Outline Studies in Biology

Protein Biosynthesis

Alan E. Smith

Protein Biosynthesis

Alan Smith

Department of Molecular Virology
Imperial Cancer Research Fund
Lincolns Inn Fields, London

LONDON
CHAPMAN AND HALL

A Halsted Press Book
JOHN WILEY & SONS, INC., NEW YORK

First published in 1976
by Chapman and Hall Ltd
11 New Fetter Lane, London EC4P 4EE

Printed in Great Britain by
William Clowes & Sons Ltd.,
London, Colchester and Beccles

ISBN 0 412 13460 8

Distributed in the U.S.A.
by Halsted Press, a Division
of John Wiley & Sons, Inc., New York

Library of Congress Cataloging in Publication Data
Smith, Alan E
Protein biosynthesis.

(Outline studies in biology series)
"A Halsted Press book".
Bibliography: P.
1. Protein biosynthesis. I. Title.
QP551.S64 1976 574.1'9296 75-41516
ISBN 0-470-80263-4

List of titles

Already published

Cell Differentiation	J.M. Ashworth
Biochemical Genetics	R.A. Woods
Functions of Biological Membranes	M. Davies
Cellular Development	D. Garrod
Brain Biochemistry	H.S. Bachelard
Immunochemistry	M.W. Steward
The Selectivity of Drugs	A. Albert
Biomechanics	R. McN. Alexander
Molecular Virology	T.H. Pennington, D.A. Ritchie
Hormone Action	A. Malkinson
Cellular Recognition	M.F. Greaves
Cytogenetics of Man and other Animals	A. McDermott
RNA Biosynthesis	R.H. Burdon
Protein Biosynthesis	A.E. Smith

In preparation

The Cell Cycle	S. Shall
Biological Energy Transduction	C. Jones
Control of Enzyme Activity	P. Cohen
Metabolic Regulation	R. Denton, C.I. Pogson
Polysaccharides	D.A. Rees
Microbial Metabolism	H. Dalton
Bacterial Taxonomy	D. Jones
Molecular Evolution	W. Fitch
A Biochemical Approach to Nutrition	R.A. Freedland
Metal Ions in Biology	P.M. Harrison, R. Hoare
Nitrogen Metabolism in Plants and Microorganisms	A.P. Sims
Cellular Immunology	D. Katz
Muscle	R.M. Simmons
Xenobiotics	D.V. Parke
Plant Cytogenetics	D.M. Moore
Human Genetics	J.H. Edwards
Population Genetics	L.M. Cook
Membrane Biogenesis	J. Haslam
Biochemical Systematics	J.H. Harborne
Biochemical Pharmacology	B.A. Callingham
Insect Biochemistry	H.H. Rees

OUTLINE STUDIES IN BIOLOGY

Editor's Foreword

The student of biological science in his final years as an undergraduate and his first years as a graduate is expected to gain some familiarity with current research at the frontiers of his discipline. New research work is published in a perplexing diversity of publications and is inevitably concerned with the minutiae of the subject. The sheer number of research journals and papers also causes confusion and difficulties of assimilation. Review articles usually presuppose a background knowledge of the field and are inevitably rather restricted in scope. There is thus a need for short but authoritative introductions to those areas of modern biological research which are either not dealt with in standard introductory textbooks or are not dealt with in sufficient detail to enable the student to go on from them to read scholarly reviews with profit. This series of books is designed to satisfy this need. The authors have been asked to produce a brief outline of their subject assuming that their readers will have read and remembered much of a standard introductory textbook of biology. This outline then sets out to provide by building on this basis, the conceptual framework within which modern research work is progressing and aims to give the reader an indication of the problems, both conceptual and practical, which must be overcome if progress is to be maintained. We hope that students will go on to read the more detailed reviews and articles to which reference is made with a greater insight and understanding of how they fit into the overall scheme of modern research effort and may thus be helped to choose where to make their own contribution to this effort. These books are guidebooks, not textbooks. Modern research pays scant regard for the academic divisions into which biological teaching and introductory textbooks must, to a certain extent, be divided. We have thus concentrated in this series on providing guides to those areas which fall between, or which involve, several different academic disciplines. It is here that the gap between the textbook and the research paper is widest and where the need for guidance is greatest. In so doing we hope to have extended or supplemented but not supplanted main texts, and to have given students assitance in seeing how modern biological research is progressing, while at the same time providing a foundation for self help in the achievement of successful examination results.

J.M. Ashworth, Professor of Biology, University of Essex.

Contents

1 Introduction

1.1 The problem

The discovery that the genetic material of living organisms is DNA, and the later demonstration that the DNA molecule is a double helix were both great milestones in twentieth century science, and formed the foundation of the new discipline of molecular biology. But even after these momentous discoveries, the detailed mechanism by which such genetic material could be expressed as the structural and catalytic proteins which play so important a role in the functioning of all living cells was still not obvious. It was only after the rigorous demonstration that a given protein is composed of a unique linear sequence of amino acids that the possibility of a relationship between the base sequence of DNA and the sequence of amino acids in a protein was realized. Once this conceptual breakthrough had been made, the complex task of unravelling the many steps in protein biosynthesis could begin in the laboratory. This task has continued apace during the past twenty years, so that today we have a fairly clear overall picture of the process. This elucidation of the many steps and intricacies of protein biosynthesis, which is now known to involve well in excess of a hundred different molecular species, must rank as one of the major and most exciting achievements of present-day molecular biology.

The purpose of this monograph is to outline in some detail our present knowledge of protein biosynthesis, describing the molecules and the basic mechanisms involved, and also the possible control processes operating to adjust protein synthesis to the needs of the cells and organism. It will be assumed that the reader has some knowledge of molecular biology in general and protein biosynthesis in particular, but by way of introduction each of the major molecules and stages of the process will be described in simple terms, and in subsequent chapters each will be discussed again in greater depth.

1.2 Overall steps in protein biosynthesis

The information encoded in the two complementary strands of the DNA of any structural gene is transcribed by an enzyme called DNA-dependent RNA polymerase. It makes a single-stranded RNA copy, complementary to one of the strands, which is called *messenger RNA* (mRNA). This attaches to a subcellular organelle called a *ribosome* which is composed of two subunits and functions as a black box upon which the mRNA is *translated.* The term translation encompasses all those steps by which the genetic content of the mRNA contained in the linear sequence of ribonucleotides is converted into a linear sequence of amino acids which may have enzymatic or other biological properties. The language of translation which dictates that a particular sequence of nucleotides (a *codon*) leads to the insertion of one particular amino acid is called the *genetic code.* A special class of adaptor molecules which are both able to read the codons of the genetic code and to carry the appropriate amino

acid to the ribosome for polymerization are composed of RNA and termed *transfer RNA* (tRNA).

The initiation of protein biosynthesis is the process by which the mRNA first attaches to a ribosome and is prepared for translation. Initiation involves a series of critical reactions, requiring a unique *initiator tRNA* and catalysed by several proteins, termed *initiation factors*, which ensure the correct alignment or *phasing* of the mRNA upon the ribosome. The formation of the *initiation complex* containing all the components mentioned above is probably one of the rate-limiting steps in the biosynthesis of a protein and it is the point at which many control elements operate.

After the successful formation of the 'initiation complex' the next amino acid in the polypeptide chain is carried to the ribosome attached to a tRNA. The binding of the appropriate *amino-acyl tRNA* is directed by the codon and is catalysed by an *elongation factor*. Peptide bond formation is then catalysed by a ribosomal protein *peptidyl-transferase* and results in the formation of a covalent bond between the amino group of the incoming amino acyl-tRNA and the carboxyl group of the amino acid immediately preceding it in the growing polypeptide. In this way the mRNA is read 5′ to 3′ and the polypeptide chain *elongated* from the NH_2 to COOH ends. The ribosome is then *translocated* along the mRNA by a second elongation factor such that a new codon is exposed and the cyclic elongation process then continues.

Once elongation has proceeded sufficiently and the ribosome passed some distance along the mRNA, the ribosome-binding site becomes free again and a second ribosome can attach. In this way several ribosomes may attach to a single mRNA molecule and the resulting structure is called a *polysome*. When a ribosome passes to the end of the coding portion of an mRNA it reaches a codon signalling *terminate* protein synthesis and in a reaction involving several *termination factors* the polypeptide is released. The nascent polypeptide chain probably begins to fold into its native secondary and tertiary structure whilst it is being synthesized and is still bound to the ribosome.

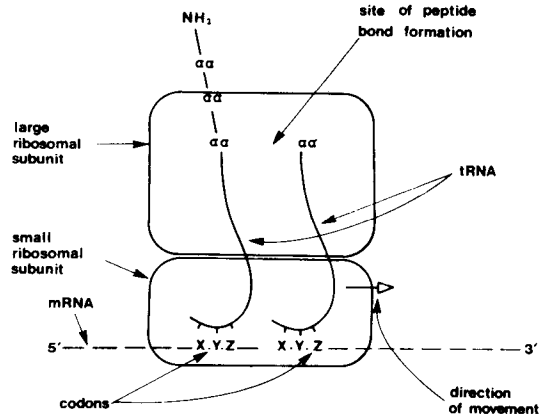

Fig. 1.1 General model of the ribosome.

Nevertheless, as released the polypeptide is not necessarily in its final form. In addition to more folding of the nascent chain and the possible formation of dilsuphide bridges, several other *post-translational modifications* may also occur.

The mechanism of protein biosynthesis is remarkably similar in different organisms. The major differences appear on moving from the prokaryotes to the eukaryotes and are mainly a consequence of the presence of the nuclear membrane in higher cells and the longer half-life of such cells. Other structural differences in the size of ribosomes and structure of the initiator tRNA, mRNA etc. are rather trivial. The genetic code itself is probably completely universal.

2 The molecules involved in protein biosynthesis

In this chapter, each of the molecules, classes of molecule or molecular components involved in protein biosynthesis will be described. Their role in protein biosynthesis, methods of isolation and results of biochemical and biophysical analysis will be discussed.

2.1 Messenger RNA

2.1.1 Discovery and isolation

Early experiments, in which rats were injected with labelled amino acids and the incorporation of radioactivity into polypeptides measured after various chase periods, indicated that amino acids were first assembled into protein in the microsomal fraction of cells, probably on ribosomes attached to membranes. The triplet nature of the genetic code was also deduced at an early date by genetic analysis, but the way in which the genetic information was transferred from DNA to the ribosome remained unclear until the concept of messenger RNA (mRNA) was crystallized in the *lac* operon theory of Jacob and Monod [1].

With this theory in mind, it was quickly established that a new species of non-ribosomal RNA was associated with *E. coli* ribosomes shortly after infection with bacteriophage T4 [2] and this added experimental support to the mRNA concept. However, it proved extremely difficult to isolate mRNA in any quantity and study its biochemical properties. At about the same time (1961), however, further support for a role for RNA in protein synthesis came with the discovery [3] that synthetic polyribonucleotides could act *in vitro* as messenger and direct the polymerization of amino acids into polypeptides. This finding was quickly exploited to elucidate the exact nature of the genetic code (see monograph by Dr. Woods for detailed discussion and Section 2.2). The visualization of polysomes in the electron microscope with a thin strand of material joining adjacent ribosomes, added further weight to the arguments in favour of the existence of a messenger molecule (Fig. 2.1).

Nevertheless, the isolation of cellular mRNA still remained a grave experimental problem, first because it constitutes only a small fraction of cellular RNA and secondly because it is very susceptible to attack by ribonucleases.

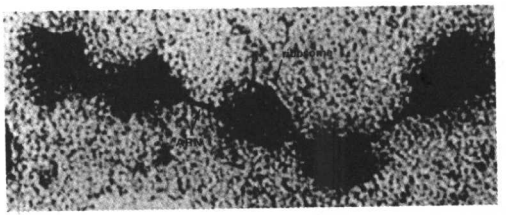

Fig. 2.1 Electron micrograph of a rabbit reticulocyte polysome.

For a long period of time putative mRNA was usually characterized as that heterogeneous population of rapidly labelled, polysome associated RNA, that could be dissociated from ribosomes by treatment with EDTA, and which had a base composition similar to that of bulk

DNA rather than ribosomes (which have a relatively high content of the ribonucleotides guanosine (G) and cytosine (C)).

The isolation of large amounts of RNA with messenger-like properties was first achieved using bacteriophage and viruses. Several coliphage (eg. f_2, R_{17}, MS_2, Q_β) contain no DNA, and their genome appears instead to consist of single stranded RNA (ssRNA), surrounded by a simple protein coat. Similar, very simple animal viruses also exist, the best known being the picornaviruses which include polio, and foot and mouth disease viruses. It was soon realized that such phage and viruses might represent one of the simplest classes of self-replicating organisms whose RNA could function both as the store of genetic information and the mRNA itself. Since the ssRNA constitutes 30% of the mass of particles, the isolation of large amounts of RNA from such sources is not difficult, and its ability to direct the synthesis of the appropriate polypeptides in a cell-free system can then be tested.

The isolation of bacterial mRNA even today presents formidable difficulties because of its inherent instability *in vivo* but one method which is being increasingly used is the enzymatic synthesis of mRNA *in vitro*. The methodology of RNA synthesis is beyond the scope of this monograph (and is discussed in Dr Burdon's book in this series), but using appropriate enzymes and cofactors the accurate and large scale production of mRNA can be achieved *in vitro* providing that a suitable DNA template is available. Early experiments used phage DNA such as T4 but as methods to amplify and purify specific bacterial genes become available, more sophiscated DNA templates are being used, such as purified *lac* operon DNA.

In the past, animal cell mRNA has been isolated from polysomes by disrupting them with a chelating agent such as EDTA [4] and separating the resulting ribosomal subunits and messenger ribonucleoprotein complex (mRNP) on a sucrose gradient. The protein can be removed from the mRNP by treatment with detergent. Many of the more complex animal viruses contain within their virion enzymes which are able to synthesize mRNA. For example, reovirus which has a fragmented double stranded RNA (dsRNA) genome contains an enzyme able to use the dsRNA as template for the synthesis of ssRNA *in vitro*. This and similar viruses have been exploited to make large amounts of eukaryotic mRNA *in vitro*. However, much the easiest and most novel method of isolating eukaryotic cell mRNA makes use of the recent discovery that a sequence of polyA is present in almost all mRNA species. Such polyA can hybridise to polyU impregnated on glass fibre filters, to polyU attached to sepharose or to oligo dT attached to cellulose, whereas other RNA species are not so bound [5]. The polyA containing RNA can be eluted by reducing the salt concentration or by using formamide, and its properties as an mRNA examined. Many eukaryotic mRNAs have been isolated in the last two or three years by utilizing the polyA tail; these include the messengers for globin, immunoglobins, myosin, ovalabumin and several viral proteins (for review see [6].)

2.1.2 Assay of mRNA

The ultimate assay of any mRNA preparation is its ability to direct the synthesis of the protein for which it codes. This can be achieved in many cases by adding the putative mRNA to an appropriate cell-free system containing radioactively labelled amino acids, and after a suitable period of incubation analysing the labelled polypeptides that have been synthesized.

Cell-free systems from bacteria are usually made from cells harvested during the exponential growth phase. These are washed and then ground with alumina in an ice-cold mortar and pestle with buffer containing mono and divalent cations (usually NH_4^+, always Mg^{++}), a sulphydryl protecting reagent and DNAse. Once

the cells are disrupted, the mixture is centrifuged at 30 000 g for 10 minutes to sediment the alumina, cell wall and debris and the supernatant (which is referred to as an S30) is removed. The S30 is next pre-incubated at 37°C under the appropriate ionic conditions in the presence of unlabelled amino acids, ATP, GTP and an energy generating system. During the pre-incubation the endogenous mRNA in the S30 is translated and destroyed. After dialysis to remove the amino acids and other low molecule weight materials, the S30 is stored in small portions at low temperature (liquid nitrogen) [3].

Similar pre-incubated S30 preparations can be made from many eukaryotic cells. One that has been extensively utilized is prepared from Krebs II cells grown as an ascitic tumour in mice [7]. A very useful cell-free system which has low levels of endogenous mRNA activity and is prepared from commercial wheat germ has also been described [8]. A slightly different cell-free system is that prepared from rabbit reticulocytes. Rabbits are made anaemic and the reticulocytes prepared and washed. The packed reticulocytes, which have a very fragile membrane, are then lysed by osmotic shock and the cell stroma removed by centrifugation. The resulting supernatant which is referred to as a lysate is extremely efficient in the cell-free synthesis of globin, the initial rate of synthesis being similar to that in whole cells. Furthermore exogenous mRNAs are also efficiently translated when added to the lysate.

Each of the cell-free extracts mentioned above contain the ribosomes, enzymes, tRNA molecules etc. required for polypeptide synthesis *in vitro*. In addition to the high molecular weight components contained in the extracts amino acids, ATP, GTP and an energy-generating source are required (Fig. 2.2). The concentration of cations is also critical; synthesis is extremely sensitive to changes in Mg^{++} concentration (Fig. 2.3) and to a lesser extent to K^+ or NH_4^+ concentration. Assuming these conditions are satisfied, addition of small amounts of mRNA (1 μg) leads to the incorporation of large amounts

Cellular extract (from bacteria, animal cells, wheat germ etc.)		5 mg/ml
Tris buffer		20 mM
$MgCl_2$		1–5 mM
KCl		60–120 mM
ATP		1 mM
GTP	energy source	0.1 mM
creatine phosphate	energy source	5 mM
creatine phosphokinase	energy source	5 μg/ml
β mercaptoethanol		5 mM
amino acids (except radioactive species)		50 μM
radioactive amino acid		10–200 μCi/ml
mRNA		20–200 μg/ml

Total volume 25 μl–1 ml
Incubation at 20–37°C
For 10–180 min

Fig. 2.2 Requirements for protein synthesis *in vitro*.

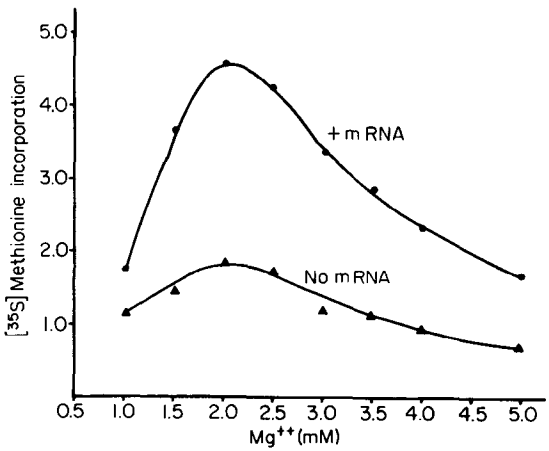

Fig. 2.3 Protein synthesis *in vitro* varies with Mg^{++} concentration (redrawn from [9]).

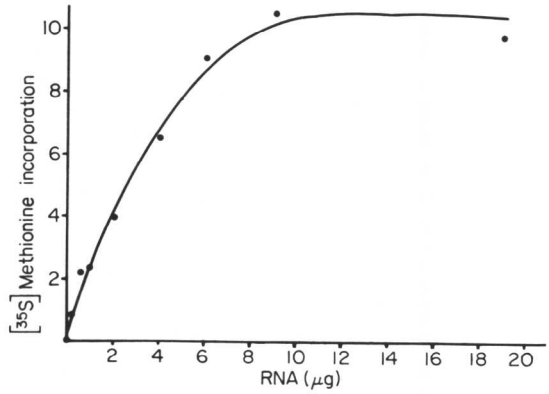

Fig. 2.4 Protein synthesis *in vitro* in response to added viral RNA (redrawn from [9]).

of labelled amino acid into polypeptide (Fig. 2.4) [9].

A somewhat different assay for mRNA activity has been described by Gurdon and his colleagues. Growing oocytes or activated eggs of *Xenopus laevis* are micro-injected with mRNA and incubated with salts and radioactive precursor. Analysis of the proteins made in the injected cells shows that protein coded for by the injected mRNA is synthesized *in ovo* for a period of several days. The oocyte system is favourable because it requires very small amounts of mRNA, the efficiency of translation is high and translation continues for very long periods [10].

The choice of the particular system to translate a given mRNA is not necessarily obvious. Homologous systems (i.e. a mouse extract to translate a mouse mRNA) are preferable in that artefacts of cell-free translation are likely to be minimized, but homologous systems are by no means essential; rabbit globin mRNA is translated quite efficiently in extracts of wheat germ! However, although such an assay can be extremely useful to identify an mRNA it does not necessarily indicate anything about the mechanism and requirements of translation *in vivo*. In general cell-free protein synthesis is extremely inefficient when compared with the process in whole cells and this must always be borne in mind when assessing the physiological significance of data obtained *in vitro*.

The incorporation of labelled amino acid into hot TCA insoluble material (polypeptide), does not indicate that mRNA is being translated accurately *in vitro*, and therefore further analysis is necessary. The commonest method used to analyse the polypeptides made in a cell-free extract is polyacrylamide gel electrophoresis. Ideally, the whole cell-free reaction is reduced and briefly boiled in a solution containing SDS. All polypeptides are thereby denatured and saturated with SDS; they subsequently migrate in the gel depending on their molecular weight. After electrophoresis the gel is stained, destained, dried and auto-

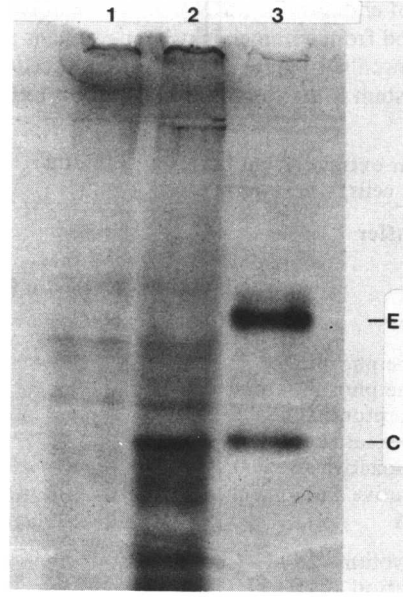

Fig. 2.5 Polyacrylamide gel of protein made *in vitro* in response to Semliki Forest Virus RNA. (1) no mRNA added (2) plus RNA (3) virus proteins; E envelope C capsid (from [9]).

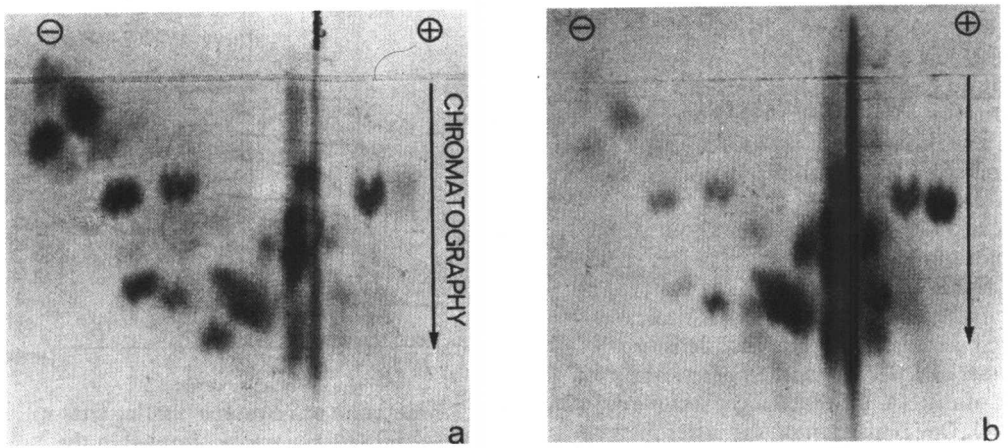

Fig. 2.6 Tryptic peptide fingerprint of (a) mouse globin (stained with ninhydrin) and (b) globin made *in vitro* using ^{14}C amino acids (autoradiograph) taken from [7].

radiographed. Comparison of the radioactively labelled polypeptides in the cell-free product with those present in similar incubations without added mRNA and with appropriate marker polypeptides is used to demonstrate the mRNA-dependent synthesis of the protein in question (Fig. 2.5).

Similarity in the molecular weight of two proteins, however, does not prove their identity and other methods must also be used. Such methods include column chromatography and immunological precipitation of protein made *in vitro*. The most rigorous method, however, is by analysis of the peptides released after digestion with a specific protease. For example trypsin cleaves polypeptides to the C terminal side of arginine and lysine residues. The cell-free polypeptide is therefore digested with trypsin, preferably after previous partial purification by one of the methods mentioned above, and the resulting peptides 'fingerprinted' by separation of the peptides in two dimensions using appropriate combinations of chromatography and/or electrophoresis (Fig. 2.6). Only after several of the techniques described have been used and proved positive can the accurate cell-free synthesis of a given protein be claimed.

2.1.3 Structure

An estimate of the size of purified mRNA can be obtained by sucrose gradient centrifugation or more accurately by polyacrylamide gel electrophoresis under denaturing conditions (for example in the presence of formamide). The size is calculated using a relationship between the distance moved by any particular molecule and the logarithm of its molecular weight. Almost invariably the molecular weight obtained for a given mRNA is greater than that expected for a molecule containing only coding information for the protein in question. For example, the number of amino acids in the α and β chains of rabbit globin are 141 and 146 respectively; allowing three nucleotides per amino acid a molecular weight of about 135 000 daltons would be predicted for the mRNA. In fact the best estimates to date suggest the true molecular weights of the mRNAs are 202 000 and 220 000 daltons [11]. Even allowing for the presence of 50 to 200

nucleotides present in polyA, and the possible loss of part of a nascent polypeptide by post-translational cleavage, all mRNAs appear, like globin, to have non-coding portions. These presumably are required to fold the mRNA, to form a site to which ribosomes can attach and to bind other transcriptional or translational control elements.

In an attempt to learn more about such features of the primary and secondary structure of mRNA, several have been analysed by nucleotide sequencing methods. Classical sequencing methods using optical density amounts of RNA are quite impractical for analysis of mRNA, but the Sanger techniques can be used. One limitation on the latter method however is that the RNA must be labelled at high specific activity with radioactive phosphorus (^{32}P) [12].

Bacteriophage mRNAs can be labelled in this way and spectacular progress has been made over the last few years in the sequence analysis of these molecules. The great length of the ssRNA phage mRNA presents enormous technical problems and in the initial sequence analysis specific portions of mRNA have been isolated and examined. Of particular interest is that region of the mRNA around the ribosome binding site. This can be isolated by first forming an initiation complex containing ribosomes, mRNA, initiator tRNA and initiation factors (but not amino acyl-tRNA or elongation factors) and then enzymatically digesting away all the mRNA not attached to and therefore

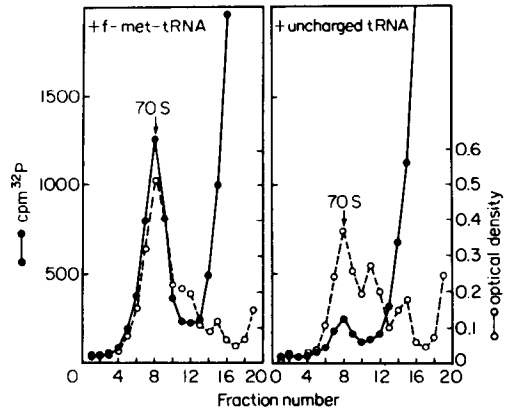

Fig. 2.7 Isolation of ribosome binding sites of R_{17} RNA. Initiation complex formed in the presence of (1) charged tRNA and (2) uncharged tRNA, digested with RNAse and separated on sucrose gradient (from [13]).

protected by the ribosome (Fig. 2.7). Several such bacterial ribosome binding sites have been isolated and sequenced (Fig. 2.8). The sequences are known to be meaningful because they all contain an AUG initiation codon followed by a nucleotide sequence which codes for the amino acids found at the N-terminus of the proteins in question. In many cases the initiating AUG codon appears at the end of a 'hairpin' loop (Fig. 2.9) where presumably it is easily recognized by ribosomes, initiation factors and initiator tRNA, and is thereby able to set the reading frame or phase of translation. Several nucleotide sequences appear common to many

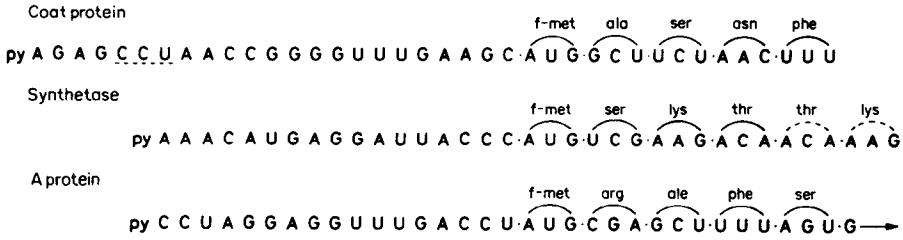

Fig. 2.8 Sequences of R_{17} ribosome binding sites (from [13]).

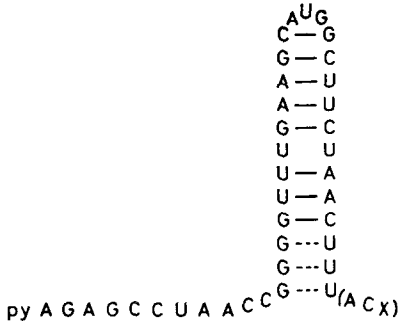

Fig. 2.9 Possible secondary structures of R17 coat protein ribosome binding site (from [13]).

of the binding sites; these may be involved as specific recognition sites. In particular the 3′ terminus of 16S ribosomal RNA (ACCUCC-UAA_{OH}) is complementary to one of these common sequences (GGAGGU) [14] and this region may be directly involved in ribosome binding (see Section 3.3.2).

ssRNA phage messenger codes for three proteins; the replicase, the coat protein and the maturation protein. The ribosome binding site of each of these genes has been sequenced. These then provide markers to order the three genes along the mRNA. The RNA is first cleaved specifically into two halves using specific RNAses, and the fragements identified as coming from either the 5′ or 3′-end of the mRNA by assaying for pppXp (which can only come from the 5′ terminus) and subsequently isolating the ribosome binding sites contained in both the fragments. Such analysis suggested the gene order along the mRNA is maturation: coat: replicase (Fig. 2.10).

Other methods of isolating specific fragments for sequence analysis include the identification of terminal oligonucleotides of the phage RNA, and the isolation by polyacrylamide gel electrophoresis of segments of the molecule partially resistant to ribonuclease attack (presumably they are protected by their secondary structure).

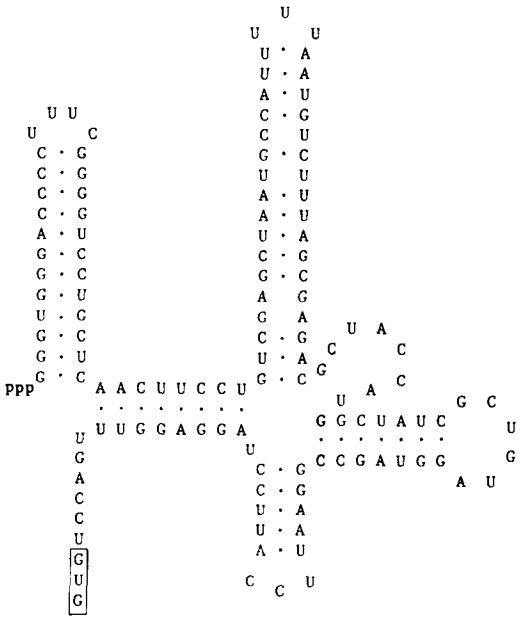

Fig. 2.11 Nucleotide sequence at the 5′terminus of MS2 RNA (from [16]).

ture). In this way much of the nucleotide sequence of several phage mRNAs have been determined. For example the initiation codon of the maturation protein gene is located 130 nucleotides from the 5′-end of MS2 RNA, thus confirming the presence of untranslated regions (Fig. 2.11). The complete nucleotide sequence of the MS2 coat protein gene has also been

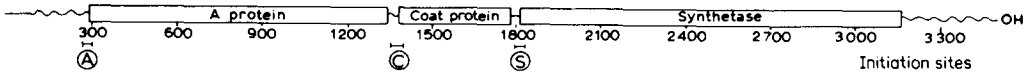

Fig. 2.10 Biochemical map of the R17 genome (from [15]).

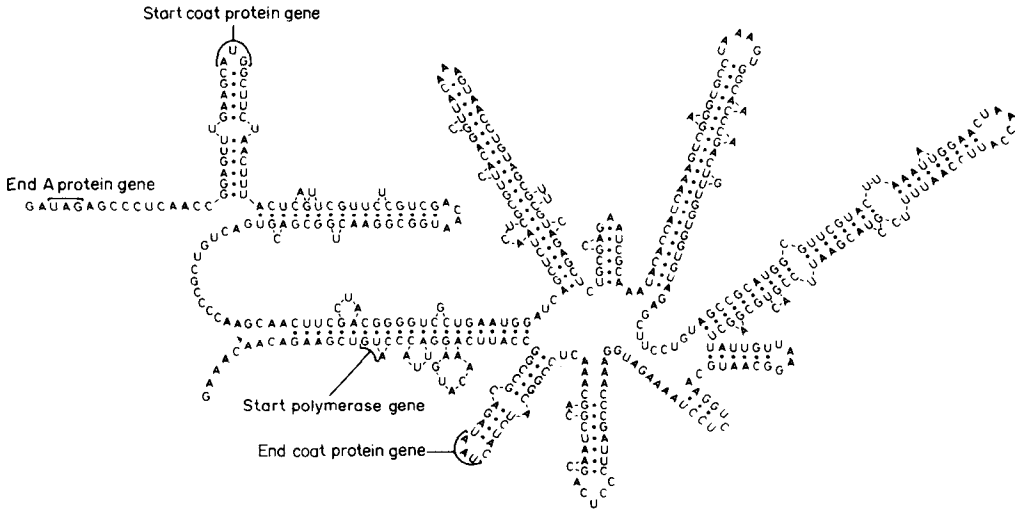

Fig. 2.12 MS2 coat protein gene (from [18]).

established. This can be compared with the amino acid sequence of the coat and this data can be used to compare the genetic code as used in natural mRNA with that obtained with synthetic polymers and other genetic methods [17].

The coat protein gene can be folded into a variety of secondary structures. Analysis of the stability of the possible structures and consideration of the susceptibility of the various parts of the molecule to digestion with RNAse lead Fiers and co-workers to propose the model shown in Fig. 2.12 for the MS2 coat protein gene. The postulated structure shows extensive 'flower-like' folding, as well as details of the initiation, termination and inter-cistronic sites of the gene. The initiation site of the polymerase gene on the mRNA is partially covered by secondary structure with the coat protein gene. This finding explains the earlier observation that ribosomes do not attach to the polymerase initiation site until some time after initiation at the coat protein gene. Presumably until a ribosome has moved down the coat protein gene and thereby disrupted its secondary structure the polymerase ribosome binding site is inaccessible.

Other ssRNA phage messengers have been analysed in this way but all these may be considered atypical of cellular mRNAs because in addition to their messenger function they have to be packaged in the phage particle.

Analysis of other mRNAs, which cannot be extracted in suficient quantity from whole cells or which, in the case of animal cells, cannot be labelled to high specific activity, requires other methods. One approach is to use enzymatic synthesis of the mRNA *in vitro* using purified DNA or DNA fragments, highly purified DNA dependent RNA polymerase and nucleotide triphosphates labelled in the α phosphate position. Alternatively, radioactive ^{32}P or ^{125}I can be introduced either enzymically or chemically into specific positions of unlabelled mRNA and the resulting material analysed by modifications of the Sanger technique [12].

mRNA isolated by gentle methods from animal cells contains associated protein. The