

Communication



Prediction of SARS-CoV-2 Omicron Variant Immunogenicity, Immune Escape and Pathogenicity, through the Analysis of Spike Protein-Specific Core Unique Peptides

Evangelos Kontopodis^{1,2}, Vasileios Pierros¹, Dimitrios J. Stravopodis² and George T. Tsangaris^{1,*}

- ¹ Proteomics Research Unit, Biomedical Research Foundation of the Academy of Athens, 11527 Athens, Greece; kontopodisv@hotmail.gr (E.K.); pierrosv@gmail.com (V.P.)
- ² Section of Cell Biology and Biophysics, Department of Biology, School of Science, National and Kapodistrian University of Athens, 15701 Athens, Greece; dstravop@biol.uoa.gr
- * Correspondence: gthtsangaris@bioacademy.gr; Tel.: +30-210-659-7075

Abstract: The recently discovered Omicron variant of the SARS-CoV-2 coronavirus has raised a new, global, awareness. In this study, we identified the Core Unique Peptides (CrUPs) that reside exclusively in the Omicron variant of Spike protein and are absent from the human proteome, creating a new dataset of peptides named as SARS-CoV-2 CrUPs against the human proteome (C/H-CrUPs), and we analyzed their locations in comparison to the Alpha and Delta variants. In Omicron, 115 C/H-CrUPs were generated and 119 C/H-CrUPs were lost, almost four times as many compared to the other two variants. At the Receptor Binding Motif (RBM), 8 mutations were detected, resulting in the construction of 28 novel C/H-CrUPs. Most importantly, in the Omicron variant, new C/H-CrUPs carrying two or three mutant amino acids were produced, as a consequence of the accumulation of multiple mutations in the RBM. These C/H-CrUPs could not be recognized in any other viral Spike variant. Our findings indicated that the virus binding to the ACE2 receptor is facilitated by the herein identified C/H-CrUPs in contact point mutations and Spike cleavage sites, while the immunoregulatory NF9 peptide is not detectably affected. Thus, the Omicron variant could escape immune-system attack, while the strong viral binding to the ACE2 receptor leads to the highly efficient fusion of the virus to the target cell. However, the intact NF9 peptide suggests that Omicron exhibits reduced pathogenicity compared to the Delta variant.

Keywords: core unique peptide; COVID-19; immune escape; infectiveness; mutation; Omicron variant; pathogenicity; SARS-CoV-2; Spike protein; Uniquome

1. Introduction

The SARS-CoV-2 virus has a high mutagenesis frequency, hitherto producing 63 different variants with 39 considered as the most predominant forms, including Delta, the dominant variant of the 4th pandemic wave [1]. Recently, a new variant, Omicron (B.1.1.529), was identified in South Africa. Omicron is characterized by 30 amino acid changes, three small deletions, and one small insertion in Spike protein, as compared to the original virus, with 15 of them residing in the Receptor Binding Domain (RBD) from 319 to 541 amino acid residues [2].

In our previous studies, we have defined as Unique Peptides (UPs) the peptides whose amino acid sequence appears only in one protein across a given proteome. We also introduced the term of Core Unique Peptides (CrUPs), which are the peptides with a minimum amino acid sequence length that appear only in one protein across a given proteome, thus having a unique signature for a particular protein identification [3]. Therefore, each peptide of any size that contains a CrUP is considered a UP. Peptides of bigger sizes than CrUPs being constructed by continuous CrUPs are considered as Composite Core Unique Peptides (CmUPs). Hitherto, our results regarding the analysis of CrUPs in



Citation: Kontopodis, E.; Pierros, V.; Stravopodis, D.J.; Tsangaris, G.T. Prediction of SARS-CoV-2 Omicron Variant Immunogenicity, Immune Escape and Pathogenicity, through the Analysis of Spike Protein-Specific Core Unique Peptides. *Vaccines* 2022, 10, 357. https://doi.org/10.3390/ vaccines10030357

Academic Editor: Vasso Apostolopoulos

Received: 21 January 2022 Accepted: 22 February 2022 Published: 24 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). different species and organisms strongly suggest that CrUPs constitute a concrete group of peptides within a given proteome, with specialized properties and functions Thereby, we have introduced the new term "Uniquome", which is defined as the total set of UPs belonging to a given proteome and serving as its unique molecular signature. Hence, to map the UP landscape of a proteome under examination, we have herein developed a novel and advanced bioinformatics tool, including big data analysis, and we have applied this tool for the analysis of Uniquome typifying all model organisms. In *Homo sapiens*, the analysis of the 20,430 reviewed proteins resulted in the identification of 7,263,888 CrUPs which construct the human Uniquome (hUniquome) ([3] and Kontopodis et al., 2022 (manuscript in preparation)).

Most importantly, to elucidate SARS-CoV-2 virus-host organism interactions, we have further designed a novel bioinformatics platform to analyze the Core Unique Peptides (CrUPs) of the SARS-CoV-2 virus against the human proteome (C/H-CrUPs) [1]. C/H-CrUPs represent a completely new set of peptides, which are the shortest in length peptides in a viral proteome that do not exist in the human proteome [3]. Based on their properties, the viral C/H-CrUPs could advance our knowledge regarding virus-host interactions, immune system response(s), and infectiveness and pathogenicity of the virus. Moreover, most importantly, they can be used as antigenic and diagnostic peptides, and likely druggable targets for successful therapeutic treatments.

In the present study, we have identified, cataloged, and analyzed Omicron-specific C/H-CrUPs in order to illuminate the mechanisms controlling infectivity, immune escape, and pathogenicity of the new variant.

2. Materials and Methods

2.1. Methods

In our previous, recent studies, we developed a bioinformatics tool that can extract the Core Unique Peptides (CrUPs) from a given proteome, thus creating its Uniquome (Figure S1) [1,3]. We have expanded this tool by introducing a new feature that can extract the CrUPs of each individual protein of a given proteome (target) versus the proteins of a reference proteome. This new feature, like the initial implementation, will split each protein in the target proteome to all possible peptides of length minimum (4 amino acids) to length maximum (100 amino acids), and search them against the reference proteome. Each search will exclude all peptides that contain a smaller peptide already identified as CrUP (Figure S2).

For the present study, we have engaged this new feature of our tool. We created a "custom" proteome consisting of sequences from all variants of the SARS-CoV-2 Spike proteins and used it as the target versus the human proteome. The tool produces as output the C/H-CrUPs per protein of the target proteome, thus revealing the CrUPs for each Spike variant versus the human proteome.

Once we obtained the desired data, we ran a meta-analysis to identify how many C/H-CrUPs remained the same, or were added or lost on each variant versus the wild-type Spike protein. For this analysis, initially we took the identified C/H-CrUPs of the wild-type sequence and checked their presence against the respective C/H-CrUPs of the other variants. We only cared for the amino acid sequence and not the position this could be found within the protein. If the sequence was found, then we considered the peptide to be the same, otherwise we considered it to be lost on the examined variant. Next, we analyzed the identified C/H-CrUPs of each variant versus the wild-type sequence. If the peptide was detected only on the variant's C/H-CrUPs, then we considered it as added. This meta-analysis also provided us with the position of each C/H-CrUP within the Spike protein, which we used to determine the area (e.g., RBD, RBM and S-cleavage site, as obtained by the Stanford COVID-19 Database) they resided in.

2.2. Databases

All proteomes and proteins were obtained from Uniprot. SARS-CoV-2 wild-type and variant sequences, and mutations were obtained from the Stanford COVID-19 Database (https://covdb.stanford.edu/page/mutation-viewer/, accessed on 23 December 2021).

3. Results and Discussion

3.1. Mapping the C/H-CrUPs Landscape of Spike Protein of the SARS-CoV-2 Omicron Variant

SARS-CoV-2 virus seems to be highly mutated, so far producing more than 60 distinct variants. Hitherto, the highest pathogenic form is the Delta variant (B.1.617.2), with 10 different sub-variants. Recently, a novel variant called Omicron has been identified. It is characterized by 30 amino acid changes, three small deletions, and one small insertion in the Spike protein area, as compared to the wild-type viral respective sequence (Figure S3) [2]. Out of these genetic changes, 15 reside in the Receptor Binding Domain (RBD) from amino acid position 318 to 541, and two are located around the S-cleavage site(s) (Figure S3).

Advanced bioinformatics analysis of the Omicron variant Spike protein showed that it contains 983 C/H-CrUPs, a number that is comparable to the one of wild-type Spike proteins (987 C/H-CrUPs) and to the mean \pm SD value of Spike protein-specific C/H-CrUPs (983 \pm 2 C/H-CrUPs) (Table 1). Omicron variant Spike protein contains 34 mutations in total, which is the highest number of identified mutations among all virus variants.

Table 1. SARS-CoV-2 Spike protein C/H-CrUPs across variants, as compared to the wild-type virus respective sequence.

	Spike Protein								
Variant	C/H-CrUPs	Same C/H-CrUPs	% of Same C/H-CrUPs	New C/H-CrUPs	% of New C/H-CrUPs	Lost C/H-CrUPs	% of Lost C/H-CrUPs		
Wild-type virus	987								
Alpha (B.1.1.7) + (Q1-Q4)	982	931	94.8	51	5.2	56	5.7		
Alpha (B.1.1.7 + E484K)	983	928	94.4	55	5.6	59	6.0		
Alpha (B.1.1.7 + L452R)	981	936	95.4	45	4.6	51	5.2		
Alpha (B.1.1.7 + S494P)	981	936	95.4	45	4.6	51	5.2		
Beta (B.1.351)	981	954	97.2	27	2.8	33	3.3		
Beta (B.1.351 + E516Q)	981	949	96.7	32	3.3	38	3.8		
Beta (B.1.351 + L18F) (B.1.351.2-3)	979	948	96.8	31	3.2	39	3.9		
Beta (B.1.351 + P384L)	980	949	96.8	31	3.2	38	3.9		
Gamma (P.1) (P.1.1 - P.1.2)	985	930	94.4	55	5.6	57	5.8		
Gamma (P1 + P681H)	985	930	94.4	55	5.6	57	5.8		
Delta (B.1.617.2)	984	948	96.3	36	3.7	39	4.0		
Delta (B.1.617.2 + E484Q)	984	945	96.0	39	3.4	42	4.3		
Delta (B.1.617.2 + K417N)	984	944	95.9	40	4.1	43	4.4		
Delta (B.1.617.2 + Q613H)	984	947	96.2	37	3.8	40	4.1		
Delta (AY.1)	984	944	95.9	40	4.7	43	4.1		
Delta (AY.2)	985	939	95.3	46	4.8	48	4.9		
Delta (AY.3 - AY.8) + (AY.12)	983	951	96.7	32	3.3	36	3.7		
Delta (AY.9)	983	951	96.7	32	3.3	36	3.6		
Delta (AY.10)	983	951	96.7	32	3.3	36	3.6		

	Spike Protein								
Variant	C/H-CrUPs	Same C/H-CrUPs	% of Same C/H-CrUPs	New C/H-CrUPs	% of New C/H-CrUPs	Lost C/H-CrUPs	% of Lost C/H-CrUPs		
Delta (AY.11)	983	951	96.7	32	3.3	36	3.6		
Eta (B.1.525)	990	956	96.5	34	3.4	31	3.1		
Iota (B.1.526)	984	960	97.5	24	2.4	27	2.7		
Kappa (B.1.617.1)	985	964	97.8	21	2.1	23	2.3		
Lambda (C.37)	982	949	96.6	33	3.4	38	3.9		
Mu (B.1.621)	983	953	96.9	30	3.1	34	3.4		
Omicron (B.1.1.529)	983	868	88.3	115	11.7	119	12.1		

Table 1. Cont.

New C/H-CrUPs is the number of new constructed peptides of each variant compared to C/H-CrUPs of wild-type virus; % of new C/H-CrUPs is the % of new constructed peptides compared to the total C/H-CrUPs number of each variant; Lost C/H-CrUPs is the number of peptides lost in each variant compared to C/H-CrUPs of wild-type virus; % of lost C/H-CrUPs is the % of lost peptides compared to the total C/H-CrUPs number of each variant.

These mutations seem to have a dramatic effect on the Spike protein C/H-CrUPs map. Compared to the wild-type Spike sequence, we found that 115 (new) C/H-CrUPs were created and 119 C/H-CrUPs were lost, almost twice as many when compared to the Alpha variant (51 and 56 C/H-CrUPs, respectively), and almost four times as many, compared to the other variants (Table 1). The distribution of these new C/H-CrUPs shows that the majority carry 6 amino acids in length (Figure S4).

3.2. Omicron-Specific C/H-CrUPs That belong to the Receptor Binding Domain

SARS-CoV-2 belongs to the β coronavirus group, which uses the plasma membrane receptor of Angiotensin-Converting Enzyme 2 (ACE2) to recognize and bind to the target cell [4]. The viral Spike protein attaches to ACE2 receptor by a Receptor Binding Domain (RBD) defined from amino acid position F318 to F541 [4,5]. The amino acid residues from W436 to Q506 inside RBD shape the Receptor Binding Motif (RBM), which carries 11 contact positions with ACE2 [5]. The RBD region has received great attention, as it seems to be a major target of antibodies against the virus and other therapeutic interventions [6–8].

In the RBD region, the Omicron variant carries 15 mutations, 10 of which are identified in the RBM area (Figure 1A). This results in the identification of the highest number of newly constructed C/H-CrUPs in the RBD/RBM region, as compared to all other previous virus variants examined (Table S1). Table 2 describes all the new, herein identified, C/H-CrUPs of Omicron variant in Spike's RBD region, in comparison to the Alpha and Delta variants, which represent two of the most predominant variants of the virus in human populations. Hence, it was proven that, in contrast to Alpha and Delta variants, at the end of Omicron variant RBM area from 440 to 508 amino acid position, 8 novel mutations were identified, resulting in the production of 28 new C/H-CrUPs. The most important finding is that in Omicron variant, for the first time, new C/H-CrUPs including two or three mutant amino acids were generated, with the peptides "QAGN*K*P", "N*K*PCN", "LK*SYS*F" and "K*SYS*FR*" being characteristic examples, as a result of the accumulation of multiple mutations in the positions 440, 446, 477, 478 and 493–505. These novel C/H-CrUPs that contain several mutated amino acids could not be found in any other virus variants previously. Taking into consideration recent data about virus infectivity, the multimutated, new, C/H-CrUP collection seems to radically change the structure and the epitope regions of end positions of the RBM area in the Omicron variant, causing a serious compromise of its antigenic capacity and facilitating the immune escape of the virus [9].

Δ	RECEPTOR BINDING DOMAIN						
<i>/</i> \	RECEPTOR BINDING MOTIE	NE9 PEPTIDE		RECEPTOR BINDING MOTIF			
Variant / AA Desition							
SARS-CoV-2 wildtyne	436 437 438 439 440 441 442 443 444 445 446 W N S N N I D S K V G	G N V N V I V B I F	8 K S N I K P F F B D I S T F I Y O A G	147/1478 1479 1480 1481 1482 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 500 1500 500 1500 500 1500 500 1500 500	V 0		
Alpha (B.1.1.7 & Q1-Q4)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVEGFNCYFPLQPYGFQPTYGVGY	Y Q		
Alpha (B.1.1.7+E484K)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGV <mark>K</mark> GFNCYFPLQ <mark>P</mark> YGFQPT <mark>Y</mark> GVGY	Y Q		
Alpha (B.1.1.7+L452R)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVEGFNCYFPLQ <mark>S</mark> YGFQPT <mark>Y</mark> GVGY	γQ		
Alpha (B.1.1.7+S494P)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	S T P C N G V E G F N C Y F P L Q P Y G F Q P T Y G V G Y	YQ		
Beta (B.1.351) Beta (B.1.351+E516O)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVKGFNCTFPLQSTGFQPTTGVGT	YO		
Beta (B.1.351+L18F & B.1.351.2-3)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVKGFNCYFPLQSYGFQPTYGYGY	Y Q		
Beta (B.1.351+P384L)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGV <mark>K</mark> GFNCYFPLQSYGFQPT <mark>Y</mark> GVGY	Y Q		
Gamma (P.1 & P.1.1 - P.1.2)	W N S N N L D S K V G	GNYNYLYRLF	R K S N L K P F E R D I S T E I Y Q A G	STPCNGV <mark>K</mark> GFNCYFPLQSYGFQPT <mark>Y</mark> GVGY	Y Q		
Gamma (P1+P681H)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVKGFNCYFPLQSYGFQPTYGVGY	Y Q		
Delta (B.1.617.2) Delta (B.1.617.2+F484O)	W N S N N L D S K V G W N S N N I D S K V G	G N Y N Y R Y R I F	R K S N L K P F F R D I S T F I Y O A G	SKPCNGVEGFNCTFPLQSTGFQPINGVGT	Y Q		
Delta (B.1.617.2+K417N)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	SKPCNGVEGFNCYFPLQSYGFQPTNGVG	Y Q		
Delta (B.1.617.2+Q613H)	W N S N N L D S K V G	G N Y N Y <mark>R</mark> Y R L F	R K S N L K P F E R D I S T E I Y Q A G	S K P C N G V E G F N C Y F P L Q S Y G F Q P T N G V G Y	Y Q		
Delta (AY.1)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	SKPCNGVEGFNCYFPLQSYGFQPTNGVGY	YQ		
Delta (AY.2)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	S K P C N G V E G F N C Y F P L Q S Y G F Q P T N G V G Y	Y Q		
Delta (AY.9)	W N S N N L D S K V G		R K S N L K P F F R D I S T F I Y O A G	SKPCNGVEGENCTEPLQSTGEQPINGVGT	v o		
Delta (AY.10)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	SK PCNGVEGFNCYFPLQSYGFQPTNGVGY	Y Q		
Delta (AY.11)	W N S N N L D S K V G	G N Y N Y R Y R L F	R K S N L K P F E R D I S T E I Y Q A G	SKPCNGVEGFNCYFPLQSYGFQPTNGVGY	Y Q		
Eta (B.1.525)	W N S N N L D S K V G	GNYNYLYRLF	R K S N L K P F E R D I S T E I Y Q A G	STPCNGV <mark>K</mark> GFNCYFPLQSYGFQPTNGVGY	γQ		
lota (B.1.526)	W N S N N L D S K V G	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVKGFNCYFPLQSYGFQPTNGVGY	Y Q		
Lambda (C.37)	W N S N N L D S K V G	G N Y N Y O Y R L F	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVEGFNCTFPLQSTGFQPINGVGT	YO		
Mu (B.1.621)	W N S N N L D S K V G	GNYNYLYRLF	R K S N L K P F E R D I S T E I Y Q A G	STPCNGVKGFNCYFPLQSYGFQPTYGVGY	YQ		
OMIKRON (8.1.1.529	W N S N K L D S K V S	G N Y N Y L Y R L F	R K S N L K P F E R D I S T E I Y Q A G	NKPCNGVAGFNCYFPLKSYSFRPTYGVGH	H Q		
R	S1	Cleavage cite					
D			S2 Cleavage site	S3 Cleavage site 👻			
D	S1 DOMAIN	BRIDGE	S2 Cleavage site	S3 Cleavage site			
Variant / AA Position	S1 DOMAIN	BRIDGE 79 680 681 682 683 684 685 686 687 68	S2 Cleavage site S2 DOMAIN 18 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706	S3 Creavage site S2 DOMAIN 707 706 700 710 110 111 111 111 111 111 111 111	19 830		
Variant / AA Position SARS-CoV-2 wildtype	S1 DOMAIN 670 671 672 673 674 675 676 677 678 67 I C A S Y Q T Q T	BRIDGE 79 680 681 682 683 684 685 686 687 68 N S P R R A R S V A	S2 Cleavage site S2 S2 DOMAIN S2 88 89 000 691 892 693 696 697 698 699 700 701 702 703 705 705 705 705 705 705 705 705 705 705	S3 LieaVaĝesite ♥ S2 DOMAIN 1207 /201 /205 /205 /205 /201 /201 /201 /201 /201 /201 /201 /201	19 830 A D		
Variant / AA Position SARS-COV-2 wildtype Alpha (8.1.1.7 & Q1-Q4) Alpha (1.1.7 & R4K)	S1 DOMAIN 670 671 672 673 674 675 676 677 678 67 I C A S Y Q T Q T I C A S Y Q T Q T I C A S Y Q T Q T	BRIDGE 79 680 682 684 686 686 687 68 N S P R R A R S V A N S P R R A R S V A N S P R R A R S V A N S P R R A R S V A N S R R A R S V A	S2 Cleavage stre \$2 DOMAIN Is an ose one (no) (no) (no) (no) (no) (no) (no) (no)	S2 DOMAIN 700 <th>19 830 A D A D</th>	19 830 A D A D		
Variant / AA Position SARS-CoV-2 wildtype Alpha (8.1.1.7 & Q1-Q4) Alpha (8.1.1.7t4528) Alpha (8.1.1.7t4528)	S1 DOMAIN 670 671 672 673 674 675 667 677 678 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T	BRIDGE 79 680 681 682 683 684 685 666 687 68 N S P R A R V A N S H R A R V A N S H R A R V A N S H R A R V A	S2 Cleavage stre \$2 OMAIN \$8 689 690 691 691 702 702 703 705 70	S3 LiceVogeSite S3 707 708 709 709 709 709 7 5 N N 5 K P S K R S F I E D L L F N K V T L A 5 K P S K R S F I E D L L F N K V T L A 5 K P S K R S F I E D L L F N K V T L A 5 K P S K R S F I E D L L F N K V T L A 5 K P S K R S F I E D L L F N K V T L A	19 830 A D A D A D		
Variant / AA Position SAR5-CoV-2 wildtype Alpha (8.1.1.7 & Q1-Q4) Alpha (8.1.1.7+E484K) Alpha (8.1.1.7+5494P) Alpha (8.1.1.7+5494P)	S1 DOMAIN 670 671 672 673 674 675 677 678 6 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T	BRIDGE BRIDGE 9 680 681 682 683 684 683 686 687 688 688 685 686 687 688 N S P R A R S V A N S P R A R S V A N S H R A R S V A N S H R A R S V A N S H R A R S V A N S H R R R S V A N S H R R R S V A	S2 Cleavage str S2 52 DOMAIN 16 450 000 001 (002 005 004 (002 004 (002 004 007 004 002 004 007 004 002 004 007 004 004 004 004 004 004 004 004	So Cleavage site So Cleavage site<	19 830 A D A D A D A D A D		
Variant / AA Position SARS-CoV-2 wildtype Alpha (8.1.1.7 & Q1-Q4) Alpha (8.1.1.7+E454K) Alpha (8.1.1.7+E454K) Alpha (8.1.1.7+E454K) Bits (8.1.351) Beta (8.1.351)	S1 DOMAIN 670 671 672 673 676 677 678 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T	P R A R S V N S P R A R S V N S P R A R S V N S P R A R S V N S H R A R S V N S H R A R S V N S H R A R S V N S H R A R S V N S H R A R S V	S2 Cleavage stre \$2 00MaiN 8 686 607 668 607 701 702 701 701 701 701 701 70	Society construction Society c	19 830 A D A D A D A D A D A D A D		
Variant / AA Position SABS-CoV-2 wildtype Alpha (8.1.1.7 & 6.1-0.4) Alpha (8.1.1.7 + 6.434K) Alpha (8.1.1.7 + 6.434K) Alpha (8.1.1.7 + 6.434P) Beta (8.1.551 + 6.551 +	S1 DOMAIN 670 671 672 671	BRIDGE BRIDGE 79 680 61 647 687 647 687 647 </th <th>S2 LiceWage Ste S2 52 DOMAIN 16 600 601 602 604 607 608 607 702</th> <th>So Cleavage site So Cleavage site<</th> <th>19 830 4 D 4 D 4 D 4 D 4 D 4 D 4 D</th>	S2 LiceWage Ste S2 52 DOMAIN 16 600 601 602 604 607 608 607 702	So Cleavage site So Cleavage site<	19 830 4 D 4 D 4 D 4 D 4 D 4 D 4 D		
Variant / AA Position SABS-CoV-2 widtype Alpha (8.1.17 & 40.1-04) Alpha (8.1.17+64444) Alpha (8.1.17+64528) Alpha (8.1.17+554847) Beta (8.1.351+65560) Beta (8.1.351+65560) Beta (8.1.351+61560)	S1 DOMAIN 670 671 672	Picket 6x2 6x3 6x6 6x5 6x6 6x7 6x7 Picket 6x3 6x5 6x6 6x5 6x6 6x7 6x7 N S P R A R S V A N S H R A R S V A N S H R A R S V A N S H R A R S V A N S H R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S	S2 Cleavage ste \$2 00Main 8 68 695 666 697 668 697 702 702 702 704 705 706 705 702 702 703 705 706 705 702 702 702 702 702 703 705<	Society approximation Society approximatis approximation Socis approximation </th <th>19 830 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D</th>	19 830 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D		
Variant / AA Position SAR5-CoV-2 wildtype Alpha (8.1.17 & Q1-Q4) Alpha (8.1.17+4584) Alpha (8.1.17+4584) Beta (8.1.351) Beta (8.1.351)	S1 DOMAIN cm gcr (cr) cr) crister (cr) crister (cr	Clearwaye size W Intermediate REIDGE Intere REIDGE	S2 Cleavage stre S2 52 Observation	So Cleavage site So Cleavage site 207 700 700 700 700 700 700 3 5 10 82 81 8 10 82 81 824 815 814 817 818 817 818 817 820 82 82 823 824 825 824 827 828 825 3 5 10 81 82 81 8 82 81 824 815 814 817 818 817 820 82 82 823 824 825 824 827 828 825 3 5 10 81 82 81 8 82 81 824 815 814 817 818 817 820 82 82 823 824 827 828 827 828 825 3 5 10 81 823 813 814 817 810 820 82 823 824 825 824 827 828 825 3 5 10 81 823 813 814 817 810 820 82 823 822 823 824 827 828 825 3 5 10 81 823 813 814 817 810 820 82 823 824 823 824 825 824	19 830 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 4 D 5 D 4 D		
Variant / AA Position SAB5.cov.2 withype Alpha (6.1.17.8 Q.1.04) Alpha (6.1.17.45244) Alpha (6.1.17.45248) Alpha (6.1.17.45248) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160) Bets (0.1.35145160)	S1 DOM AIN 670 671 672 <th>R R</th> <th>S2 Cleavage stre S2 B 600 607 608 608</th> <th>Society constraints Society cocociety constraints Society constra</th> <th>19 830 A D</th>	R R	S2 Cleavage stre S2 B 600 607 608 608	Society constraints Society cocociety constraints Society constra	19 830 A D		
Variant / AA Position 5845-CoV-2 wildtype Alpha (6.1.17-8 Q1-Q4) Alpha (6.1.17-8 Q1-Q4) Alpha (6.1.17-14528) Alpha (6.1.17-14528) deta (6.1.351-15449) deta (6.1.351-15449) deta (6.1.351-15449) deta (6.1.351-15449) deta (6.1.351-15449) deta (6.1.351-15449) deta (6.1.351-1549) deta (6.1.351-	S1 DOMAIN org org (zz) zz) zzy szy brzy brzy brzy brzy brzy brzy brzy br	Clearwaye size W Intermediate R <th>S2 Cleavage stre S2 S2 Cleavage stre S S</th> <th>So Cleavage site So Cleavage site 207 708 709 720 724 72</th> <th>19 830 A D</th>	S2 Cleavage stre S2 S2 Cleavage stre S S	So Cleavage site So Cleavage site 207 708 709 720 724 72	19 830 A D		
Variant / A A Position SAB.5-07-2 withype Alpha (B. 11.7 & 0.1-0.4) Alpha (B. 11.7 & 0.1-0.4) Alpha (B. 11.7 + 0.1484) Alpha (B. 11.7 + 0.1584) Baby (B. 11.7 + 0.1584) Betz (B. 1.331+1.18F & 8.1.351+2.18 Betz (B. 1.331+1.18F & 8.1.351+2.18 Betz (B. 1.331+1.18F & 8.1.351+2.18 Gamma (P.1 + 0.141+7.12) Gamma (P.1 + 0.141+7.12) Gamma (P.1 + 0.141+7.12) Deta (B. 1.637.2) Deta (B. 1.637.2) Deta (B. 1.637.2)	S1 DOMAIN c00 crit cri crit crit	P R	S2 Cleavage stre S2 B 600 600 600 600 600 700 701 702 703 705	So Cleavage site So Cleavage site 200 700<	19 830 4 D		
Variant 1 / AA Position SAS.Col-2 widtype Alpha (e.1.17-4.02-44) Alpha (e.1.17-4.03-44) Alpha (e.1.17-4.03-47) Barta (e.1.33-11 Barta (e.1.33-11 Barta (e.1.33-11 Barta (e.1.33-11) Barta (e.1.33-11 Barta (e.1.33-11) Barta (e.1.33-11 Barta (e.1.33-11) Barta (e.1.33-11 Barta (e.1.33-11) Barta (e.1.33-11 Barta (e.1.33-11) Barta (e.1.33-11) Barta (e.1.33-11) Barta (e.1.33-11) Data (e.1.63-12-444(e)) Data (e.1.63-12-444(e))	S1 DOMAIN org gr1 gr2 gr3	Clearwaye site W Pices R R R R S Pices R R R S Pices R R R S V R R R S V R R R S V R R R S V R R R S V R R R R S V R R R R S V R R R R S V R R R R R S V R R R R R S V R R R R R S V A R S V A R S V A R S V A R S V A R S V A R A R S V A </th <th>S2 Cleavage stre S2 1000 001 002 001 004 004 001 002 001 <td0< th=""><th>So Cleavage site So Cleavage site 207 708 709 720 724 72</th><th>19 830 4 D</th></td0<></th>	S2 Cleavage stre S2 1000 001 002 001 004 004 001 002 001 <td0< th=""><th>So Cleavage site So Cleavage site 207 708 709 720 724 72</th><th>19 830 4 D</th></td0<>	So Cleavage site So Cleavage site 207 708 709 720 724 72	19 830 4 D		
Variant 1 / AA Position SAS.Cov2 wildyge Alpha (b. 11-74 dot-40) Alpha (b. 11-74 dot-40) Alpha (b. 11-74 dot-40) Alpha (b. 11-74 dot-40) Alpha (b. 11-74 dot-40) Beta (b. 133) Beta (b. 133) Beta (b. 133) Defa (b. 147) Beta (b. 147) Aleda) Defa (b. 1617) Defa (b. 1617)	S1 DOMAIN c00 crit cri crit crit	P R R A R S 000 79 60 633 643 643 663 666 67 68 79 60 633 643 643 643 643 644 653 646 67 68 N S P R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S V A N S P R A R S V A N S R R	S2 Display Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage site So Cleavage site<</th><th>19 830 4 D</th></thdispla<></thdisplay<></thdisplay<>	So Cleavage site So Cleavage site<	19 830 4 D		
Variant 1 / AA Position SAS-Cor-2 widtype Alpha (e. 1.17-4 02-04) Alpha (e. 1.17-4 02-04) Alpha (e. 1.17-4 02-04) Alpha (e. 1.17-4 02-04) Beta (e. 1.38) Beta (e. 1.38) Deta (e. 1.617.2-164840) Deta (e. 1.617.2-164870) Deta (e. 1.617.2-164870) Deta (e. 1.617.2-164870) Deta (e. 1.617.2-164870)	S1 DOM AIN cm cr2 cr2 <th>Clearwaye site W 191 00 01 02 03</th> <th>S2 Cleavage stret S2 10 00 001 002 001 001 002 001 001 002 001 001 002 001 001 002 001 00</th> <th>So Cleavage site So Cleavage site 20 708 709 720 736 740 75 71 71 75 71 75 71 75 71 75 71 71 75 71 71 71 71 71 71 71 71 71 71 71 71 71 71</th> <th>19 830 4 D</th>	Clearwaye site W 191 00 01 02 03	S2 Cleavage stret S2 10 00 001 002 001 001 002 001 001 002 001 001 002 001 001 002 001 00	So Cleavage site So Cleavage site 20 708 709 720 736 740 75 71 71 75 71 75 71 75 71 75 71 71 75 71 71 71 71 71 71 71 71 71 71 71 71 71 71	19 830 4 D		
Variant 1 / AA Position SAS.Cov2 wildyge Alpha (b.1)-7 40:404 Alpha (b.1)-7 40:404 Alpha (b.1)-7 40:404 Outing (b.1)-7 20:404 Dotting (b.1)-7 20:404 Dotting (b.1)-7 20:401 Dotting (b.1)-7 20:401	S1 DOMAIN c00 cr cr2	P R R A R S G2 G3 G42 G43 G43 <thg43< th=""> <thg43< th=""> G43</thg43<></thg43<>	S2 Display Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage site So Cleavage site<</th><th>H9 830 A D</th></thdispla<></thdisplay<></thdisplay<>	So Cleavage site So Cleavage site<	H9 830 A D		
Variant 1 / AA Position SAS-Col-2 widtype Alpha (0.1.17-6.02-04) Alpha (0.1.17-4648) Alpha (0.1.17-4648) Alpha (0.1.17-4648) Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.351- Beta (0.1.51- Deta (0.1.51- De	S1 DOMAIN or0 6371 c22 637 64 637	Clearwaye site W 1000 01 02	S2 Cleavage stret S2 10 671 672 672 722 7	So Cleavage store So Cleavage store 20	19 830 4 D		
Variant / AA Position SAS.Cov2 wildyge Alpha (b.1):74 (b.1):40 Alpha (b.1):74 (b.1):40 Alpha (b.1):74 (b.1):40 Deta (b.1):74 (b.1):74 (b.1):11 Alpha (b.1):74 (b.1):11 Deta (b.1):74 (b.1):11 Alpha (b.1):74 (b.1):11 Deta (b.1):74 (b.1):11 Alpha (b.1):11 Deta (b.1):74 (b.1):11 Alpha (b.1):11 Deta (b.1):74 (b.1):11 Alpha (b.1):11 Deta (b.1):11 <	S1 DOMAIN c00 cr3 cr2 cr3 cr3 <td< th=""><th>Pierre Bior <</th><th>S2 Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage set e So So Cleavage set</th><th>19 830 4 0</th></thdispla<></thdisplay<></thdisplay<></th></td<>	Pierre Bior <	S2 Display Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage set e So So Cleavage set</th><th>19 830 4 0</th></thdispla<></thdisplay<></thdisplay<>	So Cleavage set e So So Cleavage set	19 830 4 0		
Variant / AA Position SAS.Col-2 widtype Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Berge (b.1.17-04-02) Alpha (b.1.17-4 02-04) Berge (b.1.17-04-02) Berge (b.1.17-04-02) Berge (b.1.17-04-02)	S1 DOMAIN or0 6371 c21 c37 e36 c36 c37 c37 1 C A S Y Q Q Q 1 C A S Y Q <	Clearways R R R R R S Clearways 191 00 01 02	S2 Cleavage stret S2 10 60	So Cleavage set e So Cleavage set e 20 08 07 70	19 830 4 0		
Variant / AA Position SAS.Cov2 withype Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Alpha (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):40 Deta (b. 1):74 (b. 1):10 Deta (b. 1):10 Deta (b. 1):10 Deta (b. 1):10 Deta (b. 1):11 Deta (b. 1):10	S1 DOMAIN c00 cr3 cr2 cr3 cr3 <td< th=""><th>Pierre Bior <</th><th>S2 Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e 207 708 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 70 8 <</th><th>19 830 4 D</th></thdispla<></thdisplay<></thdisplay<></th></td<>	Pierre Bior <	S2 Display Display <thdisplay< th=""> <thdisplay< th=""> <thdispla< th=""><th>So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e 207 708 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 70 8 <</th><th>19 830 4 D</th></thdispla<></thdisplay<></thdisplay<>	So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e So Cleavage set e 207 708 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 709 70 8 <	19 830 4 D		
Variant / AA Position SAS.Col-2 widtype Alpha (b.1.7-4 QL-24) Alpha (b.1.7-4 QL-24) Alpha (b.1.7-4 QL-24) Alpha (b.1.7-4 QL-24) Alpha (b.1.7-4 QL-24) Star (b.1.7-4 QL-24) Star (b.1.7-4 QL-24) Star (b.1.7-4 QL-24) Star (b.1.7-4 QL-24) Star (b.1.7-14 QL-24) Star (b.1.7-14 QL-24) Star (b.1.7-14 Q	S1 DOMAIN or0 6371 c21 672 678 678 c77 678 c77 678 c77 678 1 C A S Y Q Q Q 1 C A S Y Q Q Q Q 1 C A S Y Q Q Q Q 1 C A S Y Q Q Q Q 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T Q T 1 C A S Y Q T	Piccolage site Piccola	S2 Cleavage stret S2 1000 001 002 001 004 001 <td< th=""><th>So Cleavage sets So Cleavage sets 20 08 07 70</th><th>19 830 4 D</th></td<>	So Cleavage sets So Cleavage sets 20 08 07 70	19 830 4 D		
Variant / AA Position SAS.Cov2 withype Alpha (b.1.17-4 0.0-0.1) Alpha (b.1.17-4 0.0-0.1) Alpha (b.1.17-4 0.0-0.1) Betz (b.1.33 0.0-0.00.1) Batz (b.1.33 0.0-0.00.1) Deta (b.1.47.2-4 0.01.1) Deta (b.1.47.2-4 0.01.1)	S1 DOMAIN c00 cri cri <td< th=""><th>Piero R<th>S2 Delevage stret S2 12 bot loss est est</th><th>So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store 10</th><th>19 830 A D</th></th></td<>	Piero R <th>S2 Delevage stret S2 12 bot loss est est</th> <th>So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store 10</th> <th>19 830 A D</th>	S2 Delevage stret S2 12 bot loss est	So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store So Cleavage store 10	19 830 A D		
Variant / AA Position SAS.Col-2 widtype Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Alpha (b.1.17-4 02-04) Berta (b.1.351) Berta (b.1.351) Berta (b.1.351) Berta (b.1.351) Berta (b.1.351-21404) Gamma (b.1.4 b.1.351-21404) Berta (b.1.351-21404) Deita (b.1.617.2-0404) Deita (b.1.617.2-0404) Deita (b.1.617.2-0404) <t< th=""><th>S1 DOMAIN cm crit cri<</th><th>Pickerseyes skr Pickerseyes skr Pickerseyeskr Pickerseyes skr Pickerseyes</th><th>S2 Cleavage stret S2 S2 Cleavage stret S S3 S <t< th=""><th>So Cleavagester v 20 08 07 70</th><th>19 830 4 D</th></t<></th></t<>	S1 DOMAIN cm crit cri<	Pickerseyes skr Pickerseyeskr Pickerseyes skr Pickerseyes	S2 Cleavage stret S2 S2 Cleavage stret S S3 S <t< th=""><th>So Cleavagester v 20 08 07 70</th><th>19 830 4 D</th></t<>	So Cleavagester v 20 08 07 70	19 830 4 D		
Variant / AP osition SAS.507-20 withype Alpha (b. 1.17-461) Alpha (b. 1.17-464) Alpha (b. 1.17-4644) Alpha (b. 1.17-4644) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46440) Alpha (b. 1.17-46401) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11-41.2) Gamma (b. 1.46-11.2) Gata (Ard.3) Gata (Ard.3) Gata (Ard.3) Gata (Ard.3)	S1 DOMAIN c00 cr3 cr2 cr3 cr3 <td< th=""><th>Image: state Image: state<</th><th>S2 Cleavage stret S2 Cleavage stret S S2 S <th< th=""><th>So Cleavage set e 100</th><th>19 330 A D</th></th<></th></td<>	Image: state Image: state<	S2 Cleavage stret S2 Cleavage stret S S2 S <th< th=""><th>So Cleavage set e 100</th><th>19 330 A D</th></th<>	So Cleavage set e 100	19 330 A D		

Figure 1. Mutations in different virus variants. **(A)** The mutations of the Receptor-Binding Motif (RBM) included in the Receptor-Binding Domain (RBD) are presented. **(B)** The mutations around the Spike cleavage sites are presented. Purple blocks mark the point mutation sites in the variants. Green colors indicate the Universal Peptides of the Spike proteins from Figure S2. Yellow colors mark the Receptor-Binding Domain of Spike protein interaction with ACE2. Pink colors mark the Receptor-Binding Motif. Cyan colors indicate the NF9 peptide, while light blue colors mark the Bridge between S1 and S2 domains. Red-colored arrows indicate the cleavage sites. With different colors in the upper side of the alignment, the different domains of the Spike protein are presented.

	-	Alpha Varia	nt		Delta Variant					
C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	C/H-Cr	UP	Position	Mutation	New C/H-CrUPs	Position
GNYNYL	447		GNYNY <mark>R</mark>	447	PGQTC	G <mark>K</mark> I	412			
NYNYLY	448				GQTG	IA	413		GQTG <mark>N</mark> I	413
		I 452D	YNY <mark>R</mark> Y	449				K417N	QTG <mark>N</mark> IA	414
NYLYRL	450	- L432K	NY <mark>R</mark> YRL	450	TG <mark>K</mark> IA	D	415		TG <mark>N</mark> IAD	415
YLYRLF	451		Y <mark>R</mark> YRLF	451	G <mark>K</mark> IAI	DΥ	416		G <mark>N</mark> IADY	416
LYRLFR	452									
					GNYN	YL	447		GNYNY <mark>R</mark>	447
CNGVEG	480		CNGV <mark>K</mark> G	480	NYNY	LY	448			
NGVEGF	481	Елели	NGV <mark>K</mark> GF	481				I 450D	YNY <mark>R</mark> Y	449
GVEGFN	482	E404N	GV <mark>K</mark> GFN	482	NYLYI	RL	450	L452K	NY <mark>R</mark> YRL	450
			K GFNC	484	YLYRI	ĴF	451		Y <mark>R</mark> YRLF	451
					LYRLF	R	452			

Table 2. C/H-CrUPs constructed aroud the mutations in RBD of Alpha, Delta and Omicron SARS-CoV-2 variants.

		Alpha Varia	nt		Delta Variant					
C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	
YFPLQ <mark>S</mark>	489		YFPLQ <mark>P</mark>	489						
FPLQ <mark>S</mark> Y	490		FPLQ <mark>P</mark> Y	490	YQAGST	473		YQAGS <mark>K</mark>	473	
PLQ <mark>S</mark> YG	491		PLQ <mark>P</mark> YG	491			· -	QAGS <mark>K</mark> P	474	
Q <mark>S</mark> YGF	493		Q <mark>P</mark> YGF	493	AGSTPC	475	T4701/	AGS <mark>K</mark> PC	475	
S YGFQP	494		P YGFQP	494			14/0K -	GS <mark>K</mark> PCN	476	
GFQPTN	496	S494P			STPCN	477				
FQPT <mark>N</mark> G	497	N501Y					-	K PCNG	478	
QPT <mark>N</mark> GV	498									
			QPTY	498						
PT <mark>N</mark> GVG	499		PT Ƴ G	499						
TNGVGY	500		Τ <mark>Υ</mark> GV	500						
NGVGYQ	501		ƳGVG	501						
				Omicro	n variant					
C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	
NLCPF <mark>G</mark>	334		NLCPFD	334	IYQAG <mark>S</mark>	472				
LCPF <mark>G</mark> E	335		LCPFDE	335	YQAG <mark>ST</mark>	473		YQAG <mark>N</mark>	473	
PF GEVF	337	G339D	PFDEV	337			S477N T478K	QAG <mark>NK</mark> P	474	
FGEVFN	338		FDEVFN	338	AG <mark>ST</mark> PC	475				
GEVFNA	339		DEVFNA	339	STPCN	477		NK PCN	477	
								K PCNG	478	
VLYN <mark>S</mark> A	367		VLYNLAP	367						
LYN <mark>S</mark> AS	368				CNGVEG	480		CNGV <mark>A</mark> G	480	
YN <mark>S</mark> ASF	369		YNLAPF	369	NGVEGF	481	F181A -	NGV <mark>A</mark> GF	481	
NSASFST	370	. S371L .			GVEGFN	482		GV <mark>A</mark> GFN	482	
		S373P	LAPFFT	371				VAGFNC	483	
ASFSTF	372	S375F	APFFTF	372						
SFSTFK	373				CYFPLQ	488		CYFPL <mark>K</mark>	488	
			F <mark>F</mark> TFK	374	YFPLQS	489		YFPL <mark>K</mark> S	489	
STFKC	375		F TFKCY	375	FPLQSY	490		FPL <mark>K</mark> SY	490	
					PLQSYG	491	O493K			
PGQTG <mark>K</mark> I	412						G496S	L <mark>K</mark> SY <mark>S</mark> F	492	
GQTG <mark>K</mark> IA	413		GQTG <mark>N</mark> I	413	Q SY G F	493	Q498R N501Y	KSYSFR	493	
		K417N	QTG <mark>N</mark> IA	414	SY <mark>GFQ</mark> P	494	Y505H	SY <mark>SFR</mark> P	494	
TG <mark>K</mark> IAD	415		TG <mark>N</mark> IAD	415	Y <mark>GFQ</mark> PT	495		Y <mark>SFR</mark> PT	495	
G <mark>K</mark> IADY	416		GN IADY	416	GFQPTN	496				
					FQPTNG	497		FRPTY	497	

Table 2. Cont.

Omicron variant											
C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position	C/H-CrUP	Position	Mutation	New C/H-CrUPs	Position		
WNSNN	436		WNSN <mark>K</mark> L	436	Q PT N GV	498		RPTYGV	498		
SNNLDS	438		SN <mark>K</mark> LDS	438	PT <mark>N</mark> GVG	499					
NNLDSK	439		N <mark>K</mark> LDSKV	439	T <mark>N</mark> GVG <mark>Y</mark>	500		T <mark>Y</mark> GVG <mark>H</mark>	500		
		NI440IZ	KLDSKV <mark>S</mark>	440	NGVGYQ	501					
LDSKV <mark>G</mark>	441	G446S			GVG <mark>Y</mark> QP	502		GVG <mark>H</mark> Q	502		
DSKV <mark>G</mark> G	442		DSKV <mark>S</mark> G	442	VG <mark>Y</mark> QPY	503		VG <mark>H</mark> QPY	503		
KV <mark>G</mark> GNY	444		KV <mark>S</mark> GNY	444	GYQPYR	504					
V <mark>G</mark> GNYN	445		V <mark>S</mark> GNYN	445	YQPYRV	505		HQPYR	505		
G GNYNY	446										

Table 2. Cont.

The original and newly constructed C/H-CrUPs around the native and mutant sites of RBD region of SARS-CoV-2 Spike protein in Alpha, Delta and Omicron variants are presented. With the red colors, the mutant amino acids in wild-type C/H-CrUPs and in the newly constructed peptides are marked.

Remarkably, RBM area contains 11 out of the 12 contact points of viral Spike protein with the ACE2 cellular receptor. Among them, 7 contact points remained intact, while 4 mutations in positions Q493K, Q498R, N501Y and Y505H were identified, resulting in the construction of 17 new C/H-CrUPs (Table 3). N501Y mutation was found to be a major determinant of increased viral transmission, due to the improved binding affinity of Spike protein to ACE2 cellular receptor [10]. These findings indicate that virus binding to ACE2 receptor is notably affected by C/H-CrUP-specific mutations that can likely strengthen Spike-ACE2 protein–protein interaction(s).

Table 3. C/H-CrUPs around SARS-CoV-2 RBD contact positions.

	WILD-TYPE							OMICRON VARIANT				
Contact Positions	C/H-CrUPs						Mutations	6	Newly C	onstructed C/	H-CrUPs	
N439	AWNSN	WNS <mark>N</mark> N	S <mark>N</mark> NLDS	NNLDSK								
Y449	KVGGN <mark>Y</mark>	VGGN <mark>Y</mark> N	GGN <mark>Y</mark> NY	GN <mark>Y</mark> NYL	NYNYLY	-						
Y453	NYNYL <mark>Y</mark>	NYL <mark>Y</mark> RL	YL <mark>Y</mark> RLF	LYRLFR	Y RLFRK	-						
F486	NGVEG <mark>F</mark>	GVEG <mark>F</mark> N	GF NCY	FNCYF		-						
N487	GVEGF <mark>N</mark>	GF <mark>N</mark> CY	FNCYF									
Y489	GFNC <mark>Y</mark>	FNC <mark>Y</mark> F	C <mark>Y</mark> FPLQ	Y FPLQS		-						
Q493	CYFPLQ	YFPL <mark>Q</mark> S	FPL <mark>Q</mark> SY	PL <mark>Q</mark> SYG	Q SYGF		Q493K	CYFPL <mark>K</mark>	YFPL <mark>K</mark> S	FPL <mark>K</mark> SY	L <mark>K</mark> SY <mark>S</mark> F	KSYSFR
Q498	SYGF <mark>Q</mark> P	YGF <mark>Q</mark> PT	GF <mark>Q</mark> PTN	F <mark>Q</mark> PTNG	Q PTNGV		Q498R	K SYSFR	SY <mark>SFR</mark> P	Y <mark>S</mark> FRPT	F <mark>R</mark> PTY	RPTYGV
T500	YGFQP <mark>T</mark>	GFQP <mark>T</mark> N	FQP <mark>T</mark> NG	QP <mark>T</mark> NGV	P T NGVG	TNGVGY						
N501	GFQPT <mark>N</mark>	FQPT <mark>N</mark> G	QPT <mark>N</mark> GV	PT <mark>N</mark> GVG	T <mark>N</mark> GVGY	NGVGYQ	N501Y	F <mark>R</mark> PTY	RPTYGV	T <mark>Y</mark> GVG <mark>H</mark>		
Y505	TNGVG <mark>Y</mark>	NGVG <mark>Y</mark> Q	GVG <mark>Y</mark> QP	VG <mark>Y</mark> QPY	G <mark>y</mark> qpyr	Y QPYRV	Y505H	T <mark>Y</mark> GVG <mark>H</mark>	GVG <mark>H</mark> Q	VG <mark>H</mark> QPY	HQPYR	

The original and newly constructed C/H-CrUPs residing around the native and contact positions of the SARS-CoV-2 Spike protein RBD region. The C/H-CrUPs of wild-type and Omicron variant are presented. With red colors, the mutant amino acids in wild-type C/H-CrUPs and in the newly constructed peptides are marked.

Interestingly, an important amino acid sequence in the RBM area is the "*NYNYLYRLF*" peptide (from 448 to 456 position). This Tyrosine (Y)-enriched peptide carries two contact sites (Y449 and Y453), and it is known as the NF9 peptide [11]. It seems to affect antigen recognition, by being an immunodominant HLA*24:02-restricted epitope identified by CD8⁺ T cells. Of note, NF9 presents immune stimulation activity, and increases cytokine production derived from CD8⁺ T cells, such as IFN- γ , TNF- α and IL-2 [12]. In contrast to Delta, in the Omicron variant the NF9 amino acid content is not changed by any mutation

detected, thus suggesting that the NF9 peptide could induce early immune system activation and efficient cytokine production, leading to a faster immune response, and thus reducing SARS-CoV-2 virus pathogenicity.

3.3. C/H-CrUPs Altered Architecture around the Spike-Cleavage Site(s) of the Omicron Variant

The molecular mechanism of Spike protein's proteolytic activation has been shown to play a crucial role in the selection of host species, virus–cell fusion, and the viral infection of human lung cells [13–15]. Spike protein [SPIKE_SARS2 (P0DTC2)] contains three cleavage sites (known as S-cleavage sites) crucial for the virus fusion to the host cell: the $R^{685}\downarrow S$ and $R^{815}\downarrow S$ positions that serve as direct targets of the Furin protease, and the $T^{696}\downarrow M$ position that can be recognized by the TMPRSS2 protease [16–18].

In these cleavage sites, the Omicron variant carries only the critical mutation P681H, which also appears in the Alpha variant (Figure 1B). Strikingly, in contrast to the Delta variant, which contains the P681R mutation, the P681H mutation constructs several new C/H-CrUPs in the Alpha and Omicron variants, thus indicating their dispensable contribution to virus fusion to the host cell (Table 4).

Cleavage Site	Mutation	Variant	Position	New C/H-CrUPs
	P681R	Delta	680	S <mark>R</mark> RRAR↓S
D685 L C			677	QTNS <mark>H</mark>
R ⁰⁰³ ↓S	P681H	Alpha Omicron	678	TNS <mark>H</mark> R
		-	680	SHRRAR
T ⁶⁹⁶ ↓M	A701V	Beta	None	
$R^{815} \downarrow S$	None	None		None

Table 4. C/H-CrUPs arround the Spike protein cleavage sites.

C/H-CrUPs around the mutant positions of Spike protein cleavage sites are presented. Symbol " \downarrow " indicates the protein cleavage positions.

4. Conclusions

Core Unique Peptides constitute a distinct and important group of peptides within a proteome. The identification of CrUPs in an organism (e.g., virus, microbe, or mutant protein) against a distinct proteome of another organism is a completely novel approach, which could prove useful for the understanding of the action of microorganisms, the association of novel pharmacological targets with therapies, and the design of novel vaccines. It could be employed in many different kinds of diseases, such as cancer, athropozoans diseases, the design of vaccines for pathogenic viruses, and the identification of new antigenic epitopes capable for the development of new diagnostic or therapeutic antibodies. Therefore, we applied this dynamic and novel strategy, for the first time, in the identification of CrUPs derived from SARS-CoV-2 against the human proteome [1]. In that study, we analyzed all the CrUP peptides of all SARS-CoV-2 variants against the proteome of the host organism, which in our case was *Human sapiens*. Remarkably, this approach clearly revealed the immune escaping capacity, the contagious power and the high pathogenicity of Delta variant, in contrast to other variants. Notably, these findings have been confirmed by epidemiological data concerning the course of the disease.

In the present study, we engaged this approach to the analysis of the SARS-CoV-2 Omicron variant. The analysis of C/H-CrUP landscapes in the heavily mutated SARS-CoV-2 Omicron variant Spike protein unveiled that the Omicron variant, by the generation of novel multi-mutated C/H-CrUPs, could escape the immune system defense mechanisms, while these C/H-CrUP-specific mutations could facilitate more efficient virus binding to the ACE2 cellular receptor, and a more productive fusion of the virus to the host cell. Most importantly, in contrast to the Delta variant, the intact NF9 peptide in the Omicron variant, which has a known immunostimulatory effect, suggests that Omicron exhibits reduced pathogenicity as compared to Delta.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/vaccines10030357/s1. Figure S1: Uniquome creation algorithm; Figure S2: Extracting CrUPs of Target vs Reference Proteome; Figure S3: Alignment of the SARS-CoV-2 Spike protein (SPIKE_SARS2, P0DTC2) of the 26 variants, together with the wild-type Spike Protein (SPIKE_SARS2, P0DTC2); Figure S4: Length distribution of Omicron variant Spike protein C/H-CrUPs; Table S1: New C/H-CrUPs located in the RBD and RBM regions of the Spike protein across virus variants.

Author Contributions: Conceptualization, V.P., E.K. and G.T.T.; methodology, V.P. and E.K.; investigation, V.P., E.K. and G.T.T.; visualization, E.K., D.J.S. and G.T.T.; supervision, G.T.T.; writing—original draft, D.J.S. and G.T.T.; writing—review and editing, D.J.S. and G.T.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data of the present article are available in the main text or in the supplementary materials.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

UPs: Unique Peptides; CrUPs: Core Unique Peptides; C/H-CrUPs: SARS-CoV-2 Core Unique Peptides against Human Proteome; RBD: Receptor Binding Domain; RBM: Receptor Binding Motif.

References

- 1. Pierros, V.; Kontopodis, E.; Stravopodis, D.J.; Tsangaris, G.T. Unique Peptide Signatures of SARS-CoV-2 Against Human Proteome Reveal Variants' Immune Escape and Infectiveness. *bioRxiv* 2021. [CrossRef]
- Wang, L.; Cheng, G. Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. J. Med. Virol. 2022, 94, 11728–11733. [CrossRef] [PubMed]
- Kontopodis, E.; Pierros, V.; Anagnostopoulos, A.; Stravopodis, D.J.; Papassideri, I.; Vorgias, C.; Tsangaris, G.T. Data Processing Approach for the Construction and Evaluation of an Organism's UNIQUOME with Comparative Analysis for the Human, Rat and Mouse Uniquomes. In Proceedings of the XIII. Annual Congress of the European Proteomics Association: From Genes via Proteins and their Interactions to Functions, Potsdam, Germany, 24–28 March 2019.
- Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* 2020, 581, 221–224. [CrossRef] [PubMed]
- Hatmal, M.M.; Alshaer, W.; Al-Hatamleh, M.A.I.; Hatmal, M.; Smadi, O.; Taha, M.O.; Oweida, A.J.; Boer, J.C.; Mohamud, R.; Plebanski, M. Comprehensive Structural and Molecular Comparison of Spike Proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and Their Interactions with ACE2. *Cells* 2020, *9*, 2638. [CrossRef] [PubMed]
- Chen, Y.; Zhang, Y.N.; Yan, R.; Wang, G.; Zhang, Y.; Zhang, Z.-R.; Li, Y.; Ou, J.; Chu, W.; Liang, Z.; et al. ACE2-targeting monoclonal antibody as potent and broad-spectrum coronavirus blocker. *Signal Transduct. Target. Ther.* 2021, *6*, 315. [CrossRef] [PubMed]
- Zahradnik, J.; Marciano, S.; Shemesh, M.; Zoler, E.; Harari, D.; Chiaravalli, J.; Meyer, B.; Rudich, Y.; Li, C.; Marton, I.; et al. SARS-CoV-2 variant prediction and antiviral drug design are enabled by RBD in vitro evolution. *Nat. Microbiol.* 2021, 6, 1188–1198. [CrossRef] [PubMed]
- Hastie, K.M.; Li, H.; Bedinger, D.; Schendel, S.L.; Dennison, S.M.; Li, K.; Rayaprolu, V.; Yu, X.; Mann, C.; Zandonatti, M.; et al. Defining variant-resistant epitopes targeted by SARS-CoV-2 antibodies: A global consortium study. *Science* 2021, 374, 472–478. [CrossRef] [PubMed]
- Chen, J.; Wang, R.; Gilby, N.B.; Wei, G.-W. Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance. J. Chem. Inf. Model. 2022, 62, 412–422. [CrossRef] [PubMed]
- Liu, Y.; Liu, J.; Plante, K.S.; Plante, J.A.; Xie, X.; Zhang, X.; Ku, Z.; An, Z.; Scharton, D.; Schindewolf, C.; et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature* 2022, 602, 294–299. [CrossRef] [PubMed]

- Motozono, C.; Toyoda, M.; Zahradnik, J.; Saito, A.; Nasser, H.; Tan, S.T.; Ngare, I.; Kimura, I.; Uriu, K.; Kosugi, Y.; et al. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host Microbe* 2021, 29, 1124–1136.e11. [CrossRef] [PubMed]
- Kared, H.; Redd, A.D.; Bloch, E.M.; Bonny, T.S.; Sumatoh, H.; Kairi, F.; Carbajo, D.; Abel, B.; Newell, E.W.; Bettinotti, M.P.; et al. SARS-CoV-2-specific CD8+ T cell responses in convalescent COVID-19 individuals. *J. Clin. Investig.* 2021, 131, 1124–1136.e11. [CrossRef] [PubMed]
- Peacock, T.P.; Goldhill, D.H.; Zhou, J.; Baillon, L.; Frise, R.; Swann, O.C.; Kugathasan, R.; Penn, R.; Brown, J.C.; Sanchez-David, R.Y.; et al. The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat. Microbiol.* 2021, 6, 899–909. [CrossRef] [PubMed]
- 14. Whittaker, G.R. SARS-CoV-2 Spike and its Adaptable Furin Cleavage Site. Lancet Microbe 2021, 2, e488–e489. [CrossRef]
- Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. Proc. Natl. Acad. Sci. USA 2020, 117, 11727–11734. [CrossRef] [PubMed]
- Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, Y.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020, *181*, 271–280.e8. [CrossRef] [PubMed]
- Hoffmann, M.; Kleine-Weber, H.; Pohlmann, S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol. Cell* 2020, 78, 779–784.e5. [CrossRef] [PubMed]
- 18. Takeda, M. Proteolytic activation of SARS-CoV-2 spike protein. Microbiol. Immunol. 2022, 66, 15–23. [CrossRef] [PubMed]