Posttranscriptional Control of Gene Expression in Yeast

JOHN E. G. McCARTHY*

Posttranscriptional Control Group, Department of Biomolecular Sciences, University of Manchester
Institute of Science and Technology (UMIST), Manchester M60 1QD, United Kingdom

INTRODUCTION.................................................................1493
CONCEPTS OF CONTROL IN GENE EXPRESSION.................................1493
DELIVERING A FUNCTIONAL mRNA TO THE SITE OF TRANSLATION...........1494
YEAST TRANSLATION APPARATUS...............................................1495
Initiation Components.........................................................................1498
Binding of an active ternary complex to the ribosome............................1498
Association of mRNA with the cap-binding complex...............................1500
Recruitment via the poly(A) tail—an alternative route.........................1503
Selection of the translational start site................................................1504
Initiation of polypeptide synthesis.......................................................1504
Additional factors involved in translation............................................1504
Reaction pathways and kinetic control................................................1505
Translational Elongation and Termination...........................................1505
Elongation factors................................................................................1506
Release factors....................................................................................1506
Sequence contexts and termination efficiency.....................................1506
Analogous functions in prokaryotes and eukaryotes...............................1507
Termination need not be followed by ribosome release.........................1507
Mitochondrial Translation....................................................................1507

PATHWAYS OF TRANSLATIONAL INITIATION.................................1508
5'-End-Dependent Initiation..............................................................1510
Principles of the scanning model.........................................................1510
(i) AUG context................................................................................1511
(ii) AUG priority rule.........................................................................1511
(iii) Secondary structure blocks initiation..........................................1511
Mechanisms relating translation rates to mRNA structure...................1512
(i) Structure in the prokaryotic TIR..................................................1512
(ii) Structure in the eukaryotic 5'UTR.................................................1512
Mechanism of scanning......................................................................1514

Reinitiation.......................................................................................1515
Prokaryotic reinitiation.......................................................................1515
Eukaryotic reinitiation.......................................................................1515
GCN4............................................................................................1516
(i) There are essentially two types of uORF......................................1516
(ii) The spacing between sites of termination and initiation is critical...1518
(iii) Regulation is dependent on a fine balance between uORF types and intersite spacing......1518
(iv) Regulation of GCN2 kinase activity............................................1518
CPA1.............................................................................................1519
YAP1 and YAP2..............................................................................1520
CLN3 and other examples of uORF-related control............................1522
Internal Initiation in Yeast?...............................................................1522

Alternative Initiation of Open Reading Frames..................................1522
Alternative Coding Potential of Open Reading Frames.........................1523
+1 frameshifting.............................................................................1523
−1 frameshifting.............................................................................1524
Alternative decoding........................................................................1524
Protein splicing................................................................................1524

Prokaryotic and Eukaryotic Translation..............................................1525
Unresolved Issues of Quantitative Control.........................................1526

MECHANISMS OF TRANSLATIONAL REGULATION.........................1526
Modulation of Translation Factor Activities.....................................1527
eIF4E.............................................................................................1527

* Mailing address: Posttranscriptional Control Group, Department of Biomolecular Sciences, UMIST, Manchester M60 1QD, United Kingdom. Phone: (44)161-200-8916. Fax: (44)161-200-8918. E-mail: J.McCarthy@umist.ac.uk.
INTRODUCTION

Much of the excitement in research on eukaryotic gene expression in recent years has been generated by work on the steps of this process that follow transcription. Taken literally, posttranscriptional gene expression includes all of the steps downstream of transcription that are involved in the realization of the coding potential of the genome, encompassing processes from mRNA modification and processing through to protein folding, sorting, transport, and turnover. However, this review focuses on the fate of pre-mRNA and mRNA during its path through the nucleus into the cytoplasm and its subsequent translation and degradation. In particular, research on the interactions between the translational apparatus and mRNA has uncovered many forms of posttranscriptional control. Moreover, it has become increasingly apparent that many of these different types of control are, in a number of ways, interdependent or coupled to each other.

The yeast Saccharomyces cerevisiae has played a key role as subject and/or host of increasing numbers of investigations in this area and remains a popular organism because of the ease with which it lends itself to genetic manipulation and analysis, in vivo phenotypic analysis, and biochemical experimentation. Moreover, the extensive nature of current knowledge of this relatively simple eukaryote, combined with the attention it is receiving from programs of intensive analysis at the genome, “transcriptome,” and “proteome” levels (175, 176, 501, 571), places it first in line for achieving the status of being at least close to comprehensively characterized at some future date. All these points underline the importance of baker’s yeast as an organism for study in an area of research like posttranscriptional control, offering, as it does, an increasingly complete picture of how the investigated mechanisms contribute to the physiology and growth of a whole organism. At the same time, it should be remembered that there are aspects of posttranscriptional gene expression that were exclusive discoveries of the yeast research community, including mRNA-specific translational stress responses, mRNA destabilization mediated by upstream open reading frames (UORFs), positive modulation of mitochondrial mRNAs via nucleus-encoded activator proteins, and autocatalytic protein splicing (all of which are discussed in this review).

This review explores the diversity of posttranscriptional control pathways in S. cerevisiae, focusing primarily on those currently known to be mediated or influenced by ribosome-mRNA interactions. Its content and structure reflect the philosophy that the processes of posttranscriptional gene expression should be considered components of a whole rather than being isolated systems. It has therefore been a general aim to examine the interrelationships between the mechanisms underlying posttranscriptional control and their thermodynamic and kinetic consequences at the molecular level. Given that control can be understood only in quantitative terms, the first section sets the stage by considering appropriate theoretical tools for handling control data. In the body of the review, comparisons with analogous prokaryotic and higher eukaryotic systems are made where these highlight key mechanistic principles. Since this review focuses on the issue of control, it does not attempt to serve as a comprehensive compendium of the literature on the cellular components involved. This has inevitably led to the omission of direct citations of many interesting papers, but the reader is encouraged to seek access to these via the cited reviews by other authors.

CONCEPTS OF CONTROL IN GENE EXPRESSION

The rapid development of techniques of molecular biology, biochemistry, genetics, and structural biology over the last few decades has resulted in an explosive increase in the rate of generation of descriptive information relating to cellular components. However, one of the major challenges of contemporary biology is the formulation of physiologically relevant models that describe how these components function in cellular processes. This depends on a successful transition from qualitative to more quantitative representation of living systems which, in turn, requires that biological mechanisms are increasingly described in terms of their thermodynamic and kinetic properties. However, this remains an uneasy interface between the disciplines of the physical and biological sciences, a problem that is exacerbated by the lack of conventions in the use of appropriate terminology. A prime example is the concept of kinetic control as applied to gene expression. As argued previously (361), there already exists a clear definition of control within the framework of metabolic control theory (276, 277),
and this provides a suitable basis for unambiguous terminology that can be used to describe posttranscriptional events. This also allows the term regulation to be used in a consistent and unambiguous manner.

Two types of approach to the description and analysis of gene expression pathways will be considered briefly in this review. They offer complementary views that can both be helpful in planning and evaluating quantitative experiments. The systemic approach to metabolic control analysis described by Kacser and Burns (276, 277) was conceived as a means of analyzing complex multienzyme systems which are generally too complex to be amenable to accurate analysis by standard kinetic descriptions of the component reactions. This approach can be usefully applied to the partial processive reactions of pathways such as translation, and some of the concepts are introduced here so that the corresponding terminology can be used later in this review without causing confusion.

The key conceptual tools of the analysis are the coefficients used to define “control” in such a complex system. The most relevant of these in the present context is the control coefficient. This describes the relationship between the activity of each catalytic component (E) and the resulting effect on the flux. The activity variation of each E can be caused by a change in concentration, modulation of its kinetic properties, or the binding of an effector. Each change in a given E component (enzyme) can be expressed as a fractional change, \( \frac{\delta E_i}{E_i} \), and this is reflected in a shift to a new steady-state flux, expressed as \( \frac{\delta F}{F} \). The ratio of the latter to the former represents a measure of the effectiveness of the imposed change in altering the flux, and under the condition \( \delta E_i \to 0 \) it represents a definitive property of the system, called the control coefficient: \( Z_i = \frac{\ln F/\delta F}{\ln E_i} \). Z can theoretically assume any value between 1 and 0: a value near 1 means that a very strong controlling influence on the overall flux of the system, whereas a value near 0 corresponds to a comparatively minimal contribution to control and probably applies frequently to components of gene expression pathways. An important constraint defined by the Kacser and Burns analysis is the summation principle, which states that the sum of the Z values must equal 1. Most importantly, this type of analysis emphasizes that control is distributed among the respective E values, with the relative individual contributions being reflected in their respective Z values. Accordingly, the use of the term “rate-limiting” for any chosen step or E in the vast majority of living systems is likely to be misleading and can be meaningless. The Z values are determined by a range of factors; in a multienzyme pathway, for example, these factors include the distance of each enzyme from substrate saturation, enzyme concentration, the relationship between the mass-action ratio and the equilibrium constant for each step, and the role of effectors. Analogous properties of the gene expression pathway also contribute to an equivalent set of Z values. Unless exceptional forces are at play, there is likely to be selection pressure on a cellular pathway to evolve toward a system in which the Z values are not excessively unequal. For example, the provision of certain catalytic components in great excess of their required operational capacities generally makes little sense in terms of cellular energetic housekeeping. Whatever the pathway, this treatment tells us that estimates of Z for the respective components are required so that we can model the balance of control.

Of these three points, the first two remain at an early stage of characterization and provide us with only a few hints about potential sites of posttranscriptional control on mRNA transport. It is generally agreed that, with very few exceptions, only some of the reactions, the concepts and terminology are useful for even relatively qualitative descriptions of the pathways under examination. For convenience, this review continues to use the term “control” in its generally accepted sense to describe the factors determining the (constitutive) rates of biological processes. However, when applied in discussions of the kinetic details of specific pathways, “control” is applied in the sense explained above and is not intended to imply exclusive rate-limitation by any given step or entity.

The systemic type of model contrasts with the more conventional approximations of pathway kinetics based on several assumptions regarding sites of strong controlling influence (see, for example, references 173, 339, and 340). This second approach is discussed in connection with the specific posttranscriptional pathways as these are addressed in the review. As will become apparent, the latter type of model for a partially processive pathway can be regarded as a special case of the more general approach, but the assumptions on which it is based require careful scrutiny.

**DELIVERING A FUNCTIONAL mRNA TO THE SITE OF TRANSLATION**

One of the least well understood areas of gene expression is how polymerase II (PolII) transcripts are transported from the sites of their synthesis in the nucleus to the sites where they are translated (Fig. 1) (see, for example, reference 404). The significance of this problem in terms of posttranscriptional control can be seen in a number of ways. First, the rate of export to the cytoplasm influences the steady-state availability of translatable mRNA. Second, pre-mRNA and mRNA interact with a range of splicing components and/or heterogeneous nuclear ribonucleoproteins (hnRNPs) (141, 581), most of which have to be effectively replaced at some stage by ribosomes and translation factors in the cytoplasm if translation is to occur. Some hnRNPs shuttle between the nucleus and the cytoplasm (439), and these might be of particular importance to the architecture and translation of the cytoplasmic mRNPs (596). Third, mRNA export and translation may occur simultaneously, possibly with vectorial and energetic consequences for the export process.

Of these three points, the first two remain at an early stage of characterization and provide us with only a few hints about potential sites of posttranscriptional control on mRNA transport. It is generally agreed that, with very few exceptions, only
mature mRNAs (bound by RNA-binding proteins [mRNPs]) leave the nucleus (254, 255). A key feature in this respect is the 5' cap structure. PolII transcripts are capped with methylated terminal structures (516), comprising in S. cerevisiae either m'G(5')PPPpAp (relative frequency, 75%) or m'G(5')PPPpGp (25%). Cap methylation is essential for cell viability (351). The cap promotes mRNA export (198), although experiments with an S. cerevisiae strain containing a temperature-sensitive capping enzyme (Ceg1p, which transfers GMP from GTP to the 5' end of the mRNA) indicate that the cap is not essential for splicing, polyadenylation, or transport (158, 493). In another study, a hammerhead ribozyme was used to catalyze in cis cleavage of a fusion mRNA in S. cerevisiae (144). The capless downstream product was undetectable by standard blotting techniques, as would be expected if the cap is important for nuclear export and/or stability. From the above-mentioned work on CEG1 mutants (158, 493) and further studies described in the section on mRNA stability in this review, it would now seem that the cap influences stability more than transport. At least for histone mRNA, 3' end formation also stimulates the transport process (143).

A heterodimeric nuclear m'G cap-binding complex (CBC), comprising two cap-binding proteins (CBP20 and CBP80), has been identified (256), but its role is unclear and it is not essential in yeast (404). In a wider context, discussions about whether nuclear mRNA moves through a “track” (461, 601) or via “channelled diffusion” (612) and discussions about the interactions between mRNA and various nuclear components, including the nucleoskeleton (47), splicesome components (257, 324, 511, 614, 615), nucleolus (537) and nuclear membrane and/or pore complexes (238), all of which could theoretically influence the transport process, are still under way. To what degree the gene expression processes in the nucleoplasm are structurally or functionally compartmentalized is controversial (505).

The third point raises the issue of how mRNAs find their way to translationally competent ribosomes. Perhaps the most relevant data have come from studies of the Balbiani ring granule, a large RNP particle, in the dipteran Chironomus tentans (367). These results indicate that the particle generally exits the nucleus 5'-end first and is bound by ribosomes before the 3' end passes through the nuclear pore. However, there is no evidence that mRNA export per se requires translation (41). Studies of S. cerevisiae continue to yield new clues about the process of mRNA export. For example, recent evidence indicates that the ATP-dependent RNA helicase Dhp5p, which is a DEAD-box protein closely related to eIF4A, is involved in mRNA export through the nuclear pore complexes (507a, 552a).

While translatable mRNAs are undoubtedly generated primarily by PolIII promoters, there is evidence that the pathway outlined above is not the only possible route from nuclear gene to cytoplasmic protein in the eukaryotic cell. Notable in this context is the demonstration that the PolIII promoter of the adenovirus type 2 VA RNAI gene generates uncapped and nonpolyadenylated RNA, which is nevertheless translated, albeit poorly, in HeLa cells (191). This indicates that capping and polyadenylation are not essential for nuclear transport or translation, although it has yet to be determined whether the route taken by such PolIII transcripts may allow them to escape restrictions otherwise imposed on their PolII counterparts. Similarly, it has since been shown that in S. cerevisiae, HIS4 can be transcribed from a PolI promoter to generate primarily uncapped but polyadenylated mRNAs that are poorly translated and rapidly degraded (338a). Overall, these data are in accord with a theme that threads its way through a number of the processes of gene expression: redundancy of function and/or parallel routes provide alternatives for many key steps. What therefore appears to be a major pathway may not be dictated by fixed mechanistic limitations but, rather, may be guided by kinetic or thermodynamic principles of control.

It is clear from the above that there remains considerable uncertainty about the role of any of the nuclear events in posttranscriptional control. Looking beyond the nuclear membrane, there are various lines of evidence for selective distribution of exported mRNA to specific sites within a cell (129, 504, 594) or within whole organisms, for example in Drosophila embryos (504, 589). The fact that mRNA is observed associated with microtubules and actin filaments suggests that these components of the cytoskeleton may be responsible for (selective) mRNA transport (35a).

In higher eukaryotes, mRNA partitioning is involved in developmental processes including the establishment of cell polarity and morphogenesis (118, 196, 288) and in the response to signals from the cell surface (89a). Recent evidence also indicates that at least one yeast mRNA (ASH1) becomes localized within the cell by virtue of its association with the cytoskeleton (535). The 3′ untranslated region (3′UTR) of this mRNA is necessary for its transport to the distal tip of daughter buds in postanaphase cells. To what extent mRNA sorting plays a role in the yeast cell cycle remains to be established.

Posttranscriptional gene expression begins with a nascent transcript, which goes through a highly complex series of nuclear interactions before emerging into the cytoplasm. However, each transcript is much more than simply an intermediate carrier of genomic coding sequences. A single mRNA can contain several different types of signal element that contribute to one or more forms of posttranscriptional control (Fig. 2). The 5′- and 3′-terminal modifications have a number of general functions that affect the whole mRNA pool. Apart from its role in nuclear export (198), the cap is required for efficient translation (471, 484) and also influences mRNA stability (162), although the extent to which each function overlaps with the others remains unclear. At the 3′ end, the mRNA carries a poly(A) tail (initially 60 to 90 adenylate residues in yeast [186, 472]), which also influences the cytoplasmic expression and fate of yeast mRNAs (264, 476). The functions of the untranslated, flanking regions of yeast mRNAs have come under increasing scrutiny in recent years since they, unlike the mRNA modifications, can contain a number of signals that modulate the expression of specific genes in individual ways. Finally, the main reading frame of the mRNA not only constitutes a decodable codon sequence, but also can contain further information in the form of linear signals or conformational blueprints that influence posttranscriptional gene expression, although little is known about them at present. Overall, mRNA carries much more information than merely its coding sequence. The central challenge is understanding how this additional information is able to exert its influence via interactions with the cellular machineries responsible for translation and mRNA turnover.

**YEAST TRANSLATION APPARATUS**

Review articles dealing with the cellular translation apparatus have appeared very recently for mammalian cells (369, 418, 508) and plants (69, 166) and somewhat less recently for yeast (228, 332, 553), and the reader is directed to these reviews for more detailed information about the individual components of the respective systems. In this section I focus on the properties of yeast translation that have most relevance to the known
posttranscriptional control mechanisms and how the latest research has shaped the current view of them. The primary pathway of translation in *S. cerevisiae* is initiated via a cap-dependent mechanism that seems to follow broadly the main pathway that has been delineated on the basis of biochemical studies with mammalian cell extracts. Initiation is not only the most complex step of translation but also a major site for regulation of individual and global gene expression at this level. The subsequent elongation and termination of the polypeptide chain also follow the same general pattern seen in mammalian cells. The overall similarities between the translation machineries of the higher and lower eukaryotes offer the advantage that results gained with both types of system contribute to a general eukaryotic picture of translation. Nevertheless, yeast translation is by no means a carbon copy of its higher eukaryote counterpart, and the increasingly apparent differences between the respective systems can provide additional insight into the structural basis for specific functions in this process. In the following, “yeast translation” refers to protein synthesis in the cytoplasm. However, it should not be forgotten that a very small proportion of yeast proteins are synthesized in the mitochondria.

Before moving on to describe the yeast translational apparatus in more detail, it is useful to consider how the three stages of translation are kinetically related to each other (Fig. 3). Their respective kinetic characteristics are relevant to the functions of the individual components of the translation apparatus; the potential contributions of initiation, elongation, and termination to the control of flux through the whole system; and the interdependency of these three phases. A striking aspect of eukaryotic translation is the processive nature of the events occurring after initiation and the at least partly processive nature of the initiation phase. This has consequences for the ways in which rate control can be exercised on translation (Fig. 3).

For example, there are kinetic arguments why initiation can be expected to figure so prominently in terms of posttranscriptional control. Given certain apparently reasonable assumptions, it is relatively easy to arrive at a simplified (nonrigorous) model of the translation pathway (Fig. 3). This scheme summarizes some basic principles concerning potential control points in the protein synthesis pathway. For the sake of simplicity, the steps of translation are represented by flow rates (\(j\) values, as defined in Fig. 3). It is assumed that the rate of release of 40S ribosomal subunits from the mRNA during the scanning phase is negligible and that the scanning rate itself is relatively high. The maximal attainable rate of initiation of protein synthesis is dictated by the time taken for the 80S ribosome to clear the AUG region, making space for the next approaching 40S subunit. The region blocked by one ribosome is approximately 30 nucleotides (597); therefore, \(j_T\approx j_{E}/30\). The \(j_E\) term used here may be adequately described as an average elongation term, but at least in certain mRNAs it may have to be qualified as referring to only part of the open reading frame (ORF). There is evidence that pausing occurs within eukaryotic reading frames, causing ribosomes to “stack,” at least in certain regions of the mRNA (597). If termination of protein synthesis (\(j_T\)) were to proceed at a much lower rate than initiation (\(j_I\)), there would be a blockage that would feed back from the 3' end of the ORF, distorting the structure of polysomes. This is not known to happen (49), but there is evidence for a pause near the termination codon (597), and so it would seem reasonable to assume that for an mRNA with an unstructured leader, \(j_T\approx j_{L}\). The scanning component of \(j_{L}\) is unlikely to be slow compared to \(j_{L}\) on unstructured leaders, since this would result in a high loading density of 40S subunits on the
5' UTRs of all mRNAs, for which there is no evidence. Given that pausing seems to occur at the start codon (232, 260, 296, 597), we can assume that \( j_B \) for an unstructured leader. However, the \( j_B \) term will be greatly reduced in the presence of structure, thus changing this relationship to \( j_B, j_I \). The above set of principles and assumptions yields a model in which the individual steps of protein synthesis are well matched, allowing the most efficient throughput of ribosomes on an mRNA molecule.

It should be pointed out, however, that while restrictions on \( j_B, j_I \), or \( j_T \) can seriously disturb this balance, changes in \( j_E \) (at least over a certain range) can be more readily accommodated. In other words, applying the terminology of the Kacser and Burns (276, 277) approach to this processive series of reactions, the control coefficient for elongation is considerably smaller than for the other reactions. A low average rate of elongation \( (j_{E1} \text{[Fig. 3B]}) \) may not greatly affect the relative rate of production of complete polypeptide chains on a specific mRNA unless it results in restricted access to initiating ribosomes \( (j_I) \). Increasing the rate of elongation \( (j_{E2} \text{[Fig. 3C]}) \) will also have no effect on the number of polypeptide chains completed in unit time under steady-state conditions unless \( j_T \) (and/or \( j_T \)) changes. The major difference between the cases in Fig. 3B and C in the steady state will therefore be the density of ribosomal binding. This, in turn, will influence primarily the number of ribosomes bound up in the process of protein synthesis at any one time and hence the availability of free ribosomal subunits in the cellular pool. Where the rate of elongation on individual mRNAs is modulated, for example via the internal nucleotide sequence (codon usage), the impact on the cellular ribosome pool will be negligible, so that the overall steady-state rate of polypeptide production will be relatively

![Diagram](http://mmbr.asm.org/)

**FIG. 3.** Rate control exercised at different steps of translation. (A) The general scheme indicates the flow rates \( (j \) values, in events per unit time for binding and release and in nucleotides per unit time for elongation) assigned to the respective steps of 40S ribosomal subunit binding \( (j_B) \), 60S junction \( (j_I) \), elongation \( (j_E) \), and 40S/60S release \( (j_T) \). For the purpose of illustrating certain general points, the release rates for 40S and 60S are assumed here to be identical, although this is unlikely to apply to at least some mRNAs. (B) At a low relative rate of elongation \( (j_E_1) \), ribosome packing on the mRNA is high. (C) A reduced packing density occurs at a higher elongation rate \( (j_E_2) \). However, variation in \( j_E \) need not have a strong effect on overall ribosomal throughput on a given mRNA if the \( j_B \) and \( j_I \) rates are not too high. (D) On the other hand, if termination and initiation are coupled, \( j_T \) may exercise an important control function on translation as a whole.
By far the most common route is the initiation can occur in at least three ways on cellular mRNAs. On the translation pathway has been considered. Focus, they do not yet enable us to model translation adequately. Rigorous analyses reviewed so far bring the issue of control into account the differences so far identified in the yeast system primarily from experimental work performed on mammalian proteins. The pathway depicted in Fig. 4 has been adapted to take into account the differences so far identified in the yeast system. Unfortunately, the disparities between the genetic nomenclatures for the respective organisms makes this subject area highly confusing for the non-specialist. One potential solution would be to adopt a new systematic nomenclature that is related in a readily deducible fashion to the biochemical designations. A proposal of this nature has been made for Schizosaccharomyces pombe which could easily be applied generally (Table 1) (334). Of the yeast initiation factors identified so far, eIF4A4, eIF4G, and eIF5A have been found to be encoded by duplicated genes. Like many other duplicated genes in S. cerevisiae, they are at least partially phenotypically redundant (180, 333). Indeed, the eIF4A products are identical. It has been suggested that the maintenance of at least some duplicated genes may reflect past adaptation of the organism to a changing environment (595), but to what extent this type of selective pressure is directly applicable to the initiation factor genes is unknown.

**Initiation Components**

As discussed later in this review, eukaryotic translational initiation can occur in at least three ways on cellular mRNAs. By far the most common route is the 5'-end-dependent pathway, in which ribosomes apparently select the initiation site via processive scanning along the 5' region of the mRNA. The most striking aspect of the cellular system involved in this pathway is the number and complexity of its components. Apart from the subunits of the eukaryotic ribosome, there are at least 11 eukaryotic initiation factors (eIFs) comprising more than 25 polypeptides (369). These are presented in Table 1. Unfortunately, the disparities between the genetic nomenclatures for the respective organisms makes this subject area highly confusing for the non-specialist. One potential solution would be to adopt a new systematic nomenclature that is related in a readily deducible fashion to the biochemical designations. A proposal of this nature has been made for Schizosaccharomyces pombe which could easily be applied generally (Table 1) (334). Of the yeast initiation factors identified so far, eIF4A4, eIF4G, and eIF5A have been found to be encoded by duplicated genes. Like many other duplicated genes in S. cerevisiae, they are at least partially phenotypically redundant (180, 333). Indeed, the eIF4A products are identical. It has been suggested that the maintenance of at least some duplicated genes may reflect past adaptation of the organism to a changing environment (595), but to what extent this type of selective pressure is directly applicable to the initiation factor genes is unknown.

**TABLE 1. Translation initiation factors in yeast**

<table>
<thead>
<tr>
<th>Initiation factor with subunits (reference)</th>
<th>Proposed function*</th>
<th>S. cerevisiae gene name(s)</th>
<th>Proposed new gene designation(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eIF1d (106, 607)</td>
<td>Met-tRNA, and mRNA binding to 40S</td>
<td>SUI1</td>
<td>tif1</td>
</tr>
<tr>
<td>eIF1Ad (580)</td>
<td>40S-60S dissociation; Met-tRNA, binding</td>
<td>TIF11</td>
<td>tif1</td>
</tr>
<tr>
<td>eIF2, α, β, γ (94, 135, 200)</td>
<td>AUG selection</td>
<td>SU2, SU3, GCD11, GCD3</td>
<td>tif211, tif212, tif213</td>
</tr>
<tr>
<td>eIF2B, α, β, γ, δ, ε (226)</td>
<td>Guanine nucleotide exchange on eIF2</td>
<td>GCD7, GCD1, GCD2, GCD6</td>
<td>tif221, tif222, tif223, tif224, tif225</td>
</tr>
<tr>
<td>eIF3, α, β, χ, δ, ε, η, θ (169, 199, 399)</td>
<td>Met-tRNA, and mRNA binding to 40S; 40S-60S dissociation</td>
<td>PRT1, GCD10</td>
<td>tif31, tif32, tif33, tif34, tif35</td>
</tr>
<tr>
<td>eIF4A (333)</td>
<td>RNA-binding helicase, ATPase</td>
<td>TIF1, TIF2</td>
<td>tif41A, tif41B</td>
</tr>
<tr>
<td>eIF4B (13, 98)</td>
<td>RNA binding promotes helicase</td>
<td>TIF3 (STMI)</td>
<td>tif42</td>
</tr>
<tr>
<td>eIF4E (8, 63)</td>
<td>Cap binding</td>
<td>CDC33</td>
<td>tif45</td>
</tr>
<tr>
<td>eIF4G1 (p150), eIF4G2 (p130) (180)</td>
<td>Interactions with eIF3, eIF4E, and Pab1p</td>
<td>TIF4631 and TIF4632, respectively</td>
<td>tif47A and tif47B, respectively</td>
</tr>
<tr>
<td>eIF4H (458)</td>
<td>Stimulation of activities of eIF4F components and eIF4B</td>
<td>TIF5</td>
<td>tif5</td>
</tr>
<tr>
<td>eIF5 (85)</td>
<td>Ejection of eIFs</td>
<td>TIF51A, TIF51B</td>
<td>tif51A, tif51B</td>
</tr>
<tr>
<td>eIF5A (490, 620)</td>
<td>Function unclear; mutation affects mRNA stability</td>
<td>TIF5</td>
<td>tif5</td>
</tr>
<tr>
<td>eIF6 (369)</td>
<td>40S:60S dissociation</td>
<td>TIF51A, TIF51B</td>
<td>tif51A, tif51B</td>
</tr>
</tbody>
</table>

* These are functions proposed primarily on the basis of in vitro investigations of (partially) purified components.

* These are the designations proposed for the S. pombe initiation factors (334).

* Note the recently discovered existence of eIF2 (59).

* Mammalian eIF1 and eIF1A have now been implicated in start codon selection (435a).

The chief theoretical constraints required to explain how initiation can feature as a step with strong controlling influence (a large control coefficient). The processive nature of elongation means that translation can be represented by a simplified model in which the elongation process is viewed as a single step (i.e., elongation is represented by simply f_k). This and other assumptions result in a greatly simplified treatment, which has formed the basis for previous analyses of eukaryotic translation (see, for example, references 173, 339, and 416). One of the questions raised by these analyses is whether the mode of rate control leads to differential attenuation of the translation of individual mRNAs upon shifting a cell from good growth conditions to more restrictive ones. If it is assumed that the poor translation of an mRNA carrying stable secondary structure or a uORF in its leader is based on poor selection of that mRNA by the ribosome, it might be expected that the shift from saturating or near-saturating activity of the translational apparatus to a reduced capacity will affect the poorly translatable mRNAs more severely. However, as discussed below, it is not clear whether selection of an mRNA (i.e., the initial step[s] of initiation) is affected by the structural elements located in the leader. Thus, while the semirigorous analyses reviewed so far bring the issue of control into focus, they do not yet enable us to model translation adequately. This point will be revisited once further information on the translation pathway has been considered.
during the initiation process (423a). The component subunits of eIF2 (α, β, and γ) are all essential for cell viability, and mutations in them affect start codon selection by the ribosome (93, 185, 369). Moreover, Cigan et al. (93) also used mutations in the anticodon of one of the four tRNAi genes in *S. cerevisiae* to establish the key role of tRNAi-start codon interactions in initiation site selection. Appropriate compensatory mutations in the anticodon allowed initiation on the *HIS4* reading frame by using non-AUG codons that would otherwise not be recognized by the preinitiation complex. The order of the interactions between eIF2-GTP-Met-tRNAi, the ribosome, and the mRNA is not fully resolved. The bulk of the available in vitro data points to 40S–ternary-complex binding preceding mRNA binding, but Trachsel discusses conditions where this might not apply (552). Certainly, the phenomenon of reinitiation constitutes evidence that close association of 40S subunits and mRNA can occur in vivo before the binding of the ternary complex, at least downstream of a termination event (see below). Finally, earlier this year there was an exciting development in this area. *S. cerevisiae* was found to possess a homologue of *E. coli* IF2, called yIF2 (89b). The encoding gene, *FUN12*, is not essential, but its deletion imposes a severe slow-growth phenotype and a marked translation initiation defect. Biochemical experiments indicate that yIF2, like eIF2, func-
promote Met-tRNA<sub>i</sub> binding to the 40S ribosomal subunit (89b). The tantalizing question yet to be resolved is why <i>S. cerevisiae</i> uses both yIF2 and eIF2; does this represent an intermediate stage in ongoing evolution of the translation system in this organism, or are there particular reasons for the maintenance of these parallel functions?

**Association of mRNA with the cap-binding complex.** eIF4E is the cap-binding component of the initiation factor complex eIF4F, anchoring this complex to the 5' end of capped mRNAs. It is the least variable of the eIF4F components (in terms of presence in the complex and/or protein sequence) and one of the less abundant eIFs (estimated to be present at levels greatly substoichiometric with respect to those of the ribosome [see below]). eIF4E is required for efficient translational initiation in vitro (11), and it is thought that it fulfills the same functional role within the assembled eIF4F complex in vivo. In mammalian eIF4F, eIF4E is associated with two other factors, eIF4G and eIF4A, whereby eIF4G holds the respective factors together in the complex (Fig. 5). Recent reports (121, 168) describe a second human eIF4G gene (encoding eIF4GII), which is 46% identical to the initially cloned gene (eIF4G1 [605]) (Fig. 5), and also a second human eIF4E gene, which encodes a protein differing at only two positions. eIF4GII is likely to be a functional homologue of eIF4GI, and there is speculation that the different forms of this protein may be required for a differentiated response to developmental signals (181). The second eIF4E species is likely to be fully functional, and the reason for the duplication has yet to be ascertained. On the other hand, a human protein (4E homologous protein [4EHP], with 30% identity to eIF4E, has also been described, but as yet it has no apparent function (460a).

Mammalian eIF4A and eIF4F exhibit ATP-dependent bidirectional RNA helicase activity that is enhanced by eIF4B (467). There has also been a very recent addition to the eIF4 group of factors, called eIF4H (458). This factor stimulates the activities of eIF4F and eIF4B in a rabbit reticulocyte lysate, although its role in in vivo translation remains unclear. In <i>S. cerevisiae</i>, the isolatable eIF4F complex contains eIF4E and only one other type of initiation factor, eIF4G. The binding domain of the latter has a high affinity for eIF4E (the <i>K<sub>d</sub></i> is...
estimated by surface plasmon resonance analysis to be approximately 10^{-12} \text{ [449]}. However, two other entities, a protein called p20 and the poly(A)-binding protein Pab1p, are also found associated with the eIF4F components of S. cerevisiae; p20 competes with eIF4G for binding to eIF4E, and Pab1p binds to eIF4G via a site on eIF4G (10, 15, 539). Moreover, other interactions are also suspected to occur at least transiently (see the sections on eIF4A and eIF4B below). eIF4G also occurs in two forms in S. cerevisiae, eIF4G1 and eIF4G2, and the former type (p150) is larger than the latter (p130) (180) (Fig. 5).

Mammalian eIF4G is much larger than its yeast homologues and has binding sites for eIF4E, eIF3, and eIF4A (312, 346). It is therefore appealing to regard eIF4G as a sort of docking protein or adapter (214, 312, 383, 479, 509) for the assembly of the complex between mRNA and the preinitiation complex. There is also a further family of eIF4G-like proteins in mammalian systems. For example, one of them, p97, shows 28% identity to the C-terminal two-thirds of eIF4G and thus apparently lacks an eIF4E-binding site (243). Other versions of this 97-kDa protein have been isolated from various sources (see the summary in reference 383). p97 may act as a translational repressor by forming translationally inactive complexes with eIF4A and eIF3, although other properties of this protein may be involved (383). In this context, it should be noted that mammalian eIF4G can be cleaved by viral proteases (e.g., protease L of poliovirus or protease 2A of foot-and-mouth disease virus) to leave a C-terminal product lacking the eIF4E-binding domain (312, 409). The resulting C-domain can support cap-independent or internal ribosome entry site (IRES)-dependent translation in vitro (409). One of the known plant eIF4G factors, p86, is considerably smaller than the other known eIF4G proteins (6, 68, 373). This protein, also known as Elfiso-4G, seems to be lacking much of the N-terminal domain present in its yeast and mammalian counterparts (Fig. 5). It is found associated with eIF4E in the alternative plant complex referred to as eIF4E-4F (68, 373). Despite the inclusion of the much smaller version of eIF4G, the latter complex is functionally equivalent to eIF4F in a range of in vitro assays (320, 373). The other plant eIF4G protein, p220, has yet to be fully characterized (69).

It is not known in which order the eIF4F-mRNA complex is assembled in vivo, nor is the potential functional significance of any given order fully clear. However, the occurrence of cooperativity effects (see later in this section) is likely to dictate a preferred route for the formation of a cap-associated complex. Moreover, the association of eIF4E with eIF4G can be at least partially blocked by the binding of eIF4E-binding proteins (4E-BPs) (421). In S. cerevisiae, there is currently only one candidate for this role, p20 (10, 315), which in fact has a molecular mass of approximately 18 kDa (613). Whether p20 constitutes the full yeast equivalent of a mammalian 4E-BP is discussed in the section on translational regulation (see below).

The X-ray structure of an N-terminal truncated form (amino acids 28 to 217) of mouse eIF4E has recently been determined at 2.2-Å resolution (354). The structure correlates well with previous genetic and biochemical data and also provides the basis for a number of important predictions. It is therefore worthwhile considering certain of its details at this point. Since the truncated mouse protein still binds the cap analogue 7-methyl-GDP and since an equivalent N-terminally truncated form of S. cerevisiae eIF4E (amino acids 30 to 213) has been shown to support growth in an otherwise eIF4E-deficient strain of yeast (569), it is clear that the N-terminal region is not essential for the maintenance of (at least partial) structure and function by the eIF4E protein. The determined structure comprises one domain with an overall shape resembling a cupped hand. Within this domain there is an eight-stranded antiparallel β-sheet, three long α-helices, and one short α-helix. The short α-helix and the concave surface of the β-sheet form the cap-binding slot. Of particular interest are the locations of a number of the residues that are conserved between the various eIF4E sequences that have been sequenced so far (Fig. 6). The methylated base of the cap analogue fits between the tryptophan residues at positions 56 and 102, and the interaction is most likely driven by π-π stacking enthalpy as predicted previously by Ishida et al. (246–248; see also reference 9). Apart from these stacking interactions, there are hydrogen bonds or van der Waals contacts between the methylated G of the cap and Trp102, Gln123, and Trp166; direct interactions between the ribose and diphosphate groups of the cap structure and Trp56, Arg157, and Lys162; and water-mediated contacts with Trp166 and Arg112. All of these residues are either fully conserved or subject to only conservative changes between the respective eIF4E species (Fig. 6). The results of mutational studies on a number of these residues in the yeast and human eIF4E proteins, in particular the tryptophans, have generally been consistent with their playing a role in cap binding (9, 382).

Equally interesting are the absolutely conserved surface residues: a nonpolar group of Val69, Trp73, Leu131, and Gly139 and an acidic group of Glu70, Glu140, and Asp143 (354). These surface areas are potential candidates for the binding sites for other factors, including eIF4G and 4E-BP1 (see below). On the other hand, the crystal structure also reveals that Ser209 is located near the cap-binding slot. Moreover, on the basis of model building, Marcotrigiano et al. (354) propose that Lys159, which lies on the other side of the slot, could form a salt bridge with the phosphorylated form of Ser209, perhaps stabilizing the binding of mRNA in the slot. This hypothesis has been tested in vitro using the human protein. It was reported that the addition of anionic charge at the position Ser209 by means of mutation reduced the off-rate of eIF4E from the m^7GpppG cap (501a). Moreover, examination of this question in S. pombe may also prove to be highly informative, especially with regard to analysis of the physiological significance of phosphorylation at this site (see below). Overall, a key challenge generated by progress in the X-ray crystallography of initiation factors will be to establish the functional significance of the molecular features that have been identified.

The structure of the S. cerevisiae eIF4E-cap analogue complex was analysed by NMR using a protein-CHAPS micelle (358). While largely similar to the mouse eIF4E X-ray structure, the proposed yeast nuclear magnetic resonance spectroscopy (NMR) structure shows some small differences. Five of the β-strands and one α-helix are shorter in the yeast protein, and Trp58 in the cap-binding site has a different orientation from that of its mouse counterpart. Matsuo et al. (358) also described a complex between yeast eIF4E and mammalian 4E-BP2 and showed that the NMR resonances were particularly perturbed in yeast eIF4E at amino acid positions 32 to 50 and 62 to 79, although the resonances of 13 further amino acids in the region 85 to 169 were also affected. At least some of these are surface residues, and it will now be necessary to establish whether any of them participate directly in 4E-BP binding.

More recent work has combined classical genetic, immunological, and biochemical methods with surface plasmon resonance analysis to obtain information about the residues in yeast eIF4E that contribute to or influence the binding sites for eIF4G and p20 (449). Mutations at HPL37–39, W75, E72, V71, and G139 were found to decrease the affinity of eIF4E for a recombinant protein bearing the eIF4E-binding domain of
eIF4G. These residues are all highly conserved among the eIF4E proteins sequenced so far and map to a cluster on eIF4E that is located on the dorsal side of the structure relative to the cap-binding slot (Fig. 7). It is, however, striking that the binding site for p20 overlaps but is not identical to that of eIF4G (Fig. 7).

The characterization of binding sites for both eIF4E and eIF3 on eIF4G (Fig. 5) has led to the suggestion that eIF4G mediates mRNA-ribosome association by linking eIF3 and eIF4E (312, 369). Mammalian eIF4G also has two binding sites for eIF4A near its C terminus (Fig. 5) (242), whereas an equivalent site has yet to be identified in S. cerevisiae eIF4G, and

FIG. 6. Conserved sequence and structural motifs in eIF4E. Comparison of eIF4E sequences from a range of different organisms reveals the presence of many strictly conserved or conservatively maintained features (boldface type). A number of these are involved in binding the mRNA cap structure (arrows), while others are surface residues (dots), some of which have the potential to be involved in interactions with other proteins (such as eIF4G, 4E-BPs, or p20) (354, 358, 449). Asterisks mark the positions of serine residues in the respective eIF4E sequences that have either been shown or are suspected to be sites of phosphorylation. The sequences shown belong to eIF4E proteins that bind preferentially to the m7GpppX type of mRNA cap. Caenorhabditis elegans has multiple forms of the cap-binding protein (not shown here), at least two of which also recognize m3Gppp2,2,7G caps (269a). Relatively few differences in the primary sequences apparently suffice to confer this broader specificity on the C. elegans cap-binding proteins.
eIF4A has not been found to copurify with the yeast complex. It is accordingly not yet clear to what degree the mammalian and yeast eIF4F complexes should be regarded as functionally equivalent. One potentially critical feature of the formation of these various complexes may be that the interactions cause changes in the binding characteristics of the respective components. For example, cross-linking experiments have indicated that the binding of mammalian eIF4G to eIF4E increases the latter factor's affinity for the 5' mRNA cap (194), and the cap-binding affinity of *S. cerevisiae* eIF4E has been estimated to increase at least 10-fold upon binding of the eIF4E-binding domain of eIF4G and of p20 (V71 and W75) are dark grey. The other amino acids seem to influence only binding to the eIF4E-binding domain of eIF4G (E72, H37, P38, L39, and G139).

Moreover, beyond the functional complexity of eIF4F itself, there may be other translation components that can perform parallel functions. For example, at least the mammalian eIF4B protein may also be capable of mediating complex formation between mRNA and ribosomes via eIF3 and/or interactions with rRNA (371, 372). It remains to be seen whether this capability acts in concert with the equivalent functions of eIF4G or defines a parallel (alternative) means to the same end.

**Recruitment via the poly(A) tail—an alternative route?** Recent work has shown that *S. cerevisiae* eIF4G1 and eIF4G2 also have N-terminal binding sites for the poly(A)-binding protein.
(539, 541) (Fig. 5). This and other observations (165, 167, 320, 392) have revitalized the debate about the roles of the poly(A) tail and the poly(A)-binding protein in translation (see, for example, the recent review of this theme by Jacobson [264]). Speculation about possible “long-range” interactions between Pab1p and the ribosome had already been stimulated by the isolation of pab1 suppressor strains that harbored mutations in genes encoding 60S subunit proteins (474). Moreover, as discussed below, there is in vitro evidence that the poly(A) tail participates in the recruitment of ribosomes onto mRNAs via a pathway that can function independently of the cap. Investigations in vivo (445, 446, 541), however, paint a more complex picture (see below). The poly(A)-binding protein of higher and lower eukaryotes has four RNA recognition motifs. These contribute to differing degrees to poly(A)-specific and non-poly(A)-specific RNA binding (72, 114, 306, 403, 478). The second RNA recognition motif of Pab1p is required for binding to eIF4G in S. cerevisiae (284), but the actual site has yet to be defined. Most remarkable is the conclusion from mutagenesis studies that Pab1p does not require its specificity for poly(A) to perform functions necessary for cell viability (114). Moreover, while Pab1p binding is shared by wheat eIF-iso4G (320), it has not been clear whether it is also evident in human eIF4GI/II or S. pombe eIF4G (181, 383). There is, however, now agreement that in the former case the site was originally overlooked because of uncertainties about the N-terminal sequence of human eIF4G (509a). Continuing work on S. pombe eIF4G should also resolve uncertainty about Pab1p binding for this fission yeast (449a). A further study has now complicated the story somewhat, since a 480-amino-acid human poly(A)-binding protein (PABP) which shows similarity to the central region of eIF4G has been described (102). The results of continuing investigations of these various Pab1p-related proteins and interactions are awaited with interest.

Selection of the translational start site. The prokaryotic ribosome can locate a start codon via direct interactions with sequence elements such as the Shine-Dalgarno region located within a translational initiation region (TIR) that has evolved to guide and modulate the initiation process (177, 178, 362, 363). In contrast, despite the theoretical proposition that “translation-initiation promotion sites” may enhance the expression of certain genes (545), there is no experimental evidence for such rRNA-mRNA interactions mediating start site selection in the eukaryotic cell. Instead, initiation on the vast majority of cellular mRNAs involves a process currently modelled by the “scanning hypothesis” (to be considered in more detail below). Investigations of mammalian in vitro systems have indicated that the process of scanning through structural leader regions to the start codon may be driven by the helicase activity of eIF4A (which is enhanced in association with eIF4B), possibly associated with the small ribosomal subunit (509). eIF4A belongs to the DEAD-box family of proteins, which possess ATPase and ATP-dependent RNA helicase activities (161, 487). It is required for translation in a yeast cell extract (57). However, it is not included in the eIF4F complex isolated from S. cerevisiae, and it is therefore questionable whether it cycles through the yeast eIF4F complex, as has been suggested for the mammalian protein (420). On the other hand, it seems unlikely that the function of eIF4A is linked solely to its potential role in destabilizing secondary structure in the 5’UTR, since in the experiments of Blum et al. (57) it was required for the translation of an mRNA that had a relatively short, unstructured leader. Overall, the role of the RNA helicase activity of eIF4A and eIF4B in scanning remains poorly defined.

Scanning continues until the preinitiation complex has selected an AUG codon. In the apparent absence of an equivalent to the prokaryotic tRNA–Shine-Dalgarno (SD) region interaction, AUG selection by the eukaryotic ribosome is directed by the anticodon-codon specificity of Met-tRNA(M), (93). It is known that the selection process involves participation of eIF2, since mutations in this factor can alter the specificity of the codon selection process (135, 236). Moreover, very recent work by the Donahue group suggests an unexpected role for eIF5 in determining the stringency of AUG selection (236). It seems that eIF5 acts to control the fidelity of the AUG selection process (there is a functional analogy here to the role of prokaryotic IF3). A mutant form of eIF5 was found to allow recognition of UUG as a start codon in vivo and to enhance GTP hydrolysis on the 43S preinitiation complex in vitro (236). It was proposed that an abnormal eIF5 can promote GTP hydrolysis, thus triggering release of eIF2-GDP (plus other initiation factors [Fig. 4]) at a non-AUG codon. This is then thought to allow the ribosome to recognize the non-AUG codon as a start site, effectively switching it into the polypeptide initiation mode. The β subunit of eIF2 carries a binding site for eIF5, suggesting that the above functions involve direct interactions between the two factors (112a).

Initiation of polypeptide synthesis. Once the 40S subunit has located a start codon, the 60S subunit joins it to form the 80S initiation complex, which can then begin with peptide bond formation between the initial methionine and the second encoded amino acid. This ribosome-joining step is promoted by eIF5, which, as we have seen, acts to ensure the fidelity of the Met-tRNA-M-start codon interaction (85, 236, 369). A number of other factors are likely to be released at this point. Of particular interest is the dissociation of eIF2, which is now complexed with GDP after hydrolysis of the GTP that was originally bound and which will have to be recycled back to the GTP form by eIF2B in preparation for a further round of initiation. The mechanism of this GDP-GTP exchange reaction is still the subject of lively discussion (552).

Additional factors involved in translation. There is quite a collection of proteins that, like Pab1p, seem to be involved in, or influence, the initiation process but are not formally classified as eIFs. This is something of a grey area in terms of formal classification, but it may also be indicative of important but as yet uncharacterized functional interactions between the translational apparatus and other cellular components. For example, Donahue and colleagues (190, 609) isolated four unlinked genes (SSL1 to SSL4) which, when mutated, can suppress the inhibitory effect of a stem-loop structure in the 5’UTR of HIS4. SSL1 and SSL2 were more recently determined to encode components of the transcription factor TFIHH (578). SSL2 was also found to be a yeast homologue of the human ERCC-3 gene, thought to be involved in DNA repair (190). It is still unclear to what extent the observed effects of the SSL mutants on translation reflect the normal roles of these genes in wild-type cells. SIS1, which encodes a yeast homologue of Escherichia coli DnaJ, has also been linked to translation (617). The temperature-sensitive phenotype of a SIS1 mutant was suppressed by one of the ribosomal gene deletion mutants that was previously shown to suppress a pab1 mutant (474). Again, while links between translation and other cellular processes are suggested, the mechanism underlying this effect remains unknown. It is interesting that the theme of chaperones apparently influencing translation is also evident in the phenotypes of mutations in two genes encoding 70-kDa heat shock proteins (hsp70s), i.e., SSB1 and SSB2. In this case however, the translation defect was suppressed by overexpression of the HBS1 gene, which encodes a protein resembling eIF1A and eRF3 (401). The Ssb proteins may be core components of the
translating ribosome, perhaps preventing misfolding of nascent polypeptide chains (436a). Finally, one additional factor has been found that seems to be required for translation in wild-type yeast cells (91, 119). This is Dedlp, which is a DEAD-box protein originally identified as a potential suppressor of defects in pre-mRNA splicing (269) and PolIII (549). Dedlp is required for translational initiation in vitro (91), and ded1 mutants are defective in translational initiation (91, 119).

The above examples of gene products that can apparently influence translation might be telling us that the cellular translational apparatus is not adequately defined in terms of the formally classified translation factors alone. They can also be interpreted in terms of the view that functional overlaps are an essential feature of molecular evolution, so that it would be inadvisable to apply excessively rigid definitions of what might constitute “true” translational components. In this context, it is worthwhile to consider how a family of proteins, such as that sharing the DEAD-box motifs, can be involved in a range of processes associated with RNA (161, 487). Indeed, the protein encoded by FAL1 in S. cerevisiae has 55% amino acid sequence identity and 73% similarity to eIF4A but is active in pre-rRNA processing rather than translation (303). These observations raise the question whether eIF4A itself may be inadequately defined as “merely” a translation factor and whether, in a wider context, an unknown number of the other proteins currently classified as translation factors are fully “dedicated” components of the translational apparatus. Further research should reveal how the DEAD-box-type motifs can be combined with a variety of other protein domain structures to confer different functional roles on the members of this family.

Reaction pathways and kinetic control. While it is convenient to break down the initiation process into distinct steps, there is in fact little information on the spatial and temporal relationships between the respective partial reactions within the living cell. For example, the translational apparatus can hardly be envisaged as a farrago of randomly acting components, but is there a highly ordered multifactor supercomplex (a “translatosome”), or does the reality lie somewhere between these two possible extremes? It is impossible to resolve issues such as this on the basis of in vitro experiments alone, since disruption of the complex and delicate pathways in the cell may leave only partially or completely uncoupled component reactions. Effectively torn out of its natural cellular environment, translation in vitro is likely to reflect, but unlikely to reproduce, the bona fide process in vivo. This problem may be particularly applicable to the initiation step, since this encompasses not only the transition from the nuclear phase of the life of an mRNA to its recruitment by the translational machinery but also the orchestration of the most complex phase in protein synthesis. Also linked to this problem is the fact that initiation is not an autonomous process occurring independently of the other phases of translation. Statistically, it would be expected that most ribosomes initiating protein synthesis on an mRNA had recycled after previously terminating at least one other polypeptide chain. In general, this means that the pool of ribosomes available for initiation is subject to control by the termination process and posttermination events. Moreover, there are theoretically two extreme cases where this control might be exercised within the confines of a single polysome: as a consequence of reinitiation on a multicistronic mRNA (see below), and in the context of a “closed-loop” (265) type of polysome structure which might allow a ribosome terminating at the 3′ end of the mRNA to be recycled back to the 5′ end (see, for example, reference 233) or to influence de novo initiation at the 5′ end. Reinitiation is known to occur (see below), while the latter type of mechanism is feasible but has not been confirmed as a potential pathway. These considerations emphasize the importance of the cyclical nature of the operations performed by the translational machinery and thus of the relationship between initiation and the other phases of translation.

The kinetic (and thus the temporal) control of the various steps of a process as complex as translational initiation is an even more difficult problem. The apparent order of the current schemes for a major pathway is seductive but may also be misleading. There are at least three potentially serious issues. The first is that in vitro analyses of partial reactions may use conditions that distort the behavior of the translation components under study. For example, the use of an RNA-binding translation factor at excessively high ratios over the mRNA template may lead to the attribution of significant reaction rates to a process that is kinetically insignificant in vivo. Alternatively, an excess of mRNA template in an in vitro cell-free system may titrate out RNA-binding proteins that normally influence the selectivity of the translation apparatus (see, for example, reference 534). Second, reactions that may be largely temporally compartmentalized in vivo may be literally thrown together in an in vitro system, thus generating an apparently viable process that nevertheless does not accurately reflect the course of events in vivo. A hypothetical example, used here to illustrate this point, would be the role of eIF4E. Although this factor can be shown to be required for translation in a cell-free system, its major role in vivo might be restricted to the earliest stages of interaction between ribosomes and mRNA. It is not yet certain that initiation cycles on polyribosomes follow the same type of eIF4E-dependent pathway as applies to the earliest initiation events on mRNA molecules that have just left the nucleus (see also the next section). In other words, is every initiation cycle on a given mRNA mechanically equivalent? Finally, the route followed by the translational machinery under any given conditions in the cell may be dictated by kinetic control rather than the absence of mechanistic alternatives. Thus, for example, 40S binding to the mRNA might theoretically be primarily cap mediated in vivo because of the relatively rapid kinetics of at least the initial recruitment of capped mRNA into polysomes. However, uncapped mRNA may be an acceptable alternative that is normally discriminated against on competitive kinetic grounds. These and other complications with in vitro experimental work generate uncertainties with respect to the interpretation of the resulting data.

Translational Elongation and Termination

Elongation factors. The process of elongation in eukaryotic translation has generally received comparatively little attention and is assumed to function in an analogous fashion to that of its counterpart in E. coli (459). While this assumption is certainly likely to apply to the basic biochemical principles, the eukaryotic systems have their own, more complex, set of elongation factors. A highly abundant homologue of the bacterial factor EF1A (formerly called EF-Tu) is present (eEF1A). This forms a ternary complex with GTP and aminoacyl-tRNA and promotes binding of the latter to the ribosomal A site. The other eukaryotic factors (eEF1B and eEF2) do not show readily identifiable sequence homology to their prokaryotic counterparts (EF1B and EF2, formerly known as EF-Ts and EF-G, respectively). The heterotrimeric factor eEF1B catalyzes GDP-GTP exchange on eEF1A yet is dissimilar to the prokaryotic factor EF1B (EF-Ts), which performs an analogous function for EF1A. However, in yeast, eEF1B is rendered redundant if eIF1A is overexpressed (286). The other G-protein, eEF2, is thought to be required for translocation of the
peptidyl-tRNA to the P site and, by analogy to the prokaryotic system, of decacylated tRNA to the E site. eEF2 is potentially a major site of regulation mediated by phosphorylation in higher cells (395, 418). An intriguing property of the eukaryotic factors eEF1A and eEF2 is their ability to bind to cytoskeletal components (36, 500), since this may provide a mechanism for the intracellular transport of mRNA, perhaps within polysomes. eIF1A and eEF2 are encoded by duplicated genes. In both cases, the encoded proteins are identical whereas there are only minimal differences in the respective reading frames (99, 394, 433).

A remarkable feature of yeasts and fungi is that they have an additional elongation factor, eEF3 (506), that is apparently absent in mammalian systems, although it does have an apparent homologue in the Chlorella virus CIVU2 (604). eEF3 has a number of structural features, with domains similar to ATP-binding/catalytic domains of the ATP-binding cassette (ABC) superfamily of proteins, the E. coli 55 ribosomal protein, and regions of predicted interaction with tRNA, tRNA, and mRNA (40). There is considerable uncertainty about the function of eEF3, with suggestions that it is involved in stimulating either the binding of cognate aminoacyl-tRNA to the A site (381) or the release of uncharged tRNA from the ribosomal E site (402). It remains to be seen whether the function of this factor is fulfilled by an integral component of the mammalian ribosome. For obvious reasons, eEF3 constitutes a favored target for the development of specific antifungal agents by pharmaceutical companies (201, 555).

**Release factors.** There have been considerable advances in the study of eukaryotic translational termination in recent years. It is now thought that in-frame stop codons are recognized following the binding of the heterodimeric release factor (RF) complex, comprising eRF1 and eRF3, to the ribosome (517, 618). eRF1 is involved in recognition of all three stop codons (UAA, UAG, or UGA) and mediates peptidyl release (159). eRF3, a GTPase showing homology to eRF1, stimulates this reaction in a GTP-dependent, codon-independent way. The GTPase activity of eRF3 is triggered by the formation of a ternary complex with eRF1 and the ribosome (160). The current model of termination therefore envisages that eRF3, with its three GTP-binding consensus elements, binds the ribosome in a similar way to eRF1A (517). During elongation, eRF1A would bind aminoacyl-tRNA and GTP in preparation for the formation of the next peptide bond. eRF3, in contrast, can bind only eRF1 and GTP, and through this interaction on the ribosome it promotes peptidyl release. In S. cerevisiae, eRF1 and eRF3 are essential, low-abundance proteins present at a molar ratio of less than 1:20 with respect to the ribosome. eRF1 and eRF3 are essential, low-abundance proteins present at a molar ratio of less than 1:20 with respect to the ribosome.

**Sequence contexts and termination efficiency.** Current evidence points to at least two factors influencing the rate of translational termination: the context of the termination codon, and the structure of the C terminus of the peptide chain. In **S. cerevisiae**, the overall use of the three possible stop codons is biased: UAA (53.1%) > UGA (26.8%) > UAG (20.2%). The bias toward UAA is even stronger (87.2%) if the analysis is restricted to highly expressed genes. However, there is evidence that the triplet sequence alone is not solely responsible for determining termination efficiency and stop codon usage. There is clearly discernible nonrandomness in the bases observed at positions close to the stop codon in E. coli, S. cerevisiae, and mammalian cells (64–67, 442, 543). In particular, there are biases in the identity of the base immediately downstream of the stop codon. Analysis of 784 S. cerevisiae genes (66) revealed overall preferences in terms of the frequency of use of the following tetranucleotide sequences: UAAA (18.2%), UAAG (13.5%), UAAU (16.6%), UGAA (9.9%), UGUA (9.6%), and UAGA (8.4%). Moreover, the bias toward a number of these was greatly increased if only highly expressed genes were included in the analysis: UAAA (33.3%), UAAG (35.9%), and UAAU (16.7%). This type of data has been interpreted to mean that the fourth-base context influences the recognition of the triplet stop codons, perhaps via modulation of the interactions of the RFs in various organisms (543). Bonetti et al. (59) tested the influence of stop codon (fourth-base) context in S. cerevisiae by examining its ability to modulate the suppression of stop codons inserted early in the lacZ gene. They found the following apparent relative termination efficiencies for the respective stop codons: G > U > A > C (UAG), G > A > U > C (UAA), and A > U > C > G (UAG). Comparison with the previous data sets reveals a correlation with the frequency of use of tetranucleotide signals in more highly expressed genes, whereas there is evidently a wider spread of tetranucleotides used in the yeast genes as a whole. A relatively small effect on termination efficiency was also observed when the third base downstream of the stop codon was altered (59). Overall, these data suggest that a significant level of posttranscriptional control is exercised at this step of translation. At least in E. coli, the influence of the stop codon context is thought to be related to evolutionary and adaptive principles affecting the use of the genetic code (543). The role of the termination codon context in yeast has yet to be intensively studied, but there are a number of ways in which it could be significant. One aspect worthy of attention is the influence of context effects on ribosomal protein. Moreover, by virtue of its ability to bind eRF3, eRF1 apparently coaggregates with eRF3 in [PSI+] cells (422). Interestingly, there is evidence that the [PSI+] activity can be modulated by the chaperone functions of heat shock proteins (88). There is at least one other yeast prion (URE3, encoded by URE2) (592).
complex, but there are apparent correlations with properties such as the propensity to form α-helices and β-strands. As with *S. cerevisiae*, there are indications that the C-terminal codon choices also influence termination (59), and more recent investigations have revealed that the relationship between C-terminal peptide structure and termination efficiency in *S. cerevisiae* differs from that observed in *E. coli* (385).

**Analogous functions in prokaryotes and eukaryotes.** Comparisons between the respective prokaryotic and eukaryotic translation termination factors are informative (71, 398). eRF1 (Sup45p in yeast) seems to perform the same function as the *E. coli* factors RF1 and RF2. There is no extensive homology between eRF1 and RF1/2, but certain conserved sequence elements have been identified (249, 398), suggesting that there is common ancestry and/or convergence. This has provided one argument supporting the idea that RF1/RF2 and eRF1 may bind to the ribosomal A site by mimicking at least the tRNA component of the tRNA-elongation factor complexes involved in elongation (398). On the other hand, *S. cerevisiae*, *Xenopus laevis*, and *Homo sapiens* each have an eRF3 homologue (in yeast encoded by SUP35) with a C-terminal domain that contains a GTP-binding motif and shows significant similarity to apparently equivalent domains in the prokaryotic factors RF3, EF-Tu, EF-G, and the eukaryotic factor eEF1A. It is important to note that RF3 (which seems to be the prokaryotic counterpart of eRF3 [Sup35p]) resembles EF-Tu and the N-terminal domain of EF-G. It has been suggested that either the complex RF1/2/RF3, or possibly RF3 alone, can mimic the complex between EF-Tu and tRNA in a manner analogous to that proposed for EF-G, thus binding to the ribosomal A site (249, 397, 398, 408). However, unlike Sup35p, RF3 is not essential for growth and is therefore thought to act to promote the action of the other RFs. Overall, there is reason to believe that despite the original expectations of functional homology raised by the identification of regions of sequence similarity, Sup35p and RF3 do not fulfill (fully) equivalent functions.

One clue to the difference between the prokaryotic and eukaryotic termination processes may lie in the observed prokaryotic requirement for a fourth factor, the essential ribosome-recycling factor (RRF, or RF4) (270). RRF may promote the translocation event that moves deacylated tRNA out of the P site of the prokaryotic ribosome (into the E site) and/or the removal of whichever RF is in the A site, thus promoting ribosome release and recycling (398), although alternative pathways can be envisaged (71). The latest in vitro experiments have revealed that RRF, together with EF2 (EF-G), promote posttermination mobility of the ribosome, which can then lead to either release from the mRNA or reinitiation (424, 425). Ouzounis and colleagues have described a yeast gene encoding an RRF-like factor, but this protein is probably mitochondrial (417, 518). Perhaps, therefore, eRF1 and eRF3 are sufficient to bring about all of the prokaryotic RF functions (stop codon recognition in the A site, peptidyl-tRNA hydrolysis, termination complex disassembly, and ribosome release) without the involvement of further factors. Further parallel investigations of the eukaryotic and prokaryotic factors will shed light on this and other questions concerning the respective termination processes. This will also be significant to our understanding of the relationship between termination and posttermination events on the mRNA, including reinitiation and mRNA degradation processes. One area of interest will be the nature of the interactions between the various termination factors and the ribosome. Mutational analysis of the *E. coli* ribosome has revealed that tRNAs from both subunits are involved in peptidyl-tRNA hydrolysis during termination (22). Analysis of the role of a nucleotide in *S. cerevisiae* 18S rRNA that is equivalent to one of the sites mutated in the *E. coli* study has indicated that the yeast tRNAs also participate in an analogous fashion in the termination process.

**Termination need not be followed by ribosome release.** Given that translation is performed primarily by recycling ribosomes, the question arises how the transition from termination of one polypeptide chain to initiation of another is achieved. As will become evident later, this turns out to be an issue of fundamental importance to certain types of posttranscriptional control, but here it is appropriate to focus on the roles of initiation factors in preparing the ribosome for the different operations it has to perform. Work on GCN4 expression in *S. cerevisiae* has indicated that termination does not necessarily lead to rapid release of ribosomal subunits from the mRNA (see reference 227 for a review). The fate of the ribosome after termination is influenced by the sequence context of the stop codon. One option is for at least the 40S subunit to remain on the mRNA, where it seems to be capable of resuming scanning. That this is possible demonstrates that Met-tRNA binding can occur on a 40S subunit that is already associated with the mRNA. It also shows that the eukaryotic posttermination ribosome can remain much more firmly associated with the mRNA than can its prokaryotic counterpart. These observations leave a number of questions open, including whether the 60S subunit is necessarily released by a posttermination 40S subunit that is associated with the mRNA (and, if so, how rapidly) and whether, at the 3′ end of an mRNA, posttermination scanning of the ribosomal subunit(s) continues until the poly(A) tail is reached before being released. Unfortunately, the sequence of events associated with termination is still poorly understood, and much remains to be learned about the cycling of ribosomes. For example, eIF1A, eIF3, and eIF6 are all thought to promote separation of the ribosomal subunits, yet it is not known at which point they bind in the posttermination phase. Moreover, the kinetics of binding of such factors is expected to influence the potential of posttermination ribosomes to participate in (re)initiation events both before and after release from the mRNA.

**Mitochondrial Translation.** Yeast is dependent on mitochondrial translation only under conditions in which oxidative phosphorylation is required for the metabolism of respiratory substrates. There are only eight major mitochondrial mRNAs, and more than 100 nucleus-encoded proteins have to be transported into this organelle (see, for example, references 32, 208, 492, and 530 for reviews of mitochondrial protein transport; see also reference 221 for references relevant to other types of protein transport) to allow mitochondrial translation to take place. Mitochondrial translation in yeast has been reviewed recently by Fox (157), and only a few points are summarized here. Translation in the *S. cerevisiae* mitochondrion shows some striking differences from cytoplasmic translation. The mRNAs are generally uncapped, and most of them have 5′UTRs longer than 300 nucleotides. It is thought that initiation involves internal ribosome binding, but there is no clear evidence for the existence of SD-like or alternative motifs in the mRNAs that might promote the prokaryotic (*E. coli*) type of initiation pathway, although A residues are conserved at eight positions within the region from −25 to +18 flanking the start codon (155, 380). Many of the imported translation factors are homologous to prokaryotic factors, but at the same time, the mitochondrial ribosomes contain subunits that have no identifiable homology to any prokaryotic counterparts. Overall, it seems that mitochondrial translation is more prokaryotic than eukaryotic in type, but the
The mitochondrial translation apparatus also interprets certain parts of the genetic code differently from its cytoplasmic counterpart, as discussed below in the section on alternative readings of open reading frames. Of further note is that at least five, and possibly all, of the *S. cerevisiae* mitochondrial mRNAs require the presence of activator proteins for translation (157, 350, 426, 528, 557). Moreover, at least one nucleus-encoded protein (Cbp1p) stabilizes a mitochondrial mRNA (*COB*) (87, 128). Remarkably, the translational activators seem to be specific for each mRNA. The current working model is that the activator proteins tether each target mRNA to the inner mitochondrial membrane, perhaps thereby facilitating cotranslational insertion of the encoded proteins (157). The fact that translational activation is generally a comparatively rare phenomenon (363, 364), makes this unusual arrangement all the more puzzling. The explanation presumably lies somewhere in the selective forces and mechanisms that have maintained the shared coding potential of mitochondria and nuclei.

**PATHWAYS OF TRANSLATIONAL INITIATION**

Eukaryotic ribosomes enter into the translation process by at least three different routes, which can usefully be compared to the pathways of initiation in *E. coli* (Fig. 9). Where eukaryotic and prokaryotic initiation were compared in the past, it was usual to stress the immediately obvious differences between the two systems. However, there are also some striking similarities, and a balanced analysis highlights important elements.
of kinetic and thermodynamic control that play decisive roles in mRNA-ribosome interactions.

In the eukaryotic cell, the 5' and 3' mRNA modifications (Fig. 2 to 4) play important roles in translation. Recent work has made it important to examine the roles of protein interactions with these terminal structures in some detail before considering the alternative pathways of initiation. It has been known for some time that capping and polyadenylation are positive modulators of higher eukaryotic translation (264).

Analogous effects were observed on the translation of in vitro-synthesized mRNA electroporated into spheroplasts from S. cerevisiae (165) or added to S. cerevisiae translation cell extracts (172).

The poly(A)-binding protein was originally found to promote poly(A)-dependent translation in rabbit reticulocyte lysates (187). Later work then showed that polyadenylation stimulates translation even in a yeast cell extract lacking active wild-type eIF4E and that cap-dependent stimulation also functions in the absence of active Pab1p (538). Since both eIF4E and Pab1p can bind eIF4G (539, 541) (Fig. 5), it is tempting to conclude that eIF4G is the key player in both stimulatory effects. However, in the absence of further information on the physical and kinetic causality of ribosome-mRNA association, it is not yet possible to pinpoint the mechanisms involved, and this remains a challenging area of investigation.

An aspect of eIF4E-eIF4G interactions not yet understood is how these relate to the efficiency of selection of capped as opposed to uncapped mRNAs. Very recent studies with mutants of eIF4G compromised in their ability to bind eIF4E have revealed reduced selectivity for the translation of capped as opposed to uncapped mRNAs in vitro (540). In the same series of experiments, the Sachs group also found that the product of the mutant eIF4E gene cdc33-1 (defined in reference 11), which shows reduced binding to both the cap and to eIF4G, also supported lower cap selectivity. The experiments with the cdc33-1 mutant are consistent with the effects of deletion mutations in eIF4E on the selectivity of the translational apparatus toward capped and uncapped mRNAs electroporated into yeast spheroplasts (569): the cap-binding affinity of eIF4E was found to play a central role in controlling this selectivity. However, the effects of mutations in the eIF4G-binding site for eIF4E on in vitro selectivity have been interpreted to mean that eIF4E acts as a negative regulator of the capacity of eIF4G to promote the translation of uncapped mRNAs (539).

An alternative possibility is that the changes in eIF4E-eIF4G interaction skew the competitive selection of capped and uncapped mRNAs. The latter effect would be expected of a poly(A)-dependent form of eIF4G-eIF4E binding. This contrasts with the lethalities of point mutations in eIF4G that eliminate eIF4G-eIF4E binding. At 21°C in an appropriate environment, the poly(A) tail (and thus the Pab1p bound to it) may play an important role in the recycling of posttermination (40S) ribosomes through to their renewed initiation on either the same or another mRNA molecule. Another question regarding poly(A)-mediated initiation is whether it could define a cap-independent pathway in vivo that places the 40S subunit downstream of the cap structure, thus effectively mediating internal initiation. Perhaps this type of pathway contributes to the poly(A) stimulation seen in yeast cell-free systems (172, 538). The question remains, however, whether this route is kinetically favored in vivo. Interestingly, at least one plant virus (PAV barley yellow dwarf virus) uses an element in the 3'UTR to promote cap-independent translation (577).

There is a need for further characterization of the role of Pab1p in living cells. Recent studies of S. cerevisiae strains bearing a mutant pap1 gene that encodes a temperature-sensitive form of poly(A) polymerase have been informative. Using this mutation in different host backgrounds, Proweller and Butler (445, 446) have been able to explore the relationship between poly(A) length and recruitment of mRNAs into polyribosomes. At the nonpermissive temperature, pap1-1 cells contained approximately half as much total mRNA and their translational apparatus showed apparently only minimal discrimination between poly(A)+ and poly(A)- mRNAs. However, discrimination was increased in a strain in which the abundance of ribosomal subunits was reduced, creating a nearly wild-type ratio of mRNA to ribosomes. In this strain, poly(A)- mRNAs were generally associated with smaller polysomes. These results emphasize the significance of the relative activities of the components of the translational machinery in terms of rate control. They also exemplify the problem of at least quantitative disparity between the results obtained with in vivo and in vitro systems (538). This problem also arises in the interpretation of a further report of poly(A) tail-mediated stimulation of translation in vitro (442a). This study essentially confirmed earlier demonstrations of poly(A) stimulation of translation in cell extracts from S. cerevisiae (172, 538) while also finding that the translation of a noncapped, polyadenylated mRNA was resistant to inhibition by antisense 21-O-allyl-oligoribonucleotides targeted to sites in the 5'UTR more than a minimal distance upstream from the reporter gene start codon. This result was taken to mean that the poly(A) tail supports cap-independent, internal initiation. However, as indicated in the earlier discussion of the distorted conditions for translation created by the preparation of cell extracts, it is difficult to interpret such data in terms of the in vivo functional role of the poly(A) tail. Perhaps the PolII transcription system described by the Donahue laboratory (338a) could provide more direct information about the in vivo role of the poly(A) tail (on uncapped mRNAs).

The issue of in vivo versus in vitro behavior is also raised by the respective studies of the role of eIF4G-Pab1p interactions in vitro and in vivo (538, 541). While cap-independent, poly(A)-dependent translation is clearly evident in a yeast cell extract (172, 538) and eIF4G and Pab1p are capable of interacting with each other (539), more recent work has shown that the Pab1p-binding site of eIF4G is not essential for yeast viability (541). Moreover, a point mutation in eIF4G1 (Tif4631-213p) which eliminated Pab1p binding reduced the growth rate by only 20% at 30°C and still allowed the synergistic stimulation of mRNA translation by the cap and a poly(A) tail in vitro. This contrasts with the lethality of point mutations in eIF4G (540) or eIF4E (449) that eliminate eIF4G-eIF4E binding. At the same time, mutational analysis of the eIF4E- and Pab1p-binding sites in eIF4G and the demonstration of synthetically lethal interactions between the respective genes suggest that the in vivo functional roles of eIF4E and Pab1p might at least partially overlap (541). In this respect, it is relevant that in vitro experiments with wheat germ extracts indicate the occurrence of mutual cooperativity in this plant system between the binding of PABP to eIF4F (eIF-iso4F) and the binding of eIF4F...
nucleotides are therefore omitted here.


titates

a parallel here to the observation that the contribution of an (eIF-iso4F) to a cap analogue (580a). However, the significance of downstream nucleotide contexts is likely to be complex, given that these will affect the N-terminal sequence of the encoded protein. These nucleotides are therefore omitted here.

one pathway. Moreover, it should be emphasized that the given mRNA may be translated (or translatable) via more than one pathway. Intermediate between the three major classes, and a mRNA (86).

This value depends on the G content of the stem-loop structure (570).

<table>
<thead>
<tr>
<th>Organism</th>
<th>AUG context and initiation site signals</th>
<th>AUG priority</th>
<th>Sensitivity to mRNA secondary structure (90% inhibition at ca.)</th>
<th>Position effect guiding inhibition via structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>. . SD . . AUG . . AU rich or other elements</td>
<td>AUG selection efficiency determined by local TIR sequence and structure; coupling between AUGs in cis can be very influential</td>
<td>8 kcal mol$^{-1}$ in TIR</td>
<td>Structure restricting access to SD/AUG</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>Preferred: AA/UAUAG . .  (context effects)</td>
<td>Scanning dictates a 5' proximity selection gradient for multiple AUGs which is subject to modulation by context effects</td>
<td>15 kcal mol$^{-1}$ 5' of AUG</td>
<td>Degree of inhibition largely independent of position of secondary structure in 5'UTR</td>
</tr>
<tr>
<td>Vertebrates</td>
<td>Preferred: C . . A/GCCAUG . .</td>
<td>Same as in yeast, but different context effects</td>
<td>50 kcal mol$^{-1}$ 5' of AUG</td>
<td>Secondary structure more inhibitory in 5'end-proximal position</td>
</tr>
</tbody>
</table>

SD-to-AUG distances are generally 5 to 13 nucleotides. Initiation at GUG (relative frequency, ~ 8%), UUG (relative frequency, ~ 1%), or AUU (one case) possible. Other sequence elements may play a role in initiation, especially in the absence of a strong SD region (see summaries in references 362 and 363). The significance of downstream nucleotide contexts is likely to be complex, given that these will affect the N-terminal sequence of the encoded protein. These nucleotides are therefore omitted here.

The scanning model, and the evidence from mammalian systems consistent with it, have been extensively discussed by Kozak (see references 296, 300, and 301 for recent reviews of the relevant data). Since in vitro translational initiation is dependent on ATP hydrolysis (258, 292), it is assumed that the scanning process itself is either directly or indirectly ATP driven. The same basic principles of
the scanning model all seem to apply to the translation of cellular mRNAs in *S. cerevisiae*, but there is some divergence in terms of detailed sequence and structure requirements for initiation and the magnitude of their respective influences on translational efficiency. These are summarized in a comparative tabulation of related data from *S. cerevisiae*, vertebrate cells, and *E. coli* (Table 2) and are also considered below.

(i) **AUG context.** Essentially, initiation in *S. cerevisiae* is restricted to the AUG codon and is reduced in efficiency if the sequence surrounding it deviates significantly from certain preferred nucleotides (134). However, the effects of changes in the context sequence are considerably smaller in *S. cerevisiae* than in vertebrate systems (299). Even at the most influential position, −3 with respect to the AUG, the best choice (A) is maximally twofold better than other nucleotides (31, 93). Overall, analyses of large numbers of *S. cerevisiae* genes indicate a general preference for A upstream of the start codon (82, 92, 608) whereas the typical vertebrate upstream context is considerably more CG rich (82, 295). Analysis of the downstream context yields no apparent preference for A at positions +4, +5 and +6 but, rather, shows a predominance of pyrimidines (92, 228, 608). This may be linked to selective forces affecting the N-terminal sequence of the encoded protein (92, 93) (also compare the upstream biases for termination codons discussed in the earlier section on termination). It seems likely that the bias toward A in the *S. cerevisiae* upstream start codon context, which is especially pronounced in highly expressed mRNAs, is at least partly a reflection of the relatively high sensitivity of the yeast translational apparatus towards secondary structure.

(ii) **AUG priority rule.** A further factor influencing the selection of an AUG as a site of initiation is its position within the 5′UTR. In general, the first AUG downstream of the 5′ end acts as the major initiation site (134, 136, 293, 299, 338, 440, 499, 608). This can be explained most readily by an overall 5′→3′ movement of ribosomes during scanning, which effectively imposes a “priority rule” on start site selection. As explained in the section on initiation components, the selectivity of the translational apparatus is tightly controlled by structural characteristics of the initiator-tRNA and of eIF2. There is a sizeable group of natural mRNAs whose leaders contain upstream AUGs (uAUGs) or short uORFs (92, 299, 572). As expected on the basis of the scanning model, these upstream elements generally inhibit translation, albeit to greatly differing degrees. Moreover, analysis of the GCN4 leader, which has four uAUGs (see below), has provided some of the most convincing support for the scanning model of initiation. Equally consistent with the scanning model is the observation that the introduction of a uAUG into a natural yeast leader sequence, or into a completely artificial yeast leader sequence, also leads to inhibition of translation of the main ORF (92, 411, 498). This latter type of result supports the contention that there are unlikely to be any special yeast context sequences required for the priority rule to be observed. Additionally, this type of experiment, as well as other work with leader deletions (31, 565), indicates that there is no eukaryotic equivalent to a prokaryotic SD region. However, if the first AUG is located less than 15 to 20 nucleotides from the 5′ end, depending on the organism, recognition by the ribosome will be impaired, thus attenuating the initiation rate (299, 565). One explanation offered for this is that close proximity of the AUG to the 5′ end may lead to steric hindrance of 40S subunits entering the scanning pathway by 80S complexes paused at the start site (565). Alternatively, start site recognition may be compromised because the shortness of the leader prevents optimal interactions between the 40S subunit and the mRNA to be coincident with appropriate AUG positioning. Since the former explanation assumes that 80S complexes exert rate limitation at the start site, “queuing” of 40S subunits on leaders longer than the apparent minimum length might also be expected to result in a similar inhibitory effect. Although a clear-cut distinction may not be possible here, these considerations would argue for the model in which optimal start site recognition is dependent on a longer region of interaction 5′ of the AUG. At the other extreme, there is no indication that extending yeast 5′ UTR lengths beyond the average length of approximately 50 nucleotides has any detrimental effects on translation.

The factors of AUG context, AUG priority, and leader length also come into play when the 5′ UTR has more than one potential initiation site. For example, a combination of these influences allows MOD5 mRNAs whose 5′ ends are at −10 or −11 with respect to the first AUG to support translation at both the first AUG and at a second AUG at position +34 (396, 507). Indeed, all three factors are thought to be balanced in such a way as to allow some ribosomes to generate an N-terminally extended form of the tRNA-modifying enzyme Δ²-isopentenyl pyrophosphate:tRNA isopentenyl transferase, which is targeted to the mitochondria. However, other ribosomes scan through the first AUG and initiate a shorter form of the enzyme that is localized to the cytoplasm and nucleus. Although the details of the mechanisms determining this distribution of initiation events have yet to be worked out, it is clear that leader design can be used to ensure the generation of two different proteins from one and the same mRNA template. A number of examples of the use of alternative initiation codons are known in mammalian cells (562).

(iii) **Secondary structure blocks initiation.** The ability of stem-loop structures in the 5′UTR to block the progress of 40S subunits along the mRNA and thereby to inhibit translation constitutes a further indicator of the operation of a scanning mechanism in mammalian initiation (294, 296). Moreover, differential sensitivity to the presence of a stem-loop as a function of its position in the leader was taken to reflect interference by this structure in two distinct stages of the mammalian initiation pathway (297). In a rabbit reticulocyte translation system, a cap-proximal stem-loop with a stability of −30 kcal mol⁻¹ was found to be more inhibitory than an equivalent structure placed 52 nucleotides further downstream. It was argued that the cap-proximal structure interferes with initial 40S binding at the 5′ end of the mRNA whereas a downstream stem-loop is encountered by a scanning ribosome, which is evidently driven by a thermodynamic force that is sufficiently large to unwind the structure (297). On the other hand, a cap-distal stem-loop can inhibit translation by more than 85% if its stability is increased to over −50 kcal mol⁻¹ (297). These effects are not only readily explained in terms of the scanning model but also provide useful clues about the energetics of component steps of the initiation pathway. This becomes particularly evident in a comparison between the respective responses of the mammalian and yeast translational machineries to mRNA structure (Table 2). Not only is translational initiation more sensitive to leader structure in *S. cerevisiae* (31, 52, 93, 412, 570, 608), but also there is no cap-proximal potentiation effect in this organism (412, 570), so that variations in the position of a stem-loop structure in the 5′UTR hardly modulate its inhibitory effect. As discussed below, this difference is also reflected in the mechanisms of action available for translational regulation by RNA-binding proteins in higher and lower eukaryotic cells. Finally, it should be added that a stem-loop structure positioned behind an AUG can enhance the efficiency of start site recognition; the optimum gap was found to be 14 nucleotides (298). A plausible hypothesis has been offered to explain this. It is
proposed that the stem-loop causes scanning ribosomes to pause over the AUG, increasing the probability that the start codon will be recognized (298).

In conclusion, the observations of AUG context effects, the AUG priority rule, and the blocking effects of secondary structure are all consistent with the basic scanning model (299, 608), although they do not tell us much about the mechanism of ribosome movement along the mRNA. While most yeast mRNAs tend to have 5'UTRs that are relatively free of secondary structure (82, 92), approximately 10% are predicted to have sufficient structure to cause significant translational inhibition. Many of the leaders in this group have the potential to form localized hairpin structures with stabilities of up to $-10 \text{ kcal mol}^{-1}$, which can reduce translational efficiency by approximately 50% (Table 2; the level of inhibition depends on the G+C content of the stem-loop [412, 481, 570]). The corollary of the above is that, as in vertebrate cells (299), translational restriction of this kind affects the expression of a sizeable pool of natural mRNAs in the cell. Finally, there is an interesting side to this story of the thermodynamic control of translation. The mRNA-folding energies that have been found capable of restricting initiation on selected eukaryotic mRNAs are large in biomolecular terms. For example, the free energy of folding of an entire monomeric protein may be no more than approximately $-15 \text{ kcal mol}^{-1}$ (although the component enthalpy and entropy terms are much bigger [103]). Viewing this from another angle, proteins involved in the formation of mRNP may be able to destabilize structures below a certain stability threshold, so that further unwinding may not be required for translational initiation to take place. However, the free energies of protein-mRNA binding, including perhaps those of protein conformational changes, will have to be found in the reactions driving scanning that have to displace the bound proteins.

Mechanisms relating translation rates to mRNA structure. Returning to the comparison of prokaryotic and eukaryotic initiation pathways (Fig. 9 and 10; Table 2), it is useful to examine how the respective translation systems react to structured regions of initiation in the mRNA.

(i) Structure in the prokaryotic TIR. Prokaryotic 30S subunits are able to recognize internal TIRs within polycistrionic mRNAs on the basis of 16S rRNA-mRNA interactions, most importantly involving Watson-Crick base pairing with the SD region. However, as in the eukaryotes, the rate at which the ribosomes can effectively access the start codon and initiate polypeptide synthesis varies over at least a 1,000-fold range (125, 363), and a key contributory factor to this variation in rate is the mRNA structure. Even in complex polycistrionic mRNAs, the efficiency of translational initiation is determined primarily by local properties of the mRNA (121, 140, 268, 313, 365). Investigations of the E. coli atpE operon provided insight into the mechanism by which this local control can be achieved (313). Small in vitro-synthesized fragments of the atpE operon containing the TIRs of individual genes were allowed to interact with purified 30S subunits and fMet-tRNAfMet, and the resulting complexes were analyzed by toeprinting (203), sucrose gradients (189), and runoff polypeptide incorporation assays. The results revealed that structure in the atpE TIR fragment, which supported very poor initiation, interfered with 30S binding whereas the efficient 30S TIR supported both efficient initiation and strong 30S binding. These data support the idea that prokaryotic TIRs exercise control at least partly by regulating the affinity of 30S subunit binding. Consistent with this, detailed analysis of the inhibitory influence of various mutant forms of the hairpin in the RNA bacteriophage MS2 coat protein TIR revealed a correlation between the predicted free energy of folding of secondary structure in this TIR and the efficiency of coat protein synthesis (122). The function of a prokaryotic TIR can be approximately modelled on the basis of competition between folding in the TIR and 30S subunit binding to the SD region (123; compare reference 527). At 37°C, translation efficiency in a TIR controlled by structure is reduced by a factor of 10 with every increase in folding stability of 1.4 kcal mol$^{-1}$ (above a threshold of $-5$ to $-6 \text{ kcal mol}^{-1}$ [124, 305]). This form of thermodynamic control of ribosome accessibility to the initiation site is apparently achieved without any requirement for additional factors or sources of free energy.

A whole series of alternative TIR recognition sequences has been proposed since the first description of a potential cellular “enhancer” upstream of the E. coli atpE gene (365). It is possible, but generally not proven, that such sequences engage in additional interactions with the 16S rRNA of the 30S subunit, thereby promoting the formation of the initiation-competent 30S-mRNA complex (152, 362). However, these additional interactions are at least as likely as the SD-anti-SD (ASD) interaction to be modulated by secondary structure in the TIR. This obvious principle, which can easily influence many of the weak interactions that are proposed and can seriously hamper attempts to delineate precisely apparent enhancer elements (486), has often been ignored, leading to misinterpretation of much of the data on putative alternative “enhancer” sequences (363). Apart from reassessing the claims made for various sequence elements, there is a serious need for rigorous analysis of their structural and functional properties.

(ii) Structure in the eukaryotic 5'UTR. An appreciation of the molecular basis of translational inhibition via secondary structure within the 5'UTR is essential to our understanding of the scanning process. The driving force for initiation site location and stable initiation complex formation in E. coli is primarily the free energy of formation of Watson-Crick base pairing between the 30S subunit 16S rRNA and sequences in the mRNA. This explains the relatively great sensitivity of prokaryotic initiation to secondary structure involving the TIR (Table 2), although this relationship is a function of the extent and stability of the mRNA-rRNA interactions that can be achieved (see, for example, reference 123). In the absence of this direct mechanism of AUG localization, start site selection on eukaryotic cellular mRNAs is the result of scanning that depends on a relatively tight association of 40S subunits with the mRNA and apparently on coupling to ATP hydrolysis. A processive, ATP-driven scanning mechanism would be ineffective and energetically wasteful if the off-rate of ribosomal subunits on the 5'UTR were comparable to the on-rate at the 5' end. Unlike prokaryotic ribosomal subunits, therefore, eukaryotic start site selection must be preceded by a mechanism that effectively promotes tight association of the 40S subunits to the mRNA (Fig. 10). This explains why 40S subunits have to negotiate secondary structure within the 5'UTR, irrespective of its position relative to the cap or the start codon. Such a model predicts that the sensitivity of eukaryotic ribosomes to cap-distal structure in the 5'UTR is therefore determined by the specific rate of free energy input channelled into the scanning process. In S. cerevisiae, the translation rate decreases approximately 10-fold for an increase of 15 kcal mol$^{-1}$ in the stability of a discrete stem-loop structure in the 5'UTR (at 30°C) (482, 570). This value is influenced by the G+C content of the folded region. Moreover, at least mammalian 40S subunits react differently to structure that is cap proximal.

There are some apparent extremes of structure in natural yeast 5'UTRs, whose existence at first sight seems inconsistent with quantitative assessments of the range of stem-loop stabil-
ities that can be tolerated in the leader (412, 570). For example, the PMA1 mRNA leader is predicted to be able to form secondary structure with a total stability exceeding $-250 \text{ kcal mol}^{-1}$ (75). According to the relationship between stem-loop stability and translational inhibition established by Vega Laso et al. (570), this might lead to the prediction that translation should be inhibited by more than 99%. However, the expected structure would be of a very extended nature, involving primarily A-U base pairs, and the PMA1 leader was in fact found to support significant translation in vivo (481). It therefore seems likely that a minimum density of C-G base pairs is required for secondary structure to block scanning ribosomes effectively. An extensive AU-rich structure may be relatively easily opened up, analogous to a long zip fastener, possibly because the free energy per unit distance of mRNA required for disruption of hydrogen bonds is comparatively low. This would fit with a model in which the scanning ribosome acts like a molecular machine subject to a driving force whose maximum value is related by an as yet undefined “coupling factor” to the nonequilibrium free energy available from ATP hydrolysis (Fig. 10). Accordingly, internal initiation (see below) or “ribosome shunting” (133, 163, 260, 610) need not necessarily

FIG. 10. Principles of kinetic control affecting prokaryotic and eukaryotic initiation. (A) A prokaryotic 30S ribosomal subunit has direct access to the SD sequence of an mRNA with an unstructured TIR. (B) Inhibition via structure in the TIR can be adequately modelled by assuming a thermodynamic control mechanism in which the steady-state distribution of folded and unfolded TIR dictates the amount of mRNA accessible to ribosome binding. (C) In the eukaryotic case, 40S ribosomal subunit binding can occur unhindered on a leader bearing localized structure, but the structured region is thought then to inhibit the scanning process. The off-rate ($k_3$) for ribosomal subunits in this situation is unknown. Disruption of the secondary structure can be driven by an apparently ATP-dependent process, allowing resumption of scanning through the structured region. (D) Random internal binding of 40S ribosomal subunits to mRNA seems possible but is most probably not favored kinetically, and there may be kinetic or mechanistic restrictions on the ability of the subunits to become tightly associated with the mRNA. Initiation factors will influence the type of interaction entered into (compare Fig. 4). (E and F) 5'-end-dependent (E) or IRES-directed (F) initiation results in a tight association between the 40S subunit and the mRNA (the clasp is closed around the mRNA). It is not known whether scanning is normally unidirectional (E and F).
constitute the explanation for the translatability of the *PMA1* leader. In a 40S-driven unwinding model, the unknown (numerical) coupling factor would be a function of the mechanism of the device coupling ATP hydrolysis to ribosome translocation, including a term reflecting the number of ATP molecules hydrolyzed per unit length of mRNA (which might be variable), and is expected to differ between yeast and mammalian cells. Alternatively, as we shall see, the free energy available for driving scanning might be channelled, at least partly, via translation factor complexes. It follows from the above that reliance on predicted mRNA folding energies alone, especially when these have been calculated over a large stretch of nucleotides, can be misleading.

**Mechanism of scanning.** Having briefly considered the evidence which has justified adoption of the scanning model as the working hypothesis of eukaryotic 5'-end-dependent translation, it comes as something of a surprise to realize that the mechanism underlying the scanning process is still obscure. This is not for the lack of potentially feasible models of the scanning process. One alternative model used to explain how at least mammalian ribosomal subunits negotiate stable secondary structure in the 5'UTR has been referred to as the mRNA helicase hypothesis (509). This proposes that eIF4A and eIF4B pave the way for the scanning ribosome by virtue of their unwinding activity, thus deviating from the order of binding intrinsic to the schemes proposed by most other authors (see, for example, references 300 and 369). It is envisaged that these two factors initially become assembled at the 5' end of the mRNA as part of a complex with eIF4F, which in higher eukaryotes comprises eIF4E, eIF4G, and eIF4A. The ribosome is then thought to bind the 5' region of the mRNA which has been freed from inhibitory structure. Accordingly, this model foresees that mRNA lacking stable secondary structure in its 5'UTR will not require the helicase-catalyzed unwinding activity in order to be bound by ribosomes. In a more extreme version of this model, multiple copies of eIF4A and eIF4B extend from the 5' end along the mRNA and allow internal binding of 40S subunits (534). A testable prediction of this proposal is that in the presence of excess eIF4A and eIF4B, it should be possible to detect mRNA heavily loaded with both of these initiation factors.

As noted already, *S. cerevisiae* eIF4F is structured differently from its mammalian counterpart, raising the question whether it could follow the above type of model or whether eIF4A and eIF4B in yeast might function independently of eIF4F to unwind leader structure. Yeast eIF4B, like its mammalian equivalent, is an RNA-binding protein and promotes 40S and 80S binding to mRNA in vitro when present in excess over added mRNA (14). However, disruption of the *S. cerevisiae* gene encoding eIF4B is not lethal but, rather, generates a slow-growth and temperature-sensitive phenotype (13, 98). There seems to be only a single copy of the eIF4B gene in the *S. cerevisiae* genome. This would argue that unless there is a relatively distantly related protein with homologous function, eIF4B fulfills only a helper or modulator function in yeast translation. Its mammalian counterpart enhances the helicase activity of eIF4A in vitro (370, 467). In vitro investigations of purified recombinant yeast eIF4B, on the other hand, revealed that it catalyzes a slow RNA-annealing reaction at high molar protein-to-RNA ratios (at least 50:1) (14). This annealing reaction was inhibited in the presence of eIF4A, and, indeed, the complex between the two factors catalyzed the unwinding of duplex RNA. On the basis of these data, Altmann et al. (14) speculate that during translation, eIF4B switches between its annealing and duplex-melting modes as a function of (presumably controlled) interactions with eIF4A. Whether this idea reflects the true functions of these two proteins in vivo remains to be determined. Neither the kinetics observed nor the protein-to-RNA ratios required in these experiments to achieve the effects reported are reassuring in this respect. Perhaps other factors are necessary for the proposed model to work, but these have yet to be identified.

In summary, the known characteristics of the yeast translation system are not readily reconciled with the mRNA helicase hypothesis that has been proposed to describe mammalian scanning (509). If proteins such as eIF4A and eIF4B are used to free mRNA of potentially inhibitory structure, how are their activities restrained to that part of the mRNA, i.e., the 5'UTR, where they might be required? This lack of containment, so to speak, would also provide no obvious means of effectively maintaining the unwound state of a leader region up to the time of arrival of the ribosome. As discussed above, eIF4F in *S. cerevisiae* seems not to provide the anchorage that is intrinsic to the specificity required by the mRNA helicase model. Moreover, quite elaborate and unlikely schemes have to be proposed to explain start site selection in complex leaders like that of *GCN4* in terms of this model (see below). At present, it is therefore easier to envisage a ribosome-associated unwinding activity, where this might be necessary, as providing the required specificity and kinetic properties to function effectively, at least in the yeast cell. However, the Sonenberg type of model does focus our attention on the fact that there may be alternative explanations for at least some of the properties of eukaryotic (yeast) mRNAs.

Overall, there is still much to learn about the mechanism(s) of the scanning process. One of the most important questions is whether scanning is exclusively processive and unidirectional or whether eukaryotic 40S subunits can passively shuffle back and forth as do their prokaryotic counterparts. A form of discontinuous scanning or “shunting” has been described in cells of higher eukaryotes (133, 163, 610), but there is no evidence for this phenomenon in yeast. Beyond this, it will be essential to establish whether yeast eIF4A and eIF4B are obligatory components of the scanning ribosomal complex or are “targeted” to the 5' region of the mRNA by some other means. Finally, reliable estimates of the on- and off-rates of eukaryotic ribosomes at different stages of translational initiation are required. Both the latest methods of molecular biology and new biophysical techniques (such as surface plasmon resonance analysis [153] and, once certain technical problems have been overcome, time-resolved fluorescence analysis of movements on ordered mRNA molecules [using, for example, the techniques described in reference 275]) could be of help here.

Does all of the above mean that it is premature to attempt to model translational control? Only partly. A number of principles have been established that can contribute to the interpretation of experimental data. Moreover, the response of the eukaryotic translational apparatus to variables such as structure in different 5'UTRs provides information that can be used to refine the modelling process. At least in *S. cerevisiae*, the absence of a significant position effect on inhibition caused by structure in the 5'UTR indicates that the (quantitative) modelling of ribosome binding and scanning as part of the same process (Fig. 3) is justifiable. It is therefore reasonable to model the quantitative control (but not necessarily the mechanism) of initiation according to the principle of thermodynamic control of access to the start codon, much as has been done for *E. coli*. The essential mechanistic difference is that eukaryotic 40S ribosomal subunits can begin the initiation pathway on a structured leader whereas structure in a prokaryotic TIR displaces the initial binding “equilibrium” away from the complexed form of the mRNA. It should also be noted that
a stem-loop structure, once disrupted, may be held unfolded by the traffic of scanning ribosomal subunits (481) and may therefore not participate in the type of mass-action control envisaged to operate on the prokaryotic TIR.

Reinitiation

Reinitiation is a widespread phenomenon among E. coli mRNAs and plays an important role in the overall control of gene expression in this organism (50, 171, 252, 337, 365, 436, 491, 606). On the other hand, while recognized as playing a key role in the control of a few eukaryotic mRNAs (170, 226), reinitiation has tended not to be regarded as a significant factor in terms of the overall eukaryotic transcriptome. However, as discussed in this review, reinitiation is a manifestation of the posttermination behavior of eukaryotic ribosomes, which not only is of broad significance in terms of the diversity of mRNAs that it affects but also reflects important mechanistic properties of the eukaryotic translation apparatus.

Prokaryotic reinitiation. Reinitiation provides a means of coupling distinct reading frames on a prokaryotic polycistronic mRNA, but it is not the only mechanism available or required for achieving coupled expression. The terminating E. coli ribosome seems to be able to shuffle back and forth in the vicinity of the stop codon of the gene on which it has just terminated (2), and in doing so, it can bind to a fresh initiation site without leaving the mRNA, providing that this site is located within a region that is short in terms of the ribosomal off-rate (on the order of one ribosome-equivalent’s length). The start codon can either lie upstream or downstream of the previous gene’s stop codon or directly overlap with it (AUGA [see, for example, references 331 and 484]). Reinitiation at this site reaches full efficiency if there is an adequate SD region and the stop and start codons of the respective genes are close to each other (certainly less than the equivalent of one ribosome length [510]). The latter characteristic presumably reflects the relatively fast release kinetics of the ribosomal subunits. However, at the same time, the presence of an SD in an internal TIR also opens the door to de novo initiation. The extent to which alternative recognition sequences, such as the so-called downstream box (514), can participate in reinitiation seems not to have been explored. An extremely poor TIR may effectively be “activated” by means of reinitiation (253) but will not support very efficient translation. One striking consequence of prokaryotic coupling is that whole series of genes in polycistronic operons can be more or less tightly linked to the translation of the leading cistron on a polycistronic mRNA (212, 331, 455). It should, however, be noted that translational coupling can also be of a negative kind. Interference by ribosomes on an upstream cistron can be expected to become increasingly likely where termination is close to the start of an efficiently translated gene but there is no coupling pathway to ensure that ribosomes have free access to the downstream TIR (147).

In general, reinitiation itself is not necessarily directly responsible for rendering translation of the downstream gene strongly dependent on the reading frame preceding it (124, 456). This is especially true where the gap between the respective stop and start codons allows ribosomes from the cellular pool to gain access to the downstream TIR. Coupling can be imposed by the presence of structure in the TIR of the downstream ORF that inhibits initiation unless it is disrupted by ribosomes as they translate the preceding ORF. In this case, reinitiation need not be the exclusive or major mechanism for translation of the downstream ORF. Indeed, a coupled region can evolve in such a way that translation of the second gene in a coupled pair is performed almost exclusively by ribosomes entering directly from the cellular pool (the facilitated-binding mechanism [456]). Translation of the upstream ORF seems to eliminate restricted accessibility for ribosomes from the cellular pool to the second TIR. Therefore, whereas reinitiation can support maximally a 1:1 ratio of translation of two coupled genes, the facilitated-binding mechanism can allow the downstream gene in a coupled gene pair to be much more efficiently translated than its upstream partner. The ratio of translation rates in the latter case will be linked to both the stability and the (re)folding kinetics of structure in the downstream TIR (124, 456).

Whereas the driving force for the facilitated-binding mechanism is generally provided by translating ribosomes, reinitiation on prokaryotic mRNAs in which the stop and start codons are not overlapping requires nondissociative ribosome movements that resemble eukaryotic scanning. Very little is understood about this type of posttermination ribosome-mRNA interaction, but the known properties can be usefully compared with eukaryotic scanning and reinitiation. In anticipation of the discussion to come, we should consider briefly the characteristics of reinitiation as opposed to de novo initiation on the UUG codon in E. coli. This codon is a very poor de novo start site but functions quite well as a reinitiation site (2, 519). It has been suggested that IF3 proofreading may be responsible for this difference in selection of UUG as a start codon (2, 519). Thus, if IF3 is unlikely to rebind to a posttermination 3OS ribosomal subunit (or 7OS ribosome) that has not dissociated from the mRNA, its absence may prevent the proofreading function from discriminating against UUG. Therefore, the principle underlying UUG selection by a reinitiating prokaryotic ribosome may be the modulation of (30S) ribosomal subunit behaviour via a translation factor.

Finally, the relationship between termination of the polypeptide chain and the posttermination behaviour of ribosomes on the mRNA will be influenced by the RFs. Of particular note are the observations of Ryoji et al. (471) on the reinitiation capacity of ribosomes in the absence of RF4. The system studied was a mutant form of bacteriophage R17, which has an amber codon at position 7 of the coat cistron. Readthrough to produce a full-length coat protein required a suppressor tRNA. In the absence of RF4, reinitiation occurred at the codon immediately 3’ of the amber codon, producing a coat protein lacking the first 7 amino acids. Thus, RF4 presumably acts to prevent recognition of a further codon (presumably at least initially in the A site) once a stop codon has been reached. The intriguing question is how rapidly RF4 promotes dissociation of the ribosome from the mRNA subsequent to release of peptidyl-tRNA from the P site. The relative kinetics of these component steps of the termination pathway are clearly of direct significance to the ability of ribosomes to reinitiate on adjacent reading frames.

Eukaryotic reinitiation. Cellular eukaryotic mRNAs are generally monocistronic, in the sense that each carries only a single major ORF. However, according to the most recent estimates, there are likely to be at least a few hundred genes in S. cerevisiae with short uORFs (572) (see examples in Table 3), as well as many more uORF-containing mammalian mRNAs (299). Yeast mRNAs can contain one to at least six uORFs, and they can appear as distinct entities or overlapped with each other on the main reading frame. There seem to be two different modes of action of eukaryotic uORFs. In the first, the function of the uORF is independent of the coding potential of the uORF. In the second, the influence of the uORF on gene expression depends on the uORF-encoded peptide (343). The most intensively studied eukaryotic uORF-containing mRNA is that of GCN4, which has an exceptionally long
5′UTR with special properties (222, 223, 391, 548, 559). Work on this system has provided unequivocal evidence of reinitiation both as an alternative pathway to the standard scanning mechanism and as a basis for translational regulation in response to environmental stress (226) (Fig. 11). On the other hand, quite different aspects of uORF function have been revealed by studies on other mRNAs. Some of these are also covered in this section, while others are discussed below in the section on mechanisms controlling mRNA stability.

**GCN4.** A remarkable feature of the response of *S. cerevisiae* to the deprivation of amino acids or purine is that while overall protein synthesis is partially inhibited, GCN4 expression is up regulated. This induction of synthesis of Gcn4p, which is a transcriptional activator, leads to the activation of at least 40 genes involved in amino acid biosynthesis (224). The currently available evidence is consistent with a model in which the kinetic control of reinitiation underlies the posttranscriptional regulation of this gene. The two key components of this model are the special structure of the GCN4 leader (Fig. 11A) and the ability of the cell to regulate the phosphorylation status of eIF2 in response to nutrient limitation (Fig. 11B). The GCN4 leader has evidently evolved in such a way as to ensure that initiation on the main reading frame is performed almost exclusively by posttermination ribosomal subunits which have already gone through at least one cycle of initiation and termination on an uORF. The recovery of the posttermination ribosomal subunits to resume initiation competence is modulated by the availability of the active ternary complex eIF2–GTP–Met-tRNAi. Thus, uORF-containing 5′UTRs can act as sites of termination and reinitiating ribosomal subunits with the ternary complex. It is the relocation of this repriming reaction from the cellular ribosome pool(s) to the GCN4 5′UTR which is thought to couple the reinitiation process to the phosphorylation state of eIF2 (Fig. 11B).

The historical development of studies of the mechanisms regulating eIF2 activity is detailed in a number of reviews (224–226, 228), and only the essential details are summarized here. The ability of the yeast cell to respond to starvation conditions by triggering GCN4 induction was found to be affected by mutations in a large group of unlinked genes. These fell into two groups, with the phenotypes general control non-derepressible (Gcn) and general control derepressed (Gdc). The former mutations affect genes encoding positive regulators of GCN4 expression and prevent GCN4 induction, whereas the latter mutations cause constitutive derepression. The action of all of these mutations on GCN4 regulation was found to be dependent on the presence of the uORFs. The *gcn* mutations typically affect either the activity of the Gcn2 kinase or that of eIF2B, while the *gcd* mutations were generally found to affect eIF2B or eIF2 (Fig. 11B). Consistent with the proposed role of eIF2 in GCN4 regulation, other mutations in genes encoding constituent subunits of this heterotrimeric factor (SUI2 and SUI3) were also found to confer the Gcd phenotype. An additional bonus of analyzing many of the *gcd* and *gcn* mutations is that they were found to encode component subunits of eIF2, eIF2B, and eIF3 which had not previously been characterized. For the sake of clarity, the principles of both kinetic control and regulation of GCN4 expression are presented together in this one section.

The work of the Hinnebusch group on GCN4 has provided many important details of the reinitiation process in *S. cerevisiae*, and the reader is referred to a series of earlier reviews for comprehensive and incisive discussions of the large body of experimental data on this system (225–227). The analysis of the GCN4 leader has made use of a large number of substitutions, deletions, and insertions as well as *lacZ* fusions with both the main reading frame and the respective uORFs (Fig. 11A). The following key features, which are most easily interpreted in terms of the scanning model of initiation, emerged from this work.

**(i) There are essentially two types of uORF.** Each of the four uORFs is initiated upon with a similar efficiency while individually inhibiting downstream translation (of GCN4) to a greater or lesser degree under nonstarvation conditions. uORF1 and,

<table>
<thead>
<tr>
<th>Gene</th>
<th>Length of major 5′UTR (nt)*</th>
<th>No. and size of uORFs (no. of codons)</th>
<th>Product (reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBS1</td>
<td>101</td>
<td>uORF (4)</td>
<td>PET gene involved in the 5′ end processing of the cytochrome b (57)</td>
</tr>
<tr>
<td>CLN3</td>
<td>864</td>
<td>uORF (1)</td>
<td>G1 cyclin (441)</td>
</tr>
<tr>
<td>CPA1</td>
<td>244</td>
<td>uORF (26)</td>
<td>Small subunit of cytosolic carbamoyl phosphate synthetase (406)</td>
</tr>
<tr>
<td>DCD1</td>
<td>33</td>
<td>uORF (4)</td>
<td>dCMP deaminase (366)</td>
</tr>
<tr>
<td>GCN4</td>
<td>591</td>
<td>uORF1 (4), uORF2 (3), uORF3 (4), uORF4 (4)</td>
<td>Transcriptional activator of amino acid biosynthetic pathway (391)</td>
</tr>
<tr>
<td>HAP4</td>
<td>−280</td>
<td>uORF1 (10), uORF2 (4)</td>
<td>Subunit of transcriptional activator complex binding CCAAT (156)</td>
</tr>
<tr>
<td>HOL1</td>
<td>−385</td>
<td>uORF (6)</td>
<td>Major facilitator family (drug resistance subfamily) of putative transport proteins (598)</td>
</tr>
<tr>
<td>LEU4</td>
<td>85</td>
<td>uORF1 (13)</td>
<td>α-Isopropylmalate synthase (cytoplasmic) (43)</td>
</tr>
<tr>
<td>PET111</td>
<td>459</td>
<td>uORF1 (6), uORF2 (31), uORF3 (11), uORF4 (30)</td>
<td>Mitochondrial translational activator (528)</td>
</tr>
<tr>
<td>PPR1</td>
<td>50</td>
<td>uORF (6)</td>
<td>Regulatory protein controlling transcription of two genes in pyrimidine biosynthesis pathway (278)</td>
</tr>
<tr>
<td>SCHO9</td>
<td>−600</td>
<td>uORF (55)</td>
<td>Protein kinase that positively regulates the progression of yeast through G1 phase (56)</td>
</tr>
<tr>
<td>SCO1</td>
<td>−150</td>
<td>uORF (3)</td>
<td>PET gene involved in the accumulation of cytochrome c oxidase subunits I and II (304)</td>
</tr>
<tr>
<td>TIF4631</td>
<td>295</td>
<td>uORF1 (12), uORF2 (20), uORF3 (16), uORF4 (8), uORF5 (12), uORF6 (22)</td>
<td>Translation inhibitor factor p150 (180)</td>
</tr>
<tr>
<td>YAP1</td>
<td>164</td>
<td>uORF (7)</td>
<td>Stress-related transcription factor (386)</td>
</tr>
<tr>
<td>YAP2</td>
<td>157</td>
<td>uORF1 (6), uORF2 (23)</td>
<td>Stress-related transcription factor (62)</td>
</tr>
</tbody>
</table>

* nt, nucleotides.
to a lesser extent, uORF2 are significantly less inhibitory than uORF3 and uORF4. Two lines of evidence indicate that uORF1 is efficiently translated and promotes downstream reinitiation (183). First, moving uORF1 to within 50 nucleotides of GCN4 inhibited GCN4 translation by approximately 90%, indicating that leaky scanning past uORF1 does not provide the mechanism for ribosomes to reach the main ORF (Fig. 11A). Second, extending uORF1 via fusion so that the resulting uORF overlaps GCN4 by 130 nucleotides abolished GCN4 expression. Ribosomes scanning past the uORF1 AUG would be expected to have remained unaffected by this change.

Moreover, under derepressing (starvation) conditions, uORF1 and uORF2 act as positive control elements that relieve the strong negative effects of uORF3 and uORF4. It became clear during the early experiments that much of the regulatory character of the leader was maintained if only one of each of the two classes of uORF (uORF1 and uORF4) was left intact after elimination of the start codons of uORF2 and uORF3 by point mutations. Thus, with only uORF1 and uORF4 in the leader, amino acid starvation still induces GCN4 expression by a factor of 7 (228).

FIG. 11. Features of GCN4 regulation involving modulation of eIF2 activity and the roles of short uORFs (222–227). (A) Diagram indicating the lengths (in nucleotides) of the respective uORFs and noncoding regions of the GCN4 leader. Since uORF2 and uORF3 are not essential for regulation, the scheme focuses on the roles of uORF1 and uORF4. The effects of alterations in the intercistronic distances on the estimated level of reinitiation on uORF4 and the measured rate of translation of GCN4 (as a GCN4-lacZ fusion) are indicated in columns above the illustration of the GCN4 leader. Also shown is a sketch representing simple theoretical gradients of reinitiation competence as a function of time and/or nucleotide sequence travelled in the repressed and derepressed states. These gradients have been used as the basis for a model of regulation of GCN4 repression mediated by changes in the status of eIF2 phosphorylation. It is, however, still uncertain how accurately they reflect the true dynamics of change in the status of ribosomal subunits on the GCN4 leader. These will undoubtedly be a complex function of both distance and various sequence effects. (B) Proposed cycle of events regulating the level of activity of eIF2 in the yeast cell. Under starvation conditions, part of the eIF2-GDP population is "sidelined" into a form restricting GDP-GTP exchange via phosphorylation of eIF2 by GCN2 kinase in response to increased levels of uncharged tRNA. See the text for further details and references.

VOL. 62, 1998 POSTTRANSCRIPTIONAL CONTROL OF YEAST GENE EXPRESSION 1517
tion. AU-richness in the penultimate uORF codon and in the downstream region is correlated with efficient reinitiation whereas GC-richness confers an inhibitory phenotype (182). The leader sequence upstream of uORF1 was also found to influence the ability of this uORF to promote reinitiation (184), but the mechanistic basis of this result is not yet clear. It has been suggested that a nucleotide sequence environment rich in Gs and Cs might cause the ribosomes terminating on uORF4 to pause long enough to allow a factor equivalent to RF4 to bind and cause ribosomal release (226). On the other hand, by analogy to the influence of terminal amino acids in the nascent polypeptide in *E. coli* on termination efficiency (53), the C-terminal codon(s) of an uORF might exert an influence on the ability of a ribosome to reinitiate via interactions with the encoded peptide product. Progress toward an understanding of the molecular mechanism underlying the choice between ribosomal release and reinitiation will be dependent on the provision of kinetic data on the component steps of termination, including peptidyl-tRNA hydrolysis, eRF binding, and deacylated tRNA release, as a function of the context of the stop codon. A further unanswered question is to what extent the shortness of uORFs such as those in the *GCN4* leader might modify the normal course of termination, as occurs at the end of a main reading frame near the 3′ end of the mRNA, thus creating a special sensitivity to the stop codon environment.

(ii) **The spacing between sites of termination and initiation is critical.** The spacing between the respective uORFs and between them and *GCN4* is a crucial factor in *GCN4* regulation. This is because the spacing has to be matched to the rate of change of the reinitiation competence of ribosomes as they move through the leader (Fig. 11A). The positive influence of uORF1 (and uORF2) on *GCN4* expression is attributable to the ability of posttermination ribosomes to ignore the inhibitory uORFs further downstream when they have not yet reattained initiation competence through the binding of eIF2–GTP–Met-tRNA, Figure 11A includes a theoretical gradient for the change in reinitiation competence as a function of distance travelled after termination on uORF1. Under repressing (nonstarvation) conditions, reinitiation competence is regained more rapidly in the derepressed state. This explains the marked sensitivity of the system to changes in uORF1-uORF4 spacing under derepressed conditions. Reinitiation at uORF4 is strongly compromised when this spacing is reduced, whereas extending the spacing has the opposite effect (1, 182). In the latter case, the increase in uORF recognition is paralleled by a decrease in *GCN4* expression. The spacing between uORF4 and *GCN4* is also critical: reducing it restricts the space (and time) available for scanning ribosomes to regain initiation competence, thus attenuating *GCN4* expression.

(iii) **Regulation is dependent on a fine balance between uORF types and intersite spacing.** The *GCN4* posttranscriptional regulatory system can work effectively only when uORF types 1 and 4 are placed in their natural order and their spacing is kept within defined limits. The mechanism of this sensitively poised system seems generally well accounted for by the basic model proposed by Hinnebusch and colleagues (226). Derepression of *GCN4* is explained by proposing that the phosphorylation of eIF2α slows the kinetics of active ternary-complex binding to the ribosomes scanning on after termination on uORF1, thus causing reinitiation competence to peak, or reach a plateau, beyond uORF4 (Fig. 11B). However, there may be some additional forces at play which cause certain observed small deviations from the behavior that is predicted on the basis of the model. For example, even uORF1 or uORF4 alone in a *GCN4* leader otherwise devoid of uORFs was found to support a *GCN4* induction ratio of up to 2 (see, for example, references 226 and 228). Although this induction ratio is small compared to that of the intact *GCN4* system, the underlying cause may be generally relevant to the function of uORF-containing 5′ UTRs in yeast.

A more major issue was raised by a comparative study of reinitiation on uORF4 and the *GCN4* main reading frame when these are positioned downstream of uORF1 in a series of test constructs (183). Analysis of the effects of varying the distance between uORF1 and the downstream ORF revealed that ribosomes become competent to reinitiate on uORF4 over much shorter stretches of intervening sequence than they do when uORF1 is followed by *GCN4*. Moreover, the efficiency of reinitiation on uORF4 at shorter intercistronic distances seems to be much more sensitive to the availability of active eIF2. Grant et al. (183) interpret these data in terms of two types of model. The first postulates that a slow step in initiation or elongation on uORF4 restricts the rate of access of scanning ribosomes to the AUG under nonstarvation conditions, thus causing a queuing of ribosomal subunits that have a correspondingly increased period subsequent to termination on uORF1 during which they can rebind the ternary complex. The second possibility is that the sequence of the uORF4 region promotes the binding of an additional, as yet uncharacterized, factor necessary for reinitiation. This could be a factor which, like eIF3, is needed to promote the binding of eIF2. Whatever the potential role of additional factors, the former model touches upon a kinetic principle that may be of key importance to the role of uORFs. The very translation of a short uORF may be defined by a different set of kinetic parameters from those applicable to a full-sized gene. The most obvious consequence of the short length of a three- or four-codon uORF is that it cannot allow simultaneous initiation, elongation, and termination and can be translated only by a single ribosome at any one time. The potential for direct feedback from elongation and/or termination on the initiation process is therefore obvious and may represent a factor that contributed to the evolution of such small uORFs in the *GCN4* system. It would be of particular interest to know what role the termination process on the uORFs plays in controlling the reacquisition of initiation competence by scanning ribosomes. Finally, the above discussion emphasizes the importance of detailed quantitative experimentation to the analysis of such a finally balanced regulatory system.

(iv) **Regulation of GCN2 kinase activity.** Since the discovery that an intact GCN2 kinase (Gcn2p) is essential for *GCN4* derepression (224, 466), much evidence has accumulated which indicates that this kinase plays a key role in the regulation of eIF2 activity via phosphorylation. GCN2 kinase has a multidomain structure, including a protein kinase domain and a histidyl-tRNA synthetase (HisRS)-like domain (583). It also possesses a C-terminal region responsible for binding to the ribosome (452). The current model of the stimulation of GCN2 kinase-mediated phosphorylation of eIF2α (226), (Fig. 11B) involves the binding of an uncharged tRNA to the HisRS domain, which in turn allosterically activates the kinase domain. The resulting stimulation of GCN2 kinase activity leads to increased phosphorylation of Ser51 on eIF2α. There is a correlation between the GCN2 kinase-dependent phosphorylation of Ser51 and the level of induction of *GCN4* (125). Moreover, phosphorylation of this site by the mammalian kinases PKR (double-stranded RNA dependent kinase) and HRI (hemin-controlled repressor) in yeast led to constitutive induction of *GCN4* (126). Significantly, the level of phosphorylation catalyzed by these two enzymes was unphysiologically high, causing strong growth inhibition. This emphasizes the
A significant part of this model is based on recent studies of puzzle, allowing the proposal of a more detailed model (227). Change by the phosphorylated form of eIF2 (423a).

The localization of reinitiation events to a specific leader protein of the 

50% of this model is based on recent studies of puzzle, allowing the proposal of a more detailed model (227). Change by the phosphorylated form of eIF2 (423a). (GCD2) subunits mediates the inhibition of nucleotide exchange by the phosphorylated form of eIF2 (423a).

Either of these Gcn1p-Gcn20p-mediated events would then promote eIF2 kinase (227). Alternatively, overexpressing the eIF2B subunits renders eIF2 phosphorylation less inhibitory (127). Ongoing work is revealing more details of the functions of the component subunits of eIF2B. For example, a subcomplex comprising the subunits γ(GCD1) and ε(GCD6) catalyzes GDP-GTP exchange on eIF2 while a regulatory subcomplex comprising α(GCN3), β(GCD7) and δ(GCD2) subunits mediates the inhibition of nucleotide exchange by the phosphorylated form of eIF2 (423a).

Other results are providing additional pieces of the GCN4 puzzle, allowing the proposal of a more detailed model (227). A significant part of this model is based on recent studies of GCN1 and GCN20, both of which are required for activation of the GCN2 kinase in starved cells (GCD1). Overexpression of the gene encoding the glutaminase subunit of the carbamoyl-phosphate synthetase A (arginine pathway) is repressed approximately fivefold by arginine. CPA1 was one of the earliest yeast mRNAs to be found to have an uORF-AUG-UUG eliminates repressibility of CPA1 leader has a single, relatively long uORF comprising 26 codons (including the stop codon). Moreover, mutational analyses performed on it have revealed the operation of a different type of regulatory principle, indicating that it acts according to the second type of mechanism outlined at the beginning of this section (Fig. 12). Mutation of the uORF AUG to UUG eliminates repressibility by arginine while having no significant effect on the absolute level of nonrepressed expression (588). This apparent lack of any stimulation of CPA1 expression was originally taken to mean that the uORF is not efficiently translated by ribosomes (588). However, further reflection by this research group on subsequent experimental findings (110) led them to suggest that the CPA1 uORF might in fact simply promote efficient
reinitiation, analogously to uORF1 in the GCN4 leader. This is fully consistent with the fact that the CPA1 leader has an AU-rich downstream sequence, which also promotes reinitiation in the GCN4, YAP1, and YAP2 mRNAs (225, 572), but does not rule out the possibility that leaky scanning through the uORF contributes to downstream translation.

The most striking outcome of the analysis of CPA1 was that missense and nonsense mutations in the CPA1 uORF eliminate arginine repressibility whereas silent nucleotide substitutions have no effect (120, 586). It therefore seems that the peptide encoded by the uORF plays an important role in arginine regulation of CPA1 (Fig. 12). The failure of mRNAs carrying a wild-type uORF to complement mutations affecting the composition of the encoded peptide shows that the coding sequence can act only in cis. Remarkably, arginine-dependent repression is to a large extent retained when the CPA1 uORF is fused to the lacZ gene (120). This means that the encoded peptide sequence acts to repress gene expression even when fused to a large protein that is unrelated to the system.

A limited deletion analysis of the 5′ and 3′ regions surrounding the uORF indicated that at least some of the CPA1 uORF-flanking sequences are not essential for arginine repressibility (120), although this study paid scant attention to the nucleotides closest to the uORF. In an alternative type of approach, the CPA1 uORF was inserted at the site of either uORF1 or of uORF4 in a derivative GCN4 leader from which the original uORFs had been eliminated. This experiment indicated that the GCN4 uORF1 type of leader environment allows the CPA1 uORF to promote efficient reinitiation at the GCN4::lacZ fusion ORF and to mediate arginine repressibility. In contrast, at the GCN4 uORF4 position, the CPA1 uORF is strongly inhibitory and does not confer arginine repressibility on the mRNA. It should be added that the fact that the later experiments (120) evidently did not include controls for potential differences in mRNA abundance and/or stability introduces an element of uncertainty into the interpretation of the data, which remains to be addressed.

It is therefore clear that resolution of the effects of position and sequence environment on the function of the CPA1 uORF will require further investigation. While the most likely explanation of the data on the CPA1 system is that the peptide product of the uORF mediates the repressive effect of arginine, the mechanism remains undefined. The fact that the uORF only functions in cis might be explained by a requirement for interactions between the peptide product and the translational machinery during or immediately after synthesis. To what extent the stability or potential rate of diffusion of the peptide product might influence its effective range of action is unknown. Why should the uORF-encoded peptide block expression in an arginine-dependent manner? It has been suggested that the peptide is retained on the translating 80S ribosome in the presence of arginine and possibly also of the regulatory gene product CPAR, forming a complex that somehow blocks the initiation of further ribosomal subunits on CPA1 (Fig. 12). The role of CPAR is, however, unclear, because the effects of cpAR mutants on CPA1 expression may be attributable to indirect effects on ribosomal reinitiation downstream of uORFs (120a). Since the uORF-encoded peptide also functions when fused to β-galactosidase, the blocking effect is presumably imposed on translating 80S ribosomes. It would be of interest to know if this sequence also functions when inserted further 3′ of the lacZ start codon. Biochemical analyses of the proposed “pausing” behavior, as well as genetic suppression analysis starting from mutations in the uORF (or cpAR), should prove illuminating. Investigations of the human cytomegalovirus gpUL4 (gp48) mRNA provide an interesting parallel (74). Current evidence is consistent with a peptide-dependent mechanism leading to ribosomal arrest at termination on uORF2 of this transcript. Cao and Geballe could detect a species representing the uORF2 peptide still covalently bound to the tRNA\(^\text{t\(1\)}\) that decodes the penultimate codon of uORF2. Translation is therefore thought to arrest at the stop codon with the peptidyl-tRNA still bound to the ribosome (74).

How does CPA1 regulation relate to that of the GCN4 system? Apart from the distinctive role of the CPA1 uORF-encoded peptide, there may be a common principle of regulating reinitiation downstream of the uORF. Both the nonpressed inhibitory influence and the arginine repressibility of the CPA1 uORF respond to the sequence environment of GCN4 uORF1 and uORF4 in ways that reinforce this idea. Perhaps, therefore, at least one function of the CPA1 uORF-encoded peptide sequence is to regulate the fate of terminating ribosomes in an arginine-dependent fashion. The peptide-based regulatory mechanism may effectively switch the posttermination events between the two extremes otherwise typical of uORF1 and uORF4 in the GCN4 leader. This would provide an alternative model for CPA1 regulation to the “blocking” mechanism proposed originally (Fig. 12). Again, it should be possible to distinguish between such alternative models experimentally. It may, however, be necessary to consider less obvious mechanisms.

For example, the CPA1 uORF-encoded peptide might have a quite different significance, perhaps targeting the mRNA to a particular compartment or binding an as yet unidentified additional factor. Finally, since the majority of the work on this system has focused on translational regulation, it would be easy to forget that other results have identified differential mRNA stability as a potential explanation of arginine-induced CPA1 repression (101). It would therefore be premature at this stage to assume that even the broad outline of CPA1 posttranscriptional control is settled.

An analogous uORF-ORF constellation in the arg-2 mRNA encodes the carbamoyl phosphate synthetase small subunit in two other fungi (345, 579), and comparative analysis promises to provide further insight into at least some of the questions concerning this type of regulation. The arg-2 uORF encodes a peptide with a sequence closely related to that of CPA1. Toprinting analysis data were consistent with ribosomal stalling occurring at the stop codon (and perhaps one other site) of the uORF at high arginine concentrations (579). Since modification of the start codon context of the uORF to allow enhanced recognition inhibited the translation of the arg-2 uORF, it seems likely that the uORF does not support efficient reinitiation. Thus, leaky scanning and regulatable ribosomal stalling on the uORF seem to be the primary events dictating the translational behavior of the arg-2 mRNA. The mechanism by which the stalling is able to respond to changes in the arginine concentration remains to be determined.

**YAP1 and YAP2.** Two further uORF-containing mRNAs have recently been investigated (572). The YAP1 and YAP2 genes encode transcription factors manifesting strong homology to API-like factors in higher eukaryotes and to yeast Gcn4p (62, 202, 386, 599). These regulatory genes are involved in the mechanisms used by the yeast cell to protect itself in various situations of stress. For example, overexpression of YAPI and YAP2 confers general stress resistance to a variety of unrelated compounds from metal ions to different inhibitors and drugs (62, 192, 219, 230, 326, 489, 556, 599). The YAPI leader has one 7-codon uORF, whereas the YAP2 leader has one 6-codon uORF (uORF1) and an overlapping short reading frame (uORF2) of 23 codons which is positioned −1 with respect to the main reading frame (572). The uORFs of the respective mRNAs influence expression in quite different ways.
This was initially indicated by examining the effects of eliminating them from the leaders (Fig. 13A). Thus, by using centromeric plasmid constructs, it could be shown that while removal of the \textit{YAP1} uORF had little effect on the resistance of yeast to H$_2$O$_2$, elimination of the \textit{YAP2} uORF start codons greatly enhanced cellular resistance to heavy metals.

Further examination of the effects of these changes on the expression of reporter genes indicated the nature of the underlying control mechanisms. First of all, the apparent lack of influence of the \textit{YAP1} uORF was shown not to be due to the absence of recognition of this uORF by ribosomes. Instead, the \textit{YAP1} uORF shows similar functional characteristics to those of \textit{GCN4} uORF1; it is recognized by approximately half of the ribosomal subunits scanning along the \textit{YAP1} leader, and these show efficient reinitiation subsequent to termination on the uORF. Translation of the downstream main reading frame (\textit{YAP1} or a reporter gene) is therefore performed by both ribosomes that have scanned through the uORF (leaky scanning) and by reinitiating ribosomes. In contrast, the \textit{YAP2} uORFs do not allow efficient reinitiation, and act to block the expression of downstream reading frames. The \textit{YAP2} type of uORF therefore seems to fall into the same class as \textit{GCN4} uORF4. Indeed, comparison of the flanking sequences of the \textit{YAP} and \textit{GCN4} uORFs suggests that there is a correlation between the sequence environment of these uORFs and their functional influence on gene expression (572). This proposal was borne out by experiments in which the flanking sequences and the penultimate codon of the \textit{YAP} uORFs were manipulated. It was found to be possible to interconvert the characteristics of the respective uORF types so that the \textit{YAP1} uORF assumed the characteristics of the \textit{YAP2} type of uORF, and vice versa (572, 573).

Consideration of the effects of the \textit{YAP} uORFs on translation suggests that we have found additional examples of principles of translational control originally described for the \textit{GCN4} system. However, this applies only in the sense that the individual uORFs manifest similar properties in the respective systems; the \textit{YAP1} and \textit{YAP2} uORFs are not configured in such a way as to allow \textit{GCN4}-type regulation. Thus, neither of the \textit{YAP} mRNAs is subject to translational regulation coupled to eIF2 activity (573). However, the intrinsic properties of the respective \textit{YAP} uORFs are very similar to those of the \textit{GCN4} uORF1 and uORF4 types, so that, organized appropriately, they can form the basis of a regulatable system of the \textit{GCN4}.

FIG. 13. The \textit{S. cerevisiae} mRNAs \textit{YAP1} and \textit{YAP2} have different types of uORF, which affect translation and mRNA stability in distinct ways. (A) Elimination of the respective uORFs via mutation of the start codons (AUG→AAG; crosses in the construct maps) reveals striking differences which are most readily detected by using a reporter gene (in this case \textit{LUC}, encoding luciferase). Removal of the uORF from the natural \textit{YAP1} leader (puY1) has little effect on translation or mRNA stability (p\textit{D}uY1). The \textit{YAP2} uORFs, in contrast, have been found to influence translation and mRNA stability (here reflected in changed steady-state mRNA levels). Both translation and relative mRNA levels were increased upon removing either \textit{YAP2} uORF1 (p\textit{D}uY1) or \textit{YAP2} uORF2 (p\textit{D}uY2) or both uORFs (p\textit{D}u(1+2)Y2) from the 5′ UTR. The wild-type and mutant sequences of the \textit{YAP1} uORF and \textit{YAP2} uORF are indicated. (B) Mechanism by which the \textit{YAP2} uORFs affect mRNA decay. The Northern blots show the disappearance of the \textit{YAP2} mRNA signal as a consequence of inactivating the PolII promoter in a heat-sensitive \textit{rpb1} mutant. The destabilization effect is largely \textit{UPF1} independent. Reproduced from reference 572 with permission of the publisher.
type. Thus, combining the YAP1 uORF1 with YAP2 uORF1 in a newly created leader was found to support the inducibility of the downstream reporter gene in response to amino acid starvation (546). As with the GCN4 system, the degree of inducibility was found to be a function of the distances between the uORFs.

In conclusion, while the YAP uORFs have the intrinsic potential to perform adequately as building blocks in a GCN4-type regulatory system, they are not used to this effect in the wild-type YAP mRNAs, although some form of stress-related modulation of at least YAP2 has not been ruled out. However, further investigation has revealed a further property of the YAP2 type of uORF: it acts to destabilize mRNA. This new function for a natural uORF will be revisited in the section on mRNA stability.

CLN3 and other examples of uORF-related control. Since there are many further examples of uORF-containing mRNAs (441, 572), we can expect to hear more about the roles of uORFs in modulating the synthesis of a whole range of proteins. CLN3 has recently been identified as an mRNA whose translation is restricted by a four-codon uORF (441). Theycin encoded by CLN3 is required for the normal passage of the cell cycle through the G_1→S transition. By attenuating the translation of CLN3, the uORF contributes to the control of the cell cycle at this step. Elimination of the uORF start codon, for example, increased budding in late log phase, most probably due to accelerated progress through Start. The wild-type CLN3 leader is therefore capable of inhibiting cell division as growth becomes limited. An important question is whether the attenuation of CLN3 translation associated with inhibition of the overall capacity of the protein synthesis machinery is a specific effect. Polysomal fractionation experiments with a mutant strain (edc63-1) with a defective γ subunit of eIF3 revealed different degrees of shifting of the distribution of the CLN3, ACT1, and SSA2 mRNAs between monosomes and polysomes in response to the mutation (441), indicating that there are different degrees of inhibition. The extent to which translation of the CLN3 mRNA is differentially attenuated relative to other cellular mRNAs is therefore an issue worthy of more intensive investigation. The theoretical possibility that differential control of the translation of an mRNA such as CLN3 can be exercised at the selection (initial binding) step is considered in the section on prokaryotic and eukaryotic translation.

**Internal Initiation in Yeast?**

The third type of initiation described so far in eukaryotic systems occurs without any identifiable mediation by the 5’ end of the mRNA. This is mediated by IRESs. It is striking that eukaryotic internal initiation, unlike the standard prokaryotic form of TIR-guided initiation, generally does not occur on mRNAs that are polycistronic in the sense that they contain more than one structural gene (260, 261, 407, 562). In other words, despite the demonstrable capability of the internal initiation pathway to promote the translation of both genes on a dicistronic mRNA, its primary physiological role is apparently to render the translation of certain major ORFs partially independent of the main host initiation pathway. Despite the high level of interest in the mechanism of internal initiation on viral and mammalian mRNAs, its relevance to yeast remains unclear. Attempts to promote internal initiation in S. cerevisiae cells by using the IRESs of poliovirus (100) and encephalomyocarditis virus (146) have failed. On the other hand, there have been reports that cap-independent initiation can be supported in yeast cell-free systems by mammalian and plant viral IRES sequences as well as by natural yeast leaders (12, 241). Moreover, the in vitro study of Iizuka et al. (241) also describes experiments with dicistronic constructs that apparently support internal initiation. A more recent study has indicated that the capacity of the yeast translational machinery to initiate on uncapped mRNAs in vivo can be increased by mutations in the cap-binding affinity of eIF4E (569). The implication of this result is that a latent potential for cap-independent translation is present in vivo but that it becomes measurably active only after modification of the cap specificity of the translation apparatus. However, this in itself does not constitute evidence for the operation of an internal initiation mechanism.

One of the unexpected aspects of the work of Iizuka et al. (241) is the fact that the yeast 5’UTRs reported to contain IRES elements show no readily identifiable characteristics that might distinguish them markedly from the majority of other leaders in this organism (156, 488). One of these leaders (HAP4) contains uORFs (Table 3), which may impose properties affecting posttranscriptional gene expression that are unrelated to IRES function. Overall, it remains uncertain whether bona fide internal initiation, as opposed to simply weakly cap-dependent initiation, can occur in vivo in S. cerevisiae. There has been some debate whether S. pombe might be better equipped for the internal pathway, but we await further news on the ongoing work in this area. A further twist to this tale is provided by the observation that S. cerevisiae contains a 60-nucleotide RNA species (called I-RNA) which has been shown to inhibit preferentially the translation mediated by the poliovirus IRES in mammalian systems (111, 112). Cross-linking studies with HeLa cell extracts revealed that I-RNA interacts with the human La autoantigen, which is suspected to be a trans-acting factor that somehow promotes IRES-dependent translation. S. cerevisiae seems to possess an La protein homologue, but this is nonessential (90), and thus, overall, the significance of I-RNA is still a mystery. Another study has implicated the yeast La homologue in tRNA maturation (567).

Despite the uncertainty about the potential significance of internal initiation in yeast, there are aspects of this phenomenon that are relevant to the discussion of yeast translational control. One of these is the fact that 5’-end-dependent initiation seems to require the 40S ribosomal subunit to form a “tight” association with the mRNA that allows scanning to occur in the absence of a high off-rate (Fig. 10). The analogy can be drawn to the operation of a type of clasp, which in molecular terms would be closed around the mRNA. Potential mechanisms of encirclement of the mRNA by the 40S ribosomal subunit have been discussed previously (260, 300, 301). Whatever the exact mechanism involved, there is currently no reason to assume that this close association can be achieved only via the 5’ end. Mediation of this clamping effect can be envisaged to occur via internal sequences, and this may effectively be the function of IRES elements. The extent to which the localization of the early steps of initiation to the 5’ end of most mRNAs is a reflection of structural characteristics of the eukaryotic ribosome, as opposed to kinetic control exercised by other components of the translational apparatus, is not known (Fig. 10). However, the latter explanation raises the possibility that there is a background level of internal initiation at non-IRESs that is normally insignificant but that could theoretically be enhanced by changes in the availability or state of components of the translational machinery. In this context, it should be noted that even if 40S ribosomal subunits in the cellular pool have a finite affinity for internal regions of mRNA, this is normally likely to be relatively low. Overall, these considerations reveal that, rather than being restricted to a choice between (fixed) mech-
anistically distinct pathways (see, for example, reference 479), the route taken can be explained in terms of kinetic control.

A model in which kinetic control determines the dominance of 5'-end- or IRES-dependent initiation can provide a basis for explaining how the balance can be shifted to normally less favored initiation pathways. Underlying the preference for the 5'-end- and IRES-dependent initiation events could be the reduced off-rates of ribosomal subunits that become bound (by a clasp mechanism) to the mRNA via these two types of site (Fig. 10). In contrast, the combination of both a high off-rate and the minimal probability of binding to an AUG-containing sequence so as to form an initiation-competent complex may normally discriminate strongly against non-5'-end- and non-IRES-dependent initiation (Fig. 10). Changes in the activities of components of the translation machinery or in the loading of mRNA with RNA-binding proteins may be sufficient to cause a shift in emphasis to less specific initiation. One particular case where such distortions may be particularly significant is translation in cell-free systems. For example, it is known that translation in reticulocyte lysates manifests poor fidelity and an increased incidence of initiations at normally insignificant internal start sites (113). Moreover, the ratio of mRNA to general RNA-binding proteins in this system has been observed to influence the level of priority given to cap-dependent initiation (334). Again, such considerations warn against the danger of misinterpreting in vitro expression data, since these may reflect a distorted emphasis on translation pathways that are less significant in vivo.

Another relevant property of internal initiation in mammalian cells is that it involves essentially the same set of canonical factors as cap-dependent initiation (260, 435). However, work on picornavirus IREs has revealed that the interactive properties of eIF4F (Fig. 5) influence the specificity of the translation apparatus. Cleavage of eIF4G by picornavirus proteases generates an N-terminal fragment that binds eIF4E and a C-terminal fragment that contains binding sites for eIF3 and eIF4A (312) (Fig. 5). In vitro studies have indicated that cleavage of eIF4G allows efficient initiation on IRES-containing (327, 619) and uncapped (408, 619) mRNAs but disrupts cap-dependent initiation. It has been proposed that this cleavage uncouples the eIF4E-dependent cap binding of the N-terminal domain from the C-terminal functions of ribosome binding (via eIF3) and interaction with helicase activity (312). The shift away from cap dependence is not as pronounced in vivo, since eIF4G cleavage alone does not suffice for complete inhibition of host protein synthesis (60, 434). Nevertheless, it achieves a similar effect to that seen when eIF4E cap specificity is reduced in S. cerevisiae (569): prioritization for cap-mediated initiation is relaxed. This can be explained in terms of elimination of the kinetic bias toward cap-mediated initiation, which, in turn, allows the translation apparatus to focus more on cap-independent or internal initiation. The molecular basis and kinetic control of the selection of capped mRNAs are discussed in the earlier sections on the initiation components and on pathways of translational initiation.

This is consistent with the general argument that internal initiation and 5'-end-dependent initiation differ primarily with respect to the route by which the 40S ribosomal subunits are guided into a “committed” association with the mRNA. 5' end (cap) recognition and IRES recognition by ribosomes may therefore constitute two alternative initial steps leading to what is fundamentally the same overall initiation process. In neither case is the mechanism that claps the 40S ribosomal subunit tightly onto the mRNA understood, but the molecular machinery involved seems to be more or less the same. However, the existence of these two types of initial pathway seems to confirm the principle that specific recognition events secure rapid access to bona fide start codons. As yet, there is no convincing evidence that S. cerevisiae has the capacity to recognize the IRES type of binding element in vivo. Moreover, it remains to be determined whether the yeast eIF4G domains equivalent to those generated by protease cleavage in mammalian cells (Fig. 5) can be used to shift the specificity of translational initiation. Finally, random access via alternative routes dependent on relatively nonspecific interactions is normally restricted by mechanistic or structural barriers (the conformation of the 40S subunit or masking of the mRNA by proteins) and/or kinetic control (Fig. 10).

**Alternative Coding Potential of Open Reading Frames**

There is an alternative dimension to the posttranscriptional control of gene expression in yeast. As in other eukaryotic and prokaryotic organisms, specific signals can induce “programmed” shifts away from the reading frame established by initiation at the major start codon (582). Both +1 and −1 frameshifting have been described in S. cerevisiae, and each type involves a distinct mechanism (148). The mechanisms of programmed frameshifts stand out against the background of an estimated frameshift error frequency of approximately $5 \times 10^{-5}$ per codon (307). Moreover, as discussed in this section, there is at least potential scope for regulation of programmed frameshifting events.

**+1 frameshifting.** Yeast transposable (Ty) elements are retrotransposons whose life cycle requires the expression of two overlapping genes in different reading frames, TyA and TyB, which are analogues of the retroviral gag and pol genes, respectively, whereby TyB is produced as a fusion to the TyA protein after the ribosomes have shifted +1 in the overlap region (95) (Fig. 14A). Not only is the effective shifting direction in the Ty elements different from that of retroviruses, but also the mechanisms are distinct. Frameshifting in Ty1 requires two key components: a core “slippage” sequence comprising 7 nucleotides (CUU AGG C), and a low-abundance Arg-tRNA$_{\text{CCU}}$ that is thought to manifest slow decoding kinetics (39). The working model for this system is that pausing occurs with the Leu-tRNA$_{\text{UAU}}$ in the ribosome P site pairing with the CUU codon in the 0 frame and the A site poised empty over the AGG (Fig. 14A). Repairing occurs between the Leu-tRNA$_{\text{UAU}}$ and the UUA, which leads to decoding of the following GGC (glycine) in the A site, thus establishing the +1 shift. Two further aspects of this system attract particular attention in terms of control. The first is that the cellular abundance of the Arg-tRNA$_{\text{CCU}}$ strongly influences the rate of frameshifting. Increasing the copy number of the corresponding gene fivefold reduced the efficiency from more than 40% to 1% (39, 603). This result is consistent with kinetic control of the frameshifting event which depends on the relative rates of binding of Arg-tRNA$_{\text{CCU}}$, to the A site and the rate of slippage of pairing by Leu-tRNA$_{\text{UAU}}$ from CUU to UUA. The second point is that frameshifting is inhibited by proximity to the start codon (39, 95). Frameshifting was completely inhibited by reducing the distance between the essential 7-nucleotide element and the start codon to two codons and became unaffected only once the gap was extended to four codons. It is unclear why the competence to undergo a frameshift is suppressed in this early region of the reading frame. Since the translational apparatus is very decisively sets the 0 reading frame at the start codon, it will be of interest to determine how this maintenance of frame is imposed on the first three or four codons.

Another mechanism of +1 frameshifting operates in Ty3 between GAG3 and POL3 (149). Analogously to Ty1, there is
a 7-nucleotide core slippage site (GCG AGU U), but the shift mechanism does not involve re-pairing from GCG to CGA, and the 14-nucleotide stretch downstream of the shift site promotes frameshifting. It is proposed that the A base between GCG and GUU is skipped (Fig. 14B) by an as yet unidentified mechanism. A further feature shared with the Ty1 system is that overexpression of Ser-tRNAGCU, which is the cognate tRNA for the AGU codon, reduces the estimated frameshift activity approximately 10-fold. Thus, the principle of inducing a pause at the second codon in the 7-nucleotide core sequence seems to be shared by the Ty1 and Ty3 types of frameshifting mechanism. Moreover, it seems unlikely to be an accident that two such similar codons (AGG and AGU) are used (149), indicating that the codon-tRNA interaction at this site must have specific properties. However, there is a striking difference in the codon-tRNA interaction immediately upstream of the shift site. The GCG codon in the Ty3 site is decoded by Ala-tRNAAGG, which cannot slip onto the +1 frame codon, CGA. Overproduction of a synthetic Ala-tRNAAGG was found to reduce frameshifting, indicating that the codon choice is important for the mechanism (419). The Farabaugh group also investigated whether codons other than GCG can assume the same role and found that in total only eight tRNAs can stimulate the +1 frameshift (575). It has been proposed that frameshifting is most probably mediated either by direct out-of-frame binding of Val-tRNAAC to the +1 codon GUU (perhaps promoted by tRNA “neighbor” effects associated with the upstream Ala-tRNAAGG) or by means of 4-base decoding involving a tRNA capable of participating in an extended codon-anticodon interaction.

**-1 frameshifting.** Double-stranded RNA (dsRNA) viruses are widespread among strains of *S. cerevisiae*. They show many similarities to the dsRNA viruses found in cells of higher eukaryotes and are responsible for the yeast killer phenotype (590). The dsRNA viruses have the overall gag-pol structure typical of the retroviruses. For example, in the L-A viral dsRNA, ORF1 of the (+) strand comprises 681 codons and encodes the viral particle major coat protein (80 kDa) whereas ORF2 comprises 869 codons and encodes RNA-binding and RNA Pol domains. ORF1 and ORF2 overlap in the −1 frame by 130 nucleotides. Frameshifting at the by now familiar retroviral type of slippery site (257) located in the overlap region yields a 180-kDa protein which incorporates the domains encoded by both ORF1 and ORF2. The latter is equivalent to the reverse transcriptase domain encoded in an analogous position in retroviruses. The L-A virus frameshifting site possesses a number of specific features of a retroviral type of −1 shift site (Fig. 14C). The shift sequence is followed by a structural element (possibly a pseudoknot [561]) that is thought to cause the translating ribosome to pause at the appropriate position. The most likely pathway is that tRNAUGG shifts −1 back from the GUU codon, thus reestablishing its nonwobble base pairs while changing the wobble interaction from a U to a G (130, 240). The efficiency of this process is equivalent to approximately 1.8% of the total number of ribosomes passing through the site. The mechanistic details of the various types of frameshifting have not been determined, but models are discussed by Farabaugh (148). One area of special interest is the potential role of trans-acting factors. These will certainly include proteins that act as general modulators of translational accuracy (such as those discussed in references 106 and 584 to 586), but it is uncertain whether there are factors that specifically promote frameshifting at −1 (or for that matter, +1) sites.

Could the efficiency of frameshifting at this site be subject to regulation? There is certainly no evidence for this in yeast. However, an intriguing proposition is that variation in the state of modification of bases in the tRNA anticodon loop (4) in cells of higher eukaryotes could be brought about by retroviral infection (205, 206), thus modulating the ability of at least certain tRNAs to re-pair on the −1 shift site. This is, however, a controversial proposition (63a). It remains to be determined whether there is any form of modulation of the capacity of tRNAs to re-pair in yeast.

**Alternative decoding.** Recoding is not evident in *S. cerevisiae* cytoplasmic translation but does occur in the mitochondrion, in which UGA encodes Trp (instead of stop), AUA encodes Met (instead of Ile), and CUN encodes Thr (instead of Leu). The only known occurrence of cytoplasmic recoding in a yeast is the reassignment of CUG from Leu to Ser in several species of *Candida* (415, 485). Finally, there has been keen interest in the possibility that yeast species incorporate the 21st amino acid, selenocysteine, into certain proteins. This is known to occur in representatives of the prokaryotes and archaea, as well as in other eukaryotic organisms. In all cases, it involves recoding of the UGA codon. Unfortunately, no unequivocal evidence for this phenomenon has been reported for either *S. cerevisiae* or *S. pombe*.

**Protein splicing.** It was recently discovered that there is a very different way in which the encoded information of a read-

---

**FIG. 14. Different forms of frameshifting in *S. cerevisiae*.** Two types of +1 frameshifting have been described. (A) The first type, observed in the yeast transposable element Ty1, involves re-pairing from the CUU codon to the +1 UUA codon. (B) The second type, described for Ty3, does not seem to involve re-pairing by the decoding tRNA interacting with the first codon of the “slippage site.” (C) Finally, the retroviral type of −1 frameshifting occurs in the dsRNA viruses of *S. cerevisiae*. 

---


**Published by the American Society for Microbiology. All rights reserved.**
ing frame can be interpreted to generate a noncolinear product (97). Protein splicing was first observed in S. cerevisiae. The TFP1 gene theoretically encodes a protein of 119 kDa, but the protein product generated is the 69-kDa catalytic subunit of the vacuolar H^+ -ATPase (V-ATPase) (231, 279). The remarkable solution to this puzzle was found to be the processing of the initial protein precursor. A 454-amino-acid segment (the intein) is excised via a process in which the N- and C-terminal exons are joined to yield the 69-kDa mature product. Protein splicing has been observed in the prokaryotes, archaea and eukaryotes, and much current work focuses on determining its mechanism (97).

**Prokaryotic and Eukaryotic Translation**

At this point, it is informative to return to the comparative assessment of the initiation pathways available to prokaryotic and eukaryotic ribosomes. Overall, there seem to be two different strategies for start codon localization. Interactions between rRNA and mRNA sequences, of which the ASD-SD interaction is the paradigm, constitute the primary driving force in the prokaryotic system. The prokaryotic 30S ribosomal subunit can theoretically obtain access to all internal sites on the mRNA, but its binding is guided by the thermodynamics of interaction with the mRNA (and possibly by the kinetics of transition from the preinitiation complex to the initiation complex proper [460]) to those TIRs that are relatively structure free and contain SDs that manifest the appropriate range of ASD-binding energies (Fig. 10). The energetics of this type of interaction evidently allow broad flexibility in the control of translational initiation in cells where transcription and translation are tightly coupled and nascent transcripts are accessible to ribosomes. It is likely that these early ribosome-mRNA interactions influence the folding pathways and functions of nascent mRNA and also have significant effects on the temporal control of gene expression on polycistronic mRNAs (185, 456). Moreover, the fact that ribosomes can directly access internal TIRs means that a polycistronic mRNA can support greatly varying translation rates from its respective genes, whereby the coupling ratio between neighboring genes can deviate significantly in either a positive or a negative sense from the obligatory ratio of 1:1 obtained under conditions of tight reinitiation (363, 456).

Overall, the eukaryotic system has not evolved to provide this form of flexibility. It is, however, equipped to deal with particular properties of mRNAs that emerge from the nucleus having undergone modification, possibly splicing, and interactions with a host of nuclear proteins and RNAs. The process of start site selection in eukaryotic cells has shifted away from the prokaryotic type of rRNA-mRNA interaction mediated process to principles of protein-mRNA mediation. The translation apparatus (or factors physically or functionally associated with it) is energetically competent to displace at least the majority of mRNA-binding proteins that accompany the mRNA out of the nucleus, apparently irrespective of the rate at which individual mRNAs are translated. The exceptions are the proteins that serve to “mask” or “silence” mRNAs in vertebrates and amphibia at early stages of development (457, 512, 589).

The adaptation of the eukaryotic translation apparatus to form a tightly binding complex during the preinitiation stage may even be a consequence of the need for it to function on “prepackaged” mRNPs emerging from the nucleus. According to the general model outlined in Fig. 10, nonspecific initiations at internal AUGs can be avoided by kinetic prioritization of the 5′-end (or IRES)-dependent pathway, which leads to tight association of the 40S ribosomal subunits with the mRNA and allows them to “scan” over long distances. At the same time, the alternative route to internal initiation is discouraged by high off-rates (and/or low on-rates) and possibly by relatively low resistance to blockage by certain RNA-binding proteins (534). The IRES-mediated channelling of 40S ribosomal subunits into the initiation pathway may involve ribosome-mRNA interactions analogous to those of the prokaryotic system, but this has yet to be experimentally proven. However, as with the prokaryotic TIR, the efficiency of IRES-mediated internal initiation does vary over a wide range as a function of IRES structure (61).

While prokaryotic and eukaryotic ribosomal subunits may share a low-affinity “scanning” facility (Fig. 10), the tightly associated eukaryotic 40S ribosomal preinitiation complex formed during the major initiation pathway is capable of scanning through structured regions which would be impenetrable to its prokaryotic counterpart (297, 481) and may have to negotiate that structure only during the scanning phase. This may mean that the selection step (i.e., initial binding) cannot distinguish between mRNAs that can be efficiently translated and those whose translation is restricted by structural elements in the 5′ UTR. The idea that this type of selection does occur forms the basis of some of the analyses mentioned at the beginning of this review and is frequently used to explain the behavior of specific mRNAs under various translation conditions. There is, however, no unequivocal evidence for mRNA selection of this kind, and this principle of control requires further consideration.

There is indirect evidence that scanning 40S subunits pause in front of a stem-loop structure (297), and there is little doubt that unwinding of stem-loops up to a given stability (Table 2) can occur. The inhibition curve of eukaryotic translational initiation in response to stem-loops of various stabilities in the 5′ UTR indicates that this is the eukaryotic form of thermodynamic control via secondary structure (481, 570). The extent to which blocked 40S subunits are forced to queue up on a structured leader, as opposed to abandoning it by virtue of a significant kₐ (Fig. 10C), is undetermined. The release rate that pertains in this situation may be different from that applicable to unstructured leaders but may have to be comparable to the scanning rate (kₛ) in order to prevent the formation of a queue of 40S subunits back to the 5′ end. It is also highly relevant to the question of mRNA selection and rate control exerted during initiation addressed above.

There is currently considerable excitement about the mechanisms and functional roles of termination and reinitiation in eukaryotic cells. How is the release of ribosomes and/or the capacity to reinitiate controlled? By analogy to E. coli, we might expect eRF1 and eRF3 to be involved in the response of the translational apparatus to signals in the vicinity of the stop codon. For example, a recycling-factor type of activity analogous to RF4 might be activated upon termination in certain sequence environments, such as those provided by GCN4 uORF4 or YAP2 uORF1. Alternatively, rapid progression from termination to resumption of the scanning mode might be
possible in at least some AU-rich stop codon environments (as in the GCN4 uORF1 or the YAPI uORF), whereupon the capacity to reinitiate may be governed by the kinetics of re-binding of specific factors. The arguments for the involvement of eIF2 and perhaps of other factors were considered earlier in this review. At first sight, the apparent involvement of IF3 in influencing the ability of prokaryotic ribosomes to reinitiate on a UUG start codon seems to suggest that the modulation of mRNA-bound ribosomes via initiation factor binding represents a conserved theme in evolution. However, for the prokaryotes, it is thought that the inability of the factor to bind to 30S ribosomal subunits in the scanning mode is required for reinitiation. For yeast GCN4, it is not yet known why reinitiation of ribosomes on the leader is so much more sensitive to ternary-complex activity than is general initiation, but it seems likely that kinetic control and/or involvement of distinct combinations of initiation factors in the respective processes is responsible (226). Further research into these factor-ribosome interactions will probably uncover novel mechanisms of control.

Unresolved Issues of Quantitative Control

The systemic type of analysis of multistep pathways considered earlier in this review formalizes the view that control is rarely concentrated in just one component step of a pathway. The definition of control used by Kacser and Burns (276, 277) is particularly appropriate for the binding and catalytic steps of the translational pathway that are nonprocessive. It provides a framework to represent the distribution of control among a number of steps. On the other hand, the Kacser and Burns type of treatment is less useful for processive reactions, especially elongation and possibly initiation. These steps can be approximated modelled by more standard procedures, provided that a number of assumptions are made (173, 340). Given the quantification of translation in terms of the rate of production of complete polypeptide chains, protein synthesis can be considered a process whose rate is determined by the rate of initiation and/or termination.

Although insufficient data are available for translation to be accurately modelled, the use of these theoretical treatments focuses attention on important questions. For example, it has been assumed that eIF4E functions can be attributed to large “sensitivity” coefficients, meaning that changes in eIF4E activity exert strong control on the rate of translation, but the experimental data available do not justify this. Another issue is the extent to which the limited availability of ribosomal subunits restricts the translation process, thus influencing rate control and/or selectivity (see, for example, the discussion of the pap1-1 mutant). A further open question is the degree to which termination, either directly or indirectly, controls the initiation rate in individual polysomes. This aspect was not considered in earlier work. Finally, one key factor in earlier models was the so-called discriminatory factor (173, 339, 340), which was presumed to be responsible for determining relative rates of translation on individual mRNAs but was not clearly defined. There is no evidence that eIF4E binding to different capped mRNAs in S. cerevisiae is generally controlled via features of the 5’UTR (314), so that eIF4F seems unlikely to function as a discriminatory factor. It presently seems more likely that “discrimination” in terms of initiation rate is determined via the modulation of ribosome progress along the mRNA leader. These and other questions make it clear that the relationship between mechanistic and kinetic models requires further clarification.

More information is required about the “control coeffi-
cients” of the components of the translational apparatus in terms of overall rate before quantitative data can be reliably interpreted in terms of mechanistic models. There are a number of ways in which this information can be obtained. Muta-
tional analysis and biochemical experiments can provide information about the effective control coefficients of translation factors, and, for example, rRNA-processing mutants might help with controlling ribosome concentration. In performing this type of analysis, it is essential to consider appropriate procedures for relating quantitative control effects to mechanistic function in vivo. Useful information on the control of ribosome throughput on individual mRNAs can be obtained by studying the response of translation rate to structure or protein binding in the leader. Since the yeast system shows no pronounced position dependence and an apparently straightforward, partially linear dependence of inhibitory effect on structural stability, it is tempting to apply a thermodynamic model of rate control equivalent to that used for initiation in a prokaryotic TIR (123, 527). However, even when approximations like this seem to fit the data, more information is required for the scanning mechanism to be characterized. For example, in vitro investigations of ribosome on- and off-rates will be necessary to fill in the missing numbers in the schemes shown in Fig. 3 and 10.

Finally, genetic and biochemical studies with S. cerevisiae show that the interactions between components of the translational apparatus are suggestive of functional networks. Conside-

r, for example, the complex interactions involving eIF4F proteins (Fig. 15). They illustrate the extensive functional interac-
tions and overlaps that can be identified for many of the proteins involved in posttranscriptional control and warn us that a process such as initiation cannot be adequately described solely in terms of a linear mechanistic pathway.

MECHANISMS OF TRANSLATIONAL REGULATION

The interactions between the translational machinery and the more than 6,000 different mRNA species in S. cerevisiae are governed by a number of different structural attributes of these mRNAs. Returning to the scheme of a “typical” mRNA (Fig.
2), the numbered regions indicate the diversity of sites at which even small structural changes can make a significant impression on the translation of a given mRNA. These domains of the mRNA can be involved in two types of translational regulation. General regulation is exercised via modulation of the activities of components of the translational machinery which, in turn, interact with more or less precisely defined regions of each mRNA. In contrast, variations in individual mRNA sequences are linked to gene-specific posttranscriptional regulatory events, which may involve site-specific protein binding or, as with GCN4, mRNA-specific responses to changes in translation factor activity.

Modulation of Translation Factor Activities

The main site of general translational regulation of the pool of eukaryotic cellular mRNAs is thought to be the initiation process (262, 330, 364). While this view may potentially do injustice to the role of other steps in translation, there is a considerable body of data implicating initiation factors in global regulation. The majority of these data indicate that eIF4E and eIF2 constitute key regulatory targets. However, the balance of evidence pertaining to these factors is different for yeast and mammalian systems.

eIF4E. Regulation of eIF4E activity is expected to allow the cell to modulate the 5′-end-dependent binding of 40S ribosomal subunits to mRNA. Two types of regulatory mechanism have been proposed: modulation of eIF4E cap-binding activity mediated by changes in the phosphorylation state of this protein, and regulation mediated by interactions between eIF4E and the so-called 4E-binding proteins. Before considering either of these, it is important to address the general perception that eIF4E must be a “limiting factor” in initiation because it is present in substoichiometric amounts relative to the cellular content of ribosomes. The assumption that low relative abundance is a reliable indicator of a special regulatory status is unjustified without knowledge of the mechanism of action of this factor. Given that eIF4E exchanges between mRNAs, there is certainly no pressing theoretical argument for this proposition. Moreover, the earlier data on eIF4E abundance in mammalian systems, which were originally interpreted as evidence that this factor is limiting, could not be confirmed in more recent work (453). As for S. cerevisiae, studies of the influence of variations in cellular levels of eIF4E on translation have revealed no evidence of the proposed “limiting” role (314, 569).

Another type of result taken to indicate a special limiting status for eIF4E is the observation that its overproduction in various cultured cells leads to effects such as cellular transformation or changes in cell morphology (115, 319). However, such transformation effects are also seen after the overproduction of other proteins in higher cells, including a nonphosphorylatable mutant form of human eIF2 (138) and hepatitis C virus proteins including the nonstructural protein NS3 (483). In none of these cases is the causal link to transformation characterized, and even if it were, it would not necessarily confer a special “limiting status” on the overproduced molecule. Again, studies with S. cerevisiae have revealed no analogous effects of overproduction of eIF4E: no measurable change in growth is seen unless eIF4E abundance is increased by a factor of approximately 100-fold compared to the wild-type, when a slight reduction in growth rate is observed (314). As in mammalian cells, the cause of the anomalous behavior of yeast cells associated with eIF4E overproduction is not certain. A favored explanation for higher cells is that the excess eIF4E potentiates the capacity of eIF4F-supported unwinding activity that releases translational attenuation of mRNAs with structured leaders which encode proteins involved in growth regulation (483, 484). It has also been reported that eIF4E overproduction can influence the nucleocytoplasmic transport of cyclin D1 mRNA in NIH 3T3 cells (465). The principle of selective translational enhancement does not apply to S. cerevisiae (569) and has in fact not been found to apply to all highly structured eukaryotic mRNAs (51). Overall, while translation-related effects may be involved, other mechanisms have not been ruled out. Since eIF4E is both nuclear and cytoplasmic (314, 325), these could potentially be linked to a wide range of processes (447). In conclusion, it seems likely that eIF4E is present in quantities that are well matched to its role in translation (314, 569). Indeed, the yeast cell shows remarkably low sensitivity to variations in eIF4E function, and even large reductions in cap-binding affinity caused by deletion mutations can be partially or fully compensated for by overproduction of the resulting truncated versions of this protein (542). On the other hand, in vitro (172) and in vivo (35) data suggest that heat shock mRNAs may be subject to an eIF4E-related selective response under stress conditions, although the magnitude and mechanism (see also the section on prokaryotic and eukaryotic translation) of such an effect are uncertain.

The proposal that the level of eIF4E phosphorylation is related to its cellular activity is derived primarily from reported correlations between its phosphorylation status and translation rates under different cellular growth conditions (218). Direct experimental evidence is thin on the ground. One study has indicated that the affinity of mammalian eIF4E for the cap structure is enhanced approximately threefold by phosphorylation (375), whereas other work has thrown doubt on the concept that phosphorylation controls the incorporation of eIF4E into the cap-bound eIF4F complex (418). A further controversial issue has been the actual site of phosphorylation on eIF4E. This was originally thought to be Ser53 in the mammalian factor, because mutation of this residue to alanine nullified various properties of wild-type eIF4E (115, 274, 319, 470). However, there have been conflicting data on the effects of the Ser-to-Ala mutation at position 58 (280). Moreover, subsequent biochemical analyses of the mammalian protein have demonstrated that the true major site of phosphorylation is Ser209 (154, 273). Parallel studies of the S. cerevisiae factor, which lacks the C-terminal serine, revealed that phosphorylation is generally weak, apparently not regulated in response to changes in cellular growth conditions, and located primarily at the N-terminal sites Ser2 and Ser15 (613). Both of these sites can be mutated to alanine with little effect on cellular growth or protein synthesis (613). However, in S. pombe, eIF4E does possess the Ser209 site (448), although current evidence indicates that S. pombe eIF4E is poorly phosphorylated (151). It must therefore be concluded that the question whether eIF4E phosphorylation provides the basis for a physiologically significant regulatory mechanism remains unresolved.

4E-BPs. A set of mammalian 4E-BPs which can act as regulators of eIF4E function have been identified. For example, treatment of adipose cells with insulin leads to enhanced phosphorylation of rat PHAS-I (phosphorylated heat- and acid-stable protein regulated by insulin [one example of a 4E-BP] [42, 235]), which correlates with enhanced translation (328, 329, 421). Since it could be shown that 4E-BPs compete with eIF4G for binding to eIF4E (195), a model was proposed in which eIF4G-eIF4E binding is regulated by the eIF4E–4E-BP interaction, which in turn is subject to modulation by phosphorylation of the 4E-BP. Serum, growth factors, hormones, and stress can all modulate 4E-BP binding via a number of signal transduction pathways (172a, 508, 576a). There are se-
quence similarities between the eIF4E-binding domain of eIF4G (Fig. 4) and a region in PHAS-I and other 4E-BPs (346, 508).

The comparison with yeast is again an interesting exercise. Analogously to the 4E-BPs, the so-called p20 protein of *S. cerevisiae* is subject to levels of phosphorylation that vary in response to changes in growth conditions (613). The N terminus of p20 also carries a sequence resembling the motif found in both forms of *S. cerevisiae* eIF4G that is thought to comprise an essential part of the eIF4E binding site (15). These results are suggestive of a mechanism of regulation analogous to that of the 4E-BPs. It follows that the absence of p20 should allow the cap complex to achieve higher activities. However, under standard laboratory growth conditions, the observed effects have been small. One group reported a small increase in the growth rate of *S. cerevisiae* in rich medium associated with disruption of the gene encoding p20 (CAF20) (15, 315), while another group saw no significant change (119). On the other hand, the latter group found that a CAF20 disruption partially suppresses growth defects associated with mutations in other initiation factor genes. There is agreement that overexpression suppresses growth defects associated with mutations in other helicases belonging to the DEAD-box family (119). The latter rationalizes the limited inhibitory capacity of p20, we still have to explain the function of a regulatory protein which, at least in yeast, acts like a yeast 4E-BP.

A more recent study has provided information relevant to the relatively limited capacity of p20 to compete with eIF4G for binding to eIF4E in vivo (449). Surface plasmon resonance analysis was used to compare the eIF4E-binding affinities of p20 and the eIF4E-binding domain of eIF4G, p20 was estimated to bind to eIF4E with an approximately 10-fold-lower affinity ($K_d = 10^{-8}$ liter mol$^{-1}$) than it bound to the eIF4E-binding domain of eIF4G ($K_d = 10^{-9}$ liter mol$^{-1}$). A further factor in the p20 regulatory equation is the fact that the control coefficient ($Z_r$ [see above]) for eIF4E activity in terms of the translation rate is likely to be considerably smaller than 1 (569), which therefore constrains the impact on translation that can be expected to be achieved by regulating the cap-binding interaction in this way. At the same time, p20 is present at relatively low levels in the yeast cell, partly as the result of rapid degradation of this unstable protein. Having rationalized the limited inhibitory capacity of p20, we still have to explain the function of a regulatory protein which, at least under laboratory conditions, seems to exert a strong effect. It is possible that it normally acts as a fine regulator and/or that it is more potent under limiting growth conditions that are more similar to those of the native environment of natural yeast strains. There are clearly interesting issues of quantitative control here that remain to be addressed in future work. Moreover, alternative leads may have to be followed up. For example, apart from the N-terminal motif resembling part of eIF4G mentioned above, p20 also has a site of homology (amino acids 26 to 31) to eIF4E (amino acids 73 to 78). It is not yet known whether this is significant.

Are there other modulators of eIF4E function in *S. cerevisiae*? A genetic screen for extragenic suppressors of a temperature-sensitive mutant allele of *cdc35* yielded two genes, *DED1* and *DBP3*, both encoding putative ATP-dependent RNA helicases belonging to the DEAD-box family (119). The latter gene suppressed the mutant phenotype only when expressed from a multicopy plasmid. However, neither of the encoded proteins is known to be a direct modulator of eIF4E. The mutant Ded1p was found to cause distortion of cellular polysomal gradient profiles typical for mutant cells bearing a mutation in the translational initiation pathway (91, 119). It can be speculated that the identified helicase activities somehow compensate for the partial loss of eIF4A and eIF4B helicase activity that would otherwise be functionally linked to translational initiation via eIF4E. The possible existence of such a functional linkage now needs to be investigated. The complexity of helicase function is, however, underlined by the fact that *DED1* has previously been identified as a suppressor of a gene defect in nuclear pre-mRNA splicing (269). Moreover, overexpression of the p20 gene (CAF20) strongly inhibited the growth of a ded1 mutant strain (119). These and other data are possibly telling us that we have to look at the pathway leading up to translational initiation on cytoplasmic mRNA to find the functional links between eIF4E, helicases, and p20. We are also reminded that the networks of genetic and biochemical interactions seen in such studies are indicative of the existence of finely balanced functional overlaps in yeast (as shown, for example, in Fig. 15).

Finally, while possible mechanisms of regulation of yeast eIF4E are suggested by analogy to mammalian cells, it is still not clear whether this route is used to mediate a cellular response to environmental signals. It has been proposed that the TOR (target of rapamycin)-containing signal transduction pathway (308) regulates eIF4E activity via modulation of this factor’s phosphorylation status (33). However, the authors present no evidence of any such change in eIF4E activity. In cells of higher eukaryotes, rapamycin-induced inhibition of translation seems to be mediated via 4E-BP1 (46). While rapamycin causes large reductions in the general rate of protein synthesis in *S. cerevisiae*, nitrogen deprivation has been found to repress expression of the cyclin-encoding *CLN3* mRNA more specifically (164). The resulting reduction in Cln3 production is possibly one of a number of routes leading to arrest of the cell cycle in G1 phase. Both decreased *CLN3* mRNA translation and increased turnover of Cln3 protein are involved (164). However, Gallego et al. (164) found no evidence of specific translational repression mediated by the *CLN3* 5′UTR, while concluding that the effect is independent of the TOR pathway. This contrasts with previous results relating to uORF-mediated inhibition of *CLN3* translation (441). It is clear that there is much to be learned about the role of posttranscriptional control mechanisms in the cell cycle. Indeed, the Trachsel and Altman group has reported that the addition of rapamycin to *S. cerevisiae* causes degradation of eIF4G without affecting the abundance (or presumably the activity) of eIF4E (51a).

**eIF2.** While a convincing case for physiologically meaningful regulation of eIF4E via phosphorylation has yet to be presented, there is little doubt about the role of eIF2 phosphorylation in regulating protein synthesis. The key players in this form of regulation, apart from eIF2 itself, are the kinases (and phosphatases) that act upon eIF2 and the guanine nucleotide exchange factor eIF2B. eIF2 is a heterotrimeric complex whose α subunit is phosphorylated by three known eukaryotic kinases: HRI and PKR in mammals and GCN2 kinase in *S. cerevisiae* (see references 96, 226, 227, 359, and 552 for recent reviews). The phosphorylation event prevents eIF2 from undergoing GDP-GTP exchange on eIF2B. Moreover, since the phosphorylated form has a greatly increased affinity for eIF2B, it acts as an effective competitive inhibitor of the cellular regeneration of active eIF2–GTP–Met-tRNAi. The essential details of eIF2 phosphorylation and their relationship to GCN4 regulation are summarized above in the section on reinitiation, while comprehensive analyses of the literature in this area are available elsewhere (96, 142, 226). Two points remain to be mentioned here. First, given the considerable number of
uORF-bearing mRNAs in *S. cerevisiae*, there is a significant probability that GCN4 is not the only case of posttranscriptional control mediated via the status of eIF2 phosphorylation, whereby the regulatory effects in other systems may turn out to be less marked. Second, the extent to which eIF2-mediated regulation of overall translation rates features as a response by yeast to stress conditions is unclear. This situation contrasts with the range of eIF2-mediated regulatory phenomena being studied in higher eukaryotes (96).

**Further gene-specific regulatory systems linked to Gcn4p activity.** The regulation of GCN4, mediated via eIF2 activity, is also coupled to other regulatory circuits by virtue of the transcriptional activation function of Gcn4p. One example of this was provided by a study of the expression of *GCD5* (otherwise known as *KRS7*), which encodes lysyl-tRNA synthetase (316). The GCD5 gene has a consensus Gcn4p-binding site (TGACTC box) (378), and it was shown that GCD5 transcription is Gcn4p dependent (316). Analysis of the effects of a point mutation located in the lysine-binding domain of this synthetase led to the proposal of an autoregulatory model. According to this model, a reduced level of tRNA<sup>lys</sup> charging by the synthetase activates GCN2 kinase and thereby induces GCN4, the consequence of which is activation of GCD5 transcription. The resulting increased levels of synthetase stimulate tRNA<sup>lys</sup> charging, thus providing negative feedback within the regulatory loop (316). This coupling between the translational regulation of GCN4 and the transcriptional regulation of GCD5 may also apply to other genes encoding aminoacyl-tRNA synthetases.

**Gene-Specific Regulation via trans-Acting Factors**

**Regulation via RNA-binding proteins.** Studies of gene expression in *E. coli* and its bacteriophages provided the first examples of translational regulation mediated by RNA-binding proteins targeted to specific motifs in the mRNA (363). The majority of the prokaryotic regulatory circuits are of a negative type, whereby binding of the effector interferes with the 30S subunit-TIR interaction. An alternative to this competitive binding mechanism is exemplified by *E. coli* S15, whose binding to its own TIR stabilizes a pseudoknot structure. The latter structure does not prevent binding of the 30S ribosomal subunit but, rather, acts to “trap” it in an SD region-associated complex which cannot proceed to polypeptide initiation (437). Since RNA-binding proteins can alter local mRNA conformations by binding to adjacent regions of the mRNA sequence, this means that access of the 30S ribosomal subunit to binding sites in an mRNA can also be positively influenced by a regulatory protein (16, 363, 600). Functional equivalents to the negative types of prokaryotic regulation are more easily imagined for eukaryotic translation than are positive-type mechanisms, and the characterized systems have so far been consistent with the expected pattern (364). Some of the principles of RNA-protein recognition in these prokaryotic and eukaryotic systems are becoming evident (see, for example, references 27, 48, 73, 139, 282, 360, 393, 495, 563, 564, and 568).

Although this review focuses on posttranscriptional control in yeast, it is appropriate to begin a consideration of gene-specific regulation via RNA-binding proteins with the higher eukaryotic system that has been most extensively characterized to date. The iron-dependent regulation of mRNAs containing iron-responsive elements (IREs) in vertebrates and insects involves the binding of the iron-regulatory proteins (IRPs) (289, 464). There are IREs located in the 5′UTRs of the mRNAs encoding the vertebrate ferritins, the erythroid form of δ-aminolevulinic acid synthase, and succinate dehydrogenase subunit b. The 5′ location of an IRE allows translational regulation to be achieved via binding of an IRP, since this protein has a sufficiently high affinity for IREs (*Kd* = 10<sup>−10</sup> to 10<sup>−11</sup> at low iron concentrations) to form a tight complex capable of blocking initiation by vertebrate 40S ribosomal subunits (464). The IRP-IRE affinity is reduced 50- to 100-fold in the presence of iron levels in excess of the requirements of the vertebrate cell. This range of iron-related affinity change is sufficient to mediate the required iron-sensitive regulation. The other location of IREs is the 3′UTR of the mRNA encoding the vertebrate transferrin receptor, whereby IRP binding at this site regulates the stability, rather than the translation, of the target mRNA. Translational regulation of vertebrate mRNAs mediated by an IRP is subject to a strong position effect. Analogously to the position dependence of inhibition via a stem-loop structure, cap proximity is a requirement for effective translational inhibition via IRP. Increasing the distance of the IRE from the cap greatly reduces the repressive effect of IRP binding (179).

The above work demonstrated that a translational repressor can function by targeting the 5′UTR of the mRNA of a higher eukaryote. However, the situation in yeast is less clear. Evidence has been presented that is consistent with translational regulation of the mRNAs encoding catalase (197) and the ribosomal protein L32 (109) in *S. cerevisiae*. However, in neither case has the regulatory mechanism been defined, and the results reported were not indicative of the operation of a high degree of control. Moreover, the regulatory role of L32 seems to be multifunctional, with recent evidence indicating that it also inhibits pre-35S rRNA processing and the splicing of its own pre-mRNA (574). Analysis of the expression patterns of other yeast ribosomal protein genes has also provided unequivocal evidence for translational regulation (263). On the other hand, despite the relative paucity of information about endogenous translational regulation systems in yeast, recent analyses of regulatory components “imported” from other organisms into *S. cerevisiae* have been informative. Coupling an IRE-containing 5′UTR to a reporter gene in *S. cerevisiae* renders the encoded mRNA subject to translational regulation by IRP (413). Translational repression of the reporter mRNA can be achieved by adding recombinant IRP to a yeast cell extract or by expressing the IRP gene on a second plasmid in vivo. This work established that an IRP and an IRE are sufficient components to create a regulatory circuit and that this two-component system is fully active in a foreign cellular environment. Further studies were performed to investigate the regulatory properties of quite different RNA-binding proteins in yeast and higher cells. It turns out that the RNA-binding protein targeted to the 5′UTR need not normally fulfill the function of a eukaryotic translational repressor to be able to block eukaryotic translational initiation. Thus, the insertion of the respective binding motifs of the bacteriophage MS2 coat protein and the spliceosomal protein U1A into the 5′UTR of a reporter gene allowed both of these proteins to act effectively as repressors when expressed in yeast and HeLa cells (529). This indicates that in this context, it is the ability of the chosen protein to bind specifically to its corresponding target motif rather than its evolved cellular function that determines whether it is capable of repressing translation successfully. Clearly, in their original cellular environments, these RNA-binding proteins have evolved to provide specific regulatory responses under defined conditions.

A further study in *S. cerevisiae* has also contributed to our understanding of the mechanism of translational repression (290). Unlike IRP-mediated repression in cells of higher eukaryotes, the distance of the target IRE from the cap could be extended to 59 nucleotides without affecting repression in
yeast. However, other experiments indicated that the relationship between IRP binding and the degree of translational repression in yeast are comparable to those seen in the vertebrate cellular environment. In particular, a range of IRE mutations that reduced IRP binding affinity by up to 160-fold manifested a progressive attenuation in the degree of translational repression, reaching almost complete abolition of the inhibitory effect (290). This mirrors the quantitative range of iron-induced adjustment of IRP-IRE affinity in mammalian cells. It was not possible to mimic the regulation of IRP binding affinity by changing the iron concentration in yeast, most probably because the internal iron levels in this organism do not respond in the same way as vertebrate cells to changes in external iron concentrations. Overall, the distinctions in the position dependence of repression between yeast and higher eukaryotic cells are indications of differences between the translational machineries of the respective organisms. One testable model explaining this is that the binding and scanning phases (Fig. 10, \( k_b \) and \( k_s \)) of the respective higher- and lower-eukaryote 40S ribosomal subunits are driven by different thermodynamic forces. Most importantly, the scanning of a vertebrate 40S subunit may be coupled to a greater thermodynamic force than is its yeast counterpart, thus explaining the ability of the former to overcome a cap-distal IRP-IRE complex. We can conclude that translational regulation can be imposed on yeast mRNAs by the binding of RNA-binding proteins to specific motifs located in the 5′UTR. However, the parameters governing this regulation are different in yeast, reflecting key properties of its translational apparatus. It will be important to explore the general validity of these principles in whatever endogenous translational repression systems are found to function in this organism.

Regulating the expression of dsRNA genomes. The genome of the cytoplasmic dsRNA viral particles frequently found in S. cerevisiae is transcribed to generate a single-stranded (plus-strand) mRNA that serves as the template for translation of the gag and pol genes (see the section on frameshifting). This mRNA is neither capped nor polyadenylated (70, 547). The lack of the 5′ and 3′ mRNA modifications means that the viral mRNA is potentially disadvantageous in terms of translation and also is at risk of being rapidly degraded by cellular 5′→3′ exoribonuclease activities (see the section on mRNA decay). Indeed, exoribonuclease-1 is encoded by exoribonuclease activities (see the section on mRNA decay). It has accordingly been postulated that Gag effectively generates decapped mRNAs that seem to decoy Xrn1p away from at least some of the viral uncapped mRNAs (357), allowing them to survive longer in the yeast cytoplasm. Given that the viral mRNAs are not delivered to the cytoplasm through the nuclear membrane, it would be interesting to know whether this affects their “presentation” to either the degradation machinery or the translational apparatus. Perhaps differences in the compartmentalization and/or kinetic control of the decay or translation of viral mRNAs contributes to the viral survival strategy. Other genes involved in controlling or regulating dsRNA viruses (such as the MAK genes [23]) are discussed by Wickner (592).

Gene-specific regulation via antisense RNA in yeast? Studies of plasmid replication in E. coli demonstrated the principle of antisense regulation more than 17 years ago (251, 310). Since then, a whole range of cases of specific antisense RNA-mediated regulation has been uncovered in prokaryotic systems (244, 502, 576b). In general, prokaryotic antisense RNAs are between 70 and 110 nucleotides long and are capable of forming stem-loop structures with apical loops of 6 to 8 nucleotides. The stem regions are thought to confer both a clearly defined structure and stability on the RNAs, while the loop nucleotides are believed to be important for the initial interactions with the target molecules. The binding of prokaryotic antisense RNAs can modulate a number of different target functions, including RNA processing, translational initiation, transcriptional termination, and mRNA stability (502). The prokaryotic work therefore demonstrates that antisense RNA can act as a versatile modulator of gene expression.

In contrast, the role of antisense RNA in eukaryotic cells is much less clear. There are numerous examples of eukaryotic genes that are transcribed on both strands, meaning that potential antisense RNAs are undoubtedly generated in the cell. However, unequivocal evidence for natural antisense regulation is scarce. Despite this, artificial antisense RNA constructions have been used successfully in both plant and animal cells, although the mechanisms underlying the inhibitory effects observed have received little attention. The approach used has generally been based on simple empirical principles or trial and error.

S. cerevisiae is a particularly poor host for experiments of this type (23, 26), yielding only the occasional success (400). The apparent ineffectiveness of antisense RNA strategies in S. cerevisiae has led to the suggestion that at least this yeast has characteristics that somehow foil the standard procedures (26). S. pombe might be more amenable to this type of imposed gene suppression (23). However, it should be emphasized that there...
has been little systematic investigation of the possible mechanistic pathways that might lead to antisense RNA suppression in eukaryotic cells in general and in yeast in particular. The standard approach involves the generation of RNAs that are complementary to smaller or larger regions of the target mRNA via transcription of the second strand. However, comparison of the natural prokaryotic systems suggests that this strategy is unlikely to provide optimal conditions for sense–antisense regulation because it ignores the influence of antisense RNA structure on stability and the kinetics of antisense–sense interactions. Consistent with this view, recent studies with translation cell extracts derived from *S. cerevisiae* have shown that at least in vitro, a modified form of the E. coli IS10 RNA<sub>AUT</sub>/RNA<sub>IN</sub> regulatory system can be used to regulate translational initiation on a eukaryotic mRNA in a yeast environment (28). By targeting the RNA<sub>IN</sub> sequence inserted upstream of a reporter gene, it could be shown that an antisense RNA can block yeast translation via interactions with a 5′UTR target site. This suggests that the prokaryotic type of specifically targeted antisense regulation may be possible in yeast, given the appropriate set of conditions. It follows, therefore, that it may be possible to establish a set of rules that allow the design of reliable antisense RNA constructs capable of imposing regulation on selected eukaryotic (yeast) genes. However, this will depend on further characterization of the mechanistic principles required for antisense RNA regulation. At the same time, we may learn the extent to which antisense-RNA-mediated regulation could, at least theoretically, offer an alternative route to protein-mediated regulation.

**CONTROL OF mRNA DECAY**

Much of this review has been concerned with the functional lives of mRNA molecules, but these are all eventually terminated by enzyme-catalyzed degradation reactions that lead to functional inactivation and then complete hydrolysis. This turnover process would be of limited interest if it were not for the fact that mRNA decay is subject to a number of controlling influences. As a result, the estimated physical half-lives of mRNA in *S. cerevisiae* vary from less than 1 min to over 60 min (215, 267). This means that mRNA decay contributes to the differential control of gene expression. Moreover, there are strong indications that at least some mRNA turnover rates are regulated in response to environmental changes. Specialized reviews on eukaryotic mRNA degradation have been published at regular intervals (the more recent ones include references 78, 266, 352, 353, 454, 462, 463, 468, and 546), and the reader is directed to these for comprehensive coverage of the relevant literature. There is increasing interest in the role of mRNA degradation in the yeast mitochondrion (355). Information about the various methods used to study mRNA decay in yeast is available elsewhere (29, 78, 215).

The main objectives of this section are to summarize the basic principles that have emerged from studies of nonaberrant mRNA decay in the cytoplasm of *S. cerevisiae* and to consider how these fit into the overall picture of posttranscriptional control in this organism, especially in relation to translation. mRNA decay can be conveniently considered in terms of *cis*-acting mRNA determinants, *trans*-acting factors, the order and causality of decay events, and the potential mechanisms of regulation, and these themes will provide the framework for this brief overview.

**cis-Acting mRNA Determinants**

Stability determinants in the main ORF and 3′UTR. One of the major unanswered questions in the study of mRNA decay is how the specific decay rates of individual mRNAs are determined. Experimental analysis of a number of decaying mRNAs extracted from cells in which transcription has been blocked reveals that, in general, whereas full-length or near-full-length mRNAs are readily detectable, smaller intermediate products generated by endo- and exonucleolytic cleavages of natural mRNAs are comparatively short-lived. This indicates that degradation, once triggered by an early event that does not greatly change the overall structural integrity of the mRNA, is rapidly completed. Control over the observed stabilities of different mRNAs is therefore exercised early in the overall pathway. It is also important to note that the stabilities of eukaryotic mRNAs are generally defined in terms of the half-lives of physically intact mRNAs comprising their full DNA-encoded sequences, rather than in terms of functional inactivation. By analogy to the protective effects of ribosomes and RNA-binding proteins on prokaryotic mRNA (80, 245), it might be expected that the stability of eukaryotic mRNA should be related to the accessibility of potential cleavage sites in each mRNA molecule. In an extreme model, these sites would be highly redundant and would therefore be frequently represented in all mRNAs. Such a model would predict that the number of potentially accessible sites, and the relative periods of their exposure, would be influenced by the density of ribosomal loading. However, general correlations between mRNA stability and features likely to affect the number and proportion of mRNA sites occupied by ribosomes at any one time, including transcript length, codon usage, and the kinetics of translational initiation and elongation, are not identifiable (78, 215, 438, 482, 570). This in itself does not rule out any participation of such determinants in the turnover of various mRNAs, but it does suggest that other features are responsible for controlling the rates of decay. As will become apparent, much less is known about key rate-controlling features in mRNA decay than in translation. The acid test for this is to ask whether it is possible to make an educated guess at the likely stability of an mRNA on the basis of knowledge of its sequence. This will generally not be possible, whereas a number of clues to the translation efficiency of an mRNA are readily identifiable.

The poly(A) tail is a general feature of PolII mRNAs and is thought to play a general role in controlling the onset of decay in many mRNAs (see the section on decay pathways). Other stability determinants exist in various forms and act at a number of different locations within the mRNA (Fig. 16). In most cases, their influence is position dependent. The analysis of a small number of relatively unstable mRNAs (HIS3, MATα1, *MFα2*, and *STE3*) has revealed the existence of segments of the coding region or the 3′UTR that are capable of inducing rapid decay. “Cut-and-paste” experiments and deletion analyses have been used to achieve approximate delineation of the *HIS3* and *STE3* destabilizing regions (211, 213, 267). The relatively compact *MATα1* element has been characterized to a considerably greater degree than any of the other putative internal stability elements. It has been defined as an entity comprising 65 nucleotides, of which the first 33 are rich in rare codons and the C-terminal region is AU rich. Destabilization by this element seems to require translation through it, since the introduction of a stop codon 5′ of it stabilizes the *MATα1* mRNA two- to threefold (213). It has been proposed that the rare-codon region helps decelerate translating ribosomes, which then interact somehow with the AU-rich region, perhaps...
via an unknown factor. It is not known how this proposed mode of action then leads to destabilization. The 3' UTR can also contain destabilizing elements. The 3' UTR of STE3 was found to be capable of destabilizing a truncated PGK1 mRNA (211), while mutations in the MFA2 3' UTR resulted in stabilization of the MFA2 mRNA (388).

It is not clear whether there are any common features shared by these destabilizing elements. No mechanisms of action have been characterized, but certain consequences in terms of the decay process have been reported. Deadenylation and decapping were found to be accelerated in the presence of the MFA2 and MATa1 destabilizing elements (79, 389). Both events feature prominently in the proposed pathways of yeast mRNA decay (see below). Also unclear is the extent to which the potential decay behavior of a given mRNA can be defined in terms of the identified determinants alone. For example, the inactivation or removal of the MATa1 or MFA2 elements generates mRNAs with only intermediate stabilities (213). Such results suggest that other, as yet undefined, elements also contribute to the stability of each mRNA. Indeed, taking this further, could it be that a combination of structural attributes, perhaps even spread over the whole mRNA, determines stability? If so, the determination of mRNA stability may generally be a complex function of relatively large domains of molecular structure. This model would explain why at least some elements are only fully functional in certain mRNA environments (211). An alternative explanation of such context effects could be the existence of specific stabilizing sequences which, when combined with destabilizing elements, might neutralize the effects of the latter elements. However, until it can be proved that a range of appropriately placed, discrete stabilizing and/or destabilizing elements exist, there is little reason to accept the latter model. Relevant to this point is the observation that the PGK1 mRNA requires translation to maintain its

FIG. 16. Features of yeast mRNAs influencing stability. The major types of mRNA stability determinant reported so far for S. cerevisiae are shown. The majority of these elements act to destabilize the mRNAs in which they have been studied. Internal stop codons (here indicated as UAA) can cause strong destabilization in 5'-proximal positions (strong) but have less or no effect at more distal positions (weak). Apart from the cap or poly(A) tail, it is not clear whether discrete (and transferable) stabilizing elements exist.
high stability. It remains to be seen whether this behavior is attributable to the combined action of a number of discrete stability elements or to the overall structure of much of the mRNA.

Modulation of decay via the 5'UTR. The 5'UTR can be home to a number of structural elements that influence stability (Fig. 16). The effects of many of these elements depend on the nature of the mRNA downstream of their location in the sequence. The most generally relevant feature is the cap structure, which is recognized to fulfill a protective function in terms of 5'→3' exonuclease activity (78, 162). Other features, such as a stem-loop, a poly(G) stretch, or the formation of a protein-target complex in the 5'UTR, affect the stabilities of only certain types of mRNA (37, 336, 390, 482, 570, 576). They are thought to exert their influence in two ways.

First, by inhibiting translation, they may change the activity of (as yet undefined) stability determinants within the body of the mRNA (see, for example, references 76, 213, 266, and 336). Consistent with this is the observation that the degree of destabilization of the PGK1 mRNA caused by a stem-loop or IRP-IRE interaction in the 5'UTR correlates with the degree of translational inhibition imposed on the mRNA by these elements (336). However, in mutant strains whose translational apparatus is dependent on partially inactivated eIF4E proteins, the reduction in translation capacity did not correlate with destabilisation of PGK1 mRNA. This suggests that it is the disruption of the initiation process on the mRNA, rather than the reduction in the frequency of translational initiation per se, that controls the destabilization potential of the PGK1 stability determinants. Blocking the scanning process may disrupt intramolecular interactions that disturb the formation of translation-competent polysomes in a way that is distinct from the effects of reducing the frequency of ribosome-mRNA interactions via eIF4E. Alternatively, it has not been ruled out that such mutations in eIF4E affect events prior to translation, such as mRNA transport or compartmentalisation. It should, however, be pointed out that the PGK1 mRNA may be atypical in its response to elements in the 5'UTR. Stem-loop structures have no effect on the stabilities of MEA2, YAP1, or cat mRNAs in yeast (37, 336, 482, 570, 572), and they stabilize the LUC mRNA (336). These differences are likely to reflect the distinct parameters dictating the decay behavior of the respective mRNAs.

The second effect of structural elements in the 5'UTR is to inhibit directly the progress of exonuclease activities. Poly(G) stretches and, to a lesser extent, hairpin loops of stabilities sufficient to cause strong translational inhibition (−10 to −20 kcal mol−1), act to impede 5'→3' exonucleolytic degradation (390, 576). The balance between the contributions of the two types of effect discussed above will be a function of the presence or absence of stability elements and the decay pathways of individual mRNAs. This complicates the interpretation of the observed influence of structure in the 5'UTR on stability.

Further experimental data that are consistent with there being, in broad terms, a relationship between translation and mRNA decay have been obtained by imposing general blocks on translation. Cycloheximide (215, 429) or a defective tRNA nucleotidyltransferase (429) inhibits translational elongation, thereby increasing the average size of cellular polysomes. This has been found to stabilize a number of mRNAs. However, it is difficult to interpret these results reliably in terms of protective interactions between the translational apparatus and the degradation machinery on individual mRNAs because of the indiscriminate nature of the inhibition caused by these measures. Such general attenuation of protein synthesis may lead to rapid distortion of the balance of cellular components that directly or indirectly affect mRNA stability. It should also be pointed out that verrucarin A has also been found to stabilize certain mRNAs even though it greatly reduces polysome sizes (532). As discussed above, another means of inhibiting translation is to inactivate one of the translation factors. Work with a prt1 mutant strain (defective in eIF3 [Table 1]) revealed preferential destabilization of two mRNAs (SSA1 and SSA2) encoding proteins belonging to the Hsp70 heat shock protein family (34). Cereghino et al. (84) found that inactivation of prt1 led to destabilization of SDH2 mRNA but not of ACT1 or CUP1 mRNA. Such differential effects are presumably due to the influence of specific elements or structural characteristics in the respective mRNAs. However, the responses mediated by such elements may differ according to the step of translation that is disrupted. For example, it is not known whether defects in eIF4E (compare reference 336) are capable of generating comparable differential destabilization to that seen with the prt1 mutant. Mutants with mutations in GCD2 or SUI2 were found not to destabilize SDH2 mRNA (84). In contrast, a single amino acid substitution in eIF5A has been found to stabilize a number of mRNAs, causing the host cell to accumulate uncapped mRNAs (620).

An uORF constitutes a quite different type of 5'UTR element. The presence of an uORF can, but need not, lead to accelerated degradation of an mRNA (411, 469, 572). The experiments of Oliveira and McCarthy (411) showed that the already relatively unstable cat mRNA can be destabilized by approximately a factor of 4 as a result of a single nucleotide change (AAG to AUG) in the 5'UTR that generates a seven-codon uORF. Given the foreign nature of the cat mRNA, this and other results suggested that natural yeast stability elements may not be required for the destabilization process. Investigations of natural uORFs have revealed a more complex picture. The PPR1 5'UTR has been identified as having transferable destabilizing properties (438). It contains a six-codon uORF that overlaps +1 at its 3' end (AUAAUG) with the start codon of the main ORF. Fusion of this leader in the same configuration with the PGK1 ORF generates a highly unstable mRNA, while the effect is nullified if the AUG codons at positions 1 and 2 of the uORF are mutated to AAG (336). This suggests that the stability determinant of the PPR1 5'UTR that is capable of destabilizing the PGK1 mRNA includes this natural overlapping uORF. This contrasts with the earlier results of Pierrat et al. (438), who observed no change in the stability of the unfused PPR1 mRNA when the two AUGs in the overlapping uORF were mutated to AGGs (438). Resolution of this apparent discrepancy will require further analysis of the role of the PPR1 coding region in controlling decay.

Two natural 5'UTRs containing nonoverlapping uORFs that were considered in the earlier section on reinitiation have also been subject to recent investigations: the GCN4 leader, which normally has no detectable effect on the stability of its mRNA, and the YAP2 leader, which naturally destabilizes its own mRNA. It has been proposed that the GCN4 leader does not destabilize mRNA because it lacks a downstream element of the type previously reported to be required for nonsense-dependent destabilization in mRNAs carrying nonsense codons (469). To be able to assess this potential relationship between premature termination and the role of uORFs, it is necessary to briefly consider the current models that attempt to explain how the premature stop codons are dealt with in the cell.

The destabilizing effect of internal nonsense codons in mRNA was observed first in E. coli (229) and then somewhat later in yeast (342). Nonsense codon-dependent accelerated decay has since been observed in a number of yeast mRNAs...
(266, 322, 323, 427, 611, 616), and unspliced pre-mRNAs (209), but most investigations of this phenomenon have been performed with the PGK1 mRNA. Destabilization of the PGK1 mRNA was initially reported to be dependent on the presence of a downstream element comprising some 80 nucleotides 3′ of a premature termination codon (430). Two AUGs in this element were originally found to be required for destabilization. However, a later report concluded that the critical element is the sequence motif UGYYGAUGYYYYY (616), and the most recent analysis concludes that the AUG in this motif is not required for activity (469).

This leaves a highly degenerate 13-nucleotide pyrimidine-rich region as the proposed motif. Moreover, this motif was found to be nonfunctional in the absence of flanking sequences (616), and therefore the extent to which these additional regions dictate the stability of the mRNA is not clear. The motif was also found not to be necessary for nonsense codon-dependent destabilization in at least one PGK1 construct (430). Sections of more than 100 nucleotides containing regions resembling the motif were taken from the ADE3 and HIS4 genes and tested for their ability to promote nonsense codon-dependent destabilization of a “mini-PGK1” gene that carries an N-terminal amber codon but lacks most of the normal PGK1 coding region (616). These regions also supported accelerated decay. Models proposed for the mode of action of this motif have predicted either that it interacts with the 40S subunit or an undefined factor to promote translational reinitiation or that this interaction causes ribosomal pausing (432). The interaction with the ribosome was suggested to be mediated via the partial complementarity of the motif to a region of 18S rRNA (430), but this idea will require particularly careful testing, given the high degree of redundancy apparently permissible in the motif, the fact that the corresponding region in the 18S rRNA is predicted to be involved in intramolecular basepairing (110), and the observation that the AUG is nonessential (469). The lack of a requirement for an AUG also means that reinitiation within the motif plays no essential role. Moreover, the role of secondary structure in the mRNA downstream of nonsense codons remains uncharacterized. Whatever signals are involved, premature termination in PGK1 or a truncated version of it leads to accelerated decay via a mechanism involving the UPF gene products (see the discussion of trans-acting factors). This dependence on the UPF genes is a feature of nonsense codon-mediated decay in the mRNAs studied so far (266). However, UPF-dependence is apparently not exclusive to the decay of nonsense codon-containing mRNAs; at least two non-aberrant mRNAs that lack early nonsense codons are degraded via upf-dependent pathways (78; compare reference 33a).

There is a nonlinear gradient of sensitivity to the presence of premature stop codons in PGK1 (Fig. 16), so that stop codons in the last quarter of the reading frame have no effect, stop codons in the first 55% accelerate decay approximately 12-fold, and there is a transition in effect over the intervening region (431). This polarity effect in the destabilization potential of nonsense codons has been explained in terms of the existence of localized stabilizing sequences in the PGK1 reading frame that are proposed to modulate the function of the destabilizing motifs (431). The proposed stabilizing elements have yet to be characterized. The extent to which the secondary structure of regions containing the putative elements and the potential interactions between such regions play a role is a matter for discussion. A further matter worthy of consideration is that even upf-independent accelerated decay of the PGK1 mRNA, caused by inhibition of translational initiation, reacts very sensitively to changes in the translationrate (see above). As suggested in the previous section, this may mean that an unusually complex set of factors modulates PGK1 decay behavior. On the other hand, since the HIS4 and CYC1 mRNAs show similar position gradients of nonsense codon-dependent destabilization (193, 611), it is likely that certain generally acting principles are involved. Further research is needed to determine the extent to which the polarity effect is attributable to the specific action of localized stability determinants as opposed to more general principles of control related to the overall structure of mRNPs and polysomes (see Fig. 18).

Returning to the GCN4 mRNA, it has been proposed that termination on one of the uORFs of its leader can be coupled to upf-dependent destabilization via the same means as that which applies in the case of premature termination within a reading frame, provided that a downstream motif of the type described above is introduced 3′ of the uORF (469). It has been argued that the wild-type GCN4 mRNA lacks the elements required for destabilization (469). (For the sake of completeness, it should be noted that another group has described transient destabilization of the wild-type GCN4 mRNA following translational derepression [309]). A segment of more than 200 nucleotides derived from PGK1 was inserted 3′ of uORF1 or uORF4 to achieve destabilization. However, upon inspection, this region was found to have the potential to form stable secondary structure (including individual stem-loops with predicted stabilities of up to −24 kcal mol⁻¹) (573). It is possible that these structured sections in the inserted RNA or other, as yet undefined properties play a role in the destabilization caused by the PGK1 segment. In a later paper, the Peltz group reported that the action of the PGK1 downstream element can be inhibited by inserting a “stabilizing element” between the translation codon of the uORF and the destabilizing element (469a). The stabilizing element they describe is in fact an AU-rich segment derived from the GCN4 leader (nucleotides +498 to +565). It is of interest to consider these results in the light of the following discussion of studies with the YAP2 system.

Termination in the 5′ UTR. Since only modified forms of the GCN4 leader can destabilise mRNA, it has been unclear for some time to what extent termination-linked accelerated decay might play a role in the turnover of natural, non-aberrant mRNAs. This situation was changed by the observation that the YAP2 mRNA is subject to uORF-dependent destabilisation (572). Not only do the YAP2 uORFs inhibit translation, they also accelerate mRNA decay (Fig. 13B). This mRNA is therefore under further investigation because it throws light on the use of termination-linked mRNA destabilization as a means of imposing posttranscriptional control on non-aberrant cellular mRNAs (572, 573). A number of features of the YAP2 mRNA are of interest. First, destabilization by the uORFs of either the YAP2 mRNA or reporter mRNAs is to a great extent ufp independent. This distinguishes the YAP2 system from nonsense codon-dependent destabilization in aberrant mRNAs. Second, in this natural mRNA, there is no evidence that uORF-dependent destabilization is dependent on the presence of the CU-rich type of sequence element. This may reflect the lack of dependence on the UPF genes for decay but might also be attributable to a high level of redundancy in the types of mRNA element and structure capable of supporting destabilization. It has also been observed that nonsense codon-dependent decay of PGK1 mRNA can be triggered in the absence of the UC-rich motif (430). Third, there is a strong correlation between a combination of efficient recognition of the uORF start codon plus inhibition of downstream reinitiation and the destabilizing effect of the uORF. This was established by manipulating uORF structure and examining the consequences in
terms of translational inhibition and decay rate. One of the structural features promoting accelerated decay in the YAP2 system is the presence of a CG-rich sequence downstream of the YAP2 termination codon. This is a feature shared by GCN4 uORF4, which was also found to be capable of destabilizing both YAP mRNAs (573). A proposed explanation for this behavior is that the decay rate is influenced by ribosomal termination and release. A further piece of evidence in favor of this model is that the insertion of a stem-loop structure 3′ of the uORF leads to destabilization of the mRNA even if the uORF structure alone is not capable of causing destabilization (573). This effect is generally diminished if the stem-loop structure is located at a greater distance from the uORF. The influence of structure on uORF function seen here is also relevant to the experiments performed on the GCN4 leader (469), since the downstream elements inserted 3′ of the GCN4 uORFs are potentially structured (see above). Fourth, reinitiation (or suppression of the frequency of termination and ribosome release) on the uORF) acts to suppress destabilization. This may be linked to the ability of the ribosomal subunits to remain on the mRNA after termination, or to scan through the potentially destabilizing uORF, or could be coupled to the reinitiation process.

Given the existence of a considerable number of different uORF-containing mRNAs in S. cerevisiae (Table 3), the apparent role of uORFs in mRNA decay may be of more general relevance. The mechanism underlying uORF function in YAP2 has not yet been fully characterized, but certain properties of this leader have become apparent. Experiments with gcd and gen mutations have indicated that the activity of eIF2 in the cell influences both the reinitiation capacity of ribosomes terminating on YAP2 uORF1 and the destabilizing potential of post-termination events (573). This becomes evident when using one of the constructs mentioned above in which a YAP uORF was followed by a stem-loop structure. For example, in a gcd2 strain with reduced eIF2 activity, the destabilizing effect was enhanced, much as would be expected if fewer ribosomes had acquired reinitiation competence by the time they reached the structural impediment in the leader. This supports the notion that the effects of the YAP uORFs on translation and mRNA decay are tied into the same set of functional principles as those that govern GCN4 translation. Indeed, there are probably some further similarities between the modes of action of the GCN4 and YAP2 uORFs.

Two examples taken from the available experimental data illustrate this. First, it was found that the destabilizing effect of the PGK1 downstream element was reduced upon extending its distance from the uORF stop codon (469a). If the secondary structure of the large downstream element is important, this result is consistent with the effects of increasing distance between a stem-loop and the YAP uORFs (573). In both cases, the increased distance between the uORF and the element inserted downstream gives the ribosomes more time to reacquire (re)initiation competence, thus becoming resistant to release caused by the downstream structure. Second, the stabilizer element described by the Peltz group is an AU-rich sequence. When placed downstream of an uORF, this type of sequence promotes reinitiation and does not support uORF-related destabilisation (226, 572, 573). These considerations raise the question whether the segments tested in the GCN4 system really represent specific mRNA stability elements or, rather, are further examples of general structural features that influence the posttermination behavior of ribosomes. The latter concept would provide a potentially unifying model to explain the behavior of the respective uORF-containing systems.

It therefore remains to ask why, if the YAP2 uORF and GCN4 uORF4 are really of the same type, does GCN4 uORF4 not succeed in destabilizing its own mRNA? One possibility is that the GCN4 main ORF is relatively resistant to attack by the degradation machinery in an mRNP particle or polysome with a low ribosomal loading density and that more than termination on the uORF is required to destabilize this “protected” status. There is good reason to believe that there is significant variation in the susceptibility of different mRNAs to accelerated decay in disrupted or poorly translated polysomes (as exemplified, for example, by a comparison with PGK1). Further experimentation should be directed to testing this and the localized-stability-element model.

Various structural features can influence mRNA stability. Summarizing the current understanding of stability elements in yeast mRNAs is not a simple matter. It is not yet clear how numerous or diverse the proposed stability elements found in the main coding region or 3′ UTR of nonaberrant mRNAs really are. Of the stability elements studied so far, the MATa1 element has been most intensively studied. At least one type of downstream element associated with nonsense codon-dependent decay contains highly degenerate UC-rich tracts, but the requirement for other sequence components remains undefined. Overall, the current data do little to convince us that stability is related to the action of a small group of destabilizing elements. The challenge, therefore, is to establish whether discrete internal elements of the MATa1 type are a common feature of yeast mRNAs. Moreover, future work will also have to determine whether discrete stabilizing elements are present in yeast mRNAs. If so, it will be important to determine whether they are transferable, independently acting stabilizers, or (more degenerate) sequences and/or structures that act to suppress the function of destabilizing elements, perhaps by obscuring or affecting the structures in which the latter are involved.

Should the apparent lack of a readily discernible pattern in the roles of stability elements in yeast mRNAs give us cause for concern? Comparison with the situation in E. coli reassures us that this is nothing unusual. It seems likely that, as in the prokaryotic case, alternative structural options are available to achieve a given rate and/or course of decay and that clearly identifiable motifs or structural elements may be difficult to identify, especially within the body of the mRNA. A useful comparative example is the cleavage pattern of RNase E in E. coli. This enzyme cleaves at a large number of what appear to be only loosely related cleavage sites. Thus, recognition of sites that play a key role in controlling decay rates does not depend on the existence of highly conserved sequence motifs. While yeast mRNAs may not be subject to this type of decay, the elements controlling their decay may be at least as diverse in sequence and structure. This generally flexible strategy also seems likely to be advantageous, since it enables the cell to determine individual mRNA half-lives by using combinations of different, most probably varied and possibly delocalized structural features within the body of the mRNA (modelled in reference 80). Additionally, other layers of control can be imposed by more discrete elements, including the poly(A) tail, the 5′ cap, certain uORFs, and what so far seems likely to be a small number of internal discrete stability elements such as that in MATa1. Each of these elements exerts a relatively powerful effect as an independent unit. The flexibility of the overall system also allows mRNA turnover rates to be made either dependent or independent of translation rates. Interestingly, discrete consensus stability elements are more a feature of mammalian and plant mRNAs than of S. cerevisiae mRNAs (462, 463). It remains to be seen whether S. pombe is different in this respect.
The 5′UTR elements can be considered in relation to their effects on the translational initiation process, and on the decay process itself. Inhibitory elements like stem-loops or protein-target complexes affect the stability only of an mRNA which, like PGK1, has internal elements or overall structural features which allow it to react to translational inhibition at this step. As discussed below, the effects of the latter type of activity are dependent on the pathway of decay of each mRNA. Certain uORFs seem to be able to act generally to destabilize mRNAs and apparently do not necessarily require (translation-independent) internal elements to do this. The YAP2 type of uORF can apparently link termination to mRNA decay, and there is evidence that it acts (in combination with short flanking sequences) as a discrete, transferable agent. At the same time, it seems that other elements can also be involved. Studies of uORF-dependent destabilization may therefore provide an alternative route to characterizing at least certain forms of decay in terms of clearly defined mRNA elements. Moreover, as we have seen, further detailed study of these systems may help us identify a unifying model for 5′UTR-mediated decay that links disruption of the translation process to destabilization (see Fig. 18).

**Trans-Acting Factors**

A number of yeast mRNA degradation activities have been partially characterized (Table 4). The Stevens laboratory described a 5′→3′ exonuclease activity (175 kDa) that attacks mRNA (234, 520, 524) and an mRNA-decapping enzyme (522). The gene encoding the exonuclease was subsequently cloned and given the designation XRN1 (5′-exoribonuclease 1 [317, 523]). Later work described further genes, DCP1 (38, 311) and DCP2 (142a), which encode decapping activities. Gene disruption experiments have revealed that none of these genes is essential individually, but the resulting mutants are growth impaired and contain mRNAs with longer half-lives. There is, however, a second 5′→3′ exonuclease activity (5′-exoribonuclease 2; 116 kDa) in S. cerevisiae, which is encoded by the essential gene HKE1 (283, 521, 522, 525). The amino acid sequences of Xrn1p and Hke1p show significant similarities, and both proteins are multifunctional (5, 17, 281). Xrn1p is a cytoplasmic protein involved in a number of activities other than mRNA degradation, including DNA strand exchange and 5′→3′ exonucleolytic degradation of DNA (220, 272). Indeed, XRN1 has also been identified in different roles as DST2, KEM1, RAR5, SEP1, and SKI1 (234, 285, 287, 318, 520, 550). The functions of Hke1p, which is nuclear (272, 283), are less clear, but mutations that mislocalize it to the cytoplasm allow it to suppress the phenotype of an xrn1 deletion (272). HKE1 is identical to RAT1 (17) and TAP1 (132). The functional interchangeability is also reciprocal; targeting Xrn1p to the nucleus complements an hke1 mutant (272).

A number of proteins are now thought to influence mRNA decay via the 3′ end. An Rnase activity which might catalyze deadenylation in vivo has been detected in vitro (PAN). This activity is at least partially Pab1p dependent (344) but has yet to be characterized in any detail (473). Pab1p may be involved in controlling poly(A) tail lengths via interactions with Rna15p, which is one of the components of the yeast 3′-end RNA-processing complex CF1 (18, 377). Early work on Rna14p and Rna15p suggested that these proteins might be involved in the regulation of mRNA deadenylation (348, 376). However, other evidence points to a role in polyadenylation (18, 78, 349, 377), and therefore these two proteins can be given only very tentative positions among the cast of players controlling mRNA decay.

The most striking recent findings in this area have concerned a number of proteins that either catalyze or modulate 3′→5′ exonucleolytic degradation of mRNA. The Tollervey group described a heteropentameric complex in S. cerevisiae, named the exosome, which is required for 3′ processing of the 5.8S rRNA (379). Three of the exosome proteins, Rrp4p, Rrp41p, and Rrp44p, were shown to exhibit intrinsic 3′→5′ exonuclease activity. Rrp41p is identical to one of the Ski proteins (Skii6p) (see the section on dsRNA viruses, above). Moreover, sequence similarities were identified between Rrp44p and bacterial RNase II and between three other exosome components (Rrp4p, Rrp42p, and Rrp43p) and bacterial Rnase PH. These data suggest that the exosome is a multi-Rnase complex. A subsequent analysis revealed that mutations in SKI6/RRP41 and RRP4, as well as in three further SKI mutants (ski2, ski3, and ski9), attenuated 3′→5′ degradation of a poly(G) protected mRNA fragment derived from MFA2 or PGKI (20). These and other data suggest that exosome components are involved in 3′→5′ mRNA decay and that Ski2p, Ski3p, and Ski8p modulate this activity. Since Ski2p is a DEVH box protein, it is tempting to speculate that its involvement is somehow related to an intrinsic RNA helicase activity (20). This would be reminiscent of the bacterial “degradosome” (19, 374, 450, 584, 585).

<table>
<thead>
<tr>
<th>Enzymes involved in general decay</th>
<th>Function</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xrn1p</td>
<td>5′→3′ exonuclease (plus other functions)</td>
<td>234, 520, 524</td>
</tr>
<tr>
<td>Hke1p (Xrn2p)</td>
<td>5′→3′ exonuclease (plus other functions)</td>
<td>283, 521, 522, 525</td>
</tr>
<tr>
<td>Dcp1p and Dcp2p</td>
<td>Decapping enzymes</td>
<td>38, 311, 522</td>
</tr>
<tr>
<td>Rrp41p/38khp</td>
<td>3′→5′ exonuclease (5.8S rRNA processing)</td>
<td>20, 379</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other activities influencing decay</th>
<th>Function</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAN</td>
<td>Pab1p-dependent poly(A) nuclease</td>
<td>344, 475</td>
</tr>
<tr>
<td>Ski2p, Ski3p, Ski8p</td>
<td>Can bind other Uf proteins relatively strongly</td>
<td>104, 107, 210, 266</td>
</tr>
<tr>
<td>Mrt1p, Mrt3p</td>
<td>Predominantly nuclear (influences nuclear transport)</td>
<td>25, 266, 321, 323</td>
</tr>
</tbody>
</table>

**TABLE 4. trans-acting factors known or thought to be involved in mRNA decay in S. cerevisiae**

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xrn1p</td>
<td>5′→3′ exonuclease (plus other functions)</td>
<td>234, 520, 524</td>
</tr>
<tr>
<td>Hke1p</td>
<td>5′→3′ exonuclease (plus other functions)</td>
<td>283, 521, 522, 525</td>
</tr>
<tr>
<td>Dcp1p</td>
<td>Decapping enzymes</td>
<td>38, 311, 522</td>
</tr>
<tr>
<td>Rrp41p/38khp</td>
<td>3′→5′ exonuclease (5.8S rRNA processing)</td>
<td>20, 379</td>
</tr>
<tr>
<td>PAN</td>
<td>Pab1p-dependent poly(A) nuclease</td>
<td>344, 475</td>
</tr>
<tr>
<td>Ski2p, Ski3p, Ski8p</td>
<td>Can bind other Uf proteins relatively strongly</td>
<td>104, 107, 210, 266</td>
</tr>
<tr>
<td>Mrt1p, Mrt3p</td>
<td>Predominantly nuclear (influences nuclear transport)</td>
<td>25, 266, 321, 323</td>
</tr>
</tbody>
</table>
which comprises a helicase (RhIB), an endonuclease (RNase E), an exonuclease (polynucleotide phosphorylase, PNPase), and an enolase. In vitro data indicate that the helicase activity in this complex facilitates 3′→5′ degradation through structured RNA. Whether the association with enolase in some way links (localized) energy metabolism to the ATP-dependent functions of the degradosome remains to be seen. Future work should tell us whether the exosome constitutes a eukaryotic equivalent to the degradosome and whether its specificity is subject to modulation via other proteins, including those of the Ski type (20).

The relevance of Ski-modulated 3′→5′ decay to cellular mRNA metabolism seems to be underlined by the previous observation that mutations in SKI2 and SKI3 are synthetically lethal with deletions of XRN1 (271). One possible concern here might be that, given the multifunctionality of Xrn1p, we cannot assume that this interaction is attributable purely to the combined effects of the mutations on mRNA decay alone. However, reassuringly, the Parker group found that the combination of xrn1Δ and skiΔ and that of dcp1Δ with any of the ski2, ski3, or ski8 deletions were also lethal. They also showed that at a growth-permissive temperature the double mutation dcp1 ski8Δ resulted in greatly lengthened mRNA half-lives (20).

The accelerated decay associated with premature translational termination depends on the activities of three nonessential genes (UPF1, UPF2, and UPF3 [up-frameshift]). Although the encoded proteins are generally involved only in the decay of aberrant mRNAs, they are of interest because they seem to mediate one type of interaction between translational events and mRNA stability. Inactivation of any one of these genes can give rise to stabilization (in a nonadditive fashion) of mRNAs containing premature stop codons (327). The Upf proteins are not thought to have nuclease activities, and they are generally not required for the decay of wild-type mRNAs (78, 266), although there may be exceptions (such as PPR1 and CTFT13 [78, 322; see also reference 33a]). UPF genes were initially identified via a screen for suppressors of the His- phenotype caused by a + 1 frameshift in the HIS4 gene (his4-38) (107, 322, 323). This 5′-proximal frameshift leads to translational termination at an adjacent stop codon, which in turn destabilizes the mRNA. UPF1 was also identified as a high-copy suppressor of mitochondrial RNA splicing reactions (7).

Upf1p interacts with both of the other Upf proteins, but the binding to Upf2p is significantly stronger (25, 266). It is a cytoplasmic, multifunctional protein bearing a cysteine-rich region of about 500 amino acids. Its N terminus contains a transcriptional activation motif and its C terminus is involved in mRNA binding and polyadenylation. It is not clear whether the list of exo- and/or endonucleolytic activities associated with mRNAs. UPF1 was also identified as a high-copy suppressor of mitochondrial RNA splicing reactions (7).}

Upf1p has been shown to affect mRNAs differentially and whether they really are part of a generally acting decaying pathway. It is not yet clear whether the list of exo- and/or endonucleolytic activities involved in yeast mRNA turnover is complete (379, 523), and further activities may well be discovered. Finally, there may be factors that influence the stability of only certain mRNAs. For example, Ume2p and Ume5p are suspects that mediate the decay rates of mRNAs involved in yeast meiosis (see the section on regulation of the mRNA decay rate).

Pathways of mRNA Decay

Role of deadenylation. A picture of the pathways responsible for mRNA decay in yeast only began to take form over the last 5 years. There are a number of reasons to believe that deadenylation plays an important role in the degradation of at least some yeast mRNAs (78). Jacobson (264) has discussed how earlier studies in higher eukaryotic systems suggested a role for poly(A) shortening in triggering decay. A study of S. cerevisiae cells with a deletion in XRN1 revealed that they accumulated uncapable mRNAs bearing shortened poly(A) tails (234). This suggested that deadenylation might normally precede decapping, which in turn exposes mRNAs to 5′→3′ exonucleolytic degradation. Further experiments used poly(G) as a physical block to trap decay intermediates from specific mRNAs (576). After insertion of poly(G) into the 5′ UTR of MFA2, deadenylated and decapped decay products whose 5′ ends had been released of the eRFs, Upf1p forges a path for the remaining ribosome complex along the mRNA downstream of the termination event and mediates the recognition of a downstream element that triggers nonsense codon-dependent decay (108a). As we have seen, it is unclear whether a discrete consensus sequence motif is required for this triggering step (see also the next section). Alternatively, it might be envisaged that the Upf proteins are involved in the cellular response to the disruptive effects on polysome structure and/or nuclear transport caused by premature termination. Ongoing work by the respective research groups can be expected to elucidate the means by which the Upf proteins are involved in the coupling of translational termination and mRNA decay and also to clarify the full range of functions of these proteins in the cell.

A number of other proteins are known, or suspected, to be involved in mRNA degradation (Table 4). The screening of temperature-sensitive S. cerevisiae strains for the presence of mutations that affect mRNA decay has led to the identification of a number of mrt (mRNA turnover) alleles. Three of the mutations were in the XRN1 and DCP1 genes (see above), while four were localized to two further, as yet uncharacterized genes (MRT1 and MRT3) (204). The MRT1 and MRT3 mutations do not stabilize nonsense codon-disturbed mRNAs, but they do extend the half-lives of at least some noberrant mRNAs. However, stabilisation was found to be selective: mrt1-3 stabilized five different mRNAs between 1.9- and 4.6-fold but did not stabilize PAB1 mRNA, while mrt3-1 stabilized four of the tested mRNAs but had only a minimal effect on the half-life of GAL10 mRNA and slightly destabilised PAB1 mRNA (204). Both of these mrt mutations also inhibited decapping of MFA2 and PGK1 mRNAs containing poly(G) inserts in their respective 3′UTRs but did not affect the decapping of the PGK1 mRNA when this contained an early nonsense codon. The same authors found that PAB1 mRNA was also unaffected by inactivation of Dep1p. Since extracts prepared from the mrt1-3 and mrt3-1 strains contained normal levels of decapping activity, it was suggested that the Mrt proteins modulate the decapping reaction of nonaberrant mRNAs (204). It remains to be determined why the mrt alleles affect mRNAs differentially and whether they really are part of a generally acting decaying pathway.
trimmed to the 5' side of the poly(G) were detected (117, 118). Moreover, in an **xrn1** deletion strain, some of the **MFA2** mRNA accumulated as a decapped and deadenylated species. In a **dcp1** strain, a number of mRNAs were stabilized (36). The importance of the cap in stabilizing mRNA is underlined by the observation that conditional mutants of the capping enzyme Ceg1p manifest accelerated decay rates (493). These and other detailed studies (78) of two short-lived mRNAs (**MFA2** and **MATa**1) and one long-lived mRNA (**PGK1**) are consistent with the existence of a causal link between deadenylation and decapping in different types of yeast mRNA.

The concept of a major pathway comprising the causally linked steps deadenylation→decapping→5'→3' exonucleolytic decay has been formulated by the Parker group (Fig. 17) (78). This incorporates elements of direct signalling that are believed to trigger and/or modulate key steps. The nature of the causal relationship between deadenylation and further processes in decay remains uncharacterized. The respective kinetics of deadenylation and of the remaining decay process suggest that deadenylation triggers overall degradation but does not determine the rate at which it occurs (78). Other factors, including perhaps the **MRT** genes (207), are likely to control decapping and the further steps of degradation. Moreover, deadenylation-dependent decapping is apparently not the only route that has been observed in detailed studies of **PGK1** and **MFA2** decay (389, 390), meaning that there is at least a minor contribution from other pathways. For example, in **PGK1**, 3'→5' exonucleolytic decay also followed deadenylation (390). The mechanisms controlling or regulating the rate of deadenylation by poly(A) nuclease activities are not known, but current models implicate the mediation of signals in the 3'UTR or main body of the mRNA via Pab1p and/or additional factors (79, 344, 476). The triggering of further degradation (decapping) occurs once the poly(A) tail has reached an oligo(A) length and is also thought to involve Pab1p. Such a view is compatible with the suspected role of Pab1p in mediating 5'-3' interactions in mRNPs (see the sections on translation) and provides an explanation for the observations that mutations in **PAB1** result in the accumulation of decapped mRNAs with long poly(A) tails (78). It is argued that the loss of the coupling between deadenylation and decapping allows the latter to occur independently of the former, although this is not necessarily the only explanation of the observed effect.

However, attractive as this model may seem, these links between 3' and 5' events could be indirect. One way of representing this is to consider a minimum model of coupling be-
between triggering events and decay processes in which a number of possible processes that disturb the structure and function of the mRNP and polysome initiate degradation (polysome disruption model [Fig. 18] [see, for example, references 266 and 336]. It is a major objective of current work in this area to establish the extent to which more direct and/or differentiated mechanisms are responsible for the respective decay patterns of the mRNAs studied so far (78). Future efforts will have to distinguish between models of the two types illustrated here (Fig. 17 and 18).

Other triggers of decay. The accelerated decay induced by the presence of a premature stop codon can trigger rapid decapping that is apparently not dependent on deadenylation of the tail to the oligo(A) state (387). Analysis of the effects of xrn1 and dcp1 mutants indicates that the decay events observed here involve the same components of the degradation machinery as normal decay does (78). Might this reflect the fact that premature termination also triggers decay via destabilization of mRNP/polysome structure (see, for example, reference 266)? The disruptive effect seems to occur via a parallel route to that linked to deadenylation (Fig. 18).

This area has received renewed impetus from the most recent investigations of the Upf proteins, which have provided indications of how these factors might influence both translation and mRNA decay (108a). There are, however, still many questions to be answered, not least of which is how interactions between the mRNA and the posttermination ribosome complex (also called the surveillance complex [469a]) trigger accelerated decay. As discussed above, it is still not clear whether a specific type of discrete downstream element is required; certainly the pyrimidine-rich motif alone is insufficient. Moreover, the field is missing a unifying model explaining the responses of the degradation machinery both to termination on uORFs and/or the main ORFs of nonaberrant mRNAs and to aberrant termination on premature nonsense codons. Perhaps this model might be found if more was known of the kinetics governing normal and aberrant termination events, in particular the kinetics of binding and release of the eRFs and of other factors, like the Upfps, which can modulate termination. One possibility, for example, would be that peptide hydrolysis and eRF release at a subset of stop codons can proceed efficiently and be coupled to accelerated decay without positive modulation by the Upfps. This would, in turn, explain why termination at such stop codons can trigger accelerated decay.
at least partially independently of at least one of the UPF genes. In other words, the UPF dependence of termination-triggered accelerated decay may be kinetically controlled rather than constituting mechanistic necessity. Other characteristics of nonsense codon-dependent decay, such as the strong position dependence of the degradation kinetics, might also be related to the respective contributions of the eRFs and Upfs to the component steps of termination. The kinetics of the interactions between these respective factors and the ribosome might be controlled by polysome localization and/or structure. Much further work is needed to determine whether any of these speculations approximate to reality.

The other known means of triggering decay is endonucleolytic cleavage. So far, it has been suggested that endonucleolytic cleavage triggers the decay of the PGK1 (576; but see reference 390) and L2 (444) mRNAs. For PGK1, there is a considerable body of evidence indicating that 5'-3' exonuclease-driven decay from the decapped 5' end is the dominant pathway. More cases of endonucleolytic cleavage have been found in higher eukaryotes (117, 266, 462, 463). It remains to be seen whether endonucleases participate in rate-controlling steps in the decay pathways of a significant number of yeast mRNAs.

### Regulation of mRNA Decay Rate

An area of growing interest is the extent to which regulation of gene expression can be achieved by modulating the steps of mRNA decay. One potential means of achieving this is via coupling between mRNA turnover and translation. As discussed above, changes in the decay rate of an mRNA can be coupled to modulation of translational rate via the 5'UTR, provided that the mRNA has appropriate stability determinants. At present it is not clear how many yeast mRNAs show the sensitivity toward translation manifested by PGK1. On the basis of current evidence, 5'UTR-mediated translational regulation via trans-acting (site-specific) factors seems likely to be rare in yeast. However, a potential role for translation in observed regulatory changes in mRNA stability should always be considered possible until experimentally disproven. The role of uORFs has yet to be fully clarified, but stress- or even cell cycle-related modulation of decay mediated by uORFs can certainly be envisaged as a feasible, if not necessarily widely used, regulatory mechanism. Despite these considerations, it is noteworthy that the cases of stability regulation reported so far are evidently not directly mediated by translational modulation.

Five different meiotic mRNAs (early transcripts SPO13, SPO11, and IME1, the middle transcript SPO12, and the middle-late transcript DIT2) are stabilized up to approximately twofold in response to the shift from vegetative growth of *S. cerevisiae* in glucose medium to sporulation conditions in acetate medium (532, 533). This stabilization effect is dependent on proteins encoded by UME2 and UME5, whereby the product of the latter gene shows similarity to members of the *CDC28* serine/threonine-specific protein kinases. The regulation of meiotic mRNA stability could be shown to be triggered by the removal of glucose from the medium rather than by the shift to meiosis. Surosky et al. (532, 533) also showed that UME5-dependent stabilization of SPO13 does not involve any change in the rate of deadenylation. A more striking example of destabilization associated with glucose repression is the 12-fold decrease in the half-life of the *SDH2* mRNA observed upon a glycerol-to-glucose shift (84, 341). The 5'UTR of the *SDH2* mRNA is thought to play a major role in the glucose-induced destabilization effect (84). Other work has shown that glucose induces slightly accelerated decay of the *PGK1* and *FBP1* mRNAs (368) and a larger decrease in the stability of the *SUC2* mRNA (83). It remains to be seen how general this glucose-dependent destabilization of yeast mRNAs is and to what extent mRNA-specific factors are involved. So far, it is known that neither Ume5p nor most of the other known factors involved in the glucose repression pathway are required for the regulation of *SDH2* mRNA stability (83). The *REG1* gene product, on the other hand, was found to be required for regulation of the decay of this mRNA, although the mode of influence of this factor on mRNA turnover was not characterized.

There is also evidence that the stability of a number of ribosomal protein mRNAs is subject to regulation. Mild heat shock results in a transient destabilization of the S10 and L25 mRNAs (217). In a more intensively investigated case, the decay rate of the L2 mRNA was found to be regulated autogenously by the L2 subunit (443, 444). The destabilization caused by L2 requires the region of the L2 mRNA from −21 to +339 and has been shown to function in both the nucleus and the cytoplasm (444). A strategy involving the insertion of poly(G) to stabilize mRNA 3' of the region responsible for L2-dependent destabilization was used to examine the pathway of decay. The results were consistent with a model in which an endonucleolytic cleavage on the 3' side of this region constitutes the initial cleavage step (444).

### Back to Rate Control

Explaining how the trans-acting factors interact with the various structural features of mRNAs (mRNPs) to determine the pathways and rates of degradation remains a major goal of this field of research. The effective degrees of control (control coefficients) exercised by the steps of deadenylation, decapping, exonucleolytic decay, and endonucleolytic cleavage events can apparently vary from mRNA species to mRNA species. This matter is complicated by the processivity of at least some of the reaction steps (see for example, reference 379) and the fact that the causality of apparently sequential events is not understood. Moreover, one mRNA species may be degraded by more than one pathway at any one time (390). Despite this apparent complexity, mRNA turnover is not simply a free-for-all for RNases, and there is persuasive evidence that a considerable number of yeast mRNAs are degraded via a deadenylation-dependent 5'→3' pathway (Fig. 17) (78). However, this still leaves us without a general model that could explain the turnover behaviour of all mRNAs, whether aberrant or nonaberrant, and how this is influenced by the translational apparatus.

Although the polysome-disruption type of model (Fig. 18) is undoubtedly an oversimplification, the concept of a critical mRNP disruption event (or series of sequential or alternative disruption events) being initiated by a number of possible triggers allows some flexibility in explaining the relationships between mRNA structure and the kinetics of decay. Structural and/or functional interactions between the 5' and 3' ends of the mRNA may, in as yet unknown ways, contribute to the control of the decay pathway. Moreover, the roles of the respective translation factors remain unclear. Analogously to the parallel pathways of translational initiation, there are likely to be several potential routes to full hydrolysis of an mRNA. Obviously, the primary pathway followed by any given mRNA could theoretically be affected by a large number of factors. However, kinetic control and/or molecular channeling (in the context of guiding processes with the help of large molecular assemblies) can be expected to play decisive roles in mRNA stability.
degradation. For example, if normal mRNP structure inhibits decapping until an event such as deadenylation leads to exposure of the cap to attack by Dcp1p, this may explain the molecular basis of causality between these two steps (Fig. 17 and 18). It remains to be seen whether one relevant form of “disruption” caused by deadenylation is associated with its negative effect on translation. The rate of subsequent events, including decapping, may then be subject to control via mRNA structural elements that have little or nothing to do with the deadenylation phase (78). It should be possible to distinguish between the direct action of specific mRNA signals on individual steps (Fig. 17) and the mediation of polysome structure in controlling mRNA decay (Fig. 18), although both are likely to be relevant.

Finally, in this context it is of interest that prokaryotic polyadenylation may influence decay via mechanisms that, at least at first sight, appear to be different from those acting in eukaryotes. It has been demonstrated that 3′ polyadenylation of CoeE1 RNAI in E. coli destabilizes this RNA (602). Moreover, polyadenylation by a poly(A) polymerase(s) of mRNA decay intermediates generated by endonucleolytic cleavage (by RNase E) in E. coli accelerates the decay of these fragments (208a). The poly(A) tails on mRNA decay intermediates may recruit the 3′→5′ exonucleases PNPase and RNase II, perhaps as components of the prokaryotic degradosome. Interestingly, however, polyadenylation cannot be essential for the initial cleavage(s) of full-length de novo transcripts, since rapid degradation of parts of at least some polycistronic mRNAs occurs before transcription is completed (see, for example, references 365a and 490a). Thus, while the poly(A) tail is likely to be a general modulator of the translation and decay of full-length eukaryotic mRNAs, its role in E. coli seems to be limited to that of a tagging function that promotes the rapid degradation of decay intermediates. There may, however, be some common ground between the two domains; it remains to be determined whether the eukaryotic poly(A) tail acts to promote the recruitment of RNases located in “exosome” particles. Perhaps at least this recruitment property of the poly(A) tail is a feature that has been conserved by means of convergent or divergent evolution.

CONCLUSIONS AND PERSPECTIVES

The posttranscriptional control field is in a phase of rapid development, and the key role of yeast in much of the research performed so far is clearly evident. However, while focusing on yeast, this review has endeavored to emphasize the value of comparative reference to the wealth of information generated by work with bacterial, mammalian, and plant systems. Analysis of the similarities and the differences between the respective cell types will continue to provide additional insight into central principles of mechanism and control.

Recent years have seen the cloning and at least preliminary characterization of many genes and proteins involved in eukaryotic translation and mRNA decay. The three-dimensional structures of only very few of these proteins have been solved so far, but new ones will undoubtedly appear with increasing frequency, and we can also look forward to structural information on multicomponent complexes such as eIF4F. However, progress in understanding the complex process of posttranscriptional gene expression will require that structural studies be complemented by detailed quantitative investigations of intermolecular interactions and rate control. Work on the thermodynamics and kinetics of function in these systems will therefore be as least as important as the analysis of their structures, and it will continue to be essential to examine carefully the relevance of in vitro data to the conditions that exist in the cellular environment.

Any reviewer of this field cannot help but be struck by the overlaps and networks between functions of the cellular components that are involved in cytoplasmic gene expression. This principle is certainly not unexpected for evolving cellular systems, and awareness of it warns us not to be unduly eager to categorize too narrowly the roles of individual factors. Indeed, the redundancy of function becoming so evident in yeast is a significant feature worthy of analysis in its own right. In proceeding with work in this area, it is important to take note of the recent developments in large-scale yeast gene functional analysis (414). Of the approaching 1,000 genes analyzed by disruption to date, only approximately 12% have been found to be essential. Other data have indicated that up to 40% of the S. cerevisiae ORFs can be disrupted without resulting in detectable phenotypes in a range of assays. One of the reasons for this is likely to derive from the large gap between the conditions under which this organism has naturally evolved and the conditions that apply in standard laboratory experiments and assays. It is therefore essential that we begin to bridge that gap by examining regulatory phenomena in yeast under precisely specified growth-limiting conditions. Such work could potentially redefine the meaning of significant parts of the current dataset on posttranscriptional control and could also lead to an understanding of the functions of many new genes whose roles are apparent only transiently or under tightly defined conditions. This is certainly a major challenge for the future.

Multifunctionality can take the form of cellular functions (or capabilities) shared by two or more proteins (e.g., RNA helicases) or could constitute the basis for coupling between different processes. Coupling has been a common theme in this review and is also seen in other forms, for example between protein synthesis and transport (30). Beyond this versatility of the components of the cellular machineries, future work will have to address their supramolecular organization. In particular, and reminiscent of multifunctional enzyme complexes in metabolism (416), the posttranscriptional pathways are subject to levels of control and channeling that are only just becoming discernible. In considering these and other aspects in this review, it is evident that many revelations about these exquisitely balanced systems are still to come.

ACKNOWLEDGMENTS

While taking full responsibility for the views expressed, and any errors, in this review, I am grateful to Patrick Linder (Geneva, Switzerland) and Mick Tuite (Canterbury, United Kingdom) for their comments on early versions of it and to one of the reviewers, who provided valuable constructive criticism. I also thank colleagues who sent me manuscripts prior to publication. Barbara Skoyles provided assistance with the figures and references. Some of the work from my laboratory mentioned in this review was supported by the BBSRC (UK).

REFERENCES


105. Cui, Y., J. D. Dinman, and S. W. Peltz. 1996. mof-1 is an allele of the UPF1/FS2 gene which affects both mRNA turnover and -1 ribosomal frameshifting efficiency. EMBO J. 15:5726–5736.


174. Geballe, A. P.


260. Ito, K., K. Eihihara, and Y. Nakamura. Dissection of tRNA mimicry element of protein release factor eRF1 of fission yeast: eRF3 binding is not required for eRF1 function and cell viability. Submitted for publication.


266. Itoh, T., K. Eihihara, and Y. Nakamura. Dissection of tRNA mimicry element of protein release factor eRF1 of fission yeast: eRF3 binding is not necessary for eRF1 function and cell viability. Submitted for publication.


289. LaGrande, T. L., and R. Parker. 1998. Isolation and characterization of


transport of cyclin D1 mRNA are increased in cells overexpressing eukaryotic initiation factor 4E. Proc. Natl. Acad. Sci. USA 93:1065–1070.


termini of yeast mRNA. J. Biol. Chem. 251:2988–2904.


518. Stansfield, L., S. Pitts, and M. F. Tuite. Personal communication.


