



Shape Complementarity Optimization of Antibody–Antigen Interfaces: The Application to SARS-CoV-2 Spike Protein

Alfredo De Lauro^{1†}, Lorenzo Di Rienzo^{2*†}, Mattia Miotto², Pier Paolo Olimpieri³, Edoardo Milanetti^{2,3} and Giancarlo Ruocco^{2,3}

¹Department of Sciences, Roma Tre University, Rome, Italy, ²Center for Life Nano & Neuro-Science, Istituto Italiano di Tecnologia, Rome, Italy, ³Department of Physics, Sapienza University, Rome, Italy

OPEN ACCESS

Edited by:

Adriana Erica Miele,
Université Claude Bernard Lyon 1,
France

Reviewed by:

Luca Palazzolo,
University of Milan, Italy
Julien Bergeron,
King's College London,
United Kingdom

*Correspondence:

Lorenzo Di Rienzo
lorenzo.dirienzo@iit.it

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Structural Biology,
a section of the journal
Frontiers in Molecular Biosciences

Received: 11 February 2022

Accepted: 07 April 2022

Published: 20 May 2022

Citation:

De Lauro A, Di Rienzo L, Miotto M,
Olimpieri PP, Milanetti E and Ruocco G
(2022) Shape Complementarity
Optimization of Antibody–Antigen
Interfaces: The Application to SARS-
CoV-2 Spike Protein.
Front. Mol. Biosci. 9:874296.
doi: 10.3389/fmolb.2022.874296

Many factors influence biomolecule binding, and its assessment constitutes an elusive challenge in computational structural biology. In this aspect, the evaluation of shape complementarity at molecular interfaces is one of the main factors to be considered. We focus on the particular case of antibody–antigen complexes to quantify the complementarities occurring at molecular interfaces. We relied on a method we recently developed, which employs the 2D Zernike descriptors, to characterize the investigated regions with an ordered set of numbers summarizing the local shape properties. Collecting a structural dataset of antibody–antigen complexes, we applied this method and we statistically distinguished, in terms of shape complementarity, pairs of the interacting regions from the non-interacting ones. Thus, we set up a novel computational strategy based on *in silico* mutagenesis of antibody-binding site residues. We developed a Monte Carlo procedure to increase the shape complementarity between the antibody paratope and a given epitope on a target protein surface. We applied our protocol against several molecular targets in SARS-CoV-2 spike protein, known to be indispensable for viral cell invasion. We, therefore, optimized the shape of template antibodies for the interaction with such regions. As the last step of our procedure, we performed an independent molecular docking validation of the results of our Monte Carlo simulations.

Keywords: antibodies, shape complementarity, molecular interaction, SARS-CoV-2, computational modeling

1 INTRODUCTION

Cellular functioning is widely dependent on processes occurring when biological molecules recognize each other and bind (Jones and Thornton, 1996; Gromiha et al., 2017). In particular, the non-covalent protein–protein pairing proved to be essential in several biochemical pathways, ranging from biocatalysis to organism immunity or cell regulatory network construction (Gavin et al., 2002; Han et al., 2004). Not surprisingly, in the last few decades, a very high amount of effort has been devoted to developing computational tools for the structural characterization of protein–protein complexes. The aim of these methods are various, varying from binding site identification (Gainza et al., 2020; Milanetti et al., 2021a) to binding affinity prediction (Vangone and

Bonvin, 2015; Siebenmorgen and Zacharias, 2020) or protein–protein docking guide (Vakser, 2014; Kozakov et al., 2017; Geng et al., 2020; Yan et al., 2020).

In this scenario, the shape complementarity at the molecular interface is one of the most basic tasks to take into account (Katchalski-Katzir et al., 1992; Lawrence and Colman, 1993; and Jones and Thornton, 1996). Indeed, the evaluation of shape complementarity is essential for docking, both in terms of searching and evaluating the binding poses (Chen and Weng, 2003; Nicola and Vakser, 2007; Kuroda and Gray, 2016; Gromiha et al., 2017; and Yan and Huang, 2019), and represents one of the factors to take into account for binding site recognition (Gainza et al., 2020; Milanetti et al., 2021a) or to assess the binding affinity (Erijman et al., 2014).

Among the wide variety of methods developed in the last few years to describe the geometrical properties of a molecular region and to evaluate the complementarity with a putative binding partner region, using the Zernike polynomials is an effective and promising strategy (Venkatraman et al., 2009a; Di Rienzo et al., 2017; Daberdaku and Ferrari, 2019; and Di Rienzo et al., 2021a). Indeed, once extracted, the molecular surface region and its geometrical properties are summarized through a set of numerical descriptors, namely, the Zernike descriptors. The accuracy of the description is increased by enlarging the number of descriptors considered (Zernike and Stratton, 1934; Canterakis, 1999; and Novotni and Klein, 2004).

The main advantage of the Zernike formalism is that the molecular surface representation is invariant under protein rotation, constituting an absolute morphological characterization of the examined protein region. Therefore, the complementarity between two molecular regions is computed by comparing their Zernike descriptors, without the need for any preliminary superposition step (Daberdaku and Ferrari, 2018; Di Rienzo et al., 2020a).

In the last decade, the Zernike approach, in its 3D version, has been widely applied for the analysis of biomolecules (Venkatraman et al., 2009a; Venkatraman et al., 2009b; Kihara et al., 2011; Di Rienzo et al., 2017; Daberdaku and Ferrari, 2018; Daberdaku and Ferrari, 2019; Han et al., 2019; Di Rienzo et al., 2020a; Alba et al., 2020; and Di Rienzo et al., 2020b), proving its efficacy in characterizing both global and local protein properties.

We recently developed a computational protocol that allows us to employ the 2D Zernike formalism to assess the shape complementarity observed in protein–protein interfaces (Milanetti et al., 2021a). The utilization of the 2D formalism allows to sensibly decrease the computational time needed to compute the shape descriptors without a significant loss in description accuracy (Di Rienzo et al., 2021b).

In this work, we focused on antibody–antigen interactions, since these complexes represent a critical case of molecular recognition where the interface shape complementarity level is similar to the typical protein–protein interfaces (Li et al., 2003; Kuroda and Gray, 2016).

Moreover, antibodies have been the object of extensive biomedical studies since their modular architecture facilitates the engineering of novel binding sites (Gotwals et al., 2017; Saeed et al., 2017; and Singh et al., 2018). Indeed, the recognition of

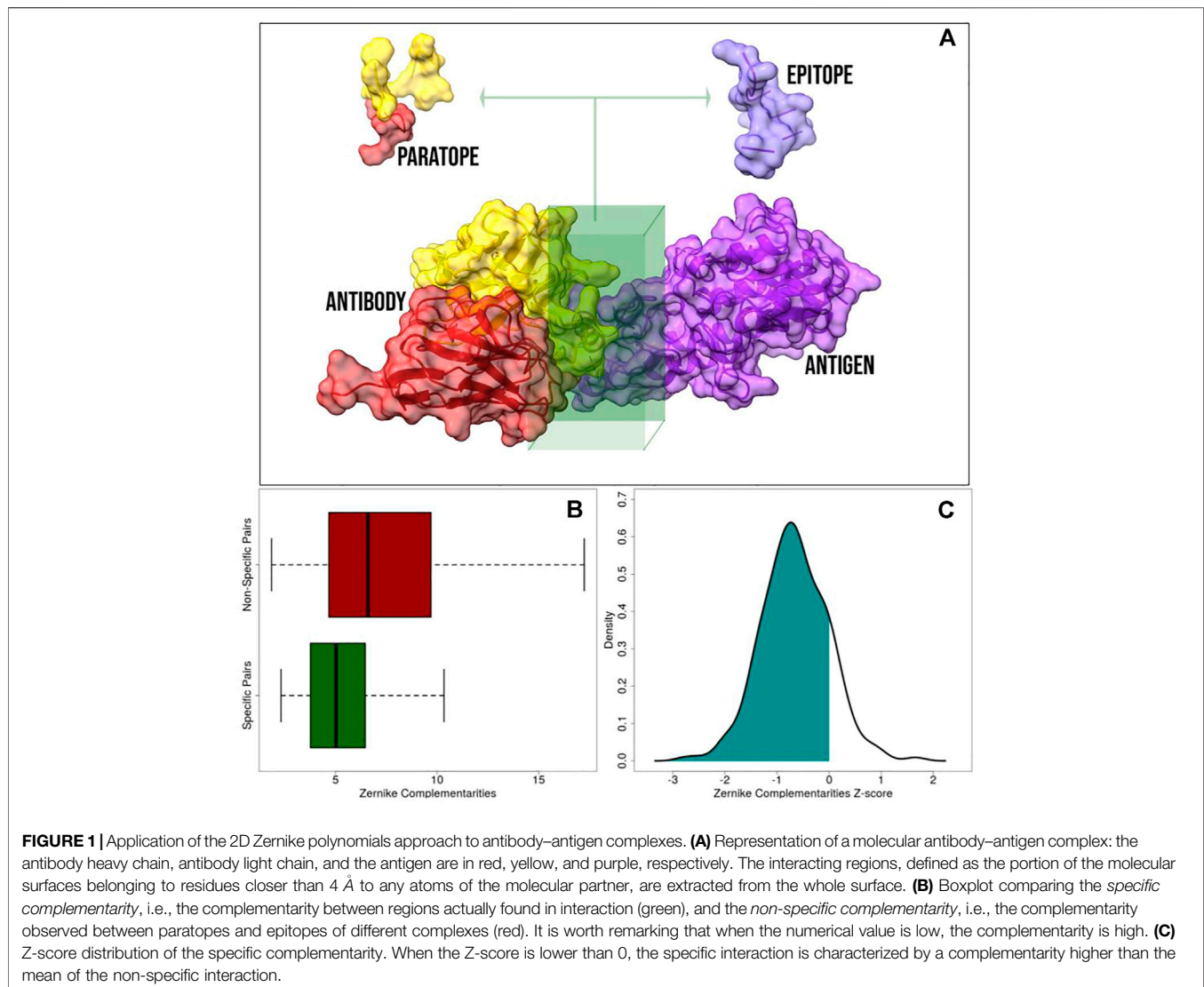
virtually any foreign antigen is due to high sequence variability in the antigen-binding site, while the overall architecture is largely conserved (Chothia and Lesk, 1987; Chothia et al., 1989; and Tramontano et al., 1990). The antigen-binding site is structurally composed of three loops of both the heavy and light chains, forming the *Complementary Determining Regions* (CDRs). Notwithstanding the variability of the CDR sequences, these loops (at least five out of six) can acquire only a limited number of structural conformations, called *canonical structures*. Moreover, studying the growing number of experimentally determined antibody structures, the relationship between the presence of given residues in certain sequence positions and the canonical structure adopted by the antibody has been demonstrated (Tramontano et al., 1990; Chothia et al., 1992; Foote and Winter, 1992; Decanniere et al., 2000; Chailyan et al., 2011; and North et al., 2011).

In this framework, thanks to the public availability of an increasing number of experimental antibody structures (Dunbar et al., 2013), several effective computational approaches—for predicting the structure of antibodies from their sequences—have been produced, often based on machine learning approaches (Dunbar et al., 2016; Lepore et al., 2017; Weitzner et al., 2017; and Abanades et al., 2022). Moreover, obtaining structural information about antibody–antigen complexes has been the object of extensive studies and it is still elusive. Many computational protocols focus on the prediction of the residues involved in partner interaction, both on the antibody and antigen side of the interface (Olimpieri et al., 2013; Potocnakova et al., 2016; Liberis et al., 2018).

All these kinds of computational tools can be used for antibody design, that is, the development of a novel molecule able to bind a given antigen (Norman et al., 2020). In particular, *ab initio* protocols are able to design paratopes integrating antibody structure prediction, molecular docking, and binding energy assessment (Pantazes and Maranas, 2010; Li et al., 2014; Lapidoth et al., 2015; and Adolf-Bryfogle et al., 2018).

Here, we collected a structural dataset of antibody–protein complexes solved in x-ray crystallography. In this work, we apply, for the first time, our recently developed method based on 2D Zernike descriptors to study the antibody–antigen interfaces. Concerning this specific kind of interaction, we demonstrate that such a fast and compact description can recognize with satisfying success the specific interaction from non-specific ones. Indeed, paratopes show a shape complementarity statistically higher toward their corresponding epitopes than toward epitopes belonging to unrelated antigens.

Based on these results, we propose here, for the first time, a new computational protocol employing 2D Zernike descriptors that, for a given target protein region, optimizes the shape complementarity of an antibody toward that region. Indeed, once a target region belonging to a protein antigen is identified and characterized with its Zernike descriptors, we compared such a region with the paratope of the antibodies in our dataset. Selecting as the starting template the antibody that has the most complementary patch, we perform a Monte Carlo (MC) simulation for the optimization of the paratope's structural conformation. Through extensive computational mutagenesis,



substituting in each step an interacting antibody residue with a different random one, we accept or reject each mutation according to the gain in shape complementarity, as evaluated by the Zernike method (Di Rienzo et al., 2020b; Di Rienzo et al., 2021b). In this work, the combined application of both the 2D Zernike formalism and a Monte Carlo simulation allows a computationally fast and effective exploration of the space of the possible mutants.

In the current pandemic situation, the interactions between SARS-CoV-2 spike protein and human cellular receptors have been extensively studied through the 2D Zernike polynomials formalism (Milanetti et al., 2021b; Bò et al., 2021; Miotto et al., 2021; and Miotto et al., 2022). Therefore, despite the generality of such an approach, we selected as a target for the optimization protocol some surface regions of the SARS-CoV-2 spike protein. We discuss here the results we got. Indeed, elucidating the interaction mechanism between antibodies and viral proteins represents a fundamental element for developing new therapies.

2 RESULTS AND DISCUSSION

2.1 Description of the Antibody–Antigen Interface Through Zernike Descriptors

In the present section, we discuss the results we obtained applying our recently developed computational protocol (Milanetti et al., 2021a) on a structural dataset composed of 229 antibody–antigen complexes (see Methods).

In particular, we firstly identified for each complex the paratope (epitope) as the set of residues with at least one atom closer than 4 Å to an antigen (antibody) atom. Therefore, after separately computing the molecular surface (Richards, 1977) for both the proteins in interaction, we extracted the portions of the molecular surfaces relating to the binding site residues to properly characterize the shape of the interacting regions of antibodies and antigens (see **Figure 1A**).

Once we identified the interacting regions, we characterized them through the 2D Zernike polynomials, summarizing their

geometrical properties in an ordered set of numerical descriptors (see Materials and Methods). By definition, two perfectly fitting surfaces are characterized by the same shape, meaning that, in principle, the difference between their Zernike descriptors is zero. Therefore, the shape complementarity between two molecular regions is compactly evaluated by such formalism. The lower is the distance between the Zernike descriptors and the higher is the shape complementarity between the corresponding protein regions (Milanetti et al., 2021a; Di Rienzo et al., 2021b).

We described all the paratopes and epitopes in the dataset with the Zernike formalism. In summary, we deal with 229 (the number of structures in our dataset) sets of 121 (the number of invariant descriptors when the order of expansion is set to 20) numerical descriptors for the paratopes and 229 for the epitopes. We thus defined the *specific complementarities* as the Euclidean distance between the descriptors of the 229 pairs of interacting (extracted from the same structure) paratope and epitope. Diversely, *Non-Specific complementarity* is the Euclidean distance between all the pairs of unrelated paratopes and epitopes (i.e. a paratope extracted from one structure and an epitope extracted from another one). In the end, we deal with 229 specific complementarities (one for each complex) and a high number (25,000) of non-specific complementarities (all the possible paratope–epitope pairs given a dataset of 229 items). In other words, in dealing with N antibody–antigen complexes, the specific complementarity, C_s , is defined as

$$D(p_i, e_j)_{i=j} = \sqrt{\sum_{k=1}^{121} (p_i^k - e_j^k)^2}_{i=j}, \quad (1)$$

where D is the Euclidean distance between the vectors of the paratope (p_i) and epitope (e_j) Zernike descriptors. Since we expanded all the paratopes and epitopes to the order 20, we dealt with 121 descriptors for each binding region. On the other hand, the non-specific complementarities, C_{ns} , can be computed as

$$D(p_i, e_j)_{i \neq j} = \sqrt{\sum_{k=1}^{121} (p_i^k - e_j^k)^2}_{i \neq j}. \quad (2)$$

In **Figure 1B**, we reported a boxplot highlighting the differences between C_s and C_{ns} distributions. As expected, the distribution of C_s is statistically lower than the distribution of C_{ns} (Kolmogorov–Smirnov test p value $< 2.2e - 16$), testifying the sensitivity of Zernike in recognizing regions actually in contact from the non-interacting ones. In the second step, we normalized the Zernike complementarities with the Z-score. In particular, for each paratope we have 1 C_s and 228 C_{ns} complementarity values, each of which is related to a specific patch pair. Normalizing over this set of numbers and looking at the Z-score regarding the specific one, we assessed the propensity that characterizes each antibody toward its specific antigen. In **Figure 1C**, we report the distribution of such specific Z-scores. As evident, they are mostly negatives (86% of the cases have a Z-score lower than 0, and 28% have a Z-score lower than -1), providing evidence that specific

interactions are characterized by lower distances (higher complementarity) than non-specific ones.

Taken together, these results confirm the ability of our method to correctly capture the main determinant of molecular shape complementarity.

2.2 Zernike-Based Monte Carlo Simulation for Molecular Interface Optimization

The Zernike formalism enjoys several advantageous features in representing the molecular surface: mainly, the invariance under rotation that makes such descriptors an “absolute” characterization of local protein morphology and the low computational cost of its calculation. Indeed, in this section, we present our algorithm that exploiting these advantages aims to optimize the shape complementarity of an antibody toward a given molecular target region. A similar procedure has already been presented and tested in our previous work (Di Rienzo et al., 2020b), and here for the first time, it is applied to antibody–antigen interaction systems.

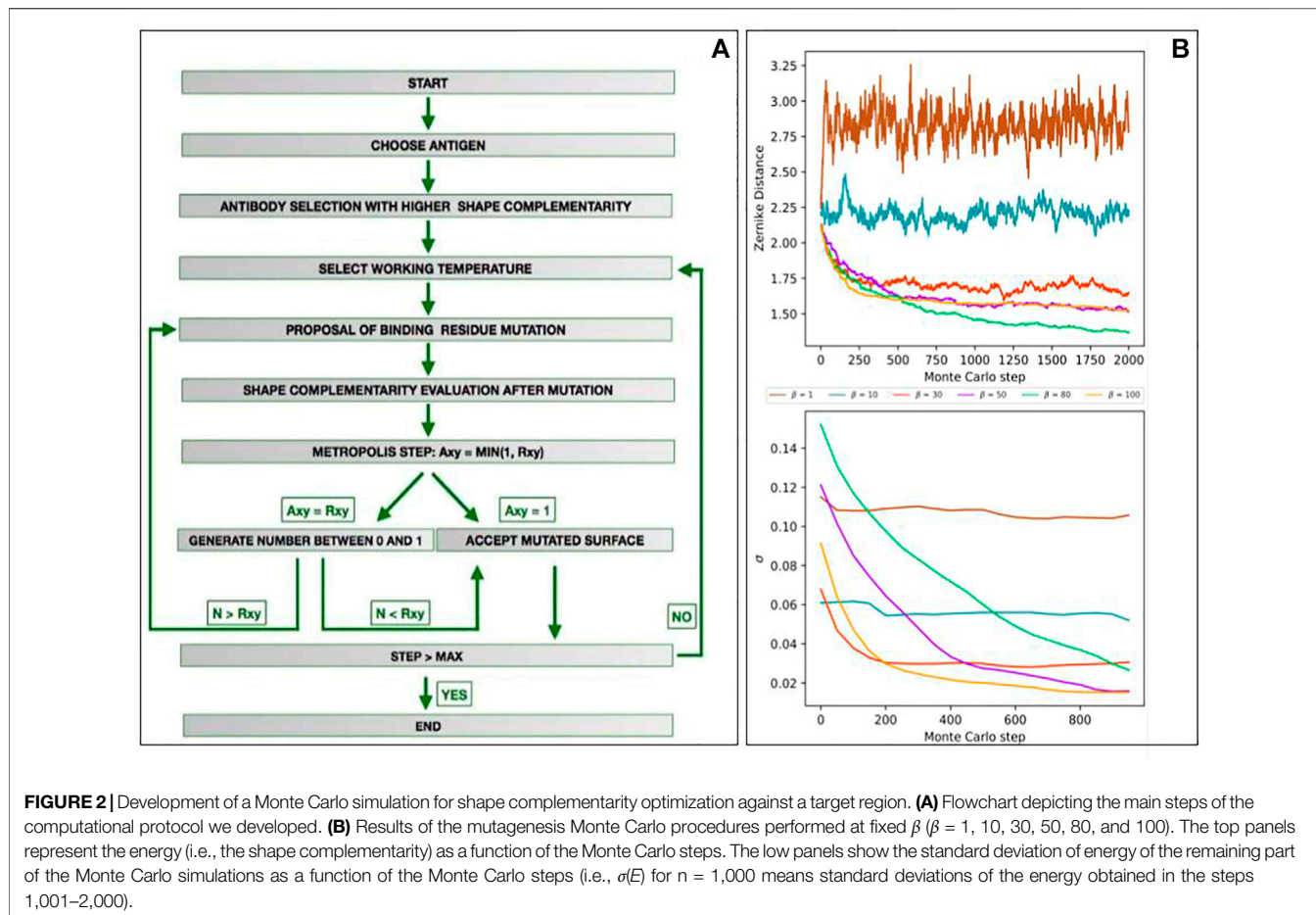
Figure 2A illustrates the main steps of the algorithm. We defined the target region as the portion of the antigen surface toward which an antibody will be optimized. It is thus necessary to identify the antibody chosen as a starting point of the algorithm. Summarized with the Zernike descriptors of the target region we compute the shape complementarity with all the paratopes of our dataset; here, the template, i.e., the antibody selected as the starting point for the mutagenesis process, can be chosen among the paratopes characterized by a high initial complementarity.

After establishing the template, we perform a Monte Carlo simulation employing computational mutagenesis on the paratope residues. In each step, we randomly selected a residue mutating it into another one of the 19 possible ones. The mutation generates a different paratope, characterized by a different shape of its molecular surface. Consequently, re-computing the Zernike descriptors, we can evaluate the effect of the mutation on the complementarity with the target region; indeed we can define the *complementarity balance* as

$$\Delta C = C_{mut} - C_{wt} = D(p_{mut}, e_{tar}) - D(p_{orig}, e_{tar}), \quad (3)$$

where p_{orig} and p_{mut} are the Zernike descriptors of the original and the mutated paratopes, respectively, while the e_{tar} represent the Zernike descriptors of the target epitope and D represent the distance between 2 sets of descriptors. Since, as said, a high complementarity is reached when D is low, $\Delta C < 0$ means a higher complementarity surface, and $\Delta C > 0$ is obtained when the mutation is deleterious since it causes a worsening of the shape complementarity.

The number of combinations of possible mutations in an interacting region, composed usually of tens of residues, is huge. Therefore, to effectively sample the space of the possible mutants, we perform a Monte Carlo Metropolis simulation, iterating the procedure described previously, where the mutation in each step is accepted according to the following rules:



$$P = \begin{cases} 1 & \text{if } \Delta C < 0 \\ e^{-\beta\Delta C} & \text{if } \Delta C \geq 0 \end{cases}, \quad (4)$$

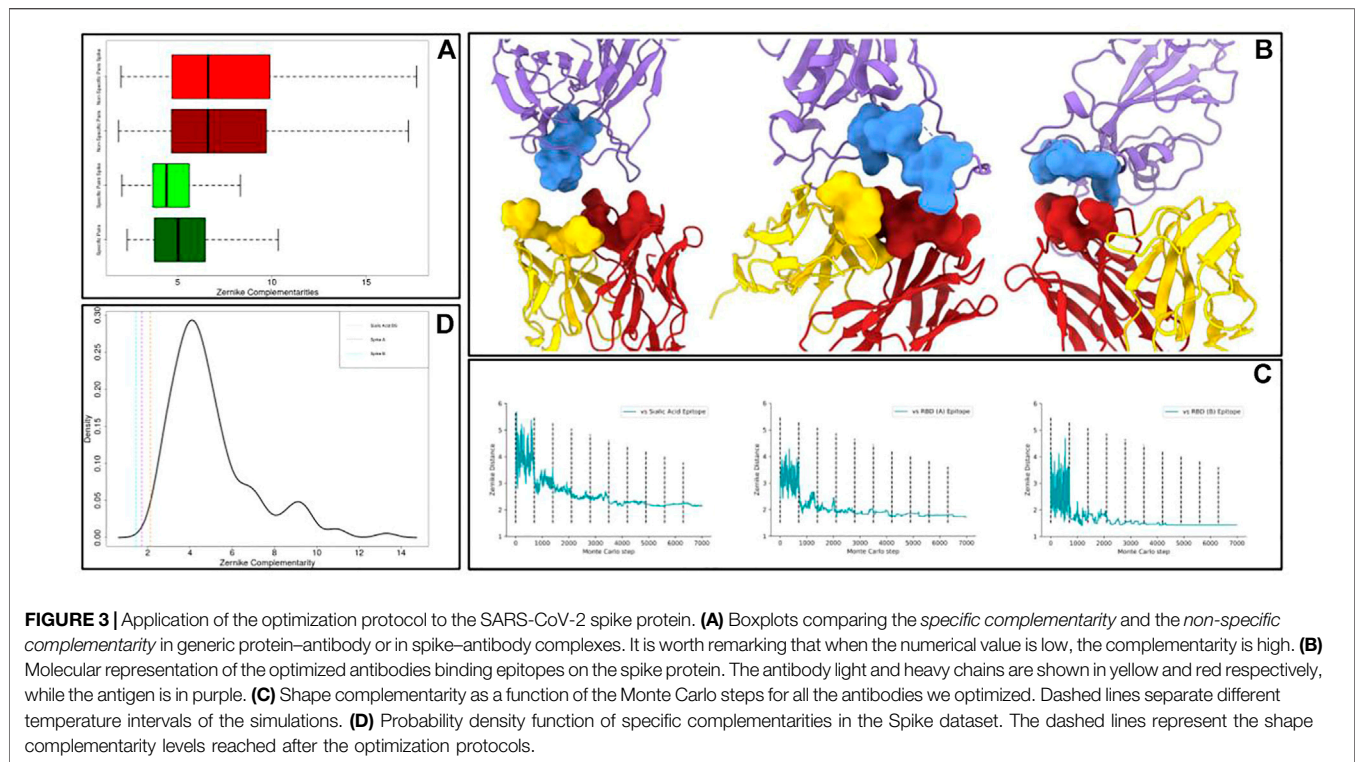
where $\beta \sim \frac{1}{T}$ is the temperature factor that determines the probability of acceptance of a step worsening complementarity. Note that this aspect is crucial to properly exploring a large number of different mutants: β is thus progressively increased during the Monte Carlo simulation, progressively confining the system to minimum energy in a simulated annealing process (Kirkpatrick et al., 1983).

To observe how many steps are necessary to reach the equilibrium for each β , we preliminary ran several fixed-temperature Monte Carlo simulations. We selected the epitope of an antigen structure in our dataset (PDB id: 1AR1) and, excluding its one, we choose as the starting template, the most complementary paratope in the structural dataset. We thus performed six different Monte Carlo simulations, each for a different β , where the acceptance probability in each mutagenesis step is given by Eq. 4. Performing 10 independent simulations of $M = 2,000$ steps for each β , the averaged results we obtained are summarized in Figure 2B. In the top panel, we reported the energy (i.e., the complementarity) as a function of the Monte Carlo steps. In the low panel, we reported the standard deviation of the energy of the remaining

part of the Monte Carlo simulations as a function of the number of steps (i.e., $\sigma(E)$ for $m = 1,000$ means standard deviations of the energy obtained in steps 1,001–2,000). As expected, for low values of β (i.e., high temperature) the system lives in a condition of indifferent equilibrium, where whatever mutation has an equal likelihood to be accepted, independently from its effect on complementarity. When, on the contrary, β is high (low temperature), the energy of the system rapidly decreases to a stationary local minimum. This trend is confirmed by looking at the stationary value of energies or, equivalently, noting that standard deviation tends progressively to zero. In the light of these results, in our protocol, we set $N = 700$ for each temperature. In this way, we preliminary allow the system to move away from the starting local conformation, thus freezing it in a new energy minimum, characterized by an increased shape complementarity with the target region.

2.3 A Case of Study: Application to the SARS-CoV-2 Spike Protein

The approach described here is general and can be applied to whatever protein. This notwithstanding that we applied it to the SARS-CoV-2 spike protein is a very relevant case of macromolecular interaction. Indeed, the severe acute



respiratory syndrome Corona virus 2 infection is still a very serious danger for public health (Huang et al., 2020; Zhu et al., 2020).

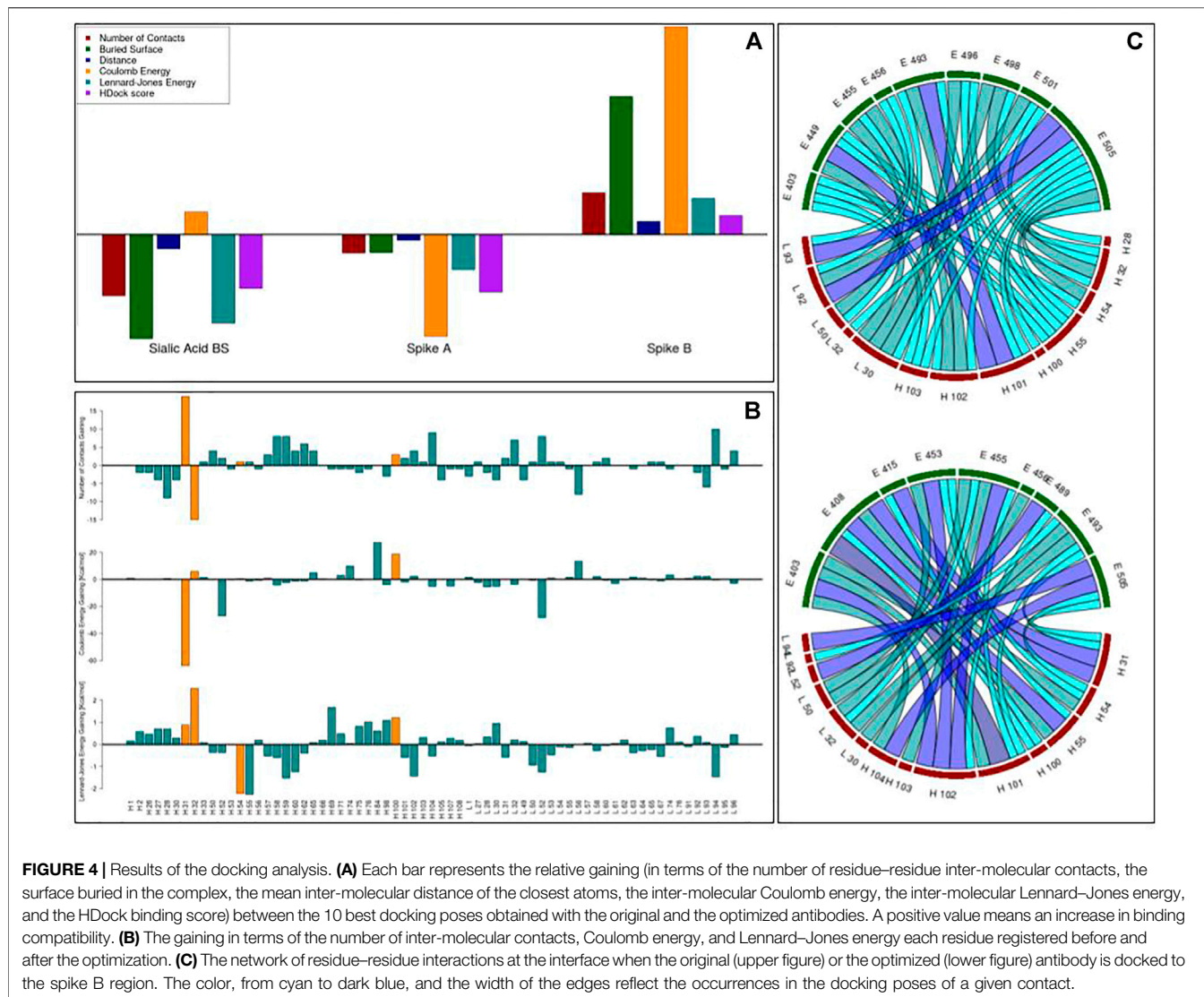
Many therapeutic strategies are devoted to the SARS-CoV-2 spike protein, protruding from the viral envelope and responsible for the cell entry mechanism (Walls et al., 2020; Wan et al., 2020; and Zhou et al., 2020). Thus, we obtained, using the dedicated section of Cov-AbDab (Raybould et al., 2021), a structural dataset of 145 spike–antibody complexes (we will call it the “Spike dataset”). We thus characterized the paratopes (antibody binding residues region) and the epitopes (various regions on the Spike) with the Zernike formalism. This allows us to compute, also for the Spike dataset, the *specific complementarities* and the *non-specific complementarities*, defined in Eqs 1, 2. The result of this analysis is shown in the following Figure 3A. These complementarities are reported as boxplots (light green for specific complementarities, light red for non-specific ones), even showing the boxplots regarding the general dataset (dark green and dark red, already shown in Figure 1B). Its results are evident that, also in the Spike dataset, specific interactions are characterized by a complementarity much higher than non-specific interactions (k.s. test p -value $< 2.2 \times 10^{-16}$). As expected, the non-specific interactions are represented by very similar distributions, since in both datasets they would represent an ensemble of non-interactions. Looking at these results, we noticed that the population of spike–antibody complexes showed the same behavior as the general population of the protein–antibody complexes.

In this framework, we selected on spike molecular surfaces, three different regions as targets for the optimization protocol.

On one hand, we targeted two different molecular regions involved in the interaction between spike and angiotensin-converting enzyme 2 (ACE-2), the well-known cellular receptor responsible for viral cell invasion. Moreover, we also optimized an antibody toward a very exposed region in the N-terminal spike domain, responsible for contacting sialic acid molecules. Indeed, such an interaction can confer to the virus, as occurred for the Middle East respiratory syndrome coronavirus (MERS-CoV) (Li et al., 2017), an additional molecular mechanism for cell intrusion. The responsible spike region represents a promising therapeutic target (Baker et al., 2020; Milanetti et al., 2021b).

We selected the residues constituting such epitopes and we characterized their molecular surfaces through Zernike formalism. Thus, we calculated the complementarity between these regions and all the antibody binding sites in our original dataset. To begin the optimization from a favorite starting point, we selected as templates antibody-binding sites characterized by the highest complementarity with each identified target.

We applied the procedure described in the previous section obtaining optimized paratopes whose molecular images are shown in Figure 3B, where we reported both the optimized antibodies and antigen interacting surfaces. In Figure 3C, we reported, for each of the Monte Carlo simulations performed, the shape complementarity as a function of the steps of the simulation, where the dashed lines enclose ranges with different β values. Each simulation significantly optimizes shape complementarity, obtaining a Zernike distance decrease of 43% on average. Significantly, all the designed binding sites are characterized by a very high final shape complementarity, in



terms of Zernike descriptors. Indeed, it is worth noting that the values obtained by all the three designed binding sites are lower than all the specific complementarities obtained in our structural dataset.

Moreover, in **Figure 3D**, we reported the probability density function of specific interactions in the Spike dataset. The vertical dashed lines represent the final complementarity values we get after the optimization procedures. It has to be noted that our protocol can effectively optimize the shape complementarity, obtaining final shape complementarity similar to the best cases observed in the Spike dataset.

The computational protocol we developed does not take into account several properties, known to be important in molecular recognition, such as electrostatics or hydrophobicity. In particular, our working hypothesis is that the shape complementarity plays a primary role as a perfect match between molecular surfaces due to an optimal structural rearrangement, which is probably caused by the compatibility of amino acid compositions of the interacting patches. However,

the relationship between shape complementarity and chemical–physical properties is not always trivial, requiring a further test for the patches proposed as interacting, to also analyze the compatibility of a chemical nature. This means that a residue substitution can, in principle, worsen the chemical compatibility between molecules, even if the shape complementarity is enhanced. For this reason, as a further step of our optimization protocol, we performed a molecular docking analysis using HDock (Yan et al., 2020). More specifically, we docked the spike protein and the antibodies, both in the original and in the optimized versions, to study the effects our computational protocol has produced. We constrained docking to interact with the residues composing the spike target epitopes and the antibodies' optimized regions. We summarized in **Figure 4** the results we obtained.

Thus, we selected the 10 best docking poses regarding both the original and the optimized antibodies. To assess whether the optimization protocol has been effective, some estimators of binding compatibility have been calculated. In particular, the

number of residue–residue inter-molecular contacts, the surface buried in the complex, the mean distance of the closest atoms between the two interfaces, the Coulomb inter-molecular energy, the Lennard–Jones inter-molecular energy, and the HDock binding score. For each of these observables, we computed the relative percentage of gaining after optimization so that the positive values indicate an increased binding tightness (See Materials and Methods). As shown in **Figure 4A**, even if in two applications we note a worsening, in one case, the optimization procedure has produced an antibody with better values of all such estimators, indicating the importance of including the molecular docking approach as a filter of selected patches based on geometric compatibility.

We focused therefore on this case and we analyzed how the introduced residue substitutions were responsible for this better compatibility. Analyzing the best docking poses, in **Figure 4B**, we reported the gaining in terms of the number of inter-molecular contacts and inter-molecular energy each residue registered before and after the optimization. Each residue is represented by a blue bar, while the residues mutated in the protocol are depicted in orange. As evident from the main effect regarding the residues “H 31” and “H 32” indeed, to increase the shape complementarity, the optimization protocol preferred to switch the exposition of these residues. It can be noted that the “H 31” residue, characterized by a high increase in the number of contacts, gains a very high amount of favorable (negative) Coulomb energy. Moreover, even if the number of inter-molecular contacts gained by H 54 is negligible, such a residue (and its neighborhood) acquired in the docking poses an increment of favorable Lennard–Jones energies.

Lastly, we assessed the difference in residue–residue inter-molecular interaction networks. In **Figure 4C** we reported the contacts between the main couples of residues, where a higher number of occurrences in the docking poses is testified by the thickness and the color of the edge. The spike residues are shown in green, and the antibody ones are in red. As further proof of the goodness of the proposed mutants, it can be noted that the interface of the optimized antibody (lower figure) is much more interconnected than the one of the original antibody (upper figure), indicating a possible effect on binding stability.

3 CONCLUSION

The binding affinity between biomolecules depends on a complex balance of several effects, including enthalpic and entropic contributions. Indeed, the substitution of even one residue at the interface could produce dramatic changes. Although many efforts were spent in this direction, predicting such effects has proven to be a difficult task and is still an open problem in computational biology.

In this scenario, the evaluation of shape complementarity between molecular regions is undoubtedly a central aspect. In this work, we focused on antibody–antigen interaction, a relevant case of molecular recognition. We applied our recently developed formalism based on the 2D Zernike polynomials to evaluate the shape complementarity with a quantitative approach. Once

summarized the topological properties of interacting regions with a set of numerical descriptors, we demonstrated that such formalism assigns to pairs of interacting region complementarities statistically higher than the ones assigned to regions not in interaction.

We thus developed a Monte Carlo-based approach for the shape optimization of an antibody towards a molecular target region. We proposed a new strategy that, potentially, could modify an antibody in order to acquire a very high shape complementarity for a given epitope of any antigen–protein.

Because of the emergence of viral variants that can eventually escape antibodies matured in vaccinated or recovered patients, the interactions between antibodies and SARS-CoV-2 spike protein are being extensively studied and still need further investigation.

For this reason, we selected three molecular regions on the spike protein as the target epitopes for our procedure. We, therefore, devised a set of antibodies characterized by a high shape complementarity toward their cognate epitopes.

However, even without considering therapeutically important elements such as immunogenicity and solubility, some other aspects have to be properly considered in our algorithm to increase the probability of identifying actually binding antibodies. Firstly, to produce more reliable mutant structures, the residue substitutions’ procedure has to account for the hypervariable loops canonical structure modeling. Moreover, we worked on antibody-bound structures: a structural conformational exploration can allow the antibodies to energetically rearrange their side chains, and to find the proper conformation able to bind the studied antigen. Finally, the binding compatibility does not depend only on shape complementarity, and thus the inclusion of terms accounting for the residues’ chemical characteristics will surely improve the method’s performance.

In conclusion, this procedure can represent a promising strategy for interface region molecular optimization, where the inclusion of the aspects discussed previously represents the necessary improvement steps. In the present work, we highlighted, with an independent molecular docking evaluation, the case when the optimization procedure has increased molecular complex compactness.

4 MATERIALS AND METHODS

4.1 Dataset

We selected 229 protein-binding antibodies with a sequence identity lower than 90% and resolution < 3.0 Å using the SabDab database (Dunbar et al., 2013). The Spike dataset, i.e., the structural dataset of the spike–antibody complexes was built using CoV-AbDab (Raybould et al., 2021). It results in 145 complex structures with a sequence identity lower than 90%, as calculated using CD-HIT (Huang et al., 2010).

The sequence of each antibody was renumbered according to the Chothia numbering scheme (Chothia and Lesk, 1987; Chothia et al., 1989) using an in-house python script.

TABLE 1 | The residue substitutions performed during the shape optimization procedure. The template structures for the sialic acid binding site, spike A, and spike B were 3bdy, 1yjd, and 1kb5, respectively. We adopted the Chotia numbering scheme.

Epitope	Chain ID	Res no.	Insert	Original res	Inserted res
Sialic acid BS	H	33	S	TYR	TRP
Sialic acid BS	H	52		TYR	TRP
Sialic acid BS	H	95		TRP	GLN
Sialic acid BS	H	100		PHE	ARG
Sialic acid BS	L	30	B	SER	ALA
Sialic acid BS	L	30	C	ILE	PHE
Sialic acid BS	L	30	D	SER	ILE
Sialic acid BS	L	31		TYR	ALA
Sialic acid BS	L	50		TRP	GLY
Sialic acid BS	L	91		HIS	SER
Sialic acid BS	L	92		TYR	GLY
Sialic acid BS	L	93		THR	HIS
Sialic acid BS	L	94		THR	VAL
Spike A	H	31		SER	LYS
Spike A	H	96		HIS	MET
Spike A	H	97		TYR	PHE
Spike A	H	98		GLY	LEU
Spike A	H	99		LEU	PRO
Spike A	H	100	S	ASP	HIS
Spike A	H	100	T	TRP	LYS
Spike A	L	30		TYR	ILE
Spike A	L	32		TRP	VAL
Spike A	L	91		GLY	ARG
Spike A	L	92		GLN	ARG
Spike B	H	31		GLY	GLU
Spike B	H	32		TYR	CYS
Spike B	H	53		TYR	PHE

The structure of the SARS-CoV-2 spike protein used for the identification of the ACE-2 interacting region has the PDB code 6vw1. When we investigated the N-terminal domain, we used the structure 7jji.

We identified on the spike protein two epitopes in the ACE-2 binding region: spike A and spike B. spike A epitope is constituted by the residues: “TYR 453, LEU 455, PHE 456, ALA 475, GLY 476, PHE 486, ASN 487, TYR 489, and GLN 493.” Spike B epitope is constituted by the residues: “TYR 449, GLY 496, GLN 498, THR 500, ASN 501, GLY 502, and TYR 505.” We identified another epitope in the N-terminal domain, in the region involved in sialic acid binding. That region is defined as the set of residues whose CA atoms are closer than 8Å to the TRP 258 CA (Milanetti et al., 2021b). Such an epitope is constituted by these residues: “LEU 244, HIS 245, ARG 246, SER 247, TYR 248, LEU 249, THR 250, PRO 251, GLY 252, ASP 253, SER 254, SER 255, SER 256, GLY 257, TRP 258, THR 259, and ALA 260”.

4.2 Surface Construction

Using as reference the experimental structures, computational mutagenesis has been performed using SCWRL4 (Krivov et al., 2009).

For each protein structure, solvent accessible surface is computed using DMS software with the standard option (Richards, 1977). The interacting surface is constituted by the surface points belonging to interacting residues, defined as the set

of residues having at least one atom closer than 4 Å to any atoms of the molecular partner.

4.3 Zernike Descriptors

Given a 2D function $f(r, \phi)$ in the unitary circle (region $r < 1$), it can be expanded in the Zernike polynomials basis. Therefore,

$$f(r, \phi) = \sum_{n=0}^{\infty} \sum_{m=0}^{m=n} c_{nm} Z_{nm}, \quad (5)$$

where

$$c_{nm} = \frac{(n+1)}{\pi} \langle Z_{nm} | f \rangle = \frac{(n+1)}{\pi} \int_0^1 dr r \int_0^{2\pi} d\phi Z_{nm}^*(r, \phi) f(r, \phi) \quad (6)$$

are the expansion coefficients (Zernike moments). The complex functions $Z_{nm}(r, \phi)$ are the Zernike polynomials, each composed of a radial and an angular part:

$$Z_{nm} = R_{nm}(r) e^{im\phi}. \quad (7)$$

The radial dependence, given n and m , can be written as follows:

$$R_{nm}(r) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} r^{n-2k}. \quad (8)$$

For each couple of polynomials, the following rule holds:

$$\langle Z_{nm} | Z_{n'm'} \rangle = \frac{\pi}{(n+1)} \delta_{nm'} \delta_{mm'}. \quad (9)$$

Therefore, the set of polynomials forms a basis. Knowing the coefficients, $\{c_{nm}\}$ allows the reconstruction of the original function. The level of the detail can be modified by modulating the order of expansion, $N = \max(n)$.

The norm of each coefficient ($z_{nm} = |c_{nm}|$) does not depend on the phase; therefore, it is invariant under rotations around the origin.

The shape complementarity between two regions can be evaluated by comparing their Zernike invariants. In particular, we measured the complementarity between regions i and j as the Euclidean distance between the invariant vectors, i.e.,

$$d_{ij} = \sqrt{\sum_{k=1}^{M=121} (z_i^k - z_j^k)^2}. \quad (10)$$

We adopted $N = 20$, therefore, dealing with 121 invariant descriptors for each patch.

4.4 Analysis of Docking Poses

In Table 1, we report the mutations proposed as a result of the three Monte Carlo simulations.

We docked the three original and the three optimized antibody structures with spike using HDOCK (Yan et al., 2020), indicating as interacting residues the one written in the Dataset section.

We selected, for all the 6 docking simulations, the best 10 poses according to the Hdock binding score, an iterative knowledge-based scoring function. For each pose we get the following:

- The number of inter-molecular residue–residue contacts. Two residues are in contact if the minimum distance between their atoms is less than 4 Å.
- The surface is buried in the complex. The surface buried is defined as the difference between the sum of the monomers' area and the complex area. For this calculation, we use DMS software (Richards, 1977).
- The mean of the lowest 100 atom–atom inter-molecular distances.
- The sum of the Coulomb energy of the interactions occurring between antibody and spike atoms. We used the CHARMM27 force field (MacKerell et al., 2002).
- The sum of the Lennard–Jones energy of the interactions occurring between antibody and spike atoms. We used the CHARMM27 force field (MacKerell et al., 2002).
- The pose Hdock binding score.

The comparisons between the results regarding original and optimized antibodies are performed so as a positive value means an increase in binding compatibility after optimization. Therefore, the relative percentage of gaining is defined as follows:

- Number of contacts: $\frac{\langle \text{Cont} \rangle_{\text{opt}} - \langle \text{Cont} \rangle_{\text{orig}}}{\langle \text{Cont} \rangle_{\text{orig}}}$
- Buried area: $\frac{\langle \text{Surf} \rangle_{\text{opt}} - \langle \text{Surf} \rangle_{\text{orig}}}{\langle \text{Surf} \rangle_{\text{orig}}}$
- Distance: $\frac{\langle \text{Dist} \rangle_{\text{orig}} - \langle \text{Dist} \rangle_{\text{opt}}}{\langle \text{dist} \rangle_{\text{orig}}}$

REFERENCES

- Abanades, B., Georges, G., Bujotzek, A., and Deane, C. M. (2022). ABlooper: Fast Accurate Antibody CDR Loop Structure Prediction with Accuracy Estimation. *Bioinformatics*. 38, 1877. doi:10.1093/bioinformatics/btac016
- Adolf-Bryfogle, J., Kalyuzhnyi, O., Kubitz, M., Weitzner, B. D., Hu, X., Adachi, Y., et al. (2018). RosettaAntibodyDesign (RABD): A General Framework for Computational Antibody Design. *Plos Comput. Biol.* 14, e1006112. doi:10.1371/journal.pcbi.1006112
- Alba, J., Di Rienzo, L., Milanetti, E., Acuto, O., and D'Abramo, M. (2020). Molecular Dynamics Simulations Reveal Canonical Conformations in Different pMHC/TCR Interactions. *Cells*. 9, 942. doi:10.3390/cells9040942
- Baker, A. N., Richards, S.-J., Guy, C. S., Congdon, T. R., Hasan, M., Zwetsloot, A. J., et al. (2020). The SARS-CoV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device. *ACS Cent. Sci.* 6, 2046–2052. doi:10.1021/acscentsci.0c00855
- Bò, L., Miotto, M., Di Rienzo, L., Milanetti, E., and Ruocco, G. (2021). Exploring the Association Between Sialic Acid and SARS-CoV-2 Spike Protein Through a Molecular Dynamics-Based Approach. *Frontiers in Medical Technology* 2, 24.
- Canterakis, N. (1999). "Scandinavian Conference on Image Analysis," in Proceedings of the 11th Scandinavian Conference on Image Analysis, Kangerlussuaq (Greenland: Pattern Recognition Society of Denmark).

- Coulomb energy: $\frac{\langle E_c \rangle_{\text{orig}} - \langle E_c \rangle_{\text{opt}}}{\langle E_c \rangle_{\text{orig}}}$
- Lennard–Jones energy: $\frac{\langle E_{lj} \rangle_{\text{orig}} - \langle E_{lj} \rangle_{\text{opt}}}{\langle E_{lj} \rangle_{\text{orig}}}$
- Hdock score: $\frac{\langle \text{Score} \rangle_{\text{orig}} - \langle \text{Score} \rangle_{\text{opt}}}{\langle \text{Score} \rangle_{\text{orig}}}$

where the subscripts “orig” and “opt” refer to the poses obtained with antibodies before and after the optimization, respectively.

DATA AVAILABILITY STATEMENT

All the protein structures are available on Protein Data Bank. The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AD: methodology, software, and analysis. LD: conceptualization, analysis, methodology, software, and writing. MM: conceptualization, writing, analysis, and methodology. PO: conceptualization, revision, and dataset. EM: conceptualization, analysis, methodology, and writing. GR: conceptualization, methodology, supervision, and revision.

FUNDING

The research leading to these results was also supported by European Research Council Synergy grant ASTRA (no. 855923).

- Chailyan, A., Marcatili, P., Cirillo, D., and Tramontano, A. (2011). Structural Repertoire of Immunoglobulin λ Light Chains. *Proteins*. 79, 1513–1524. doi:10.1002/prot.22979
- Chen, R., and Weng, Z. (2003). A Novel Shape Complementarity Scoring Function for Protein–Protein Docking. *Proteins*. 51, 397–408. doi:10.1002/prot.10334
- Chothia, C., and Lesk, A. M. (1987). Canonical Structures for the Hypervariable Regions of Immunoglobulins. *J. Mol. Biol.* 196, 901–917. doi:10.1016/0022-2836(87)90412-8
- Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., et al. (1992). Structural Repertoire of the Human VH Segments. *J. Mol. Biol.* 227, 799–817. doi:10.1016/0022-2836(92)90224-8
- Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., et al. (1989). Conformations of Immunoglobulin Hypervariable Regions. *Nature*. 342, 877–883. doi:10.1038/342877a0
- Daberdaku, S., and Ferrari, C. (2019). Antibody Interface Prediction with 3D Zernike Descriptors and SVM. *Bioinformatics* 35, 1870–1876. doi:10.1093/bioinformatics/bty918
- Daberdaku, S., and Ferrari, C. (2018). Exploring the Potential of 3D Zernike Descriptors and SVM for Protein–Protein Interface Prediction. *BMC bioinformatics*. 19, 35. doi:10.1186/s12859-018-2043-3
- Decanniere, K., Muyldermans, S., and Wyns, L. (2000). Canonical Antigen-Binding Loop Structures in Immunoglobulins: More Structures, More Canonical Classes? *J. Mol. Biol.* 300, 83–91. doi:10.1006/jmbi.2000.3839

- Di Rienzo, L., Milanetti, E., Alba, J., and D'Abramo, M. (2020a). Quantitative Characterization of Binding Pockets and Binding Complementarity by Means of Zernike Descriptors. *J. Chem. Inf. Model.* 60, 1390–1398. doi:10.1021/acs.jcim.9b01066
- Di Rienzo, L., Milanetti, E., Testi, C., Montemiglio, L. C., Baiocco, P., Boffi, A., et al. (2020b). A Novel Strategy for Molecular Interfaces Optimization: The Case of Ferritin-Transferrin Receptor Interaction. *Comput. Struct. Biotechnol. J.* 18, 2678–2686. doi:10.1016/j.csbj.2020.09.020
- Di Rienzo, L., Milanetti, E., Lepore, R., Olimpieri, P. P., and Tramontano, A. (2017). Superposition-Free Comparison and Clustering of Antibody Binding Sites: Implications for the Prediction of the Nature of Their Antigen. *Scientific Rep.* 7, 1. doi:10.1038/srep45053
- Di Rienzo, L., Milanetti, E., Ruocco, G., and Lepore, R. (2021a). *Frontiers in Molecular Biosciences*. United Kingdom: King's College London London, 933.
- Di Rienzo, L., Monti, M., Milanetti, E., Miotto, M., Boffi, A., Tartaglia, G. G., et al. (2021b). Computational Optimization of Angiotensin-Converting Enzyme 2 for Sars-Cov-2 Spike Molecular Recognition *Comput. Struct. Biotechnol. J.* 19, 3006. doi:10.1016/j.csbj.2021.05.016
- Dunbar, J., Krawczyk, K., Leem, J., Baker, T., Fuchs, A., Georges, G., et al. (2013). SAbDab: the Structural Antibody Database. *Nucl. Acids Res.* 42, D1140–D1146. doi:10.1093/nar/gkt1043
- Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., et al. (2016). SAbPred: a Structure-Based Antibody Prediction Server. *Nucleic Acids Res.* 44, W474–W478. doi:10.1093/nar/gkw361
- Erijman, A., Rosenthal, E., and Shifman, J. M. (2014). How Structure Defines Affinity in Protein-Protein Interactions. *PLOS one* 9, e110085. doi:10.1371/journal.pone.0110085
- Foote, J., and Winter, G. (1992). Antibody Framework Residues Affecting the Conformation of the Hypervariable Loops. *J. Mol. Biol.* 224, 487–499. doi:10.1016/0022-2836(92)91010-m
- Gainza, P., Sverrisson, F., Monti, F., Rodolà, E., Boscaini, D., Bronstein, M. M., et al. (2020). Deciphering Interaction Fingerprints from Protein Molecular Surfaces Using Geometric Deep Learning. *Nat. Methods* 17, 184–192. doi:10.1038/s41592-019-0666-6
- Gavin, A.-C., Bösch, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al. (2002). Functional Organization of the Yeast Proteome by Systematic Analysis of Protein Complexes. *Nature* 415, 141–147. doi:10.1038/415141a
- Geng, C., Jung, Y., Renaud, N., Honavar, V., Bonvin, A. M. J. J., and Xue, L. C. (2020). iScore: a Novel Graph Kernel-Based Function for Scoring Protein-Protein Docking Models. *Bioinformatics* 36, 112–121. doi:10.1093/bioinformatics/btz496
- Gotwals, P., Cameron, S., Cipolletta, D., Cremasco, V., Crystal, A., Hewes, B., et al. (2017). Prospects for Combining Targeted and Conventional Cancer Therapy with Immunotherapy. *Nat. Rev. Cancer* 17, 286–301. doi:10.1038/nrc.2017.17
- Gromiha, M. M., Yugandhar, K., and Jemimah, S. (2017). Protein-protein Interactions: Scoring Schemes and Binding Affinity. *Curr. Opin. Struct. Biol.* 44, 31–38. doi:10.1016/j.sbi.2016.10.016
- Han, J.-D. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., et al. (2004). Evidence for Dynamically Organized Modularity in the Yeast Protein-Protein Interaction Network. *Nature* 430, 88–93. doi:10.1038/nature02555
- Han, X., Sit, A., Christoffer, C., Chen, S., and Kihara, D. (2019). A Global Map of the Protein Shape Universe. *Plos Comput. Biol.* 15, e1006969. doi:10.1371/journal.pcbi.1006969
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *The Lancet* 395, 497–506. doi:10.1016/s0140-6736(20)30183-5
- Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a Web Server for Clustering and Comparing Biological Sequences. *Bioinformatics* 26, 680–682. doi:10.1093/bioinformatics/btq003
- Jones, S., and Thornton, J. M. (1996). Principles of Protein-Protein Interactions. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13–20. doi:10.1073/pnas.93.1.13
- Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A., Aflalo, C., and Vakser, I. A. (1992). Molecular Surface Recognition: Determination of Geometric Fit between Proteins and Their Ligands by Correlation Techniques. *Proc. Natl. Acad. Sci. U.S.A.* 89, 2195–2199. doi:10.1073/pnas.89.6.2195
- Kihara, D., Sael, L., Chikhi, R., and Esquivel-Rodriguez, J. (2011). Molecular Surface Representation Using 3D Zernike Descriptors for Protein Shape Comparison and Docking. *Curr. Protein. Pept. Sci.* 12, 520–530. doi:10.2174/138920311796957612
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*. 220, 671–680. doi:10.1126/science.220.4598.671
- Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., et al. (2017). The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* 12, 255–278. doi:10.1038/nprot.2016.169
- Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L., Jr (2009). Improved Prediction of Protein Side-Chain Conformations with SCWRL4. *Proteins*. 77, 778–795. doi:10.1002/prot.22488
- Kuroda, D., and Gray, J. J. (2016). Shape Complementarity and Hydrogen Bond Preferences in Protein-Protein Interfaces: Implications for Antibody Modeling and Protein-Protein Docking. *Bioinformatics*. 32, 2451–2456. doi:10.1093/bioinformatics/btw197
- Lapidath, G. D., Baran, D., Pszolla, G. M., Norn, C., Alon, A., Tyka, M. D., et al. (2015). AbDesign: An Algorithm for Combinatorial Backbone Design Guided by Natural Conformations and Sequences. *Proteins*. 83, 1385–1406. doi:10.1002/prot.24779
- Lawrence, M. C., and Colman, P. M. (1993). Shape Complementarity at Protein/protein Interfaces. *J. Mol. Biol.* 234, 946–950. doi:10.1006/jmbi.1993.1648
- Lepore, R., Olimpieri, P. P., Messih, M. A., and Tramontano, A. (2017). PIGSPRO: Prediction of immunoglobulin Structures V2. *Nucleic Acids Res.* 45, W17–W23. doi:10.1093/nar/gkx334
- Li, T., Pantazes, R. J., and Maranas, C. D. (2014). OptMAVEN - A New Framework for the De Novo Design of Antibody Variable Region Models Targeting Specific Antigen Epitopes. *PLoS one* 9, e105954. doi:10.1371/journal.pone.0105954
- Li, W., Hulswit, R. J., Widjaja, I., Raj, V. S., McBride, R., Peng, W., et al. (2017). Identification of Sialic Acid-Binding Function for the Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein. *Proc. Natl. Acad. Sci.* 114, E8508. doi:10.1073/pnas.1712592114
- Li, Y., Li, H., Yang, F., Smith-Gill, S. J., and Mariuzza, R. A. (2003). X-ray Snapshots of the Maturation of an Antibody Response to a Protein Antigen. *Nat. Struct. Mol. Biol.* 10, 482–488. doi:10.1038/nsb930
- Liberis, E., Veličković, P., Sormanni, P., Vendruscolo, M., and Liò, P. (2018). Parapred: Antibody Paratope Prediction Using Convolutional and Recurrent Neural Networks. *Bioinformatics*. 34, 2944–2950. doi:10.1093/bioinformatics/bty305
- MacKerell, A. D., Jr, Brooks, B., Brooks, C. L., III, Nilsson, L., Roux, B., Won, Y., et al. (2002). *CHARMM: The Energy Function and its Parameterization*, 1. Hoboken, NJ: Wiley Online Library.
- Milanetti, E., Miotto, M., Di Rienzo, L., Monti, M., Gosti, G., and Ruocco, G. (2021a). 2D Zernike Polynomial Expansion: Finding the Protein-Protein Binding Regions. *Comput. Struct. Biotechnol. J.* 19, 29–36. doi:10.1016/j.csbj.2020.11.051
- Milanetti, E., Miotto, M., Rienzo, L. D., Nagaraj, M., Monti, M., Golbek, T. W., et al. (2021b). In-Silico Evidence for a Two Receptor Based Strategy of SARS-CoV-2. *Front. Mol. Biosciences*. 8, 690655. doi:10.3389/fmolb.2021.690655
- Miotto, M., Di Rienzo, L., Gosti, G., Bo, L., Parisi, G., Piacentini, R., et al. (2022). Inferring the Stabilization Effects of SARS-CoV-2 Variants on the Binding with ACE2 Receptor. *Commun. Biol.* 5, 1. doi:10.1038/s42003-021-02946-w
- Miotto, M., Rienzo, L. D., Bò, L., Boffi, A., Ruocco, G., and Milanetti, E. (2021). Molecular Mechanisms Behind Anti SARS-CoV-2 Action of Lactoferrin. *Front. Mol. Biosciences* 8, 607443. doi:10.3389/fmolb.2021.607443
- Nicola, G., and Vakser, I. A. (2007). A Simple Shape Characteristic of Protein Protein Recognition. *Bioinformatics* 23, 789–792. doi:10.1093/bioinformatics/btm018
- Norman, R. A., Ambrosetti, F., Bonvin, A. M. J. J., Colwell, L. J., Kelm, S., Kumar, S., et al. (2020). Computational Approaches to Therapeutic Antibody Design: Established Methods and Emerging Trends. *Brief. Bioinformatics* 21, 1549–1567. doi:10.1093/bib/bbz095
- North, B., Lehmann, A., and Dunbrack, R. L., Jr (2011). A New Clustering of Antibody CDR Loop Conformations. *J. Mol. Biol.* 406, 228–256. doi:10.1016/j.jmb.2010.10.030
- Novotni, M., and Klein, R. (2004). Shape Retrieval Using 3D Zernike Descriptors. *Computer-Aided Des.* 36, 1047–1062. doi:10.1016/j.cad.2004.01.005
- Olimpieri, P. P., Chailyan, A., Tramontano, A., and Marcatili, P. (2013). Prediction of Site-specific Interactions in Antibody-Antigen Complexes: the proABC

- Method and Server. *Bioinformatics* 29, 2285–2291. doi:10.1093/bioinformatics/btt369
- Pantazes, R. J., and Maranas, C. D. (2010). OptCDR: a General Computational Method for the Design of Antibody Complementarity Determining Regions for Targeted Epitope Binding. *Protein Eng. Des. Selection* 23, 849–858. doi:10.1093/protein/gzq061
- Potocnakova, L., Bhide, M., and Pulzova, L. B. (2016). An Introduction to B-Cell Epitope Mapping and in Silico Epitope Prediction. *J. Immunol. Res.* 2016, 6760830. doi:10.1155/2016/6760830
- Raybould, M. I. J., Kovaltsuk, A., Marks, C., and Deane, C. M. (2021). CoV-AbDab: the Coronavirus Antibody Database. *Bioinformatics* 37, 734–735. doi:10.1093/bioinformatics/btaa739
- Richards, F. M. (1977). Areas, Volumes, Packing, and Protein Structure. *Annu. Rev. Biophys. Bioeng.* 6, 151–176. doi:10.1146/annurev.bb.06.060177.001055
- Saeed, A. F. U. H., Wang, R., Ling, S., and Wang, S. (2017). Antibody Engineering for Pursuing a Healthier Future. *Front. Microbiol.* 8, 495. doi:10.3389/fmicb.2017.00495
- Siebenmorgen, T., and Zacharias, M. (2020). Computational Prediction of Protein–Protein Binding Affinities. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 10, e1448. doi:10.1002/wcms.1448
- Singh, S., Kumar, N. K., Dwiwedi, P., Charan, J., Kaur, R., Sidhu, P., et al. (2018). Monoclonal Antibodies: A Review. *Ccp* 13, 85–99. doi:10.2174/1574884712666170809124728
- Tramontano, A., Chothia, C., and Lesk, A. M. (1990). Framework Residue 71 Is a Major Determinant of the Position and Conformation of the Second Hypervariable Region in the VH Domains of Immunoglobulins. *J. Mol. Biol.* 215, 175–182. doi:10.1016/s0022-2836(05)80102-0
- Vakser, I. A. (2014). Protein-Protein Docking: From Interaction to Interactome. *Biophysical J.* 107, 1785–1793. doi:10.1016/j.bpj.2014.08.033
- Vangone, A., and Bonvin, A. M. (2015). Contacts-based Prediction of Binding Affinity in Protein–Protein Complexes. *elife* 4, e07454. doi:10.7554/elife.07454
- Venkatraman, V., Sael, L., and Kihara, D. (2009). Potential for Protein Surface Shape Analysis Using Spherical Harmonics and 3D Zernike Descriptors. *Cell Biochem Biophys* 54, 23–32. doi:10.1007/s12013-009-9051-x
- Venkatraman, V., Yang, Y. D., Sael, L., and Kihara, D. (2009). Protein-protein Docking Using Region-Based 3D Zernike Descriptors. *BMC bioinformatics* 10, 407. doi:10.1186/1471-2105-10-407
- Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281. doi:10.1016/j.cell.2020.02.058
- Wan, Y., Shang, J., Graham, R., Baric, R. S., and Li, F. (2020). Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* 94, e00127-20. doi:10.1128/JVI.00127-20
- Weitzner, B. D., Jeliakzov, J. R., Lyskov, S., Marze, N., Kuroda, D., Frick, R., et al. (2017). Modeling and Docking of Antibody Structures with Rosetta. *Nat. Protoc.* 12, 401–416. doi:10.1038/nprot.2016.180
- Yan, Y., and Huang, S.-Y. (2019). Pushing the Accuracy Limit of Shape Complementarity for Protein-Protein Docking. *BMC bioinformatics*. 20, 1. doi:10.1186/s12859-019-3270-y
- Yan, Y., Tao, H., He, J., and Huang, S.-Y. (2020). The HDock Server for Integrated Protein-Protein Docking. *Nat. Protoc.* 15, 1829–1852. doi:10.1038/s41596-020-0312-x
- Zernike, F., and Stratton, F. J. M. (1934). Diffraction Theory of the Knife-Edge Test and its Improved Form, the Phase-Contrast Method. *Monthly Notices R. Astronomical Soc.* 94, 377–384. doi:10.1093/mnras/94.5.377
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A Pneumonia Outbreak Associated With a New Coronavirus of Probable Bat Origin. *Nature* 579, 270. doi:10.1038/s41586-020-2012-7
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New Engl. J. Med.* 382, 727. doi:10.1056/NEJMoa2001017

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 De Lauro, Di Rienzo, Miotto, Olimpieri, Milanetti and Ruocco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.