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

Vasileios Pierros^{1,†}, Evangelos Kontopodis^{1,2,†}, Dimitrios J. Stravopodis² and George Th. Tsangaris^{1*}

Affiliations:

¹Proteomics Research Unit, Biomedical Research Foundation of the Academy of Athens; 11527 Athens, Greece

²Section of Cell Biology and Biophysics, Department of Biology, School of Science, National and Kapodistrian University of Athens; 15701 Athens, Greece

[†]These authors contributed equally to this work

*Correspondence: gttsangaris@bioacademy.gr (GThT)

Summary

The recently discovered Omicron variant of the SARS-CoV-2 corona virus has raised a new, global, awareness, since it is considered as a new variant of concern from all major health organizations, including WHO and ECDC. Omicron variant is characterized by 30 amino acid changes, three small deletions and one small insertion in the Spike protein. In this study, we have identified the Core Unique Peptides (CrUPs) that reside exclusively in the Omicron variant of Spike protein and are absent from the human proteome, thus creating a new dataset of peptides named as C/H-CrUPs. Furthermore, we have analyzed their protein locations and compared them with the respective ones of Alpha and Delta SARS-CoV-2 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. From position 440 to position 508, at the Receptor Binding Motif (RBM), 8 mutations were detected, resulting in the construction of 28 novel C/H-CrUPs. Most importantly, in 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. Remarkably, these Omicron-derived C/H-CrUPs that bear several mutated amino acids could not be recognized in any other viral Spike variant. We suggest that 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. Taken together, our findings indicate that Omicron variant contains intrinsic abilities to escape immune-system attack, while its mutations can mediate strong viral binding to the ACE2 receptor, leading to highly efficient fusion of the virus to the target cell. However, the intact NF9 peptide suggests that Omicron exhibits reduced pathogenicity compared to Delta variant.

Keywords

Core Unique Peptide, COVID-19, Immune escape, Infectiveness, Mutation, Omicron variant, Pathogenicity, SARS-CoV-2, Spike protein, Uniquome

Introduction

SARS-CoV-2 virus has been presented with an ability of high mutagenesis frequency, hitherto producing 63 different variants with 39 of them being considered as the most predominant forms, including Delta, the dominant variant of the 4th pandemic wave (1). Recently, a new variant, the 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).

To elucidate the SARS-CoV-2 virus-host organism interactions, we have designed a novel bioinformatics approach 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 viral proteome that do not exist in human proteome. Based on their properties, the viral C/H-CrUPs could advance our knowledge regarding virus-host interactions, immune-system response(s), infectiveness and pathogenicity of the virus, while, 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 the Omicron-specific C/H-CrUPs, in order to illuminate the mechanisms controlling infectivity, immune escape and pathogenicity of the new variant.

Materials and Methods

Methods

For our recent studies (1, 3), we have developed a bioinformatics tool that can extract from a given proteome the Core Unique Peptides (CrUPs) (thus creating its Uniquome) (Supp. Fig. 1). 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 (Supp. Fig. 2).

For the present study, we have engaged this new feature of our tool. We have 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 have the desired data, we run a meta-analysis to identify how many C/H-CrUPs remain the same, or are added or lost on each variant versus the wild-type Spike protein. For this analysis, initially we take the identified C/H-CrUPs of the wild-type sequence and check their presence against the respective C/H-CrUPs of the other variants. We only care for the amino acid sequence and not the position this could be found within the protein. If the sequence is found, then we consider the peptide to be the same, otherwise we consider it to be lost on the examined variant.

Next, we analyze the identified C/H-CrUPs of each variant versus the wild-type sequence. If the peptide is detected only on variant's C/H-CrUPs, then we consider it as added. This meta-analysis also provides us with the position of each C/H-CrUP within the Spike protein, which we have used to determine the area (i.e. RBD, or S-cleavage site) they reside in.

Databases

All proteomes and proteins were obtained from [Uniprot](https://www.uniprot.org/). SARS-CoV-2 wild-type and variant sequences, and mutations were obtained from Stanford COVID Database (<https://covdb.stanford.edu/page/mutation-viewer/>).

Results and Discussion

Mapping the C/H-CrUPs landscape of Spike protein of SARS-CoV-2 Omicron variant

SARS-CoV-2 virus seems to be highly mutated, producing so far 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 (Supp. Fig. 3) (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) (Supp. Fig. 3).

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 protein (987 C/H-CrUPs) and to the mean $\pm$ SD value of Spike protein-specific C/H-CrUPs (983 ± 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.

These mutations seem to have a dramatic effect on 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 other variants (Table 1). The distribution of these new C/H-CrUPs shows that their majority carry 6 amino acids in length (Supp. Fig. 4).

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 (Walls *et al.*, 2020; Wang *et al.*, 2020). The viral Spike protein attaches to ACE2 receptor by a Receptor Binding Domain (RBD) defined from amino acid position F318 to F541 (4). 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, Omicron variant carries 15 mutations, 10 of which are identified in the RBM area (Fig. 1A). This results in 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 (Supp. Table 1). Table 2, describes all the new, herein identified, C/H-CrUPs of Omicron variant in Spike's RBD region, in comparison to Alpha and Delta variants, which represent two of the most

predominant variants of the virus in human populations. Hence, it proved 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 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 till today. Taking into consideration recent data about virus infectivity, the multimutated, new, C/H-CrUP collection seems to change radically the structure and the epitope regions of end positions of the RBM area, in Omicron variant, causing serious compromise of its antigenic capacity and facilitating the immune escape of the virus (9).

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 improved binding affinity of Spike protein to ACE2 cellular receptor (12). These findings indicate that virus binding to ACE2 receptor is notably affected by C/H-CrUP-specific mutations that can likely strengthen the Spike-ACE2 protein-protein interaction(s). Interestingly, an important amino acid sequence in 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 (10). 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 (11). In contrast to Delta, in Omicron variant the NF9 amino acid content is not changed by any mutation detected, thus suggesting that the NF9 peptide could induce an early immune system activation and efficient cytokine production, leading to a faster immune response, and thus reducing SARS-CoV-2 virus pathogenicity.

C/H-CrUPs altered architecture around the Spike-cleavage site(s) of 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 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 R⁶⁸⁵↓S and R⁸¹⁵↓S positions that serve as direct targets of the Furin protease, and the T⁶⁹⁶↓M position that can be recognized by the TMPRSS2 protease (16-18).

In these cleavage sites Omicron variant carries only the critical mutation P681H, which also appears in the Alpha variant (Fig. 1B). Strikingly, in contrast to 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).

Conclusions

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

Acknowledgments

Funding: No funds were received for this work.

Author contributions: Conceptualization: VP, EK, GThT; Methodology: VP, EK; Investigation: VP, EK, GThT; Visualization: EK, DJS, GThT; Supervision: GThT; Writing - original draft: DJS, GThT; Writing - review & editing: DJS, GThT.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All data are available in the main text or in the supplementary materials.

References

1. Pierros, V., Kontopodis, E., Stravopodis, D. and Tsangaris, G.Th. (2021). Unique Peptide Signatures of SARS-CoV-2 Against Human Proteome Reveal Variants' Immune Escape and Infectiveness. *BioRxiv*. <https://doi.org/10.1101/2021.10.03.462911>.
2. Wang L, Cheng G. (2021). Sequence Analysis of the Emerging SARS-CoV-2 Variant Omicron in South Africa. *J. Med. Virol.* 2021 doi: 10.1002/jmv.27516.
3. Kontopodis, E., Pierros, V., Anagnostopoulos, A., Stravopodis, D., Papassideri, I., Vorgias C. and Tsangaris, G.Th. (2019). Data Processing Approach for the Construction and Evaluation of an Organism's UNIQUOME with Comparative Analysis for the Human, Rat and Mouse Uniquomes. Paper presented at XIII. Annual Congress of the European Proteomics Association: From Genes via Proteins and their Interactions to Functions, Potsdam, Germany, P194 March 24-28.
4. Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A. and Li, F. (2020). Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature*, 581, 221-224.
5. 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. and Plebanski, M. (2020) Comprehensive Structural and Molecular Comparison of Spike Proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and Their Interactions with ACE2. *Cells* 9, 2638. <https://doi.org/10.3390/cells9122638>.
6. Chen, Y., Zhang, Y. N., Yan, R., Wang, G., Zhang, Y., Zhang, Z.R., Li, Y., Ou, J., Chu, W., Liang, Z., Wang, Y., Chen, Y.L., Chen, G., Wang, Q., Zhou, Q., Zhang, B., and Wang, C. (2021). ACE2-targeting Monoclonal Antibody as Potent and Broad-spectrum Coronavirus Blocker. *Signal Transduction and Targeted Therapy*, 6(1), 315. <https://doi.org/10.1038/s41392-021-00740-y>.

7. Zahradník, J., Marciano, S., Shemesh, M., Zoler, E., Harari, D., Chiaravalli, J., Meyer, B., Rudich, Y., Li, C., Marton, I., Dym, O. *et al.* (2021). SARS-CoV-2 Variant Prediction and Antiviral Drug Design are Enabled by RBD *in vitro* Evolution. *Nat. Microbiol.* 6, 1188-1198.
8. 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.* (2021). Defining Variant-resistant Epitopes Targeted by SARS-CoV-2 Antibodies: A Global Consortium Study. *Science* 374, 472-478. <https://doi.org/10.1126/science.abh2315>.
9. Chen J, Wang R, Gilby N.B., Wei G.W. Omicron (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance. *ArXiv [Preprint]*. 2021 Dec 1:arXiv:2112.01318v1. PMID: 34873578; PMCID: PMC8647651.
10. Motozono, C., Toyoda, M., Zahradnik, J., Saito, A., Nasser, H., Tan, T.S., Ngare, I., Kimura, I., Uriu, K., Kosugi, Y., Yue, Y., Shimizu, R., Ito, J., Torii, S., Yonekawa, A., Shimono, N., Nagasaki, Y., Minami, R., Toya, T., Sekiya, N., ... Sato, K. (2021). SARS-CoV-2 Spike L452R Variant Evades Cellular Immunity and Increases Infectivity. *Cell Host & Microbe*, 29(7), 1124-1136.e11. <https://doi.org/10.1016/j.chom.2021.06.006>.
11. 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.* (2021). SARS-CoV-2-specific CD8+ T Cell Responses in Convalescent COVID-19 Individuals. *J. Clin. Invest.* 131, e145476. doi: 10.1172/JCI145476.
12. Liu, Y., Liu, J., Plante, K.S., Plante, J.A., Xie, X., Zhang, X., Ku, Z., An, Z., Scharon, D., Schindewolf, C., Widen, S.G., Menachery, V.D., Shi, P.Y. and Weaver, S.C. (2021). The N501Y Spike Substitution Enhances SARS-CoV-2 Infection and Transmission. *Nature*. 2021, Nov. 24. doi: 10.1038/s41586-021-04245-0.
13. 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.* (2021). The Furin Cleavage Site in the SARS-CoV-2 Spike Protein is Required for Transmission in Ferrets. *Nat. Microbiol.* 6, 899-909.
14. Whittaker, G.R. (2021). SARS-CoV-2 Spike and its Adaptable Furin Cleavage Site. *Lancet Microbe*. [https://doi.org/10.1016/S2666-5247\(21\)00174-9](https://doi.org/10.1016/S2666-5247(21)00174-9).
15. Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A. and Li, F. (2020). Cell Entry Mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* 117, 11727-11734.
16. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A. *et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271-280.
17. Hoffmann, M., Kleine-Weber, H., Pöhlmann, S. (2020). A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 is Essential for Infection of Human Lung Cells. *Mol. Cell*. 78, 779-784.
18. Takeda, M. (2021). Proteolytic Activation of SARS-CoV-2 Spike Protein. *Microbiology and Immunology*, 10.1111/1348-0421.12945. <https://doi.org/10.1111/1348-0421.12945>.

A

		RECEPTOR BINDING MOTIF										NF9 PEPTIDE										RECEPTOR BINDING MOTIF																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Variant / AA Position	430	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
SARS-CoV-2 spike	W	N	S	N	N	N	D	S	A	V	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G																																				

B

Variant / AA Position	S1 DOMAIN										BRIDGE										S2 DOMAIN										S3 DOMAIN																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
Alpha (H.3.2.2.2009)	C	A	S	Y	G	T	Q	T	N	P	R	A	S	V	A	S	Q	S	I	T	A	Y	T	M	S	L	G	A	E	N	S	V	A	Y	S	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

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

Variant	Spike protein						
	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
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

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 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 of each variant

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

Alpha variant					Delta variant					Omicron variant				
C/H-CrUP	Posistion	Mutation	New C/H-CrUPs	Posistion	C/H-CrUP	Posistion	Mutation	New C/H-CrUPs	Posistion	C/H-CrUP	Posistion	Mutation	New C/H-CrUPs	Posistion
GNYN YL	447	L452R	GNYN R	447	PGQTGK I	412	K417N			NLCP FG	334	G339D	NLCP FD	334
NYN LY	448				GQTGK IA	413		GQTG NI	413	LCPF GE	335		LCPF DE	335
			YNY RY	449				QTGN IA	414	PFG EVF	337		P FD EV	337
NYLYRL	450		NYR YRL	446	TGKIAD	415		TGNIAD	415	FG EVFN	338		F DE VFN	338
YLYRLF	451		YRYRLF	447	GKIADY	416		GNIADY	416	GEV FNA	339		D EV FNA	339
LYRLF R	452													
CNGV EG	480	E484K	CNGV KG	480	GNYN YL	447	L452R	GNYN R	447	VLYN SA	367	S371L S373P S375F	VLYN LAP	367
NGV EGF	481		NGV KGF	481	NYN LY	448				LYNS AS	368			
GVE GFN	482		G VK GFN	482				YNY RY	449	YNS ASF	369		YN LAPF	369
			KGFNC	484	NYLYRL	450		NYR YRL	450	NS ASFST	370			
					YLYRLF	451		YRYRLF	451				LAP FFT	371
					LYRLF R	452				AS FSTF	372		AP FFT F	372
YFPL QS	489	S494P N501Y	YFPL QP	489	YQAG ST	473	T478K	YQAG SK	473	ST FTK	373			
FPL QSY	490		FPL QPY	490				QAGS KP	474	STFKC	375			
PLQ SYG	491		PLQ PYG	491	AG STPC	475		AGS KPC	475					
Q SYGF	493		Q PYGF	493				GSK PCN	476					
SY GFQP	494		PY GFQP	494	ST PCN	477								
GFQ PTN	496							KPCN G	478					
FQ PTN G	497									PGQTGK I	412	K417N		
QPTN GV	498									GQTGK IA	413		GQTG NI	413
			QPT Y	498									QTGN IA	414
PTN GVG	499		PTY G	499						TGKIAD	415		TGNIAD	415
TN GVGY	500		TY GV	500						GKIADY	416		GNIADY	416
NGVGY Q	501		YGV G	501										
										WNS NN	436	N440K G446S	WNS NKL	436
										SN NLDS	438		SN KLDS	438
										NN LDSK	439		NK LDSKV	439
													K LDSKV S	440
										LDSK VG	441			
										DSK VGG	442		DSK VSG	441
										KV G GNY	444		KV S GNY	443
										V G GNYN	445		V S GNYN	444
										G G NYNY	446			
										IYQAG S	472	S477N T478K		
										YQAG ST	473		YQAG N	473
													QAGN KP	474
										AG STPC	475			
										ST PCN	477		NK PCN	476
													KPCN G	478
										CNGV EG	480	E484A	CNGV AG	480
										NGV EGF	481		NGV AGF	481
										GVE GFN	482		GV AGFN	482
													VAG FNC	483
										CYFPL Q	488	Q493K G496S Q498R N501Y Y505H	CYFPL K	488
										YFPL QS	489		YFPL KS	489
										FPL QSY	490		FPL KSY	490
										PLQ SYG	491			
													LK SYSF	492
										Q SYGF	493		K SYSFR	493
										SY GFQP	494		SY SFRP	494
										YGFQ PT	495		Y SFRPT	495
										GFQ PTN	496			
										FQ PTN G	497		FRPT Y	497
										QPTN GV	498		RPT YGV	498
										PTN GVG	499			
										TN GVGY	500		TYGV GH	500
										NGVGY Q	501			
										GVGY QP	502		GVGH Q	502
										VGYQ PY	503		VGH QPY	503
										GYQ PYR	504			
										YQPYR V	505		HQPYR	505

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

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

	WILD-TYPE							OMICRON VARIANT					
Contact positions	C/H-CrUPs						Mutations	New constructed C/H-CrUPs					
N439	AWNSN	WNSNN	SNNLDS	NNLDSK									
Y449	KVGGNY	VGGNYN	GGNYNY	GNYNYL	NYNYLY								
Y453	NYNYLY	NYLYRL	YLYRLF	LYRLFR	YRLFRK								
F486	NGVEGF	GVEGFN	GFNCY	FNCYF									
N487	GVEGFN	GFNCY	FNCYF										
Y489	GFNCY	FNCYF	CYFPLQ	YFPLQS									
Q493	CYFPLQ	YFPLQS	FPLQSY	PLQSYG	QSYGF		Q493K	CYFPLK	YFPLKS	FPLKSY	LKSYSF	KSYSFR	
Q498	SYGFQP	YGFQPT	GFQPTN	FQPTNG	QPTNGV		Q498R	KSYSFR	SYSFRP	YSFRPT	FRPTY	RPTYGV	
T500	YGFQPT	GFQPTN	FQPTNG	QPTNGV	PTNGVG	TNGVGY							
N501	GFQPTN	FQPTNG	QPTNGV	PTNGVG	TNGVGY	NGVGYQ	N501Y	FRPTY	RPTYGV	TYGVGH			
Y505	TNGVGY	NGVGYO	GVGYOP	VGYQPY	GYQPYR	YQPYRV	Y505H	TYGVGH	GVGHQ	VGHOPY	HOPLYR		

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

TABLE 4. C/H-CrUPs around the Spike protein cleavage sites.

Cleavage Site	Mutation	Variant	Position	New C/H-CrUPs
R ⁶⁸⁵ ↓S	P681R	Delta	680	SRRRAR↓S
	P681H	Alpha Omicron	677	QTNSH
			678	TNSHR
			680	SHRRAR
T ⁶⁹⁶ ↓M	A701V	Beta	None	
R ⁸¹⁵ ↓S	None	None	None	

The C/H-CrUPs around the mutant positions of Spike protein cleavages sites were presented. ↓ indicates the protein cleavage position.