Production of high-quality SARS-CoV-2 antigens: impact of bioprocess and storage on glycosylation, biophysical attributes, and ELISA serologic tests performance

Rute Castro[1], Lígia S. Nobre[1], Rute P. Eleutério[1], Mónica Thomaz[1,2], António Pires[1], Sandra M. Monteiro[1], Sónia Mendes[1], Ricardo A. Gomes[1], João J. Clemente[1], Marcos F. Q. Sousa[1,2], Filipe Pinto[1], Ana C. Silva[1], Micael C. Freitas[1], Ana R. Lemos[1], Onome Akpogheneta[3], Lindsay Kosack[3], Marie-Louise Bergman[3], Nadia Duarte[3], Paula Matoso[3], Júlia Costa[2], Tiago M. Bandeiras[1,2], Patricia Gomes-Alves[1,2], Carlos Penha Gonçalves[3], Jocelyne Demengeot[3] and Paula M. Alves[1,2*]

[1]*iBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780-901, Oeiras, Portugal.*

[2]*ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.*

[3]*IGC, Instituto Gulbenkian de Ciência, Rua da Quinta Grande, 6, 2780-156 Oeiras, Portugal.*

***Corresponding author:** P. M. Alves

E-mail: marques@ibet.pt; Tel: +351 21 446 9416

Abstract

SARS-CoV-2 is an RNA coronavirus that causes severe acute pneumonia, also known as COVID-19 disease. The World Health Organization declared the COVID-19 outbreak in January 2020 and a pandemic 2 months later. Serological assays are valuable tools to study virus spread among the population and, importantly, to identify individuals that were already infected and would be potentially immune to a virus re-infection. SARS-CoV-2 Spike protein and its Receptor Binding Domain (RBD) are the antigens with higher potential to develop SARS-CoV-2 serological assays. Moreover, structural studies of these antigens are key to understand the molecular basis for Spike interaction with angiotensin converting enzyme 2 receptor, hopefully enabling the discovery and development of COVID-19 therapeutics. Thus, it is urgent that significant amounts of this protein became available at the highest quality.

In this work we evaluated the impact of different and scalable bioprocessing approaches on Spike and RBD production yields and, more importantly, in these antigens' quality attributes. Using negative and positive sera collected from human donors, we show an excellent performance of the produced antigens, assessed in serologic ELISA tests, as denoted by the high specificity and sensitivity of the test. We have shown that, despite of the human cell host and the cell culture strategy used, for production scales ranging from 1 L to up to 30 L, final yields of approx. 2 mg and 90 mg per liter of purified bulk for Spike and RBD, respectively, could be obtained. To the best of our knowledge these are the highest yields for RBD production reported to date.

An in-depth characterization of SARS-CoV-2 Spike and RBD proteins was also performed, namely the antigens oligomeric state, glycosylation profiles and thermal stability during storage. The correlation of these quality attributes with ELISA performance show equivalent reactivity to SARS-CoV-2 positive serum, for all Spike and RBD produced, and for all the storage conditions tested.

Overall, we provide herein straightforward protocols to produce high-quality SARS-CoV-2 Spike and RBD antigens, that can be easily adapted to both academic and industrial settings; and integrate, for the first time, studies on the impact of bioprocess with an in-deep characterization of these proteins, correlating antigens glycosylation and biophysical attributes to performance of COVID-19 serologic tests. We strongly believe that our work will contribute to advance the current and recent knowledge on SARS-CoV-2 proteins and support the scientific society that is persistently searching for solutions for COVID-19 pandemics.

Key words: SARS-CoV-2, COVID-19, serologic assay, ELISA, Spike, RBD, bioprocess, production and purification, thermal stability during storage, glycosylation.

Introduction

The new **CO**rona**VI**rus **D**isease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first detected in the Wuhan region in China, in December 2019 (Zhou et al., 2020). In March 2020, COVID-19 outbreak was declared pandemic by the world health organization (WHO) and, by mid-October 2020, the virus was responsible for the infection of more than 38 million people and caused 1 million deaths worldwide (WHO, Coronavirus Disease Dashboard). Currently, there are several therapeutics and vaccines in clinical trials, but there is no antiviral therapy or prophylaxis available for SARS-CoV-2 virus. Thus, COVID-19 prevention strategy relies on the implementation of social distancing measures, that have economic and social impact.

The development of serological assays to study the human response to SARS-CoV-2 have been reported (Amanat et al., 2020; Weiss et al., 2020; Perera et al., 2020; Okba et al., 2020). It is well known that some infected individuals are asymptomatic, therefore a broad application of serological assays will provide clear epidemiological data regarding the SARS-CoV-2 infection rate among a population, as well as the real mortality rates for COVID-19. Moreover, the identification of individuals that were already infected, and therefore, would possibly be immune to virus re-infection, has important social and economic impact. Serological assays based on SARS-CoV-2 Spike protein and its Receptor Binding Domain (RBD), present good sensitivity and specificity (Amanat et al., 2020; Okba et al., 2020).

SARS-CoV-2 Spike glycoprotein mediates virus entry in the target cells via its binding to the angiotensin converting enzyme 2 (ACE2) receptor. The determination of SARS-CoV-2 Spike protein structure, provided good indications for the development of vaccines and inhibitors (Walls et al., 2020; Wrapp et al., 2020; Wang et al., 2020). Additionally, despite the high structural similarity between Spike proteins from SARS-CoV-2 and SARS-CoV viruses, no antibody cross-reactivity has been detected (Wrapp et al., 2020). Characterization of Spike glycosylation profile as been the subject of several studies, due to its perceived importance on the development of COVID-19 therapies or prophylaxis (Shajahan et al., 2020; Watanabe et al., 2020). Indeed, mapping of SARS-CoV-2 Spike glycosylation using a cryo-EM structure of the protein suggested the shielding of receptor binding domain by proximal glycans (Watanabe et al., 2020).

With the ultimate goal of providing high-quality substrates to perform SARS-CoV-2 serological assays we investigated the production process for Spike and RBD antigens, using two human cells hosts, HEK293-E6 (Durocher, 2002) and Expi293F™. Different cell culturing approaches and production scales were evaluated. The impact of downstream processing steps and distinct storage temperature conditions were also assessed. An in-depth characterization of the antigens was performed correlating oligomeric state, glycosylation profile and thermal stability with the bioprocess set-up and the storage conditions.

Finally, the quality of the antigens was assessed by evaluating their performance in ELISA serological tests using human serum control samples.

Materials and Methods

Recombinant proteins

Plasmid DNA for the expression of SARS-CoV-2 Spike and Spike's receptor binding domain (RBD) was kindly provided by Prof. Florian Krammer (Icahn School of Medicine at Mount Sinai, NY, USA). Soluble Spike protein presents a T4 foldon trimerization domain, a C-terminal hexahistidine tag, two stabilizing mutations and includes the removal of polybasic cleavage site (further details described by Amanat and colleagues (Amanat et al., 2020)). Soluble RBD includes the signal peptide and C-terminal hexahistidine tag. Transfection grade plasmids were obtained from 2.5 L cultures of *E. coli* DH5α transformed with Spike or RBD expression vectors, using the Qiagen Giga Prep kit or equivalent, following the manufacturer instructions.

Cell lines, culture conditions and cell concentration determination

HEK293-E6 cells (Durocher, 2002) were cultured in suspension in FreeStyle™ F17 expression medium, supplemented with 4 mM Glutamax, 0.1 % Pluronic F-68 and 25 µg/mL of Geneticin, in shake-flasks at 37 ˚C in an humidified atmosphere of 5 % CO_2 in air, and stirring rates of 75 or 90 rpm. Expi293F™ cells (Thermo Fisher Scientific, MA, USA) were cultivated in Expi293™ Expression Medium, according to the manufacturer instructions. All media and cell culture supplements were from Thermo Fisher Scientific.

Cell concentration and viability was determined by the trypan blue (Gibco, Life Technologies Ltd, Paisley, UK) exclusion method using a 0.1 % (v/v) solution prepared in Dulbecco's phosphate-buffered saline (DPBS; Gibco) and counting cells in a Fuchs Rosenthal haemacytometer (Brand Wertheim, Germany) using an inverted microscope (Olympus CK40, Tokyo, Japan). Viable cell concentration was also monitored using NucleoCounter® NC-200™ (Chemometec, Allerod, Denmark).

Spike and RBD production in Human cell lines

Exponentially growing HEK293-E6 cells were transfected with 1 mg of plasmid DNA per liter of culture, complexed with polyethylenimide (PEI, Polysciences, Warrington, PA, USA), in a DNA:PEI ratio of 1:2. Six hours post-transfection, 0.5 mM of valproic acid (Merck KGaA, Darmstadt, Germany) was added to the cultures. Cell concentration and viability was monitored every day and cultures were harvested 3 to 5 days post-transfection.

Three different cell culturing strategies were tested: 2.5 L culture volume in 5 L in shake-flasks (Corning, NY, USA), stirred tank bioreactors of 2 and 5 L (STB, Sartorius, Gottingen, Germany) and wave bioreactors up to 30 L (Sartorius). In STB (Biostat DCU-3) dissolved oxygen (DO) was kept at 40 % (in air) by sequentially varying stirring rate and the percentage of oxygen in gas inlet (sparger in the bottom of the vessel), at a constant aeration rate of 0.01 vvm. pH was controlled at 7.2 using the addition of CO_2 or $NaHCO_3$ and temperature was controlled at 37 °C using a heating jacket. The wave bioreactor (Biostat Cultibag RM) cultures were performed at rocking angle of 8°, 18 rocks/min, and a continuous supply of air with 5 % of CO_2 through the headspace, at a rate of 0.02 vvm. Temperature was maintained at 37 °C. Process control and monitoring was carried out using Multi Fermenter Control Software (Sartorius).

For the Expi293F™ cells, shake-flask cultures at 1 L scale were used, according to the manufacturer instructions. Cell concentration and viability was monitored every day and the cultures were harvested at 3 days post-transfection. The bulk from 5 cultures (5 L), *per* Spike or RBD production run, was pooled together and purified as described below.

The impact of decreasing the temperature to 32 °C during protein expression and lowering the coding DNA amount to 50 % (0.5 µg/ml of culture) was also evaluated at small scale (25 ml shake-flasks cultures). For the 32 °C experiments, cells were transfected as described above and, 18 h post-transfection, the cultures were moved to an incubator at 32 °C. In the 50 % DNA experiments, transfection was performed with a total of 1 µg DNA/mL of culture consisting in 0.5 µg of Spike or RBD expression vector and 0.5 µg of empty pTT5™ expression vector per mL of culture. Cell concentration and viability was monitored every day and culture supernatants were analyzed by SDS-PAGE.

For all the experimental production set-ups glucose, lactate, glutamine, glutamate and ammonia were quantified in culture supernatants using Cedex Bio analyzer (Roche, Basel, Switzerland).

Protein purification and quantification

At day 3 – 5 post-transfection, cell culture bulks were clarified by centrifugation at 2000 x g for 20 minutes at 4 °C, followed by filtration using 0.2 µm filters (Sartopore 2, Sartorius). Tangential flow

filtration (TFF) was used to concentrate and dialyse the clarified supernatants to 50 mM Sodium Phosphate supplemented with 300 mM NaCl and 20 mM Imidazole, at pH 7.4 (binding buffer). Membranes of 10 or 30 kDa (Sartorius) were used for RBD or Spike, respectively.

After the TFF step, a chromatographic step was performed in an Äkta (GE Healthcare, IL, USA) using HisTrap HP columns (GE Healthcare), previously equilibrated with binding buffer. Two washing steps with 35 and 50 mM imidazole were performed, and proteins were eluted with a linear gradient up to 500 mM Imidazole. Spike and RBD eluted from the nickel affinity chromatography (AC) were concentrated using Vivaflow 200 crossflow device (Sartorius) and subjected to size exclusion chromatography (SEC) using Superdex 200 or Superdex 75 columns (GE Healthcare), respectively, previously equilibrated with phosphate buffered saline at pH 7.4 (PBS, formulation buffer). The proteins eluted from SEC were concentrated, filtered, fast freeze in liquid nitrogen and stored at - 80 °C.

In the STB producion runs, to reduce further processing time and product losses, the SEC step was replaced by buffer exchange using Vivaflow 200 crossflow devices and performing dialysis with minimal of 10 volumes of formulation buffer. For purification of RBD produced in Expi293F™ cells, an additional washing step at 68 mM imidazole was included before elution with linear gradient to 500 mM imidazole, to improve protein purity. Additionally, the protein fractions eluted from AC were desalted to formulation buffer, using G25 Sephadex desalting column (GE Healthcare).

Alternativelly, wave bioreactor bulks containing Spike protein were also purified performing the clarification step with high capacity filters of 1.8 m^2 filtration area and 0.45 µm / 0.2 µm pore size (Sartopore 2 MaxiCaps, Sartorius), and the protein eluted from AC was subjected to concentration and dialysis using Vivaflow 200 crossflow device (Sartorius). This way the centrifugation and SEC steps were avoided.

Protein concentration was determined by A_{280nm} combined with the specific extinction coefficients, using MySpec spectrofotometer (VWR, Radnor, PA, USA). Final pure products were also quantified by Thermo Fisher Scientific Pierce™ BCA Protein Assay Kit and Pierce™ Coomassie Plus (Bradford) Assay reagent.

SDS-PAGE analysis was performed by loading protein samples on 4 - 12% Bis-Tris NuPAGE gels (Thermo Fisher Scientific) using 3-(N-morpholino)propanesulfonic acid (MOPS) buffer and standard running conditions. For reducing conditions, 200 mM of dithiothreitol (DTT) was added to loading buffer and protein samples were heated at 100 °C for 3 minutes. Protein bands were revealed by incubation with InstantBlue™ (Expedeon Protein Solutions, Cambridgeshire, UK) and SDS-PAGE de-staining was performed in water.

Spike and RBD Characterization

High-Performance Liquid Chromatography (HPLC) Analysis

Spike and RBD samples were analysed in an HPLC system equiped with Photodiode Array Detector (Waters, MA, USA). Spike samples were injected in XBridge BEH 200 Å SEC 3.5 µm or XBridge BEH 450 Å SEC 3.5 µm HPLC columns (Waters), at 0.86 ml/min, using as mobile phase PBS pH 7.4. RBD samples were injected in XBridge BEH 125 Å SEC 3.5 µm HPLC column (Waters) at 0.86 ml/min using as mobile phase PBS with 0.5 M of Arginine at pH 7.4. Twenty µg of protein was injected in each HPLC run.

Differential Scanning Fluorimetry (DSF)

DSF was performed in MicroAmp™ EnduraPlate™ Optical 96-Well Clear Reaction Plates with Barcode (Applied Biosystems, Life Technologies, California, USA) using a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). Protein samples were centrifuged for 15 minutes before preparation. The final reaction mixture (20 µL of total volume) contained 4 µg of either Spike or RBD protein and Protein Thermal Shift™ Dye (1000 X stock, Applied Biosystems) diluted 1:250 in PBS pH 7.4. Melting curve data was recorded from 15 to 90 °C with an increment rate of 0.016 °C.s^{-1}. Excitation and emissions filters were applied for Protein Thermal Shift™ Dye (470 nm and 520 nm, respectively) and for ROX reference dye (580 nm and 623 nm, respectively). The melting temperatures were obtained by calculating the midpoint of each transition, using the Protein Thermal Shift Software™ version 1.3. All samples were tested in duplicates.

Nano Differential Scanning Fluorimetry (NanoDSF)

NanoDSF was performed on a Prometheus NT.48 instrument (NanoTemper Technologies GmbH, Munich, Germany). Protein samples were centrifuged for 15 minutes before preparation. The final reaction mixture contained 4 µg of either Spike or RBD proteins diluted in PBS pH 7.4. High sensitivity capillaries (NanoTemper Technologies) were filled with 10 µL of sample and placed on the sample holder. A temperature gradient of 1 °C.min^{-1} was applied from 15 to 95 °C and the intrinsic protein fluorescence at 330 and 350 nm was recorded. Data was analysed using either the value of fluorescence at 330 nm (for Spike protein) or the derived ratio 350/330 value (for RBD protein). All samples were tested in duplicates.

Dynamic Light Scattering (DLS) Analysis

A SpectroLight 610 (Xtal Concepts GmbH, Hamburg, Germany) was used to carry out serial DLS measurements. All samples were centrifuged (15 - 30min, 4 °C, 17200 × g) in a benchtop centrifuge before measurements and were pipetted (each sample in duplicate, 1 µL per well) onto a 96 well Vapor Batch Plate (Jena Bioscience GmbH, Jena, Germany). Prior to usage, the plates were filled with paraffin oil (Cat N. 18512, Merck) to protect sample solutions from drying out. The used laser wavelength was 660 nm at a power of 100 mW. The scattering angle for placement of the detector was fixed at 142°. All tested samples were kept in PBS buffer, therefore, the refractive index (1.34) and viscosity (1.006) of water were used for calculations. All samples were measured at constant 20 °C, one scan per drop with 20 measurements of 20 seconds each.

Glycosylation pattern analysis by LC-MS, Lectin blotting and glycosidase digestion

Proteins were analyzed by SDS-PAGE and transferred to polyvinyledene fluoride membranes, which were blocked with 3 % BSA biotin free (Carl-Roth, Karlsruhe, Germany) in Tris Buffered Saline supplemented with tween 20 (TBST) for 1 hour. Then, membranes were incubated with lectins from *Wisteria floribunda* agglutinin (WFA; Vector Laboratories, Burlingame, CA, USA), *Maackia amurensis* lectin (MAL), *Sambucus nigra* agglutinin (SNA), *Aleuria aurantia* lectin (AAL) and *Galanthus nivalis* agglutinin (GNA) (Galab Technologies, Geesthacht, Germany), at 1 µg/ml in TBST for 1 hour. Blots were washed with TBST and incubated with 0.1 µg/mL streptavidin-peroxidase (Merck KGaA) for 1h. Blots were washed, and detection was performed with the Immobilon Western chemiluminescent HRP substrate (Millipore).

Glycoproteins were digested with peptide N-glycosidase F (PNGase F; ProZyme, Agilent, Santa Clara, CA, USA), endoglycosidase H (Endo H; Roche) and sialidase from *Arthrobacter urefaciens* (Roche) as previously described (Escrevente et al., 2011; Machado et al., 2011).

The glycosylation profiling of Spike and RBD proteins was further assessed using LC-MS analysis of glycopeptides. Thirty micrograms of protein sample were subjected to trypsin digestion. Briefly, protein sample was denatured with 6 M guanidine hydrochloride, reduced with 10 mM DTT (Merck KGaA) for 15 min at 56 ˚C followed by alkylation with 20 mM iodoacetamide (Merck KGaA) for 30 min in the dark. Excessive iodoacetamide was quenched by further incubation with DTT (10 mM for 10 min in the dark). A step of buffer exchange was performed using Zeba Spin desalting plates (Thermo Fisher Scientific). The resulting sample was digested overnight with trypsin (Proteomics grade from Promega, Madison, WI, USA) at 37 ˚C (1:50 protein/trypsin ratio). Trypsin digestion was stopped with the addition of formic acid (1 % final concentration).

Glycopeptides were analysed by LC-MS using an X500B-QTOF mass spectrometer (SCIEX, Framingham, MA, USA) connected to an ExionLC AD UPLC system. LC separation was achieved through reversed-phase chromatography using an XBridge BEH C18, 2.5 µm 2.1 x 150 mm (Waters). Separation was performed at 200 µl/min with 0.1 % formic acid in water LC-MS grade as solvent A and 0.1 % formic acid in acetonitrile as solvent B, and column temperature was set to 40 ˚C. The LC gradient was as follows: 0-5 min, 1 % B; 5-50 min, 1-35 % B; 50-55 min, 35-90 % B; 55-56 min, 60-90 % B; 56-60 min, 90 % B; 60-62 min, 90-1 % B; 62-64 min, 1 % B.

Peptides were sprayed into the MS through the twin sprayer ion source with the following parameters: 50 GS1, 50 GS2, 30 CUR, 5.5 keV ISVF, 450 ^{o}C TEM, 80 V declustering potential and 10 V collision energy. An information dependent acquisition (IDA) method was set with a TOF-MS survey scan of 350-2000 m/z for 250 msec. The 12 most intense precursors were selected for subsequent fragmentation MS/MS mode (150-1800 m/z for 100 msec each). The selection criteria for parent ions included dynamic background subtraction and counts above a minimum threshold of 300 counts per second. Ions were excluded from further MSMS analysis for 5 s. Fragmentation was performed using rolling collision energy with a collision energy spread of 5.

MS data were analysed using the BioPharmaView software (Version 3.0, SCIEX) using the protein sequences of Spike and RBD proteins (Amanat et al., 2020). For glycans identification, N-glycans described in (Watanabe et al., 2020) were considered.

Enzyme-linked immunosorbent assay (ELISA)

Anti-Spike and RBD ELISA assay implemented followed a checkerboard strategy, whereby both the antigen and the positive sera were serially titrated. The ELISA was performed as described in (Stadlbauer et al., 2020) with minor modifications. Briefly, high binding 96-well plates (Corning) were coated with either RBD or Spike as capture antigen along a 1:2 dilution in PBS starting at 2 µg/ml, and blocked with PBS supplemented with 2 % BSA. Reference positive sera were submitted to a 1:3 serial dilution starting at 1 in 50 and reference negative sera were used at 1 in 50 dilution. Bound IgG was revealed with goat anti-Human IgG Fc-HRP (Abcam, Cambridge, UK) followed by incubation with 3,3′,5,5′-Tetramethylbenzidine (TMB, BD OptEIA™, BD Biosciences, Franklin Lakes, NJ, USA). The colorimetric assay was read at 450 nm. Reference sera were collected at least 7 days post the first PCR SARS-CoV-2 diagnostic (positive sera) or at least 3 years before the COVID-19 pandemic (negative sera). Positive and negative sera were obtained upon informed consent in the frame of the projects "Genetic susceptibility factors and immunologic protection in COVID-19", and “Genetic variance in Portuguese population: candidate genes in COVID-19", both approved by the IGC Ethic Committee (reference H004.2020 and H002.2020, respectively).

Results

SARS-CoV-2 Spike and RBD production

In this work we have used two human cell lines, HEK293-E6 and Expi293F™, for the production of SARS-CoV-2 Spike and RBD. The experimental set-up is summarized in Figure 1. Different approaches for the Upstream and Downstream processes were evaluated. The impact of these production strategies and scales in antigens quality was assessed by an in-depth biochemical and biophysical characterization. Thermal stability during storage and the impact of freeze-thaw cycles was also studied. Ultimately, antigen quality was confirmed by measuring the reactivity in ELISA COVID-19 serologic tests using human sera.

Spike and RBD global bioprocess final yields, obtained after purification (mg protein per L of harvested cell culture bulk) are presented in Figure 2A, that also shows, for comparison, data obtained from the literature (Amanat et al., 2020; Li et al., 2020; Esposito et al., 2020; Herrera et al., 2020; Claudia et al., 2020; Johari et al., 2020; Stuible et al., 2020).

Overall, for Spike, independently of the cell culture system (shake-flasks, stirred tank or wave bioreactors) or the scale (5 - 30 L) used, the day of harvest is the parameter that has the most impact in the antigens final yields, approx. 1 and 2 mg/L at Day 3 and 4, respectively. The one day extension of the culture, that allows for duplication of the productivity, has not compromised neither the quality of the protein, as assessed by SDS-PAGE and performance in the ELISA serologic tests (left panel of Figure 2B and 2C). Small-scale (200 mL) feasibility studies performed in shake-flasks indicated that one extra day of culture, i.e. harvesting at Day 5, led to protein degradation resulting in lower global final yields (data not shown). In fact, Spike degradation was only detected for the wave bioreactor runs (Figure 2B), being more evident for the first run. Additionally, Spike produced in the first wave bioreactor run presented lower performance in SARS-CoV-2 ELISA (Figure 2C).

Similar results were obtained for RBD in what concerns the impact on protein degradation and performance at ELISA tests, i.e. no significant degradation occurred for all the culture systems and scales tested and equivalent ELISA performance was observed. In fact, no degradation was detected for RBD purified from bulks harvested at Day 5, as assessed by SDS-PAGE (Figure 2B right panel).

For the different downstream strategies evaluated for Spike and RBD purification (see Materials and Methods) there was no major impact in global production yields. Affinity Chromatography (AC), in particular the washing and elution steps, allowed for important reduction of protein impurity profile with only single bands visible in SDS-PAGE (Figure 2B). This high purity was achieved independently of performing or not the size exclusion chromatography as polishing step.

The studies with different cell hosts, Expi293F™ and HEK293-E6, show that for similar culture conditions harvested at Day 3, as described in (Amanat et al., 2020), there is no significant effect in productivity for Spike. In contrast, higher yields were obtained for RBD when Expi293F™ cells were used (approx. 90 mg/L) only achieved by HEK293-E6 cells when the culture was extended two extra days (harvesting at Day 5, Figure 2A).

For all parameters studied, both for Spike and RBD, no statistically significant effect was observed in the performance of ELISA serologic tests (excluding the wave bioreactor run 1); in contrary, as expected, this performance is affected by the amount of antigen used to coat the ELISA plates (Figure 2C black, grey and white bars).

It is worth to mention that all HEK293-E6 cultures harvested after 4 days post-transfection were exhausted for glucose, in particular, the stirred tank bioreactor cultures, were depleted of glucose already at 2 days post-transfection (data not shown). In contrast, cultures of Expi293F™ cells expressing Spike or RBD had about 10 mM of glucose at harvesting time (Day 3 post-transfection) because according to the manufacturer instructions, an enhancer solution is added at 18 h post-transfection (data not shown). These results indicate the potential of further increase in protein expression yields in stirred tank bioreactor HEK293-E6 cultures by implementing a fed-batch operation mode with glucose supplementation.

Parallel small-scale expression screen runs were conducted with the ultimate goal of evaluating if the amount of coding DNA used for cell transfection and temperature shifts during production impacts antigen production yields. The results (Figure 3) show that lowering the expression temperature to 32 ˚C was only beneficial for Spike expression in Expi293F™, resulting in a significant increase in productivity (approx. 5-Fold as measured by densitometry). This productivity was further improved when the temperature shift to 32 °C was combined with lowering the coding DNA amount by 50 %. (Figure 3 top right). For HEK293-E6 cells, no improvement in antigens expression was obtained. Curiously, when Spike is produced at 32 ˚C, the SDS-PAGE bands migrate slightly less, suggesting differences in Spike glycosylation profiles.

Overall the results summarized in Figure 2 and 3 show that we were able to develop scalable bioprocesses and to produce and purify at large scale (up to 30 L) high-quality Spike and RBD antigens, maintaining the production yield reported in the literature for smaller scales using human derived cell hosts (Figure 2A).

Characterization of produced Spike and RBD

The extensive glycosylation of SARS-CoV-2 Spike protein is well described in the literature (Grant et al., 2020; Henderson et al., 2020; Shajahan et al., 2020; Watanabe et al., 2020). In this work, we investigated Spike and RBD glycosylation patterns, focusing on putative differences between the two expression hosts cells used.

We first performed lectin blotting analysis to evaluate the presence of specific glycans in Spike and RBD samples produced in HEK293-E6 and Expi293F™ cells. All proteins were detected with SNA and MAL, which indicated the presence of α2,6- and α2,3-linked Neu5Ac (Figure 4). Neu5Ac was largely present in N-glycans as evaluated from decrease in signal after digestion with PNGase F (Suppl. Figure 2A). As control, sensitivity to *A. urefaciens* sialidase supported the signal specificity (Suppl. Figure 2A). Proteins were also detected with GNA, therefore supporting the presence of high mannose structures, particularly evident in Spike protein samples. Specificity was confirmed by a decrease in signal after digestion with Endo H (Suppl. Figure 2B).

Proteins were also detected with WFA (binds terminal GalNAc), therefore, indicating the presence of the GlcNAcGalNac (LacdiNAc) structure; this structure was present in N-glycans as inferred from PNGase F sensitivity (Figure 4; Suppl. Figure 2C). Proteins were also strongly detected with AAL (Figure 4), which indicated proximal fucosylation and possibly peripheral fucosylation.

Spike and RBD have 22 and 2 potential N-glycosylation sites, respectively. We have performed LC-MS analysis to screen for the presence of N-glycans previously identified by others in SARS-CoV-2 Spike protein (Watanabe et al., 2020). A detailed description of the different glycoforms detected in each site from tryptic peptides of Spike and RBD produced in HEK293-E6 cells and Expi293F™ cells is presented in Supplementary data (Files “Glycans comparison” and “MS data”).

To investigate site-specific glycosylation and differences between HEK293-E6 and Expi293F™ cells comparison of N-glycan compositions was also done (Suppl. File “Glycans comparison”). For a qualitative evaluation, structures (compatible with the monosaccharide compositions, the N-glycan biosynthetic pathway and the lectin blotting above) have been proposed using GlycoWorkbench (Ceroni et al., 2008; Suppl. Table 2). Glycoforms of each identified site from Spike and RBD, produced either in HEK293-E6 or Expi293F™ cells, were compared based on the following structural features: presence of high mannose glycans, fucosylation (proximal and peripheral), sialylation, sialylation and fucosylation and detection of sialyl-LacdiNAc (Table 1 and Table 2).

SARS-CoV-2 Spike and RBD from the two host cells displayed high mannose, paucimannose, hybrid and complex di-, tri- and tetraantennary glycans. Many fucosylated structures were detected, which included proximal fucose and also peripheral fucose (the latter inferred from the presence of more than one fucose residue in glycan compositions, e.g., structures 47, 72, 81, etc). Complex

glycans were partially or completely sialylated. The proteins bound the lectin WFA (Figure 4), which indicated the presence of LacdiNAc, compatible with proposed structures 49, 50, 52-54, 56, 57, 77, 78, 80- 82, 86-89, 91, 107-109. The presence of LacdiNac was further supported by the unequivocal finding of sialylated LadiNAc in structures 27, 51, 55, 59, 79 and 90. It should be considered that bisecting GlcNAc-containing structures are also compatible with the detected masses (Suppl. Table 2).

Glycoform profiles of individual Spike sites were quite distinct. Most striking was: the abundance of high mannose glycans at N234, low heterogeneity at N17 with a relatively high proportion of hybrid and paucimannose structures but undetectable high mannose structures; high heterogeneity at sites N61/N74 and N331/N343 (which could be attributed to the presence of two N-glycosylation sites in the same tryptic peptide).

When Spike glycoform profiles were compared between host cells some differences were found, but not striking (Table 1, suppl. files “Glycans comparison Spike samples”). On the other hand, for RBD glycoforms, which has only two N-glycosylation sites (N331 and N343 both present in the same tryptic peptide), a remarkable difference was found between HEK293-E6 cells and Expi293F™ cells derived protein (Table 2). The heterogeneity was very high in RBD produced in HEK293-E6 cells, with the detection of large high mannose glycans ($Man_9GlcNAc_2$ and $Man_8GlcNAc_2$) and several sialylated/ fucosylated complex glycans. By contrast RBD produced in Expi293F™ cells exhibited fewer glycoforms, which included the small high mannose glycan (Man_5GlNAc_2), several paucimannose/hybrid structures, and fewer complex sialylated/fucosylated glycans.

Besides the analysis of Spike and RBD glycosylation profile, we also proceeded with the assessment of overall oligomeric state of the proteins. SARS-CoV-2 Spike forms trimers that interact with ACE2 receptor at the RBD region (Wrapp et al., 2020; Esposito et al., 2020; Walls et al., 2020). We have started by performing size-exclusion HPLC analysis of Spike protein in a Xbridge BEH200 column. The results indicated that Spike has a molecular weight of approximately 600 kDa, confirming its trimeric conformation (Figure 5A). Since we were interested in evaluating the presence of protein aggregates, we have run Spike samples in a Xbridge BEH450, an SE-HPLC column suitable for the analysis of high-molecular weight proteins (Figure 5B). Unexpectedly, the Spike peak was resolved into two separated peaks eluting at 9.7 and 10.8 min, that corresponded to apparent molecular weight of approximately 1200 kDa and 600 kDa, respectively. This particular elution profile was present in all Spike samples, produced in HEK293 or Expi293F™ cells, in shake-flask, stirred tank bioreactor or wave bioreactor, although the relative contribution of each Spike peak changed (data not shown). A small peak of putative aggregates was also detected in all Spike samples at retention time of approximately 8.5 min (Figure 5B).

Analysis of RBD by size exclusion HPLC validated the observations performed during RBD purification runs. Interestingly, RBD production in HEK293-E6 cells resulted in a mixture between RBD monomer and dimer, with up to 45 % of dimer, whereas in Expi293F™ production run, 93 % of RBD is in monomeric state (Figure 5C).

Spike and RBD thermal stability

The ultimate goal of this work is to develop scalable bioprocess to produce and supply high-quality antigens to prepare SARS-CoV-2 ELISA serology assays. Considering that automation will be needed for ELISA plates preparation and that its global distribution will be required, it is critical to evaluate Spike and RBD thermal stability, namely during storage and when subjected to freeze-thaw cycles. We have analyzed Spike and RBD thermal denaturation by DSF (Figure 6 and Suppl. Table 1). Spike presented a melting temperature of 44.7 °C, whereas RBD shows slightly improved thermal stability, with a melting temperature of 49.4 °C. Additionally, RBD produced in Expi293F™, being mainly in monomeric state, presents slightly increased melting temperature (Suppl. Table 1). All Spike samples analyzed, independently of cell host or cell culture process, presented the same melting temperature, as assessed by nanoDSF (Suppl. Table 1).

Size exclusion HPLC was used to investigate the impact of temperature during storage on Spike and RBD conformation. Antigens obtained from STB runs (HEK293-E6 cells), were incubated at 4 °C or at room-temperature (RT, 20 to 22 °C), up to 14 days or 24 hours. Analysis of Spike and RBD during storage at - 80 °C was also performed for up to 58 days.

For Spike, we observed conversion between SEC species when incubated at 4 °C or RT (Figure 7A). After 24 or 6 hours at 4 °C or RT, respectively, only one peak was detected at approximately 9.7 min. During storage at - 80 °C, the proportion between SEC species is comparable for up to 58 days. Putative Spike aggregates were also detected in all conditions, and accounted for up to 10 % of total SEC species during incubation of Spike at 4 °C or during storage at – 80 °C, or up to 20 % after 6 hours of incubation at RT (data not shown). The conversion between SEC species described for Spike protein was also observed for Spike produced in Expi293F™ cells, when incubated at 4 °C (data not shown).

Analysis of RBD samples after incubation at 4 °C, RT or during storage at – 80 °C, showed that the monomer/dimer moieties were maintained in all conditions (Figure 7B). Likewise, the monomeric portion of RBD, when produced in Expi293F™ cells, was maintained up to 14 days at 4 °C (data not shown).

Importantly, our results show that performing 3 freeze-thaw cycles to either Spike or RBD did not result in significant changes in the antigens conformation or oligomeric state (Figure 7).

Moreover, no protein degradation was detected by SDS-PAGE when samples of Spike and RBD were analyzed after 14 days at 4 °C or, only for Spike, after 24 hours incubation at RT (data not shown).

Dynamic light scattering analysis was also performed to further investigate the impact of temperature in Spike and RBD proteins (Table 3). The polydispersity index (PDI) determined for Spike samples ranged from 13 to 20 % and the estimated molecular weight was approximately 600 - 700 kDa, a value in agreement with the expected size of glycosylated Spike trimer. No major differences in PDI or molecular weight were observed for the several Spike sample treatments or for the different Spike tested (HEK293-E6 stirred tank bioreactor, Expi293F™ cells or wave bioreactor). Most of the RBD results presented in the table are relative to the protein obtained from stirred tank bioreactor, which consist of a quasi-equimolar mixture of RBD monomer and dimer. These samples presented PDI between 24 and 35 % and estimated molecular weight around 70 or 120 kDa, higher than the expected molecular weight of either RBD monomer or dimer. When RBD was subjected to three freeze-thaw cycles, two particles of different sizes were detected (Table 3). Lower PDI of 11 - 18 % and lower estimated molecular weight around 40 kDa were determined for RBD monomer samples, obtained from Expi293F™ or HEK293-E6 productions.

The observation that Spike undergoes putative conformational changes when incubated at 4 °C or RT (Figure 7), raises the question whether these changes impact the performance of Spike as antigen for the detection of SARS-CoV-2 specific antibodies. Therefore, we performed ELISA assays using as plate coating Spike and RBD samples after incubation at 4 °C, RT or after freeze-thaw cycles. The results show that the serum reactivity to both antigens, as well as the assay specificity, was maintained despite of the treatments performed (Figure 8, Suppl. Figure 3). Furthermore, the data reveals good reproducibility for Spike samples coated at different antigen concentrations, whereas for RBD, coating at low concentration of 0.125 µg/ml increases the ELISA serology test variability.

Discussion

COVID-19 pandemic triggered a combined and global effort of the scientific community to understand the new SARS-CoV-2 virus and its key players during infection. Thus, a worldwide demand for SARS-CoV-2 proteins was launched. Spike protein, being responsible for the binding to the ACE2 receptor, and consequently for the virus entry in the host cells, has been intensively studied and is one of the most clinically relevant SARS-CoV-2 proteins. In this work, we have studied the production of full-length Spike and its receptor binding domain (RBD), with the intention of using the two Spike protein formats as antigens in SARS-CoV-2 serological assays.

We have used two human derived cell lines, HEK293-E6 and Expi293F™ cells, for Spike and RBD production and we also studied different cell culturing modes and antigen production scale-up (Figure 1). We showed that, for both Spike and RBD productions, the yields obtained with HEK293-E6 cells, using regular transfection reagents and protocols (Durocher, 2002), were in the same order of magnitude of those obtained with highly improved Expi293F™ expression system (Figure 2A, Amanat et al., 2020; Esposito et al., 2020; Herrera et al., 2020). Additionally, we shown that equivalent final Spike production yields could be obtained for the three different cell culturing strategies tested: shake-flask, stirred tank bioreactors and wave bioreactors, with productions scales ranging from 1 L up to 30 L. By analyzing the media metabolites through the culture time (data not shown) we can anticipate that Spike and RBD production yields can be further improved in HEK293-E6 cultures by the implementation of feeding strategies consisting mainly in glucose re-feeds (Stuible et al., 2020; Pham et al., 2005; Jäger et al., 2015). Performing such feed would be of particular importance in stirred tank bioreactor cultures, since we observed early glucose depletion, as a result of improved cell growth (data not shown).

Aiming at decreasing the antigens downstream processing time and reduce protein losses with unnecessary purification steps we evaluated the impact of process adjustments on final product purity and ultimately on the antigen performance in SARS-CoV-2 specific ELISA assays. The polishing step by size exclusion chromatography could be replaced by dialysis or desalting steps without impacting overall antigen quality, the same conclusion was recently published by other groups (Esposito et al., 2020; Herrera et al., 2020; Stadlbauer et al., 2020). Moreover, the protocol established for product elution from affinity chromatography allowed for the achievement of high pure products, with only one purification step (Figure 2B and Figure 5). Additionally, Spike and RBD purity, oligomeric state, thermal stability and performance on ELISA assays, were similar for the different lots produced (Figure 2B and C, Table 3, Suppl. Table 1). The only exception being the Spike produced in the first run of wave bioreactor, that presented slightly lower reactivity to SARS-CoV-2 positive serum as assessed by ELISA (Figure 2C). This lower performance on ELISA test is most likely the result of Spike degradation, as observed in SDS-PAGE (Figure 2B) and this Spike protein degradation was due to extended downstream processing times. Therefore, we implemented changes in the DSP of large-scale production batches that would, not only reduce processing time, as also allow for further scaling-up of antigen production process in an industrial setup. On the subsequent Spike production using wave bioreactor, we bypassed large scale culture centrifugation by using high capacity filters to clarify culture bulk, and the final size exclusion chromatography step was replaced by sample dialysis. The overall processing time was drastically reduced and analysis of resulting Spike lot suggested lower degree of protein degradation and undistinguishable SARS-CoV-2

positive serum reactivity to Spike, as compared to Spike from smaller production scales (Figure 2B and 2C).

Transient gene expression using mammalian cells has been established for decades (Durocher, 2002; Baldi et al., 2012; Nettleship et al., 2010; Pham et al., 2006). Several approaches have been described for the improvement of transient protein expression, in particular for difficult-to express proteins (Estes et al., 2015; Lin et al., 2015; Mason et al., 2014; Simone et al., 2003), namely decreasing both the expression temperature and coding DNA amount. We obtained a significant improvement of Spike expression in Expi293F™ cells when the cell culture temperature was reduced to 32 ˚C (Figure 3), also in agreement with recently published data for SARS-CoV-2 Spike protein production in the same cells (Esposito et al., 2020; Herrera et al., 2020). Additionally, we show that combining temperature decrease and reduction of coding DNA concentration to 50 % further improves Spike production in Expi293F™ cells. Surprisingly, in HEK293-E6 cells, decreasing of temperature to 32 ˚C had only marginal effects on Spike expression (Figure 3) suggesting that other parameters may be limiting Spike expression. One can speculate that the lack of a carbon source, as discussed above, might be one the reason that can be overcome by implementing fed-batch strategies combined with a fine tune of temperature and DNA shifts. In fact, recent studies report high titers of Spike expression when using CHO cells and optimized production protocols that combine glucose feedings, hypothermia and modulation of coding DNA amounts, among other parameters (Figure 2A, Johari et al., 2020; Stuible et al., 2020). However, to the best of our knowledge, the impact of these optimization production strategies in the antigens quality, namely on post-translational modification, was not deeply discussed.

The expression screen performed showed that reduction of expression temperature or coding DNA did not improve RBD expression in either Expi293F™ or HEK293-E6 cells. Analysis Spike and RBD thermal stability by differential scanning fluorimetry showed 5 ˚C difference in the calculated protein melting temperatures (Tm), with Spike displaying lower temperature stability with a Tm of 44.7 ˚C (Figure 6). We hypothesize that the improvement of Spike expression at 32 ˚C in Expi293F™ cultures, is due to not only to changes in the cell metabolic state but also to improved Spike stability during expression.

Altogether, the protocols reported herein for Spike and RBD production, that range from shake-flask cultures to large scale bioreactors, can be easily implemented in academic laboratories, in clean rooms in health centers or in large scale industrial setups.

To better understand the impact of cell hosts and bioprocess in antigens quality and performance in ELISA serology tests a deep characterization analysis was performed. One of the main quality attributes of reagents destined to COVID-19 diagnosis and, most importantly, COVID-19 prophylaxis

and therapeutics, is the human-derived post-translational modifications, that should closely mimic the natural host for SARS-CoV-2 replication. Using complementary glycomics approaches with lectin blotting and LC-MS, we showed that Spike and RBD produced in HEK293-E6 and Expi293F™ cells contained high mannose and complex glycans with proximal and peripheral fucose, α2,3- and α2,6-linked sialic acid, LacdiNAc and sLacdiNAc. The N-glycosylation of protein S ectodomain (Watanabe et al., 2020) and S1/S2 subdomains (Shajahan et al., 2020) produced in HEK293 cells has recently been described in the literature, where a large diversity of high mannose, hybrid and complex structures were assigned to individual sites. Here, we further identify LacdiNAc/sialyl-LacdiNAc structures and the type of sialic acid linkage. In agreement, LacdiNAc has already been detected in HEK293 cells (André et al., 2007; Costa et al., 2018) and a possible positive impact of this structure on the pharmacokinetics of recombinant proteins has been advanced (Chin et al., 2019). Furthermore, α2,3/6-linked sialic acid has been detected in HEK293 cells N-glycans (Costa et al., 2018); sialylation has a recognized impact in the *in vivo* half-life of therapeutic proteins.

Only few reports have addressed the N-glycosylation properties of Expi293F™ cells (González-Feliciano et al., 2020). Since they are derived from the HEK293 cell line they would be expected to share glycosylation properties as we report here. Curiously, the N-glycosylation of RBD produced Expi293F™ cells showed lower diversity and a tendency towards smaller glycans ($Man_5GlcNAc_2$, paucimannose, hybrid and partially processed glycans) in comparison to RBD produced in HEK293-E6 cells (Table 2). In this context, glycan size is particularly relevant since smaller glycans increase the accessible protein area for antibody development. Indeed, it has been recently reported that paucimannose N-glycans (e.g., $Man_3GlcNAc_2$) shield lower area at the surface of the protein than high mannose $Man_9GlcNAc_2$ or unsialylated proximally fucosylated diantennary glycan (Grant et al., 2020). For protein produced from HEK293 cells an average antibody accessible surface area of 48% has been predicted, which is lower than the shielding for gp120 from HIV (Grant et al., 2020). Additionally, the fact that the RBD from HEK293-E6 analyzed by MS was that produced in stirred tank bioreactor, might have contributed to the high glycosylation heterogeneity found between the different RBD samples tested. Further MS analysis should be performed to infer whether the differences in N-glycosylation correlates with different culturing modes or differences in host cells used.

In contrast to RBD, glycosylation profiles of individual sites in Spike protein showed more resemblances between HEK293-E6 and Expi293F™ cells (Table 1). Thus, further quantitative analysis of individual glycoforms is required to elucidate this issue.

Most sites of Spike protein displayed a high heterogeneity, but in some cases a tendency could be found. For example, N234 displayed predominantly high mannose glycans similarly to the reported

610 before, which could be due to lower accessibility during glycan processing (Watanabe et al., 2020).
611 On the other hand, predominance of more processed glycoforms would be expected in more
612 accessible and loop regions. Glycosylation heterogeneity may be relevant in interactions between
613 glycans and the protein surface, whereas glycosylation at the RBD and adjacent regions is expected
614 to play a role in binding to the ACE2 receptor.

615 HPLC analysis of Spike protein using Xbridge BEH450 column showed two main forms of Spike
616 protein, eluting at 9.7 and 10.8 min, which corresponds to apparent molecular weight of
617 approximately 1200 kDa and 600 kDa, respectively (Figure 5B). This particular Spike elution profile
618 has been recently reported and most likely reflects differences in Spike conformations, rather than
619 oligomerization (Herrera et al., 2020). Moreover, our DLS analysis together with literature data
620 corresponding to the same Spike protein construct in the same buffer conditions, substantiates
621 Spike's trimeric form with approximately 600 kDa (Table 3, Esposito et al., 2020; Herrera et al.,
622 2020). The melting temperature determined for Spike and RBD are in agreement with literature data
623 and suggest low temperature stability for both proteins, in particular for full-length Spike protein
624 (Figure 6, Herrera et al., 2020; Li et al., 2020). So, we set out to study the impact of temperature on
625 Spike and RBD conformation (Figure 7). Interestingly, we show here that temperature disturbs the
626 equilibrium between the two forms of Spike detected. Time-course analysis of Spike protein after
627 storage at 4 ˚C or at room temperature showed interconversion between the two HPLC peaks
628 (Figure 7A). After 24 h or 6 h at 4 ˚C or RT, respectively, the 10.8 min peak was fully converted to the
629 9.7 min peak, which was maintained stable for up to 14 days at 4 ˚C or for 24 h at RT. The observed
630 interconversion between Spike forms after 14 days storage at 4 ˚C has also been described recently
631 (Herrera et al., 2020).

632 Dynamic light scattering analysis is a powerful tool to study protein stability, in particular, protein
633 aggregation in solution. We have used this technique to characterize Spike and RBD samples and to
634 investigate the impact of storage at 4 ˚C on both proteins. The polydispersity index (PDI) obtained
635 for Spike samples is higher than that typically obtained for monodispersed proteins (usually below
636 10%) (Table 3). This is not surprising taking into consideration the high degree of glycosylation
637 detected for Spike samples (Table 1). Similar DLS results on polydispersity index and estimated
638 molecular weight were obtained for either different Spike samples or after sample treatment
639 (storage at 4 ˚C and performing three cycles for freeze-thaw). This suggested that sample treatment
640 did not impacted Spike protein. Likewise, storage of RBD at 4 ˚C or performing freeze-thaw cycles,
641 did not caused changes in the sample polydispersity nor in estimated molecular weight, as assessed
642 by DLS (Table 3). Nevertheless, analysis of different RBD samples, constituted of different monomer/
643 dimer proportions, showed differences in PDI and estimated molecular weight. Homogeneous RBD

monomer samples, from either HEK293-E6 or Expi293F™ productions, presented lower PDI and lower estimated molecular weight as compared to RBD lot with equimolar contribution of dimer and monomer (from HEK293-E6 stirred tank bioreactor). Moreover, the PDI values between 24 and 35 % observed for RBD produced in HEK293-E6 stirred tank bioreactor, correlates well with the high degree of glycosylation detected for this sample (Table 2 and Table 3).

Taking into consideration the large-scale production of ELISA tests for the detection of SARS-CoV-2 specific antibodies, most likely with the use of automated equipment to prepare the ELISA plates, together with plate storage and wide-range distribution, it is anticipated that the antigen proteins will be exposed to temperature fluctuations. So, in this work we have evaluated the performance of Spike and RBD samples as antigens in SARS-CoV-2 specific ELISA assays, after incubation at 4 °C or room temperature, or after performing three cycles of freeze-thaw (Figure 8). Our results clearly showed that the reactivity of Spike and RBD samples to SARS-CoV-2 positive serum was maintained despite of the treatments performed. These observations are reassuring in the view that wide-range serological testing of the population is paramount to understand SARS-CoV-2 true infection rate and to adopt social measures that might mitigate the impact of COVID-19 pandemic.

In summary, our work shows that different culturing approaches can be used for scalable production in Expi293F™ or HEK293-E6 cells of SARS-CoV-2 Spike and RBD, without impacting the final production yield nor compromising the performance of ELISA serologic COVID-19 tests. Moreover, our profound analysis of Spike and RBD glycosylation contributes to the advance of current knowledge SARS-CoV-2 Spike glycosylation patterns critical for the design and development of therapeutics to fight COVID-19. Finally, we have performed a thorough characterization of Spike and RBD in terms of thermal stability and we have shown that Spike and RBD performance as antigens in ELISA assays is maintained even after suffering from temperature fluctuations. This data is also relevant to aid industrial set-ups that will use these antigens for production of ELISA serologic tests or to supply R&D pipelines for drug discovery.

It is our believe that the results described herein are relevant for the scientific community, both in academic and industrial settings and contribute to the current knowledge on SARS-CoV-2 proteins supporting the research and clinical programs to find solutions for COVID-19 pandemics.

Acknowledgements

This work was developed within the scope of the Serology4COVID consortium, Portugal. The authors thank Florian Krammer from Icahn School of Medicine at Mount Sinai (NY, USA), for useful discussions and for kindly supply of Spike and RBD plasmids, and patients and healthy volunteers for

providing SARS-CoV-2 positive and negative serum. Ana Barbas, Cristina Peixoto, Pedro Cruz and António Roldão are acknowledged for useful discussions, and Tiago Nunes, Inês Santos and Katia Ribeiro-de-Jesus are acknowledged for their support in some experiments. Mass Spectrometry data presented was obtained by the UniMS – Mass Spectrometry Unit, ITQB/IBET, Oeiras, Portugal. The authors thank the financial support provided by Oeiras City Council, Portugal.

Conflict of Interest

The authors declare no conflict of interest.

References

Amanat F, Stadlbauer D, Strohmeier S, Nguyen THO, Chromikova V, McMahon M, Jiang K, Arunkumar GA, Jurczyszak D, Polanco J, Bermudez-Gonzalez M, Kleiner G, Aydillo T, Miorin L, Fierer DS, Lugo LA, Kojic EM, Stoever J, Liu STH, Cunningham-Rundles C, Felgner PL, Moran T, García-Sastre A, Caplivski D, Cheng AC, Kedzierska K, Vapalahti O, Hepojoki JM, Simon V, Krammer F. 2020b. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**:1033–1036.

André M, Morelle W, Planchon S, Milhiet PE, Rubinstein E, Mollicone R, Chamot-Rooke J, Le Naour F. 2007. Glycosylation status of the membrane protein CD9P-1. *Proteomics* **7**:3880–95. doi: 10.1002/pmic.200700355.

Baldi L, Hacker DL, Meerschman C, Wurm FM. 2012. Protein Expression in Mammalian Cells **801**:13–26. http://link.springer.com/10.1007/978-1-61779-352-3.

Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. 2008. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7**:1650–1659. https://doi.org/10.1021/pr7008252.

Chin CL, Goh JB, Srinivasan H, Liu KI, Gowher A, Shanmugam R, Lim HL, Choo M, Tang WQ, Tan AHM, Nguyen-Khuong T, Tan MH, Ng SK. 2019. A human expression system based on HEK293 for the stable production of recombinant erythropoietin. *Sci. Rep.* **9**:16768. https://doi.org/10.1038/s41598-019-53391-z.

Claudia A, Auge G, Bredeston L, Corapi ES, Craig PO, Leandro A, Elias F, Gorojovsky N, Gudesblat GE, Herrera MG, Consortium AA, Arbeitman CR, Auge G, Blaustein M, Bredeston L, Corapi ES, Craig PO, Cossio LA, Dain L, Alessio CD', Elias F, Fernández NB, Gasulla J, Gorojovsky N, Gudesblat GE, Herrera MG, Ibañez LI, Idrovo T, Randon MI, Kamenetzky L, Nadra AD, Noseda DG, Paván CH, Pavan MF, Pignataro MF, Roman E, Ruberto LAM, Rubinstein N, Santos J, Duarte FV, Zelada AM. 2020. Structural and Functional Comparison of SARS-CoV-2-Spike Receptor Binding Domain Produced in Pichia pastoris and Mammalian Cells. *bioRxiv*.

Costa J, Gatermann M, Nimtz M, Kandzia S, Glatzel M, Conradt HS. 2018. N-Glycosylation of Extracellular Vesicles from HEK-293 and Glioma Cell Lines. *Anal. Chem.* **90**:7871–7879. https://doi.org/10.1021/acs.analchem.7b05455.

Durocher Y. 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res.* **30**:9e – 9.

Escrevente C, Keller S, Altevogt P, Costa J. 2011. Interaction and uptake of exosomes by ovarian cancer cells. *BMC Cancer* **11**:108. https://doi.org/10.1186/1471-2407-11-108.

720 Esposito D, Mehalko J, Drew M, Snead K, Wall V, Taylor T, Frank P, Denson JP, Hong M, Gulten G,
721 Sadtler K, Messing S, Gillette W. 2020. Optimizing high-yield production of SARS-CoV-2 soluble
722 spike trimers for serology assays. *Protein Expr. Purif.* **174**:105686.
723 https://doi.org/10.1016/j.pep.2020.105686.

724 Estes B, Hsu YR, Tam LT, Sheng J, Stevens J, Haldankar R. 2015. Uncovering methods for the
725 prevention of protein aggregation and improvement of product quality in a transient
726 expression system. *Biotechnol. Prog.* **31**:258–67. doi: 10.1002/btpr.2021.

727 González-Feliciano JA, Akamine P, Capó-Vélez CM, Delgado-Vélez M, Dussupt V, Krebs SJ, Wojna V,
728 Polonis VR, Baerga-Ortiz A, Lasalde-Dominicci JA. 2020. A recombinant gp145 Env glycoprotein
729 from HIV-1 expressed in two different cell lines: Effects on glycosylation and antigenicity. *PLoS*
730 *One* **15**:e0231679. doi: 10.1371/journal.pone.0231679.

731 Grant OC, Montgomery D, Ito K, Woods RJ. 2020. Analysis of the SARS-CoV-2 spike protein glycan
732 shield: implications for immune recognition. *bioRxiv*. doi: 10.1101/2020.04.07.030445.

733 Henderson R, Edwards RJ, Mansouri K, Janowska K, Stalls V, Kopp M, Haynes BF, Acharya P. 2020.
734 Glycans on the SARS-CoV-2 Spike Control the Receptor Binding Domain Conformation. *bioRxiv*.
735 doi: 10.1101/2020.06.26.173765.

736 Herrera NG, Morano NC, Celikgil A, Georgiev GI, Malonis RJ, Lee JH, Tong K, Vergnolle O, Massimi AB,
737 Yen LY, Noble AJ, Kopylov M, Bonanno JB, Garrett-Thomson SC, Hayes DB, Bortz RH,
738 Wirchniaski AS, Florez C, Laudermilch E, Haslwanter D, Fels JM, Dieterle ME, Jangra RK, Barnhill
739 J, Mengotto A, Kimmel D, Daily JP, Pirofski L, Chandran K, Brenowitz M, Garforth SJ, Eng ET, Lai
740 JR, Almo SC. 2020. Characterization of the SARS-CoV-2 S Protein: Biophysical, Biochemical,
741 Structural, and Antigenic Analysis. *bioRxiv*. doi: 10.1101/2020.06.14.150607.

742 Johari YB, Jaffé SR, Scarrott JM, Johnson AO, Mozzanino T, Pohle TH, Maisuria S, Bhayat-Cammack A,
743 Brown AJ, Lan Tee K, Jackson PJ, Seng Wong T, Dickman MJ, Sargur R, James DC. 2020.
744 Production of Trimeric SARS-CoV-2 Spike Protein by CHO Cells for Serological COVID-19 Testing.
745 *medRxiv*.

746 Ju B, Zhang Q, Ge X, Wang R, Yu J, Shan S, Zhou B, Song S, Tang X, Yu J, Ge J, Lan J, Yuan J, Wang H,
747 Zhao J, Zhang S, Wang Y, Shi X, Liu L, Wang X, Zhang Z, Zhang L. 2020. Human neutralizing
748 antibodies elicited by SARS-CoV-2 infection. *Nature* **584**:115–119.
749 https://doi.org/10.1038/s41586-020-2380-z.

750 Li T, Zheng Q, Yu H, Wu D, Xue W, Zhang Y, Huang X, Zhou L, Zhang Z, Zha Z, Chen T, Wang Z, Chen J,
751 Sun H, Deng T, Wang Y, Chen Y, Zhao Q, Zhang J, Gu Y, Li S, Xia N. 2020. Characterization of the
752 SARS-CoV-2 Spike in an Early Prefusion Conformation. *bioRxiv*. doi:
753 10.1101/2020.03.16.994152.

754 Lin CY, Huang Z, Wen W, Wu A, Wang C, Niu L. 2015. Enhancing protein expression in HEK-293 cells
755 by lowering culture temperature. *PLoS One* **10**:e0123562.
756 https://doi.org/10.1371/journal.pone.0123562.

757 Machado E, Kandzia S, Carilho R, Altevogt P, Conradt HS, Costa J. 2011. N-Glycosylation of total
758 cellular glycoproteins from the human ovarian carcinoma SKOV3 cell line and of recombinantly
759 expressed human erythropoietin. *Glycobiology* **21**:376–86. doi: 10.1093/glycob/cwq170.

760 Mason M, Sweeney B, Cain K, Stephens P, Sharfstein S. 2014. Reduced Culture Temperature
761 Differentially Affects Expression and Biophysical Properties of Monoclonal Antibody Variants.
762 *Antibodies* **3**:253–271.

763 Nettleship JE, Assenberg R, Diprose JM, Rahman-Huq N, Owens RJ. 2010. Recent advances in the
764 production of proteins in insect and mammalian cells for structural biology. *J. Struct. Biol.*
765 **172**:55–65. doi: 10.1016/j.jsb.2010.02.006.

766 Okba NMA, Muller MA, Li W, Wang C, GeurtsvanKessel CH, Corman VM, Lamers MM, Sikkema RS,
767 Bruin E de, Chandler FD, Yazdanpanah Y, Hingrat Q Le, Descamps D, Houhou-Fidouh N, Reusken
768 CBEM, Bosch B-J, Drosten C, Koopmans MPG, Haagmans BL. 2020. SARS-CoV-2 specific
769 antibody responses in COVID-19 patients. *medRxiv*.

770 Perera RAPM, Mok CKP, Tsang OTY, Lv H, Ko RLW, Wu NC, Yuan M, Leung WS, Chan JMC, Chik TSH,
771 Choi CYC, Leung K, Chan KH, Chan KCK, Li KC, Wu JT, Wilson IA, Monto AS, Poon LLM, Peiris M.
772 2020. Serological assays for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),
773 March 2020. *Eurosurveillance* **25**:pii=2000421. https://doi.org/10.2807/1560-
774 7917.ES.2020.25.16.2000421.

775 Pham PL, Perret S, Cass B, Carpentier E, St.-Laurent G, Bisson L, Kamen A, Durocher Y. 2005.
776 Transient gene expression in HEK293 cells: Peptone addition posttransfection improves
777 recombinant protein synthesis. *Biotechnol. Bioeng.* **90**:332–44. doi: 10.1002/bit.20428.

778 Pham PL, Kamen A, Durocher Y. 2006. Large-scale transfection of mammalian cells for the fast
779 production of recombinant protein. In: . *Mol. Biotechnol.*, pp. 34(2):225-37. doi:
780 10.1385/MB:34:2:225.

781 Shajahan A, Supekar NT, Gleinich AS, Azadi P. 2020. Deducing the N- and O-glycosylation profile of
782 the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*:cwaa042.
783 https://doi.org/10.1093/glycob/cwaa042.

784 Simone MS, Randolf JK, Gabriele G, Martina K, Friedrich D, Friedrich S. 2003. Higher expression of fab
785 antibody fragments in a CHO cell line at reduced temperature. *Biotechnol. Bioeng.* **84**:433–8.
786 doi: 10.1002/bit.10793.

787 Stadlbauer D, Amanat F, Chromikova V, Jiang K, Strohmeier S, Arunkumar GA, Tan J, Bhavsar D,
788 Capuano C, Kirkpatrick E, Meade P, Brito RN, Teo C, McMahon M, Simon V, Krammer F. 2020.
789 SARS−CoV−2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen
790 Production, and Test Setup. *Curr. Protoc. Microbiol.* **57**:e100. doi: 10.1002/cpmc.100.

791 Stuible M, Gervais C, Lord-Dufour S, Perret S, Abbe DL', Schrag J, St-Laurent G, Durocher Y. 2020.
792 Rapid, high-yield production of full-length SARS-CoV-2 spike ectodomain by transient gene
793 expression in CHO cells. *bioRxiv*.

794 Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. 2020. Structure, Function, and
795 Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**:281-292.e6.
796 https://doi.org/10.1016/j.cell.2020.02.058.

797 Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, Lu G, Qiao C, Hu Y, Yuen KY, Wang Q, Zhou H, Yan J,
798 Qi J. 2020. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*
799 **181**:894-904.e9.

800 Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. 2020. Site-specific glycan analysis of the
801 SARS-CoV-2 spike. *Science (80-.).* **369**:330–333. doi: 10.1126/science.abb9983.

802 Weiss S, Klingler J, Hioe C, Amanat F, Baine I, Kojic EM, Stoever J, Liu S, Jurczyszak D, Bermudez-
803 Gonzalez M, Simon V, Krammer F, Zolla-Pazner S. 2020. A High Through-put Assay for
804 Circulating Antibodies Directed against the S Protein of Severe Acute Respiratory Syndrome
805 Coronavirus 2 (Sars-Cov-2). *J Infect Dis*:jiaa531. doi: 10.1093/infdis/jiaa531.

806 Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. 2020.
807 Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (80-.).*
808 **367**:1260–1263. doi: 10.1126/science.abb2507.

809 Zhou P, Yang X Lou, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J,
810 Luo Y, Guo H, Jiang R Di, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F,

811 Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**:270–273.

813 **Table 1.** Comparison of glycan composition in Spike produced in HEK293-E6 or Expi293F™ cells.

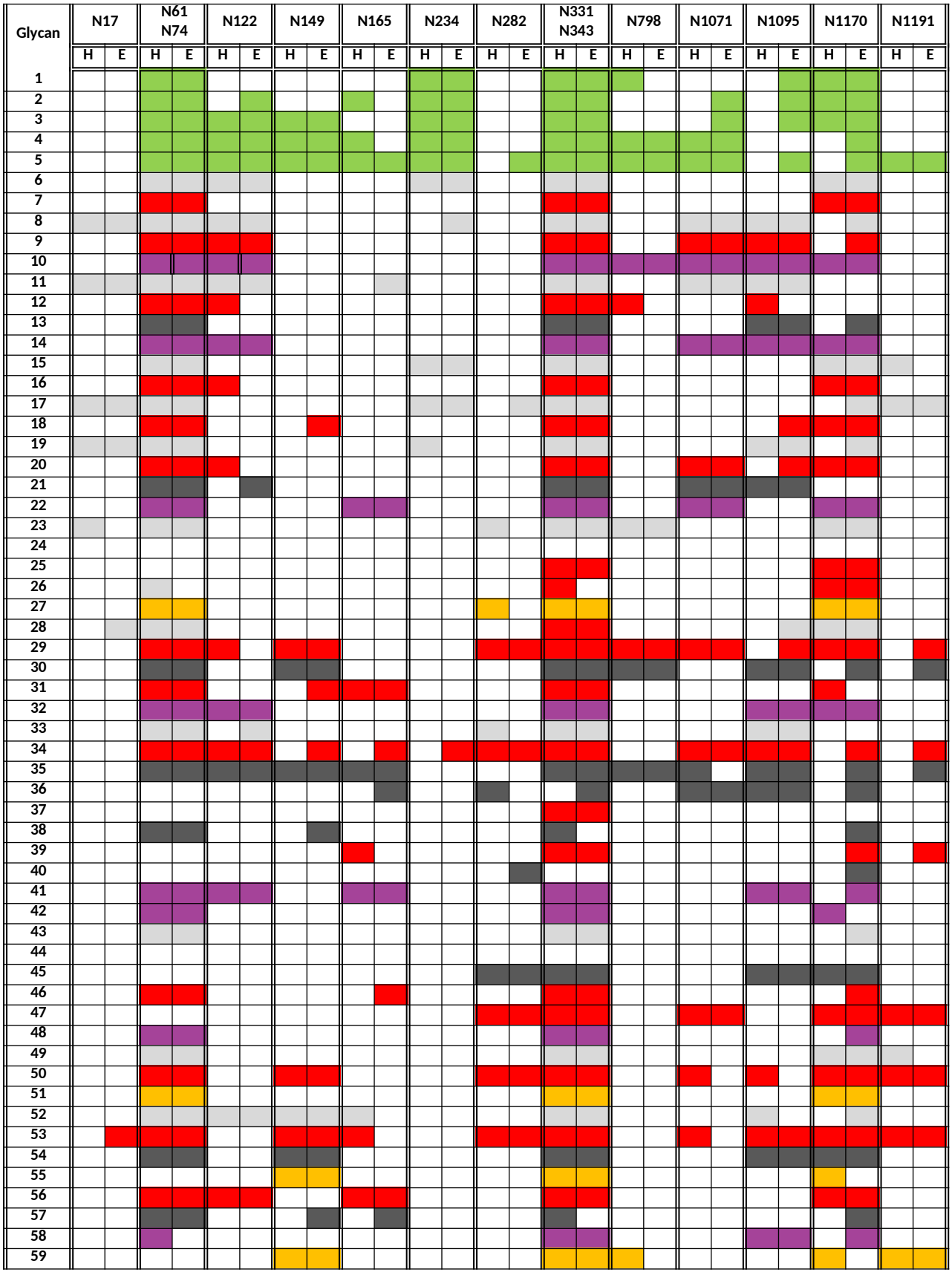
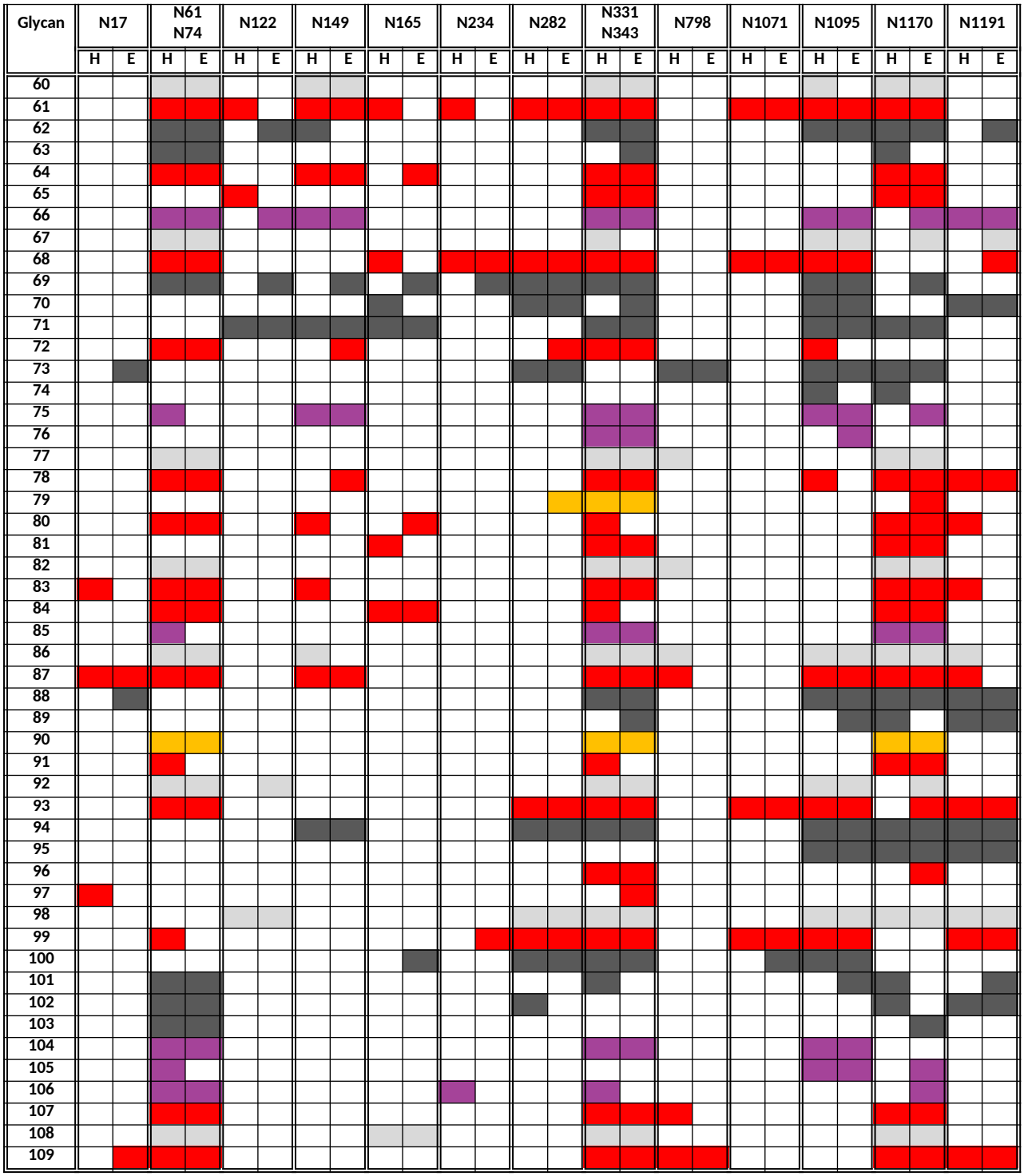
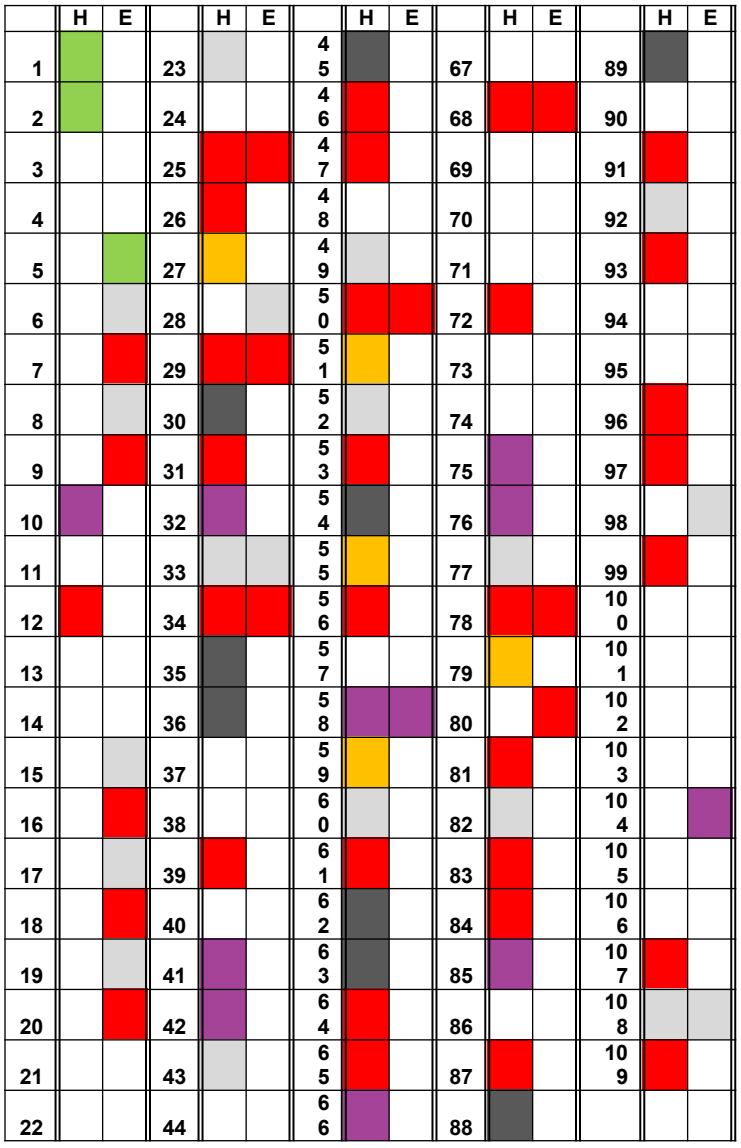


Table 1 (cont.). Comparison of glycan composition in Spike produced in HEK293-E6 or Expi293F™ cells.



Results were obtained with Spike samples from HEK293-E6 shake flask day 3 (H) or Expi293F™ cells shake flask day 3 (E). MS data was screened considering the 109 N-glycans structures previously identified in SARS-CoV-2 Spike (Watanabe et al., 2020, Suppl. Table 2). Colors represent the following N-glycans structural features: High mannose (green), Fucosylated (red), Sialylated (purple), Sialylated/fucosylated (dark grey), Sialyl-LacdiNac (yellow) and others (light grey).

Table 2. Comparison of glycan composition in RBD produced in HEK293-E6 or Expi293F™ cells.



Results were obtained with RBD samples from HEK293-E6 stirred tank bioreactor (H) or Expi293F™ cells shake flask day 3 (E). MS data was screened considering the 109 N-glycans structures previously identified in SARS-CoV-2 Spike (Watanabe et al., 2020, Suppl. Table 2). Colors represent the following N-glycans structural features: High mannose (green), Fucosylated (red), Sialylated (purple), Sialylated/fucosylated (dark grey), Sialyl-LacdiNac (yellow) and others (light grey).

Table 3. Dynamic Light Scattering analysis of Spike and RBD.

Condition		Spike		RBD	
		PDI (%)	MW (KDa)	PDI (%)	MW (KDa)
- 80 °C (HEK293-E6)	D2	14.7	581	23.9	111
	D30	17.1	657	35.2	154
	D44 (monomer)	-	-	10.9	41
	D1 (wave2)	20.4	751	-	-
- 80 °C (Expi293F™)	D1	19.6	695	-	-
	D57	-	-	18.4	48
4 °C (HEK293-E6)	40 h*	13.3	630	32.4	139
	D7	14.1	617	-	-
	D14	15.1	701	26.5	79
	16h* (wave2)	19.2	732	-	-
3x Freeze-thaw (HEK293-E6)		15.8	641	10.5/29.8**	74/158**

Results were obtained with Spike and RBD samples from HEK293-E6 stirred tank bioreactor (h_STB_d4) and Expi293F™ shake flask day 3 (e_SF_d3), RBD monomer from HEK293-E6 shake flask day 5 run (h_SF_d5) and Spike from HEK293-E6 wave bioreactor run2 (h_w2_d4). * Samples analyzed without fast freezing in liquid nitrogen. ** Two species detected by DLS.

Figure legends

Figure 1. Experimental set-up for production, characterization and storage of SARS-CoV-2 Spike and RBD recombinant proteins. Top illustration of SARS-CoV-2 and Spike/RBD were adapted from Centers for Disease Control and Prevention, US (CDC ID#23311, Alissa Eckert, MSMI; Dan Higgins, MAMS, 2020) and Wrapp et al., 2020 (Wrapp et al., 2020), respectively.

Figure 2. SARS-CoV-2 Spike and RBD production. (A) Impact of Harvest Day, Bioreactor Type and Cell Host on Spike and RBD Final Production Yields and Quality, accessed by **(B)** SDS-PAGE, and **(C)** ELISA SARS-CoV-2 positive serum reactivity.

Panel **A** includes a comparison with data recently published in the literature (Amanat et al., 2020; Esposito et al., 2020; Herrera et al., 2020; Li et al., 2020; Claudia et al., 2020; Johari et al., 2020; Stuible et al., 2020). When production scale was not specified, namely for the insect and CHO cell cultures (Amanat et al., 2020; Li et al., 2020; Johari et al., 2020; Stuible et al., 2020) an assumption of < 2 L was made; Error bars represent standard deviation. In Panel **B**, 3 µg of each protein was loaded per lane under reducing conditions. The results shown in panel **C** were obtained in ELISA plates coated with 2, 0.5 and 0.125 µg/ml of Spike or RBD (see M&M for details). ELISA reactivity was assessed by OD 450 nm using SARS-CoV-2 positive serum collected 14 days post PCR diagnostic and diluted 1:1350. The ELISA assay specificity was assessed by analysis of negative serum reactivity to Spike and RBD (Suppl. Figure 1). Error bars represent 20 % error of the ELISA method.

Figure 3. Effect of temperature shift and coding DNA amount on Spike and RBD production. Reducing SDS-PAGE analysis of Expi293F™ or HEK293-E6 culture supernatants expressing Spike and RBD at 32 °C or 37 °C, after transfection with 1 µg coding DNA/mL (100%) or 0.5 µg coding DNA/mL (50%). 20 µl of Spike expressing culture supernatants were loaded in each lane. 5 µl and 10 µl of RBD supernatants from Expi293F™ and HEK293-E6 cultures, respectively, were loaded in each lane. Spike

and RBD expression at 37 °C was analysed from day 3 to day 5 post-transfection in HEK293-E6 cultures or day 2 to day 4 post-transfection in Expi293F™ cultures.

Figure 4. Effect of host cell in Lectin blotting of Spike and RBD. RBD produced in Expi293F™ cells (E) or HEK293-E6 cells (H) and Spike produced in Expi293F™ cells (E) or HEK293-E6 cells (H) have been detected with the lectins GNA, WFA, AAL, MAL and SNA. Each lane contained 200 ng protein. Preferred glycan specificities of lectins are shown and are according to the suppliers and to the literature (Cummings et al., 2017). Glycan representation is according to the Consortium of Functional Glycomics. Results are representative of at least 3 blots.

Figure 5. Effect of host cell in Spike and RBD oligomeric state. Size-exclusion HPLC analysis of Spike (A and B) and RBD (C) purified proteins produced in Expi293F™ and HEK293-E6. Spike samples were analysed with Xbridge BEH200 **(A)** or Xbridge BEH450 **(B)** HPLC columns and RBD samples were analysed in Xbridge BEH125 column **(C)**. Protein standard mixes were injected in each column, under the same conditions, and the arrows indicate the respective retention time. Protein standards used: thyroglobulin Dimer (1320 kDa), thyroglobulin (660 kDa), IgG (150 kDa), ovalbumin (44.2 kDa) and ribonuclease A (13.7 kDa).

Figure 6. Spike and RBD thermal stability. Differential Scanning Fluorimetry analysis of Spike from HEK293-E6 stirred tank bioreactor (h_STB_d4) and RBD monomer sample from HEK293-E6 shake flask day 5 (h_SF_d5 monomer). The average melting temperature was determined considering the two replicate measurements shown in the figure.

Figure 7. Impact of storage temperature on Spike and RBD conformation. Size-exclusion HPLC analysis of Spike protein using Xbridge BEH450 **(A)** or RBD protein using Xbridge BEH 125 **(B).** Results were obtained with the Spike and RBD produced in HEK293-E6 stirred tank bioreactor (h_STB_day4). Protein samples were stored at 4 °C, room temperature (RT, 20 -22 °C) or at – 80 °C, for several time-points. The starting point of the analysis (0 H) consists in injecting the sample into HPLC column immediately after thawing from – 80 °C. Samples FT3x were subjected to three freeze-thaw cycles using liquid nitrogen. The error is represented in bars.

Figure 8. Effect of storage temperature on SARS-CoV-2 Spike and RBD performance in ELISA serologic tests. Evaluation of SARS-CoV-2 positive serum reactivity to Spike and RBD samples after incubation at 4 ˚C, room temperature (RT, 20-22 ˚C), or after three freeze-thaw cycles using liquid nitrogen (FT3x). Results were obtained with Spike and RBD produced in HEK293-E6 stirred tank bioreactor (h_STB_d4). Control condition corresponds to coating the antigen after thawing from – 80 ˚C. Serial dilutions of positive serum and antigen were analysed; the data presented corresponds to SARS-CoV-2 positive serum collected 25 days post PCR diagnostic diluted 1:1000 and Spike or RBD coating with 2, 0.5 and 0.125 µg/ml. The assay specificity was assessed by analysis of negative serum reactivity to Spike and RBD samples (Suppl. Figure 3). Error bars represent 20 % error of the ELISA method.

939

940

941

942