Quarterly Reviews of Biophysics **37**, 3/4 (2004), pp. 197–284. © 2004 Cambridge University Press
doi:10.1017/S0033583505004026 Printed in the United Kingdom

Translation initiation: structures, mechanisms and evolution

Assen Marintchev and Gerhard Wagner*
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, USA

Abstract. Translation, the process of mRNA-encoded protein synthesis, requires a complex apparatus, composed of the ribosome, tRNAs and additional protein factors, including aminoacyl tRNA synthetases. The ribosome provides the platform for proper assembly of mRNA, tRNAs and protein factors and carries the peptidyl-transferase activity. It consists of small and large subunits. The ribosomes are ribonucleoprotein particles with a ribosomal RNA core, to which multiple ribosomal proteins are bound. The sequence and structure of ribosomal RNAs, tRNAs, some of the ribosomal proteins and some of the additional protein factors are conserved in all kingdoms, underlying the common origin of the translation apparatus. Translation can be subdivided into several steps: initiation, elongation, termination and recycling. Of these, initiation is the most complex and the most divergent among the different kingdoms of life. A great amount of new structural, biochemical and genetic information on translation initiation has been accumulated in recent years, which led to the realization that initiation also shows a great degree of conservation throughout evolution. In this review, we summarize the available structural and functional data on translation initiation in the context of evolution, drawing parallels between eubacteria, archaea, and eukaryotes. We will start with an overview of the ribosome structure and of translation in general, placing emphasis on factors and processes with relevance to initiation. The major steps in initiation and the factors involved will be described, followed by discussion of the structure and function of the individual initiation factors throughout evolution. We will conclude with a summary of the available information on the kinetic and thermodynamic aspects of translation initiation.

* Author for correspondence: G. Wagner, Department of Biological Chemistry and Molecular Pharmacology, 240 Longwood Avenue, Harvard Medical School, Boston, MA 02115, USA.
Tel.: (617) 432 3213; Fax: (617) 432 4383; E-mail: wagner@hms.harvard.edu

1. Introduction

The field of translation and translation initiation in particular has experienced an unprecedented growth in recent years, both in terms of accumulation of new data and of much deeper understanding of the underlying processes. We now have insights into the structure and location of most translation initiation factors (IFs) and can discuss their roles on a structural and mechanistic level.

Here, we have attempted to summarize at least a fraction of the landslide of new information and present the emerging picture of various aspects of translation initiation. The main focus of this review is on the mechanism of translation initiation, the structure and function of the IFs and the organization of the initiation complexes (ICs). A look on translation initiation from an evolutionary perspective emphasizes both the common origins and organization of translation and the great diversity among species. We give special attention to the organization of the translational apparatus in the cell, the concept of channeling of factors and intermediates, and their implications. The endless variety of mechanisms of translation regulation and alternative initiation are beyond the scope of this review and are only included where they have a direct relation to our understanding of the general mechanisms of initiation.

Most of our knowledge about translation comes from eubacteria and eukaryotes. In recent years, the archaeal system has started to attract more attention, in part because it is remarkably similar to that in eukaryotes, but much simpler and involves fewer translation factors. As archaeal translation is related to eukaryotic translation, statements about eukaryotes will be assumed throughout this review to apply to archaea as well and vice versa, unless otherwise specified. The organellar translational apparatus is evolutionarily related to its eubacterial counterpart, but has undergone long independent evolution in its specific environment and will not be discussed here.

Section 2 contains a brief overview of the structure of the ribosome. In Section 3 we present an overview of translation, with emphasis on factors and processes with relevance for our understanding of translation initiation. The mechanism of translation initiation is discussed in Section 4. Section 5 contains a summary of our knowledge about the structures of individual IFs and the organization of the ICs. Finally, in Section 6 we try to look at translation initiation and its regulation from a kinetic perspective.

It was not humanly possible to discuss all individual reports on any subject (or even only the ones we are aware of). Therefore, we have tried to present what we see as the prevailing views and refer the reader to recent specialized reviews for details. While browsing through the sea of sometimes contradictory publications, we tried to follow some general 'guidelines':

(1) With the risk of ignoring groundbreaking discoveries, we rarely mention isolated reports, contradicting the consensus from the rest of the field, unless the results appear sound and unambiguous. On some occasions, we have discussed controversies, mainly to emphasize that a 'mainstream' concept has been seriously challenged and promote the broader acceptance of the alternative.
(2) Detection of relatively weak interactions depends on the limitations of the method used, concentrations and experimental conditions. Even relatively strong interactions in the submicromolar range can be lost during centrifugation, a method routinely used in translation studies. Therefore, we have tried to be cautious with negative binding results, and especially

Abbreviations: IC, initiation complex; IF, translation initiation factor; eIF, translation eukaryotic initiation factor; EF, elongation factor; eEF, eukaryotic elongation factor; aa-tRNA, aminoacyl-tRNA; aaRS, aminoacyl-tRNA synthetase; cryo-EM, cryo-electron microscopy; A site (of the ribosome), aminoacyl-tRNA site; P site, peptidyl-tRNA site; E site, exit site; ASL, anticodon stem-loop (of the tRNA); PTC, peptidyl transferase center; GAC, GTPase-associated center; SRL, sarcin/ricin-binding loop (in the large ribosomal subunit); GEF, guanine nucleotide exchange factor; GAP, GTPase activating protein; SD, Shine–Dalgarno sequence; RBS, ribosome-binding site; ORF, open reading frame; UTR, untranslated region; IRES, internal ribosome entry site; Gcn^-, general (translational) control non-derepressible; Gcd^-, general control derepressed; Sui, suppresor of initiator codon mutation; NTD, N-terminal domain; CTD, C-terminal domain; ZBD, Zn-binding domain; RRM, RNA recognition motif; WT, wild type.

with reports that a mutation completely abolishes binding. On the other hand, interactions observed only *in vitro* at non-physiological concentrations (especially if they were mainly electrostatic) were only considered reliable if it was known that the factors involved are brought in proximity via other interactions.

(3) A modification or mutation often has an effect in some systems, and some mRNAs, but not others, and the effect may depend largely on the experimental conditions. Therefore, if effects were seen in some studies but not in others, we have generally favored the presence of an effect, unless there were clear contradictions.

(4) Data obtained under non-physiological conditions were considered with extreme caution.

(5) If a factor is reported to bind to other factors both individually and in combination, it is hard to know if the interactions are cooperative, anti-cooperative or independent, without quantitative binding data. If a factor is found to bind to two other factors simultaneously, but not individually, more often than not, it also binds to each individual factor, even if the binding was too weak to detect by the method used.

(6) Similarly to point (2) above, reports of mRNA binding by RNA-binding proteins were also subject to scrutiny, because any sequence- or structure-specific RNA-binding protein usually has a fairly high non-specific RNA-binding affinity, which cannot be taken as proof that it actually binds to mRNA *in vivo*.

(7) On the other hand, there are numerous reports of high-affinity sequence/structure-specific mRNA binding by 'non-specific' factors, such as eukaryotic initiation factor 4A (eIF4A) and eIF4G, for example. Although such high-affinity binding sites may be absent from most mRNAs, it cannot be assumed that a 'non-specific' mRNA-binding eIF binds all mRNAs with the same affinity. Furthermore, as it appears that any imaginable regulation mechanism is in fact used somewhere, it is likely to find more and more viral or cellular mRNAs using high-affinity 'non-specific' factor-binding sites for translation initiation.

We apologize for any omissions of important work and views due to space limitations or ignorance on our part.

2. Ribosome structure and organization of the translation apparatus

2.1 Nomenclature

The ribosomes, ribosomal subunits and ribosomal RNAs (rRNAs) are identified by their sedimentation coefficients: the intact ribosomes are 70S in eubacteria and archaea, and 80S in eukaryotes; the small subunits are 30S and 40S respectively; and the large subunits are 50S and 60S respectively. The rRNA in the eubacterial and archaeal small subunit is 16S (18S in eukaryotes), and the rRNAs in the large subunit are 23S (26S and 5·8S in *S. cerevisiae*, and 28S and 5·8S in human respectively) and 5S. The ribosomal proteins are given designation 'S' and a number for proteins in the small ribosomal subunit and 'L' and a number for those belonging to the large ribosomal subunit. For example, S1 is small ribosomal subunit protein 1. As mentioned above, part of the bacterial and eukaryotic ribosomal proteins are homologous to each other and it is accepted that homologous ribosomal proteins have similar location on the ribosome and probably perform similar functions in ribosome biogenesis and/or translation. Unfortunately, for historic reasons, ribosomal proteins conserved between bacteria and eukaryotes do not have the same names in both nomenclatures, and alternative nomenclatures exist for the eukaryotic ribosomes (for a comparison of the nomenclatures of yeast ribosomal

proteins and their relationships to mammalian, archaeal and eubacterial ribosomal proteins see Planta & Mager, 1998).

Additional designations exist in eukaryotic translation for the pre-IC of the small ribosomal subunit with IFs, before it is bound to mRNA (43S), and for the IC on the mRNA (48S), also based on their respective sedimentation coefficients.

2.2 Ribosome structure

In recent years, several X-ray structures of ribosomal subunits and of the 70S bacterial ribosome were determined, some at a resolution as high as 2·4 Å. Some of these structures contained tRNAs and translation factors. In addition, cryo-electron microscopy (cryo-EM) reconstructions are also available for structures and complexes corresponding to various steps along the translation pathway (reviewed in Ramakrishnan, 2002). Whereas X-ray structures of the small ribosomal subunit and of the intact ribosome are only available from eubacteria, a high-resolution X-ray structure of an archaeal 50S subunit has also been determined (Ban *et al.* 2000) and the high degree of similarity to its eubacterial counterpart reinforces the universal conservation of ribosome structure throughout evolution. Only cryo-EM information is available for the eukaryotic ribosome. The resolution of the cryo-EM reconstruction of the yeast 80S ribosome (∼15 Å) was sufficient to fit common structural elements from the X-ray structures of the bacterial 30S and archaeal 50S subunits and model homologous proteins. Regions of unassigned electron density provided indications where ribosomal proteins without bacterial homologs could be located, although the identity of the proteins could not be inferred (Spahn *et al.* 2001).

The structure of the small subunit can be subdivided into head, neck, platform, and body, which have obvious relationships to the structural domains of the 16S rRNA: 5′-domain, central domain, 3′-major domain, and 3′-minor domain (Fig. 1). The 5′-domain corresponds to the body; the central domain – to most of the platform; and the 3′-major – to the head. The neck provides a relatively flexible connection between the head and the rest of the small subunit. The 3′-minor domain consists of the last two helices (44 and 45) and the 3′-end of the rRNA. It was noted that the structural domains are 'nearly structurally autonomous' and that 'this organization immediately suggests that the domains are designed to move relative to one another during protein synthesis'. The long helix 44 lies across the body and ends into the platform and neck. Therefore, helix 44 is connected to all major domains and could relay conformational changes and movement along the entire small subunit. In contrast, the large subunit consists of a rigid core and mobility is restricted primarily to segments on the periphery (Yusupov *et al.* 2001).

The mRNA binding site of the ribosome is on the small subunit, along the neck region between the head and the body, whereas the peptidyl transferase center (PTC) is on the large subunit (Fig. 1*c*). The ribosome has three binding sites for tRNA, shared between the two subunits. The aminoacyl (A) site has high affinity for aminoacyl-tRNA (aa-tRNA); the peptidyl (P) site has high affinity for peptidyl-tRNA; and the exit (E) site has high affinity for deacylated tRNA. The anticodon stem-loop (ASL) of the tRNA is oriented toward the mRNA on the small subunit, whereas the acceptor end of the tRNA, to which the amino acid is attached, binds to the large subunit, and the acceptor ends of the tRNAs in the A and P sites are in the PTC. The GTPase-associated center (GAC) and the sarcin/ricin-binding loop (SRL) on the large ribosomal subunit are important for stimulation of GTP hydrolysis by several translation factors

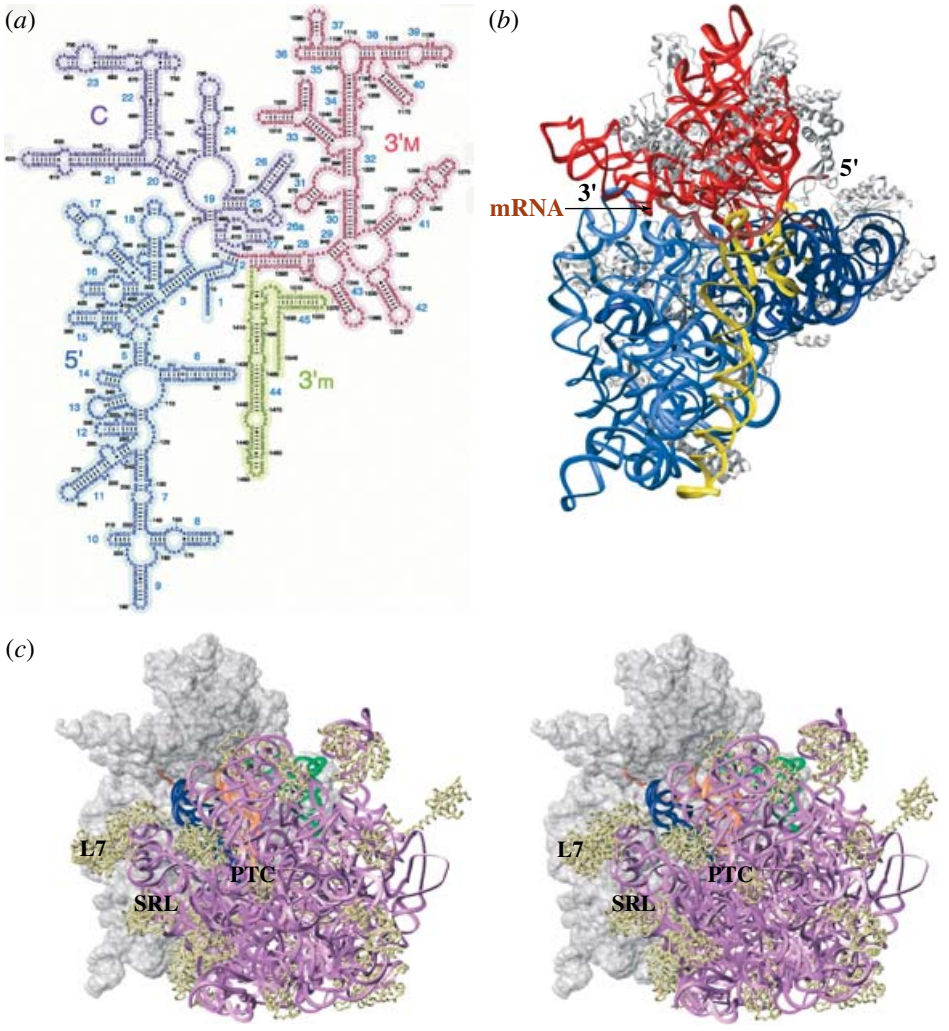

Fig. 1. Ribosome structure. (*a*) Domain organization of the 16S rRNA (from Fig. 3a, Yusupova *et al.* 2001, with permission). The 5′-domain (blue) corresponds to the body, the central domain (purple) to the platform, the 3′-major domain (red) to the head, and the 3′-minor domain (yellow) – to the last two helices (44 and 45) and the 3′-end of the rRNA. (*b*) The structure of the 30S ribosomal subunit from the structure of the 70S ribosome, PDB code 1JGP (Yusupova *et al.* 2001). 16S rRNA is shown as a ribbon and colored as in (*a*). Ribosomal proteins are shown in grey. The mRNA is in brown and its 3′-end (in the 'Entry' channel) and 5′-end (in the 'Exit' channel) are shown. (*c*) Stereo view of the structure of the 70S ribosome, PDB codes 1JGP and 1GIY (Yusupova *et al.* 2001). The small ribosomal subunit is in semitransparent surface representation and colored in grey. The mRNA is in brown, the A-, P-, and E-site tRNAs are in blue, coral and green respectively. The 23S and 5S rRNAs from the large subunit are shown as ribbon and colored in violet, and the large ribosomal proteins are in beige. The peptidyl-transferase centre (PTC), the sarcin/ricin-binding loop (SRL), and the L7 ribosomal protein are labeled.

(Fig. 1*c*). They are located at the base of the L7/L12 stalk near the aa-tRNA entry site (the nomenclatures concerning the GAC vary and the term is often used with a broader meaning to include both the L7/L12 stalk and the SRL). The nascent peptide exits through a channel in the large subunit (reviewed in Ramakrishnan, 2002; Tenson & Ehrenberg, 2002).

3. Overview of translation

The process of protein synthesis can be subdivided into several major stages: initiation, elongation, termination and recycling. Translation initiation will be discussed in depth in the following sections and, therefore, only a brief description is presented in this section. In discussing elongation, termination and recycling, specific attention will be paid to processes and factors with relevance to initiation. Some highlights from this section, which we would like to bring to the reader's attention are: (1) The elongation factors (EFs) delivering aa-tRNA to the ribosome are homologous to the γ subunit of the eukaryotic initiation factor 2 (eIF2). (2) The large domain rearrangement in elongation factor EF1A (formerly EF-Tu) upon GTP hydrolysis is an exception rather than the rule for this family of G proteins. Accordingly, the nearly 1000-fold lower affinity of EF1A·GDP for aa-tRNA, compared to EF1A·GTP, may also be an exception. (3) In eukaryotes, some events within translation are organized at a higher level, which is termed channeling; tRNA, factors and intermediates are predominantly channeled along the translation pathway and rarely able to diffuse freely. This is in part also true for yeast, especially with respect to channeling of tRNAs between aminoacyl-tRNA synthetases (aaRS), eukaryotic elongation factor 1A (eEF1A), and ribosomes. As part of closing the tRNA channeling cycle, eEF1B, the exchange factor (GEF) for eEF1A, forms a stable complex with eEF1A·GTP and is only released upon aa-tRNA binding to eEF1A·GTP, whereas free eEF1A·GDP has nanomolar affinity for unacylated tRNA.

3.1 Translation initiation

Translation initiation covers all the steps between subunit dissociation upon termination in the previous translation cycle, and the assembly at an mRNA start codon of a ribosome ready for elongation. During translation initiation, the ribosome, with an initiator aa-tRNA in the P-site, is assembled on mRNA, with the help of a set of IFs. The main tasks that are performed by the translation apparatus during initiation (not necessarily in this order) are: (1) subunit dissociation and anti-association, (2) selection of the initiator aa-tRNA, (3) selection of the correct translation start site, and (4) subunit joining at the start codon. At the end of initiation, the ribosome is ready to accept the first elongator tRNA and form the first peptide bond, which marks the beginning of the next stage, elongation (Fig. 2).

3.2 Translation elongation

Translation elongation is the process of synthesis of the polypeptide chain, by the ribosome assembled at the start codon, until a stop codon is reached.

3.2.1 Mechanism

During elongation, an aa-tRNA is first bound to the A-site and if proper base-pairing between the mRNA codon in the A-site and the tRNA anticodon is established, a peptide bond is formed with the peptide attached to the tRNA in the P-site, accompanied by transfer of the peptide (now 1 amino acid longer) to the A-site tRNA. Then, the peptidyl tRNA is moved from the A- to the P-site, and the deacylated tRNA from the P-site is moved to the E-site, displacing from there the tRNA deacylated in the previous cycle. The mRNA is coordinately translocated by one codon. Thus, tRNA-mRNA base-pairing and the correct reading frame are retained (reviewed in Merrick & Nyborg, 2000; Ramakrishnan, 2002).

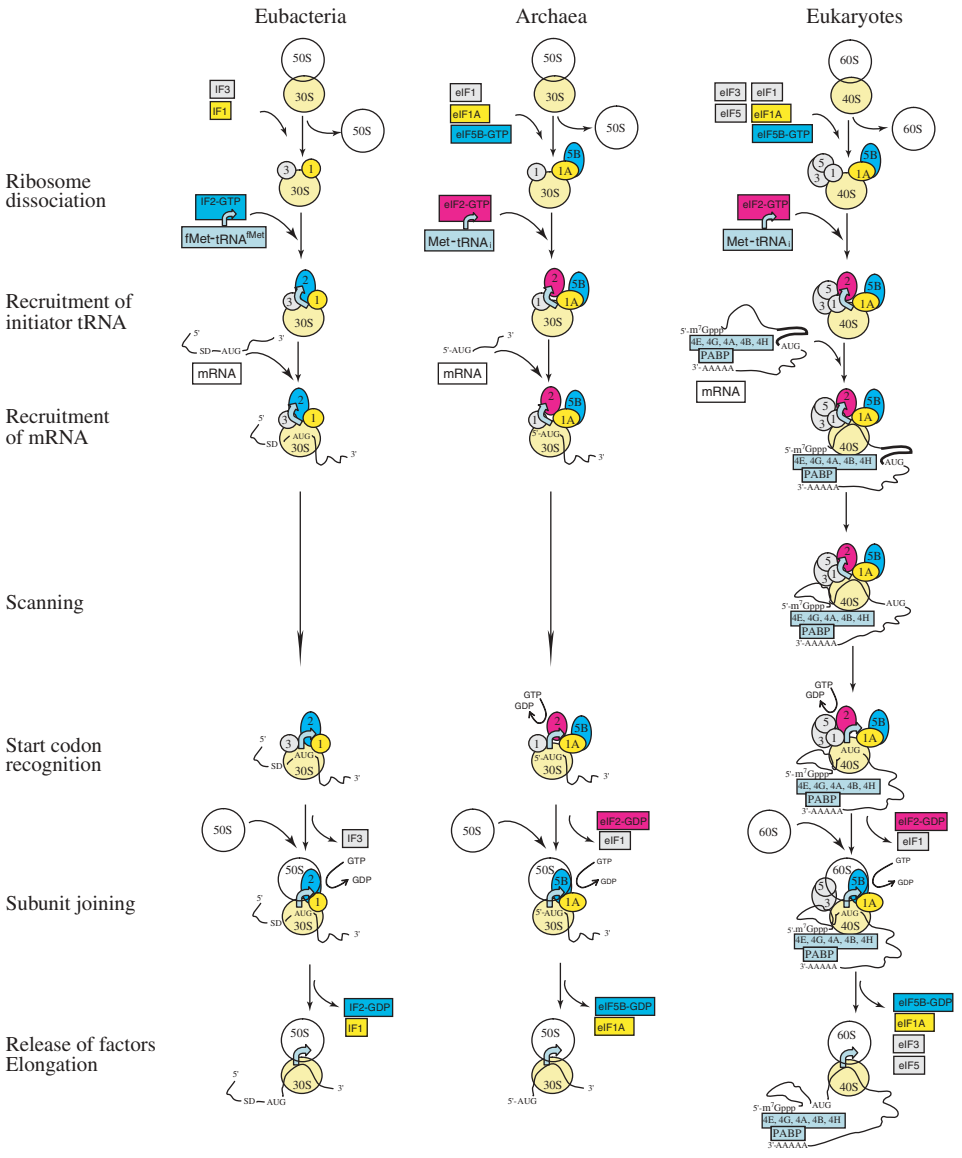

Fig. 2. Translation initiation. Schematic representation of translation initiation in eubacteria, archaea and eukaryotes. The universally conserved pairs of proteins IF1/eIF1A and IF2/eIF5B are in yellow and blue respectively. eIF2 (present only in archaea and eukaryotes) is in red. The cap/poly-A-binding complex (present only in eukaryotes) is in light blue. The rest of the initiation factors (IFs) are in grey. The 5′- and 3′-ends of mRNA are labeled. SD, Shine–Dalgarno sequence. Every effort has been made to provide a correct temporal and spatial representation of the events; however, the exact timing of recruitment and release of factors is not always known. Furthermore, the recruitment of IFs and RNAs need not follow a precise order, but may be a stochastic process. Note that GTP hydrolysis by IF2/eIF5B occurs after subunit joining and is required for release of IF2/eIF5B from the ribosome. Only the 5′-end-dependent initiation mechanism is shown for archaea, but internal SD-dependent initiation is also used in these organisms on polycistronic mRNAs. The scheme for eukaryotic initiation presumes that the scanning 43S ribosomal complex remains associated with the 5′-cap (see text for details). eIF1 and eIF2 and bacterial IF3 need to be displaced from their original positions for subunit joining to occur (and are shown as 'leaving' before subunit joining), but could remain associated with the ribosome. The other IFs remain associated with the ribosome during subunit joining and some even early in elongation (see text for details).

In bacteria, the aa-tRNA is brought to the ribosome as part of an EF1A·GTP·aa-tRNA ternary complex. Elongation factor EF1A (formerly EF-Tu) is universal and binds to most combinations of tRNAs and the amino acids attached to their 3′-end. The binding affinity of EF1A is determined by its affinities for the tRNA portion and for the aminoacyl portion of the aa-tRNA and the lack of apparent specificity is achieved through combinations of high affinity for the tRNA and low – for the amino acid, and vice versa. The affinity of EF1A is lower for certain aa-tRNAs, such as the initiator tRNA fMet-tRNA^{fMet} (recognized by IF2), the selenocysteine-tRNA (Sec-tRNA^{Sec}, recognized by a specialized EF SelB), and some aa-tRNA combinations that are intermediates for further modification of the attached amino acid, like conversion of Asp-tRNA^{Asn} into Asn-tRNA^{Asn} in some species. The latter group gains high affinity for EF1A after modification. The same recognition mechanism of aa-tRNA is used by the eukaryotic EF1A homolog, eEF1A (reviewed in Francklyn *et al.* 2002).

Initial binding of the EF1A·GTP·aa-tRNA ternary complex to the ribosome near the GAC places the aa-tRNA in a hybrid A/T site, where the ASL of the tRNA is near the A-site mRNA codon in the decoding center of the small subunit, but the rest of the tRNA is not yet positioned in the A-site. EF1A·GTP·aa-tRNA ternary complexes containing non-cognate tRNA have equal chance to bind to the ribosome as complexes containing the correct tRNA complementary to the codon in the A-site. After initial binding of the EF1A·GTP·aa-tRNA ternary complex to the ribosome, selection against the incorrect tRNAs is performed at two stages. First, coordinated conformational changes in the ternary complex and the ribosome allow the anticodon of the aa-tRNA to contact the mRNA codon in the A site. The codon-anticodon pairing has dual roles: (1) It stabilizes the complex between EF1A·GTP·aa-tRNA and the ribosome. The affinity of ternary complexes containing non-cognate tRNA is low and they quickly dissociate without GTP hydrolysis by EF1A. (2) Discrimination between cognate tRNAs and near-cognate tRNAs (forming non-canonical base pairs) is performed by 'inspection' of the geometry of the minor groove in the first two base pairs of the codon in the A-site. Non-canonical base pairs (e.g. G·U) are tolerated in the third, 'wobble' position. The discrimination in the first two positions is mediated by the universally conserved nucleotides 530, 1492 and 1493, whose bases become inserted in the minor groove of the codon-anticodon base pairs (Ogle *et al.* 2001). The binding of a cognate tRNA to the codon in the A-site promotes a 'closed' conformation of the small subunit required for ribosome-stimulated GTP hydrolysis by EF1A.

Upon GTP hydrolysis, EF1A is released and the aa-tRNA is accommodated in the A-site, with the acceptor end being inserted into the PTC of the large subunit. A second round of selection occurs at this stage: the rates of accommodation of the near-cognate aa-tRNAs in the PTC ($0{\cdot}1\ s^{-1}$) are much slower than their rates of dissociation (rejection): $\sim 6\ s^{-1}$, leading to ~100-fold discrimination. In contrast, cognate aa-tRNAs bind more tightly with negligible ($<0{\cdot}3\ s^{-1}$) dissociation rates, compared to their higher rates of accommodation ($\sim 7\ s^{-1}$). In summary, the discrimination against non-cognate aa-tRNAs is achieved predominantly at the first selection step, before GTP hydrolysis, whereas discrimination against near-cognate aa-tRNAs is achieved both at the first step before GTP hydrolysis (10- to 100-fold), and at the second step, after GTP hydrolysis (~100-fold), yielding overall misincorporation rates for near-cognate aa-tRNAs in the order of 10^{-3} to 10^{-4}. Such high fidelity could not have been achieved based only on differences in binding affinities of the cognate *versus* near-cognate aa-tRNAs, but instead rely on induced fit, where only the binding of a cognate aa-tRNA leads to acceleration of rate-limiting structural rearrangement steps (reviewed in Rodnina &

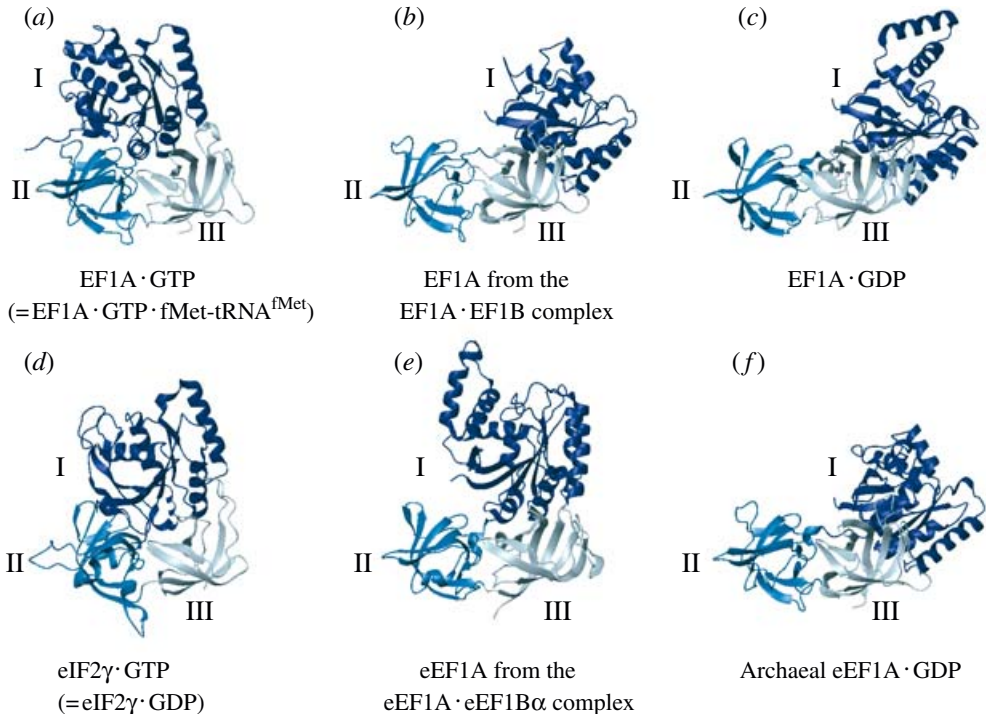

Fig. 3. Structures of proteins from the EF1A (EF-Tu) family in different conformations. Domain I (the G domain) is in dark blue, domain II is in blue, and domain III is in light blue. Domains II and III from all proteins were structurally aligned, in order to illustrate the different orientations of domain I with respect to the rest of the protein. (*a*) Structure of *T. aquaticus* EF1A (EF-Tu) in complex with the non-hydrolyzable GTP analog GDPNP, PDB code 1EFT (labeled as EF1A·GTP for simplicity). The structure is superimposable with the structure of EF1A from the EF1A·GTP·Cys-tRNA^{Cys} ternary complex, PDB code 1B23 (not shown). (*b*) Structure of *T. thermophilus* EF1A from the complex with the nucleotide exchange factor EF1B (EF-Ts), PDB code 1AIP. (*c*) Structure of *T. aquaticus* EF1A·GDP, PDB code 1TUI. (*d*) Structure of *P. abyssi* eIF2γ·GTP, PDB code 1KK1. The domain orientation is identical in the structures of *P. abyssi* eIF2γ·GDP, PDB code 1KK2, and apo-eIF2γ, PDB code 1KK0, and very similar ($\sim$14° rotation of domain I) in *M. jannaschii* apo-eIF2γ, PDB code 1S0U (not shown). (*e*) Structure of yeast eEF1A from the complex with the nucleotide exchange factor eEF1Bα-CTD, PDB code 1G7C. (*f*) Structure of *S. solfataricus* eEF1A·GDP, PDB code 1JNY. Note that the structure of EF1A in (*b*) is similar to the EF1A·GDP and eEF1A·GDP structures in (*c*) and (*f*), whereas the structure of eEF1A in (*e*) is closer to that of EF1A·GTP in (*a*). eEF1Bα is not related to bacterial EF1B and the two proteins bind to different surfaces of eEF1A and EF1A respectively (not shown).

Wintermeyer, 2001). The steric restrictions based on the geometry of canonical *versus* non-canonical base pairs are very important in discrimination between cognate and near-cognate tRNAs. This becomes especially clear in cases where a near-cognate codon-anticodon pair has the same or even higher binding energy than the cognate pair, but is still efficiently eliminated during translation (Ogle *et al.* 2001).

The structures of bacterial EF1A (EF-Tu) have been determined in a GTP- and GDP-bound form and with GTP and bound aa-tRNA (Fig. 3). EF1A is composed of three domains. The G-domain (domain I) binds the nucleotide and is packed against domains II and III. Two segments, called 'Switch 1' and 'Switch 2' are important for regulation of GTP binding and hydrolysis. The aa-tRNA binds across domains II and III with the amino-acid moiety binding to domain II, at the interface with domain I. Domains I and II are the most conserved, whereas

domain III is not conserved in some members of the EF1A family. Cryo-EM reconstructions of EF1A·GTP·aa-tRNA binding to the ribosome indicate that the sarcin/ricin-binding loop (SRL) interacts with the G-domain near the nucleotide-binding site, whereas the L7/L12 stalk is on the opposite side of the G-domain (reviewed in Merrick & Nyborg, 2000; Ramakrishnan, 2002).

After aa-tRNA binding in the A-site and peptide bond formation, a second factor, EF2·GTP (formerly EF-G) binds to the same site on the ribosome as EF1A, and triggers translocation of the A-site tRNA to the P-site with concomitant translocation of the mRNA by one codon. EF2 hydrolyzes GTP in the process and the resulting EF2·GDP dissociates from the ribosome, leaving the A-site open for binding another EF1A·GTP·aa-tRNA ternary complex. The translocation is thought to go through hybrid, 'A-P' and 'P-E' states of the tRNAs, in which the acceptor ends of the tRNAs move first, followed by simultaneous translocation of the mRNA and the anticodon ends of the tRNAs. The translocation involves a 'ratchet-like' rotation of the small subunit with respect to the large subunit. The structure of EF2·GDP bears a remarkable resemblance with the EF1A·GTP·aa-tRNA complex and has sparked a long-lasting search for other cases of molecular mimicry in translation. It appears, however, that most claims for mimicry are not as obvious and the important characteristics are interactions with common sites and fitting into the same cavities on the ribosome, which not always require extensive mimicry in shape and structure (reviewed in Lancaster *et al.* 2002; Ramakrishnan, 2002; Valle *et al.* 2003).

EF1A·GDP is recycled to its active GTP-bound form by a guanine nucleotide exchange factor (GEF), EF1B (formerly EF-Ts). No GEF has been reported for EF2 (EF-G) and it has been assumed that spontaneous dissociation of GDP from EF2 is fast enough to allow equilibration between the GTP- and the GDP-bound forms. Under optimal conditions, elongation in bacteria proceeds at a rate of 10–15 amino acids per second and with fairly low error rate of 10^{-3} to 10^{-4} (reviewed in Merrick & Nyborg, 2000; Rodnina & Wintermeyer, 2001).

Elongation is well conserved among all kingdoms of life and the eukaryotic factors eEF1A (formerly eEF1α) and eEF2 are homologous to bacterial EF1A (EF-Tu) and EF2 (EF-G) respectively. The structures of GTP- and GDP-bound EF1A demonstrate that a large rearrangement occurs upon GTP hydrolysis, involving almost 90° rotation between the G-domain (domain I) and domains II and III (Fig. 3). This rearrangement has been ascribed to the entire family of EF1A-like G proteins. It is likely that the same is true for eEF1A: the structure of archaeal eEF1A·GDP (Vitagliano *et al.* 2001) resembles that of EF1A·GDP, whereas the structures of the complex of the eukaryotic EF1A homolog, eEF1A with its exchange factor eEF1B, with or without GDP or GDPNP, are closer to the 'active' GTP-bound structure of EF1A ($\sim$25° rotation), than to the GDP- and EF1B-bound conformations of EF1A ($\sim$60° rotation) (Andersen *et al.* 2000, 2001). The structures of archaeal eEF1A in complex with eEF1B, or of eukaryotic eEF1A·GDP are not known, but the sequence identities between archaeal and eukaryotic eEF1A and eEF1B are $\sim$50 and $\sim$20% respectively, indicating, that the corresponding structures are likely to be similar in both kingdoms. It has been suggested, however, that the structure of eukaryotic eEF1A·GDP could be similar to that of the eEF1A·eEF1B complex (Andersen *et al.* 2000, 2001). The truth may be somewhere in the middle, because one group has found indications for a flexible, extended conformation of free eEF1A (Budkevich *et al.* 2002).

Structural data from other members of the family, however, suggest that this large-scale rearrangement could be the exception, rather than the rule (Fig. 3). No significant domain

rearrangements are seen in the structures of another EF1A homolog, eIF2γ in apo-form (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004), GDP-bound, and GTP-bound forms (Schmitt *et al.* 2002), all of which were found to be close ($\sim$15°) to the 'active' EF1A·GTP conformation. A more distant member of the same G protein family, IF2/eIF5B (which will be discussed in detail in Section 5) displays only 8° of rotation between the GTP- and GDP-bound forms (Roll-Mecak *et al.* 2000). A direct implication of these findings is that the relative affinities of the GTP- and GDP-bound forms of the above factors for their ligands need not necessarily be drastically different and, thus, the release of the GDP-bound proteins may not be instantaneous.

Despite the homology between the bacterial and eukaryotic EFs, bacterial EFs cannot work with eukaryotic ribosomes and vice versa. It was found, however, that if the proteins of the L7/L12 stalk (the GTPase-associated center) of the bacterial ribosome were removed *in vitro* and replaced with their eukaryotic counterparts, then these modified bacterial ribosomes could use rat eEF1A and eEF2, but not the bacterial EFs (Uchiumi *et al.* 2002).

In fungi, there is an additional translation elongation factor eEF3, not found in other eukaryotes, which is essential *in vivo* and required for each cycle of elongation *in vitro* (reviewed in Belfield *et al.* 1995). The structure of eEF3 (residues 1–980) from *S. cerevisiae* consists of four domains, including an N-terminal HEAT domain and two ABC domains (Andersen *et al.* 2004).

GTP hydrolysis by EF1A (EF-Tu) is accompanied with $\sim$1000-fold reduction in its affinity for aa-tRNA, whereas eEF1A·GDP retains significant affinity for aa-tRNA (Crechet & Parmeggiani, 1986). Another interesting difference between EF1A and eEF1A is that their GEFs are unrelated and even bind to different regions of the proteins. The GEF for bacterial EF1A is EF1B (EF-Ts), which binds to domains I (the G domain) and III of EF1A (reviewed in Merrick & Nyborg, 2000). eEF1B (formerly eEF1β), the GEF for eEF1A, binds predominantly to domain II of eEF1A, as well as to domain I – from almost the opposite side compared to the binding site of bacterial EF1B (EF-Ts) to EF1A (EF-Tu) (Andersen *et al.* 2000).

3.2.2 Higher order organization of the eukaryotic translational apparatus and channeling of tRNA

Channeling is a phenomenon, characteristic for multi-step enzymic processes, where reaction intermediates are not allowed to diffuse freely in the medium, but are passed on from one active center to the next. Among the benefits of channeling for the overall efficiency of the process are higher effective concentration of the intermediates at the enzyme active site and protection of unstable, highly reactive or insoluble intermediates from contact with the environment (Fersht, 1998).

It has been found, that the translational apparatus in eukaryotes is highly organized, to the extent that most components are not able to diffuse freely out of permeabilized cells. Such permeabilized cells were able to sustain high rates of translation over long periods of time, if supplied with only amino acids and energy sources. The aa-tRNAs are highly sensitive to deacylation and it appears that they are never free in the cell, but instead are transferred from the aaRS directly to eEF1A. What was even more remarkable is that even the relatively stable deacylated tRNAs were not able to diffuse freely in cells from higher eukaryotes (Negrutskii *et al.* 1994; Stapulionis & Deutscher, 1995). Both eEF1A and eEF1B have been found to bind to F-actin. It is not clear to what extent this phenomenon applies to yeast, but there is strong

evidence at least for the channeling of aa-tRNAs from the aaRS directly to eEF1A. As explained above, eEF1B binds tightly to eEF1A·GTP and is only released upon aa-tRNA binding. Furthermore, eEF1A·GTP stimulates the activity of the aaRS (reviewed in Negrutskii & El'skaya, 1998).

The kinetic aspects of eIF2B-catalyzed nucleotide exchange on eIF2 is discussed in more detail in Section 5.4.2. Here, we will discuss nucleotide exchange by eEF1B and EF1B (EF-Ts) in the context of channeling of aa-tRNA. The stable binding of eEF1B to its product, eEF1A·GTP slows down significantly the rate of exchange in the absence of aa-tRNA, indicating that eEF1B is in fact 'optimized' for channeling, and not for working as a stand-alone enzyme. In addition to ensuring protection of aa-tRNA from deacylation, such behavior of eEF1B has other potential advantages: an enzyme cannot change the equilibrium in a reaction, unless it does not dissociate from the product. The higher affinity of eEF1B for the product eEF1A·GTP, than for the substrate eEF1A·GDP, combined with the high concentrations of eEF1B relative to its substrate, indicates that the equilibrium between eEF1A·GTP and eEF1A·GDP is shifted toward the GTP-bound form in the complex eEF1B–eEF1A. Of course, the original equilibrium would be restored if eEF1B dissociated from eEF1A·GTP. This increases the concentration of eEF1A·GTP (in the form of eEF1B–eEF1A·GTP) available for binding to aa-tRNA. Furthermore, eEF1A and all three subunits of eEF1B have been reported to interact with individual aaRSs and it has been proposed that eEF1B has an additional role in facilitating the transfer of aa-tRNA (Bec *et al.* 1994; Sang Lee *et al.* 2002). If eEF1A and the tRNA are brought together before they are even converted into their 'active' forms, eEF1A·GTP and aa-tRNA respectively, then their binding to each other is transformed into a first-order, concentration-independent reaction. To complete the cycle, eEF1A·GDP binds to both deacylated tRNA and to aaRS with nanomolar affinity (Petrushenko *et al.* 2002). Thus, eEF1A·GDP can re-bind the tRNA and assist in its delivery to the aaRS.

It is clear, that channeling in translation cannot be absolute, because some steps require a certain degree of diffusion. One such example is the delivery of aa-tRNA to the ribosome by the universal eEF1A, where the 'correct' and 'incorrect' eEF1A·GTP·aa-tRNA complexes bind randomly to the ribosome and can be distinguished only after binding to the ribosome. One can only speculate what additional benefits can be obtained from the organization of the translational apparatus in higher order structures. One such possibility is that, if a tRNA exiting the ribosomal E-site is picked by the corresponding aaRS and aminoacylated, the resulting aa-tRNA will be transferred back to eEF1A·GTP in the vicinity of the same codon of mRNA, to which it was basepaired in the previous cycle. Then, the probability of that same aa-tRNA binding to the same codon in the context of the next ribosome is increased. The result of this purely hypothetical scenario is that the frequency of futile cycles of binding and release of 'incorrect' aa-tRNAs may be decreased and the overall efficiency of translation increased.

The situation is quite different with bacterial EF1A (EF-Tu) and EF1B (EF-Ts). EF1B appears 'optimized' for stand-alone operation, because its dissociation from the product EF1A·GTP is not much slower than that from the substrate EF1A·GDP, although it still is ~5-fold slower (Gromadski *et al.* 2002). There are no reports, to our knowledge, of binding of EF1B to any aaRS (of course, the presence of weak transient interactions cannot be excluded). aa-tRNAs in bacteria are obviously as sensitive to deacylation as they are in eukaryotes and it is ensured that they are at least predominantly protein-bound. It appears, however,

that this is done 'kinetically' – through quick binding to EF1A·GTP, rather than via physical channeling. It would be interesting to know (but maybe hard to test) whether EF1A·GTP and the aa-tRNA could bind to each other before being released from EF1B and aaRS respectively.

The structures of the eEF1A·eEF1B complex (Andersen *et al.* 2000) and the EF1A·EF1B (EF-Tu·EF-Ts) complex (Kawashima *et al.* 1996) provide an explanation for the different properties of the two GEFs. eEF1B stabilizes an active-like conformation, similar to that of EF1A·GTP (Andersen *et al.* 2000), thus 'preparing' it for GTP binding (Fig. 3). Accordingly, eEF1B binds more tightly to the 'product' eEF1A·GTP, than to the 'substrate' eEF1A·GDP (Crechet & Parmeggiani, 1986; Janssen & Moller, 1988). This is not the case in bacteria, where EF1A in the EF1A·EF1B resembles free EF1A (Kawashima *et al.* 1996), and EF1B binds almost equally well to EF1A·GTP and EF1A·GDP (Gromadski *et al.* 2002).

3.3 Translation termination and recycling

Translation termination is the process of recognition of an in-frame stop codon in the mRNA, release of the nascent polypeptide and dissociation of the ribosomal complexes.

Recognition of a stop codon in the A-site is performed by two 'class I' release factors in bacteria, RF1 and RF2. RF1 recognizes the UAA and UAG stop codons, whereas RF2 recognizes UAA and UGA. Eukaryotes have only one class I termination factor eRF1, which recognizes all three stop codons. A class II release factor, RF3·GDP in bacteria, binds to the ribosome in the presence of a class I factor. Upon release of the nascent peptide, the GDP bound to RF3 is exchanged to GTP, accompanied by conformational changes and dissociation of the class I factor. GTP hydrolysis, in turn, causes dissociation of RF3. A ribosome-recycling factor, RRF, together with EF2 (formerly EF-G) then completes the process by dissociating the two subunits. Eukaryotes do not have an RRF, but unlike the non-essential bacterial RF3, eRF3 is essential and could also be fulfilling the role of an RRF (reviewed in Merrick & Nyborg, 2000; Ramakrishnan, 2002; Kisselev *et al.* 2003).

4. Translation initiation

In this section, we discuss the process of translation initiation: its mechanism and the roles of individual IFs (Fig. 2, Tables 1 and 2). The next section is dedicated to the structural aspects of initiation, whereas the last section contains a view at translation initiation from an enzyme kinetics perspective. Although we have tried to discuss every topic only once, in the most appropriate context, a certain degree of redundancy was inevitable. Some of the highlights of this section are as follows: (1) eIF4F remains associated with the IC during scanning and even transiently during elongation. In more general terms, IFs appear to be recruited earlier than previously thought and to be released much later than previously thought: factors are often displaced from their original locations, but remain associated with the IC, although more weakly. (2) The affinity of eIF2·GDP for the initiator tRNA (Met-tRNA$_i$) is not much lower than that of eIF2·GTP and the difference appears to involve mainly recognition of the Met moiety by eIF2·GTP. (3) The IFs do not directly 'inspect' the identity of the start codon and recognition is mediated by the complementarity and geometry of the interaction of the initiator tRNA and mRNA in the context of the ribosome.

Table 1. *Functions in translation initiation and the factors involved*

Task	Bacteria	Archaea	Eukaryotes
Subunit dissociation	**IF3**	**eIF1**(?)	**eIF3**
and anti-association	IF1	eIF1A(?), eIF2(?)	eIF1, eIF1A (=IF1), eIF2
Initiator tRNA binding	**IF2**, **IF3**	**eIF2**, **eIF1**	**eIF2**, **eIF1**
and selection		eIF5B(?)	eIF5B(?)
mRNA binding by the	**IF1**, **IF3**	**eIF1A** (=**IF1**)	**eIF1, eIF1A, eIF3,**
ribosome, and scanning	n.a.	n.a.	**eIF4E, G, A, B, H, PABP**
Start site selection	**IF3**	**eIF1, eIF2**	**eIF1, eIF2, eIF5,**
			eIF3, eIF4G
Subunit joining	**IF2**	**eIF5B** (=**IF2**)	**eIF5B** (=**IF2**)
	IF1(?)	eIF1A(?)	eIF1A(?)

Factors in bold have a principal role.
(?) indicates function not proven.

Table 2. *Interactions involving translation initiation factors in eukaryotes*[a]

		mRNA	tRNA	1[a]	1A	2	2B	3	4A	4B	4E	4G	4H	5	5B	PABP
Rib.		++[b]	++	+	+±	++	±	++						+	++	
mRNA	++		+			±		++	+	+	+	++	+			+
tRNA	++	+		±		++									±	
1	+		±			+		++[c]				+		+		
1A	+±					+		+						±	++	
2	++	±	++	+	+		++	+						+±		
2B	±					++										
3	++	++		++	+	+				+		+		++		
4A		+								±		++	±			
4B		+						+	±							+
4E		+										+				
4G		++		+				+	++		+			+		+
4H		+							±							
5	+			+	±	+±		++				+			+	
5B	++		±		++									+		
PABP		+								+		+				
Other[d]	++	++	+			++	++	++	++		++	++			+	++

[a] 'eIF' is omitted from the names of the factors.
[b] '+', One reported interaction; '++', more than one interactions; ±, interaction not proven.
[c] Interactions reported in at least one species are listed, whether present in all eukaryotes or not.
[d] Interactions with other proteins, whether with or without a direct role in translation. The list is likely incomplete and some of the interactions may not be referenced in the text.

4.1 General translation initiation factors

The relationships between bacterial, archaeal and eukaryotic IFs can be found in Table 1 and Fig. 2, and will be discussed in depth in Section 5. All archaeal initiation factors (aIFs) have eukaryotic homologs and are designated in the literature as either aIFs or eIFs. In this review, we will call them eIFs, both for simplicity and to underline the relationship with their eukaryotic homologs.

Briefly, bacterial IF1 is homologous to eIF1A, IF2 is homologous to eIF5B, and IF3 has no archaeal or eukaryotic homolog. IF3 was previously thought to correspond to eIF3, but recent data clearly indicate that it is functionally similar to eIF1, present in both archaea and eukaryotes. Some bacteria have an eIF1 homolog (YciH in *E. coli*), but its function is unclear and its absence from a number of bacterial species makes it an unlikely candidate for a general IF.

Archaea and eukaryotes also have eIF2 (unrelated to IF2). There is an eIF4A homolog in all kingdoms, named W2 (not related to the W2 domain found in some eIFs), or more recently IF4A in bacteria. However, the role of IF4A/eIF4A in bacteria and archaea has not been completely understood. A number of IFs have only been found in eukaryotes: eIF2B, eIF3, eIFs 4B, 4E, 4G, 4H, eIF5 and the poly-A-binding protein (PABP).

The distinction between general IFs and proteins with a regulatory role in translation is not always clear-cut. The function of a protein called eIF2A is unclear, but its deletion in yeast has no obvious phenotype and it is thus unlikely to act as a general IF (Zoll *et al.* 2002). On the other hand, PABP, which is not always considered a general IF, has an important role on polyadenylated mRNAs, which are the majority of the mRNAs in the cell, and is considered here a general IF.

Eukaryotic initiation factor 5A (eIF5A) and its bacterial homolog, elongation factor P (EF-P) are essential factors with a number of roles in translation. One role shared between the bacterial and eukaryotic homologs is stimulation of peptide bond formation, especially the first peptide bond, in an amino acid-dependent manner (reviewed in Ganoza *et al.* 2002). eIF6 is found in archaea and eukaryotes and appears to be involved in ribosome biogenesis and (in eukaryotes) nuclear export and signaling (Ceci *et al.* 2003).

4.2 Subunit dissociation/anti-association

Dissociation of ribosomes at the end of termination is an active process involving a combination of termination, elongation and initiation factors. Anti-association activity involves binding to already dissociated subunits and prevention of subunit association, and is mainly mediated by IFs. The role of anti-association is not only to provide a pool of ribosomal subunits for initiation, but also to prevent premature assembly of translationally inactive ribosomes during initiation as well as ribosome assembly at an incorrect site.

In bacteria, IF3 is responsible for subunit dissociation and anti-association. Its C-terminal domain (IF3-CTD) binds to the subunit interface surface of the platform of the small subunit, thus directly competing with the large subunit for binding. Binding of IF3-CTD to the small subunit is stabilized by the N-terminal domain (IF3-NTD) and/or the inter-domain linker (Dallas & Noller, 2001). The binding of IFs 1, 2, 3 and the initiator tRNA fMet-tRNA^{fMet} to the small ribosomal subunit is strongly cooperative (Weiel & Hershey, 1982; Zucker & Hershey, 1986) and, therefore, all other factors also contribute to IF3's anti-association activity. In eukaryotes, eIF3 (unrelated to IF3, see above and Table 2) carries the main subunit dissociation and anti-association function, and similarly to bacteria, its activity (and affinity for the small subunit) is enhanced in the presence of other factors: eIF1, eIF2·GTP·Met-tRNA$_i$, mRNA and small U-rich RNAs (Kolupaeva *et al.* 2005). eIF1 (unrelated to IF1 or IF3) binds to the same surface of the small subunit as IF3-CTD in bacteria, and its binding is cooperative with eIF3 binding (Pestova & Kolupaeva, 2002; Lomakin *et al.* 2003). Binding of eIF1 and eIF1A (homologous to IF1) to the small ribosomal subunit is cooperative (Maag & Lorsch, 2003)

and, thus, eIF1A can also indirectly promote anti-association. The ternary complex of eIF2 (unrelated to IF2) with GTP and Met-tRNA$_i$ stabilizes eIF3 binding and also provides a steric block against subunit joining. As archaea do not have eIF3, subunit anti-association would have to rely on binding of eIF1, eIF1A and the ternary complex eIF2·GTP·Met-tRNA$_i$, with eIF1 and the ternary complex providing a steric block against subunit association.

4.3 Initiator aa-tRNA recognition

The selection of the initiator tRNA involves several tasks: (1) recognition of the initiator tRNA (proper charging) by the aaRS; (2) discrimination against the initiator tRNA by EFs; (3) discrimination against uncharged or mischarged initiator tRNA by IFs; and (4) discrimination against elongator tRNAs by IFs.

The first task involves interaction of the aaRS with not only the acceptor end of the tRNA, but also directly with the anticodon. No elongation or IF directly binds the anticodon of the tRNA. In bacteria, there is a second enzymic step – formylation of Met-tRNA^{fMet}, which has multiple roles: it increases the stability of the resulting fMet-tRNA^{fMet} to deacylation, reduces the affinity for EF1A (task 2) and increases the affinity for IF2 (task 3). In eukaryotes, task 2 is accomplished by post-transcriptional modification of tRNA$_i$. Recognition of properly charged initiator tRNA (tasks 3 and 4) is described below.

In bacteria, the initiator tRNA is tRNA^{fMet}, and in eukaryotes it is tRNA$_i$ (or tRNA$_i^{Met}$) – specific for methionine, but distinct from the methionine-specific elongator tRNA (tRNA^{Met}). The initiator tRNA must be charged with the correct amino acid, formylmethionine (fMet-tRNA^{fMet}) in bacteria, and methionine (Met-tRNA$_i$) in eukaryotes. In bacteria, formylation increases the stability of fMet-tRNA^{fMet} to deacylation, whereas the free eukaryotic Met-tRNA$_i$ and the elongator aa-tRNAs are fairly unstable.

In bacteria, selection for fMet-tRNA^{fMet} occurs at two levels. The G protein IF2 binds specifically to the fMet moiety and the acceptor end of the tRNA with moderate affinity (K_D of $\sim$0·5 μM) and has negligibly low affinity (K_D $>$1 mM) for the deacylated tRNA and free fMet (Guenneugues *et al.* 2000). IF2 and fMet-tRNA^{fMet} bind cooperatively to the small ribosomal subunit. Whereas GTP or GDP binding and fMet-tRNA^{fMet} binding to free IF2 are independent in solution, off the ribosome, it was reported that IF2·GTP stabilizes fMet-tRNA^{fMet} binding to the small subunit more than IF2·GDP does (Antoun *et al.* 2003).

IF3 is also involved in tRNA^{fMet} selection, but does not distinguish between acylated and deacylated tRNA^{fMet}. IF3 does, however discriminate against elongator tRNAs and has been proposed to act indirectly – through the ribosome. The initiator tRNA^{fMet} has three conserved GC base pairs (nucleotides 29–31 and 39–41) in the anticodon stem, which, upon IF3-induced conformational changes in the small subunit, could be directly inspected by the conserved nucleotides G1338 and A1339 in the head (Dallas & Noller, 2001).

In archaea and eukaryotes, eIF2 (unrelated to IF2) is responsible for selection and recruitment of Met-tRNA$_i$ to the ribosome. eIF2 is a heterotrimer and its biggest, γ subunit is homologous to EF1A, eEF1A and SelB/eEFSec (Leibundgut *et al.* 2005). Like the EFs, eIF2 binds to Met-tRNA$_i$ more tightly in its GTP-bound form, recognizing determinants from both the tRNA and the Met moiety. As discussed in the previous section, the distinction between 'specific' and 'non-specific' binding is only quantitative and aa-tRNA recognition by the universal factors, EF1A/eEF1A and specific factors such as SelB/eEFSec is similar. The affinity of eIF2 for Met-tRNA$_i$ is $\sim$10 nM (Kapp & Lorsch, 2004), much higher than the affinity of bacterial

(structurally unrelated) IF2 for fMet-tRNA^{fMet} and the release of eIF2 from Met-tRNA$_i$ requires GTP hydrolysis by eIF2. The affinity of eIF2 for Met-tRNA$_i$ drops only ~20-fold upon GTP hydrolysis, with ~100-fold increase in dissociation rate; therefore, eIF2·GDP dissociation from Met-tRNA$_i$ may not be instantaneous (Kapp & Lorsch, 2004). After eIF2 hydrolyzes GTP and is released from the IC, the acceptor end of Met-tRNA$_i$ is free and could potentially interact with eIF5B (the eukaryotic homolog of IF2), which at that stage needs to be properly oriented to promote subunit joining (see Section 4.5). Such an interaction could stabilize Met-tRNA$_i$ binding and provide an additional level of selection for Met-tRNA for the time interval between eIF2 release and subunit joining. Unfortunately, direct binding between eIF5B and Met-tRNA$_i$ has not been reported and there is only indirect data in support of such an interaction (Choi *et al.* 1998, 2000; Marintchev *et al.* 2003). The putative interaction of eIF5B with Met-tRNA$_i$ would be expected to be weak, in order for eIF5B not to compete with eIF2 off the ribosome. An additional difficulty is that Met-tRNA$_i$ is both unstable and difficult to prepare and label in large amounts.

There is no indication that eIF1, which is a functional analog of IF3, is involved in direct or indirect selection for tRNA$_i$. However, the GC base pairs in the initiator tRNA and the corresponding nucleotides in the head of the small subunit, proposed to be involved in initiator tRNA selection (see above) are conserved between bacteria and eukaryotes, raising the possibility that at some stage, or under specific conditions, the eukaryotic small subunit could be involved in discrimination against elongator tRNAs.

4.4 Start site recognition

4.4.1 Overview

This point in translation initiation is the most divergent among kingdoms and one of the reasons why for a long time it was thought that eukaryotic and bacterial translation initiation are unrelated. Bacteria have a conserved sequence motif, called Shine–Dalgarno (SD) or ribosome-binding site (RBS), several nucleotides upstream of the start codon. The SD is complementary to the 3′-end of the 16S rRNA and the 5–7 nt spacer allows the start codon to be positioned in the P-site of the decoding center of the small subunit. The mRNAs in eubacteria usually contain more than one open reading frame (ORF), i.e. encode more than one protein, and the ribosomes assemble directly at the translation start sites. Regulation of translation initiation in bacteria is usually mediated by mRNA secondary structures and proteins binding at or near the SD element or the start codon (reviewed in Hershey & Merrick, 2000; Jackson, 2000). An interesting 'riboswitch control' mechanism was discovered recently, where small molecules bind directly to the 5′-leader sequence of mRNAs encoding enzymes involved in their metabolism, causing rearrangement of secondary structures. Depending on the position and nature of the affected secondary structure elements, the metabolites stimulate or inhibit the expression of the enzymes at the level of transcription or translation initiation (reviewed in Nudler & Mironov, 2004). No structures are available yet for riboswitches regulating translation initiation, but the first structure of a riboswitch that regulates transcription was recently published: the purine-binding domain of the guanine riboswitch in complex with hypoxanthine (Batey *et al.* 2004).

In eukaryotes, the majority of mRNAs contain only one ORF; their 5′-end is 'capped' with an m^{7}G-cap through a reverse, 5′—5′ bond; and they have a poly-A tail at their 3′-end. Both the 5′-cap and the 3′-poly-A tail are important for efficient translation. The 43S IC is first recruited to the 5′-cap through interactions with a set of eIFs, called eIF4F or cap-binding complex. The

IC then scans along the mRNA until the start codon. This is usually, but not always, the first AUG, as the nucleotide context around the AUG significantly influences initiation efficiency. The optimal sequence context for the AUG start codon in higher eukaryotes is GCCA/GCCAUGG, (Kozak, 1986, 1987). The most important nucleotides are the A or G at position −3 (where the A of the AUG codon is +1) and the G at +4. The consensus sequence context in plants and other eukaryotes is similar to that in vertebrates, although it may be quite different in some organisms, such as *S. cerevisiae* (reviewed in Kozak, 1991). The length and the presence or absence of secondary structure in the 5′-untranslated region (5′-UTR) are major determinants of translation efficiency. In addition to the 3′-poly-A tail, sequences in the 3′-untranslated region (3′-UTR) often regulate translation, usually serving as binding sites for translation regulators. Regulatory proteins binding to the 5′-UTR have also been found. A number of eukaryotic mRNAs, including many viral mRNAs, have an alternative mode of translation: the ribosome is recruited directly to an internal site at or near the start codon, called internal ribosome entry site (IRES). The IRESs can be fairly long and structured RNA segments. There are several types of IRES, differing in both length and structure of the RNA. Translation initiation at an IRES is typically independent of the presence of a 5′-cap and requires only a subset of the eIFs, involved in canonical cap-dependent translation initiation. The factor requirements vary among different types of IRESs (reviewed in Jackson, 2000; Pestova *et al.* 2001). On the extreme is the case of the cricket paralysis virus IRES, which does not even require Met-tRNA$_i$ (Pestova & Hellen, 2003). In addition to the canonical and IRES-dependent translation initiation, the scanning IC has been found to skip RNA segments and resume scanning on certain mRNAs. Since the 3′-end of the eukaryotic 18S rRNA is not complementary to the bacterial SD sequence (or to any sequence near the start site of the eukaryotic mRNA), the eukaryotic apparatus cannot recognize a bacterial translation start site. Similarly, bacteria cannot recognize the start signals in eukaryotic mRNAs. For an in-depth review of start site recognition mechanisms in bacteria and eukaryotes, and their variations see Jackson (2000).

Archaeal translation is not as well studied as translation in eubacteria and eukaryotes. Like eubacteria, archaea have SD-like sequences and polycistronic mRNAs, which are not capped or polyadenylated. However, it appears that in most mRNAs, the first start codon is at or near the 5′-end and is not preceded by an SD sequence. SD elements are used mainly in polycistronic mRNAs for initiation at internal start sites. mRNAs, where the start site is at or within a few nucleotides from the 5′-end, are called 'leaderless' mRNAs. They have also been found in eubacteria and eukaryotes, and, unlike other types of mRNA, leaderless mRNAs can be translated in cell extracts derived from all kingdoms (reviewed in Moll *et al.* 2002). The properties of leaderless mRNAs make them particularly interesting from evolutionary and mechanistic perspectives and will be discussed in more detail in Section 4.7 below.

4.4.2 Roles of individual factors

In addition to the small ribosomal subunit, which recognizes the SD element, bacterial IF3 has a central role in start site selection: it can dissociate ICs assembled on non-canonical codons or on canonical codons located at or near the 5′-end. Although AUG is the predominant start codon in bacteria (∼90%), GUG and UUG are also considered 'canonical' and represent ∼8% and ∼1% of the start codons respectively. IF3 cannot dissociate ICs preformed on canonical start sites and is released upon start site recognition (reviewed in Hershey & Merrick, 2000). This indicates that the small subunit can spontaneously undergo the conformational changes

associated with start site recognition, leading to a state with lower affinity for IF3. The discrimination between 'good' and 'bad' start sites is indirect and based on the balance between the stabilities of the two alternative conformations: IF3 is able to reverse the changes in the IC, unless they are stabilized by proper interactions of the small subunit, the initiator tRNA, and the mRNA (Dallas & Noller, 2001). As no IF in either bacteria or eukaryotes directly 'inspects' the start codon, the identity of AUG as the start codon is defined by the anticodon of the initiator tRNA (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000).

In eukaryotes, the dynamic discrimination between 'good' and 'bad' start sites is taken to a new level of complexity and multiple factors are involved. As translation initiation on most mRNAs involves scanning from the 5′-cap to the start codon, selection needs to be achieved in the context of an IC sliding along mRNA and efficient discrimination against incorrect initiation sites must be sustained over longer periods of time. The processes of 5′-cap and 3′-poly-A recognition and scanning do not have bacterial or archaeal counterparts and will be discussed separately in Section 4.6 below.

eIF1 discriminates against non-AUG codons and its mechanism of action is probably similar to that proposed for IF3 (see above), as is its binding site on the small ribosomal subunit (Lomakin *et al.* 2003). In eukaryotes, only AUG is a 'canonical' start codon, but UUG and GUG appear to be the preferred 'non-canonical' start codons (see Section 4·9 below). eIF1 discriminates also between 'good' and 'bad' nucleotide context of the start codon (the 'Kozak' consensus element). This function of eIF1 is performed indirectly – via conformational changes in the small ribosomal subunit and may not be based on sequence-specific recognition by the ribosome, but rather on the propensity of the mRNA segment for the conformation required to fit in the mRNA-binding groove of the small subunit. eIF1, similarly to bacterial IF3, also prevents formation of ICs on AUG codons located at or near the 5′-end (reviewed in Hershey & Merrick, 2000; Pestova *et al.* 2001).

Other factors directly involved in start codon selection are eIF2 and eIF5. eIF2 in its GTP-bound form brings Met-tRNA$_i$ to the IC. Upon start codon recognition, eIF2 hydrolyzes GTP and is subsequently released from the IC. In eukaryotes, GTP hydrolysis requires the presence of eIF5, which serves as a GTPase-activating protein (GAP) for eIF2. GTP hydrolysis and the subsequent release of eIF2 are required for progression of the ICs toward the next stage – subunit joining. Accordingly, an increase in the rate of GTP hydrolysis by eIF2 or decrease of its affinity for Met-tRNA$_i$ leads to higher error rates of initiation (reviewed in Donahue, 2000; Hershey & Merrick, 2000; Pestova *et al.* 2001). Conversely, any factor that stabilizes binding of eIF1 to the IC, like eIF3, for example, would be expected to promote higher fidelity of initiation.

As mentioned above, archaea have SD-like elements, but also many leaderless mRNAs. Furthermore, although archaea have eIF1 and eIF2, they do not have eIF3 and eIF5. The absence of eIF3 suggests that eIF1 binding to the IC may be weaker than in eukaryotes. Consequently, eIF1-mediated start site discrimination could be less efficient.

A parallel between archaea and eukaryotes suggests that eIF2 GTP hydrolysis and nucleotide exchange are more stringently controlled in eukaryotes, hence the need for a GAP (eIF5) and a GEF (eIF2B). Given the high degree of homology between eukaryotic and archaeal eIF2, the eukaryotic factor probably has retained the intrinsic ability to hydrolyze GTP, but this activity is efficiently repressed. According to this scenario, an alternative role of eIF5 could be to help derepress the GTPase activity of eIF2, rather than (or in addition to) acting as a classical GAP factor, stabilizing a transition state in GTP hydrolysis. In support of such an interpretation,

eIF5 is homologous to the two core domains of eIF2β, covering almost the entire length of archaeal eIF2β, except for a short eIF2γ-binding segment that is absent in eIF5 (eukaryotic eIF2β and eIF5 also have additional, non-homologous segments, responsible for mutual binding, see Section 5). Furthermore, mutations in, or deletion of the second of the eIF2β domains shared with eIF5 cause increased rate of spontaneous GTP hydrolysis by eIF2 (Donahue, 2000; Hashimoto *et al.* 2002).

4.5 Subunit joining and factor release

Subunit joining depends on the proper orientation of the initiator aa-tRNA, achieved upon start codon recognition. A small subunit with an initiator aa-tRNA in the P-site base-paired to the start AUG codon of mRNA can bind to the large subunit and form a translationally active ribosome, in the absence of additional factors. In all kingdoms subunit joining is promoted by a universally conserved G-protein, called IF2 in bacteria and eIF5B in eukaryotes, which, like the initiator tRNA, needs to be properly positioned. The other universally conserved factor, IF1/eIF1A (in bacteria and eukaryotes respectively), when bound alone to the small subunit, stimulates the rates of both subunit joining and dissociation, and also stabilizes binding of IF2/eIF5B to the small subunit. The two factors appear to be coordinately released after subunit joining in all kingdoms (Benne *et al.* 1973; Choi *et al.* 2000; Olsen *et al.* 2003) and therefore IF1/eIF1A is likely also involved in subunit joining at the end of translation initiation.

In addition to proper positioning of the initiator tRNA, subunit joining requires certain IFs to be released or at least displaced from their original location. The event that triggers factor release in all kingdoms is start site recognition. As explained above, both bacterial IF3 and eukaryotic eIF1 bind to the interface surface of the small subunit and need to be displaced from there before subunit joining. In eukaryotes, start site recognition also induces GTP hydrolysis by eIF2 and release of eIF2·GDP. There is no obvious need for the IFs to physically dissociate from the 40S subunit in order for subunit joining to occur, as long as they do not block the interaction with the 60S subunit either directly or indirectly. After subunit joining, eIFs 1, 2, 3 and 5 do not co-sediment with the ICs in sucrose gradient centrifugation and could already be released at that stage. Alternatively, the remaining interactions are too weak to withstand centrifugation. For example, eIF3 is involved in several interactions with the 40S subunit, some of which involve solvent-exposed surfaces and can be retained even after subunit joining. Therefore, eIF3 and other eIFs could still be associated with the ICs during and even after subunit joining (Jackson, 2000). eIF3 also interacts with mRNA, with affinity dependent on RNA structure and/or sequence. It was found that eIF3 stays associated with the 40S subunit after the release of eIF2, and eIF1 could also be associated with such complexes through its interaction with eIF3 (Unbehaun *et al.* 2004). Recently, direct evidence for transient presence of IFs after the onset of elongation was obtained in mammals (Poyry *et al.* 2004), as well as more indirect indication in yeast (Rajkowitsch *et al.* 2004). Both reports demonstrated time-dependence of the ability of ribosomes to reinitiate after translating a short ORF. The former group demonstrated direct requirement for the presence of eIF4F or at least the central domain of eIF4G and eIF4A in the first IC (for the short ORF) and that these eIFs could not bind after the short ORF was already translated. As eIF4G binding to the 48S IC is mediated by eIF3, the above results implied that eIF3, and possibly most other factors associated with it, can remain on the 80S for a short period of time. Clearly, eIF2·GTP·Met-tRNA$_i$ needs to be

regenerated from eIF2·GDP and $tRNA_i$, before reinitiation can occur. There was no apparent need for either eIF4E or PABP (Poyry *et al.* 2004). As discussed in Section 4.6 below, it is not clear if and with what rates eIF4E dissociates from the cap and/or eIF4G. There is also no indication whether binding to eIF4E or PABP affects the association of eIF4G with the IC during scanning, at the start codon or after subunit joining. There are somewhat controversial reports that eIF2 stays on the 40S subunit even after being released from Met-$tRNA_i$, and is later transferred to the 60S subunit (Ramaiah *et al.* 1992), but other authors found the association to be unstable under physiological conditions (Chakrabarti & Maitra, 1992). In view of the above results, transient association of eIF2·GDP with the ribosomes does not seem surprising.

The dependence of reinitiation on the presence of a set of eIFs on the terminating ribosome provides an explanation for the rather unexpected finding that deletion or mutations of eIF5B that slow down subunit joining or release of eIF5B after subunit joining, inhibit reinitiation (Lee *et al.* 2002; Shin *et al.* 2002). Both slow subunit joining and slow eIF5B release delay the onset of elongation and extend the time interval between GTP hydrolysis by eIF2 and termination, thus allowing the eIFs to dissociate. This interpretation relies on the assumption that the association of at least part of the eIFs required for reinitiation with the 48S complex is destabilized upon GTP hydrolysis by eIF2, which is supported by the different stability and composition of 48S complexes subjected to centrifugation before and after GTP hydrolysis by eIF2.

IF2/eIF5B and IF1/eIF1A are coordinately released after subunit joining (Benne *et al.* 1973; Choi *et al.* 2000; Olsen *et al.* 2003). The release is triggered by GTP hydrolysis by IF2/eIF5B. The GTPase activity of IF2/eIF5B (like the GTPase activities of the elongation factors EF1A/eEF1A and EF2/eEF2) is stimulated by interaction with the GAC on the large ribosomal subunit. IF2/eIF5B·GDP has lower affinity for the ribosome than IF2/eIF5B·GTP and dissociates quickly (Benne *et al.* 1973; Pestova *et al.* 2000; Lee *et al.* 2002; Shin *et al.* 2002; Antoun *et al.* 2003).

One kinetic study using *E. coli* ribosomes and factors and *B. stearothermophilus* IF2 found no difference in the activity of IF2·GTP *versus* IF2·GDP, and no role of GTP hydrolysis in release of IF2 upon subunit joining, contradicting previous data and casting doubts over the similarity of subunit joining between bacteria and eukaryotes (Tomsic *et al.* 2000). A recent report from the Ehrenberg group reconfirmed the importance of GTP binding and hydrolysis by IF2 for subunit joining (Antoun *et al.* 2003). It was proposed that the conflicting results in (Tomsic *et al.* 2000) may have been due to possible contamination with GTP during the preparation of IF2 (Antoun *et al.* 2003). IF2 from the thermophilic *B. stearothermophilus*, used in (Tomsic *et al.* 2000), is known to bind more tightly to the *E. coli* ribosome than the endogenous IF2 from *E. coli* (Brombach *et al.* 1986; Severini *et al.* 1990). The use of a foreign protein with high affinity for the *E. coli* ribosome is convenient for biochemical assays (Brombach *et al.* 1986; Severini *et al.* 1990), but hardly appropriate for kinetic analysis of translation initiation in *E. coli*, which affects the validity of the obtained results. A detailed explanation of the expected effects of such an IF2 variant/mutant with high affinity for the ribosome in its GDP-bound state can be found in Antoun *et al.* (2003). Briefly, if a mutant (or exogenous) IF2·GDP binds too tightly to the ribosome, it will have a slower rate of dissociation after subunit joining, which can, depending on the 50S subunit concentration, make IF2·GDP dissociation the rate-limiting step. The end result is that no difference in the activity of the mutant IF2·GTP and IF2·GDP will be observed, because the processes will be equally slow.

4.6 Processes specific for eukaryotic translation initiation

A distinct feature of most eukaryotic mRNAs is the presence at their 5′-end of an m^7G-cap formed through a reverse, 5′–5′ bond, and of a poly-A tail at their 3′-end. The cap and the poly-A tail are brought together by protein factors and play important roles in mRNA stability to degradation and recruitment of the 43S pre-IC. On the vast majority of eukaryotic mRNAs, the small ribosomal subunit is first recruited to the 5′-cap-proximal region, after which it scans along the mRNA in a 5′–3′ direction, with the help of IFs, until the start codon is reached.

4.6.1 Cap binding and scanning

Capping of mRNA serves several important functions. It allows discrimination between intact and 5′-truncated mRNAs. Furthermore, it provides protection from nucleolytic degradation by blocking the 5′-end. The cap-binding complex stimulates recruitment of the small ribosomal subunit to a 5′-proximal location immediately downstream from the cap. In that respect, cap-dependent translation initiation resembles bacterial SD-dependent initiation and eukaryotic IRES-dependent initiation, except that upon recruitment to mRNA, the IC starts scanning. Most of the eukaryote-specific IFs, not found in archaea, are involved directly or indirectly in 5′-cap recognition, 3′-poly-A recognition and scanning. The eIF4F complex, composed of the cap-binding protein eIF4E, the scaffold protein eIF4G and the ATP-dependent RNA helicase eIF4A, assembles on the 5′-cap. The assembly is mediated by eIF4G, which has binding sites for both eIF4E and eIF4A. The recruitment of the small subunit to the cap is mediated by an interaction between eIF4G and eIF3 in human. No such interaction has been reported in *S. cerevisiae*, where eIF4G was instead reported to bind to eIF5 and eIF1 in the 43S complex. If the mRNA 3′-end is polyadenylated, eIF4G can bridge the 5′- and the 3′-ends via an interaction with the PABP. The interactions of eIF4E with the cap and eIF4G, and of eIF4G with eIF4E and PABP have been reported to be cooperative (Haghighat & Sonenberg, 1997; Ptushkina *et al.* 1998; Wei *et al.* 1998; Luo & Goss, 2001). The eIF4A helicase can disrupt secondary structure (if any) in the vicinity of the 5′-cap, thus promoting binding by the small ribosomal subunit. eIF4A is also required for scanning on mRNAs with structured 5′-UTRs and stimulates scanning on unstructured 5′-UTRs. eIF4A alone is not a processive helicase. Its activity is higher when part of the eIF4F complex, and is further stimulated by eIF4B and eIF4H. eIF4E, 4G, 4B, and 4H all bind RNA and it is not clear if stimulation of eIF4A helicase activity by eIF4G is due to anchoring to mRNA and/or to changes in the conformation of eIF4A itself. It has been found that eIF4F-bound eIF4A can exchange with free eIF4A (Yoder-Hill *et al.* 1993), and the stability of association of eIF4A with eIF4G is not constant during the process of RNA unwinding, indicating that eIF4A could shuttle in and out of eIF4F (Pause *et al.* 1994b), or that at least some of the interactions between eIF4G and eIF4A may be dynamically lost and regained. eIFs 4B and 4H have ssRNA binding and annealing activities, reminiscent of the ssDNA-binding proteins (SSB), known to stimulate the activity of DNA helicases by stabilizing the ssDNA. Curiously, while free eIF4A can translocate along RNA bidirectionally, scanning appears to be unidirectional from 5′ to 3′ (reviewed in Hershey & Merrick, 2000; Jackson, 2000; Preiss & Hentze, 2003), although it has been noted, that unidirectional scanning is difficult to distinguish from scanning that is only predominantly unidirectional (Berthelot *et al.* 2004).

A subset of the IFs (eIF2·GTP·Met-tRNA$_i$ and eIFs 1, 1A, 3, 5, and 5B) assembles on the 40S ribosomal subunit to form the 43S complex, while the cap-binding complex assembles at the 5′-end of mRNA (Fig. 2). eIFs 1, 3, 5 and the eIF2·GTP·Met-tRNA$_i$ ternary complex have been reported to form a complex in yeast off the ribosome and are likely coordinately recruited. eIF1A appears to bind directly to the small subunit and its affinity (~30 nM) is sufficient to ensure binding under physiological concentrations, which are estimated to be in the micromolar range. Furthermore, binding of eIF1A is stabilized in the presence of eIF1 (Maag & Lorsch, 2003), and this cooperativity is likely maintained when eIF1 is part of the large eIF1,3,5,eIF2·GTP·Met-tRNA$_i$ complex. In yeast, eIF5B was found to bind to eIF1A in solution and the two factors were proposed to be recruited coordinately to the small subunit (Choi *et al.* 2000; Olsen *et al.* 2003). Human eIF5B appears to bind to eIF1A less tightly than its yeast homolog (Marintchev *et al.* 2003), but instead was found to be associated with eIF5 and is likely to be recruited together with it, instead (A. Marintchev, T. V. Pestova & G. Wagner, unpublished results).

Once the small subunit is recruited to mRNA near the 5′-cap, it can start scanning. Both eIF1A and eIF1 are required for efficient scanning. Like its bacterial homolog IF1, eIF1A stimulates mRNA binding to the small ribosomal subunit. And like its bacterial analog IF3, eIF1 discriminates against 'incorrect' start sites. However, whereas the bacterial small subunit is bound at the SD element, the eukaryotic small subunit scans along from the 5′-end. eIF1A ensures stable association with mRNA and thus processive scanning, and eIF1 prevents the IC from remaining 'stuck' at the 5′-end or from stopping at an incorrect codon. eIF3 stabilizes binding of eIF1 to the IC and can bind directly to mRNA, providing additional stabilization of the scanning complex on mRNA. The eIF2·GTP·Met-tRNA$_i$ ternary complex is required for detection of, and stable binding to the correct start codon. As already explained above, the eIF4A helicase disrupts secondary structure in the path of the scanning IC, and its activity is stimulated by eIF4B and eIF4H; and eIF4G both stimulates eIF4A activity and tethers it to the IC.

Therefore, although scanning is unique to eukaryotes, it relies to some extent on IFs and activities shared among all kingdoms. The mutual stabilization of binding of the individual factors ensures that there is little premature dissociation of factors from the IC and of the IC from mRNA (reviewed in Hershey & Merrick, 2000; Jackson, 2000; Preiss & Hentze, 2003).

It is not clear whether or not the IC remains tethered to the 5′-cap during scanning. The two alternatives would have quite different implications on the overall mechanism of translation initiation in eukaryotes. If the IC remains attached to the 5′-end, a new 40S subunit cannot be recruited until after the start codon has been reached. Alternatively, if the interaction with the cap is lost once the 43S complex is recruited to the mRNA and starts scanning, a new 43S complex could bind immediately. If the scanning complexes do not remain tethered to the 5′-cap, a new question arises – which interaction(s) is being disrupted: eIF4E-cap, eIF4E-4G, or eIF4G-eIF3 (in human) or eIF4G-eIF5/eIF1 (in yeast)? The available evidence is in favor of the model where the scanning complex remains anchored to the cap, at least on most mRNAs, most of the time. But the situation may be more complex: the ICs could sometimes, but not always, dissociate from the 5′-cap, with different rate, depending on the mRNA; and the process need not be identical between yeast and human (see below).

Translation initiation on some mRNAs with long structured 5′-UTRs could take minutes, which means that, in order to be maintained throughout 43S complex recruitment and scanning on such mRNAs, an interaction would have to have a lifetime in that range and not be significantly

weakened during any stage of the process. The secondary structures in the 5′-UTR of mRNA pose another question: it is known that they significantly lower the rate of translation, but do they slow down the scanning complex or cause premature dissociation of a significant fraction of the ICs, or both?

The recent reports on the mechanism of reinitiation (see above) indicate that eIF4G remains associated with the scanning complexes, at least in mammals (Poyry *et al.* 2004). Although no similar direct evidence for the yeast system was presented (Rajkowitsch *et al.* 2004), yeast eIF4G is likely also associated with the scanning complex and required for reinitiation. The fate of eIF4E, remains unclear, but the interaction of eIF4E with eIF4G involves co-folding of large segments of eIF4G and eIF4E, stable for minutes (Gross *et al.* 2003) and is likely retained during scanning. eIF4E binding to the 5′-cap is stabilized by eIF4G binding, and is further stabilized when PABP binds to eIF4G (even though there is no direct interaction between PABP and eIF4E). Thus, the eIF4E-cap and the eIF4E-4G interactions are likely to be sufficiently stable, at least on short 5′-UTRs.

As stated above, eIF4G mediates recruitment of 43S ICs to the 5′-cap in both yeast and human, but while the interaction is mediated by eIF3 in human, in yeast it was reported to involve eIF5 and eIF1 instead, and it is not clear how stable either interaction is. The interaction of eIF4G with eIF3 (or eIF5/eIF1 in yeast) could be relatively weak, but sufficient for 43S complex recruitment in the context of multiple protein-mRNA, ribosome-mRNA interactions at the 5′-cap, and scanning over short distances. Such a weak interaction may not (always) be retained during scanning through long structured 5′-UTRs. Loss of eIF4G would also lead to loss of eIF4A and prevent the IC from reaching the start codon. Alternatively, the interaction could be strong, making it more likely to survive during scanning. As eIF4G and eIF3 in human (or eIF5/eIF1 in yeast) have not been reported to exist as a stable complex off mRNA, such putative strong interaction would have to depend on structural rearrangements upon eIF4G binding to eIF4E, 43S complex recruitment to mRNA and/or stabilization by mRNA binding. It has been noted, that short upstream ORFs (uORFs) are more frequent in mammals than in yeast, which suggests that reinitiation may be more efficient in mammals (Poyry *et al.* 2004), indirectly pointing toward a possibly stronger association of mammalian eIF4G with eIF3, than of yeast eIF4G with eIFs 1 and 5, at least after start site recognition.

It should be noted that the presence of an interaction during scanning does not in itself prove that the interaction is important for scanning. For example, it was recently found that N-terminal deletions of yeast eIF4E destabilize binding to eIF4G and cause a temperature-sensitive phenotype and reduced growth rates, but only moderate reduction of polysome content *in vivo*. The K_D of these eIF4E mutants for eIF4G is weakened by ~100-fold, to ~200 nM, with complex life times in the order of seconds, instead of minutes. Furthermore, the cooperativity of eIF4E binding to the cap and to eIF4G was affected in the eIF4E deletion mutants (Gross *et al.* 2003), and the same could be true for the cooperativity of eIF4G binding to eIF4E and PABP. This suggests that the maintenance of a stable interaction between eIF4G and eIF4E during scanning is not essential on the majority of mRNAs. Instead, the *in vivo* phenotype could be due mainly to the eIF4E-binding protein (4EBP), which competes with eIF4G for eIF4E binding and inhibits translation (Pause *et al.* 1994a; Altmann *et al.* 1997; Ptushkina *et al.* 1998). 4E-BP can bind equally well to the mutant and WT eIF4E, and inhibit initiation predominantly on 'weak' mRNAs.

Another difference between human and yeast is that human eIF4A binds to eIF4G with higher affinity, whereas yeast eIF4A is not stably associated with the eIF4G-eIF4E complex. Human

eIF4G has two additional domains at its C-terminus, absent in yeast. One of the domains provides a new, second binding site for eIF4A and the other binds the MNK kinases, which phosphorylate eIF4E. The role of eIF4E phosphorylation by MNK is somewhat controversial: it has been reported to stimulate translation by some groups, but others have found that it inhibits translation (reviewed in Scheper & Proud, 2002). Interestingly, phosphorylated eIF4E binds more weakly to the 5′-cap, but any possible effects of eIF4E phosphorylation on the cooperativities in the cap-eIF4E-eIF4G-PABP interaction network have not been studied, to our knowledge. One attractive interpretation is that MNK can access eIF4E and phosphorylate it only after recruitment of the 43S IC. The lower affinity of phosphorylated eIF4E for the 5′-cap would then allow release of the complex from the cap, and binding of a new eIF4E:eIF4G:PABP complex, which could recruit a new 43S complex while the first is still scanning along mRNA. Such 'recycling' role of eIF4E phosphorylation requires that the lifetime of the eIF4E-cap complex be of the same order as scanning, and would have to be mRNA-dependent. Alternatively, eIF4E phosphorylation could have a more indirect regulatory role, such as changing the affinity of the eIF4F complex for a (yet unknown) factor (reviewed in Dever, 2002; Scheper & Proud, 2002; Preiss & Hentze, 2003).

Recruitment of the ICs to the 5′-cap is one of the two major targets of regulation of translation initiation. eIFs 4A, 4E, 4G, and PABP have all been reported to interact with translational regulators. For example, the eIF4E-binding proteins (4EBPs) compete with eIF4G for eIF4E binding and inhibit translation. Interactions and modifications could affect the stability of the IC. Alternatively, their main role could be to promote or prevent binding of specific translational regulators, rather than directly modulate the rates of translation initiation. An interesting example in this respect is a report about the inhibition of translation of non-polyadenylated mRNAs by Ski2p and Slh1p in yeast (Searfoss & Wickner, 2000; see next section).

4.6.2 Polyadenylation

Another specific feature of eukaryotic translation is the regulation of the rate of initiation by polyadenylation of the mRNA 3′-end. Poly-A is added post-transcriptionally, by a template-independent poly-A polymerase, which recognizes a specific polyadenylation signal near the 3′-end. Polyadenylation has a number of different roles. (1) It is important for export of the mRNA from the nucleus. (2) It stimulates translation initiation. (3) It increases the stability of the mRNA to nucleolytic degradation. (4) It is the binding site for PABP, which in turn interacts with translational regulators, such as the PABP-interacting proteins 1 (PAIP1) and 2 (PAIP2), which stimulate (PAIP1) or inhibit (PAIP2) translation (Craig *et al.* 1998; Khaleghpour *et al.* 2001). As mentioned above, an interaction between PABP and eIF4G can bring the 5′- and 3′-ends of mRNA together and stabilizes binding of eIF4E to the 5′-cap. The exact mechanisms of stimulation of translation initiation by PABP remain controversial and the degree of stimulation varies greatly depending on the system and experimental conditions. Early work had suggested that PABP stimulates subunit joining, but the discovery of a direct PABP-eIF4G interaction provided strong support for the view that PABP promotes 43S IC recruitment. Recent reports showed that when bound to poly-A oligonucleotides, PABP can stimulate translation *in trans*. These results indicate that circularization itself is not required for stimulation of translation *in vitro* and that probably the role of the poly-A tail is to allow PABP association with the IC. A number of proteins, both activators and inhibitors of translation, have been reported to bind to PABP. The length of the poly-A tail is sufficient for binding of multiple

PABP molecules and, therefore, various factors could be simultaneously associated with the 3′-end of an mRNA. In higher eukaryotes, one of the proteins that bind to PABP is eRF3, suggesting a possible role for PABP in channeling of terminating ribosomes back to the 5′-end of mRNA. It should be noted, that such channeling may not be observable *in vitro* (reviewed in Sachs, 2000; Preiss & Hentze, 2003).

An additional issue with circularization is how and when it occurs. Considering that the length of the mRNAs can be thousands of base-pairs and the total cellular mRNA concentrations can be near 1 μM (von der Haar & McCarthy, 2002), the effective concentration of 3′-ends from other mRNA molecules in the vicinity of the 5′-end of an mRNA molecule could easily be higher than that of its own 3′-end and would depend on local mRNA concentrations. Although the secondary and tertiary structures in mRNA could bring the two ends closer together, they could also keep them apart. Therefore, it is hard to imagine how circularization can occur spontaneously in the cytoplasm. It is known that mRNAs are processed co-transcriptionally and exported from the nucleus as ribonucleoprotein particles (mRNPs) (reviewed in Reed, 2003; Erkmann & Kutay, 2004; Vinciguerra & Stutz, 2004). Therefore, circularization most likely takes place in the nucleus and nuclear factors, such as the nuclear cap-binding complex (CBC) probably participate in the initial circularization. For example, CBC stimulates polyadenylation (Flaherty *et al.* 1997). A more practical aspect of the problem with circularization is what happens *in vitro* upon addition of naked mRNA. If only one type of mRNA is added, it may not be functionally important whether mRNAs are circularized or connected with each other in a string. However, that is not necessarily the case if a large amount of competitor mRNAs is added simultaneously.

It has been reported that in yeast, PABP binding to poly-A prevents the inhibition of translation mediated by two nonessential proteins, Ski2p and Slh1p. The most striking findings in that work were that: (1) In a yeast strain lacking both proteins, non-polyadenylated mRNA is translated as well as polyadenylated mRNA, indicating that at least in that case polyadenylation is necessary to prevent inhibition, rather than to directly stimulate translation initiation. (2) PABP binding to poly-A prevents the Ski2p- and Slh1p-mediated inhibition (Searfoss & Wickner, 2000; Searfoss *et al.* 2001). It was proposed, that the inhibition of translation initiation by Ski2p and Slh1p is not at the stage of 43S complex recruitment, but instead, at the stage of eIF5/eIF5B-mediated subunit joining (Searfoss *et al.* 2001). However, as also found by others, the authors' own results show synergistic effect between polyadenylation and capping – a strong indication that the poly-A tail and the cap affect the same step. Furthermore, evidence from eIF5 mutant and eIF5B deletion strains was interpreted based on the assumption that eIF5 and eIF5B both act at the stage of subunit joining (Searfoss *et al.* 2001), which we now know is not the case for eIF5. Ski2p and Slh1p are part of a multiprotein complex, which also contains nucleases. The main pathway for mRNA degradation in eukaryotes starts with an endonuclease cleavage, producing an accessible non-capped 5′-end, followed by exonucleolytic degradation in a 5′ to 3′ direction, and PABP can similarly be involved in protection of mRNA from endonucleolytic cleavage near the 5′-end.

In summary, the poly-A tail plays important roles in stabilization of eIF4F binding to the cap, recruitment of the 43S complex to mRNA, protection from inhibitors of translation, and protection from nucleolytic degradation. Although circularization of mRNA *per se* does not appear to be required for stimulation of translation *in vitro*, it is necessary in order to transfer signals from the 3′-end to the cap. And, at least in higher eukaryotes, it may be important *in vivo* to channel newly dissociated 40S subunits back to the cap (reviewed in Sachs, 2000; Preiss & Hentze, 2003).

4.7 Leaderless mRNAs – a minimal 'universal' system

mRNAs, where the start codon is at or near the 5′-end are called leaderless. Leaderless mRNAs are found in various organisms and, as explained above, such mRNAs can be translated in cell extracts derived from all kingdoms. 5′-terminal AUG codons, however, are 'non-canonical' and both bacteria and eukaryotes have IFs that can discriminate against them – IF3 and eIF1 respectively, although translation of leaderless mRNAs does not appear to be completely repressed. Thus, leaderless mRNAs can be viewed as a highly simplified, but still biologically relevant model system for translation initiation.

In bacteria, translation of leaderless mRNAs is regulated by the ratio between IF2 and IF3. IF3 association with the ICs appears to be dynamic, with multiple cycles of binding and release. At lower IF3 concentrations, there is a greater opportunity for formation of an IC, and upon IF2-promoted subunit joining, the IC becomes resistant to dissociation by IF3. It has been noted, that some bacterial species have much greater number of leaderless mRNAs than others, which could be related to differences in the activity and steady-state concentrations of IF3 and other IFs (reviewed in Moll *et al.* 2002).

It is not clear what role leaderless mRNAs play in eukaryotes. ICs can form directly at a 5′-terminal AUG codon in the absence of eIF1. When the 5′-end of the mRNA is not capped and is free of secondary structure, the Cap-binding complex eIF4F is not necessary or inhibitory (Pestova & Kolupaeva, 2002). eIF1 forms a complex with the multisubunit eIF3 (and, at least in yeast, with eIF4G and eIF5), which stabilizes its binding to the small ribosomal subunit. But eIF1 is not an integral part of eIF3 and, therefore, eIF1 and eIF3 could bind individually. It is not clear how efficiently eIF1 can inhibit initiation at a 5′-terminal codon in the absence of eIF3, but eIF3 could increase the inhibitory effect of eIF1. Therefore, the concentrations of eIF1, eIF3 and their ratios to the ribosomes are likely to modulate the translation efficiency of leaderless mRNAs.

Archaea appear to have a great number of leaderless mRNAs. Consistent with the above discussion, although archaea have an eIF1 homolog, they do not have eIF3. We are not aware of evidence whether or not archaeal eIF1 discriminates against 5′-terminal AUG codons as its eukaryotic counterpart. However, even if archaeal eIF1 is as active in selecting against leaderless mRNAs as eukaryotic eIF1, in the absence of eIF3 it will likely be cycling on and off the small ribosomal subunit, similarly to bacterial IF3. It will be especially interesting to know more about the properties of archaeal eIF1: whether it discriminates against 5′-terminal AUG codons, and if yes – whether it does that more or less efficiently than discrimination against non-AUG codons or against 5′-proximal non-AUG codons. In bacteria, the small subunit is able to slide up to 40 nt in the absence of secondary structure (Adhin & van Duin, 1990). Therefore, even start codons that are located a short distance from the 5′-end may be accessible for initiation.

From an evolutionary standpoint, leaderless mRNAs could be viewed as the precursors of eukaryotic scanning. The ribosome, especially in the presence of eIF1 and eIF1A, may be able to slide ('scan') over short distances if the mRNA near the 5′-end is unstructured. The missing link between archaea and eukaryotes is provided by some primitive eukaryotes. It was recently reported that the protozoan Giardia lamblia has leaderless mRNAs with 0–14 nt before the start codon, has no eIF3 or eIF4G homolog and does not use scanning. However, the mRNAs are capped and polyadenylated (Li & Wang, 2004).

It should be noted, that intact 70S ribosomes can bind to, and initiate translation from a leaderless mRNA, at least *in vitro*, although initiation and binding are inhibited by IFs, and the

in vivo significance of these findings is not clear (O'Donnell & Janssen, 2002; Udagawa *et al.* 2004).

4.8 Reinitiation and leaky scanning

Reinitiation is a process, where the ribosome initiates translation of a subsequent ORF, without dissociating from the mRNA. Studies of reinitiation and initiation at non-AUG start codons (discussed in Section 4.9) have provided valuable information about the mechanism of translation initiation. Furthermore, the abundant genetic information about these processes, accumulated over the years, provides new insights in the context of new structural data about the factors involved. Here we will provide descriptions of reinitiation, initiation at non-AUG codons, and of the main genetic systems, whereas individual mutants will be discussed in Section 5, in the context of the structure/function of the IFs.

In a number of bacterial mRNAs, the stop and start codons of two consecutive ORFs are overlapping, producing the tetranucleotide AUGA. It was found that the ribosome reinitiates efficiently on such mRNAs, whether or not the second ORF has an SD element, and the process is called translational coupling. In polycistronic bacterial mRNAs, translation from the second ORF is typically ~70% of that from the first ORF; translation from the third ORF is ~70% of that from the second ORF, etc. Translationally coupled ORFs, on the other hand, usually have approximately equal rates of translation initiation, which could allow coordinated expression of the encoded proteins in stoichiometric amounts. Ribosomes can also reinitiate translation if the new start codon is a short distance from the site of termination, but less efficiently, and the detailed mechanism and requirements may differ from those on mRNAs with overlapping stop and start codons (reviewed in Kozak, 1999; Jackson, 2000). It is not clear if concentrations of initiator tRNA, IFs or ratios between them regulate reinitiation, or if the process is as efficient in stationary phase or during starvation as it is in exponentially growing cells.

Reinitiation occurs on some eukaryotic mRNAs and has important regulatory functions. Like the bacterial ribosome, the eukaryotic ribosome can also reinitiate on overlapping stop-start codons (Peabody & Berg, 1986), and this reinitiation mechanism has been described for viral mRNAs (Horvath *et al.* 1990; Meyers, 2003). A second reinitiation mechanism exists in the eukaryotic system, where there is no stop-start codon overlap, but instead, the small subunit reinitiates scanning in the 3′-direction until it reaches a new start codon. As explained above, scanning depends on a number of eIFs. With the exception of eIF4E and PABP, responsible for cap binding and poly-A binding respectively, all factors normally required for scanning and initiation are required for reinitiation. Furthermore, most of the factors must have been present during the first initiation event (and remained bound through elongation and termination) and cannot re-bind (Poyry *et al.* 2004). And as it appears that at least some of the factors are only loosely associated with the ribosome after start site recognition, reinitiation is efficient only after short ORFs, where the eIFs are still present at termination (Poyry *et al.* 2004; Rajkowitsch *et al.* 2004). The eIF2·GTP·Met-tRNA$_i$ ternary complex needs to be regenerated and does not appear necessary for scanning itself, but is required for start site recognition. An IC, that has not yet 'picked' eIF2·GTP·Met-tRNA$_i$, will scan through an AUG codon, even if the codon is in a proper context. Therefore, the efficiency of reinitiation at a given start codon depends both on the distance from the stop codon, where the small subunit resumed scanning, and the concentration of eIF2·GTP·Met-tRNA$_i$ in the cell. The fates of eIF1A and eIF5B are less clear. Unlike other factors, eIF1A and eIF5B appear to dissociate from the A-site upon subunit joining.

However, at least eIF1A has a role in scanning and would have to re-bind quickly. It is possible that eIF1A and/or eIF5B also remain weakly associated with the ribosomes at secondary sites, e.g. via their N-terminal positively charged regions. And, in higher eukaryotes, eIF5B can associate with eIF5 (see below).

The role of reinitiation as a sensor for the levels of eIF2·GTP·Met-tRNA$_i$ *in vivo* has been studied for many years in yeast, and similar systems have been described in higher eukaryotes as well. The regulatory regions of the mRNAs often contain a set of consecutive and alternative ORFs, some of which have suboptimal sequence context or a non-AUG start codon, adding to the complexity of the systems and the opportunities for fine-tuning and regulation they present. Depending on the organization of the mRNA 5′-region, translation of the main ORF can be up-, or down-regulated to various degrees by only moderate decrease in the levels of eIF2·GTP·Met-tRNA$_i$, which does not affect general translation (reviewed in McCarthy, 1998; Hinnebusch, 2000; Dever, 2002).

Historically, translation reinitiation systems, and the *GCN4* mRNA in particular, have been a valuable tool for genetic studies of the roles of individual eIFs in initiation. GCN4 is a transcriptional activator responsible for up-regulation of the expression of key enzymes involved in amino-acid synthesis. The *GCN4* mRNA contains four short, 3- or 4-codon upstream ORFs (uORFs). Under non-starvation conditions, the majority of ribosomes initiate at uORF 1, ~50% of them (the 40S subunits) resume scanning after termination, reinitiate at uORF 4 and fail to reinitiate again at the *GCN4* start codon. Under conditions where the levels of eIF2·GTP·Met-tRNA$_i$ are lowered, the majority of ribosomes fail to reinitiate at uORF 4, because by the time they reach the start codon of uORF 4, they have not yet 'picked' a new eIF2·GTP·Met-tRNA$_i$. Instead, the ribosomes reinitiate at the *GCN4* start codon. Therefore, the expression of GCN4 is increased when the levels of eIF2·GTP·Met-tRNA$_i$ are decreased.

Mutations that inactivate or decrease the activity of proteins required for maintenance of high eIF2·GTP·Met-tRNA$_i$ levels lead to constitutive derepression of GCN4 – a general control derepressed (Gcd$^-$) phenotype. Mutations that interfere with the regulatory network prevent GCN4 derepression – a general control non-derepressible (Gcn$^-$) phenotype. Gcd$^-$ phenotypes have been associated with mutations in subunits of eIF2 and eIF2B (reviewed in Hinnebusch, 2000), but also mutations in eIF1A (Olsen *et al.* 2003), eIF3 (the Prt1-1 mutant in yeast eIF3b; Valasek *et al.* 2004) and eIF1 (Singh *et al.* 2004a). Gcn$^-$ mutations have been found in eIF2α and the α, β, and δ subunits of eIF2B (forming the regulatory subcomplex of eIF2B). Of course, inactivation of the target protein GCN4 also leads to a Gcn$^-$ phenotype (reviewed in Hinnebusch, 2000). Gcn$^-$ phenotypes have also been reported for mutations in eIF5B (Lee *et al.* 2002; Shin *et al.* 2002) and eIF3 (Nielsen *et al.* 2004).

Gcd$^-$ phenotype can be caused by various factors, such as decreased concentrations of charged Met-tRNA$_i$ or of eIF2·GTP. The Met-tRNA$_i$ levels depend on the concentration of methionine and the activity of the Met-tRNA$_i$ synthetase, as well as on the tRNA$_i$ levels. In turn, eIF2·GTP levels depend on the GTP/GDP ratio, on the activity of the eIF2 GEF, eIF2B as well as on the levels of eIF2. eIF2B is one of the main targets for regulation of translation initiation. Its activity is regulated through phosphorylation of its substrate eIF2, as well as directly – through binding to, or phosphorylation of eIF2B itself. Phosphorylation of the conserved Ser51 in the eIF2α subunit converts eIF2·GDP into eIF2(α-P)·GDP – a competitive inhibitor of eIF2B. eIF2α is phosphorylated by several kinases (GCN2 in yeast and GCN2, PERK, PKR and HRI in mammals) under various types of stress or amino-acid starvation. The eIF2B activity is not always limiting and the sensitivity of eIF2B to eIF2α

phosphorylation (and thus the potential of eIF2α phosphorylation to regulate eIF2B activity) depends on the ratio between eIF2 and eIF2B: while eIF2 is always in excess, this excess varies significantly and can be only 2-fold in some cells (reviewed in McCarthy, 1998; Hinnebusch, 2000; Dever, 2002). The exchange of eIF2-bound nucleotide by eIF2B will be discussed in detail in Section 5.4.2.

A Gcd$^-$ phenotype is also produced by mutations that destabilize the binding of eIF2·GTP·Met-tRNA$_i$ to the IC, that destabilize binding of Met-tRNA$_i$ or GTP to eIF2, or mutations that increase the rate of spontaneous GTP hydrolysis by eIF2. This last class of mutations can cause premature release of eIF2 from the IC and may also affect start site selection (see next section).

As mentioned above, initiation from at least some start codons is not 100% efficient and a fraction of the ICs continue scanning. The phenomenon is known as leaky scanning. The frequency of leaky scanning past a given start codon depends on the nucleotide context, as well as on presence or absence of secondary structures, which could slow down scanning. ICs that scan through the first start codon could initiate at a downstream start codon, if one exists and if scanning is not blocked by secondary structures, or could be 'stuck' on mRNA and eventually dissociate. If dissociation from the cap is faster than dissociation from mRNA, it may be possible to start a new round of initiation and have more than one IC on the same mRNA simultaneously. Although queuing of 40S subunits has been observed by footprinting under somewhat artificial conditions, it is very difficult to 'visualize' the fairly unstable scanning 40S subunits (Jackson, 2000).

4.9 Initiation at non-AUG codons and the stringency of start codon selection

As discussed above, in bacteria, although AUG is the predominantly used start codon, GUG and UUG are also considered 'canonical'. On rare occasions, other, 'non-canonical' start codons are also found in bacteria. In eukaryotes, AUG is used almost exclusively, but UUG and GUG, although considered 'non-canonical' in eukaryotes, are used more efficiently than the other 'non-canonical' codons. IFs in both bacteria and eukaryotes act only indirectly in start codon selection, but do not directly 'inspect' the start codon. Thus, a Met-tRNA$_i$ (or fMet-tRNA^{fMet} in bacteria) with a mutation in the anticodon allows initiation at the complementary codon, but not at AUG (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000). Therefore, the preference for UUG and GUG in both kingdoms must be due to the architecture of the codon–anticodon interaction in the context of the small subunit. It should be noted, that, whereas bacterial IF3 does not discriminate against UUG and GUG, its functional eukaryotic analog, eIF1 does discriminate against UUG and GUG, and even against AUG in 'bad' nucleotide context.

When present at concentrations equal to or higher than the concentration of the small ribosomal subunit, IF3 discriminates against initiation from non-canonical codons. At sub-stoichiometric concentrations, however, IF3 stimulates translation initiation indiscriminately, indicating that IF3 goes through multiple cycles of binding and release from the small subunit. It appears that IF3 activates the small subunit by inducing conformational changes, some of which are retained after its release. At the same time, IF3 needs to be present on the ribosome in order to prevent initiation at a suboptimal start site. These peculiar properties of IF3 are utilized in *E. coli* for translational autoregulation of IF3 levels: the ORF for IF3 starts with an AUU codon and its translation is inversely proportional to the IF3 concentration in the cell.

Similarly, translation of other mRNAs with non-canonical start codons is subject to regulation by the relative levels of IF3, the other IFs and fMet-tRNA^{fMet} (reviewed in Hershey & Merrick, 2000; Petrelli *et al.* 2001).

Although the stringency of start codon selection is higher in eukaryotes, initiation can still occur with low efficiency at a non-AUG codon. Screening for suppressors of initiation codon mutation (Sui) in yeast has yielded mutations in a number of eIFs that lower the stringency of discrimination against non-AUG codons (reviewed in Donahue, 2000).

Here, we will try to present a general overview of the Sui phenotypes, while individual mutations will be discussed in more detail in Section 5, in the context of the structure and function of the respective proteins. Analysis of Sui mutants has provided valuable information about the mechanisms of start site selection and the factors involved. Typically, the studies have been done using the *HIS4* gene with a start codon mutation. Yeast carrying such a mutation do not grow on minimal medium without histidine. The *HIS4* mRNA has an in-frame UUG codon in an appropriate context and, in the absence of the AUG codon, only several-fold increase of initiation from this alternative start site is sufficient for viability. Transcription of *HIS4* mRNA is upregulated by the GCN4 transcriptional activator and, therefore, increases in the presence of Gcd$^-$ mutations (see the previous section). This phenomenon has dual effects. On one hand, it requires that a singenic strain with a WT *HIS4* gene be used as a control to account for potential changes in mRNA levels. On the other hand, it allows detection of a weaker Sui phenotype, if accompanied by a Gcd$^-$ phenotype, because the increase in translation of the reporter protein is multiplied by the increase in the levels of its mRNA.

Sui mutations have been found in eIF1, all three subunits of eIF2, and in eIF5. As explained above, start site recognition involves structural changes in the IC, GTP hydrolysis by eIF2, and release of eIF2·GDP. eIF1 opposes the above structural changes and the efficiency of initiation at any given codon relies on the dynamic balance between the 'forward' and 'reverse' processes. Sui mutations can shift the balance toward initiation in several ways:

(1) As eIF1 does not directly inspect the start codon, the Sui phenotype of the eIF1 mutants arises from impaired ability of eIF1 to dissociate ICs formed on suboptimal start sites.
(2) Any mutation that increases the rate of spontaneous or IC-dependent GTP hydrolysis by eIF2 is likely to also increase the frequency, with which eIF2 'escapes' the eIF1-mediated control of start site selection. Such mutations have been found in both eIF2 and its GAP, eIF5.
(3) Sui mutations in eIF2, that decrease Met-tRNA$_i$ binding, have also been found. The interpretation of their phenotype is not straightforward. Such mutations could cause spontaneous dissociation of eIF2·GTP from the ICs, without GTP hydrolysis, but it is not clear if eIF2 release alone would be sufficient to promote initiation. If the reorientation of Met-tRNA$_i$ during start codon recognition puts strain on its interaction with eIF2, then lower affinity of eIF2 for Met-tRNA$_i$ could lower the energy barrier for Met-tRNA$_i$ reorientation. The result could be an increased rate of IC-dependent eIF2 GTP hydrolysis, eIF2·GTP dissociation or both (Huang *et al.* 1997; Donahue, 2000). Clearly, the above scenarios are not mutually exclusive.

Recently, a weak Sui phenotype was reported for a mutation in yeast eIF4G, consistent with the observed interactions of yeast eIF4G with eIF1 and eIF5 (He *et al.* 2003). Mutations in eIF3c also showed a Sui phenotype, whereas other mutations in the same region of the

protein suppressed the effect of Sui mutants of eIF5 and eIF2β – a 'suppressor of Sui' (Ssu) phenotype (Nielsen *et al.* 2004).

Sui mutants (as well as the Gcd$^-$ and Gcn$^-$ mutants discussed in the previous section) can be either dominant or recessive, that is the phenotype may be retained or lost if the mutants are co-expressed with the wild type (WT) proteins. Dominant mutants are designated with all-capital letters (e.g. SUI) and recessive mutants – with small letters (e.g. Sui), but the above convention is not always followed. Dominance and recessivity, as well as other genetic interactions, can be useful for understanding the effect of the mutations in eIFs on a mechanistic level, especially now that the structures of most factors are known.

Mutations that destabilize or eliminate the binding of a factor to the IC, or to a subcomplex, are likely to be recessive, because the WT protein will replace the mutant, yielding an active complex. Such mutations could affect binding directly and/or affect the concentration or structural integrity of the factor. Mutations that do not affect the association of the factor with a complex, but affect its function as part of the complex can be dominant, but the interpretation in this latter case is not always straightforward. For example, mutations in eIF2β, which destabilize its binding to eIF2γ and cause increased rate of spontaneous GTP hydrolysis, have a recessive Sui phenotype, whereas mutations in eIF2β, which cause increased rate of spontaneous GTP hydrolysis without affecting binding to eIF2γ, have a dominant SUI phenotype. Similarly, an S51A mutation in eIF2α has a Gcn$^-$ phenotype, because it prevents phosphorylation of eIF2α and tight association of eIF2 with the regulatory α,β,δ-subcomplex of eIF2B upon amino-acid starvation, but the phenotype is recessive, because if WT eIF2α is co-expressed, it can be phosphorylated and the resulting eIF2(α-P)·GDP will be able to bind to eIF2B and inhibit its activity. An S51D mutation, mimicking eIF2α phosphorylation causes stable binding of eIF2·GDP to the regulatory α,β,δ-subcomplex of eIF2B, whether or not WT eIF2α is present, and, therefore, has a dominant GCD$^-$ phenotype. Some, but not all dominant phenotypes can be suppressed if the WT protein is overexpressed to levels much higher than those of the mutant. In the above example, overexpression of WT eIF2β will displace most of the dominant mutant eIF2β from the trimeric eIF2 complexes by mass action (reviewed in Hinnebusch, 2000; Dever, 2002).

Sui, Gcd$^-$, and Gcn$^-$ mutations can also affect directly or indirectly cell growth and viability: they can be lethal, conditionally lethal, or cause a slow growth phenotype. Such additional phenotypes, as well as genetic interactions with other factors and mutations also provide insights into the functions of the affected eIFs and the effects of the individual mutations. The Gcd$^-$ and Gcn$^-$ phenotypes are not sufficient to affect growth in rich medium by themselves. Therefore, if any growth defects are observed, they will be due to additional functions being affected by the mutation. The Gcd$^-$ (Gcn$^-$) phenotype, in turn, can often be caused by an impaired essential function. The Sui phenotype, however, is caused by impaired start codon selection and severe Sui phenotypes cannot be tolerated and are lethal. The above considerations point toward certain limitations of *in vivo* screens, especially for Sui mutations: neither a recessive mutation, that is also lethal, nor a dominant mutation, that is also dominant lethal under the screening conditions, can be obtained. Dominant mutants that are also recessive lethal however, have been obtained when screens were performed in the presence of the WT proteins (reviewed in Hinnebusch, 2000; Donahue, 2000; Dever, 2002). In more general terms, Sui mutations obtained from *in vivo* screens are unlikely to affect the residues most important for start codon selection, as such mutations would also probably be lethal.

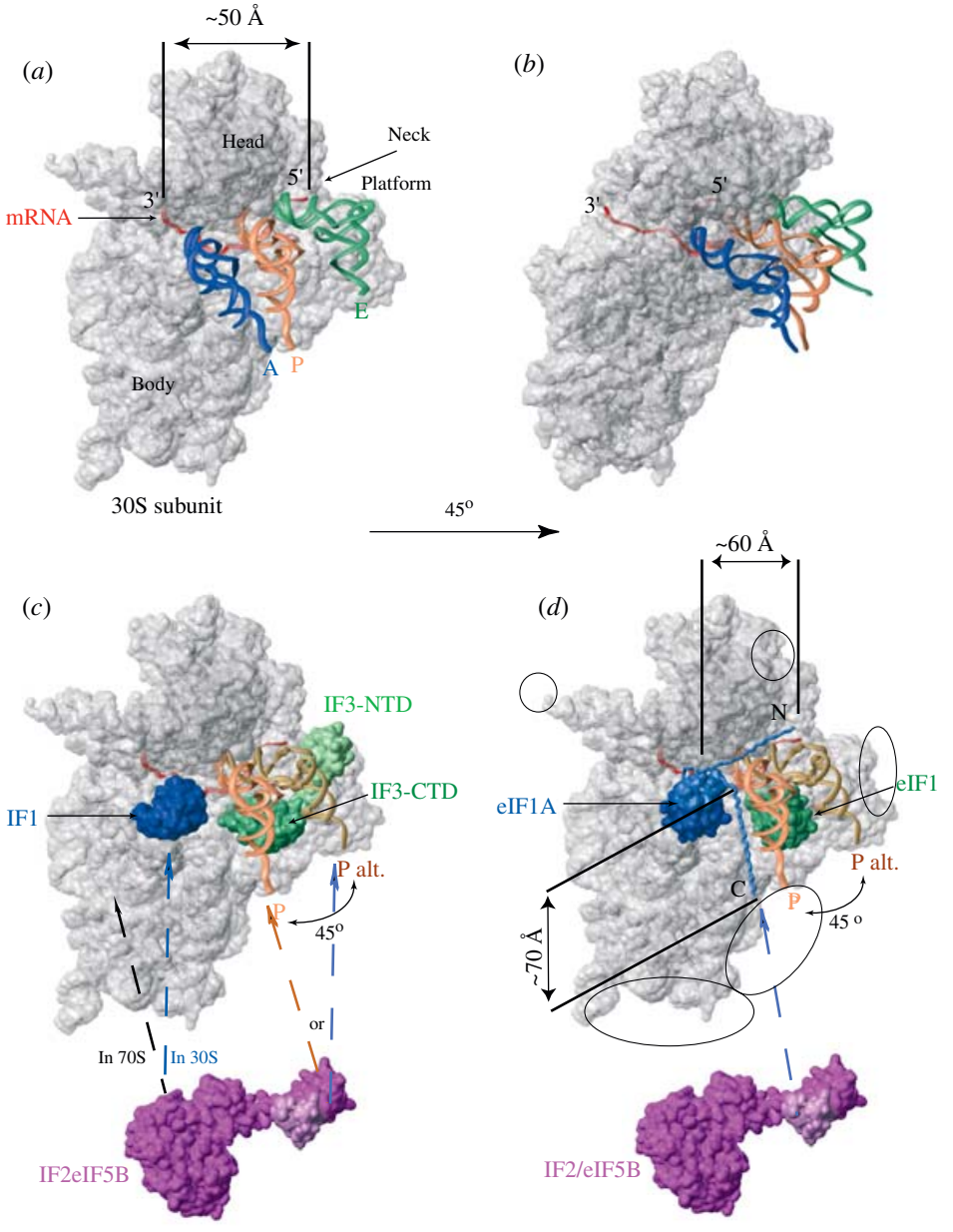

Fig. 4. Locations of tRNA and (e)IFs on the small ribosomal subunit. (*a*) Bacterial 30S ribosomal subunit (grey) with mRNA (brown) and A-, P-, and E-site tRNAs (in dark blue, coral and green respectively), from the structure of the 70S ribosome, PDB code 1JGP (Yusupova *et al.* 2001). The 30S subunit is semi-transparent, in order to display the path of the mRNA. The head, neck, platform and body are labeled. The mRNA 3′-end (in the 'entry' channel) and 5′-end (in the 'exit' channel), as well as the distance between the entry and exit sites of mRNA are shown. (*b*) Same as (*a*), but rotated 45°. (*c*) Location of IF1 (dark blue) – from the structure of the 30S-IF1 complex, PDB code 1HR0 (Carter *et al.* 2001), and IF3 (Dallas & Noller, 2001). A possible alternative orientation for the P-site tRNA early in initiation is shown in dark brown and labeled 'P alt' (see text). IF3 was placed on the 30S subunit manually, and therefore, the present orientation may not be identical to that in Dallas & Noller (2001). The interactions and possible location of IF2 are shown with arrows. The structure of the archaeal IF2/eIF5B homolog (magenta) is from

5. Structure/function of initiation factors

In the sections below, we will try to summarize the available information about the structure and function of individual IFs, their interactions with each other and with the ribosome. Where appropriate, interactions between factors will be discussed in more detail in the section describing the second interacting partner. Some of the highlights from this section are listed below: (1) Structural information is now available for most of the IFs. Some notable exceptions remain, like eIF3 and large parts of eIF2B. (2) The positions of some of the factors (IF1/eIF1A, IF3, eIF1) in the 43S IC on the small ribosomal subunit are known; the general locations of others (IF2/eIF5B, eIF2, and to some extent eIF5) can be predicted. (3) We are close to understanding the overall organization of the heterotrimeric eIF2, and its interactions with other factors. However, the study of the detailed molecular mechanisms of action and regulation of eIF2 function in initiation may prove difficult. (4) The function of eIF2B extends beyond simple stimulation of nucleotide exchange on eIF2: eIF2B is 'optimized' for channeling and accompanies eIF2 during several steps up to the loading of the eIF2·GTP·Met-tRNA$_i$ complex on the 43S IC. (5) Unexpected structural homologies are found between factors that appear unrelated on a sequence level, like between eEF1Bα-CTD and eIF2α-CTD, or eIF4G and CBP80. Even the most sensitive algorithms for sequence homology searches are unable to detect proteins with conserved structures if their sequences have diverged. Therefore, more examples of structural homology are to be expected and could provide new insights into the functions of the respective proteins.

We will start with a brief discussion of the orientation of the initiator tRNA on the small ribosomal subunit. The only known orientation for the P site tRNA is the one in the 70S ribosome (Fig. 1*c*, Fig. 4*a*, *b*), where the ASL of the tRNA is sandwiched between the head and the body of the small subunit, while the acceptor end points away from the small subunit and into the PTC of the large subunit. Naturally, when modeling ICs, the P-site tRNA is used to infer the orientation of the initiator tRNA. While such an orientation is probably correct for a late stage IC, it is likely that the position of the initiator tRNA changes during initiation. Start site recognition by the tRNA is accompanied by major conformational changes in the IC, preparing it for subunit joining. The proper orientation of the initiator tRNA toward the P-site cavity of the incoming large subunit is expected to play a major role and is, therefore, likely to occur only after start site recognition.

The available biochemical data are difficult to interpret unambiguously, because it is not always clear whether the orientation of the initiator tRNA corresponds to a stage before or after start codon recognition, or if there is a mixture of both. The anticodon of the initiator tRNA is near the mRNA in the decoding center of the small subunit, but at the same time the tRNA would also need a degree of mobility in order to 'inspect' the mRNA for

---

Roll-Mecak *et al.* (2000), with the last two helices, absent in bacteria, painted in violet. (*d*) Locations of eIF1 and eIF1A. The bacterial 30S ribosomal subunit, mRNA, and P-site tRNA are as in (*c*). The location of eIF1 (green) is from (Lomakin *et al.* 2003), and its N-terminal tail is not shown. The location of eIF1A (in blue) was obtained by structural alignment to its bacterial homolog IF1, from the 30S-IF1 complex, PDB code 1HR0 (Carter *et al.* 2001), as previously proposed (Li & Hoffman, 2001; Marintchev *et al.* 2003). The segment of eIF1A, homologous to IF1 is in dark blue. The unstructured N- and C-terminal tails of eIF1A are shown as extended wires, in order to compare their lengths with the distances to other eIFs. Regions that are larger in the eukaryotic 40S subunit (Spahn *et al.* 2001) are circled. The eukaryote-specific interaction between the C-termini of eIF5B and eIF1A is shown with an arrow.

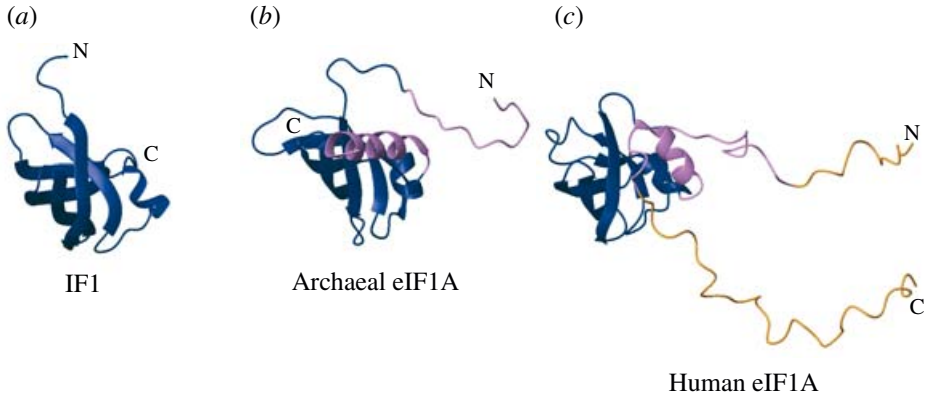

Fig. 5. Structures of IF1 and eIF1A. (*a*) Structure of *E. coli* IF1 (in dark blue), PDB code 1AH9. (*b*) Structure of archaeal eIF1A from *M. janaschii*, PDB code 1JT8. The domain homologous to bacterial IF1 is in dark blue, and segments absent from bacterial IF1 are painted in violet. (*c*) Structure of human eIF1A, PDB code 1D7Q. The coloring is as in (*b*). Segments missing from the archaeal eIF1A are in yellow. The N- and C-termini of the proteins are labeled. Note that the N-terminal tail of archaeal eIF1A (*b*) and the N-terminal and C-terminal tails of human eIF1A (*c*) are unstructured in solution.

complementarity. The tRNA is anchored to the small subunit with the help of IFs: IF2 in bacteria and the unrelated eIF2 in eukaryotes. The available biochemical data indicate that the tRNA is near the head, body and platform of the small subunit and likely interacts directly with one or more of them.

Site-directed hydroxyl radical cleavage data suggest that in the presence of IF3 in bacteria and eIF1 in eukaryotes, the initiator tRNA may be slightly rotated toward the ribosomal E-site (see below). Furthermore, there are no data in any of the kingdoms, to indicate that the initiator tRNA could be approaching from the A-site direction. It is therefore most likely, that the acceptor end of the initiator tRNA is initially oriented toward the E-site (Fig. 4*c*, *d*), and upon start site recognition, it is rotated toward the 'canonical' P-site orientation with concomitant displacement of IF3 (or eIF1 in eukaryotes). It would be difficult to speculate whether the above hypothetical orientation of the initiator tRNA would have any relation to the hybrid P-E orientation of tRNA in elongation, where the acceptor end of the deacylated tRNA has migrated toward the E site, but the anticodon end is still in the P site (see Section 3 above).

5.1 Universally conserved factors

5.1.1 IF1/eIF1A

Bacterial IF1 (Fig. 5*a*) is a ~7 kDa protein, composed of a single oligonucleotide/oligosaccharide-binding fold (OB) domain (Sette *et al.* 1997). The main roles of IF1 are prevention of tRNA binding to the A-site; stabilization of mRNA binding by the 30S subunit; and mutual stabilization of binding with IF2, fMet-tRNA^{fMet}, and IF3 to the 30S subunit. IF1 increases the rates of both subunit association and dissociation and stimulates the antiassociation activity of IF3 during early stages of initiation (reviewed in Hershey & Merrick, 2000). IF1 however, is present in the IC after IF3 has been released, allowing it to stimulate subunit joining at the late stages of initiation. IF1 interacts with IF2 on (but not off) the ribosome and this interaction may be important not only for binding of the factors to the 30S subunit,

but also for their coordinated release upon subunit joining (Benne *et al.* 1973; Boileau *et al.* 1983; Zucker & Hershey, 1986 and see below).

The crystal structure of IF1 bound to the 30S ribosomal subunit was also reported (Carter *et al.* 2001), providing the first structure of a ribosome-bound IF at atomic resolution (Fig. 4*c*). IF1, soaked in 30S subunit crystals, was found to bind in the ribosomal A-site, near the mRNA-binding channel, consistent with biochemical data. IF1 binding induced long-range conformational changes in the 30S subunit, involving tilting of the head and platform toward the A-site, which could have been even greater in the absence of crystal packing. IF1 covers the mRNA-binding channel of the 30S subunit but does not appear to contact mRNA directly. The tilting of the head and platform could tighten the grip of the small subunit on mRNA and provides a possible structural basis for an increased affinity for mRNA (Carter *et al.* 2001).

Two bases in helix 44 of the 16S rRNA: 1492 and 1493, known to be protected by IF1 binding in footprinting experiments, were found to be flipped out and buried into IF1. A third base, 530 also protected by IF1, was exposed in the complex. The authors argued that interaction of IF1 with nt. 530 may have required more extensive conformational changes in the 30S subunit, but was prevented by crystal packing (Carter *et al.* 2001). Some of the IF1-induced structural changes have been proposed to be similar to those occurring during subunit joining, thus lowering the energy of a transition state. In this context, it is interesting to note that the three bases protected and flipped out by IF1 are also protected and flipped out by tRNA binding to the A-site, and directly inspect the geometry of the codon-anticodon basepairs in the A site (Ogle *et al.* 2001). The similarities in binding between IF1 and the A-site tRNA may be due to the two molecules exploiting the same properties of the A-site for binding. An attractive alternative is also possible, where IF1 and the tRNA specifically induce similar conformational changes in the A-site, and the small subunit as a whole. Thus, IF1, which is clearly not a tRNA mimic, could be mimicking the effects of tRNA binding in the A-site and thus 'preparing' the small subunit for interaction with the large subunit (reviewed in (Ramakrishnan, 2002)).

Archaeal eIF1A (Fig. 5*b*) is homologous to IF1 over the entire length of IF1, but contains an additional 10–15 residue N-terminal tail and a helix at its C-terminus, which appears to form a rigid body with the OB domain (Li & Hoffman, 2001). In eukaryotic eIF1A, a short helix and an extended segment are added to the C-terminal helix. Eukaryotic eIF1A (Fig. 5*c*) has two long unstructured tails: a conserved positively charged N-terminal tail of $\sim$25 residues and a negatively charged C-terminal tail of $\sim$20–35 residues, whose middle segment is not well conserved, except for the negative charge. Biochemical data from eukaryotes indicate that eIF1A, like IF1, binds in the A-site and NMR data indicate a role for the new subdomain in RNA binding (Battiste *et al.* 2000). If eIF1A is superimposed with IF1 in the IF1-30S complex, the new subdomain appears solvent-exposed and points toward the head of the small subunit, which could interact with it if tilted more toward the A-site (Fig. 4*d*). The ASL of the P-site tRNA (in an orientation taken from the structure of the 70S ribosome) is not too far away and with moderate structural changes could also be contacted by eIF1A. Interestingly, a C-terminal deletion of yeast eIF1A reaching into the helical subdomain had a Gcd$^-$ phenotype, but any effect on eIF2·GTP·Met-tRNA$_i$ binding could have been indirect. The mutant protein was unstable, which makes an interpretation even harder (Olsen *et al.* 2003).

The positively charged N-terminal tail of eIF1A is likely to bind RNA, but its location in the IC has not been studied. Its length is sufficient to extend as far as $\sim$60 Å from the eIF1A binding site and could reach not only rRNA, but also tRNA or mRNA (Fig. 4*d*). If binding of eIF1A to the ribosome is similar to that of IF1, the direction in which the N-terminal tail 'exits'

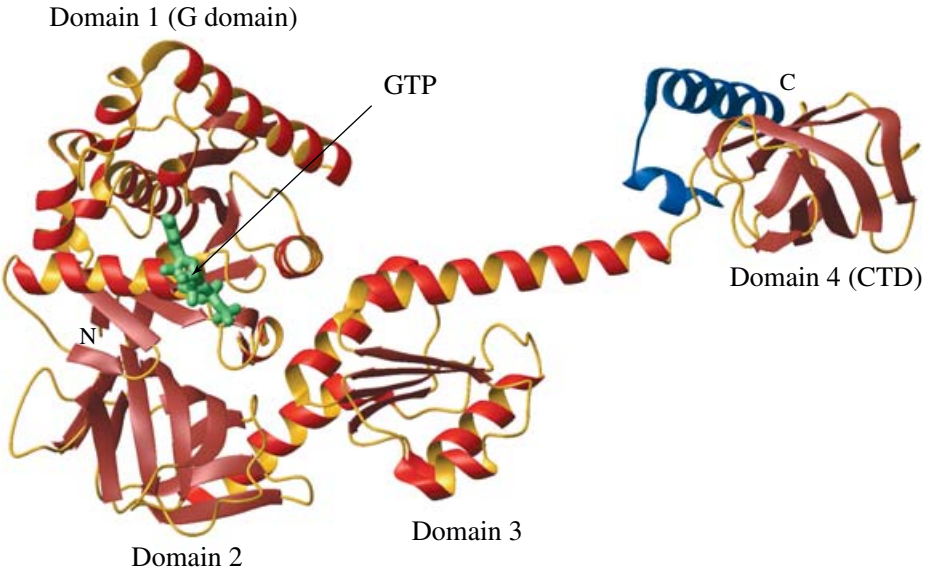

Fig. 6. Structure of eIF5B. Structure of archaeal eIF5B from *M. thermoautotrophicum* in complex with GDPNP, PDB code 1G7T. β-strands are colored in brown, α-helices and loops are colored in red and orange. The last two helices of the CTD, absent from bacterial IF2, are in blue. The domains of eIF5B and the nucleotide (shown in green) are labeled. The N- and C-termini of the protein are labeled.

the OB domain is consistent with any of these possibilities. The N-terminus of eIF1A was reported to be involved in weak interactions with eIF2 and eIF3, which could be observed in solution only if eIF1A was overexpressed. An N-terminal deletion mutant was found to be defective at low temperature, in a step following eIF2·GTP·Met-tRNA$_i$ recruitment, indicating that the interactions of eIF1A with eIF2 and eIF3 are important for a later step in initiation (Olsen *et al.* 2003). The interactions of eIF1A with eIF2 and eIF3 will be discussed in Sections 5.3 and 5.5 respectively.

The negatively charged C-terminal tail of eIF1A contains conserved patches of more hydrophobic regions. It was found to bind to the C-terminal domain (CTD) of eIF5B in both yeast and human (Marintchev *et al.* 2003; Olsen *et al.* 2003). This interaction, as well as its variations and implications, will be discussed in the next section.

5.1.2 IF2/eIF5B

The principal role of the other universally conserved IF, IF2/eIF5B, is to promote subunit joining. Upon subunit joining, the GAC of the large subunit induces IF2/eIF5B GTP hydrolysis, leading to the release of IF2/eIF5B·GDP, together with IF1/eIF1A (Benne *et al.* 1973; Choi *et al.* 2000; Olsen *et al.* 2003). In bacteria, IF2 is also responsible for selection of fMet-tRNA^{fMet} and stabilizes its binding to the 30S ribosomal subunit. No binding of eIF5B with Met-tRNA$_i$ has been reported in eukaryotes, although there is indirect data in support of such interaction (Choi *et al.* 1998; Marintchev *et al.* 2003).

The structure of archaeal eIF5B was recently determined (Roll-Mecak *et al.* 2000). It consists of four domains and has a chalice-like shape (Fig. 6). The G-domain (domain I) and domain II are homologous in both sequence and structure to the first two domains of the G proteins of

the EF1A (EF-Tu) family, but domain III, although similarly positioned as domain III of EF1A, is unrelated. The C-terminal domain IV (eIF5B-CTD) is connected to the rest of the protein by a long helix and is homologous to domain II.

In this review, we have adopted the above nomenclature based on the structure of the archaeal eIF5B (Roll-Mecak *et al.* 2000) for all IF2/eIF5B homologs. As archaeal eIF5B does not have the long positively charged N-terminal segment found in IF2 from most bacteria and in eukaryotic eIF5B, this segment of up to ~600 residues is not numbered. The correspondence of the above nomenclature with existing nomenclatures of IF2 (see e.g. Moreno *et al.* 1999; Marzi *et al.* 2003) is as follows: (1) in the 6-domain nomenclature for *E. coli* IF2, domains I, II, and III correspond to the N-terminal segment missing in archaea. Domain IV of *E. coli* IF2 corresponds to domain I above (the G domain), domain V – to domain II, and 'domain' VI (the C domain) is in fact composed of two domains (C1 and C2) and corresponds to archaeal domains III and IV. In *B. stearothermophilus* IF2, the N-terminal segment is much shorter than in *E. coli* and designated as one domain (N). Domains G1 and G2 correspond to domains I and II in archaeal eIF5B, and domains C1 and C2 – to archaeal domains III and IV.

There is extensive biochemical information about the location of IF2 on the ribosome (Moreno *et al.* 1999). On the small subunit, domain II of IF2 (the domain immediately following the G domain) interacts with IF1 in the A-site, whereas IF2-CTD binds the acceptor end of fMet-tRNA^{fMet} and upon start site recognition, IF2 is likely coordinately oriented with fMet-tRNA^{fMet} to promote efficient subunit joining (see above and Fig. 4*c*). In the context of the 70S ribosome, IF2 is known to interact with several regions of the large subunit, mainly in and near the GAC. The G domain (domain I) of IF2 interacts with the L7/L12 stalk and the SRL of the large subunit in an orientation similar to that of the elongation factors EF1A and EF2. Even in the absence of direct structural information, the abundant biochemical data has allowed to model the orientation of IF2 on the ribosome, based on the above constraints (Moreno *et al.* 1999; Ramakrishnan, 2002; Marintchev *et al.* 2003) (Fig. 4*c*). A recent report utilizing site-directed hydroxyl radical cleavage is consistent with the present models and clearly demonstrated the different orientations of IF2 between 30S and 70S complexes, but the data were insufficient to provide much further detail (Marzi *et al.* 2003).

Certain questions have not been addressed by the current models, especially with respect to the location of the acceptor end of fMet-tRNA^{fMet}.

(1) As the initial orientation of fMet-tRNA^{fMet} is unlikely to be identical to that after start codon recognition, the initial position of IF2-CTD would also have to be different.
(2) The cavity of the PTC of the large subunit appears too narrow to allow IF2-CTD to be inserted together with the acceptor end of fMet-tRNA^{fMet}. Unless significant conformational changes occur in the PTC, it is unlikely that the acceptor end of fMet-tRNA^{fMet} would be able to be properly positioned in the P-site of the large subunit without being released from IF2. Therefore, at least upon subunit joining, IF2-CTD (and IF2 as a whole) would have to be displaced, compared to the location obtained by placing IF2-CTD in contact with the acceptor end of the P-site tRNA as found in the 70S ribosome. The orientation of the acceptor end of fMet-tRNA^{fMet} remains an open question, as well.

It has been found, that the affinity of IF2·GTP for the A-site of 70S ribosomes depends on the presence of fMet-tRNA^{fMet} in the P-site, indicating that IF2 either directly or indirectly senses the presence of fMet on the P-site tRNA. In the same study, the nature of the P-site tRNA (fMet-, peptidyl-tRNA, or deacylated tRNA) regulated the binding and GTP hydrolysis by

several translation factors, consistent with an indirect read-out. On the other hand, GTP hydrolysis by IF2 was stimulated by both fMet-tRNA^{fMet} and deacylated tRNA in the P-site, but IF2-GMPNP bound stably only when fMet-tRNA^{fMet} was in the P-site (Zavialov & Ehrenberg, 2003).

Sequence comparison of archaeal eIF5B with bacterial IF2 indicates that a small helical subdomain is added to the C-terminus of eIF5B, packing against the C-terminal domain (domain IV) and the connecting helix (H12) (Fig. 6). The other more significant difference is a patch on the surface of the G domain (domain I) of archaeal eIF5B, formed by several insertions brought close in space in the three-dimensional structure of eIF5B. Remarkably, these insertions do not affect significantly the overall shape and dimensions of IF2/eIF5B, consistent with the high degree of conservation of the corresponding surfaces on the ribosome. There are no significant insertions or deletions between archaeal and eukaryotic eIF5B over the entire archaeal eIF5B sequence, with the possible exception of the extreme C-terminus of eIF5B, which is a few residues longer in some archaeal species and disordered in the available crystal structure of archaeal eIF5B (Roll-Mecak *et al.* 2000). The degree of conservation in the structure of IF2/eIF5B and its function in subunit joining supports the idea that the orientation of eIF5B on the ribosome is similar to that of IF2, which would allow Met-tRNA$_i$ (after start site recognition and release of eIF2) to interact with and help orient eIF5B for subunit joining (Roll-Mecak *et al.* 2000, 2001; Marintchev *et al.* 2003; Olsen *et al.* 2003).

Bacterial IF2 (with the exception of some, mostly thermophilic bacteria) and eukaryotic eIF5B have a long N-terminal region of up to ~600 residues, which is absent in archaeal eIF5B. The N-terminus of IF2/eIF5B is positively charged and a large part of it is likely unfolded, although the structure of a small RNA binding domain in bacterial IF2 was recently reported (Laursen *et al.* 2003). The N-terminus of IF2/eIF5B was found to interact with the ribosome, at least in bacteria, and was proposed to bind across the small subunit, extending from the A-site to the E-site (Moreno *et al.* 1999). Consistent with its absence in archaea and some bacteria, the N-terminus of IF2/eIF5B is not essential, but its deletion has been reported to confer a growth defect (Laalami *et al.* 1991). A possible role of the N-terminus of IF2/eIF5B could be ribosome remodeling and facilitation of conformational changes in the small subunit during initiation. The recent finding that a number of IFs may remain transiently associated with the elongating ribosome (Poyry *et al.* 2004; see Section 4.8), raises an interesting question about the fate of eIF1A and eIF5B: the behavior of these two factors was not studied in the above work and they must be removed from their original positions, in order to free the ribosomal A-site. However, both factors are necessary for reinitiation and thus are expected to be present in the reinitiation complexes. Therefore it is possible that one function of the tails of eIF5B and eIF1A is to anchor the proteins to the IC through interactions with the ribosome and/or other factors.

As explained above, IF1 interacts on the ribosome with domain II of IF2 and a similar interaction is likely established in eukaryotes as well. In eukaryotes, the C-terminus of eIF1A (unique to eukaryotes) binds to eIF5B-CTD, providing a second interaction interface, located ~50 Å away on the eIF5B structure (Marintchev *et al.* 2003; Olsen *et al.* 2003). The eIF5B-CTD binding motif of human eIF1A binds to the helical subdomain of eIF5B-CTD and is conserved at the extreme C-termini of most known eukaryotic eIF1A sequences. The length of the C-terminal tail of eIF1A is sufficient to allow a simultaneous interaction with domain II of eIF5B (as in bacteria) and with eIF5B-CTD. The eIF1A peptide binds to the C-terminal helical subdomain of eIF5B-CTD, which first appears in archaea and is not present in bacteria (Marintchev *et al.* 2003).

Certain differences between the eIF1A-eIF5B interactions in human and budding yeast (*S. cerevisiae*) are worth noting. In human, as few as seven C-terminal residues of eIF1A are sufficient for maximal binding to eIF5B, but the interaction is relatively weak (~30 μM) and could not be detected by pull-down assay (Marintchev *et al.* 2003; and A. Marintchev & G. Wagner, unpublished data). The interaction between the yeast proteins has not been quantitated, but appears to be stronger; the two proteins have been co-immunoprecipitated from yeast extracts; and are likely coordinately recruited to the small subunit. Furthermore, the last 24 residues of yeast eIF1A were necessary for stable binding (Olsen *et al.* 2003). The extreme C-termini of human and yeast eIF1A are conserved, but it appears that the C-terminal tail of human eIF1A may be lacking a second binding determinant, present in the yeast protein. Yeast eIF5B-CTD was sufficient for stable binding to both full-length eIF1A and its C-terminal 24 amino-acid fragment, residues 130–153 (Choi *et al.* 2000; Olsen *et al.* 2003). The exact location of the yeast-specific interface remains to be mapped.

In higher eukaryotes (but not in yeast), eIF5 has a C-terminal tail, similar to that of eIF1A, which binds to the same surface of eIF5B-CTD (see Section 5.4.1 below). Human eIF5B co-purifies with eIF5, but not eIF1A, from cell extracts (T. V. Pestova, personal communication). Therefore, the eukaryote-specific interactions between the C-termini of eIF1A and eIF5B are involved in the coordinated recruitment and release of these two factors in *S. cerevisiae* (Olsen *et al.* 2003), but likely only in coordinated release of human eIF1A and eIF5B upon subunit joining, whereas recruitment appears to be coordinated with eIF5.

Although yeast eIF5B performs a number of important functions and its deletion causes severe slow growth, the gene is not essential (Choi *et al.* 1998), unlike IF2, for example (Laalami *et al.* 1991). eIF5B is, however, essential in *Drosophila* (Carrera *et al.* 2000).

5.2 IF3 and eIF1

5.2.1 IF3

IF3 is a ~21 kDa protein composed of two domains, connected by a flexible linker. The structures of both domains of IF3 have been determined (Biou *et al.* 1995; Garcia *et al.* 1995) (Fig. 7*a*). As discussed above, the main roles of IF3 are subunit dissociation/antiassociation, $tRNA^{fMet}$ selection, and start codon selection. The C-terminal domain of IF3 (IF3-CTD) at higher concentrations can perform all functions of full-length IF3 (reviewed in Petrelli *et al.* 2001). The binding site of IF3-CTD on the small ribosomal subunit was mapped to the interface surface of the platform by both cryo-EM (McCutcheon *et al.* 1999) and directed hydroxyl radical cleavage and footprinting (Dallas & Noller, 2001), consistent with the available biochemical data (Fig. 4*c*). IF3-CTD thus appears to compete directly with the large subunit for binding.

Binding of IF3 to the 30S subunit induced significant conformational changes, as seen by cryo-EM (McCutcheon *et al.* 1999), making it difficult to assign unambiguously IF3-NTD to a region of positive electron density in the differential cryo-EM map. IF3-NTD was initially placed on the head, near the neck (McCutcheon *et al.* 1999), but later, directed hydroxyl radical cleavage data were used to assign IF3-NTD to a different region of positive density – on the platform of the 30S subunit, contiguous with IF3-CTD (Fig. 4*c*). Binding of IF3-NTD to the platform was proposed to contribute to steric prevention of tRNA binding to the E-site. The positive density on the head was then ascribed to rotation of the head toward the platform, consistent with directed cleavage and other biochemical data. Directed hydroxyl radical

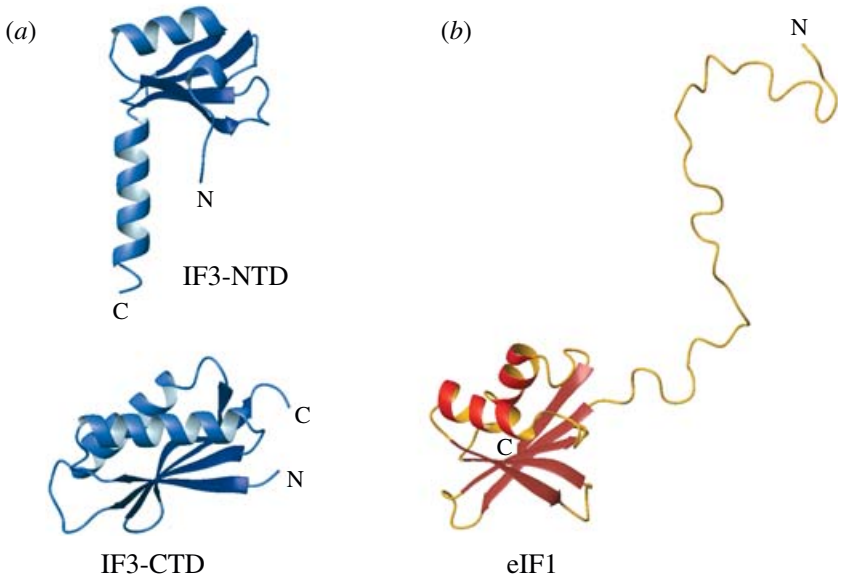

Fig. 7. Structures of IF3 and eIF1. (*a*) Structure of *B. stearothermophilus* IF3. (Top) IF3-NTD, PDB code 1TIF; (bottom) IF3-CTD, PDB code 1TIG. (*b*) Structure of human eIF1, PDB code 2IF1. Note that the linker between the two domains of IF3 [absent from the structures in (*a*)] and the N-terminal tail of eIF1 (*b*) are unstructured in solution. The N- and C-termini of the proteins are labeled.

cleavage data placed IF3 in close contact with fMet-tRNA^{fMet}, consistent with the observed coupling between start codon recognition and IF3 release (Dallas & Noller, 2001).

A reported crystal structure of an IF3-30S subunit complex placed IF3-CTD on the solvent-exposed surface of the small subunit (Pioletti *et al.* 2001). In that study, IF3 was soaked into preformed 30S subunit crystals and crystal packing would have prevented IF3-CTD binding to the location identified by cryo-EM and directed cleavage. Therefore, the location of IF3-CTD in the crystals is most probably an artifact due to the RNA-binding properties of IF3-CTD (Dallas & Noller, 2001). IF3-NTD was also present but not resolved in the above crystals.

IF3 has been crosslinked to IF2 in *E. coli* (Boileau *et al.* 1983). The current models for the location of IF2 on the ribosome place it (and especially its domains II and III) not too far from IF3-CTD, but the crosslinking could instead be to the long N-terminal segment of *E. coli* IF2, which was proposed to bind across the A-, P-, and even E-sites of the small subunit (Moreno *et al.* 1999).

5.2.2 eIF1

eIF1 is a ~12 kDa protein consisting of a single domain and an unstructured N-terminal tail (Fletcher *et al.* 1999) (Fig. 7*b*). eIF1 is important for scanning and start codon selection and cooperates with eIF3 in subunit dissociation/antiassociation (Pestova & Kolupaeva, 2002). Remarkably, eIF1 was recently found to bind to approximately the same site on the small subunit as IF3-CTD in bacteria (Lomakin *et al.* 2003) (Fig. 4*d*). Both IF3 and eIF1 can discriminate against non-AUG codons or AUG codons located at or near the 5′-end. Thus, although eIF1 is not related in sequence or structure to IF3, the two factors have common functions (Pestova & Kolupaeva, 2002; Lomakin *et al.* 2003).

At the same time, eIF1 and IF3 differ in several respects: (1) IF3, but not eIF1, is involved in initiator tRNA selection. Both IF3 and eIF1 induce conformational changes in the small subunit, but only a subset of the changes appears similar. tRNA selection by IF3 is proposed to be indirect and mediated by the head of the small ribosomal subunit (Dallas & Noller, 2001; see Section 5.2.1 above), and there is no indication that such movement of the head occurs upon eIF1 binding (Lomakin *et al.* 2003). (2) eIF1, but not IF3, can discriminate (most likely indirectly) between AUG codons on the basis of their nucleotide context: the Kozak sequence in eukaryotic mRNA. Thus, the conformational changes induced by eIF1 must be consistent with scanning and recognition of the Kozak sequence around the start codon (Pestova & Kolupaeva, 2002; Lomakin *et al.* 2003).

eIF1 converts the 43S complex from a 'closed' conformation, able to bind to any codon even with partial basepairing of the codon and anticodon, to an 'open' conformation, that can only bind to an AUG codon in a proper sequence context (Pestova & Kolupaeva, 2002). A different way to view the 'open/closed' phenomenon is that the 43S complexes are able to bind any codon in a 'closed' conformation, whereas eIF1 binding reverses and prevents this 'closed' conformation and only AUG codons in a 'good' context are able to resist dissociation by eIF1. During scanning, the 43S complex slides along mRNA constantly (but unsuccessfully) attempting to 'close', until the proper start codon is reached. Then a stable 'closed' complex is formed, eIF2 hydrolyzes GTP, and eIF2·GDP is released. Initiation on non-canonical sites occurs when the 43S complex escapes the control of eIF1 and the large subunit binds to the 'closed' complex before eIF1 has been able to dissociate it. Once an 80S IC is formed, it is resistant to eIF1 action.

A number of mutations in eIF1 have been reported. Start site selection by human eIF1 was not affected by deletion of the N-terminal tail, at least *in vitro*. However, eIF1 proteins with single point mutations on several surfaces of the folded domain were defective to various degrees, indicating that eIF1 is finely optimized for start site selection and even minor changes can have an effect (Lomakin *et al.* 2003). Sui mutations have been reported in yeast eIF1 (Yoon & Donahue, 1992). They all map to a conserved solvent-exposed surface of eIF1 (Fletcher *et al.* 1999) and are not expected to directly contact the 40S subunit (Lomakin *et al.* 2003). Accordingly, their Sui phenotype is recessive, indicating that the recruitment of eIF1 is affected (see Section 4.9 above). These mutations could affect binding of eIF1 to eIF3, required for stable association of eIF1 with the scanning IC. Yeast eIF1 was found to bind eIF5, but the interaction was localized to the unstructured N-terminus of eIF1. Surprisingly, one of the Sui mutations in yeast, D88G was reported to weaken eIF1 binding to eIF4G, but not to eIF3 (He *et al.* 2003). The factors involved in the interaction between the cap-binding complex and the 43S complex appear to differ between yeast and mammals (see Section 4.6.1 above). Furthermore, it is not clear if mammalian eIF1 binds to eIF4G. At the same time, human eIF1 can functionally substitute for its yeast homolog (Cui *et al.* 1998). It is thus interesting to know the effects of mutations in human eIF1, corresponding to the yeast Sui mutations.

Homologs of eIF1 (YciH in *E. coli*) are present in some bacteria, but their functions have not been elucidated. The structure of YciH has been determined (Cort *et al.* 1999). Not only the structure of YciH, but also its surface charge distribution, is conserved with eIF1, lending support for the hypothesis, that YciH could be an alternative IF in bacteria (its absence from most bacterial genomes rules out the possibility that it is a general IF). One could only speculate whether or not the ancestors of bacteria and eukaryotes have possessed both IF3 and eIF1 homologs but have retained one or the other.

5.3 eIF2

eIF2 is a heterotrimer, comprised of one each of an α, β and γ subunits. Its primary functions are selection and recruitment of Met-tRNA$_i$ to the 40S ribosomal subunit and control of start site recognition. eIF2·GTP binds specifically to Met-tRNA$_i$, whereas GTP hydrolysis or loss of the methionine moiety weakens the interaction of eIF2 with the initiator tRNA. The eIF2·GTP·Met-tRNA$_i$ ternary complex can bind stably to the 40S ribosomal subunit and the interaction is further stabilized by other factors. Upon start site recognition, eIF2 hydrolyzes GTP and is released as eIF2·GDP from the Met-tRNA$_i$. The basal GTPase activity of eIF2 is very low and is significantly increased only when eIF2·GTP·Met-tRNA$_i$, as part of the IC, is at the translation start site, most probably through conformational changes in the complex. In eukaryotes, but not in archaea, GTP hydrolysis also requires the GAP factor eIF5. Eukaryotic eIF2·GDP is recycled to eIF2·GTP by the GEF eIF2B, whereas archaea do not seem to have an eIF2 GEF (reviewed in Hinnebusch, 2000; Dever, 2002).

Structures are available for archaeal eIF2γ (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004), eIF2α (Nonato *et al.* 2002; Dhaliwal & Hoffman, 2003; Ito *et al.* 2004), and most of archaeal eIF2β (Cho & Hoffman, 2002; Gutierrez *et al.* 2004) (Fig. 8). The largest subunit, eIF2γ is a G protein, homologous to the elongation factors EF1A (EF-Tu), eEF1A and the selenocysteine-specific factor SelB/eEFSec (Leibundgut *et al.* 2005). The structure of archaeal eIF2γ was determined in the apo-form (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004) and GTP- and GDP-bound forms (Schmitt *et al.* 2002) and not surprisingly, was found to be very similar to those of the EFs. The high degree of structural conservation allowed the authors to model the interaction of eIF2γ with Met-tRNA$_i$. In addition to the three-domain core, corresponding to the EFs, archaeal eIF2γ has a small Zn-binding domain (ZBD) inserted in the G domain The corresponding inserted segment in eukaryotic eIF2γ is even longer and essential for viability in yeast (Erickson *et al.* 1997), but the lack of obvious conservation of all cysteines makes it uncertain if it also binds zinc. Yeast eIF2γ, but not human eIF2γ, also has a long N-terminal tail, whose deletion is tolerated *in vivo*, but a point mutation in it causes a growth defect (Erickson *et al.* 1997).

The crystal structures of archaeal eIF2γ indicate a major difference to EF1A: the apo- and GDP-bound forms of eIF2γ have the same closed 'active'-like domain orientation as the GTP-bound form of EF1A (and eIF2γ), instead of the open conformation found in apo-EF1A and EF1A·GDP (see Section 3.2.1 and Fig. 3 above). The orientation is unlikely to be due to crystal packing, because it is found in the structures from two different organisms, in different crystal forms (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004). Consistent with the structural data, recent biochemical experiments showed that the affinity of eIF2·GDP for Met-tRNA$_i$ is only $\sim$20-fold lower than that of eIF2·GTP. Furthermore, the affinities of both eIF2·GTP and eIF2·GDP for deacylated tRNA$_i$ were approximately equal to that of eIF2·GDP for Met-tRNA$_i$. Therefore, the discrimination between GDP- and GTP-bound eIF2 and between Met-tRNA$_i$ and deacylated tRNA$_i$ is mainly through interactions with the methionine, whereas the rest of the binding interface appears unperturbed, at least in solution (Kapp & Lorsch, 2004). The Switch 1 and Switch 2 regions of the G-proteins, including those of the EF1A family, respond to ligand and nucleotide binding. In the context of the above discussion, it seems that the cooperativity between Met-tRNA$_i$ and GTP binding to eIF2 is mediated by their direct interactions with, and induction of conformational changes in, the switch regions.

Although the structure of eIF2γ indicates that it carries the necessary determinants for both Met-tRNA$_i$ and GTP binding, eIF2 exists as a heterotrimer and all three subunits are essential

[if the eIF2 β and γ subunits are overexpressed, the α subunit is no longer essential (Erickson *et al.* 2001)]. The eIF2 α and β subunits bind to the γ subunit, but not to each other. Bacterial EF1A (EF-Tu) is a monomer and eEF1A in its GDP- and GTP-bound forms is in complex with the GEF eEF1B, but eEF1B is released upon aa-tRNA binding. One distinctive feature of eIF2 is that, unlike the EFs or IF2/eIF5B, GTP hydrolysis by eIF2 is not activated by the GAC of the large ribosomal subunit, but is coupled to start site recognition on the small subunit. The eIF2 α and β subunits are likely to participate in one or more of the following activities: modulation of Met-tRNA$_i$ binding, GTP binding and hydrolysis in free eIF2, regulation and coordination of ribosome binding, start codon recognition, GTP hydrolysis and release of eIF2·GDP from the IC. It was recently reported, that archaeal eIF2γ binds specifically Met-tRNA$_i$. The binding was not affected by the presence or absence of the β subunit. The α subunit or its CTD stabilized Met-tRNA$_i$ binding by $\sim$50-fold (Yatime *et al.* 2004). In contrast, yeast eIF2α only stabilized Met-tRNA$_i$ binding to eIF2$\beta\gamma$ by 5- to 10-fold (Nika *et al.* 2001). The eIF2β subunit is at least in proximity to the tRNA, as it has been cross-linked to it (Gaspar *et al.* 1994). Some of the cross-links were to the positively charged N-terminal tail of eIF2β, which is not present in archaea and could participate in Met-tRNA$_i$ binding in eukaryotes.

Structures for the NTD of eIF2α have been published (Nonato *et al.* 2002; Dhaliwal & Hoffman, 2003) and the structure of full-length eIF2α was recently determined in our laboratory (Ito *et al.* 2004) (Fig. 8*a*). The NTD is composed of two subdomains: an S1-like OB fold, followed by a helical subdomain. The conserved S51, responsible for regulation of eIF2B activity (see Section 5.4.2 below) is located on a flexible loop. The full-length eIF2α is composed of the above NTD, a CTD and a long C-terminal tail that appears mostly unstructured, but contains an α-helix (Ito *et al.* 2004).

Although lacking sequence homology, the eIF2α-CTD is structurally homologous to the CTD of eEF1Bα (eEF1Bα-CTD). eEF1Bα is the GEF for eEF1A. Furthermore, eEF1Bα-CTD corresponds to the entire archaeal eEF1B and is sufficient for the GEF activity in eukaryotes. eEF1Bα-CTD binds to domain II of eEF1A, near the binding site for the aminoacyl end of the aa-tRNA, and interacts with domain I near the nucleotide-binding pocket (Andersen *et al.* 2000; see Section 3 above). Thus, the eEF1A/eEF1Bα-CTD and eIF2γ/eIF2α-CTD complexes appear to form structurally homologous pairs (Ito *et al.* 2004). Consistently, the eIF2α-interacting region of eIF2γ was mapped to the surface of domain II of eIF2γ, corresponding to the eEF1Bα-binding surface of eEF1A (Roll-Mecak *et al.* 2004), and eIF2α-CTD was found to mediate the interaction with eIF2γ (Yatime *et al.* 2004). If the orientation between eIF2α and γ is similar to that between eEF1A and eEF1Bα, the eIF2α-NTD would be positioned near the nucleotide-binding site of eIF2γ, similarly to the regulatory eEF1Bα-NTD. Furthermore, eIF2α has been reported in one study to be in proximity to the eIF2-bound GTP (Bommer *et al.* 1988). Unfortunately, no detailed studies of GTP exchange and hydrolysis by the isolated eIF2γ subunit or the eIF2$\alpha\gamma$ and eIF2$\beta\gamma$ subcomplexes have been reported.

Unlike eEF1B, which dissociates upon aa-tRNA binding to eEF1A, eIF2α and γ remain stably associated and the α subunit even stimulates Met-tRNA$_i$ binding by eIF2γ. It is difficult to draw the parallel between eEF1B and eIF2α further, to suggest a GEF function for eIF2α in eukaryotes, because a yeast eIF2$\beta\gamma$ subcomplex lacking eIF2α was reported to have a moderate defect in Met-tRNA$_i$ binding, but near-WT rates of nucleotide exchange (Nika *et al.* 2001). Furthermore, whereas the surface of eIF2α-CTD expected to interact with domain II of eIF2γ is highly conserved, there is no strong conservation of the opposite surface of eIF2α-CTD, where the GEF activity would reside. It is possible, that in archaea, eIF2α-CTD does have

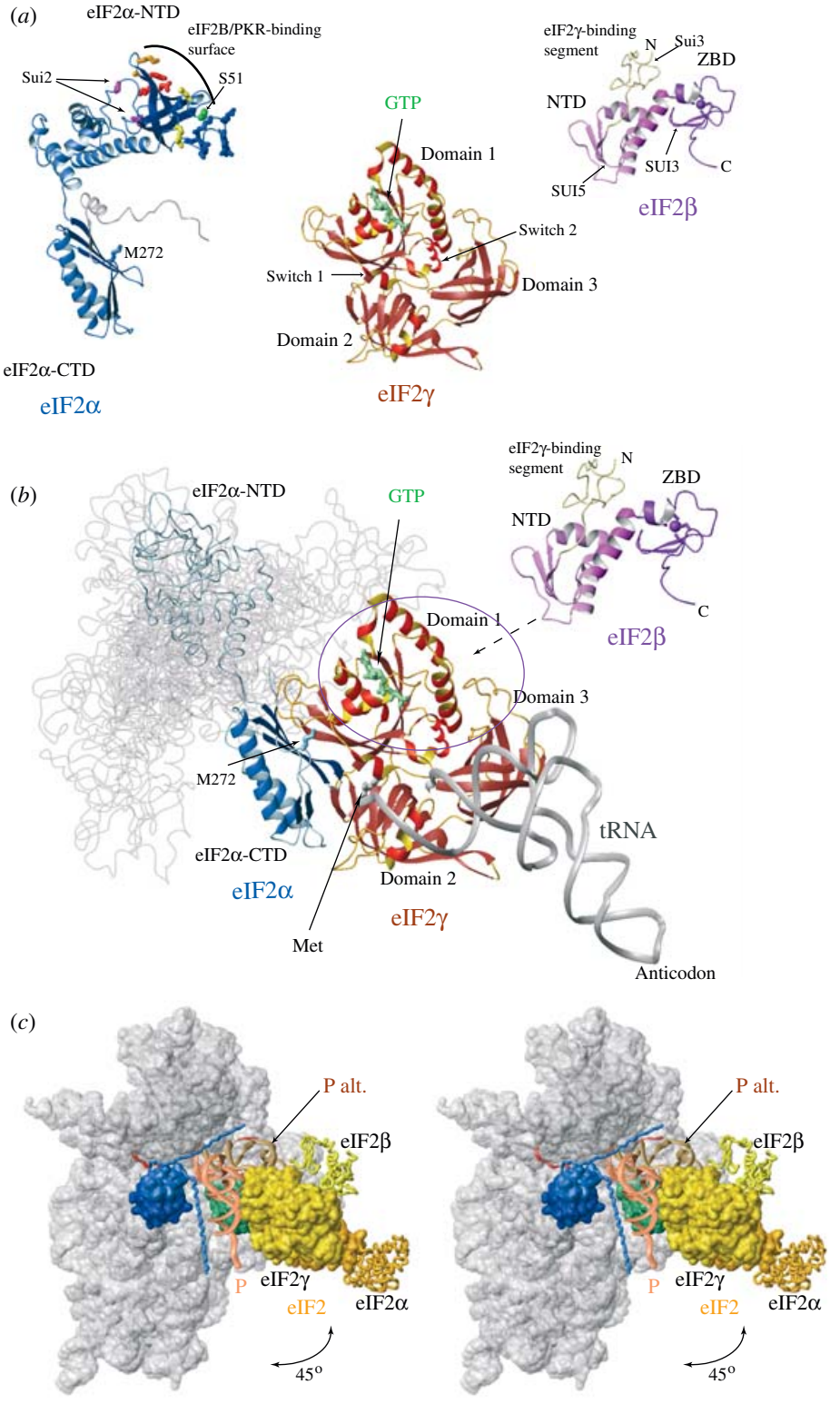

Fig. 8. For legend see opposite page.

a GEF activity, modulated by the presence of Met-tRNA$_i$ and/or the NTD. Consistent with this hypothesis, in the model for the interaction of eIF2α and γ (Ito *et al.* 2004), M272 is directed toward Switch 1 of eIF2γ. Met272 is conserved in eukaryotes, located in a bulge and its side chain is solvent-exposed. In addition, similar to the catalytic K205 in eEF1Bα-CTD, a conserved basic residue is found in archaea at the corresponding position near the C-terminus. Although K205 from eEF1Bα corresponds to V277 in eIF2α, located several Å away, the relative orientation of domains I and II differs between eEF1A and eIF2γ. Therefore, M272 is structurally better positioned for contact with the nucleotide. The unusual bulged conformation around M272 in eukaryotic eIF2α could allow it to retain the same orientation as in archaea, where it

Fig. 8. eIF2. (*a*) Structures of eIF2 subunits. *Left*, human eIF2α, PDB code 1Q8K, is displayed in blue. The C-terminal tail, absent in archaea, is in grey. The side-chain of S51 (the target for phosphorylation) is in green and labeled). Note that both the NTD and the C-terminal tail are mobile with respect to the CTD [see text and panel (*b*)]. The surface involved in interactions with PKR and eIF2B is labeled; the side-chains of residues important for PKR binding are displayed in red, those important for eIF2B binding in yellow, and those important for both in orange. Residues, forming a conserved positively charged patch on the NTD, are displayed in blue. The residues, whose mutations lead to Sui phenotype are shown in magenta and labeled. Met272, which could be involved in nucleotide exchange in archaeal eIF2 (see text) is shown in light blue and labeled. *Center*, archaeal eIF2γ from *P. abyssi* in complex with GTP, PDB code 1KK1, is shown in red and orange, and GTP in green. The three domains and Switch regions 1 and 2 are labeled. *Right*, the structure of eIF2β from *M. thermoautotrophicum*, PDB code 1NEE, is shown in violet (the NTD) and purple (the Zn-binding domain, ZBD). The N-terminus, which is involved in eIF2γ binding and unstructured in free eIF2β, is in beige. The zinc in the ZBD is shown as a sphere, and the coordinating cysteine side-chains are also shown. The N- and C-termini of the proteins are labeled, except for eIF2γ. The general locations of Sui mutations in the N-terminus and SUI mutations in the ZBD of eIF2β (Sui3 and SUI3 respectively) are shown with arrows. The location of the SUI mutation in the homologous eIF5 (SUI5) is also shown. Eukaryotic eIF2β has an additional long N-terminal segment, missing in archaea. Note that the orientation between the NTD and the ZBD of eIF2β is taken from 1NEE, but the two domains are mobile with respect to each other. (*b*) Model for the organization of the eIF2·GTP·Met-tRNA$_i$ complex. The structures of eIF2γ and eIF2β are displayed as in (*a*). eIF2α-CTD is as in (*a*). The CTDs of the ensemble of fifteen NMR structures of eIF2α, PDB code 1Q8K, are aligned (but not shown). One NTD is displayed as a wire and the rest of the NTDs are displayed as semi-transparent wires, in order to illustrate the space accessible to the eIF2α-NTD with respect to eIF2α-CTD and the rest of eIF2. Coloring of the eIF2 subunits is as in (*a*). The C-terminal tail of eIF2α [grey in panel (*a*)] is not shown. The orientation of Met-tRNA$_i$ (shown in grey) is modeled as in (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004): domain 2 of eIF2γ was aligned to domain 2 of EF1A (EF-Tu) from a ternary complex with GTP and Cys-tRNA^{Cys}, PDB code 1B23, and the resulting position of the tRNA was not modified. The Cys residue was replaced by Met and is labeled. The orientation of eIF2α-CTD is from the model in (Ito *et al.* 2004). eIF2β is likely to contact part of the remaining conserved surface of eIF2γ near the nucleotide-binding site (circled). (*c*) Stereo view of the possible orientation of the eIF2·GTP·Met-tRNA$_i$ complex. The model for the eIF2·GTP·Met-tRNA$_i$ complex from (*b*) was oriented with respect to the small ribosomal subunit by aligning the tRNA (brown) with the alternative orientation of the initiator tRNA in Fig. 4 above, consistent with biochemical data. Aligning Met-tRNA$_i$ in the eIF2·GTP·Met-tRNA$_i$ complex to the P-site tRNA (shown in coral) from the 70S ribosome would have projected eIF2 away from the ribosome surface. The bacterial 30S ribosomal subunit (in semi-transparent surface representation, colored in grey), the mRNA (brown) and the P-site tRNA (coral) are from the structure of the 70S ribosome, PDB code 1JGP (Yusupova *et al.* 2001). The locations of eIF1 (green) and eIF1A (blue) are as in Fig. 4*d* above. The segment of eIF1A, homologous to IF1 is in dark blue. The unstructured N- and C-terminal tails of eIF1A are shown as extended wires, in order to compare their lengths with the distances to other eIFs. eIF2γ is in yellow, eIF2α in orange, and eIF2β in light yellow. eIF2γ and eIF2α-CTD are in surface representation. The eIF2α-NTD (whose orientation is flexible) and eIF2β (whose orientation with respect to the rest of eIF2 is unknown) are shown as wires, to give a perspective for the overall size of eIF2. The C-terminal tail of eIF2α (absent in archaea) is not shown.

is near the C-terminus, but the role of M272 in eukaryotes is likely restricted to interaction with Switch 1.

The GEFs of the EFs have other functions, in addition to nucleotide exchange. They have been found, for example, to assist in aa-tRNA loading from the aa-tRNA synthetase directly to eEF1A (see Section 3 above). By analogy, eIF2α could also be involved in Met-tRNA$_i$ transfer from the synthetase to eIF2, but such a role would not be discovered *in vitro* using free Met-tRNA$_i$.

The NMR data indicate that the eIF2α NTD and CTD do not bind to each other, but are not completely independent, either. The mobility of the two domains with respect to one another is restricted by the short linker and could be further limited in the context of the eIF2 heterotrimer (Ito *et al.* 2004). These findings could provide structural explanation for recently reported data, that deletion of (most of) the CTD of eIF2α makes *in vitro* binding of eIF2α-NTD to eIF2B independent of S51 phosphorylation (Krishnamoorthy *et al.* 2001), indicating that an eIF2B-binding surface is buried or sterically occluded by the eIF2α-CTD. S51 phosphorylation could increase the affinity of eIF2α-NTD for eIF2B or act indirectly by changing the electrostatic environment between the NTD and CTD, thus increasing the accessibility of the NTD. However, NMR data showed only minor effects on the CTD in the presence of the phosphomimetic mutation S51D, at least at high salt concentration (Ito *et al.* 2004). The interactions between eIF2 and eIF2B will be discussed later in Section 5.4.2, in the context of eIF2B structure and function.

The function of the C-terminal tail of eIF2α (after the CTD) is not known. This segment is not present in archaeal eIF2α and is likely involved in regulation. It was reported to contain a caspase cleavage site and that eIF2α is cleaved during apoptosis (Satoh *et al.* 1999; Bushell *et al.* 2000; Clemens *et al.* 2000; Marissen *et al.* 2000). One of the groups found that caspase cleavage of eIF2α inactivated eIF2 and made nucleotide exchange independent of eIF2B (Marissen *et al.* 2000). Yeast eIF2α (but not human eIF2α) has phosphorylation sites at its C-terminus, and their mutation exacerbated the growth defects of Gcd$^-$ eIF2Bα and δ mutants (Feng *et al.* 1994). The structure of the full-length eIF2α indicated, that the C-terminal tail could be near the eIF2α-NTD (Ito *et al.* 2004).

eIF2β contains two domains conserved in all species: an NTD and a ZBD (Cho & Hoffman, 2002) (Fig. 8*a*), as well as an eIF2γ-binding segment immediately preceding the NTD (Thompson *et al.* 2000). Eukaryotic eIF2β also contains a positively charged N-terminal tail, which binds to the GAP and GEF, eIF5 and eIF2B respectively (reviewed in Dever, 2002). The structures of the two domains of archaeal eIF2β were determined (Cho & Hoffman, 2002; Gutierrez *et al.* 2004). The two domains are connected by a relatively flexible helical region. No interactions were reported between the eIF2β-NTD and eIF2β-ZBD (Cho & Hoffman, 2002; Gutierrez *et al.* 2004).

The available structural and other information has finally started to provide clues about the detailed mechanism of action of eIF2 and especially the roles of the α and β subunits. The function of the α subunit was discussed above. There is some information about the properties of eIF2β-depleted eIF2$\alpha\gamma$ subcomplexes *in vitro*. The properties of eIF2β-depleted eIF2$\alpha\gamma$ subcomplexes differed among studies in whether a defect in Met-tRNA$_i$ binding was observed, but there were no reports of any serious nucleotide binding defect (Flynn *et al.* 1993 and references therein). This is in sharp contrast with the properties of eIF2 proteins with point mutations in the ZBD or in the eIF2γ-binding region, which led to increase in spontaneous eIF2 GTP hydrolysis and some of which were lethal *in vivo* (Donahue, 2000; Thompson *et al.* 2000;

Hashimoto *et al.* 2002). There are different ways to rationalize this discrepancy. It is theoretically possible that eIF2β-depleted subcomplexes contained undetected proteolytic fragments of eIF2β, sufficient to retain some activity; or the spontaneous GTP hydrolysis rates in eIF2β-depleted subcomplexes may not have been studied. It is also possible that the main roles of eIF2β are associated with processes like start site recognition occurring in the context of the IC. In eukaryotes, β-less eIF2 is certain to have a defect due to its inability to stably interact with eIF5 and eIF2B (which bind to the N-terminal tail of eIF2β), but this cannot be the only role of eIF2β, because it is present in archaea, which do not have eIF5 or eIF2B.

Two groups of Sui mutations have been described in eIF2β: mutations in and around the Zn-binding motif (dominant SUI phenotype), and in the eIF2γ-binding region (recessive Sui phenotype). Both types of mutations caused elevated rates of spontaneous GTP hydrolysis. The Sui mutations in the γ-binding region caused moderate decrease in binding to eIF2γ, whereas mutations with more severe defect were lethal (Donahue, 2000; Hashimoto *et al.* 2002). Thus, it appeared that the ZBD down-regulates GTP hydrolysis and its function is impaired in the Sui mutants. Weakening of the eIF2β-eIF2γ interaction could loosen the association of the ZBD with its target (the γ subunit and/or GTP). Increased rate of GTP hydrolysis by eIF2 on the small subunit could relax the checkpoint for start codon selection, even in the presence of eIF1, and facilitate initiation at suboptimal sites. A more severe defect of the eIF2β-γ interaction would cause destabilization of the interactions of eIF2 with eIF2B, eIF5 and the IC, any of which is sufficient to cause lethality (Donahue, 2000; Thompson *et al.* 2000; Hashimoto *et al.* 2002). Thus, there is clear indication that, while eIF2 is intrinsically able to hydrolyze GTP, the rate of GTP hydrolysis is strictly regulated both on and off the ribosome. In summary, GTP hydrolysis coupled with start codon recognition must involve a form of derepression of the intrinsic ability of eIF2 to hydrolyze GTP, and not just activation by eIF5, an internal region of eIF2 or by another component of the IC.

The interaction interface between the β and γ subunits of eIF2 is not known, but multiple segments of eIF2β are in proximity to tRNA and/or GTP: residues from the NTD, the ZBD and the N-terminal tail have all been cross-linked to tRNA (Gaspar *et al.* 1994); and a C-terminal region of the β subunit has been cross-linked to GTP (Bommer *et al.* 1991a). The structure of eIF2γ illustrates that extensive surfaces in the vicinity of both Met-tRNA$_i$ and the nucleotide are highly conserved from archaea to eukaryotes (see for example, Fig. 4 in Roll-Mecak *et al.* 2004). With portions of these conserved regions already ascribed to Met-tRNA$_i$ and eIF2α binding (Schmitt *et al.* 2002; Ito *et al.* 2004; Roll-Mecak *et al.* 2004; Yatime *et al.* 2004), the remaining conserved surface in domain I near the nucleotide most likely contains the eIF2β-binding site. One such speculative model of eIF2β interaction with eIF2γ·GTP·Met-tRNA$_i$ was recently proposed (Gutierrez *et al.* 2004), placing eIF2β in contact with domain I of eIF2γ, although it is not consistent with all available experimental data. While the main eIF2γ-binding site on eIF2β is known (Thompson *et al.* 2000) and the eIF2β-binding site on eIF2γ will probably be known soon, sites of additional, functional interactions between the two subunits may be more difficult to define and could include all regions of eIF2β. Thus, the bulk of the α, β and γ subunits appear to form a more or less compact assembly around the acceptor end of the initiator tRNA, possibly with protruding segments from individual subunits (Fig. 8*b*).

The α and γ subunits of eIF2 have been cross-linked to several 40S ribosomal proteins: S3, S3a, S6, S13/16, S15 and S15a (Westermann *et al.* 1979). S3a and S6 also cross-linked to Met-tRNA$_i$ (Westermann *et al.* 1981). Of these ribosomal proteins, only S15 has a bacterial

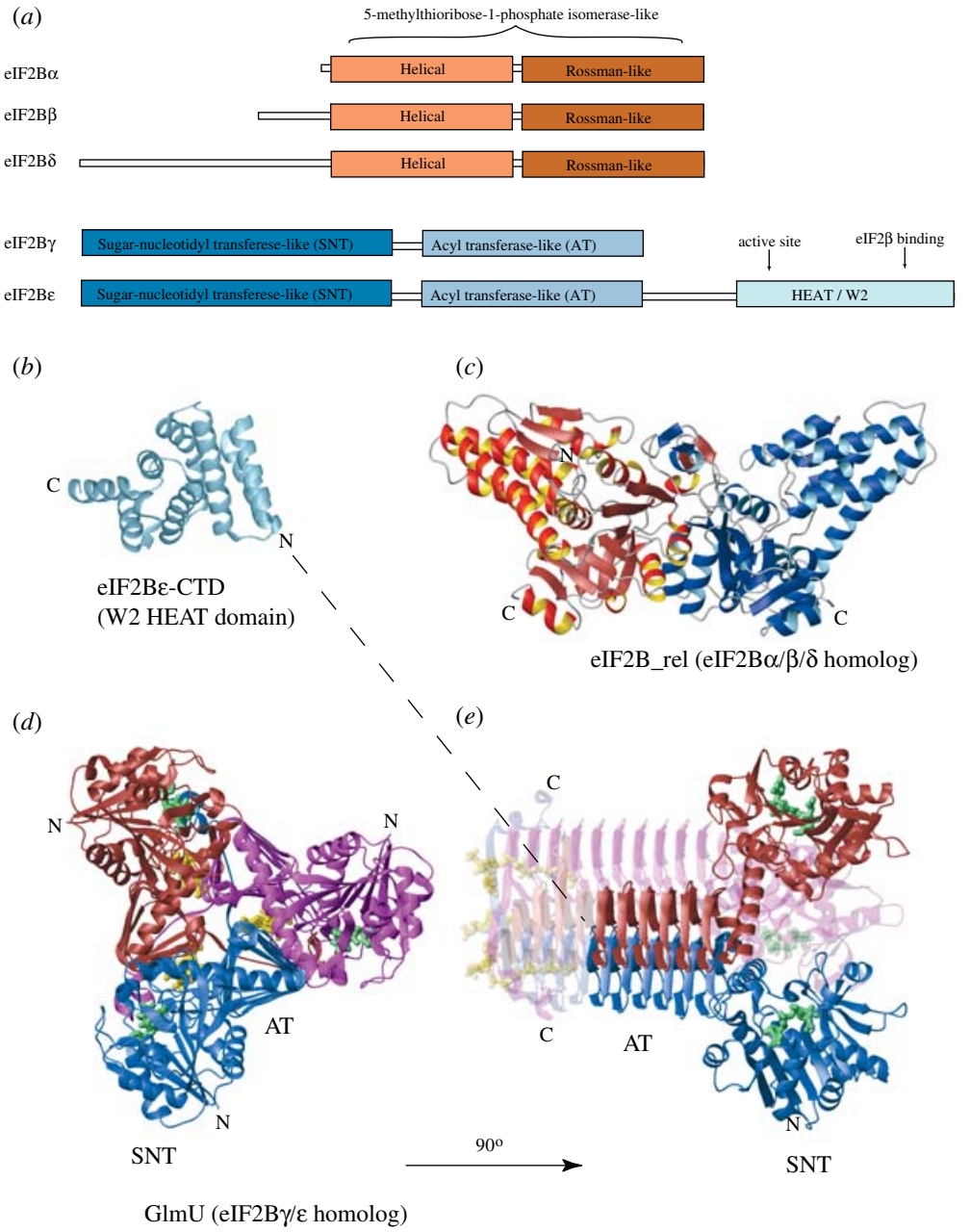

Fig. 9. Structures of eIF2Bε-CTD and eIF2B homologs. (*a*) Domain organization of the eIF2B subunits. (*b*) Structure of the W2 HEAT domain at the C-terminus of eIF2Be, PDB code 1PAQ. The N- and C-termini of the protein are labeled. (*c*). Structure of a member of the eIF2B_rel subfamily of proteins homologous to the eIF2B*α*/*β*/*δ* subunits, from the archaeon *A. fulgidus* (PDB code 1T5O). One subunit is colored brown (strands) and red/yellow (helices), the second dark blue (strands) and blue/light blue (helices). The N- and C-termini of the proteins are labeled, where visible. (*d*) Structure of the bifunctional enzyme N-acetyl-glucosamine-1-phosphate uridyltransferase (GlmU) in complex with acetyl-coenzyme A (AcCoA) and UDP-N-acetylglucosamine (UDP-GlcNAc), PDB code 1HM9. GlmU is homologous to the sugar-nucleotidyl transferase-like (SNT) and acyl transferase-like (AT) domains of eIF2B*γ* and *ε*. The three GlmU subunits are colored blue, brown and magenta respectively. UDP-GlcNAc is in green, and AcCoA is

homolog, S19, which is located on the head. In the cryo-EM reconstruction of the yeast ribosome, the regions of unassigned electron density (suggested as possible locations of eukaryote-specific ribosomal proteins) closest to the P-site of the 40S subunit are on solvent-exposed surfaces of the head and the platform (Spahn *et al.* 2001) (Fig. 4*d*), indicating that eIF2 could be in proximity to one or more of these regions. Consistently, although the exact locations of the eukaryote-specific ribosomal proteins are not known, S3a was reported to be close to the 3′-end of the 18S rRNA. The available information about the structure and interactions of eIF2 suggest a possible general orientation of the eIF2.GTP.Met-tRNA$_i$ ternary complex on the small ribosomal subunit (Fig. 8*c*). While eIF2 could bind to either the head, or the platform, or both, an attractive possibility is that it contacts both. If that is the case, movement of Met-tRNA$_i$ with respect to either the head or the body, and movement of the head with respect to the rest of the 40S subunit (such as upon codon-anticodon base-pairing) could cause breaking or rearrangement of eIF2–Met-tRNA$_i$, eIF2–ribosome, eIF2–eIF5 interactions, or interactions within eIF2, thus promoting coordinated GTP hydrolysis.

Yeast eIF2 was also found to bind to eIF3c (NIP1) (Valasek *et al.* 2002) and to the N-terminal tail of eIF1A, which also interacts with eIF3 (Olsen *et al.* 2003; see Section 5.5 below).

5.4 Eukaryotic factors required for eIF2 function

5.4.1 eIF5

As explained above, in eukaryotes eIF5 acts as a GAP for eIF2, but only on the ribosome. eIF5 is composed of three domains. The first two domains are homologous to the two domains of eIF2β whose structure is known (the NTD and the ZBD, Fig. 8*a*). The length and sequence of the linker between the two domains are also conserved. A unique feature of the NTD of eIF5 is an insertion of a few residues, whose sequence is conserved among eIF5 proteins and slightly resembles a nucleotide-binding motif. eIF5 does not have the eIF2γ-binding segment of eIF2β or the eukaryote-specific N-terminal tail of eIF2β, but instead has a CTD, which binds to the N-terminal tail of eIF2β and is important for association with eIF2. eIF5-CTD (W2 domain) is homologous to the CTDs of eIF2Bε and mammalian eIF4G. The eIF5-CTD is connected to the ZBD through a long mainly negatively charged linker (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Pestova *et al.* 2000, 2001; Dever, 2002). The structure of the eIF2Bε W2 domain was recently reported (Boesen *et al.* 2004), showing unexpectedly, that the W2 domain is a member of the HEAT domain family, as is the middle domain of eIF4G (Marcotrigiano *et al.* 2001) (see Fig. 9*b*). Accordingly, we will call the W2 domains 'W2 HEAT' domains to emphasize this newly established structural homology. The W2 HEAT domain of eIF2Bε also binds to the eIF2β N-terminal tail (Asano *et al.* 1999).

eIF5 was reported by several groups to act as a classical GAP – by stabilizing a transition state in GTP hydrolysis (Asano *et al.* 2001; Das *et al.* 2001; Paulin *et al.* 2001), but direct proof is difficult to obtain, because eIF5 does not stimulate GTP hydrolysis of free eIF2. eIF5 could be involved in derepressing the intrinsic eIF2 GTPase activity in addition to, or even instead of,

---

in yellow. (*e*) Same as in (*d*), but the structure was rotated 90°. One of the subunits and the C-terminal segments of the other two (beyond the region corresponding to eIF2B$\gamma\varepsilon$) are semi-transparent in order to illustrate what an eIF2B$\gamma\varepsilon$ heterodimer might look like, and the location, to which eIF2Be-CTD would be connected is marked with dashed line. The SNT and AT domains of one of the subunits are labeled. The N- and C-termini of the proteins are labeled, where visible.

acting as a classical GAP. In favor of direct GTPase activation by eIF5, a SUI mutation was isolated in the eIF5-NTD – in the conserved eIF5-specific loop (Huang *et al.* 1997; see above). The conserved lysine residue, K15 near the N-terminus of eIF5, which was proposed to be responsible for the GAP activity of eIF5 (Das *et al.* 2001; Paulin *et al.* 2001) is located a few residues before the eIF2β-NTD in a region that appears to be flexible, at least in free eIF2β (Cho & Hoffman, 2002; Gutierrez *et al.* 2004), and its orientation is impossible to infer. A second conserved lysine, K48 proposed to have a role in catalysis (Paulin *et al.* 2001) is on the opposite surface of the NTD with respect to the position of the eIF5 SUI mutation, arguing against direct involvement of K48 in nucleotide exchange.

A more complex role for eIF5 is supported by the fact that both the ZBD (which in eIF2β is responsible for repression of spontaneous GTP hydrolysis) and the NTD (which in eIF5 is proposed to be involved in GTPase activation) are conserved between eIF2β and eIF5. Homologous domains from eIF5 and eIF2β could compete for interaction with GTP, eIF2γ and/or the 40S subunit. Furthermore, if the NTDs of eIF2β and eIF5 are 'activating' (or 'derepressing') they could, in turn compete with the 'repressing' ZBD.

In higher eukaryotes, eIF5-CTD is followed by a $\sim$40-residue long negatively charged tail, whose extreme C-terminus resembles that of eIF1A. The 9-residue C-terminus of human eIF5 binds to the same surface of eIF5B-CTD as the C-terminus of eIF1A, but $\sim$2-fold tighter ($K_D = 15$ μM). A 39-residue C-terminal fragment and full-length human eIF5 bind with even higher affinity ($\sim$3 μM) and eIF5B-CTD is sufficient for the binding (A. Marintchev, T. V. Pestova & G. Wagner, unpublished data). This appears different from human eIF1A, whose extreme C-terminus binds to eIF5B as tightly as full-length eIF1A (Marintchev *et al.* 2003), but is reminiscent of the yeast eIF1A–eIF5B interaction, that also requires a longer C-terminal tail of eIF1A for maximum binding affinity (Olsen *et al.* 2003). There is, however, no obvious sequence similarity between the C-terminal tails of human eIF5 and yeast or human eIF1A beyond the extreme C-termini, except an overall negative charge. The C-terminal tail of human eIF5 is unstructured, but contains a region of significant hydrophobicity, which has negligible affinity for eIF5B (A. Marintchev & G. Wagner, unpublished data), but could be involved in an (yet unidentified) interaction with another factor. It should, however, be noted that the above studies of eIF5B–eIF1A and eIF5B–eIF5 interactions were performed with a truncated eIF5B lacking its long N-terminal region.

5.4.2 eIF2B

eIF2B, the guanine nucleotide exchange factor (GEF) for eukaryotic eIF2, is a heteropentamer, composed of α, β, γ, δ, and ε subunits (Fig. 9*a*) and is among the largest known GEFs. The C-terminal $\sim$20 kDa region of the ε subunit contains the eIF2β-binding W2 HEAT domain (Fig. 9*b*) and the active site, and can catalyze eIF2 nucleotide exchange, suggesting that the rest of the $>$250 kDa complex has mainly accessory and regulatory functions (Gomez *et al.* 2002; Boesen *et al.* 2004). All subunits are conserved between yeast and human and, with the exception of the α subunit, are essential in yeast. The α, β and δ subunits form a 'regulatory' subcomplex, and the γ and ε subunits – a 'catalytic' subcomplex, which is also involved in regulation (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Dever, 2002).

Regulation of eIF2B activity. eIF2B is one of the two major targets for regulation of translation initiation. The best studied pathway of eIF2B inhibition is by phosphorylation of S51 of eIF2α (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Proud, 2001; Dever, 2002; see also

Section 5.3 above). eIF2(α-P)·GDP is a competitive inhibitor of eIF2B. It binds to eIF2B with higher affinity than the unphosphorylated eIF2·GDP, but forms a non-productive complex with the $\alpha\beta\delta$ subcomplex of eIF2B, which cannot be converted to the catalytically active complex with the active site of eIF2Bε. One study reported that the K_D of phosphorylated eIF2 is $\sim$50-fold lower than that of unphosphorylated eIF2, but others found only 2- to 3-fold difference. It appears that the α, β, δ subcomplex of eIF2B interacts weakly also with unphosphorylated eIF2α, but that the interaction does not prevent formation of the catalytically active complex and in fact stimulates catalysis by lowering the K_M for eIF2·GDP. Phosphorylation of eIF2B itself has been reported to regulate its activity. eIF2B activity is also under the control of a number of allosteric regulators. Two molecules of NADPH are bound per eIF2B hetero-pentamer and are important for structural stability. NADP, NAD, ATP, heparin, fusil alcohols have all been reported to inhibit eIF2B activity, whereas polyamines, sugar-phosphates and nucleotidyl-sugars stimulate it (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Proud, 2001; Dever, 2002).

Structural organization of eIF2B. eIF2Bγ is homologous to eIF2Bε, but lacks the C-terminal catalytic region (Fig. 9*a*). eIF2Bε-CTD is a W2 HEAT domain (Boesen *et al.* 2004) (Fig. 9*b*), homologous to eIF5-CTD, and also binds to the N-terminal tail of eIF2β. eIF2Bε and eIF2Bγ contain one copy each of a sugar-nucleotidyl transferase (SNT)-like and an acyl transferase (AT)-like domain (Koonin, 1995). A number of enzymes homologous to eIF2Bε and eIF2Bγ have been found to also have an SNT domain, followed by an AT domain. Some of these enzymes, like GlmU, are bifunctional, whereas others, like the ADP-glucose pyrophosphorylase (ADPGlc PPase) for example, have a truncated AT domain and only SNT activity. eIF2Bε and eIF2Bγ appear to have a truncated version of the AT domain, like the second group (Koonin, 1995). There is no indication that eIF2Bε and eIF2Bγ have even the SNT activity, although the γ subunit in yeast has the complete SNT signature motif. The structures of several SNTs and ATs are known, including the structure of the bifunctional GlmU in the presence of substrates (Kostrewa *et al.* 2001; Sulzenbacher *et al.* 2001) (Fig. 9*d*, *e*). The structure of the SNT domain resembles a dinucleotide-binding fold in its N-terminal (nucleotide-binding) half, whereas the C-terminal half of the domain binds the sugar and deviates from the typical dinucleotide-binding fold. The AT domains (also known as Ile-rich repeat domains) form a left-handed β-helix (LβH), where each turn is formed of three short β-strands. Of the enzymes with known oligomerization state, the AT domains of all proteins with AT activity, form a trimer, and the SNT domains (if present) are clustered on the same end of the trimer. SNT enzymes with truncated AT domains and no AT activity have been reported to be tetramers: either homotetramers (a_4) or heterotetramers (a_2b_2) (Takata *et al.* 1997), but no structures of such monofunctional SNT enzymes are available.

The homology of eIF2Bε and eIF2Bγ to SNT and AT enzymes is of interest not only because it allows homology modeling of the corresponding domains, but also because some of these SNT enzymes have been found to be allosterically regulated, and most of the allosteric regulators of eIF2B (Singh & Wahba, 1996) are similar to the allosteric regulators, substrates, and/or products of SNT enzymes (Leung *et al.* 1986). Although the binding sites of most regulators on eIF2B are not known, one can speculate that the sugars and at least some of the nucleotides act through the SNT and AT domains of eIF2Bγ and eIF2Bε. The homology to SNT and AT enzymes also suggests a model for the structure of the eIF2$\gamma\varepsilon$ subcomplex: heterodimerization can be mediated by the AT domains of the γ and ε subunits. The available data indicates that eIF2B is a heteropentamer of $\sim$250 kDa, indicating that eIF2γ and eIF2ε form a

heterodimer (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Dever, 2002). This would be the first example of an AT domain dimer, unlike the known proteins with an AT oligomerization domain, which are trimers or tetramers. However, the possibility that eIF2B is a heterodecamer with two copies of each of the five subunits cannot be formally excluded.

The α, β, and δ subunits are homologous to each other over at least a $\sim$200-residue region, but the homology may extend over the entire length of eIF2Bα ($\sim$300 residues). No structural information is available for any of these subunits, but within the last year, several structures of distant homologs of eIF2B$\alpha/\beta/\delta$ from yeast, archaea and bacteria were determined [(Bumann *et al.* 2004; Kakuta *et al.* 2004), as well as two more unpublished structures available from the PDB database, PDB codes 1T5O and 1T9K]. All four structures belong to the eIF2B-related (eIF2B_rel) subfamily and show high degree of homology to the enzyme 5-methylthioribose-1-phosphate isomerase from *B. subtilis*, which is the first member of the subfamily with known function (Ashida *et al.* 2003). Although some of the members of the eIF2B_rel subfamily from archaea and bacteria have been proposed to be eIF2B subunits involved in translation initiation, they are closely related to the above 5-methylthioribose-1-phosphate isomerase and most likely are also enzymes involved in the methionine salvage pathway (Bumann *et al.* 2004). It should also be noted, that archaea do not have a homolog of the catalytic eIF2Bε subunit or of the eIF2Bγ subunit, and bacteria do not even have eIF2.

The eIF2B_rel subfamily members provide some interesting insights into the structure and function of the eIF2Bα, β and δ subunits. They consist of two subdomains, an N-terminal mostly helical domain and a CTD, resembling a Rossman fold (Fig. 9*c*). The proteins are dimeric and the same surface of the CTDs is responsible for dimerization, although the angles between monomers vary among structures. It is intriguing that, like eIF2Bγ/ε, which are homologous to sugar-nucleotidyl transferase (SNT) enzymes (see above), the eIF2B$\alpha/\beta/\delta$ subunits also show homology to enzymes, whose substrates are sugar-phosphates – 5-methylthioribose-1-phosphate isomerase. Even though the eIF2B subunits have probably lost the enzymic activities of their homologs, they may have retained the ability to bind some of the ligands, and could serve as sensors for various aspects of the metabolic state of the cell.

Although eIF2Bε and even its C-terminus alone are sufficient for catalytic activity (Gomez *et al.* 2002), the β, γ and δ subunits of eIF2B are also essential in yeast. The eIF2Bγ subunit has been proposed to function in complex assembly. eIF2Bα is required for inhibition by eIF2α phosphorylation and is not essential. Therefore, the β and δ subunits are likely to have additional function(s) beyond eIF2(α-P) binding. It has been proposed, that eIF2·GDP could remain associated with the ribosome and eIF2B could be required for its release (reviewed in Hinnebusch, 2000; Dever, 2002). Alternatively, eIF2Bβ and δ could be involved in coordination of Met-tRNA$_i$ transfer from the aaRS directly to the eIF2·BeIF2·GTP complex (Erickson *et al.* 2001; see below). It is also possible that the eIF2Bβ and δ subunits are important for catalysis in the presence of eIF5, which competes with eIF2Bε for binding to the same site on eIF2 (see below).

The regulatory eIF2B$\alpha\beta\delta$ subcomplex binds tightly to a conserved surface of the S1 subdomain of eIF2α, when S51 on eIF2α is phosphorylated (reviewed in Hinnebusch, 2000; Dever, 2002). Two Sui mutations have been reported in yeast eIF2α. The mutations were recessive and found to impair the stability of eIF2α (Cigan *et al.* 1989). Both mutations affect residues in the NTD: one is buried in the structure and the other is a conserved proline, consistent with the proposed defects in folding and stability of the mutant eIF2α proteins (Ito *et al.* 2004). Both Sui mutants also had a Gcd$^-$ phenotype, indicating defects in eIF2·GTP·Met-tRNA$_i$

binding to the IC. Interestingly, at least one of the mutants was synthetic lethal with deletion of the nonessential eIF2Bα subunit (Cigan *et al.* 1989; Williams *et al.* 1989). These data support the existence of a functionally important interaction between eIF2α and the regulatory eIF2B$\alpha\beta\delta$ subcomplex in the absence of eIF2α phosphorylation, and a weakened interaction in the absence of eIF2Bα.

Catalytic mechanism of eIF2B and channeling. The mechanism of catalysis of eIF2 nucleotide exchange by eIF2B has been the subject of a long-standing controversy over the last 20 years (Manchester, 2001). Some groups have reported substituted enzyme mechanism, outlined in Fig. 10, whereas others suggest a ternary complex mechanism. The practical implications of these two alternatives are that a ternary complex mechanism involves an intermediate complex where both substrates eIF2·GDP and GTP are simultaneously bound to eIF2B and predicts direct GTP binding to eIF2B. Accordingly, eIF2B has been reported to bind GTP (Dholakia & Wahba, 1989; Nika *et al.* 2000), but it is not clear if GTP binding is directly involved in catalysis or has a modulatory role. It was recently shown, that eIF2B can catalyze nucleotide exchange by the substituted enzyme mechanism, i.e. via an eIF2B-eIF2 intermediate (Williams *et al.* 2001). It is not clear if this is the only mechanism or both mechanisms occur in parallel (Manchester, 2001). It is highly unlikely that the controversy surrounding the catalytic mechanism of eIF2B will be resolved any time soon, because as mentioned above, eIF2B is the target of a complex regulatory network, involving phosphorylation of both eIF2 and eIF2B (with some of the kinases being able to use GTP as a substrate), as well as by allosteric regulators, including nucleotides (reviewed in Hinnebusch, 2000). Any of these forms of regulation can affect the outcome of a kinetic study.

An important issue, overshadowed by the above controversy, is that eIF2B does not release the 'product' eIF2·GTP. Instead, the nucleotide exchange is coupled with Met-tRNA$_i$ binding (Salimans *et al.* 1984; Panniers *et al.* 1988; Gross *et al.* 1991). A comparison with the EFs and their GEFs shows that, similar to eEF1B, and unlike EF1B (EF-Ts), eIF2B is 'optimized' for channeling and not for stand-alone operation, because it dissociates very slowly from its 'product', eIF2·GTP. In fact, eIF2B appears to take the channeling one step further than eEF1B: the quaternary eIF2B·eIF2·GTP·Met-tRNA$_i$ complex was found to be stable during centrifugation in sucrose density gradients and eIF2B stimulated (and was released upon) binding of eIF2·GTP·Met-tRNA$_i$ to 43S complexes (Salimans *et al.* 1984). This transfer probably involves eIF5, which shares a common binding site on eIF2β with eIF2B. The explanation why eIF2B extends channeling one step further than eEF1B is that there is only one type of initiator tRNA to be delivered to the P site during initiation, but multiple aa-tRNAs need to have random access to the ribosomal A-site during elongation.

By analogy with eEF1A nucleotide exchange and aa-tRNA binding (see Section 3 above), one can hypothesize that *in vivo* Met-tRNA$_i$ binding occurs via direct transfer from the Met-tRNA$_i$ synthetase to eIF2 (Fig. 10), and a role for eIF2B in channeling has been proposed, based on genetic data (Erickson *et al.* 2001).

Channeling was discussed in Section 3.2.2 above. It has been noted long ago, that under physiologic conditions the affinity of eIF2 for GDP is ~100-fold higher than that for GTP, whereas the concentration of GTP is only ~10-fold higher. Therefore, the equilibrium between free eIF2·GTP and eIF2·GDP would be shifted toward the inactive eIF2·GDP. Upon Met-tRNA$_i$ binding, the equilibrium ratio between eIF2·GTP·Met-tRNA$_i$ and eIF2·GDP·Met-tRNA$_i$ is ~1:1 and the eIF2·GTP·Met-tRNA$_i$ complex is stable for minutes. Thus,

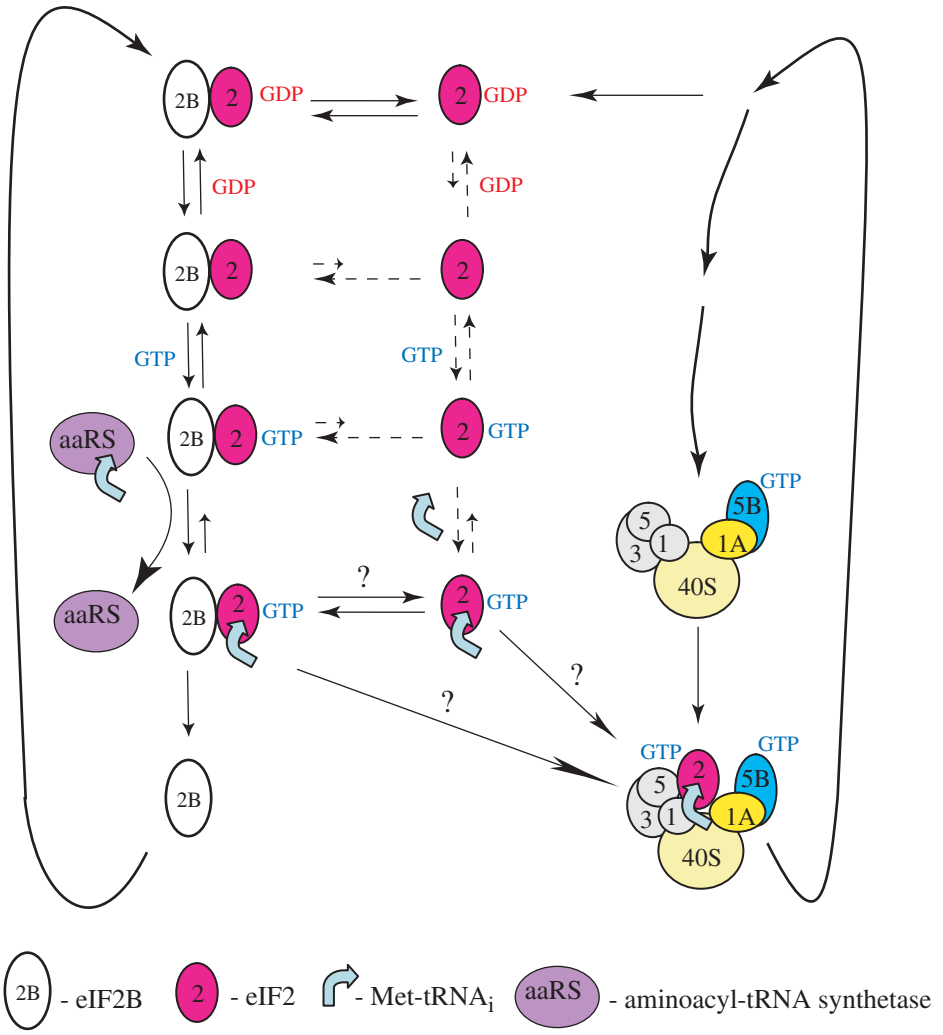

Fig. 10. eIF2B-mediated nucleotide exchange on eIF2. Schematic representation of the role of eIF2B in recycling of eIF2 during translation initiation, and channeling of eIF2·GTP and Met-tRNA$_i$. The initiation complexes to the right are as in Fig. 2 above. Short arrows indicate that the corresponding reaction is slow; dashed arrows indicate that the corresponding branches in the pathway are likely underpopulated and insignificant *in vivo*. Note for example, that in this scheme, eIF2·GTP is considered an unwanted side-product which would quickly convert back to eIF2·GDP. The question marks indicate that either the quaternary complex eIF2B·eIF2·GTP·Met-tRNA$_i$ or the ternary complex eIF2·GTP·Met-tRNA$_i$, or both could be involved in the transfer of eIF2·GTP·Met-tRNA$_i$ to the initiation complex (or eIF5), because both complexes are stable. The rate of dissociation of eIF2B from the eIF·2B·eIF2·GTP·Met-tRNA$_i$ complex has not been reported, but in at least one study, eIF2B stimulated eIF2·GTP·Met-tRNA$_i$ binding to the initiation complexes (Salimans *et al.* 1984), supporting a complete channeling cycle for eIF2B (see text).

channeling by eIF2B provides at least 10-fold stimulation of the process (Panniers *et al.* 1988). Completing the channeling cycle to include binding of eIF2·GTP·Met-tRNA$_i$ to the 43S complex (Salimans *et al.* 1984) results in further stimulation by unidirectional flow of intermediates, because the equilibrium is never reached. Many of the sharp discrepancies between the biochemical properties of mutations and deletions of eIF2 and eIF2B subunits *in vitro* and their *in vivo* phenotypes are likely due to their effects on channeling. For example, although

deletion of the essential eIF2B in yeast can be rescued by overexpression of eIF2 and $tRNA_i$, the resulting strain grows very slowly (Erickson *et al.* 2001). Therefore, many mutants may need further experimental testing and reevaluation. A detailed analysis is beyond the scope of this review, but we will illustrate the problem with two examples. (1) A mutation that affects dissociation of GDP from eIF2 may have different effects on the overall exchange process, depending on its effect on GTP binding and release. Therefore, studying only exchange of labeled GDP with unlabeled GDP cannot provide reliable information about the properties of a mutant. (2) Exchange of GDP with GTP, however, is very difficult, because eIF2·GTP is not a product but an intermediate and its dissociation from eIF2·BeIF2·GTP is very slow. In fact, a mutation that causes faster release of eIF2·GTP from eIF2·BeIF2·GTP will appear as 'stimulating' the nucleotide exchange, but may in fact be deleterious, because it will abolish channeling and may make the overall process less efficient. To our knowledge, exchange rates in the presence of Met-$tRNA_i$, have seldom been measured (Gross *et al.* 1991); a possible role for the synthetase has not been studied and presents serious technical challenges.

The dependence of the eIF2-binding affinities of eIF5 and eIF2B as a function of GDP, GTP, and Met-$tRNA_i$ binding has not been extensively compared. One could argue that eIF2B may have higher affinity for eIF2·GDP (and eIF2·GTP?), and eIF5 – higher affinity for eIF2·GTP·Met-$tRNA_i$. There are reasons to believe, that the role of the interaction between the N-terminal tail of eIF2β and the W2 HEAT domains of eIF2Bε and eIF5 may be more than simple anchoring: (1) the GEF catalytic site of eIF2Bε is at the N-terminus of the W2 HEAT domain, indicating that the W2 HEAT domain itself comes in the vicinity of the nucleotide. (2) The N-terminus of eIF2β has been cross-linked to tRNA in the eIF2·GTP·Met-$tRNA_i$ complex. Therefore, it is possible, that the interactions of the eIF2β N-terminus with the W2 HEAT domains of eIF5 and eIF2Bε are differentially affected by tRNA binding. Other mechanisms can also be proposed, for example, additional interactions of eIF2 with eIF2B and eIF5, affected differently by GDP-, GTP-, and Met-$tRNA_i$-binding. Other IFs are also likely to affect the shuttling of eIF2 between eIF5 and eIF2B. For example, it was found recently that in yeast, eIF2 and eIF3 bind cooperatively to eIF5-CTD (Singh *et al.* 2004b), whereas no interaction between eIF2B and eIF3 has yet been reported. The N-terminal tail of eIF2β contains three lysine-rich boxes (K-boxes) and mutating the K-boxes individually or in various combinations had different effects on eIF5 and eIF2B binding (Asano *et al.* 1999). This complicates the system even further, because it is not clear which K-box(es) of eIF2β are bound by eIF5 and eIF2B and makes it possible for eIF5 and eIF2B to contact eIF2β simultaneously, at least as a transient intermediate.

In summary, one can envision the following hypothetical scenario for the roles of the individual eIF2 subunits and of eIF2B and eIF5: (1) eIF2γ carries the principal Met-$tRNA_i$-, nucleotide- and ribosome-binding functions, similar to its homolog eEF1A. (2) Like eEF1A, eIF2γ needs a GAP and a GEF. eIF2α-CTD has served as the GEF in the primordial eIF2 and probably still does in archaea. (3) In eukaryotes, eIF2α appears to have lost its GEF activity to eIF2B, but is still involved in the process through interactions with eIF2B. (4) Unlike eEF1A, eIF2γ cannot use the GAC on the large ribosomal subunit (which is simply not there yet). The evidence points toward eIF2β serving as a latent GAP: with the NTD carrying the GAP activity, and the ZBD repressing GTP hydrolysis. The repression is then relieved through structural rearrangements upon start site recognition. (5) In eukaryotes, eIF5 has emerged through gene duplication from eIF2β and has taken on (most of?) the GAP function, leaving (most of?) the

repressing function to the ZBD of eIF2β. Obviously, multiple alternative explanations exist, especially for the last two points.

5.5 eIF3

eIF3 is a multisubunit factor, involved in dissociation/antiassociation of ribosomal subunits, IC formation and recruitment to mRNA, and scanning. eIF3 has been reported to interact with the 40S subunit, mRNA and several factors, although some of these interactions appear to vary among species, as does its polypeptide composition (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Browning *et al.* 2001).

eIF3 is composed of 5–6 subunits in yeast (*S. cerevisiae*) but as many as 12 in mammals, designated in order of decreasing size as eIF3a through eIF3l (Browning *et al.* 2001). eIF3 binds to the small ribosomal subunit, mRNA, eIF1, eIF2 and eIF4B in both yeast and mammals and to eIF4G in mammals, but not (or at least not reported) in yeast (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000). In yeast, eIF3 has been found to interact with eIF1A and eIF5 (He *et al.* 2003; Olsen *et al.* 2003) and it is not yet known if these interactions also exist in human.

Little is known about the structure and function of the individual eIF3 subunits. Sequence homology searches detect a number of conserved domains (Asano *et al.* 1997; Naranda *et al.* 1997; Hofmann & Bucher, 1998; Aravind & Ponting, 1998; Aravind & Koonin, 2000; Hershey & Merrick, 2000; Chang & Schwechheimer, 2004; Kim *et al.* 2004).

The human eIF3a, c and e subunits share a putative PINT/PCI domain (with distant homology) among each other and with regulatory subunits of the signalosome. Of them, eIF3a (TIF32) and eIF3c (NIP1) are also present in yeast, where they are shorter than in mammals by $\sim$400 and $\sim$100 residues respectively, affecting mainly segments of low sequence complexity.

More can be inferred about the structure of eIF3b (PRT1) and eIF3i (TIF34). Almost the entire eIF3i is composed of a total of 7 WD40 repeats and it is known to interact with the TGF-β receptor. The WD40 proteins with known structures form a propeller ring, composed of 7 repeats. The diameter of the propeller is $\sim$50 Å and the height is $\sim$30 Å. The eIF3b (PRT1) subunit has a RNA recognition motif (RRM) domain, between residues $\sim$200 and $\sim$300, followed by a 500–600 residue WD40 repeat segment, which most likely has 14 WD40 repeats, but the exact number is difficult to predict with certainty. Therefore, eIF3b appears to have two tandem WD40 propellers, although the possibility for one or one-and-a-half propellers cannot be excluded. The last alternative, however, may require that another protein dimerizes with eIF3b, to form a complete second propeller. WD40 proteins with as many as 16 WD40 repeats have been reported and it has been proposed that WD40 propeller domains can be composed of between 4 and 8 repeats, but the only known structures of propellers with more than 7, or fewer than 7 'blades' belong to non-WD40 proteins. The WD40 region of eIF3b (PRT1) shows homology to eIF2A (a protein, which is not essential in yeast and whose function in translation initiation is unknown). The $\sim$200-residue N-terminus of eIF3b is Pro-rich and shorter in yeast. RRMs are common among ribosomal proteins and IFs and often (although not always) bind RNA. WD40 domains are often involved in protein–protein interactions and found in proteins from various pathways, such as cell-signaling, RNA processing and transcription (Li & Roberts, 2001). eIF3b and eIF2A appear more closely related to WD40s of transcription factors, than to eIF3i.

eIF3d is not found in *S. cerevisiae* and is not homologous to any other known proteins.

eIF3f and eIF3h (not present in *S. cerevisiae*) are homologous to proteasome subunits over their entire lengths, and (to a lesser degree) to each other and to signalosome subunits over the N-terminal two-thirds of their sequences. A putative Mov34/JAB_MPN domain of ~150 residues is found at their N-termini. A mutant of plant eIF3h was recently found to have a defect in reinitiation (Kim *et al.* 2004). eIF3k is also homologous to a proteasome regulatory subunit. Its structure was recently determined making it the first eIF3 subunit with known structure (Wei *et al.* 2004). eIF3k consists of a HEAT-like domain and a winged helix-like (WH) domain. eIF3k was found to interact directly with eIF3c, eIF3g and eIF3j *in vitro* (Mayeur *et al.* 2003). A putative PINT/PCI domain has recently been predicted in eIF3k (reviewed in Chang & Schwechheimer, 2004). If correct, this finding would provide information about the structure of the PINT/PCI domains.

The only recognizable features of eIF3g (TIF35) are a RRM domain at the C-terminus and a putative Zn-binding motif.

eIF3j was predicted to contain a putative PINT/PCI domain (Valasek *et al.* 2001a; Chang & Schwechheimer, 2004). Its yeast homolog, HCR1 was found to be associated with eIF3 in substoichiometric amounts (Valasek *et al.* 2001b) and to have a role in ribosome biogenesis in addition to translation initiation (Valasek *et al.* 2001a). eIF3j stabilizes binding of eIF3 to the 40S ribosomal subunit (Fraser *et al.* 2004) and is released upon binding of the 43S IC to mRNA (Unbehaun *et al.* 2004).

eIF3l was reported to have a PINT/PCI, tetratricopeptide repeat (TPR), and a Pumilio FBF RNA binding domain repeat (Puf R) motifs; and to bind to eIF3e (Morris-Desbois *et al.* 2001).

In summary, eIF3 from both yeast and human contains at least two RRM and three WD40 repeat domains, known to be involved in protein–protein interactions. The presence of three (and up to six in human) putative PINT/PCI domains with distant homology to regulatory subunits of the signalosome and the proteasome, suggests possible common origins with components of these multiprotein complexes. One could hypothesize possible involvement of the PCI domains in eIF3 in protein–protein interactions and/or regulation. Among the other conserved domains found in eIF3 subunits are two putative Mov34/JAB_MPN domains and at least one HEAT domain, also involved in protein–protein interactions.

Interestingly, three of the human eIF3 subunits, absent in *S. cerevisiae*, are close homologs of proteasomal regulatory subunits, indicating that they were 'borrowed' from the proteasome more recently or are even 'shared'. Furthermore, the overall organization of the signalosome and the proteasome regulatory complex was proposed to be similar and possibly related to that of eIF3 (Chang & Schwechheimer, 2004).

The interactions among yeast eIF3 subunits and with other eIFs and the ribosome have been extensively studied in recent years (see Valasek *et al.* 2002; Nielsen *et al.* 2004, 2003, and references therein; Table 2). The interactions among the yeast eIF3 subunits appear conserved in human eIF3 (Fraser *et al.* 2004). The roles of the predicted conserved domains in eIF3 subunits or their exact boundaries have not been studied extensively, but a comparison of the yeast *in vivo* and *in vitro* data with the conserved domains in eIF3 indicates the following possible correlations:

(1) The PINT/PCI domain of eIF3a (TIF32) overlaps with the eIF3c (NIP1) binding site. The PINT/PCI domain of eIF3c (NIP1) has been proposed to interact with the 40S subunit (without it being clear whether the interaction is with ribosomal proteins and/or rRNA).

(2) The RRM of eIF3b (PRT1) binds to eIF3a (TIF32) and eIF3j (HCR1), but its role in RNA binding has not been studied. The function of the other RRM, that of eIF3g (TIF35) is also unknown. The C-terminal tail of eIF3b at the end of the WD40 region interacts with eIF3i (TIF34) and eIF3g (TIF35). The interaction of the WD40 eIF3i with a region immediately adjacent to the WD40s of eIF3b suggests that it is possible that the three WD40 domains could form a flat platform, which could have either a triangular shape with a side of ~100 Å or a rod-like shape with length of ~150 Å and width of ~50 Å. EM studies of mammalian eIF3 bound to the 40S subunit have proposed a flat triangular prism with ~150 Å sides and ~70 Å height (Bommer *et al.* 1991b) and an elongated, bilobed structure with a length of ~170 Å (Srivastava *et al.* 1992), which is consistent with either orientation. It should be noted, however, that in addition to the uncertainty of these EM models, mammalian eIF3 is much larger than yeast eIF3.

EM reconstructions indicated that mammalian eIF3 binds to the solvent-exposed surface of the platform and the 'back' of the 40S ribosomal subunit, and does not directly occlude the 60S–40S subunit interface (Srivastava *et al.* 1992). At the same time, eIF3 has been reported to interact with eIFs 1, 1A, 2, and 5. All of these factors bind to the interface surface of the 40S subunit. In yeast, eIF3a (TIF32) and eIF3c (NIP1) have been found to interact with ribosomal protein Rps0A, which is on the 'back' of the 40S subunit with respect to the 60S interface. eIF3a was found to interact with a segment of domain I of the 18S rRNA, encompassing helices 16–18, near the A-site (Valasek *et al.* 2003). Parts of the latter segment are accessible from either the interface, or the solvent-exposed surface of the 40S subunit. The C-terminal region of yeast eIF3a (TIF32), to which the binding was mapped, is highly charged and likely unstructured, and its interaction with rRNA does not impose significant restraints on the location of eIF3.

As many of the interactions between eIF3 and eIFs 1, 1A, 2 and 5 are likely to be retained on the 40S subunit (or even only occur when the factors are bound to the 40S subunit), individual subunits of eIF3 need to protrude close to the interface surface, at least in the presence of other factors. The shortest paths from the eIF3 binding site on the 40S subunit to the targets in the A- and P-sites would be around the neck, between the head and the body: either from the platform, or from the A-site. eIF3 binding to the solvent-exposed surface of the 40S subunit and/or access to the interface surface may induce movement of the head with respect to the rest of the 40S subunit.

As already explained, eIF3 contacts proteins located in both the A-site and the P-site. Of these, the interaction with eIF1 is the easiest to interpret. The location of eIF1 on the interface surface of the platform was mapped and it is known that the unstructured N-terminal tail of eIF1 is not important for eIF3 binding (Lomakin *et al.* 2003). Taking into account the rRNA and tRNA contacts of eIF1, two surfaces, exposed in the model for the eIF1-40S complex (Lomakin *et al.* 2003), are most likely to mediate the interaction with eIF3: the surface oriented away from the P-site mRNA and the contiguous surface facing the A-site. Interestingly, the Sui mutations in yeast eIF1 map to the latter, indicating that it is the most likely eIF3-binding site. At odds with this prediction, one of the yeast eIF1 Sui mutants was reported to be defective in eIF4G, instead of eIF3 binding (He *et al.* 2003). On the eIF3 side of the interaction, eIF3c has been found to bind eIF1 in both yeast and eukaryotes. Yeast eIF3a (TIF32) was also reported to bind eIF1. Neither of the above interactions was easy to reconcile with a model for the putative orientation of eIF3 subunits on the yeast 40S subunit and it was suggested that on the 40S subunit, eIF3 no longer binds eIF1 (Valasek *et al.* 2003). The cooperativity in binding and subunit antiassociation between eIF1 and eIF3, as well as other recent data, however favor

interaction between the two factors on the small subunit (Pestova & Kolupaeva, 2002; Unbehaun *et al.* 2004).

Although the locations of eIF1A and eIF2 are fairly well established, their interactions with eIF3 and with each other are harder to localize in space. The eIF2- and eIF3-binding region of eIF1A is an unstructured ~20-residue N-terminal tail. As the tail is highly positively charged, it is likely to interact with RNA on the ribosome, rather than remain free, but if extended, its length could allow it to reach as far away as ~50 Å. The eIF1A-interacting subunit of eIF2 is not known, but the interaction of eIF2 with eIF3 was mapped to the eIF2β subunit (Valasek *et al.* 2002) and that with eIF5-CTD was further mapped to the ~100-residue-long N-terminus of eIF2β (Asano *et al.* 1999). It is not known if the N-terminus of eIF2β is folded in free eIF2β or in eIF2, but two sets of evidence indicate that it may be localized near the rest of eIF2: (1) a fragment comprising residues 43–61 of eIF2β has been cross-linked to Met-tRNA$_i$ in the ternary eIF2·GTP·Met-tRNA$_i$ complex (Gaspar *et al.* 1994); (2) the catalytic site of eIF2Bε is at the N-terminal end of its W2 HEAT domain, which binds the N-terminus of eIF2β (Gomez *et al.* 2002; Boesen *et al.* 2004).

eIF1A is thought to bind in the A-site of the 40S subunit (Li & Hoffman, 2001; Roll-Mecak *et al.* 2001). It is not known which eIF3 subunit contacts eIF1A, but eIF3c (NIP1), which binds to eIF1, as well as eIF3a (TIF32), which also binds to eIF1, eIF2β, and to helices 16–18 of 18S rRNA, are candidates. The distance from the site where the N-terminal tail of eIF1A 'exits' the folded domain to the nearest surface of eIF1 is ~40 Å, and to the farthest surface of the folded eIF1 domain – ~70 Å. Therefore, the N-terminus of eIF1A could theoretically reach an eIF3 subunit bound to almost any surface of eIF1.

eIF2γ is expected to bind tRNA in the same way as its homologs EF1A and eEF1A, contacting mainly the acceptor end of the tRNA (Schmitt *et al.* 2002; Roll-Mecak *et al.* 2004; and see above). Therefore, for an eIF1A-eIF2 interaction, the N-terminus of eIF1A and/or a protrusion from an eIF2 subunit may have to cross over the length of the ASL of the Met-tRNA$_i$, which can be estimated to ~40 Å when the Met-tRNA$_i$ is oriented as the P-site tRNA in the 70S ribosome, but could be longer if the tRNA is rotated toward the E-site. In either case, the distance from eIF2 to eIF1 would be much shorter, even with possible direct contact between them. The length of the N-terminus of eIF1A alone could account for simultaneous interactions on the ribosome of eIF1 with eIF3, of eIF1A with eIFs 2 and 3, and between eIFs 2 and 3, although protruding flexible segments from eIF2β or other eIF2 subunits may also be involved. As the C-terminal W2 HEAT domain of eIF5 has been reported to bind simultaneously eIF2β and eIF3c [and weakly to eIF1 (He *et al.* 2003)], it is likely also positioned in the same general area, which would allow the N-terminal two domains of eIF5 to be near the GTP on eIF2γ, in order to promote GTP hydrolysis.

In summary, the space around the A- P- and E-sites on the 40S subunit is the site of a network of interactions of proteins and RNAs: eIF1A and eIF5B are in the A-site; eIF2·GTP·Met-tRNA$_i$, eIF1, and eIF5 occupy the P-site and maybe parts of the E-site; whereas eIF3 subunits 'reach' through the E- and/or the A-site.

5.6 Factors involved in cap binding and scanning

5.6.1 eIF4E

eIF4E binds to the 5′-cap of mRNA and the scaffold protein eIF4G, thus recruiting the rest of the cap-binding complex and the 43S IC to the 5′-end. eIF4E is one of the major targets

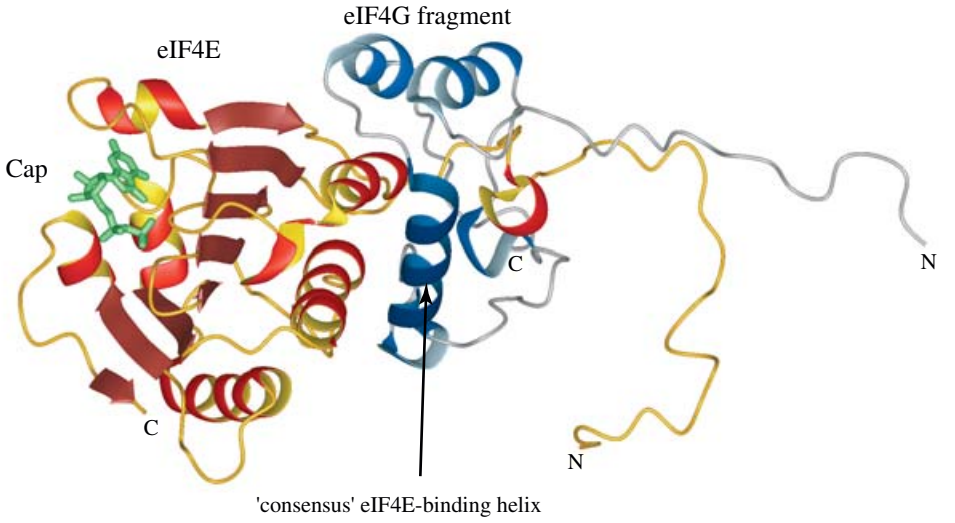

Fig. 11. eIF4E-eIF4G interaction. Structure of the complex of yeast eIF4E with a cap analog and a ~100-residue fragment of eIF4G, PDB code 1RF8. eIF4E is in red and orange, eIF4G is in blue, and the cap analog is in green. Note that eIF4G is wrapped around the N-terminus of eIF4E. The 'consensus' eIF4E-binding helix of eIF4G is in dark blue. The N- and C-termini of the proteins are labeled.

of regulation of translation initiation. eIF4E-binding proteins (4E-BP), in their unphosphorylated/hypophosphorylated form, compete with eIF4G for binding to eIF4E and inhibit initiation. eIF4E is limiting in many tissues and is overexpressed in many cancer cell lines (reviewed in Dever, 2002; Preiss & Hentze, 2003).

The structures of eIF4E, bound to a cap analog, from yeast and mouse have been determined, as well as the complexes of eIF4E with small consensus peptides from 4E-BP and eIF4G (Marcotrigiano *et al.* 1997a, b; Matsuo *et al.* 1997). The structure of eIF4E resembles a baseball glove, with the cap binding to the concave surface. Specific recognition of the 7-methyl-G of the cap is achieved by stacking between conserved tryptophan rings and direct contacts with the methyl group. N7-methylated guanosine nucleotides bind eIF4E with a 4- to 6-fold higher affinity than the non-methylated analogs. This can be attributed to an enhanced π–π-stacking interaction of the methylated base with the tryptophans. The methylation leads to electron deficient π-orbitals that interact favorably with the electron-rich π-orbitals of the tryptophan indols (Ishida *et al.* 1988; Carberry *et al.* 1990; Ueda *et al.* 1991). The consensus peptides bound to the dorsal surface of eIF4E. The consensus peptides, while sufficient for binding, did not account for all features of the eIF4E·4E-BP and eIF4E·eIF4G interactions, such as the stabilization of eIF4E – cap binding by eIF4G (Haghighat & Sonenberg, 1997) or the role of 4E-BP phosphorylation for down-regulation of eIF4E binding (Haghighat *et al.* 1995). The structure of a complex between eIF4E and a ~100 residue fragment of eIF4G was recently determined in our laboratory, providing new insights into the eIF4E·4G and eIF4E·4E-BP interactions (Gross *et al.* 2003) (Fig. 11). The eIF4G fragment was folded upon binding and wrapped around the N-terminus of eIF4E; consistent with mutational and biochemical data (Hershey *et al.* 1999). The binding of the longer fragment was orders of magnitude more stable than that of the consensus peptide and caused stabilization of cap-binding by eIF4E. The tighter cap binding by eIF4E was not accompanied by significant structural changes in the cap-binding site, indicating that eIF4G binding indirectly stabilizes the cap-bound state of eIF4E. The observed larger

binding interface allowed to propose an explanation for the role of the ordered phosphorylation of 4E-BPs in their dissociation from eIF4E (Gross *et al.* 2003). The large-scale conformational changes in eIF4G upon eIF4E binding are in line with reported cooperativity of eIF4G binding to the 4E-cap complex, PABP, and the 43S IC (reviewed in Hershey & Merrick, 2000; Jackson, 2000).

5.6.2 eIF4G

eIF4G (Fig. 12) is a large scaffolding protein, responsible for the assembly of the cap-binding complex and recruitment of the 43S IC to the 5′-cap of mRNA. It has binding sites for eIF4E, eIF4A (one in yeast, two in human), PABP, mRNA and other factors. Human eIF4G also binds to eIF3 whereas yeast eIF4G has not been found to bind eIF3, but was reported to bind eIF1 and eIF5. The N-terminal ~600 residue region of eIF4G contains the binding sites for PABP and eIF4E, as well as other factors. This region is followed by a HEAT domain (NIC/MIF4G), involved in binding to eIF4A, RNA and eIF3. Human eIF4G has at least two additional domains: a second eIF4A-binding domain (MI/MA3) and a C-terminal W2 HEAT domain, which is homologous to the W2 HEAT-domains of eIF2Bε and eIF5, but binds to the MNK kinases and not to eIF2β. The N-terminal region of eIF4G may be unstructured in the free protein, but seems to at least functionally interact with the region around the first HEAT domain, because binding of eIF4E or removal of the N-terminus of eIF4G has an effect on interactions involving the middle segment of eIF4G (reviewed in Hershey & Merrick, 2000; Jackson, 2000; Prevot *et al.* 2003; Sonenberg & Dever, 2003).

The structure of the first HEAT domain of human eIF4GII was determined (Marcotrigiano *et al.* 2001) (Fig. 12*b*). The MI/MA3 domain was found to be distantly related to the first HEAT domain (Ponting, 2000) and is probably also a HEAT domain. The structure of the W2 HEAT domain of eIF2Bε, homologous to the C-terminal W2 domain of human eIF4G, was also recently determined (Boesen *et al.* 2004) (Fig. 12*c*). Thus, whereas yeast eIF4G has only one HEAT domain, eIF4G from higher eukaryotes appears to contain three consecutive HEAT domains. This suggests that the structural homology between eIF4G and the nuclear cap-binding complex protein 80 (CBP80) (Mazza *et al.* 2001) extends far beyond the first HEAT domain and over the entire length of CBP80 (Fig. 12). In support of this interpretation, the first HEAT domain of human eIF4G (Marcotrigiano *et al.* 2001) is most structurally homologous to the first HEAT domain of CBP80 (Mazza *et al.* 2001), and the W2 HEAT domain of eIF2Bε (with sequence homology to the C-terminal W2 HEAT domain of human eIF4G) is most structurally homologous to the third domain of CBP80 (Boesen *et al.* 2004). It should be noted, that the region of eIF4G between the first HEAT domain and the third, W2 HEAT domain is ~400 residues, whereas the corresponding segment of CBP80 (the second HEAT domain and the long interdomain loops) is only ~250 residues. Therefore, the connecting segments between the HEAT domains could be quite different between eIF4G and CBP80.

The above findings have a number of interesting implications. (1) The observed high degree of similarity between the W2 HEAT domain of eIF2Bε and domain 3 of CBP80 (Dali Z score of 12 and RMSD of only 2 Å) indicates that the latter is structurally related to the W2 HEAT domain family, even though no sequence homology can be detected. (2) The structure of the predicted second HEAT domain of eIF4G (the MI domain) can be modeled after that of domain 2 from CBP80, and the structure of the W2 HEAT domain of eIF4G – after that of the eIF2Bε W2 HEAT domain. (3) The fairly compact structure of CBP80 can suggest a possible

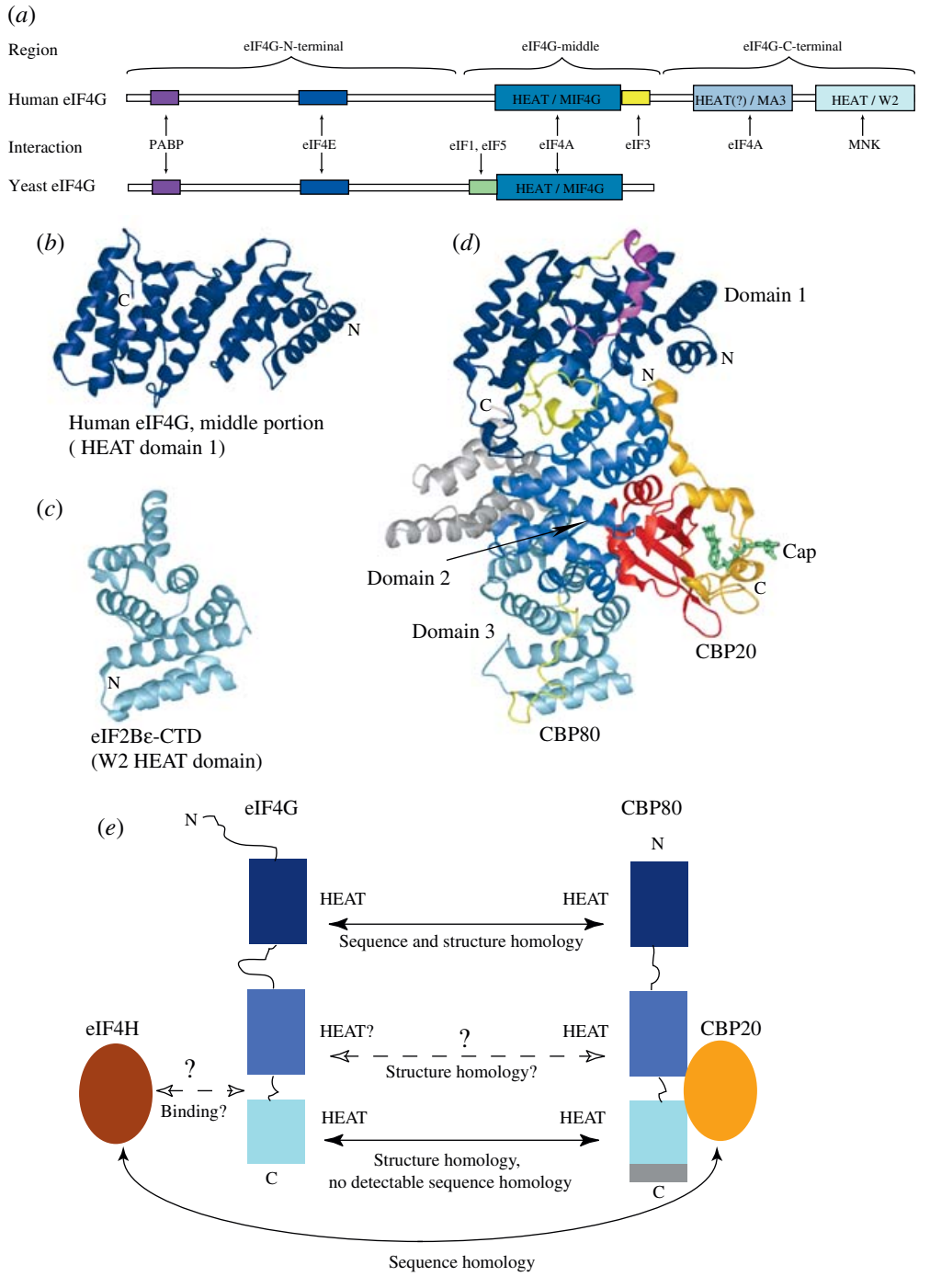

Fig. 12. eIF4G and homologous HEAT domain proteins. (*a*) Domain organization and interactions of human and yeast eIF4G. (*b*) Structure of the middle portion (HEAT domain 1) of human eIF4GII, PDB code 1HU3, is colored dark blue. (*c*) Structure of the C-terminal W2 HEAT domain of eIF2Bε, PDB code 1PAQ, homologous to the CTD of eIF4G colored light blue. (*d*) Structure of the human nuclear cap-binding complex, PDB code 1H2T. The cap analog is in green, the RRM domain of CBP20 is in red, the segments of CBP20 that are unstructured in the absence of a cap are in orange. HEAT domain 1 of

orientation of the HEAT domains in the C-terminal two-thirds of human eIF4G. This could allow comparisons between experimental data available for eIF4G or CBP80 (A. Marintchev & G. Wagner, unpublished observations).

An ~100-residue segment in yeast eIF4G folds into a five-helix 'bracelet' upon binding to eIF4E and the mode of interaction is likely conserved with higher eukaryotes (Gross *et al.* 2003). Similarly, the PABP-binding region of human eIF4G is folded upon binding to the rotavirus NSP3 protein, and the same is probably true upon PABP binding (Groft & Burley, 2002).

The HEAT domain of yeast eIF4G and the first two HEAT domains of human eIF4G bind eIF4A (Lamphear *et al.* 1995; Imataka & Sonenberg, 1997; Dominguez *et al.* 1999; Neff & Sachs, 1999). The information about the boundaries of the eIF4A-binding regions of eIF4G is somewhat contradictory (Korneeva *et al.* 2000, 2001; Lomakin *et al.* 2000; Morino *et al.* 2000). The first HEAT domain is sufficient for binding (Lomakin *et al.* 2000; Marcotrigiano *et al.* 2001). However, adjacent regions appear to be involved in, or at least modulate eIF4A binding, because, deletions of N-terminal segments have been reported to abolish binding (Morino *et al.* 2000).

The binding site for eIF3 was mapped to the region immediately C-terminal to the first HEAT domain, and eIF3 and eIF4A were found to bind cooperatively (Korneeva *et al.* 2000). Similar to the contradictory eIF4A binding data (see above), deletions and mutations outside of the minimal eIF3-binding site of eIF4G were found to affect eIF3 binding (Morino *et al.* 2000). It should be noted, that although no binding between yeast eIF4G and eIF3 has been found, the C-terminus of yeast eIF4G extends beyond the HEAT domain (homologous to the first HEAT domain of human eIF4G) and parts of it are conserved among lower eukaryotes. The C-terminus of yeast eIF4G could correspond to the eIF3-binding segment of human eIF4G. There is no detectable sequence homology between the two segments; but there is no obvious sequence homology between the PABP-binding sites of human and yeast eIF4G, either. Therefore, it is possible, that yeast eIF4G may interact weakly with eIF3 (or that the interaction may require the binding sites to be exposed upon binding to other factors). Alternatively, the C-terminus of yeast eIF4G could have a different function, not related to eIF3. Yeast eIF4G was reported to bind to eIF1 and eIF5, and these interactions were proposed to act as alternative to eIF4G-eIF3 binding in higher eukaryotes (He *et al.* 2003). As with eIF4A binding, the HEAT domain and flanking segments of eIF4G appeared involved in the interaction. However, the HEAT domain alone did not bind either factor, and the main binding site appeared to lie

---

CBP80 (homologous to the first HEAT domain of eIF4G) is in dark blue, HEAT domain 2 is in blue, and the first four helical hairpins of HEAT domain 3 (structurally homologous to the C-terminal W2 HEAT domains of eIF2Bε and eIF4G) are in light blue. The last two hairpins of domain 3, absent in the C-terminal W2 HEAT domains of eIF2Bε and eIF4G, are in grey. The first 17 residues of the linker between domains 1 and 2 are colored magenta; the rest of the interdomain segments are in yellow. Note that the interdomain linkers interact with, and wrap around the HEAT domains. The C- and N-termini of the proteins are indicated. The structures in (*b*) and (*c*) were aligned to their corresponding domains in CBP80 in (*d*), using DALI (Holm & Sander, 1993) and translated to the left, while retaining the same orientation. The domain of eIF4G between the first HEAT domain and the C-terminal W2 HEAT domain was predicted to also be a HEAT domain (Ponting, 2000), and may correspond to domain 2 of CBP80. Seventeen residues, present in the protease-resistant fragment used to determine the structure in (*a*), were not resolved in the crystals (Marcotrigiano *et al.* 2001). This segment could interact with the HEAT domain, similar to the corresponding region of CBP80, colored in magenta in (*d*). (*e*) A cartoon representation of the hypothesis for extended structural homology between CBP80 and the C-terminal two-thirds of eIF4G (A. Marintchev & G. Wagner, unpublished observations). The coloring of the HEAT domains is as in panels (*a*), (*b*) and (*c*). Predictions not confirmed experimentally are shown with dashed arrows and labeled with '?'.

N-terminal form the HEAT domain (He *et al.* 2003). As deletion of the C-terminal region after the HEAT domain diminished binding, at least to eIF1, the role of this segment in yeast could be in eIF1 binding. It is not known if human eIF4G also binds eIF1 and eIF5.

A number of the interactions involving eIF4G are cooperative: binding of eIF4E and PABP; of eIF4E and RNA; of eIF4E and 43S; of eIF4A and eIF3, etc. In most cases, there is no physical overlap between the individual binding sites on eIF4G and thus, no obvious reason for the observed cooperativity. This indicates, that eIF4G is not just a string of several independent binding sites, but rather a dynamic scaffold with extensive cross-talk among different parts of the molecule (Prevot *et al.* 2003). If, as proposed above, the domain organization of the C-terminal two-thirds of eIF4G is indeed similar to that of CBP80, the three HEAT domains would be expected to form a compact structure, with the N- and C-termini not far from each other. Furthermore, intervening segments may interact with the HEAT domains, as seen in the CBP80 crystal structure (Mazza *et al.* 2001). Regions N-terminal to the first HEAT domain, including the binding sites for eIF4E and PABP, may also be in proximity to, or even interact weakly with the HEAT domains, consistent with the ability of the MNK kinase bound to the C-terminal W2 HEAT domain to phosphorylate eIF4E, bound $\sim$1000 residues away along the eIF4G sequence.

Several translation regulators found in higher eukaryotes contain regions homologous to the C-terminal two-thirds of eIF4G (the translation repressor p97/NAT1/DAP-5), only to the first HEAT domain (the translation stimulator PAIP1), or to the second HEAT domain (the translation repressor Pdcd4, containing two copies of the second HEAT domain). These proteins often retain the interactions, in which the corresponding domains of eIF4G are involved. In line with the roles of the individual domains of eIF4G, some of the above translation repressors were found to inhibit cap-dependent translation but support IRES-dependent translation. Remarkably, several viral proteases and Caspase 3 can cleave eIF4G, producing similar fragments (reviewed in Hershey & Merrick, 2000; Jackson, 2000; Prevot *et al.* 2003; Sonenberg & Dever, 2003).

5.6.3 eIF4A

eIF4A is an ATP-dependent RNA helicase, belonging to the DEAD box family of Superfamily II helicases. eIF4A is responsible for unwinding mRNA secondary structures near the 5′-cap, thus helping binding of the 40S ribosomal subunit to the cap-proximal region of mRNA. Once the 43S IC is recruited, eIF4A promotes scanning to the start codon and is indispensible if the 5′-UTR contains even minor secondary structure elements. Alone, eIF4A is not a processive helicase and its activity is stimulated by eIF4G, eIF4B and eIF4H. eIF4A homologs are found in all kingdoms, but the roles of the archaeal and bacterial eIF4A homologs are less defined. For example, bacterial W2/IF4A (not related to the W2 domain found in some eIFs) may be involved in both initiation and elongation (reviewed in Hershey & Merrick, 2000; Hinnebusch, 2000; Ganoza *et al.* 2002; Prevot *et al.* 2003).

The structures of eIF4A from both yeast (Caruthers *et al.* 2000) and archaea (Story *et al.* 2001) have been determined (Fig. 13). Both structures consist of two RecA-like domains, forming the conserved core of superfamily I and II helicases. Unlike other helicases, however, eIF4A does not have any additional domains, underscoring the dependence of its helicase activity on accessory proteins. In the structure of yeast eIF4A, the two domains do not interact with each other (Fig. 13*b*). While this indicates the ability of eIF4A to exist in an 'open' conformation, the

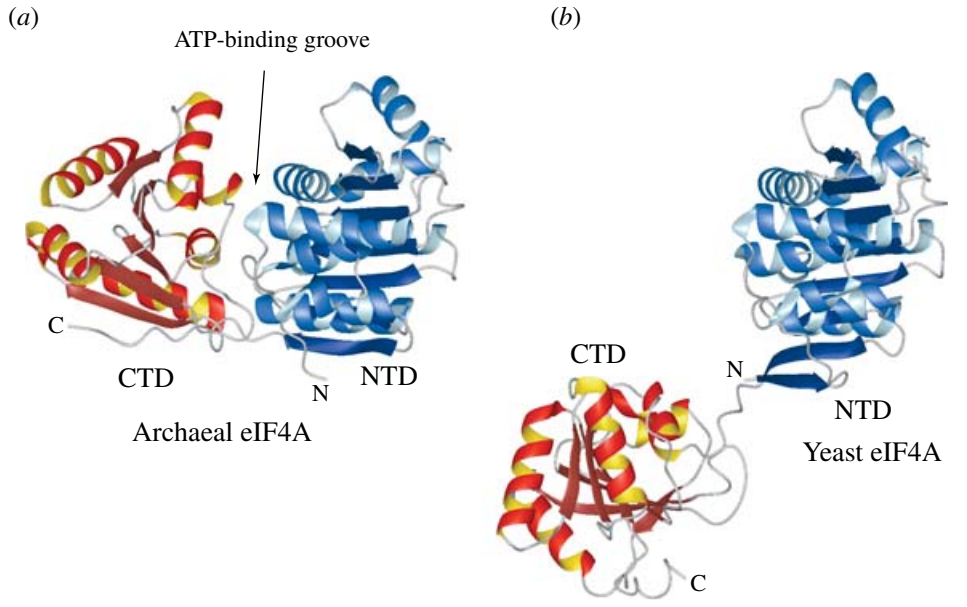

Fig. 13. eIF4A. (*a*) Structure of the archaeal eIF4A from *M. jannaschii*, PDB code 1HV8, in a closed 'active' conformation. The N-terminal domain (NTD) is in blue, the C-terminal domain (CTD) in red. The ATP-binding groove formed between the two domains is indicated with an arrow. The C- and N-termini of the protein are indicated. (*b*) The structure of the homologous eIF4A from *S. cerevisiae*, PDB code 1FUU, in an open conformation. The coloring and labels are as in (*a*). The NTD has the same orientation as the NTD in (*a*).

dumbbell-like orientation in the yeast eIF4A structure was proposed by the authors to be due to crystal packing and should not be taken as an actual interdomain orientation. For the helicase activity of eIF4A the two domains must dynamically come together in the typical 'closed' orientation, as seen in the archaeal eIF4A structure (Story *et al.* 2001) (Fig. 13*a*) and in the structures of other helicases. Most of the helicase motifs are found at the interface between the two RecA-like domains and are highly conserved in eIF4A from different species. Many of the motifs were unresolved in the crystal structure of yeast eIF4A (Caruthers *et al.* 2000) and are unstructured in isolated human eIF4A-CTD (M. Oberer, A. Marintchev & G. Wagner, unpublished observations). For a recent review on helicase structure and function see Delagoutte & von Hippel (2002, 2003).

The presence of two eIF4A-binding sites in eIF4G (see the previous section) poses a new question: does eIF4G bind one or two eIF4A molecules simultaneously? The answer is subject to controversy, with one group reporting a 1:2 binding (Korneeva *et al.* 2001), and others reporting 1:1 binding mode (Li *et al.* 2001). Another unresolved issue is whether eIF4A interacts with the two HEAT domains of eIF4G using the same or different surfaces. The former group reported, that one of the eIF4A-binding sites of eIF4G can compete with eIF4A binding to the other site, indicating anticooperativity between simultaneous binding of the same eIF4A molecule with both sites on eIF4G (Korneeva *et al.* 2001). Others have reported, that another eIF4A-binding protein, Pdcd4 competed for eIF4A binding with the second, but not the first HEAT domain of eIF4G (Yang *et al.* 2003). Pdcd4 contains two copies of a domain homologous to the second HEAT domain of eIF4G and the results in (Yang *et al.* 2003) are consistent with the existence of two separate surfaces on eIF4A, interacting with the two binding sites

on eIF4G. The middle domain of eIF4G (the first HEAT domain) binds predominantly to eIF4A-CTD, but both domains of eIF4A are necessary for maximum binding (M. Oberer, A. Marintchev & G. Wagner, unpublished observations) and the dynamic orientation between the two domains of eIF4A adds to the complexity of the problem.

The association of eIF4A with eIF4G and eIF4E near the cap poses a question about the orientation of eIF4A with respect to the small ribosomal subunit: both upon recruitment to mRNA and during scanning. On the one hand, it is simple to place all three proteins on the 5′-side of the small subunit. On the other hand, mechanistically it is expected that the eIF4A helicase must be in front of the scanning complex, with the rest of the translation machinery following and stabilizing the single-stranded conformation of mRNA behind eIF4A. The path of mRNA through the ribosome, and the directions, in which mRNA 'exits' the small subunit (Yusupova *et al.* 2001), as well as the overall sizes of eIF4A, G and E, indicate that such an arrangement is structurally feasible. Consistent with this idea, it was found that on some IRESs, eIF4G and eIF4A are stably bound upstream of the 40S-binding site, but promote destabilization of downstream secondary structure in an ATP-dependent manner (Kolupaeva *et al.* 2003).

Most organisms have more than one isoform of eIF4A. For example, human has three isoforms. The presence of multiple gene copies and isoforms is fairly common among translation factors. In some cases, the individual sequences are identical or nearly identical, allowing higher levels and/or differential regulation of expression. In other cases, as with isoforms I and II of human eIF4A, the sequences have diverged significantly, but the proteins appear to have similar (if not identical) functions and properties. Certain isoforms have different sets of interacting partners and different functions. eIF4A-III, for example, does not support translation *in vitro*, appears to bind only to the first HEAT domain of eIF4G (Li *et al.* 1999) and was reported to have a role in nonsense-mediated decay (Palacios *et al.* 2004), even though it has $>60\%$ sequence identity with the other two eIF4A isoforms.

5.6.4 eIF4B and eIF4H

eIF4B ($\sim$70 kDa) and eIF4H ($\sim$25 kDa) have similar function in stimulation of the helicase activity of eIF4A. eIF4H is homologous to part of eIF4B, including a conserved RRM domain near the N-terminus of both proteins. eIF4B has been reported to dimerize through a DRYG repeat region, C-terminal from the RRM. Both eIF4B and eIF4H bind RNA, and regions outside the RRMs are required for maximum binding (reviewed in Hershey & Merrick, 2000). The reported ssRNA-binding and strand annealing properties of eIF4B and eIF4H suggest a possible mechanism for their stimulatory effect on eIF4A: similar to ssDNA-binding proteins (SSBs) associated with other helicases, they could stabilize newly formed ssRNA and contribute to the processivity of the eIF4A helicase. Although mechanistically direct binding of SSBs to the helicase is not absolutely required, it allows better coordination and stimulation, and SSBs often interact with the associated helicase, at least weakly (for a review see, e.g. Delagoutte & von Hippel, 2002, 2003). No direct physical interaction has been reported between eIF4B (or eIF4H) and either eIF4A or eIF4G, but weak interactions may be hard to detect, especially if they depend on RNA binding. eIF4B, however, has been found to bind to eIF3 (reviewed in Hershey & Merrick, 2000). Thus, at least this factor is physically associated with the scanning 43S IC.

The structural homology between CBP80 and eIF4G (see Section 5.6.2 above) can be extended even further. CBP20, the cap-binding subunit of the nuclear cap-binding complex (CBC), is homologous to the RRMs and flanking regions of eIF4H ($\sim$30% identity). The entire

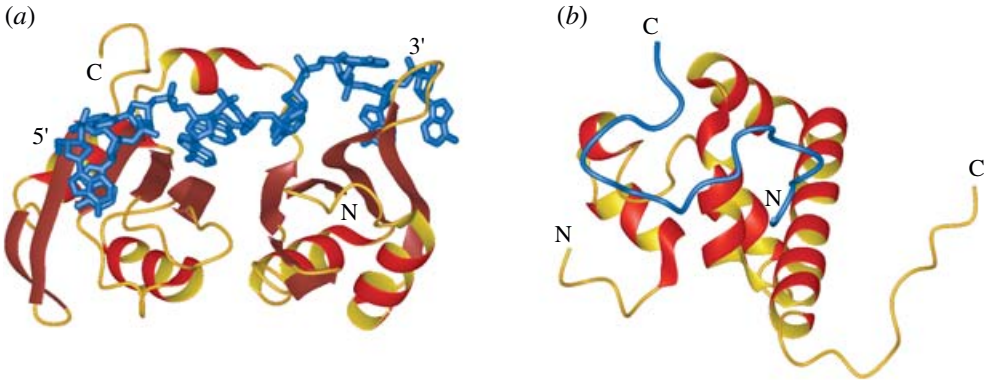

Fig. 14. Poly-A binding protein (PABP). (*a*) Structure of the complex of the first two RRMs of human PABP with RNA, PDB code 1CVJ. The protein backbone is shown as ribbon and colored red and orange. The RNA is in blue. (*b*) Structure of the complex of the PABP C-terminal domain (PABC) with a peptide from Paip2, PDB code 1JGN. PABC is in red and orange, and the peptide is in blue. The C- and N-termini of the proteins, and the 5′- and 3′-ends of the RNA are indicated.

CBP20, and not just the RRM, is required for cap binding. Most of the residues important for cap binding by CBP20 (Calero *et al.* 2002; Mazza *et al.* 2002) are conserved in eIF4H, but one of the aromatic residues 'sandwiching' the cap is missing. Parts of CBP20 outside the RRM were unstructured in apo-CBC, but folded upon cap binding. Furthermore, free CBP20 was found to be unstructured (Calero *et al.* 2002; Mazza *et al.* 2002). Although direct interaction between eIF4G and eIF4H has not been reported, the two proteins interact functionally in stimulating the helicase activity of eIF4A. It is tempting to speculate, that human eIF4G and eIF4H could interact similar to their structural homologs, CBP80 and CBP20 (Fig. 12*e*). The same parallel cannot be extended to yeast, because yeast eIF4G lacks the second and third HEAT domains, corresponding to the principal CBP20-binding site on CBP80. It is interesting to note, that free eIF4H is highly unstable (Hershey & Merrick, 2000) and appears at least partly unstructured (H. Matsuo & G. Wagner, unpublished observations).

The sequence of the RRM of eIF4H shows stronger homology to CBP20 and eIF4B (>30% identity), than CBP20 and eIF4B to each other. The structure of the RRM from eIF4B was recently published (Fleming *et al.* 2003) and, although the coordinates have not been deposited by the authors, the published figures and the homology with the RRM from CBP20 (~25% sequence identity) make it safe to assume, that the structure of the RRM of eIF4B is very similar to that of CBP20 and eIF4H.

5.7 PABP

PABP binds specifically to the poly-A tail of mRNA and bridges the 5′- and 3′-ends through interaction with eIF4G. As explained above, the interactions in the 5′-cap·eIF4E·eIF4G·PABP·3′-poly-A network are cooperative. The main functions of PABP in translation initiation are recruitment of the 43S IC to the mRNA, transmission of signals between the 3′- and 5′-ends, and protection of the IC from nucleases and translation repressors (reviewed in Hershey & Merrick, 2000; Sachs, 2000; Preiss & Hentze, 2003; Prevot *et al.* 2003).

PABP consists of four RRMs, a Gln-rich segment, and a CTD (PABC). At least the first two RRMs are involved in poly-A binding. The structure of the first two RRMs of PABP in complex

with poly-A was determined (Deo *et al.* 1999) (Fig. 14*a*). The structures of the PABC domain from several species, as well as its complex with a consensus peptide have also been solved (Kozlov *et al.* 2001, 2002, 2004; Siddiqui *et al.* 2003) (Fig. 14*b*). A structure was also solved for a close PABC homolog found in HYD – a protein that is not linked to translation, but may share interacting partners with PABP (Deo *et al.* 2001).

A number of proteins containing a consensus sequence called PABP-binding motif-1 (PAM1) bind to the RRMs of PABP. The Gln-rich segment resembles prion-like domains and has been implicated in dimerization/oligomerization. The PABC domain is involved in a number of interactions with proteins containing the PAM2 consensus sequence. The length of the poly-A tail is sufficient to allow binding of multiple PABP molecules and, therefore, more than one of the PABP interaction partners can be associated simultaneously with the same poly-A tail. eIF4G has a PAM1 motif, whose binding to PABP is cooperative with PABP binding to mRNA. At least two proteins, Paip1 and Paip2, have both the PAM1 and PAM2 motifs and thus have two binding sites on PABP. Paip2 was found to inhibit translation; it bound to PABP in a 2:1 ratio and the binding to the RRMs was anticooperative with PABP binding to mRNA. Free Paip2 appeared unstructured and the consensus peptides were sufficient for near-maximum binding. Conversely, Paip1 was found to stimulate translation; to bind in a 1:1 stoichiometry with PABP, without interfering with mRNA binding. Paip1 contains a HEAT domain, homologous to the first HEAT domain of eIF4G, immediately after its PAM2 motif and the presence of the HEAT domain was required for high affinity binding to PABP. In higher eukaryotes, but not in yeast, the termination factor eRF3 contains a PAM2 motif and binds to PABP. This finding strongly supports a proposed role for PABP in bringing the stop codon near the 5′- and 3′-end, and channeling 40S subunits from the terminating ribosomes back to the 5′-cap of the same mRNA (reviewed in Sachs, 2000; Preiss & Hentze, 2003; Prevot *et al.* 2003; Sonenberg & Dever, 2003).

6. Kinetic aspects of translation initiation and its regulation

The process of initiation is schematically presented in Fig. 2. It can easily be appreciated that translation initiation has little in common with the simple enzymic reactions taking place under ideal Michaelis–Menten conditions. Instead, it is a multistep process, involving a number of factors and subject to a complex regulatory network. To complicate matters even further, the convention is to write equations from the standpoint of mRNA as the enzyme and the ribosomal subunits and all IFs – as co-factors. This is dictated by practical reasons: the large ribosomal subunit, which carries the catalytic activity, joins near the end of initiation and then spends minutes through elongation and termination, after which it does not necessarily come back to the same mRNA. At the same time, multiple initiation events occur on the mRNA during one elongation/termination cycle. There are few studies dedicated to the kinetics of translation initiation. However, every time we try to interpret changes in expression of a reporter protein or rates of initiation, we are forced to deal with enzyme kinetics in an extremely complex system.

In the absence of an extensive thermodynamic framework for initiation, we have tried to promote certain general approaches in dealing with, and making sense of the scarce and often controversial kinetic data. The available information about the kinetics of translation initiation is very limited and for the most part concerns the indirect estimation of initiation as a whole, whereas little is known about the individual steps. It appears that in most cases initiation, and not elongation, is rate-limiting. Initiation frequency can be estimated from the polysome

size on a given mRNA and the time it takes for one round of translation. Elongation rates of ~1–8 codons/s have been measured in eukaryotes, and a few-fold higher in bacteria. Actual elongation rates would vary depending on conditions, mRNA codon composition and secondary structure. Initiation rates of once every few seconds have been measured experimentally on highly expressed mRNAs under optimal conditions in both bacteria and eukaryotes (reviewed in Mathews *et al.* 2000).

6.1 General considerations

The kinetics of translation initiation is probably one of the most common reasons why different groups have reported conflicting results about the roles of individual factors, or have proposed conflicting interpretations for the same results. As it has been postulated that the number of people reading a book is inversely proportional to the number of equations in it (Hawking, 1998), we will try to present a pictorial, rather than mathematical, description of the kinetics of translation initiation. We will start with several hypothetical examples, which could help visualize the results of scenarios relevant to initiation.

In a multistep process, such as translation initiation, the overall time for completion of the process is the sum of the times for all individual *sequential* steps. If multiple *physical* steps can take place in parallel, the length of the corresponding *kinetic* step would be the time it takes to complete the slowest *physical* step required, before the next sequential step can begin (rather than the sum of the times for the individual parallel *physical* steps). Thus, whereas any step can be rate-limiting, not all steps are simultaneously rate-limiting under a given set of conditions. This can be illustrated by the examples given below for a two-step process: step 1 from complex A to complex B, and step 2 – from complex B to complex C.

(1) If step 1 takes 40 s and step 2 takes 2 s, the overall time is $40+2=42$ s. If both factors x and y act independently and increase 10-fold the rate of step 1, the time for step 1 will be reduced to 4 s in the presence of any factor alone. The overall time for the entire process will then be $4+2=6$ s. In the presence of both factors x and y the time for the first step will be $0{\cdot}4$ s, and the overall time for the entire process will be $0{\cdot}4+2=2{\cdot}4$ s. In other words, each factor alone stimulates the rate of the process ~7-fold (42/6) when added alone, but only ~2·5-fold (6/2·4), when added to a reaction already containing the other factor, because the second step becomes rate-limiting in the presence of both factors.

(*a*) $\mathrm{A} \xrightarrow[40\,\mathrm{s}]{} \mathrm{B} \xrightarrow[2\,\mathrm{s}]{} \mathrm{C} \qquad \mathrm{A} \xrightarrow[42\,\mathrm{s}]{} \mathrm{C} \qquad \text{Rate: } 0{\cdot}024\,\mathrm{s}^{-1}$

(*b*) $\mathrm{A} \xrightarrow[4\,\mathrm{s}]{+x} \mathrm{B} \xrightarrow[2\,\mathrm{s}]{} \mathrm{C} \qquad \mathrm{A} \xrightarrow[6\,\mathrm{s}]{} \mathrm{C} \qquad \text{Rate: } 0{\cdot}17\,\mathrm{s}^{-1}$

(*c*) $\mathrm{A} \xrightarrow[4\,\mathrm{s}]{+y} \mathrm{B} \xrightarrow[2\,\mathrm{s}]{} \mathrm{C} \qquad \mathrm{A} \xrightarrow[6\,\mathrm{s}]{} \mathrm{C} \qquad \text{Rate: } 0{\cdot}17\,\mathrm{s}^{-1}$

(*d*) $\mathrm{A} \xrightarrow[0{\cdot}4\,\mathrm{s}]{+xy} \mathrm{B} \xrightarrow[2\,\mathrm{s}]{} \mathrm{C} \qquad \mathrm{A} \xrightarrow[2{\cdot}4\,\mathrm{s}]{} \mathrm{C} \qquad \text{Rate: } 0{\cdot}42\,\mathrm{s}^{-1}$

(2) If factor z stimulates 10-fold the rate of step 2, its addition will have no apparent effect under conditions where step two is not rate-limiting: in the absence of factors x and y, the overall time will only drop from $40+2=42$ s, to $40+0{\cdot}2=40{\cdot}2$ s. In the presence of either

x or *y*, the overall time of the process will only drop from 4 + 2 = 6 s, to 4 + 0·2 = 4·2 s, or 30%, which is usually within the experimental error. In the presence of both *x* and *y*, step 2 becomes rate-limiting and addition of *z* has a significant effect: the overall time of the entire process will now drop from 0·4 + 2 = 2·4 s, to 0·4 + 0·2 = 0·6 s, or 4-fold – compare example 2(*a–d*) below with example 1(*a–d*) above.

(*a*) $A \xrightarrow[40\,s]{} B \xrightarrow[0{\cdot}2\,s]{+z} C \qquad A \xrightarrow[40{\cdot}2\,s]{} C \qquad \text{Rate: } 0{\cdot}025\,s^{-1}$

(*b*) $A \xrightarrow[4\,s]{+x} B \xrightarrow[0{\cdot}2\,s]{+z} C \qquad A \xrightarrow[4{\cdot}2\,s]{} C \qquad \text{Rate: } 0{\cdot}24\,s^{-1}$

(*c*) $A \xrightarrow[4\,s]{+y} B \xrightarrow[0{\cdot}2\,s]{+z} C \qquad A \xrightarrow[4{\cdot}2\,s]{} C \qquad \text{Rate: } 0{\cdot}24\,s^{-1}$

(*d*) $A \xrightarrow[0{\cdot}4\,s]{+xy} B \xrightarrow[0{\cdot}2\,s]{+z} C \qquad A \xrightarrow[0{\cdot}6\,s]{} C \qquad \text{Rate: } 1{\cdot}7\,s^{-1}$

The conclusions from the above examples may seem counterintuitive. If a factor has similar effect on translation initiation in both presence and absence of another factor, it is concluded that the two factors act independently, which is often extended to mean that they act independently and on different steps. It is obvious from the above examples that the two factors may act independently, but on the same step (which has to be the rate-limiting step). For example, in many systems, capping or polyadenylation of mRNAs alone bring 5- to 50-fold stimulation of translation, and together they act synergistically, indicating that they both act on the same step (recruitment of 43S IC to mRNA), which is the rate limiting step. However as in example (2) above, if the eIF2 GTP hydrolysis step is slow because of a mutation in eIF5, or when the subunit joining step is slow in the absence of eIF5B (see e.g. Searfoss *et al.* 2001), then polyadenylation has no significant effect on translation initiation, because recruitment of the 43S IC is no longer the rate limiting step.

(3) Only in the extreme case, where two steps have comparable rates, stimulating any of the two steps could have a moderate (only up to 2-fold) effect on the overall rate. If, in the example above, steps 1 and 2 take 1 s each (and the overall time of the process is 1 + 1 = 2 s), increasing 10-fold the rate of step 1 by factor *x*, or increasing 10-fold the rate of step 2 by factor *z* will each lead to almost 2-fold reduction of the overall time: 1 + 0·1 = 1·1 s. Adding both *x* and *z* together will lead to 10-fold stimulation: 0·1 + 0·1 = 0·2 s. Such apparent cooperative effect (compare 1·8 × 1·8 = **3·24** to the observed **10**-fold stimulation) can easily be misinterpreted as proof for the factors *x* and *z* acting together on the same step. It is, therefore, useful to remember that a kinetic experiment can disprove a mechanism, can be consistent with an alternative mechanism, but alone cannot prove a mechanism, even in a simple system. The effects of inhibitors are even more difficult to interpret, because any step can be made rate limiting, if it is severely inhibited.

(*a*) $A \xrightarrow[1\,s]{} B \xrightarrow[1\,s]{} C \qquad A \xrightarrow[2\,s]{} C \qquad \text{Rate: } 0{\cdot}5\,s^{-1}$

(*b*) $A \xrightarrow[0{\cdot}1\,s]{+x} B \xrightarrow[1\,s]{} C \qquad A \xrightarrow[1{\cdot}1\,s]{} C \qquad \text{Rate: } 0{\cdot}9\,s^{-1}$

(*c*) $A \xrightarrow[1\,s]{} B \xrightarrow[0{\cdot}1\,s]{+z} C$ $\qquad A \xrightarrow[1{\cdot}1\,s]{} C$ $\qquad$ Rate: $0{\cdot}9\,s^{-1}$

(*d*) $A \xrightarrow[0{\cdot}1\,s]{+x} B \xrightarrow[0{\cdot}1\,s]{+z} C$ $\qquad A \xrightarrow[0{\cdot}2\,s]{} C$ $\qquad$ Rate: $5\,s^{-1}$

The reason we spend so much time on these examples is that they illustrate the dependence of the effects of IFs on the system, mRNA, other factors and conditions. If the rate-limiting step for translation initiation on a given mRNA, under given conditions is known, the effect (or lack thereof) of a factor can provide insights about the step(s) it affects and the types of mRNA it acts on. Conversely, if the exact function of a factor is known, its effects can help identify the rate-limiting step(s) for the specific combination of mRNA and conditions. Furthermore, well-characterized mutant IFs can be used to manipulate the kinetics of translation initiation, identify rate-limiting steps and compare rates.

Kinetic steps can be either concentration-independent reactions occurring in a complex with a rate determined by a first-order rate constant with a unit of s^{-1}, or bi- (or multi-) molecular binding steps, whose rate depends on the concentrations of the molecules involved. The rate of a step, where a molecule binds to the enzyme, is directly proportional to the corresponding rate constant (with unit of $M^{-1}\,s^{-1}$) and the concentration of the molecule. Even for a multistep process like translation initiation, one can define an apparent K_M for any molecule, corresponding to the concentration that gives half-maximum rate of initiation. That K_M, however, will depend on a network of rates of individual steps, concentrations of other molecules, and the action of regulators. In other words, the K_M of a factor can be completely different for different types of mRNA under the same conditions and for the same type of mRNA under different conditions, and thus published K_M values can seldom be safe to use directly. Accordingly, changing the concentration of an IF will affect the rate of initiation on a given mRNA if the concentration of the IF is lower than or comparable to the K_M, but will have no effect on another mRNA with much lower K_M for the same IF. Therefore, lack of effect from depleting a factor on a given mRNA does not prove that initiation from that mRNA is independent of that factor. Several very interesting examples from eukaryotic initiation can be found in (Jackson, 2000).

6.2 Translation initiation in bacteria

(1) It has been generally accepted that mRNA is in excess with respect to ribosomes and/or IFs under a number of circumstances. The affinity of the IC for different mRNAs varies significantly, and the translation of 'weakly binding' mRNAs is stimulated under conditions where 'strongly binding' mRNAs are saturated with ICs.
(2) Increasing stability of the IC on mRNA stimulates the rate of initiation only to the point where the lifetime of the complex is sufficient to allow initiation. Greater stability would not increase initiation rates further and could even slow down the clearance from the start site after initiation. A major determinant in ribosome recruitment in bacteria is the SD sequence. The optimal stability of the SD–rRNA interaction is start-site dependent: a start site containing suboptimal spacing between the AUG and the SD element or secondary structure may need longer time for initiation, and thus stronger SD–rRNA binding.
(3) Most interactions have affinities in the nM range and, therefore, at physiological concentrations (typically in the μM range), the equilibrium will be shifted toward the bound state.

A second set of interactions are weak and not detected by standard biochemical methods in solution under physiological concentrations (e.g. between IF1 and IF2). These could exist mainly in the context of larger complexes and could have a role in the cooperativity of complex formation and/or in structural rearrangements in the IC. The cooperativity in complex formation shifts the equilibrium toward complexes containing a full set of factors and ensures that the IC has lifetime sufficient for initiation to take place.

(4) The potential kinetic rate-limiting steps are: (i) First order reactions that occur in the IC: rearrangement upon start codon recognition, GTP hydrolysis, release of IF2, and formation of the first peptide bond. (ii) Binding of factors to the IC and of the IC to mRNA, where the rates are dependent on the factor concentrations.

(5) If ribosomes and all necessary factors are present in excess with respect to their apparent K_M values, higher order rate constants for the binding steps are converted into pseudo-first order and the binding steps are not rate-limiting. The main rate-limiting steps then are the first-order reactions occurring in the IC upon mRNA binding and initiation operates at a near-maximum rate. The apparent K_M values for individual factors would vary among different start sites, depending on the binding affinity of the ribosomes and on the maximum rate of initiation (see above). Thus in the same cell, some mRNAs can be translated at near-maximum rates, while initiation from other may occur at only a fraction of the maximum rate.

(6) The scenario becomes completely different if special cases are considered. For example, as discussed above, initiation on noncanonical start codons is stimulated by low, but inhibited by higher concentrations of IF3 (reviewed in Petrelli *et al.* 2001). Here, the outcome will be determined by competition between IF2·GTP-promoted subunit joining (which makes the IC resistant to IF3) and IF3-promoted dissociation of the IC from the non-cognate codon.

(7) Kinetic studies of the events occurring after mRNA binding *in vitro* have yielded rate constants of $\sim 1 \times 10^6$ M^{-1} s^{-1} for subunit joining (Antoun *et al.* 2003) and 0·2–0·3 s^{-1} for formation of the first peptide bond (Tomsic *et al.* 2000). Other parameters were also reported by the second group, but as discussed in Section 4.5 above, may have been influenced by the use of exogenous IF2 from *B. stearothermophilus*, which is known to bind more tightly to the *E. coli* ribosome than the endogenous IF2. Subunit joining is a bimolecular reaction and its rate will depend on the concentration of 50S subunits. The two first-order rate constants, corresponding to GTP hydrolysis and IF2·GDP dissociation, appear faster than the formation of the first peptide bond, but should add some delay to the overall rate of initiation. In the context of reported initiation rates of the order of 0·3 s^{-1} in *E. coli*, the formation of the first peptide bond appears to be a major rate-limiting step under conditions of near-maximum rate of initiation. It should be noted, that the rate of formation of the peptide bond depends on the nature of the amino acid following fMet and may be influenced by EF-P/IF5A. Thus, it could be significantly different on individual mRNAs *in vivo* (reviewed in Mathews *et al.* 2000; Ganoza *et al.* 2002).

6.3 Translation initiation in eukaryotes

The above general considerations also apply to eukaryotic initiation. The main features specific to eukaryotes are cap-binding complex-mediated binding of the IC to mRNA and the presence of a new step, scanning. The extra GTP hydrolysis step (by eIF2) does not add a new kinetic step, but rather makes the existing step of start site recognition more complex and potentially

slower. As in bacteria, the K_Ms of eIFs and ribosomes vary significantly depending on the mRNA and thus initiation at different mRNAs operates under different kinetic conditions.

(1) Under conditions of excess of factors and ribosomes, with respect to their K_M values for a given mRNA type, 43S complex formation is not rate-limiting; a cap-binding complex is present at the 5′-end of all mRNA molecules of this type all the time; and binding of the 43S complex at the cap is fast, compared to later steps in initiation.
(2) On mRNAs with short 5′-UTR, the 43S complex will be recruited at or near the start codon, minimizing the time required for scanning. Furthermore, the analysis of initiation at such mRNAs will not depend on the answers of the questions whether the cap-binding complex remains associated with the small subunit during scanning or whether scanning is random or predominantly 5′ to 3′ (see Section 4.6.1 above). Thus, under optimal conditions, most potential rate-limiting steps will be reactions occurring in the IC, similar to those in prokaryotes, except that start codon recognition will be more complex and involve a GTP hydrolysis step by eIF2. Accordingly, the highest initiation rates in eukaryotes are not much slower than those in bacteria (Ganoza *et al.* 2002; Mathews *et al.* 2000).
(3) The two major targets for regulation of initiation are the cap-binding complex and the concentration of the eIF2·GTP·Met-tRNA$_i$ complex. The mechanisms of regeneration of the eIF2·GTP·Met-tRNA$_i$ complex and its regulation were described in Section 5.4.2 above. Moderate reduction of eIF2·GTP·Met-tRNA$_i$ levels has mostly regulatory role of maintaining adequate supply of amino acids through feedback loops (see Section 4.8). More significant reduction decreases the concentration of functional 43S IC and will cause overall inhibition of translation, up to nearly complete shut-off. At the same time, initiation from certain mRNAs will be differentially affected and could even be stimulated.
(4) There is little doubt that the cap-binding complex stimulates binding of the 43S IC to mRNA and that inhibition of its assembly at the cap reduces the rate of 43S binding to mRNA, making it the rate-limiting step (if it was not already). A simple way to look at the effect of decreasing eIF2·GTP·Met-tRNA$_i$ levels is that it leads to decrease in the concentration of 43S ICs and in essence inhibits the same step in initiation as inhibition of cap-binding complex assembly. Unfortunately, few aspects of translation initiation are simple and the reduction of eIF2·GTP·Met-tRNA$_i$ levels could also have other consequences. If, as discussed above, ICs lacking eIF2·GTP·Met-tRNA$_i$ can scan upon reinitiation, before acquiring eIF2·GTP·Met-tRNA$_i$, it is possible that they could also initiate scanning. In this case, however, the corresponding IC should have lower stability and weaker affinity for mRNA. If such an IC scans through the start site before acquiring eIF2·GTP·Met-tRNA$_i$, the IC and the mRNA will be unavailable for initiation until the IC dissociates, at least from the cap.
(5) Limiting eIF4E, 4G, and/or PABP will have differential effects on different mRNAs, depending on their binding affinities, presence and length of a poly-A tail, etc. Limiting eIF4E would lead to not all 5′-caps being bound. eIF4E could be limiting for some caps even if present in excess with respect to mRNA – depending on the ratio between concentration of free eIF4E in the cytoplasm and its affinity for the caps. As discussed in Sections 4.6 and 5.6.2 above, the interactions among the 5′-cap, eIF4E, eIF4G, PABP, and the 3′-poly-A tail are mutually cooperative. Therefore, the affinities of eIF4E, eIF4G and PABP for any given mRNA, and of eIF4G for the 43S pre-IC will depend on a network of concentrations and affinities of all other factors for that mRNA. Thus, the first mRNAs to be affected by

low eIF4E concentrations may not always be those whose 5′-cap regions have lowest affinity for free eIF4E (e.g. 'masked' by RNA secondary structure). The analysis is further complicated by the presence of multiple binding sites for PABP on a single poly-A tail.

(6) The roles of eIF4A, eIF4B and eIF4H on mRNAs with structured 5′-UTRs can be subdivided into: (i) removal of secondary structures near the 5′-end, which would increase the rate of 43S binding to mRNA, and (ii) removal of mRNA secondary structures during scanning, which would increase the rate of scanning through structured 5′-UTRs. These factors also appear to be an integral part of the eIF4E-4G-PABP 'network'.

(7) If the start codon is not in 'good' context (Kozak sequence), then a fraction of the ICs will scan through and may spend a long time on the mRNA. Unlike the 'simplified' scenario above, the effect of leaky scanning on further initiation events depends on the stability of the cap-eIF4E-eIF4G-43S complex interactions: a new round of initiation will not start until the IC is released from the cap, due to either initiation or breaking of a bond. On the flip side, such mRNAs with a leaky start site are a powerful tool for studying the stability, roles, and interplay of eIF4E-cap, eIF4E-eIF4G, eIF4G-PABP, and eIF4G-43S complex interactions. Under conditions of multiple rounds of initiation, mRNAs containing ICs that have scanned through the start codon will accumulate, allowing quantitation of the effects of the stability of the cap-IC interactions on initiation. 5′-caps with lower or higher affinity for eIF4E can be designed and used in the presence or absence of a poly-A tail, and at varying factor concentrations. The mRNA can be designed with and without a downstream ORF, secondary structure, etc.

(8) As stated above, on mRNAs and under conditions where 43S binding and scanning are not rate-limiting, the remaining potential rate-limiting steps are start-site recognition, subunit joining and synthesis of the first peptide bond. Start site recognition involves GTP hydrolysis by eIF2, triggered by (and possibly triggering) conformational changes upon codon-anticodon basepairing, and dissociation of eIF2 from the IC. The exact rates of GTP hydrolysis or of eIF2 dissociation are not known. As the maximum rates of initiation in eukaryotes are in the order of once every several seconds, at least in such cases the time from start-site recognition to release of eIF2·GDP should be in the order of seconds or faster. While it is not clear if GTP hydrolysis can be rate limiting *in vivo*, a mutation in yeast eIF5, which artificially slows down eIF2·GTP hydrolysis (Donahue, 2000), does make this step rate-limiting on most mRNAs (on which the other steps, such as 43S binding to mRNA, are fast). As discussed above, although eIF2·GDP must dissociate from Met-tRNA$_i$, it is not clear how fast this happens, whether other factors influence the off rate, or even whether eIF2·GDP is released completely from the IC or only from Met-tRNA$_i$.

(9) The mechanisms of subunit joining and synthesis of the first peptide bond are highly conserved with bacteria and it is thus likely that their rates are of the same order of magnitude. The rate of subunit joining is proportional to the concentration of 60S subunits and depends on the presence of eIF5B·GTP. As with the eIF5 mutant above, in the absence of eIF5B, subunit joining becomes rate limiting for most mRNAs (Choi *et al.* 1998; Pestova *et al.* 2000). It is probably safe to assume that, as in bacteria (Antoun *et al.* 2003), GTP hydrolysis and subsequent release of eIF5B·GDP are not major rate-limiting steps with wild type eIF5B, but become rate limiting with some eIF5B mutants (Lee *et al.* 2002; Shin *et al.* 2002; Antoun *et al.* 2003). The synthesis of the first peptide bond is probably rate limiting on highly expressed mRNAs under optimal conditions: like in

bacteria, it may take seconds and is likely dependent on the second amino acid (the first being Met) and eIF5A (Mathews *et al.* 2000; Ganoza *et al.* 2002).

(10) The role of PABP is more complex: it has a major role in protection of polyadenylated mRNA from nucleolytic degradation (reviewed in Dever, 2002; Preiss & Hentze, 2003; Prevot *et al.* 2003). PABP was found to prevent inhibition of translation by two proteins, Ski2p and Slh1p in yeast, rather than (or in addition to) directly stimulating translation (Searfoss *et al.* 2001). In view of the above discussion, it would appear, that Ski2p and Slh1p themselves would have to inhibit the assembly of the cap-binding complex and/or 43S recruitment [assuming that the possible effects of the nucleases in the Ski2p/Slh1p-containing multiprotein complex were accounted for in Searfoss *et al.* (2001)].

(11) A number of translation regulators affect the assembly of the cap-binding complex and/or the recruitment of the 43S IC. 4E-BPs compete with eIF4G for binding to eIF4E. The translational activator Paip1 binds to PABP and could bind simultaneously with eIF4G, but to different PABP molecules at the 3′-end of the mRNA, providing extra binding sites for eIF4A, for example. The translational inhibitor Paip2 binds to PABP and appears to interfere with poly-A binding. Proteins like DAP5, homologous to the C-terminal part of eIF4G contain binding sites not only for eIF4A, but also for eIF3. They do not, however contain binding sites for eIF4E or PABP and can compete with the cap-binding complex for binding to the 43S. The corresponding fragment of eIF4G can be produced by proteolytic cleavage (e.g. by some viral proteases) and both it and DAP5 could support IRES-dependent initiation (reviewed in Dever, 2002; Preiss & Hentze, 2003; Prevot *et al.* 2003). It is obvious that the effects of all these factors also need to be looked at in the context of the 4E-4G-PABP 'network', because the affinity of the competing interaction will be 'network'-determined. For example, the binding of 4E-BP to eIF4E is not cooperative with eIF4E binding to the cap and, therefore, 4E-BPs will have a better chance competing for free eIF4E.

7. Concluding remarks

In the above sections, we have tried to summarize the current knowledge about translation initiation. Here, we will try to focus briefly on some of the major gaps in that knowledge. From a structural biology perspective, the structures of the bacterial ribosome and two IFs have been solved, although the structure of a large portion of IF2 remains unknown. The structures of a large number of eukaryotic translation factors are available and there are cryo-EM reconstructions of 40S and 80S ribosomes, but high resolution structures of eukaryotic ribosomes, subunits and ICs are still lacking. Furthermore, there is only limited structural information for the largest eukaryotic IFs, like eIF3 and eIF2B.

The main steps in translation initiation and the functions of the individual factors are now known, but a detailed understanding of the underlying mechanism on a molecular level seems still far away. Advances have been made in identifying the interactions among eukaryotic IFs and ribosomes, and new interactions are being discovered at increasing rates. There are multiple reports of cooperative and competing interactions among IFs. The steps in initiation, at which individual factors are recruited and released from the ICs, are becoming increasingly clear. However, the organization and dynamics of the interaction network in eukaryotic translation initiation present a special challenge and the available information is still limited. Weak

interactions that are formed only transiently and/or found only in the context of higher order complexes, are especially difficult to study and may require development of radically new approaches.

In summary, knowledge about the interaction network and the structures of ICs would allow detailed mechanistic studies in translation initiation at a level that is now only achieved for translation elongation.

8. Acknowledgments

We thank Tatyana Pestova for helpful discussions and for communicating results before publication, and the members of the Wagner laboratory for their comments and input. A.M. was funded by a post-doctoral fellowship PF-02-117-01-GMC from the American Cancer Society. G.W. acknowledges support by the National Institutes of Health (CA 68262, U19CA87427, GM47467, and RR-00995).

9. References

Adhin, M. R. & van Duin, J. (1990). Scanning model for translational reinitiation in eubacteria. *Journal of Molecular Biology* **213**, 811–818.

Altmann, M., Schmitz, N., Berset, C. & Trachsel, H. (1997). A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. *EMBO Journal* **16**, 1114–1121.

Andersen, C. F., Anand, M., Boesen, T., Van, L. B., Kinzy, T. G. & Andersen, G. R. (2004). Purification and crystallization of the yeast translation elongation factor eEF3. *Acta Crystallographica. Section D, Biological Crystallography* **60**, 1304–1307.

Andersen, G. R., Pedersen, L., Valente, L., Chatterjee, I., Kinzy, T. G., Kjeldgaard, M. & Nyborg, J. (2000). Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1Balpha. *Molecular Cell* **6**, 1261–1266.

Andersen, G. R., Valente, L., Pedersen, L., Kinzy, T. G. & Nyborg, J. (2001). Crystal structures of nucleotide exchange intermediates in the eEF1A-eEF1Balpha complex. *Nature Structural Biology* **8**, 531–534.

Antoun, A., Pavlov, M. Y., Andersson, K., Tenson, T. & Ehrenberg, M. (2003). The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO Journal* **22**, 5593–5601.

Aravind, L. & Koonin, E. V. (2000). Eukaryote-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system. *Genome Research* **10**, 1172–1184.

Aravind, L. & Ponting, C. P. (1998). Homologues of 26S proteasome subunits are regulators of transcription and translation. *Protein Science* **7**, 1250–1254.

Asano, K., Kinzy, T. G., Merrick, W. C. & Hershey, J. W. (1997). Conservation and diversity of eukaryotic translation initiation factor eIF3. *Journal of Biological Chemistry* **272**, 1101–1109.

Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G. D. & Hinnebusch, A. G. (1999). Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO Journal* **18**, 1673–1688.

Asano, K., Shalev, A., Phan, L., Nielsen, K., Clayton, J., Valasek, L., Donahue, T. F. & Hinnebusch, A. G. (2001). Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO Journal* **20**, 2326–2337.

Ashida, H., Saito, Y., Kojima, C., Kobayashi, K., Ogasawara, N. & Yokota, A. (2003). A functional link between RuBisCO-like protein of Bacillus and photosynthetic RuBisCO. *Science* **302**, 286–290.

Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2·4 A resolution. *Science* **289**, 905–920.

Batey, R. T., Gilbert, S. D. & Montange, R. K. (2004). Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432**, 411–415.

Battiste, J. L., Pestova, T. V., Hellen, C. U. & Wagner, G. (2000). The eIF1A solution structure reveals a large RNA-binding surface important for scanning function. *Molecular Cell* **5**, 109–119.

Bec, G., Kerjan, P. & Waller, J. P. (1994). Reconstitution in vitro of the valyl-tRNA synthetase-elongation factor (EF) 1 beta gamma delta complex.

Essential roles of the NH2-terminal extension of valyl-tRNA synthetase and of the EF-1 delta subunit in complex formation. *Journal of Biological Chemistry* **269**, 2086–2092.

Belfield, G. P., Ross-Smith, N. J. & Tuite, M. F. (1995). Translation elongation factor-3 (EF-3): an evolving eukaryotic ribosomal protein? *Journal of Molecular Evolution* **41**, 376–387.

Benne, R., Naaktgeboren, N., Gubbens, J. & Voorma, H. O. (1973). Recycling of initiation factors IF-1, IF-2 and IF-3. *European Journal of Biochemistry* **32**, 372–380.

Berthelot, K., Muldoon, M., Rajkowitsch, L., Hughes, J. & McCarthy, J. E. (2004). Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Molecular Microbiology* **51**, 987–1001.

Biou, V., Shu, F. & Ramakrishnan, V. (1995). X-ray crystallography shows that translational initiation factor IF3 consists of two compact alpha/beta domains linked by an alpha-helix. *EMBO Journal* **14**, 4056–4064.

Boesen, T., Mohammad, S. S., Pavitt, G. D. & Andersen, G. R. (2004). Structure of the catalytic fragment of translation initiation factor 2B and identification of a critically important catalytic residue. *Journal of Biological Chemistry* **279**, 10584–10592.

Boileau, G., Butler, P., Hershey, J. W. & Traut, R. R. (1983). Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane. *Biochemistry* **22**, 3162–3170.

Bommer, U. A., Kraft, R., Kurzchalia, T. V., Price, N. T. & Proud, C. G. (1991a). Amino acid sequence analysis of the beta- and gamma-subunits of eukaryotic initiation factor eIF-2. Identification of regions interacting with GTP. *Biochimica et Biophysica Acta* **1079**, 308–315.

Bommer, U. A., Lutsch, G., Stahl, J. & Bielka, H. (1991b). Eukaryotic initiation factors eIF-2 and eIF-3: interactions, structure and localization in ribosomal initiation complexes. *Biochimie* **73**, 1007–1019.

Bommer, U. A., Salimans, M. M., Kurzchalia, T. V., Voorma, H. O. & Karpova, G. G. (1988). Affinity labeling by a photoreactive GTP analogue of the alpha-subunit of eukaryotic initiation factor eIF-2 in different initiation complexes. *Biochemistry International* **16**, 549–557.

Brombach, M., Gualerzi, C. O., Nakamura, Y. & Pon, C. L. (1986). Molecular cloning and sequence of the *Bacillus stearothermophilus* translational initiation factor IF2 gene. *Molecular and General Genetics* **205**, 97–102.

Browning, K. S., Gallie, D. R., Hershey, J. W., Hinnebusch, A. G., Maitra, U., Merrick, W. C. & Norbury, C. (2001). Unified nomenclature for the subunits of eukaryotic initiation factor 3. *Trends in Biochemical Science* **26**, 284.

Budkevich, T. V., Timchenko, A. A., Tiktopulo, E. I., Negrutskii, B. S., Shalak, V. F., Petrushenko, Z. M., Aksenov, V. L., Willumeit, R., Kohlbrecher, J., Serdyuk, I. N. & El'skaya, A. V. (2002). Extended conformation of mammalian translation elongation factor 1A in solution. *Biochemistry* **41**, 15342–15349.

Bumann, M., Djafarzadeh, S., Oberholzer, A. E., Bigler, P., Altmann, M., Trachsel, H. & Baumann, U. (2004). Crystal structure of yeast Ypr118w, a methylthioribose-1-phosphate isomerase related to regulatory eIF2B subunits. *Journal of Biological Chemistry* **279**, 37087–37094.

Bushell, M., Wood, W., Clemens, M. J. & Morley, S. J. (2000). Changes in integrity and association of eukaryotic protein synthesis initiation factors during apoptosis. *European Journal of Biochemistry* **267**, 1083–1091.

Calero, G., Wilson, K. F., Ly, T., Rios-Steiner, J. L., Clardy, J. C. & Cerione, R. A. (2002). Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nature Structural Biology* **9**, 912–917.

Carberry, S. E., Darzynkiewicz, E., Stepinski, J., Tahara, S. M., Rhoads, R. E. & Goss, D. J. (1990). A spectroscopic study of the binding of N-7-substituted cap analogues to human protein synthesis initiation factor 4E. *Biochemistry* **29**, 3337–3341.

Carrera, P., Johnstone, O., Nakamura, A., Casanova, J., Jackle, H. & Lasko, P. (2000). VASA mediates translation through interaction with a Drosophila yIF2 homolog. *Molecular Cell* **5**, 181–187.

Carter, A. P., Clemons Jr., W. M., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T. & Ramakrishnan, V. (2001). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* **291**, 498–501.

Caruthers, J. M., Johnson, E. R. & McKay, D. B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proceedings of the National Academy of Sciences USA* **97**, 13080–13085.

Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L. A., Offenhauser, N., Marchisio, P. C. & Biffo, S. (2003). Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579–584.

Chakrabarti, A. & Maitra, U. (1992). Release and recycling of eukaryotic initiation factor 2 in the formation of an 80 S ribosomal polypeptide chain initiation complex. *Journal of Biological Chemistry* **267**, 12964–12972.

Chang, E. C. & Schwechheimer, C. (2004). ZOMES III: the interface between signalling and proteolysis. Meeting on the COP9 signalosome, proteasome and eIF3. *EMBO Report* **5**, 1041–1045.

Cho, S. & Hoffman, D. W. (2002). Structure of the beta subunit of translation initiation factor 2 from the archaeon *Methanococcus jannaschii*: a representative of the eIF2beta/eIF5 family of proteins. *Biochemistry* **41**, 5730–5742.

Choi, S. K., Lee, J. H., Zoll, W. L., Merrick, W. C. & Dever, T. E. (1998). Promotion of met-tRNAiMet

binding to ribosomes by yIF2, a bacterial IF2 homolog in yeast. *Science* **280**, 1757–1760.

Choi, S. K., Olsen, D. S., Roll-Mecak, A., Martung, A., Remo, K. L., Burley, S. K., Hinnebusch, A. G. & Dever, T. E. (2000). Physical and functional interaction between the eukaryotic orthologs of prokaryotic translation initiation factors IF1 and IF2. *Molecular and Cellular Biology* **20**, 7183–7191.

Cigan, A. M., Pabich, E. K., Feng, L. & Donahue, T. F. (1989). Yeast translation initiation suppressor sui2 encodes the alpha subunit of eukaryotic initiation factor 2 and shares sequence identity with the human alpha subunit. *Proceedings of the National Academy of Sciences USA* **86**, 2784–2788.

Clemens, M. J., Bushell, M., Jeffrey, I. W., Pain, V. M. & Morley, S. J. (2000). Translation initiation factor modifications and the regulation of protein synthesis in apoptotic cells. *Cell Death and Differentiation* **7**, 603–615.

Cort, J. R., Koonin, E. V., Bash, P. A. & Kennedy, M. A. (1999). A phylogenetic approach to target selection for structural genomics: solution structure of YciH. *Nucleic Acids Research* **27**, 4018–4027.

Craig, A. W., Haghighat, A., Yu, A. T. & Sonenberg, N. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392**, 520–523.

Crechet, J. B. & Parmeggiani, A. (1986). Characterization of the elongation factors from calf brain. 2. Functional properties of EF-1 alpha, the action of physiological ligands and kirromycin. *European Journal of Biochemistry* **161**, 647–653.

Cui, Y., Dinman, J. D., Kinzy, T. G. & Peltz, S. W. (1998). The Mof2/Sui1 protein is a general monitor of translational accuracy. *Molecular and Cellular Biology* **18**, 1506–1516.

Dallas, A. & Noller, H. F. (2001). Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Molecular Cell* **8**, 855–864.

Das, S., Ghosh, R. & Maitra, U. (2001). Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *Journal of Biological Chemistry* **276**, 6720–6726.

Delagoutte, E. & von Hippel, P. H. (2002). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Quarterly Reviews of Biophysics* **35**, 431–478.

Delagoutte, E. & von Hippel, P. H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: integration of helicases into cellular processes. *Quarterly Reviews of Biophysics* **36**, 1–69.

Deo, R. C., Bonanno, J. B., Sonenberg, N. & Burley, S. K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98**, 835–845.

Deo, R. C., Sonenberg, N. & Burley, S. K. (2001). X-ray structure of the human hyperplastic discs protein: an ortholog of the C-terminal domain of poly(A)-binding protein. *Proceedings of the National Academy of Sciences USA* **98**, 4414–4419.

Dever, T. E. (2002). Gene-specific regulation by general translation factors. *Cell* **108**, 545–556.

Dhaliwal, S. & Hoffman, D. W. (2003). The crystal structure of the N-terminal region of the alpha subunit of translation initiation factor 2 (eIF2alpha) from *Saccharomyces cerevisiae* provides a view of the loop containing serine 51, the target of the eIF2alpha-specific kinases. *Journal of Molecular Biology* **334**, 187–195.

Dholakia, J. N. & Wahba, A. J. (1989). Mechanism of the nucleotide exchange reaction in eukaryotic polypeptide chain initiation. Characterization of the guanine nucleotide exchange factor as a GTP-binding protein. *Journal of Biological Chemistry* **264**, 546–550.

Dominguez, D., Altmann, M., Benz, J., Baumann, U. & Trachsel, H. (1999). Interaction of translation initiation factor eIF4G with eIF4A in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry* **274**, 26720–26726.

Donahue, T. F. (2000). Genetic approaches to translation initiation in *Saccharomyces cerevisiae*. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 487–502. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Erickson, F. L., Harding, L. D., Dorris, D. R. & Hannig, E. M. (1997). Functional analysis of homologs of translation initiation factor 2 gamma in yeast. *Molecular and General Genetics* **253**, 711–719.

Erickson, F. L., Nika, J., Rippel, S. & Hannig, E. M. (2001). Minimum requirements for the function of eukaryotic translation initiation factor 2. *Genetics* **158**, 123–132.

Erkmann, J. A. & Kutay, U. (2004). Nuclear export of mRNA: from the site of transcription to the cytoplasm. *Experimental Cell Research* **296**, 12–20.

Feng, L., Yoon, H. & Donahue, T. F. (1994). Casein kinase II mediates multiple phosphorylation of *Saccharomyces cerevisiae* eIF-2 alpha (encoded by SUI2), which is required for optimal eIF-2 function in *S. cerevisiae*. *Molecular and Cellular Biology* **14**, 5139–5153.

Fersht, A. (1998). In *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. New York: W. H. Freeman.

Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W. & Gilmartin, G. M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3′ processing. *Proceedings of the National Academy of Sciences USA* **94**, 11893–11898.

Fleming, K., Ghuman, J., Yuan, X., Simpson, P., Szendroi, A., Matthews, S. & Curry, S. (2003). Solution structure and RNA interactions of the RNA

recognition motif from eukaryotic translation initiation factor 4B. *Biochemistry* **42**, 8966–8975.

Fletcher, C. M., Pestova, T. V., Hellen, C. U. & Wagner, G. (1999). Structure and interactions of the translation initiation factor eIF1. *EMBO Journal* **18**, 2631–2637.

Flynn, A., Oldfield, S. & Proud, C. G. (1993). The role of the beta-subunit of initiation factor eIF-2 in initiation complex formation. *Biochimica et Biophysica Acta* **1174**, 117–121.

Francklyn, C., Perona, J. J., Puetz, J. & Hou, Y. M. (2002). Aminoacyl-tRNA synthetases: versatile players in the changing theater of translation. *RNA* **8**, 1363–1372.

Fraser, C. S., Lee, J. Y., Mayeur, G. L., Bushell, M., Doudna, J. A. & Hershey, J. W. (2004). The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40 S ribosomal subunits in vitro. *Journal of Biological Chemistry* **279**, 8946–8956.

Ganoza, M. C., Kiel, M. C. & Aoki, H. (2002). Evolutionary conservation of reactions in translation. *Microbiology and Molecular Biology Reviews* **66**, 460–485.

Garcia, C., Fortier, P. L., Blanquet, S., Lallemand, J. Y. & Dardel, F. (1995). ^{1}H and ^{15}N resonance assignments and structure of the N-terminal domain of Escherichia coli initiation factor 3. *European Journal of Biochemistry* **228**, 395–402.

Gaspar, N. J., Kinzy, T. G., Scherer, B. J., Humbelin, M., Hershey, J. W. & Merrick, W. C. (1994). Translation initiation factor eIF-2. Cloning and expression of the human cDNA encoding the gamma-subunit. *Journal of Biological Chemistry* **269**, 3415–3422.

Gomez, E., Mohammad, S. S. & Pavitt, G. D. (2002). Characterization of the minimal catalytic domain within eIF2B: the guanine-nucleotide exchange factor for translation initiation. *EMBO Journal* **21**, 5292–5301.

Groft, C. M. & Burley, S. K. (2002). Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization. *Molecular Cell* **9**, 1273–1283.

Gromadski, K. B., Wieden, H. J. & Rodnina, M. V. (2002). Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu. *Biochemistry* **41**, 162–169.

Gross, J. D., Moerke, N. J., von der Haar, T., Lugovskoy, A. A., Sachs, A. B., McCarthy, J. E. & Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115**, 739–750.

Gross, M., Rubino, M. S. & Hessefort, S. M. (1991). The conversion of eIF-2.GDP to eIF-2.GTP by eIF-2B requires Met-tRNA(fMet). *Biochemical and Biophysical Research Communications* **181**, 1500–1507.

Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C. L., Boelens, R. & Gualerzi, C. O. (2000). Mapping the fMet-tRNA(f)(Met) binding site of initiation factor IF2. *EMBO Journal* **19**, 5233–5240.

Gutierrez, P., Osborne, M. J., Siddiqui, N., Trempe, J. F., Arrowsmith, C. & Gehring, K. (2004). Structure of the archaeal translation initiation factor aIF2 beta from *Methanobacterium thermoautotrophicum*: implications for translation initiation. *Protein Science* **13**, 659–667.

Haghighat, A., Mader, S., Pause, A. & Sonenberg, N. (1995). Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. *EMBO Journal* **14**, 5701–5709.

Haghighat, A. & Sonenberg, N. (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5′-cap structure. *Journal of Biological Chemistry* **272**, 21677–21680.

Hashimoto, N. N., Carnevalli, L. S. & Castilho, B. A. (2002). Translation initiation at non-AUG codons mediated by weakened association of eukaryotic initiation factor (eIF) 2 subunits. *Biochemical Journal* **367**, 359–368.

Hawking, S. (1998). *A Brief History of Time*. New York: Bantam.

He, H., von der Haar, T., Singh, C. R., Ii, M., Li, B., Hinnebusch, A. G., McCarthy, J. E. & Asano, K. (2003). The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. *Molecular and Cellular Biology* **23**, 5431–5445.

Hershey, J. W. B. & Merrick, W. C. (2000). The pathway and mechanism of initiation of protein synthesis. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews,), pp. 33–88. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Hershey, P. E., McWhirter, S. M., Gross, J. D., Wagner, G., Alber, T. & Sachs, A. B. (1999). The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. *Journal of Biological Chemistry* **274**, 21297–21304.

Hinnebusch, A. G. (2000). Mechanism and regulation of initiator methionyl-tRNA binding to ribosomes. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 185–244. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Hofmann, K. & Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. *Trends in Biochemical Science* **23**, 204–205.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *Journal of Molecular Biology* **233**, 123–138.

Horvath, C. M., Williams, M. A. & Lamb, R. A. (1990). Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. *EMBO Journal* **9**, 2639–2647.

Huang, H. K., Yoon, H., Hannig, E. M. & Donahue, T. F. (1997). GTP hydrolysis controls stringent

selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *Genes & Development* **11**, 2396–2413.

Imataka, H. & Sonenberg, N. (1997). Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Molecular and Cellular Biology* **17**, 6940–6947.

Ishida, T., Doi, M., Ueda, H., Inoue, M. & Scheldrick, G. M. (1988). Specific ring stacking interaction on the tryptophan-7-methylguanine system: comparative crystallographic studies of indole derivatives-7-methylguanine base, nucleoside, and nucleotide complexes. *Journal of the American Chemical Society* **110**, 2286–2294.

Ito, T., Marintchev, A. & Wagner, G. (2004). Solution structure of human initiation factor eIF2alpha reveals homology to the elongation factor eEF1B. *Structure* (*Cambridge*) **12**, 1693–1704.

Jackson, R. J. (2000). Comparative view of initiation site selection mechanisms. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 127–184. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Janssen, G. M. & Moller, W. (1988). Kinetic studies on the role of elongation factors 1 beta and 1 gamma in protein synthesis. *Journal of Biological Chemistry* **263**, 1773–1778.

Kakuta, Y., Tahara, M., Maetani, S., Yao, M., Tanaka, I. & Kimura, M. (2004). Crystal structure of the regulatory subunit of archaeal initiation factor 2B (aIF2B) from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3: a proposed structure of the regulatory subcomplex of eukaryotic IF2B. *Biochemical and Biophysical Research Communications* **319**, 725–732.

Kapp, L. D. & Lorsch, J. R. (2004). GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *Journal of Molecular Biology* **335**, 923–936.

Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. (1996). The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2·5 A resolution. *Nature* **379**, 511–518.

Khaleghpour, K., Kahvejian, A., De Crescenzo, G., Roy, G., Svitkin, Y. V., Imataka, H., O'Connor-McCourt, M. & Sonenberg, N. (2001). Dual interactions of the translational repressor Paip2 with poly(A) binding protein. *Molecular and Cellular Biology* **21**, 5200–5213.

Kim, T. H., Kim, B. H., Yahalom, A., Chamovitz, D. A. & von Arnim, A. G. (2004). Translational regulation via 5′ mRNA leader sequences revealed by mutational analysis of the *Arabidopsis* translation initiation factor subunit eIF3h. *Plant Cell* **16**, 3341–3356.

Kisselev, L., Ehrenberg, M. & Frolova, L. (2003). Termination of translation: interplay of mRNA, rRNAs and release factors? *EMBO Journal* **22**, 175–182.

Kolupaeva, V. G., Lomakin, I. B., Pestova, T. V. & Hellen, C. U. (2003). Eukaryotic initiation factors 4G and 4A mediate conformational changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. *Molecular and Cellular Biology* **23**, 687–698.

Kolupaeva, V. G., Unbehaun, A., Lomakin, I. B., Hellen, C. U. & Pestova, T. V. (2005). Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA* **11**, 470–486.

Koonin, E. V. (1995). Multidomain organization of eukaryotic guanine nucleotide exchange translation initiation factor eIF-2B subunits revealed by analysis of conserved sequence motifs. *Protein Science* **4**, 1608–1617.

Korneeva, N. L., Lamphear, B. J., Hennigan, F. L., Merrick, W. C. & Rhoads, R. E. (2001). Characterization of the two eIF4A-binding sites on human eIF4G-1. *Journal of Biological Chemistry* **276**, 2872–2879.

Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. & Rhoads, R. E. (2000). Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *Journal of Biological Chemistry* **275**, 41369–41376.

Kostrewa, D., D'Arcy, A., Takacs, B. & Kamber, M. (2001). Crystal structures of Streptococcus pneumoniae N-acetylglucosamine-1-phosphate uridyltransferase, GlmU, in apo form at 2·33 A resolution and in complex with UDP-N-acetylglucosamine and Mg(2+) at 1·96 A resolution. *Journal of Molecular Biology* **305**, 279–289.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292.

Kozak, M. (1987). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *Journal of Molecular Biology* **196**, 947–950.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *Journal of Biological Chemistry* **266**, 19867–19870.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187–208.

Kozlov, G., De Crescenzo, G., Lim, N. S., Siddiqui, N., Fantus, D., Kahvejian, A., Trempe, J. F., Elias, D., Ekiel, I., Sonenberg, N., O'Connor-McCourt, M. & Gehring, K. (2004). Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. *EMBO Journal* **23**, 272–281.

Kozlov, G., Siddiqui, N., Coillet-Matillon, S., Trempe, J. F., Ekiel, I., Sprules, T. & Gehring, K. (2002). Solution structure of the orphan PABC domain from *Saccharomyces cerevisiae* poly(A)-binding protein. *Journal of Biological Chemistry* **277**, 22822–22828.

Kozlov, G., Trempe, J. F., Khaleghpour, K., Kahvejian, A., Ekiel, I. & Gehring, K. (2001). Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proceedings of the National Academy of Sciences USA* **98**, 4409–4413.

KRISHNAMOORTHY, T., PAVITT, G. D., ZHANG, F., DEVER, T. E. & HINNEBUSCH, A. G. (2001). Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. *Molecular and Cellular Biology* **21**, 5018–5030.

LAALAMI, S., PUTZER, H., PLUMBRIDGE, J. A. & GRUNBERG-MANAGO, M. (1991). A severely truncated form of translational initiation factor 2 supports growth of *Escherichia coli*. *Journal of Molecular Biology* **220**, 335–349.

LAMPHEAR, B. J., KIRCHWEGER, R., SKERN, T. & RHOADS, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *Journal of Biological Chemistry* **270**, 21975–21983.

LANCASTER, L., KIEL, M. C., KAJI, A. & NOLLER, H. F. (2002). Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell* **111**, 129–140.

LAURSEN, B. S., MORTENSEN, K. K., SPERLING-PETERSEN, H. U. & HOFFMAN, D. W. (2003). A conserved structural motif at the N terminus of bacterial translation initiation factor IF2. *Journal of Biological Chemistry* **278**, 16320–16328.

LEE, J. H., PESTOVA, T. V., SHIN, B. S., CAO, C., CHOI, S. K. & DEVER, T. E. (2002). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proceedings of the National Academy of Sciences USA* **99**, 16689–16694.

LEIBUNDGUT, M., FRICK, C., THANBICHLER, M., BOCK, A. & BAN, N. (2005). Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO Journal* **24**, 11–22.

LEUNG, P., LEE, Y. M., GREENBERG, E., ESCH, K., BOYLAN, S. & PREISS, J. (1986). Cloning and expression of the Escherichia coli glgC gene from a mutant containing an ADPglucose pyrophosphorylase with altered allosteric properties. *Journal of Bacteriology* **167**, 82–88.

LI, D. & ROBERTS, R. (2001). WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cellular and Molecular Life Sciences* **58**, 2085–2097.

LI, L. & WANG, C. C. (2004). Capped mRNA with a single nucleotide leader is optimally translated in a primitive eukaryote, *Giardia iamblia*. *Journal of Biological Chemistry* **279**, 14656–14664.

LI, Q., IMATAKA, H., MORINO, S., ROGERS JR., G. W., RICHTER-COOK, N. J., MERRICK, W. C. & SONENBERG, N. (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Molecular and Cellular Biology* **19**, 7336–7346.

LI, W., BELSHAM, G. J. & PROUD, C. G. (2001). Eukaryotic initiation factors 4A (eIF4A) and 4G (eIF4G) mutually interact in a 1:1 ratio in vivo. *Journal of Biological Chemistry* **276**, 29111–29115.

LI, W. & HOFFMAN, D. W. (2001). Structure and dynamics of translation initiation factor aIF-1A from the archaeon *Methanococcus jannaschii* determined by NMR spectroscopy. *Protein Science* **10**, 2426–2438.

LOMAKIN, I. B., HELLEN, C. U. & PESTOVA, T. V. (2000). Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Molecular and Cellular Biology* **20**, 6019–6029.

LOMAKIN, I. B., KOLUPAEVA, V. G., MARINTCHEV, A., WAGNER, G. & PESTOVA, T. V. (2003). Position of eukaryotic initiation factor eIF1 on the 40S ribosomal subunit determined by directed hydroxyl radical probing. *Genes Development* **17**, 2786–27897.

LUO, Y. & GOSS, D. J. (2001). Homeostasis in mRNA initiation: wheat germ poly(A)-binding protein lowers the activation energy barrier to initiation complex formation. *Journal of Biological Chemistry* **276**, 43083–43086.

MAAG, D. & LORSCH, J. R. (2003). Communication between eukaryotic translation initiation factors 1 and 1A on the yeast small ribosomal subunit. *Journal of Molecular Biology* **330**, 917–924.

MANCHESTER, K. L. (2001). Catalysis of guanine nucleotide exchange on eIF2 by eIF2B: can it be both a substituted enzyme and a sequential mechanism? *Biochemical and Biophysical Research Communications* **289**, 643–646.

MARCOTRIGIANO, J., GINGRAS, A. C., SONENBERG, N. & BURLEY, S. K. (1997a). Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* **89**, 951–961.

MARCOTRIGIANO, J., GINGRAS, A. C., SONENBERG, N. & BURLEY, S. K. (1997b). X-ray studies of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Nucleic Acids Symposium Series*, 8–11.

MARCOTRIGIANO, J., LOMAKIN, I. B., SONENBERG, N., PESTOVA, T. V., HELLEN, C. U. & BURLEY, S. K. (2001). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Molecular Cell* **7**, 193–203.

MARINTCHEV, A., KOLUPAEVA, V. G., PESTOVA, T. V. & WAGNER, G. (2003). Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: a new interaction between old partners. *Proeedings of the National Academy of Sciences USA* **100**, 1535–1540.

MARISSEN, W. E., GUO, Y., THOMAS, A. A., MATTS, R. L. & LLOYD, R. E. (2000). Identification of caspase 3-mediated cleavage and functional alteration of eukaryotic initiation factor 2alpha in apoptosis. *Journal of Biological Chemistry* **275**, 9314–9323.

MARZI, S., KNIGHT, W., BRANDI, L., CASERTA, E., SOBOLEVA, N., HILL, W. E., GUALERZI, C. O. &

LODMELL, J. S. (2003). Ribosomal localization of translation initiation factor IF2. *RNA* **9**, 958–969.

MATHEWS, M. B., SONENBERG, N. & HERSHEY, J. W. B. (2000). Origins and principles of translational control. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 1–32. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

MATSUO, H., LI, H., MCGUIRE, A. M., FLETCHER, C. M., GINGRAS, A. C., SONENBERG, N. & WAGNER, G. (1997). Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nature Structural Biology* **4**, 717–724.

MAYEUR, G. L., FRASER, C. S., PEIRETTI, F., BLOCK, K. L. & HERSHEY, J. W. (2003). Characterization of eIF3k: a newly discovered subunit of mammalian translation initiation factor elF3. *European Journal of Biochemistry* **270**, 4133–4139.

MAZZA, C., OHNO, M., SEGREF, A., MATTAJ, I. W. & CUSACK, S. (2001). Crystal structure of the human nuclear cap binding complex. *Molecular Cell* **8**, 383–396.

MAZZA, C., SEGREF, A., MATTAJ, I. W. & CUSACK, S. (2002). Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO Journal* **21**, 5548–5557.

MCCARTHY, J. E. (1998). Posttranscriptional control of gene expression in yeast. *Microbiology and Molecular Biology Reviews* **62**, 1492–1553.

MCCUTCHEON, J. P., AGRAWAL, R. K., PHILIPS, S. M., GRASSUCCI, R. A., GERCHMAN, S. E., CLEMONS W. M., JR., RAMAKRISHNAN, V. & FRANK, J. (1999). Location of translational initiation factor IF3 on the small ribosomal subunit. *Proceedings of the National Academy of Sciences USA* **96**, 4301–4306.

MERRICK, W. C. & NYBORG, J. (2000). The protein biosynthesis elongation cycle. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 89–126. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

MEYERS, G. (2003). Translation of the minor capsid protein of a calicivirus is initiated by a novel termination-dependent reinitiation mechanism. *Journal of Biological Chemistry* **278**, 34051–34060.

MOLL, I., GRILL, S., GUALERZI, C. O. & BLASI, U. (2002). Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Molecular Microbiology* **43**, 239–246.

MORENO, J. M., DRSKJOTERSEN, L., KRISTENSEN, J. E., MORTENSEN, K. K. & SPERLING-PETERSEN, H. U. (1999). Characterization of the domains of *E. coli* initiation factor IF2 responsible for recognition of the ribosome. *FEBS Letters* **455**, 130–134.

MORINO, S., IMATAKA, H., SVITKIN, Y. V., PESTOVA, T. V. & SONENBERG, N. (2000). Eukaryotic translation initiation factor 4E (eIF4E) binding site and the middle one-third of eIF4GI constitute the core domain for cap-dependent translation, and the C-terminal one-third functions as a modulatory region. *Molecular and Cellular Biology* **20**, 468–477.

MORRIS-DESBOIS, C., RETY, S., FERRO, M., GARIN, J. & JALINOT, P. (2001). The human protein HSPC021 interacts with Int-6 and is associated with eukaryotic translation initiation factor 3. *Journal of Biological Chemistry* **276**, 45988–45995.

NARANDA, T., KAINUMA, M., MACMILLAN, S. E. & HERSHEY, J. W. (1997). The 39-kilodalton subunit of eukaryotic translation initiation factor 3 is essential for the complex's integrity and for cell viability in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **17**, 145–153.

NEFF, C. L. & SACHS, A. B. (1999). Eukaryotic translation initiation factors 4G and 4A from Saccharomyces cerevisiae interact physically and functionally. *Molecular and Cellular Biology* **19**, 5557–5564.

NEGRUTSKII, B. S. & EL'SKAYA, A. V. (1998). Eukaryotic translation elongation factor 1 alpha: structure, expression, functions, and possible role in aminoacyl-tRNA channeling. *Progress in Nucleic Acid Research and Molecular Biology* **60**, 47–78.

NEGRUTSKII, B. S., STAPULIONIS, R. & DEUTSCHER, M. P. (1994). Supramolecular organization of the mammalian translation system. *Proceedings of the National Academy of Sciences USA* **91**, 964–968.

NIELSEN, K. H., SZAMECZ, B., VALASEK, L., JIVOTOVSKAYA, A., SHIN, B. S. & HINNEBUSCH, A. G. (2004). Functions of eIF3 downstream of 48S assembly impact AUG recognition and GCN4 translational control. *EMBO Journal* **23**, 1166–1177.

NIKA, J., RIPPEL, S. & HANNIG, E. M. (2001). Biochemical analysis of the eIF2beta gamma complex reveals a structural function for eIF2alpha in catalyzed nucleotide exchange. *Journal of Biological Chemistry* **276**, 1051–1056.

NIKA, J., YANG, W., PAVITT, G. D., HINNEBUSCH, A. G. & HANNIG, E. M. (2000). Purification and kinetic analysis of eIF2B from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **275**, 26011–26017.

NONATO, M. C., WIDOM, J. & CLARDY, J. (2002). Crystal structure of the N-terminal segment of human eukaryotic translation initiation factor 2alpha. *Journal of Biological Chemistry* **277**, 17057–17061.

NUDLER, E. & MIRONOV, A. S. (2004). The riboswitch control of bacterial metabolism. *Trends in Biochemical Science* **29**, 11–17.

O'DONNELL, S. M. & JANSSEN, G. R. (2002). Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. *Journal of Bacteriology* **184**, 6730–6733.

OGLE, J. M., BRODERSEN, D. E., CLEMONS JR., W. M., TARRY, M. J., CARTER, A. P. & RAMAKRISHNAN, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902.

Olsen, D. S., Savner, E. M., Mathew, A., Zhang, F., Krishnamoorthy, T., Phan, L. & Hinnebusch, A. G. (2003). Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. *EMBO Journal* **22**, 193–204.

Palacios, I. M., Gatfield, D., St Johnston, D. & Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* **427**, 753–757.

Panniers, R., Rowlands, A. G. & Henshaw, E. C. (1988). The effect of Mg^{2+} and guanine nucleotide exchange factor on the binding of guanine nucleotides to eukaryotic initiation factor 2. *Journal of Biological Chemistry* **263**, 5519–5525.

Paulin, F. E., Campbell, L. E., O'Brien, K., Loughlin, J. & Proud, C. G. (2001). Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Current Biology* **11**, 55–59.

Pause, A., Belsham, G. J., Gingras, A. C., Donze, O., Lin, T. A., Lawrence Jr., J. C. & Sonenberg, N. (1994a). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5′-cap function. *Nature* **371**, 762–767.

Pause, A., Methot, N., Svitkin, Y., Merrick, W. C. & Sonenberg, N. (1994b). Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. *EMBO Journal* **13**, 1205–1215.

Peabody, D. S. & Berg, P. (1986). Termination-reinitiation occurs in the translation of mammalian cell mRNAs. *Molecular and Cellular Biology* **6**, 2695–2703.

Pestova, T. V. & Hellen, C. U. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes & Development* **17**, 181–186.

Pestova, T. V. & Kolupaeva, V. G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes & Development* **16**, 2906–2922.

Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I. & Hellen, C. U. (2001). Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences USA* **98**, 7029–7036.

Pestova, T. V., Lomakin, I. B., Lee, J. H., Choi, S. K., Dever, T. E. & Hellen, C. U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* **403**, 332–335.

Petrelli, D., LaTeana, A., Garofalo, C., Spurio, R., Pon, C. L. & Gualerzi, C. O. (2001). Translation initiation factor IF3: two domains, five functions, one mechanism? *EMBO Journal* **20**, 4560–4569.

Petrushenko, Z. M., Budkevich, T. V., Shalak, V. F., Negrutskii, B. S. & El'skaya, A. V. (2002). Novel complexes of mammalian translation elongation factor eEF1A.GDP with uncharged tRNA and aminoacyl-tRNA synthetase. Implications for tRNA channeling. *European Journal of Biochemistry* **269**, 4811–4818.

Pioletti, M., Schlunzen, F., Harms, J., Zarivach, R., Gluhmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A. & Franceschi, F. (2001). Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO Journal* **20**, 1829–3189.

Planta, R. J. & Mager, W. H. (1998). The list of cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Yeast* **14**, 471–477.

Ponting, C. P. (2000). Novel eIF4G domain homologues linking mRNA translation with nonsense-mediated mRNA decay. *Trends in Biochemical Science* **25**, 423–426.

Poyry, T. A., Kaminski, A. & Jackson, R. J. (2004). What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes & Development* **18**, 62–75.

Preiss, T. & Hentze, M. W. (2003). Starting the protein synthesis machine: eukaryotic translation initiation. *Bioessays* **25**, 1201–1211.

Prevot, D., Darlix, J. L. & Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biology of the Cell* **95**, 141–156.

Proud, C. G. (2001). Regulation of eukaryotic initiation factor eIF2B. *Progress in Molecular and Subcellular Biology* **26**, 95–114.

Ptushkina, M., von der Haar, T., Vasilescu, S., Frank, R., Birkenhager, R. & McCarthy, J. E. (1998). Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5′ cap in yeast involves a site partially shared by p20. *EMBO J* **17**, 4798–4808.

Rajkowitsch, L., Vilela, C., Berthelot, K., Ramirez, C. V. & McCarthy, J. E. (2004). Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *Journal of Molecular Biology* **335**, 71–85.

Ramaiah, K. V., Dhindsa, R. S., Chen, J. J., London, I. M. & Levin, D. (1992). Recycling and phosphorylation of eukaryotic initiation factor 2 on 60S subunits of 80S initiation complexes and polysomes. *Proceedings of the National Academy of Sciences USA* **89**, 12063–12067.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572.

Reed, R. (2003). Coupling transcription, splicing and mRNA export. *Current Opinion in Cell Biology* **15**, 326–331.

Rodnina, M. V. & Wintermeyer, W. (2001). Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annual Review of Biochemistry* **70**, 415–435.

Roll-Mecak, A., Alone, P., Cao, C., Dever, T. E. & Burley, S. K. (2004). X-ray structure of translation initiation factor eIF2gamma: implications for tRNA

and eIF2alpha binding. *Journal of Biological Chemistry* **279**, 10634–10642.

Roll-Mecak, A., Cao, C., Dever, T. E. & Burley, S. K. (2000). X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. *Cell* **103**, 781–792.

Roll-Mecak, A., Shin, B. S., Dever, T. E. & Burley, S. K. (2001). Engaging the ribosome: universal IFs of translation. *Trends in Biochemical Science* **26**, 705–709.

Sachs, A. (2000). Physical and functional interactions between the mRNA cap structure and the poly(A) tail. In *Translational Control* (eds. N. Sonenberg, J. W. B. Hershey & M. B. Mathews), pp. 447–466. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Salimans, M., Goumans, H., Amesz, H., Benne, R. & Voorma, H. O. (1984). Regulation of protein synthesis in eukaryotes. Mode of action of eRF, an eIF-2-recycling factor from rabbit reticulocytes involved in GDP/GTP exchange. *European Journal of Biochemistry* **145**, 91–98.

Sang Lee, J., Gyu Park, S., Park, H., Seol, W., Lee, S. & Kim, S. (2002). Interaction network of human aminoacyl-tRNA synthetases and subunits of elongation factor 1 complex. *Biochemical and Biophysical Research Communications* **291**, 158–164.

Satoh, S., Hijikata, M., Handa, H. & Shimotohno, K. (1999). Caspase-mediated cleavage of eukaryotic translation initiation factor subunit 2alpha. *Biochemical Journal* **342**, 65–70.

Scheper, G. C. & Proud, C. G. (2002). Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *European Journal of Biochemistry* **269**, 5350–5359.

Schmitt, E., Blanquet, S. & Mechulam, Y. (2002). The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. *EMBO Journal* **21**, 1821–1832.

Searfoss, A., Dever, T. E. & Wickner, R. (2001). Linking the 3′ poly(A) tail to the subunit joining step of translation initiation: relations of Pab1p, eukaryotic translation initiation factor 5b (Fun12p), and Ski2p-Slh1p. *Molecular and Cellular Biology* **21**, 4900–4908.

Searfoss, A. M. & Wickner, R. B. (2000). 3′ poly(A) is dispensable for translation. *Proceedings of the National Academy of Sciences USA* **97**, 9133–9137.

Sette, M., van Tilborg, P., Spurio, R., Kaptein, R., Paci, M., Gualerzi, C. O. & Boelens, R. (1997). The structure of the translational initiation factor IF1 from *E. coli* contains an oligomer-binding motif. *EMBO Journal* **16**, 1436–1443.

Severini, M., Choli, T., La Teana, A. & Gualerzi, C. O. (1990). Proteolysis of *Bacillus stearothermophilus* IF2 and specific protection by GTP. *FEBS Letters* **276**, 14–16.

Shin, B. S., Maag, D., Roll-Mecak, A., Arefin, M. S., Burley, S. K., Lorsch, J. R. & Dever, T. E. (2002). Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity. *Cell* **111**, 1015–1025.

Siddiqui, N., Kozlov, G., D'Orso, I., Trempe, J. F. & Gehring, K. (2003). Solution structure of the C-terminal domain from poly(A)-binding protein in *Trypanosoma cruzi*: a vegetal PABC domain. *Protein Science* **12**, 1925–1933.

Singh, C. R., He, H., Ii, M., Yamamoto, Y. & Asano, K. (2004a). Efficient incorporation of eukaryotic initiation factor 1 into the multifactor complex is critical for formation of functional ribosomal preinitiation complexes in vivo. *Journal of Biological Chemistry* **279**, 31910–31920.

Singh, C. R., Yamamoto, Y. & Asano, K. (2004b). Physical association of eukaryotic initiation factor (eIF) 5 carboxyl-terminal domain with the lysine-rich eIF2(beta) segment strongly enhances its binding to eIF3. *Journal of Biological Chemistry* **279**, 49644–49655.

Singh, L. P. & Wahba, A. J. (1996). Regulation of protein synthesis in eukaryotic cells by the guanine nucleotide exchange factor and chain initiation factor 2. *SAAS Bulletin, Biochemistry and Biotechnology* **9**, 1–8.

Sonenberg, N. & Dever, T. E. (2003). Eukaryotic translation initiation factors and regulators. *Current Opinion in Structural Biology* **13**, 56–63.

Spahn, C. M., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G. & Frank, J. (2001). Structure of the 80S ribosome from *Saccharomyces cerevisiae*–tRNA-ribosome and subunit-subunit interactions. *Cell* **107**, 373–386.

Srivastava, S., Verschoor, A. & Frank, J. (1992). Eukaryotic initiation factor 3 does not prevent association through physical blockage of the ribosomal subunit-subunit interface. *Journal of Molecular Biology* **226**, 301–304.

Stapulionis, R. & Deutscher, M. P. (1995). A channeled tRNA cycle during mammalian protein synthesis. *Proceedings of the National Academy of Sciences USA* **92**, 7158–7161.

Story, R. M., Li, H. & Abelson, J. N. (2001). Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proceedings of the National Academy of Sciences USA* **98**, 1465–1470.

Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F. & Bourne, Y. (2001). Crystal structure of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridyltransferase bound to acetyl-coenzyme A reveals a novel active site architecture. *Journal of Biological Chemistry* **276**, 11844–11851.

Takata, H., Takaha, T., Okada, S., Takagi, M. & Imanaka, T. (1997). Characterization of a gene cluster for glycogen biosynthesis and a heterotetrameric ADP-glucose pyrophosphorylase from *Bacillus stearothermophilus*. *Journal of Bacteriology* **179**, 4689–4698.

Tenson, T. & Ehrenberg, M. (2002). Regulatory nascent peptides in the ribosomal tunnel. *Cell* **108**, 591–594.

THOMPSON, G. M., PACHECO, E., MELO, E. O. & CASTILHO, B. A. (2000). Conserved sequences in the beta subunit of archaeal and eukaryal translation initiation factor 2 (eIF2), absent from eIF5, mediate interaction with eIF2gamma. *Biochemical Journal* **347**, 703–709.

TOMSIC, J., VITALI, L. A., DAVITER, T., SAVELSBERGH, A., SPURIO, R., STRIEBECK, P., WINTERMEYER, W., RODNINA, M. V. & GUALERZI, C. O. (2000). Late events of translation initiation in bacteria: a kinetic analysis. *EMBO Journal* **19**, 2127–2136.

UCHIUMI, T., HONMA, S., NOMURA, T., DABBS, E. R. & HACHIMORI, A. (2002). Translation elongation by a hybrid ribosome in which proteins at the GTPase center of the *Escherichia coli* ribosome are replaced with rat counterparts. *Journal of Biological Chemistry* **277**, 3857–3862.

UDAGAWA, T., SHIMIZU, Y. & UEDA, T. (2004). Evidence for the translation initiation of leaderless mRNAs by the intact 70S ribosome without its dissociation into subunits in eubacteria. *Journal of Biological Chemistry* **279**, 8539–8546.

UEDA, H., MARUYAMA, H., DOI, M., INOUE, M., ISHIDA, T., MORIOKA, H., TANAKA, T., NISHIKAWA, S. & UESUGI, S. (1991). Expression of a synthetic gene for human cap binding protein (human IF-4E) in Escherichia coli and fluorescence studies on interaction with mRNA cap structure analogues. *Journal of Biochemistry* (*Tokyo*) **109**, 882–889.

UNBEHAUN, A., BORUKHOV, S. I., HELLEN, C. U. & PESTOVA, T. V. (2004). Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. *Genes & Development* **18**, 3078–3093.

VALASEK, L., HASEK, J., NIELSEN, K. H. & HINNEBUSCH, A. G. (2001a). Dual function of eIF3j/Hcr1p in processing 20 S pre-rRNA and translation initiation. *Journal of Biological Chemistry* **276**, 43351–43360.

VALASEK, L., MATHEW, A. A., SHIN, B. S., NIELSEN, K. H., SZAMECZ, B. & HINNEBUSCH, A. G. (2003). The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome in vivo. *Genes & Development* **17**, 786–799.

VALASEK, L., NIELSEN, K. H. & HINNEBUSCH, A. G. (2002). Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. *EMBO Journal* **21**, 5886–5898.

VALASEK, L., NIELSEN, K. H., ZHANG, F., FEKETE, C. A. & HINNEBUSCH, A. G. (2004). Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. *Molecular and Cellular Biology* **24**, 9437–9455.

VALASEK, L., PHAN, L., SCHOENFELD, L. W., VALASKOVA, V. & HINNEBUSCH, A. G. (2001b). Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO Journal* **20**, 891–904.

VALLE, M., ZAVIALOV, A., SENGUPTA, J., RAWAT, U., EHRENBERG, M. & FRANK, J. (2003). Locking and unlocking of ribosomal motions. *Cell* **114**, 123–134.

VINCIGUERRA, P. & STUTZ, F. (2004). mRNA export: an assembly line from genes to nuclear pores. *Current Opinion in Cell Biology* **16**, 285–292.

VITAGLIANO, L., MASULLO, M., SICA, F., ZAGARI, A. & BOCCHINI, V. (2001). The crystal structure of *Sulfolobus solfataricus* elongation factor 1alpha in complex with GDP reveals novel features in nucleotide binding and exchange. *EMBO Journal* **20**, 5305–5311.

VON DER HAAR, T. & MCCARTHY, J. E. (2002). Intracellular translation initiation factor levels in *Saccharomyces cerevisiae* and their role in cap-complex function. *Molecular Microbiology* **46**, 531–544.

WEI, C. C., BALASTA, M. L., REN, J. & GOSS, D. J. (1998). Wheat germ poly(A) binding protein enhances the binding affinity of eukaryotic initiation factor 4F and (iso)4F for cap analogues. *Biochemistry* **37**, 1910–1916.

WEI, Z., ZHANG, P., ZHOU, Z., CHENG, Z., WAN, M. & GONG, W. (2004). Crystal structure of human eIF3k, the first structure of eIF3 subunits. *Journal of Biological Chemistry* **279**, 34983–34990.

WEIEL, J. & HERSHEY, J. W. (1982). The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30 S ribosomal subunits determined by fluorescence polarization. *Journal of Biological Chemistry* **257**, 1215–1220.

WESTERMANN, P., HEUMANN, W., BOMMER, U. A., BIELKA, H., NYGARD, O. & HULTIN, T. (1979). Crosslinking of initiation factor eIF-2 to proteins of the small subunit of rat liver ribosomes. *FEBS Letters* **97**, 101–104.

WESTERMANN, P., NYGARD, O. & BIELKA, H. (1981). Crosslinking of Met-tRNAf to eIF-2 beta and to the ribosomal proteins S3a and S6 within the eukaryotic inhibition complex, eIF-2.GMPPCP.Met-tRNAf.small ribosomal subunit. *Nucleic Acids Research* **9**, 2387–2396.

WILLIAMS, D. D., PRICE, N. T., LOUGHLIN, A. J. & PROUD, C. G. (2001). Characterization of the mammalian initiation factor eIF2B complex as a GDP dissociation stimulator protein. *Journal of Biological Chemistry* **276**, 24697–24703.

WILLIAMS, N. P., HINNEBUSCH, A. G. & DONAHUE, T. F. (1989). Mutations in the structural genes for eukaryotic initiation factors 2 alpha and 2 beta of *Saccharomyces cerevisiae* disrupt translational control of GCN4 mRNA. *Proceedings of the National Academy of Sciences USA* **86**, 7515–7519.

YANG, H. S., JANSEN, A. P., KOMAR, A. A., ZHENG, X., MERRICK, W. C., COSTES, S., LOCKETT, S. J., SONENBERG, N. & COLBURN, N. H. (2003). The transformation suppressor Pdcd4 is a novel eukaryotic translation initiation factor 4A binding protein that inhibits translation. *Molecular and Cellular Biology* **23**, 26–37.

YATIME, L., SCHMITT, E., BLANQUET, S. & MECHULAM, Y. (2004). Functional molecular mapping of archaeal translation initiation factor 2. *Journal of Biological Chemistry* **279**, 15984–15993.

YODER-HILL, J., PAUSE, A., SONENBERG, N. & MERRICK, W. C. (1993). The p46 subunit of eukaryotic initiation factor (eIF)-4F exchanges with eIF-4A. *Journal of Biological Chemistry* **268**, 5566–5573.

YOON, H. J. & DONAHUE, T. F. (1992). The suil suppressor locus in *Saccharomyces cerevisiae* encodes a translation factor that functions during tRNA(iMet) recognition of the start codon. *Molecular and Cellular Biology* **12**, 248–260.

YUSUPOV, M. M., YUSUPOVA, G. Z., BAUCOM, A., LIEBERMAN, K., EARNEST, T. N., CATE, J. H. & NOLLER, H. F. (2001). Crystal structure of the ribosome at 5·5 A resolution. *Science* **292**, 883–896.

YUSUPOVA, G. Z., YUSUPOV, M. M., CATE, J. H. & NOLLER, H. F. (2001). The path of messenger RNA through the ribosome. *Cell* **106**, 233–241.

ZAVIALOV, A. V. & EHRENBERG, M. (2003). Peptidyl-tRNA regulates the GTPase activity of translation factors. *Cell* **114**, 113–122.

ZOLL, W. L., HORTON, L. E., KOMAR, A. A., HENSOLD, J. O. & MERRICK, W. C. (2002). Characterization of mammalian eIF2A and identification of the yeast homolog. *Journal of Biological Chemistry* **277**, 37079–37087.

ZUCKER, F. H. & HERSHEY, J. W. (1986). Binding of *Escherichia coli* protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization. *Biochemistry* **25**, 3682–3690.