal abundance, would still be about 90 percent of its level in the polysomes. The curves shown in Fig. 26b rule out such a possibility, since the average concentration of <u>in vitro</u>-transported sequences is

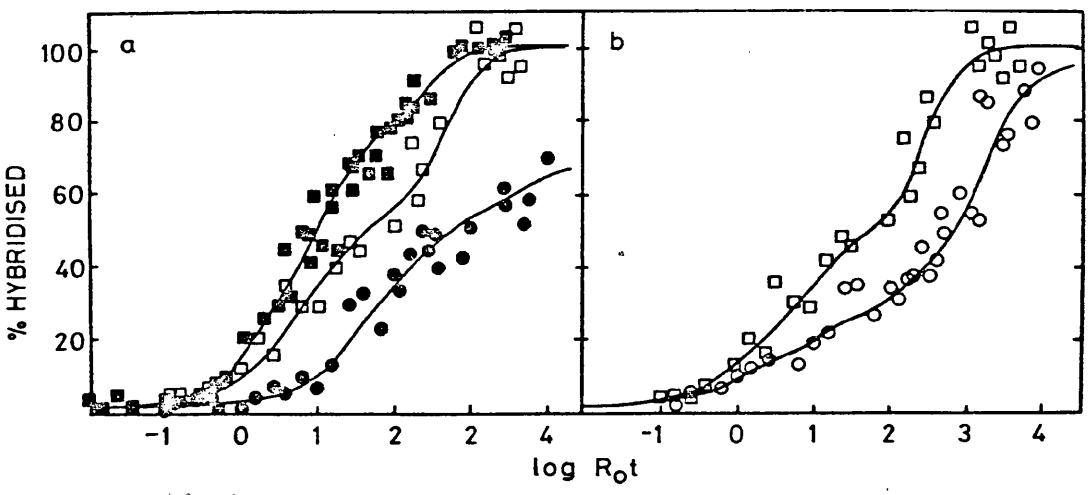


FIGURE 26. Hybridisation characteristics of HTC cell nuclear, polysomal and in vitro-transported cDNAs with HTC cell poly(A)$^+$ nuclear and polysomal RNAs. Poly(U) was present, as previously, in reactions which were carried out in HB/S at 70°. All data are normalised for probe hybridisability and S1 backgrounds have been subtracted.

(b) ivt cDNA hybridised with poly(A)$^+$ nuclear (O) and poly(A)$^+$ polysomal (□) RNAs at 10 µg/ml, 1 mg/ml and 10 mg/ml. RNA excess at the highest $R_o t$ values was 30,000 fold, and at the lowest $R_o t$ values was 1070-fold (high specific activity cDNA labelled with ^{32}P was used).

(a) ivt cDNA (□), nuclear cDNA (●), and polysomal cDNA (■) hybridised with 10 µg/ml, 100 µg/ml, 500 µg/ml (polysomal cDNA only), 5 mg/ml (polysomal cDNA only) and 10 mg/ml HTC cell polysomal poly(A)$^+$ RNA. The former data are reproduced from Fig. 26a, the latter data from Fig. 7d. RNA-excess in the hybridisation of HTC nuclear cDNA, which was also ^{32}P labelled, ranged from 1070-fold to 30,000-fold.

only 10-fold, not 90-fold greater in polysomes than nuclei. The data suggest that the contribution of detached adherent RNA to the poly(A)$^+$ RNA transported *in vitro* is also unlikely to exceed 10 percent, as predicted from the kinetics of RNA release considered in relation to the use of a pre-incubation step. Moreover, this upper limit is calculated on the assumption that 'non-spuriously' transported RNA still has the pattern of relative abundances of steady-state hnRNA, an assumption which also proves to be untenable in the light of studies using cloned cDNAs (see Section 4.8.5). Therefore, the actual contribution of adherent RNA release may be considerably less than 10 percent.

The composition of poly(A)$^+$ RNA transported *in vitro* from HTC cell nuclei cannot be accounted for by a simple combination of artifacts. Furthermore, its pattern of relative abundance would appear to be intermediate between that of steady-state nuclear and polysomal RNA. Despite the possible contributions of a low-level of non-specific leakage, and of adherent RNA release, *in vitro* mRNA transport does appear to be sequence selective, excluding those poly(A)-adjacent sequences which are normally confined to the nucleus, and boosting the relative levels of sequences destined to be at high abundance on the polysomes, processing and transporting them preferentially over sequences which are at more similar abundances in nuclei and polysomes. This does not prove, however, that such selectivity is genuinely physiological. This point will be returned to in discussion. If the selectivity of processing and transport *in vitro* does resemble that which pertains in vivo, then it would appear that intranuclear events cannot account entirely for post-transcriptional modulation of mRNA abundances. At least some of the higher abundance HTC messengers would also need to be selectively stabilised in the cytoplasm. The extent to which this might reflect different levels of control of different (sets of) mRNAs will also be considered subsequently (Sections 4.8.5 and 5.5).

4.8.4. Hybridisation of in vitro-transported RNA with nucleus-confined single-copy DNA

An alternative strategy for assessing the qualitative composition of *in vitro*-transported RNA, and one which is highly sensitive to the presence of sequences normally restricted to the nucleus, is the use of a single-copy DNA probe enriched for nucleus-confined sequences by successive cycles of hybridisation and fractionation. Rat liver single-copy DNA was therefore 'gap'-translated (Section 3.5.2) in the presence of ^{32}P-labelled dCTP to a specific radioactivity of 10^7 cpm/μg DNA, and hybridised with a 50-fold excess of mercurated HTC cell poly(A)$^+$ polysomal RNA to $R_o t$ = 10,000 moles.s.l^{-1}. Unhybridised material was recovered by thiol-sepharose chromatography, and re-hybridised with a 100-fold excess of mercurated HTC cell poly(A)$^+$ nuclear RNA to $R_o t$ = 12,500 moles.s.l^{-1}. The recovered, hybridised material was reacted analytically with a large excess each of HTC cell nuclear, polysomal and *in vitro*-transported poly(A)$^+$ RNAs, and the extent of hybridisation assayed by S1 nuclease resistance (Fig. 27). The very poor hybridisability (with nuclear RNA) was due partly to the incomplete nature of the fractionation, and also to shortening of the probe as a result of radioactive disintegrations. Nevertheless, the results indicate clearly that the probe is specific for nucleus-confined sequences (the reaction with polysomal RNA being no more than the S1 background), and that such sequences are detectable in *in vitro*-transported RNA, although they represent no more than about 30 percent of the nucleus-confined complexity.

This can be readily understood in terms of the size of *in vitro*-released heterogeneous RNA (Fig. 23). Since this is comparable with that of messenger RNA, and since the bulk of nucleus-confined complexity is probably to be found in intron sequences very distal to poly(A), poly(A)$^+$ RNA transported *in vitro*, even by non-specific leakage, is unlikely to contain a significant proportion of this poly(A)-distal

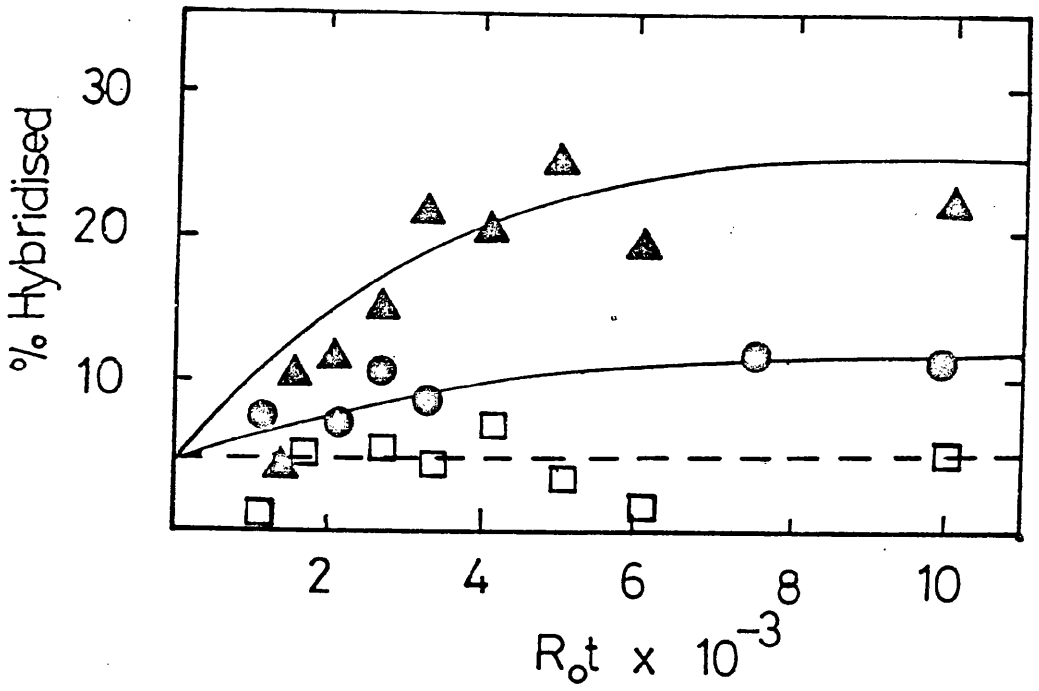


FIGURE 27. Hybridisation of nucleus-confined single-copy DNA with HTC cell poly(A)$^+$ nuclear (▲), polysomal (□) and *in vitro*-transported (○) RNAs, all at 10 mg/ml in HB/S at 70°. Analysis by S1 nuclease. S1 background, not subtracted from data, shown as dashed line (5 percent). RNA excess was 6,700-fold in all cases.

complexity. A homogeneous population of sequences, of the complexity determined for the nucleus-confined information detected in *in vitro*-transported RNA (i.e. about 9×10^7 nucleotides) should hybridise single-copy rat DNA with first-order kinetics, some 7×10^4 times slower than the homologous hybridisation of globin cDNA, i.e. with a $R_o t_{1/2}$ of about 280 moles.s.l^{-1}. The hypothesis that non-specifically leaking nucleus-confined sequences account for 10 percent or less of the mass of *in vitro*-transported RNA would increase the observed $R_o t_{1/2}$ for the hybridisation of nucleus-confined single-copy DNA by *in vitro*-transported RNA by a factor of 10, to about 2800 moles.s.l^{-1}. Unfortunately, the data of Fig. 27 are insufficient to permit an accurate assessment of how realistic is this estimate, although it does not appear to be wildly in error.

Non-specific leakage appears therefore to be occurring at a low level against a background of specific mRNA transport, as inferred from cDNA studies. It remains, however, to demonstrate whether the specificity of RNA transport *in vitro* is similar to that which pertains *in vivo*. A pilot study, using cloned cDNAs, which is described in the following section points the way towards how this may be accomplished.

4.8.5. Effect of cytosol on specificity of mRNA transport in vitro: hybridisations using heterogeneous and cloned cDNA probes for post-transcriptionally regulated sequences

Previous investigations have indicated that the spectrum of repetitive sequence transcripts transported from hepatocyte nuclei under similar conditions to those employed here, depends on the source of cytosol used, and that the patterns of post-transcriptional sequence selection are similar to those which operate in the cell-type of origin of the cytosol (389, 537). The results presented in Sections 4.1 - 4.5 indicated that as far as poly(A)$^+$ mRNA sequences are concerned, only

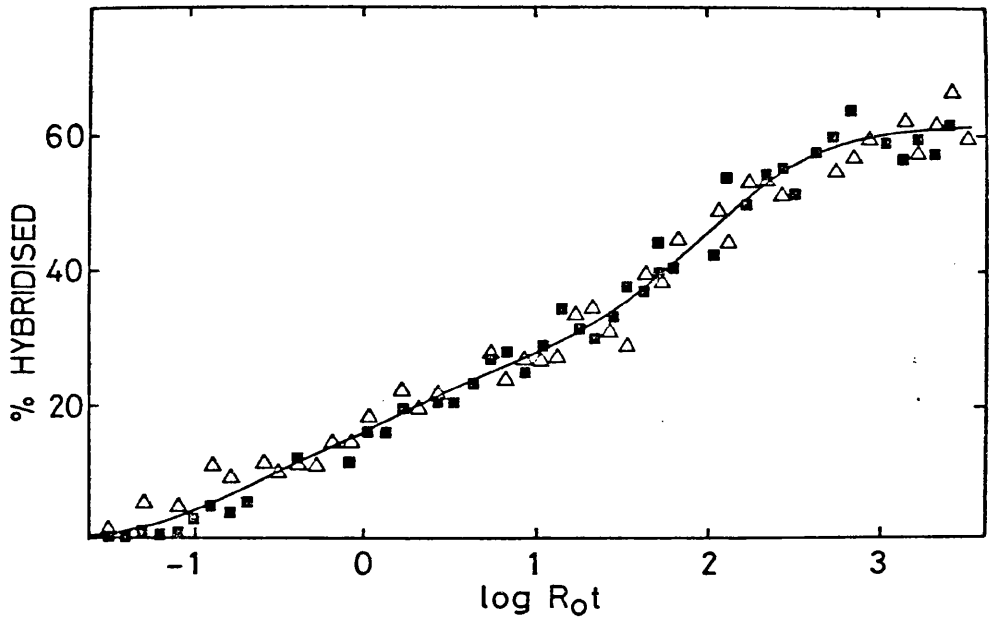


FIGURE 28. Hybridisation of abundant rat liver polysomal cDNA with <u>in-vitro</u>-transported poly(A)+ RNAs obtained using liver (Δ) and HTC cell (■) cytosols. RNA concentrations were 50 µg/ml, 1 mg/ml and 5 mg/ml (RNA excess of 2200-fold at highest $R_o t$ values), and reactions were carried out in HB/F at 43°. S1 background was subtracted, and values normalised for probe hybridisability (67 percent). Poly(U) was included, as previously.

quantitative post-transcriptional changes are detectable between liver and hepatoma cells. The results presented in section 4.7 and the remainder of Section 4.8, indicated that despite low level non-specific release of total nuclear and adherent cytoplasmic RNA, the population of mRNA sequences transported *in vitro* from isolated HTC cell nuclei differs in its relative composition from both steady-state nuclear and polysomal RNA, implying sequence-specific processing and transport. If *in vitro* mRNA transport has the same selectivity as *in vivo* mRNA transport, then a contribution of post-nuclear modulation to polysomal abundances is also implied by these data . Since *in vivo* studies have provided hybridisation probes for the sequences most dramatically regulated post-transcriptionally between rat liver and HTC cells (Section 4.6), it is possible to determine whether post-transcriptional mRNA sequence selectivity in HTC cell nuclei *in vitro* is subject to the same cytosol-mediated regulation as observed previously for repetitive sequence transcripts. If so, a direct approach is possible to the mechanisms of this regulation.

HTC cell nuclei, prepared as previously, and pre-incubated in homologous cytosol to eliminate the bulk of adherent RNA, were incubated for 75 minutes in media containing either HTC cell or rat liver cytosol, at the same protein concentration, and which were otherwise identical. Yields of *in vitro*-transported poly(A)$^+$ RNA from the parallel incubations were 24 µg and 21 µg, each from approximately 4 x 10^9 nuclei. The only other observable difference was that aggregation of nuclei was somewhat less in the presence of liver cytosol, and no macroscopic aggregates formed. *In vitro*-transported RNAs were hybridised in excess with abundant liver polysomal cDNA, which represents largely a population of sequences under post-transcriptional modulation between the cell-types. Fig. 28 demonstrates that the hybridisation kinetics of the probe with both samples of *in vitro*-transported RNA are the same, and not unlike the kinetics of hybridisation of the

same probe with poly(A)$^+$ polysomal RNA from HTC cells (Fig. 12a). Several interpretations of this finding are possible, and will be formulated subsequent to the presentation of results using cloned cDNAs.

^{32}P-labelled ivt cDNAs were prepared from the RNAs transported in the presence of each of the cytosol preparations. These highly labelled probes were then hybridised with filter-bound recombinant plasmid DNAs under conditions identical with those used previously (see Section 4.5) - in fact, the hybridisations were carried out simultaneously. The results are summarised in Table 9, together with relevant data abstracted or derived from Table 6. Again, no significant cytosol-induced specificity is evident. The mass fractions in ivt cDNA of the two recombinants corresponding to abundant liver mRNAs post-transcriptionally depleted in hepatoma, pRR 83 and pRR 5B, are less than those found previously in steady-state nuclear cDNA. Post-transcriptional depletion of these mRNAs in HTC cells therefore seems to be reproduced in vitro, regardless of the source of cytosol. Further inspection of the levels of the clones in ivt cDNA, in comparison with their levels in steady-state nuclear and polysomal cDNAs, reveals profound differences in the relative in vitro-processing and transport efficiencies of sequences at similar polysomal abundances. Furthermore, they indicate once more that the composition of in vitro-transported cDNA cannot be explained as a simple mixture of steady-state nuclear and polysomal RNAs, arising from different artifacts. Clone pRR 117, for example is at a mass fraction in ivt cDNA which could be accounted for, roughly, by a 2:1 mixture of polysomal and nuclear RNA. However, the mass fraction of clone pRR 133 could only arise if the relative proportions were more like 1:20. This supports the conclusions already derived from Fig. 26, although it does not constitute proof

TABLE 9. Studies of relative mRNA abundances in poly(A)$^+$ RNAs transported from HTC cell nuclei *in vitro* in the presence of liver and HTC cell cytosols, using selected recombinants from a liver polysomal cDNA library.

DESIGNATION	pRR 83	pRR 5B	pRR 117	pRR 133
Mass fraction in ivt cDNA (HTC cytosol) = I	0.0092	0.025	0.21	0.056
Mass fraction in ivt cDNA (liver cytosol) = II	<0.02 (c)	<0.04 (c)	0.17	0.065
Mass fraction in HTC mRNA = III(a)	0.0072	0.018	0.27	0.63
Mass fraction in HTC hnRNA= IV(a)	0.014	0.034	0.042	0.033
Ratio I/III (b)	1.3	1.4	0.78	0.089
Ratio I/IV (b)	0.66	0.74	5.0	2.0
Abundance ratio liver/HTC mRNA(a)	62	84	0.22	0.69

All values quoted to 2 significant figures. SE's in mass fractions $\pm$ 18% computed as in Table 6.

(a) taken from Table 6.

(b) as a measure of how far the abundance of each of these sequences in ivt cDNA resembles that in either mRNA or hnRNA.

(c) not significant above background in t-test (low number of input counts).

All mass fractions are percentages.

that the selectivity of processing and transport *in vitro* is actually physiological. If the observed specificity were to be regarded as being physiological, however, it would imply that different sequences are raised to high abundance by quite independent mechanisms, with different relative contributions of transcription, processing/transport, and cytoplasmic stability.

5. DISCUSSION

5.1. Methodological considerations

5.1.1. Limits of interpretation of hybridisation experiments

The different types of information which can be derived from saturation and kinetic hybridisation methods have already been referred to. The results serve to illustrate the importance of using both approaches in order to characterise RNA populations adequately. The use of either method alone can lead to serious misinterpretations. For example, the relatively small proportion (by weight) of nuclear poly(A)-adjacent sequences which are absent from, or substantially rarer in polysomes (see Figs. 8 and 10) contrasts with the very large differences between poly(A)$^+$ nuclear and polysomal RNAs in overall sequence complexity (Figs. 4,5). Conversely, the very small proportion, if any, of polysomal sequences specific to either cell-type contrasts with the large differences in relative mRNA abundances, affecting a considerable proportion of the polysomal sequences by weight (compare Figs. 4 and 11). The results shown in Figs. 26 and 27 demonstrate the necessity of using both methods in judging the composition of *in vitro*-transported RNA, obtained by a novel procedure. The kinetic studies alone might suggest a complete absence of contamination with non-specifically leaking nuclear RNA, whereas consideration of the single-copy DNA saturation experiment (Fig. 27) in isolation would lead to an erroneous supposition of its severity. Radical biological conclusions drawn from the use of just one of these methodologies, of which there are many examples in the literature, should therefore be treated with the utmost caution. In addition, as shown in Fig. 8b, and as suggested by studies of cloned cDNAs in comparison with the data of Figs. 11-12, apparent qualitative differences between RNA populations may be due, in practice, to large differences in relative abundance: estimates of homology between RNA populations are therefore

very sensitive to the amount of RNA-excess used to drive cDNA cross-hybridisations. In the absence of clear evidence that sufficient RNA-excess has been used to drive the reaction to a definitive plateau, estimates of homology based upon this method should be regarded as uncertain.

Four further points should be noted. Firstly, considerable disparity in saturation and kinetic estimates of complexity derived here (see Tables 2-3) is similar to that found by many other investigators (Table 1) and predicted on theoretical grounds by Kiper (183). However, the large percentage error in kinetic complexity estimates, deduced from computer analyses, coupled with the uncertainties in the absolute saturation estimates inherent in the nature of the probe and the assay method, mean that neither estimate should be regarded as being 'accurate' to within more than about half an order of magnitude. Furthermore, it is now clear that functionally distinct mRNAs can share portions of their nucleotide sequence (420), so the precise number of different mRNA species implied by these estimates of complexity remains uncertain.

Secondly, the use of cloned cDNA probes illustrates the arbitrary nature of the division of mRNA populations into abundance classes, on the basis of curve-fitting to the kinetics of heterogeneous cDNA hybridisation reactions. Despite this, the evidence from cDNA cross-hybridisations, for groups of mRNA sequences which change in abundance between intracellular compartments or between cell-types in a similar way, is supported by studies using cloned cDNAs. This supports the interpretation of such changes in terms of shared phenotypic specification, for example, as genes coding for specialist secretory products or growth-related functions, whose abundances appear to be co-ordinately regulated, although not necessarily through a common mechanism.

Thirdly, it is clear that subfractionation of heterogeneous cDNAs into 'abundant' and 'rare' components can aid in the dissection of relative abundance patterns and how these change between RNA populations. Ideally, further subfractionation could yield even more detailed information. However, this is a labour of diminishing returns due to the progressive loss of hybridisability, especially of fractions successively enriched for rare sequences, coupled with the uncertainty attached to the proportion of the RNA actually driving the reaction, which renders a strict mathematical analysis impossible. However, an unbiassed sampling of a sufficiently large number of recombinants from a cDNA library makes such procedures redundant.

Fourthly, as has already been indicated, the small number of recombinants studied here, although not constituting such an unbiassed sample, do illustrate that the patterns of relative abundances, and the shifts which they experience between nucleus and polysomes on the one hand, and different cell-types on the other, inferred on the basis of heterogeneous cDNA hybridisations, do correspond with the properties of real mRNAs, and are not the result of methodological artifacts or a spurious averaging process.

5.1.2. Suitability and validity of the liver-hepatoma comparison

At this point it is appropriate to re-emphasise that the choice of biological system was based on its likely applicability to the development of an *in vitro* system for studying post-transcriptional controls. The large number of parameters along which the cell-types compared were known to differ was regarded as an advantage from this viewpoint, even though it makes the interpretation of *in vivo* studies in relation to phenotype that much more hazardous. In particular, it should be stressed that at no point is this system advanced as a

model for understanding the mechanisms of carcinogenesis or differentiation. The results do, nevertheless, have considerable bearing on theories of how gene expression is related to each of these processes. These aspects will be discussed in a subsequent section (5.7). In the first instance, however, it is important to raise the question of exactly which cell-types are being compared in this study, in order to permit a meaningful assessment of the results.

The HTC line is only one of a very large number of transplantable rat hepatocarcinomas which could have been used in this study. As a minimum-deviation tumour (610), it was considered particularly suitable, since the possession of specific differentiation-related markers (587) indicates with certainty the cell's lineage. Furthermore, the phenotypic differences from the normal differentiated cell are sufficiently undramatic so as to make it unlikely that altered gene expression in the tumour is the result of a gross lesion, or a large number of multiple, unrelated changes. The passage of the tumour in cell culture may, however, have caused some deviations of this type, particularly mutations leading to an enhanced rate of proliferation (thus conferring selective advantage) and abnormalities in gene dosage. HTC cells exhibit glucocorticoid inducibility of tyrosine aminotransferase (587) at the level of mRNA (632). However, unlike some other hepatomas, such as Morris hepatoma 7777 (438, 611) HTC cells are in a demonstrably less differentiated (less specialised) state than normal hepatocytes, since they have a very low level of albumin synthesis (612), and of marker enzymes of the adult liver, such as tryptophan oxidase (612). In these respects HTC cells recall the pattern of gene expression of immature hepatocytes. These phenotypic features of HTC cells correlate well with the levels of specific translatable mRNAs (612,632). Various observations suggest that hepatic gene expression generally is independent of gross translation level controls. The levels of albumin and α-fetoprotein

synthesis (436-439, 611), and the activities of many liver enzymes, such as PEP carboxykinase (613) are correlatable with the levels of their particular mRNAs during development (613), hepatocarcinogenesis (438,439,611) and hormone administration (436,437), although Ferritin (614) and $\alpha 2_u$-globulin (604) are subject to fine translational controls. The dramatic increase in protein synthesis which occurs during liver regeneration is due to increased mRNA synthesis, and does not involve preferential recruitment of informosomal mRNAs into polysomes (615). Increased mRNA levels in both regenerating liver (615) and hepatomas (616) compared with normal liver have been attributed (at least in part) to increased post-transcriptional conservation at the levels of processing, transport, or cytoplasmic turnover, and these same mechanisms may result in the modulation of mRNA abundances inferred from the in vivo studies. Therefore, the HTC line, in comparison with normal rat liver, appears to be an appropriate system for studies of post-transcriptional, but pre-translational regulation of gene expression, leading to demonstrable changes in phenotype involving both proliferation and cell specialisation.

The degree to which cell culture conditions may have influenced the pattern of gene expression in HTC cells, thus biassing the results, is difficult to assess. However, recent reports of alterations in the mRNA population between normal liver, carcinogen-treated liver, carcinogen-induced hepatomas in vivo and ascitic hepatomas cultured in vivo (617,146,150,618) lead to conclusions strikingly similar to those reported here. Although the hepatoma lines used in these studies were different from the one used here, they nevertheless exhibit depletion of abundant liver mRNAs in the hepatomas, in the absence of large changes in the overall mRNA complexity. Changes induced by carcinogen treatment alone are effectively those of liver regeneration (resulting, presumably, from acute toxic effects). Unless the similarities in the various findings are purely coincidental, gene expression in hepatomas would appear to be largely unaffected by cell culture conditions

or by secondary physiological changes in carcinogen-treated or tumour-bearing rats. Although this leaves unanswered the question of whether cultured HTC cells can be regarded simply as transformed hepatocytes, it nevertheless indicates that the observed changes in the pattern of mRNA abundances between liver and HTC cells are likely to be due to differences in their proliferative capacity, degree of functional specialisation and neoplastic (as opposed to normal) state, not to growth in culture.

A more serious objection arises, however, from the heterogeneity of cell-types in the liver, since only about 70 percent of the cells are in fact hepatocytes (619) although they constitute 90 percent of the cell mass. Some of the abundant liver mRNAs depleted in HTC cells could, conceivably have been contributed by a minority of non-hepatocytes. However, if these accounted for a substantial proportion of the depleted (low complexity) messengers, then the affected mRNAs would have to account for almost the entire mRNA complement of the cell-type(s) concerned. Such a situation has only previously been observed in terminally differentiated cells such as reticulocytes (261), and there is no obvious sub-population of liver cells with such characteristics. Blood cells can effectively be excluded since all liver preparations were perfused <u>in vivo</u> prior to excision and homogenisation. The procedure employed for tissue homogenisation and polysome preparation, being particularly gentle, probably selects heavily for hepatocytes as against other cell-types. This assumption is supported indirectly by the results of single-copy DNA hybridisations (Figs. 4-5): the effectively complete overlap of expressed sequences in the two cell-types, both in nuclear and polysomal poly(A)$^+$ RNA, contrasts with the observation that cell-types of different lineages do express a detectable proportion of tissue-specific RNAs (see Section 2.3.5).

5.2. Metabolic relationship of nuclear and polysomal poly(A)$^+$ RNAs

Evidence has been presented in support of qualitative and quantitative post-transcriptional selection of mRNA sequences, on the basis of nuclear and polysomal cDNA cross-hybridisations (Section 4.3) in rat liver. In the light of the comparative study of liver and hepatoma mRNA sequences (Sections 4.4 and 4.5) it is clear that there is no detectable change in qualitative aspects of post-transcriptional selection between the cell-types. Effectively the same sequences are expressed on liver and HTC cell polysomes (Fig. 4) and almost all the mass of poly(A)-adjacent sequences in HTC cell mRNA is represented in liver polysomes (Fig. 11). Nucleus-confined information accounts for similar proportions of nuclear complexity, and similar proportions by weight of the nuclear poly(A)-adjacent RNA in the two cell-types (Figs. 8 and 26). As in the case of HeLa (215) or mouse Friend cells (346) these probably represent a class of low abundance transcripts of complexity greater than that found adjacent to poly(A) in polysomal RNA. The failure to detect qualitative changes in the composition of nucleus-confined sequences between liver and hepatoma does not rule out the possibility that some, at least, of these sequences have coding potential, and that they are expressed in other cell-types. It does support the view, however, that qualitative post-transcriptional selection, of the type observed to operate between different tissues or developmental stages in sea urchins, frogs, and tobacco (178, 177, 186), is not responsible for differential gene expression in mammalian cells in different states of cell proliferation or specialisation.

The reported expression, in regenerating liver, hepatomas, and livers of carcinogen-treated rats, of repetitive sequence transcripts which are confined to the nucleus in normal liver (388,389,617) is not, therefore, applicable to poly(A)$^+$ mRNAs transcribed from single-copy genes. Knöchel *et al* (617)

have reported that the cellular relocation of these transcripts is still detectable in livers of carcinogen-treated rats, where there is concomitantly only a minimal quantitative alteration in the poly(A)$^+$ mRNA population. There is no reason, thus, to cast doubt on the earlier observations, which may have some specific relevance to the promotion of proliferation or regeneration. However, they are unlikely to involve protein-coding sequences directly, unless these are confined to the non-polyadenylated compartment. Histone mRNAs could conceivably be included in this category, although their presence in the nuclei of quiescent cells is uncertain.

The precise nature and role of nucleus-confined poly(A)-adjacent sequences remain to be elucidated. If they are not, in fact, potential coding sequences, they may represent spuriously or non-specifically polyadenylated transcripts which are not recognised as mRNA precursors, and hence degraded within the nucleus (375). Alternatively, the 3' (poly(A)-proximal) regions of the primary transcripts of many genes may be intron-like, and spliced out during processing, relocating poly(A) next to the 3' end of the mature mRNA sequence. Such a mechanism would explain the paradoxical conservation of poly(A) from nucleus to cytoplasm, despite the demonstrable presence of poly(A)-adjacent nucleus-confined sequences of great complexity. Whether such sequences may play a role in the efficiency or selectivity of processing remains an open question.

An altered pattern of quantitative post-transcriptional selection between liver and hepatoma is implied by the cDNA hybridisation data presented in Sections 4.4 and 4.5. As already noted, the cross hybridisation of nuclear and polysomal cDNAs (Figs. 8-9) reveals that in liver, mRNA sequences are in a somewhat more restricted frequency range in hnRNA than mRNA. The postulate that this is due principally to the

preferential post-transcriptional stabilisation of the most abundant class of mRNAs is supported by the use of cloned cDNAs (Table 6). This shows that the disparity in nuclear and polysomal abundances is greatest for the highly abundant mRNAs represented in clones pRR 83, pRR 5B and pRR 133. An mRNA of more modest abundance, such as pRR 117, can be at a similar level in steady-state nuclear RNA as those which ultimately are to be raised to very high abundance on the polysomes. Special mechanisms operating at multiple levels (transcription, processing, transport, stability) to promote the relative abundance of differentiation-related mRNAs are a feature common to many biological systems, as has already been noted (Sections 2.3.4 and 2.3.6). The findings reported here seem to conform to this general pattern, especially if the observed sequence-selectivity of in vitro RNA transport is in fact physiological. Such an arrangement enables massive disparities in abundance to be built up by the compounding of processes which in themselves are only modestly selective. The advantages of this, in evolutionary terms, can only be speculated about: clearly an extra element of flexibility is introduced, allowing the rate of 'luxury' protein synthesis to respond to a number of independent variables in the physiological status of the organism, which may be acting over different time-scales, without placing an undue restraint on the evolution of the regulatory properties of the enzymes of RNA metabolism and other factors involved in the control of gene activity.

A further important finding of this work is that some relatively abundant sequences adjacent to poly(A) in nuclear RNA are found at a much lower concentration in polysomal RNA (see Figs. 8 and 10). Such sequences, under negative post-transcriptional selection in rat liver, may be hypothesised to represent mRNAs whose expression needs to be rapidly activated under certain circumstances. Relaxation of this negative control would then permit their rapid accumulation.

The function of such a set of genes is unknown, although obvious candidates would be those involved in cell proliferation or in the repair of cellular integrity (such as the heat-shock genes). The 'accessibility' of such mRNAs might be crucial in the response to cellular damage, or the induction of a proliferative response following organ damage. Another set of genes which might be controlled in this way are those encoding major metabolic enzymes whose expression is under relatively short-term hormonal control, by glucagon and insulin for example.

Rare polysomal mRNAs may be, in general, slightly less abundant in polysomal than in nuclear RNA, such as is the case for those represented in clones pRR 83 and pRR 5B in HTC cells. This may not be the case for those housekeeping genes, however, whose mRNAs are rare in all cell-types.

5.3. Physiological equivalence of mRNA transport from HTC nuclei in vitro

Previous investigations have used wholly inadequate criteria for judging the degree to which the physiological specificity of mRNA processing and transport is reproduced *in vitro*. In this study, the population of sequences transported from isolated rat hepatoma nuclei has been analysed using hybridisation assays for the presence, and relative abundance, of sequences regulated post-transcriptionally *in vivo*. Two aspects of physiological equivalence have been considered: firstly, whether the levels of different sequences in *in vitro*-transported RNA can be related to those postulated in the population of newly-transported mRNA *in vivo*, in a given cell-type; secondly, whether cytosol from a heterologous cell can influence the pattern of post-transcriptional processing and transport of mRNA sequences regulated post-transcriptionally between the cell-types considered. Before either of these questions can be satisfactorily addressed, it

is necessary to eliminate simple artifactual processes as possible explanations for the observed transport of messenger-like RNA.

The first of these is the possibility that the release of adherent RNA of cytoplasmic origin, which co-purifies with nuclei as an unavoidable contaminant, will swamp out the release of genuinely transported RNA, causing its composition to be judged as falsely similar to that of cytoplasmic RNA. The observation that contaminant-RNA-release is ATP-dependent, and is considerable in volume, means that, despite the use of a pre-incubation step, the possible effect of such an artifact on the results should be carefully scrutinised. Consideration of the results suggests that it cannot account for more than about 10 percent of the transported RNA. However, two uncertainties remain. The first concerns the likely composition of the adherent RNA. If it does not consist of a random set of polysomal sequences, its contribution to RNA transport may have been underestimated. This would be the case if abundant polysomal sequences were depleted in it, due to some sequence-selectivity in its adherence or in its release, or if abundant polysomal sequences were at a lower level in total cytoplasmic than in polysomal RNA, due to translational specificity. The second uncertainty arises from the pre-incubation step itself. If the processing and transport of some abundant polysomal mRNAs were particularly efficient, then a pre-incubation step might leave the population of mRNAs transported in the subsequent long incubation significantly depleted in them. The degree to which adherent-RNA release replenished them in total released RNA would again be likely to have been underestimated. Although very rapid processing times are only documented for ultra-high abundance mRNAs, such as those encoding immunoglobulins in myeloma cells (609), this possibility cannot be entirely discounted. One possible solution to this problem is the use of pre-labelled nuclei. However, there are considerable

problems with this approach also. Firstly, if processing times for abundant mRNAs really are very short in these cells, then the labelling time needed to raise nuclear RNA to sufficiently high specific activity for cDNA-excess hybridisation analysis to be feasible would see significant accumulation of label into efficiently processed mRNAs in the cytoplasm. The release of adherent cytoplasmic RNA would still, therefore, be problematic. A second difficulty is that unlabelled adherent RNA would still contribute to the RNA mass, so that very great DNA excess would be required in heterogeneous hybridisations, which would be exceedingly costly in material. It is also doubtful whether nuclear RNA could be labelled to sufficient specific activity to make accurate detection of moderately rare messenger sequences practical, using the filter-bound DNA technique. Apart from these considerations is the possibility that the labelling regime may distort the normal physiological processing of mRNA, especially as it is necessary to submit the cells to extremely high doses of radioactivity.

The second type of artifact which may act to bias the results is the possible release of hnRNA sequences nonspecifically, due to nuclear lysis or leakage. Despite the evidence that most pre-labelled RNA which is released is messenger-like by criteria of size, poly(A)-content, and association into informosomal particles, it is clear from the results of single-copy DNA hybridisation analysis of transported RNA that some non-specific release is occurring, albeit at a relatively low level. This will tend to make the composition of <u>in vitro</u>-transported RNA, in terms of relative sequence abundances, appear falsely similar to that of nuclear RNA. However, for the reasons already elaborated, the low level of contamination detected from single-copy and cDNA hybridisation techniques is unlikely to have influenced the pattern very profoundly.

The results presented indicate that these artifacts are not responsible for the bulk of *in vitro* RNA transport, and that the sequence selectivity of *in vitro* transport is therefore 'real', even if it is not physiological. They indicate that the artifacts, nonetheless, cannot be disregarded entirely, especially in respect of the claims of other investigators who have not taken specific steps to determine their extent under different conditions from those used here. For example, the apparently correct nucleus-restriction pattern of repetitive sequence transcripts, in liver nuclei incubated *in vitro* with homologous cytosol (537), could be explained as being due, largely, to the release of adherent RNA of cytoplasmic origin. The breakdown of this correct restriction pattern induced by heterologous (hepatoma) cytosol (537) could similarly be explained as being due to increased non-specific leakage, resulting from protease or ribonuclease action.

The pattern of relative abundances in *in vitro*-transported RNA indicates a degree of sequence selectivity, distorted to a small, but possibly significant degree, by the operation of the artifacts discussed above. Even if the effects of this biassing could be eliminated entirely, it is uncertain whether the observed selectivity of processing and transport *in vitro* is the same as that which operates *in vivo*. All that can be said is that its apparent effect, being to boost those relatively abundant mRNA sequences which constitute the bulk of polysomal mRNA part of the way towards their high polysomal frequencies, is as predicted from various other studies which envisage a cascade of selective mechanisms establishing the relative abundance pattern of polysomal mRNAs, operating at multiple levels in gene expression (110,609). Furthermore, the fact that different mRNAs, boosted ultimately to similar polysomal abundances, are subject to quite different degrees of selection in *in vitro* transport, is fully compatible with the idea that controls are exerted over the expression of

(sets of) genes at different levels. These observations do not prove that sequence selectivity *in vitro* is physiological, merely that it is consistent with postulated physiological specificity.

A more definitive study of whether the pattern of relative abundances in *in vitro*-transported RNA could be the result of physiological mechanisms requires an independent determination of the contribution of intranuclear and postnuclear processes to the polysomal abundances of post-transcriptionally regulated mRNAs. This can be achieved using a representative set of cDNA clones corresponding to individual mRNAs (e.g. from an HTC cell polysomal library), whose cytoplasmic half-lives can be measured directly in cell culture. Furthermore, if the corresponding mRNAs are of sufficiently high abundance, it should prove possible to measure their relative concentrations *in vivo*, in labelled, newly-transported RNA. It should at least be possible to establish the average contribution to abundance of post-nuclear events, by measuring the kinetics of cDNA-driven hybridisation of labelled, newly released RNA *in vivo*, as has been done for other cell-types (363,423, 424). If the data reported here are the result of *in vivo*-like specificity, then it can be predicted that the polysomal frequencies of the mRNAs corresponding to clones pRR 117, and possibly pRR 83 and pRR 5B, are largely determined by intranuclear selection, whereas that of the messenger corresponding to clone pRR 133 would be significantly boosted by specific cytoplasmic stabilisation. Verification of the *in vivo*-like specificity of *in vitro* processing and transport will require an extensive study using many more clones than these, however.

Even if the properties of HTC nuclei incubated with homologous cytosol do, indeed, prove to resemble those which pertain *in vivo*, this still leaves the failure to demonstrate cytosol-induced specificity of *in vitro* RNA transport open to multiple

interpretations. A number of relatively trivial technical causes must be considered, in the first instance, although they are not without biological implications. Firstly, the determinants of specificity may already reside in the nuclei, whether they be protein or RNA factors, or transcriptionally or post-transcriptionally determined features of the structures of pre-messenger molecules. A possibility along similar lines is that specificity-factors which remain associated with nuclei are able to over-ride those supplied from cytosol (whatever its origin): the operation of a particular gene regulatory 'switch' being determined by the overall balance between opposing factors. A second possibility is that the regulatory factors may have been lost from, or inactivated in the cytosol preparation(s) used, whether they be low molecular weight, labile to dialysis against distilled water or to other aspects of the preparation, or present only in high molecular weight (pelletable) aggregates. It is also possible that the conditions of incubation are inappropriate for their functioning. Fourthly, *trans*-acting regulatory factors may be unable to penetrate into the nuclei, due to conformational abnormalities caused by the removal of nuclear envelope phospholipid, ionic conditions, or absence of low molecular weight cofactors. Finally, if the regulatory factors represent a 'negative control' the use of a pre-incubation in homologous cytosol may re-introduce them into nuclei which would otherwise be sufficiently depleted of them to exhibit cytosol-mediated specificity.

If *in vitro* mRNA transport is indeed physiological, in the presence of both homologous and heterologous cytosols, this would indicate that the abundant, post-transcriptionally modulated mRNAs are not regulated at the level of nuclear RNA processing, but must be determined cytoplasmically by differential mRNA turnover. This can, in principle, be ascertained by direct determination of mRNA half-lives *in vivo*. If this albeit unlikely possibility proved to be the case, the cell-free system as developed would only be

applicable to the study of the mechanisms of post-transcriptional abundance modulation between liver and hepatoma, for mRNAs other than those encoding the differentiation-related functions suppressed in the hepatoma. (The failure to detect cytosol-induced specificity in respect of mRNAs corresponding to clones pRR 133 and pRR 117 in this system is not surprising: indeed they constitute a control for the detection of specificity relating to the other clones studied, since the mRNA corresponding to pRR 133 is hardly modulated between the cell-types at all, and that represented by pRR 117 appears to be transcriptionally controlled). In this eventuality, the study of the ability of cytosol to reproduce the cell-specific pattern of intranuclear post-transcriptional abundance modulation would have to proceed using cloned cDNAs for messengers truly regulated at this level, even subtly, between liver and hepatoma. There is no firm evidence provided in the whole of this study, however, that such sequences exist. They would only be likely to be isolated, assuming that they do exist, by a detailed screening of recombinants in liver and HTC cell polysomal cDNA libraries, coupled with detailed measurements, for many of these clones, of synthesis and decay rates, and relative levels <u>in vivo</u> in steady-state nuclear and polysomal RNAs and in newly-transported RNA.

A more plausible explanation for the results is that for one of the technical reasons already hypothesised, the determinants of processing selectivity are not active in heterologous cytosol. This is in no way inconsistent with the demonstrable cytosol-dependence of mRNA transport, nor with the possibility that (at least with homologous cytosol) <u>in vivo</u> specificity is maintained: cytosol may be required merely to contribute (or replenish) some vital, general component of the mRNA processing and transport machinery.

Truly physiological sequence selectivity, whether in respect of relative processing rates in the presence of homologous cytosol, or in respect of the effects of putative regulatory

factors in heterologous cytosol, might be regarded as elusive, on the grounds that the many different steps in RNA processing may show different optima *in vitro*. In terms of the variables of the incubation, conditions near optimal for one step may be far from it for others. Further refinement of the system towards *in vivo*-equivalence, by manipulating the conditions so as to eliminate the residual artifactual RNA release and optimise for faithful overall sequence selectivity, is likely therefore to be an iterative process. Factors affecting the overall and relative efficiencies of mRNA processing are likely to be: ionic conditions, temperature, cytosol protein concentration (and its composition as determined by the method used to prepare the cytosol), conformation of nuclear envelope pore complexes (as determined by the method of nuclear preparation) and nuclear suspension density. This optimisation process is not, it should be stressed, designed to achieve a maximal overall rate of mRNA transport, but to achieve the closest approach to the observed specificity of transport *in vivo*, as determined by studies with a representative sample of cDNA clones. *In vivo*, this specificity may be achieved by regulatory mechanisms which ensure that many sequences are processed and transported much less efficiently than the maximal rate at which they can be transported under unphysiological conditions *in vitro*.

The conditions originally adopted in this study were arrived at somewhat arbitrarily by investigators whose principal concern (519) was the satisfaction of very crude properties. With the availability of highly sensitive probes for the relative levels of different post-transcriptionally regulated sequences (in the form of cDNA clones), with which *in vitro*-transported RNA obtained under a variety of conditions may be readily analysed, it is possible to envisage a far more rigorous characterisation of the most appropriate conditions for physiological equivalence. One of the most pressing problems to be tackled in such an extension of this work would be the method of nuclear preparation, with a view to eliminating the problems caused by residual nuclear instability

(manifested as non-specific leakage), contamination with adherent RNA of cytoplasmic origin, and nuclear aggregation during incubation, which affects RNA processing and transport in ways which remain to be elucidated. A second priority is clearly the method of cytosol preparation, particularly from the viewpoint of its effect on the processing and transport of sequences modulated between the cell-types. A third area requiring considerable refinement is the composition of the incubation medium itself.

Regarding methods of nuclear preparation, attention may be focussed on the effects on the specificity of RNA transport of the composition of the cell homogenisation medium (particularly the content of divalent cations and the overall osmolarity), high-speed centrifugation, detergent treatment, the effects of hypo- and hyperosmolarity at various stages of preparation, and the effects of different osmotic stabilisers. The cell homogenisation medium is also of importance in the investigation of the effects of different cytosol preparation methods. Also to be examined here are the method of concentration and dialysis, including the buffer against which dialysis is performed, and storage conditions. As regards variables of the incubation medium all those mentioned above, as well as the possible influence of other biological small molecules (e.g. nucleotides other than ATP, cAMP, polyamines, glutathione) can be usefully included in such an investigation. A further possible variable, which might lead to more fruitful reproduction of cell-specific effects, would be the use of hnRNP-free nucleoplasm extracts as a supplement to (or alternative to) cytosol.

At the beginning of Section 4.8 three questions were posed, which may now be answered in summary. Firstly, the qualitative and quantitative composition of <u>in vitro</u>-transported poly(A)-adjacent RNA is indeed similar to that which would be predicted for a population of newly-transported messengers. Nucleus-confined sequences are at a low level, and may only be detected

by sensitive single-copy DNA methodology. The relative abundance pattern of the *in vitro*-transported RNA is intermediate between that of steady-state nuclear and polysomal RNA, and sequence-selectivity is evident at the level of individual mRNAs. Secondly, artifactual release of RNA from isolated nuclei accounts only for a minor fraction of the transported RNA, and cannot explain its relative abundance pattern. Thirdly, physiological selectivity, in response to heterologous cell cytosol, is not maintained *in vitro*, suggesting that the physiological equivalence of the system is, at best, partial. Alternatively, the results could be interpreted indicating that some hepatocyte mRNAs are modulated post-transcriptionally within the nucleus, whereas others, possibly including the differentiation-linked abundant mRNAs encoding specialist liver functions, are regulated at the level of cytoplasmic stability. Distinguishing these will require further studies *in vivo*, notable measurements of the stabilities, and mass fractions in newly transported RNA, of specific cloned cDNAs.

5.4. Implications of in vitro studies for mechanisms of RNA transport

RNA transport from isolated HTC cell nuclei has been shown in this investigation to be dependent upon ATP and cytosol, and to be unimpaired by pre-treatment of the nuclei, during isolation, with a non-ionic detergent. The ATP-dependence of pulse-labelled (messenger-like) RNA transport contradicts the findings of Schumm and Webb (526) who found that hepatoma nuclei, unlike those of liver (519) or indeed, other neoplastic cell-types (520,521) were only weakly stimulated in RNA transport by ATP. Although this discrepancy may arise from the difference in the precise cell-line used, it is most likely to result from differences in the method of preparing nuclei. Detergent treatment was found to increase radically the yield of nuclei on high-speed centrifugation through a dense sucrose pad. This suggests

that quite apart from considerations of contamination with adherent RNA, untreated nuclei are likely to be heavily contaminated with membranous tags containing cytoplasmic material. The poor ATP-dependence of RNA transport from untreated nuclei (526) may therefore result from the retention of significant amounts of ATP in adherent cytoplasm. In a preliminary study similarly poor ATP-dependence was observed for pulse-labelled RNA release from HTC nuclei prepared without detergent treatment, and with only low-speed centrifugation (data not shown). As in other systems where nuclear integrity is maintained (522,518,529,523) ATP-dependent pulse-labelled RNA transport was found here to be dependent on ATP hydrolysis, as judged by the inhibitory effect of a non-hydrolysable analogue. This is compatible with the involvement of nuclear pore-complex nucleoside triphosphatase in RNA transport (530,529,531), although ATP-requiring reactions may equally well be taking place in the nucleoplasm or within processing micro-organelles. ATP-requiring reactions of RNA processing include polyadenylation, and splicing (which has been demonstrated at least in the case of the yeast tRNAs (575)). The release of ribosomal subunits as evidenced here by the transport properties of steady-state labelled RNA from HTC cell nuclei also appears to be ATP-dependent, although in this case, non-hydrolysable analogues were found to be not only uninhibitory, but able to act as effective substitutes for ATP. This implies either that rRNA transport requires only pyrophosphate cleavage of ATP (which could be tested by using an equivalent α,β-blocked analogue of ATP), or that it does not require direct energisation by ATP hydrolysis, but rather depends on ATP-induced conformational changes at sites on the nuclear envelope. Ishikawa et al (524) made similar observations using liver nuclei pre-labelled in vivo for 3 hours, sufficiently long for much of the label to consist of pre-rRNA, in a medium which had previously been shown to favour rRNA release (620). Ribosomal subunits and mRNPs would

therefore appear to be transported by entirely independent mechanisms.

The role of cytosol in promoting RNA release from hepatoma nuclei appears to be essentially restrictive in nature, since gross (ATP-independent) nonspecific RNA release occurs if it is omitted. The size characterisation of the RNA released under such circumstances strongly suggests that this release is the result of generalised ribonuclease action, which is blocked by some component present in cytosol. The ability of cytosol prepared from hypotonically lysed cells, or of nuclear extracts, to restore viral RNA processing activity to nucle prepared hypotonically (540543,544) suggests that cytosol may in fact be replacing nuclear components which leach out of them during isolation. One possible explanation, therefore, for the gross degradation and spurious release of RNA in nuclei incubated without cytosol is that the nuclei, although having retained the excision activities associated with trimming and splicing, have been depleted either in the factors which confer specificity on these activities, or else of the RNA ligase activity with which they are normally associated, which can, however, be restored from cytosol. The dialysed cytosol preparations used here and in related studies are necessarily very crude, and their activity is more likely due to a relatively trivial non-specific effect, however.

The fact that HTC nuclei, treated as here with a non-ionic detergent during isolation, are still functional in rRNA and mRNA transport with very similar properties to those reported for untreated nuclei from other cell-types, but which maintain correct nuclear-restriction of the bulk of nucleus-confined information (albeit against a low level background of nonspecific leakage), suggests that the intactness of the lipid portion of the nuclear membrane is unimportant for these processes. This confirms the observations of Roy et al (514) who found that after Triton

X-100 treatment rat liver nuclei were still able to synthesise, process and transport messenger-like RNA in the form of informosomal particles. These properties of lipid-denuded nuclei imply that features of nuclear architecture other than the phospholipid of the nuclear envelope are responsible for limiting macromolecular access to, and release from, the genetic apparatus. It cannot be excluded, however, that detergent treatment increases the frequency with which correct nuclear RNA restriction breaks down, promoting a higher level of non-specific leakage than would otherwise be the case. Moreover, it may lead to other, more subtle alterations in the selectivity of RNA processing and transport, which may render the specificity observed here unphysiological. It may also be responsible for the considerable degree of nuclear aggregation which occurs during incubation, which is undesirable from the point of view of determining easily reproduced and well controlled conditions for the study of RNA transport with physiological specificity.

5.5. Implications of in vitro studies for gene control mechanisms

The significance of the *in vitro* studies, as has already been stated, depends critically upon the degree of physiological equivalence of the cell-free system. Since the assumption that the system's characteristics are a reasonable approximation to those of HTC nuclei *in vivo* remains unproven, all deductions about gene control mechanisms which depend upon it are necessarily speculative. Since, however, the ultimate objective of this project was to use the system to probe the mechanisms of post-transcriptional control beyond the crude characterisation possible from studies of *in vivo* RNA populations, it is useful to consider the possible implications of the results in this light. At the outset it must be recognised that the characterisation of the system has not reached a stage where direct studies of 'mechanism' have been possible. Nevertheless, the results do suggest certain general principles underlying gene control mechanisms in higher eukaryotic cells. In the

course of this discussion, it will be assumed that the observed specificity of in vitro mRNA transport is close to physiological.

Firstly, the level of individual abundant mRNAs seems to be regulated post-transcriptionally both within the nucleus and in the cytoplasm. Secondly, these controls are not (necessarily) co-ordinate. In other words, sequences which are strongly selected at one level, for example cytoplasmic stability, such as may be the case for the mRNA corresponding to clone pRR 133, may not be under strong positive selection at other levels (such as transcription, processing or transport). The converse is seen for clone pRR 117. Thirdly, rare-class sequences may not be under direct post-transcriptional modulation, (neither positive nor negative) their relative abundances being determined by 'basal' processing rates and stabilities. This would apply to the mRNAs represented by clones pRR 83 and pRR 5B in HTC cells where they are rare, although conclusions drawn from them may be misleading, since they represent differentiation-linked mRNAs whose expression may differ fundamentally from that of housekeeping genes. Nevertheless it is striking that the relative abundances of these sequences decrease from nuclear RNA, through in vitro-transported RNA to the polysomes, by an amount expected for sequences that were merely suffering a modest shift-down in abundance due to the selective amplification of the relative levels of mRNAs occupying the more abundant classes. The abundant sequences increase from representing but a tiny fraction of the total mass in hnRNA, to occupying about 70 percent of it on the polysomes. Fourthly, the post-transcriptionally regulated abundance changes which operate between liver and HTC cells, to modulate the levels of differentiation-linked mRNAs, do not alter the abundances of all the abundant mRNAs whose high polysomal frequencies are due to post-transcriptional stabilisation at the relevant level. For instance, if changes in cytoplasmic mRNA stability are responsible, at

least in part, for altered mRNA abundances between the cell-types, there remain sequences whose high abundance depends on specific cytoplasmic stabilisation, such as that of pRR 133, which are unaffected.

As has already been indicated, the failure to detect cytosol-induced specificity may be (and most probably is) technical rather than biological. If the system is, however, faithfully reproducing appropriate rates of relative mRNA processing and transport in the presence of the heterologous cytosol, then post-transcriptional modulation of the relevant mRNAs would be exclusively at the cytoplasmic level. Nevertheless, it is not possible to describe the alteration in gene expression between liver and hepatoma merely in terms of a single change in the specificity of mRNA turnover, since other genes appear to be modulated transcriptionally (such as may be the case of that represented by clone pRR 117) and possibly at other levels in addition.

5.6. Evaluation of the isolated nuclei system for in vitro studies of gene control mechanisms

Assuming that the remaining uncertainties concerning the physiological equivalence of the isolated hepatoma nuclei system can be overcome (although it is recognised that this is a considerable task), it is pertinent to ask how useful is such a system, in comparison with other *in vitro* systems for studying mechanisms of gene control. Its clear advantage over gene transfer systems and systems based on soluble extracts is that it permits examination of processing and transport mechanisms under conditions very close to those which pertain *in vivo*. With suitable assays, as may be provided by a battery of recombinants from a cDNA library, it provides a way of testing the effects of putative gene regulatory factors on the intranuclear post-transcriptional selectivity of gene expression, identifying what these factors are and how they work. However, in order to assess

the true viability of the system for studies of this nature, it is first necessary to demonstrate that regulatory macromolecules can enter the nuclei, and can effect the changes in the pattern of gene expression which are inferred from in vivo studies. The failure in this study to demonstrate any cytosol-specific effects is not encouraging from this point of view. The type of system described and developed here, although it encompasses a whole succession of steps in the 'gene expression pathway', is less problematic in this respect than gene transfer systems or systems based on the incorporation of a radioactive label into RNA synthesised, processed and subsequently transported from nuclei. In the system developed here, new transcriptional initiation can account for only a tiny fraction of the in vitro-transported RNA, and the vast bulk of transported sequences derives from mRNA precursors whose transcription was either in progress or complete when the nuclei were isolated. The system is therefore particularly appropriate for studying regulatory events within the limited panorama of intranuclear RNA processing and mRNA transport to the cytoplasm, to the exclusion of transcriptional initiation and of cytoplasmic RNA metabolism. Further dissection may become possible (although with the introduction of new uncertainties) with the use of specific inhibitors capable of blocking polyadenylation, capping, cap-methylation, internal methylation, splicing, transport and so on.

The major disadvantage of this system is that it is not easily susceptible to studying the nature of signals in primary nucleotide sequence that are the substrates for processing and processing-regulatory macromolecules. The surrogate genetics approach that is proving so successful in this respect in gene transfer and soluble-extract systems cannot be used here.

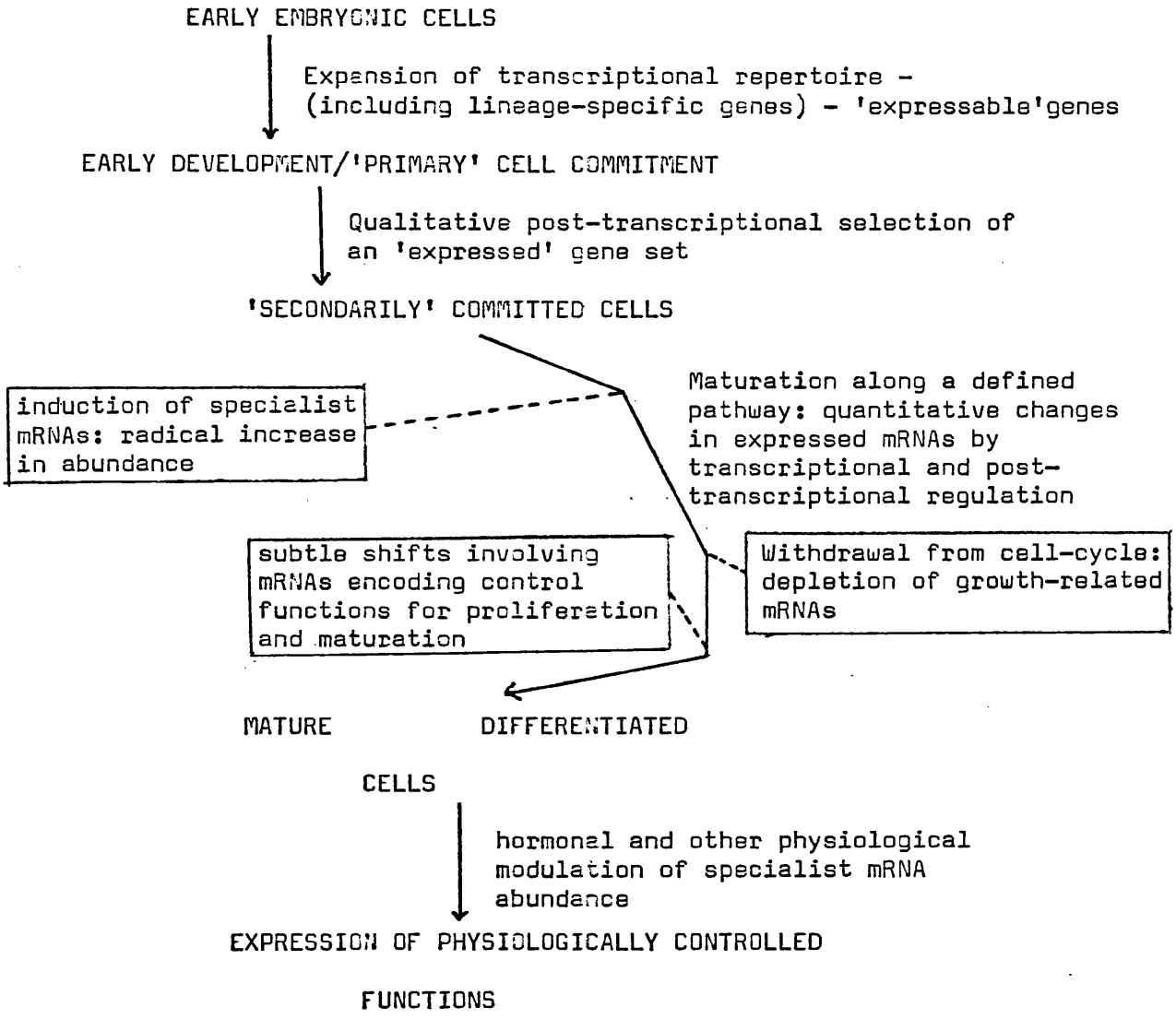
5.7. Significance of results in relation to differential gene expression in development, growth, differentiation and carcinogenesis

5.7.1. Modulation of mRNA abundances associated with hepatocyte differentiation and proliferation

The HTC cell differs from the fully differentiated hepatocyte by virtue of its decreased functional specialisation, its capacity to proliferate and its neoplastic nature. The differences observed here, in the composition of their polysomal mRNA populations, can be regarded as a superposition of the effects of gene controls associated with each of these phenotypic parameters. Within the limits of detection, the differences in mRNA populations are essentially quantitative: shifts in relative mRNA abundance rather than alterations in the set of genes expressed, are therefore responsible for the markedly different phenotypes of these two cell-types. This finding, taken together with the detailed analysis of the abundance changes and the inferences concerning the levels at which such changes are brought about, strongly reinforces the pattern of gene regulation deduced tentatively from other studies (see Sections 2.3.5 and 2.3.6). This pattern will now be restated explicitly in the form of a general model: it is important to state, at the outset, that whilst the data presented here do not prove such a model, or any of its elements, they do contribute significant evidence in support of it. The general validity of such a model can only be satisfactorily judged following the accumulation of a great deal of experimental data of the type presented here, using diverse biological systems. The uncertainties regarding the homogeneity and developmental origins of adult tissues and cell-lines in higher eukaryotes will further complicate this analysis. Doubtless, certain features of the model to be presented will require modification or qualification, to accommodate further experimental findings as they arise.

The essential elements of the model are summarised in Chart 2. Clearly the data presented here do not contribute any information directly on gene regulation in early development or relating to cell commitment. However, the results of single-copy DNA hybridisations support the idea that committed cells of a given lineage express a defined gene set which does not alter qualitatively as the differentiation pathway is traversed. This is supported by measurements of mRNA complexity and homology of other hepatoma lines, such as the Novikoff hepatoma (150), Chang's hepatoma (617) and Morris hepatoma 7777 (611) and also regenerating rat liver (145,374). These cell-types which may represent hepatocytes in various stages of maturation, express a set of polysomal mRNAs indistinguishable from that expressed in the fully differentiated hepatocyte. (Single-copy DNA methodology was not used in every case, however). The possible exception is the Novikoff hepatoma, where polysomal mRNA complexity seems to be smaller than that found in liver, although this finding was based solely on homologous cDNA hybridisation kinetics (146). Unlike HTC cells, or regenerating liver, the Novikoff hepatoma may also express a more limited set of sequences in poly(A)$^+$ nuclear RNA (146). The Novikoff hepatoma may represent a transformed multipotential hepatocyte precursor cell, which has still to pass through a final commitment stage at which the transcriptional repertoire is expanded. Alternatively, the Novikoff hepatoma may have lost the potential for expressing part of its normal mRNA set during serial passage in tissue-culture. The only reliable evidence for qualitative differences between messenger populations comes from comparisons between cell-types which diverged early in development (174,171,159,222,186) between different stages in embryogenesis itself, when cell commitment is taking place (179,177,313). The concept that progress along a differentiation pathway involves quantitative rather than qualitative changes in mRNA expression is supported by recent studies in other systems. No qualitative

CHART 2: Model for patterns of gene regulation in higher eukaryotes during development and differentiation.



differences in gene expression are detectable by single-copy DNA saturation, between hormone stimulated and hormone withdrawn chick oviduct (175),and a preliminary study using clones from a chick myoblast cDNA library has shown that changes which occur in the mRNA population during myogenesis are essentially quantitative, with shifts in mRNA abundance taking place in both directions (621). The results obtained in the liver/hepatoma system do not indicate whether the regulatory mechanism specifying the expression of a defined gene set in a committed cell-line throughout its differentiation, is transcriptional or post-transcriptional. Although there is good evidence, as already discussed, for post-transcriptional selection as the basis for this in a number of phylogenetically disparate eukaryotes, the question is unresolved as far as mammals and birds are concerned, since only very limited data are available (286,345). Much of the scepticism regarding this may be traced to the relatively low estimates of hnRNA complexity reported in a number of earlier studies (see Table 1). As has been indicated, the use here of a method which prepares undegraded poly(A)$^+$ hnRNA yields considerably increased estimates of hnRNA complexity, which could easily accommodate all of the extra tissue-specific complexity of, say, brain mRNA. The idea that the discovery of introns necessarily invalidates the proposal that tissue-specific mRNAs are ubiquitously represented in nuclear RNA, is probably untenable. However, further investigation is needed in order to fix the precise extent of tissue-specificity of mRNA and hnRNA sequence sets.

Quantitative variations within the set of expressed mRNAs in hepatocytes have here been shown to arise both transcriptionally and post-transcriptionally. The alterations in relative mRNA levels between different stages of hepatocytic differentiation which appear to be brought about by regulatory mechanisms acting at both these levels conform to the model shown in Chart 2.

As in numerous other comparisons of committed cells in different states or stages of differentiation along a given pathway (154,309,369,190,380), the more highly specialised liver cell expresses its specialist functions in the form of a set of very highly abundant mRNAs, which are found at a much lower frequency in the less differentiated homologue. These mRNAs, which are located on both free and membrane-bound polysomes,presumably direct the synthesis of the major secretory (serum) proteins manufactured in the liver, and of the principal enzymes of liver cellular function, such as those of glycogen metabolism. As in other cases, the levels of these differentiation-linked mRNAs seems to be controlled both transcriptionally (at least to a small degree), and post-transcriptionally (369,408,362,381,432,396,397,485, 484,395,510,493,506,511). If the cDNA clones corresponding to mRNAs of this category which have been studied (pRR 83 and pRR 5B) are typical, the predominant control exerted over their abundances must be post-transcriptional. This contrasts somewhat with data presented recently by other investigators, who have studied the composition of Novikoff hepatoma mRNA by heterogeneous cDNA hybridisation (146,150). They concluded that the contribution of post-transcriptional controls to depletion of abundant liver mRNAs in the hepatoma was relatively minor. Several possible reasons for this discrepancy can be advanced. In the first instance, the studies referred to (146,150) did not involve direct measurements of the level of polysomal messenger sequences (either heterogeneous or cloned) represented in hnRNA, which may have biassed the interpretation of the results. Secondly, a discrepancy may have arisen from the use of cytoplasmic (150) rather than polysomal poly(A)$^+$ RNA. Since the former is likely to be unavoidably contaminated with hnRNPs, the distinct difference in the abundance shifts in nuclear and polysomal RNAs which has been noted here, may have been masked. Thirdly, the Novikoff hepatoma may represent a distinct differentiation state from that of HTC cells. Its content of high abundance

mRNAs appears to be even more dramatically lowered (146,150), and it contains no detectable albumin mRNA (Y. Capetanaki, personal communication). It is possible, therefore, that the level at which the synthesis of differentiation-linked mRNAs is repressed is different in the two cell-lines, with complete shut-off of their transcription in Novikoff hepatoma, but with the HTC cell representing a more advanced stage in hepatocytic differentiation, where differentiation-linked messengers are efficiently transcribed, but where post-transcriptional controls boosting them to high abundance have not yet been switched on.

A recent comparative study of poly(A)$^+$ RNA populations in mouse mammary gland and chemically induced mammary carcinoma (in vivo) (206) has generated very similar findings to those reported here. The tumour was found to be highly depleted in super-abundant differentiation-specific mRNAs coding for milk proteins but to show a very modest increase in abundance of some sequences in the opposite direction. Single-copy DNA and cDNA methodologies failed to detect any qualitative differences. Chemically transformed embryonic hamster cells have also been reported recently, as expressing an indistinguishable gene set from the untransformed cell-line (173).

The apparently post-transcriptional depletion of abundant liver mRNAs in HTC cells could be due to regulation at a number of different levels in the cell. Perhaps surprisingly, the results of in vitro studies of mRNA transport suggest that the sole, or principal regulation might be cytoplasmic, although the interpretation of these results is subject to assumptions about their physiological equivalence. The processing/transport efficiency of the mRNA represented by clone pRR 133 in the (homologous) in vitro system certainly suggests that some highly abundant mRNAs in these cell-types owe their abundance very largely to selective cytoplasmic stability. Altered mRNA stability may underlie

abundance modulation during sea urchin development (624), and a number of differentiation-linked mRNAs encoding specialist functions in mammalian and avian systems have been shown to be regulated by differential turnover rates (396,381,432). More trivial explanations cannot, however, be entirely ruled out. The most obvious would be the specific removal of high abundance liver mRNAs from polysomes by a translational control mechanism in the hepatoma. Reihers and Busch have shown that this is not the case in the Novikoff hepatoma, where these mRNAs are massively depleted at the level of total cytoplasmic poly(A)$^+$ RNA (150). The normally abundant polysomal mRNAs are not present in the form of untranslated cytoplasmic RNPs. In HTC cells both albumin and α2 globulin mRNAs have been shown to be absent from cytoplasmic RNA, on the basis of a translational assay (512,627). Nevertheless, the possibility that some of the mRNAs depleted in HTC cells persist in an untranslated form cannot be entirely excluded, and at least one abundant liver mRNA (that coding for ferritin) has been shown to be regulated at this level in response to iron administration (614). Moreover, albumin mRNA has been found to be particularly efficient in translational initiation in a wheat-germ system (623), and it is therefore plausible that 'fine tuning' of the rate of differential protein synthesis in hepatocytes is brought about by translational regulation of the type postulated by Lodish (101), over and above other post-transcriptional controls. The possibility that abundant mRNA depletion in hepatoma is due to specific destabilisation of secretory mRNAs, resulting from a shutdown of translational recruitment into membrane-bound polysomes, would seem to be ruled out by the use of fractionated liver RNAs to drive hybridisations of polysomal cDNAs (see Section 4.4.5). A further possibility is that the depleted poly(A)$^+$ mRNAs are present in HTC cells in a non-adenylated form, and so not detected by the methodology adopted. Again this cannot be rigorously excluded, although there is no precedent for alterations in the state of polyadenylation of a whole set

of differentiation-linked mRNAs, in relation to the degree of cellular specialisation. Abundant mRNA depletion in the Novikoff hepatoma is evident in poly(A)$^-$ hnRNA to the same extent as in poly(A)$^+$ hnRNA (150).

Clearly, a minority of high abundance liver mRNAs are not depleted in HTC cells, such as the one represented by recombinant pRR 133. It is not necessary to regard such messengers as being unrelated to differentiation, or specialist hepatic function, although clone pRR 133 is represented at a high level in reticulocytes (R. Shott, unpublished observations). Despite the apparently uniform nature of the depletion of high abundance mRNAs between liver and HTC cells, there is no reason to assume that the expression of all differentiation-linked mRNAs is co-ordinate. Messengers such as that represented by clone pRR 133 may begin to be expressed at an earlier stage of hepatodifferentiation than the majority of specialist genes. Even the mRNAs depleted in HTC cells need not necessarily be regulated by a common mechanism. Albumin mRNA, for example, is known to be regulated independently of other abundant liver mRNAs in response to insulin levels (437), and $\alpha 2_u$ globulin messenger is specifically controlled by androgens (624) post-transcriptionally (625).

As well as postulating a dramatic increase in abundance during cell maturation of the specialist mRNAs encoding the physiological functions of highly differentiated cells, the model shown in Chart 2 also proposes more subtle shifts in the abundance of sequences involved in directing the maturation process itself, in cell proliferation and in controlling the cell's ability to respond to proliferative, anti-proliferative or developmental stimuli. The observation that mRNAs other than those coding for dramatically regulated differentiation-specific products are modulated between liver and HTC cells supports this aspect of the model, and recalls the situation noted in differentiating Friend cells (370), myogenesis (156,621) and pancreatic development (151), where sets of

sequences at moderate abundance early in the maturation process decline to very low levels subsequently.

The 'growing versus quiescent' aspect of the liver/HTC comparison may be considered in the light of other studies where this has been the sole or predominant variable (374,145, 371,141). A modest increase in abundance of a set of previously rarer mRNAs has been postulated to accompany the shift to the growing state (371,374) and such a hypothesis is supported by the data presented here. The bulk of HTC cell polysomal mRNA sequences are present on liver polysomes at slightly decreased average concentration, and as evidenced by the use of a fractionated cDNA enriched for abundant sequences, this conceals abundance changes operating in opposing directions. Whereas HTC cell polysomes are highly depleted in abundant, differentiation-linked liver mRNAs, their place has been taken by a group of sequences rather more subtly altered in abundance, of which clone pRR 117 may be a typical example. The data as presented do not permit an accurate estimate of the number of growth-inducible sequences, nor of the degree to which they are, individually, raised in abundance in proliferating cells. If the abundance shift observed for clone pRR 117 in HTC cells (or in regenerating liver) is typical of this class, then the kinetic shifts observed in heterogeneous cDNA cross-hybridisations must have been due to changes in frequency of many hundreds of messengers. Proliferation-associated mRNAs would, if this assumption proves correct, account for the vast majority of high abundance HTC cell messengers. Whether they are controlled transcriptionally or post-transcriptionally is again not possible to say with certainty from the data, although clone pRR 117 does appear to alter in abundance purely as a result of transcriptional selectivity. However, it must be stressed that early post-transcriptional events are also likely to contribute to the composition of steady-state hnRNA.

5.7.2. Implication of results for mechanisms of carcinogenesis

The data are not without interest in terms of theories of the etiology of carcinogenesis. In other systems, where relatively unspecialised cells already showing considerable proliferation capacity have been transformed with either viruses (385,386) or with chemical carcinogens (387,173), there is very little alteration in poly(A)$^+$ mRNA populations. This suggests that carcinogenesis involves changes in the expression of a relatively small number of genes, probably of a purely quantitative nature. However, these events may be atypical of transformation in vivo, in that viral action, in particular, may be highly specific for a few key growth-regulatory genes. A more usual course of events leading to tumour formation in vivo might be the accumulation of random perturbations in gene expression, which ultimately have a high probability of freeing the cell's proliferative capacity from the normal controls. This 'random-hit' theory is really a more generalised version of earlier somatic mutation hypotheses (627). The expression, in tumour cells, of inappropriate enzymes or antigens, often characteristic of a cell's developmental lineage, but which are repressed in the normal differentiated cell, seems to support this idea (628). The data presented here and elsewhere, on gene expression at the mRNA level in hepatomas (146,150,617,611) and other tumours (206), suggest that the genes whose expression is altered in tumour cells are not random, but constitute distinct sets of differentiation or growth-related mRNAs. Unless there is a high probability that lesions will occur in the expression of genes controlling the co-ordinate expression of other genes, then the 'random-hit' idea is not particularly favoured by the data.

The data are, however, compatible with the notion that carcinogenesis may result from a more specific lesion in the gene(s) responsible for maintaining the cell's differentiated state (to which cell proliferation may be linked). The

pattern of mRNA abundances in HTC cells does resemble that expected for a hepatocyte which has reverted to a de-differentiated phenotype characteristic of its earlier development. A phenotypically very similar type of de-differentiation to that seen in hepatomas has been observed in hepatocytes cultured on collagen gel-coated nylon meshes (629), which re-enter the cell cycle, express a number of fetal markers and repress differentiated cell functions such as albumin secretion (although the cells are not transformed as such). Furthermore, a number of rat hepatomas can revert, in a single step, to a phenotype in which the full set of differentiated functions is restored (433). This reversion is not enhanced by mutagenesis, suggesting that it results from a low probability re-programming event in the gene expression machinery. An alternative view, with which the data are equally compatible, is that tumours arise by uncontrolled proliferation of a committed cell whose maturation has become arrested (630,631). The expression of a specific set of fetal antigens in different tumours is seen as reflecting the developmental stage at which maturation arrest occurred. These two hypotheses - stepwise de-differentiation and maturation arrest - probably cannot be distinguished on the basis of studies of mRNA populations. Both would predict that the mRNA population of a newly transformed cell would resemble that of a normal progenitor cell; both theories can also accommodate the existence of tumours which have deviated irreversibly from the normal developmental pathway: mutational events subsequent to transformation might confer selective advantages in a rapidly growing cell population. Cells thus mutated would rapidly come to dominate that population, and the potential of the tumour to revert to (or progress on to) the more specialised phenotype would be lost.

Despite the close (inverse) relationship between the extent of cell specialisation and proliferation which emerges from data presented here and elsewhere, it is clear that controls exerted over these processes are to some extent separable, as is evident in the case of regenerating rat liver. Despite a rapid proliferation rate, and a large percentage of cycling cells, regenerating rat liver expresses differentiation-linked mRNAs at an almost unchanged abundance in both nuclear and polysomal poly(A)$^+$ RNA (145,374, Table 6). It appears, however, that messenger RNAs coding for growth related functions (such as may be represented by clone pRR 117) are induced to almost the same extent as in hepatoma. The processes leading to enhanced expression of such genes are therefore able to operate independently of those regulating the level of specialist mRNAs. Since the hyperplasia of liver regeneration ceases when the tissue damage is repaired, it could be argued that specialisation is not linked with proliferation per se, but with the ability to respond appropriately to proliferative or anti-proliferative stimuli. De-differentiation would not therefore be expected to accompany liver regeneration since the controlled regenerative response is part of the phenotypic repertoire of the normal differentiated hepatocyte. Alternatively, regenerating liver may represent a proliferative state akin to that of the very highly differentiated hepatomas, such as Morris hepatoma 7777 (611), where there is only a minimal depletion of high abundance mRNAs. On this view, only the very last step in hepatocytic differentiation might be coupled to the control of cell proliferation. HTC cells would represent a much earlier stage of development, whose further specialisation would normally occur as far as the stage of Morris hepatoma 7777, before losing the ability to proliferate. A final step-up in the levels of differentiation-linked mRNAs, concomitant with the withdrawal from the cell-cycle, may be envisaged as an indirect consequence of a much decreased rate of mRNA production, analogous to the mechanism of translational control envisaged by Lodish (101). This would only require that the differentiation-linked mRNAs were particularly high-affinity substrates for the rate-limiting step in mRNA synthesis.

6. REFERENCES

1. Vasil V, Hildebrant AC. (1965) Science 150, 889-892.

2. Gurdon J. (1974) The control of gene expression in animal development. Oxford: OUP.

3. Sala-Trepat J, Sargent TD, Sell S, Bonner J. (1979) Proc. Nat. Acad. Sci. USA 76, 695-699.

4. Brown DD, Dawid IB. (1968) Science 160, 272-280.

5. Spradling AC, Mahowald AP. (1980) Proc. Nat. Acad. Sci. USA 77, 1096-1100.

6. Alt FW, Kellems RE, Bertino JR, Schimke RT (1978) J. Biol. Chem. 253, 1357-1370.

7. Hoejimakers JHJ, Frasch ACC, Bernards A, Borst P, Cross GAM. (1980) Nature 284, 78-80.

8. Hicks J, Strathern J, Klar A. (1979) Nature 282, 478-483.

9. Sakano H, Huppi K, Heinrich G, Tonegawa S. (1979) Nature 280, 288-294.

10. Honjo T, Kataoka T. (1978) Proc. Nat. Acad. Sci. 75, 2140-2144.

11. Coleclough C, Cooper D, Perry RP. (1980) Proc. Nat. Acad. Sci. 77, 1422-1426.

12. Early P, Huang H, Davis M, Calame K, Hood L. (1980) Cell 19, 981-992.

13. Sakano H, Maki R, Kurosawa K, Roeder W, Tonegawa S. (1980) Nature 286, 676-283.

14. Dickinson DG, Baker RF. (1978) Proc. Nat. Acad. Sci. 75, 5627-5630.

15. Strom CM, Moscona M, Dorfman A. (1978) Proc. Nat. Acad. Sci. 75, 4451-4454.

16. Felsenfeld G. (1978) Nature 271, 115-122.

17. Adolph KW, Cheng SM, Laemmli UK. (1977) Cell 12, 805-816.

18. Kornberg RD (1977) Ann. Rev. Biochem. 46, 931-954.

19. Littau VC, Allfrey VG, Frenster JH, Mirsky AE (1964) Proc. Nat. Acad. Sci. USA 52, 93-100.

20. Scheer U. (1978) Cell 13, 535-549.

21. McKnight SL, Bustin M, Miller OL (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 741-754.

22. Ashburner M, Richards G. (1976). In: Insect development, Lawrence PA, ed. New York: Wiley. pp.203-225.

23. Ritossa FM. (1964) Exp. Cell Res. 36, 515-523.

24. Jamrich M, Greenleaf AL, Bautz EKF. (1979) Proc. Nat. Acad. Sci. USA 74, 2079-2083.

25. Tissières A, Mitchell HK, Tracy UM. (1974) J. Mol Biol. 84, 389-398.

26. Spradling A, Pardue ML, Penman S. (1977) J. Mol. Biol. 109, 559-587.

27. Schedl P, Artavanis-Tsakonis S, Steward R, Gehring W, Mirault M-E, Goldschmidt-Clermont M, Moran L, Tissières A. (1978) Cell 14, 921-929.

28. Weintraub H, Groudine M. (1976) Science 193, 848-858.

29. Garel A, Axel R. (1976) Proc. Nat. Acad. Sci. USA 73, 3966-3970.

30. Bellard M, Kuo MT, Dretzen G, Chambon P. (1980) Nucleic Acids Res. 8, 2737-2750.

31. Stalder J, Groudine M, Dodgson JB, Engel JD, Weintraub H. (1980) Cell 19, 973-980.

32. Varshavsky A, Sundin O, Bahn M. (1979) Cell 16, 453-466.

33. Stalder J, Larsen A, Engel JD, Dolan M, Groudine M, Weintraub H. (1980) Cell 20, 451-460.

34. Kuo T, Mandel J, Chambon P. (1979) Nucleic Acids Res. 7, 2105-2113.

35. Wu C, Bingham PM, Livak KJ, Holmgren R, Elgin SCR. (1979) Cell 16, 797-806.

36. Weintraub H. (1980) Nucleic Acids Res. 8, 4745-4753.

37. Pfeiffer W, Zachau HG (1980) Nucleic Acids Res. 8, 4621-4638.

38. Paul, J, Zollner E, Gilmour RS, Birnie GD. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 597-603.

39. Levy-W B,Dixon GH. (1977) Nucleic Acids Res. 4, 883-894.

40. Garel A, Zolan M, Axel R. (1977) Proc. Nat. Acad. Sci. USA 74, 4867-4871.

41. Miller DM, Turner P, Nienhuis AW, Axelrod DE, Gopalakrishnan TV.(1978) Cell 14, 511-521.

42. Gottesfeld JM, Partington GA. (1977) Cell 12, 953-962.

43. Weisbrod S, Weintraub H. (1979) Proc. Nat. Acad. Sci. 76, 630-634.

44. Weisbrod S, Groudine M, Weintraub H. (1980) Cell 19, 289-301.

45. Levy-Wilson B, Gjenset RA, McCarthy BJ. (1977) Biochim. Biophys. Acta 475, 168-175.

46. Davie JR, Candido PM. (1978) Proc. Nat. Acad. Sci. USA 75, 3574-3577.

47. Gottesfeld JM, Melton DA. (1978) Nature 273, 318-319.

48. Mayfield JE, Serunian LA, Silver LM, Elgin SCR. (1978) Cell 14, 539-544.

49. Lee D, Castro E, Jagodzinski LL, Sevall JS. (1979) Biochemistry 18, 3160-3165.

50. O'Malley BW, Means AR. (1974) Science 183, 610-620.

51. Schwartz RJ, Tsai M-J, Tsai SY, O'Malley BW (1975) J. Biol. Chem. 250, 5175-5182.

52. Vidali G, Boffa LC, Bradbury EM, Allfrey VG.(1978) Proc. Nat. Acad. Sci. USA 75, 2239-2243.

53. Marushige K. (1976) Proc. Nat. Acad. Sci. USA 73, 3937-3941.

54. Johnson EM, Allfrey VG (1978) In: Biochemical Actions of Hormones; Litwack G, ed. New York: Academic Vol.5, pp.1-53.

55. Lawson GM, Tsai M-J, O'Malley BW. (1980) Biochemistry 19, 4403-4411.

56. Fritsch EF, Lawn RM, Maniatis T. (1980) Cell 19, 959-972.

57. Royal A, Garapin A, Cami B, Perrin F, Mandel JF, Le Meur M, Brégégère F, Gannon F, Le Pennec JP, Chambon P, Kourilsky P. (1979) Nature 279, 125-132.

58. Wadsworth SC, Craig EA, McCarthy BJ (1980) Proc. Nat. Acad. Sci. USA 77, 2134-2137.

59. Zimmerman CR, Orr WC, Leclerc RF, Barnard EC, Timberlake WE. (1980) Cell 21, 709-715.

60. Jones CW, Kafatos FC. (1980) Nature 284, 635-638.

61. Hughes SH, Stubblefield E, Payvar F, Engel JD, Dodgson JB, Spector D, Cordell B, Schimke RT, Varmus HE. (1979) Proc. Nat. Acad. Sci. USA 76, 1348-1352.

62. van der Ploeg LHT, Flavell RA. (1980) Cell 19, 947-958.

63. Mandel JL, Chambon P. (1979) Nucleic Acids Res. 7, 2081-2103.

64. Bird AP, Taggart MH, Smith BA. (1979) Cell 18, 889-901.

65. Jacob F, Ullman A, Monod J. (1964)C R Acad Sci. Paris 258, 3125-3128.

66. Chamberlin MJ. (1974) Ann. Rev. Biochem. 43, 721-775.

67. Lee F, Yanofsky C. (1977) Proc. Nat. Acad. Sci. USA 74, 4365-4369.

68. Lamb MM, Daneholt B. (1979) Cell 17, 835-848.

69. Thompson EW, Lane BG (1980) J. Biol. Chem. 255, 5965-5970.

70. Kimchi A, Shulman L, Schmidt A, Chernayovsky Y, Fradin A, Revel M. (1979) Proc. Nat. Acad. Sci. USA 76, 3208-3212.

71. Tomkins GM, Gelehrter TD, Granner D, Martin DT, Samuels HH, Thompson EB (1969) Science 166, 1474-1480.

72. Chaterjee B, Hopkins J, Dutchak D, Roy AK. (1979) Proc. Nat. Acad. Sci. USA 76, 1833-1837.

73. Sehgal PB, Dobberstein B, Tamm I. (1977) Proc. Nat. Acad. Sci. USA 74, 3409-3413.

74. Crick FHC. (1979) Science 204, 264-271.

75. Shatkin AJ. (1976) Cell 9, 645-653.

76. Darnell JE, Jelinek WR, Molloy GR. (1973) Science 181, 1215-1221.

77. Karn J,Vidali G, Boffa LC, Allfrey VG. (1977) J. Biol. Chem. 252, 7307-7322.

78. Berger SL, Cooper HL. (1978) Biochim. Biophys. Acta 517, 84-98.

79. Derman E, Goldberg S, Darnell JE. (1976) Cell 9, 465-472.

80. McKnight SL, Miller OL. (1976) Cell 8, 305-319.

81. Goldberg S, Schwartz H, Darnell JE.(1977) Proc. Nat. Acad. Sci. USA 74, 4520-4523.

82. Salditt-Georgieff M, Harpold M, Chen-Kiang S, Darnell JE, (1980) Cell 19, 69-78.

83. Perry RP, Kelley DE. (1976) Cell 8, 433-442.

84. Beach LR, Ross J. (1978) J. Bio. Chem. 253, 2628-2632.

85. Puckett L, Chambers S, Darnell JE. (1975) Proc. Nat. Acad. Sci. USA 72, 389-393.

86. Perry RP, Kelley DE, La Torre J. (1974) J. Mol. Biol. 82, 315-331.

87. Johnson LF, Williams JG, Abelson HT, Green H, Penman S. (1973) Cell 4, 69-75.

88. Wu RS, Wilt FH. (1974) Dev. Biol. 41, 352-370.

89. Nevins JR, Darnell JE. (1977) Cell 15, 1477-1493.

90. Sawicki JG, Jelinek WR, Darnell JE. (1977) J. Mol. Biol. 117, 219-235.

91. Huez G, Marbaix G, Hubert E, Leclerq M, Nudel U, Soreq H, Salmon R, Lebleu B, Revel M, Littauer U. (1974) Proc. Nat. Acad. Sci. USA 71, 3143-3146.

92. Huez G, Marbaix G, Gallwitz D, Weinberg E, Devos R, Hubert E, Cleuter Y. (1978) Nature 271, 572-573.

93. Sehgal PB, Soreq H, Tamm I. (1978) Proc. Nat. Acad. Sci. USA 75, 5030-5033.

94. Wilson MC, Sawicki SG, White PA, Darnell JE. (1978) J. Mol. Biol. 126, 23-36.

95. Claisse M, Slonimski PP, Johnson J, Mahler HR. (1980) Mol. Gen. Genet. 177, 375-385.

96. Khoury G, Gruss P, Dhar R, Lai C-J. (1979) Cell 18, 85-92.

97. Stahl DA, Walker TA, Meyhack B, Pace NR. (1979) Cell 19, 1133-1143.

98. Pederson T. (1974) J. Mol. Biol. 83, 163-183.

99. Beyer AL, Miller OL, McKnight SL. (1980) Cell 20, 75-84.

100. Jain SK, Sarkar S. (1979) Biochemistry 18, 745-753.

101. Lodish HF. (1976) Ann. Rev. Biochem. 45, 39-72.

102. Bester AJ, Durrheim G, Kennedy DS, Heywood SM. (1979) Biochem. Biophys. Res. Commun. 92, 524-531.

103. Rosenthal ET, Hunt T, Ruderman J. (1980) Cell 20, 487-494.

104. Braude P, Pelham H, Flach G, Lobatto R. (1979) Nature 282, 102-105.

105. Macleod C, Firtel RA, Papkoff J. (1980) Dev. Biol. 76, 263-275.

106. Mozer TJ. (1980) Cell 20, 479-485.

107. Friderici K, Kaehler M, Rottman F. (1976) Biochemistry 15, 5234-5241.

108. Allende CC, Allende JE, Firtel RA. (1974) Cell 2, 189-196.

109. Pitot HC, Peraino C. (1964) Biochem. Clin. 3, 139-149.

110. Tobin AJ. (1979) Dev. Biol. 66, 47-58.

111. Marmur J, Rownd R, Schildkraut CL. (1963) Prog Nucleic Acid Res 1, 231-300.

112. Subirana JA, Doty P. (1966) Biopolymers 4, 171-187.

113. Britten RJ, Kohne DE. (1966) Carnegie Institute of Washington Yearbook 65, 75-106.

114. Wetmur JG, Davidson N. (1968) J. Mol. Biol. 31, 349-370.

115. Marmur J, Doty P. (1961) J. Mol. Biol. 3, 585-594.

116. Tso POP, Helmkamp GK, Sander C. (1962) Proc. Nat. Acad. Sci. USA 48, 686-698.

117. Zimmer C, Venner H. (1962) Naturwissenschaften 49, 86-87.

118. Marmur J, Doty P. (1962) J. Mol. Biol. 5, 109-118.

119. Geiduschek EP (1962) J. Mol. Biol. 4, 467-487.

120. McConaughy BL, Laird CD, McCarthy BJ. (1969) Biochemistry 8, 3289-3295.

121. Levine L, Murakami WT, Van Vunakis H, Grossman L. (1960) Proc. Nat. Acad. Sci. USA 46, 1038-1043.

122. Casey J, Davidson N. (1977) Nucleic Acids Res. 4, 1539-1552.

123. Schildkraut CL, Marmur J, Doty P. (1961) J. Mol. Biol. 3, 595-617.

124. Britten RJ, Kohne DE (1968) Science 161, 529-540.

125. Cairns J. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 43-46.

126. Efstradiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissmann S, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) Cell 21, 653-668.

127. Bautz EKF, Bautz FA (1964) Proc. Nat. Acad. Sci. USA 52, 1476-1481.

128. Shearer RW, McCarthy BJ (1967) Biochemistry 6, 283-289.

129. Brown DD, Dawid IG. (1969) Ann. Rev. Genet. 3, 127-154.

130. McCarthy BJ. (1967) Bacteriol. Rev. 31, 215-229.

131. Davidson EH, Hough BR. (1969) Proc. Nat. Acad. Sci. USA 63, 342-349.

132. Gelderman AH, Rake AV, Britten RJ. (1971) Proc. Nat. Acad. Sci. USA 68 172-176.

133. Holmes, DS, Bonner J. (1974) Biochemistry 13, 841-848.

134. Hahn WE, Laird CD. (1971) Science 173, 158-161.

135. Galau GA, Britten RJ, Davidson EH. (1974) Cell 2, 9-20.

136. Chikaraishi DM, Deeb SS, Suoeka N. (1978) Cell 13, 111-120.

137. Davidson EH, Hough BR. (1971) J. Mol. Biol. 56, 491-506.

138. Brown IR, Church RB. (1971) Biochem. Biophys Res. Commun. 42, 850-856.

139. Dale RMK, Martin F, Livingston DC, Ward DC. (1975) Biochemistry 14, 2447-2457.

140. Brown TDK and Balmain A. (1979) Nucleic Acids Res. 7, 2357-2368.

141. Williams JG, Penman S. (1975) Cell 6, 197-206.

142. Bishop JO, Morton JG, Rosbash M, Richardson M. (1974) Nature 250, 199-204.

143. Holland CA, Mayrand S, Pederson T. (1980) J. Mol. Biol. 138, 755-778.

144. Savage MJ, Sala-Trepat J, Bonner J. (1978) Biochemistry 17, 462-467.

145. Wilkes PR, Birnie GD, Paul J. (1979) Nucleic Acids Res. 6, 2193-2208.

146. Capetanaki YG, Alonso A. (1980) Nucleic Acids Res. 8, 3193-3214.

147. Chikaraishi D. (1979) Biochemistry 18, 3249-3256.

148. Grady LJ, Campbell WP, North AB. (1979) Nucleic Acids Res. 4, 259-269.

149. Manson P-E, Harris JE (1979). Biochemistry 18, 2073-2078.

150. Reiners JJ, Busch H. (1980) Biochemistry 19, 833-841.

151. Harding JD, MacDonald RJ, Przybyla AE, Chirgwin JM, Pictet RL, Rutter WJ. (1977) J. Biol. Chem. 252, 7391-7397.

152. Aziz S, Balmain A, Knowler J. (1979) Eur. J. Biochem. 100, 85-94.

153. Higgins SJ, Parker MG, Fuller FM, Jackson PJ. (1979) Eur. J. Biochem. 102, 431-440.

154. Parker MG, Mainwaring WIP. (1977) Cell 12, 401-407.

155. Lewin B. (1975) Cell 4, 77-93.

156. Leibovitch M-P, Leibovitch S-A, Harel J, Kruh J. (1979) Eur. J. Biochem. 97, 321-326.

157. Holmes DS, Bonner J. (1974) Biochemistry 13, 841-848.

158. Grady LJ, North AB, Campbell WP. (1978) Nucleic Acids Res. 5, 697-711.

159. Hastie ND, Bishop JO. (1976) Cell 9, 761-774.

160. Levy-W B, Johnson CB, McCarthy BJ.(1976) Nucl. Acids. Res. 3, 1777-1789.

161. Bantle JA, Hahn WE. (1976) Cell 8, 139-150.

162. Hahn WE, Van Ness J, Maxwell IH. (1980) Nature 283, 601.

163. Van Ness J, Maxwell IH, Hahn WE. (1979) Cell 18, 1341-1349.

164. Jacquet M, Affara NA, Robert B, Jakob H, Jacob F, Gros F. (1978) Biochemistry 17, 69-79.

165. Kleiman L, Birnie GD, Young BD, Paul J. (1977) Biochemistry 16, 1218-1223.

166. Birnie GD, MacPhail E, Young BD, Getz MJ, Paul J. (1974) Cell Differ. 3, 221-232.

167. Getz MJ, Birnie GD, Young BD, Paul J. (1975) Cell 4, 121-129.

168. Samal B, Bekhor I. (1979) Eur. J. Biochem. 94, 51-57.

169. Getz MJ, Elder PK, Benz EW, Stephen RE, Moses HL. (1976) Cell 7, 255-265.

170. Affara NA, Jacquet M, Jakob H, Jacob F, Gros F. (1977) Cell 12, 509-520.

171. Ryffel GU, McCarthy BJ. (1975) Biochemistry 14, 1379-1385.

172. Schrier BK, Zubairi MY, Letendre CH, Grouse LD. (1978) Differentiation 12, 23-30.

173. Moyzis RK, Grady DL, Li DW, Mirvis SE, Tso POP. (1980) Biochemistry 19, 821-832.

174. Axel R, Feigelson P, Schütz G. (1976) Cell 7, 247-254.

175. Mizuno S, Cox RF. (1979) Biochemistry 18, 2049-2056.

176. Ryffel GU. (1976) Eur. J. Biochem. 62, 417-423.

177. Shepherd GW, Flickinger R. (1979) Biochim. Biophys. Acta 563, 413-421.

178. Wold BJ, Klein WH, Hough-Evans BR, Britten RJ, Davidson EH. (1978) Cell 14, 941-950.

179. Galau GA, Klein WH, Davis MM, Wold BJ, Britten RJ, Davidson EH. (1976) Cell 7, 487-505.

180. Hough-Evans BR, Ernst SG, Britten RJ, Davidson EH. (1979) Dev. Biol. 69, 258-269.

181. Hough BR, Smith MJ, Britten RJ, Davidson EH. (1975) Cell 5, 291-299.

182. Hough-Evans BR, Jacobs-Lorena M, Cummings MR, Britten RJ, Davidson EH. (1980) Genetics (in press).

183. Levy-W B, McCarthy BJ. (1976) Biochemistry 14, 2440-2444.

184. Goldberg RB, Timberlake WE. (1980) Nature 283, 601-602.

185. Goldberg RB, Hoschek G, Kamalay JC, Timberlake WE. (1978) Cell 14, 121-131.

186. Kamalay JC, Goldberg RB. (1980) Cell 19, 935-946.

187. Kiper M, Bartels D, Kochel H. (1979) In: Genome, Chromatin and Karyotype: Evolution and Function; Nagl W, Hemleben V, Ehrendorfer, F. eds. New York: Springer.

188. Kiper M. (1979) Nature 278, 279-280.

189. Silflow CD, Hammett JR, Key JL (1979) Biochemistry 18, 2725-2731.

190. Auger S, Baulcombe D, Verma DPS. (1979) Biochim. Biophys. Acta 563, 496-507.

191. Curtis SE, Rawson JRY. (1979) Biochemistry 18, 5299-5304.

192. Zantinge B, Dons D, Wessels JGH. (1979) Eur. J. Biochem. 101, 251-260.

193. Rozek CE, Orr WC, Timberlake WE. (1978) Biochemistry 17, 716-722.

194. Timberlake WE, Schumard DS, Goldberg RB. (1977) Cell 10, 623-632.

195. Hereford LM, Rosbash M. (1977) Cell 10, 453-462.

196. Aviv H, Leder P. (1972) Proc. Nat. Acad. Sci. 69,1408-1412.

197. Ostrow RS, Woods WG, Vosika GJ, Faras AJ. (1979) Biochim. Biophys. Acta 562, 92-102.

198. Davidson EH, Hough BR, Amenson CS, Britten RJ. (1973) J. Mol. Biol. 77, 1-23.

199. Davidson EH, Britten RJ. (1973) Q Rev. Biol. 48, 565-613.

200. Davidson EH, Hough BR, Klein WH, Britten RJ. (1975) Cell 4, 217-238.

201. Kuroiwa A, Natori S. (1979) Nucleic Acids Res. 7, 751-764.

202. Goldberg RB, Timberlake WE. (1980) Nature 283, 601-602.

203. Ross J, Aviv H, Scolnik E, Leder P. (1972) Proc. Nat. Acad. Sci. USA 69, 264-268.

204. Kacian DL, Spiegelman S, Bank A, Terada M, Metafora S, Dow L, Marks PA. (1972) Nature (New Biol) 235, 167-169.

205. Verma IM, Temple GF, Fan H, Baltimore D. (1972) Nature (New Biol) 235, 163-167.

206. Supowit SC, Rosen JM. (1980) Biochemistry 19, 3452-3460.

207. Taylor JM, Illmensee R, Summers J. (1976) Biochim. Biophys. Acta 442, 324-330.

208. Birnstiel ML, Sells BH, Purdom IF. (1972) J. Mol. Biol. 63, 21-39.

209. Nygaard AP, Hall BD. (1964) J. Mol. Biol. 9, 125-142.

210. Bishop, JO. (1969) Biochem J. 113, 805-811.

211. Young BD, Harrison PR, Gilmour RS, Birnie GD, Hell A, Humphries S, Paul, J. (1974) J. Mol. Biol. 84, 555-568.

212. Laird JG, McCarthy BJ. (1968) Genetics 60, 303-312.

213. Young BD, Paul J. (1973) Biochem J. 135, 573-576.

214. Kinniburgh AJ, McMullen MD, Martin TE. (1979) J. Mol. Biol. 132, 695-708.

215. Herman RC, Williams JG, Penman S. (1976) Cell 7, 429-437.

216. Baltimore D, Smoler D. (1971) Proc. Nat. Acad. Sci. USA 68, 1507-1511.

217. Baltimore D, Smoler D. (1972) J. Biol. Chem. 247, 7282-7287.

218. Leis JP, Berkower I, Hurwitz J. (1973) Proc. Nat. Acad. Sci. USA 70, 466-470.

219. Youvan DC, Hearst JE. (1979) Proc. Nat. Acad. Sci. USA 76, 3751-3754.

220. Lasky L, Nozick ND, Tobin AJ. (1978) Dev. Biol. 67, 23-39.

221. Levy-W B, Johnson CB, McCarthy BJ. (1976) Nucleic Acids Res. 3, 1777-1789.

222. Young BD, Birnie GD, Paul J. (1976) Biochemistry 15, 2823-2829.

223. Perlman S, Rosbash M. (1978) Dev. Biol. 63, 197-217.

224. Sippel AE, Hynes NE, Groner B, Schütz G. (1977) Eur. J. Biochem. 77, 141-151.

225. Wu R, ed. Methods in Enzymol. Vol. 68: Recombinant DNA. New York: Academic.

226. Sinsheimer RL. (1977) Ann. Rev. Biochem. 46, 415-438.

227. Maniatis T, Hardison RC, Lacy E, Lauer J, O'Connel C, Quan D. (1978) Cell 15, 687-701.

228. Higuchi R, Paddock GV, Wall R, Salser W. (1976) Proc. Nat. Acad. Sci. USA 73, 3146-3150.

229. Humphries P, Cochet M, Krust A, Gerlinger P, Kourilsky P, Chambon P. (1977) Nucleic Acids Res. 4, 2389-2406.

230. Tilghman JM, Tiemeier DC, Polsky F, Edgell MH, Seidman JG, Leder A, Enquist LW, Norman B, Leder P. (1977) Proc. Nat. Acad. Sci. USA 74, 4406-4410.

231. Harpold MM, Evans RM, Salditt-Georgieff M, Darnell JE. (1979) Cell 17, 1025-1035.

232. Sim GK, Kafatos FC, Jones CW, Kochler MD, Efstradiadis A, Maniatis T. (1979) Cell 18, 1303-1316.

233. Benton WD, Davis RW. (1977) Science 196, 180-182.

234. Grunstein M, Hogness DS. (1975) Proc. Nat. Acad. Sci. USA 72, 3961-3965.

235. Costantini FD, Scheller RH, Britten RJ, Davidson EH. (1978) Cell 15, 173-187.

236. Scheller RH, Costantini FD, Kozlowski MR, Britten RJ, Davidson EH. (1978) Cell 15, 189-203.

237. Stark GR, Williams JG. (1979) Nucleic Acids Res. 6, 195-203.

238. Kafatos FC, Jones CW, Efstradiadis A. (1979) Nucleic Acids Res. 7, 1541-1552.

239. Southern EM. (1975) J. Mol. Biol. 58, 503-517.

240. Jeffreys AJ, Flavell RA. (1977) Cell 12, 1097-1108.

241. Alwine R, Kemp DJ, Stark GR. (1977) Proc. Nat. Acad. Sci. USA 74, 5350-5354.

242. Thomas M, White RL, Davis RW. (1976) Proc. Nat. Acad. Sci. USA 73, 2294-2298.

243. Berk AJ, Sharp PA. (1978) Cell 12, 721-732.

244. Davidson EH, Galau GA, Angerer RC, Britten RJ. (1975) Chromosoma 51, 253-259.

245. Genome Organisation and Expression in Plants (1979): various authors. (NATO-FEBS Advanced Studies Institute, Edinburgh) New York: Plenum.

246. Crain WR, Davidson EH, Britten RJ. (1976) Chromosoma 59, 1-12.

247. Manning JE, Schmid CW, Davidson N. (1975) Cell 4, 141-155.

248. Schachat F, O'Connor DJ, Epstein HF. (1978) Biochim. Biophys. Acta 520, 688-692.

249. Eden,FC, Hendrick JP. (1978) Biochemistry 17, 5838-5844.

250. Firtel RA, Kindle K. (1975) Cell 5, 401-411.

251. Hinnebusch AG, Klotz LC, Immergut E, Loeblich AR. (1980) Biochemistry 19, 1744-1755.

252. Hudspeth MES, Timberlake WE, Goldberg RB. (1977) Proc. Nat. Acad. Sci. USA 74, 4332-4335.

253. Timberlake WE. (1978) Science 202, 973-975.

254. Krumlauf R, Marzluff GA. (1979) Biochemistry 18, 3705-3713.

255. Cavalier-Smith T. (1978) J. Cell Sci. 34, 247-278.

256. Wimpee CF, Rawson JRY (1979) Biochim. Biophys. Acta 562, 192-206.

257. Lee DC, McKnight SL, Palmiter RD. (1980) J. Biol. Chem. 255, 1442-1450.

258. Suzuki Y, Suzuki E. (1974) J. Mol. Biol. 88, 393-407.

259. Rosbash M, Ford PJ, Bishop JO. (1974) Proc. Nat. Acad. Sci. USA 71, 3746-3750.

260. Grady LJ, Campbell WP. (1975) Nature 254, 356-358.

261. Hunt JA. (1974) Biochem. J. 138, 499-510.

262. Jones CW, Rosenthal N, Rodakis G, Kafatos FC. (1979) Cell 18, 1317-1332.

263. Fyrberg EA, Kindle KL, Davidson N. (1980) Cell 19, 369-378.

264. Tata JR, Baker BS, Deeley J. (1980) J. Biol. Chem. 255, 6721-6726.

265. Kedes L, Birnstiel ML. (1971) Nature (New Biol) 230, 165-169.

266. Hereford LM, Fahrner K, Woodford J, Rosbash M, Kaback DB. (1979) Cell 18, 1261-1271.

267. Wilson MI, Melli M. (1977) J. Mol. Biol. 110, 511-535.

268. Cavalier-Smith T. (1980) Nature 285, 617-618.

269. Garcia-Bellido A, Ripoll P. (1978) Nature 273, 399-400.

270. Ohta T, Kimura M. (1971) Nature 233, 118-119.

271. Dover G. (1980) Nature 285, 618-619.

272. Orgel LE, Crick FHC. (1980) Nature 284, 604-607.

273. Doolittle WF, Sapienza C. (1980) Nature 284, 601-604.

274. Flamm WG. (1972) Int.Rev. Cytol. 32, 2-51.

275. Kidd SJ, Glover DM. (1980) Cell 19, 103-119.

276. Houck CM, Rinehart FP, Schmid CW. (1979) J. Mol. Biol. 132, 289-306.

277. Kramerov DA, Grigoryan AA, Ryskov AP, Georgiev GP. (1979) Nucleic Acids Res. 6, 697-713.

278. Jelinek WR, Toomey TP, Leinward L, Duncan CH, Biro PA, Choudary CV, Weissmann SM, Rubin CM, Houck CM, Deininger PL, Schmid CW. (1980) Proc. Nat. Acad. Sci. USA 77, 1398-1402.

279. Tokarskaya ON, Tchurikov NA, Ivanov PL, Kramerov DA, Ryskov AP. (1980). Nucleic Acids Res. 8, 425-440.

280. Benoist C, O'Hare K, Breathnach R, Chambon P. (1980) Nucleic Acids Res. 8, 127-142.

281. Both GW. (1979) FEBS Lett 101, 220-224.

282. Gallwitz D, Sures I. (1980) Proc. Nat. Acad. Sci. USA 77, 2546-2550.

283. Busslinger M, Portmann R, Irminger JC, Birnstiel ML. (1980) Nucleic Acids Res. 8, 957-977.

284. Fiddes JC, Seeburg PH, DeNoto FH, Hallewell RA, Baxter JD, Goodman HM. (1979) Proc. Nat. Acad. Sci. USA 76, 4294-4298.

285. Britten RJ, Davidson EH. (1969) Science 165, 349-357.

286. Davidson EH, Britten RJ. (1979) Science 204, 1052-1059.

287. Dugaiczyk A, Woo SLC, Colbert DA, Lai EC, Mace ML, O'Malley BW. (1979) Proc. Nat. Acad. Sci. USA 76, 2253-2257.

288. Shen C-K, Maniatis T. (1980) Cell 19, 379-391.

289. Coggins LW, Grindlay GJ, Vass JK, Slater AA, Montague P, Stinson MA, Paul J. (1980). Nucleic Acids Res. 8, 3319-3333.

290. Leder A, Miller HI, Hamer DH, Seidman JG, Norman B, Sullivan M, Leder P. (1978) Proc. Nat. Acad. Sci. USA 75, 6187-6191.

291. Chaudhari N, Craig SP. (1979) Proc. Nat. Acad. Sci. USA 76, 6101-6105.

292. Graham DE, Schanke KE. (1980) Nucleic Acids Res. 8, 3875-3894.

293. Wensink PC, Tabata S, Pachl C. (1979) Cell 18, 1231-1246.

294. Eden FC. (1980) J. Biol. Chem. 255, 4854-4863.

295. Jelinek WR, Evans R, Wilson M, Salditt-Georgieff M, Darnell JE. (1978) Biochemistry 17, 2776-2783.

296. Mayfield JE, Bonner J. (1971) Proc. Nat. Acad. Sci. USA 68, 2652-2655.

297. Moore GP, Costantini FD, Posakony JW, Davidson EH, Britten RJ (1980) Science 208, 1046-1048.

298. Smith MJ, Hough BR, Chamberlin ME, Davidson EH. (1974) J. Mol. Biol. 85, 103-126.

299. Costantini FD, Britten RJ, Davidson EH. (1980) Nature 287, 111-117.

300. Jelinek W. (1978) Proc. Nat. Acad. Sci. USA 75, 2679-2683.

301. Kramerov DA, Ryskov AP, Georgiev GP. (1977) Biochim. Biophys Acta 475, 461-475.

302. Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA. (1980) Nature 283, 220-224.

303. Howard EF. (1978) Biochemistry 17, 3228-3236.

304. Seifert H, Schleuler M, Northemann W, Heinrich PC. (1979) Biochim. Biophys. Acta 564, 55-66.

305. Maxwell ES, Fischer MS. (1979) Biochim. Biophys. Acta 562, 302-318.

306. Matthews MB. (1980) Nature 285, 575-577.

307. Kaplan BB, Schachter BS, Osterburg HH, de Vellis JS, Finch CE. (1978) Biochemistry 17, 5516-5524.

308. Hastie ND, Held WA. (1978) Proc. Nat. Acad. Sci. USA, 75, 1217-122.

309. Paterson BM, Bishop JO. (1977) Cell 12, 751-765.

310. Dworkin MB, Dawid IG. (1980) Dev. Biol. 76, 435-448.

311. Williams JG, Lloyd MM. (1979) J. Mol. Biol. 129, 19-35.

312. Lev Z, Thomas TL, Lee AS, Angerer RC, Britten RJ., Davidson EH. (1980) Dev. Biol. 76, 322-340.

313. Biesseman H. (1980) Dev. Biol. (in press).

314. Galau GA, Klein WH, Britten RJ, Davidson EH. (1977) Arch Biochem. Biophys. 179, 584-599.

315. Gilbert W. (1978) Nature 271, 501.

316. Nunberg JH, Kaufman RJ, Chang ACY, Cohen SN, Schimke RT. (1980) Cell 19, 355-364.

317. Tsujimoto Y, Suzuki Y. (1979) Cell 18, 591-600.

318. Smith M, Leung DW, Gillan S, Astell CR, Montgomery DL, Hall BD. (1979) Cell 16, 753-761.

319. Holmgren R, Livak K, Morimoto R, Freund F, Meselson M. (1979) Cell 18, 1359-1370.

320. Craik CS, Buchman SR, Beychok S. (1980) Proc. Nat. Acad. Sci. USA 77, 1384-1388.

321. Weinstock R, Sweet R, Weiss M, Cedar H, Axel R. (1978) Proc. Nat. Acad. Sci. USA 75, 1299-1303.

322. Berget SN, Moore C, Sharp PA (1977) Proc. Nat. Acad. Sci. USA 74, 3171-3175.

323. Aloni Y, Dhar O, Laub M, Horowitz M, Khoury G. (1977) Proc. Nat. Acad. Sci. 74, 3686-3690.

324. Crawford LV, Colec CN, Smith AE, Paucha E, Tegtmeyer P, Rundell K, Berg P. (1978) Proc. Nat. Acad. Sci. USA 75, 117-121.

325. Mellon P, Duesberg H. (1977) Nature 270, 631-634.

326. Lindenmaier W, Nguyen-Huu MC, Lurz R, Stratmann M, Blin N, Wurtz T, Hauser HJ, Sippel AE, Schütz G. (1979) Proc. Nat. Acad. Sci. USA 76, 6196-6200.

327. Stein JP, Catterall JF, Kristo P, Means AR, O'Malley BW. (1980) Cell 21, 681-687.

328. Gough MM, Kemp DJ, Tyler BM, Adams JM, Cory S. (1980) Proc. Nat. Acad. Sci. USA 77, 554-558.

329. Tiemeier DC, Tilghman SM, Polsky FI, Seidman JG, Leder A, Edgell MH, Leder P. (1978) Cell 14, 237-245.

330. Horowitz M, Laub O, Bratosin S, Aloni Y. (1978) Nature 275, 558-559.

331. May E, Kress M, May P. (1978) Nucleic Acids Res. 5, 3083-3099.

332. Nevins JR. (1979) J. Mol. Biol. 130, 493-506.

333. Schibler U, Marcu KB, Perry RP. (1978) Cell 15, 1495-1509.

334. Tilghman SM, Curtis PJ, Tiemeier DC, Leder P, Weissmann C. (1978) Proc. Nat. Acad. Sci. USA 75, 1309-1313.

335. Roop DR, Nordstrom SY, Tsai Y, Tsai M-J, O'Malley BW. (1978) Cell 15, 671-685.

336. Piper P, Wardale J, Crew J. (1979) Nature 282, 686-691.

337. Tsai M-J, Tsai SY, O'Malley BW. (1979) Science 204, 314-316.

338. Roop DR, Tsai M-J, O'Malley BW. (1980) Cell 19, 63-68.

339. Acheson N. (1978) Proc. Nat. Acad. Sci. USA 75, 4754-4758.

340. Sippel AE, Groner B, Hynes N, Schutz G. (1977) Eur. J. Biochem. 52, 153-164.

341. Peunova NI, Samarina OP, Georgiev GP. (1979) Nucleic Acids Res. 6, 3625-3640.

342. van den Berg J, van Ooyen A, Mantei N, Schamböck A, Grosveld G, Flavell RA, Weissmann C. (1978) Nature 276, 37-44.

343. Kleene KC, Humphreys T. (1977) Cell 12, 143-155.

344. Balmain A, Minty AJ, Birnie GD. (1980) Nucleic Acids Res. 8, 1643-1660.

345. Lasky L, Tobin AJ. (1979) Biochemistry 18, 1594-1598.

346. Minty AJ. (1978) Ph.D. Thesis: University of Glasgow.

347. Stein JL, Thrall CL, Park WB, Mans RJ, Stein GS. (1975) Science 189, 557-558.

348. Borun TW, Gabrielli F, Ajiro K, Sweidler A, Baglioni C. (1975) Cell 4, 59-67.

349. Stein GS, Borun TW. (1972) J. Cell Biol. 52, 292-307.

350. Melli M, Spinelli G, Arnold E. (1977) Cell 12, 167-174.

351. Wilkes PR, Birnie GD, Old RW. (1978) Exp. Cell Res. 115, 441-444.

352. Dekte S, Lichtler A, Phillips I, Stein J, Stein G. (1979) Proc. Nat. Acad. Sci. USA 76, 4995-4999.

353. Groppi VE, Coffino P. (1980) Cell 21, 195-204.

354. Humphries S, Windass J, Williamson R. (1976) Cell 7, 267-277.

355. Perlman SM, Ford PJ, Rosbash M. (1977) Proc. Nat. Acad. Sci. USA 74, 3835-3839.

356. Harrison PR, Gilmour RS, Affara NA, Conkie D, Paul J. (1974) Cell Differ. 3, 23-30.

357. Therwath A, Scherrer K. (1978) Proc. Nat. Acad. Sci. USA 75, 3776-3780.

358. Beltz GA, Flint SJ. (1979) J. Mol. Biol. 131, 353-373.

359. Mauron A, Spohr G. (1978) Nucleic Acids Res. 5, 3013-3032.

360. Jacquet M, Gros F. (1979) Nucleic Acids Res. 6, 1639-1655.

361. Hofer E, Alonso A, Krieg L, Schatzle V, Sekeris CE. (1979) Eur. J. Biochem. 97, 455-462.

362. Soreq H, Harpold M, Wilson M, Darnell JE. (1980). Biochem. Biophys. Res. Commun. 92, 485-491.

363. Herman RC. (1979) Biochemistry 18, 916-920.

364. Nemer M, Graham M, Dubroff LM (1974) J. Mol. Biol. 89, 435-454.

365. Kaufman Y, Milcarek C, Berissi H, Penman S. (1977) Proc. Nat. Acad. Sci. USA 74, 4801-4805.

366. Milcarek C. (1979) Eur. J. Biochem. 102, 467-476.

367. Ordahl CP, Caplan AI. (1978) J. Biol. Chem. 253, 7683-7691.

368. Ernst S, Britten RJ, Davidson EH. (1979) Proc. Nat. Acad. Sci. USA 76, 2209-2212.

369. Minty AJ, Birnie GD, Paul J. (1978) Exp. Cell Res. 115, 1-14.

370. Affara N, Daubas P. (1979) Dev. Biol. 72 , 110-125.

371. Hereford LM, Rosbash M. (1977) Cell 10, 463-467.

372. Towle HC, Dillmann WH, Oppenheimer JH. (1979) J. Biol. Chem. 254, 2250-2257.

373. Miller WL, McCarthy BJ. (1978) J. Biol. Chem. 253, 742-748.

374. Scholla CA, Tedeschi MV, Fausto N. (1980) J. Biol. Chem. 255, 2855-2860.

375. Balmain A, Birnie GD. (1980) Biochemistry (m/s submitted).

376. Galau GA, Lipson ED, Britten RJ, Davidson EH. (1977) Cell 10, 415-432.

377. Rodgers WH, Gross PR. (1978) Cell 14, 279-288.

378. Goldstein ES, Arthur CG. (1979) Biochim. Biophys. Acta 565, 265-274.

379. Spector DH, Smith K, Padgett T, McCombe P, Roulland-Dusoix D, Moscovici C, Varmus HE, Bishop JM. (1978) Cell 13, 371-379.

380. Hynes NE, Groner B, Sippel AE, Nguyen-Huu MC, Schutz G. (1977) Cell 11, 923-932.

381. Cox RF. (1977) Biochemistry 16, 3433-3443.

382. Obinata M, Kameji R, Uchiyama Y, Ikawa Y. (1979). In: Oncogenic viruses and host cell genes; Ikawa Y, Odaka T. eds. New York: Academic. pp.429-438.

383. Aziz S, Balmain A, Knowler J. (1979) Eur. J. Biochem. 100, 95-100.

384. Felsani A, Berthelot F, Gros, F, Croizat B. (1978) Eur. J. Biochem. 92, 569-577.

385. Rolton HA, Birnie GD, Paul J. (1977) Cell Differ. 6, 25-39.

386. Williams JG, Hoffmann R, Penman S. (1977) Cell 11, 901-908.

387. Getz MJ, Reiman HM, Siegal GP, Quinlan TJ, Proper J, Elder PK, Moses HL. (1976) Cell 17, 241-254.

388. Church RB, McCarthy BJ. (1969) Proc. Nat. Acad. Sci. USA 58, 1548-1555.

389. Shearer RW, Smuckler EA (1972) Cancer Res. 32, 339-342.

390. Grady LJ, Campbell WP. (1975) Cancer Res. 35, 1559-1562.

391. Ashburner M, Bonner JJ. (1979) Cell 17, 241-254.

392. Mirault M-E, Goldschmidt-Clermont M, Moran L, Arrigo AP, Tissières A. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 819-827.

393. Kelley PM, Aliparti G, Schlesing MJ. (1980) J. Biol. Chem. 255, 3230-3233.

394. Nudel U, Salmon J, Fibach E, Terada M, Rifkind R, Marks PA, Bank A. (1977) Cell 12, 463-469.

395. Ganguly R, Mehta NM, Ganguly N, Bannerjee MR. (1979) Proc. Nat. Acad. Sci. USA 76, 6466-6470.

396. Matusik RJ, Rosen JM (1978) J. Biol. Chem. 253, 2343-2347.

397. Tsai SY, Roop DR, Tsai M-J, Stein JP, Means AR, O'Malley BW. (1978) Biochemistry 17, 5773-5780.

398. Lee DC, McKnight SL, Palmiter RD. (1978) J. Biol. Chem. 253, 3494-3503.

399. Ryan R, Shupnik M, Gorski J. (1979) Biochemistry 18, 2044-2048.

400. Toole JJ, Hastie ND, Held WA. (1979) Cell 17, 441-448.

401. Hastie ND, Held WA, Toole JJ. (1979) Cell 17, 449-457.

402. Savouret J-F, Loosfelt H, Atger M, Milgrom E. (1980) J. Biol. Chem. 255, 4131-4136.

403. Tsai SY, Tsai M-J, Lin C-T, O'Malley BW. (1979) Biochemistry 18, 5726-5731.

404. Dworkin MB, Dawid IB. (1980) Dev. Biol. 76, 449-465.

405. Dym HP, Kennedy DS, Heywood SM. (1979) Differentiation 12, 145-155.

406. Kunkel NS, Weinberg ES. (1978) Cell 14, 313-326.

407. Sala-Trepat JM, Dever J, Sargent TD, Thomas K, Sell S, Bonner J. (1979) Biochemistry 18, 2167-2178.

408. Aviv H, Bastos R, Volloch Z. (1979) In: Oncogenic viruses and host cell genes; Ikawa Y, Odaka T, eds. New York, Academic. pp.449-460.

409. Zitomer RS, Montgomery DL, Nichols DL, Hall BD. (1979) Proc. Nat. Acad. Sci. USA 76, 3627-3631.

410. Evans R, Weber J, Ziff E, Darnell JE. (1979) Nature 278, 367-370.

411. Tamm I. (1979) Proc. Nat. Acad. Sci. USA 74, 5011-5015.

412. Laub O, Jakobovits EB, Aloni Y (1980) Proc. Nat. Acad. Sci USA 77, 3297-3301.

413. Alestrom P, Akusjarvi G, Perricaudet M, Mathews MB, Klessig DF, Pettersson U. (1980) Cell 19, 671-681.

414. Benoff S, Bernardo N-G. (1979) Proc. Nat. Acad. Sci. USA 76, 1853-1857.

415. Segal S, Levine AJ, Khoury G. (1979) Nature 280, 335-337.

416. Segal S, Khoury G. (1979) Proc. Nat. Acad. Sci. USA 76, 5611-5615.

417. Fraser N, Ziff E. (1978) J. Mol. Biol. 124, 27-51.

418. Cory S, Jackson J, Adams JM. (1980) Nature 285, 450-456.

419. Singer PA, Singer HH, Williamson AR. (1980) Nature 285, 294-300.

420. Lee AS, Thomas TL, Lev Z, Britten RJ, Davidson EH. (1980) Proc. Nat. Acad. Sci. USA 77, 3259-3263.

421. Lewis , Bonewald , Lewis (1980) Biochem. Biophys. Res. Commun. 92, 511-

422. Wallis M. (1980) Nature 284, 512.

423. Lenk R, Herman R, Penman S. (1978) Nucleic Acids Res. 5, 3057-3070.

424. Meyuhas O, Perry RP. (1979) Cell 16, 139-148.

425. Lengyel JA, Penman S. (1979) Dev. Biol. 57, 243-253.

426. Spradling A, Hui H, Penman S. (1975) Cell 4, 131-137.

427. Berger SL, Cooper HL (1975) Proc. Nat. Acad. Sci. USA 72, 3873-3877.

428. Lowenhaupt K, Lingrel JB. (1978) Cell 14, 337-344.

429. Palmiter RD, Schimke RT. (1973) J. Biol. Chem. 248, 1502-1512.

430. Buckingham ME, Caput D, Cohen A, Whalen RG, Gros F. (1974) Proc. Nat. Acad. Sci. USA 71, 1466-1470.

431. Suzuki Y, Brown DD. (1972) J. Mol. Biol. 63, 409-429.

432. Hynes NE, Groner B, Sippel AE, Jeep S, Wurtz T, Nguyen-Huu MC, Giesecke K, Schütz G. (1979) Biochemistry 18, 616-624.

433. Walters RA, Yandell PM, Enger MD. (1979) Biochemistry 18, 4254-4261.

434. Goldstein ES. (1978) Dev. Biol. 63, 59-66.

435. Levy-W B, Rizzino A. (1977) Exp. Cell Res. 100, 377-380.

436. Keller GH, Taylor JM. (1979) J. Biol. Chem. 254, 276-278.

437. Peavy DE, Taylor JM, Jefferson LS. (1978) Proc. Nat. Acad. Sci. USA 75, 5879-5883.

438. Sell S, Thomas K, Michalson M, Sala-Trepat J, Bonner J. (1979) Biochim. Biophys Acta 564, 173-178.

439. Innis MA, Miller DL. (1977) J. Biol. Chem. 252, 8469-8475.

440. Harding JD, Rutter WJ. (1978) J. Biol. Chem. 253, 8736-8740.

441. Newrock KM, Cohen LH, Hendricks MB, Donnelly RJ, Weinberg ES. (1978) Cell 14, 327-336.

442. Zähringer J, Baliga BS, Drake RL, Munro HL. (1977) Biochim. Biophys. Acta 474, 234-244.

443. Harris H. (1974) Nucleus and cytoplasm (3rd ed.) Oxford: Clarendon.

444. Ringertz NA, Savage RE. (1976) Cell hybrids. New York: Academic.

445. Brown JE, Weiss MC (1975) Cell 6, 481-494.

446. Orkin SH, Hanosi FL, Leder P. (1975) Proc. Nat. Acad. Sci. USA 72, 98-102.

447. Allan M, Harrison PR. (1980) Cell 19, 437-447.

448. Pyati J, Kucherlapati RS, Skoultchi AI. (1980) Proc. Nat. Acad. Sci. USA 77, 3435-3439.

449. Lucas JJ, Kates JR. (1976) Cell 7, 397-405.

450. Lipsich LA, Kates JR, Lucas JJ. (1979) Nature 281, 74-76.

451. Gurdon JB, De Robertis EM, Partington G. (1976) Nature 260, 116-120.

452. De Robertis EM, Longthorne R, Gurdon JB. (1978) Nature 272, 254-256.

453. De Robertis EM, Gurdon JB. (1977) Proc. Nat. Acad. Sci. USA 74, 2470-2473.

454. Etkin LD. (1976) Dev. Biol. 52, 201- 209.

455. Furusawa M, Yamaizumi M, Nishimura T, Uchida T, Okada Y. (1976) Methods in Cell Biol. 14, 73-80.

456. Yamaizumi M, Uchida T, Okada Y, Furusawa M, Mitsui H. (1978) Nature 273, 782-784.

457. De Robertis EM, Olson MV. (1979) Nature 278, 137-143.

458. Melton DA, De Robertis EM, Cortese R. (1980) Nature 284, 143-148.

459. De Robertis EM, Mertz JE. (1977) Cell 12, 175-182.

460. Rungger D, Turler H. (1978) Proc. Nat. Acad. Sci. USA 75, 6073-6077.

461. Probst E, Kressmann A, Birnstiel ML. (1979) J. Mol. Biol. 135, 709-732.

462. Brown DD, Gurdon JB. (1978) Proc. Nat. Acad. Sci. USA 75, 2849-2853.

463. Arthur CG, Weide CM, Vincent WS, Goldstein ES. (1979) Exp. Cell Res. 123, 87-94.

464. Gurdon JB, De Robertis EM, Laskey RA, Wyllie AH. (1978) In: Miami Winter Symposium Vol. 15; Ahmad F, Schultz, J., Russell TR, Werner R, eds. New York: Academic. pp.457-466.

465. Grosschedl R, Birnstiel ML. (1980) Proc. Nat. Acad. Sci. USA 77, 1432-1436.

466. Wickens MP, Woo S, O'Malley BW, Gurdon JB. (1980) Nature 285, 628-634.

467. Hamer DH, Leder P. (1979) Nature 281, 35-40.

468. Hamer DH, Leder P. (1979) Cell 18, 1299-1302.

469. Mantei N, Boll W, Weissmann C. (1979) Nature 281, 40-46.

470. Wigler M, Sweet R, Sim GK, Wold B, Pellicer A, Lacy E, Maniatis T, Silverstein S, Axel R. (1979) Cell 16, 777-785.

471. Beggs JD, van der Berg J, van Ooyen A, Weissmann C. (1980) Nature 283, 835-840.

472. Ng R, Abelson J. (1980) Proc. Nat. Acad. Sci. USA 77, 3912-3916.

473. Breathnach R, Mantei N, Chambon P. (1980) Proc. Nat. Acad. Sci. USA 77, 740-744.

474. Lai EC, Woo SLC, Bordelon-Riser ME, Fraser TH, O'Malley BW (1980) Proc. Nat. Acad. Sci. USA 77, 244-248.

475. Wigler M, Perucho M, Kurtz D, Dana S, Pellicer A. Axel R, Silverstein S. (1980) Proc. Nat. Acad. Sci. USA 77, 3567-3570.

476. Klobutcher LA, Ruddle FH. (1979) Nature 280, 657-660.

477. Howell, N, Sager R. (1980) Proc. Nat. Acad. Sci. USA 77, 2844-2847.

478. Smith MM, Reeve AE, Huang RCC. (1978) Cell 15, 615-626.

479. Winicov I. (1979) Biochemistry 18, 1575-1582.

480. Detke S, Stein JL, Stein GS. (1978) Nucleic Acids Res. 5, 1515-1528.

481. Zimmer SG, Goldenberg CJ, Carlson DP, Craig EA, Raskas H. (1978) Biochemistry 17, 4207-4213.

482. Yang VW, Binger MH, Flint SJ. (1980) J. Biol. Chem. 255, 2097-2108.

483. Ickeson I, Yaniv M. (1979) Eur. J. Biochem. 95, 377-382.

484. McKnight GS, Palmiter RD. (1979) J. Biol. Chem. 254, 3050-3058.

485. Swaneck GE, Nordstrom JL, Kreuzaler F, Tsai M-J, O'Malley BW. (1979) Proc. Nat. Acad. Sci. USA 76, 1049-1053.

486. Panyim S, Ohno T, Jost J-P. (1978) Nucleic Acids Res. 5, 1353-1370.

487. Bastian C. (1978) Biochem. Biophys. Res. Commun 83 ,893-900.

488. Bastian C. (1980) Biochem. Biophys. Res. Commun 91 , 80-88.

489. Crampton JM, Woodland HR. (1979) Dev. Biol. 70, 453-466.

490. Yasuda Y, Shiokawa K, Yamana K. (1977) Biochim. Biophys. Acta 475, 453-460.

491. Gross RH, Ringler JC. (1979) Biochemistry 18, 4923-4927.

492. Sumner-Smith M, Phillips JP (1979) Insect Biochem. 9, 55-60.

493. Williams JG, Lloyd MM, Devine JM. (1979) Cell 17, 903-913.

494. Manley JL, Sharp PA, Gefter ML. (1979) Proc. Nat. Acad. Sci. USA 76, 160-164.

495. Mory Y, Gefter ML. (1978) Nucl. Acids Res. 5, 3899-3912.

496. Kieras RM, Almendinger RJ, Edmonds M. (1978) Biochemistry 17, 3221-3228.

497. Biswas DK, Martin TFJ, Tashijan AH. (1976) Biochemistry 15, 3270-3280.

498. Wu G-J, Zubay G. (1974) Proc. Nat. Acad. Sci. USA 71, 1803-1807.

499. McNamara DJ, Racevskis J, Schumm DE, Webb TE. (1975) Biochem. J. 147, 193-197.

500. Compton JL, McCarthy BJ. (1978) Cell 14, 191-201.

501. Taylor RN, Smith RG. (1979) Biochem. Biophys. Res. Commun. 91, 136-142.

502. Ide T, Whelly S, Baserga R. (1977) Proc. Nat. Acad. Sci. USA 74, 3189-3192.

503. Brown TDK, Pragnell IB, Paul J. (1980) Eur. J. Biochem. 104, 459-467.

504. Wydro R, Brot N, Weissbach H. (1980) Arch Biochem. Biophys. 201, 73-80.

505. Kozlov AP, Plyusnin AZ, Evtushenko VI, Seits IF. (1979) Biokhimiya (English Translation) 44, 27-32.

506. Orkin SH. (1978) Biochemistry 17, 487-492.

507. Tsai SY, Roop DR, Stumph WE, Tsai M-J, O'Malley BW. (1980) Biochemistry 19, 1755-1761.

508. Schafer KP. (1977) Nucleic Acids Res. 4, 4465-4475.

509. Ganguly R, Bannerjee MR (1978) Nucleic Acids Res. 5, 4463-4477.

510. McKnight SL, Lee DC, Palmiter RD. (1980) J. Biol. Chem. 255, 148-153.

511. Chan K-M, Kurtz DT, Feigelson P. (1978) Biochemistry 17, 3092-3096.

512. Chisick MC, Brennessel BA, Biswas DK (1979) Biochem. Biophys. Res. Commun. 91, 1109-1116.

513. Biswas DK. (1978) Biochemistry 17, 1131-1136.

514. Roy AK, Lau AS, Munro HL, Baliga BS, Sarkar S. (1979) Proc. Nat. Acad. Sci. USA 76, 1751-1755.

515. Mory YY, Gefter ML. (1977) Nucleic Acids Res. 4, 1739-1757.

516. Worofka R, Sauermann G. (1978) Biochim. Biophys. Acta 518, 61-80.

517. Cooper DL, Marzluff WF. (1978) J. Biol. Chem. 253, 8375-8380.

518. Schumm DE, McNamara DJ, Webb TE. (1973) Nature 'New Biol) 245, 201-203.

519. Schumm DE, Webb TE. (1972) Biochem. Biophys. Res. Commun. 48, 1259-1265.

520. Chang MJW, Koestner A, Palayoor T, Schumm DE, Webb TE. (1980) Biochem. Biophys. Res. Commun. 92,1340-1346.

521. Agutter PS, McCaldin (1979) Biochem. J. 180, 371-378.

522. Agutter PS. (1980) Biochem J. 188, 91-97.

523. Clawson GA, Koplitz M, Castler-Schechter B, Smuckler EA. (1978) Biochemistry 17, 3747-3752.

524. Ishikawa K, Sato-Odani S, Ogata K. (1978) Biochim. Biophys. Acta 521, 650-661.

525. Suzuki I. (1978) Biochem. J. 174, 174-177.

526. Schumm DE, Webb TE. (1975) Biochem. Biophys. Res. Commun. 67,706-713.

527. Smuckler EA, Koplitz M. (1974) Cancer Res. 34, 827-838.

528. Schumm DE, Hanausek-Walaszek M, Yannarell A, Webb TE. (1977) Eur. J. Cancer 13, 139-147.

529. Clawson GA, James J, Woo CH, Friend DS, Moody D, Smuckler EA. (1980) Biochemistry 19, 2748-2756.

530. Agutter PS, McArdle J, McCaldin B. (1976) Nature 263, 165-167.

531. Agutter PS, McCaldin B, McArdle HJ. (1979) Biochem. J. 182, 811-819.

532. Schumm DE, Webb TE. (1978) J. Biol. Chem. 253, 8513-8517.

533. Heilmann W, Hardeland R. (1977) Naturwissenschaften 64, 644-645.

534. Schumm DE, Webb TE. (1974) Biochem. J. 139, 191-196.

535. Yannarell A, Niemann M, Schumm DE, Webb TE. (1977) Nucleic Acids Res. 4, 503-511.

536. Herlan G, Giese G, Wunderlach F. (1979) Exp. Cell Res. 118, 305-309.

537. Schumm DE, Morris HP, Webb TE. (1973) Cancer Res. 33, 1821-1828.

538. Yu LC, Racevskis J, Webb TE. (1972) Cancer Res. 32, 2314-2321.

539. Racevskis J, Webb TE. (1974) Eur. J. Biochem. 49, 53-100.

540. Hamada H, Igarashi T, Muramatsu M. (1980) Nucleic Acids Res. 8, 587-599.

541. Schumm DE, Webb TE. (1975) J. Nat. Cancer Inst. 54, 123-128.

542. Schumm DE, Webb TE. (1975) Nature 256, 508-509.

543. Blanchard J-M, Weber J, Jelinek W, Darnell JE. (1978) Proc. Nat. Acad. Sci. USA 75, 5344-5348.

544. Goldenberg CJ, Raskas HJ. (1980) Biochemistry 19, 2719-2723.

545. Giesecke K, Sippel AE, Nguyen-Huu MC, Groner B, Hynes NE, Wurtz W, Schütz G. (1977) Nucleic Acids Res. 4, 3943-3958.

546. Pays E, Donaldson D, Gilmour RS. (1979) Biochim. Biophys. Acta 562, 112-130.

547. Crouse GF, Fodor E, Doty P. (1976) Proc. Nat. Acad. Sci. USA 73, 1564-1567.

548. Konkel DA, Ingram VM. (1977) Nucleic Acids Res. 4, 1237-1252.

549. Crouse GF, Fodor EJB, Doty P. (1978) Nucleic Acids Res. 5, 371-383.

550. Bachelor LT, Smith KD. (1976) Biochemistry 15, 3281-3290.

551. Dierks-Ventling C, Stalder J, Gautschi J. (1978) Nucleic Acids Res. 5, 2643-2656.

552. Beebee TJB, Carty DS. (1979) Biochem. Biophys. Res. Commun. 86, 199-205.

553. Birkenmeier EH, Brown DD, Jordan E. (1978) Cell 15, 1077-1086.

554. Sakonju S, Bogenhagen DF, Brown DD. (1980) Cell 19, 13-25.

555. Bogenhagen DF, Sakonju S, Brown DD. (1980) Cell 19, 27-35.

556. Engelke DR, Ng S-Y, Shastry BS, Roeder RG. (1980) Cell 19, 717-728.

557. Pelham HRB, Brown DD (1980) Proc. Nat. Acad. Sci. USA 77, 4170-4174.

558. Mattoccia E, Baldi MI, Carrara G, Fruscoloni P, Benedetti P, Tocchini-Valentini GP. (1979) Cell 18, 643-648.

559. Kressmann A, Hofsteller H, Di Capua E, Grosschedl R, Birnstiel ML. (1979) Nucleic Acids Res. 7, 1749-1763.

560. Duncan C, Biro PA, Choudary PV, Elder JT, Wang RCC, Forget B, de Riel JK, Weissmann SM. (1979) Proc. Nat. Acad. Sci. USA 76, 5095-5099.

561. Wasylyk B, Kedinger L, Corden J, Brison O, Chambon P. (1980) Nature 285, 367-373.

562. Manley JL, Fire A, Cano A, Sharp PA, Gefter ML. (1980) Proc. Nat. Acad. Sci. USA 77, 3855-3859.

563. Weil PA, Luse DS, Segall J, Roeder RG. (1979) Cell 18, 469-484.

564. Luse DS, Roder RG. (1980) Cell 20, 691-699.

565. Koch S, Niessing J. (1978) FEBS Lett. 96, 354-356.

566. Rose KM, Jacob ST. (1980) Biochemistry 19, 1472-1476.

567. Rose KM, Jacob ST. (1979) J. Biol. Chem. 254, 10256-10261.

568. Rose KM, Jacob ST, Kumar A. (1979) Nature 279, 260-262.

569. Wei C-M, Moss B. (1977) Proc. Nat. Acad. Sci. USA 74, 3758-3761.

570. Groner Y, Gilboa E, Aviv H. (1978) Biochemistry 17, 977-982.

571. Mizumoto K, Lipmann F. (1979) Proc. Nat. Acad. Sci. USA 76, 4961-4965.

572. Venkatesan S, Gershowitz A, Moss B. (1980) J. Biol. Chem. 255, 2829-2834.

573. Knapp G, Beckmann JS, Johnson PF, Fuhrman SA, Abelson J. (1978) Cell 14, 221-236.

574. Ogden RC, Beckmann JS, Abelson J, Kang HS, Söll D, Schmidt O. (1979) Cell 17, 399-406.

575. Peebles CL, Ogden RC, Knapp G, Abelson J. (1979) Cell 18, 27-35.

576. Knapp G, Ogden RC, Peebles CL, Abelson J. (1979) Cell 18, 37-45.

577. Roberts BE, Paterson BM. (1973) Proc. Nat. Acad. Sci. USA 70, 2330-2334.

578. Pelham HRB, Jackson RJ (1976) Eur. J. Biochem. 67, 247-257.

579. Kosower NS, Vanderhoff GA, Benerofe B, Hunt T, Kosower EM. (1971) Biochem. Biophys. Res. Commun. 45, 816-821.

580. Ehrenfeld E, Hunt T. (1971) Proc. Nat. Acad. Sci. USA 68, 1075-1078.

581. Ernst V, Levin DH, London IM. (1978) Proc. Nat. Acad. Sci. USA 75, 4110-4114.

582. Farrell PJ, Balkow K, Hunt T, Jackson RJ, Trachsel H. (1977) Cell 11, 187-200.

583. Hunt T. (1978) In: Miami Winter Symposia Vol. 16; Russell TR, Brew K, Faber A, Schultz J. eds. New York: Academic. pp.321-346.

584. Baglioni C, Minks MA, Maroney PA. (1978) Nature 273, 684-686.

585. Farrell PJ, Sen GC, Dubois MF, Ratner L, Slattery E, Lengyel P. (1978) Proc. Nat. Acad. Sci. (USA) 75, 5893-5897.

586. Strauss AW, Donohue AM, Bennett DC, Rodkey J, Alberts AW. (1977). Proc. Nat. Acad. Sci. USA 74, 1358-1362.

587. Thompson EB, Tomkins GM, Curran JF. (1966) Proc. Nat. Acad. Sci. USA 56, 296-303.

588. Gribnay AAM, Schoenmakers JGG, Bloemendal H. (1969) Arch Biochem. Biophys. 130, 48-52.

589. Gamulin S, Gray CH, Norman MR. (1972) Biochim. Biophys. Acta 259, 239-242.

590. Leffert HL, Koch KS, Moran T, Williams M. (1979) Methods in Enzymol. 58, 536-544.

591. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ. (1979) Biochemistry 18, 5294-5299.

592. Balmain A, Birnie GD. (1979) Biochim. Biophys. Acta 561, 155-166.

593. Studier FW. (1965) J. Mol. Biol. 11, 373-390.

594. Dale RMK, Ward DC. (1975) Biochemistry 14, 2458-2469.

595. Brown TDK, Balmain A. (1979) Nucleic Acids Res. 7, 2357-2368.

596. Laemmli UK. (1970) Nature 227, 680-685.

597. Anderson CW, Baum PR, Gesteland RF. (1973) J. Virol. 12, 241-252.

598. Bradford M. (1976) Anal. Biochem. 72, 248-254.

599. Sutcliffe RG. (1978) Nucleic Acids Res. 5, 2721-2728.

600. Birnie GD. In: Centrifugal separations in cell and molecular biology; Birnie GD, ed. London: Butterworths. pp.210-211.

601. Denhardt DT. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.

602. Monahan JJ, Harris SE, O'Malley BW. (1977) In: Receptors and hormone action. Vol. 1; O'Malley BW, Birnbaumer L. eds. New York: Academic. pp.297-329.

603. Chamberlin ME, Galau GA, Britten RJ, Davidson EH. (1978) Nucleic Acids Res. 5, 2073-2094.

604. Kurtz DT, Chan KM, Feigelson P. (1978) Cell 15, 743-750.

605. Ramsey JC, Steele WJ. (1976) Biochemistry 15, 1704-1709.

606. Smuckler EA, Koplitz M, Smuckler DE. (1976) In: Subnuclear components; Birnie GD, ed. London: Butterworths. pp.1-57.

607. Georgiev GP. (1967) In: Enzyme cytology; Roodbyn DB, ed. London: Academic Press.p.27.

608. Blobel G, Potter VR. (1966) Science 154, 1662-

609. Perry RP, Schibler U, Meyuhas O. (1979) In: Miami Winter Symposia Vol. 16; Russell TR, Brew K, Faber H, Schultz J, eds. New York: Academic. pp.187-201.

610. Potter UR (1961) Cancer Res. 21, 1331-1333.

611. Tse TPH, Morris HP, Taylor JM. (1978) Biochemistry 17, 3121-3128.

612. Hofer E, Land H, Sekeris C, Morris HP. (1978) Eur. J. Biochem. 91, 223-229.

613. Ruiz JPG, Ingram R, Hanson RW. (1978) Proc. Nat. Acad. Sci USA 75, 4189-4193.

614. Zähringer J, Baliga BS, Munro HL. (1979) FEBS Lett. 108, 317-320.

615. Atryzek V, Fausto N. (1979) Biochemistry 18, 1281-1287.

616. Zborovskaya IB, Alekhina RP, Likhtenstein AV, Shapot VS (1979) Biokhimiya (English Translation) 43, 1193-1200.

617. Knöchel W, Patel NT, Holoubek V. (1980) Biochim. Biophys. Acta 606, 67-75.

618. Atryzek V, Tamaoki T, Fausto N. (1980) Cancer Res. 40, 3713-3718.

619. Bucher NLR. (1963) Int. Rev. Cytol. 15, 245-300.

620. Sato T, Ishikawa K, Ogata K. (1977) Biochim. Biophys. Acta 474, 536-548.

621. Ordahl CP, Kioussis D, Tilghman SM, Ovitt CE, Fornwald J. (1980) Proc. Nat. Acad. Sci. USA 77, 4519-4523.

622. Feigelson P, DeLap L, Chen CC, Chan K, Kurtz D. (1979) In: The cell nucleus. Vol. 7: Chromatin, Part D; Busch H. ed. New York: Academic. pp.229-257.

623. Sonsenheim BE, Brawerman G. (1977) Biochemistry 16, 5445-5448.

624. Shepherd GW, Nemer M. (1980) Proc. Nat. Acad. Sci. USA 77, 4653-4656.

625. Sippel AE, Feigelson P, Roy AK. (1975) Biochemistry 14, 825-829.

626. Chaterjee B, Hopkins J, Dutchak D, Roy AK. (1979) Proc. Nat. Acad. Sci. USA 76, 1833-1837.

627. Boveri TH. (1914) Zur Frage der Entstchung maligner tumoren. Jena (GDR): Fischer.

628. Shields R. (1977) Nature 269, 752-753.

629. Sirica AE, Richards W, Tsukada Y, Sattler CA, Pitot HC. (1979) Proc. Nat. Acad. Sci. USA 76, 283-287.

630. Greaves MF, Janossy G. (1978) Biochim. Biophys. Acta 516, 193-220.

631. Graf T, Beug H. (1978) Biochim. Biophys. Acta 516, 269-299.

632. Ivarie RD, O'Farrell PH. (1978) Cell 13, 41-56.

616. Zborovskaya IB, Alekhina RP, Likhtenstein AV, Shapot VS (1979) Biokhimiya (English Translation) 43, 1193-1200.

617. Knöchel W, Patel NT, Holoubek V. (1980) Biochim. Biophys. Acta 606, 67-75.

618. Atryzek V, Tamaoki T, Fausto N. (1980) Cancer Res. 40, 3713-3718.

619. Bucher NLR. (1963) Int. Rev. Cytol. 15, 245-300.

620. Sato T, Ishikawa K, Ogata K. (1977) Biochim. Biophys. Acta 474, 536-548.

621. Ordahl CP, Kioussis D, Tilghman SM, Ovitt CE, Fornwald J. (1980) Proc. Nat. Acad. Sci. USA 77, 4519-4523.

622. Feigelson P, DeLap L, Chen CC, Chan K, Kurtz D. (1979) In: The cell nucleus. Vol. 7: Chromatin, Part D; Busch H. ed. New York: Academic. pp.229-257.

623. Sonsenheim BE, Brawerman G. (1977) Biochemistry 16, 5445-5448.

624. Shepherd GW, Nemer M. (1980) Proc. Nat. Acad. Sci. USA 77, 4653-4656.

625. Sippel AE, Feigelson P, Roy AK. (1975) Biochemistry 14, 825-829.

626. Chaterjee B, Hopkins J, Dutchak D, Roy AK. (1979) Proc. Nat. Acad. Sci. USA 76, 1833-1837.

627. Boveri TH. (1914) Zur Frage der Entstchung maligner tumoren. Jena (GDR): Fischer.

628. Shields R. (1977) Nature 269, 752-753.

629. Sirica AE, Richards W, Tsukada Y, Sattler CA, Pitot HC. (1979) Proc. Nat. Acad. Sci. USA 76, 283-287.

630. Greaves MF, Janossy G. (1978) Biochim. Biophys. Acta 516, 193-220.

631. Graf T, Beug H. (1978) Biochim. Biophys. Acta 516, 269-299.

632. Ivarie RD, O'Farrell PH. (1978) Cell 13, 41-56.