

# Enhancement of SARS-CoV-2 Receptor Binding Domain -CR3022 Human Antibody Binding Affinity via in Silico Engineering Approach

**Fateme Sefid**

Shahid Sadoughi University of Medical Sciences and Science and Art University

**Zahra Payandeh** (✉ [zpayandeh58@yahoo.com](mailto:zpayandeh58@yahoo.com))

Tabriz University of Medical Sciences

**Ghasem Azamirad**

Yazd University

**Behzad Mansoori**

Tabriz University of Medical Sciences and University of Southern Denmark

**Behzad Baradaran**

Tabriz University of Medical Sciences

**Mohammad Reza Sadeghi**


Tabriz University of Medical Sciences

---

## Research

**Keywords:** Coronavirus, SARS-COV-2, Antibody, bioinformatics, affinity maturation

**DOI:** <https://doi.org/10.21203/rs.3.rs-56411/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** The nCoV-2019 is a cause of COVID-19 disease. The surface spike glycoprotein (S), which is necessary for virus entry through the intervention of the host receptor and it mediates virus-host membrane fusion, is the primary coronavirus antigen (Ag). The angiotensin-converting enzyme 2 (ACE2) is reported to be the effective human receptor for SARS-CoVs 2. ACE2 receptor can be prevented by neutralizing antibodies (nAbs) such as CR3022 targeting the virus receptor-binding site. Considering the importance of computational docking, and affinity maturation we aimed to find the important amino acids of the CR3022 antibody (Ab). These amino acids were then replaced by other amino acids to improve Ab-binding affinity to a receptor-binding domain (RBD) of the 2019-nCoV spike protein. Finally, we measured the binding affinity of Ab variants to the Ag.

**Result:** Our findings disclosed that several variant mutations could successfully improve the characteristics of the Ab binding compared to the normal antibodies.

**Conclusion:** The modified antibodies may be possible candidates for stronger affinity binding to Ags which in turn can affect the specificity and sensitivity of antibodies.

## Introduction

Coronaviruses, a single-stranded RNA virus, generally classified into four major genera including Alpha, Beta, Gamma, and Delta coronavirus. Severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory disorder coronavirus (MERS-CoV) classified as  $\beta$ -coronaviruses and novel coronavirus SARS-CoV-2 belong to this family which caused coronavirus disease 2019 (COVID-19) disease. This virus has originally begun in China since 2019 and a global pandemic has been started in March 2020 (1, 2).

The key surface protein of this virus involves in entry into the host cells, is facilitated by envelope-anchored trimeric spike (S) protein. A potential human receptor for 2019-nCoV is the angiotensin-converting enzyme 2 (ACE2). This enzyme as a type I membrane protein is expressed on the membrane of different cells such as lung, kidney, intestine and heart cells and reduces the blood pressure, and it has a role in vasoconstriction. Furthermore, it has appeared that decreased ACE2 expression is related to cardiovascular diseases (3–6).

The virus enters into animal cells can be avoided by neutralizing antibodies (nAbs) focusing on the virus receptor sites. The novel coronavirus entry facilitated by, the spike (S) glycoprotein interaction by a receptor through the receptor-binding domain (RBD). RBD is a fine target for neutralizing the antibodies that can avoid virus entry by blocking the attachment of ACE2. Two partitioned groups of researchers recently reported this receptor is also involved in an entry of a new SARS-CoV-coronavirus to the human body. The receptor-binding domain (RBD) of 2019-nCoV spike protein has 6 amino acid substitutions compared to the RBD of SARS spike protein, so the ACE2 could be the receptor for 2019-nCoV (SARS-CoVs 2) entry (3, 4, 7). Recent studies identified a number of potent monoclonal antibodies against SARS coronavirus spike protein. Neutralize antibodies against SARS-CoV produced by experimental vaccination or monoclonal antibody (mAb) technology demonstrated in passive transfer experiments that protect mice from infection by reducing viral load (8–12).

Several SARS-CoV-specific antibodies have been expressed and purified by Tian et al. They reported CR3022 antibody (Ab) targeting on RBD has strong neutralizing activities. The binding ability was measured by ELISA and BLI. Notably, CR3022 Ab bound with 2019-nCoV RBD potently (13). Ter Meulen et al. reported that CR3022 could be developed as a candidate therapeutics Ab, alone or in combination with other neutralizing antibodies, for the prevention and treatment of 2019-nCoV infections (13, 14).

Regarding the formerly mentioned reasons about the advantages of CR3022 Ab and the impact of its indicate, affinity improvement on its practicability, biological action, and treatment costs, we aimed to, progress the binding affinity of neutralizing CR3022 Ab to RBD in SARS-CoVs 2 virus by protein engineering.

## Methods

### 2.1. CR3022 complementarity-determining Region Prediction

The Ab binding activity is principally medicated by the complementarity-determining region (CDR). The web server named Paratome(15) predicts the antigen-binding regions (ABRs) of an Ab, by submitting the amino acid sequence or protein 3D structure. Paratome was designed by the structural alignment of a non-redundant set of all established Ab-antigen (Ag) complexes in the PDB.

### 2.2. CR302 conservation of amino acid positions evolution

Estimating the evolutionary conservation of amino acid (AA) positions in protein macromolecule assessed by a bioinformatic tool named ConSurf server (16). The estimation of AA position performed based on the phylogenetic relations between homologous sequences. The degree of AA is evolutionary conservation (i.e., its evolutionary rate) is heavily dependent on its structural and functional value. Therefore, the importance of each AAs position for structure or function was assessed by conservation analysis of AAs position among the family members. The rate of evolutionary is calculated based on the evolutionary similarity between the protein and its homologues.

### 2.3. CR3022 Interfaces prediction

For the prediction of the protein-protein interface of a monomeric query protein applying its adjusted and remote structural neighbours, we used PredUs (17) web server, which is a flexible, interactive, template-based tool. The other tool which we used for predicting protein-protein interaction sites was cons-PPISP

(consensus Protein-Protein Interaction Site Predictor) (18) which is a consensus neural network. It predicts the residues that will probably form the binding domain for another protein.

The solvent accessibility and sites deleterious mutations were predicted by WESA (<https://pipe.rcc.fsu.edu/wesa/>) with an accuracy of 80% (19). The decisive prediction for each residue is based upon a weighted total of the specific predictions. A residue is predicted as either a hidden or exposed (have got a surface area bigger than 20% of the maximum area for that kind of AA).

The novel method for prediction of partner-specific protein interfaces from pdb files or input sequences was assessed by Extreme Gradient Boosting (xgBoost) (20). This method is based on Interface Prediction of Specific Partner Interactions (BIPSPI) and available at <http://bipspi.cnb.csic.es/xgbPredApp/>. In this model.

#### 2.4. CR3022 binding sites and Pocket detection

To identifying multi-scale pockets on the surface of the protein, we used mathematical morphology via the GHECOM (Grid-based HECOMi finder) program (21).

The protein-protein interaction site predicted by Meta-PPISP at <https://pipe.rcc.fsu.edu/meta-ppisp.html>(22, 23). Several matching methods have been established for calculating protein-protein interaction sites. Meta-PPISP sought to enhance prediction reliability and accuracy by integrating results from separate predictors, and report here a Meta web server, meta-PPISP, that is developed on three separate web servers including Promate, cons-PPISP, and PINUP. A linear regression statistical method, employing the raw scores of the three servers as input, has been trained on a set of 35 nonhomologous proteins. Cross-authentication showed that meta-PPISP outperforms all three separate servers.

#### 2.5. Significant residues selection

Some AAs selected as major residues in CR3022 structure by utilizing the outcome of different software. These residues sited in one of three CDR regions were estimated by the Paratome server.

Chosen residues in BIPSPI software which have a score above 0.5, have a score above 0.00 in Cons PPISP software and above 4 scores in GHecom software. In connection with this, PredUS, Meta-PPISP and WESA predicted residues exploration to pick the major AAs.

#### 2.6. SIFT analyses

To find if an AA replacement affects protein function, we used the SIFT server (<http://sift.jcvi.org/>). The server predicts the degree of conservation of AA residues in sequence alignments originated from closely linked to sequences gathered through PSI-BLAST. Also, it could be applied to naturally-occurring nonsynonymous polymorphisms or missense mutations induced in the laboratory (24).

#### 2.7. CR3022variants sketching

71 variants containing mutations in at the very least one of three ABRs were designed. The three-dimensional structure of all submitted variants was decided by SAbPred (25) (Antibody simulation tool and prediction application from the Oxford Protein Informatics Group (OPIG)).

#### 2.8. Molecular docking analyses

The three-dimensional structure of each variant and bap component provides input for High Ambiguity Driven protein-protein DOCKing (HADDOCK) (26) HADDOCK is an information-driven flexible docking method for the simulation of biomolecular complexes

## Results

### 3.1. CR3022CDR prediction

Paratome is a bioinformatic tool for the recognition of ABRs in antibodies This server predicts 3 regions as ABRs in CR3022 heavy chain and 3 regions as ABRs in CR3022 light chain. These regions include YGFITYWIG (27-35) as ABR1 and WMGLIYPGDSETRY (47-60) and GGSGISTPMDV (98-108) as ABR2 and 3 respectively in CR3022 heavy chain and QSVLYSSINKNYLA (27-40), LLIYWASTRES (52-62) and QQYYSTPY (95-102) as ABR1, ABR2 and ABR3 respectively in CR3022 light chain. Paratome results are displayed in Table 1.

### 3.2. CR302 conservation of amino acid positions evolution

The nine-color conservation scores are being projected onto the three-dimensional structure of the Ab and coloured protein structure is displayed by FirstGlance in Jmol (Figure 1). The normalized score calculated for each amino acid position calculated by the ConSurf server. The color scale represented by the conservation scores (9 - conserved, 1 - variable) is shown in Table 1.

Light chain    Heavy chain

**Figure 1: CR302 conservation of amino acid positions evolution by the ConSurf server.** conservation scores projected onto the three-dimensional structure of the Ab with nine colours. The schematic structure of the coloured protein displayed by FirstGlance in Jmol.

**Table 1: Paratome, predUS, cons-PPISP, GHECOM, WESA and ConSurf server Predictions**

Chain: Residue	H:55	H:56	H:57	H:58	H:59	H:60	H:98	H:99	H:100	H:101	H:102	H:103	H:104	H:105	F
Paratome	<b>D</b>	<b>S</b>	<b>E</b>	<b>T</b>	<b>R</b>	<b>Y</b>	<b>G</b>	<b>G</b>	<b>S</b>	<b>G</b>	<b>I</b>	<b>S</b>	<b>T</b>	<b>P</b>	<b>M</b>
Pred us	D	S	E	T	R	Y	G	G	S	G	I	S	P	M	T
Cons ppisp	0.073	0.084	0.084	0.067	0.065	0.09	0	0	0.069	0.044	0.044	0.036	0.04	0	0
Consurf	1	1	1	3	1	8	4	1	1	1	1	1	1	1	1
Ghecom	0.43	0	0	0	1.26	0.02	0	0.1	0.89	0.46	1.61	4.2	3.6	0	0
WESA	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
BIPSPI	0.4305	0.3976	0.2848	0.3277	0.2426	0.1917	0.6202	0.5643	0.4652	0.2215	0.2111	nan	nan	nan	nan
Meta-PPISP	N	N	N	N	N	N	-	-	P	P	N	P	P	-	-

Chain: Residue	L:27	L:28	L:29	L:30	L:31	L:32	L:33	L:34	L:35	L:36	L:37	L:38	L:39	L:40	L:52	I
Paratome	<b>Q</b>	<b>S</b>	<b>V</b>	<b>L</b>	<b>Y</b>	<b>S</b>	<b>S</b>	<b>I</b>	<b>N</b>	<b>K</b>	<b>N</b>	<b>Y</b>	<b>L</b>	<b>A</b>	<b>L</b>	<b>I</b>
Pred us	Q	Y	V	S	L	S	S	I	N	K	N	Y	L	A	L	I
Cons ppisp	0.02	0.077	0	0.065	0.063	0.058	0.048	0.059	0.072	0.05	0.051	0.039	0	0	0	0
Consurf	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
Ghecom	0.15	0	0.29	1.69	1.07	0	0	0.09	2.41	1.92	3.42	1.53	0.02	0	4.81	0
WESA	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0
BIPSPI	0.2704	0.6071	0.4767	0.4071	0.2915	0.3562	0.2482	nan	nan	nan	nan	nan	nan	nan	0.2368	0
Meta-PPISP	N	N	-	P	P	P	P	P	P	P	P	P	-	-	-	-

Chain: Residue	H:27	H:28	H:29	H:30	H:31	H:32	H:33	H:34	H:35	H:47	H:48	H:49	H:50	H:51	H:52
Paratome	<b>Y</b>	<b>G</b>	<b>F</b>	<b>I</b>	<b>T</b>	<b>Y</b>	<b>W</b>	<b>I</b>	<b>G</b>	<b>W</b>	<b>M</b>	<b>G</b>	<b>I</b>	<b>I</b>	<b>Y</b>
Pred us	Y	G	F	I	T	Y	W	I	G	W	M	G	I	I	Y
Cons ppisp	0.14	0.124	0	0.081	0.112	0.099	0.068	0	0	0	0	0	0	0	0.063
Consurf	5	3	6	4	1	1	1	4	1	8	5	3	1	7	1
Ghecom	0.37	0.62	2.76	2.99	0.88	1.32	0	0	0	3.08	1.69	0	0.02	0	0.82
WESA	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
BIPSPI	0.417	0.5276	0.5765	0.6946	1	0.6402	0.6246	0.4596	nan	nan	nan	nan	0.2091	0.3352	0.5082
Meta-PPISP	P	P	-	P	P	P	P	-	-	-	-	-	-	-	P

Chain: Residue	L:55	L:56	L:57	L:58	L:59	L:60	L:61	L:62	L:95	L:96	L:97	L:98	L:99	L:100	L:101
Paratome	<b>Y</b>	<b>W</b>	<b>A</b>	<b>S</b>	<b>T</b>	<b>R</b>	<b>E</b>	<b>S</b>	<b>Q</b>	<b>Q</b>	<b>Y</b>	<b>Y</b>	<b>S</b>	<b>T</b>	<b>P</b>
Pred us	Y	W	A	S	T	R	E	S	Q	Q	Y	Y	S	T	P
Cons ppisp	0.053	0.044	0	0.034	0.056	0.062	0.061	0.086	0	0	0	0.039	0.047	0.063	0.098
Consurf	3	1	1	1	1	5	1	7	1	1	1	1	1	1	1
Ghecom	1.94	3.06	1.25	1.41	0.15	0.04	6.6	3.78	0	0.06	2.31	1.44	2.03	1.86	2.82
WESA	0	1	0	1	1	1	1	1	0	0	0	1	1	1	0
BIPSPI	0.2802	0.2489	0.1975	nan	nan	nan	nan	nan	0.2459	0.2307	0.1949	nan	nan	nan	nan
Meta-PPISP	P	P	-	P	P	N	N	N	-	-	-	P	N	N	N

### 3.3. CR3022 Interfaces prediction

Potential interfacial residues identified through PredUs are presented in table 1. Residue 33W in ABRs I of a heavy chain (H chain ABR I), residue W47, I50, R59 and Y60 in ABRs II of a heavy chain (H chain ABR II), residue I102, S103 and D107 in ABRs III of a heavy chain (H chain ABR III), Residue Q27 in ABRs I of a light chain (L chain ABR I), residue L52, Y55, W56 and S58-E61 in ABRs II of a light chain (L chain ABR II) and residue Y97-S99, P101 and Y102 in ABRs III of a light chain predicted as a possible interfacial residue.

cons-PPISP calculates a score of neural network for every residue. This score estimated if a residue involved in the protein-protein interaction location or not. The interred and not predicted residues presume as 0 score. The score above 0 considers as an interaction residue and below 0 score non-interaction residue (Table 1).

WESA a Weighted Ensemble Solvent Accessibility predictor identified a number of AAs as Solvent Accessible residues. WESA entire results are shown in Table 1. WESA computes six different scores for each residue. These scores include Bayesian Statistics (BS), Multiple Linear Regression (MLR), Decision Tree (DT), Neural Network (NN), Support Vector Machine (SVM) and Weighted Ensemble (WE). The WE score predicts whether a residue is exposed (=1) or Buried (=0).

BIPSPI can be used to predict partner-specific protein-protein interfaces by utilizing sequences or structural patterns. An individual threshold for antigen (Ag) and Ab predictions can be set. These thresholds are suitable in order to examine various expected accuracy/recall values. A prediction of the likely accuracy at the set of thresholds is presented as a threshold. Residues whose score predicted accuracy is more or equal than the precision threshold (0.500) are highlighted in green. The interactive visualization of predicted residues in the Ab structure is shown in Figure 2. The predicted interface scores for the residues of Ab is displayed in Table 1

#### Figure 2. BIPSPI Interactive visualization of predicted residues in the antibody structure.

Residues whose score has an expected precision greater or equal than the precision threshold (0.500) are highlighted in green. Interface Residues Prediction by BIPSPI whose score has an expected precision greater or equal than the precision threshold (0.500) are listed below the picture.

### 3.4. CR3022 binding sites and Pocket detection

GHECOM server discovers five pockets on protein shells by utilizing mathematical morphology. In this respect, it calculates a pockets score (sum of  $1/[R_{pocket}] / (1/[R_{min}] * [vol \text{ of shell}])$ ) for each residue. A residue in a deeper and bigger pocket has a greater value of pockets. The pockets of small-molecule binding and active locations were greater than the be an average of value; in particular, the values for the active locations were much greater. This implies that pockets are contributing to the prediction of binding and active sites from protein structures (Figure 3) (Table 1)[24]. Residue H:S103, H:D107, H:V108, L:L52 and L:E61 have GHECOM score above 4.

Meta-PPISP is a metamethod for Protein-Protein Interaction Site Prediction. Meta-PPISP is created on three individual web servers including PINUP, cons-PPISP, and Promate. The interface residues which predicted are shown as residue ID: chain ID)

1:H, 3:H, 25:H, 26:H, 27:H, 28:H, 30:H, 31:H, 32:H, 33:H, 52:H, 53:H, 73:H, 76:H, 96:H, 97:H, 99:H, 100:H, 27C:L, 27D:L, 27E:L, 27F:L, 28:L, 29:L, 30:L, 31:L, 32:L, 49:L, 50:L, 52:L, 53:L, 92:L

Meta-PPISP calculated cons-PPISP, PINUP, Promate, and meta-PPISP scores.

In this regard Meta-PPISP was Predicted whether the residue is in an interface (P = Positive; N = Negative; - = Buried and not predicted) Note: P corresponds to a score > 0.34 in meta-PPISP score. Meta-PPISP entire results are shown in Table 1.

#### Figure 3: GHECOM results showing Jmol view of a pocket structure. Jmol view of pocket structure based on pockets colour.

#### Figure 4. Illustration of mutations sequences

### 3.5. Significant residues selection

We select H:30I, H:31T, H:33W, H:54G, H:103S, L:56W, L:58S, L:59T, L:61E and L:98Y residues by using the results of various software. These residues situated in one of three CDR regions were predicted by the Paratome. The specially selected residues confirm by at least four software. Cons PPISP score above 0.00, the scores above 0.5 and 4 for BIPSPI and GHecom were the considered threshold, respectively.

In this regard, PredUS, Meta-PPISP and WESA predicted residues research to select the significant AAs.

Residue H:30I was predicted by Cons PPISP, WESA, BIPSPI and Meta-PPISP servers. Residue H:31T was predicted by Cons PPISP, WESA, BIPSPI and Meta-PPISP servers. Residue H:33W was predicted by Cons PPISP, PredUs, BIPSPI and Meta-PPISP servers. Residue H:54G was predicted by Cons PPISP, WESA, BIPSPI and Meta-PPISP servers. Residue H:103S was predicted by Cons PPISP, PredUs, Ghecom, WESA and Meta-PPISP servers. Residue L:56W was predicted by Cons PPISP, PredUs, WESA and Meta-PPISP servers. Residue L:58S was predicted by Cons PPISP, PredUs, WESA and Meta-PPISP servers. Residue L:59T was predicted by Cons PPISP, PredUs, WESA and Meta-PPISP servers. Residue L:61E was predicted by Cons PPISP, PredUs, WESA and Ghecom servers. Residue L:98Y was predicted by Cons PPISP, PredUs, WESA and Meta-PPISP servers.

### 3.6. SIFT analyses

Sorts intolerant from tolerant (SIFT) is a sequence homology-based program that amino acid switches and predicts this change in protein will have a phenotypic impact. SIFT assumes that protein development is associated with protein function. Positions that are essential for the function should be conserved in an alignment of the protein family, although insignificant positions should be displayed diverse in an alignment.

SIFT brings a query sequence and utilizes multiple alignment information that will predict tolerated and detrimental alterations in each position of the query sequence. SIFT is a multistep process that (1) searches for analogous sequences, (2) selects closely linked sequences that may share parallel function to the query sequence, (3) obtains the alignment of these chosen sequences, and (4) computes normalized possibilities for all changes from the alignment. Positions with normalized possibilities lower than 0.05 are expected to be deleterious; those higher than or equal to 0.05 are expected to be tolerated (Table 2).

**Table 2. SIFT results for selected residues**

Innate features of Amino acid (AA) showed in different colours including black: nonpolar, green: uncharged polar, red: basic, blue: acidic. Initial capitals imply AAs showing in the alignment, lower case letters arise from prediction.

Position	Seq Rep	Predict Tolerated
H:30I	0.89	H p l V R q N D G l A k E T S
H:31T	0.89	h i v p l q A K R E G N D T S
H:33W	0.89	C p m D E q W K N R G i T S V A H L F Y
H:54G	0.85	c p M W Q E i R K T D A H F V L N S Y G
H:55D	0.84	m h F i P L V r q T A k E G S N D
H:57E	0.81	c F Y m h l P V L G N R Q T D S A K E
H:102I	0.56	c W P D m E K q N G R T S l A V H L F Y
H:103S	0.65	w C m P D l q N G r K h E T V F L S A Y
L:56W	1.00	c p m D Q N W K i E G R T S V H A L F Y
L:58S	1.00	N S
L:59T	0.99	f c Y m h l p v L g D Q N A S R K E T
L:61E	0.98	l G T V S K Q H D Y P A E
L:98Y	0.93	c W p D m E K Q N G R l T S V A H L F Y

Figure 5. HADDOCK Score of all variants. The variants with scores below the threshold (control score is -138.4) are predicted to have an enhanced affinity toward the CR3022 human Ab. The 5-best mutated CR3022 human Ab variants coloured in red on the graph.

### 3.7. CR3022variants sketching

71 variants including mutations in at least one of 3 ABRs were offered. H:30I, H:31T, H:33W, H:54G, H:103S, L:56W, L:58S, L:59T, L:61E and L:98Y residues which were established by various programs mutated in proposed variants at random. In addition H:30I, H:33W, H:57E, H:55D, H:102I and L:56W were introduced by Meng Yuan et al. (27) as interactive residues in Crystal structure of CR3022 in complex with SARS-CoV-2 RBD. In this regard, we also mutated H:57E, H:55D and H:102I to access a more diverse variant. Mutation sequences were aligned and illustrated in Figure 4.

### 3.8. protein-protein docking is based upon biochemical or biophysical data

HADDOCK server evaluates ligand and receptor integration based on biochemical and/or biophysical data. Table 3 represents the information of variants with the HADDOCK score is more than the control. The Van der Waals and electrostatic energy values, in addition to the interred surface between the two complexes, are shown.

**Table 3- Docking between the normal and 5 best mutated CR3022 human antibody variants with RBD Ag**

Control	variant45	variant60	variant67	variant69	variant71	
HADDOCK score	-138.4 +/- 1.6	-167.3 +/- 3.2	-167.5 +/- 6.5	-161.6 +/- 4.0	-173.0 +/- 6.0	-169.8 +/- 7.0
Cluster size	71	40	47	48	32	39
RMSD from the overall lowest-energy structure	1.0 +/- 0.9	3.2 +/- 0.3	1.2 +/- 1.0	0.9 +/- 0.5	1.8 +/- 1.1	3.3 +/- 0.0
Van der Waals energy	-54.5 +/- 8.7	-68.9 +/- 5.6	-64.4 +/- 0.8	-80.1 +/- 5.3	-63.8 +/- 10.4	-63.3 +/- 4.1
Electrostatic energy	-341.7 +/- 53.7	-338.4 +/- 12.1	-466.5 +/- 19.9	-247.3 +/- 27.7	-431.1 +/- 39.3	-435.3 +/- 43.7
Desolvation energy	-21.9 +/- 6.9	-35.6 +/- 3.3	-14.5 +/- 3.2	-38.8 +/- 6.7	-27.5 +/- 7.5	-21.6 +/- 4.6
Restraints violation energy	63.7 +/- 15.68	49.1 +/- 18.23	46.7 +/- 28.23	67.1 +/- 16.10	45.4 +/- 36.81	22.3 +/- 20.17
Buried Surface Area	1809.7 +/- 119.0	2242.2 +/- 46.5	2221.8 +/- 77.1	2335.3 +/- 104.3	2136.4 +/- 99.7	2165.1 +/- 105.8
Z-Score	-1.7	-1.5	-1.2	-2.2	-1.7	-1.6

The categorization of complex structures is based on HADDOCK scores. HADDOCK provides a fully flexible scoring plan since the weight of the different power terms can be determined separately for each stage of the docking. The scoring is achieved according to the weighted sum (HADDOCK score) of the following terms: such as electrostatic, van der Waals, a radius of gyration restraint, distance restraints, intervention projection angle restraints, direct RDC restraint, pseudo contact shift restraint, dihedral angle restraints, diffusion anisotropy, symmetry restraints, binding and desolvation energy. The structure with the lowest weighted sum will be classified top.

## Discussion

Protein engineering is a potent method for developing ideal therapeutic proteins (28). Contemporary bioinformatics techniques are commonly used in protein engineering (29, 30). In the case of diseases like COVID-19 where time is important, these tools could save a lot of time and effort. To entry virus into the cells, binding between ACE2 and RBD in spike protein is essential. So, many studies try to inhibit the binding of them.

Recently Changhai et al. published a recombinant ACE2-Ig with possible applications for the 2019-nCoV diagnosis, prophylaxis and treatment (31). Other researchers have also produced an engineered ACE2 protein using bioinformatics techniques that are expected to bind RBD with higher affinity while being more thermostable and lacking enzymatic activity (32).

In very promising research, it was shown that human Ab CR3022 binds to SARS-CoV-2 RBD in a cryptic epitope (13). In addition, we found no studies that directed the design of these antibodies against different epitopes of SARS-CoV-2 proteins, such as targeting the spike protein and subsequently preventing its binding to human ACE2. Since passive immunization with polyclonal antibodies has been reported to curb outbreaks of the hepatitis A virus and prevent varicella-zoster infection (33), mAb prophylaxis can be an important way to contain an outbreak of SARS. This computational research aimed to advance the binding affinity of CR3022 Ab neutralization to ACE2 (RBD) in SARS-CoVs 2 virus by focusing on AA mutagenesis. We believed that affinity maturation increases the performance of diagnostic antibodies due to enhanced specificity at reduced Ab concentrations. We used site-directed mutagenesis approaches for this order. Site-directed mutagenesis (SDM) methods are used to produce cloned DNAs with changed sequences to investigate the significance of different residues in the structure and function of the proteins (34, 35). Our first goal is to find structurally and functionary important AAs. We ought to identify sites in the complementarity-determining regions (CDRs) that are permissive to mutagenesis while maintaining Ag binding. Next, we selected several residues by various methods. First, we predict antigen-binding regions of CR3022 Ab using the Paratome web server, second using ConSurf, PredUs, Meta-PPISP to predict protein-protein interaction sites. Third applying new method BIPSPI and WESA predicted partner-specific protein interfaces. BIPSPI predictions from two input structural models (Ab and Ag) were performed using PREDICT from structural data option. By employing the results of different software, we selected H:30I, H:31T, H:33W, H:54G, H:103S, L:56W, L:58S, L:59T, L:61E and L:98Y residues. Based on the prediction by Paratome web server, these residues placed in one of CDR regions. The specially selected residues were confirmed by at least four software. Cons PPISP score above 0.00, BIPSPI score above 0.5 and GHecom score above 4 were the considered threshold. In this regard PredUS, Meta-PPISP and WESA predicted residues to choose the major AAs. Finally, the SIFT server were used to predict whether an AA switch interferes with the protein act.

Various studies have been designed using this method and experimental results were confirmed (29, 30, 36).

Based on these results, we substituted AAs and then did docking, which plays a significant role in targeted medicine designing since this is the only theoretical method that explicitly models physical interactions between proteins. To this end, HADDOCK has been added. HADDOCK software has also been developed enabling versatility of side chains and/or backbone.

HADDOCK 2.1 does the docking of variants CR3022 Ab and SARS-CoV 2 Spike Receptor-Binding Domain, totally flexible, utilizing molecular dynamics simulations. Docking findings for both normal and mutant antibodies indicate that engineered mutations have changed and strengthened the binding and energy properties in mutated antibodies to the Ags. One of the criteria for the comparison of docking results is the amount of energy ( $\Delta G$ ). The smaller energy complexes have greater stability and greater binding affinity.

Based on the results, variants 45,60,67,69 and 71 antibodies have smaller bonding energy and their stability has also been increased compared to the wild type. Based on obtained scores which are -138.4 for control and mutated variants: Variant 45: -167.3, Variant 60: -167.5, Variant 67: -161.6, Variant 69: -173.0

and Variant 71: -169.8 have more affinity to bind to SARS-CoV2 Spike Receptor-Binding Domain than the natural Ab. The buried area between the two complexes in these variants was more positive than the natural state. This means that these mutations have modified and increased the Ab binding properties relative to normal antibodies.

## Conclusion

These results suggest that the CR3022 Ab could potentially be developed as a candidate therapeutic alone or in combination with other neutralizing antibodies for the prevention or treatment of nCoV-2019 infections. Optimization of the CR3022 properties to boost their binding characteristics and maturation of affinity to avoid cross-reactivity and selection of suitable antibodies by adding new mutations to create the desired Ab is advantageous.

## Abbreviations

COVID-19

Coronavirus disease-19; nAbs:neutralizing antibodies, mAb:monoclonal antibody, ABRs:antigen-binding regions.SARS-CoV-2:Severe acute respiratory syndrome coronavirus 2; 2019-nCoV:Novel coronavirus; ACE:Angiotensin converting enzyme; MERS-CoV:Middle East Respiratory Syndrome Coronavirus; RBD:Receptor-binding domain; ACE2:Angiotensin-converting enzyme 2

## Declarations

### Acknowledgments

The authors wish to thank Immunology Research Center, Tabriz- Iran and Shahid Sadoughi University of Medical Science, Yazd, Iran for their support by unrestricted free access to the web site for data collection.

### Authors' contributions:

All authors participated in writing. Writing-review and language correction.

### Funding:

This work was financially supported by grants from the Iran National Science Foundation (INSF), Iran (project no. 97021150) and Tabriz University of Medical Sciences, Tabriz, Iran (project no. 61892). We gratefully acknowledge them for their contribution to this study.

### Availability of data and materials:

Data presented in this manuscript is available upon request.

### Ethics approval and consent to participate:

This article does not contain any studies with human participants or animals performed by any of the authors

### Consent for publication:

All authors have read and approved the final version of the manuscript.

### Conflict of interests

All authors declare no potential conflicts of interest.

### Competing interests:

The authors declare that they have no competing interests.

### Author details:

1- Department of Medical Genetics, Shahid Sadoughi University of Medical Science, Yazd, Iran

2- Department of biology, Science and Art University, Yazd, Iran

3- Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

4- School of Mechanical Engineering, Yazd University, Yazd, Iran

5- Department of Cancer and Inflammation Research, Institute for Molecular Medicine, University of Southern Denmark, 5000, Odense, Denmark

6- Department of Molecular Medicine, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

## References



1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. 2020;395(10223):497–506.
2. of the International CSG. (2020) The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*.1.
3. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, et al. A novel angiotensin-converting enzyme–related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1–9. *Circulation research*. 2000;87(5):e1–9.
4. Boehm M, Nabel EG. Angiotensin-converting enzyme 2—a new cardiac regulator. *N Engl J Med*. 2002;347(22):1795–7.
5. Zhang H, Kang Z, Gong H, Xu D, Wang J, Li Z, et al. (2020) The digestive system is a potential route of 2019-nCoV infection: a bioinformatics analysis based on single-cell transcriptomes. *BioRxiv*.
6. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. (2020) Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCoV. *BioRxiv*.
7. Ge X-Y, Li J-L, Yang X-L, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*. 2013;503(7477):535–8.
8. Der Sarkissian S, Grobe JL, Yuan L, Narielwala DR, Walter GA, Katovich MJ, et al. Cardiac overexpression of angiotensin converting enzyme 2 protects the heart from ischemia-induced pathophysiology. *Hypertension*. 2008;51(3):712–8.
9. Coleman CM, Liu YV, Mu H, Taylor JK, Massare M, Flyer DC, et al. Purified coronavirus spike protein nanoparticles induce coronavirus neutralizing antibodies in mice. *Vaccine*. 2014;32(26):3169–74.
10. Chen W-H, Chag SM, Poongavanam MV, Biter AB, Ewere EA, Rezende W, et al. Optimization of the production process and characterization of the yeast-expressed SARS-CoV recombinant receptor-binding domain (RBD219-N1), a SARS vaccine candidate. *Journal of pharmaceutical sciences*. 2017;106(8):1961–70.
11. Pang J, Wang MX, Ang IYH, Tan SHX, Lewis RF, Chen JI-P, et al. Potential rapid diagnostics, vaccine and therapeutics for 2019 novel coronavirus (2019-nCoV): a systematic review. *Journal of clinical medicine*. 2020;9(3):623.
12. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. (2020) Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host & Microbe*.
13. Tian X, Li C, Huang A, Xia S, Lu S, Shi Z, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerging Microbes Infections*. 2020;9(1):382–5.
14. Ter Meulen J, Van Den Brink EN, Poon LL, Marissen WE, Leung CS, Cox F, et al. (2006) Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS medicine*.3(7).
15. Kunik V, Ashkenazi S, Ofra Y. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. *Nucleic acids research*. 2012;40(W1):W521-W4.
16. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research*. 2016;44(W1):W344-W50.
17. Zhang QC, Deng L, Fisher M, Guan J, Honig B, Petrey D. PredUs: a web server for predicting protein interfaces using structural neighbors. *Nucleic acids research*. 2011;39(suppl\_2):W283-W7.
18. Chen H, Zhou HX. Prediction of interface residues in protein–protein complexes by a consensus neural network method: test against NMR data. *Proteins: Struct Funct Bioinf*. 2005;61(1):21–35.
19. Fernández-Recio J. (2011) Prediction of protein binding sites and hot spots. *Wiley Interdisciplinary Reviews: Computational Molecular Science*.1(5):680 – 98.
20. Sanchez-Garcia R, Sorzano COS, Carazo JM, Segura J. BIPSPi: a method for the prediction of partner-specific protein–protein interfaces. *Bioinformatics*. 2019;35(3):470–7.
21. Kawabata T. Detection of multiscale pockets on protein surfaces using mathematical morphology. *Proteins: Struct Funct Bioinf*. 2010;78(5):1195–211.
22. Qin S, Zhou H-X. meta-PPISP: a meta web server for protein-protein interaction site prediction. *Bioinformatics*. 2007;23(24):3386–7.
23. Zhou H-X, Qin S. Interaction-site prediction for protein complexes: a critical assessment. *Bioinformatics*. 2007;23(17):2203–9.
24. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*. 2009;4(7):1073.
25. Dunbar J, Krawczyk K, Leem J, Marks C, Nowak J, Regep C, et al. SAbPred: a structure-based antibody prediction server. *Nucleic acids research*. 2016;44(W1):W474-W8.
26. De Vries SJ, Van Dijk M, Bonvin AM. The HADDOCK web server for data-driven biomolecular docking. *Nature protocols*. 2010;5(5):883.
27. Yuan M, Wu NC, Zhu X, Lee C-CD, So RT, Lv H, et al. (2020) A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. *Science*.
28. Jones DS, Silverman AP, Cochran JR. Developing therapeutic proteins by engineering ligand–receptor interactions. *Trends in biotechnology*. 2008;26(9):498–505.
29. Payandeh Z, Rajabibazl M, Mortazavi Y, Rahimpour A, Taramchi AH. Ofatumumab monoclonal antibody affinity maturation through in silico modeling. *Iranian biomedical journal*. 2018;22(3):180.
30. Payandeh Z, Rajabibazl M, Mortazavi Y, Rahimpour A, Taramchi AH, Dastmalchi S. Affinity maturation and characterization of the ofatumumab monoclonal antibody. *Journal of cellular biochemistry*. 2019;120(1):940–50.

31. Lei C, Fu W, Qian K, Li T, Zhang S, Ding M, et al. (2020) Potent neutralization of 2019 novel coronavirus by recombinant ACE2-Ig. *Biorxiv*.
32. Alhenc-Gelas F, Druke TB. (2020) Blockade of SARS-CoV-2 infection by recombinant soluble ACE2. *Kidney International*.
33. Keller MA, Stiehm ER. Passive immunity in prevention and treatment of infectious diseases. *Clin Microbiol Rev*. 2000;13(4):602–14.
34. Antikainen NM, Martin SF. Altering protein specificity: techniques and applications. *Bioorg Med Chem*. 2005;13(8):2701–16.
35. Edelheit O, Hanukoglu A, Hanukoglu I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol*. 2009;9(1):61.
36. Eyvazi S, Payandeh Z, Khalili S, Sefid F, Ouladsahebmadarek E. Bevacizumab Antibody Affinity Maturation to Improve Ovarian Cancer Immunotherapy: In Silico Approach. *Int J Pept Res Ther*. 2019;25(4):1417–30.

## Figures

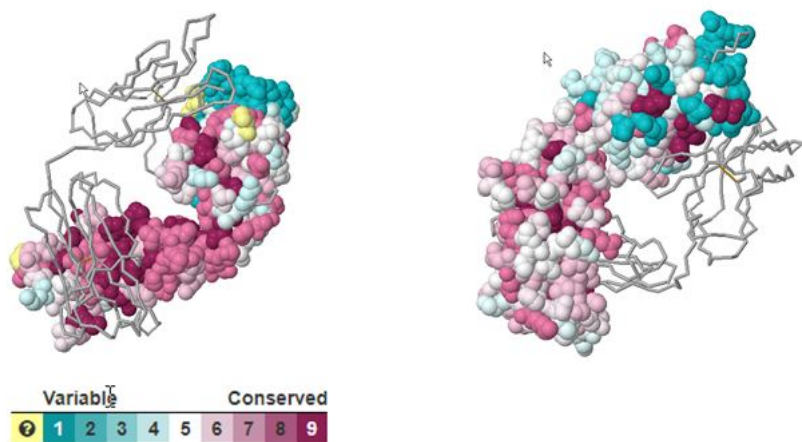
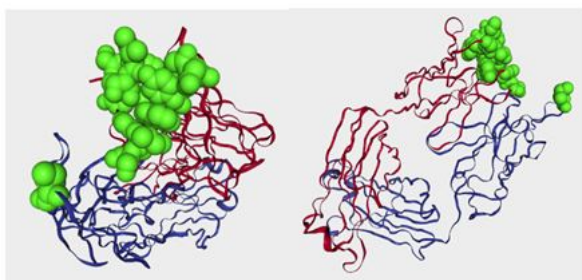


Figure 1

CR302 conservation of amino acid positions evolution by the Consurf server. conservation scores projected onto the three-dimensional structure of the Ab with nine colours. The schematic structure of the coloured protein displayed by FirstGlance in Jmol.



Chain	Residue	Score	Expected precision
H	THR31	6.0353	1
H	ILE30	2.1479	0.6946
H	GLY54	2.0132	0.6796
H	TYR32	1.0832	0.6402
H	TRP33	0.9694	0.6246
H	ASP55	0.7617	0.6202
H	ILE102	0.7564	0.6202
L	ILE34	0.6683	0.6071
H	PHE29	0.4376	0.5765
H	GLY101	0.4356	0.5765
H	SER103	0.3547	0.5643
H	PRO53	0.2697	0.5403
H	GLY28	0.2412	0.5276
H	TYR52	0.1972	0.5082

Figure 2

BIPSPI Interactive visualization of predicted residues in the antibody structure. Residues whose score has an expected precision greater or equal than the precision threshold (0.500) are highlighted in green. Interface Residues Prediction by BIPSPI whose score has an expected precision greater or equal than the precision threshold (0.500) are listed below the picture.

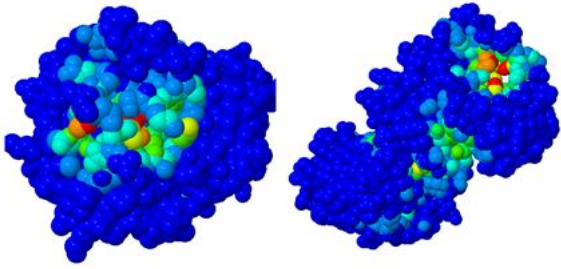


Figure 3

GHECOM results showing Jmol view of a pocket structure. Jmol view of pocket structure based on pockets colour.

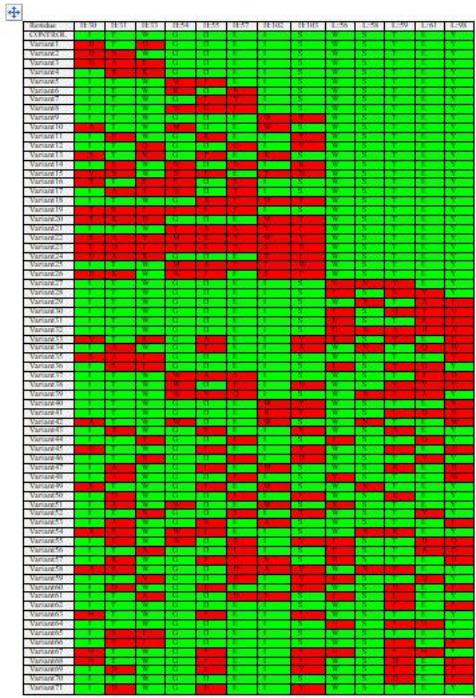


Figure 4

Illustration of mutations sequences

L:56W	1.00	c p m D Q N W K i E G R T S V H A L F Y
L:58S	1.00	N S
L:59T	0.99	f c Y m h l p v L g D Q N A S R K E T
L:61E	0.98	l G T V S K Q H D Y P A E
L:98Y	0.93	c W p D m E K Q N G R I T S V A H L F Y

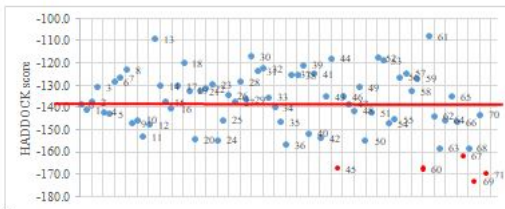


Figure 5

HADDOCK Score of all variants. The variants with scores below the threshold (control score is -138.4) are predicted to have an enhanced affinity toward the CR3022 human Ab. The 5-best mutated CR3022 human Ab variants coloured in red on the graph.