

Revised manuscript for publication in the *Journal of General Virology*

Rapid identification of coronavirus replicase inhibitors using a selectable replicon RNA

Tobias Hertzog¹, Elke Scandella², Barbara Schelle¹, John Ziebuhr¹, Stuart G.
Siddell³, Burkhard Ludewig², and Volker Thiel^{2*}

¹Institute of Virology and Immunology, University of Würzburg, Würzburg, Germany,

²Research Department, Cantonal Hospital St.Gallen, St.Gallen, Switzerland

³Department of Pathology and Microbiology, School of Medical Sciences, University of Bristol,
Bristol, United Kingdom.

Running title: Coronavirus replicon RNA

***Corresponding author:**

Volker Thiel; Cantonal Hospital St.Gallen, Research Department, 9007 St.Gallen, Switzerland.

Phone: +41-71-4941074; fax: +41-71-4946321; e-mail: volker.thiel@kssg.ch

Manuscript information: 24 pages (excluding figures and table), 3 figures, 1 table, 150 words
in the abstract, 4303 words in the text.

Summary

A previously unknown coronavirus (CoV) is the etiologic agent causing severe acute respiratory syndrome (SARS), for which an effective antiviral treatment is urgently needed. To enable the rapid and biosafe identification of coronavirus replicase inhibitors, we have generated a non-cytopathic, selectable replicon RNA (based on human CoV 229E) that can be stably maintained in eukaryotic cells. Most importantly, the replicon RNA mediates reporter gene expression as a marker for coronavirus replication. We have used a replicon RNA-containing cell line to test the inhibitory effect of several compounds that are currently being assessed for SARS treatment. Amongst those, interferon- ζ displayed the strongest inhibitory activity. Our results demonstrate that coronavirus replicon cell lines provide a versatile and safe assay for the identification of coronavirus replicase inhibitors. Once this technology is adapted to SARS-CoV replicon RNAs, it will allow high throughput screening for SARS-CoV replicase inhibitors without the need to grow infectious SARS-CoV.

Introduction

Coronaviruses are enveloped, vertebrate viruses that are associated mainly with respiratory and enteric diseases in humans, livestock and companion animals (Ziebuhr & Siddell, 2002). Human coronaviruses (HCoV) are one of the causative agents of common cold. HCoV infections are generally mild, last only a few days and are rarely associated with severe symptoms such as headache, fever or diarrhea. However, it is now clear that SARS is caused by a previously unknown coronavirus, termed SARS coronavirus (SARS-CoV) (Drosten et al., 2003, Fouchier et al., 2003, Ksiazek et al., 2003, Kuiken et al., 2003, Marra et al., 2003, Peiris et al., 2003, Rota et al., 2003). SARS-CoV emerged in late 2002 and spread within a few months from its likely origin in Guangdong Province, China, to more than 30 countries. The rapid transmission and high mortality rate made SARS a global threat for which no efficacious therapy is available (Donnelly et al., 2003). Most likely, this newly recognized pathogen has crossed the species barrier from small animals, such as masked palm civets, to humans (Guan et al., 2003, Martina et al., 2003). Given the presence of an animal reservoir, and the recent re-emergence of SARS in 2004 in Guangzhou (Parry, 2004), China, the need for efficacious strategies to prevent and control coronavirus infections is evermore urgent.

Although the knowledge of coronavirus replication at the molecular level is still rudimentary, the general principles of coronavirus replication have been elucidated (Lai & Holmes, 2001, Sawicki & Sawicki, 1998, Thiel et al., 2003a, Ziebuhr et al., 2000). Coronavirus gene expression starts with the translation of the replicase gene from the infectious genomic RNA. The replicase gene is comprised of two large open reading frames (ORFs), designated ORF1a and ORF1b, that are located at the 5' end, and extend over more than two-thirds, of the genome. The upstream ORF1a encodes a polyprotein of 450-500 kDa, termed polyprotein (pp) 1a, whereas ORF1a and ORF1b together encode pp1ab (750-800 kDa), which is synthesized by a

(-1) ribosomal frameshift during translation. Virally encoded proteinases, namely papain-like and 3C-like proteinases, process the coronavirus polyproteins pp1a and pp1ab to generate at least 15 – 16 end-products and an unknown number of intermediate products. These non-structural proteins (nsp) assemble to form the functional replication/transcription complex in the cytoplasm of the infected cell. The replication/transcription complex mediates the replication of the coronavirus genomic RNA and the transcription of multiple subgenomic mRNAs. Although it has been shown that the replicase gene products suffice for coronavirus transcription (Thiel et al., 2001b), it is still a matter of debate whether the replicase gene products may also suffice for coronavirus genome replication.

A recent analysis of coronavirus replicase genes by comparative genomics revealed a new view of the (putative) enzymatic functions encoded by coronaviruses and identified a number of conserved replicase domains that remain uncharacterized (Snijder et al., 2003). Thus, the list of potential targets for antiviral intervention, such as the viral RNA-dependent RNA polymerase, helicase and proteases, has been extended to newly predicted enzymes that might be associated with RNA processing and metabolism. However, since the function of most coronavirus replicase gene products is still unknown, the identification of antiviral drugs that interfere with coronavirus replication will, most probably, be dependent on random screening of compound libraries. In contrast, the rational design of antivirals may be limited to few well characterized target enzymes (Anand et al., 2002, Anand et al., 2003, Davidson & Siddell, 2003, Tanner et al., 2003, Thiel et al., 2003a, Yang et al., 2003).

The recent development of reverse genetic systems for coronaviruses (Almazan et al., 2000, Casais et al., 2001, Thiel et al., 2001a, Yount et al., 2000, Yount et al., 2003, Yount et al., 2002) provides an opportunity to carry out an extensive biological characterization of the viral replicative proteins and functions. Hopefully, these analyses will elucidate the unique features of coronavirus replication at the molecular level. In the long term, these studies may lead also to the

development of new strategies to prevent and control coronavirus infection. In the short term, however, the reverse genetic systems can be used to facilitate the identification of coronavirus replicase inhibitors. For example, the concept of using autonomously replicating RNAs (replicon RNAs) has been explored in a number of positive strand RNA virus systems and has led to the establishment novel antiviral screening assays (Bartenschlager, 2002, Frolov et al., 1996, Khromykh, 2000, Lo et al., 2003, Randall & Rice, 2001). Stable cell lines containing non-cytopathic, selectable replicon RNAs are currently used to assess the efficacy of candidate inhibitors of viruses that can not be propagated efficiently in tissue culture, such as hepatitis C virus (Bartenschlager, 2002, Randall & Rice, 2001). Moreover, since no structural genes and, therefore, no infectious viruses are formed, replicon-based assays represent an attractive tool for the identification of antivirals if the pathogenicity of the virus is a concern.

We report here the generation and analysis of the first selectable, coronavirus-based replicon RNA. We have introduced the gene for a selectable marker into the HCoV 229E replicase gene and we have replaced three structural genes, namely the S (spike), E (envelope) and M (membrane) genes, and two genes of unknown function (genes 4a and 4b), by a reporter gene encoding green fluorescent protein (GFP). Our data indicate that this approach facilitates the rapid and biosafe evaluation of candidate coronavirus replicase inhibitors.

Methods

Cells and viruses. MRC-5, CV-1 and BHK-21 were purchased from the European Collection of Cell Cultures, D980R cells were a kind gift from G. L. Smith, Imperial College, London, UK. All cells were maintained in minimum essential medium (MEM) supplemented with HEPES (25mM), foetal bovine serum (5-10%) and antibiotics unless otherwise indicated. HCoV 229E and recombinant vaccinia virus (VV) were propagated, titrated and purified as described (Thiel et al., 2001a).

Cloning of plasmid DNAs and recombinant viruses.

Recombinant VV vHCoV-inf-1 containing the full-length cDNA of HCoV 229E (Thiel et al., 2001a) was used as a basis to construct the recombinant VV vHCoV-Rep-1 cDNA encoding the HCoV 229E-based replicon RNA. The construction of vHCoV-Rep-1 involved several steps of VV-mediated recombination using the E.coli guanine phosphoseryltransferase (*gpt*) as a marker for positive or negative selection (Isaacs et al., 1990, Kerr & Smith, 1991).

To perform VV-mediated homologous recombination, the plasmid DNAs described below were constructed using standard procedures. The precise details, plasmid maps and sequences are available from the authors upon request.

To construct the plasmid pGPT-1 a 2.1 kb *HpaI* - *SphI* fragment, derived from pTM-3 (Elroy-Stein et al., 1989) and encoding the *gpt* gene downstream of a VV promoter, was inserted into the *SmaI* restriction site of pBluescriptII KS+ (Stratagene). Plasmid pRec-1 is based on pBR322 (New England Biolabs) and a DNA insert has been cloned between the *EcoRI* – *EagI* sites of pBR322 comprised of the HCoV 229E-derived nucleotides 19601 - 21145, a 2.1 kb *EcoRI* – *XbaI* fragment derived from pGPT-1 (containing the *gpt* gene) and HCoV 229E-derived nucleotides 24201 - 25874. pVec-GN is based on pBR322 and a DNA insert has been cloned at

the BamHI site of pBR322 comprised of the HCoV 229E-derived nucleotides 19485 - 20568, one C nucleotide, the GFP gene and the HCoV 229E-derived nucleotides 25654 - 27273. pRec-2 is based on pGPT-1 and contains a 500 bp HindIII - NotI fragment derived from VV vNotI/tk (Merchinsky & Moss, 1992), the 2.1 kb fragment derived from pTM-3 and HCoV 229E-derived nucleotides 2177 -3323. pTNI is based on pRec-2 and contains the 500 bp HindIII - NotI fragment derived from VV vNotI/tk, the bacteriophage T7 RNA polymerase promoter, one G nucleotide, the HCoV 229E-derived nucleotides 1 - 615, the sequence encoding the TaV 2A-like element (Donnelly et al., 2001), the neo gene derived from pTET/ON (Clontech), the EMCV-IRES element derived from pTM-3 and the HCoV 229E-derived nucleotides 626 - 3323.

VV-mediated homologous recombination was done as follows. 5×10^5 CV-1 cells were infected (moi=1) with the respective recombinant VV indicated below followed by transfection of 5 σ g plasmid DNA one hour post infection using Lipofectin transfection reagent (Invitrogen). Two days later, the cells were harvested and recombinant VV were isolated by three rounds of plaque purification under gpt-positive or negative selection as appropriate. In order to isolate gpt-containing VV, gpt-positive selection was done on CV-1 cells in the presence of mycophenolic acid (25 σ g/ml, Sigma), xanthine (250 σ g/ml, Calbiochem) and hypoxanthine (15 σ g/ml, Calbiochem). Gpt-negative selection was done on D980R cells in the presence of 6-thioguanine (1 σ g/ml, Sigma). The identity of all recombinant VV were confirmed by southern blot and sequencing analysis.

The insertion of the GFP gene into the HCoV 229E-derived cDNA of vHCoV-inf-1 was done in two steps as follows. First, recombinant VV vRec-1 was isolated after homologous recombination of vHCoV-inf-1 with plasmid DNA pRec-1 using gpt-positive selection. Second, recombinant VV vVec-GN was isolated after homologous recombination of vRec-1 with plasmid DNA pVec-GN using gpt-negative selection. As a result of this cloning procedure the HCoV 229E genes S, 4a, 4b, E and M were replaced by the GFP gene.

The insertion of the sequence encoding the TaV 2A-like element, the neo gene and the EMCV-IRES element at the sequence encoding the HCoV 229E nsp1/nsp2 polyprotein cleavage site was done in two steps as follows. First, recombinant VV vRec-2-GN was isolated after homologous recombination of vVec-GN with plasmid DNA pRec-2 using gpt-positive selection. Second, recombinant VV vRep-1 was isolated after homologous recombination of vRec-2-GN with plasmid DNA pTNI using gpt-negative selection. The recombinant VV vRep-1 represents the source of template DNA used to generate the replicon RNA Rep-1 (figure 1).

In vitro transcription, electroporation and generation of a stable cell line containing Rep-1 RNA. Rep-1 RNA was prepared by in vitro transcription using purified genomic DNA from vRep-1 and 10 σ g of Rep-1 RNA was used for electroporation of 1×10^7 BHK-21 cells as described (Thiel et al., 2001a, Thiel et al., 2001b). After two days, G418 (400-800 σ g/ml) was added to the tissue culture medium and two weeks later individual cell clones were isolated and expanded until analysis or storage at liquid nitrogen.

Analysis of replicon RNA-containing BHK cells. Individual cell clones obtained after G418 selection of Rep-1 RNA-transfected BHK-21 cells were analysed by fluorescence microscopy using a Leica DM R fluorescence microscope and Leica IM 1000 software (Leica, Heerbrugg, Switzerland). Flow cytometry analysis was done using a FACSCalibur and CellQuest software (BD Pharmingen).

Poly(A)-containing RNA from BHK-Rep-1 cells or HCoV 229E-infected MRC-5 cells was isolated using oligo(dT)₂₅ Dynabeads (Dyna) (Thiel et al., 1997). Northern blot analysis of poly(A)-containing RNA involved electrophoresis and transfer to nylon membranes as described (Thiel et al., 2003a). To detect RNAs derived from HCoV 229E and Rep-1 a ³²P-labelled probe corresponding to the HCoV 229E nucleotides 26297 - 27273 was produced using the Multiprime

DNA-labelling system (Amersham). A complete sequencing analysis of re-isolated Rep-1 RNA from BHK-Rep-1 cells was done by standard RT-PCR sequencing methods (Thiel et al., 2003a).

Inhibition of coronavirus replication. BHK-Rep-1 cells were seeded out in 96-well plates at a concentration of 5×10^3 per well in MEM medium containing 10% FCS and 400 σ g/ml G418. 24 hours later, the medium was replaced by fresh medium without G418 and graded doses of IFN- ζ , ribavirin or glycyrrhizin (all purchased from Sigma) were added. As the parental BHK-21 cells are derived from hamster kidneys an IFN- ζ hybrid constructed from recombinant human interferons ζ A and ζ D that crosses the species barrier was used for the assay (Sigma, product No. I 4401). An untreated well served as control. Tissue culture media and inhibitors were replenished after 24 h and 48 h, and the cells were harvested and analyzed by flow cytometry on day three. Inhibition of reporter gene expression was calculated as reduction of the mean fluorescence intensity (MFI) of GFP-positive cells by setting the MFI of untreated BHK-Rep-1 cells as 0% inhibition and the complete down regulation of GFP to background levels as 100% inhibition. Cells that were subjected to fluorescence microscopy were harvested on day two, transferred to glass slides (Falcon) and cultivated for another 24 hours.

Cytotoxicity of candidate inhibitors was assayed using BHK-21 cells that were grown in MEM medium containing 10% FCS. Cell proliferation was analysed 2 and 3 days after treatment with inhibitors using the CellTiter 96 AQueous One Kit (Promega).

Results

Selectable coronavirus replicon RNAs. Our initial attempts to generate a selectable coronavirus replicon RNA were based on the assumption that, as for many positive strand RNA viruses, only the replicase gene and the 5'- and 3'-genomic termini are needed for autonomous replication of the viral RNA. Although we could demonstrate that the replicase gene products suffice for coronavirus transcription (Thiel et al., 2001b), we repeatedly failed to detect efficient replication using various RNA constructs. Indeed, during these studies (Thiel et al., 2001b, Thiel et al., 2003b), we found that the nucleocapsid (N) protein may be required for efficient coronavirus genome replication. Therefore, this observation has been included in our strategy to establish selectable replicon RNAs. In order to provide a selectable marker gene, we chose to use the neomycin resistance (neo) gene, which has been proven as a functional selection marker in other positive strand RNA virus replicons. Since we aimed to insert the neo gene into the replicase-coding region, we first defined a convenient integration site. Using a reverse genetic approach we have been able to rescue two recombinant human coronaviruses that contained the coding sequence of a '2A-like' autoprocessing peptide derived from *Thosea asigna* virus (TaV-2A) (Donnelly et al., 2001), or the internal ribosomal entry site derived from Encephalomyocarditis virus (EMCV-IRES) (data not shown). Both elements have been introduced at the sequence encoding the aminoterminal polyprotein cleavage site nsp1/nsp2 within the replicase polyproteins pp1a and pp1ab and in both constructs three carboxyterminal glycine residues of nsp1 were deleted. Since both recombinant viruses replicated to wild-type titres, we concluded that insertion of heterologous sequences at this site did not affect virus replication in tissue culture. Therefore, the nsp1/nsp2 junction of the HCoV 229E replicase polyproteins has been chosen to insert a "selection cassette" comprised of the TaV-2A element,

the neo gene and the EMCV-IRES element. The overall structure of the human coronavirus replicon RNA is illustrated in figure 1.

Generation and analysis of a stable cell line containing HCoV replicon RNA. To clone a cDNA encoding the HCoV replicon RNA, we modified the full-length HCoV 229E cDNA, which is cloned and propagated in a vaccinia virus (VV) vector (Thiel et al., 2001a). VV-mediated homologous recombination was used to replace the structural genes S, E and M and non-structural genes 4a and 4b by the GFP gene as well as to insert the TaV-2A/neo/EMCV-IRES sequence. The genomic DNA of the resulting recombinant VV was used as template for *in vitro* transcription to synthesize the replicon RNA, designated Rep-1. Two days after transfection of Rep-1 RNA into BHK-21 cells, green fluorescent cells could be detected, which is indicative of replicon RNA-mediated GFP expression. To establish a stable cell line containing autonomously replicating Rep-1 RNA (figure 1b), transfected cells were selected with Geneticin (G418). Several clones displaying green fluorescence have been obtained and analysed for the presence of replicon RNAs. Northern blot analysis of poly(A)-containing RNA derived from several individual clones revealed the presence of full-length Rep-1 RNA and Rep-1-derived transcripts encoding GFP and N (figure 2a). GFP expression of one clone, designated BHK-Rep-1, has been analysed in detail by fluorescence microscopy and flow cytometry for a period of 4 months (over 50 passages under G418 selection). Throughout this period the percentage of green fluorescent cells remained at a constant level of 40 – 60 % (figure 2b,c). Furthermore, BHK-Rep-1 cells that have been stored at liquid nitrogen and re-cultured still displayed green fluorescence indistinguishable from cells that have been passaged continuously. A complete sequence analysis of the replicon RNA, re-isolated from BHK-Rep-1 cells that had been passaged for two months in tissue culture, showed five nucleotide changes that were not encoded in the Rep-1 cDNA. (table 1). The nucleotide changes are located in the replicase gene (in the coding sequences of nsp1 and nsp2), in the EMCV-IRES element (two changes) and,

interestingly, in the N gene (one change). Since these nucleotide changes were not encoded by the cloned cDNA, they, most likely, have been acquired during replication in tissue culture and thus, indicate the capability of autonomous replication of Rep-1 RNA. Whether they contribute to the establishment of autonomously replicating Rep-1 RNA in BHK cells remains to be elucidated.

Inhibition of coronavirus replication. One of the major goals in this study was to develop a fast and convenient antiviral screening protocol for the detection of coronavirus replicase inhibitors. It should be possible to assess inhibitory effects of candidate compounds by simply seeding out the BHK-Rep-1 cells and assaying for reporter gene expression levels. Decreasing reporter gene expression should indicate the antiviral activity of a particular compound. To test this hypothesis, we have assessed the antiviral activity of three compounds, namely ribavirin, glycyrrhizin and interferon- ζ (IFN- ζ), all of which have been evaluated in the context of SARS-CoV inhibition in tissue culture (Cinatl et al., 2003a, Cinatl et al., 2003b). The compounds were applied in different concentrations on BHK-Rep-1 cells and three days later GFP expression was monitored by FACS analysis and fluorescent microscopy. Untreated cells served as negative controls and the cytotoxic effect of each compound was assessed in parallel on parental BHK-21 cells. As shown in figure 3a, incubation of BHK-Rep-1 cells with IFN- ζ (10000 U/ml) reduced not only the level of reporter gene expression but also the overall percentage of green fluorescent cells. Importantly, cell viability was not affected by IFN- ζ -treatment, even when the maximal dose had been applied. Titration of IFN- ζ showed that the inhibition of coronavirus replication was dose-dependent and half-maximal inhibition could be achieved with only 10 U/ml (figure 3b). In contrast, the maximal inhibitory effects of ribavirin- and glycyrrhizin did not exceed 10% at the non-toxic concentrations of 300 σ g/ml and 3000 σ g/ml, respectively (figure 3c). It is important to note that concentrations of 3000 σ g/ml ribavirin

and 10000 σ g/ml glycyrrhizin reduced cell viability below 50% (data not shown). These results indicate that the therapeutic efficacy of these compounds may be limited because their specific inhibitory effect is rather low and accompanied with significant cytotoxicity. Taken together, our data indicate that IFN- ζ appears to represent a promising candidate for the inhibition of coronavirus replicase function and furthermore, that the inhibition of coronavirus replication can be monitored using coronavirus replicon RNA-containing cell lines.

Discussion

During the first outbreak of SARS in 2002/2003, the causal pathogen was identified, isolated and sequenced in a remarkable short period of time. Soon after, detailed information on the organization of the SARS-CoV genome organization, phylogenetic relationships and the biochemical and structural characteristics of a number of its key replicative enzymes became available. Despite these remarkable achievements, the development of effective strategies to prevent or treat coronavirus infections remains a significant challenge. The rational design of antiviral drugs will certainly be one approach but potential target enzymes have to be characterized in great detail (Anand et al., 2003, Tanner et al., 2003, Yang et al., 2003). Therefore, it seems likely that the identification of coronavirus antivirals will include the random screening of small compound libraries. Fast and convenient screening assays have to be established and, as the accidental laboratory infections that occurred in Taiwan and Singapore in 2003 made clear, the biosafety of the screening procedure is a matter of importance.

In order to provide a convenient and safe assay for the identification of coronavirus replicase inhibitors, we have established a protocol that is based on the use of replicon RNAs generated by reverse genetic techniques. The replicon RNAs have been modified by the introduction of the neomycin resistance gene into the replicase gene and the introduction of a downstream reporter gene (GFP) that is expressed via the replicase-mediated synthesis of a subgenomic mRNA. This strategy enabled us to select for stable cell lines containing coronavirus-derived, autonomously replicating RNAs that mediate the expression of GFP as a marker for coronavirus replication.

Our results clearly demonstrate that the coronavirus replicon system facilitates the identification of replicase inhibitors in tissue culture. We could show that IFN- ζ greatly reduces the replicon-mediated GFP expression in our system. This finding confirms a study by Cinatl and

colleagues (Cinatl et al., 2003b), who showed that IFN- ζ effectively inhibited SARS-CoV replication in tissue culture. The observed inhibition of coronavirus replication by IFN- ζ in our assay reached a maximum of about 80%, indicating that a small number of cells still displayed green fluorescence. Whether these cells do not respond to IFN- ζ -treatment or may contain resistant replicon RNAs remains to be determined.

Our finding, that ribavirin only moderately affected coronavirus replication is also in agreement with reports showing that SARS patients may not benefit from ribavirin treatment and that ribavirin does not inhibit SARS-CoV replication in tissue culture (Booth et al., 2003, Cinatl et al., 2003a, So et al., 2003). Although it has been claimed that ribavirin forces some RNA viruses into "error catastrophe" (Crotty et al., 2001, Crotty et al., 2000), our data further support the notion that coronaviruses may be resistant to this drug. The mechanisms of this resistance are unclear but may be linked to a recently identified 3'-to-5' exonuclease domain (ExoN) encoded by coronaviruses (Snijder et al., 2003). It has been proposed that the ExoN activity may be linked to mechanisms such as RNA proofreading, repair, and/or recombination which ensure the genetic stability of the extraordinary large coronavirus RNA genome. Although this idea has still to be proven, our data that re-isolated replicon RNAs after two months in tissue culture showed only five nucleotide changes already indicate a low error frequency of coronavirus replication.

Unlike IFN- ζ and ribavirin, there are few clues to the antiviral mechanism of glycyrrhizin. Our data indicate that, as for ribavirin, glycyrrhizin only moderately affects coronavirus replicase functions. However, in contrast to ribavirin, glycyrrhizin has been shown to inhibit SARS-CoV replication in tissue culture (Cinatl et al., 2003a). This indicates that glycyrrhizin may not target the coronavirus replication machinery and that antiviral effects may be exerted, for example, during virus adsorption or release.

The replicon system we describe here can be used to screen for coronavirus inhibitors that display antiviral activity against cellular and viral targets involved in viral RNA replication,

including proteases, RNA-dependent RNA polymerase, NTPase/helicase, and various (putative) functions that have been identified recently (i.e. poly(U)-specific endonuclease, ExoN, S-adenosylmethionine-dependent ribose 2'-O-methyltransferase, adenosine diphosphate-ribose 1"-phosphatase and cyclic phosphodiesterase). In addition, the impact of drugs that modify the host cell can be studied in detail. However, because the structural genes S, E and M are not included in the replicon RNA, the assay does not cover cellular and viral targets that are involved in receptor binding, virus entry, genome encapsidation and virus release. On the other hand, since no infectious virus is formed, the assay represents a safe protocol that can be performed in biosafety level 2 laboratories. Another advantage of the coronavirus replicon system is that HCoV 229E vector and replicon RNAs are able to mediate reporter gene expression in a variety of human cells. Thus, the efficacy of candidate inhibitors can be evaluated in human cells from different tissues, irrespectively of their susceptibility to SARS-CoV infection in tissue culture. Furthermore, it should be noted that the replicon system could be further improved, since the expression of other reporter proteins that would improve sensitivity, such as chloramphenicol-acetyltransferase or firefly luciferase, are compatible with coronavirus-based transcription (Thiel et al., 2003b).

The replicon-based system will also facilitate the analysis of viral resistance as it may occur under selection pressure. For example, once an efficient inhibitor has been identified, the replicon cell line can be used to select for drug resistant cells under continuous inhibitor and G418 treatment. Replicon RNA, re-isolated from such cells, can then be analysed for marker mutations related to drug resistance. The detection of such mutations would provide valuable information on the viral target protein(s) and the mechanism of inhibition if the basis of antiviral activity of a particular compound is unknown.

The generation of the first selectable coronavirus replicon RNA represents a novel system and, clearly, there are a number of questions that have to be addressed in future studies. For

example, do coronavirus replicon RNAs display any cytotoxicity? Do the observed nucleotide changes in the re-isolated replicon RNA represent adaptive mutations necessary for continuous replication in selected cell lines? Using HCoV-229E vector RNAs we have observed heterologous gene expression for at least one week in tissue culture without any indication for vector-related cytotoxicity (unpublished data). On the other hand, the amount of replicon RNA in stable BHK-Rep-1 cell lines is at least 10-fold reduced compared to the amount of viral genomes in coronavirus-infected cells (figure 2a), indicating that (co-) selection (replicon/cell) may have taken place. Another important question is whether any host factors are required for coronavirus replication and/or transcription. Are there, for example, eukaryotic cells that are not permissive for coronavirus replication? Studies aimed to identify and analyse such cell lines are currently performed in our laboratory.

Phylogenetic, biochemical and structural data indicate that the mechanisms and enzymes involved in coronavirus genome expression are, to a high degree, similar amongst different coronaviruses (Anand et al., 2003, Gonzalez et al., 2003, Snijder et al., 2003, Thiel et al., 2003a, Yang et al., 2003, Ziebuhr et al., 2000). We, therefore, predict that inhibitors identified using the HCoV-229E replicon system described here will, most likely, also be effective against SARS-CoV replication. Nevertheless, our data should encourage the development of SARS-CoV replicons that would allow the identification of inhibitors that are specific for SARS-CoV. A reverse genetic system for the construction of SARS-CoV replicons is available (Yount et al., 2003), and the strategy we have used to generate selectable HCoV 229E replicons should be fully applicable to SARS-CoV.

Acknowledgements. T.H. and E.S. contributed equally to this work. This study was supported by the Gebert-Rüf Foundation and the Deutsche Forschungsgemeinschaft.

References

- Almazan, F., Gonzalez, J. M., Penzes, Z., Izeta, A., Calvo, E., Plana-Duran, J. & Enjuanes, L. (2000).** Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc Natl Acad Sci U S A* **97**, 5516-5521.
- Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J. & Hilgenfeld, R. (2002).** Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *Embo J* **21**, 3213-3224.
- Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. (2003).** Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763-1767.
- Bartenschlager, R. (2002).** Hepatitis C virus replicons: potential role for drug development. *Nat Rev Drug Discov* **1**, 911-916.
- Booth, C. M., Matukas, L. M., Tomlinson, G. A. & 18 other authors (2003).** Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. *Jama* **289**, 2801-2809.
- Casais, R., Thiel, V., Siddell, S. G., Cavanagh, D. & Britton, P. (2001).** Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J Virol* **75**, 12359-12369.
- Cinatl, J., Morgenstern, B., Bauer, G., Chandra, P., Rabenau, H. & Doerr, H. W. (2003a).** Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus. *Lancet* **361**, 2045-2046.
- Cinatl, J., Morgenstern, B., Bauer, G., Chandra, P., Rabenau, H. & Doerr, H. W. (2003b).** Treatment of SARS with human interferons. *Lancet* **362**, 293-294.
- Crotty, S., Cameron, C. E. & Andino, R. (2001).** RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A* **98**, 6895-6900.

Crotty, S., Maag, D., Arnold, J. J., Zhong, W., Lau, J. Y., Hong, Z., Andino, R. & Cameron, C. E. (2000). The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med* **6**, 1375-1379.

Davidson, A. & Siddell, S. (2003). Potential for antiviral treatment of severe acute respiratory syndrome. *Curr Opin Infect Dis* **16**, 565-571.

Donnelly, C. A., Ghani, A. C., Leung, G. M., & 16 other authors (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet* **361**, 1761-1766.

Donnelly, M. L., Hughes, L. E., Luke, G., Mendoza, H., ten Dam, E., Gani, D. & Ryan, M. D. (2001). The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J Gen Virol* **82**, 1027-1041.

Drosten, C., Gunther, S., Preiser, W., & 23 other authors (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**, 1967-1976.

Elroy-Stein, O., Fuerst, T. R. & Moss, B. (1989). Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proc Natl Acad Sci U S A* **86**, 6126-6130.

Fouchier, R. A., Kuiken, T., Schutten, M., van Amerongen, G., van Doornum, G. J., van den Hoogen, B. G., Peiris, M., Lim, W., Stohr, K. & Osterhaus, A. D. (2003). Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* **423**, 240.

Frolov, I., Hoffman, T. A., Pragai, B. M., Dryga, S. A., Huang, H. V., Schlesinger, S. & Rice, C. M. (1996). Alphavirus-based expression vectors: strategies and applications. *Proc Natl Acad Sci U S A* **93**, 11371-11377.

Gonzalez, J. M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A. E. & Enjuanes, L. (2003). A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Arch Virol* **148**, 2207-2235.

Guan, Y., Zheng, B. J., He, Y. Q. & 15 other authors (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**, 276-278.

Isaacs, S. N., Kotwal, G. J. & Moss, B. (1990). Reverse guanine phosphoribosyltransferase selection of recombinant vaccinia viruses. *Virology* **178**, 626-630.

Kerr, S. M. & Smith, G. L. (1991). Vaccinia virus DNA ligase is nonessential for virus replication: recovery of plasmids from virus-infected cells. *Virology* **180**, 625-632.

Khromykh, A. A. (2000). Replicon-based vectors of positive strand RNA viruses. *Curr Opin Mol Ther* **2**, 555-569.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S. & 23 other authors (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953-66.

Kuiken, T., Fouchier, R. A., Schutten, M. & 19 other authors. (2003). Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* **362**, 263-270.

Lai, M. M. C. & Holmes, K. V. (2001). Coronaviridae: the viruses and their replication. In *Fields Virology*, 4 edn, pp. 1163-1185. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams and Wilkins.

Lo, M. K., Tilgner, M. & Shi, P. Y. (2003). Potential high-throughput assay for screening inhibitors of West Nile virus replication. *J Virol* **77**, 12901-12906.

Marra, M. A., Jones, S. J., Astell, C. R. and 56 other authors (2003). The Genome sequence of the SARS-associated coronavirus. *Science* **300**, 1399-1404.

Martina, B. E., Haagmans, B. L., Kuiken, T., Fouchier, R. A., Rimmelzwaan, G. F., Van Amerongen, G., Peiris, J. S., Lim, W. & Osterhaus, A. D. (2003). Virology: SARS virus infection of cats and ferrets. *Nature* **425**, 915.

Merchinsky, M. & Moss, B. (1992). Introduction of foreign DNA into the vaccinia virus genome by in vitro ligation: recombination-independent selectable cloning vectors. *Virology* **190**, 522-526.

- Parry, J. (2004).** WHO confirms SARS in Chinese journalist. *Bmj* **328**, 65.
- Peiris, J. S., Lai, S. T., Poon, L. L. and 13 other authors (2003).** Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319-1325.
- Randall, G. & Rice, C. M. (2001).** Hepatitis C virus cell culture replication systems: their potential use for the development of antiviral therapies. *Curr Opin Infect Dis* **14**, 743-747.
- Rota, P. A., Oberste, M. S., Monroe, S. S. & 32 other authors (2003).** Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**, 1394-1399.
- Sawicki, S. G. & Sawicki, D. L. (1998).** A new model for coronavirus transcription. *Adv Exp Med Biol* **440**, 215-219.
- Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J. & Gorbalenya, A. E. (2003).** Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* **331**, 991-1004.
- So, L. K., Lau, A. C., Yam, L. Y., Cheung, T. M., Poon, E., Yung, R. W. & Yuen, K. Y. (2003).** Development of a standard treatment protocol for severe acute respiratory syndrome. *Lancet* **361**, 1615-1617.
- Tanner, J. A., Watt, R. M., Chai, Y. B., Lu, L. Y., Lin, M. C., Peiris, J. S., Poon, L. L., Kung, H. F. & Huang, J. D. (2003).** The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *J Biol Chem* **278**, 39578-39582.
- Thiel, V., Herold, J., Schelle, B. & Siddell, S. G. (2001a).** Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J Gen Virol* **82**, 1273-1281.
- Thiel, V., Herold, J., Schelle, B. & Siddell, S. G. (2001b).** Viral replicase gene products suffice for coronavirus discontinuous transcription. *J Virol* **75**, 6676-6681.

- Thiel, V., Ivanov, K. A., Putics, A. & 9 other authors (2003a).** Mechanisms and enzymes involved in SARS coronavirus genome expression. *J Gen Virol* **84**, 2305-2315.
- Thiel, V., Karl, N., Schelle, B., Disterer, P., Klagge, I. & Siddell, S. G. (2003b).** Multigene RNA vector based on coronavirus transcription. *J Virol* **77**, 9790-9798.
- Thiel, V., Rashtchian, A., Herold, J., Schuster, D. M., Guan, N. & Siddell, S. G. (1997).** Effective amplification of 20-kb DNA by reverse transcription PCR. *Anal Biochem* **252**, 62-70.
- Yang, H., Yang, M., Ding, Y. & 12 other authors (2003).** The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **100**, 13190-13195.
- Yount, B., Curtis, K. M. & Baric, R. S. (2000).** Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J Virol* **74**, 10600-10611.
- Yount, B., Curtis, K. M., Fritz, E. A., Hensley, L. E., Jahrling, P. B., Prentice, E., Denison, M. R., Geisbert, T. W. & Baric, R. S. (2003).** Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **100**, 12995-13000.
- Yount, B., Denison, M. R., Weiss, S. R. & Baric, R. S. (2002).** Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* **76**, 11065-11078.
- Ziebuhr, J. & Siddell, S. G. (2002).** Nidovirales. In *The encyclopaedia of life sciences*, pp. 190-198. Edited by R. Atlas, W. F. Bynum & M. Cox. London: Stockton Press.
- Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. (2000).** Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* **81**, 853-879.

Figure legends:

Figure 1.

Structure of HCoV replicon RNA (a) The structural relationship of the HCoV 229E genome and HCoV replicon RNA Rep-1 is shown. ORFs are indicated as boxes. The heterologous sequence inserted at the HCoV 229E nsp1/nsp2 junction is depicted. Note that the nsp1 coding sequence is truncated by three glycine codons and nsp2 translation initiates at an extra AUG inserted upstream of the HCoV 229E nsp2 coding sequence. The position of the translational termination and reinitiation site within the TaV 2A-like sequence is indicated by an arrow (Donnelly et al., 2001). (b) The predicted Rep-1 RNA-mediated gene expression in BHK-Rep-1 cells is shown. Initiation of translation at the authentic HCoV 229E ORF1a start codon is predicted to result in the synthesis of a fusion protein comprised of a carboxyterminally truncated nsp1 protein fused to the TaV 2A-like element. Translation is predicted to be interrupted at the glycine/proline junction (Donnelly et al., 2001) of the TaV 2A-like sequence, resulting in the expression of the neo protein preceded by an additional proline residue. Translational initiation mediated by the EMCV-IRES element is predicted to result in the synthesis of replicase polyproteins pp1a and pp1ab starting with an additional methionine residue at the nsp2-coding region. Replicase gene products are predicted to be involved in (i) the replication of the Rep-1 RNA and (ii) in the transcription of two Rep-1 RNA-derived subgenomic mRNAs encoding GFP and N, respectively.

Figure 2.

Analysis of replicon RNA-containing cells. (a) A northern blot analysis of poly(A)-containing RNA from HCoV 229E infected MRC-5 cells and from two independent BHK-Rep-1 cell clones is shown. The sizes of HCoV 229E genomic and subgenomic RNAs are indicated in kilobases

(kb). Arrows indicate Rep-1 RNA and Rep-1-derived transcripts encoding GFP and N. **(b)** A FACS analysis of BHK-Rep-1 cells is shown. Indicated values represent the percentages of non-fluorescent and green fluorescent cells, respectively. **(c)** Fluorescence (left panel) and phase contrast (right panel) microscopy of BHK-Rep-1 cells is shown.

Figure 3.

Inhibition of coronavirus replication. **(a)** The top panel shows a FACS analysis of untreated (dark grey) and IFN- ζ -treated (light grey, 10000 U/ml) BHK-Rep-1 cells. Indicated values represent the mean fluorescence intensity of gated (bar) cells. Also shown is a fluorescence microscopy analysis of untreated and IFN- ζ -treated (10000 U/ml) BHK-Rep-1 cells. **(b)** The graph shows inhibition of GFP expression of IFN- ζ -treated BHK-Rep-1 cells from three independent experiments. Bar, 95% confidence intervals. **(c)** Inhibition of GFP expression of BHK-Rep-1 cells treated with IFN- ζ (10000 U/ml), ribavirin (300 σ g/ml) and glycyrrhizin (3000 σ g/ml) observed in three (IFN- ζ) or two (ribavirin and glycyrrhizin) independent experiments. Bar, 95% confidence intervals.

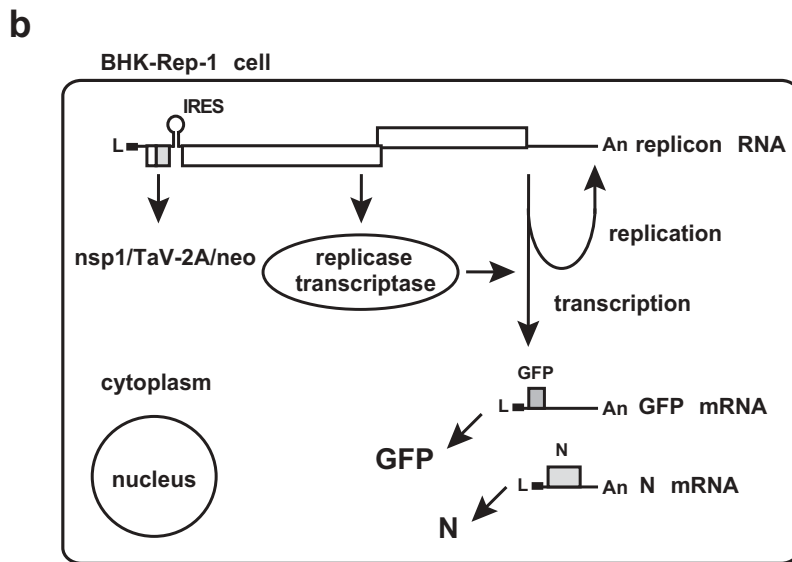
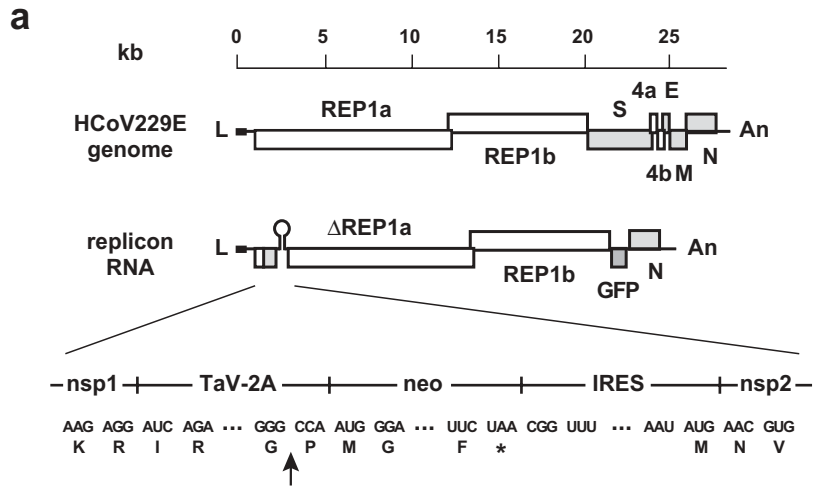


Figure 1

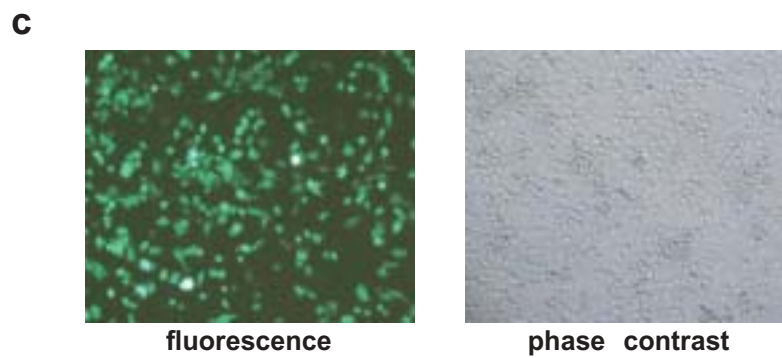
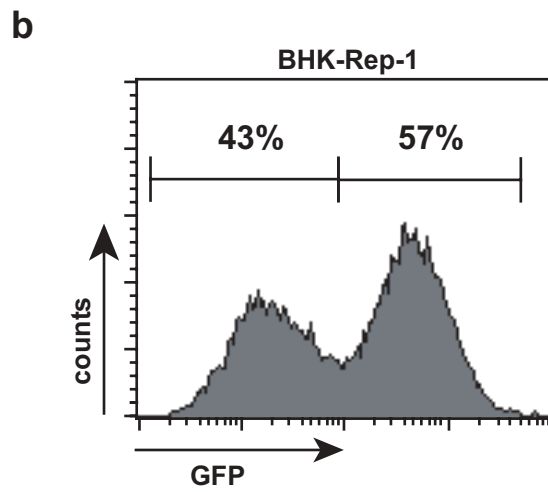
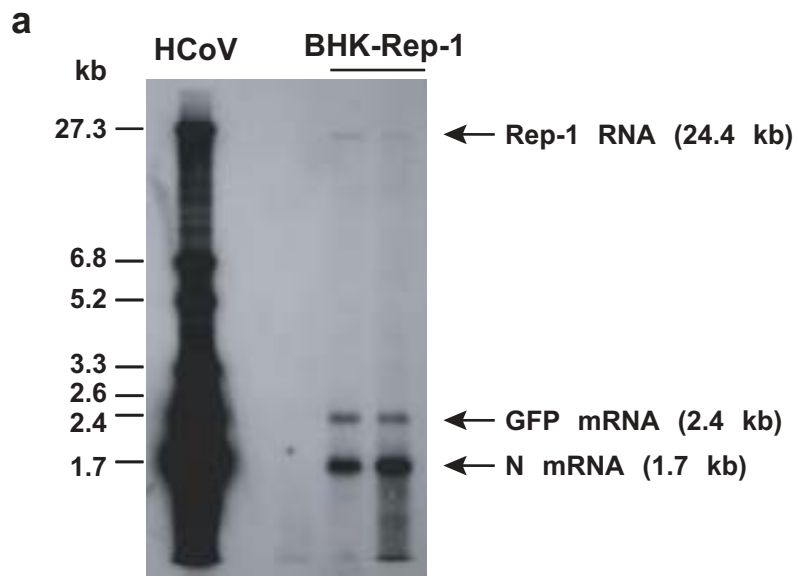
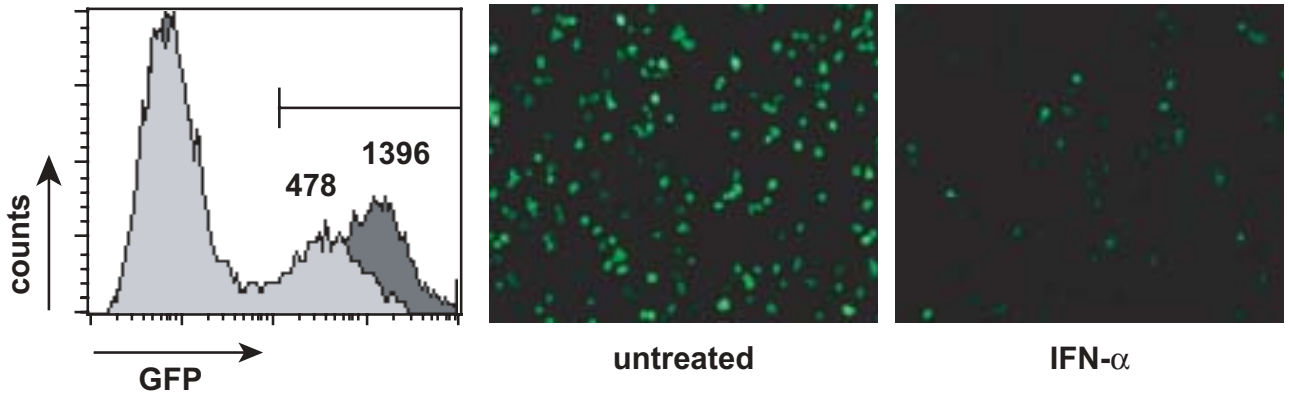
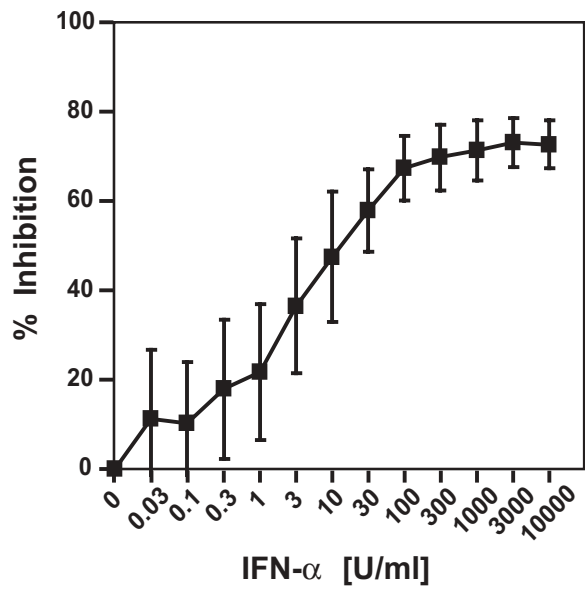


Figure 2

a



b



c

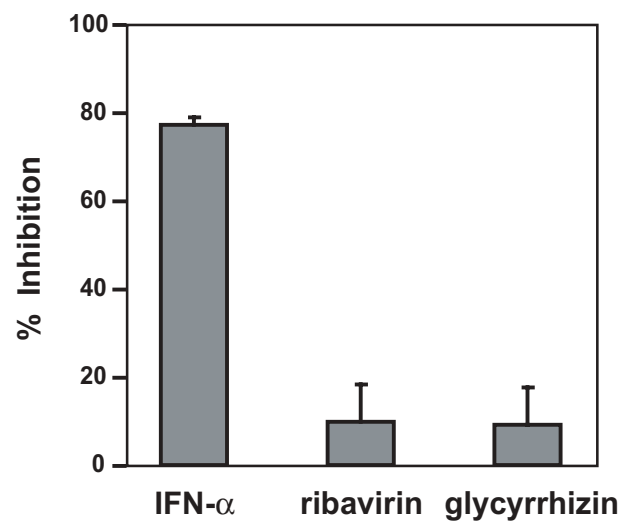


Figure 3

Table I. Nucleotide changes encoded in re-isolated Rep-1 RNA.

Rep-1 RNA nucleotide	HCoV 229E nucleotide	nucleotide change	amino acid change	encoded domain
536	536	C to A	P to T	replicase nsp1
2012	-	A insertion	-	EMCV-IRES
2063	-	A to G	-	EMCV-IRES
3350	1897	A to G	S to S	replicase nsp2
23661	26572	U to C	V to A	nucleocapsid