Eukaryotic aspects of translation initiation brought into focus

Christopher H. S. Aylett and Nenad Ban

Institute of Molecular Biology and Biophysics, Eidgenössische Technische Hochschule (ETH) Zürich, Otto-Stern-Weg 5, Zürich 8093, Switzerland

In all organisms, mRNA-directed protein synthesis is catalysed by ribosomes. Although the basic aspects of translation are preserved in all kingdoms of life, important differences are found in the process of translation initiation, which is rate-limiting and the most important step for translation regulation. While great strides had been taken towards a complete structural understanding of the initiation of translation in eubacteria, our understanding of the eukaryotic process, which includes numerous eukaryotic-specific initiation factors, was until recently limited owing to a lack of structural information. In this review, we discuss recent results in the field that provide an increasingly complete molecular description of the eukaryotic initiation process. The structural snapshots obtained using a range of methods now provide insights into the architecture of the initiation complex, start-codon recognition by the initiator tRNA and the process of subunit joining. Future advances will require both higher-resolution insights into previously characterized complexes and mapping of initiation factors that control translation on an additional level by interacting only peripherally or transiently with ribosomal subunits.

This article is part of the themed issue 'Perspectives on the ribosome'.

1. Eubacterial translation initiation

At its most basic level, the process of translation initiation encompasses only a few key steps. The mRNA molecule undergoing translation must bind to the mRNA channel on the small ribosomal subunit such that the start-codon is positioned into the P-(peptidyl) decoding site. A charged initiator tRNA, in prokaryotes generally (formyl) methionyl-tRNAi, must then be bound at the aforementioned start-codon, decoding the first triplet of the encoded protein establishing the frame in which the eventual addition of the following tRNA adaptor molecules will be incorporated. Finally, the large ribosomal subunit joins the small subunit to form ribosomal complex with initiator tRNA in the P site of the ribosome primed for the elongation stage of protein synthesis during which protein is synthesized based on the message encoded in the RNA.

Owing to the large investment of energy involved in translation, this entire process is tightly regulated at the stage of initiation, reducing the wasteful expenditure of resources. In eubacteria, translation initiation is facilitated by two universally conserved factors, IFs (initiation factors) 1 (note that this homologue of eubacterial IF1 in eukaryotes is named eIF1A: eIF1 is a separate factor) and 2, and a region of rRNA dubbed the anti-Shine–Dalgarno sequence [1,2]. The position of the first codon within the decoding site is established by base pairing between the Shine–Dalgarno sequence, located upstream of the first codon within the mRNA, and its complement within the ribosomal rRNA [2]. The separation between the two mRNA elements then defines the start site for translation [2]. Recruitment of charged fMet-tRNAi and correct formation of the codon–anticodon triplet pair are monitored by IF1 and IF2. IF1 binds within the A-site, preventing its premature occupation by a further tRNA and favouring mRNA association, while IF2 stabilizes recruited formyl-methionyl-tRNAi within the P-site and remains bound, blocking the initiation process from continuing until the large subunit joins the small subunit. The selection of cognate tRNAi and...
formation of the proper codon–anticodon interaction are kinetically favoured, but reversible, occurring stochastically [1]. Joining of the large subunit leads to the only irreversible step in eubacterial translation initiation, GTP hydrolysis by IF2, releasing GDP and Pi; the factors then depart, allowing translation elongation to begin [1].

2. Challenges in the eukaryotic setting

Whereas prokaryotic transcription and translation frequently occur concurrently, eukaryotic transcription is separated from the process of translation by splicing, mRNA quality control, nuclear export and transcript relocalization. The canonical mRNA processing and export pathways result in the addition of a 5′-m7G cap and a 3′-poly-A tail to the untranslated regions (UTRs) of the mRNA, the first of which acts as the binding site for translation initiation. This enforces the completion of the previous mRNA processing steps, but precludes the correct placement of the cognate start-codon within the decoding site of initiating ribosomal subunits due to the lack of a defined spacing between the cap and the open reading frame. The small subunit must instead pass through the 5′UTR of the mRNA, bridging or sampling structured elements in a process referred to as ‘scanning’ until it recognizes the cognate start-codon where it binds more tightly in order to allow subunit joining and progression into the translation elongation cycle [3,4]. Such a selection process implies a concerted, irrevocable, codon–anticodon recognition event, coupled to conformational changes in the small subunit, which is controlled by the expanded set of regulatory and initiation factors [3].

Of the many eukaryotic initiation factors that are required for translation in vivo, several are ‘core’ factors conserved from the eubacteria or archaea; eIF1, eIF1A (homologous to eubacterial IF1), eIF2 (a heterotrimeric GTPase that binds aminoacyl-tRNAi) and eIF5B (homologous to eubacterial IF2), being essential for proper initiation in all cases. Many others may be dispensed of under certain artificial conditions in vitro (minimal 5′UTRs and saturation of the core initiation factors), implying that they are required for ‘peripheral’ regulatory or organizational roles [5,6] (figure 1). Among these accessory initiation factors are ATP-dependent helicases, such as DHX29 and eIF4a, the roles of which are presumed to lie in unwinding the 5′UTR to provide access for scanning by the small subunit, and also eIF5, which acts as the GTPase-activating protein (GAP) for eIF2 [3]. The most notable, however, are eIF3 and eIF4, both of which are involved in the delivery of other components to the 40S. eIF3 is more tightly associated with the 40S than the mRNA, and vice versa in the case of eIF4. These massive multi-protein complexes have been referred to as ‘scaffolds’, given that they bind many of the other components [4]. Their functions remain only partially understood; however, recent work has shed light on the organization of many of these peripheral factors, both in yeast and in mammalian systems.

3. eIF3 and eIF4: worthy scaffolds

The initiation of translation in eukaryotes in vivo requires two massive scaffolds, eIF3 and eIF4, each consisting of a number of polypeptides, which orchestrate the initiation process on the 40S subunit and on the 5′-m7G cap (and most probably also the 3′-poly-A tail) of the mRNA molecule, respectively [11]. Their role appears to be to ensure the concerted assembly and continued presence of the complete cast of necessary factors for initiation, both on the message and on the ribosome. Insights into the roles of eIF3 and eIF4 as structural scaffolds have been obtained from biochemical experiments, revealing that they co-purify with accessory initiation factors [3,4,12]. These are poly-A binding protein in all eukaryotes and eIF4a and eIF4b in yeast in the case of eIF4 [11], and eIF3j, eIF5, eIF2 and eIF1 in the case of eIF3 [12]. eIF3, in particular, has been implied to form a ‘multifactor-complex’ interacting with the majority of

---

**Figure 1.** Overview of 40S and 60S subunits and the core translation initiation factors. The molecular structures of the ‘core’ protein and RNA factors absolutely required for translation initiation in yeast in vitro are shown in surface representation separated from the 40S and 60S ribosomal subunits; their known binding sites are indicated by dotted arrows. Note that the factors indicated are arranged and populated differently during different stages of translation initiation. Key features of the 40S and 60S subunits that are highlighted in the text are also labelled to provide context for the reader according to the nomenclature developed in reference [7]. The structures shown were adapted from references [8–10].
the other important initiation factors to bind as a single unit to the recycled 40S subunits [12]. Although the exact order of initiation factor recruitment to a recycled 40S subunit is still unclear, it is established that both eIF3 and eIF4 scaffold assemblies are central to the process of translation initiation. eIF4 is known to bind both the cap and, through poly-A binding protein, the tail of the translated mRNA, a pattern suspected to result in the proximity of the translational termination and start sites with obvious consequences for the efficiency of recycling of terminating ribosomes [11]; however, its exact structural arrangement with respect to the 40S remains one of the gaping holes in our knowledge. With the advent of higher-resolution cryo-electron microscopy (EM), it has proved possible, however, to obtain a relatively complete intermediate-resolution description of the organization of eIF3 from both yeast [8,9,13] and mammals [14–16] (figure 2).

eIF3 circumnavigates the entire 40S ribosomal subunit [8,13], recruiting and coordinating essentially every element of the translational apparatus. Two key conserved sub-complexes occupy the vicinity of each end of the mRNA channel on the solvent-exposed surface of the small 40S subunit [8,15], stabilizing mRNA interacting helicases and contacting eIF4, while from both sides of the ribosomal subunit protein extensions of eIF3 interact with and recruit the evolutionarily conserved machinery that operates within the mRNA-binding channel and the decoding site. The complex varies in complexity between six subunits in yeast and 13 subunits at the more massive extreme in mammals [4]. The main difference between the smaller and larger variants of eIF3 is that the mammalian system maintains the full complement of PCI/MPN (Proteasome, COP9, eIF3/Mpr1, Pad1 N-terminal) proteins giving this sub-complex of eIF3 a similar architecture to that of the COP9 signalosome or the proteasome cap [15], whereas in yeast only two of these proteins, eIF3a and 3c, are present. This supports a more versatile regulatory role for the larger structure, providing a variety of different binding sites for proteins and post-translational modifications such as phosphorylation.

The PCI proteins form an arc of elongated α-solenoidal ‘fingers’ that join at a single hub point [8,13,14]. This architecture accounts for the relative ease by which the extension or reduction of the complex, through the addition or subtraction of new α-solenoidal ‘fingers’, is accomplished in different species. The ‘fingertips’ of the PCI/MPN core, comprising the universally conserved eIF3a and eIF3c subunits, contact the reverse of the 40S platform [13,14]. The angle at which they depart the ribosomal subunit varies between organisms, presumably because of the difference in mass, which would entail clashes should the mammalian complex not tilt to a different angle [13]. They occupy a small area on the 40S surface despite their large mass (approx. 600 kDa in mammals), and overlap minimally with the subunit interface, although enough to retain anti-association activity [13,14]. The extended amino-terminal region of eIF3c is known to bind and recruit together eIF1, eIF5 and presumably the eIF2-ternary complex from the platform side of the subunit, where it has been observed extended across the groove running up to the platform adjacent to helix 44 [9]. A further peripheral eIF3 component present in some organisms but not others, eIF3d, forms a flexible link between the PCI/MPN complex of eIF3 and the head of the eukaryotic ribosome in the vicinity of RACK1, which presumably stabilizes the interaction with the 40S subunit and also topologically constrains the mRNA channel [14,15] (figure 2).

The remainder of eIF3 includes an additional absolutely conserved group of proteins, eIF3b, eIF3i and eIF3g, which form a structurally independent sub-complex, and eIF3j, which is a less well-conserved, sometimes dispensable protein that appears to perform a more peripheral function [8,13]. The eIF3big sub-complex is connected to the core PCI/MPN sub-complex through the extended α-helical carboxy-terminal tail of eIF3a. In mammals this region remains unresolved [14]; however, in yeast it has been observed bound across the solvent-exposed surface of the 40S, presumably contributing to the interaction surface with which eIF3 binds the 40S.
subunit [13]. The eIF3a carboxy-terminal tail binds across the surface of eIF3b, the core protein of the eIF3big sub-complex [13,14]. eIF3b itself binds on the solvent-exposed surface of the ribosomal subunit in a cleft between helix 16 and the extended tip of rRNA expansion segment ES6A [13,14]. The amino-terminal domain of eIF3b recruits the sub-stoichiometric peripheral factor eIF3j, while the other members of the sub-complex are localized immediately adjacent to eIF3b protruding away from the ribosomal surface [13,14]. eIF3i is again linked to eIF3b through the extension of an α-helical carboxy-terminal segment across its surface, while the small, weakly structured protein eIF3g lies between the two, bound within this flexible connection [8]. The eIF3big sub-complex has consistently been implicated in scanning, the aforementioned mRNA unwinding process which is supported by the activity of ATP-dependent helicases that denature the secondary structure of the 5′-UTR [4]. Its role in this process has been corroborated by the visualization of a complex formed with DHX29 [14,15], a DExD box helicase required for the translation of certain mRNA transcripts with considerable secondary structure. DHX29 binds to the solvent-exposed face of the ribosomal subunit adjacent to helix h16, while eIF3big cradles both sides of the helical domains of DHX29. The position of eIF3b is essentially unchanged between structures resolved in the presence or absence of DHX29, suggesting that other helicases might be recruited to the preformed initiation complex in a similar manner [13,14] (figure 2).

Notably the amino-terminus of DHX29 forms a small folded domain at the end of an α-helix that occupies identical space to that taken up by eIF3j, suggesting that these two regions might perform similar roles [13,14]. eIF3j itself protrudes onto the subunit interface, occupying space adjacent to the primordially conserved factor eIF1A. As a consequence, although eIF3 does not form extensive contacts with the surface of the 40S subunit, it encircles it by establishing a series of contacts between factors; as mentioned above, the amino-terminal region of eIF3c contacts eIF1 on one side of helix 44 close to the P site on the 40S subunit, whereas the amino-terminal tail of eIF3b contacts eIF3j, which in turn contacts eIF1A located close to the A site on the other side of helix 44. Since the binding of eIF1 and eIF1A is cooperative because of a conformational change in helix 44, binding of the multifactor complex can be destabilized if any of the small factors depart the 40S, which is desirable for the concerted departure of the initiation complex (figure 2).

4. Start-codon selection: to stay or to go

Once mRNA has been incorporated into the 40S subunit, the multifactor initiation complex organized by eIF3 through a bridging interaction with the mRNA/cap/poly-A tail edifice formed by the eIF4 scaffold, the ribosomal complex scans through the 5′-UTR towards the cognate start-codon [3]. Two important but opposing processes must be regulated during this stage of translation initiation; on the one hand, the decoding site must move over the mRNA to sample codon−anticodon pairs until cognate start-codon−anti-start-codon recognition occurs, whereas, on the other hand, start-codon−anti-start-codon recognition must form a stable pairing and prevent further movement of the initiation complex along the mRNA. In this context, the localization of helicases in the vicinity but not immediately within the mRNA entrance site of the 40S subunit [14–16] makes perfect sense. Firstly, they unfold mRNA secondary structures as near as possible to the site in which the start-codon to be recognized lies. Secondly, their processive movement over the mRNA molecule would not exert a substantial force upon the ribosome once it reaches the start site, since mRNA can simply be looped out between the helicase and the mRNA channel.

On the inter-subunit side of the 40S subunit the primordially conserved initiation factors cooperate to recognize and arrest the initiation complex upon start-codon−anticodon recognition before the ribosome scans past its target. Our structural understanding of this initiation intermediate is a montage of initial high- [17,18] and intermediate-resolution [19] crystal structures of small protein factors and tRNAi, fleshed out to completeness through cryo-EM structures spanning a wide range of resolutions [8,9,13–15,19,20]. While eIF1, eIF1A, tRNAi and parts of eIF2a are resolved to a molecular level, several other factors (eIF5, parts of eIF3 on the subunit interface) are of such low resolution that their identity must be considered tentative rather than final (figure 3).

Two small, universally conserved key proteins, eIF1 and eIF1A, bind to the 40S ribosomal subunit on either side of the proofreading helix h44 close to the decoding sites within the mRNA channel [17–19]. eIF1A occupies part of the A-site, while eIF1 occludes part of the P-site. Binding by eIF1A keeps the RNA bases within the decoding site in a flipped-out conformation, facilitating mRNA scanning by minimizing the interaction between the mRNA and rRNA in the mRNA channel [18], and eIF1 protrudes with a conserved loop into the mRNA channel, ready to sense cognate codon−anticodon base pairing [17,19]. The ternary complex, which consists of charged Met-tRNAi in a selective complex in which the acceptor stem is cradled by eIF2, is incorporated directly into the P-site on its recruitment by the multifactor complex [9,14,15,20]. eIF2 is a GTPase, which binds in a state in which GTP has already been hydrolysated. It releases phosphate as a consequence of the stable binding of the tRNAi anticodon to the start-codon. As observed for the scaffolding factors and 40S, eIF2 also makes minimal contact with the 40S, maintaining approximately 34 Å separation from the body of the subunit and passing above eIF1 and eIF1A [9,14,15,20]. The only significant contact is made with the head within the E-site adjacent to the P-site occupied by the tRNA itself [20]. This ensures that the ternary complex moves with the flexible 40S head in relation to the body of the subunit, while the contacts made with eIF1 and eIF1A will vary with head movement [9]. The eIF2 GAP eIF5 was resolved only extremely weakly as a cloud of density abutting eIF2 [20]. This is not unexpected since eIF5 represents the flexible ‘handle’ by which eIF3 may recruit the ternary complex, eIF3 having substantially greater affinity for eIF5 than eIF2, ensuring that the GTPase is active and primed due to its GAP activity. This implies that eIF5 would make only secondary contacts with the 40S subunit and therefore be relatively disordered with respect to it. The eIF3big complex has been weakly resolved in contact with eIF2 on the subunit interface in the presence of mRNA, a position that is considerably different compared with its position on the 40S solvent-exposed surface in complexes in the absence of mRNA, suggesting a conformational change during scanning [9,21]. While there is currently no mechanistic explanation for this observation, it appears plausible that might be involved in the communication of the stable incorporation of the ternary complex and mRNA to other factors bound by the scaffold (figure 3).
The structural results outlined above reveal that the 40S head motion plays a key role in the scanning process. When the head and the associated ternary complex containing tRNAi is tilted back, away from the mRNA channel, sometimes referred to as the ‘open’ state, the ‘latch’ (h18) region found below the beak of the 40S opens more widely, facilitating the entry of mRNA. During active scanning, the head tilts towards the groove closing the latch and positioning the anticodon stem-loop down onto the mRNA codon, which facilitates sampling of the codon–anticodon pairing [9]. Once the start-codon is encountered, the accommodate tRNAi induces a kink in the mRNA molecule, which facilitates sampling of the codon–anticodon pairing [9]. Once the start-codon is encountered, the accommodate tRNAi induces a kink in the mRNA molecule, which facilitates sampling of the codon–anticodon pairing [9]. Once the start-codon is encountered, the accommodate tRNAi induces a kink in the mRNA molecule, which facilitates sampling of the codon–anticodon pairing [9]. Once the start-codon is encountered, the accommodate tRNAi induces a kink in the mRNA molecule, which facilitates sampling of the codon–anticodon pairing [9].

Upon start-codon recognition the multifactor complex departs, during which several distinct mechanistic events occur. It has been suggested that the termination of the activity of the helicases in active scanning might be linked to the start-codon—anti-start-codon recognition event through the movement of the head of the 40S subunit [15]; however, it is clear that the departure of the multifactor complex would lead to dissociation of helicases recruited by it. A mechanism for eIF1 departure is apparent, while release of Pi would also effect the departure of eIF2 and presumably eIF5. For eIF3, the departure of several of its key binding factors from the subunit interface may decrease its affinity for subunit binding, although its presence on the solvent-exposed surface is known to persist stochastically for some time. It is important to note that eIF1A is the only interface-bound factor believed not to be part of the multifactor complex recruited by eIF3, in agreement with its role in recruiting eIF5B and thereby is extremely short-lived and therefore remains unresolved, signalling start-codon recognition to the remainder of the complex and rendering it irrevocable (figure 3).

Figure 3. Approaching a complete structural overview of eukaryotic translation initiation. Our current understanding of the eukaryotic translation initiation pathway is presented through a series of structural snapshots as detailed within the text. Recycled 40S ribosomal subunits may either be repopulated by the full multifactor complex or undergo a series of independent recruitment events, culminating in the recruitment of the ternary complex to form a scanning-competent pre-initiation complex. Handover of mRNA proceeds through the action of eIF4 by an as-yet unresolved process, resulting in scanning of the pre-initiation complex through the 5′UTR in an ATP/helicase-dependent process. The first start-codon—anti-start-codon recognition event in good Kozak sequence context results in pausing of the pre-initiation complex with Met-tRNAi-AUG in the P-site. At this point eIF1 is released from the complex, resulting in the release of Pi by the eIF2-GTPase active site and the concomitant departure of eIF2 and eventually eIF3 and eIF5. The newly exposed platform of the 40S and tail of eIF1A are then free to recruit eIF5B and the 60S subunit, while the A-site is kept clear by the presence of eIF1A. Once Met-tRNAi, eIF1A and the saricin—ricin loop are all bound by eIF5B, GTP hydrolysis is stimulated and eIF1A and eIF5B depart, leaving a translating ribosome with Met-tRNAi in the P-site. Because of the resolution of almost the complete pathway of structures from yeast, this organism has been chosen as the basis for the figure.
promoting the subunit joining that occurs after tRNAi accommodation and the departure of the rest of the multifactor complex components from the subunit interface.

5. Incorporating the large subunit: coming full circle

Following the departure of eIF2, the 60S subunit must be recruited to form the 80S ribosomal complex. This process is undertaken by eIF5B, a partial homologue of eubacterial IF2, but more closely related to archaeal elf5B. The structural details of how elf5B could concurrently check for both the presence of tRNAi and the sarcin–ricin loop in the context of the large ribosomal subunit had only previously been established at low resolution for the partial bacterial homologue, IF2, which functions in a somewhat different manner [22,23]. Two key cryo-EM structures [24,25], in conjunction with an important crystallographic study of the ribosome [26], have revealed the mechanism by which this takes place.

Elf5B is recruited in part by the large subunit, binding in a very similar mode to that adopted by the eukaryotic homologues of the prokaryotic elongation factors, eEF-1 and eEF-2, at the sarcin–ricin loop, in the same position and mode in which these factors support the addition of successive aminoacyl-tRNA molecules. Nevertheless, many detailed questions remain unresolved in these snapshots. The effect of the Kozak consensus sequence around the start-codon, which promotes recognition, still escapes a rigorous structural description, while the means by which the conformational changes at the decoding site are propagated outwards to the periphery, resulting in the release of successive factors, is difficult to describe based on the current structural data. Furthermore, many important initiation factors such as eIF5, sections of eIF2 and the majority of eIF3, are flexibly attached at the periphery of the initiation complex and are therefore relatively poorly resolved in their bound states. Finally, large holes remain in our knowledge on how eIF4 interacts with the scanning ribosome.

### Table 1. Deposited models from sub-nanometre structural studies of ribosomal subunit-bound canonical eukaryotic translation initiation factors.

<table>
<thead>
<tr>
<th>composition</th>
<th>res. (Å)</th>
<th>technique</th>
<th>key insights</th>
<th>last author</th>
<th>year</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>40S/1</td>
<td>3.9</td>
<td>X-ray diffraction</td>
<td>40S/eIF1 activity</td>
<td>Ban</td>
<td>2011</td>
<td>[17]</td>
</tr>
<tr>
<td>60S/6</td>
<td>3.5</td>
<td>X-ray diffraction</td>
<td>60S/eIF6 anti-association</td>
<td>Ban</td>
<td>2011</td>
<td>[10]</td>
</tr>
<tr>
<td>40S/1/1A</td>
<td>3.7</td>
<td>X-ray diffraction</td>
<td>eIF1/1A and h44</td>
<td>Ban</td>
<td>2013</td>
<td>[18]</td>
</tr>
<tr>
<td>40S/1</td>
<td>7.8</td>
<td>X-ray diffraction</td>
<td>mammalian eIF1</td>
<td>Steitz</td>
<td>2013</td>
<td>[19]</td>
</tr>
<tr>
<td>40S/1/1A</td>
<td>7.0</td>
<td>X-ray diffraction</td>
<td>mammalian eIF1/1A</td>
<td>Steitz</td>
<td>2013</td>
<td>[19]</td>
</tr>
<tr>
<td>40S/1A/mRNA tRNAi</td>
<td>7.0</td>
<td>X-ray diffraction</td>
<td>tRNAi/elf1A binding</td>
<td>Steitz</td>
<td>2013</td>
<td>[19]</td>
</tr>
<tr>
<td>80S/5B/tRNAi</td>
<td>6.6</td>
<td>cryo-EM</td>
<td>elf5B subunit joining</td>
<td>Scheres</td>
<td>2013</td>
<td>[24]</td>
</tr>
<tr>
<td>40S/1A mRNA</td>
<td>3.8</td>
<td>cryo-EM</td>
<td>elf1/1A-mRNA</td>
<td>Ramakrishnan</td>
<td>2014</td>
<td>[20]</td>
</tr>
<tr>
<td>40S/1/1A/2/5 tRNAi/mRNA</td>
<td>4.0</td>
<td>cryo-EM</td>
<td>eIF1 release/eIF5</td>
<td>Ramakrishnan</td>
<td>2014</td>
<td>[20]</td>
</tr>
<tr>
<td>40S/1/1A/3</td>
<td>6.5</td>
<td>cryo-EM</td>
<td>yeast eIF3 architecture</td>
<td>Ban</td>
<td>2015</td>
<td>[8]</td>
</tr>
<tr>
<td>40S/1/1A/2/3 tRNAi/mRNA</td>
<td>4.9</td>
<td>cryo-EM</td>
<td>eIF2β/eIF3 WD40</td>
<td>Ramakrishnan</td>
<td>2015</td>
<td>[9]</td>
</tr>
<tr>
<td>40S/1/1A/2/3 tRNAi/mRNA</td>
<td>6.0</td>
<td>cryo-EM</td>
<td>tRNAi movement</td>
<td>Ramakrishnan</td>
<td>2015</td>
<td>[9]</td>
</tr>
<tr>
<td>80S/5B/tRNAi HCV-IRES</td>
<td>8.2</td>
<td>cryo-EM</td>
<td>elf5B–tRNA interaction</td>
<td>Spahn</td>
<td>2015</td>
<td>[25]</td>
</tr>
<tr>
<td>40S/2/3 DHX29/tRNAi</td>
<td>6.0</td>
<td>cryo-EM</td>
<td>mammalian eIF3 architecture</td>
<td>Hashem</td>
<td>2015</td>
<td>[14]</td>
</tr>
</tbody>
</table>

6. Unanswered questions: the past as a prologue

Structural breakthroughs have thrown many aspects of eukaryotic translation initiation into sharp relief within only a few years, revolutionizing our molecular understanding of the process. Initially the constraints of crystallography held back progress on less well-ordered peripheral factors; however, the advent of cryo-EM has seen these barriers give way precipitously (table 1). We now have a truly molecular view into the details of tRNA accommodation and the outward communication of codon–anticodon recognition, as well as a clear map of the peripheral scaffolding formed by eIF3 and associated proteins. Nevertheless, many detailed questions remain unresolved in these snapshots. The effect of the Kozak consensus sequence [27] around the start-codon, which promotes recognition, still escapes a rigorous structural description, while the means by which the conformational changes at the decoding site are propagated outwards to the periphery, resulting in the release of successive factors, is difficult to describe based on the current structural data. Furthermore, many important initiation factors such as eIF5, sections of eIF2 and the majority of eIF3, are flexibly attached at the periphery of the initiation complex and are therefore relatively poorly resolved in their bound states. Finally, large holes remain in our knowledge on how elf4 interacts with the scanning ribosome.
The future appears bright, however, as it should prove possible to resolve many of the most important remaining questions within the field using tools currently available or under development. At a conformational level, while we now have some first hints of the means by which the Kozak sequence might contribute to start-codon recognition within the P-decoding site [20], further high-resolution structural information of complexes with a variety of different sequences could help us to fully understand this process, and such studies are well within the grasp of current techniques. The same approaches will be applicable for investigating the triggering of Pi release on elf1 departure, the mRNA loading complex and remaining, as of yet poorly understood, initiation factor complexes. Structural studies of characterized mutants that ‘freeze’ the pre-initiation complex in a certain conformation will be very important for better understanding the details of the process. Furthermore, while at present we look only at steady-state snapshots of a continuous process, recent developments in temporally resolved sample preparation promise to allow investigations of the short-lived structural states in future through production of cryo-EM samples mixed for a controlled timespan using microfluidics followed by directly spraying the sample onto a grid [28].

One of the abiding problems that remains to be tackled is that of differential local order. In many of the initiation complex structures resolved by cryo-EM, the most important and interesting factors are resolved only to much lower local resolution than the core ribosomal subunit, which is used to orient the structure. For the majority of the structures resolved to date, incorporating crystallographic or other higher-resolution information into a larger model has been an important step. When the reduced local resolution in EM is due to flexible attachment of small factors (less than approx. 150 kDa), as would appear to be the case for elf5, the problem can only be addressed by filtering maps locally to prevent over-interpretation. Owing to developments in detector technology and software, however, larger (more than approx. 150 kDa) regions of weakly ordered density can now be sorted based on conformation, and reconstructed separately after computational subtraction of the main mass of the complex to yield composite structures at higher resolution. This approach should yield substantial dividends in resolving partially ordered factors such as elf2 and elf3. In the case of wholly separated complexes, however, it remains unfeasible. This would appear to be the case for elf4 and will necessitate thinking, quite literally, outside of the box. Tomographic, or hybrid single-particle approaches, in which higher-resolution information is combined with positional information obtained at lower resolution, are becoming capable of reaching atomic resolution, promising a feasible route to structures of less well-organized complexes [29].

While cryo-EM cannot solve every structural problem for us, it has revolutionized, and will continue to revolutionize, the field of translation initiation. The most desirable prospect for the future of the field, and one that is very plausible with recent advances, would be for cryo-EM to become a structural tool for the resolution of the component states of any particular translation initiation assay, providing positional and temporal information and quantification of different conformations in addition to studying pure molecular structure. Great strides have been made in our understanding of eukaryotic translation initiation; however, it may well prove that the most exciting discoveries remain to come.

Authors’ contributions. C.H.S.A. and N.B. jointly conceived of the subject matter, researched and wrote the article, and produced the figures.

Competing interests. The authors declare no competing interests.

Funding. C.H.S.A. was supported by an ETH Zürich Marie Curie postdoctoral fellowship in conjunction with a European Molecular Biology Organization long-term postdoctoral fellowship. N.B. is supported by the European Research Council and by the Swiss National Science Foundation through the National Centre of Excellence in RNA and Disease.

Acknowledgment. The authors thank Daniel Boehringer, Basel Greber, Ahmad Jomaa, Melanie Weisner and in particular Marc Leibundgut for profitable discussions of the subject area.

References


