Expression and novel alternative purification of the recombinant nucleocapsid (N) protein of SARS-CoV-2 in *Escherichia coli* for the serodiagnosis of COVID-19

Jose David Rosales^{1*}, William Quintero², Jhon Cruz², Balbino Perdomo², Militza Quintero², Marcos Bastidas², Jose Domingo Lugo², Keila Rivas Rodriguez², Juan Carlos Freitas-Perez³, Annie Castillo⁴

^{1*}Center of immunoproduction of antigens (CIMA), Institute For Advance Studies Foundation (IDEA)

²Laboratory of Biology and Experimental Medicine (LaBioMEx) University of the Andes (ULA)

³Laboratory of Molecular Biology. Center of Agriculture and Food Sovereignty, Institute for Advanced Studies Foundation (IDEA)

⁴Interdisciplinary Doctorate in Sciences. Simón Bolívar University (USB)

* Correspondence of the author:

Center for Immunoproduction of Antigens (CIMA)

Institute for Advanced Studies Foundation (IDEA)

Hoyo de la Puerta National Highway

Sartenejas Valley Caracas-Venezuela

Phone: 005802129035172

Email: jdrr55@gmail.com

Abstract

The SARS-CoV-2 coronavirus causes severe acute respiratory syndrome and has caused a global pandemic by causing the COVID-19 disease. To monitor and control it, diagnostic methods such as molecular and serological tests are necessary. The serological approach uses SARS-CoV-2 antigens to detect the antibodies present in patients using quantitative techniques such as enzyme-linked immunosorbent assay (ELISA) or qualitative rapid tests such as lateral flow chromatography (RDT's). The main antigens used are the spike protein (S) and the nucleocapsid protein (N). Both proteins are obtained in different expression systems, in eukaryotic cells, their production is expensive, so in this work we chose a simpler and cheaper system such as prokaryotic to express and purify the N protein. Therefore, the nucleotide sequence had to be optimized to be expressed in *Escherichia coli*. The protein N is sensitive to *E.coli* proteases and also has the ability to self-proteolyze under native conditions, degrading into different fragments. However, under denaturing conditions, using urea and at pH 5.3 it is stable and efficiently purified using metal exchange chromatography (IMAC). In our purification strategy, we surprisingly found that by not using a sonicator, a homogeneous and time-stable preparation of the recombinant antigen is obtained. An approximate yield of 200 mg / L was obtained. It was then tested with healthy sera and sera from COVID-19 convalescent patients in Western-blot tests that were able to recognize it. Our work provides a novel strategy to produce the SARS-CoV-2 protein N so that it can be used as an input in the development and innovation of serological tests in the diagnosis of COVID-19.

Keywords: Nucleocapsid, recombinant protein, SARS-CoV-2, COVID-19

Introduction

The etiological agent of the severe acute respiratory syndrome (SARS-CoV-2) that emerged in 2019 has caused 250 million cases and 5 million deaths until November 2021 (WHO, 2021), being identified as a new type of coronavirus with a genome of single-stranded positive sense RNA (+ ssRNA) of 29,903 nucleotides, and its genomic sequence being reported for the first time in March 2020 (1), Genome ID: MN908947.3. Since then, thousands of genomes have been sequenced and potentially more contagious and serious new viral variants have been identified such as: Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Epsilon (B.1.427 and B.1.429), Eta (B.1.525), Iota (B.1.526), Kappa (B.1.617.1), Mu (B.1.621, B.1.621.1), Zeta (P.2) and delta (B.1.617) according to the Center for Disease Control and Prevention (5).

RNA detection is performed by reverse transcription polymerase chain reaction (RT-PCR), although there are other very promising molecular diagnostic methods including those based on Regularly Inter-spaced Clustered Short Palindromic Repeats (CRISPR), reverse transcription loop-mediated isothermal amplification (RT-LAMP), digital PCR, microarray assays, and next-generation sequencing (NGS) (3).

Molecular techniques are highly sensitive, consume a lot of resources, time and in certain conditions, where patients suffer from heart disease, immunosuppression and diabetes, a false negative result could be obtained (4), for which serological and clinical methods have been used to complement the analysis and help in containing the pandemic (5-6).

Serological tests provide rapid results by detecting antibodies against SARS-CoV-2 in blood, plasma or serum sample in most infected patients in the first 15 days after the start of the clinic and up to 3 months after infection, although, this time will depend on the immune response of each person.

This means that there is a window period where molecular techniques will stop detecting the presence of SARS-CoV-2 and will give negative PCR, while the anti-SARS-CoV-2 IgG antibodies will remain and give positive results (7). These tests will depend on the selected antigen, such as spike protein (S) and nucleocapsid (N) (8-9). Having recombinant antigens for serological diagnosis makes it possible to analyze the immune response to SARS-CoV-2, detect and characterize specific antibodies, and carry out large-scale epidemiological studies within the population. These serological tests will also help to evaluate the efficacy of the different drugs and vaccines available (10).

Protein N is a protein made up of 419 amino acids, with two domains rich in arginine and lysine, whose molecular weight varies between 45 and 60 kDa, depending on its post-translational modifications (11), and it plays an important role in the process of assembly of particles of virus, enveloping complete genomic RNA (12). Protein N is a powerful antigen for the diagnosis of SARS-CoV-2 infection by detecting IgM and IgG antibodies and is used in different immunological techniques such as chemiluminescent microparticle immunoassay (CMIA) (13), ELISA (14), Western blot (15) and rapid tests (16). Recombinant protein N is obtained from *E. coli*, baculovirus and mammalian cells (17-20). In our work we provide a simple strategy to produce SARS-CoV-2 Protein N in *Escherichia coli* and we demonstrate that it is recognized by positive sera from COVID-19 patients.

Materials and methods

Synthetic construction N-pET20b

The sequence of the SARS-CoV-2 recombinant protein N gene is obtained from the GenBank accession: MN908947.3 region 28274-29533. The optimized sequence is obtained synthetically and cloned into the prokaryotic expression vector pET20b (GenScript, USA). The promoter T7 oligos were used: AATACGACTCACTATAGG and terminator T7: GCTAGTTATTGCTCAGCGG, and a DNA polymerase (Promega., USA) for PCR, using the following parameters: heating at 95 ° C for 5 min; 32 cycles of denaturation at 95 ° C for 1 min, hybridization at 45 ° C for 1 min and extension at 72 ° C for 2 min.

Culture media:

Peptone, casein, yeast extract, NaCl (Promega, USA). Isopropyl-β-D-thiogalactopyranoside (IPTG), anti-human IgG conjugate (SIGMA-ALDRID, USA). Ni-NTA agarose resin (Invitrogen, USA).

Transformation and verification of the expression vector in *E. coli*:

The plasmid pET20b-N (4ug) is resuspended in 100 uL of Megapure water and 5 ng is used to transform the *E. coli* Top10 (Invitrogen, USA) to maintain the plasmid and in the strain of *E. coli* BL21 (DE3) (Novagen, USA) for your expression. Briefly, the transformation was carried out with the CaCl₂ chemical method and was placed in an LB plate containing 100 µg / ml of ampicillin and incubated at 37°C in a 12-16 h incubator. A PCR is performed from a colony to verify the insert. The fragment was visualized and identified by 1% DNA agarose gel electrophoresis.

Expression of protein N in *E. coli* cepa BL21 (DE3)

The construction of the recombinant SARS-CoV-2 N protein in the pET20b vector encodes a protein of 667 amino acids with a weight of 51 kDa. Induction was carried out in LB medium at 18° C with 1mM IPTG. Bacteria were collected by centrifugation (12000 g for 15 min at 4° C) and frozen at -20°C until use. Expression was verified on 15% SDS-PAGE gels, according to Laemmli (21). When verifying the soluble and insoluble fractions, all the protein was found forming inclusion bodies.

Recombinant nucleocapsid purification

Hybrid Conditions

Bacterial cells were resuspended in a solution consisting of Tris (100 mM) pH 8.0, NaCl (300 mM) and lysozyme (1 mg / ml), PMSF (1mM) and 10 uL of a protease inhibitor cocktail. It was incubated at room temperature for 60 min and frozen for 16 h, thawed and the pellet was separated from the supernatant by centrifugation at 12000 rpm. Two conditions were used. In the first (hybrid condition) the pellet was resuspended in 100 mM phosphate buffer (pH 8.0), 300 NaCl, 10mM Imidazole and 8M guanidine. It was incubated for 16 h with shaking and centrifuged at 12000 rpm. The supernatant was used to load a column with IMAC ProBond™ resin (Thermofisher, USA) and equilibrated with the same buffer. 5 column volumes

were washed with PBS with 10 mM imidazole and 8M guanidine, to remove non-specific proteins. Then it was washed with 5 column volumes with PBS without guanidine. The protein is then eluted with PBS containing 250 mM imidazole. In the second condition (denaturing condition) the pellet resulting from cell lysis was resuspended in PBS buffer with 6M urea, shaken for 2 hours and centrifuged. The supernatant was stored at -20° C. Subsequently, the fraction solubilized with urea is subjected to chromatography on metal affinity columns (IMAC). Unlike hybrid conditions, in denaturing conditions, the protein is eluted with PBS containing 6M urea. with a gradient of pH 7.8; pH 6.0; pH 5.3 and pH 4. The protein is purified at pH 4.0. It was then stored at 4°C until use.

SDS-PAGE and Western Blot Assays

The purified protein N was brought to 95° C for 10 min and diluted in 4 times Laemmli sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris HCl, pH 6.8). After electrophoresis, the gels were stained for one hour with Coomassie Brilliant Blue (Coomassie Brilliant Blue R-250 0.1%, acetic acid 10%, methanol 50%) and subsequently discolored with a solution of cold methanol. 50%. In Western-blot assays, the protein was transferred to Immobilon-P membranes (Milipore, USA) in transfer buffer (Tris 25 mM, glycine 192 mM and methanol 20%) at 2 mA / cm². for 30 min, using a semi-dry transfer kit (Bio-Rad, USA). The membrane was blocked with skim milk 5%, for 1 h at 4° C, washed three times with PBS-Tween 20 0.05%. In some tests, the membrane was cut into strips to evaluate individual sera. It was incubated with 3 sera positive people tested with the IgG / IgM rapid kit for COVID-19 (Abbott), the dilution of the sera was 1:50. Three other membranes were incubated with 3 healthy sera using the same dilution, in blocking buffer at room temperature (25°C) for 30 min. The membrane was washed with PBS-Tween 20 0.05% and incubated with peroxidase-conjugated anti-human IgG antibody (Sigma, USA) at a dilution of 1: 1000, in blocking buffer for 1 h at 37°C, and finally wash with PBS-Tween 20 0.05%. The protein was detected with a revealing solution of diaminobenzidine-H₂O₂. The positive and negative sera were provided by a health center, with prior informed consent to the patients.

Results and Discussion

Synthetic construction of nucleocapsid N

The nucleocapsid sequence of SARS-CoV-2 was obtained from accession number MN908947.3, which corresponds to the complete genome of the new coronavirus 2 of the isolated severe acute respiratory syndrome coronavirus 2 Wuhan-Hu-1. From the genome, the coding sequence of the nucleocapsid (region 28274-29533), was selected. The sequence is 1289 bp which code for a theoretical 45.6 kDa protein. In order to express the gene in *Escherichia coli*, the codon bias was changed (Figure 1) and it was obtained by chemical synthesis, and it was cloned in the expression plasmid pET-20b. This vector has an export sequence to the periplasm (pelb) that allows the protein to be purified by osmotic shock and a tag of six histidines at the C-terminal end that allows it to be purified by metal exchange chromatography (IMAC). When comparing the synthetic sequence with the wild one, 67.54% changes in the triplets are observed.

The plasmid is used as a template to amplify the insert using the forward and reverse oligos of T7, using the following parameters: heating at 95° C for 5 min; 32 cycles of denaturation at 95°C for 1 min, hybridization at 45°C for 3 min and extension at 72°C for 1 min, and a final elongation at 72°C for 5 min, a band of 1610 bp was obtained, (Figure 2), which corresponds to the insert plus part of the vector.

Expression of protein N in *E. coli* BL21 (DE3)

The *E. coli* strain BL21 (DE3) was used for recombinant N expression tests. A band with a molecular weight of 50 kDa was obtained, evidenced in 15% SDS-PAGE gels. Protein N was located in the insoluble fraction, forming inclusion bodies, so the osmotic shock strategy was not suitable. The intense band of molecular weight of 15 kDa corresponds to the lysozyme used in the disruption of bacteria (Figure 3).

Purification and Western-blot of the recombinant nucleocapsid (protein N)

Since it was not possible to use native conditions for purification, due to the nucleocapsid forming inclusion bodies, purification methodologies were carried out under conditions; hybrid and denaturing. The hybrid condition uses guanidine to solubilize the inclusion bodies, after centrifugation, the supernatant is loaded onto the metal affinity column (IMAC), washed with a guanidine buffer to eliminate unfixed proteins and then washed with PBS buffer without guanidine. The protein is eluted with 250 mM imidazole (Figure 4A). loss of protein by degradation is observed and, in addition, when performing the Western-blot an antigenic protein of *E.coli*, of approximately 60 kDa, is evidenced, which in case of using the ELISA technique would cause false positives (Figure 4C).

The methodology under denaturing conditions, includes urea in all solutions used to elute the nucleocapsid attached to the metal affinity column (IMAC), the elution is achieved with a stepped pH gradient. At pH 4.0, the nucleocapsid is eluted with a high degree of homogeneity (Figure 4B).

It is important to note that the fraction obtained under hybrid conditions, exhibited little homogeneity and was unstable, degrading even in the presence of protease inhibitors at –20°C, and it also included *E. coli* contaminants observed in western blots (Figure 4C). When examining the profile of the western blot with the isolated fraction under hybrid conditions, the autoproteolytic degradation recently reported is evidenced (22), due to the bands that appear in the reactions with positive serum, which exhibit different molecular weights and are not achieved. in reaction with COVID-19 negative serum (Figure 4C).

However, unlike purification under hybrid conditions, the denaturing condition yielded a fraction with a high degree of homogeneity (fig 4B), and also with very good stability. Stable solutions of the nucleoprotein are still preserved, even after two months of purification. The nucleoprotein profile is unaltered and reproducible, the reaction exhibited in western blot is highly specific as shown in Figure 4D.

When comparing the previous methodologies; The advantage resulting from the use of the denaturing condition is evident. It is then possible to use this methodology routinely to obtain recombinant nucleoprotein antigen, conveniently useful in diagnostic assays, given the complete absence of reaction with COVID-19 negative sera and its high stability, making it possible to conceive the development of a diagnostic assay. on a large scale such as ELISA kits or rapid tests. The strategy described under denaturing conditions is a simplified purification alternative to that described by De Marco Verissimo C (14) and Djukic et al (19) who report its expression in soluble form and Li (17) where it is obtained by forming bodies of inclusion.

Wester-blot with positive and negative covid-19 sera

The fraction eluted under denaturing conditions and at pH 5.3 was evaluated with 3 positive sera and 3 negative sera in western blot (5 ug / line), observing reaction bands in the 3 positive sera for COVID-19 (Figure 5A). There was no reaction with the COVID-19 negative sera (Figure 5B). The difference in intensity of the bands may be due to the antibody titers, since these depend on the immunological window period the patient is in.

Conclusion

An alternative method of purification of the recombinant nucleocapsid of SARS-CoV-2, expressed in *E.coli* and using only IMAC and denaturing conditions, is proposed. Unlike most established purification protocols, it does not use strong solubilizing or sonication agents. The procedure is relatively simple and inexpensive and could be carried out on a large scale to obtain an antigen that could be used in the development of serological diagnostic methods for COVID-19.

Acknowledgment:

Mauro Herrera President of CIDCITEI INTERNACIONAL, for the donation of the plasmid pET20b-N.

Financing: Project 2007001425 FONACIT "Development and / or adaptation of supplies to be applied to kits for serological diagnosis for infectious diseases"

MN908947.3	atg	tct	gat	aat	gga	ccc	caa	aac	cag	aga	aac	gca	ccc	cgt	att	acc	ttt	gac	gga	ccc	tca	gat	tca	act	ggt	agt	aac	cag	aat	gga	gaa	cgt	agt	gga	cgt	tca	aaa	caa	cgt		
Optimized sequence	atg	tca	gat	aac	gga	ccc	caa	aac	cag	agg	aac	gca	ccc	aga	atc	acc	ttc	gac	gga	ccc	aga	gat	aga	acc	ggt	aga	aac	cag	aac	gga	gaa	cgt	agt	gga	cgt	tca	aaa	caa	cgt		
Nucleocapsid	M	S	D	N	G	P	Q	N	Q	R	N	A	P	R	I	T	F	G	G	P	S	D	S	T	G	S	N	Q	N	G	E	R	S	G	A	R	S	K	Q	R	
	cgg	ccc	caa	ggg	tta	ccc	aaa	aat	act	ggc	tct	tgg	ttc	acc	ggt	ctc	act	caa	cat	ggg	aaa	gaa	gac	ctt	aaa	ttc	cct	gga	gga	ggg	ggt	cca	att	aac	acc	aat	agg	agg	cca		
	aga	cca	caa	ggg	ctg	ccg	aac	aac	ccg	ggg	aga	tgg	ttt	acc	ggg	tgg	acc	cag	cat	ggg	aaa	gaa	gat	ctg	aaa	ttc	ccg	cgt	ggg	cag	ggg	ggt	ccg	att	aac	acc	aat	tcc	tcc	cgg	
	R	P	Q	G	L	P	N	N	T	A	S	W	F	T	A	L	T	Q	H	G	K	E	D	L	K	F	P	R	G	Q	G	V	P	I	N	T	N	S	S	P	
	gat	gac	caa	att	ggg	tac	tac	cga	aga	gct	acc	aga	cga	att	cgt	ggg	ggg	gac	ggg	aaa	atg	aaa	gat	ctc	agg	cca	aga	tgg	tat	ttc	tac	tac	ctc	gga	act	ggg	cca	gaa	gct	gga	
	gat	gac	cag	atc	ggg	tat	tac	cag	cgt	ggc	act	cgt	cgt	atc	cgc	ggg	ggg	gac	ggg	aaa	atg	aaa	ggc	tgg	tcc	cag	cgt	tgg	tac	ttc	tac	tac	ctg	ggg	act	ggg	cct	ggg	gca	ggc	
	D	D	Q	I	G	Y	Y	R	R	A	T	R	R	I	R	G	G	D	G	K	M	K	D	L	S	P	R	W	Y	F	Y	Y	L	G	T	G	P	E	A	G	
	ctt	ccc	tac	ggg	ggc	aac	aaa	gac	ggc	atc	ata	tgg	ggt	gga	act	gag	gga	gac	tgg	aat	aca	cca	aaa	gat	cac	att	ggc	ccc	cgc	aat	cct	gct	aac	aat	gct	gca	atc	ggg	ctg	caa	
	ctg	ccg	tat	ggg	ggc	aac	agg	gat	ggc	att	att	tgg	ggt	gga	acc	gaa	ggg	ggg	ctt	aac	agg	cgc	agg	gac	cac	att	ggg	act	cgc	aac	ccg	ggg	aat	aac	ggc	ggc	atc	ggt	ctg	cag	
	L	P	Y	G	A	N	K	D	G	I	I	W	V	A	T	E	G	A	L	N	T	P	K	D	H	I	G	T	R	N	P	A	N	N	A	A	I	V	L	Q	
	ctt	cct	caa	gga	aca	aca	tgg	cca	aaa	ggg	ttc	tac	gca	gaa	ggg	agg	gga	ggc	ggc	agg	caa	ggc	tct	tct	cgt	tcc	tca	tca	cgt	agg	cgc	aac	agg	tca	aga	aat	tca	act	cca	ggc	
	ctg	ccg	cag	ggg	aca	acc	ctg	ccg	aaa	ggg	ttt	tac	gca	gag	ggc	agg	cgt	ggg	ggc	agg	caa	ggc	tct	tct	cgt	cgc	tcc	tct	cgc	tcc	cgt	aac	agg	cgc	tct	cgc	aat	tcc	acc	cgg	
	L	P	Q	G	T	T	L	P	K	G	F	Y	A	E	G	S	R	G	G	S	Q	A	S	S	R	S	S	S	R	S	S	R	N	S	S	R	N	S	T	P	G
	agg	agg	gga	ggg	act	tct	cct	gct	aga	atg	gct	ggg	aat	ggt	ggg	ggt	gct	gct	ctt	gct	tgg	ctg	ctg	ctt	gac	aga	tgg	aac	cag	ctt	gga	agg	aaa	atg	tct	ggt	aaa	ggt	caa	caa	
	agg	tct	cgt	ggc	acc	agg	cgc	gct	cgt	atg	gca	ggg	aac	ggg	ggg	gac	gca	ggc	ctg	gca	ctg	tta	ctc	gat	cgc	ctg	aac	cag	tgg	gaa	agg	aaa	atg	agg	ggg	aaa	ggt	cag	caa		
	S	S	R	G	T	S	P	A	R	M	A	G	N	G	G	D	A	A	L	A	L	L	L	L	D	R	L	N	Q	L	E	S	K	M	S	G	K	G	Q	Q	
	caa	caa	ggg	caa	act	gct	act	agg	aaa	tct	gct	gct	ggg	gct	tct	agg	agg	cct	cgg	caa	aaa	ggt	act	gct	act	aaa	gca	tac	aat	gta	aca	caa	gct	ttc	ggc	agg	cgt	ggg	cca	gaa	
	cag	caa	ggg	caa	acc	ggg	acc	aaa	aaa	ggc	ggc	ggc	ggg	tcc	aat	ggg	ccc	cgg	caa	aaa	ggc	acc	ggg	acc	agg	ggc	tat	aac	ggg	act	gag	ccc	ggg	acc	cgt	ggg	cgg	ggg			
	Q	Q	G	Q	T	V	T	K	K	S	A	A	E	A	S	K	K	P	R	Q	K	R	T	A	T	K	A	Y	N	V	T	Q	A	F	G	R	R	G	P	E	
	caa	acc	caa	ggg	aat	ttt	ggg	gac	gag	gaa	cta	atc	aga	caa	ggg	act	gat	tac	aaa	cot	tgg	cgg	caa	att	ggc	ccc	agg	ggt	tca	ggc	ttc	ttc	ggg	atg	tgg	cgc	att				
	cag	acc	caa	ggc	aac	ttt	ggg	gat	caa	ggg	ctg	att	cgt	caa	ggc	acc	gat	tat	aaa	cot	tgg	cga	cag	att	gga	caa	ttt	ggt	cgg	tca	ggt	agg	gac	ttc	ttt	ggg	atg	agg	cgt	atc	
	Q	T	Q	G	N	F	G	D	Q	E	L	I	R	Q	G	T	D	Y	K	H	W	P	Q	I	A	Q	F	A	P	S	A	S	A	F	F	G	M	S	R	I	
	ggg	atg	gaa	ggc	aca	cct	ggc	ggg	ggg	tgg	ttg	acc	tac	aca	ggg	ggc	atc	aaa	tgg	gat	gac	aaa	gat	cca	aat	ttc	aaa	gat	caa	ggc	atc	ttg	ctg	aat	agg	cat	act	ggc	gga	tac	
	ggg	atg	gaa	ggt	acc	cca	tct	ggg	acc	tgg	ctg	acc	tat	acc	ggg	ggc	atc	aaa	ctg	ggc	gat	aaa	ggc	cgg	aat	ttc	aaa	gac	cag	ggc	atc	ctg	ttg	aat	agg	cac	atc	gat	ggg	tac	
	G	M	E	V	T	P	S	G	T	W	L	T	Y	T	G	A	I	K	L	D	D	K	D	P	N	F	K	D	Q	V	I	L	L	N	K	H	I	D	A	Y	
	aaa	aca	ttt	cca	cca	aca	gag	tct	aaa	agg	gac	aaa	agg	agg	agg	gct	gat	gaa	act	caa	ggc	tta	ccg	cag	aga	cag	agg	aaa	cag	caa	act	ggg	act	ctt	ctt	ctt	gct	gga	gat	ttg	
	aaa	acc	ttt	ccg	cct	acc	gag	ccg	agg	agg	gac	aaa	agg	agg	agg	ggg	gat	gaa	acc	caa	ggc	ctg	ccg	cag	cgt	caa	aaa	agg	cag	caa	acc	ggg	acc	ctg	tta	ccg	ggg	gct	gac	ttg	
	K	T	F	P	P	T	E	P	K	K	D	K	K	K	K	A	D	E	T	Q	A	L	P	Q	R	Q	K	K	Q	Q	T	V	T	L	L	P	A	A	D	L	
	gat	gat	ttc	tcc	aaa	caa	tgg	caa	caa	tcc	atg	agg	agg	gct	gac	tca	act	cag	ggc																						
	gac	gac	ttc	agg	agg	cag	ctg	cag	cag	agg	atg	agg	agg	ggg	gat	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	agg	
	D	D	F	S	K	Q	L	Q	Q	S	M	S	S	S	A	D	S	T	Q	A																					

Figure 1. Codon bias of the SARS-CoV-2 nucleocapsid sequence, MN908947.3: 28274-29533. The triplets that were substituted to obtain an optimized sequence to be expressed in *Escherichia coli* are indicated in yellow. Below the sequences of each codon its corresponding amino acid is indicated.

Figure 2. Product of the amplification of the N insert from vector pET20b.
Line 1: Molecular weight marker.

Figure 3. Polyacrylamide gels with S.D.S at 15% of the fractions obtained by inducing the recombinant nucleocapsid with 1 mM IPTG at 37 oC for 16 h. PM: Molecular weight marker. Channel 1-2: Cell extract; Channel 3-4: Soluble fraction. Channels 5-8: Insoluble fraction. 50 ng of protein was used per lane. The apparent molecular weight of recombinant N is 50 kDa.

Figure 4. Analysis by SDS-PAGE 12% of the fractions obtained by purifying the recombinant N in: (A). Hybrid condition 1: lane 1: Induced extract. Lane 2: Insoluble fraction. Lane 3: eluted. Lane 4: fraction eluted with imidazole 50 mM, Lane 5: fraction eluted with imidazole 100 mM. Lane 6 fraction eluted with imidazole 250 mM. Lane 7: fraction eluted with imidazole 500 mM. Lane 8: fraction eluted with imidazole 1000 mM. (B). Denaturing condition: Lanes 1 and 2: induced fraction, Lane 3: washing. Lane 4-6: Elution with, pH 7.0; 6.0 and 5.3. Lane 7: Elution at pH 4.0, Lane 8: elution at pH 2.0. (C). western blot of positive and negative sera for COVID 19 using the fraction eluted with imidazole 250 mM under hybrid conditions, the samples were placed in duplicate; (D) Western blot of positive and negative sera for COVID 19 using the fraction eluted at pH 5.3 under denaturing conditions, the samples were placed in duplicate. The apparent molecular weight of recombinant N is 50 kDa. PM: Molecular weight marker.

Figure 5. Western-blot showing the presence and absence of antibodies against recombinant N in positive sera A, and negative sera B.

References

1. World Health Organization. 2020. World Health Organization coronavirus disease 2019 (COVID-19) situation report. <https://covid19.who.int/> ...Accessed 07 November 2021
2. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020 Mar;579(7798):265-269. doi: 10.1038/s41586-020-2008-3. Epub 2020 Feb 3. Erratum in: *Nature*. 2020 Apr;580(7803):E7. PMID: 32015508; PMCID: PMC7094943.
3. <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html> ... Accessed 07 November 2021
4. Habibzadeh, P., Mofatteh, M., Silawi, M., Ghavami, S., & Faghihi, M. A. (2021). Molecular diagnostic assays for COVID-19: an overview. *Critical reviews in clinical laboratory sciences*, 58(6), 385–398. <https://doi.org/10.1080/10408363.2021.1884640>
5. Caramello V, Macciotta A, De Salve AV, Mussa A, De Leo AM, Bar F, Panno D, Nota F, Ling CYG, Solitro F, Ricceri F, Sacerdote C, Boccuzzi A. False-negative real-time polymerase chain reaction tests in COVID-19 patients: an epidemiological analysis of 302 patients. *Public Health*. 2021 Sep 22;200:84-90. doi: 10.1016/j.puhe.2021.09.010. Epub ahead of print. PMID: 34710718; PMCID: PMC8455250.

6. Pallett, S.J.C., Denny, S.J., Patel, A. et al. Point-of-care SARS-CoV-2 serological assays for enhanced case finding in a UK inpatient population. *Sci Rep* 11, 5860 (2021). <https://doi.org/10.1038/s41598-021-85247-w>
7. Langa, L. S., Sallent, L. V., & Díez, S. R. (2021). Interpretación de las pruebas diagnósticas de la COVID-19. *FMC : Formación médica continuada en atención primaria*, 28(3), 167–173. <https://doi.org/10.1016/j.fmc.2021.01.005>
8. Naqvi, A., Fatima, K., Mohammad, T., Fatima, U., Singh, I. K., Singh, A., Atif, S. M., Hariprasad, G., Hasan, G. M., & Hassan, M. I. (2020). Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochimica et biophysica acta. Molecular basis of disease*, 1866(10), 165878. <https://doi.org/10.1016/j.bbdis.2020.165878>
9. Long, S. W., Olsen, R. J., Christensen, P. A., Bernard, D. W., Davis, J. J., Shukla, M., Nguyen, M., Saavedra, M. O., Yerramilli, P., Pruitt, L., Subedi, S., Kuo, H. C., Hendrickson, H., Eskandari, G., Nguyen, H., Long, J. H., Kumaraswami, M., Goike, J., Boutz, D., Gollihar, J., ... Musser, J. M. (2020). Molecular Architecture of Early Dissemination and Massive Second Wave of the SARS-CoV-2 Virus in a Major Metropolitan Area. *mBio*, 11(6), e02707-20. <https://doi.org/10.1128/mBio.02707-20>
10. Marfe G, Perna S, Shukla AK. Effectiveness of COVID-19 vaccines and their challenges (Review). *Exp Ther Med*. 2021 Dec;22(6):1407. doi: 10.3892/etm.2021.10843. Epub 2021 Oct 5. PMID: 34676000; PMCID: PMC8524740.
11. Nitin T. Supekar, Asif Shajahan, Anne S. Gleinich, Daniel Rouhani, Christian Heiss, Parastoo Azadi. SARS-CoV-2 Nucleocapsid protein is decorated with multiple N- and O-glycans. *bioRxiv*. 2020.08.26.269043; doi: <https://doi.org/10.1101/2020.08.26.269043>
12. Ye, Q., West, A., Silletti, S., & Corbett, K. D. (2020). Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *bioRxiv : the preprint server for biology*, 2020.05.17.100685. <https://doi.org/10.1101/2020.05.17.100685>
13. Weber, M. C., Risch, M., Thiel, S. L., Grossmann, K., Nigg, S., Wohlwend, N., Lung, T., Hillmann, D., Ritzler, M., Ferrara, F., Bigler, S., Egli, K., Bodmer, T., Imperiali, M., Salimi, Y., Fleisch, F., Cusini, A., Heer, S., Renz, H., Paprotny, M., ... Kahlert, C. R. (2021). Characteristics of Three Different Chemiluminescence Assays for Testing for SARS-CoV-2 Antibodies. *Disease markers*, 2021, 8810196. <https://doi.org/10.1155/2021/8810196>
14. De Marco Verissimo C, O'Brien C, López Corrales J, Dorey A, Cwiklinski K, Lalor R, Doyle JM, Field S, Masterson C, Ribes Martinez E, Hughes G, Bergin C, Walshe K, McNicholas B, Laffey JG, Dalton JP, Kerr C, Doyle S. Improved diagnosis of SARS-CoV-2 by using nucleoprotein and spike protein fragment 2 in quantitative dual ELISA tests. *Epidemiol Infect*. 2021 Jun 8;149:e140. doi: 10.1017/S0950268821001308. PMID: 34099081; PMCID: PMC8207563.

15. Smits VAJ, Hernández-Carralero E, Paz-Cabrera MC, Cabrera E, Hernández-Reyes Y, Hernández-Fernaudo JR, Gillespie DA, Salido E, Hernández-Porto M, Freire R. The Nucleocapsid protein triggers the main humoral immune response in COVID-19 patients. *Biochem Biophys Res Commun*. 2021 Mar 5;543:45-49. doi: 10.1016/j.bbrc.2021.01.073. Epub 2021 Jan 22. PMID: 33515911; PMCID: PMC7825866.

16. Mahmoudinobar, F., Britton, D., & Montclare, J. K. (2021). Protein-based lateral flow assays for COVID-19 detection. *Protein engineering, design & selection: PEDS*, 34, gzab010. <https://doi.org/10.1093/protein/gzab010>
17. Li, G., Li, W., Fang, X., Song, X., Teng, S., Ren, Z., Hu, D., Zhou, S., Wu, G., & Li, K. (2021). Expression and purification of recombinant SARS-CoV-2 nucleocapsid protein in inclusion bodies and its application in serological detection. *Protein expression and purification*, 186, 105908. <https://doi.org/10.1016/j.pep.2021.105908>

18. de Camargo, B. R., da Silva, L. A., de Oliveira, A. S., & Ribeiro, B. M. (2021). An easy pipeline for one-step purification of SARS-CoV-2 nucleocapsid protein from insect cell suspension culture. *Journal of virological methods*, 299, 114341. Advance online publication. <https://doi.org/10.1016/j.viromet.2021.114341>

19. Djukic T, Mladenovic M, Stanic-Vucinic D, Radosavljevic J, Smiljanic K, Sabljic L, Devic M, Cujic D, Vasovic T, Simovic A, Radomirovic M, Cirkovic Velickovic T. Expression, purification and immunological characterization of recombinant nucleocapsid protein fragment from SARS-CoV-2. *Virology*. 2021 May;557:15-22. doi: 10.1016/j.virol.2021.01.004. Epub 2021 Feb 9. PMID: 33582454; PMCID: PMC7871913.

20. García-Cordero J, Mendoza-Ramírez J, Fernández-Benavides D, Roa-Velazquez D, Filisola-Villaseñor J, Martínez-Frías SP, Sanchez-Salguero ES, Miguel-Rodríguez CE, Maravillas Montero JL, Torres-Ruiz JJ, Gómez-Martín D, Argumedo LS, Morales-Ríos E, Alvarado-Orozco JM, Cedillo-Barrón L. Recombinant Protein Expression and Purification of N, S1, and RBD of SARS-CoV-2 from Mammalian Cells and Their Potential Applications. *Diagnostics (Basel)*. 2021 Sep 30;11(10):1808. doi: 10.3390/diagnostics11101808. PMID: 34679506; PMCID: PMC8534734.

21. U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *nature* 227 (1970) 680-685.

22. Corinne A. Lutonski, Tarick J. El-Baba, Jani R. Bolla, Carol V. Robinson. Autoproteolytic Products of the SARS-CoV-2 Nucleocapsid Protein are Primed for Antibody Evasion and Virus Proliferation. *bioRxiv* 2020.10.06.328112; doi: <https://doi.org/10.1101/2020.10.06.328112>

23. De Marco Verissimo C, O'Brien C, López Corrales J, Dorey A, Cwiklinski K, Lalor R, Doyle JM, Field S, Masterson C, Ribes Martinez E, Hughes G, Bergin C, Walshe K, McNicholas B, Laffey JG, Dalton JP, Kerr C, Doyle S. Improved diagnosis of SARS-CoV-2 by using nucleoprotein and spike protein fragment 2 in quantitative dual ELISA tests. *Epidemiol Infect*. 2021 Jun 8;149:e140. doi: 10.1017/S0950268821001308. PMID: 34099081; PMCID: PMC8207563.