Received: 25 May 2020 | Revised: 22 June 2020 | Accepted: 2 July 2020

DOI: 10.1002/jcc.26383

RAPID COMMUNICATION

Journal of COMPUTATIONAL CHEMISTRY WILEY

Atomistic simulation reveals structural mechanisms underlying D614G spike glycoprotein-enhanced fitness in SARS-COV-2

I. Olaposi Omotuyi[1,2] | Oyekanmi Nash[3] | O. Basiru Ajiboye[4] | C. Gift Iwegbulam[3] | E. Babatunji Oyinloye[4,5] | O. Abimbola Oyedeji[2] | Z. Abimbola Kashim[3] | Kunle Okaiyeto[6]

[1]Chemo-genomics Research Unit, Department of Biochemistry, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria

[2]Bio-Computing Research Unit, Mols and Sims, Ado-Ekiti, Ekiti State, Nigeria

[3]Center for Genomics Research and Innovation, National Biotechnology Development Agency, Abuja, Nigeria

[4]Phytomedicine, Biochemical Toxicology and Biotechnology Research Laboratories, Department of Biochemistry, College of Sciences, Afe Babalola University, Ado-Ekiti, Nigeria

[5]Biotechnology and Structural Biology (BSB) Group, Department of Biochemistry and Microbiology, University of Zululand, Kwa-Dlangezwa, South Africa

[6]Applied and Environmental Microbiology Research Group (AEMREG), Department of Biochemistry and Microbiology, University of Fort Hare. Alice campus, Alice, South Africa

Correspondence
I. Olaposi Omotuyi, Chemo-genomics Research Unit, Department of Biochemistry, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria.
Email: olaposi.omotuyi@aaua.edu.ng

Abstract

D614G spike glycoprotein (sgp) mutation in rapidly spreading severe acute respiratory syndrome coronavirus-2 (SARS-COV-2) is associated with enhanced fitness and higher transmissibility in new cases of COVID-19 but the underlying mechanism is unknown. Here, using atomistic simulation, a plausible mechanism has been delineated. In G614 sgp but not wild type, increased D(G)614-T859 Cα-distance within 65 ns is interpreted as S1/S2 protomer dissociation. Overall, ACE2-binding, post-fusion core, open-state and sub-optimal antibody-binding conformations were preferentially sampled by the G614 mutant, but not wild type. Furthermore, in the wild type, only one of the three sgp chains has optimal communication route between residue 614 and the receptor-binding domain (RBD); whereas, two of the three chains communicated directly in G614 mutant. These data provide evidence that D614G sgp mutant is more available for receptor binding, cellular invasion and reduced antibody interaction; thus, providing framework for enhanced fitness and higher transmissibility in D614G SARS-COV-2 mutant.

KEYWORDS

COVID-19, molecular dynamics simulation, mutation, SARS-COV-2, spike glycoprotein

1 | BACKGROUND

COVID-19 is global pandemic caused by severe acute respiratory syndrome coronavirus-2 (SARS-COV-2). This disease was first reported in patients presenting pneumonia-like symptom in Wuhan province of China but has now spread to other countries, infecting more than 5 million and with case-to-fatality ratio of ~1.38%.[1] The race to find cure had relied on the genome sequence and with the release of Wuhan reference strain, diagnostic kits, candidate drugs and vaccine development had commenced.[2] Vaccine development may be the hope as most of the re-purposed drug candidates have not produced the desired outcome in clinical settings. Most of the promising vaccine candidates under investigation recognize and bind the spike glycoproteins.[3,4] Therefore, SARS-COV-2 genetic drift, especially in the region encoding spike glycoprotein has implications on its virulence, and more importantly, the potency or applicability of target vaccines in clinical settings. A major finding from a recent study is guanine-to-adenine transition in position 23,403 in the Wuhan reference strain leading to aspartate-to-glycine mutation in spike glycoprotein residue 614 (D614G). Although, the G614 is thought to originate from China or Europe, its intercontinental spread and prevalence in sequenced samples in other continents would suggest strongly that the mutation confers a selective advantage and enhanced fitness over D614.[5]

Among the mechanisms put forward by Korber et al. (2020)[5] to explain the enhanced fitness; perhaps, the most compelling came from

the resolved structure of the spike glycoprotein itself. D614 carboxylic functional group located on the S1 protomer, proximal to the N-terminal domain (NTD) and receptor binding domain (RBD) forms hydrogen bond with hydroxyl group of T859 of viral membrane bound S2 protomer; thus making D614-T859 interaction important in spike protein oligomeric state formation.[6] The loss of this critical interaction in G614 is thought to promote viral glycoprotein shedding. In order to test this hypothesis, following two independent 150 ns, atomistic simulations, the minimal distance between D(G)614 and T859 were monitored with time along the trajectories. Indeed, both were on the average of 1.5 nm apart at the beginning of the simulation, at around 65 ns, G614-T859 inter-residue distance began to increase to 2.0 nm while D614-T859 decreased to approximately 0.5 nm which was maintained throughout the simulation. Notably, at 75 ns, the G614-T859 inter-residue distance further increased to ~2.5 nm and was maintained till about 135 ns but moved back to its initial state of 1.0 nm at the end of the simulation; thus, providing the first evidence in support of S1 shedding from viral-membrane-bound S2 following D614G mutation (Figure 1a). Theoretically, increased shedding, interpreted as increased S1/S2 unit distance in this study, will ultimately facilitate viral attachment to host cells, hence, enhanced fitness, and increased transmissibility.

Another plausible mechanism proposed is its influence on RBD conformation. RBD is required for the receptor angiotensin converting enzyme-2 (ACE2) recognition and binding.[7] Indeed, such binding is known to be preceded by closed-to-open state conformational change in the RBD. How the mutation in position 614 would influence this transitional event is yet to be investigated. To provide further insight, root-mean-square deviation (rmsd) tool and population distribution plot were used to investigate the conformations sampled by the two biosystems in comparison with the open (PDB ID: 6vyb)[2] and closed (PDB ID: 6vxx) states.[2] The results showed that the open state was preferentially sampled by G614 biosystem (Figure 1b) while the D614 biosystem sampled the closed state conformation (Figure 1c). Generally, while these data do not explicitly give information about the communication between 614 residue and RBD, it, however, provide a proof that events at 614 influence conformations around the RBD. To further establish these findings, the conformations sampled were also compared with that of S2-post-fusion core (PDB ID: 6lxt).[8] Recall that once S1 unit disaggregates (shedding), from the spike glycoprotein trimer, and attaches through ACE2 to host cells, the S2 unit re-assembles into a unit leading to the virus-host membrane fusion; the entire fusion process, relies on conformational changes within the heptad repeats (HR1 and HR2; fusion core) of the S2 unit. The conformations sampled by S2 unit in the G614 biosystem were predominantly those of the post-fusion core (Figure 1d), while the wild type sampled the pre-fusion core conformation (data not shown); and clearly, the conformations of the RBD of G614 biosystem were those previously characterized to bind ACE2 (PDB ID: 6lzg, Figure 1e) but not the wild type.[7] These data strongly lend credence to the fitness and selectivity of G614-clade above D614-clade.

Another key consideration in this study was to evaluate how well antibody will bind the conformations sampled by both biosystems. Given that ACE2 and antibody binding regions and RBD are overlapped

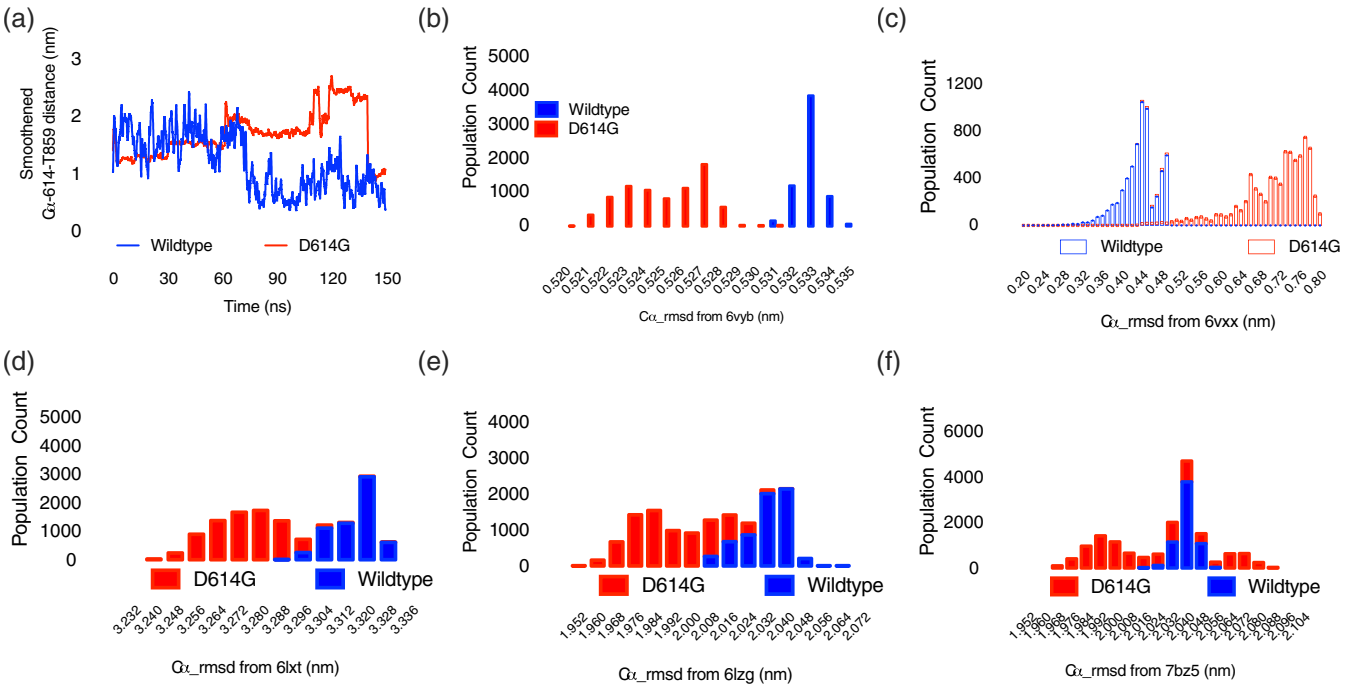

FIGURE 1 Sampled population and dynamics of wild type and D614G SARS-COV-2 spike glycoprotein . (a) Distance-time plot of Ca-distance between residue 614 and T859 during atomistic simulation. (b–f) Population count-rmsd histogram plots; see the x-axes for the proteins used in comparison (unless otherwise stated, each plot represents the mean of 2 independent simulations) [Color figure can be viewed at wileyonlinelibrary.com]

in most cases, it may be taken for granted that since G614 biosystem preferentially sampled ACE2 binding conformation, same may apply for antibody-binding conformation, and its transmissibility success will be lessened due to immune response. The results obtained from comparison with antibody (B38, PDB ID: 7bz5)[9]-bound conformation revealed a rather unusual pattern. G614 biosystems sampled wide range of conformations; those close to the antibody binding conformation (rmsd <2.008 nm) and those non-recognizable by the antibody (> 2.056 nm) while the wild type sampled a very narrow range of conformations (rmsd between 2.016 ~ 2.056 nm) (Figure 1f).

To answer the most important question of how a mutation at a remote site (614) dictates the conformational events at the RBD, network of communities and communication analyses were performed on the trajectories. Community clusters formed by residues is a direct reflection of internal motions and functional dynamics of protein complexes.[10] In the wild type biosystem, one defining feature is the presence of a large community at the N-terminal domain (NTD, blue edge-node; (Figure 2a) while every other regions present a near continuous homogenous community interspaced by small communities. G614 mutant presents a rather unique pattern, revealing three (3) large communities around the NTD, RDB and the region bordering G614 residue towards the viral membrane-binding region while the rest (Figure 2b); just like the wild type is a near continuous stretch of community.

The coronavirus spike glycoprotein NTD confers receptor specificity[11]; the seeming communication with the RBD (G614) may represent key adaptive mechanism for RBD optimization for receptor binding and the diffusion of such information toward the viral membrane may provide the necessary signal for host invasion; thus, providing a good evidence that G614 clade may be more clinically successful than the wild type.

Communication between residue 614 and RBD (T500) was also established; and surprisingly, in the wild type, such optimal communication route could only be found on only one chain (chain B); on the other hand, in G614 biosystem, two of the three chains (chains B and C) have optimal communication route between residue 614 and T500; further providing evidence that 614 mutations have positive implication on the receptor recognition, binding and overall fitness of G614 clade SARS-COV-2.

2 | METHODS

The spike glycoprotein trimer coordinates used for atomistic simulation had been previously reported.[2] The structure is essentially the NTD, RBD, and S2 (1 ~ 1,230) domain; since the fibritin C superfamily region (1,231 ~ 1,270) required for membrane attachment was not

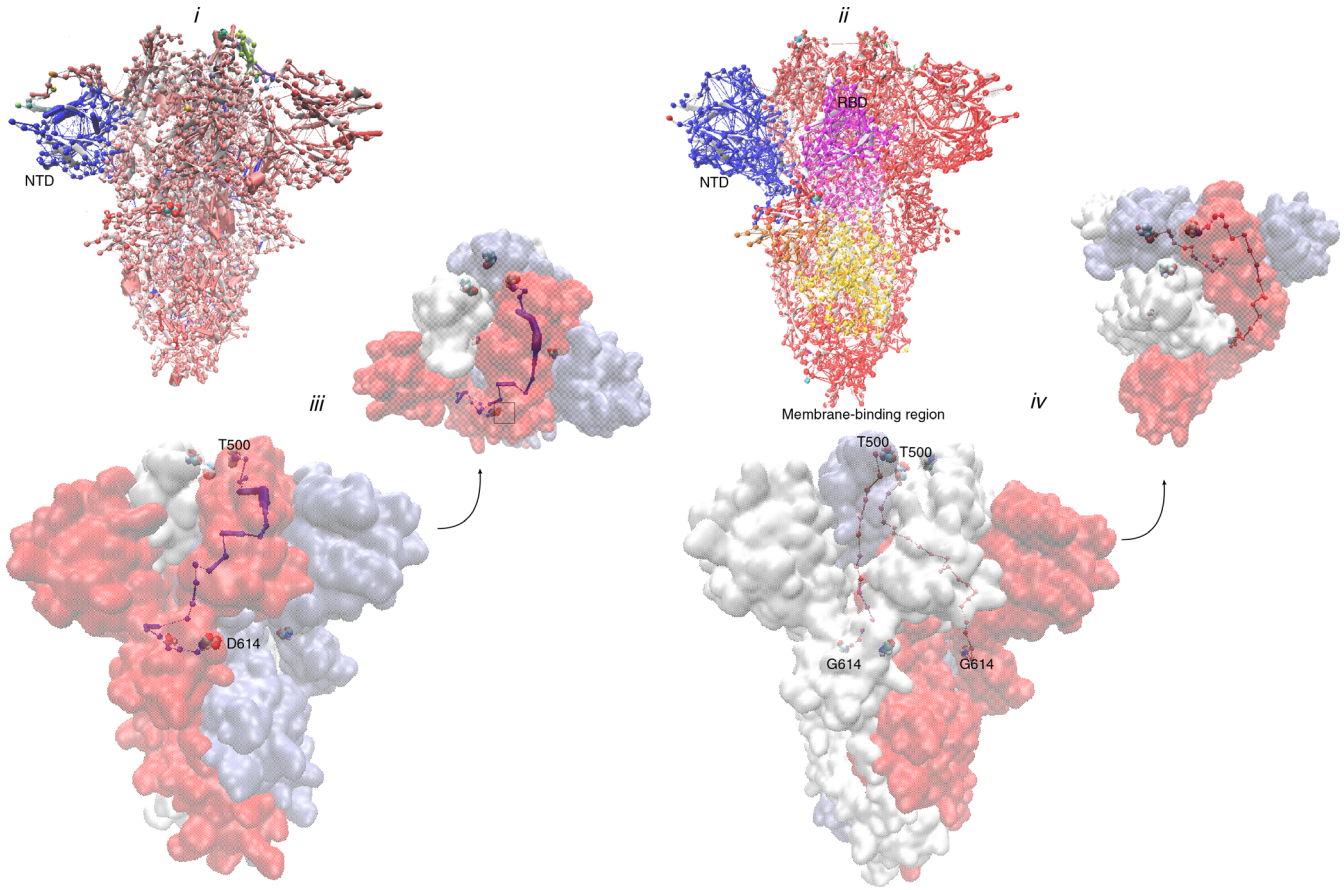

FIGURE 2 Protein network and communication analyses. (a) Communities evolved by the wild type (b) and D614G mutant during simulation. (c) Surface representation of spike glycoprotein trimer showing the optimal communication path between residue 614 and T500 (RBD) in wild type and D614G mutant (d) [Color figure can be viewed at wileyonlinelibrary.com]

captured in the starting structure, it was therefore treated as the soluble fraction without a membrane. Broken chains and incomplete residue re-modeling and glycan removal were performed using DeepView suite. D614G mutant was generated using mutagenesis plugin in PyMol.

TIP3P-solvated, and NaCl-neutralized biosystems were prepared using HTMD platform with charmm36 forcefield parameters.[12] Equilibration (80 ns) and production simulation (150 ns) were subsequently performed[13] using ACEMD3 software[14]. Prior to analysis, all biosystems were checked for convergence (data not shown). RMSD from selected structures (see text) were performed using GROMACS inbuilt *gmx rms* tool, protein community and network analyses were performed using networkTools in VMD software (see supplementary methods for details).[15] All line graphs were plotted as mean of two independent simulations using GraphPad prism (ver. 7.0).

ORCID

I. Olaposi Omotuyi https://orcid.org/0000-0002-5473-745X

REFERENCES

[1] J. Bedford, D. Enria, J. Giesecke, D. L. Heymann, C. Ihekweazu, G. Kobinger, H. C. Lane, Z. Memish, M. D. Oh, A. A. Sall, A. Schuchat, K. Ungchusak, L. H. Wieler, W. H. O. Strategic, *H. Lancet* **2020**, *395*, 1015.
[2] A. C. Walls, Y. J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, D. Veesler, *Cell* **2020**, *181*, 281.
[3] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu, Z. Xiang, Z. Mu, X. Chen, J. Chen, K. Hu, Q. Jin, J. Wang, Z. Qian, *Nat Commun* **2020**, *11*, 1620.
[4] M. Yuan, N. C. Wu, X. Zhu, C. D. Lee, R. T. Y. So, H. Lv, C. K. P. Mok, I. A. Wilson, *Science* **2020**, *368*, 630.
[5] B. Korber, W. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, B. Foley, E. Giorgi, T. Bhattacharya, M. Parker, D. Partridge, C. Evans, T. de Silva, C. LaBranche, D. Montefiori, *bioRxiv* **2020**, 2020.2004.2029.069054.
[6] D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C. L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, *Science* **2020**, *367*, 1260.
[7] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K. Y. Yuen, Q. Wang, H. Zhou, J. Yan, J. Qi, *Cell* **2020**, *181*, 894 e899.
[8] S. Xia, M. Liu, C. Wang, W. Xu, Q. Lan, S. Feng, F. Qi, L. Bao, L. Du, S. Liu, C. Qin, F. Sun, Z. Shi, Y. Zhu, S. Jiang, L. Lu, *Cell Res* **2020**, *30*, 343.
[9] Y. Wu, F. Wang, C. Shen, W. Peng, D. Li, C. Zhao, Z. Li, S. Li, Y. Bi, Y. Yang, Y. Gong, H. Xiao, Z. Fan, S. Tan, G. Wu, W. Tan, X. Lu, C. Fan, Q. Wang, Y. Liu, J. Qi, G. F. Gao, F. Gao, L. Liu, *medRxiv* **2020**, 2020.2005.2001.20077743.
[10] K. Blacklock, G. M. Verkhivker, *PLoS One* **2013**, *8*, e71936.
[11] J. C. Tsai, B. D. Zelus, K. V. Holmes, S. R. Weiss, *J Virol* **2003**, *77*, 841.
[12] J. Huang, A. D. MacKerell Jr., *J Comput Chem* **2013**, *34*, 2135.
[13] O. I. Olaposi, N. Oyekanmi, A. A. Ojo, G. O. Eniafe, *Drug Res (Stuttg)* **2019**, *69*, 643.
[14] M. J. Harvey, G. Giupponi, G. D. Fabritiis, *J Chem Theory Comput* **2009**, *5*, 1632.
[15] O. I. Omotuyi, J. Nagai, H. Ueda, *Sci Rep* **2015**, *5*, 13343.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: IO Omotuyi, O Nash, OB Ajiboye, et al. Atomistic simulation reveals structural mechanisms underlying D614G spike glycoprotein-enhanced fitness in SARS-COV-2. *J Comput Chem*. 2020;41:2158–2161. https://doi.org/10.1002/jcc.26383