Short Communication

Insertion of EGFP into the replicase gene of Semliki Forest virus results in a novel, genetically stable marker virus

Nele Tamberg, Valeria Lulla, Renos Fragkoudis, Aleksei Lulla, John K. Fazakerley and Andres Merits

1Estonian Biocentre, Tartu, Estonia
2Centre for Infectious Diseases, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, UK
3Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

Correspondence
Andres Merits
andres.merits@ut.ee

Received 31 July 2006
Accepted 10 December 2006

Alphavirus-based vector and replicon systems have been extensively used experimentally and are likely to be used in human and animal medicine. Whilst marker genes can be inserted easily under the control of a duplicated subgenomic promoter, these constructs are often genetically unstable. Here, a novel alphavirus construct is described in which an enhanced green fluorescent protein (EGFP) marker gene is inserted into the virus replicase open reading frame between nsP3 and nsP4, flanked by nsP2 protease-recognition sites. This construct has correct processing of the replicase polyprotein, produces viable virus and expresses detectable EGFP fluorescence upon infection of cultured cells and cells of the mouse brain. In comparison to parental virus, the marker virus has an approximately 1 h delay in virus RNA and infectious virus production. Passage of the marker virus in vitro and in vivo demonstrates good genetic stability. Insertion of different markers into this novel construct has potential for various applications.

Infectious cDNA (icDNA) clones and corresponding expression vectors have been developed for several alphaviruses (Liljestrom & Garoff, 1991; Strauss & Strauss, 1994; Simpson et al., 1996). In most alphavirus-based expression vectors, the viral structural genes are replaced by foreign genes. These vectors are capable of only one round of infection and are thus non-mobilizable, a property that has severely hampered their use in studies that require a spreading, propagating infectious process. Marker and other recombinant alphaviruses have been produced by expressing the foreign gene from an internal ribosomal entry site element or from a duplicated 26S promoter (Raju & Huang, 1991; Hahn et al., 1992; Pugachev et al., 2000; Vaha-Koskela et al., 2003). Double-subgenomic vectors containing marker genes allow direct observation of infection and have been used successfully in animal experiments (Levine et al., 1996; Cook & Griffin, 2003). Unfortunately, these vectors tend to suffer from genome instability, probably because the inserted genes are introduced as separate transcription units and have no selective value (Pugachev et al., 1995, 2000).

An alternative strategy is to insert genes into natural gene expression units. A marker gene has been inserted successfully into the Sindbis virus (SIN) structural gene (Thomas et al., 2003). SIN genomes with markers inserted into the non-structural replicase to produce nsP3-fusion proteins have also been produced (Bick et al., 2003; Frolova et al., 2006). Production of non-structural–marker fusion proteins has also been successful in other viruses, including Poliovirus, Hepatitis C virus, Equine arteritis virus and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.

An SFV nsP3–EGFP fusion protein virus was engineered by subcloning and PCR-based mutagenesis. The C residue at position 5447 of pSFV4 (Liljestrom et al., 1991) was changed to G and a 12 bp long sequence, GGGCCCATAGGATCC, was inserted after the modified codon. The resulting construct was designated pSFV(3F)4 and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.

An alternative strategy is to insert genes into natural gene expression units. A marker gene has been inserted successfully into the Sindbis virus (SIN) structural gene (Thomas et al., 2003). SIN genomes with markers inserted into the non-structural replicase to produce nsP3-fusion proteins have also been produced (Bick et al., 2003; Frolova et al., 2006). Production of non-structural–marker fusion proteins has also been successful in other viruses, including Poliovirus, Hepatitis C virus, Equine arteritis virus and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.

An alternative strategy is to insert genes into natural gene expression units. A marker gene has been inserted successfully into the Sindbis virus (SIN) structural gene (Thomas et al., 2003). SIN genomes with markers inserted into the non-structural replicase to produce nsP3-fusion proteins have also been produced (Bick et al., 2003; Frolova et al., 2006). Production of non-structural–marker fusion proteins has also been successful in other viruses, including Poliovirus, Hepatitis C virus, Equine arteritis virus and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.

An alternative strategy is to insert genes into natural gene expression units. A marker gene has been inserted successfully into the Sindbis virus (SIN) structural gene (Thomas et al., 2003). SIN genomes with markers inserted into the non-structural replicase to produce nsP3-fusion proteins have also been produced (Bick et al., 2003; Frolova et al., 2006). Production of non-structural–marker fusion proteins has also been successful in other viruses, including Poliovirus, Hepatitis C virus, Equine arteritis virus and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.

An alternative strategy is to insert genes into natural gene expression units. A marker gene has been inserted successfully into the Sindbis virus (SIN) structural gene (Thomas et al., 2003). SIN genomes with markers inserted into the non-structural replicase to produce nsP3-fusion proteins have also been produced (Bick et al., 2003; Frolova et al., 2006). Production of non-structural–marker fusion proteins has also been successful in other viruses, including Poliovirus, Hepatitis C virus, Equine arteritis virus and several filamentous plant viruses (Mueller & Wimmer, 1998; Moradpour et al., 2004; Rajamaki et al., 2005; van den Born et al., 2005). Here, we report that Semliki Forest virus (SFV) nsP3–enhanced green fluorescent protein (EGFP) fusion protein marker virus is viable, but genetically unstable; however, by the novel strategy of placing the EGFP insert between nsP3 and nsP4 flanked by duplications of the nsP3/4 nsP2 protease-recognition site, we have generated a viable, genetically stable virus with only minor changes in phenotype.
transcripts as described by Liljestrom et al. (1991). The primary virus stocks (P1) were collected after 24 h, titrated and used to infect fresh BHK-21 cells (m.o.i., 0.1). The second-passage stocks (P2) were also harvested at 24 h. SFV(3F)4-EGFP virus was viable and expressed the expected nsP3–EGFP fusion protein (Fig. 2a). However, plaque purification followed by analysis of the fluorescence produced by individual plaques and Western blotting demonstrated that this virus was genetically unstable; a truncated form of the nsP3–EGFP fusion protein was present even in cells transfected directly with in vitro-synthesized RNA. The genetic instability of this virus is most probably due to defect(s) in the formation or functioning of replication complexes. As with SIN nsP3-fusion
designated pSFV(3H)4-EGFP and SFV(3H)4-EGFP, pSFV(3F)4-EGFP. The resulting clone and virus were created by PCR amplification and cloned into terminus of nsP3 with indicated downstream sequence (Fig. 1a). The region encoding the C stabilizing amino acid (Lulla et al., 2001). The first amino acid residue after the inserted Gly was followed by aa 2–7 from the N position and Gly represents a 9 cleavage point was also changed from Tyr to Gly; nsP2 protease-cleavage sites. Starting with pSFV(3F)4-EGFP, the C EGFP was placed between nsP3 and nsP4 flanked by nsP2 and nsP4 or EGFP are indicated with a solid arrow; bands corresponding to nsP3–EGFP and EGFP–nsP4 fusion proteins are marked with an open arrow. (b) Proteins synthesized in infected cells were labelled metabolically at 3 h post-infection and chased for 1, 3, 8 or 24 h. Cell lysates were immunoprecipitated with anti-nsP1, anti-nsP3 or anti-EGFP antibodies and analysed by SDS-PAGE.

Infectious-centre assays (Gorchakov et al., 2004) demonstrated that 3.1 × 10^5 to 4.4 × 10^5 plaques (μg transfected RNA)^−1 were obtained for both pSFV(3H)4-EGFP and pSFV4 transcripts, with no statistically significant difference between them. In one-step growth studies (m.o.i., 20), relative to SFV4, production of infectious SFV(3H)4-EGFP was delayed by approximately 1 h (Fig. 1b). However, as SFV(3H)4-EGFP was able to replicate to 10^9 p.f.u. ml^−1 within 10 h, it can be concluded that insertion of EGFP between nsP3 and nsP4 did not affect virus replication substantially.

To determine whether the EGFP insertion in SFV(3H)4-EGFP was inherited stably, 96 virus plaques were purified from each of five in vitro passages (P1–P5) in BHK-21 cells (m.o.i., 0.1) and each of five in vivo passages (M1–M5) in mouse brains. After each passage, the percentage of viruses expressing EGFP was assessed by random selection of 96 plaques followed by determination of EGFP expression. All plaques from the P2 stock of SFV(3H)4-EGFP were EGFP-positive and >90 % were EGFP-positive after the fifth passage. Even greater stability was observed for in vivo-propagated stocks: after the fifth passage, only one plaque (1/96) was EGFP-negative. As an EGFP-negative phenotype could result from EGFP inactivation by point mutation or by deletion, we RT-PCR-amplified and sequenced the corresponding regions for several of the EGFP-negative (in vitro-passaged) plaque-purified viruses. In all cases, these genomes contained large in-frame deletions, indicating that deletions in the marker gene do occur upon passage of SFV(3H)4-EGFP. Whilst the frequency of this process may be low, given their growth advantage, deleted genomes are likely to increase in the viral population following multiple passages.

In SFV-infected cells, nsP3 is known to associate with modified endosomes and lysosomes (Froshauer et al., 1988). In contrast, free EGFP is typically localized diffusely in the cytoplasm and more abundantly in the nucleus. In SFV(3F)4-EGFP-infected BHK-21 cells, EGFP co-localized with nsP3 to punctate cytoplasmic structures, presumably virus replicase complexes (Fig. 1c). In contrast, SFV(3H)4-EGFP-infected cells showed granular cytoplasmic staining for nsP3 and diffuse, predominantly nuclear staining for EGFP (Fig. 1d). The independent localization of nsP3 and EGFP in the SFV(3H)4-EGFP-infected cells indicates that EGFP was released from the replication complexes.

To determine the phenotype of SFV(3H)4-EGFP in vivo, groups of six 5–6-week-old female BALB/c mice were inoculated (20 μl) intracerebrally with SFV4(3H)-EGFP or SFV4. All animal experiments were carried out under the authority of a UK Home Office licence and were approved by the University of Edinburgh ethical-review process. Mice in both groups had clinical signs of encephalitis at day 2 and were sampled. Half brains were fixed in 4 % neutral-buffered formalin for 16 h, processed through graded sucrose solutions, frozen in OCT and cut into sections (12 μm). EGFP-positive cells were observed readily in all (n=6) brains infected with SFV(3H)4-EGFP (Fig. 1e).
Expression of the ns proteins in SFV(3H)4-EGFP- and SFV(3F)4-EGFP-infected BHK-21 cells was examined by Western blotting (Fig. 2a). At 6 h post-infection, cells were lysed in Laemmli buffer and samples corresponding to 100 000 cells were separated by SDS-PAGE. Monospecific rabbit polyclonal antibodies were used to detect nsP1, nsP3 and nsP4, and a mouse monoclonal antibody was used to detect nsP2 (Peränen et al., 1988; Laakkonen et al., 1994; Rikkonen et al., 1994; Kujala et al., 1997). Individual nsP1, nsP2, nsP3, nsP4 and EGFP were observed readily in SFV(3H)4-EGFP-infected cells. No nsP3–EGFP fusion protein was detected, but a small amount of unprocessed EGFP–nsP4 was detected by antibodies to both nsP4 and EGFP. In SFV(3F)4-EGFP-infected cells, nsP3 was, as expected, present predominantly as an nsP3–EGFP fusion protein. A small amount of free nsP3 was also present; this may have resulted from the rapid generation of EGFP-negative genomes.

To study the dynamics of cleavage of the SFV(3H)4-EGFP polyprotein, infected cells (m.o.i., 100) were pulse labelled metabolically for 15 min with 50 μCi (1.85 MBq) [35S]methionine and [35S]cysteine, and nsP3 and its precursors were immunoprecipitated with antiserum against nsP3 and Protein A–Sepharose. Immunocomplexes were separated by SDS-PAGE and radiolabelled proteins were visualized by autoradiography. In pulse-labelled cells, the P123–EGFP, P3–EGFP–4 and P3–EGFP processing intermediates were all detected, whereas in chased samples, processing intermediates were not detected or were present in lower quantities, and the amount of fully processed nsP3 was increased (data not shown). Thus, all expected processing products, except the very rapidly processed P123–EGFP–4, were detected and no apparent disturbance of processing was observed.

In BHK-21 cells, EGFP is generally quite stable, with an estimated half-life of 24 h. However, in BHK-21 cells infected with SFV(3H)4-EGFP, the EGFP signal detected late in infection by fluorescence microscopy had decreased considerably, suggesting low EGFP stability. To assess the stability of the EGFP and ns proteins, BHK-21 cells infected with SFV(3H)4-EGFP or SFV4 were labelled metabolically, chased and EGFP, nsP1 and nsP3 were immunoprecipitated. For both viruses, only a small decrease in nsP1 and nsP3 was detected over 24 h (Fig. 2b). In contrast, by 24 h, the amount of EGFP was below the detection limit of this analysis (Fig. 2b). Thus, the low amounts of EGFP observed by microscopy indeed represent rapid degradation of EGFP. The processed EGFP has a Gly residue at its N terminus that, according to the N-end rule, is stabilizing. Most likely, EGFP instability results from the duplicated C-terminal sequence of nsP3 (Fig. 1a), which, as a result of the construction strategy, remains attached to the C terminus of EGFP. If this is the case, then the effect of this same sequence in its native location (C terminus of nsP3) is either different or is suppressed by interaction(s) with other components of the replicase complex.

Virus RNA synthesis by SFV4 and SFV(3H)4-EGFP was compared by [3H]uridine labelling. Relative to SFV4-infected

---

**Fig. 3.** (a) Synthesis of virus RNA. BHK-21 cells were infected with SFV4 or SFV(3H)4-EGFP; at 1 h post-infection, actinomycin D (2 μg ml⁻¹) and, at 2 h, 25 μCi (0.925 MBq) [3H]uridine, were added. Cells were collected at 3, 4, 5, 6, 7 and 8 h post-infection. Total RNA was precipitated with trichloroacetic acid. Incorporated radioactivity was measured by liquid scintillation. Each bar represents the mean of six replicates; error bars indicate SD. (b) Shutdown of cellular translation in BHK-21 cells infected with SFV4 or SFV(3H)4-EGFP. Proteins synthesized in infected cells were labelled metabolically at 1 h intervals and cell lysates were analysed by SDS-PAGE. M, Mock-infected cells.
cells, SFV(3H)4-EGFP-infected cells had a delay in virus RNA synthesis (Fig. 3a); this was most obvious at 3–5 h post-infection and less so at later time points. To analyse whether insertion of the marker gene altered the temporal expression of SFV structural proteins or the shutdown of host-cell translation, protein synthesis in infected BHK-21 cells was studied by metabolic labelling. As observed with virus growth (Fig. 1b) and RNA synthesis (Fig. 3a), production of viral proteins C, E1 and p62 in SFV(3H)4-EGFP-infected cells started approximately 1 h later than in SFV4-infected cells. A longer, approximately 2 h, delay was observed for host-cell translational shutdown (Fig. 3b).

In all measurements of virus replication and growth, SFV(3H)4-EGFP was slightly slower than SFV4. The possible reasons for this include effects resulting from larger ns polyprotein sizes and delays or defects in replicase-complex formation. However, these changes are unlikely to preclude use of this virus, as it replicates to high titres and remains neurovirulent. Despite the small duplication of viral sequence necessitated by the addition of the nsP2 processing site upstream of EGFP, SFV(3H)4-EGFP demonstrated remarkably improved genetic stability relative to SFV(3F)4-EGFP. However, it should be noted that, at late passages in vitro, some genomes with a deletion in the marker gene were present. Attempts to reduce this by removing the duplicated 6 aa from the N terminus of EGFP, or by decreasing the length of the nsP3 C-terminal fragment fused to the C terminus of EGFP, resulted in viable viruses, but did not increase their genetic stability (data not shown). Importantly, following replication and spread of SFV(3H)4-EGFP in the mouse brain, no deletion variants were observed until passage M5, demonstrating the utility of this virus for in vivo pathogenesis studies.

Acknowledgements

We are grateful to the research group of Leevi Kääriäinen for their help at the start of this project and to Eva Zubinaite and Kaja Kiiuer for their experimental help. This research was supported by grants from the European Union SVectors programme, grant 067575 from the Wellcome Trust and grants 5055 and 6609 from the Estonian Science Foundation.

References


