

# Modulating the Function of the Measles Virus RNA-Dependent RNA Polymerase by Insertion of Green Fluorescent Protein into the Open Reading Frame

W. Paul Duprex,\* Fergal M. Collins, and Bert K. Rima

School of Biology and Biochemistry, The Queen's University of Belfast, Belfast BT9 7BL, Northern Ireland, United Kingdom

Received 3 December 2001/Accepted 12 April 2002

**Measles virus (MV) is the type species of the *Morbillivirus* genus and its RNA-dependent RNA polymerase complex is comprised of two viral polypeptides, the large (L) and the phospho- (P) proteins. Sequence alignments of morbillivirus L polymerases have demonstrated the existence of three well-conserved domains (D1, D2, and D3) which are linked by two variable hinges (H1 and H2). Epitope tags (c-Myc) were introduced into H1 and H2 to investigate the tolerance of the variable regions to insertions and to probe the flexibility of the proposed domain structures to spatial reorientation. Insertion into H1 abolished polymerase activity whereas introduction into H2 had no effect. The open reading frame of enhanced green fluorescent protein was also inserted into the H2 region of the MV L gene to extend these observations. This resulted in a recombinant protein that was both functional and autofluorescent, although the overall polymerase activity was reduced by over 40%. Two recombinant viruses which contained the chimeric L genes EdtagL(MMc-mycM) and EdtagL(MMEGFPM) were generated. Tagged L proteins were detectable, by indirect immunofluorescence in the case of EdtagL(MMc-mycM) and by autofluorescence in the case of EdtagL(MMEGFPM). We suggest that D3 enjoys a limited conformational independence from the other domains, indicating that the L polymerases of the *Mononegavirales* may function as multidomain proteins.**

*Measles virus* (MV) is the type species of the *Morbillivirus* genus, which is one of the seven genera of the family *Paramyxoviridae*. The *Paramyxoviridae*, *Bornaviridae*, *Filoviridae*, and *Rhabdoviridae* make up the order *Mononegavirales*, and other well-known prototypes of this grouping include Sendai, Borna disease, Ebola, and rabies viruses. The viruses are enveloped and contain single-stranded, negative-sense RNA genomes of between 9 and 19 kb. Virion RNA is complexed in a helical nucleocapsid comprised of the nucleocapsid (N), phosphoprotein (P), and large (L) proteins. This ribonucleoprotein (RNP) complex is the obligatory template for viral transcription and replication and unencapsidated viral genomic RNA is transcriptionally inactive and not infectious. RNA-dependent RNA polymerase (RdRp) activity is associated with the L and P proteins (12). Genes are expressed from a single promoter at the 3' end of the genome. Although at present there is no direct evidence for its direct role in RNA synthesis, it is assumed that the L protein contains the polymerase activity and the P protein has an accessory role. Genetic or biochemical studies have also demonstrated other activities such as capping, methylation, and polyadenylation (1, 10, 13).

It is unlikely that the L protein exists as a single globular structure because of its size, 2,183 amino acids (~250 kDa) for MV. Sequence alignments of the morbillivirus L proteins revealed two highly variable regions, which we termed hinges. The first (H1) extends from residues 607 to 650 and the second (H2) is comprised of residues 1695 to 1717; the hinges may

form domain boundaries (17). Alignments of more distantly related L proteins from the *Paramyxoviridae* and *Rhabdoviridae* have identified six blocks of conserved sequence (21). That study identified H1 as a region of low homology and variable length but did not allow the precise location of a second variable region in the morbilliviruses, identified by McIlhatton et al. (17), to be pinpointed. Complementation studies with vesicular stomatitis virus temperature-sensitive mutants defined several intragenic complementation groups in the L protein. This is also consistent with the existence of multiple functional domains in this protein (22). The functionality of chimeric L proteins with a mixture of canine distemper virus (CDV) and MV domains further suggests that this is the case (Collins et al., unpublished data).

Green fluorescent protein (GFP) has dramatically enhanced studies of many molecular events in living cells. Detection of autofluorescence permits the real-time observation of such diverse processes as cytoskeletal assembly (16), chromosome dynamics (32), and cell signaling (4). A widely used approach is to fuse GFP to either the amino or the carboxy terminus of the protein of interest and to examine autofluorescence within the cells over time and thereby gain a more complete understanding of changes in the intracellular localization. Recently we have used a modified form of this reporter gene, enhanced GFP (EGFP), to study virus cell-to-cell spread in vitro and in vivo (7, 8).

Transcription and replication of the *Mononegavirales* have been extensively studied using synthetic genome analogues comprised of the viral terminal *cis*-acting sequences flanking an antisense reporter gene. Expression of the reporter gene permits the L protein function to be quantified (26). Based on the alignment of morbillivirus L proteins we chose to focus on the hinge regions of MV in an attempt to probe the tolerance

\* Corresponding author. Mailing address: School of Biology and Biochemistry, The Queen's University of Belfast, Medical Biology Centre, 97 Lisburn Rd., Belfast BT9 7BL, Northern Ireland, United Kingdom. Phone: 44 28 9027 2060. Fax: 44 28 9023 6505. E-mail: p.duprex@qub.ac.uk.

to insertions of this highly conserved catalytic protein. We describe the insertion of epitope tags into H1 and H2 and the insertion of the EGFP open reading frame (ORF) into H2. Insertion of a c-Myc tag into H1 abolishes polymerase functionality whereas insertion in H2 has no effect. Insertion of the entire EGFP ORF into the same region of the polymerase significantly reduced but did not abolish protein function. The availability of an MV rescue system (23, 25) enabled modified L genes to be introduced into a full-length MV plasmid. Rescued recombinant viruses were detectable, either indirectly using immunofluorescence or directly due to the autofluorescence of the L protein. This enables visual tracking of the fluorescent L protein in infected cells.

## MATERIALS AND METHODS

**Cell culture and viruses.** Vero and HeLa cells were grown as described previously (6). Recombinant MVs, Edtag, the unmodified rescued Edmonston virus (23), and MVeGFP, a recombinant MV expressing EGFP from an additional transcription unit (7), were used as controls. Modified vaccinia virus Ankara, which expresses T7 RNA polymerase (MVA-T7), was grown as described before (6).

**Growth analysis.** This was carried out essentially as outlined before except that supernatant- and cell-associated viruses were collected every 4 or 8 h (6). Samples were stored at  $-70^{\circ}\text{C}$  and titers were determined by 50% tissue culture infective dose.

**Construction of pEMC-La Eco47III.** The vector pEMC-La, a plasmid from which the L gene is transcribed, was generated to facilitate the rescue of recombinant MVs (23). The vector backbone is based on pTM1 (19), in which an *NcoI* site overlaps the ATG start codon of the L ORF, a T7 promoter directs transcription, and an internal ribosome entry site directs translation. A unique *Eco47III* site was inserted at the 3' terminus of the L ORF by overlap PCR (18) using primers containing *Eco47III* sites (in italics) (emcl-2 [5'-GTC CTT AAT CAG AGC GCT GTA TCC GAC TAA-3'] and emcl-4 [5'-TTA GTC GGA TAC AGC GCT CTG ATT AAG GAC-3']) and two additional primers, emcl-1 (5'-CGA AGT CAC GTG GGT AGG CA-3'), representing nucleotides 14,789 to 14,808 of the MV genome, and emcl-3 (5'-TAG CTC GAC ATT GTC ACC TG-3'), corresponding to sequences 8,785 to 8,755 of pTM1. The mutagenized PCR product was digested with the restriction enzymes *AocI* and *AspI* and was ligated into similarly treated pEMC-La vector to generate pEMC-La Eco47III.

**Construction of pEMC-La(MMM).** Unique restriction sites were introduced into the hinge regions of the L gene by sequentially assembling the complete gene domain by domain. Cloning was carried out in an insertion vector and the resultant plasmid was named pEMC-La(MMM) to indicate that the three domains were derived from MV. This vector was generated by removing almost the complete ORF of the L gene from pEMC-La Eco47III by restriction digestion with *NcoI* and *Eco47III*. Two complementary oligonucleotides containing the restriction sites *BssHII* (italics) and *AatII* (underlined) were synthesized. These were designated LCU (5'-CAT GGA TAT TTG AGC GCG CTG TGT ATG ACG TCG CAT ACC GAG C-3') and LCL (5'-GCT CGG TAT GCG ACG TCA TAC ACA GCG CGC TCA AAT ATC-3'). They were annealed (6) to generate overhangs which were compatible with the *NcoI*- and *Eco47III*-cleaved pEMC-La Eco47III to produce the plasmid pEMCassette. Domain 1 of MV L was generated by PCR amplification from pEMC-La using primers MVL *NcoI*F (5'-GAT AAT ACC ATG GAC TCG CTA TCT GTC-3') and MVL *BssHII*R (5'-TTT TGC TGC GCG CAC GTT-3'). Upon treatment with *NcoI* and *BssHII* this fragment could be ligated into the similarly treated pEMCassette, creating the plasmid pEMC-La(M). Domain 2 of MV L was generated by PCR amplification from pEMC-La using primers MVL *BssHII*F (5'-GAA CGT GCG CGC AGC AAA AGG GTT TAT AGG-3') and MVL *AatII*R (5'-TGA GAT GAC GTC GCT GCC GAT CTT TGG-3'). Following digestion with *BssHII* and *AatII* this fragment was ligated into the similarly treated pEMC-La(M), creating the plasmid pEMC-La(MM). Domain 3 of MV L was generated by PCR amplification of pEMC-La using primers MVL *AatII*F (5'-GGC AGC GAC GTC ATC TCA AAT ATG AGC ATC-3') and MVL *Eco47III*R (5'-CCT TAA TCA GAG CGC TGT ATC CGA CTA AC-3'). Upon restriction digestion with *AatII* and *Eco47III* this fragment was ligated into the similarly treated pEMC-La(MM), creating the plasmid pEMC-La(MMM), which contained all three domains of MV L and unique restriction sites in H1 and H2.

**Insertion of epitope tags into hinge regions of MV L.** Sequences encoding epitope tags were inserted into the hinge regions of the L genes by annealing overlapping oligonucleotides (uppercase denotes c-Myc epitope coding sequence). Oligonucleotides *BssHII* upper (5'-cg cgc atg GCA TCA ATG CAG AAG CTG ATC TCA GAG GAG GAC CTG ctt g-3') and c-myc *BssHII* lower (5'-g tac CGT AGT TAC GTC TTC GAC TAG AGT CTC CTC CTG GAC gaa cgc gc-3') generated *BssHII*-compatible ends. Oligonucleotides c-myc *AatII* upper (5'-c atg GCA TCA ATG CAG AAG CTG ATC TCA GAG GAG GAC CTG ctt gac gt-3') and c-myc *AatII* lower (5'-tg cag tac CGT AGT TAC GTC TTC GAC TAG AGT CTC CTC CTG GAC gaa c-3') generated *AatII*-compatible ends. The resulting double-stranded linkers were ligated into either the *BssHII* site (H1) or the *AatII* site (H2) of pEMC-La(MMM). This resulted in the generation of pEMC-La(Mc-mycMM) and pEMC-La(MMc-mycM).

**Insertion of EGFP into hinge 2 of MV L.** The complete EGFP ORF was amplified from pMeGFPNV (7) using GFP *AatII*F (5'-GAA TAC GAC GTC ATG GTT AGC AAG GGC-3') and GFP *AatII*R (5'-CGA TAC GAC GTC ACC CTT GTA CAG CTC-3') (restriction sites are in italics). The resulting fragment was digested with *AatII* and ligated into the similarly treated pEMC-La(MMM), resulting in the plasmid pEMC-La(MMEGFP). The inserted EGFP fragment was sequenced across the *AatII* joins to verify that the orientation and reading frame were correct.

**Assessment of L protein function.** HeLa cells grown to a confluency of 90% in 35-mm-diameter wells were infected with MVA-T7 to a multiplicity of infection (MOI) of 0.5. Transfections were carried out using Lipofectin transfection reagent (Life Technologies) as outlined by Schneider et al. (25). Plasmids pEMC-Na (1.2  $\mu\text{g}$ ) and pEMC-Pa (1.2  $\mu\text{g}$ ), the L plasmid under assessment (0.4  $\mu\text{g}$ ), and the MV mini-replicon p107MV(-)CAT (4.0  $\mu\text{g}$ ) were transfected into HeLa cells over 18 h. Chloramphenicol acetyltransferase (CAT) activity in cell lysates was assayed overnight using  $^{14}\text{C}$ -labeled chloramphenicol. Samples were monitored for the presence of acetylated products by liquid scintillation counting (28).

**Generation of c-Myc- and EGFP-tagged recombinant MVs.** The subclone pscMV(*Clal*) (6) was digested with *NcoI* and *BbrPI* to remove a fragment containing the region encoding H2 within the L gene. This was replaced with the corresponding region from pEMC-La(MMc-mycM) to generate pscMV(*Clal*-mycL). The plasmid was digested with *Clal* and the 10,624-bp fragment, now containing a c-myc tag within the L coding region, was introduced into the full-length plasmid p(+)MV to generate p(+)MVmycL. The sequence encoding the epitope tag was removed from p(+)MVmycL by digestion with *AatII*. This was replaced with the EGFP ORF, obtained by restriction digestion of pEMC-La(MMEGFP) with *AatII*, to generate p(+)MVEGFP. Rescue of recombinant MVs was carried out as described previously (6) using the unmodified MV L plasmid, pEMC-La, to support the replication and transcription of the in vitro-generated RNPs.

**Immunofluorescence, autofluorescence, and confocal laser microscopy.** Vero cells were grown on glass coverslips in 35-mm-diameter petri dishes to 80% confluency. These were either infected at an MOI of 0.01 or transfected as described above. Cells were fixed and processed as described before (7). The primary antibody, mouse anti-c-Myc monoclonal antibody (Chemicon International Inc.), was used at a dilution of 1:1,000. A polyclonal rabbit anti-CAT antibody (Sigma) was used at a dilution of 1:500. Secondary antibodies were either fluorescein isothiocyanate-conjugated rabbit anti-mouse (Sigma) at a 1:40 dilution or CY3-conjugated goat anti-rabbit (Sigma) at a dilution of 1:500. EGFP autofluorescence was examined in paraformaldehyde-fixed cells as described previously (7).

**Western blotting.** Vero cells were infected at an MOI of 0.1 and the monolayers were incubated for 3 days to allow syncytia to develop. Cells were lysed in sodium dodecyl sulfate-polyacrylamide gel electrophoresis sample buffer, and Western blots were performed by following standard procedures. EGFP was detected using an anti-EGFP rabbit immunoglobulin G fraction polyclonal antibody (Molecular Probes) at a 1:500 dilution. An anti-rabbit immunoglobulin G (Fc) alkaline phosphatase conjugate (Promega) was used as a secondary antibody at a 1:6,000 dilution. Bound antibody was detected by using the Western Blue Stabilised substrate (Promega).

## RESULTS

**Expression of L polymerases containing epitope tags.** The L gene of MV is 6,644 nucleotides long and there are a limited number of unique restriction sites that can be used to modify the gene. There are also no unique restriction sites in the

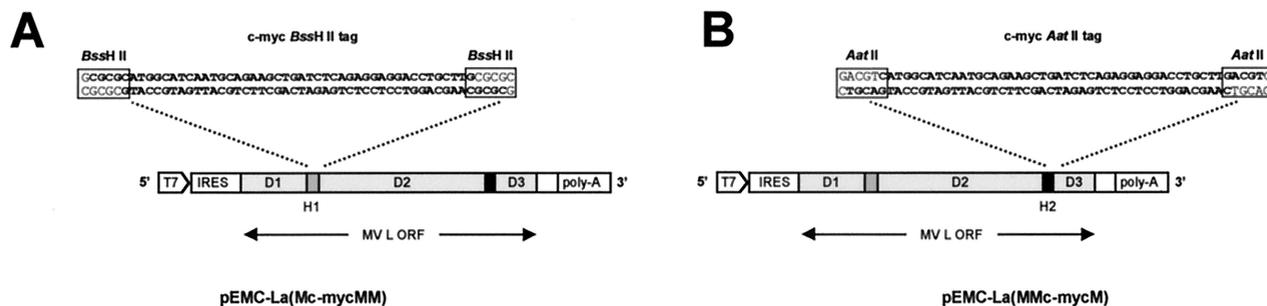


FIG. 1. Insertion of epitope tags into the hinge regions of the MV L protein. Complementary oligonucleotides encoding a c-Myc epitope (bold) were annealed to generate either *Bss*HII (A)- or *Aat*II (B)-compatible ends. The resulting fragments were inserted into the *Bss*HII site present in hinge 1 or the *Aat*II site present in hinge 2 to generate pEMC-La(Mc-mycMM) (A) and pEMC-La(MMc-mycM) (B), respectively.

variable hinge regions. To overcome this difficulty restriction sites were introduced sequentially, using a cassette vector-based strategy, into H1, H2, and the 3' end of the L gene, thereby generating the construct pEMC-La(MMM). Introducing *Bss*HII and *Eco*47III sites resulted in no changes to the L protein, whereas two amino acids in H2, L1708N→T and L1709N→S, were altered due to the introduction of the *Aat*II site. These two amino acids are present in the hinge region of CDV (17).

Epitope tags were introduced into H1 and H2 for two reasons. Firstly, they were introduced in order to investigate if the hinges are located in regions of the polymerase which are flexible and could thereby tolerate short insertions. This was of interest because the length of the morbillivirus proteins is completely conserved except for a single, terminal amino acid. Secondly, they were added to provide a means to study the cellular localization of the L protein in transfected cells. This is important as it has been difficult to generate an MV L-specific antiserum and little is known of its intracellular localization in infected cells. The c-Myc tag was chosen because it is highly immunoreactive and relatively small and thus less likely to interfere with protein function. Two constructs, pEMC-La(Mc-mycMM) and pEMC-La(MMc-mycM) (Fig. 1), were generated. The constructs were transfected into MVA-T7-infected cells and polymerase expression was observed by indirect immunofluorescence (data not shown).

**Functionality of c-Myc-tagged L polymerases.** The modified polymerase proteins were used to rescue a minigenome plasmid containing a CAT reporter gene in order to determine the effect that insertion of c-Myc epitope tags into H1 or H2 had on protein function. Unmodified pEMC-La and modified pEMC-La(MMM) were used as positive controls. Plasmid encoding L was excluded from the negative control to verify that any CAT activity obtained was solely due to its presence. Transfections were repeated three times and the results are expressed as a percentage of the counts per minute obtained using the unmodified positive control (Table 1). It was clear that introduction of the *Aat*II restriction site into pEMC-La(MMM) and the resulting amino acid alterations had no effect on the function of the protein. Presence of the c-Myc tag in H1 completely abolished function whereas presence of the same sequence in H2 had no effect on the ability of the protein to replicate and transcribe the minigenome RNP. Antisense insertion of the c-Myc epitope tag sequences into the H2 re-

gion led to the generation of pEMC-La(MMcym-cM). This construct has a premature stop codon in the L gene and the resulting polymerase protein completely lacks D3. Minigenome rescue experiments demonstrated that the truncated L protein was not functional (Table 1). These results suggest that H2 can tolerate small insertions, that D3 enjoys a degree of conformational freedom from D1 and D2, and that D3 is absolutely required to maintain L protein function.

**Insertion of EGFP into the L polymerase.** Based on the observations made with the c-Myc epitope tag a large insertion was made into H2 to further probe its plasticity to insertion. The gene encoding EGFP was amplified using oligonucleotides containing *Aat*II restriction sites. An additional glycine residue replaced the stop codon to maintain the "rule of six" (3). Plasmid pEMC-La(MMEGFPM) was generated upon insertion into the *Aat*II restriction site in H2. Autofluorescence was observed in MVA-T7-infected cells which were transfected with this construct (data not shown), indicating that the EGFP protein was expressed within cells. However, this gave no indication of the catalytic activity of the polymerase. The ability of pEMC-La(MMEGFPM) to rescue the CAT-containing minigenome construct was assayed using a range of amounts of L polymerase expressing plasmid, with pEMC-La(MMM) used as a positive control (Fig. 2). Very surprisingly, CAT activity was detected in cell lysates, albeit at lower levels (approximately 60%) than the control. This demonstrates that the normally highly conserved polymerase ORF can be extended by 11% and indicates that the third domain of the polymerase can be reoriented relative to the rest of the protein. Furthermore, it shows that this lengthening has a limited effect on the other

TABLE 1. Relative activity of MV polymerase proteins determined by production of CAT in minigenome rescue

Polymerase-expressing plasmid	% CAT activity ( $\pm$ SD) obtained in minigenome assay <sup>a</sup>
pEMC-La	100 $\pm$ 4.8
pEMC-La(MMM)	98.5 $\pm$ 2.5
pEMC-La(Mc-mycMM)	1.4 $\pm$ 0.28
pEMC-La(MMc-mycM)	98.3 $\pm$ 2.4
pEMC-La(MMcym-cM)	2.5 $\pm$ 0.37
No plasmid	1.1 $\pm$ 0.1

<sup>a</sup> 100% activity represents approximately 21,000 cpm of acetylated <sup>14</sup>C-chloramphenicol produced using one-third of the lysate obtained from each minigenome transfection experiment in this series of experiments.

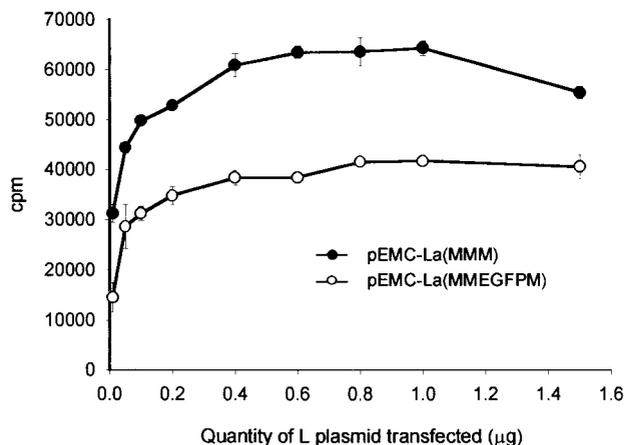


FIG. 2. Relative activities of unmodified and EGFP-modified polymerase in minigenome rescue. Various amounts of pEMC-La(MMM) (closed circles) and pEMC-La(MMEGFPM) (open circles) were used in minigenome transfection assays. The amount of acetylated  $^{14}\text{C}$ -chloramphenicol produced was determined by scintillation counting and is expressed in counts per minute.

critical protein-protein interactions essential for transcription and replication, those of L with P and of L and P with RNP. Finally, this illustrates that it is possible to modulate the activity of the RdRp of MV.

#### Generation and characterization of recombinant viruses.

Rescue of virus minigenomes and detection of levels of reporter gene expression have been useful in the understanding of viral promoters. However, in studies which seek to dissect the function of virus proteins it is usually more desirable to examine the modified protein in virus-infected cells. To this end two full-length antigenome constructs, p(+)MVC-mycL and p(+)MVEGFPL, which contained the modified polymerase genes were generated. Recombinant viruses EdtagL(MMc-mycM) and EdtagL(MMEGFPM) were recovered following transfection into HeLa cells using the unmodified MV L-expressing construct as a support plasmid. Interestingly, particularly in light of the reduced activity of the EGFP L polymerase, the viruses were no more difficult to rescue than usual (ca. 1 to 5 rescue events per primary transfection). The growth characteristics were examined by single-step growth analysis (Fig. 3), in which the titers were determined for both cell-associated and supernatant virus. It was clear that the virus with the c-Myc tag in H2 grew to an equivalent titer and at an equivalent rate to the Edtag parental virus (6, 23). From the data obtained in the rescue of the minigenome construct (Fig. 2) it could be expected that the virus containing EGFP in H2 would be growth impaired, and this is reflected in the one-step growth curve. EdtagL(MMEGFPM) titers from supernatants and cell-associated samples reached  $10^{5.5}$  and lagged behind EdtagL(MMc-mycM) by approximately 5 h.

It was important to confirm that autofluorescence in infected cells is due to the presence of EGFP within the polymerase protein. Hence, Vero cells were infected at an MOI of 0.1 with Edtag, MVEGFP, and EdtagL(MMEGFPM). Lysates of the infected cells and uninfected Vero cells were analyzed by Western blotting (Fig. 4). A 27-kDa band, corresponding to EGFP, was readily detectable in MVEGFP lysates. The high

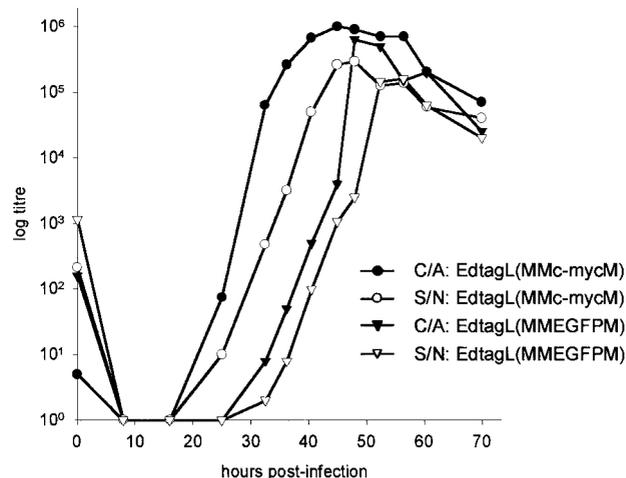


FIG. 3. Growth kinetics of recombinant MVs. Vero cells were infected at an MOI of 0.1 with EdtagL(Mc-mycMM) and EdtagL(MMc-mycM). Titers of cell-associated (C/A) and supernatant (S/N) viruses were determined in triplicate and lines join the average titers obtained.

level of expression is consistent with the promoter-proximal location of the additional EGFP transcription unit. A band of approximately 250 kDa was detected in EdtagL(MMEGFPM) lysates which corresponds to the L protein containing EGFP within the polymerase. The protein was detected at a significantly lower level, which is consistent with the position of the L gene in the virus genome. Unfortunately, due to the unavailability of an L-specific antiserum, it was not possible to dually stain the immunoblot for unmodified L to visualize the difference in the molecular weights.

**Examination of the intracellular localization of the modified polymerases.** MVA-T7-infected cells were transiently transfected with the plasmid encoding the EGFP-tagged polymerase. Unfixed cells were examined directly for EGFP autofluorescence and a high level of protein expression was detected in intracytoplasmic inclusions (data not shown).

In most cases CAT assays are performed on lysed monolayers of transfected cells. We were interested to determine if it

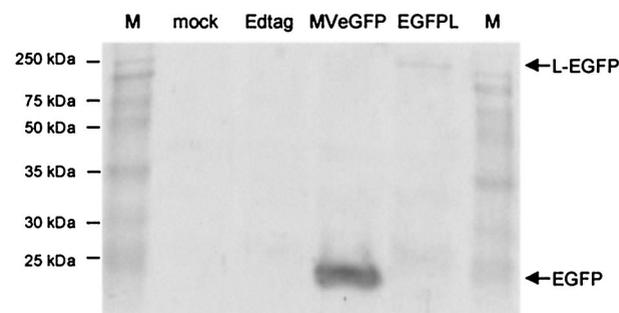


FIG. 4. Detection of EGFP in recombinant virus-infected cell lysates by Western blotting. Vero cells were infected at an MOI of 0.1 with Edtag, MVEGFP, and EdtagL(MMEGFPM). An uninfected Vero cell lysate was used as a negative control. Samples were separated on a 10% polyacrylamide gel and proteins were transferred to nitrocellulose by electroblotting. EGFP was detected using a polyclonal antiserum and broad-band molecular mass markers (M) were used to determine the approximate sizes of the proteins.

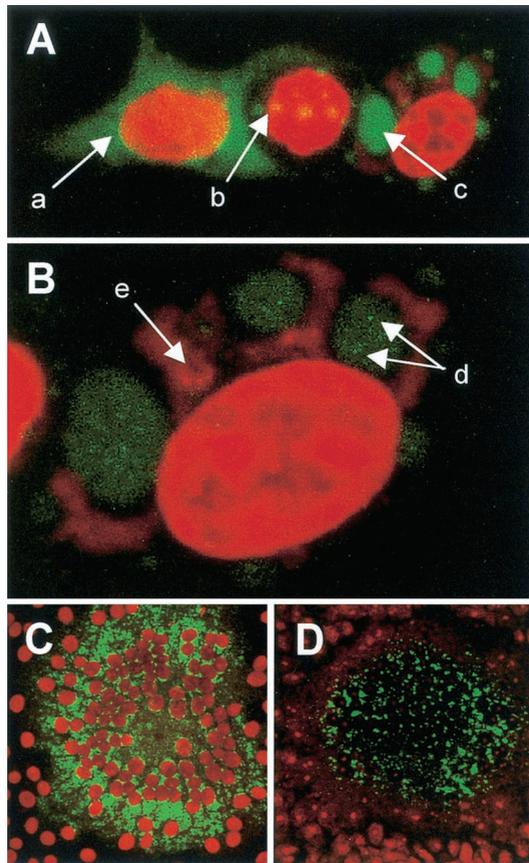


FIG. 5. Intracellular localization of epitope- and EGFP-tagged MV polymerases in transfected and infected cells. CAT and c-Myc were detected by indirect immunofluorescence using rabbit polyclonal and mouse monoclonal antisera, respectively, and the appropriate fluorescein isothiocyanate-conjugated secondary antibody. EGFP was visualized by virtue of its autofluorescence and nuclei were counterstained with propidium iodide (red). (A and B) Cells were transfected with plasmids expressing the N and P proteins of MV, pEMC-La(MMc-mycM) and p107MV(-)CAT. Epitope-tagged L protein (green) and CAT (red) were detected in the cytoplasm of the transfected cells. Magnification,  $\times 500$  (A) and  $\times 2,320$  (B). (C) Vero cells were infected at an MOI of 0.1 with EdtagL(MMc-mycM) and the epitope-tagged polymerase was detected in a syncytium (green). Magnification,  $\times 200$ . (D) Vero cells were infected at an MOI of 0.1 with EdtagL(MMEGFP) and the autofluorescent polymerase was detected in a syncytium (green). Magnification,  $\times 200$ .

was possible to detect CAT in situ combined with an assessment of polymerase localization in cells replicating a minigenome. This could be used to further the understanding of virus transcription and replication at the single-cell level. Minigenome rescue was performed in transfected HeLa cells using pEMC-La(MMc-mycM) for 48 h, after which time the cells were fixed and stained for CAT by indirect immunofluorescence and c-Myc-tagged polymerase. Nuclei were counterstained with propidium iodide. Variability was observed in the intracellular localization of the polymerase and CAT protein (Fig. 5). In a number of cells the polymerase was distributed diffusely throughout the cytoplasm and no CAT expression was detected (arrow a). An accumulation of the polymerase into intracytoplasmic inclusions was noted in some cells, and again CAT expression was not detectable (arrow b). Larger inclu-

sions (arrow c) were observed in other cells which, when examined at a higher magnification (Fig. 5B), had granular accumulations of the polymerase within the inclusion (arrows d). CAT antigen was detected in the cytoplasm of these cells in close proximity to the polymerase-containing inclusions (arrow e).

Polymerase expression levels were also examined in cells infected with the recombinant viruses. Vero cells were infected at a low MOI and syncytia of up to 100 cells formed after 2 days. Epitope-tagged polymerase was detected within intracytoplasmic inclusions in fixed EdtagL(MMc-mycM) syncytia (Fig. 5C). EGFP autofluorescence was detected directly in both unfixed (data not shown) and fixed (Fig. 5D) EdtagL(MMEGFP) syncytia. EGFP autofluorescence within the syncytium is less intense than that observed indirectly using the fluorescein isothiocyanate secondary antibody conjugate.

## DISCUSSION

One of the few features all RNA viruses, with the exception of retroviruses, have in common is that they encode RdRps which are employed in transcription and/or replication of the viral genome (31). Manipulation of the hinge regions of the RdRp of MV and the assessment of the function of the modified polymerases both in minigenome assays and within recombinant viruses give us the first insight into the structure of this relatively large protein. This study adds weight to the idea that the polymerase exists as a concatenated or multidomain protein and we believe it is reasonable to consider the EGFP  $\beta$ -barrel structure as an additional, inserted domain. It is clear that the presence of EGFP does not impinge drastically on the catalytic activities of the polymerase, which suggests that it is located in a region of the protein which is not in contact with either the P protein or the RNP. The presence of the autofluorescent reporter within the protein also permits the localization of the tagged polymerase in living cells.

It was necessary to introduce unique restriction sites into the hinge regions of the L gene and to verify that a fully functional L polymerase was generated upon reassembly. One of the ideals in any mutagenesis strategy is that the protein sequence remains unaltered; however, this can be difficult to achieve in large plasmids. Additionally, as the variable regions were identified solely by sequence alignments it was not possible to determine precisely the beginning and end of the hinges. Therefore, it was considered reasonable to introduce the restriction sites as near to the center of the predicted hinge as possible. Two amino acids were changed in H2 of the L polymerase. No significant decrease in the function of the polymerase was detected when pEMC-La(MMM) was used to rescue a minigenome containing the CAT gene. This gave the first indication that minor alterations could be engineered into a hinge region and confirmed what is observed in sequence alignments of MV strains (2, 17). In order to extend this observation epitope tags were introduced into the hinges to probe their tolerance to larger insertions. Introduction of the c-Myc epitope into H1 abolished the function of the polymerase, demonstrating that the insertion had a significant effect on the structure of the protein. This was surprising, as it is a highly variable region of the primary sequence both in terms of length and amino acid composition. Sendai virus, a *Respirovirus*, has a

hinge of 37 amino acids, whereas H1 of Newcastle disease virus, a *Rubulavirus*, is 18 amino acids in length. Two members of the more distantly related *Rhabdoviridae*, rabies virus and vesicular stomatitis virus, completely lack this variable region (21). Loss of polymerase function is most likely due to the reorientation of the critical catalytic domain (D2) in the three-dimensional structure of the protein. No decrease in polymerase function was detected when an insertion was made into H2, illustrating that the *Morbillivirus* alignments were informative in this instance. Interestingly, this region is more highly conserved between the morbilliviruses than H1 and therefore it could have been expected that insertions into this region would be more likely to abolish polymerase activity (21). The insertion of the epitope tag increases the length of H2 by 65% and the length of the polymerase by 0.68%. In a comparison of nucleotide sequences of the L proteins of ten MV strains, not a single change in H2 was observed in genotype A MV L proteins. However, this was the most variable area of the non-genotype A MV L proteins (2). Insertion of small sequences into a polymerase is not without precedent. Insertions have been made into the  $\beta$ 3- $\beta$ 4 hairpin loop of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase in order to explore the influence of the fingers subdomain in polymerase processivity. The resulting polymerases demonstrated 25 to 30% greater processivity than the wild-type enzyme (15). Very interestingly, a recent study reports the isolation of an HIV-1 mutant with an insertion in the same surface loop from a patient who had not responded to nucleoside analogue inhibitors. This indicates that modulating polymerase function has a clear role to play in adaptive evolution when strong selective pressures are present (24).

GFP is comprised of eleven  $\beta$ -sheets which form a barrel-shaped  $\beta$ -can, with a section of coaxial  $\alpha$ -helix running through the middle of the barrel (20, 30). The amino and carboxy termini protrude from the same end of the barrel and are separated by  $\sim 18.5$  Å in the crystal structure. The last nine amino acids of the carboxy terminus are not visible at atomic resolution, which shows that this region is flexible. Based on the fact that it was possible to insert a c-Myc tag into H2 we postulated that it might be feasible to insert the whole ORF of EGFP into the polymerase. Following transfection and assessment of CAT activity it was clear that the chimeric polymerase was functional, albeit at a lower level ( $\sim 60\%$  activity) than the fully functional control protein. A range of plasmid concentrations was used in CAT assays to ascertain if this decrease in polymerase activity was observed when higher amounts than normal ( $>0.4$   $\mu$ g) of the plasmids were used. It was clear that over this range the chimeric polymerase consistently retained its lower activity compared to the control. Hence, increasing the L protein concentration was not able to compensate for the diminished function. When the separate domains of MV and CDV are aligned D2 shows the highest degree of conservation (78.9%). This conservation possibly reflects its catalytic role. However, it could also give an indication that D2 may interact with other domains of the polymerase or elements of the RNP. Domain 3 is much more variable (61.6% homology) and if this hypothesis for D2 is correct it could explain the highly improbable results obtained following the insertion of EGFP into the RdRp.

Expression of proteins in vaccinia virus-infected cells and

assessment of replication and transcription in minigenome assays are important starting points in the assessment of polymerase function. It is, however, more informative to examine activity in cells infected with a recombinant virus. Surprisingly, the two corresponding viruses, EdtagL(MMc-mycM) and EdtagL(MMEGFPM), were readily rescued. Interestingly, both viruses grew to appreciable titers and, as expected, EdtagL(MMEGFPM) was slightly growth impaired (6). This demonstrates that it was feasible to use EGFP as an additional "silent" domain in studies of multidomain protein structure with the added advantage that the resulting chimeric protein is readily detected in living cells. For example, insertion of EGFP into the similarly sized agrin ( $>210$  kDa), a large multidomain glycoaminoglycan-containing glycoprotein, could help us to understand domain relationships within this modular protein (9, 11). The utility of the approach developed in this study is evidenced by the fact that the intracellular localization of the polymerase can be examined for the first time. Furthermore, Western blotting can be used to detect the chimeric protein. Many heterologous proteins have been introduced into the genomes of the *Mononegavirales* within additional transcription units, and the transgenes are stably maintained. Since homologous recombination does not occur in the *Mononegavirales* (22), the EGFP is essentially locked into the protein, as any premature stop codons or frameshift mutations would generate a nonfunctional polymerase.

Introduction of EGFP into RdRps may provide a novel means to irreversibly attenuate viruses by permanently modulating transcription and replication, and we are presently applying this observation to an in vivo model of virulence. It is clear that even small mutations in the polymerase can have profound effects on the virus phenotype. Three amino acid substitutions in the L protein are sufficient to give rise to the temperature-sensitive, cold adaptation, and attenuation phenotypes of human parainfluenza virus type 3 in hamsters (27). Recombinant respiratory syncytial viruses based on cold-adapted mutants (5) that have small numbers of changes in the polymerase protein are presently being tested as novel vaccines (14, 29).

This study has wide applications for the study of other RdRps of the *Mononegavirales* and possibly reverse transcriptases. To the best of our knowledge, it is the first report of EGFP being inserted into a protein as opposed to being fused to the amino or carboxy terminus and we believe this strategy could be used effectively in a diversity of cell biology studies. The results of this study add weight to the idea that the 250-kDa MV polymerase exists as a multidomain protein and suggest that the carboxy-terminal domain shows a degree of conformational independence. Modulating RdRp function using this approach may have significant implications in the future for the rational attenuation of virus replication.

#### ACKNOWLEDGMENTS

W.P.D. and F.M.C. contributed equally to this study.

We thank Paula Haddock for excellent technical assistance and acknowledge the help of Uta Gassen in the critical reading of the manuscript. This work was supported by the Wellcome Trust (grant 047245) of the European Social Fund.

## REFERENCES

1. Banerjee, A. K. 1987. Transcription and replication of rhabdoviruses. *Microbiol. Rev.* **51**:66–87.
2. Bankamp, B., W. J. Bellini, and P. A. Rota. 1999. Comparison of L proteins of vaccine and wild-type measles viruses. *J. Gen. Virol.* **80**:1617–1625.
3. Calain, P., and L. Roux. 1993. The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* **67**:4822–4830.
4. Chiesa, A., E. Rapizzi, V. Tosello, P. Pinton, M. de Virgilio, K. E. Fogarty, and R. Rizzuto. 2001. Recombinant aequorin and green fluorescent protein as valuable tools in the study of cell signalling. *Biochem. J.* **355**:1–12.
5. Crowe, J. E., Jr., C. Y. Firestone, S. S. Whitehead, P. L. Collins, and B. R. Murphy. 1996. Acquisition of the ts phenotype by a chemically mutagenized cold-passaged human respiratory syncytial virus vaccine candidate results from the acquisition of a single mutation in the polymerase (L) gene. *Virus Genes* **13**:269–273.
6. Duprex, W. P., I. Duffy, S. McQuaid, L. Hamill, S. L. Cosby, M. A. Billeter, J. Schneider-Schaulies, V. ter Meulen, and B. K. Rima. 1999. The H gene of rodent brain-adapted measles virus confers neurovirulence to the Edmonston vaccine strain. *J. Virol.* **73**:6916–6922.
7. Duprex, W. P., S. McQuaid, L. Hangartner, M. A. Billeter, and B. K. Rima. 1999. Observation of measles virus cell-to-cell spread in astrocytoma cells by using a green fluorescent protein-expressing recombinant virus. *J. Virol.* **73**:9568–9575.
8. Duprex, W. P., S. McQuaid, B. Roscic-Mrkic, R. Cattaneo, C. McCallister, and B. K. Rima. 2000. In vitro and in vivo infection of neural cells by a recombinant measles virus expressing enhanced green fluorescent protein. *J. Virol.* **74**:7972–7979.
9. Gesemann, M., A. J. Denzer, and M. A. Ruegg. 1995. Acetylcholine receptor-aggregating activity of agrin isoforms and mapping of the active site. *J. Cell Biol.* **128**:625–636.
10. Hercyk, N., S. M. Horikami, and S. A. Moyer. 1988. The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities. *Virology* **163**:222–225.
11. Hoch, W., J. T. Campanelli, S. Harrison, and R. H. Scheller. 1994. Structural domains of agrin required for clustering of nicotinic acetylcholine receptors. *EMBO J.* **13**:2814–2821.
12. Horikami, S. M., and S. A. Moyer. 1995. Alternative amino acids at a single site in the Sendai virus L protein produce multiple defects in RNA synthesis in vitro. *Virology* **211**:577–582.
13. Hunt, D. M., and K. L. Hutchinson. 1993. Amino acid changes in the L polymerase protein of vesicular stomatitis virus which confer aberrant polyadenylation and temperature-sensitive phenotypes. *Virology* **193**:786–793.
14. Juhasz, K., S. S. Whitehead, C. A. Boulanger, C. Y. Firestone, P. L. Collins, and B. R. Murphy. 1999. The two amino acid substitutions in the L protein of cpts530/1009, a live-attenuated respiratory syncytial virus candidate vaccine, are independent temperature-sensitive and attenuation mutations. *Vaccine* **17**:1416–1424.
15. Kew, Y., L. R. Olsen, A. J. Japour, and V. R. Prasad. 1998. Insertions into the beta3-beta4 hairpin loop of HIV-1 reverse transcriptase reveal a role for fingers subdomain in processive polymerization. *J. Biol. Chem.* **273**:7529–7537.
16. Ludin, B., and A. Matus. 1998. GFP illuminates the cytoskeleton. *Trends Cell Biol.* **8**:72–77.
17. McIlhatton, M. A., M. D. Curran, and B. K. Rima. 1997. Nucleotide sequence analysis of the large (L) genes of phocine distemper virus and canine distemper virus (corrected sequence). *J. Gen. Virol.* **78**:571–576.
18. Moeller, K., I. Duffy, W. P. Duprex, B. K. Rima, R. Beschoner, R. Meyer-mann, S. Niewiesk, V. ter Meulen, and J. Schneider-Schaulies. 2001. Recombinant measles viruses expressing altered hemagglutinin (H) genes: functional separation of mutations determining H antibody escape from neurovirulence. *J. Virol.* **75**:7612–7620.
19. Moss, B. 1991. Vaccinia virus: a tool for research and vaccine development. *Science* **252**:1662–1667.
20. Ormo, M., A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. 1996. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**:1392–1395.
21. Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo. 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J. Gen. Virol.* **71**:1153–1162.
22. Pringle, C. R. 1982. The genetics of vesiculoviruses. *Arch. Virol.* **72**:1–34.
23. Radecke, F., and M. A. Billeter. 1996. The nonstructural C protein is not essential for multiplication of Edmonston B strain measles virus in cultured cells. *Virology* **217**:418–421.
24. Sato, H., Y. Tomita, K. Ebisawa, A. Hachiya, K. Shibamura, T. Shiino, R. Yang, M. Tatsumi, K. Gushi, H. Umeyama, S. Oka, Y. Takebe, and Y. Nagai. 2001. Augmentation of human immunodeficiency virus type 1 subtype E (CRF01\_AE) multiple-drug resistance by insertion of a foreign 11-amino-acid fragment into the reverse transcriptase. *J. Virol.* **75**:5604–5613.
25. Schneider, H., P. Spielhofer, K. Kaelin, C. Dotsch, F. Radecke, G. Sutter, and M. A. Billeter. 1997. Rescue of measles virus using a replication-deficient vaccinia-T7 vector. *J. Virol. Methods* **64**:57–64.
26. Sidhu, M. S., J. Chan, K. Kaelin, P. Spielhofer, F. Radecke, H. Schneider, M. Masarekar, P. C. Dowling, M. A. Billeter, and S. A. Udem. 1995. Rescue of synthetic measles virus minireplicons: measles genomic termini direct efficient expression and propagation of a reporter gene. *Virology* **208**:800–807.
27. Skiadopoulos, M. H., A. P. Durbin, J. M. Tatem, S. L. Wu, M. Paschalis, T. Tao, P. L. Collins, and B. R. Murphy. 1998. Three amino acid substitutions in the L protein of the human parainfluenza virus type 3 cp45 live attenuated vaccine candidate contribute to its temperature-sensitive and attenuation phenotypes. *J. Virol.* **72**:1762–1768.
28. Sleight, M. J. 1986. A nonchromatographic assay for expression of the chloramphenicol acetyltransferase gene in eukaryotic cells. *Anal. Biochem.* **156**:251–256.
29. Whitehead, S. S., C. Y. Firestone, R. A. Karron, J. E. J. Crowe, W. R. Elkins, P. L. Collins, and B. R. Murphy. 1999. Addition of a missense mutation present in the L gene of respiratory syncytial virus (RSV) cpts530/1030 to RSV vaccine candidate cpts248/404 increases its attenuation and temperature sensitivity. *J. Virol.* **73**:871–877.
30. Yang, F., L. G. Moss, and G. N. Phillips, Jr. 1996. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**:1246–1251.
31. Zanotto, P. M., M. J. Gibbs, E. A. Gould, and E. C. Holmes. 1996. A re-evaluation of the higher taxonomy of viruses based on RNA polymerases. *J. Virol.* **70**:6083–6096.
32. Zink, D., and T. Cremer. 1998. Cell nucleus: chromosome dynamics in nuclei of living cells. *Curr. Biol.* **8**:R321–R324.