Heliyon 8 (2022) e11568

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Structural analysis of a simplified model reproducing SARS-CoV-2 S RBD/ACE2 binding site

Michela Buonocore [a,b,1], Angelo Santoro [a,c,1], Manuela Grimaldi [a], Verdiana Covelli [a,d], Mohammad Firoznezhad [a,e], Manuela Rodriquez [a], Matteo Santin [f], Anna Maria D'Ursi [a,*]

[a] University of Salerno, Department of Pharmacy, Via Giovanni Paolo II, 132-84084 Fisciano, Salerno, Italy
[b] University of Naples Federico II, Department of Veterinary Pathology, via Federico Delpino 1, 80137, Naples, Italy
[c] University of Salerno, Scuola di Specializzazione in Farmacia Ospedaliera, via Giovanni Paolo II, 132, 84084, Fisciano, Italy
[d] University of Salerno, Department of Chemistry and Biology "A. Zambelli", via Ponte Don Melillo, 84084, Fisciano, Italy
[e] University of Pisa, Department of Pharmacy, via Bonanno, 6, 56126, Pisa, Italy
[f] Centre for Regenerative Medicine and Devices, School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, BN2 4GJ, UK

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2 S RBD
ACE2
NMR
CD
MD
Structural biology
Peptides

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an RNA virus identified as the cause of the coronavirus outbreak in December 2019 (COVID-19). Like all the RNA viruses, SARS-CoV-2 constantly evolves through mutations in its genome, accumulating 1–2 nucleotide changes every month, giving the virus a selective advantage through enhanced transmissibility, greater pathogenicity, and the possibility of circumventing immunity previously acquired by an individual either by natural infection or by vaccination. Several SARS-CoV-2 variants of concern (VoC) have been identified, among which we find Alpha (Lineage B.1.1.7), Beta (Lineage B.1.351), and Gamma (Lineage P.1) variants. Most of the mutations occur in the spike (S) protein, a surface glycoprotein that plays a crucial role in viral infection; the S protein binds the host cell receptor, the angiotensin-converting enzyme of type 2 (ACE2) via the receptor binding domain (RBD) and catalyzes the fusion of the viral membrane with the host cell. In this work, we present the development of a simplified system that would afford to study the change in the SARS-CoV-2 S RBD/ACE2 binding related to the frequent mutations. In particular, we synthesized and studied the structure of short amino acid sequences, mimicking the two proteins' critical portions. Variations in the residues were easily managed through the one-point alteration of the sequences. Nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopies provide insights into ACE2 and SARS-CoV-2 S RBD structure with its related three variants (Alpha, Beta, and Gamma). Spectroscopy data supported by molecular dynamics lead to the description of an ACE2/RBD binding model in which the effect of a single amino acid mutation in changing the binding of S protein to the ACE2 receptor is predictable.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified as the cause of the coronavirus outbreak in December 2019 (COVID-19), when a group of atypical pneumonia cases was reported in Wuhan, China [1]. Molecular epidemiology and viral phylogeny studies suggested a zoonotic origin of the infection and proved that SARS-CoV-2 was already circulating from October to November 2019 [2,3]. The high transmissibility of the pathogen, combined with collective unconsciousness, led to record an ever-increasing number of infections worldwide, for which the World Health Organization (WHO) declared a pandemic in March of 2020.

A plethora of evidence has shown that SARS-CoV-2 infects lung cells through the surface viral glycoprotein spike (S), which binds the human angiotensin-converting enzyme 2 (hACE2), expressed in type II pneumocytes, myocardial cells, cholangiocytes, enterocytes, and oral mucosal epithelium [4, 5, 6]. S is a trimeric protein where each monomer is composed of S1 and S2 subunits. S1 contains a receptor binding domain (RBD) responsible for ACE2 recognition, while S2 mediates the virus and host cell membrane fusion via a conformational modification.

* Corresponding author.
E-mail address: dursi@unisa.it (A.M. D'Ursi).
[1] Equal contribution.

https://doi.org/10.1016/j.heliyon.2022.e11568
Received 10 May 2022; Received in revised form 7 September 2022; Accepted 7 October 2022

Since the pandemic outbreaks, enormous efforts have focused on developing new efficacious antiviral drugs, vaccines, and diagnostic tools to cope with health emergencies. As a result, several vaccines have been approved and are currently commercially available to immunize the worldwide population. However, despite these extraordinary results, a threat of the COVID-19 pandemic remains due to the rise and diffusion of mutated SARS-CoV-2 lineages, for which the newly developed vaccines, therapeutics, and diagnostic tools may become potentially inefficacious. Indeed SARS-CoV-2, like all RNA viruses, constantly evolves through mutations in its genome, accumulating 1–2 nucleotide changes every month [7]. While most mutations may be irrelevant, some may induce in the virus greater transmissibility, pathogenicity, and the ability to circumvent immunity previously acquired by natural infection or vaccination. Therefore, these virus variants have become a public health concern and must be monitored closely [8].

Five variants of concern (VoC) and numerous variants of interest (VoI) have been identified to date (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/). Among the VoC defined by WHO and European Center for Disease Prevention and Control (ECDC), Lineage B.1.1.7 (Alpha variant, e.d.s in UK), proved to have higher transmissibility [9]; Lineage B.1.351 (Beta variant, e.d.s. in South Africa, 2020) and Lineage P.1 (Gamma variant, e.d.s. in Brazil, 2021) showed a

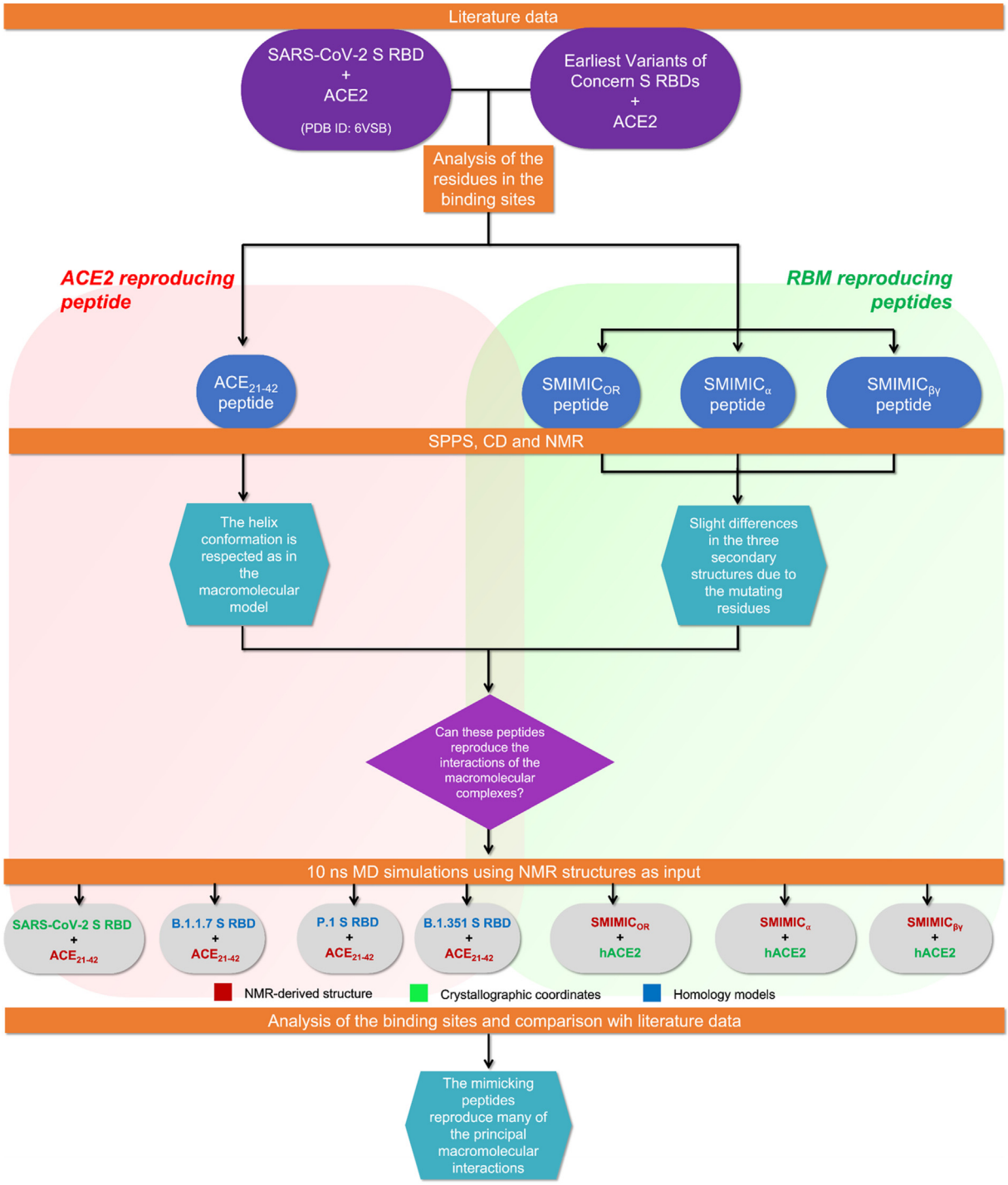

Figure 1. Workflow followed for the identification of ACE2 and S RBD mimicking peptides.

propensity for reinfection [10, 11, 12]. More recently, two additional VoC, Lineage B.1.617 (Delta variant, e.d.s in India, 2021) and Lineage B.1.1.529 (Omicron variant, e.d.s. in multiple countries, 2021), demonstrated a combination of immune evasion and increased transmissibility [13] that seem to escape most of the known neutralizing antibodies [14, 15, 16]. However, studies are still underway to confirm the efficacy of vaccines on the market against these variants. Meanwhile, their spread is faster than the production and distribution of *ad hoc* diagnostics, vaccines, and therapeutics, allowing variants to escalate ongoing COVID-19 outbreaks [17]. Therefore, it is necessary to design therapeutic strategies that are effective toward the original strain and its variants to slow down the spread of the virus and avoid the development of new lineages.

Among the targets considered to design anti-COVID-19 drugs and vaccines, the macromolecular complex composed of S and ACE2 proteins received great attention thanks to the available structural information [6, 18]. However, the design of drugs and vaccines goes hand in hand with the need to update the structural information of the S protein due to the high occurrence of point mutations in critical domains of the protein. In this context, we propose to study an S/ACE2 interaction model consisting of peptides mimicking their binding site. The ability of peptides to mimic larger protein domains has been extensively explored and proposed as a valuable tool in structural biology. Short peptides indeed can be synthesized as precise replicas of protein moieties and easily modified in their sequences using non-standard residues, allowing a broader interpretation of the function of a specific domain, and empowering the analysis at the level of individual amino acids [19, 20, 21, 22]. This strategy presents limitations, mainly when the binding pocket is buried inside the protein and is composed of amino acids deriving from different chains or subunits; however, it can be a helpful strategy when the binding site is composed of consecutive amino acids and is located at the interface between the two macromolecules [23], like in S/ACE2 complex.

On this basis, starting from the analysis of the SARS-CoV-2 S RBD/ACE2 crystallographic complex, we selected a 22-residues long sequence deriving from the N-terminus of ACE2 ($ACE_{21–42}$) and three 25-residues sequences corresponding to the S protein $^{482}G–Q^{506}$ residues of original SARS-CoV-2 S ($SMIMIC_{OR}$), Lineage B.1.1.7 S ($SMIMIC_{\alpha}$), Lineage B.1.351 S ($SMIMIC_{\beta}$) and Lineage P.1 S ($SMIMIC_{\gamma}$).

The sequences were synthesized, and their structures were studied by CD and NMR spectroscopies. The interaction of each sequence with the cognate ligand was examined by molecular dynamics (MD) simulations: $ACE_{21–42}$ was tested in complex with the RBDs of SARS-CoV-2 S protein and the three variants aforementioned, whereas $SMIMIC_{OR}$, $SMIMIC_{\alpha}$, and $SMIMIC_{\beta\gamma}$ were tested in complex with the crystal model of hACE2 protein (Figure 1).

Our data show that punctiform mutations induce structural perturbation extending throughout the binding site. This may have a dramatic effect on the SARS-CoV-2 S-ACE2 interaction at the origin of SARS-CoV-2 infection, thus providing a structural explanation of the critical change in the pathological profile of the VoCs.

1.1. Identification of ACE2 mimicking peptide and RBM mimicking peptide

1.1.1. Analysis of structural data

Analyzing the crystal structure of SARS-CoV-2 S RBD/ACE2 complex (PDB ID: 6M0J) [18], it appears that the interaction of ACE2 with S protein involves, on one side, the ACE2 N-terminal helix – residues ^{24}Q, ^{30}D, ^{37}E, ^{42}Q – and on the other side the SARS-CoV-2 S receptor binding motif (RBM 437–508) which is transiently exposed when S protein passes from a "down" to an "up" conformation while interacting with ACE2 (Figure 2) [6].

In particular, the stability of the S/ACE2 complex is due to a high number of negatively charged and hydrophobic residues interacting with the ACE2 N-terminus [24, 25, 26].

Based on this evidence, we first identified the sequences, including ACE2 N-terminus $^{21}I–Q^{42}$ ($ACE_{21–42}$) and RBM $^{437}G–Q^{508}$ (SMIMIC), as ACE2 and S binding domain mimicking peptides, respectively.

On the other hand, when the variants of concern began to spread out, *in silico* approaches were applied to understand if and how the one-point alterations in the S sequence might influence the binding with ACE2. In particular, the N501Y mutation in the RBD, present in the Alpha, Beta, and Gamma variants, has been found to enhance the binding affinity with ACE2 by affecting the electrostatic interactions and decreasing the repulsions between many other residues in the neighborhood [27]. Furthermore, the E484K mutation showed increased contact with ACE2, in particular with the ^{75}E residue; moreover, the two E484K and N501Y mutations together have been found to change the conformation of the binding loop compared to the wild-type sequence and thus stabilize an electrostatic interaction between E484K and ^{75}E sidechain [28].

Based on all these observations, we decided to focus on four peptide sequences and study them by combining spectroscopic and *in silico* data. The four selected sequences are:

- ^{21}IEEQAKTFLDKFNHEAEDLFYQ^{42} deriving from N-terminus of hACE2, named $ACE_{21–42}$;

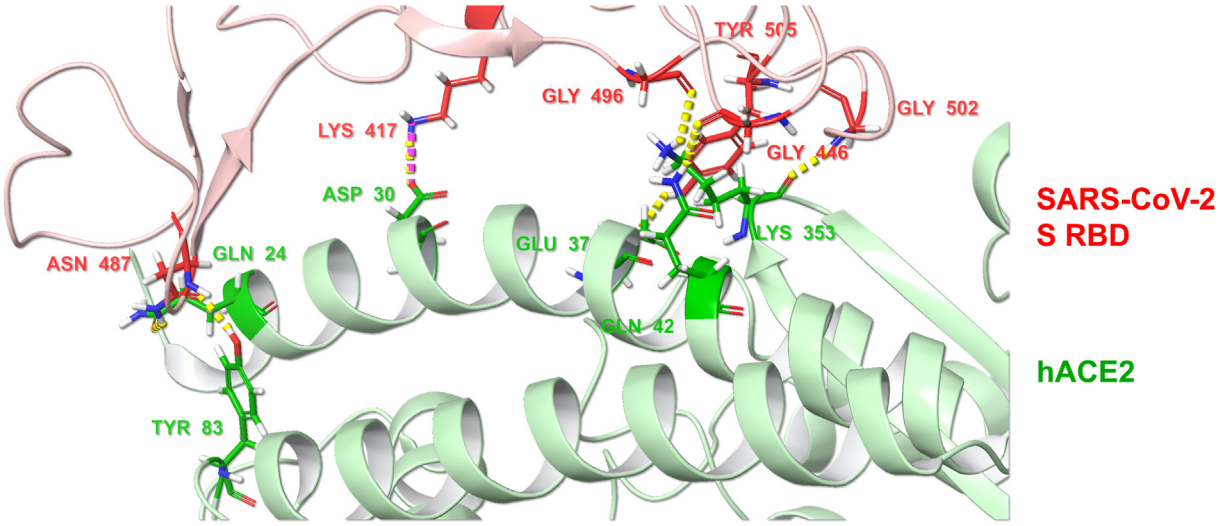

Figure 2. Ribbon representation of the complex SARS-CoV-2 S RBD (red)/ACE2 (green) (PDB ID: 6M0J) [18]. The residues involved in the interaction are labeled and shown in stick representation. H-bonds and salt bridge interactions are represented as yellow and pink dashed lines, respectively.

- ^{482}GVEGFNCYFPLQSYGFQPTNGVGYQ^{506} deriving from S RBM of the original lineage, named SMIMIC$_{OR}$;
- ^{482}GVEGFNCYFPLQSYGFQPTYGVGYQ^{506} containing the N501Y mutation typical of the B.1.1.7 lineage (Alpha) S RBM, named SMIMIC$_{\alpha}$;
- ^{482}GVKGFNCYFPLQSYGFQPTYGVGYQ^{506} containing the E484K and N501Y mutations typical of B.1.351 and P.1 lineages (Beta and Gamma) S RBMs, named SMIMIC$_{\beta\gamma}$.

The four peptides were synthesized and subjected to conformational analysis in hexafluoisopropanol (HFIP)/water 50/50 v/v using CD and NMR spectroscopies. Mixtures of water and organic fluorinated solvents are compatible with CD and NMR conformational analyses to stabilize secondary structures in peptides without affecting the intrinsic conformational attitude induced by the amino acid sequence [29, 30]. Finally, the complexes composed of peptides with the cognate macromolecules were subjected to MD simulations to evaluate the significance of the structural predictions.

2. Results

2.1. CD and NMR analysis of ACE$_{21–42}$

ACE$_{21–42}$ peptide was synthesized following the standard procedure of solid-phase peptide synthesis (SPPS) [31]. The secondary structure of the ACE$_{21–42}$ peptide was studied by CD and NMR techniques in a solution mixture of HFIP/water 50/50 v/v [32]. Taking advantage of the physical-chemical properties of the fluorinated solvents, HFIP/water mixtures have been extensively used to study the conformation of partially hydrophobic peptides [29, 33, 34, 35, 36].

Quantitative estimation of ACE$_{21–42}$ CD curve (DICHROWEB website, CONTIN algorithm) in HFIP/water 50/50 v/v (Figure S1) [37] indicates that the peptide assumes 56.1% α-helix, 8.9% β-sheet, 15% turn, and 20% random coil conformations.

NMR structure of ACE$_{21–42}$ was determined by collecting 2D TOCSY, NOESY, and ^{1}H–^{13}C HSQC spectra (Figure S2) in HFIP/water mixture 50/50 v/v (500 MHz Bruker Avance III at 298 K). ^{1}H and ^{13}C chemical shift assignments (Table S1) were carried out according to the Wüthrich procedure [38] by iteratively analyzing 2D spectra using SPARKY software [39].

Analysis of NOE connectivities reported in Figure 3A indicates sequential and medium-range correlations typical of regular secondary structures. NOE data were translated into interprotonic distances using the CALIBA routine of CYANA 2.1 software [40] and used as restraints in the structure calculations. Figure 3B shows the NMR structure bundle of ACE$_{21–42}$ as derived from the CYANA and TALOS+ [41] calculations based on 292 sequential and short-range NOE distances and 84 backbone dihedral angle restraints (Table S2). PROMOTIF analysis of dihedral angles using Kabsh and Sanders parameters indicates ACE$_{21–42}$ assuming an α-helix encompassing the residues ^{22}E–Q^{24} and ^{30}D–D^{38} and a 3$_{10}$ helix in the residues ^{25}A–L^{29} [42]. The structure was deposited on Protein Data Bank (http://www.wwpdb.org/) with the PDB ID: 7P55.

2.2. CD and NMR analysis of SMIMIC peptides

SMIMIC peptides were synthesized by SPPS and studied by CD and NMR in HFIP/water solution 50/50 v/v.

CD spectra of SMIMIC$_{OR}$, SMIMIC$_{\alpha}$, and SMIMIC$_{\beta\gamma}$ in HFIP/water, 50/50 v/v are shown in Figure S3A. Evaluation of CD curves indicates that all SMIMIC peptides assume incipient helix conformations, with a significant content of random coil conformations in SMIMIC$_{\alpha}$ and SMIMIC$_{\beta\gamma}$ (Figure S3B).

SMIMICs were studied by NMR in HFIP/water mixture 50/50 v/v. The chemical shift assignment was carried out as previously described (Figures S4–S6, Tables S3–S5). Figure 4 shows sequential and medium-range NOE effects collected in the NOESY experiments and used for structure calculation. SMIMIC$_{OR}$, SMIMIC$_{\alpha}$ and SMIMIC$_{\beta\gamma}$ assume type II turn in the ^{482}G–G^{485} and 3$_{10}$ helix on the ^{491}P–Q^{498} segments, respectively. The small region connecting these two segments is unordered, while the C-terminal ^{499}P–N^{501}/^{499}P–Y^{501} and ^{502}G–G^{504} assume specific conformations in each of the three peptides (Figure 4).

Figure 4 shows NMR structure bundles of SMIMIC peptides. Each bundle includes up to 50 structures calculated using CYANA software based on NOE restraints [40] and selected according to the lowest values of the target function. The structures are superimposed on the backbone heavy atoms of the residues ^{491}P–P^{499} and filtered for RMSD values <0.8 Å. Statistics for the final NMR ensembles are reported in Table S6.

PROMOTIF analysis of the bundles confirms the presence of the 3$_{10}$ helix in the ^{492}L–Y^{495} portion. N- and C-terminal extremities assume different conformations in the correspondence of the punctiform

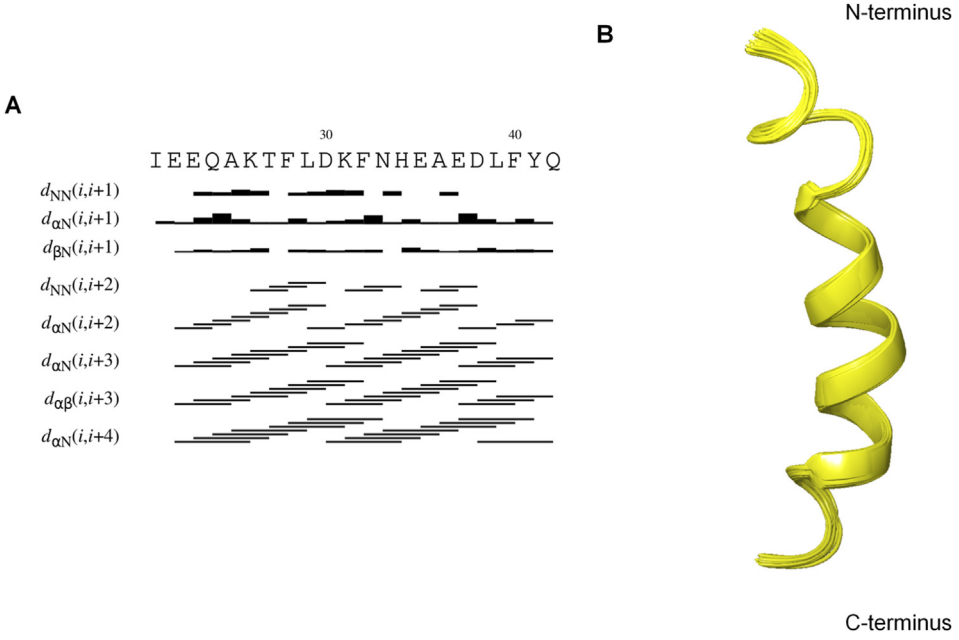

Figure 3. (A) An overview of the sequential and medium-range nuclear Overhauser enhancements (NOEs) used to calculate the ACE$_{21-42}$ structure ensemble; (B) ribbon visualization of the representative structures of the calculated ensemble.

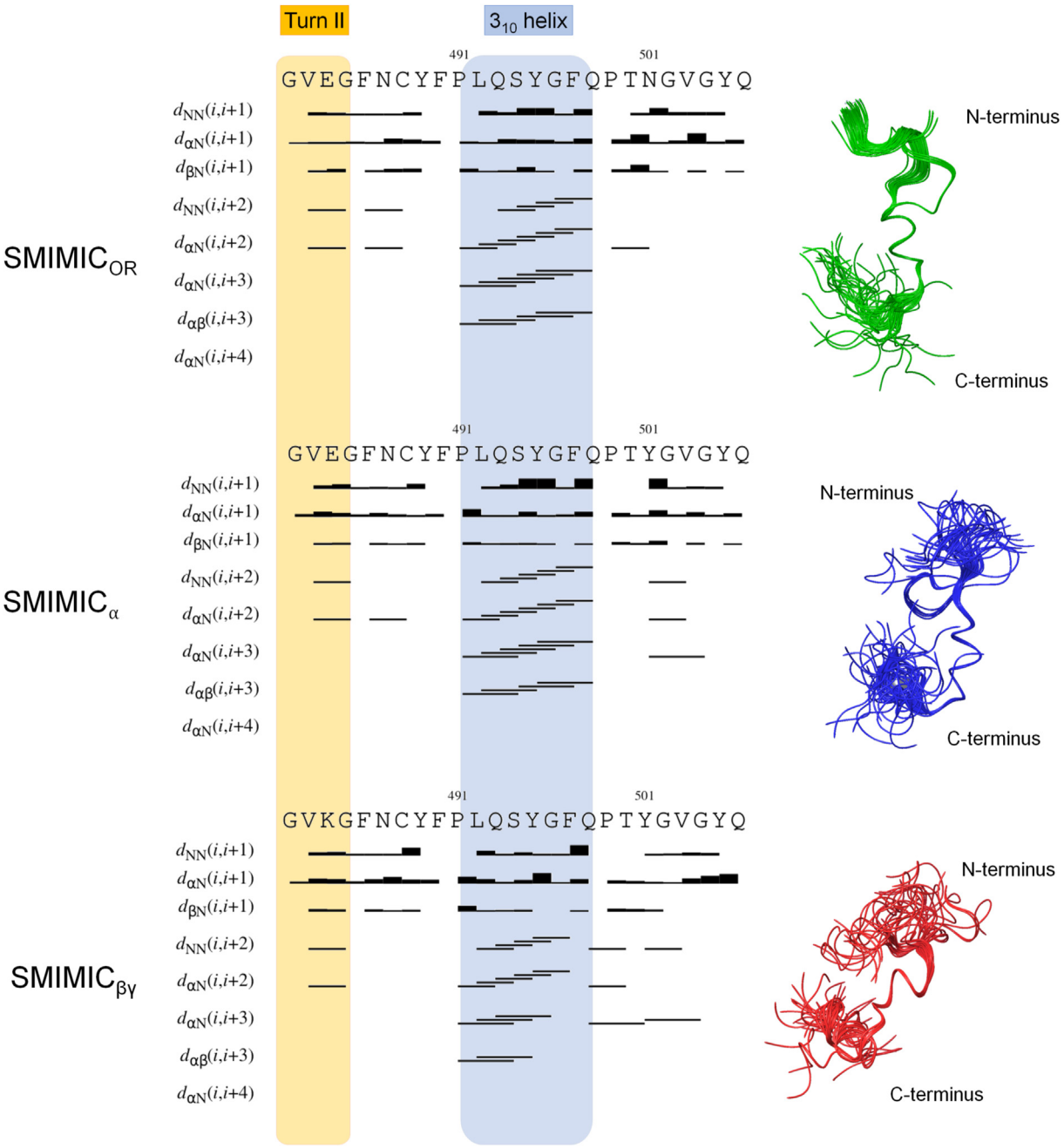

Figure 4. Sequential and medium-range interprotonic connectivities collected in the NOESY spectra of SMIMIC$_{OR}$, SMIMIC$_{\alpha}$, and SMIMIC$_{\beta\gamma}$ in HFIP/water 50/50 v/v. The NMR structure bundle is shown for each SMIMIC peptide on the right.

mutations. Interestingly, these data show that the mutation causes a conformational effect on the whole amino acid sequence rather than the mutation site. SMIMIC NMR structures were deposited on Protein Data Bank (http://www.wwpdb.org/) with the PDB IDs: 7P5G (SMIMIC$_{OR}$), 7P5Q (SMIMIC$_{\alpha}$), and 7P5S (SMIMIC$_{\beta\gamma}$).

2.3. MD simulations

To understand if the selected sequences are useful for the construction of a model reproducing the main interactions in the S RBD/ACE2 binding site, two sets of MD simulations were carried out:

- ACE$_{21-42}$ in complex with the full-length RBDs of SARS-CoV-2 S and the Alpha (B.1.1.7), Beta (B.1.351), and Gamma (P.1) lineages;
- SMIMIC$_{OR}$, SMIMIC$_{\alpha}$, SMIMIC$_{\beta\gamma}$ in complex with the hACE2 protein.

2.3.1. Simulation of ACE$_{21-42}$ with SARS-CoV-2 S RBD and its variants

As previously reported, the interaction of S protein with ACE2 consists of a dense hydrogen bonding network, including ^{446}G/^{42}Q, ^{487}N/^{24}Q–^{83}Y, ^{496}G–^{502}G/^{353}K, ^{505}Y/^{37}E, and ^{417}K/^{30}D [24, 25, 26].

The structural model of S protein in Alpha (B.1.1.7), Beta (B.1.351), and Gamma (P.1) lineages of SARS-CoV-2 variants were calculated using homology modeling (ExPASy SWISS-MODEL) [43, 44] (Figure S7). SARS-CoV-2 S RBD structural coordinates were used (residues 319–541) as templates (Table S7). The models were energy minimized and subjected to 2 ns MD simulation in water to stabilize the system; the PDB coordinates of the three RBDs were saved from the last frame of the trajectories and then used to build the complexes with the ACE2 crystal structure.

MD simulations were carried out using GROMACS/2020.6 [45,46]. A 10 ns MD simulation (CHARMM36 force field [47, 48]) was carried out in water using SARS-CoV-2 S RBD crystal structure (PDB ID: 6M0J) [18], and the RBDs

Table 1. Critical residues in SARS-CoV-2 original lineage, B.1.1.7, B.1.351 and P.1 S RBDs/$ACE_{21–42}$ interaction resulting from 10 ns MD simulations.

SARS-CoV-2 S RBD	Interaction	$ACE_{21–42}$	B.1.1.7 S RBD	Interaction	$ACE_{21–42}$	B.1.351 S RBD	Interaction	$ACE_{21–42}$	P.1 S RBD	Interaction	$ACE_{21–42}$
K417	H-Bond, salt bridge	E35	F456	π–π stacking	F32	Y489	π–π stacking	F40	K484	H-bond, salt bridge	E23
E484	H-bond, salt bridge	K31	K458	H-bond, salt bridge	E35	Q493	H-bond	F40	Q493	H-bond, salt bridge	E22, E23, Q24
Y489	H-bond	H34	Y473	H-bond	F28, E35	K484	H-bond, salt bridge	E37			
Q493	H-bond	Q24	F490	H-bond	Q24						

of B.1.1.7, B.1.351 and P.1 variants, previously calculated by homology modeling. RMSD plots show that the systems reach equilibrium after 2 ns and are stable during the simulation (Figures S5-8A). Table 1 reports the critical interactions stabilizing the four complexes during the simulations. As demonstrated by the trajectory analysis, the stability of the complex is mainly based on electrostatic interactions involving the side chains of several Lys in RBD and Glu in $ACE_{21–42}$. Moreover, an extended network of H-bonds is evident with an additional contribution of π–π stacking interactions. A detailed list of interactions is reported in the Supplementary material (Tables S8–S11).

Figure 5 shows the interaction of SARS-CoV-2 S and its variants with $ACE_{21–42}$ peptide derived from the MD simulations' last frames. In addition, the model of SARS-CoV-2 S and hACE2 protein interaction model is also shown for reference. Analysis of the complexes indicates that $ACE_{21–42}$ is oriented as in the corresponding region of hACE2 protein when in complex with SARS-CoV-2 S or with SARS-CoV-2 S B.1.1.7: the charged residues are exposed to the solvent and the interaction is driven by H-bonds and hydrophobic contacts. However, $ACE_{21–42}$ orientation changes in binding P.1 and B.1.351 S RBDs as the negatively charged residues of $ACE_{21–42}$ peptide are directed against the positively charged ^{484}K (Figure 6).

2.3.2. Simulations of SMIMICs with hACE2

SMIMICs NMR structures in complex with the crystal structure of hACE2 (PDB ID: 6M0J [18]) were subjected to 10 ns MD simulation as previously described. In the preparation of the system, SMIMIC peptides were positioned in proximity of the hACE2 moiety including the residues 21–42. RMSD plot indicates that the systems reach equilibrium after 2 ns and are stable during all the simulations (Figures S12–S14A). Furthermore, the trajectory analysis indicates that several bonds stabilize the SMIMICs/ACE2 complexes throughout the simulation (Table 2). By comparing the critical interactions, it is possible to note that as $SMIMIC_{OR}$ and $SMIMIC_{βγ}$ steadily interact with hACE2 throughout the entire simulations especially with their residues in the N-terminus, $SMIMIC_{α}$ establishes a heterogeneous set of networks with its whole sequence, managing also to reach residues in hACE2 that are not included in the examined 21–42 region. The detailed list of interactions for all the time steps is reported in the Supplementary material (Figures S12–S14, Tables S12–S14).

Figure 7 shows the binding surfaces for each SMIMIC with hACE2. As evident, negatively charged residues on $SMIMIC_{OR}$ and $SMIMIC_{α}$ are exposed to engage electrostatic interactions with the Lys positively charged side chains. Noticeably several Lys residues in ACE2 N-terminus take part in the interaction with SARS-CoV-2 S.

When the E484K mutation occurs in the $SMIMIC_{βγ}$ sequence, the peptide points to the negatively charged portion of ACE2. Nevertheless, the stability of the complex is preserved thanks to a tight network of H-bonds and hydrophobic contacts.

3. Discussion

Despite the numerous therapies and vaccines developed in record time to manage the SARS-CoV-2 outbreak, the COVID-19 pandemic

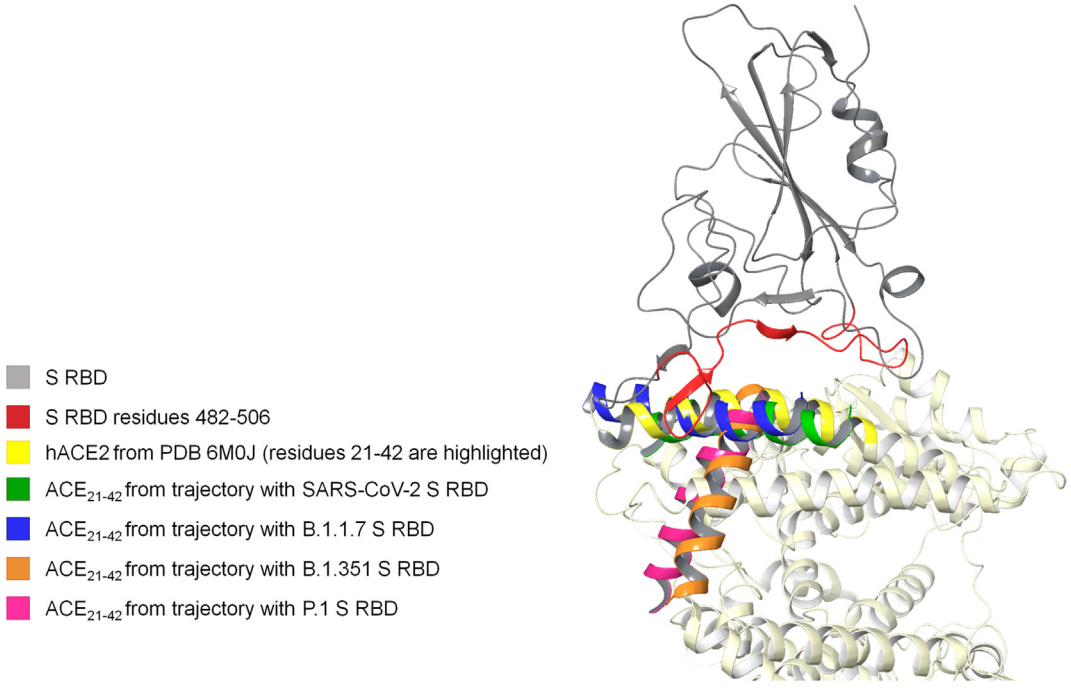

Figure 5. Ribbon representation of $ACE_{21–42}$ poses taken from the last frames in the MD simulations in complex with SARS-CoV-2 S RBD (green), B.1.1.7 S RBD (blue), B.1.351 S RBD (orange) and P.1 S RBD (pink) and superimposed with hACE2 (yellow ribbons, residues 21–42 are highlighted). Reference RBD is shown as a grey ribbon, with the 482–506 residues colored in red.

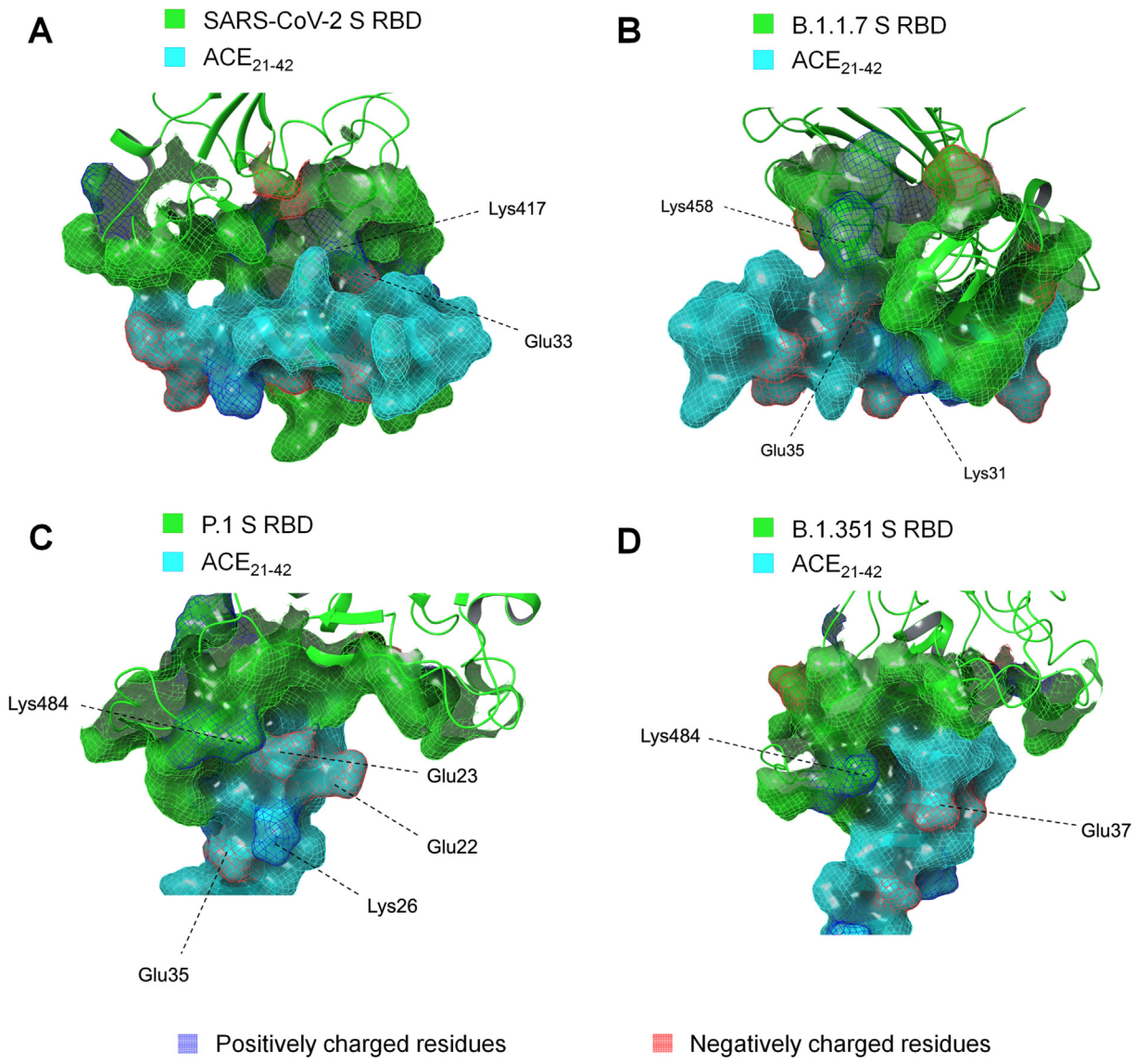

Figure 6. Visualization of the binding surfaces from the last frame of 10 ns MD simulations between $ACE_{21–42}$ (light blue ribbon and surface) and (A) SARS-CoV-2, (B) B.1.1.7, (C) P.1, and (D) B.1.351 S RBDs (green ribbons and surfaces). Positively and negatively charged residues are reported as blue and red meshes, respectively.

Table 2. Critical residues in $SMIMIC_{OR}$, $SMIMIC_{\alpha}$ and $SMIMIC_{\beta\gamma}$/hACE2 interaction resulting from 10 ns MD simulations.

$SMIMIC_{OR}$	Interaction	hACE2	$SMIMIC_{\alpha}$	Interaction	hACE2	$SMIMIC_{\beta\gamma}$	Interaction	hACE2
G482	H-bond, salt bridge	E23	G482	H-bond, salt bridge	D30	Y489	H-bond	Q24
E484	H-bond, salt bridge	S19	Y501	H-bond	D30, E37	S494	H-bond	D30
V483	H-bond	T27	Y501	π–π stacking	H34	Q493	H-bond	H34
			T500	H-bond	E37	L492	H-bond	T27
			F497, Q498	H-bond	L351			
			G496	H-bond	R357			
			Y495	H-bond	D355			

cannot yet be considered an ended chapter. SARS-CoV-2 is a coronavirus that tends to replicate and mutate quickly, as demonstrated by several variants that have been identified to date. Therefore, it is necessary to design therapeutic strategies that are effective toward the original strain and its variants to slow down the spread of the virus and avoid the development of new lineages.

One of the most studied targets for drug or COVID-19 vaccine design is the membrane spike glycoprotein (S). To design S protein targeting drugs and vaccines, it is strategic to have information on its structure when interacting with the ACE2 receptor.

Indeed, SARS-CoV-2 S protein has been extensively studied, and several structural models derived from crystallographic and electron microscopy data are currently deposited on online databases [6, 18]. However, determining the structural coordinates of a macromolecular complex consisting of two proteins, as in the case of ACE2 and SARS-CoV-2 S, is extremely expensive and time-consuming, while on the

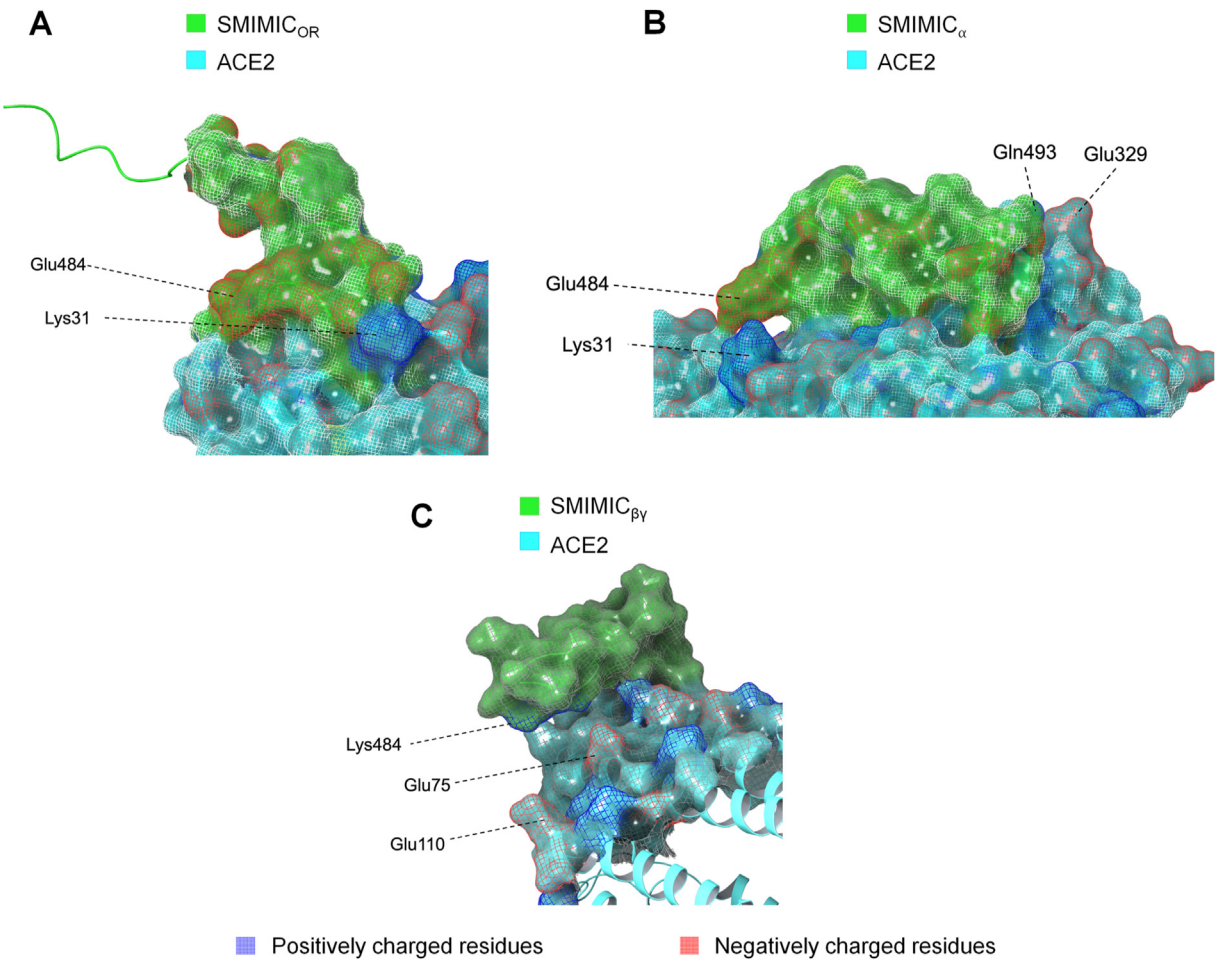

Figure 7. Visualization of the binding surfaces from the last frame of 10 ns MD simulations between (A) SMIMIC$_{OR}$, (B) SMIMIC$_{\alpha}$ and (C) SMIMIC$_{\beta\gamma}$ (green ribbons and surfaces) and hACE2 (light blue ribbon and surface). Positively and negatively charged residues are reported as blue and red meshes, respectively.

other hand, SARS-CoV-2 tends to mutate S protein sequence with extreme rapidity.

We propose simplified models consisting of peptide mimics of the S RBD/ACE2 binding domain to acquire valuable structural information focusing on sequences composed of consecutive amino acids, whose critical role was previously established [18, 24, 25, 26, 27, 28]. The role of peptides in mimicking larger domains has been extensively investigated. Moreover, given their chemical nature and ease in manipulating their sequence, peptides are considered suitable for the reproduction of protein-binding sites, especially for those concerning large protein-protein interfaces [19, 20, 21, 22]. To test the validity of this model, we identified four peptides deriving from ACE2 and SARS-CoV-2 RBM (residues 437–508), studied their structures, and simulated by molecular dynamics (MD) the reliability in reproducing ACE2/SARS-CoV-2.

Specifically, we identified four peptides mimicking the binding sites of ACE2/SARS-CoV-2: (i) ACE$_{21–42}$ as mimicking of ACE2 protein, (ii) SMIMIC$_{OR}$, mimicking SARS-CoV-2 S RBM, corresponding to ^{482}G–Q^{506} of S protein, (iii) SMIMIC$_{\alpha}$ mimicking the Alpha (B.1.1.7) variant of SARS-CoV-2 S RBM, corresponding to ^{482}G–Q^{506} of S protein and including N501Y mutation, (iv) SMIMIC$_{\beta\gamma}$ mimicking the Beta (B.1.351) and Gamma (P.1) variants of SARS-CoV-2 S RBM and including both E484K and N501Y mutations [18, 26].

CD and NMR conformational data show that, while ACE$_{21–42}$ assumes a helix conformation similar to the one in the corresponding region of the parent macromolecule, the SMIMICs include two conserved secondary structure motifs only in the N-terminal (β turn), and the central part (3_{10} helices). The segment connecting these two motifs is different and represents an area of structural flexibility due to the presence of glycine residues and the E484K mutation in SMIMIC$_{\beta\gamma}$. The region downstream of the 3_{10} helix also assumes variable secondary structures or is disordered. E484K and N501Y mutations induce changes in conformation that are not localized to the mutation sites but propagate throughout the sequence giving each peptide peculiar molecular shapes.

To test the stability of the complex and to explore the local neighbourhood around the interaction sites we performed 10 ns MD simulations. The peptides, sampled from the NMR conformations, were simulated in complex with the respective molecular target, in particular, (i) ACE$_{21–42}$ with RBDs of SARS-CoV-2, B.1.1.7, P.1, and B.1.351 and (ii) SMIMIC$_{OR}$, SMIMIC$_{\alpha}$, and SMIMIC$_{\beta\gamma}$ in complex with hACE2 protein structure. The N-terminal α-helix 21–42 of the ACE2 protein is primarily involved in the interaction with RBD of SARS-CoV-2. Conformational and MD data show that the helix structure is reproduced in the ACE$_{21–42}$ peptide and is highly stable, as shown by the MD data. The binding poses of ACE$_{21–42}$ with SARS-CoV-2 S and B.1.1.7 S RBD is superimposable with the corresponding fragment in the crystal structure of the whole ACE2 protein. Conversely, E484K mutation in B.1.351 and P.1 S RBD alters the orientation of ACE$_{21–42}$, favoring electrostatic interactions with the positively charged side chain of ^{484}K (Figure 5). Moreover, E484K and N501Y mutations change the chemical-physical properties of the binding surface; in the absence of E484K mutation, as in ACE$_{21–42}$ interacting with SARS-CoV-2 or the Alpha variant S RBDs, binding occurs predominantly through H-bonds and hydrophobic interactions, while the charged residues are exposed to the solvent. On the other hand, when

E484K mutation is present as in the interaction of $ACE_{21–42}$ with Beta and Gamma variants RBDs, the charged residues, i.e. ^{484}K, are engaged in electrostatic interactions, whereby the binding interface is mainly uncharged (Figure 6).

MD simulations show that the three SMIMIC peptides have conformational flexibility to fit the hACE2 binding site. Analysis of SMIMICs binding poses derived from the last frame of MD indicates that the peptides interact with the hACE2 N-terminal α-helix through a tight network of H-bonds; moreover, the binding involves ^{357}R, ^{353}K, and ^{355}D, residues that lie outside the N-terminal helix (Table 2). Even MD simulation confirms that E484K mutation is determinant in orienting $SMIMIC_{βγ}$ peptide to a negatively charged area in proximity of the N-terminus and in changing the polarity of the binding surface (Figure 7).

4. Conclusions

In this work, we identified, synthesized, and characterized a peptide derived from the residues 21–42 of hACE2 ($ACE_{21–42}$) and three peptides, each derived from the RBM of S protein and three of its variants ($SMIMIC_{OR}$, $SMIMIC_{α}$, and $SMIMIC_{βγ}$). CD and NMR data show that $ACE_{21–42}$ assumes a helix conformation superimposable to the corresponding region in the full-length hACE2 protein. This structure preserves stability when subjected to MD in water, in the presence of the cognate wild-type S RBD receptor. E484K mutation in the S RBD receptor induces a change in $ACE_{21–42}$ helix orientation to engage contacts with negatively charged receptor sites.

On the other hand, according to NMR conformational analysis, S mimicking peptides – $SMIMIC_{OR}$, $SMIMIC_{α}$, and $SMIMIC_{βγ}$ – assume 3_{10} helix conformation in the conserved residues but are flexible in the neighbors of the mutation sites. MD simulations show that this flexibility affects the binding with ACE2 and is functional for the peptide to reach charged residues on the receptor. Specifically, $SMIMIC_{OR}$ and $SMIMIC_{α}$ bind $ACE_{21–42}$ through H-bonds and hydrophobic interactions, while $SMIMIC_{βγ}$ through electrostatic interactions. Consequently, a change in the characteristics of the contact surface occurs as the E484K mutated protein points its charged residues to the binding pocket, while $SMIMIC_{OR}$ and $SMIMIC_{α}$ point the charged residues to the solvent.

In conclusion, we can state that the strategy proposed might offer the opportunity to reproduce large macromolecule patterns and focus on the contribution of each residue in the binding site, in the case of interacting residues at the interface, using a combination of spectroscopy and computational data. We believe that these findings define a valuable strategy in the perspective of building a model for the rapid development of molecules targeting SARS-CoV-2 S and its variants. This model is helpful in a preliminary step for optimizing new lead compounds or peptide sequences capable of interacting with SARS-CoV-2 S RBD with high affinity and specificity, applicable in versatile devices for diagnostic and therapeutic use.

5. Material and methods

5.1. Peptides synthesis

$ACE_{21–42}$, $SMIMIC_{OR}$, $SMIMIC_{α}$ and $SMIMIC_{βγ}$ were manually synthesized using Fmoc/tBu solid-phase peptide synthesis (SPPS) following the standard Merrifield strategy [31]. Fmoc-protected amino acids were coupled using 1-hydroxybenzotriazole and O-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (four-fold excess) as coupling reagents. A six-fold excess of N, N-diisopropylethylamine was added to the solution as a scavenger. Wang resin was used for peptide synthesis. Cleavage and side chain deprotection were carried out using a 90% trifluoroacetic acid (TFA), 5% water, and 5% triisopropylsilane (TIS) solution for 3 h. After the cleavage step, the functionalized resin was filtered using a cold diethyl ether solution to precipitate the peptide. Raw peptides were purified by reversed-phase chromatography (HPLC) using the Phenomenex C18 column. Peptides were characterized on a Finningan L.C.Q. Deca ion trap instrument equipped with an electrospray source (LCQ Deca Finnigan, San José. CA, USA). The samples were directly infused in the ESI source using a syringe pump at a 5.0 mL/min flow rate. The data were analyzed using the Xcalibur software. The sample purity was >98%.

5.2. CD experiments

The CD spectra were obtained using a JASCO J-810 spectropolarimeter, with a 1 mm long quartz cell and working at a temperature of 25 °C. The CD spectra were obtained by an average of 4 scans, in a measuring range 260-190 nm, at a bandwidth of 1 nm, and a scanning speed of 10 nm/min. Each spectrum was processed by subtracting the solvent spectrum. The analysis of the CD curves was performed using the CONTIN algorithm of the online platform DICHROWEB [37, 49]. All CD spectra were acquired with a concentration of 0.15 mM of peptide. The solvent system used was HFIP/water 50/50 v/v. Each spectrum was processed by subtracting the solvent spectrum.

5.3. NMR data recording and processing

All NMR spectra were acquired at 298 K on Bruker AVANCE III 500 MHz spectrometer. $ACE_{21–42}$, $SMIMIC_{OR}$, $SMIMIC_{α}$, and $SMIMIC_{βγ}$ peptides (2.5 mM) were dissolved in a mixture of HFIP/water 50/50 v/v. All NMR samples were added of 10% (v/v) D_2O. 2D $^{1}H–^{1}H$ homonuclear TOCSY and NOESY and $^{1}H–^{13}C$ heteronuclear HSQC experiments were transformed and visualized in TopSpin 3.5 (Bruker Biospin). The water signal was suppressed [50]. TOCSY and NOESY experiments were acquired using 80 and 200 ms mixing times, respectively. Chemical shifts assignment was obtained using the standard approach described by Wüthrich [38]. 2D TOCSY, NOESY, and HSQC spectra were analyzed using SPARKY software [39]. The acquisition of 2D NOESY experiments made possible to collect intramolecular distance restraints derived from Nuclear Overhauser Enhancements (NOEs). The assigned chemical shift values of backbone ^{15}N, $^{13}Cα$, and $^{13}Cβ$ were used as input for the TALOS+ software [41] to predict backbone dihedral angles.

5.4. Structure calculations

Assigned peaks of NMR spectra were integrated using the Gaussian fit integration method of SPARKY software. Peak volumes deriving from the assignment were translated into upper distance bounds with the CALIBA routine from the CYANA 2.1 software package [40]. Redundant and duplicate constraints were discarded for each sample, and the final list of constraints was used to generate a set of 50 structures using the CYANA protocol of simulated annealing in torsion angle space (50000 steps). Entries presenting the lowest target function value (2–12) and irrelevant residual violation (maximum violation = 0.71 Å) were analyzed using Schrödinger's Maestro 12.5.139 [51].

5.5. Molecular dynamics

MD simulations were performed with GROMACS/2020.6 [45,46] using the High-Performance Computer Marconi100, whose access was kindly provided by CINECA, within the "COVID-19 Fast access to HPC supercomputing facilities" Call for Proposals organized by Associazione Big Data. The topology files were generated using CHARMM36 all-atom force field [47, 48]. The complexes were solvated in cubic boxes with the TIP4P water model. Na^{+} and Cl^{-} ions were added to neutralize the charge of the system. After these steps, the energy minimization of the system was performed using the steepest descent integrator, and then the system was equilibrated using NVT and NpT runs. The system's temperature and pressure were kept constant at 300 K and 1.01325 bar using the Berendsen weak coupling method [52]. The results were used for an MD simulation using Particle Mesh Ewald for long-range electrostatics under NpT conditions. Coordinates were saved every 100 ps. Trajectory files

containing the coordinates of the receptor-ligand complex at different time steps (from 100 ps to 1 ns) were fitted in the box and converted in PDB coordinates by using *trjconv* tool of GROMACS package. The structures were visualized with Maestro 12.5.139 [51]. Analyses of RMSD and the number of bonds were carried out for the MD simulations of each system using *rms* and *hbond* tools of GROMACS package.

5.6. Homology model

Homology models were built using ExPASy SWISS-MODEL [43, 44]. B.1.1.7, B.1.351 and P.1 S RBD protein templates were chosen according to three parameters: (i) sequence similarity; (ii) global model quality estimate (GMQE) value, a quality estimation combining properties from the target–template alignment, and the template search method. It is expressed as a number between 0 and 1, reflecting the accuracy of the model (the higher the value, the higher the reliability); (iii) qualitative model energy analysis (QMEAN) value [53, 54] a composite estimator based on different geometrical properties which provide global and local absolute quality estimates based on one single model. The PDB models generated were prepared by adding hydrogen atoms and removing the N-Acetylglucosamine (NAG) residues and used for the MD calculation.

Declarations

Author contribution statement

Michela Buonocore, Angelo Santoro: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Manuela Grimaldi: Analyzed and interpreted the data.

Verdiana Covelli, Mohammad Firoznezhad, Manuela Rodriquez: Performed the experiments.

Matteo Santin, Anna Maria D'Ursi: Conceived and designed the experiments; Contributed reagents, materials, analysis tools and data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data associated with this study has been deposited at Protein Data Bank under the accession number 7P55 7P5G 7P5Q 7P5S.

Declaration of interest's statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e11568.

Acknowledgements and Funding

We acknowledge CINECA for awarding us access to Marconi100 based at Bologna, Italy. The support of Alessandro Grottesi and Paola Alberigo from CINECA for the technical work is greatly acknowledged.

References

[1] H. Lu, C.W. Stratton, Y.W. Tang, Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle, J. Med. Virol. 92 (2020) 401.
[2] A. Lai, A. Bergna, C. Acciarri, M. Galli, G. Zehender, Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2, J. Med. Virol. 92 (2020) 675–679.
[3] M. Buonocore, C. Marino, M. Grimaldi, A. Santoro, M. Firoznezhad, O. Paciello, F. Prisco, A.M. D'Ursi, New putative animal reservoirs of SARS-CoV-2 in Italian fauna: a bioinformatic approach, Heliyon 6 (2020), e05430.
[4] M. Gheblawi, K. Wang, A. Viveiros, Q. Nguyen, J.-C. Zhong, A.J. Turner, M.K. Raizada, M.B. Grant, G.Y. Oudit, Angiotensin-converting enzyme 2: SARS-CoV-2 receptor and regulator of the renin-angiotensin system: celebrating the 20th anniversary of the discovery of ACE2, Circ. Res. 126 (2020) 1456–1474.
[5] J. Luan, Y. Lu, X. Jin, L. Zhang, Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection, Biochem. Biophys. Res. Commun. 526 (2020) 165–169.
[6] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science 367 (2020) 1260–1263.
[7] N.D. Grubaugh, E.B. Hodcroft, J.R. Fauver, A.L. Phelan, M. Cevik, Public health actions to control new SARS-CoV-2 variants, Cell 184 (2021) 1127–1132.
[8] A. Sanyaolu, C. Okorie, A. Marinkovic, N. Haider, A.F. Abbasi, U. Jaferi, S. Prakash, V. Balendra, The emerging SARS-CoV-2 variants of concern, Therapeut. Adv. Infect. Dis. 8 (2021), 20499361211024372.
[9] S.E. Galloway, P. Paul, D.R. MacCannell, M.A. Johansson, J.T. Brooks, A. MacNeil, R.B. Slayton, S. Tong, B.J. Silk, G.L. Armstrong, Emergence of SARS-CoV-2 b. 1.1. 7 lineage—United States, december 29, 2020–january 12, 2021, Morb. Mortal. Wkly. Rep. 70 (2021) 95.
[10] N. Zucman, F. Uhel, D. Descamps, D. Roux, J.-D. Ricard, Severe reinfection with South African SARS-CoV-2 variant 501Y. V2: a case report, Clin. Infect. Dis.: Off. Publ. Infect. Dis. Soc. Am. (2021).
[11] T. Staub, V. Arendt, E.C.L. de la Vega, P. Braquet, C. Michaux, M. Kohnen, C. Tsobo, T. Abdelrahman, A. Wienecke-Baldacchino, J.-H. Francois, Case series of four re-infections with a SARS-CoV-2 B. 1.351 variant, Luxembourg, February 2021, Euro Surveill. 26 (2021), 2100423.
[12] F. Naveca, C. da Costa, V. Nascimento, V. Souza, A. Corado, F. Nascimento, Á. Costa, D. Duarte, G. Silva, M. Mejía, Three SARS-CoV-2 Reinfection Cases by the New Variant of Concern (VOC) P. 1/501Y. V3, 2021.
[13] P. Mlcochova, S.A. Kemp, M.S. Dhar, G. Papa, B. Meng, I.A. Ferreira, R. Datir, D.A. Collier, A. Albecka, S. Singh, SARS-CoV-2 B. 1.617. 2 Delta variant replication and immune evasion, Nature 599 (2021) 114–119.
[14] J. Chen, R. Wang, N.B. Gilby, G.-W. Wei, Omicron (B. 1.1. 529): Infectivity, Vaccine Breakthrough, and Antibody Resistance, ArXiv (2021).
[15] F. Fratev, The High Transmission of SARS-CoV-2 Omicron (B. 1.1. 529) Variant Is Not Only Due to its hACE2 Binding: A Free Energy of Perturbation Study, bioRxiv (2021).
[16] Y. Cao, J. Wang, F. Jian, T. Xiao, W. Song, A. Yisimayi, W. Huang, Q. Li, P. Wang, R. An, Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies, Nature (2021) 1–9.
[17] R. Noor, A review on the effectivity of the current COVID-19 drugs and vaccines: are they really working against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants? Curr. Clin. Microbiol. Rep. (2021) 1–8.
[18] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, Nature 581 (2020) 215–220.
[19] J. Eichler, Peptides as protein binding site mimetics, Curr. Opin. Chem. Biol. 12 (2008) 707–713.
[20] A. Groß, C. Hashimoto, H. Sticht, J. Eichler, Synthetic peptides as protein mimics, Front. Bioeng. Biotechnol. 3 (2016) 211.
[21] O. Marin, F. Meggio, B. Boldyreff, O.-G. Issinger, L.A. Pinna, Dissection of the dual function of the β-subunit of protein kinase CK2 ('casein kinase-2'): a synthetic peptide reproducing the carboxyl-terminal domain mimicks the positive but not the negative effects of the whole protein, FEBS Lett. 363 (1995) 111–114.
[22] M. Grimaldi, I. Stillitano, G. Amodio, A. Santoro, M. Buonocore, O. Moltedo, P. Remondelli, A.M. D'Ursi, Structural basis of antiviral activity of peptides from MPER of FIV gp36, PLoS One 13 (2018), e0204042.
[23] P. Morales, M. Bruix, M.A. Jiménez, Structural insights into β-arrestin/cb1 receptor interaction: NMR and CD studies on model peptides, Int. J. Mol. Sci. 21 (2020) 8111.
[24] A. Ali, R. Vijayan, Dynamics of the ACE2–SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms, Sci. Rep. 10 (2020) 1–12.
[25] S. Gao, J. Luan, H. Cui, L. Zhang, ACE2 isoform diversity predicts the host susceptibility of SARS-CoV-2, Transbound. Emerg. Dis. 68 (2021) 1026–1032.
[26] F. Curreli, S.M. Victor, S. Ahmed, A. Drelich, X. Tong, C.-T.K. Tseng, C.D. Hillyer, A.K. Debnath, Stapled peptides based on Human Angiotensin-Converting Enzyme 2 (ACE2) potently inhibit SARS-CoV-2 infection in vitro, mBio 11 (2020) e02451-2420.
[27] F. Ali, A. Kasry, M.J.M.i.d.d. Amin, The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant 10, 2021, 100086.
[28] G. Nelson, O. Buzko, P.R. Spilman, K. Niazi, S. Rabizadeh, P.R. Soon-Shiong, Molecular Dynamic Simulation Reveals E484K Mutation Enhances Spike RBD-ACE2 Affinity and the Combination of E484K, K417N and N501Y Mutations (501Y. V2 Variant) Induces Conformational Change Greater than N501Y Mutant Alone, Potentially Resulting in an Escape Mutant, BioRxiv (2021).
[29] N. Österlund, J. Luo, S.K. Wärmländer, A. Gräslund, Membrane-mimetic systems for biophysical studies of the amyloid-β peptide, Biochim. Biophys. Acta, Proteins Proteomics 1867 (2019) 492–501.
[30] M. Scrima, A. Le Chevalier-Isaad, P. Rovero, A.M. Papini, M. Chorev, A.M. D'Ursi, CuI-Catalyzed azide-alkyne intramolecular i-to-(i+4) side-chain-to-side-chain cyclization promotes the formation of helix-like secondary structures, Eur. J. Org Chem. 2010 (2010) 446–457.
[31] R.B. Merrifield, The synthesis of biologically active peptides and proteins, JAMA 210 (1969) 1247–1254.
[32] D. Seebach, M. Overhand, F.N. Kühnle, B. Martinoni, L. Oberer, U. Hommel, H. Widmer, β-Peptides: synthesis by Arndt-Eistert homologation with concomitant

peptide coupling. Structure determination by NMR and CD spectroscopy and by X-ray crystallography. Helical secondary structure of a β-hexapeptide in solution and its stability towards pepsin, Helv. Chim. Acta 79 (1996) 913–941.
[33] A. Santoro, M. Grimaldi, M. Buonocore, I. Stillitano, A.M. D'Ursi, Exploring the early stages of the amyloid aβ (1–42) peptide aggregation process: an NMR study, Pharmaceuticals 14 (2021) 732.
[34] G. Vitiello, M. Grimaldi, A.M. D'Ursi, G. D'Errico, The iAβ5p β-breaker peptide regulates the Aβ (25–35) interaction with lipid bilayers through a cholesterol-mediated mechanism, Biochem. Biophys. Res. Commun. 417 (2012) 88–92.
[35] R. Oliva, A. Emendato, G. Vitiello, A. De Santis, M. Grimaldi, A.M. D'Ursi, E. Busi, P. Del Vecchio, L. Petraccone, G. D'Errico, On the microscopic and mesoscopic perturbations of lipid bilayers upon interaction with the MPER domain of the HIV glycoprotein gp41, Biochim. Biophys. Acta, Biomembr. 1858 (2016) 1904–1913.
[36] A. Santoro, M. Grimaldi, M. Buonocore, I. Stillitano, A. Gloria, M. Santin, F. Bobba, M.S. Saponetti, E. Ciaglia, A.M. D'Ursi, New Aβ (1–42) ligands from anti-amyloid antibodies: design, synthesis, and structural interaction, Eur. J. Med. Chem. (2022), 114400.
[37] L. Whitmore, B. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (2004) W668–W673.
[38] K. Wüthrich, NMR with proteins and nucleic acids, EuroPhys. News 17 (1986) 11–13.
[39] T. Goddard, D. Kneller, SPARKY 3, University of California, San Francisco, 2004, p. 15.
[40] P. Guntert, Automated NMR structure calculation with CYANA, Methods Mol. Biol. 278 (2004) 353–378.
[41] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR 44 (2009) 213–223.
[42] E.G. Hutchinson, J.M. Thornton, PROMOTIF—a program to identify and analyze structural motifs in proteins, Protein Sci. 5 (1996) 212–220.
[43] S. Bienert, A. Waterhouse, T.A. de Beer, G. Tauriello, G. Studer, L. Bordoli, T. Schwede, The SWISS-MODEL Repository—new features and functionality, Nucleic Acids Res. 45 (2017) D313–D319.
[44] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303.
[45] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1 (2015) 19–25.
[46] H. Bekker, H. Berendsen, E. Dijkstra, S. Achterop, R. van Drunen, D. der Spoel, H. Bekker, E. Dijkstra, D. Van Der Spoel, A. Sijbers, Gromacs: A Parallel Computer for Molecular Dynamics Simulations, 1993.
[47] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 31 (2010) 671–690.
[48] R.B. Best, X. Zhu, J. Shim, P.E. Lopes, J. Mittal, M. Feig, A.D. MacKerell Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 dihedral angles, J. Chem. Theor. Comput. 8 (2012) 3257–3273.
[49] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, Biopolymers: Orig. Res. Biomol. 89 (2008) 392–400.
[50] T. Parella, P. Adell, F. Sánchez-Ferrando, A. Virgili, Effective multiple-solvent suppression scheme using the excitation sculpting principle, Magn. Reson. Chem. 36 (1998) 245–249.
[51] Schrödinger, Maestro, LLC, New York, NY, 2020.
[52] A. Onufriev, D.A. Case, D. Bashford, Effective Born radii in the generalized Born approximation: the importance of being perfect, J. Comput. Chem. 23 (2002) 1297–1304.
[53] P. Benkert, M. Biasini, T. Schwede, Toward the estimation of the absolute quality of individual protein structure models, Bioinformatics 27 (2011) 343–350.
[54] G. Studer, C. Rempfer, A.M. Waterhouse, R. Gumienny, J. Haas, T. Schwede, QMEANDisCo—distance constraints applied on model quality estimation, Bioinformatics 36 (2020) 1765–1771.