Independent Structural Domains in *Paramyxovirus* Polymerase Protein**

Melanie Dochow†, Stefanie A. Krumm‡, James E. Crowe, Jr.§, Martin L. Moore¶, and Richard K. Plemper*†

From the †Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia 30322, the ‡Department of Pediatrics and Pathology, Vanderbilt University Medical Center, Nashville, Tennessee 37232, and §Children’s Healthcare of Atlanta, Atlanta, Georgia 30322

**Background:** RNA-dependent RNA polymerases of nonsegmented negative-strand RNA viruses (Mononegavirales) harbor multiple catalytic activities. However, our insight into the organization of these different enzymatic activities into a bioactive tertiary structure remains rudimentary. Fragments of Mononegavirales polymerases analyzed to date cannot restore bioactivity through trans-complementation, unlike the related L proteins of segmented NSVs. We investigated the domain organization of phylogenetically diverse *Paramyxovirus* L proteins derived from measles virus (MeV), Nipah virus (NiV), and respiratory syncytial virus (RSV). Through a comprehensive *in silico* and experimental analysis of domain intersections, we defined MeV L position 615 as an interdomain candidate in addition to the previously reported residue 1708. Only position 1708 of MeV and the homologous positions in NiV and RSV L also tolerated the insertion of epitope tags. Splitting of MeV L at residue 1708 created fragments that were unable to physically interact and trans-complement, but strikingly, these activities were reconstructed by the addition of dimerization tags to the fragments. Equivalently split fragments of NiV, RSV, and MeV L oligomerized with comparable efficiency in all homo- and heterotypic combinations, but only the homotypic pairs were able to trans-complement. These results demonstrate that synthesis as a single polypeptide is not required for the Mononegavirales polymerases to adopt a proper tertiary conformation. *Paramyxovirus* polymerases are composed of at least two truly independent folding domains that lack a traditional interface but require molecular compatibility for bioactivity. The functional probing of the L domain architecture through trans-complementation is anticipated to be applicable to all Mononegavirales polymerases.

**Results:** Domain screening and trans-complementation of *Paramyxovirus* polymerase fragments with dimerization tags identifies independent folding-competent domains.

**Conclusion:** *Paramyxovirus* polymerases harbor at least two independent domains that lack high-affinity interfaces but require molecular compatibility for bioactivity.

**Significance:** First demonstration of Mononegavirales polymerase trans-complementation sets the stage for structural analyses of folding domains.

All enzymatic activities required for genomic replication and transcription of nonsegmented negative strand RNA viruses (or Mononegavirales) are believed to be concentrated in the viral polymerase (L) protein. However, our insight into the organization of these different enzymatic activities into a bioactive tertiary structure remains rudimentary. Fragments of Mononegavirales polymerases analyzed to date cannot restore bioactivity through trans-complementation, unlike the related L proteins of segmented NSVs. We investigated the domain organization of phylogenetically diverse *Paramyxovirus* L proteins derived from measles virus (MeV), Nipah virus (NiV), and respiratory syncytial virus (RSV). Through a comprehensive *in silico* and experimental analysis of domain intersections, we defined MeV L position 615 as an interdomain candidate in addition to the previously reported residue 1708. Only position 1708 of MeV and the homologous positions in NiV and RSV L also tolerated the insertion of epitope tags. Splitting of MeV L at residue 1708 created fragments that were unable to physically interact and trans-complement, but strikingly, these activities were reconstructed by the addition of dimerization tags to the fragments. Equivalently split fragments of NiV, RSV, and MeV L oligomerized with comparable efficiency in all homo- and heterotypic combinations, but only the homotypic pairs were able to trans-complement. These results demonstrate that synthesis as a single polypeptide is not required for the Mononegavirales polymerases to adopt a proper tertiary conformation. *Paramyxovirus* polymerases are composed of at least two truly independent folding domains that lack a traditional interface but require molecular compatibility for bioactivity. The functional probing of the L domain architecture through trans-complementation is anticipated to be applicable to all Mononegavirales polymerases.

The *Paramyxovirus* family comprises major human and animal pathogens including measles virus (MeV), mumps virus, Newcastle disease virus (NDV), and respiratory syncytia virus (RSV). Collectively, members of the family are responsible for substantial morbidity and mortality worldwide. Belonging to the Mononegavirales, the Paramyxoviridae store all genetic information in a single RNA genome of negative polarity, which is encapsidated by the viral nucleocapsid (N) protein in a protein–RNA (ribonucleoprotein) complex (1). Cytosolic transcription and replication of the viral genome is mediated by the viral RNA-dependent RNA polymerase (RdRp) complex, which includes the ribonucleoprotein and viral P and ~2200-residue L proteins. Besides genome replication, the latter is responsible for viral mRNA synthesis, requiring catalysis of RNA capping, methylation, and polyadenylation in addition to phosphodiester bond formation (1).

Different catalytic activities mediated by a single polypeptide are well compatible with a multidomain architecture, in which individually folding structural domains with discrete functions are connected through linker regions with low intrinsic structure, thus ensuring local concentration of enzymatic activities in a flexible three-dimensional scaffold. Consistent with this model, analysis of purified L proteins of segmented and nonsegmented NSV proteins by electron microscopy supported a structural organization in discernible domains (2, 3). Furthermore, a systematic bioinformatics-based analysis of the L pro-

---

* This work was supported, in whole or in part, by National Institutes of Health Grants AI071002 and AI085328 (NIAID to R. K. P.).

† This article contains supplemental Tables 1 and 2.

‡ To whom correspondence should be addressed: Dept. of Pediatrics and Children’s Healthcare of Atlanta, Emory University School of Medicine, 2015 Uppergate Dr., Atlanta, GA 30322. Tel.: 404-27-1605; Fax: 404-727-9223; E-mail: rplempe@emory.edu.

§ The abbreviations used are: MeV, measles virus; NiV, Nipah virus; L, polymerase protein; N, nucleocapsid protein; NSV, nonsegmented-negative-strand RNA virus; RSV, respiratory syncytial virus; VSV, vesicular stomatitis virus; RdRp, RNA-dependent RNA polymerase, CR, conserved region; LR, large region; HPIV, human parainfluenza virus; Edm, MeV-Edmonston; CAT, chloramphenicol; nt, nucleotide; P, phosphoprotein.
tein of Lassa virus, a distantly related segmented NSV of the Arenavirus family, has revealed that the protein can be split at two positions into distinct fragments that are capable of reconstituting RdRp bioactivity through trans-complementation when co-expressed (4). Thus, Lassa virus polymerase is composed of at least three distinct, independently folding-compe-
tent structural domains. Comparable trans-complementation studies with Mononegavirales-derived L proteins are limited to date to the L protein of vesicular stomatitis virus (VSV) of the Rhabdoviridae. By contrast to Lassa virus L, a recent report found N- and C-terminal VSV polymerase fragments were unable to trans-complement each other functionally (3), although these VSV L subunits tolerated separation by an enhanced GFP moiety (5). This was interpreted to reflect that proper tertiary folding of the Mononegavirales polymerase may require synthesis of the protein as a single polypeptide (3).

Based on sequence alignments between different NNSV fam-
ily representatives, six conserved regions (CR I to CR VI) of higher sequence homology have been identified in the L pro-
teins (6, 7). Of these, the N-terminal sections harboring CR I have been implicated in L oligomerization (8–10) and/or interaction with P (8, 9, 11–14), CR III in RNA polymerization (15, 16), and CR VI in methyltransferase activity (3, 6, 17). A conserved GXXT,HR motif in NNSV L CR V, which was first identified in VSV L, is furthermore thought to mediate unusual capping of the viral mRNAs through transfer of 5′-monophosphate-mRNA onto GDP (18, 19). This polyribonucleotide- 
dytransferase activity differs from that of eukaryotic mRNA capping through guanylyltransferases, which transfer guano-
sine monophosphate to pp-mRNA to form the cap structure (20). Surprisingly, Paramyxoviruses, but not rhabdovirus, L pro-
teins also may contain traditional guanylyltransferase activity. Purified polymerase of rinderpest virus, a member of the Mor-
bullivirus genus within the Paramyxoviridae, reportedly formed covalent guanosine monophosphate L intermediates in vitro (21), and a largely conserved guanylyltransferase consensus motif required for transcriptase activity was identified near the C terminus of the human parainfluenza virus (HPIV) type 2 L protein (22).

Within the Paramyxovirus family, sequence alignments of different Morbillivirus genus polymerases suggested three large regions (LR I to LR III) separated by variable connectors (23, 24). Of these, LR I harbors CR I and II and thus the oligomeri-
ization domain, and LR II contains CR III-CR V including the predicted polymerase and polyribonucleotide- 
dytransferase activities, whereas LR III is considered to encompass the methyltransferase and recently proposed guanylyltransferase functions. Similar to recent reports for VSV L (5), L proteins of MeV, the archetype of the Morbillivirus genus, and rinderpest virus furthermore tolerated insertions into the LR II/LR III but not the LR I/LR II junction (23, 25), consistent with at least a two-domain organization. However, a comprehensive analysis of the Paramyxovirus L domain architecture and direct assess-
ment of whether Paramyxovirus L fragments located on either side of the LR II/LR III intersection remain competent for fold-
ing when expressed separately remains elusive. In analogy to the findings obtained for VSV L (3), synthesis as an intact poly-
peptide may constitute a generally conserved prerequisite for Mononegavirales L bioactivity.

To dissect fundamental principles that govern NNSV poly-
merase folding, we examined in this study three phylogeneti-
cally diverse Paramyxovirus L proteins. Commencing with MeV L, we employed in silico structure predictions combined with experimental evaluation to identify individual folding domain candidates. Guided by this screen, we generated split MeV, NiV, and RSV L pairs and explored their ability to reconsti-
tute RdRp activity in biochemical and functional assays. Our results demonstrate that Paramyxovirus polymerases are com-
posed of at least two discrete structural domains that are capa-
bile of independent folding. Neither section contains a high affinity protein-protein interface, but close physical proximity and homotypic origin from the same L protein are essential for restoring RdRp bioactivity. These findings illuminate basic principles of NNSV polymerase architecture and provide a tan-
gible path toward characterizing the structural organization of distantly related Mononegavirales polymerases beyond the Paramyxovirus family. They guide the identification of mean-
ingful NNSV substructures that may be more suitable for x-ray crystallographic analysis than full-length L proteins.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Transfection**—Baby hamster kidney (BHK-21) cells stably expressing T7 polymerase (BSR-T7/5) (26) were maintained at 37 °C and 5% CO₂ in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum and incubated at every third passage in the presence of G-418 (Geneti-
cin) at a concentration of 100 µg/ml. Lipofectamine 2000 (Invitrogen) was used for cell transfections.

**Replicon Reporter Systems**—Base vectors for all MeV replicon experiments were previously reported plasmids containing the MeV-Edmonston (Edm) strain-derived L, N, or P open reading frames under the control of the T7 promoter (27). To generate an MeV luciferase replicon reporter construct, the terminal untranslated regions of the MeV genome were added to the firefly luciferase open reading frame (Promega) through recom-
bination PCR (all oligonucleotide primers used in this study are listed in supplemental Table 1, entries #1–7) followed by replacement of the chloramphenicol (CAT) reporter cassette in the previously reported MeV-CAT replicon plasmid (27) with the sequence-confirmed recombination product. For NiV, pre-
viously described plasmids containing the NiV L, N, or P open reading frames and a NiV-CAT reporter construct were used as starting material (28). Sequence-optimized cDNA copies encoding the RSV L, N, P, or M2–1 genes based on the strain RSV A2 were synthesized in vitro (GeneArt; codon-optimized sequences are shown in supplemental Table 2) and cloned into the pcDNA3.1 expression vector (Invitrogen). The RSV mini-
genome reporter pRSVlucM5 was constructed reminiscent of a previously described RSV minigenome (29). Four overlapping oligonucleotides were annealed to form a 238-bp DNA frag-
ment containing a terminal BamHI site, the upstream 32 nucle-
otide (nt) nonstructural protein 1 (NS1) nontranslated region, the 10-nt RSV NS1 gene start signal, a 44-nt RSV leader sequence, a 94-bp hammerhead ribozyme, a 47-bp T7 termina-
tor, and a NotI compatible end (supplemental Table 1; entries
Structural Domains of Paramyxovirus Polymerase

A 191-bp DNA fragment containing a terminal HindIII site, a 155-nt RSV trailer sequence, the 12-nt RSV L gene end sequence, a 12-nt nontranslated region of RSV L, and an XhoI site was synthesized in vitro (Integrated DNA Technologies; supplemental Table 1, entry #12). These two fragments were ligated along with a BamHI/XhoI fragment of firefly luciferase cDNA (pGEM-luc, Promega) into the NotI and HindIII sites of pcDNA3.1 such that an antisense copy of luciferase flanked by RSV leader and trailer regulatory elements is produced by T7 polymerase transcription.

Mutagenesis and Generation of Expression Plasmids—For linker insertion scanning, the MeV L expression plasmid was subjected to PCR amplification using primers that duplicate the codon at the insertion site and introduce an AfeI restriction site (supplemental Table 1, entries #13–36). Amplicons were subjected to DpnI digest to remove template material followed by digestion with AfeI and ligation. Correct insertion of the linker domain was confirmed by DNA sequencing. Analogous strategies were employed to introduce streptactin SII-triple FLAG tandem tags into MeV L or triple HA tags into MeV, NiV, and RSV L (supplemental Table 1, entries #37–40 for MeV L SII-FLAG tagging; entries #41–44 for MeV L HA tagging; entries #45–48 for NiV L HA tagging; entries #49–52 for RSV L HA tagging).

To generate the MeV L\textsubscript{N-frag} expression construct, a stop codon was introduced in-frame downstream of the SII-FLAG tag in the MeV L\textsubscript{FLAG} expression plasmid (supplemental Table 1, entry #53) using directed mutagenesis after the QuikChange protocol (Stratagene). Correct insertion was confirmed by sequence analysis and immunoblotting. To generate the MeV L\textsubscript{C-frag} construct, a SpeI site followed by an ATG start codon was introduced in MeV L\textsubscript{HA} frame upstream of the HA tag (supplemental Table 1, entry #54) through PCR amplification of the C-terminal fragment. A SpeI/Sall fragment of the sequence-confirmed amplicon was then ligated into the SpeI/Sall-digested MeV L expression vector, thus replacing the L gene with the new insert.

To add a GCN4 tag to the MeV L\textsubscript{N-frag} construct, the BstBI site in the SII-FLAG cassette was first removed through directed mutagenesis (supplemental Table 1, entry #55). Using appropriate primers (supplemental Table 1, entries #56–57), the GCN4 tag was introduced through PCR amplification followed by DpnI digest of the template DNA and recirculation of the amplicon after BstBI digest. An equivalent strategy was employed for GCN4 tagging of the MeV L\textsubscript{C-frag} construct using primers annealing at the third copy of the HA tag and the ATG start codon, respectively (supplemental Table 1, entries #58–59). To generate SII-FLAG-tagged L\textsubscript{N-frag-GCN4} expression constructs of NiV and RSV L, the SII-FLAG-GCN4 cassette of MeV L\textsubscript{N-frag-GCN4} was amplified followed by joining in frame to the N termini of the L fragments through recombination PCR (supplemental Table 1, entries #60–63 for NiV; entries #66–69 for RSV). Sequence-confirmed amplicons were transferred into the NiV or RSV L expression plasmids. HA-tagged L\textsubscript{C-frag-GCN4} expression constructs of NiV and RSV L were generated through PCR-mediated insertion of the GCN4 sequence (supplemental Table 1, entries #64–65 for NiV L\textsubscript{HA}; entries #70–71 for RSV L\textsubscript{HA}). Silent BstBI (NiV L) or BplI (RSV L) restriction sites introduced into the GCN4-encoding sequences enabled re-circularization of the amplicons after DpnI-mediated removal of the template DNA and digestion with BstBI or BplI, respectively. All constructs were sequence confirmed before further experimentation.

Minireplicon Reporter Assays of RdRp Activity—BSR-T7/5 cells (2.5 × 10\textsuperscript{5} per well in a 12-well plate format) were transfected with plasmid DNAs encoding the viral RdRp components (unless otherwise specified, DNA amounts were for MeV: MeV-L (1.1 μg), MeV-N (0.4 μg), and MeV-P (0.3 μg); for NiV: NiV-L (0.2 μg), NiV-N (0.63 μg), and NiV-P (0.4 μg); for RSV: RSV-L (0.1 μg), RSV-M2 (0.15 μg), RSV-N (0.17 μg), and RSV-P (0.15 μg)), and 1 μg of the MeV or RSV luciferase replicon reporter or the NiV CAT replicon reporter plasmid. Control wells included identical amounts of reporter and helper plasmids but lacked the plasmids harboring the respective L gene or, for some experiments, L fragment subunits. In all cases vector (pUC19) DNA was added as necessary so that all transfection mixtures contained identical amounts of total DNA. Thirty-eight hours post-transfection, cells were lysed, and luciferase activities in cleared lysates were determined using Bright-Glo luciferase substrate (Promega) and an Envision Multilabel microplate reader (PerkinElmer Life Sciences), or CAT concentrations were assessed using a CAT-ELISA assay system (Roche Applied Science). Statistical significance of results was determined where indicated using Student’s t test.

SDS-PAGE and Immunoblotting—BSR-T7/5 cells (5 × 10\textsuperscript{5}) were transfected with 3 μg of plasmid DNA encoding MeV, NiV, or RSV L protein or L protein fragments as specified. Thirty-eight hours post-infection, cells were washed in phosphate-buffered saline (PBS), lysed for 10 min at 4 °C in lysis buffer (50 mM Tris, pH 8.0, 62.5 mM EDTA, 0.4% deoxycholate, 1% Igepal (Sigma)) containing protease inhibitors (Complete mix (Roche Applied Science)) and 1 mM phenylmethylsulfonyl fluoride (PMSF), and centrifuged at 20,000 × g for 10 min at 4 °C. Cleared lysates were mixed with equal volumes of urea buffer (200 mM Tris-Cl, pH 6.8, 8 mM urea, 5% sodium dodecyl sulfate (SDS), 0.1 mM EDTA, 0.03% bromphenol blue, 1.5% dithiothreitol) and incubated for 30 min at 50 °C. Samples were fractionated on SDS-polyacrylamide gels, blotted to polyvinylidene difluoride membranes (Millipore), and subjected to enhanced chemiluminescence detection (Amersham Biosciences) using specific antibodies directed against the FLAG (M2; Sigma) or HA (16B12; Abcam) epitopes, cellular GAPDH (Ambion), or MeV P (Chemicon) as specified. Blots were developed using a ChemiDoc XRS digital imaging system (Bio-Rad), and signals were assessed with the Image Lab software package.

Co-immunoprecipitation—BSR-T7/5 cells (5 × 10\textsuperscript{5}/well) were transfected with plasmid DNA encoding MeV, NiV, or RSV L, different L\textsubscript{N-frag} or L\textsubscript{C-frag} subunits, or MeV P as specified in the individual experiments. At 38 h post-transfection, cells were washed 5 times with cold PBS and lysed in co-immunoprecipitation buffer (10 mM Tris, pH 7.4, 150 mM NaCl, 1% deoxycholate, 1% Triton X-100, 0.1% sodium dodecyl sulfate (SDS), protease inhibitors (Roche Applied Science), and 1 mM PMSF). Cleared lysates (20,000 × g; 30 min; 4 °C) were incubated with specific antibodies directed against the FLAG or HA epitopes as specified at 4 °C, followed by precipitation with

6880 JOURNAL OF BIOLOGICAL CHEMISTRY
immobilized protein G (Pierce) at 4 °C. Precipitates were washed 3 times each in buffer A (100 mM Tris, pH 7.6, 500 mM lithium chloride, 0.1% Triton X-100), then buffer B (20 mM HEPES, pH 7.2, 2 mM EGTA, 10 mM magnesium chloride, 0.1% Triton X-100) followed by resuspension in urea buffer. Denatured samples were fractionated on SDS-polyacrylamide gels (4–20% gradient or 10% depending on the antigenic material assessed) followed by immunoblotting and chemiluminescence detection using specific antibodies directed against the FLAG or HA epitopes, cellular GAPDH, or MeV P as described above.

In Silico Assessment of Protein Domain Architecture—For DomCut (30)-based identification of candidate linker insertion sites, Paramyxovirus L protein sequences were aligned using the ClustalW2 (31) and MUSCLE (32) algorithm as alternative solutions in three distinct settings representing 1) different MeV genotypes, 2) a diverse panel of morbilliviruses, or 3) members of all Paramyxovirus genera. For each individual input sequence, interdomain linkers were predicted, and relative DomCut propensity scores then averaged separately based on the different sequence alignments. DomCut values for all residues that lacked an L-Edm homologue were excluded. Input L sequences were derived from MeV-Edm (genotype A) (33), MeV-Gambia (genotype B2) (34), MeV-Gambia (genotype B3.2) (35), MeV-Toulon (genotype C2) (36), MeV-Illinois (genotype D3) (34), MeV-Alaska (genotype H2) (37) for (1), MeV-Edm, MeV-Gambia B2, MeV-Gambia B3.2, MeV-Toulon, MeV-Illinois, MeV-Alaska, rinderpest virus-KabeteO (38), rinderpest virus-RBOK (39), canine distemper virus Understepoort (40), canine distemper virus 5804 (41), peste des petits ruminants virus Turkey 2000 (42), dolphin Morbillivirus (43) for (2), or MeV-Edm, canine distemper virus 5804, NIV (44), HPIV type 1 C35 (45), HPIV type 3 LZZ22 (46), HPIV type 2 (47), HPIV type 4 SKPIV4 (48), NDV-IG0210 (GenBankTM JF340367), Tupaia Paramyxovirus (49), RSV A2 (50), human metapneumovirus Sabana (51) for (3). MeDor (52) was employed to independently predict disordered domains in L-Edm using the IUPred (53), GlobPlot2 (54), DisEMBL (55), FoldIndex (56), and RONN (57) algorithms. In addition, MeV L-Edm was submitted to PONDR-FIT (58) and DRIP-PRED (59) for disorder predictions. To assess the consensus of all algorithms quantitatively, GlobPlot2, FoldIndex, and DomCut average values were shifted to positive integers, and positive output scores of all algorithms were normalized for identical hit cut-off values, averaged, transformed to a 0–10 scale, and plotted as a function of L-Edm residues. MeV L-Edm secondary structure prediction was based on the StrBioLib library of the Pred2ary program (60, 61) embedded in the MeDor package.

RESULTS

In Silico Domain Analysis of Paramyxovirus L Protein—Based on the concept that protein domains are the smallest autonomously folding-competent units within a protein structure (62), linker regions connecting individual domains are expected to show little structural order and comparably low sequence conservation. A variety of predictive algorithms have been developed that seek to identify intrinsically disordered proteins or disordered sections within a protein by means of specific sequence signatures of unstructured regions (63, 64). Because combining different predictors that explore discrete biophysical parameters reportedly boosts the accuracy of the prediction (52, 65, 66), we employed a panel of algorithms for a comprehensive in silico analysis of candidate interdomain boundaries within the Paramyxovirus L protein (Fig. 1, A and B).

The DomCut package predicts candidate linker regions based on a data set of domain/linker segments (30). We targeted three groups of viruses in these studies. Focusing on MeV L as the initial target for experimental assessment, we generated average DomCut interdomain propensity scores for L protein sequences representing a variety of MeV genotypes. We then examined different members of the Morbillivirus genus and finally members of each genus of the Paramyxovirinae and Pneumovirinae subfamilies of the Paramyxovirinae in this way.

DomCut output scores then were cross-referenced quantitatively (Fig. 1A) and graphically (Fig. 1B) with the MeDor (METAserver of Disorder (52)), POND-FIT meta-predictor (58), and DRIP-PRED (59) predictors of unstructured sections, resulting in a total application of eight algorithms. These studies highlighted 12 candidate regions in MeV L that were located between CR I and CR VI, received combined propensity scores of ≥4, and were identified by at least two of the algorithms employed. Focusing on the L core, potential interdomain regions located in the N-terminal first 408-amino acid section of MeV L, which mediates both L interaction with the viral P protein and polymerase oligomerization (8, 14), and candidates positioned downstream of the LR II/LR III junction (residues 1695–1717 (23, 24)) were not considered for experimental evaluation.

Linker Insertion Analysis of MeV Polymerase Organization—To assess the quality of the in silico predictions, 10-residue hydrophilic linker peptides were engineered into each of the selected regions. In addition, two L variants were generated that contain HA or streptactin-FLAG epitope insertions at position 1708 in the LR II/LR III junction, guided by a previously described epitope-tagged MeV L construct (23, 25). To determine the effect of linker insertions on L bioactivity, we employed a replicon reporter system that contains a firefly luciferase reporter gene flanked by the MeV non-coding genome termini. Co-expression of this construct with MeV N, P, and L proteins generated N-encapsidated RNA of negative polarity that served as the template for luciferase mRNA synthesis by the viral RdRp complex. Variation of the molar ratio of transfect L-encoding plasmid DNA relative to N- and P-encoding plasmids revealed a steep optimum curve with peak luciferase activities measured at 0.14±0.09 and 0.06 pmol (L:N and P) of plasmid DNA transfected (Fig. 1C, signal:background, 320 at peak luciferase activity). Subsequent experiments were performed at this molar ratio.

When examining the different L variants with linker insertions in this assay, we found that the L-Arg-615 linker construct returned significant luciferase activity equivalent to ∼60% that of standard L, whereas all other L variants lacked appreciable bioactivity (Fig. 1D). Replacement of the linker domain at residue 615 with streptactin-FLAG or HA epitope tags, however, substantially reduced L bioactivity to <1% that of standard L. Consistent with previous experiments (23), insertion of these
epitope tags at position 1708 in the C-terminal hypervariable region did not abolish L activity (Fig. 1D).

A previous molecular characterization of MeV L regions necessary for interaction with the viral P protein highlighted the N-terminal L residues 1–408 as directly required for P binding (14). To explore whether epitope insertion at L position 615 compromises the interaction with P, we examined the physical interaction of the proteins biochemically. Both bioactive L-Asn-1708-HA and inactive L-Arg-615-HA efficiently and specifically co-immunoprecipitated the P protein (Fig. 1E), suggesting unperturbed hetero-oligomerization of the proteins.

These results support a model in which the overall MeV L architecture comprises at least three sections that tolerate further separation through linker insertion without eliminating bioactivity. Loss of L activity after insertion of FLAG or HA epitope tags at position 615 suggests, however, that this region stands overall under tighter structural scrutiny than the LR II/LR III junction. Although P is still bound efficiently, introduction of a higher content of charged (FLAG) or aromatic/charged (HA) residues is overall not tolerated at position 615.

Independently Expressed L Fragments Assume a Physiological Conformation—To assess whether L domains located up- or downstream of residue 1708 are capable of truly independent folding or require synthesis as a single polypeptide to assume their physiological conformation, we split the L gene at this position and generated separate expression plasmids for the corresponding N- and C-terminal fragments (termed L_{N-frag} and L_{C-frag}). Strepactin-FLAG and HA epitope tags were added

---

**FIGURE 1. Identification of MeV L folding domains through linker insertion analysis.** A, average disorder propensity scores obtained through a combination of eight algorithms (listed in B) are plotted as a function of MeV L. Conserved regions I–VI (green shading); annotations according to predicted activities are detailed in the Introduction, the postulated catalytic center for RNA polymerization (GDNO), and the area and score cut-off for experimental testing (red line) are highlighted. Black arrows mark residues that were chosen for subsequent linker insertion analysis, the red arrow identifies the position in the LR II/LR III section that was previously demonstrated to tolerate epitope tag insertions (23), and the black asterisk marks a hit candidate that was not tested due to its immediate proximity to the GDNO catalytic motif. Pol, polymerase domain. B, shown is a graphic representation of individual predictions. To identify candidate disordered domains in MeV L, the MeDor (encompassing RONN, DisEMBL REM465, IUPRED, GLOBPLOT2, and FoldIndex) meta-analysis tool, PONDR-FIT, DRIP-PRED, and DomCut algorithms were used. Secondary structure predictions (SSP) are based on the Pred2ary program embedded in MeDor and show postulated α-helical (red) or β-sheet (blue) regions. Graphic predictions were aligned with the propensity score averages shown in A. C, plasmid titration to determine activity maxima of a firefly luciferase MeV RdRp reporter system is shown. Peak luciferase activities were observed at a molar ratio of 1:0.6 and 0.4 (L-encoding plasmid DNA:N- and P-encoding plasmid DNA). Total amounts of DNA transfected/well were kept constant, and values represent averages of four independent experiments ± S.D. D, linker insertion analysis to map MeV L domain intersections is shown. Twelve 10-residue linkers of the specified sequence were inserted at the indicated positions (corresponding to the black arrows in A). In addition, HA and tandem SII-FLAG epitope tags were inserted after L residues 615 and 1708. Results are expressed relative to luciferase activity observed after transfection of standard L-Edm and represent averages of three independent experiments ± S.D. Statistical analysis probes activity deviation relative to control transfections that received vector DNA in place of the L expression plasmid (***, p < 0.01; ***, p < 0.001; NS, not significant; ND, not determined). E, inactive MeV L_{615-HA} and bioactive L_{1708-HA} efficiently interact with MeV P. Whole cell lysates (WCL) of BSR-T7/5 cells co-expressing MeV P and the different L constructs as indicated were subjected to gel fractionation and immunoblotting (IB) or to reciprocal immunoprecipitation (IP) using specific antibodies directed against the HA epitopes followed by anti-P immunostaining. GAPDH was detected as internal standard. Control samples expressed untagged standard MeV L-Edm or MeV P alone.
to the newly generated N and C termini, respectively, to facilitate the biochemical characterization of the discrete L fragments.

To test the hypothesis that the reciprocal affinity of independently expressed L domains may be low, even if folding-competent, we generated a second set of L expression plasmids in which additional GCN4 affinity tags (67) were added to the N and C termini of the L fragments (detailed in Fig. 2A). These tags were expected to induce intracellular dimerization of the independently synthesized constructs. All four L variants were synthesized and showed the anticipated mobility pattern when expressed individually followed by SDS-PAGE and immunodetection using antibodies directed against the different epitope tags (Fig. 2B).

Despite efficient fragment expression, co-transfection of plasmids encoding LN-frag and LC-frag without the GCN4 tags at equimolar DNA ratios returned essentially background luciferase activities in the replicon reporter assay. This finding was independent of the combined amount of L fragment-encoding DNA added (Fig. 2C). In striking contrast, the presence of the
GCN4 dimerization domains at the L fragments significantly restored L bioactivity, resulting in a luciferase activity optimum curve largely mimicking that described above for standard L (Fig. 2C). Successful trans-complementation was dependent on GCN4-mediated oligomerization of LN-frag GCN4 and LC-frag GCN4, as alternative complementation attempts between LN-frag and LC-frag GCN4 or between LN-frag GCN4 and LC-frag did not restore bioactivity (data not shown).

Reciprocal co-immunoprecipitation of the GCN4-tagged or untagged fragments, either through precipitation with specific antibodies directed against the HA epitope followed by immunodetection with anti-FLAG antibodies or through anti-FLAG precipitation and anti-HA detection, yielded results consistent with the bioactivity data; efficient co-precipitation of the fragments was observed only in the presence of the additional GCN4 dimerization domains (Fig. 2D).

These data demonstrate that the MeV L sections located upstream or downstream of residue 1708 have no biochemically detectable affinity for each other, arguing against the presence of a traditional protein-protein interface between the LN-frag and LC-frag subunits. Both fragments are capable, however, of folding into a physiological conformation when expressed separately, as reinstating the physical proximity of the L subunits post-translationally through added GCN4 dimerization domains restored RdRp activity.

Dominant-negative Effect of LN-frag and LC-frag Subunits on Full-length L Bioactivity—To assess the full extent of functional complementation, we generated replicon-based RdRp activity curves for a set of different LN-frag GCN4 and LC-frag GCN4 plasmid DNA ratios ranging from 3.5:1 to 1:3.5. Bioactivity peaked when cells received equimolar amounts of L fragment-encoding plasmid DNA or a slight (1:1.75) excess of LC-frag GCN4, as indicated by the luciferase activity optimum curve (Fig. 3A). Statistical analysis determines the significance of activity deviation relative to the optimum curve (molar ratio of LN-frag to LC-frag plasmid DNA 1:1.75; *, p < 0.05; **, p < 0.01; ***, p < 0.001).

Dominant-negative Effect of overexpressed L fragments on RdRp activity. A, L trans-complementation efficiency peaks in the presence of a slight excess of the C-terminal L fragment. Trans-complementation MeV replicon activity assays were carried out as outlined for Fig. 2 but after transfection of different molar ratios of the LN-frag and LC-frag expression plasmids as indicated. Values are expressed relative to the L fragment combination that returned the overall highest luciferase activities (molar ratio of LN-frag to LC-frag plasmid DNA 1:1.75, total amount of L fragment-encoding plasmid DNA 0.14 pmol) and represent averages of at least three independent experiments ± S.D. Statistical analysis determines the significance of activity deviation relative to the optimum curve (molar ratio of LN-frag to LC-frag plasmid DNA 1:1.75; *, p < 0.05; **, p < 0.01; ***, p < 0.001).
whereas increasing the relative amount of L<sub>N-frag</sub>GCN4 resulted in a significant reduction of RdRp activity (Fig. 3A). Co-expression of an excess of the L<sub>N-frag</sub> subunit with standard L confirmed a dominant-negative effect of this fragment on RdRp activity (Fig. 3B). Overexpression of the L<sub>C-frag</sub> subunit relative to standard L at a 4:1 plasmid DNA ratio did not significantly affect bioactivity of standard L. Surprisingly, we observed a reduction in activity by ∼50% at a 10:1 L<sub>C-frag</sub>:L ratio, suggesting a dominant negative effect of L<sub>C-frag</sub> at high relative excess.

Consistent with previous studies locating the P binding domain in the N-terminal 408-residue section of MeV L (14), we found the L<sub>N-frag</sub> but not the L<sub>C-frag</sub> subunit to be capable of efficient co-precipitation with P (Fig. 3C). Competition for P binding between L<sub>N-frag</sub> and standard L thus emerges as a possible basis for the dominant-negative phenotype associated with this fragment in the RdRp activity assays. By analogy, the inhibitory effect of the L<sub>C-frag</sub> subunit seen at high excess could be due to direct interaction of the fragment with the corresponding C-terminal domain in standard L. To test this hypothesis, we attempted co-immunoprecipitation of L<sub>C-frag</sub> and full-length L. As seen before (Fig. 3D), whereas the GCN4-tagged L<sub>N-frag</sub> and L<sub>C-frag</sub> subunits added for control efficiently co-precipitated. This result was independent of whether cells were transfected with all plasmids of the minireplicon system (shown in Fig. 3D) or with plasmids encoding the L variants and P protein (data not shown).

These observations re-emphasize proper folding of the L<sub>N-frag</sub> subunit, enabling it to compete for interaction with RdRp complex components. Unexpectedly, the data also suggest the possibility of weak interactions between the L<sub>C-frag</sub> subunit and full-length L or P proteins, which may be responsible for the dominant negative effect associated with the fragment at high molar excess but cannot be detected biochemically.

**Epitope Tag Insertions in RSV and NiV L**—To explore whether the organization of MeV L into at least two independently folding functional domains constitutes a general feature of Paramyxovirus L proteins, we identified residues homologous to MeV L amino acid 1708 in polymerase proteins of two additional Paramyxoviruses in search of sites suitable for L epitope insertion. For this proof-of-concept approach, NiV and RSV L proteins were chosen based on their diverse phylogenetic proximity to MeV; NiV, like MeV a member of the Paramyxoviridae subfamily albeit of the Henipavirus genus, is closely related to MeV. By contrast, RSV is a member of the Pneumovirus subfamily and thus a far more distantly related representative of the Paramyxoviridae (1). Separate sequence alignments of a panel of Paramyxovirus L proteins using ClustalW2 (31) and MUSCLE (32) as alternative algorithms posited NiV and RSV L proteins to exhibit different patterns relative to MeV polymerase (Fig. 4, A and B). In the case of RSV L, in particular, residues predicted to be homologous to MeV L 1708 differed by 54 amino acids, exceeding the total length of the LR II/LR III junction in MeV L.

We, therefore, first prescreened the different predictions experimentally through HA epitope insertion in analogy to the

![FIGURE 4. Identification of HA epitope tag insertion sites in NiV and RSV L. A and B, sequence alignments of L proteins derived from MeV-Edm, NiV (88), HPIV2 (89), MuV-Miyahara (90), HPIV3–14702 (91), NDV-AF2240 (80), and RSV-A2 using the ClustalW2 (A) and MUSCLE (B) algorithms are shown. Regions corresponding to part of the LR II/LR III intersection in MeV L are shown, and numbers reflect amino acid positions. Gray boxes mark the insertion sites of the epitope tags. C, shown is expression analysis of the newly generated NiV and RSV L variants. Whole cell lysates (WCL) of BSR-T7/5 cells transfected with the different L expression plasmids were gel fractionated and immunostained (IB) using specific antibodies directed against the HA epitope or cellular GAPDH. Control cells (mock) received empty vector in place of L-encoding plasmid. D and E, shown is activity testing of the HA-tagged NiV and RSV L variants using specific replicon reporter systems for NiV (D, CAT reporter) and RSV (E, firefly luciferase reporter). Values are expressed relative to activities of standard NiV or RSV L and reflect the averages of four independent experiments ± S.D.]
showed significant differences in bioactivity ranging from ~60% that of standard RSV L in the case of RSV L_{HA-(1749)} to only 10% in the case of RSV L_{HA-(1695)} (Fig. 4E).

Thus, we identified at least one position for each L protein, at which significant RdRp bioactivity was retained in the reporter assays after HA epitope insertion. In the case of RSV L, Clust-
alW-based predictions were superior to those obtained through MUSCLE-driven alignment. Based on these results, residues 1763 in NiV L and 1749 in RSV L were chosen for L protein splitting and subsequent trans-complementation experiments.

_Homo- and Heterotypic Trans-complementation of MeV, NiV, and RSV L_—All RSV and NiV L fragments were generated with added FLAG or HA epitope tags and terminal GCN4 dimerization domains in analogy to the MeV L\_N-frag and L\_C-frag subunits. When subjected to replicon-based quantification of RdRp activity, we observed that both the NiV and RSV L fragment pairs were capable of significant trans-complementation of bioactivity similar to our initial finding with the MeV L fragments (Fig. 5, A–C). In contrast, co-expressing the different L fragments in all possible heterotypic combinations did not result in the recovery of any significant L bioactivity, indicating the inability of L fragments to interact productively with each other when derived from different Paramyxoviruses (Fig. 5, A–C).

To test whether the lack of productive interaction was due to inefficient GCN4 domain-mediated L subunit dimerization under heterotypic conditions, we subjected all fragment combinations examined to co-immunoprecipitation experiments relying on the additionally added FLAG and HA epitope tags for precipitation and immunodetection. As expected, all L subunits were synthesized successfully when expressed alone or in combination with the different counterpart fragments (Fig. 5, D and E, whole cell lysate panels). Importantly, all C-terminal L fragments were also capable of efficiently co-precipitating the different N-terminal fragments regardless of their virus origin (Fig. 5, D and E, immunoprecipitation panels), confirming full functionality of the GCN4 domains under both homo- and heterotypic conditions and hence equally productive dimerization of the different L fragment combinations.

These data demonstrate that a molecular organization into at least two independent folding domains constitutes a general feature of Paramyxovirus polymerases, as all homotypic L fragment combinations tested returned significant bioactivity in the replicon assays. Failure of trans-complementation between the heterotypic L pairs despite efficient, GCN4-mediated fragment hetero-oligomerization reveals that molecular compatibility between the subunits is required for RdRp activity. Although a biochemically appreciable high affinity protein-protein interface is lacking between the Paramyxovirus L N- and C-terminal fragments, physical proximity of the subunits alone is necessary but not sufficient for RdRp bioactivity.

**DISCUSSION**

Reflecting their central position in the life cycle of all RNA viruses, RNA-dependent RNA polymerases are determinants for viral pathogenesis and constitute attractive targets for antiviral therapeutics (68, 69). Crystal structures, either free or complexed with nucleic acid substrates, are available for RdRps derived from a variety of different viral families (70–72). These structures revealed a conserved fundamental organization of the proteins into a geometry resembling a “right-hand” shape in which the “fingers,” “palm,” and “thumb” domains are thought to ensure the correct positioning of substrates and metal ion cofactors (71, 73). Although it has been proposed that this basic shape may be conserved in all nucleotide polymerases (74), our basic understanding of Mononegavirales L protein structure and organization remains in its infancy.

Examining different L proteins derived from representatives of the Paramyxovirinae and Pneumovirinae subfamily, our study illuminates three principles of Paramyxovirus polymerase folding. First, the proteins are composed of a set of at least two truly independent structural units that are competent of proper folding. Second, assuming an enzymatically active tertiary conformation does not require synthesis of the protein as a single polypeptide and, consequently, does not adhere to an integrated folding concept. Third, molecular compatibility between the discrete subunits is essential for functionality, but the subunits lack a high affinity protein-protein interface. Several lines of evidence support these conclusions.

Our bioinformatics analysis has highlighted multiple candidate regions in the Paramyxovirus L protein that may represent domain boundaries. The identification of the MeV L LR II/LR III intersection as a candidate zone by several of the algorithms used supports the validity of the in silico predictions, as successful expansion of this region was reported previously for Morbillivirus L proteins (23, 25). Our subsequent linker insertion scan confirmed these data and, by extending successful epitope insertion to distantly related L proteins, illuminated a conserved Paramyxovirus polymerase organization into at least two distinct sections. In addition, we identified an N-terminal position in MeV L, residue 615 located downstream of the N-terminal L-P and L-L oligomerization domain (8, 14), that likewise tolerates enlargement without eliminating L bioactivity. This finding did not extend, however, to acceptance of FLAG or HA epitope tags, which contain a high density of charged or bulky aromatic amino acids (75, 76) and are thus expected to be structurally more active than the 10-residue linker used in the primary scan. Interestingly, L proteins of the

---

**FIGURE 5. Functional trans-complementation mandates homotypic L fragment combinations.** A–C, shown are homo- and hetero-complementation assays after co-transfection of plasmids encoding GCN4 domain-tagged MeV, NiV, or RSV L\_N- and C-terminal fragments in all combinations in the context of the MeV (A), NiV (B), or RSV (C) replicon reporter systems. Values show relative luciferase or CAT activities and represent the averages of at least three independent experiments ± S.D. Control cells received the respective L\_N-frag or L\_C-frag encoding plasmids. Statistical analysis assesses the significance of deviation of the different L fragment combinations from the controls (**, p < 0.001; NS, not significant). D–F, all homo- and heterotypic L\_N- and C-fragment combinations analyzed in A–C efficiently co-immunoprecipitate. Whole cell lysates (WCL) of BSR-T7/5 cells expressing the GCN4-tagged homo- and heterotypic L fragment pairs as specified were subjected to immunoprecipitation (IP) with anti-HA antibodies followed by gel fractionation and immunoblotting (IB) using anti-FLAG antibodies. In parallel, lysates were directly analyzed using HA-, FLAG-, or cellular GAPDH-specific antibodies. Controls received expression plasmids encoding standard, untagged L in place of the epitope-tagged L variants.
Structural Domains of Paramyxovirus Polymerase

 Arenavirus family reportedly comprise at least three independent folding domains (4).

Based on our data, we conclude with confidence that at least two defined structural domains are present in Paramyxovirus polymerases. The N-terminal fragment harbors P binding, polymerase and polyribonucleotidyltransferase activities, whereas the C-terminal section contains methyltransferase and guanylyltransferase motifs and thus contributes predominantly to the generation of proper mRNA cap structures. Although comparative evaluations of different algorithms have demonstrated that a combination of discrete predictors has the highest prospect to maximize the overall accuracy of the prediction (52, 65, 66), our study also highlights the limitations of currently available in silico tools when applied to very large polypeptides such as Mononegavirales L proteins. Therefore, additional domain intersections may well exist in the Paramyxovirus L protein that were not detected by our in silico analysis and experimental evaluation.

We have shown that the addition of GCN4 dimerization domains to independently expressed Paramyxovirus L N- and C-terminal fragments results in significant trans-complementation of RdRp bioactivity. These data demonstrate that post-translational interaction of individually synthesized polypeptides is sufficient for the formation of a catalytically active tertiary structure. Because the replicon reporter system used in our study first generates T7-polymerase-driven RNA molecules of negative polarity, luciferase reporter activity confirms full transcriptase functionality, comprising RNA polymerization, polyadenylation, and capping and methylation of the non-covalently linked N- and C-terminal L fragments. Differently composed NNSV polymerase oligomers are thought to carry out replicase and transcriptase activities. Although replicase activity is mediated by an L-P-N core complex, the transcriptase comprises cellular proteins in addition to L and P (77–79). Although in all cases statistically significant, luciferase activity levels representing reconstituted complexes did not reach the level of those obtained with standard L. This finding does not necessarily reflect, however, that non-covalently linked complexes are inherently less bioactive. A major contributing factor to lower activity may be that GCN4 mediated interactions are not intrinsically geared toward mediating enzymatically productive L\(_{N\text{-frag}}\)/L\(_{C\text{-frag}}\) dimerization events but will most likely drive the formation of L\(_{N\text{-frag}}\) and L\(_{C\text{-frag}}\) homodimers with equal affinity. Our observation that a 3.5-fold molar excess of either fragment is sufficient to reduce RdRp activity substantially in the replicon trans-complementation assay supports this notion.

Last, the co-precipitation experiments and functional RdRp assays performed in the absence of additional GCN4 tags demonstrate that the L fragments, albeit competent for folding, have no biochemically appreciable inherent affinity for each other. The previously reported lack of VSV L trans-complementation (3) suggests that this may, in fact, constitute a general theme of Mononegavirales L domain interaction, which stands in stark contrast to the strong inherent domain affinities observed for L proteins of segmented negative strand RNA viruses of the Arenaviruses family (4). Successful splitting of all three Paramyxovirus polymerase proteins analyzed in our study at homologous positions and the presence of comparable catalytic motifs in methyltransferase in CRV VI (6, 17, 81, 82) and guanylyl-transferase near the C terminus (21, 22) in the MeV, NIV, and RSV L\(_{C\text{-frag}}\) subunits support that the different L C-terminal fragments harbor equivalent enzymatic activities. Clearly, bringing the Paramyxovirus L N- and C-terminal fragments into close physical proximity is a prerequisite for reconstituting RdRp activity. If proximity alone were sufficient for complementation, however, heterotypic trans-complementation between L fragments derived from different Paramyxovirus family members and brought into proximity through GCN4-mediated dimerization should have been successful. Because heterotypic complementation did not restore functionality despite efficient biochemical interaction of the various fragments, our data suggest an architectural model of Mononegavirales L proteins in which independently folding subdomains do not share a traditional protein-protein interface but require low affinity molecular compatibility to achieve functionality.

Mammalian cells are not capable of RNA-dependent RNA polymerization, rendering the viral RdRp complexes attractive targets for antiviral therapies. Novel nucleoside and allosteric inhibitors of hepatitis C virus polymerase, for instance, are at different stages of clinical evaluation and show strong drug potential (83–85). Experimentally tested allosteric blockers of Paramyxovirus polymerase likewise combine high potency with minimal cytotoxicity, opening a desirably large therapeutic window (86, 87). High resolution information about the Mononegavirales polymerase core structure would pave the path for structure-guided de novo drug design efforts, the informed optimization of existing inhibitor scaffolds, and the proactive design of inhibitor protected against rapid viral escape from inhibition. At present, high resolution structures of L have not been solved for any of the NNSV family members. Identifying structurally meaningful fragments that assume a physiological conformation when expressed independently emerges as a valid strategy toward obtaining structural insight. Confirming proper folding of candidate Paramyxovirus L fragments through successful trans-complementation after the addition of GCN4 dimerization tags provides a novel approach toward achieving this goal, which is most likely transferable to the analysis of Mononegavirales L proteins that do not originate from members of the Paramyxovirus family.

Acknowledgments—We thank A. L. Hammond for critical reading of the manuscript, P. A. Rota for expression plasmids encoding NIV N, P, and L and the NIV replicon reporter construct, and J. P. Snyder and A. Prussia for helpful discussions concerning implementation of the DomCut algorithm.

REFERENCES


 Structural Domains of Paramyxovirus Polymerase


Structural Domains of Paramyxovirus Polymerase


59. Deleted in proof.


66. Ferron, F., Rancurel, C., Longhi, S., Camilllau, C., Henriussat, B., and Ca-

*Structural Domains of Paramyxovirus Polymerase*
### Supplement Table 1: Oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Entry #</th>
<th>Oligonucleotide Sequence</th>
<th>Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>atggagacgccaaacaaatcaggg</td>
<td>MeV-Luc-Replicon</td>
</tr>
<tr>
<td>2</td>
<td>ttcaatggatttcttgcg</td>
<td>MeV-L418</td>
</tr>
<tr>
<td>3</td>
<td>cattctagctgcacgagg</td>
<td>MeV-L517</td>
</tr>
<tr>
<td>4</td>
<td>ggcagagagagagag</td>
<td>MeV-L1416</td>
</tr>
<tr>
<td>5</td>
<td>ggcggagacaccggaaagcagctcctcgatcttgtcctccagctgcttcatccactagtgagctccccggg</td>
<td>MeV-L418</td>
</tr>
<tr>
<td>6</td>
<td>attagtggggttcgcctcgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctgtgcctg</td>
<td></td>
</tr>
</tbody>
</table>
### Supplement Table 2: Codon-optimized sequences of RSV N, P, M2-1 and L genes.

#### RSV-N

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>gtttaaagaggagatcttcac                                             GCN4</td>
</tr>
<tr>
<td>62</td>
<td>tgttttaaataggaggattctctac                                           NiV-Lc.05*</td>
</tr>
<tr>
<td>63</td>
<td>gtagagaatcctcctattaaaacatggagccacccgcagtttg                         GCN4</td>
</tr>
<tr>
<td>64</td>
<td>ggtcgacttaattaattaggcctctcgagctgcagggattcattatctctcgccgatcagcttc   NiV-Lc.05*</td>
</tr>
<tr>
<td>65</td>
<td>gtagatcttcgaaagcagctcctcgatcttgtcctccagctgcttcatccatggtattatcgtgttttc RSV-Lc.05*</td>
</tr>
<tr>
<td>66</td>
<td>ggagacccaagctggctagc                                              GCN4</td>
</tr>
</tbody>
</table>

#### RSV-P

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>gtttaaagaggagatcttcac                                             GCN4</td>
</tr>
<tr>
<td>62</td>
<td>tgttttaaataggaggattctctac                                           NiV-Lc.05*</td>
</tr>
<tr>
<td>63</td>
<td>gtagagaatcctcctattaaaacatggagccacccgcagtttg                         GCN4</td>
</tr>
<tr>
<td>64</td>
<td>ggtcgacttaattaattaggcctctcgagctgcagggattcattatctctcgccgatcagcttc   NiV-Lc.05*</td>
</tr>
<tr>
<td>65</td>
<td>gtagatcttcgaaagcagctcctcgatcttgtcctccagctgcttcatccatggtattatcgtgttttc RSV-Lc.05*</td>
</tr>
<tr>
<td>66</td>
<td>ggagacccaagctggctagc                                              GCN4</td>
</tr>
</tbody>
</table>

#### RSV-M2-1

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>gtttaaagaggagatcttcac                                             GCN4</td>
</tr>
<tr>
<td>62</td>
<td>tgttttaaataggaggattctctac                                           NiV-Lc.05*</td>
</tr>
<tr>
<td>63</td>
<td>gtagagaatcctcctattaaaacatggagccacccgcagtttg                         GCN4</td>
</tr>
<tr>
<td>64</td>
<td>ggtcgacttaattaattaggcctctcgagctgcagggattcattatctctcgccgatcagcttc   NiV-Lc.05*</td>
</tr>
<tr>
<td>65</td>
<td>gtagatcttcgaaagcagctcctcgatcttgtcctccagctgcttcatccatggtattatcgtgttttc RSV-Lc.05*</td>
</tr>
<tr>
<td>66</td>
<td>ggagacccaagctggctagc                                              GCN4</td>
</tr>
</tbody>
</table>

#### RSV-L

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>gtttaaagaggagatcttcac                                             GCN4</td>
</tr>
<tr>
<td>62</td>
<td>tgttttaaataggaggattctctac                                           NiV-Lc.05*</td>
</tr>
<tr>
<td>63</td>
<td>gtagagaatcctcctattaaaacatggagccacccgcagtttg                         GCN4</td>
</tr>
<tr>
<td>64</td>
<td>ggtcgacttaattaattaggcctctcgagctgcagggattcattatctctcgccgatcagcttc   NiV-Lc.05*</td>
</tr>
<tr>
<td>65</td>
<td>gtagatcttcgaaagcagctcctcgatcttgtcctccagctgcttcatccatggtattatcgtgttttc RSV-Lc.05*</td>
</tr>
<tr>
<td>66</td>
<td>ggagacccaagctggctagc                                              GCN4</td>
</tr>
</tbody>
</table>
Structural Domains in Paramyxovirus Polymerase