I. ADDITIONAL FIGURES AND TABLES

| Complex | Dimer | Structure | Complex | Dimer | Structure | Complex | Class | Structure |
|---------|-------|-----------|---------|-------|-----------|---------|-------|-----------|
| 1h89    | SBR-hom | SS       | 3b33    | SBR-hom | HH       | 4s5h    | SBR-hom | HH       |
| 1aa8    | SBR-hom | HH       | 3axv    | SBR-hom | HH       | 4ycl    | SBR-hom | HH       |
| 1a09    | nIBR-hom | HH      | 3j9u    | nIBR-het | HH      | 4mck    | nIBR-hom | SS       |
| 1amu    | HH       | HH       | 3aju    | nIBR-hom | HH      | 4jcy    | nIBR-hom | HH       |
| 1arp    | HH       | HH       | 3t3u    | IBR-hom | HH      | 4bic    | nIBR-het | HH       |
| 1a8f    | HH       | HH       | 3pvi    | SBR-hom | HH       | 4y9u    | nIBR-het | HH       |
| 1a8j    | HH       | HH       | 3b5h    | HH       | HH       | 4u4n    | SBR-het | HH       |
| 1a05    | HH       | HH       | 3wpe    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1b75    | HH       | HH       | 3jkl    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3h    | HH       | HH       | 3huy    | HH       | HH       | 4w3h    | SBR-het | HH       |
| 1h0q    | HH       | HH       | 3w7g    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1blb    | HH       | HH       | 3y6b    | HH       | HH       | 4y9u    | nIBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1m16    | HH       | HH       | 3x9z    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1i3q    | HH       | HH       | 3y6b    | HH       | HH       | 4y7s    | SBR-het | HH       |
| 1j3l    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |
| 1jps    | HH       | HH       | 3n7x    | HH       | HH       | 4v8h    | SBR-het | HH       |

TABLE I: ‘Human’ dataset. For each of the 199 considered complexes the name of the pdb, the classification of the dimer and the classification of the secondary structure are reported respectively in the Complex, Dimer and Structure columns.
| Complex | $B_a$ | Complex | $B_a$ | Complex | $B_a$ |
|---------|-------|---------|-------|---------|-------|
| 1jtd    | -10.57 | 1y34    | -9.28 | 2uuy    | -8.25 |
| 1bvn    | -11.05 | 1wr1    | -4.83 | 2ptt    | -5.4  |
| 1j7d    | -5.7   | 1wrd    | -3.39 | 2uyz    | -7.09 |
| 1emv    | -13.62 | 1ugh    | -10.92 | 2rn    | -6.82 |
| 1grn    | -6.41  | 1y33    | -8.77 | 2omw    | -5.0  |
| 1jiw    | -11.4  | 1y4a    | -8.92 | 2qc1    | -11.0 |
| 1e96    | -5.57  | 1y3d    | -9.25 | 2wp3    | -6.1  |
| 1j7y    | -10.46 | 1yx6    | -4.14 | 2v9t    | -8.09 |
| 1ay7    | -6.0   | 1zgu    | -4.01 | 2vlq    | -10.64 |
| 1lp1    | -5.7   | 1y4d    | -8.92 | 2wvk    | -5.89 |
| 1lzw    | -6.48  | 1y3c    | -10.6 | 2wo3    | -5.64 |
| 1l8c    | -8.15  | 1y6k    | -9.0  | 2wy8    | -6.44 |
| 1m10    | -8.24  | 1yx5    | -3.46 | 3bh6    | -7.02 |
| 1kac    | -7.83  | 1y4x    | -8.77 | 2z58    | -7.8  |
| 1lx5    | -8.92  | 2a9k    | -7.22 | 2vlp    | -12.27 |
| 1lw6    | -11.7  | 1zvy    | -10.0 | 3knb    | -6.03 |
| 1ry7    | -6.64  | 1zv5    | -8.0  | 3ixe    | -5.64 |
| 1ta3    | -8.05  | 2b12    | -4.7  | 3kuc    | -6.35 |
| 1t0p    | -4.6   | 2a78    | -7.22 | 3kud    | -5.77 |
| 1sq0    | -7.52  | 2hrk    | -8.05 | 3flp    | -3.92 |
| 1op9    | -9.15  | 2jti    | -6.23 | 3hct    | -5.83 |
| 1ri8    | -8.54  | 2k2s    | -7.28 | 3k1r    | -9.0  |
| 1r8u    | -7.89  | 2hev    | -7.21 | 3bn3    | -4.7  |
| 1te1    | -8.47  | 2jy6    | -4.7  | 3kw5    | -6.41 |
| 1tm5    | -10.77 | 2jt4    | -4.4  | 3elz    | -10.44 |
| 1tdq    | -7.92  | 2hth    | -3.98 | 3fpu    | -9.92 |
| 1tm1    | -11.52 | 2f4m    | -7.19 | 3fu    | -8.8  |
| 1tm7    | -11.48 | 2fuh    | -3.52 | 3nc    | -0.67 |
| 1u0s    | -6.64  | 2k6d    | -3.77 | 3n0p    | -2.29 |
| 1tm3    | -10.72 | 2k3s    | -5.57 | 3me2    | -10.17 |
| 1to1    | -9.34  | 2k5b    | -5.4  | 3m18    | -9.74 |
| 1thl    | -3.52  | 2k8c    | -2.74 | 3ncb    | -1.03 |
| 1tm4    | -9.89  | 2k8b    | -2.74 | 3nv    | -8.03 |
| 1tba    | -9.0   | 2k79    | -3.17 | 3n06    | -2.51 |
| 1vet    | -7.89  | 2k7a    | -3.17 | 3mzg    | -2.3  |
| 1wq1    | -4.77  | 2o3b    | -11.49 | 3qq8    | -5.82 |
| 1y1k    | -9.52  | 2knb    | -4.96 | 3tnf    | -11.0 |
| 1xg2    | -8.3   | 2nqd    | -10.41 | 3oma    | -6.17 |
| 1wqj    | -6.06  | 2omu    | -9.22 | 3t04    | -7.28 |
| 1nus    | -2.52  | 2ka4    | -7.24 | 3qc8    | -4.95 |
| 1y3b    | -9.51  | 2omx    | -6.92 | 3uyo    | -5.17 |

**TABLE II:** 'Affinity' dataset. For each of the 123 considered complexes the name of the pdb and the $B_a$ are reported.
FIG. 1: Charge properties of the dataset. a) On each interacting surface residue a sphere of radius R is built, and the number of interacting residues on the partners surface included in the sphere is counted. The bar plot shows, for increasing values of R (as reported by the labels on the right) and for both the whole dataset and each of the four classes, the fraction of positively or negatively charged residues that can be found close to positive residues, respectively in yellow or ochre. In grey, the fraction of negative residues closed to a negative amino acid. b) For increasing values of R (as reported by the labels on the right) and for both the whole dataset and each of the four classes, the fraction of positively, negatively or null residues that can be found close to non charged residues, respectively in brown, beige and grey.
FIG. 2: Structural classification of the 'Human' dataset, amino acid composition and charge properties of the classes. 

a) The complexes in the dataset are divided into SS, HH and SH. The colored boxes report an example for each category. The same colors are used to indicate in the pie chart each class abundance in the dataset. 

b) For each protein, the sum of the charges of all its residues and only the interacting residues on the surface is computed. For each complex, these total and interacting charges from the two interacting partners are multiplied. The bar plot shows, for the whole dataset and each class, the percentage of complexes whose total (in orange) and interacting (in blue) products are negative. 

c) On each interacting surface residue a sphere of radius R is built, and the number of interacting residues on the partners surface included in the sphere is counted. The bar plot shows, for increasing values of R (as reported by the labels on the right) and for both the whole dataset and each of the three classes, the fraction of positively or negatively charged residues that can be found close to positive residues, respectively in yellow or ochre. In grey, the fraction of negative residues closed to a negative amino acid. 

d) For increasing values of R (as reported by the labels on the right) and for both the whole dataset and each of the three classes, the fraction of positively, negatively or null residues that can be found close to non charged residues, respectively in brown, beige and grey. 

e) The relative abundances of each of the twenty natural amino acids considering all the residues (orange), only the interacting ones (in green), and only the solvent-exposed residues (brown) are shown. The results are divided into the three classes.
FIG. 3: Electrostatic complementarity contribution in protein-protein complexes. Distributions of the F values of the interacting patches in complexes from the IBR-hom (blue) and nIBR-hom (green) classes. In the insert the corresponding ROC curves.

FIG. 4: 2D Zernike polynomials to compare surfaces regions. a) Distributions of the distances between the Zernike vectors describing the molecular surface of nIBR-het and SBR-hom interacting (red and yellow respectively) and random (grey) patches in the Human dataset. The distribution of all the patches in the dataset is shown in orange. b) ROC curves of the distributions in a) and corresponding AUC (in the legend) computed against the random distribution. c) For each patch the distance between the Zernike vectors describing the electrostatic potential surface in that region is computed. Then the same analysis and classification as in a) is performed. d) ROC curves of the distributions in c) and corresponding AUC (in the legend) computed against the random distribution.
II. PREDICTION TESTING

To test the observed anti-correlation between binding affinity and electrostatic complementarity we selected a third dataset, that we call ‘Variants’ dataset. This dataset includes five SARS-CoV-2 variants of concerns (VOCs) (alpha, beta, gamma, delta and omicron) with known dissociation constant ($K_d$) [1]. The ‘Variants’ dataset was obtained starting from the experimentally resolved structure of the wild-type (WT) spike protein bound to Angiotensin-Converting Enzyme 2 (ACE2) (pdb id: 6M0J). Since not all the VOCs have an available experimental structure, WT was subjected to computational mutagenesis using the dedicated tool provided in the PyMol software [2]. We selected the ACE2 residues from 19 to 615 in complex with spike residues from 333 to 526, and we only considered the mutations in the spike Receptor Binding Domain (RBD), including residues from 319 to 541, because those are the interacting regions [3, 4]. For each complex, we performed with Gromacs 2020.6 [5] a 100-ns-long molecular dynamics simulation and extracted configurations of the system every 1 ns. For this dataset, to reduce the computational time, the centers of all the frames’ interacting regions were defined using the starting structure of the spike protein original version. We super-positioned each structure with the original spike protein, and selected the points closest to the binding region on this original version. To increase the volume of our data despite the low number of complexes, for each frame we defined N pairs of interacting patches, where N corresponds to the 5% of the points forming the surface mesh included in that interacting region. To avoid redundancy in the analysis we defined the patches with a radius of 6 Å. Since for this dataset we only performed the Zernike-based complementarity evaluation, this radius value was already shown to result in the highest efficiency [6]. The so-defined interacting patches include the points of the electrostatic surfaces that are projected in the electrostatic matrices. The electrostatic surfaces and the electrostatic matrices were instead obtained as described in the Methods in the main text.

We then applied the Zernike formalism to measure the complementarity between ACE2 and each variants for all the simulation frames. Table III shows for each variant the experimental $K_d$ and the shape ($Z_s$ column) and electrostatic ($Z_{el}$ column) complementarities measured in terms of Euclidean distances between the Zernike descriptors, as discussed in the main text. As expected, complexes with lower $K_d$ (higher binding stability) have higher shape complementarity (smaller $Z_s$ values) because the role of the Lennard-Jones potential predominates. On the other hand, complexes with higher $K_d$ (lower binding affinity) tend to have a higher electrostatic complementarity (smaller $Z_{el}$ values) because they exploit Coulombic complementarity to acquire specificity. This would seem to confirm that electrostatic complementarity has greater role in less stable complexes.

| Variant | $K_d$ | $Z_s$ | $Z_{el}$ |
|---------|-------|-------|---------|
| Alpha   | 5.40  | 2.97  | 17.88   |
| Gamma   | 11.00 | 2.96  | 17.96   |
| Beta    | 13.83 | 2.96  | 17.94   |
| Delta   | 25.07 | 3.01  | 17.40   |
| Omicron | 31.40 | 3.05  | 16.87   |

**TABLE III:** Dissociation constant and shape and electrostatic complementarity for five SARS-CoV-2 variants. List of the SARS-CoV-2 variants considered in this study with their $K_d$ as measured by Han, P. et al [1] in nM, and the shape ($Z_s$ column) and electrostatic ($Z_{el}$ column) complementarities measured in terms of Euclidean distances between the Zernike descriptors.
[1] Pengcheng Han, Linjie Li, Sheng Liu, Qisheng Wang, Di Zhang, Zepeng Xu, Pu Han, Xiaomei Li, Qi Peng, Chao Su, et al. Receptor binding and complex structures of human ace2 to spike rbd from omicron and delta sars-cov-2. Cell, 185(4):630–640, 2022.

[2] Warren L DeLano et al. Pymol: An open-source molecular graphics tool. CCP4 Newsl. Protein Crystallogr, 40(1):82–92, 2002.

[3] Renhong Yan, Yuanyuan Zhang, Yaning Li, Lu Xia, Yingying Guo, and Qiang Zhou. Structural basis for the recognition of sars-cov-2 by full-length human ace2. Science, 367(6485):1444–1448, 2020.

[4] Gabriele Cerutti, Yicheng Guo, Tongqing Zhou, Jason Gorman, Myungjin Lee, Micah Rapp, Eswar R Reddem, Jian Yu, Fabiana Bahna, Jude Bimela, et al. Potent sars-cov-2 neutralizing antibodies directed against spike n-terminal domain target a single supersite. Cell host & microbe, 29(5):819–833, 2021.

[5] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E Mark, and Herman JC Berendsen. Gromacs: fast, flexible, and free. Journal of computational chemistry, 26(16):1701–1718, 2005.

[6] Edoardo Milanetti, Mattia Miotto, Lorenzo Di Rienzo, Michele Monti, Giorgio Gosti, and Giancarlo Ruocco. 2d zernike polynomial expansion: Finding the protein-protein binding regions. Computational and Structural Biotechnology Journal, 19:29–36, 2021.