Structural insights into the transcriptional and translational roles of Ebp1

Tom P. Monie[1], Andrew J. Perrin[1], James R. Birtley[1], Trevor R. Sweeney[1], Ioannis Karakasiliotis[3], Yasmin Chaudhry[3], Lisa O. Roberts[4], Stephen Matthews[2], Ian G. Goodfellow[3]and Stephen Curry[1]*

[1]Division of Cell and Molecular Biology and [2]Division of Biomolecular Science, Imperial College, South Kensington Campus, London SW7 2AZ, UK.

[3]Department of Virology, Faculty of Medicine, Imperial College London, St Mary's Campus, London W2 1PG, UK

[4]School of Biomedical and Molecular Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK

*Corresponding author. Email: s.curry@imperial.ac.uk; Tel: +44-20-7594-7632; Fax: +44-20-7589-0191.

Journal Categories: Structural Biology / RNA

Running Title: Structure of Ebp1

Abstract

The ErbB3 binding protein 1 (Ebp1) is an important regulator of transcription, affecting eukaryotic cell growth, proliferation, differentiation and survival. Ebp1 can also affect translation and co-operates with the polypyrimidine tract binding protein (PTB) to stimulate the activity of the internal ribosome entry site (IRES) of foot-and-mouth disease virus (FMDV). We report here the crystal structure of murine Ebp1 (p48 isoform), providing the first glimpse of the architecture of this versatile regulator. The structure reveals a core domain that is homologous to methionine aminopeptidases, coupled to a C-terminal extension that contains important motifs for binding proteins and RNA. It sheds new light on the conformational differences between the p42 and p48 isoforms of Ebp1, the disposition of the key protein-interacting motif (^{354}LKALL^{358}) and the RNA binding activity of Ebp1. We show that the primary RNA binding site is formed by a Lys-rich motif in the C-terminus and mediates the interaction with the FMDV IRES. We also demonstrate a specific functional requirement for Ebp1 in FMDV IRES-directed translation that is independent of a direct interaction with PTB.

Keywords (5): Ebp1/ITAF$_{45}$/PA2G4, crystal structure, IRES, RNA-binding, translation initiation

Introduction

Eukaryotic protein translation is predominantly initiated via a cap-dependent scanning mechanism (Pestova et al., 2001). However, during viral infection or cell death cap-dependent translation initiation may be disabled and cap-independent mechanisms adopted instead. Cap-independent initiation requires direct recruitment of ribosomal subunits to an RNA sequence known as an internal ribosome entry site (IRES). This element, which often contains significant secondary structure elements, is usually contained within the 5′-untranlated region of the mRNA (Hellen and Sarnow, 2001). Although IRESes were originally identified in the positive-sense, single-stranded RNA genomes of picornaviruses (Jang et al., 1988; Pelletier and Sonenberg, 1988), they have since been found in other RNA viruses (Hellen and Sarnow, 2001); it has been estimated that up to 10% of cellular mRNAs may contain IRES elements (Stoneley and Willis, 2004).

In most cases only specific subsets of the canonical eukaryotic translation initiation factors (eIFs) are used for cap-independent mechanisms (Jackson, 2005), but a variety of other host-cell proteins, termed IRES-specific cellular trans-acting factors (ITAFs), may also be required for efficient IRES function. For example, the rhinovirus IRES and several cellular IRESes require both the polypyrimidine tract binding protein (PTB) and upstream of N-ras (unr) for activity (Hunt et al., 1999; Mitchell et al., 2003), while the IRES from foot-and-mouth disease virus (FMDV)—another picornavirus—requires both PTB and the proliferation associated factor $ITAF_{45}$ (Pilipenko et al., 2000).

In general ITAFs are RNA binding proteins recruited from their normal cellular functions and are believed to help in folding or stabilising the active form of the IRES to stimulate translation initiation (Hellen and Sarnow, 2001; Stoneley and Willis, 2004). $ITAF_{45}$ was originally identified as a member of the PA2G4 family of proliferation regulated proteins (Radomski and Jost, 1995; Lamartine et al., 1997). However, the protein is now better known as Ebp1 (ErbB3 binding protein 1) as a result of more recent investigations that have revealed its interaction with the ErbB3 receptor (Yoo et al., 2000) and probed its role in mediating growth and proliferation signals. The Ebp1-ErbB3 interaction is regulated by the growth factor heregulin, which promotes the dissociation of Ebp1 and its relocation to the nucleus (Yoo et al., 2000) where it appears to function as a transcriptional repressor of genes regulated by E2F1 (Zhang et al., 2003; Zhang and Hamburger, 2004) and the androgen receptor (AR) (Zhang et al., 2002; Zhang et al., 2005a; Zhang et al., 2005b). This activity is linked to the ability of Ebp1 to inhibit proliferation and promote differentiation of breast and prostate cancer cell lines (Lessor et al., 2000; Zhang et al., 2005b). Lately, Ebp1 has also been shown to regulate cell growth and proliferation in plants (Horvath et al., 2006).

The transcriptional repression mediated by Ebp1 depends on its participation in multi-protein complexes that interact with DNA promoter sequences of regulated genes (Xia et al., 2001; Zhang et al., 2003). With the exception of the co-repressor Sin3A (Zhang et al., 2005a) details of its interactions with other proteins have yet to be precisely determined. More recently Ebp1 has also been ascribed a role in regulating cell survival since its interaction with Akt kinase suppresses apoptosis (Ahn et al., 2006; Liu et al., 2006). This interaction depends on the phosphorylation of Ebp1 on Ser 360, which is also required for its interaction with ErbB3 (Lessor and Hamburger, 2001; Liu et al., 2006).

In addition to its ability to interact with various proteins, Ebp1 binds DNA (Yamada et al., 1994; Zhang and Hamburger, 2004) and an array of RNA targets including the FMDV IRES (Pilipenko et al., 2000), the 3′-untranslated region of bcl-2 mRNA (Bose et al., 2006) and ribosomal RNA (Squatrito et al., 2004), suggesting that Ebp1 can also regulate protein synthesis at a number of different levels. A further dimension to this role has appeared

recently with a report that Ebp1 can promote translation initiation by interacting with PKR and inhibiting the phosphorylation of the eIF2α subunit of eIF2 (Squatrito et al., 2006).

Despite burgeoning interest in the role of Ebp1 in several important cellular functions, little is known about either the molecular structure of the protein or the interactions involved in ligand binding. To address these questions we have determined the crystal structure of Ebp1. We show that, as predicted (Pilipenko et al., 2000), the protein is structurally homologous to the type II methionine aminopeptidases (MAPs) but has a novel C-terminal extension containing binding motifs for proteins and RNA. Moreover, we provide direct evidence, using siRNA knockdown, of the specific involvement of Ebp1 in FMDV IRES mediated translation initiation and we have determined that the ability of Ebp1 to bind to the FMDV IRES maps to a lysine-rich sequence in the C-terminus of the protein**.** Together these findings give new insights into the structure and function of Ebp1 and provide a valuable framework for further dissection of its molecular mechanism.

Results

Crystal structure of Ebp1

As an aid to crystallisation, sequence alignments with homologous type II MAPs (Pilipenko et al., 2000) of known structure were used to identify potentially unstructured termini of Ebp1 (Supp. Fig. S1). Following a series of deletion mutagenesis experiments, the construct Ebp1(8-360), which lacks the first seven and the last 34 amino acids of the protein, was found to be the most complete fragment that could be over-expressed solubly in *E. coli* and crystallised (Supp. Table S1). The 2.5 Å crystal structure of Ebp1(8-360) was phased by molecular replacement and refined to yield an $R_{free}$of 21.6% (Materials and methods; Table I). The refined model contains all the Ebp1 residues present in the construct.

As expected, the protein fold is very similar to the two-domain structure of type II MAPs. The main or 'catalytic' domain (non-functional in Ebp1 - see below) is conserved in type I and II MAPs in prokaryotes, archaea and eukaryotes. In Ebp1 it is composed of residues 8-243 and 307-337 and adopts a 'pita-bread' fold (Bazan et al., 1994) formed from two anti-parallel β-sheets each flanked by a pair of α-helices (Figure 1A).This domain has a curved, globular structure and in MAPs the active site is contained within a deep pocket on its concave surface. The smaller insertion domain, one of the distinguishing features of type II MAPs (Liu et al., 1998; Tahirov et al., 1998), normally has a βαβααβ topology and forms a small anti-parallel β-sheet backed by a three-helix bundle. However, in Ebp1 this domain (residues 244-306) is slightly simplified by deletion of the middle helix (Fig. 1A).

Superposition of Ebp1 with the structures of type II MAPs using SSM (Krissinel and Henrick, 2004) revealed rms differences in C_α positions of 1.7 Å for both human MAP2 (hMAP2; PDB ID 1kq9) and *P. fu.* MAP (PDB ID 1xgs) emphasising the overall similarity of these proteins. Nevertheless there are several notable differences between Ebp1 and type II MAPs. In particular, Ebp1 contains two extended loop insertions (loop 1: residue 63-71; loop 2: residues 128-134), one helical insertion (residues 205-213) and a C-terminal extension (residues 338-394) that are not present in type II MAPs (Fig. 1B). These features are generally conserved in metazoan orthologues of Ebp1 (Supp. Fig. S2) and all occur on the surface of the protein. The novel C-terminal extension of Ebp1 wraps around the concave surface of the main domain and is stabilized by interactions with loop 2 and hydrophobic contacts between the ^{354}LxxLL^{358} motif in the C-terminal helix and the body of the protein

(see Discussion). In contrast, hMAP2 contains a long N-terminal extension (residues 1-150) and a β-hairpin structure that are absent from Ebp1 and *P. fu.* MAP (Fig. 1B).

Despite the overall similarity with type II MAPs, the structure of Ebp1 suggests that it does not retain methionine aminopeptidase activity since key residues in the active site are not conserved. This prediction was verified experimentally in methionine aminopeptidase assays that confirmed the activity of hMAP2 but found no detectable activity for Ebp1 (Supp Fig. S3). In functional MAPs the active site contains a pair of divalent cations co-ordinated by a strictly conserved set of residues (Asp 251, Asp 262, His 331, Glu 364, Glu 459 - hMAP2 numbering) and a pair of flanking His residues (His 231 and His 339), that aid catalytic activity (Lowther and Matthews, 2002) (Fig. 1C). The crystal structure of Ebp1 reveals that this active site architecture is only partially conserved; in particular, Asp 262, Glu 459 and His 339 from human MAP2 are substituted by Asn 120, Lys 320 and Asp 196 respectively in Ebp1 (Fig. 1D, E). The electron density map shows that these changes are sufficient to abrogate metal binding, thus accounting for the absence of aminopeptidase activity. This is most likely due to the fact that the sidechain amino group of Lys 320 in Ebp1 occupies a position equivalent to one of the two cation sites in hMAP2, though other differences at the active site may also contribute to disruption of metal binding (Fig. 1D, E).

A Lys-rich motif in the C-terminus of Ebp1 binds RNA

Sequence alignments with homologous MAPs (Supp. Fig. S1) revealed two lysine rich regions, one (residues 65-72) within loop 1 and one (residues 364-373) within the C-terminal extension, both of which are well conserved in Ebp1 orthologues (Supp. Fig. S2). To probe the role of these Lys-rich motifs in RNA binding, electrophoretic mobility shift assays (EMSAs) were performed using radiolabelled full-length FMDV IRES RNA (Materials and Methods). Full-length Ebp1 bound the FMDV IRES with an apparent K_D ~ 1-2 μM (Fig. 2A). Mutation of the N-terminal lysine rich motif ^{65}KKEKEMKK^{72} to ^{65}SSSSSSSS^{72} in the context of the full-length protein [Ebp1(1-394LM)] had no effect on RNA binding activity (Fig. 2B). In marked contrast, the construct Ebp1(1-360), in which the C-terminal Lys-rich sequence has been removed, was unable to bind the FMDV IRES (Fig. 2D) although a lesser deletion [Ebp1(1-375)] which retains the Lys-rich motif had wild-type binding activity (Fig. 2C). These results demonstrate the critical importance of the sequence ^{361}SAS<u>RKTQKKKKKK</u>AS^{375}, containing the C-terminal Lys-rich region (underlined), for binding to FMDV IRES RNA by Ebp1.

Toe-printing studies of Ebp1 previously indicated that the protein bound to a site located at the basal region of domain I of the FMDV IRES (Pilipenko et al., 2000). To further probe the specificity of the Ebp1-IRES interactions, we compared the binding of the protein to RNA transcripts encoding domains H (nt 280-336), I (nt 337-551) and JKL (nt 552-716) of the FMDV IRES. EMSAs performed using full-length Ebp1 indicated that the protein binds with broadly similar affinities (apparent K_D ~ 0.2-1.0 μM) to all three RNA constructs (Fig. 2E-G), although some minor variations were observed. For example, binding of Ebp1 to domains I and JKL was about 2-fold stronger than to domain H and the interaction of Ebp1 with domains JKL resulted in the formation of two distinct complexes with different affinities (Fig. 2G). Overall however, these data suggest that under our experimental conditions the binding of RNA by Ebp1 is non-specific. Further experiments confirmed this idea by showing that Ebp1 bound to the FMDV and EMCV IRESes (which differ in nucleotide sequence by about 50% (Drew and Belsham, 1994)) with equal affinity (Supp. Fig. S4); this result is consistent with previous reports that Ebp1 could bind to both IRESes (Pilipenko et al., 2000). Moreover, the non-specific nature of the interaction is underscored by the finding

that the protein has broadly similar affinity for single and double-stranded RNA and for double-stranded DNA targets (Supp. Fig. S5).

Nevertheless, for all nucleic acid targets tested, deletion of the C-terminal Lys-rich sequence severely impaired binding of Ebp1, confirming that this is the binding motif for both RNA and DNA (Fig. 2; Supp. Figs S4, S5). The dependence of Ebp1 on such a motif readily accounts for its apparent lack of sequence specificity in binding RNA or DNA, at least in assays using purified components. This nevertheless raises interesting questions regarding the functional specificity of the protein (see Discussion).

The interaction between the ITAFs PTB and Ebp1 is RNA dependent

Ebp1 appears to operate synergistically with PTB in stimulating formation of 48S initiation complexes on the FMDV IRES (Pilipenko et al., 2000). We initially tried to probe the molecular basis of this phenomenon using dual-protein EMSAs in which RNA was pre-bound to either Ebp1 or PTB and then incubated with the other protein, but this invariably resulted in extensive aggregation of the sample (data not shown). We therefore attempted to detect a direct, RNA-independent Ebp1-PTB interaction using pull-down assays. Ebp1 constructs were transcribed and translated *in vitro* (Fig. 3A) and then pulled-down with either GST-PTB or, as a negative control, GST-Sex Lethal (GST-Sxl). The construct Raver1(1-455) was included as a positive control with known RNase sensitivity in its interaction with PTB (Rideau et al., 2006). Both Ebp1(1-394) and Ebp1(1-375) were pulled down with GST-PTB. However, this interaction was clearly abrogated by treatment with RNase to remove mRNAs generated during *in vitro* transcription (Fig.3B, bottom panel). Moreover, the construct Ebp1(1-360), which lacks the C-terminal RNA binding motif, was not pulled-down significantly above background levels, even in the absence of RNase treatment (Fig.3B, bottom panel). These results indicate that the positive interactions observed with Ebp1(1-394) and Ebp1(1-375) do not result from direct interaction between PTB and Ebp1 but are the consequence of both proteins binding to the same RNA molecule, consistent with the fact that both proteins have RNA-binding activity. It remains possible nevertheless that RNA-binding by either protein may induce conformational changes that permit formation of direct protein-protein contacts.

Role of Ebp1 in translation initiation

Previous work demonstrating a role for Ebp1 in stimulating 48S initiation complex formation on the FMDV IRES relied exclusively on the use of *in vitro* biochemical toe-printing analysis (Pilipenko et al., 2000). To confirm a role for Ebp1 in FMDV IRES mediated translation initiation, we examined the effect of Ebp1 specific siRNAs on IRES activity *in vivo* (Fig. 4). Transfection of human embryonic kidney cells with Ebp1-specific siRNAs resulted in a significant reduction of Ebp1 levels, whereas PTB or GAPDH levels were unaffected (Fig. 4B). RNAi mediated reduction of Ebp1 levels resulted in a 55% reduction in FMDV IRES activity whereas EMCV IRES and cap-dependent translation were unaffected (Fig 4A, C), consistent with previous observations on 48S complex formation (Pilipenko et al., 2000). This confirms that Ebp1 exhibits sequence-specific stimulation of the FMDV IRES *in vivo*.

Discussion

The crystal structure of Ebp1 provides new molecular insights into the transcriptional and translational roles of the protein, allowing us to re-interpret some of the existing functional data and to probe its mechanism more precisely. As anticipated from sequence alignments (Pilipenko et al., 2000), Ebp1 is structurally similar to the homologous type II MAPs (Fig. 1), though there are some significant differences. The structure reveals that the identity and conformation of the active site residues in MAPs is not maintained in Ebp1, resulting in the loss of aminopeptidase activity (Supp. Fig. S3. More significantly, the structure of Ebp1 shows that at least part of the novel C-terminal extension in Ebp1 (residues 338-360) lies in a relatively extended conformation across the concave dorsal surface of the protein and is apparently available for interactions with proteins and RNA (see below).

Predicted structural differences between the p42 and p48 isoforms of Ebp1

There are two major isoforms of Ebp1 (p42 and p48) which have distinct cellular functions (Liu et al., 2006). The longer p48 isoform (used in our study), which localises to the cytoplasm and nucleolus, suppresses apoptosis but promotes cellular differentiation. In contrast the shorter p42 isoform, which lacks the N-terminal 54 amino acids present in p48, is confined predominantly to the cytoplasm where it acts to promote differentiation and suppress proliferation (Liu et al., 2006). The crystal structure shows that the shorter isoform is missing the whole of helix α1 and half of α2 (Fig. 5A), structural features that are conserved in all known MAPs. The truncation of p42 removes K20 and K22, which form part of the nucleolar localisation signal (Squatrito et al., 2004). Such signals generally involve extended basic sequences (Fujiwara et al., 2006) and the structure of p48 Ebp1 shows that these residues are situated at solvent exposed positions close to the highly basic loop1 (^{63}IFKKEKEMK^{71} **;** Fig 5A); this suggests that although loop1 has no demonstrated role in RNA binding (Fig. 2), it may contribute to a bipartite nucleolar localisation signal.

The truncation of p42 also exposes of a large hydrophobic surface to solvent and may cause destabilisation of the protein, perhaps accounting for its significantly lower levels of expression (Liu et al., 2006). Consistent with this, we were unable to express Ebp1(55-394) in *E. coli* (Supp. Table S1). In eukaryotic cells the shorter isoform may be stabilised by refolding of the N-terminal region, perhaps involving the nearby C-terminal extension (Fig. 5A, B), or by recruitment of novel protein partners. It is tempting to speculate that the conformational changes associated with the absence of residues 1-54 account for the ability of p42 Ebp1, but not p48, to interact with ErbB3 (Liu et al., 2006).

Structure of the conserved LxxLL protein-binding motif in Ebp1

Ebp1 has an important role in transcriptional repression which is reported to involve interaction with components of repressive complexes such as Rb (Xia et al., 2001), AR (Zhang et al., 2002) and Sin3A (Zhang et al., 2005a). In each case the C-terminal 70-80 residues of Ebp1, which incorporates the novel extension identified in the structure (residues 338-394; Fig. 1), have been shown to be required for these interactions. However, it is not yet clear if these interactions all involve direct contact with Ebp1 or are mediated by other components of the complex. Evidence for a direct protein-protein contact has been reported only for the Sin3A-Ebp1 interaction (Zhang et al., 2005a). In particular it has been shown that this is primarily mediated by the conserved LxxLL motif in the C-terminus of Ebp1 (^{354}LKALL^{358}) (Zhang et al., 2005a). This motif has also been shown to be important for the interaction of Ebp1 with AR, which is known to have a binding site for the LxxLL motif (Hur

et al., 2004); however, a direct Ebp1-AR interaction remains to be demonstrated experimentally.

The crystal structure of Ebp1 reveals that the ^{354}LKALL^{358} sequence adopts the helical structure common for LxxLL motifs, allowing all three Leu sidechains to interact with a shallow hydrophobic pocket on the back-side of Ebp1 (Fig. 5C). The binding pocket for the motif in Ebp1 is primarily formed by amino acid sidechains from helices α3 and α4 that are completely apolar or have substantial hydrophobic moieties (L148, A152, R155, L156, A169, K172 and V173). Interestingly helices α3 and α4 adopt a different architecture to that observed in the helical LxxLL binding sites of nuclear receptors (Fig. 5C,D) revealing that the LxxLL motif can interact with structurally diverse binding sites. The interaction appears to at least partially conserve the "charge-clamp" mechanism previously detected in the ligand binding domains of nuclear receptors that bind this motif (Nolte et al., 1998; Hur et al., 2004; Plevin et al., 2005). Normally, a basic and an acidic amino acid sidechain serve to neutralise the partial charges associated with the helix dipole (Plevin et al., 2005). In the case of Ebp1, Arg 155 is positioned to interact with the negatively charged C-terminal end of the LxxLL helix, while Asp 351 interacts with the positively charged N-terminal end. Note however that Asp 351 is immediately upstream in the primary sequence from the helix so, while this interaction can stablise the helix, it may not otherwise contribute significantly to its interaction with the body of the protein. This weakening of the charge-clamp may be advantageous since it is evident that the ^{354}LKALL^{358} motif from Ebp1 must peel away from the protein surface to interact with a binding partner such as Sin3A (Zhang et al., 2005a). Intriguingly, dissociation of the ^{354}LKALL^{358} helix from the dorsal surface of Ebp1 makes the binding pocket available for another LxxLL motif; however, a binding partner that might take advantage of this site has not yet been identified.

The structure also shows that Ser 360, the phosphorylation site of PKC (Liu et al., 2006), is solvent exposed and lies very close to the C-terminal end of the LxxLL helix (Fig. 5D,E). This residue is therefore likely to contact binding partners that interact with the LxxLL motif, suggesting that such interactions may be regulated by phosphorylation. Phosphorylation of this residue enhances binding to ErbB3 and Akt (Lessor and Hamburger, 2001; Liu et al., 2006), but further work is required to establish a definitive role for the LxxLL motif in interactions with these binding partners.

Ebp1 does not bind to PTB

Ebp1 operates synergistically with PTB in forming 48S initiation complexes on the FMDV IRES (Pilipenko et al., 2000). Although PTB is known to interact with a variety of protein partners, particularly in its role as a regulator of alternative splicing [*e.g.* Raver1 (Gromak et al., 2003), Nova (Polydorides et al., 2000)], we were unable to detect a direct PTB-Ebp1 interaction. This is consistent with the fact that Ebp1 lacks the conserved [S/G][I/L]LGxxP PTB-binding motif that has recently been identified within Raver1 (Rideau et al., 2006). Our results do show that PTB and Ebp1 can interact by binding to the same RNA molecule so their synergistic effect on FMDV translation initiation is likely to derive from concerted effects on the IRES structure.

RNA binding activity of Ebp1

The ability of Ebp1 (and its orthologues) to bind nucleic acid is well documented (Yamada et al., 1994; Radomski and Jost, 1995; Pilipenko et al., 2000; Squatrito et al., 2004; Zhang and Hamburger, 2004; Bose et al., 2006; Squatrito et al., 2006). With regard to RNA binding,

previous sequence analyses suggested that Ebp1 contains two conserved RNA binding domains, a σ^{70}-like domain (residues 46-64) (Squatrito et al., 2004) and a double-stranded RNA binding domain (residues 91-156) (Squatrito et al., 2006), both of which were apparently confirmed when deletion of these domains abrogated RNA binding. Strikingly however, the crystal structure reveals that neither of these domains is present in Ebp1 (Supp. Fig. S6), but shows that deletion of the residues predicted to form them would cause loss of function by severely disrupting the protein fold. This is in fact the second time that the predicted helix-turn-helix motif of so-called σ^{70}-like domains has been contradicted by a crystal structure (Ng et al., 2005), thus calling into question the functionality of this hypothetical motif (Wehner and Baserga, 2002).

In contrast to the above studies, we showed that a lysine rich sequence (^{364}RKTQKKKKKK^{373}), situated within the portion of the C-terminus that had to be removed for crystallisation (suggesting that it may be relatively unstructured in the absence of ligand), is likely to be the key determinant for RNA binding (Fig. 2A-D). It is interesting to note that mutations within this motif partially inhibit nucleolar localisation (Squatrito et al., 2004), suggesting that RNA binding activity contributes to this function.

The attribution of the primary RNA binding activity to a Lys-rich sequence suggests that RNA binding by isolated Ebp1 may not be sequence specific since a series of basic amino acids sidechains may primarily interact with the phosphate backbone of RNA ligands. Consistent with this idea, we observed similar binding by Ebp1 to a variety of RNA and DNA targets, including FMDV and EMCV IRES RNA (Supp. Fig. S4), distinct sub-domains of the FMDV IRES (Fig. 2E-G) and ssRNA, dsRNA and dsDNA samples (Supp. Fig. S5).

The absence of detectable specific interactions with RNA or PTB throws into sharp relief the observation that Ebp1 nevertheless has a specific stimulatory effect on FMDV IRES-directed translation initiation but not on translation initiation directed by the closely related EMCV IRES (Figure 4). At present it is unclear how this functional specificity arises. Specific activation of the FMDV IRES by Ebp1 may depend on the intact RNA structure and also the presence of other ITAFs, including PTB. Consistent with this line of thinking, it has been suggested previously that Ebp1 and PTB may stabilise a conformation of the FMDV IRES that favours eIF4A/eIF4G binding (Pilipenko et al., 2000), thereby enhancing translation initiation, though alternative modes of ITAF action may operate (Mitchell et al., 2003). Indeed, the functional specificity probably only emerges once the full complement of initiation factors has assembled on the IRES.

The plurality of interactions with proteins and RNA made by Ebp1 means that further work is required to dissect its roles both as a transcriptional regulator — and potential target for prostate cancer therapy (Zhang et al., 2005b) — and as an enhancer of IRES-mediated translation initiation. The structure reported here has begun to illuminate these functions in greater detail and will facilitate the design of more incisive functional investigations.

Materials and Methods

Plasmid construction: The plasmids pQE9-PTB1 (Conte et al., 2000) and pSP64 FMDV polyA (Stassinopoulos and Belsham, 2001) have been described previously. Ebp1 expression plasmids encoding the larger p48 isoform of the protein (Liu et al., 2006) were generated following cDNA amplification by PCR from the plasmid pET(His_6-$ITAF_{45}$) (Pilipenko et al., 2000) using the primers detailed in Table S2. cDNA was digested with either *Bsp*HI/*Hind*III for constructs beginning at residue 1, or *Nco*I/*Hind*III for constructs beginning at residue 8,

and inserted into *Nco*I/*Hind*III digested pETM-11 (EMBL) creating constructs with an N-terminal TEV cleavable His-Tag (Table S1). Ebp1(1-394LM) was generated by overlap PCR in the context of Ebp1(1-394) using the primers described (Supp. Table S2).

Protein expression and purification: PTB proteins were expressed in SG13009 *E. coli* and purified on TALON resin (BD Biosciences) essentially as described previously (Conte et al., 2000). Ebp1 constructs were expressed in Rosetta™(DE3) *E.coli* (Novagen). Cell pellets were lysed by sonication in 50 mM HEPES (pH7.0), 500 mM NaCl, 5 mM β-mercaptoethanol supplemented with 0.1% Triton X-100 and a protease inhibitor cocktail (Sigma) and the protein purified using TALON (BD Biosciences) affinity resin. Samples for crystallographic screening were further purified as follows. The His-tag was removed with His-tagged TEV protease (Nunn et al., 2005) (1:10 TEV:target ratio) during overnight dialysis against 25 mM Tris (pH 8.0), 100mM NaCl, 5 mM β-mercaptoethanol. The cleaved His-tag and the His-TEV were subsequently removed using TALON resin. Further impurities were removed by passage through a 6 ml ResourceQ (Amersham) column (which did not bind Ebp1) in 25 mM Tris (pH 8.0), 100mM NaCl, 10 mM β-mercaptoethanol followed by gel filtration using a HR 30/10 Superdex75 (Amersham) column in 25 mM HEPES (pH 7.7), 100 mM NaCl, 10 mM β-mercaptoethanol. Purified Ebp1 was concentrated to 10 mg/mL in 10 mM HEPES (pH 7.7), 10 mM β-mercaptoethanol. GST-PTB (Pérez et al., 1997)and GST-Sxl were expressed in *E. coli* and purified using glutathione sepharose 4B resin (Amersham) following the manufacturer's protocol.

Crystallisation and data collection: Crystals of Ebp1 (residues 8-360) were grown in 4-5 days at 4°C from vapour-diffusion sitting drops at a protein concentration of 10 mg/mL in 10 mM HEPES (pH 7.7), 10 mM β-mercaptoethanol. The reservoir solution contained 20% PEG 8000 in 0.1 M HEPES (pH 7.5) and the drop was supplemented with 0.56 mM CYMAL-6. Diffraction data to 2.5 Å resolution, collected at room temperature using a single capillary-mounted crystal on beamline X11 at EMBL/DESY Hamburg, were processed and scaled using the CCP4 program suite (Collaborative Computer Project No. 4, 1994). The crystals belong to space-group $P2_12_12_1$ and have one molecule in the asymmetric unit. Data collection statistics are summarised in Table 1.

Structure solution and refinement: The diffraction data were phased using the molecular replacement program PHASER v1.2 (Storoni et al., 2004) with a search model generated by superposition of the structures of the homologous human and Pfu MAPs (PDB IDs 1kq9 and 1xgs respectively). Regions with non-conserved sidechains were removed from the search models, which were weighted according to the degree of identity (human MAP - 18%; Pfu MAP - 22%). This yielded a single solution with a Z-score of 10. $3F_o$-$2F_c$ and F_o-F_c electron density maps calculated using this solution showed clear features not present in the phasing model and allowed initiation of model building using ARP/wARP (Morris et al., 2003) starting with a truncated Ebp1 homology model generated using SWISSMODEL (Schwede et al., 2003) from the superposition described above. Manual model building and correction of protein register errors introduced during sequence alignment were performed using O (Jones et al., 1991). Structure refinement was completed using CNS (Brunger et al., 1998).

RNA transcription and binding assays: FMDV IRES RNA (nts 280-717; numbering according to Pilipenko et al. (2000)) was generated using T7 RNA polymerase following linearization of pSP64-FMDV-polyA with EcoRI (Stassinopoulos and Belsham, 2001). Transcription templates for the FMDV IRES domain truncations were generated by PCR amplification from pSP64-FMDV-polyA with the primer pairs described in Table S2. All forward primers contained a T7 promoter. Radiolabelled RNA was produced either by incorporation of [α-^{32}P] UTP during transcription, or by addition of [γ-32P] ATP to the 5' end

of the RNA using T4 polynucleotide kinase (Promega). The resultant RNA transcripts corresponded to – Domain H (nts 280-336); Domain I (nts 337-551) and Domain JKL (nts 552-716).

Prior to use in EMSAs all RNA constructs were refolded by heat denaturation for 5 min at 70 °C followed by slow cooling to room temperature. EMSAs were performed essentially as described previously (Simpson et al., 2004) using 3-10 nM RNA and final binding conditions of 10 mM HEPES (pH 7.0), 3 mM MgCl2, 5 mM DTT, 100 mM KCl, 50 µg/ml tRNA, 40 µg/ml HSA and 5% glycerol. Samples were analysed in 1% 0.5xTBE agarose gels and visualised by phosphomaging with a Fuji FLA5000.

GST-pulldown assays: ^{35}S-labelled proteins were generated from plasmid DNA using coupled transcription-translation assays (Promega) as per the manufacturer's instructions. Equal activities of the ^{35}S-labelled proteins were rotated for 3 hrs at 4°C with 35 pmoles of either glutathione-Sepharose 4B (Amersham) bound GST-Sxl or GST-PTB fusion protein in pull-down buffer (20 mM HEPES (pH 7.9), 100 mM KCl, 0.5 mM DTT, 10% glycerol, 0.2% Tween-20). The total reaction volume was 500 µl. After incubation beads were transferred to Zeba^{TM} micro spin columns (Pierce) washed three times with pull-down buffer, and protein eluted with 2X SDS-sample buffer before analysis by 12% SDS-PAGE. Bound proteins were visualised by phosphoimaging with a Fuji FLA5000. For reactions requiring RNAse treatment 10 units of RNAse One (Promega) were added to the ^{35}S-labelled protein immediately post translation and a further 10 units to each individual reaction mix.

RNAi experiments and translation assays: siRNAs targeting either EGFP (target sequence GGCTACGTCCAGGAGCGCACC) or Ebp1 (target sequence GTGAGGTGGAAAGGCGTTT) were purchased from Eurogentec. siRNA mediated reduction of Ebp1 in human embryonic kidney cells (293) was carried out using Lipofectamine 2000 according to the manufacturers instructions. 48 hours post transfection with siRNAs, cap and IRES-dependant translation was monitored by transfection of 5' capped and 3' polyadenylated *in vitro* transcribed dicistronic reporter mRNA of the form cap-CAT-IRES-Luc (Roberts et al., 1998). *In vitro* transcribed reporter RNAs were capped by the incorporation of cap analogue [m^{7}G(5′)ppp(5′)G; Promega] during transcription and subsequently polyadenylated using *E.coli* poly-A polymerase (Ambion). CAT and luciferase levels were determined 24 hours post transfection with reporter RNA using a CAT-specific ELISA (Roche) and luciferase assay (Promega) respectively.

Acknowledgements

We thank C. Hellen and T. Pestova (SUNY, USA) for providing the original plasmid encoding Ebp1, C. Smith, C. Gooding and A. Rideau (Cambridge University, UK) for plasmids encoding GST-PTB, GST-Sxl and Raver1 1-455 and S. Djordjevic (University College London, UK) for the kind gift of the plasmid for His-TEV^{pro}. We are indebted to Brian Matthews and Anthony Addlagatta (University of Oregon, USA) for the hsMAP2 plasmid and to Núria Roqué-Rosell and Robin Leatherbarrow (Imperial College) for invaluable assistance with aminopeptidase assays. We are grateful to staff on beamline X11 at EMBL/DESY (Hamburg, Germany) for assistance with data collection. This work was funded by grant support from the Wellcome Trust (SC/SJM). IG is the recipient of a Wellcome Trust Career Development fellowship.

Supplementary information is available at *The EMBO Journal* Online.

References

Ahn, J.Y., Liu, X., Liu, Z., Pereira, L., Cheng, D., Peng, J., Wade, P.A., Hamburger, A.W. and Ye, K. (2006) Nuclear Akt associates with PKC-phosphorylated Ebp1, preventing DNA fragmentation by inhibition of caspase-activated DNase. *EMBO J*, **25**, 2083-2095.

Bazan, J.F., Weaver, L.H., Roderick, S.L., Huber, R. and Matthews, B.W. (1994) Sequence and structure comparison suggest that methionine aminopeptidase, prolidase, aminopeptidase P, and creatinase share a common fold. *Proc Natl Acad Sci U S A*, **91**, 2473-2477.

Bose, S.K., Sengupta, T.K., Bandyopadhyay, S. and Spicer, E.K. (2006) Identification of Ebp1 as a component of cytoplasmic bcl-2 mRNP (messenger ribonucleoprotein particle) complexes. *Biochem J*, **396**, 99-107.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54**, 905-921.

Collaborative Computer Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760-763.

Conte, M.R., Grüne, T., Ghuman, J., Kelly, G., Ladas, A., Matthews, S. and Curry, S. (2000) Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *EMBO J*, **19**, 3132-3141.

Drew, J. and Belsham, G.J. (1994) Trans complementation by RNA of defective foot-and-mouth disease virus internal ribosome entry site elements. *J Virol*, **68**, 697-703.

Fujiwara, T., Suzuki, S., Kanno, M., Sugiyama, H., Takahashi, H. and Tanaka, J. (2006) Mapping a nucleolar targeting sequence of an RNA binding nucleolar protein, Nop25. *Exp Cell Res*, **312**, 1703-1712.

Gromak, N., Rideau, A., Southby, J., Scadden, A.D., Gooding, C., Huttelmaier, S., Singer, R.H. and Smith, C.W. (2003) The PTB interacting protein raver1 regulates alpha-tropomyosin alternative splicing. *EMBO J*, **22**, 6356-6364.

Hellen, C.U. and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev*, **15**, 1593-1612.

Horvath, B.M., Magyar, Z., Zhang, Y., Hamburger, A.W., Bako, L., Visser, R.G., Bachem, C.W. and Bogre, L. (2006) EBP1 regulates organ size through cell growth and proliferation in plants. *EMBO J*, **25**, 4909-4920.

Hunt, S.L., Hsuan, J.J., Totty, N. and Jackson, R.J. (1999) unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA. *Genes Dev*, **13**, 437-448.

Hur, E., Pfaff, S.J., Payne, E.S., Gron, H., Buehrer, B.M. and Fletterick, R.J. (2004) Recognition and accommodation at the androgen receptor coactivator binding interface. *PLoS Biol*, **2**, E274.

Jackson, R.J. (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem Soc Trans*, **33**, 1231-1241.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. and 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**, 2636-2643.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.

Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*, **60**, 2256-2268.

Lamartine, J., Seri, M., Cinti, R., Heitzmann, F., Creaven, M., Radomski, N., Jost, E., Lenoir, G.M., Romeo, G. and Sylla, B.S. (1997) Molecular cloning and mapping of a human cDNA (PA2G4) that encodes a protein highly homologous to the mouse cell cycle protein p38-2G4. *Cytogenet Cell Genet*, **78**, 31-35.

Lessor, T.J. and Hamburger, A.W. (2001) Regulation of the ErbB3 binding protein Ebp1 by protein kinase C. *Mol Cell Endocrinol*, **175**, 185-191.

Lessor, T.J., Yoo, J.Y., Xia, X., Woodford, N. and Hamburger, A.W. (2000) Ectopic expression of the ErbB-3 binding protein ebp1 inhibits growth and induces differentiation of human breast cancer cell lines. *J Cell Physiol*, **183**, 321-329.

Liu, S., Widom, J., Kemp, C.W., Crews, C.M. and Clardy, J. (1998) Structure of human methionine aminopeptidase-2 complexed with fumagillin. *Science*, **282**, 1324-1327.

Liu, Z., Ahn, J.Y., Liu, X. and Ye, K. (2006) Ebp1 isoforms distinctively regulate cell survival and differentiation. *Proc Natl Acad Sci U S A*, **103**, 10917-10922.

Lowther, W.T. and Matthews, B.W. (2002) Metalloaminopeptidases: common functional themes in disparate structural surroundings. *Chem Rev*, **102**, 4581-4608.

Mitchell, S.A., Spriggs, K.A., Coldwell, M.J., Jackson, R.J. and Willis, A.E. (2003) The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. *Mol Cell*, **11**, 757-771.

Morris, R.J., Perrakis, A. and Lamzin, V.S. (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol*, **374**, 229-244.

Ng, C.L., Waterman, D., Koonin, E.V., Antson, A.A. and Ortiz-Lombardia, M. (2005) Crystal structure of Mil (Mth680): internal duplication and similarity between the Imp4/Brix domain and the anticodon-binding domain of class IIa aminoacyl-tRNA synthetases. *EMBO Rep*, **6**, 140-146.

Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K. and Milburn, M.V. (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature*, **395**, 137-143.

Nonato, M.C., Widom, J. and Clardy, J. (2006) Human methionine aminopeptidase type 2 in complex with L- and D-methionine. *Bioorg Med Chem Lett*, **16**, 2580-2583.

Nunn, C.M., Jeeves, M., Cliff, M.J., Urquhart, G.T., George, R.R., Chao, L.H., Tscuchia, Y. and Djordjevic, S. (2005) Crystal structure of tobacco etch virus protease shows the protein C terminus bound within the active site. *J Mol Biol*, **350**, 145-155.

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

Pérez, I., McAfee, J.G. and Patton, J.G. (1997) Multiple RRMs contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein. *Biochemistry*, **36**, 11881-11890.

Pestova, T.V., Kolupaeva, V.G., Lomakin, I.B., Pilipenko, E.V., Shatsky, I.N., Agol, V.I. and Hellen, C.U. (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A*, **98**, 7029-7036.

Pilipenko, E.V., Pestova, T.V., Kolupaeva, V.G., Khitrina, E.V., Poperechnaya, A.N., Agol, V.I. and Hellen, C.U. (2000) A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev*, **14**, 2028-2045.

Plevin, M.J., Mills, M.M. and Ikura, M. (2005) The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem Sci*, **30**, 66-69.

Polydorides, A.D., Okano, H.J., Yang, Y.Y., Stefani, G. and Darnell, R.B. (2000) A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc Natl Acad Sci U S A*, **97**, 6350-6355.

Radomski, N. and Jost, E. (1995) Molecular cloning of a murine cDNA encoding a novel protein, p38-2G4, which varies with the cell cycle. *Exp Cell Res*, **220**, 434-445.

Rideau, A.P., Gooding, C., Simpson, P.J., Monie, T.P., Lorenz, M., Huttelmaier, S., Singer, R.H., Matthews, S., Curry, S. and Smith, C.W. (2006) A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain. *Nat Struct Mol Biol*, **13**, 839-848.

Roberts, L.O., Seamons, R.A. and Belsham, G.J. (1998) Recognition of picornavirus internal ribosome entry sites within cells; influence of cellular and viral proteins. *RNA*, **4**, 520-529.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*, **31**, 3381-3385.

Simpson, P.J., Monie, T.P., Szendroi, A., Davydova, N., Tyzack, J.K., Conte, M.R., Read, C.M., Cary, P.D., Svergun, D.I., Konarev, P.V., Curry, S. and Matthews, S. (2004) Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure (Camb)*, **12**, 1631-1643.

Squatrito, M., Mancino, M., Donzelli, M., Areces, L.B. and Draetta, G.F. (2004) EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. *Oncogene*, **23**, 4454-4465.

Squatrito, M., Mancino, M., Sala, L. and Draetta, G.F. (2006) Ebp1 is a dsRNA-binding protein associated with ribosomes that modulates eIF2alpha phosphorylation. *Biochem Biophys Res Commun*, **344**, 859-868.

Stassinopoulos, I.A. and Belsham, G.J. (2001) A novel protein-RNA binding assay: functional interactions of the foot-and-mouth disease virus internal ribosome entry site with cellular proteins. *RNA*, **7**, 114-122.

Stoneley, M. and Willis, A.E. (2004) Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene*, **23**, 3200-3207.

Storoni, L.C., McCoy, A.J. and Read, R.J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr*, **60**, 432-438.

Tahirov, T.H., Oki, H., Tsukihara, T., Ogasahara, K., Yutani, K., Ogata, K., Izu, Y., Tsunasawa, S. and Kato, I. (1998) Crystal structure of methionine aminopeptidase from hyperthermophile, Pyrococcus furiosus. *J Mol Biol*, **284**, 101-124.

Wehner, K.A. and Baserga, S.J. (2002) The sigma(70)-like motif: a eukaryotic RNA binding domain unique to a superfamily of proteins required for ribosome biogenesis. *Mol Cell*, **9**, 329-339.

Xia, X., Cheng, A., Lessor, T., Zhang, Y. and Hamburger, A.W. (2001) Ebp1, an ErbB-3 binding protein, interacts with Rb and affects Rb transcriptional regulation. *J Cell Physiol*, **187**, 209-217.

Yamada, H., Mori, H., Momoi, H., Nakagawa, Y., Ueguchi, C. and Mizuno, T. (1994) A fission yeast gene encoding a protein that preferentially associates with curved DNA. *Yeast*, **10**, 883-894.

Yoo, J.Y., Wang, X.W., Rishi, A.K., Lessor, T., Xia, X.M., Gustafson, T.A. and Hamburger, A.W. (2000) Interaction of the PA2G4 (EBP1) protein with ErbB-3 and regulation of this binding by heregulin. *Br J Cancer*, **82**, 683-690.

Zhang, Y., Akinmade, D. and Hamburger, A.W. (2005a) The ErbB3 binding protein Ebp1 interacts with Sin3A to repress E2F1 and AR-mediated transcription. *Nucleic Acids Res*, **33**, 6024-6033.

Zhang, Y., Fondell, J.D., Wang, Q., Xia, X., Cheng, A., Lu, M.L. and Hamburger, A.W. (2002) Repression of androgen receptor mediated transcription by the ErbB-3 binding protein, Ebp1. *Oncogene*, **21**, 5609-5618.

Zhang, Y. and Hamburger, A.W. (2004) Heregulin regulates the ability of the ErbB3-binding protein Ebp1 to bind E2F promoter elements and repress E2F-mediated transcription. *J Biol Chem*, **279**, 26126-26133.

Zhang, Y., Wang, X.W., Jelovac, D., Nakanishi, T., Yu, M.H., Akinmade, D., Goloubeva, O., Ross, D.D., Brodie, A. and Hamburger, A.W. (2005b) The ErbB3-binding protein Ebp1 suppresses androgen receptor-mediated gene transcription and tumorigenesis of prostate cancer cells. *Proc Natl Acad Sci U S A*, **102**, 9890-9895.

Zhang, Y., Woodford, N., Xia, X. and Hamburger, A.W. (2003) Repression of E2F1-mediated transcription by the ErbB3 binding protein Ebp1 involves histone deacetylases. *Nucleic Acids Res*, **31**, 2168-2177.

Figure Legends

Figure 1: Molecular Structure of Ebp1 and comparison with a type II human methionine aminopeptidase (hMAP2). (**A**) Ribbon diagram of the crystal structure of Ebp1 8-360; α-helices are coloured pink and β-strands blue. (**B**) Superposition of Ebp1 in orange and hMAP2 in cyan [PDB 1kq9; (Nonato et al., 2006)]; insertions in Ebp1 are coloured dark red and those in hMAP2 dark blue. (**C-E**) Comparison of the active site of hMAP2 with the corresponding region in Ebp1. Sidechains of selected residues are shown as sticks for (**C**) hMAP2, (**D**) a superposition of hMAP2 and Ebp1 and (E) Ebp1. The colour coding is the same as for panel B.

Figure 2: The C-terminal Lys-rich motif of Ebp1 binds the FMDV IRES RNA. Electrophoretic mobility shift assays were performed using ^{32}P-labelled FMDV IRES RNA (nt 280-716) and two-fold dilutions of various Ebp1 proteins over the range 0.125-8 μM. (**A**) Ebp1(1-394); (**B**) Ebp1(1-394LM); (**C**) Ebp1(1-375) and (**D**) Ebp1(1-360). Ebp1 binds multiple regions of the FMDV IRES. Electrophoretic mobility shift assays were performed using ^{32}P-labelled RNA and two-fold dilutions of full-length Ebp1 (residues 1-394) over the range 0.125-32 μM. (**E**) FMDV IRES Domain H (nts 280 – 336); (**F**) Domain I (nts 337 – 551) and (**G**) Domain JKL (nts 552-716). (**H**) Schematic secondary structure of the FMDV IRES.

Figure 3: Ebp1 interacts with PTB in an RNA dependent manner **(A)** ^{35}S Methionine labelled coupled transcription-translation products for Ebp1 and control constructs. **(B)** GST pull-down assays between ^{35}S-labelled Ebp1 constructs and 35 picomoles of either GST-Sxl (top panel) or GST-PTB (bottom panel). Reactions were performed in the absence (-) and presence (+) of RNAse treatment.

Figure 4: Ebp1 is required for efficient translation from the FMDV IRES. **(A)** Schematic representation of reporter RNAs used to analyse the effect of Ebp1 specific siRNAs on IRES mediated translation. **(B)** Western blot analysis of cells transfected with either EGFP or Ebp1 specific siRNAs. Blots were probed with antisera to Ebp1, PTB or GAPDH. **(C)** Levels of the reporter proteins CAT and luciferase produced in cells transfected with either EGFP or Ebp1 specific siRNAs. Levels are expressed as a percentage of that observed in the EGFP siRNA transfected control. Assays were carried out in triplicate and the error bars represent the observed standard deviation.

Figure 5: Structural features at the N- and C-termini of Ebp1. (**A**) The structure shows that the predicted p42 isoform (left) which starts at Met 55 lacks one and a half helices at the N-terminus of the p48 isoform (indicated in grey in the structure on the right). This helix makes extensive hydrophobic contacts with the body of Ebp1 (coloured by atom type: carbon - orange; nitrogen - blue oxygen - red; sulphur - yellow); its removal exposes a large hydrophobic cleft on one face of the protein. The structure of p48 Ebp1 also illustrates the proximity of K20 and K22 to the lys-rich loop 1; together these features may constitute a bipartite nucleolar localisation signal (Squatrito et al., 2004; Fujiwara et al., 2006). (**B**) Position of the ^{354}LKALL^{358} protein-binding motif at the C-terminus of Ebp1. Colouring is the same as in Fig. 1B except that residues from the motifs are highlighted in green. The surface of Ebp1 up to residue 337 is shown. Close-up views (in similar orientations) of the LxxLL motif from (**C**) Ebp1 and (**D**) the androgen receptor ((Hur et al., 2004); PDB - 1t7f). Residues from Ebp1 are colour coded as described above. Carbon atoms of the LxxLL motif of the peptide ligand of AR are cyan.

Table I: Data collection and refinement statistics

DIFFRACTION DATA	
Space-group	$P2_12_12_1$
a (Å)	49.15
b (Å)	72.78
c (Å)	115.62
Wavelength (Å)	0.8123
Resolution range (Å)	38.41 - 2.50
Independent reflections	14714
Multiplicity[a]	3.0 (3.0)
Completeness (%)	98.9 (98.9)
Mean I/(std. dev.)	10.4 (3.3)
R_{merge} (%)[b]	8.0 (41.0)
MODEL REFINEMENT	
Nonhydrogen atoms/waters	2729/55
R_{model} [c] / R_{free} [d] (%)	18.3/21.6
r.m.s deviation from ideal bond lengths (Å)	0.006
r.m.s deviation from ideal angles (°)	1.27
Ramachandran plot, favoured/additional (%)	91.8/7.6
Average B-factor (Å^2)	35.2
PBD ID	2v6c

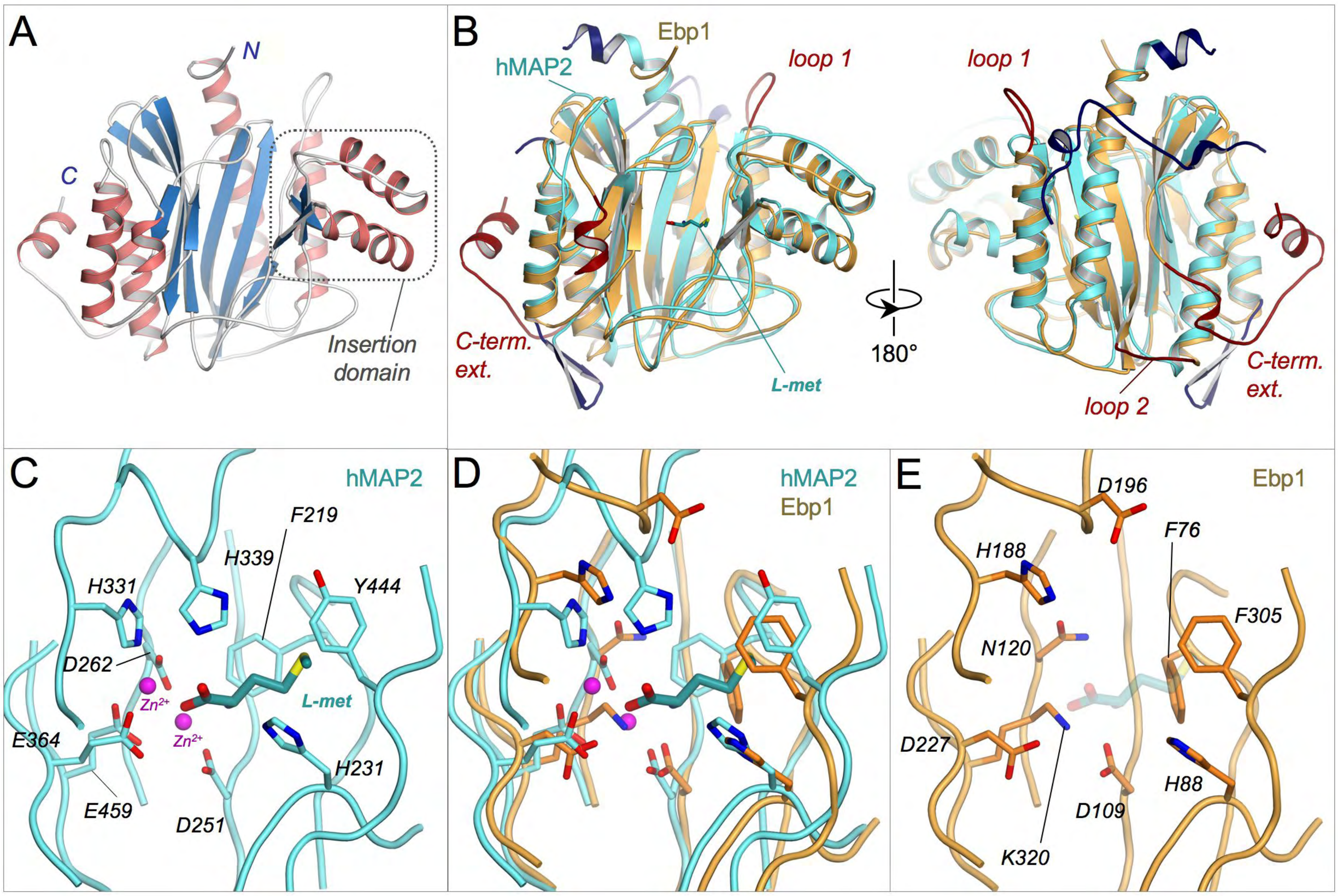
[a]Values for the outermost resolution shell (2.64-2.50 Å) are given in parentheses.

[b] $R_{merge} = 100x\Sigma_h\Sigma_i|I_{hj} - I_h|/\Sigma_h\Sigma_i I_{hj}$ where I_h is the weighted mean intensity of the symmetry related reflections I_{hj}

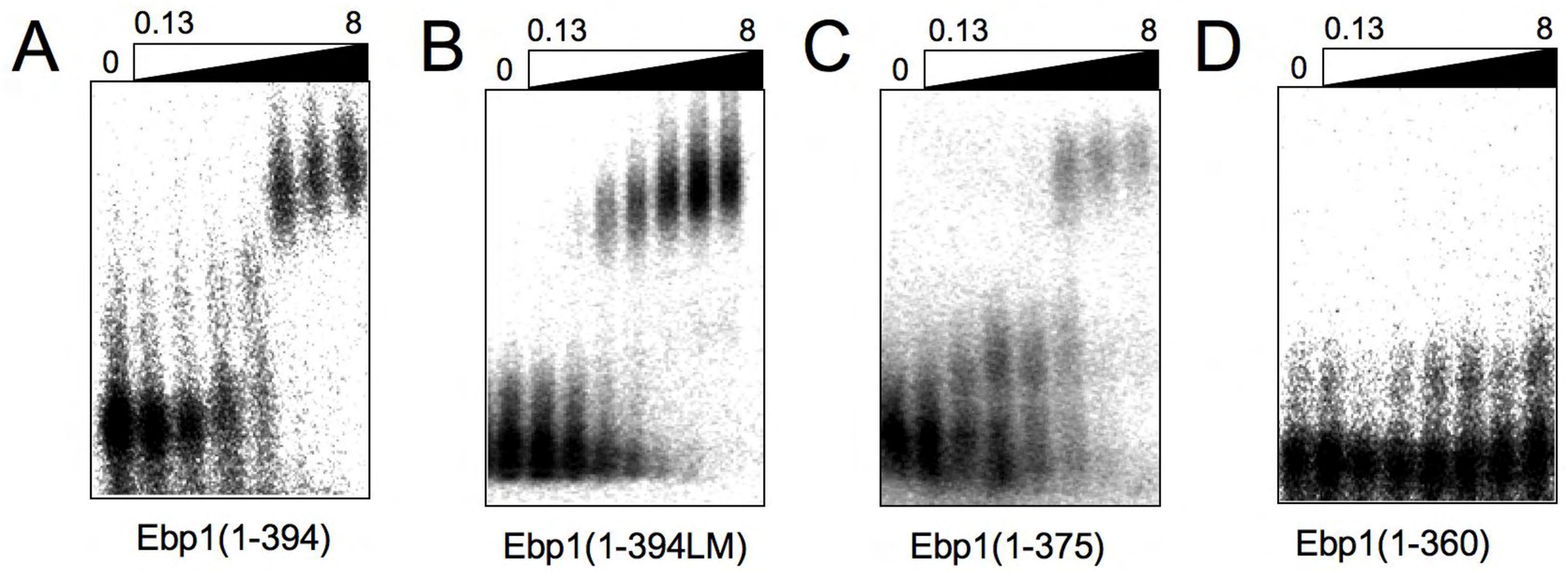
[c] $R_{model} = 100x\Sigma_h|F_{obs} - F_{calc}|/\Sigma_h F_{obs}$ where F_{obs} and F_{calc} are the observed and calculated structure factors respectively.

[d]R_{free} is the R_{model} calculated using a randomly selected 5% sample of reflection data omitted from the refinement.

Monie et al. - Fig 1

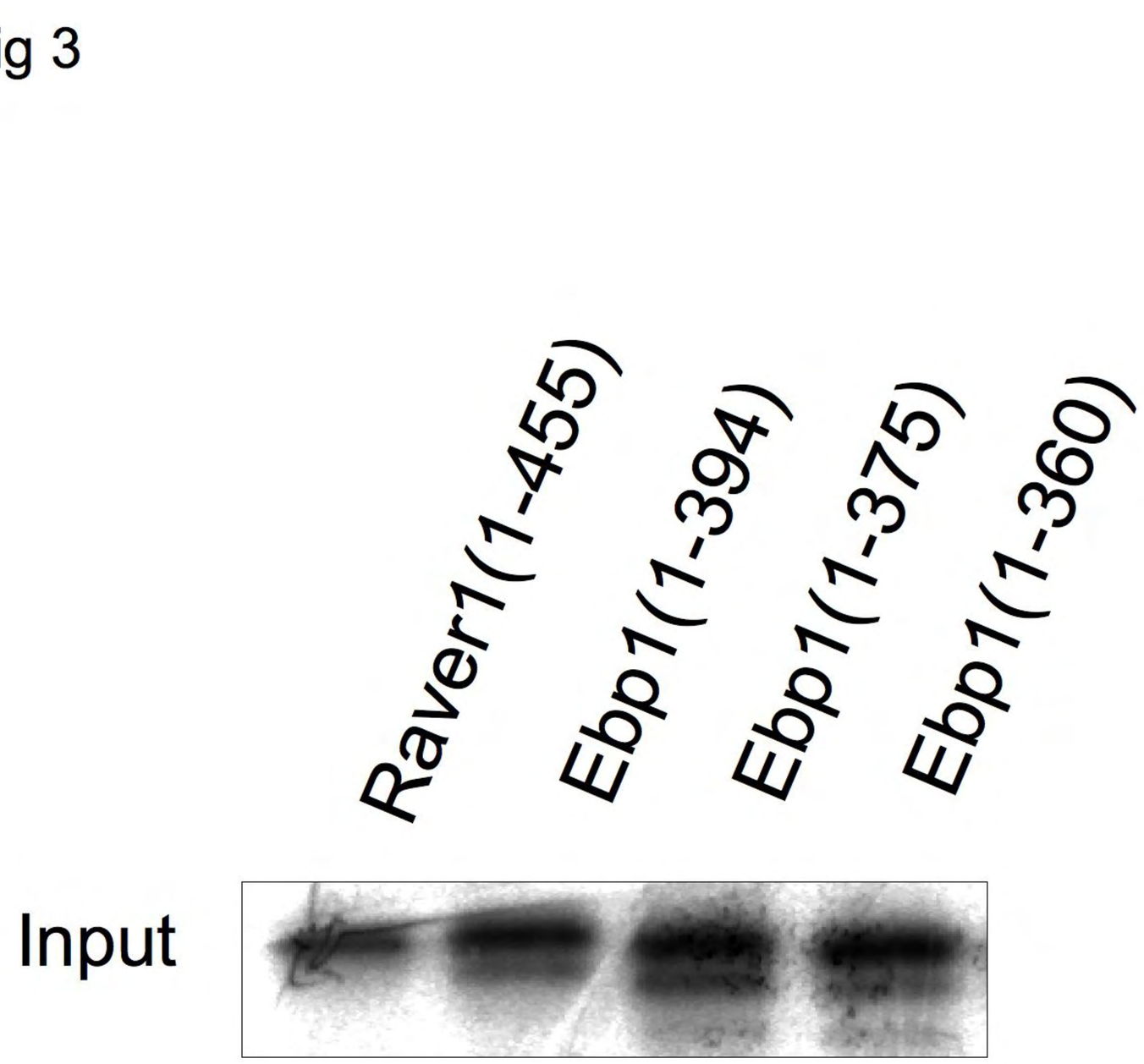


Monie et al. - Fig 2

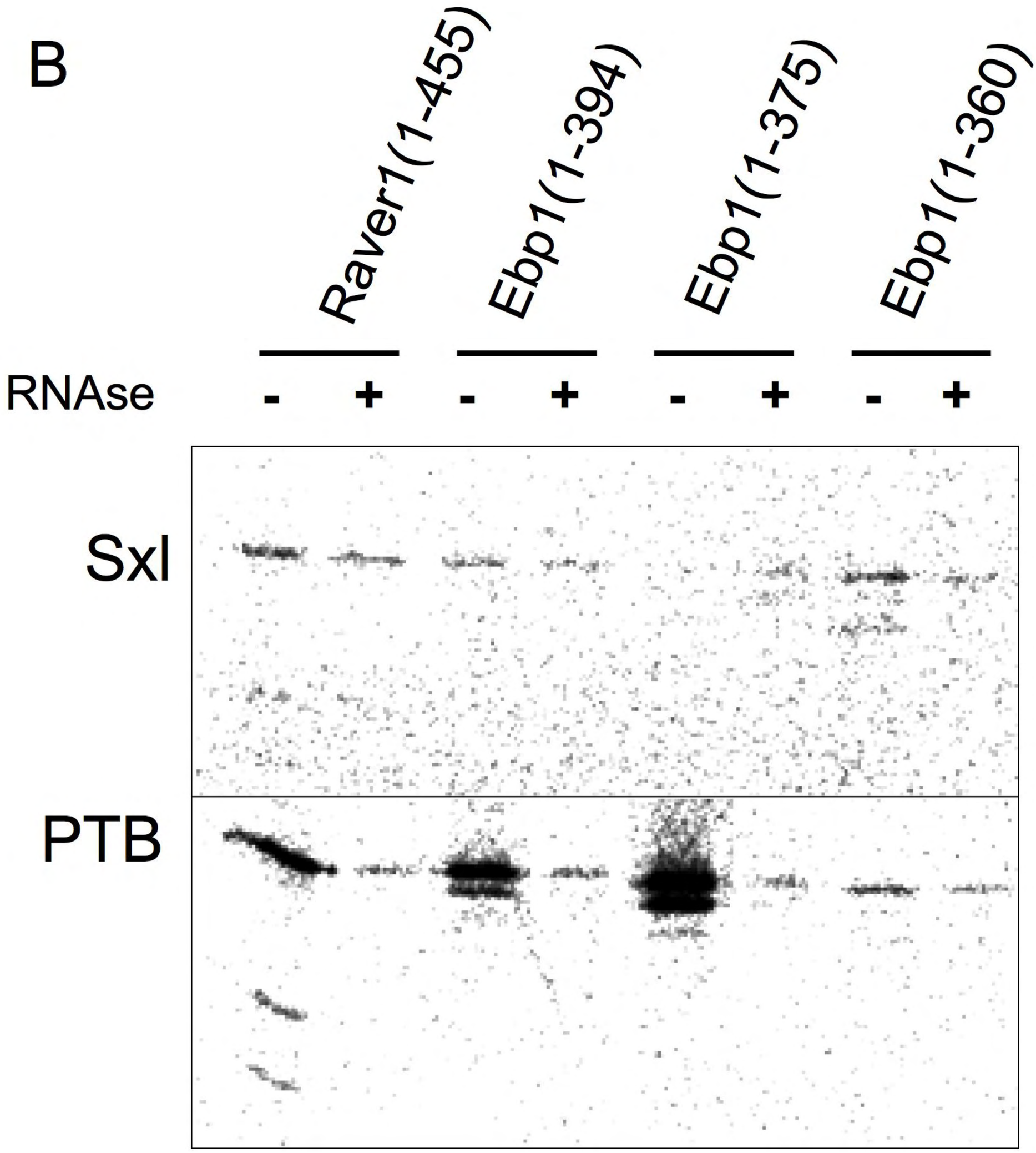


Monie et al. - Fig 3

A



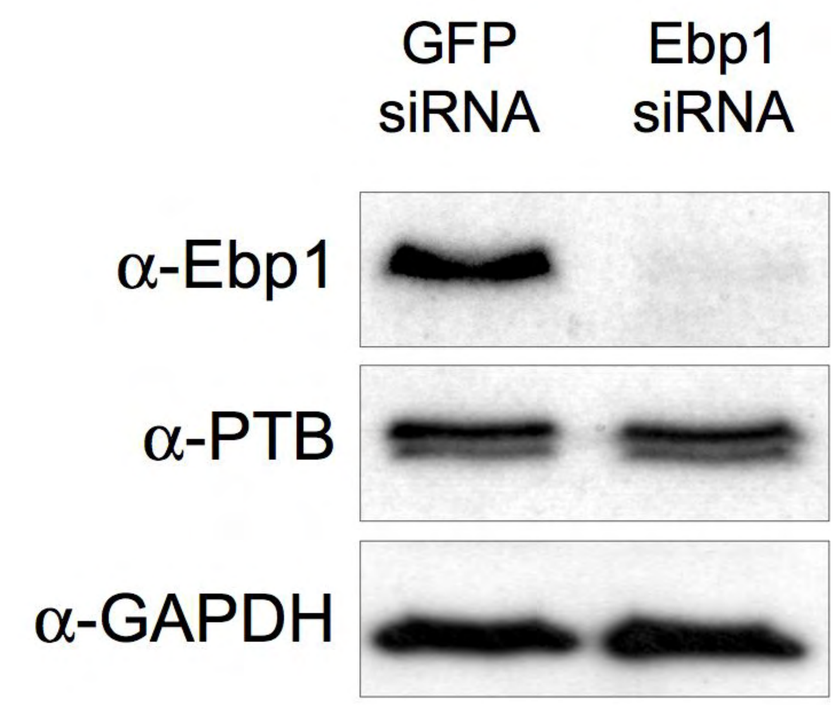
B



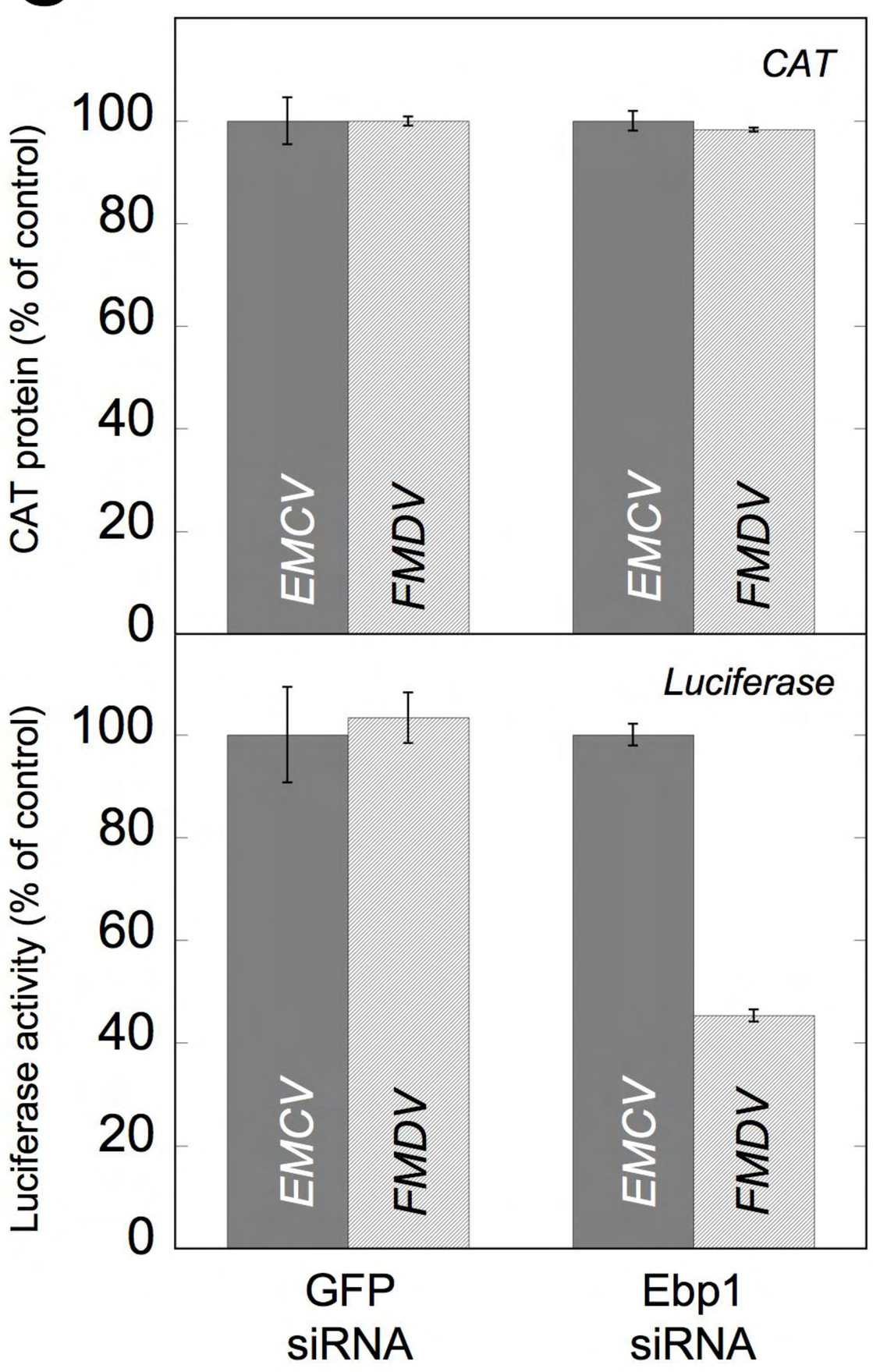
Monie et al. - Fig 4

A

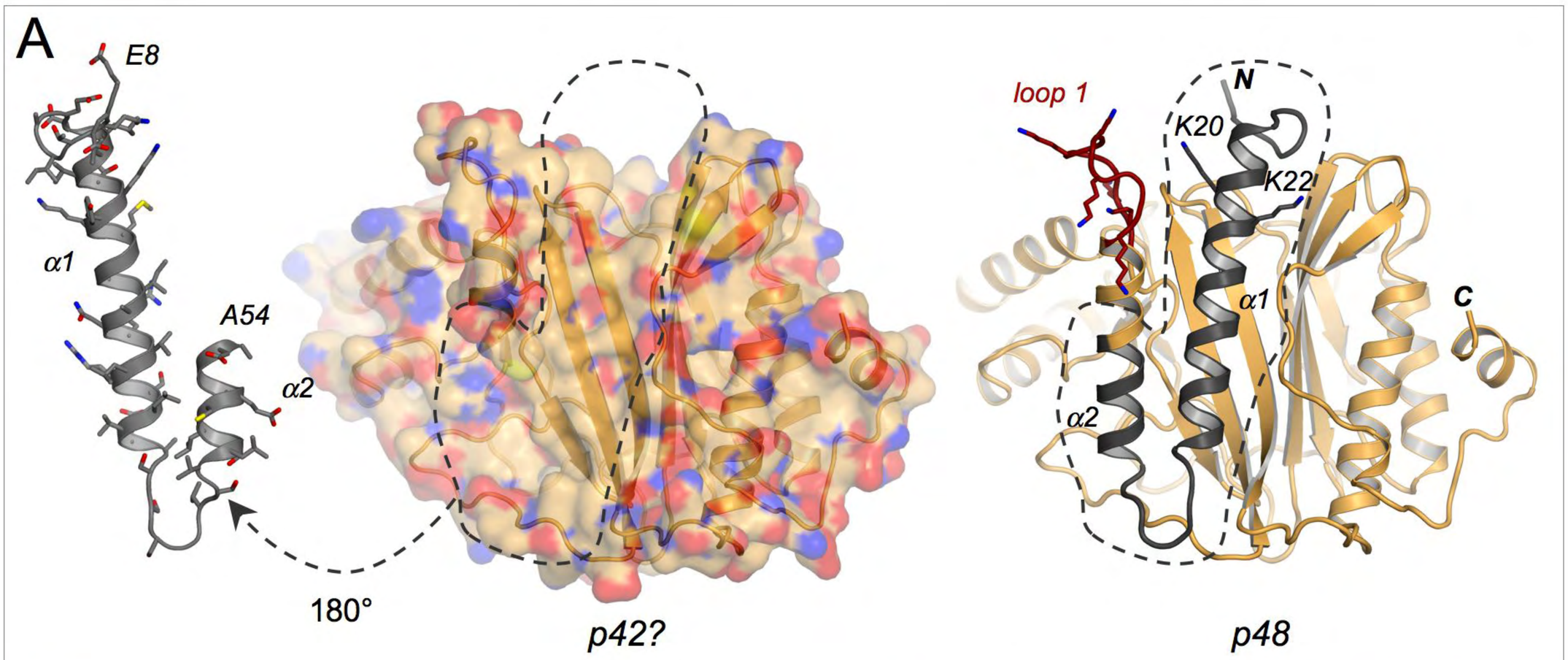
B

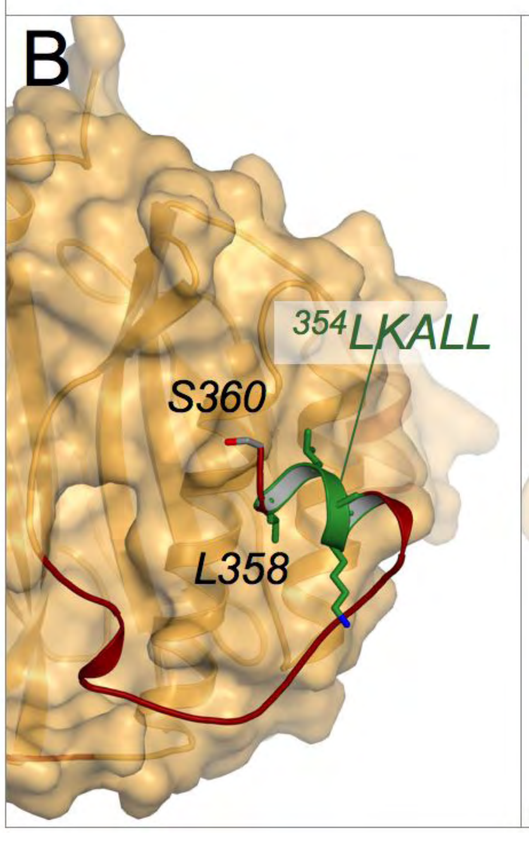


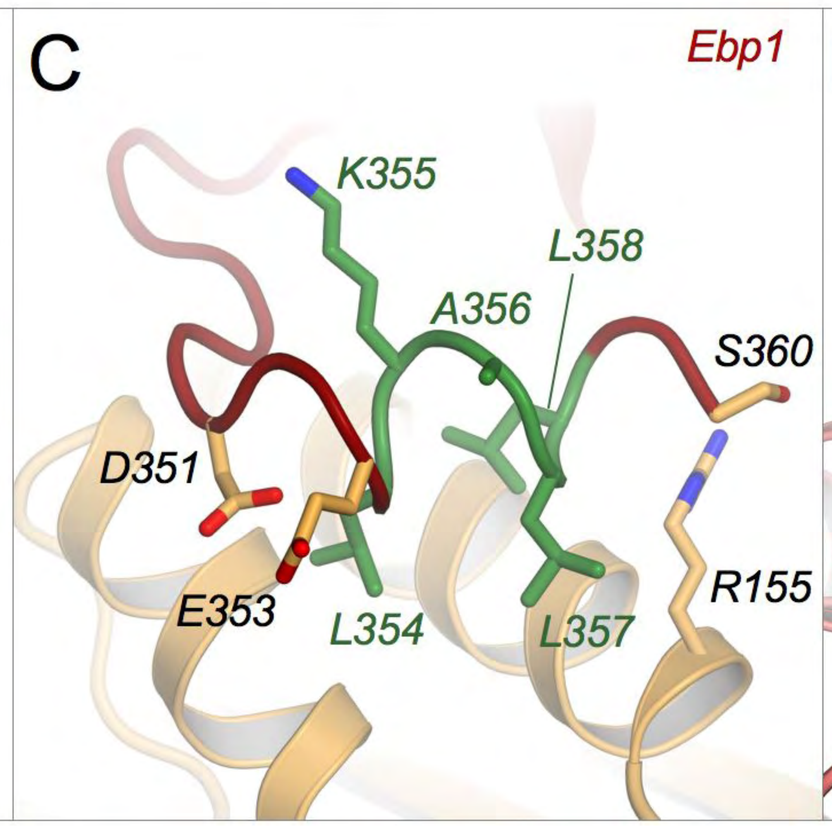
C

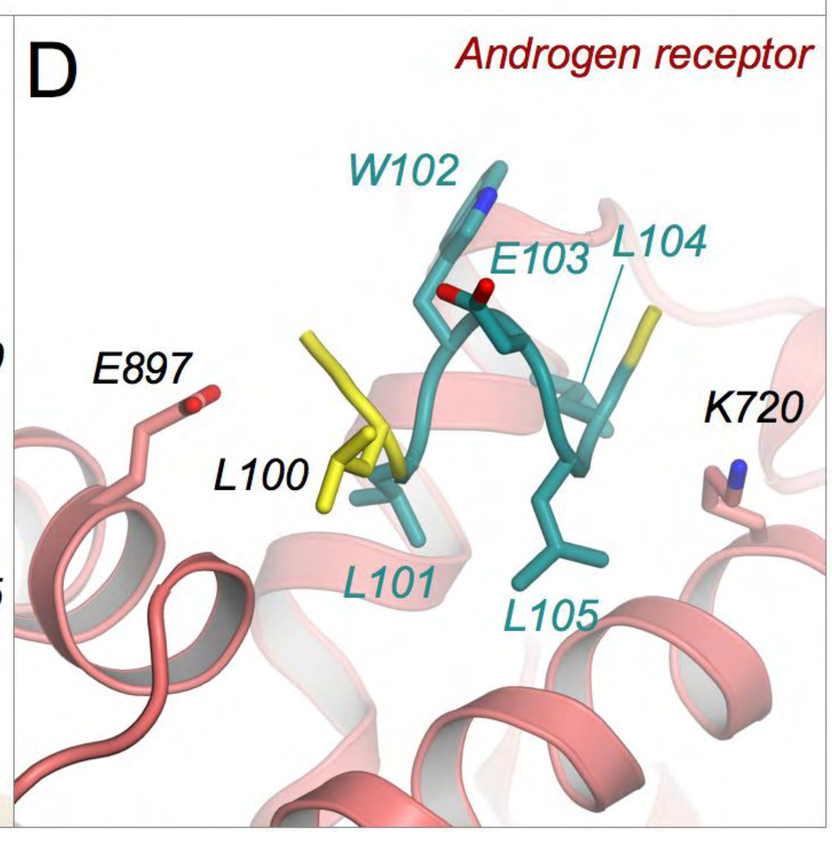


Monie et al. - Fig 5









Supplementary Information for:

Structural insights into the transcriptional and translational roles of Ebp1

Tom P. Monie[1], Andrew J. Perrin[1], James R. Birtley[1], Trevor R. Sweeney[1], Ioannis Karkasiliotis[3], Yasmin Chaudry[3], Lisa O. Roberts[4], Stephen Matthews[2], Ian G. Goodfellow[3]and Stephen Curry[1]*

[1]Division of Cell and Molecular Biology and [2]Division of Biomolecular Science, Imperial College, South Kensington Campus, London SW7 2AZ, UK.

[3]Department of Virology, Faculty of Medicine, Imperial College London, St Mary's Campus, London W2 1PG, UK

[4]School of Biomedical and Molecular Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK

*Corresponding Author. Email: s.curry@imperial.ac.uk; Tel: +44-20-7594-7632; Fax: +44-20-7589-0191.

Supplementary Table S1: Expression of Ebp1 constructs in *E. coli*

Construct	Expression	Solubility	Comments
Ebp1(1-394)	✔	✔	No crystals obtained
Ebp1(8-394)	✔	✔	No crystals obtained
Ebp1(8-360)	✔	✔	Crystallised
Ebp1(8-346)	✔	✖	-
Ebp1(8-336)	✔	✖	-
Ebp1(18-394)	✔	✔	No evidence of TEV cleavage
Ebp1(23-394)	✔	✖	-
Ebp1(23-360)	✔	✖	-
Ebp1(23-346)	✔	✖	-
Ebp1(23-336)	✔	✖	-
Ebp1(55-394)	✖	✖	-
Ebp1(55-360)	✖	✖	-

Constructs were cloned into pETM-11 for expression with an N-terminal TEV^{pro}-cleavable His-tag.

Wild-type Ebp1 sequence (murine):

```
1   MSGEDEQQEQ TIAEDLVVTK YKMGGDIANR VLRSLVEASS SGVSVLSLCE
51  KGDAMIMEET GKIFKKEKEM KKGIAFPTSI SVNNCVCHFS PLKSDQDYIL
101 KEGDLVKIDL GVHVDGFIAN VAHTFVIGVA QGTQVTGRKA DVIKAAHLCA
151 EAALRLVKPG NQNTQVTEAW NKVAHSFNCT PIEGMLSHQL KQHVIDGEKT
201 IIQNPTDQQK KDHEKAEFEV HEVYAVDVLV SSGEGKAKDA GQRTTIYKRD
251 PSKQYGLKMK TSRAFFSEVE RRFDAMPFAL RAFEDEKKAR MGVVECAKHE
301 LLQPFNVLYE REGEFVAQFK FTVLLMPNGP MRITSGPFEP DLYKSEMEVQ
351 DAELKALLQS SASRKTQKKK KKKASKTVEN ATSGETLEEN GAGD
```

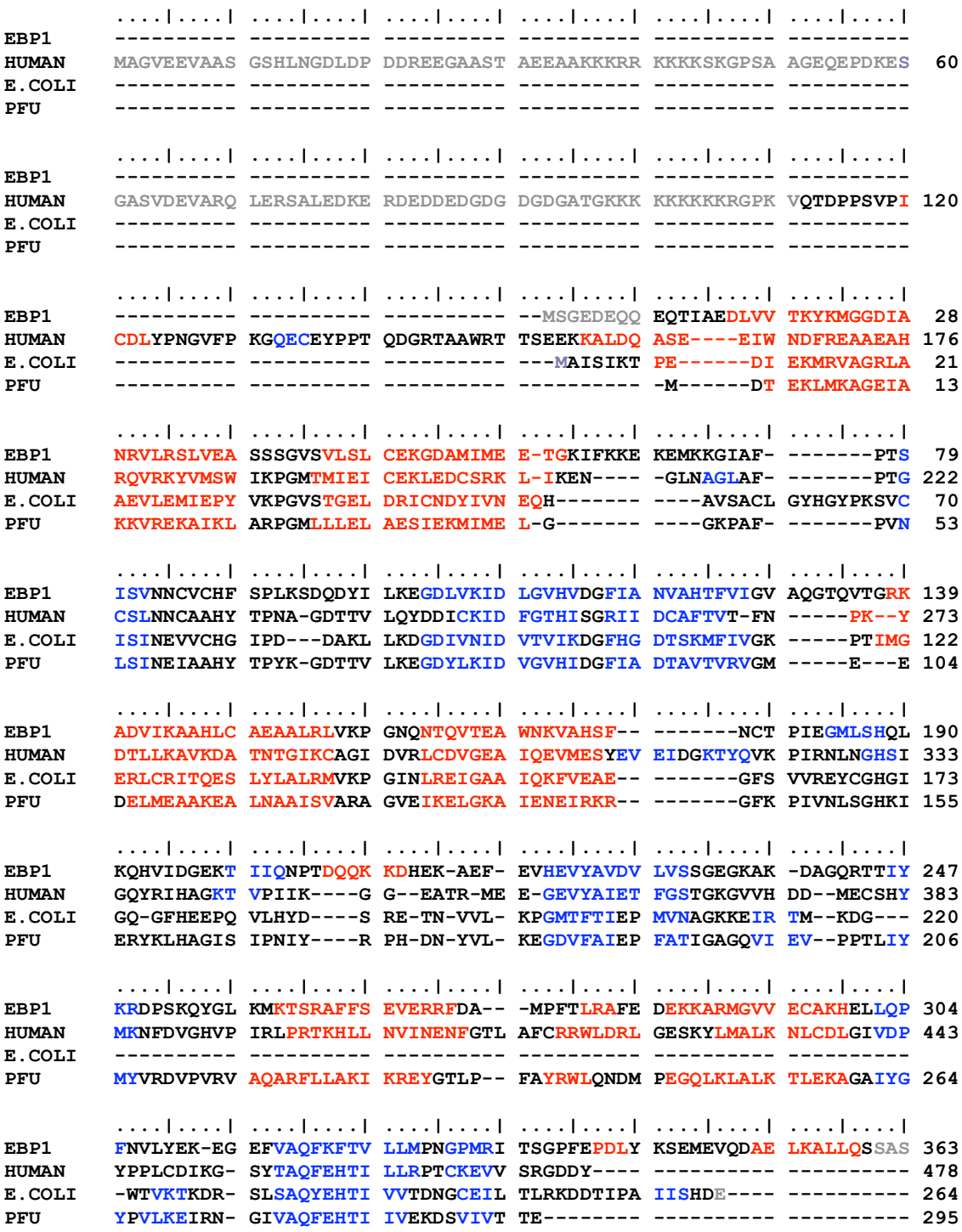
As reported previously (Pilipenko et al., 2000), the sequence contains two amino acid substitutions (T279A and K311R - underlined) in comparion to the original sequence reported for this protein (then known as Mpp1 - accession number NP_035249) (Nakagawa et al., 1997).

Supplementary Table S2: DNA primers and templates used for the generation of Ebp1 expression constructs, truncated FMDV IRES constructs, ssRNA and dsRNA transcription and preparation of dsDNA samples.

Construct	Primers
Ebp1(1-394)	5'–GTAGTAGCAGTAGGATCATGACGGGCGAAGACG–3' 5'–GAGGGGACAAGCTTCAGTCCCCAGCTCC–3'
Ebp1(1-375)	5'–GTAGTAGCAGTAGGATCATGACGGGCGAAGACG–3' 5'–TCCACTGGTGGCATTCTCAAGCTTTTAGGAGGCCT–3'
Ebp1(1-360)	5'–GTAGTAGCAGTAGGATCATGACGGGCGAAGACG–3' 5'–GGTTTTTAAGCTTGCTTAACTCTGGAGAAGAGC–3'
Ebp1(8-360)	5'–GGGCGAAGCCATGGAGCAGGAGCAAACTATCG–3' 5'–GGTTTTTAAGCTTGCTTAACTCTGGAGAAGAGC–3'
Ebp1(1-394LM)	5'–GTAGTAGCAGTAGGATCATGACGGGCGAAGACG-3' 5'–ACCTGAACTCGACGAGCTTGACGAGCTGAAGATCTTCCCTGTC–3' 5'–AGCTCGTCAAGCTCGTCGAGTTCAGGTATTGCCTTTCCTACC–3' 5'–GAGGGGACAAGCTTCAGTCCCCAGCTCC–3'
FMDV IRES Domain H (280 – 336)	5'–TAATACGACTCACTATAGGGTGCAACTTGAAACTCCGCCTGG–3' 5'-AGTACAAAATGTTACCCCTCTAGACC–3'
FMDV IRES Domain I (337 – 551)	5'–TAATACGACTCACTATAGGGTGTTTGACTCCACGCTCGA–3' 5'–GTGTTTGAGTACCAGTATCAATGTCACC–3'
FMDV IRES Domain JKL (552 – 716)	5'–TAATACGACTCACTCACTATAGGGTGGTGACAGGCTAAGGATGC–3' 5'–CATGTGTTCAGTGGTTATAAAAGGAAAGGTGCCTTCCTCCGG–3'
ssRNA (T7 primer and template)	5´TAATACGACTCACTATAGGGAG 3´ATTATGCTGAGTGATATCCCTCTAAAAAAAAAAAAAAAAAAAAACGTTAAAAAAAAAAAAAAAAAAAAA
dsRNA (T7 primer and template)	5´TAATACGACTCACTATAGGGAG 3´ATTATGCTGAGTGATATCCCTCTTTTTTTTTTTTTTTTTTTTCGTTAAAAAAAAAAAAAAAAAAAAA
dsDNA	5´-CAAGGATATCGTGGTTTTAATGGACGGAGACACCATGCCTGGCC-3´ 3´-GTTCCTATAGCACCAAAATTACCTGCCTCTGTGGTACGGACCGG-5

NB: The cloning strategy used in the generation of the plasmid Ebp1(8-360) resulted in the substitution mutation Gln8Glu. Each forward primer for the FMDV IRES truncations incorporates a T7 promoter for direct transcription.

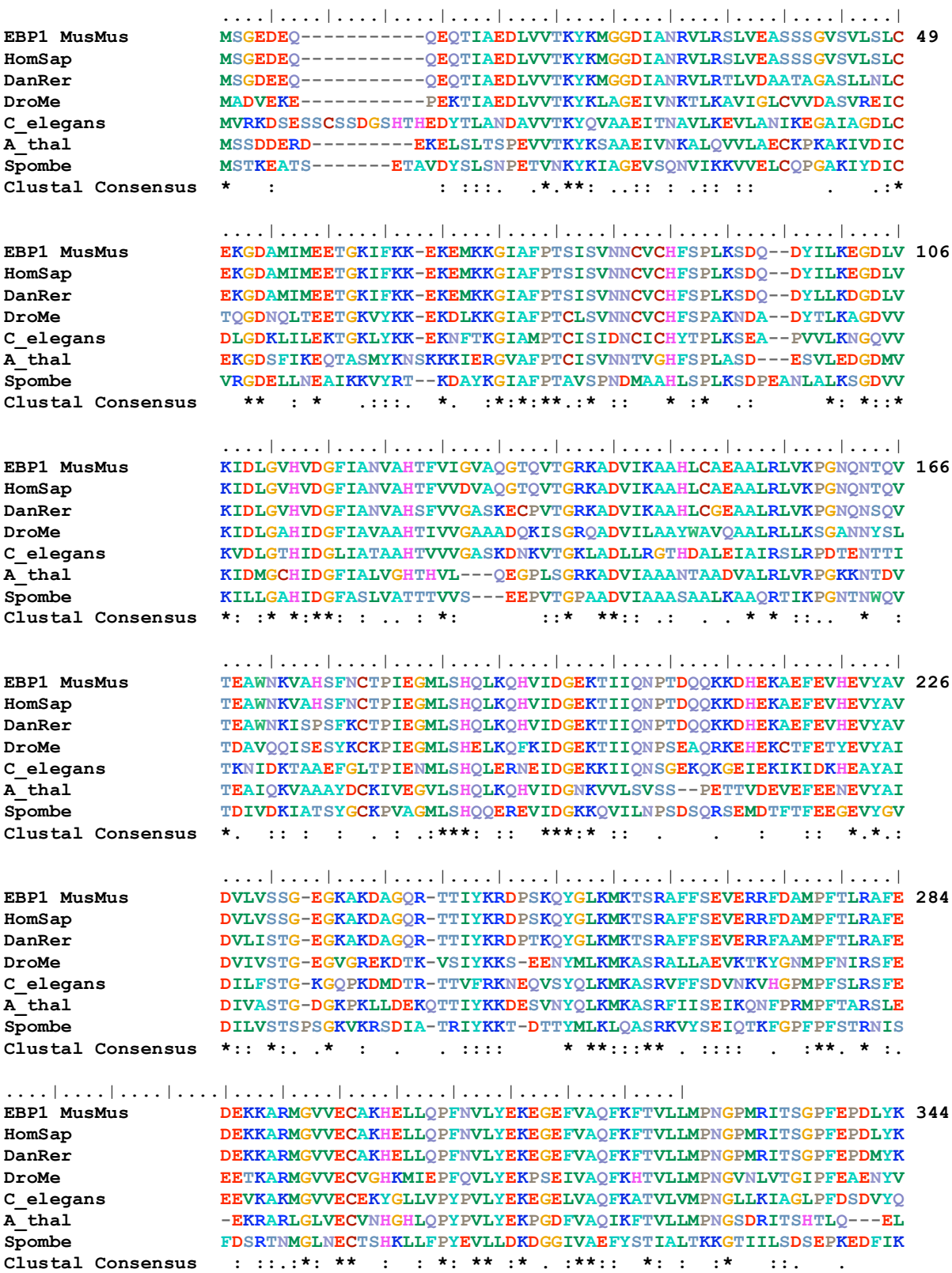
Figure S1: Structure based sequence alignment of EBP1 with representative methionine aminopeptidases (MAPs).

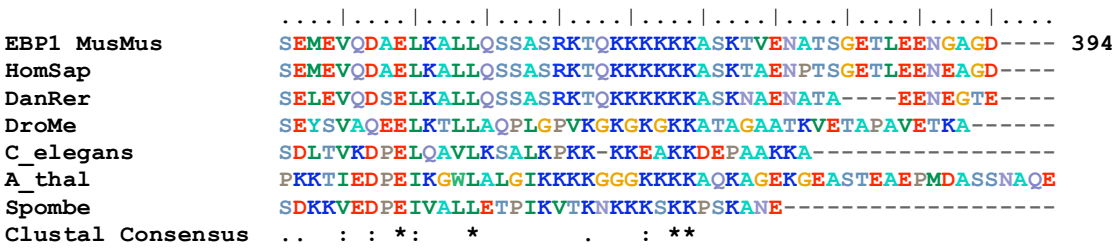


```
          ....|....| ....|....| ....|....| .
EBP1      RKTQKKKKKK ASKTVENATS GETLEENGAG D   394
HUMAN     ---------- ---------- ---------- -
E.COLI    ---------- ---------- ---------- -
PFU       ---------- ---------- ---------- -
```

PDB files for Ebp1, human p67MAP2 (1kq9), *P. fu* MAP (1xgs) and *E. coli* MAP (2mat) were submitted for multiple 3D Secondary Structure Matching (SSM) at http://www.ebi.ac.uk/msd-srv/ssm/. Aligned PDB files were visualised in Pymol and the identified secondary structure mapped on to the sequence alignment file generated by SSM. Alpha helices are shown in red, beta sheets in blue. Regions in grey were not present in the input PDB files.

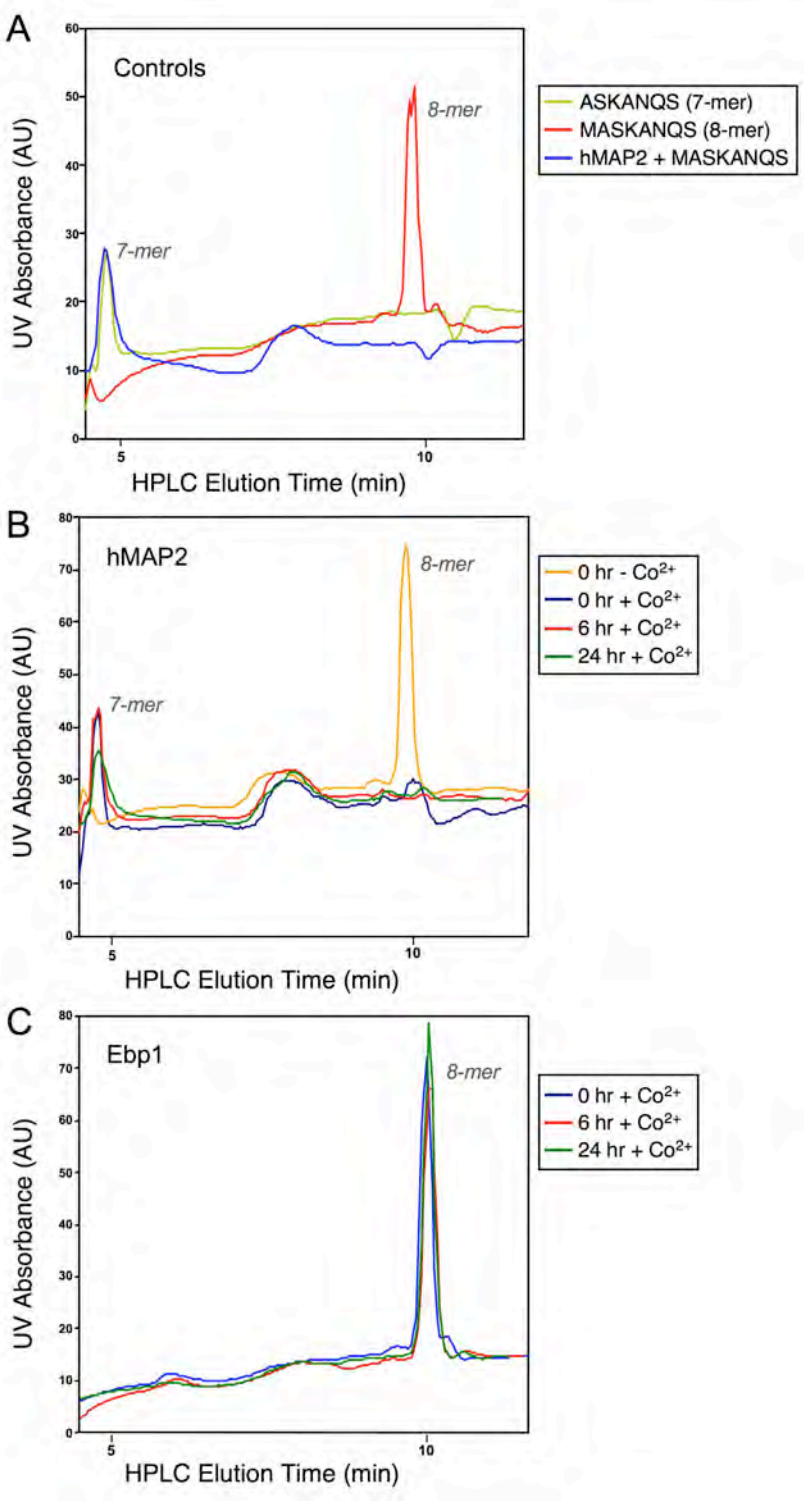
Figure S2: Sequence alignment of Ebp1 orthologues.





Orthologues to the murine Ebp1 sequence (P50580) were identified by Fasta analysis and the sequences aligned using ClustalW. Gaps in the sequences are represented by dashes. The Clustal consensus sequence highlights identity and conservation. Numbering relates to the Ebp1 sequence.

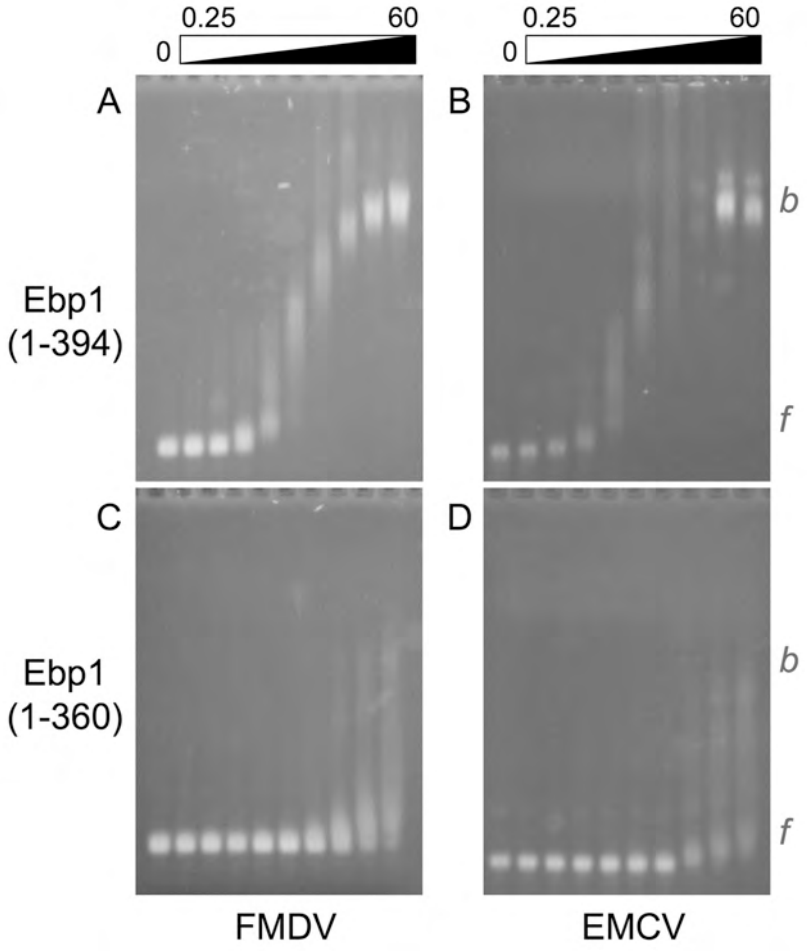
Figure S3: Comparative aminopeptidase activity of hMAP2 and Ebp1.



Both proteins were analysed by HPLC for their ability to remove the N-terminal Met from an octameric peptide substrate (MASKANQS), yielding the heptameric product (ASKANQS). The plasmid encoding a truncated but active form of hMAP2 (lacking residues 1-89) was kindly provided by Anthony Addlagatta

and Brian Matthews (University of Oregon, USA). The protein was expressed and purified essentially as described (Addlagatta et al., 2005). Methionine aminopeptidase assays were performed by incubating the protein at 4 μM with 300 μM peptide substrate at 25°C for 0, 6 and 24 hrs in a buffer composed of 10 mM HEPES (pH 7.5), 100 mM NaCl and 1mM $CoCl_2$. Reaction volumes were 200 μL. At the end of the incubation 50 μL of the reaction mix was diluted with an equal volume of water and applied to an ACE 5 C18 reverse phase HPLC column (250x4.6 mm) (Hichrom Limited, Berckshire, UK). (A) The UV trace for the heptameric and octameric peptides in the absence of added protein at concentrations of peptide equal to those present in samples of the reaction solution. A trace showing the heptameric product in the presence of hMAP2 is also shown. (B) UV traces of samples taken at 0 hrs, 6 hrs and 24 hrs from an incubation of the octameric substrate with hMAP2 as described above. In the presence of 1 mM $CoCl_2$ cleavage is rapid, with a sharp decline in the substrate peak and a concomitant increase in the product peak at the earliest time point. In the absence of added $CoCl_2$ no such change is observed. (C) UV traces of samples taken at 0 hrs, 6 hrs and 24 hrs from an incubation of the octameric substrate with full length Ebp1. Incubation with Ebp1 had no affect on the concentration of octameric substrate in the reaction even after 24 hr incubation. Similarly, no product peak is observed.

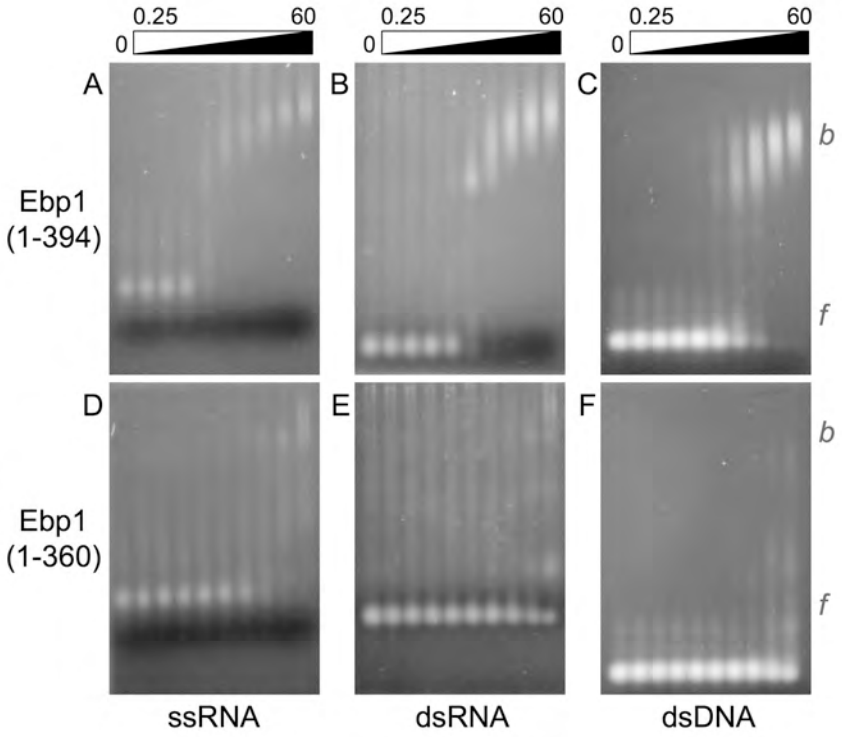
Figure S4: Comparative binding of Ebp1(1-394) and Ebp1(1-360) to FMDV and EMCV IRES RNA using electromobility shift assays (EMSAs).



The full-length protein, Ebp1(1-394) binds to the IRES from (A) FMDV and (B) EMCV with essentially the same affinity. Free and bound forms of RNA are indicated by 'f' and 'b' respectively on the right hand side of the figure. In each case, significant shifts are evident at 2-4 μM. (C,D) In contrast, only a marginal shift of both IRESes is observed at 60 μM Ebp1(1-360), the highest concentration used. Clearly, deletion of the C-terminal Lys-rich motif severely impairs RNA binding to both IRESes (see main text).

Methods: The IRES RNA samples were prepared by unlabelled *in vitro* transcription using the Promega RiboMAX kit following the manufacturers instructions. Following DNAseI digestion, RNA transcripts were purified using RNAeasy columns (Qiagen), precipitated using ethanol/sodium acetate and dissolved in RNase-free water to a concentration of 1.5 mg/ml. The RNA was annealed before use in binding experiments by heating to 70 °C for 5 minutes and slow-cooling to room temperature over 1 hr. Binding experiments were performed in 10 μL reaction volumes by mixing 100 ng of RNA (approx. 67 nM) with increasing concentrations of protein (0, 0.25, 0.5, 1, 2, 4, 8, 16, 32 and 60 μM) in 10 mM HEPES (pH 7.0), 3 mM $MgCl_2$, 5 mM DTT, 100 mM KCl, and 5% glycerol and incubating at 25 °C for 10 minutes. Samples were then mixed with 2 μL of 50% (v/v) glycerol and 0.1 % (w/v) Bromophenol Blue and separated on 1% (w/v) agarose gels run in Tris-Borate (TB) (pH 8.3).

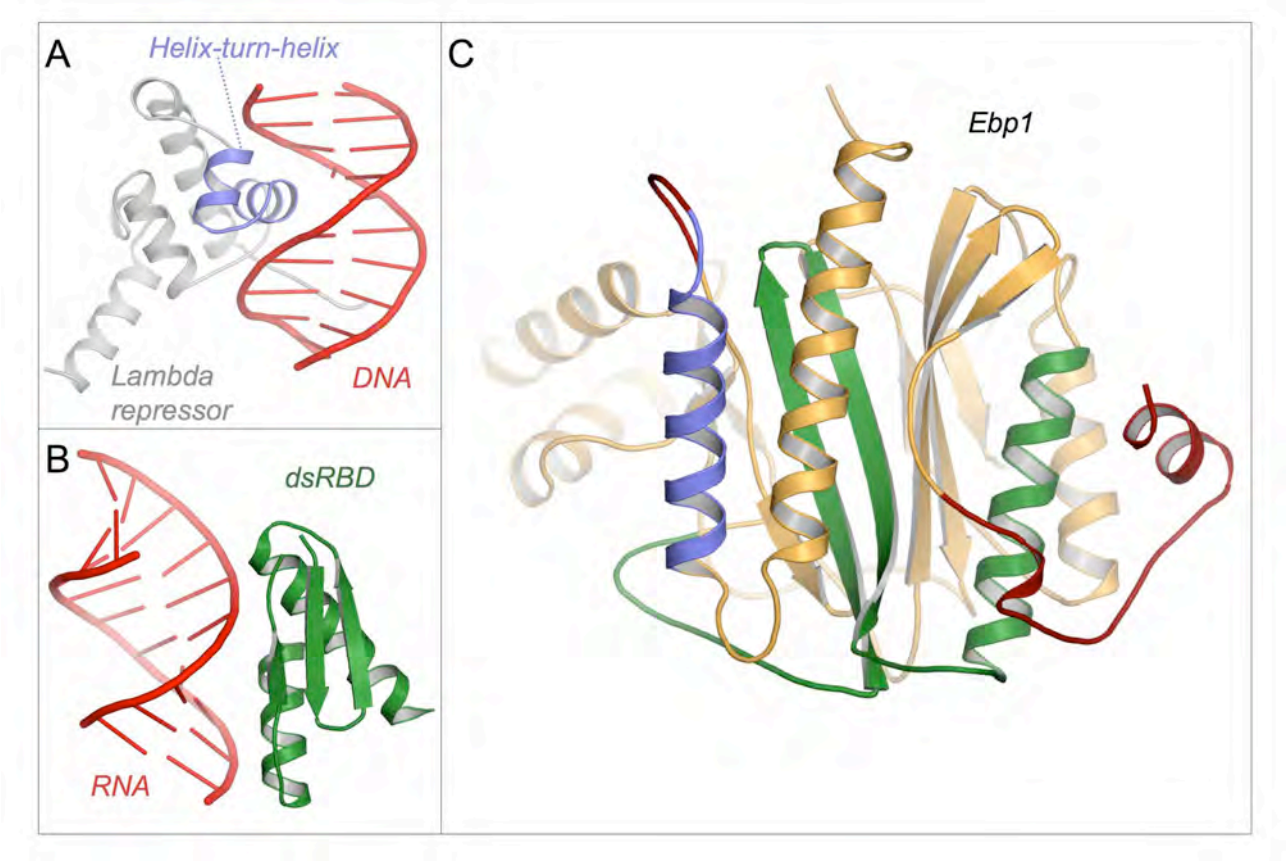
Figure S5: Comparative binding of Ebp1(1-394) and Ebp1(1-360) to single-stranded RNA (ssRNA), double-stranded RNA (dsRNA) and dsDNA using EMSAs.



The full-length protein, Ebp1(1-394), is able to interact with (A) singled-stranded RNA, (B) double-stranded RNA and (C) double stranded DNA with similar affinities. Free and bound forms of RNA are indicated by 'f' and 'b' respectively on the right hand side of the figure. Appreciable shifts are evident at 2-4 μM, similar to the concentrations required to shift the FMDV and EMCV IRES RNAs under these conditions. In contrast, only weak binding activity is exhibited by Ebp1(1-360) with (D) ssRNA, (E) dsRNA and (F) dsDNA; of these different types of ligand, ssRNA appears to bind tightest.

Methods: RNA samples were prepared by *in vitro* transcription using the synthetic DNA templates and T7 primers detailed in Table S2. Transcription reactions were performed using MEGAshortscript T7 kit (Ambion). DNaseI-treated RNA transcripts were purified on G25 microspin columns (GE Lifesciences) with a concentration 1 mg/ml. Transcription of the template for dsRNA yields a molecule with the sequence GGGAG(A_{20})GCAA(U_{20}) that can fold into a stem-loop structure containing 20 A-U basepairs in the stem, topped by a GCAA tetraloop. The template for ssRNA is a modified version of the dsRNA template and yields a similarly sized RNA transcript in which all the A's are replaced by U's and has the sequence GGGAG(U_{20})GCAA(U_{20}). Binding experiments were performed in 10 μL reaction volumes by mixing 100 ng of RNA (approx. 250-630 nM) with increasing concentrations of protein (0, 0.25, 0.5, 1, 2, 4, 8, 16, 32 and 60 μM) in 10 mM HEPES (pH 7.0), 3 mM $MgCl_2$, 5 mM DTT, 100 mM KCl, and 5% glycerol and incubating at 25 °C for 10 minutes. Samples were then mixed with 2 μL of 50% (v/v) glycerol and 0.1 % (w/v) Bromophenol Blue and separated on 2% (w/v) agarose gels run in Tris-Borate (TB) (pH 8.3).

Figure S6: Ebp1 does not contain the previously predicted σ^{70}-like motif or double-stranded RNA binding domain (dsRBD).



σ^{70}-like domains are predicted to adopt a helix-turn-helix motif (Wehner and Baserga, 2002), similar to the structure observed in some DNA binding proteins, such as the lambda repressor whereas the dsRBD forms an autonomous αββββα-fold (Doyle and Jantsch, 2002).

(A) Ribbon diagram of a single monomer of the Lambda repressor [(Beamer and Pabo, 1992); PDB 1lmb] bound to DNA. The helix-turn-helix motif that binds in the major groove is shown in blue.

(B) A double-stranded RNA binding domain (green) bound to RNA [(Ryter and Schultz, 1998); PDB 1di2].

(C) Structure of Ebp1. The portions of the structure that were predicted to contain RNA-binding domains composed of a σ^{70}-like helix-turn-helix motif (Squatrito et al., 2004) or a dsRBD (Squatrito et al., 2006) are shown in blue and green respectively. It is clear from the structure that these sequences adopt quite different folds within Ebp1. In fact the primary RNA-binding domain maps to a Lys-rich sequence (^{361}SASRKTQKKKKKKAS^{375}) that lies immediately beyond the C-terminus of the construct that was crystallised.

References

Addlagatta, A., Hu, X., Liu, J.O. and Matthews, B.W. (2005) Structural basis for the functional differences between type I and type II human methionine aminopeptidases. *Biochemistry*, **44**, 14741-14749.

Beamer, L.J. and Pabo, C.O. (1992) Refined 1.8 A crystal structure of the lambda repressor-operator complex. *J Mol Biol*, **227**, 177-196.

Doyle, M. and Jantsch, M.F. (2002) New and old roles of the double-stranded RNA-binding domain. *J Struct Biol*, **140**, 147-153.

Nakagawa, Y., Watanabe, S., Akiyama, K., Sarker, A.H., Tsutsui, K., Inoue, H. and Seki, S. (1997) cDNA cloning, sequence analysis and expression of a mouse 44-kDa nuclear protein copurified with DNA repair factors for acid-depurinated DNA. *Acta Med Okayama*, **51**, 195-206.

Pilipenko, E.V., Pestova, T.V., Kolupaeva, V.G., Khitrina, E.V., Poperechnaya, A.N., Agol, V.I. and Hellen, C.U. (2000) A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev*, **14**, 2028-2045.

Ryter, J.M. and Schultz, S.C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *Embo J*, **17**, 7505-7513.

Squatrito, M., Mancino, M., Donzelli, M., Areces, L.B. and Draetta, G.F. (2004) EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. *Oncogene*, **23**, 4454-4465.

Squatrito, M., Mancino, M., Sala, L. and Draetta, G.F. (2006) Ebp1 is a dsRNA-binding protein associated with ribosomes that modulates eIF2alpha phosphorylation. *Biochem Biophys Res Commun*, **344**, 859-868.

Wehner, K.A. and Baserga, S.J. (2002) The sigma(70)-like motif: a eukaryotic RNA binding domain unique to a superfamily of proteins required for ribosome biogenesis. *Mol Cell*, **9**, 329-339.