Recognition of mRNA cap structures by viral and cellular proteins

Pierre Fechter and George G. Brownlee

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Most cellular and eukaryotic viral mRNAs have a cap structure at their 5’ end that is critical for efficient translation. Cap structures also aid in mRNA transport from nucleus to cytoplasm and, in addition, protect the mRNAs from degradation by 5’ exonucleases. Cap function is mediated by cap-binding proteins that play a key role in translational control. Recent structural studies on the cellular cap-binding complex, the eukaryotic translation initiation factor 4E and the vaccinia virus protein 39, suggest that these three evolutionarily unrelated cap-binding proteins have evolved a common cap-binding pocket by convergent evolution. In this pocket the positively charged N7-methylated guanine ring of the cap structure is stacked between two aromatic amino acids. In this review, the similarities and differences in cap binding by these three different cap-binding proteins are discussed. A comparison with new functional data for another viral cap-binding protein – the polymerase basic protein (PB2) of influenza virus – suggests that a similar cap-binding mechanism has also evolved in influenza virus.

Introduction

It is well known that cap structures at the 5’ end of mRNA (Fig. 1) are required for efficient translation of mRNA on the ribosome and to protect mRNA from degradation by cellular 5’ exonucleases. Cap-binding proteins have evolved to bind to this cap structure. In turn, these cap-binding proteins are associated with other binding proteins that regulate binding of the mRNA and the cap-binding proteins themselves. Details of how cap structures are recognized at the molecular level are now available from extensive structural and functional data on different cap-binding proteins derived from unrelated cellular or viral origins.

Two cellular cap-binding complexes are well known. CBC is a cap-binding complex present in the nucleus and consists of two subunits CBP20 (the cap-binding protein) in association with an ancillary protein CBP80. The complex is thought to aid transport of pre-mRNAs through the different maturation processes (for a review see Lewis & Izaurralde, 1997). The eukaryotic translation initiation factor, eIF4E, is another well known cellular cap-binding protein that is essential for the initiation of translation of capped mRNA on the ribosome. eIF4E associates with at least two other proteins, eIF4A (a helicase) and eIF4G (a scaffold protein) to form the eIF4F complex to facilitate cap binding (for reviews see Gingras et al., 1999; von der Haar et al., 2004).

Some viruses, e.g. retroviruses, utilize the host cellular capping mechanism (Coffin, 1990). Most viruses, however, have evolved their own, virus-specific, capping mechanism. They can be divided into two main groups. The first group includes viruses such as poxviruses, coronaviruses, flaviviruses and reoviruses that replicate in the cytoplasm of cells. A well-studied example of such viruses is vaccinia virus, a double-stranded DNA poxvirus (Moss, 2001). Thus VP39 – the vaccinia viral protein 39 – recognizes its own capped mRNAs. After binding to the cap structure, the enzyme acts as a 2’-O-methyltransferase transferring a methyl group to the 2’ carbon atom of the ribose moiety of the first residue of the cap structure (see Fig. 2 and Condit & Niles, 2002; Moss, 2001; Smith et al., 2002). Baculoviruses, like flaviviruses, also appear to have evolved cap-dependent 2’-O-methyltransferases, (Wu & Guarino, 2003) but the mechanism of cap binding by the flavivirus 2’-O-methyltransferase domain may differ from that of VP39 (Egloff et al., 2002). The second group includes negative-strand RNA viruses replicating either in the nucleus, like orthomyxoviruses (e.g. influenza virus) or in the cytoplasm, like bunyaviruses. These viruses have not evolved capping enzymes as such. Instead they have evolved a mechanism to ‘snatch’ the cap structure from host cell mRNAs (Fodor & Brownlee, 2002; Lamb & Krug, 2001). The viral polymerase binds to cellular capped mRNA in infected cells. An endonuclease activity associated with the polymerase then cleaves the capped oligonucleotides, which subsequently act as primers to initiate viral transcription. Fig. 3 illustrates this mechanism for the influenza virus.
virus life cycle (Fodor & Brownlee, 2002; Lamb & Krug, 2001; Portela et al., 1999).

The purpose of this review is to compare and contrast structural and functional data obtained on the molecular mechanism of cap binding by three evolutionary unrelated proteins, CBC, eIF4E and VP39. We also compare the cap-binding mechanism of the other well studied viral cap-binding protein, the PB2 subunit of influenza RNA polymerase, with CBC, eIF4E and VP39. Inside the cell, cap binding is subject to numerous regulatory processes that modulate its efficiency and these are discussed below.

Structural studies on VP39, eIF4E and CBC suggest a common aromatic ‘sandwich’ motif binds cap structures

How do cap-binding proteins discriminate between capped RNAs and non-capped ones? This is of paramount importance to the functions of the specific cap-binding proteins. In 1997, the NMR or crystal structures of complexes between short-cap analogues and two evolutionary unrelated cap-binding proteins were solved. These were the complexes between eIF4E and m7GDP (Marcotrigiano et al., 1997; Matsuo et al., 1997), and between VP39 and m7GDP (Hodel et al., 1997). Because a similar aromatic cap-binding pocket was present in both proteins, this suggested that convergent evolution of the cap-binding pocket had occurred (reviewed by Quiocho et al., 2000).

Fig. 1. Cap structures. Cap 0 structures lack the 2′-O-methyl residue of the ribose attached to bases 1 and 2. Cap 1 structures have a 2′-O-methyl residue at base 1, whilst cap 2 structures have a 2′-O-methyl residues attached to both bases 1 and 2.

Fig. 2. Replication of vaccinia virus (adapted from Moss, 2001). Vaccinia virus enters the cell releasing cores into the cytoplasm. Early mRNAs are then transcribed from the core particle. After DNA replication, late mRNAs are transcribed. Both early and late mRNAs are transcribed by the viral RNA polymerase and capped by a viral-capping enzyme. Viral mRNAs with cap 0 structures are subsequently recognized by the viral 2′-O-methyltransferase, VP39, and 2′-O-methylation occurs. The virus is assembled in the cytoplasm and enveloped with additional membranes. Fusion with the plasma membrane then occurs allowing release of virus. Note that this figure makes no distinction between the different enveloped forms of vaccinia virus, namely the intracellular mature virus and the extracellular enveloped virus (for review see Smith et al., 2002).
The structural description of cap binding was later improved by crystallization of both proteins with longer cap analogues (m’GpppA in the case of eIF4E, a capped hexamer in the case of VP39) (Hodel et al., 1998; Tomoo et al., 2003). More recently, the crystal structure of the complex between the cap analogue m’GpppG and CBC has revealed a similar aromatic cap-binding pocket, adding support to the convergent evolution hypothesis (Calero et al., 2002; Mazza et al., 2002). Specific binding of purine residues is common to another major class of enzymes – the nucleotidyltransferases (e.g. DNA and RNA ligases, RNA capping enzymes). Interestingly, they also recognize the purine residues of ATP or GTP by the formation of a hydrophobic sandwich between a conserved aromatic and a conserved aliphatic residue (Shuman & Lima, 2004).

Here, we will summarize the main features of the interactions described between these different cap-binding proteins and the cap structure. The reader is referred to the detailed structural papers above for a fuller and more complete description of the interactions involved since not all interactions of eIF4E, CBC and VP39 with the cap are mentioned below.

Although the overall structure of the three cap-binding proteins – eIF4E, CBC and VP39 – differ widely due to their evolutionary unrelated origin, the cap-binding pockets are essentially very similar, although there are some differences in the detail (Fig. 4). In addition to the two aromatic amino acids, the pocket comprises an acidic area to accommodate the positively charged π-ring system of the m’G, and a basic area to accommodate the triphosphate moiety of the cap structure. Overall five main features may be identified that contribute to the specific recognition of the m’GpppN of the cap in the three proteins.

First – the most striking feature – there is a pocket which accommodates the m’G aromatic ring stacked between two aromatic residues (W102/W56 in eIF4E, Y20/Y43 in CBC and Y22/F180 in VP39) (Fig. 4). The presence of an aromatic ‘sandwich’ was suggested earlier from crystallographic, spectroscopic and theoretical studies on small molecule models, like tryptophan-containing peptides and/or indole derivatives (Darzynkiewicz & Lonnberg, 1989; Ishida et al., 1988; Ishida & Inoue, 1981; Stolarski et al., 1996; Ueda et al., 1991). The almost perfect alignment of the three aromatic rings, a nearly optimal interplanar...
distance [between 3.2 and 3.6 Å (0.32–0.36 nm)], substantial areas of overlap in the two stacked rings and the delocalization of the positive charge arising from the N7-methylation of the guanine all contribute to strong interactions between the π-electrons of the stacked aromatic rings (Hu et al., 1999). The strength of these specific interactions explains the low affinity of cap-binding proteins for non-methylated cap analogues (>100-fold difference in affinity compared with N7-methylated ones) (Hodel et al., 1997; Izuarralde et al., 1992; Niedzwiecka et al., 2002).

Perturbation of the aromaticity of the sandwich by mutation of the aromatic residues either to Ala (Hodel et al., 1997; Mazza et al., 2001) or to Leu (Morino et al., 1996) decreases both the affinity and specificity for cap structures.

Second – the delocalized positive charge of the m7G purine ring, due to the methylation of N7, it is generally believed, allows both salt bridges and hydrogen bonds to form between N1 and/or N2 of m7G and the carboxylate groups of Glu or Asp residues in an acidic cavity. This cavity contains at least one critical acidic amino acid – E103 in eIF4E, D116 in CBC and E233 in VP39 (Fig. 4). Mutations of these critical acidic residues to Ala decrease affinity as well as specificity (Hodel et al., 1997; Mazza et al., 2002; Morino et al., 1996). There are conflicting results in the literature of the VP39 phenotype resulting from the mutation of E233 to Glu, which lead to different interpretations of the importance of salt bridges in stabilizing cap binding (Hsu et al., 2000; Hu et al., 1999; Quiocio et al., 2000).

Additional contacts are made between O6, N1 and N2 of m7G (see Figs 1 and 4) and other amino acids. Thus, W102 is involved in these additional contacts in the case of eIF4E, although this residue is also involved in the aromatic stacking with m7G – see above. D114, R112 and W115 are involved in these additional contacts in the case of CBC. D182 and water molecules (not shown) make contacts with m7G in the case of VP39.

Third – the triphosphate moiety is bound in a cavity by residues forming salt bridges with the oxygen atoms of the phosphate groups (Fig. 4). R157, K162 and water molecules (not shown) form a binding cavity in the case of eIF4E. In CBC, R135, R127, Y20, Q133 and V134 are involved, while in VP39, R177, E207 and S205 form the cavity.

Fourth – the N7-methyl group of the guanine is stabilized by van der Waals or weak polar interactions with W166 in the case of eIF4E and Y204 in the case of VP39. In CBC, however, no interaction with the N7-methyl group has yet been reported in the crystal structure.

Finally, interactions with base 1 of the cap structure (see Fig. 1) also contribute to binding, although non-specifically, to accommodate any base at that position. Base 1 is stacked with Y138 in CBC (not shown in Fig. 4), or is stabilized by hydrogen bonds and electrostatic interactions with residues T205, S207 and K206 in eIF4E (not shown in Fig. 4). No structural data, however, have been described detailing the interaction of cap-binding proteins with the 2'-O-methyl groups present of bases 1 and 2 in cap 1 and cap 2 structures (see Fig. 1).

Functional studies on the mechanism of cap recognition

Further biophysical and functional studies have proved controversial. Initially binding of m7GTP to eIF4E was studied by stopped flow dynamics and fluorometric titration (Blachut-Okrasinska et al., 2000; Niedzwiecka et al., 2002). A two-step binding mechanism was proposed, the
binding of the triphosphate moiety being the primary anchor of the cap structure, enabling the subsequent specific interactions with m7G. In contrast, Hu et al. (2003) proposed a simple one-step binding mechanism for VP39, suggesting that the binding to the triphosphate was not contributing to the fast rate of association. Binding depended, it was suggested, almost entirely on interactions of m7G with the two aromatic residues in a cation-π sandwich. From the results of the pH-dependence of the association rate constant, Hu et al. (2003) further proposed that the keto tautomer of the cap structure (the keto form is represented in Fig. 1) rather than the enol form is bound.

Interestingly, Hsu and co-workers (Hsu et al., 2000) conducted mutational studies on both VP39 and eIF4E, showing that different combinations of aromatic residues in the aromatic sandwich could support high affinity for cap structures, providing that at least one Tyr or one Trp was present. These results agreed with those obtained later by Mazza and co-workers (Mazza et al., 2001, 2002) on CBC. Thus, there is a specific interaction between the cap and the aromatic side chains of either Tyr or Trp which would not be possible, or somewhat weaker with Phe. It was suggested that the second stacking residue may only provide a flat complementary surface for lower affinity van der Waals interactions with the cap and may thus accommodate any aromatic amino acid. Further studies on different cap-binding proteins are needed to confirm this hypothesis (Hsu et al., 2000). Interestingly, single or double Trp substitutions of the Tyr or Phe residues of the aromatic sandwich in VP39 increased affinity for cap structures, but not for bona fides substrates containing RNA downstream of the cap (Hu et al., 2002). Another protein, 4EHP, with homology to eIF4E, has been described. It has an aromatic sandwich formed from Trp and Tyr but its function is unknown (Rom et al., 1998).

The specificity of the different cap-binding proteins for caps and cap analogues

The specificity of the different cap-binding proteins for cap structures has been addressed by studies of their affinity for a broad set of capped or uncapped substrates, using a variety of approaches. A preliminary comparison of the data obtained by different authors (see below) seems to suggest some differences in the binding mechanisms by the different cap-binding proteins. All three proteins studied (eIF4E, CBC and VP39) had different affinities for m7GTP. CBC had the strongest affinity, in the range of 10 nM (Wilson et al., 1999), while eIF4E had an affinity of about 260–280 nM (Niedzwiecka et al., 2002; Schepers et al., 2002). VP39 had a lower affinity in the micromolar range (Hu et al., 1999). Interestingly the different affinities correlate with the number of hydrogen bonds made by the three proteins with m7GTP. CBC making more hydrogen bonds than eIF4E, and eIF4E more than VP39. Adding 1 nt residue to the cap analogue used in these studies affected the three proteins in different ways. Thus, the affinity of eIF4E for m7GTP or m7GpppG was not significantly increased (Carberry et al., 1989, 1991). Similarly, no significant difference in affinity was detected between VP39 and m7GDP or m7GpppG (Hu et al., 1999). CBC, on the other hand, had a 100-times higher affinity for m7GpppG than for m7GTP (Izaurralde et al., 1992). In the crystal structure it was observed that the second G of the m7GpppG is stacked on Y138 of CBP20. It has recently been confirmed by fluorescence studies that this stacking interaction with the first base of the transcribed RNA enhances the affinity for pyrimidines but not for purines. However, this stacking does not fully account for the extra affinity of m7GpppG compared with m7GTP (S. Cusack, E. Darzynkiewicz and R. Stolarski, personal communication). The relative affinity of longer oligonucleotides (> 20 mer) has also been studied. While VP39 had an almost 100-fold increased affinity (Lockless et al., 1998), the affinity of both eIF4E and CBC increased but to a lesser extent (<6-fold) (Goss et al., 1990; Mazza et al., 2002; Wilson & Cerione, 2000). The significance of the 2′-O-methylation of base 1 and base 2 in cap binding has, however, not been addressed, in either structural or functional studies.

The relative contribution of the different chemical groups of the m7GTP analogue was tested by binding studies of different derivatives. The important contribution of the triphosphate moiety for binding to eIF4E, but not to VP39, has been shown by comparing the binding of m7GTP, m7GDP, m7GMP or m7G (Cai et al., 1999; Hooker et al., 2003; Hu et al., 1999; Niedzwiecka et al., 2002). The study of other methylated analogues in which guanine was substituted with adenine, cytidine (Hu et al., 1999) or other analogues (Cai et al., 1999; Hooker et al., 2003) showed that different bases, e.g. adenine derivatives, but not acyclic nucleosides, could replace guanine. Both eIF4E and CBC were able to bind guanine substituted by ethyl or benzyl groups at N7 (Cai et al., 1999; Carberry et al., 1990; Darzynkiewicz & Lonnerb, 1989; Izaurralde et al., 1992; Niedzwiecka et al., 2002). The crystal structures of eIF4E and CBC are consistent with these results, as there is sufficient space to accommodate these longer side chains at N7 (Calero et al., 2002; Marrotigiano et al., 1997; Matsuo et al., 1997; Mazza et al., 2002). The binding cavity for N7-methyl, however, is smaller in the case of VP39, suggesting that VP39 would not bind longer side chains at N7, but this remains to be tested (Hodel et al., 1997). Taken together, these different cap analogues have helped establish further details of the specificity of cap-binding proteins for the m7G moiety of the cap structure. Nevertheless, further quantitative data with different analogues of m7GTP are still needed to accurately compare the properties of the different cap-binding proteins.

Cap binding by the PB2 subunit of influenza RNA polymerase, a fourth example of an aromatic sandwich?

Another viral cap-binding protein – the influenza RNA-dependant RNA polymerase – has been the subject of
extensive studies. This polymerase is a heterotrimer, approximately 250 kDa, composed of three subunits, PB1, PB2 and PA. Influenza RNA-dependant RNA polymerase is unable to cap its own mRNAs. Instead it has evolved a specific cap-snatching mechanism that allows it to bypass this step (Fig. 3). Cap snatching can be divided into three steps. First, the polymerase binds the virion RNA promoter, formed by circularization of the 5' and 3' ends of each of the eight influenza RNA segments, thought to exist in a local secondary structure called a ‘corkscrew’ (Flick et al., 1996; Leaby et al., 2001). The polymerase is likely to undergo a conformational change on binding of the virion RNA promoter (Cianci et al., 1995; Li et al., 1998). Second, the cap structure of host mRNAs is specifically recognized by the PB2 subunit, but only in the context of the trimeric holoenzyme (Blaas et al., 1982; Bouloy et al., 1979; Ulmanen et al., 1981). Third, the bound host mRNA is cleaved 9–17 nt from the cap structure by an endonuclease activity of the polymerase complex (Plotch et al., 1981), an activity requiring all three subunits of the RNA polymerase (Fodor et al., 2002; Li et al., 2001; Shi et al., 1995, 1996).

The specificity of cap binding by the PB2 subunit of the influenza RNA polymerase has been studied extensively. The RNA polymerase is essentially inactive in transcription, but not in replication (Vreede et al., 2004), in the absence of a cap structure (Bouloy et al., 1979). Omission of the N7-methyl group of the guanine induces a 5–20-fold loss in binding, depending on the substrate (Bouloy et al., 1980; Hooker et al., 2003). This decrease in binding affinity is less, however, than observed for eIF4E, CBC or VP39, which are typically >100-fold (see above). Thus, the contribution of aromatic stacking interactions to cap binding in the case of PB2 is probably less than in eIF4E, CBC or VP39. Influenza RNA polymerase is the only cap-binding complex for which the contribution of 2'-O-methyl group in cap binding has been shown to be critical, since deletion of this group led to >10-fold decrease in cap binding (Bouloy et al., 1980; Brownlee et al., 1995). The length of the capped RNAs is also crucial for cap binding. m7GpppG is a poor binder (Ulmanen et al., 1981), and 4 nt seem to be the minimal length for efficient binding (Chung et al., 1994). Binding efficiency is increased further if 12–15 nt-long capped oligonucleotides were used (Chung et al., 1994; Fechter et al., 2003). Hooker et al. (2003) compared the affinity of influenza polymerase and eIF4E for different cap analogues. Based on a comparison of the binding of acycloguanosine and guanosine to PB2 and to eIF4E, they argued that ribose was perhaps less significant for the specificity of binding for PB2 than for eIF4E. This work allowed the design of the first influenza inhibitor (Fig. 5) that has been described in the literature that is specific for cap binding (Hooker et al., 2003).

The region(s) within the PB2 subunit of the influenza RNA polymerase involved in cap binding have also been studied in detail. Early studies showed that cap binding was a function of PB2, although all three subunits (PB1, PB2 and PA) of the RNA polymerase were required for cap binding (Blaas et al., 1982; Bouloy et al., 1979; Ulmanen et al., 1981). More recently, however, it has emerged that PB1 is also required, in addition to PB2, for efficient UV cross-linking of cap-containing oligonucleotides longer than four residues in length. This suggests that PB1 cooperates closely with PB2 in binding the capped oligonucleotides (Fechter et al., 2003).

Several studies have attempted to define the region of PB2 involved in cap binding. (i) Honda et al. (1999) studied the 32P-labelled, V8 protease peptides of PB2 derived by UV cross-linking of the influenza ribonucleoprotein complex to a m7G32ppp-labelled capped oligonucleotide. They concluded that residues 242–282 and a second region 538–577, were involved. (ii) Li et al. (2001) extended this approach by UV cross-linking a 4 thioU-containing, 32P-labelled, capped oligonucleotide. A peptide, SVLVNTYQWIIRNW (residues 544–557) was identified after V8 protease digestion. Mutation of W552 to Ala reduced cap binding to 25 % of wild-type levels. Given the proximity of the 4 thioU residue (at residue 2 of the oligonucleotide) to the labelled cap structure, it was proposed that the isolated peptide must be close to the aromatic sandwich. The authors concluded that one or other of the nearby aromatic residues, i.e. W337, Y550, W557 or W564, form the aromatic sandwich with the cap structure, similar to eIF4E or VP39. However, they excluded position W552 as being directly involved in the aromatic sandwich, because the observed 25 % activity of the A552 mutant in cap binding is inconsistent with direct binding. No direct evidence for the involvement of W537, Y550, W557 or W564 was possible because Ala mutants at these positions in PB2 failed to assemble into a functional complex. (iii) Fechter et al. (2003) adopted a systematic mutation approach of all evolutionary conserved aromatic residues of PB2 identified in an alignment between influenza A, B, C and Thogotovirus. Twenty-seven aromatic residues were mutated

![Fig. 5. Structure of a specific inhibitor of influenza cap binding (from Hooker et al., 2003, with permission).](https://example.com/figure5.png)
either to an Ala or to a conservative amino acid (W, Y or F) on the assumption that aromatics directly involved in cap binding would retain cap-binding activity if a conservative mutation were introduced, but would lose binding if an Ala mutation was present. F363 and F404, it was proposed, were directly involved in cap binding in an aromatic sandwich.

The contradictory results (above) suggest there are design flaws in the different approaches. Although it may be argued that cross-linking is a more direct approach than the mutagenesis approach, the length of the oligonucleotide probe used (14 nt), the distance of the 4 thioU residue from the cap [which could be 10–15 Å (1–1.5 nm)] and the unknown folding of PB2 could result in the labelling of a peptide distant from the aromatic pocket. Equally, the mutagenesis approach may fail because some conservative mutations may interfere in cap binding and some Ala mutations may interfere with the formation of a functional polymerase complex. A potential criticism of the emphasis placed on aromatic residues, W537, Y550, W557 or W564 by Li et al. (2001) arises since these residues are not completely conserved as aromatics between influenza A, B, C and Thogotovirus. It seems likely that once recognition of cap structures by PB2 had evolved in a progenitor influenza virus, this feature would have been retained by natural selection in the subsequent evolution of influenza viruses. There remain other discrepancies. First, Fechter et al. (2003) failed to confirm the reduced cap-binding efficiency of the Y550A mutant (Li et al., 2001). Second, Honda et al. (1999) reported two regions of PB2 involved in cap binding whereas Li et al. (2001), reported one. Presumably differences in the conditions of UV cross-linking and/or the V8 protease digestion conditions are responsible.

Interestingly, mutagenesis data on elf4E shows that a cap-binding pocket formed from two Phe residues has a lower affinity for caps than one formed from two Trp residues (Hsu et al., 2000). This result is consistent with the suggestion that two Phe residues comprise the cap-binding pocket in influenza RNA polymerase (Fechter et al., 2003) and recent data (Hooker et al., 2003), which imply that the influenza RNA polymerase is about 100-fold less potent at binding cap analogues than elf4E.

Structural studies are obviously needed to confirm that influenza RNA polymerase binds capped RNAs through an aromatic sandwich, to detail further its cap-binding mechanism and to resolve the discrepancies highlighted above. In particular, the nature of the acidic residues predicted to interact with the m^G ring, and the basic residues predicted to be involved in binding the triphosphate linkage of the cap structure remain unknown. Such studies could improve our knowledge of the detailed mechanism of cap binding, and potentially provide new targets for the generation of novel specific antivirals against influenza viruses.

Regulation of cap-binding proteins
Relatively little is known about the intracellular regulation of cap-binding proteins with the exception of elf4E, which has been extensively studied (Clemens, 2001; Kimball, 2001; McKendrick et al., 1999; Proud, 2002; Raught & Gingras, 1999; Sonenberg & Gingras, 1998; reviewed by von der Haar et al., 2004). elf4E together with elf4G (scaffold) and elf4A (helicase) form a large translation initiation complex, known as elf4F (Fig. 6). It is thought that elf4E interacts with mRNA cap structures only as part of the elf4F complex (Haghighat & Sonenberg, 1997). Interaction of elf4G with elf4E induces a conformational change, enhancing the affinity for capped RNAs (Gross et al., 2003; Haghighat & Sonenberg, 1997; Ptushkina et al., 1998; von der Haar et al., 2000). Different proteins can compete with elf4G for elf4E binding, and thus inhibit the formation of the active elf4F initiation complex. Inhibition occurs with elf4E-binding proteins (4E-BP) (Haghighat & Sonenberg, 1997; Lawrence et al., 1997; Lin et al., 1994; Pause et al., 1994; von der Haar et al., 2000) or an elf4E-transporter (Dostie et al., 2000). Marcotrigiano et al. (1999) defined a short amino acid motif in elf4G that is involved in interaction with elf4E. Interestingly this motif was also present in 4E-BP suggesting that molecular mimicry is involved in controlling initiation of translation. The negative regulation of translation mediated by 4E-BP can be reversed by phosphorylation in response to extracellular stimuli (Kimball, 2001; Lin et al., 1994; Pause et al., 1994). Other translation stimulatory proteins, like the poly(A)-binding protein (PABP) (Fig. 6), increase mRNA affinity for elf4E by binding to the active elf4F complex, (Borman et al., 2000; Le et al., 1997; Wei et al., 1998). Finally, affinity of elf4E for caps is partially inhibited by phosphorylation of S209 of elf4E (Flynn & Proud, 1995; Joshi et al., 1995), correlating with the stage in the cell cycle (reviewed by Kleijn et al., 1998; Proud, 2002). Phosphorylation decreases affinity of elf4E for mRNAs, possible because of electrostatic repulsion between the negatively charged phosphate group attached to S209 and the triphosphate group of the cap structure (Zuberek et al., 2003). This may allow recycling of the elf4F complex at the level of the ribosome (Scheper & Proud, 2002; Scheper et al., 2002).
Regulation of mRNA binding by the nuclear cap-binding complex, CBC, is less well understood than eIF4E. CBC is a dimer composed of CBP20 and CBP80, CBP20 being the cap-binding protein. Though it does not dissociate from CBP20, CBP80 plays a similar role to that of eIF4G. Indeed, it serves as a scaffold for CBP20, enhancing its affinity for capped RNAs (Calero et al., 2002). CBP80 may also be the site of multiple regulatory steps, affecting affinity for mRNAs. Thus, its N terminus is reported to have a binding site for ε-importin, whose binding results in an increased affinity for cap structures (Calero et al., 2002). Thus, though of quite different evolutionary origins, eIF4F and CBC share, in a general way, similar cap binding and regulatory processes.

Up to now, no regulatory pathways affecting the well studied viral cap-binding proteins (the vaccinia virus, VP39, and the PB2 subunit of influenza RNA polymerase) have been described. However, both viral proteins exist as complexes with other viral proteins or nucleic acids, e.g. VP39 with VP55 in the case of vaccinia virus, and PB2 in association with PB1 and PA and viral RNA, in the case of influenza virus. These associated proteins and nucleic acids are essential cofactors for cap-binding activity in the case of influenza virus. Thus, the PB1 and PA subunits of influenza RNA polymerase along with viral RNA may be regarded as playing a similar ‘scaffold role’ for cap binding, as eIF4G does for eIF4E or CBP80 does for CBP20.

Other regulatory pathways probably remain to be discovered. For example, it is known that capped mRNAs, initially bound by CBC in the nucleus, must be subsequently bound to eIF4F in the cytoplasm. Few details are known about this process, which could be controlled by undiscovered proteins (Lewis & Izaurralde, 1997). Another puzzle is how viral cap-snatching proteins, such as the influenza virus polymerase, compete effectively with cellular cap-binding complexes for mRNAs. This may be a significant problem considering the fact that the affinity of the influenza virus cap binding for mRNAs is much lower than that of cellular cap-binding proteins (Hooper et al., 2003; Niedzwiecka et al., 2002; Scheper et al., 2002; Wilson et al., 1999).

Concluding remarks

Many cap-binding proteins, other than CBC, eIF4E and VP39 or the PB2 subunit of the influenza virus RNA polymerase, must have evolved in viruses that replicate in the cytoplasm, e.g. coronaviruses, flaviviruses and reoviruses. These viruses need to have their own capping enzymes and probably O-methylases if they are to effectively ‘take-over’ the host-cell translation machinery, unless they have evolved specialist IRES sequences or other mechanisms that bypass the need for caps. We predict that an aromatic sandwich motif will be found in newly discovered O-methylases, such as has recently been described in the SARS coronavirus. Nevertheless, there is a need to experimentally verify such predictions by studying the 3-dimensional structure of these cap-recognizing proteins in the future and comparing their structure with the structures already known for eIF4E, CBP and VP39. Only armed with this knowledge will it be possible to fully assess the subtleties that have evolved in cap-binding proteins. Moreover, such information allows the rational design of a new class of antiviral agents that targets viral cap-binding proteins whilst avoiding cap-binding proteins, such as eIF4E and CBP, needed for normal cellular function.

Acknowledgements

P.F. and G.G.B. were supported by an MRC component grant (G9901312) and an MRC programme grant (G9523972). The authors gratefully acknowledge the assistance of Drs Elspeth Garman and Susan Lea for preparing Fig. 4, and Drs Ervin Fodor, Stephen Burley, Stephen Cusack, Klaus Klumpp, Stewart Shuman and three anonymous referees for critically reviewing the manuscript.

References


Carberry, S. E., Darzynkiewicz, E. & Goss, D. J. (1991). A comparison of the binding of methylated cap analogues to wheat


Wei, C. C., Balasta, M. L., Ren, J. & Goss, D. J. (1998). Wheat germ poly(A) binding protein enhances the binding affinity of eukaryotic


