

## Co-translational, Intraribosomal Cleavage of Polypeptides by the Foot-and-mouth Disease Virus 2A Peptide\*

Received for publication, November 15, 2002, and in revised form, January 7, 2003  
Published, JBC Papers in Press, January 8, 2003, DOI 10.1074/jbc.M211644200

Pablo de Felipe, Lorraine E. Hughes, Martin D. Ryan, and Jeremy D. Brown<sup>‡§</sup>

From the School of Biology, Centre for Biomolecular Sciences, Biomolecular Sciences Building, University of St. Andrews, North Haugh, St. Andrews KY16 9ST, United Kingdom and the <sup>‡</sup>School of Cell and Molecular Biosciences, The Medical School, University of Newcastle, Framlington Place, Newcastle upon Tyne NE2 4HH, United Kingdom

**During co-translational protein import into the endoplasmic reticulum ribosomes are docked onto the translocon. This prevents inappropriate exposure of nascent chains to the cytosol and, conversely, cytosolic factors from gaining access to the nascent chain. We exploited this property of co-translational translocation to examine the mechanism of polypeptide cleavage by the 2A peptide of the foot-and-mouth disease virus. We find that the scission reaction is unaffected by placing 2A into a co-translationally targeted protein. Moreover, the portion of the polypeptide C-terminal to the cleavage site remains in the cytosol unless it contains its own signal sequence. The pattern of cleavage is consistent with the proposal that the 2A-mediated cleavage reaction occurs within the ribosome itself. In addition, our data indicate that the ribosome-translocon complex detects the break in the nascent chain and prevents any downstream protein lacking a signal sequence from gaining access to the endoplasmic reticulum.**

Positive-strand RNA viruses typically encode polyproteins that are cleaved by viral or host-encoded proteinases (proteolytic processing) to produce mature, individual proteins (reviewed in Refs. 1 and 2). Alternatively proteins may be generated by translational effects such as ribosomal frameshifting or read-through of “leaky” stop codons. Such programmed alterations of translation are not virus-specific but widespread (although rare) mechanisms of gene expression (reviewed in Refs. 3 and 4). In foot-and-mouth disease virus (FMDV)<sup>1</sup> and some other picornaviruses the oligopeptide (~20 amino acid) 2A region of the polyprotein mediates cleavage at its own C terminus to release it from the 2B region. 2A is also active when placed between reporter proteins and, therefore, cleavage requires no viral (proteinase) sequences outside this short peptide (5). Similarly, no host proteinases are known that cleave the 2A/2B site.

Scission of 2A-containing polyproteins requires the correct protein rather than mRNA sequence (6). However, synthetic peptides containing 2A and “2A-like” sequences from other

viruses do not autoproteolyse (7). Furthermore, on translation *in vitro* the portion of a polyprotein N-terminal to 2A typically accumulates in excess over the C-terminal portion. This imbalance is not because of protein degradation or nonspecific transcription/translation termination (6, 8). Modeling of 2A and 2A-like sequences indicates that the majority of each peptide can form an amphipathic helix whereas the amino acids immediately preceding the “cleavage” site (-NPG↓P-) form a tight turn (7). A co-translational model for the cleavage reaction has been proposed (6, 7) in which the conformation of 2A places strain on the peptidyltransferase center of the ribosome, repositioning the peptidyl(2A)-tRNA ester linkage. This steric effect prohibits nucleophilic attack by the incoming (prolyl)-tRNA amide nitrogen that normally creates the new peptide bond. Instead, the N-terminal product is released from the ribosome by hydrolysis of the peptidyl(2A)-tRNA ester bond. A proportion of ribosomes then cease translation, while the remainder continue, effectively “initiated” by the prolyl-tRNA, to produce the downstream product as a discrete (“cleaved”) entity.

Although published data support the proposal that 2A acts within the ribosome no direct evidence has been provided for this. Intraribosomal cleavage would not require cytosolic factors such as proteinases, and demonstration of this requires translation of 2A-containing proteins in a situation where cytosolic factors have no access to the nascent chain. Such “screening” occurs during co-translational translocation into the endoplasmic reticulum (ER). Here the nascent chain is shielded from cytosolic factors by first the ribosome, then the translocation apparatus, before being partitioned into the ER lumen. Establishment of co-translational translocation requires the signal recognition particle (SRP; reviewed in Refs. 9 and 10). SRP binds *cis*-acting hydrophobic signal sequences at the N terminus of nascent ER-targeted proteins as they emerge from the ribosome and concomitantly slows translation by the ribosome, a phenomenon termed elongation arrest (11–13). This ensures that ribosome-nascent chain complexes targeted to the translocon by SRP arrive with a short length of cytosolically exposed nascent chain. Once the ribosome-translocon junction is established the ribosomal nascent chain exit site lies directly over the translocon (14–16). This junction is sufficiently tight to protect the nascent chain from externally added proteases (17, 18) and even excludes small ions (19, 20). Thus comparing the translation products of 2A-containing proteins with and without SRP-dependent signal sequences will reveal if 2A functions without the influence of cytosolic factors.

The translocon has been proposed to have a signal sequence recognition function independent from that of targeting factors such as SRP (21, 22). Co-translational translocation of a 2A-containing polypeptide will result in amino acids C-terminal to 2A being presented to the translocon immediately after they

\* This work was supported by Wellcome Trust Research Career Development and UK Medical Research Council Senior NonClinical Research Fellowships (to J. D. B.) and Wellcome Trust Grant 064814/Z/01/Z (to P. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>‡</sup> To whom correspondence should be addressed. Tel.: 44-191-222-7470; Fax: 44-191-222-7424; E-mail: Jeremy.Brown@ncl.ac.uk.

<sup>1</sup> The abbreviations used are: FMDV, foot-and-mouth disease virus; ER, endoplasmic reticulum; SRP, signal recognition particle; GFP, green fluorescent protein; CPY, carboxypeptidase Y.

have emerged from the ribosome. Therefore if 2A is active in a co-translationally targeted protein we reasoned that we could use this to ask two interdependent questions. First, where does cleavage take place? Second, can the translocon discern the presence or absence of a signal sequence on a nascent chain presented to it directly by the ribosome and *in vivo*? If cleavage does not take place inside the ribosome the portion of the protein downstream of 2A will arrive in the lumen of the ER. However, if the nascent chain cleaves within the ribosome a gap will occur in the polypeptide. The translocon may then "detect" this discontinuity in the nascent chain as it does the normal termination of translation, closing, and excluding the downstream protein from the ER. In contrast addition of a signal sequence to the N terminus of protein downstream of the 2A site and expression *in vivo* should result in reopening of the translocon and translocation of this protein into the ER.

We chose the yeast *Saccharomyces cerevisiae* for these studies. We demonstrate that the FMDV 2A sequence is functional in yeast, and that targeting through the SRP-dependent co-translational translocation pathway does not impair the cleavage reaction. In addition the released downstream product remains in the cytosol unless it contains its own signal sequence. Thus nascent chain cleavage takes place before the C-terminal portion of the protein initiates translocation and therefore within the ribosome. These data provide significant support for the proposal that 2A modifies the activity of the ribosome to promote scission of the nascent chain (6, 7) and are consistent with the notion that the translocon itself examines nascent chains for the presence of signal sequences.

#### EXPERIMENTAL PROCEDURES

**Strains, Constructs, and General Methods**—Yeast strains were JDY6 (*MATa*, *trp1-Δ99*, *his3-Δ200*, *ura3-Δ99*, *leu2-Δ1*, *ade2-101*, *cir<sup>o</sup>*), JDY13 (*trp1*, *his3*, *ura3*, *ade2*, *lys2*, *sec65-1*, *MATa*), JDY365 (*trp1-Δ99*, *his3-Δ200*, *ura3-Δ99*, *leu2-Δ1*, *ade2-101*, *sec63-201-HIS3*, *cir<sup>o</sup>*) both (13) and JDY37 (*trp1-1*, *his3-11*, *-15*, *ura3-1*, *leu2-3*, *-112*, *ade2-1*, *can1-100*, *pep4Δ::TRP1*). Yeast transformations were performed by the lithium acetate method (23) and growth media and temperatures were as indicated. Plasmids were constructed as follows. Sequences encoding ppαF, ssΔαF, or D<sub>N</sub>αF were amplified from pDJ100 (24) or pJD75 (25) and cloned as *Bam*HI-*Xba*I fragments, along with an *Xba*I-*Apa*I fragment encoding the 19-amino acid 2A/2B FMDV sequence from plasmid pMR90 (5) and an *Apa*I-*Nsi*I GFP(S65T) fragment with the ER retention motif HDEL appended to its C terminus into pGEM11zf+ (Promega). This yielded plasmids encoding ppαF-2A-GFP, D<sub>N</sub>αF-2A-GFP, and ssΔαF-2A-GFP. D<sub>N</sub>αF-2A-Kar2-GFP and ssΔαF-2A-Kar2-GFP were constructed in the same way but using a PCR product that contained the Kar2 signal sequence in addition to GFP. Digestion of these plasmids with *Apa*I and *Cla*I removed the Kar2 signal sequence allowing it to be replaced with the carboxypeptidase Y (CPY) signal sequence. To generate noncleavable fusions 2A was replaced with 2A\* incorporating a change that alters the essential proline at position 17 of 2A to alanine (26). Fragments consisting of the whole of each fusion were transferred to pMW20 (27) for expression in yeast. *PHO8*-containing constructs were assembled directly into pMW20 from *PHO8* lacking its stop codon amplified from genomic DNA as an *Eco*RI-*Xba*I fragment, 2A or 2A\* and EGFP (Clontech) amplified as an *Apa*I-*Sac*I fragment.

Mouse monoclonal anti-GFP were from Roche Molecular Biochemicals, sheep polyclonal anti-2A antibodies were raised against a peptide corresponding to the FMDV 2A sequence (QLLNFDLLKLAGD-VESNPG). Anti-CPY, Kar2p, and Pho8p antibodies were as previously described (25).

**In Vitro Transcription-Translation**—Coupled transcription/translation reactions were performed as per the manufacturer's instructions (Promega). Briefly, rabbit reticulocyte lysates (10 μl) were mixed with [<sup>35</sup>S]methionine (10 μCi; Amersham Biosciences) and 0.1 μg of unrestricted plasmid DNA and incubated at 30 °C for 90 min.

**Microscopy**—For visualization of GFP cultures were concentrated to 5–10 A<sub>600</sub>/ml, mixed 1:1 with molten 0.8% (w/v) low melting point agarose, and spotted onto slides. Vacuolar membranes were visualized with FM4-64 (Molecular Probes Inc.) (28). Cells were prepared for immunofluorescence (29) using anti-GFP and anti-Kar2p antibodies at 1:50 and 1:10,000 dilution, respectively. Alexa 488 and 594 dye-coupled

fluorescent secondary antibodies (Molecular Probes Inc.) were used at 1:200. Live and fixed cells were viewed and images were captured using a Zeiss Axiovert 200 microscope equipped with Plan-Apochromat ×100 1.4NA DIC objective, Zeiss AxioCam monochrome camera, and Zeiss Axiovision software using Zeiss filter sets 10 (GFP and Alexa 488) and 31 (FM4-64 and Alexa 594). Phase images were collected in the 4,6-diamidino-2-phenylindole channel.

**Cell Labeling and Fractionation**—Cell labeling with [<sup>35</sup>S]LPromix (Amersham Biosciences), preparation of non-native extracts and immunoprecipitation were as described (25). Quantification was carried out using a Fuji BAS1500 PhosphorImager and TINA software (Raytest). To allow comparison between cleaved and uncleaved species, values obtained were divided by the number of labeled amino acids in each. Native extracts were prepared by breaking cells using zirconium beads (Bio-Spec Products) in a ribolyzer (Hybaid) with two pulses on setting 5.5 for 20 s with 100 A<sub>600</sub> units of cells in 1 ml of 20 mM HEPES-KOH, pH 7.4, 1 mM EDTA, 0.8 M sorbitol. The lysate was spun at 500 × g for 5 min to remove nonbroken cells and at 100,000 × g for 1 h in a Beckman SW50.1 rotor to produce the cytosol fraction (supernatant). The pellet of the second spin was resuspended in 1 ml of 20 mM HEPES-KOH, pH 7.4, 1 mM EDTA, 150 mM KOAc, 2.1 M sucrose, overlaid with 1 ml of the same buffer at 1.9 M sucrose and 3 ml without sucrose and spun for 5 h at 190,000 × g in a Beckman SW50.1 rotor yielding a single interface/membrane band. Cytosol and membrane fractions were adjusted to 15% (w/v) trichloroacetic acid, proteins were recovered by centrifugation in a microfuge and washed in acetone before electrophoresis on SDS-PAGE gels.

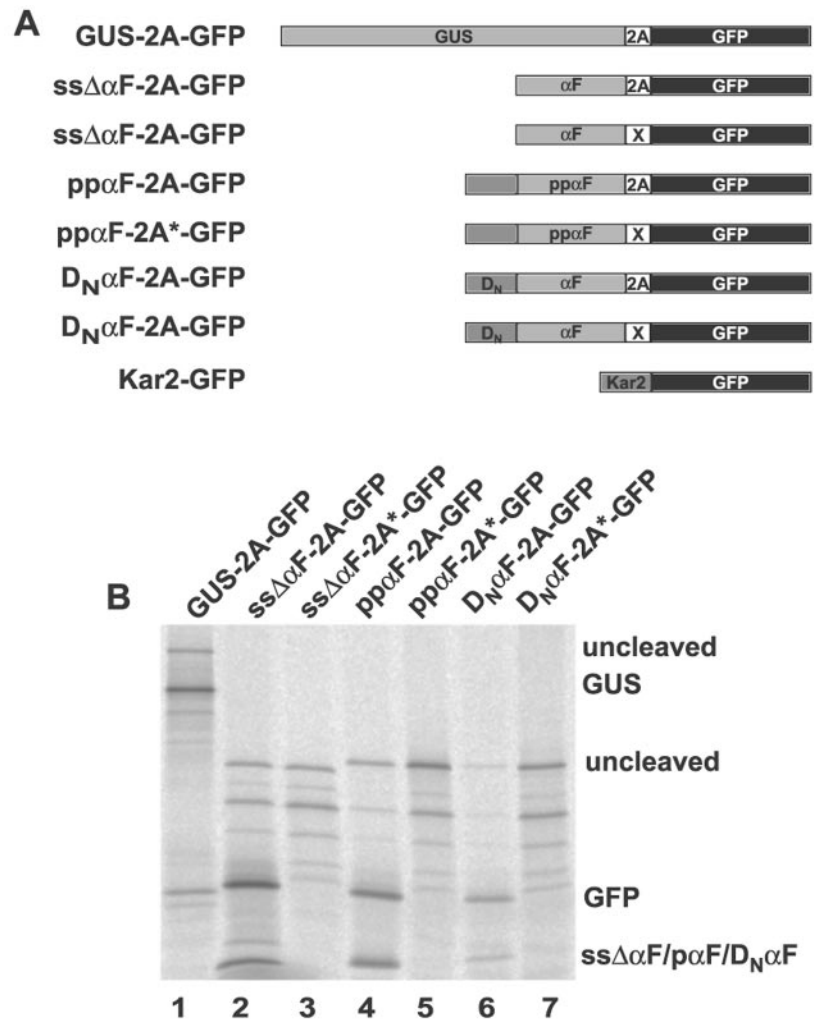
#### RESULTS

**The FMDV 2A Sequence Functions during Co-translational Translocation**—Two targeting pathways to the ER operate side-by-side in yeast, the SRP-dependent co-translational route and an SRP-independent post-translational route (25, 30). The hydrophobicity of the signal sequence determines which route is used and changing the signal sequence of a protein can target it into a different pathway. To examine whether 2A functions in yeast, and further during co-translational translocation, we tested a series of constructs (Fig. 1A) each encoding the same core artificial polyprotein consisting pro-α factor-2A-GFP appended by the ER retention motif HDEL. Three variants had the Dap2p signal sequence (D<sub>N</sub>αF-2A-GFP), the native α-factor signal sequence (ppαF-2A-GFP), or no signal sequence (ssΔαF-2A-GFP). D<sub>N</sub>αF is a well characterized SRP-dependent translocation substrate (25), whereas ppαF is translocated post-translationally relying on cytosolic chaperones for this, indicating cytosolic exposure (31–33). Without a signal sequence ssΔαF remains in the cytosol. A requirement for cytosolic factors in 2A-dependent cleavage would be revealed by cleavage of ppαF-2A-GFP and ssΔαF-2A-GFP but not D<sub>N</sub>αF-2A-GFP. If cytosolic factors are not required then all constructs would be cleaved. Controls were (i) the same three polyproteins with a nonfunctional 2A variant (26) termed 2A\* hereafter, (ii) GFP alone preceded by the signal sequence from the ER luminal hsp70 orthologue Kar2p (recognized by both ER targeting pathways) and appended by HDEL (Kar2-GFP), and (iii) the previously characterized GUS-2A-GFP fusion (6).

*In vitro* translation reactions using rabbit reticulocyte lysate confirmed that 2A was active in the context of the new artificial polyproteins (Fig. 1B). Similar to GUS-2A-GFP (lane 1) most of the translation products were of sizes corresponding to cleavage products (ssΔαF-ppαF/D<sub>N</sub>αF-2A and GFP; lanes 2, 4, and 6) indicating that 2A was active. Immunoprecipitation with anti-GFP and anti-2A antibodies confirmed the identity of the proteins (data not shown). As expected, equivalent reactions in which 2A\*-containing polyproteins were synthesized (lanes 3, 5, and 7) yielded predominantly full-length fusion proteins and minor, lower molecular weight products likely from internal initiation of translation.

Next we examined the proteins produced from the constructs *in vivo* in yeast. As the N-terminal portions of D<sub>N</sub>αF-2A-GFP and ppαF-2A-GFP encode secreted α-factor peptides we did not

**FIG. 1. 2A fusion constructs.** A, constructs are drawn N to C terminus, left to right, with the N-terminal portions of the fusions in *pale gray* and the C-terminal GFP *black*. Signal sequences are in *dark gray*, active 2A is a *clear box*, whereas the inactive mutant (2A\*) is indicated by a *cross*. For details of how constructs were assembled, see "Experimental Procedures." *GUS*,  $\beta$ -glucuronidase; *pp $\alpha$ F*, prepro- $\alpha$ -factor; *D<sub>N</sub>*, the dipeptidyl aminopeptidase B N-terminal cytosolic tail and signal anchor region; *ss $\Delta\alpha$ F*, prepro- $\alpha$ -factor lacking its signal sequence; *Kar2*, the signal sequence of Kar2p. B, activity of 2A constructs in coupled *in vitro* transcription and translation reactions. Reactions were carried out in rabbit reticulocyte extract using plasmids encoding the fusions shown and the proteins labeled with [<sup>35</sup>S]methionine. Reaction products were analyzed by SDS-PAGE and fluorography. Full-length proteins (uncleaved) and cleavage products are indicated.



expect them to be stable. We therefore pulse-labeled cells expressing the various proteins with [<sup>35</sup>S]methionine/cysteine and isolated 2A- and GFP-containing species from cell lysates by immunoprecipitation (Fig. 2A). N- and C-terminal portions of *ss $\Delta\alpha$ F-2A-GFP*, *pp $\alpha$ F-2A-GFP*, and *D<sub>N</sub> $\alpha$ F-2A-GFP* (lanes 3, 5, and 7) were immunoprecipitated from the lysates along with some uncleaved full-length proteins. Thus the FMDV 2A sequence is functional in yeast cells. Pulse-chase analysis (e.g. Fig. 2B) revealed that the proportion of full-length protein to cleavage product remained constant during the chase period, the amount of both reducing similarly over time, presumably because of turnover. Therefore, similar to the situation in other systems (5) 2A-dependent cleavage is closely coupled to protein synthesis, taking place rapidly and only during the labeling period. On the assumption that 2A antibodies immunoprecipitated full-length and cleavage products with equal efficiency, we determined the approximate percentage of each protein that was cleaved. Averages of two independent experiments gave 88% for *D<sub>N</sub> $\alpha$ F-2A-GFP*, 82% for *pp $\alpha$ F-2A-GFP*, and 76% for *ss $\Delta\alpha$ F-2A-GFP*, similar to the cleavage efficiency determined previously using *in vitro* transcription/translation for a fusion protein containing this 2A sequence (26).

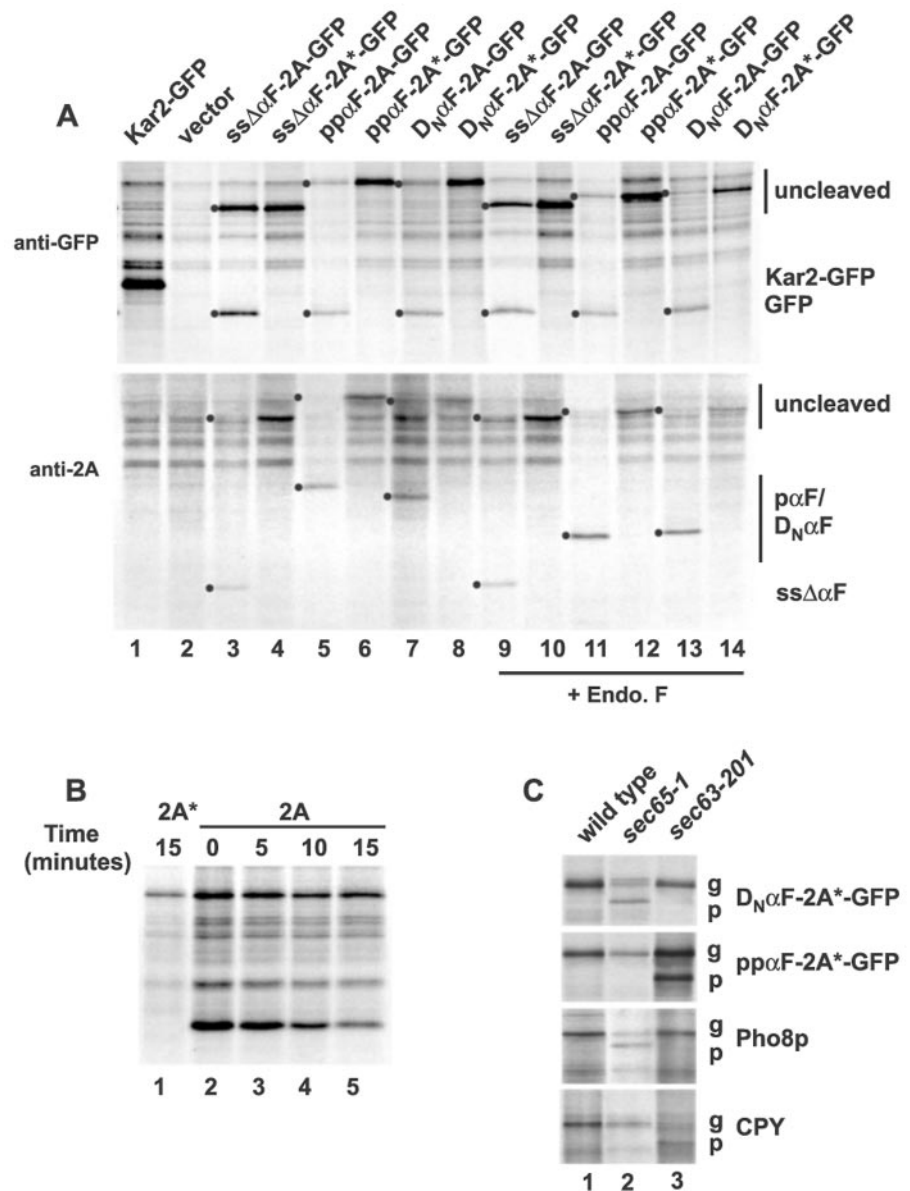
To confirm that 2A is functional when shielded from the cytosol it was necessary to demonstrate that *D<sub>N</sub> $\alpha$ F-2A-GFP* was indeed targeted to the ER, and through the expected SRP-dependent co-translational targeting pathway. As *pp $\alpha$ F*-derived sequences in the polyproteins contain sites for N-linked glycosylation the proteins will be glycosylated if they enter the ER lumen. Treatment of cell lysates with endoglycosidase F

prior to immunoprecipitation increased the mobility of both full-length *D<sub>N</sub> $\alpha$ F-2A-GFP* and *pp $\alpha$ F-2A-GFP* fusion proteins and the N-terminal *D<sub>N</sub> $\alpha$ F-2A* and *pp $\alpha$ F-2A* fragments derived from these fusions in SDS-PAGE. This confirmed that they had been glycosylated and hence translocated into the ER (Fig. 2A, compare lanes 5–8 with 11–14). In contrast the mobility of the *ss $\Delta\alpha$ F*-containing proteins were not affected by endoglycosidase treatment (compare lanes 3 and 4 with 9 and 10) indicating that, as expected, these remained in the cytosol. As GFP is not modified on entering the ER the fate of the GFP portions of the fusion proteins could not be assessed by this method.

To verify that polyproteins were targeted by the expected pathways, *D<sub>N</sub> $\alpha$ F-2A\*-GFP* and *pp $\alpha$ F-2A\*-GFP* were expressed in wild type yeast and strains defective in either co-translational (*sec65-1*) (34, 35) or post-translational (*sec63-201*) (36) translocation. Fig. 2C shows the results of immunoprecipitations with anti-GFP (top 2 panels) and control (anti-CPY and Pho8p) antibodies from extracts of these cells pulse-labeled with [<sup>35</sup>S]methionine/cysteine. As expected *sec65-1* cells revealed a defect in the glycosylation and thus translocation of the co-translationally translocated Pho8p, but not the post-translational substrate CPY (compare lane 2 with the wild type in lane 1). A significant proportion of *D<sub>N</sub> $\alpha$ F-2A\*-GFP* immunoprecipitated from the *sec65-1* cell extract was also untranslocated, whereas all *pp $\alpha$ F-2A\*-GFP* was translocated. The results obtained with *sec63-201* cells were opposite to those obtained with *sec65-1* cells, these revealing defects in translocation of CPY and *pp $\alpha$ F-2A\*-GFP* but not Pho8p or *D<sub>N</sub> $\alpha$ F-2A\*-GFP* (lane 3). Thus the pathway specificity of the *pp $\alpha$ F* and



**FIG. 2. 2A is active in yeast.** *A*, wild type yeast cells carrying plasmids encoding the fusions indicated were grown overnight at 30 °C in media containing raffinose as the sole carbon source. Expression of the fusion proteins was induced by addition of galactose to 2% (w/v). After 5 h cells were labeled with [<sup>35</sup>S]methionine/cysteine, extracts were made and immunoprecipitations performed with anti-GFP (*upper panel*) or anti-2A (*lower panel*) antibodies on either untreated extracts (*lanes 1–8*) or following treatment with endoglycosidase F (*lanes 9–14*). The immunoprecipitated proteins were analyzed by SDS-PAGE and fluorography, full-length (uncleaved) and cleavage products are indicated with *dots* to the left of the lanes. Note that the glycosylated products in *lanes 5–8* run at the same size as the anti-2A Ig heavy chain and thus are not as clearly resolved as other species. *B*, pulse-chase analysis. Cells expressing D<sub>N</sub>αF-2A\*-GFP (*lane 1*) or D<sub>N</sub>αF-2A-GFP (*lanes 2–5*) were labeled as in *A*, and then excess cold methionine and cysteine were added. Samples were removed at this time (the zero time point), and subsequently at the times indicated and immunoprecipitations were carried out as in *A* with anti-GFP antibodies. *C*, pathway specificity of constructs. Cells as indicated expressing either D<sub>N</sub>αF-2A\*-GFP or ppaF-2A\*-GFP were treated as in *A* except that *sec65-1* cells were incubated at 23 °C and then at 37 °C for 30 min prior to labeling. Antibodies used in each immunoprecipitation were anti-GFP (*upper two panels*), anti-Pho8p, or anti-CPY as indicated and relevant portions of each autoradiogram are shown.



Dap2p signal sequences were maintained. We conclude that D<sub>N</sub>αF-2A-GFP is co-translationally translocated and, therefore, that cytosolic factors are not necessary for 2A-dependent cleavage.

**Protein Sequences following 2A Are Excluded from the ER Lumen**—As discussed above if a nascent 2A-containing polypeptide “cleaves” within the ribosome an expectation might be that protein C-terminal to the cleavage site remains in the cytosol. We therefore examined the localization of the GFP portion of D<sub>N</sub>αF-2A-GFP by fluorescence microscopy. As the GFP was appended by the ER retention motif HDEL, we expected an ER localization pattern (Fig. 3D, Kar2-GFP) if the released GFP was translocated, and a cytosolic signal if the protein was not (Fig. 3C, ssΔαF-2A-GFP). Cells expressing D<sub>N</sub>αF-2A-GFP revealed cytosolic fluorescence (Fig. 3A), as did the post-translationally targeted ppaF-2A-GFP (Fig. 3B), the major observable exclusion being from the vacuole in both cases. Thus even though translation and translocation are coupled for D<sub>N</sub>αF-2A-GFP, the released GFP did not gain access to the ER. This indicates both intraribosomal cleavage of the 2A containing protein and recognition of the gap in the nascent chain by the translocon.

A caveat to the above experiment is that a proportion of the

D<sub>N</sub>αF-2A-GFP polyprotein does not cleave and some intact (uncleaved) fusion protein remains in cells. Intact, translocated D<sub>N</sub>αF-2A-GFP fusion protein should reside in the ER because of the ER retention (HDEL) motif at its C terminus. Thus we expected some ER fluorescence in cells expressing D<sub>N</sub>αF-2A-GFP regardless of the localization of the released GFP. As we did not see this (Fig. 3A), this raised the possibility that the signal from the uncleaved D<sub>N</sub>αF-2A-GFP in the ER was masked by the strong cytoplasmic GFP signal. If this were the case then some released GFP could also have been translocated and its ER fluorescence again hidden. We therefore carried out two further experiments. First we examined cells expressing the noncleaving pD<sub>N</sub>αF-2A\*-GFP and ppaF-2A\*-GFP fusions. Unexpectedly no GFP signal was detected, despite robust expression of these proteins (data not shown). However, immunofluorescence using antibodies against GFP revealed that the translocated D<sub>N</sub>αF-2A\*-GFP and pαF-2A\*-GFP proteins were localized to the ER as expected (Fig. 4, A and B). This explained the lack of ER signal from intact D<sub>N</sub>αF-2A-GFP protein in Fig. 3A. Second, we separated extracts of cells expressing the various fusions into cytosolic and membrane fractions. Fig. 4C shows that GFP released from D<sub>N</sub>αF-2A-GFP and ppaF-2A-GFP was cytosolic, whereas intact D<sub>N</sub>αF-2A\*-GFP was exclu-

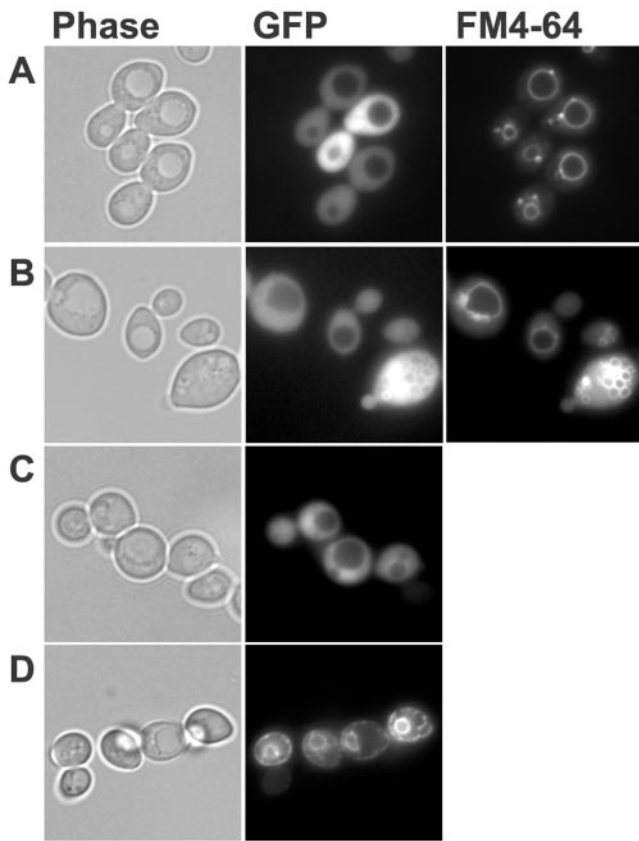


FIG. 3. **GFP cleaved from  $D_N\alpha F$ -2A-GFP is cytosolic.** Phase images, and the fluorescence signals of GFP and vacuolar membranes stained with FM4-64 were captured and processed from live cells as described ("Experimental Procedures"). Cells were grown and protein expression was induced as described in the legend to Fig. 2 before mounting on slides. Cells expressed  $D_N\alpha F$ -2A-GFP (A),  $pp\alpha F$ -2A-GFP (B),  $ss\Delta\alpha F$ -2A-GFP (C), or Kar2-GFP (D).

sively in the membrane fraction, consistent with its ER localization by immunofluorescence (Fig. 4A).

Taken together the results of these experiments confirm that cytosolic fluorescence in cells expressing  $D_N\alpha F$ -2A-GFP is an accurate reflection of the subcellular localization of GFP released by cleavage of this protein. The translocon therefore "perceives" the gap in the nascent chain and closes, effectively and quantitatively excluding the signal sequence-deficient GFP from the ER.

**2A Is Active in a  $Pho8$ -2A-GFP Fusion**—We sought to extend examination of co-translationally targeted 2A-containing proteins. Although SRP-mediated elongation arrest slows translation and thus reduces the amount of each nascent chain exposed to the cytosol, translation and translocation are not immediately coupled for all individual polypeptides targeted by SRP *in vivo*. This stochastic aspect to the coupling of targeting and translocation was revealed in experiments in which ubiquitin was inserted into proteins targeted to the ER by SRP, the "ubiquitin translocation assay" (13, 37). These were translocated into the ER intact if ubiquitin was >30 amino acids from the signal sequence. However, if ubiquitin was closer than this, a proportion of the proteins were proteolytically processed by ubiquitin-dependent proteases C-terminal to ubiquitin, indicating cytosolic exposure of the ubiquitin. The signal sequence to 2A distance in  $D_N\alpha F$ -2A-GFP is significantly greater than that required to prevent a ubiquitin-containing protein from being processed (210 amino acids compared with ~100 amino acids for a ubiquitin-containing protein). We therefore believed that translation and translocation of  $D_N\alpha F$ -2A-GFP would be coupled when the 2A peptide was synthesized. However, we considered it important to demonstrate that an SRP-dependent protein with a significantly greater distance between the signal sequence and 2A (and hence a statistically greater chance of tight coupling between translation and translocation for all nascent chains) would still cleave efficiently.

We therefore fused 2A-GFP and 2A\*-GFP (in these cases

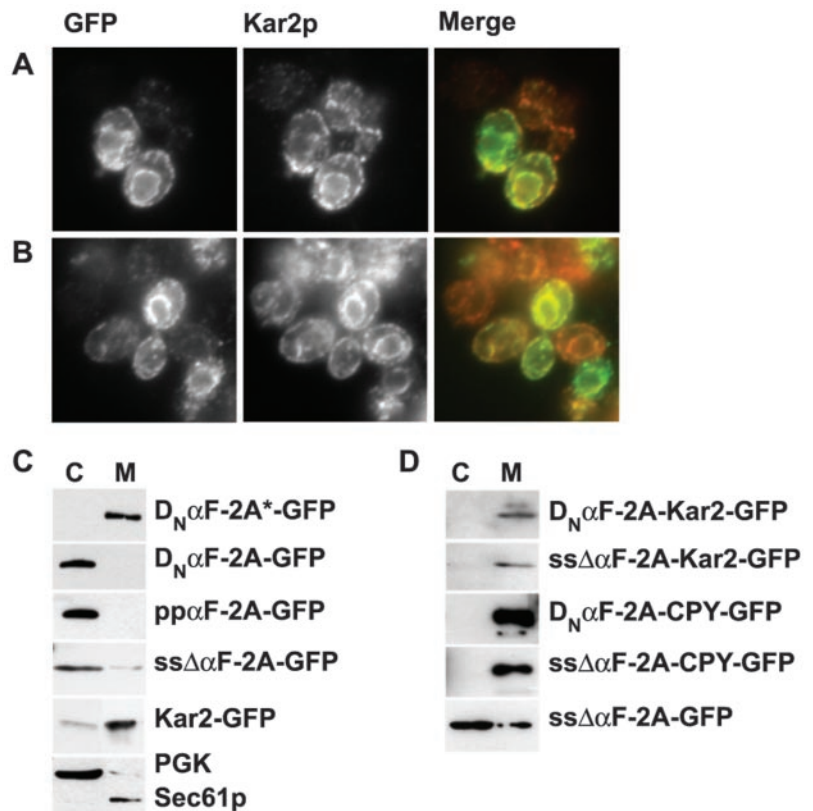
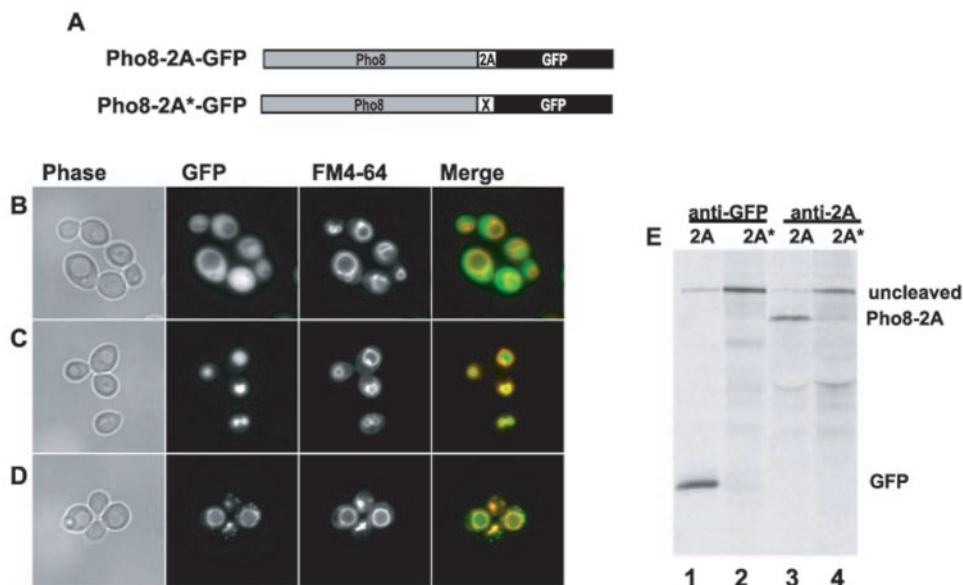


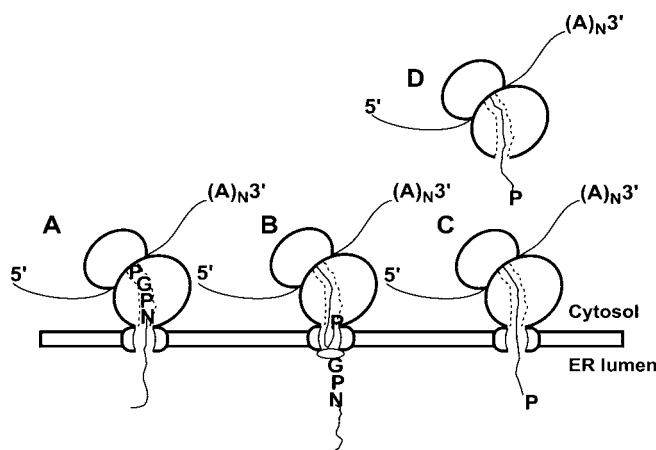
FIG. 4. **A and B**, localization of the non-cleaved polyproteins  $D_N\alpha F$ -2A\*-GFP and  $pp\alpha F$ -2A\*-GFP to the ER. Wild type cells transformed with plasmids encoding either of the noncleaved proteins  $D_N\alpha F$ -2A\*-GFP (A) or  $pp\alpha F$ -2A\*-GFP (B) were grown and protein expression was induced as described in the legend to Fig. 2. Cells were then fixed and proteins were localized by indirect immunofluorescence using either anti-GFP or anti-Kar2p antibodies as described ("Experimental Procedures"). In the merged image GFP is shown in green, Kar2p in red. **C and D**, fractionation of cells extracts. Cells expressing the indicated proteins were fractionated as described ("Experimental Procedures") and probed with antibodies against GFP, Sec61p, or phosphoglycerate kinase (PGK). Bound antibodies were revealed by enhanced chemiluminescence (Amersham Biosciences). Relevant portions of exposed films are shown: full-length  $D_N\alpha F$ -2A\*-GFP and GFP released from fusions containing active 2A.



**FIG. 5. 2A is active in Pho8-2A-GFP.** *A*, constructs drawn as in Fig. 1*A*. *B–D*, localization of GFP. Wild type cells transformed with plasmids encoding Pho8-2A-GFP (*B*) or Pho8-2A\*-GFP (*C*) and *pep4* cells transformed with Pho8-2A\*-GFP (*D*) were grown, protein expression was induced as in Fig. 2 and images captured from them as in Fig. 3 are shown. *GFP*, vacuolar membranes stained with FM4-64, merged images and phase images are shown for each. In the merged images GFP is green and FM4-64 is red. *E*, Pho8-2A-GFP is cleaved. Cells from the same cultures as in *B* and *C* were labeled with [<sup>35</sup>S]methionine/cysteine, extracts made from them and proteins immunoprecipitated as in Fig. 2 using the antibodies are shown. Positions of the intact fusion proteins and cleaved portions are indicated.

lacking a C-terminal HDEL motif) to the 566-amino acid vacuolar type II membrane protein Pho8 (Fig. 5*A*). Examination of cells expressing Pho8-2A-GFP revealed, as with  $D_N\alpha F$ -2A-GFP, a strong cytoplasmic fluorescence (Fig. 5*B*). In contrast, Pho8-2A\*-GFP was localized to the vacuolar lumen (Fig. 5*C*). Pho8p is a vacuolar membrane protein, but its maturation involves removal of a C-terminal propeptide (38). Thus GFP fluorescence within the vacuolar lumen was not surprising for Pho8-2A\*-GFP. To confirm that Pho8-2A\*-GFP reached the vacuole intact we expressed it in a processing protease-deficient (*pep4*) strain. In this case the GFP signal was largely coincident with the vacuolar membrane (Fig. 5*D*). Pulse labeling and immunoprecipitation with anti-GFP and anti-2A antibodies (Fig. 5*E*) confirmed that Pho8-2A-GFP cleaved efficiently (96% in the experiment shown) whereas, as expected, Pho8-2A\*-GFP did not. Thus positioning 2A in the context of a co-translationally translocated protein of much greater length than  $D_N\alpha F$ -2A-GFP had no effect on the cleavage reaction and these data confirm the findings with the shorter protein and our conclusion that 2A-dependent cleavage is intraribosomal.

**A Signal Sequence after 2A Is Recognized Allowing Translocation of the Released C-terminal Protein**—A prediction of the “gating” model for translocon function is that if a substrate is presented to it directly by the ribosome it should recognize signal sequences as well as prevent nonsignal sequence containing proteins from accessing the ER-lumen. We modified two of our initial substrates ( $D_N\alpha F$ -2A-GFP and  $ss\Delta\alpha F$ -2A-GFP) to contain either the Kar2 or CPY signal sequence after the 2A site. Expression of these proteins in yeast resulted, in all cases, in released GFP being translocated and thus associated with the membrane rather than cytosolic fraction of cell extracts (Fig. 4*D*). With the exception of a proline at their N terminus, Kar2-GFP and CPY-GFP released from these proteins are structurally normal translocation substrates. As the N-terminal portions of  $ss\Delta\alpha F$ -2A-Kar2-GFP and  $ss\Delta\alpha F$ -2A-CPY-GFP are not targeted to the ER the ribosomes synthesizing them would be expected to be free in the cytosol. Thus for the C-terminal fragments released from these polyproteins we



**FIG. 6. Co-translational translocation and intraribosomal nascent chain cleavage by 2A.** A model is presented to incorporate the findings in this paper into established models of translocation. *A*, as 2A is synthesized by the ribosome the cleavage reaction separates 2A from the 2B proline residue creating a gap in the nascent chain (26). *B*, the first portion of the cleaved polyprotein passes across the ER membrane and the translocon is closed, at least partly by the action of Kar2p (*oval* at bottom of translocon in *B*) (53). At this point or soon after the portion of the cleaved polyprotein C-terminal to the cleavage site engages with the translocon and is examined for the presence of a signal sequence. *C* and *D*, depending on whether a signal sequence is present on the downstream protein the translocon either reopens to allow passage of the downstream protein into the ER or remains closed, resulting in dissociation of the ribosome into the cytosol (46).

concluded that the proline at their N terminus had no effect on either targeting to or translocation into the ER. For the equivalent fragments released from  $D_N\alpha F$ -2A-Kar2-GFP or  $D_N\alpha F$ -2A-CPY-GFP we could not distinguish whether the translocon had opened directly on presentation of Kar2-GFP/CPY-GFP by the ribosome or whether the ribosome had dissociated and then been re-targeted. However, as discussed below, we consider it likely that these proteins may be recognized directly by the translocon without the need for re-targeting.



## DISCUSSION

The 2A sequence of FMDV directs production of separate proteins from one polyprotein open reading frame in a variety of higher eukaryotic and plant systems. As such it has great potential as a biotechnological tool, reducing the need for multiple vectors (e.g. Refs. 39–42). Previously the activity of FMDV 2A had not been examined in yeast. We found that it is active *in vivo* in yeast (Fig. 2). Thus, whatever features of higher eukaryotic cells allow 2A to promote scission of the polypeptide chain at the 2A C terminus are conserved to this “simpler” organism. This opens up the possibility of examining, in a genetically tractable system, how 2A functions and what *trans*-acting factors are required for or influence the cleavage reaction. A similar approach has been successfully applied to analysis of ribosomal frameshifting (43), indicating that such examination of alterations to normal ribosome function in yeast is feasible.

The proportion of 2A-containing proteins that were cleaved in yeast was similar to that seen previously in *in vitro* translation reactions, and the 2A-dependent reaction was unaffected by sequestering the nascent chain from cytosolic factors. As discussed above although SRP-dependent targeting functions similarly in yeast to higher eukaryotes (13, 44), translation and translocation are not efficiently coupled at short nascent chain lengths (37). Both SRP-dependent substrates used in our analysis are of sufficient length to expect tight coupling between targeting and translocation for the great majority of individual protein chains. Thus if cytosolic factors were required for 2A-dependent cleavage, a significant reduction of the proportion of these proteins that were cleaved would have been expected. Instead the efficiency of cleavage was maintained when the protein containing it was co-translationally translocated and we conclude that extraribosomal cytosolic factors such as proteases are not necessary for the 2A reaction.

Determination of the fate of the separated C-terminal (GFP) portions of 2A-containing polyproteins yielded information both on (i) where cleavage took place and (ii) the gating function of the translocon. Even when polyproteins were co-translationally translocated into the ER, GFP remained cytosolic if it lacked a signal sequence. This exclusion was, as far as we could determine from fractionation experiments, quantitative and our view of how this occurs is shown in Fig. 6. For GFP to be excluded from the ER the cleavage reaction must have taken place before the N terminus of GFP was translocated and therefore our data demonstrate that the proposal that the break in the nascent chain is generated within the ribosome is correct (6, 7). We suggest that the N terminus of GFP is treated as a new translocation substrate and is “surveyed” by the translocon for the presence of a signal sequence. Without a signal sequence the translocon remains closed and the ribosome-translocon complex dissociates. When a signal sequence is added to the N terminus of GFP (substrates D<sub>N</sub>αF-2A-Kar2-GFP and D<sub>N</sub>αF-2A-CPY-GFP) the translocon recognizes this and reopens. In addition to indicating that cytosolic proteases are not responsible for the cleavage reaction the cytoplasmic location of GFP released from ER-targeted proteins refutes other models of 2A activity such as one 2A sequence acting in *trans* on another nascent 2A-containing protein to cleave it, or 2A requiring release from the ribosome and folding before it becomes active in *cis*. Each of these possibilities would result in cleavage within the ER lumen, and thus the GFP portion of the polyproteins would have been found in the ER.

The model above and in Fig. 6 is consistent with the recent finding that, upon completion of co-translational translocation *in vitro*, ribosomes remain attached to the translocon, competent to initiate translation of new mRNA species (45, 46). If

these ribosomes encode proteins with signal sequences they are translocated directly into the ER independent of targeting by SRP, whereas if they do not contain a signal sequence the translating ribosomes detach and release the protein into the cytosol. This has been argued to represent a physiologically important mechanism by which the cell ensures correct partitioning of cytosolic and ER-targeted proteins. Assuming that the *in vitro* experiments (45, 46) reflect *in vivo* events, then translocation of Kar2-GFP and CPY-GFP released from D<sub>N</sub>αF-2A-Kar2-GFP and D<sub>N</sub>αF-2A-CPY-GFP likely represents direct recognition of signal sequences *in vivo* by the translocon. Thus 2A may provide a means by which we can, for the first time, examine the signal sequence recognition and gating activities of the translocon *in vivo* independent of targeting.

Analysis of 2A and 2A-like sequences has revealed that the FMDV 2A reaction is far from unique. It is a strategy for creating separate proteins from one polypeptide chain used by a variety of picornaviruses and insect viruses, and active 2A-like sequences are also found in repeated sequences within the genomes of *Trypanosoma* species (8, 26, 47). Neither is 2A the only example of a nascent peptide chain that modifies the activity of the ribosome while within the ribosome exit tunnel (reviewed in Refs. 48 and 49). Other examples are sequences within the bacteriophage T4 gene 60 that promote “hopping” of the ribosome along the mRNA (50), the TnaC leader peptide of the *Escherichia coli* tryptophanase operon that, in the presence of tryptophan, causes ribosomes to stall allowing expression of the downstream open reading frames (51) and a peptide within SecM that causes ribosomes to stall unless the protein is engaged with the protein export machinery. Mutations that suppress the stalling activity of SecM have been identified in both RNA and protein components of the ribosome, these mapping to the nascent chain exit tunnel (52). Thus interactions between the nascent chain and the ribosome can have significant effects on translation and we expect that our further analysis may reveal similar interactions important for the activity of the 2A peptide.

*Acknowledgments*—We thank members of the Brown laboratory for discussions and suggestions, and Janet Quinn, Simon Whitehall, and Susan Farrington for comments on the manuscript. We are grateful to Davis Ng, Peter Walter, Tom Stevens, Joachim Li, Caroline Shamu, Nils Johnsson, and Reid Gilmore for reagents that were used during this study.

## REFERENCES

- Ryan, M. D., Monaghan, S., and Flint, M. (1998) *J. Gen. Virol.* **79**, 947–959
- Seipelt, J., Guarne, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., and Skern, T. (1999) *Virus Res.* **62**, 159–168
- Baranov, P. V., Gurvich, O. L., Fayet, O., Prere, M. F., Miller, W. A., Gesteland, R. F., Atkins, J. F., and Giddings, M. C. (2001) *Nucleic Acids Res.* **29**, 264–267
- Farabaugh, P. J. (1996) *Annu. Rev. Genet.* **30**, 507–528
- Ryan, M. D., and Drew, J. (1994) *EMBO J.* **13**, 928–933
- Donnelly, M. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., and Ryan, M. D. (2001) *J. Gen. Virol.* **82**, 1013–1025
- Ryan, M. D., Donnelly, M. L., Lewis, A., Mehrotra, A. P., Wilke, J., and Gani, D. (1999) *Bioorg. Chem.* **27**, 55–79
- Donnelly, M. L., Gani, D., Flint, M., Monaghan, S., and Ryan, M. D. (1997) *J. Gen. Virol.* **78**, 13–21
- Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) *Annu. Rev. Biochem.* **70**, 755–775
- Bui, N., and Strub, K. (1999) *Biol. Chem.* **380**, 135–145
- Walter, P., and Blobel, G. (1981) *J. Cell Biol.* **91**, 557–561
- Wolin, S. L., and Walter, P. (1989) *J. Cell Biol.* **109**, 2617–2622
- Mason, N., Ciuflo, L. F., and Brown, J. D. (2000) *EMBO J.* **19**, 4164–4174
- Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997) *Science* **278**, 2123–2126
- Beckmann, R., Spahn, C. M., Eswar, N., Helmers, J., Penczek, P. A., Sali, A., Frank, J., and Blobel, G. (2001) *Cell* **107**, 361–372
- Menetret, J., Neuhof, A., Morgan, D. G., Plath, K., Radermacher, M., Rapoport, T. A., and Akey, C. W. (2000) *Mol. Cell* **6**, 1219–1232
- Matlack, K. E., and Walter, P. (1995) *J. Biol. Chem.* **270**, 6170–6180
- Connolly, T., Collins, P., and Gilmore, R. (1989) *J. Cell Biol.* **108**, 299–307
- Crowley, K. S., Liao, S., Worrell, V. E., Reinhart, G. D., and Johnson, A. E. (1994) *Cell* **78**, 461–471
- Crowley, K. S., Reinhart, G. D., and Johnson, A. E. (1993) *Cell* **73**, 1101–1115

21. Jungnickel, B., and Rapoport, T. A. (1995) *Cell* **82**, 261–270
22. Mothes, W., Jungnickel, B., Brunner, J., and Rapoport, T. A. (1998) *J. Cell Biol.* **142**, 355–364
23. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153**, 163–168
24. Hansen, W., Garcia, P. D., and Walter, P. (1986) *Cell* **45**, 397–406
25. Ng, D. T. W., Brown, J. D., and Walter, P. (1996) *J. Cell Biol.* **134**, 269–278
26. Donnelly, M. L., Hughes, L. E., Luke, G., Mendoza, H., ten Dam, E., Gani, D., and Ryan, M. D. (2001) *J. Gen. Virol.* **82**, 1027–1041
27. Zieler, H. A., Walberg, M., and Berg, P. (1995) *Mol. Cell. Biol.* **15**, 3227–3237
28. Vida, T. A., and Emr, S. D. (1995) *J. Cell Biol.* **128**, 779–792
29. Ciuffo, L. F., and Brown, J. D. (2000) *Curr. Biol.* **10**, 1256–1264
30. Feldheim, D., and Schekman, R. (1994) *J. Cell Biol.* **126**, 935–944
31. Deshaies, R. J., Koch, B. D., Werner, W. M., Craig, E. A., and Schekman, R. (1988) *Nature* **332**, 800–805
32. Caplan, A. J., Cyr, D. M., and Douglas, M. G. (1992) *Cell* **71**, 1143–1155
33. Becker, J., Walter, W., Yan, W., and Craig, E. A. (1996) *Mol. Cell. Biol.* **16**, 4378–4386
34. Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R., and Schekman, R. (1992) *Mol. Biol. Cell* **3**, 129–142
35. Stirling, C. J., and Hewitt, E. W. (1992) *Nature* **356**, 534–537
36. Ng, D. T., and Walter, P. (1996) *J. Cell Biol.* **132**, 499–509
37. Johnsson, N., and Varshavsky, A. (1994) *EMBO J.* **13**, 2686–2698
38. Klionsky, D. J., and Emr, S. D. (1989) *EMBO J.* **8**, 2241–2250
39. Halpin, C., Cooke, S. E., Barakate, A., El Amrani, A., and Ryan, M. D. (1999) *Plant J.* **17**, 453–459
40. Thomas, C. L., and Maule, A. J. (2000) *J. Gen. Virol.* **81**, 1851–1855
41. Varnavski, A. N., and Khromykh, A. A. (1999) *Virology* **255**, 366–375
42. De Felipe, P., and Izquierdo, M. (2000) *Hum. Gene Ther.* **11**, 1921–1931
43. Harger, J., Meskauskas, A., and Dinman, J. (2002) *Trends Biochem. Sci.* **27**, 448–454
44. Hann, B. C., and Walter, P. (1991) *Cell* **67**, 131–144
45. Potter, M. D., and Nicchitta, C. V. (2000) *J. Biol. Chem.* **275**, 33828–33835
46. Potter, M. D., and Nicchitta, C. V. (2002) *J. Biol. Chem.* **277**, 23314–23320
47. Hahn, H., and Palmenberg, A. C. (1996) *J. Virol.* **70**, 6870–6875
48. Tenson, T., and Ehrenberg, M. (2002) *Cell* **108**, 591–594
49. Morris, D. R., and Geballe, A. P. (2000) *Mol. Cell. Biol.* **20**, 8635–8642
50. Herr, A. J., Gesteland, R. F., and Atkins, J. F. (2000) *EMBO J.* **19**, 2671–2680
51. Gong, F., and Yanofsky, C. (2002) *Science* **297**, 1864–1867
52. Nakatogawa, H., and Ito, K. (2002) *Cell* **108**, 629–636
53. Hamman, B. D., Hendershot, L. M., and Johnson, A. E. (1998) *Cell* **92**, 747–758