ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

COVID-19 PROJECT

LISV II BA4

Mathieu Reibel, Bilel El-Ghallali, Jade Therras

B-12

Contents

1 Background **2**

2 Aim 1: Design PCR tests for SARS-CoV-2 diagnosis and detection of new variants **5**
- 2.1 RNA isolation . . . 5
- 2.2 cDNA synthesis . . . 5
- 2.3 Primers validation . . . 6
- 2.4 Annealing Tenperature for PCR . . . 7
- 2.5 Agarose gel concentration . . . 7
- 2.6 proposed PCR test versus qRT-PCR for SARC-CoV-2 detection . . . 8
- 2.7 RdRP,N, and E genes targetting for SARC-Cov-2 detection . . . 8
- 2.8 B.1.1.7 variant : mutations . . . 9

3 Aim 2: Design protein-protein interaction assays to measure the binding affinity of the Spike protein to its surface protein receptor, ACE2 **12**
- 3.1 Mammalian transfection experiment . . . 12
- 3.2 Determination of the transfection efficiency . . . 13
- 3.3 Purification of RBD and RBD-N501Y proteins . . . 13
- 3.4 SDS-PAGE of the recombinant RBD and RBD-N501Y . . . 14
- 3.5 BCA assay to measure the protein concentration of your affinity purified RBD and RBD-N501Y proteins . . . 16
- 3.6 SPR experiment . . . 17

4 Aim 3: Vaccine investigation and proposals to adapt for the emergence of new variants **20**
- 4.1 Vaccine strategies . . . 20
- 4.2 Impacts of the variants on the effectiveness of the vaccines . . . 22
- 4.3 Proposals to adapt for the emergence of new variants . . . 24

5 Conclusion **25**

PART 1

BACKGROUND

Covid-19 is a high contagious disease caused by the SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2) : a positive-sense single-stranded RNA virus. It appeared in November 2019 in Wuhan (in China) before spreading across the world until the World Health Organization declared a pandemic on 11 March 2020. [1]

The SARS-CoV-2 belongs to a category of viruses named coronaviruses and particularly Betacoronaviruses. It is an 120 nm (diameter) enveloped virus with 3 membrane proteins : S (responsible for the crown shape), M and E. Its helicoidal nucleocapsid contains a single positive stranded RNA with a five prime cap and a poly A tail composed of around 29 kb. This RNA codes for 22 to 29 proteins.

The receptor on which the virus binds on a cell is ACE2, a surface protease that can be found especially in lung and small intestine epithelium.

The S protein has 2 sites :

- S1 (binds to ACE2)
- S2 (for the fusion)

The protease TMPRSS2 cleaves the virus between S1 and S2. This leads to the activation of the S2, triggering the fusion with the host membrane.

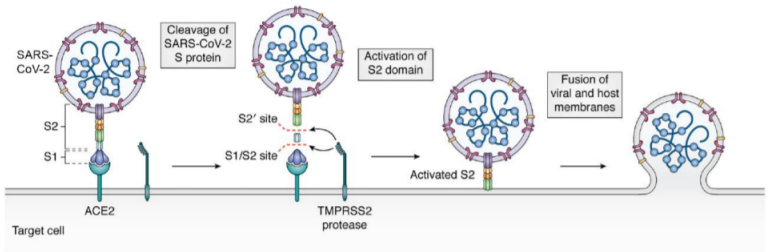


Figure 2. Mechanism of SARS-CoV-2 viral entry. The SARS-CoV-2 S protein engages with the host ACE2 receptor and is subsequently cleaved at S1/S2 and S2′ sites by TMPRSS2 protease. This leads to activation of the S2 domain and drives fusion of the viral and host membranes. See section on 'viral entry' for details. *Source : J. Biol. Chem. (2020) 295(37) 12910–12934*

FIGURE 1.1
Mechanism of SARC-coV-2 viral entry

Then, NonStructural Proteins are made by translation from the genome and a polymerase. Structural proteins are made after replication from subgenomic RNA. Moreover, some proteins are synthesized as precursors polyprotein and cleaved by viral proteases.

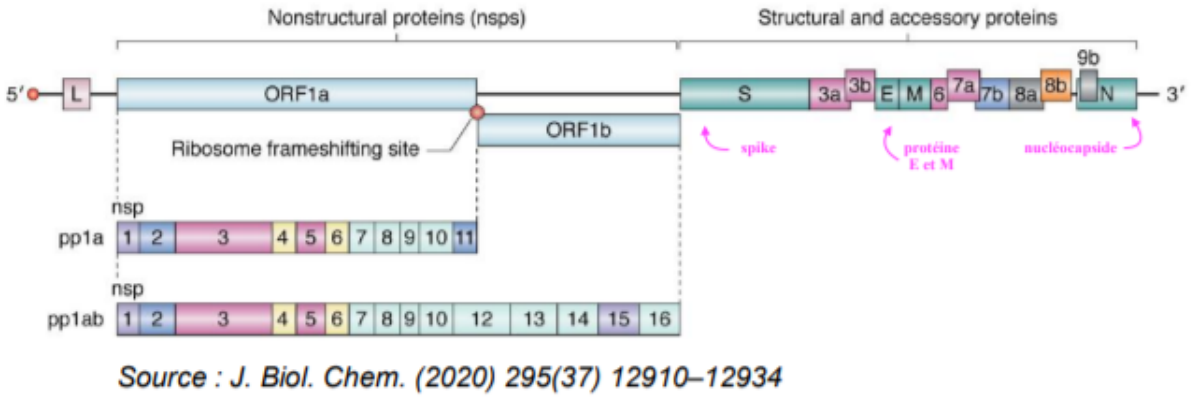


FIGURE 1.2
genome organisation of SARC-coV-2

For the structural proteins, the cycle begins with the original positive RNA copied into two negative RNAs, a complete copy, and an incomplete copy. This is the result of polymerase "jumps" and leads to the production of subgenomic negative RNA. Finally, subgenomic positive RNAs are made from the previous pieces of negative ones and allows the synthesis of structure proteins S, M and E.

These proteins follow the RE-GOLGI path and are grouped in a ERGIC (endoplasmic reticulum–Golgi intermediate compartment) and the nucleocapsids assembled in the cytoplasm join them. Once the nucleocapsid is in ERGIC, the virus is released. [2]

With its 2% of death rate [3], SARS-Cov-2 is not the most lethal virus that we encountered but it is extremely contagious. To inhibit its propagation, prevent additional contamination and give treatment, infected individuals need to be detected and isolated form healthy individuals. In addition, knowledge about the localisation and the proportion of infection allow cluster detection and following of the virus propagation, in order to take appropriate measures. Therefore, an efficient testing strategy is necessary.

To date, we know many ways to detect SARS-Cov-2, one way is by looking for the virus genetic material in a sample. This includes reverse transcription polymerase chain reaction (RT-PCR) testing, CRISPR, isothermal nucleic acid amplification, digital polymerase chain reaction, microarray analysis, next-generation sequencing...[4] And antigen tests whose purpose is to find viral proteins on the surface of the virus, but only the RT-PCR and the antigen test are commonly used.

The RT-PCR is based on the fact that if RNA can be found in a sample. It will first use reverse transcriptase to convert it into cDNA before the PCR which uses the principle of DNA replication :

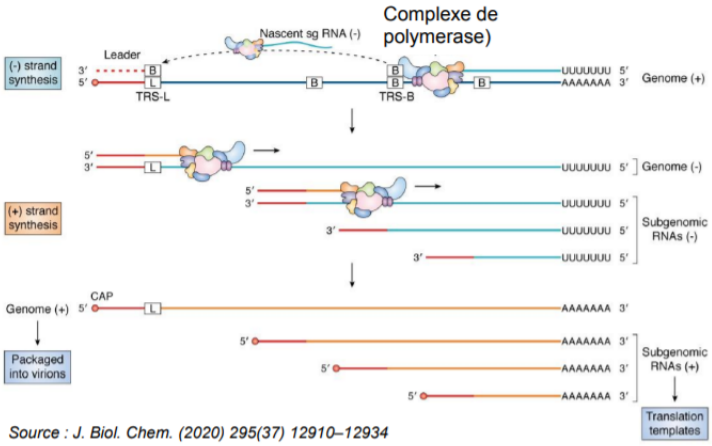


FIGURE 1.3
SARC-cov-2 replication

double stranded DNA is separated into single stranded DNA that will replicate itself with DNA polymerase following a cycle of 3 steps repeat multiple times: denaturation, annealing, and extension. This allows us to get large quantities of DNA (easier to detect) with a few picograms at first. Although it is very efficient to amplify fragments of DNA, only the result is available, but this flaw has been overcome with real time quantitative PCR. It gives information during the process because probes bind to the fragment of interest and emit fluorescent light when they are bound to DNA (single or double stranded). The measurement of light emission allows the quantification of DNA all along. [5]

Mutations can occur randomly on the sequence of the SARS-Cov-2, eventually provide advantage to the virus propagation and rendering the testing and treatment devices useless. Since covid 19 is highly contagious, the probability that mutations occur is pretty high too. The most iconic example is the variant that appeared in late January 2020 because of a D614G substitution in the gene that codes for the spike protein, this mutation effects aren't fully understood but structural analysis have shown that the receptor binding domains occupy a higher percentage of the virus' surface, giving it more chances to bind to ACE2 and then infecting the host. In June 2020, this variant replaced the original Wuhan strain as the most widespread, hopefully this mutation only increased its infectivity and transmission but not its effects on the human organism.

Developing tools to identify variants has become a major concern. Sequencing, PCR with specific primer for mutation, or imaging are use to detect variants. Organizations like the center for disease control and prevention regularly receive samples of SARS-CoV-2 for genomic sequencing allowing scientists to track the changes in the virus and understand how it might affect the human body. [1]

PART 2

AIM 1: DESIGN PCR TESTS FOR SARS-COV-2 DIAGNOSIS AND DETECTION OF NEW VARIANTS

2.1 RNA ISOLATION

Before extracting the RNA, destruction of the cells by a lysis buffer is needed so that nucleic acids be released.

Then, RNA isolation can be performed by the MagNA Pure 96 System (Roche, Penzberg, Germany) which uses magnetism. [6]

Magnetic particles purification uses para-magnetic beads with specific silica surfaces to bind to RNA. The RNA bound can then be easily separated from the aqueous phase with a magnetic field. With this process, centrifugation, is avoided. This step is time-consuming and generates shear forces that can damage RNA. Thus, using magnetism makes RNA isolation more efficient. It is a high-throughput isolation. [7]

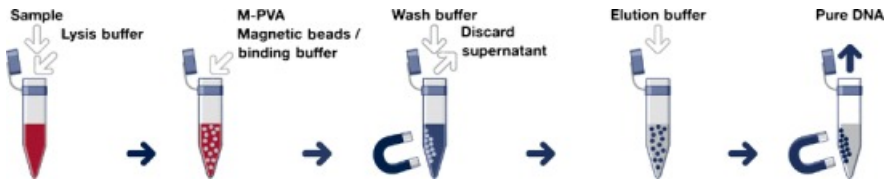


FIGURE 2.1
Schematic procedure for nucleic acid purification by magnetic bead technology [8]

2.2 CDNA SYNTHESIS

Viral RNA can be converted into cDNA by retrotranscription. This method use RNA primer.

With Superscript III One-step RT-PCR can be performed to convert viral RNA into cDNA, and amplify sequences at the same time. This method needs specific primers.[9]

For an RNA virus of unknown sequence we will firstly do a retro transcription with a random primer pool. This method amplify randomly part of the virus RNA. [10]

Here, we know the sequence of the virus. In addition, with a specific method we can process one-step RT-PCR that prevent contamination. We will choose specific primer.

2.3 PRIMERS VALIDATION

Target gene	Oligo	Sequence With all base pairs defined	Melting temp (°C)	GC %	Amplicon size (bp)
RdRP gene	RdRp_SARSrF	gtgaaatggtcatgtgtggcgg	58.8	54.55	100
	RdRp_SARSrR	caaatgttaaaaacactattagcata	49.1	23.08	
E gene	E_Sarbeco_F	acaggtacgttaatagttaatagcgt	54.1	34.62	113
	E_Sarbeco_R	atattgcagcagtacgcacaca	57.1	45.45	
N gene	N_Sarbeco_F	cacattggcacccgcaatc	56.8	57.89	128
	N_Sarbeco_R	gaggaacgagaagaggcttg	54.5	55.00	

TABLE 2.1
primer's characteristics

all primers are valids. Amplicons size, primers sequence and melting temperature are reasonable. Some contains low GC%, but a low GC aren't a problem, instead of high GC that make primer unspecific.

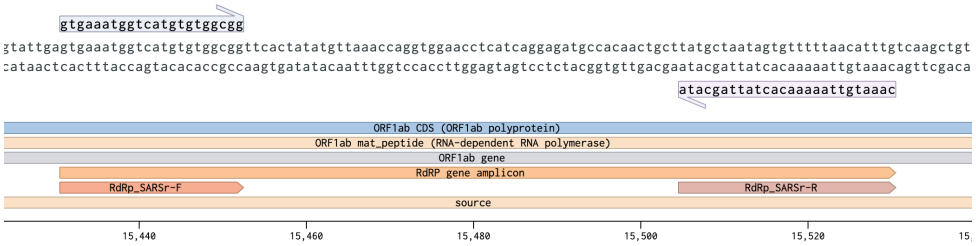


FIGURE 2.2
primer alignement and amplicon for rdrp gene [11]

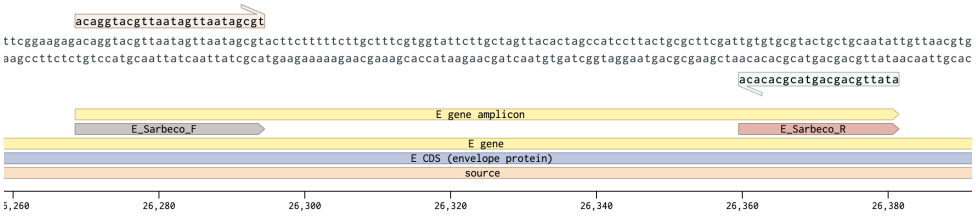


FIGURE 2.3
primer alignement and amplicon for E gene [11]

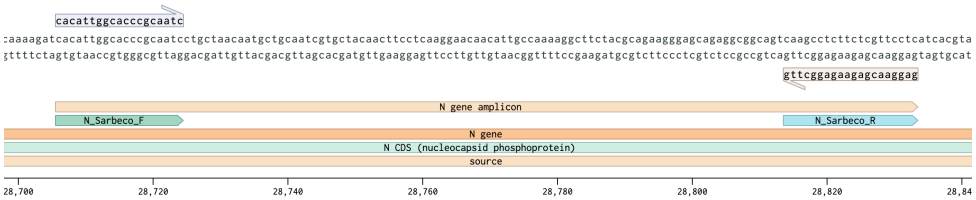


FIGURE 2.4
primer alignement and amplicon for N gene [11]

2.4 ANNEALING TENPERATURE FOR PCR

The annealing temperature depends on the composition of the reaction buffer and the sequences of the template and primers. A general rule of thumb is to begin with an annealing temperature 3-5°C lower than the lowest Tm of the primers. [12] Thus, the temperature of this step should be 20°C even if it seems quite low.

As a positive control we can use the known sequence of viral RNA (if we know it) sequence with the primers. A low temperature increases the specificity of the step but lower the sensitivity and a high temperature maximizes the sensitivity to the detriment of specificity.

As a negative control, we use just a buffer mixed with diluant and the primers to check if we have a contamination with nucleic acids or if we have interactions between the primers, for example self-annihilation.

2.5 AGAROSE GEL CONCENTRATION

Migration of double stranded DNA molecules through the gel matrix is inversely proportional to the log10 of the number of base pairs. Larger molecules will migrate slower than smaller ones because they enter less efficiently gel pores.

High concentration of agarose resolve better small molecules but take a lot of time. Low concentration can't be use for too small molecule cause the difference in the run will be to low to determinate mass difference, but is faster.

Our Amplicons size are included between 100 and 128 bp, according to Thermo Scientist reconmmandation, 2% agarose gel is appropriate. This concentration is the lowest, so the fastest, of whose allowing us to differentiate them.. [13]

Table 1. Recommended Agarose Gels for Electrophoretic Separation of DNA Fragments.

Agarose gel, %	Range of effective separation, bp	Approximate positions of tracking dyes, bp*			
		Bromophenol blue		Xylene cyanol FF	
		TBE buffer	TAE buffer	TBE buffer	TAE buffer
0.5	2000-50000	750	1150	13000	16700
0.6	1000-20000	540	850	8820	11600
0.7	800-12000	410	660	6400	8500
0.8	800-10000	320	530	4830	6500
0.9	600-10000	260	440	3770	5140
1.0	400-8000	220	370	3030	4160
1.2	300-7000	160	275	2070	2890
1.5	200-3000	110	190	1300	1840
2.0	100-2000	65	120	710	1040
3.0	25-1000	30	60	300	460
4.0	10-500	18	40	170	260
5.0	10-300	12	27	105	165

FIGURE 2.5
Agarose gel concentration recommended by thermofisher [13]

2.6 PROPOSED PCR TEST VERSUS QRT-PCR FOR SARC-COV-2 DETECTION

If we compare qRT-PCR with the conventional RT-PCR (use for the covid diagnostic), the qRT-PCR allows to have the quantity of DNA during the process and to know the amount at the beginning and thus, the viral charge of the patient. That is why, qRT-PCR is more precise (reduced risk of false positive) and has a better sensitivity.

However, the machines for qRT-PCR are more expensive and the fact that this method is more specific gives a restricted spectrum of pathogen detection and detecting the variants would be more difficult. [14]

2.7 RDRP,N, AND E GENES TARGETTING FOR SARC-COV-2 DETECTION

RdRp, N and E genes are three really conserved gene in SARC-CoV-2 genome. Targeting multiple region provide false positive.

If only one of the primer show a positive result and the patients present no symptoms, we suspect a false positive : the virus isn't present. The primer may be contaminated or another virus or human RNA show the same sequence. In contrary if the patient is negative but show symptoms, the virus can be present but the sample doesn't contain the virus, or RNA has been degraded for example by RNase before processing retro-transcription.

Internal extraction controls detects that nucleic acid are present. Ensure that negative sample aren't negative only because of lack of nucleic acid. negative extraction controls detect contamination during extraction process and validate extraction reagents and the success of RNA extraction.

If control fail or in case of doubt, it's necessary to retry the test from extraction step to confirm the result.

If we have a false positive, it can be explained by contamination after extraction or by traces of viral RNA in the nose if the sensitivity is high. If the temperature is high, the specifity is lowered, other elements can be amplified by the PCR and create the false positive. In case of false negative, lack of nucleic acid, RNA degradation or fails in PCR could explain it. [15] [16]

2.8 B.1.1.7 VARIANT : MUTATIONS

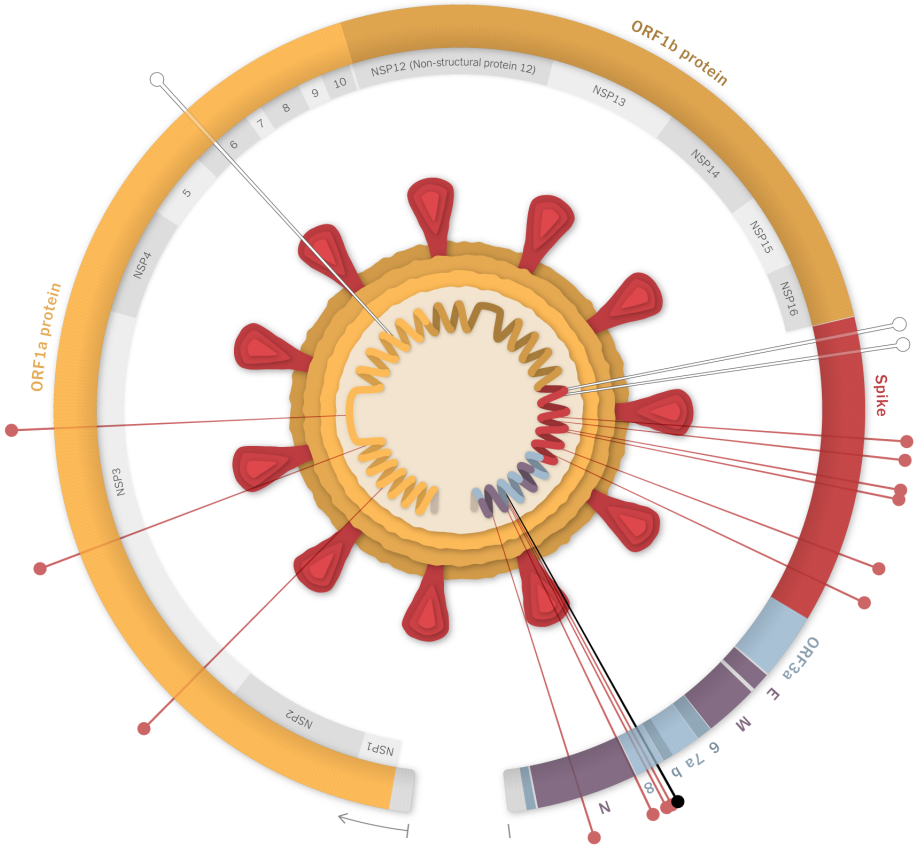
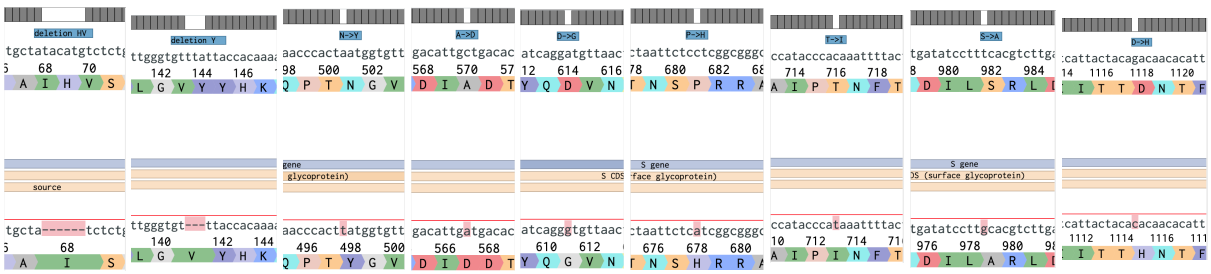


FIGURE 2.6
Map of the mutations of the B.1.1.7 (English) variant of COVID-19 [17]



mutation type	number	position (bp)	effect	notation
out of coding sequence	4	241	c->t in 5'UTR	
		28.048	g->a, after stop codon, R->I	
		28.111	a->g, after stop codon, I->C	
		28.271	deletion, not in a gene	
substitution non-sens	1	27.972	insert TAA	
deletion	3	11.288	miss SGF, no change in reading frame	
		21.765	miss HV, no change in reading frame	HV69-70del
		21.991	miss Y, no change in reading frame	Y144del
substitution silence	7	913	c->t, stay S	
		3.037	c->t, stay F	
		5.266	t->a, stay I	
		5.986	c->t, stay F	
		14.408	c->t, stay L	
		16.468	c->t, stay H	
		28.924	t->g, stay A	
substitution missence	17	3.267	c->t, T->I	
		5.388	c->a, A->D	
		6.954	t->c, I->T	
		14.676	c->t, P->L	
		15.279	c->t, T->I	
		16.176	t->c, L->P	
		17.615	a->g, S->G	
		23.063	a->t, N->Y	N501Y
		23.271	c->a, A->D	A570D
		23.403	a->g, D->G	D614G
		23.604	c->a, P->H	P681H
		23.709	c->t, T->I	T716I
		24.506	t->g, S->A	S982A
		24.914	g->c, D->H	D1118H
		28.280	gat->cta, D->L	
		28.881	ggg->aac, RG->KR	
		28.977	c->t, S->F	

TABLE 2.2
report of all mutations in B.1.1.7 variant [11]

mutation in S gene

A total of 32 mutations in the variant genome, including silence, non-sens, missence and deletion mutation. If we focus only in effective mutation (in the coding sequence, not silence), we count 21 mutations.

These graphic highlight that the region between 21.000 and 25.000 bases pairs is the more mutated. In fact we count 9 mutations in 4.000 bp, that correspond to 43% of mutations concentrated on 13% of the genome : mutations in B.1.1.7 variant are enriched in S gene (21563-25384bp). [18]

Mutations are selected by virus transmission, efficacy and resistance. The S gene encode for spike protein, that created the corolla of the virus and permit to catch cells, in order to infect it. Mutation in this gene affect the transmitiability of the virus.

The mutation that will affect the behavior of the virus the most is N501Y in the S gene. asparagine is

polar and uncharged, tyrosine is aromatic and hydrophobic. This mutation change the receptor-binding motif, the part of the receptor-binding domain of Spike that bind human ACE2. This mutation change binding specificity and recognition : we expect change in transmitiability and in antibody recognition.

We have to take care that we amplify a mutated region, and not a human sequence. If our primer can annealing with human sequence, then we can't extract the virus. We also use mutations in primers, to be specific to the variant and not to all strain. We propose two set of primer, one other that use two important mutation, like N501Y, in primer sequence, and one other that use two mutation deletion in primer sequence. These sets don't anneal with human sequence or with not mutated sequence.

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GGTGTTGGTTACCAACCATA	Plus	20	22997	23016	55.25	45.00	11.00	6.00
Reverse primer	GGACAGCATCAGTAGTGTCA	Minus	20	23222	23203	56.96	50.00	4.00	3.00
Product length	226								

FIGURE 2.7
primer set, in 2 mutation N501Y and A570D [19]

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GTTCCATGCTATCTCTGGGA	Plus	20	21694	21713	56.40	50.00	4.00	3.00
Reverse primer	TTGTGGTAAACACCCAAAAA	Minus	20	21933	21914	53.69	35.00	5.00	1.00
Product length	240								

FIGURE 2.8
primer set, in 2 deletions site of gene S [19]

PART 3

AIM 2: DESIGN PROTEIN-PROTEIN INTERACTION ASSAYS TO MEASURE THE BINDING AFFINITY OF THE SPIKE PROTEIN TO ITS SURFACE PROTEIN RECEPTOR, ACE2

3.1 MAMMALIAN TRANSFECTION EXPERIMENT

STEP 1 :

- Negative control : only Opti-MEM (medium without serum) because we will proceed to a mock transfer to check if the transfection is successful.

- Sample one : we want to proceed to a co-transfection with pHLsec-RBD and pGFP in order to highlight the success (or failure) of the transfection of this sequence.

- Sample two : we want to proceed to a co-transfection with pHLsec-RBD-N501Y and pGFP for the same reason of sample one.

STEP 2 :

Volume (µl)	Negative control	Sample 1	Sample 2
Opti-MEM	500	500	500
pGFP	0,00	8,00	8,00
pRBD	0,00	10,0	0,00
pRBD-N501Y	0,00	0,00	6,15

TABLE 3.1
Values to prepare DNA mix (4 µg of experimental plasmid and 4 µg of pGFP needed for each well)

STEP 3 :

Volume (µl)	Negative control	Sample 1	Sample 2
Opti-MEM	500	500	500
lipofectamine	20,0	20,0	20,0

TABLE 3.2
Values to prepare lipofectamine mix

STEP 4 :

- Incubate for 5min the mixtures, blend softly the DNA solution with the Lipofectamine solution and incubate for 20 minutes at room temperature.

- View cultures in 6 well plates by using an inverted microscope to check confluence and check if a contamination with bacteria or fungi occurred.

- Add 500 l of each transfection mix to the cells on two different plates. This step will be done under the laboratory fum. Blend smoothly by rocking the plate back and forth.

- Incubate the cells at 37°C in a CO2 incubator for 18-48 hours. After this step, we can test if the experiment was a success.

3.2 DETERMINATION OF THE TRANSFECTION EFFICIENCY

If the transfection was a success, cells will integrate two type of plasmid : the plasmid of interest and pGFP. GFP is a exceptionally stable fluorescence protein, is expression can be visualise and quantified. by excitation with UV or blue light, we can check GFP presence, and conclude for transfection success and efficiency, by comparaison with negative control reaction to light. Another technical is flow cytometry by labeling a reporter plasmid. [20] Finally, after affinity purification, detection of RBD and RBD-N501Y protein suppose success of transfection.

3.3 PURIFICATION OF RBD AND RBD-N501Y PROTEINS

- The RBD protein is secreted through the eukaryotic secretory pathway. While the protein is beeing formed during the translation, the N-terminal is targeted by proteins that drives the ribosome with RNA to the ER. In the organite, the synthesis continue until the protein is formed. Then, the COPII enzyme allows the protein to be transported to the cis side of the Golgi apparatus to undergo post-translational modifications (especially o-/n-glycosylation). When the RBD arrived to the trans side, a vesicle containing the protein buds (thanks to clathrin) and fused with the plasmic membrane to release the RBD. [21]

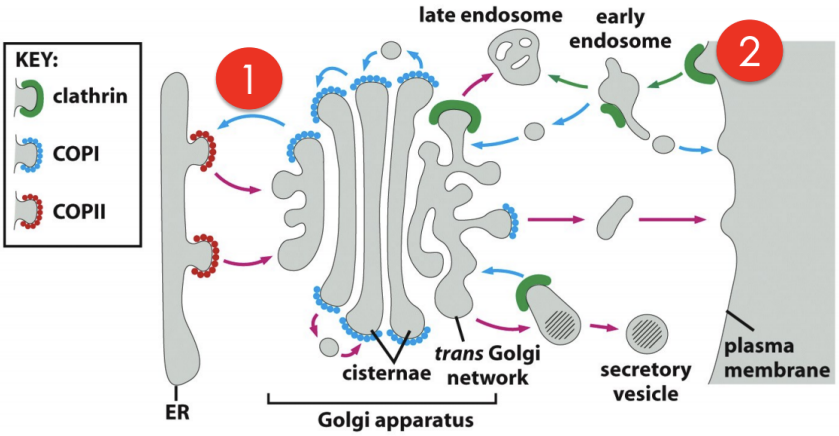
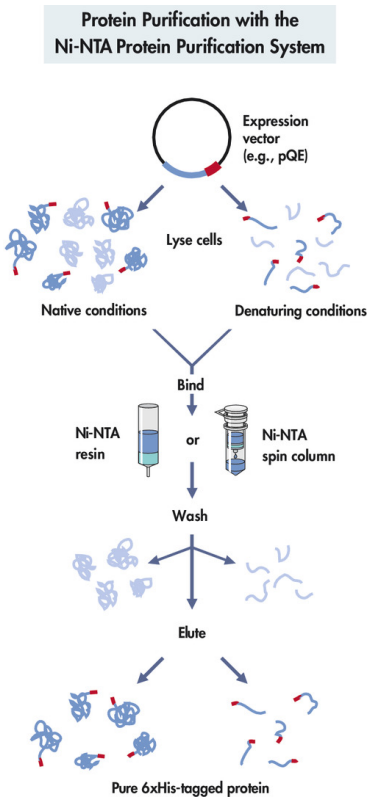


FIGURE 3.1
Picture of the eukaryotic secretory pathway taken from the page 674 of the book : Molecular biology of the cell, 6th edition [21].
The picture highlights two steps of the secretory pathway : the arriving from the ER to the Golgi and from the Golgi to the plasmic membrane

- His-Tag affinity purification is a binding-washing-elution process which is performed under native or denaturing conditions. Choosing native or denaturing conditions depends on several factors : location and solubility of the protein, accessibility of the His tag, wanted application. Purifications of His-tagged proteins can be proceeded by a single-step affinity chromatography, immobilized metal ion affinity chromatography (IMAC) that are available in different types of formats in trade. Ni-NTA matrices are the most used. The protocols provided describe protein purification in the batch binding mode and apply gravity-assisted flow in disposable columns; this procedure is easy to perform and robust. IMAC purification can also be performed in prepackaged columns by using liquid chromatography instruments or magnetic bead-based methods. [22] [23]



His tag fonctionnement [22]

3.4 SDS-PAGE OF THE RECOMBINANT RBD AND RBD-N501Y

We have to estimate RBD and RBD-N501Y molecular weight. Using pHLsec-IgK leader peptide-RDB-His tag and pHLsec-IgK leader peptide-RDBN501Y-His tag files, we take the number of amino acid for

this protein. Notice that the signal sequence is cleaved.

$$M = \frac{number\ of\ base}{3\ per\ amino\ acid} * 110 \frac{da}{amino\ acid}$$

then, we have

$$M(RBD) = M(RBD - N501Y) = \frac{702}{3} * 110 \frac{da}{amino\ acid} = 25.740\ kda$$

By verification with bioinformatics software [24], we find

$$M(RBD) = 26.36574\ kda \qquad M(RBD - N501Y) = 26.41481\ kda$$

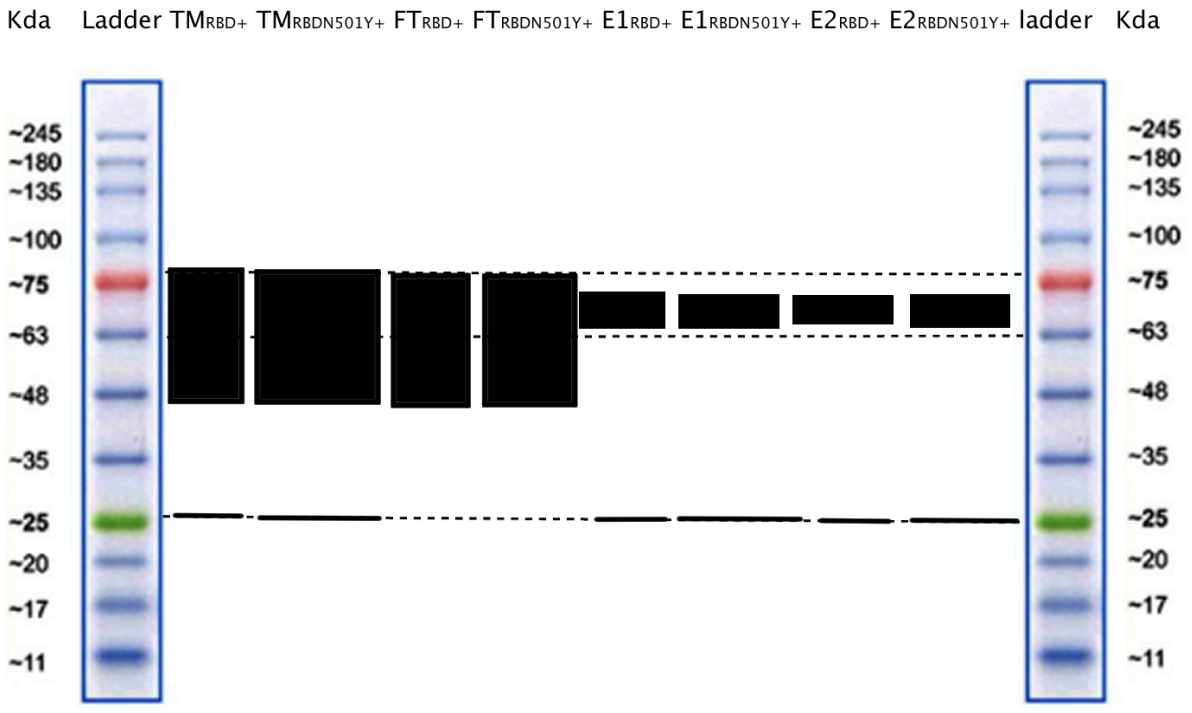


FIGURE 3.2
scheme of SDS-PAGE

1. Prestained size marker (10 µl)

   ladder of size comparaison

2. TM RBD+
3. TM RBD- N501Y+

   total medium, obtained with cell medium supplemented with 50mM Na2HPO4, 150 mM NaCl, 10 mM imidazole, use as base for affinity purification. Expected all secreted proteins from cells and medium proteins.

4. FT RBD+
5. FT RBD- N501Y+

   Flow through affinity column. Expected all protein that don't bind to the column (medium proteins for example)

6. E1 RBD+
7. E1 RBD- N501Y+

   first elution affinity column. Expected proteins from the part that bind to the column (his-tagged protein). Expect our protein of interest and maybe albumin, a important protein in medium.

8. E2 RBD+
9. E2 RBD- N501Y+

   second elution affinity column, Expected the rest of his-tagged proteins and maybe albumin.

10. Prestained size marker (10 µl)

    one more ladder at the end, to be able to determine protein size.

abbreviations : L = Lysate, M = Medium, TM = Total Medium, FT = Flow through, E = Elution.

3.5 BCA ASSAY TO MEASURE THE PROTEIN CONCENTRATION OF YOUR AFFINITY PURIFIED RBD AND RBD-N501Y PROTEINS

The BCA method is a colorimetric protein assay based on bicinchoninic acid. Amino acids residues including cysteine, cystine, tyrosine and tryptophan in the protein reduce the copper ion Cu (II) to Cu (I) in an alkaline medium. Bicinchoninic acid is a highly specific chromogenic reagent for Cu (I), which forms a purple complex with maximum optical absorption at 562nm (can be read between 540 and 590nm). A drawback of the BCA assay regarding the Bradford test is that it is sensitive to some substances in protein samples, like reducing agents (such as beta-mercaptoethanol) and high concentration buffers. [12]

Conc. µg/ml	**OD562**	**OD562 minus blank**
Blank	0,0945	0,0
25	0,1275	0,033
125	0,225	0,1305
250	0,3195	0,225
500	0,559	0,4645
750	0,7905	0,696
1000	0,9925	0,898
1500	1,501	1,4065
2000	1,9515	1,857
RBD(1/10)	1,689	1,5945
RBD-N501Y (1/10)	1.876	1,7815

Table of Values of OD562 and OD562-blank

To compute the concentrations, we use the equation obtained :

$$A = 0,0009 * C(g/L) + 0,0017$$

and we multiply by the factor of dilution 10. Thus, we have the formula

$$C(g/L) = 10 * (A - 0,0017)/0,0009.$$

For RBD, OD562-blank is 1,5945 so C=17,700g/L and for RBD-N501Y, OD562-blank is 1,7815 so C=19,780g/L

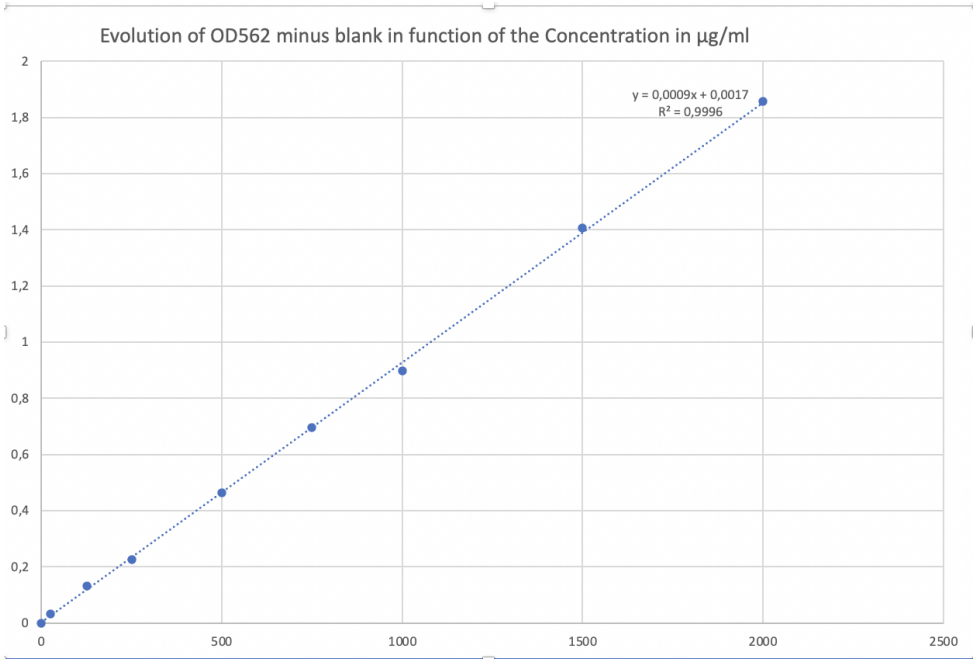


FIGURE 3.3
Curve of the absorbance in function of the concentration (on Excel)

3.6 SPR EXPERIMENT

A.

Thanks to the last part, we obtained a concentration of C(RBD)=17,700g/L, and C(RBD-N501Y) = 19,780 g/L. Then, since the molecular weight of RBD and RBD-N501Y is known, we use the formula

$$Cm = C(g/L)/m(Da = g/mol)$$

Therefore, we obtain for our two samples :

$$Cm(RBD) = 17,700/26365.74 = 0,671mM$$

$$Cm(RBD - N501Y) = 19,780/26414,81 = 0,749mM$$

B.

We diluate 1000X to obtain the concentration value :

$$Cm(RBD) = 0,671uM = 671nM$$

$$Cm(RBD - N501Y) = 0,749uM = 749nM$$

$$Vi = \frac{C_f V_f}{C_i}$$

Therefore,

$$Vi(RBD) = \frac{100}{671} C_f$$

$$Vi(RBD - N510Y) = \frac{100}{749} C_f$$

Sample	RBD		RBD-N501Y	
(nM)	(uL)	SPR buffer (uL)	(uL)	SPR buffer (uL)
0	0	100	0	100
2	0.30	99.70	0.27	99.73
4	0.60	99.40	0.53	99.47
8	1.19	98.81	1.07	98.93
16	2.38	97.62	2.14	97.86
32	4.77	95.23	4.27	95.73
64	9.54	90.46	8.54	91.46

C.

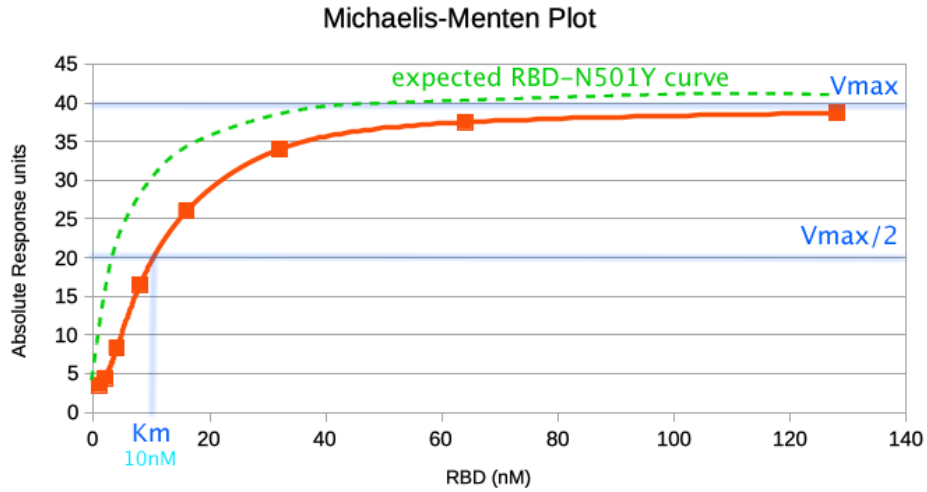


FIGURE 3.4
Michaelis-Menten plot for RBD-ACE2 binding measurements

Since N501Y is a gain of function mutation for the virus, that is more infectious, we expect RBD-N501Y-ACE2 binding affinity to be stronger.

the equilibrium dissociation constant Kd is commonly use to reprensent the binding affinity. The smaller the KD value is, the greater the binding affinity is. In Michaelis-Menten plot, the plateau of response correspond to the maximum for the velocity of the reaction. At one half of this velocity, the concentration correspond to Kd, according to Michaelis-Menten equation. [12] With the plot, we determine

$$Km = Kd = 10nM.$$

D.

In the same way, if the relationship between response units and concentration of analyte is linear, the concentration of analyte at one half of the maximal velocity is Kd, representing the binding affinity.

We find, using Michaelis-Menten plot

$$Kd(RBD) = 10\,nM,\ Kd(RBD - N501Y) = 5\,nM$$

We also try with Lineweaver-Burk Plot, but this method is too sensitive to give consistant result. Notice that Kd(RBD-N501Y) is approximately one half of Kd(RBD). Has expected, the mutation permit a better affinity for the ACE2 protein.

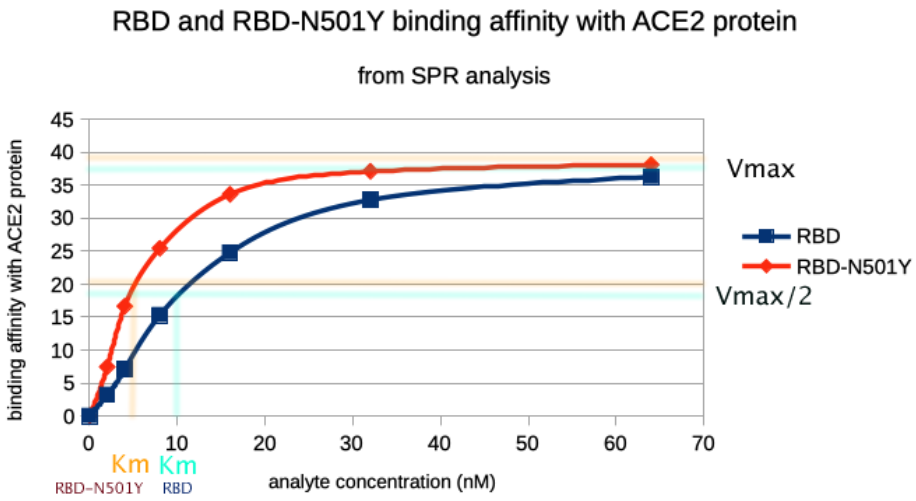


FIGURE 3.5
ACE2 binding affinity for RBD and mutant RBD

E.

A alternative method to measure the binding affinity can be Isothermal titration calorimetry (ITC). Since binding interactions have a enthalpy change, ITC measure the heat change in order to characterize affinity. It's a non-optical or spectroscopic methods, that can be performed in sample in solution without immobilization or denaturation. [25]

PART 4

AIM 3: VACCINE INVESTIGATION AND PROPOSALS TO ADAPT FOR THE EMERGENCE OF NEW VARIANTS

4.1 VACCINE STRATEGIES

Methods	How they work
Vaccines produced with SARS-CoV-2 : Virus weakened (Codagenix)	Attenuated viruses are obtained under laboratory conditions so that after infection, they can still replicate but cause no or very mild disease but still trigger immune response, and the mobilization of T cells, helper T cells and antibody-producing B cells, that lasts until the viruses is completely absent from the body, leaving plenty of time for memory cells to develop.
Vaccines produced with SARS-CoV-2 : Virus rendered uninfectious (CoronaVac)	Viruses are inactivated by destroying their genetic material, the immune response is then caused by the proteins contained in the virus, the immune system is able to react to them. Adjuvants are added in order to stimulate the immune response.
Nucleic acid vaccines (Pfizer, Moderna)	The virus' genetic material is isolated (DNA or RNA) and a specific gene that encodes for a specific antigen is identified and injected in human cells so that it is read by their protein-making material to finally produce antigens that can trigger immune response.
Viral vector-based vaccines (AstraZeneca)	A harmless virus (adenovirus, vaccinia virus...) is selected, then a sequence coding for a specific antigen of the SARS-Cov-2 (often it's the spike protein) is inserted in the virus' vector which is used a mean of transport to get in the patient's cells that will read the previous sequence and then produce the antigen to trigger immune response.
Protein subunit vaccine (Spy Biotech/Serum Institute of India)	For COVID-19, protein subunit is the only type of subunit vaccine produced. So fragments of proteins which have been selected to trigger a large immune response are injected inside the patient's organism (not inside the cells, so only an antibody mediated response is triggered).

TABLE 4.1
Table of the methods of the vaccines

Methods	**Pros**	**Cons**
Vaccines produced with SARS-CoV-2 : Virus weakened (Codagenix)	• Well-known technology • Strong immune response • Immune response mobilizes B cells and T cells • Pretty easy to produce	• Cannot be used on patients with weak immune systems • Can cause diseases in very rare case • Very sensitive to temperature
Vaccines produced with SARS-CoV-2 : Virus rendered uninfectious (CoronaVac)	• Well-known technology • Can be used on patients with weak immune system • No live components, so no risk of the vaccine causing any disease • Pretty easy to produce	• Booster shots can be required • Adjuvants can be required
Nucleic acid vaccines (Pfizer, Moderna)	• Immune response mobilizes B cells and T cells • No live components, so no risk of the vaccine causing any disease • Pretty easy to produce (Moderna vaccine was already at the human-test stage two months after the genome of SARS-CoV-2 was sequenced)	• May require very cold storage (Pfizer : -70°C) • Never been used on humans before the pandemic • Booster shots can be required
Viral vector-based vaccines (AstraZeneca)	• Well-known technology • Strong immune response • Immune response mobilizes B cells and T cells	• Hard to produce, vectors are grown with cells attached to a substrate, which is hard to do on a large scale, assembling them with the gene sequence is also complex, and all those steps require additional testing procedures, increasing the cost • If the vector used has already been met by the patient's organism, it could reduce its effectiveness
Protein subunit vaccine (Spy Biotech/Serum Institute of India)	• Well known technology • Can be used on patients with weak immune system • No live components, so no risk of the vaccine causing any disease	• Hard to produce since subunits are made thanks to living organisms such as bacteria that grow on substrates, which makes them expensive to produce • Lacks pathogen-associated molecular patterns which makes the immune response weaker, so it requires additional components • Booster shots can be required • Adjuvants can be required • It takes time to find the best antigen combination

TABLE 4.2
Table of the advantages and drawbacks of each method

One of the most important challenges in the struggle against the pandemic is the adaptation to new variants, so our favorite vaccine strategy would be the one that could be adapted to new variants very quickly. The RNA vaccine strategy appears to be the one to fulfill this function the best. Once the virus' genome has been sequenced (which is a pretty common procedure nowadays) it is very easy to identify a sequence that codes for a specific antigen and then use it to design a vaccine. This type of vaccine strategy has the advantage of being almost the same for any virus (including variants of SAR-CoV-2) whereas a strategy like inactivated viruses or subunit vaccines must be specifically adapted to the actual virus you are trying to treat. Vector based vaccines do not have this problem, but they are harder to design than nucleic acid ones, so RNA and DNA based vaccines remain the best solution for now.[26] [27] [28]

4.2 IMPACTS OF THE VARIANTS ON THE EFFECTIVENESS OF THE VACCINES

One of the most important mutation of the SARS-CoV-2 that led to the advent of new variants is the E484K, it is considered as an escape mutation since it allows the variants that contain it to pass through the immune response, therefore, rendering some vaccines ineffective since the principle of a vaccine is to trigger immune response. Indeed, the replacement of a lysine at the 484th position by a glutamic acid changes the affinity of the S protein of the SARS-CoV-2 with the ACE2 human protein, an affinity on which most vaccines are based. The majority of the vaccines trigger immune response using one of the ways we described earlier to create a contact between the S protein and the cells of the immune system in order to start the production of antibodies that would bind with the S protein instead of the ACE2 of human cells. To understand the cause of this change, we will analyze an experiment performed by a team of researchers in China [29]

MD trajectories	ΔE_{vdw}	ΔE_{ele}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta E_{vdw} + \Delta G_{nonpolar}$	$\Delta E_{ele} + \Delta G_{polar}$	ΔG_{bind}
WT[b]-1	-88.82	-619.25	662.75	-12.66	-101.48	43.5	-57.98
WT-2	-89.04	-675.15	721.90	-12.81	-101.85	46.75	-55.10
WT-3	-89.31	-650.30	696.32	-12.90	-102.21	46.02	-56.19
WT-4	-92.57	-650.33	691.09	-13.27	-105.84	40.76	-65.08
WT-5	-93.44	-678.66	724.33	-13.40	-106.84	45.67	-61.17
WT-ave[c]	-90.64	-654.74	699.28	-13.01	-103.64	44.54	-59.10
MT[d]-1	-93.46	-1124.49	1164.43	-13.34	-106.80	39.94	-66.86
MT-2	-96.23	-1134.52	1169.19	-14.36	-110.59	34.67	-75.92
MT-3	-91.78	-1111.58	1146.05	-13.35	-105.13	34.47	-70.66
MT-4	-92.81	-1105.55	1137.30	-13.71	-106.52	31.75	-74.77
MT-5	-90.59	-1092.43	1131.55	-12.99	-103.58	39.12	-64.46
MT-ave	-92.97	-1113.71	1149.70	-13.55	-106.52	35.99	-70.53

[a]The unit of the values is kcal/mol; [b]WT represents wild-type; [c]ave denotes average; [d]MT represents mutant.

FIGURE 4.1
Table of the binding free energies of the spike-RBD/hACE2 complex for wild type and E484K mutant

"Gbind = EMM + Gpolar + GnonPolar T S, where EMM is the gas phase interaction energy including the electrostatic and the van der Waals interactions; GPolar is the polar contributions to the desolvation free energy, which is calculated with the GB equation; GnonPolar is the nonpolar component of the desolvation free energy, which is estimated empirically by using the solvent accessible surface area; TS represents the change of the conformational entropy" 5 experiments for each type were led and one result is obvious, the binding free energies of the mutant RBD-hACE2 complex are lower than those of the wild-type (average binding free energies : 59.10 kcal/mol for the wild-type and 70.53 kcal/mol for the mutant) So it appears clear that the E484K mutation changes the spike protein in some way that allows it to have higher binding affinity with ACE2 than with the antibodies, let's have a look on those

changes and how they affect the binding of the S protein : First let's interpret the results of the electrostatic interactions calculations : Here, the hydrophobic interactions between the binding partners are represented by this equation Evdw + GnonPolar while the sum of the gas-phase electrostatic interactions and the polar desolvation energy, represented by this equation, Eele + GPolar informs us about the burial of the charged or polar groups during binding. By relying on these results we can conclude that for both types, the hydrophobic interactions are favorable and the burial of the charged and polar groups is unfavorable. But for the mutant complex, the Eele + GPolar is lowered from 44.54 kcal/mol to 35.99 kcal/mol meaning that the gas-phase electrostatic interactions compensate for the the unfavorable energies for the burial of the charged and polar groups during the binding phase. This is the main reason why the mutant spike binds better to ACE2 than the wild-type. -Now let's take a look at the interface between ACE2 and its contenders : Analyzing ACE2's, we can notice that it is overall negatively charged due to the presence of Glu35, Asp38 and Glu75 that have negative charges, since Glu484 of the spike S is negatively charged too, it gives a disadvantage to the virus, but the E484K replaces the Glu with a Lys that is positively charged which overcomes this obstacle. Therefore the average electrostatic energy decreases from 654.74 kcal/mol to 1113.71 kcal/mol. -Finally, looking at the difference between the structures of the two complexes: They noticed that Glu484 lies on a flexible loop of the RBD which means that the mutation to a Lys at this position caused conformational movements of this loop that led it to the receptor hACE2.

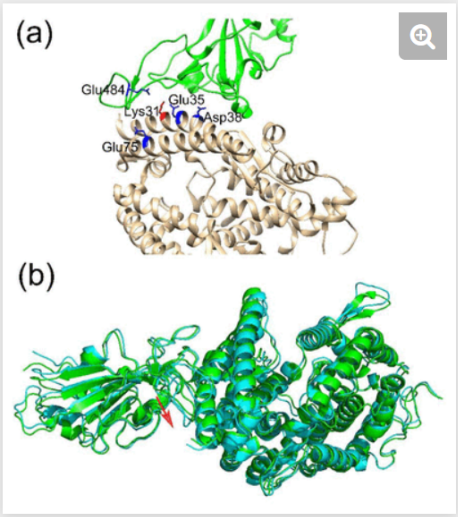


Fig. 1. Download figure | Open in new tab

(a) The charged residues on hACE2 locating nearby the mutated residue E484K. In this figure, RBD and hACE2 are shown in green and brown colors, respectively. The positively and negatively charged residues are displayed with the stick model in red and blue colors, respectively. (b) The superposition of the average structure in the MD simulations between the wild-type complex and the E484K mutant. The average structures of the wild-type complex and the mutant are displayed in green and canyon colors, respectively. The red arrows in the figure represent the local structural movements of the mutant system compared with the wild-type.

FIGURE 4.2
MD simulation of ACE2 and the wild-type S protein (a) and of the superposition of the wild and mutant type of S protein (b).

However, the E484K mutation doesn't only increase the affinity between ACE2 and S, it also reduces the affinity between S and its antibodies.

Looking at these results, we can clearly see that the binding energy is always increasing with the E484K mutation, therefore, it reduces the binding affinity with these antibodies. Furthermore, we can see that the electrostatic energy of every one of these systems is increased with the E484K mutation which is also limiting the binding of these antibodies with the mutant RBD. Regarding the structural facts, it appears that changing the Glu484 into a Lys affects all the antibodies in the same way, it changes the electrostatic forces equilibrium leading to a loss of binding affinity with the mutant RBD : For the antibody BD23,

Neutralizing antibodies	RBD	ΔE_{vdw}	ΔE_{ele}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta E_{vdw} + \Delta G_{nonpolar}$	$\Delta E_{ele} + \Delta G_{polar}$	ΔG_{bind}
S2M11	WT[b)]	-71.59	-74.16	97.06	-9.33	-80.92	22.90	-58.02
	MT[c)]	-72.79	31.18	8.59	-8.50	-81.29	39.77	-41.52
BD23	WT	-71.28	-122.61	149.54	-9.89	-81.17	26.93	-54.24
	MT	-75.06	44.91	7.04	-9.85	-84.91	51.95	-32.96
BD368-2	WT	-74.68	-92.73	128.59	-9.59	-84.27	35.86	-48.41
	MT	-70.23	3.98	35.93	-9.21	-79.44	39.91	-39.53
nanobody Nb20	WT	-82.75	-237.52	268.30	-12.09	-94.84	30.78	-64.06
	MT	-75.69	-47.65	85.08	-10.53	-86.22	37.43	-48.79
Nanobody H11-D4	WT	-62.60	-2.90	26.09	-8.36	-70.96	23.19	-47.77
	MT	-64.60	271.64	-236.30	-7.89	-72.49	35.34	-37.15
Nanobody MR17-K99Y	WT	-102.33	-299.26	344.51	-13.63	-115.96	45.25	-70.71
	MT	-106.09	-226.00	274.81	-13.94	-120.03	48.81	-71.22

[a)]The unit of the values is kcal/mol; [b)]WT represents wild-type RBD; [c)]MT represents the mutant RBD.

FIGURE 4.3
Table of the binding free energies of the S protein (wild and mutant type) with 6 antibodies

the appearance of the positive Lys484 creates repulsive interactions with the positive Arg107 moving the two chains of amino acids apart. For the nanobody H11-D4, it's the exact same situation but with Arg52 For the antibody BD368-2, Glu484 (-) is stuck in a sandwich made of Arg100 (+) and Arg102 (+) which results in a very stable structure, replaced with a Lys, the structure falls apart too. For the nanobody Nb20, the attractive interactions Glu484 had with Arg31 and Aeg97 are changed to repulsive. For the nanobody MR17-K99Y, it's the same situation as the previous but with Arg59 and Arg33. For the antibody S2M11, Glu484 forms strong hydrogen bonds with a lot of polar amino acids, Ser55, Asn52 and Tyr33, but the change into a Lys breaks most of them. To conclude, the term escape mutation seems appropriate since it enhances the capacity of the S-protein to bind more with its original target and to "escape" the antibodies designed to counter it. [30]

4.3 PROPOSALS TO ADAPT FOR THE EMERGENCE OF NEW VARIANTS

As said earlier, the current challenge for vaccines is to adapt to variants, otherwise any vaccine just becomes outdated once a new variant arises. For now, vaccines that aim to make the human body produce the spike protein (Pfizer, Moderna, AstraZeneca…) could be updated to make it produce the new mutant version of the antigen and also the old one to be sure that the vaccine would be effective against any strain that we already know. But a more efficient solution would be to design vaccines that aim to produce another antigen that would be less vulnerable to mutations since it is obvious that the spike protein is very sensitive to them, it could be another surface protein of the SARS-CoV-2. Flexibility of the vaccines will be the key to limit the spread of the virus, laboratories should focus on vaccination techniques that are easy to modify, like nucleid acid vaccines that do not require huge modifications to be effective against a new variant, only a modification of the RNA or DNA sequence. A "template" of the vaccine could be kept and completed for every new variant appearance.

PART 5

CONCLUSION

In view of what we have studied, we can notice with the measurement of infectivity of the variants and the proposal to adapt the vaccines for the emergent of new variants that they attract our attention. The media evoked the variant B.1.1.7 (British), the B.1.351 (South African), P.1 (Brazilian), B.1.427 and B.1.429 (Californian). They are known to spread faster and more easily than the original version and can increase the number of cases, the resources to contain the pandemic and the number of death people in case of the saturation of the hospital. We can explain the process of emerging of variant by the fact that when a virus replicates, mutations can occur. Thus, the more the virus circulates, the more it may mutate. These changes can occasionally result in a virus variant that is better adapted to its environment compared to the original virus. This process of changing and selection creates the "variants". Some mutations can lead to changes in a virus's characteristics like reproduction rate or severity of the symptoms. Some viruses change quickly and others more slowly. Coronaviruses (SARS-CoV-2 included), tend to change more slowly than others like retroviruses (HIV or influenza). [1] [31]

As a conclusion, many methods are used for the vaccines : virus weakened, viral vector-based vaccines or Nucleic Acid vaccines (which has not been used on humans before the pandemic). For the moment, the vaccines are effective and we can notice it in countries where the population has been vaccinated early like Israel.[32][33] However, the vaccines raise questions about the duration of the effectiveness of the vaccines and the effectiveness on the variants. The risks are to restart from stretch if a resistant variant emerge or that the vaccines do not last enough to allow the containment of the pandemics on the long term if they become obsolete. Therefore, the key to combat with the other variants is the flexibility of the vaccine. A strategy would be to design vaccines whose purpose would be to produce another antigen that would be less vulnerable to mutation. Another would be to focus on vaccination technics that are easy to modify, such as Nucleic Acid vaccines. Thus, it would allow to be adapted quickly if a new variant become problematical.

BIBLIOGRAPHY

[1] *SARS-CoV-2 Spike protein (RBD, His Tag) Protein - GenScript*. URL: https://www.genscript.com/protein/Z03479-SARS_CoV_2_Spike_protein_RBD_His_Tag_.html (visited on 26th May 2021).

[2] American Society of Biological Chemists. *J. Biol. Chem.* An optional note. May 2020.

[3] CSSEGISandData. *CSSEGISandData/COVID-19*. original-date: 2020-02-04T22:03:53Z. 2nd June 2021. URL: https://github.com/CSSEGISandData/COVID-19 (visited on 2nd June 2021).

[4] CDC. *Labs*. Centers for Disease Control and Prevention. 11th Feb. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/lab/index.html (visited on 2nd June 2021).

[5] Biology with Animations. *Coronavirus Test: Real time RT-PCR - Animation video*. 2020. URL: https://www.youtube.com/watch?v=ThG_02miq-4 (visited on 2nd June 2021).

[6] *MagNA Pure 96 Instrument*. URL: https://lifescience.roche.com/en_fr/products/magna-pure-96-instrument-382411-1.html#documents (visited on 26th May 2021).

[7] Sonja Berensmeier. 'Magnetic particles for the separation and purification of nucleic acids'. In: *Appl Microbiol Biotechnol* 73.3 (2006), pp. 495–504. ISSN: 0175-7598. DOI: 10.1007/s00253-006-0675-0. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7080036/ (visited on 26th May 2021).

[8] Germany chemagen Biopolymer-Technology AG. *magnetic bead technology*. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7080036/figure/Fig1/.

[9] *SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase*. en. URL: https://www.thermofisher.com/order/catalog/product/12574026 (visited on 26th May 2021).

[10] Willard M. Freeman, Stephen J. Walker and Kent E. Vrana. 'Quantitative RT-PCR: Pitfalls and Potential'. en. In: *BioTechniques* 26.1 (Jan. 1999), pp. 112–125. ISSN: 0736-6205, 1940-9818. DOI: 10.2144/99261rv01. URL: https://www.future-science.com/doi/10.2144/99261rv01 (visited on 26th May 2021).

[11] *Cloud-Based Informatics Platform for Life Sciences R&D*. en-US. URL: https://www.benchling.com/ (visited on 30th May 2021).

[12] Barbara Grisoni-Neupert. *Laboratoire intégré en sciences de la vie I et II*. School of Life Sciences.

[13] *General Recommendations for DNA Electrophoresis*. en. URL: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012614_Gen_Recommend_DNA_Electrophoresis_UG.pdf.

[14] *What is Real-Time PCR (qPCR)? | LSR | Bio-Rad*. URL: https://www.bio-rad.com/fr-fr/applications-technologies/what-real-time-pcr-qpcr?ID=LUSO4W8UU (visited on 4th June 2021).

[15] Alejandro Lopez-Rincon et al. 'Design of Specific Primer Set for Detection of B.1.1.7 SARS-CoV-2 Variant using Deep Learning'. en. In: *bioRxiv* (Dec. 2020). Publisher: Cold Spring Harbor

Laboratory Section: New Results, p. 2020.12.29.424715. DOI: 10.1101/2020.12.29.424715. URL: https://www.biorxiv.org/content/10.1101/2020.12.29.424715v1 (visited on 26th May 2021).

[16] 'COVID-19 RT-PCR Test - EUA Summary'. en. In: (2021), p. 30.

[17] *Inside the B.1.1.7 Coronavirus Variant - The New York Times*. URL: https://www.nytimes.com/interactive/2021/health/coronavirus-mutations-B117-variant.html (visited on 30th May 2021).

[18] *Lineage B.1.1.7*. en. Page Version ID: 1022966066. May 2021. URL: https://en.wikipedia.org/w/index.php?title=Lineage_B.1.1.7&oldid=1022966066 (visited on 26th May 2021).

[19] *primer designing tool*. en. URL: https://www.ncbi.nlm.nih.gov/tools/primer-blast/.

[20] Wenli Mu et al. 'A Flow Cytometric Method to Determine Transfection Efficiency'. In: *Bio Protoc* 9.10 (May 2019). ISSN: 2331-8325. DOI: 10.21769/BioProtoc.3244. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7449265/ (visited on 30th May 2021).

[21] Michael Maroli. 'Molecular Biology of the Cell 6th Edition - Copy 11.32.40 PM'. en. In: (). URL: https://www.academia.edu/36686286/Molecular_Biology_of_the_Cell_6th_Edition_Copy_11_32_40_PM (visited on 30th May 2021).

[22] Anne Spriestersbach et al. 'Purification of His-Tagged Proteins'. eng. In: *Methods Enzymol* 559 (2015), pp. 1–15. ISSN: 1557-7988. DOI: 10.1016/bs.mie.2014.11.003.

[23] *Ni-NTA Agarose*. en. URL: https://www.qiagen.com/us/products/discovery-and-translational-research/protein-purification/tagged-protein-expression-purification-detection/ni-nta-agarose/ (visited on 26th May 2021).

[24] *ExPASy - Compute pI/Mw tool*. URL: https://web.expasy.org/compute_pi/ (visited on 26th May 2021).

[25] *Isothermal titration calorimetry*. en. Page Version ID: 1017848023. Apr. 2021. URL: https://en.wikipedia.org/w/index.php?title=Isothermal_titration_calorimetry&oldid=1017848023 (visited on 30th May 2021).

[26] *What are protein subunit vaccines and how could they be used against COVID-19?* en. URL: https://www.gavi.org/vaccineswork/what-are-protein-subunit-vaccines-and-how-could-they-be-used-against-covid-19 (visited on 26th May 2021).

[27] *Volume 580 Issue 7805, 30 April 2020*. en. ISSN: 1476-4687. URL: https://www.nature.com/nature/volumes/580/issues/7805 (visited on 26th May 2021).

[28] *How does the SARS-CoV-2 E484K mutation increase viral fitness?* en. Section: Medical News. Feb. 2021. URL: https://www.news-medical.net/news/20210219/How-does-the-SARS-CoV-2-E484K-mutation-increase-viral-fitness.aspx (visited on 26th May 2021).

[29] | *bioRxiv*. URL: https://www.biorxiv.org/node/ (visited on 2nd June 2021).

[30] Wei Bu Wang et al. 'E484K mutation in SARS-CoV-2 RBD enhances binding affinity with hACE2 but reduces interactions with neutralizing antibodies and nanobodies: Binding free energy calculation studies'. en. In: *bioRxiv* (Feb. 2021). Publisher: Cold Spring Harbor Laboratory Section: New Results, p. 2021.02.17.431566. DOI: 10.1101/2021.02.17.431566. URL: https://www.biorxiv.org/content/10.1101/2021.02.17.431566v1 (visited on 26th May 2021).

[31] CDC. *Coronavirus Disease 2019 (COVID-19)*. Centers for Disease Control and Prevention. 11th Feb. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant.html (visited on 2nd June 2021).

[32] 'Covid-19: Israel records no daily deaths for the first time in 10 months'. In: *BBC News* (24th Apr. 2021). URL: https://www.bbc.com/news/world-middle-east-56868383 (visited on 2nd June 2021).

[33] Hannah Ritchie et al. 'Coronavirus Pandemic (COVID-19)'. In: *Our World in Data* (26th May 2020). URL: https://ourworldindata.org/coronavirus/country/israel (visited on 2nd June 2021).