

Efficient production of protein complexes in mammalian cells using a poxvirus vector

Robert Drillien (✉ rdrillien@orange.fr)

Institut of Genetics and Molecular and Cellular Institut de Genetique et de Biologie Moleculaire et Cellulaire <https://orcid.org/0000-0002-3571-931X>

Karine Pradeau-Aubret

Institut de Génétique et de Biologie Moléculaire et Cellulaire: Institut de Genetique et de Biologie Moleculaire et Cellulaire

Julien Batisse

Institut de Génétique et de Biologie Moléculaire et Cellulaire: Institut de Genetique et de Biologie Moleculaire et Cellulaire

Joëlle Mezher

Institut de Biologie Moléculaire et Cellulaire: Institut de Biologie Moleculaire et Cellulaire

Emma Schenckbecher

Institut de Biologie Moléculaire et Cellulaire: Institut de Biologie Moleculaire et Cellulaire

Justine Marguin

Institut de Biologie Moléculaire et Cellulaire: Institut de Biologie Moleculaire et Cellulaire

Eric Ennifar

Institut de Biologie Moléculaire et Cellulaire: Institut de Biologie Moleculaire et Cellulaire

Marc Ruff

Institut de Génétique et de Biologie Moléculaire et Cellulaire: Institut de Genetique et de Biologie Moleculaire et Cellulaire

Research article

Keywords:

Posted Date: March 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1227811/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The production of full length, biologically active proteins in mammalian cells is critical for a wide variety of purposes ranging from structural studies to preparation of subunit vaccines. Prior research has shown that Modified vaccinia virus Ankara encoding the bacteriophage T7 RNA polymerase (MVA-T7) is particularly suitable for high level expression of proteins upon infection of mammalian cells. The expression system is safe for users and 10-50 mg of full length, biologically active proteins may be obtained in their native state, from a few litres of infected cell cultures.

Results: Here we report further improvements which allow an increase in the ease and speed of recombinant virus isolation, the scale-up of protein production and the simultaneous synthesis of several polypeptides belonging to a protein complex using a single virus vector. Isolation of MVA-T7 viruses encoding foreign proteins was simplified by combining positive selection for virus recombinants and negative selection against parental virus, a process which eliminated the need for tedious plaque purification. Scale-up of protein production was achieved by infecting a BHK 21 suspension cell line and inducing protein expression with previously infected cells instead of virus, thus saving time and effort in handling virus stocks. Protein complexes were produced from infected cells by concatenating the Tobacco Etch Virus (TEV) N1A protease sequence with each of the genes of the complex into a single ORF, each gene being separated from the other by twin TEV protease cleavage sites. We report the application of these methods to the production of a complex formed on the one hand between the HIV-1 integrase and its cell partner LEDGF and on the other between the HIV-1 VIF protein and its cell partners APOBEC3G, CBF β , Elo B and Elo C.

Conclusions: The strategies developed in this study should be valuable for the overexpression and subsequent purification of numerous protein complexes.

Background

Overexpression of proteins as well as protein complexes in higher eukaryotic cells is performed for a variety of applications in research, diagnostics and therapeutics. Such strategies are particularly valuable for proteins from higher eukaryotes that require their natural environment for folding, post-translational processing and optimal solubility. Many alternative strategies have been developed, each with its own advantages and disadvantages. Virus vectors are particularly suitable because they naturally amplify their genomes in infected cells thus providing multiple copies of the genes to be expressed. Furthermore, viruses are readily propagated in a single host cell line thereby avoiding the need to handle distinct cell lines for each gene and viruses are less cumbersome to store and recover from storage than cell lines. As compared to transient expression upon nucleic acid transfection, a powerful method for protein production, virus infection more readily allows expression in entire cell cultures without the need for transfection agents or large amounts of nucleic acid. Thus our focus has been to take advantage of the ability of a member of the Poxvirus family, namely Modified Vaccinia Ankara engineered to encode the bacteriophage T7 RNA polymerase (MVA-T7), to amplify its genome within infected hamster cells and

induce expression of any gene that has been integrated into the viral genome downstream of a bacteriophage T7 promoter [1, 2]. Gene expression from this viral vector is tightly regulated by an *E. coli* lac operator and lac repressor switch under the control of the inducer Isopropyl β -D-1-thiogalactopyranoside (IPTG) [3, 4]. Using this strategy our laboratory has overexpressed and purified several full length cellular and viral proteins for functional and structural studies [5–8]. Despite its efficiency, the current MVA expression system still has a number of drawbacks such as the time required to isolate virus recombinants, the difficulty to scale up protein production and to co-express multiple proteins belonging to a protein complex upon infection with a single virus. Here we report strategies to address each of these issues.

Results

Streamlining recombinant virus isolation

In studies conducted previously we inserted foreign genes to be expressed into the MVA-T7 genome by transfecting infected cells with a plasmid containing the gene downstream of a T7 promoter and adjacent to the gene encoding the *E. coli* xanthine-guanine phosphoribosyl transferase (GPT). *In vivo* recombination between the MVA-T7 genome and the plasmid resulted in the simultaneous insertion of both genes into the viral genome and allowed selection of virus recombinants on the basis of their ability to multiply in medium containing mycophenolic acid, xanthine and hypoxanthine [9], hereafter referred to as MPA medium. Despite this selective pressure, multiple steps of plaque purification were necessary to isolate a clonal population of recombinant virus devoid of parental contaminants. To speed up the process of recombinant virus isolation and simplify it we decided to introduce a negative selection pressure against the parental virus in addition to the positive selection pressure for virus recombinants provided by the GPT gene. For this purpose we added to the parental viral genome (MVA-T7) the fusion gene GyrB-PKR comprising residues 1–220 of the *E. coli* gyrase B fused to the kinase domain (residues 258–551) of human dsRNA dependent protein kinase (PKR) following a procedure devised by White et al. [10] Expression of this GyrB-PKR fusion gene is inhibitory for viral multiplication only in the presence of coumermycin which induces dimerization of the fusion protein, activation of PKR and inhibition of protein synthesis [10, 11]. The addition of the GyrB-PKR fusion gene to the viral genome was ensured by simultaneous insertion of another fusion gene (mCherry-NeoR) which confers both red fluorescence and resistance to G418 [12]. Both the GyrB-PKR and mCherry-NeoR fusion genes were inserted together at the viral hemagglutinin locus which is routinely targeted for insertion of foreign genes. We thus obtained an MVA-T7 vector named MVA-Koom which multiplied well in BHK 21 cells under standard conditions but was inhibited when coumermycin was added to the medium.

Virus recombinants encoding a variety of genes under the control of a T7 promoter were generated by infecting cells with MVA-Koom and transfecting the same cells with plasmids encoding the genes to be expressed under the control of a T7 promoter flanked by the recombination arms of the hemagglutinin locus (figure 1A). To isolate the rare recombination events, a first round of selective pressure was carried out in the presence of MPA medium. Both red (mCherry positive) and green (GFP positive) plaques were

visible under a fluorescent microscope as the MVA Koom virus encodes both mCherry and GFP (adjacent to the gene encoding the T7 RNA polymerase). Virus from this first round of selection was then passaged twice on BHK 21 cells in the presence of MPA medium containing coumermycin. Under these conditions, plaques displaying green fluorescence but no red fluorescence were observed indicating that negative selection against the parental virus had occurred because a new gene had replaced the mCherry-NeoR and GyrB-PKR fusion genes. Under the microscope, virus plaques formed by the MVA-Koom parental virus displayed weak green fluorescence and strong red fluorescence in the absence of the inducer IPTG or in its presence (figure 1B). Virus plaques formed by a recombinant virus encoding YFP downstream of a T7 promoter, isolated following the combined positive and negative selection method, no longer displayed red fluorescence but now displayed yellow fluorescence due to induction of the T7 RNA polymerase by IPTG and strong green fluorescence due to an overlap of the YFP signal in the wavelength selected by the GFP filter. In the absence of IPTG induction, no yellow fluorescence or strong green fluorescence was observed. This isolation strategy, which involves no virus plaque picking, has been used routinely to obtain dozens of virus recombinants with no failures although as discussed below with variable levels of expression of individual proteins.

Protein Production In Suspension Cell Cultures

In prior studies we have used adherent BHK 21 cells for protein production after infection with MVA recombinants [1, 2, 5]. However, adherent cells are not practical for handling large cell volumes particularly for expression and recovery of proteins that remain intracellular as is the case for the gene products we have expressed so far. Adherent cells must be treated with trypsin at each passage when they reach confluence and their recovery after infection requires scraping which is a time-consuming process that can damage cells. We therefore sought to use a suspension cell line (BHK 21-C13-2P), derived from the original BHK 21 cell line, which we cultivated in cylindrical flasks of various sizes and submitted to orbital shaking under standard cell culture conditions (5% CO₂ at 37°C). Initial experiments demonstrated that the amount of protein produced in suspension cells on a per cell basis was about 2.5 fold lower than the amount produced in adherent cells. However, the suspension cells could be grown to higher densities so that the same volume of cells from either type of culture led to equivalent protein yields. Since suspension cells were also easier to handle in large volumes we chose to further explore the potential of this cell culture system. To determine the optimal virus concentration for expression over a 24-hour period, suspension cell cultures were infected in the presence of the inducer IPTG with an MVA-T7 recombinant virus encoding β glucuronidase at a range of multiplicities of infection (MOI). Protein yields were highest when the MOI was 1 or 2 PFU per cell (figure 2A). Extending the time of expression to 36 hours or 48 hours did not result in any increase in yield but rather a slight decrease (not shown) suggesting that cell lysis, known to occur over time, results in protein escape. The MOI required for optimal protein production being relatively high and the viral stocks that could be produced having at best a titer in the range of 10⁷ PFU/ml implied that it would be necessary to produce large amounts of virus stock to induce protein expression in several liters of cell culture as is often necessary to obtain

sufficient amounts for downstream analysis. We therefore examined whether a virus saving method could be set up where cells are infected for protein production with infected cells rather than virus as has been successful for baculovirus expression [13]. For this purpose, suspension cells were infected with an MVA recombinant encoding a model protein (GST-tagged HIV-1 integrase) at a low MOI (0.1 PFU/cell) in the absence of any inducer and virus amplification allowed to proceed for two days. At this point the cell culture was checked for successful infection by ensuring that all cells displayed GFP fluorescence. Infected cells were then mixed with a suspension culture of uninfected cells at various ratios from 1:50 up to 1:2.5 (figure 2B). A control infection with virus instead of infected cells showed that the GST-integrase was detected by SDS-PAGE of the total cell extract in the presence of the IPTG inducer but not in its absence. Interestingly, cells that had been mixed with infected cells produced roughly as much GST-integrase as cells directly infected with virus at all cell ratios indicating that the virus saving procedure was promising. To more clearly determine the relative amount of GST-integrase produced we purified GST-integrase using a glutathione resin and compared the relative protein yields. We could confirm under these conditions that the GST-integrase was produced in the mixed cell cultures at levels comparable to those in the infected cell culture (figure 2C). A ratio of 1 infected cell per 10 uninfected cells was sufficient to produce the highest protein yield so this ratio has been routinely used in numerous protein production runs involving up to 24 liters of infected cell culture. A schematic representation of the overall production procedure is shown in supplementary figure 1.

The polyprotein precursor strategy as applied to two HIV protein complexes

In an initial attempt to express multiple proteins from a single virus recombinant, two genes were inserted into the MVA genome downstream of identical T7 promoters positioned at two distinct viral loci. Although both proteins were produced, only poor expression was obtained (not shown). We therefore chose an alternative strategy based on expression of cleavable polyproteins that have been successfully produced after expression in *E. coli* [14, 15], in insect cells infected with baculovirus vectors [16, 17] and in mammalian cells [18, 19]. The genes encoding several proteins known to assemble into a complex were fused together into a single open reading frame downstream of an IPTG inducible T7 promoter. Each open reading frame, devoid of translation stop codons, was separated from the adjacent one by sequences encoding twin TEV cleavage sites (figure 3A and 4A). Taking advantage of codon degeneracy, the protease cleavage site codons were designed to be unique at each ORF boundary to avoid recombination during virus propagation that could result in gene deletion. The gene encoding the TEV protease was inserted at the 5' end of the construct immediately downstream of the T7 promoter so that the polyprotein would be cleaved upon expression and finally a gene encoding the yellow fluorescent protein (YFP) was inserted at the 3' end to provide a readily visible marker of polyprotein expression. Gene assembly was performed from synthetic or PCR amplified ORFs using a BioBrick strategy [20, 21] or Gibson cloning [22]. Expression plasmids were transfected into MVA-T7 infected cells and virus recombinants were selected with MPA and coumermycin counter-selection. In this manner two protein complexes involved in replication of the human immunodeficiency virus HIV1 were overexpressed and purified.

The first protein complex, the HIV-1 preintegration complex, plays a key role in the integration of HIV cDNA into the host cell genome and is comprised of both HIV and host-cell encoded proteins. Two of the key protein players are the 32 kDa viral integrase (IN) and the 75 kDa host lens epithelium-derived growth factor (LEDGF) which tethers the preintegration complex to chromatin through its binding to IN. Further studies of this complex and its interaction with DNA are crucial for a full understanding of HIV infection. Our previous experiments on the IN-LEDGF complex have demonstrated that increased solubility and biochemical activity are observed when mammalian cells rather than *E. coli* or Baculovirus expression vectors are used to produce these proteins [6]. However, the process of producing and purifying individual proteins that are subsequently assembled into a complex is tedious. To express complexes preassembled in cells infected with a single recombinant virus, the genes encoding IN and LEDGF were fused together into one open reading frame (ORF) downstream of an IPTG inducible T7 promoter. Two copies of the gene encoding HIV-1 IN with distinct codons were introduced because previous studies established that the stoichiometry of the IN/LEDGF complex was two to one [23, 24]. The 10His tag fused to the N-Terminus of LEDGF was used to purify the complex by affinity chromatography. Subsequent gel filtration led to production in the 10 mg range starting from 12 liters of suspension cells (figure 3B and 3C). The purified IN/LEDGF complex displayed characteristic 3' processing activity on a DNA sequence that mimics the U5 end of HIV-1 cDNA but no activity in the presence of non-specific DNA (figure 3D).

To further confirm the usefulness of the MVA-T7 platform for expression of a cleavable polyprotein, a complex involved in overcoming cell resistance to HIV infection was produced. The 23 kDa HIV-1 viral infectivity factor (VIF) targets the host restriction factor cytidine deaminase (APOBEC) for degradation in the proteasome. For this to occur VIF enters into a complex in particular with the 46 kDa APOBEC3G protein and three cell proteins namely Core binding factor β (CBF β , 22.2 kDa), Elongin B (EloB, 13.4 kDa) and Elongin C (EloC, 11 kDa), constituting part of the E3 ubiquitin ligase complex [25–27]. Thus, we constructed an MVA-T7 virus encoding the polyprotein CBF β -VIF-EloB-EloC- APOBEC3G, referred to as the VIF-APOBEC complex (figure 4A). Whereas APOBEC3G, expressed on its own from the MVA-T7 genome displayed cytidine deaminase activity, expression of the VIF-APOBEC complex no longer allowed detection of such activity (figure 4B) suggesting that association of APOBEC3G within the complex was inhibitory for its activity as expected from previous studies [28]. We then purified the Vif-APOBEC complex from 2 liters of infected cells (6×10^9 cells) using the Twin-Strep tag at the N-terminus of APOBEC3G. Elution from the Strep-Tactin column allowed detection of a peak fraction where the most prominent Coomassie blue stained proteins corresponded in size to those predicted for APOBEC3G, CBF β , Vif, EloB and EloC respectively (figure 4C) while mass spectrometry confirmed their identity as well as their precise cleavage by the TEV protease. The complex was nevertheless contaminated with a few extraneous polypeptides, particularly Hsp70. Although further studies are necessary to determine whether a complex devoid of these additional species can be obtained, the results indicate that a protein complex composed of 5 different polypeptides can be assembled and purified from cells infected with an MVA-T7 recombinant virus using the polyprotein strategy.

Discussion

The purpose of these studies has been to make the isolation of MVA recombinants faster and easier so as to facilitate the scale up of protein production in mammalian cells and enable the simultaneous synthesis of several polypeptides that interact to form a complex. There are several paths that can be taken to achieve such goals for instance the use of bacterial artificial chromosomes (BACs) containing the entire viral genome [29, 30], CRISPR-Cas9 recombination [31, 32] and even chemical synthesis of the entire MVA genome [33]. Whereas each of these methods has their own advantages and limitations we have found that the combination of positive selection for the GPT gene and negative selection against the GyrB-PKR gene with visual control of virus purification by following the loss of mCherry has considerably simplified the isolation of virus recombinants in our hands and resulted in consistent success. Negative coumermycin selection is attractive because coumermycin has no cytopathic effect at all over a two-day selection period whereas MPA visibly damages cell monolayers. The ease of the combined selection system makes it feasible to isolate in parallel a number of virus recombinants simultaneously and we have done so for dozens of independent viruses. One could consider automating this system entirely to isolate hundreds of virus recombinants in one workflow. The overall time involved to have a recombinant virus stock for expression studies once a plasmid construct is available is 2 to 4 weeks but because tedious plaque isolation is not required, a limited amount of lab work is involved.

Production of large amounts of infected cells is critical for protein yields. The availability of a very fast growing BHK 21 suspension cell line (12 hour doubling time) and the establishment of a simple method for amplification and infection have allowed us so far to achieve as much as 40 mg yields of a protein complex from 12 liters of infected cell culture. Obviously, yields are highly dependent on the proteins being expressed, their solubility, stability and the efficiency of purification. We have focused particularly on the methods required to infect large cell volumes and found that adding virus to suspension cell cultures without any prior adsorption step in a reduced volume is simple and effective. We also report a simple way to infect large cell volumes by mixing uninfected cells with cells that have been previously infected at a low multiplicity. This considerably reduces the amount of virus required initially and saves time making virus stocks as well as storage space. A tradeoff is carrying out protein production over a 3-day period. For handling volumes beyond the 24 liters we have processed so far it could be useful to investigate other methods. The fact that cells are ultimately killed by the infection may appear to be an intrinsic limitation of the system however it is clearly feasible to grow cells to high densities for MVA production[34] as well as to produce MVA using continuous stirred tank procedures[35, 36] raising the option of adapting these methods to industrial scale protein production with MVA vectors.

The ultimate aim, in our case to produce pure protein complexes for functional and structural studies, has been achieved by concatenating multiple ORFs into a single one, each being separated from the other by twin TEV cleavage sites, with the TEV protease being included in the polyprotein. This strategy has been inspired by prior work in other systems with the knowledge that alternative methods such as the use of picornavirus polyprotein processing may be equally suitable. In fact, Bahar et al. recently demonstrated the use of either a Foot and Mouth Disease Virus-2A sequence or a poliovirus internal ribosome entry site as alternative strategies to obtain individual proteins from a polyprotein sequence encoded by MVA [37]. However, these authors resorted to mixed infection of an MVA virus encoding the T7 RNA polymerase

with an MVA virus encoding the polyprotein precursor. The method involving a single recombinant virus as reported here avoids the requirement to fine tune virus titers for mixed infections and thus facilitates volume scaleup.

Using the strategies developed in this report, a biochemically functional Integrase-LEDGF complex was produced in amounts sufficient for numerous enzymatic assays as well as structural studies while saving considerable time and effort. The alternative means involving the production of integrase and LEDGF on their own, then reconstituting the complex, has also been carried out but found to be considerably more time consuming. In addition, we have been successful in producing the Vif-APOBEC complex comprised of 5 distinct polypeptides although further studies will be required to improve the purity of the complex. Overall, the strategies described in this report should be valuable for the production of a variety of protein complexes difficult to produce in non-mammalian host systems and in many instances requiring post-translational modifications as well as a proper environment for folding.

Material And Methods

Cells and their propagation

Adherent Syrian hamster BHK 21 cells and suspension-adapted BHK 21 C13-2P cells (European Collection of Cell Cultures, catalogue number 84111301) were both cultured in G-MEM supplemented with 10% fetal calf serum, 1,5 g/L bacto tryptose phosphate and 40 mg/L gentamicin.

Plasmids for the construction of virus recombinants

The pKoom plasmid used to generate the MVA-Koom virus was constructed from pGP-GNR, pRSET-mCherry and pVote2 kindly provided by James Jancovich, Jean-Marie Garnier and Bernard Moss respectively (supplementary figure 2). The BioBrick strategy was set up as follows. Twin TEV cleavage sites, purification tags (10His, Flag10His, 10HisFlag, TwinStrep, FlagTwinStrep and TwinStrepFlag) fused to a P3C cleavage site (Supplementary figure 3A), and primers were purchased from Sigma. Synthetic DNA encoding the HIV-1 codon modified IN, LEDGF and YFP were purchased from GeneArt or GeneScript respectively. PCR was used to amplify TEV, APOBEC3G, CBF β , EloB, EloC, HIV-1 Vif and IN (pNL4.3) with the primers listed (Supplementary figure 3B). Translational stop signals were removed from all ORFs and replaced with a series of degenerate codons for twin TEV cleavage sites (Supplementary figure 3C) except the YFP stop codon. All ORFs were flanked at their 5' end by EcoRI and XbaI sites and at their 3' end by SpeI and PstI sites (Supplementary figure 3D) in order to carry out a reiterative BioBrick cloning strategy and assemble a polyprotein coding sequence [20]. Briefly, as schematically represented in supplementary figure 4, ORFs were prepared with EcoRI and PstI digestion and inserted into pUC19 that was digested with EcoRI and PstI, to generate a pUC_BB plasmid. Each pUC_BB was further linearised with SpeI and PstI to add a downstream BioBrick generated by XbaI and PstI restriction, or linearized with EcoRI and XbaI to add an upstream BioBrick generated by EcoRI and SpeI restriction. The polyprotein coding sequences were then removed from the pUC_BB constructs by XbaI and PstI digestion and inserted into the pENTR_BB-TEV-YFP digested with SpeI and PstI. Finally, the pENTR_BB plasmids were

recombined with the pVote0GW plasmid using Gateway recombination to generate a plasmid for insertion of the polyprotein sequences into the MVA genome as previously described [2].

Selection and amplification of virus recombinants

To isolate the MVA-Koom virus 10^6 adherent BHK 21 cells were infected with MVA-T7 at about 2 PFU/cell and transfected in OptiMEM medium with 1 μ g pKoom plasmid DNA using Lipofectamine 2000 (Thermo Fisher Scientific). Cells were frozen at -20°C one day later and subsequently freeze-thawed several times to release virus. MVA-Koom was selected by serial passage in the presence of 1 mg/ml G418 and visualization of mCherry positive plaques under a fluorescence microscope. To generate virus recombinants encoding genes under the control of the T7 promoter, 10^6 adherent BHK 21 cells were infected with MVA-Koom at about 2 PFU/cell and transfected as above. Cells were frozen at -20°C one day later and subsequently freeze-thawed several times to release virus. Released virus was diluted and used to infect fresh BHK 21 monolayers in the presence of 25 μ g/ml mycophenolic acid, 15 μ g/ml hypoxanthine and 250 μ g/ml xanthine and/or 100 ng/ml coumermycin. Two days later, cell cultures were checked for green and red fluorescence using a fluorescence imaging system. This process was repeated several times in the presence of coumermycin only until all viral plaques displayed green fluorescence and no red fluorescence. At the end of the selection cycles, virus was amplified on BHK 21 cells in the absence of any selection pressure. Final virus stocks were prepared from infected cells displaying GFP expression within the entire cell culture as well as an overall cytopathic effect visible under a light microscope. Infected cells cultures were submitted to three cycles of freeze-thawing at -20° then centrifuged at 2000 g for 5 minutes. The supernatants were recovered, frozen at -20°C and used as crude virus stocks for protein expression experiments.

Protein production and purification

Medium scale protein production was carried out in adherent BHK 21 cells grown in 175 cm^2 plastic bottles or in suspension BHK 21 C13-2P cells grown up to $1\text{--}1.5 \times 10^6$ cells/ml in cylindrical flasks submitted to orbital shaking. In the case of adherent cells, cell culture medium was removed and crude virus stocks were added to the cultures at about 2 PFU/cell for one hour incubation at 37°C . At this time, fresh medium containing 5-10% serum and 1 mM IPTG were added and incubation continued at 37°C for 24 hours. Infected cells were pelleted at 2000 g for 5 minutes, suspended in PBS and pelleted again at 2000 g for 5 minutes. Finally, the cell pellets were stored at -80° until their use for protein analysis or purification. GST-tagged proteins were purified from infected cell pellets using a batch resin method (Glutathione-Sepharose-4B) and a protocol provided by GE Healthcare.

For large scale production of proteins (Supplementary figure 1), suspension BHK 21 C13-2P cells ($1\text{--}1.5$ cells/ml) were infected as above with about 0.1 PFU/cell in 5 liters cylindrical flasks containing 1.2 liters of cell culture medium. Two days later, infected cells were mixed with uninfected cells at a 1:10 ratio in six 5 liters cylindrical flasks containing 1.5 to 2 liters of cell culture at $1\text{--}1.5 \times 10^6$ cells/ml. IPTG (1 mM

final) was added at the time of cell mixture. Incubation at 37°C was pursued for another 24 hours. Cells were then pelleted and washed as above and the pellets stored at -80°C until use.

For purification of Twin-Strep tagged proteins (Vif/APOBEC complex), cell pellets were lysed by sonication in a buffer containing 50 mM Tris-HCl (pH8), 500 mM NaCl, 1mM EDTA, 10 mM CaCl₂, 1mM MgCl₂, 7 mM β mercaptoethanol, 0,5% tween 20 and a protease inhibitor cocktail. The lysates were then treated for one hour with RNase and DNase at 4°C. Insoluble proteins were removed by ultracentrifugation at 100,000 g and the supernatants obtained were filtered through 5µm and 0.2 µm cellulose membranes. Tagged proteins were then retained on a 1ml Strep Tactin column, washed in three successive steps (1 M NaCl, 1 mM EDTA, 50 mM Tris-HCl pH 8 then 150 mM NaCl, 1 mM EDTA, 100 mM Tris-HCl pH8 then 10 mM NaCl, 50 mM Tris-HCl pH8, 5% glycerol, 10 µM ZnCl₂ and eluted with 5 mM desthiobiotin in 50 mM Tris-HCl pH 8, 10mM NaCl, 10 µM ZnCl₂, 5% glycerol. Peak fractions were collected, pooled together, concentrated on a 100 kDa ultracentrifugal filter unit by centrifugation at 4000 rpm for 5 min and analyzed by SDS PAGE and Coomassie staining.

For large scale purification of His-tagged IN/LEDGF complex, cell pellets were suspended in buffer A (50 mM Hepes pH 7.5, 400 mM NaCl, 2 mM MgCl₂, 2 mM β mercaptoethanol, 1 tablet of a protease inhibitor cocktail (Roche)) containing 10 mM imidazole at a ratio of 5ml buffer for 1 g cell pellet. Cell suspensions were sonicated 15 minutes with a 2 seconds on/off cycle. The lysed cells were then centrifuged at 100,000 g for one hour and the supernatants recovered and filtered through a 5 µm cellulose acetate filter. The soluble proteins were then run through a 5 ml HisTrap excel Ni Sepharose column equilibrated with buffer A containing 10 mM imidazole. A 3-step wash was then performed with 10 column volumes of the same buffer (A with 10 mM Imidazole), then buffer A with 20 mM imidazole and buffer A with 40 mM imidazole. Finally, elution of His-tagged proteins was carried out with an imidazole gradient from 40 mM to 500 mM. Peak fractions were collected and analyzed by SDS PAGE and Coomassie staining. Fractions of interest were then pooled together and concentrated to 5 ml on a 100 kDa ultracentrifugal filter unit by centrifugation at 4000 rpm for 10 min. Proteins were then purified by size exclusion on a Hi Load 16/60 superdex-200 column (GE Healthcare) pre-equilibrated in buffer A. Peak fractions were collected and analyzed by SDS PAGE and Coomassie staining.

Measurement of cytidine deaminase and integrase activities

The cytidine deaminase activity of APOBEC3G alone or in a multi-protein complex was determined following a previously described method with a few modifications[38]. Briefly, 40 µM of the 39 base pair oligonucleotide 5' AAAGAGAAAGAGAACCCAAAGAGGAAAGGTGAGGAGGA 3', where the third C in the CCCA motif may be deaminated, was incubated in buffer A (20 mM Tris-HCl pH 7.5, 30 mM NaCl, 10 µM MgCl₂ and 5 mM DTT) in the presence of infected cell lysates for two hours at 37°C. DNA was then extracted with phenol-chloroform, ethanol precipitated, solubilized in buffer A and treated with 10 units uracil glycosylase for two hours at 37°C to remove the uracils created by cytidine deaminase activity. Finally alkaline hydrolysis of the abasic oligonucleotide for 20 min at 95°C in 0.2 M NaOH cut the

abasic 39 oligonucleotides into 17 and 21 long oligonucleotides which were separated on an anion exchange column (Nucleopac PA-100 from Dionex).

The IN 3'processing activity was monitored by fluorescence anisotropy as described previously[39]. The reaction was done in 96 well-plates. One well contained 100 µL of reaction mix composed of 25 mM BisTris pH 6.5, 10 mM MgCl₂, 5 mM DTT, 50 nM DNA and 200 nM protein complex. The DNA substrate was a 40 base pair double strand DNA (5'-GACTACGGTTCAAGTCAGCGTGTGGAAAATCTCTAGCAGT-3'), mimicking the U5 end of HIV-1 DNA and 3' modified by 6-fluorescein. After homogenization, 25µL of paraffin oil was added on the top of the well to avoid evaporation. Fluorescence anisotropy measurements were performed on a PHERAstarPlus (BMGLab) spectrophotofluorimeter with an excitation polarized wavelength of 470 nm. The reaction was monitored for 5 hours at 37°C.

Abbreviations

MVA
Modified Vaccinia Ankara, BHK:Baby hamster kidney
TEV
Tobacco etch virus
GPT
xanthine-guanine phosphoribosyl transferase
GyrB
gyrase B
PKR
dS RNA dependent protein kinase
NeoR
neomycin resistance
YFP
yellow fluorescent protein
IPTG
isopropyl -D-1-thiogalactopyraniside
GFP
green fluorescent protein
moi
multiplicity of infection
PFU
plaque forming units
GST
glutathione S transferase
IN
integrase

LEDGF

lens epithelium-derived growth factor

VIF

viral infectivity factor

APOBEC

apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like

CBF

core binding factor

Elo

Elongin.

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material: Data supporting the findings in this study are included in this published article and its supplementary files. Materials described are available upon request.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by grants from ANRS, SIDACTION and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01, and Instruct as part of the European Strategy Forum on Research Infrastructures (ESFRI).

Authors contributions: RD, KPA, EE and MR conceived the experiments. RD, KPA, JB, JM, ES and JM carried out experiments and performed data analysis. RD and MR wrote the manuscript. All authors have read and approved the manuscript.

Acknowledgments: We are grateful to Lauriane Kuhn, Philippe Wolff and Philippe Hammann on the Proteomics platform at the IBMC and to members of the IGBMC Structural Biology and Genomics platform as well as members of the IGBMC common services.

References

1. Hebben M, Brants J, Birck C, Samama JP, Wasylyk B, Spehner D, Pradeau K, Domi A, Moss B, Schultz P, et al. High level protein expression in mammalian cells using a safe viral vector: modified vaccinia virus Ankara. *Protein Expr Purif*. 2007;56(2):269–78.
2. Pradeau-Aubret K, Ruff M, Garnier JM, Schultz P, Drillien R. Vectors for recombinational cloning and gene expression in mammalian cells using modified vaccinia virus Ankara. *Anal Biochem*. 2010;404(1):103–5.

3. Fuerst TR, Fernandez MP, Moss B. Transfer of the inducible lac repressor/operator system from *Escherichia coli* to a vaccinia virus expression vector. *Proc Natl Acad Sci U S A*. 1989;86(8):2549–53.
4. Moss B, Elroy-Stein O, Mizukami T, Alexander WA, Fuerst TR. Product review. New mammalian expression vectors. *Nature*. 1990;348(6296):91–2.
5. Osz J, Pradeau-Aubret K, Drillien R, Troffer-Charlier N, Kolb-Cheynel I, Poterszman A, Ruff M, Moras D, Rochel N. Expression of functional full-length hSRC-1 in eukaryotic cells using modified vaccinia virus Ankara and baculovirus. *Anal Biochem*. 2012;426(2):106–8.
6. Levy N, Eiler S, Pradeau-Aubret K, Maillot B, Stricher F, Ruff M. Production of unstable proteins through the formation of stable core complexes. *Nat Commun*. 2016;7:10932.
7. Bedez C, Lotz C, Batisse C, Broeck AV, Stote RH, Howard E, Pradeau-Aubret K, Ruff M, Lamour V. Post-translational modifications in DNA topoisomerase 2alpha highlight the role of a eukaryote-specific residue in the ATPase domain. *Sci Rep*. 2018;8(1):9272.
8. Vanden Broeck A, Lotz C, Drillien R, Haas L, Bedez C, Lamour V. Structural basis for allosteric regulation of Human Topoisomerase IIalpha. *Nat Commun*. 2021;12(1):2962.
9. Falkner FG, Moss B. *Escherichia coli* gpt gene provides dominant selection for vaccinia virus open reading frame expression vectors. *J Virol*. 1988;62(6):1849–54.
10. White SD, Conwell K, Langland JO, Jacobs BL. Use of a negative selectable marker for rapid selection of recombinant vaccinia virus. *Biotechniques*. 2011;50(5):303–9.
11. Ung TL, Cao C, Lu J, Ozato K, Dever TE. Heterologous dimerization domains functionally substitute for the double-stranded RNA binding domains of the kinase PKR. *EMBO J*. 2001;20(14):3728–37.
12. Franke CA, Rice CM, Strauss JH, Hruby DE. Neomycin resistance as a dominant selectable marker for selection and isolation of vaccinia virus recombinants. *Mol Cell Biol*. 1985;5(8):1918–24.
13. Nie Y, Bellon-Echeverria I, Trowitzsch S, Bieniossek C, Berger I. Multiprotein complex production in insect cells by using polyproteins. *Methods Mol Biol*. 2014;1091:131–41.
14. Zheng N, Huang X, Yin B, Wang D, Xie Q. An effective system for detecting protein-protein interaction based on in vivo cleavage by PPV Nla protease. *Protein Cell*. 2012;3(12):921–8.
15. Yang J, Xie X, Xiang N, Tian ZX, Dixon R, Wang YP. Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. *Proc Natl Acad Sci U S A*. 2018;115(36):E8509–17.
16. Bieniossek C, Imasaki T, Takagi Y, Berger I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci*. 2012;37(2):49–57.
17. Vijayachandran LS, Viola C, Garzoni F, Trowitzsch S, Bieniossek C, Chaillet M, Schaffitzel C, Busso D, Romier C, Poterszman A, et al. Robots, pipelines, polyproteins: enabling multiprotein expression in prokaryotic and eukaryotic cells. *J Struct Biol*. 2011;175(2):198–208.
18. Chen X, Pham E, Truong K. TEV protease-facilitated stoichiometric delivery of multiple genes using a single expression vector. *Protein Sci*. 2010;19(12):2379–88.

19. Puckette M, Smith JD, Gabbert L, Schutta C, Barrera J, Clark BA, Neilan JG, Rasmussen M. Production of foot-and-mouth disease virus capsid proteins by the TEV protease. *J Biotechnol.* 2018;275:7–12.
20. Ho-Shing O, Lau KH, Vernon W, Eckdahl TT, Campbell AM. Assembly of standardized DNA parts using BioBrick ends in *E. coli*. *Methods Mol Biol.* 2012;852:61–76.
21. Liu S, Ruban L, Wang Y, Zhou Y, Nesbeth DN. Establishing elements of a synthetic biology platform for Vaccinia virus production: BioBrick design, serum-free virus production and microcarrier-based cultivation of CV-1 cells. *Heliyon.* 2017;3(2):e00238.
22. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 2009;6(5):343–5.
23. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J Biol Chem.* 2003;278(1):372–81.
24. Michel F, Crucifix C, Granger F, Eiler S, Mouscadet JF, Korolev S, Agapkina J, Ziganshin R, Gottikh M, Nazabal A, et al. Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J.* 2009;28(7):980–91.
25. Zhou X, Evans SL, Han X, Liu Y, Yu XF. Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBFbeta and CRL5 E3 ubiquitin ligase components. *PLoS One.* 2012;7(3):e33495.
26. Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z. Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature.* 2014;505(7482):229–33.
27. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med.* 2003;9(11):1404–7.
28. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature.* 2002;418(6898):646–50.
29. Domi A, Moss B. Cloning the vaccinia virus genome as a bacterial artificial chromosome in *Escherichia coli* and recovery of infectious virus in mammalian cells. *Proc Natl Acad Sci U S A.* 2002;99(19):12415–20.
30. Cottingham MG, Andersen RF, Spencer AJ, Saurya S, Furze J, Hill AV, Gilbert SC. Recombination-mediated genetic engineering of a bacterial artificial chromosome clone of modified vaccinia virus Ankara (MVA). *PLoS One.* 2008;3(2):e1638.
31. Di Gioia C, Yuan M, Wang Y. Vaccinia Virus Genome Editing Using CRISPR. *Methods Mol Biol.* 2019;2023:109–17.
32. Gowripalan A, Smith S, Stefanovic T, Tschärke DC. Rapid poxvirus engineering using CRISPR/Cas9 as a selection tool. *Commun Biol.* 2020;3(1):643.
33. Chiuppesi F, Salazar MD, Contreras H, Nguyen VH, Martinez J, Park Y, Nguyen J, Kha M, Iniguez A, Zhou Q, et al. Development of a multi-antigenic SARS-CoV-2 vaccine candidate using a synthetic poxvirus platform. *Nat Commun.* 2020;11(1):6121.

34. Vazquez-Ramirez D, Genzel Y, Jordan I, Sandig V, Reichl U. High-cell-density cultivations to increase MVA virus production. *Vaccine*. 2018;36(22):3124–33.
35. Tapia F, Jordan I, Genzel Y, Reichl U. Efficient and stable production of Modified Vaccinia Ankara virus in two-stage semi-continuous and in continuous stirred tank cultivation systems. *PLoS One*. 2017;12(8):e0182553.
36. Granicher G, Tapia F, Behrendt I, Jordan I, Genzel Y, Reichl U. Production of Modified Vaccinia Ankara Virus by Intensified Cell Cultures: A Comparison of Platform Technologies for Viral Vector Production. *Biotechnol J*. 2021;16(1):e2000024.
37. Bahar MW, Porta C, Fox H, Macadam AJ, Fry EE, Stuart DI. Mammalian expression of virus-like particles as a proof of principle for next generation polio vaccines. *NPJ Vaccines*. 2021;6(1):5.
38. Rausch JW, Chelico L, Goodman MF, Le Grice SF. Dissecting APOBEC3G substrate specificity by nucleoside analog interference. *J Biol Chem*. 2009;284(11):7047–58.
39. Maillot B, Levy N, Eiler S, Crucifix C, Granger F, Richert L, Didier P, Godet J, Pradeau-Aubret K, Emiliani S, et al. Structural and functional role of INI1 and LEDGF in the HIV-1 preintegration complex. *PLoS One*. 2013;8(4):e60734.

Figures

Figure 1

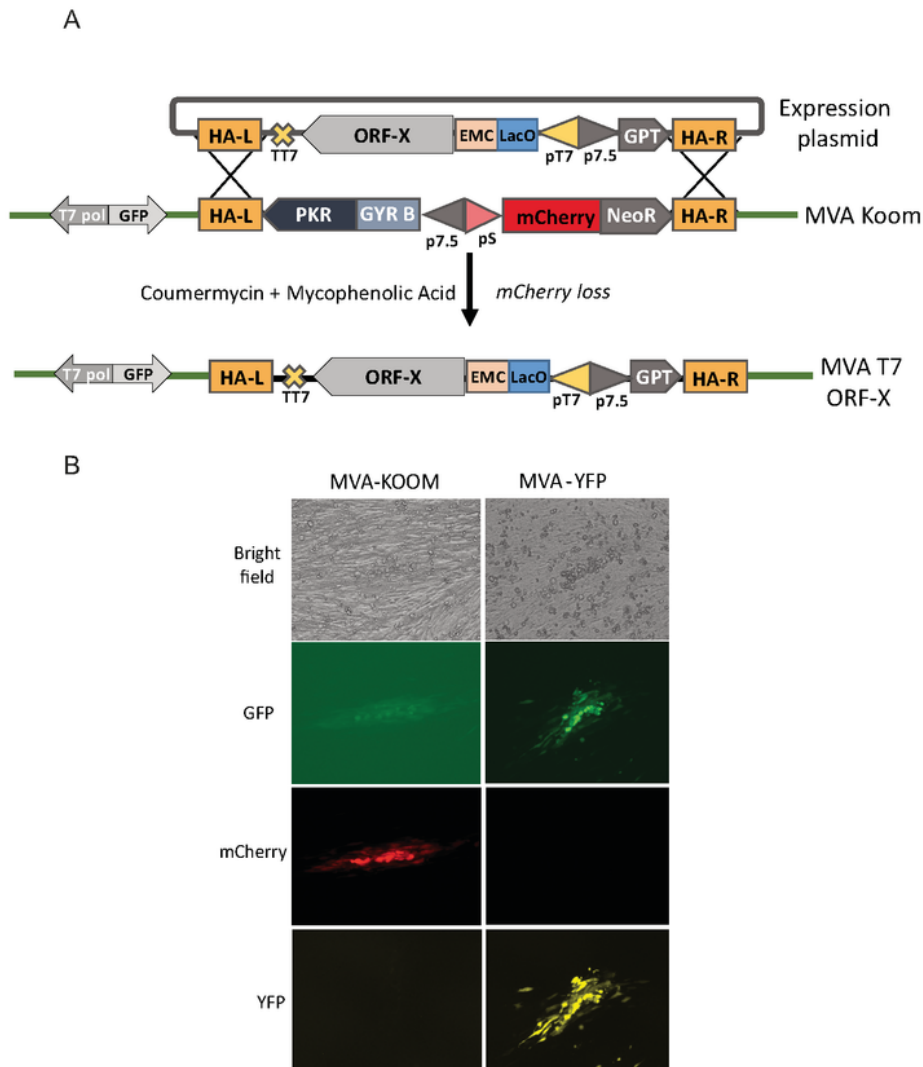


Figure 1

Construction of virus recombinants by combined positive and negative selection

A. A vaccinia virus expression plasmid encoding an ORF of interest (ORF-X) is cloned downstream of a bacteriophage T7 promoter (pT7) under the additional transcriptional control of a Lac operator (LacO), an encephalomyocarditis (EMC) virus leader sequence and a bacteriophage transcriptional stop sequence

(TT7). A positive selection marker (GPT behind a 7.5 early/late vaccinia virus promoter is included on the plasmid. Sequences on the left and right sides of the vaccinia virus hemagglutinin gene (HA-L and HA-R) are used in vivo to recombine with the MVA-Koom virus which encodes T7 RNA polymerase, GFP as well as the GYR-PKR and the mCherry-NeoR fusion genes. Replacement of the fusion genes by OFF-X is selected by the ability of the recombinant virus to multiply in the presence of MPA and coumermycin resulting in loss of mCherry expression. Note that the sequences are represented according to the standard orientation of the virus genome.

B. Analysis of recombinant virus plaques by light and epifluorescence microscopy. BHK 21 cells were infected in the presence of IPTG with MVA-Koom or an MVA-T7 recombinant virus derived from MVA-Koom by insertion of a polyprotein sequence encoding YFP at its 3' end (see figure 3). One representative virus plaque for each infection was observed with an EVOS Cell Imaging System using from top to bottom a white light cube (top panel), a GFP light cube (470/22, 510/42), a YFP light cube 500/24, 524/27 or an RFP light cube (531/40, 593/40).

Figure 2

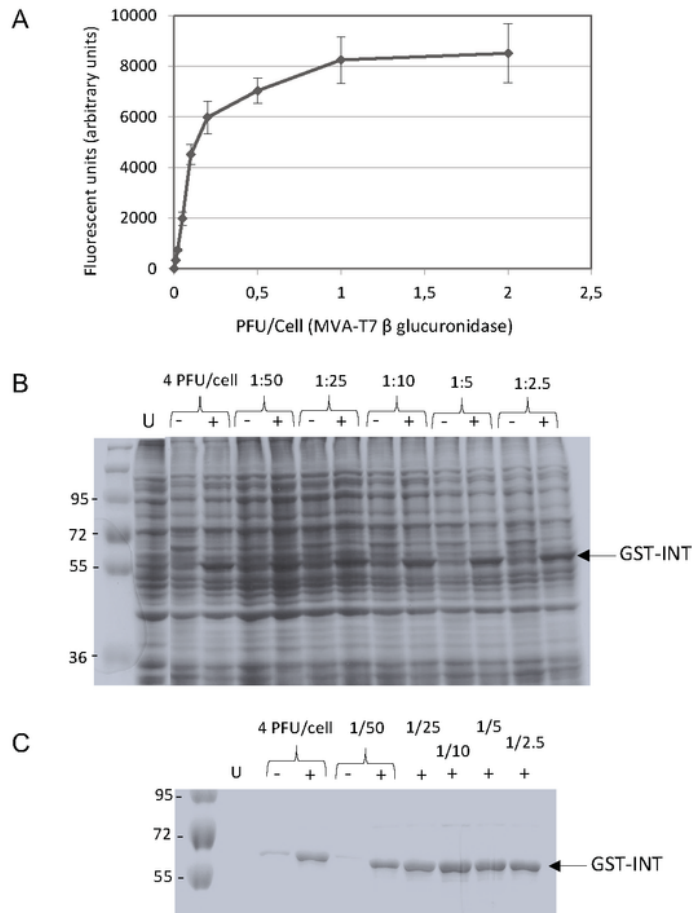


Figure 2

Protein production in suspension cell cultures

A. BHK 21 suspension cells (1.2×10^6 cells/ml) were infected in 50 ml cell culture medium with MVA-T7 encoding β glucuronidase at increasing MOIs. After 24 hours infection cells were pelleted and enzymatic activity determined in duplicate.

B. BHK 21 suspension cells (10^6 cells/ml) were infected in a 100 ml cell culture medium with an MVA-T7 virus encoding a GST-tagged HIV-1 integrase at 0.1 PFU/cell. After 48 hours 1 ml, 2 ml, 5 ml, 10 ml or 20 ml of infected cells were added to 50 ml of uninfected cells (10^6 cells/ml). Cell culture medium was added so that the culture volumes were identical in all samples and protein expression was induced by the simultaneous addition of IPTG. Control infections were conducted with 4PFU/cell of the MVA-T7 virus encoding GST-integrase in the presence or absence of IPTG. 24 hours later 1.5ml of the infected cell cultures were recovered for analysis of total protein by PAGE. U uninfected cells. – and + indicates the presence or absence of IPTG in the infected cell cultures.

C. 45 ml samples of the infected BHK 21 suspension cultures in panel B were pelleted, sonicated and the GST-integrase purified and examined by PAGE. U uninfected cells. – and + indicates the absence or presence of IPTG in the infected cell cultures

Figure 3

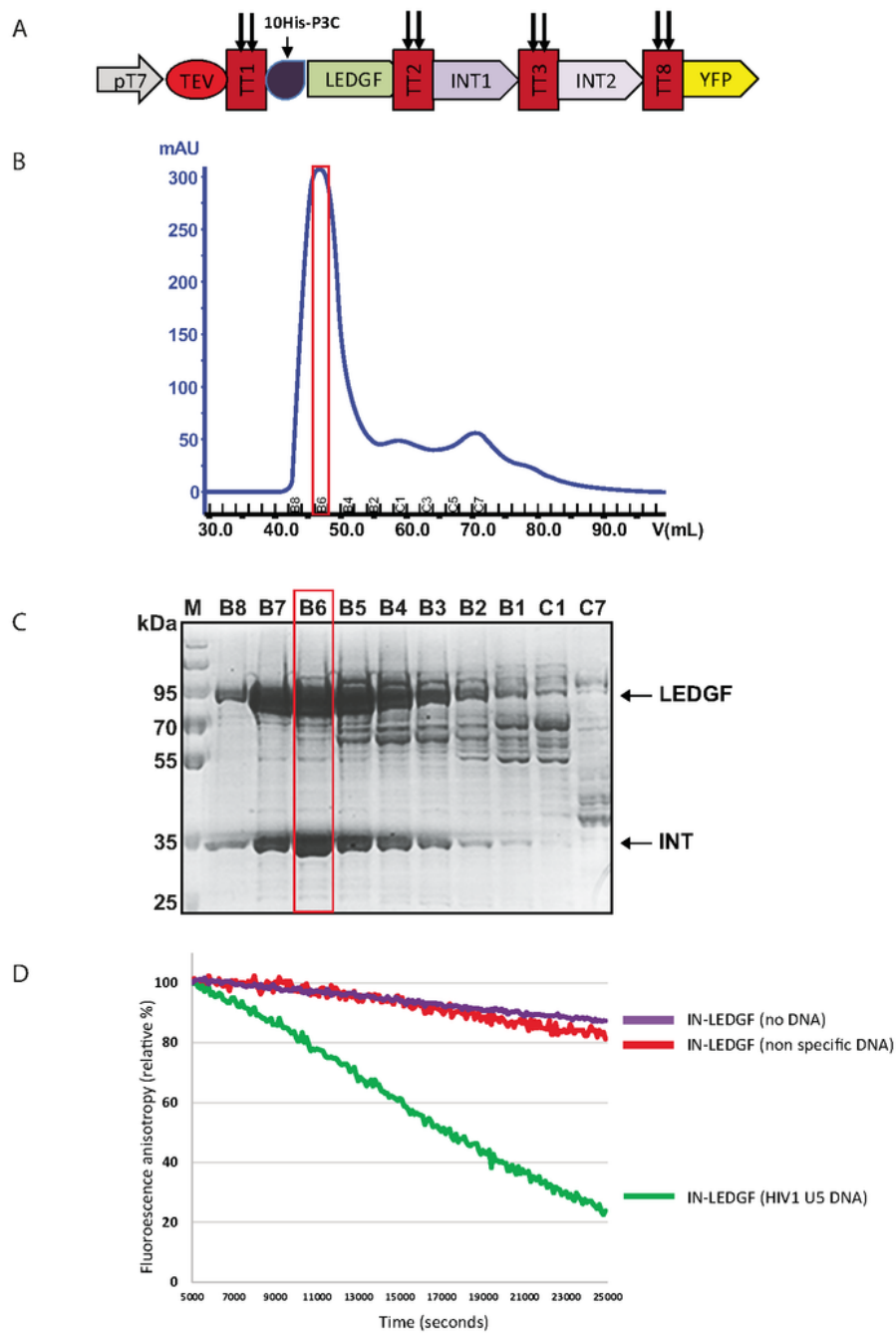


Figure 3

Expression of the IN/LEDGF complex

A. Schematic drawing of the ORFs inserted into the MVA-T7 virus behind a T7 promoter and encoding a polyprotein including from 5' to 3' the TEV protease, LEDGF, two distinct copies of the HIV-1 IN and YFP. Each ORF is separated from the other by twin TEV cleavage sites (perpendicular arrows).

- B. Purification of the IN-LEDGF complex from 12×10^9 suspension BHK 21 cells infected with an MVA recombinant encoding the complex. The His-tagged complex was first purified by HPLC on a HisTrap excel Ni sepharose column (not shown) then the pooled fractions were concentrated and separated by size fractionation on a Hi Load 16/60 Superdex-200 column. The position of the B6 peak fraction is shown.
- C. PAGE was performed on aliquots of the fractions recovered from the Superdex-200 column. Arrows point to the LEDGF and IN proteins stained with Coomassie blue.
- D. 3' processing test of the IN/LEDGF complex. The release of GT-fluorescent dinucleotide was monitored by following fluorescence anisotropy as a function of time for the IN-LEDGF complex in the absence of DNA; in the presence of non-specific DNA or in the presence of HIV-1 U5 DNA. A control reaction was performed with IN alone in the presence of HIV-1 U5 DNA.

Figure 4

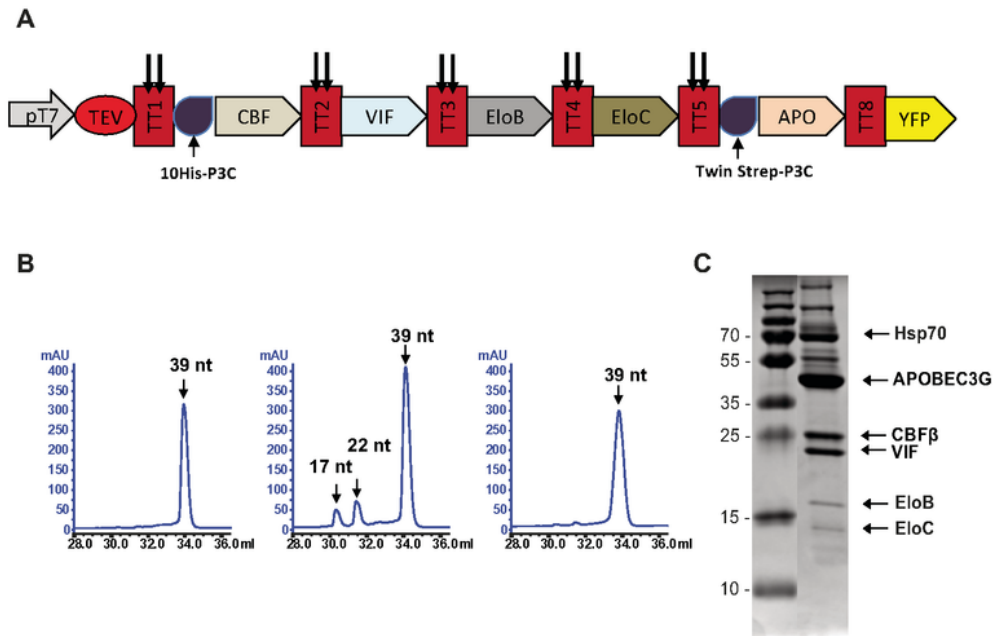


Figure 4

Expression of the Vif-APOBEC complex

A. Schematic drawing of the ORFs inserted into the MVA-T7 virus behind a T7 promoter and encoding a polyprotein including from 5' to 3' the TEV protease, CBF β , VIF, EloB, EloC, APOBEC3G and YFP. Each ORF is separated from the other by twin TEV cleavage sites (perpendicular arrows).

B. Cytidine deaminase activity in the absence of cell extract (left) or in the presence of cell extracts from BHK21 cells infected with MVA-T7 encoding APOBEC3G (middle) or MVA-T7 encoding the Vif-APOBEC complex (right). Enzymatic activity was determined in a two-step reaction where the appearance of the 17 and 21 oligonucleotide products was indicative of deaminase activity.

C. Purification of the Vif-APOBEC complex from 2 liters of suspension BHK 21 C13-2P cells. Cell lysates clarified by ultracentrifugation at 100,000 g were injected onto a 1 ml StrepTactin® column and washed extensively. Peak fractions were combined, concentrated and examined by PAGE (right lane). Molecular weight markers are displayed in the left lane. Arrows point to proteins migrating with the mobility expected for APOBEC3G, CBF β and Vif. The identity of these proteins was confirmed by mass spectrometry. Proteins were stained with Coomassie blue.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [LegendstoSupplementaryfigures.docx](#)
- [Supplementaryfigure1.pdf](#)
- [Supplementaryfigure2.pdf](#)
- [Supplementaryfigure3.pdf](#)
- [Supplementaryfigure4.pdf](#)
- [SupplementaryFigure5.pdf](#)