



# From oncoproteins to spike proteins: the evaluation of intramolecular stability using hydrophobic force field

Federica Agosta<sup>1</sup> · Glen E. Kellogg<sup>2</sup> · Pietro Cozzini<sup>1</sup>

Received: 26 July 2022 / Accepted: 15 September 2022 / Published online: 31 October 2022  
© The Author(s) 2022

## Abstract

Evaluation of the intramolecular stability of proteins plays a key role in the comprehension of their biological behavior and mechanism of action. Small structural alterations such as mutations induced by single nucleotide polymorphism can impact biological activity and pharmacological modulation. Covid-19 mutations, that affect viral replication and the susceptibility to antibody neutralization, and the action of antiviral drugs, are just one example. In this work, the intramolecular stability of mutated proteins, like Spike glycoprotein and its complexes with the human target, is evaluated through hydrophobic intramolecular energy scoring originally conceived by Abraham and Kellogg based on the “Extension of the fragment method to calculate amino acid zwitterion and side-chain partition coefficients” by Abraham and Leo in *Proteins: Struct. Funct. Genet.* 1987, 2:130–52. HINT is proposed as a fast and reliable tool for the stability evaluation of any mutated system. This work has been written in honor of Prof. Donald J. Abraham (1936–2021).

**Keywords** Intramolecular stability · Scoring function ·  $\text{Log}P_{o/w}$  · Mutations

## Abbreviations

|           |   |
|-----------|---|
| ACE2      | Angiotensin-Converting Enzyme 2                 |
| HINT      | Hydrophobic Interaction                         |
| $P_{o/w}$ | Partition Coefficient between octanol and water |
| RAF       | Rapidly Accelerated Fibrosarcoma Protein        |
| RAS       | Rat Sarcoma Virus protein                       |
| RBD       | Receptor Binding Domain                         |
| SASA      | Solvent Accessible Surface Area                 |

## Introduction

Intramolecular interactions play a central role in the polypeptide folding process leading to a protein native conformation that represents the most energetically favourable structure in a given environment [1], [2]. The analysis of the three-dimensional protein structure and the resulting protein stability can allow an understanding of its molecular function, predict a plausible mechanism of activation, and reveal new strategies for drug discovery or optimization [3], [4].

Native structure modifications can lead to conformational variations that are inherent in the protein’s mechanism of action but are also influenced by single-point alterations due to mutations. These substitutions of native amino acids with others can be characterized by different chemical and physical properties of the protein and alter its ligand-binding mode, function and/or overall stability [5], [6].

In order to analyse protein stability, several computational methods have been developed ranging from molecular mechanics force-field to the most recent machine-learning tools [7]. However, evaluating the relative stability of proteins after mutation is experimentally difficult and can be time-consuming, especially when considering a huge number of potential mutants. The stability of mutated proteins is shown to be influenced by intramolecular interactions, the

✉ Pietro Cozzini  
pietro.cozzini@unipr.it

<sup>1</sup> Molecular Modeling Laboratory, Food and Drug Department, University of Parma, Parco Area delle Scienze 17/A, 43124 Parma, Italy

<sup>2</sup> Department of Medicinal Chemistry and Institute for Structural Biology, Drug Discovery and Development, Virginia Commonwealth University, 3298-0133 Richmond, VA, USA

solvent-accessible surface area, and amino-acid hydrophobicity [8]. These three key aspects are taken into account within HINT.

HINT, Hydropathic INteractions, is a program designed for quantifying hydrophobic and polar interactions between or within molecules in the biological environment [9–11]. It was motivated by the work of Abraham and Leo [12] that extended the fragment method for predictions of  $\log P_{o/w}$  (partition coefficient for solute transfer in 1-octanol/water) to amino acid residues. Briefly, HINT uses these  $\log P_{o/w}$  data as the basis for a forcefield describing intermolecular and intramolecular interactions.

The HINT intramolecular energy function can be rapidly calculated from the structure in terms of a score summing atom-atom interactions. In previous work, Koparde et al. [13] showed that this score can be used as a target function to guide an X-ray crystallography refinement protocol that showed significant improvement for low-resolution data (poorer than 3.0 Å), as evaluated with the global Ramachandran score and the MolProbity clashscore [14].

This work applies the intramolecular HINT force field to evaluate the stability of mutated proteins with a particular focus on the COVID-19 Spike glycoprotein whose mutations are known to modify the affinity toward the human ACE2 and/or several antibodies. The ability of the function to estimate small energy changes related to stability induced by mutations is firstly evaluated by examining a set of oncoproteins whose mutation effects on protein stability and biological behavior are widely known. In particular, the RAS-RAF-MEK-ERK signal transduction cascade is a chain of intramolecular kinase proteins that control cellular proliferation, differentiation, and apoptosis [15]. RAS proteins exist in two different conformations: a GDP-bound state (inactive) and a GTP-bound one (active) [16]. Mutations that occur in codons 12, 13, or 61 are able to alter this conformational equilibrium with a significant stabilization of the active conformation leading to a permanently activated protein and hyperproliferative disorders [17], while mutations to RAF increase its kinase activity [18]. The Intramolecular HINT scoring of this work is applied to estimate the effects of all described point mutations on conformational stability with respect to available experimental data.

## Material and Methods

### Protein Preparation

Protein structures were downloaded from the Protein Data Bank (<https://www.rcsb.org/>). Since not all mutated structures were available, oncoprotein and spike protein mutations were introduced into the wild-type structure using the

Pymol mutagenesis module (<https://pymol.org/2/>). Mutated proteins were solvated in a triclinic water box in a radius of 15 Å using the TIP3P water model, neutralized and minimized in Gromacs v.2019.4 (<https://www.gromacs.org/>), choosing the Amber force field. Hydrogens necessary for the partition step were introduced into the topology generation.

### Protein log P Partitioning

Protein partitioning, i.e., assigning atomic hydropathic parameters to each atom, was carried out using HINT [10], [11] with the “Dictionary” option and choosing the “semi-essential” hydrogen treatment that explicitly includes polar, unsaturated, and alpha to heteroatom hydrogens. All polar and unsaturated  $\text{CH}_x$  hydrogens were considered as possible H-bond donors. This option treats these hydrogens as potential hydrogen bond donors without excessively diluting the hydrophobic density. Proteins were partitioned under a neutral pH solvent condition as in a normal physiological environment. The hybridized pi/lone pairs directionality vector, derived from the geometry of the atom and its attachments with tetrahedral, trigonal bipyramidal, or octahedral classes, was chosen in order to optimize the approach direction of two interaction atoms. Applying a vector focus = 1 forces a smaller and tighter cone of interaction to be scored favourably. Optimum interactions occur when explicit ion pairs and pi orbitals of polar atoms and unsaturated atoms are oriented toward suitable electron acceptor atoms, while nonpolar atoms are treated as spherical.

### Intramolecular stability scoring

The Intramolecular HINT score was calculated as a summation of hydropathic interactions between all atom pairs ( $\sum \sum b_{ij}$ ,  $i=1$  to  $N$ ,  $j=i+1$  to  $N$ ) – excluding those in 1–2 (bonded) and 1–3 (angle) sets – considering the solvent-accessible surface area ( $S$ ), the hydrophobic atom constant ( $a$ ) and the functional distance behavior for the interaction of two different atoms ( $R_{ij}$ ) [10]:

$$B = \sum \sum b_{ij}$$

$$b_{ij} = S_i a_i S_j a_j R_{ij}$$

Since only the heavy atoms are known exactly and hydrogen atom positions are modeled, hydrogen bonds were described as heavy atom-heavy atom distances. Above 3.65 Å the interaction was classified as Acid/Base. Positive values represent favorable binding situations such as hydrophobic-hydrophobic, hydrogen bond, acid-base, and Coulombic, while a negative value represents unfavorable

interactions such as acid-acid, base-base, and hydrophobic-polar, which represents desolvation.

### Total computational cost

The minimization phase for large systems, such as the mutated RBD-ACE2 structure which consists of 12,531 protein atoms, requires 30 min of CPU time in a local HPC system, while protein partitioning and intramolecular calculation less than min of computing time.

## Results and discussion

Small protein structural variations can be related to missense mutations generated by single nucleotide polymorphisms, which is the most common human genetic alteration. These mutations, however, occur with different frequencies [19] leading to possibly altered responses to pathogens, chemical agents, or drugs, or altered structural and/or functional characteristics of encoded proteins [20]. Mutations in the Covid-19 Spike glycoprotein are often associated with higher transmissibility, high virulence, and reduced susceptibility to antibody neutralization [21]. Characterizing the stability of thus mutated proteins and their intramolecular connections are important prerequisites for quickly assessing the possible effects of such mutations on the mechanism of action or drug susceptibility.

In order to verify the sensitivity of HINT to estimate these small structural changes, we first applied it to oncoproteins, some of the best-known examples of proteins where altered structures are associated with hyperproliferative developmental disorders and cancer. RAS and RAF are intramolecular kinase proteins that control cellular proliferation, differentiation, and apoptosis [15]. RAS proteins are a small GTPase protein family [22] that exist in two different states: the GTP-bound state (active) and GDP-bound state (inactive) [16]. They possess N, K, and H isoforms; mutations at codons 12, 13, or 61 promote GTP binding and produce constitutive activation of RAS, producing uncontrolled cell proliferation [17]. RAF kinases are a family of three serine/threonine kinases proteins of which B-RAF has the highest basal kinase activity [23]. Sequence analyses of B-RAF genes have identified different mutations in the kinase domain related to this increased kinase activity [18].

All the 18 known mutations of RAS proteins were taken into account and introduced into the wildtype structure of each isoform considering both active and inactive conformations. In total, 76 mutated structures of K- and N-RAS were thus generated. Crystallographic structures of H-RAS have already revealed higher flexibility of the two adjacent regions delimiting the binding pocket known as switch I

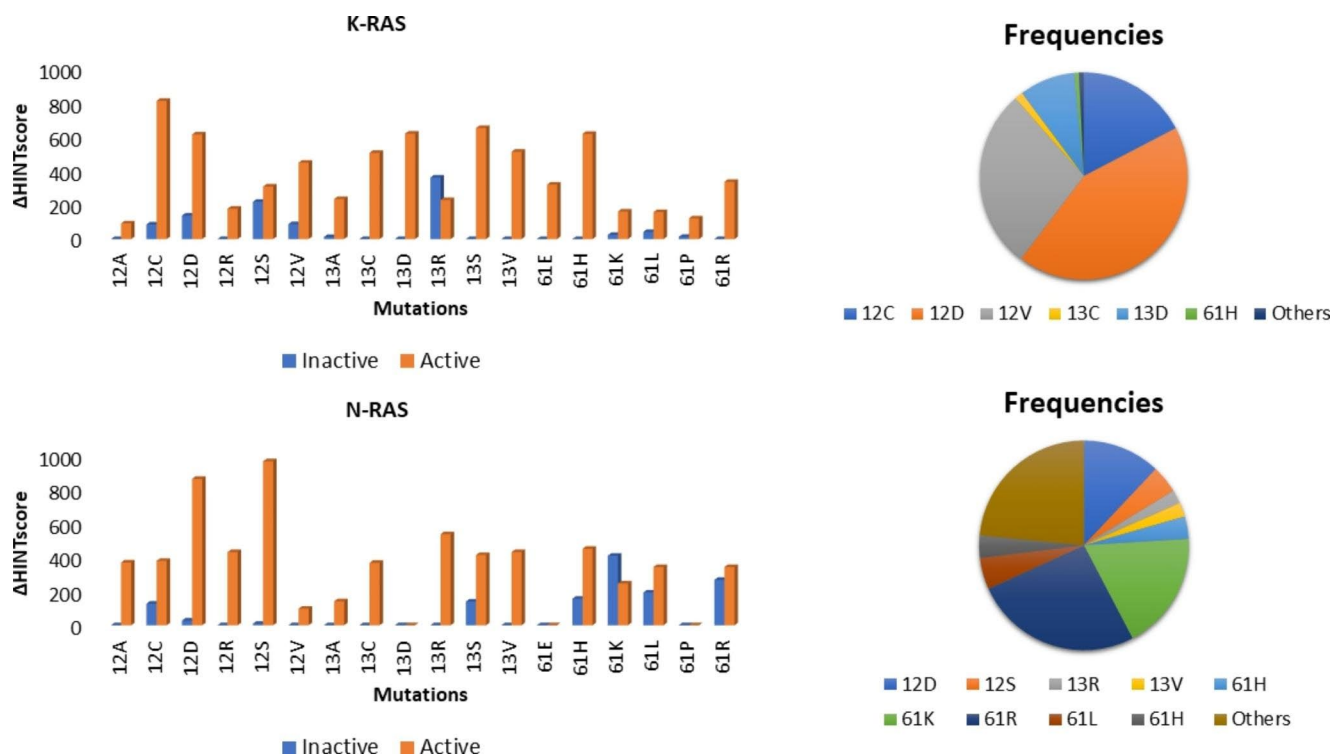
(residues 32–38) and switch II (residues 60–75) [24]. The first is a single loop while the latter consists of a loop and an  $\alpha$ -helix. The active conformation (GTP-bound) shows two possible different orientations of these structural elements, leading to two possible states called state 1 and state 2 in a dynamic equilibrium. Since the recruitment of the effector proteins induces a shift of the conformational equilibrium toward state 2, it is described as the real active state [25]. Considering 18 possible mutations and three different conformations, a total of 57 H-RAS proteins were generated by molecular modeling in our study. Lastly, 18 different B-RAF mutations were introduced into the kinase domain to evaluate their possible effects on intramolecular stability. All mutated structures, after solvation and neutralization, were energy minimized and the Intramolecular HINT scores were calculated for each.

The mutated *inactive* K- and N-RAS conformations show minimal variations of intramolecular HINT score compared to the wildtype with a not significant  $\Delta$ HINT score between mutated and wildtype form as shown in Fig. 1. However, these mutations stabilize the *active* conformations that exhibit higher intramolecular values than the wild type. So, if the intramolecular HINT score of the mutated state is higher than that of the wildtype, the active-inactive conformational equilibrium is shifted towards the *active* conformation that is *more stable* than the wildtype (Fig. 1). The stabilization of the closed conformation would thus explain the ability of high-scoring mutations to permanently activate the protein. Also of note is that the mutants with the highest experimental frequency are the most stable (Fig. 1) [16].

Considering the three different known H-RAS conformations, we observed that mutations destabilized the *inactive* conformations and generated a stabilization of the *active state 2* conformations. This result is in line with the explained mechanism of action of H-RAS for which only the state 2 conformation is able to recruit co-activator proteins [25]. According to our calculations, G12S, G12V, G13R, and Q61R are the most stable mutants and present the highest experimental frequency as shown in Fig. 2 [16].

Although a general stabilization of all RAS mutated active conformation is observed, different mutations produce a different effect on the intramolecular stability due to the nature of the amino-acid substitution and more complex biological mechanism that involve the residues Q61 that show a more significant conformational change when the active conformation dimerize with effector proteins [26].

The intramolecular HINT scores reveal that all *active* mutated B-RAF proteins are as stable as, or more than, the wildtype. Wan et al. [18] describe the experimental kinase activity of B-RAS mutants. V599D is the registered mutation that most greatly increased kinase activity; our



**Fig. 1** K and N-RAS Intramolecular HINT score. K-RAS: The intramolecular stabilities of the mutated inactive (PDB ID: 5W22) and active (PDB ID: 6GOD) conformations of K-RAS were calculated. G12C, G12D, G12V, G13D, G13C, and Q61H are the most stable K-RAS closed conformations and present the highest experimental frequency.

calculations also show that it shows increased protein stability as estimated by the highest intramolecular HINT score value (Fig. 3).

Protein stabilization after mutation is often related to local and generally minimal structural variations that include the creation of one or more hydrogen bonds, acquisition of hydrophobic interactions, or reduction of steric bulk. The intramolecular HINT function is sensitive to each of these small changes in the protein's energetic stability, but also records and accounts for unfavorable interactions such as changes in desolvation energy (unfavorable hydrophobic-polar score) and repulsive Coulombic interactions. In general, the intramolecular HINT function score produces results quite quickly, and these results can be correlated with experimental measurements such as demonstrated above. Also, analysis of the output on a residue pair-by-residue pair basis or even atom-by-atom basis often reveals specific information on the energetics involved in the mutation as shown in Fig. 4.

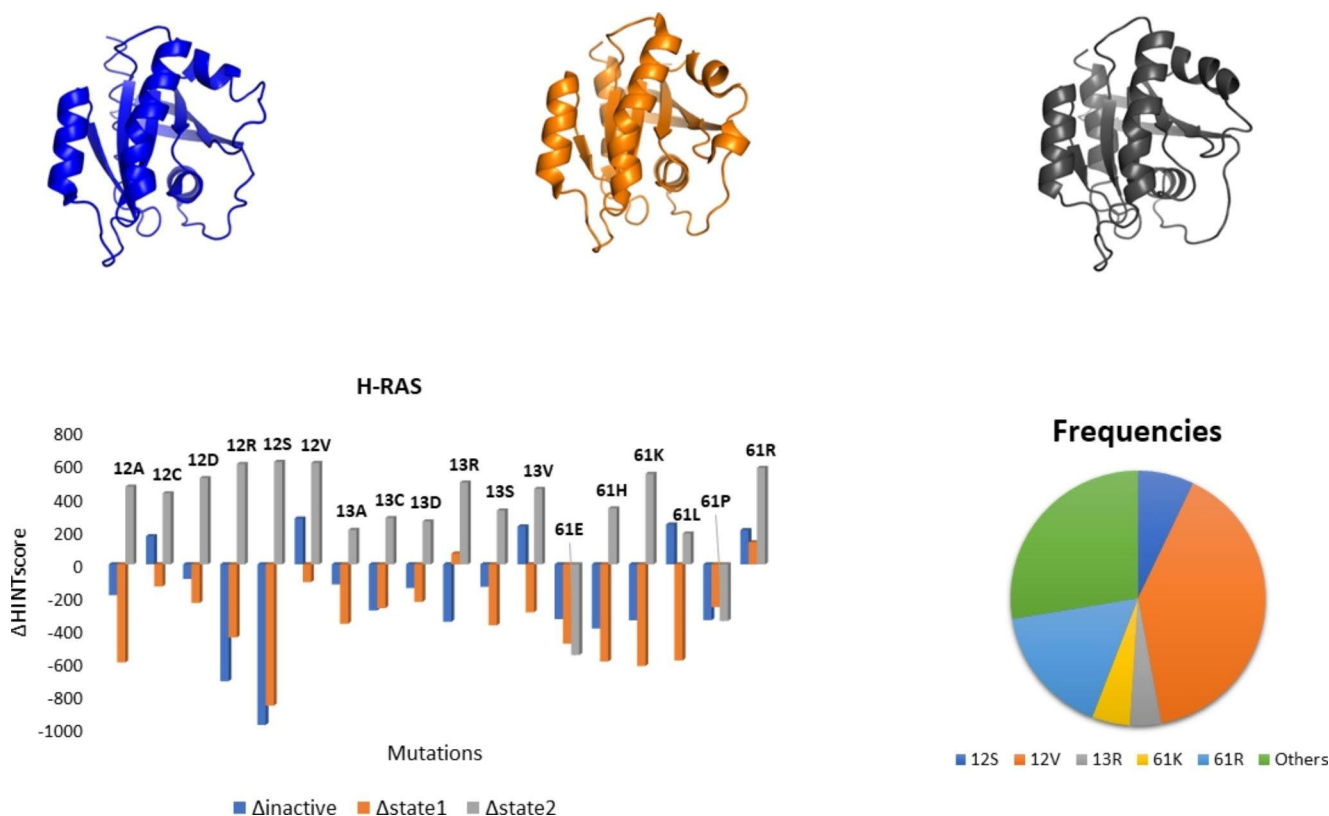
### Sars Cov2 Spike Intramolecular stability

The spike glycoprotein is a well-known and increasingly studied Covid-19 structural protein involved in virus

H-RAS: The intramolecular stability of the mutated inactive (PDB ID: 6WGH) and active (PDB ID: 5UHV) conformations reveal that G12D, G12S, and Q61R are the most stable and frequent N-RAS closed conformations [16].

replication due to the interaction between its receptor-binding domain and the human angiotensin-converting enzyme 2 [27]. Several mutations have been identified in different domains that can affect this interaction with the human target or lead to a reduced response to antibodies [28]. Since the spike glycoprotein is one of the main targets of vaccines, antibodies and drugs, understanding the effect of these mutations is necessary to guarantee their long-term efficacy and to monitor the clinical impact of the variant's diffusion. Previous studies have already demonstrated that protein stability is necessary for survival and diffusion [29] by monitoring the thermodynamic effect of a single mutation on the spike/ACE2 complex stability [30]. Evaluating the effect of a single mutation could be reductive since each variant of concern is characterized by several simultaneous structural mutations. Our aim is to use the intramolecular HINT score to evaluate the stability of the known variants in the closed trimeric conformation of spike glycoprotein, the effect of mutations on the stability of the receptor-binding domain as is, and in complex with the human target.

Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2), Gamma (P.1), and Omicron (BA.1) are the principal circulating variants, defined as variants of concern by the U.S. Centers for Disease Control and Prevention (CDC) (<https://>



**Fig. 2** H-RAS structures and Intramolecular HINT score. Three different conformations of H-RAS were identified: an inactive GDP-bound conformation (blue) (PDB ID: 4Q21), an active state 1 conformation (orange) (PDB ID: 3RSO), and an active state 2 conformation (grey)

(PDB ID: 5P21). The  $\Delta$ HINT score between the mutated and the wild-type structure was calculated considering all three possible conformations. Only state 2 active conformation is stabilized by mutations

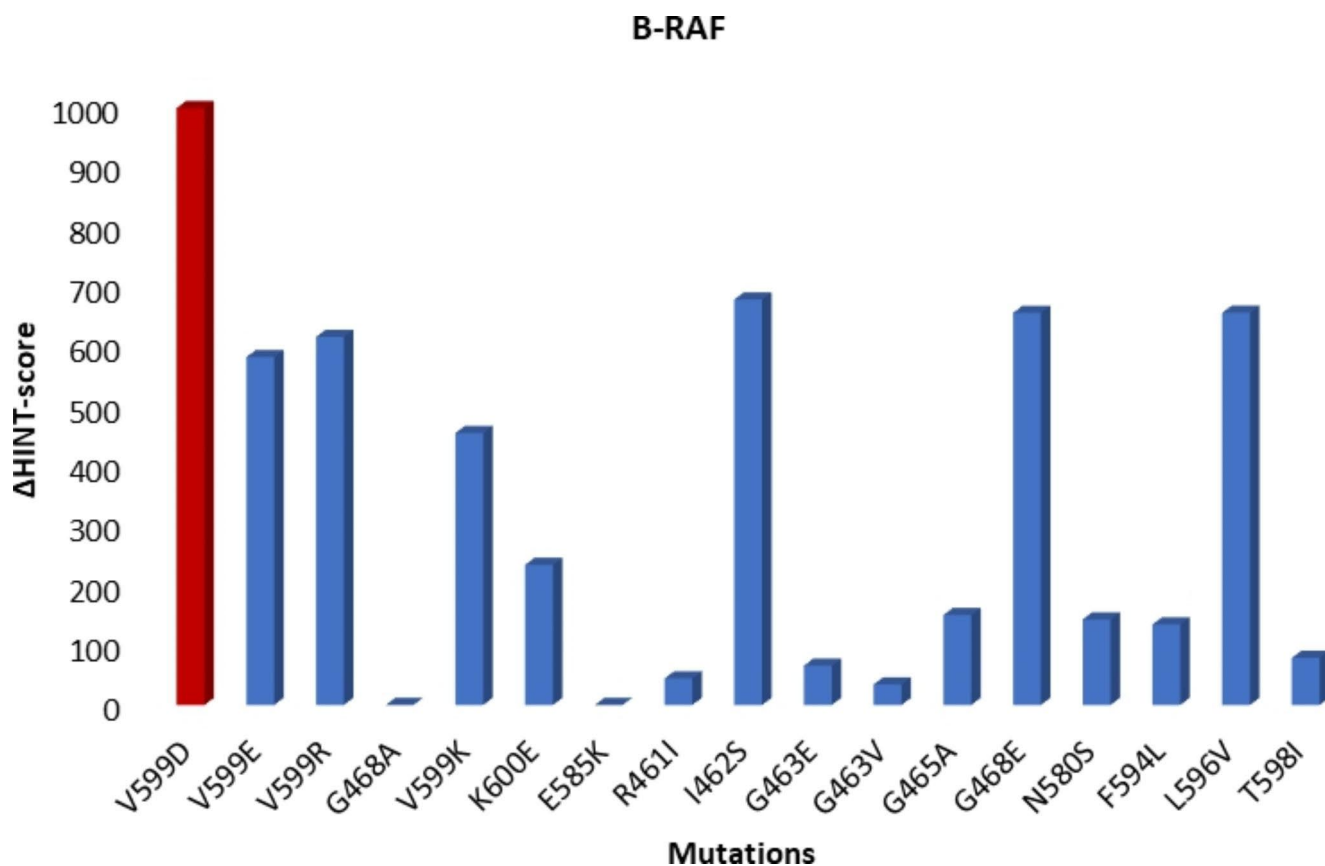
[www.cdc.gov/](http://www.cdc.gov/)). Their structural characteristics were retrieved from GISAID [31]. Mutations were introduced in the wild-type structures and minimized. The Intramolecular HINT score evaluation shows that all mutated trimeric closed structures are at least as stable as the wildtype and all mutated receptor binding domains (RBDs) are more stable than the wildtype (Fig. 5). Alpha, Beta, and Gamma variants are characterized by one to three mutations in the receptor-binding domain with consequently minimal local variations in the structural stability. The BA.1 (Omicron) variant, despite a huge number of mutations, is stable. The Delta variant's receptor binding domain presents a lower Intramolecular HINT score than other variants due to L452R mutation, which involves the substitution of a hydrophobic residue (leucine) with a basic and sterically bulky one (arginine) in one of the  $\beta$ -sheets that connect the flexible loop to the domain core.

The RBD-ACE2 complexes are all as or more stable than wildtype ones – thus underlining the variant's higher affinity toward the human target. This effect is related to the increased local interactions generated by some substitutions like N501Y that characterized the RBD of Alpha, Beta, and Gamma variants, by acquiring a new  $\pi$ - $\pi$  interaction. The

Omicron-ACE2 complex is one of the most stable since the Omicron mutations introduce local positive charges responsible for electrostatic interaction with ACE2, which is negatively charged [32].

## Conclusion

Intramolecular HINT energy scoring allows calculating all hydrophobic interactions (hydrophobic and polar) occurring within a biological macromolecule in evaluating their thermodynamic stability. It allows the evaluation of the effect of point mutations on the stability of different conformations of the same protein with a good level of agreement with the experimental data. Given its sensitivity in estimating small energy differences due to the loss or acquisition of single bonds, this function is used to evaluate the stability of the COVID\_19 spike glycoprotein of the main variants of concern considering the effect of the multiple mutations occurring in the closed trimeric conformation and in their receptor-binding domain even in complex with the human ACE2.



**Fig. 3** B-RAF Intramolecular HINT scores. The  $\Delta$ HINT score between mutated and wildtype structure (PDB ID: 6XFP) was calculated. All mutated forms are more stable than the wild-type ones. V599D is the most stable and presents the higher experimental kinase activity

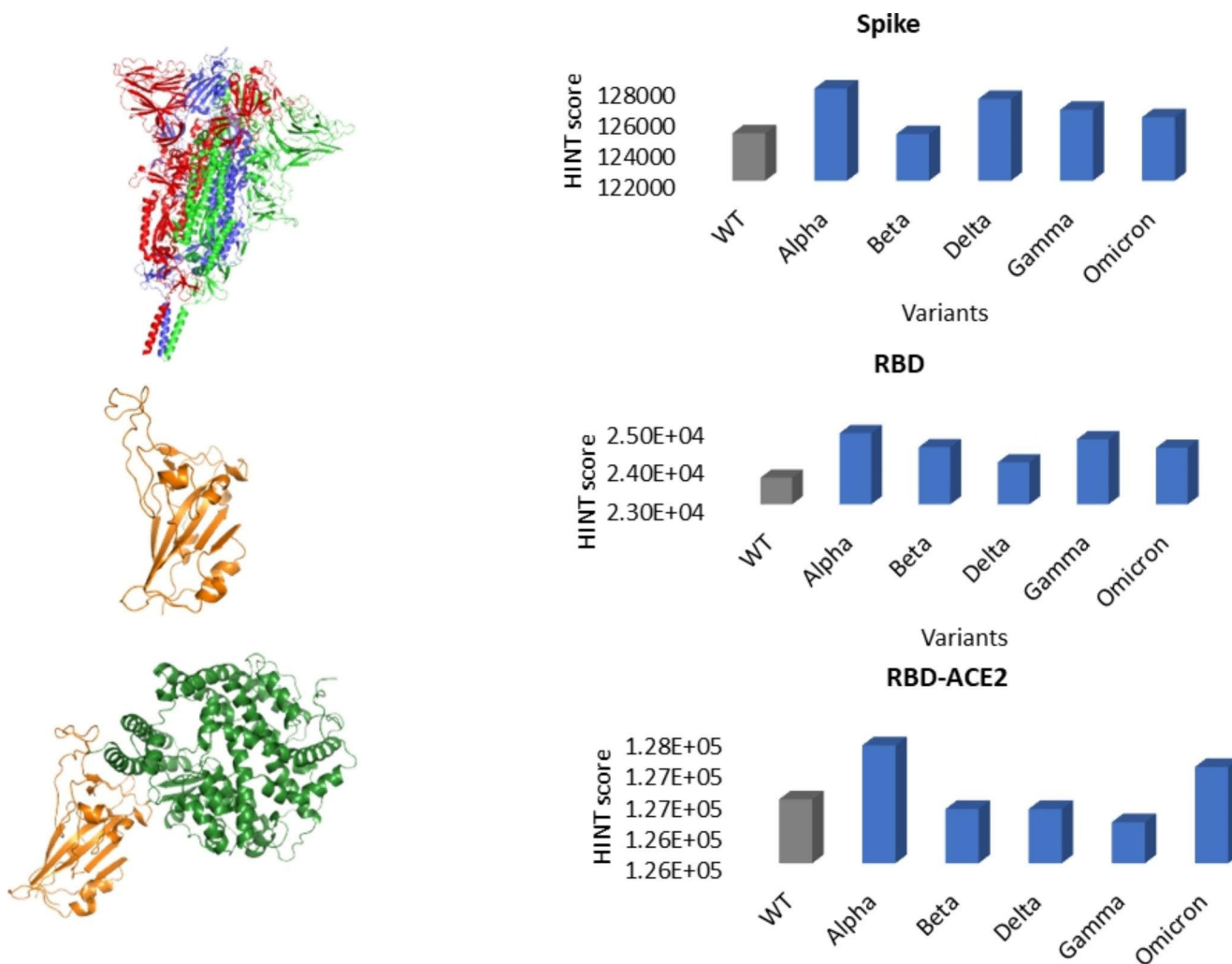
| Monomer Type                 | Atom Name | Atom Type | Hydrophobic Atom Const. | SASA         | Monomer Type | Atom Name | Atom Type | Hydrophobic Atom Const. | SASA  | Distance Angstroms | VDW   | Interaction Score | Interaction Type |     |                |
|------------------------------|-----------|-----------|-------------------------|--------------|--------------|-----------|-----------|-------------------------|-------|--------------------|-------|-------------------|------------------|-----|----------------|
| ....                         |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| ALA                          | ALA11     | CB        | C.3:3                   | 0.810        | 75.00        | VAL       | VAL14     | CG1                     | C.3:3 | 0.818              | 61.50 | 4.118             | 121.11           | 51  | Hydrophobic    |
| ALA                          | ALA11     | CB        | C.3:3                   | 0.810        | 75.00        | VAL       | VAL14     | CG2                     | C.3:3 | 0.818              | 61.50 | 5.465             | 160.74           | 13  | Hydrophobic    |
| ALA                          | ALA11     | CB        | C.3:3                   | 0.810        | 75.00        | SER       | SER89     | CB                      | C.3:1 | 0.489              | 51.00 | 3.733             | 109.80           | 25  | Hydrophobic    |
| ALA                          | ALA11     | CB        | C.3:3                   | 0.810        | 75.00        | SER       | SER89     | OG                      | 0.3   | -0.898             | 31.00 | 4.932             | 154.13           | -17 | Hydroph./Polar |
| ALA                          | ALA11     | CB        | C.3:3                   | 0.810        | 75.00        | ASP       | ASP92     | CB                      | C.3:2 | 0.489              | 41.00 | 4.644             | 136.60           | 13  | Hydrophobic    |
| ARG                          | ARG12     | N         | N.am                    | -0.297       | 32.00        | GLU       | GLU62     | OE1                     | O.co2 | -3.310             | 36.00 | 3.465             | 113.62           | 141 | Hydrogen Bond  |
| ARG                          | ARG12     | N         | N.am                    | -0.297       | 32.00        | GLU       | GLU62     | OE2                     | O.co2 | -3.310             | 36.00 | 4.127             | 135.30           | 89  | Hydrogen Bond  |
| ARG                          | ARG12     | O         | 0.2                     | -1.915       | 27.00        | GLY       | GLY13     | CA                      | C.3:1 | 0.480              | 53.00 | 2.837             | 88.65            | -16 | Hydroph./Polar |
| ARG                          | ARG12     | O         | 0.2                     | -1.915       | 27.00        | GLY       | GLY13     | O                       | 0.2   | -1.915             | 30.00 | 4.222             | 140.74           | -14 | Base/Base      |
| ARG                          | ARG12     | O         | 0.2                     | -1.915       | 27.00        | VAL       | VAL14     | N                       | N.am  | -0.703             | 32.00 | 3.415             | 111.98           | 45  | Hydrogen Bond  |
| ....                         |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| Total Interaction Constant   |           |           |                         | 2.682601e+04 |              |           |           |                         |       |                    |       |                   |                  |     |                |
| Contributions:               |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| 3.758318e+04 Hydrogen Bond   |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| 1.344848e+04 Acid/Base       |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| 1.384005e+04 Hydrophobic     |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| -1.048634e+04 Acid/Acid      |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| -1.793159e+03 Base/Base      |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |
| -2.576621e+04 Hydroph./Polar |           |           |                         |              |              |           |           |                         |       |                    |       |                   |                  |     |                |

**Fig. 4** Intramolecular HINT output file. The output file shows each atom-atom interaction and the total intramolecular score with a detailed description of the single energy contribution (hydrogen bond,

electrostatic and hydrophobic interactions, and negative contributions). In this capture, for example, the mutated ARG12 establishes two different hydrogen bonds with the residue GLU62.

Intramolecular HINT scoring can therefore have multiple applications and is presented as a fast, effective, and

versatile tool.



**Fig. 5** Intramolecular HINT score of Spike proteins. All mutants generated from the trimeric wildtype closed conformation (PDB ID: 6VXX) are as stable as the wild type. RBD HINT score reveals higher

stability of mutants than wildtype. Alpha and Omicron variants present a higher affinity toward ACE2, and their complexes are the most stable

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10822-022-00477-y>.

**Acknowledgment** This present research work is funded by L.A.V.-Lega Anti Vivisezione and by a grant from High-Performance Computing Facilities of Centro di Calcolo di Ateneo, University of Parma.

**Author contributions** All authors contributed equally to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript.

**Funding** Open access funding provided by Università degli Studi di Parma within the CRUI-CARE Agreement.

**Data and Software availability** All structural data was extracted from the Protein Data Bank. All algorithms and formulas for calculations we performed and reported are presented within this manuscript and/or in reference within. Readers with questions or who wish to access HINT are encouraged to contact the program's author, Glen E. Kellogg (gkellogg@vcu.edu).

## Declarations

**Supporting Information** The HINTscores related to all conformational protein structures evaluated in this work are provided in a .xls format: S1. RAS HINTscores (K-N and H isoforms) show the effect of all mutations on protein stability; S2. B-RAF HINTscores; S3. SPIKE proteins HINT score considering the trimeric conformation, the receptor-binding domain, and the RBD-ACE2 complexes.

**Statements and Declaration** The authors have no competing interests to declare.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted

use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anfinsen CB (1973) “Principles that Govern the Folding of Protein Chains,” *Science* (1979), vol. 181, no. 4096, pp. 223–230, Jul. <https://doi.org/10.1126/science.181.4096.223>
- Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) “Funnel, pathways, and the energy landscape of protein folding: A synthesis,” *Proteins: Structure, Function, and Genetics*, vol. 21, no. 3, pp. 167–195, Mar. <https://doi.org/10.1002/prot.340210302>
- Godoy-Ruiz R, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM (Feb. 2004) Relation Between Protein Stability, Evolution and Structure, as Probed by Carboxylic Acid Mutations. *J Mol Biol* 336(2):313–318. <https://doi.org/10.1016/j.jmb.2003.12.048>
- Goldstein RA (2011) “The evolution and evolutionary consequences of marginal thermostability in proteins,” *Proteins: Structure, Function, and Bioinformatics*, vol. 79, no. 5, pp. 1396–1407, May <https://doi.org/10.1002/prot.22964>
- Kucukkal TG, Petukh M, Li L, Alexov E (2015) “Structural and physico-chemical effects of disease and non-disease nsSNPs on proteins,” *Curr Opin Struct Biol*, vol. 32, pp. 18–24, Jun. <https://doi.org/10.1016/j.sbi.2015.01.003>
- Petukh M, Kucukkal TG, Alexov E (May 2015) On Human Disease-Causing Amino Acid Variants: Statistical Study of Sequence and Structural Patterns. *Hum Mutat* 36(5):524–534. <https://doi.org/10.1002/humu.22770>
- Kulshreshtha S, Chaudhary V, Goswami GK, Mathur N (May 2016) Computational approaches for predicting mutant protein stability. *J Comput Aided Mol Des* 30(5):401–412. <https://doi.org/10.1007/s10822-016-9914-3>
- Caldararu O, Blundell TL, Kepp KP (Apr. 2021) Three Simple Properties Explain Protein Stability Change upon Mutation. *J Chem Inf Model* 61(4):1981–1988. <https://doi.org/10.1021/acs.jcim.1c00201>
- Kellogg GE, Semus SF, Abraham DJ (1991) “HINT: A new method of empirical hydrophobic field calculation for CoMFA,” *J Comput Aided Mol Des*, vol. 5, no. 6, pp. 545–552, Dec. <https://doi.org/10.1007/BF00135313>
- Eugene Kellogg G, Abraham DJ (2000) “Hydrophobicity: is LogPo/w more than the sum of its parts?,” *Eur J Med Chem*, vol. 35, no. 7–8, pp. 651–661, Aug. [https://doi.org/10.1016/S0223-5234\(00\)00167-7](https://doi.org/10.1016/S0223-5234(00)00167-7)
- Sarkar A, Kellogg G (2010) “Hydrophobicity - Shake Flasks, Protein Folding and Drug Discovery,” *Curr Top Med Chem*, vol. 10, no. 1, pp. 67–83, Jan. <https://doi.org/10.2174/156802610790232233>
- Abraham DJ, Leo AJ (1987) Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. *Proteins Struct Funct Genet* 2(2):130–152. <https://doi.org/10.1002/prot.340020207>
- Koparde VN, Scarsdale JN, Kellogg GE (Jan. 2011) Applying an Empirical Hydrophobic Forcefield in Refinement May Improve Low-Resolution Protein X-Ray Crystal Structures. *PLoS ONE* 6(1):e15920. <https://doi.org/10.1371/journal.pone.0015920>
- Davis JP, Foegeding EA (May 2006) Foaming and Interfacial Properties of Polymerized Whey Protein Isolate. *J Food Sci* 69(5):C404–C410. <https://doi.org/10.1111/j.1365-2621.2004.tb10706.x>
- Hilger RA, Scheulen ME, Strumberg D (2002) The Ras-Raf-MEK-ERK Pathway in the Treatment of Cancer. *Oncol Res Treat* 25(6):511–518. <https://doi.org/10.1159/000068621>
- Prior IA, Hood FE, Hartley JL (2020) “The Frequency of Ras Mutations in Cancer,” *Cancer Res*, vol. 80, no. 14, pp. 2969–2974, Jul. <https://doi.org/10.1158/0008-5472.CAN-19-3682>
- Quinlan MP, Settleman J (Feb. 2009) Isoform-specific ras functions in development and cancer. *Future Oncol* 5(1):105–116. <https://doi.org/10.2217/14796694.5.1.105>
- Wan PTC et al (Mar. 2004) Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. *Cell* 116(6):855–867. [https://doi.org/10.1016/S0092-8674\(04\)00215-6](https://doi.org/10.1016/S0092-8674(04)00215-6)
- Collins FS, Guyer MS, Chakravarti A (1997) “Variations on a Theme: Cataloging Human DNA Sequence Variation,” *Science* (1979), vol. 278, no. 5343, pp. 1580–1581, Nov. <https://doi.org/10.1126/science.278.5343.1580>
- Wang Z, Moutl J (2001) “SNPs, protein structure, and disease,” *Hum Mutat*, vol. 17, no. 4, pp. 263–270, Apr. <https://doi.org/10.1002/humu.22>
- Harvey WT et al (Jul. 2021) SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* 19(7):409–424. <https://doi.org/10.1038/s41579-021-00573-0>
- Mo SP, Coulson JM, Prior IA (2018) “RAS variant signalling,” *Biochem Soc Trans*, vol. 46, no. 5, pp. 1325–1332, Oct. <https://doi.org/10.1042/BST20180173>
- Morrison DK (May 2001) KSR: a MAPK scaffold of the Ras pathway? *J Cell Sci* 114(9):1609–1612. <https://doi.org/10.1242/jcs.114.9.1609>
- Geyer M, Schweins T, Herrmann C, Prinsner T, Wittinghofer A, Kalbitzer HR (1996) “Conformational Transitions in p21<sup>ras</sup> and in Its Complexes with the Effector Protein Raf-RBD and the GTPase Activating Protein GAP,” *Biochemistry*, vol. 35, no. 32, pp. 10308–10320, Jan. <https://doi.org/10.1021/bi952858k>
- Matsumoto K et al (2011) “Critical Roles of Interactions among Switch I-preceding Residues and between Switch II and Its Neighboring  $\alpha$ -Helix in Conformational Dynamics of the GTP-bound Ras Family Small GTPases,” *Journal of Biological Chemistry*, vol. 286, no. 17, pp. 15403–15412. <https://doi.org/10.1074/jbc.M110.204933>
- Buhrman G, Wink G, Mattos C (2007) “Transformation Efficiency of RasQ61 Mutants Linked to Structural Features of the Switch Regions in the Presence of Raf,” *Structure*, vol. 15, no. 12, pp. 1618–1629, Dec. <https://doi.org/10.1016/j.str.2007.10.011>
- Wan Y, Shang J, Graham R, Baric RS, Li F (2020) “Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus,” *J Virol*, vol. 94, no. 7, Mar. <https://doi.org/10.1128/JVI.00127-20>
- Salleh MZ, Derrick JP, Deris ZZ (Jul. 2021) Structural Evaluation of the Spike Glycoprotein Variants on SARS-CoV-2 Transmission and Immune Evasion. *Int J Mol Sci* 22(14):7425. <https://doi.org/10.3390/ijms22147425>
- Laha S, Chakraborty J, Das S, Manna SK, Biswas S, Chatterjee R (Nov. 2020) Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect Genet Evol* 85:104445. <https://doi.org/10.1016/j.meegid.2020.104445>
- Teng S, Sobitan A, Rhoades R, Liu D, Tang Q (2021) “Systemic effects of missense mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity,” *Brief Bioinform*, vol. 22, no. 2, pp. 1239–1253, Mar. <https://doi.org/10.1093/bib/bbaa233>
- Shu Y, McCauley J (Mar. 2017) GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* 22(13). <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>
- Giordano D, de Masi L, Argenio MA, Facchiano A (Aug. 2021) Structural Dissection of Viral Spike-Protein Binding of SARS-CoV-2 and SARS-CoV-1 to the Human Angiotensin-Converting Enzyme 2 (ACE2) as Cellular Receptor. *Biomedicines* 9(8):1038. <https://doi.org/10.3390/biomedicines9081038>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.