8844–8858 *Nucleic Acids Research, 2017, Vol. 45, No. 15* *Published online 14 June 2017*
doi: 10.1093/nar/gkx528

Identification of *Arabidopsis thaliana* upstream open reading frames encoding peptide sequences that cause ribosomal arrest

Noriya Hayashi[1], Shun Sasaki[1], Hiro Takahashi[2], Yui Yamashita[1], Satoshi Naito[1,3] and Hitoshi Onouchi[1,*]

[1]Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan, [2]Graduate School of Horticulture, Chiba University, Chiba 263-8522, Japan and [3]Graduate School of Life Science, Hokkaido University, Sapporo 060-0810, Japan

Received February 02, 2017; Revised May 26, 2017; Editorial Decision June 05, 2017; Accepted June 07, 2017

ABSTRACT

Specific sequences of certain nascent peptides cause programmed ribosomal arrest during mRNA translation to control gene expression. In eukaryotes, most known regulatory arrest peptides are encoded by upstream open reading frames (uORFs) present in the 5′-untranslated region of mRNAs. However, to date, a limited number of eukaryotic uORFs encoding arrest peptides have been reported. Here, we searched for arrest peptide-encoding uORFs among *Arabidopsis thaliana* uORFs with evolutionarily conserved peptide sequences. Analysis of *in vitro* translation products of 22 conserved uORFs identified three novel uORFs causing ribosomal arrest in a peptide sequence-dependent manner. Stop codon-scanning mutagenesis, in which the effect of changing the uORF stop codon position on the ribosomal arrest was examined, and toeprint analysis revealed that two of the three uORFs cause ribosomal arrest during translation elongation, whereas the other one causes ribosomal arrest during translation termination. Transient expression assays showed that the newly identified arrest-causing uORFs exerted a strong sequence-dependent repressive effect on the expression of the downstream reporter gene in *A. thaliana* protoplasts. These results suggest that the peptide sequences of the three uORFs identified in this study cause ribosomal arrest in the uORFs, thereby repressing the expression of proteins encoded by the main ORFs.

INTRODUCTION

During the translation of certain mRNAs, specific nascent peptide sequences cause ribosomal arrest by interacting with components of the ribosomal exit tunnel (1–13). Such programmed ribosomal arrest caused by nascent arrest peptides is involved in gene expression control in both prokaryotes and eukaryotes (14–17). In eukaryotes, most known regulatory arrest peptides are encoded by upstream open reading frames (uORFs) located in the 5′-untranslated regions (5′-UTRs) of mRNAs. The presence of a uORF can negatively affect translation of the downstream main ORF in a sequence-independent manner if ribosomes translate the uORF at a significant frequency and dissociate from the mRNA after translating the uORF. In fact, many uORFs not encoding an evolutionarily conserved peptide sequence have a repressive effect on main ORF translation to a moderate extent (18). On the other hand, specific peptides encoded by the uORFs of certain genes cause ribosomal arrest and strongly repress translation of the main ORF by blocking ribosomal scanning. The arginine attenuator peptides (AAPs), encoded by the uORFs of the *Neurospora crassa arg-2* gene and its *Saccharomyces cerevisiae* ortholog *CPA1*, cause ribosomal arrest in response to arginine. The main ORFs of the *arg-2* and *CPA1* genes encode a subunit of an enzyme involved in arginine biosynthesis; therefore, the arginine-induced ribosomal arrest plays a role in the feedback regulation of arginine biosynthesis (19,20). The arrest peptides encoded by uORFs of mammalian and plant *S*-adenosylmethionine decarboxylase genes (21,22), which are involved in polyamine biosynthesis, also act in feedback regulation. The peptide sequences encoded by these uORFs cause ribosomal arrest in response to polyamine (23,24). Recently, a uORF of the *A. thaliana bZIP11* gene was shown to cause ribosomal arrest in response to sucrose in a sequence-dependent manner (25). In another example, a sequence-dependent arrest uORF is involved in stress-responsive translational regulation. The uORF-encoded peptide sequence of the mammalian *CHOP* gene causes ribosomal arrest and represses translation of the main ORF under normal conditions (26). Endoplasmic

*To whom correspondence should be addressed. Tel: +81 11 706 3887; Fax: +81 11 706 4932; Email: onouchi@abs.agr.hokudai.ac.jp

reticulum (ER) stress-induced phosphorylation of eukaryotic initiation factor 2α (eIF2α) promotes ribosomal bypass of the uORF to alleviate its repressive effect, thereby enhancing the translation of the *CHOP* main ORF (27). Another example of uORF sequence-dependent ribosomal arrest is found in the human cytomegalovirus *gp48* gene, in which the peptide encoded by the second uORF causes ribosomal arrest and represses translation of the main ORF (28).

To date, only a small number of uORFs have been shown to encode a regulatory arrest peptide. As a step towards a comprehensive identification of arrest peptide-encoding uORFs, genome-wide searches for uORFs with evolutionarily conserved peptide sequences, referred to as 'conserved peptide uORFs (CPuORFs)', were performed in various organisms including plants (29–31), mammals (32), and insects (33), using comparative genomic approaches. In *Arabidopsis thaliana*, 76 CPuORF-containing genes were identified by these bioinformatic studies. The identified CPuORFs, which included paralogous ones, were classified into 43 homology groups (HGs) based on similarity of the peptide sequences encoded by the CPuORFs (29–31,34). Nineteen CPuORFs belonging to distinct HGs were analyzed for their sequence-dependent effects on main ORF expression, and nine CPuORFs that exert a sequence-dependent repressive effect were identified among them (22,35–38). However, to date, only two of them were shown to possess the ability to cause ribosomal arrest (24,25). Besides plants, four uORF families encoding arrest peptides were identified in eukaryotes (19,20,23,26,28). However, a genome-wide search for uORFs encoding arrest peptides has not been reported in any eukaryotic organisms.

The aim of the present study was to identify novel sequence-dependent arrest uORFs in *A. thaliana*. For this purpose, we analyzed 22 uncharacterized *A. thaliana* CPuORFs for their ability to cause ribosomal arrest using an *in vitro* translation system and identified three novel uORFs that cause ribosomal arrest in a sequence-dependent manner. We determined the positions of the ribosomal arrest in these uORFs by stop codon-scanning mutagenesis and toeprint analysis, and found that two of the three uORFs cause ribosomal arrest during translation elongation, whereas the other one causes ribosomal arrest during translation termination. Furthermore, transient expression assays showed that the three newly identified uORFs exerted a strong sequence-dependent repressive effect on main ORF expression in protoplasts from *A. thaliana* cultured cells. Thus, this study identified three novel regulatory uORFs that cause ribosomal arrest during translation elongation or termination in the uORFs, thereby negatively controlling the expression of proteins encoded by the main ORFs.

MATERIALS AND METHODS

Plant material and growth conditions

Arabidopsis thaliana MM2d suspension cells (39) were cultured as previously described (37).

Plasmid construction

Plasmids pNH001 and pNH002 contain HA and glutathione *S*-transferase (GST) tag sequences in the pSP64 Poly(A) vector (Promega), respectively. To construct pNH001, oligonucleotides SP6HAfor and HABKXSrev (Supplementary Table S2) were annealed and filled in using KOD-Plus-Neo DNA polymerase (Toyobo). The resulting fragment containing a HA tag sequence was digested with SalI and SacI and inserted between the SP6 promoter and the 30 nucleotides (nt) poly(A) sequence of pSP64 Poly(A) digested with the same enzymes to yield pNH001. To generate pNH002, the GST tag sequence was amplified from pYN10 (40) by PCR with primers GSTfor and GSTrev (Supplementary Table S2). The amplified GST fragment was digested with SalI and BamHI and ligated into SalI/BamHI-digested pNH001 to replace the HA tag sequence.

Plasmids pNH01 to pNH22 carry the *GST:CPuORF(WT)* constructs, in which the GST tag sequence was fused in-frame to the second codon of each CPuORF analyzed in this study, in pSP64 Poly(A) (Supplementary Table S1). To construct these plasmids, the regions containing each CPuORF and its downstream sequence of approximately 100 nt (see Supplementary Figure S1 for the exact amplified regions) were amplified from the 5′-UTRs of the 22 genes using the primers listed in Supplementary Tables S1 and S2. For cloning of the AT2G27230, AT3G12010, and AT1G68550 CPuORFs, total RNA was extracted from *A. thaliana* (Col-0 ecotype) seedlings using the RNeasy Mini kit (Qiagen). cDNA was prepared from the total RNA using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) and an oligo(dT) primer (Life Technologies) and used as template for PCR. AT2G27230 has two splice variants with different CPuORF sequences. The reverse transcription PCR fragment amplified from the more abundant splice variant was used for cloning of the AT2G27230 CPuORF (Supplementary Figure S1A). For cloning of the other CPuORFs, genomic DNA extracted from *A. thaliana* (Col-0 ecotype) seedlings was used as template for PCR because the amplified regions contain no intron and their nucleotide sequences are consistent with those of the corresponding regions of cDNA. The amplified DNA fragments were digested with BamHI and XbaI and ligated into BamHI/XbaI-digested pNH002 to yield pNH01 to pNH22.

Plasmids pNH38 and pNH40 harbor the *HA:AT1G70780-CPuORF(WT)* and *HA:AT4G36990-CPuORF(WT)* constructs, respectively, in which the HA tag sequence was in-frame fused to each of the AT1G70780 and AT4G36990 CPuORFs, in pSP64 Poly(A). To construct these plasmids, the BamHI–SacI fragments containing each CPuORF from pNH10 and pNH15 were separately ligated into BamHI/SacI-digested pNH001. Plasmid pNH42 harbors the *HA:AT5G53590-CPuORF(WT)* construct, in which the HA tag sequence, the second to 76th codons of the yeast dipeptidyl-aminopeptidase B (DAP2) coding sequence, and the AT5G53590 CPuORF are fused in-frame. To create this plasmid, the HA tag sequence was amplified from pNH001 using primers SP6for and HArev, and the partial *DAP2* sequence was amplified

from genomic DNA of the *S. cerevisiae* strain Y187 with primers DAP2-75aaF and DAP2-75aaR (Supplementary Table S2). These PCR fragments were fused by the overlap extension PCR (41), digested with SalI and BamHI, and then ligated into SalI/BamHI-digested pNH001 to generate pNH003. The BamHI–SacI fragment containing the AT5G53590 CPuORF from pNH16 was ligated into BamHI/SacI-digested pNH003 to generate pNH42.

Plasmid pNH005 contains eight consecutive methionine codons and the 49th to 98th codons of the yeast *DAP2* coding sequence in pSP64 Poly(A) (Supplementary Table S1). To generate this plasmid, oligonucleotides M8for and M8rev (Supplementary Table S2) were annealed and ligated into SalI/BamHI-digested pNH001. The resulting plasmid was used as a template to amplify the eight consecutive methionine codon sequence, using primers SP6for and M8rev2 (Supplementary Table S2). The partial *DAP2* coding sequence was amplified from genomic DNA of the *S, cerevisiae* strain Y187 with primers DAP2-50aaF and DAP2-50aaR (Supplementary Table S2). These two PCR fragments were fused by the overlap extension PCR method using primers listed in Supplementary Tables S1 and S2, digested with SalI and BamHI, and then ligated into SalI/BamHI-digested pNH001 to generate pNH005. Plasmids pNH44, pNH55 and pNH65 carry the *M8:CPuORF(WT)* constructs, in which each of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs was fused in-frame downstream of the partial *DAP2* coding sequence of pNH005. To construct these plasmids, the BamHI–SacI fragments containing each CPuORF from pNH10, pNH15 and pNH16 were separately ligated into BamHI/SacI-digested pNH005.

Plasmid pKM56 harbors the cauliflower mosaic virus 35S RNA (35S) promoter, the emerald luciferase coding sequence with a PEST sequence (ELuc-PEST), and the polyadenylation signal of the *A. thaliana HSP18.2* gene in pUC19. To construct this plasmid, we amplified the ELuc-PEST coding sequence from pELuc(PEST)-test (Toyobo) with primers ELucfor and ELucrev (Supplementary Table S2). The amplified fragment was digested with XbaI and SacI and ligated into the XbaI and SacI sites between the 35S promoter and the polyadenylation signal of the *Agrobacterium tumefaciens* nopaline synthase (*NOS*) gene of plasmid pIE0 (37) to yield plasmid pKM6. The region containing the polyadenylation signal of the *HSP18.2* gene was amplified from *A. thaliana* (Col-0 ecotype) genomic DNA by PCR with primers HSPterF and HSPterR (Supplementary Table S2). The amplified fragment was digested with SacI and EcoRI and ligated into SacI/EcoRI–digested pKM6 to replace the *NOS* polyadenylation signal sequence.

Plasmids pNH80, pNH84, and pNH88 carry the AT1G70780, AT4G36990, and AT5G53590 5′-UTRs, respectively, between the 35S promoter and the ELuc-PEST coding region, and were used as reporter plasmids in transient expression assays. To generate these plasmids, the regions containing the 35S promoter and each 5′-UTR were amplified from plasmid pIE0 and cDNA described above, respectively, with the primers listed in Supplementary Tables S1 and S2, and were fused using the overlap extension PCR. The primers for amplifying the entire 5′-UTRs were designed based on full-length cDNA sequence information and transcription start site data available at the Arabidopsis Information Resource (TAIR) (http://arabidopsis.org/) and plant promoter database 3.0 (http://ppdb.agr.gifu-u.ac.jp/ppdb/cgi-bin/index.cgi), respectively. The fused DNA fragments were digested with EcoRV and SalI and ligated into EcoRV/SalI-digested pKM56 to yield pNH80, pNH84 and pNH88.

Plasmid pKM5 harbors the 35S promoter, the *Renilla* luciferase (RLuc) coding sequence, and the *NOS* polyadenylation signal in pUC19, and was used as an internal control plasmid in transient expression assays. To create this plasmid, the RLuc coding sequence was amplified from pIE0 using primers RLucATGfor and NOSterR (Supplementary Table S2). The amplified fragment was digested with XbaI and SacI and ligated into XbaI/SacI-digested pBI221 (Clontech) to replace the coding sequence of the *uidA* gene.

Mutations were introduced into the CPuORFs on the plasmids described above by the overlap extension PCR method using primers listed in Supplementary Tables S1 and S2. For all constructs, the integrity of the PCR-amplified regions was confirmed by sequence analysis.

In vitro transcription

To prepare template DNAs for *in vitro* transcription, pKM04 (42), which carry a GST coding sequence in the pSP64 poly(A) vector, and the plasmids containing a *GST:CPuORF*, *HA:CPuORF* or *M8:CPuORF* construct in the pSP64 poly(A) vector were linearized with EcoRI and purified by ethanol precipitation, except for the analysis of the AT3G12010 CPuORF. Because the AT3G12010 CPuORF contains an EcoRI site, a linear DNA template containing this CPuORF was prepared by PCR. The region containing the SP6 promoter, the *GST:AT3G12010-CPuORF(WT)* construct, and the 30 nt poly(A) tail was amplified from pNH6 (Supplementary Table S1) with KOD-FX (Toyobo) and primers SP6for and SP63′rev (Supplementary Table S2). The PCR product was purified using the QIAquick Gel Extraction kit (Qiagen). *In vitro* transcription in the presence of a cap analog $m^7G[5']ppp[5']GTP$ was performed using the AmpliCap SP6 High Yield Message Maker kit (Cellscript), and transcribed RNA was purified using the RNeasy Mini kit (Qiagen).

In vitro translation

The wheat germ extract (WGE) *in vitro* translation system was prepared according to Ericson *et al.* (43) using wheat germ purchased from Sigma-Aldrich. *In vitro* translation reactions were performed in 10 μl of a modified WGE reaction mixture containing 4 μl WGE, 36 mM HEPES–KOH (pH 7.6), 10 mM creatine phosphate, 50 μg ml^{-1} creatine kinase, 5 mM dithiothreitol (DTT), 2.1 mM $Mg(OAc)_2$, 53 mM KOAc, 0.4 mM spermidine, 1.2 mM ATP, 0.1 mM GTP, 80 μM amino acid mixture minus leucine and cysteine (Promega), 40 μM leucine and cysteine, and 0.8 units μl^{-1} RNase inhibitor (Promega) at 25°C, except for experiments with [^{35}S]methionine labeling. For [^{35}S]methionine labeling, 80 μM amino acid mixture minus methionine (Promega),

0.1 μM L-methionine, and 0.1 μl of [^{35}S]methionine (37 TBq mmol^{-1}, Perkin Elmer) were added to the WGE reaction mixture instead of the amino acid mixture used for other experiments. For analyses with immunoblotting and [^{35}S]methionine labeling, 1 pmol of RNA was translated. For toeprint analysis, 200 fmol of RNA was translated.

SDS-PAGE and immunoblot analyses of *in vitro* tranlasion products

The *in vitro* translation reaction was stopped after 15 or 30 min by addition of 40 μl SDS-PAGE sample buffer (62.5 mM Tris–HCI, pH 6.8, 2% SDS, 100 mM DTT, 5% glycerol and 0.002% bromophenol blue) and boiling for 3 min. For reactions treated with RNase A, RNase A was added to the *in vitro* translation reaction mixture to a final concentration of 0.5 mg ml^{-1}. The reaction mixture was further incubated at 37°C for 15 min, followed by addition of 39 μl SDS-PAGE sample buffer and boiling for 3 min. [^{35}S]methionine-labelled translation products from *M8:CPuORF* RNAs were separated on a NuPAGE 4–12% Bis–Tris gel with MES running buffer (Life Technologies), dried, and visualized using a FLA-7000 image analyzer (Fuji Photo Film). Translation products from *GST:CPuORF* and *HA:CPuORF* RNAs were separated on a NuPAGE 4–12% Bis–Tris gel with MOPS–SDS and MES–SDS running buffers (Life Technologies), respectively, and were transferred to an Immobilon-P membrane (Millipore). Immunoblotting was performed with a polyclonal anti-GST antibody (Santa Cruz Biotechnology) or a polyclonal anti-HA antibody (Santa Cruz Biotechnology), and bands were visualized using the Immobilon Western Chemiluminescent HRP Substrate (Millipore).

Toeprint analysis

Toeprint analysis was performed as described by Wang and Sachs (19) with the following modifications. The oligonucleotides AT1G70780TP2, HsfB1TP1 and SAUR30TP1 (Supplementary Table S2) were radiolabeled at their 5′ ends with T4 polynucleotide kinase (Takara Bio) and [γ-^{32}P] ATP (110 TBq mmol^{-1}, Perkin Elmer). The ^{32}P-labeled oligonucleotides were purified using a QIAquick Nucleotide Removal kit (Qiagen). After *in vitro* translation of *GST:CPuORF* RNA in WGE, a 1-μl aliquot of the translation reaction mixture was mixed with 7.7 μl of precooled reverse transcription buffer (19). For analysis of the AT1G70780 CPuORF, the mixture was heated at 50°C for 3 min. This denaturing step was omitted for analysis of the AT4G36990 and AT5G53590 CPuORFs because the result did not change irrespective of the presence or absence of this step in preliminary experiments. After adding 1 μl of ^{32}P-labeled gene-specific primer AT1G70780TP2, HsfB1TP1 or SAUR30TP1, each mixture was incubated at 37°C for 3 min. Then, 60 units of SuperScript III RNase H^{-} reverse transcriptase (Thermo Fisher Scientific) were added, and the reaction mixture was incubated at 37°C for 30 min. The reverse transcription products were purified by phenol–chloroform extraction and ethanol precipitation. Sequence ladders were generated from plasmids pNH10, pNH15 and pNH16 (Supplementary Table S1) with Terminator DNA Polymerase (New England BioLabs) using the ^{32}P-labeled primers AT1G70780TP2, HsfB1TP1 and SAUR30TP1, respectively. The reverse transcription products and the sequence ladders were separated on an 8% polyacrylamide/7M urea gel. The gel image was analyzed using a FLA-7000 image analyzer (Fuji Photo Film).

Transient expression assay

To prepare protoplasts, MM2d suspension cells (39) were collected by centrifugation on the third day after transfer to fresh media, and suspended in modified LS medium containing 1% (w/v) cellulase Onozuka RS (Yakult Pharmaceutical Industry), 0.5% (w/v) pectolyase Y23 (Seishin Pharmaceutical), and 0.4 M mannitol, and incubated at 26°C with gentle shaking until the suspension became turbid with protoplasts. The protoplasts were then washed five times with wash buffer (0.4 M mannitol, 5 mM $CaCl_2$, and 0.5 M 2-(*N*-morpholino)ethanesulfonic acid, pH 5.8) and suspended in MaMg solution (5 mM morpholinoethanesulfonic acid, 15 mM $MgCl_2$ and 0.4 M mannitol, pH 5.8). Reporter plasmid DNA (1.5 μg) and the internal control plasmid DNA (1.5 μg) were mixed with 1.5×10^5 protoplasts in 100 μl of MaMg solution and 103 μl of polyethylene glycol (PEG) solution (40% PEG4000, 0.1 mM $CaCl_2$ and 0.2 M mannitol). This mixture was incubated for 15 min at room temperature, and diluted by adding 800 μl of wash buffer. The protoplasts were centrifuged and resuspended in 1 ml of the modified LS medium containing 0.4 M mannitol. After 24 h of incubation at 22°C in the dark, cells were harvested and disrupted in 200 μl of extraction buffer [100 mM$(NaH_2/Na_2H)PO_4$ and 5 mM DTT, pH 7] by sonication on ice with a Branson Sonifier 250. A Dual-Luciferase Reporter Assay kit (Promega) was used to measure the RLuc and ELuc activities.

RESULTS

Identifying novel sequence-dependent arrest uORFs in *A. thaliana*

For analysis of the arrest-causing ability, we selected 22 *A. thaliana* CPuORFs belonging to distinct HGs in which a sequence-dependent regulatory effect of any member has not been investigated. We excluded two uncharacterized CPuORF families, HG22 and HG30 (the numbering of HG follows Jorgensen and Dorantes-Acosta (34)), from the analysis. The amino acid sequences of HG22 CPuORFs are conserved only among the same order (29), whereas those of the other CPuORF families are conserved among a wide range of angiosperms or dicots (29,31). The amino acid sequences of HG30 CPuORFs are similar to those of mammalian and fungal CDC26, a component of the anaphase-promoting complex/cyclosome; the peptides encoded by the HG30 CPuORFs are therefore unlikely to act as arrest peptides (30,31,34). In HGs containing more than one *A. thaliana* CPuORF, the CPuORF with the most highly conserved amino acid sequence was selected.

First, the *in vitro* translation products of the 22 CPuORFs were analyzed. If a translation reaction is arrested at a specific site before the completion of translation, a specific

peptidyl-tRNA is expected to accumulate during the *in vitro* translation reaction, as reported previously (24,25,40,44–47). To detect the CPuORF translation products, a GST tag sequence was fused in-frame to each CPuORF. The resulting *GST:CPuORF(WT)* constructs were transcribed *in vitro* to yield capped RNAs with poly(A) tails of 30 nt. *GST:CPuORF(WT)* RNAs containing each of the 22 CPuORFs were translated for 30 min using a WGE *in vitro* translation system. Translation products were separated by SDS-PAGE and detected by immunoblotting with an anti-GST antibody. Although bands with higher molecular weight than the corresponding full-length (FL) peptide product were detected in all the CPuORFs examined, 15 CPuORFs showed one or more specific bands with higher intensity (Figure 1B, open arrowheads) than the non-specific band of approximately 40 kDa (Figure 1B, closed arrowhead), which was observed even when the GST tag sequence alone was translated. The sizes of these specific bands were consistent with the expected sizes of peptidyl-tRNAs accumulated owing to translation arrest in the CPuORFs, because the apparent molecular weight of tRNA on SDS-PAGE is approximately 20 kDa. Therefore, we selected these 15 CPuORFs as candidate arrest uORFs.

Next, to identify sequence-dependent arrest uORFs among the 15 candidates, we tested the effect of altering the peptide sequences of the 15 CPuORFs on the accumulation of the putative peptidyl-tRNAs. To alter many amino acid residues in the conserved region of these CPuORFs with minimum nucleotide changes, a +1 or −1 frameshift (fs) mutation was introduced upstream or within the conserved region of each CPuORF, and another fs mutation was introduced before the stop codon to allow the reading frame to be shifted back to the original frame (Supplementary Figure S1). Comparison of *in vitro* translation products between the WT and fs mutant versions of each *GST:CPuORF* RNA by immunoblotting identified a band with clearly higher intensity in the WT CPuORF than in the fs mutant CPuORF in each of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs (Figure 1C, open arrowheads). As shown in Figure 1D, these bands disappeared when the *in vitro* translation products were treated with RNase A before the immunoblot analysis. This indicated that the WT CPuORF sequence-dependent bands contained an RNA moiety, suggesting that they were peptidyl-tRNAs.

To further establish the sequence dependence of the effects of the three CPuORFs on the accumulation of the peptidyl-tRNAs, five amino acids within the highly conserved region of each CPuORF were scrambled (Figures 2 and 3A). For the analysis of the effect of this mutation, an HA tag sequence was fused to the WT and scramble (sc) mutant versions of the CPuORFs (Figure 3A) to rule out the possibility that the effects of these CPuORFs on the accumulation of the peptidyl-tRNAs were artifacts due to fusing the GST tag sequence to the CPuORFs. For the AT5G53590 CPuORF, the second to 76th codons of the yeast *DAP2* coding sequence was fused in-frame between the HA tag sequence and the CPuORF as a linker sequence (Figure 3A) because the intensity of immunobloting signals was very low when the HA tag sequence was directly fused to the AT5G53590 CPuORF (data not shown). When RNAs containing the HA tag sequence and each of the three WT CPuORFs were translated in WGE and translation products were detected by immunoblotting with an anti-HA antibody, a band with higher molecular weight than the FL band by approximately 20 kDa was observed in each case (Figure 3B, lanes 1, 5 and 9, open arrowheads). These bands disappeared after treatment of the translation products with RNase A, suggesting that they were peptidyl-tRNAs (Figure 3B, lanes 2, 6 and 10). The sc mutation introduced in each CPuORF reduced the accumulation of the peptidyl-tRNA (Figure 3B, lanes 3, 7 and 11), although the effect of the sc mutation in the AT5G53590 CPuORF was relatively weak. Together with the effects of the fs mutations on the accumulation of the peptidyl-tRNAs, these results suggested that the AT1G70780, AT4G36990, and AT5G53590 CPuORFs cause translation arrest at a specific site in a sequence-dependent manner.

Identification of critical regions of the newly identified arrest uORFs for the ribosomal arrest

To determine important regions of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs for the ribosomal arrest, N-terminal truncation mutagenesis and stop codon-scanning mutagenesis were performed. In the stop codon-scanning mutagenesis, the stop codon of each CPuORF was displaced to several positions around the 3′ end of the highly conserved region of each CPuORF, by substituting a stop codon for an amino acid-specifying codon (Figures 2 and 4). In the *51A/52* and *37A/38* mutants of the AT1G70780 and AT4G36990 CPuORFs, in which the stop codon was moved one codon downstream from the original position, an alanine codon was substituted for the original stop codon, and the three nucleotides located immediately downstream of the original stop codon was changed to a stop codon (Figure 4A and B). For these analyses, we used [^{35}S]methionine labeling to visualize *in vitro* translation products of the CPuORFs instead of immunoblotting, because alteration in the size of the *in vitro* translation products by truncation mutations may affect the transfer efficiency of translation products to the nitrocellulose membrane used for immunoblotting. Eight consecutive methionine codons and a linker sequence were fused in-frame to each of the CPuORFs (Figure 4) to facilitate the detection of *in vitro* translation products by [^{35}S]methionine labeling. The resulting construct was designated as *M8:CPuORF*. Before analyzing the truncation and stop codon-scanning mutants, we compared *in vitro* translation products of the WT and fs and sc mutant versions of *M8:CPuORF* RNAs containing each of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs. The *M8:CPuORF* RNAs were translated in WGE supplemented with [^{35}S]methionine, and the translation products were analyzed by SDS-PAGE followed by autoradiography. In each of the CPuORFs, in addition to the WT sequence-dependent band with higher molecular weight than the FL band by approximately 20 kDa (Figure 4A–C, lane 1, open arrowheads), non-specific bands were observed around the WT sequence-dependent band (Figure 4A–C, lanes 1–3). The presence of these non-specific bands may make it difficult to assess the effects of truncation and stop codon-scanning mutations on the sequence-dependent peptidyl-

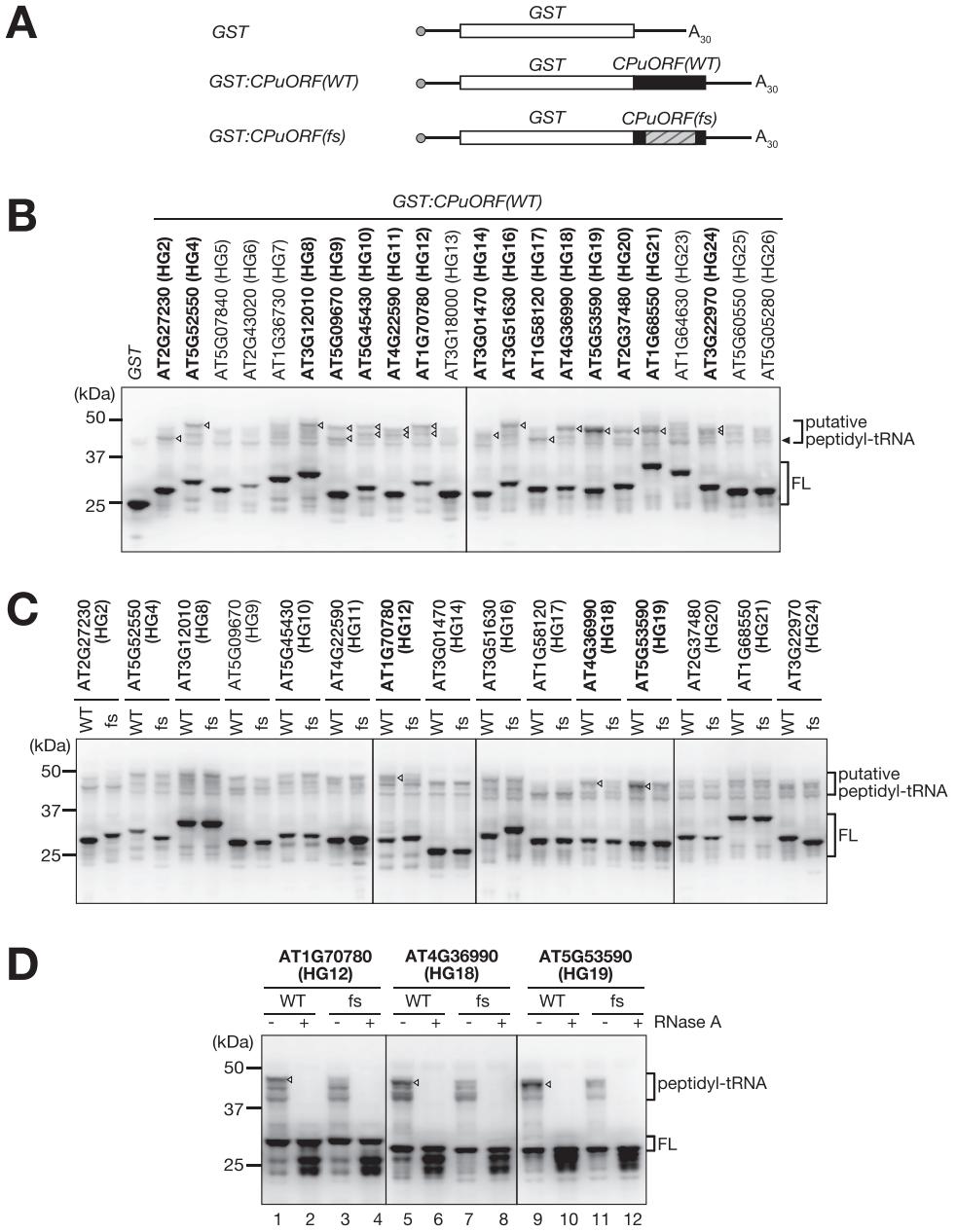

Figure 1. Search for sequence-dependent arrest uORFs by analyzing *in vitro* translation products of *A. thaliana* CPuORFs. RNAs harboring the *GST:CPuORF* constructs were translated in WGE for 30 min. Translation products were separated by SDS-PAGE and immunoblotted with an anti-GST antibody. The positions of molecular mass standards and full-length translation products (FL) are indicated on the left and right of each immunoblot, respectively. For each panel, a representative result of at least two independent experiments is shown. (**A**) Schematic representation of RNAs used for *in vitro* translation. The hatched box in the *GST:CPuORF(fs)* RNA indicates the frame-shifted region. See Supplementary Figure S1 for the exact positions of the fs mutations in each CPuORF. (**B**) Analysis of *in vitro* translation products of 22 CPuORFs. *GST* and 22 *GST:CPuORF(WT)* RNAs were translated in WGE. The open and closed arrowheads mark the specific and non-specific bands with higher molecular weight than the FL products, respectively. The AGI codes of the genes containing the analyzed CPuORFs are indicated, and those in which one or more specific bands were detected with higher intensity than the non-specific band observed in the same lane are shown in bold. (**C**) Effects of the CPuORF sequences on accumulation of *in vitro* translation products. The *GST:CPuORF(WT)* and *GST:CPuORF(fs)* RNAs containing each of 15 CPuORFs were translated in WGE. The open arrowheads indicate WT CPuORF sequence-dependent bands. The AGI codes of the genes containing the CPuORFs in which WT sequence-dependent bands were observed are shown in bold. (**D**) RNase A treatment of the *in vitro* translation products. After *GST:CPuORF(WT)* and *GST:CPuORF(fs)* RNAs containing each of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs were translated in WGE, the translation mixtures were treated with (+) or without (−) RNase A before separation by SDS-PAGE. The open arrowheads indicate WT CPuORF sequence-dependent peptidyl-tRNAs.

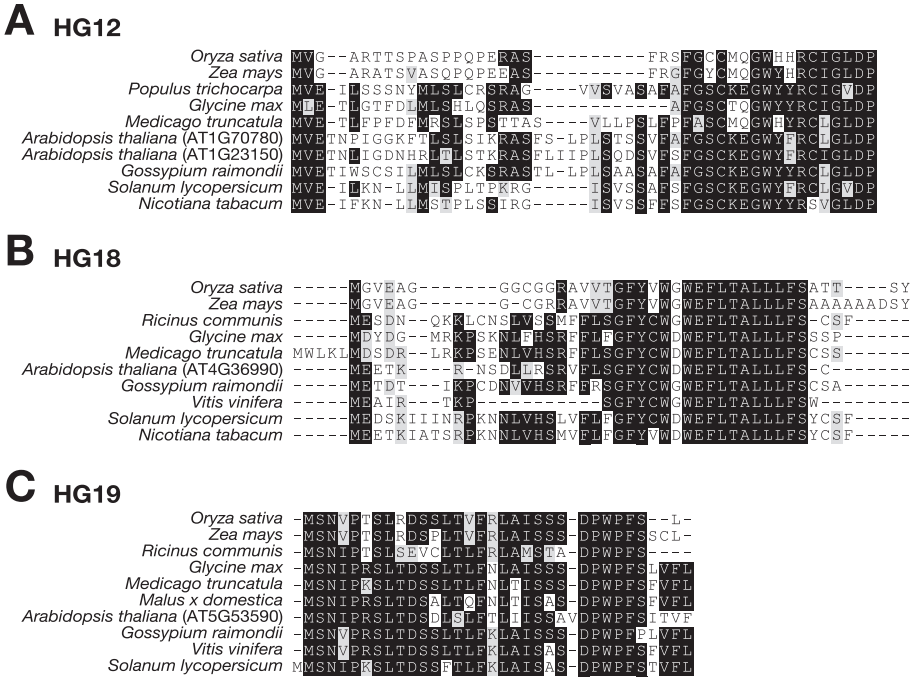

Figure 2. Peptide sequence alignments of CPuORF HGs of the newly identified arrest uORFs. (A–C) Amino acid sequence alignments of HG12 (**A**), HG18 (**B**), and HG19 (**C**) CPuORFs are shown. Putative orthologous genes of AT1G70780 (A), AT4G36990 (B) and AT5G53590 (C) in other plant species were identified using BLASTP. All deduced peptide sequences of the CPuORFs are derived from RefSeq sequences (accession numbers are indicated in Supplementary Table S3). The peptide sequences were aligned using Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and visualized using Jalview 2.9 (http://www.jalview.org).

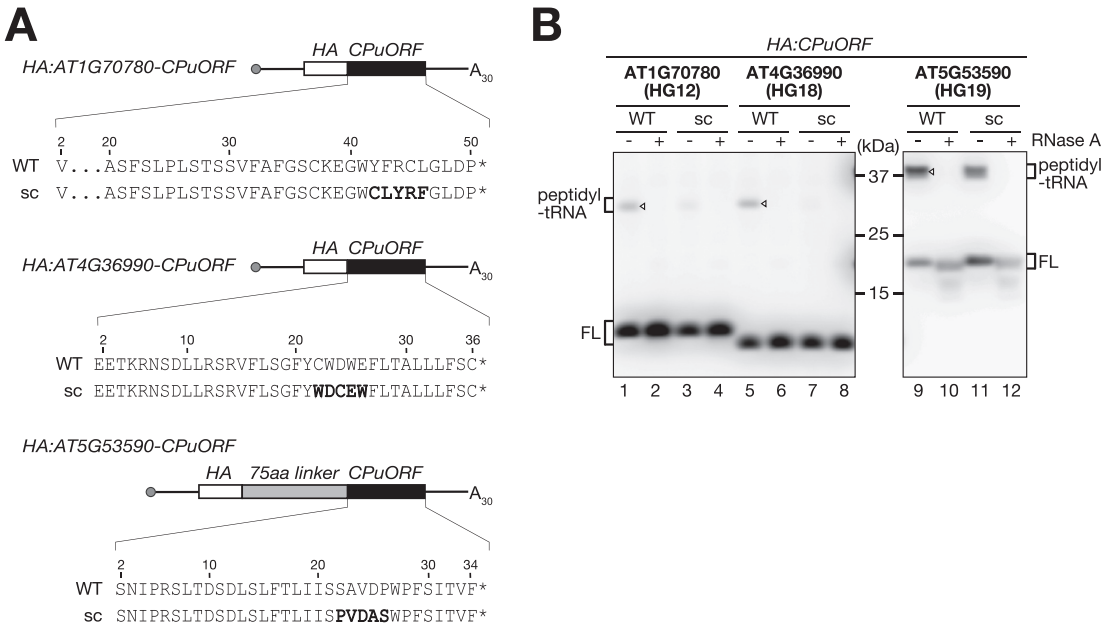

Figure 3. Effects of scrambling the CPuORF amino acid sequences on translation arrest. (**A**) Schematic representation of RNAs used for *in vitro* translation. The amino acid sequences of the WT and scramble (sc) mutant versions of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs are indicated. The scrambled amino acids in the sc mutant sequences are shown in bold. (**B**) RNAs encoding the constructs shown in (A) were translated in WGE for 15 min. Translation products were treated with (+) or without (−) RNase A, and then separated by SDS-PAGE and immunoblotted with an anti-HA antibody. The positions of molecular mass standards and full-length translation products (FL) are indicated. The open arrowheads mark the WT CPuORF sequence-dependent peptidyl-tRNAs. For each panel, a representative result of two independent experiments is shown.

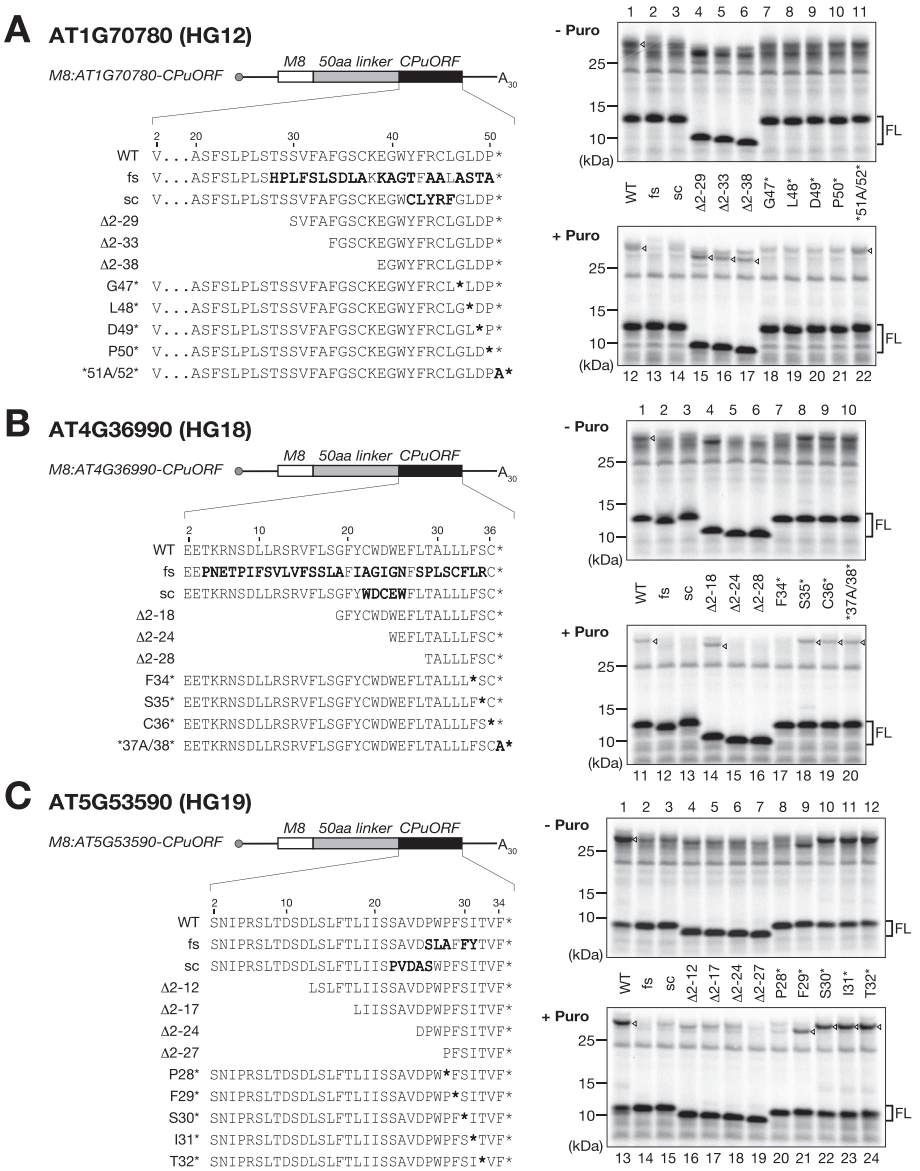

Figure 4. Identification of critical regions of the arrest uORFs. (A–C) N-terminal truncation and stop codon scanning mutagenesis. Schematic representation of *M8:CPuORF* RNAs and the amino acid sequences of the WT and mutated versions of the AT1G70780 (**A**), AT4G36990 (**B**), and AT5G53590 (**C**) CPuORFs are indicated. Asterisks represent stop codons. Altered amino acids in the mutants are shown in bold. *M8:CPuORF* RNAs containing the WT and mutated versions of each CPuORF were separately translated in WGE supplemented with [^{35}S]methionine for 15 min. Translation products were treated with (+ Puro) or without (− Puro) 2 mM puromycin for 30 s, separated by SDS-PAGE and visualized by autoradiography. The positions of molecular mass standards and full-length translation products (FL) are indicated. The open arrowheads mark the WT CPuORF sequence-dependent peptidyl-tRNAs. For each panel, a representative result of two independent experiments is shown.

tRNA accumulation. To reduce non-specific bands and to detect the sequence-dependent bands more clearly, the *in vitro* translation mixtures were treated with puromycin, a translation elongation inhibitor, after the translation reaction. Puromycin acts as an aminoacyl-tRNA analog and accepts a nascent peptide in the peptidyl transfer reaction, resulting in the releases of the nascent peptide from a peptidyl-tRNA (48,49). It has been shown in many systems that peptidyl-tRNAs bound to ribosomes stalled by arrest peptides have a lower reactivity to puromycin than those on ribosomes that are not stalled (3,14,25,47,50–54). When the *in vitro* translation mixtures were treated with 2 mM puromycin for 30 s after the translation reaction, non-specific translation products around the sequence-dependent bands were reduced, and therefore the sequence-dependent bands were more clearly detected as puromycin-resistant bands (Figure 4A, lanes 12–14, 4B, lanes 11–13, 4C, lanes 13–15). When the translation products were treated with RNase A, the sequence-dependent bands disappeared, suggesting that they were peptidyl-tRNAs (Supplementarry Figure S2). It should be noted that the sc mutation in the AT5G53590 CPuORF showed a weaker effect than the fs mutation, consistent with the weak effect of the sc mutation observed in

the immunoblot analysis using the HA tag (Figure 4C, lanes 13–15, Supplementarry Figure S2C, lanes 7, 9 and 11).

To investigate the effects of N-terminal truncation and stop codon displacement of the CPuORFs on the ribosomal arrest, the *M8:CPuORF* RNAs containing the WT or mutated versions of each CPuORF were translated for 30 min in the presence of [^{35}S]methionine, and then the translation mixtures were treated with puromycin for 30 s. As shown in Figure 4A, even when the AT1G70780 CPuORF was truncated up to codon 38 (Δ2–38), the intensity of the specific puromycin-resistant peptidyl-tRNA band did not appreciably decrease, although the band size was smaller than that of WT because of the truncation (lanes 12 and 15–17, open arrowheads). When the stop codon of the AT1G70780 CPuORF was moved upstream from the original position (G47*, L48*, D49* and P50*), the intensity of the specific puromycin-resistant band was diminished, whereas moving the stop codon one codon downstream (*51A/52*) did not affect the band intensity (lanes 12 and 18–22, open arrowheads). These results suggest that the region from codons 39 to 50 is sufficient for the full ability of the AT1G70780 CPuORF to cause ribosomal arrest. The importance of this region is supported by the fact that the five codons scrambled in the sc mutant are included in this region.

For the AT4G36990 CPuORF, deletion of the region from codons 2 to 24 (Δ2–24) or 28 (Δ2–28) reduced the intensity of the specific puromycin-resistant band to the background level, whereas deletion of the region from codons 2 to 18 (Δ2–18) did not (Figure 4B, lanes 11 and 14–16). When a stop codon was substituted for codon 34 (F34*), the intensity of the specific puromycin-resistant band markedly decreased, whereas moving the stop codon up to codon 35 (S35*) did not affect the band intensity (Figure 4B, lanes 11 and 17–20). These results suggest that the region from codons 19 to 34 of the AT4G36990 CPuORF is sufficient for the sequence-dependent ribosomal arrest. Consistent with this conclusion, this region includes the five codons that were scrambled in the sc mutant.

For the AT5G53590 CPuORF, deletion of the region from codons 2 to 12 (Δ2–12) reduced the intensity of the specific puromycin-resistant band, and truncation up to codon 27 (Δ2–27) further decreased the band intensity (Figure 4C, lanes 13 and 16–19). When a stop codon was substituted for codon 29 (F29*), the intensity of the specific puromycin-resistant band was diminished and the position of the band was shifted to a lower molecular weight (Figure 4C, lane 21, open arrowhead). This result indicates that the AT5G53590 CPuORF needs to be translated up to codon 29 to cause ribosomal arrest at the same positon as WT, but that ribosomal arrest occurs to some extent at a different position even when a stop codon was substituted for codon 29. Together with the result of the N-terminal truncation mutagenesis, this suggests that the region from codons 2 to 29 of the AT5G53590 CPuORF is important for the full ability to cause ribosomal arrest.

Determination of ribosome stalling sites in the identified arrest uORFs

To obtain further evidence of ribosomal arrest and to determine the positions of the stalled ribosomes in the AT1G70780, AT4G36990, and AT5G53590 CPuORFs, we performed toeprint analysis (primer extension inhibition assay). This analysis enables the detection of stalled ribosomes as obstacles blocking reverse transcription, and their positions on mRNAs can be determined by analyzing the sizes of the aborted reverse transcription products (55). *GST:CPuORF* RNAs containing the WT and fs and sc mutant versions of each of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs were translated separately for 30 min in WGE. The translation reaction mixtures were then subjected to reverse transcription with a ^{32}P-labeled primer (see Supplementary Figure S1J, O and P for the primer positions), and the reverse transcription products were analyzed on a sequencing gel. In some previous studies of ribosome stalling using toeprint analysis, stalled ribosomes needed to be fixed with a translation elongation inhibitor before the reverse transcription reaction to obtain toeprint signals (24,40,56). In these analyses, hygromycin B could fix stalled ribosomes at the translation elongation and termination steps. Therefore, in the present study, we performed toeprint analysis with and without addition of hygromycin B at 0 min (negative control) or 30 min after starting the *in vitro* translation reaction.

As shown in Figure 5A, a WT-specific toeprint signal was detected at nt 166 of the AT1G70780 CPuORF (i.e. 15 nt downstream of the of the first nucleotide of the stop codon) when hygromycin B was added after the translation reaction (lanes 4–6, open arrowheads). The intensity of this signal decreased to the background level when hygromycin B was added to WGE prior to the translation reaction (lanes 7–9), indicating that ribosomes were responsible for the toeprint signal. In previous studies, uORF peptide-mediated ribosome stalling during translation termination gave toeprint signals 13–17 nt downstream of the first nucleotide of the stop codon (19,23,24,28). Therefore, together with the result of the stop codon scanning mutagenesis, which indicates that the AT1G70780 CPuORF needs to be translated up to the last amino acid-coding codon for the full ability to cause ribosomal arrest, this result suggests that the AT1G70780 CPuORF-encoded peptide sequence causes ribosome stalling predominantly during translation termination (Figure 5D).

For the AT4G36990 CPuORF, WT-specific toeprint signals were detected at nt 117 and nt 119 of the CPuORF (Figure 5B, lanes 1–6, open arrowheads). In previous studies, when eukaryotic translation initiation complexes and 80S ribosomes were stalled at an initiation codon by a translation inhibitor, toeprint signals were observed 15–17 nt downstream of the first nucleotide of the initiation codon (23,57,58). This relationship between the positions of toeprint signals and stalled ribosomes suggests that the signal at nt 117 represents a ribosome stalled with its P site at codon 34 (Figure 5E). This stalling site is consistent with the result of the stop codon scanning mutagenesis, which indicates that the AT4G36990 CPuORF needs to be translated up to codon 34 to cause ribosomal arrest (Figure 4B).

For the AT5G53590 CPuORF, WT-specific toeprint signals were detected at nt 102 and nt 103 of the CPuORF (Figure 5C, lanes 1–6). This result suggests that the WT sequence of the AT5G53590 CPuORF causes ribosome stalling when the P site of a ribosome is located at codon 29 (Fig-

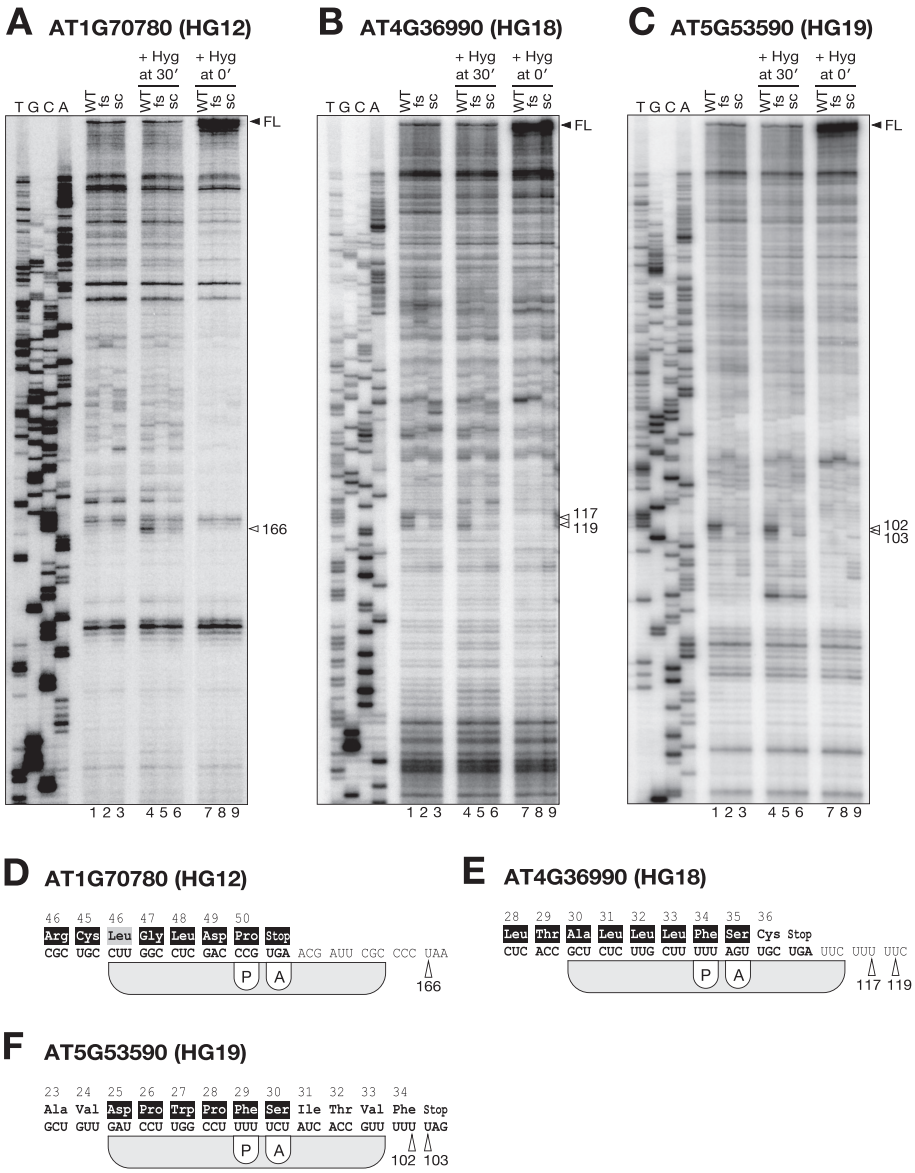

Figure 5. Determination of the positions of stalled ribosomes. (A–C) Toeprint analysis of ribosome stalling. *GST:CPuORF* RNAs containing each of the WT (lanes 1, 4 and 7) and fs (lanes 2, 5 and 8) and sc (lanes 3, 6 and 9) mutant versions of the AT1G70780 (**A**), AT4G36990 (**B**), and AT5G53590 (**C**) CPuORFs were translated separately for 30 min in WGE, and then the translation reaction mixtures were subjected to reverse transcription with ^{32}P-labeled gene-specific primers, which anneal between the CPuORFs and the poly(A) tails (see Supplementary Figure S1J, O and P for the exact positions). In lanes 4–9, hygromycin B (Hyg) was added to a final concentration of 2 mM at 30 min (lanes 4–6) or 0 min (lanes 7–9) after starting the translation reaction, and the reverse transcription reaction was performed in the presense of 2 mM hygromycin B. The closed arrowheads indicate the position of the full-length (FL) primer extension products. The open arrowheads denote WT CPuORF sequence-dependent toeprint signals. The sequence ladders (shown in the sense strand sequence) were synthesized using the same primers as those used for toeprinting. For each panel, a representative result of at least two independent experiments is shown. (**D**–**F**) Schematic representation of the positions of stalled ribosomes in the AT1G70780 (D), AT4G36990 (E), and AT5G53590 (F) CPuORFs. The nucleotide and deduced amino acid sequences of the region around the stalled positions are shown. Based on the allignments shown in Figure 2, amino acid residues conserved among angiosperm are shown in a black background, and amino acid residues with conservative substitutions are shown in a gray background. The amino acid residue numbers are indicated above the amino acid sequences. The open arrowheads mark the positions of the WT CPuORF sequence-dependent toeprint signals. The positions of stalled ribosomes and their P and A sites deduced from the specific toeprint signals are depicted.

ure 5F). This stalling site is consistent with the result of the stop codon-scanning mutagenesis, which indicates that the AT5G53590 CPuORF needs to be translated up to codon 29 for the full ability to cause ribosomal arrest (Figure 4C).

In all the three CPuORFs, the posision of P or A site of the stalled ribosome deduced from the WT-specific toeprint signals is consistent with the position of the last amino acid-coding codon of the highly conserved region (Figures 2 and 5D–F). Thus, the ribosome stalling sites determined by the toeprint analyses are in good agreement with the conservation patterns of the amino acid sequences of the CPuORFs.

The identified arrest uORFs strongly repress main ORF expression in a peptide sequence-dependent manner

Next, the effects of the newly identified arrest CPuORFs on the expression of proteins encoded by the main ORFs were investigated using a transient expression assay in *A. thaliana* protoplasts. For this analysis, the entire 5′-UTR sequences of the AT1G70780, AT4G36990, and AT5G53590 genes were fused to the ELuc-PEST coding sequence and placed under the control of the 35S promoter to generate the WT reporter constructs (Figure 6, Supplementary Figure S1J, O and P). The effects of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs on main ORF expression were tested by generating ΔAUG reporter constructs, in which the initiation codon of the CPuORF was removed from each WT reporter construct by changing the AUG initiation codon to AAG (Figure 6). In addition, to assess the importance of the peptide sequences for the effects of these three CPuORFs on main ORF expression, fs mutations were introduced into the WT and ΔAUG reporter constructs at the same position as those in the *GST:CPuORF(fs)* and *M8:CPuORF(fs)* constructs (Figures 4 and 6, Supplementary Figure S1J, O and P).

These reporter constructs were separately transfected into protoplasts prepared from *A. thaliana* MM2d suspension-cultured cells. After 24 h of incubation, cells were harvested and disrupted for analysis of luciferase activity. In all three genes analyzed, the removal of the CPuORF initiation codon markedly upregulated the expression of the reporter gene, indicating strong repressive effects of these CPuORFs on main ORF expression (Figure 6). The fs mutations introduced into these CPuORFs also strongly enhanced reporter gene expression, although to a lesser extent than elimination of the initiation codon in each case (Figure 6). This suggested that the peptide sequences are in part responsible for the strong repressive effects of these CPuORFs. By contrast, the fs mutations did not significantly enhance reporter gene expression in the absence of the initiation codons of the CPuORFs, indicating that the effects of the fs mutations depend on translation of the CPuORFs (Figure 6). These results suggest that the peptide sequences encoded by the AT1G70780, AT4G36990, and AT5G53590 CPuORFs have a strong repressive effect in protoplast cells, implying that these CPuORFs cause ribosome stalling not only in cell-free systems but also in cells and thereby strongly repress main ORF expression.

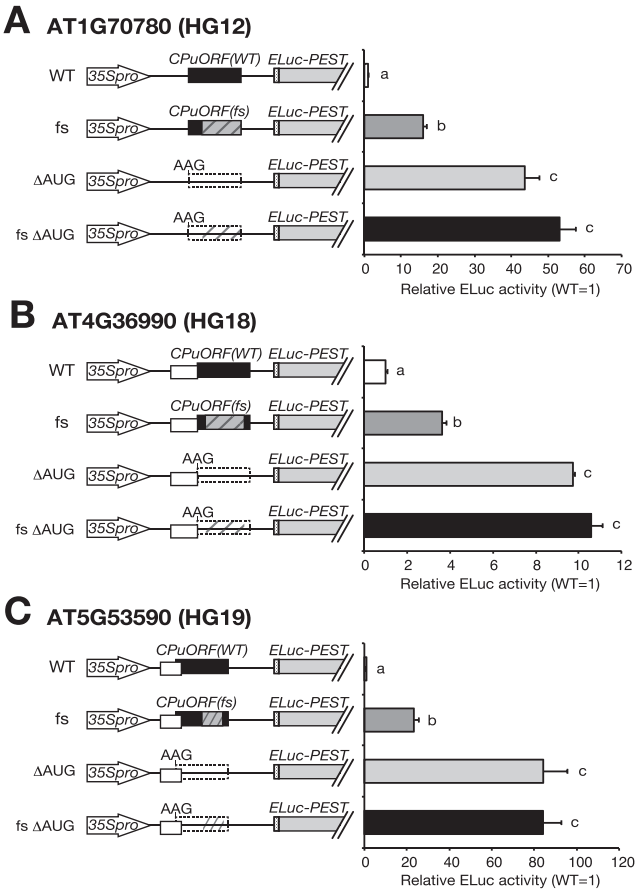

Figure 6. Effects of the newly identified arrest uORFs on main ORF expression in protoplast cells. (A–C) Transient expression assays were performed to test the regulatory effects of the AT1G70780 (**A**), AT4G36990 (**B**), and AT5G53590 (**C**) CPuORFs. Schematic structures of the reporter constructs used are shown on the left of each graph. The 5′-UTR sequence containing each CPuORF was inserted between the 35S promoter (*35Spro*) and the ELuc-PEST coding sequence. The closed boxes in the WT and fs constructs represent the CPuORFs. The hatched boxes in the fs and fs ΔAUG constructs show the frame-shifted region. The dotted lines in the ΔAUG and fs ΔAUG constructs indicate the CPuORFs eliminated by changing the ATG initiation codon to AAG. In (B) and (C), the open boxes represent uORFs other than the CPuORFs. The dotted boxes represent the first five nt of the main ORFs of the AT1G70780, AT4G36990, and AT5G53590 genes. Each reporter plasmid containing the 5′-UTR with a WT or mutated CPuORF was co-transfected with the *35S::RLuc* internal control plasmid into MM2d protoplasts by PEG treatment. After 24 h of incubation, cells were collected and disrupted for dual luciferase assay. ELuc activity was normalized to RLuc activity, and the relative activity to that of the corresponding WT reporter construct was calculated. Means ± SD of three biological replicates are shown. *P* values were calculated by Student's *t*-test and adjusted by Bonferroni's correction for multiple comparisons. Columns with different letters are significantly different at $P < 0.05$. Representative results of three experiments using independently prepared protoplasts are shown.

DISCUSSION

Only two sequence-dependent arrest uORF have been reported in plants to date. In the present study, we identified three novel *A. thaliana* uORFs that cause ribosomal arrest

in a sequence-dependent manner, thereby repressing the expression of proteins encoded by the main ORFs. In addition, we found that the peptide sequences encoded by two of these uORFs cause ribosomal arrest during translation elongation, whereas the other one causes ribosomal arrest during translation termination.

Identification of novel sequence-dependent arrest uORFs in *A. thaliana*

In the present study, analysis of *in vitro* translation products of 22 *A. thaliana* CPuORFs identified three uORFs that caused uORF sequence-dependent accumulation of specific peptidyl-tRNAs when translated *in vitro* in WGE (Figures 1 and 3). Toeprint analyses confirmed that ribosomes were stalled in these uORFs in a sequence-dependent manner (Figure 5). Thus, this study identified three novel sequence-dependent arrest uORFs in *A. thaliana*.

The present study focused on uORFs that cause ribosomal arrest in WGE under normal *in vitro* translation conditions. However, many of the previously identified arrest peptides require a low molecular weight effector molecule, such as amino acids, polyamines and antibiotics, to cause ribosomal arrest (15,19,20,23–25,40,59). In the analysis of the *in vitro* translation products of the 22 CPuORFs presented in Figure 1, 19 CPuORFs showed no sequence-dependent effect on the peptidyl-tRNA accumulation or caused only low levels of peptidyl-tRNA accumulation, despite their evolutionarily conserved amino acid sequences (29). It is possible that the peptides encoded by some of these CPuORFs cause ribosomal arrest under certain specific conditions, for example, in the presence of an effector molecule that was not contained enough in the *in vitro* translation system used in this study.

The stop codon scaning mutagenesis and toeprint analyses presented here identified the ribosome stalling sites in the newly identified arrest uORFs. These analyses showed that the AT1G70780, AT4G36990, and AT5G53590 CPuORFs causes ribosome stalling when the A sites of ribosomes are positioned at the stop, Ser-35, and Ser-30 codons, respectively (Figure 5D to F). Despite that the stop codon position of the AT1G70780 CPuORF and the serine residues encoded by codon 35 of the AT4G36990 CPuORF and codon 30 of the AT5G53590 CPuORF are highly conserved among angiosperm (Figure 2), the stop codon scaning mutagenesis showed that alterlation of these A site codons at the stalling sites to an alanine or stop codon did not affect the efficiency of the ribosomal arrest (Figure 4A, lane 22, 4B, lane 18 and 4C, lane 22). This discrepancy can be accounted for by previous studies on prokaryotic regulatory nascent peptides. In some prokaryotic translational regulation systems involving nascent peptide-mediated translation elongation arrest, the efficiency of ribosomal arrest depends on the nature of the amino acid of the aminoacyl-tRNA bound to the A site of a stalled ribosome, but the A site codon can be replaced by some other amino acid-coding codons or a stop codon without appreciably affecting the efficiency of ribosomal arrest (52,60,61). Therefore, although changing the A site codons at the stalling sites to an alanine or stop codon is tolerated in the AT1G70780, AT4G36990, and AT5G53590 CPuORFs, changing these codons to some other codons might affect the ribosomal arrest.

The positions of ribosome stalling in the three CPuORFs suggest that the AT1G70780 CPuORF causes ribosome stalling during translation termination, whereas the AT4G36990 and AT5G53590 CPuORFs cause ribosome stalling during translation elongation (Figure 5). In prokaryotes, most of the known arrest peptides cause translation elongation arrest (14,15). By contrast, among the six previously identified uORF families encoding arrest peptides in eukaryotes, only one causes translation elongation arrest (26), whereas the others show ribosome stalling at the uORF termination codon (19,20,23–25,28). The present study added two novel examples of uORFs that cause translation elongation arrest and provided the first example in plants, suggesting that translation elongation arrest caused by uORF-encoded peptides is not an exceptional mechanism even in eukaryotes.

The results of the N-terminal truncation and stop codon-scanning mutagenesis shown in Figure 4 indicate that the amino acid sequences encoded within the regions comprising 12, 16, and 28 codons of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs, respectively, are important for the ribosomal arrest. In all the three CPuORFs, the amino acid sequences encoded by the critical regions are highly conserved among angiosperm (Figure 2). In addition, the toeprint analyses showed that ribosomes are stalled immediately after translating the critical region of each of the three CPuORFs (Figures 5). These findings suggest that, if nascent peptides encoded by these CPuORFs cause the ribosomal arrest, the critical amino acid residues of the nascent peptides should act within the ribosomes that synthesized them, as reported in previously characterized arrest peptides (14), because a eukaryotic ribosomal exit tunnel holds 30–40 amino acid residues (62,63). Among the previously identified arrest uORFs, the C-terminal regions comprising 16 and 19 amino acid residues in the *gpUL4* and *arg-2* uORF-encoded nascent peptides, respectively, interact with exit tunnel components when they cause ribosomal arrest (7). The most likely possibility is that the critical amino acid residues encoded by the newly identified arrest uORFs also interact with exit tunnel components to cause ribosomal arrest in a similar fashion to the *gpUL4* and *arg-2* uORF-encoded peptides.

Regulatory roles of the arrest uORFs

In the present study, transient expression assays showed that the three newly identified arrest uORFs strongly repress main ORF expression in a peptide sequence-dependent manner (Figure 6), suggesting that ribosome stalling caused by these uORFs inhibits protein synthesis from the main ORFs. Ribosome stalling in these uORFs probably inhibits main ORF translation *via* the 'roadblock' mechanism proposed by Wang and Sachs (19), in which a ribosome stalled in a uORF blocks ribosomal scanning, thereby preventing other ribosomes from reaching the initiation codon of the main ORF.

In the three uORFs, transient assay experiments showed that removal of the uORF initiation codon had a greater effect on increasing the expression of the main ORF than

the fs mutation (Figure 6), suggesting that these uORFs exert both sequence-dependent and -independent repressive effects on main ORF expression. This can happen if the translation initiation efficiencies of these uORFs are not low. This possibility is supported by the fact that the initiation contexts of these uORFs are in part consistent with the optimal context for efficient translation initiation. In *A. thaliana*, a purine (A or G) at position −3 and a G at position +4, where the A of an AUG initiation codon is defined as +1, are the optimal context for efficient translation initiation (64,65), as established in mammals by Kozak (66,67). The nucleotides at either position −3 or +4 of the AT1G70780, AT4G36990, and AT5G53590 CPuORFs are consistent with the optimal context (Supplementary Figure S1J, O and P), suggesting that the translation initiation efficiencies of these CPuORFs are modest; however, another uORF located upstream of the AT4G36990 and AT5G53590 CPuORFs may affect the translation initiation efficiencies of these CPuORFs (Supplementary Figure S1O and P).

As mentioned above, many of the previously identified arrest peptides cause ribosomal arrest in response to low molecular weight effector molecules. By contrast, it is likely that the arrest uORFs identified in this study cause ribosomal arrest constitutively, because these uORFs caused ribosomal arrest in WGE under normal *in vitro* translation conditions (Figures 1, 3, 4 and 5) and strongly repressed main ORF expression in protoplasts cultured under normal conditions (Figure 6). Alternatively, the extent of the ribosomal arrest caused by these uORFs may be reduced under certain conditions to induce main ORF translation. Among the previously identified arrest uORFs, the peptide encoded by the second uORF of the cytomegalovirus *gp48* gene causes ribosomal arrest constitutively (28,44). The peptide encoded by the mammalian *CHOP* uORF also causes ribosomal arrest constitutively when the uORF is translated (26). However, phosphorylation of eIF2α in response to ER stress leads to reduced translation of the *CHOP* uORF, resulting in ribosomal bypass of the uORF (27). Because of bypassing the inhibitory uORF, translation of the *CHOP* main ORF is promoted under ER stress conditions. An example of a regulated release of ribosomal arrest in response to a low molecular weight molecule is found in the *S. cerevisiae OAZ1* gene. In this gene, the nascent polypeptide encoded by the main ORF causes ribosomal arrest, which inhibits ribosomal frameshifting necessary for production of the intact OAZ1 protein, and the ribosomal arrest is alleviated when cellular polyamine concentration is high (68).

Of the three genes containing the newly identified arrest uORFs, AT4G36990 encodes a heat shock transcription factor, HsfB1, which is involved in heat and biotic stress responses and whose expression is induced by heat shock and bacterial infection (69–72). One possible regulatory role of the AT4G36990 CPuORF is stress-responsive induction of main ORF translation, as seen in the translational regulation of the mammalian *CHOP* gene. This possibility is supported by previous observations that the repressive effect of the AT4G36990 CPuORF on main ORF translation is alleviated under heat or biotic stress conditions (72,73). Another possible regulatory role is the negative modulation of HsfB1 protein levels to prevent HsfB1 overproduction when its expression is induced. Transgenic *A. thaliana* plants overexpressing the HsfB1 protein exhibit a dwarf phenotype under normal growth conditions (70), and overexpression of HsfB1 induces cell death in *Nicotiana benthamiana* leaves (73). Thus, overproduction of the HsfB1 protein may be harmful for plants.

AT5G53590 encodes SMALL AUXIN UP RNA 30 (SAUR30), a member of SAUR family proteins. SAUR family genes were originally identified as genes whose expression was rapidly induced by exogenous auxin, but expression of all of the *A. thaliana* SAUR family genes is not induced in response to auxin (74). The transcript level of *SAUR30* is upregulated in response to abscisic acid (ABA), but not to auxin (75). Recent studies suggested physiological and developmental roles of several SAUR proteins (74). However, the function of SAUR30 has not been reported. The function of the protein encoded by the AT1G70780 main ORF is also unknown. Although it is difficult to surmise physiological roles of the CPuORFs of these genes until the functions of the proteins encoded by the downstream main ORFs are elucidated, one possible role is to modulate main ORF translation to an appropriate level by constitutively repressing translation. Another possibility is that the repressive effects of these CPuORFs are alleviated under specific conditions to induce main ORF translation. For example, ABA might induce expression of SAUR30 at the translational level in addition to the transcriptional level by alleviating the repressive effect of the AT5G53590 CPuORF.

CONCLUSIONS

The present study identified three novel sequence-dependent arrest uORFs in *A. thaliana* and found that two of them cause ribosome stalling during translation elongation, which suggests that uORF peptide-mediated translation elongation arrest is a more prevalent regulatory mechanism in eukaryotes than previously thought. It is likely that the identified arrest uORFs do not require any effector molecule to cause ribosomal arrest. Currently, however, the physiological roles of this type of arrest uORFs remain unclear, except for the ER stress-responsive translational regulation of the *CHOP* gene. Further studies on the physiological function of the arrest uORFs identified in this study will provide a better understanding of the roles of constitutive ribosomal arrest and may uncover novel regulatory roles of uORF-encoded arrest peptides.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Ms. Maki Mori and Ms. Kazuko Harada for general assistance. We used the Radioisotope Laboratory and the DNA sequencing facility of the Graduate School of Agriculture, Hokkaido University.

FUNDING

Grants-in-Aid for Scientific Research [16K07387 to H.O. and 22119006 and 15H01525 to S.N.] from the Ministry

of Education, Culture, Sports, Science and Technology of Japan. Funding for open access charge: Japan Society for the Promotion of Science [Grant-in-Aid for Scientific Research (C) (16K07387) to H.O.].

Conflicts of interest statement. None declared.

REFERENCES

1. Nakatogawa,H. and Ito,K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell*, **108**, 629–636.
2. Cruz-Vera,L.R., Rajagopal,S., Squires,C. and Yanofsky,C. (2005) Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of tna operon expression. *Mol. Cell*, **19**, 333–343.
3. Vazquez-Laslop,N., Thum,C. and Mankin,A.S. (2008) Molecular mechanism of drug-dependent ribosome stalling. *Mol. Cell*, **30**, 190–202.
4. Lawrence,M.G., Lindahl,L. and Zengel,J.M. (2008) Effects on translation pausing of alterations in protein and RNA components of the ribosome exit tunnel. *J. Bacteriol.*, **190**, 5862–5869.
5. Yap,M.N. and Bernstein,H.D. (2009) The plasticity of a translation arrest motif yields insights into nascent polypeptide recognition inside the ribosome tunnel. *Mol. Cell*, **34**, 201–211.
6. Chiba,S., Lamsa,A. and Pogliano,K. (2009) A ribosome-nascent chain sensor of membrane protein biogenesis in *Bacillussubtilis*. *EMBO J.*, **28**, 3461–3475.
7. Bhushan,S., Meyer,H., Starosta,A.L., Becker,T., Mielke,T., Berninghausen,O., Sattler,M., Wilson,D.N. and Beckmann,R. (2010) Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Mol. Cell*, **40**, 138–146.
8. Wu,C., Wei,J., Lin,P.J., Tu,L., Deutsch,C., Johnson,A.E. and Sachs,M.S. (2012) Arginine changes the conformation of the arginine attenuator peptide relative to the ribosome tunnel. *J. Mol. Biol.*, **416**, 518–533.
9. Martinez,A.K., Shirole,N.H., Murakami,S., Benedik,M.J., Sachs,M.S. and Cruz-Vera,L.R. (2012) Crucial elements that maintain the interactions between the regulatory TnaC peptide and the ribosome exit tunnel responsible for Trp inhibition of ribosome function. *Nucleic Acids Res.*, **40**, 2247–2257.
10. Martínez,A.K., Gordon,E., Sengupta,A., Shirole,N., Klepacki,D., Martinez-Garriga,B., Brown,L.M., Benedik,M.J., Yanofsky,C., Mankin,A.S. *et al.* (2014) Interactions of the TnaC nascent peptide with rRNA in the exit tunnel enable the ribosome to respond to free tryptophan. *Nucleic Acids Res.*, **42**, 1245–1256.
11. Arenz,S., Meydan,S., Starosta,A.L., Berninghausen,O., Beckmann,R., Vázquez-Laslop,N. and Wilson,D.N. (2014) Drug sensing by the ribosome induces translational arrest via active site perturbation. *Mol. Cell*, **56**, 446–452.
12. Arenz,S., Ramu,H., Gupta,P., Berninghausen,O., Beckmann,R., Vázquez-Laslop,N., Mankin,A.S. and Wilson,D.N. (2014) Molecular basis for erythromycin-dependent ribosome stalling during translation of the ErmBL leader peptide. *Nat. Commun.*, **5**, 3501.
13. Sohmen,D., Chiba,S., Shimokawa-Chiba,N., Innis,C.A., Berninghausen,O., Beckmann,R., Ito,K. and Wilson,D.N. (2015) Structure of the Bacillus subtilis 70S ribosome reveals the basis for species-specific stalling. *Nat. Commun.*, **6**, 6941.
14. Ito,K. and Chiba,S. (2013) Arrest peptides: *cis*-acting modulators of translation. *Annu. Rev. Biochem.*, **82**, 171–202.
15. Lovett,P.S. and Rogers,E.J. (1996) Ribosome regulation by the nascent peptide. *Microbiol. Rev.*, **60**, 366–385.
16. Morris,D.R. and Geballe,A.P. (2000) Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.*, **20**, 8635–8642.
17. Cruz-Vera,L.R., Sachs,M.S., Squires,C.L. and Yanofsky,C. (2011) Nascent polypeptide sequences that influence ribosome function. *Curr. Opin. Microbiol.*, **14**, 160–166.
18. Calvo,S.E., Pagliarini,D.J. and Mootha,V.K. (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 7507–7512.
19. Wang,Z. and Sachs,M.S. (1997) Ribosome stalling is responsible for arginine-specific translational attenuation in *Neurosporacrassa*. *Mol. Cell. Biol.*, **17**, 4904–4913.
20. Wang,Z., Gaba,A. and Sachs,M.S. (1999) A highly conserved mechanism of regulated ribosome stalling mediated by fungal arginine attenuator peptides that appears independent of the charging status of arginyl-tRNAs. *J. Biol. Chem.*, **274**, 37565–37574.
21. Ruan,H., Shantz,L.M., Pegg,A.E. and Morris,D.R. (1996) The upstream open reading frame of the mRNA encoding *S*-adenosylmethionine decarboxylase is a polyamine-responsive translational control element. *J. Biol. Chem.*, **271**, 29576–29582.
22. Hanfrey,C., Elliott,K.A., Franceschetti,M., Mayer,M.J., Illingworth,C. and Michael,A.J. (2005) A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J. Biol. Chem.*, **280**, 39229–39237.
23. Law,G.L., Raney,A., Heusner,C. and Morris,D.R. (2001) Polyamine regulation of ribosome pausing at the upstream open reading frame of *S*-adenosylmethionine decarboxylase. *J. Biol. Chem.*, **276**, 38036–38043.
24. Uchiyama-Kadokura,N., Murakami,K., Takemoto,M., Koyanagi,N., Murota,K., Naito,S. and Onouchi,H. (2014) Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the *AdoMetDC1* gene triggers nonsense-mediated mRNA decay in *Arabidopsis thaliana*. *Plant Cell Physiol.*, **55**, 1556–1567.
25. Yamashita,Y., Takamatsu,S., Glasbrenner,M., Becker,T., Naito,S. and Beckmann,R. (2017) Sucrose sensing through nascent peptide-meditated ribosome stalling at the stop codon of *Arabidopsis bZIP11* uORF2. *FEBS Lett.*, **591**, 1266–1277.
26. Young,S.K., Palam,L.R., Wu,C., Sachs,M.S. and Wek,R.C. (2016) Ribosome elongation stall directs gene-specific translation in the integrated stress response. *J. Biol. Chem.*, **291**, 6546–6558.
27. Palam,L.R., Baird,T.D. and Wek,R.C. (2011) Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance *CHOP* translation. *J. Biol. Chem.*, **286**, 10939–10949.
28. Cao,J. and Geballe,A.P. (1996) Coding sequence-dependent ribosomal arrest at termination of translation. *Mol. Cell. Biol.*, **16**, 603–608.
29. Hayden,C.A. and Jorgensen,R.A. (2007) Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol.*, **5**, 32.
30. Vaughn,J.N., Ellingson,S.R., Mignone,F. and Arnim,A. (2012) Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA*, **18**, 368–384.
31. Takahashi,H., Takahashi,A., Naito,S. and Onouchi,H. (2012) BAIUCAS: a novel BLAST-based algorithm for the identification of upstream open reading frames with conserved amino acid sequences and its application to the *Arabidopsis thaliana* genome. *Bioinformatics*, **28**, 2231–2241.
32. Crowe,M.L., Wang,X.Q. and Rothnagel,J.A. (2006) Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides. *BMC Genomics*, **7**, 16.
33. Hayden,C.A. and Bosco,G. (2008) Comparative genomic analysis of novel conserved peptide upstream open reading frames in *Drosophila melanogaster* and other dipteran species. *BMC Genomics*, **9**, 61.
34. Jorgensen,R.A. and Dorantes-Acosta,A.E. (2012) Conserved peptide upstream open reading frames are associated with regulatory genes in angiosperms. *Front. Plant Sci.*, **3**, 191.
35. Imai,A., Hanzawa,Y., Komura,M., Yamamoto,K.T., Komeda,Y. and Takahashi,T. (2006) The dwarf phenotype of the *Arabidopsis acl5* mutant is suppressed by a mutation in an upstream ORF of a bHLH gene. *Development*, **133**, 3575–3585.
36. Rahmani,F., Hummel,M., Schuurmans,J., Wiese-Klinkenberg,A., Smeekens,S. and Hanson,J. (2009) Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. *Plant Physiol.*, **150**, 1356–1367.
37. Ebina,I., Takemoto-Tsutsumi,M., Watanabe,S., Koyama,H., Endo,Y., Kimata,K., Igarashi,T., Murakami,K., Kudo,R., Ohsumi,A. *et al.* (2015) Identification of novel *Arabidopsis thaliana* upstream open reading frames that control expression of the main coding sequences in a peptide sequence-dependent manner. *Nucleic Acids Res.*, **43**, 1562–1576.
38. Noh,A.L., Watanabe,S., Takahashi,H., Naito,S. and Onouchi,H. (2015) An upstream open reading frame represses expression of a tomato homologue of Arabidopsis *ANAC096*, a NAC domain

transcription factor gene, in a peptide sequence-dependent manner. *Plant Biotechnol.*, **32**, 157–163.
39. Menges,M. and Murray,J.A. (2002) Synchronous *Arabidopsis* suspension cultures for analysis of cell-cycle gene activity. *Plant J.*, **30**, 203–212.
40. Onouchi,H., Nagami,Y., Haraguchi,Y., Nakamoto,M., Nishimura,Y., Sakurai,R., Nagao,N., Kawasaki,D., Kadokura,Y. and Naito,S. (2005) Nascent peptide-mediated translation elongation arrest coupled with mRNA degradation in the *CGS1* gene of *Arabidopsis*. *Genes Dev.*, **19**, 1799–1810.
41. Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
42. Murota,K., Hagiwara-Komoda,Y., Komoda,K., Onouchi,H., Ishikawa,M. and Naito,S. (2011) Arabidopsis cell-free extract, ACE, a new in vitro translation system derived from Arabidopsis callus cultures. *Plant Cell Physiol.*, **52**, 1443–1453.
43. Erickson,A.H. and Blobel,G. (1983) Cell-free translation of messenger RNA in a wheat germ system. *Methods Enzymol.*, **96**, 38–50.
44. Cao,J. and Geballe,A.P., (1996) Inhibition of nascent-peptide release at translation termination. *Mol. Cell. Biol.*, **16**, 7109–7114.
45. Raney,A., Baron,A.C., Mize,G.J., Law,G.L. and Morris,D.R. (2000) *In vitro* translation of the upstream open reading frame in the mammalian mRNA encoding *S*-adenosylmethionine decarboxylase. *J. Biol. Chem.*, **275**, 24444–24450.
46. Nakatogawa,H. and Ito,K. (2001) Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Mol. Cell*, **7**, 185–192.
47. Gong,F., Ito,K., Nakamura,Y. and Yanofsky,C. (2001) The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-tRNA(Pro). *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 8997–9001.
48. Konevega,A.L., Fischer,N., Semenkov,Y.P., Stark,H., Wintermeyer,W. and Rodnina,M.V. (2007) Spontaneous reverse movement of mRNA-bound tRNA through the ribosome. *Nat. Struct. Mol. Biol.*, **14**, 318–324.
49. Schmeing,T.M., Huang,K.S., Strobel,S. and Steitz,T. (2005) An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature*, **438**, 520–524.
50. Wei,J., Wu,C. and Sachs,M.S. (2012) The arginine attenuator peptide interferes with the ribosome peptidyl transferase center. *Mol. Cell. Biol.*, **32**, 2396–2406.
51. Cruz-Vera,L.R., Gong,M. and Yanofsky,C. (2006) Changes produced by bound tryptophan in the ribosome peptidyl transferase center in response to TnaC, a nascent leader peptide. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 3598–3603.
52. Muto,H., Nakatogawa,H. and Ito,K. (2006) Genetically encoded but nonpolypeptide prolyl-tRNA functions in the A site for SecM-mediated ribosomal stall. *Mol. Cell*, **22**, 545–552.
53. Chiba,S. and Ito,K. (2012) Multisite ribosomal stalling: a unique mode of regulatory nascent chain action revealed for MifM. *Mol. Cell*, **47**, 1–10.
54. Yamashita,Y., Kadokura,Y., Sotta,N., Fujiwara,T., Takigawa,I., Satake,A., Onouchi,H. and Naito,S. (2014) Ribosomes in a stacked array: elucidation of the step intranslation elongation at which they are stalled during *S*-adenosyl-L-methionine-induced translation arrest of *CGS1* mRNA. *J. Biol. Chem.*, **289**, 12693–12704.
55. Hartz,D., McPheeters,D.S., Traut,R. and Gold,L. (1988) Extension inhibition analysis of translation initiation complexes. *Methods Enzymol.*, **164**, 419–425.
56. Tanaka,M., Sotta,N., Yamazumi,Y., Yamashita,Y., Miwa,K., Murota,K., Chiba,Y., Hirai,M.Y., Akiyama,T., Onouchi,H. *et al.* (2016) The minimum open reading frame, AUG-stop, induces boron-dependent ribosome stalling and mRNA degradation. *Plant Cell*, **28**, 2830–2849.
57. Anthony,D.D. and Merrick,W.C. (1992) Analysis of 40 S and 80 S complexes with mRNA as measured by sucrose density gradients and primer extension inhibition. *J. Biol. Chem.*, **267**, 1554–1562.
58. Sachs,M.S., Wang,Z., Gaba,A., Fang,P., Belk,J., Ganesan,R., Amrani,N. and Jacobson,A. (2002) Toeprint analysis of the positioning of translation apparatus components at initiation and termination codons of fungal mRNAs. *Methods*, **26**, 105–114
59. Ramu,H., Mankin,A. and Vazquez-Laslop,N. (2009) Programmed drug-dependent ribosome stalling. *Mol. Microbiol.*, **71**, 811–824.
60. Ramu,H., Vázquez-Laslop,N., Klepacki,D., Dai,Q., Piccirilli,J., Micura,R. and Mankin,A.S. (2011) Nascent peptide in the ribosome exit tunnel affects functional properties of the A-site of the peptidyl transferase center. *Mol. Cell*, **41**, 321–330.
61. Arenz,S., Bock,L.V., Graf,M., Innis,C.A., Beckmann,R., Grubmüller,H., Vaiana,A.C. and Wilson,D.N. (2016) A combined cryo-EM and molecular dynamics approach reveals the mechanism of ErmBL-mediated translation arrest. *Nat. Commun.*, **7**, 12026.
62. Malkin,L.I. and Rich,A. (1967) Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. *J. Mol. Biol.*, **26**, 329–346.
63. Blobel,G. and Sabatini,D.D. (1970) Controlled proteolysis of nascent polypeptides in rat liver cell fractions. I. Location of the polypeptides within ribosomes. *J. Cell Biol.*, **45**, 130–145.
64. Kawaguchi,R. and Bailey-Serres,J. (2005) mRNA sequence features that contribute to translational regulation in Arabidopsis. *Nucleic Acids Res.*, **33**, 955–965.
65. Sugio,T., Matsuura,H., Matsui,T., Matsunaga,M., Nosho,T., Kanaya,S., Shinmyo,A. and Kato,K. (2010) Effect of the sequence context of the AUG initiation codon on the rate of translation in dicotyledonous and monocotyledonous plant cells. *J. Biosci. Bioeng.*, **109**, 170–173.
66. Kozak,M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.
67. Kozak,M. (1997) Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *EMBO J.*, **16**, 2482–2492.
68. Kurian,L., Palanimurugan,R., Gödderz,D. and Dohmen,R.J. (2011) Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. *Nature*, **477**, 490–494.
69. Kumar,M., Busch,W., Birke,H., Kemmerling,B., Nürnberger,T. and Schöffl,F. (2009) Heat shock factors HsfB1 and HsfB2b are involved in the regulation of *Pdf1.2* expression and pathogen resistance in *Arabidopsis*. *Mol. Plant*, **2**, 152–165.
70. Ikeda,M., Mitsuda,N. and Ohme-Takagi,M. (2011) Arabidopsis HsfB1 and HsfB2b act as repressors of the expression of heat-inducible *Hsfs* but positively regulate the acquired thermotolerance. *Plant Physiol.*, **157**, 1243–1254.
71. Pick,T., Jaskiewicz,M., Peterhänsel,C. and Conrath,U. (2012) Heat shock factor HsfB1 primes gene transcription and systemic acquired resistance in Arabidopsis. *Plant Physiol.*, **159**, 52–55.
72. Pajerowska-Mukhtar,K.M., Wang,W., Tada,Y., Oka,N., Tucker,C.L., Fonseca,J.P. and Dong,X. (2012) The HSF-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. *Curr. Biol.*, **22**, 103–112.
73. Zhu,X., Thalor,S.K., Takahashi,Y., Berberich,T. and Kusano,T. (2012) An inhibitory effect of the sequence-conserved upstream open-reading frame on the translation of the main open-reading frame of HsfB1 transcripts in Arabidopsis. *Plant Cell Environ.*, **35**, 2014–2030.
74. Ren,H. and Gray,W.M. (2015) SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Mol. Plant*, **8**, 1153–1164.
75. Nemhauser,J.L., Hong,F. and Chory,J. (2006) Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. *Cell*, **126**, 467–475.