Native hydrophobic interaction chromatography hyphenated to multi-angle light scattering detection for in-process control of SARS-CoV-2 nucleocapsid protein produced in *Escherichia coli*

Jelle De Vos[1,2], Patricia Pereira Aguilar[3*], Christoph Köppl[3], Andreas Fischer[3], Clemens Grünwald-Gruber[4], Mark Dürkop[2,5], Miriam Klausberger[2], Juergen Mairhofer[6], Gerald Striedner[2,3,6], Monika Cserjan-Puschmann[2,3], Alois Jungbauer[2,3], Nico Lingg[2,3]

[1]Vrije Universiteit Brussel, Department of Chemical Engineering, 1050 Brussels, Belgium

[2]Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna (BOKU), 1190 Vienna, Austria

[3]acib - Austrian Centre of Industrial Biotechnology, 1190 Vienna, Austria

[4]BOKU Core Facility Mass Spectrometry, University of Natural Resources and Life Sciences, Vienna (BOKU), 1190 Vienna, Austria

[5]Novasign GmbH, 1190 Vienna, Austria

[6]enGenes Biotech GmbH, 1190 Vienna, Austria

(*) Corresponding author

E-mail: patricia.aguilar@boku.ac.at

Abstract

The nucleocapsid protein (NP) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is critical for several steps of the viral life cycle, and is abundantly expressed during infection, making it an ideal diagnostic target protein. This protein has a strong tendency to dimerization and interaction with nucleic acids. A native hydrophobic interaction chromatography hyphenated to multi-angle light scattering detection (HIC-MALS) method was established for in-process control, in particular, to monitor product fragmentation and multimerization throughout the purification process. High titers of the nucleocapsid protein were expressed in *E. coli* with a CASPON tag, using a growth-decoupled protein expression system. Purification was accomplished by nuclease treatment of the cell homogenate and a sequence of chromatographic steps. 730 mg purified NP per liter of fermentation could be produced by the optimized process, corresponding to a yield of 77%. The HIC-MALS method was used to demonstrate that the NP product can be produced with a purity of 95%. The molecular mass of the main NP fraction is consistent with dimerized protein as was verified by a complementary native size-exclusion separation (SEC)-MALS analysis. Peptide mapping mass spectrometry and host cell specific enzyme-linked immunosorbent assay confirmed the high product purity, and the presence of a minor endogenous chaperone explained the residual impurities. The HIC-MALS method enables to monitor the purity of the product and simultaneously access its molecular mass.

Keywords: 2019 novel coronavirus; SARS-CoV-2; nucleoprotein; hydrophobic interaction chromatography; SEC-MALS; fusion-protein; downstream processing; CASPON technology

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid protein (NP) is an attractive diagnostic marker to coronavirus disease 2019 (COVID-19)[1–3] since it exhibits a lower mutation rate and is abundantly expressed during infection[4–7]. Vaccines against SARS-CoV-2 are almost exclusively based on the virus spike protein[8]. Hence, diagnostic markers based on NP will be essential to discriminate between infection- and vaccination-conferred immunity[9]. This protein has a strong tendency for dimerization and interaction with DNA when expressed in a bacterial host[10]. Moreover, coronaviral nucleoproteins are known to undergo concentration-dependent multimerization[3,11,12] and fragmentation[10]. In fact, most of the preparations of coronaviral nucleocapsid proteins are dealing with either the N-terminal or the C-terminal domain separately[13,14]. When full-length NP is produced, the sizes usually correspond to various multimerization states[12,13] with wide size distributions, suggesting a heterogeneous product[12,13,15]. Therefore, a fast in-process control method is needed to monitor product purity and molecular mass, especially during downstream processing (DSP) development. While polyacrylamide gel electrophoresis is commonly used for monitoring product quality, it suffers from a limited linear dynamic range, making it unsuitable for quantitative protein analysis[16]. Native high-performance liquid chromatography methods, such as hydrophobic interaction chromatography (HIC) hold promise for a rapid screening method with high resolving power[17–19]. Unfortunately, traditional HIC analysis with UV detection does not allow for monitoring the molecular mass of the separated analytes. A powerful analytical method to characterize protein mass in native conditions is to use multi-angle light scattering (MALS) detection, as was demonstrated already for ion exchange and size exclusion chromatography[20,21]. Combining HIC with MALS allows for separation by

hydrophobicity while simultaneously obtaining information about the molecular mass and relative quantity of the analytes.

Although NP is post-translationally modified by phosphorylation[22,23] and potentially glycosylation[24] in its native host, it can be expressed in *E. coli* as a soluble protein with high titers without compromising antigenicity and diagnostic performance in serological assays[9]. The *E. coli* strains BL21(DE3) and the recently developed enGenes-X-press[25] are all based on T7 expression[26], whereby with the latter recombinant protein synthesis can be decoupled from cell growth to exclusively utilize metabolic resources for synthesis of the protein of interest. Microbial expression is often combined with fusion of the target protein to an affinity tag to simplify downstream process development. A hexa-histidine (6H) tag is commonly used in combination with immobilized metal affinity chromatography (IMAC) because of the wide availability of resins and the easy capture of target proteins, even from crude solutions. These tags can be removed when the native target protein is required, and several enzymes are available to perform this task such as Tobacco etch virus protease, thrombin or the newly developed circularly permuted caspase-2 (cpCasp2)[27,28]. While most proteases are either too unspecific or leave residual amino acids on the N-terminus, cpCasp2 combines high specificity with the ability to fully remove the affinity CASPON tag, regardless of the N-terminus of the target protein. The CASPON tag combines a 6H affinity tag, a cleavage site and a bacteriophage T7 based solubility tag[28,29].

A robust biotechnological production platform, including protein expression and DSP, is essential for the reliable recombinant production of SARS-CoV-2 NP in sufficient quantity and with acceptable product quality. To ensure the absence of host cell-derived proteins that could otherwise lead to erroneous results in antibody tests, additional unit operations are usually performed after capture, such as ion exchange or hydrophobic interaction chromatography. HIC is

a potent purification method to separate highly similar proteins, such as protein variants or fragments[30,31]. This purification step also aids in removal of host cell proteins still present after IMAC[32].

In this paper, we report the successful expression and purification of recombinant SARS-CoV-2 NP from *E. coli*. A growth-decoupled *E. coli* fermentation strategy in which NP was expressed as a fusion-protein allowed for high titers[25]. A dedicated DSP strategy was identified to produce NP with high yield and product purity, using the CASPON technology for fusion-protein production[28,29]. A novel HIC-MALS method was developed to monitor the purity of the NP end product and simultaneously access its molecular mass, allowing to evaluate different DSP strategies. Several complementary analytical techniques, including MS/MS peptide mapping and SEC-MALS, were performed to confirm the qualitative and quantitative data derived from HIC-MALS.

EXPERIMENTAL SECTION

Chemicals and Reagents. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA). The buffers for analytical HPLC-MALS analysis were prepared using HQ-H_2O (18.2 MΩ·cm), filtered through a 0.1 μm filter and degassed prior to use. A variant of T7AC-6H-cpCasp2 was used for the enzymatic tag removal[28].

Gene Construct and Bacterial Strains. The full-length SARS-CoV-2 NP sequence (GenBank: NC_045512.2) from the first human isolate Wuhan-1[33] was fused to either a non-removable C-terminal hexa-histidine tag or a completely removable N-terminal CASPON tag[29] consisting of the

negatively charged T7AC solubility tag[28], a hexa-histidine tag, a short linker (GSG) and the caspase-2 cleavage site (VDVAD) resulting in the sequence MLEDPERNKERKEAELQAQTAEQHHHHHHGSGVDVAD. Additionally, for periplasmic expression the OmpA signal sequence was fused to the CASPON tag. The SARS-CoV-2 NP sequence was amplified via PCR using the qPCR control plasmid (2019-nCoV_N) from Integrated DNA Technologies (Coralville, Iowa, USA). The cDNA encoding the fusion proteins (Supporting Information) were cloned into the pET30a*cer* or pET30a*cer*-CASPON expression vector[25,34] with unique restriction sites. The POI is under the control of a T7 promoter. The resulting expression vectors (pET30a*cer*-CASPON-NP) and (pET30a*cer*-NP-6H) were transformed into chemically competent *E. coli* NEB-5α cells purchased from New England Biolabs (NEB, Ipswich, MA, USA) for screening on kanamycin plates. The positive constructs were sequenced and then transformed into different *E. coli* production strains, BL21(DE3) and two special enGenes-X-press strains (V1 and V2), for growth-decoupled recombinant protein production[25,34]. All cloning procedures were performed using Phusion High-Fidelity DNA Polymerase and restriction enzymes from NEB according to supplier instructions. Primers were purchased from Sigma Aldrich (St. Louis, USA). The following constructs were used: BL21(DE3)(pET30a*cer*-CASPON-NP), BL21(DE3)(pET30a*cer*-ompA-CASPON-NP), enGenes-X-press V2 (pET30a*cer*-CASPON-NP), enGenes-X-press V2 (pET30a*cer*-ompA-CASPON-NP), enGenes-X-press V1 (pET30a*cer*-NP-6H) and enGenes-X-press V2 (pET30a*cer*-NP-6H).

Expression of the SARS-CoV-2 nucleocapsid protein. For selection of the best expression system regarding soluble volumetric titer, shake flask experiments were performed. Cells from research cell banks of the different constructs were inoculated in semi-synthetic medium[35] and

grown on an orbital shaker at 200 rpm at 30 °C. Expression was induced at a cell density of OD_{600} = 1 by addition of 0.5 mM IPTG and additionally 100 mM arabinose for the enGenes-X-press strains.

For lab-scale production of NP, growth-decoupled recombinant protein production was performed according to Stargardt *et al.*[25]. Cells were grown in fed-batch mode in a 1.0 L (0.5 L batch volume, 0.5 L feed) DASGIP® Parallel Bioreactor System (Eppendorf AG, Hamburg, DE) equipped with standard probes (pH, pDO). The pH was maintained at 7.0 ± 0.05, temperature was maintained at 37 ± 0.5 °C during the batch phase and decreased to 30 ± 0.5 °C in the beginning of the feed phase. The dissolved oxygen (O_2) level was stabilized at > 30%. The composition of the media and the pre-culture were described elsewhere[25,35]. All media components were added in relation to the grams of calculated cell dry mass (CDM) to be produced and for calculation the required yield coefficient YX/S of 0.3 g/g specific for BL21(DE3) was used. Feeding was initiated when the culture, grown to 4 g/L CDM in 0.5 L batch medium, entered the stationary phase. The carbon-limited fed-batch regime was divided into three separate phases. An exponential substrate feed providing a constant growth rate of 0.13 h^{-1} was selected for the first 15 h, followed by two linear feed profiles at 0.4 g medium/min and 0.245 g medium/min for four and 15 h respectively, resulting in a final CDM of about 30 g/L in 1.2 L. Induction of NP production was facilitated at feed hour 19 with the addition of 0.1 mM IPTG and 100 mM arabinose. For off-line analysis (OD_{600}, CDM, product), samples were withdrawn from the bioreactor prior to induction and after induction at two-hourly intervals. To describe cell growth, OD_{600} and CDM was determined according to Cserjan-Puschmann *et al.*[36]. For determination of NP titers by SDS-PAGE, purified NP with the concentrations 75 µg/mL, 50 µg/mL and 25 µg/mL were used as standards to generate calibration curves via linear regression.

Purification of SARS-CoV-2 nucleocapsid protein. The cells were harvested by centrifugation (Beckman Avanti JXN-26 with JLA-10.500 rotor, Krefeld, Germany). Following the harvest of the cells, 20 g of cell wet mass was solubilized in 250 mL of lysis buffer (50 mM $NaPO_4$, 450 mM NaCl, pH 7.4). Cell lysis was performed using high-pressure homogenization (Panda PLUS 2000, Gea, Düsseldorf, Germany) for two cycles with a first and second stage pressure of 1000 and 100 bar, respectively. The lysed cells were clarified by centrifugation and filtration (0.22 µm). The centrifugation parameters differed between DSP strategies. Moreover, starting from strategy DSP #2, a nuclease treatment was included. Denarase was acquired from c-LEcta GmbH (Leipzig Germany) and Salt Active Nuclease High Quality was acquired from ArcticZymes Technologies ASA (Tromsø, Norway). Centrifugation settings and nuclease digest parameters can be found in Table S1.

All chromatographic purification steps were performed on an Äkta Pure 25 (Cytiva, Austria). The outlet was monitored at 254, 280 and 320 nm. The compositions of the mobile phases can be found in Table S2. All columns were packed Tricorn columns with 10 mm internal diameter, with varying length depending on the column volume (CV), which can be found in Table S2. The stationary phase resins were acquired from Cytiva, Bio-Works Technologies (Uppsala, Sweden) and Tosoh Corporation (Griesheim, Germany). For the capture step clarified cell lysis supernatant was subsequently loaded on an equilibrated WorkBeads 40 Ni NTA column. The residence time for the whole capture chromatography run was kept constant at 2 minutes. After loading, the column was washed with equilibration buffer for 10 CV to remove weakly bound impurities and CASPON-NP was eluted using a linear gradient to elution buffer in 10 CV. The subsequent unit operations differed for the five DSP strategies.

For DSP #1 the elution fraction of IMAC capture was buffer exchanged to remove imidazole using 15 mL Amicon 10 kDa ultrafiltration/diafiltration (UF/DF) units. This was done to ensure binding of impurities in the subtractive immobilized affinity chromatography (sIMAC) intermediate purification step. A protease digestion was used to remove the affinity fusion-tag and obtain native NP. CASPON-NP with a concentration of 1 mg/mL was incubated with 0.035 mg/mL T7AC-6H-cpCasp2 variant (100:1 molar ratio) for 2 hours at room temperature. The protein solution was subsequently loaded onto the equilibrated sIMAC column at a residence time of 2.5 minutes. Native NP was collected in the flow-through fraction, while remaining CASPON-NP enzyme, previously co-purified host cell proteins and free tag were bound to the column and removed. The sIMAC flow-through fraction was directly loaded to a subtractive anion exchange chromatography (sAEX) column at a residence time of 2 minutes. NP was collected from the flow-through fraction.

For DSP #2, the enzymatic tag removal conditions were changed to 5 mg/mL CASPON-NP with 0.07 mg/mL T7AC-6H-cpCasp2 variant (50:1 molar ratio) incubated over night at 4 °C. The sIMAC step was performed as in DSP #1 and the sAEX polishing step was omitted.

DSP #3 was performed the same as DSP #2, but with an additional cation exchange chromatography (CEX) step performed after sIMAC. The sIMAC flow-through fraction was loaded to an equilibrated SP Sepharose FF column at a residence time of 2 minutes and eluted at a residence time of 5 minutes using a 5 CV linear gradient.

DSP #4 added an additional HIC polishing step to DSP #3. The CEX eluate was conditioned by adding ammonium sulfate to a final concentration of 900 mM. The conditioned CEX eluate was loaded to an equilibrated Butyl Toyopearl 650-M column at a residence time of 3 minutes. NP was eluted using a 10 CV linear gradient at a residence time of 3 minutes.

For DSP #5, the whole process was streamlined and unit operations were performed in a different order. A HIC step was used for removing imidazole after the capture step, instead of UF/DF. For this, the IMAC eluate was conditioned by adding ammonium sulfate to a final concentration of 720 mM while keeping the CASPON-NP concentration under 1.5 mg/mL to avoid precipitation. The conditioned eluate was loaded to an equilibrated Butyl Sepharose HP column at a residence time of 2 minutes. The residence time was increased to 5 minutes during elution for which a linear gradient from 0-60% B in 9 CV followed by 60-100% B in 2 CV was used. The HIC eluate with a CASPON-NP concentration of 5-7 mg/mL was incubated with T7AC-6H-cpCasp2 variant at a 50:1 molar ratio for 3 hours at room temperature. The digest was conditioned to 10 mM imidazole and 150 mM NaCl for the sIMAC polishing step which was performed on a smaller column at a residence time of 2.5 minutes.

The final product fractions of all DSP strategies were buffer exchanged into PBS using UF/DF. For DSP #5 this UF/DF formulation step was performed using tangential flow filtration on an Äkta Flux using a Pellicon 3 Ultracel 10 kDa membrane (Merck Millipore). For calculation of DSP yield, a cell lysis and clarification yield of 28% was estimated.

Quantification of dsDNA, Endotoxin, and host cell protein content. The analytical assays for host cell protein (HCP) determination via ELISA, dsDNA quantification via PicoGreen assay and Endotoxin quantification via recombinant Factor C assay were performed as previously described by Sauer *et al*[37].

Gel electrophoresis. NP samples were qualitatively analyzed by sodium dodecyl sulfate poly(acrylamide) gel electrophoresis (SDS-PAGE) using NuPAGE 4–12% Bis–Tris Precast Gels (Thermo Fisher Scientific). NuPAGE MES running buffer (Thermo Fisher Scientific) was used to prepare the gel and 15 µL of sample was loaded in the appropriate wells. The precast gel was run at 200 V for 45 min. SeeBlue Plus2 (ThermoFisher Scientific) was used as a protein ladder.

Analytical hydrophobic interaction chromatography and multi-angle light scattering detection. All the HIC-HPLC measurements were performed using an Agilent 1260 Infinity II series instrument equipped with a Quat Pump, Multisampler, and VWD detector (set at 280 nm). Multi-angle light scattering data was acquired from a DAWN 8 MALS detector (Wyatt Technologies, Santa Barbara, CA, USA). OpenLab CDS ChemStation edition software (Agilent Technologies, Santa Clara, CA, USA) and ASTRA VII software was used for the data interpretation. The NP samples were analyzed by using a dn/dc value of 0.185 mL/g and a UV extinction coefficient of 0.962 mL/mg·cm as input for the MW calculation.

A MAbPac HIC-10, 4.6 mm i.d. × 250 mm column length (5 µm particles, 1000 Å, 4.15 mL CV) was acquired from Thermo Fisher Scientific to perform analytical HIC-HPLC measurements of the NP samples. The auto-sampler was set at 4°C, 30 µL sample was injected, and the mobile-phase flow rate was set at 1 mL/min. A 10 min linear gradient was applied from 100% 1.2 M ammonium sulfate containing 0.1 M phosphate buffer (pH 7) to 0.1 M phosphate buffer (pH 7), next a 4 min washing step at 100% 0.1 M phosphate buffer was applied to assure that all bound impurities were removed, followed by 10 min column equilibration at gradient starting mobile phase conditions.

Analytical size exclusion chromatography (SEC) and multi-angle light scattering (MALS) - refractive index (RI) detection. A Dionex UltiMate 3000 RSLC system (Thermo Fisher Scientific, Germering, Germany) was used to perform the SEC-HPLC experiments. The instrument consisted of a membrane degasser, a Dionex Ultimate LPG-3400SD pump module, and a WPS-3000 TSL analytical split-loop well plate autosampler with a 100 µL sample loop installed, and a DAD-3000 diode array detector equipped with a 10 µL analytical flow cell. UV detection was carried out at λ = 280, 260 and 254 nm and the data collection rate was set at 50 Hz with a response time of 0.1 s. Viper MP35N fittings (Thermo Fisher Scientific, Germering, Germany) were used to make the fluidic connections. MALS data was acquired using a DAWN HELEOS 18-angle detector (Wyatt Technologies, Santa Barbara, CA, USA), which was coupled to an Optilab refractive index detector (Wyatt Technologies, Santa Barbara, CA, USA). Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, Germering, Germany) and ASTRA V software were used for data collection. The NP samples were analyzed by using a dn/dc value of 0.185 mL/g and a UV extinction coefficient of 0.962 mL/mg·cm as input for the MW calculation.

A Superdex 200 Increase 10/300 GL, 10 mm i.d. × 300 mm column length (GE Healthcare, Uppsala, Sweden) column was used for the SEC experiments. The auto-sampler was set at 4°C, 90 µL of sample were injected, and the mobile-phase flow rate was set at 0.5 mL/min. A 60 min isocratic analysis was performed using 0.1 M phosphate buffer (pH 7) containing 0.3 M sodium chloride as mobile phase.

Peptide mapping using reversed-phase-mass spectrometry (RPLC-MS). Purified NP sample was digested in solution. The proteins were S-alkylated with iodoacetamide and digested with LysC/GluC (Roche/Promega) or Chymotrypsin (Roche). The digested samples were analyzed by

RPLC-MS using an UltiMate 3000 system (Thermo Fisher Scientific, Germering, Germany) hyphenated to a QTOF MS (Bruker maXis 4G, Bruker) equipped with the standard ESI source in positive ion, DDA mode. MS-scans were recorded (range: 150-2200 Da) and the six highest peaks were selected for fragmentation. Instrument calibration was performed using ESI calibration mixture (Agilent).

A BioBasic C18 column, 0.32 mm i.d. × 150 mm (5 µm particles) acquired from Thermo Scientific was used for RPLC-MS analysis. 80 mM ammonium formate buffer was used as the aqueous solvent and 80:20 v/v % acetonitrile: water as B solvent. 2 µg of NP tryptic digest was injected. A linear gradient was applied from 5% B to 40% B in 30 min, followed by a 6 min gradient from 40% B to 95% B to facilitate the elution of large peptides, at a flow rate of 6 µL/min. The analysis files were converted (using Data Analysis, Bruker) to mgf files, which are suitable for performing a MS/MS ion search with MASCOT. The files were searched against database containing the target sequences, HCPs and contaminates.

RESULTS

Expression of SARS-CoV-2 NP and optimization of the downstream processing protocol. An overview of the different steps for developing a DSP strategy based on the CASPON process with stable quality attributes for diagnostic applications can be found in Figure 1A. Different host strain and affinity fusion tag combinations with and without signal peptides for periplasmic or cytoplasmic expression were designed. The N-terminally fused CASPON tag, which contains a 6H-tag, can be removed with a specific protease to generate an authentic N-terminus[28,29], whereas the C-terminally fused 6H-tag is not removable. Evaluation at the level of shake flask expression identified enGenes-X-press V2 [25,34] (pET30a*cer*-CASPON-NP) as the most promising candidate

for high level of soluble NP expression (Figure S1). This candidate was therefore further scaled up to 1L fed-batch cultivations, using the enGenes-X-press process that was optimized by Stargardt *et al.* [25]. The process was run in quadruplicates and showed reproduceable high NP titers (3.7 ± 0.3 g/L; n = 4) and reproducible biomass concentration (33.7 ± 0.9 g/L; n = 4) (Figure 1B).

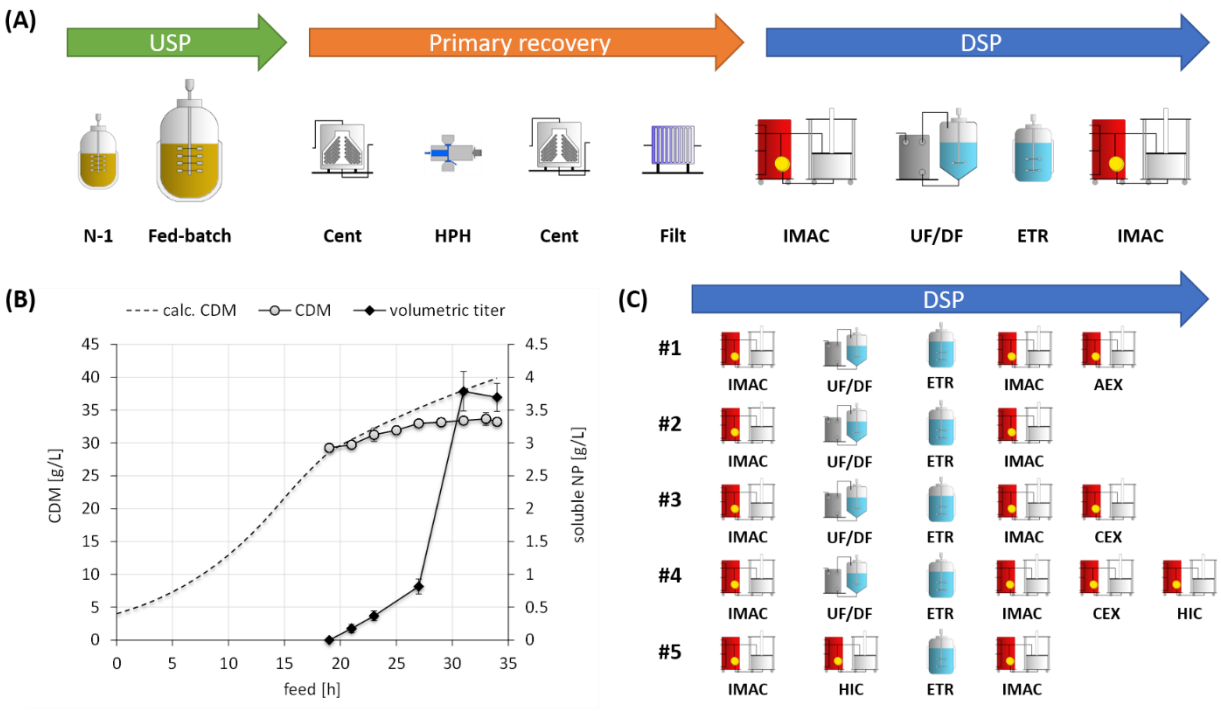


Figure 1: Production of SARS-CoV-2 NP. (A) Schematic overview of the optimized biotechnological production platform based on the CASPON platform process. (B) Process characteristic and product formation kinetics of *E. coli* enGenes-X-press V2 (pET30a*cer*-CASPON-NP) during carbon limited fed-batch bioreactor cultivation. Error bars indicate the standard deviation of four biological replicates (n = 4). (C) DSP variants that were used to purify NP.

Abbreviations used in Figure 1: USP: upstream process; N-1: pre-culture; Cent: centrifugation; HPH: high pressure homogenization; Filt: Filtration; DSP: downstream process; IMAC: immobilized metal affinity chromatography; UF/DF: ultrafiltration/diafiltration; ETR: enzymatic

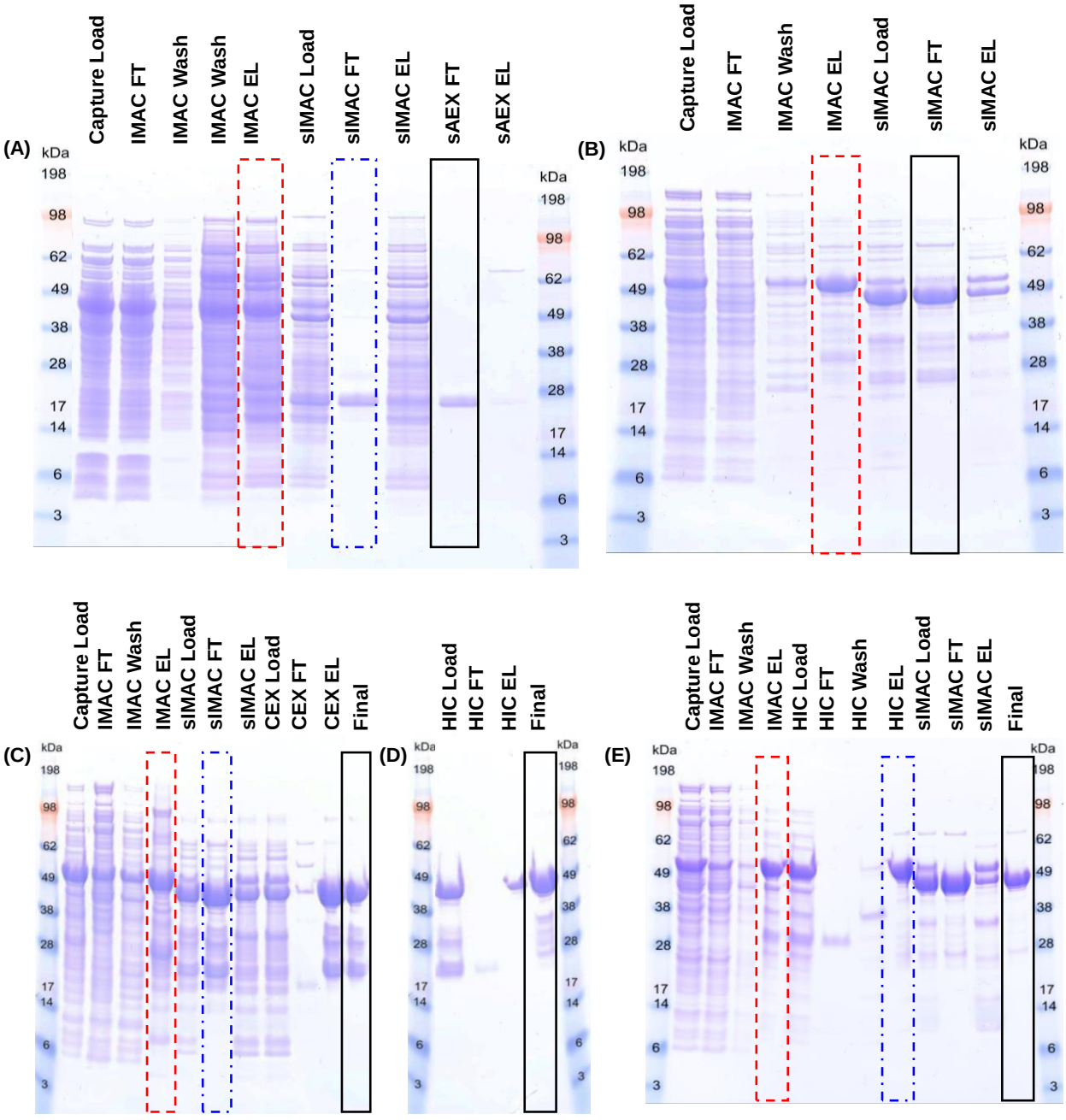
tag removal; AEX: anion exchange chromatography; CEX: cation exchange chromatography; HIC: hydrophobic interaction chromatography.

An overview of the applied chromatography purification steps can be found in Figure 1C. Screening of the process binding and elution conditions, as well as comparing different chromatographic media was performed by SDS-PAGE (Figure 2). Based on the primary sequence, NP product has a theoretical molecular weight of 49.9 kDa before and 45.6 kDa after removal of the CASPON tag. Full-length NP was not detected when using strategy DSP #1, as the final DSP product fraction (black rectangle in Figure 2A) consisted mainly of an NP fragment with a mass of ~25 kDa, which was confirmed by LC-MS/MS after in-gel digestion (data not shown). The process yield of DSP #1 was very low (Table 1) with most of the losses already occurring during capture (Figure 2A, IMAC FT and Wash fractions), where most of NP did not bind. The calculated dynamic binding capacities, even when altering the residence time, were not consistent with a 49.9 kDa protein. It was hypothesized that these losses were due to diffusional hindrance where NP bound to large nucleotides is prevented from entering the pores of the stationary phase.

To reduce the size of the nucleotides available for NP binding the capture load was treated with a nuclease in DSP #2. This reduced the losses during capture chromatography (Figure 2B IMAC FT and Wash fractions), confirming the diffusional hindrance hypothesis. The final product also had a higher ratio of full-length NP to fragment (Figure 2B, sIMAC FT fraction). To further improve the full-length content, several polishing steps were tested. DSP #3 included a CEX polishing step to remove low pI HCPs. This removed proteins >50 kDa (compare Figure 2B sIMAC FT and Figure 2C, Final fraction) but did not remove the fragments with a molecular mass between

20 and 40 kDa. An additional HIC step was used in DSP #4, which resulted in a highly pure final product (Figure 2D, Final fraction), but lowered the process yield (Table 1).

In order to reduce the number of unit operations and increase process yield, HIC purification was used as the intermediate purification step and sIMAC as polishing step in DSP #5 (Figure 2E). The ion exchange steps were omitted, since the nucleotides removed in these steps were of no concern for the intended use of NP as an antigen. This resulted in the purest NP product, providing a process yield of 77 % after cell lysis. The purity of NP regarding dsDNA, HCP and endotoxin was quantified for DSP #5 at different process steps (Table 2). The protein of interest already has a very low HCP concentration after the capture step, which was further reduced during the subsequent processing steps (Table 2). In all five DSP strategies, the CASPON enzymatic tag removal step could be performed in ~3 hours with yields ranging between 70% to 90%.



339

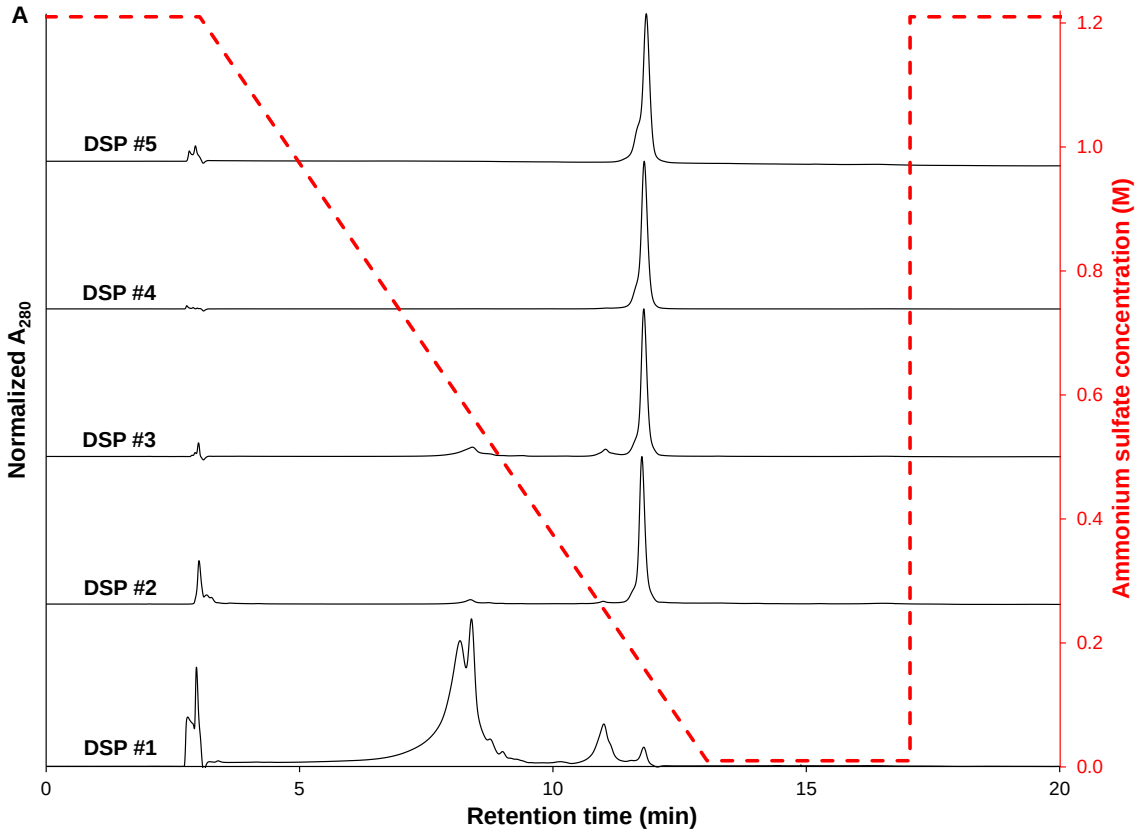
340 **Figure 2.** SDS-PAGE highlighting the influence of the different downstream processing
341 purification strategies on the target NP product: (A) DSP #1, (B) DSP #2, (C) DSP #3, (D) DSP #4
342 and (E) DSP #5. The capture elution fraction is marked with a red dashed rectangle, whereas the
343 product fraction of the intermediate chromatography purification step is denoted with a dash-dotted

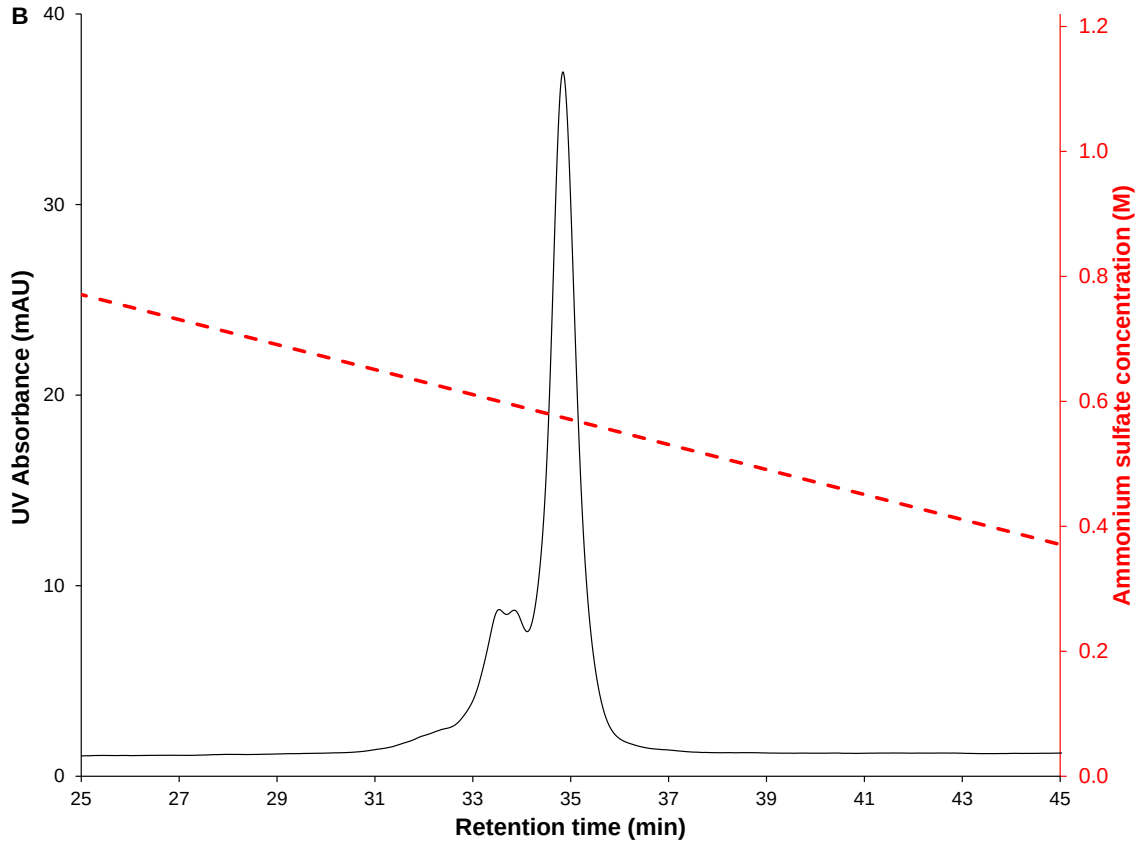
blue rectangle. The SDS-PAGE lane showing the final product is marked with a black solid rectangle.

Quality assessment of the purified NP product using HIC. The process purity could not be assessed by the SDS-PAGE analysis, hence a native analytical HIC-HPLC screening method was established to assess the purity of the different DSP approaches. The salt concentration at the start of the gradient was optimized to separate the impurities eluting between 5-10 min (Figure 3A DSP #1) from the main NP product peak (eluting just before 12 min). The starting concentration of ammonium sulfate was lowered from 2 M to 1.2 M to reduce the total run time. This also improves the resolution of the late-eluting hydrophobic compounds (*i.e.*, the main NP peak), which is a known effect in HIC chromatography[38]. The pump was set at 1 mL/min and a 10 min gradient was selected as the best-compromise condition between sample-throughput and adequate resolution for the purified NP product resulting from the different DSP strategies. The purity of NP product (defined by the main peak) based on the HIC method is reported in Table 1. The recovery of the method was verified by comparing the peak area of an injected NP sample in column bypass with an experiment where the sample is injected on the column using elution buffer. A recovery of 97% for 5 µL injected sample and 98% when injecting 50 µL was observed.

The unbound flow-through fraction (eluting at 3 min) is mainly attributed to oligo- and polynucleotides, based on the high A_{254}/A_{280} ratio at the peak-maximum of 1.5 for DSP #5. The main NP peak in contrast has an A_{254}/A_{280} ratio of 0.4, indicative for proteins. The lower-molecular mass species around 28 kDa, that appear in the fractions marked by the black box in the SDS-PAGE (Figure 2A-C), are also observable in the chromatograms for DSP #1, DSP #2, and DSP #3. A very pure NP product peak (purity >70%) is observed in all DSP protocols except for DSP #1.

The same product quality could be obtained when comparing DSP #4 with DSP #5, even though the latter is simpler. The NP peak features a small shoulder of co-eluting species that comprises about 20% of the peak area. When extending the gradient time to 60 min, the co-eluting analytes could be better separated from the main target peak (Figure 3B). The fact that this high-resolution gradient separation could still not fully resolve these closely eluting species indicates that the analytes exhibit very similar binding characteristics and hence also similar hydrophobicity. The shoulder seen in Figure 3B could consist of a hetero-dimer of full-length NP with a fragment of NP. Putative fragments are likely to originate from fragmentation processes in the serine/arginine-rich region between N-terminal and C-terminal domain (NTD and CTD, respectively)[10]. The NTD of NP has an aliphatic index of 45.8 to 50.2, depending on length of the fragment (185 to 203 amino acids), whereas the CTD of NP has an aliphatic index between 54.4 and 58.9 (216 to 234 amino acids). This is very similar to the aliphatic index of full-length NP (52.53), which can explain the similar HIC retention behavior.

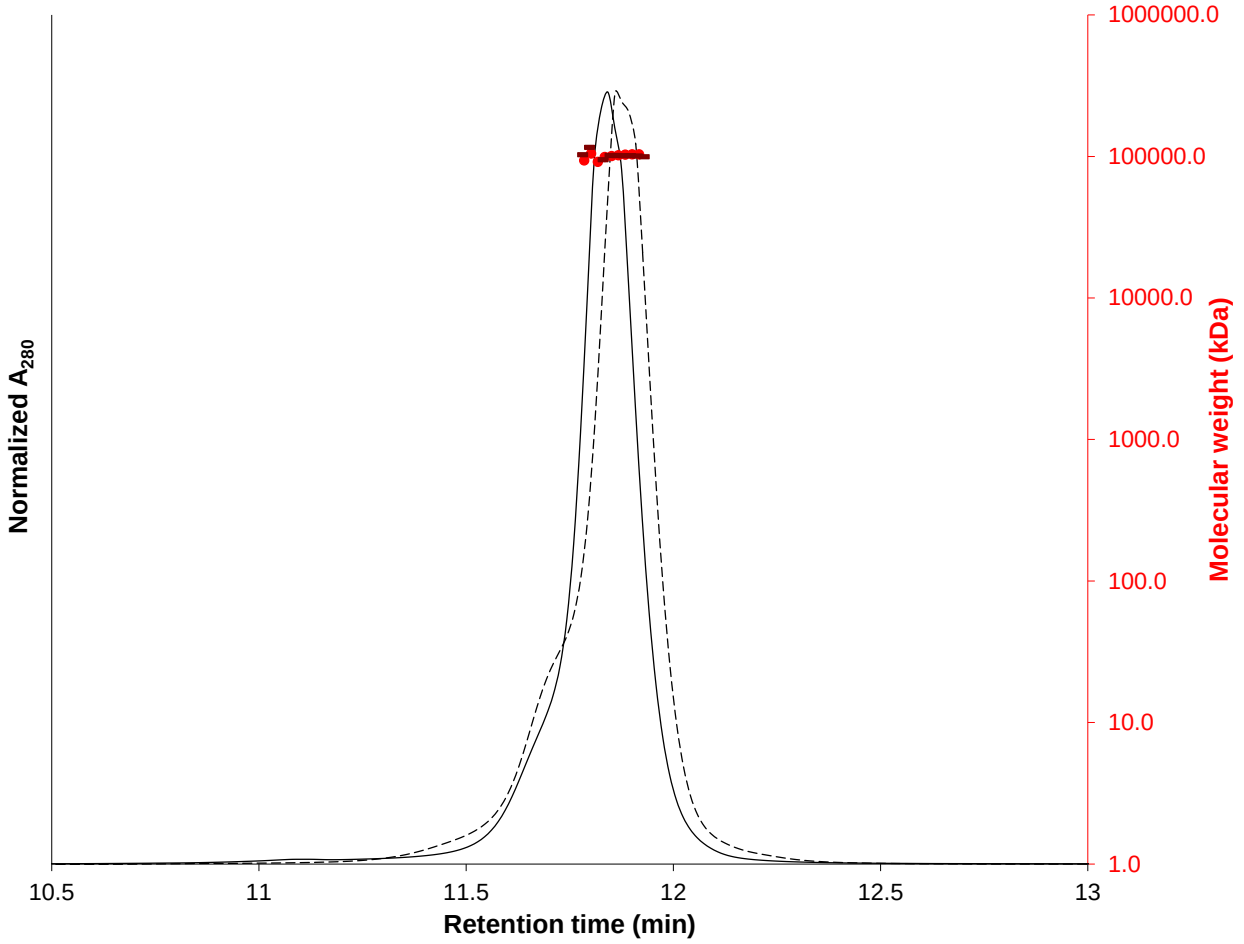

A
DSP #5
DSP #4
DSP #3
DSP #2
DSP #1
Normalized A280
Ammonium sulfate concentration (M)
Retention time (min)
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0
1.2


B
UV Absorbance (mAU)
Ammonium sulfate concentration (M)
Retention time (min)
0
10
20
30
40
25
27
29
31
33
35
37
39
41
43
45
0.0
0.2
0.4
0.6
0.8
1.0
1.2

382

Figure 3. (A) HIC analysis of the final product from the five different DSP strategies, revealing the differences in purity resulting from the purification process. The chromatograms are normalized to the highest peak in the run. The red dotted line shows the applied gradient profile, and the ammonium sulfate concentration is shown on the secondary vertical axis to the right. (B) High-resolution HIC separation of DSP #5 (recorded at 280 nm) using a 60 min gradient time, revealing that species with a similar hydrophobicity are co-eluting with the main peak associated with NP.

In order to explain the identity of the shoulder and the main peak, the HIC method was coupled to an on-line multi-angle light scattering (MALS) detector. Bovine Serum Albumin (BSA) was used as a test protein to optimize the HIC-MALS set-up and to normalize the MALS detectors for accurate molecular mass calculation (Figure S2). NP product from DSP #4 and DSP #5 was analyzed using HIC-MALS, and the main peak has a weight-average molar mass of 100.0 and 99.2 kDa, respectively (Figure 4). This value is consistent with the molar mass of two monomeric NP entities and therefore we assume that mainly dimerized NP is present in the final purified sample solution.



398

399 **Figure 4.** HIC-MALS data of the NP product from DSP #4 (solid black line, solid red circles) and
400 DSP #5 (dashed black line, dark red dashes). Both product purity as MW information can be
401 directly attained from the analysis.

402

403 **Native SEC-MALS-RI of the NP product.** In order to verify the newly established HIC-MALS
404 results, we additionally performed native SEC-MALS experiments. BSA was again used as a
405 model protein to calibrate for exact molecular mass and a good agreement with the HIC
406 measurements was observed. The optimized SEC-MALS-RI method was then used to analyze NP
407 from DSP #5 (Figure 5). The main peak has a molecular mass of 99.9 kDa when using RI as a
408 concentration source. Some lower molecular mass impurities eluting at 30 min could also be

detected, which are likely to be the same as the non-binding fraction in HIC and are associated with nucleotides present in the sample. These are also faintly visible in the final product fraction in SDS-PAGE (Figure 2E, around 28 kDa). The SEC-MALS data supports the MW analysis from the HIC-MALS method, showing an excellent degree of overlap between the results obtained from both methods. The SEC elution profile is comparable with HIC, however, SEC results in a poorer resolution between the main peak and the co-eluted fraction. Despite optimization of the mobile-phase flow rate and column selection was performed to achieve the best resolution, this drawback inherently makes high-resolution screening with SEC-MALS very challenging.

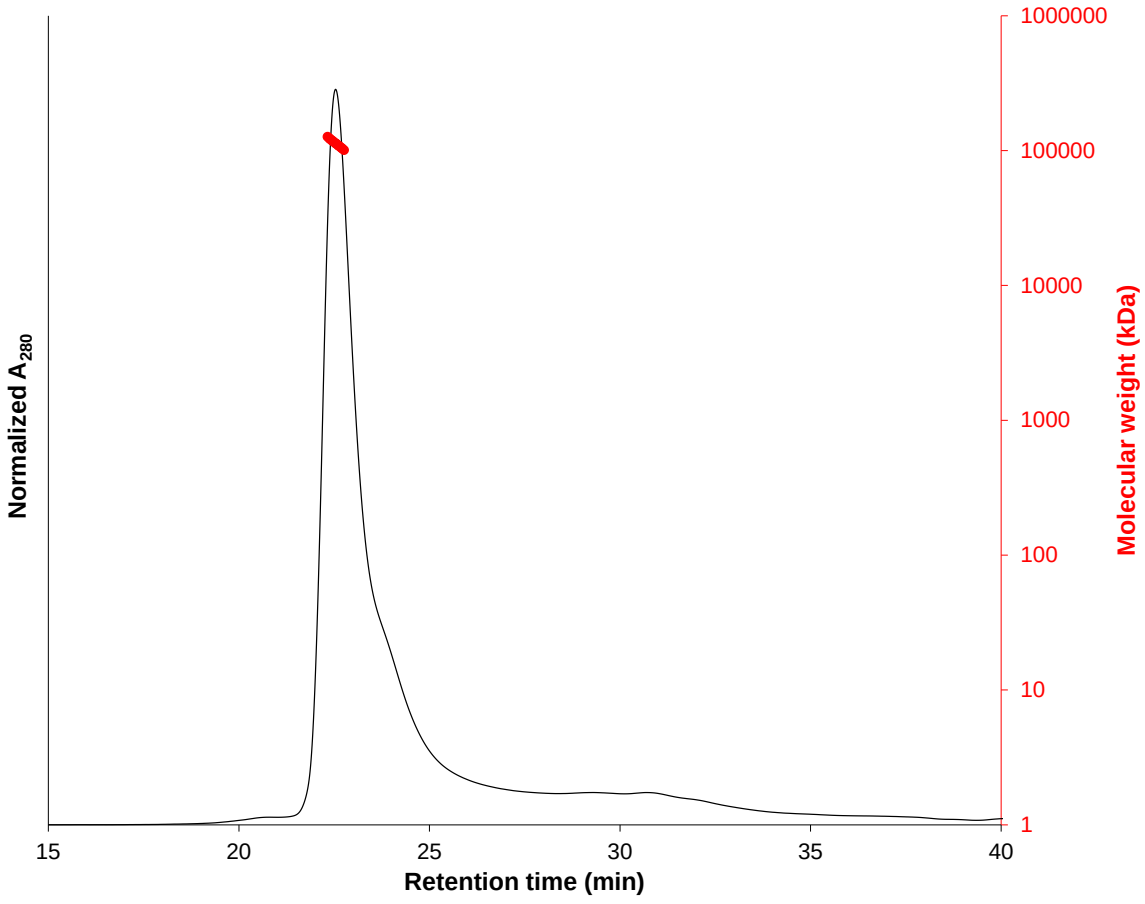


Figure 5. SEC-MALS data of NP originating from DSP #5. The main peak is related to the dimer of NP, and some lower-molecular weight species are visible around 30 min.

RPLC-MS peptide mapping of NP. A tryptic digest of the purified NP sample from DSP #5 was analyzed using RPLC-MS, applying a data-independent database search. The LC-MS total ion chromatogram (TIC) obtained from the mass spectrometer with associated peak assignment of the peptides can be seen in Figure S3A. The MS/MS data was searched against a protein sequence database by MASCOT including the host (*E. coli*) and the target sequence. The sequence coverage map (Figure S3B) shows the identified peptides (Table S3) based on color coding. The sequences which are highlighted in red are identified by MS/MS. The grey lines represent the peptides using greyscale shades to indicate the intensities of the precursor ions. The matched b- and y-ions are shown as red squares. N-glycosylation sites are highlighted in yellow. Based on the coverage of the C- and N-terminus, it can be concluded that the whole protein had been expressed. NP product from DSP #5 features nucleocapsid protein from human SARS-CoV-2 with high sequence coverage, supporting previous results obtained from native chromatography. MS/MS analysis also revealed the minor presence of chaperone protein DnaK from *E. coli*, with a molecular weight of 69.1 kDa. The final product fraction in Figure 2E depicts a faint band with a MW > 62 kDa, which is likely to be associated with DnaK.

CONCLUSIONS

HIC-MALS was developed for in-process analysis of a protein with intrinsic tendency for multimerization and fragmentation and was applied to evaluate product quality from downstream

processing. The new method is a powerful analytical tool that was used for evaluation of the end product resulting from a biotechnological production platform that features growth-decoupled expression in the *E. coli* enGenes-X-press strain, and the CASPON platform process[28,29]. Using this novel technology, it was shown that nucleocapsid protein from SARS-CoV-2 with high process yield (730 mg of purified NP per liter fermentation) and purity (95%) could be reliably produced. Complementary characterization tools, including, bioanalytical assays, SDS-PAGE, native SEC-MALS, and peptide mapping RPLC-MS were used to verify the results from the newly established HIC-MALS method. The NP product mainly contains dimerized protein (100 kDa) and has a co-eluting fraction which is likely associated with NP fragments that were also detected in SDS-PAGE. In contrast with SEC, which has limited resolution, the developed HIC method has the advantage that the resolving power can be improved by tuning the mobile-phase composition and gradient time. Combined with the molecular-mass information from MALS detection, HIC-MALS has proven to be a powerful tool to guide process development of proteins produced from recombinant origin under native conditions. With this novel analytical method, the availability of high-quality antigens for further research and diagnostic purposes can be accelerated, especially in highly demanding times, such as the current COVID-19 pandemic.

AUTHOR INFORMATION

Corresponding author

*E-mail: patricia.aguilar@boku.ac.at

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

We thank our company partners at Boehringer-Ingelheim RCV Process Bioscience for their collaboration and fruitful discussions. We would also like to acknowledge Novasign GmbH for making NP product available for this research. Samples of the pure NP product can be requested for research purpose from https://portal.boku-covid19.at.

JDV acknowledges the Research Foundation Flanders (FWO) for support by grants 12J6520N and V443719N, and the OEAD for scholarship ICM-2019-14929. PPA, AF, CK, MC and NL acknowledge the COMET center acib: Next Generation Bioproduction is funded by BMK, BMDW, SFG, Standortagentur Tirol, Government of Lower Austria und Vienna Business Agency in the framework of COMET - Competence Centers for Excellent Technologies. The COMET-Funding Program is managed by the Austrian Research Promotion Agency FFG. MD received funding by the Austrian Promotion Agency (FFG) [grant number 859219]. This work has also been partially funded by the Vienna Science and Technology Fund (WWTF) through project COV20-

016. The MS equipment for peptide analysis was kindly provided by the EQ-BOKU VIBT GmbH and the BOKU Core Facility for mass spectrometry.

REFERENCES

(1) Che, X. Y.; Hao, W.; Wang, Y.; Di, B.; Yin, K.; Xu, Y. C.; Feng, C. Sen; Wan, Z. Y.; Cheng, V. C. C.; Yuen, K. Y. Nucleocapsid Protein as Early Diagnostic Marker for SARS. *Emerg. Infect. Dis.* **2004**, *10* (11), 1947–1949. https://doi.org/10.3201/eid1011.040516.

(2) Diao, B.; Wen, K.; Chen, J.; Liu, Y.; Yuan, Z.; Han, C.; Chen, J.; Pan, Y.; Chen, L.; Dan, Y.; Wang, J.; Chen, Y.; Deng, G.; Zhou, H.; Wu, Y. Diagnosis of Acute Respiratory Syndrome Coronavirus 2 Infection by Detection of Nucleocapsid Protein. **2020**. https://doi.org/10.1101/2020.03.07.20032524.

(3) Surjit, M.; Lal, S. K. The SARS-CoV Nucleocapsid Protein : A Protein with Multifarious Activities. *Infect. Genet. Evol.* **2008**, *8*, 397–405. https://doi.org/10.1016/j.meegid.2007.07.004.

(4) Drosten, C.; Günther, S.; Preiser, W.; Van der Werf, S.; Brodt, H. R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A. M.; Berger, A.; Burguière, A. M.; Cinatl, J.; Eickmann, M.; Escriou, N.; Grywna, K.; Kramme, S.; Manuguerra, J. C.; Müller, S.; Rickerts, V.; Stürmer, M.; Vieth, S.; Klenk, H. D.; Osterhaus, A. D. M. E.; Schmitz, H.; Doerr, H. W. Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* **2003**, *348* (20), 1967–1976. https://doi.org/10.1056/NEJMoa030747.

(5) Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R. H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **2020**, *27* (4), 671-680.e2. https://doi.org/10.1016/j.chom.2020.03.002.

(6) Zhu, Y.; Liu, M.; Zhao, W.; Zhang, J.; Zhang, X.; Wang, K.; Gu, C.; Wu, K.; Li, Y.; Zheng, C.; Xiao, G.; Yan, H.; Zhang, J.; Guo, D.; Tien, P.; Wu, J. Isolation of Virus from a SARS Patient and Genome-Wide Analysis of Genetic Mutations Related to Pathogenesis and Epidemiology from 47 SARS-CoV Isolates. *Virus Genes* **2005**, *30* (1), 93–102. https://doi.org/10.1007/s11262-004-4586-9.

(7) Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W.; Rollin, P. E.; Dowell, S. F.; Ling, A. E.; Humphrey, C. D.; Shieh, W. J.; Guarner, J.; Paddock, C. D.; Roca, P.; Fields, B.; DeRisi, J.; Yang, J. Y.; Cox, N.; Hughes, J. M.; LeDuc, J. W.; Bellini, W. J.; Anderson, L. J. A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* **2003**, *348* (20), 1953–1966. https://doi.org/10.1056/NEJMoa030781.

(8) Krammer, F. SARS-CoV-2 Vaccines in Development. *Nature* **2020**, *586* (October). https://doi.org/10.1038/s41586-020-2798-3.

(9) Klausberger, M.; Dürkop, M.; Haslacher, H.; Wozniak-Knopp, G.; Cserjan-Puschmann, M.; Perkmann, T.; Lingg, N.; Aguilar, P. P.; Laurent, E.; De Vos, J.; Hofer, M.; Holzer, B.; Stadler, M.; Manhart, G.; Vierlinger, K.; Egger, M.; Milchram, L.; Gludovacz, E.; Marx, N.; Köppl, C.; Tauer, C.; Beck, J.; Maresch, D.; Grünwald-Gruber, C.; Strobl, F.; Satzer, P.; Stadlmayr, G.; Vavra, U.; Huber, J.; Wahrmann, M.; Eskandary, F.; Breyer, M.-K.; Sieghart, D.; Quehenberger, P.; Leitner, G.; Strassl, R.; Egger, A. E.; Irsara, C.; Griesmacher, A.; Hoermann, G.; Weiss, G.; Bellmann-Weiler, R.; Loeffler-Ragg, J.; Borth, N.; Strasser, R.; Jungbauer, A.; Hahn, R.; Mairhofer, J.; Hartmann, B.; Binder, N. B.; Striedner, G.; Mach, L.; Weinhäusl, A.; Dieplinger, B.; Grebien, F.; Gerner, W.; Binder, C. J.; Grabherr, R. A Comprehensive Antigen Production and Characterization Study for Easy-to-Implement,

Highly Specific and Quantitative SARS-CoV-2 Antibody Assays. *medRxiv* **2021**, 2021.01.19.21249921. https://doi.org/10.1101/2021.01.19.21249921.

(10) Mark, J.; Li, X.; Cyr, T.; Fournier, S.; Jaentschke, B.; Hefford, M. A. SARS Coronavirus : Unusual Lability of the Nucleocapsid Protein. *Biochem. Biophys. Res. Commun.* **2008**, *377* (2), 429–433. https://doi.org/10.1016/j.bbrc.2008.09.153.

(11) Chang, C.; Chen, C. M.; Chiang, M.; Hsu, Y.; Huang, T. Transient Oligomerization of the SARS-CoV N Protein – Implication for Virus Ribonucleoprotein Packaging. **2013**, *8* (5). https://doi.org/10.1371/journal.pone.0065045.

(12) Luo, H.; Ye, F.; Sun, T.; Yue, L.; Peng, S. In Vitro Biochemical and Thermodynamic Characterization of Nucleocapsid Protein of SARS. **2004**, *112*, 15–25. https://doi.org/10.1016/j.bpc.2004.06.008.

(13) Tang, T.; Wu, M. P.-J.; Chen, S.-T.; Hou, M.; Hong, M.-H.; Pan, F.-M.; Yu, H.-M.; Chen, J.-H.; Yao, C.-W.; Wang, A. H.-J. Biochemical and Immunological Studies of Nucleocapsid Proteins of Severe Acute Respiratory Syndrome and 229E Human Coronaviruses. *Proteomics* **2005**, *5*, 925–937. https://doi.org/10.1002/pmic.200401204.

(14) Kang, S.; Yang, M.; Hong, Z.; Zhang, L.; Huang, Z.; Chen, X.; He, S.; Zhou, Z. Crystal Structure of SARS-CoV-2 Nucleocapsid Protein RNA Binding Domain Reveals Potential Unique Drug Targeting Sites. *Acta Pharm. Sin. B* **2020**, *10* (7), 1228–1238. https://doi.org/10.1016/j.apsb.2020.04.009.

(15) Zeng, W.; Liu, G.; Ma, H.; Zhao, D.; Yang, Y.; Liu, M.; Mohammed, A.; Zhao, C.; Yang, Y.; Xie, J.; Ding, C.; Ma, X.; Weng, J.; Gao, Y.; He, H.; Jin, T. Biochemical Characterization of SARS-CoV-2 Nucleocapsid Protein. *Biochem. Biophys. Res. Commun.* **2020**, *527* (3),

618–623. https://doi.org/10.1016/j.bbrc.2020.04.136.

(16) Gauci, V. J.; Wright, E. P.; Coorssen, J. R. Quantitative Proteomics: Assessing the Spectrum of in-Gel Protein Detection Methods. *J. Chem. Biol.* **2011**, *4* (1), 3–29. https://doi.org/10.1007/s12154-010-0043-5.

(17) Baca, M.; De Vos, J.; Bruylants, G.; Bartik, K.; Liu, X.; Cook, K.; Eeltink, S. A Comprehensive Study to Protein Retention in Hydrophobic Interaction Chromatography. *J. Chromatogr. B* **2016**, *1032*, 182–188. https://doi.org/10.1016/j.jchromb.2016.05.012.

(18) Fekete, S.; Veuthey, J. L.; Beck, A.; Guillarme, D. Hydrophobic Interaction Chromatography for the Characterization of Monoclonal Antibodies and Related Products. *J. Pharm. Biomed. Anal.* **2016**, *130*, 3–18. https://doi.org/10.1016/j.jpba.2016.04.004.

(19) Chen, B.; Peng, Y.; Valeja, S. G.; Xiu, L.; Alpert, A. J.; Ge, Y. Online Hydrophobic Interaction Chromatography-Mass Spectrometry for Top-Down Proteomics. *Anal. Chem.* **2016**, *88* (3), 1885–1891. https://doi.org/10.1021/acs.analchem.5b04285.

(20) Amartely, H.; Avraham, O.; Friedler, A.; Livnah, O.; Lebendiker, M. Coupling Multi Angle Light Scattering to Ion Exchange Chromatography (IEX-MALS) for Protein Characterization. *Sci. Rep.* **2018**, *8* (1), 1–9. https://doi.org/10.1038/s41598-018-25246-6.

(21) Ye, H. Simultaneous Determination of Protein Aggregation, Degradation, and Absolute Molecular Weight by Size Exclusion Chromatography-Multiangle Laser Light Scattering. *Anal. Biochem.* **2006**, *356* (1), 76–85. https://doi.org/10.1016/j.ab.2006.05.025.

(22) Chang, C.; Hou, M.; Chang, C.; Hsiao, C.; Huang, T. The SARS Coronavirus Nucleocapsid Protein – Forms and Functions. *Antiviral Res.* **2014**, *103*, 39–50. https://doi.org/10.1016/j.antiviral.2013.12.009.

(23) Lim Tung, H. Y.; Limtung, P. Mutations in the Phosphorylation Sites of SARS-CoV-2 Encoded Nucleocapsid Protein and Structure Model of Sequestration by Protein 14-3-3. *Biochem. Biophys. Res. Commun.* **2020**, *532* (January), 134–138.

(24) Supekar, N. T.; Shajahan, A.; Gleinich, A. S.; Rouhani, D.; Heiss, C.; Azadi, P. SARS-CoV-2 Nucleocapsid Protein Is Decorated with Multiple N- and O-Glycans. *bioRxiv* **2020**, 2020.08.26.269043. https://doi.org/10.1101/2020.08.26.269043.

(25) Stargardt, P.; Feuchtenhofer, L.; Cserjan-Puschmann, M.; Striedner, G.; Mairhofer, J. Bacteriophage Inspired Growth-Decoupled Recombinant Protein Production in Escherichia Coli. *ACS Synth. Biol.* **2020**, *9* (6), 1336–1348. https://doi.org/10.1021/acssynbio.0c00028.

(26) Huang, C.; Lin, H.; Yang, X. Industrial Production of Recombinant Therapeutics in Escherichia Coli and Its Recent Advancements. *J. Ind. Microbiol. Biotechnol.* **2012**, *39*, 383–399. https://doi.org/10.1007/s10295-011-1082-9.

(27) Waugh, D. S. An Overview of Enzymatic Reagents for the Removal of Affinity Tags. *Protein Expr. Purif.* **2011**, *80* (2), 283–293. https://doi.org/10.1016/j.pep.2011.08.005.

(28) Cserjan-Puschmann, M.; Lingg, N.; Engele, P.; Krößß, C.; Loibl, J.; Fischer, A.; Bacher, F.; Frank, A.; Öhlknecht, C.; Brocard, C.; Oostenbrink, C.; Berkemeyer, M.; Schneider, R.; Striedner, G.; Jungbauer, A. Production of Circularly Permuted Caspase-2 for Affinity Fusion-Tag Removal: Cloning, Expression in Escherichia Coli, Purification, and Characterization. *Biomolecules* **2020**, *10* (12), 1592. https://doi.org/doi:10.3390/biom10121592.

(29) Lingg, N.; Cserjan-Puschmann, M.; Fischer, A.; Engele, P.; Krößß, C.; Schneider, R.; Brocard, C.; Berkemeyer, M.; Striedner, G.; Jungbauer, A. Advanced Purification Platform

Using Circularly Permuted Caspase-2 for Affinity Fusion-Tag Removal to Produce Native Fibroblast Growth Factor 2. *J. Chem. Technol. Biotechnol.* https://doi.org/https://doi.org/10.1002/jctb.6666.

(30) Sauer, D. G.; Mosor, M.; Jungbauer, A.; Dürauer, A. Separation of Truncated Basic Fibroblast Growth Factor from the Full-Length Protein by Hydrophobic Interaction Chromatography. *Sep. Purif. Technol.* **2021**, *254* (April 2020), 117564. https://doi.org/10.1016/j.seppur.2020.117564.

(31) Pereira Bresolin, I. R. A.; Lingg, N.; Bresolin, I. T. L.; Jungbauer, A. Hydrophobic Interaction Chromatography as Polishing Step Enables Obtaining Ultra-Pure Recombinant Antibodies. *J. Biotechnol. X* **2020**, *6* (January), 100020. https://doi.org/10.1016/j.btecx.2020.100020.

(32) Lingg, N.; Öhlknecht, C.; Fischer, A.; Mozgovicz, M.; Scharl, T.; Oostenbrink, C.; Jungbauer, A. Proteomics Analysis of Host Cell Proteins after Immobilized Metal Affinity Chromatography: Influence of Ligand and Metal Ions. *J. Chromatogr. A* **2020**, *1633*, 461649. https://doi.org/10.1016/j.chroma.2020.461649.

(33) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579* (7798), 265–269. https://doi.org/10.1038/s41586-020-2008-3.

(34) Stargardt, P.; Striedner, G.; Mairhofer, J. Tunable Expression Rate Control of a Growth - Decoupled T7 Expression System by l - Arabinose Only. *Microb. Cell Fact.* **2021**, 1–17.

https://doi.org/10.1186/s12934-021-01512-7.

(35) Marisch, K.; Bayer, K.; Cserjan-puschmann, M.; Luchner, M.; Striedner, G. Evaluation of Three Industrial Escherichia Coli Strains in Fed-Batch Cultivations during High-Level SOD Protein Production. *Microb. Cell Fact.* **2013**, *12* (1), 1. https://doi.org/10.1186/1475-2859-12-58.

(36) Kramer, M. C. W.; Duerrschmid, E. Metabolic Approaches for the Optimisation of Recombinant Fermentation Processes. **1999**, 43–50.

(37) Sauer, D. G.; Mosor, M.; Frank, A. C.; Weiß, F.; Christler, A.; Walch, N.; Jungbauer, A.; Dürauer, A. A Two-Step Process for Capture and Purification of Human Basic Fibroblast Growth Factor from E. Coli Homogenate: Yield versus Endotoxin Clearance. *Protein Expr. Purif.* **2019**, *153* (August 2018), 70–82. https://doi.org/10.1016/j.pep.2018.08.009.

(38) Kleinmann, I.; Plicka, J.; Šmídl, P.; Svoboda, V. Hydrophobic Interaction Chromatography of Proteins on Separon HEMA: I. The Effect of an Initial Salt Concentration on the Separation of Proteins. *J. Chromatogr. A* **1989**, *479*, 327–334. https://doi.org/https://doi.org/10.1016/S0021-9673(01)83347-1.

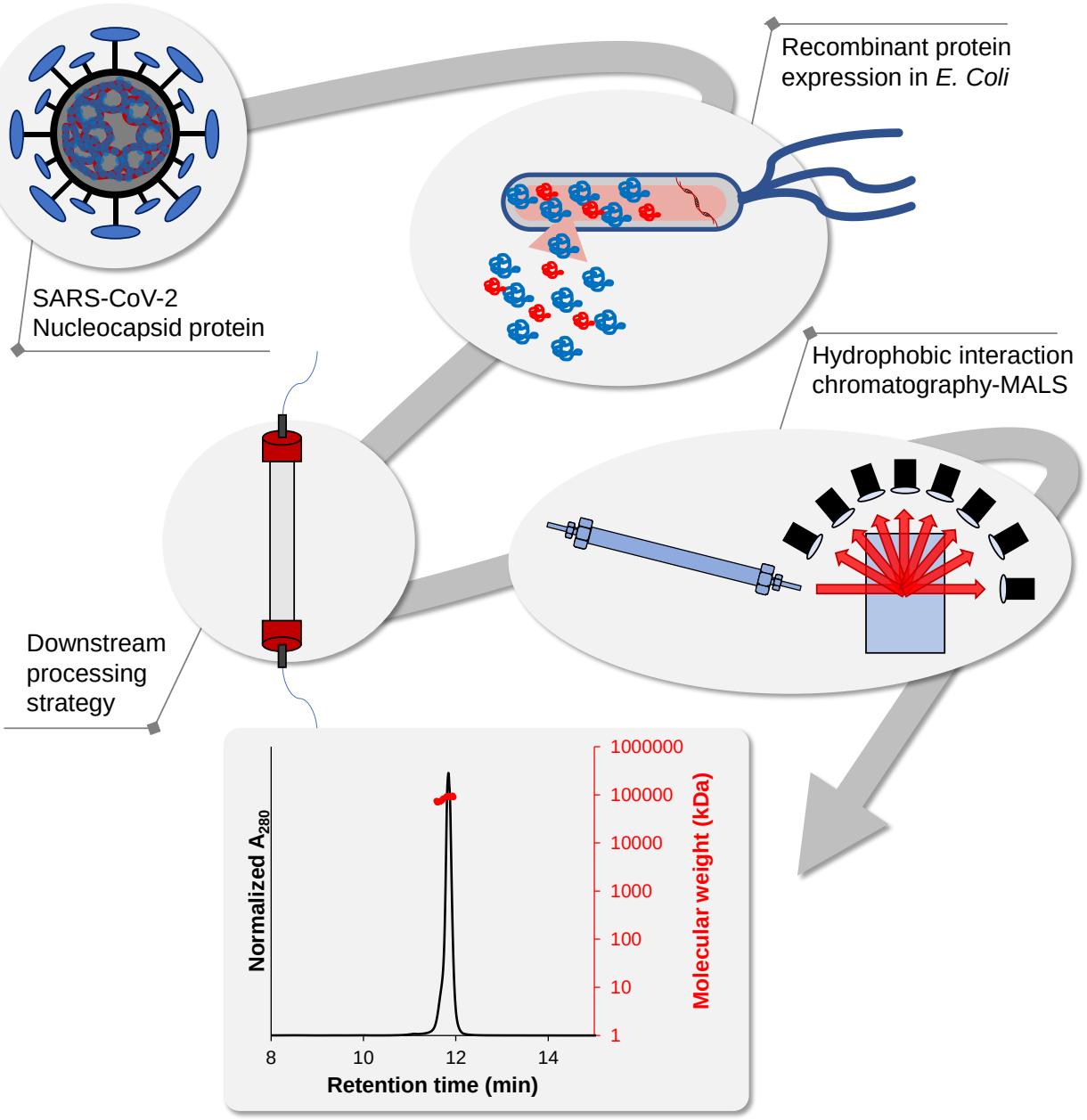
Table 1. *Overview of the conditions applied in the different downstream processing campaigns to purify SARS-CoV-2 nucleocapsid protein.*

DSP strategy	Combination of DSP	Pre-processing	Yield (%)	NP purity from HIC (%)
DSP #1	IMAC / UF/DF / ETR / IMAC / AEX / UF/DF	None	1.9	2.6
DSP #2	IMAC / UF/DF / ETR / IMAC / UF/DF	Denarase	60	73.4
DSP #3	IMAC / UF/DF / ETR / IMAC / CEX / UF/DF	Denarase	26	72.9
DSP #4	IMAC / UF/DF / ETR / IMAC / CEX / UF/DF / HIC / UF/DF	Denarase	8.5	93.6
DSP #5	IMAC / HIC / ETR / IMAC / UF/DF	Salt active nuclease	77	94.6

Table 2. *Purity parameters of NP at different steps of DSP #5. The concentrations of impurities are given relative to NP. The values for the homogenate are estimated based on SDS-PAGE quantification and literature values for typical E. coli fermentations*[37].

Process step	NP concentration (g/L)	dsDNA (µg/mg)	HCP (ng/mg)	Endotoxin (EU/mg)
Homogenate	~0.5	~650	~10^7	~10^6
Capture	6.0	6.0	4.0	12,731
Intermediate	6.6	4.4	1.9	209
Polishing	5.1	1.0	0.9	113

642 **for TOC only**



643