UAB
Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona

Escola d'Enginyeria
Departament d'Enginyeria Química, Biològica i Ambiental
Programa de Doctorat en Biotecnologia

BIOPROCESS ENGINEERING FOR HIV-1 GAG VLP PRODUCTION IN HEK293 CELLS

Jesús Lavado García

Advisors

Dr. Francesc Gòdia Casablancas
Dr. Laura Cervera Gracia

Engineering Prof. Francesc Gòdia Casablancas and post-doctoral researcher from the Chemical, Biological and Environmental Engineering Department at Universitat Autònoma de Barcelona (UAB) PhD. Laura Cervera Gracia

Hereby certify:

That Jesús Lavado García, graduate in Biotechnology by Pablo de Olavide University (Seville) has successfully completed his PhD entitled "**Bioprocess engineering for HIV-1 Gag VLP production in HEK293 cells**" in the Chemical, Biological and Environmental Engineering Department at Universitat Autònoma de Barcelona under our supervision. This PhD thesis is hereby presented in this document in order to opt to the PhD title as Doctor in Biotechnology by the Universitat Autònoma de Barcelona.

Thereby and for the appropriate purposes, we sign this document in Bellaterra, May 2021.

Jesús Lavado García   Laura Cervera Gracia   Francesc Gòdia Casablancas
(author)   (co-director)   (co-director)

"I believe that if you, dear reader, can extend your patience for just a moment, you will find that there is a method to this tale of madness"

-Cloud Atlas

CONTENT

       Multiplexed quantitative proteomic analysis of HEK293 provides insights of molecular changes associated to the cell density effect, transient transfection and virus-like particles production.

       Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles.

       Culture media screening and selection of optimal transfection conditions. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures. Bioprocess analytics and perfusion-based continuous VLP harvest.

ABBREVIATIONS

AIDS: Acquired immune deficiency syndrome

APC: Antigen-presenting cell

ATF-CB: Combination of a continuous upstream process using ATF coupled with continuous capture and batch polishing

ATF: Alternating tangential flow

BHK: Baby hamster kidney.

BV: Baculovirus

CCD: Central composite design

CDE: Cell density effect

CHO: Chinese hamster ovary

COVID-19: Coronavirus disease

CSPR: Cell specific perfusion rate

DoE: Design of experiments

EGE: Extended gene expression

eGFP: Enhanced green fluorescent protein

EMA: European Medicines Agency

ER: Endoplasmatic reticulum

ESCRT: Endosomal sorting complex required for transport

EV: Extracellular vesicle

FB-CB: Combination of a fed-batch upstream process coupled with continuous capture and batch polishing

FDA: Food and Drug Administration

FucT: α1,3-fucosyltransferase

Gag::eGFP: Translational fusion of Gag and eGFP proteins.

GalNAc: N-Acetylgalactosamine

GlcNAc: N-Acetylglucosamine

HBsAg: Hepatitis B surface antigen

HCD: Higher-energy collisional dissociation

HEK: Human embryonic kidney

HFM: Hollow fiber module

HIB: Haemophilus influenza type B

HILIC: Hydrophilic interaction liquid chromatography

HIV-1: Human immunodeficiency virus 1

HPLC: High-performance liquid chromatography

Hpt: Hours post transfection

HPV: Human papillomavirus

ILV: Intralumenal vesicles

iTRAQ: Isobaric tags for absolute and relative quantification

LC-MS/MS: Liquid chromatography coupled to tandem mass spectrometry

LTR: Long terminal repeat

MALDI: Matrix assisted laser desorption/ionization

MDCK: Madin-Darby canine kidney

MHC: Major histocompatibility complex

MHRA: Medicines and Healthcare products Regulatory Agency

MOI: Multiplicity of infection

mPES: Modified polyethersulfone

MS: Mass spectrometry

MS/MS: Tandem mass spectrometry

MVB: Multivesicular bodies

NMT-1: N-Myristoyltransferase

nSMase: Neutral sphingomyelinase

PE: Polyethylene

PEI: Polyethyleneimine

PES: Polyethersulfone

PGC: Porous graphitized carbon

PRM: Parallel reaction monitoring

PS: Polysulfone

PTM: Post-translational modification

RMCE: Recombinase-mediated cassette exchange

RNAi: RNA interference

RSM: Reponse surface methodology

RV: Reactor volume

RV/day: Reactor volume per day

SARS-CoV: Severe acute respiratory syndrome coronavirus

SGE: Stable gene expression

shRNA: Short hairpin RNA

SIV: Simian immunodeficiency virus

SPE: Solid- phase extraction

SPR: Surface plasmon resonance

SU: Single-use

TFF: Tangential flow filtration

TGE: Transient gene expression

TMT: Tandem mass tags

ToF: Time of flight

UTR: Untranslated region

VCD: Viable cell density

VHU: Viral harvesting unit

VLP: Virus-like particle

XylT: β1,2-xylosyltransferase

SUMMARY

The importance of vaccine technology has recently become a matter of high interest due to the outbreak of the COVID-19 pandemic. However, the development of vaccine technology against infectious diseases throughout history has never stopped. It has influenced the way the population interacts and the establishment of social rules. The research in the field of vaccine technology is deeply tied to the research in bioprocess engineering, as in a globalized world, an effective vaccine is the one that can be produced at large scale, reaching as many people as possible through efficient vaccination policies. This PhD thesis is focused on the study of the bioprocess for the production of human immunodeficiency virus (HIV-1) virus-like particles (VLPs) based on the HIV-1 polyprotein Gag using HEK293 cells and transient transfection. These VLPs are non-infectious membrane-bound nanostructures whose conformation resembles the native HIV-1 virus albeit lacking the viral genetic material. In order to use these particles as potential vaccines, they can be later modified presenting exogenous antigens in their membrane to elicit an immune response against different diseases. In order to optimize VLP production, multiple aspects of the bioprocess must be studied, from molecular factors like cellular pathways to the modulation of operational parameters at bioreactor scale. In this work, this undertaken challenge can be divided in three main parts.

The first part, comprising chapters one and two, focused on the optimization of the VLP production process at cellular level through proteomics and metabolic engineering. Here, the effects of transient transfection and production of VLPs was studied at molecular level using multiplexed quantitative proteomics to identify metabolic bottlenecks influencing VLP production. Once these pathways were identified, metabolic engineering and design of experiments (DoE) approaches were applied to optimize transfection, cell budding efficiency and VLP production.

The second part, comprising chapter three, focused on the intensification and optimization of VLP production at bioreactor level, implementing a perfusion-based bioprocess. The culture media used, the DNA complexation agent and culture conditions in the bioreactor were studied. Also, operational parameters like the cell specific perfusion rate (CSPR), the amount of DNA used for transfection and the

time of retransfection were also optimized using a DoE approach. Moreover, different cell retention devices were tested to work towards the implementation of a continuous harvest mode of operation.

The third and last part, comprising chapters four, five and six, focused on the characterization of the product aiming to design an efficient downstream process. Parallel reaction monitoring (PRM) was used to establish a method of determining Gag stoichiometry in the VLPs for different production platforms, paving the way for improvements in quality control and process analytics technologies (PAT) in a future large scale bioprocess. Moreover, the coproduced extracellular vesicles (EVs) were studied through multiplexed quantitative proteomics to better understand the nature of the main source of impurities in a VLP preparation. Additionally, EV and VLP glycosylations were analyzed, to characterize the two nanostructures glycosylation signature and help gain some insight towards the design of a specific separation method for VLP purification.

The results hereby presented contribute to the development of a promising vaccine platform based on Gag VLPs and pave the way towards the advance and evolution of vaccine manufacturing.

RESUMEN

La importancia de las nuevas tecnologías de producción de vacunas se ha convertido en un asunto de gran interés recientemente debido a la aparición de la pandemia COVID-19. Sin embargo, el desarrollo de las tecnologías de producción de vacunas contra la afección de enfermedades infecciosas a lo largo de la historia nunca ha cesado. Este es uno de los hechos que ha influenciado la forma de relacionarse e interactuar de la población influyendo en el establecimiento de nuevas normas sanitarias y de conducta. La investigación en el campo de las tecnologías de producción de vacunas está altamente relacionada con la investigación en el desarrollo de bioprocesos, debido a que en un mundo globalizado, la eficiencia de una vacuna viene medida por la posibilidad de producirla a gran escala, llegando a la mayor cantidad posible de personas mediante políticas de vacunación efectivas. Esta tesis doctoral se centra en el estudio del bioproceso para producir partículas semejantes a virus (VLPs) del virus de la inmunodeficiencia humana (VIH) basadas en la poliproteína Gag del mismo virus, usando células HEK293 y transfección transitoria. Estas VLPs son nanoestructuras no infecciosas rodeadas de membrana celular cuya conformación se asemeja a la del virus nativo del VIH, pero sin su material genético. Posteriormente, la membrana que las rodea se puede modificar, añadiendo antígenos de otras enfermedades para estimular el sistema inmune y usarlas como potenciales vacunas. Para poder optimizar la producción de VLPs, es necesario el estudio de múltiples factores, que van desde rutas metabólicas a nivel celular hasta el diseño de las condiciones y parámetros de operación a escala de biorreactor. En el trabajo que se presenta en esta tesis doctoral, este reto puede dividirse en tres partes principales.

La primera parte, presentada en los capítulos uno y dos, se centra en la optimización del proceso de producción de VLPs a nivel celular, mediante proteómica e ingeniería metabólica. Aquí, se han estudiado los efectos de la transfección transitoria y la producción de VLPs a nivel molecular usando proteómica cuantitativa para identificar cuellos de botella en el metabolismo influenciando la producción de VLPs. Una vez identificados, se han aplicado estrategias de ingeniería metabólica usando diseño de experimentos (DoE) para optimizar la transfección y el proceso de generación de partículas.

La segunda parte, presentada en el capítulo tres, se centra en la intensificación y optimización de la producción de VLPs a nivel de biorreactor, para establecer un proceso en perfusión. Para ello se ha estudiado el medio de cultivo, los reactivos para formar los complejos de DNA y las condiciones de cultivo en el biorreactor. Además, usando estrategias como el DoE se han optimizado parámetros operacionales como el recambio específico de medio (CSPR), la cantidad de DNA usada para transfectar y los tiempos de retransfección. Asimismo, se han probado diferentes equipos de retención celular para llegar a implementar un sistema de perfusión en continuo.

La tercera y última parte, que comprende los capítulos cuatro, cinco y seis, se centra en la caracterización del producto, dirigida a diseñar un procesos de purificación eficiente. Para establecer un método analítico que determine la estequiometría de la proteína Gag en las VLPs producidas en distintos tipos celulares, se ha usado la técnica *parallel reaction monitoring* (PRM), contribuyendo a la mejora de futuros métodos de control de calidad y analíticos (PAT) en un futuro bioproceso a gran escala. Además, la vesículas extracelulares (EVs) que se coproducen con las VLPs han sido estudiadas mediante proteómica cuantitativa para entender la naturaleza de las que son las principales impurezas de las preparaciones de VLPs. De igual forma se ha caracterizado el perfil de glicosilaciones de las VLPs y las EVs para ampliar el conocimiento que ayude a diseñar un proceso de purificación específico para VLPs.

Los resultados que aquí se presentan contribuyen al desarrollo de una prometedora plataforma para la producción de vacunas basadas en VLPs de Gag y allanan el camino para el avance y la evolución del campo de la producción de vacunas.

RESUM

La importància de les noves tecnologies de producció de vacunes s'ha convertit recentment en un assumpte de gran rellevància a causa de l'aparició de la pandèmia COVID-19. No obstant això, el desenvolupament de les tecnologies de producció de vacunes com a mesura per a combatre les malalties infeccioses mai ha cessat. Aquest és un dels fets que ha influenciat la manera de relacionar-se i interactuar de la població i ha influït en l'establiment de normes socials. La recerca en el camp de les tecnologies de producció de vacunes està altament relacionada amb la recerca en enginyeria de bioprocessos, pel fet que en un món globalitzat, l'eficiència d'una vacuna ve donada per la possibilitat de produir-la a gran escala, arribant a la major quantitat possible de persones mitjançant polítiques de vacunació efectives. Aquesta tesi doctoral se centra en l'estudi del bioprocés per a produir partícules semblants a virus (VLPs) del virus de la immunodeficiència humana (VIH) basades en la poliproteïna Gag del mateix virus, usant cèl·lules HEK293 i transfecció transitòria com a plataforma productora. Aquestes VLPs són nanoestructures no infeccioses envoltades de membrana cel·lular la conformació de la qual s'assembla a la del virus natiu del VIH, però sense el seu material genètic. Posteriorment, la membrana que les envolta es pot modificar, afegint antígens d'altres malalties per a estimular el sistema immune i usar-les com a potencials vacunes. Per a poder optimitzar la producció de VLPs, és necessari l'estudi de múltiples factors, que van des de rutes metabòliques a nivell cel·lular fins al disseny de les condicions i paràmetres d'operació a escala de bioreactor. En el treball que es presenta en aquesta tesi doctoral, aquest repte pot dividir-se en tres parts principals.

La primera part, presentada en els Capítols ú i dos, se centra en l'optimització del procés de producció de VLPs a nivell cel·lular, mitjançant proteòmica i enginyeria metabòlica. Aquí, els efectes de la transfecció transitòria i la producció de VLPs ha estat estudiada a nivell molecular usant proteòmica quantitativa per a identificar colls d'ampolla en el metabolisme influenciant la producció de VLPs. Una vegada identificats, s'han aplicat estratègies d'enginyeria metabòlica usant disseny d'experiments (DoE) per a optimitzar la transfecció i el procés de generació de partícules.

La segona part, presentada en el Capítol tres, se centra en la intensificació i optimització de la producció de VLPs a nivell de bioreactor, establint un procés en perfusió. Per a això s'ha estudiat el medi de cultiu, els agents per a formar els complexos de DNA i les condicions de cultiu en el bioreactor. A més, usant estratègies com el DoE s'han optimitzat paràmetres operacionals com el recanvi específic de medi per cèl·lula (CSPR), la quantitat de DNA usada per a transfectar i els temps de retransfecció. Així mateix, s'han provat diferents equips de retenció cel·lular per a arribar a implementar un sistema de perfusió en continu. La tercera i última part, que comprèn els Capítols quatre, cinc i sis, se centra en la caracterització del producte, dirigida a dissenyar un procés de purificació eficient. Per a establir un mètode analític que determini l'estequiometria de la proteïna Gag en les VLPs produïdes en diferents tipus cel·lulars, s'ha usat la tècnica *parallel reaction monitoring* (PRM), contribuint a la millora de futurs mètodes de control de qualitat i analítics (PAT) per a un futur bioprocés a gran escala. A més, la vesícules extracel·lars (EVs) que es coprodueixen amb les VLPs han estat estudiades mitjançant proteòmica quantitativa per a entendre la seva naturalesa ja que són les principals impureses de les preparacions de VLPs. D'igual forma s'ha caracteritzat el perfil de glicosilacions de les VLPs i les EVs per a ampliar el coneixement que i permetre dissenyar un procés de purificació específic per a VLPs.

Els resultats que aquí es presenten contribueixen al desenvolupament d'una prometedora plataforma per a la producció de vacunes basades en VLPs de Gag i aplanen el camí per a l'avanç i l'evolució del camp de la producció de vacunes.

INTRODUCTION

1. Vaccine Development
    1.1 First Generation
          1.1.1 Live attenuated virus
          1.1.2 Inactivated virus
    1.2 Second Generation
          1.2.1 Subunit vaccines
          1.2.2 Conjugate vaccines
          1.2.3 Recombinant vaccines
    1.3 Third Generation
          1.3.1 DNA vaccines
          1.3.2 RNA vaccines

2. HIV-derived VLPs
    2.1 HIV-1 virus
    2.2 Gag VLP assembly and budding pathways
    2.3 The use of HIV-1 Gag VLPs
    2.3 Extracellular vesicles
    2.4 Gag VLP glycosylation

3. HIV-1 Gag VLP production
    3.1 Plant cells
    3.2 Insect cells
    3.3 Mammalian cells
        3.3.1 HEK293
    3.4 Production strategies for HEK293 cultures
          3.4.1 Stable gene expression (SGE)
          3.4.2 Transient gene expression (TGE)
          3.4.3 Extended gene expression (EGE)
    3.5 Continuous biomanufacturing of VLPs
          3.5.1 Perfusion for process intensification
          3.5.2 Perfusion rate
          3.5.3 Hollow fiber membranes
    3.6 Downstream processing of VLPs

4. Bioprocess Engineering

In 1796, deep in the English countryside, a physician was working in a devastating skin disease causing disfiguring scars, fluid-filled bumps, blisters and eventually death. Smallpox, a disease produced by the *Variola* virus, was a long-standing acquaintance for humanity and had caused an uncountable number of deaths throughout the world. The physician Edward Jenner was trying to ease the disastrous effects of the disease which had decimated England. It was him who first noticed a recurrent pattern among Gloucestershire farmers, more specifically among milkmaids. In a Europe ravaged by smallpox, these women seemed to maintain a clear skin, untouched by the disease. After strolling across the fields observing the milkmaids' environment and checking their condition, he suggested that they had contracted the cowpox, a milder and close-related variant of the smallpox. He proposed that having contracted the cowpox had conferred upon them immunity to the deadly smallpox[1]. Edward Jenner was also a scientist, so he intended to test his hypothesis. He took a young boy, James Phipps, and produced a small cut in his arm, where he injected pus from Sarah Nelmes, a milkmaid infected with cowpox. The boy developed a moderate infection from which he rapidly recovered. A few days later Jenner scrubbed some scabs from a smallpox-carrying patient and injected the material into the boy. No signs of the disease. After several attempts, the boy never got infected. This is the first reported immunization scientific study and led to the establishment of the term vaccine, from the Latin *vaccinus*, meaning from the cow[2]. Nevertheless, the old and classic medical book of Zhang Lu described immunization-like practices also focused on protecting people against smallpox in the Imperial China during the XVIIth century[3]. The smallpox was declared eradicated in 1979, being the first infectious diseases to disappear from the world thanks to purposefully human action[4]. The remarkable importance of this milestone in human history is currently more relevant than ever, when emerging anti-vaccine movements threaten decades of scientific progress. In January 2019 a measles outbreak, a disease nearly eradicated in developed countries, was reported in the United States due to anti-vaccine viral messages on social media, spreading missinformation[5]. Even now, when spread of COVD-19 is representing a turning point in global healthcare management, innovation and development, the anti-vaccine movement threatens global heath and recovery from the current pandemic. The emergence of the hitherto unknown severe acute respiratory syndrome coronavirus 2 or SARS-CoV-2 in Wu Han, China and the respiratory disease it caused, COVID-19,
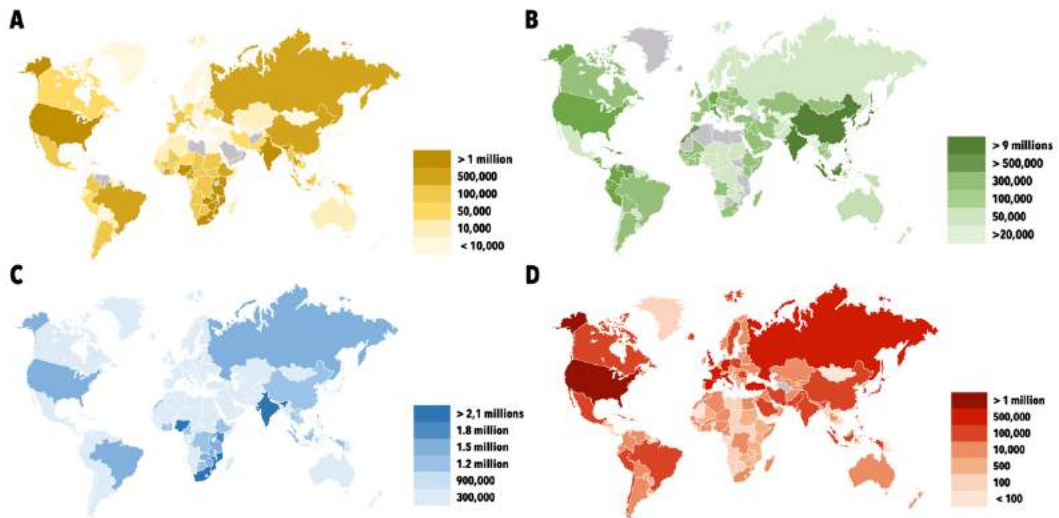
Figure 1. World distribution of the most relevant pandemics. **A)** Affected global area by smallpox infection until its eradication in 1979. **B)** Distribution of Spanish flu (H1N1) during 1918-1919 causing an estimated total death of 39.3 million people. **C)** Prevalence of HIV-1 in the population in 2005, when this pandemic reached the peak of infection. **D)** Total confirmed COVID-19 cases by 1st May 2020. Grey represents no available data.

became the world's most recent pandemic. Causing a global sanitary, social and economic crisis forcing governments all over the world to implement home confinement resulting in half of humanity locked down at home[6]. All efforts were relentlessly focused towards the development of a vaccine. By January 2021, the scientific community was working on 93 different SARS-CoV-2 vaccine candidates all over the world. The importance of vaccine research has never been clearer. Pandemics have constantly affected the world and among the most relevant, smallpox, the Spanish flu (H1N1)[7] or the human immunodeficiency virus 1 (HIV-1)[8] (**Figure 1**) had such remarkable effect on society that contributed to profound changes in our lifestyle, improving sanitary conditions, or implementing new habits and conducts to exert protection such as raising awareness of sexual health due to HIV-1 outbreak[9]. The observed death toll trend has been reduced over time, from 300 millions of estimated deaths caused form for smallpox[10], to roughly 2 million by January 2021 in the current COVID-19 outbreak[11]. This reduction is due to healthcare improvements, medicine advances and in long term to increasing human knowledge along in vaccine technology and thus, allowing a stronger and more prepared response, which represents one of the greatest accomplishments in the history of medicine and human progress[12]. This response consisted in the

development of better vaccines in a shorter period of time with high efficacy and safety, the deployment of industrial production for manufacturing a large number of doses and the establishment of vaccination programs. Throughout history, great progress has been made from the first injections of pus-like material to genetic engineering, industrial screenings and controlled culture conditions for the development of safer, more stable and more efficacious vaccines.

1. Vaccine development
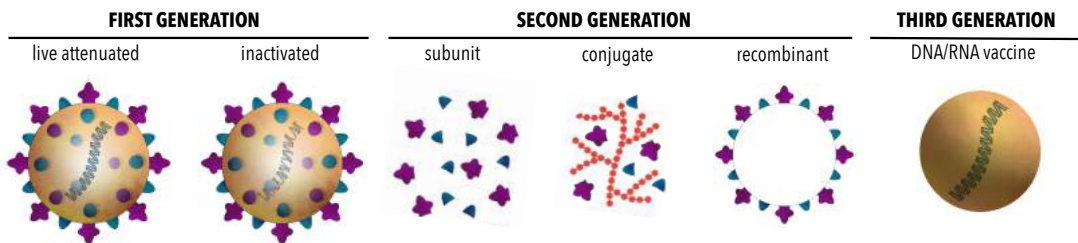
Figure 2. Vaccine generations throughout their development.

Along the process of vaccine development, vaccines were organized into three different generations according to their characteristics and technology used to create them (**Figure 2**).

1.1 First generation

The first generation of vaccines are live attenuated or inactivated viruses. These are conventional vaccines, made from a less virulent variant of the specific pathogen we want to be protected against. They are highly immunogenic and elicit a strong immune response for long term protection. They are easy to produce therefore, their production costs are relatively low[13,14]. The production of this type of vaccines requires additional biosecurity measures to avoid a potential infection in the manufacturing process.

1.1.1 Live attenuated

For this type of vaccines, natural weak pathogenic strains can be used, like cowpox for smallpox immunization. However, the process of attenuation could be performed in controlled conditions in the laboratory, via subsequent rounds of replications in a series of cellular passages to introduce mutations or exposing the pathogen to extreme conditions like cold[15]. Currently genetic engineering is also being used to design weaker strains[16]. Vaccines against measles, mumps, rubella (the triple vaccine), influenza, yellow fever, and hepatitis or poliovirus are different examples of live attenuated viruses. They are highly effective, so much that from the three different strains of poliovirus, type 2 was declared eradicated in 2015, type 3 in 2019 and by April 2020 there were only 48 reported cases in the world of type 1[17,18]. The recurrence of vaccine rejection and hesitancy could threaten this huge success[19].

1.1.2 Inactivated viruses

Complete inactivation of the virus led to safer immunization candidates, but still entailing the risk of reactivation and the consequent development of the disease. Also, strict inactivation meant faster clearance of the inoculated agent in the body, as it does not replicate. Reducing exposure time subsequently reduced the immunogenic capacity in some cases[20]. Inactivation can be achieved by exposure to heat, or chemical compounds like formaldehyde[21]. Examples of this type of vaccines also include influenza and polio vaccine as well as leishmaniasis, cholera or pertussis.

1.2 Second generation

However, being composed mainly of the native virus bears the possibility of developing symptoms, the infection and finally the disease[22]. In order to avoid these risks, vaccines composed of subunits, recombinant elements and protein antigens avoiding viral genetic material, were envisaged[23]. These are known as second-generation vaccines. Although overcoming some of their predecessor's drawbacks, their production costs are much higher, since more complex production conditions are required. Nowadays these type of vaccines have proven to be highly effective and are the most used for the prevention of

diseases such as hepatitis B, influenza, human papillomavirus, herpes virus, diphtheria, pertussis, anthrax and foot-and-mouth disease among others[20]. Currently, many efforts are being made to reduce their production costs and promote their distribution reaching further[24].

1.2.1   Subunit vaccines

In order to reduce production costs and still maintain a high immunogenic response, subunit vaccines were developed. Individual antigenic proteins from the disease-causing virus or bacterial toxins previously inactivated by chemical treatment are the most common. The latter are known as toxoids, coming from bacterial exotoxins inactivated by formaldehyde, which causes cross-link bonds in the toxin rendering it unable to interact with cellular receptors and cause disease[25]. These vaccines have been successfully implemented for cholera, tetanus or diphtheria. However, long term humoral immunity is not well achieved by this type of vaccine[26].

1.2.2   Conjugate vaccines

Some infections are driven by interactions with polysaccharides from the host. This represented a problem for long term immunization against these diseases, since polysaccharides do not interact with antigen-presenting cells (APC) and interact straight with B-cells to produce a rapid and restricted antibody response. Since they follow a T-cell-independent pathway, immunological memory is not created. This is why polysaccharides are considered weak antigens. In order to design an effective vaccine eliciting a potent and long-lasting immune response, polysaccharides are usually combined with proteins, which are excellent antigens and can be used as a carrier. Glycoconjugate vaccines or conjugate vaccines are composed of polysaccharides and antigenic proteins creating a strong immune response. Haemophilus influenzae type b (HIB), meningococcal or streptococcus pneumonia are examples of this type of vaccines[27].

1.2.3   Recombinant vaccines

The development of molecular biology, genetic engineering and recombinant DNA technology led to the transfection of heterologous antigen-coding DNA in known, modified and optimized host organisms, like

Escherichia coli, *Pichia pastoris* or animal cells, producing the antigenic proteins at a low cost that could be later purified and employed as a vaccine[28,29]. These antigenic proteins can be used as subunit vaccines but interestingly, some viral structural proteins can self-assemble and oligomerize in host organisms independently of the rest of the viral components forming a structure that resembles the native form of the virus but lacks genetic material. These structures are known as virus-like particles (VLPs). Unlike subunit vaccines or polysaccharides, these three-dimensional structures comprise a repetitive protein pattern highly effective for recognition by the immune system[30]. The first approved VLP vaccine was against hepatitis B, composed of hepatitis B surface antigen (HBsAg) VLPs. Research on VLPs rapidly took off and many capsid and structural proteins from many viruses were reported to be able to self-assemble and form VLPs. Vaccines against human papillomavirus[31], polyomavirus[32], parvovirus[33], hepatitis E[34], influenza H5N1[35,36], influenza H1N1[37] influenza H3N2[38] and even SARS-CoV-1[39] among others are either commercialized, in preclinical, clinical trials or haven been successfully demonstrated at research level. VLPs can be made of one protein or the combination of two, three and up to four proteins. A relevant type of VLPs are enveloped VLPs, which carry an external lipid envelope, part of the cell membrane from the producer cell. In order to produce enveloped VLPs, like HIV-1 VLPs, different cell lines can be used, containing the right cell machinery, membrane proteins and post-translational modifications. Plant[40], insect[41,42] or mammalian[43,44] cell-based platforms are usually employed.

1.3 Third generation

There is a third generation of vaccines, known as genetic or DNA vaccines. Unlike any other type of vaccine, these are composed of genetic material, usually DNA, codifying for antigenic proteins so that the immune system can generate an immune response against the foreign proteins synthesized by our cells.

1.3.1   DNA vaccines

DNA vaccines are successful in veterinary applications, for instance, against canine melanoma or the West Nile virus in horses. Human DNA vaccines have been recently approved for human use, like COVID-19 Oxford-AstraZeneca adenovirus vaccine. They achieve immunization by *in vivo* transiently transfecting APC

with foreign DNA, inducing the production of the desired antigen which is subsequently presented in the MHC-I and MHC-II. This generates CD8+ and CD4+ T cells responses in parallel, leading to B cell interaction and the consequent antibody production against the desired antigen. Another pathway to reach this effect is transfecting keratinocytes or myocytes, which by the production of extracellular vesicles or apoptotic bodies target the plasmids to APC[45,46]. The main concern for the use of this type of vaccines is the potential random genome integration of the foreign DNA, causing mutations and leading to tumours and dysregulation of genomic expression. A DNA vaccine was the first prophylactic candidate against HIV-1 to reach phase 3 in clinical trials. This mosaic vaccine entered a large-scale phase 3 efficacy study in 2020, promoted by the pharmaceutical company Janssen Vaccines & Prevention B.V.[47,48].

1.3.2 RNA vaccines

The state-of-the-art research in the field of genetic vaccines is led by RNA vaccines. Circumventing the issues raised by DNA vaccines, they are usually presented in the form of mRNA which can be non-replicating or self-amplifying RNA[49]. These self-amplifying RNA sequences contain, in addition to the gene of interest encoding the desired antigen, the genes coding for the RNA replication machinery. Higher number of RNA copies improve antigen expression and also triggered both cellular and humoral response, as they mimic the metabolic behaviour of a virus infection. RNA can therefore be engineered to enhance stability and expression levels, adding different promoters, usually T7, 5' and 3' untranslated regions (UTRs), poly(A) tails and 5' caps[50]. To achieve a pure enough preparation of RNA, purification protocols are established based on high-performance liquid chromatography (HPLC) to remove contaminants and double strand RNA (dsRNA)[51]. Due to the fragile stability of RNA and the ubiquitous presence of ribonucleases, RNA strands are often delivered using lipid-bound vesicles as carriers. A lipid membrane allows the particle to easily enter the cell and deliver the cargo[52]. The lipid-bound vesicles used for this task can be synthetic lipid vesicles or extracellular vesicles (EVs) produced by any cell type and whose biological function is cell-to-cell communication[53-55]. RNA delivery can be used not only for vaccines, but also for many other types of therapeutics, loading different kinds of siRNA, miRNA onto EVs. This opened a whole new field for mRNA and delivery optimization[56]. On the 02/12/2020, the first RNA vaccine for human use

against SARS-CoV-2 was approved by the Medicines and Healthcare products Regulatory Agency (MHRA) of the United Kingdom, followed by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Astonishingly developed in 9-10 months as an emergency response to the COVID-19 pandemic.

2. HIV-derived VLPs

One of the HIV-1 components, Gag polyprotein, is one of the viral structural proteins able to self-assemble and form recombinant VLPs. The importance and overwhelming effects of the HIV-1 pandemic started with the first case of death by HIV-caused acquired immune deficiency syndrome (AIDS), reported in 1959. Since then, it has driven many efforts in the understanding of the mechanics of this virus[57]. The mechanisms taking place in HIV infection and replication have been deeply characterized. Industriously, this knowledge has allowed the creation of a Gag-VLP based platform for vaccine therapies.

2.1 HIV-1 virus - Infection cycle and its use in vaccine technology

HIV-1 is a lentivirus belonging to the *Retroviridae* family. It was reported to be transmitted from simians to humans in the early 1910's and to be derived from the simian immunodeficiency virus (SIV)[58]. Its relevance lies within the devastating disease it causes, AIDS, which is responsible for a total of over 37 million deaths and around 38 million people living with HIV-1 worldwide by 2019[59]. AIDS undermines the immune system since the virus replicates within CD4+ T lymphocytes, progressively reducing their number, rendering the patient vulnerable to opportunistic diseases that eventually lead to death, like tuberculosis[60]. Its genome is composed of single-strand positive RNA encoding 9 genes: *gag, pol, env* (structural), *tat, rev* (regulatory), *vif, vpu, vpr and nef* (accessory). The gene *gag* encodes Gag polyprotein (p55), composed by six different parts: matrix protein p17 (MA), capsid protein p24 (CA), spacer peptides (SP1, SP2), nucleocapsid protein p7 (NC) and p6[61]. The gene *pol* encodes three enzymes: a retrotranscriptase (RT), an integrase (IN) and a protease (PR). The gene *env* encodes the surface proteins gp120 and gp41. Tat and Rev proteins are involved in activation of transcription and nuclear export of viral RNA to the cytoplasm
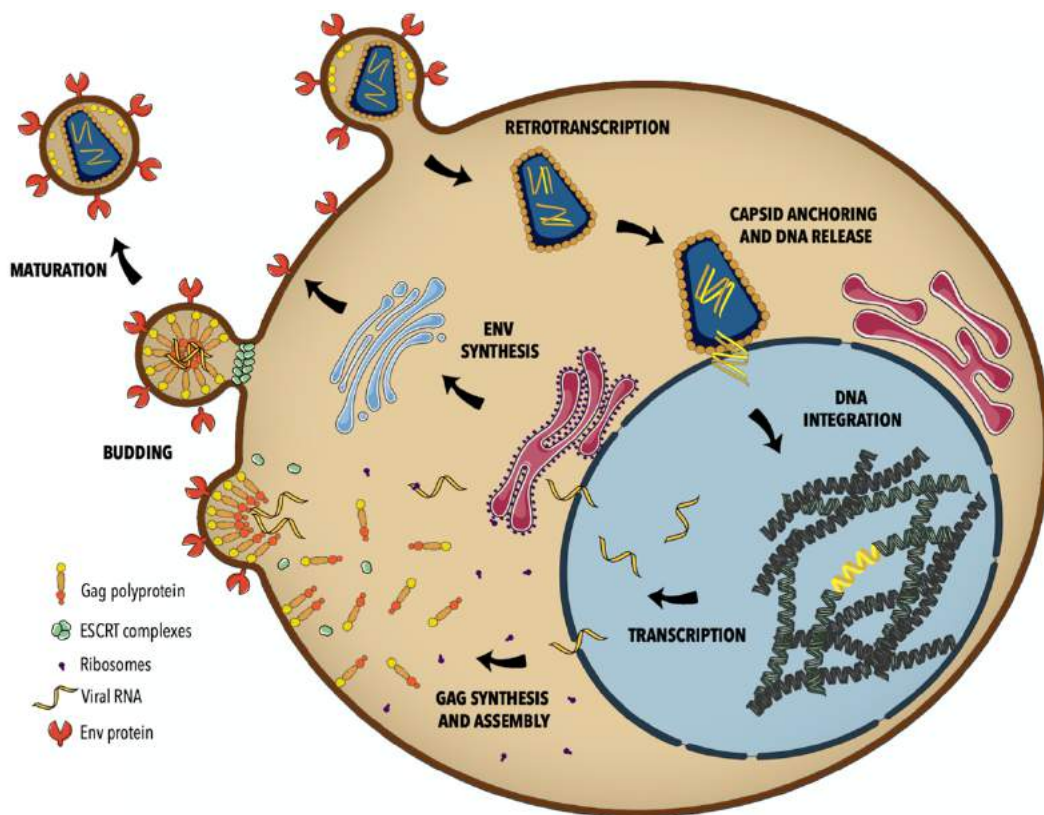
Figure 3. Representation of HIV-1 life cycle.

respectively. Vif inhibits the APOBEC, interfering with DNA retrotranscription[62], Vpr inhibits the cellular mechanisms of foreign DNA silencing and arrests the cell cycle at G2/M[63]. Vpu is involved in the trafficking of new CD4 to the membrane and Nef inhibits the renewal and presentation MHC-I, also degrading pre-existing CD4[64]. Nef has also been reported to promote cell-to-cell transmission of the virus in infected tissues mediated by exosomes[65,66].

The HIV life cycle starts when the envelope glycoprotein Env (gp120 and gp41) interacts with CD4 receptor and CCR5 or CXCR4 co-receptors in CD4+ T lymphocytes. A conformation change brings the enveloped virus and the host cell membrane together allowing their fusion. The viral capsid enters the cell and rapidly the retrotranscriptase starts synthesizing cDNA from the viral RNA+ strand. The capsid is transported via microtubule networks to the nucleus where it uncoats, allowing the cDNA to enter the nucleus. Early uncoating leads to abortive infections[67,68]. Once in the nucleus, cDNA integrates in long terminal repeat
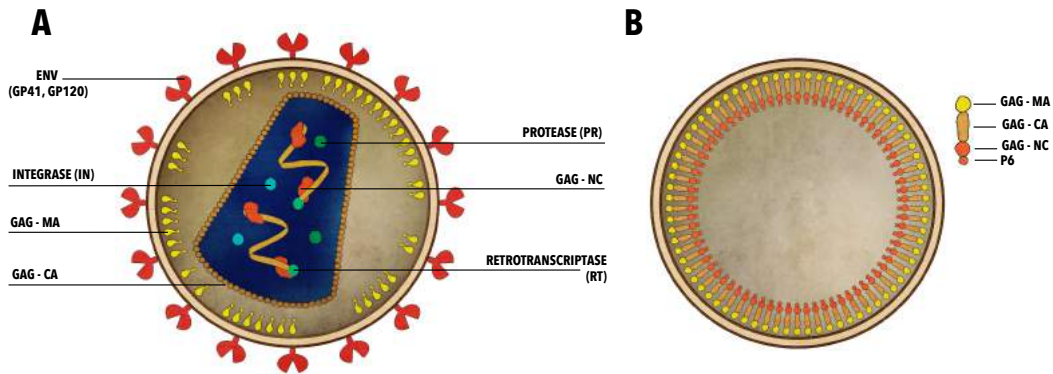
Figure 4. A) Structure of a HIV-1 native virus **B)** Structure of a HIV-1 Gag VLP

(LTR) sequences by the integrase coded by *pol* gene. Here it can remain dormant as a provirus for years, known as latency. When transcribed, mRNA is spliced for all the different proteins mentioned above. Gag MA domain undergoes a post-translational modification and N-terminal myristate groups are added which target Gag to the cell membrane[69]. Gag polyprotein then travels to the cell membrane, oligomerizing in areas containing microdomains enriched in lipid rafts and cholesterol[69,70]. Gag assembly induces the coalescence of tetraspanin-rich domains, containing proteins like CD81, CD63 or CD9 and glycosphingolipids like ceramide[71-73]. Gag assembly oligomerization, and later scission is carried out by the endosomal sorting complexes required for transport (ESCRT)[74-77]. Two motifs from Gag p6 interact with TSG-101 from complex ESCRT-I and ALIX[78], the coordinator protein between ESCRT-I and ESCRT-III. This interaction leads to the recruitment of subunits of ESCRT-III[79] that are placed to form a ring that strangles the plasma membrane, forming the new out-coming enveloped particle. ESCRT-III is mainly responsible for scission[76]. The final step to release the new particle is carried out by the effector protein Vps4 which upon ATP hydrolyzation performs the cleavage[80]. The role of ESCRT-II is still uncertain as it appears to be not involved in HIV budding[81], although some examples proving its requirement for the process have been shown[82], its role remains unclear. Alternative accessory proteins also play a role in sending Gag to the membrane. It is the case of the ubiquitin ligase NEDD4L[83,84]. Ubiquitinated proteins interact with domains of proteins from ESCRT-0 and might interact with ESCRT-II in the process of Gag transport to the membrane[85]. N-Myristylation of Gag MA has also been reported to have targeting effects[78]. The assembly

of Gag in the plasma membrane is accompanied by the recruitment of viral proteins and viral RNA to be encapsulated. Viral RNA interacts with NC domain of Gag polyprotein and it is described to serve as scaffold for virion formation[86,87]. After excision Gag polyprotein suffers a process of maturation comprising proteolytic cleavage of its parts by the protease encoded in *pol* gene[88]. The reorganization forms the characteristic cone-shaped capsid protecting the viral RNA and leaves an outer shell of MA subunits below the plasma membrane. The new HIV-1 virions of ~145 nm[89] are ready to begin the cycle again (**Figure 3**). The process of Gag assembly, oligomerization and budding from the membrane are triggered by the mere presence of Gag polyprotein which can self-assemble independently of the rest of the viral proteome. This led to the expression of Gag by recombinant DNA technology in different cell lines to produce Gag VLPs. A platform with high potential for vaccine development. Gag VLPs have the same size of ~145 nm but differ from HIV-1 virions not only because they do not contain viral genetic material, therefore being non-infectious like all recombinant vaccines, but also in their protein structure (**Figure 4**). Gag VLPs lack the proteolytic maturation of Gag polyprotein, forming a shell underneath the cell membrane conforming the protein core.

2.2 Gag VLP assembly and budding pathways

Gag VLPs assemble in cell membranes microdomains enriched in lipid rafts containing cholesterol and ceramide and tetraspanins like CD81, CD63 and CD9. These features are present not only in the cell membrane, but also in membranes of intracellular compartments or endosomes. These endosomes originate from the cell membrane by a process of endocytosis[90], therefore sharing the above mentioned membrane features. The originated vesicles gather together creating what is called early endosome. This intracellular compartment has different functions. It can be used to recycle plasma membrane proteins, so small scissions of this early endosome go back to the plasma membrane and fuse, placing new proteins at the cell surface. This mechanism is widely used in nature for maintaining homeostasis and controlling protein display at the plasma membrane, like GLUT4 in the regulation of glucose uptake by insulin[91]. Another function of the early endosomes is to undergo a series of changes and become late endosomes, also known as multivesicular bodies (MVB)[92]. The membrane of the MVB can invaginate towards the lumen

of the endosome forming intralumenal vesicles (ILV)[93]. The process of ILV formation is similar to VLP budding as both originate from a cell membrane, budding off into extracellular o intraluminal space taking a part of the cell membrane with them. There are different pathways for the generation of ILVs. The most described pathway is using the ESCRT, similarly to Gag budding process[94]. The same mechanics of VLP budding in the cell membrane govern this process. However, interestingly there are ESCRT-independent pathways for the formation of ILVs. Therefore, opening new pathways for molecular VLP production intensification. In this work, these pathways have been explored (**Figure 5**). One alternative pathway is via CD63, where this tetraspanin has been described to promote the loading cargo, release and formation of ILVs independently of ESCRT[95]. Another pathway uses the ceramide-enriched microdomains in the cell membrane to generate ILVs mediated by the enzyme sphingomyelin phosphodiesterase[96].

2.3 The use of HIV-1 Gag VLPs

Along the process of developing VLPs for vaccine therapies, the fact of incorporating part of the host cell membrane allows the further functionalization of the particle (also known as pseudotyping), adapting it for the presentation of different antigens (**Figure 6**). Example of these hybrid or chimeric VLPs are Influenza Gag VLPs, H1N1[97] or H5N1[98]. There are also reported Gag VLPs specifically pseudotyped for dengue, west Nile virus, HPV, equine herpes virus, rabies and pseudorabies[99-103]. The expression of proteins or *in vitro* addition of peptides, like using click chemistry methods[104], could target the VLPs to specific tissues. This led VLPs to harbor the potential to also act as RNA vaccines, as Gag packs cellular RNA in the process of VLP formation[105]. Leading to the use of a VLP-based platform also for drug delivery[106]. Modified Gag VLPs expressing the enzyme Cre-recombinase or even IFN-γ have been designed and successfully tested for targeted protein delivery[107]. Their versatility has been also exploited for cancer therapies, demonstrating to be able to elicit the CD8+ cytotoxic T-cell response and neutralizing antibodies against tumors, showing the great potential of this platform[108,109]. These traits made HIV-1 Gag VLPs a promising alternative for the design and development of a SARS-CoV-2 vaccine, which is already ongoing.
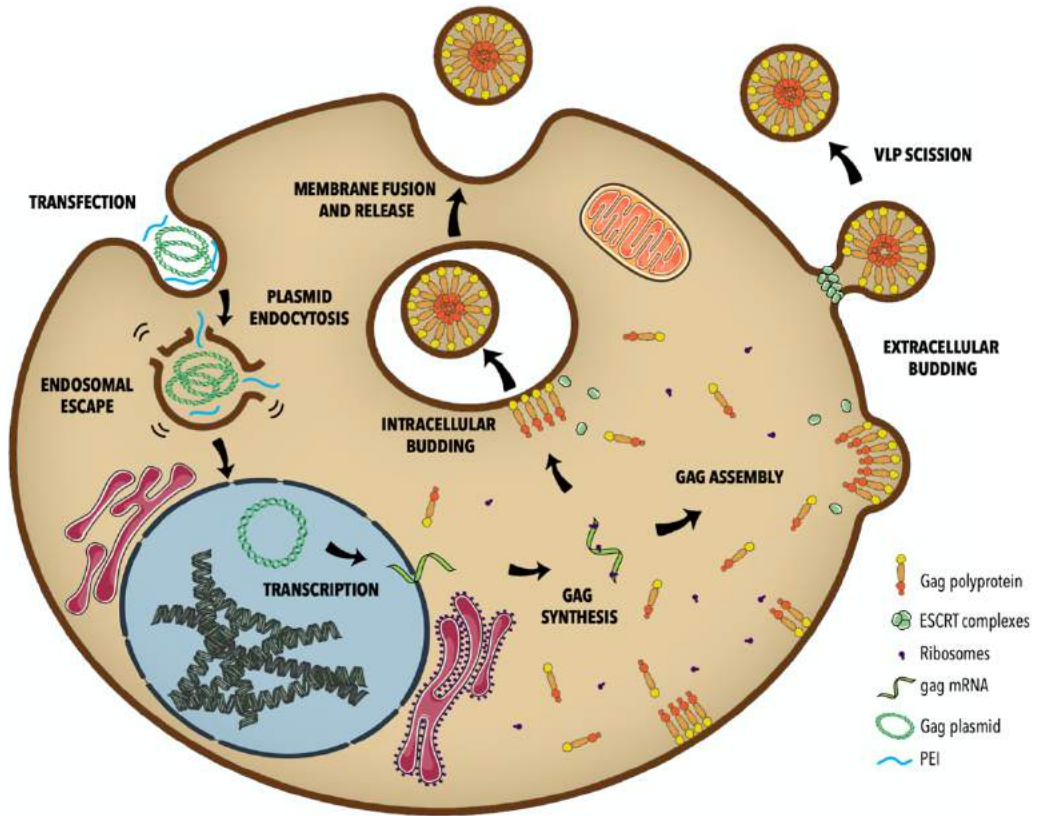
Figure 5. Representation of Gag VLP assembly and production pathways.

2.4 Extracellular vesicles

In order to create vaccines that are suitable for commercialization, VLP preparations need to attain high levels of purity. To successfully separate VLPs from the rest of biological by-products they need to display differential physicochemical properties for the separation methods to be based on. However, viral mechanisms use cellular machinery and pathways for natural production of extracellular vesicles. These microvesicles and exosomes are membrane-bound vesicles sharing the same density and physicochemical properties than VLPs, hindering the following separation downstream processes. Although this represents a major issue in bioprocess development, increasing knowledge in what kind of EVs are coproduced with VLPs would provide the insight to engineer and use them as a tool in the design of a VLP-based therapy. The characterization of EVs is further explored in this work. This offers as well the possibility of using EVs or functionalizing their membrane as adjuvants.
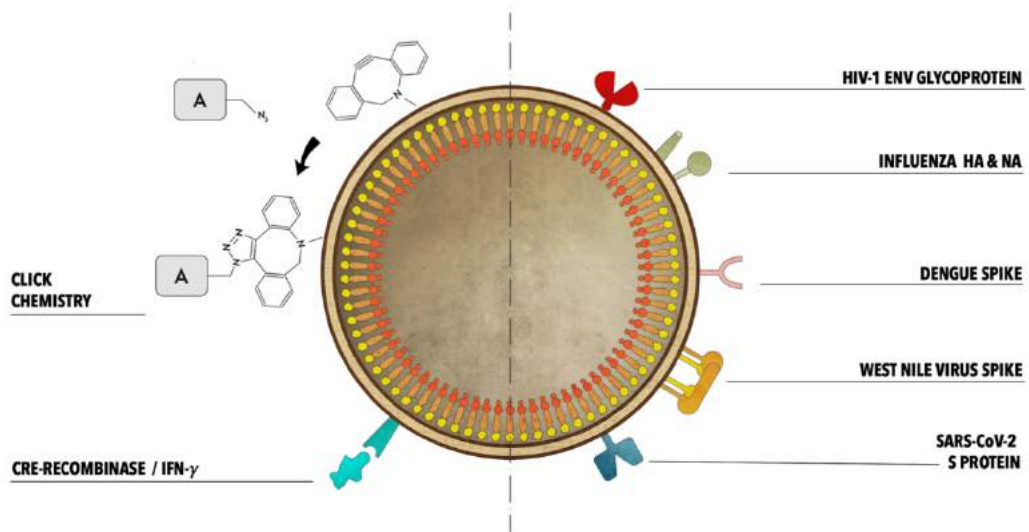
Figure 6. Gag VLP pseudotyping. On the right, membrane modification with surface proteins from different viruses. These modifications include monomers like HA and NA from influenza or dimer structures like the west Nile virus spike. This platform offers potential to harbor the SARS-CoV-2 S protein to develop a COVID-19 vaccine. On the left, strategies to modify the membrane in order to include peptides, enzymes or different compounds denoted as "A".

2.5 Gag VLP glycosylation

Glycans are structural components conjugated to lipids or proteins, in the form of glycoproteins present in many biological components. They are present in the membrane of VLPs and EVs uniquely designed for specific interactions, working as information carriers. They are involved in the quality control of protein folding, protein-protein interactions, subcellular targeting, cell-cell recognition events and communication. The capacity of these molecules to modulate molecular interactions is given by their complex structure. Lipids and proteins can be modified with branched carbohydrates (glycans) with a structured pattern of repetition allowing a singular conformation depending on their molecular function. Their function varies depending on the size of the conjugated carbohydrate structure, its biochemical composition, the repetitions found in the pattern and the type of bonds found in the branching. Glycans and proteoglycans are used for recognition at cellular level, to control intracellular trafficking and to maintain the quality and integrity of biomolecules, such us antibodies. Therefore, even when the
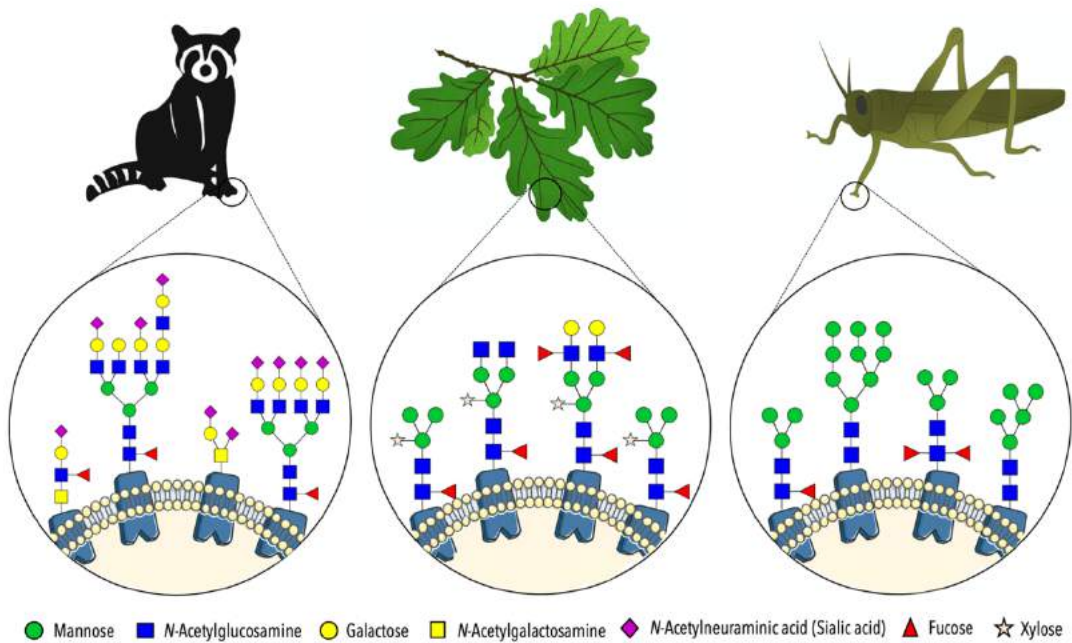
Figure 7. Different types of glycosylation patterns in mammalian, plant and insect cells. Mammalian glycosylation is characterized for the N-acetylneuraminic acid capping and only α1,6-fucosylations . Plant glycosylation is characterized by the presence of α1,3-fucosylations and β1,2-xyloses. Insect glycosylation is characterized by the presence of α1,3-fucosylations and high mannose glycans.

physicochemical properties of EVs and VLPs are similar, their glycan and proteoglycan patterns may introduce relevant differences to both populations of biomolecules. There are two major types of glycosylation, N-linked and O-linked glycosylations depending on the aminoacidic motif where it is attached to. N-linked glycosylation comprises a paucimannose core attached to the amide nitrogen of asparagine in the motif Asn-X-Ser/Thr and complex branching of different sugar moieties depending on the eukaryotic organism (**Figure 7**). O-linked glycosylation takes place when the sugar molecules, usually N-acetylgalactosamine (GalNAc) and N-acetylglucosamine (GlcNAc), are attached to the oxygen of a Ser or Thr. In humans, regardless of the type of glycosylation, branched GlcNAc and GalNAc are capped by N-acetylneuraminic acid or sialic acid. The pattern of branching, presence and abundance of one glycosyl species or another strictly depend on the organism. Protein glycosylation is also a key factor for viral infection and the relation of different viruses with our immune system. Many viruses, like HIV-1, are

enveloped by a part of the host cell plasmatic membrane, taking with it many glycoproteins that allow the virus to interact with the receptors of the target cell and the immune system. The viral structure takes advantage of the complex system of these information-carrying molecules and have used them throughout evolution as a part of their tools to keep changing and adapting to new hosts[110]. For instance, the HIV-1 envelope protein (Env, gp120) is one of the most heavily glycosylated proteins ever studied. The high frequency of mutation, constantly altering the *N*-glycan branching pattern associated to this protein was suggested to be one of the mechanisms of HIV-1 from escaping neutralizing antibodies while still having the ability to bind receptors and cause infection[111].

The glycosylic characteristics of Gag VLPs are investigated in this work and their potential differences with the rest of EVs may contribute to design specific separation methods and allow obtaining higher purity VLP preparations. Lectin microarrays and high-resolution mass spectrometry (MS) techniques can be applied for the analysis of the glycome of Gag VLPs and EVs[112].

3. HIV-1 Gag VLP production

In 1931, a new era in vaccine mass production began, based on a principle discovered by Louis Pasteur in 1880. He discovered that a pathogen replicated within an alternative host species, resulting in live-attenuated viruses that could be used as vaccines[113]. This principle led scientists to develop a robust platform for live-attenuated vaccine production: embryonated chicken eggs. The inoculation of viruses in the allantoic cavity of developing chicken eggs allowed the replication of the virus and posterior harvest, obtaining high titers[114]. This method allowed the first vaccine mass productions against yellow fever in 1937[115] and influenza in 1943[116]. Although obtaining an excellent yield, embryonated chicken eggs as vaccine platforms required long time (13-14 days for batch), high cost in infrastructure due to a liner scale-up and came at the expense of animal welfare. Due to the success of this technique for production of live-attenuated and inactivated viruses, a method using *in vitro* chicken embryo fibroblasts cultures was later developed[117]. Nowadays, both embryonated chicken eggs and *in vitro* production using cell lines are being used for vaccines mass production. The choice depends on the characteristics of the type of vaccine that is going to be produced and the advantages offered by each system in each specific case.

In order to produce recombinant HIV-1 Gag VLPs and to be able to achieve mass-production, different approaches must be undertaken. Gag VLP production requires the producer cell to be able to perform cell budding in order to generate enveloped particles. Other important requirements are the correct protein folding, the right coordination for oligomerization of protein assembly and the proper post-translational modifications (PTMs). PTMs are the fingerprint of the producer cell, as proteins from insect cells will have different species of oligosaccharides decorating their proteins compared to human cells (**Figure 7**). The mere compliance of these requirements discards bacteria platforms. Although recent findings describe some species being capable of performing some PTMs like phosphorylation, acetylation or even glycosylation, these PTMs are significantly different from the ones found in eukaryotic cells[118,119]. Yeasts are another efficient production platform. However, they cannot perform cell budding due to their cell wall. At research scale, removing the cell wall allowed the generation of enveloped VLPs[120]. Nevertheless, yeast

PTMs are also considerably different from human's, thus envisaging mass production renders this platform inviable. Due to these different factors, Gag VLPs are usually produced in plant, insect or mammalian cells.

3.1 Plant cells

Plant cells offer the advantage of performing PTMs similar enough to human PTMs. Engineered plants like *Arabidopsis thaliana* can produce human-like glycosylation, lacking xylose and core α1,3 fucose (**Figure 7**), upon genetic modification to supress β1,2-xylosyltransferase (XylT) and core α1,3-fucosyltransferase (FucT) activity[121]. Plant cells can be genetically modified using *Agrobacterium tumefaciens*-based expression systems. This bacterium infects the plant and transfer the so-called Ti plasmid to plant cells. This plasmid can be engineered to achieve transient gene expression or to generate stable transgenic plants. *Nicotiana benthamiana* or *Nicotiana tabacum* are among the most used species to produce Gag VLPs. Nonetheless, the characteristic feature of having cell wall makes plant cells a singular production platform since VLPs remain trapped in the apoplastic space between the cell wall and the plasmatic membrane. In order to harvest VLPs from the culture media, removal of the cell wall is necessary, producing protoplasts that can bud off VLPs directly to the extracellular medium. This requires high extraction and downstream processing costs. Additionally, this platform showed low productivity and low yields, achieving a maximum of 22mg/Kg when gp41 was coproduced with Gag[40]. A different approach using plant cells is to produce peptides or fragments of MA-CA or only the CA domain aiming for a subunit vaccine. The yields from this approach ranges from 0.5-50 μg/kg[122]. These subunit vaccines are conceived as edible vaccines, made to be accumulated in the plant tissue for later human consumption. Edible vaccines offer better distribution because they do not need to maintain the cold chain and could reach further into developing countries. However, they present issues regarding standardization of the vaccine concentration in the plant tissue, leading to difficulties in the establishment of dosage[123]. It is difficult to standardize not only between tissues, but also between plant parts (leaf, fruit), and between plant generations, rendering mass production highly difficult[124]. Moreover, if cooked, vaccine concentration is reduced, posing a problem for plants like potatoes, which are great protein producers[124]. A different problem is gastric denaturalization and inactivation or the appearance of immune tolerance for the specific proteins since low concentrations

in the plant would lead to eat more of them to be able to reach the dose[125]. However, possibly another problem for this approach would be soil intensification and reducing field area for farming, posing a dilemma when balancing priorities for crop seeding.

3.2 Insect cells

Insect cells are a well-studied platform for Gag VLP production. They can grow in suspension cultures, releasing VLPs to the extracellular medium. Therefore, the corresponding production bioprocesses can be scaled-up to meet mass production demands, being easily transferred to GMP manufacturing. However, their glycosylation pattern differs from human cells (**Figure 7**). Insect moieties are composed of paucimannosidic residues (highly-branched mannose) and present $\alpha 1,3$ fucose residues[126] whose recognition by the immune system could interfere with the expected immune response as fucosylation is critical in cell-cell interaction and recognition events[127]. Two insect cell lines are widely used for VLP production, Sf9 (*Spodoptera frugiperda*) and High Five (*Trichoplusia ni*), being the latter known for high-yield recombinant protein production in larvae tissues[128]. VLP production in Sf9 and High Five cell lines can be achieved by baculovirus (BV) infection or by generation of a stable cell lines. In spite of high titers being obtained when using BV-based systems, BV capsids remain in the media together with VLPs, hindering posterior purification downstream processes and increasing the costs when choosing these systems. The combination of multiplicity of infection (MOI) and cell density are two key parameters that need to be optimized when working with BV-based expression system in insect cells to maximize VLP production. This is due to the fact that there is no linear relation with MOI, cell density and final VLP concentration. Therefore, prior optimization depending on the production platform is needed. In order to generate stable cell lines, *piggyBac* transposon-based system, Insect Genjuce® or recombinase-mediated cassette exchange (RMCE) technology among others have been used. The yield of BV-based system ranges from 2-20 µg/mL of p55 (Gag)[41,129,130]. Regarding stable cell lines, Sf9 has recently been reported to yield ~14 ng/mL p24 (MA)[131] and High five ~2000 ng/mL p24 (MA)[132].

3.3 Mammalian cells

The ability to produce recombinant proteins in mammalian-cell-based systems had an extensive effect in the field of research, biopharmaceutical production, the therapeutics sector and biotechnology industry in general, raising more than US$120-150 billion per year[133]. Despite of mammalian cells providing lower yields when compared to the rest of cell-based platforms, the shift to mammalian cell lines for production of biologics was driven by the demand of biosynthetic complexity in the manufactured proteins, including the intricate network of PTMs required for the treatment of human diseases. The right glycosylation pattern is essential to grant the correct pharmacodynamics, thus conferring the target protein different properties depending on variations on the glycosylation pattern. Factors like the type of cell line, the bioreactor, the culture media, or the operational mode influence the final PTMs[134].

There are many mammalian cell lines, derived from baby hamster kidney cells (BHK-21), Madin-Darby canine kidney cells (MDCK), African green monkey kidney epithelial cells (Vero), Henrietta Lacks cervical cancer cells (HeLa) or monkey kidney cells (COS), among others. However market dominance is imposed by the extensive use of Chinese hamster ovary cells (CHO), with more than 30 biopharmaceuticals approved by FDA[135]. Mostly monoclonal antibodies and hormones. However, regarding Gag VLP production, CHO cell line presents a key drawback, since the MA domain from Gag polyprotein is not correctly myristoylated, preventing monomers from assembly and oligomerization to form VLPs[136]. Different mammalian cell lines are used for Gag VLP production. One is CAP-T cell line (Cevec Pharmaceuticals), derived from human amniocytes and adapted to suspension and chemically-defined media. Titers of ~23 μg/mL have been achieved at laboratory scale using this platform[137,138]. Yet, the most widely used mammalian cell line for Gag VLP production is human embryonic kidney cells (HEK293).

3.3.1   HEK293

HEK293 cells are characterized by their ability to be easily transfected. They are typically cultured in adherence but can also be cultured in suspension, using chemically-defined and protein-free media, offering the possibility to transfer the bioprocess to a GMP-compliant manufacturer. Since they contain

adenoviral genes, like E1A, E1B or E3 they are widely used for safe adenoviral vector propagation which lack these genes[139]. The main drawback is still the low yields and productivity achieved by this platform. In this work, genetic and metabolic engineering and bioprocess intensification have been used to improve yields and take HEK293 cell line further as a Gag VLP production platform.

3.4 Production strategies for HEK293 cultures

There are two main strategies to achieve VLP production using HEK293: stable gene expression (SGE) and transient gene expression (TGE). SGE is based on the integration of the protein-coding gene in the genome, whereas TGE relies on episomal DNA expression. Whereas SGE is commonly used for well-established and characterized bioprocesses[140], TGE still prevails among previous stages of clinical trials because of the versatility of the technique, the need for performing different screening tests, and the reduced cost and time requirements of implementing TGE compared to SGE[141,142].

3.4.1   Stable gene expression (SGE)

At industrial scale, when the product has already been characterized and selected, the use of stable cell lines is widely implemented[143]. In order to generate a cell line which constitutively produces a protein of interest, the plasmid containing the gene encoding the protein and different selection markers, is delivered to cell by transfection. This construct can randomly integrate by nonsite-specific, homology-independent mechanisms in the cell genome with low probability. To increase the likelihood of this event, the constructs usually carry homology sequences to promote integration. After transfection is carried out, cells are individually sorted and screened for positive clones. As the number of copies that are integrated can vary, different rounds of screening are performed increasing selective pressure to obtain the most stable and productive clones[143]. The generation, screening, selection and evaluation of the final clones is a costly and time-consuming process that can take up to 12 months[144]. Therefore, in order to reduce time and randomness and increase efficiency, methods based on CRISPR-Cas9 and recombinase mediated cassette exchange (RMCE) have been developed to target the integration to an active and stable region

(hotspot) for the gene of interest to be actively expressed. The construct is composed of a reporter gene flanked by integration target sequences using recombinase systems (such as FRT/Flp or Cre/Loxp) so that once the construct is successfully integrated in a hotspot, the gene of interest can be easily exchanged[145]. This method reduced the required time for the generation of the stable cell line to be reduced down to 3-4 months and moderately improve yields[146].

3.4.2   Transient gene expression (TGE)

TGE offers the possibility of obtaining relatively high titers, reducing costs, time and increasing the platform versatility as the cells can be engineered and modulated for different types of products. HEK293 is easily transfected by chemical methods, becoming a promising platform for the development of TGE strategies. For the DNA to enter the cell, it needs to be combined with a transfection reagent. There are different transfection reagents, each relying on a different method to aid DNA entering the cell. Calcium phosphate, cationic lipids or cationic polymers are the most used. Nevertheless, the recent advances in mammalian cell technology are driving the field to the implementation of GMP-compliant bioprocesses with the aim of further large-scale manufacturing application. These requirements are imposing to the advantages of high transfection efficiency provided by calcium phosphate, since it requires culture media containing serum[147]. Transfection reagents based on cationic lipids, like Lipofectamine, introduce the DNA into the cell by the formation of liposomes. The negatively charged DNA is previously mixed with the cationic liposome formulation, producing positively charged vesicles containing the DNA of interest. These vesicles can easily fuse with the negatively charged cell membrane, releasing the DNA into de cytoplasm[148]. Cationic polymers, like polyethyleneimine (PEI) polymers rely on a different method. These positively charged molecules drive DNA to coalesce forming positively charged complexes that interact with the cell membrane. They enter the cell by pinocytosis and the DNA remain in an endosomal compartment within the cell[149]. Through a non-well characterized process, the DNA is released into the cytoplasm by endosomal membrane disruption. The most theorized hypothesis is the so-called *proton sponge effect,* claiming that the extraordinary buffer capacity of PEI allows the inward flow of protons and subsequent Cl⁻ ions and water molecules into the endosome, causing its collapse, releasing its content[150]. Nevertheless, in order for the
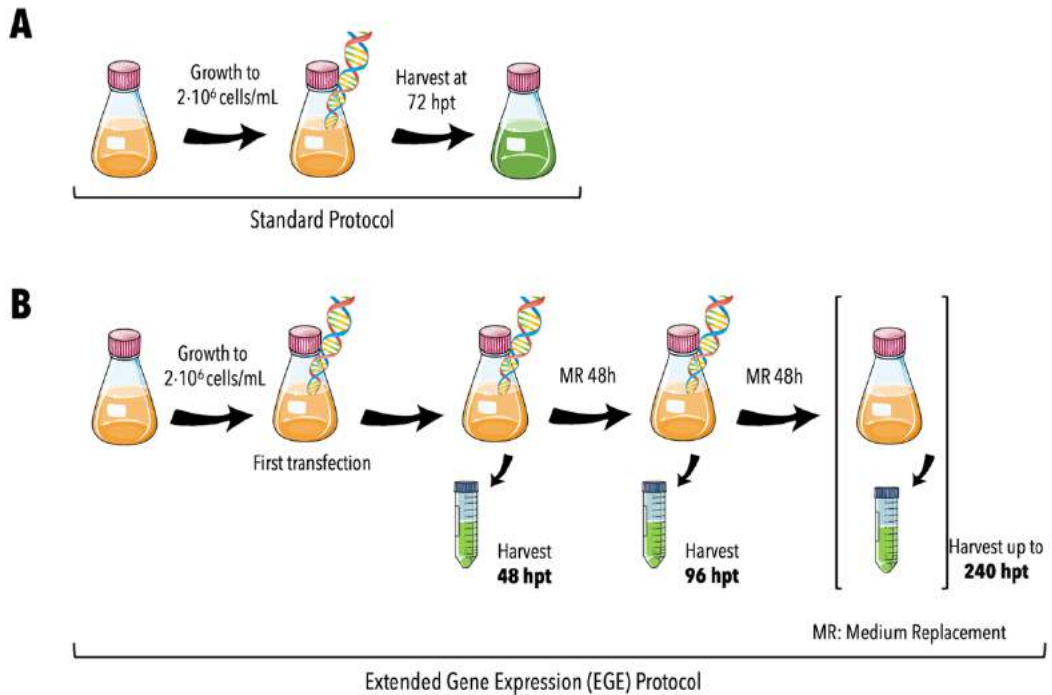
Figure 8. Transient transfection protocols. **A)** Standard transfection protocol for VLP production. Cells are seeded at $0.5 \cdot 10^6$ cells/mL and transiently transfected when reached $2 \cdot 10^6$ cells/mL. DNA is complexed with PEI. The complex mixture is added to the cell culture without media replacement. At 72 hpt VLPs are harvested from the extracellular medium. **B)** Extended gene expression (EGE) protocol for VLP production. The first transfection is performed like in the standard protocol. Then, it is followed by consecutive media replacements every 48 hours and two retransfections at 48 and 96 hpt.

gene of interest to be expressed, the coding DNA needs to reach the nucleus. Both lipofectamine and PEI-based strategies take advantage of cell replication events and the DNA is incorporated into the nucleus during G2 phase of the cell cycle as the main route[151]. DNA incorporation into the nucleus was also described through nuclear import mechanisms, involving nuclear transport proteins or the incorporation of the plasmid to the endoplasmatic reticulum (ER) system[152].

25 kDa PEI-based transfection is the most commonly used in HEK293. Nonetheless, due to its polymeric nature, PEI solutions contain molecules that can vary in size, chain lengths and monomer branching when synthesized in the lab. In order to obtain a uniform solution, designed for transfection and optimized for a most reproducible response to DNA:PEI complex formation, a commercial PEI (PEIpro) is also available overcoming the limitations found when in house preparing PEI solution.

For Gag VLP production, the plasmid encoding the gene *gag* is complexed with PEI and added to the cell culture. At 1.5 hours post-transfection (hpt) the DNA:PEI complexes have already entered the cell and at 4 hpt the complexes have reached the nucleus. Then, the episomal plasmid is transcribed and the mRNA, leaving the nucleus and synthesizing Gag polyproteins from 4 to 10 hpt[153]. These monomers then oligomerize following the VLP production pathways already described (**Figure 5**). The formation of DNA:PEI complexes depend upon many factors: DNA:PEI ratio, time of DNA and PEI interaction, the medium where complexes are formed, their concentration or even the centrifugal force applied while forming. These parameters have been previously optimized for HEK293 transient transfection of a plasmid coding for Gag polyprotein fused to an enhanced GFP protein (Gag::eGFP)[43,149]. In this work, TGE has been thoroughly explored to design an intensified bioprocess for HIV-1 Gag VLP production.

The basal standard protocol for TGE that has been used in this work (**Figure 8A**) consists in transfecting at $2 \cdot 10^6$ cells/mL with using the previously optimized DNA:PEI ratio and concentrations and harvesting the produced VLPs at 72hpt. Since the total VLP production depends on the subpopulation of transfected cells, transfecting a high cell density culture might seem a quick way to achieve this. However, the impossibility or difficulty to transfect high cell density cultures is known as the cell density effect (CDE). When the viable cell density (VCD) exceeds $4 \cdot 10^6$ cells/mL, transfection yields abruptly decrease. Analyzing the transfectable state of the cell in order to improve transfection yields and promote VLP production became crucial to exploit the TGE potential. Established strategies have already proved to improve TGE yield, like the methodology of extended gene expression (EGE).

3.4.3   Extended gene expression (EGE)

The limiting factor for production using TGE is the number of transfected cells. Therefore, in order to tackle the untransfected subpopulation of cells, a series of retransfections are carried out, increasing the probability of these untransfected cells to acquire the plasmid. If the total amount of DNA that is incorporated to the cell culture increases, so it does the total amount of PEI. This presents a major issue as the cytotoxic nature of PEI negatively affects cell viability[154]. To avoid cytotoxic accumulation, consecutive media replacements are performed every 48 hours, to avoid cytotoxic effects, remove metabolic waste and

to renew fresh culture media. The combination of retransfections and media replacements was optimized for shake flask production improving 12 fold the yields obtained in the standard protocol for transient transfection[155]. The optimized EGE protocol is shown in **Figure 8B,** which includes two retransfections, at 48 and 96 hpt and consecutive harvests up to 240hpt. However, constant media replacement led to constant cell growth, so when the EGE protocol was transferred to bioreactor scale and adapted to perfusion, the CDE appeared, as the VCD achieved in the bioreactor test was much higher than in the shake flask. Therefore, preventing the second retransfection from successfully increasing productivity[156]. In this work, in order to transfer the EGE protocol to a perfusion-based system at bioreactor scale and discard this drawback, an optimization was performed focusing on the transfectable state of the cells, tending to reach a continuous and intensified biomanufacturing process.

3.5 Continuous biomanufacturing of VLPs

To be able to achieve mass production of VLPs, production bioprocesses need to be scaled-up and transferred to bioreactor scale. However, transferring the EGE protocol to bioreactor scale required a different production mode. In order to perform constant media replacements, a continuous system needed to be implemented. Recently, the biopharmaceutical industry is moving toward the implementation of continuous manufacturing, since bioreactor size can be reduced, volume productivity increased and infrastructure cost are also reduced due to overall process intensification[157]. Thus, perfusion has gained interest for many viral particle bioprocesses that are currently adapting their production and aiming to attain high titers. This would allow a step forward in animal cell technology and vaccine development.

3.5.1 Perfusion for process intensification

Nutrients consumption presents a crucial restraint in batch and even in fed-batch processes, limiting the VCD up to $20\text{-}30\cdot10^6$ cells/mL[158–160], thus constraining the total achievable amount of viral particles produced. Perfusion is commonly used to achieve high cell densities and overcome this barrier, allowing mammalian cell cultures to reach VCD ranging from $150\text{-}200\cdot10^6$ cells/mL[161,162]. Constant addition of fresh

media and removal of metabolic waste and cytotoxic compounds has allowed mammalian cells-based bioprocesses to shift from a production of $9·10^3$ PFU/mL in shake flasks to $3.9·10^7$ PFU/mL implementing a perfusion system in a bioreactor[163]. For adherent cells, fixed bed bioreactors offers a solution where surface area is remarkably increased for a reduced size, like iCELLis® bioreactor. However, for cells growing in suspension, a cell retention device is required to operate in continuous perfusion. There are different technologies in the market to achieve this goal, mostly based on sedimentation or filtration. Operational technology from different engineering areas is being successfully applied to bioprocess development, this is the case of the inclined settler. Commonly used in water treatment, it is used to clarify drinking water from particulates. It is composed of inclined parallel plates, increasing the surface area for the particles to deposit and sediment. Maintaining a constant flow, the particles interact with the plates and sediment depending on their settling velocity. Perfusion of CHO cultures has been successfully achieved using this system[164]. Hydrocyclones are a different device that relies on the difference in centripetal force of particles in a liquid suspension. A simple device that has been adapted to animal cell culture and also been proved successful for CHO cells[165]. The first attempt to bring the EGE protocol to the bioreactor scale was performed using another sedimentation-based device, an acoustic filter[156]. The acoustic filter generates acoustic waves that once reflected within the device chamber, form nodes where cells are retained and sediment by the effect of gravity. However, the most used technology in cell-based production platforms operating in perfusion mode is filtration. Hollow fiber membranes (**Figure 9**) or hollow fiber modules (HFM) are widely used in the recombinant protein production industry. These HFM are easy to scale-up and have offered great results in the last few years. The different filtration technologies are classified according to how the liquid circulates along the membranes. Unidirectional flow along the membrane increases the probability of fouling, which occurs when the membrane pores are clogged by cellular debris, DNA, extracellular vesicles, proteins or aggregates, reducing the flux and membrane permeability, decreasing overall bioprocess performance and reducing production yields[166]. Tangential flow filtration (TFF) technology overcomes this problem by operating with a flow tangential to the membrane surface, instead of the previous system that operated orthogonally. In order to reduce membrane fouling while operating, the alternating tangential flow (ATF) technology allowed to operate in
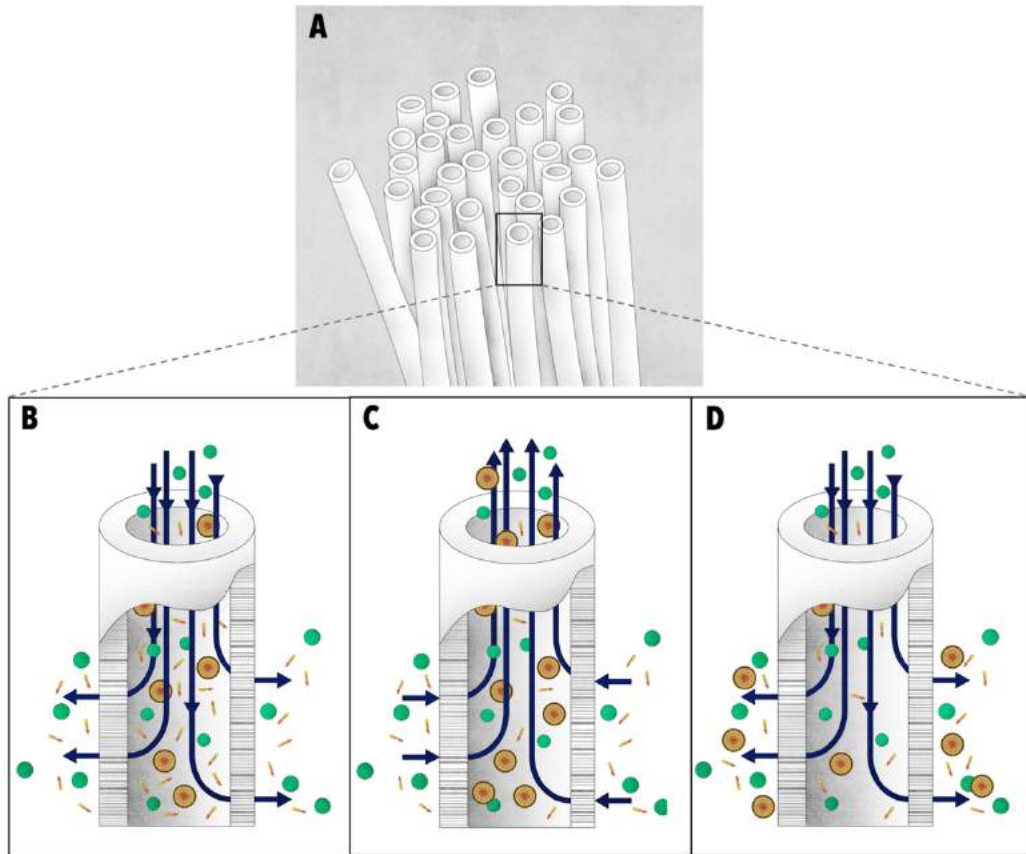
Figure 9. Hollow fiber module (HFM) representation. **A)** Generic bundle of hollow fiber membranes. **B)** Pressure phase of an alternating tangential flow (ATF) cycle. The membrane pore size threshold retains cells and different cell products. **C)** Exhaust phase of an ATF cycle. **D)** Different membranes pore sizes, materials and characteristics allow the filtration of specific products allowing separation, purification and downstream processing.

bidirectional flow (**Figure 9B-9C**). Using a diaphragm pump, the system relies on a compressor that creates a cycle of two sequential phases, pressurizing and exhausting air to drive the diaphragm, changing the direction of filtration and subsequently reducing fouling (**Figure 9B-9C**).

The ATF system was proved more suitable than TFF for animal cell cultures operating in continuous perfusion mode as its main advantage is reducing shear stress in the cells by 50%[167]. In this work, the ATF system was used to intensify the EGE protocol in continuous perfusion. However, the process needed to maintain cells in a transfectable state to avoid reaching very high cell density, thus avoiding the CDE. This resulted in the design of an alternative perfusion approach optimizing the retransfection times, the DNA

concentration and the perfusion rate. The perfusion rate is a key element for the design of a perfusion-based bioprocess.

3.5.2   Perfusion rate

The continuous inflow of fresh media into the bioreactor is measured and controlled by the perfusion rate. This parameter is closely related to cell growth as it paces the rhythm of incoming nutrients and removal of toxic by-products, but it is also key for the bioprocess design. In perfusion processes that aim to achieve the maximum possible VCD, the perfusion rate is usually restrained as media usage increases along perfusion rate, thereby raising operational costs and rendering the large-scale implementation less attractive. Bioprocesses are often optimized for the minimum perfusion rate required.

There are different strategies when implementing perfusion rate optimization. The simplest one is adjusting the number of reactor volumes (RV) of fresh media that are being incorporated per day (RV/d). This measurement is useful for practical estimates of media consumption and for comparability between processes and along scale changes[168]. However, it often overestimates the required amount of fresh media as in the initial stages of the bioprocess, low cell concentrations would require fewer media replacements than in late stages, when cell concentrations is higher. Coupling the perfusion rate to cell concentration led to the implementation of the cell specific perfusion rate (CSPR). Maintaining constant the amount of media fed per cell, the overall cellular and metabolic state can be preserved. This is particularly useful when bringing the EGE protocol to a perfusion-based bioprocess since the transfectable metabolic state of the cells needs to be maintained and prolonged in time. This optimization approach is also advantageous since the use of media is adjusted to cell requirements, reducing potential media waste. For a more stable and tight control, a loop with an on-line measurement of VCD is usually implemented[169], refining the CSPR at each time point. Similarly to cell concentration, perfusion rates can be coupled to a limiting substrate or target metabolite that is required to remain stable along the process[170].

As for the operational details, usually the total bioreactor volume is maintained using a scale measuring the bioreactor weight in a control loop with the feeding pump. The pump adjusting the outflow rate is fixed

and manually modified or controlled by VCD probes or specific metabolite probes[171,172]. For conventional perfusion bioprocesses, once the desired cell concentration is achieved, steady state is maintained controlling the perfusion rate and the bleed flowrate. Bleeding of cells takes place once the cell density has surpassed the desired value and part of the cells are removed from the culture to re-establish the set point cell concentration. This operation can also be controlled by a cell density probe placed in the bioreactor and directly coupled to VCD measurements[171].

3.5.3   Hollow fiber membranes

Hollow fiber membrane properties like type of material, pore size, or functionalization depend on the requirements of the basic operation for which it is implemented. As a cell retention device in perfusion-based bioprocesses, HFM are often autoclavable and made of polymers like polyethylene (PE), polysulfone (PS) or polyethersulfone (PES)[173]. These polymers are relatively inert to protein interaction, avoiding adsorption and reducing fouling. The pore size also varies depending on the membrane functionality. For perfusion, commercial available membranes present pore sizes with a cut-off ranging from 0.6-0.2 μm, retaining cells and allowing most of recombinant proteins and products to be filtered and harvested in the permeate fraction[157]. To successfully adapt a bioprocess to perfusion mode using HFM, choosing the right material and cut-off pore size according to the product is essential[174]. HFM are used not only for perfusion, but also for the subsequent purification steps since the type, pore size and characteristics of the fiber can be used for VLP filtration and separation (**Figure 9D**). For specific products, the PES membrane can be further modified and functionalized with different chemical groups, known as modified polyethersulfone (mPES), to create interactions with the retentate and improve separation and purification of the product[175].

3.6 Downstream processing of VLPs

For VLP purification, the downstream process is similar to other enveloped viral particles. Since particles are released to the extracellular medium, the first step is clarification from cells, cellular debris and large aggregates, which is usually achieved by low-speed centrifugation[176]. For further concentration, ultracentrifugation using sucrose cushions has been reported to successfully concentrate VLPs[177,178]. However, aiming to design a scalable, integrated continuous bioprocess, different approaches based on filtration or capture chromatography have been implemented. Depth filtration is widely used for the first step of clarification since this type of filters are able to retain large elements like cells, debris or big particulates being less prone to clogging[179]. TFF technology can also be used for clarification and concentration as the versatile and inert HFM can be selected to filter the desired product choosing the correct pore size. TFF technology is easily scaled-up and has been used in industrial manufacturing of biologics produced in mammalian cell cultures for decades[180]. Recently, scalable methods based on capture chromatography have been explored for the concentration step. Anion exchange monolith columns or Mustang® Q have been reported to successfully concentrate viral vectors up to 400 fold[181,182]. Currently, efforts are being made towards the achievement of an efficient separation of VLPs and EVs using chromatographic methods. For the final step of diafiltration, polishing and sterile filtration and TFF technology can also be used.

4. Bioprocess engineering

When optimizing mammalian cell-based production bioprocesses, there are two main aspects that need to be addressed. The optimization at operational level, regarding culture media, additives, bioreactor conditions and culture mode and the optimization at molecular and cellular level, regarding cellular homeostasis and metabolic pathways controlling the biological synthesis of the desired product. To tackle the first challenge, many variables should be considered (e.g. the effect of different additives enhancing production, different perfusion rates…etc). Therefore, testing all the possible variables and their interactions, may present a long, poorly cost-effective, time consuming and sometimes virtually impossible process, as it would require vast amounts of scarce reagents. A useful methodology to test these operational variables and to optimize their performance is design of experiments (DoE) and response surface methodology (RSM)[183]. RSM analyzes the relationship between the different independent tested variables and the response arisen when these are combined. In order to construct a representative model based on these estimations, the tested variables need to be carefully chosen. For this choice, different DoE methodologies offer different combinations of experiments to perform so the experimental conditions are reduced to the minimum necessary tests to create a representative model. Box-Behnken design establishes 15 experimental conditions, each of them analyzing three independent variables. Each combination explores three different levels of the independent variables. Therefore, placed in a way that they define an experimental space where the model is rendered (**Figure 10A**). The response from each of the 15 combinations is statistically analyzed by RSM and the response model is generated (**Figure 10A**). There are different DoE approaches depending on the number of independent variables to be analyzed and the levels at which the variables need to be tested. This strategy is a powerful tool for bioprocess optimization at operational level, being applicable to bioreactor design variables, culture media additives and other relevant parameters. Additionally to bioprocess performance characterization and optimization, the understanding of cell metabolism is the second block to consider in process optimization. Over the past few years, the importance of a holistic understanding of cell metabolism and molecular pathways governing the cell homeostasis and ultimately the synthesis of the desired product, have changed the
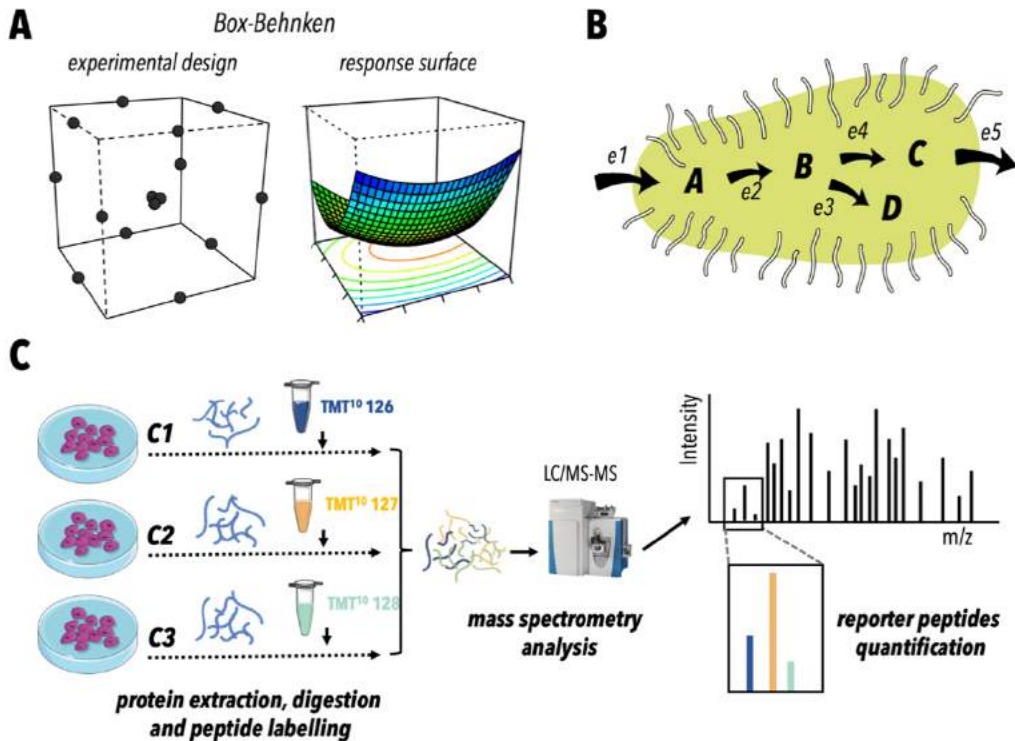
Figure 10. Bioprocess optimization tools. **A)** Representation of design of experiments (DoE) and response surface methodology (RSM). The 15 experimental conditions designed by the Box-Behnken approach and their position in the experimental space (left) and the model rendered by the response obtained (right) . **B)** Representation of the possible manipulations in a metabolic pathway to enhance the production of a desired metabolite. A,B,C and D represent the different intermediates or metabolites in a metabolic pathway. e1,e2,e3,e4 and e5 represent the different enzymes that participate in the metabolic pathway and can be modulated. The first is modulating the uptake of nutrients or essential compounds or for the synthesis of the envisaged product (A). This can be achieved by overexpressing transport enzymes (e1). The second strategy is overexpressing the enzyme (e2) that synthesizes the precursor of our product (B). The third strategy is to inhibit enzymes (e3) that diverge the pathway towards by-products (D). Therefore, maximizing the use of resources and targeting them towards the synthesis of the desired product (C). A forth strategy is to heterologously express new enzymes (e4) from a different organism that could transform potential precursors and produce our metabolite (C). The last and fifth strategy is to enhance the product secretion, by overexpression of the direct export proteins or the enzymes involved in the outward mechanism transport (e5). **C)** Representation of a generic workflow for isobaric labeling using tandem mass tags (TMT). Proteins from different conditions (C1, C2 and C3) are digested and peptides are labeled using different molecules whose mass is identical but their composition, in terms of heavy isotopes distribution, varies between studied condition. Each tag is composed of a reporter mass (126,127,128) and a balance mass linked by a specific cleavage site that is fragmented under higher-energy collisional dissociation (HCD) conditions in a mass spectrometer. Therefore, when the peptide is identified and sequenced by tandem mass spectrometry (MS/MS), the reporter tags can be relatively quantified in the fragmentation spectrum. The ratios between the reporter tags reflect the changes between the different studied conditions or proteomes (C1, C2 and C3).

focus from single reactions to the modulation of the network as the unit of action. The realization that simple changes in cell homeostasis could ripple throughout the entire network has shifted the efforts to

the identification of the nodes and links forming the relevant metabolic cluster for our bioprocess. The field of metabolic engineering provided tools for overexpression, downregulation or modulation of genes and enzymes to promote the desired output and the synthesis of the desired metabolite. There are five main strategies when optimizing a metabolic route, as illustrated in **Figure 10B**. The first is modulating the uptake of nutrients or essential compounds for the synthesis of the envisaged product. This can be achieved by overexpressing transport enzymes or permeabilizing the cell membrane with additives. The second strategy is overexpressing the enzymes responsible for the synthesis of the precursor of the product, increasing concentration and shifting the equilibrium towards the synthesis of the desired product, increasing the production rate. The third strategy is to inhibit enzymes that diverge the pathway towards by-products. Therefore, maximizing the use of resources and targeting them towards the synthesis of the desired product. This can be achieved using RNA interference (RNAi), directly modulating the enzyme by protein inhibitors or via regulating their parallel cofactor recycling pathways. A fourth strategy is to heterologously express new enzymes from a different organism that could transform potential precursors and produce our metabolite. The last and fifth strategy is to enhance the product secretion, by overexpression of the direct export proteins or the enzymes involved in the outward mechanism transport (like the machinery responsible of cell budding for the production of VLPs). Usually, the final engineering endeavor is a combination of some of the described strategies[184]. In the recent years, more holistic approaches are being undertaken, engineering transcription factors to modulate the transcription of several enzymes at the same time or designing from scratch the whole pathway, including genes, promoters and terminators. Complex synthetic biology can be applied for the task of metabolic engineering[185]. In order to select the optimal cellular manipulation, understanding and analyzing the metabolic network involved is essential to be able to predict potential results. The rise of the "-omics" techniques enabled fast characterization of several different pathways at the same time on the basis of powerful tools like multiplexed quantitative proteomics[186]. The proteome is the entire set of proteins found in an organism in a given moment or under a given condition. Depending on the developmental stage or manipulation applied to an organism, its proteome changes. The identification and quantification of these variations allows to define functional changes, since almost every single perturbation in the cell is reflected

as a change in the proteome. The reconstruction of all these changes in the cellular metabolic network offers a complete understanding of the consequences of that perturbation in the cell as a whole. The required technology to study these phenomena is mass spectrometry and isobaric labeling. In isobaric labeling, proteins from different conditions are digested and peptides are labeled using different molecules whose mass, or molecular weight, is identical but their composition, in terms of heavy isotopes distribution, varies between the studied conditions (**Figure 10C**). Each tag is composed of a reporter mass and a balance mass linked by a specific cleavage site that is fragmented under higher-energy collisional dissociation (HCD) conditions in a mass spectrometer. Therefore, when the peptide is identified and sequenced by tandem mass spectrometry (MS/MS), the reporter tags can be relatively quantified in the fragmentation spectrum (**Figure 10C**). The ratios between the reporter tags reflect the changes between the different studied conditions or proteomes. The most common used approaches for relative quantification using isobaric labeling are isobaric tags for absolute and relative quantification (iTRAQ) which can compare up to 8 different conditions and tandem mass tags (TMT) which can compare up to 10 different conditions. This is achieved by changing the distribution of the number of atoms of ^{13}C and ^{14}N in the same molecule, obtaining 10 different combinations (126, 127N, 127C, 128N, 128C, 129N, 129C, 130N, 130C and 131). Likewise, unlabeled peptides can be monitored using MS. Under this approach, the parallel reaction monitoring (PRM), synthesized peptides can be used to construct a standard calibration curve using MS identification to absolutely quantify the endogenous peptide in different conditions.

The use of proteomics tools started to be applied in recent years [187]. In this work, multiplexed quantitative proteomics technology is applied as one of the tools for bioprocess optimization for the production of Gag VLPs in HEK293 cultures. First, to characterize the implications of TGE in the cell proteome and second, to characterize the coproduced EVs found in the VLP secretome. The results obtained from these metabolic network analyses have been used to design strategies for improvement of VLP budding and transfection efficiency. Furthermore, tools like DoE have been used to optimize the bioprocess at operational level, proposing a perfusion-based culture mode to intensify VLP production. Finally, the structure of VLPs have been characterized by PRM, electron microscopy techniques and glycomic techniques.

REFERENCES

1. Jenner, E. *An inquiry into the causes and effects of the variolae vaccinae: a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox. Springfield [Mass.] : Re-printed for Dr. Samuel Cooley, by Ashley & Brewer, 1802* (1802). doi:http://resource.nlm.nih.gov/2559001R

2. Riedel, S. Edward Jenner and the History of Smallpox and Vaccination. *Baylor Univ. Med. Cent. Proc.* **18**, 21–25 (2005).

3. Leung, A. K. C. "Variolation" and Vaccination in Late Imperial China, Ca 1570–1911. in *History of Vaccine Development* 5–12 (Springer New York, 2011). doi:10.1007/978-1-4419-1339-5_2

4. Fenner, F. *et al.* Smallpox and its eradication / F. Fenner ... [et al.]. (1988).

5. Benecke, O. & DeYoung, S. E. Anti-Vaccine Decision-Making and Measles Resurgence in the United States. *Glob. Pediatr. Heal.* **6**, 2333794X1986294 (2019).

6. Chakraborty, I. & Maity, P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Sci. Total Environ.* **728**, 138882 (2020).

7. Pyle, K. D. P. G. F. The Geography and Mortality of the 1918 Influenza Pandemic - PubMed. **65**, 4–21 (1991).

8. Number of people living with HIV, 2017. Available at: https://ourworldindata.org/grapher/number-of-people-living-with-hiv?year=latest. (Accessed: 1st May 2020)

9. Geary, C. W. *et al.* Personal involvement of young people in HIV prevention campaign messages: the role of message format, culture, and gender. *Health Educ. Behav.* **35**, 190–206 (2008).

10. Ochmann, S. & Roser, M. Smallpox. *Our World Data* (2018).

11. Coronavirus Pandemic (COVID-19) – the data - Our World in Data. Available at: https://ourworldindata.org/coronavirus-data. (Accessed: 1st May 2020)

12. Bloom, D. E., Canning, D. & Weston, M. *The Value of Vaccination*.

13. Noazin, S. *et al.* First generation leishmaniasis vaccines: A review of field efficacy trials. *Vaccine* **26**, 6759–6767 (2008).

14. Huang, D. B., Wu, J. J. & Tyring, S. K. A review of licensed viral vaccines, some of their safety concerns, and the advances in the development of investigational viral vaccines. *Journal of Infection* **49**, 179–209 (2004).

15. Cox, R. J., Brokstad, K. A. & Ogra, P. Influenza Virus: Immunity and Vaccination Strategies. Comparison of the Immune Response to Inactivated and Live, Attenuated Influenza Vaccines. *Scand. J. Immunol.* **59**, 1–15 (2004).

16. Vignuzzi, M., Wendt, E. & Andino, R. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.* **14**, 154–161 (2008).

17. Salk, J. Polio vaccines and polioviruses. *British Medical Journal* **2**, 765 (1977).

18. World Health Organization. Global Wild Poliovirus 2015 - 2020. Available at: http://polioeradication.org/wp-content/uploads/2020/02/Weekly-GPEI-Polio-Analyses-WPV-20200220.pdf. (Accessed: 8th May 2020)

19. Dubé, E. *et al.* Vaccine hesitancy: An overview. *Human Vaccines and Immunotherapeutics* **9**, 1763–1773 (2013).

20. Tahamtan, A., Charostad, J., Hoseini Shokouh, S. J. & Barati, M. An Overview of History, Evolution, and Manufacturing of Various Generations of Vaccines. *J. Arch. Mil. Med.* **In Press**, (2017).

21. Stauffer, F., El-Bacha, T. & Da Poian, A. Advances in the Development of Inactivated Virus Vaccines. *Recent Pat. Antiinfect. Drug Discov.* **1**, 291–296 (2008).

22. Robinson, H. L & Amara, R. R. T cell vaccines for microbial infections. *Nat. Med.* **11**, S25 (2005).

23. Plotkin, S. A. Vaccines: The fourth century. *Clinical and Vaccine Immunology* **16**, 1709–1719 (2009).

24. Nguyen, A. & Schwalbe, N. Apples and oranges? Can second generation vaccines become as low cost as generic medicines? *Vaccine* **37**, 2910–2914 (2019).

25. Moloney, P. J. THE PREPARATION AND TESTING OF DIPHTHERIA TOXOID (ANATOXINE-RAMON). *Am. J. Public Health* **16**, 1208–1210 (1926).

26. Amanna, I. J., Carlson, N. E. & Slifka, M. K. Duration of humoral immunity to common viral and

vaccine antigens. *N. Engl. J. Med.* **357**, 1903–1915 (2007).

27.   Goldblatt, D. Conjugate vaccines. *Clinical and Experimental Immunology* **119**, 1–3 (2000).

28.   Tan, M., Zhong, W., Song, D., Thornton, S. & Jiang, X. E. coli-expressed recombinant norovirus capsid proteins maintain authentic antigenicity and receptor binding capability. *J. Med. Virol.* **74**, 641–649 (2004).

29.   Saraswat, S. *et al.* Expression and Characterization of Yeast Derived Chikungunya Virus Like Particles (CHIK-VLPs) and Its Evaluation as a Potential Vaccine Candidate. *PLoS Negl. Trop. Dis.* **10**, (2016).

30.   Roldão, A., Mellado, M. C., Castilho, L. R., Carrondo, M. J. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev Vaccines* **9**, 1149–1176 (2010).

31.   A, F., AP, L., S, B. & RJ, L. Second-generation Prophylactic HPV Vaccines: Current Options and Future Strategies for Vaccines Development. *Minerva Med.* **107**, 26–38 (2016).

32.   Sasnauskas, K. *et al.* Generation of recombinant virus-like particles of human and non-human polyomaviruses in yeast Saccharomyces cerevisiae. in *Intervirology* **45**, 308–317 (Intervirology, 2002).

33.   Shelly D, C. V. Parvovirus B19 VLP Vaccine Manufacturing. *Genet Eng Biotechnol News* **29**, 1–4 (2009).

34.   Niikura, M. *et al.* Chimeric recombinant hepatitis E virus-like particles as an oral vaccine vehicle presenting foreign epitopes. *Virology* **293**, 273–280 (2002).

35.   Crevar, C. J. & Ross, T. M. Elicitation of protective immune responses using a bivalent H5N1 VLP vaccine. *Virol. J.* **5**, (2008).

36.   Kang, S. M. *et al.* Induction of long-term protective immune responses by influenza H5N1 virus-like particles. *PLoS One* **4**, (2009).

37.   Matassov, D., Cupo, A. & Galarza, J. M. A novel intranasal virus-like particle (VLP) vaccine designed to protect against the pandemic 1918 influenza A virus (H1N1). *Viral Immunol.* **20**, 441–452 (2007).

38.   Prel, A., Le Gall-Reculé, G. & Jestin, V. Achievement of avian influenza virus-like particles that

could be used as a subunit vaccine against low-pathogenic avian influenza strains in ducks. *Avian Pathol.* **37**, 513–520 (2008).

39. Ho, Y., Lin, P. H., Liu, C. Y. Y., Lee, S. P. & Chao, Y. C. Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochem. Biophys. Res. Commun.* **318**, 833–838 (2004).

40. Kessans, S. A., Linhart, M. D., Matoba, N. & Mor, T. Biological and biochemical characterization of HIV-1 Gag/dgp41 virus-like particles expressed in Nicotiana benthamiana. *Plant Biotechnol. J.* **11**, 681–690 (2013).

41. Puente-Massaguer, E., Lecina, M. & Gòdia, F. Application of advanced quantification techniques in nanoparticle-based vaccine development with the Sf9 cell baculovirus expression system. *Vaccine* **38**, 1849–1859 (2020).

42. Puente-Massaguer, E., Lecina, M. & Gòdia, F. Nanoscale characterization coupled to multi-parametric optimization of Hi5 cell transient gene expression. *Appl. Microbiol. Biotechnol.* **102**, 10495–10510 (2018).

43. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

44. Cervera, L. *et al.* Selection and optimization of transfection enhancer additives for increased virus-like particle production in HEK293 suspension cell cultures. *Appl Microbiol Biotechnol* **99**, 9935–9949 (2015).

45. Hobernik, D. & Bros, M. DNA vaccines–How far from clinical use? *International Journal of Molecular Sciences* **19**, (2018).

46. Ghaffarifar, F. Plasmid DNA vaccines: Where are we now? *Drugs of Today* **54**, 315–333 (2018).

47. Baden, L. R. *et al. Safety and immunogenicity of two heterologous HIV vaccine regimens in healthy, HIV-uninfected adults (TRAVERSE): a randomised, parallel-group, placebo-controlled, double-blind, phase 1/2a study.* **7**, (2020).

48. Johnson & Johnson Announces New Public-Private Partnership to Support First Phase 3 Efficacy Study of Janssen's Investigational Prophylactic HIV Vaccine | Johnson & Johnson. Available at:

https://www.jnj.com/johnson-johnson-announces-new-public-private-partnership-to-support-first-phase-3-efficacy-study-of-janssens-investigational-prophylactic-hiv-vaccine. (Accessed: 9th January 2021)

49. Zhang, C., Maruggi, G., Shan, H. & Li, J. Advances in mRNA vaccines for infectious diseases. *Frontiers in Immunology* **10**, 594 (2019).

50. Jackson, N. A. C., Kester, K. E., Casimiro, D., Gurunathan, S. & DeRosa, F. The promise of mRNA vaccines: a biotech and industrial perspective. *npj Vaccines* **5**, 1–6 (2020).

51. Weissman, D., Pardi, N., Muramatsu, H. & Karikó, K. HPLC purification of in vitro transcribed long RNA. *Methods Mol. Biol.* **969**, 43–54 (2013).

52. Lamichhane, T. N. & Jay, S. M. Production of extracellular vesicles loaded with therapeutic cargo. in *Methods in Molecular Biology* **1831**, 37–47 (Humana Press Inc., 2018).

53. Fitzpatrick, Z., György, B., Skog, J. & Maguire, C. A. Extracellular vesicles as enhancers of virus vector-mediated gene delivery. *Hum. Gene Ther.* **25**, 785–786 (2014).

54. György, B. & Maguire, C. A. Extracellular vesicles: nature's nanoparticles for improving gene transfer with adeno-associated virus vectors. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **10**, e1488 (2018).

55. Doyle, L. & Wang, M. Overview of Extracellular Vesicles, Their Origin, Composition, Purpose, and Methods for Exosome Isolation and Analysis. *Cells* **8**, 727 (2019).

56. Pardi, N., Hogan, M. J., Porter, F. W. & Weissman, D. mRNA vaccines-a new era in vaccinology. *Nature Reviews Drug Discovery* **17**, 261–279 (2018).

57. Zhu, T. *et al.* An African HIV-1 sequence from 1959 and implications for the of the epidemic. *Nature* **391**, 594–597 (1998).

58. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS pandemic. *Cold Spring Harb. Perspect. Med.* **1**, (2011).

59. Whiteside, A. & Wilson, D. Health and AIDS in 2019 and beyond. *African J. AIDS Res.* **17**, iii–v (2018).

60. Moir, S., Chun, T.-W. & Fauci, A. S. Pathogenic Mechanisms of HIV Disease. *Annu. Rev. Pathol.*

Mech. Dis. **6**, 223–248 (2011).

61.   Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136–144 (2013).

62.   Vasudevan, A. A. J. *et al.* Structural features of antiviral DNA cytidine deaminases. *Biological Chemistry* **394**, 1357–1370 (2013).

63.   Kim, K. *et al.* Vpr-Binding Protein Antagonizes p53-Mediated Transcription via Direct Interaction with H3 Tail. *Mol. Cell. Biol.* **32**, 783–796 (2012).

64.   Garcia, J. V. & Miller, A. D. Serine phosphorylation-independent downregulation of cell-surface CD4 by nef. *Nature* **350**, 508–511 (1991).

65.   Lenassi, M. *et al.* HIV Nef is secreted in exosomes and triggers apoptosis in bystander CD4+ T cells. *Traffic* **11**, 110–22 (2010).

66.   Raymond, A. D. *et al.* HIV Type 1 Nef Is Released from Infected Cells in CD45 [+] Microvesicles and Is Present in the Plasma of HIV-Infected Individuals. *AIDS Res. Hum. Retroviruses* **27**, 167–178 (2011).

67.   McDonald, D. *et al.* Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* **159**, 441–452 (2002).

68.   Lee, K. E. *et al.* Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).

69.   Saad, J. S. *et al.* Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl. Acad. Sci.* **103**, 11364–11369 (2006).

70.   Hogue, I. B., Grover, J. R., Soheilian, F., Nagashima, K. & Ono, A. Gag Induces the Coalescence of Clustered Lipid Rafts and Tetraspanin-Enriched Microdomains at HIV-1 Assembly Sites on the Plasma Membrane. *J. Virol.* **85**, 9749–9766 (2011).

71.   Florin, L. & Lang, T. Tetraspanin Assemblies in Virus Infection. *Front. Immunol.* **9**, 1140 (2018).

72.   Grigorov, B. *et al.* A role for CD81 on the late steps of HIV-1 replication in a chronically infected T cell line. *Retrovirology* **6**, 28 (2009).

73.   Patters, B. J. & Kumar, S. The role of exosomal transport of viral agents in persistent HIV

pathogenesis. *Retrovirology* **15**, 79 (2018).

74. Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

75. von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

76. Henne, W. M., Buchkovich, N. J. & Emr, S. D. The ESCRT pathway. *Dev Cell* **21**, 77–91 (2011).

77. Wollert, T. *et al.* The ESCRT machinery at a glance. in *J Cell Sci* **122**, 2163–2166 (2009).

78. Fujii, K., Hurley, J. H. & Freed, E. O. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nat Rev Microbiol* **5**, 912–916 (2007).

79. McCullough, J., Fisher, R. D., Whitby, F. G., Sundquist, W. I. & Hill, C. P. ALIX-CHMP4 interactions in the human ESCRT pathway. *Proc Natl Acad Sci U S A* **105**, 7687–7691 (2008).

80. Alonso, Y. A. M., Migliano, S. M. & Teis, D. ESCRT-III and Vps4: a dynamic multipurpose tool for membrane budding and scission. *Febs j* **283**, 3288–3302 (2016).

81. Langelier, C. *et al.* Human ESCRT-II complex and its role in human immunodeficiency virus type 1 release. *J Virol* **80**, 9465–9480 (2006).

82. Meng, B., Ip, N. C., Prestwood, L. J., Abbink, T. E. & Lever, A. M. Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. *Retrovirology* **12**, 72 (2015).

83. Chung, H. Y. *et al.* NEDD4L overexpression rescues the release and infectivity of human immunodeficiency virus type 1 constructs lacking PTAP and YPXL late domains. *J Virol* **82**, 4884–4897 (2008).

84. Usami, Y., Popov, S., Popova, E. & Gottlinger, H. G. Efficient and specific rescue of human immunodeficiency virus type 1 budding defects by a Nedd4-like ubiquitin ligase. *J Virol* **82**, 4898–4907 (2008).

85. Hurley, J. H. ESCRT complexes and the biogenesis of multivesicular bodies. *Curr Opin Cell Biol* **20**, 4–11 (2008).

86. Khorchid, A., Halwani, R., Wainberg, M. A. & Kleiman, L. Role of RNA in Facilitating Gag/Gag-Pol Interaction. *J. Virol.* **76**, 4131–4137 (2002).

87. Muriaux, D., Mirro, J., Harvin, D. & Rein, A. RNA is a structural element in retrovirus particles. *Proc. Natl. Acad. Sci.* **98**, 5246–5251 (2001).

88. Könnyű, B. *et al.* Gag-Pol Processing during HIV-1 Virion Maturation: A Systems Biology Approach. *PLoS Comput. Biol.* **9**, e1003103 (2013).

89. Briggs, J. A. G. *et al.* The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–675 (2004).

90. Hessvik, N. P. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci.* **75**, 193–208 (2018).

91. Brewer, P. D., Habtemichael, E. N., Romenskaia, I., Coster, A. C. F. & Mastick, C. C. Rab14 limits the sorting of Glut4 from endosomes into insulin-sensitive regulated secretory compartments in adipocytes. *Biochem. J.* **473**, 1315–1327 (2016).

92. Gruenberg, J. & Stenmark, H. The biogenesis of multivesicular endosomes. *Nat. Rev. Mol. Cell Biol.* **5**, 317–323 (2004).

93. Pant, S., Hilton, H. & Burczynski, M. E. The multifaceted exosome: Biogenesis, role in normal and aberrant cellular function, and frontiers for pharmacological and biomarker opportunities. *Biochem. Pharmacol.* **83**, 1484–1494 (2012).

94. Vietri, M., Radulovic, M. & Stenmark, H. The many functions of ESCRTs. *Nat. Rev. Mol. Cell Biol.* **21**, 25–42 (2020).

95. van Niel, G. *et al.* The Tetraspanin CD63 Regulates ESCRT-Independent and -Dependent Endosomal Sorting during Melanogenesis. *Dev. Cell* **21**, 708–721 (2011).

96. Trajkovic, K. *et al.* Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes. *Science (80-.).* **319**, 1244–1247 (2008).

97. Venereo-Sánchez, A. *et al.* Characterization of influenza H1N1 Gag virus-like particles and extracellular vesicles co-produced in HEK-293SF. *Vaccine* **37**, 7100–7107 (2019).

98. Giles, B. M. & Ross, T. M. A computationally optimized broadly reactive antigen (COBRA) based H5N1 VLP vaccine elicits broadly reactive antibodies in mice and ferrets. *Vaccine* **29**, 3043–3054 (2011).

99. Chua, A. J. *et al.* A novel platform for virus-like particle-display of flaviviral envelope domain III: Induction of Dengue and West Nile virus neutralizing antibodies. *Virol. J.* **10**, 129 (2013).

100. Di Bonito, P. *et al.* Anti-tumor CD8+ T cell immunity elicited by HIV-1-based virus-like particles incorporating HPV-16 E7 protein. *Virology* **395**, 45–55 (2009).

101. Osterrieder, N., Wagner, R., Brandmüller, C., Schmidt, P. & Wolf, H. Protection against EHV-1 challenge infection in the murine model after vaccination with various formulations of recombinant glycoprotein gp14 (gB). *Virology* **208**, 500–510 (1995).

102. Garnier, L. *et al.* Incorporation of pseudorabies virus gD into human immunodeficiency virus type 1 Gag particles produced in baculovirus-infected cells. *J. Virol.* **69**, 4060–4068 (1995).

103. Fontana, D. *et al.* Chimeric vlps based on hiv-1 gag and a fusion rabies glycoprotein induce specific antibodies against rabies and foot-and-mouth disease virus. *Vaccines* **9**, (2021).

104. Hein, C. D., Liu, X. M. & Wang, D. Click chemistry, a powerful tool for pharmaceutical sciences. *Pharmaceutical Research* **25**, 2216–2230 (2008).

105. Rulli, S. J. *et al.* Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. *J. Virol.* **81**, 6623–6631 (2007).

106. Lua, L. H. *et al.* Bioengineering virus-like particles as vaccines. *Biotechnol Bioeng* **111**, 425–440 (2014).

107. Kaczmarczyk, S. J., Sitaraman, K., Young, H. A., Hughes, S. H. & Chatterjee, D. K. Protein delivery using engineered virus-like particles. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16998–17003 (2011).

108. Deml, L., Schirmbeck, R., Reimann, J., Wolf, H. & Wagner, R. Recombinant human immunodeficiency Pr55(gag) virus-like particles presenting chimeric envelope glycoproteins induce cytotoxic T-cells and neutralizing antibodies. *Virology* **235**, 26–39 (1997).

109. Mohsen, M. O., Speiser, D. E., Knuth, A. & Bachmann, M. F. Virus-like particles for vaccination against cancer. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **12**, (2020).

110. Vigerust, D. J. & Shepherd, V. L. Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol.* **15**, 211–218 (2007).

111. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).

112. Williams, C. *et al.* Glycosylation of extracellular vesicles: current knowledge, tools and clinical perspectives. *J. Extracell. vesicles* **7**, 1442985 (2018).

113. Tak W. Mak, M. E. S. *The Immune Response: Basic and Clinical Principles*. (2006).

114. Brauer, R. & Chen, P. Influenza virus propagation in embryonated chicken eggs. *J. Vis. Exp.* **2015**, (2015).

115. Frierson, J. G. The yellow fever vaccine: A history. *Yale J. Biol. Med.* **83**, 77–85 (2010).

116. Huang, S. S. H. *et al.* Immunity toward H1N1 influenza hemagglutinin of historical and contemporary strains suggests protection and vaccine failure. *Sci. Rep.* **3**, 1–9 (2013).

117. Freire, M. S. *et al.* Production of yellow fever 17DD vaccine virus in primary culture of chicken embryo fibroblasts: Yields, thermo and genetic stability, attenuation and immunogenicity. *Vaccine* **23**, 2501–2512 (2005).

118. Macek, B. *et al.* Protein post-translational modifications in bacteria. *Nature Reviews Microbiology* **17**, 651–664 (2019).

119. Eichler, J. & Koomey, M. Sweet New Roles for Protein Glycosylation in Prokaryotes. *Trends in Microbiology* **25**, 662–672 (2017).

120. Sakuragi, S., Goto, T., Sano, K. & Morikawa, Y. HIV type 1 Gag virus-like particle budding from spheroplasts of saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 7956–7961 (2002).

121. Strasser, R., Altmann, F., Mach, L., Glössl, J. & Steinkellner, H. Generation of Arabidopsis thaliana plants with complex N-glycans lacking **β**1,2-linked xylose and core **α**1,3-linked fucose. *FEBS Lett.* **561**, 132–136 (2004).

122. Meyers, A. *et al.* Expression of HIV-1 antigens in plants as potential subunit vaccines. *BMC Biotechnol.* **8**, 53 (2008).

123. Concha, C. *et al.* Disease prevention: An opportunity to expand edible plant-based vaccines? *Vaccines* **5**, (2017).

124. Mishra, N., Gupta, N., Khatri, K., Goyal, A. K. & Vyas, S. P. *Edible vaccines: A new approach to oral immunization. Indian Journal of Biotechnology* **7**, (2008).

125. Kurup, V. M. & Thomas, J. Edible Vaccines: Promises and Challenges. *Molecular Biotechnology* **62**, 79–90 (2020).

126. Shi, X. & Jarvis, D. Protein N-Glycosylation in the Baculovirus-Insect Cell System. *Curr. Drug Targets* **8**, 1116–1125 (2007).

127. Li, J., Hsu, H. C., Mountz, J. D. & Allen, J. G. Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. *Cell Chemical Biology* **25**, 499–512 (2018).

128. Kovaleva, E. S., O'Connell, K. P., Buckley, P., Liu, Z. & Davis, D. C. Recombinant protein production in insect larvae: Host choice, tissue distribution, and heterologous gene instability. *Biotechnol. Lett.* **31**, 381–386 (2009).

129. Pillay, S., Meyers, A., Williamson, A. L. & Rybicki, E. P. Optimization of chimeric HIV-1 virus-like particle production in a baculovirus-insect cell expression system. in *Biotechnology Progress* **25**, 1153–1160 (Biotechnol Prog, 2009).

130. Puente-Massaguer, E., Lecina, M. & Gòdia, F. Integrating nanoparticle quantification and statistical design of experiments for efficient HIV-1 virus-like particle production in High Five cells. *Appl. Microbiol. Biotechnol.* **104**, 1569–1582 (2020).

131. Vidigal, J. *et al.* RMCE-based insect cell platform to produce membrane proteins captured on HIV-1 Gag virus-like particles. *Appl. Microbiol. Biotechnol.* **102**, 655–666 (2018).

132. Tagliamonte, M. *et al.* Constitutive expression of HIV-VLPs in stably transfected insect cell line for efficient delivery system. *Vaccine* **28**, 6417–6424 (2010).

133. Bandaranayake, A. D. & Almo, S. C. Recent advances in mammalian protein production. *FEBS Lett* **588**, 253–260 (2014).

134. Marth, J. D. & Grewal, P. K. Mammalian glycosylation in immunity. *Nature Reviews Immunology* **8**, 874–887 (2008).

135. Zhu, J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.* **30**, 1158–1170 (2012).

136. Reed, M. *et al.* Chimeric Human Immunodeficiency Virus Type 1 Containing Murine Leukemia Virus Matrix Assembles in Murine Cells. *J. Virol.* **76**, 436–443 (2002).

137. Gutiérrez-Granados, S., Cervera, L., Segura, M. . L., Wölfel, J. & Gòdia, F. Optimized production of HIV-1 virus-like particles by transient transfection in CAP-T cells. *Appl Microbiol Biotechnol* **100**, 3935–3947 (2016).

138. Gutiérrez-Granados, S. *et al.* Production of HIV virus-like particles by transient transfection of CAP-T cells at bioreactor scale avoiding medium replacement. *J. Biotechnol.* **263**, 11–20 (2017).

139. He, T. C. *et al.* A simplified system for generating recombinant adenoviruses. in *Proceedings of the National Academy of Sciences of the United States of America* **95**, 2509–2514 (National Academy of Sciences, 1998).

140. Bussow, K. Stable mammalian producer cell lines for structural biology. *Curr Opin Struct Biol* **32**, 81–90 (2015).

141. Merten, O. W. *et al.* Large-scale manufacture and characterization of a lentiviral vector produced for clinical ex vivo gene therapy application. *Hum Gene Ther* **22**, 343–356 (2011).

142. Geisse, S. & Fux, C. Recombinant protein production by transient gene transfer into Mammalian cells. *Methods Enzym.* **463**, 223–238 (2009).

143. Lai, T., Yang, Y. & Ng, S. K. Advances in mammalian cell line development technologies for recombinant protein production. *Pharmaceuticals* **6**, 579–603 (2013).

144. Bandaranayake, A. D. & Almo, S. C. Recent advances in mammalian protein production. *FEBS Letters* **588**, 253–260 (2014).

145. Turan, S., Kuehle, J., Schambach, A., Baum, C. & Bode, J. Multiplexing RMCE: Versatile Extensions of the Flp-Recombinase-Mediated Cassette-Exchange Technology. *J. Mol. Biol.* **402**, 52–69 (2010).

146. Turan, S. *et al.* Recombinase-mediated cassette exchange (RMCE): Traditional concepts and current challenges. *Journal of Molecular Biology* **407**, 193–221 (2011).

147. Geisse, S. Reflections on more than 10 years of TGE approaches. *Protein Expr Purif* **64**, 99–107 (2009).

148. Dalby, B. *et al.* Advanced transfection with Lipofectamine 2000 reagent: Primary neurons, siRNA, and high-throughput applications. *Methods* **33**, 95–103 (2004).

149. González-Domínguez, I., Cervera, L., Gòdia, F. & Roldán, M. Quantitative colocalization analysis of DNA delivery by PEI-mediated cationic polymers in mammalian cells. *J. Microsc.* **273**, 53–64 (2019).

150. Benjaminsen, R. V., Mattebjerg, M. A., Henriksen, J. R., Moghimi, S. M. & Andresen, T. L. The possible "proton sponge" effect of polyethylenimine (PEI) does not include change in lysosomal pH. *Mol. Ther.* **21**, 149–157 (2013).

151. Grosse, S., Thévenot, G., Monsigny, M. & Fajac, I. Which mechanism for nuclear import of plasmid DNA complexed with polyethylenimine derivatives? *J. Gene Med.* **8**, 845–851 (2006).

152. Elouahabi, A. & Ruysschaert, J. M. Formation and intracellular trafficking of lipoplexes and polyplexes. *Molecular Therapy* **11**, 336–347 (2005).

153. Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

154. Carpentier, E., Paris, S., Kamen, A. A. & Durocher, Y. Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *J Biotechnol* **128**, 268–280 (2007).

155. Cervera, L., Gutiérrez-Granados, S., Berrow, N. S., Segura, M. M. & Gòdia, F. Extended gene expression by medium exchange and repeated transient transfection for recombinant protein production enhancement. *Biotechnol Bioeng* **112**, 934–946 (2015).

156. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).

157. Gutiérrez-Granados, S., Gòdia, F. & Cervera, L. Continuous manufacturing of viral particles. *Curr. Opin. Chem. Eng.* **22**, 107–114 (2018).

158. Wiegmann, V., Giaka, M., Martinez, C. B. & Baganz, F. Towards the development of automated fed-batch cell culture processes at microscale. *Biotechniques* **67**, 238–241 (2019).

159. Yongky, A. *et al.* Process intensification in fed-batch production bioreactors using non-perfusion seed cultures. *MAbs* **11**, 1502–1514 (2019).

160. Mellahi, K. *et al.* Assessment of fed-batch cultivation strategies for an inducible CHO cell line. *J. Biotechnol.* **298**, 45–56 (2019).

161. Clincke, M. F. *et al.* Very high density of CHO cells in perfusion by ATF or TFF in WAVE bioreactor. Part I. Effect of the cell density on the process. *Biotechnol Prog* **29**, 754–767 (2013).

162. Clincke, M. F. *et al.* Very high density of Chinese hamster ovary cells in perfusion by alternating tangential flow or tangential flow filtration in WAVE Bioreactor™-part II: Applications for antibody production and cryopreservation. *Biotechnol Prog* **29**, 768–777 (2013).

163. Nikolay, A., Castilho, L. R., Reichl, U. & Genzel, Y. Propagation of Brazilian Zika virus strains in static and suspension cultures using Vero and BHK cells. *Vaccine* **36**, 3140–3145 (2018).

164. Searles, J. A., Todd, P. & Kompala, D. S. Viable Cell Recycle with an Inclined Settler in the Perfusion Culture of Suspended Recombinant Chinese Hamster Ovary Cells. *Biotechnol. Prog.* **10**, 198–206 (1994).

165. Pinto, R. C. V., Medronho, R. A. & Castilho, L. R. Separation of CHO cells using hydrocyclones. *Cytotechnology* **56**, 57–67 (2008).

166. Hanemaaijer, J. H., Robbertsen, T., van den Boomgaard, T. & Gunnink, J. W. Fouling of ultrafiltration membranes. The role of protein adsorption and salt precipitation. *J. Memb. Sci.* **40**, 199–217 (1989).

167. Karst, D. J., Serra, E., Villiger, T. K., Soos, M. & Morbidelli, M. Characterization and comparison of ATF and TFF in stirred bioreactors for continuous mammalian cell culture processes. *Biochem. Eng. J.* **110**, 17–26 (2016).

168. Venereo-Sanchez, A. *et al.* Process intensification for high yield production of influenza H1N1 Gag virus-like particles using an inducible HEK-293 stable cell line. *Vaccine* **35**, 4220–4228 (2017).

169. Dowd, J. E., Jubb, A., Kwok, K. E. & Piret, J. M. Optimization and control of perfusion cultures using a viable cell probe and cell specific perfusion rates. *Cytotechnology* **42**, 35–45 (2003).

170. Karst, D. J. *et al.* Intracellular CHO Cell Metabolite Profiling Reveals Steady-State Dependent Metabolic Fingerprints in Perfusion Culture. *Biotechnol. Prog.* **33**, 879–890 (2017).

171.   Bielser, J. M., Wolf, M., Souquet, J., Broly, H. & Morbidelli, M. Perfusion mammalian cell culture for recombinant protein manufacturing – A critical review. *Biotechnology Advances* **36**, 1328–1340 (2018).

172.   Butler, M. Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. *Appl Microbiol Biotechnol* **68**, 283–291 (2005).

173.   Pinnau, I. & Freeman, B. D. Formation and modification of polymeric membranes: Overview. *ACS Symposium Series* **744**, 1–22 (1999).

174.   Nikolay, A., de Grooth, J., Genzel, Y., Wood, J. A. & Reichl, U. Virus harvesting in perfusion culture: Choosing the right type of hollow fiber membrane. *Biotechnol. Bioeng.* (2020). doi:10.1002/bit.27470

175.   Cowan, S. & Ritchie, S. Modified polyethersulfone (PES) ultrafiltration membranes for enhanced filtration of whey proteins. in *Separation Science and Technology* **42**, 2405–2418 (2007).

176.   Cervera, L. *et al.* Production of HIV-1-based virus-like particles for vaccination: achievements and limits. *Appl. Microbiol. Biotechnol.* **103**, 7367–7384 (2019).

177.   Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

178.   Steppert, P. *et al.* Purification of HIV-1 gag virus-like particles and separation of other extracellular particles. *J Chromatogr A* **1455**, 93–101 (2016).

179.   Liu, H. F., Ma, J., Winter, C. & Bayer, R. Recovery and purification process development for monoclonal antibody production. *mAbs* **2**, 480–499 (2010).

180.   van Reis, R., Leonard, L. C., Hsu, C. C. & Builder, S. E. Industrial scale harvest of proteins from mammalian cell culture by tangential flow filtration. *Biotechnol. Bioeng.* **38**, 413–422 (1991).

181.   Steppert, P. *et al.* Quantification and characterization of virus-like particles by size-exclusion chromatography and nanoparticle tracking analysis. *J Chromatogr A* **1487**, 89–99 (2017).

182.   McNally, D. J., Darling, D., Farzaneh, F., Levison, P. R. & Slater, N. K. H. Optimised concentration and purification of retroviruses using membrane chromatography. *J. Chromatogr. A* **1340**, 24–

32 (2014).

183. Montgomery, D. C. *Design and analysis of experiments*.

184. Yadav, V. G., De Mey, M., Giaw Lim, C., Kumaran Ajikumar, P. & Stephanopoulos, G. The future of metabolic engineering and synthetic biology: Towards a systematic practice. *Metab. Eng.* **14**, 233–241 (2012).

185. García-Granados, R., Lerma-Escalera, J. A. & Morones-Ramírez, J. R. Metabolic engineering and synthetic biology: Synergies, future, and challenges. *Front. Bioeng. Biotechnol.* **7**, (2019).

186. Edwards, A. & Haas, W. Multiplexed quantitative proteomics for high-throughput comprehensive proteome comparisons of human cell lines. in *Methods in Molecular Biology* **1394**, 1–13 (Humana Press Inc., 2016).

187. Zamani, L. *et al.* High Cell Density Perfusion Culture has a Maintained Exoproteome and Metabolome. *Biotechnol. J.* **13**, (2018).

OBJECTIVES

1. Analyze molecular changes associated to transient transfection and VLP production in HEK293 in order to characterize the cells' metabolic state.

2. Identify cell growth, transfection and VLP production bottlenecks subject to manipulation for VLP production improvement.

3. Modulate the HEK293 metabolic network via metabolic engineering to improve VLP budding and VLP production.

4. Optimize and intensify VLP production at bioreactor scale using an alternating tangential flow (ATF) system to operate in perfusion mode.

5. Characterize the produced VLPs in order to determine Gag stoichiometry.

6. Characterize the HEK293 secretome upon Gag::eGFP transient transfection and the coproduced extracellular vesicles.

7. Characterize the glycome of Gag VLPs and coproduced extracellular vesicles.

RESULTS

1. **CHAPTER ONE**

   Multiplexed quantitative proteomic analysis of HEK293 provides insights of molecular changes associated to the cell density effect, transient transfection and virus-like particles production.

2. **CHAPTER TWO**

   Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles.

3. **CHAPTER THREE**

   Part I. Culture media screening and selection of optimal transfection conditions.

   Part II. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures.

   Part III. Bioprocess analytics and perfusion-based continuous VLP harvest.

4. **CHAPTER FOUR**

   Characterization of HIV-1 virus-like particles and determination of Gag stoichiometry for different production platforms.

5. **CHAPTER FIVE**

   Molecular characterization of the coproduced extracellular vesicles in HEK293 during virus-like particle production.

6. **CHAPTER SIX**

   Differential *N*- and *O*-glycosylation signature of HIV-1 Gag virus-like particles and coproduced extracellular vesicles.

CHAPTER ONE

Multiplexed quantitative proteomic analysis of HEK293 provides insights of molecular changes associated to the cell density effect, transient transfection and virus-like particles production

ABSTRACT

The production of virus-like particles (VLP) has gained importance over the last few years owing to the benefits they provide compared to conventional vaccines. The biopharmaceutical industry is currently searching for safer candidates based on VLPs for new and existing vaccines and implementing new methods of manufacturing, thus allowing a more sustainable, effective and species-specific production. Despite achieving lower yields compared to traditional platforms, the use of mammalian cells provides the right post-translational modifications, and consequently, the intensification of bioprocesses using mammalian cells platforms has become a matter of pressing concern. One of the methods subject of intensification is transient gene expression (TGE), which has been proven to be highly effective regarding VLP production for preclinical or even clinical trials. In this work, a multiplexed quantitative proteomic approach has been applied to study the molecular characteristics of HEK293 cell cultures when growing at cell densities higher than $4 \cdot 10^6$ cells/mL, and the effects related to cell transfection and to VLP production. The obtained results revealed a set of functional and metabolic profiles of HEK293 under these three different conditions that allowed the identification of physiological bottlenecks regarding VLP production. Regarding the cell density effect (CDE), molecular alterations in the cell biology were proposed helping explain the difficulty for the cells to be transfected at higher densities. In addition, an overall disruption of cellular homeostasis after transfection was observed based on altered biological processes and after identifying potential pathways liable to be optimized via metabolic engineering, different solutions are proposed to improve VLP production.

ABBREVIATIONS

CDE: Cell Density Effect; **DTT**: Dithiothreitol; **EDTA**: Ethylenediaminetetraacetic acid; **eGFP**: enhanced green fluorescence protein; **F17**: FreeStyle F17™ cell culture media; **FASP**: Filter Aided Sample Preparation; **Gag::eGFP**: translational fusion of HIV-Gag protein and eGFP; **hpt**: hours post transfection; **LC-MS/MS**: Liquid Chromatography coupled with tandem mass spectrometry; **MVB**: Multi Vesicular Bodies; **PBS**: Phosphate-buffered saline; **PEI**: polyethyleneimine; **PVDF**: Polyvinylidene fluoride; **RFU**: Relative fluorescence units; **ROS**: Reactive Oxygen Species ; **SDS**: Sodium Dodecyl Sulfate; **SGE**: Stable Gene Expression; **TEAB**: Triethylammonium bicarbonate; **TFA**: Trifluoroacetic acid; **TGE**: transient gene expression; **THF**: Tetrahydrofolic acid; **TMT**: Tandem Mass Tag; **VLP**: virus-like particle.

INTRODUCTION

The industry of biopharmaceutical manufacturing has widely adopted the use of animal cell cultures as the platform production due to the complexity of the biological products that are being used for new therapies. The eukaryotic cellular machinery catalyzes the correct folding of these macromolecular structures, coordinates the assembly of different subunits and incorporates post-translational modifications, essential for the final biomolecule to exert the proper biological function[1]. Recombinant subunit vaccines are one of the biopharmaceuticals produced in animal cell cultures currently growing in the market[2,3]. This approach is much safer and efficient than traditional vaccines built on live-attenuated or inactivated viruses containing no viral genetic material and thus, being non-infectious. Recombinant subunits of viral proteins have been used as vaccines since 1986 when the recombinant hepatitis B surface antigen (HBsAg) proved to auto-assemble into particles mimicking the native structure of the virus that had immunogenic capacity. Therefore, conferring immunity against the virus[4].

This is the beginning of the so-called Virus-like particles (VLPs). Since then, there has been a rapid increase in the development of new forms of production, designs for new antigens and new targeted viruses, becoming a potential platform for unprecedented therapies. One of these novel candidates is the Human

Immunodeficiency Virus (HIV) whose core protein can undergo auto-assembly and coordinates to form a VLP. This core is formed by Gag polyprotein, which comprises six different parts: matrix protein (MA), capsid protein (CA), spacer peptides (SP1, SP2), nucleocapsid protein (NC) and p6[5]. The native HIV is an enveloped virus, meaning that it leaves the cell by budding, carrying with it part of the host cell membrane. In order to produce an enveloped HIV-1 VLP, the producer cell needs to express the gene coding for Gag polyprotein and allow cell budding. The cell machinery responsible for sending Gag to the membrane and for the later excision and release of the particles via budding is the endosomal sorting complexes required for transport (ESCRT)[6–10]. Two motifs from Gag p6 interact with TSG101 from complex ESCRT-I and ALIX (PDCD6IP)[11], the coordinator protein between ESCRT-I and ESCRT-III. This interaction leads to the recruitment of subunits of ESCRT-III[12] that are placed forming a ring that strangles the plasma membrane forming the new out coming enveloped particle[10]. ESCRT-III is the main responsible for excision[8]. The final step to release the new particle is carried out by the effector protein Vps4 which upon ATP hydrolyzation performs the cleavage[13]. The role of ESCRT-II is still uncertain as it appears it is not involved in HIV budding[14] although some examples proving its requirement for the process have been shown[15]. Its role remains unclear. Alternative accessory proteins also play a role in sending Gag to the membrane. It is the case of the ubiquitin ligase NEDD4L[16–18]. Ubiquitinated proteins interact with domains of proteins from ESCRT-0 and might interact with ESCRT-II in the process of sending Gag to the membrane[19]. N-myristylation of Gag MA has also been reported to have targeting effects[6,11].

Mammalian cell lines like Chinese Hamster Ovary (CHO) or Human Embryonic Kidney 293 (HEK 293) are being exploited for this purpose thanks to their ability to produce recombinant proteins in suspension cultures upon transfection[20]. In order to achieve this, two main strategies have been developed, stable gene expression (SGE) and transient gene expression (TGE). SGE is based on the integration of the protein-coding gene in the genome while TGE relies on episomal DNA expression. While SGE is commonly used for well-established and characterized bioprocesses[21], TGE still prevails among previous stages of clinical trials due to the versatility of the technique, the need of performing different screening tests and the reduced costs of implementing TGE compared to SGE[22,23]. TGE involves transient transfection of the gene

of interest, which enters the cell complexed by compounds like the cationic polymer polyethylenimine (PEI)[24]. The cell culture begins producing at around 2 hours post-transfection (hpt) and the product is further harvested between 48 and 72 hpt[25]. TGE has been described as a potent method of recombinant protein production[26] together with the use of HEK293 owing to the suitability of this cell line for transfection.

Although the implementation of TGE has been demonstrated at larger scale, where it has been reported to work at 100L scale in stirred tanks[27], it still remains predominant at preclinical stages since transfection yield decreases in mammalian cell cultures. These lower yields could be overcome via strategies involving transfection at higher densities. This presents important problems due to the cell density effect. Transfection is usually carried out at cell densities of $2 \cdot 10^6$ cells/mL and it is considered high cell density from $4 \cdot 10^6$ cells/mL up to higher orders of magnitude. If transfection is carried out when growing cells exceed $4 \cdot 10^6$ cells/mL, it becomes more difficult to successfully transfect them and the percentage of transfected cells abruptly decreases. Many strategies have been implemented to surpass this obstacle as diluting the cells to $2 \cdot 10^6$ cells/mL prior to transfection and changing the media to avoid spent media interference but the nature of problem remains unclear. Centrifuging cells at $2 \cdot 10^6$ cells/mL to concentrate before transfection proved to work[28] suggesting the referred cell density effect might be related and conditioned by the intrinsic metabolic state of the cells. The same problematic appears when working with insect cell lines[29].

With the implementation of perfusion in bioprocessing productions, cell cultures are able to reach very high densities up to $50\text{-}150 \cdot 10^6$ cells/mL [30,31]. The rapid advances in the field are making essential the need to characterize metabolic bottlenecks regarding transfection and production and some insights of the metabolic causes of the cell density effect. This work has focused on the study of the changes in the HEK293 proteome when growing without transfection, when transfected with an empty plasmid and when transfected with Gag-coding gene to produce Gag VLPs, demonstrating the differential proteins changes in each cellular state. These findings will provide new keys for further metabolic engineering and optimization of VLP production.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from McGill University (McGill, Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Culture media was FreeStyle™ F17 Expression Media (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) supplemented with 8 mM GlutaMAX™ (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA), 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and IGF-1 at a final concentration of 50µg/L.

Cell counts and viability were performed using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

Glucose and lactate were measured using YSI 2700 select glucose/lactate analyzer (YSI, Yellow Springs, OH, USA) according to manufacturer's instructions.

TRANSIENT TRANSFECTION

Transfections were carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 µg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used.

The plasmid used contained a gene coding for HIV-Gag protein fused in frame to eGFP (Gag::eGFP). This is a translational gene fusion of HIV-*gag* gene and *egfp* gene under a CMV enhancer and CMV promoter with no IRES sequence between them. Briefly, pGag::eGFP plasmid was diluted with supplemented FreeStyle™ F17 media and vortexed for 10 seconds. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, then the mixture was incubated for 15 min at room temperature and added to the cell culture.

FLOW CYTOMETRY

Samples were taken at each time point and cells were fixed using formaldehyde 2% during 10 minutes, centrifuged and then resuspended in PBS for FACS analysis. The percentage of GFP positive cells was assessed using a BD FACS Canto flow cytometer (BD Biosciences, San Jose, CA, USA). Laser 488 was used for GFP measurement. The results were analyzed with FACS DIVA software (BD Biosciences, San Jose, CA, USA).

HIV-1 GAG VLP QUANTIFICATION

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using an in-house developed and validated quantification assay[32]. VLP containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples. There is a linear correlation between fluorescence intensity and p24 values determined using the INNOTEST ELISA HIV antigen mAb (Innogenetics NV, Gent, Belgium). RFU values can be converted to Gag::eGFP concentration values using the following equation:

$$Gag::eGFP\,(ng/mL) = (3.245 \times RFU - 1.6833) \times 36 \qquad (1)$$

where Gag::eGFP is the estimated concentration of polyprotein and RFU is the measured GFP fluorescence intensity in the samples. The first term is the correlation equation between fluorescence values and p24 concentrations determined by ELISA and 36 is a correction factor that takes into account the difference in molecular weight between p24 and Gag::eGFP and an underestimation arising from using the p24 ELISA to estimate p55 Gag concentrations. Assuming that a single VLP contains 2,500 Gag::eGFP molecules[33] and that one Gag::eGFP is 84kDa ($1,39 \cdot 10^{-10}$ ng), the concentration of VLPs can be calculated.

PROTEIN SAMPLES PREPARATION FOR MASS SPECTROMETRY ANALYSES

Pellets of each tested condition were obtained by centrifugation of 500 µL of the cell culture at 10000xg for 15 minutes at 4ºC. Pellets were stored at −80ºC.

Protein extraction was performed using extraction buffer (100 mM Tris-HCl pH=8.8, 2mM EDTA, 4% SDS, 100mM DTT) of which 100µL were added to the cell pellet of each condition. Samples were sonicated for 5 minutes and then boiled for another 5 minutes. Protein extracts from pellets samples were quantified with RC/DC Protein Assay (Bio-Rad, Hercules, CA, USA) and stored in -20 ºC until the tryptic digestion process. Protein digestion was performed as previously described[34]. Proteins were digested using sequencing grade trypsin (Promega, Madison, WI, USA) and the filter-assisted sample preparation technology (FASP, Expedeon, San Diego, CA, USA), and the resulting peptides were subjected to TMT-10 plex labelling (AB Sciex, Framingham, MA, USA), joined and desalted. A total of 150µg from samples of each condition was diluted to a final concentration of 100mM of TEAB labeled with TMT-10 plex according to the manufacturer. Protein samples were labeled by adding 41µL of TMT isobaric tag diluted in anhydrous acetonitrile, followed by a 1h-incubation step at room temperature. To quench the reaction, 5% hydroxyl-amine (8µL per sample) was added, incubated 15 min at room temperature and mixed together followed by addition of TFA 1% to lower pH at 3. TMT-labeled samples were equally mixed. Pooled mix was purified using Oasis HLB C18 column (Waters, MA, USA).

TMT-labeled mix was fractionated using High pH reversed-phase peptide fractionation kit (Thermo Scientific, San Jose, CA, USA) according to manufacturer's instructions into 5 fractions for further LC-MS/MS analysis.

LIQUID CHROMATOGRAPHY TANDEM MASS SPECTROMETRY ANALYSIS

The tryptic peptide mixtures were subjected to LC-MS/MS analysis on a nano-HPLC Easy nLC 1000 liquid chromatograph (Thermo Scientific, San Jose, CA, USA) coupled to a QExactive mass spectrometer (Thermo Scientific, San Jose, CA). Peptides were suspended in 0.1% formic acid, loaded onto a C18 reverse-phase trapping column (Acclaim PepMap100, 75-µm internal diameter, 3-µm particle size and 2-cm length, Thermo Scientific), and separated on an analytical C18 nano-column (EASY-Spray column PepMap RSLC

C18, 75-µm internal diameter, 3-mm particle size and 50-cm length, Thermo Scientific), in a continuous gradient (8–31%B in 240 min, 31–90%B in 2min, 90%B in 7 min, and 2%B in 30min; where buffer A is 0.1% formic acid in HPLC grade H_2O and buffer B is 100% ACN, 0.1% formic acid in HPLC grade H_2O). Spectra were acquired using full ion scan mode over the mass-to-charge (*m/z*) range 390–1500, 70,000 FT-resolution was performed on the top 15 ions in each full MS scan using the data-dependent acquisition mode with 45s dynamic exclusion enabled.

PROTEIN IDENTIFICATION AND QUANTIFICATION

Protein identification was performed over the raw files using the SEQUEST HT algorithm integrated in the Proteome Discoverer 2.1 (Thermo Finnigan). MS/MS scans were matched against a human database (UniProtKB/Swiss-Prot 2017_10 Release). Sequence of Gag::eGFP protein was added to the selected database to enable identification.

For database searching, parameters were selected as follows: trypsin digestion with 2 maximum missed cleavage allowed, precursor mass tolerance of 800 ppm, fragment mass tolerance of 0.02 Da. TMT-10 plex labeling at N-terminal and lysine (+229.62932 Da) as well as cysteine carbamidomethylation (+57.021 Da) were chosen as static modifications whereas methionine oxidation (+15.994915 Da) was chosen as dynamic modification. The same MS/MS spectra collections were searched against inverted database constructed from the same target database. SEQUEST results were analyzed by the probability ratio method[35]. False discovery rate (FDR) for identified peptides was calculated in the inverted database search results using the refined method[36].

Quantitative information of TMT reporter ions was extracted from MS/MS spectra for relative quantification of protein abundance to characterize dynamic protein expression profiles in the selected conditions.

STATISTICAL ANALYSIS

For the comparative analysis of the protein abundance changes we applied Weighted Scan- Peptide-Protein (WSPP) statistical workflow[37,37], using SanXoT package[39]. It uses as an input a list of quantifications in the form of log2-ratios (for example a condition versus control sample) with their statistical weights and

generates the standardized forms of the original variables computing the quantitative values expressed in units of standard deviation around the averages. The quantitative information is obtained from the spectra and used to quantify the peptides from which the spectra are produced and then, proteins that generate these peptides. In other words, the quantitative information is integrated from the spectrum level to the peptide level and then from the peptide level to the protein level[40]. These standardized variables, (Zq), express the quantitative values in units of standard deviation[41]. For the protein functional analysis, Systems Biology Triangle (SBT) model[42] was used. This algorithm estimates weighted functional category averages (Zc) from the protein values by performing the protein to category integration. After the integration from spectra to peptide, the integration from peptide to protein and protein to category represent a higher level. This integration allows the detection of changes in functional categories produced by the coordinated behavior of their proteins[40]. The quantified proteins were functionally annotated using the Gene Ontology database[43,44]. For further Gene Ontology annotation, DAVID[45,46] was used to perform functional enrichment analysis and to extract *p-values* for the enriched processes. To help analyze and comprehend the data, the online software for reactions, proteins and pathways analysis REACTOME[47] was used.

WESTERN BLOTS

RC/DC protein quantification was used to normalize protein used for western blot. A total of 20 g of each condition was separated on SDS-PAGE and transferred onto a polyvinylidene difluoride (PVDF) membrane for 7 minutes using the system Trans-Blot® Turbo™ Transfer System (Bio-Rad, Bio-Rad, Hercules, CA, USA) as described in the instructions. Membranes were incubated overnight with diluted primary antibody in 5% (w/v) non-fat dry milk 1x TBS 0.1% Tween-20 at 4°C with gentle shaking. Primary antibodies used for protein validation were anti-Histone H2A antibody (L88A6, Cell Signaling Technology, 1:1000), anti-Histone H3 antibody (96C10, Cell Signaling Technology 1:1000), anti-Importin4 antibody (ab181037, Abcam, UK, 1:10000), anti-HMGCS1 antibody (ab155787, Abcam, UK, 1:3000), anti-ACSF2 antibody (ab180068, Abcam, UK, 1:1000), mouse Anti-TSG101 antibody (612696, BD Biosciences), mouse monoclonal antibody to HIV-1 p24 (A2-851-100, Icosagen, Estonia, 1:1000). After primary incubation, a

secondary incubation was performed using anti-Rabbit IgG coupled with alkaline phosphatase antibody produced in goat (A9919, Sigma Aldrich, San Luis, MO, USA) and anti-Mouse IgG coupled with alkaline phosphatase antibody produced in goat (A3562, Sigma Aldrich, San Luis, MO, USA) as required in 2.5% (w/v) non-fat dry milk 1x TBS 0.1% Tween-20 for 1 hour at room temperature. Proteins were visualized using NBT-BCIP® solution (Sigma Aldrich, San Luis, MO, USA) incubating the membrane for 2-3 minutes. Membranes were let to dry and then scanned at 400 bpi and analyzed using the software ImageJ[48].

EXPERIMENTAL DESIGN

For the multiplexed quantitative proteomics experiment based on TMT-10 plex labeling, three conditions were tested in this study: no transfection (N) transfection with empty mock plasmid (M) and transfection following the standard protocol (S) previously described. The mock plasmid shares the same backbone as the plasmid coding for Gag::eGFP but lacks the whole *gag::egfp* gene. Samples of each condition were taken at 2 hours post-transfection (2 hpt) and 72 hours post-transfection (72 hpt). As it is depicted in **Figure 1A**, the TMT-based isobaric labelling quantification was performed with three biological replicates for each condition. A peptide pool from the non-transfected condition at 2hpt was used as an internal control for the following statistical analyses.

RESULTS

The following analyses are described using the following notation: non-transfected condition (**N**), mock condition (**M**) and standard Gag::eGFP transfected condition (**S**). The standardized log2-ratios (from analyzed condition versus the internal control) measured in units of deviation (**Zq**) was used to perform the following analyses. The same type variable but at the level of functional category was also used (**Zc**).

Cell culture growth, VLP production and quantification

Cell cultures were inoculated at $0.5 \cdot 10^6$ cells/mL and transfected when reached $2 \cdot 10^6$ cells/mL. As shown in **Figure 1B**, after transfection (at time point 0), transfected conditions (M and S) stopped cell growth and

maintained their concentration at $2 \cdot 10^6$ viable cells/mL while cell viability started to slightly decrease from 98% to 78% at 72 hpt. Conversely, the non-transfected control condition (N) continued growing reaching $6 \cdot 10^6$ cells/mL at 72 hpt maintaining viability at 96%. Both M and S conditions behave similarly, regardless of the nature of the plasmid that the cells have been transfected with. To monitor transfection, samples of each time point from the conditions transfected with pGag::eGFP were analyzed by flow cytometry to quantify the percentage of GFP positive cells. At 72 hpt cells reached the highest percentage of transfection (*p-value* = 0.02) with 95% of GFP positive cells as shown in **Figure 1C**. Production of VLPs was assessed by fluorimetry and converted to VLP concentration (VLP/mL) using an in-house developed protocol[32]. **Figure 1D** shows the significant increase (*p-value* = 0.0002) of particle production from 24 to 48 hpt and were harvested at its maximum at 72 hpt when particle concentration reached $1.6 \cdot 10^{10}$ VLPs/mL.

Transfection at different cell densities was tested from 2 to $5 \cdot 10^6$ cells/mL. Transfection efficiency dramatically decreases if transfection is performed at cell densities higher than $3 \cdot 10^6$ cells/mL. When transfected at $2 \cdot 10^6$ cells/mL, transfection efficiency reaches 92% at 72 hpt whereas if transfected at $4 \cdot 10^6$ cells/mL, percentage of transfection drops to 3%. Transfection was already reduced at $3 \cdot 10^6$ cells/mL, hindered at $4 \cdot 10^6$ cells/mL and completely inhibited at $5 \cdot 10^6$ cells/mL (**Figure 1E**), even using media able to sustain cell densities of $7\text{-}8 \cdot 10^6$ cells/mL where glucose was maintained above 1g/L up to 5 days of culture (**Figure 1F**). The same effect was observed using HyCell™ TransFx-H (HyClone) media (**Supplementary Figure S1**).

Proteomic characterization of HEK293 at 2 and 72 hours post HIV-1 Gag transfection

A comparative analysis of the HEK293 cell proteome without transfection, transfected with an empty plasmid and transfected with the plasmid containing the gene *gag::egfp* at 2 and 72 hours post transfection as carried out by high-throughput multiplexed quantitative proteomic analysis. A total of 7262 proteins were identified and quantified in this study (FDR < 0.01), with 4403 proteins quantified with more than one peptide (**Supplementary Table S1)** and 3335 shared in both time points (**Figure 2A**).
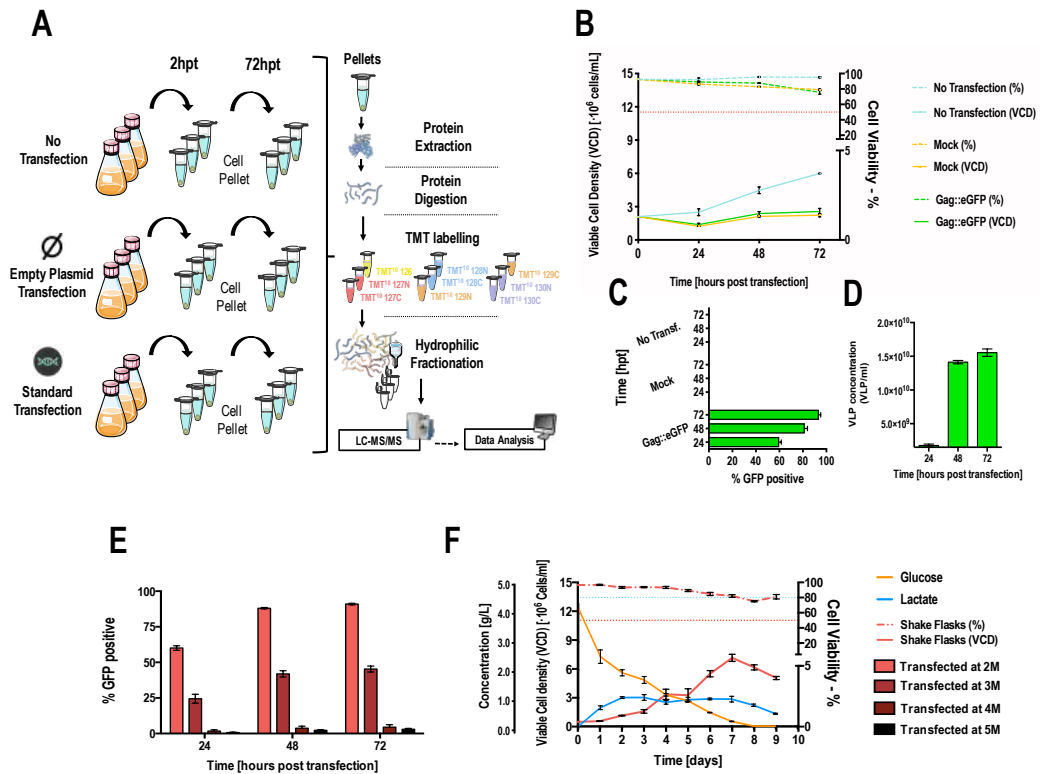
Figure 1. Experimental design and cell growth and VLP production data. **A)** Experimental workflow. Three biological replicates of HEK293 cells were cultured in three different conditions. No transfected condition, transfected with an empty plasmid denoted as 'mock' and transfected using the standard protocol with the plasmid containing the gene gag::egfp gene. Both transfections were performed at the same cell density of $2 \cdot 10^6$ cells/mL. At the time point of 2 hours post transfection (hpt) cultured samples were taken, centrifuged and cellular pellets were stored at -80°C. Same procedure was repeated at 72hpt. Proteins were extracted from the pellet samples, digested and peptides were labelled using Tandem Mass Tag (TMT) labelling. Labelled peptides were fractionated and analysed via LC-MS/MS. **B)** Viable cell density and cell viability graphs along the time course. The red dotted line indicates 50% of cell viability. Cells were transfected at time point 0. **C)** Flow cytometry analysis of the transfection percentage (showed in GFP positive percentage of cells) in the different conditions along the time course. **D)** VLP concentration in the supernatants of the standard transfection condition along the time course, showing the production of VLPs. **E)** Flow cytometry analysis of the transfection percentage (showed in GFP positive percentage of cells) of different conditions each transfected at a different cell density from 2 to $5 \cdot 10^6$ cells/mL. **F)** Growth curve of HEK293 in batch culture using FreeStyle F17 medium. Blue and red dotted lines indicate 80% and 50% of cell viability respectively. Solid orange and blue lines represent glucose and lactate respectively monitored throughout the culture.

The functional enrichment analysis performed on the 667 unique proteins (**Figure 2B**) found at 2hpt confirmed that at that time point the most significant biological process is protein transport (*p-value* = $7.2 \cdot 10^{-8}$). However, this enrichment was not due to Gag production since no Gag::eGFP was detected at 2hpt (**Figure 2D**). At 2hpt, the DNA/PEI complexes are still being taken up by the cell and

transported to the nucleus but no expression of *gag::egfp* is yet detected[49]. Another significantly enriched biological processes were proteasome-mediated ubiquitin-dependent protein catabolic process, establishment of protein localization, protein ubiquitination and protein K63-linked deubiquitination. A similar enrichment analysis was performed on the unique 401 proteins found at 72hpt. Coherently, at this time point, when VLPs are being produced, Gag::eGFP is detected by western blot (**Figure 2D**). Gag::eGFP is the quantified protein showing the highest significance (FDR=0). The significantly enriched biological processes found are related to the secretion of proteins into the extracellular space and cell division, such as regulation of small GTPase mediated signal transduction, endocytosis, cytoskeleton-dependent intracellular transport, microtubule-based movement and chromosome segregation (**Figure 2C**). Interestingly, the endocytosis process is related to the formation of the Multi Vesicular Bodies (MVB) component, which allows intracellular budding of VLPs[50] and later cell release as exosomes. To further characterize the differences between 2 and 72hpt, proteins and processes were grouped depending on their up or downregulation at each timepoints (**Figure 2E**). Proteins were selected with Zq greater or equal than 1.5 for upregulations and lower or equal than -1.5 for downregulations. Processes identified and quantified with more than 4 proteins were selected with Zc greater or equal than 1.5 for upregulations and lower or equal than -1.5 for downregulations. It can be drawn from this analysis that 86% of all altered processes are identified and quantified at 72hpt. Regardless of the studied condition, there is a clear molecular and metabolic shift that is taking place from 2 to 72hpt.

These results are an overall view of the initial and final state of HEK293 cells along the transient transfection time course, albeit more detailed analyses can be drawn out focusing on individual protein changes in each condition. Protein up and downregulations presented in all three conditions (N, M and S) were attributed to cell growth, regardless of transfection, providing insight of the metabolic alterations that could be found in HEK293 cultures growing at high cell density. Changes presented in each transfected condition, regardless of the plasmid used, could be explaining the transient transfection process. Finally, changes found only in the condition transfected with the plasmid pGag::eGFP could be deciphering the effect of VLP production on the cell physiology. The next analyses were carried out following these guidelines.
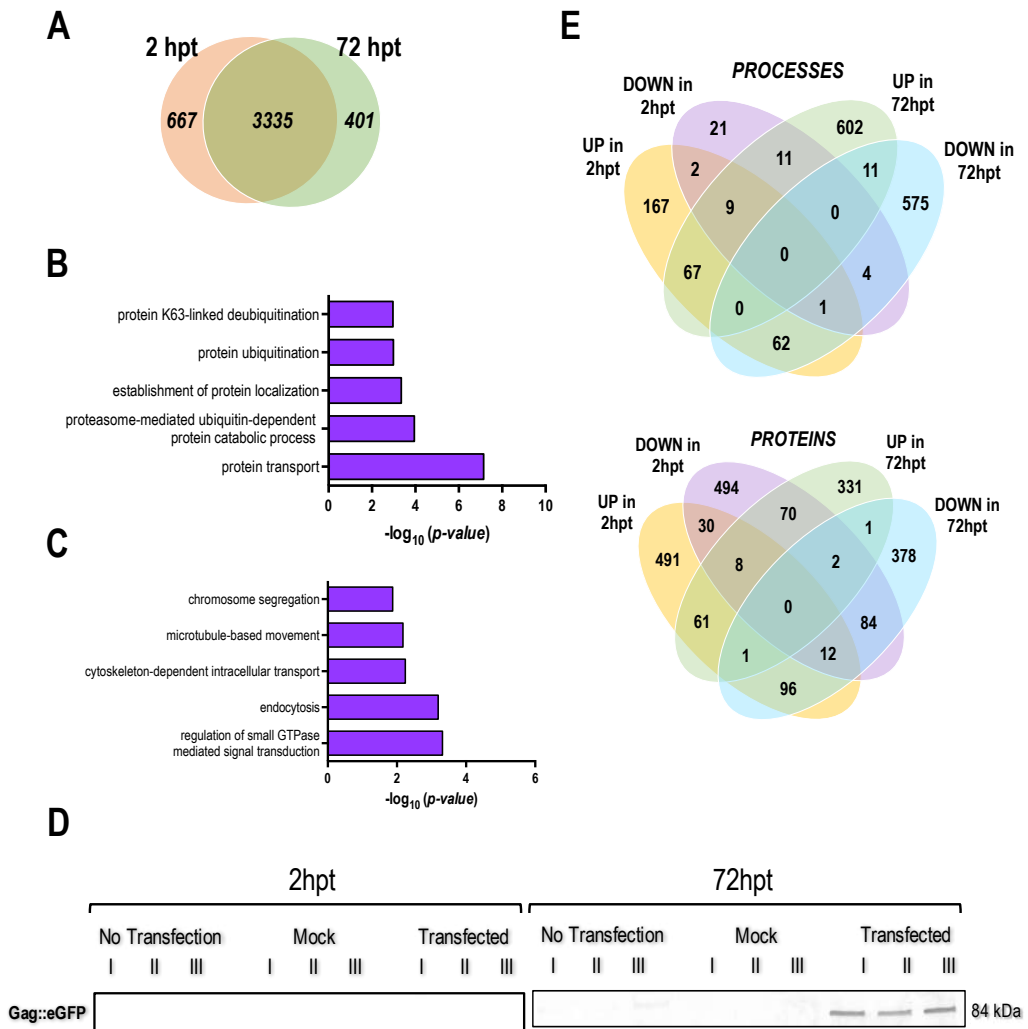
Figure 2. Analysis of relevant differences between 2hpt and 72hpt . **A)** Venn diagram of proteins identified with more than 1 peptide at 2 hpt (667 proteins), 72 hpt (401 proteins) and both (3335 proteins) all present in the three studied conditions. **B)** Biological processes significantly enriched related to the 667 unique proteins found at 2hpt. **C)** Biological processes significantly enriched related to the 401 unique proteins found at 72hpt. **D)** Western blot of p24 of Gag polyprotein. **E)** Venn diagram representing up and downregulated proteins and processes at 2 and 72hpt.

Proteomic analysis of HEK293 cell growth

Proteins showing a common pattern of alteration in all three conditions comparing the time points 2 and 72hpt using Zq, were classified in four groups such as transport proteins (importins and exportins), lipids and metabolites biosynthesis, mitochondrial and histones proteins. **Figure 3** shows protein abundance changes (Zq) of all the proteins showing the same alteration in all conditions. In **Figure 3A** it can be observed that the four selected group of proteins are the ones showing the highest significance and greatest abundance changes. All identified importins showed a downregulation at 72hpt mainly indicated by importin subunit alpha-1, 8 and 4, importin-11, importin-5 and importin-4. Moreover, all identified exportins also presented downregulation at 72hpt from which exportin-T, exportin-1 and exportin-2. Lipids and metabolites biosynthesis also showed downregulation at 72hpt. Proteins like fatty acid synthase, cytoplasmic hydroxymethylglutaryl-CoA synthase, squalene synthase, glutamine synthetase, ATP-citrate synthase, thymidylate synthase and up to 18 biosynthesis related enzymes showed a significant decrease at 72hpt. Likewise, an increase in mitochondrial protein composition can be found at 72hpt. Up to 30 mitochondrial proteins presented upregulation compared to 2hpt (**Figure 3A**). A more significant increase appeared in the non-transfected condition (N). In addition to this, histones (H2A, H3) and histones-related proteins like lysine-methyltransferases were also significantly increased at 72hpt. Validation of relevant and representative proteins of each group were performed using western blot (**Figure 3B-C**).

Proteomic analysis of HEK 293 transfection

Upon transfection, cellular viability starts decreasing (**Figure 1B**). A broader overview of the changes in cell physiology upon transient transfection is listed in **Figure 4**. Unlike the changes related to the cell density effect, transfection seems to influence homeostasis in the entire cell, affecting a wide range of pathways. This systemic alteration was analyzed using biological processes GO terms. At 72 hours post transfection, there was an overall decrease in several metabolic and physiological processes, as approximately 79% of all analyzed processes were downregulated (**Figure 4C**). The upregulated
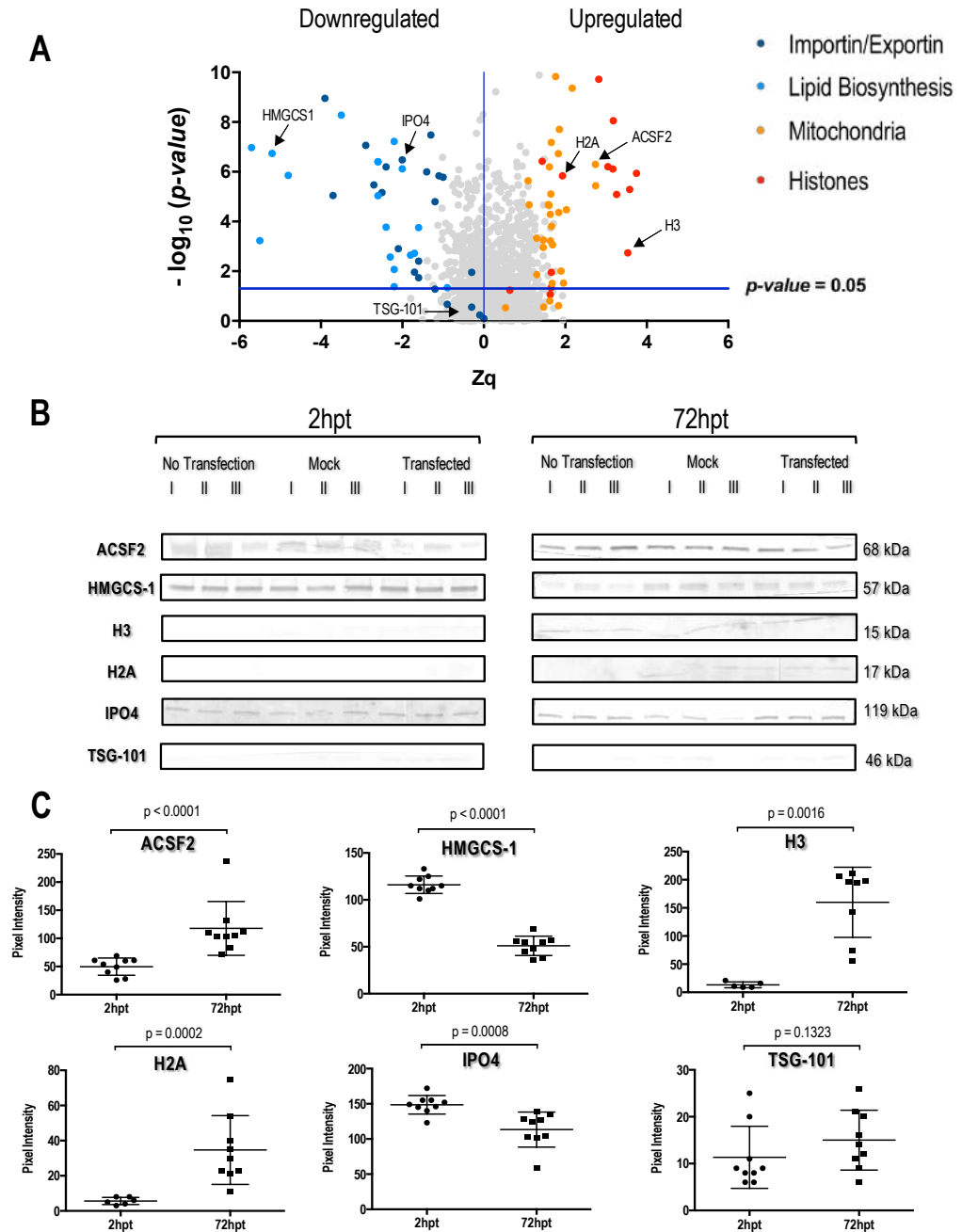
Figure 3. Protein changes related to cell growth. **A)** Representation of the proteins showing up or downregulations in the three studied conditions: N (no transfected), M (transfected with mock) and S (standard transfection) at the same time, from 2 to 72hpt. Different colours represent the different analysed groups of proteins. **B)** Western blots of relevant proteins from each group of interest. I, II and III represent the biological replicates. **C)** Dot plots showing the change in expression from 2hpt to 72hpt from these relevant proteins. Variations are calculated using the 9 biological replicates. Medians are represented by horizontal bars. *p-values* are calculated using Mann Whitney U test.

processes corresponded to the remaining 21%, involving gene expression, protein production and microvesicles/exosomes machinery activation. Interestingly, both ER and cytosolic ribosome protein translation were upregulated. Regarding exocytosis, protein trafficking towards the cell membrane was pointed out in the analysis as GO terms such as phosphatidylinositol-4, vesicle docking and the cytoskeleton are increased (**Figure 4A**). This suggested that the network cytosolic protein production and ER-Golgi network trafficking were both enhanced upon transfection.

Curiously, there was a group of processes that not only were downregulated upon transfection, but also increased in the non-transfected condition (N) at 72hpt compared to 2hpt (**Figure 4D**), suggesting that transfection prevented cells from upregulation interfering especially in these processes. Coherently to the phenomena observed in **Figure 1B**. The complexation agent (PEI) used for transfection prevents cell growth due to its cytotoxic effects. Therefore, processes related to DNA regulation are downregulated upon transfection. DNA repair, nuclear chromosome formation and histone exchange were downregulated. Metabolic processes such as glycosphingolipid metabolism and NADP+/ADP ribosyltransferase activity also belong to this group of altered processes. The rest of downregulated processes indicated in **Figure 4** cover some homeostatic pathways mostly related to detoxification and damage repair. Homeostasis is disrupted at different levels. At DNA level, processes like DNA binding, mismatch edition or telomere maintenance were downregulated. At enzyme level, many metabolic enzymes whose function depend on metal ion binding such as Co^{2+} binding, THF, pyrimidines or xenobiotic binding and detoxification were decreased. Peptidase activity is reduced as well. Redox homeostasis is also disrupted. Glutathione-related processes like glutathione biosynthesis, reductase activity, dehydrogenase activity, and glutathione binding are decreased. Cell signaling and the response to stimulus also appeared to be altered since Ca^{2+} binding proteins, Ca^{2+}-related processes like Ca^{2+} sequestration, release, regulation, channel localization and ryanodine receptors involved in signaling pathways, all showed downregulation at 72hpt in transfected conditions compared to non-transfected ones.
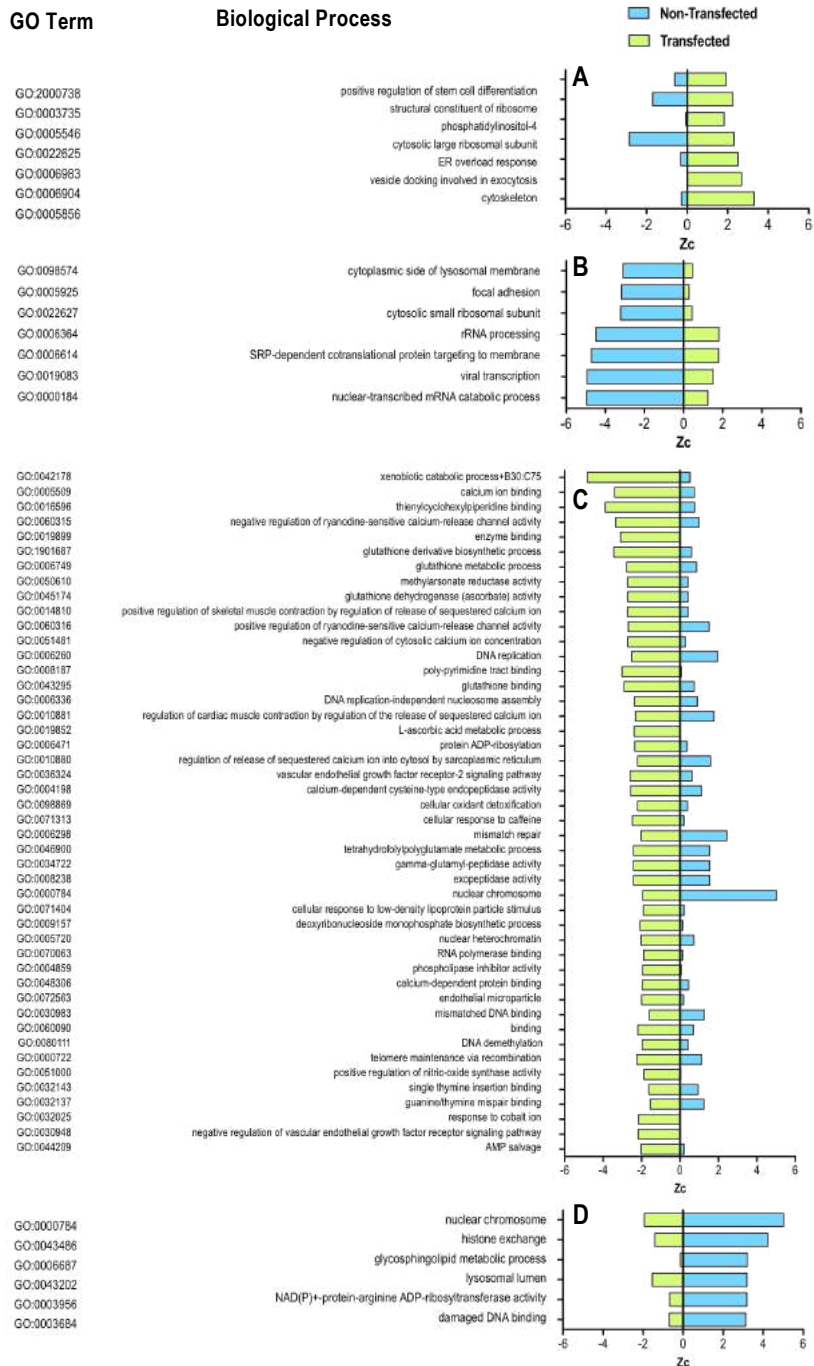
Figure 4. Altered biological processes due to transfection. **A)** Processes upregulated upon transfection. **B)** Processes whose downregulation is prevented by transfection. **C)** Processes downregulated by transfection. **D)** Processes whose upregulation is prevented by transfection. Blue and green represent non-transfected and transfected condition respectively.

Proteomic analysis of HEK293 VLP production

HIV-Gag VLPs were successfully identified being Gag::eGFP is the protein identified with the highest intensity in the whole study. Regarding cellular physiology, compared to the alterations observed after transfection, VLP production had a minor effect on cell physiology (**Figure 5)** compared to the homeostatic alterations due to transfection. The most remarkable change was the increase of Gag::eGFP protein at 72hpt while the rest of the proteins showing alterations were only detected at 2hpt. Within these altered proteins, two groups were observed. The first group were proteins upregulated in conditions non-transfected (N) and transfected with mock plasmid (M) while showing no change in VLP-producing condition (S). Belonging to this group, ubiquitin ligases of RNF family and accessory proteins like Sharpin were found. Cell adhesion and proteins relevant in the extracellular matrix composition like protocadherin and collagens were also altered. The second group is formed by proteins showing downregulation in S compared to the rest. Similar proteins as the ones found in the previous group, were also found here. Different ubiquitin ligases from RNF family, transcription repressors, zinc fingers, Ca^{2+} -dependent proteins and voltage channels were present. There were no significant processes showing upregulation when producing VLPs.

The ESCRT machinery

The machinery responsible for processing and secreting Gag-based VLPs is the one comprising the ESCRT complexes[8]. The heat map of all ESCRT proteins in the different conditions at 2 and 72hpt is shown in **Figure 6A**. All ESCRT proteins were identified and quantified. However, no significant changes were observed regarding the components of the ESCRT complexes comparing 2 and 72hpt. Although no protein showed individual significant increase or decrease, the overall variation of the whole ESCRT-I showed a significant decrease (p-value < 0.0001) from 2 to 72hpt, showing a coordinated behavior of all of its proteins (**Figure 6B**). ESCRT-II and ESCRT-III complexes did not show significant variations. Variation of TSG-101 was assessed by western blot and showed no significant difference from 2 to 72 hpt (**Figure 3C**) validating the previous results. Accessory proteins like NEDD8 and NEDD4L ubiquitin ligases, also

Figure 5. Heat map showing main changes upon VLP production. Colours represent Zq of each protein in each condition. Conditions are denoted by N (no transfected), M (transfected with mock) and S (standard transfection).

important for sending Gag to the cell membrane and promoting cell budding were quantified. Interestingly, at 72hpt this group of proteins were found to be significantly downregulated.

DISCUSSION

The Cell Density Effect (CDE)

Regardless of any external distress, the mere cell growth influences the metabolic and physiological state of the cell affecting the later processes of transfection and production. It has been widely reported that productivity in cell cultures decreases when transfection is carried out at higher cell densities[29,51-53]. This phenomenon is known as Cell Density Effect (CDE). It affects different production approaches: transfection[54], infection[29], mammalian[52,53,55] and insect cells[56]. The limitation imposed by this constraint prevents the production process from being thoroughly exploited when scaled up. If cultures need to be maintained at early exponential phase ($1\text{-}2\cdot10^6$ cells/mL) to be effectively transfected, the potential of scaling up is reduced[57]. During the last decade different studies regarding the CDE has been conducted and showed some metabolic variations in high cell density cultures, mostly regarding energy[53,56,58]. CDE was thought to be mainly a metabolic issue, concerning the limited resources available when high concentrations of cells are reached. ATP was described to help attenuate the CDE[59] as perfusion cultures started to be implemented[30] since a continuous addition of fresh media provides the necessary source of carbon, preventing glucose from depletion and maintaining ATP production. However, when carrying out transient transfection at high cell densities operating in perfusion mode, difficulties to transfect the cells still appeared[60]. Encountering the CDE in perfusion mode suggests that molecular mechanism are influencing this phenomenon beyond energetic issues and media depletion is not a determining factor in the CDE. Cell division requires energy and gene expression. As cells reach high densities along with cells which are recovering from a cytotoxic event like transfection, the need of energy increases. Our results showed an upregulation of histones, chromosomal proteins, gene expression and mitochondrial components that reflects the increase in energy demand and it is intimately related to the CDE. However,

A

B
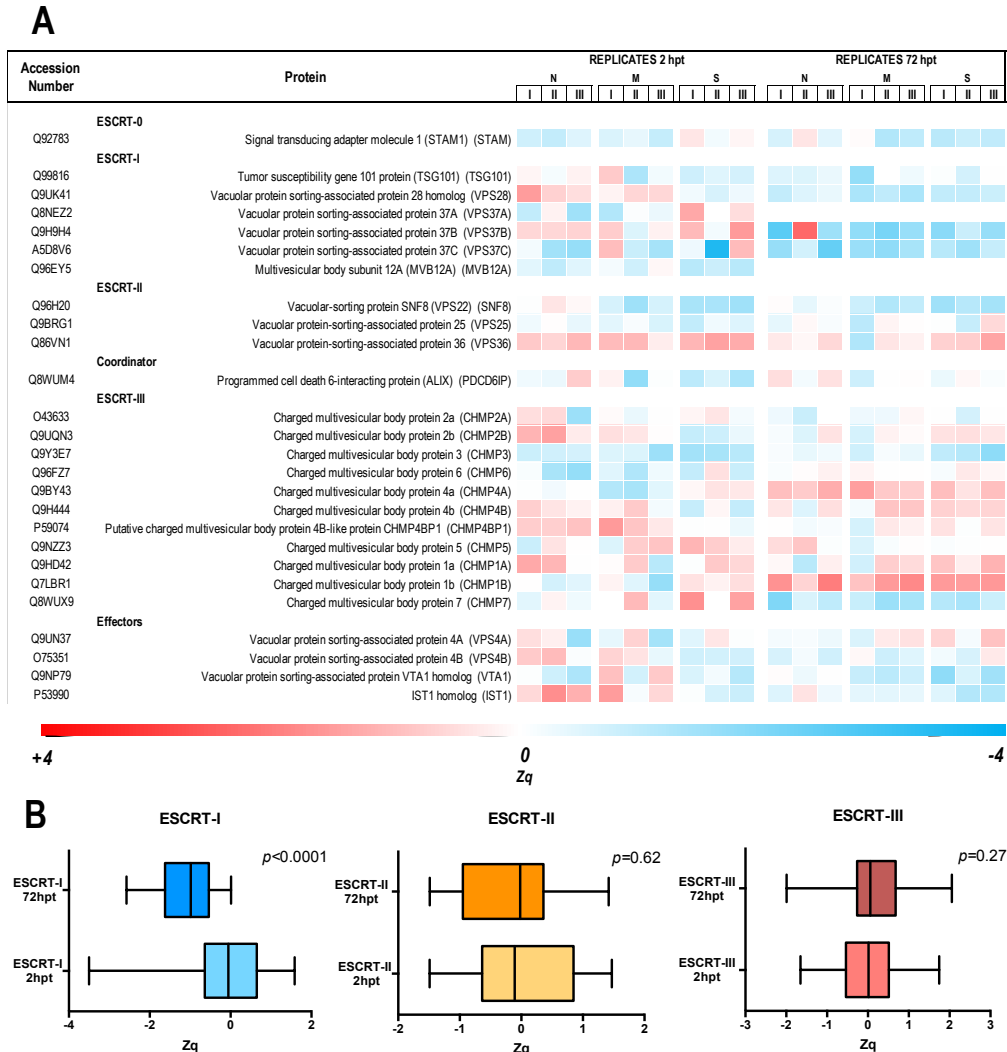
Figure 6. A) Heat map showing all components of the different ESCRT complexes at 2 and 72 hpt. Colours represent Zq of each protein in each condition. Conditions are denoted by N (no transfected), M (transfected with mock) and S (standard transfection). **B**) Box plots representations of the Zq values composing the different ESCRT complexes at 2 and 72hpt. Medians are represented by vertical bars and whiskers extend to extreme data points. *p-values* are calculated using Mann Whitney U test. *p-values* are calculated using Mann Whitney U test).

energy is not the only element behind the CDE and different physiological causes remain unknown[53].

Using media able to sustain densities of 10^7 cells/mL, the same effect was observed at $3\text{-}4 \cdot 10^6$ cells/mL, suggesting there are more cellular processes playing a role rather than being only an energy issue.

Maintaining glucose above 1g/L in the conditions where a decrease in transfection efficiency was observed indicates that energy depletion is not a limiting and influencing factor.

In addition to this, how the cell cycle affects transfection efficiency was already reported[25] and showed that when transfecting HEK293 at high cell densities cell cycle does not present a limiting factor and does not influence the difficulty found to achieve high transfection efficiency. Still, no molecular contribution has been reported so far regarding this issue. This study helps cast some light on some molecular mechanisms that are being altered at densities above $4 \cdot 10^6$ cells/mL which are strongly related to the event of transfection and processing of the exogenous DNA. When DNA/PEI complexes are added for transfection, they first encounter various cell membranes until they reach the nucleus, where the plasmid stays as episomal DNA[49]. To be able to be transported into the nucleus the machinery of nuclear importin/exportin play a crucial role. DNA binds to importins and only then is able to cross the nuclear pores, following the unbinding of DNA/importing complex coupled with the hydrolysis of GTP into GDP[61]. Exportins work similarly transporting cargos from the nucleus to the cytoplasm. The downregulation of the whole nucleus transport machinery (exportins and importins showed in **Figure 3**) at high cell densities might explain the decrease of transfection efficiency and thus be a relevant component of the CDE. Our work showed an essential component of the plasmid up taking pathway hitherto omitted in similar studies. Not only the energy balance is important to describe and characterize this effect, but also molecular mechanisms affecting the cell biology.

Another important factor is the decrease in lipid biosynthesis. A change in the lipid composition of the cell membrane might be essential to prevent the DNA/PEI complexes from entering the cell and block transfection. In addition, budding from the cell membrane is the secretory pathway used by HIV-VLPs and therefore for this platform. This implies a continuous remodeling and regeneration of the cell membrane to be able to cope with the constant production and release of VLPs. If the overall lipid biosynthesis is reduced, cells that recently acquired the plasmid might not engage in producing VLPs and cells that already are producing could start to slow the production rate. This, coupled with cell division and the subsequently loss of the plasmid will concur in a decrease of productivity from 72hpt along time[25].

Physiological consequences upon transfection

Transient transfection is a well-defined cytotoxic event. The addition of DNA/PEI complexes have proved to reduce cellular viability[62]. To avoid this problem, optimization strategies need to be found to improve PEI-mediated transfection[63]. Cytotoxicity of PEI has already been described[64]. However, the physiological cause wreaking havoc on cell culture viability still remained undescribed. Here, several altered biological processes have been identified showing homeostasis disruption and casting some light on the reasons for the cell viability to be reduced. Glycosphingolipid biosynthesis has previously been reported to increase during cell growth and even described as one of the metabolic causes of the cell density effect[53,65]. In non-transfected conditions, glycosphingolipid metabolic processes significantly increase. Yet upon transfection it is hindered. Glycosphingolipids act in the intracellular vesicle trafficking to promote cholesterol trafficking from cytosol to the membrane[66]. It is reported that its loss from the cytosolic cell membranes and lipid rafts alter vesicle trafficking[67] and the correct cholesterol transport and consequently correct lipid metabolism[68]. Interestingly, this cholesterol homeostasis disruption coincides with the downregulation of lipid biosynthesis previously described. The loss of relevant lipids from the cell membrane and the loss of lipid trafficking and lipid homeostasis prevents the cell from displaying a healthy signaling to the rest of the cell population[69] and more importantly, it alters essential physiological processes, contributing to the reduction of cellular viability. Glycosphingolipid metabolic process category comprises proteins like prosaposin, beta-hexosaminidase, sphingomyelin phosphodiesterase and acid ceramidase among others. This suggests that breaking down ceramide and ceramide-derived lipids is blocked after transfection. Consequently, the release of fatty acids and other lipids from their glycosphingolipid precursors is hampered. Enhancing these pathways or adding fatty acids and other lipids, like cholesterol, when transfecting may clear this metabolic bottleneck. An additional approach in order to improve cell homeostasis upon transfection could be overexpressing enzymes promoting glycosphingolipids metabolism. An improvement in cellular homeostasis could enhance VLP production and grant the acquisition of higher protein titers.

Another important factor for cell survival is maintaining redox homeostasis. DNA/PEI complexes are highly charged molecules which when entering the cell modify cytosolic compartment polarization as calcium-

dependent processes happen to be altered. The change in intracellular polarization is crucial for the right protein signaling and protein-protein interaction. Polarization changes also affects the mitochondria which can consequently alter energy homeostasis and increase reactive oxygen species (ROS)[70] damaging the cell. Curiously, proteins helping detoxifying the cell from ROS such as glutathione are also downregulated therefore suggesting redox homeostasis disruption caused by DNA/PEI complexes also contributes to decrease cellular viability. These series of changes may affect protein integrity, nevertheless some peptidases were also found to be decreased, hampering correct protein degradation. Accumulation of harmful undegraded proteins present a high risk for cellular homeostasis. Although PEI could be detoxified by the cell, this general homeostasis and protein disruption also affects xenobiotics catabolic processes, as well hindered. Homeostasis disruption might not be the only cause for cellular viability to decrease upon transfection. Exogenous DNA might be also playing a role. Cellular machinery to repair, modify and bind DNA was downregulated. These processes are essential for DNA renewal in exponentially growing cells. If the cell cannot use this sort of mechanisms, it will find solutions for DNA detoxification. All of the biological processes found to be upregulated upon transfection point to the biosynthesis and release of exosomes. Interestingly, the production of exosomes has recently been described as a powerful mechanism of excreting DNA from the cell[71]. Upon a homeostasis-disruptive event, the cell engages mechanisms to eliminate the source of cellular stress generation. This upregulation in exocytosis suggests that DNA/PEI complexes could be being excreted via exosomes production in order to maintain homeostasis.

Physiological consequences of producing HIV-1 VLPs

Cells transfected with pGag::eGFP plasmid shared all previously described physiological effects caused by transient transfection. However, when cells start producing VLPs, they display some specific physiological changes related to the expression of episomal DNA and remodeling of the membrane proteins. At this scenario, most of the cell resources are being used either to fight homeostasis disruption or to produce VLPs. This translates into the downregulation of production and release of growth-signaling proteins. Cell communication is essential and VLP-producing cells need to focus their resources on surviving and

producing so cell growth is slowed down. Production requires DNA expression, so specific zinc fingers, transcriptional repressors and DNA regulators were found downregulated, allowing expression of the exogenous DNA. There were not upregulated ribosomal processes, indicating that the cell does not enhance protein synthesis pathways, but instead, uses the constitutive protein synthesis activity to channel production towards VLPs. This might be the first bottleneck in VLP production that can be found. After the monomers of Gag are produced, they need to be sent to the membrane for assembly. The phase of sending Gag monomers to the membrane and their assembly is really important to determine physiological bottlenecks for the following optimization and improvement of VLP production. Previous work has already reported that the limiting step in VLP production is the transport of Gag monomers to the cell membrane[49] but the responsible molecular elements remained unknown. In this work, a close analysis of the ESCRT molecular pathway has been performed in order to identify the limiting step. The comparative of the different ESCRT complexes at 2 and 72hpt indicates that ESCRT-I presents a bottleneck in VLP production. ESCRT-I is responsible for sending Gag to the membrane[8,9] and at 72hpt showed an overall downregulation. This indicates that Gag is being produced but is not being sent to the cytoplasmic membrane at its fullest capacity. No alteration of ESCRT-II and ESCRT-III imply that even if it is not presenting any hindrance, it can be further optimized if any upregulation is achieved. In order to tackle this issue, accessory proteins like NEDD8 and NEDD4L which were found downregulated at 72hpt could be overexpressed to promote transportation of Gag to the membrane and enhance cell budding. Another possibility to improve VLP production is to promote intracellular budding. Biological processes like endocytosis or late endosomal-related pathways were quantified at 72hpt. VLPs are described to be produced also by intracellular budding[50,72] and overexpressing proteins like Citron-Rho Kinase could be a solution to promote this pathway. Once VLPs are being produced, secreted from the cell membrane, they take part of it with them during excision. VLPs are charged particles, meaning their constant production will interfere with cell membrane polarization, altering Ca^{2+} import. More interestingly, Ca^{2+} has been described to play an important role regarding the multivesicular bodies (MVB) pathway, also leading to VLP intracellular budding. The downregulation in Ca^{2+} import observed when producing VLPs can be overcome by adding Ca^{2+} together with ionomycin, improving VLP production via intracellular budding
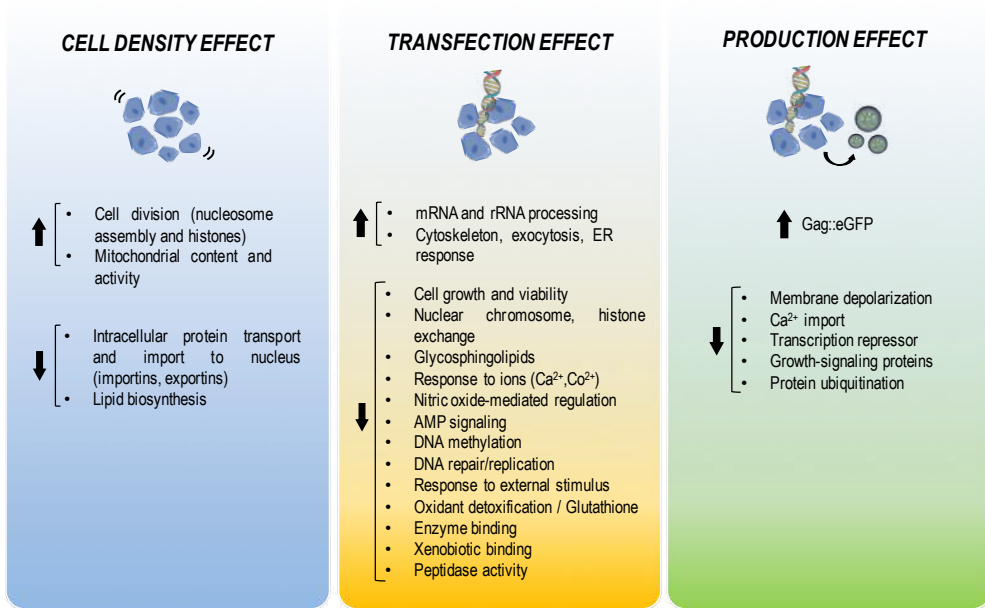
Figure 7. Summary of quantified changes present in all conditions (Cell density effect), present in transfected conditions (transfection effect) and present only in the VLP-producing condition (production effect).

by eight fold[73]. The combination of Citron-Rho kinase overexpression and the addition of free intracellular calcium and ionomycin could help increase VLP budding. Collagens and other extracellular proteins were as well altered, suggesting extracellular matrix composition itself is being modified along VLP production. Specific ubiquitin-ligase downregulated activity at 2hpt might indicate that ubiquitination is being redirected to a few processes, which could be directly related to Gag trafficking to the membrane.

CONCLUSIONS

Increasing the knowledge of cell physiology is key to understand and characterize the bioprocess that is trying to be optimized. The cellular causes for some well-known and observable effects in animal cell bioprocessing still remain not fully understood. In this work, some physiological causes are identified related to the cell density effect, the reduction of cell viability upon transient transfection and VLP production. A summary of the altered processes in each condition is shown in **Figure 7.** The difficulty to achieve transfection when it is carried out at cell densities greater than $4 \cdot 10^6$ cells/mL is known as the cell density effect. This is partially explained by a downregulation in lipid biosynthesis and intracellular protein transport to the nucleus. Transient transfection depends on the physiological state of the cells and thus it is directly related to cell density. The decrease in cellular viability upon transfection is caused by a systemic disruption of homeostasis due to many levels of homeostasis control being altered, such us glycosphingolipids metabolism, calcium regulation, oxidant detoxification, xenobiotic metabolism, peptidases activity and DNA detoxification. When producing VLPs, these traits are maintained and specific modifications in the extracellular matrix and membrane calcium channels are evidenced.

These findings could support designing less stressful processes enabling the productivity to increase, to reduce bottlenecks and help establish fundamental cell physiological basis for future, more effective bioprocesses.

SUPPORTING INFORMATION
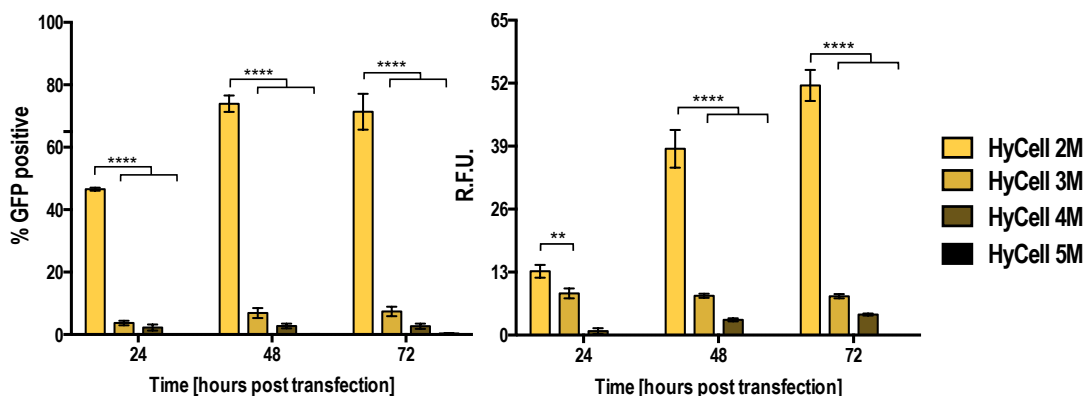
Supplementary Figure S1: Decrease in transfection efficiency at high cell density in HyCell™ TransFx-H media. Flow cytometry analysis of the transfection percentage (showed in GFP positive percentage of cells) and RFU measurement of different conditions each transfected at a different cell density from 2 to $5 \cdot 10^6$ cells/mL in HyCell™ TransFx-H (HyClone) media.

Supplementary Table S1. (List of all Identified proteins in this study) can be found at:

https://pubs.acs.org/doi/10.1021/acs.jproteome.9b00601?goto=supporting-info

DATA AVAILABILITY

The raw mass spectrometry data has been submitted to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) with the dataset identifier PXD014746.

REFERENCES

1. Bandaranayake, A. D. & Almo, S. C. Recent advances in mammalian protein production. *FEBS Lett* **588**, 253–260 (2014).

2. Roldão, A., Mellado, M. C., Castilho, L. R., Carrondo, M. J. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev Vaccines* **9**, 1149–1176 (2010).

3. Nascimento, I. & Leite, L. Recombinant vaccines and the development of new vaccine strategies. in *Braz J Med Biol Res* **45**, 1102–1111 (2012).

4. Gurramkonda, C. *et al.* Purification of hepatitis B surface antigen virus-like particles from recombinant Pichia pastoris and in vivo analysis of their immunogenic properties. *J Chromatogr B Anal. Technol Biomed Life Sci* **940**, 104–111 (2013).

5. Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136–144 (2013).

6. Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

7. von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

8. Henne, W. M., Buchkovich, N. J. & Emr, S. D. The ESCRT pathway. *Dev Cell* **21**, 77–91 (2011).

9. Wollert, T. *et al.* The ESCRT machinery at a glance. in *J Cell Sci* **122**, 2163–2166 (2009).

10. Scourfield, E. J. & Martin-Serrano, J. Growing functions of the ESCRT machinery in cell biology and viral replication. *Biochem Soc Trans* **45**, 613–634 (2017).

11. Fujii, K., Hurley, J. H. & Freed, E. O. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nat Rev Microbiol* **5**, 912–916 (2007).

12. McCullough, J., Fisher, R. D., Whitby, F. G., Sundquist, W. I. & Hill, C. P. ALIX-CHMP4 interactions in the human ESCRT pathway. *Proc Natl Acad Sci U S A* **105**, 7687–7691 (2008).

13. Alonso, Y. A. M., Migliano, S. M. & Teis, D. ESCRT-III and Vps4: a dynamic multipurpose tool for membrane budding and scission. *Febs j* **283**, 3288–3302 (2016).

14. Langelier, C. *et al.* Human ESCRT-II complex and its role in human immunodeficiency virus type

1 release. *J Virol* **80**, 9465–9480 (2006).

15. Meng, B., Ip, N. C., Prestwood, L. J., Abbink, T. E. & Lever, A. M. Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. *Retrovirology* **12**, 72 (2015).

16. Chung, H. Y. *et al.* NEDD4L overexpression rescues the release and infectivity of human immunodeficiency virus type 1 constructs lacking PTAP and YPXL late domains. *J Virol* **82**, 4884–4897 (2008).

17. Usami, Y., Popov, S., Popova, E. & Gottlinger, H. G. Efficient and specific rescue of human immunodeficiency virus type 1 budding defects by a Nedd4-like ubiquitin ligase. *J Virol* **82**, 4898–4907 (2008).

18. Pincetic, A. & Leis, J. The Mechanism of Budding of Retroviruses From Cell Membranes. *Adv Virol* **2009**, 6239691–6239699 (2009).

19. Hurley, J. H. ESCRT complexes and the biogenesis of multivesicular bodies. *Curr Opin Cell Biol* **20**, 4–11 (2008).

20. Dyson, M. R. Fundamentals of Expression in Mammalian Cells. *Adv Exp Med Biol* **896**, 217–224 (2016).

21. Bussow, K. Stable mammalian producer cell lines for structural biology. *Curr Opin Struct Biol* **32**, 81–90 (2015).

22. Ausubel, L. J. *et al.* Production of CGMP-Grade Lentiviral Vectors. *Bioprocess Int* **10**, 32–43 (2012).

23. Merten, O. W. *et al.* Large-scale manufacture and characterization of a lentiviral vector produced for clinical ex vivo gene therapy application. *Hum Gene Ther* **22**, 343–356 (2011).

24. Geisse, S. & Fux, C. Recombinant protein production by transient gene transfer into Mammalian cells. *Methods Enzym.* **463**, 223–238 (2009).

25. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

26.   Geisse, S. Reflections on more than 10 years of TGE approaches. *Protein Expr Purif* **64**, 99–107 (2009).

27.   Tuvesson, O., Uhe, C., Rozkov, A. & Lullau, E. Development of a generic transient transfection process at 100 L scale. *Cytotechnology* **56**, 123–136 (2008).

28.   Backliwal, G., Hildinger, M., Hasija, V. & Wurm, F. M. High-density transfection with HEK-293 cells allows doubling of transient titers and removes need for a priori DNA complex formation with PEI. *Biotechnol Bioeng* **99**, 721–727 (2008).

29.   Huynh, H. T., Tran, T. T., Chan, L. C., Nielsen, L. K. & Reid, S. Effect of the peak cell density of recombinant AcMNPV-infected Hi5 cells on baculovirus yields. *Appl Microbiol Biotechnol* **99**, 1687–1700 (2015).

30.   Genzel, Y. *et al.* High cell density cultivations by alternating tangential flow (ATF) perfusion for influenza A virus production using suspension cells. *Vaccine* **32**, 2770–2781 (2014).

31.   Clincke, M. F. *et al.* Very high density of CHO cells in perfusion by ATF or TFF in WAVE bioreactor. Part I. Effect of the cell density on the process. *Biotechnol Prog* **29**, 754–767 (2013).

32.   Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

33.   Cheeks, M. C., Edwards, A. D., Arnot, C. J. & Slater, N. K. H. Gene transfection of HEK cells on supermacroporous polyacrylamide monoliths: A comparison of transient and stable recombinant protein expression in perfusion culture. *N. Biotechnol.* **26**, 289–299 (2009).

34.   Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

35.   Martínez-Bartolomé, S. *et al.* Properties of average score distributions of SEQUEST: the probability ratio method. *Mol. Cell. Proteomics* **7**, 1135–45 (2008).

36.   Bonzon-Kulichenko, E., Garcia-Marques, F., Trevisan-Herraz, M. & Vazquez, J. Revisiting peptide identification by high-accuracy mass spectrometry: problems associated with the use of narrow mass precursor windows. *J Proteome Res* **14**, 700–710 (2015).

37. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J Proteome Res* **13**, 1234–1247 (2014).

38. Martínez-Acedo, P. *et al.* A novel strategy for global analysis of the dynamic thiol redox proteome. *Mol. Cell. Proteomics* **11**, 800–13 (2012).

39. Trevisan-Herraz, M. *et al.* SanXoT: a modular and versatile package for the quantitative analysis of high-throughput proteomics experiments. *Bioinformatics* **35**, 1594–1596 (2019).

40. Garcia-Marques, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol Cell Proteomics* **15**, 1740–1760 (2016).

41. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **13**, 1234–47 (2014).

42. García-Marqués, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol. Cell. Proteomics* **15**, 1740–60 (2016).

43. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

44. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

45. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2009).

46. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).

47. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

48. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–5 (2012).

49. Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

50. Ding, J., Zhao, J., Sun, L., Mi, Z. & Cen, S. Citron kinase enhances ubiquitination of HIV-1 Gag protein and intracellular HIV-1 budding. *Arch Virol* **161**, 2441–2448 (2016).

51. Ferreira, T. B., Carrondo, M. J. & Alves, P. M. Effect of ammonia production on intracellular pH: Consequent effect on adenovirus vector production. *J Biotechnol* **129**, 433–438 (2007).

52. Le Ru, A. *et al.* Scalable production of influenza virus in HEK-293 cells for efficient vaccine manufacturing. *Vaccine* **28**, 3661–3671 (2010).

53. Petiot, E., Cuperlovic-Culf, M., Shen, C. F. & Kamen, A. Influence of HEK293 metabolism on the production of viral vectors and vaccine. *Vaccine* **33**, 5974–5981 (2015).

54. Yang, S., Zhou, X., Li, R., Fu, X. & Sun, P. Optimized PEI-based Transfection Method for Transient Transfection and Lentiviral Production. *Curr Protoc Chem Biol* **9**, 147–157 (2017).

55. Robert, M. A. *et al.* Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. *Biotechnol J* **12**, (2017).

56. Bernal, V., Carinhas, N., Yokomizo, A. Y., Carrondo, M. J. & Alves, P. M. Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism. *Biotechnol Bioeng* **104**, 162–180 (2009).

57. Rajendra, Y., Balasubramanian, S. & Hacker, D. L. Large-Scale Transient Transfection of Chinese Hamster Ovary Cells in Suspension. *Methods Mol Biol* **1603**, 45–55 (2017).

58. Bereiter-Hahn, J., Munnich, A. & Woiteneck, P. Dependence of energy metabolism on the density of cells in culture. *Cell Struct Funct* **23**, 85–93 (1998).

59. Henry, O., Perrier, M. & Kamen, A. Metabolic flux analysis of HEK-293 cells in perfusion cultures for the production of adenoviral vectors. *Metab Eng* **7**, 467–476 (2005).

60. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).

61. Bai, H., Lester, G. M. S., Petishnok, L. C. & Dean, D. A. Cytoplasmic transport and nuclear import of plasmid DNA. *Biosci Rep* **37**, (2017).

62. Chahal, P. S., Schulze, E., Tran, R., Montes, J. & Kamen, A. A. Production of adeno-associated virus (AAV) serotypes by transient transfection of HEK293 cell suspension cultures for gene

delivery. *J Virol Methods* **196**, 163–173 (2014).

63.   Fuenmayor, J., Cervera, L., Gutierrez-Granados, S. & Godia, F. Transient gene expression optimization and expression vector comparison to improve HIV-1 VLP production in HEK293 cell lines. *Appl Microbiol Biotechnol* **102**, 165–174 (2018).

64.   Hall, A., Lachelt, U., Bartek, J., Wagner, E. & Moghimi, S. M. Polyplex Evolution: Understanding Biology, Optimizing Performance. *Mol Ther* **25**, 1476–1490 (2017).

65.   Vukelic, Z. & Kalanj-Bognar, S. Cell density-dependent changes of glycosphingolipid biosynthesis in cultured human skin fibroblasts. *Glycoconj J* **18**, 429–437 (2001).

66.   Sillence, D. J. *et al.* Glucosylceramide modulates membrane traffic along the endocytic pathway. *J Lipid Res* **43**, 1837–1845 (2002).

67.   Kamau, S. W., Kramer, S. D., Gunthert, M. & Wunderli-Allenspach, H. Effect of the modulation of the membrane lipid composition on the localization and function of P-glycoprotein in MDR1-MDCK cells. *Vitr. Cell Dev Biol Anim* **41**, 207–216 (2005).

68.   Lippincott-Schwartz, J. & Phair, R. D. Lipids and cholesterol as regulators of traffic in the endomembrane system. *Annu Rev Biophys* **39**, 559–578 (2010).

69.   Lingwood, C. A. Glycosphingolipid functions. *Cold Spring Harb Perspect Biol* **3**, (2011).

70.   Gorlach, A., Bertram, K., Hudecova, S. & Krizanova, O. Calcium and ROS: A mutual interplay. *Redox Biol* **6**, 260–271 (2015).

71.   Takahashi, A. *et al.* Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nat Commun* **8**, 15287 (2017).

72.   Loomis, R. J. *et al.* Citron Kinase, a RhoA Effector, Enhances HIV-1 Virion Production by Modulating Exocytosis. *Traffic* **7**, 1643–1653 (2006).

73.   Perlman, M. & Resh, M. D. Identification of an intracellular trafficking and assembly pathway for HIV-1 gag. *Traffic* **7**, 731–45 (2006).

CHAPTER TWO

Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles

ABSTRACT

HIV-1 Gag virus-like particles (VLPs) are promising candidates for the development of future vaccines. Recent viral outbreaks have manifested the need of robust vaccine production platforms able to adapt to new challenges while achieving mass production capacity. For the rapid production of VLPs, the method of transient gene expression (TGE) have proved highly efficient. Based on a previous characterization of the HEK293 cell line upon transient transfection using multiplexed quantitative proteomics, molecular production bottlenecks and metabolic pathways likely to be optimized were identified. In this work, these molecular components and metabolic pathways have been explored and modulated via transient metabolic engineering using approaches like design of experiments (DoE) to fully exploit and optimize VLP production and transfection and budding efficiency. Upon overexpression of endosomal sorting complex required for transport (ESCRT) accessory proteins like NEDD4L and CIT, VLP production increased 3.3 and 2.9-fold respectively. Overexpression of glycosphingolipid precursor enzyme UGCG improved transfection efficiency by 17% and knocking-down the Gag-binding protein CNP improved 2.5-fold VLP specific productivity. Combining CNP inhibition and UGCG overexpression further improved budding efficiency by 37.3%. Modulating VLP production and accessory pathways like intracellular budding, demonstrated the potential of metabolic engineering to optimize and intensify the development of robust production platforms for future vaccines.

ABBREVIATIONS

CCD: Central composite design, **CIT**: Citron Rho-interacting kinase, **CNP**: 2′,3′-Cyclic-Nucleotide 3′-Phosphodiesterase, **DMSO**: Dimethyl sulfoxide, **DoE**: Design of experiments, **eGFP**: enhanced green fluorescence protein, **ESCRT**: Endosomal sorting complex required for transport, **FA**: Fatty acids, **Gag::eGFP**: translational fusion of HIV-Gag protein and eGFP, **Gag(E40K)::eGFP**: translational fusion of HIV-Gag protein and eGFP with a change in codon 40, from glutamate to lysine, **HEK**: Human embryonic kidney, **HIV-1**: Human immunodeficiency virus, **HMGCS-1**: hydroxylmethylglutaryl-CoA-synthase 1, **Hpt**: hours post transfection, **MA**: Matrix domain of gag gene, **MVB**: Multivesicular bodies, **NMT-1**: N-Myristoyl transferase, **NTA**: Nanoparticle tracking analysis, **PBS**: Phosphate-buffered saline, **PCR**: Polymerase chain reaction, **PEI**: polyethyleneimine, **RFU**: Relative fluorescence units, **RT**: Room temperature, **SGE**: Stable gene expression, **shRNA**: short-hairpin RNA, **TEM**: Transmission electron microscopy, **TGE**: Transient gene expression, **UGCG**: UDP-glucose ceramide glucosyltransferase, **VCD**: Viable cell density, **VLPs**: Virus-like particles

INTRODUCTION

The recent outbreak of the new virus SARS-CoV-2 has evidenced the relevance of optimizing our current technology for recombinant vaccine production. HIV-1 virus-like particles (VLPs) are a promising alternative in the development of new vaccines[1] due to their immunogenic properties[2]. Moreover, their nature of enveloped particles provides them with the ability to be functionalized with exogenous antigens targeting different diseases[3,4]. This strategy has been proved successful for the development of hybrid or chimeric H1N1 Gag VLPs[5]. HIV-1 Gag polyprotein self-assembles when heterologously expressed in mammalian cell-based production platforms like HEK293 cells[6]. Gag polyprotein is transported to the cell membrane by the endosomal sorting complexes required for transport (ESCRT)[7–9]. Accessory pathways, like protein ubiquitination have been reported to promote interaction with ESCRT-0 and ESCRT-II proteins, participating in targeting Gag to the cell membrane[10]. Moreover, the addition of a myristoyl group as a post-translational modification (PTM) to the MA domain of Gag polyprotein also contributes to directing

Gag to the cell membrane[8,11]. Monomers of Gag are then oligomerized, forming a protrusion in the cell membrane. To complete the budding process, proteins from ESCRT-III strangle the cell membrane and together with the action of Vps4 release the new particle[12]. Being familiar with the complex metabolic network governing the cellular pathways for VLPs synthesis is essential to be able to determine optimal manipulations aiming to optimize and increase production. The complex topology and intricate map of links and nodes characterizing the metabolic network requires tools for holistic characterization, like multiplexed quantitative proteomics[13]. A previous study characterizing the molecular changes in HEK293 revealed up and downregulations in specific metabolic pathways upon transient transfection and HIV-1 VLP production[14]. Therefore, several targets were identified, being potential subjects of improvement via metabolic engineering to promote cell budding and VLP production. In agreement with previous reported studies of Gag trafficking[15], the main bottleneck was reported in the process of directing Gag to the cell membrane, as the ESCRT-I complex was downregulated at 72hpt compared to 2hpt. Accessory proteins like NEDD8 or NEED4L ubiquitin ligases described to interact with Gag[16-19] were also significantly downregulated. Overexpressing the ESCRT complexes has already been reported not to improve VLP production as these complexes are an important node connecting many pathways with different biological functions[20]. In this work, strategies overexpressing the accessory proteins NEDD8 and NEDD4L have been tested as well as promoting secondary ubiquitination routes reported to act on Gag VLP budding, like overexpressing the protein citron rho-interacting kinase (CIT). NEDD4L and CIT proteins have been described to have synergetic effects on the formation of Gag VLPs[21,22]. Regarding VLP budding, the protein 2′,3′-cyclic-nucleotide 3′-phosphodiesterase (CNP) was also identified in the aforementioned proteomic study. This protein has been described to bind to Gag molecules in the plasma membrane and prevent the formation of HIV-1 Gag VLPs[23]. Consequently, a strategy of expressing short-hairpin RNAs (shRNAs) aiming to knock down the expression of this protein has also been investigated. The shRNA strategy has been previously and successfully used in the enhancement of Gag VLP production, validating the suitability of the methodology to modulate HEK293 cells metabolism[24]. In addition to this, the increase of Gag N-myristoylation has been also tested overexpressing the N-myristoyl transferase 1 (NMT-1). Altogether, exploring supplementary pathways to improve the traffic of Gag to the cell membrane. N-myristoylation is

crucial for the correct VLP formation, being responsible for the inability of CHO cells to produce Gag VLPs, as they lack this type of PTM[25]. Moreover, the reported proteomic analysis addressed an overall disruption in cellular homeostasis upon transient transfection, whose effects rippled throughout the entire cellular system[14]. One important element was lipid biosynthesis, downregulated at 72hpt. To evaluate the influence on VLP production and based on previous reported studies of lipid supplementation[26], the addition of synthecol, tocopherol and fatty acids was also tested in this work. These lipids have proven to benefit HEK293 cells and contribute to TGE and VLP production[26,27]. Another remarkable change upon transfection was the downregulation of glycosphingolipid metabolism. The assessment of the role of glycosphingolipids in VLP production was tested modulating the glycosphingolipid precursor, ceramide, and overexpressing the UDP- glucose ceramide glucosyltransferase (UGCG), the first enzyme in the glycosphingolipids biosynthesis pathway. Ceramide is an interesting sphingolipid, able to target the production of small extracellular vesicles using a ESCRT-independent pathway[28,29], an unexplored pathway in VLP production. Therefore, after identifying effective molecular modulators, a design of experiments (DoE) approach was carried out to achieve an optimal combination to maximize production and VLP budding. Improving the HEK293 VLP production platform via metabolic engineering would allow the acquirement of higher VLP titers, providing insights to the development of industrial production of VLP-based vaccines.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from McGill University (McGill, Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell culture media was HyCell™ TransFx-H media from HyClone™ (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™

(Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA).

Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

When required according to planned experiments, culture media was supplemented with 0.12 mg/mL cholesterol (SyntheChol®, Sigma-Aldrich, Steinheim, Germany) diluted in PBS (HyClone, Little Chalfont, GE Healthcare, UK), fatty acids (1:2200) (F7050, SAFC Sigma-Aldrich) diluted in PBS, 2 mg/L tocopherol (T1157, Sigma-Aldrich) diluted in PBS or 2 mg/L ceramide (860052P, Avanti, Sigma-Aldrich) dissolved in dimethyl sulfoxide 99% (DMSO) (W387520, Sigma-Aldrich).

PLASMIDS

The pGag::eGFP plasmid codes for a Rev-independent HIV-1 Gag polyprotein fused in frame to the enhanced green fluorescent protein (eGFP). Plasmids pCIT and pNMT-1 were generated *in house*. Briefly, *cit* and *nmt-1* genes, originally derived from pENTR223.1 from DNASU (plasmid ID: HsCD00353950, Hemet, CA, USA) and pMON-HsNMT1 from Addgene (Addgene plasmid #67475, Watertown, Massachusetts, USA) were amplified using the primers 1-4 in **Supplementary Table S1**, bounded by restriction enzymes recognition sequences. PCR products were checked by gel electrophoresis and then digested with BssHII (R0199S, New England Biolabs, Ipswich, Massachusetts, USA) and NotI-HF® (R3189S, New England Biolabs) for pCIT, and EcoRI-HF (R3101S, New England Biolabs) and NotI-HF for pNMT-1. Fragments were cloned into pGag::eGFP vector using T4 DNA ligase (M0202S, New England Biolabs). pCI HA NEDD4L from Addgene (Addgene plasmid #27000), HA-NEDD8 from Addgene (Addgene plasmid #18711) and pcDNA3.1+/C-(K)DYK encoding UGCG from GenScript (Piscataway, NJ, USA) (Clone ID: OHu61224, Homo sapiens UDP-glucose ceramide glucosyltransferase (UGCG)) were also used.

To compare VLP production efficiencies, a total concentration of 1µg/mL of DNA was used, as previously optimized[26]. To maintain a constant total concentration of DNA, 0.5µg/mL of the plasmid encoding Gag::eGFP protein and 0.5µg/mL of the tested plasmid were used in co-transfection tests. For the control, 0.5µg/mL of pGag::eGFP plasmid was used together with an empty plasmid (Mock) at 0.5µg/mL.

For conditions combining two tested plasmids and in order to be comparable, a total of 1µg/mL and 0.5µg/mL of pGag::eGFP was maintained. Therefore, each of the tested plasmids was used at a concentration of 0.25µg/mL.

The shRNA used against CNP had been previously reported[30]. Here, it was obtained as ssDNA ultramer (5 in **Supplementary Table S1**) flanked by regions with sequence complementarity to the reported adapted U6 pGag::eGFP backbone[24]. It was cloned downstream from the U6 promoter following a Gibson Assembly protocol: 1-hour incubation at 50 °C using the Hi-Fi Master Mix 2x (New Englands Biolabs Ipswich, Massachusetts, USA[31]. The molar ratio of backbone:insert used was 1:5. The plasmid encoding Gag::eGFP and CNP shRNA was used at a concentration of 1µg/mL. The plasmid pGag(E40K)::eGFP, which encodes Gag::eGPF with an amino acid change from glutamic (E) to lysine (K) at position 40, was generated following the point mutagenesis protocol described by Zheng et al. 2004[32]. Briefly, the plasmid pGag::eGFP was amplified using two complementary oligos (6, 7 in **Supplementary Table S1**)**.** Once amplified, the PCR product was digested with the enzyme DpnI, a methylation-dependent endonuclease that only digests the non-PCR-amplified plasmid. DH5α *E. coli* was transformed with the restriction product. Then, colonies were randomly selected and grown in LB medium, followed by plasmid purification and sequencing. Plasmids containing both *cit* and *gag::egfp* genes as well as *nedd4l* and *gag::egfp* were designed adding the gene of interest flanked by CMV promoter and polyA sequence to the pGag::eGFP standard plasmid and constructed by GenScript (Piscataway, NJ, USA).

TRANSIENT TRANSFECTION

Transfections were carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 µg/mL. DNA:PEI complexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used. Briefly, the corresponding DNA mixture was diluted with supplemented HyCell™ media and vortexed for 10 seconds. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, then the mixture was incubated for 15 min at room temperature and then added to the cell culture.

FLOW CYTOMETRY

Samples were taken every 24h after transfection and cells were fixed using formaldehyde 2% during 10 minutes, centrifuged and then resuspended in PBS for FACS analysis. The percentage of GFP positive cells was assessed using a BD FACS Canto flow cytometer (BD Biosciences, San Jose, CA, USA). Laser 488 was used for GFP measurement. The results were analyzed with FACS DIVA software (BD Biosciences, San Jose, CA, USA).

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using an *ad hoc* developed and validated quantification assay[33]. VLP-containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Pellets underwent 3 freeze/thaw cycles for cell lysis, freezing for 2h at -80°C and thawing at 37°C for 30 min. After 3 cycles, samples were vortexed and resuspended in 500µL of PBS. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples.

WESTERN BLOTS

Proteins from each analyzed condition were separated on SDS-PAGE and transferred onto a polyvinylidene difluoride (PVDF) membrane for 7 minutes using the system Trans-Blot® Turbo™ Transfer System (Bio-Rad, Bio-Rad, Hercules, CA, USA) as described in the instructions. Membranes were incubated overnight with diluted primary antibody in 5% (w/v) non-fat dry milk 1x TBS 0.1% Tween-20 at 4°C with gentle shaking. Primary antibodies used for protein validation were mouse anti-HIV-1 p24 (A2-851-100, Icosagen, Estonia, 1:1000), rabbit anti-CNPase antibody (ab183500, Abcam, UK, 1:1000), rabbit anti-CIT antibody (ab86782, Abcam, UK, 1:2000), rabbit anti-NEDD4-2 antibody (ab46521, Abcam, UK, 1:1000) and mouse anti-ß-actin antibody (MA5-15739, Thermo Fisher Scientific, 1:5000) as a loading control. After primary incubation, a secondary incubation was performed using anti-rabbit IgG coupled with alkaline phosphatase antibody produced in goat (A9919, Sigma Aldrich, San Luis, MO, USA) and anti-mouse IgG coupled with alkaline phosphatase antibody produced in goat (A3562, Sigma Aldrich, San Luis, MO, USA)

as required in 2.5% (w/v) non-fat dry milk 1x TBS 0.1% Tween-20 for 1 hour at room temperature. Proteins were visualized using NBT-BCIP® solution (Sigma Aldrich, San Luis, MO, USA) incubating the membrane for 2-3 minutes. Membranes were let to dry and then scanned at 400 bpi and analyzed using the software ImageJ[34].

OPTIMIZATION OF shCNP:UGCG RATIO USING DESIGN OF EXPERIMENTS (DoE)

CNP shRNA and UGCG ratio was optimized to maximize VLP specific productivity (VLP·cell^{-1}·day^{-1}) and budding efficiency per cell (cell^{-1}). A central composite design (CCD) was selected in order to determine the optimal concentration for both responses to the two tested plasmids. Adjusted specific productivity (P_{sp}) and specific budding efficiency (B_{sp}) were calculated as presented in **Equation 1** and **Equation 2**, where C_{VLP} is the VLP concentration (VLP/mL), C_x is the concentration of viable transfected cells (10^6 cells/mL), v is the cell viability (%), d is the day of the cell culture harvest (d=3), RFU_{SN} is the normalized relative fluorescent units in the supernatant and RFU_{Pellet} is the normalized fluorescent units in the cell pellet. Both equations were adjusted considering the viability (v) of the cell culture as a factor to positively promote the conditions where viability remained high independently from the calculated specific productivity and budding efficiency per cell.

$$P_{sp} = \frac{\frac{C_{VLP} \cdot v}{C_x}}{d} \qquad (1) \qquad\qquad B_{sp} = \frac{\frac{RFU_{SN}}{RFU_{Pellet}} \cdot v}{C_x} \qquad (2)$$

The two variables were screened at five levels ($\sqrt{2}$, +1, 0, -1, -$\sqrt{2}$). CCD experimental results were fitted to a second order polynomial equation described below (Eq. 3) by non-linear regression analysis: where Y is the corresponding response, β_0 is the offset term, β_i is the linear coefficient, β_{ii} is the quadratic coefficient, β_{ij} is the interaction coefficient, and X_i and X_j are the independent variables.

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ii} X_i^2 + \sum \beta_{ij} X_i X_j \qquad (3)$$

This equation was used to predict the optimal values of the independent variables using Design-Expert® software (Design-Expert 11.1.2.0, StatEase®, Minneapolis, MN, USA). Statistical analyses expressing the quality, fitting parameters and significance of the model were also performed using Design-Expert® software.

HIV-1 GAG VLP QUANTIFICATION BY NANOPARTICLE TRACKING ANALYSIS (NTA)

Nanoparticle Tracking Analysis (NTA) was also used to quantify fluorescent particles. NTA measurements were performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag-GFP VLPs and neutral density filter for total particle by light scattering. Data were analyzed with NanoSight® NTA 3.1 software. Briefly, samples were diluted in sterile PBS (HyClone, Little Chalfont, GE Healthcare, UK) prior to injection into the device chamber to obtain a concentration of approximately 10^8 particles/mL. Particles were identified and tracked by their Brownian motion at room temperature. Capture settings were recorded with an sCMOS camera (camera level of 8 for Gag::eGFP VLP samples, and 11 for controls, viscosity: 0.9 cP) and analyzed with a detection threshold of 4.

TRANSMISSION ELECTRON MICROSCOPY (TEM)

To determine the effect of the different genes on intracellular and extracellular VLP budding, at 72 hpt, cells were fixed for 12h at 4 °C in a solution containing 1.6 % (v/v) glutaraldehyde (Merck), 1 % (w/v) OsO_4 (TAAB Lab) and 0.1 M phosphate buffer. Cells then were rinsed three times in a solution containing 0.1 M cacodylate and incubated for 30 min in 0.2 M cacodylate. The cells were then dehydrated in a graded ethanol series with increasing concentration (from 25 to 100 %) before being embedded in epoxy resin at 60 °C for 48h[21]. Ultrathin sections (70nm) of selected areas of semi-thin sections (1 μm) were cut on a Leica Ultracut UCT microtome (Leica Microsystems, Germany) placed on 200 mesh copper grids, and contrasted with uranyl acetate (30 min) and lead citrate (5 min) solutions. For the qualitative approach, 10 randomly selected areas of each grid were observed with a TEM Jeol JEM-1400 equipped with a CCD Gatan ES1000W Erlangshen camera.

RESULTS

The addition of lipids does not improve VLP specific productivity

As aforementioned, a previous proteome study evidenced that lipid biosynthesis pathways were reported to be downregulated at 72hpt[14]. Metabolic pathways like glycosphingolipid synthesis and enzymes like hydroxymethylglutaryl-CoA-synthase (HMGCS-1), the first enzyme in the biosynthetic pathway of cholesterol; fatty acid synthase, squalene synthase, acetoacetyl-CoA synthase, and various sterol synthases were also downregulated. Aiming to compensate this downregulation, transient transfections were carried out adding synthecol, fatty acids, tocopherol and ceramide to the culture media at concentrations already described to have an effect on cell metabolism[26]. Regarding cell growth, the addition of synthecol and fatty acids reduced viable cell density (VCD) at 72hpt as well as viability compared to the control while the addition of tocopherol increased VCD at 72 hpt up to $7.7 \cdot 10^6$ cells/mL (**Figure 1A**). Transient transfection was not affected by the addition of any lipid additive, showing the same percentage of transfection along the studied time course (**Figure 1B**). However, VLP production was significantly reduced (*p<0.001*) upon addition of synthecol and fatty acids while the conditions containing tocopherol and ceramide showed no improvement compared to the standard Gag::eGFP production (**Figure 1C**). Taking into account the VCD of each condition, specific productivities (**Table 1**) showed that lipid addition did not improve the Gag::eGFP productivity of $(3.5 \pm 0.1) \cdot 10^3$ VLP·cell^{-1}·day^{-1}.

Overexpression of UGCG, NEDD4L and CIT improves transfection, production and VLP budding

The transport of Gag monomers to the cell membrane has been described to be a bottleneck in the process of VLP production[15]. Therefore, to test the overexpression of previously identified potential enhancers in the HEK293 production system[14], like accessory proteins NEDD8, NEDD4L, NMT-1 and CIT, double transient transfections were carried out using plasmid coding for Gag::eGFP and the corresponding enzyme. In order to target the downregulation of the glycosphingolipid pathway, the precursor enzyme UGCG was also overexpressed aiming to improve cellular homeostasis. The optimal concentration and ratio
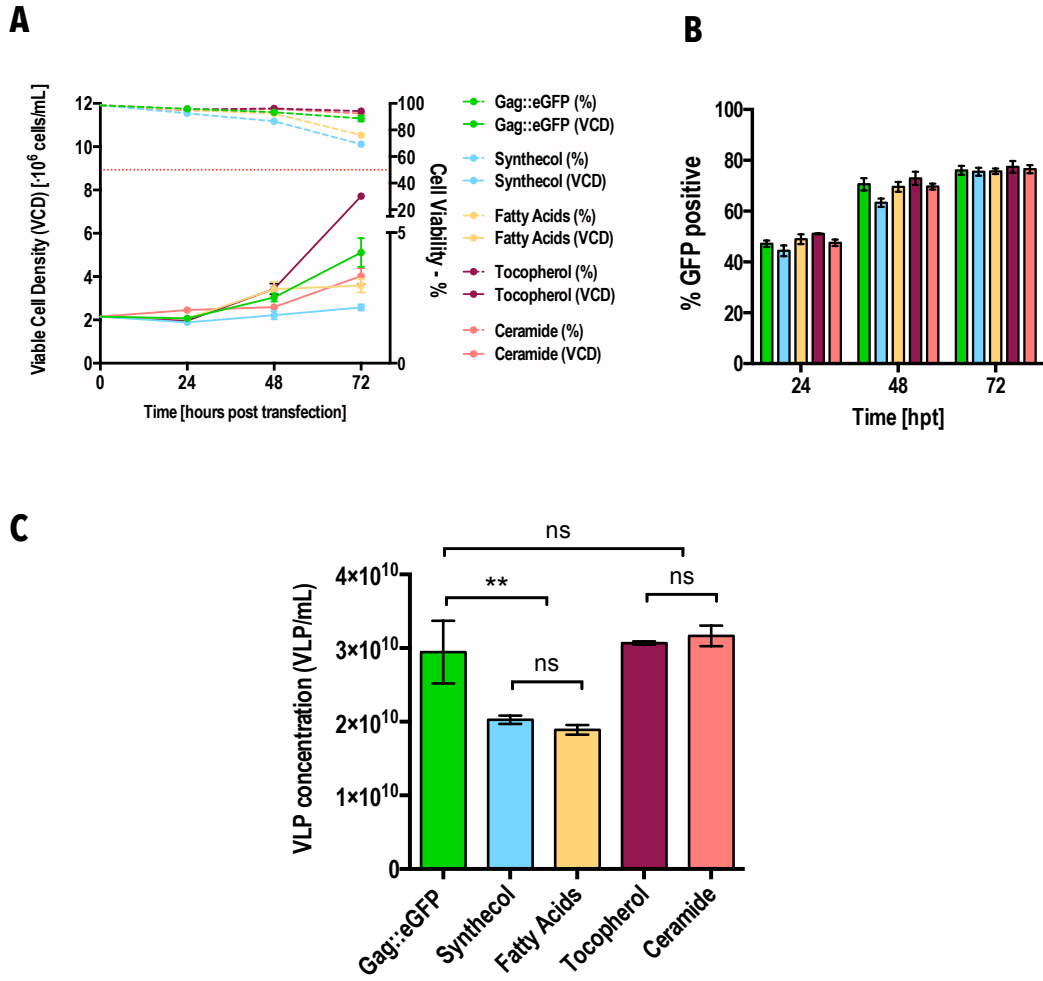
Figure 1. Analysis of HEK293 growth, transfection and VLP production upon the addition of different lipid metabolites. **A)** Cell growth and viability of the different studied conditions. Red dotted line represents 50% of viability. **B)** Percentage of transfection by flow cytometry analysis measured in % of GFP positive cells along the studied time course. **C)** Produced VLP concentration at 72hpt upon lipid addition in each studied condition. Significance was calculated using tow-way ANOVA tests.

of DNA:PEI had been previously optimized[26] and a total of 1µg/mL of DNA and a ratio (w/w) DNA:PEI of 1:2 was used for all conditions. To cotransfect the plasmid encoding Gag:.eGFP and the enzyme of interest, 0.5µg/mL of DNA from each plasmid was used, while an empty plasmid (mock) up to 1µg/mL was used for the control. Cell viability along the time course was not affected by the overexpression of any accessory enzyme and VCD slightly increased upon UGCG overexpression. More interestingly, transfection efficiency was significantly ($p < 0.0001$) improved upon UGCG overexpression, reaching ~70% of transfected cells

only at 24hpt, where the rest of conditions remained at ~40%. At 72hpt, the overexpression of UGCG led to a percentage of transfection of ~90%, improving 17% the standard transfection efficiency. As for VLP production the overexpression of NEDD4L, CIT, UGCG and NEDD8 led to an improvement in supernatant VLP concentration of 3.3, 2.9, 2.4 and 1.5-fold respectively. NMT-1 overexpression completely hindered VLP production, and no fluorescence in the supernatant could be detected (**Figure 2A**). Combinations of NEDD4L, CIT and UGCG, the three enzymes which contributed to a higher improvement in VLP specific productivity (**Table 1**), were tested at a concentration of 0.25µg/mL of DNA each, together with 0.5µg/mL of pGag::eGFP. These combinations did not have a significant effect on viability and VCD. Interestingly, all three pairs of overexpressed enzymes improved transfection efficiency. NEDD4L + CIT, CIT + UCGC and NEDD4L + UCGC improved transfection efficiency 18, 16 and 19% respectively, at 72hpt from the studied time course. VLP production was also improved, being NEDD4L + CIT the best combination, improving 2.0 fold the VLP concentration (**Figure 2B**). VLP specific productivity improvement was observed to depend on the final concentration of DNA used, coherently with the transient transfection process, where more copies of the transfected DNA would induce a higher protein expression. The total amount of DNA for transient transfection is restricted by the concentration of complexation agent, PEI, which is reported to be cytotoxic at high concentrations[35]. In order to analyze the budding efficiency, the amount of Gag::eGFP within the cell and in the supernatant was measured by fluorescence intensity. It was observed that 60% of total fluorescence in the Gag::eGFP standard production condition was found within the cell, obtaining only a 40% of extracellular fluorescent signal. The ratio of [intracellular:extracellular] fluorescence was calculated to be 1:0.67 in the control condition. Upon overexpression of NEDD4L, CIT and UGCG, this ratio improved to 1:1.05, 1:1.55 and 1:1.13, increasing the extracellular fluorescence signal to 51, 61 and 53% respectively (**Figure 2D**).

CNP knock-down improves production and budding

CNP protein has been described to prevent the assembly of viral components in the cell membrane or the cytosol, subsequently preventing viral replication of viruses like HIV-1, HIV-2 or Hepatitis B[23,30]. In order to test the effects on Gag VLP production upon the partial inhibition of the membrane-associated protein CNP,

a shRNA was designed to knock-down CNP expression and cloned into the pGag::eGFP, plasmid that could be used to test the effects of the knock-down at a final DNA concentration of 1µg/mL. VCD, viability and transfection efficiency remained unaffected while VLP production significantly ($p<0.0001$) increased 62% or 2.7 fold compared to the control transfected with 1µg/mL of the standard pGag::eGFP (**Figure 2C**). This time, being able to carry out the VLP production improvement at 1µg/mL of Gag::eGFP-coding plasmid, the specific productivity increased 59% compared to the control, reaching up to $(8.5 \pm 0.8) \cdot 10^3$ VLP·cell^{-1}·day^{-1} (**Table 1**). The quantification of fluorescence intensity within the cell and in the supernatant showed a ratio of 1:3.1, reflecting that the 76% of the total Gag::eGFP production was located in the extracellular medium. Remarkably, it was determined that CNP knock-down increased 36% VLP budding compared to the standard Gag::eGFP transfection (**Figure 2D**). The inhibition of viral particle assembly by the interaction between CNP and Gag polyprotein was described to be prevented by a single point mutation altering codon 40 in the MA domain (p17) from Gag changing glutamate to lysine (E40K)[23]. This mutation rendered some viruses resistant to CNP natural antiviral activity through evolution. In order to test if through this aminoacidic change we could obtain similar VLP concentration titers to the ones obtained via CNP knock-down, a plasmid encoding a mutated Gag(E40K)::eGFP was constructed and tested via transient transfection. Compared to the standard transfection of Gag::eGFP, no changes were observed in VCD, viability and transient transfection along the time course. However, VLP concentration was significantly ($p<0.0001$) reduced, obtaining 44% less than the standard production (**Figure 2C**). Interestingly, the budding ratio increased up to 61% (**Figure 2D**), suggesting that the E40K mutation interfered with the proper protein synthesis, folding or interaction with the cell membrane, albeit once assembled, most of the Gag monomers where found forming VLP in the extracellular medium. Nevertheless, the specific productivity of the E40K mutant was reduced down to $(2.0 \pm 0.2) \cdot 10^3$ VLP·cell^{-1}·day^{-1} (**Table 1**), indicating that the CNP shRNA approach was more effective in terms of VLP production. Plasmids encoding both Gag::eGFP and NEDD4L and CIT respectively were constructed and tested at 1µg/mL (**Figure 2C**). Surprisingly, VLP production was significantly reduced (p=0.0012) compared to the control Gag::eGFP. Although NEDD4L and CIT overexpression were confirmed (**Supplementary Figure S1**), the fact of transfecting a larger plasmid at the same total DNA concentration reduced the number of
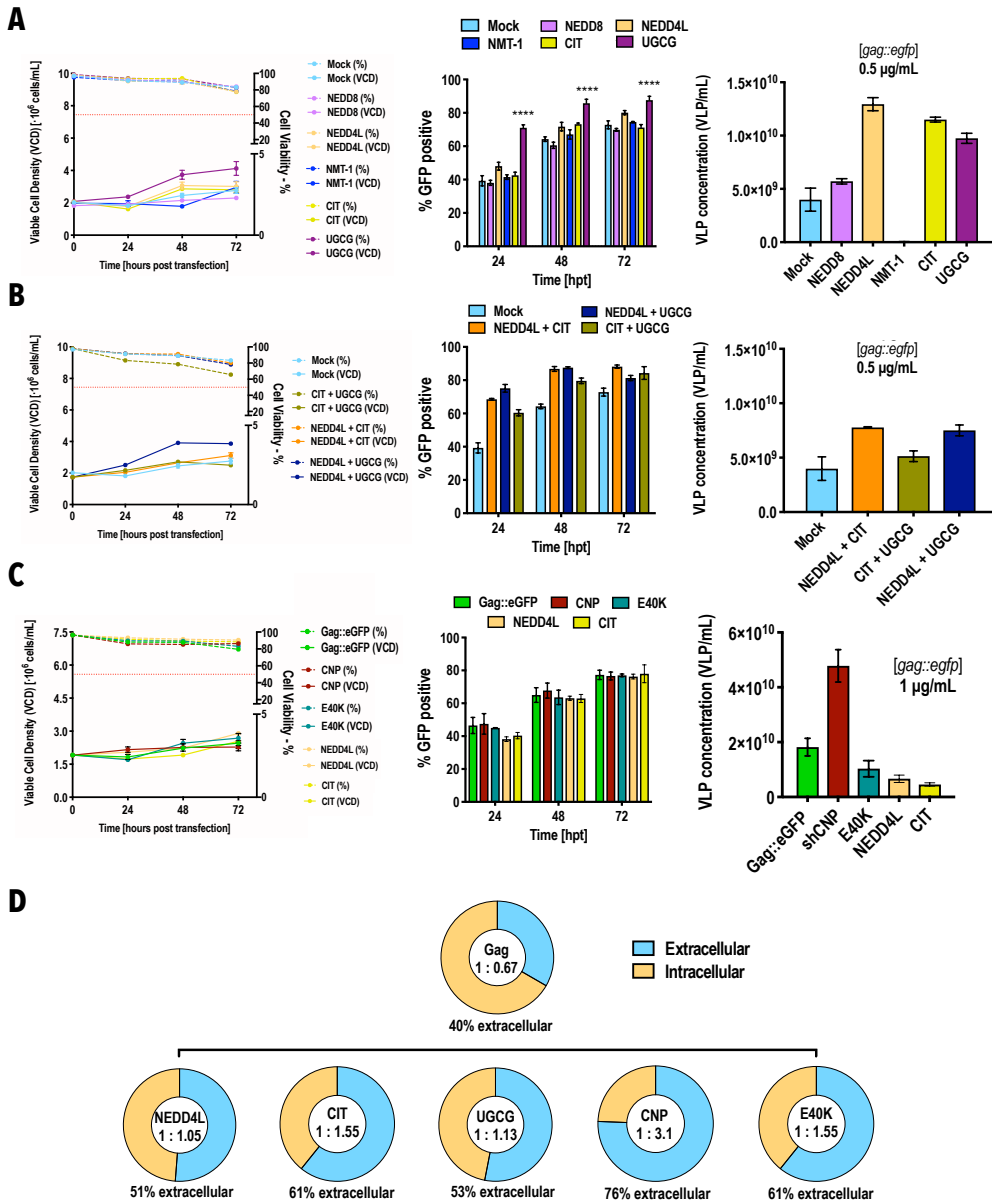
Figure 2. Analysis of HEK293 transfection and VLP production upon the different gene overexpressions. **A)** Cell growth and viability, flow cytometry analysis and VLP concentration plots upon NEDD8, NEDD4L, NMT-1, CIT and UGCG genes overexpression. Mock denotes the condition transfected with 0.5 mg/ml of *gag::egfp* gene and 0.5 mg/ml of an empty plasmid. The conditions denoted as the overexpressed gene are transfected with 0.5 mg/ml of *gag::egfp* gene and 0.5 mg/ml of the gene of interest. Significance was calculated using non-parametric *t-test*. **B)** Cell growth and viability, flow cytometry analysis and VLP concentration plots upon combinations of the enhancer genes overexpression. For the gene combinations, transfections were carried out with 0.5 mg/ml of *gag::egfp* and 0.25 mg/ml of each of the genes of interest in each combination. **C)** Cell growth and viability, flow cytometry analysis and VLP concentration plots upon CNP knock-down using shRNA encoded in the pGag::eGFP. The red dotted line in the cell viability plots represents 50% of cell viability. **D)** Representation of the different intracellular/extracellular ratios and percentage of extracellular fluorescent signal obtained using different budding-enhancing genes.

Table 1. Specific productivities for all studied metabolic modulations

Condition	Specific Productivity (VLP·cell^{-1}·day^{-1}) ·10^3
1 µg/mL of *gag::egfp*	
Gag::eGFP	3.5 ± 0.1
Synthecol	3.4 ± 0.1
Fatty acids	2.3 ± 0.2
Tocopherol	1.7 ± 0.1
Ceramide	3.5 ± 0.2
shRNA CNP	8.5 ± 0.8
E40K	2.0 ± 0.2
CNP:UGCG (2:1) [a]	3.1 ± 0.3
0.5 µg/mL of *gag::egfp*	
Mock	0.8 ± 0.1
NEDD8	1.19 ± 0.07
NEDD4L	1.8 ± 0.1
CIT	1.9 ± 0.1
UGCG	1.46 ± 0.09
NEDD4L+ CIT	1.4 ± 0.4
CIT + UGCG	0.8 ± 0.5
NEDD4L + UGCG	0.9 ± 0.4

[a] 1.05 µg/mL for CNP shRNA-coding plasmid and 0.53 µg/mL for UGCG plasmid.

gag::egfp copies being transfected. The cellular resources diversion due to having one CMV for each gene of the DNA:PEI complex formation for substantially larger plasmids could account for this reduction in VLP production.

CNP shRNA achieves the highest specific productivity and the optimal combination with UGCG overexpression improves specific VLP budding

Therefore, after having screened possible metabolic modulators of the HEK293 cell-based production platform, the knock-down of CNP using shRNA and the overexpression of UGCG were selected for further

optimization combined using a central composite design (CCD) approach. These two modulators were selected as they achieved the highest improvements in VLP production, budding and transfection efficiency (**Figure 3A**). Working ranges for each variable were selected based on the aforementioned experiments and are presented in **Table 2.** The total amount of transfected DNA in each condition was limited to 1.6 µg/mL by the concentration of complexation agent PEI due to its cytotoxic effect. The DNA:PEI ratio was maintained at [1:2] for all the experiments following the standard procedure. A DNA concentration range of [1.2-0.2] µg/mL was selected for the plasmid containing shRNA against CNP as the highest specific productivity was obtained at 1µg/mL (**Figure 2A**). For the plasmid containing the *ugcg* gene, a DNA concentration range of [0.6-0.1] µg/mL was selected since 0.5µg/mL was enough to improve transient transfection (**Figure 2A**). Using these working ranges, a 13-experiment matrix was defined in which the central point was tested five times. Data was fitted to two second-order model non-linear regression looking for possible synergic effects between CNP knock-down the and UGCG overexpression in order to optimize the adjusted specific productivity (P_{sp}) as described in Equation 1 and the specific budding efficiency (B_{sp}) as described in Equation 2. The statistical significance of both models was confirmed by ANOVA analysis, is showed in **Table 2** and **Supplementary Figure S2**. Data from the model was used to construct contour plots and response surface graphs where the interaction between factors can be observed.

Regarding specific productivity, the statistical analyses indicated that only the contribution of CNP shRNA was significant to the model, while UGCG did not contribute to increase specific productivity (**Table 2**). This is observed in the contour plot (**Figure 3B**) and in the response surface graph (**Figure 3C**), where maximum values for P_{sp} were obtained combining the maximum value for CNP shRNA and the minimum for the plasmid encoding UGCG, indicating that for P_{sp} optimization, CNP shRNA single expression is favored. The estimated specific productivity using these parameters was lower than the obtained when using CNP shRNA alone. Therefore, in order to maximize specific productivity, the optimal metabolic modulation for the HEK293 system via transient transfection is to exclusively use the plasmid coding for CNP shRNA. Regarding specific budding efficiency, the statistical analyses indicated that both CNP shRNA and UGCG were significant to the model (**Table 2**). However, only the linear contributions were significant,

indicating a linear relation between the two variables and the B_{sp} response. In the contour plot (**Figure 3D**) and response surface graph (**Figure 3E**) it can be observed that B_{sp} increases with increasing concentrations of both plasmids. The optimal value offered by the model is subsequently the use of maximum concentration of both plasmids. However, this is restricted by the total concentration of PEI. The maximum DNA concentration that allows transient transfection maintaining cell culture viability above 50% is ~1.6μg/mL of total DNA, which corresponds to 1.05μg/mL of CNP shRNA-coding plasmid and 0.53μg/mL of UGCG plasmid. This condition offered a total of 77.3% of extracellular fluorescence signal, a [intracellular:extracellular] ratio of 1:3.4, improving the extracellular Gag::eGFP obtained using 1μg/mL of CNP shRNA-coding plasmid in 1.7%. The ratio [1:2] of UGCG:CNP shRNA plasmids at high total DNA concentrations improved the specific budding efficiency. The positive contribution of both plasmids to the specific budding efficiency can be explained considering both individual effects, since CNP shRNA improves VLP production and budding and UGCG overexpression improves transfection efficiency. In order to characterize the produced VLPs, samples from the studied conditions were analyzed by nanoparticle tracking analysis (NTA) where the size distribution curve and the percentage of fluorescent particles from the total diffracting particles could be assessed (**Figure 4**). All conditions improving VLP production increased the 145 nm of diameter particle subpopulation. CNP knock-down proved to be the best candidate for HEK293-based VLP production optimization, presenting the highest VLP concentration and percentage of fluorescent particles.

Intracellular VLP production is observed upon CIT and NEDD4L overexpression

Budding from the plasma membrane is the main pathway for HIV-1 Gag VLP production[15]. Previous studies have described and quantified this process, henceforth called extracellular budding. Here, Gag molecules oligomerize in specific plasma membrane domains, rich in lipid rafts and tetraspanins[36,37] where they self-assemble and form a protrusion. Upon the ESCRT complexes action, this membrane evagination is excised, forming a VLP[20]. The same series of actions can take place in membranes from intracellular or endosomal compartments, since their membrane share the same properties as the plasma membrane. Intracellular budding takes place when Gag molecules oligomerize and bud off into these intracellular compartments.
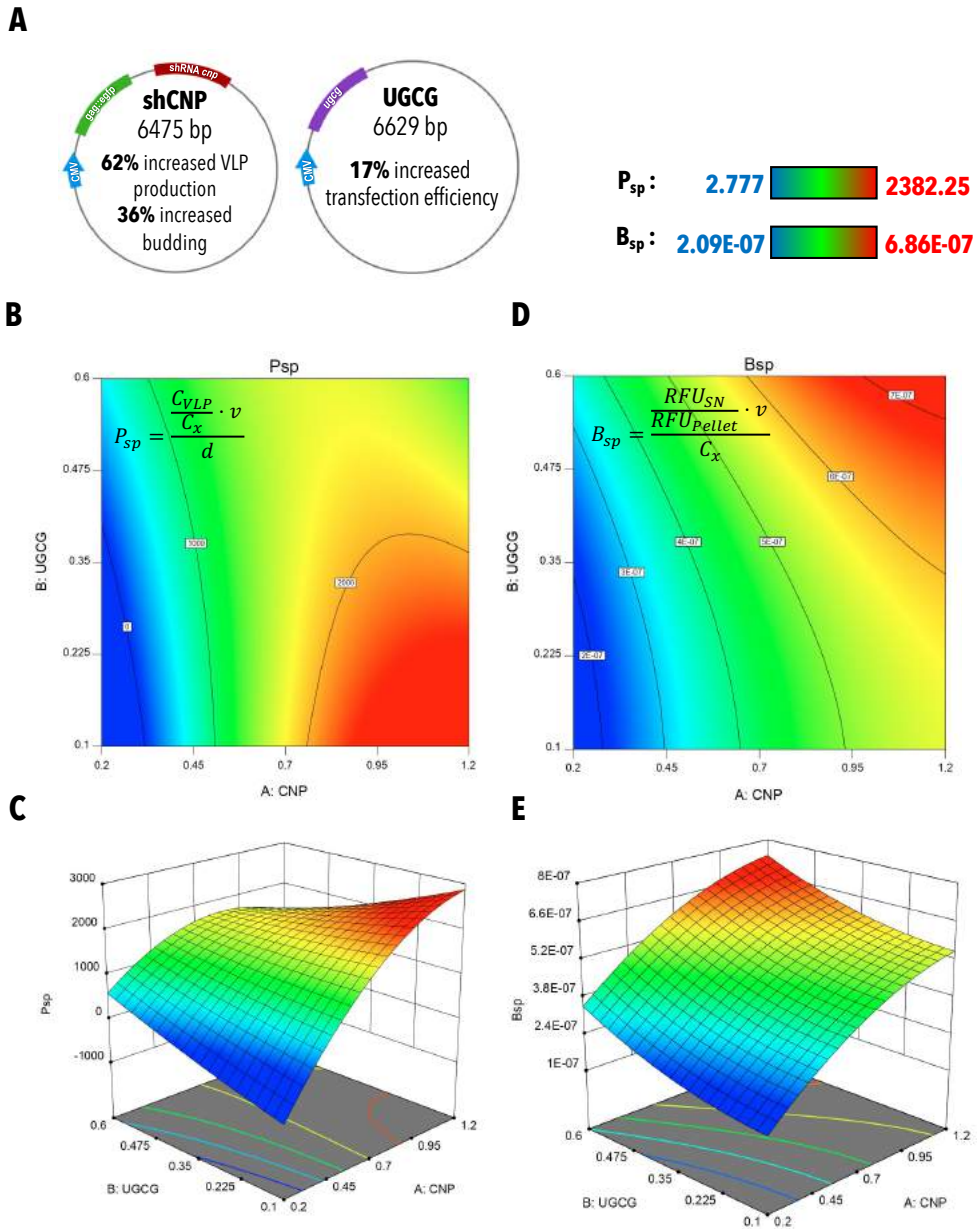
Figure 3. Central composite design (CCD) analysis [...] each individual effect in the HEK293 VLP production system via transient transfection. **B)** Contour profile for the adjusted specific productivity response (P_{sp}) measured in VLP·cell^{-1}·day^{-1} based on CCD experimental results as a function of the concentration of the two analyzed plasmids. **C)** Response surface graphs for P_{sp}. **D)** Contour profile for the adjusted specific budding efficiency (B_{sp}) measured in cell^{-1} based on CCD experimental results as a function of the concentration of the two analyzed plasmids. **E)** Response surface graphs for B_{sp}. $\sqrt{2}$, +1, 0, -1, and -$\sqrt{2}$ correspond to 1.2, 1.05, 0.7, 0.35, 0.2 µg/mL and 0.6, 0.53, 0.35, 0.17, 0.1 µg/mL for the CNP shRNA plasmid and the UGCG plasmid respectively. shCNP: shRNA against CNP. C_{VLP} is the VLP concentration (VLP/mL), C_x is the concentration of viable transfected cells (10^6 cells/mL), v is the cell viability (%), d is the day of the cell culture harvest (d=3), RFU_{SN} is the normalized relative fluorescent units in the supernatant and RFU_{Pellet} is the normalized fluorescent units in the cell pellet.

Table 2. Central composite design, results and ANOVA analyses for optimization of the CNP shRNA and UGCG plasmids ratio

	Coding Levels				
	$\sqrt{2}$	1	0	-1	$-\sqrt{2}$
shCNP(µg/mL)[a]	1.2	1.05	0.7	0.35	0.2
UGCG (µg/mL)	0.6	0.53	0.35	0.17	0.1

Experimental Run		CNP	UGCG	P_{sp}	B_{sp}
1		1	1	1.78E+03	6.86E-07
2		0	$-\sqrt{2}$	2.00E+03	4.75E-07
3		$\sqrt{2}$	0	2.07E+03	6.22E-07
4		0	0	1.92E+03	4.94E-07
5		0	0	1.54E+03	5.03E-07
6		0	0	1.62E+03	4.72E-07
7		1	-1	2.38E+03	4.90E-07
8		$-\sqrt{2}$	0	2.78E+00	2.09E-07
9		0	$\sqrt{2}$	1.62E+03	5.76E-07
10		-1	1	9.01E+02	4.24E-07
11		0	0	1.59E+03	4.18E-07
12		-1	-1	9.02E+01	2.21E-07
13		0	0	1.73E+03	5.03E-07

Model	R^2	Adjusted R^2	F value	p value[b]	Lack of fit[c]
P_{sp}	0.9677	0.9446	41.91	< 0.0001	0.3161
B_{sp}	0.9354	0.8892	20.26	0.0005	0.2149

Model	Parameters	Coefficient	SE	F value	p value
	Constant	1679.3	76.1		
	[CNP]	760.93	60.16	159.96	**< 0.0001**
	[UGCG]	-40.05	60.16	0.4431	0.527
P_{sp}	[CNP]·[UGCG]	-353.38	85.08	17.25	**0.0043**
	[CNP]^2	-355.91	64.52	30.43	**0.0009**
	[UGCG]^2	31.61	64.52	0.24	0.6392
	Constant	4.78E-07	2.01E-08		
	[CNP]	1.39E-07	1.59E-08	76.82	**< 0.0001**
	[UGCG]	6.77E-08	1.59E-08	18.14	**0.0038**
B_{sp}	[CNP]·[UGCG]	-1.75E-09	2.25E-08	0.0061	0.9402
	[CNP]^2	-3.51E-08	1.71E-08	4.23	0.0788
	[UGCG]^2	1.99E-08	1.71E-08	1.37	0.2806

[a] DNA/PEI ratio was always maintained at [1:2].

[b] p values under 0.05 are considered statistically significant with 95% confidence, and p values under 0.1 are considered statistically significant with 90%.

[c] p values associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.
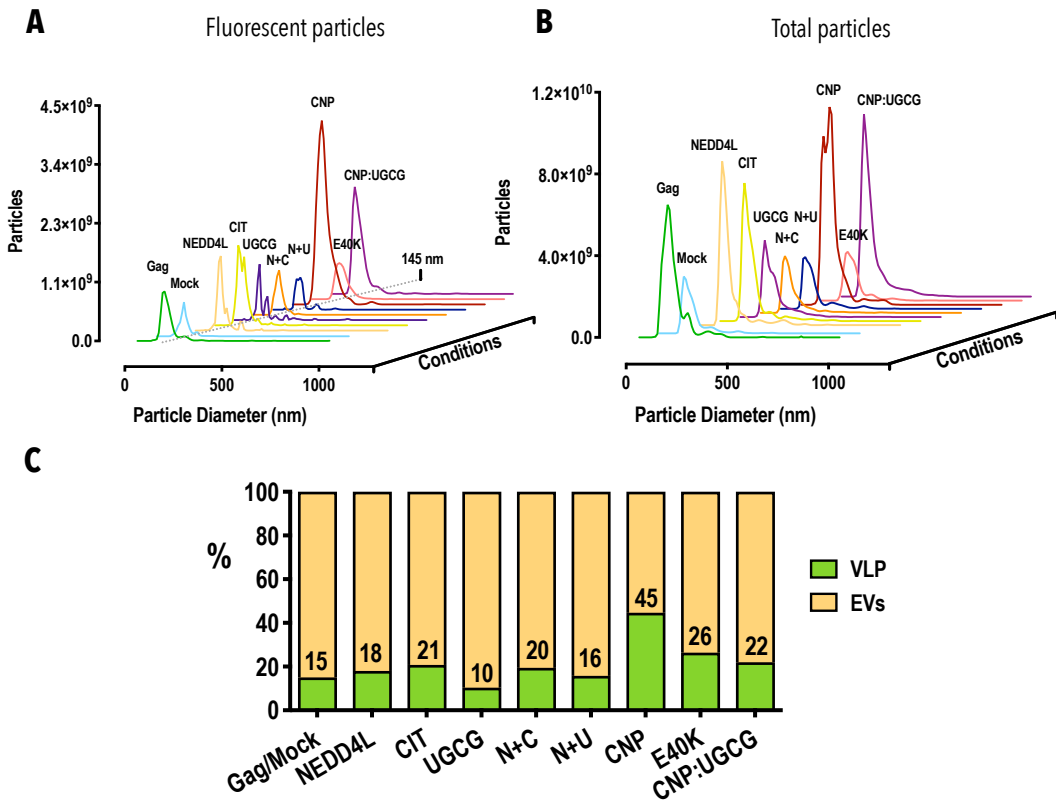
SE: Standard error

Figure 4. Nanoparticle tracking analysis (NTA) from the different studied conditions. **A)** Size distribution curves from fluorescent particles. All conditions show a peak around 145nm. **B)** Size distribution curves from total observed diffracting particles **C)** Bar charts showing the percentage of fluorescent particles in each condition. The CNP:UGCG condition correspond to a CNP:UGCG ration of 2:1 and a DNA concentration of 1.05 μg/mL for CNP shRNA-coding plasmid and 0.53 μg/mL for UGCG plasmid.

Once VLPs are within these endosomal compartments, they are released to the extracellular medium as exosomes (**Figure 5**). Interestingly, the overexpression of CIT and NEDD4L induces the generation of intracellular compartments and increases Gag ubiquitination[21,22], tagging Gag molecules and promoting their transport to the membrane. As a result, the exosomal pathway for VLP production is enhanced[22]. In this work, cells overexpressing CIT and NEDD4L were visualized via transmission electron microscopy (TEM) and intracellular budding was observed (**Figure 5**). Round structures of ~140nm of diameter could be seen at high concentration in the extracellular medium (**Figure 5A**), as well as emerging from the plasma membrane (**Figure 5B,E**). These structures, also localized enclosed in intracellular compartments, presented concentric ring-shaped areas of higher density, indicating the presence of protein following this pattern (**Figure 5C**). This is a characteristic feature of Gag VLPs. Intracellular budding was confirmed,

133
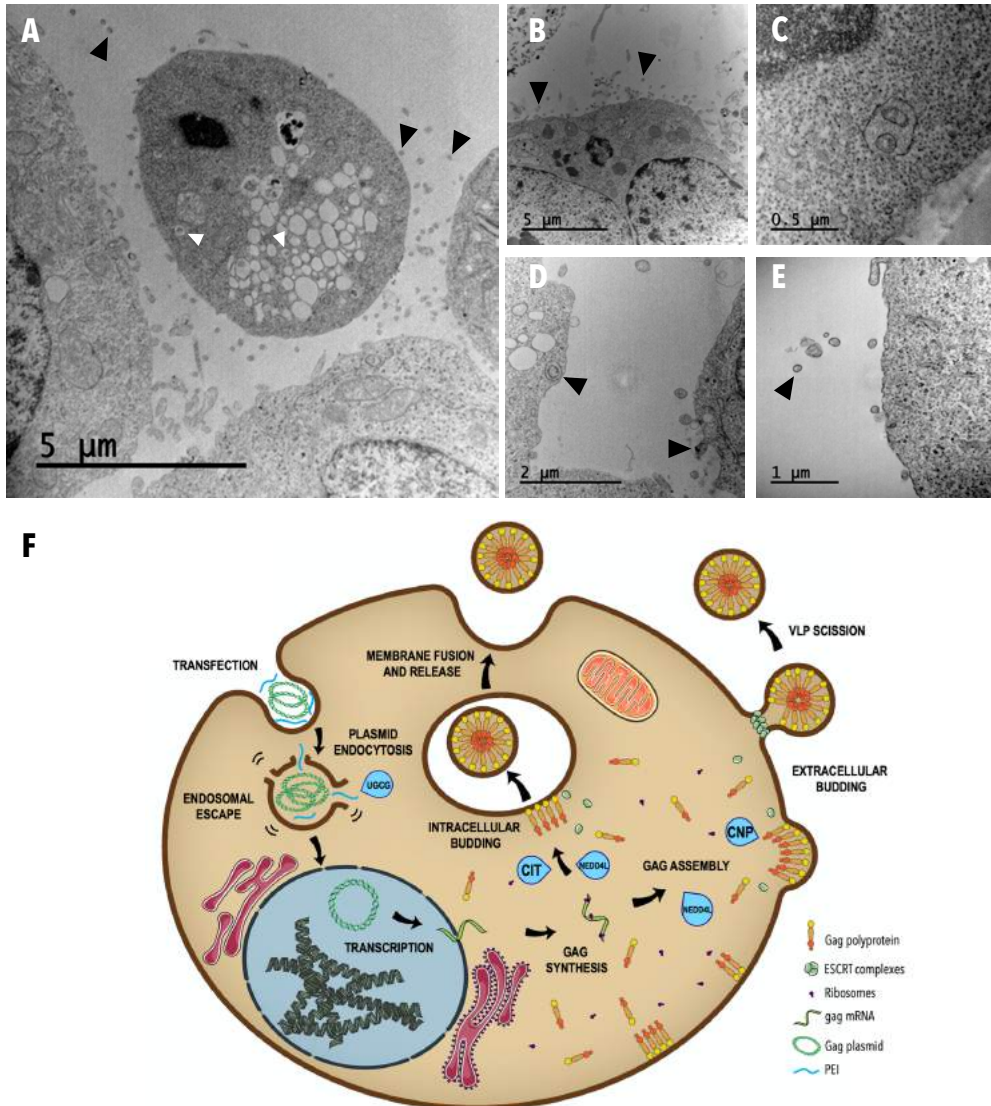
Figure 5. Study of CIT effect on intracellular and extracellular budding via transmission electron microscopy (TEM) visualization of transfected HEK293 cultures using the different budding-enhancing combinations. **A)** HEK293 cells expressing Gag::eGFP and overexpressing CIT. The intracellular vesicle trafficking system or endosomal system and extracellular VLPs can be observed. VLPs found inside endosomal compartments pointed by white arrows. Extracellular VLPs pointed by black arrows. **B)** Same condition overexpressing CIT and NEDD4L where extracellular budding can be observed. VLPs and extracellular budding pointed by black arrows. **C)** Detail of VLPs inside an endosomal compartment (intracellular budding). **D)** Detail of intracellular (left) and extracellular budding (right). **E)** Extracellular VLPs pointed by black arrow. **F)** Schematic representation of the process of VLP production, showing the steps of transient transfection, Gag synthesis, the different intracellular and extracellular pathways for Gag VLP budding and the cellular processes affected by UGCG, NEDD4L and CIT enzyme overexpression and CNP knock-down.

being to the best of our knowledge, the first time that this VLP production pathway is being targeted and

explored to increase VLP concentration as a way to optimize the VLP production bioprocess.

DISCUSSION

The study of metabolic modulations to enhance the VLP production capacity of HEK293 cells upon transient transfection has been approached in this work on the basis of the results obtained in a previous quantitative proteomic study in which VLP transfection and production bottlenecks were identified[14]. The observed downregulation at 72hpt in lipid biosynthesis was tackled by adding lipids that were described to benefit HEK293 cells[26] such as synthecol, fatty acids, tocopherol and ceramide. None of the lipid additions improved VLP production nor specific productivity (**Table 1**). One possible explanation to this result is that the VLP production process is coupled with cell growth. Since there is no distinction between growth and production phase, cellular resources and energy are being diverted to promote growth, like observed upon tocopherol addition (**Figure 1A**) or to compensate the overall homeostasis disruption caused by the transfection event[14]. Thus, reducing VLP production or specific productivity. The next targeted bottleneck was the transport of Gag molecules to the specific budding sites in the cell membrane. This is carried out by PTMs like N-myristoylation or monoubiquitination. Contrarily to what was expected, overexpressing NMT-1 enzyme showed to completely block VLP production (**Figure 2A**). When Gag polyprotein MA (p17) domain is myristoylated, it interacts with membrane sites enriched in phosphatidylinositol-(4,5)-bisphosphate[38]. This lipid act as a membrane link and trigger a conformational change in the myristoyl molecule known as the myristoyl switch[39,40]. Myristoyl molecules change from a sequestered conformation in a monomeric state to a exposed conformation in a multimeric state, anchoring to membrane domains enriched in lipid rafts[41]. Myristoylated MA domains are found in the cell in an equilibrium between sequestered monomeric state and exposed trimeric species. Our findings suggest that overexpressing N-myristoylation might cause a disequilibrium in the myristoyl switch preventing Gag binding and assembly to the plasma membrane or inducing N-myristoyl PTMs in different molecules that block the phosphatidylinositol binding site, inhibiting the release of VLPs. Conversely, the overexpression of accessory proteins like NEDD8, NEDD4L and CIT led to an increase in VLP production (**Figure 2B**). These proteins contribute to Gag ubiquitination, which is essential for ESCRT-mediated VLP production[42]. CIT colocalizes and interact with Hrs, a protein that belongs to ESCRT-0 complex and NEDD8

and NEDD4L have been described as TSG-101-mediated ESCRT-I trans-activators[17,43]. The observed increased in VLP production and budding, in the plasma membrane or intracellular compartments, is produced by enhancing the ESCRT pathway. Exploring the possibility of knocking-down ESCRT-0 Hrs component in future works might improve further VLP production as it is described that Hrs inhibition strategies led to an increase of CIT effects on Gag transport[43]. Another identified metabolic bottleneck likely to be modulated was the glycosphingolipid biosynthetic pathway. Apart from the previously tested addition of ceramide, UGCG enzyme was overexpressed. It improved not only VLP titers, but interestingly, also transfection efficiency (**Figure 2A**). UGCG is the only enzyme that produces *de novo* glucosylceramide, the precursor of all glycosphingolipids. Curiously, UGCG has been described as essential for early endosomal trafficking and entry of different viruses, such as severe fever with thrombocytopenia syndrome virus (SFTSV), Influenza, Ebola and measles virus among others[44,45]. UGCG is reported to act on the late stages of entry of these viruses, when they require endosomal acidification to release the virus into the cytoplasm. UGCG knock-down has been shown to lead to inhibition of virus entry and accumulation in endosomal compartments[44]. The same effect have been reported for Influenza infection[45]. This process is highly similar to the endosomal escape that takes place upon transient transfection. DNA:PEI complexes enter the cell via macropinocytosis and endocytosis and then are released upon acidification of the endosomal compartments[46,47]. The effect of UGCG overexpression might be leading to a more efficient endosomal escape, improving transient transfection. UGCG also alters glycosphingolipid-enriched microdomains in the cell membrane[48], also affecting Gag VLP assembly. In order to fully exploit the effect of UGCG overexpression, it could be combined with different DNA:PEI complex formation time and ratios to explore the possibilities of a modified plasma membrane and endocytic scenario. A different approach undertaken in this work was to knock-down CNP, a protein that binds to Gag monomers and inhibit assembly. This was the most efficient modulation, improving VLP production, specific productivity and extracellular budding. There are other proteins that share this antiviral property of binding Gag molecules like tetherin or MOV10[49,50]. Testing and analyzing the knock-down of these proteins might contribute to further improvement in VLP production. More interestingly, modulations of human CNP has been demonstrated to work in many species of lentivirus including HIV-2, SIV$_{MAC}$239, SIV$_{AGM}$SAB[23]. This makes

CNP knock-down a promising modulation for the further establishment of a large-scale lentiviral production platform. In this work, all these metabolic engineering modulations have been tested in a transient transfection HEK293 model, limiting the results as one of the main components of the transfection was the Gag::eGFP plasmids. However, once screened for the most efficient modifications to improve VLP production and budding, they could be applied to stable cell lines to further improve HEK293 performance and develop a robust platform for the mass production of future vaccines.

CONCLUSIONS

After having screened for several metabolic engineering modulations of HEK293 cells upon TGE for Gag VLP production, some proteins were found to improve VLP titers upon modification. Overexpression of NEDD4L, CIT and UGCG improved VLP production 3.3, 2.9, and 2.4-fold. At the same time, budding percentage increased from 40% in a standard production of Gag::eGFP via TGE up to 51, 61 and 53% respectively. CNP knock-down increased VLP production 2.7-fold and budding percentage up to 76%. DoE analyses demonstrated that only knocking-down CNP was enough to obtain the optimal specific productivity among the tested conditions. Combining CNP and UGCG at a ratio of [2:1] at 1.6 µg/mL of total DNA improved budding percentage up to 77.3%. The best percentage of fluorescent particles was achieved only knocking-down CNP, reaching 45% of VLPs with respect of the total number of diffracting particles produced. The further testing and analysis of the HEK293 platform could reveal more metabolic points likely to be modified to improve VLP production. This represent a step forward into the development of a robust platform for VLP-based vaccines.

SUPPORTING INFORMATION
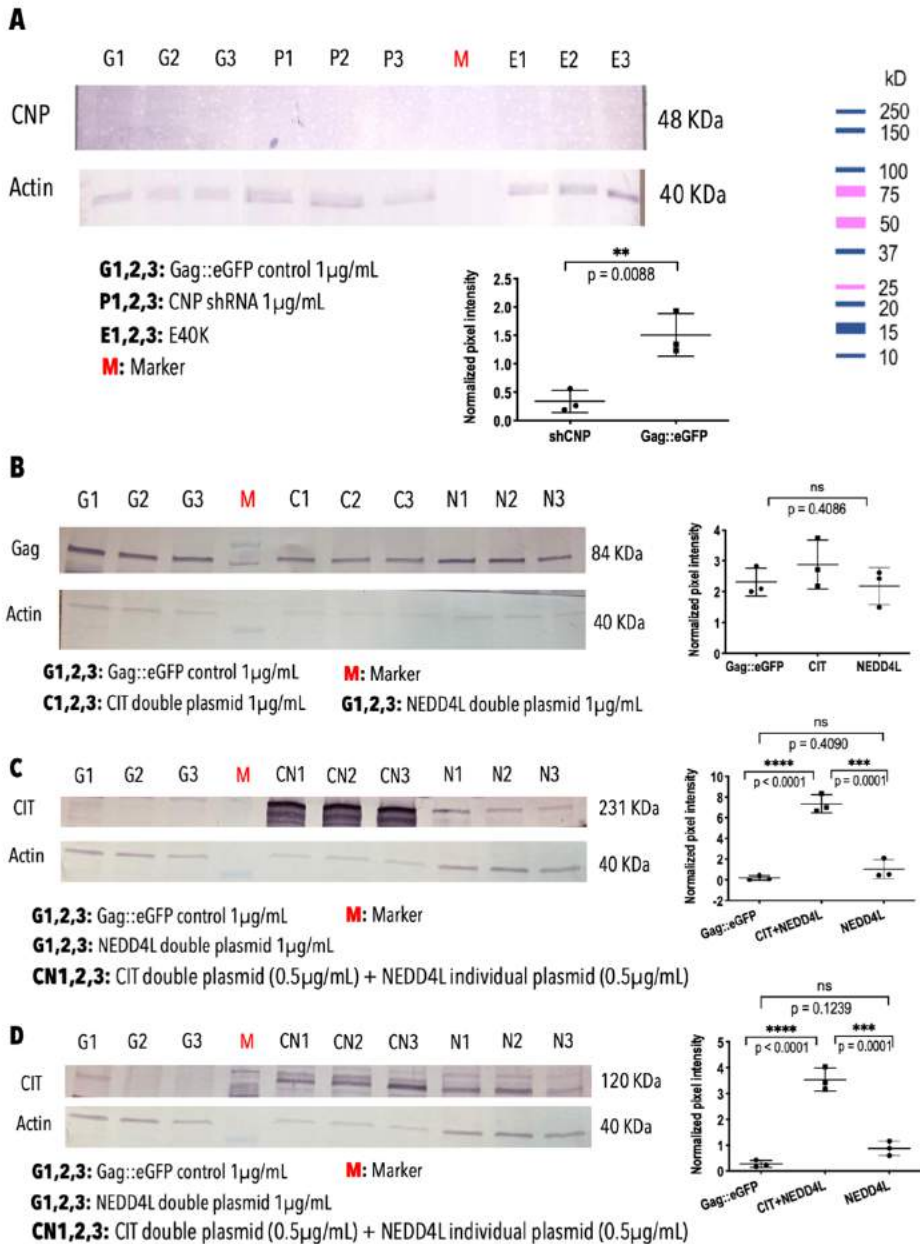
Supplementary Figure S1: Western blot analyses. **A)** WB for CNP protein **B)** WB for Gag polyprotein **C)** WB for CIT protein **D)** WB for NEDD4L protein. All pixel intensities were normalized to the corresponding actin pixel intensity value. Significances were calculates using one-way ANOVA.

A

Predicted vs. Actual

B
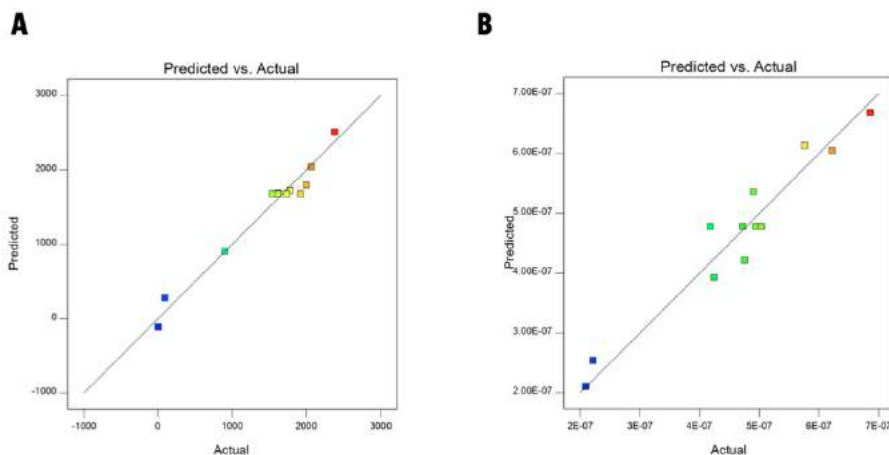
Predicted vs. Actual

Supplemetary Figure S2: Predicted versus experimental (actual) data for the different CCD generated models. A) Predicted vs actual plot for the model calculating the P_{sp} response. B) Predicted vs actual plot for the model calculating the B_{sp} response. The straight line is driven by the equation $y = x$.

Supplementary Table S1. Primers used in the generation of plasmids

Code	Name	Sequence 5' - 3'
1	CIT_F	AGAAGCGCGCGTAAAACGACGGCCAG
2	CIT_R	AGAAGCGGCCGCCAGGAAACAGCTATGACC
3	NMT-1_F	AGAAGAATTCAGAAGGAGATATACATATGGCGGAC
4	NMT-1_R	AAAAGCGGCCGCGCTGGGATGGCGCTAAT
5	CNP ssDNA ultramer	GTGGAAAGGACGAAACACCGTGTTCTCACCACCACTTATGCTCGAGCATAAGTGGTGGTGAGAACACTT TTTCATGGAGTTCCGCGTTACATA
6	E40K_F	TAGCTTCCTGCTTGCCCATAGC
7	E40K_R	GCAGGAAGCTAGAACGATTCGCA

REFERENCES

1. Lua, L. H. *et al.* Bioengineering virus-like particles as vaccines. *Biotechnol Bioeng* **111**, 425–440 (2014).

2. Cervera, L. *et al.* Production of HIV-1-based virus-like particles for vaccination: achievements and limits. *Appl. Microbiol. Biotechnol.* **103**, 7367–7384 (2019).

3. Kuate, S. *et al.* Immunogenicity and efficacy of immunodeficiency virus-like particles pseudotyped with the G protein of vesicular stomatitis virus. *Virology* **351**, 133–144 (2006).

4. Kueng, H. J. *et al.* General strategy for decoration of enveloped viruses with functionally active lipid-modified cytokines. *J. Virol.* **81**, 8666–76 (2007).

5. Venereo-Sanchez, A. *et al.* Hemagglutinin and neuraminidase containing virus-like particles produced in HEK-293 suspension culture: An effective influenza vaccine candidate. *Vaccine* **34**, 3371–3380 (2016).

6. Dyson, M. R. Fundamentals of Expression in Mammalian Cells. *Adv Exp Med Biol* **896**, 217–224 (2016).

7. von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

8. Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

9. Henne, W. M., Buchkovich, N. J. & Emr, S. D. The ESCRT pathway. *Dev Cell* **21**, 77–91 (2011).

10. Hurley, J. H. ESCRT complexes and the biogenesis of multivesicular bodies. *Curr Opin Cell Biol* **20**, 4–11 (2008).

11. Fujii, K., Hurley, J. H. & Freed, E. O. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nat Rev Microbiol* **5**, 912–916 (2007).

12. Alonso, Y. A. M., Migliano, S. M. & Teis, D. ESCRT-III and Vps4: a dynamic multipurpose tool for membrane budding and scission. *Febs j* **283**, 3288–3302 (2016).

13. Edwards, A. & Haas, W. Multiplexed quantitative proteomics for high-throughput comprehensive proteome comparisons of human cell lines. in *Methods in Molecular Biology* **1394**, 1–13

(Humana Press Inc., 2016).

14. Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed Quantitative Proteomic Analysis of HEK293 Provides Insights into Molecular Changes Associated with the Cell Density Effect, Transient Transfection, and Virus-Like Particle Production. *J. Proteome Res.* **19**, 1085–1099 (2020).

15. Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

16. Pincetic, A. & Leis, J. The Mechanism of Budding of Retroviruses From Cell Membranes. *Adv Virol* **2009**, 6239691–6239699 (2009).

17. Chung, H. Y. *et al.* NEDD4L overexpression rescues the release and infectivity of human immunodeficiency virus type 1 constructs lacking PTAP and YPXL late domains. *J Virol* **82**, 4884–4897 (2008).

18. Shirk, A. J., Anderson, S. K., Hashemi, S. H., Chance, P. F. & Bennett, C. L. SIMPLE interacts with NEDD4 and TSG101: Evidence for a role in lysosomal sorting and implications for Charcot-Marie-Tooth disease. *J. Neurosci. Res.* **82**, 43–50 (2005).

19. Usami, Y., Popov, S., Popova, E. & Gottlinger, H. G. Efficient and specific rescue of human immunodeficiency virus type 1 budding defects by a Nedd4-like ubiquitin ligase. *J Virol* **82**, 4898–4907 (2008).

20. Vietri, M., Radulovic, M. & Stenmark, H. The many functions of ESCRTs. *Nat. Rev. Mol. Cell Biol.* **21**, 25–42 (2020).

21. Ding, J., Zhao, J., Sun, L., Mi, Z. & Cen, S. Citron kinase enhances ubiquitination of HIV-1 Gag protein and intracellular HIV-1 budding. *Arch Virol* **161**, 2441–2448 (2016).

22. Loomis, R. J. *et al.* Citron Kinase, a RhoA Effector, Enhances HIV-1 Virion Production by Modulating Exocytosis. *Traffic* **7**, 1643–1653 (2006).

23. Wilson, S. J. *et al.* Inhibition of HIV-1 particle assembly by 2′,3′-cyclic- nucleotide 3′-Phosphodiesterase. *Cell Host Microbe* **12**, 585–597 (2012).

24. Fuenmayor, J., Cervera, L., Rigau, C. & Gòdia, F. Enhancement of HIV-1 VLP production using gene inhibition strategies. *Appl. Microbiol. Biotechnol.* **102**, 4477–4487 (2018).

25. Reed, M. *et al.* Chimeric Human Immunodeficiency Virus Type 1 Containing Murine Leukemia Virus Matrix Assembles in Murine Cells. *J. Virol.* **76**, 436–443 (2002).

26. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

27. Cervera, L. *et al.* Selection and optimization of transfection enhancer additives for increased virus-like particle production in HEK293 suspension cell cultures. *Appl Microbiol Biotechnol* **99**, 9935–9949 (2015).

28. Trajkovic, K. *et al.* Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes. *Science (80-.).* **319**, 1244–1247 (2008).

29. Colombo, M., Raposo, G. & Théry, C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol.* **30**, 255–289 (2014).

30. Ma, H. *et al.* 2′,3′-Cyclic Nucleotide 3′-Phosphodiesterases Inhibit Hepatitis B Virus Replication. *PLoS One* **8**, e80769 (2013).

31. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

32. Zheng, L., Baumann, U. & Reymond, J. L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, (2004).

33. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

34. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–5 (2012).

35. Carpentier, E., Paris, S., Kamen, A. A. & Durocher, Y. Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *J*

Biotechnol **128**, 268–280 (2007).

36.   Hogue, I. B., Grover, J. R., Soheilian, F., Nagashima, K. & Ono, A. Gag Induces the Coalescence of Clustered Lipid Rafts and Tetraspanin-Enriched Microdomains at HIV-1 Assembly Sites on the Plasma Membrane. *J. Virol.* **85**, 9749–9766 (2011).

37.   Jolly, C. & Sattentau, Q. J. Human Immunodeficiency Virus Type 1 Assembly, Budding, and Cell-Cell Spread in T Cells Take Place in Tetraspanin-Enriched Plasma Membrane Domains. *J. Virol.* **81**, 7873–7884 (2007).

38.   Alfadhli, A., Barklis, R. L. & Barklis, E. HIV-1 matrix organizes as a hexamer of trimers on membranes containing phosphatidylinositol-(4,5)-bisphosphate. *Virology* **387**, 466–472 (2009).

39.   Resh, M. D. A myristoyl switch regulates membrane binding of HIV-1 Gag. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 417–418 (2004).

40.   Tang, C. *et al.* Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 517–22 (2004).

41.   Ghanam, R. H., Samal, A. B., Fernandez, T. F. & Saad, J. S. Role of the HIV-1 matrix protein in Gag intracellular trafficking and targeting to the plasma membrane for virus assembly. *Front. Microbiol.* **3**, (2012).

42.   Sette, P., Nagashima, K., Piper, R. C. & Bouamr, F. Ubiquitin conjugation to Gag is essential for ESCRT-mediated HIV-1 budding. *Retrovirology* **10**, (2013).

43.   Ding, J., Su, L. & Gao, G. Hrs inhibits citron kinase-mediated HIV-1 budding via its FYVE domain. *Protein Cell* **2**, 470–476 (2011).

44.   Drake, M. J. *et al.* A role for glycolipid biosynthesis in severe fever with thrombocytopenia syndrome virus entry. *PLoS Pathog.* **13**, e1006316 (2017).

45.   Drews, K. *et al.* Glucosylceramide synthase maintains influenza virus entry and infection. *PLoS One* **15**, e0228735 (2020).

46.   González-Domínguez, I., Grimaldi, N., Cervera, L., Ventosa, N. & Gòdia, F. Impact of physicochemical properties of DNA/PEI complexes on transient transfection of mammalian cells.

N. Biotechnol. **49**, 88–97 (2019).

47.   Benjaminsen, R. V., Mattebjerg, M. A., Henriksen, J. R., Moghimi, S. M. & Andresen, T. L. The possible "proton sponge " effect of polyethylenimine (PEI) does not include change in lysosomal pH. *Mol. Ther.* **21**, 149–157 (2013).

48.   Schömel, N. *et al.* UGCG overexpression leads to increased glycolysis and increased oxidative phosphorylation of breast cancer cells. *Sci. Rep.* **10**, 8182 (2020).

49.   Liberatore, R. A., Mastrocola, E. J., Powell, C. & Bieniasz, P. D. Tetherin Inhibits Cell-Free Virus Dissemination and Retards Murine Leukemia Virus Pathogenesis. *J. Virol.* **91**, (2017).

50.   Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).

CHAPTER THREE

Part I.

Culture media screening and selection of optimal transfection conditions

ABSTRACT

The first step in bioprocess development is choosing the culture media which provides the best results for the process that is going to be implemented. For the development of a transient transfection-based bioprocess, a culture media sustaining both cell growth and transfection is required. Protein-free, chemically defined culture media, suitable for GMP-compliant bioprocess development and causing no interference with the transfection step have been developed to offer a unique element for the upstream bioprocess design. This eliminate steps of complete media exchange separating phases of cell growth and production. Therefore, more versatile bioprocess approaches can be conceived. In order to select the most convenient culture media for a perfusion-based transient transfection bioprocess, five different protein-free, chemically defined culture media were tested for transient transfection performance: FreeStyle® 293, FreeStyle® F17, HyCell, HEKTF and BalanCD®HEK. Together with the media selection, elements concerning transfection were also evaluated: the DNA/PEI complexation medium, the complexation agent (PEIpro®, in-house prepared PEI 25kDa solution and FECTOpro®) and the cell density at the moment of transfection. The optimal combination of these elements will be implemented for the further intensification and optimization of a VLP production bioprocess.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from McGill University (Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). The different tested cell culture media were: HyCell™ TransFx-H media from HyClone™ (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco, Life Technologies). FreeStyle™ 293 Expression Medium (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) supplemented with 0.1% Pluronic™ F-68 non-ionic surfactant, transferrin and insulin (Merck Millipore, Kankakee, IL, USA), a lipid mix composed of synthetic cholesterol (100x Synthecol®, Sigma) tocopherol acetate (0.2 mg/mL T1157, Sigma), fatty acids (100x F7050, Sigma) and glucose (G8270, Sigma) at a concentration previously reported[1]. FreeStyle™ F17 Expression Medium (Gibco, Life Technologies) supplemented with 8 mM GlutaMAX™ (Gibco, Life Technologies), 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, Life Technologies) and IGF-1 at a final concentration of 50μg/L. HEK TF (Xell AG, Bielefeld, Germany) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies), 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, Life Technologies). BalanCD® HEK293 (Irvine Scientific, Santa Ana, CA, USA) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies), 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, Life Technologies).

TRANSIENT TRANSFECTION

Transfection was carried out at different cell densities and using different components for DNA/PEI complex formation. Standard transfection was established at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 μg/mL, according to previous work[2]. DNA/PEI complexes were formed by adding PEI to plasmid DNA diluted in complexation medium (10% of the total culture volume to be transfected). Media for complexation tested were HyCell, FreeStlye F17, FreeStyle 293, 150 mM NaCl solution and 125 mM

NaCl solution. Three different transfection reagents were tested: PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France), PEI 25 kDa (408727, Sigma-Aldrich) and FECTOpro® (Polyplus). The plasmid used contained a gene coding for HIV-Gag protein fused to eGFP (Gag::eGFP). Briefly, pGag::eGFP plasmid is diluted in the complexation medium and vortexed for 10 seconds. Then, the transfection reagent is added in 1:2 (w/w) ratio and vortexed three times, then the mixture is incubated for 15 min at room temperature and then added to the cell culture. This standard protocol varied for the condition of 125 mM NaCl which was incubated 10 min. This standard procedure was also applied for transfection at high cell densities of $3 \cdot 10^6$, $4 \cdot 10^6$ and $5 \cdot 10^6$ cells/mL at the moment of transfection.

FLOW CITOMETRY

Samples were taken every 24h after transfection and cells were fixed using formaldehyde 2% for 10 minutes, centrifuged and then resuspended in PBS for FACS analysis. The percentage of GFP positive cells was assessed using a BD FACS Canto flow cytometer (BD Biosciences, San Jose, CA, USA). Laser 488 was used for GFP measurement. The results were analyzed with FACS DIVA software (BD Biosciences).

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using a validated quantification assay[3]. Supernatants containing VLPs were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples.

BIOREACTOR CULTURE CONDITIONS AND SETUP DESCRIPTION

A BioStat B Plus bioreactor (Sartorius Stedim Biotech, Goettingen, Germany) equipped with a 3-blade segment dual impeller with UP-DP configuration[4] was used for HEK293 cell cultivation. The agitation was set a 200 rpm. Temperature was set a 37 °C. The pH was set at 7.1 and controlled with CO_2 and $NaHCO_3$ (7.5%w/v). The dissolved oxygen was controlled at 40% of air saturation by sparging air at a constant flow of 0.1L/min and additional pure oxygen when needed. HEK293 growing exponentially in disposable

polycarbonate 1 L shake flasks (Corning®) were used to seed the bioreactor at $0.5 \cdot 10^6$ cells/mL in 1 L. Samples were taken every 24 h for cell counting and viability determination. After transfection, assessment of the percentage of GFP positive cells and VLP quantification was also performed every 24h.

RESULTS AND DISCUSSION

MEDIA SCREENING

HEK293 cells were grown in a total of five different culture media, transiently transfected and their performance was subsequently compared. Regarding cell growth, supplemented FreeStyle® 293 (FreeMix, n=6), HyCell (n=21), FreeStyle® F17 (F17, n=21) and BalanCD®HEK (CDHEK, n=12) reported a similar trend, reaching $2\text{-}4 \cdot 10^6$ cells/mL at 72hpt. HEKTF (n=6) media allowed cells to reach up to $8 \cdot 10^6$ cells/mL at 72hpt (**Figure 1A**). Viability was maintained above 80% in all cases. Regarding VLP production inferred by supernatant fluorescence intensity, at 24hpt all culture media, except HEKTF, showed no significant differences in relative fluorescence units (RFU). HEKTF showed no VLP production. At 48hpt, HyCell and F17 media significantly stood out from the rest reaching ~50 RFUs. The same trend was observed at 72 hpt, when HyCell and F17 media significantly showed higher VLP production compared to the rest (**Figure 1B**). HEKTF showed no VLP production at any time. Although F17 media showed high levels of fluorescence, readings showed high variability among all replicates (n=21) with a standard deviation of 24.9 RFU compared to the 4.8 RFU of standard deviation of HyCell. Regarding transfection efficiency, at 24hpt all culture media, except HEKTF, reported ~50% of GFP positive cells. HEKTF showed ~10%. The significantly low percentage of transfection can explain the higher cell densities reached using this culture media as culture transfection reduces the cell growth rate[5]. At 48 hpt, FreeMix and F17 media showed significantly higher percentage of GFP positive cells, reaching ~80%. At 72 hpt, reaching ~85%, the same trend is observed, with FreeMix and F17 presenting the highest percentage of transfected cells (**Figure 1C**). Since HyCell and F17 reached the highest production values, these two culture media were selected for further testing, optimizing the different elements for DNA/PEI complex formation and cell density at the time of transfection.
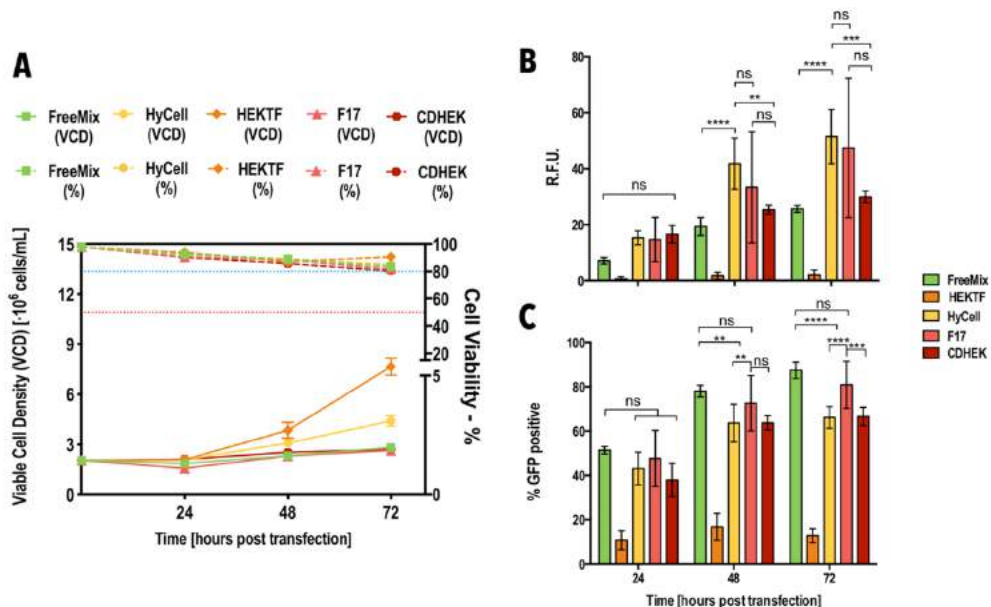
Figure 1. Culture media screening. **A)** Viable cell density (VCD) and cell viability (%) monitoring along the transient transfection process. Culture media analyzed are supplemented as described in Materials and Methods section: FreeStyle® 293 (FreeMix), HyCell, HEKTF, FreeStyle® F17 (F17) and BalanCD® HEK (CDHEK). Blue and red dotted line represent 80% and 50% cell viability respectively. **B)** Relative fluorescent units (RFU) measured at each analyzed time point in each of the different tested culture media. **C)** Transfection efficiency measured in percentage of GFP positive cells (%). Significance calculated using one-way ANOVA.

HYCELL CULTURE MEDIA OPTIMIZATION

Several alternatives for DNA/PEI complex formation and complexation agent were tested for transient transfection in HyCell culture media. The standard protocol for HEK293 transient transfection was developed using FreeMix, so it was also included in the tested conditions to compare the obtained yields. Three different complexation reagents were used: commercial PEIpro®, in-house prepared PEI solution (PEI 25 kDa) and commercial FECTOpro®. The different combinations of tested conditions are described in **Table 1**. All conditions showed a similar trend regarding cell growth. FECTOpro condition reached slightly lower cell densities ($\sim 4\cdot10^6$ cells/mL) and PEI25 slightly higher ($\sim 6\cdot10^6$ cells/mL), with viabilities always above 80% (**Figure 2A**). Regarding VLP production, at 24hpt only PEI25 and FECTOpro conditions showed significantly lower RFUs levels compared to PEIpro. Regardless of the complexation medium used (FreeStyle, HyCell or 120 mM NaCl solution), fluorescence intensity obtained at 24hpt was the same
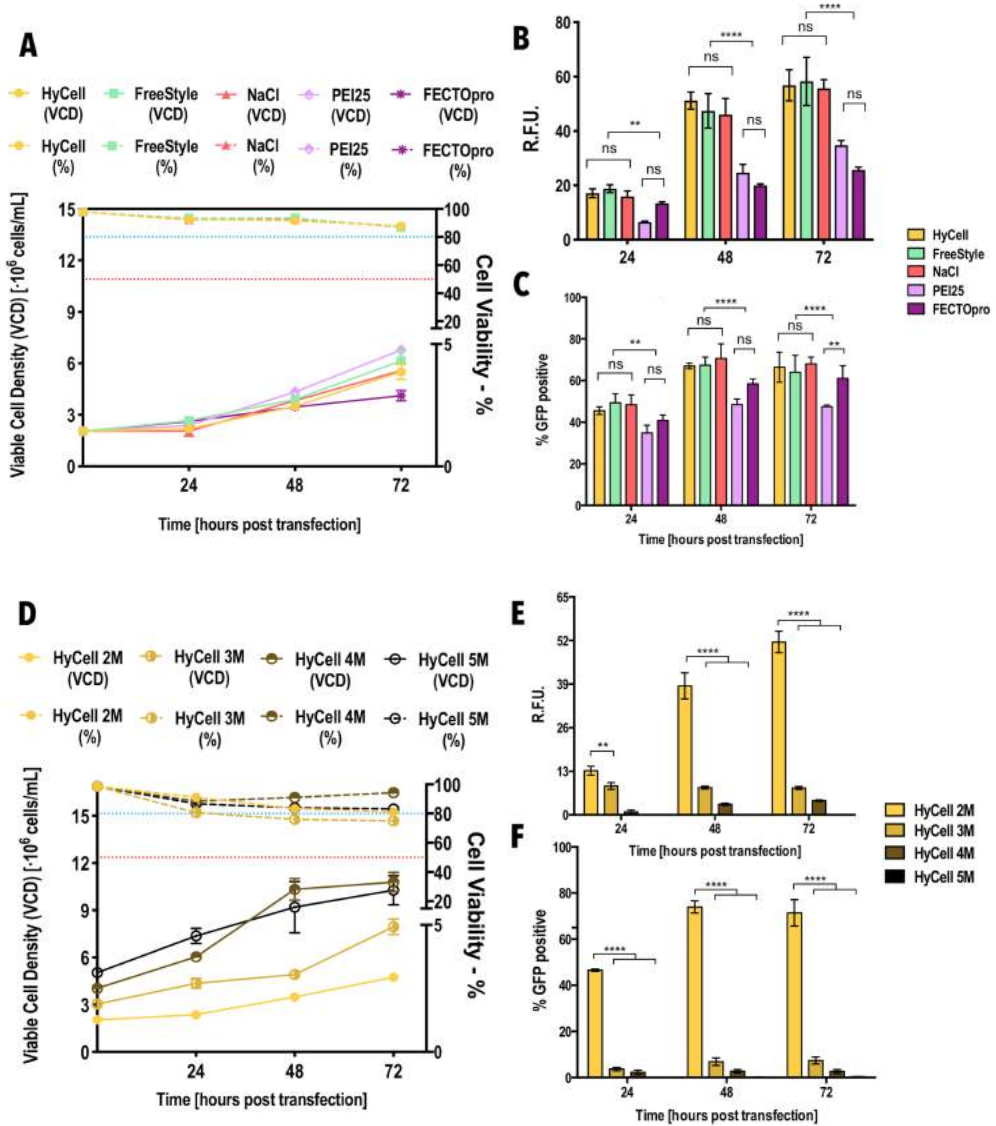
Figure 2. HyCell optimization. **A)** Viable cell density (VCD) and cell viability (%) monitoring along the transient transfection process using the different combinations for complex formation. FreeStyle® 293 culture media (FreeStyle), HyCell and 150 mM NaCl solution as complexation medium. PEIpro®, PEI 25 and FECTOpro® as complexation reagents. Blue and red dotted line represent 80% and 50% cell viability respectively. **B)** Relative fluorescent units (RFU) measured at different time point for each of the different conditions. **C)** Transfection efficiency measured in percentage of GFP positive cells (%). Significance calculated using one-way ANOVA. **D)** VCD and cell viability (%) monitoring along the transient transfection process starting a different cell densities: $2 \cdot 10^6$, $3 \cdot 10^6$, $4 \cdot 10^6$ and $5 \cdot 10^6$ cells/mL. **E)** RFUs measurements for these conditions **F)** Transfection efficiency measurements for these conditions.

Table I. Combinations of the different elements for DNA/PEI complex formation process in HyCell.

Complexation Medium	Complexation Agent	Condition	
HyCell	PEI-pro (Polyplus)	Hycell:	HyCell
FreeStyle 293	PEI-pro (Polyplus)	FreeStyle:	FreeStyle
NaCl 150 mM	PEI-pro (Polyplus)	NaCl:	NaCl
FreeStyle 293	PEI 25 kDa (Sigma-Aldrich)	PEI25:	PEI25
FreeStyle 293	FECTOpro (Polyplus)	FECTOpro:	FECTOpro

(18 RFUs) using PEIpro. The same pattern is observed at 48 and 72 hpt, with these three conditions reaching ~65 RFUs and PEI25 and FECTOpro conditions reaching significantly fewer, around 30 RFUs (**Figure 2B**). As for transfection efficiency, the same patter was also preserved. The conditions using HyCell, Freemix and 150 mM NaCl solution as complexation medium and PEIpro® as complexation reagent reached significantly higher transfection efficiency values compared to conditions using PEI25 and FECTOpro as complexation agent. At 72 hpt, HyCell, FreeMix and 150 mN NaCl conditions reached ~75% of GFP positive cells (**Figure 2C**). This suggests that the essential parameter driving the significant changes is the complexation agent, regardless of the complexation medium among the tested ones. PEIpro® was proved to achieve better transfection efficiencies and better VLP production values. Additionally, different cell densities were studied at the moment of transfection: $2 \cdot 10^6$, $3 \cdot 10^6$, $4 \cdot 10^6$ and $5 \cdot 10^6$ cells/mL. The condition of $2 \cdot 10^6$ cells/mL provided consistent results compared to the previous tests. The rest of conditions showed proportional cell growth, with higher cell concentration or higher initial cell concentrations (**Figure 2D**). On the other hand, the higher the cell density at the moment of transfection, the less fluorescence and subsequently VLP production, was observed along the studied time course. A significant decrease of VLP production when cells are transfected at $3 \cdot 10^6$ and $4 \cdot 10^6$ cells/mL was observed and no production when transfected at $5 \cdot 10^6$ cells/mL (**Figure 2E**). The same pattern was observed regarding transfection efficiency, where GFP positive cells steeply dropped to less than 5% in any condition transfected at more than $2 \cdot 10^6$ cells/mL at any time point (**Figure 2F**). This can be explained by the cell density effect (CDE), proving that high cell densities present an intrinsic hindrance to be transfected.

FREESTYLE® F17 CULTURE MEDIA OPTIMIZATION

The same variables were tested for FreeStyle® F17 culture media (F17). Here, only PEIpro® was used as the complexation agent since the previous experiments demonstrated that the main differences observed between conditions were caused by this factor. F17, FreeStyle® 293 (FreeStyle), 150 mM NaCl and 125 mM NaCl were tested as complexation medium (**Table 2**). No significant differences were observed regarding cell growth and viability between conditions (**Figure 3A**). As for VLP production, the 125 mM NaCl condition reached significantly lower production at any time point (**Figure 3B**). This can be explained since less than 5% of the cell population was successfully transfected in this condition (**Figure 3C**). For the remaining conditions, no significant differences were observed. Only in VLP production at 48 hpt when the condition using F17 as complexation medium showed a slightly increase compared to the rest. These results correlated with those obtained when optimizing HyCell culture media, since choosing the same media for culturing and complexation provided the best results.

In this case, the experiment transfecting at different cell densities was repeated with F17 and the same results compared to the high cell density tests using HyCell were obtained (**Figure 3D-3F**). Here, transfection efficiency and VLP production dropped as the cell density at the moment of transfection increased. The CDE was again confirmed using F17 culture media.

Table II. Combinations of the different elements for DNA/PEI complex formation process in FreeStyle® F17.

Complexation Medium	Complexation Agent	Condition		
F17	PEI-pro (Polyplus)	F17:		F17
NaCl 150 mM	PEI-pro (Polyplus)	NaCl 150 mM:		NaCl 150mM
NaCl 125 mM	PEI-pro (Polyplus)	NaCl 125 mM:		NaCl 125mM
FreeStyle 293	PEI-pro (Polyplus)	FreeStyle:		FreeStyle
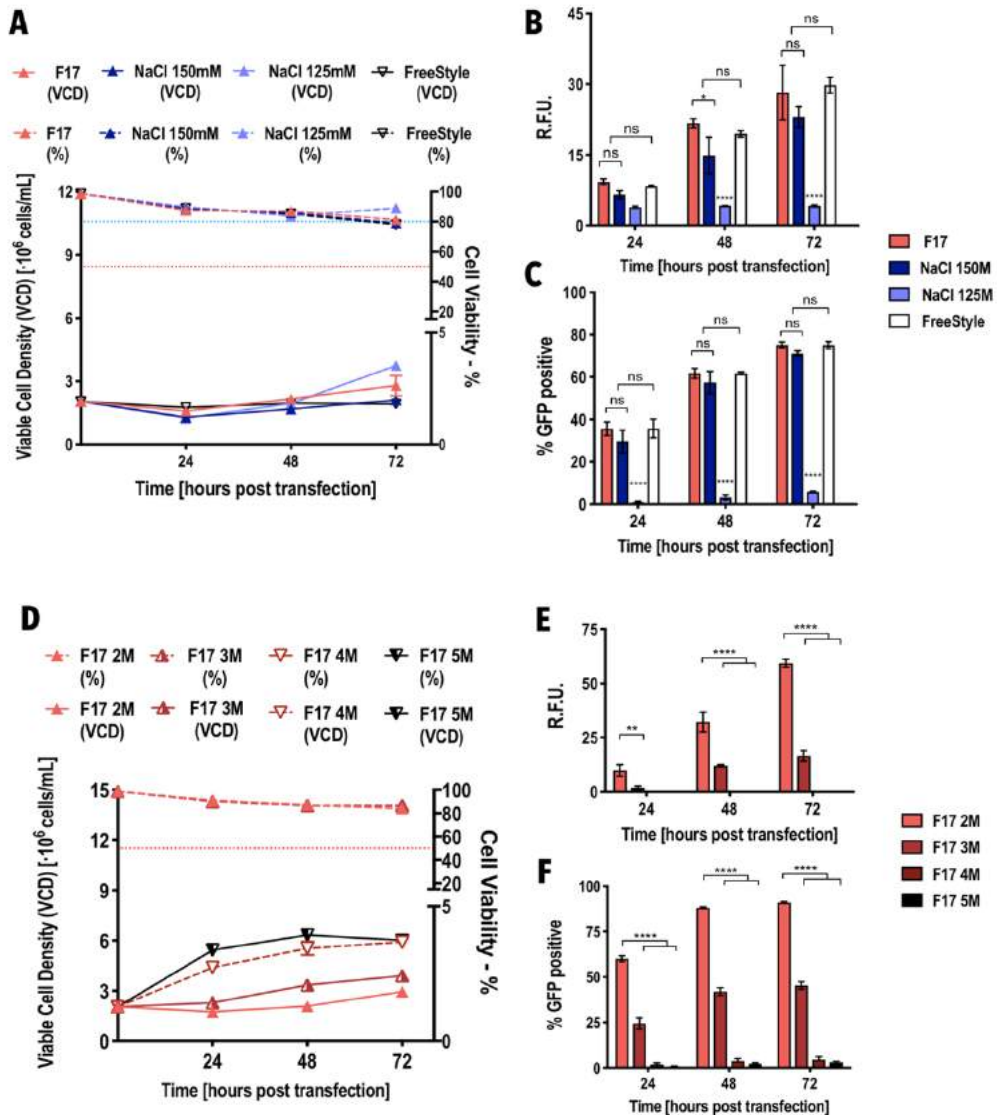
Figure 3. FreeStyle® F17 optimization. **A)** Viable cell density (VCD) and cell viability (%) monitoring along the transient transfection process using the different combinations for complex formation. FreeStyle® F17 culture media (F17), FreeStyle® 293 (FreeStyle), 150 mM NaCl and 125 mM NaCl solution as complexation medium. All conditions used PEIpro® as complexation reagent. Blue and red dotted line represent 80% and 50% cell viability respectively. **B)** Relative fluorescent units (RFU) measured at each time point for each of the different conditions. **C)** Transfection efficiency measured in percentage of GFP positive cells (%). Significance calculated using one- way ANOVA. **D)** VCD and cell viability (%) monitoring along the transient transfection process starting a different cell densities: $2 \cdot 10^6$, $3 \cdot 10^6$, $4 \cdot 10^6$ and $5 \cdot 10^6$ cells/mL. **E)** RFUs measurements for these conditions **F)** Transfection efficiency measurements for these conditions.

CONCLUSIONS

All these data were considered together with variables like specific productivity per million of transfected cells, reproducibility and robustness of the observed results and economic factors like the price of each culture media, which substantially varied between the two, as F17 almost doubles the price of HyCell. The main elements taken into consideration are presented in **Figure 4**. Finally, HyCell was selected as the culture media and also for DNA/PEI complexation in order to reduce the number of reagents, reducing costs. PEIpro® was selected as complexation agent. These elements were selected for the development of a VLP production bioprocess and its subsequent optimization and intensification.
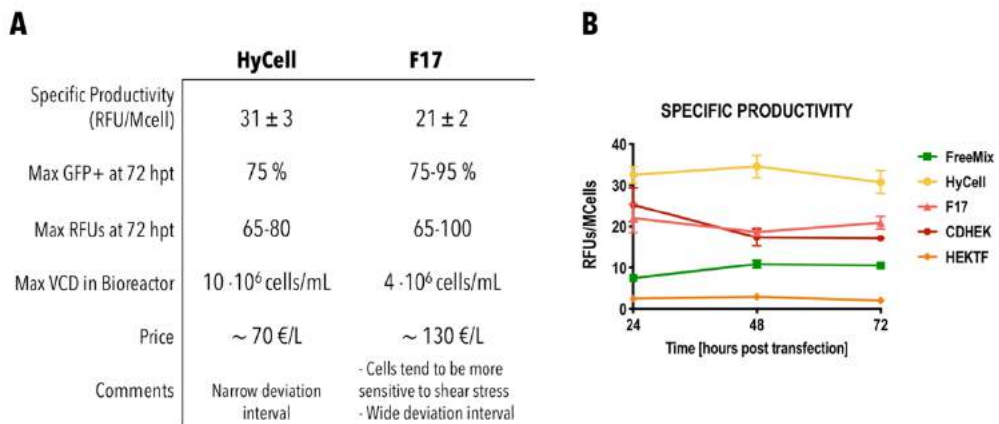
Figure 4. Culture media comparison. **A)** HyCell vs F17 comparison. **B)** Specific productivity of all tested culture media.

BATCH TEST AND PRODUCTION ANALYSIS

The selected conditions were used for a batch test, using transient transfection for VLP production. Cells were culture at a working volume of 1L in a 1.5L bioreactor and transfected using the previously optimized parameters and reagents at $2 \cdot 10^6$ cells/mL. The bioprocess main parameters are presented in **Figure 5A**. Here, it can be observed the pO_2 and pH profile along the studied time course up to 72hpt when VLPs were harvested. After transfection, cell growth is slowed down until ~48hpt when growth rate increased again. This already described tendency was also observed in batch cultivation mode at bioreactor scale. Cell viability always surpassed 80%. Regarding transfection efficiency, the same percentage of GFP positive cells acquired in the parallel shake flask cultures was obtained in the bioreactor (**Figure 5B**). However, the VLP concentration achieved in batch at bioreactor scale was significantly lower than the obtained in shake flasks. The overall batch performance is presented in **Table 3**.

The difference in VLP production suggested that the process at bioreactor scale needed to be optimized and intensified. In order to do this, the strategy of extended gene expression (EGE) that had already proven effective in shake flasks, needed to be optimized at bioreactor scale. Altogether, the elements previously selected in the media and transfection optimization phase and the implementation of EGE at bioreactor scale set the basis for the development of a perfusion-based bioprocess and its subsequent optimization and intensification presented in Part II.

Table III. Batch overall process performance

Accumulated VLPs	1.60E+13
VLP concentration in reactor at 72hpt (VLP/mL)	1.60E+10
VLP specific productiviy (VLP·cell^{-1}·day^{-1})	8.90E+02
Volumetric productivity (VLP·L^{-1}·day^{-1})	2.70E+12
Total volume of spent media (L)	1
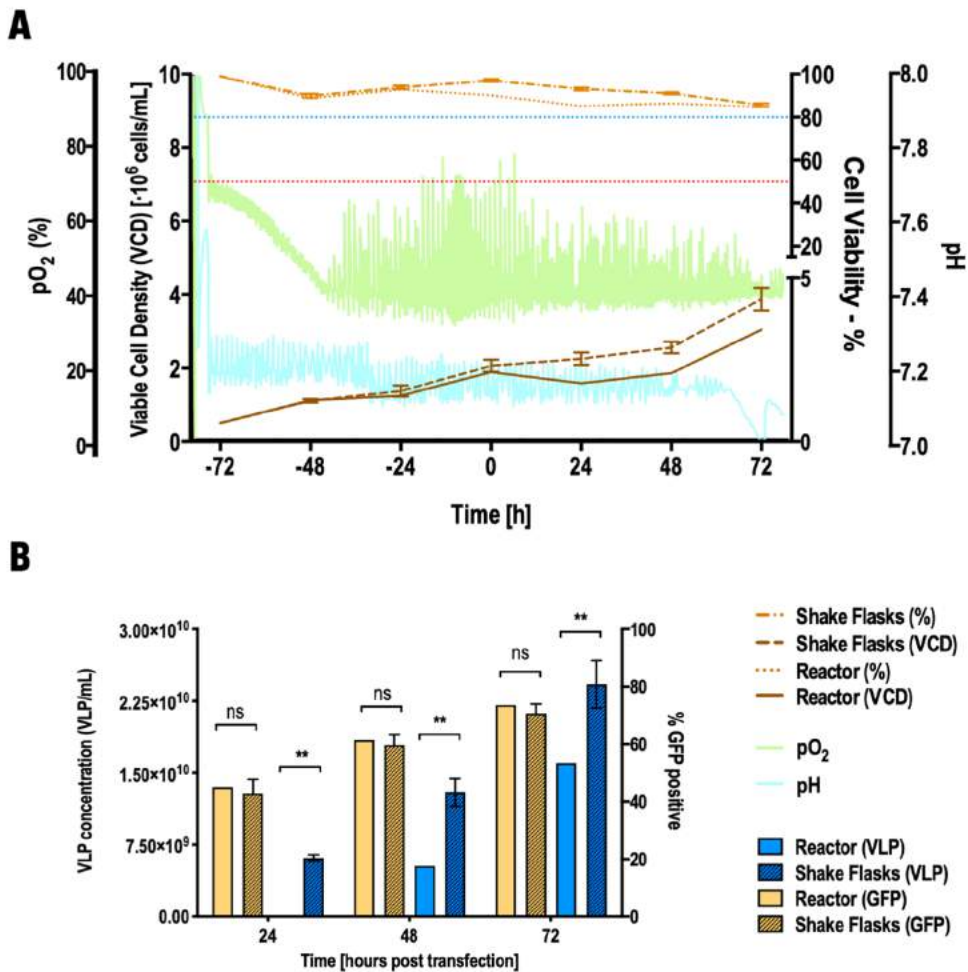
Figure 5. Batch VLP production. **A)** Viable cell density (VCD), cell viability, pH and pO_2 graphs along the studied process. The red and blue dotted line represent 50 and 80% of cell viability respectively. Cells were transfected at time point 0h. Agitation was set at 200 rpm and temperature at 37°C. **B)** Transfection efficiency measured in percentage of GFP positive cells and VLP concentration assessed by fluorimetry measurements and calculation using the previously described quantification method.

REFERENCES

1. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

2. Cervera, L. *et al.* Selection and optimization of transfection enhancer additives for increased virus-like particle production in HEK293 suspension cell cultures. *Appl Microbiol Biotechnol* **99**, 9935–9949 (2015).

3. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

4. Buffo, M. M. *et al.* Influence of dual-impeller type and configuration on oxygen transfer, power consumption, and shear rate in a stirred tank bioreactor. *Biochem. Eng. J.* **114**, 130–139 (2016).

5. Carpentier, E., Paris, S., Kamen, A. A. & Durocher, Y. Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *J Biotechnol* **128**, 268–280 (2007).

CHAPTER THREE

Part II.

An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures

ABSTRACT

Virus-like particles (VLPs) have gained interest over the last years as recombinant vaccine formats, as they generate a strong immune response and present storage and distribution advantages compared to conventional vaccines. Therefore, VLPs are being regarded as potential vaccine candidates for several diseases. One requirement for their further clinical testing is the development of scalable processes and production platforms for cell-based viral particles. In this work, the extended gene expression (EGE) method, which consists in consecutive media replacements combined with cell retransfections, was successfully optimized and transferred to a bioreactor operating in perfusion. A process optimization using design of experiments (DoE) was carried out to obtain optimal values for the time of retransfection, the cell specific perfusion rate (CSPR) and transfected DNA concentration, improving 86.7% the previously reported EGE protocol in HEK293. Moreover, it was successfully implemented at 1.5L bioreactor using an ATF as cell retention system achieving concentrations of $6.8 \cdot 10^{10}$ VLP/mL. VLP interaction with the ATF hollow fibers were studied via confocal microscopy, field emission scanning electron microscopy and nanoparticle tracking analysis to design a bioprocess capable of separating unassembled Gag monomers and concentrate VLPs in one step.

ABBREVIATIONS

ATF: Alternating tangential flow; **CCD**: Central composite design; **CDE**: Cell density effect; **CSPR**: Cell specific perfusion rate; **DoE**: Design of experiments; **EGE**: Extended gene expression; **eGFP**: enhanced green fluorescence protein; **Gag::eGFP**: translational fusion of HIV-Gag protein and eGFP; **HFM**: Hollow fiber module; **Hpt**: hours post transfection; **NTA**: Nanopaticle tracking analysis; **PBS**: Phosphate-buffered saline; **PEI**: polyethyleneimine; **PES**: Polyethersulfone; **RFU**: Relative fluorescence units; **RSM**: Response surface methodology; **FESEM**: Field emission scanning electron microscopy; **SF**: Shake Flasks; **SPR**: Sequential Perfusion Recovery; **STR**: Stirred tank bioreactor; **TGE**: Transient gene expression; **VLPs**: Virus-like particles

INTRODUCTION

Since the human immunodeficiency virus 1 (HIV-1) was first reported in 1959, 74.9 million people have been infected[1]. Due to the high impact and importance of this disease, many efforts have been devoted to understand and characterize the mechanisms taking place in HIV infection and replication[2-5]. For replication and virion production, HIV-1 Gag polyprotein play a key role[6]. Upon expression of this polyprotein, it self-assembles forming the virion protein core which can be later packed with the genetic material and bud off the host cell[2,4,7]. This characteristic can be exploited to obtain protein core structures comprising Gag monomers and lacking genetic material. The use of Gag to generate virus-like particles (VLPs) has become a promising option to develop alternative platforms for recombinant vaccines being able to target several diseases[8,9]. Having multimeric structures such as Gag-VLPs helps induce a stronger immune response[10] and since they are produced by a budding process, the lipid envelope can be further used for particle pseudotyping[11-13]. The addition of conjugated antigens can be used to target different diseases, such as influenza[14]. The use of mammalian cell cultures to express HIV-Gag in order to produce VLPs offers the advantage of providing the necessary cellular machinery to produce enveloped particles via membrane budding and performing the correct post-translational modification[15]. One methodology to express heterologous proteins like Gag, is transient gene expression (TGE)[16]. It consists in the introduction

of a plasmid carrying the gene coding for the protein of interest that will remain episomal in the nucleus, while it is being expressed. This expression method allows production of relatively high titers of protein in a short period of time, reduces time costs related to the generation of a stable cell line and offers high versatility when developing a production platform regardless of the final product, as only the plasmid of interest needs to be changed for a new product. Nevertheless, this method is limited in its timespan, since the plasmid is not integrated in the cell genome and upon consecutive rounds of cell division it is eventually lost. To overcome this issue and envisioning a future step towards industrialization, a protocol named Extended Gene Expression (EGE) protocol was designed allowing to advance from a typical harvest at 72 hours post transfection (hpt) to a continuous harvest up to 240 hpt[17]. The EGE protocol increased protein production up to 12-fold and consisted of a combination of media replacements every 48 hours and two retransfections at 48 and 96 hpt. The fact of having a constant rate of media replacement makes the EGE protocol suitable for continuous biomanufacturing using cell retention devices. Continuous production allows the constant renewal of nutrients and depletion of cell waste, toxic by-products or metabolites that might interfere with the culture homeostasis[18]. As a result, higher volumetric productivity can be reached compared to batch and fed-batch culture. Indeed, the reduction of bioreactor size and facilities infrastructure, with a reduction in production costs, is facilitating the adoption of continuous processes in the biopharmaceutical industry over the last few years[19]. One of the main advantages of continuous bioprocessing is that this mode of operation allows the support of high cell density cultures, dramatically increasing volumetric productivity as the product is being constantly harvested and operation times extended. The fact of achieving higher cell densities in shorter times was one of the issues encountered in the first attempt of transferring the EGE protocol to a bioreactor operated in continuous mode without further optimization[20]. Indeed, cells need to be maintained around $2 \cdot 10^6$ cells/mL for an efficient transfection. Higher cell densities are more difficult to transfect due to the widely described cell density effect (CDE)[21-23]. Transfection efficiency and productivity decrease when performed at higher cell densities. The energetic demand is solved with the constant addition of fresh media but the CDE comprise molecular causes that cannot be solved only by perfusion[21]. In this work, we propose overcoming this problem by changing the approach of perfusion using cell retention devices, like alternating tangential

flow (ATF), not to achieve high cell densities, but to ensure the constant media replacement in order to maintain the cell culture in a re-transfectable state for a longer period of time in the bioreactor and therefore, prolong production. This perfusion approach is not intended to operate at steady state and cell bleeding is not necessary, as continuous perfusion is used to constantly renew media while cells are retransfected to keep a high proportion of cells producing VLPs, up to a certain limited process span. To achieve a scalable optimization of the EGE protocol, key parameters important to improve production yields were considered. The time of retransfection, the cell specific perfusion rate (CSPR), and the DNA concentration added in each transfection were the three variables optimized using a Design of Experiments (DoE) method. VLP production was assessed and a new approach to develop the EGE protocol at bioreactor scale using ATF is suggested. This intensified and optimized procedure can reduce costs and increase VLP production for a future potential implementation at industrial scale, offering a production platform for novel VLP-based vaccines.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from McGill University (Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell culture media was HyCell™ TransFx-H media from HyClone™ (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA).

Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

For the pseudoperfusion experiments, the total culture volume was 20mL and media replacement (MR) was carried out centrifuging the culture at 300xg for 5 min every 12h ensuring that the proportional

volume of media was replaced depending on the condition. To maintain a MR rate of 2 reactor volume per day (RV/day), 20 mL were replaced every 12h. For a rate of 1 RV/day, 10 mL were replaced every 12h and for a rate of 0.5 RV/day, 5 mL were replaced every 12h.

BIOREACTOR CULTURE CONDITIONS AND SETUP DESCRIPTION

A BioStat B Plus bioreactor (Sartorius Stedim Biotech, Goettingen, Germany) equipped with a 3-blade segment dual impeller with UP-DP configuration[24] was used for HEK293 cell cultivation. The agitation was set at 200 rpm. Temperature was set at 37 °C. The pH was set at 7.1 and controlled with CO_2 and $NaHCO_3$ (7.5%w/v). The dissolved oxygen was controlled at 40% of air saturation by supplementing air by sparger at a constant flow of 0.1L/min and additional pure oxygen when needed. HEK293 growing exponentially in disposable polycarbonate 1 L shake flasks (Corning®) were used to seed the bioreactor at $0.5 \cdot 10^6$ cells/mL in 1.5 L. Samples were taken every 24 h for cell counting and viability determination. After transfection, assessment of the percentage of GFP positive cells and VLP quantification was also performed every 24h. Perfusion was achieved using an alternating tangential flow (ATF) cell retention device (Repligen, Waltham, MA, USA) with 0.5 and 0.2 μm pore size and 0.13 m^2 of filtration area hollow fiber modules (Repligen, Waltham, MA, USA) and a ATF flow rate of 0.6 L/min. When performing transfection, perfusion was stopped to incubate the cells with the DNA/PEI solution and reestablished 2 hours after transfection[25]. To carry out media replacement, the filtration rate was set at 0.26 mL/min at the beginning of the process using a MasterFlex L/S peristaltic pump (MasterFlex Group, Gelsenkirchen, Germany) and was modified every day depending on the viable cell density to maintain a cell specific perfusion rate (CSPR) of 30 pL/cell/day. The bioreactor and the ATF system were placed over a scale (MSE36200S-000-D0 Cubis, Sartorius, Goettingen, Germany) to control the mass displacement created due to the filtration of spent media. The scale was connected to a Scilog peristaltic pump (Scilog tandem 1081, Parker, Oxnard, CA, USA) which controlled the addition of fresh media upon detection of mass displacement to maintain constant volume in the reactor.

TRANSIENT TRANSFECTION

The initial transfection (t_1) was carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 µg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used.

The plasmid used contained the gene coding for HIV-Gag protein fused to eGFP (Gag::eGFP). Briefly, the corresponding DNA was diluted with supplemented HyCell™ media and vortexed for 10 seconds. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, then the mixture was incubated for 15 min at room temperature and then added to the cell culture.

In the different series of perfusion experiments performed, the first and second retransfections (t_2 and t_3) were carried out following the previous protocol and incorporating some variations depending on each of the optimization experiment. The specific conditions used in each case are detailed in the *Results and Discussion* section.

FLOW CYTOMETRY

Samples were taken every 24h after transfection and cells were fixed using formaldehyde 2% for 10 minutes, centrifuged and then resuspended in PBS for FACS analysis. The percentage of GFP positive cells was assessed using a BD FACS Canto flow cytometer (BD Biosciences, San Jose, CA, USA). Laser 488 was used for GFP measurement. The results were analyzed with FACS DIVA software (BD Biosciences, San Jose, CA, USA).

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using a developed and validated quantification assay[26]. VLP containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples.

HIV-1 GAG VLP QUANTIFICATION BY NANOPARTICLE TRACKING ANALYSIS (NTA)

Nanoparticle Tracking Analysis (NTA) was also used to quantify fluorescent particles. NTA measurements were performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag::eGFP VLPs and neutral density filter for total particle by light scattering. Data was analyzed with NanoSight® NTA 3.1 software. Briefly, samples were injected, and independent analyses were carried out. Three video recordings of 60sec duration were taken for each sample. Capture settings were recorded with an sCMOS camera (camera level of 8 for Gag::eGFP VLP samples, and 11 for controls, viscosity: 0.9 cP) and analyzed with a detection threshold of 4.

FIELD EMISSION SCANNING ELECTRON MICROSCOPY (FESEM) VISUALIZATION

Morphometry of the hollow fiber inner layer and fluorescence detection at nanoscale were determined by Field Emission Scanning Electron Microscopy (FESEM). The analyzed samples were 0.5 μm pore size hollow fiber samples exposed to non-transfected HEK293 as a control, 0.5μm and 0.2μm pore size hollow fibers exposed to VLP-producing HEK293 cell cultures. Hollow fibers were longitudinally cut in pieces of approximately 2-3 mm^2 and deposited in carboncoated gold grids (200 mesh) during 1 min, air dried and observed in a FESEM Zeiss Merlin (Zeiss, Jena, Germany) operating at 1.5 kV and 3.4 mm of working distance. Samples were then randomly checked with an in-lens secondary electron detector for morphology and with a Back-scattered Electron (BSE) detector for fluorescence detection. Representative images were obtained at a wide range of high magnifications (from 200,000x to 500,000x).

CONFOCAL MICROSCOPY VISUALIZATION

The visualization of the hollow fiber modules was performed using a FluoView®FV1000 confocal microscope (Olympus, Tokyo, Japan) at sampling speed of 2 μm/pixel, excitation at 488 nm and detection at 500-600 nm. Step size was 2 μm/slice using XYZ scan mode. Transversal and longitudinal cuts of the hollow fiber were made from samples of 0.5μm pore size exposed to non-transfected HEK293 as a control and 0.5μm and 0.2μm pore size hollow fibers exposed to VLP-producing HEK293 cell cultures. For the

longitudinal cuts, objective lens UPLAPO 20x, NA: 0.75 and optical zoom x3 was used. For the transversal cuts, objective lens UPLAPO 10x, NA: 0.40 and optical zoom 1x was used.

OPTIMIZATION OF RETRANSFECTION PARAMETERS USING DESIGN OF EXPERIMENTS

Retransfection parameters were optimized in order to maximize VLP specific productivity. The three variables chosen for optimization were the time of retransfection, the cell specific perfusion rate (CSPR) and the DNA concentration added in each retransfection. The adjusted VLP specific productivity (P_{sp} : VLP·10^6 cells^{-1}·days^{-1}) for the DoE response was calculated taking into account the volume of media replacement and the cell culture viability in each day along the culture using the equation showed below (**Eq. 1**):

$$P_{sp} = \frac{\sum_{i=1}^{d}(\frac{C_{VLP}}{C_x} \cdot K_{MR_i} \cdot v_i)}{d} \qquad (1)$$

$$K_{MR} = \begin{cases} \frac{MR}{RV} < 1 & 1 - \frac{MR}{RV} \\ \frac{MR}{RV} > 1 & \frac{MR}{RV} \end{cases} \qquad (2)$$

where C_{VLP} is the VLP concentration (VLP/mL), C_x is the concentration of viable transfected cells (10^6 cells/mL), v_i is the cell viability (%), d is the number of culture days after the first transfection and K_{MR} is the media replacement coefficient, proportional to the total volume replaced each day. K_{MR} is calculated as shown in **Equation 2**, where MR is the replaced media and RV is the reactor volume. Thus, estimating the final volume at the time extracellular fluorescence was measured. A Box-Behnken design was selected to define the optimal value for each variable. The three variables were screened at three levels: a low level coded as -1, an intermediate level coded as 0, and a high level coded as $+1$ as indicated in **Table 1**. Box-Behnken experimental results were fitted to a second order polynomial equation described below (**Eq. 3**)

by non-linear regression analysis: where Y is the response in VLP·cell^{-1}·days^{-1}, β_0 is the offset term, β_i is the linear coefficient, β_{ii} is the quadratic coefficient, β_{ij} is the interaction coefficient, and X_i and X_j are the independent variables. This equation was used to predict the optimal values of the independent variables

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ii} X_i^2 + \sum \beta_{ij} X_i X_j \qquad (3)$$

using R software (RStudio, Inc., Boston, MA, USA). The quality of the fit of the model equation is expressed by the coefficient R^2 and *p-values* obtained by regression analysis. Additionally, a lack of fit test was performed to compare the experimental error to the prediction error. The overall significance of the model was determined by analysis of variance (ANOVA) F test, whereas the significance of each coefficient was determined by the corresponding t test.

RESULTS AND DISCUSSION

STUDY OF THE CELL GROWTH UPON DIFFERENT MEDIA REPLACEMENT RATES

Monitoring cell growth under three different media replacement rates showed that maximum cell concentration, reaching $22 \cdot 10^6$ cells/mL, was achieved in the shake flask replacing media at 1RV/day and no difference was observed at 2RV/day (**Figure 1A**). This suggests that 1RV/day is enough to meet the nutrient demand of the culture and further cell growth could be hindered by other factors such as oxygen diffusion in the flask. Cell growth monitoring was ceased when viability reached 50% or below, and no more data from that condition was acquired. The observed fluctuations are due to the process of centrifugation and resuspension, which added an intrinsic error to the monitoring process. However, this fluctuation does not affect the following CSPR calculation. When the CSPR is calculated for every time point in each condition (**Figure 1B**) it can be observed that it tends to converge to a range going from 30-60 pL/cell/day range, suggesting that maintaining the media replacement rate within this interval is enough to maintain cell growth.

A

B
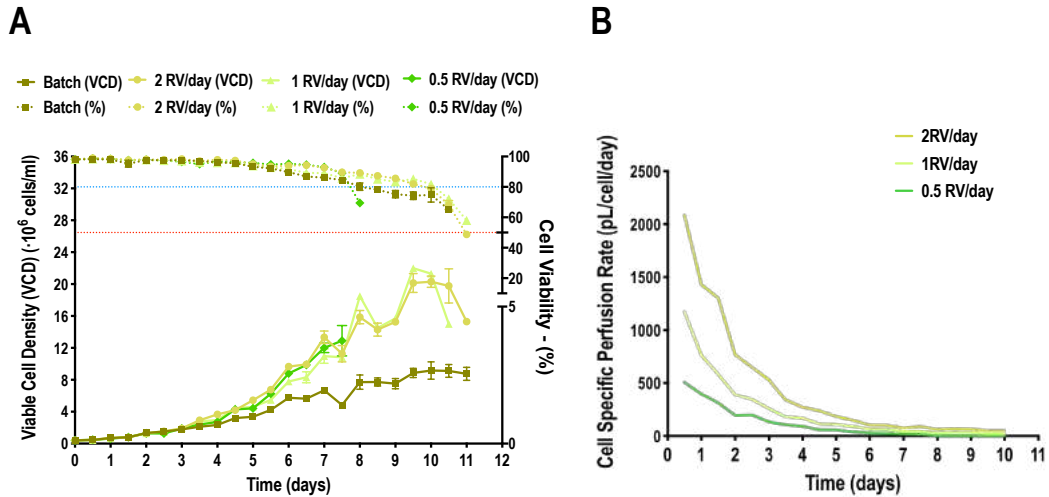
Figure 1. Pseudoperfusion growth curves A) HEK293 cell culture growth curves in shake flasks where medium replacement was performed at a rate of 0.5, 1 or 2 reactor volume per day (RV/day) and a batch condition where no medium replacement was carried out. Blue and red dotted lines represent 80% and 50% of cell viability respectively. Experiment performed n=3. B) Representation of the cell specific perfusion rate (CSPR) from cell cultures with different rate of medium replacement.

OPTIMIZATION OF RETRANSFECTION BY DESIGN OF EXPERIMENT (DOE) METHOD

In order to optimize the EGE protocol, the continuous variables *time of retransfection*, *CSPR* and the *DNA concentration* used to retransfect, were selected. These variables were studied by means of a design of experiments (DoE). A response surface methodology (RSM), specifically a Box-Behnken design, was selected as it allows to reduce number of experimental runs and yet obtain statistically relevant information[27]. The parameters of the first transfection point had been already optimized in previous work and fixed in $2 \cdot 10^6$ cells/mL and 1μg/mL of DNA[28,29]. The variables of [DNA:PEI] ratio, nature of the complexation agent and time for the maturation of the DNA:PEI complexes were also individually determined in previous works[30,31]. The range of DNA was also selected based on this previous experience and most importantly, restrained by the PEI cytotoxicity. In addition to this, multiple variable optimization was selected for the presented study since the VLP production wanted to be tested combined with CSPR and time of retransfection to consider possible synergic effects and simulate the closest condition to the bioreactor. The ratio [DNA:PEI] was fixed to [1:2] and maintained for all the experiments following the

standard procedure. Regarding the time of retransfection, the working range was set from 24hpt to 72hpt, based on the EGE protocol where the first retransfection was performed at 48hpt. As for the CSPR, the working range was set based on the previous study of the cell growth upon different media replacement rates. The upper limit was set at 1000 and the lower limit at 30 pL/cell/day. Since the DNA concentration used for retransfection in the original EGE protocol was 0.5 μg/mL, the lower limit was set at this value. The upper limit for the DNA was set in 2μg/mL due to the cytotoxic effects of PEI[28,32]. Working ranges for each variable are presented in **Table 1**. A three-factor three-level Box-Behnken design was constructed using the selected ranges for each variable as the design space boundaries. A 15-experiment matrix was defined in which the central point was triplicated, and each experimental run was duplicated to assess the pure experimental error. This led to a 30-experiment matrix (**Table 1**). Each condition was retransfected at its corresponding time and harvested 72 hours after. The response variable considered in this analysis was the adjusted VLP specific productivity (P_{sp}) calculated as described in **Equation 1**. This equation is proposed to optimize specific productivity considering two aspects of VLP production. A media replacement coefficient (K_{MR}) was added, reflecting how many times a reactor volume of media was replaced every 12 hours, making it directly proportional to the amount of replaced media to favor the conditions where the VLP concentration in the supernatant was diluted due to a higher replacement rate. A term reflecting cell viability (v) was also added to penalize conditions with low viability which are incrementing extracellular relative fluorescent units (RFUs) due to the release of unassembled free monomer caused by cell death. This is one of the reasons why specific productivity was chosen over total volumetric productivity as the DoE response to be based on. Otherwise conditions reporting high fluorescence intensity, but low cell densities and viabilities had been favored. Data was fitted to a second-order model by non-linear regression analysis. The statistical significance of the nonlinear regression was confirmed by ANOVA analysis, also showed in **Table 1**. Data from the model was used to construct response surface plots where the interaction between factors can be observed (**Figure 2**). Noticeably, lower CSPRs led to higher adjusted specific productivities (P_{sp}) (**Figure 2D-2F**). This condition allows to minimize media consumption while increasing productivity. Higher CSPRs led to the same cell densities

Table I. Box-Behnken design, results and ANOVA analyses for optimization of the extended gene expression (EGE) protocol for VLP production.

		-1	0	1
Retransfection time (hpt)		24	48	72
CSPR (pL/cell/day)		30	515	1000
DNA (mg/mL)[a]		0.5	1.25	2

Experimental Run	Retransfection time	CSPR	DNA/mL	P_{sp}
1	1	0	-1	1.83E+09
2	1	0	-1	2.20E+09
3	-1	0	-1	2.68E+09
4	-1	0	-1	3.29E+09
5	1	0	1	1.85E+09
6	1	0	1	1.92E+09
7	0	-1	-1	5.06E+09
8	0	-1	-1	4.33E+09
9	0	1	-1	2.32E+09
10	0	1	-1	2.36E+09
11	0	1	1	2.10E+09
12	0	1	1	2.47E+09
13	-1	1	0	4.64E+09
14	-1	1	0	2.91E+09
15	1	-1	0	5.21E+09
16	1	-1	0	5.56E+09
17	0	0	0	2.07E+09
18	0	0	0	1.92E+09
19	0	0	0	1.72E+09
20	0	0	0	1.89E+09
21	0	-1	1	5.19E+09
22	0	-1	1	4.10E+09
23	1	1	0	2.25E+09
24	1	1	0	1.94E+09
25	-1	0	1	2.40E+09
26	-1	0	1	4.07E+09
27	-1	-1	0	5.65E+09
28	-1	-1	0	5.19E+09
29	0	0	0	2.12E+09
30	0	0	0	2.38E+09

Model	Multiple R^2	p value [b]	Lack of fit [c]
	0.9126	1.06E-08	0.9316

Parameters	Coeficient		t	p value [b]
Constant	2.02E+09		10.225	<0.0001
[Time]	-5.04E+08		-4.1761	0.0005
[CSPR]	-1.21E+09		-9.9874	<0.0001
[DNA]	1.88E+06		0.0155	0.988
[Time]·[CSPR]	-4.11E+08		-2.4077	0.026
[Time]·[DNA]	-9.50E+07		-0.5562	0.584
[CSPR] · [DNA]	-1.25E+06		-0.0073	0.994
$[Time]^2$	5.95E+08		3.3492	0.003
$[CSPR]^2$	1.56E+09		8.7562	<0.0001
$[DNA]^2$	-8.21E+07		-0.4617	0.649

Optimal values				
	Time of Retransfection	CSPR	DNA	P_{sp}
	-1	-1	0.5977154	5.49E+09
	At 24hpt	**30 pL/cell/day**	**1.7 $\mu g/ml$**	

[a] DNA/PEI ratio was always maintained at [1:2]

[b] *p values* under 0.05 are considered statistically significant with 95 % confidence, and *p values* under 0.1 are considered statistically significant

[c] *p values* associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.

hpt: hours post transfection

at the moment of transfection, but the transfection percentage dropped faster in these conditions as the subpopulation of non-transfected cells grew more rapidly than the transfected one, losing the plasmid more rapidly (**Figure 2J-2L**). This can be explained as overall cell homeostasis is disrupted upon transfection, reducing cell growth and downregulating the cell cycle[33]. Regardless of the time of retransfection, decreasing CSPR led to an increase in the specific productivity. It should also be considered that, due to cytotoxicity of the complexation agent (PEI), an increase in the DNA concentration compromises viability. The model was used to predict an optimal combination of the three factors to maximize P_{sp}, reaching a value of $5.5 \cdot 10^9$ VLP·Mcell^{-1}·days^{-1}. The optimal set of parameters for the first retransfection point was to perform it at 24hpt using 1.7 µg/mL of DNA and maintaining a CSPR of 30 pL/cell/day. In order to validate the model, a confirmation experiment (n=4) was performed using the set of parameters determined by the optimal solution predicted from the model. A P_{sp} of $(5.7 \pm 0.8) \cdot 10^9$ VLP·Mcell^{-1}·days^{-1} was obtained, confirming the model adequacy.

OPTIMIZATION OF THE SECOND RETRANSFECTION PARAMETERS. SHAKE FLASKS STUDY

Following the EGE protocol and after optimizing the first retransfection, a similar approach was performed to optimize the second re transfection. This time, based on the results from the first optimization, a CSPR of 30 pL/cell/day was maintained for all the experiments. In order to estimate the optimal values for the retransfection time and DNA concentration, a central composite design (CCD) was used. The same working range as in the Box-Behnken design was maintained for the DNA and for the time of retransfection, the working range was set from 24 to 72 hours post retransfection. A 10-experiment matrix was defined, and each experimental run was duplicated to assess the experimental error. In this case, the temporal proximity of the three transfections was so close that cell toxicity caused by PEI made unfeasible the completion of the experiment for the tested conditions. Therefore, no statistically significant model was obtained (data not shown). As this approach did not succeed to optimize the second retransfection, a further analysis of the culture behavior once transferred to the bioreactor was carried out as discussed further.
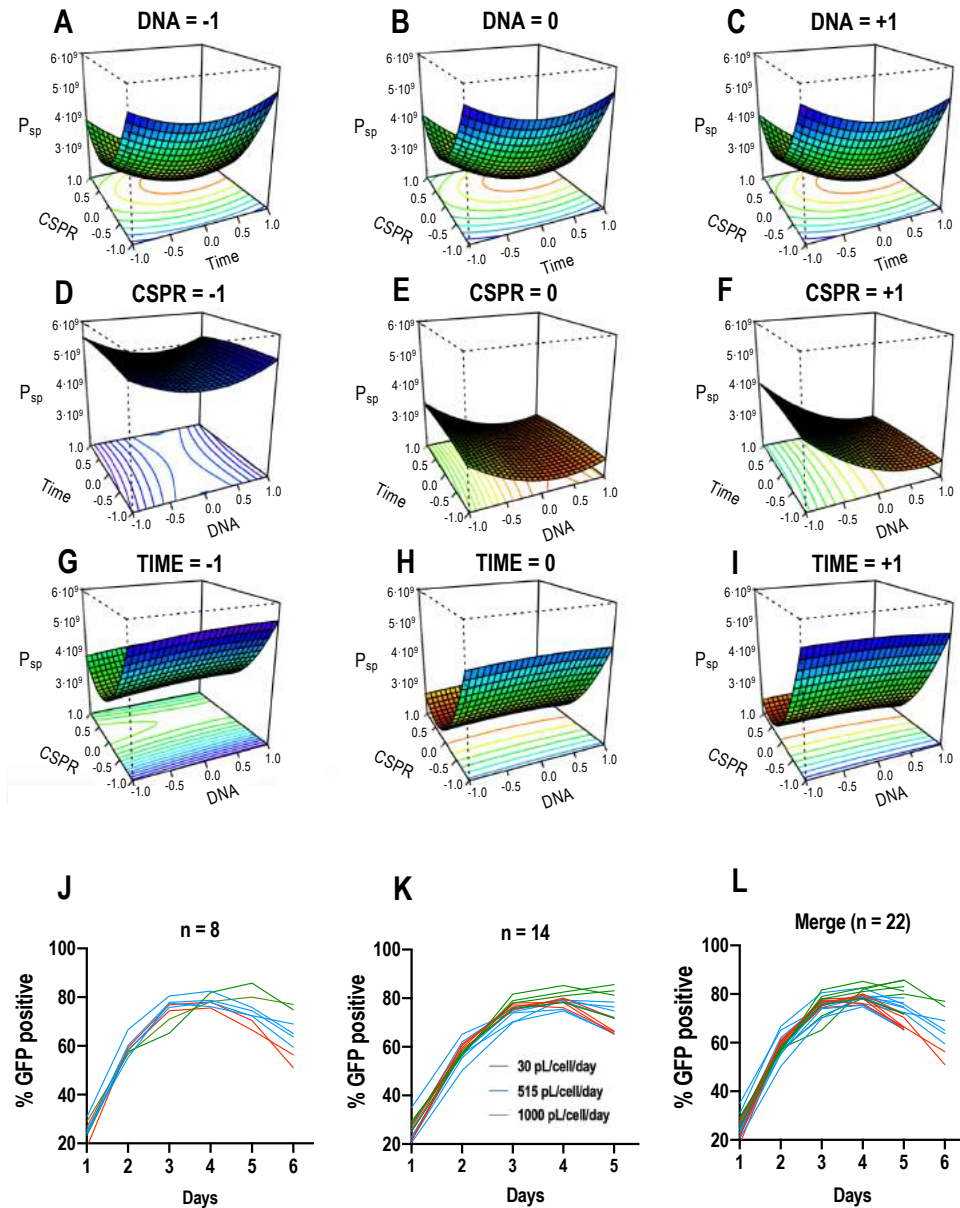
Figure 2. Response surface graphs based on Box-Behnken experimental results. Maximum VLP specific production in cell culture supernatants as a function of **A**, **B**, **C** CSPR vs. Time; **D**, **E**, **F** time vs. DNA; and **G**, **H**, **I** CSPR vs. DNA. The graphs were constructed by depicting two variables at a time and maintaining the third one at a fixed level. +1, 0 and -1 correspond to 0.5, 1.25 and 2 mg/mL for the DNA concentration; 30, 51.5 and 100 pL/cell/day for the CSPR and 24, 48 and 72 hpt for the time of retransfection. **J,K,L** represent the changes of the percentage of transfection measured in percentage of GFP positive cells along the studied time course for conditions reaching the sixth day of the process (**J**) and conditions reaching the fifth (**K**). **L** shows the merged plot.

INTENSIFICATION OF THE OPTIMIZED PROTOCOL IN BIOREACTOR USING ATF

The ATF configuration previously described in the Methods section and illustrated in **Figure 3A** was implemented to apply the optimal solution obtained using the Box-Behnken design to the stirred tank bioreactor (STR) culture operating in perfusion mode. Transfecting at $2 \cdot 10^6$ cells/mL at 0 hours and then retransfecting at 24hpt reduced cell growth, making the viable cell density (VCD) to remain stable at 3-$4 \cdot 10^6$ cells/mL up to 144hpt (**Figure 3B**). Up to this point viability steadily dropped from 98 to 50%, following the same trend in both shake flasks (SF) and in the STR. From 144hpt viability started to restore and cell growth recovered exponential rate until 264hpt when the culture was stopped. In this phase of the culture, the STR reached $22\text{-}23 \cdot 10^6$ cells/mL compared to the $10 \cdot 10^6$ cells/mL observed in SF. Although the same CSPR was used for both systems, in the STR, perfusion was automatically and continuously operated while in SF a pseudoperfusion was carried out, replacing the corresponding amount of media manually every day. Operating in continuous mode proved to be more efficient as nutrients were continuously replenished and waste removed. The percentage of transfection and the fluorescence intensity were monitored daily. There was a peak in transfection at 96-120hpt reaching 90% of transfected cells (**Figure 3D**) and a peak of production at 168hpt reaching 115-120 RFUs in the bioreactor (**Figure 3E**). RFU were converted to VLP concentration using a quantification method by fluorimetry previously described[26]. Considering the VLP concentration in the STR, the harvest and the corresponding CSPR, the total VLP production per day was calculated. As it can be observed in **Figure 3F**, the system ceased producing VLPs at 168hpt, indicating the 13 days process can be reduced to 9 days subsequently reducing time, resources and costs. Cumulative VLP production (**Figure 3G**) achieved was $6.8 \cdot 10^{10}$ VLPs/mL, improving in 86.7% or 7.54 fold the previous reported results[17] using the original EGE protocol in shake flasks. Comparing the EGE protocol carried out in bioreactor prior to the developed optimization[20], the presented work achieved a reactor and media volumetric productivity of $7.1 \cdot 10^{12}$ and $2.7 \cdot 10^{12}$ VLP$\cdot$L$^{-1}\cdot$day^{-1} respectively, improving 26.8% or 1.36 fold and 67.8% or 3.1 fold respectively the reported results. This improvement led to the achievement of an average VLP specific productivity of 3000 VLP$\cdot$cell$^{-1}\cdot$day^{-1} (**Figure 3C**). The same VLP concentration, specific productivity and trend in transfection efficiency were
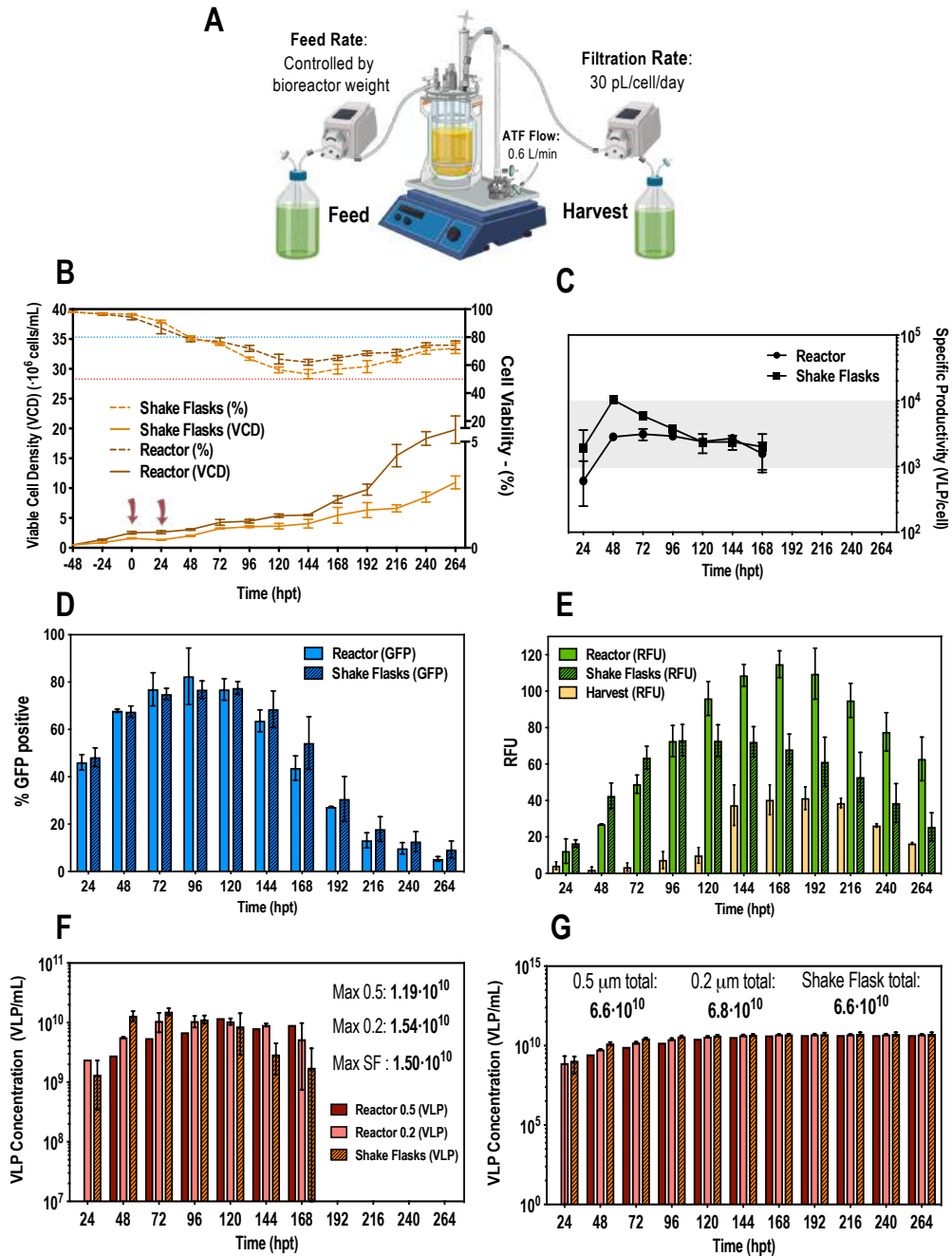
Figure 3. A) Schematic representation of the implemented ATF perfusion system. **B)** Viable cell density and cell viability graphs along the studied time course. The blue red dotted lines indicate 80% and 50% of cell viability respectively. Cells were transfected at time point 0 and 24 hours post-transfection (hpt) noted by red arrows. **C)** VLP specific productivity rate along the studied time course. **D)** Flow cytometry analysis of the transfection percentage (showed in GFP positive percentage of cells) along the time course. **E)** Measurements of fluorescence in relative fluorescent units (RFUs) from the reactor, the harvest from the reactor and the shake flasks. **F)** Total VLP concentration produced per day. Two different pore size for the hollow fiber module were used in the ATF system: 0.5 and 0.2 mm denoted as 0.5 and 0.2 respectively. **G)** Total cumulative VLP concentration. Experiments performed n=3.

achieved in SF and STR, proving that the optimized EGE protocol was successfully transferable to the 1.5L bioreactor.

OPTIMIZATION OF THE SECOND RETRANSFECTION PARAMETERS. BIOREACTOR STUDY

After the difficulties of the first attempt of optimizing a second retransfection at small scale using pseudoperfusion, a different approach was taken. Considering the trends in cell viability, a second retransfection following the standard protocol with 1μg/mL of DNA, was carried out at 168hpt. At this point, viability surpassed 60% and started recovering. Cells need to be growing to be efficiently transfected as they need to be dividing for exogenous DNA to enter the nucleus[25]. Therefore, at 168hpt their recovering state would allow a new retransfection. In addition to this, at 168hpt VLP production ceased. At this moment the VCD was 7-8·10^6 cells/mL. **Figure 4** illustrates the different studied process variables such as specific VLP production, percentage of transfection and total VLP production. Interestingly, this second retransfection had no effect. VCD and cell viability maintained the trend previously observed and no additional changes in transfection and production were observed. As a result, a second retransfection was discarded. This can be explained by the cell density effect (CDE), which shows the difficulty to transfect cell cultures at high cell densities[21]. Indeed, these results confirmed the CDE, which was originally reported to be related to energy depletion[21,23,34]. However, perfusion cultures are fed with a continuous addition of fresh media and removal of metabolic waste, preventing energy depletion. Encountering the CDE in continuous cultures has been already reported [20,35] and different molecular and metabolic factors have been described to influence the CDE[33]. Therefore, using perfusion to achieve very high cell densities, as implemented when working with stable gene expression (SGE) cell-based platforms, is no longer useful for bioprocesses using TGE. Instead, the ATF perfusion system can be used to maintain a transfectable cell density and perform perfusion to maintain the constant media replacement in order to prolong production a defined period of time.
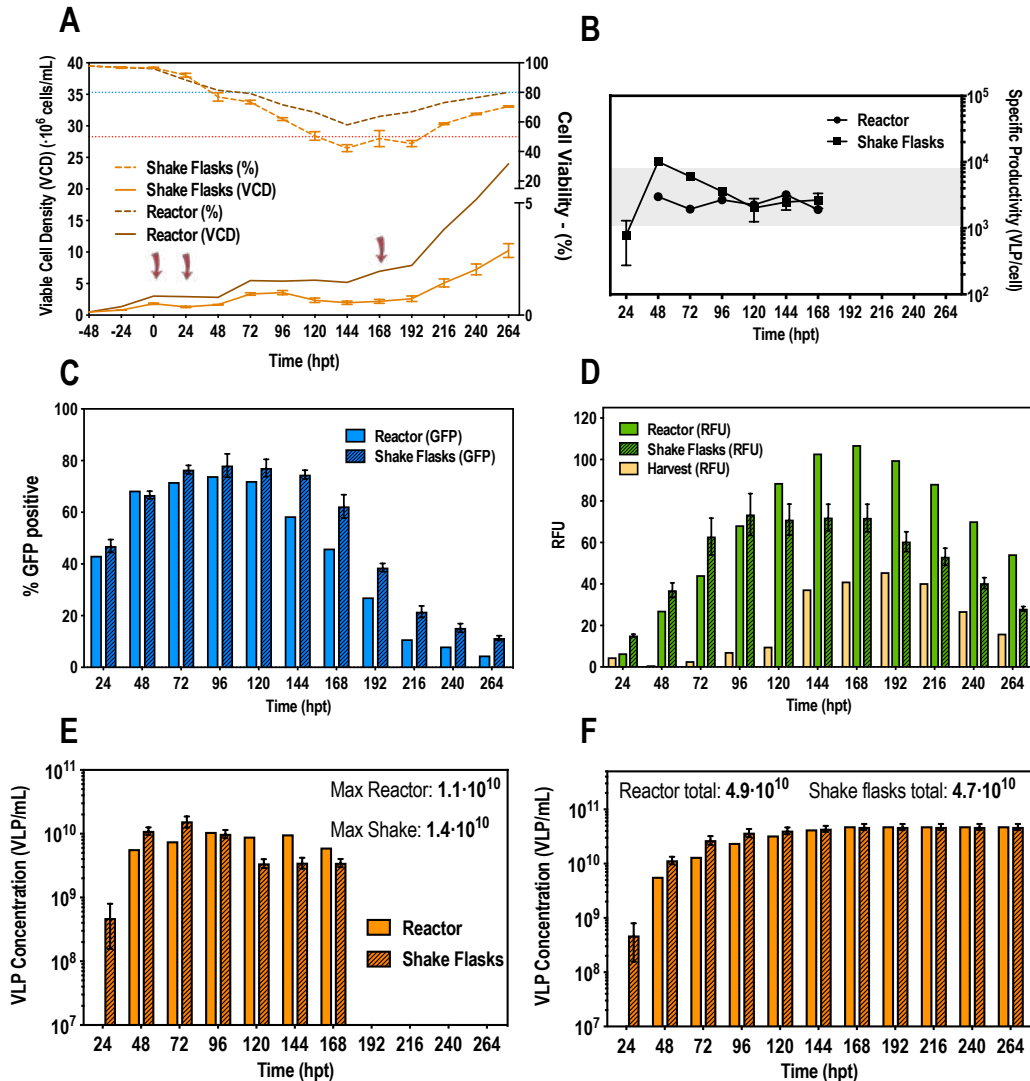
Figure 4. Second retransfection at 168hpt in bioreactor. **A)** Viable cell density and cell viability graphs along the studied time course. The blue red dotted lines indicate 80% and 50% of cell viability respectively. Cells were transfected at time point 0, 24 and 168 hours post-transfection (hpt) noted by red arrows. **B)** VLP specific productivity rate along the studied time course. **C)** Flow cytometry analysis of the transfection percentage (showed in GFP positive percentage of cells) along the time course. **D)** Measurements of fluorescence in relative fluorescent units (RFUs) from the reactor, the harvest from the reactor and the shake flasks. **E)** Total VLP concentration produced per day **F)** Total cumulative VLP concentration.

VLP RETENTION BY THE ATF HOLLOW FIBER MODULE

For the ATF configuration, hollow fiber modules (HFM) with pore sizes of 0.5 and 0.2 μm were used in order to compare their VLP recovery efficiency. Regardless of the HFM, transfection efficiency and total VLP

production remained constant (**Figure 3F-3G**). However, despite having a diameter of 140 nm, VLPs were clearly retained inside the bioreactor. Therefore, a detailed study of VLP retention by the HFM was carried out. In order to characterize the interaction between VLPs and the HFM, confocal and field emission scanning electron microscopy (FESEM) studies were performed. In these analyses for characterization of VLP retention in the ATF system, a sample from fibers exposed to conditioned media free of VLPs from non-transfected HEK293 culture was used as a control. Longitudinal cuts of the lumen of the hollow fibers from the three conditions were visualized under FESEM to evaluate the morphology and presence of nanoparticles deposited on the fiber (**Figure 5**). Detection of secondary electrons (SE) is routinely used to evaluate size, size distribution and topography while detection of backscattered electrons (BSE) reveals differences in chemical composition and allows the detection of nanoparticles[36,37] and substrates with different composition deposited on the fiber surface. Contrast of BSE for imaging has been widely used for metal nanoparticles but it can also be used to detect fluorophores or molecules with electron excitation states whose electronic rearrangement after being exited under the electron beam, behave like a heavier atom. BSE originate from the elastic scattering of primary electrons upon interaction with the nucleus of the atoms in the sample. Therefore, atoms with heavier nuclei cause stronger scattering, providing a more intense signal for the BSE detector. Fluorescent molecules produce a different BSE signal compared to the non-fluorescent background[38–41]. The SEM analyses of the different fiber samples revealed that there is a deposit of particles in the fibers exposed to VLPs. The morphology of polyethersulfone (PES) fibers showed no difference but the BSE analyses showed clusters of particles of different chemical composition and a higher BSE intensity compared to the inert PES. In order to confirm the retention and deposit of VLPs in both 0.5 and 0.2 HFM, transversal and longitudinal cuts were also visualized using confocal microscopy (**Figure 6**). The transversal cuts revealed that fluorescent deposits appeared in the luminal space of the fibers in the conditions exposed to VLPs. Interestingly, the longitudinal cuts analysis showed fluorescent deposits following the same clustering pattern previously detected using FESEM. Coherently, in both transversal and longitudinal cuts, the detected fluorescence intensity increased in the 0.2 μm HFM due to smaller pore and higher retention. Therefore, it is proved that VLPs are being retained in the HFM forming
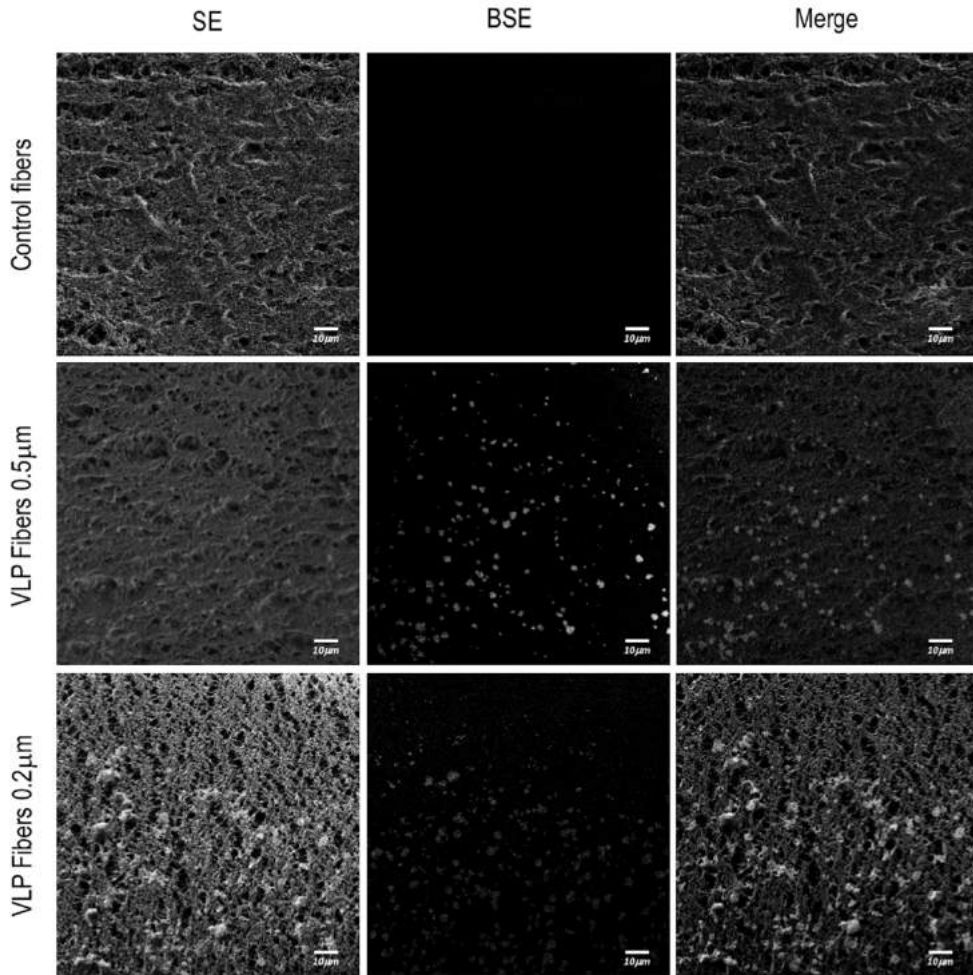
Figure 5. Visualization of a ATF hollow fiber inner side from a longitudinal cut using field emission scanning electron microscopy (FESEM). Rows represent the different studied conditions: control fibers used in non-producing-VLPs HEK293 cultures and fibers used in VLP-producing HEK293 cultures of 0.5 and 0.2 mm of pore diameter. Images were taken at 3.4 mm of working distance and 129,000x. The first column shows the morphology of the fibers using scanning of secondary electron (SE) detection. The second column shows back-scattered electron (BSE) detection for fluorescence detection from the same area. The third column shows the merge of the SEM image and the electron back scattering acquisition.

clusters of particle aggregates of approximately 2.5 μm and 5μm of diameter in the 0.5 and 0.2 μm HFM respectively, with consistent results via confocal microscopy and SEM. Nevertheless, the amount of deposited VLP was negligible as VLP production in 0.2 and 0.5 conditions were equivalent (**Figure 3G**). The use of HFM has been previously reported to lead to particle retention. Especially when manufacturing viral particles, as hollow fibers are used to concentrate VLPs[42]. Significant retention of viruses has been
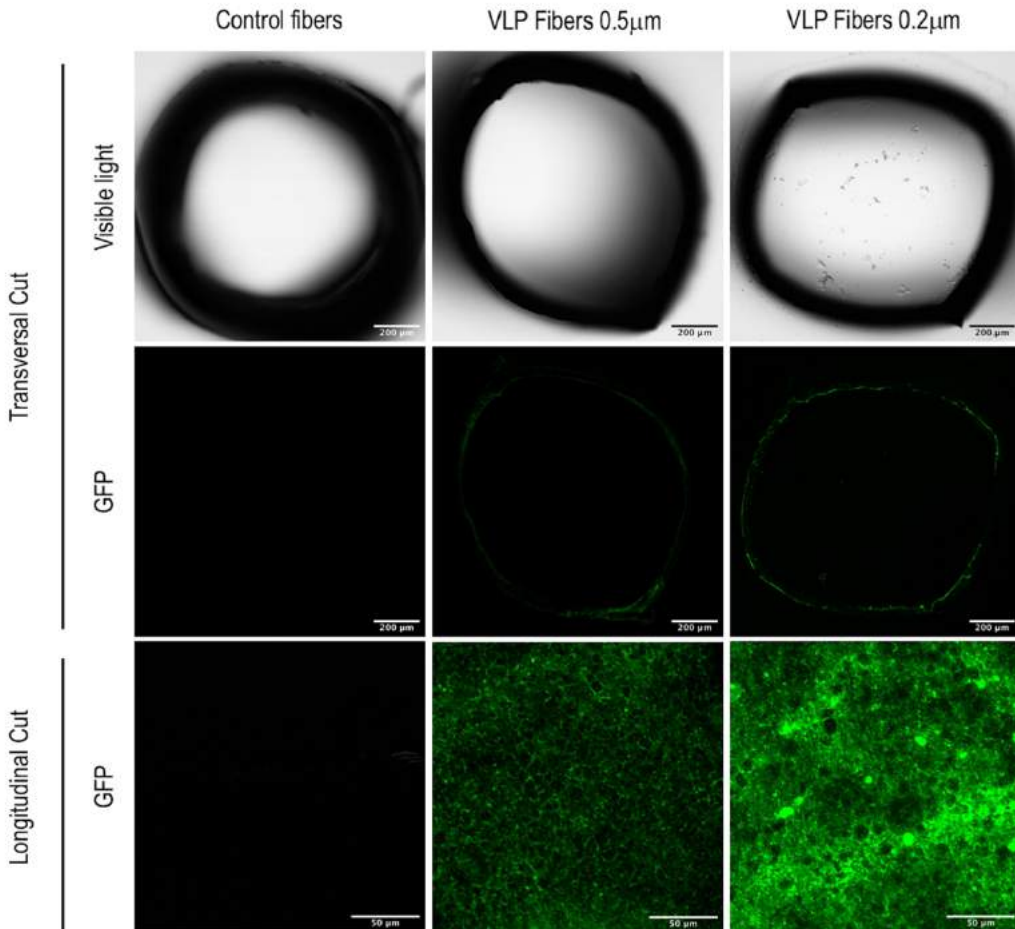
Figure 6. Visualization of ATF hollow fibers using confocal microscopy. Columns represent the different studied conditions: control fibers used in non-producing-VLPs HEK293 cultures and fibers used in VLP-producing HEK293 cultures of 0.5 and 0.2 mm of pore diameter. The first row shows transversal cuts of the fibers under visible light. The second row shows GFP fluorescence of the same transversal cut. The third row shows GFP fluorescence from a longitudinal cut of the luminal space.

reported when using ATF as the retention device for production of viral particles. In the case of viruses like Influenza A which are about 80-100 nm[43,44], modified vaccinia Ankara (MVA) virus[45] or even the production of yellow fever and Zika virus, both about 50 nm, recoveries < 1% of the virus titer in the harvest fraction have been reported[46]. Therefore, VLP retention observed in this work agrees with the already described results of viral particle retention.

QUANTIFICATION OF VLP RETENTION

NTA analyses of samples from the bioreactor and from the harvest fraction were performed at different days to evaluate VLP quality and percentage of fluorescent particles in the different fractions and to determine the percentage of retention (**Figure 7A**). The size distribution curve of the fluorescent particles in the bioreactor, regardless of the HFM used, described particles of 140 nm and was maintained along the whole bioprocess, confirming the stability of the particles. More VLPs were retained in the 0.2 compared to the 0.5 μm HFM along the process. The measurement of total particles revealed that, as expected, the 0.2 μm HMF was retaining a greater total amount of particles inside the bioreactor compared to the 0.5 μm HFM. However, 12-15% of fluorescent particles was observed in the bioreactor in both cases. Surprisingly, when measuring the fluorescent particles found in the harvest fractions, the concentration found was significantly lower than expected. The size distribution curve showed an heterogeneous population of particles, representing only the 0.08-0.1% of the total particles found in the harvest, in both HFM used. Calculations for both HFM revealed a 99% of VLP retention inside the bioreactor.

Comparing the value of VLP concentration obtained by the fluorimetry-based and NTA methods of quantification, ratios reflecting the presence of extracellular unassembled free fluorescent Gag monomers can be assessed. Samples with a high concentration of free monomer will provide a high fluorescence intensity signal while presenting a low particle concentration by NTA. Inside de bioreactor, the observed ratio VLP:monomer is around 1:4 while in the harvest fraction can reach up to 1:1762. These ratios are showed in **Figure 7B**. The free monomer ratio observed in the bioreactor did not depend on the HFM, as the difference found comparing 0.5 and 0.2 μm HFM was not significant. On the other hand, the ratio found in the harvest was around 10^3 and significantly changed depending on the HFM used. The bigger pore of the 0.5 μm HFM facilitated the filtration of free monomer together with VLPs than in comparison to the 0.2 μm HFM, reducing the ratio VLP/free monomer to 1:816 while in the 0.2μm HFM the ratio significantly increased to 1:1762 due to the separation of VLP and free monomer via its filtration through the fiber. This can explain the low percentage of fluorescent particles found in the harvest and can be used for the direct separation of unassembled free Gag monomer, providing substantial bioprocessing advantages as discussed further.
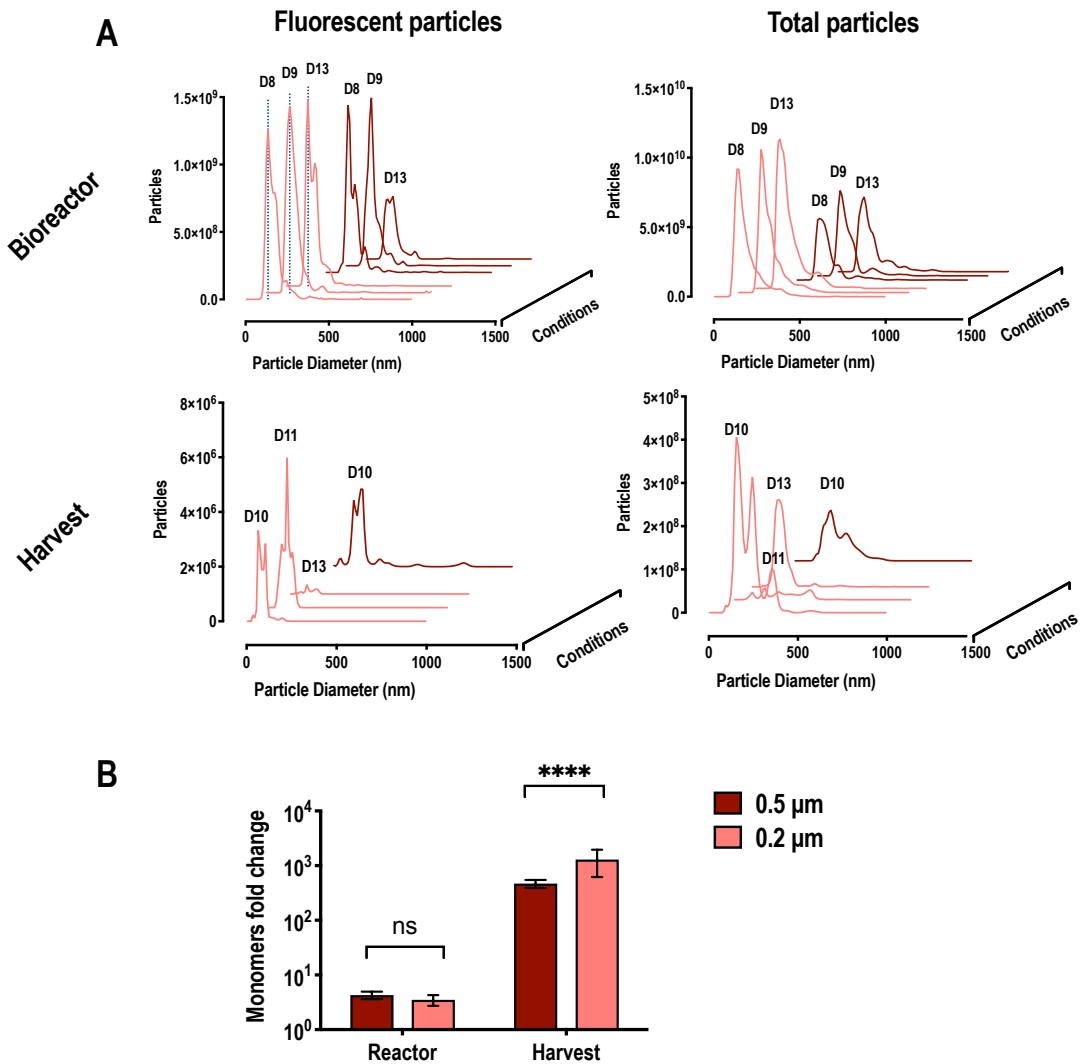
Figure 7. A) Nano particle tracking analysis (NTA) from cultures where hollow fiber modules (HFM) with pore size of 0.5 and 0.2 mm were used. The left columns show the particle size distribution of the detected fluorescent particles (VLPs) along different days of the culture, in the reactor and harvest. The right column shows the particle size distribution of all particles detected (VLPs and extracellular vesicles) along different days in the reactor and harvest. Dotted blue lines indicate particle size of 140 nm. D labels indicate the time (day of the bioprocess) **B)** Fold change ratio between the estimated particle concentration using fluorimetry and the detected actual particle concentration using NTA. This fold change suggests the presence of unassembled free Gag monomers. Significance calculated using two-way ANOVA, DF= 6 and n=3 in each condition.

SELECTION OF CONDITIONS FOR A VLP PRODUCTION PLATFORM

Considering the results obtained in the previous sections, the conditions for the implementation of the intensified and optimized EGE method at bioreactor scale using an ATF system for perfusion could be defined. The use of 0.2 μm HFM provided the most advantages for the development of a VLP-based production platform. The retention of the product within the reactor allowed the one-step production and concentration of VLPs, achieving 12-15% VLPs, compared to the total of extracellular vesicles, increasing from the previously reported 1%[20]. Thus, eliminating the problem of having the total amount of VLPs diluted in the harvest fraction, increasing media volumetric productivity 67.8% or 3.1-fold. The problem of having the product retained might concern VLP stability but particles were proved stable at 37°C throughout the process, overcoming the stability issue. This might be due to the nature of enveloped particles. Biological membrane-bound extracellular vesicles, like exosomes and microvesicles, are stable at physiological conditions such as temperature and pH[47,48], equivalent to the bioreactor conditions. In addition to this, the 0.2 μm HFM allows the separation of unassembled free Gag monomer in the same step. Hence, improving the consecutive downstream processes not only by increasing the initial VLP concentration regarding the total number of particles, such as exosomes or microvesicles, but also regarding the quality, separating VLPs from free extracellular monomers. Therefore, overcoming challenges typically associated with VLP purification. However, the CDE prevented production from being extended in time, not being able to prolong gene expression further than 168hpt, when cells ceased producing. However, product retention in the bioreactor is forcing to stop the process at a harvest time point and to be restarted, as in batch culture mode. In order to operate continuously and not being restricted to the batch process limitations, a continuous approach is proposed. Future work will be performed to demonstrate the validity of such an approach. When reaching the VLP harvest point at 168hpt, a bleeding step might be added to harvest most of the culture and maintain the corresponding number of cells to seed the bioreactor again at $0.5 \cdot 10^6$ cells/mL with fresh media and to recover the initial state, restarting the optimized EGE transfection process.

This novel perfusion approach is different to the conventional use of perfusion for achieving high cell densities and maintaining steady state through purge or bleeding. Here, continuous perfusion is used to

keep a constant media replacement while bleeding would be used to recover the product every 168hpt and continue the process. Although the recovered culture should undergo a clarification process to separate the cell mass from the VLPs in the supernatant, this method would allow the acquirement of higher volumetric productivities while maintaining the constant media replacement. Thus, overcoming the batch and fed-batch limitations, the implementation of Sequential Perfusion recovery (SPR) is the proposed mode of operation. Conventional perfusion processes using ATF has been reported to successfully run for 50-80 days[49-51]. Therefore, the ATF membrane integrity should not be compromised for several 168-hours cycles, depending on the desired VLP titer.

Table II. Optimized and intensified process performance

Accumulated VLPs	8.7E+13
Max. VLP concentration in a day (VLP/mL)	1.5E+10
VLP concentration in reactor (VLP/mL)	6.8E+10
Mean VLP specific productiviy (VLP·cell^{-1}·day^{-1})	3.0E+03
Volumetric productivity / Reactor (VLP·L^{-1}·day^{-1})	7.1E+12
Volumetric productivity / Spent medium (VLP·L^{-1}·day^{-1})	2.7E+12
Total volume of spent media (L)	3.8

CONCLUSIONS

The implementation of the extended gene expression protocol has been achieved at bioreactor scale in perfusion mode circumventing limitations previously observed, such as a unfavorable effect on the purity of the product. In this work, an approach based on the design of experiments method was applied to optimize the retransfection time, the CSPR and the DNA concentration. The optimal solution for the variables of retransfection at 24hpt, a CSPR of 30pL/cell/day and a DNA concentration of 1.7 μm g/mL were used to perform the process in a 1.5L stirred tank bioreactor. The same volumetric productivity, specific productivity and VLP concentration were achieved in the parallel shake flask culture and the bioreactor. The process was successfully operated at a 1.5L, reaching a concentration of $6.8 \cdot 10^{10}$ VLP/mL and a total cumulative particle amount of $8.7 \cdot 10^{13}$ VLPs in the bioreactor, improving the original EGE protocol 86.7% or 7.54-fold and the first bioreactor approach using acoustic filter 67.8% or 3.1-fold. A summary of the process variables is presented in **Table 2**. Hollow fiber modules of 0.5 and 0.2 μm of pore size were tested when implementing the ATF perfusion system. A 99% of VLP retention in the bioreactor was observed in both alternatives, confirmed by FESEM, confocal microscopy and NTA analyses. Bearing this in mind, 168hpt proved to be the optimal VLP harvesting time from the bioreactor. Interestingly, the 0.2 μm HFM provided the advantage of separating unassembled Gag monomers from correctly assembled VLPs and concentrating them in the bioreactor, facilitating the subsequent downstream processing. The use of the ATF system also allowed the intensification of the process, increasing production while reducing spent media. Although there is no previous work in VLP production by TGE using an ATF system for perfusion, the proposed platform reduce spent media from 9-10 to 3.8L and therefore resources and costs compared to the only additional reported study regarding TGE at bioreactor scale[52]. Due to product retention, a proposed procedure of controlled cell bleeding at 168hpt to harvest and restart the culture at an initial cell density of $0.5 \cdot 10^6$ is proposed under the name of Sequential Perfusion Recovery (SPR). The cycle could be repeated as many times as needed to achieve the desired titer. This strategy is an intensified and optimized VLP production platform that could encourage the potential further development VLP-based therapies, potentially applicable to several diseases.

REFERENCES

1.	Whiteside, A. & Wilson, D. Health and AIDS in 2019 and beyond. *African J. AIDS Res.* **17**, iii–v (2018).

2.	Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

3.	Pincetic, A. & Leis, J. The Mechanism of Budding of Retroviruses From Cell Membranes. *Adv Virol* **2009**, 6239691–6239699 (2009).

4.	von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

5.	Patters, B. J. & Kumar, S. The role of exosomal transport of viral agents in persistent HIV pathogenesis. *Retrovirology* **15**, 79 (2018).

6.	Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136–144 (2013).

7.	Jolly, C. & Sattentau, Q. J. Human Immunodeficiency Virus Type 1 Assembly, Budding, and Cell-Cell Spread in T Cells Take Place in Tetraspanin-Enriched Plasma Membrane Domains. *J. Virol.* **81**, 7873–7884 (2007).

8.	Cervera, L. *et al.* Production of HIV-1-based virus-like particles for vaccination: achievements and limits. *Appl. Microbiol. Biotechnol.* **103**, 7367–7384 (2019).

9.	Lua, L. H. *et al.* Bioengineering virus-like particles as vaccines. *Biotechnol Bioeng* **111**, 425–440 (2014).

10.	Meador, L. R. *et al.* A heterologous prime-boosting strategy with replicating Vaccinia virus vectors and plant-produced HIV-1 Gag/dgp41 virus-like particles. *Virology* **507**, 242–256 (2017).

11.	Sharma, J., Uchida, M., Miettinen, H. M. & Douglas, T. Modular interior loading and exterior decoration of a virus-like particle. *Nanoscale* **9**, 10420–10430 (2017).

12.	Kueng, H. J. *et al.* General strategy for decoration of enveloped viruses with functionally active lipid-modified cytokines. *J. Virol.* **81**, 8666–76 (2007).

13.	Kueng, H. J., Schmetterer, K. G. & Pickl, W. F. Lipid rafts, pseudotyping, and virus-like particles:

Relevance of a novel, configurable, and modular antigen-presenting platform. *International Archives of Allergy and Immunology* **154**, 89–110 (2011).

14. Venereo-Sanchez, A. *et al.* Process intensification for high yield production of influenza H1N1 Gag virus-like particles using an inducible HEK-293 stable cell line. *Vaccine* **35**, 4220–4228 (2017).

15. Bandaranayake, A. D. & Almo, S. C. Recent advances in mammalian protein production. *FEBS Lett* **588**, 253–260 (2014).

16. Gutiérrez-Granados, S., Cervera, L., Kamen, A. A. & Gòdia, F. Advancements in mammalian cell transient gene expression (TGE) technology for accelerated production of biologics. *Crit. Rev. Biotechnol.* **38**, 918–940 (2018).

17. Cervera, L., Gutiérrez-Granados, S., Berrow, N. S., Segura, M. M. & Gòdia, F. Extended gene expression by medium exchange and repeated transient transfection for recombinant protein production enhancement. *Biotechnol Bioeng* **112**, 934–946 (2015).

18. Butler, M. Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. *Appl Microbiol Biotechnol* **68**, 283–291 (2005).

19. Gutiérrez-Granados, S., Gòdia, F. & Cervera, L. Continuous manufacturing of viral particles. *Curr. Opin. Chem. Eng.* **22**, 107–114 (2018).

20. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).

21. Petiot, E., Cuperlovic-Culf, M., Shen, C. F. & Kamen, A. Influence of HEK293 metabolism on the production of viral vectors and vaccine. *Vaccine* **33**, 5974–5981 (2015).

22. Le Ru, A. *et al.* Scalable production of influenza virus in HEK-293 cells for efficient vaccine manufacturing. *Vaccine* **28**, 3661–3671 (2010).

23. Bernal, V., Carinhas, N., Yokomizo, A. Y., Carrondo, M. J. & Alves, P. M. Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism. *Biotechnol Bioeng* **104**, 162–180 (2009).

24. Buffo, M. M. *et al.* Influence of dual-impeller type and configuration on oxygen transfer, power

consumption, and shear rate in a stirred tank bioreactor. *Biochem. Eng. J.* **114**, 130–139 (2016).

25. Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

26. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

27. Montgomery, D. C. *Design and analysis of experiments*.

28. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

29. Fuenmayor, J., Cervera, L., Gutierrez-Granados, S. & Godia, F. Transient gene expression optimization and expression vector comparison to improve HIV-1 VLP production in HEK293 cell lines. *Appl Microbiol Biotechnol* **102**, 165–174 (2018).

30. González-Domínguez, I., Grimaldi, N., Cervera, L., Ventosa, N. & Gòdia, F. Impact of physicochemical properties of DNA/PEI complexes on transient transfection of mammalian cells. *N. Biotechnol.* **49**, 88–97 (2019).

31. González-Domínguez, I., Cervera, L., Gòdia, F. & Roldán, M. Quantitative colocalization analysis of DNA delivery by PEI-mediated cationic polymers in mammalian cells. *J. Microsc.* **273**, 53–64 (2019).

32. Gutiérrez-Granados, S., Cervera, L., Segura, M. . L., Wölfel, J. & Gòdia, F. Optimized production of HIV-1 virus-like particles by transient transfection in CAP-T cells. *Appl Microbiol Biotechnol* **100**, 3935–3947 (2016).

33. Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed quantitative proteomic analysis of HEK293 provides insights of molecular changes associated to the cell density effect, transient transfection and virus-like particles production. *J. Proteome Res.* acs.jproteome.9b00601 (2020). doi:10.1021/acs.jproteome.9b00601

34. Bereiter-Hahn, J., Munnich, A. & Woiteneck, P. Dependence of energy metabolism on the density of cells in culture. *Cell Struct Funct* **23**, 85–93 (1998).

35. Genzel, Y. *et al.* High cell density cultivations by alternating tangential flow (ATF) perfusion for influenza A virus production using suspension cells. *Vaccine* **32**, 2770–2781 (2014).

36. Österreicher, J. A., Grabner, F., Schiffl, A., Schwarz, S. & Bourret, G. R. Information depth in backscattered electron microscopy of nanoparticles within a solid matrix. *Mater. Charact.* **138**, 145–153 (2018).

37. Kowoll, T. *et al.* Contrast of backscattered electron SEM images of nanoparticles on substrates with complex structure. *Scanning* **2017**, (2017).

38. Vancová, M. & Nebesářová, J. Correlative fluorescence and scanning electron microscopy of labelled core fucosylated glycans using cryosections mounted on carbon-patterned glass slides. *PLoS One* **10**, (2015).

39. Fokkema, J. *et al.* Fluorescently Labelled Silica Coated Gold Nanoparticles as Fiducial Markers for Correlative Light and Electron Microscopy. *Sci. Rep.* **8**, (2018).

40. Garming, M. W. H. *et al.* Nanoparticle discrimination based on wavelength and lifetime-multiplexed cathodoluminescence microscopy. *Nanoscale* **9**, 12727–12734 (2017).

41. Seras-Franzoso, J., Sánchez-Chardi, A., Garcia-Fruitós, E., Vázquez, E. & Villaverde, A. Cellular uptake and intracellular fate of protein releasing bacterial amyloids in mammalian cells. *Soft Matter* **12**, 3451–3460 (2016).

42. Negrete, A., Pai, A. & Shiloach, J. Use of hollow fiber tangential flow filtration for the recovery and concentration of HIV virus-like particles produced in insect cells. *J. Virol. Methods* **195**, 240–246 (2014).

43. Coronel, J. *et al.* Influenza A virus production in a single-use orbital shaken bioreactor with ATF or TFF perfusion systems. *Vaccine* **37**, 7011–7018 (2019).

44. Genzel, Y. *et al.* High cell density cultivations by alternating tangential flow (ATF) perfusion for influenza A virus production using suspension cells. *Vaccine* **32**, 2770–2781 (2014).

45. Vázquez-Ramírez, D., Genzel, Y., Jordan, I., Sandig, V. & Reichl, U. High-cell-density cultivations

to increase MVA virus production. *Vaccine* **36**, 3124–3133 (2018).

46. Nikolay, A., Léon, A., Schwamborn, K., Genzel, Y. & Reichl, U. Process intensification of EB66® cell cultivations leads to high-yield yellow fever and Zika virus production. *Appl. Microbiol. Biotechnol.* **102**, 8725–8737 (2018).

47. Madison, M. N. & Okeoma, C. M. Exosomes: Implications in HIV-1 pathogenesis. *Viruses* **7**, 4093–4118 (2015).

48. Ellwanger, J. H., Veit, T. D. & Chies, J. A. B. Exosomes in HIV infection: A review and critical look. *Infection, Genetics and Evolution* **53**, 146–154 (2017).

49. Clincke, M. F. *et al.* Very high density of Chinese hamster ovary cells in perfusion by alternating tangential flow or tangential flow filtration in WAVE Bioreactor™-part II: Applications for antibody production and cryopreservation. *Biotechnol Prog* **29**, 768–777 (2013).

50. Schwarz, H. *et al.* Small-scale bioreactor supports high density HEK293 cell perfusion culture for the production of recombinant Erythropoietin. *J. Biotechnol.* **309**, 44–52 (2020).

51. Zhang, Y., Stobbe, P., Silvander, C. O. & Chotteau, V. Very high cell density perfusion of CHO cells anchored in a non-woven matrix-based bioreactor. *J Biotechnol* **213**, 28–41 (2015).

52. Hong, J. *et al.* Development of an alternating tangential flow (ATF) perfusion-based transient gene expression (TGE) bioprocess for universal influenza vaccine. *Biotechnol. Prog.* (2019). doi:10.1002/btpr.2831

CHAPTER THREE

Part III.

Bioprocess analytics and perfusion-based
continuous VLP harvest

ABSTRACT

One of the disadvantages of the perfusion process established in Part II was the retention of VLPs within the bioreactor and the impossibility of a continuous harvest. Here, an alternative filtering module, the viral harvesting unit (VHU®) from Artemis Biosystems is evaluated. The same perfusion process based on transient transfection and carrying out the optimized extended gene expression methodology was performed with the only change of incorporating the new cell retention device. This time, VLPs could be recovered in the harvest fraction, moving forward towards the establishment of a conventional perfusion process. Moreover, the bioprocess described in Part II was analyzed by a proteomic study to identify molecular and cellular changes throughout the perfusion process, allowing the characterization of protein variations as transfection states develop in the cell culture.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from McGill University (Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell culture media was HyCell™ TransFx-H media from HyClone™ (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco, Life Technologies).

Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

BIOREACTOR CULTURE CONDITIONS AND SETUP DESCRIPTION

A BioStat B Plus bioreactor (Sartorius Stedim Biotech, Goettingen, Germany) equipped with a 3-blade segment dual impeller with UP-DP configuration[1] was used for HEK293 cell cultivation. The agitation was set at 200 rpm. Temperature was set at 37 °C. The pH was set at 7.1 and controlled with CO_2 and $NaHCO_3$ (7.5%w/v). The dissolved oxygen was controlled at 40% of air saturation by supplementing air by sparger at a constant flow of 0.1L/min and additional pure oxygen when needed. HEK293 cells growing exponentially in disposable polycarbonate 1 L shake flasks (Corning®) were used to seed the bioreactor at $0.5 \cdot 10^6$ cells/mL in 1.5 L. Samples were taken every 24 h for cell counting and viability determination. After transfection, VLP quantification was performed every 24h. Perfusion was performed using an alternating tangential flow (ATF) and a viral harvesting unit (VHU2) as cell retention device (Artemis Biosystems, Quincy, MA, USA) with 0.05 m^2 of filtration area and a ATF flow rate of 1.0 L/min. When performing transfection, perfusion flow was stopped during 2 hours, to allow the contact of cells with DNA/PEI complexes and then reestablished after transfection[2]. To carry out media replacement, the filtration rate was set at 0.26 mL/min at the beginning of the process using a MasterFlex L/S peristaltic pump (MasterFlex
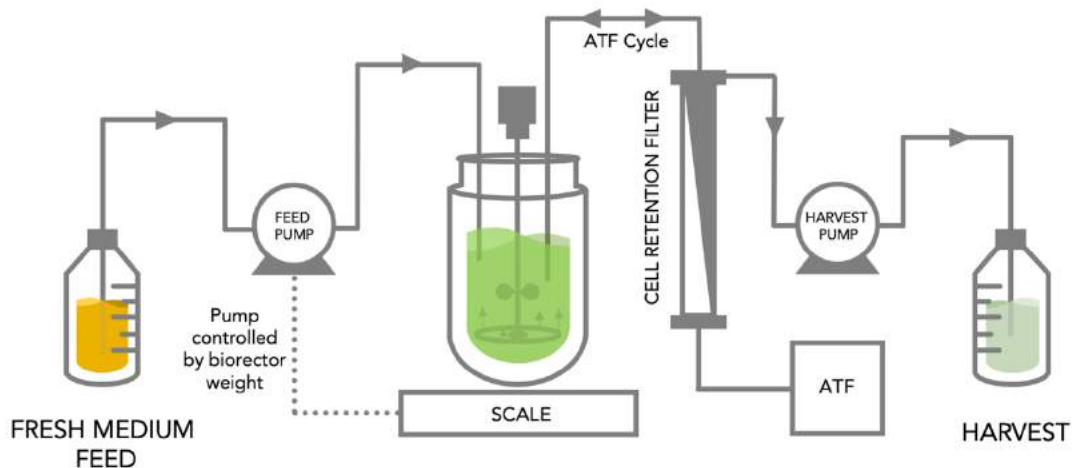
Figure 1. Schematic bioprocess diagram for the performed perfusion-based continuous VLP harvest. The harvest pump was set at a constant rate of 0.26 mL/min at the beginning of the process and modified every day depending on the viable cell density to maintain a cell specific perfusion rate of 30 pL/cell/day. Constant filtration created a mass displacement that was registered by the scale, activating the feed pump to add fresh culture media to the bioreactor. ATF flow rate was maintained constant during the whole process at 0.6L/min for XCell ATF2 hollow fiber module and at 1.0L/min for the VHU2 filter.

Group, Gelsenkirchen, Germany) and was modified every day depending on the viable cell density to maintain a cell specific perfusion rate (CSPR) of 30 pL/cell/day. The bioreactor and the ATF system were placed over a scale (MSE36200S-000-D0 Cubis, Sartorius, Goettingen, Germany) to measure the change in mass created due to the filtration of spent media. The scale was connected to a Scilog peristaltic pump (Scilog tandem 1081, Parker, Oxnard, CA, USA) which controlled the addition of fresh media upon detection of mass decrease, in order to maintain constant volume in the reactor (**Figure 1**).

TRANSIENT TRANSFECTION

The initial transfection (t_1) was carried out at a cell density of $2·10^6$ cells/mL using a final DNA concentration of 1 µg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used. The plasmid used contained the gene coding for HIV-Gag protein fused to eGFP (Gag::eGFP). Briefly, the corresponding DNA was diluted with supplemented HyCell™ media and vortexed for 10 seconds. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, then the mixture was incubated for 15 min at room temperature and then added to the cell culture.

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using a developed and validated quantification assay[3]. VLP containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected culture supernatants.

HIV-1 GAG VLP QUANTIFICATION BY NANOPARTICLE TRACKING ANALYSIS (NTA)

Nanoparticle Tracking Analysis (NTA) was also used to quantify fluorescent particles. NTA measurements were performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag::eGFP VLPs and neutral density filter for total particle by light scattering. Data was analyzed with NanoSight® NTA 3.1 software. Briefly, samples were injected, and independent analyses were carried out. Three video recordings of 60sec duration were taken for each sample. Capture settings were recorded with an sCMOS camera (camera level of 8 for Gag::eGFP VLP samples, and 11 for controls, viscosity: 0.9 cP) and analyzed with a detection threshold of 4.

PROTEIN SAMPLES PREPARATION FOR MASS SPECTROMETRY ANALYSES

Pellets of each tested condition were obtained by centrifugation of 500 μL of the cell culture at 10000xg for 15 minutes at 4°C and stored at −80°C.

Protein extraction was performed using extraction buffer (100 mM Tris-HCl pH=8.8, 2mM EDTA, 4% SDS, 100mM DTT) of which 100μL were added to the cell pellet of each condition. Samples were sonicated for 5 minutes and then boiled for another 5 minutes. Protein extracts from pellet samples were quantified with RC/DC Protein Assay (Bio-Rad, Hercules, CA, USA) and stored in -20 °C until the tryptic digestion process. Protein digestion was performed as previously described[4]. Proteins were digested using sequencing grade trypsin (Promega, Madison, WI, USA) and the filter-assisted sample preparation technology (FASP, Expedeon, San Diego, CA, USA), and the resulting peptides were subjected to TMT-10 plex labelling (AB Sciex, Framingham, MA, USA), joined and desalted. A total of 150μg from samples of each condition was diluted to a final concentration of 100mM of TEAB labeled with TMT-10 plex according

to manufacturer's instructions. Protein samples were labeled by adding 41µL of TMT isobaric tag diluted in anhydrous acetonitrile, followed by a 1h-incubation step at room temperature. To quench the reaction, 5% hydroxyl-amine (8µL per sample) was added, incubated 15 min at room temperature and mixed together followed by addition of TFA 1% to lower pH at 3. TMT-labeled samples were equally mixed. Pooled mix was purified using Oasis HLB C18 column (Waters, MA, USA).

TMT-labeled mix was fractionated using High pH reversed-phase peptide fractionation kit (Thermo Scientific, San Jose, CA, USA) according to manufacturer's instructions into 5 fractions for further LC-MS/MS analysis.

LIQUID CHROMATOGRAPHY TANDEM MASS SPECTROMETRY ANALYSIS

The tryptic peptide mixtures were subjected to LC-MS/MS analysis on a nano-HPLC Easy nLC 1000 liquid chromatograph (Thermo Scientific, San Jose, CA, USA) coupled to a QExactive mass spectrometer (Thermo Scientific, San Jose, CA). Peptides were suspended in 0.1% formic acid, loaded onto a C18 reverse-phase trapping column (Acclaim PepMap100, 75-µm internal diameter, 3-µm particle size and 2-cm length, Thermo Scientific), and separated on an analytical C18 nano-column (EASY-Spray column PepMap RSLC C18, 75-µm internal diameter, 3-mm particle size and 50-cm length, Thermo Scientific), in a continuous gradient (8–31%B in 240 min, 31–90%B in 2min, 90%B in 7 min, and 2%B in 30min; where buffer A is 0.1% formic acid in HPLC grade H_2O and buffer B is 100% ACN, 0.1% formic acid in HPLC grade H_2O). Spectra were acquired using full ion scan mode over the mass-to-charge (*m/z*) range 390–1500, 70,000 FT-resolution was performed on the top 15 ions in each full MS scan using the data-dependent acquisition mode with 45s dynamic exclusion enabled.

PROTEIN IDENTIFICATION AND QUANTIFICATION

Protein identification was performed over the raw files using the SEQUEST HT algorithm integrated in the Proteome Discoverer 2.1 (Thermo Finnigan). MS/MS scans were matched against a human database (UniProtKB/Swiss-Prot 2017_10 Release). Sequence of Gag::eGFP protein was added to the selected database to enable identification.

For database searching, parameters were selected as follows: trypsin digestion with 2 maximum missed cleavage allowed, precursor mass tolerance of 800 ppm, fragment mass tolerance of 0.02 Da. TMT-10 plex labeling at N-terminal and lysine (+229.62932 Da) as well as cysteine carbamidomethylation (+57.021 Da) were chosen as static modifications whereas methionine oxidation (+15.994915 Da) was chosen as dynamic modification. The same MS/MS spectra collections were searched against inverted database constructed from the same target database. SEQUEST results were analyzed by the probability ratio method[5]. False discovery rate (FDR) for identified peptides was calculated in the inverted database search results using the refined method[6]. Quantitative information of TMT reporter ions was extracted from MS/MS spectra for relative quantification of protein abundance to characterize dynamic protein expression profiles in the selected conditions.

STATISTICAL ANALYSIS

For the comparative analysis of the protein abundance changes, the Weighted Scan- Peptide-Protein (WSPP) statistical workflow was applied[7,37], using SanXoT package[9]. It uses as an input a list of quantifications in the form of log2-ratios (for example a condition versus a control sample) with their statistical weights and generates the standardized forms of the original variables computing the quantitative values expressed in units of standard deviation around the averages. The quantitative information is obtained from the spectra and used to quantify the peptides from which the spectra are produced and then, proteins that generate these peptides. In other words, the quantitative information is integrated from the spectrum level to the peptide level and then from the peptide level to the protein level[10]. These standardized variables, (Zq), express the quantitative values in units of standard deviation[11]. For the protein functional analysis, Systems Biology Triangle (SBT) model[12] was used. This algorithm estimates weighted functional category averages (Zc) from the protein values by performing the protein to category integration. After the integration from spectra to peptide, the integration from peptide to protein and protein to category represent a higher level. This integration allows the detection of changes in functional categories produced by the coordinated behavior of their proteins[10]. The quantified proteins were functionally annotated using the Gene Ontology database[13,14]. For further Gene Ontology annotation,

DAVID[15,16] was used to perform functional enrichment analysis and to extract *p-values* for the enriched processes. To help analyze and comprehend the data, the online software for reactions, proteins and pathways analysis REACTOME[17] was used.

EXPERIMENTAL DESIGN

For the multiplexed quantitative proteomics experiment based on TMT-10 plex labeling, three time points were compared and tested: 24hpt, 96hpt and 168hpt, from a perfusion process. The TMT-based isobaric labelling quantification was performed with three biological replicates for each condition, as it is depicted in **Figure 2A.** The three perfusion processes were carried out using XCell ATF2 hollow fiber membranes with 0.2µm. Samples for the TMT-based experiments were taken from the triplicate bioprocess shown in **Chapter three Part II**. A peptide pool from the 24hpt condition was used as an internal control for the following statistical analyses.

CONFOCAL MICROSCOPY VISUALIZATION

The visualization of the VHU2 module was performed using a FluoView®FV1000 confocal microscope (Olympus, Tokyo, Japan) at sampling speed of 2 µm/pixel, excitation at 488 nm and detection at 500-600 nm. Step size was 2 µm/slice using XYZ scan mode. Longitudinal, transversal and luminal cuts were made from one of the tubular filers of a VHU2 module exposed to VLP-producing HEK293 cell cultures. Objective lens UPLAPO 10x, NA: 0.40 and optical zoom 1x was used.
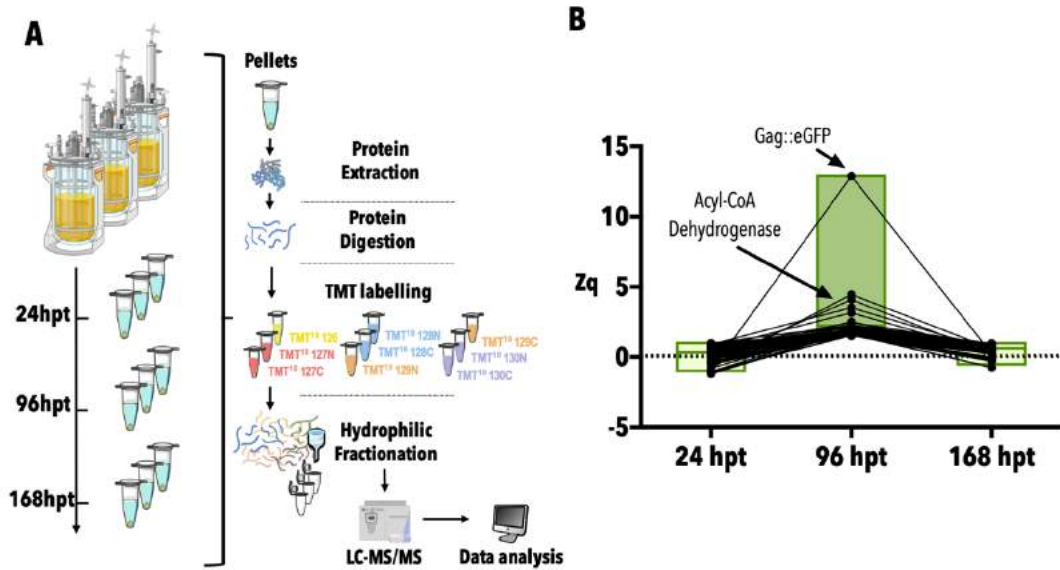
Figure 2. Experimental design and protein upregulation results . **A)** Experimental workflow. Three time points of three biological replicates of HEK293 cells perfusion cultures. Samples from 24 hours post-transfection (hpt), 96hpt and 168hpt from each perfusion culture were compared. Transfections were performed at cell density of $2 \cdot 10^6$ cells/mL. At the corresponding time point, samples were taken, centrifuged and cellular pellets were stored at -80°C. Proteins were extracted from the pellet samples, digested and peptides were labelled using Tandem Mass Tag (TMT) labelling. Labelled peptides were fractionated and analysed via LC-MS/MS. **B)** Results of the upregulated proteins at 96hpt. Gag::eGFP protein reached its maximum upregulation value at 96hpt.

RESULTS AND DISCUSSION

BIOPROCESS PROTEOMIC ANALYTICS

Determining the changes in protein expression throughout the perfusion process, using the XCell ATF2 HFM, can provide valuable information about the metabolic state of the cell culture and the state of the transient transfection. In this multiplexed TMT-based experiment, a total of 5420 proteins were identified and quantified with more than 1 peptide. Most of these proteins did not show any significant up or downregulation in any of the three studied time points (24, 96 and 168hpt). Nevertheless, a group of proteins followed a specific pattern, increasing at 96hpt and returning to a basal level of expression at 168hpt (**Figure 2B**), indicating that within the transient transfection state, there are some changes in protein expression that could provide insight for further optimization. Within this group, the protein showing the highest upregulation was Gag::eGFP, whose production was induced. This protein was

Table I. Upregulated proteins at 96hpt

Accesion number	Protein	Zq
-	Gag::eGFP	12.89
O94929	Actin-binding LIM protein 3	4.45
Q8WTQ7	Rhodopsin kinase GRK7	4.09
Q6PJ61	F-box only protein 46	3.52
P28330	Long-chain specific acyl-CoA dehydrogenase	3.15

identified and quantified with 45 peptides. This agreed with the results presented in **Chapter three Part II**, as the transfection efficiency decreased over time from 96hpt and at 168hpt it reached the same transfection efficiency values than 24hpt (~45% of GFP positive cells). Interestingly, another four proteins followed this pattern (**Table 1**). These four protein markers could be further used as transfection efficiency indicators, if no fluorescent reporter is used. The increase of acyl-CoA dehydrogenase also agreed with previous results in **Chapter one**, indicating the increase in energy demand when Gag::eGFP production peaked. Moreover, F-box protein belongs to the E3 ubiquitin ligase complexes, also indicating the increase in Gag::eGFP synthesis where ubiquitin ligases contribute. These protein markers could be used in the monitoring of a perfusion process for Gag VLP or different biopharmaceuticals production and could also be used as reporters to optimize the phase within transfection where protein synthesis is increased. Effects like lipid biosynthesis downregulation can also be observed at 96hpt in accordance with the results of **Chapter one**. Enzymes like fatty acid synthase or Hydroxymethylglutaryl-CoA synthase (HMGS-1) presented a higher downregulation at 96hpt, although not significant compared to the rest of the time points. Interestingly, at 168hpt, when transfection efficiency is decreasing, cell viability is increasing and the cell culture is being reestablished, the biological process of 'glycosphingolipid metabolic process' with GO:0006687 is significantly increased, compared to the downregulation observed at 24 and 96hpt, when transfection efficiency is increasing. Coherently with the exposed results in **Chapter one**, transfection hindered the correct metabolism of glycosphingolipids. When the cell culture can recover thanks to the continuous addition of fresh media and removal of waste in perfusion mode, the episomal plasmid is lost, transfection efficiency decreases and the disrupted glycosphingolipid metabolism is restored.

PERFUSION-BASED CONTINUOUS VLP HARVEST

The implementation of the EGE protocol at bioreactor scale presented in **Chapter three Part II** introduced the concept of sequential perfusion recovery (SPR), consisting in accumulating the VLPs within the bioreactor up to 168hpt when they can be collected and the culture set to $0.5 \cdot 10^6$ cells/mL to restart the process. However, one of the benefits of implementing a perfusion approach is the possibility of continuous harvest of the product along the bioprocess duration. This alternative was not possible due to the high (>99.9%) retention of VLPs in both hollow fiber module (HFM) tested: $0.5 \mu m$ and $0.2 \mu m$ pore size from XCell®. In this section, a different filter module (VHU2®) has been tested aiming to achieve a continuous harvest of VLPs. The implementation of an efficient continuous upstream production would contribute to the development of a complete integrated bioprocess for HIV-1 Gag VLPs including a less complex downstream processing as the first clarification step would be already integrated in the production phase.

In order to test the VHU2 filter module, the bioprocess presented in **Chapter three Part II** was reproduced only with one change, using VHU2 module instead of XCell ATF2 HFM. Due to time constraints, a limited number of filter modules to test and a technical problem in the run, the process could only be extended up to 96hpt (**Figure 3**). Nevertheless, the data obtained up to this point could be compared to the data obtained (only up to 96hpt) using XCell ATF2 HFM. The cell culture was transfected at $2 \cdot 10^6$ cells/mL and retransfected at 24hpt with the corresponding DNA/PEI complexes amount reported in **Chapter three Part II.** The viability dropped after the two transfection events down to ~65-70% (**Figure 3A**) in agreement with the results obtained using the ATF2 HFM. Up to 96hpt, the presence of VLPs in the harvest fraction was significantly higher than in the ATF2 run, obtaining around 20 RFUs using the VHU2 (**Figure 3B**) compared to an average of 5 RFU with ATF2. Interestingly, when analyzing the VLP content of the bioreactor and harvest fraction up to 96hpt, the ratio of unassembled free monomer vs monomers forming VLPs (calculated as previously described in **Chapter three Part II**) was <1 for the harvest. While the VLP:monomer ratio for the ATF2 run was 1:1762 in the harvest fraction, using the VHU2 module the VLP:monomer ratio is 1:0.5 in the harvest (**Figure 3C**). These promising preliminary results indicate that the VHU2 module might be not only filtering VLPs therefore allowing a continuous harvest, but also
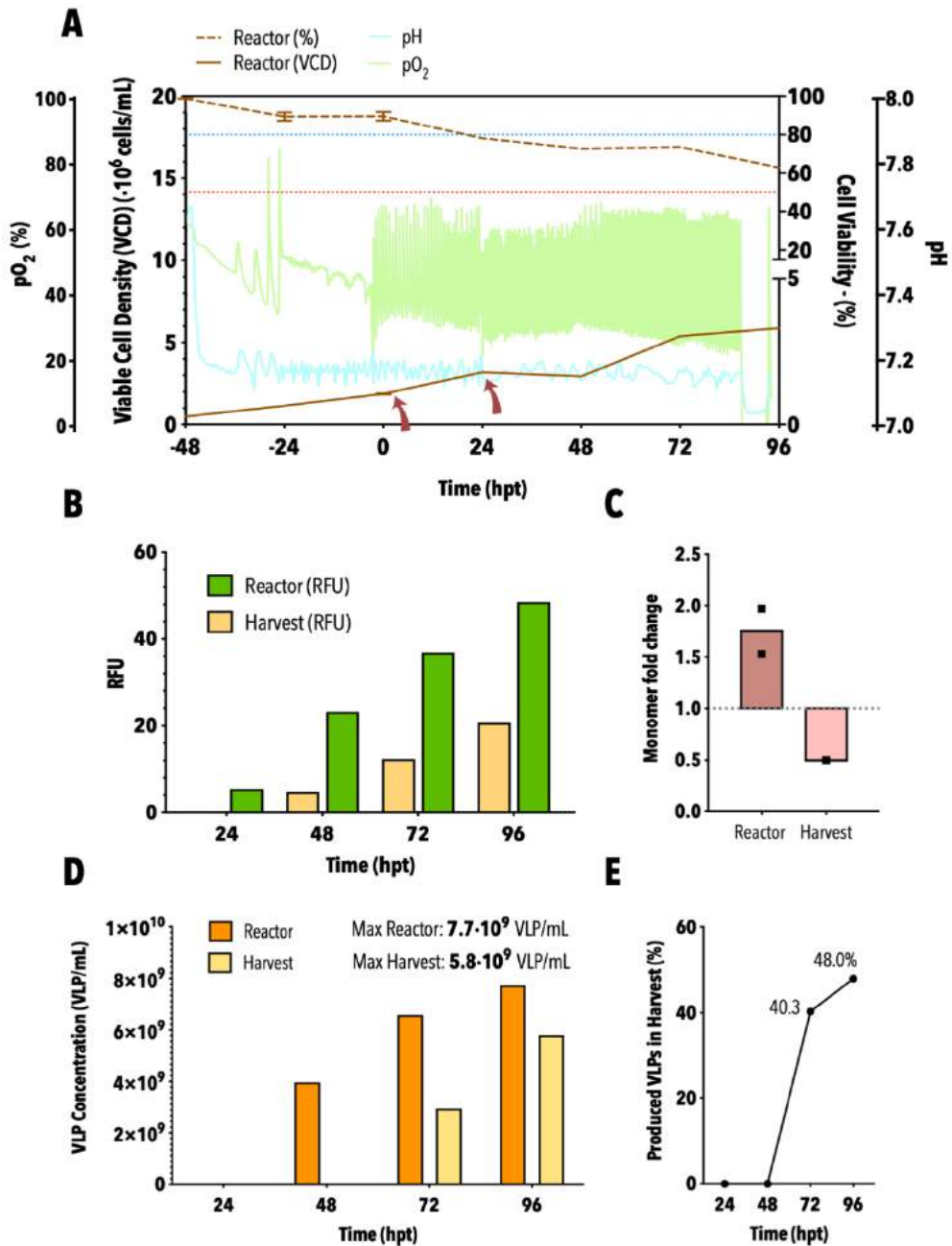
Figure 3. Perfusion process using VHU2 as cell retention device. **A)** Viable cell density (VCD), cell viability, pH and pO_2 graphs along the studied process. The red and blue dotted line represent 50 and 80% of cell viability respectively. Cells were transfected at time point 0h and 24hpt (indicated by red arrows). **B)** Measurements of fluorescence in relative fluorescent units (RFUs) from the reactor and the harvest fraction. **C)** Fold change ratio between the estimated particle concentration using fluorimetry and the detected actual particle concentration using NTA. This fold change suggests the presence of unassembled free Gag monomers. **D)** Total VLP concentration produced per day in the bioreactor and harvest fractions. **E)** Percentage of recovered VLPs in the harvest fraction compared to the total number of VLPs produced each day.
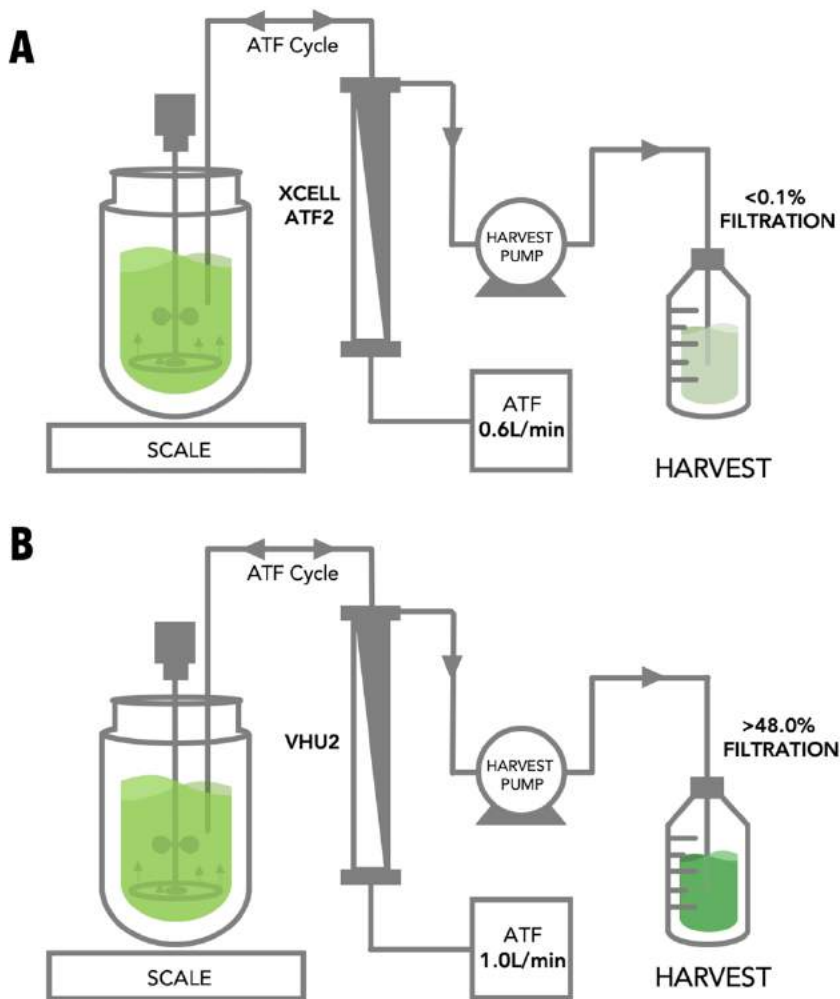
Figure 4. Sequential perfusion recovery and continuous harvest. **A)** Schematic bioprocess diagram for sequential perfusion recovery (SPR) described in Chapter three Part II. The use of XCell ATF2® hollow fiber module retained >99.9% of VLPs and allowed the separation of unassembled monomers. **B)** Schematic bioprocess diagram for the continuous harvesting alternative using VHU2® retention filter module, allowing filtration up to 48.0% of produced VLPs at 96hpt.

reducing the amount of free unassembled monomers in the harvest, possibly due to the nature of the filtration material. This would increase the purity of the harvest fraction and contribute to a better posterior downstream processing. The measured VLP concentration at 96hpt in the bioreactor and the harvest fraction (**Figure 3D**) indicate that VLPs are indeed being filtered and recovered in the harvest. Comparing the total number of VLP produced and the total number of VLPs filtered in the harvest fraction at each time point, the percentage of filtration can be calculated. This percentage was not constant and increased along

Table II. Comparison of the two studied cell retention devices

	ATF2	VHU2
Material	PS, PES	PS, PP
Effective membrane surface area (m^2)	0.13	0.05
Pore size (μm)	0.2 - 0.5	-[a]
Length (cm)	60	60
Diameter (cm)	1.50	2.85
ATF flow rate (L/min)	0.6	1.0
Cell retention (%)	>99.9	>99.9
VLP retention (%)	>99.9	54[b]
VLP concentration in bioreactor (VLP/mL)	6.80E+10	7.74E+09
VLP concentration in harvest (VLP/mL)	1.29E+07	5.79E+09
Produced VLPs in harvest fraction at 96hp (%)	>0.1	48.0

PS: Polysulfone, PP: Polypropylene, PES: Polyethersulfone

(a) Confidential information not provided by the manufacturer

(b) Up to 96hpt

the process, reaching 48% at 96hpt (**Figure 3E**). Further confirmation and measurements up to 168hpt are still required to successfully assess the overall performance of the VHU2 module. The same set up to couple both XCell ATF2 and VHU2 modules with the bioreactor was used (**Figure 4**). Both in ATF mode and using the same XCell™ C24V3 controller for pressure and exhaust pump cycles. Due to the structural and dimensional differences of the different modules (**Table 2**), the ATF flow had to be set at a different rate, 0.6 L/min for XCell ATF2 and 1.0 L/min for VHU2 (**Figure 4, Table 2**). The difference in ATF flow rate directly influences the induction of cellular shear stress. During the initial 48h before transfection, before initializing continuous addition and removal of culture media, the bioreactor operates in batch mode until a VCD of 2·10^6 cells/mL is reached. As the ATF cycles are running during this 48h period, shear stress can be observed, although negligible at 0.6 L/min as cell viability is maintained at ~97% (**Chapter three Part II**). However, at an ATF flow rate of 1.0 L/min, cell viability decreases in this 48h period (**Figure 3A**) down to ~90%. Nevertheless, the main difference between both retention devices is the percentage of VLP retention. The fact that VLPs do not permeate XCell ATF2 HFM implied that this retention device can be used to concentrate VLPs in the bioreactor while separating unassembled monomers and the VHU2 module can be used for a continuous harvest of VLPs, implementing a conventional perfusion process.

VHU2 FILTER MODULE CHARACTERIZATION

The first macroscopic noticeable difference between filters is size, since XCell ATF2 presents an internal diameter of 1mm and a wall 120µm thick, while the VHU2 module presents a variable internal diameter of 2-3mm and a wall 2mm thick. Moreover, in order to observe the microscopic internal structure and characterize the nature of the VLP retention, one of the tubular filters from a VHU2 module was visualized using confocal microcopy. Different sections were individually analyzed. The luminal wall of the VHU2 filter component was observed using reflected light and fluorescence emitted by GFP excitation (**Figure 5A, Supplementary Figure S1**). The observed structure of the luminal wall presented characteristic features, a repeated pattern of seemingly regular forms such as pentagons and hexagons forming a grid or mesh in which the VLPs are encased or embedded along the edges. This is significantly different to the structure of the XCell ATF2 HFM reported in **Chapter three Part II**, showing a more irregular pattern. Also the intensity of the detected fluorescence was different comparing both retention devices. The detected fluorescence in the VHU2 module was 12-times lower than the one observed in the XCell ATF2 HFM. This indicated that fewer VLPs are being retained in the filter, coherently with the filtration efficiency results reported **Table 2**. Also, the fact that the filtration percentage increased with time (**Figure 3E**) could be explained observing the transversal cut of the fiber (**Figure 5B**). Although overall fluorescence intensity was low in most of the analyzed areas in transversal cuts, the luminal edge always presented higher fluorescence intensity suggesting that VLP retention is mostly found in the luminal edge and sequentially decreased towards the exterior of the fiber. This gradient might account for the observed sequential increase of VLP filtration percentage. At the beginning of the process, VLP retention over the luminal wall and in the edges of the mesh-like structure might be higher and once this interface is covered, retention decreases, allowing a higher filtration percentage. The fluorescent gradient in the section of the tube wall and the luminal wall can be observed in **Figure 5C**, depicting a longitudinal cut of a VHU2 fiber. Here, a complete section of the tube wall of ~2.0 mm is shown together with the interface and luminal wall. All stacks acquired using confocal microscopy have been merged to show the continuous wall.
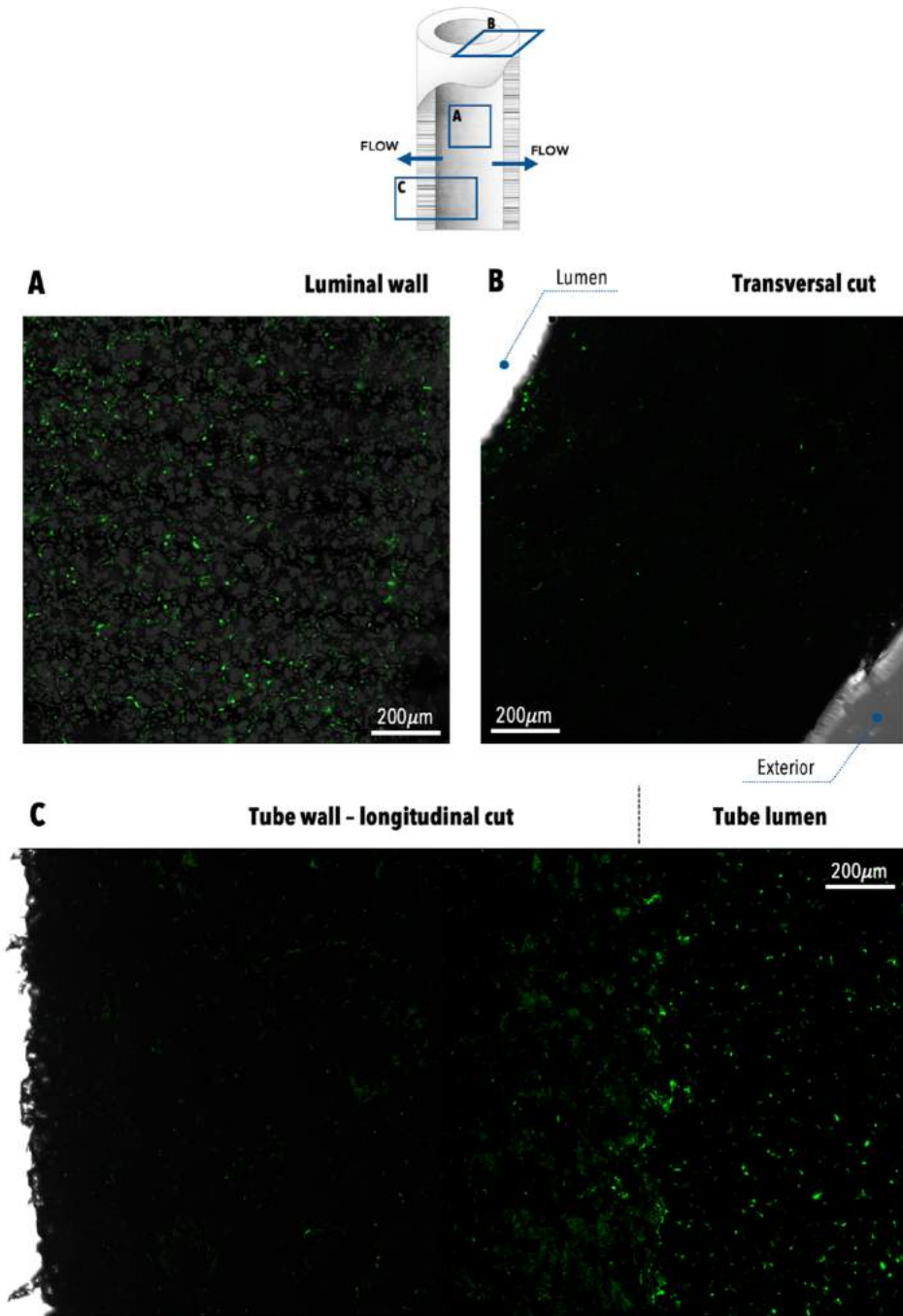
Figure 5. Characterization of the VHU2 filter module using confocal microscopy. **A)** Merge of reflected light and fluorescence emission for a cut of the luminal wall of the tube. **B)** Merge of transmitted light and fluorescent emission for a transversal cut of the filter wall. **C)** Merge of transmitted light and fluorescence emission of a longitudinal cut of the tube wall. Here, it can be observed the tube wall and the luminal wall and their subsequent fluorescence patterns.

CONCLUSIONS

The proteomic analysis of the established perfusion process allowed the identification of cellular markers throughout the process correlating with transfection efficiency like GRK7 or acyl-CoA dehydrogenase. These markers could be eventually used to monitor a production process with no fluorescent or protein reporter associated. Furthermore, the VHU® retention system for continuous harvest recovery was evaluated observing a total VLP filtration of 46% up to 96hpt. The filter module presented a saturation profile, increasing the filtration efficiency with time along the process. The highest observed filtration percentage up to 96hpt was 48%. This saturation profile could be explained by the internal microscopic structure of the filter module whose inner wall presented a gradient of retention from the luminal edge to the exterior. The higher percentage of VLP filtration of the VHU2 module allowed the establishment of a conventional perfusion process with continuous harvest, which needs to be further evaluated and optimized.

REFERENCES

1. Buffo, M. M. *et al.* Influence of dual-impeller type and configuration on oxygen transfer, power consumption, and shear rate in a stirred tank bioreactor. *Biochem. Eng. J.* **114**, 130–139 (2016).

2. Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

3. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

4. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

5. Martínez-Bartolomé, S. *et al.* Properties of average score distributions of SEQUEST: the probability ratio method. *Mol. Cell. Proteomics* **7**, 1135–45 (2008).

6. Bonzon-Kulichenko, E., Garcia-Marques, F., Trevisan-Herraz, M. & Vazquez, J. Revisiting peptide identification by high-accuracy mass spectrometry: problems associated with the use of narrow mass precursor windows. *J Proteome Res* **14**, 700–710 (2015).

7. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J Proteome Res* **13**, 1234–1247 (2014).

8. Martínez-Acedo, P. *et al.* A novel strategy for global analysis of the dynamic thiol redox proteome. *Mol. Cell. Proteomics* **11**, 800–13 (2012).

9. Trevisan-Herraz, M. *et al.* SanXoT: a modular and versatile package for the quantitative analysis of high-throughput proteomics experiments. *Bioinformatics* **35**, 1594–1596 (2019).

10. Garcia-Marques, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol Cell Proteomics* **15**, 1740–1760 (2016).

11. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **13**, 1234–47 (2014).

12.   García-Marqués, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol. Cell. Proteomics* **15**, 1740–60 (2016).

13.   Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

14.   The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

15.   Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2009).

16.   Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).

17.   Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

18.   Carpentier, E., Paris, S., Kamen, A. A. & Durocher, Y. Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *J Biotechnol* **128**, 268–280 (2007).

CHAPTER FOUR

Characterization of HIV-1 virus-like particles and determination of Gag stoichiometry for different production platforms

ABSTRACT

The importance of developing new vaccine technologies towards versatile platforms that can cope with global virus outbreaks has been evidenced with the most recent SARS-CoV-2 pandemic. Virus-like particles (VLPs) are a highly immunogenic, safe and robust approach that can be used to base several vaccine candidates on. Particularly, HIV-1 Gag VLPs are a flexible system comprising a Gag core surrounded by a lipid bilayer that can be modified to present diverse types of membrane proteins or antigens against several diseases, like Influenza, dengue, west Nile virus or HPV, where it has been proven successful. The size distribution, and structural characteristics of produced VLPs vary depending on the cell line used to produce them. In this work we established an analytical method of characterization for the Gag protein core and clarify the current variability of Gag stoichiometry in HIV-1 VLPs depending on the cell-based production platform, directly determining the number of Gag molecules per VLP in each case. Three Gag peptides have been validated to quantify the number of monomers using parallel reaction monitoring (PRM), an accurate and fast, mass-spectrometry-based method that can be used to assess the quality of the produced Gag VLPs regardless of the cell line used. An average of 3617 ± 17 monomers per VLP was obtained for HEK293, substantially varying between platforms, including mammalian and insect cells. This offers a key advantage in quantification and quality control methods to characterize VLP production at large scale to accelerate new recombinant vaccine production technologies.

ABBREVIATIONS

eGFP: enhanced green fluorescence protein, **ESCRT**: Endosomal sorting complex required for transport, **Gag::eGFP**: translational fusion of HIV-Gag protein and eGFP, **hpi**: hours post infection, **hpt**: hours post transfection, **HPV**: Human papilloma virus, **MOI**: Multiplicity of infection, **NTA**: Nanoparticle tracking analysis, **PAT**: Process analytical technologies, **PBS**: Phosphate-buffered saline, **PEI**: polyethyleneimine, **PRM**: Parallel reaction monitoring, **PSM**: Peptide-spectrum match, **RFU**: Relative fluorescence units, **RMCE**: Recombinase-mediated cassette exchange, **SGE**: Stable gene expression, **TGE**: Transient gene expression, **VLPs**: Virus-like particles

INTRODUCTION

Virus-like particles (VLPs) are a suitable platform for the design and production of novel vaccines for emerging diseases. They present a protein core that can be later modified adding immunogenic epitopes from different diseases, what is known as pseudotyping. This core is formed by protein subunits that self-assemble within the cell. It is the case of HIV-1 Gag VLPs, it is composed of multiple Gag molecules. When recombinant Gag polyprotein is produced in mammalian cell cultures, like HEK293, Gag monomers are transported to the plasma membrane where the endosomal sorting complex required for transport (ESCRT) coordinates their assembly and oligomerization, later promoting excision of the evaginated portion of the plasma membrane, forming VLPs[1,2]. Gag polyprotein is composed by six different parts: matrix protein p17 (MA), capsid protein p24 (CA), spacer peptides (SP1, SP2), nucleocapsid protein p7 (NC) and p6[3]. In case of HIV-1 native virions, Gag suffers a process of maturation comprising proteolytic cleavage of its parts after budding off from the cell[4]. This reorganization forms the characteristic cone-shaped capsid protecting the viral RNA and leaves an outer shell of MA subunits below the plasma membrane. For VLPs, the outer shell is formed by monomers of complete and unmatured Gag polyproteins, lacking the following processing steps. The structure and stoichiometry of mature HIV-1 virions is well characterized[5] but for Gag VLPs or immature HIV-1 virions, controversial results ranging from 750 to 5000 Gag molecules per particle have been reported[6-9]. These values become even more variable when comparing among different cell-

based production platforms. The reason for this variability is the method of estimation, mostly based on mature HIV virion analysis instead of uncleaved Gag VLPs. Most approaches determine the number of Gag monomers through indirect methods like CA quantification, via ratio RNA to CA or envelope glycoprotein visualization by EM[8,10,11]. Direct observations of Gag lattice from immature particles reported up to 5000 monomers per particle considering the membrane surface was fully packed with Gag[7]. However, Gag distribution was reported to be uneven throughout the membrane, finding patches of Gag and areas less densely packed[5]. Assuming 60% of membrane surface coverage, a theoretical estimation around 3000 monomers was reported[12]. The differences of particle size depending on the producing cell line added variability to this conundrum. Determining the number of Gag monomers is an elemental factor for the field of bioprocess development since most of fast and routine estimation of product titer, quality control and process monitoring is based on fluorimetry, using reporters like GFP; or ELISA, obtaining ng/mL of Gag[13]. In order to translate these results into particle concentration, the number of monomers per VLP is essential. No direct measurement of the number of Gag molecules has been reported before to the best of our knowledge. Due to the needs for a quantification tool for rapid VLP characterization, this work describes analytical methods for fast and accurate VLP absolute quantification, a tool for process monitoring using MS-based parallel reaction monitoring (PRM) validating the use of three Gag peptides for correct quantification, as well as the characterization of the Gag shell, visualizing the VLP structure via electron cryo-tomography where the distribution and measurements of Gag::eGFP monomers can be observed. These techniques have been applied for the characterization of Gag VLPs produced in different production platforms, such as mammalian and insect cells, to determine their differences and composition as well as giving insight about their stoichiometry. All in all, this work contributes to refine the existing knowledge on the structure of Gag-based embedded VLPs and to their further application as vaccines.

MATERIALS AND METHODS

CELL LINES AND CULTURE CONDITIONS

Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®, NY, USA) at 37°C, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Culture media for HEK293 cells was FreeStyle™ 293 Expression Medium (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) supplemented with 0.1% Pluronic™ F-68 non-ionic surfactant, transferrin and insulin (Merck Millipore, Kankakee, IL, USA), a lipid mix composed of synthetic cholesterol (100x Synthecol®, Sigma, San Luis, MO, USA) tocopherol acetate (0.2 mg/mL T1157, Sigma) and fatty acids (100x F7050, Sigma) and glucose (G8270, Sigma) at a concentration previously reported[14]. For validation in different culture media, HyCell™ TransFx-H media (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ and 0.1% Pluronic™ F-68 Non-ionic Surfactant (Gibco) and FreeStyle™ F17 Expression Medium (Gibco) supplemented with 8 mM GlutaMAX™, 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco) and IGF-1 at a final concentration of 50 g/L. Stable cell line noted as SGE-Random was generated by PEI-mediated transfection of HEK293 and random integration of pGag::eGFP plasmid, followed by clonal selection under Geneticin™ selective antibiotic G418 (Gibco). Stable cell line noted as SGE-RMCE was generated by infection (MOI = 0.01) with a lentivirus containing the recombinase-mediated cassette exchange (RMCE) which consists of CMV-Gag::eGFP-pA construct surrounded by two flipase recognition targets (FRT). Selection was performed by clonal selection under Puromicyn (P8833, Millipore, Merck). These cell lines were cultured in FreeStyle™ 293 media supplemented as described before. Sf9 (71104, Merck, Darmstadt, Germany) and High Five (B85502, ThermoFisher) cells were cultured ExpiSf™ CD Medium (Gibco) using an orbital shaker (Stuart, Stone, UK) and maintained at 27 °C in an incubator. Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

TRANSIENT TRANSFECTION AND INFECTION

For HEK293 cells, transient transfection was carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 µg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used. The plasmid used contained a gene encoding HIV-Gag protein fused to eGFP (Gag::eGFP). Briefly, the corresponding DNA mixture was diluted with culture media and vortexed for 10 seconds. Then, PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times. The mixture was then incubated for 15 min at room temperature and then added to the cell culture. Cell cultures were harvested at 72hpt, centrifuged at 3.000xg for 10 min, pellet discarded and supernatant stored at -80ºC for further analysis. For Sf9 and High Five cells, baculovirus stock preparation and titration were previously prepared and performed as previously described[15]. Infection was carried out at $2 \cdot 10^6$ cells/mL using a multiplicity of infection (MOI) of 1. Cell cultures were harvested at 48 hpi by centrifugation at 3000×g for 5 min. Pellets were discarded and supernatants stored at -80ºC for further analysis.

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

Supernatants were centrifuged at 1000×g for 5 min. Green fluorescence was measured at room temperature using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) set as follows: λex = 488 nm (slit 5 nm), λem = 510 nm (slit 10 nm). Relative fluorescence units values (RFU) were calculated by subtracting fluorescence units (FU) values of non-transfected control samples.

ULTRACENTRIFUGATION

Concentrated and purified HIV-1 Gag VLPs were obtained by single cushion ultracentrifugation. Briefly, a volume of 35 mL of clarified supernatant from HEK293 transiently transfected cell culture was layered on top of a 30% sucrose cushion of 3mL and centrifuged at 26.000 rpm for 2.5h at 4ºC using a SW32 rotor in a Beckman Optima L100XP centrifuge (Brea, CA, USA). After ultracentrifugation, supernatant was

discarded and the obtained pellet was resuspended in 1 mL of PBS (Hyclone, GE HeathCare, Chicago, IL, USA). The concentrated material was stored at −80ºC for future studies.

HIV-1 GAG VLP QUANTIFICATION BY NANOPARTICLE TRACKING ANALYSIS

Nanoparticle Tracking Analysis (NTA) was also used to quantify fluorescent particles. NTA measurements were performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag::eGFP VLPs and neutral density filter for total particle by light scattering. Data were analyzed with NanoSight® NTA 3.1 software. NanoSight® settings are described elsewhere[16,17].

HIV-1 GAG VLP QUANTIFICATION BY ELISA

Gag::eGFP polyprotein concentrations were determined by p24 ELISA using the commercially available kit Innotest® HIV antigen mAb (Innogenetics NV, Gent, Belgium). The assay was performed according to manufacturer's instructions.

PROTEIN SAMPLES PREPARATION FOR LC-MS/MS ANALYSIS

Protein extraction was performed from stored supernatants using extraction buffer (50 mM Tris-HCl pH=8.5, 4% SDS, 50mM DTT) 1:1 (v/v) added to the sample of each condition. Samples were sonicated for 5 minutes, boiled for another 5 minutes and then centrifuged at 13.000xg for 15 min. Protein extracts were quantified with RC/DC Protein Assay (Bio-Rad, Hercules, CA, USA) and stored in -20 ºC until the tryptic digestion process. Protein digestion was performed as previously described[18]. Briefly, proteins were digested using sequencing grade trypsin (Promega, Madison, WI, USA) and the filter-assisted sample preparation technology (FASP, Expedeon, San Diego, CA, USA). After digestion, three synthetic Gag peptides ETINEEAAEWDR*, FAVNPGLLETTSEGCR*, and TLNAWVK* labelled with ^{13}C and ^{14}N in Lys (K*) and Arg (R*) were added to the digested pool of peptides at different known concentrations. A standard curve was analyzed using 0, 10, 100, 250, 500, 1000, 2000 and 5000 fmol of each peptide. These peptides were selected from previous studies[19,20] among 79 different Gag peptides based on best identification scores,

number of peptide-spectrum match (PSM), highest intensities MS/MS spectra and suitability for PRM (**Supplementary Table S1**). The peptides ETINEEAAEWDR and TLNAWVK are found within the CA p24 domain and FAVNPGLLETTSEGCR within the MA p17 domain of Gag polyprotein (**Supplementary Figure S1**). The peptide mix was purified using Oasis HLB C18 column (Waters, Milford, MA, USA), dried and stored at -20°C for further LC-MS/MS analysis.

PARALLEL REACTION MONITORING (PRM) ANALYSIS

Tryptic peptide mixtures were subjected to LC-MS/MS analysis on a nano-HPLC Easy nLC 1000 liquid chromatograph (Thermo Scientific, San Jose, CA, USA) coupled to a QExactive HF mass spectrometer (Thermo Scientific). Peptides were suspended in 0.1% formic acid, loaded onto a C18 reverse-phase trapping column (Acclaim PepMap100, 75-μm internal diameter, 3-μm particle size and 2-cm length, Thermo Scientific), and separated using two different analytical C18 nano-columns to validate reproducibility (EASY-Spray column PepMap RSLC C18, 75-μm internal diameter, 3-μm particle size and 50-cm length and Acclaim PepMap100, 75-μm internal diameter, 3-μm particle size, 50-cm length, Thermo Scientific), in a continuous gradient (8–28%B in 60 min, 28–90%B in 2min, 90%B in 5min, 90–2%B in 3min and 2%B in 30min; where buffer A is 0.1% formic acid in HPLC grade H_2O and buffer B is 100% ACN, 0.1% formic acid in HPLC grade H_2O). Spectra were acquired using the full ion scan mode followed by three parallel reaction monitoring (PRM) analyses along the chromatographic run. For the full MS scan mode, range 390–1500 *m/z* was selected and 120,000 FT. For the PRM scan mode, the same m/z range was selected, resolution was set to 15,000, isolation window of 1.0 *m/z*, maximum injection time of 100 ms and HCD fragmentation was performed at 30% of normalized collision energy. The monitored monoisotopic masses are presented in **Supplementary Table S2**. The chromatograms and fragmentation patterns of the different peptides are shown in **Supplementary Figure S2**. Chromatograms were analyzed using the FreeStyle™ software package from Xcalibur™ Software (Thermo Scientific). Peak areas were quantified using ICIS algorithm and Gaussian smoothing of 7. Two standard curves were analyzed, one for particles produced in FreeStyle™ 293 Expression Medium and one for particles produced in ExpiSf™ CD Medium.

MONOMERS CALCULATION

VLP samples from different production platforms were analyzed by PRM in order to calculate the number of Gag molecules per particle. Using the standard curve of 0, 10, 100, 250, 500, 1000, 2000 and 5000 fmol, the total amount of fmol of each peptide can be calculated. It is assumed that the number of fmol

$$\frac{Monomers}{VLP} = \frac{fmol_{Pep} \cdot N_A}{10^{15} \cdot C_{VLP} \cdot V_{DIGES}} \cdot \frac{F_D}{F_F} \quad (1)$$

from each peptide is equal to the number of Gag::eGFP fmol. Using **Equation 1**, the number of monomers per particle can be calculated. N_A is the Avogadro constant ($6.022 \cdot 10^{23}$), C_{VLP} is the VLP concentration of fluorescent diffracting particles measured by NTA (VLP/mL), V_{DIGES} is the total sample volume used for protein digestion (mL) and F_D and F_F are the dilution factor for the digestion process and the factor for the fluorimetry:NTA ratio (as further described) respectively.

ELECTRON CRYO-TOMOGRAPHY

VLPs from ultracentrifugation sample were mixed with 6 nm BSA-tracker colloidal gold particles (Aurion) as fiducials, applied to glow discharged 200 mesh R2/2 Quantifoil Cu-Rh grids (Quantifoil, Großlöbichau, Germany) and plunge frozen using a Vitrobot (Thermofisher-Scientific). Samples were imaged under cryo conditions at 200 kV on a Talos Arctica (xFEG) electron microscope (FEI, Hillsboro, OR, USA). equipped with a Falcon III DDD. Cryo-ET acquisition was performed using a 0° start, 3° step up to ± 60°. Tilt series datasets were collected under low-dose conditions with a cumulative electron dose of ~110 electrons/ $\mathring{A}^2$. Tilt series were collected using EPU software (FEI Company, Hillsboro, OR). Defocus was applied from -2.5 to -3.5 μm with a pixelsize of 1.81 $\mathring{A}$/px[21] (nominal magnification 57000X). Individual VLPs from 2D cryo images were further analyzed by 2D image analysis using Fiji software[22].

IMAGE PROCESSING

Images were binned to 3.62 Å/px and aligned by using the gold particles as fiducial markers in IMOD[21,23]. SIRT Tomo3D[24] method was used to get final 3D volumes reconstructions (tomograms). These reconstructions were filtered using Tomobflow[25] software. Semiautomatic segmentation was performed using Amira™ software (Thermofisher). Segmented volumes were visualized using Chimera-USFC[26].

RESULTS

VLP production and experimental design

VLPs produced by transient gene expression (TGE) transfecting HEK293 cells at $2 \cdot 10^6$ cells/mL and harvested at 72 hpt were subsequently concentrated by ultracentrifugation. Purified Gag VLPs where then quantified by fluorimetry, NTA and p24 ELISA to establish correlations between the different techniques and design a fast and accurate quantification method for VLPs. These purified samples were also used to determine the distribution and configuration of Gag::eGFP monomers in the polyprotein shell within VLPs via electron cryo-tomography. Additionally, to directly quantify the molecules of Gag per particle, the analytical method of PRM was performed. In order to determine Gag stoichiometry for different production platforms, HEK293 stable gene expression (SGE), produced by random integration and RMCE, as well as Sf9 and High Five insect cells cultures were also analyzed. VLPs from SGE HEK293 cells were harvested after 72 h of culture at $4\text{-}5 \cdot 10^6$ cells/mL and insect cells, Sf9 and High Five, were infected with baculovirus and harvested at 40 hpi (**Figure 1A**). The analysis of VLPs directly from the clarified supernatant allowed the implementation of a robust and accurate analytic method that could be implemented as a routine industrial product characterization procedure.

Correlations between ELISA, fluorimetry and NTA quantification measurements

VLPs from clarified supernatants from HEK293 via TGE were ultracentrifuged using a 30% sucrose cushion purifying and concentrating VLPs 18.3-fold (**Figure 1B**), discarding unassembled free Gag::eGFP monomers that could interfere in fluorescence detection. A series of increasing dilutions from this purified
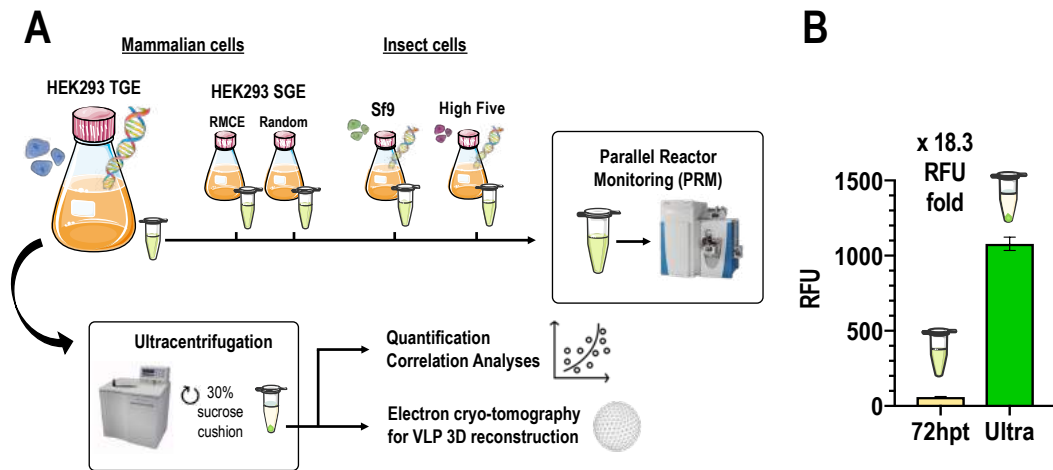
Figure 1. Experimental design, workflow and production data. **A**) Experimental design. VLPs produced in HEK293 cells by transient gene expression (TGE) were harvested at 72hpt and ultracentrifuged from cell cultures using a 30% sucrose cushion to separate VLPs. VLPs were quantified using different methods and correlation analysis were performed between the quantification methods. VLPs were also visualized via electron cryo-tomography to characterize their structure and Gag::eGFP monomers distribution. Culture samples were also centrifuged to remove cells and the supernatants were used to performed absolute quantification of Gag polyprotein via via LC-MS/MS analysis using parallel reaction monitoring (PRM). Results provided an estimation of the number of Gag molecules per particle in four production platforms that can be compared to analyze the stoichiometry in the different platforms: HEK293 stable gene expression (SGE) via random integration (random), SGE via recombinase-mediated cassette exchange (RMCE), insect cell Sf9 and Hi five via infection. **B**) Relative fluorescent units measured in HEK293 TGE centrifuged culture at 72hpt and after ultracentrifugation. The separation process achieved to increase 18.3 times the fluorescence intensity.

sample was used to obtain correlations between p24 ELISA, fluorimetry and NTA measurements within the linear range of the different techniques based on previous data[13]. The correlation plots are presented in **Figure 2A-C.** Most ELISA kits used for p24 detection are optimized to detect the cleavage form of CA p24 **(Supplementary Figure S1**), as it is present in mature HIV-1 particles. Previous work showed that the measurement obtained from these kits needed to be corrected by 10-fold to correctly report the amount of p24 in the unprocessed form of Gag polyprotein[13]. In order to obtain a Gag::eGFP (84kDa) measure instead of p24 (24kDa), another 3.6-fold must be added to the ELISA-RFU correlation (**Figure 2A**). The resulted equivalence is shown in **Equation 2**. This equation can also be used for p24 determination in ng/mL removing the x36 factor from the correlation. For fast quantification of VLP titers, fluorescence is routinely used. NTA-RFU correlation is shown in **Figure 2B**. Here, an expression to convert RFU to VLP concentration is calculated (**Equation 3**). This expression implies a limit of quantification of 14 RFUs,

impeding the use of this method for platforms with low production, where there was little difference between the fluorescence given by the product and by the medium background. The correlation NTA-ELISA is shown in **Figure 2C**. Here, the value obtained from ELISA assay can be converted to VLP concentration using **Equation 4**. The validity of the correlation was tested in FreeStyle™ 293, FreeStyle™ F17 and HyCell™ TransFx-H media. Direct measurements from clarified supernatants showed that a correction factor was needed when quantifying from unpurified samples. The presence of extracellular free unassembled Gag::eGFP monomers in the supernatant caused the NTA-RFU correlation to overestimate VLP concentration. The VLP concentration calculated via fluorimetry was magnified by 4-fold (**Figure 2D**). Therefore, when assessing VLP concentration from direct harvested supernatant fluorimetry measurements (n = 12), a correction factor of ¼ to the previous correlation (**Eq. 3**) was needed. This correction factor has already reported as the ratio of VLP:monomer in previous studies of VLP production at bioreactor scale and calculated as an average value for different harvest times and different final cell culture viabilities[16]. This factor was introduced in **Eq.1** as $F_F=4$ for the analytical calculation of monomers. The aforementioned set of correlations could be used to theoretically calculate the number of Gag::eGFP monomers per particle.

Theoretical calculation of the number of Gag molecules

In order to calculate the number of Gag molecules, **Eq. 2** and **Eq. 3** were used. First, an interval of 30-90 RFUs was taken as a reference since it is the most common RFU interval found in Gag::eGFP VLP harvested cultures in HEK293 cells using FreeStyle™ 293, FreeStyle™ F17 and HyCell™ TransFx-H media. RFUs were converted to ng/mL of Gag::eGFP using **Eq. 2**. Then, considering that one monomer of Gag::eGFP of 84kDa corresponds to $1.39 \cdot 10^{-10}$ ng, the total number of monomers in the sample was calculated. Thus, calculating the VLP concentration using **Eq. 3**, the number of monomers in one VLP can be assessed. The same procedure was followed using **Eq. 2** to convert RFU to p24 concentration and considering that one monomer of p24 corresponds to $3.98 \cdot 10^{-11}$ ng, in a different approach the number of monomers was again calculated. Considering both results for each RFU value taken as reference, an estimation of 3696 ± 32 Gag molecules per VLP was calculated.

A

p24 Correlation

Gag::eGFP (ng/mL) = (4.768 · RFU + 9.541) · 36

$R^2 = 0.9944$

B

VLP/mL Correlation

VLPs/mL = (4.448 · RFU - 63.30) · 10^8

$R^2 = 0.9938$

C

VLP/mL Correlation

VLPs/mL = (9.292 · p24 - 694.6) · 10^7

$R^2 = 0.9914$

D

Fold: 4:1
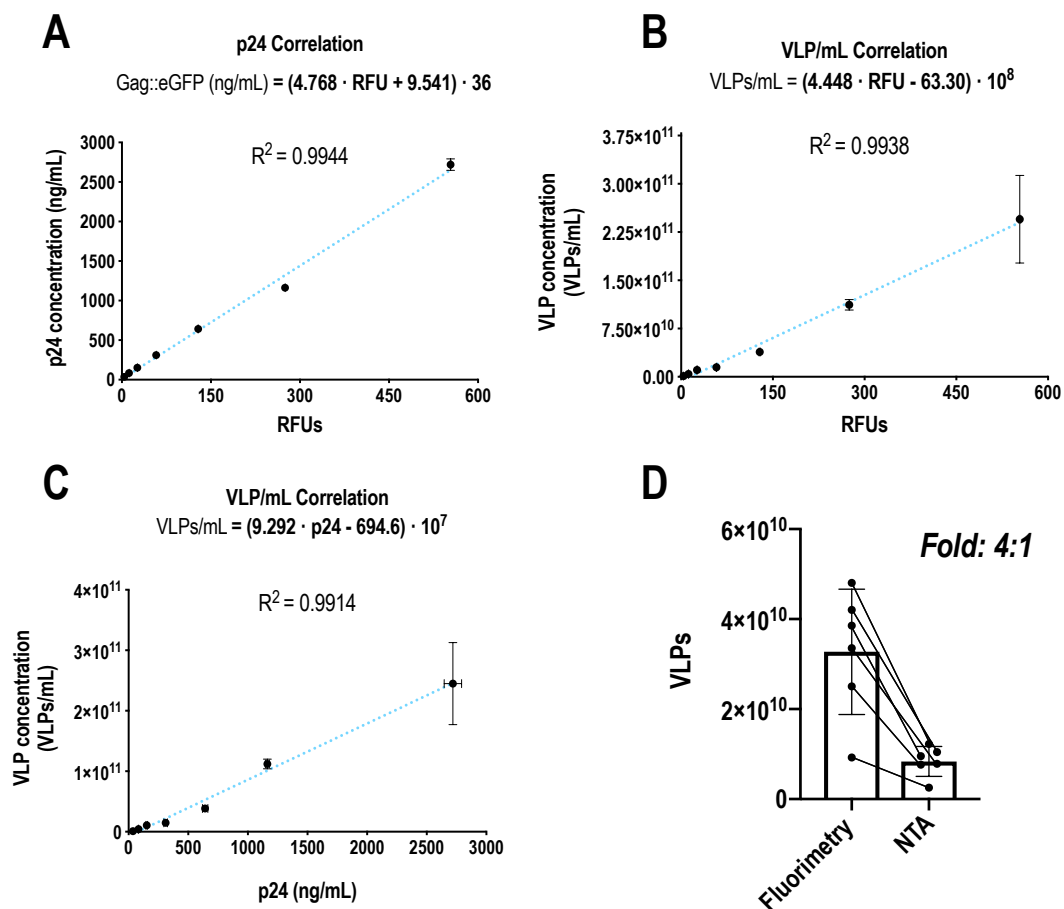
Figure 2. Correlation analyses for VLP quantification. **A)** Correlation between relative fluorescent units (RFUs) and p24 concentration of a purified VLP fraction measured by ELISA. A correction factor of 36 was used to transform ng/mL of p24 to ng/mL of Gag::eGFP. **B)** Correlation between RFUs and VLP concentration of a purified VLP fraction measured by nanoparticle tracking analysis (NTA). **C)** Correlation between ng/mL of p24 measured by ELISA and VLP concentration by NTA of a purified VLP fraction. **D)** Comparison of estimated VLP concentration in a cell culture via fluorimetry correlation and direct NTA measurement. In order to quantify the VLP concentration of a harvested cell culture (72hpt), a fold of 1/4 was applied to the fluorimetry correlation (B) for the final quantification model.

Parallel reaction monitoring (PRM) for analytical determination of Gag stoichiometry

PRM was performed for an absolute quantification of the number of monomers per particle and to establish a method of process quality assessment for Gag VLP-based bioprocesses. Four different Gag peptide ions species were monitored: ETINEEAAEWDR (charge 2+), TLNAWVK (charge 2+) and FAVNPGLLETTSEGCR (charge 2+ and 3+) from CA and MA Gag domains respectively. Synthetically

synthesized peptides labelled with ^{13}C and ^{14}N in arginine (R) and lysine (K) were added to previously digested VLP sample in order quantified the endogenous peptides (**Figure 3A**). Different known amounts of each labelled peptide were added creating a standard curve of 0, 10, 100, 250, 500, 1000, 2000 and 5000 fmol (**Table 1**) and obtaining a linear correlation with $R^2 > 0.95$ in all cases (**Figure 3B**). Endogenous peptides were quantified and the ratio of monomers per particle calculated using **Eq. 1**. The results are shown in **Table 1**. A final number of monomers per particle was calculated considering the results from all four peptides, resulting in 3617 ± 63 Gag molecules per VLP produced by TGE in HEK293 and harvested at 72hpt. This calculation was validated using three different culture media (**Supplementary Figure S3**) and agreed with the previous theoretical result, thus validating the quantification method by fluorimetry. The observation that only (25 ± 2) % of the total fluorescence corresponded to assembled Gag VLPs was also validated since PRM quantification was carried out directly from harvested supernatants. Therefore, all quantifiable Gag::eGFP monomers came from VLPs and unassembled free monomers. The quantification equation (**Eq. 1**) required the addition of the term of $F_F = 4$ associated to these unassembled free monomers. This term balanced with the dilution factor of 4 introduced due to the digestion and desalting process. For the application of this method to GFP-free Gag VLPs, **Eq. 4** can be used to determine VLP concentration and to experimentally determine the F_F factor in each specific case.

Analysis of VLPs from different HEK293 cell production platforms

HEK293 cell line can be used to produce VLPs in suspension cultures by two main strategies. TGE is based on episomal expression of exogenous DNA upon transfection while SGE is based on the integration of the protein-coding DNA in the cell genome for prolonged expression. The integration can be achieved by a method of random integration, like RMCE. VLPs produced by these three different platforms (TGE, SGE-random and SGE-RMCE) have been analyzed by NTA and quantified by PRM. Interestingly, the average size of the produced VLPs varied depending on the cell platform where they had been produced. VLPs produced in SGE systems presented smaller sizes than the ones produced by TGE (**Figure 4A**). For these analyses, only the main subpopulation in each condition was considered. The particle size distribution in SGE-RMCE
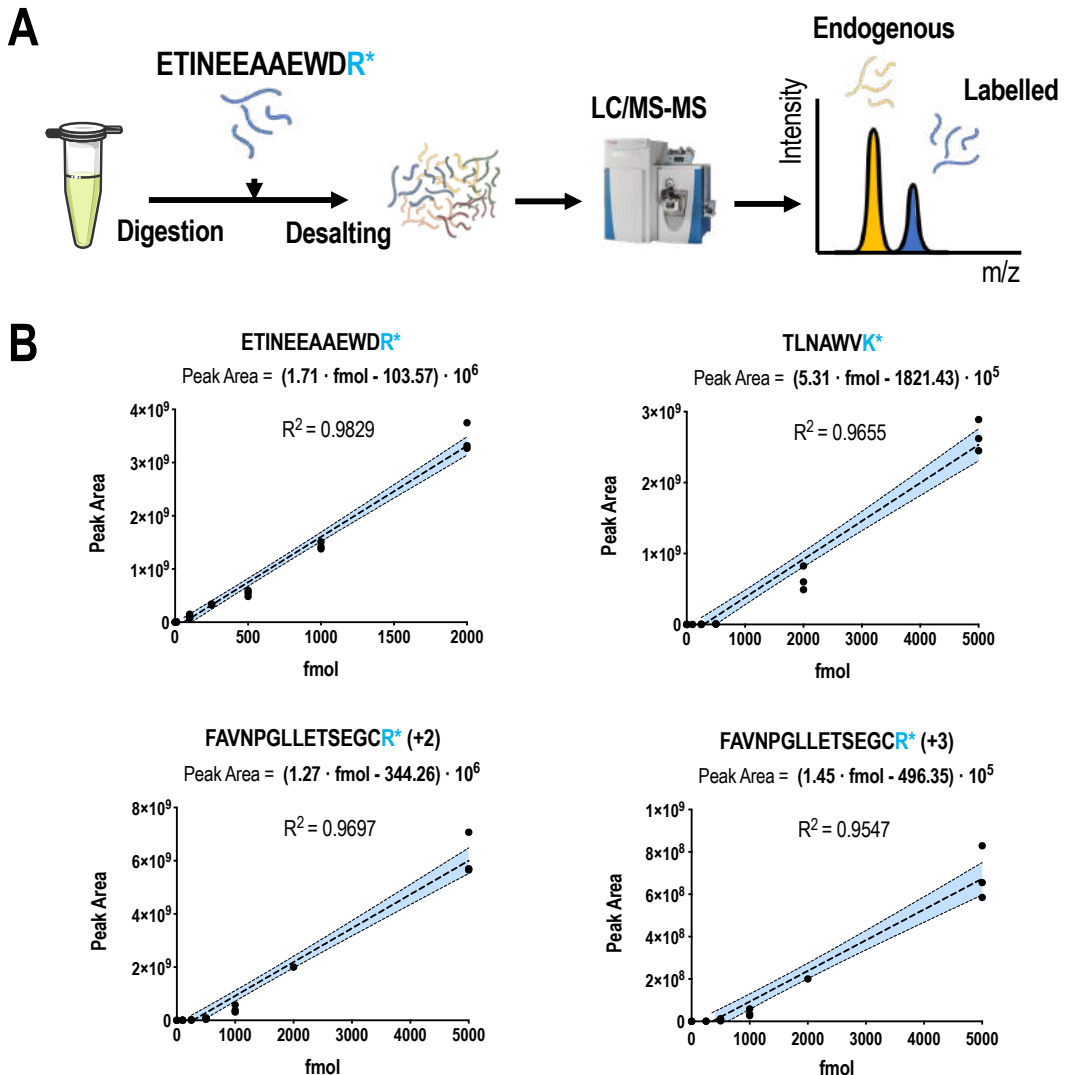
Figure 3. Parallel reaction monitoring for absolute Gag quantification. **A)** Experimental workflow. A known concentration of Gag peptides labelled with ^{13}C and ^{14}N in Lys (K*) and Arg (R*), was added to digested protein samples from harvested cell cultures and analyzed using LC-MS/MS. The peak areas from the labelled peptide were used for absolute Gag quantification. **B)** Correlation for ETINEEAAEWDR, TLNAWVK and FAVNPGLLETSEGCR peptide with charge 2+ and 3+. 95% confidence bands represented by blue area.

shifted to 123 ± 1 nm of particle diameter from the 145 ± 4 nm of TGE. The SGE-random condition produced even smaller particles, presenting a diameter of 109 ± 6 nm. This condition not only produced smaller particles but also a lower VLP concentration (**Figure 4B**). At 72 hours of harvest, SGE-Random produced a mean of $(3.8 \pm 0.3) \cdot 10^8$ VLPs/mL, significantly lower (*p-value* < 0.0001) than a harvest at 72

Table 1. Parallel reaction monitoring for Gag::eGFP peptides in harvested VLPs from HEK293 cell line cultured in FreeStyle™ 293 Expression Medium. Analyzed peptides presented charge 2 + except peptide FAVNPGLLETSEGCR whose charge 3 + especies was also analyzed. The number of replicates is denoted as "n". Coefficient of variation was calculated as %CV = 100·SD/$\bar{x}$. Standard error is calculated as ε=SD/$\sqrt{n}$

		ETINEEAAEWDR			FAVNPGLLETSEGCR (2+)			FAVNPGLLETSEGCR (3+)			TLNAWVK	
n	fmol	Mean Peak Area	%CV	fmol	Mean Peak Area	%CV	fmol	Mean Peak Area	%CV	fmol	Mean Peak Area	%CV
3	0	0	-	0	0	-	0	0	-	0	0	-
3	10	4.13E+06	14.61	10	0	-	10	-	-	10	4.11E+04	n=1[a]
5	100	1.07E+08	36.44	100	7.67E+06	31.85	100	-	-	100	3.90E+05	n=1[a]
3	250	3.33E+08	2.43	250	1.72E+07	24.12	250	1.16E+06	16.95	250	3.15E+06	29.15
3	500	5.44E+08	10.61	500	6.90E+07	37.71	500	7.21E+06	22.24	500	6.91E+06	16.32
4	1000	1.43E+09	4.47	1000	4.24E+08	33.31	1000	4.02E+07	40.96	1000	6.05E+07	17.59
3	2000	3.45E+09	7.53	2000	9.56E+08	12.02	2000	9.03E+07	24.22	2000	6.39E+08	26.68
3	5000	-	-	5000	6.14E+09	13.10	5000	6.90E+08	18.24	5000	2.65E+09	8.47
Endogenous Quantificatio	**n**	**Monomers/VLP**	**St. Error**	**n**	**Monomers/VLP**	**St. Error**	**n**	**Monomers/VLP**	**St. Error**	**n**	**Monomers/VLP**	**St. Error**
	26	3022	58	24	3919	120	24	4064	49	15	3366	40

[a]Only one replicate was obtained for these measurements

hpt from TGE condition, which produced $(4.9 \pm 0.5) \cdot 10^9$ VLPs/mL. Samples from these conditions were analyzed by PRM to determine the total amount of Gag and calculate the number of Gag molecules per particle. Coherently, VLPs coming from SGE-random production platform contained fewer monomers per particle (**Figure 4C**). TGE VLPs showed a number of monomers of 3617 ± 63 while SGE-RMCE particles with 123 ± 1 nm of diameter, contained 3315 ± 100 monomers, a coherent number of monomers, given the particle average size. Particles from SGE-random platform, being 109 ± 6 nm of diameter were composed by 2305 ± 297 molecules per particle. Furthermore, the ratio of monomer per nm^3 was calculated for each platform and interestingly, all particles produced by TGE regardless of the culture media used presented a ratio of $(2.5 \pm 0.2) \cdot 10^{-3}$ monomers per nm^3 (**Supplementary Figure S3**), while all particles produced by SGE regardless of the integration method presented a significant higher (p=0.0022) ratio of $(3.4 \pm 0.1) \cdot 10^{-3}$ monomers per nm^3 (**Figure 4C**).

VLPs produced in insect cell not always share the same stoichiometry

Fluorimetry, NTA and PRM analyses were conducted for VLPs produced in Sf9 and High Five cells via the baculovirus (BV) infection system. Ratios between fluorimetry and NTA results were again calculated obtaining a factor of $F_F = 2.8$ for Sf9 and $F_F = 5.7$ for High Five cells, which were later applied for the calculation of the number of monomers using **Eq 1**. Regarding PRM, due to the differences of the proteome background present in these samples, a new standard correlation curve was constructed for the labelled peptides following the same procedure used for HEK293. This time, the charge 3+ species of peptide FAVNPGLLETTSEGCR could not be successfully quantified and peptide monitoring was carried out using only the charge 2+ labelled peptide species. The correlation curves of each studied peptide are shown in **Figure 5A**, obtaining a linear correlation with $R^2 > 0.97$ in all cases. The data corresponding to the PRM quantification for the standard curves in Sf9 can be found in **Table 2**. Using these correlations, a total of 3755 ± 174 monomers per VLP could be calculated for Sf9. The main peak of fluorescent diffracting particles at 145 nm observed by NTA analysis of the VLP produced in Sf9 contained a tailing deviation (**Figure 5B**) in agreement with what has been previously reported in VLP produced Sf9 via BV infection[27] corresponding to aggregates of the 145 nm particles and not to larger individual particles. Considering
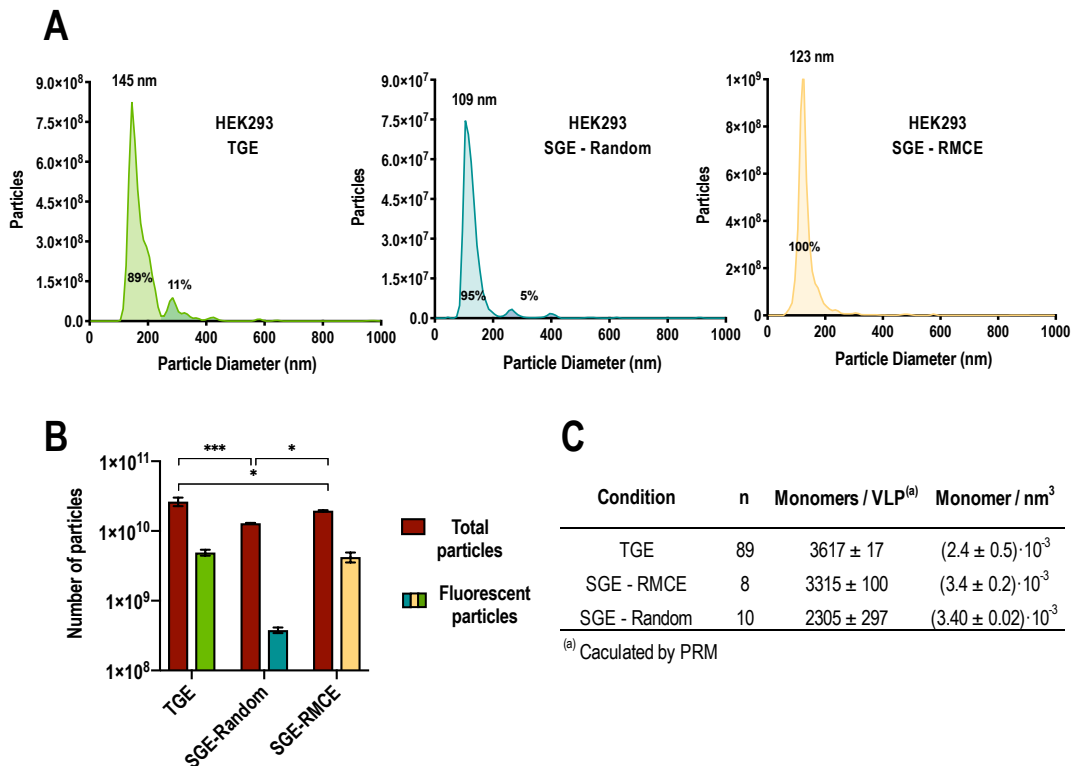
Figure 4. Analysis of VLPs produced in different HEK293 cell-based production platforms. **A)** NTA size distribution curves for fluorescent particles produced by transient gene expression (TGE), stable gene expression by recombinase-mediated cassette exchange (SGE-RMCE) and by stable gene expression by random integration (SGE-random). Each subpopulation presents their contribution to the total number of particles expressed in percentage. **B)** Total number of particles analysed by dispersion and total number of fluorescent particles measured in NTA. Significance calculated by two-way ANOVA, DF=12. **C)** Number of Gag::eGFP molecules (monomers) per VLP in HEK293 for each condition calculated by parallel reaction monitoring (PRM). Standard error is calculated as $\varepsilon = SD/\sqrt{n}$.

this particle size, a $(2.4 \pm 0.1) \cdot 10^3$ monomers/nm^3 ratio was calculated, interestingly sharing the stoichiometry observed in VLPs produced by TGE in HEK293. The total number of diffracting particles was also measured for insect-based production platforms and it was observed that Sf9 and High Five cells produced significantly more ($p<0.0001$) extracellular vesicles than HEK293 platforms observed by NTA, being High Five cells the production platform producing the highest number of extracellular vesicles, in agreement with previous works[15,28,29] (**Figure 5C**). As for VLPs produced in the High Five BV platform, the size distribution curve observed via NTA analysis presented two main distinct subpopulations of 195 and 275 nm, therefore, to successfully calculate the number of monomers of each subpopulation a

A ETINEEAAEWDR*
Peak Area = (3.7 · fmol + 512.3) · 10^5
$R^2 = 0.9869$

FAVNPGLLETSEGCR* (+2)
Peak Area = (1.85 · fmol - 205.5) · 10^6
$R^2 = 0.9739$

TLNAWVK*
Peak Area = (5.18 · fmol - 185.6) · 10^5
$R^2 = 0.9821$

B Sf9 BV — 145 nm, 91%, 9%

High Five BV — 195 nm, 275 nm, 62%, 38%

C Significance: ****, ns — Sf9, Hi5 — Total particles, Fluorescent particles

D

Condition	n	Monomers / VLP[a]	Monomer / nm^3
Sf9	77	3755 ± 174	$(2.4 ± 0.1)·10^{-3}$
High Five A[b]	6	2184 ± 179	$(5.6 ± 0.5)·10^{-4}$
High Five B	6	6126 ± 502	$(5.6 ± 0.5)·10^{-4}$

[a] Caculated by PRM
[b] Subpopulation A corresponds to 195 nm and B to 275 nm

Figure 5. Stoichiometric analysis for VLPs produced in insect Sf9 and High Five cell lines. **A)** Standard correlations for ETINEEAAEWDR, TLNAWVK and FAVNPGLLETSEGCR (+2) peptides analysed in ExpiSf™ CD Medium. 95% confidence bands represented by blue area. **B)** NTA size distribution curves for fluorescent particles produced by baculovirus (BV) infection of Sfe9 and High Five cell lines. Each subpopulation presents their contribution to the total number of particles expressed in percentage. **C)** Total number of diffracting and fluorescent particles analysed by NTA. Significance calculated by two-way ANOVA, DF=12. **D)** Number of Gag::eGFP molecules (monomers) per VLP in each condition calculated by parallel reaction monitoring (PRM) based on the correlation plots in A). Standard error is calculated as $ε=SD/\sqrt{n}$.

Table 2. Parallel reaction monitoring for Gag::eGFP peptides in Sf9 cell line cultured in ExpiSf™ CD Medium. All analyzed peptides presented charge 2+. The number of replicates is denoted as "n". Coefficient of variation was calculated as %CV = $100 \cdot SD / \bar{x}$.

		ETINEEAAEWDR			FAVNPGLLETSEGCR			TLNAWVK	
n	fmol	Mean Peak Area	%CV	fmol	Mean Peak Area	%CV	fmol	Mean Peak Area	%CV
4	0	0	-	0	0	-	0	0	-
3	10	6.77E+06	23.56	10	1.77E+07	27.46	10	-	-
3	100	4.60E+07	21.50	100	1.04E+08	13.17	100	5.14E+07	n=1[a]
3	250	1.46E+08	n=1[a]	250	1.71E+08	18.23	250	1.21E+08	n=1[a]
5	500	2.80E+08	5.10	500	3.25E+08	53.51	500	1.49E+08	n=1[a]
3	1000	5.06E+08	18.52	1000	2.06E+09	45.31	1000	4.66E+08	38.13
3	2000	8.29E+08	n=1[a]	2000	3.48E+09	26.82	2000	1.06E+09	10.32
3	5000	1.86E+09	n=1[a]	5000	9.01E+09	16.86	5000	2.56E+09	13.53
	n	Monomers/VLP	St. Error	n	Monomers/VLP	St. Error	n	Monomers/VLP	St. Error
Endogenous	24	4008	286	36	3548	279	17	3404	407

(a) Only one replicate was obtained for these measurements

mathematical restriction was imposed. The ratio monomers/volume was assumed constant for both subpopulations and **Eq.1** was used for the calculation of number of monomers considering the percentages of each subpopulation. A total of 2184 ± 179 and 6126 ± 502 monomers per particle were calculated for each subpopulation respectively (**Figure 5D**). Both maintaining a ratio monomers/nm^3 of $(5.6 \pm 0.5) \cdot 10^{-4}$ significantly lower ($p<0.0001$) than the one obtained for VLPs from Sf9 cells and HEK293 cell-based platforms.

Visualization of Gag::eGFP VLP structures by electron cryo-tomography

To further characterize the structure and distribution of Gag::eGFP within the VLP, electron cryo-microscopy visualizations were performed to define the Gag::eGFP shell and the configuration of the monomers. Samples of VLPs produced by TGE in HEK293 cells after purification by sucrose cushion were taken to be observed by electron cryo-microscopy. Morphology and size of VLPs were analyzed to determine structural changes and characteristics. Individual VLPs (N=115) were analyzed, presenting a population composed of 91.3% of spherical particles of 140-145 nm of diameter, 2.6% and 6.1% of smaller and larger than 140 nm respectively. The size of 140-145 nm agreed with previous results reported of VLP production in HEK293 by TGE[13,19,27,30]. VLPs presented a clear ordered ring-shaped structure underneath the plasmatic membrane which corresponded to concentric shells formed by Gag::eGFP oligomerization (**Figure 6A**). The lipid bilayer was well resolved in the cryo-micrographs, being able to perfectly distinguish both membrane leaflets along the different tomogram slices and also in the density profile (**Figure 6B**). Density profile analyses of the series of concentric rings from Gag::eGFP shell allowed the determination of different peaks corresponding to the different domains from which Gag polyprotein is composed: MA, CA and NC. The presence of eGFP is attributed to the peak following the previous domains. Measurements of different density profiles showed that the plasmatic membrane is 4.6 ± 0.3 nm and the Gag series of layers about 10 ± 2 nm, in agreement with previous data[12,31,32]. The measurements regarding eGFP also agreed with previously reported data showing a size of 4.2 nm[33]. This makes the complete Gag::eGFP shell ~16 nm. The 84 kDa Gag::eGFP polyprotein is illustrated in **Figure 6C** as a structural schematic together with a model of the analyzed VLPs showing the concentric ring observed in the cryo-micrographs. The
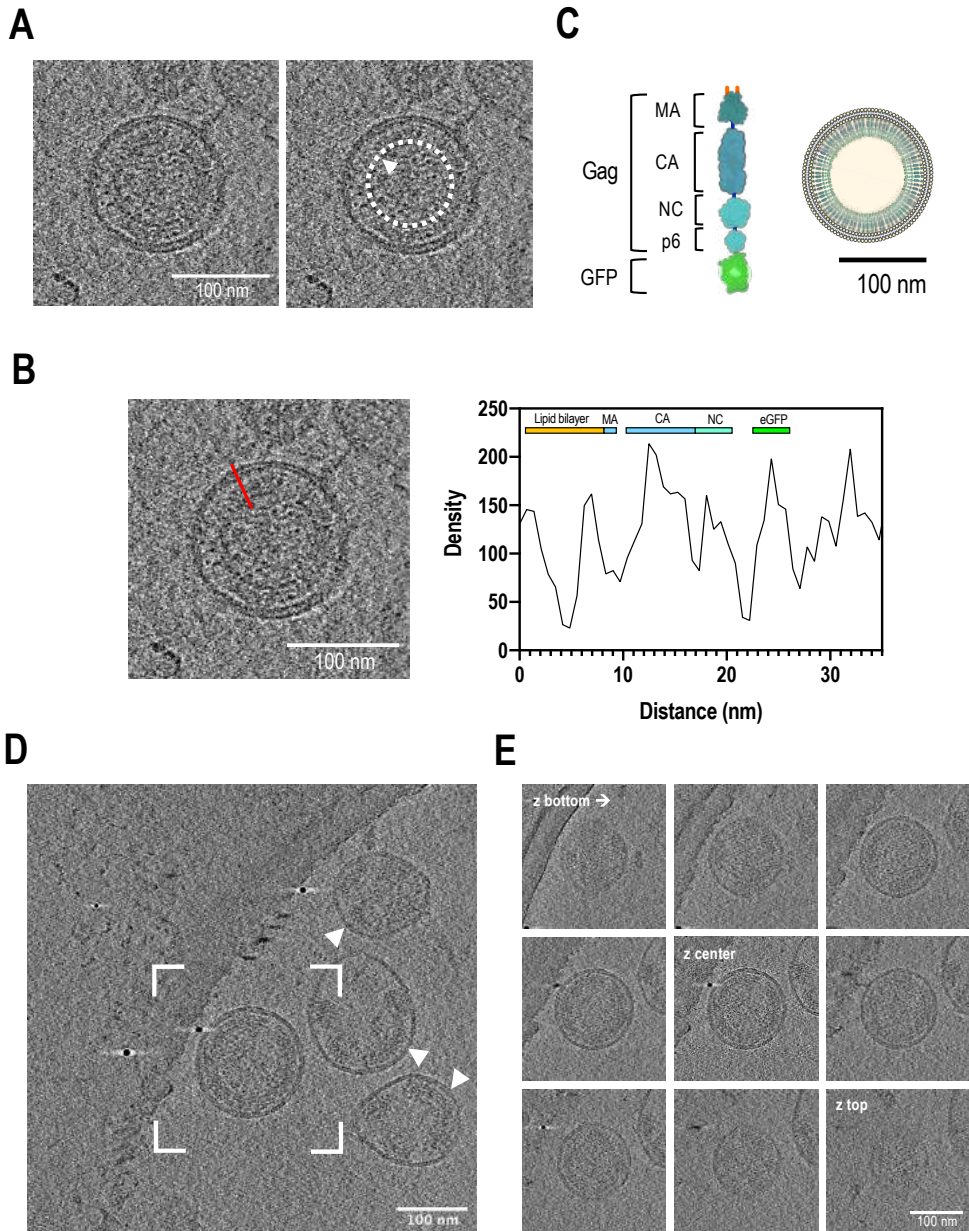
Figure 6. Electron cryo-tomograms of purified HEK293 VLPs produced by TGE. **A)** (Left) Central slice from a tomogram containing a VLP clearly showing an ordered structure under the cell membrane. (Right) White arrow and dotted line pointing to the boundary of the shell underneath the membrane. **B)** Density profile of the Gag::eGFP shell. (Left) The red line indicates the example area from where VLP slices were analyzed. (Right) Representative density plot showing density peaks in the micrograph corresponding to the different parts of the Gag::eGFP polyprotein. **C)** Structural schematics for the Gag::eGFP shell in Gag VLPs. MA, CA, NC and p6 are the different domains of Gag polyprotein. **D)** Central slice form a tomogram containing 4 different VLPs. Bracketed VLP presents the whole Gag shell. The rest three VLPs contain partial Gag shells forming patches and not covering the whole membrane. White arrows indicate areas of Gag oligomerization underneath the cell membrane. The membrane of the lower VLP is noticeably broken, contributing to release free unassembled monomers. **E)** 2D image slices as iterative z heights from the bottom, the center and to the top surface from the bracketed VLP in panel C.
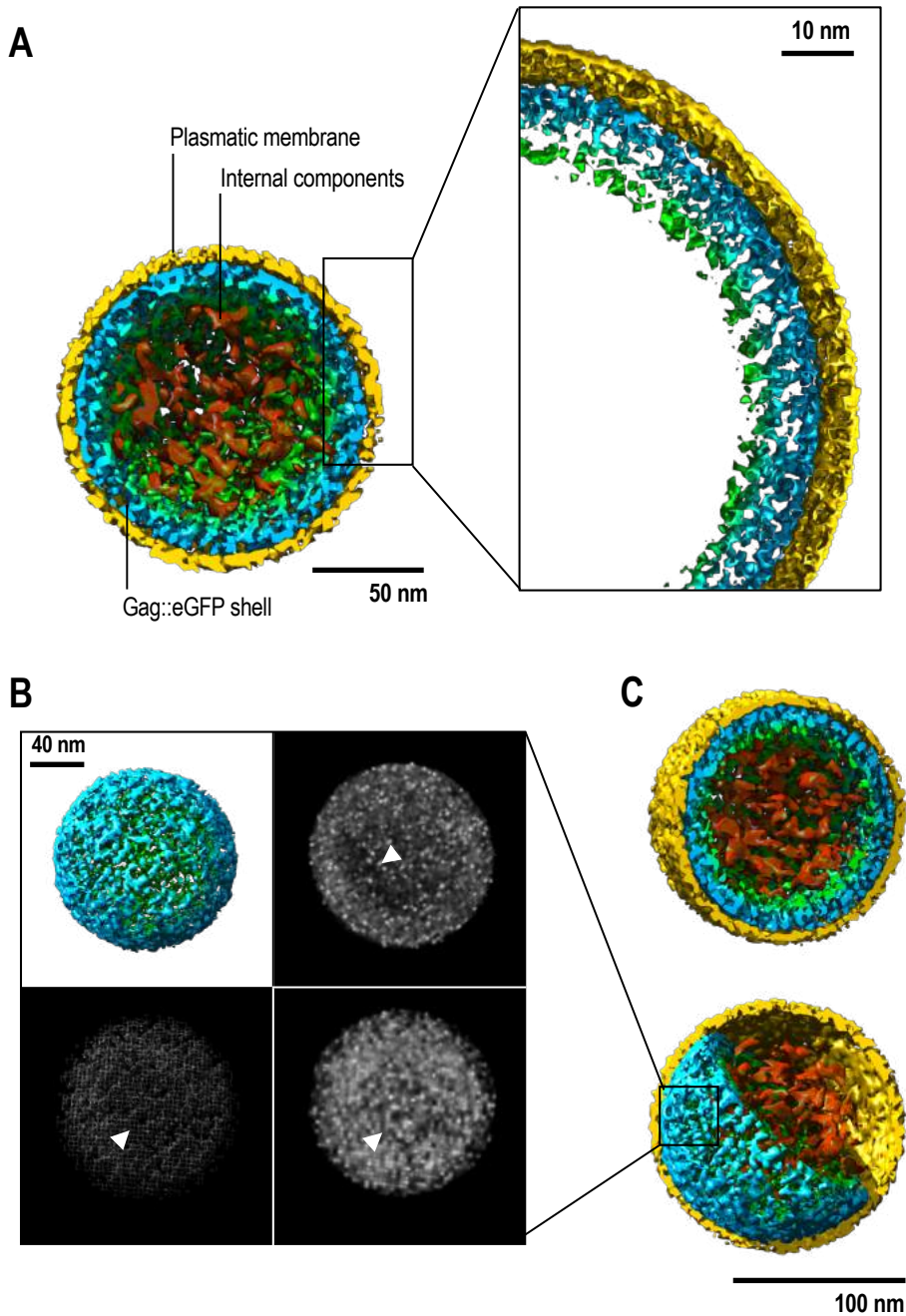
Figure 7. 3D reconstruction and analysis of HEK293 VLP produced by TGE from electron cryo-tomography. **A)** VLP cross section showing the plasmatic membrane on yellow, internal components in red and the Gag::eGFP shell in a gradient of blue (Gag) to green (eGFP). This color coding is maintained for the rest of the panels. The lipid bilayer and Gag::eGFP monomers can be observed in the detailed zoomed area. **B)** Analysis of the Gag::eGFP shell. Upper left panel shows the 3D reconstruction of Gag::eGFP shell. Upper left panel showing a section of the Gag::eGFP shell density map. White arrow pointing towards a less dense area of the shell. Lower left and right panels show a different section of the density maps and white arrows point to the 8 nm spacing characteristic of Gag hexagonal lattice. Representations of solid (right) and mesh surface model (left). **C)** Axonometric view of a VLP.

analyzed VLPs did not show a clear hexagonal lattice in the outer MA shell in most of the VLP surface area, again in agreement with reported data of immature HIV virions analysis[12]. Broken VLP membranes can be observed in some cryo-micrographs, contributing to the release of unassembled free Gag::eGFP monomers. These VLPs showed patches of clear ordered Gag oligomerization, not covering the whole surface underneath the membrane (**Figure 6D**) and counted for ~10% of all analyzed VLPs. To further study the conformation of the Gag::eGFP shell, from all the cryo-micrograph taken, 25 were selected to reconstruct a 3D volume (tomogram). In order to generate the tomogram, a series of cryo-micrographs were taken at sequential varying angles (from -60° to 60°) slicing through the volume (**Figure 6E**). From these 25 tomograms, 11 VLPs were selected and further segmented to allow rendering of 3D surface volume reconstructions (**Supplementary Figure S4 A-C**) to visualize a structural model (**Figure 7A**). This reconstruction allowed the visualization of individual Gag::eGFP monomers within the Gag::eGFP shell, anchored to the plasmatic membrane and pointing towards the center of the particle. Segmentation and reconstruction also allowed to identify and visualize internal components of the VLP. All analyzed VLPs contained material inside the particle, visually showing a relative constant amount. This material was observed to be irregularly distributed in the volume contained within the Gag::eGFP shell. The lack of a pattern in the distribution of the internal components was also constant in all analyzed particles. Segmentation also allowed the analysis in detail of the Gag::eGFP shell showing that the characteristic hexagonal lattice with a spacing of 8nm[7,12,34] was not present in the majority of the volume surface, only in rare occasions this structure was observed in the complete reconstructed shell. Most of the volume surface of Gag::eGFP distribution was irregular, with density points randomly distributed over the surface showing no clear periodic order. Some areas of lower densities in the volume surface were observed when analyzing different cross sections of the Gag::eGFP shell (**Figure 7B**). VLP renderings offered a detailed model of the Gag::eGFP particle (**Figure 7C**) and monomers distribution.

DISCUSSION

Advances in vaccine development are crucial nowadays when new virus outbreaks can jeopardize the world in a matter of months. The design of versatile vaccine-production platforms that can rapidly adapt for the generation of new vaccine candidates and their subsequent implementation at industrial scale is essential to ensure global access to new vaccines. HIV Gag VLPs are a robust platform proven successful for pseudotyping engineering against different viruses like Influenza, dengue, HPV or equine herpes virus[35–38]. The variability in size observed when producing Gag VLPs ranging from 100-200 nm and being dependent on the cell line produced presents a problem for the implementation of conventional process analytical technologies (PAT) to control the production bioprocess at larger scale. Quality control of the produced VLP titers requires knowledge of the number of Gag molecules per particles for a successful quantification. This work describes a fast, accurate, direct and mass-spectrometry-based method to quantify the number of monomers present in a VLP, being the first analytical method for detailed absolute Gag monomers quantification using three peptides from Gag polyprotein, one from p17 and two from p24 domain. Although fluorescence emission of eGFP was used for the calculation of the number of monomers (**Eq. 1**), it was only used as a normalization or weighing factor to compensate the contribution of the unassembled free monomer. Absolute fluorescence intensity was not used or compared between platforms. This method has been applied to compare the particle size and monomer composition of Gag VLPs produced in different cell-based platforms (**Table 3**). Since all previously reported works about monomers calculation for Gag VLPs described the variability based on CA lattice characteristics or quantification of other HIV-1 viral components, a constant ratio of monomer/volume was assumed, larger particles would have a higher number of monomers[7,39–41]. However, comparing the composition and stoichiometry of Gag VLPs produced in different cell-based platforms, the ratio monomer/volume was not constant. In fact, it has been found to depend on the cell line and the methodology used for their production (**Figure 8**). Interestingly, cell-based platforms with a lower particle production rate showed a higher monomer/volume ratio, like HEK293 SGE conditions, while platforms that produced a significantly higher number of total particles, like High Five insect cells, showed a lower ratio. The molecular and

biochemical reason for this difference remains unknown. Nevertheless, we hypothesize that the extracellular vesicle production rate might be the driving force for this difference. While HEK293 SGE-based platforms produced fewer total particles in a given period of time, allowing a highly dense packing of Gag molecules, High Five cell line which produced an abundant number of total particles in the same time, probably accounting for the lower monomer/volume ratio in their VLPs. The internal structure and distribution of Gag::eGFP monomers has also been explored in detail for the first time. The interaction of eGFP in the formation of Gag VLPs was recently reported to affect the monomer distribution and arrangement in the VLP[27]. However, when compared to structural analysis of Gag VLPs[5,12,42], Gag::eGFP VLPs analyzed via electron cryo-tomography in this work presented the same organization and arrangement, not altering the VLP structure. The interference in the ordered disposition of monomers when using Gag::eGFP observed in previous reported works may be due to low resolution or to the analysis of disordered and irregular VLPs, totally lacking the ring-shaped structure beneath the lipid bilayer which also contributed to the total population of VLPs in a ~8% in this work but it is also present in preparations of Gag VLPs, lacking eGFP. This demonstrates that the presence of some irregular VLPs does not depend on the use of the reporter eGFP and it is inherent to cell-based VLP production.

Table 3. Summary of results of Gag::eGFP monomers per VLP in different production platforms. Monomer/nm^3 errors calculated by partial derivatives. The number of replicates is denoted as "n".

	Number of monomers	n	nm Diameter	Monomer / nm^3
HEK293 TGE (HyCell)	3762 ± 144	5	145 ± 1	$(2.36 ± 0.09){\cdot}10^{-3}$
HEK293 TGE (F17)	3786 ± 136	5	140 ± 3	$(2.6 ± 0.1){\cdot}10^{-3}$
HEK293 TGE (FreeMix)	3604 ± 63	79	145 ± 4	$(2.4 ± 0.5){\cdot}10^{-3}$
HEK293 SGE (Random)	2220 ± 244	8	109 ± 6	$(3.3 ± 0.4){\cdot}10^{-3}$
HEK293 SGE (RCME)	3014 ± 232	10	123 ± 1	$(3.1 ± 0.2){\cdot}10^{-3}$
Sf9	3755 ± 174	77	145 ± 2	$(2.4 ± 0.1){\cdot}10^{-3}$
High Five (Subpopulation A)	2184 ± 179	6	195 ± 30	$(5.6 ± 0.5){\cdot}10^{-4}$
High Five (Subpopulation B)	6126 ± 502	6	275 ± 30	$(5.6 ± 0.5){\cdot}10^{-4}$
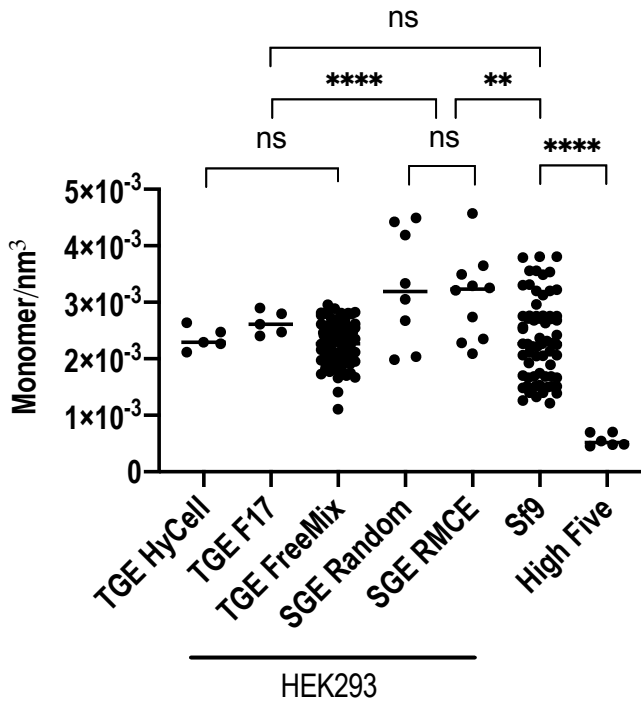
Figure 8. Summary of the different stoichiometry for Gag::eGFP molecules per VLP in different production platforms. The ratio of monomer/nm^3 is represented for each studied condition. Significances calculated by Mann–Whitney *U test*.

CONCLUSIONS

In this work, three Gag peptides, ETINEEAAEWDR, FAVNPGLLETTSEGCR, and TLNAWVK have been validated for mass spectrometry quantification and analysis of Gag stoichiometry for the main cell-based production platforms for HIV-1 VLPs. The results of the reference production platform, HEK293, producing VLPs by TGE, indicate that VLPs with an average size of 145 nm contain 3617 ± 17 monomers of Gag::eGFP. The monomer/volume ratio was constant in HEK293 producing via TGE regardless of the culture media and changed when producing via SGE, obtaining a higher ratio. In the case of Sf9 insect cells, when producing via the BV-infection methodology, VLPs contained 3755 ± 174 monomers of Gag::eGFP. In this case, the monomer/volume ratio was the same than the HEK293 TGE system. For High Five cells, the ratio was significantly lower and a different distribution of VLP size was observed shifting from the average size 145 nm to 195 and 275 nm. Regarding the Gag::eGFP model for VLP characterization, the reporter eGFP did not interfere in VLP formation and structure. VLPs produced in HEK293 showed a consistent pattern, strengthen their value for further development of HIV-1 Gag VLPs in vaccine technology.

DATA AVAILABILITY

The raw mass spectrometry data has been submitted to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) with the dataset identifier PXD021278.

SUPPORTING INFORMATION

Supplementary Figure S1: Gag::eGFP sequence. Selected peptides for PRM colored in blue, spacer sequence underlined and eGFP sequence in green. MA p17 domain from amino acid 1-132. CA p24 domain from amino acid 132-363.

1	MGARASVLSG	GELDRWEKIR	LRPGGKKKYK	LKHIVWASRE	LERFAVNPGL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN
81	TVATLYCVHQ	RIEIKDTKEA	LDKIEEEQNK	SKKKAQQAAA	DTGHSNQVSQ	NYPIVQNIQG	QMVHQAISPR	TLNAWVKVVE
161	EKAFSPEVIP	MFSALSEGAT	PQDLNTMLNT	VGGHQAAMQM	LKETINEEAA	EWDRVHPVHA	GPIAPGQMRE	PRGSDIAGTT
241	STLQEQIGWM	TNNPPIPVGE	IYKRWIILGL	NKIVRMYSPT	SILDIRQGPK	EPFRDYVDRF	YKTLRAEQAS	QEVKNWMTET
321	LLVQNANPDC	KTILKALGPA	ATLEEMMTAC	QGVGGPGHKA	RVLAEAMSQV	TNSATIMMQR	GNFRNQRKIV	KCFNCGKEGH
401	TARNCRAPRK	KGCWKCGKEG	HQMKDCTERQ	ANFLGKIWPS	YKGRPGNFLQ	SRPEPTAPPE	ESFRSGVETT	TPPQKQEPID
481	KELYPLTSLR	SLFGNDPSSQ	NRNGDPPVAT	MVSKGEELFT	GVVPILVELD	GDVNGHKFSV	SGEGEGDATY	GKLTLKFICT
561	TGKLPVPWPT	LVTTLTYGVQ	CFSRYPDHMK	QHDFFKSAMP	EGYVQERTIF	FKDDGNYKTR	AEVKFEGDTL	VNRIELKGID
641	FKEDGNILGH	KLEYNYNSHN	VYIMADKQKN	GIKVNFKIRH	NIEDGSVQLA	DHYQQNTPIG	DGPVLLPDNH	YLSTQSALSK
721	DPNEKRDHMV	LLEFVTAAGI	TLGMDELYK					

Supplementary Figure S2: **A)** Chromatograms for each analyzed peptide. Peptides labelled with with ^{13}C and ^{14}N in Lys and Arg are marked with (*). Peptide E: ETINEEAAEWDR, peptide F: FAVNPGLLETTSEGCR, peptide T: TLNAWVK. On top of each chromatographic peak there is the retention time. m/z values from each peptide is marked in its corresponding chromatogram. **B)** Fragmentation pattern of each monitored peptide.
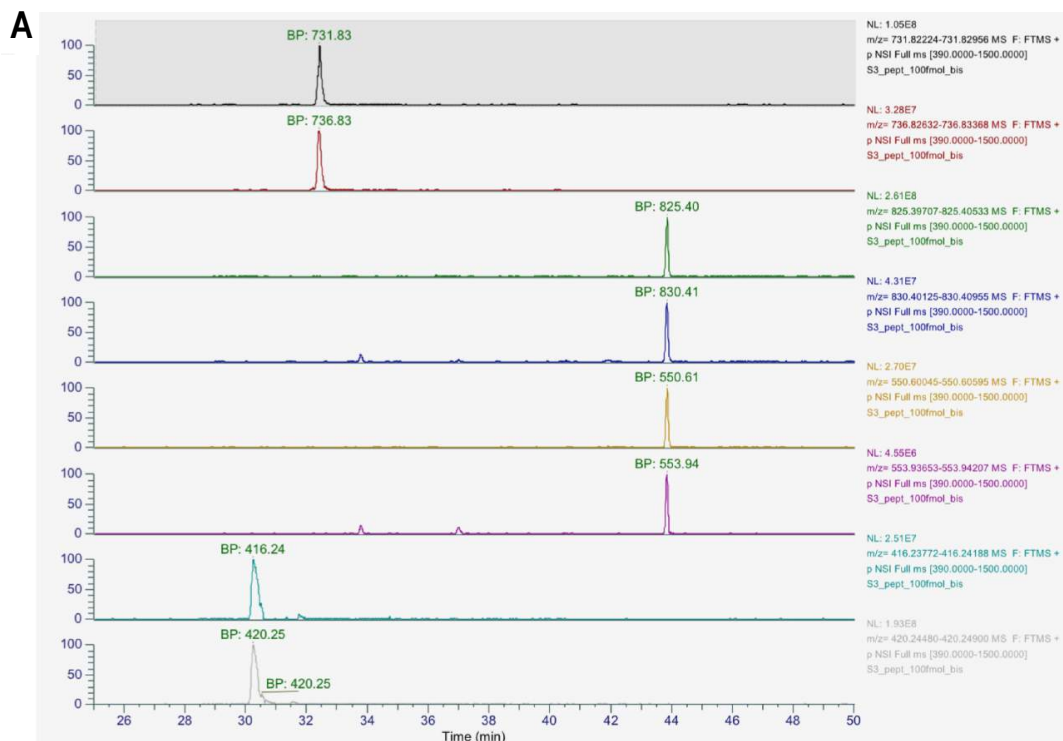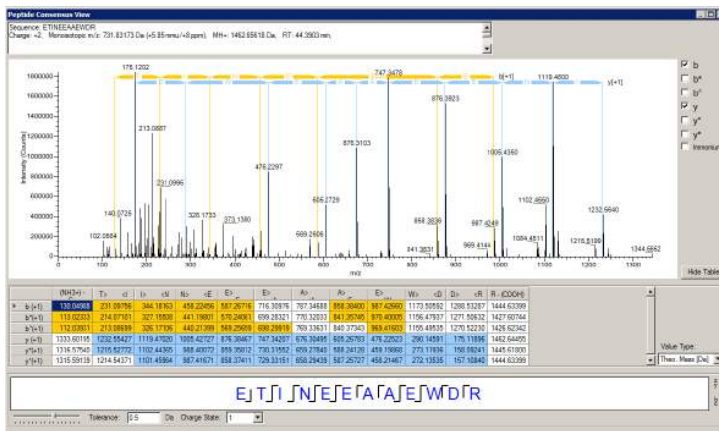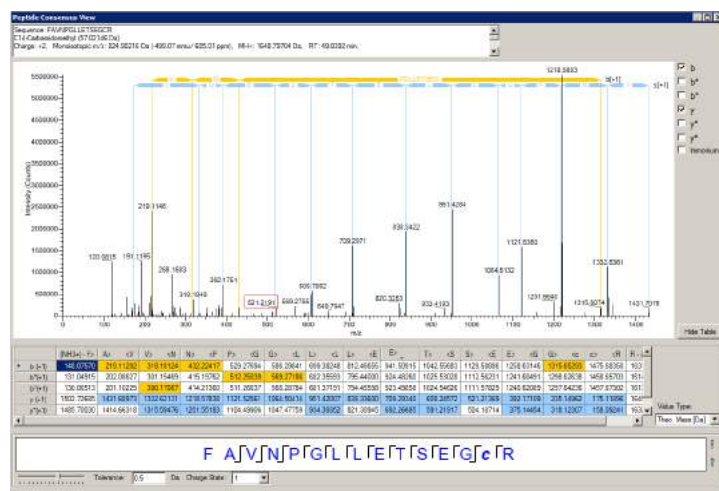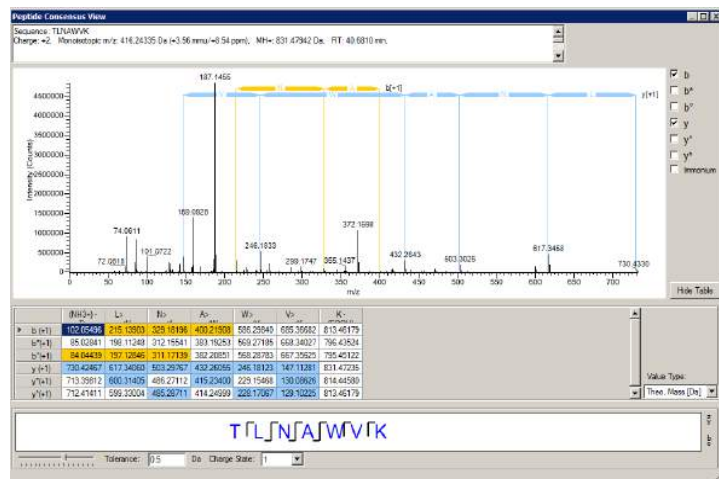
Supplementary Figure S3: Validation of parallel reaction monitoring (PRM) analysis using HyCell and FreeStyle® F17 culture media. **A)** NTA size distribution curves for fluorescent particles produced by transient gene expression (TGE), in HyCell media. **B)** NTA size distribution curves for fluorescent particles produced by transient gene expression (TGE), in F17 media. **C)** Number of Gag::eGFP molecules (monomers) per VLP in each condition calculated by PRM. Standard error is calculated as ε=SD/$\sqrt{n}$. Estimation of number of monomers per nm of diameter. Errors calculated by error propagation theory using partial derivatives.
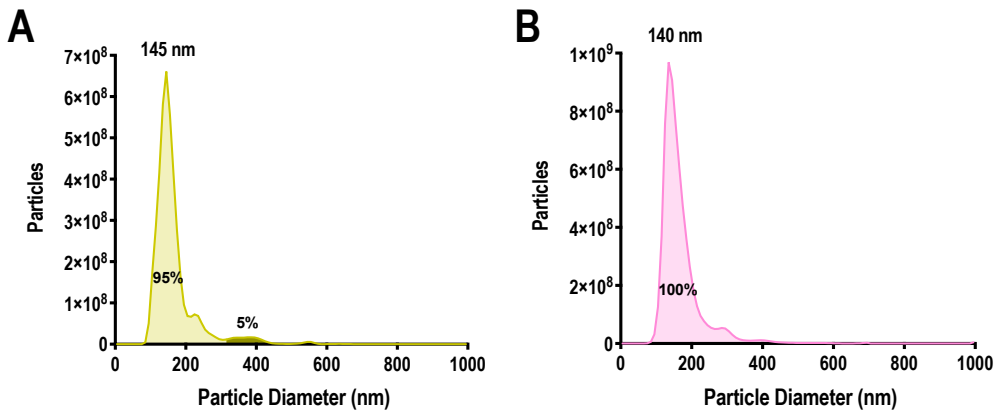
	Number of monomers	n	nm Diameter	Monomer / nm^3
HEK293 TGE (HyCell)	3762 ± 144	5	145 ± 1	(2.36 ± 0.09) ·10^{-3}
HEK293 TGE (F17)	3786 ± 136	5	140 ± 3	(2.6 ± 0.1) ·10^{-3}

Supplementary Figure S4: **(A-C)** Process in 3D modelling for VLPS. **A)** Central slice of a tomogram generated by electron cryo-tomography. **B)** Segmentation of the different areas. **C)** 3D surface volume reconstruction. **D)** Complete zoomed section of a VLP. Lipid bilayer in yellow, Gag::eGFP shell showing Gag shell in blue and eGFP in green.
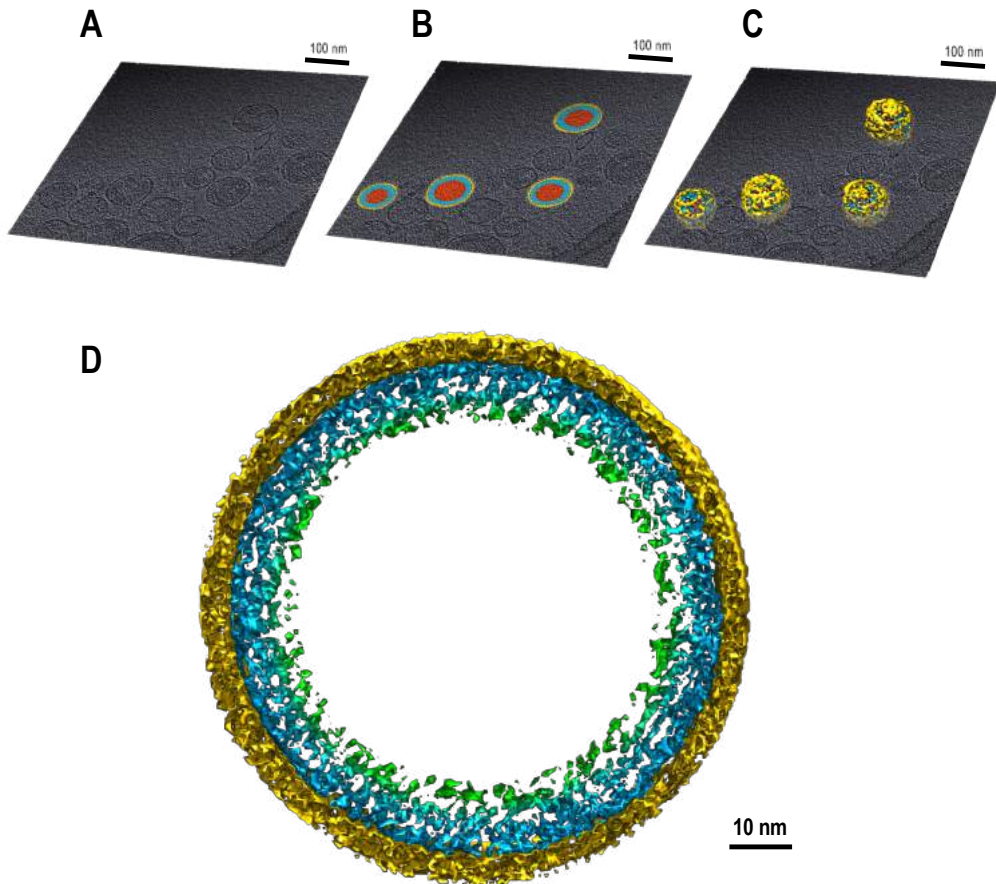
Supplementary Table S1: List of all 77 previously identified Gag peptides. Raw mass spectrometry data is available via ProteomeXchange with identifier PXD014746. Methionine oxidation denoted as M# and TMT-labelled lysine denoted as K@. The three selected peptides for PRM are highlighted in light green.

https://onlinelibrary.wiley.com/doi/abs/10.1002/bit.27786

Supplementary Table S2. Monoisotopic *m/z* masses of the analyzed petide ions depending on their charges. R* and K* represent ^{13}C and ^{14}N labelled Arg and Lys. Monitored masses are highlithed in light green.

	Peptide	mono 1+	mono 2+	mono 3+
Endogenous	ETINEEAAEWDR	1462.645	731.8259	488.2197
	FAVNPGLLETSEGCR	1649.795	825.4012	550.6032
	TLNAWVK	831.4723	416.2389	277.829
Labelled	ETINEEAAEWDR*	1472.653	736.83	491.5558
	FAVNPGLLETSEGCR*	1659.804	830.4054	553.9393
	TLNAWVK*	839.4865	420.2469	280.5004

REFERENCES

1. Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

2. von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

3. Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136–144 (2013).

4. Könnyű, B. *et al.* Gag-Pol Processing during HIV-1 Virion Maturation: A Systems Biology Approach. *PLoS Comput. Biol.* **9**, e1003103 (2013).

5. Benjamin, J., Ganser-Pornillos, B. K., Tivol, W. F., Sundquist, W. I. & Jensen, G. J. Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography. *J. Mol. Biol.* **346**, 577–588 (2005).

6. Chen, Y., Wu, B., Musier-Forsyth, K., Mansky, L. M. & Mueller, J. D. Fluorescence fluctuation spectroscopy on viral-like particles reveals variable Gag stoichiometry. *Biophys. J.* **96**, 1961–1969 (2009).

7. Briggs, J. A. G. *et al.* The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–675 (2004).

8. Layne, S. P. *et al.* Factors underlying spontaneous inactivation and susceptibility to neutralization of human immunodeficiency virus. *Virology* **189**, 695–714 (1992).

9. Jouvenet, N., Simon, S. M. & Bieniasz, P. D. Visualizing HIV-1 assembly. *Journal of Molecular Biology* **410**, 501–511 (2011).

10. Piatak, M. *et al.* High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science (80-.).* **259**, 1749–1754 (1993).

11. Zhu, P. *et al.* Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15812–15817 (2003).

12. Wright, E. R. *et al.* Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J.* **26**, 2218–2226 (2007).

13. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

14. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol* **166**, 152–165 (2013).

15. Puente-Massaguer, E., Lecina, M. & Gòdia, F. Integrating nanoparticle quantification and statistical design of experiments for efficient HIV-1 virus-like particle production in High Five cells. *Appl. Microbiol. Biotechnol.* **104**, 1569–1582 (2020).

16. Lavado-García, J., Cervera, L. & Gòdia, F. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures. *Front. Bioeng. Biotechnol.* **8**, 617 (2020).

17. Lavado-García, J. *et al.* Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus–like particles. *Biotechnol. Bioeng.* (2021). doi:10.1002/bit.27679

18. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

19. Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed Quantitative Proteomic Analysis of HEK293 Provides Insights into Molecular Changes Associated with the Cell Density Effect, Transient Transfection, and Virus-Like Particle Production. *J. Proteome Res.* **19**, 1085–1099 (2020).

20. Lavado-García, J. *et al.* Molecular Characterization of the Coproduced Extracellular Vesicles in HEK293 during Virus-Like Particle Production. *J. Proteome Res.* **19**, 4532 (2020).

21. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).

22. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–5 (2012).

23. Mastronarde, D. N. Correction for non-perpendicularity of beam and tilt axis in tomographic

reconstructions with the IMOD package. in *Journal of Microscopy* **230**, 212–217 (J Microsc, 2008).

24. J. I. Agulleiro, J. J. F. Fast tomographic reconstruction on multicore computers. *Bioinformatics* **27**, 582–583

25. Fernandez, J. J. TOMOBFLOW: Feature-preserving noise filtering for electron tomography. *BMC Bioinformatics* **10**, 178 (2009).

26. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

27. González-Domínguez, I., Puente-Massaguer, E., Cervera, L. & Gòdia, F. Quality assessment of virus-like particles at single particle level: A comparative study. *Viruses* **12**, (2020).

28. González-Domínguez, I., Puente-Massaguer, E., Cervera, L. & Gòdia, F. Quantification of the HIV-1 virus-like particle production process by super-resolution imaging: From VLP budding to nanoparticle analysis. *Biotechnol. Bioeng.* **117**, 1929–1945 (2020).

29. Puente-Massaguer, E., Lecina, M. & Gòdia, F. Nanoscale characterization coupled to multi-parametric optimization of Hi5 cell transient gene expression. *Appl. Microbiol. Biotechnol.* **102**, 10495–10510 (2018).

30. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).

31. Fuller, S. D., Wilk, T., Gowen, B. E., Kräusslich, H. G. & Vogt, V. M. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.* **7**, 729–738 (1997).

32. Wilk, T. *et al.* Organization of Immature Human Immunodeficiency Virus Type 1. *J. Virol.* **75**, 759–771 (2001).

33. Hink, M. A. *et al.* Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *J. Biol. Chem.* **275**, 17556–17560 (2000).

34. Nermut, M. V. *et al.* Further evidence for hexagonal organization of HIV gag protein in prebudding assemblies and immature virus-like particles. *J. Struct. Biol.* **123**, 143–149 (1998).

35. Venereo-Sanchez, A. *et al.* Hemagglutinin and neuraminidase containing virus-like particles

produced in HEK-293 suspension culture: An effective influenza vaccine candidate. *Vaccine* **34**, 3371–3380 (2016).

36. Chua, A. J. *et al.* A novel platform for virus-like particle-display of flaviviral envelope domain III: Induction of Dengue and West Nile virus neutralizing antibodies. *Virol. J.* **10**, 129 (2013).

37. Di Bonito, P. *et al.* Anti-tumor CD8+ T cell immunity elicited by HIV-1-based virus-like particles incorporating HPV-16 E7 protein. *Virology* **395**, 45–55 (2009).

38. Osterrieder, N., Wagner, R., Brandmüller, C., Schmidt, P. & Wolf, H. Protection against EHV-1 challenge infection in the murine model after vaccination with various formulations of recombinant glycoprotein gp14 (gB). *Virology* **208**, 500–510 (1995).

39. Müller, B., Tessmer, U., Schubert, U. & Kräusslich, H.-G. Human Immunodeficiency Virus Type 1 Vpr Protein Is Incorporated into the Virion in Significantly Smaller Amounts than Gag and Is Phosphorylated in Infected Cells. *J. Virol.* **74**, 9727–9731 (2000).

40. Franke, E. K., Yuan, H. E. H. & Luban, J. Specific incorporation of cyclophilin a into HIV-1 virions. *Nature* **372**, 359–362 (1994).

41. Welker, R., Hohenberg, H., Tessmer, U., Huckhagel, C. & Kräusslich, H.-G. Biochemical and Structural Analysis of Isolated Mature Cores of Human Immunodeficiency Virus Type 1. *J. Virol.* **74**, 1168–1177 (2000).

42. Woodward, C. L., Cheng, S. N. & Jensen, G. J. Electron cryotomography studies of maturing HIV-1 particles reveal the assembly pathway of the viral core. *J. Virol.* **89**, 1267–77 (2015).

CHAPTER FIVE

Molecular characterization of the coproduced extracellular vesicles in HEK293 during virus-like particle production

ABSTRACT

Vaccine therapies based on virus-like particles (VLPs) are currently in the spotlight due to their potential for generating high immunogenic responses while presenting fewer side effects than conventional vaccines. These self-assembled nanostructures resemble the native conformation of the virus but lack genetic material. They are becoming a promising platform for vaccine candidates against several diseases due to the ability of modifying their membrane with antigens from different viruses. The coproduction of extracellular vesicles (EVs) when producing VLPs is a key phenomenon currently still under study. In order to characterize this extracellular environment, a quantitative proteomics approach has been carried out. Three conditions were studied: non-transfected, transfected with an empty plasmid as control, and transfected with a plasmid coding for HIV-1 Gag polyprotein. A shift in EV biogenesis has been detected upon transfection, changing the production from large to small EVs. Another remarkable trait found was the presence of DNA being secreted within vesicles smaller than 200 nm. Studying the protein profile of these biological nanocarriers, it was observed that EVs were reflecting an overall energy homeostasis disruption via mitochondrial protein deregulation. Also, immunomodulatory proteins like ITGB1, ENO3, and PRDX5 were identified and quantified in VLP and EV fractions. These findings provide insight on the nature of the VLP extracellular environment defining the characteristics and protein profile of EVs, with potential to develop new downstream separation strategies or using them as adjuvants in viral therapies.

ABBREVIATIONS

BH: Benjamini−Hochberg, **DTT**: dithiothreitol, **EDTA**: ethylenediaminetetraacetic acid, **eGFP**: enhanced green fluorescence protein, **ESCRT**: endosomal sorting complex required for transport, **EVs**: extracellular vesicles, **F17**: FreeStyle F17 cell culture medium, **FASP**: filter aided sample preparation, **Gag::eGFP**: translational fusion of HIV-Gag protein and eGFP, **hpt**: hours post transfection, **ILV**: intralumenal vesicles, **LC-MS**: liquid chromatography coupled with tandem mass spectrometry, **lEVs**: large extracellular vesicles, **MVs**: micro- vesicles, **MVBs**: multivesicular bodies, **MVA**: Modified Vaccinia Ankara, **PBS**: phosphate-buffered saline, **PEI**: polyethylenimine, **RFUs**: relative fluorescence units, **SBT**: Systems Biology Triangle, **SDS**: sodium dodecyl sulfate, **sEVs**: small extracellular vesicles, **TEAB**: tetraethylammonium tetrahydroborate, **TFA**: trifluoroacetic acid, **TMT**: tandem mass tag, **VLPs**: virus-like particles.

INTRODUCTION

The manufacturing of viral particles for vaccine development is currently an important field of research to explore novel therapies for emerging diseases. Virus-like particles (VLPs) are nanostructures, typically composed of recombinant viral proteins whose outer conformation is identical to the native virus structure but lacking genetic material, hence being noninfectious. Thus, creating a versatile platform for the generation of vaccines. The recombinant expression of HIV-1 Gag polyprotein in human cell cultures like HEK293, is enough for its self-assembly and release as enveloped VLPs[1]. Gag VLPs are a good platform for the development of vaccines for several diseases. These structures can be used as a scaffold to add immunogenic epitopes decorating its cell membrane, what is known as pseudotyping[2]. Another important fact in the production of VLPs is the coproduction of extracellular vesicles (EVs), naturally secreted by cell cultures. EVs are cell membrane-derived nanovesicles that have recently become a major target of research due to their hitherto unknown extent of implication in many diseases[3-6]. EVs are a heterogeneous population of vesicles produced by all cell types[7] to deliver cargos, send signals, excrete harmful materials, maintain homeostasis and regulate many biological processes[8-11]. This phenomenon can be regarded as a way of cell-to-cell communication[12,13]. When EVs are produced, they leave the cell taking part of the its

membrane, similar to Gag VLPs. When reaching the recipient cell, both cell membranes fuse and the contents of the vesicle enter the cytosol and the membrane of the vesicle becomes part of the recipient cell membrane. Through this process cells can exchange all sort of biomolecules such as DNA, RNA, transcription factors, transport proteins, tetraspanins, heat shock proteins, lipid-associated proteins, cytokines, metabolic enzymes…etc[7,14,15]. It is widely reported that different kinds of EVs are used by the cell for different purposes[16,17]. An efficient separation of EVs from VLPs is required to obtain high purity VLP preparations. However, both VLPs and EVs are membrane-bound nanoparticles sharing the same physicochemical properties, including density, and displaying the same families of membrane proteins[18-20], which strongly difficult their separation. Rounds of centrifugation and density gradients are the current methods for VLPs and EVs separation but still do not reach good enough purity for vaccine standards[18]. Up to now, there is no previous work on the molecular characterization of the EVs subpopulations copurified with VLPs. Characterizing the protein profile of these EVs subpopulations will help cast some light on potential separation approaches. Apart from helping develop a strategy for EV separation, VLP-like EVs could be used as an adjuvant, modifying the membrane composition to promote an immunogenic response or even presenting new antigens via membrane engineering. Furthermore, these EVs could be useful as nanocages for drug delivery via metabolic engineering. In this work, a proteomic approach is used to analyze the copurified subpopulations of EVs when producing VLPs and characterize the protein profiles as well as the changes in EV generation when comparing non-transfected and transfected conditions in order to understand the cellular reaction to the production of VLPs regarding EVs.

MATERIALS AND METHODS

HEK 293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK 293 cell line (HEK293SF-3F6, NRC, Canada) kindly provided by Dr. Amine Kamen from McGill University (Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning®) at 37°C, 5% of CO2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Culture medium was

FreeStyle™ F17 Expression Medium (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) supplemented with 8 mM GlutaMAX™ (Gibco, Life Technologies), 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, Life Technologies) and IGF-1 at a final concentration of 50 g/L.

Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

TRANSIENT TRANSFECTION

Transfections were carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 µg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture medium (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used.

The plasmid used contained a gene coding for HIV-Gag protein fused to eGFP (Gag::eGFP). Briefly, pGag::eGFP plasmid is diluted with supplemented FreeStyle™ F17 medium and vortexed for 10 seconds. As a transfection control, a plasmid sharing the same backbone but lacking the Gag::eGFP gene was used and noted as mock. Then PEI is added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, then the mixture is incubated for 15 min at room temperature and then added to the cell culture.

ULTRACENTRIFUGATION

VLP containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Then, concentrated and purified HIV-1 Gag VLPs were obtained by double cushion ultracentrifugation. Briefly, a volume of 15 mL of clarified supernatant from every condition was layered on top of a 25% (w/v) sucrose cushion (5 mL) and 60% (w/v) sucrose cushion (8 mL) and centrifuged at 31.000 rpm for 2.5h at 4°C using a SW32 rotor in a Beckman Optima L100XP centrifuge. Ultracentrifuge tubes were filled with PBS (Hyclone, GE HeathCare, Chicago, IL, USA). The 25-60% sucrose interphase was extracted for each condition. The concentrated material was stored at −80°C for future studies.

HIV-1 GAG VLP QUANTIFICATION

The concentration of HIV-1 Gag VLPs was assessed by fluorimetry using an developed and validated quantification assay[21]. VLP containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples. There is a linear correlation between fluorescence intensity and p24 values determined using the INNOTEST ELISA HIV antigen mAb (Innogenetics NV, Gent, Belgium). RFU values can be converted to Gag::eGFP concentration values using the following equation:

$$\text{Gag::eGFP (ng/mL)} = (3.245 \times \text{RFU} - 1.6833) \times 36 \qquad (1)$$

where Gag::eGFP is the estimated concentration of polyprotein and RFU is the measured GFP fluorescence intensity in the samples. The first term is the correlation equation between fluorescence values and p24 concentrations determined by ELISA and 36 is a correction factor that takes into account the difference in molecular weight between p24 and Gag::eGFP and an underestimation arising from using the p24 ELISA to estimate p55 Gag concentrations. Assuming that a single VLP contains 2,500 Gag::eGFP molecules[22] and that one Gag::eGFP is 84kDa ($1,39 \cdot 10^{-10}$ ng), the concentration of VLPs can be calculated.

PROTEIN SAMPLES PREPARATION FOR MASS SPECTROMETRY ANALYSES

Protein extraction was performed from ultracentrifuged supernatants using extraction buffer (100 mM Tris-HCl pH 8.8, 2mM EDTA, 4% SDS, 50mM DTT) from which 100µL were added to the sample of each condition. Samples were sonicated for 5 min and then boiled for another 5 min. Protein extracts were quantified with RC/DC Protein Assay (Bio-Rad, Hercules, CA, USA) and stored in -20 ºC until the tryptic digestion process. Protein digestion was performed as previously described[23]. Briefly, proteins were digested using sequencing grade trypsin (Promega, Madison, WI, USA) and the filter-assisted sample preparation technology (FASP, Expedeon, San Diego, CA, USA), and the resulting peptides were subjected to TMT-10 plex labelling (AB Sciex, Framingham, MA, USA), joined and desalted. A total of 150µg of

protein from samples of each condition was diluted to a final concentration of 100mM of TEAB labeled with TMT-10 plex according to the manufacturer. Protein samples were labeled by adding 41µL of TMT isobaric tag diluted in anhydrous acetonitrile, followed by a 1hour-incubation step at room temperature. To quench the reaction, 5% (v/v) hydroxyl-amine (8µL per sample) was added, incubated 15 min at room temperature and mixed together followed by addition of TFA 1% to lower pH at 3. TMT-labeled samples were equally mixed. Pooled mix was purified using Oasis HLB C18 column (Waters, Milford, MA, USA).

TMT-labeled peptides were fractionated using High pH reversed-phase peptide fractionation kit (Thermo Scientific, San Jose, CA, USA) according to manufacturer's instructions into 5 fractions for further LC-MS/MS analysis.

LIQUID CHROMATOGRAPHY TANDEM MASS SPECTROMETRY ANALYSIS

The tryptic peptide mixtures were subjected to LC-MS/MS analysis on a nano-HPLC Easy nLC 1000 liquid chromatograph (Thermo Scientific) coupled to a QExactive mass spectrometer (Thermo Scientific). Peptides were suspended in 0.1% formic acid, loaded onto a C18 reverse-phase trapping column (Acclaim PepMap100, 75-µm internal diameter, 3-µm particle size and 2-cm length, Thermo Scientific), and separated on an analytical C18 nano-column (EASY-Spray column PepMap RSLC C18, 75-µm internal diameter, 3-mm particle size and 50-cm length, Thermo Scientific), in a continuous gradient (8–31%B in 240 min, 31–90%B in 2min, 90%B in 7 min, and 2%B in 30min; where buffer A is 0.1% formic acid in HPLC grade H_2O and buffer B is 100% ACN, 0.1% formic acid in HPLC grade H_2O). Spectra were acquired using full ion scan mode over the mass-to-charge (m/z) range 390–1500, 70,000 FT-resolution was performed on the top 15 ions in each full MS scan along the chromatographic run, using the data-dependent acquisition mode with 45s dynamic exclusion enabled. HCD fragmentation as performed at 30% of normalized collision energy.

PROTEIN IDENTIFICATION AND QUANTIFICATION

Protein identification was performed over the raw files using the SEQUEST HT algorithm integrated in the Proteome Discoverer 2.1 (Thermo Finnigan, Thermo Scientific). MS/MS scans were matched against a

human database (UniProtKB/Swiss-Prot 2017_10 Release). Sequence of Gag::eGFP protein was added to the selected database to enable identification.

For database searching, parameters were selected as follows: trypsin digestion with 2 maximum missed cleavage allowed, precursor mass tolerance of 800 ppm, fragment mass tolerance of 0.02 Da. TMT-10 plex labeling at N-terminal and lysine (+229.62932 Da) as well as cysteine carbamidomethylation (+57.021 Da) were chosen as static modifications, whereas methionine oxidation (+15.994915 Da) was chosen as dynamic modification. The same MS/MS spectra collections were searched against inverted database constructed from the same target database. SEQUEST results were analyzed by the probability ratio method[24]. False discovery rate (FDR) for identified peptides was calculated in the inverted database search results using the refined method[25].

TMT reporter ion intensities were extracted from MS/MS spectra for relative quantification of protein abundance to characterize dynamic protein expression profiles in the selected conditions.

STATISTICAL ANALYSIS

For the comparative analysis of the protein abundance changes we applied Weighted Scan- Peptide-Protein (WSPP) statistical workflow[26], using SanXoT package[27]. It uses as an input a list of quantifications in the form of log2-ratios (for example a condition versus control sample) with their statistical weights and generates the standardized forms of the original variables computing the quantitative values expressed in units of standard deviation around the averages. The quantitative information is obtained from the spectra and used to quantify the peptides from which the spectra are produced and then, proteins that generate these peptides. In other words, the quantitative information is integrated from the spectrum level to the peptide level and then from the peptide level to the protein level[28]. These standardized variables, (Zq), express the quantitative values in units of standard deviation[29]. For the protein functional analysis, Systems Biology Triangle (SBT) model[30] was used. This algorithm estimates weighted functional category averages (Zc) from the protein values by performing the protein to category integration. After the integration from spectra to peptide and peptide to protein, this integration represents a higher level, from protein to category.

The integration allows the detection of changes in functional categories produced by the coordinated behavior of their proteins[28]. Together with each Zq and Zc, the corresponding FDR was calculated. 5% FDR was considered significant. The quantified proteins were functionally annotated using the Gene Ontology database[31,32]. For further Gene Ontology annotation, DAVID[33,34] was used to perform functional enrichment analysis and to extract *Benjamini-Hochberg (BH) adjusted p-values* for the enriched processes. To help analyze and comprehend the data, the online software for reactions, proteins and pathways analysis REACTOME[35] was used.

INTERACTION ANALYSIS

Proteins identified in the study and described in the online database Vesiclepedia were first subjected to an enrichment analysis using DAVID bioinformatic tool[34,36]. Based on their main biological process GO term, five different groups of proteins were selected for further analysis: viral processes, vesicle transport, immune response, DNA and RNA and cellular response. Each group then followed a protein-protein interaction network analysis using STRING database[37]. The interaction networks and clusters resulting from STRING were edited with the software Cytoscape[38].

PARTICLE SIZE MEASUREMENT

Dynamic light scattering (DLS) experiments were performed using a Zetasizer Nano ZS instrument (Malvern instruments, Malvern, UK) with a He/Ne 633 nm laser at 173°. The hydrodynamic diameter, particle size distribution in volume, derived count rate (dCR) and polydispersity index (PDI) were calculated with cumulative fit correlation at 25°C and 0.887 cP or 2.448 cP for concentrated samples by ultracentrifugation, respectively. Briefly, 50 μL of sample was placed in disposable plastic cuvettes (UV-Cuvette micro, BRAND GMBH, Germany) followed by automated experimental data collection. Technical triplicates with 12 scans of 10 s were performed in each independent measurement.

Nanoparticle tracking analysis (NTA) was performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag-GFP VLPs and neutral density filter for total particle by light scattering. Data were analyzed with NanoSight® NTA 3.2 software.

Briefly, samples were injected, and three technical replicate analyses were carried out. Three video recordings of 60sec length were made for each sample. Subsequently, particles were identified and tracked by their Brownian motion at room temperature. Capture settings were recorded with an sCMOS camera (camera level of 8 for Gag::eGFP VLP samples, and 11 for controls, viscosity: 0.9 cP) and analyzed with a detection threshold of 4.

CRYO-TRANSMISSION ELECTRON MICROSCOPY (CRYO-TEM)

A 2-3 µL amount of sample was blotted onto holey carbon grids (Quantifoil Micro Tools, Großloebichau, Germany and Micro to Nano, Haarlem, Netherlands) previously glow discharged in a PELCO easiGlow glow discharger unit. The samples were subsequently plunged into liquid ethane at −180°C using a Leica EM GP cryo workstation and observed in a JEM-2011 electron microscope (JEOL Ltd., Tokyo, Japan) operating at 200 kV. During imaging, samples were maintained at −173°C, and pictures were taken using a CCD-multiscan camera (Gatan Inc., Pleasanton, CA, USA).

FLOW VIROMETRY

Flow Virometry experiments were performed with CytoFLEX LX (Beckman Coulter Inc., Brea, CA, USA) with Violet SSC (VSSC) 405 nm filter, as reported previously[39]. Instrument was standardized using Megamix-Plus SSC and FSC fluorescent polystyrene beads (0.1, 0.16, 0.20, 0.24, and 0.5 µm; Biocytex, Marseille, France) as a quality control tool. Threshold of the trigger signal (VSSC) was manually adjusted to 1500 and gains were set as 95, 9 and 115 for FSC, VSSC and B525-FITC lasers, respectively. Samples were diluted with PBS 1X until abort rate value below 2%. 300.000 events were analyzed at a flow rate of 10 µL/min per sample. VSSC-H vs B525-FITC density plots were used to gate the different particle populations (i.e. small EVs, large EVs, fluorescent particles and HIV-1 Gag::eGFP VLPs). Gating was adjusted manually for each channel. Events after 50 seconds were taken for analysis. The results were analyzed with CytExpert software (Beckman Coulter, Brea, CA, USA). Quantitative values were calculated with equation (2):

$$\text{Particle concentration (Events/mL)} = (\text{Events}) \cdot (\mu L/mL) \cdot \text{Dilution} \qquad (2)$$

NUCLEASE AND RNase ASSAY

Absorbance was measured at 260 nm using NanoDrop™ 1000 Spectrophotometer (ThermoFisher, San Jose, CA, USA) of samples from non-transfected, transfected with mock and following the standard transfection protocol conditions. Then, 100 μL of each biological replicate sample was treated with Benzonase® nuclease (Millipore, Merck, Burlington, MA, USA) diluted in 100 mM Tris-HCl, 20 mM Mg_2Cl at pH 8 and a final concentration of 30 U/mL. After 30 min of incubation at 37°C and gentle shaking, absorbance was measured again at 260 nm. Three technical replicates were measured for each sample. For the RNase assay, vesicle lysis and RNase treatment were performed using a Miniprep Kit (Quiagen, Hilden, Germany) to extract DNA. Lysed samples were again measured at 260 nm.

WESTERN BLOTS

RC/DC protein quantification was used to normalize protein used for western blot assay. A total of 35 g of protein from each condition was separated on SDS-PAGE and transferred onto a polyvinylidene difluoride (PVDF) membrane for 7 minutes using the system Trans-Blot® Turbo™ Transfer System (Bio-Rad, Hercules, CA, USA) as described in the instructions. Membranes were incubated overnight with diluted primary antibody in 5% (w/v) non-fat dry milk 1x TBS 0.1% (v/v) Tween-20 at 4°C with gentle shaking. Primary antibodies used for protein validation were mouse anti-Histone H2A antibody (L88A6, Cell Signaling Technology, 1:1000), mouse anti-Histone H3 antibody (96C10, Cell Signaling Technology 1:1000), mouse anti-TSG101 antibody (612696, BD Biosciences, 1:1000), mouse anti-ALIX antibody (MCA2493, Bio-Rad, 1:1000), rabbit anti-CD63 antibody (ab134045, Abcam, UK, 1:1000), mouse anti-CD81 antibody (ab79559, Abcam, UK, 1:1000), mouse anti-HIV-1 p24 antibody (A2-851-100, Icosagen, Estonia, 1:1000). After primary incubation, a secondary incubation was performed using anti-Mouse (A3562, Sigma Aldrich) or anti-Rabbit (A9919, Sigma Aldrich) antibodies coupled with alkaline phosphatase antibody produced in goat and anti-Mouse IgG coupled with alkaline phosphatase antibody produced in goat as required in 2.5% (w/v) non-fat dry milk 1x TBS 0.1% (v/v) Tween-20 for 1 hour at room temperature. Proteins were visualized using NBT-BCIP® solution (Sigma Aldrich) incubating the membrane for 2-3 min. Membranes were let to dry, scanned at 400 bpi and then analyzed using ImageJ software[40].

EXPERIMENTAL DESIGN AND STATISTICAL RATIONALE

For the multiplexed quantitative proteomics experiment based on TMT-10 plex labeling, three conditions were tested, with three biological replicates (independent cell cultures) from each: no transfection, transfection of empty mock plasmid, and transfection with the plasmid coding for Gag::eGFP protein following the standard protocol previously described. Samples of each condition were taken at 72 hours post-transfection (hpt) followed by an ultracentrifugation to purify HIV-1 Gag VLPs and extracellular vesicles (EVs) with the same density. As it is depicted in **Figure 1A**, the TMT-based isobaric labelling quantification was performed with the three biological replicates for each condition. Protein Zq values comparing the standard transfection vs mock condition was notated as "S condition". Protein Zq values comparing the non-transfected vs mock condition was notated as "N condition"

RESULTS

HEK293 VLP production analysis

Cell cultures for the standard transfection condition were transfected with a plasmid coding for HIV-1 Gag polyprotein fused to GFP (Gag::eGFP), while mock transfection with an empty plasmid served as a transfection control. An extra condition including non-transfected cells was added to characterize the basal EV production. All transfections were carried out at $2 \cdot 10^6$ cells/mL and supernatants from all three conditions were harvested at 72 hours post-transfection (hpt). NTA quantification of fluorescent VLPs were used to assess the purification efficiency. VLPs were concentrated around 15 times from $(1.5 \pm 0.1) \cdot 10^{10}$ to $(2.3 \pm 0.4) \cdot 10^{11}$ VLPs/mL (**Figure 1B**) upon ultracentrifugation. The same purification protocol was carried out with the rest of the samples and henceforth all the analyses are referred to the purified samples. The total number of diffracting particles was assessed by NTA showing that transfection does not influence EV production, as the total number of particles in non-transfected and mock condition did not present any significant difference (*p-value* = 0.99) , with a mean of $(1.82 \pm 0.6) \cdot 10^{11}$ and $(1.86 \pm 0.4) \cdot 10^{11}$ of total particles respectively (**Figure 1C**). Reasonably, upon *gag::egfp* transfection, the total number of particles significantly increased (*p-value* = 0.0008), up to $(3.44 \pm 0.6) \cdot 10^{11}$ as VLPs are being produced.
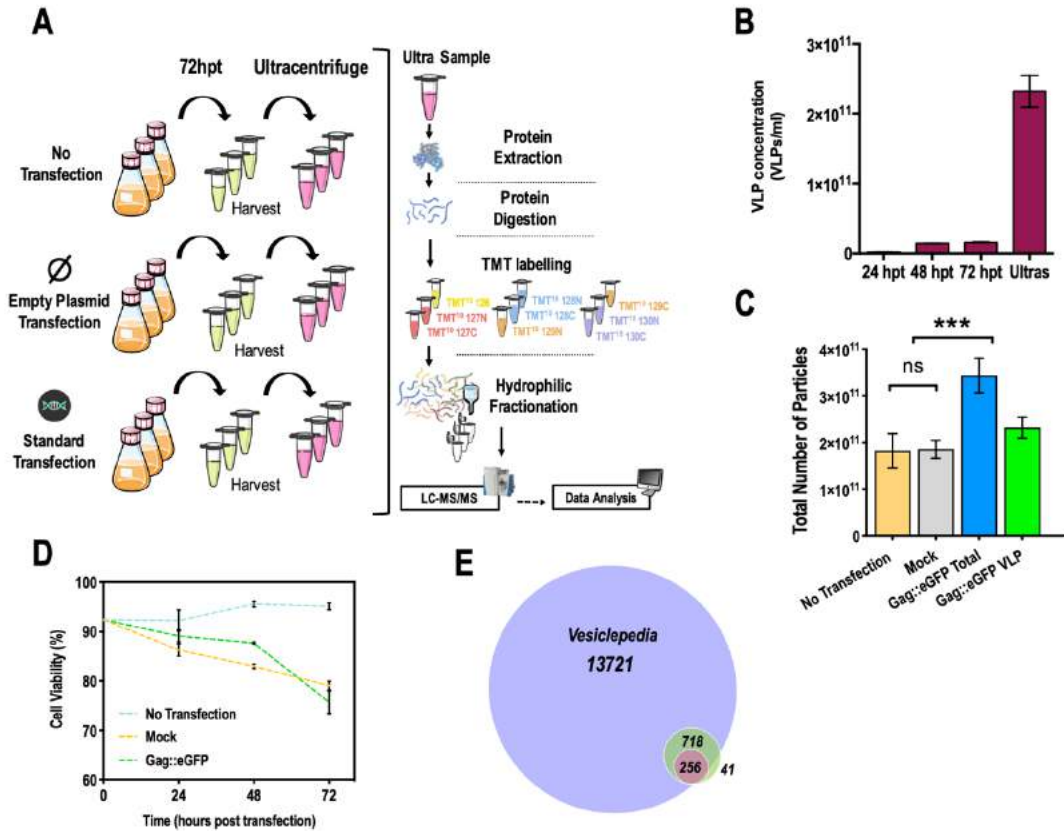
Figure 1. Experimental design and particle production characterization. **A)** Proteomic experimental workflow. Three biological replicates of HEK293 cells were cultured in three different conditions. No transfected condition, transfected with an empty plasmid denoted as "mock", and transfected using the standard protocol with the plasmid containing the gene *gag::egfp*. At 72 h post transfection (hpt), sample was taken from each replicate and centrifuged at 1000 × g and the pellet was discarded. The stored supernatant was ultracentrifuged at 31000 × g using a sucrose double cushion. The interphase containing VLPs and extracellular vesicles of the same density was extracted. Proteins were extracted from the ultracentrifugation samples and digested, and peptides were labeled using tandem mass tag (TMT) labeling. Labeled peptides were fractionated and analyzed via LC-MS/MS. **B)** VLP concentration in the supernatants of the standard transfection condition in the samples before and after ultracentrifugation. **C)** Total number of particles measured by NTA in the three different conditions. For the condition transfected with gag::egfp, fluorescent particles were also quantified. Significance calculated using one-way ANOVA, F = 18.82, DF = 8, and n = 3 for each condition. **D)** Cell viability throughout the time course of the three different conditions. **E)** Venn diagram for the total number of quantified proteins (718) and the number of proteins quantified with more than 1 peptide (256) within the reported proteins in the Vesiclepedia database (13721). There is a total of 41 identified proteins not found in Vesiclepedia from which 33 were identified only with 1 peptide.

Comparison of fluorescent over total diffracting particles revealed that the VLP fraction represented the (68 ± 4) % of the total number of diffracting particles. Thus, copurified EVs are still present in the ultracentrifugation sample. Due to the previous ultracentrifugation step, only copurified vesicles with similar density to VLPs were analyzed. The total number of EVs produced was not assessed in this study.

Proteomic analysis of HEK293 secretome

In order to characterize the changes in EV biogenesis and to understand the pathways involved in this process, a multiplexed quantitative proteomics approach based on TMT-10 plex labeling was used. A total of 718 proteins were identified in this study at 1% FDR, from which 264 with more than 1 peptide. The list of all identified proteins can be found in Supplementary **Table S1**. The protein database Vesiclepedia was used to match the identified proteins, containing 256 out of the 264 proteins (**Figure 1E**). The following analyses were made using the control condition mock to identify changes in the secretome when producing VLPs via transient transfection and in the non-transfected cultures. The SBT model provides functional category averages (Zc) for each biological process derived from the proteins present in the analysis. The ratio of Zc values from the standard transfection and mock condition is notated as "S condition". The ratio of Zc values from non-transfected and mock condition is notated as "N condition". A total of 592 biological processes gene ontology (GO) terms were quantified with more than 5 proteins in our study, from which 59 and 54 were significantly altered (5% FDR) in the S condition and N condition, respectively (**Figure 2A**) . Therefore, to delimit and filter the most relevant processes, a complementary enrichment analysis was carried out using DAVID over the 256 proteins identified with more than 1 peptide. From this enrichment analysis, 81 biological process GO terms showed a 5% *BH adjusted p-value* **(Figure 2A).** Considering SBT model and DAVID analysis together, a total of 16 significantly enriched biological process GO terms were significantly up or downregulated in the extracellular environment by the two different models (**Figure 2B**). The average Zc value of these 16 enriched biological processes resulting from the SBT model are plotted in **Figure 2C**. A complete overview of the alterations in EV physiology when producing VLPs in cell cultures can be observed in this plot. RNA processing and protein translation pathways were increased when producing VLPs (S condition). This correlated with the metabolic state of producing cells, which was enhanced for protein production and required more energy for it. Conversely, microtubule-based processes, localization of Cajal bodies, protein folding and viral processes GO terms were downregulated in S condition. Analyzing closely the proteins that were used to annotate these GO terms by the different enrichment tools from DAVID and STRING databases, it was noticed that the observed downregulation in viral process in S condition came from a downregulation in nuclear transport proteins,

A

Proteomic Data	
Total quantified proteins	718
Quantified proteins > 1 pep	256
Total quantified GO terms	4190
Quantified GO terms > 5 prots	529
Significant GO terms (S)	59
Significant GO terms (N)	54
Enriched GO terms by DAVID	81
Common enriched GO terms	16

B

C
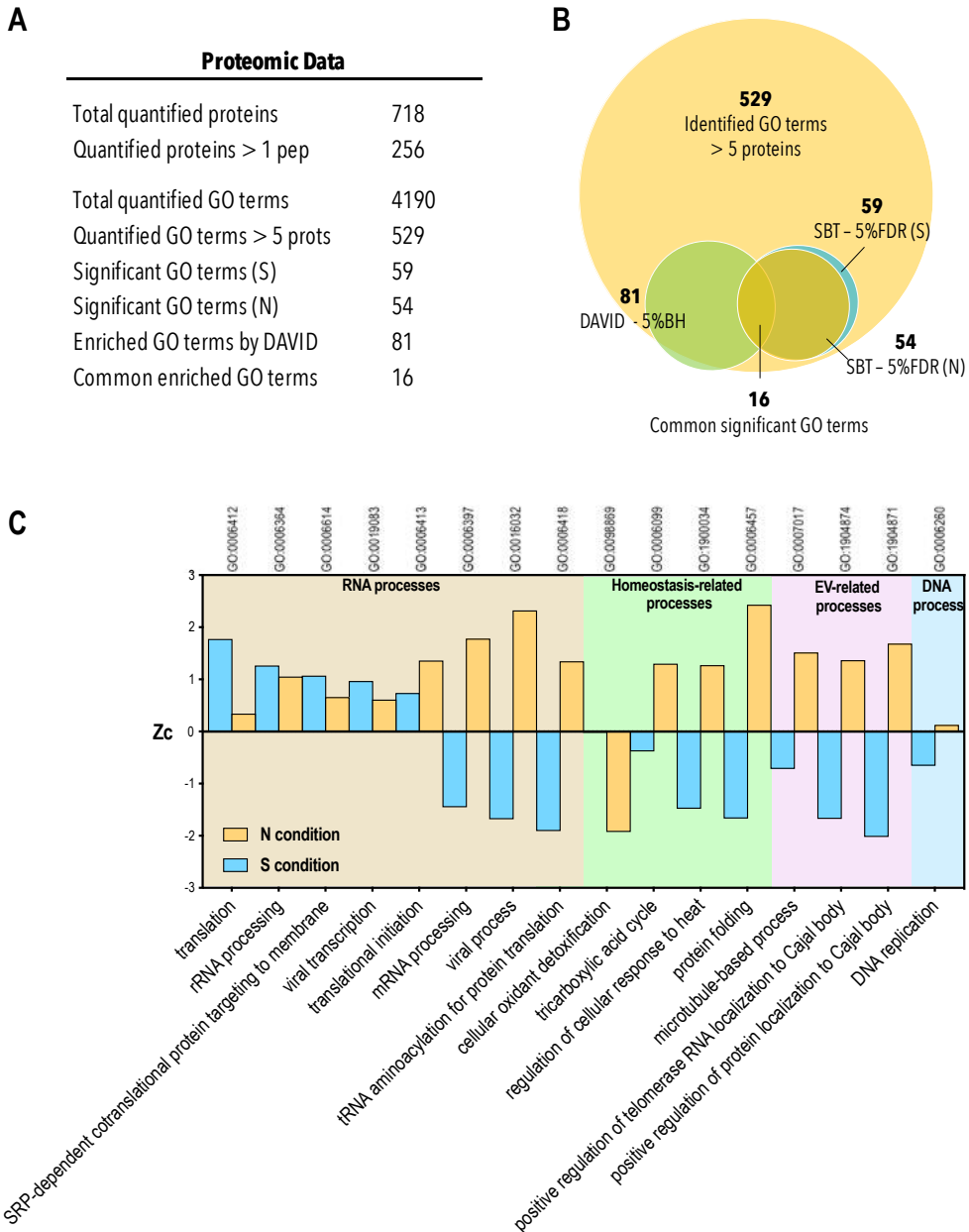
Figure 2. A) Main proteomic data from the study. Zc values for the S condition were calculated as the ratio of the Zc from standard transfection and mock condition. Zc values for the N condition were calculated as the ratio of the non-transfected and the mock condition. Significant GO terms in the S and N conditions were calculated by the SBT model with 5% FDR. Significant GO terms by DAVID enrichment analysis were calculated using the 256 proteins identified with more than 1 peptide and 5% Benjamini–Hochberg adjusted p-value. **B)** Venn diagrams showing the main group of analyzed GO terms. The 16 significant common GO terms obtained via DAVID analysis and the SBT model were selected for further analysis. **C)** Altered biological processes in VLP-copurified extracellular vesicles at 72 hpt. The X-axis represents the 16 significantly enriched biological process GO terms. 5% BH adjusted p-value given by DAVID analysis and 5% FDR by the SBT model. The different Zc values corresponding to each of these GO terms are plotted in the left Y-axis. The yellow series represents the given Zc value for the non-transfected condition, and the blue series, for the standard transfection condition.

like NUP155 and NUP160 and importins like IPO7 among others (**Supplementary Figure S1**). The downregulation in nuclear transport has been reported to be triggered upon transfection, in agreement with these results[41]. Within the proteins annotated in the viral process GO term, those involved in protein translation are upregulated (**Supplementary Figure S1**). Also, analyzing localization of Cajal bodies GO term, common proteins like the chaperonins CCTs, heat shock proteins like HSP90 and HSP70, signaling proteins like PPIA, calnexin and COPB2 were found in all previous mentioned biological processes. T-Complex chaperonin proteins (CCTs) are involved in actin and tubulin folding[42] as well as in the BBsome formation[43]. The BBsome complex plays a role in microtubule-based intracellular transport and is involved in loading cargo into microvesicles[44,45]. COPB2 is also involved in intracellular vesicle transport and calnexin is an ER and microvesicle marker. These findings pointed to the fact that microtubule-based vesicle transport, or microvesicle transport, was downregulated when producing VLPs (S condition). On the contrary, in the N condition, these biological processes were upregulated, showing a shift from non-transfected to VLP-producing in microtubule-based vesicle transport. Oxidative stress-related GO terms, like oxidant detoxification, were downregulated in N condition compared to S condition, showing an increase in oxidative stress-related proteins in the EV environment, reflecting the disruption of homeostasis taking place in the cells when VLP production is engaged. Biological processes related to immune activation like antigen processing and presentation increased in S compared to N suggesting that the proteins present in the cellular membrane of VLPs and vesicles copurified with VLPs have the potential to act as adjuvants in immunization strategies. In addition to this, DNA replication GO term suggests the presence of histones in the VLP-copurified EVs.

Shift in extracellular vesicle biogenesis: From large to small EVs

The total number of particles did not change due to transfection. However, it increased when producing VLPs (**Figure 1C**). In order to characterize the composition and distribution of particles in each condition, experiments using flow virometry, DLS and NTA were performed (**Figure 3**). Using flow virometry, subpopulations of small EVs (sEVs: 30-200 nm of diameter) and large EVs (lEVs: 200-1000 nm of diameter)

A

B

C
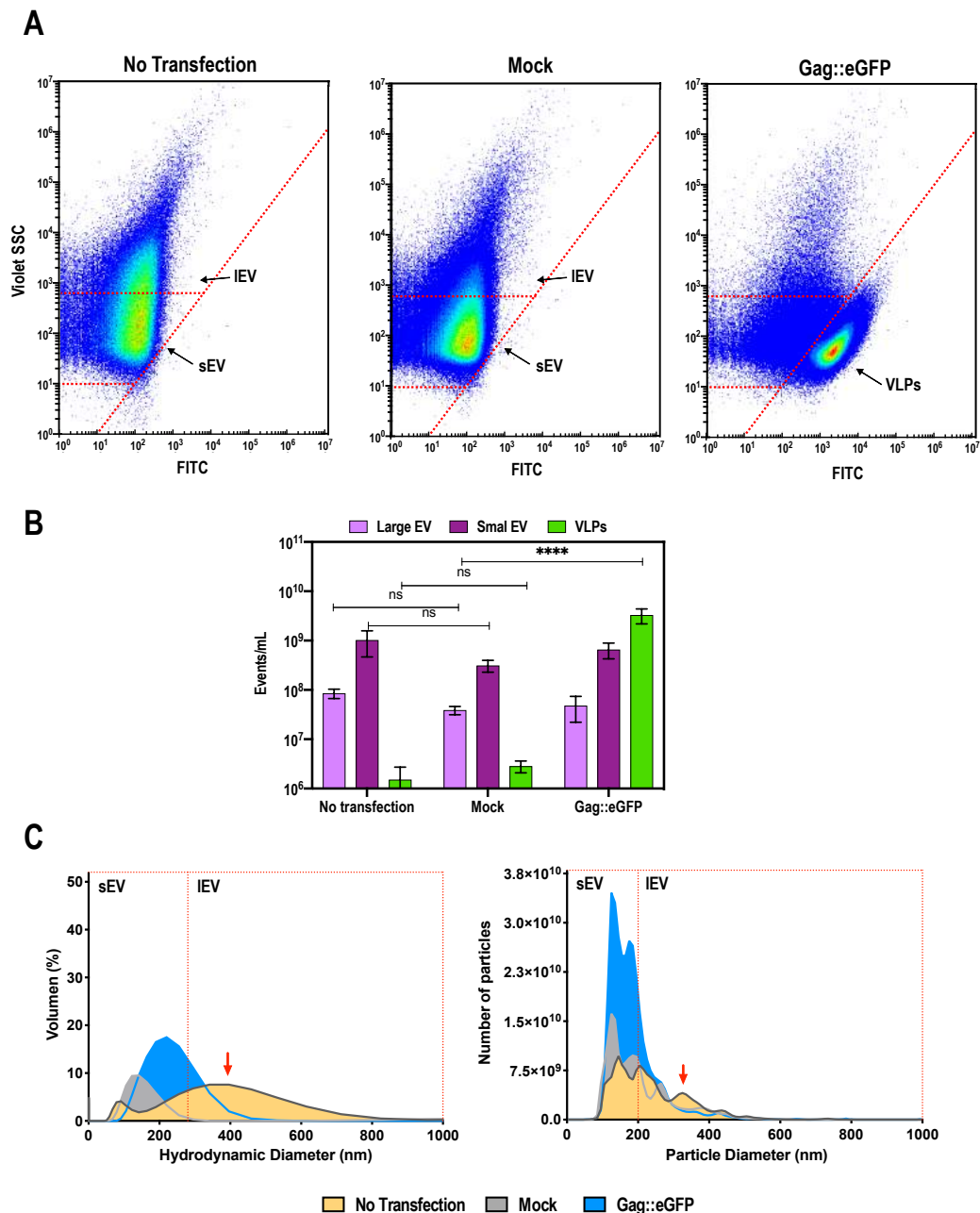
Figure 3. Extracellular vesicle distribution characterization. **A)** Flow virometry density plots showing size (violet side scatter) and green fluorescence (FITC) in the three different studied conditions. Different regions correspond to different particle sizes. Regions corresponding to small extracellular vesicles (sEVs) and to large extracellular vesicles (lEVs) are delimited. In the standard transfection condition, the population of green fluorescent particles corresponds to Gag::eGFP VLPs. **B)** Quantification of flow virometry subpopulations of large EV, small EV, and VLPs in each of the studied conditions. Significance is calculated via two-way ANOVA, DF = 18. **C)** Relative quantification of particles measured by DLS (% of sample volume). Absolute quantification of particles measured by NTA. Red arrows point to the population of large extracellular vesicles.

were defined (**Figure 3A**) using control beads. In standard transfection condition a clear subpopulation of GFP positive nanoparticles was observed. This subpopulation was present in the sEV area suggesting that, indeed, these were VLPs (140-150 nm). Interestingly, although the same total number of particles was measured in non-transfected and the mock condition, a change in their composition was assessed by this technique. Cytoflex quantification analyses (**Figure 3B**) proved a significant increase in VLPs in the Gag::eGFP condition. Likewise, the presence of large EVs seemed to decrease when cells were transfected. To confirm the increase in sEVs and the decrease of lEVs, particle distribution in DLS and NTA was analyzed (**Figure 3C**). Here, a subpopulation of large EVs of 350-400 nm of diameter clearly appeared in N condition. Likening this condition to the ones which underwent transfection, this large EV subpopulation was no longer observed. Yet, an increase in sEVs was definitely noted in mock and standard transfection condition. In both conditions, two main sEV subpopulations whose diameter was around 140 nm and 180 nm were highlighted. Consistently, the condition producing VLPs presented a total number of particles which almost doubled the one in mock condition, in absolute number of particles and also in percentage of total volume, measured by NTA and DLS respectively. Therefore, taking into consideration that EVs and VLPs shared the same density, these findings suggested a shift in EV biogenesis upon transfection, from larger to smaller EVs. In order to characterize this shift in vesicle biogenesis, proteins involved in microvesicle (MV), exosomes biogenesis and vesicle-transport-related processes were analyzed. As previously observed in **Figure 2**, the main GO analysis revealed that microvesicle-mediated transport was downregulated in S. For a deeper study of the proteins involved, an interaction network analysis using STRING was carried out for the proteins associated with vesicle-transport biological function GO term. The interaction network resulting from this analysis is shown in **Figure 4A,** whose Zq value corresponded to standard transfection vs mock condition. Here, it was observed that proteins involved in MV budding like ARF5 or AP3M1 were downregulated in standard transfection condition. Other downregulated proteins like CFL1, SPTBN1 and SPTBN2 are involved in cytokinesis and cytoskeleton-related contraction processes, promoting MV budding. MV biomarkers like CANX were also downregulated. On the other hand, proteins involved in the formation of the late endosomal compartment or MVB and its later fusion to the cell membrane, were upregulated in S condition. Proteins involved in endocytosis like PACSIN2, SH3GL1,

A

B

C

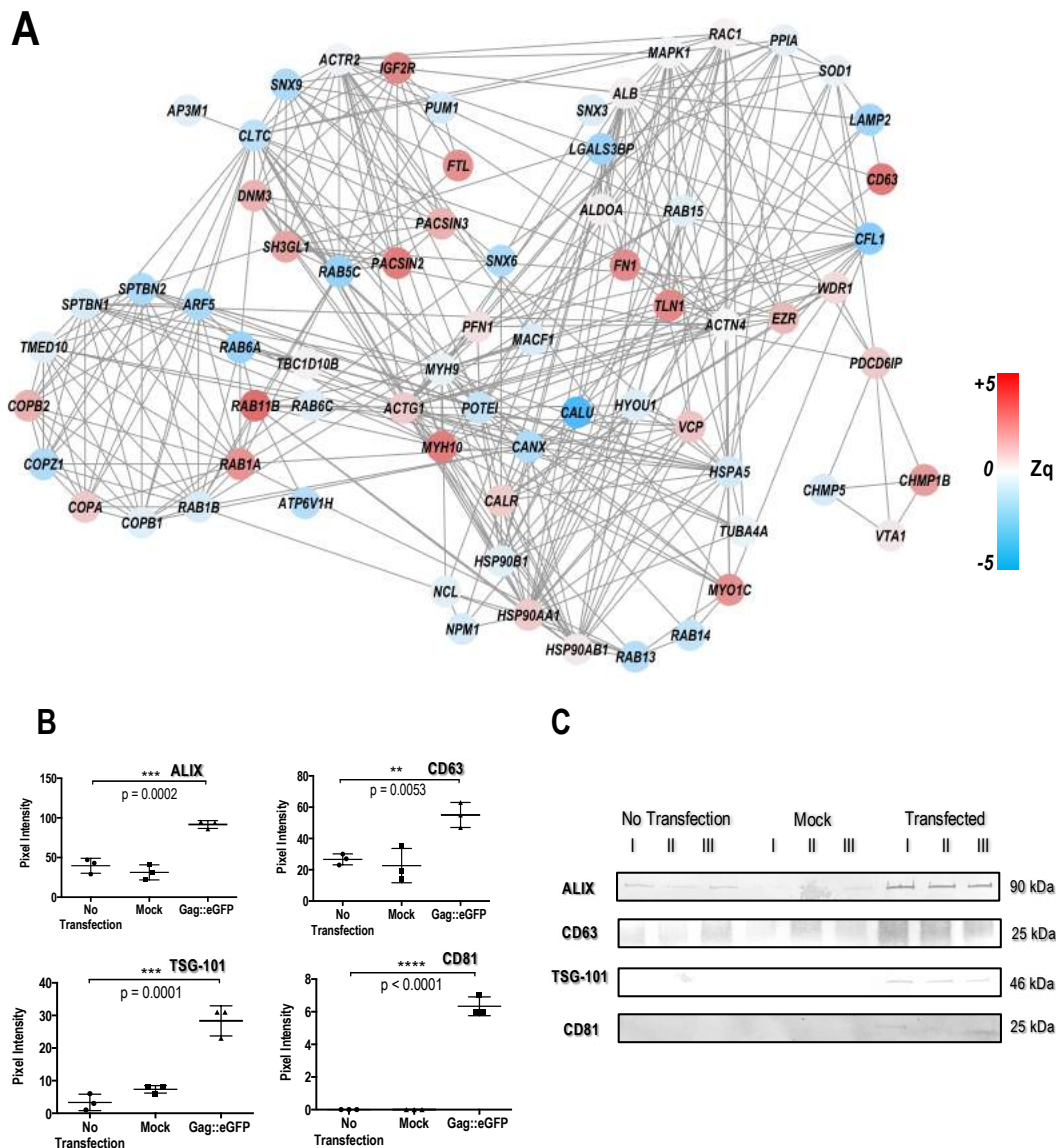ImageJ. Bar plots showing the change in expression of ea[...]
whiskers extend to extreme data points. p-Values are calcul[...]
(F = 47.16, DF = 8), TSG-101 (F = 55.97, DF = 8), CD63 ([...]
membranes. I, II, and III represent the three biological repli[...]

IGFR2, RAB11, RAB1A and DNM3 were increased in S, suggesting that the MVB pathway for production of exosomes was upregulated, coinciding with the increase in the sEVs fraction observed by DLS and NTA. On top of that, proteins reported to be exosome biomarkers were found to be upregulated, like FN1, TLN1, FTL, CD63 and some components of the ESCRT complexes, like, CHMPB1, PDC6CI (ALIX) and TSG-101 (**Figure 4B-C**). The protein interaction analysis supported that there was shift from MV biogenesis, to the upregulation of the MVB pathway, producing sEVs in mock and standard transfection condition. When visualizing a total of 176 particles using cryogenic transmission electron microscopy, in non-transfected samples, vesicles larger than 200 nm were found (**Figure 5A**). In mock and standard transfection condition (**Figure 5B-C**), vesicles smaller than 200 nm were observed and coherently with the increase in particles reported by NTA, the concentration of particles in the standard transfection condition for the same analyzed area was much higher than in mock condition. Interestingly, VLPs are shown as electrodense particles compared to other EVs. It can be observed that there is a shell of ~20 nm that might be corresponding to the Gag::eGFP shell below the plasma membrane. Therefore, it was corroborated that the increase in number of particles upon transfection was not only due to the increment of electrodense particles (**Figure 5D**), but also to the increase of non-electrodense particles smaller than 200 nm (sEVs) as shown in **Figure 5C**, validating the fact that exosome biogenesis was indeed incremented upon transfection. This shift was also observed analyzing the particle size distribution given by NTA. In non-transfected samples, (47 ± 3)% of total particles were smaller than 200 nm and (53 ± 4)% larger than 200 nm. Upon transfection with mock plasmid, these percentages shifted to (66 ± 9)% of particles smaller than 200 nm and (34 ± 9)% of particles larger than 200 nm (lEVs). This tendency continued when producing Gag::eGFP (**Figure 5F**), where (88 ± 6)% of all particles were under 200 nm, from which (68 ± 4)% were fluorescent (VLPs) and the rest were sEVs. Here, only (22 ± 6)% were particles above 200 nm, evidencing the shift from lEVs to sEVs biogenesis (**Figure 5E**).
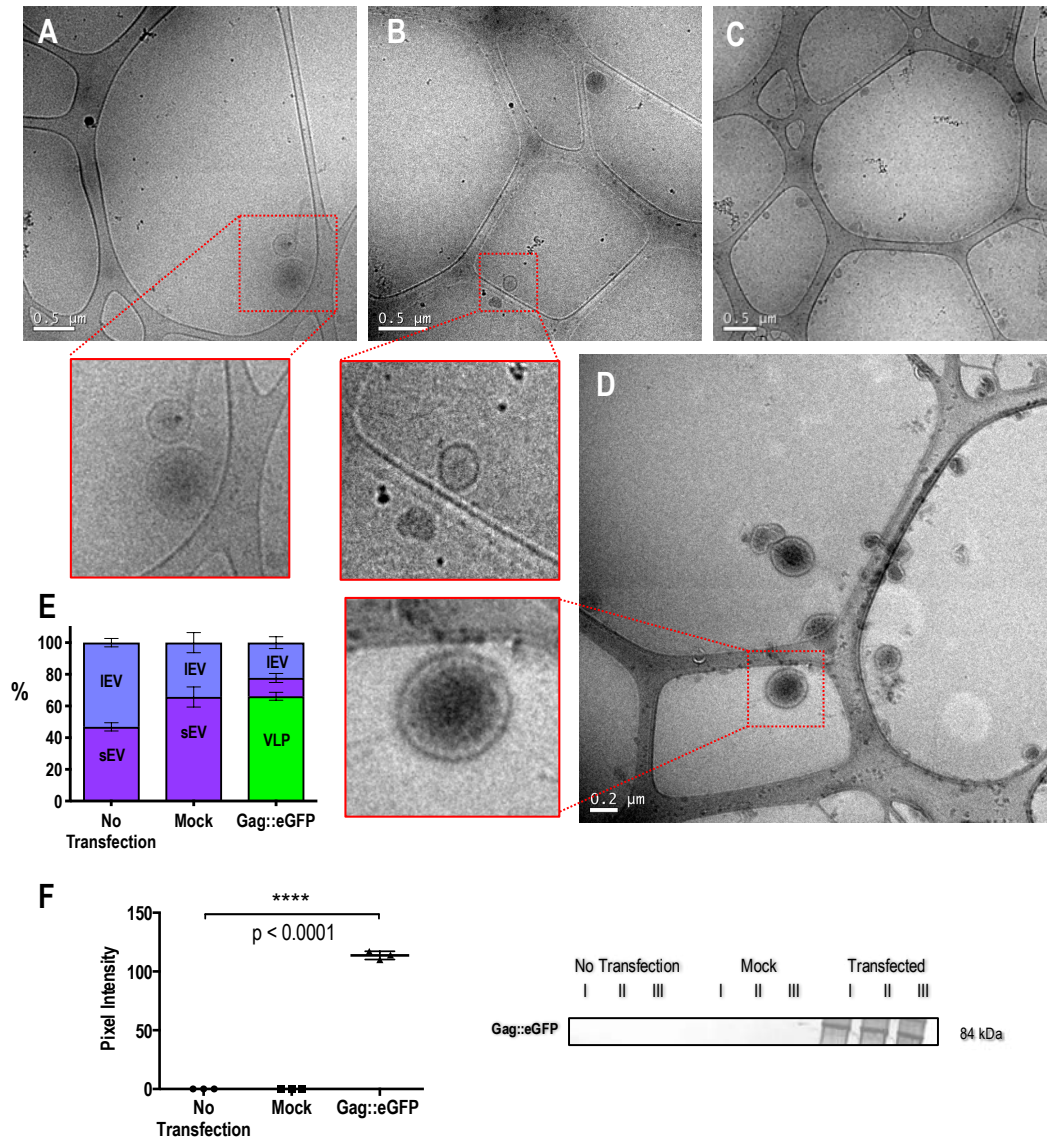
Figure 5. Cryogenic transmission electron microsco[...]e studied condition. **A)** No transfected. The zoomed area depic[...]Mock. The zoomed area depicts an exosome, with a diamete[...]Standard transfection. The zoomed area depicts a VLP, more e[...]EV), large EV (lEV), and virus-like particles (VLPs) in each conditi[...])Western blot of p24 protein, present only in standard transf[...]zed using ImageJ. p-Values of the bar plot showing the change[...], n = 3, in each condition, F = 3143, DF = 8.

The increase in sEVs correlates with DNA secretion

Similar to the interaction analyses performed for vesicle transport-related proteins, a DNA and RNA-related processes interaction analysis was carried out for the identified proteins. As it is observed in **Figure 6A**, most of the proteins related to DNA and RNA processes were downregulated upon VLP production. These proteins are involved in biological functional processes like DNA replication, DNA repair and mRNA splicing (**Figure 6A**). The event of transient transfection substantially impacts intracellular homeostasis, disrupting then processes such as DNA repair and maintenance. Consequently, a downregulation in these proteins was observed upon transfection and VLP production. Since in S condition Gag polyprotein was being largely produced, ribosomal proteins like some proteins belonging to the RSP and RLP protein family were upregulated, corresponding to the process of translation initiation.

Nevertheless, the highest upregulation observed in this group of proteins was found for histones. The presence of histones in copurified EVs during VLP production has been reported before[46,47]. Coherently, histones were upregulated in samples corresponding to S condition (**Figure 6B**), suggesting DNA was present inside sEVs or VLPs, since this upregulation in histones was found in the purified samples of EVs and VLPs. DNA can be secreted within these structures, or even owing to the sticky nature of DNA due to its electric charge, it can be adhered to the outside surface. Absorbance at 260 nm was measured in samples from non-transfected, mock and standard transfection conditions before and after performing a nuclease treatment. This technique was used to remove potential DNA contamination that might be present outside the vesicles and VLPs. Absorbance readings after the assay showed that the nuclease treatment did not reduce the signal at 260 nm. To further asses DNA presence within these structures and discard that this signal is coming from RNA presence, a DNA purification protocol was performed, comprising vesicle lysis and RNase treatment. Afterwards, absorbance at 260 nm was again measured showing no significant difference (**Figure 6C**). Considering the particle concentration of each condition, the ratio of absorbance
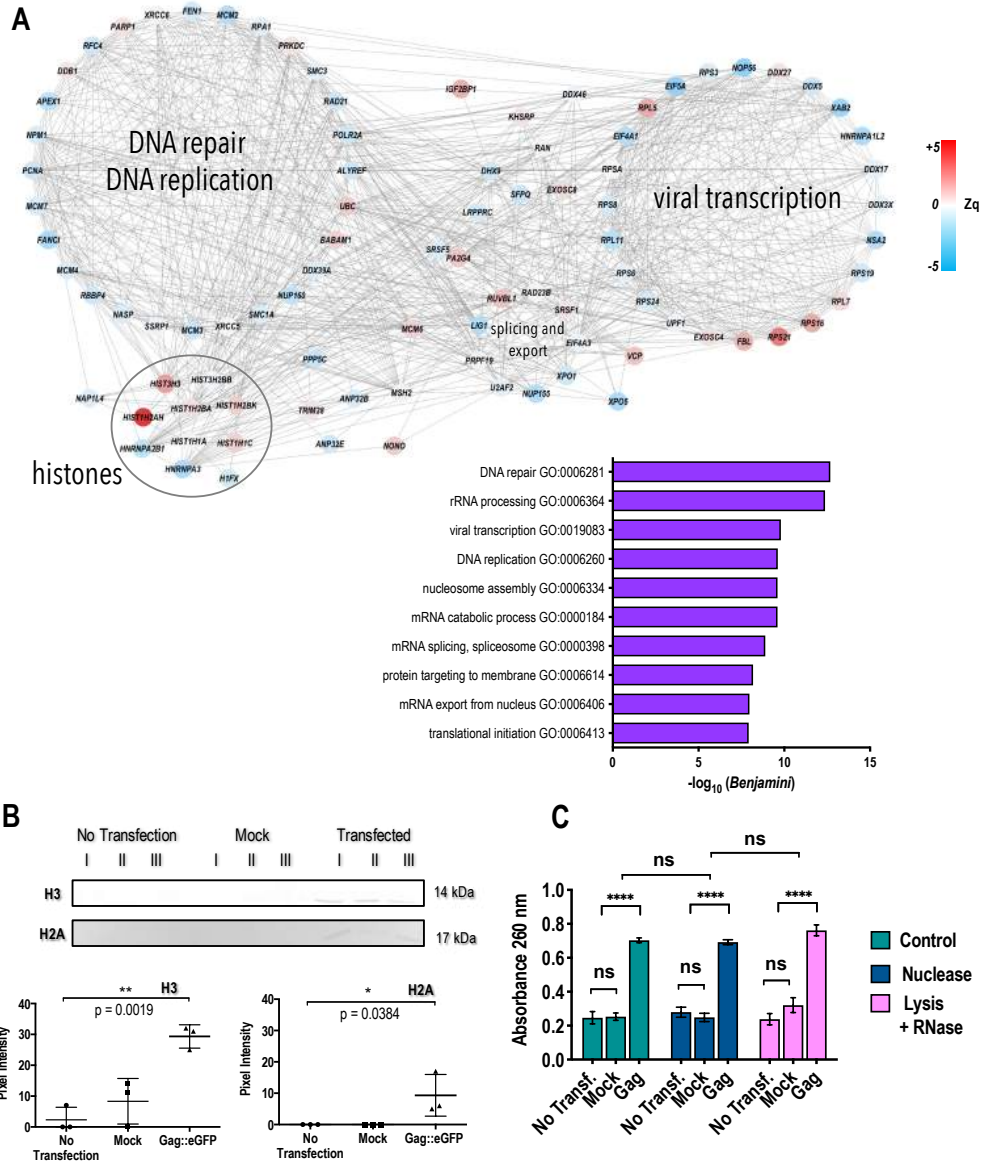
A) Interaction network of proteins having DNA- and RNA-related GO ... represents the value of Zq (S condition) of the different proteins in the ... nt clusters represents the enriched GO term corresponding to the proteins ... and their corresponding Benjamini p-value from the group of proteins ... ons are presented in the bar chart. **B)** Western blots of non-transfected, ... conditions for H3 and H2A proteins. The results from these blots were ... in expression of each condition. Medians are represented by horizontal ... values are calculated using one-way ANOVA test, n = 3, in each condition ... uclease and RNase assays. Absorbance at 260 nm was measured before any treatment (control), after a nuclease, and after lysis and RNase treatment. Significance is calculated using two-way ANOVA test, DF = 18.

per particle was calculated as 0.13, 0.13 and 0.20 (Abs units/ 10^{11} particle) in non-transfected, mock and standard transfection conditions respectively. A significant increase (*p-value* = 0.0351) was observed in standard transfection condition. Therefore, it could be concluded that DNA was present inside EVs structures and it correlates with the previously observed increase in sEVs. Curiously, in non-transfected samples, there was a basal DNA presence, which might be due to the fact that the cell uses exosomes to excrete harmful DNA[10]. This basal presence, or absorbance at 260 nm, did not increase upon transfection. Mock condition samples showed the same level of DNA presence. However, there was a significant increment of DNA in VLP production condition, being coherent with the upregulation in histones found in the proteomic analysis. Yet, the relative contribution of VLPs and other sEVs to the DNA presence remains unclear.

Energy homeostasis disruption is reflected in EVs

Extracellular vesicles are normally produced as a way of cell-to-cell communication. The contents of these structures reflect the physiological state of the cell generating them. Analyzing the changes in metabolic proteins in EVs we can get a glimpse of how homeostasis is regulated in every condition.

Protein ubiquitination was upregulated in S condition. Many of proteasome subunits were found to be upregulated in EVs upon transfection and VLP production (**Figure 7**). This suggested that protein degradation was upregulated. The reason triggering this upregulation could be found in the mitochondria. Mitochondrial content was upregulated upon VLP production, creating an increment of ROS that needed to be detoxified. There was an increase in peroxiredoxins, thioredoxins and other oxidant detoxifiers in S condition. The increment of mitochondria correlated with an increment in proteins involved in glucose metabolism, reflecting the high energy demand. However, TCA proteins and different ATP metabolism-related proteins were downregulated (**Figure 2**). This evidenced that the cell is communicating energy homeostasis disruption via extracellular vesicles. This energy homeostasis disruption agrees with previous reported results[41]. Apoptotic markers like thrombospondin-2 (THB-2) were found to be upregulated together with proteasome subunits, RALB or PRDX5, also indicating the redox detoxification event (**Figure 7B**).
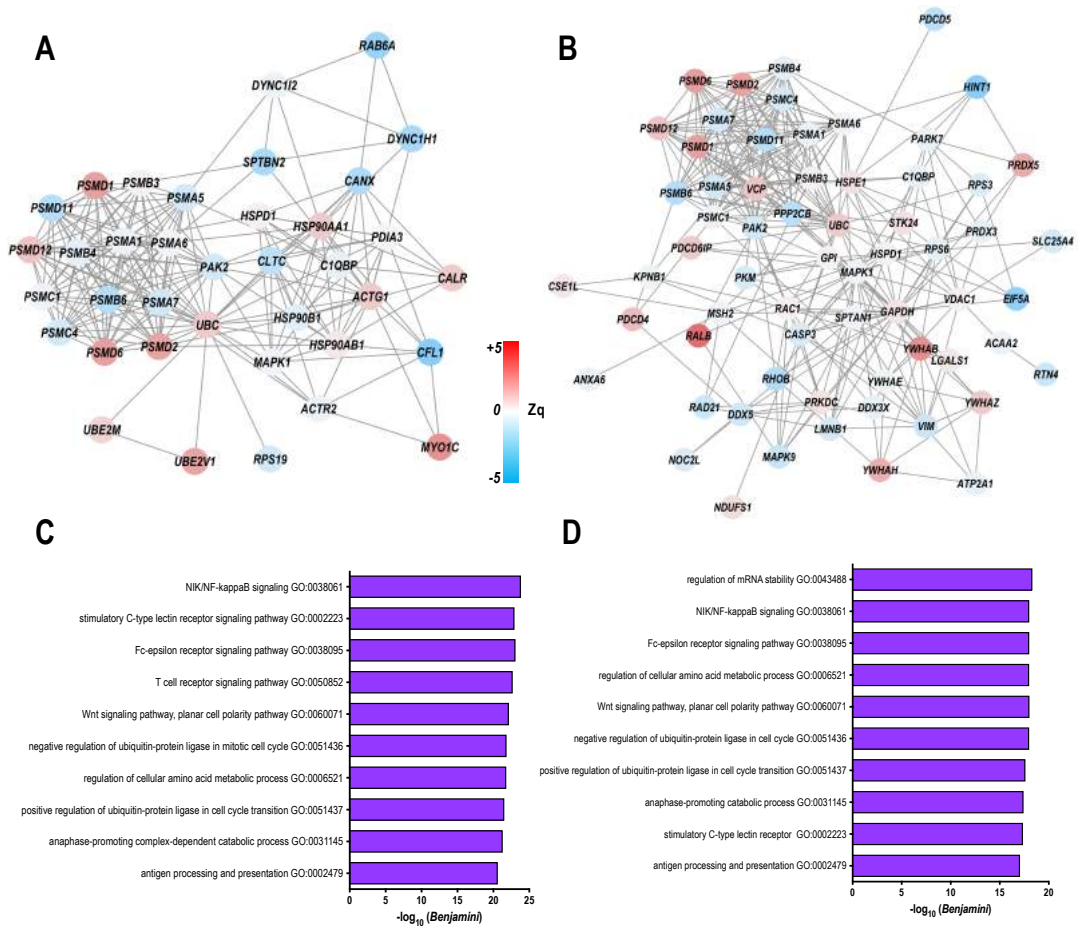
Figure 7. Protein interaction networks corresponding to the immune system process and the cellular response gene ontology annotation. **A)** Identified proteins present in Vesiclepedia with immune response GO annotation. **B)** Identified proteins present in Vesiclepedia with cellular response GO annotation. Color coding represents the value of Zq (S condition) of the different proteins in the standard transfection condition. **C)** Enrichment analysis of GO terms from the group of proteins annotated with the immune response biological function. **D)** Enrichment analysis of GO terms from the group of proteins annotated with the cellular response biological function.

Proteins related to the immune system found in copurified vesicles

Immune activation, inflammation and oxidative stress response are biological processes influencing the development of an immunogenic response when patients are vaccinated. When producing VLPs, copurified extracellular vesicles and VLPs presented proteins associated with these immune system-related processes. Proteins like ACTG1, ITGB1, ENO3, PRDX5, GAPDH, YWHAB, RAC1, TXN, HSPA1L and 10 other proteins from the HSP family are associated with stress response, oxidative stress and immune system

modulation (**Figure 7A-7B**). A enrichment analysis of these proteins showed that they are involved in T cell receptor signaling, stimulation of C-type lectin receptor signaling and antigen processing and presentation processes among others (**Figure 7C-7D**). These processes have been reported to be stimulated and activated by adjuvants in current vaccine therapies[48,49].

DISCUSSION

Extracellular vesicles, including VLPs, are produced by different cellular pathways. Identifying and characterizing how cells behave and modulate EV biogenesis upon transfection can be used to define optimization strategies to enhance VLP production. Depending on the biogenesis pathway used, EVs present different characteristics. Microvesicles (MV) are particles of 50-2000 nm diameter[50]. Their size range is so broad that overlaps with the rest of EVs. Microvesicles are produced by direct outward budding of the plasma membrane. Contraction of actin and myosin structures help redistribute the components of the plasma membrane, creating a membrane evagination and thus completing the budding process[51]. These structures present marker proteins like calnexin or ARF5, ARF6 [3,52] and lack others like ALIX or TSG-101, which are highly present in exosomes[53]. Exosomes are nanoparticles of 30-100 nm diameter which are produced by different mechanisms, like the multivesicular bodies (MVB)[54] pathway. Here, cargo is loaded into intracellular endosomal compartments forming intralumenal vesicles (ILV)[16]. These are invaginations of the membrane towards the lumen of the endosome. The ILV are ultimately bound by plasma membrane, but, as the result of two invaginations, they are contained within a larger endosome. This late endosome then fuses with the plasma membrane, releasing its contents of ILV, to the extracellular space and becoming exosomes. The cargo present in exosomes can be DNA, RNA, enzymes, cytokines…although many of the mechanisms of loading these cargos are still unknown. On the other hand, the process of budding into the luminal face of the endosome has been well reported. First, micro domains rich in tetraspanins like CD81, CD9 and CD63 seem to play an important role, being the anchoring site for proteins to dock and start the budding process[20,55,56]. This makes them widely accepted exosome biomarkers. The Endosomal Sorting Complex Required for Transport (ESCRT) is responsible for

loading the cargos and inducing the ILV formation[57]. Cargos interact with TSG-101 and ALIX, components of complexes ESCRT-I and ESCRT-II respectively[58,59]. This promotes the recruitment of the cargo and the components of the ESCRT-III, responsible for excision. HIV-1 virions have been described to budd off directly from the plasma membrane as well as via the MVB pathway[60,61]. The *Trojan exosome hypothesis*, proposes that HIV uses the existing exosome biogenesis pathway to produce infectious particles[61]. This hypothesis is supported by experiments that showed viral particles presenting exosome biomarkers like tetraspanins CD63, CD81, CD9[62] as well as endosome components[63], showing that HIV uses the MVB vesicle production mechanism to bud off from the cell. More evidence backed up this model as proteins like TSG-101 and ALIX were found in HIV virions[64]. However, HIV mainly buds off directly from the cell membrane recruiting the ESCRT complexes as well as using the tetraspanin-enriched microdomains of the cell membrane[65]. Gag polyprotein has been observed to interact with TSG-101, ALIX, tetraspanin-enriched domains, lipid rafts and other endosome components. This suggests that both pathways are used by the virus to release from the host cell and therefore for VLP production. The results from this work showed that there is a shift from large to small EVs and based on the protein profile observed it could be associated to a change from microvesicle to exosome biogenesis in HEK293 upon transfection. This could be due to the fact that Gag VLPs leave the cell using ESCRT machinery which is present in the MVB pathway as well as in the plasma membrane, reducing cell resources and physical space at the plasma membrane for microvesicle production. The increased recruitment of ESCRT due to the high Gag intracellular concentration might be promoting the exosomes biogenesis pathway since it is also used for VLP production. The evidence suggesting the increase in intracellular trafficking pathways is crucial to explore the optimization of intracellular Gag VLP production.

Transfection and VLP production generate a state of disrupted energy and redox homeostasis within the cell. The cell communicates this state to its extracellular environment by releasing EVs containing proteins reflecting the disruption. This homeostasis disruption is observed in transfected cultures, where VLP production is enhanced. The uptake of these produced EVs could influence the homeostasis of the recipient cell, promoting VLP production. Once homeostasis is disrupted, the cell tries to excrete harmful material causing this disruption. Exosomes are the main resource for the cell to secrete unwanted material like

harmful proteins that cannot be degraded, foreign DNA, enzymes…etc [10]. This could be another reason for the observed shift from microvesicles to exosomes generation. This fact could also explain the presence of DNA within sEVs. It is all triggered by VLP production. Cells have been transfected with large amounts of DNA. Therefore, cells can be using sEVs, including VLPs, to excrete it. Considering a VLP of ~140 nm of diameter, being the Gag::eGFP shell of ~20 nm thick according to the cryo electron micrographs, and considering a nucleosome of 11 x 5.5 nm[66], it can also be possible that DNA is secreted together with histones forming nucleosomes within VLPs as one nucleosome would only represent 0.2% of the available volume. The NC domain of Gag polyprotein is reported to interact with nucleic acids and serve as a scaffold in virion assembly[67,68]. Lacking viral RNA, host cell DNA might be serving as scaffold in VLP formation. The NC domain might be the docking point of nucleic acids. Therefore, if a DNA-free vaccine is envisaged, an engineered NC domain, preventing nucleic acids from docking and being incorporated to the VLPs might be a future study objective.

The presence of EVs has been a topic of interest in many HIV-1 Gag VLP studies[39,69–73]. Mostly regarded as contaminants as the efficiency of VLP production was calculated by the percentage of VLPs achieved in the proposed bioprocess compared to the total number of diffracting particles produced (including EVs). However, their presence in the copurified fraction of the VLPs could be used as an adjuvant for future therapies. The potential use of exosomes conjugated with VLPs and other different nanoparticles as a vaccine is currently being explored since many cancer treatments have proven to be improved by using exosomes as adjuvants[74]. Exosomes are considered a key element subject to further study in order to use them as a potential adjuvant for a VLP-based therapy.

The future combined use of VLPs and EVs could offer a new alternative for the further development of a HIV-1 prophylactic vaccine. This is still a major challenge for science nowadays. Several strategies have been developed for a protective HIV-1 vaccine, such as MPER-V3 HIV-1 VLP, modified Vaccinia Ankara (MVA)- based vaccine candidates or even DNA vaccines[75,76]. However, none showed significant results in humans. Although eliciting a strong immune response, no strategy was able to generate quality antibodies to achieve protection. The main limitation hitherto encountered is the selection of the immunogen or antigen which leads to the generation of high-quality neutralizing antibodies. The combined use of EVs

offers new possibility to combine antigens with immunomodulatory proteins to try the generation of a more specific immune response.

Whether EVs are used as adjuvants or considered impurities in VLP-based therapies, methods of separation are required to reach high purity preparations. The EV characterization performed in this work shows that ultracentrifugation is not enough to fully separate exosomes/microvesicles from VLPs. Chromatography strategies have also been implemented to purify VLPs[77]. However, the low yields achieved suggest the need for a more specific separation protocol in order to be implemented at large scale. Considering the biogenesis pathways of the different EVs copurified with VLPs and the characterization of the defined protein profile it is possible to design metabolic engineering strategies to further develop a specific downstream process. Affinity chromatography could be designed and affine proteins could be overexpressed in these vesicles. The actual impossibility to efficiently separate these subpopulations remains one of the main challenges to overcome.

Considering the great potential of exosomes to complement the effect of VLPs, the next envisaged step would require vesicle engineering in order to isolate exosomes containing proteins which could contribute to improve the immunogenic response. Proteins promoting oxidative stress protection and anti-inflammatory proteins like ENO3 or PRDX5 could be engineered to be coproduced with VLPs, cotransfecting their coding genes together with *gag::egfp*. Another possibility could be metabolic engineering to neutralize the homeostasis disruption signaling which takes place in the cell cultures when producing VLPs. In order to do that, a further study of the glycome of EVs and VLPs would be required. Apart from modifying the outer layer of the vesicle membrane customizing glycoproteins and specific signaling proteins, another interesting target would be the VLP cargos. Gag polyprotein could be used as the anchor to interact with cargos and load them into the VLPs. RNA is described to interact with Gag and to be necessary as a scaffold for Gag multimerization and further VLP formation[67,78,79]. This particularity could be used to design miRNAs with protective effects, such as anti-inflammatory or immune activation effects to be loaded to the VLPs. In order to do that, a further study in the selectiveness of Gag for different species of RNA and targeting strategies would be required and should be assessed to successfully develop an effective VLP-based therapy.

CONCLUSIONS

Characterizing the extracellular environment of the produced VLPs is essential in order to design future VLP-based therapies. In this work, a multiplexed quantitative proteomic approach has been used to determine changes taking place in the secretome of VLP-copurified EVs. When producing VLPs, there is a shift from microvesicle to exosome biogenesis, increasing in 30% the production of vesicles smaller than 200 nm. This change comes together with a disruption of energy and redox homeostasis. DNA is secreted in these small EVs, whose presence is increased when producing VLPs. Another remarkable trait of the extracellular environment accompanying VLPs is the presence of immunomodulatory proteins in these vesicles. Proteins associated with immune activation, anti-inflammatory processes and response to oxidative stress were identified in this coproduced fraction. The characterization of EVs which copurify with VLPs after a first basic purification step is key to develop further, more specific downstream purification processes or to use these copurified vesicles in our advantage to design exosome-based adjuvants.

SUPPORTING INFORMATION
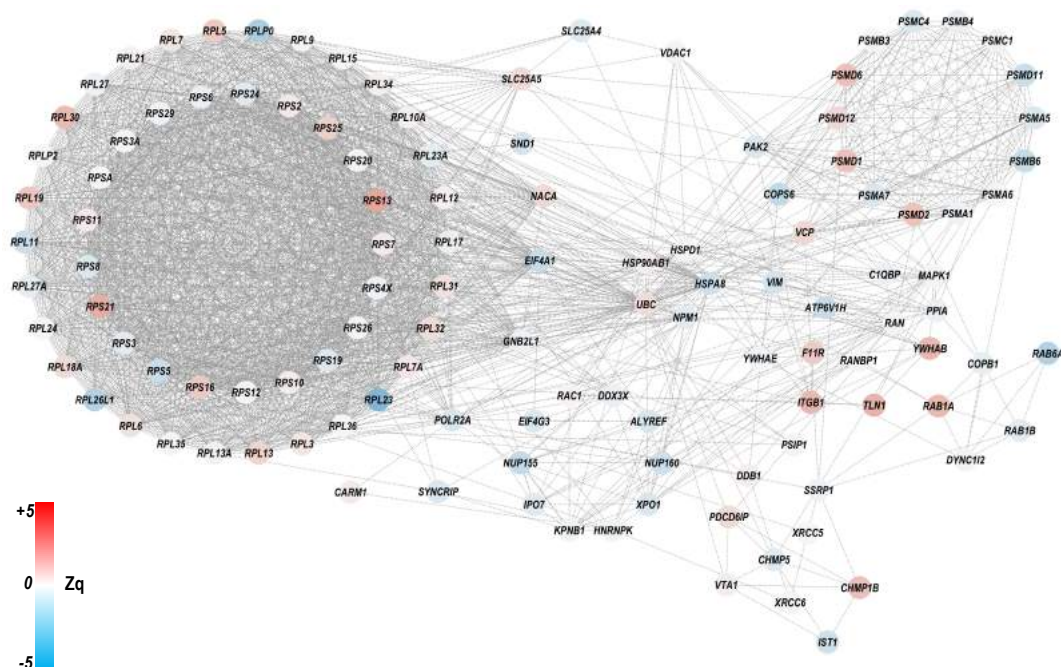
Supplementary Figure S1: Interaction network of proteins having "viral processes" GO annotations and present in Vesiclepedia. Colour coding represents the value of Zq (S condition vs control mock) of the different proteins in the standard transfection condition.

Supplementary Table S1. (List of all Identified proteins in this study) can be found at: https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00581?goto=supporting-info

DATA AVAILABILITY

The raw mass spectrometry data has been submitted to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) with the dataset identifier PXD014746.

REFERENCES

1.   Sundquist, W. I. & Kräusslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine* **2**, (2012).

2.   Charlton Hume, H. K. *et al.* Synthetic biology for bioengineering virus-like particle vaccines. *Biotechnology and Bioengineering* **116**, 919–935 (2019).

3.   Willms, E. *et al.* Cells release subpopulations of exosomes with distinct molecular and biological properties. *Sci. Rep.* **6**, 22519 (2016).

4.   Reiter, K. *et al.* Separation of virus-like particles and extracellular vesicles by flow-through and heparin affinity chromatography. *J. Chromatogr. A* (2018). doi:10.1016/j.chroma.2018.12.035

5.   Lee, Y., EL Andaloussi, S. & Wood, M. J. A. Exosomes and microvesicles: extracellular vesicles for genetic information transfer and gene therapy. *Hum. Mol. Genet.* **21**, R125–R134 (2012).

6.   Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).

7.   Vlassov, A. V., Magdaleno, S., Setterquist, R. & Conrad, R. Exosomes: Current knowledge of their composition, biological functions, and diagnostic and therapeutic potentials. *Biochim. Biophys. Acta - Gen. Subj.* **1820**, 940–948 (2012).

8.   Scourfield, E. J. & Martin-Serrano, J. Growing functions of the ESCRT machinery in cell biology and viral replication. *Biochem Soc Trans* **45**, 613–634 (2017).

9.   Chettimada, S. *et al.* Exosome markers associated with immune activation and oxidative stress in HIV patients on antiretroviral therapy. *Sci. Rep.* **8**, 7227 (2018).

10.   Takahashi, A. *et al.* Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nat Commun* **8**, 15287 (2017).

11.   Zhu, Y. *et al.* A Comprehensive Proteomics Analysis Reveals a Secretory Path- and Status-Dependent Signature of Exosomes Released from Tumor-Associated Macrophages. *J. Proteome Res.* **14**, 4319–31 (2015).

12.   Sun, D. *et al.* Exosomes are endogenous nanoparticles that can deliver biological information

between cells. *Adv. Drug Deliv. Rev.* **65**, 342–347 (2013).

13. Ung, T. H., Madsen, H. J., Hellwinkel, J. E., Lencioni, A. M. & Graner, M. W. Exosome proteomics reveals transcriptional regulator proteins with potential to mediate downstream pathways. *Cancer Sci.* **105**, 1384–1392 (2014).

14. Simpson, R. J., Jensen, S. S. & Lim, J. W. E. Proteomic profiling of exosomes: Current perspectives. *Proteomics* **8**, 4083–4099 (2008).

15. Hessvik, N. P. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci.* **75**, 193–208 (2018).

16. Pant, S., Hilton, H. & Burczynski, M. E. The multifaceted exosome: Biogenesis, role in normal and aberrant cellular function, and frontiers for pharmacological and biomarker opportunities. *Biochem. Pharmacol.* **83**, 1484–1494 (2012).

17. Simons, M. & Raposo, G. Exosomes – vesicular carriers for intercellular communication. *Curr. Opin. Cell Biol.* **21**, 575–581 (2009).

18. Ellwanger, J. H., Veit, T. D. & Chies, J. A. B. Exosomes in HIV infection: A review and critical look. *Infection, Genetics and Evolution* **53**, 146–154 (2017).

19. Florin, L. & Lang, T. Tetraspanin Assemblies in Virus Infection. *Front. Immunol.* **9**, 1140 (2018).

20. Sims, B. *et al.* Tetraspanin blockage reduces exosome-mediated HIV-1 entry. *Arch. Virol.* **163**, 1683–1689 (2018).

21. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

22. Cheeks, M. C., Edwards, A. D., Arnot, C. J. & Slater, N. K. H. Gene transfection of HEK cells on supermacroporous polyacrylamide monoliths: A comparison of transient and stable recombinant protein expression in perfusion culture. *N. Biotechnol.* **26**, 289–299 (2009).

23. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

24. Martínez-Bartolomé, S. *et al.* Properties of average score distributions of SEQUEST: the

probability ratio method. *Mol. Cell. Proteomics* **7**, 1135–45 (2008).

25. Bonzon-Kulichenko, E., Garcia-Marques, F., Trevisan-Herraz, M. & Vazquez, J. Revisiting peptide identification by high-accuracy mass spectrometry: problems associated with the use of narrow mass precursor windows. *J Proteome Res* **14**, 700–710 (2015).

26. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J Proteome Res* **13**, 1234–1247 (2014).

27. Trevisan-Herraz, M. *et al.* SanXoT: a modular and versatile package for the quantitative analysis of high-throughput proteomics experiments. *Bioinformatics* **35**, 1594–1596 (2019).

28. Garcia-Marques, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol Cell Proteomics* **15**, 1740–1760 (2016).

29. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **13**, 1234–47 (2014).

30. García-Marqués, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol. Cell. Proteomics* **15**, 1740–60 (2016).

31. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

32. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

33. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2009).

34. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).

35. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

36. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

37. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association

networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

38.   Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).

39.   González-Domínguez, I., Puente-Massaguer, E., Cervera, L. & Gòdia, F. Quality assessment of virus-like particles at single particle level: A comparative study. *Viruses* **12**, (2020).

40.   Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–5 (2012).

41.   Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed Quantitative Proteomic Analysis of HEK293 Provides Insights into Molecular Changes Associated with the Cell Density Effect, Transient Transfection, and Virus-Like Particle Production. *J. Proteome Res.* **19**, 1085–1099 (2020).

42.   Freund, A. *et al.* Proteostatic Control of Telomerase Function through TRiC-Mediated Folding of TCAB1. *Cell* **159**, 1389–1403 (2014).

43.   Seo, S. *et al.* BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. *Proc. Natl. Acad. Sci.* **107**, 1488–1493 (2010).

44.   Sheffield, V. C. The blind leading the obese: the molecular pathophysiology of a human obesity syndrome. *Trans. Am. Clin. Climatol. Assoc.* **121**, 172–81; discussion 181-2 (2010).

45.   Jin, H. *et al.* The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia. *Cell* **141**, 1208–19 (2010).

46.   Venereo-Sánchez, A. *et al.* Characterization of influenza H1N1 Gag virus-like particles and extracellular vesicles co-produced in HEK-293SF. *Vaccine* **37**, 7100–7107 (2019).

47.   Steppert, P. *et al.* Purification of HIV-1 gag virus-like particles and separation of other extracellular particles. *J Chromatogr A* **1455**, 93–101 (2016).

48.   Decout, A. *et al.* Rational design of adjuvants targeting the C-type lectin Mincle. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2675–2680 (2017).

49.   McKee, A. S. & Marrack, P. Old and new adjuvants. *Current Opinion in Immunology* **47**, 44–51 (2017).

50. Akers, J. C., Gonda, D., Kim, R., Carter, B. S. & Chen, C. C. Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies. *J. Neurooncol.* **113**, 1–11 (2013).

51. McConnell, R. E. *et al.* The enterocyte microvillus is a vesicle-generating organelle. *J. Cell Biol.* **185**, 1285–1298 (2009).

52. Muralidharan-Chari, V. *et al.* ARF6-Regulated Shedding of Tumor Cell-Derived Plasma Membrane Microvesicles. *Curr. Biol.* **19**, 1875–1885 (2009).

53. Chen, L., Chen, R., Kemper, S. & Brigstock, D. R. Pathways of production and delivery of hepatocyte exosomes. *J. Cell Commun. Signal.* **12**, 343–357 (2018).

54. Gruenberg, J. & Stenmark, H. The biogenesis of multivesicular endosomes. *Nat. Rev. Mol. Cell Biol.* **5**, 317–323 (2004).

55. Zöller, M. Tetraspanins: push and pull in suppressing and promoting metastasis. *Nat. Rev. Cancer* **9**, 40–55 (2009).

56. Hemler, M. E. Tetraspanin Proteins Mediate Cellular Penetration, Invasion, and Fusion Events and Define a Novel Type of Membrane Microdomain. *Annu. Rev. Cell Dev. Biol.* **19**, 397–422 (2003).

57. Hurley, J. H. ESCRT complexes and the biogenesis of multivesicular bodies. *Curr Opin Cell Biol* **20**, 4–11 (2008).

58. Fujii, K., Hurley, J. H. & Freed, E. O. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nat Rev Microbiol* **5**, 912–916 (2007).

59. McCullough, J., Fisher, R. D., Whitby, F. G., Sundquist, W. I. & Hill, C. P. ALIX-CHMP4 interactions in the human ESCRT pathway. *Proc Natl Acad Sci U S A* **105**, 7687–7691 (2008).

60. Pincetic, A. & Leis, J. The Mechanism of Budding of Retroviruses From Cell Membranes. *Adv Virol* **2009**, 6239691–6239699 (2009).

61. Gould, S. J., Booth, A. M. & Hildreth, J. E. K. The Trojan exosome hypothesis. *Proc. Natl. Acad. Sci.* **100**, 10592–10597 (2003).

62. Nguyen, D. G., Booth, A., Gould, S. J. & Hildreth, J. E. K. Evidence That HIV Budding in Primary

Macrophages Occurs through the Exosome Release Pathway. *J. Biol. Chem.* **278**, 52347–52354 (2003).

63.   Kramer, B. *et al.* HIV interaction with endosomes in macrophages and dendritic cells. *Blood Cells, Mol. Dis.* **35**, 136–142 (2005).

64.   von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).

65.   Meng, B. & Lever, A. M. Wrapping up the bad news: HIV assembly and release. *Retrovirology* **10**, 5 (2013).

66.   Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS Letters* **589**, 2914–2922 (2015).

67.   Cimarelli, A., Sandin, S., Höglund, S. & Luban, J. Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J. Virol.* **74**, 3046–57 (2000).

68.   Muriaux, D., Mirro, J., Harvin, D. & Rein, A. RNA is a structural element in retrovirus particles. *Proc. Natl. Acad. Sci.* **98**, 5246–5251 (2001).

69.   Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).

70.   Cervera, L., Gonzalez-Dominguez, I., Segura, M. M. & Godia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol Bioeng* **114**, 2507–2517 (2017).

71.   Gutiérrez-Granados, S., Cervera, L., Segura, M. . L., Wölfel, J. & Gòdia, F. Optimized production of HIV-1 virus-like particles by transient transfection in CAP-T cells. *Appl Microbiol Biotechnol* **100**, 3935–3947 (2016).

72.   Lavado-García, J., Cervera, L. & Gòdia, F. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures. *Front. Bioeng. Biotechnol.* **8**, 617 (2020).

73.   Fuenmayor, J., Cervera, L., Rigau, C. & Gòdia, F. Enhancement of HIV-1 VLP production using gene inhibition strategies. *Appl. Microbiol. Biotechnol.* **102**, 4477–4487 (2018).

74.   Seifalian, A., De La Peña, H. & Seifalian, A. M. The application of exosomes as a nanoscale cancer

vaccine. *Int. J. Nanomedicine* **5**, 889 (2010).

75. Tohidi, F., Sadat, S. M., Bolhassani, A. & Yaghobi, R. Construction and Production of HIV-VLP Harboring MPER-V3 for Potential Vaccine Study. *Curr. HIV Res.* **15**, (2017).

76. Chea, L. S. & Amara, R. R. Immunogenicity and efficacy of DNA/MVA HIV vaccines in rhesus macaque models. *Expert Review of Vaccines* **16**, 973–985 (2017).

77. Pereira Aguilar, P. *et al.* Polymer-grafted chromatography media for the purification of enveloped virus-like particles, exemplified with HIV-1 gag VLP. *Vaccine* **37**, 7070–7080 (2019).

78. Rulli, S. J. *et al.* Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. *J. Virol.* **81**, 6623–6631 (2007).

79. Khorchid, A., Halwani, R., Wainberg, M. A. & Kleiman, L. Role of RNA in Facilitating Gag/Gag-Pol Interaction. *J. Virol.* **76**, 4131–4137 (2002).

CHAPTER SIX

Differential *N*- and *O*-glycosylation signature of HIV-1 Gag virus-like particles and coproduced extracellular vesicles

ABSTRACT

HIV-1 virus-like particles (VLPs) are nanostructures derived from the self-assembly and cell budding of Gag polyprotein. Mimicking the native structure of the virus and being non-infectious, they represent promising candidates for the development of new vaccines as they elicit a strong immune response. In addition to this, the bounding membrane can be functionalized with exogenous antigens to target different diseases. Protein glycosylation depends strictly on the production platform and expression system used and displaying the correct glycosylation pattern is essential for the achievement of the desired response. One of the main challenges for the development of Gag VLP production bioprocess is the separation of VLPs and coproduced extracellular vesicles (EVs). In this work, porous graphitized carbon separation method coupled with mass spectrometry was used to characterize the N- and O- glycosylation profiles of Gag VLPs produced in HEK293 cells. We have identified differential glycan signatures between VLPs and EVs that could pave the way for further separation and purification strategies in order to optimize downstream processing and move forward in VLP-based vaccine production technology.

ABBREVIATIONS

APC: Antigen-presenting cells, **CHO**: Chinese hamster ovary, **CLR**: C-type lectin receptor, **EIC**: Extracted ion chromatogram, **EV**: Extracellular vesicle, **Gag::eGFP**: translational fusion of HIV-1 Gag polyprotein and enhanced GFP, **GlcNAc**: N-Acetylglucosamine, **GnT-III**: N-Acetylglucosaminyltransferase-III, **HEK**: Human embryonic kidney, **Hex**: Hexose, **HexNAc**: N-acetylhexosamine, **HIV-1**: Human immunodeficiency virus 1, **LC-MS/MS**: Liquid chromatography coupled to tandem mass spectrometry, **Man**: Mannose, **NeuAc**: N-Acetylneuraminic acid, **NeuGc**: N-Glycolylneuraminic acid, **NTA**: Nanoparticle tracking analysis, **NTA**: Nanoparticle tracking analysis, **PBS**: Phosphate-buffered saline, **PEI**: polyethyleneimine, **PGC**: Porous graphitized carbon, **RFU**: Relative fluorescence unit, **sEV**: Small extracellular vesicles, **VLP**: Virus-like particle

INTRODUCTION

Protein glycosylation represents a crucial feature in the design and development of new vaccines. Glycans play essential roles in cellular processes such as signaling, cell-cell communication, vesicle transport regulation, pathogen recognition, immune responses and allergic reactions[1–3]. Therefore, the design and characterization of the different glycoforms present in a biopharmaceutical product has tremendous impact in its development and further safety, efficacy, tolerability and bioavailability optimization[4,5]. Harboring the correct glycosylation of biopharmaceuticals like monoclonal antibodies has been demonstrated to improve the stability, solubility and it is important for safety and efficacy[6,7]. In the case of influenza viral particles, high titers of IgG were produced when HA proteins carried high mannose glycans while HA inhibition and neutralizing antibody titers were higher with complex glycans[8]. The usefulness of polysaccharides in eliciting a complete immune response is widely valued in conjugate vaccines where they are used to interact directly with B cells through a T cell-independent pathway to produce a rapid and restricted antibody response[9]. Thus, future glycoengineered viral particles could lead the way to self-adjuvanted viral vaccines[10]. However, the most important element representing implications in vaccine development is choosing the best-suited cell-based platform for viral particle production, one of the most

relevant challenges nowadays. Viruses depend almost entirely on the host cell glycosylation machinery[11]. Thus, the recombinant production of viral particles based on viral component, such as HIV-1 Gag-based virus-like particles (VLPs), relies on the host cell biosynthesis pathways for post-translational modifications[12]. Different expression systems like plant or insect cells are among the most used platforms for viral vaccine production due to high titers[13] while producing glycans that are substantially different from human glycosylation, influencing the immunological properties of the subsequent viral particles[14]. Although displaying specific exogenous glycan forms, like fungal glycans, can activate C-type lectin receptors (CLRs) and promote an immune response[15], the use of biopharmaceuticals produced in non-human expression systems for clinical applications requires special attention. Exogenous immunologically active glycans may cause allergy reactions, therapy rejection or other different side effects[16,17]. Still within mammalian cell systems, non-human platforms like Chinese hamster ovary (CHO) cells present complex *N*-glycans capped with *N*-glycolylneuraminic acid (NeuGc), largely absent in human cells[18,19]. Also, human cell lines can generate very different glycosylation patterns depending on their progeny, culture conditions or immortalization methods[20]. Altogether these considerations make the glycomic characterization of the biotherapeutic derived from the chosen production platform a requirement for further clinical applications. In the case of HIV-1 Gag VLPs, the human embryonic kidney (HEK) cell line is considered as a reference for its production, since they are easily transfected and can be cultured in suspension, resulting in good scalability[21,22]. VLPs are nanostructures resembling the nature conformation of the virus but lacking genetic information, hence they are non-infective while carrying viral antigens to elicit an immune response. Gag VLPs are produced via budding from the cell membrane, carrying the host cell membrane glycoproteins. Understanding the glycosylation displayed in VLPs produced in HEK293 cells could help assess and study their immunogenic potential. Moreover, characterizing the glycan profile of Gag VLPs produced in HEK293 cells could offer relevant information that could be used to design purification strategies and new methods to separate VLPs and extracellular vesicles (EVs), a current remaining challenge[23,24]. In this work, we report for the first time, to the best of our knowledge, the characterization of the *N*- and *O*-glycan profile of Gag VLPs produced in HEK293 cells. In order to do this, the analysis by porous graphitized carbon (PGC) liquid chromatography coupled to tandem mass spectrometry (LC-

MS/MS) has been selected due to the advantages of the method. PGC-LC-MS/MS is able to separate structural and linkage isomers, providing a detailed and specific glycosylation fingerprint[25–28]. The insight of these analyses may help further optimize the design of VLP-based vaccines via glycoengineering approaches and provide new ideas for purification steps that could pave the way towards better bioprocesses, vaccines and therapies.

MATERIALS AND METHODS

HEK293 MAMMALIAN CELL LINE, CULTURE CONDITIONS

The cell line used in this work is a serum-free suspension-adapted HEK293 cell line (HEK293SF-3F6, NRC, Canada) kindly provided by Prof. Amine Kamen from McGill University (Montreal, Canada). Cells were cultured in disposable polycarbonate 250 mL flasks with vent cap (Corning®) at 37°C, 5% of CO2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell culture media was HyCell™ TransFx-H media from HyClone™ (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, Life Technologies, ThermoFisher, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, Life Technologies). Cell concentration and viability were determined using the NucleoCounter®NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

TRANSIENT TRANSFECTION

Transfections were carried out at a cell density of $2 \cdot 10^6$ cells/mL using a final DNA concentration of 1 μg/mL. PEI/DNA complexes were formed by adding PEI to plasmid DNA diluted in fresh culture medium (10% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus-transfection, Illkirch-Graffenstaden, France) was used. Briefly, the corresponding plasmid was diluted with supplemented HyCell™ culture media and vortexed for 10 s. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, incubated for 15 min at room temperature (RT) and then added to the cell culture. The plasmid

used encoded a HIV-Gag protein fused in frame to eGFP (Gag::eGFP). As a transfection control, a plasmid sharing the same backbone but lacking the Gag::eGFP gene was used and noted as mock.

HIV-1 GAG VLP QUANTIFICATION BY FLUORIMETRY

The extracellular concentration of HIV-1 Gag VLPs was assessed by fluorimetry using an *ad hoc* developed and validated quantification assay[29]. VLP-containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Relative fluorescence unit values (RFU) were calculated by subtracting fluorescence unit (FU) values of non-transfected negative control samples.

ULTRACENTRIFUGATION

VLP-containing supernatants were recovered by cell culture centrifugation at 1000×g for 5 min. Then, concentrated and purified HIV-1 Gag VLPs were obtained by single sucrose cushion ultracentrifugation. Briefly, a volume of 35 mL of clarified supernatant from HEK293 transiently transfected cell culture was layered on top of a 30% sucrose cushion of 3 mL and centrifuged at 31.000 rpm for 2.5h at 4°C using a SW32 rotor in a Beckman Optima L100XP centrifuge (Brea, CA, USA). After ultracentrifugation, supernatant was discarded and the obtained pellet was resuspended in 1 mL of PBS (Hyclone, GE HeathCare, Chicago, IL, USA). The concentrated material was stored at −80°C for future studies.

EXTRACELLULAR VESICLES AND VIRUS-LIKE PARTICLES LYSIS

Total protein concentration was quantified using Micro BCA™ Protein Assay Kit (23235, ThermoFisher Scientific, Waltham, MA, USA) according to manufacturer's instructions. Membrane lysis was performed using lysis buffer without detergents: 50 mM Tris-HCl Sigma-Aldrich (Steinheim, Germany), 100 mM NaCl (Sigma-Aldrich), 1mM EDTA (Sigma-Aldrich) and one tablet of protease inhibitor cocktail (Complete ULTRA Tablets Mini, Roche, Basel Switzerland). Briefly, the corresponding volume of purified samples containing 20 μg of protein was mixed at a ratio 9:1 with concentrated lysis buffer, vortexed for 30 s and followed by two cycles of heating at 60°C for 15 min and sonication for 30 min in an ultrasound bath at RT.

GLYCOPROTEIN AND VESICLE LYSATE IMMOBILIZATION, DENATURATION AND REDUCTION

A specific method for glycan release and reduction was carried out[27]. A 96-well plate with hydrophobic Immobilon-P PVDF membrane (MultiScreen® HTS 96 multiwell plates, pore size 0.45 μm, Millipore, Burlington, MA, USA) was preconditioned with three times 200 μL 70% ethanol followed by equilibration three times with 200 μL water. Previously prepared lysates as well as intact EVs and VLPs were applied to the PVDF membrane directly. Purified fetuin glycoprotein (Sigma-Aldrich) was used as a standard and also directly applied to the PVDF membrane followed by horizontal shaking for 30 min. Protein denaturation was achieved by applying 75 μL of denaturation mix (72.5 μL 8 M GuHCl (ThermoFisher Scientific) and 2.5 μL 200 mM DTT (Sigma-Aldrich)) in each well, followed by 10 min shaking and incubating at 60°C in a moisture box for 30 min. Subsequently the unbound material was removed by washing three times with water each time followed by centrifugation at 500 × g for 1 min. Any residual liquid was subsequently removed from the membrane by pipetting and discarded.

ENZYMATIC *N*-GLYCAN RELEASE, REDUCTION AND PURIFICATION

N-glycan release was performed as previously described[30] with small modifications. PNGase F (Roche Diagnostics, Mannheim, Germany, 2 U of enzyme diluted with water up to 15 μL) was added to each well followed by 10 min shaking at RT. The plate was incubated in a humidified box for 15 min at 37 °C. Subsequently, another 15 μL of water was added in each well, including 10 ng of maltoheptaose (DP7; M7753, Sigma-Aldrich) as an internal standard. The plate was then incubated overnight in a humidified plastic box at 37 °C to avoid evaporation of the digestion solution. Released *N*-glycans were recovered from the PVDF plate by three rounds of centrifugation at 1000 × g for 2 min and washing with 40 μL of water. To hydrolyze *N*-glycan glycosylamine forms, 10 μL of 100 mM ammonium acetate (Sigma-Aldrich) (pH 5) were added followed by a 1 h incubation at RT. Samples were subsequently dried in a SpeedVac concentrator 5301 (Eppendorf, Hamburg, Germany) at 40°C. Collected *N*-glycans were then reduced and desalted followed by PGC cleanup using a 96-well plate based protocol previously reported[25]. Briefly, for the reduction, 20 μL of 1 M NaBH$_4$ (Sigma-Aldrich) in 50 mM KOH (Sigma-Aldrich) was added to each well followed by incubation for 3 hours at 60°C in a humidified plastic box. To quench the reaction, 3 μL of

glacial acetic acid (Honeywell Fluka. Charlotte, CA, USA) was added to each sample. Desalting of the samples was performed using a strong acidic cation exchange resin Dowex 50W-X8 (Sigma-Aldrich) which was self-packed into 96-well filter plates. Briefly, 100 μL of resin slurry in methanol (50/50, v/v) was added to each well in the filter plate. The columns were preconditioned by 3 × 100 μL of 1 M HCl, followed by 3 × 100 μL methanol and 3 × 100 μL water each time removed by centrifuging at 500 × g. The samples containing N-glycans were loaded onto the columns and eluted two times with 40 μL of water followed by centrifugation at 500 × g. The combined flow-through and wash were collected and dried in a SpeedVac concentrator at 40°C. The remaining borate was removed by several rounds of co-evaporation using 100 μL methanol in the SpeedVac concentrator at 40°C. SPE cleaning step was performed by packing 60 μL (approximately 6 mg) of bulk sorbent Carbograph slurry (Grace Discovery Sciences, Columbia, SC, USA) in methanol into a 96-well filter plate. The columns were preconditioned by 3 × 100 μL of 80% acetonitrile in water containing 0.1% TFA, and 3 × 100 μL water with 0.1% TFA. After loading the sample, the columns were washed three times with 80 μL of 0.1% TFA, followed by N-glycan elution by 3 × 40 μL of 60% acetonitrile in water containing 0.1% TFA. Samples were dried in a SpeedVac concentrator directly in polymerase chain reaction (PCR) plates for injection and re-dissolved in 10 μL of water prior to PGC nano-LC-ESI-MS/MS analysis.

O-GLYCAN RELEASE AND PURIFICATION

After removal of N-glycans, O-glycans were released from the same PVDF membrane immobilized sample via reductive ß-elimination. Briefly, 50 μL of 0.5 M $NaBH_4$ in 50 mM KOH was applied onto each PVDF membrane well after rewetting with 3 μL of methanol. 10 ng of maltopentaose (DP5; M8128, Sigma-Aldrich) was added as an internal standard. Plates were then shaken for 15 min on a horizontal shaker and incubated in a humidified plastic box for 16 h at 55°C. After incubation and cooling to RT, released O-glycans were recovered by centrifugation at 1000 × g for 2 min into 96-well collection plates. The wells were rewetted by 3 μL of methanol and washed three times with 50 μL of water with 10 min incubation steps on a horizontal shaker prior to centrifugation at 500 × g for 2 min. Prior to desalting, the collected samples were concentrated to approximately 30 μL under vacuum in a SpeedVac concentrator at 40°C for

2 h. Subsequently, 3 μL of glacial acetic acid was added to acidify the reaction followed by brief centrifugation to collect the sample at the bottom of the well. The following high throughput desalting and PGC SPE purification were performed as described in the *N*-glycan preparation section. The purified *O*-glycan alditols were re-suspended in 10 μL of water prior to PGC nano-LC-ESI-MS/MS analysis.

ANALYSIS OF RELEASED *N*- AND *O*-GLYCAN ALDITOLS USING PGC NANO-LC-ESI-MS/MS

The analysis of *N*-glycans and *O*-glycans was performed on an Ultimate 3000 UHPLC system (Thermo Fisher Scientific) equipped with a Hypercarb PGC trap column (5 μm Hypercarb Kappa, 32 μm × 30 mm, Thermo Fisher Scientific) and a Hypercarb PGC nano-column (3 μm Hypercarb 75 μm × 100 mm, Thermo Fisher Scientific) coupled to an amaZon ETD speed ion trap (Bruker, Bremen, Germany). Mobile phase A consisted of 10 mM ammonium bicarbonate (ABC), while mobile phase B was 60% (v/v) acetonitrile/10 mM ABC. To analyze glycans, 2 μL of prepared sample was injected and trapped on the trap column using a 6 μL /min loading flow in 2% buffer B for *N*-glycan and 1% buffer B for *O*-glycans for 5 min. Separation was achieved with a multi-step gradient of B: 2–9% in 1 min and 9–49% in 80 min for *N*-glycan and 1–52% over 72 min for *O*-glycans followed by a 10 min wash step using 95% of B at a flow of rate of 0.6 μL /min. The column was held at a constant temperature of 45°C. Ionization was achieved using the nanoBooster source (Bruker) with a capillary voltage of 1000 V applied, dry gas temperature of 280°C at 5 L/min and nebulizer at 3 psi. Isopropanol enriched dopant nitrogen was used. MS spectra were acquired from *m/z* 500–1850 for *N*-glycans and 380–1850 for *O*-glycans in enhanced mode, smart parameter setting (SPS) was set to *m/z* 1200 for *N*-glycans and *m/z* 900 for *O*-glycans; ion charge control (ICC) to 4×10^3 and maximum acquisition time to 200 ms. MS/MS spectra were generated using collision-induced dissociation over a *m/z* range from 100 to 2500 on the top three most abundant precursors, applying an isolation width of 3 Thomson. The fragmentation cut-off was set to 27% with 100% fragmentation amplitude using the Enhanced SmartFrag option from 30–120% in 32 ms and ICC was set to 150,000.

DATA PROCESSING

Glycan structures were assigned based on glycan composition obtained from accurate mass, relative PGC elution position, MS/MS fragmentation pattern in negative-ion mode and general glycobiological knowledge[31,32] with help of Glycoworkbench[33] and Glycomod[34] software tools. Extracted ion chromatograms were used to integrate the area under the curve for each individual glycan isomer using Compass Data Analysis software v.5.0 (Bruker). The most abundant peaks in the glycan profile were manually picked and integrated. Relative quantitation of individual glycans was performed on the total area of all included glycans within one sample normalized to 100%.

PARTICLE SIZE MEASUREMENT

Nanoparticle tracking analysis (NTA) was performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) equipped with a blue laser module (488 nm) to quantify HIV-1 Gag::eGFP VLPs and neutral density filter for total particle by light scattering. Data were analyzed with NanoSight® NTA 3.2 software. Briefly, samples were injected, and three technical replicate analyses were carried out. Subsequently, particles were identified and tracked by their Brownian motion at RT. Capture settings were recorded with a sCMOS camera (camera level of 8 for Gag::eGFP VLP samples, and 11 for controls, viscosity: 0.9 cP) and analyzed with a detection threshold of 4.

EXPERIMENTAL DESIGN AND STATISTICAL RATIONALE

For the comparative relative quantification glycomic experiments, three conditions were tested, with two biological replicates (independent cell cultures) from each: no transfection (N), transfection with mock plasmid (M), and transfection with the plasmid coding for Gag::eGFP protein following the standard protocol previously described (S). Samples of each condition were taken at 72 hours post-transfection (hpt) followed by an ultracentrifugation to purify HIV-1 Gag VLPs and extracellular vesicles (EVs) with the same density. As it is depicted in **Figure 1A**, three technical replicates were analyzed from each sample. Fold changes were calculated from the average relative abundance value of each glycan species in the three studied conditions, noted as N, M and S as previously described.

A

B

C
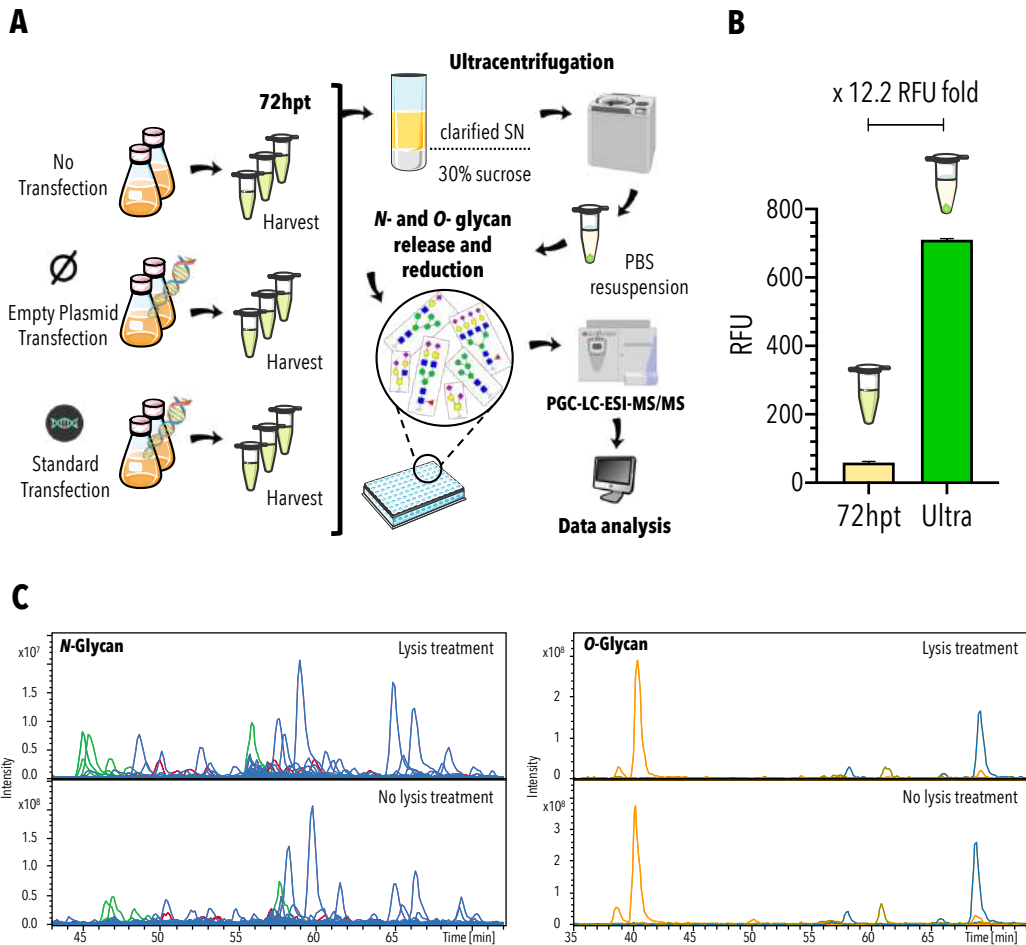
Figure 1. Experimental design and VLP production analysis. **A)** Experimental **w**orkflow. Two biological replicates of HEK293 cell cultures were grown under three different conditions: no transfected condition, transfected with an empty plasmid denoted as "mock," and transfected using the standard protocol with the plasmid coding for the Gag::eGFP polyprotein. Both transfections were performed at the same cell density of 2×10^6 cells/mL. At the time point of 72 hpt, cultured samples were taken and centrifuged and supernatants were ultracentrifuged at 31,000 rpm using a 30% sucrose cushion. The resulting pellet was resuspended in 1mL of PBS. The concentrated sample was subjected to N- and O-glycan release and reduction. Resulted alditols were analyzed by porous grafitized carbon separation coupled to liquid chromatography, electrospray ionization and tandem mass spectrometry (PGC-LC-ESI-MS/MS) and further analyzed to assign the glycosylation structures. **B)** Concentration via ultracentrifugation measured in relative fluorescence units (RFU). Harvested VLPs were concentrated 12.2-fold in the analytical ultracentrifugation step. **C)** Extracted ion chromatograms (EICs) profiles of samples that underwent lysis treatment and samples that contained intact VLPs for glycan analysis (no lysis treatment).

RESULTS

HEK293 VLP production

After transfecting HEK293 cells with PEI:DNA complexes containing the plasmid encoding Gag::eGFP polyprotein, cell cultures were harvested at 72hpt. The clarified supernatants containing VLPs were then concentrated using analytical ultracentrifugation. The resulting pellet was resuspended in PBS, increasing fluorescent VLP concentration in 12.2-fold (**Figure 1B**). These concentrated VLP samples were used for further glycan analyses.

Analysis of N-glycans

In order to analyze the *N*-glycans species present in the Gag VLP membrane, some samples underwent a lysis treatment to disrupt the particles and others maintained the complete structures prior to *N*-glycan release and LC-MS/MS analysis. These first measurements revealed highly comparable *N*-glycan profiles of lysed *versus* intact particles (**Figure 1C**). Relative abundance and elution patterns of glycans were maintained, with differences in signal intensity. All identified *N*-glycans were observed in both lysed and intact sample conditions. A total of 51 *N*-glycan species were identified and 48 quantified. All characteristics and quantitative data relating to these 51 species are presented in **Supplementary Table S1**. Overall, 38 glycan species were complex-type. Of these, 31 were sialylated, 1 contained a sulfate group and 7 presented Lewis epitopes (**Figure 2A-B**). Lewis epitopes or antigens are fucosylated oligosaccharides derived from Galβ1-3/4(Fucα1,4/3)GlcNAcβ1- structures. The Gal residue can be sialylated, sulfated, also fucosylated or remain unsubstituted. From the total 51 identified species, 7 were high mannose *N*-glycans. In these structures, the α(1,3) and α(1,6) arms contained only mannoses whose specific branching pattern was not resolved. Additionally, hybrid *N*-glycans were detected, which is a group of glycans with a mannosidic α(1,6) arm and a complex α(1,3) arm containing GlcNAc, Gal and NeuAc. This group accounted for 6 of the 51 *N*-glycans (**Figure 2B**) and 3 out of these 6 hybrid glycans presented NeuAc as terminal modifications.

A

B

C

● Mannose ■ *N*-Acetylglucosamine ● Galactose □ *N*-Acetylgalactosamine ◆ *N*-Acetylneuraminic acid (Sialic acid) ▲ Fucose
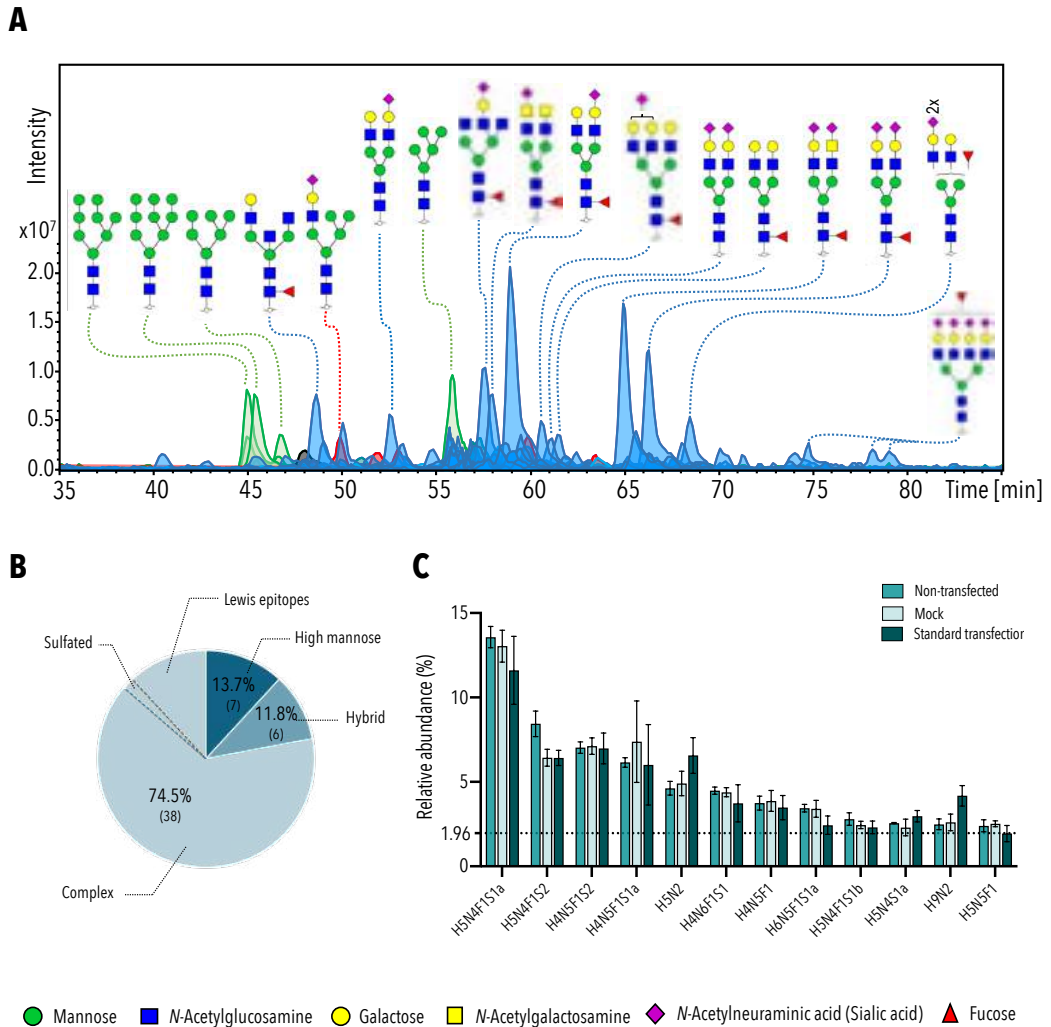
Figure 2. Analysis of *N*-glycan species from Gag VLPs. **A)** Combined extracted ion chromatograms (EIC) of 51 *N*-glycan species released from VLP samples. Structures of the 17 most abundant glycans are depicted. Complex glycans represented in blue, hybrid glycans represented in red and oligomannoses or high mannose glycans represented in green. **B)** Classification of the different types of *N*-glycan species identified in the sample. Glycans presenting sulfate and Lewis epitopes represent a subdivision within complex glycans. **C)** Relative abundance distribution of the most abundance twelve *N*-glycan species identified in the VLP sample. Error bars represent standard deviation. The average value of all 51 glycans (1.96) is represented by the dotted line. H: hexose, N: *N*-Acetylhexosamine, F: fucose, S: *N*-Acetylneuraminic acid, Su: Sulfate.

The measured glycans differed greatly in relative abundance. The most abundant glycan, accounting for ~13% of relative abundance was H5N4F1S1a observed at *m/z* 1038.89 (charge 2-) which may be represented as $Hex_2 HexNAc_2 Deoxyhexose_1 NeuAc_1 Man_3 GlcNAc_2$ (**Figure 2A-2C**).
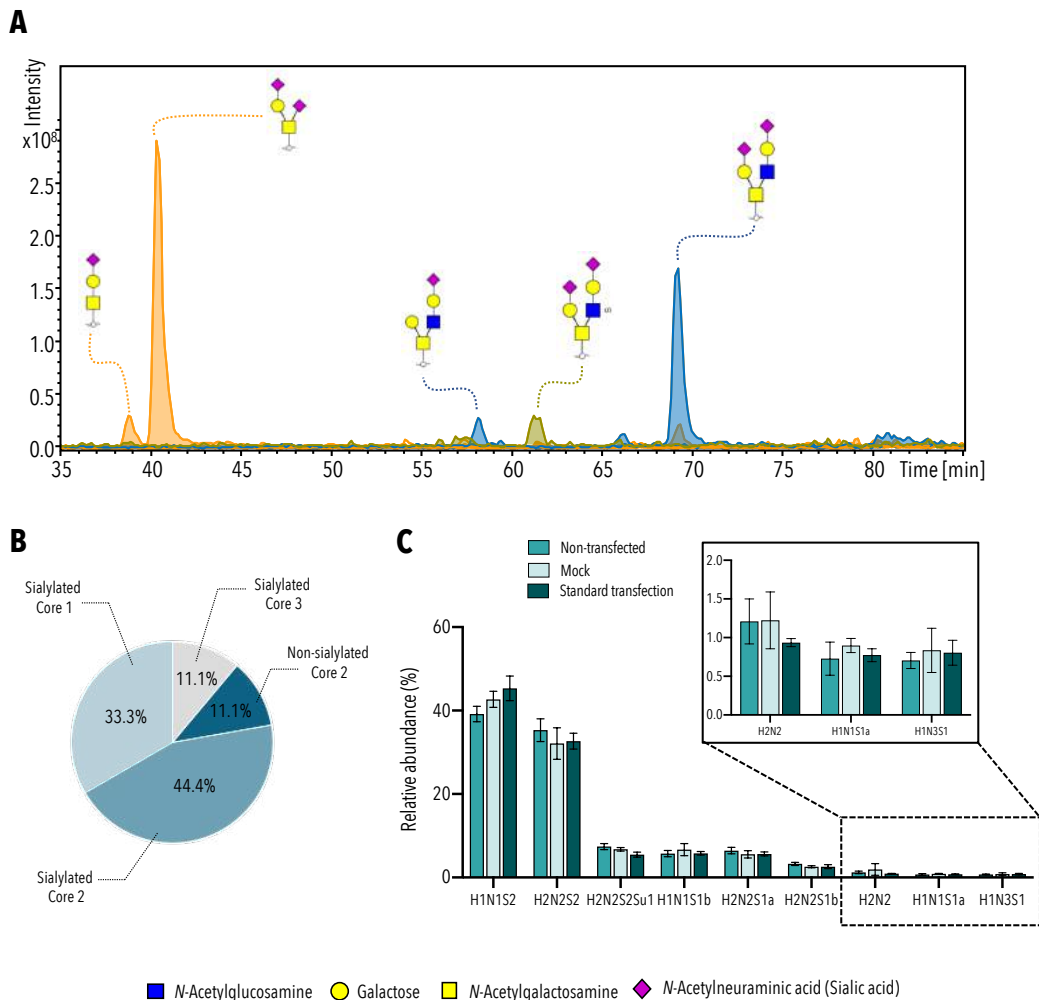
Figure 3. Analysis of *O*-glycan species from Gag VLPs. **A)** Combined extracted ion chromatograms (EIC) of 9 *O*-glycan species released from VLP samples in which the top five most abundant glycans account for 94% of relative abundance. **B)** Classification of the different types of *O*-glycan species identified in the sample. **C)** Relative abundance distribution of the nine *O*-glycan species identified in the VLP sample. Error bars represent standard deviation. Significance calculated by two-way ANOVA. H: hexose, N: *N*-Acetylhexosamine, S: *N*-Acetylneuraminic acid, Su: Sulfate.

Analysis of O-glycans

The *O*-glycan profile of VLP sample is exemplified in **Figure 1C**. A total of 9 *O*-glycans were identified, from which 8 were sialylated, presenting core 1, 2 and core 3-type structures. The only non-sialylated *O*-glycan presented core 2-type structure and one of the lowest abundances (**Figure 3A-B**). Isobaric *O*-glycan structural isomer differentiation was successfully carried out by PCG-LC separation method coupled with

Figure 4. Exemplification of the PGC-LC separation of *O*-glycan structural isomers and MS/MS structural features. **A)** Extracted ion chromatogram (EIC) of the *m/z* 1040.52^{1-} (Hex$_2$ HexNAc$_2$ NeuAc$_1$) *O*-glycan isomers species. **B)** MS2 spectrum of the *m/z* 1040.52^{1-} isomer observed at RT 58.1 min. **C)** MS2 spectrum of the *m/z* 1040.52^{1-} isomer observed at RT 66.1 min. Diagnostic ions denoted in red. H: hexose, N: *N*-Acetylhexosamine, S: *N*-Acetylneuraminic acid.

negative ion tandem mass spectrometry (**Figure 4**). For instance, this method allowed to resolve the position of the sialic acid in the *m/z* 1040.52^{1-} based on the identification of characteristic diagnostic ions (**Figure 4B-C**). Within the total of 9 identified *O*-glycans species, the top two most abundant species accounted for the 76% of relative abundance, being the most abundant the tetrasaccharide H1N1S2 with an observed *m/z* of 966.4 followed by the hexasaccharide H2N2S2 of *m/z* 1331.6 (**Figure 3C**). These two species are common *O*-glycosylations found in many human proteins. No significant differences were observed between non-transfected, transfected with mock and transfected with Gag::eGFP conditions (**Figure 3B-C**). Similar to *N*-glycans, *O*-glycan profiles were not influenced by lysis (**Figure 1C**). All the identified and quantified *O*-glycans are presented in **Supplementary Table S2**.

Changes in glycan signature upon transient transfection and Gag VLP production

The most abundant *N*- and *O*-glycan species identified in the analyzed samples did not show significant changes in the non-transfected, transfected with mock and transfected with the plasmid coding for Gag::eGFP studied conditions. The *N*-glycosylation patterns were very similar for the different samples (**Figure 2C**). Still, some *N*-glycans were found to show small, significant differences in relative abundance. One of the four isomers (H7N6F1S4a) of *N*-glycan H7N6F1S4 showed a significant increase in the relative abundance fold change in M and S condition (**Figure 5A-B**). This sialylated tetrantennary *N*-glycan accounted for ~0.5% of relative abundance among all *N*-glycans in N and M conditions shifted to ~2%. Isomer H7N6F1S4c showed lower intensities in M and S compared to N condition. A shift was observed in H7N6F1S4c (retention time 78.2 min) from ~1.0% to ~0.5% upon transfection. Isomers H7N6F1S4b and H7N6F1S4d hardly differed in all studied conditions. It has been reported that upon recombinant production of human Protein C in HEK293 cells, the structure of H7N6F1S4 glycan shifted from presenting all α(2-3) linkage to a combination of α(2-3) and α(2-6) linkage of sialic acid[35]. The structural differences of isomers H7N6F1S4a and H7N6F1S4c might also be related to a change in sialic acid linkage combinations. Also, the isomer H3N5F1b showed a significant increase in S and M condition (**Supplementary Table S1**). Moreover, different *N*-glycans were observed to increase their relative abundance fold change only in S condition (**Figure 5A-B**). The biantennary H3N5F1a, bearing a bisecting GlcNAc, a β (1,4)GlcNAc attached to the innermost mannose residue, was observed to slightly increase its relative abundance in VLPs, from ~0.4% to ~0.7%. Structures presenting a bisecting GlcNAc have been reported to be involved in relevant immunological processes like immune tolerance and immune supression[36]. The correct separation and identification of structural isomers H3N5F1a and H3N5F1b is illustrated in **Figure 6.** The characteristic diagnostic ions from the bisecting structure (D-221) and from the non-bisecting isomer (D and D-18 ions) allowed the correct assignment and identification (**Figure 6B-C**). Other *N*-glycans presenting an increased relative abundance in S condition were three of the identified high mannose *N*-glycans; H7N2a, H9N2 and H8N2. High mannose *N*-glycans residues have been reported to influence pharmacokinetics of therapeutic molecules like IgGs, reducing their half life[37,38]. Interestingly, the other *N*-glycan increasing in S condition and therefore in the VLP fraction, was H5N5 (**Figure 5A-B**), a

Figure 5. Relative abundance changes in glycans of produced particles upon transfection and VLP production. **A)** Heatmap illustrating relative abundance and fold change ratios of *N*-glycans showing changes in the different studied conditions (left). Replicates are denotated as I, II, III and IV. **B)** Bar plots presenting the changes group by categories. Error bars represent standard deviation. Significance calculated via one-way ANOVA. Gag transfection (S) and mock transfection (M) enrichment and dehancement represent changes due to transfection. S enrichment and dehancement represent changes due to VLP production. **C)** Heatmap illustrating relative abundance and fold change ratios of *O*-glycans showing changes in the different studied conditions (left). Replicates are denotated as I, II, III and IV. H: hexose, N: N-Acetylhexosamine, F: fucose, S: *N*-Acetylneuraminic acid, Su: Sulfate.
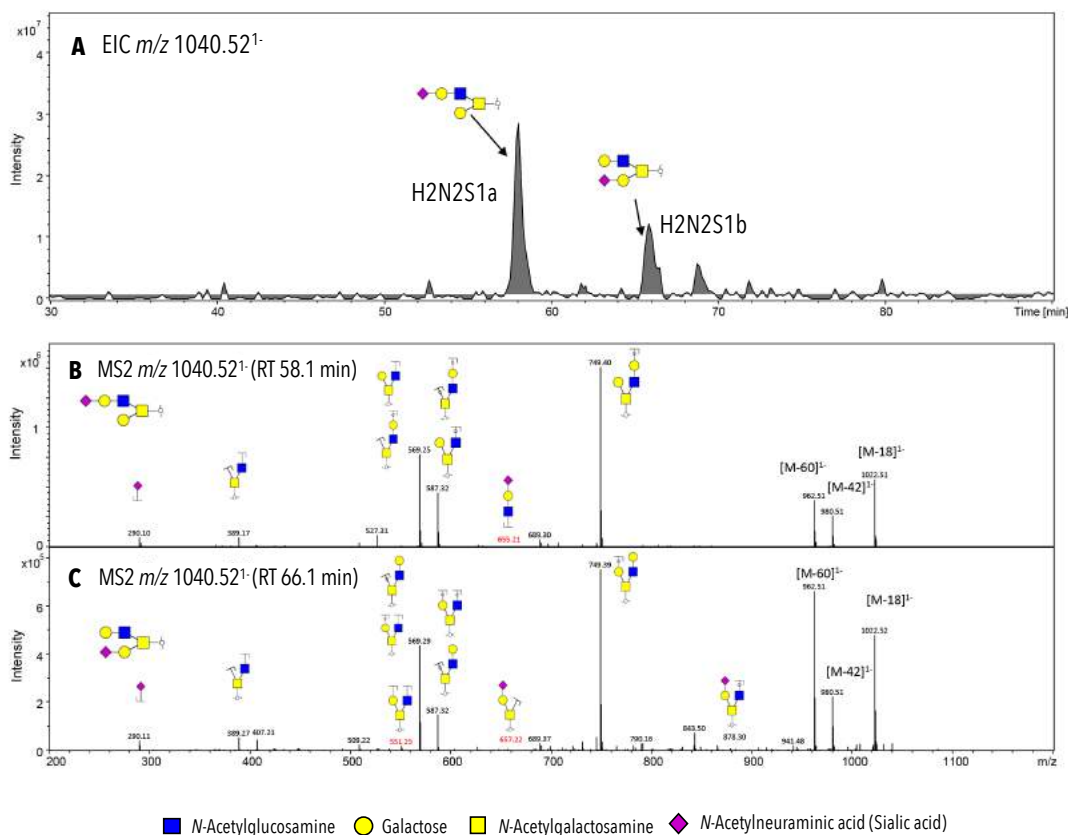
Figure 6. Exemplification of the PGC-LC separation of *N*-glycan structural isomers and MS/MS structural features. **A)** Extracted ion chromatogram (EIC) of the *m/z* 832.81^{2-} (Hex$_3$ HexNAc$_5$ Fuc$_1$) *N*-glycan isomers species. **B)** MS2 spectrum of the *m/z* 832.81^{2-} isomer observed at RT 49.8 min. **C)** MS2 spectrum of the *m/z* 832.81^{2-} isomer observed at RT 59.9 min. Diagnostic ions denoted in red. H: hexose, N: *N*-Acetylhexosamine, F: fucose.

glycosylation found in glycoprotein CD59[39,40]. Glycoprotein CD59 is commonly present in HIV-1 virions preventing the action of the complement cascade and the lysis of the viral particles[41]. There are also *N*-glycan species that decreased upon transfection and VLP production. For instance, the second most abundant *N*-glycan, H5N4F1S2 (**Figure 5A-B**). Among the *N*-glycans that showed lower intensity only upon VLP production, the greatest decrease, 0.3-fold, was found in H5N5F1S1b, having a $\alpha(1,3)$ sialic acid linkage. The hybrid H7N3F1 glycan also presented a reduction of 0.5-fold in relative abundance in S condition. As for the *O*-glycan species, no significant differences were observed between non-transfected, transfected with mock and transfected with Gag::eGFP conditions (**Figure 5C**).

VLPs present lower glycan density per particle than EVs

Particle concentration and the size distribution curve in each condition were measured by nanoparticle tracking analysis (NTA) showing a 5.4-times increased concentration of particles in the S condition compared to N and M conditions (**Figure 7A**). From the total $2.88 \cdot 10^{11}$ particles/mL measured in S condition, 82.3% corresponded to fluorescent particles (Gag::eGFP VLPs). The remaining 17.7% corresponded to coproduced EVs. We next aimed at estimating the glycan content per particle. To get an estimate of the glycan amount in each condition, glycan intensity was normalized to the intensity of the internal standard DP7. Then, assuming complete release of glycans and similar response factors between condition and DP7 standard, the number of glycans per particle was estimated (**Figure 7B**). The condition containing the VLP fraction (S condition) presented a significantly lower glycan density than N and M condition. Interestingly, M condition also presented a reduction of glycan density compared to N condition (**Figure 7B**). A reduction from an average of 2400 and 1620 *N*-glycan molecules per particle for the N and M condition respectively to 380 for the S condition was assessed. For *O*-glycans, it dropped from 6830 and 4300 *O*-glycans/particle for the N and M condition respectively to 1150 for the S condition. This is schematically represented in **Figure 7C**.

DISCUSSION

The analysis and characterization of the glycosylation signature of biopharmaceuticals is essential to understand their effect on eliciting a correct immune response. The glycosylation pattern depends exclusively on the cell-based platform where it has been produced. For human application, it is fundamental to assure that the final bioproduct does not present exogenous and allergenic glycoforms that could neutralize and handicap the envisaged function of the biopharmaceutical. Although some xeno-glycomotifs can have a desirable effect of boosting the immune response, others can cause allergy reactions[42]. HIV-1 Gag VLPs conceived as vaccine candidates are commonly produced in insect and plant cells expression systems in order to obtain higher titers. However, the presence of α1,3-fucosylations in the core of *N*-glycans produced in insect cells or the presence of both allergenic α1,3-fucosylations and
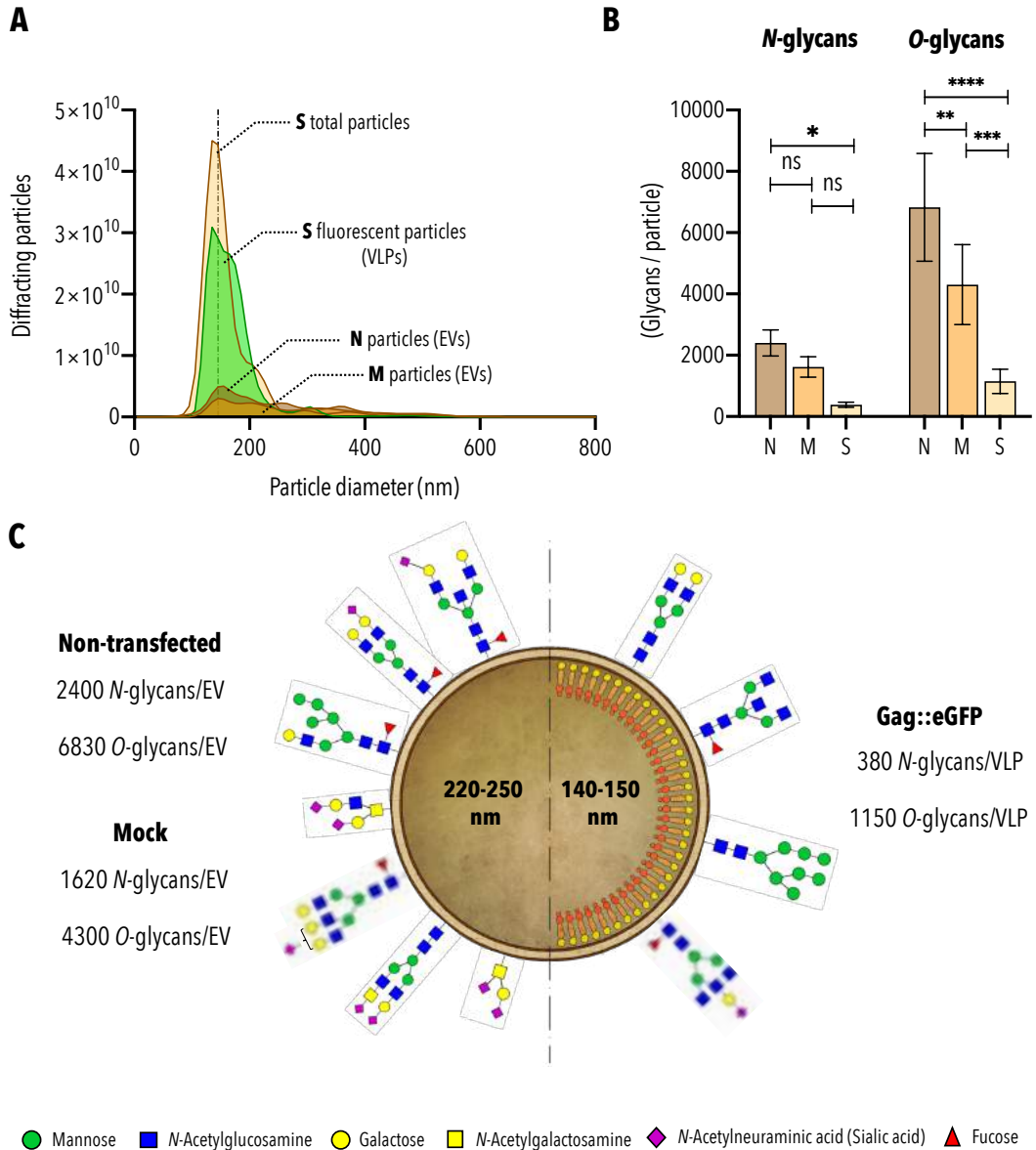
Figure 7. Analysis of diffracting particles. **A)** Size distribution curve of diffracting particles in each condition measured by nanoparticle tracking analysis (NTA). The vertical grey line represents N: non-transfected condition, M: condition transfected with mock plasmid, S: condition transfected with plasmid encoding Gag::eGFP polyprotein, VLPs: virus-like particles, EVs: extracellular vesicles. **B)** Glycan density per particle. The total intensity area of identified glycans in each condition was normalized by their corresponding internal standard DP7 intensity area and divided by the corresponding number of particles in the sample. The number of particles was calculated using the particle concentration value obtained by NTA. Significance was calculated using two-way ANOVA. **C)** Schematic representation of the glycan density in EVs (left) and Gag VLPs (right). The mean value of the particle diameter measured by NTA is represented within the particle. The represented glycan species in EVs (left) are the most abundant and the ones showing downregulation in S condition. The represented glycans in VLPs (right) are the ones showing an increase in S condition. The values showing the number glycan/particle are the approximate value calculated assuming the same response factor of the standard DP7 with known concentration, used in each condition.

β1,2-xyloses in *N*-glycans produced in plant cells increment the complexity of the bioprocesses using these expression systems. Additional steps to remove these glycoforms are needed as the final purified VLPs cannot contain them for human applications. In this work, we have characterized the glycosylation profile of Gag VLPs produced in HEK293 cells, showing that the displayed glycans presented structures of high mannose, hybrid and mostly complex glycans. For instance, Chikungunya VLPs produced in Sf9 insect cells presented predominantly high mannose structures[14], needing to complement the VLPs with adjuvants to achieve the same neutralizing antibody titers as VLPs produced in mammalian cells[43]. High mannose structures have been described to produce fewer neutralizing antibody titers than complex glycans[8]. The lower neutralizing antibody titers generated by high mannose glycans may be explained by the glycan shielding effect that they provide to viral particles. Often, immunogenic surface proteins from viruses like Ebola GP[44], Influenza HA[45], HIV-1 Env glycoprotein[46–48] or SARS-CoV-2 S protein[49,50] are masked in high mannose structures preventing detection from the immune system. This is one of the main reasons for vaccine strategies to target the surface protein, not considering the glycan shield and leading to low neutralizing titers. Displaying not only high mannose but also hybrid and complex glycans contributes to a stronger immunological response[8]. Most importantly, VLPs produced in HEK293 cells harbored sialic acid terminations in 63.5% of all identified glycans, between *N*- and *O*-glycans. This feature improves the recognition of VLPs by macrophages and further presentation to antigen-presenting cells (APC). Macrophage recognition is promoted via Siglec-1 receptor, which selectively binds to $\alpha(2,3)$-linked sialic acids, enhancing the immune response[51].

The relative abundance of *N*- and *O*-glycans identified in this study has been observed to change upon transfection and VLP production. The changes observed upon transfection correlated with the generation of small EVs (sEVs), previously reported[52]. The increase in the generation of vesicles smaller than 200 nm upon transient transfection in HEK293 cells correlates with the increase of large tetrantennary glycans like H7N6F1S4b. This could be related to the generation of EVs via de multivesicular bodies pathway and release as exosomes. This phenomenon could also be related to the reduction in glycan density in mock-derived EVs and VLPs. Since, sEVs and VLPs are smaller particles, they might deplete the cell of more membrane surface at a higher rate, reducing the cell capacity of replenishing the glycan reservoir. In fact,

the VLP specific productivity is 4.5 times higher producing VLPs than in mock condition, from (3.5 ± 0.1)$\cdot 10^3$ to (0.8 ± 0.1)$\cdot 10^3$ VLP$\cdot$cell$^{-1}\cdot$day^{-1} respectively[53]. Secreting vesicles a lower rate might provide more time for glycoproteins to be transported to the membrane. The reduction in vesicle size and increase in production rate of the generated vesicles might account for the reduction in glycan density. More interestingly, the specific changes observed in VLP glycans can shed some light on one of the greatest challenges of Gag-VLP bioprocess development: the separation of VLPs from the coproduced EVs. This is an essential part of the downstream purification process that is still under research. The usual purification process of VLPs start with clarifying the recovered supernatant to remove cells and debris followed by capture and concentration, typically performed by chromatography. The current capture technologies that are being used for VLPs are anion-exchange monolith or Capto™-Heparin columns, yielding from 2 to 13% at elution[54,55]. These current methods do not efficiently separate VLPs and EVs. The study of VLPs and coproduced EVs glycosylations can contribute to the design of more efficient affinity methods for the downstream processing step of capture, separating both population of nanoparticles. In this work, we have described differences in glycosylation abundance in VLPs and EVs, highlighting the difference in glycoforms of CD59. H5N5 *N*-glycans from CD59 significantly increased in VLPs. These changes could open the door to the design of affinity method based on these glycans to chromatographically separate both EV and VLP populations. In addition to this, the glycans that significantly decreased in VLPs, like H5N5F1S1b with α(2,3) sialic acid linkage could be used to create affinity separation methods to allow the flow-through of VLPs while retaining EVs. Either using affinity resins to capture VLPs for latter elution or retaining EVs while filtering VLPs could provide more efficient solutions for VLP purification. Apart from designing a separation method for downstream processing, knowing the glycosylation signature of Gag VLPs paves the way for glycoengineering strategies to enhance VLP immunogenicity towards the creation of self-adjuvanted and cell-specific VLP-based vaccines. One example is the use of multivalent glycan display for cell-specific CLR targeting. Specific fungal high mannose or Lewis-B or Lewis-X antigens selectively interact with CLR CD209, eliciting a robust CD4+ and CD8+ immune response[56]. This strategy has proved to achieve immunization using lentiviral nanoparticles[57], an approach that could be applied to Gag VLPs.

Moreover, the presence of *N*-glycans presenting structures containing a bisecting GlcNAc has been reported to influence immune tolerance, shielding the cells that possessed bisecting GlcNAc structures from the immune system[58,59]. In addition to this, bisecting GlcNAc has been reported to suppress further terminal modification such as sialic acid, fucose or Lewis epitopes. Losing the bisecting GlcNAc made *N*-glycans better substrates for many glycosyltransferases, increasing the number of terminal modification[60]. Therefore, in order to increase immunogenic terminal modifications and reduce shielding the VLPs from the immune system, a HEK293 *N*-acetylglucosaminyltransferase-III (GnT-III) knock out production platform could be established. This would provide VLPs rich in immunogenic terminal modifications as better vaccine candidates. Another identified feature of the analyzed VLPs is the presence of high mannose *N*-glycans. A high concentration of this type of *N*-glycan has also been reported to influence immune shielding. One relevant and illustrating example is HIV-1 Env glycoprotein. This is one of the most heavily glycosylated proteins ever studied, containing almost solely high mannose *N*-glycans[61,62]. This high concentration of high mannose glycans offer shielding to HIV-1 Env glycoprotein from the recognition of antibodies and hamper the immune response. High mannose to complex *N*-glycans ratio has also been described to directly influence the stimulation of antibody synthesis and recognition[46,63]. Consequently, in order to reduce high mannose glycans in the VLP surface, particles can be treated *in vitro* with α-mannosidases to hydrolyze terminal mannoses. However, a better option reducing the bioprocess final costs and complexity could be to establish an engineered HEK293 platform overexpressing Golgi α-mannosidases coded by genes *MAN1A1* and *MAN1A2.* Knocking out these genes has been reported to generate cells with only high mannose surface *N*-glycans[64] and overexpressing them would trim mannoses from uncapped glycans with no terminal modifications, increasing complex species and reducing the high mannose to complex *N*-glycans ratio. A different glycoengineering approach to enhance VLP immunogenicity is the expression of the α-Gal epitope. The trisaccharide structure Gal-α(1,3)-Gal-β(1,4)-GlcNAc-R is an epitope found in all non-primate mammals and absent in humans. The presence of α-Gal epitope elicits a strong immune response and has been used to boost the elicited immune response of vaccines against Influenza[65], HIV-1[66] and SARS-CoV-2[67]. These modulations based on glycoengineering could be implemented towards the development of an improved vaccine platform based on Gag VLPs.

CONCLUSIONS

The characterization of protein glycosylation in biopharmaceuticals is essential for the correct development of further clinical trials to reach human applications. HIV-1 Gag-based VLPs are promising candidates for the development of new vaccine therapies. In this work, we have characterized the glycome of VLPs, showing that they present the required complex glycosylations for a correct interaction with the immune system. Moreover, in order to design an integrated production bioprocess, key protein glycosylations have been identified for purification and separation strategies of VLPs and coproduced EVs. *N*-glycans like H5N5 and H5N5F1S1b showed significant differences in VLPs compared to EVs. This could pave the way for affinity-based separations leading to specific purification methods that would allow the optimization of downstream processing of VLPs.

DATA AVAILABILITY

The raw mass spectrometry data has been submitted and is available in GlycoPost repository with ID number GPST000169 (https://glycopost.glycosmos.org/entry/GPST000169).

SUPPORTING INFORMATION

Supplementary Table S1. All identified and quantified N-glycans. Observed N-glycan ions have been reduced. Mass deviation is calculated with the major observed charge state. Replicates are denoted as I,II,III and IV. Structural elucidation was based annotation of spectral fragments using common knowledge of biosynthetic pathways. The presence of additional structural isomers cannot be excluded. n.d., not detected. Glycans presenting N/A values for relative abundance and fold change ratios were identified in all samples but did not present a signal/noise ratio suitable for quantification.

Supplementary Table S2. All identified and quantified O-glycans. Observed O-glycan ions have been reduced. Mass deviation is calculated with the major observed charge state. Replicates are denoted as I,II,III and IV. Structural elucidation was based annotation of spectral fragments using common knowledge of biosynthetic pathways. The presence of additional structural isomers cannot be excluded. n.d., not detected.

Supplementary Table S1. All identified and quantified N-glycans. Observed N-glycan ions have been reduced. Mass deviation is calculated with the major observed charge state. Replicates are denoted as I, II, III ans IV. Structural elucidation was based annotation of spectral fragments using common knowledge of biosynthetic pathways. The presence of additional structural isomers cannot be excluded. n.d., not detected. Cyclans presenting N/A values for relative abundance and fold change ratios were identified in all samples but did not present a signal/noise ratio suitable for quantification.

Observed ions			Theoretical m/z			Deviation	Glycan Number	Name	Composition	Probable structure	Type	Retention time (min)	Relative abundance (%)												Relative abundance (Average %)			Fold change Ratios		
[M-H]-	[M-2H]²-	[M-3H]³-	[M+H]+calc	[M-2H]²-calc	[M-3H]³-calc	Δ[M-H]-							Non-transfected (N) I	II	III	IV	Mock transfected (M) I	II	III	IV	Gag transfected (S) I	II	III	IV	AV N	AV M	AV S	S vs N	S vs M	M vs N
1235.37	n.d.	n.d.	1235.4418	617.2209	411.147267	0.07180	Glycan #1	H5N2	(Hex)2 (Man)3 (GlcNAc)2		High mannose	55.9	5.11	4.12	4.64	4.62	5.04	4.46	4.26	5.88	6.21	6.71	5.41	7.93	4.62	4.91	6.57	1.42	1.34	1.06
n.d.	698.25	n.d.	1397.4946	698.2473	465.164867	0.00270	Glycan #2	H6N2	(Hex)3 (Man)3 (GlcNAc)2		High mannose	46.7	1.26	1.24	1.28	0.97	1.56	1.79	1.05	2.03	1.58	1.94	1.70	2.12	1.19	1.61	1.84	1.54	1.14	1.35
n.d.	779.26	n.d.	1559.5474	779.2737	519.182467	0.01370	Glycan #3a	H7N2a	(Hex)4 (Man)3 (GlcNAc)2		High mannose	45.2	0.73	0.69	0.77	0.67	1.03	0.70	0.84	0.17	1.37	1.49	1.53	0.59	0.72	0.68	1.24	1.74	1.82	0.95
n.d.	779.26	n.d.	1559.5474	779.2737	519.182467	0.01370	Glycan #3b	H7N2b	(Hex)4 (Man)3 (GlcNAc)2		High mannose	46.6	0.19	0.31	0.26	0.30	0.23	0.40	0.42	0.32	0.31	0.74	0.44	0.13	0.27	0.34	0.41	1.52	1.18	1.29
n.d.	799.76	n.d.	1600.574	799.787	532.858	0.02700	Glycan #4	H6N3	(Hex)3 (HexNAc)1 (Man)3 (GlcNAc)2		Hybrid	51.1	0.71	0.59	0.67	0.77	0.73	0.69	0.64	0.44	0.84	0.74	1.04	0.17	0.68	0.62	0.69	1.02	1.11	0.91
n.d.	820.3	n.d.	1641.6005	820.30025	546.5335	0.00025	Glycan #5	H5N4	(Hex)2 (HexNAc)2 (Man)3 (GlcNAc)2		Complex	49.3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
n.d.	832.84	n.d.	1666.6322	832.8161	554.8774	0.02390	Glycan #6a	H3N5F1a	(HexNAc)3 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Complex	45.5	0.41	0.42	0.39	0.46	0.52	0.36	0.45	0.30	1.30	0.72	0.74	0.43	0.42	0.41	0.80	1.89	1.96	0.97
n.d.	832.84	n.d.	1666.6322	832.8161	554.8774	0.02390	Glycan #6b	H3N5F1b	(HexNAc)3 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Complex	53	0.20	0.23	0.15	0.16	0.43	0.20	0.26	0.43	0.46	0.57	0.32	0.38	0.19	0.33	0.43	2.32	1.31	1.77
n.d.	840.81	n.d.	1682.6271	840.81355	560.209033	0.00355	Glycan #7	H4N5	(Hex)1 (HexNAc)3 (Man)3 (GlcNAc)2		Complex	40.5	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
n.d.	853.29	n.d.	1705.6054	852.3027	567.868467	0.01270	Glycan #8	H7N2F1	(Hex)4 (Deoxyhexose)1 (Man)3 (GlcNAc)2		High mannose	52	0.49	0.45	0.34	0.40	0.39	0.59	0.44	0.45	0.50	0.41	0.56	0.53	0.42	0.47	0.50	1.19	1.06	1.11
n.d.	860.3	n.d.	1721.6003	860.30015	573.2001	0.00015	Glycan #9	H8N2	(Hex)5 (Man)3 (GlcNAc)2		High mannose	45.1	2.56	2.05	2.01	1.88	2.53	2.92	2.41	1.59	4.55	3.74	3.58	2.27	2.12	2.36	3.54	1.66	1.50	1.11
n.d.	885.31	n.d.	1771.6635	885.33175	589.887833	0.02175	Glycan #10	H4N4F2	(Hex)1 (HexNAc)2 (Deoxyhexose)2 (Man)3 (GlcNAc)2		Hybrid	51.9	0.78	0.82	0.88	0.77	1.27	1.14	1.40	1.05	0.81	1.05	1.02	0.37	0.81	1.22	0.81	1.00	0.67	1.50
n.d.	893.33	n.d.	1787.6585	893.32925	595.2195	0.00075	Glycan #11	H5N4F1	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Complex	61.5	1.99	2.05	2.08	1.76	2.17	2.36	2.12	1.99	2.01	1.45	2.04	1.79	1.97	2.16	1.82	0.92	0.84	1.10
n.d.	913.83	n.d.	1828.685	913.8425	608.895	0.01250	Glycan #12	H4N5F1	(Hex)1 (HexNAc)3 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Complex	48.6	3.29	4.27	3.84	3.58	4.11	4.12	4.30	2.95	3.51	4.12	3.82	2.47	3.75	3.87	3.48	0.93	0.90	1.03
n.d.	921.84	n.d.	1844.6799	921.83995	614.226633	0.00005	Glycan #13	H5N5	(Hex)2 (HexNAc)3 (Man)3 (GlcNAc)2		Complex	41.3	0.28	0.35	0.33	0.34	0.34	0.28	0.18	0.32	0.52	0.19	0.37	0.69	0.32	0.28	0.44	1.37	1.60	0.86
n.d.	937.34	n.d.	1875.6745	937.33725	624.558167	0.00275	Glycan #14a	H5N3F1S1a	(Hex)2 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 (Man)3 (GlcNAc)2		Hybrid	53.2	1.20	0.87	0.96	0.96	0.80	0.98	0.94	1.06	1.05	1.15	1.47	0.96	1.00	0.95	1.16	1.16	1.22	0.95
n.d.	937.34	n.d.	1875.6745	937.33725	624.558167	0.00275	Glycan #14b	H5N3F1S1b	(Hex)2 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 (Man)3 (GlcNAc)2		Hybrid	63.5	0.85	0.67	0.68	0.63	0.53	0.53	0.39	0.64	0.66	0.70	0.79	0.69	0.71	0.52	0.71	1.01	1.36	0.74
n.d.	941.3	n.d.	1883.6531	941.32655	627.2177	0.02655	Glycan #15	H9N2	(Hex)6 (Man)3 (GlcNAc)2		High mannose	45.4	2.85	2.60	2.40	2.10	2.70	2.97	2.86	1.88	4.22	4.97	4.01	3.51	2.49	2.60	4.18	1.68	1.60	1.05
n.d.	945.34	n.d.	1891.6694	945.3347	629.8898	0.00530	Glycan #16	H6N3S1a	(Hex)3 (HexNAc)1 (NeuAc)1 (Man)3 (GlcNAc)2		Hybrid	50	1.49	1.44	1.41	1.26	1.10	1.35	1.48	2.00	1.29	1.55	1.98	1.76	1.40	1.48	1.65	1.18	1.11	1.06
n.d.	953.8	n.d.	1908.6847	953.84235	635.561567	0.04235	Glycan #17	H7N3F1	(Hex)4 (HexNAc)1 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Hybrid	58.3	1.69	1.86	1.67	1.79	1.76	1.49	1.60	1.68	0.91	1.25	0.80	0.72	1.75	1.63	0.92	0.53	0.56	0.93
n.d.	957.8	n.d.	1916.701	957.8505	638.233667	0.05050	Glycan #18a	H4N4F1S1a	(Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 (Man)3 (GlcNAc)2		Complex	51.8	0.90	0.90	1.00	0.98	0.66	0.56	1.03	0.55	0.71	0.70	0.76	0.79	0.95	0.70	0.74	0.78	1.06	0.74
n.d.	957.8	n.d.	1916.701	957.8505	638.233667	0.05050	Glycan #18b	H4N4F1S1b	(Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 (Man)3 (GlcNAc)2		Complex	56.9	1.82	0.82	0.69	0.66	0.65	0.67	1.31	0.54	1.62	1.11	1.69	0.32	1.00	0.79	1.19	1.18	1.50	0.79
n.d.	965.84	n.d.	1932.696	965.848	643.565333	0.00800	Glycan #19a	H5N4S1a	(Hex)2 (HexNAc)2 (NeuAc)1 (Man)3 (GlcNAc)2		Complex	52.9	2.53	2.61	2.53	2.57	2.65	2.20	2.70	1.65	2.72	3.37	3.11	2.67	2.56	2.30	2.97	1.16	1.29	0.90
n.d.	965.84	n.d.	1932.696	965.848	643.565333	0.00800	Glycan #19b	H5N4S1b	(Hex)2 (HexNAc)2 (NeuAc)1 (Man)3 (GlcNAc)2		Complex	62.5	0.75	0.83	0.75	0.78	0.85	0.72	0.55	1.25	0.65	0.86	0.87	0.48	0.78	0.84	0.71	0.92	0.85	1.08
n.d.	986.67	n.d.	1974.7429	986.87145	657.580967	0.00145	Glycan #20	H4N5F2	(Hex)1 (HexNAc)3 (Deoxyhexose)2 (Man)3 (GlcNAc)2		Complex	58.5	1.11	0.96	0.79	0.93	1.81	1.42	1.09	0.13	1.15	1.18	1.55	0.37	0.95	1.11	1.06	1.12	0.95	1.17
n.d.	994.84	n.d.	1990.7378	994.8689	662.9126	0.02890	Glycan #21	H5N5F1	(Hex)2 (HexNAc)3 (Deoxyhexose)1 (Man)3 (GlcNAc)2		Complex	50.4	2.37	2.28	2.05	2.90	2.60	2.46	2.70	2.30	2.04	2.18	2.30	1.23	2.40	2.51	1.94	0.81	0.77	1.05

	Obs. mass 1		Calc. mass 1	Calc. mass 2	Calc. mass 3	value	Glycan ID	Code	Composition	Type	%																							
n.d.	1005.81	n.d.	2012.6528	1005.8264	670.2176	0.01640	Glycan #22	H5N4S1Su1	Hex5 (HexNAc4)2 (NeuAc1 (Sulph1 (Man3)[GlcNAc2	Complex	53.1	1.33	1.35	1.64	1.25	1.31	1.31	1.51	1.51	1.92	1.56	1.84	0.97	1.15	0.79		1.39	1.58	1.19		0.85	0.75	1.13
n.d.	1038.89	n.d.	2078.7539	1038.87695	692.2513	0.01305	Glycan #23a	H5N4F1S1a	Hex5 (HexNAc4)2 (DeoxyHexose1 (NeuAc1 (Man)3[GlcNAc2	Complex	59.1	14.05	12.65	13.85	13.74	12.42	12.42	12.79	12.52	14.44	10.45	10.48	10.90	14.61		13.57	13.04	11.61		0.86	0.89	0.96
n.d.	1038.89	n.d.	2078.7539	1038.87695	692.2513	0.01305	Glycan #23b	H5N4F1S1b	Hex5 (HexNAc4)2 (DeoxyHexose1 (NeuAc1 (Man)3[GlcNAc2	Complex	68.5	2.30	3.08	2.74	3.07	2.53	2.53	2.66	2.45	2.15	2.50	1.93	2.07	2.73		2.80	2.45	2.31		0.83	0.94	0.88
n.d.	1059.39	n.d.	2119.7804	1059.3902	705.92568	0.00020	Glycan #24a	H4N5F1S1a	Hex41 (HexNAc5)3 (DeoxyHexose1 (NeuAc1 (Man)3[GlcNAc2	Complex	57.6	6.21	5.87	6.51	6.01	5.87	5.87	6.41	6.27	10.99	3.89	3.89	5.51	5.21	9.42	6.15	7.38	6.01	0.98	0.81	1.20
n.d.	1059.39	n.d.	2119.7804	1059.3902	705.92568	0.00020	Glycan #24b	H4N5F1S1b	Hex41 (HexNAc5)3 (DeoxyHexose1 (NeuAc1 (Man)3[GlcNAc2	Complex	59.6	1.54	1.38	1.53	1.04	1.90	1.90	1.59	1.35	2.44	3.89	1.27	0.95	1.41		1.37	1.82	1.88		1.37	1.03	1.32
n.d.	1075.86	n.d.	2152.7907	1075.89535	716.930233	0.03535	Glycan #25	H6N5F1	Hex3 (HexNAc3)3 (DeoxyHexose1 (Man3)[GlcNAc2	Complex	63.7	0.84	0.99	1.05	0.87	1.39	1.39	0.89	0.95	0.22	0.77	1.01	0.75	0.41		0.94	0.86	0.74		0.78	0.86	0.92
n.d.	1079.9	n.d.	2160.807	1079.9035	719.602333	0.00350	Glycan #26	H3N6F1S1	Hex3 N4 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc2	Complex	58	4.40	4.68	4.22	4.62	4.62	4.62	4.47	4.45	4.00	4.19	4.42	4.24	2.09		4.48	4.38	3.73		0.83	0.85	0.98
n.d.	1111.38	n.d.	2223.7914	1111.39157	740.597133	0.01570	Glycan #27	H5N6S2	Hex5 (HexNAc4)2 (NeuAc2 (Man3)[GlcNAc2	Complex	61.1	0.90	0.88	1.01	0.96	1.26	1.26	1.16	1.10	0.62	1.13	1.34	1.28	0.92		0.94	1.04	1.17		1.24	1.12	1.11
n.d.	1132.29	n.d.	2265.8383	1132.41915	754.612767	0.02915	Glycan #28	H4N5F2S1	Hex51 (HexNAc3) (DeoxyHexose2 (NeuAc1 (Man3)[GlcNAc-U2	Complex	65.3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		N/A	N/A	N/A		N/A	N/A	N/A
n.d.	1140.4	n.d.	2281.8332	1140.4166	759.9444	0.01660	Glycan #29a	H5N5F1S1a	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc2	Complex	49	1.56	1.87	1.29	2.01	1.55	1.55	1.29	1.55	1.58	1.47	1.74	1.42	1.33		1.68	1.49	1.49		0.89	1.00	0.89
n.d.	1140.4	n.d.	2281.8332	1140.4166	759.9444	0.01660	Glycan #29b	H5N5F1S1b	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc2	Complex	54.7	0.38	0.52	0.75	0.90	0.58	0.58	0.56	0.54	0.34	0.00	0.00	0.60	0.26		0.64	0.50	0.22		0.34	0.43	0.79
n.d.	1184.41	n.d.	2369.8493	1184.42465	789.2831	0.01465	Glycan #30	H5N4F2	Hex5 (HexNAc4)2 (DeoxyHexose2 (Man3)[GlcNAc2	Complex	66.2	7.71	8.45	8.11	9.48	6.54	6.54	6.81	6.68	5.70	6.17	6.93	5.94	6.65		8.44	6.43	6.42		0.76	1.00	0.76
n.d.	1204.91	n.d.	2410.8758	1204.93179	802.9586	0.02790	Glycan #31	H4N5F2	Hex41 (HexNAc5)3 (DeoxyHexose2 (Man3)[GlcNAc2	Complex	65.1	7.02	6.73	7.51	6.86	6.98	6.98	7.57	7.43	6.50	8.00	7.21	6.94	5.79		7.03	7.12	6.98		0.99	0.98	1.01
n.d.	1221.42	n.d.	2443.8861	1221.44305	813.962033	0.02305	Glycan #32a	H6N5F1S1a	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc-U2	Complex	60.8	3.65	3.58	3.38	3.15	3.34	3.34	3.84	3.71	2.73	2.26	1.97	2.29	3.22		3.44	3.41	2.44		0.71	0.72	0.99
n.d.	1221.42	n.d.	2443.8861	1221.44305	813.962033	0.02305	Glycan #32b	H6N5F1S1b	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc-U2	Complex	68.1	1.76	1.92	1.91	1.46	1.88	1.88	1.95	2.04	1.85	1.18	1.63	1.53	2.13		1.76	1.93	1.62		0.92	0.84	1.09
n.d.	1241.92	1226.7	2484.9126	1241.95.63	827.637533	0.03630	Glycan #33	H5N6F1	Hex5 (HexNAc4)4 (DeoxyHexose1 (NeuAc1 (Man3)[GlcNAc2	Complex	59.2	0.77	0.91	1.03	0.82	1.41	1.41	1.15	1.08	1.34	0.83	0.67	1.24	3.94		0.88	1.25	1.67		1.89	1.34	1.41
n.d.	1285.92	n.d.	2572.9287	1285.96435	856.976233	0.04435	Glycan #34	H5N5F1S2	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc2 (Man3)[GlcNAc2	Complex	55.1	1.45	1.42	0.77	1.84	1.60	1.60	0.94	1.53	1.35	1.77	0.74	1.33	0.79		1.37	1.36	1.15		0.84	0.85	0.99
n.d.	1366.92	n.d.	2734.9815	1366.99075	910.993833	0.07075	Glycan #35a	H6N5F1S2a	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc22 (Man3)[GlcNAc-U2	Complex	66.4	1.03	1.09	1.08	1.39	0.92	0.92	0.78	0.87	1.09	0.46	0.82	0.91	1.22		1.15	0.91	0.76		0.67	0.84	0.80
n.d.	1366.92	n.d.	2734.9815	1366.99075	910.993833	0.07075	Glycan #35b	H6N5F1S2b	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc2 (Man3)[GlcNAc-U2	Complex	67.5	0.13	0.40	0.34	0.26	0.32	0.32	0.25	0.22	1.25	0.44	0.48	0.50	0.49		0.28	0.51	0.48		1.69	0.93	1.81
n.d.	1366.92	n.d.	2734.9815	1366.99075	910.993833	0.07075	Glycan #35c	H6N5F1S2c	Hex5 (HexNAc4)3 (DeoxyHexose1 (NeuAc2 (Man3)[GlcNAc-U2	Complex	70.1	1.77	1.74	1.61	1.49	1.02	1.02	1.12	1.47	1.94	0.91	1.07	0.86	1.79		1.65	1.39	1.16		0.70	0.83	0.84
n.d.	1512.45	n.d.	3026.0769	1512.53845	1008.02563	0.08945	Glycan #36	H6N5F3	Hex5 (HexNAc4)3 (DeoxyHexose3 (Man3)[GlcNAc-U2	Complex	71.3	0.84	1.14	1.30	1.49	0.99	0.99	1.07	0.65	0.19	0.73	0.50	0.81	0.24		1.19	0.72	0.57		0.48	0.79	0.61
n.d.	n.d.	1226.7	3682.3045	1840.65225	1226.76817	0.06817	Glycan #37a	H7N6F1S4a	Hex N4 (HexNAc4 (DeoxyHexose1 (NeuAc4)6(Man3)[GlcNAc-U2	Complex	55.9	0.53	0.59	1.07	0.32	1.16	1.16	1.22	1.59	2.24	2.10	2.17	2.17	1.87		0.63	1.55	2.08		3.33	1.34	2.48
n.d.	n.d.	1226.7	3682.3045	1840.65225	1226.76817	0.06817	Glycan #37b	H7N6F1S4b	Hex N4 (HexNAc4 (DeoxyHexose1 (NeuAc4)6(Man3)[GlcNAc-U2	Complex	74.8	1.64	2.28	2.04	1.73	1.19	1.19	1.96	1.61	1.31	1.47	2.01	1.59	1.21		1.92	1.52	1.57		0.82	1.04	0.79
n.d.	n.d.	1226.7	3682.3045	1840.65225	1226.76817	0.06817	Glycan #37c	H7N6F1S4c	Hex N4 (HexNAc4 (DeoxyHexose1 (NeuAc4)6(Man3)[GlcNAc-U2	Complex	78.2	0.83	1.44	0.89	1.12	0.59	0.59	0.52	0.92	0.90	0.42	0.48	0.69	0.78		1.07	0.73	0.59		0.56	0.81	0.69
n.d.	n.d.	1226.7	3682.3045	1840.65225	1226.76817	0.06817	Glycan #37d	H7N6F1S4d	Hex N4 (HexNAc4 (DeoxyHexose1 (NeuAc4)6(Man3)[GlcNAc-U2	Complex	79.1	0.52	0.40	0.42	0.51	0.58	0.58	0.57	0.50	0.45	0.50	0.24	0.72	0.60		0.46	0.53	0.52		1.11	0.98	1.13

Supplementary Table S2. All identified and quantified O-glycans. Observed O-glycan ions have been reduced. Mass deviation is calculated with the major observed charge state. Replicates are denoted as I, II, III ans IV. Structural elucidation was based annotation of spectral fragments using common knowledge of biosynthetic pathways. The presence of additional structural isomers cannot be excluded. n.d., not detected.

Observed ions		Theoretical m/z		Deviation	Glycan Number	Name	Composition	Probable structure	Type	Retention time (min)	Relative abundance (%) Non-transfected (N)				Mock transfection (M)				Gag transfection (S)				Relative abundance (Average %)			Fold change Ratios		
[M-H]⁻	[M-2H]²⁻	[M-H]⁻ calc	[M-2H]²⁻ calc	Δ[M-H]⁻							I	II	III	IV	I	II	III	IV	I	II	III	IV	AV.N	AV.M	AV.S	S vs N	S vs M	M vs N
675.35	n.d.	675.2466	337.1233	0.1034	Glycan #1a	H1N1S1a	(Hex)1 (HexNAc)1 (NeuAc)1		Sialylated Core 1	30.7	1.03	0.70	0.65	0.53	0.78	0.98	0.87	0.96	0.66	0.82	0.76	0.85	0.73	0.90	0.77	1.06	0.86	1.23
675.35	n.d.	675.2466	337.1233	0.1034	Glycan #1b	H1N1S1b	(Hex)1 (HexNAc)1 (NeuAc)1		Sialylated Core 1	38.7	6.61	6.07	4.99	5.21	6.61	5.72	5.62	8.73	5.34	5.53	6.18	6.15	5.72	6.67	5.80	1.01	0.87	1.17
749.42	n.d.	749.2833	374.14165	0.1367	Glycan #2	H2N2	(Hex)2 (HexNAc)2		Nonsialylated Core 2	50.2	1.63	1.02	1.01	1.18	1.46	1.41	0.80	3.96	0.86	0.94	0.98	0.96	1.21	1.91	0.94	0.77	0.49	1.57
966.48	n.d.	966.342	482.671	0.1380	Glycan #3	H1N1S2	(Hex)1 (HexNAc)1 (NeuAc)2		Sialylated Core 1	40.2	38.56	38.70	41.86	37.64	40.97	42.58	41.79	45.41	47.97	42.94	42.58	47.80	39.19	42.69	45.32	1.16	1.06	1.09
1040.5	n.d.	1040.3787	519.68935	0.1213	Glycan #4	H2N2S1a	(Hex)2 (HexNAc)2 (NeuAc)1		Sialylated Core 2	58.1	6.70	6.47	7.24	5.37	6.39	5.59	5.90	4.36	5.07	5.59	6.26	5.63	6.45	5.56	5.64	0.87	1.01	0.86
1040.5	n.d.	1040.3787	519.68935	0.1213	Glycan #5	H2N2S1b	(Hex)2 (HexNAc)2 (NeuAc)1		Sialylated Core 2	65.9	2.94	3.60	3.59	2.95	2.83	2.60	2.60	2.23	2.13	2.74	3.17	2.29	3.27	2.56	2.58	0.79	1.01	0.78
1081.43	n.d.	1081.4053	540.20265	0.0247	Glycan #6	H1N3S1	(Hex)1 (HexNAc)3 (NeuAc)1		Sialylated Core 3	54.1	0.62	0.79	0.80	0.61	0.66	0.67	0.75	1.26	1.04	0.69	0.71	0.78	0.71	0.83	0.80	1.14	0.96	1.18
1331.66	665.33	1331.4742	665.2371	0.1858	Glycan #7	H2N2S2	(Hex)2 (HexNAc)2 (NeuAc)2		Sialylated Core 2	69	34.93	34.18	32.93	39.22	33.03	33.48	35.31	26.65	31.28	34.82	33.70	30.93	35.32	32.12	32.68	0.93	1.02	0.91
n.d.	705.33	1411.431	705.2155	0.1145	Glycan #8	H2N2S2Su1	(Hex)2 (HexNAc)2 (NeuAc)2 (Sulph)1		Sialylated Core 2	60.9	6.97	8.46	6.92	7.29	7.27	6.97	6.36	6.44	5.65	5.95	5.66	4.60	7.41	6.76	5.47	0.74	0.81	0.91

REFERENCES

1. Varki, A. Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).

2. Johannssen, T. & Lepenies, B. Glycan-Based Cell Targeting To Modulate Immune Responses. *Trends in Biotechnology* **35**, 334–346 (2017).

3. Vigerust, D. J. & Shepherd, V. L. Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol.* **15**, 211–218 (2007).

4. Yehuda, S. & Padler-Karavani, V. Glycosylated Biotherapeutics: Immunological Effects of N-Glycolylneuraminic Acid. *Frontiers in Immunology* **11**, (2020).

5. Duivelshof, B. L. *et al.* Glycosylation of biosimilars: Recent advances in analytical characterization and clinical implications. *Anal. Chim. Acta* **1089**, 1–18 (2019).

6. Gupta, S. K. & Shukla, P. Glycosylation control technologies for recombinant therapeutic proteins. *Applied Microbiology and Biotechnology* **102**, 10457–10468 (2018).

7. Mastrangeli, R., Palinsky, W. & Bierau, H. Glycoengineered antibodies: Towards the next-generation of immunotherapeutics. *Glycobiology* **29**, 199–210 (2018).

8. Lin, S. C. *et al.* Different Immunity Elicited by Recombinant H5N1 Hemagglutinin Proteins Containing Pauci-Mannose, High-Mannose, or Complex Type N-Glycans. *PLoS One* **8**, (2013).

9. Goldblatt, D. Conjugate vaccines. *Clinical and Experimental Immunology* **119**, 1–3 (2000).

10. Schön, K., Lepenies, B. & Goyette-Desjardins, G. Impact of Protein Glycosylation on the Design of Viral Vaccines. in *Advances in biochemical engineering/biotechnology* (Adv Biochem Eng Biotechnol, 2020).

11. Watanabe, Y., Bowden, T. A., Wilson, I. A. & Crispin, M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica et Biophysica Acta - General Subjects* **1863**, 1480–1497 (2019).

12. Roldão, A., Mellado, M. C., Castilho, L. R., Carrondo, M. J. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev Vaccines* **9**, 1149–1176 (2010).

13. Cervera, L. *et al.* Production of HIV-1-based virus-like particles for vaccination: achievements and limits. *Appl. Microbiol. Biotechnol.* **103**, 7367–7384 (2019).

14. Lancaster, C. *et al.* Characterization of N-glycosylation profiles from mammalian and insect cell derived chikungunya VLP. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1032**, 218–223 (2016).

15. Lepenies, B., Lee, J. & Sonkaria, S. Targeting C-type lectin receptors with multivalent carbohydrate ligands. *Advanced Drug Delivery Reviews* **65**, 1271–1281 (2013).

16. Tretter, V., Altmann, F., Kubelka, V., März, L. & Becker, W. M. Fucose **α**1,3-linked to the core region of glycoprotein n-glycans creates an important epitope for ige from honeybee venom allergic individuals. *Int. Arch. Allergy Immunol.* **102**, 259–266 (1993).

17. Md, A. *et al.* Purification and molecular characterization of a truncated-type Ara h 1, a major peanut allergen: oligomer structure, antigenicity, and glycoform. *Glycoconj. J.* **38**, (2021).

18. Dhar, C., Sasmal, A. & Varki, A. From 'Serum Sickness' to 'Xenosialitis': Past, Present, and Future Significance of the Non-human Sialic Acid Neu5Gc. *Frontiers in immunology* **10**, 807 (2019).

19. Ghaderi, D., Taylor, R. E., Padler-Karavani, V., Diaz, S. & Varki, A. Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat. Biotechnol.* **28**, 863–867 (2010).

20. Colgrave, M. L. *et al.* Site occupancy and glycan compositional analysis of two soluble recombinant forms of the attachment glycoprotein of Hendra virus. *Glycobiology* **22**, 572–584 (2012).

21. Ausubel, L. J. *et al.* Production of CGMP-Grade Lentiviral Vectors. *Bioprocess Int* **10**, 32–43 (2012).

22. Lavado-García, J., Cervera, L. & Gòdia, F. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures. *Front. Bioeng. Biotechnol.* **8**, 617 (2020).

23. Effio, C. L. & Hubbuch, J. Next generation vaccines and vectors: Designing downstream processes for recombinant protein-based virus-like particles. *Biotechnol J* **10**, 715–727 (2015).

24. Seifalian, A., De La Peña, H. & Seifalian, A. M. The application of exosomes as a nanoscale cancer vaccine. *Int. J. Nanomedicine* **5**, 889 (2010).

25. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. Structural analysis of N- and O-glycans

released from glycoproteins. *Nat. Protoc.* **7**, 1299–1310 (2012).

26.   Stavenhagen, K., Kolarich, D. & Wuhrer, M. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography–Mass Spectrometry. *Chromatographia* **78**, 307–320 (2014).

27.   Zhang, T. *et al.* Development of a 96-well plate sample preparation method for integrated: N - And O -glycomics using porous graphitized carbon liquid chromatography-mass spectrometry. *Mol. Omi.* **16**, 355–363 (2020).

28.   Thomsson, K. A., Bäckström, M., Holmén Larsson, J. M., Hansson, G. C. & Karlsson, H. Enhanced detection of sialylated and sulfated glycans with negative Ion mode nanoliquid chromatography/mass spectrometry at high pH. *Anal. Chem.* **82**, 1470–1477 (2010).

29.   Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J Virol Methods* **193**, 85–95 (2013).

30.   Holst, S., Belo, A. I., Giovannetti, E., Van Die, I. & Wuhrer, M. Profiling of different pancreatic cancer cells used as models for metastatic behaviour shows large variation in their N-glycosylation. *Sci. Rep.* **7**, (2017).

31.   Karlsson, N. G., Schulz, B. L. & Packer, N. H. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J. Am. Soc. Mass Spectrom.* **15**, 659–672 (2004).

32.   Karlsson, N. G. *et al.* Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun. Mass Spectrom.* **18**, 2282–2292 (2004).

33.   Ceroni, A. *et al.* GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7**, 1650–1659 (2008).

34.   Cooper, C. A., Gasteiger, E. & Packer, N. H. GlycoMod - A software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* **1**, 340–349 (2001).

35.   Yan, S. B., Chao, Y. B. & Van Halbeek, H. Novel Asn-linked oligosaccharides terminating in GalNAc**β**(1→4)[fuc**α**(1→3)]GlcNAc**β**(1→•) are present in recombinant human protein C

expressed in human kidney 293 cells. *Glycobiology* **3**, 597–608 (1993).

36. Chen, Q., Tan, Z., Guan, F. & Ren, Y. The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology. *Frontiers in Chemistry* **8**, 511 (2020).

37. Alessandri, L. *et al.* Increased serum clearance of oligomannose species present on a human IgG1 molecule. *MAbs* **4**, 509–520 (2012).

38. Goetze, A. M. *et al.* High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* **21**, 949–959 (2011).

39. Meri, S., Lehto, T., Sutton, C. W., Tyynelä, J. & Baumann, M. Structural composition and functional characterization of soluble CD59: Heterogeneity of the oligosaccharide and glycophosphoinositol (GPI) anchor revealed by laser-desorption mass spectrometric analysis. *Biochem. J.* **316**, 923–935 (1996).

40. Rudd, P. M. *et al.* The glycosylation of the complement regulatory protein, human erythrocyte CD59. *J. Biol. Chem.* **272**, 7229–7244 (1997).

41. Amet, T. *et al.* Glycosylphosphatidylinositol Anchor Deficiency Attenuates the Production of Infectious HIV-1 and Renders Virions Sensitive to Complement Attack. *AIDS Res. Hum. Retroviruses* **32**, 1100–1112 (2016).

42. Jin, C., Hantusch, B., Hemmer, W., Stadlmann, J. & Altmann, F. Affinity of IgE and IgG against cross-reactive carbohydrate determinants on plant and insect glycoproteins. *J. Allergy Clin. Immunol.* **121**, (2008).

43. Wagner, J. M. *et al.* Enhanced Production of Chikungunya Virus-Like Particles Using a High-pH Adapted Spodoptera frugiperda Insect Cell Line. *PLoS One* **9**, e94401 (2014).

44. Beniac, D. R. & Timothy, B. F. Structure of the Ebola virus glycoprotein spike within the virion envelope at 11 Å resolution. *Sci. Rep.* **7**, (2017).

45. An, Y., McCullers, J. A., Alymova, I., Parsons, L. M. & Cipollo, J. F. Glycosylation analysis of engineered H3N2 influenza A virus hemagglutinins with sequentially added historically relevant glycosylation sites. *J. Proteome Res.* **14**, 3957–3969 (2015).

46. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).

47.   Lee, J. H., Ozorowski, G. & Ward, A. B. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science (80-.).* **351**, 1043–1048 (2016).

48.   Seabright, G. E., Doores, K. J., Burton, D. R. & Crispin, M. Protein and Glycan Mimicry in HIV Vaccine Design. *Journal of Molecular Biology* **431**, 2223–2247 (2019).

49.   Casalino, L. *et al.* Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **6**, 1722–1734 (2020).

50.   Grant, O. C., Montgomery, D., Ito, K. & Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **10**, (2020).

51.   Jobe, O., Kim, J. & Rao, M. The role of Siglec-1 in HIV-1/macrophage interaction. *Macrophage* **4**, (2018).

52.   Lavado-García, J. *et al.* Molecular Characterization of the Coproduced Extracellular Vesicles in HEK293 during Virus-Like Particle Production. *J. Proteome Res.* **19**, 4532 (2020).

53.   Lavado-García, J. *et al.* Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles. *Biotechnol. Bioeng.* (2021). doi:10.1002/bit.27679

54.   Reiter, K. *et al.* Separation of virus-like particles and extracellular vesicles by flow-through and heparin affinity chromatography. *J. Chromatogr. A* **1588**, 77–84 (2019).

55.   Pereira Aguilar, P. *et al.* Capture and purification of Human Immunodeficiency Virus-1 virus-like particles: Convective media vs porous beads. *J. Chromatogr. A* **1627**, (2020).

56.   Unger, W. W. J. *et al.* Glycan-modified liposomes boost CD4 + and CD8 + T-cell responses by targeting DC-SIGN on dendritic cells. *J. Control. Release* **160**, 88–95 (2012).

57.   Yang, L. *et al.* Engineered lentivector targeting of dendritic cells for in vivo immunization. *Nat. Biotechnol.* **26**, 326–334 (2008).

58.   Miwa, H. E., Song, Y., Alvarez, R., Cummings, R. D. & Stanley, P. The bisecting GlcNAc in cell growth control and tumor progression. *Glycoconj. J.* **29**, 609–618 (2012).

59.   Link-Lenczowski, P. *et al.* The glycomic effect of N-acetylglucosaminyltransferase III overexpression in metastatic melanoma cells. GnT-III modifies highly branched N-glycans.

Glycoconj. J. **35**, 217–231 (2018).

60. Nakano, M. *et al.* Bisecting GlcNAc is a general suppressor of terminal modification of N-glycan. *Mol. Cell. Proteomics* **18**, 2044–2057 (2019).

61. Bonomelli, C. *et al.* The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. *PLoS One* **6**, (2011).

62. Doores, K. J. *et al.* Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13800–13805 (2010).

63. Raska, M. *et al.* Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. *J. Biol. Chem.* **285**, 20860–20869 (2010).

64. Jin, Z. C. *et al.* Genetic disruption of multiple 1,2-mannosidases generates mammalian cells producing recombinant proteins with high-mannose–type N-glycans. *J. Biol. Chem.* **293**, 5572–5584 (2018).

65. Abdel-Motal, U. M., Guay, H. M., Wigglesworth, K., Welsh, R. M. & Galili, U. Immunogenicity of Influenza Virus Vaccine Is Increased by Anti-Gal-Mediated Targeting to Antigen-Presenting Cells. *J. Virol.* **81**, 9131–9141 (2007).

66. Abdel-Motal, U. M. *et al.* Increased immunogenicity of HIV-1 p24 and gp120 following immunization with gp120/p24 fusion protein vaccine expressing **α**-gal epitopes. *Vaccine* **28**, 1758–1765 (2010).

67. Galili, U. Amplifying immunogenicity of prospective Covid-19 vaccines by glycoengineering the coronavirus glycan-shield to present **α**-gal epitopes. *Vaccine* **38**, 6487–6499 (2020).

DISCUSSION

1. **Concluding remarks**

   Overall analysis of the context in which this work needs to be considered. Discussion of the current state of application of the technology used in this work in the industry of vaccine development.

2. **Future work**

   New lines of research opened from the results of this work are analysed in this section. The results and conclusions of each chapter are extended and new approaches to continue the work are proposed.

The work presented in this PhD thesis represents new steps in VLP-based vaccine innovation bridging research and development with the industrial production of vaccines. Always bearing in mind the objective of mass production, large-scale implementation and distribution. The concepts of scalability, cost-effectiveness, easiness of implementation and applicability range are the driving force and tightly connected to every step taken when moving forward throughout this work. In order to understand how these chapters fit in with the current knowledge and how they may help towards the improvement of vaccine technology, an overall analysis of the industrial context in which this work is conceived needs to be considered. The following concluding remarks and the proposed future work are thought to complement this PhD thesis and present tools for the further applications of the results obtained in the presented chapters.

1. Concluding Remarks

Integrated continuous bioprocessing at industrial scale for the production of biopharmaceuticals has regained focus in the last years due to the advance and innovation of cell retention devices. Being able to more efficiently recirculate cells to the bioreactor, reducing technical failure rates and incorporating single-use technology have encourage biopharmaceutical companies to re-evaluate continuous bioprocess technologies[1]. Moreover, in order to holistically evaluate the production process and its costs downstream processing also has to be assessed, to fully understand the advantages of an integrated continuous biomanufacturing processes.

The current preferred mode of operation for production of biopharmaceuticals like VLPs at large scale is fed-batch. This is the operation mode of choice due to its ease of operation, robustness, lower failure rates and possibilities of achieving high titers with feed optimization[2,3]. For downstream processing, the current purification steps are based on semicontinuous chromatography, relying on new technologies for continuous capture and further batch polishing[4]. However, there are many factors to consider as large scale biomanufacturing cannot be compared to the limitations and requirements of the establishment of a small-size company for commercial purposes. The implementation of a continuous mode of operation

based on perfusion for process intensification not only contributes to increase titers, but also provide intrinsic game-changer aspects that go beyond economic benefits. The environmental footprint left by a continuous biomanufacturing process is lower than conventional fed-batch[5]. The intensification provided by perfusion lead to a reduction in the size of the facility, with the subsequent reduction in energy consumption. Also, it reduces the number of cycles and the accessory processes like cleaning and sterilization in place and preparation that mandatorily are required for each cycle. The reduction in culture media consumption and derived production costs to achieve a given amount of final product (**Table 1**), also affect the environmental aspect. One of the problems for the implementation of continuous technology was the high failure rates and the non-optimal technology regarding cell retention and recirculation. The advances in tangential flow filtration and the establishment of ATF and TFF for upstream and downstream processing have allowed and encouraged the exploration of continuous technologies [6]. Apart from economic and environmental, operational aspects such as robustness and easiness of control implementation also need to be considered. Weighting up all these factors, the combination of a continuous upstream process using ATF coupled with continuous capture, such as simulated continuous moving bed and batch polishing (ATF-CB) have proven the optimal bioprocess sequence structure for small-size companies while fed-batch combined with continuous capture and batch polishing (FB-CB) is the optimal for mass-production large-scale companies where operational feasibility and large scale technology readiness impose limitations on the establishment of continuous biomanufacturing[6]. Another element to increase flexibility and reduce costs and facility footprint together with the implementation of continuous manufacturing is the single-use (SU) technology. The biopharmaceutical industry is moving towards production on demand, increasing the number of different products one facility needs to produce. To adapt to the new changing situation where personalized medicine and multi-product facilities need to find effective approaches for best biomanufacturing practices, upstream and downstream SU technology can provide a solution and offer flexibility adding or removing unit operations on demand without wasting capacity or resources at a affordable price for the final consumer[7,8].

The work presented in this PhD thesis is delimited within a scope of work defined for the development of a bioprocess at laboratory scale, also applicable for a small-size company. Two main strategies have been

Table I. Bioprocess comparison of overall performance

	Batch	**Perfusion**
Accumulated VLPs	1.60E+13	8.70E+13
VLP specific productiviy (VLP·cell^{-1}·day^{-1})	8.90E+02	3.00E+03
Volumetric productivity (VLP·L^{-1}·day^{-1})	2.70E+12	7.10E+12
Total volume of spent media (L)	1	3.8
Spent media/produced VLP (pL·VLP^{-1})	0.063	0.044

explored (**Table 1**), presenting the proposed perfusion process using SU XCell ATF2 HMF a substantial improvement in total cumulative produced VLPs, specific and volumetric productivity. This process presented the drawback of retaining VLPs within the bioreactor and not allowing a continuous harvesting alternative. Therefore, further research must be undertaken following the perfusion-based bioprocess developed in **Chapter three Part II** and regarding it as the starting point for future optimization, as discussed later. This represents the first step in the design of a complete integrated HIV-1 Gag VLP production bioprocess. Moreover, in order to acquire the advantages that a perfusion mode provides, the bioprocess must aim towards the achievement of high cell densities and maintain it over days or weeks. Focusing on the methodology of VLP production, in **Chapter three** the optimization of the EGE protocol has proven that the CDE imposes a limitation when a bioprocess based on TGE is implemented using a perfusion approach. High cell densities cannot be achieved before transfection as high cell density cultures cannot be transfected (as observed in **Chapter one and three**). In addition, if high cell densities are achieved after transfection, the percentage of transfected cells rapidly decreases, losing the plasmid of interest, thus hindering production. Therefore, in order to fully exploit the benefits of perfusion, a cell line stably expressing Gag polyprotein needs to be used. This stable cell line could be used in a perfusion bioprocess using the SU VHU2 module which allowed up to 48% of VLP filtration and recovery in the harvest fraction, and expected to be higher for further hpts (**Chapter three Part III**). Moreover, further research, testing and optimization must be undertaken regarding the use of the VHU2 module to work towards the achievement of a higher filtration percentage. The implementation of a system allowing a continuous VLP recovery from the clarified harvest is the key for the establishment of a future ATF-CB VLP

production platform. The VLP-containing clarified harvest can be directly concentrated and purified using continuous capture, eliminating the clarification step and reducing the consequent loss of VLPs in the sequential purification steps, increasing the final yield. The key element allowing the VLPs to be continuously recovered is the cell retention device. Currently, to the best of our knowledge, no commercial cell retention device in the market allows total VLP recovery in the membrane permeate. Innovation in the development of new materials suitable for cell retention and VLP filtration is essential in the future establishment of a complete integrated VLP production bioprocess.

Innovation in metabolic engineering is also crucial to intensify the process and reduce future facility size and environmental footprint. In this work, cellular processes like transient transfection and VLP budding have been studied and characterized at molecular level. The knowledge regarding cellular metabolism and pathways involved VLP production that has been gained in the development of this thesis has allowed the increase of VLP specific productivity 3-fold from ~$3 \cdot 10^3$ to ~$9 \cdot 10^3$ VLP·cell^{-1}·day^{-1} using specific modulations like knocking down CNP. It has also contributed to improve the standard transient transfection efficiency, reaching ~90% of transfected cells when overexpressing UGCG protein. Moreover, the molecular characterization, including the glycosylation signature, of produced VLPs and coproduced EVs has also contributed to gain knowledge in order to design better strategies for membrane functionalization, downstream separation processes or even adjuvant-related applications to optimize the presence of these coproduced EVs.

The combination of metabolic intensification, smart antigen design, continuous manufacturing and SU elements could pave the way towards improvements in vaccine development. It could enable the implementation of smaller, more sustainable, versatile and flexible facilities, being capable of producing a wide range of bioproducts and vaccines. The future of vaccine technology is unquestionably tied to the improvement in adaptability of current and new facilities. Therefore, these advantages have become fundamental and indispensable to address situations like new viral outbreaks.

2. Future Work

Aa a characteristic feature of the scientific work, many new topics have arisen from the development of this thesis and several new lines of research have been unwound. The next steps following this work represent possible starting points for whole new projects to take vaccine development further into the future.

The proteomic study performed in **Chapter one** resulted in 4403 proteins identified and quantified with more than 1 peptide that presented changes upon transfection and VLP production. In **Chapter two**, 6 proteins (NEDD8, NEDD4L, NMT-1, UGCG, CIT and CNP) were studied in depth to engineer the HEK293 production system. This represents 0.14% of all the proteins presenting up or downregulations. The potential of the information generated in this protein database is tremendous. This resource may offer support for many future experiments concerning the modulation of the HEK293 cell-based production system. It can either serve as a database to check the behaviour of involved proteins and biological processes in future experiments or serve as a starting point to continue analysing different unexplored molecular and cellular pathways. Not only to intensify and optimize VLP production, but also for different purposes concerning the metabolic network of HEK293 cells. One example of this is the Hrs protein, belonging to ESCRT-0 complex. Although this ESCRT complex did not show significant changes upon transfection and VLP production, as reported in **Chapter one**, it has been described to interact with CIT protein (**Chapter two**), reported to significantly act on transporting Gag to the cell membrane[9]. The pathway of Hrs-CIT interaction is a potential target, yet unexplored, for the improvement of the HEK293 system for VLP production. Proteins with reported antiviral activity are another target of study. There are different proteins whose effect on Gag are analogous to the one exerted by CNP. Proteins like tetherin or MOV10 have shown to interact with Gag polyprotein preventing their oligomerization and subsequent budding into VLPs[10,11]. Strategies combining the knocking down of these proteins could substantially improve the HEK293 system. An approach combining in one single plasmid different shRNAs against CNP, tetherin and/or MOV10 could maximize production efficiency for any lentiviral production platform. Moreover, taking a step further, the combination of transfecting this enhancer plasmid to a cell line stably expressing Gag polyprotein could extensively increase VLP production as the maximum genetic dose could

be reached only with the enhancer plasmid. Another example of how the up and downregulation protein database could be exploited is the phosphodiesterase family. Previous research have demonstrated that adding caffeine to HEK293 cell cultures increase VLP production[12]. This effect can be explained by the interaction of caffeine and the family of phosphodiesterase proteins. Caffeine has been reported to be an inhibitor of phosphodiesterases and moreover, to act on HIV-1 VLP production through this pathway[13,14]. However, the phosphodiesterase family comprises many proteins, each of them acting on specific tissues and under different expression conditions. The identification of the correct phosphodiesterase target is crucial for the development of the correct strategy and a complete expression database of the phosphodiesterases identified in the HEK293 system represents the tool to achieve it. The most abundant phosphodiesterase identified and quantified resulting from the work of **Chapter one** was the retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit delta (PDEδ). Targeting this protein can contribute to improving VLP production. These examples are only a few cases showing the potential of such expression database for the further improvement of the HEK293 system.

Likewise, in **Chapter one**, another clear limiting factor was identified, the cell density effect (CDE). The molecular characterization of the cellular factors influencing the CDE was only the first step towards the fully development of a more in-depth study on how to improve transient transfection at high cell densities based on the insight gained in this work. The first step on the study of molecular modulations towards this aim would be to test the most relevant identified genes influencing the transport of plasmid DNA to the nucleus. To undertake this strategy, genes like IPO4, IPO7, IPO11, IPO9, KPNA1, KPNB1… identified in **Chapter one**, can be overexpressed and their effects on plasmid transport tested. The correct cell membrane composition is also crucial for the incorporation of the plasmid from the extracellular medium. Therefore, modulating metabolic enzymes, like HMGCS-1, involved in lipid biosynthesis could affect the uptake of DNA:PEI complexes. Transient gene expression (TGE) can be used to test all different overexpressions and analyse the cellular elements governing this effect. However, there are other factors concerning culture media that can be affecting transfection efficiency. DNA:PEI complexes are negatively charged and interact with positively charged membranes. In **Chapter two, three, five and six** it has been identified that extracellular vesicles (EVs) are constitutively being produced together with VLPs,

representing 85% of the total number of particles in a standard production (**Chapter two**). This high concentration of cell membrane-bound vesicles may be interacting with the complexes and preventing transfection at high cell densities, where the concentration of EVs is substantially higher. A study analysing EV content versus transfection efficiency may offer a new insight on the CDE. Combining the study of spent media and correlating the EV content in different transfection tests may help cast some light on the troublesome CDE. Nevertheless, to fully study the far-reaching implications of the CDE, model cell lines should be stablished in order to decipher the molecular mechanism. The most promising candidate to construct a new model cell line is UGCG gene. This gene has been studied in **Chapter two**, where its overexpression led to an increase of transfection efficiency up to 17%. Creating a HEK293 cell line stably overexpressing UGCG could offer the possibility to study at high cell densities the observed effects of UGCG. Understanding how a basal improvement of transfection would behave at high cell densities could provide useful insight on how transfection is prevented at this cell densities. Together, combining the analysis on EV content and spent media and the UGCG model represent great tools for the in-depth study of the CDE. On the other hand, to continue studying the strategy of TGE and in order to further optimize it, the UGCG modulation could also provide new advantages compared to the traditional method of transfection. The UGCG-expressing stable cell line could be used to adjust the DNA:PEI ratio, maximizing transfection in the case that the effect of UGCG enzyme potentiated different routes for the complexes to enter the cell. All these different strategies based on metabolic modulations would expand our knowledge of the cellular production and budding of Gag-based VLPs, allowing to maximize the molecular potential of the HEK293 system.

Along with these aspects of VLP production optimization, the operation conditions at bioreactor level can also be further optimized to fully intensify the bioprocess. As observed in **Chapter two**, using the strategy of shRNA against CNP increased VLP specific productivity, budding efficiency and the percentage of VLPs compared to the total number of particles produced. The combination of this CNP shRNA and perfusion mode of operation could further intensify the system. Moreover, as observed in **Chapter three**, the limitation of the proposed perfusion strategy was the CDE, as the culture could no longer be transfected when cell density increased. This limitation can be overcome using stable cell lines and intensifying the

subsequent perfusion process. Moreover, using a stable cell line for the production of Gag, the gene dose limiting the co-transfection of modulators can also be overcome, being able to use CNP shRNA and UGCG modulators to maximize VLP budding efficiency (**Chapter two**). In addition to this, the perfusion process using the VHU2 module for continuous harvest must be further optimized to increase the filtration efficiency and VLP recovery in the harvest fraction. The implementation of a perfusion process using a stable cell line and the VHU2 filter module could allow the establishment of a conventional perfusion process, reaching high cell densities coupled with continuous harvest.

A different complementary study, necessary to undertake further steps in the functionalization of the VLPs, is the characterization of their content. As observed in **Chapter four**, VLPs produced in the HEK293 system are filled with material whose biochemical composition remains undefined. Currently, there is an extensive research in the field of defining the biochemical features and content of EVs. As for the protein content of EVs and VLPs, the proteomic study in **Chapter five** allowed the identification, with more than one peptide, of 256 proteins included in Vesiclepedia database. Most of these proteins are enzymes reflecting changes in metabolic homeostasis in the HEK293 system. However, histones were upregulated in VLPs, indicating that compared to EVs, VLP content was enriched in histones. In order to investigate whether histones are secreted individually or forming nucleosomes, DNA sequencing techniques could be applied to purified VLP samples. The use of microarrays or immunoprecipitation assays using antibodies against H2A and H3 could provide new information towards the characterization of a more complete biochemical description of VLP content. Lipids are also key to understand how VLP membranes differ from the rest of EVs. The level of enrichment in sphingomyelin, phosphatidylserine, cholesterol, fatty acids, ceramide and derivatives could be assessed by gas chromatography, or a metabolomic study. This is particularly interesting regarding lipids rafts. Lipid rafts micro-domains in the cell membrane are described to promote spots for VLP budding as well as harbouring tetraspanins and glycoproteins for further functionalization[15–17]. Identifying the differences between EV and VLP membrane biochemical composition could pave the way for better functionalization. Another relevant biochemical composition trait is RNA. Gag polyprotein requires RNA for the correct assembly and budding, a natural biological function used for retroviruses to pack their genetic material into the new virions[18–20]. Even more interesting,

Gag selectively binds to specific RNA sequences which are the ones later packed into the VLPs[21,22]. A clear line of work can be derived from this as nucleic acids were indeed reported as VLP material content in **Chapter five**. Sequencing the RNAs that are being incorporated into VLPs and identifying the packaging sequences that are being used would provide the required information to proceed to the next and more interesting step: designing RNAs of interest to be incorporated into VLPs for target delivery. RNA-based biopharmaceuticals are the future of personalized medicine and vaccine development. A new generation of vaccines has taken off in December 2020 with SARS-CoV-2 vaccine. A rapid, cost-effective and highly efficacious vaccine has opened the door for RNA-based technologies to conquer the biopharmaceutical market. A clear line of research on designing RNAs to use Gag-based VLPs as nanocarriers is ought to be explored.

Closely related to VLP biochemical composition, a different future study can be extracted from **Chapter two and five**. VLP intracellular budding takes place in multivesicular bodies (MVBs) and part of their membrane will serve to create the secreted VLPs. Membrane composition of MVBs within the cell is key to define the pathway that their intralumenal VLPs will follow[23]. Apart from the already studied ESCRT pathway, there are ESCRT-independent unexplored pathways that could be exploited to increase VLP production. It was demonstrated that the synthesis of ceramide by sphingomyelinases nSMase, hydrolysing sphingomyelin to ceramide, triggered the generation of intralumenal vesicles (ILVs) independently from ESCRT mechanisms[24]. After characterizing the lipid composition of VLP membrane, this ceramide-dependent pathway can be explored by metabolic engineering overexpressing the neutral sphingomyelinase nSMase. Another ESCRT-independent and sphingomyelinase-independent pathway requires the tetraspanin CD63 for ILV biogenesis via the MVB pathway[25]. Modulating the overexpression of CD63 can also be a target to increase VLP production and budding. The ESCRT-independent pathways for exosome production have also been linked to NEDD4L, since the protein complex SIMPLE, a regulator of ESCRT-independent exosome biogenesis requires the interaction with accessory proteins like NEDD4L[26–28]. This ties to **Chapter two** where NEDD4L overexpression was shown to induce an increase in VLP production and budding efficiency, through regular cell membrane and intracellular budding.

Apart from optimizing membrane composition to trigger different biogenesis pathways for VLPs, membrane proteins harbour more potential for future research on VLP functionalization and purification. The VLP glycosylation study presented in **Chapter six** provided information about enrichment of CD59 glycosylations in VLPs. Apart from using PGC–LC–MS/MS for the analysis of complex glycan mixtures it is also applied as a tool for the in-depth characterization of specific glycoproteins. CD59 can be characterized in more detail performing a glycoproteomic study[29]. The purification of CD59 glycoprotein can be achieved by 2D SDS-PAGE electrophoresis and MALDI-ToF-MS identification. The use of specific glycosidases could provide information about the specific branching of the glycoforms displayed in CD59. Moreover, in order to specifically map the glycosylation sites and motifs, CD59 glycoprotein could undergo a trypsin treatment and further HILIC-SPE separation to characterize intact glycopeptides. After characterizing and quantifying CD59 glycopeptides in EVs and VLPs, a better understanding of the differences found would lead to a better design of a separation and purification strategy. This type of glycoproteomic study could also be applied in parallel to the previously glycoengineering approaches proposed in **Chapter six**. In order to design an efficient separation method for EVs and VLPs, CD59 glycosylations are promising candidates, as a prominent differential feature, for chromatographic affinity resins to be based on. For the development of chromatographic columns, the correct ligand or resin must be previously tested and their proper interaction with the target glycosylations must be validated. This type of interaction can be assessed by surface plasmon resonance (SPR)[30]. This kinetics and affinity analysis technique can be used to screen for different ligands and study their interaction with VLPs. The SPR biosensor (**Figure 1**) is typically composed of a metal surface where the ligand biomolecules to be tested are immobilized, tethered to the metal surface. Then, the analyte of interest, in this case, VLPs, are flowed across the functionalized metal surface. The interaction between VLPs and the tested ligand would disrupt the conducting electrons resonant oscillations (plasmon wave) at the interface of these two conducting phases. This adsorption of molecules to the conducting metal surface is enough to disturb the highly sensitive oscillations of the plasmon wave. In order to excite surface plasmon waves in a resonant manner, an incident light beam can be used, channelled through a glass prism to achieve resonance at a given wavelength and angle. This technique is widely used to study binding kinetics and affinity on a wide range of biosensor applications[31,32]. Another
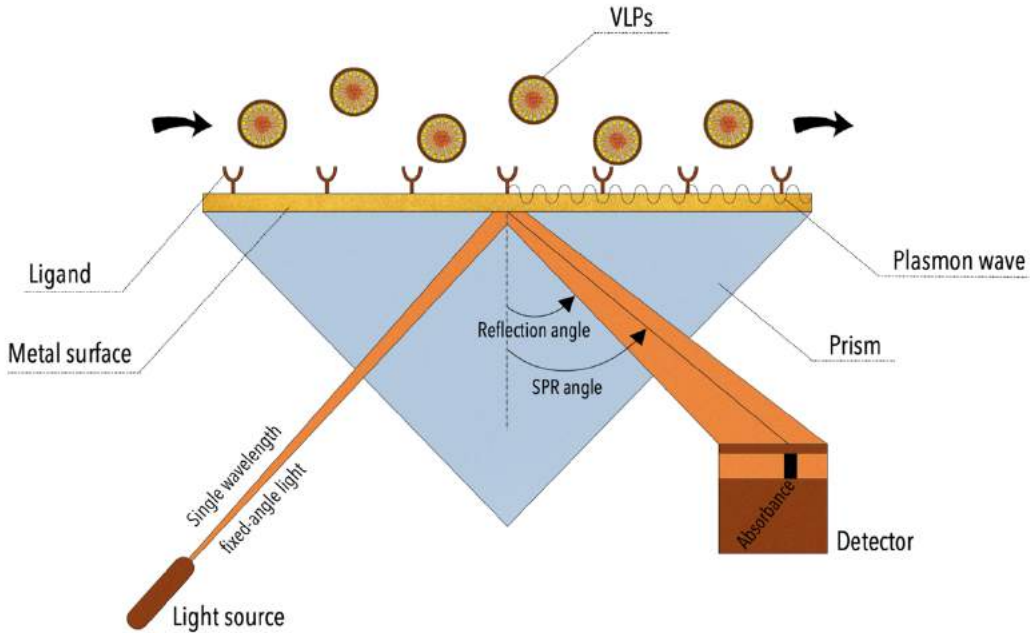
Figure 1. Schematic representation of a surface plasmon resonance (SPR) biosensor. A binding compound is immobilized on a metal surface (typically gold) and an analyte of interest is flowed across it. The change in the refractive index of the solution at the interface is monitored by a detector. Binding kinetics can be studied using surface plasmon resonance.

interesting conclusion from **Chapter four, five and six** is that upon transfection and VLP production, there is an increase in the biogenesis of smaller vesicles compared to the basal constitutive vesicle production in a non-transfected cell culture. The increase in the production rate of smaller vesicles implies reducing the number of membrane proteins in each particle as the cell has been depleted of surface membrane at a higher rate and membrane proteins and glycoproteins might not be replenished at a sufficient rate. The implications of this observed phenomenon establish a limitation for the production of Gag VLP if functionalization wants to be maximized via transient transfection. If VLP production is maximized, the number of membrane proteins to be functionalized is reduced and vice versa. To surpass this limitation in order to maximize VLP production and functionalization, one of the two modifications must be stably expressed. The use of stable Gag VLP producer HEK293 cell lines is required to produce functionalized VLPs transiently transfecting for the overexpression of the synthesis of the membrane protein of interest. Otherwise, when attempting to transiently transfect genes encoding both proteins, the gene dose and PEI limitations will hinder the possibility of overcoming the previously mentioned balance.

Having a stable Gag VLP producer cell line engineered with membrane protein modifications would also allow the future study of self-adjuvanted VLPs, including immunogenic antigens and glycoforms expressed in the VLP surface. Moreover, the presence of membrane proteins may also be used for click-chemistry approaches for functionalization. The research on maximizing membrane protein abundance in the VLP surface is one of the most promising next steps to be undertaken.

Bringing together all of these proposed studies would certainly drive and inspire the future of the design and engineering of Gag VLPs.

REFERENCES

1. Warikoo, V. *et al.* Integrated continuous production of recombinant therapeutic proteins. *Biotechnol. Bioeng.* **109**, 3018–3029 (2012).

2. Kelley, B. Very large scale monoclonal antibody purification: The case for conventional unit operations. *Biotechnology Progress* **23**, 995–1008 (2007).

3. Farid, S. S. Established bioprocesses for producing antibodies as a basis for future planning. *Advances in Biochemical Engineering/Biotechnology* **101**, 1–42 (2006).

4. Marichal-Gallardo, P. A. & Álvarez, M. M. State-of-the-art in downstream processing of monoclonal antibodies: Process trends in design and validation. *Biotechnol. Prog.* **28**, 899–916 (2012).

5. Pollock, J., Ho, S. V. & Farid, S. S. Fed-batch and perfusion culture processes: Economic, environmental, and operational feasibility under uncertainty. *Biotechnol. Bioeng.* **110**, 206–219 (2013).

6. Pollock, J., Coffman, J., Ho, S. V. & Farid, S. S. Integrated continuous bioprocessing: Economic, operational, and environmental feasibility for clinical and commercial antibody manufacture. *Biotechnol. Prog.* **33**, 854–866 (2017).

7. Jacquemart, R. *et al.* A Single-use Strategy to Enable Manufacturing of Affordable Biologics. *Comput. Struct. Biotechnol. J.* **14**, 309–318 (2016).

8. Junne, S. & Neubauer, P. How scalable and suitable are single-use bioreactors? *Curr. Opin. Biotechnol.* **53**, 240–247 (2018).

9. Ding, J., Su, L. & Gao, G. Hrs inhibits citron kinase-mediated HIV-1 budding via its FYVE domain. *Protein Cell* **2**, 470–476 (2011).

10. Liberatore, R. A., Mastrocola, E. J., Powell, C. & Bieniasz, P. D. Tetherin Inhibits Cell-Free Virus Dissemination and Retards Murine Leukemia Virus Pathogenesis. *J. Virol.* **91**, (2017).

11. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).

12. Cervera, L. *et al.* Selection and optimization of transfection enhancer additives for increased

virus-like particle production in HEK293 suspension cell cultures. *Appl Microbiol Biotechnol* **99**, 9935–9949 (2015).

13. Choi, O. H., Shamim, M. T., Padgett, W. L. & Daly, J. W. Caffeine and theophylline analogues: Correlation of behavioral effects with activity as adenosine receptor antagonists and as phosphodiesterase inhibitors. *Life Sci.* **43**, 387–398 (1988).

14. Wilson, S. J. *et al.* Inhibition of HIV-1 particle assembly by 2′,3′-cyclic- nucleotide 3′-Phosphodiesterase. *Cell Host Microbe* **12**, 585–597 (2012).

15. Hogue, I. B., Grover, J. R., Soheilian, F., Nagashima, K. & Ono, A. Gag Induces the Coalescence of Clustered Lipid Rafts and Tetraspanin-Enriched Microdomains at HIV-1 Assembly Sites on the Plasma Membrane. *J. Virol.* **85**, 9749–9766 (2011).

16. Kueng, H. J., Schmetterer, K. G. & Pickl, W. F. Lipid rafts, pseudotyping, and virus-like particles: Relevance of a novel, configurable, and modular antigen-presenting platform. *International Archives of Allergy and Immunology* **154**, 89–110 (2011).

17. Jolly, C. & Sattentau, Q. J. Human Immunodeficiency Virus Type 1 Assembly, Budding, and Cell-Cell Spread in T Cells Take Place in Tetraspanin-Enriched Plasma Membrane Domains. *J. Virol.* **81**, 7873–7884 (2007).

18. Khorchid, A., Halwani, R., Wainberg, M. A. & Kleiman, L. Role of RNA in Facilitating Gag/Gag-Pol Interaction. *J. Virol.* **76**, 4131–4137 (2002).

19. Muriaux, D., Mirro, J., Harvin, D. & Rein, A. RNA is a structural element in retrovirus particles. *Proc. Natl. Acad. Sci.* **98**, 5246–5251 (2001).

20. Comas-Garcia, M. *et al.* Efficient support of virus-like particle assembly by the HIV-1 packaging signal. *Elife* **7**, 1–11 (2018).

21. Rulli, S. J. *et al.* Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. *J. Virol.* **81**, 6623–6631 (2007).

22. Goff, S. P. *et al.* Dissection of specific binding of HIV-1 Gag to the 'packaging signal' in viral RNA. (2017). doi:10.7554/eLife.27055.001

23. Colombo, M., Raposo, G. & Théry, C. Biogenesis, Secretion, and Intercellular Interactions of

Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol.* **30**, 255–289 (2014).

24. Trajkovic, K. *et al.* Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes. *Science (80-.).* **319**, 1244–1247 (2008).

25. van Niel, G. *et al.* The Tetraspanin CD63 Regulates ESCRT-Independent and -Dependent Endosomal Sorting during Melanogenesis. *Dev. Cell* **21**, 708–721 (2011).

26. Shirk, A. J., Anderson, S. K., Hashemi, S. H., Chance, P. F. & Bennett, C. L. SIMPLE interacts with NEDD4 and TSG101: Evidence for a role in lysosomal sorting and implications for Charcot-Marie-Tooth disease. *J. Neurosci. Res.* **82**, 43–50 (2005).

27. Zhu, H. *et al.* Mutation of SIMPLE in Charcot-Marie-Tooth 1C alters production of exosomes. *Mol. Biol. Cell* **24**, 1619–1637 (2013).

28. Putz, U. *et al.* The tumor suppressor PTEN is exported in exosomes and has phosphatase activity in recipient cells. *Sci. Signal.* **5**, (2012).

29. Stavenhagen, K., Kolarich, D. & Wuhrer, M. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography–Mass Spectrometry. *Chromatographia* **78**, 307–320 (2014).

30. Zeng, S., Baillargeat, D., Ho, H. P. & Yong, K. T. Nanomaterials enhanced surface plasmon resonance for biological and chemical sensing applications. *Chemical Society Reviews* **43**, 3426–3452 (2014).

31. Jin, C., Hantusch, B., Hemmer, W., Stadlmann, J. & Altmann, F. Affinity of IgE and IgG against cross-reactive carbohydrate determinants on plant and insect glycoproteins. *J. Allergy Clin. Immunol.* **121**, (2008).

32. Rich, R. L. & Myszka, D. G. Higher-throughput, label-free, real-time molecular interaction analysis. *Analytical Biochemistry* **361**, 1–6 (2007).

CONCLUSIONS

Form the work performed in this thesis and the interpretation of the obtained experimental results, the following conclusions are proposed:

1. At high cell densities, HEK293 cells present an upregulation in mitochondrial content and activity and cell division proteins including nucleosome-associated proteins and histones. At the same time, lipid biosynthesis and intracellular protein transport to the nucleus are downregulated.

2. Transient transfection in HEK293 cells triggers a systemic disruption of homeostasis, reducing cellular viability and including a downregulation in glycosphingolipids metabolism and alterations in calcium regulation, oxidant detoxification, xenobiotic metabolism, peptidases activity and DNA detoxification.

3. When producing VLPs, Ca^{2+} import-related proteins are downregulated influencing cell membrane depolarization.

4. Overexpression of NEDD4L, CIT, and UGCG improve VLP production 3.3, 2.9, and 2.4-fold and budding efficiency up to 11%, 21%, and 13%, respectively. CNP knockdown increases VLP production 2.7-fold and budding percentage up to 76%. Combining CNP and UGCG at a ratio of [2:1] at 1.6 µg/ml of total DNA improves budding efficiency 37.3% compared to the Gag::eGFP control transfection. Harvested VLPs achieve 45% purity in the supernatant only knockingdown CNP.

5. The EGE methodology was optimized at bioreactor scale retransfecting at 24 hpt and a DNA concentration of 1.7 µm g/mL maintaining a CSPR of 30pL/cell/day. The perfusion-based intensified bioprocess achieves a volumetric productivity of $7.1 \cdot 10^{12}$ VLPs·L^{-1}·day^{-1} improving the original EGE protocol 86.7% or 7.54 fold and the first bioreactor approach using acoustic filter 67.8% or 3.1 fold respectively.

6. XCell® ATF2 hollow fiber module of 0.5 and 0.2 μm cutoff retain 99% of VLPs up to 168 hpt.

7. VHU2® filter module presents an average of 54% of VLP retention up to 96hpt.

8. Three Gag peptides, ETINEEAAEWDR, FAVNPGLLETTSEGCR, and TLNAWVK were validated for monitoring Gag stoichiometry in VLPs. VLPs produced by TGE in HEK293 cells have an average size of 145 nm and contain 3617 ± 17 monomers of Gag. The monomer/volume ratio is constant in HEK293 producing via TGE regardless of the culture media. Producing VLPs via SGE increases this ratio. For Sf9, when producing via the BV-infection system, VLPs contain 3755 ± 174 monomers of Gag::eGFP and the same monomer/volume ratio as HEK293 cells via TGE. For High Five cells, the ratio is significantly lower and a different distribution of VLP size is observed shifting from the average size 145 nm to 195 and 275 nm.

9. When producing VLPs, there is a shift from large EVs to small EVs biogenesis, increasing in 30% the production of vesicles smaller than 200 nm. In this EV fraction, a disruption of energy and redox homeostasis is observed. DNA is secreted in these small EVs, and present proteins associated with immune activation, anti-inflammatory processes, and response to oxidative stress.

10. HIV-1 Gag VLPs present 74.5% of complex, 13.7% high mannose and 11.8 hybrid *N*-glycans in the cell membrane. They also present 44.4% sialylated core 2, 33.3% sialylated core 1, 11.1% sialylated core 3 and 11.1% non-sialylated core 2 *O*-glycans. *N*-glycans like H5N5, H7N6F1S4a, H7N6F1S4b and H5N5F1S1 show significant differences in VLPs compared to EVs.

ACKNOWLEDGEMENTS

CURRICULUM VITAE

Jesús Lavado García

Bioprocess engineering
development and optimization

📍 Barcelona, Spain

WORK EXPERIENCE

Research – PhD in Biotechnology 📅 April 2017 - Present

PhD Thesis: Bioprocess engineering for HIV-1 Gag VLP production in HEK293 cells. Using proteomics tools to characterize transfection and production, design of experiments and perfusion to optimize and intensify the upstream process. Funded by **La Caixa** PhD scholarship (LCF/BQ/ES17/11600003) in the Autonomous University of **Barcelona** (UAB).
Research stay at Leiden University Medical Center (LUMC) in the Center for Proteomics and Metabolomics (**Leiden**) Sep2020-Jan2021 funded **by EMBO**.

Research – CNIC 📅 Sep 2015 – Jan 2017

Carrying out a project based on proteomic analysis of plasma from *Sus scrofa* after ischemia-reperfusion using avant-garde proteomics technology in **Madrid.**

Research - Internship CNIC 📅 June 2015 – Aug 2015

Identification of new biomarkers for heart- attack disease using the avant-garde proteomics technology. Collaborating in the PESA project. From June 2014-Aug 2014 Researching in the identification of Clec9a ligand and effect of *C.albicans* infection and neutrophil infiltration in *M.musculus.* in **Madrid**.

Research - Internship in CABD 📅 Oct 2014 – June 2015

Internship in Andalusian Centre of Developmental Biology (CABD) in **Seville**, collaborating in the project: Interaction of the mitochondria and cellular signaling pathways in the regulation of lifespan in *C.elegans.* Funded by **JAE Intro grant**, by the Ministry of Spain in collaboration with the Scientific Council (CSIC).

Designer in "MoleQla" magazine 📅 Sep 2011 – June 2015

Designer in the Science magazine "MoleQla" at University Pablo de Olavide (**Seville**). Section: "MoleQla Ambiental" and "MoleQla Industrial".

PERSONAL SKILLS

Spanish
Native
● ● ● ●

English
Proficiency (C2)
● ● ● ●

French
Independent (B2)
● ● ◐ ○

Mandarin (中文)
Independent (B2)
● ● ◐ ○

Computer Skills

- Experience in scripting, modeling, simulation and bioprocessing software: R, Expert Design, ImageJ, GraphPad (Prism).

Additional Courses

- Online course of 8 weeks in nuclear physics "A look at Nuclear Science and Technology" by the University of Pittsburgh, May 2014 with Distinction.
- Online course of 16 weeks in Calculus & Differential Equations by the University of Boston.
- Online course of 8 weeks in Graphene Science and Technology by Chalmers University of Technology.
- Online course of 10 weeks in Chinese Thought: Ancient Wisdom Meets Modern Science and 7 weeks in The Science of Religion by the University of British Columbia.

EDUCATION AND TRAINING

Master Degree in Industrial and Environmental Biotechnology   📅 Sep 2015 – Mar 2017

University Complutense of Madrid (Madrid). Applied biotechnology in industrial production of pharmaceuticals and other products of industrial interest.
Average grade: 9,3 (out of 10), 1st of my promotion.

Degree in Biotechnology   📅 Sep 2011 – June 2015

University Pablo de Olavide (Seville)
Average grade: 3,3 (out of 4) or 9,65 (out of 10), 1st of my promotion

Atlánticus Programme: Washington & Jefferson College, USA   📅 Aug 2013 – Dec 2013

Courses taken in this University: Animal Physiology (With honors), Intermediate Organic Chemistry, Advanced Chemical Measurements (With honors), Bioinformatics investigation of Manduca sexta muscle development genes.

SCIENTIFIC PUBLICATIONS

- **Lavado-García, J**., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed quantitative proteomic analysis of HEK293 provides insights of molecular changes associated to the cell density effect, transient transfection and virus-like particles production. J. Proteome Res. 19, 1085–1099 (2020). DOI:10.1021/acs.jproteome.9b00601

- **Lavado-García, J**., Cervera, L. & Gòdia, F. An alternative perfusion approach for the intensification of virus-like particle production in HEK293 cultures. Front. Bioeng. Biotechnol. 8, 617 (2020). DOI: 10.3389/fbioe.2020.00617

- **Lavado-García, J**., González-Domínguez, I., Cervera, L., Jorge, I., Vázquez, J. & Gòdia, F. Molecular characterization of the co-produced extracellular vesicles in HEK293 during virus-like particle production. J. Proteome Res. 19, 4516–4532 (2020). DOI: 10.1021/acs.jproteome.0c00581

- **Lavado-García, J**. et al. Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles. Biotechnol. Bioeng. (2021). DOI:10.1002/bit.27679

- **Lavado-García, J**. et al. Characterization of HIV-1 virus-like particles and determination of Gag stoichiometry for different production platforms. Biotechnol. Bioeng. (2021). DOI:10.1002/bit.27786

SCIENCE DISSEMINATION PUBLICATIONS

- "Medicina Nuclear. Tomografía por Emisión de Positrones", (MoleQla ISSN-e-2173-0903 N°4, 2011, pp 109-111)
- "EDTA y la terapia de Quelación" (MoleQla ISSN-e-2173-0903 N°6, 2012, pp 27-29)

- "Policaprolactona (PCL) y su aplicación a la regeneración tisular" (MoleQla   ISSN-e-2173-0903 N°10, 2013 , pp 71-73)
- "Death Regulation Beyond Apoptosis: Necroptosis" (MoleQla   ISSN-e-2173-0903 N°17, 2015 , pp 15-79)

HONOURS AND AWARDS

- Award Academic Trajectory (2011) by Ayuntamiento de Sevilla.
- Distiction in pre-University studies by Ayuntamiento de Sevilla.
- Award to the best academic record in Sevilla in 2011-2012 academic year by University Pablo de Olavide.
- Awartd to the best grade to Access Degree in Biotechnology by University Pablo de Olavide.
- Award to the best grade during the first year of Biotechnology by University Pablo de Olavide.
- Distinction being in the Dean's List of the Fall semester in the University Washington & Jefferson College, PA, USA
- Degree honors in: Cell Biology, General Chemistry, Organic Chemistry, Physics, Biochemistry, Animal & Plant Biology, Thermodynamics & Chemical Kinetics, Microbial metabolism and physiology, Microbiology, Separation operations, Bioinformatics, Economics & Business administration.
- Poster Award in the BIOIberoamerica 2016 Biotechnology congress by SEBiot.
- Award to best Academic Record of the Degree in Biotechnology, promotion of 2015. UPO-Sevilla.
- Award to the best Academic Record in faculty of Experimental Sciences year 2014-2015 in Universidad Pablo de Olavide- Sevilla.
- National Award to one of the best academic record in Biotechnology year 2014-2015 in Spain.

GRANTS AND FELLOWSHIPS

- CSIC "Introduction to Research" **JAE Intro** 2014.
- CNIC "**Cicerone**" program 2014 and 2015.
- CNIC "Beca Máster" fellowship for academic year 2015-2016 and 2016-2017.
- **La Caixa PhD Fellowship** (2017), code: LCF/BQ/ES17/11600003.
- European Molecular Biology Organization (**EMBO**) Short-term Fellowship STF 8613 (2020).

CONGRESSES

- Participation in IInd Congress of Educational Success "Análisis del Bajo Guadalquivir" as a presenter representing IES Macarena.
- Participation in the Poster Session in Washington & Jefferson College, PA, USA presenting the project "Analysis of a Venetian painting using advanced chemical measurements" carried out during the semester.
- Participation and attendance to the conferences of the American Chemical Society ACS "On the road" Conference; "Aqueous Electrolyte Polyionic Electrochemical Cells For Scaled Stationary Energy Storage" Oct 8, 2013.
- Participation in the Biotech Annual Congress 2014 (BAC2014) in Barcelona.
- Participation in poster session and award reception in the BIOIberoamerica 2016 Biotechnology congress.

- Participation in Integrated Continuous Biomanufacturing III Congress September 17-21, 2017 Cascais, Portugal.
- Participation in the 26th European Society Animal Cell Technology (ESACT) Congess in Copenhagen, May 4-9 2019.

MEMBERSHIPS

- ASBAN – Asociación de Biotecnólogos de Andalucía (Association of Biotechnologists of Andalusia)
- Participation in the Innovation Project: "MoléQla. Un proyecto de Innovación Docente Universitaria en el Espacio Europeo de Educación Superior " Action 2, Summon (2012-2013)
- SEBIOT – Sociedad Española de Biotecnología

涟漪在永恒