

The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related

Eric J. Snijder^{*+}, Johan A. den Boon⁺, Peter J. Bredenbeek, Marian C. Horzinek, René Rijnbrand and Willy J. M. Spaan⁺

Institute of Virology, Veterinary Faculty, State University of Utrecht, Yalelaan 1, 3584 CL Utrecht, The Netherlands

Received April 3, 1990; Revised and Accepted June 28, 1990

EMBL accession no. X52374

ABSTRACT

Sequence analysis of the 3' part (8 kb) of the polymerase gene of the torovirus prototype Berne virus (BEV) revealed that this area contains at least two open reading frames (provisionally designated ORF1a and ORF1b) which overlap by 12 nucleotides. The complete sequence of ORF1b (6873 nucleotides) was determined. Like the coronaviruses, BEV was shown to express its ORF1b by ribosomal frameshifting during translation of the genomic RNA. The predicted tertiary RNA structure (a pseudoknot) in the toro- and coronaviral frameshift-directing region is similar. Analysis of the amino acid sequence of the predicted BEV ORF1b translation product revealed homology with the ORF1b product of coronaviruses. Four conserved domains were identified: the putative polymerase domain, an area containing conserved cysteine and histidine residues, a putative helicase motif, and a domain which seems to be unique for toro- and coronaviruses. The data on the 3' part of the polymerase gene of BEV supplement previously observed similarities between toro- and coronaviruses at the level of genome organization and expression. The two virus families are more closely related to each other than to other families of positive-stranded RNA viruses.

INTRODUCTION

Berne virus (BEV) is the prototype of the proposed family Toroviridae, a recently characterized group of enveloped, positive-stranded RNA viruses (1) which have been found in association with enteric infections in horses, cattle and man (2).

The BEV genome consists of a single, infectious, polyadenylated RNA molecule (3). From U.V. transcription

mapping data its size was calculated to be between 25 and 30 kilobases (kb; 4). The genome organization and expression of BEV are comparable to those of the coronaviruses (for a review see 5). Sequence analysis has revealed at least five open reading frames (ORFs) on the BEV genome (4). These ORFs are expressed from a nested set of mRNAs, each containing one of the ORFs at their 5' end. The BEV mRNAs are 3' coterminal, but, in contrast to the coronaviral mRNAs, no common leader sequence has been detected at their 5' end (4).

In addition to similarities at the level of genome organization and expression, the presence of amino acid sequence homologies in replicase proteins is considered indicative for ancestral relationships between groups of viruses (6, 7). Recently, we have demonstrated that the large ORF which is located upstream of the BEV peplomer protein gene is only present on genomic RNA. Therefore this ORF most likely encodes the viral RNA-dependent RNA polymerase (POL) or replicase (4). Since the termination codon of this ORF maps at 6603 nucleotides (nt) upstream of the poly(A) tail, the complete replicase encoding region of BEV is extremely large (18 to 23 kb). Among positive-stranded RNA viruses, its size is comparable only to the POL gene of the coronaviruses (8, 9).

The coronaviral POL gene has been shown to contain two ORFs (8) of about 12 kb (ORF1a, formerly named F1) and 8 kb (ORF1b, formerly named F2). The 3' end of ORF1a overlaps with the 5' end of ORF1b, which is expressed by ribosomal frameshifting. This process involves RNA sequence elements in the ORF1a/ORF1b overlap region, including an RNA pseudoknot (10, 11, 12).

In this paper the sequence analysis of a substantial part of the BEV polymerase gene is reported. Both the primary structure of the BEV polymerase and the mechanism used to express the 3' part of the POL gene indicate that toro- and coronaviruses are indeed evolutionarily related.

* To whom correspondence should be addressed

⁺ Present address: Institute of Medical Microbiology, Department of Virology, State University of Leiden, Postbus 320, 2300 AH Leiden, The Netherlands

MATERIALS & METHODS

cDNA synthesis and cloning

The preparation of a randomly primed cDNA library from intracellular poly(A)-selected BEV RNA and the selection of BEV-specific cDNA clones have been described previously (4).

Sequence analysis and computer analysis of sequence data

Nucleotide sequence analysis and assembly of sequence data were performed as described previously (4). Amino acid sequence similarity searches were carried out using the FASTA program (13) and the NBRF protein identification resource (release 20.0). Dot matrix comparisons, sequence alignments and RNA secondary structure predictions were generated using the COMPARE, GAP and FOLD options from the software provided by the Computer Genetics Group/University of Wisconsin (version 5, 1989; 14).

Construction of pBSFS

To study ribosomal frameshifting in the ORF1a/ORF1b overlapping region the expression vector pBSFS was constructed. The unique EcoRI site of plasmid pBS (Stratagene) was digested and sticky ends were filled using the Klenow fragment of DNA polymerase I (15). An XhoI linker (5' CCTCGAGG 3') was inserted into the filled EcoRI site, resulting in pBSXho. An XhoI DNA fragment containing the matrix protein (M) gene of mouse hepatitis virus strain A59 (MHV-A59) (a gift from Peter Rottier) was cloned into the XhoI site of pBSXho to generate construct pBSM. pBSM was digested with SmaI and SalI, sticky ends were filled as described above, and the vector was religated. In this way the BamHI site in the multiple cloning region of pBSM was deleted. Subsequently, BglII (5' CAGATCTG 3') and BamHI (5' CGCGGATCCGCG 3') linkers were inserted into the filled AccI and StyI sites, respectively, of the MHV-A59 M gene in pBSM (nt positions 656 and 693 in the sequence as presented in Fig. 1 of reference 16), which resulted in construct pBSMBB. The cDNA insert of BEV clone 129 (Fig. 1) was digested with Sau3A and a 675 bp fragment (nt positions 563 to 1237 in Fig. 2) was purified. pBSFS was generated by cloning this Sau3A fragment, which contains the ORF1a/ORF1b overlapping region, between the digested and dephosphorylated BamHI and BglII sites of pBSMBB. The orientation of the insert and the nucleotide sequences at the MHV M/BEV POL junctions were examined by sequence analysis. The hybrid gene in pBSFS is under the control of the pBS T7 promoter.

In vitro transcription and translation

Plasmid DNA of construct pBSFS was linearized by digestion with HindIII. *In vitro* transcription, using T7 RNA polymerase, and *in vitro* translation of the resulting RNA transcripts in a rabbit reticulocyte lysate were performed as described previously (17).

In vivo expression of pBSFS

Hela cells (1×10^6) were infected (m.o.i. of 5 p.f.u.) with recombinant vaccinia virus vTF7-3, which contains the T7 polymerase gene under the control of a vaccinia promoter (18). After a 90 minute incubation, the cells were transfected with pBSFS (5 μ g) as described by Gorman (19). The culture supernatant was removed at 13½ hours post infection (p.i.) and replaced by methionine-free medium. A 30 minute labelling with 60 μ Ci/ml [35 S]methionine was performed at 14 hours p.i. Cells were lysed in 300 μ l lysis buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100). Lysates were

clarified by centrifugation at 10,000 g for 60 minutes (4°C) and stored at -70°C.

Immunoprecipitation and analysis of pBSFS translation products

To immunoprecipitate pBSFS translation products, a monoclonal antibody (J.1.3; a gift from John Fleming and Stephen Stohlman; 20) directed against the N-terminal region of the MHV-A59 M protein and an antipeptide serum raised against the C-terminal 18 amino acids (aa) of the same protein (Rottier, manuscript in preparation) were used. Antiserum and SDS (see below) were added to *in vitro* translation mixtures or transfected cell lysates and the mixture was diluted in lysis buffer to give a final volume of 300 μ l. Following an overnight incubation at 4°C, 50 μ l of a 3M KCl solution and 25 μ l Pansorbin (Calbiochem) were added. After an additional hour at 4°C, immune complexes were spun down and washed three times in lysis buffer (500 μ l). Throughout the entire procedure SDS was present at a concentration of 0.1% (N-terminal serum) or 1% (C-terminal serum). Immunoprecipitated proteins were suspended in Laemmli sample buffer and analyzed on 12.5% SDS-polyacrylamide gels.

RESULTS

Partial cloning and nucleotide sequence analysis of the BEV POL gene

The cloning and sequence analysis of about 10 kb of the BEV genome (starting at the poly(A) tail) and the identification of about 3 kb of the putative BEV POL gene have been reported previously (4). To extend the BEV sequence towards the 5' end of the viral RNA, cDNA clones were selected from a random-primed cDNA library. Clones covering an additional 5 kb of sequence were obtained (Fig. 1) and used to determine the sequence of a substantial part of the BEV polymerase gene (Fig. 1 and 2).

Analysis of the obtained POL sequence (with a total length of about 8 kb) revealed that the 3' part of the BEV POL gene contains two ORFs. At this moment we cannot exclude the presence of additional ORFs in the 5' part of the gene. However to facilitate the comparison with the coronaviral POL gene (see below), the presently described ORFs have provisionally been designated ORF1a and ORF1b (Fig. 1 and 2).

The 6873 nt ORF1b, which potentially encodes a 261K protein (2291 aa), covers the 3' part of the POL gene (Fig. 1 and 2). The first potential translation initiation codon in ORF1b is located at 174–176 nt downstream of its 5' end. The 3' terminal 1053 nt of ORF1a were determined (Fig. 2) and this ORF was found to overlap with ORF1b over a distance of 12 nt, the latter being in a -1 reading frame with respect to ORF1a.

Analysis of the deduced amino acid sequence of the BEV ORF1b product

A dot matrix comparison between the deduced amino acid sequences of the BEV ORF1b product and the corresponding sequences of the coronaviruses infectious bronchitis virus (IBV, strain Beaudette (M42); 8) and mouse hepatitis virus (MHV, strain A59; 12) revealed a number of domains with significant sequence similarity (Fig. 3). The overall identity between the products of ORF1b of BEV and either IBV or MHV is approximately 19%, which is much lower than the 56% observed when IBV and MHV are compared (12). However, four domains seem to be well conserved among toro- and coronaviruses and contain a much higher percentage of identical amino acids (see Fig. 4 and Discussion).

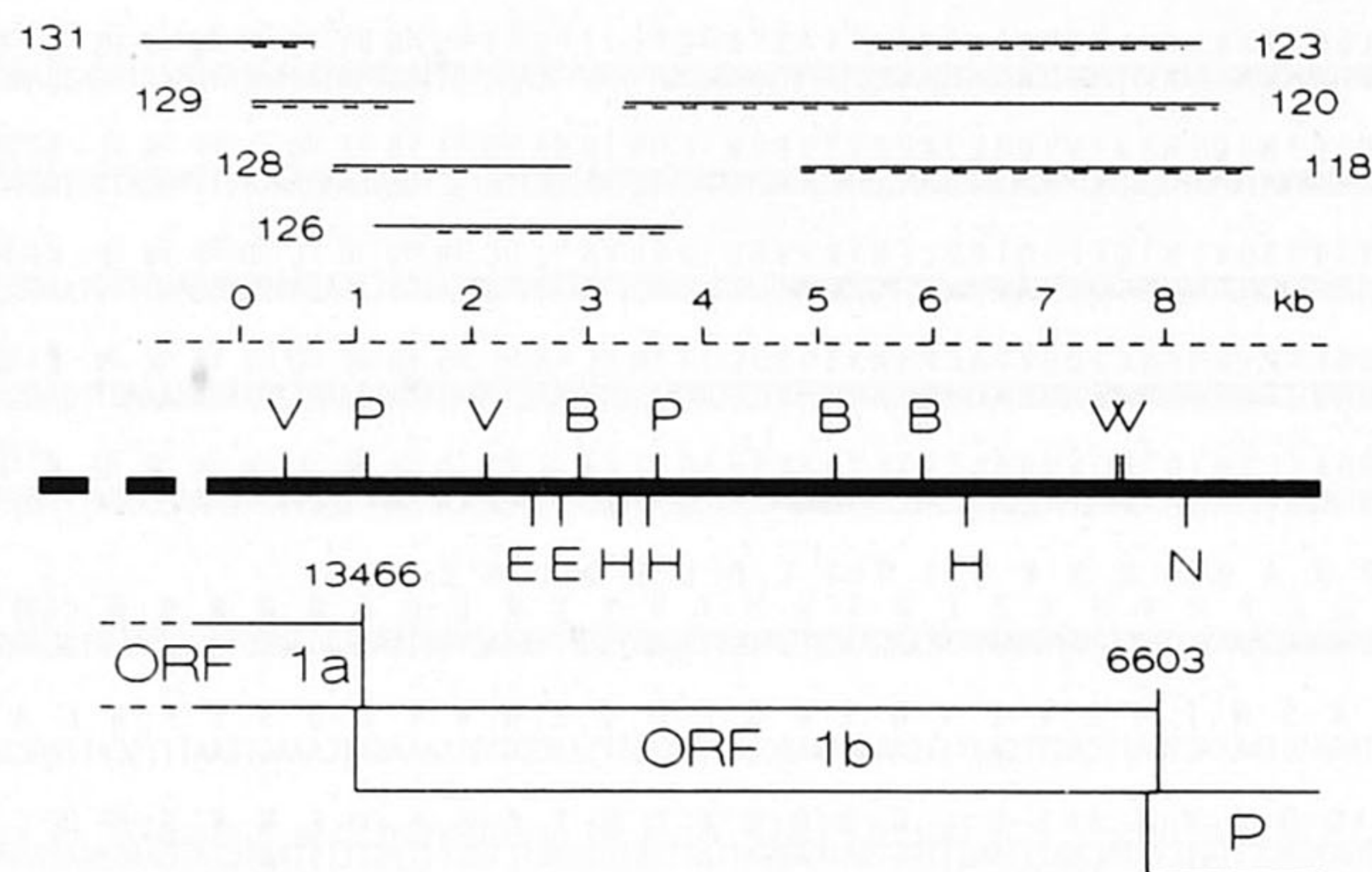


Fig. 1. Cloning and sequencing of the BEV ORF1b region. A number of restriction sites used in sequence analysis is indicated: B = BglII, E = EcoRI, H = HindIII, N = NdeI, P = PstI, and V = PvuII. The numbered lines indicate the position of the most important cDNA clones; dashed lines indicate the sequenced areas. Open boxes represent open reading frames. The ORF marked P is the BEV peplomer gene (4). Numbers indicate the distance from the start and end of ORF1b to the start of the poly(A)-tail.

Detailed analysis of these well-conserved sequence motifs revealed that two of them (domains 1 and 3 in Fig. 4) are common to RNA polymerases. Domain 1 represents the so-called 'GDD' motif (21). It has been shown to be present in viral RNA-dependent RNA polymerases and RNA-dependent DNA polymerases encoded by retroviral elements (22) and is thought to be a 'polymerase module'. As in the coronavirus polymerase motif, a serine residue precedes the two conserved aspartic acid residues in the core of the BEV domain. The percentage of identical amino acids in domain 1 of BEV and IBV/MHV is 45–50% and includes a 13 residue perfect match (KPGGTSSGDATTA) at about 70 aa upstream of the SDD core of the motif.

The amino acid sequence indicated as domain 3 (Fig. 4) is known as the 'NTP-binding' or 'helicase' motif. Putative NTP-binding domains have been identified in proteins of both viral and cellular origin (23, 24). The viral NTP-binding proteins are all thought to be involved in genome replication and/or transcription (possibly by exerting an NTP-dependent RNA or DNA duplex unwinding activity). The putative BEV helicase domain (Fig. 4) is most similar to the coronavirus one. It is remarkable that the order of the polymerase and helicase domains in both toro- and coronavirus polymerases is reversed when compared to the polymerases of other positive-stranded RNA viruses.

The ORF1b domain indicated as motif 2 (Fig. 4) contains a striking number of conserved cysteine and histidine residues. On the basis of the IBV sequence, the formation of three metal-binding 'finger' structures (25) in this region was proposed (26). However, comparison of the ORF1b sequences of BEV, MHV and IBV indicates that the metal-binding cysteine and histidine residues are conserved for only one of the three proposed finger structures. Consequently, the formation of only one conserved metal-binding finger in domain 2 remains possible, implying that the proposed 'multi-finger' model for IBV (26) requires modification.

The fourth conserved domain is located near the C-terminus of the ORF1b product (Fig. 4) and seems to be unique for toro- and coronaviruses since no significant amino acid sequence similarity with other proteins could be detected.

Analysis of the ORF1a/ORF1b overlapping region

In addition to the amino acid sequence homologies discussed above, our data indicated that the mechanism of expression of BEV ORF1b could also be comparable to the coronavirus system: the 3' part of the POL gene had been found to contain two partially overlapping ORFs and no BEV subgenomic mRNA of about 13.5 kb, which would be required for the independent translation of ORF1b, had been detected in BEV-infected cells (3, 4).

Recent studies on the gene expression of retro- (27) and coronaviruses (10, 11, 12) have identified sequence elements and RNA structures which can promote ribosomal frameshifting. First, certain heptanucleotide RNA sequences, the so-called 'slippery' sequences, have been shown to function as the actual frameshifting site. The heptanucleotide sequence 5' U UUA AAC 3' has been identified as the 'slippery' sequence involved in IBV ORF1a/ORF1b ribosomal frameshifting (11); it is also present in the corresponding region of the MHV genome (12). The same sequence was found from 10 to 4 nt upstream of the ORF1a termination codon in BEV (Fig. 2, nt positions 1045–1051, and Fig. 5).

Second, in the case of IBV (11) it has been shown that the formation of an RNA pseudoknot (28) immediately downstream of the 'slippery' sequence, is required for efficient frameshifting. Similar tertiary RNA structures are predicted for the frameshift-directing regions of MHV (12) and a number of retroviruses (11). A potential stem-loop structure was also identified just downstream of the BEV 'slippery' sequence (Fig. 5) and two sequence elements which can participate in the formation of an RNA pseudoknot were detected: nucleotides number 1099–1103 (pseudoknot 1; PK1) and 1157–1161 (pseudoknot 2; PK2) form stretches of 5 nt which are complementary to a part of the loop of the BEV hairpin (Fig. 2 and 5). From the data presented above, it was clear that the BEV ORF1a/ORF1b overlapping region contains all elements required for efficient ribosomal frameshifting.

The BEV ORF1a/ORF1b overlapping region directs ribosomal frameshifting *in vitro* and *in vivo*

To provide experimental evidence that the BEV ORF1a/ORF1b overlapping region directs ribosomal frameshifting, construct pBSFS was generated. In pBSFS a restriction fragment from BEV clone 129 (Fig. 1) containing the ORF1a/ORF1b overlap was cloned into a copy of the MHV-A59 26K membrane protein (M) gene, which is under the control of the T7 promoter (Fig. 6A). The 5' end of the MHV M ORF in pBSFS is fused in frame to the 3' end of BEV ORF1a. Termination of translation at the ORF1a termination codon will give rise to a 41K fusion protein (Fig. 6A). Since the 5' end of ORF1b was cloned in frame with the 3' end of the MHV M ORF, ribosomal frameshifting should result in the synthesis of an additional fusion protein of 50K. The two predicted products can be identified by immunoprecipitation with antibodies directed against the N- and C-terminal part of the MHV M protein (Fig. 6A).

Ribosomal frameshifting was studied *in vitro* by translation of pBSFS transcripts in a rabbit reticulocyte lysate. An *in vivo* analysis was performed by transfection of vaccinia virus recombinant vTF7-3-infected HeLa cells with pBSFS DNA. The results of both experiments were identical: translation of pBSFS RNA indeed led to the synthesis of proteins of 41K and 50K. Both products were recognized by the antiserum directed against the N-terminus of the MHV M protein, whereas only the 50K

1 V R Y I T L N L T D H L Q K Q S L K P R R R L Q Y R Q Q G V R L G G V N L G E H Q A F S N E L I S T V G Y T T W V S S T
1 AGTACGTTATATTACATTAATTAACAGACCATTTACAAAAACAAAGCCTTAAACCAGTAGACGACTCAGTATAGACAACAGGGTGTAGATTGGGAGGTGTTAATTTAGGTGAACATCAAGCCTTTTCAATGAGTTAATTTCAACTGTTGGTTATACTACATGGGTTTCATCAAC

61 V C R D N T H K H P W F V Q I P V N E K D P E W F M H N T Q L K D N Q W V V D L K P T H W L V N A D T G E Q L F A L S L
181 TGTGTGTAGAGATAACTCACAAGCATCCATGGTTTGTTCAGATACCTGTTAATGAAAAAGATCCTGAATGGTTTATGCAACAACACTCAGCTGAAAGATAATCAGTGGGTAGTTGACCTCAAACCACTCACTGGTTGGTTAATGCTGATAGCTGGTGAACAACTTTTTCATTGTCCTT

121 T D E Q A L K A E A I L Q K W S P I T Q D V E C W F K D L K G Y Y T V S G F Q P L W P V C P V N I C N V R L D P V F K P
361 AACTGATGAACAGGCTTGAAGCAGAGGCCATTTTGCAAGAGTGGTCACCAATAACACAGGATGTTGAGTGTGGTTTAAAGATTTAAAGGCTATTATACAGTTAGTGGATTCAACCACTGTGGCTGTTTGTCTGTTAATTTGTAATGTTAGACTTGATCCAGTTTAAAGCC

181 Q S I V Y A D D P T H F L S L P V V N K N F L A A F Y D L Q E G F P G K K Q V A P H I S L T M L K L S D E D I E K V E D
541 TCAGTCTATAGTTATGCTGATGATCCACTCTTCTTCTTACCTGTTGTAATAAAATTTCTTGCTGCATTTTATGATTTACAGAGGGTTTCCAGGTAAGAAACAGTGCACACATATTTCTCTCAATGTTGAAGTTGAGTGATGAAGATATAGAGAAAGTTGAAGA

241 I L D E M V L P N S W V T I T N P H M M G K H Y V C D V E G L D S L H D E V V S V L R E H G I A C D Q K R L W K P H L T
721 CATCTTGATGAGATGGTTTGCCTAATTCATGGTAACCAACCACTCATATGATGGGTAACATTATGTTGTGATGTTGAGGCTTGATTCATTCATGATGAGTGTGTCGGTGCTTAGAGAGCATGGCATTGCTGTGACCAAAAGCCCTTTGAAACCCCATTTGAC

301 I G E L N D V S F D K F K D F A I S C K L E D C D F V K L G A P K A N A R Y E F I T T L P L G D L N C *
901 TATTGGTGAGTTAATGATGTAAGTTTGTATAAATTAAGATTTGCAATTAGTTGTAAGTTAGAGGATTGTGATTGTTGTAATTTGGGTGCGCTAAGGCAATGCTAGGTATGAATTTATAACAACCTTCTCTTGGGTGATTAACTGTTGAGAGTGCCTGGAGCGCTGCAGG

13 H L C F Q N G A Y Q S S R S K H Y I D L A T E Y N A G I V K V N K S N T H S V E Y Q S K R F M I K R V K D Q S E F A L A
1081 CATCTCTGTTTCAAAATGGCGCATACCGCTTCAAGGTCAAAACATTATATGATTTGGCACTGAGTATAATGACAGGCTGTAAGAGTTAATAAATCAATACTCATTGAGTGAATATCAGAGTAAACGCTTTATGATTAAGCGTGTGAAGATCAAGTGAATTTGCATTAGCA

73 K T A F L P S I I P H H M E K O N G E W F L I R G P T S Q W S L G D L V Y A I W L G D Q D Y L S E C G F V F N P S R D E
1261 AAGACCGGCTTTTACCATCAATAATACCTCACCATATGGAAGCAAAATGGTGAATGGTTTGTATGACGGGCTTACATCTCAGTGGAGTCTGGTGACCTGTCTATGCAATTTGGCTGGTGATCAGGATTTTGTCTGAGTGGTGTGTTTGTATCCATCAAGAGATGAG

133 F L D D A N Q R S F L A N L L E P A I L N F S H I Y W Q V K M C K V P Y K L T L D N V D L N G Q L Y D F G D Y P C P N S
1441 TTTTGGATGATGCTAATCAAGGTCATTTTGGCAATTTGCTTGAGCCTGCAATTTTAAATTTTCTCATATCTATTGGCAAGTTAAATGTGAAGTCCCTACAAGTTAATGTTGAATGTGGATTGAATGGTGAGCTTTATGATTTGGTGATTATCTTGCCAAATAGT

193 V D N Q S A L F V L A E V W S M T R R P F P V A F A R L L A N E M E I P T D Y Q M F F Q N I L L S G S Y L D K A L C L N
1621 GTAGATAATCAATCTGCGCTTTTGTGTTGGCTGAAGTTGGTCTATGACTCGCAGACCATTTCCAGTAGCTTTGCAAGGCTTCTGCTAATGAGATGGAATACCAACAGACTATCAGATGTTTTCAGAACATTTTATGCTGGTAGCTACTTGGATAAAGCACTTTGTTGAAT

253 N V R P F L S D P A N L T T T P F F S Q H N G V W T H F Y N P I Y G L V E C N L D E F A E L P E V L Q Q L V T V Q G P I
1801 AATGAAGACCGTTTTGAGCGATCCAGCAATTTAAACACTACACCATTTTTCACAACATAATGGTGTGGACACATTTTATAACCCCATTTATGGTTGGTGAATGAATTTAGATGAATTTGCTGAATGCGTGAAGTTTACAACAATTTGAATGTGAGGCGCCCAATA

313 T N N M T P A I S V G E G V Y A A N V P S A S A T K Q K I P F Y D V G L C Q E L T D A G V D C G E A F K Y F Y Y L S N P
1981 ACTAATAATGACACCGCTATATCAGTTGAGAGAGTGTCTATGAGCTAATGTTCCATCCGATCTGCTACTAAACAGAAGATACCTTTTATGATGTTGGTTTGGTGAATGAATTTAGATGAATTTGCTGAATGCGTGAAGTTTACAACAATTTGAATGTGAGGCGCCCAATA

373 A G A L A D V C Y Y D Y Q G T G F Y S P K L L A G V Y D F M K R V T E C Y R I N E R F T Y E Q A K P R K S S M G I N I T
2161 GCTGGCGATGGCTGATGTTTGTCTATTATGATTATCAGGCTACTGGGTTTATTTCCCAAGTTATTAGCAGGTGTTTATGATTTATGAAAAGGTAACAGAATGTTATAGGATAATGAGAGATTACATATGAGCAAGCAAGCAAGCAAGTATCAATGGGAATAAATAAACA

433 G Y Q Q D A V Y R A L G P E N I A R L F E Y A Q K A P L P F C T K I I T K F A L S A K A R A R T V S S C S F I A S T I F
2341 GGATATCAACAAGATGAGTTTATGCTGCTTTGGGCGCTGAAATATTGCAAGGTTGTTGAATATGACAGAAAGCTCCATTGCCATTTGTACAAAAATAAACAAGTTGCAATTTGTCAGCTAAAGCAAGCAAGCAAGTTCCTCATGCTCAITTTATAGCATTCAACAATATT

493 R F A H K P V T S K M V E V A Q N S G G F C L I G V S K Y G L K F S K F L K D K Y G A I E G F D V F G S D Y T K C D R T
2521 AGATTGACATAAACCAGTAACATCTAAATGGTTGAAGTAGCAGAGAATTCAGGTGGCTTTGCCCTATTGGTGATTCAGATGTTTAAAGTTTCAAGTTTGTGAAGATAAGTACGGTGCTATTGAGGGTTTGTATGTTTGGTAGTGATTACACCAAGTGTGACCGTACA

553 F P L S F R A L T A A L L Y E L G E W D E K S W L Y L N E V N S Y M L D T M L C D G M L L N K P G G T S S G D A T T A H
2701 TTTCCACTGCTATTAGAGCTTTGACAGCGCACCTCTCTATGAATTAGGGAGTGGGATGAGAAGTCTGGTTATATCTTAATGAGGTGAATCTCATGTTGGATCTATGCTGTGTGATGAATGCTACTTAAACAACCTGGAGGACATCATCTGGTGATGCCAACATGCCAC

613 S N T F Y N Y M V H Y V V A F K T I L S D L S E G N K V M R I A A H N A Y T T G D Y Q V F N T L L E D Q F Q T N Y F L N
2881 TCAAACTATTTTATAATATATGTTGCAITATGTTGAGCTTTTAAAACAATTTTGTGAGATCTTTCAGAGGCAATAAGGTTATGAGAATTGACGACATAACGCTATACAACTGGCGATTATCAAGTGTTAATACGCTGTTGAAGATCAATTTACAGCAAAATTTTATAAT

673 F L S D D S F I F S K P E A L K I F T C E N F S N K L Q T I L H T K V D Q T K S W S T K G H I E E F C S A H I I K T D G
3061 TTTTGTCTGATGATCTTTTATATTTTCAAAACCTGAGGCTTAAAAATTTTACTTGTGAAAAATTTTCAATAAGCTACAACTATTTTACATACAAAGTAGACCAACAAATATGTTGCTACGAAAGGTCATATTGAGGAATTTTGTCTAGCACATATAATTAACACAGATGGA

733 E Y H F L P S R G R L L A S L L I L D K L S D V D I Y Y M R F V A I L C E S A V Y S R Y Q P E F F N G L F Q V F L D K V
3241 GAGTATCATTTTTACATCTAGGGCAGATTGTTGGCTTCTGCTGATCTTGATAAGCTTCAGATGTTGATTTATACATGAGATTGTTGCAATCTTGTTGAATCTGCTGTTTATCAAGATATCAGCTGAGTTTATTAAGTGTGTTTCAAGTGTCTTGATAAAGTA

793 Q Q F R K D Y C C D P C P P Q L L E R E F Y E N L V F T S N S E V G I V D C Y L E N F K L Q C E F K Q Q A N F D K V C F
3421 CAACAGTTTAGAAGGATTATGCTGTGATCCGTGCCACCTCAACTCTTAGAGCGTGAGTCTATGAGAATTTAGTTTATACATCAATAGTGAGGTGGGTTATGTTGATTGCTATCTAGAGAAATTTAAGCTTCAATGTGAGTTAAACAACAGGCAATTTTGATAAGTGTGTTT

853 C C P N P A V S V C E E C Y V P L P L C A Y C Y Y V H V V I S N H S K V E D K F K C F C G Q D N I R E L Y I V L N N S I
3601 TGTGTCCCAATCTGCAGTTTCAGTTTGTGAAGAGTGTATGTTCCCTTCCCTTTGTGCGTATTGTTATATGTCATGTTGTAATTAGTAATCATAGTAAGTAGAGGATAAGTTAAATGTTTGTGGTCAAGATAATATTAGGAGTGTATATAGTATTAATAATAGTATT

913 C M Y Q C K N C V E S D R L R I S L L S D V D Q I V R L P G F K S N S A S I A K N G V A Q L L T S V D N V D V S L D W N
3781 TGTATGATCAGTGTAAAAATGTTGATAGAGTGTATAGTAAGAATTAGTTTGTGAAGTGTGATCAATTTGTTAGGTTACCAGGTTTAAAGTCAATTTGCGCTCCATTGCAAAAAATGAGTGTGCTCAACTTTAATCAGTTGATAATGTTGATGATCTTTAGACTGGAAT

973 Y Q E S V Q Q N V A R I V Y H S A N M T Q M S I E V V Y V S F T L V R N D G S S A I L D I P N F K C P D T S Y C L F Y K
3961 TATCAAGAAAGTGTCAACAGAATGATGCTCGCATGTTTATCATTCTGCAATATGACTCAATGTCTATTGAAGTTGTATGATGAGCTTTACGCTTGTGCGTAATGATGTTCTAGTGAATCTTGATATACCAATTTTAAATGTCAGATACATCTTATGTTTATTATAAG

1033 P G K S G V L K F T G K G T L T S C Y D N K N L T W F K V T C P D F N Q P W R L A T C F V I Q Q H D V V Y P P I K A T Q
4141 CCTGTGAAAGGTGTTGTTGAAGTTTACAGGAAAGGACTTAATCATGCTATGATAAAGAATTTGACCTGGTTTAAAGTTACATGTCAGATTTAACAACCATGGGCTTGTCTACATGTTTGTATTCAACAGCATGATGTTGTTACCCCAATTAAGCAACACAG

1093 Y E N V T F V M G P P G T G K T T F V Y D T Y L S K A S S S N R F V Y C A P T H R L V G D M D E K V D G A V V V S A Y N
4321 TATGAGAATGTACCTTTGTTATGGGACACCTGGAACCTGGCAAAACCACTTTGTTTATGATACATATCTCTGAAAGCATCTTCTCGAACCGTTTGTCTATTGTGACCAACACATAGACTTGTGGAGATATGGATGAGAAGGTGATGTTGTCAGTGTGTTTTCAGCATATAAC

1153 D R T Y R N P V W N K D D S Y G V L L C T H N T L P F I K S A V L I A D E V S L I P P H V M I K I L S M G F K K V V L L
4501 GATAGGACTTACCGCAATCCAGTTTGAATAAGGATGATCTTACGGTGTCTACTTTGCACACATAACACGTTGCCATTTATTAAGGTGCTGTGTTGATTGCTGATGAAGTTTCCCTAATACCACCATGATGATGATTAATTTTGTCCATGGGTTTAAAAAGGTGTTGTTGTTA

1213 G D P F Q L S P V Y K N H K V H F K Y D T F Y L L Q L A T Q K R Y L T A C Y R C P P Q I L S A F S K P Y C D V G V D L V
4681 GGAGATCCATTTTCAGTTAAGCCAGTTTATAAAATCATAAAGTGATTTTAAATATGATCTTTTATTTATACAGCTTGCAACCAAAACGCTATCTTACAGCATGCTATGATGTCACCAACAACTTAAAGTCTTTTCAAGCCTTATTGTGATGTTGGTGTGATTAGTT

1273 S F N N K P G K F D I I V S K Q L A N I Q D F S V L S V L S K E Y P G Y V I L V N Y R A A V D Y A M Q N G L G D V T T I
4861 TCTTTTAAACAAACAGGTAAGTTTATGATAATGTTAGTAAGACGCTTGCCAAATATCAAGATTTAGTGACTTTCTGATTTGCAAAAGAGTACCTGGTTATGTTATTTAGTGAATATAGAGCTGCTGTTGATATGCAATGCAAAATGGGCTTGGTGTGACAACTATT

1333 D S S Q G T T A A N H L L V L F G A S N F S K T V N R V I V G C S R S T T H L V V V C C P E L F K H F Q P I L N W P E P
5041 GATTCATCTCAAGGAACACCGCTGCAATCATTTGCTGGTATTATTGGAGCATCCAATTTTCTAAACAGTAATCGTGTATTGTAGGCTGTCAAGATCTACACCATTTAGTAGTGGTTGTTGCCAGAACTTTTAAAGCATTTCAACCAATTTTAAATGGCCAGAACCA

1393 K Y R Y F G M E K Q S D F N I I P E V S S L V F C D I E F W H Y K A D P N S K T R T V Y P G Q I A V V T S Q T L Q L Y L
5221 AAGTATCGTTATTTGGCATGGAAGAGCAGTCTGATTTTAAATAATACCAGAAGTGTCATCTCTGTGTTTGTGATATAGAGTTTGGCATTATAAGCAGACCCCAATTCAAAACCTCGTACAGTTTATCTGGGCAATTGCTGTGTAACATCTCAAACTTTACAGTTATACCTT

1453 G V F D D T G Y K S A L R G L P K D V Y V P P N W V W M R K H Y P S Y E Q H A Y N M Q R L F K F I I D T T F G Q P W F I
5401 GGAGTCTTTGATGACATGGCTATAAGTCAGCGCTCGGGGCTACCAAGGAGTGTATGTACCACCTAATGGGTTGGATGAGAAAACATTATCCATCTTATGAACAACATGCTTATAATATGAGCGGCTGTTAAGTTTATAATGATACAACTTTGGACAACCATGGTTTAT

1513 L Y S C S N D L K S L K F Y V E F D T C Y F C S C G E M A I C L M R D G N Y K C R N C Y G G M L I S K L V N C K Y L D V
5581 TTGTATAGTTGTAGTAATGATTTAAAGAGTTTAAAGTTTATGTGGAATTTGATACCTGTTATTTTGTAGTTGGTGAAATGGCCATTGCTAATGCGTGATGGCAATTAATAATGAGAAATGTTATGGTGAATGTTGATAAGTAATAGTTAATTTGTAATATTTAGATGTG

1573 Q K E R V K L Q D A H D A I C Q Q F H G D S H E A L C D A V M T K C L Y L A S Y E A A F K D T I H V K Y K D L C L E I Q
5761 CAAAAGAAAGAGTTAAGTTACAGATGCCACGATGCTATTGTGACGAATTTATGAGAGATTCATGAGGATTTGTGATGCTGTGATGACAAAGTGCCTGATTGGCAAGCTATGAGCTGCAATTAAGATACTATACATGTGAATATAAAGATCTTGTCTAGAAATACAG

1633 Y K I T S S F V R Y D S V H K R Y L Y R D H G A M Y Y F R T P R S P M Q N V Y K Y E V G S H A E Y S I N I C T S Y E G C
5941 TATAAAATTAAGATCTTTTGTTCGATATGATGTTTCAAAACGCTATTATATCGTGATCATGGTGCCATGATTTTATAGAACACCTAGGCTCCAATGCAAAATGTTGATAAGTATGAAGTTGGTAGTCATGACAGATAGTATAAATCATCTGCTAGTTATGAAGGATG

1693 Q S F G K T C T K C I H I H C I V E Q F M A D E R F K E F I L V S V V K S D Y V E Q A L S P A A K A L M L T V T K V E D
6121 CAATCATTTGGTAAACAGTGCAAAAGTGATACATATACATTCGATTTTGAGCAGTTTATGGCAGATGAAAGATTAAAGAAATTTATACCTGTTTCTGTTGTTAAGTCAGATTATGTGAACAAGCATTGTCAACAGCAGCAAAAGCTTTGATGTTAACGGTTACGAAGGTGGAGGAC

1753 K S F Y I S N G V R Y D L Y D Y D L S K S V M R V V N S N V K P L P L Y S V I V G L G I N C T V G C V L P N V P M K L K
6301 AAAAGTTTTATATTTCAATGGTTCGCTATGATCTCTATGATTATGACTTGAGTAATCAGTCATGAGAGTAGTGAATCCAATGTTAAACCACTCCCTTGTATTCCGTTATTGGGCTTGGAAATTAATGTACTGTTGGCTGTGTTTACCTAATGTGCCAATGAAATTAAGA

1813 D E L L I T D V P L S T L R L D L Q T W Y Y I S W P T L S N K N S R W K L A G A Q V Y D C S V H I Y I E A T G E Q P L Y
6481 GATGAATGCTTATTACAGATGACCAATTATCAACCTACGGTTAGATTGCAACGCTGGTACTATATTCTCGGCCACATTAAGTAATAAAATTCGTTGGAAATTAGCTGGAGCAGAGTTTATGATTGTTGTGATATTTATATAGAGGCTACTGGAGAACACCACTTTAT

1873 Y L Q Q G K G E S L F E L P D T L F S T G R L Y N L D H D A A Q N F N V K Q L A I E T M P N N H H V F S G D F T E V G T
6661 TATTTACACAAGGTAAGGTGAAGCCTTTTGAACCTGCTGATACATTTTTCACCTGGCAGATTGTATAACTTAGACCATGATGCTGCACAGAAATTAATGTTAAACAGTAGCCATTGAAACAATGCCAAATATCATCATGTTTTCAGGAGATTTTACTGAAGTTGGAAC


```

1933 D I G G V H H V V A L N G Y K G S I I P N Y V K P I A T G L I N V G R A V K R T T L V D V C A N Q L Y E K V K Q Q L E G
6841 GATATAGGTGGAGTACATCATGTTGTTGATTGAATGGCTATAAGGGTAGTATAATCCAAATTATGTCAAGCCAATAGCAACAGGTTTGATAAATGTTGGAAGAGCTGTAACCTGTTGATGTTGTGCTAATCAGTATATGAGAAAGTGAACAACAGTTAGAAGGA
1993 V K V S K V I F V N I D F Q D V Q F M V F A N G E D D I Q T F Y P Q K D F V R S Y Y E W P N I L P Q I E S H Y D L K N Y
7021 GTTAAAGTGAGCAAGTAATTTTGTAAATATTGATTTTCAAGATGTACAGTTTATGGTATTTGCCAATGGTGAAGATGATATACAGACTTTTACCCCAAAAGGATTTTGAAGTCTGATTATGAGTGGCCAAATATTTACCACAAATTGAGTCCCATATGACCTTAAATAT
2053 G Q N P T F M P Q P V N F A K Y T Q I C T F I Q D H V K V A R N A L V W H L G A A G V D G C S P G D I V L S S F F K E C
7201 GGTCAAAATCCAACATTATGCCGCAACCTGTGAATTTTGTAAATATACACAAATTTGTACCTTTTATCAAGATCATGTGAAGTGGCAGAAATGCCCTGGTTTGGCATTAGGTGCGGCTGGTGTGATGGTTGTCACCTGGTGATATTGTACTGCTCTTTTAAAGAAATGT
2113 L V Y S W D I K D Y S T L L D K H S Y D C N F R P N L I V S D I Y N V S S N V S E V L D D C V H R L A L G G T I V F K T
7381 TTAGTTTATAGTTGGGATATTAGGATTATAGTACATTATTAGATAAGCATAGTTATGATTGTAATTTAGGCCTAATTTAATAGTAGTATATATAATGTAAGTAGTAAATGAAGTGAAGTTTGGATGATTGTGTCATCGACTTGCAATTGGGAGGAACAATAGTTTAAACA
2173 T E S S R P D I Q L S Q F T K Y F S A V Q F F T A G V N T S S S E V F V V L K Y K L Y S E P I G E E L C S P N I L S R I
7561 ACTGAAAGTTCAAGACCAGATATACAATTAAGTCAGTTTACAAAATATTTTACAGCTGTACAATTTTACAGCTGGTGAATACCTCTCTCGGAGGTTTGTGCTCTTGAATATAAATGTATTCCGAACCAATAGGTGAGGAATTGTGTTCTCCCAATATTTAAGTCGTATT
1
2233 A A Y R N K L C I V P N F K V F S T S F S Y K Y S G V K F V Q K C F Y V S V P R Q F C A S G L I Q E V P M L C Q M E H *
7741 GCTGCTTATAGGAATAAGTTATGTATAGTGCCTAATTTTAAAGTGTTAGTACTAGTTTGTAGTTAAGTATAGTGGTGAAGTTTGTGCAAAATGTTTATGTTCTGTGCGCGACAGTTTGTGCTTCTGGATTAATTCAGGAGGTGCCGATGTTGTGCCAATGGAACATTAA

```

Fig. 2. Sequence and translation of BEV ORF1b and the 3' end of ORF1a. Since ORF1b briefly overlaps with the P gene, the translation of the last block of the figure also includes the N-terminal 28 aa residues of the BEV peplomer protein. The nucleotide sequence data in this figure have been assigned the EMBL nucleotide sequence database accession number X52374.

product was precipitated when the C-terminal antiserum was used (Fig. 6B). After excision of protein bands from gel and scintillation counting, a frameshifting efficiency which ranged from 20 to 30% was calculated.

DISCUSSION

In this paper we present a substantial part of the polymerase sequence of the torovirus prototype, Berne virus. Our data on the organization and expression of the 3' part of the BEV polymerase gene supplement the previously observed similarities between toro- and coronaviruses at the level of gene order and translation strategy (4).

The detected amino acid sequence similarities strongly suggest an ancestral relationship between toro- and coronaviruses, but the two virus groups are not very closely related: whereas the ORF1b products of relatively distant coronaviruses like the avian IBV and the mammalian MHV still contain an average of 56% identical amino acids (12), the overall identity between the BEV ORF1b product and either of the corresponding coronaviral proteins is only about 19%. However, on the basis of the sequences included in Fig. 4, the percentage of identical amino acids in BEV and IBV/MHV can be calculated to rise to 45–50% for the polymerase and helicase domains (80–90% between IBV and MHV) and 40–45% for domain 4 (67% between IBV and MHV).

An analysis of the polymerase domain of toro- and coronaviruses does not reveal a closer relationship to either the 'Poliovirus-like' or the 'Sindbisvirus-like' superfamily (22). However, from a recent survey of the putative helicases of positive-stranded RNA plant viruses (29) it may be concluded that the toro- and coronaviral domain 3 is more closely related to the 'Sindbisvirus-like' helicase: almost all conserved amino acid residues from the 'Sindbisvirus-like' domain (especially from the group designated A2; 29) are conserved in the toro- and coronaviral domain 3 (Fig. 4), whereas a number of conserved residues from the 'picornavirus-like' helicase is conspicuously absent.

At the nucleotide level, the sequence of the BEV frameshift region illustrates the evolutionary distance between toro- and coronaviruses. The sequence in the frameshift region of IBV and MHV is well conserved (12). Except for the 'slippery' heptanucleotide sequence U UUA AAC, no significant sequence similarity between BEV and coronaviruses was detected in this area. Nevertheless, the RNA structure proposed for both

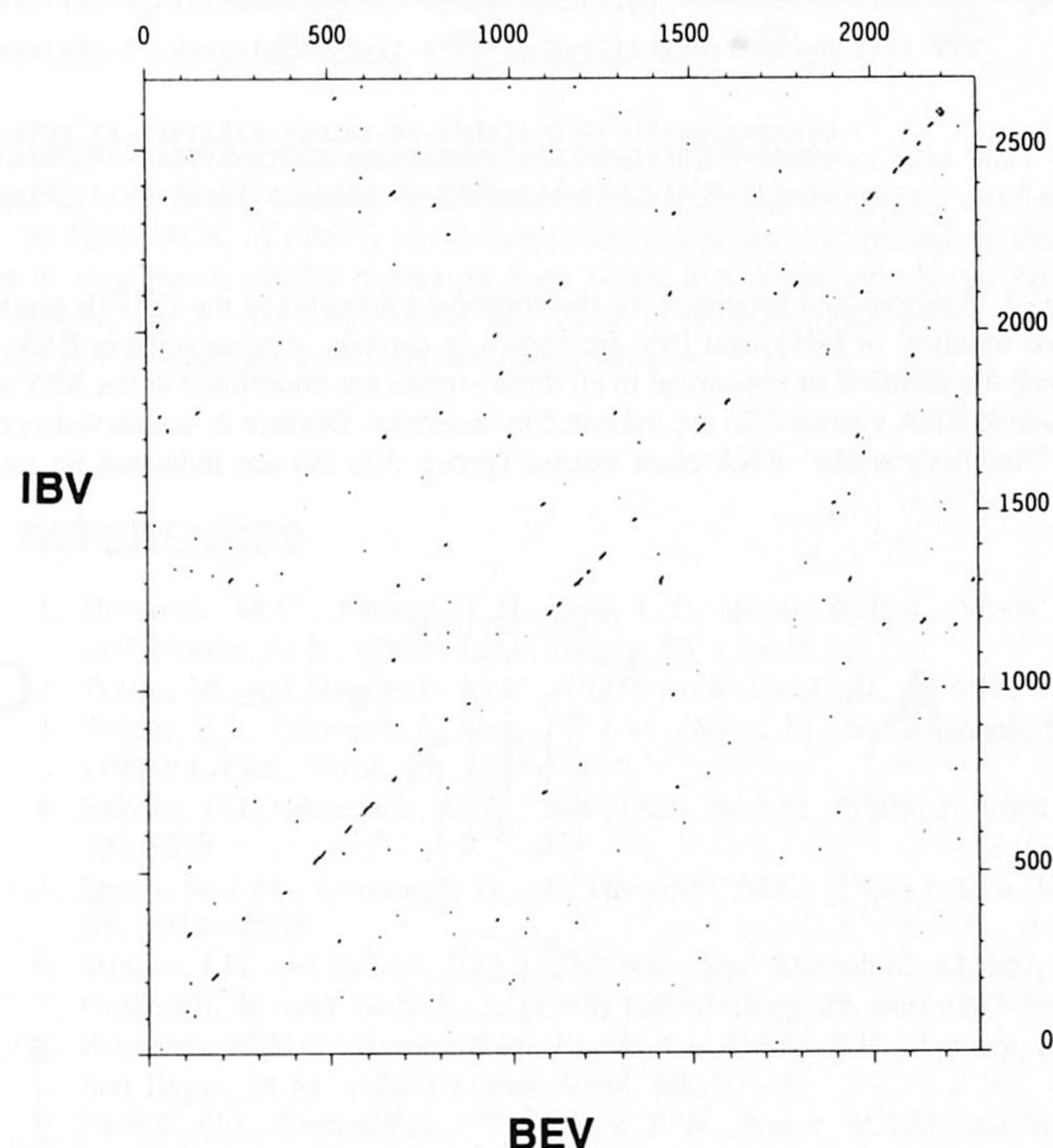


Fig. 3. Proportional dot matrix comparison of the amino acid sequence of the predicted BEV ORF1b product and the corresponding polypeptide of IBV (8). The plot was generated using the COMPARE program with a window size of 21 and a stringency of 13.

coronaviruses seems to be conserved in the BEV frameshift region. The predicted pseudoknot structure is essential for efficient ribosomal frameshifting in the coronavirus IBV (11) and can also be folded in the frameshift-directing sequences of MHV (12) and a number of retroviruses (27). The two basepaired regions S1 and S2 (28) of the proposed pseudoknot are predicted to form a quasi-continuous helix (28; Fig. 5B and 5D). This helix is thought to promote frameshifting at the 'slippery' sequence (11, 30). S1 and S2 are connected by loops L1 and L2 which bridge the deep and shallow grooves of the RNA helix (28; Fig. 5). In IBV and MHV the L2 region consists of 32 and 33 nt, respectively. In BEV two domains which can basepair with a part of the loop of the hairpin were identified at 12–16 nt and 70–74 nt downstream of the hairpin's base (Fig. 5). Although an L2 length of 11 nt (PK1; Fig. 5A and 5B) does not pose



Fig. 4. Positions and sequences of the conserved domains in the ORF1b product of the POL genes of BEV, IBV Beaudette (M42; 8) and MHV-A59 (12). Amino acids identical in MHV and IBV are shown in capitals. Amino acids in BEV which are identical in BEV, MHV and IBV are also shown in capitals. Amino acids which are identical or conserved in all three viruses are underlined in the BEV sequence. Domain 1: amino acids described to be conserved in the majority of positive-stranded RNA viruses (22) are indicated by asterisks. Domain 2: conserved cysteine and histidine residues are marked by asterisks. Domain 3: amino acids conserved in 'Sindbisvirus-like' RNA plant viruses (group A2; 29) are indicated by asterisks.

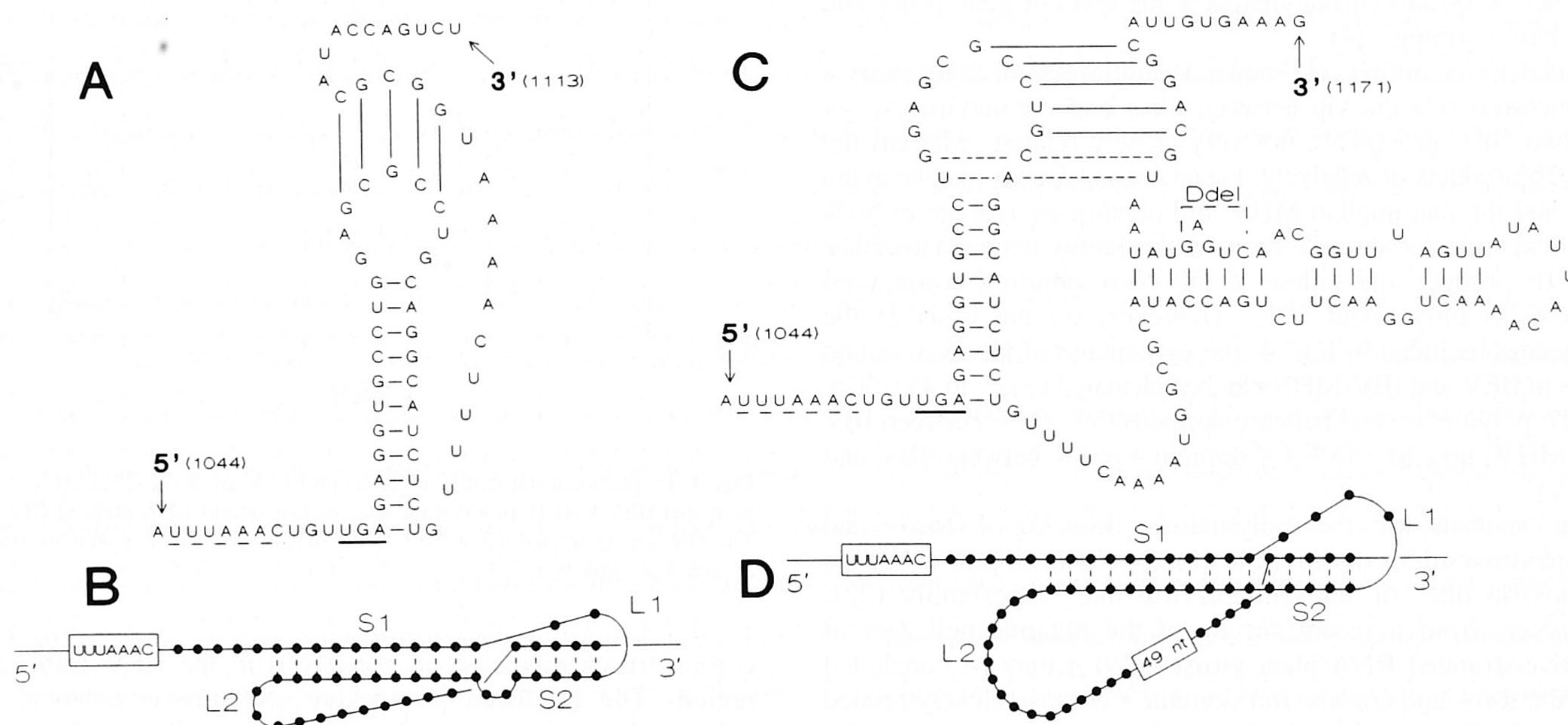


Fig. 5. Predicted secondary and tertiary RNA structure of the BEV ORF1a/ORF1b overlap region. A) Predicted structure using the most upstream region (nt 1099–1103) which is complementary to the loop of the hairpin to form an RNA pseudoknot (PK1). The 'slippery' sequence U UUA AAC is indicated by a dashed line. The ORF1a termination codon is underlined. B) Schematic representation of PK1. The 'slippery' sequence U UUA AAC is boxed. The basepaired stem structures S1 & S2 and the connecting loops L1 & L2 (28) are indicated. C) Alternative model using the sequence from nt 1157–1161 for pseudoknot folding (PK2). The DdeI site used to generate construct pBSFS δ is indicated. D) Schematic representation of PK2. The 49 nt box indicates the possible internal stem-loop structure in L2 as shown in panel C.

any sterical constraints upon pseudoknot formation (28), the involvement of the more downstream complementary region (PK2; Fig. 5C and 5D) appeared to be an interesting alternative: the resulting 69 nt L2 contains possibilities for internal folding

(Fig. 5C), which could play a role in the formation or stabilization of the pseudoknot structure.

An attempt was made to establish whether the PK1 or PK2 is involved in the formation of the BEV pseudoknot structure.

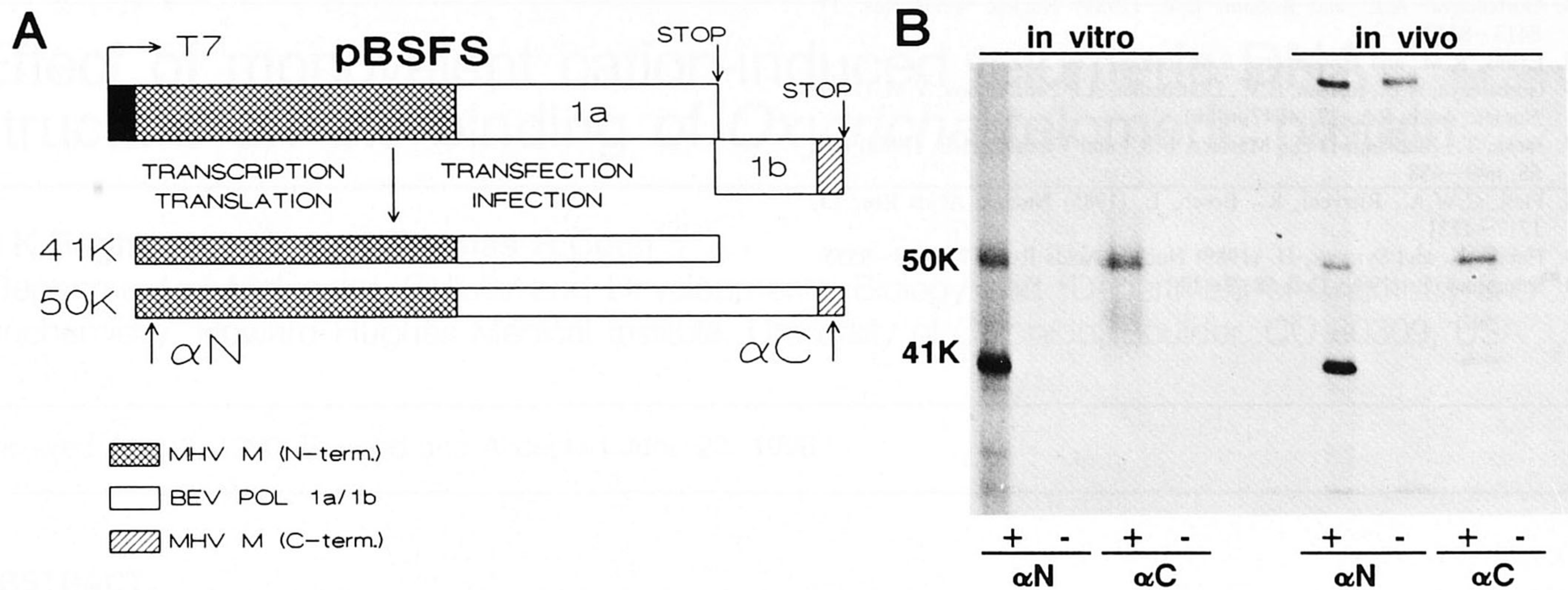


Fig. 6. Analysis of BEV ribosomal frameshifting *in vitro* and *in vivo* using construct pBSFS. A) Schematic representation of the pBSFS construct. The black box represents the T7 RNA polymerase promoter in pBS. Open boxes indicate BEV sequences; N- and C-terminal MHV M sequences are represented by differently hatched boxes. The composition of the two predicted translation products is shown. B) SDS-PAGE of pBSFS translation products immunoprecipitated by the N-terminal (αN) or C-terminal (αC) MHV M antisera. In the lanes marked '+' in the *in vitro* panel, pBSFS transcripts were added to a rabbit reticulocyte lysate. The lanes marked '-' show a control where no RNA was added to the translation system. The *in vivo* panel shows the results of an experiment with a pBSFS-transfected, vTF7-3-infected Hela cell lysate (lanes marked '+') and an untransfected vTF7-3-infected control lysate (lanes marked '-').

For this purpose construct pBSFS δ was generated (using a DdeI site at position 1145–1149 in the sequence; Fig. 5C) which contains a smaller part of BEV ORF1b and thereby lacks the downstream complementary domain of PK2 (nt position 1155–1161 in Fig. 2). However, the deletion did not result in a reduction of the frameshifting efficiency (data not shown), suggesting that PK2 is not required for efficient frameshifting. It should be noticed, however, that our present data do not formally exclude a role for this structure in translational frameshifting. If both pseudoknots are involved in the process, one structure could maintain the frameshifting efficiency in the absence of the second pseudoknot.

In conclusion, our data on the expression, organization and translation product of the 3' part of the BEV polymerase gene provide additional evidence that toro- and coronaviruses are evolutionarily more closely related to each other than to any other family of positive stranded RNA viruses. If the toroviruses are to be classified as a separate virus family, they must be considered members of a new superfamily of positive-stranded RNA viruses (6, 7), together with the coronaviruses. Alternatively, our data would justify a taxonomic proposal in which corona- and toroviruses form two separate genera of a new family of positive-stranded RNA viruses.

ACKNOWLEDGMENTS

The authors wish to thank Willem Luytjes for assistance during computer analysis, Peter Rottier for making the MHV M gene and related antisera available, and Marianne Weiss for assistance during a part of the sequence analysis. E.S. was supported by the Division for Health Research TNO in cooperation with the Organization for Scientific Research (NWO) (project number 900-502-081). P.B. received financial aid from the Organization for Scientific Research (NWO).

REFERENCES

- Horzinek, M.C., Flewett, T.H., Saif, L.J., Spaan, W.J.M., Weiss, M., and Woode, G.N. (1987) Intervirology **27**, 17–24
- Weiss, M. and Horzinek, M.C. (1987) Arch. Virol. **92**, 1–15
- Snijder, E.J., Ederveen, J., Spaan, W.J.M., Weiss, M., and Horzinek, M.C. (1988) J. Gen. Virol. **69**, 2135–2144
- Snijder, E.J., Horzinek, M.C., and Spaan, W.J.M. (1990) J. Virol. **64**, 331–338.
- Spaan, W.J.M., Cavanagh, D., and Horzinek, M.C. (1988) J. Gen. Virol. **69**, 2939–2952
- Strauss, J.H. and Strauss, E.G. (1988) Ann. Rev. Microbiol. **42**, 657–683
- Goldbach, R. and Wellink, J. (1988) Intervirology **29**, 260–267
- Bourisnell, M.E.G., Brown, T.D., Foulds, I.J., Green, P.H., Tomley, F.M., and Binns, M.M. (1987) J. Gen. Virol. **68**, 57–77
- Pachuk, C.J., Bredenbeek, P.J., Zoltick, P.W., Spaan, W.J.M., and Weiss, S.R. (1989) Virology. **71**, 141–148
- Brierley, I., Bourisnell, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D., and Inglis, S.C. (1987) EMBO J. **6**, 3779–3785
- Brierley, I., Diggard, P., and Inglis, S.C. (1989) Cell **57**, 537–547
- Bredenbeek, P.J., Pachuk, C.J., Noten, J.F.H., Charité, J., Luytjes, W., Weiss, S.R., and Spaan, W.J.M. (1990) Nucleic Acids Res. **18**, 1825–1832
- Pearson, W.R. and Lipman, D.J. (1988) Proc. Nat. Acad. Sci. USA **85**, 2444–2448
- Devereux, J., Haerberli, P., and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor University Press. Cold Spring Harbor.
- Armstrong, J., Niemann, H., Smeekeens, S., Rottier, P.J.M., and Warren, G. (1984) Nature **308**, 751–752
- Snijder, E.J., Den Boon, J.A., Spaan, W.J.M., Verjans, G.M.G.M., and Horzinek, M.C. (1989) J. Gen. Virol. **70**, 3363–3370
- Fuerst, T.R., Niles, E.G., Studier, F.W., and Moss, B. (1986) Proc. Natl. Acad. Sci. **83**, 8122–8126
- Gorman, C. (1985) In Glover, D.M. (ed.), DNA Cloning—A Practical Approach. IRL Press, Oxford, Vol. II, pp. 143–165
- Fleming, J.O., Stohlman, S.A., Harmon, R.C., Lai, M.M.C., Frelinger, J.A., and Weiner, L.P. (1983) Virology **131**, 296–307
- Kamer, G. and Argos, P. (1984) Nucleic Acids Res. **18**, 7269–7282
- Poch, O., Sauvaget, I., Delarue, M., Tordo, N. (1989) EMBO J. **8**, 3867–3874
- Hodgman, T.C. (1988) Nature **335**, 22–23

24. Gorbatenya, A.E. and Koonin, E.V. (1989) *Nucleic Acids Res.* **17**, 8413-8440
25. Evans, R.M. and Hollenberg, S.M. (1988) *Cell* **52**, 1-3
26. Gorbatenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1989) *Nucleic Acids Res.* **17**, 4847-4861
27. Jacks, T., Madhani, D.H., Masiarz, F.R., and Varmus, H.E. (1988) *Cell* **55**, 449-458
28. Pleij, C.W.A., Rietveld, K., Bosch, L. (1985) *Nucleic Acids Res.* **13**, 1717-1731
29. Habili, N. and Symons, H. (1989) *Nucleic Acids Res.* **17**, 9543-9555
30. Schimmel, P. (1989) *Cell* **58**, 9-12