The Multifaceted Eukaryotic Cap Structure

Jerry Pelletier[1,2,3,4], T. Martin Schmeing[1,4], and Nahum Sonenberg[1,3]

[1]Dept of Biochemistry, [2]Dept of Oncology, [3]Rosalind and Morris Goodman Cancer Research Centre, and [4]Centre de Recherche en Biologie Structurale,
McGill University, Montreal, Quebec, Canada

ABSTRACT

The 5' cap structure is added onto RNA polymerase II transcripts soon after initiation of transcription and modulates several post-transcriptional regulatory events involved in RNA maturation. It is also required for stimulating translation initiation of many cytoplasmic mRNAs and serves to protect mRNAs from degradation. These functional properties of the cap are mediated by several cap binding proteins (CBPs) involved in nuclear and cytoplasmic gene expression steps. The role that CBPs play in gene regulation, as well as the biophysical nature by which they recognize the cap, is quite intricate. Differences in mechanisms of capping as well as nuances in cap recognition, speak to the potential of targeting these processes for drug development. In this review, we focus on recent findings concerning the cap epitranscriptome, our understanding of cap binding by different CBPs, and explore therapeutic targeting of CBP-cap interaction.

INTRODUCTION

The cap is a unique structure positioned at the 5' end of all RNA polymerase (Pol) II transcripts (Fig. 1a). It is present on messenger RNAs (mRNAs), long non-coding (linc) RNAs, many precursor microRNAs, and small nuclear (sn) and nucleolar (sno) RNAs. It is an essential post-transcriptional modification and participates in several facets of RNA biology – including splicing, polyadenylation, RNA transport, mRNA translation and stability. A previous review in this Journal extensively covered several aspects of cap function including the molecular mechanisms of cap binding by effector proteins, the role of the cap in splicing, mRNA export, nonsense-mediated decay (NMD), and translation initiation (Topisirovic, Svitkin, Sonenberg, & Shatkin, 2011). As well, comprehensive reviews on mRNA de-capping addressed the downstream processes of cap removal and mRNA degradation (Grudzien-Nogalska & Kiledjian, 2017) and aspects of viral 5'-end maturation mechanisms (Picard-Jean et al., 2013). We refer the reader to these excellent reviews and herein seek to provide an update on cap structure/activity relationships.

Capping Enzymes: Writers that Impart Function

The three basic cellular enzymatic capping activities are: (i) an RNA 5'-triphosphatase (RTPase) responsible for removing the 5' γ-terminal phosphate group from the first nucleotide of the nascent RNA, (ii) a guanylyltransferase (GTase) responsible for transferring a GMP moiety onto the 5' end of the RNA, and (iii) an RNA N7-guanine methyltransferase (RNMT) responsible for

methylating the terminal guanosine - the end result being the m^7GpppN structure (where N is any nucleoside and m is a methyl group) termed Cap 0 (Fig. 1b).

RNA Guanylyltransferase and 5' Phosphatase (RNGTT). In vertebrates, RNGTT has two catalytic activities, separated into distinct triphosphatase and guanylyltransferase domains. In lower eukaryotes, these activities can be found separated and encoded by different genes. The RTPase domain shares homology to the cysteine phosphatase superfamily where a conserved cysteine residue within the $HC(X)_4R(S/T)$ signature motif is required for cleavage of the β-γ phosphate linkage of the nascent pppN-terminus (Changela, Ho, Martins, Shuman, & Mondragon, 2001). The GTase activity on the other hand utilizes a catalytic lysine residue to cleave the α-β linkage of GTP and forms a covalent intermediate with GMP, which is then transferred to the 5' diphosphate mRNA end to yield GpppN (Chu et al., 2011) (Fig 1b).

RNA Guanylyl-N7 methyltransferase (RNMT). In mammals, RNMT is a separately encoded activity that catalyzes the transfer of a methyl group from S-adenosylmethionine (SAM) to the GpppN mRNA terminus to produce m^7GpppN (Cap 0) and S-adenosylhomocysteine (SAH) (Pillutla, Yue, Maldonado, & Shatkin, 1998; Saha, Schwer, & Shuman, 1999; Tsukamoto, Shibagaki, Niikura, & Mizumoto, 1998) (Fig. 1b). Importin-α interacts with RNMT to promote the specific binding of RNMT to GpppN-termini and this increases MTase activity 10-fold (Wen & Shatkin, 2000). A second interacting protein, RNMT-activating miniprotein (RAM), stimulates the binding of SAM to RNMT, enhances RNA recruitment to RNMT, and increases RNMT cap MTase activity by over four-fold (Gonatopoulos-Pournatzis, Dunn, Bounds, & Cowling, 2011). The RNMT:RAM co-crystal complex, together with biochemical analysis and molecular dynamics simulation, indicate that RAM serves to stabilize components of the RNMT structure to favor substrate binding and to stimulate methyltransferase activity (Varshney et al., 2016).

RNGTT activity is controlled by the c-MYC proto-oncogene (Lombardi, Varshney, Phillips, & Cowling, 2016). MYC promotes recruitment of RNGTT to MYC-target genes and deregulated c-MYC expression increases the dependency of target gene expression on RNGTT. Ectopic overexpression of RNMT in TERT-immortalized human mammary epithelial cells enhances their cellular transformation and inhibition of RNGTT can selectively block survival of high MYC-

expressing cells (Cowling, 2010; Lombardi et al., 2016). These findings show that RNMT activity, and hence cap methylation, is a highly regulated process.

The majority of cytoplasmic-replicating RNA viruses encode their own capping enzymes. The molecular and genetic organization of the enzymatic activities however can differ significantly from host mammalian cell machinery, but in general the mechanism is as described above (Issur, Picard-Jean, & Bisaillon, 2011). However, there are rare examples of non-canonical capping mechanism. In alphaviruses, nsP1 methylates GTP first followed by formation of an m^7GMP-nsP1 adduct. The mRNA 5' triphosphate residue is trimmed by nsP2 to a 5' diphosphate end, to which m^7GMP is transferred to form Cap 0 mRNAs (Ahola & Kaariainen, 1995). In vesicular stomatitis virus (VSV), the L protein reacts with the 5' end of the RNA (pppApApCpA…) to form a covalent complex (L protein-pApApCpA…). A molecule of GDP is subsequently transferred to the viral 5' end to form GpppApApCpA… (Abraham, Rhodes, & Banerjee, 1975; Ogino & Banerjee, 2007).

Cap Methyltransferases (CMTRs). In mammals, the 2'-OH ribose of the first and second transcribed nucleosides are further methylated by mRNA (nucleoside-2'-O)-methyltransferases (CMTRs) to yield $m^7GpppN^1_m$ (Cap 1) and $m^7GpppN^1_mN^2_m$ (Cap 2) structures (Belanger, Stepinski, Darzynkiewicz, & Pelletier, 2010; Werner et al., 2011) (Fig. 1b). Messenger RNAs with Cap 1 and Cap 2 structures are protected from recognition by innate immune sensor and interferon-induced factors, IFITs (Interferon induced protein with tetracopeptide repeats). The IFIT family arose by gene duplication and although conserved in vertebrates their exact number can vary between and within species. Humans and most mammals encode five family members (IFIT1, IFIT1B, IFIT2, IFIT3, IFIT5), which bind to predominantly 5'ppp- and Cap 0-mRNAs to inhibit their translation by blocking ribosome recruitment (Abbas et al., 2017; Hyde & Diamond, 2015). Differences in the 5' end substrate specificities between some IFITs are apparent. For example, IFIT1 and IFIT1B bind to 5' Cap 0-blocked mRNAs with high affinity ($K_{1/2,\,app}$ ~9 – 23 nM). IFIT1 does not bind to RNA with Cap 1 structures whereas IFIT1B does, although quite weakly ($K_{1/2,\,app}$ ~450 nM) (Kumar et al., 2014).

A cap-specific adenosine N6-methyltransferase, CAPAM [aka PCIF1], modifies penultimate adenine residues, yielding $m^7Gpppm^6A_m$ 5' ends (Akichika et al., 2019; Sendinc et al., 2019; Sun, Zhang, Li, Bai, & Yi, 2019). Whereas CAPAM is not essential and there is little impact of CAPAM knockout on cellular proliferation under normal culture conditions, its loss leads to cell growth defects in the presence of oxidative stress, implying a function for $m^7Gpppm^6A_m$ capped mRNAs in this response (Akichika et al., 2019). Ribosome profiling and RNA-seq comparing translational efficiency

profiles from CAPAM knockout and wt HEK293 cells revealed a slight increase in the levels of mRNAs with m^6Am in the absence of CAPAM and documented a decrease in translation of mRNAs with m^7Gpppm^6Am modifications that appeared eIF4E-independent (Akichika et al., 2019). In their analysis of CAPAM function, Sendinc et al. (Sendinc et al., 2019) reported that loss of PCIF1 has no impact on transcription or mRNA stability, but rather m^6Am-containing mRNAs were translated less efficiently than those mRNAs not normally harboring this modification. In contrast, Boulias et al. (Boulias et al., 2019) reported that depleting PCIF1 does not significantly impact mRNA translation but rather is associated with reduced stability of a subset of m^6Am-containing mRNAs. Loss of PCIF1 in the mouse is of minimum consequence - with no effect on viability or fertility and only reductions in body weight noted (Pandey et al., 2020). RNA-seq analysis showed up- or down-regulation of hundreds (in spleen and brain) to thousands (in testis) of transcripts, with pseudogenes comprising a significant proportion of the top 100 upregulated genes. Ribosome profiling also identified a proportion of mRNAs whose translation appeared to be selectively affected, but this appeared to be a secondary effect of PCIF1 loss as the affected mRNAs did not have any particular preference for the transcription start site nucleotide (Pandey et al., 2020). It may be that the disparate effects on mRNA stability and translation noted in these different studies is a consequence of different requirements on PCIF1 or adaptive mechanisms unique among the cells or tissues under study.

The most complex cap structure is cap 4, $m^7Gpppm_2^6A_mA_mC_mm^3U_m$, present in Trypanosoma and Leishmania parasites (Perry, Watkins, & Agabian, 1987). There are three distinct 2'-O-methyltransferases required for these ribose modifications and their targeted ablation has indicated that they are essential. Loss of Cap 3 and Cap 4 ribose methylation leads to reduced translation rates (Zamudio, Mittra, Campbell, & Sturm, 2009).

Transcription-Dependent Capping

In addition to a direct interaction, there exists an intimate crosstalk between early transcription factors and the capping apparatus since the cap is added co-transcriptionally to RNA Pol II transcripts (Ho et al., 1998; McCracken et al., 1997; Yue et al., 1997). Immediately following transcription initiation, a pause is induced upon the recruitment of the negative elongation factor (NELF) to RNA Pol II by the DRB sensitivity-inducing factor, DSIF (Yamaguchi et al., 1999). During this period, the capping enzymes are mustered to the Pol II C-terminal domain (CTD). It is the phosphorylation pattern of the Pol II CTD that dictates recruitment of the capping activities (Komarnitsky, Cho, & Buratowski, 2000; Wen & Shatkin, 1999). Specifically, TFIIH (Cdk7) and Spt5 (a subunit of DSIF)

phosphorylate Ser^5 within the heptad YSPTSPS repeat embedded within the CTD and this serves as the mobilization signal for the capping activities. The capping complex in turn aids in the recruitment of the positive transcription elongation factor (P-TEFb) (Lenasi, Peterlin, & Barboric, 2011; Mandal et al., 2004), which phosphorylates Ser^2 of the Pol II CTD heptad repeat, as well as DSIF and NELF (E. J. Cho, Kobor, Kim, Greenblatt, & Buratowski, 2001; Guiguen et al., 2007; Ivanov, Kwak, Guo, & Gaynor, 2000; Kim & Sharp, 2001), leading to release of NELF and enabling escape from the transcription pause.

CMTR1 is also recruited to the p-Ser^5 RNA Pol II CTD (Haline-Vaz, Silva, & Zanchin, 2008), indicating that Cap 1 formation is a nuclear event. Modification of the second transcribed nucleotide at the 2'-O ribose is undertaken by a different enzyme, CMTR2. CMTR2 is present in both the nucleus and the cytoplasm, making it unclear in which cellular compartment Cap 2 formation occurs. A crystal structure of the CMTR1 catalytic domain complexed with a capped oligoribonucleotide and SAM has revealed that the cellular enzyme interacts with the cap very differently than viral methyltransferase homologs and this could provide the basis for the development of specific anti-viral compounds (Smietanski et al., 2014; Werner et al., 2011). CAPAM also directly associates with the p-Ser^5 Pol II CTD tail (Sendinc et al., 2019).

The Cap Epitranscriptome Landscape

Recent systems level mass spectrometry approaches, CapQuant and CAP-MAP, have been developed for quantitative assessment of the cap epitranscriptome (Galloway et al., 2020; J. Wang et al., 2019). CapQuant combines HPLC enrichment of cap nucleotides with mass spectrometry detection approaches to enable quantitation of atto- to femto-mole levels of RNA cap structures. CAP-MAP directly analyzes oligo d(T) selected, P1 digested mRNA, thus faithfully maintaining the relative amounts of cap variants.

CapQuant was used to analyze the cap epitranscriptome in human acute lymphoblastic leukemia CCRF-SB cells; where it was found that the most common 5' structures were m^7GpppG_m = m^7GpppC_m > $m^7Gpppm^6A_m$ = m^7GpppA_m=m^7Gpppm^6A >> m^7GpppU_m (J. Wang et al., 2019). The same set of structures were present in RNA isolated from murine kidney and liver tissues although there were differences in relative distribution (Fig. 2). Analysis of *S. cerevisiae* caps confirmed the presence of only Cap 0 structures with the abundance distribution being m^7GpppA > m^7GpppG > m^7GpppU = m^7GpppC (Fig. 2).

CAP-MAP has been used to define cap variants containing a penultimate adenosine or guanosine residue (those containing a cytosine or uridine were not analyzed) in murine activated T cells, liver, heart, brain and human HeLa cells (Galloway et al., 2020). Consistent with what was reported with CapQuant, the most frequent variants were $m^7GpppGm$ (30-52% of the total structures analyzed) and m^7Gpppm^6Am (40-60%), followed by $m^7GpppAm$ (4-10%). A small amount of Cap 0 and unmethylated GpppAm/GpppGm structures were also detected (1.6-2.3% of total), and may represent synthesis or turnover intermediates.

Previous studies using antibody-based approaches had proposed the existence of $m^7Gpppm^1A_m$ and m^7Gpppm^1A caps with mRNAs harboring these caps being more efficiently translated (Dominissini et al., 2016; X. Li et al., 2017). However, these structures were not detected by CapQuant in the cells or tissues analyzed and so their existence remains to be confirmed (J. Wang et al., 2019).

When adenosine is the first transcribed nucleotide, methylation at the N6-position was quite frequently detected by CapQuant, and its occurrence was equally present in Cap 0 (m^7Gpppm^6A) and Cap 1 (m^7Gpppm^6Am) structures in mRNAs from CCRF-SB cells (J. Wang et al., 2019). However, the Cap 0 m^7Gpppm^6A structure was quite rare in murine liver tissue and not detected in murine kidney poly(A)$^+$ RNA, where the predominant structures were $m^7GpppGm$ and m^7Gpppm^6Am. When cap distribution was assessed by CAP-MAP, levels of m^7GpppA_m appeared to be only ~1 - 3.9% of the total variant population analyzed in liver, heart and brain (Galloway et al., 2020), consistent with a previous report indicated the distribution of A-containing caps in HEK293 cells to be 92% $m^7Gpppm^6A_m$ and 8% m^7GpppA_m, with no Cap 0 m^7Gpppm^6A caps present (Akichika et al., 2019). CapQuant does not define precursor/product relationships and it may be that the m^7Gpppm^6A Cap 0 structures detected in CCRF-SB cells represent intermediates that accumulate due to limiting CMTR1 activity.

Nucleotide metabolites (nicotinamide adenine dinucleotide (NAD), flavin adenine dinucleotide (FAD), uridine diphosphate glucose (UDP-Glc), uridine diphosphate N-acetylglucosamine (UDP-GlcNAc), and 3'-dephospho-coenzyme A (dpCoA)) are all capable of serving as initiating nucleotides for bacterial RNA polymerase *in vitro* (Bird et al., 2016; Julius & Yuzenkova, 2017). An NAD-captureSeq approach was used to demonstrate the presence of NAD caps in yeast, plant and human cells (Jiao et al., 2017; Walters et al., 2017; Y. Wang et al., 2019). In humans, NAD caps have been reported on pre-mRNA, spliced and polyadenylated mRNAs, and small nucleolar RNAs (snoRNAs) (Jiao et al., 2017; Walters et al., 2017). NAD-capped transcripts in

Arabidopsis have been found in polysomes, suggesting these are actively translated (Y. Wang et al., 2019). Jiao *et al.* (Jiao et al., 2017) however reported that NAD capped mRNA does not support translation but rather that the NAD moiety promotes mRNA decay (Kiledjian, 2018; Wu et al., 2019). For the first time, CapQuant enabled quantification of the levels of metabolites caps (J. Wang et al., 2019). The frequency of metabolic caps (NAD, FAD, UDP-Glc, UDP-GlcNAc) was extremely low relative to canonical cap structures across all analyzed samples: CCRF-SB cells (0.3%), murine liver (4%), murine kidney (5.4%), and yeast (1%) (J. Wang et al., 2019) (Fig. 2). No CoA cap structures were detected. The paucity of metabolic caps in eukaryotes brings into question their role in significantly contributing to overall gene expression, although specialized roles for specific mRNAs cannot be excluded.

The cap structures of several small nuclear RNAs (snRNAs), snoRNAs, telomerase RNA, some viral mRNAs and selenoprotein mRNAs are further modified by trimethylguanosine synthase 1 (TGS1), which methylates the exocyclic N2 of the m^7G cap to generate $m_3^{2,2,7}G$ caps (TMG: trimethylguanosine) (Franke, Gehlen, & Ehrenhofer-Murray, 2008; Monecke, Dickmanns, & Ficner, 2009; Mouaikel, Verheggen, Bertrand, Tazi, & Bordonne, 2002; Wurth et al., 2014). m^7G-capped snRNAs are exported into the cytoplasm where they undergo modification by TGS1. In complex with Sm core proteins, TMG-capped snRNAs are then transported back to the nucleus by snurportin, a transporter that interacts with the TMG structure and Sm proteins, but not m^7G-capped mRNAs (Hamm, Darzynkiewicz, Tahara, & Mattaj, 1990; Huber et al., 1998). Several mRNAs also harbor hypermethylated caps. Sindbis and Semliki Forest virus late mRNAs contain m^7G, $m_2^{2,7}G$, and $m_3^{2,2,7}G$ caps (HsuChen & Dubin, 1976; van Duijn, Kasperaitis, Ameling, & Voorma, 1986). In the case of Semliki Forest virus, mRNAs with hypermethylated caps are not as abundantly associated with polysomes as m^7G-capped mRNA (van Duijn et al., 1986), likely a reflection of the reduced affinity of eIF4E (see below) for hypermethylated caps (W. Liu et al., 2011). During HIV infection, TMG caps are found on unspliced and partially spliced HIV RNAs (Yedavalli & Jeang, 2010). Among cellular mRNAs, selenoprotein mRNAs harbor TMG, and this does not appear to impair their ability to recruit ribosomes and be translated (Wurth et al., 2014).

Cap Binding Proteins: Readers that Mediate Function from Birth to Death

CBP20. A number of cap binding protein mediate the biological effects of the cap. The first cellular cap binding protein to interact with the capped nascent Pol II RNA product is the nuclear cap binding complex (nCBC). This complex consists of a 20 kDa cap-binding protein (CBP 20) and an auxiliary

partner, CBP80. nCBC plays a critical role in RNA biogenesis as it facilitates splicing, polyadenylation, and nuclear export (Topisirovic et al., 2011). It also has been implicated in NMD, an mRNA quality control pathway that ensures mRNAs with premature termination codons are shunted for degradation.

In CBP20, the m^7G residue is sandwiched between two tyrosines (Fig 3a). Guanine specificity is imparted by hydrogen bonding interactions to O6, N1, and N2. Extensive additional contacts are made with the hydroxyls of the ribose and the triphosphate bridge (not shown) explaining the high affinity of the CBC for m^7GpppG (K_d = ~13 nM) (Mazza, Ohno, Segref, Mattaj, & Cusack, 2001; Mazza, Segref, Mattaj, & Cusack, 2002). Binding of the CBC to the cap is a co-operative, induced fit process that requires CPB20 to undergo significant conformational changes (Mazza et al., 2001; Mazza et al., 2002).

eIF4E. eIF4E participates in translation initiation as a subunit of the heterotrimeric complex, eIF4F, in which it directly interacts with the eIF4G subunit. In turn, eIF4G is complexed with the eIF4A DEAD-box RNA helicase. Structural insight from eIF4E:cap complexes demonstrate that only the m^7Gppp moiety of the cap interacts with eIF4E (Marcotrigiano, Gingras, Sonenberg, & Burley, 1997). Binding of eIF4F to the mRNA is further stabilized by eIF4G, which possesses two RNA binding domains (Marcotrigiano et al., 2001; Yanagiya et al., 2009). There is evidence from *in vitro* RNA Bind-n-Seq experiments that yeast eIF4G can preferentially bind to unstructured RNA motifs enriched for oligo-uridine content (Zinshteyn, Rojas-Duran, & Gilbert, 2017). Interaction of poly(A) binding protein (PABP) with eIF4G also stabilizes eIF4F at the cap and leads to circularization of the mRNA – events that have been linked to a significant enhancement of translation (Gallie, 1991; Kahvejian, Svitkin, Sukarieh, M'Boutchou, & Sonenberg, 2005). The role that eIF4A plays in the initiation process has been extensively reviewed – its helicase activity is essential for ribosome recruitment to most mRNAs (Pelletier & Sonenberg, 2019). eIF4A is needed to resolve complex structural features, as well as RNA-bound protein complexes that can sterically block ribosome recruitment, and this requirement is an important feature in defining translation rates observed among cellular mRNAs (Svitkin, Ovchinnikov, Dreyfuss, & Sonenberg, 1996; Svitkin et al., 2001).

The structural basis of eIF4E:cap interaction is well understood at the atomic level. The N7-methyl modification favors an *anti*-conformation and generates a positive charge that mediates cation-π stacking interactions with two tryptophan residues (W56 and W102 in mammalian eIF4E) (Marcotrigiano et al., 1997; Matsuo et al., 1997) (Fig 3b). Additional interactions between E103 and

N1 and N2 of the guanine base, the backbone amide of M101 and O6, as well as between charged amino acids and the triphosphate bridge (not shown) further stabilize the interaction. No interactions are made with the two ribose hydroxyl groups, unlike what is observed in CBP20, which explains why eIF4E can be purified using affinity columns where m^7GDP or m^7GTP are linked to a solid support matrix through the ribose moiety (Edery, Altmann, & Sonenberg, 1988).

eIF4E Homology Protein (4EHP) and La-Related Protein 1 (LARP1). The cap can also serve as an inhibitory feature of translation. 4EHP (aka eIF4E2) functions by interacting with both the cap and proteins tethered to the coding sequence and the 3' UTR of the same mRNA. Since 4EHP does not interact with eIF4G (Joshi, Cameron, & Jagus, 2004; Rosettani, Knapp, Vismara, Rusconi, & Cameron, 2007; Zuberek et al., 2007), this leads to a circularized mRNA template onto which ribosomes are unable to be recruited. This paradigm of cap-mediated inhibition was first described in *Drosophila* to explain the translation repression of *caudal* mRNA by bicoid (which is 3' UTR anchored and bound to 4EHP) during early development (P. F. Cho et al., 2005). A similar repressive mechanism has been reported for miRNA-induced silencing where 4EHP is tethered to the mRNA 3' UTR via 4E-T which in turn is bound to the CCR4-CNOT complex (Chapat et al., 2017). The affinity of 4EHP (in the presence of 4E-T) for the cap is increased and efficiently blocks cap-dependent translation initiation (Chapat et al., 2017). 4EHP has also been implicated in a ribosome-associated quality control pathway designed to protect cells from proteotoxic stress arising from translation of aberrant mRNAs. Stalled, or collided mRNA-bound ribosomes are detected by the GIGYF2-sensor, a protein that recruits 4EHP to inhibit translation initiation (Hickey et al., 2020; Juszkiewicz et al., 2020). Binding of human 4EHP to the m^7G of the cap is similar to eIF4E, with cation-π stacking interactions occurring between tryptophan residue W124 and tyrosine residue Y78 (Marcotrigiano et al., 1997; Matsuo et al., 1997), N1 and N2 interacting with E125, and O6 hydrogen bonds to the backbone amide of M123 (Fig 3c).

Another protein that utilizes the cap as an inhibitory mark, which mediates the profound translational response of mRNAs harboring unique 5' terminal oligopyrimidine (TOP) sequences to mTORC1 is LARP1 (Fonseca, Lahr, Damgaard, Alain, & Berman, 2018; Meyuhas & Kahan, 2015; Tcherkezian et al., 2014). The TOP signature consists of an invariant 5' cytosine as the first transcribed nucleotide followed by a track of 4 – 14 pyrimidines. The cap binding pocket of LARP1 is more extensive than other CBPs and interacts not only with the m^7G terminal residue but also the next four downstream bases. Specific interaction with these bases rationalizes the polypyrimidine specificity in binding to TOP sequences (Lahr et al., 2017). The inhibition of translation mediate by LARP1 is a

consequence of cap sequestration, interference with eIF4E:cap binding and therefore impairment of translation initiation. LARP1 binds to TOP mRNAs when mTORC1 signalling is inhibited. When mTORC1 signalling is stimulated, LARP1 binds the raptor subunit of mTORC1, resulting in its phosphorylation and release from mRNAs to enable their translation (Fonseca et al., 2015; Hong et al., 2017; Philippe, Vasseur, Debart, & Thoreen, 2018). LARP1 interactions with the cap are the result of cation-π stacking with two tyrosine residues (Y883 and Y922) that arise from neighbouring alpha helices and which are not observed in the other cap binding proteins discussed here. Hydrogen bonding between the guanine base and LARP1 is seen with N1 interacting with E886, N2 interacting with E886, and both the side chain and backbone carbonyl of S882 (Fig 3d).

eIF4E3. eIF4E3 is a CBP that, like eIF4E, can also associate with eIF4G (Joshi et al., 2004). In a tethering assay, eIF4E3 can potently stimulate translation in an eIF4G- and eIF4A-dependent manner (Robert, Cencic, Cai, Schmeing, & Pelletier, 2020). These results allude to the presence of additional eIF4F complexes where eIF4E is replaced by eIF4E3 as a cap binding subunit. In eIF4E3, there is only one aromatic amino acid that is appropriately positioned to stack with m^7G. Accordingly, binding of eIF4E3 to m^7GDP or m^7GTP is 10-40 fold weaker relative to eIF4E, bringing into question the ability of eIF4E3 to compete with eIF4E in translation initiation (Osborne et al., 2013). The low levels and tissue-restricted expression (predominantly in heart, lung, and skeletal muscle) (Joshi et al., 2004) of eIF4E3 suggest a more specialized role in translation initiation.

Decapping proteins. Removal of the cap is a critical regulatory step in gene expression that impacts on mRNA stability (Grudzien-Nogalska & Kiledjian, 2017; Kramer & McLennan, 2019). The first discovered, and best characterized decapping enzyme is Dcp2 (Lykke-Andersen, 2002; van Dijk et al., 2002; Z. Wang, Jiao, Carr-Schmid, & Kiledjian, 2002). Dcp2 cleaves the cap between the α-β phosphate bonds to generate m^7GDP and 5' monophosphate-terminated RNA (Lykke-Andersen, 2002; Z. Wang et al., 2002). Both the cap and downstream RNA sequences are required for decapping by Dcp2 (Piccirillo, Khanna, & Kiledjian, 2003). A number of proteins which stimulate decapping function either as general- or mRNA-specific activators; underscoring the complex regulation of this process (Grudzien-Nogalska & Kiledjian, 2017; Kramer & McLennan, 2019). Dcp2 is developmentally regulated, displays tissue-restricted expression, and is expected to exhibit RNA sequence/structure discrimination based on its RNA binding activity (Grudzien-Nogalska & Kiledjian, 2017).

There are additional decapping enzymes. The NUDIX (NUcleoside DIphosphate linked to another moiety X) family of enzymes hydrolyze phosphodiester bonds and several have been implicated in decapping. For example, NUDIX hydrolase (Nudt) 16 is a decapping enzyme that shows selectivity in its repertoire of target mRNA substrates (Lu et al., 2011; Song, Li, & Kiledjian, 2010). Dcp2 and Nudt16 may function in different decay pathways with NMD employing Dcp2, but turnover of ARE-containing transcripts appear more reliant on Nudt16 (Y. Li, Song, & Kiledjian, 2011). Additional NUDIX proteins exhibit RNA decapping activity *in vitro*: Nudt17 and Nudt19 generate m^7GDP and monophosphate-terminated RNA by hydrolyzing the α-β phosphate linkage of the cap, Nudt12 and Nudt15 cleave the β-γ bond of the cap to generate m^7GMP and diphosphate-terminated RNA, and Nudt2 and Nudt3 cleave both α-β and β-γ bonds with similar efficiencies (Song, Bail, & Kiledjian, 2013). Nudt12 has been implicated in decapping NAD-capped RNAs (Wu et al., 2019). This complexity in RNA decapping speaks to an intricate program of gene expression regulation that is poorly understood and remains to be explored in terms of target specificity determinants and regulation.

Cap Hydrolysis. Following removal of the cap from the 5' end of the mRNA, the m^7GpppN residue itself can be further processed. The scavenger decapping enzyme DcpS utilizes a conserved HIT motif to cleave the pyrophosphate linkage within the cap to release m^7GMP (S. W. Liu et al., 2004). Homozygous loss of function mutations in DcpS are associated with Al-Raqad syndrome, a rare disorder characterized by growth retardation, craniofacial anomalies, altered skin coloration, intellectual disability, and neuromuscular defects (Alesi et al., 2018; Ng et al., 2015). Cells from these patients lack detectable DcpS activity (Ng et al., 2015). Whether elevated levels of intracellular cap structures in $DcpS^{-/-}$ cells impact on cellular homeostasis by interfering with the activity of any of the aforementioned CBPs remains an open question.

In yeast, there are two enzymes implicated in the quality control of the capping process. Railp releases pyrophosphate from pppN-terminated RNA (homologs from other organisms can also release pppN and/or GpppN) which are present on transcripts not fully processed by the capping machinery (Jiao et al., 2010; V. Y. Wang, Jiao, Kiledjian, & Tong, 2015; Xiang et al., 2009). A second enzyme, Dxo1p, removes GpppN from the 5' ends of mRNAs when N7 methylation fails to take place (Chang et al., 2012). In contrast, mammalian cells possess one activity implicated in quality control – DXO (aka DOM3Z). DXO cleaves the α-β bonds of pppN-terminated RNA, removes

GpppN residues from partially capped transcripts, and exhibits 5'-3' exonucleolytic activity (Doamekpor, Gozdek, Kwasnik, Kufel, & Tong, 2020; Jiao, Chang, Kilic, Tong, & Kiledjian, 2013). DXO also shows deNADing activity and removes NAD caps to promote decay of the target mRNAs (Jiao et al., 2017). The promiscuity of DXO in binding to RNAs with different 5' ends is explained by the crystal structure of DXO in complex with m^7GDP where there are no stabilizing interactions made with the m^7G moiety (Jiao et al., 2013).

The large number of decapping and cap hydrolysis proteins allude to a complex system established for fine-tuned regulation. Indeed, different activities have been implicated in mRNA decapping followed by 5'-3' degradation, decapping of incomplete mRNA transcripts, and cap turnover. Understanding the regulation of these events and defining potential cross-talk will be important to our understanding of the mRNA decay pathway.

Cap-dependent Viral Gene Expression. Efficient viral replication is dependent on the ability of the mRNA to access the host cellular translation apparatus. Discrimination between mRNAs with Cap 0 versus those having Cap 1 and Cap 2 structures is used by IFITs as a measure to block viral replication (Abbas et al., 2017; Hyde & Diamond, 2015). Consequently, viruses have evolved mechanisms to ensure that their transcripts bypass this defense mechanism. DNA viruses and retroviruses that utilize cellular RNA pol II for transcription of their genomes are co-transcriptionally capped in the same manner as cellular mRNAs. On the other hand, some cytoplasmic replicating viruses encode their own capping enzymes, whereas others (eg, picornaviruses) link a small polypeptide (VPg) to the mRNA 5' end and recruit ribosomes via a cis-acting internal ribosome entry site (IRES) (Jang et al., 1988; Pelletier & Sonenberg, 1988). Caliciviruses and potyviruses also attach a viral-encoded peptide to the mRNA 5' end but in these instances this protein directly recruits eIF4E or eIF4G for translation initiation, functionally replacing the cap structure in the process of ribosome recruitment (Chaudhry et al., 2006; Chung et al., 2014; Goodfellow et al., 2005). All of these are mechanisms to efficiently evade the IFIT-based anti-viral defenses.

Several viral CBPs have been extensively characterized. The vaccinia VP39 protein functions as a 2'-O methyl transferase to convert Cap 0 to Cap 1 (as well as in the maturation of the poly(A) tail) (Gershon, Ahn, Garfield, & Moss, 1991; Schnierle, Gershon, & Moss, 1992). The cap binding cleft of VP39 can accommodate not only the cap structure but also the first three transcribed nucleotides via a series of hydrogen bond and electrostatic interactions (Hodel, Gershon, Shi, & Quiocho, 1996), which explains the ~100-fold increase in affinity observed when capped RNA

oligonucleotides are used as competitors of cap binding, compared to m^7GTP or m^7GpppG (Lockless, Cheng, Hodel, Quiocho, & Gershon, 1998). As with eIF4E, the m^7G ring stacks between two aromatic residues (Y22 and F180) within VP39 (Fig. 3e). Binding is further stabilized by m^7G N2 interactions with E223 and D182, and N1 with E223.

Influenza A and B also encode a CBP, PB2, which is one of three subunits of the flu RNA polymerase complex (which also includes the PA endonuclease). This complex is involved in the "cap snatch" mechanism whereby short capped oligonucleotides are cleaved from host nuclear transcripts and used to prime transcription from the viral minus strands (Plotch, Bouloy, Ulmanen, & Krug, 1981). Influenza A PB2 interacts with the cap through cation-π stacking, but in a manner distinct from that observed for eIF4E, 4EHP, CBP20, and VP39 (Guilligay et al., 2008) (Fig 3f, g). For influenza A PB2, m^7G is stacked between F404 which is held at a 30° tilt and H357 which lies parallel to the base (Guilligay et al., 2008) (Fig 3f). The N1 and N2 interact with E361 and O6 with K376. There are no direct contacts made to the ribose 2' and 3' hydroxyl groups. The lower affinity of influenza A PB2 for cap analogs and relatively poor discrimination between m^7GTP versus GTP (only a 5-fold difference in affinities) is thought to be a consequence of the imperfect stacking with H357. Replacement of H357 by a tryptophan residue increases cap binding affinity (Guilligay et al., 2008) and this modification forms the basis for cap recognition by influenza B PB2 (Xie et al., 2016) (Fig 3g).

Giant viruses of amoebae, initially so called because their physical size allowed them to be visualized by light microscopy, are complex microorganisms whose genomes can encode >100's of proteins and mRNAs. Several members have also been found to encode components of the protein synthesis machinery (F. Schulz et al., 2017). These include a large repertoire of tRNAs, aminoacyl tRNA synthetases, and initiation, elongation, and release factors. Among the initiation factors, homologs of all three eIF4F subunits have been described. Phylogenetic analysis of eIF4E from different giant viruses indicate that eIF4E appears to have been independently acquired by different giant virus family members (Koonin & Yutin, 2019). This parallel acquisition may reflect evolutionary pressure to optimize viral translation independent of host and/or the implementation of measures by which host translation can be modulated during the viral life cycle. Functional studies are now required to determine if these can bind the cap and what role they play in modulating viral and host gene expression.

Targeting Capping and Cap Recognition for Therapeutic Purposes

Targeting the capping process. Cap methylation of a subset of mRNA transcripts has been shown to be stimulated by the MYC and E2F1 transcription factors (Cole & Cowling, 2009). This appears to be, at least in part, mediated by the upregulation of SAHH (S-adenosyl homocysteine hydrolase)[aka AHCY] (Fernandez-Sanchez, Gonatopoulos-Pournatzis, Preston, Lawlor, & Cowling, 2009). Accumulation of SAH during the capping reaction from SAM can result in product feedback inhibition and SAHH alleviates this by converting SAH to homocysteine and adenosine. MYC also increases the recruitment of RNGTT to RNA Pol II to stimulate the capping reaction (Lombardi et al., 2016). Therefore, the capping process appears to be upregulated in tumors with elevated MYC levels and this may represent a tumor-selective vulnerability, and indeed this has been explored. Inhibition of SAHH using an adenosine analogue, tubercidin, inhibited cap methylation and impaired proliferation of Rat1A fibroblasts ectopically expressing c-MYC, compared to control cells (Fernandez-Sanchez et al., 2009). However, a reduction in SAHH activity can have pleiotropic effects as it results in adenosine depletion, activation of the DNA damage response, and cell cycle arrest (Beluzic et al., 2018), so it will be important to link SAHH inhibition and any anti-cancer effects to impairment of cap methylation in future studies. Suppression of RNGTT or RNMT expression by siRNA/shRNAs in murine or human cells has also been explored, but this leads to cell death, attesting to the essential nature of these two gene products, and likely making it difficult to achieve a therapeutic index by targeting these specific activities (Chu & Shatkin, 2008).

Impairing CBP-cap interaction. eIF4E plays a central role in dictating translation responses to changes in MAPK and mTOR signalling. Deregulation of these pathways, as occurs in many cancers, not only leads to an altered cellular proteome but also a dependency on elevated eIF4E levels (and hence eIF4F activity) that in some settings can be MYC-driven. There is thus much interest in targeting eIF4E (and eIF4F) for therapeutic purposes. Strategies that have been explored include suppression of eIF4E expression using antisense oligonucleotides and siRNAs, interfering with eIF4E-eIF4G interaction, blocking eIF4E:cap interaction, and inhibiting upstream mitogen-activated protein kinase (MAPK) signal-integrating kinases (MNK) 1 and 2 (Bhat et al., 2015).

Cap analogs have been extensively used *in vitro* to study CBP-dependent biochemical processes (Stepinski & Darzynkiewicz, 2014). Their use *in cellula* however is hampered by their poor membrane permeability. To overcome this limitation, Wagner and colleagues have developed 4Ei-1, an N7-benzyl GMP tryptamine phosphoamidate pronucleotide (Ghosh et al., 2009). This pro-drug is cell permeable and is converted to N7-benzyl GMP by endogenous histidine triad nucleotide binding proteins

(HINTs). 4Ei-1 was shown to block epithelial-to-mesenchymal transitions in zebrafish (Ghosh et al., 2009), sensitize lung and breast cancer cells to gemcitabine (S. Li et al., 2013) and mesothelial cells to methotrexate (E. Z. Chen et al., 2014). Several lysine and arginine residues in eIF4E interact with the phosphate backbone of cap analogs and this interaction network extends along the entire triphosphate chain, explaining the rank order potency of $m^7GMP<m^7GDP<m^7GTP$ as eIF4E inhibitors (Marcotrigiano et al., 1997; Niedzwiecka et al., 2002; Tomoo et al., 2002). The low potency of 4Ei-1 is thus likely due to the presence of only a single phosphate in the intracellular released drug. As well, the reliance on HINT activity for intracellular prodrug metabolism may lead to variations in inhibition across cell types.

An under-explored opportunity for drug development lies in leveraging information on the shape of the cap-binding pocket of CBPs. A comparison of the cap binding pocket of the seven CBPs that utilize a shared m^7G stacking mechanism for cap binding (ie, eIF4E, CBP20, 4EHP, LARP1, vaccinia VP39, and influenza A and B PB2) shows quite different spatial geometries (Figs 4 and 5). In CBP20, the N7 methyl group is facing towards the exterior of the protein from its binding location and this could accommodate larger substituents (Figs. 4a, 5a). As well, there could be room for accommodating changes to the O6 and N2 positions, but there is little space to accommodate modifications to the 2' and 3' hydroxyl groups of the ribose (Figs 4a, 5a). In eIF4E, there are distinct pockets that can accommodate larger N7 substituents (Figs 4b, 5b) and explains why compounds such as N7-benzy GMP and 7-(*p*-flurobenzyl)-GMP are able to function as potent cap analogs (C. J. Brown, McNae, Fischer, & Walkinshaw, 2007). This feature was explored by Chen et al. (X. Chen et al., 2012) who generated and tested analogs with N7 extensions, obtaining a compound showing a ~50-fold increase in binding affinity for eIF4E, compared to m^7GMP. Unfortunately, selectivity of the developed compound towards eIF4E versus other CBPs was not directly assessed and this compound suffered from poor solubility - precluding its use in cells. In eIF4E and 4EHP, there is no room to accommodate substituents to O6, unlike the scenario for vaccinia VP39 (Figs 4b, c and 5b, c). The structure of cap-bound LARP1 shows that substituents to the exocyclic N2 position are likely to disrupt binding (Figs 4d, 5d). Cap binding to VP39 is not expected to be tolerant of N7 substituents (Figs 4e, 5e). The cap binding pocket of influenza A and B PB2 appears quite restrictive in not being able to tolerate modifications at N7 or O6 (Figs 4f, g and 5f, g). Exploiting the information gleaned from structural studies of cap-bound CBPs holds promise to generate designer cap analogs selective for a specific CBP.

RNA aptamers hold significant potential as biologicals to overcome some of the serendipity associated with small molecule discovery. High affinity and highly selective RNA aptamers for a variety of molecular ligands can be generated by SELEX, an *in vitro* selection approach that entails reiterative cycles of nucleic acid probe partitioning followed by amplification (Tuerk & Gold, 1990). This approach has been used to generate RNA aptamers that bind to immobilized cap structures and could selectively inhibit translation of capped mRNA reporters *in vitro* (Haller & Sarnow, 1997). However, one drawback of this approach is the lack of specificity – the cap binding activity of all CBPs is likely to be affected.

MNK1 and MNK2. In human cells, the Mnk1 and Mnk2 kinases each give rise to two alternatively spliced products differing at their C-termini (the longer "a" and shorter "b" isoforms). The "a" and "b" isoforms both interact with p38MAPK, but only the "a" isoforms are capable of additionally interacting with ERK (achieved through their unique C-termini). Ablation of both Mnk1 and Mnk2 genes in the mouse is not deleterious to general health and demonstrated that they are the only cellular kinases that phosphorylate eIF4E (Ueda, Watanabe-Fukunaga, Fukuyama, Nagata, & Fukunaga, 2004). The MNKs interact with the carboxy-terminus of eIF4G, which promotes phosphorylation of eIF4E on Ser^{209} (Pyronnet et al., 1999; Shveygert, Kaiser, Bradrick, & Gromeier, 2010). Homozygous $Eif4e^{S209A}$ mice do not exhibit gross abnormalities demonstrating that eIF4E phosphorylation is not required for normal cellular homeostasis or development (Furic et al., 2010).

Surface plasmon resonance studies have shown that Ser^{209} phosphorylation decreases the affinity of eIF4E for the cap by enhancing the off-rate (Scheper et al., 2002) whereas other studies have attributed the diminished affinity of phosphorylated eIF4E for the cap to a reduced on-rate (Slepenkov, Darzynkiewicz, & Rhoads, 2006). Thus, eIF4E phosphorylation may decrease the time eIF4E associates with mRNA templates leading to faster recruitment of mRNAs into active polysomes. Increased translational output of a subset of total mRNAs, some of which encode pro-tumorigenic factors, appears to be the consequence of eIF4E S209 phosphorylation (Furic et al., 2010).1

Several MNK inhibitors have been described and their development is driven by the possibility that these may exhibit anti-cancer activity. CGP57380 was first reported by Novartis as an MNK inhibitor but suffers from low potency (EC_{50}'s are in the uM range), off-target activity, and lack of *in vivo* activity. Subsequently, cercosporamide was found to potently inhibit MNK1 (EC_{50} = 116 nM) and MNK2 (EC_{50}=11 nM) *in vitro*, with off-target activity exhibited towards Jak3 (EC_{50}=31 nM) (Konicek

et al., 2011). Oral dosing of cercosporamide in mice revealed suppression of B16 melanoma pulmonary metastasis as well as HCT116 colon cancer xenograft tumors (Konicek et al., 2011). Recently, eFT508 (aka Tomivosertib) was rationally designed based on structural information of the MNK1 and MNK2 ATP binding pocket (Reich et al., 2018). eFT508 is an extremely potent and selective inhibitor (1 -2 nM towards MNK1 and MNK2) and shows efficacy *in vivo* in a number of xenograft cancer models (Reich et al., 2018). At the time of writing, this compound is currently under evaluation in clinical trials in patients with castrate-resistant prostate cancer and in combination with PD-1/PD-L1 inhibitors (eg, NCT03616834, NCT03690141).

tRNA and m^7G.

N-7 methyl guanosine is also present in tRNAs (Tomikawa, 2018). In most instances, the modification occurs predominantly at position 46 in the variable region and is the result of tRNA (m^7G46) methyltransferase activity. This modification is widespread in tRNAs and has been documented in eukaryotes, eubacteria, and some archaea. Whether the m^7G residue is present in other RNA species is not known but can now be addressed by profiling approaches, such as AlkAniline-Seq a chemical approach coupled to deep sequencing that allows detection of m^7G (and m^3C) in RNA (Marchand et al., 2018). An interesting question is whether internal m^7G modification compete with 5' m^7G capped mRNAs in some biological processes or recruit some of the CBPs in moonlighting functions for tRNA biology.

Caps in Synthetic Biology.

There is much interest in developing mRNA expression technologies for stem cell reprogramming, vaccination, and therapeutic protein expression. The use of RNA for these technologies holds several advantages over DNA that include the fact that RNA need not be delivered to the nucleus for expression and there is little concern for insertional mutagenesis resulting from genome integration. However, in contrast to DNA, expression is transient due to the significantly shorter half-life of RNA. Accordingly, several structural aspects of mRNAs have been assessed to search for improvements that could extend half-life and increase expression. One key feature that has been extensively investigated is the nature of the cap structure and its impact on expression.

Currently, three different types of cap analogs are used in mRNA pre-clinical research: (i) m^7GpppG, (ii) anti-reverse cap analogs (ARCAs: $m_2^{7,3'\text{-}O}GpppG$), and iii) modified ARCAs. The ability to synthesize RNA *in vitro* was made possible by the purification of phage RNA polymerases (notably

from SP6, T3, and T7), which proved to be a game changer for RNA structure/function studies (Melton et al., 1984). Soon thereafter, it was shown the m^7GpppG cap analog could be used as primer *in vitro* to initiate transcription (Konarska, Padgett, & Sharp, 1984; Pelletier & Sonenberg, 1985). However, during this process a proportion of the cap analog is mis-incorporated in the reverse orientation to yield Gpppm^{7}G-terminated RNA. To overcome this shortcoming, ARCAs were designed in which the 3'-OH of the m^7G ribose is blocked by a methyl group to prevent its use as priming site during transcription initiation (Fig. 6a). ARCA-modified mRNAs exhibit better translation than mRNAs harboring traditional caps (Grudzien-Nogalska et al., 2007). Improvements on the ARCA design have included the incorporation of 5' phosphorothiolate modifications, which do not impact on eIF4E recognition, but do confer resistance to Dcp2 decapping (Wojtczak et al., 2018).

Additional modifications to the cap aimed at improving stability have been the use of locked nucleic acid (LNA)-modified cap analogues ($m^{7(LNA)}$GpppG), which function similarly to ARCAs in ensuring appropriate cap orientation during mRNA synthesis (Kore, Shanmugasundaram, Charles, Vlassov, & Barta, 2009) (Fig. 6b). mRNAs with boranophosphate analogs ($m_2^{7,\,2'\text{-}O}Gpp_{BH3}pG$; two diasteriomers are possible) increased protein output by 1-7 – 2.2 fold when introduced into dendritic cells, compared to $m_2^{7,3'\text{-}O}$GpppG capped mRNA (Kowalska et al., 2014; Su et al., 2011). Recently mRNAs capped with a dichloromethylene tetraphosphonate bridge were shown to be resistant to decapping by Dcp2 with no impairment of translation efficiency *in vitro* (Rydzik et al., 2017) (Fig 6c).

Numerous cap analogs have been tested as inhibitors of cap-dependent translation and an interesting group of these contain modifications to the exocyclic amine group (Cai et al., 1999; X. Chen et al., 2012; Jia et al., 2010). Aromatic substituents (benzyl (bz), p-methoxybenzyl, and triazole) at N2 were found to exhibit superior inhibitory properties (Kocmik et al., 2018). The incorporation of the benzyl and 4-methoxybenzyl features within the mRNA cap structures was also undertaken to assess the consequences on mRNA function (Kocmik et al., 2018). Although no differences in translational efficiency was noted in *in vitro* experiments (relative to $m_2^{7,3'\text{-}O}$GpppG capped mRNA), when bz-modified capped mRNAs were introduced into HEK293 cells, $bn^2m_2^{7,3'\text{-}O}$GpppG, (p-OCH_3bn)$^2m_2^{7,3'\text{-}O}$GpppG, and $bn^7m_2^{7,3'\text{-}O}$GpppG-capped transcripts out performed controls by 2.4 – 3.3 fold (Kocmik et al., 2018). That the N2 modifications are tolerated can be rationalized from the eIF4E-cap structure where the exocyclic N2 is facing the solvent (Fig. 5b). Unfortunately, these cap modifications rendered the mRNA templates more susceptible to Dcp2 decapping. It will be

interesting to combine these N2 substituents with modifications that protect the cap against Dcp2-mediated cleavage to see if additive effects on expression can be obtained.

Using the Cap to Probe RNA Function

As a specific cross-linkable site, strategies to modify the cap on mRNAs have been developed as a prelude to functional studies (Nowakowska et al., 2014). A variant of the Giardia lamblia trimethylguanosine synthase 2 (GlaTgs2) methylates the N2 position of the cap but is somewhat promiscuous in its substrate specificity. GlaTgs2 has been used to transfer alkyne, azido, and 4-vinylbenzene groups to the N2 guanine position (Holstein, Stummer, & Rentmeister, 2015; Muttach, Muthmann, & Rentmeister, 2017; D. Schulz, Holstein, & Rentmeister, 2013). The derivitized guanines can then be further modified by biorthogonal reactions to append specific functional groups. mRNAs harboring anthranilol-caps have also been synthesized and this modification does not significantly impact translational efficiency (Domashevskiy, Rodriguez, Gunawardana, & Goss, 2016). We await the use of these mRNAs in biophysical and single molecule studies where they could provide a deeper understanding into the mechanism of cap-dependent initiation.

Conclusions and Perspectives.

In this review, we have extended a previously published comprehensive treatise on the cap and CBPs (Topisirovic et al., 2011). Our understanding of the role that the cap, and other 5' end modifications, play in regulating gene expression is far from complete. What is needed are approaches that would enable the visualization of single mRNA molecules from synthesis to degradation, coupled with the ability to directly detect binding of individual CBPs. As well, studies on viral CBPs and cap modifying activities provide exciting opportunities for developing small molecule inhibitors that could selectively block viral replication.

The coordination of promoter utilization and translational output is another area that remains to be explored. Given that transcription initiation of particular genes can occur at multiple sites and that this will lead to the production of mRNAs with different 5' caps and cap-proximal nucleotide context, one (out of many) area that remains to be explored is whether (and if so, how) these subtle differences in transcription initiation site utilization impact on gene expression. In yeast, transcription start site selection has been shown to quantitatively impact on translation activity of different mRNA isoforms encoding the same protein (Rojas-Duran & Gilbert, 2012). Along these lines, Tamarkin-Ben-Harush et al. (Tamarkin-Ben-Harush, Vasseur, Debart, Ulitsky, & Dikstein, 2017) have investigated

the relationship between the identity of the 5'-terminal nucleotide in mRNA and translation in MEFs under normal and glucose starvation conditions. They found no differences in protein expression under normal conditions, but reduced translation of mRNAs containing 5' terminal cytosines under starvation stress. They also documented that alternative promoter usage at the Ser-tRNA synthetase gene, resulting in transcripts differing in only 5 nucleotides, yet these showed dramatically different responses to glucose starvation. Brown et al. (J. D. Brown et al., 2020) have shown that HIV produces 5' capped RNAs with one (^{Cap}1G), two (^{Cap}2G) or three (^{Cap}3G) terminal guanosines and that the 5' end of the ^{Cap}1G transcript adopts a dimeric complex hairpin structure that sequesters the cap from eIF4E, renders the mRNA resistant to decapping, and serves as substrate that is packaged into virions. In contrast, the ^{Cap2}G and ^{Cap3}G mRNAs adopt a different conformation and their caps are recognized by eIF4E. This provides a stunning example of how subtle transcriptional choices of the first transcribed nucleotide can influence RNA function.

The epitranscriptome diversity of the second transcribed nucleotide (Cap 2 modifications) is currently undefined and technological advances are required to obtain a more complete picture of that landscape. From its humble beginnings as "non-nucleoside material (NNM)" uncovered by Yasuhiro Furuichi during his analysis of the 5' end of cytoplasmic polyhedrosis virus (Furuichi, 2015), the cap has provided enormous insight into mechanisms of gene expression regulation. We predict that it will continue to do so, as single molecule and systems-level approaches to understanding regulation are applied to study its function, and as details into the cross-talk between different CBPs emerge.

Funding Information

Support for work in the authors' labs is from a Canada Research Chair to TMS and the Canadian Institutes of Health Research to J.P. (FDN-148366) and NS (FDN-148423).

Graphical Abstract

Cap interplay with gene regulatory mechanisms associated with RNA biology.

Figure Legends

Figure 1. RNA cap and capping mechanisms. **a.** The cap consists of 7-methylguanosine linked to the first transcribed nucleoside of the mRNA through a 5'-5 inverted triphosphate bridge. The N7 methyl residue is highlighted in magenta. The 2'-O-methyl groups of the 1st and 2nd nucleosides are highlighted in turquoise and constitute Cap 1 and Cap 2 structures, respectively. **b.** Canonical capping pathway. The canonical pathway involves the activity of an RNA triphosphatase (RTPase) required to hydrolyze the γ-phosphate from the nascent mRNA. A guanylytransferase (GTase) then transfers an GMP molecule (via the formation of a GTase-GMP covalent intermediate) onto the 5' terminal β-phosphate. In vertebrates, both the RTPase and GTase activities are present within RNGTT. This is followed by (guanylyl-N7)-methyltransferase (RNMT) transferring a methyl group from S-adenosyl-L-methionine to the N7 position of the terminal guanine to form a Cap 0 structure. Subsequent (nucleoside-2'-O)-mehtyltransferases form Cap 1 (CMTR1) and Cap 2 (CMTR2) structures. When the penultimate nucleoside is adenosine, CAPAM will transfer a methyl residue to the N6 position of adenine. N refers to the nitrogenous bases adenine, cytosine, guanine, and uracil.

Figure 2. Cap composition, as defined by CapQuant, in mouse tissue and yeast. The relative abundance of different cap structures found in mouse liver and kidney, as well as *S cerevisiae,* are shown as pie charts. Mean values were taken from Wang et al. (J. Wang et al., 2019).

Figure 3. Binding to the m^7G moiety of the cap by various cap-binding proteins. Diagrams were generated in PyMOL software using the following PDB accession numbers: CBP20 (1H2T), eIF4E (1L8B), 4EHP (2JGB), LARP1 (5V4R), VP39 (1AV6), influenza A H5N1 PB2 (4CB4), and influenza B Lee PB2 (5EFA).

Figure 4. Surface representation of cap binding pockets of the indicated CBPs. The N7-methyl position of the cap is denoted by a small grey nob.

Figure 5. Cut away view of cap binding pockets of the indicated CBPs. The N7-methyl position of the cap is denoted by a small grey nob. Note that in some of the images some of the phosphate groups are outside the depth of field shown.

Figure 6. Structure of cap analogues that have been designed with the aim of avoiding mispriming during transcription initiation and improving resistance to cellular decapping activities.

References

Abbas, Y. M., Laudenbach, B. T., Martinez-Montero, S., Cencic, R., Habjan, M., Pichlmair, A., . . . Nagar, B. (2017). Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proc Natl Acad Sci U S A, 114*(11), E2106-E2115. doi:10.1073/pnas.1612444114

Abraham, G., Rhodes, D. P., & Banerjee, A. K. (1975). The 5' terminal structure of the methylated mRNA synthesized in vitro by vesicular stomatitis virus. *Cell, 5*(1), 51-58. doi:10.1016/0092-8674(75)90091-4

Ahola, T., & Kaariainen, L. (1995). Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci U S A, 92*(2), 507-511. doi:10.1073/pnas.92.2.507

Akichika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., . . . Suzuki, T. (2019). Cap-specific terminal N (6)-methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science, 363*(6423). doi:10.1126/science.aav0080

Alesi, V., Capolino, R., Genovesea, S., Capriati, T., Loddo, S., Calvieri, G., . . . Dallapiccola, B. (2018). An additional patient with a homozygous mutation in DCPS contributes to the delination of Al-Raqad syndrome. *Am J Med Genet A, 176*(12), 2781-2786. doi:10.1002/ajmg.a.40488

Belanger, F., Stepinski, J., Darzynkiewicz, E., & Pelletier, J. (2010). Characterization of hMTr1, a human Cap1 2'-O-ribose methyltransferase. *J Biol Chem, 285*(43), 33037-33044. doi:M110.155283 [pii]10.1074/jbc.M110.155283

Beluzic, L., Grbesa, I., Beluzic, R., Park, J. H., Kong, H. K., Kopjar, N., . . . Vugrek, O. (2018). Knock-down of AHCY and depletion of adenosine induces DNA damage and cell cycle arrest. *Sci Rep, 8*(1), 14012. doi:10.1038/s41598-018-32356-8

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. *Nat Rev Drug Discov, 14*(4), 261-278. doi:10.1038/nrd4505

Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvik, I., Greene, L., . . . Nickels, B. E. (2016). The mechanism of RNA 5' capping with NAD+, NADH and desphospho-CoA. *Nature, 535*(7612), 444-447. doi:10.1038/nature18622

Boulias, K., Toczydlowska-Socha, D., Hawley, B. R., Liberman, N., Takashima, K., Zaccara, S., . . . Greer, E. L. (2019). Identification of the m(6)Am Methyltransferase PCIF1 Reveals the Location and Functions of m(6)Am in the Transcriptome. *Mol Cell, 75*(3), 631-643 e638. doi:10.1016/j.molcel.2019.06.006

Brown, C. J., McNae, I., Fischer, P. M., & Walkinshaw, M. D. (2007). Crystallographic and mass spectrometric characterisation of eIF4E with N7-alkylated cap derivatives. *J Mol Biol, 372*(1), 7-15. doi:10.1016/j.jmb.2007.06.033

Brown, J. D., Kharytonchyk, S., Chaudry, I., Iyer, A. S., Carter, H., Becker, G., . . . Summers, M. F. (2020). Structural basis for transcriptional start site control of HIV-1 RNA fate. *Science, 368*(6489), 413-417. doi:10.1126/science.aaz7959

Cai, A., Jankowska-Anyszka, M., Centers, A., Chlebicka, L., Stepinski, J., Stolarski, R., . . . Rhoads, R. E. (1999). Quantitative assessment of mRNA cap analogues as inhibitors of in vitro translation. *Biochemistry, 38*(26), 8538-8547.

Chang, J. H., Jiao, X., Chiba, K., Oh, C., Martin, C. E., Kiledjian, M., & Tong, L. (2012). Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat Struct Mol Biol, 19*(10), 1011-1017. doi:10.1038/nsmb.2381

Changela, A., Ho, C. K., Martins, A., Shuman, S., & Mondragon, A. (2001). Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J, 20*(10), 2575-2586. doi:10.1093/emboj/20.10.2575

Chapat, C., Jafarnejad, S. M., Matta-Camacho, E., Hesketh, G. G., Gelbart, I. A., Attig, J., . . . Sonenberg, N. (2017). Cap-binding protein 4EHP effects translation silencing by microRNAs. *Proc Natl Acad Sci U S A, 114*(21), 5425-5430. doi:10.1073/pnas.1701488114

Chaudhry, Y., Nayak, A., Bordeleau, M. E., Tanaka, J., Pelletier, J., Belsham, G. J., . . . Goodfellow, I. G. (2006). Caliciviruses differ in their functional requirements for eIF4F components. *J Biol Chem, 281*(35), 25315-25325.

Chen, E. Z., Jacobson, B. A., Patel, M. R., Okon, A. M., Li, S., Xiong, K., . . . Kratzke, R. A. (2014). Small-molecule inhibition of oncogenic eukaryotic protein translation in mesothelioma cells. *Invest New Drugs, 32*(4), 598-603. doi:10.1007/s10637-014-0076-7

Chen, X., Kopecky, D. J., Mihalic, J., Jeffries, S., Min, X., Heath, J., . . . Wang, Z. (2012). Structure-guided design, synthesis, and evaluation of guanine-derived inhibitors of the eIF4E mRNA-cap interaction. *J Med Chem, 55*(8), 3837-3851. doi:10.1021/jm300037x

Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., & Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev, 15*(24), 3319-3329. doi:10.1101/gad.935901

Cho, P. F., Poulin, F., Cho-Park, Y. A., Cho-Park, I. B., Chicoine, J. D., Lasko, P., & Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell, 121*(3), 411-423. doi:10.1016/j.cell.2005.02.024

Chu, C., Das, K., Tyminski, J. R., Bauman, J. D., Guan, R., Qiu, W., . . . Shatkin, A. J. (2011). Structure of the guanylyltransferase domain of human mRNA capping enzyme. *Proc Natl Acad Sci U S A, 108*(25), 10104-10108. doi:10.1073/pnas.1106610108

Chu, C., & Shatkin, A. J. (2008). Apoptosis and autophagy induction in mammalian cells by small interfering RNA knockdown of mRNA capping enzymes. *Mol Cell Biol, 28*, 5829-5836.

Chung, L., Bailey, D., Leen, E. N., Emmott, E. P., Chaudhry, Y., Roberts, L. O., . . . Goodfellow, I. G. (2014). Norovirus translation requires an interaction between the C Terminus of the genome-linked viral protein VPg and eukaryotic translation initiation factor 4G. *J Biol Chem, 289*(31), 21738-21750. doi:10.1074/jbc.M114.550657

Cole, M. D., & Cowling, V. H. (2009). Specific regulation of mRNA cap methylation by the c-Myc and E2F1 transcription factors. *Oncogene, 28*(9), 1169-1175. doi:10.1038/onc.2008.463

Cowling, V. H. (2010). Enhanced mRNA cap methylation increases cyclin D1 expression and promotes cell transformation. *Oncogene, 29*(6), 930-936. doi:10.1038/onc.2009.368

Doamekpor, S. K., Gozdek, A., Kwasnik, A., Kufel, J., & Tong, L. (2020). A novel 5'-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes. *Nucleic Acids Res, 48*(1), 349-358. doi:10.1093/nar/gkz1107

Domashevskiy, A. V., Rodriguez, D. J., Gunawardana, D., & Goss, D. J. (2016). Preparation of Functional, Fluorescently Labeled mRNA Capped with Anthraniloyl-m(7)GpppG. *Methods Mol Biol, 1428*, 61-75. doi:10.1007/978-1-4939-3625-0_4

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., . . . He, C. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature, 530*(7591), 441-446. doi:10.1038/nature16998

Edery, I., Altmann, M., & Sonenberg, N. (1988). High-level synthesis in Escherichia coli of functional cap-binding eukaryotic initiation factor eIF-4E and affinity purification using a simplified cap-analog resin. *Gene, 74*(2), 517-525.

Fernandez-Sanchez, M. E., Gonatopoulos-Pournatzis, T., Preston, G., Lawlor, M. A., & Cowling, V. H. (2009). S-adenosyl homocysteine hydrolase is required for Myc-induced mRNA cap methylation, protein synthesis, and cell proliferation. *Mol Cell Biol, 29*(23), 6182-6191. doi:10.1128/MCB.00973-09

Fonseca, B. D., Lahr, R. M., Damgaard, C. K., Alain, T., & Berman, A. J. (2018). LARP1 on TOP of ribosome production. *Wiley Interdiscip Rev RNA*, e1480. doi:10.1002/wrna.1480

Fonseca, B. D., Zakaria, C., Jia, J. J., Graber, T. E., Svitkin, Y., Tahmasebi, S., . . . Damgaard, C. K. (2015). La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). *J Biol Chem, 290*(26), 15996-16020. doi:10.1074/jbc.M114.621730

Franke, J., Gehlen, J., & Ehrenhofer-Murray, A. E. (2008). Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J Cell Sci, 121*(Pt 21), 3553-3560. doi:10.1242/jcs.033308

Furic, L., Rong, L., Larsson, O., Koumakpayi, I. H., Yoshida, K., Brueschke, A., . . . Sonenberg, N. (2010). eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc Natl Acad Sci U S A, 107*(32), 14134-14139. doi:1005320107 [pii]10.1073/pnas.1005320107

Furuichi, Y. (2015). Discovery of m(7)G-cap in eukaryotic mRNAs. *Proc Jpn Acad Ser B Phys Biol Sci, 91*(8), 394-409. doi:10.2183/pjab.91.394

Gallie, D. R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev, 5*(11), 2108-2116.

Galloway, A., Atrih, A., Grzela, R., Darzynkiewicz, E., Ferguson, M. A. J., & Cowling, V. H. (2020). CAP-MAP: cap analysis protocol with minimal analyte processing, a rapid and sensitive approach to analysing mRNA cap structures. *Open Biol, 10*(2), 190306. doi:10.1098/rsob.190306

Gershon, P. D., Ahn, B. Y., Garfield, M., & Moss, B. (1991). Poly(A) polymerase and a dissociable polyadenylation stimulatory factor encoded by vaccinia virus. *Cell, 66*(6), 1269-1278. doi:10.1016/0092-8674(91)90048-4

Ghosh, B., Benyumov, A. O., Ghosh, P., Jia, Y., Avdulov, S., Dahlberg, P. S., . . . Wagner, C. R. (2009). Nontoxic chemical interdiction of the epithelial-to-mesenchymal transition by targeting cap-dependent translation. *ACS Chem Biol, 4*(5), 367-377. doi:10.1021/cb9000475

Gonatopoulos-Pournatzis, T., Dunn, S., Bounds, R., & Cowling, V. H. (2011). RAM/Fam103a1 is required for mRNA cap methylation. *Mol Cell, 44*(4), 585-596. doi:10.1016/j.molcel.2011.08.041

Goodfellow, I., Chaudhry, Y., Gioldasi, I., Gerondopoulos, A., Natoni, A., Labrie, L., . . . Roberts, L. (2005). Calicivirus translation initiation requires an interaction between VPg and eIF 4 E. *EMBO Rep, 6*(10), 968-972. doi:10.1038/sj.embor.7400510

Grudzien-Nogalska, E., & Kiledjian, M. (2017). New insights into decapping enzymes and selective mRNA decay. *Wiley Interdiscip Rev RNA, 8*(1). doi:10.1002/wrna.1379

Grudzien-Nogalska, E., Stepinski, J., Jemielity, J., Zuberek, J., Stolarski, R., Rhoads, R. E., & Darzynkiewicz, E. (2007). Synthesis of anti-reverse cap analogs (ARCAs) and their applications in mRNA translation and stability. *Methods Enzymol, 431*, 203-227. doi:S0076-6879(07)31011-2 [pii]10.1016/S0076-6879(07)31011-2

Guiguen, A., Soutourina, J., Dewez, M., Tafforeau, L., Dieu, M., Raes, M., . . . Hermand, D. (2007). Recruitment of P-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast. *EMBO J, 26*(6), 1552-1559. doi:10.1038/sj.emboj.7601627

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., . . . Cusack, S. (2008). The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol, 15*(5), 500-506. doi:10.1038/nsmb.1421

Haline-Vaz, T., Silva, T. C., & Zanchin, N. I. (2008). The human interferon-regulated ISG95 protein interacts with RNA polymerase II and shows methyltransferase activity. *Biochem Biophys Res Commun, 372*(4), 719-724. doi:10.1016/j.bbrc.2008.05.137

Haller, A. A., & Sarnow, P. (1997). In vitro selection of a 7-methyl-guanosine binding RNA that inhibits translation of capped mRNA molecules. *Proc Natl Acad Sci U S A, 94*(16), 8521-8526.

Hamm, J., Darzynkiewicz, E., Tahara, S. M., & Mattaj, I. W. (1990). The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell, 62*(3), 569-577. doi:10.1016/0092-8674(90)90021-6

Hickey, K. L., Dickson, K., Cogan, J. Z., Replogle, J. M., Schoof, M., D'Orazio, K. N., . . . Kostova, K. K. (2020). GIGYF2 and 4EHP Inhibit Translation Initiation of Defective Messenger RNAs to Assist Ribosome-Associated Quality Control. *Mol Cell, 79*(6), 950-962 e956. doi:10.1016/j.molcel.2020.07.007

Ho, C. K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., & Shuman, S. (1998). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem, 273*(16), 9577-9585. doi:10.1074/jbc.273.16.9577

Hodel, A. E., Gershon, P. D., Shi, X., & Quiocho, F. A. (1996). The 1.85 A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. *Cell, 85*(2), 247-256. doi:10.1016/s0092-8674(00)81101-0

Holstein, J. M., Stummer, D., & Rentmeister, A. (2015). Engineering Giardia lamblia trimethylguanosine synthase (GlaTgs2) to transfer non-natural modifications to the RNA 5'-cap. *Protein Eng Des Sel, 28*(6), 179-186. doi:10.1093/protein/gzv011

Hong, S., Freeberg, M. A., Han, T., Kamath, A., Yao, Y., Fukuda, T., . . . Inoki, K. (2017). LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *Elife, 6*. doi:10.7554/eLife.25237

HsuChen, C. C., & Dubin, D. T. (1976). Di-and trimethylated congeners of 7-methylguanine in Sindbis virus mRNA. *Nature, 264*(5582), 190-191. doi:10.1038/264190a0

Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., & Luhrmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J, 17*(14), 4114-4126. doi:10.1093/emboj/17.14.4114

Hyde, J. L., & Diamond, M. S. (2015). Innate immune restriction and antagonism of viral RNA lacking 2-O methylation. *Virology, 479-480*, 66-74. doi:10.1016/j.virol.2015.01.019

Issur, M., Picard-Jean, F., & Bisaillon, M. (2011). The RNA capping machinery as an anti-infective target. *Wiley Interdiscip Rev RNA, 2*(2), 184-192. doi:10.1002/wrna.43

Ivanov, D., Kwak, Y. T., Guo, J., & Gaynor, R. B. (2000). Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol, 20*(9), 2970-2983. doi:10.1128/mcb.20.9.2970-2983.2000

Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., & Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol, 62*(8), 2636-2643.

Jia, Y., Chiu, T. L., Amin, E. A., Polunovsky, V., Bitterman, P. B., & Wagner, C. R. (2010). Design, synthesis and evaluation of analogs of initiation factor 4E (eIF4E) cap-binding antagonist Bn7-GMP. *Eur J Med Chem, 45*(4), 1304-1313. doi:10.1016/j.ejmech.2009.11.054

Jiao, X., Chang, J. H., Kilic, T., Tong, L., & Kiledjian, M. (2013). A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol Cell, 50*(1), 104-115. doi:10.1016/j.molcel.2013.02.017

Jiao, X., Doamekpor, S. K., Bird, J. G., Nickels, B. E., Tong, L., Hart, R. P., & Kiledjian, M. (2017). 5' End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell, 168*(6), 1015-1027 e1010. doi:10.1016/j.cell.2017.02.019

Jiao, X., Xiang, S., Oh, C., Martin, C. E., Tong, L., & Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature, 467*(7315), 608-611. doi:10.1038/nature09338

Joshi, B., Cameron, A., & Jagus, R. (2004). Characterization of mammalian eIF4E-family members. *Eur J Biochem, 271*(11), 2189-2203. doi:10.1111/j.1432-1033.2004.04149.x

Julius, C., & Yuzenkova, Y. (2017). Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors. *Nucleic Acids Res, 45*(14), 8282-8290. doi:10.1093/nar/gkx452

Juszkiewicz, S., Slodkowicz, G., Lin, Z., Freire-Pritchett, P., Peak-Chew, S. Y., & Hegde, R. S. (2020). Ribosome collisions trigger cis-acting feedback inhibition of translation initiation. *Elife, 9*. doi:10.7554/eLife.60038

Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M.-N., & Sonenberg, N. (2005). Mammalian poly(A)-binding protien is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes & Dev, 19*, 104-113.

Kiledjian, M. (2018). Eukaryotic RNA 5'-End NAD(+) Capping and DeNADding. *Trends Cell Biol, 28*(6), 454-464. doi:10.1016/j.tcb.2018.02.005

Kim, J. B., & Sharp, P. A. (2001). Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J Biol Chem, 276*(15), 12317-12323. doi:10.1074/jbc.M010908200

Kocmik, I., Piecyk, K., Rudzinska, M., Niedzwiecka, A., Darzynkiewicz, E., Grzela, R., & Jankowska-Anyszka, M. (2018). Modified ARCA analogs providing enhanced translational properties of capped mRNAs. *Cell Cycle, 17*(13), 1624-1636. doi:10.1080/15384101.2018.1486164

Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev, 14*(19), 2452-2460.

Konarska, M. M., Padgett, R. A., & Sharp, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell, 38*(3), 731-736.

Konicek, B. W., Stephens, J. R., McNulty, A. M., Robichaud, N., Peery, R. B., Dumstorf, C. A., . . . Graff, J. R. (2011). Therapeutic inhibition of MAP kinase interacting kinase blocks eukaryotic initiation factor 4E phosphorylation and suppresses outgrowth of experimental lung metastases. *Cancer Res, 71*(5), 1849-1857. doi:0008-5472.CAN-10-3298 [pii]10.1158/0008-5472.CAN-10-3298

Koonin, E. V., & Yutin, N. (2019). Evolution of the Large Nucleocytoplasmic DNA Viruses of Eukaryotes and Convergent Origins of Viral Gigantism. *Adv Virus Res, 103*, 167-202. doi:10.1016/bs.aivir.2018.09.002

Kore, A. R., Shanmugasundaram, M., Charles, I., Vlassov, A. V., & Barta, T. J. (2009). Locked nucleic acid (LNA)-modified dinucleotide mRNA cap analogue: synthesis, enzymatic incorporation, and utilization. *J Am Chem Soc, 131*(18), 6364-6365. doi:10.1021/ja901655p

Kowalska, J., Wypijewska del Nogal, A., Darzynkiewicz, Z. M., Buck, J., Nicola, C., Kuhn, A. N., . . . Jemielity, J. (2014). Synthesis, properties, and biological activity of boranophosphate analogs of the mRNA cap: versatile tools for manipulation of therapeutically relevant cap-dependent processes. *Nucleic Acids Res, 42*(16), 10245-10264. doi:10.1093/nar/gku757

Kramer, S., & McLennan, A. G. (2019). The complex enzymology of mRNA decapping: Enzymes of four classes cleave pyrophosphate bonds. *Wiley Interdiscip Rev RNA, 10*(1), e1511. doi:10.1002/wrna.1511

Kumar, P., Sweeney, T. R., Skabkin, M. A., Skabkina, O. V., Hellen, C. U., & Pestova, T. V. (2014). Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5'ppp- mRNAs. *Nucleic Acids Res, 42*(5), 3228-3245. doi:10.1093/nar/gkt1321

Lahr, R. M., Fonseca, B. D., Ciotti, G. E., Al-Ashtal, H. A., Jia, J. J., Niklaus, M. R., . . . Berman, A. J. (2017). La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *Elife, 6*. doi:10.7554/eLife.24146

Lenasi, T., Peterlin, B. M., & Barboric, M. (2011). Cap-binding protein complex links pre-mRNA capping to transcription elongation and alternative splicing through positive transcription elongation factor b (P-TEFb). *J Biol Chem, 286*(26), 22758-22768. doi:10.1074/jbc.M111.235077

Li, S., Jia, Y., Jacobson, B., McCauley, J., Kratzke, R., Bitterman, P. B., & Wagner, C. R. (2013). Treatment of breast and lung cancer cells with a N-7 benzyl guanosine monophosphate tryptamine phosphoramidate pronucleotide (4Ei-1) results in chemosensitization to gemcitabine and induced eIF4E proteasomal degradation. *Mol Pharm, 10*(2), 523-531. doi:10.1021/mp300699d

Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., . . . Yi, C. (2017). Base-Resolution Mapping Reveals Distinct m(1)A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol Cell, 68*(5), 993-1005 e1009. doi:10.1016/j.molcel.2017.10.019

Li, Y., Song, M., & Kiledjian, M. (2011). Differential utilization of decapping enzymes in mammalian mRNA decay pathways. *RNA, 17*(3), 419-428. doi:10.1261/rna.2439811

Liu, S. W., Jiao, X., Liu, H., Gu, M., Lima, C. D., & Kiledjian, M. (2004). Functional analysis of mRNA scavenger decapping enzymes. *RNA, 10*(9), 1412-1422. doi:10.1261/rna.7660804

Liu, W., Jankowska-Anyszka, M., Piecyk, K., Dickson, L., Wallace, A., Niedzwiecka, A., . . . Davis, R. E. (2011). Structural basis for nematode eIF4E binding an m(2,2,7)G-Cap and its implications for translation initiation. *Nucleic Acids Res, 39*(20), 8820-8832. doi:10.1093/nar/gkr650

Lockless, S. W., Cheng, H. T., Hodel, A. E., Quiocho, F. A., & Gershon, P. D. (1998). Recognition of capped RNA substrates by VP39, the vaccinia virus-encoded mRNA cap-specific 2'-O-methyltransferase. *Biochemistry, 37*(23), 8564-8574. doi:10.1021/bi980178m

Lombardi, O., Varshney, D., Phillips, N. M., & Cowling, V. H. (2016). c-Myc deregulation induces mRNA capping enzyme dependency. *Oncotarget, 7*(50), 82273-82288. doi:10.18632/oncotarget.12701

Lu, G., Zhang, J., Li, Y., Li, Z., Zhang, N., Xu, X., . . . Yan, J. (2011). hNUDT16: a universal decapping enzyme for small nucleolar RNA and cytoplasmic mRNA. *Protein Cell, 2*(1), 64-73. doi:10.1007/s13238-011-1009-2

Lykke-Andersen, J. (2002). Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol, 22*(23), 8114-8121. doi:10.1128/mcb.22.23.8114-8121.2002

Mandal, S. S., Chu, C., Wada, T., Handa, H., Shatkin, A. J., & Reinberg, D. (2004). Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci U S A, 101*(20), 7572-7577. doi:10.1073/pnas.0401493101

Marchand, V., Ayadi, L., Ernst, F. G. M., Hertler, J., Bourguignon-Igel, V., Galvanin, A., . . . Motorin, Y. (2018). AlkAniline-Seq: Profiling of m(7) G and m(3) C RNA Modifications at Single Nucleotide Resolution. *Angew Chem Int Ed Engl, 57*(51), 16785-16790. doi:10.1002/anie.201810946

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell, 89*(6), 951-961.

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. *Mol Cell, 7*(1), 193-203.

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. *Nat Struct Biol, 4*(9), 717-724.

Mazza, C., Ohno, M., Segref, A., Mattaj, I. W., & Cusack, S. (2001). Crystal structure of the human nuclear cap binding complex. *Mol Cell, 8*(2), 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 J, 21*(20), 5548-5557. doi:10.1093/emboj/cdf538

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., . . . Bentley, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev, 11*(24), 3306-3318. doi:10.1101/gad.11.24.3306

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., & Green, M. R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res, 12*(18), 7035-7056. doi:10.1093/nar/12.18.7035

Meyuhas, O., & Kahan, T. (2015). The race to decipher the top secrets of TOP mRNAs. *Biochim Biophys Acta, 1849*(7), 801-811. doi:10.1016/j.bbagrm.2014.08.015

Monecke, T., Dickmanns, A., & Ficner, R. (2009). Structural basis for m7G-cap hypermethylation of small nuclear, small nucleolar and telomerase RNA by the dimethyltransferase TGS1. *Nucleic Acids Res, 37*(12), 3865-3877. doi:10.1093/nar/gkp249

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., & Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell, 9*(4), 891-901. doi:10.1016/s1097-2765(02)00484-7

Muttach, F., Muthmann, N., & Rentmeister, A. (2017). Chemo-enzymatic modification of eukaryotic mRNA. *Org Biomol Chem, 15*(2), 278-284. doi:10.1039/c6ob02144a

Ng, C. K., Shboul, M., Taverniti, V., Bonnard, C., Lee, H., Eskin, A., . . . Reversade, B. (2015). Loss of the scavenger mRNA decapping enzyme DCPS causes syndromic intellectual disability with neuromuscular defects. *Hum Mol Genet, 24*(11), 3163-3171. doi:10.1093/hmg/ddv067

Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszynska, A., Dadlez, M., . . . Stolarski, R. (2002). Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol, 319*(3), 615-635. doi:10.1016/S0022-2836(02)00328-5S0022-2836(02)00328-5 [pii]

Nowakowska, M., Kowalska, J., Martin, F., d'Orchymont, A., Zuberek, J., Lukaszewicz, M., . . . Jemielity, J. (2014). Cap analogs containing 6-thioguanosine--reagents for the synthesis of mRNAs selectively photo-crosslinkable with cap-binding biomolecules. *Org Biomol Chem, 12*(27), 4841-4847. doi:10.1039/c4ob00059e

Ogino, T., & Banerjee, A. K. (2007). Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell, 25*(1), 85-97. doi:10.1016/j.molcel.2006.11.013

Osborne, M. J., Volpon, L., Kornblatt, J. A., Culjkovic-Kraljacic, B., Baguet, A., & Borden, K. L. (2013). eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. *Proc Natl Acad Sci U S A, 110*(10), 3877-3882. doi:10.1073/pnas.1216862110

Pandey, R. R., Delfino, E., Homolka, D., Roithova, A., Chen, K. M., Li, L., . . . Pillai, R. S. (2020). The Mammalian Cap-Specific m(6)Am RNA Methyltransferase PCIF1 Regulates Transcript Levels in Mouse Tissues. *Cell Rep, 32*(7), 108038. doi:10.1016/j.celrep.2020.108038

Pelletier, J., & Sonenberg, N. (1985). Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell, 40*(3), 515-526.

Pelletier, J., & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature, 334*(6180), 320-325.

Pelletier, J., & Sonenberg, N. (2019). The Organizing Principles of Eukaryotic Ribosome Recruitment. *Annu Rev Biochem, 88*, 307-335. doi:10.1146/annurev-biochem-013118-111042

Perry, K. L., Watkins, K. P., & Agabian, N. (1987). Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. *Proc Natl Acad Sci U S A, 84*(23), 8190-8194. doi:10.1073/pnas.84.23.8190

Philippe, L., Vasseur, J. J., Debart, F., & Thoreen, C. C. (2018). La-related protein 1 (LARP1) repression of TOP mRNA translation is mediated through its cap-binding domain and controlled by an adjacent regulatory region. *Nucleic Acids Res, 46*(3), 1457-1469. doi:10.1093/nar/gkx1237

Picard-Jean, F., Tremblay-Létourneau, M., Serra, E., Dimech, C., Schulz, H., Asnselin, M., . . . Bisaillon, M. (2013). RNA 5'-end Maturation: A Crucial Step in the Replication of Viral Genomes. In V. Romanowski (Ed.), *Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications* (pp. 27 - 56). London: IntechOpen.

Piccirillo, C., Khanna, R., & Kiledjian, M. (2003). Functional characterization of the mammalian mRNA decapping enzyme hDcp2. *RNA, 9*(9), 1138-1147. doi:10.1261/rna.5690503

Pillutla, R. C., Yue, Z., Maldonado, E., & Shatkin, A. J. (1998). Recombinant human mRNA cap methyltransferase binds capping enzyme/RNA polymerase IIo complexes. *J Biol Chem, 273*(34), 21443-21446. doi:10.1074/jbc.273.34.21443

Plotch, S. J., Bouloy, M., Ulmanen, I., & Krug, R. M. (1981). A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell, 23*(3), 847-858. doi:10.1016/0092-8674(81)90449-9

Pyronnet, S., Imataka, H., Gingras, A. C., Fukunaga, R., Hunter, T., & Sonenberg, N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J, 18*(1), 270-279. doi:10.1093/emboj/18.1.270

Reich, S. H., Sprengeler, P. A., Chiang, G. G., Appleman, J. R., Chen, J., Clarine, J., . . . Webster, K. R. (2018). Structure-based Design of Pyridone-Aminal eFT508 Targeting Dysregulated Translation by Selective Mitogen-activated Protein Kinase Interacting Kinases 1 and 2 (MNK1/2) Inhibition. *J Med Chem, 61*(8), 3516-3540. doi:10.1021/acs.jmedchem.7b01795

Robert, F., Cencic, R., Cai, R., Schmeing, T. M., & Pelletier, J. (2020). RNA-tethering assay and eIF4G:eIF4A obligate dimer design uncovers multiple eIF4F functional complexes. *Nucleic Acids Res, 48*(15), 8562-8575. doi:10.1093/nar/gkaa646

Rojas-Duran, M. F., & Gilbert, W. V. (2012). Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA, 18*(12), 2299-2305. doi:10.1261/rna.035865.112

Rosettani, P., Knapp, S., Vismara, M. G., Rusconi, L., & Cameron, A. D. (2007). Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. *J Mol Biol, 368*(3), 691-705. doi:10.1016/j.jmb.2007.02.019

Rydzik, A. M., Warminski, M., Sikorski, P. J., Baranowski, M. R., Walczak, S., Kowalska, J., . . . Jemielity, J. (2017). mRNA cap analogues substituted in the tetraphosphate chain with CX2: identification of O-to-CCl2 as the first bridging modification that confers resistance to decapping without impairing translation. *Nucleic Acids Res, 45*(15), 8661-8675. doi:10.1093/nar/gkx569

Saha, N., Schwer, B., & Shuman, S. (1999). Characterization of human, Schizosaccharomyces pombe, and Candida albicans mRNA cap methyltransferases and complete replacement of the yeast capping apparatus by mammalian enzymes. *J Biol Chem, 274*(23), 16553-16562. doi:10.1074/jbc.274.23.16553

Scheper, G. C., van Kollenburg, B., Hu, J., Luo, Y., Goss, D. J., & Proud, C. G. (2002). Phosphorylation of eukaryotic initiation factor 4E markedly reduces its affinity for capped mRNA. *J Biol Chem, 277*(5), 3303-3309.

Schnierle, B. S., Gershon, P. D., & Moss, B. (1992). Cap-specific mRNA (nucleoside-O2'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. *Proc Natl Acad Sci U S A, 89*(7), 2897-2901. doi:10.1073/pnas.89.7.2897

Schulz, D., Holstein, J. M., & Rentmeister, A. (2013). A chemo-enzymatic approach for site-specific modification of the RNA cap. *Angew Chem Int Ed Engl, 52*(30), 7874-7878. doi:10.1002/anie.201302874

Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., . . . Woyke, T. (2017). Giant viruses with an expanded complement of translation system components. *Science, 356*(6333), 82-85. doi:10.1126/science.aal4657

Sendinc, E., Valle-Garcia, D., Dhall, A., Chen, H., Henriques, T., Navarrete-Perea, J., . . . Shi, Y. (2019). PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression. *Mol Cell, 75*(3), 620-630 e629. doi:10.1016/j.molcel.2019.05.030

Shveygert, M., Kaiser, C., Bradrick, S. S., & Gromeier, M. (2010). Regulation of eukaryotic initiation factor 4E (eIF4E) phosphorylation by mitogen-activated protein kinase occurs through modulation of Mnk1-eIF4G interaction. *Mol Cell Biol, 30*(21), 5160-5167. doi:10.1128/MCB.00448-10

Slepenkov, S. V., Darzynkiewicz, E., & Rhoads, R. E. (2006). Stopped-flow kinetic analysis of eIF4E and phosphorylated eIF4E binding to cap analogs and capped oligoribonucleotides: evidence for a one-step binding mechanism. *J Biol Chem, 281*(21), 14927-14938. doi:10.1074/jbc.M601653200

Smietanski, M., Werner, M., Purta, E., Kaminska, K. H., Stepinski, J., Darzynkiewicz, E., . . . Bujnicki, J. M. (2014). Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat Commun, 5*, 3004. doi:10.1038/ncomms4004

Song, M. G., Bail, S., & Kiledjian, M. (2013). Multiple Nudix family proteins possess mRNA decapping activity. *RNA, 19*(3), 390-399. doi:10.1261/rna.037309.112

Song, M. G., Li, Y., & Kiledjian, M. (2010). Multiple mRNA decapping enzymes in mammalian cells. *Mol Cell, 40*(3), 423-432. doi:10.1016/j.molcel.2010.10.010

Stepinski, J., & Darzynkiewicz, E. (2014). mRNA and snRNA Cap Analogs: Synthesis and Applications. *Chemical Biology of Nucleic Acids: Fundamentals and Clinical Applications*, 511-561. doi:10.1007/978-3-642-54452-1_28

Su, W., Slepenkov, S., Grudzien-Nogalska, E., Kowalska, J., Kulis, M., Zuberek, J., . . . Rhoads, R. E. (2011). Translation, stability, and resistance to decapping of mRNAs containing caps substituted in the triphosphate chain with BH3, Se, and NH. *RNA, 17*(5), 978-988. doi:rna.2430711 [pii]10.1261/rna.2430711

Sun, H., Zhang, M., Li, K., Bai, D., & Yi, C. (2019). Cap-specific, terminal N(6)-methylation by a mammalian m(6)Am methyltransferase. *Cell Res, 29*(1), 80-82. doi:10.1038/s41422-018-0117-4

Svitkin, Y. V., Ovchinnikov, L. P., Dreyfuss, G., & Sonenberg, N. (1996). General RNA binding proteins render translation cap dependent. *EMBO J, 15*(24), 7147-7155.

Svitkin, Y. V., Pause, A., Haghighat, A., Pyronnet, S., Witherell, G., Belsham, G. J., & Sonenberg, N. (2001). The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA, 7*(3), 382-394.

Tamarkin-Ben-Harush, A., Vasseur, J. J., Debart, F., Ulitsky, I., & Dikstein, R. (2017). Cap-proximal nucleotides via differential eIF4E binding and alternative promoter usage mediate translational response to energy stress. *Elife, 6*. doi:10.7554/eLife.21907

Tcherkezian, J., Cargnello, M., Romeo, Y., Huttlin, E. L., Lavoie, G., Gygi, S. P., & Roux, P. P. (2014). Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5′TOP mRNA translation. *Genes Dev, 28*(4), 357-371. doi:10.1101/gad.231407.113

Tomikawa, C. (2018). 7-Methylguanosine Modifications in Transfer RNA (tRNA). *Int J Mol Sci, 19*(12). doi:10.3390/ijms19124080

Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., . . . Miura, K. (2002). Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem J, 362*(Pt 3), 539-544.

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA, 2*(2), 277-298. doi:10.1002/wrna.52

Tsukamoto, T., Shibagaki, Y., Niikura, Y., & Mizumoto, K. (1998). Cloning and characterization of three human cDNAs encoding mRNA (guanine-7-)-methyltransferase, an mRNA cap methylase. *Biochem Biophys Res Commun, 251*(1), 27-34. doi:10.1006/bbrc.1998.9402

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science, 249*, 505-510.

Ueda, T., Watanabe-Fukunaga, R., Fukuyama, H., Nagata, S., & Fukunaga, R. (2004). Mnk2 and Mnk1 Are Essential for Constitutive and Inducible Phosphorylation of Eukaryotic Initiation Factor 4E but Not for Cell Growth or Development. *Mol Cell Biol, 24*(15), 6539-6549.

van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., & Seraphin, B. (2002). Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J, 21*(24), 6915-6924. doi:10.1093/emboj/cdf678

van Duijn, L. P., Kasperaitis, M., Ameling, C., & Voorma, H. O. (1986). Additional methylation at the N(2)-position of the cap of 26S Semliki Forest virus late mRNA and initiation of translation. *Virus Res, 5*(1), 61-66. doi:10.1016/0168-1702(86)90065-1

Varshney, D., Petit, A. P., Bueren-Calabuig, J. A., Jansen, C., Fletcher, D. A., Peggie, M., . . . Cowling, V. H. (2016). Molecular basis of RNA guanine-7 methyltransferase (RNMT) activation by RAM. *Nucleic Acids Res, 44*(21), 10423-10436. doi:10.1093/nar/gkw637

Walters, R. W., Matheny, T., Mizoue, L. S., Rao, B. S., Muhlrad, D., & Parker, R. (2017). Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A, 114*(3), 480-485. doi:10.1073/pnas.1619369114

Wang, J., Alvin Chew, B. L., Lai, Y., Dong, H., Xu, L., Balamkundu, S., . . . Dedon, P. C. (2019). Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res, 47*(20), e130. doi:10.1093/nar/gkz751

Wang, V. Y., Jiao, X., Kiledjian, M., & Tong, L. (2015). Structural and biochemical studies of the distinct activity profiles of Rai1 enzymes. *Nucleic Acids Res, 43*(13), 6596-6606. doi:10.1093/nar/gkv620

Wang, Y., Li, S., Zhao, Y., You, C., Le, B., Gong, Z., . . . Chen, X. (2019). NAD(+)-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated. *Proc Natl Acad Sci U S A, 116*(24), 12094-12102. doi:10.1073/pnas.1903682116

Wang, Z., Jiao, X., Carr-Schmid, A., & Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A, 99*(20), 12663-12668. doi:10.1073/pnas.192445599

Wen, Y., & Shatkin, A. J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev, 13*(14), 1774-1779. doi:10.1101/gad.13.14.1774

Wen, Y., & Shatkin, A. J. (2000). Cap methyltransferase selective binding and methylation of GpppG-RNA are stimulated by importin-alpha. *Genes Dev, 14*(23), 2944-2949. doi:10.1101/gad.848200

Werner, M., Purta, E., Kaminska, K. H., Cymerman, I. A., Campbell, D. A., Mittra, B., . . . Bujnicki, J. M. (2011). 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res, 39*(11), 4756-4768. doi:10.1093/nar/gkr038

Wojtczak, B. A., Sikorski, P. J., Fac-Dabrowska, K., Nowicka, A., Warminski, M., Kubacka, D., . . . Jemielity, J. (2018). 5'-Phosphorothiolate Dinucleotide Cap Analogues: Reagents for Messenger RNA Modification and Potent Small-Molecular Inhibitors of Decapping Enzymes. *J Am Chem Soc, 140*(18), 5987-5999. doi:10.1021/jacs.8b02597

Wu, H., Li, L., Chen, K. M., Homolka, D., Gos, P., Fleury-Olela, F., . . . Pillai, R. S. (2019). Decapping Enzyme NUDT12 Partners with BLMH for Cytoplasmic Surveillance of NAD-Capped RNAs. *Cell Rep, 29*(13), 4422-4434 e4413. doi:10.1016/j.celrep.2019.11.108

Wurth, L., Gribling-Burrer, A. S., Verheggen, C., Leichter, M., Takeuchi, A., Baudrey, S., . . . Allmang, C. (2014). Hypermethylated-capped selenoprotein mRNAs in mammals. *Nucleic Acids Res, 42*(13), 8663-8677. doi:10.1093/nar/gku580

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J. L., & Tong, L. (2009). Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. *Nature, 458*(7239), 784-788. doi:10.1038/nature07731

Xie, L., Wartchow, C., Shia, S., Uehara, K., Steffek, M., Warne, R., . . . Ma, X. (2016). Molecular Basis of mRNA Cap Recognition by Influenza B Polymerase PB2 Subunit. *J Biol Chem, 291*(1), 363-370. doi:10.1074/jbc.M115.693051

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., . . . Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell, 97*(1), 41-51. doi:10.1016/s0092-8674(00)80713-8

Yanagiya, A., Svitkin, Y. V., Shibata, S., Mikami, S., Imataka, H., & Sonenberg, N. (2009). Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol, 29*(6), 1661-1669. doi:MCB.01187-08 [pii]10.1128/MCB.01187-08

Yedavalli, V. S., & Jeang, K. T. (2010). Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proc Natl Acad Sci U S A, 107*(33), 14787-14792. doi:10.1073/pnas.1009490107

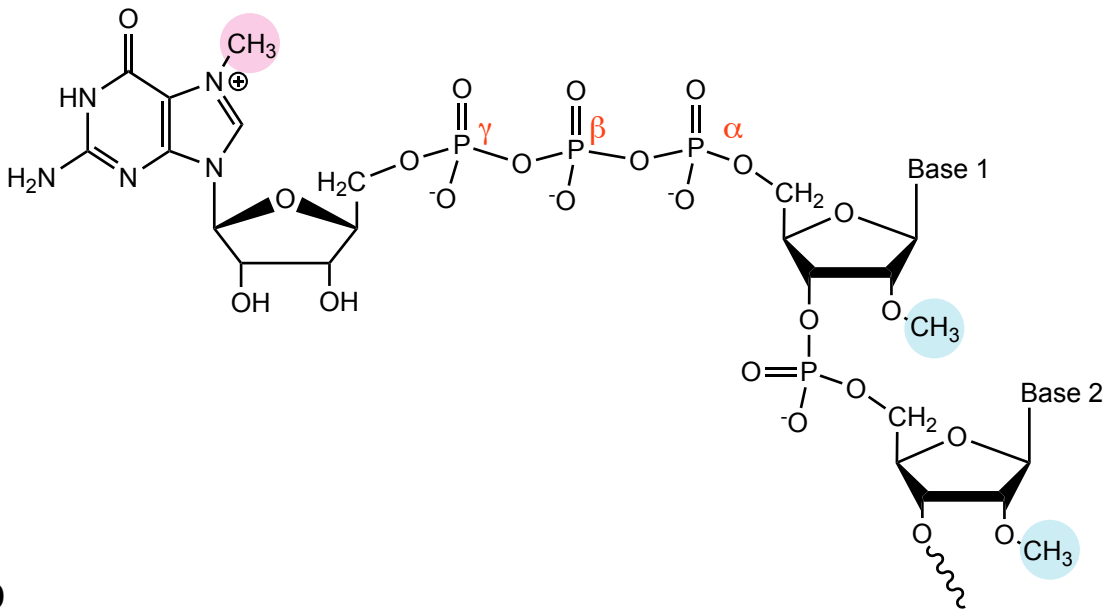
Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D., & Shatkin, A. J. (1997). Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc Natl Acad Sci U S A, 94*(24), 12898-12903. doi:10.1073/pnas.94.24.12898

Zamudio, J. R., Mittra, B., Campbell, D. A., & Sturm, N. R. (2009). Hypermethylated cap 4 maximizes Trypanosoma brucei translation. *Mol Microbiol, 72*(5), 1100-1110. doi:10.1111/j.1365-2958.2009.06696.x

Zinshteyn, B., Rojas-Duran, M. F., & Gilbert, W. V. (2017). Translation initiation factor eIF4G1 preferentially binds yeast transcript leaders containing conserved oligo-uridine motifs. *RNA, 23*(9), 1365-1375. doi:10.1261/rna.062059.117

Zuberek, J., Kubacka, D., Jablonowska, A., Jemielity, J., Stepinski, J., Sonenberg, N., & Darzynkiewicz, E. (2007). Weak binding affinity of human 4EHP for mRNA cap analogs. *RNA, 13*(5), 691-697. doi:10.1261/rna.453107

a


CH3
HN
N
H2N
N
N
O
H2C
O
P
O
-O
γ
P
O
-O
β
P
O
-O
α
O
CH2
Base 1
O
O
CH3
O
P
O
-O
CH2
Base 2
O
O
CH3
OH
OH

b

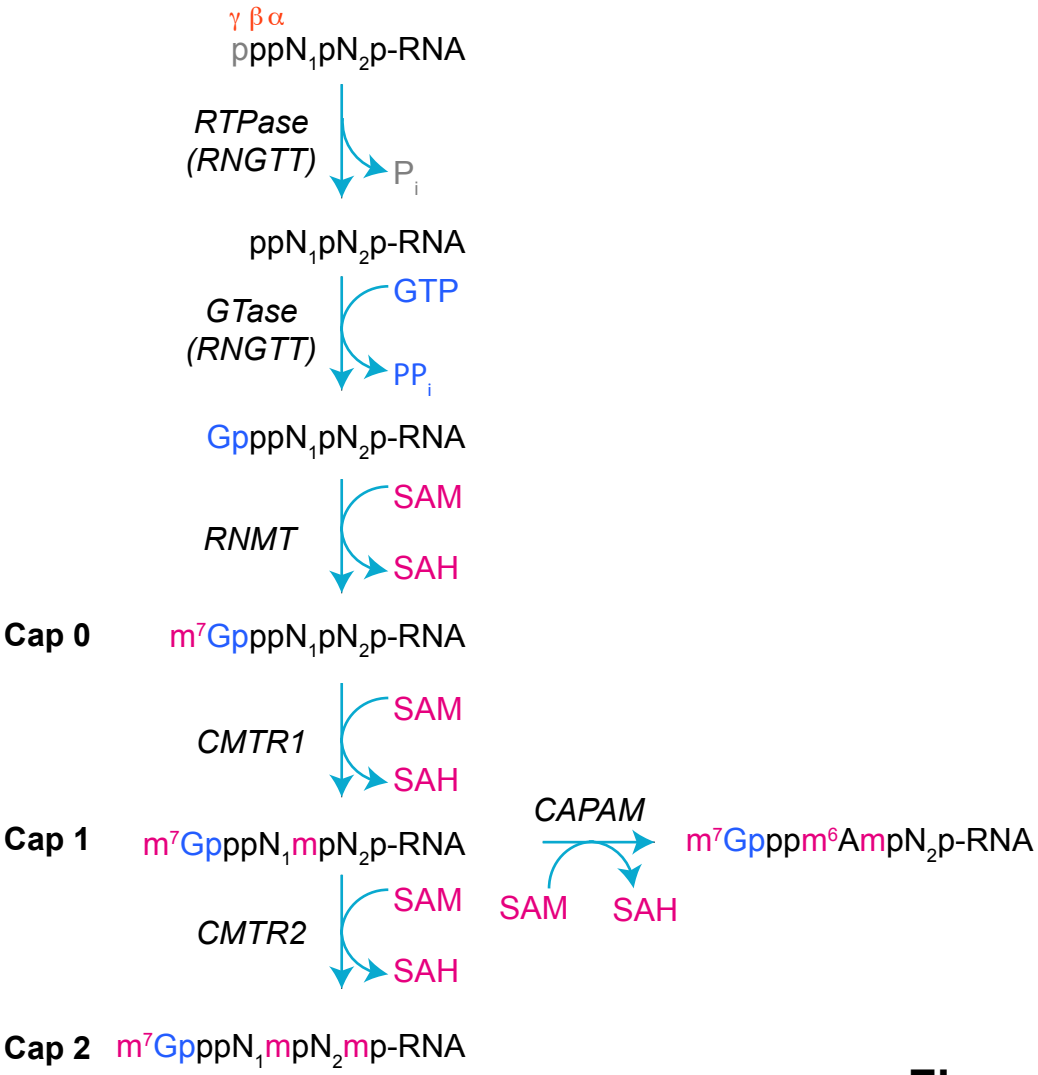

γ β α
pppN1pN2p-RNA
RTPase
(RNGTT)
Pi
ppN1pN2p-RNA
GTase
(RNGTT)
GTP
PPi
GpppN1pN2p-RNA
RNMT
SAM
SAH
Cap 0
m7GpppN1pN2p-RNA
CMTR1
SAM
SAH
Cap 1
m7GpppN1mpN2p-RNA
CMTR2
SAM
SAH
Cap 2
m7GpppN1mpN2mp-RNA
CAPAM
SAM
SAH
m7Gpppm6AmpN2p-RNA

Figure 1

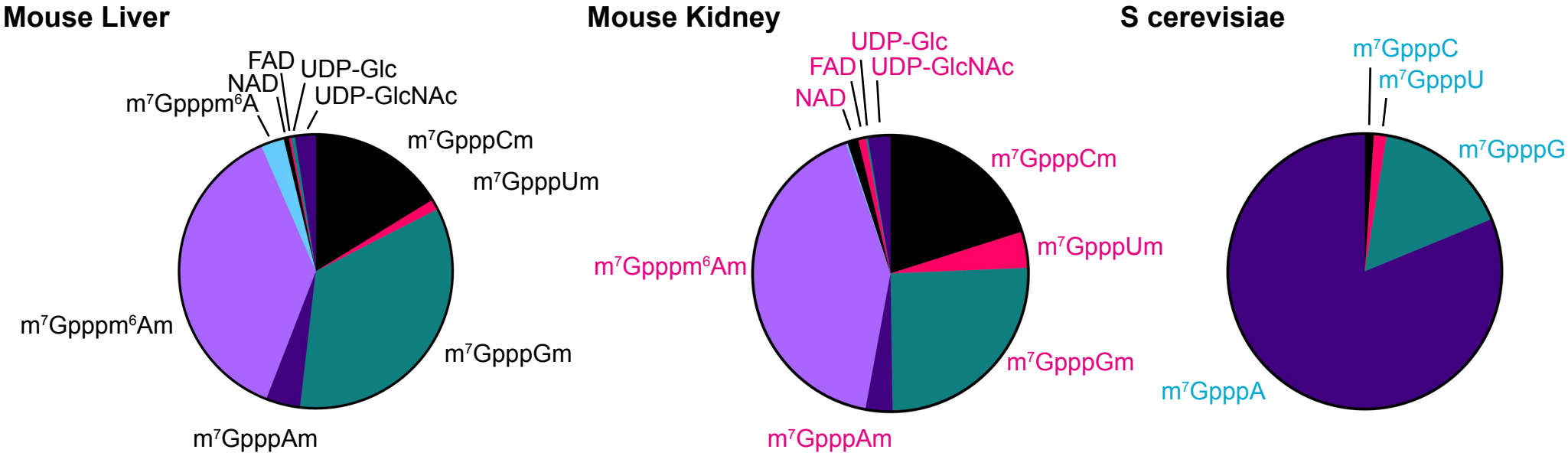
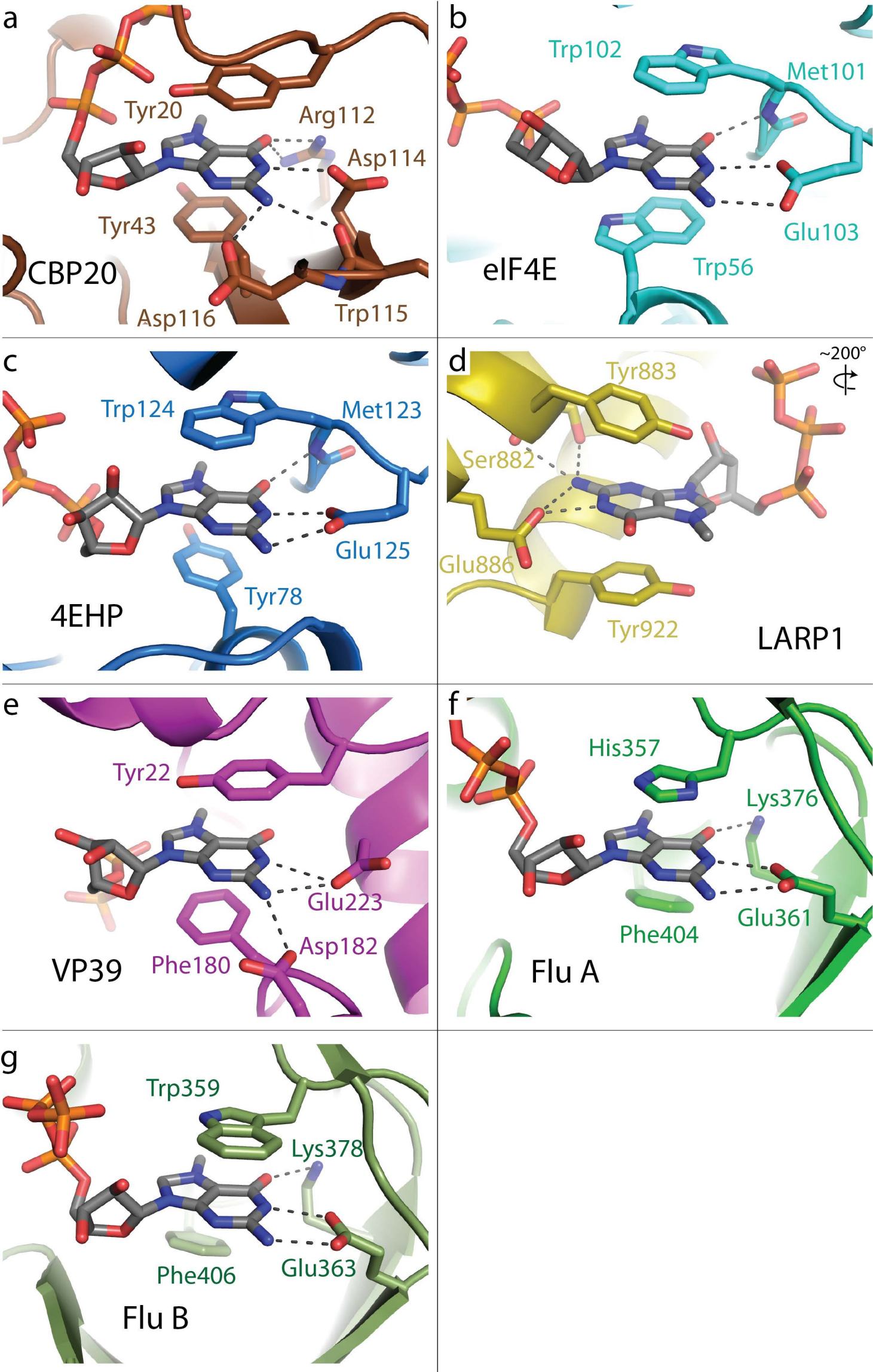
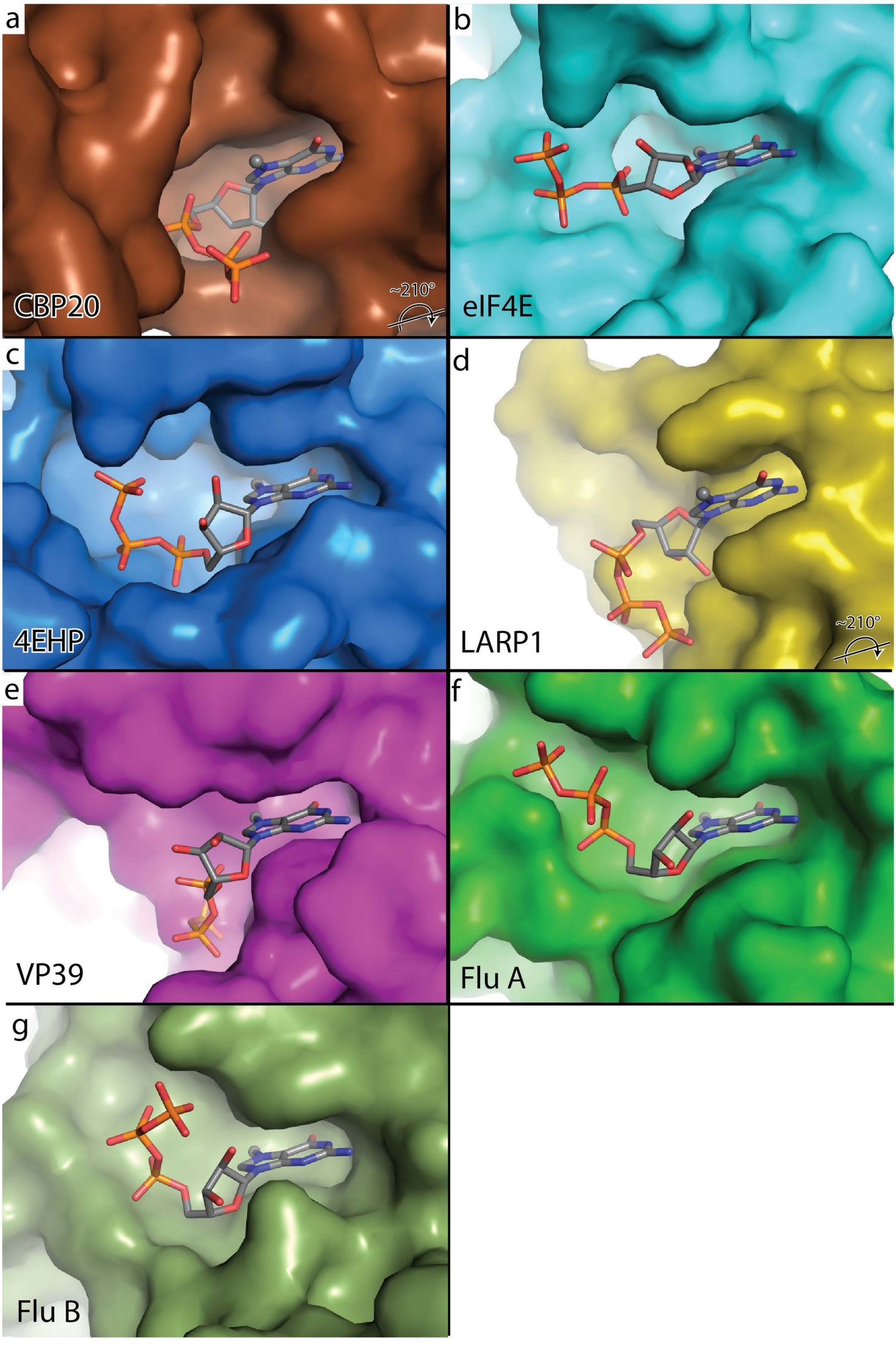
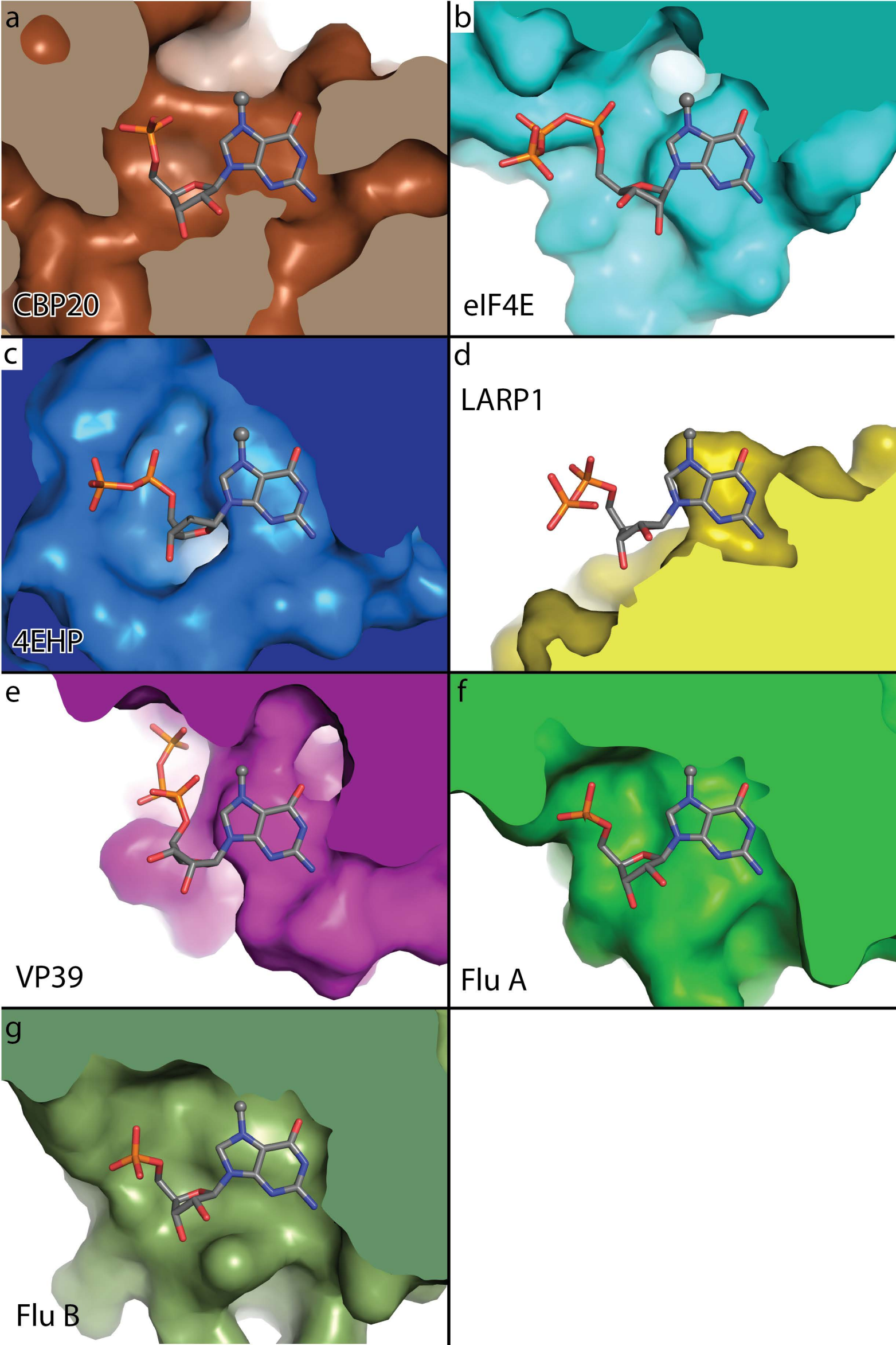

Mouse Liver
FAD
NAD
UDP-Glc
UDP-GlcNAc
m7Gpppm6A
m7GpppCm
m7GpppUm
m7Gpppm6Am
m7GpppGm
m7GpppAm
Mouse Kidney
UDP-Glc
FAD
UDP-GlcNAc
NAD
m7GpppCm
m7GpppUm
m7Gpppm6Am
m7GpppGm
m7GpppAm
S cerevisiae
m7GpppC
m7GpppU
m7GpppG
m7GpppA

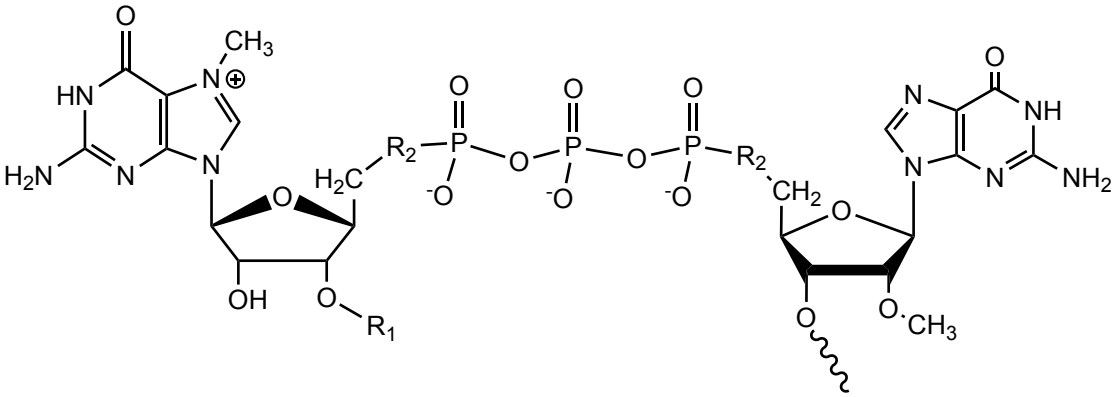
Figure 2


a
Tyr20
Arg112
Asp114
Tyr43
CBP20
Asp116
Trp115
b
Trp102
Met101
Glu103
eIF4E
Trp56
c
Trp124
Met123
Glu125
Tyr78
4EHP
d
~200°
Tyr883
Ser882
Glu886
Tyr922
LARP1
e
Tyr22
Glu223
Asp182
VP39
Phe180
f
His357
Lys376
Glu361
Phe404
Flu A
g
Trp359
Lys378
Phe406
Glu363
Flu B


a
CBP20
~210°
b
eIF4E
c
4EHP
d
LARP1
~210°
e
VP39
f
Flu A
g
Flu B


a
CBP20
b
eIF4E
c
4EHP
d
LARP1
e
VP39
f
Flu A
g
Flu B

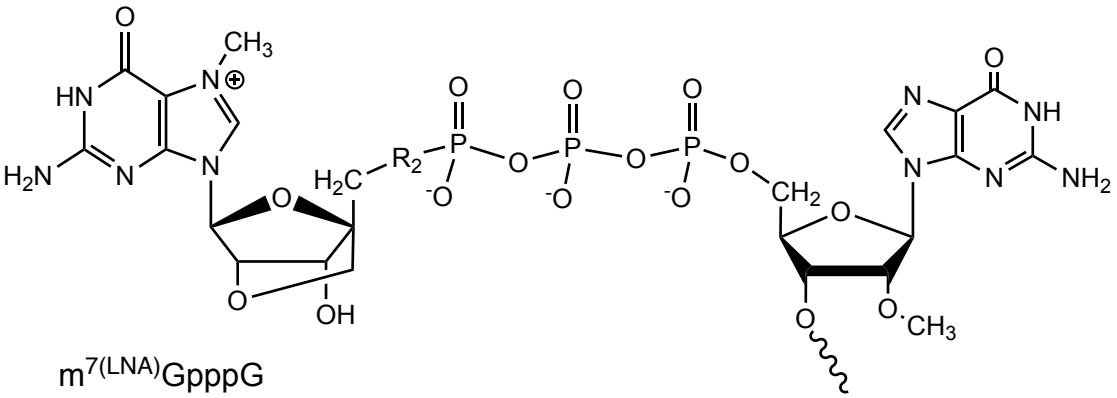
a



$m_2^{7,3'\text{-}O}$GpppG (ARCA): R1 = CH3, R2 = O

Phosphorothiolate Cap: R1=OH, R2 = S

b



$m^{7(LNA)}$GpppG

c

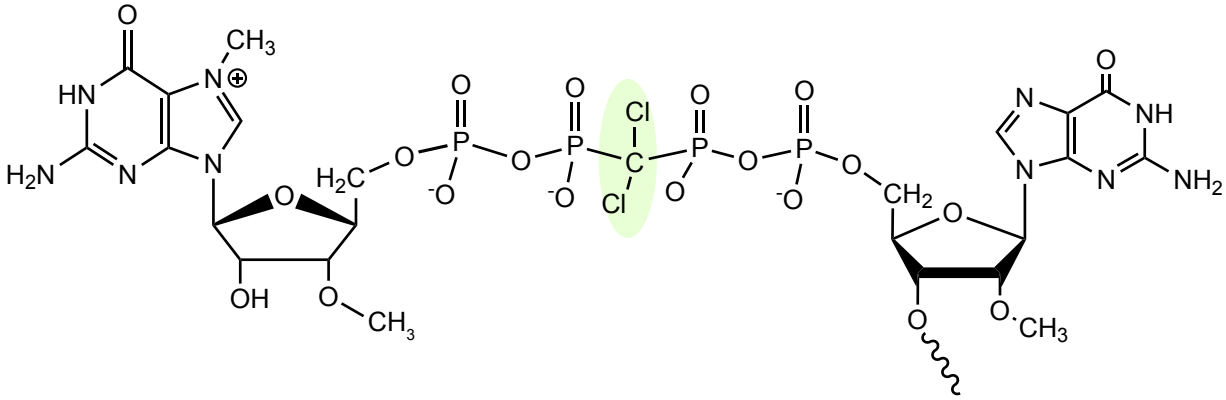


Figure 6