In Silico immunogenic prediction of SARS-CoV-2 spike protein in Iraqi variants

Majeed A. Sabbah[1*], Hala K. Ibrahim[1], Manal A. Almonem[2], Wisam J. Mohammed[2], Ahmed K. Qasim[2], Ziadoon Mohammed[2]

[1] Higher institute of forensic sciences/Al-Nahrain University, Baghdad/Iraq
[2] Ministry of Health, Baghdad/Iraq

*Correspondence email: majeedbio@gmail.com

Received: 21/5/2024
Accepted: 27/8/2024
Online: 12/10/2025

ABSTRACT

Background: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of the COVID-19 pandemic worldwide since 2019. Many variants were recorded due to a change in the virus genome. Variants associated with the severity of disease. So, investigating protein structure and immune response is essential. The aim of this work is to decipher the *in silico* prediction of the immunological function of the SARS-CoV-2 spike protein in Iraqi variants. **Methodology**: Sixteen SARS-CoV-2 genomes previously completely sequenced in the Central Public Health Laboratory, Baghdad, Iraq, were used. The physicochemical and Immunogenic Properties of the spike protein are compared among Iraqi variants and with the Wuhan variant *in silico*. **Results**: Four COVID-19 variants show two main mutations against the Wuhan variant. The Iraqi variants were found to be amphipathic and highly stable. CTLpred and ABCpred analyses were led for the detection of B and T cell epitopes. **Conclusion**: We conclude that Iraqi SARS-CoV-2 genomes exhibit numerous variants, which may impact the severity of the disease, these variants could serve as targets for Iraqi-specific vaccines or therapies.

Keywords: COVID-19, Spike, WGS, SARS-CoV-2, *in silico.*
https://doi.org/10.24126/jobrc.2025.19.3.859

INTRODUCTION

The epidemiology of the coronavirus, SARS-CoV-2, has shed light on the importance of the spike protein for entering the cell and causing infection (1). The SARS-CoV-2 spike (S) protein is a glycoprotein located at the viral surface, consisting of two subunits (S1 and S2) (2,3).

The entrance of the virus into the cell is initiated by the interaction of the S1 subunit, specifically through its receptor-binding domain (RBD), with the host cell receptor, ACE2 (angiotensin-converting enzyme 2). This interaction is a crucial step in the viral life cycle, facilitating cellular entry and subsequent infection. The S2 subunit is responsible for membrane fusion, which leads to the entry of the virus's genetic material into the host cell (4).

The spike protein is the most important part of vaccine development. Current mRNA-based vaccines, such as those from Moderna and Pfizer-BioNTech, utilize some of the genetic material from the spike protein to stimulate an immune response. Additionally, vector-based vaccines, such as those developed by AstraZeneca and Johnson & Johnson, utilize harmless viruses to deliver the spike protein gene into cells, thereby stimulating an immune response. The effectiveness of these vaccines underscores the significance of the S protein and its capacity to confer protection against severe disease and mortality (5). The development of antiviral therapeutics requires a comprehensive understanding of the structure and function of the spike protein. To inhibit viral entry into the target cell or impede membrane fusion, specific regions of the spike protein should be targeted to prevent its interaction with ACE2 (6,7). The aim of this work is to in silico predict the immunological function of the SARS-CoV-2 spike protein in Iraqi variants.

METHODOLOGY

Samples

At the onset of the pandemic in Iraq, the Baghdad Central Public Health Laboratory initiated real-time PCR testing for COVID-19 in patients from Iraq. Nasopharyngeal swabs were collected in 2020 using a VTM (viral transport medium) according to WHO guidelines. All participants provided written consent to participate in this study.

Viral RNA Extraction and Real-Time PCR:

RNA extraction of the SARS-CoV-2 virus was achieved using the ExiPrepTM Viral DNA/RNA Kit (Bioneer, Korea). Identification and confirmation of the virus was carried out using a real-time PCR-based virus detection system, AccuPower® SARS-CoV-2 Real-Time RT-PCR Kit (Bioneer, Korea).

Whole genome Sequencing (WGS) of the SARS-CoV-2 genome:

The whole-genome sequencing of previously sequenced isolates was performed at the Baghdad Central Public Health Laboratory using the MiSeq (Illumina, USA). Indexed library preparation was carried out using the AmpliSeq from Illumina. Analysis of the sequencing data was performed using the Illumina DRAGEN COVID Pipeline Software. Viral genomes were aligned using ClustalW online software. Physicochemical properties computed using the ExPASy ProtParam tool. Online tools, including ABCpred (8) for predicting B cell epitopes in an antigen sequence and CTLpred (9,10) for predicting CTL epitopes crucial in subunit vaccine design, were used for in silico prediction of Immunogenic Properties. A score of more than 0.8 indicates a high probability of epitope presence.

RESULTS

In this study, the differences between the Wuhan Spike protein and the Iraqi variant Spike proteins were manually investigated (Table 1). In order to compare the S protein of variants, viral genomes were aligned using ClustalW software. The results indicate that there are two mutations that result in changes to amino acids (at positions 22311 and 21707).

Table (1): Differences between the SARS-CoV-2 spike protein in Iraqi variants

No. of variant	location	Mutation	Amino acid change	Amino acids NO.	Chemical Change
11	21707	A to G	Threonine ACA to alanine GCA	29	Nonpolar aliphatic to polar uncharged
13	21707	A to G	Threonine ACA to alanine GCA	29	Nonpolar aliphatic to polar uncharged
13	22311	C to T	Threonine ACT to ATT isoleucine	250	Nonpolar aliphatic to polar uncharged
10	21647	T to C	alanine GCA to alanine GCC		-
	22311	C to T	Threonine ACT to ATT isoleucine	250	Nonpolar aliphatic to polar uncharged
15	21647	T to C	alanine GCA to alanine GCC		-
	22311	T to C	Threonine ACT to ATT isoleucine	250	Nonpolar aliphatic to polar uncharged

The physicochemical properties of the Wuhan spike protein compared to the Iraqi variants' Spike proteins (Table 2). The results show significant differences in molecular weight, number of atoms, aliphatic index, and grand average of hydropathicity, as well as the instability index.

Table (2): Physicochemical properties of the Wuhan spike from SARS-CoV-2 computed using the ExPASy ProtParam tool.

NO.	Properties	Wuhan spike	Iraqi SARS-CoV-2 spike with 21707 variation	Iraqi SARS-CoV-2 spike with 22311 variation
1	Number of amino acids	1273	1273	1273
2	Molecular weight (kDa)	141178.47	141148.44	141190.53
3	Formula	C6336H9770N1656O1894S54	C6335H9768N1656O1893S54	C6338H9774N1656O1893S54
4	Total number of atoms	19710	19706	19715
5	Theoretical pI	6.24	6.24	6.24
6	Extinction coefficients	148960	148960	148960
7	Aliphatic index	84.67	84.75	84.98
8	Grand average of hydropathicity	-0.079	-0.077	-0.075
9	Instability index	33.01 (protein is stable)	33.38 (protein is stable)	32.99 (protein is stable)

The In Silico Prediction of Immunogenic Properties of the Wuhan spike protein compared to Iraqi variants Spike proteins (table 3,4,5). The results show that, despite differences in amino acids, they share similar immunological properties.

Table (3): Prediction of epitopes (B and CTL), IFN-γ response, and probable antigen of SARS-CoV-2 Wuhan spike protein.

B-cell epitope				**CTL epitope**				**Antigen potential**	**Chemokine**
Rank	Position	Sequences/IFN-γ response	Score	Rank	Position	Sequences/IFN-γ response	Score	Prediction score	Prediction
1	879	AGTITSGWTFGAGAAL	0.97	1	318	FRVQPTESI	1	Epitope	Chemokine
2	594	GVSVITPGTNTSNQVA	0.95	2	448	NYNYLYRLF	1	Epitope	Chemokine

Table (4): Prediction of epitopes (B and CTL), IFN-γ response, and probable antigen of SARS-CoV-2 21707

B-cell epitope				**CTL epitope**				**Antigen potential**	**Chemokine**
Rank	Position	Sequences/IFN-γ response	Score	Rank	Position	Sequences/IFN-γ response	Score	Prediction score	Prediction
1	879	AGTITSGWTFGAGAAL	0.97	1	318	FRVQPTESI	1	Epitope	Chemokine
2	594	GVSVITPGTNTSNQVA	0.95	2	448	NYNYLYRLF	1	Epitope	Chemokine

Table (5): Prediction of epitopes (B and CTL), IFN-γ response, and probable antigen of SARS-CoV-2 22311

B-cell epitope				**CTL epitope**				**Antigen potential**	**Chemokine**
Rank	Position	Sequences/IFN-γ response	Score	Rank	Position	Sequences/IFN-γ response	Score	Prediction score	Prediction
1	879	AGTITSGWTFGAGAAL	0.97	1	236	TRFQTLLAL	1	Epitope	Chemokine
2	594	GVSVITPGTNTSNQVA	0.95	2	318	FRVQPTESI	1	Epitope	Chemokine

DISSCUSION

Iraqi SARS-CoV-2 variants caused infections of varying illness severity in addition to many deaths in Iraq, as in other countries. With time, new variants emerged that varied in location across the viral genome, especially in the S protein. With the increasing threat of the virus, the need arose to develop a vaccine that takes into account variations in the S protein. Using Bioinformatics tools reduces the time required to design the vaccine, thereby accelerating its production.

In the current study, the spike protein of sixteen whole virus genomes from Iraqi variants was compared to the Wuhan variant. Two mutations were identified in 16 variants, resulting in amino acid changes from polar to nonpolar amino acids. These mutations in the S protein can affect the interaction between the virus and host cells and may lead to alterations in the Immune response. S protein mutations affect T cell response and immunological memory (11). It has an impact on antibody binding and Neutralization (12).

Bioinformatics tools ABcpred and CTLpred were used to predict B cells, IFN-γ, and cytotoxic T cells stimulation. Two sequences were identified in the Wuhan and Iraqi variants *in silico,* predicted to have immune stimulation properties. One variant has a different stimulation sequence compared to the Wuhan variant. These differences may affect the development of a highly efficient vaccine. The escape of the vaccine is the main concern of vaccine developers, and this issue has led companies to modify their vaccines in response to changes in the S protein (13). Usually, whole inactivated virus, viral vectors, live attenuated virus, or mRNA methodology is used for designing and developing spike-protein-based vaccines (14). Many factors, including safety, efficacy, cost-effectiveness, and scalability, determine the selection of technology.

CONCLUSION

According to the results, we can conclude that Iraqi SARS-CoV-2 variants have specific mutations that affect the structure and may affect the severity of immune stimulation in patients. This may lead to the effectiveness of current vaccines.

REFERANCES

1. Yao H., Song Y., Chen Y., Wu N., Xu J., Sun C., Zhang J., Weng T., Zhang Z., Wu Z., Cheng L., Shi D., Lu X., Lei J., Crispin M., Shi Y., Li L., Li S. Molecular Architecture of the SARS-CoV-2 Virus. Cell. (2020); Oct 29,183(3): 730-738.e13.

2. Belouzard S., Chu V.C., Whittaker G.R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc. Natl. Acad. Sci. USA. (2009); 106: 5871–5876.

3. Li F., Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol. (2016); 3: 237-261

4. Lu R., Zhao X., Li J., Niu P., Yang B., Wu H., Wang W., Song H., Huang B., Zhu N., Bi Y., Ma X., Zhan F., Wang L., Hu T., Zhou H., Hu Z., Zhou W., Zhao L., Chen J., Meng Y., Wang J., Lin Y., Yuan J., Xie Z., Ma J., Liu WJ., Wang D., Xu W., Holmes E.C., Gao GF., Wu G., Chen W., Shi W., Tan W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. (2020); Feb 22,395(10224): 565-574.

5. Jackson C.B., Farzan M., Chen B., Choe H. Mechanisms of SARS-CoV-2 entry into cells. Nat Rev Mol Cell Biol. (2022); Jan,23(1): 3-20.

6. Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y., Zhang L., Fan G., Xu J., Gu X., Cheng Z., Yu T., Xia J., Wei Y., Wu W., Xie X., Yin W., Li H., Liu M., Xiao Y., Gao H., Guo L., Xie J., Wang G., Jiang R., Gao Z., Jin Q., Wang J., Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. (2020); Feb 15,395(10223): 497-506.

7. Wang Z., Schmidt F., Weisblum Y., Muecksch F., Barnes C.O., Finkin S., Schaefer-Babajew D., Cipolla M., Gaebler C., Lieberman J.A., Oliveira T.Y., Yang Z., Abernathy ME., Huey-Tubman K.E., Hurley A., Turroja M., West K.A., Gordon K., Millard K.G., Ramos V., Da Silva J., Xu J., Colbert RA., Patel R., Dizon J., Unson-O'Brien C., Shimeliovich I., Gazumyan A., Caskey M., Bjorkman P.J., Casellas R., Hatziioannou T.,

Bieniasz PD., Nussenzweig MC. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature. (2021); Apr, 592(7855):616-622.

8. Saha S., Raghava G. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins: Structure, Function, and Bioinformatics, (2006); vol. 65, no. 1: pp. 40–48.

9. Lata S. , Raghava G. Prediction and classification of chemokines and their receptors, Protein Engineering, Design & Selection, (2009); vol. 22, no. 7: pp. 441–444.

10. Srivastava V., Kaushik S., Bhargava G., Jain A., Saxena J., Jyoti A. A Bioinformatics Approach for the Prediction of Immunogenic Properties and Structure of the SARS-COV-2 B.1.617.1 Variant Spike Protein, BioMed Research International, vol. (2021); Article ID 7251119, 8 pages, 2021.

11. Mushebenge A.G.-A., Ugbaja S.C., Mbatha N.A., Khan R.B., Kumalo H.M. A Comprehensive Analysis of Structural and Functional Changes Induced by SARS-CoV-2 Spike Protein Mutations. COVID (2023); 3: 1454-1472.

12. Cameroni E., Bowen J.E., Rosen L.E., Saliba C., Zepeda S.K., Culap K., Pinto D., VanBlargan L.A., De Marco A., di Iulio J., Zatta F., Kaiser H., Noack J., Farhat N., Czudnochowski N., Havenar-Daughton C., Sprouse K.R., Dillen J.R., Powell A.E., Chen A., Maher C., Yin L., Sun D., Soriaga L., Bassi J., Silacci-Fregni C., Gustafsson C., Franko N.M., Logue J., Iqbal N.T., Mazzitelli I., Geffner J., Grifantini R., Chu H., Gori A., Riva A., Giannini O., Ceschi A., Ferrari P, Cippà P.E, Franzetti-Pellanda A, Garzoni C, Halfmann P.J, Kawaoka Y., Hebner C., Purcell L.A., Piccoli L., Pizzuto M.S., Walls A.C., Diamond M.S., Telenti A., Virgin H.W., Lanzavecchia A., Snell G., Veesler D., Corti D. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. Nature. (2022); Feb,602(7898): 664-670.

13. Moss P. The T cell immune response against SARS-CoV-2. Nat. Immunol. (2022); 23: 186-193.

14. Chavda V.P., Pandya R., Apostolopoulos V. DNA vaccines for SARS-CoV-2: Toward third-generation vaccination era. Expert Rev. Vaccines. (2021); 20: 1549–1560.

التنبؤ المناعي لبروتين الشائك في المتغيرات العراقية لفايروس SARS-CoV-2

مجيد ارشيد سباح[1] ، هالة خالد[1] ، منال عبد المنعم[2] ، وسام جاسم محمد[2] ، احمد كمال قاسم [2] ، زيدون محمد[2]

المعهد العالي للعلوم العدلية/ جامعة النهرين / بغداد / العراق
وزارة الصحة / بغداد / العراق

الخلاصة:

خلفية عن الموضوع: يعد فايروس كورونا -2 المرتبط بالمتلازمة التنفسية الحادة هو المسبب لجائحة كوفيد-19 في جميع أنحاء العالم منذ عام 2019. وقد تم تسجيل العديد من المتغيرات بسبب تغير جينوم الفايروس. المتغيرات مرتبطة بخطورة المرض. لذا فإن فحص بنية البروتين والاستجابة المناعية أمر ضروري. **الهدف من الدراسة:** دراسة تركيب والتنبؤ بالوظيفة المناعية لبروتين SARS-CoV-2 في المتغيرات العراقية. **المواد و طرق العمل:** استخدامت ستة عشر جينومًا لـ SARS-CoV-2 تم اجراء تسلسلها بالكامل سابقًا في العراق. درست الخصائص الفيزيوكيميائية والمناعية لبروتين الشائك مقارنا بين المتغيرات العراقية ومع متغير ووهان في االحاسوب (انسيليكو). **النتائج:** تُظهر أربعة متغيرات طفرتين رئيسيتين مقارنة بمتغير ووهان. وجد أن بروتين المتغيرات العراقية لها جزء محب وجزء كاره للماء ومستقرة للغاية. تم إجراء تحليل CTLpred وABCpred للكشف عن مستضدات الخلايا B وT. **الاستنتاج:** جينومات SARS-CoV-2 العراقية تحتوي على العديد من المتغيرات التي قد تؤثر على شدة المرض، ويمكن أن تكون المتغيرات هدفًا للقاحات أو علاجات عراقية محددة.

الكلمات المفتاحية: كوفيد19 ، الشائك ، تسلسل الجينوم الكامل ، كورونا -2، في الحاسوب.