Local and Distant Sequences Are Required for Efficient Readthrough of the Barley Yellow Dwarf Virus PAV Coat Protein Gene Stop Codon

Chris M. Brown
Iowa State University

S. P. Dinesh-Kumar
Iowa State University

W. Allen Miller
Iowa State University, wamiller@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/plantpath_pubs
Part of the Agricultural Science Commons, Cell and Developmental Biology Commons, Molecular Biology Commons, and the Plant Pathology Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/plantpath_pubs/50. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

This Article is brought to you for free and open access by the Plant Pathology and Microbiology at Iowa State University Digital Repository. It has been accepted for inclusion in Plant Pathology and Microbiology Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Local and Distant Sequences Are Required for Efficient Readthrough of the Barley Yellow Dwarf Virus PAV Coat Protein Gene Stop Codon

Abstract
Many viruses use stop codon readthrough as a strategy to produce extended coat or replicase proteins. The stop codon of the barley yellow dwarf virus (PAV serotype) coat protein gene is read through at a low rate. This produces an extended polypeptide which becomes part of the virion. We have analyzed the cis-acting sequences in the barley yellow dwarf virus PAV genome required for this programmed readthrough in vitro in wheat germ extracts and reticulocyte lysates and in vivo in oat protoplasts. Two regions 3' to the stop codon were required. Deletion of sections containing the first 5 of the 16 CCN NNN repeats located 3' of the stop codon greatly reduced readthrough in vitro and in vivo. Surprisingly, readthrough also required a second, more distal element that is located 697 to 758 bases 3' of the stop codon within the readthrough open reading frame. This element also functioned in vivo in oat protoplasts when placed more than 2 kb from the coat protein gene stop in the untranslated region following a GUS reporter gene. This is the first report of a long-range readthrough signal in viruses.

Disciplines
Agricultural Science | Biochemistry, Biophysics, and Structural Biology | Cell and Developmental Biology | Molecular Biology | Plant Pathology

Comments
This article is from Journal of Virology 70 (1996): 5884. Posted with permission.

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/plantpath_pubs/50
Local and Distant Sequences Are Required for Efficient Readthrough of the Barley Yellow Dwarf Virus PAV Coat Protein Gene Stop Codon†

CHRIS M. BROWN, S. P. DINESH-KUMAR‡ AND W. ALLEN MILLER*

Departments of Plant Pathology and Biochemistry & Biophysics, and Molecular, Cellular and Developmental Biology Program, Iowa State University, Ames, Iowa 50011

Received 12 February 1996/Accepted 17 May 1996

Many viruses use stop codon readthrough as a strategy to produce extended coat or replicase proteins. The stop codon of the barley yellow dwarf virus (PAV serotype) coat protein gene is read through at a low rate. This produces an extended polypeptide which becomes part of the virion. We have analyzed the cis-acting sequences in the barley yellow dwarf virus PAV genome required for this programmed readthrough in vitro in wheat germ extracts and reticulocyte lysates and in vivo in oat protoplasts. Two regions 3′ to the stop codon were required. Deletion of sections containing the first 5 of the 16 CCN NNN repeats located 3′ of the stop codon greatly reduced readthrough in vitro and in vivo. Surprisingly, readthrough also required a second, more distal element that is located 697 to 758 bases 3′ of the stop codon within the readthrough open reading frame. This element also functioned in vivo in oat protoplasts when placed more than 2 kb from the coat protein gene stop in the untranslated region following a GUS reporter gene. This is the first report of a long-range readthrough signal in viruses.

Viruses in the luteovirus group utilize a variety of unusual translational control mechanisms to express their essential genes (reviewed in references 36 and 38). These include ribosomal frameshifting (7), stop codon readthrough (16), leaky scanning (17), and the use of a 3′ translational enhancer (64). Luteoviruses can be divided into two subgroups (34, 38), with barley yellow dwarf virus (BYDV) PAV serotype being the best characterized member of subgroup I. In all luteoviruses, there is a block of three open reading frames (ORFs) which are expressed from a single subgenomic RNA (sgRNA1) and make up the 3′ half of the genome. These encode the coat protein (CP), a protein of about 17 kDa (the 17K protein) that is nested within the coat protein sequence but in a different reading frame, and an ORF of about 50 kDa following the coat protein (RT) (39). RT is expressed by readthrough of the CP stop codon as a CP-RT fusion protein (16). The CP and a carboxy-terminally truncated form of CP-RT are the structural proteins of the virion (11, 18, 20, 63), with the RT domain located on the surface of the virion (46). The ratio of fusion protein to CP in purified virus preparations differs markedly between luteoviruses, serotypes of BYDV (20, 63), and individual virus preparations but has been estimated to be between 1:100 and 1:4 (2, 8, 11, 18, 20, 63). The CP-RT fusion protein is required for the aphid transmission of members of both subgroups of luteoviruses (8, 10) and the luteo-like enamovirus pea enation mosaic virus (PEMV) (13).

Stop codon readthrough is used as a regulatory strategy by a large number of plant, animal, and bacterial viruses (24, 67). The eukaryotic release factor complex normally decodes stop signals efficiently (37, 52, 53, 69). However, at certain stop codons, the competition between the eukaryotic release factor complex and suppressor tRNA for the stop signal shifts in favor of the tRNA. The tRNAs involved are likely to be naturally occurring isoacceptors with near-cognate anticodons (25, 57, 68). This phenomenon was first reported for an eukaryotic virus in tobacco mosaic virus (41) and has been subsequently observed or proposed for members of the alpha-, carmo-, enamo-, furo-, hordei-, luteo-, machlo-, necro-, tobamo-, tobra-, tombus-, tymo- and retroviruses (6, 24, 25, 38, 65, 67). Interference with the readthrough process in plant or human pathogenic viruses may also interfere with the viral life cycle.

In a few cases, the signals that permit or promote readthrough have been identified, but the mechanisms by which they act are not understood. The simplest signal is that of the mammalian alphaviruses (54). In vitro, a UGA C stop signal is sufficient to permit 10% readthrough, with the C providing an important determinant (33). In vivo, each of the three stop codons can be suppressed at this location in the Sindbis virus genome (32). Interestingly, it has been shown that termination is profoundly affected by the base following the stop codon (5, 37, 42). Stop codons followed by pyrimidine residues were poorly recognized as stop codons, with UGA C being one of the weakest signals (37). Thus, the combination of this weak stop signal and the presence of a UGA-suppressing tRNA might permit readthrough of the alphavirus stop (56). For tobacco mosaic virus RNA, a larger signal, UAG CAR YYA, is necessary and sufficient for 5% readthrough in vivo (49, 50, 58). The wild-type signal, UAG CAA UUA, also promotes 21% readthrough in yeast cells and 2% readthrough in mammalian cells (51) but does not function in Escherichia coli (50).

A more elaborate signal is required for UAG readthrough by murine leukemia virus, including a pseudoknot spaced 8 nucleotides from the UAG stop, as well as specific sequences within the spacer and loops (19, 65).

A large number of luteoviral structural gene sequences are available because of the economic importance of this group of viruses.
viruses. In all known luteovirus sequences, the CP stop codon is flanked by the sequence CN AAA UAG GUA GAC (Fig. 1). This conserved block is entirely different from all known readthrough contexts. After a short spacer of 6 to 15 bases, this sequence is followed by a C-rich block containing 7 to 16 tandem repeats of CCN NNN (Fig. 1; also see Fig. 6). The length of the spacer puts the CCN of these repeats in frame, encoding prolines. This region, in which every second amino acid is proline, has been proposed to form a hinge between the CP and the remainder of the RT domain (21). In all luteoviruses, at least one of the CCN NNN repeats contains the sequence CCCCA (38). No conserved secondary structures have been identified within this region. Here, we determine which of these elements following the stop are required for readthrough in vitro and in vivo. Two regions, part of the C-rich block and an additional element located some distance from the CP stop codon, were found to be required for readthrough.

**MATERIALS AND METHODS**

Full-length infectious clones. All plasmids were constructed by standard techniques (47) and site-specific mutagenesis (30). Plasmid pPAVGUSRT1 is derived from a full-length BYDV PAV infectious clone pPAV6 (14) and contains the uidA reporter gene (without its initiation codon), which encodes β-glucuronidase (GUS). The Apal (in the N-terminal GUS coding region)-HpaI (following the GUS stop codon, blunt) fragment of pAGUS (49) was used to replace bases 3593 to 3787 (HpaI) of pPAV6 to generate pPAVGUSRT1 (see Fig. 4). In pPAVGUSRT2, the fragment was placed after nucleotide 3477, and in pPAVGUSRT15, it was placed after nucleotide 3534. In pPAVGUSRT3, pPAVGUSRT4, and pPAVGUSRT21, GAG sense codons replace the TGA stop codon of pPAVGUSRT1, pPAVGUSRT2, pPAVGUSRT20, respectively. Plasmids pPAVGUSRT9 and pPAVGUSRT10 contain deletions, in the coat protein ORF, of bases 2985 to 3345 and 2985 to 3423, respectively. pPAVGUS RT0 is a derivative of pPAVGUSRT15 with a T inserted after 3475 and the final T (base 3534) after the Apal site deleted to maintain the frame. pPAVGUS RT6, pPAVGUSRTH, pPAVGUSRTNA, and pPAVGUSRTAS are derivatives of pPAVGUSRT1 with deletions of 3788 (HpaI blunt) to 4515 (Scal blunt), 3788 (HpaI blunt) to 4120 (Ndel filled), 4123 (Ndel filled) to 4154 (NcoI filled), and 4155 (NcoI filled) to 4515 (Scal blunt), respectively.

**Plasmids derived from full subgenomic RNA1.** The parent vector pCB18 contains the complete sgRNA1 sequence adjacent to a T promoter. Maps are shown in Fig. 1. A T7 promoter was added to the subgenomic RNA start site (italics) by PCR with primer SG1 (GGTCTAGATAACGACTCATCATAG TGAAGGTTAGACAGCTCACCATG) and a 3′ primer. An XbaI-SalI fragment of this product was ligated into XbaI-SalI-cut pBS18 (16). In pGAG, the CP stop codon, UAG, was altered to GAG. Templates for in vitro transcription of sRNAs in which the start codons of the CP or the 17K protein had been altered to AGC, pCP, and p17K, were generated by PCR with SG1 and a primer complementary to the 3′ end of the genome (5′-GGGTCGCGCAACGTCCTCTTTC-3′) from full-length genomic clones harboring these mutations, PAV31 and PAV33 (40). pSTU contains UAG GACCTTG in place of the UAG GUA GAC at the coat protein stop, adding a SalI site (italics). The SalI-HpaI fragment of pSTU was deleted to give pSH. The BspMI (3-base overhang, filled)-HpaI fragment of pCB18 was deleted to give pBHC. Two pairs of Cs (3486 to 3489, 3494 to 3495) in pBHC were altered to Ts in pBHT. The StuI-BspMI (3-base overhang, filled) fragment of pSTU was deleted in pSB. The HpaI-SalI fragment was deleted in pHS. The SalI-BspMI fragments of pPAVGUSRT20 and pPAVGUSRT21 were subcloned into pCB18 to generate pFT and pFG, respectively. pRNA and pFAS were derived by subcloning BglII-Pol fragments from pPAVGUSRTNA and pPAVGUSRTAS into pFT. pFT was cut with HpaI-Ndel and filled with Klenow DNA polymerase, which fortuitously produced a mutant containing a slightly larger than expected deletion (3788 to 4120), pBHN. pFG contains a GAG sense codon in place of the TAG stop. A template for transcription truncated after position 4219 was generated by PCR with primers SG1 and 4219 (TTCAAGGCTGCTCGTTATGTC). All plasmids were sequenced in relevant regions with an Applied Biosystems model 377 automated sequencer at the Iowa State University Nucleic Acids Facility.

**In vitro transcription.** Transcripts were synthesized from linearized plasmid templates or PCR products with T7 RNA polymerase by the method described by Promega (43). Capped transcripts were synthesized with the ImageMethyl kits (Ambion). The RNA concentration was determined with a GeneQuant spectrophotometer and by ethidium bromide staining following electrophoresis.

**In vitro translation.** Translation in wheat germ extracts was done essentially as recommended by the manufacturer (Ambion). RNAs (usually 0.4 µg) were translated in 20-µl reaction mixtures containing 150 mM potassium acetate, 2.5 mM magnesium acetate, 15 U of RNasin (Promega), 1 µl of master mix minus methionine, 10 µl of wheat germ extract, and [35S]methionine (New England Nuclear). Translation products (1 µl) were separated on 6% stacking–15% resolving polyacrylamide gels as described previously (29), except that the concentration of resolving gel buffer was used (0.75 M). Three sets of M1 markers were used (Bio-Rad, Bethesda Research Laboratories, and Amersham); they were separated on a polyacrylamide gel, and the M1 was determined for each marker from each set. The mean of these three determinations was used as the M1 indicated (Fig. 2). Gels were fixed for 6 h before being processed for fluorography. Autoradiographs were analyzed by densitometry using ImageQuant software (Molecular Dynamics). Readthrough percentages given in the text represent the mean of three or more independent experiments. Adjustments were made for the relative number of methionines in each protein (17K protein, 3; CP, 4). Translation in reticulocyte lysates was done as described previously (16).

**Protoplast transformation and analysis.** oat protoplasts were isolated, transfected, and analyzed as described previously (17, 40). They were electroporated with 20 µg of full-length transcript and then incubated in the dark for 24 h before being harvested for analysis. Total GUS activity (picomoles of 4-methylumbellifere (MU) per minute per milligram of protein) was determined as described previously (26). GUS activity was detected following collection of total protein, separation on a semidenaturing 7.5% polyacrylamide gel, and autoradiography (Fig. 2). Protein concentrations were determined with Bradford protein assay kits (Bio-Rad). RNA analysis was performed as described previously (40). Total RNA was extracted by a procedure in which ariuchiralcarboxylic acid was used as an RNAse inhibitor (62), separated on 1% denaturing gels, transferred to a nylon membrane, and probed with an RNA probe complementary to the 3′ end of the BYDV PAV genome (from pSP10 [17]). Double-antibody sandwich enzyme-linked immunosorbant assays (ELISA) were done as described previously (22).

**RESULTS**

Translation in vitro. For these studies, uncapped BYDV PAV subgenomic RNA1 derivatives were transcribed and translated in vitro. Different 5′ ends of sgRNA1 had been reported for the Illinois (nucleotide 2769) and Australian (nucleotide 2670) isolates of BYDV PAV (16, 28). The 5′ end
The sizes of the CP-RT products are 72 kDa (lanes 1 to 5), 23 kDa (lane 18), 71 kDa (lane 9), 35 kDa (lane 10), and 25 kDa (lanes 11 and 12). The sizes of the M_r markers are indicated on the right.

### A second distal element.

Surprisingly, deletion of a more distal region (3788 to 4515), nearly 700 bases from the stop, did not affect the amount of readthrough (pFG). These data indicate that a portion of the C-rich element is required for readthrough in vitro.

To investigate whether the CCN NNN repeats needed to be in frame to function, a single U was inserted after base 3475, 15 nucleotides after the stop (pFT). This places the repeats in a different reading frame (NCC NNN). A product of the expected size (M_r 24,821) was visible (Fig. 2A, lane 11). Lane 12 contained a derivative of pFT in which the UAG was changed to GAG (pFG). These data indicate that the C-rich region does not need to encode prolines or be in a particular reading frame to function.

BYDV PAV readthrough had been noted previously (16), and leaky scanning has been characterized (17) in rabbit reticulocyte lysates. For comparison, the mutant sgRNA1 transcripts were also translated in these lysates (Fig. 2B). Translation was only one-quarter as efficient on rabbit ribosomes. Translation of sgRNA1 gave the same three products, but their relative amounts differed (Fig. 2B, lane 1). There were also additional products, which are likely to result from initiation at AUGs within the RT ORF (also noted in reference 16). In reticulocyte lysate, the ratio of CP to 17K protein was reversed and the percent readthrough (at about 10%) was higher. Deletions and alterations in the sgRNA had similar effects on readthrough in reticulocyte lysates. Deletion of the C-rich element (lanes 3 and 6) or an alteration to it (lane 5 compared with lane 4) also reduced readthrough in reticulocyte lysates, and changing the conserved bases after the stop had no effect (lane 2).

### Sequences required for readthrough.

The 6-base sequence following the stop is the same in all luteoviruses (GUA GAC). To test if it is required for readthrough, we altered five of the bases to create GCC UUG (this introduces a unique StuI site). Readthrough was not reduced; indeed, a modest increase in readthrough was observed (Fig. 2A, lane 5).

We made deletions and alterations in the C-rich region following the stop to test if these sequences were needed for readthrough. Bases 3478 to 3787 (numbered as in reference 39), which include 14 of the 16 CCN NNN repeats, were deleted (pSH). This alters the reading frame, giving an expected RT product with an M_r of only 22,902. No RT product was detected from this construct. However, a smaller deletion (3500 to 3787) did not affect the amount of readthrough (pBHC). This construct encoded a slightly smaller than wild-type CP-RT product that maintains the wild-type frame (M_r 61,466 [lane 7]). It retained five CC pairs, four of which are evenly spaced as CCN NNN repeats. Alteration of the last two of these CC doublets in this construct to UU caused a significant reduction in readthrough (to 40% of the wild-type control; pBHT [lane 8]). The effect of deletion of a region including these five CC pairs (3462 to 3499) was also tested. This placed the remaining 12 pairs closer to the stop and maintained the reading frame (pSB, M_r 70,641 [lane 9]). This alteration also reduced the amount of readthrough product (to 10%). These data indicate that a portion of the C-rich element is required for readthrough in vitro.

To investigate whether the CCN NNN repeats needed to be in frame to function, a single U was inserted after base 3475, 15 nucleotides after the stop (pFT). This places the repeats in a different reading frame (NCC NNN). A product of the expected size (M_r 24,821) was visible (Fig. 2A, lane 11). Lane 12 contained a derivative of pFT in which the UAG was changed to GAG (pFG). These data indicate that the C-rich region does not need to encode prolines or be in a particular reading frame to function.
Three internal deletions were made spanning the distal region (3788 to 4515). A deletion (Fig. 3A, 4155 to 4515; pFTAS) that encompassed the 3' border of the distal element (4219) reduced readthrough to less than 10% of wild-type levels (Fig. 3C, lanes 3 and 4). A 32-base central deletion (4123 to 4154; pFTNA) had no effect (Fig. 3C, lanes 5 and 6), whereas a larger 5’ deletion (3788 to 4129; pFTHN) reduced readthrough modestly (to 65% of wild-type levels [lanes 7 and 8]). This indicates that a distal RNA element located between nucleotides 4155 and 4219 (697 to 758 nucleotides from the CP stop) is required for readthrough. Furthermore, this element can function when located within or 3' of the readthrough ORF.

**Replicating viral constructs in vivo.** To determine whether the same readthrough signals mapped above are required in vivo, we inserted a modified GUS reporter gene downstream of the CP stop codon in a BYDV PAV full-length infectious clone (14). The GUS ORF lacking the start codon (1894 bases) was inserted in place of bases 3593 to 3787 of the RT ORF (194 bases) (Fig. 4A). This region is dispensable for efficient readthrough in vitro (Fig. 2, pBHC). This places the distal readthrough element in the noncoding region following the GUS reporter gene, 2.4 rather than 0.7 kb from the stop codon. This also places the 3' translation enhancer 1.7 kb further away (64). The resulting hybrid genome of 7.4 kb is 30% larger than the 5.7-kb BYDV PAV genome. Expression of GUS in these constructs requires viral replication for synthesis of its mRNA, sgRNA1. The GUS reporter gene can be expressed only by readthrough of the CP stop codon.

Full-length genomic RNAs were transcribed in vitro and electroporated into oat protoplasts. Accumulation of viral RNAs was measured by Northern (RNA) blot analysis (Fig. 4B), accumulation of virus particles was measured by ELISA (Fig. 4B), and the relative amount of readthrough was measured by monitoring GUS activity (Fig. 4C). Hybrid viral RNA containing the GUS reporter gene without other alterations replicated (pPAVGUSRT1 [Fig. 4B, lane 3]), but to lower levels than that of the full-length infectious clone (pPAV6 [Fig. 4B, lane 6]). The probe was complementary to the 3' end of the genome and detected the full-length genomic RNA and sgRNA1 at the expected greater lengths (Fig. 4B, 7.4 and 4.7 kb [lane 3]; cf. 5.7 and 3.0 kb [lane 6]). Viral antigen (coat protein) accumulated to half the control levels (Fig. 4B, lane 3 compared with 6). These data indicate that the larger hybrid genome can replicate in oat protoplasts, although not as efficiently as the wild-type genome. This also verifies previous observations that the RT ORF is not required for BYDV PAV RNA replication in protoplasts (20, 40).

Readthrough of the CP stop was detected by assaying for GUS activity. An extract from protoplasts infected with the reporter construct gave 89 GUS units versus 1 unit in mock-infected cells (Fig. 4C; lane 5 compared with lane 3). To ensure that the GUS activity was not due to internal initiation or translation of fragmented RNA, the size of the protein giving GUS activity was determined following separation of infected protoplast proteins on a semidenaturing polyacrylamide gel (Fig. 4C). GUS activity was found at a mobility corresponding to the expected size of the full length CP-GUS fusion protein (95 kDa [Fig. 4C, lane 5]). These data indicate that readthrough occurred in vivo in infected protoplasts.

To determine if the C-rich element was required in vivo, a construct lacking all but two of the CC doublets was tested (pPAVGUSRT2 [Fig. 4A]). It accumulated viral antigen and RNAs to levels similar to that of the construct containing the C-rich element (Fig. 4B, lane 2 compared with lane 3). However, GUS activity was only 9% of the control (Fig. 4C, lane 2)
FIG. 4. Replication and GUS expression of full-length infectious reporter constructs in vivo. (A) Schematic diagram of the pPAVGUSRT series of constructs. The genome organization of pPAVGUSRT1 and pPAVGUSRT2 is shown with the region following the CP stop expanded to show sites of GUS expression. The 3′ end of the RT ORF following the GUS ORF is indicated as a dashed box labeled RT. (B) Northern blot hybridization of total RNA from cells transfected with reporter constructs. Cells were electrophoresed with salmon sperm DNA (lane 1) or with T7 polymerase transcripts from the following Smal linearized plasmids: pPAVGUSRT2 (lane 2), pPAVGUSRT1 (lane 3), pPAVGUSRT4 (lane 4), pPAVGUSRT3 (lane 5), and pPAV6 (lane 6). The presence (+) or absence (−) of the C-rich region in these constructs is indicated (C-rich). The position of the CP UAG codon (+) or replacement with GAG (−) is indicated on the next line (STOP). The accumulation of viral antigen was measured in duplicate samples by ELISA (A500). (C) Total protein from protoplasts transfected with transcripts from the following plasmids was separated by polyacrylamide gel electrophoresis and stained for GUS activity (7). The presence of the region following the CP stop expanded to show sites of GUS expression. The 3′ end of the RT ORF following the GUS ORF is indicated as a dashed box labeled RT. (B) Northern blot hybridization of total RNA from cells transfected with reporter constructs. Cells were electrophoresed with salmon sperm DNA (lane 1) or with T7 polymerase transcripts from the following Smal linearized plasmids: pPAVGUSRT2 (lane 2), pPAVGUSRT1 (lane 3), pPAVGUSRT4 (lane 4), pPAVGUSRT3 (lane 5), and pPAV6 (lane 6). The presence (+) or absence (−) of the C-rich region in these constructs is indicated (C-rich). The position of the CP UAG codon (+) or replacement with GAG (−) is indicated on the next line (STOP). The accumulation of viral antigen was measured in duplicate samples by ELISA (A500). (C) Total protein from protoplasts transfected with transcripts from the following plasmids was separated by polyacrylamide gel electrophoresis and stained for GUS activity (7): pPAVGUSRT4 (lane 1), pPAVGUSRT3 (lane 2), salmon sperm DNA (lane 3), pPAVGUSRT2 (lane 4), pPAVGUSRT1 (lane 5), and pPAUS1 (lane 6). The GUS activity (picomoles of MU per minute per milligram of protein) in a lysate of duplicate infected samples is also shown. Fifteen times as much lysate was loaded in lanes 2 and 5 as indicated (15%). The positions of the GUS protein (GUS) and different CP-RT-GUS fusion proteins are indicated on the left. The expected mobilities are 90 kDa in lanes 1 and 2, 95 kDa in lanes 4 and 5, and 73 kDa in lane 6.

compared with lane 5), indicating that the C-rich element was required for readthrough in vivo. The expected size of the product was 90 kDa.

We constructed control in-frame derivatives of these two constructs in which the UAG stop was changed to a GAG sense codon (pPAVGUSRT3 and pPAVGUSRT4). These control viral RNAs should produce no free CP, only a CP-RT-GUS fusion. Surprisingly, these alterations reduced accumulation of viral full-length and subgenomic RNAs dramatically (Fig. 4B, lanes 4 and 5). No intact virions were detected by ELISA (Fig. 4B, lanes 4 and 5) which migrated at the expected Mv, indicating the presence of CP-RT-GUS protein (Fig. 4C, lanes 1 and 4). Because the total GUS activity was lower in constructs with the stop, 15 times as much sample was loaded to give approximately the same activity in lanes 4 and 5 (Fig. 4C, lanes 2 and 5). Because the GAG constructs did not replicate as well as UAG constructs, we could not calculate the absolute rate of readthrough. However, we can compare relative readthrough activities of constructs with the CP stop codon.

Other deletions and alterations were tested in replicating viral RNAs (Fig. 5). The GUS activities measured in different experiments are shown, with the average value presented as a percentage of the wild-type value in the final column. Deletions of most of the coat protein did not affect readthrough significantly (Fig. 5A). Deletion of the entire C-rich element reduced readthrough, as indicated above (Fig. 4), but the first 10 CCN NNN repeats were sufficient (Fig. 5B). Placing the C-rich region in the +1 frame so that it no longer encoded prolines did not affect the readthrough rate (Fig. 5C), as observed in vitro (Fig. 2).

Deletion of a large section containing the distal element identified in vitro reduced readthrough 10-fold (Fig. 5D, pPAVGUSRT6). Viral antigen accumulated to 76% of wild-type levels (data not shown), indicating that replication and virion formation were occurring. Effects of smaller deletions (Fig. 5D) were similar to those seen in vitro. The greatest reduction was caused by a deletion of nucleotides 4155 to 4515 (to 7%), whereas lesser effects were seen with deletions from 4123 to 4154 (to 61%) and 3788 to 4129 (to 42%). Deletion of both the C-rich and distal elements reduced readthrough 10-fold (8%, Fig. 5E). These results indicate that this distal element detected in vitro is also required in vivo and that key determinants lie between bases 4154 and 4515.

**DISCUSSION**

The C-rich element. The conserved block around the luteoviral stop codons (CN AAA UAG GUA GAC, Fig. 6) is not well conserved in PEMV (13) and was not sufficient for readthrough in vivo (Fig. 4). A C-rich region following the CP stop codon was required for readthrough. All the luteovirus and PEMV sequences contain multiple CCN NNN repeats located 6 to 15 bases after the CP stop codon, but there is little homology other than the CC doublets (Fig. 6). For BYDV PAV, the sequence containing the first 5 CCN NNN repeats was sufficient, but the sequence containing the next 11 repeats would not functionally replace it. Also, although mutation of the fourth and fifth CC doublets to UU weakened the element, they did not completely abolish its activity. These data may indicate a requirement for optimal spacing from the stop or for sequences between the CC doublets. Because the element functioned when placed in a different reading frame, the signal is in the RNA itself rather than in the proline-rich protein it encodes.

It had been noted that for luteoviruses, at least one of the repeats contained CCCCA (38). This sequence is also located at variable distances after the suppressible stops of other plant viruses (38). This was not sufficient for readthrough. Deletion of regions containing the CCCCA had no effect on readthrough. Deletion of the natural C-rich region, which fortuitously placed the same sequence (CCCCCA) within the reporter gene (GUS) at the same distance from the stop codon, resulted in a mutant that had no read-through activity (pPAVGUSRT2 [Fig. 4 and 5]).

The distal readthrough element. A region normally located nearly 700 bases 3′ of the stop was also required for readthrough. The unexpected distal element is more distant than any readthrough signal hitherto reported and functioned over a wide range of distances. It functioned in vitro when located...
0.4 to 0.7 kb after the stop in the 3′ untranslated region (3′ UTR) of a truncated RT protein (Fig. 3) and in vivo when located in the 3′ UTR 2.4 kb from the CP stop codon (Fig. 5).

The 5′ portion of the distal element (4154 to 4219) is well conserved in all luteoviruses and PEMV (Fig. 7), as is the amino acid sequence (13, 36). It is possible that the distal element interacts with the C-rich element by long-distance base pairing. Downstream secondary structures are needed for other types of readthrough (65). There are two conserved GG doublets in the 5′ region of the motif which could potentially base pair with CC doublets in the C-rich region. There is also potential base-pairing between bases 4184 to 4203 in the distal element and 3466 to 3485 in the C-rich element (bases indicated in Fig. 6 and 7). This pairing is conserved in the PAV and MAV strains of BYDV but not in SGV strains or other luteoviruses.

The CP-RT protein is required for aphid transmissibility (8, 10), so that alterations to the elements identified in this study could reduce the amount of RT product and aphid transmissibility. For example, deletion of the Hpa-I-Sca-I segment (which contains the distal element) from pPAV6 gave no detectable truncated CP-RT protein from Western blot (immunoblot) analysis of infected protoplasts and no aphid transmission (10). We looked for a correlation between the sequence of the homologous elements from seven potato leafroll virus (PLRV) isolates and aphid transmissibility (27) but found none. They differed in one base in the C-rich element (underlined in Fig. 6) and three in the distal element (underlined in Fig. 6) and three in the distal element (underlined...
in Fig. 7). They also differed in protein sequences in several places in the RT protein, and this may account for the alterations in transmissibility as proposed previously (27). For PEMV, the sequence of a strain that does not read through differs at two positions (13) in the distal-element homolog (underlined in Fig. 7). These changes do not alter amino acid sequence, but might abolish readthrough.

One example is known in which decoding of a stop codon is affected by the sequence located kilobases downstream in the 3′ untranslated region (3′ UTR) (4). This is the incorporation of selenocysteine (SeCys) at UGA “sense” codons in specific mammalian cellular genes. This might be considered a special case of readthrough (23). A large bulged stem-loop located in the 3′ UTR directs SeCys incorporation at UGA codons within mammalian mRNAs (3, 48). This structure is thought to specifically bind a complex of selenocysteiny1-tRNA and a special elongation factor, channeling this complex into the decoding site on the ribosome. This mechanism is specific for SeCys and UGA codons, but an analogous mechanism might be possible for readthrough. However, computer analysis did not detect any large conserved stem-loops in the distal readthrough element of the BYDV-PAV genome.

Translation in vitro in wheat germ lysates. We used the wheat germ translation system because wheat is a natural host of the virus and some plant translational control signals function only in plants. The predominance of CP over the 17K protein and this may account for the alterations in transmissibility as proposed previously (27). For PEMV, the sequence of a strain that does not read through differs at two positions (13) in the distal-element homolog (underlined in Fig. 7). These changes do not alter amino acid sequence, but might abolish readthrough.

One example is known in which decoding of a stop codon is affected by the sequence located kilobases downstream in the 3′ untranslated region (3′ UTR) (4). This is the incorporation of selenocysteine (SeCys) at UGA “sense” codons in specific mammalian cellular genes. This might be considered a special case of readthrough (23). A large bulged stem-loop located in the 3′ UTR directs SeCys incorporation at UGA codons within mammalian mRNAs (3, 48). This structure is thought to specifically bind a complex of selenocysteiny1-tRNA and a special elongation factor, channeling this complex into the decoding site on the ribosome. This mechanism is specific for SeCys and UGA codons, but an analogous mechanism might be possible for readthrough. However, computer analysis did not detect any large conserved stem-loops in the distal readthrough element of the BYDV-PAV genome.

Translation in vitro in wheat germ lysates. We used the wheat germ translation system because wheat is a natural host of the virus and some plant translational control signals function only in plants. The predominance of CP over the 17K protein and this may account for the alterations in transmissibility as proposed previously (27). For PEMV, the sequence of a strain that does not read through differs at two positions (13) in the distal-element homolog (underlined in Fig. 7). These changes do not alter amino acid sequence, but might abolish readthrough.

One example is known in which decoding of a stop codon is affected by the sequence located kilobases downstream in the 3′ untranslated region (3′ UTR) (4). This is the incorporation of selenocysteine (SeCys) at UGA “sense” codons in specific mammalian cellular genes. This might be considered a special case of readthrough (23). A large bulged stem-loop located in the 3′ UTR directs SeCys incorporation at UGA codons within mammalian mRNAs (3, 48). This structure is thought to specifically bind a complex of selenocysteiny1-tRNA and a special elongation factor, channeling this complex into the decoding site on the ribosome. This mechanism is specific for SeCys and UGA codons, but an analogous mechanism might be possible for readthrough. However, computer analysis did not detect any large conserved stem-loops in the distal readthrough element of the BYDV-PAV genome.

Translation in vitro in wheat germ lysates. We used the wheat germ translation system because wheat is a natural host of the virus and some plant translational control signals function only in plants. The predominance of CP over the 17K protein and this may account for the alterations in transmissibility as proposed previously (27). For PEMV, the sequence of a strain that does not read through differs at two positions (13) in the distal-element homolog (underlined in Fig. 7). These changes do not alter amino acid sequence, but might abolish readthrough.

One example is known in which decoding of a stop codon is affected by the sequence located kilobases downstream in the 3′ untranslated region (3′ UTR) (4). This is the incorporation of selenocysteine (SeCys) at UGA “sense” codons in specific mammalian cellular genes. This might be considered a special case of readthrough (23). A large bulged stem-loop located in the 3′ UTR directs SeCys incorporation at UGA codons within mammalian mRNAs (3, 48). This structure is thought to specifically bind a complex of selenocysteiny1-tRNA and a special elongation factor, channeling this complex into the decoding site on the ribosome. This mechanism is specific for SeCys and UGA codons, but an analogous mechanism might be possible for readthrough. However, computer analysis did not detect any large conserved stem-loops in the distal readthrough element of the BYDV-PAV genome.

The percent readthrough in phloem cells, in which virus replication occurs, has not been determined accurately. For the subgroup II luteovirus PLRV, the readthrough efficiency was estimated to be 0.9 to 1.3% in vivo in tobacco and potato protoplasts. The DNA reporter constructs contained only 18 bases before and 21 bases after the CP stop (55), so that these constructs lacked both elements we identified in this report. By using DNA reporter constructs similar to those of Tacke et al., <0.2% readthrough of BYDV PAV CP stop codon was detected in oat protoplasts (15).

Replication of chimeric viruses in protoplasts. We found that hybrid BYDV PAV viruses containing the GUS gene were able to replicate and accumulate viruses in oat protoplasts. Deletion of most of the RT protein does not affect replication (10, 40). Unexpectedly, replication of full-length infectious transcripts containing the GAG sense codons in place of the UAG stop was barely detectable by Northern blot analysis or ELISA. These clones were constructed to make none of the normal CP of the virophage, only the CP-RT–GUS fusion. The construct with the C-rich deletion would have only four RT-encoded amino acids and so would be essentially a CP-GUS fusion. Similarly, it was shown recently that a mutant BYDV PAV construct which contained a UAC sense codon in place of the CP stop also did not accumulate viral RNAs or antigen (20). This could be interpreted to indicate that free CP is required for virus replication. However, mutants that do not express CP accumulate significant levels of RNA (40), albeit at a reduced level, unlike the stop to sense codon mutants. Thus, 100% fusion of CP to RT reduced RNA accumulation more than the lack of CP alone. However, some replication must be occurring, because GUS activity was detected.

In vivo and in both in vitro translation systems, both readthrough elements were required, indicating that they are biologically relevant and do not function exclusively in plants. We have identified them as two equally essential elements; however, they may function in concert, possibly linked directly by base pairing or through an additional factor. We are currently identifying more precisely the key features of these elements.

ACKNOWLEDGMENTS

C.M.B. is a long-term fellow of the Human Frontier Science Program Organisation. This study was supported by Hatch Act and State of Iowa funds. S.P.D.K. was supported by USDA National Research Initiative grant 91-37303-6424 to W.A.M.

REFERENCES


64. Wang, S., and W. A. Miller. 1995. A sequence located 4.5 to 5 kilobases from the 5’ end of the barley yellow dwarf virus (PAV) genome strongly stimulates translation of uncapped mRNA. J. Biol. Chem. 270:13446–13452.