

# First-Principles Simulation of Dielectric Function in Biomolecules

Puja Adhikari <sup>1</sup>, Rudolf Podgornik <sup>2,3,4,5,\*</sup>, Bahaa Jawad <sup>1</sup> and Wai-Yim Ching <sup>1,\*</sup>

- <sup>1</sup> Department of Physics and Astronomy, University of Missouri-Kansas City, Kansas City, MO 64110, USA; paz67@umkc.edu (P.A.); bajrmd@mail.umkc.edu (B.J.)
- <sup>2</sup> School of Physical Sciences, Kavli Institute of Theoretical Science, University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>3</sup> CAS Key Laboratory of Soft Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100090, China
- <sup>4</sup> Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou 325000, China
- <sup>5</sup> Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia
- \* Correspondence: rudipod@gmail.com (R.P.); Chingw@umkc.edu (W.-Y.C.)

**Abstract:** The dielectric spectra of complex biomolecules reflect the molecular heterogeneity of the proteins and are particularly important for the calculations of electrostatic (Coulomb) and electrodynamic (van der Waals) interactions in protein physics. The dielectric response of the proteins can be decomposed into different components depending on the size, structure, composition, locality, and environment of the protein in general. We present a new robust simulation method anchored in rigorous ab initio quantum mechanical calculations of explicit atomistic models, without any indeterminate parameters to compute and gain insight into the dielectric spectra of small proteins under different conditions. We implement this methodology to a polypeptide RGD-4C (1FUV) in different environments, and the SD1 domain in the spike protein of SARS-COV-2. Two peaks at 5.2–5.7 eV and 14.4–15.2 eV in the dielectric absorption spectra are observed for 1FUV and SD1 in vacuum as well as in their solvated and salted models.

**Keywords:** electrostatic interaction; dielectric function; biomolecules; ab initio simulation; random phase approximation



**Citation:** Adhikari, P.; Podgornik, R.; Jawad, B.; Ching, W.-Y. First-Principles Simulation of Dielectric Function in Biomolecules. *Materials* **2021**, *14*, 5774. <https://doi.org/10.3390/ma14195774>

Academic Editor: Claudio Canale

Received: 3 September 2021  
Accepted: 28 September 2021  
Published: 2 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The dielectric properties of proteins in aqueous solutions have been studied for several decades [1] and play a crucial role in the calculations of protein electrostatic Coulomb interactions [2] as well as electrodynamic dispersion van der Waals interactions [3,4], specifically to calculate the protein–protein and protein–nucleic acid interactions, characterize the folding pathways, and investigate solution behavior and stability. In fact, quite recently electrostatic contributions have been invoked as a possible source of the differences between the binding free energy of the spike proteins of SARS-CoV-2 and SARS-CoV to the ACE2 human receptor [5], as well as being the principal interaction component between the spike proteins of SARS-CoV-2 virus and the charged electret fibers in personal protective gear [6].

The problem of the macroscopic protein dielectric “constant” is exacerbated by the fact that, as in other biomolecular systems such as lipid membrane bilayers composed of different segregated molecular components, the underlying microscopic dielectric properties are heterogeneous [7] and the continuum assumptions become—to say the least—problematic on a molecular level [8]. Consequently, the protein dielectric constant appears rather as a phenomenological parameter, or even a phenomenological function of the position inside the protein core [9], that depends on the model used to describe the heterogeneous molecular structure of proteins, then a single universal constant grounded in some fundamental theory [10]. These are important issues as many continuum computational approaches [11]

as well as simulation methodologies to characterize protein electrostatics [12] hinge crucially on the numerical choice for the dielectric constant of the protein interior, which seems to matter not only quantitatively but in many cases even qualitatively [13].

Detailed knowledge of the dielectric response in the frequency domain is particularly important for the calculation of the dispersion van der Waals interactions [14], which are expressed as a non-local functional of the dielectric spectrum at imaginary frequencies in the Lifshitz macroscopic theory [3]. Notably, the first computation of van der Waals interaction with full spectral resolution was performed for a biophysical system of interacting phospholipid membranes [15]. Non-covalent dispersion van der Waals interactions are also thought to play an important role in protein folding [16] since they make a substantial contribution to protein–protein as well as protein–water interactions [17]. Additionally, in applications tied to organic optical and electronic devices as well as biotechnology, the contribution of van der Waals interactions to the stability of proteinaceous films deposited on a dielectric substrate has been recently noted [18,19].

Even though the dielectric properties of various metals, semiconductors, and insulators, organic as well as inorganic, have been critically assessed in great detail [14], the dielectric spectrum data for proteins and polypeptides mostly lack the relevant details regarding the optical properties in a wide frequency range from zero frequency to the far ultraviolet [18]. In the absence of relevant experimental data, the calculated ultraviolet frequency region dielectric spectra for the cyclic tripeptide RGD-4C (1FUV) [20], based on an *ab initio* quantum mechanical (QM) computational scheme, were successfully used to calculate the oscillator strengths, oscillator frequencies, and the relaxation parameters that enter the calculations of the van der Waals interactions in proteinaceous systems [19].

Motivated by the past successful calculations of the electronic part of the dielectric function at full spectral resolution for many crystals and amorphous solids of infinite extent, based on the Bloch Theorem [21] considered to be the cornerstone of the electronic structure theory of condensed matter physics, we now apply the same methodology to proteins where by assumption the unit cell encloses the whole protein. We use the *ab initio* quantum mechanical (QM)-based computational scheme, dubbed the quantum mechanical random phase approximation (QMRPA) method, anchored in random phase approximation (RPA) [22], to calculate the optical transitions from occupied to unoccupied states with explicit inclusion of the dipole transition matrix from the *ab initio* wave functions. In what follows, we apply this methodology to calculate the full spectral resolution of the dielectric response for protein molecules, as well as to calculate the relevant partial charges and to generalize the concept of bond order parameter from the usual inter-atomic domain to be able to also characterize the bonding of complete amino acids (AAs). There are of course major differences between the standard implementation of this methodology in solid state physics and its application to biomolecules, specifically to proteins, which are due to their finite size, the fact that they possess no periodically replicated unit cell, and finally that they are usually immersed in a bathing aqueous medium of its own molecular structure.

Here, we implement the *ab initio* QMRPA approach to obtain the dielectric spectra, the partial charges and the AA bond order to different biomolecules such as polypeptides and proteins, specifically the small RGD (1FUV) polypeptide in different environments as well as the subdomain SD1 of the SARS-CoV-2 spike protein [23]. Our calculations are free of adjustable parameters and scalable for larger biomolecules. While some *ab initio* calculations of molecular polarizability in response to an external electric field have been attempted in the past, they were mostly focused restrictively to very small molecules such as single AAs [24–26], and may have included electronic as well as vibrational effects. To our knowledge, *ab initio* calculation has not been attempted for small proteins consisting of a substantial number of different AAs, which are of course far more challenging [27].

## 2. Methods

### 2.1. Model Construction

The starting structure of RGD-4C peptide was downloaded from the RCSB protein data bank (PDB) with ID: 1FUV [28] based on nuclear magnetic resonance (NMR) data. 1FUV contains 11 AAs with 135 atoms, including hydrogen atoms. From the initial dry model, two solvated models with 80 and 100 H<sub>2</sub>O molecules and two salted models with H<sub>2</sub>O and salt ions at concentration 0.15 M were obtained. The solvated models were generated by the three-point charge TIP3P model [29] using LEaP program included in the AMBER 18 package [30].

In total, five RGD-4C models were constructed: one dry, two solvated, and two salted models with monovalent salt at 0.15 M concentration (4 Na<sup>+</sup> and 3 Cl<sup>−</sup> ions). These ions are systematically placed in a shell surrounding the protein by using a Coulombic potential on a grid with the program LEaP, while the number of ions is determined from the volume of the cell. For the dry model of SD1 in the spike protein of SARS-CoV-2, its fully relaxed structure was obtained from PDB [23] with ID: 6VSB [31]. The solvated SD1 model is built with 300 water molecules using the same approach as in the 1FUV cases. The selection of 300 water molecules is based on the UCSF Chimera program [32]. All models including 1FUV and SD1 were optimized using VASP.

### 2.2. Structure Optimization Using VASP

This initial atomic-scale structure for the individual small protein is then fully optimized by using the Vienna ab initio simulation package (VASP) [33]. The pseudopotential plane-wave based VASP package is known for its efficiency in relaxation to the equilibrium structure at minimal energy.

Our experience and tests suggest the use of the following input parameters in VASP: Energy cut-off energy at 500 eV, electronic convergence of 10<sup>−4</sup> eV; force convergence for ionic steps at −10<sup>−2</sup> eV/Å and a single k-point sampling. After the optimization for SD1 with 24 amino acids and 391 atoms, the total energy decreases from −2370.90 to −2379.21 eV or 2.05 kJ/mol per atom. The VASP-relaxed structure is used for the electronic structure, interatomic bonding and dielectric function calculations using OLCAO method described below.

We used the projector augmented wave (PAW) method with Perdew–Burke–Ernzerhof (PBE) exchange correlation potential [34] within the generalized gradient approximation (GGA).

### 2.3. DFT Calculations Using OLCAO

For electronic structure, interatomic interactions, and dielectric function calculations, a different DFT method is used, the all-electron orthogonalized linear combination of atomic orbitals (OLCAO) method [35], developed in-house. The combination of these two different DFT codes is extremely efficient for large complex materials and is well documented [20,36–38] especially for complex proteins such as those of the SARS-CoV-2 virus [23,39,40]. The key feature of the OLCAO method is the possibility it provides to calculate the effective charge ( $Q^*$ ) on each atom as well as the bond order values  $\rho_{\alpha\beta}$  between any ( $\alpha\beta$ ) pair of atoms.

The partial charge (PC) or ( $\Delta Q_\alpha = Q_\alpha^0 - Q_\alpha^*$ ) is the deviation of the effective charge  $Q_\alpha^*$  from the neutral atomic charge  $Q_\alpha^0$  on the same atom  $\alpha$ . The bond order (BO) values are obtained from the ab initio wave functions with atomic basis expansion calculated quantum mechanically:

$$Q_\alpha^* = \sum_i \sum_{m,occ} \sum_{j,\beta} C_{i\alpha}^{*m} C_{j\beta}^m S_{i\alpha,j\beta} \quad (1)$$

$$\rho_{\alpha\beta} = \sum_{m,occ} \sum_{i,j} C_{i\alpha}^{*m} C_{j\beta}^m S_{i\alpha,j\beta} \quad (2)$$

In the above equations,  $S_{i\alpha,j\beta}$  are the overlap integrals between the  $i^{th}$  orbital in  $\alpha^{th}$  atom and the  $j^{th}$  orbital in the  $\beta^{th}$  atom.  $C_{j\beta}^m$  are the eigenvector coefficients of the  $m^{th}$

occupied molecular orbital levels. The BO values  $\rho_{\alpha\beta}$  in Equation (2) can be calculated for every pair of atoms ( $\alpha, \beta$ ) in the optimized structure with precise atomic positions. The BO quantifies the strength of the bond between two atoms and generally scales with the bond length (BL) being also influenced by the surrounding atoms. The sum of BO of all pairs in the system gives the total bond order, TBO. The calculation of PC and BO are based on the Mulliken scheme [41,42], and hence are basis-dependent. Comparisons of BO values using different methods should be treated with caution.

In proteins, the focus is rather on AAs or their individual residues, then on the individual atoms. AAs in fact contain different atoms in different molecular configurations and orientations. Strictly speaking, assigning the distance of separation between two AAs in a protein to describe their interaction is a vague and arbitrary parameter. However, with the quantum mechanically based OLCAO method and with the interatomic interaction between all atoms readily available, we can define the bond order between two AAs  $u$  and  $v$  without any ambiguity, and denote it as *amino acid bond pair* (AABP) [40]:

$$\text{AABP}(u, v) = \sum_{\alpha \in u} \sum_{\beta \in v} \rho_{\alpha i, \beta j} \quad (3)$$

where the summations are over atoms  $\alpha$  in AA  $u$  and atoms  $\beta$  in AA  $v$ . This is a rigorously defined quantity and can be further extended to different units or (sub)groups of AAs if necessary. The merit of the above scheme is that AABP includes all possible bonding between two amino acids such as covalent, ionic, hydrogen bonding (HB) and even their intermediate mixtures [43].

This single quantitative parameter reflects the internal bonding strength among different AAs in a protein, or AAs between different proteins. While proteins are generally characterized by a primary sequence of nearest neighbor (NN) AAs, non-NN amino acids in the 3D structure of a protein can also interact and the AABP can be resolved into the local NN part and the non-local bonding part, providing much more penetrating details on inter AA bonding.

#### 2.4. Optical Transition and Random Phase Approximation

We use OLCAO method to calculate the optical transitions in a biomolecule from occupied to the unoccupied molecular states within the random phase approximation [22]. The transition explicitly involves the dipole transition matrix elements calculated quantum mechanically from the ab initio wave functions which automatically obey the transition rules involved. We name the procedure thus outlined as the quantum mechanical random phase approximation (QMRAP) method.

The complex frequency dependent dielectric function,  $\varepsilon(\omega)$ , is given by the standard decomposition:

$$\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega) \quad (4)$$

where

$$\varepsilon_2(\omega) = \frac{e^2}{\pi m \omega^2} \int_{\text{BZ}} dk^3 \sum_{n,l} |\langle \Psi_n(\vec{k}, \vec{r}) | P | \Psi_l(\vec{k}, \vec{r}) \rangle|^2 \delta(E_n(\vec{k}) - E_l(\vec{k}) - E) \quad (5)$$

Here,  $\psi_n(\vec{k}, \vec{r})$  is the Bloch wave function for the  $n^{\text{th}}$  band with energy  $E_n(\vec{k})$  at Brillouin zone point  $k$ . Momentum matrix elements  $\langle \Psi_n(\vec{k}, \vec{r}) | P | \Psi_l(\vec{k}, \vec{r}) \rangle$  from occupied valence band states ( $l$ ) to empty conduction band states ( $n$ ) are calculated from ab initio wave functions. The  $E_l$  and  $E_n$  are the energy of occupied state and unoccupied state, respectively.  $E = \hbar\omega$  is the photon energy.

The real part  $\varepsilon_1(\omega)$  can be derived from  $\varepsilon_2$  by using the Kramers–Kronig relation [44]:

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{s \varepsilon_2(s)}{s^2 - \omega^2} ds, \quad (6)$$

while the energy loss function  $F(\omega)$  can be obtained as:

$$F(\omega) = \text{IM} \left( -\frac{1}{\varepsilon(\omega)} \right) = \frac{\varepsilon_2(\omega)}{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} \quad (7)$$

In crystalline solids, the position of the peak of the energy loss function  $F(\omega)$  at  $\omega_0$  is identified as the plasmon frequency, which is the energy for the collective excitation of the electrons in the unoccupied bands.

### 2.5. Past Record of Using above Methods

The QMRPA method for inorganic crystalline and amorphous materials has been successfully used by us over the last 30 years [35]. Notable examples are the timely calculations of the dielectric functions of two exceptional materials, the high temperature ceramic superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  (YBCO) in 1987 [45,46] and the Buckminster fullerene  $\text{C}_{60}$  FCC crystals in 1991 [47,48]. More recent examples of complex materials include mixed inorganic glass  $(\text{a-SiO}_2)_{1-x}(\text{GeO}_2)_x$  [49] and amorphous zeolitic imidazolate framework (a-ZIF) showing the metal-insulator transition under compression [50,51].

The calculation of dielectric functions of biomaterials started nearly 20 years ago. For example, the calculation of optical transition in vitamin  $\text{B}_{12}$  Adenosyl-cobalamin has shown good agreement with experimental spectra [52,53]. Similar calculations of the electronic structure and X-ray absorption spectra were carried out on hydroxyapatite and different calcium apatite crystals [54,55]. A particularly interesting example is the calculation of the imaginary part of the dielectric function  $\varepsilon_2(\omega)$  of herapathite, a large complex dichroic crystal  $(\text{C}_{20}\text{H}_{24}\text{N}_2\text{O}_2\text{H}_2)_4 \cdot \text{C}_2\text{H}_4\text{O}_2 \cdot 3\text{SO}_4 \cdot 2\text{I}_3 \cdot 6\text{H}_2\text{O}$  [56], verifying the strong linear dichroism with anisotropy factor of 385 that has numerous applications for its polarizing properties [57].

Other examples include the application of QMRPA method to dispersion interactions based on dielectric spectra for biomolecular systems [58–60] and the effect of optical absorption of cytochrome c (Cyt c) macromolecules at the interface or surface for high performance photodetection [61,62]. Last but not the least is the special peptide (RGD-4C) that will be highlighted in later sections. In the present paper, we focus on the QMRPA method as applied to the RGD (1FUUV) polypeptide.

## 3. Results on RGD (1FUUV) Peptide

This section focuses on the calculation of electronic structure, dielectric function, and partial charge of a simpler biomolecule RGD (1FUUV) peptide as a specific example. Other results including the extension to a more complex biomolecule SD1 domain of spike protein in SARS-CoV-2 and other relevant properties are separately described in Sections 4 and 5 that follow.

### 3.1. RGD (1FUUV) Peptide

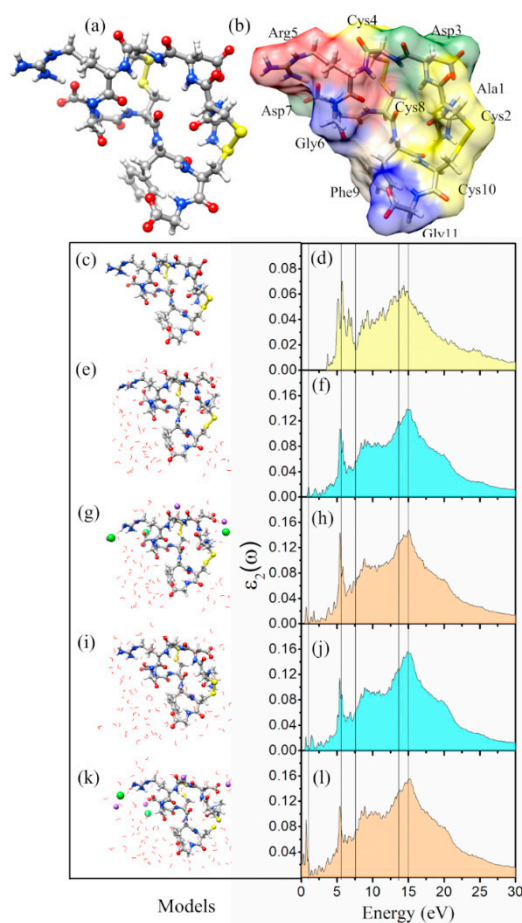
Peptide sequences with the arginine-glycine-aspartic acid (RGD) motif display a strong affinity and selectivity for a membrane protein called *integrin*, which is a key cell surface receptor mediating cell adhesion to extracellular matrices (ECM) [63,64]. Integrins are actively expressed on vascular endothelial cell surfaces and play a key role in tumor metastasis, leukocyte migration, and angiogenesis [65], making them an ideal target for treating inflammatory diseases and cancer. Because integrin receptors recognize RGD as a primary sequence, RGD peptides are used to target cancer cells [66]. In addition, RGD peptides provide numerous applications in biological and biomedical devices, being frequently incorporated into biomaterials designed to facilitate wound healing [67], serving as candidates for radiotracers in imaging [68], being used in implantable medical devices [69], and mimicking the activities of the ECM proteins in culture cells [70].

We use the QMRPA method to calculate the dielectric function of a cyclic tripeptide RGD-4C to demonstrate its applicability in the case of a small protein. The initial structure of the RGD-4C was obtained from the Protein data bank (PDB ID: 1FUUV) based on nuclear



magnetic resonance (NMR) data [28]. The electronic structure, dielectric response, and the solvent accessible surface partial charge distribution of the RGD peptide has been studied already [20], while here we present improved accuracy calculations of the dielectric function of dry 1FUV as well as for the 1FUV solvated and salted models.

1FUV has 11 AAs (ACDCRGDCFCG) of six different types: Gly (two), Cys (four), Phe (one), Asp (two), Arg (one), and Ala (one). For its 135 atoms, including H atoms, it is somewhat clustered but otherwise a rather typical small peptide. It is one of the most used RGD variants because of the presence of the four Cys residues which allow for the two rare disulfide bonds. Figure 1a–c show a ball and stick structure of 1FUV after VASP relaxation. The 11 AAs in 1FUV are marked in Figure 1b. There are considerable non-local interactions between non-nearest neighbor (NN) AAs in the primary sequence. This aspect of 1FUV will be elaborated further in Section 5.



**Figure 1.** Dry, solvated, and salted models for 1FUV with their  $\epsilon_2(\omega)$  spectra and  $\epsilon_1(0)$  values. Ball and stick illustration of (a,c) dry 1FUV with (b) surface shown with different colors.  $\epsilon_2(\omega)$  for: (d) dry 1FUV; (f) 1FUV with 80 H<sub>2</sub>O, (h) 1FUV with 80 H<sub>2</sub>O and 0.15 NaCl, (j). 1FUV with 100 H<sub>2</sub>O, and (l) 1FUV with 100 H<sub>2</sub>O and 0.15 NaCl. The ball and stick figure for respective cases are in (e,g,i,k) the left panel. Gray: C, red: O, blue: N, white: H, green: Cl, and purple: Na.

Figure 1d shows the final imaginary part  $\epsilon_2(\omega)$  for up to 30 eV in the case of dry 1FUV and will be discussed again in the following section together with the spectra (Figure 1e,f,i,j) for the solvated and (Figure 1g,h,k,l) for solvated and salted models. As can be seen, the general feature of  $\epsilon_2(\omega)$  of 1FUV is the broader peak at 14.4 eV and discernable multiple sharp peaks below 5.7 eV. We now proceed to the solvated and salted models with 80 and 100 water molecules and added 0.15 M NaCl (4 Na and 3 Cl ions).

### 3.2. Local Environmental Effect on RGD (1FUV) Peptide

For biomolecular systems and proteins in particular, environment effects such as the presence of solvent, fixed pH reservoir, fixed salt content and *protonation/deprotonation equilibrium* etc. must ideally be considered. In the explicit QMRPA calculation, which is a single point calculation, a specific model for each of the environmental effects needs to be constructed individually with increased size and complexity.

The construction of four solvated and salted models is described in the Method Section 2.1. The atomic structures of these four solvated models are further optimized using VASP and the final ball and stick sketch for these four models are displayed in the left panel of Figure 1. These 2D projections of the solvated structures of 1FUV clearly show that the water molecules are evenly distributed with random orientation and no penetration of water into the interior of the protein. The locations of the Na<sup>+</sup> and Cl<sup>-</sup> ions depend on the VASP optimization routines, corresponding to the most energetically favorable positions.

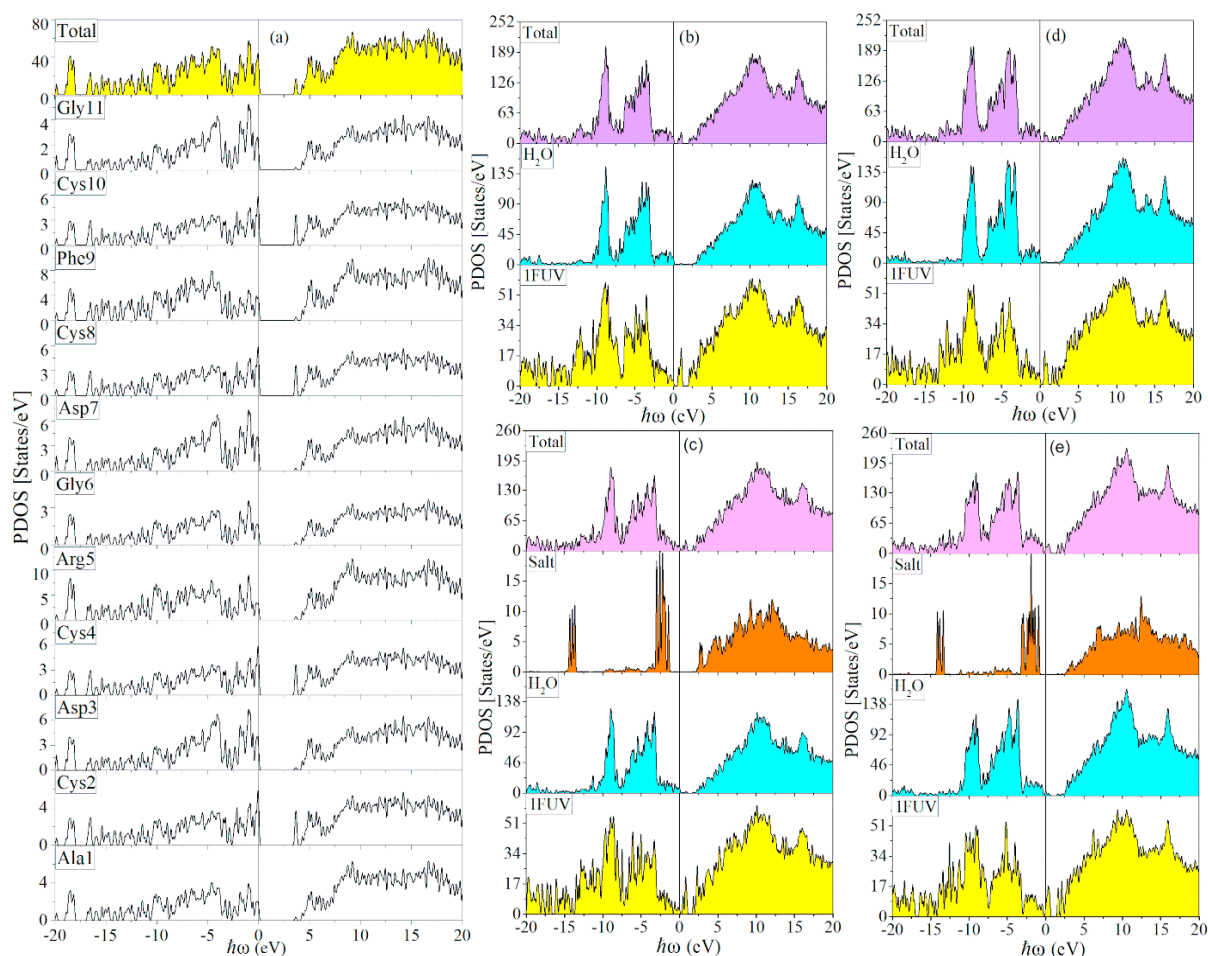
We now present the results for these four solvated and salted 1FUV models together with the dry model. The calculated dielectric spectra  $\epsilon_2(\omega)$  of the four solvated 1FUV models are shown in Figure 1f,h,j,l. It can be seen that the general features in the solvated models are similar to the dry model in Figure 1d, except for the appearance of many small transition peaks below 6 eV arising from the transitions from or to the gap states, as evidenced in the density of states (DOS) presented in Figure 2. When the salts Na<sup>+</sup> and Cl<sup>-</sup> ions are added, the  $\epsilon_2(\omega)$  and  $\epsilon_1(\omega)$  spectra are changed even further in a more complex manner depending on the actual location of the gap states and the strength of their dipole transition matrix elements (Equation (5) in Methods Section 2.4).

Figure 2 shows the calculated DOS and partial DOS (PDOS) for all five models. For the dry 1FUV, the total DOS (TDOS) is also resolved into 11 different AAs. The four S-containing Cys residues have separate peaks in the LUMO, which is not the case for other individual AAs. This feature comes from the interaction of Cys with the other AAs in 1FUV. We also observe that there are no *gap states* between the HOMO-LUMO gap in the dried 1FUV model as shown in Figure 2a. In addition, dry 1FUV has a wider band gap ( $E_g$ ) in comparison to solvated and salted models as shown in Table 1. In Figure 2b–e for the solvated and salted models, the PDOS are resolved into those from 1FUV, the H<sub>2</sub>O molecules and the salt ions. It is of interest to note that the TDOS in Figure 2b–e does have many gap states, but the PDOS can show it arises due to the interaction between different AAs in solvated 1FUV, the individual H<sub>2</sub>O molecules and the Na<sup>+</sup> or Cl<sup>-</sup> ions of the dissolved salts.

The bond order (BO) for every pair of atoms calculated from Equation (2) in Method Section 2.3), is an integral part of the description of biomolecular materials [35,39,40]. Figure 3 shows BO versus bond length (BL) plots for the dry 1FUV model and the more complex solvated model with 80 and 100 H<sub>2</sub>O molecules and salted models with 80 H<sub>2</sub>O and 100 H<sub>2</sub>O molecules. The bond order includes contributions within each AA, between different AAs, and between atoms in each of the AA with H<sub>2</sub>O molecules or salt ions in the solvated models.

**Table 1.** Total bond order (TBO) and band gap ( $E_g$ ) for five 1FUV models.

Models	TBO (e)	$E_g$ (eV)
1FUV	54.27	3.61
1FUV + 80 H <sub>2</sub> O	98.49	0.68
1FUV + 80 H <sub>2</sub> O + NaCl	99.95	0.47
1FUV + 100 H <sub>2</sub> O	109.64	0.57
1FUV + 100 H <sub>2</sub> O + NaCl	110.30	0.15

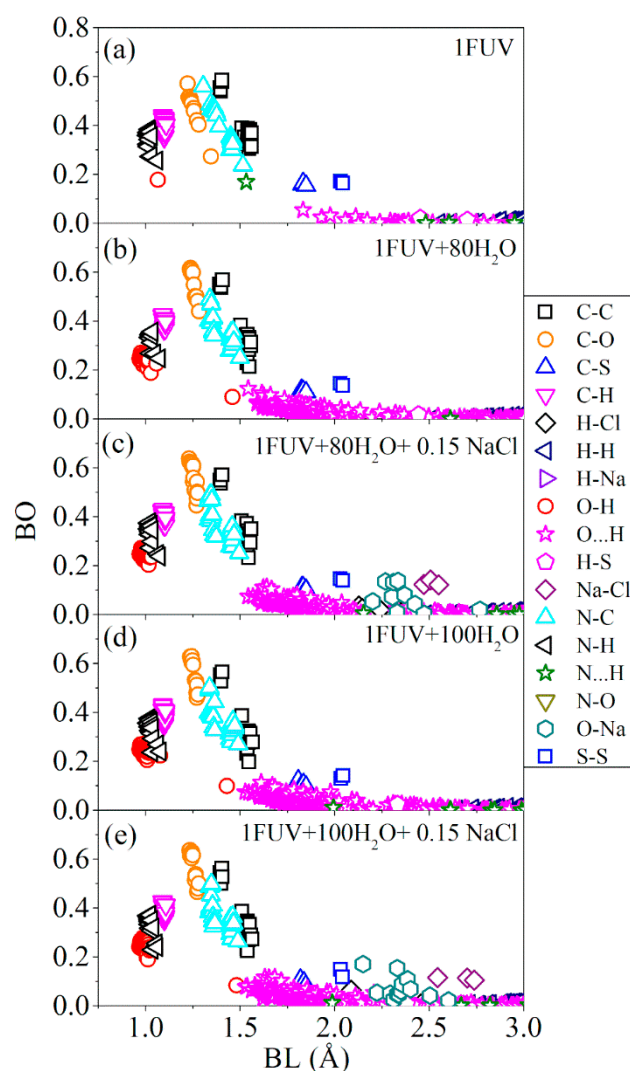


**Figure 2.** DOS and PDOS of 5 models for 1FUV: (a) Dry model further resolved for each amino acid in a separate panel; (b) with 80 water molecules resolved into 1FUV and 80 H<sub>2</sub>O; (c) with 80 H<sub>2</sub>O molecules, and 4 Na and 3 Cl ions resolved into 1FUV, 80 H<sub>2</sub>O, and 0.15 NaCl (d) with 100 water molecules resolved similar to (b,e) with 100 H<sub>2</sub>O molecules, and 0.15 NaCl resolved similar to (c).

For clarity, we briefly describe the dry 1FUV model and the solvated and salted 1FUV model with 100 H<sub>2</sub>O molecules and salt atoms of Figure 3.

*Dry 1FUV:* the strongest bonds are of course the covalent bonds containing C (C-C, C-O, N-C, C-H). Their BL varies slightly but BO values can vary rapidly, scaling with BL. The BO depends on the nature of the covalent bond (single or double), the atoms it bonds and, to a lesser extent, the local environment of the bonded pair. The covalent and C-H bonds within or between different AAs have a BL slightly larger than 1.0 Å and exhibit varying BO values. In Figure 3a, there is a singular O-H bond that can be traced to O in residue Asp7 and H in residue Arg5 which can also be interpreted as an exceptionally strong HB. The hydrogen bonds (HB) O···H or N···H are of vital importance in any biological system, being weaker than covalent bonds but ubiquitous especially in the presence of solvents. The BLs of HBs generally range between 1.5 and 3.0 Å and can have BO values larger than 0.1 e<sup>-</sup> in some cases. Figure 3a shows the presence of a very strong N···H bond with a BO of 0.016 e<sup>-</sup> and a BL of 1.532 Å with H and N, pertaining to the same residue Arg5. The rest of the HBs are all O···H bonds with a much longer BL above 1.6 Å. The other important bonds are the two S-S bonds near 2.2 Å and short C-S covalent bonds at 1.75 Å involving S atom in Cys residues. These are all very stable bonds with BO values of 0.16 e<sup>-</sup> in different environments.





**Figure 3.** Distribution of BO vs. BL in the three 1FUV models: (a) dry 1FUV; (b) 1FUV with 80 H<sub>2</sub>O, (c) 1FUV with 80 H<sub>2</sub>O and 0.15 NaCl; (d). 1FUV with 100 H<sub>2</sub>O, and (e) 1FUV with 100 H<sub>2</sub>O and 0.15 NaCl.

*Solvated 1FUV with 100 H<sub>2</sub>O and salt ions:* Figure 3e shows the BO vs. BL plots for the largest and more complicated 1FUV model with 100 H<sub>2</sub>O molecules and 7 salt ions (4 Na and 3 Cl). The main difference is the appearance of the internal O-H covalent bond in H<sub>2</sub>O together with the new bonds (H-Cl, H-Na, Na-Cl and O-Cl) with the salt ions at much larger BL. Other changes are less remarkable, but it is notable that the single O-H bond and the strong N...H HB in Figure 3e have both disappeared, showing the changes in the atomic positions in residues Asp and Arg due to solvation. The appearance of many more O...H HBs with shorter HB distances is expected. In addition, there are some minor changes in the N-C bonds while the C-S and S-S bonds become weaker but remain at similar separations. The C-H bonds retain similar separation since the AAs in 1FUV remain relatively intact. The number of O-H bonds is observed to increase with concomitant increase in the BO values. This could imply the possible occurrence of protonation when some H atoms in water molecules rearrange themselves under increased complexity of the environment, although we do not yet have solid evidence for this interpretation. All these larger number of interactions mentioned above lead to its higher total bond order (TBO) in comparison to other four 1FUV models (shown in Table 1).

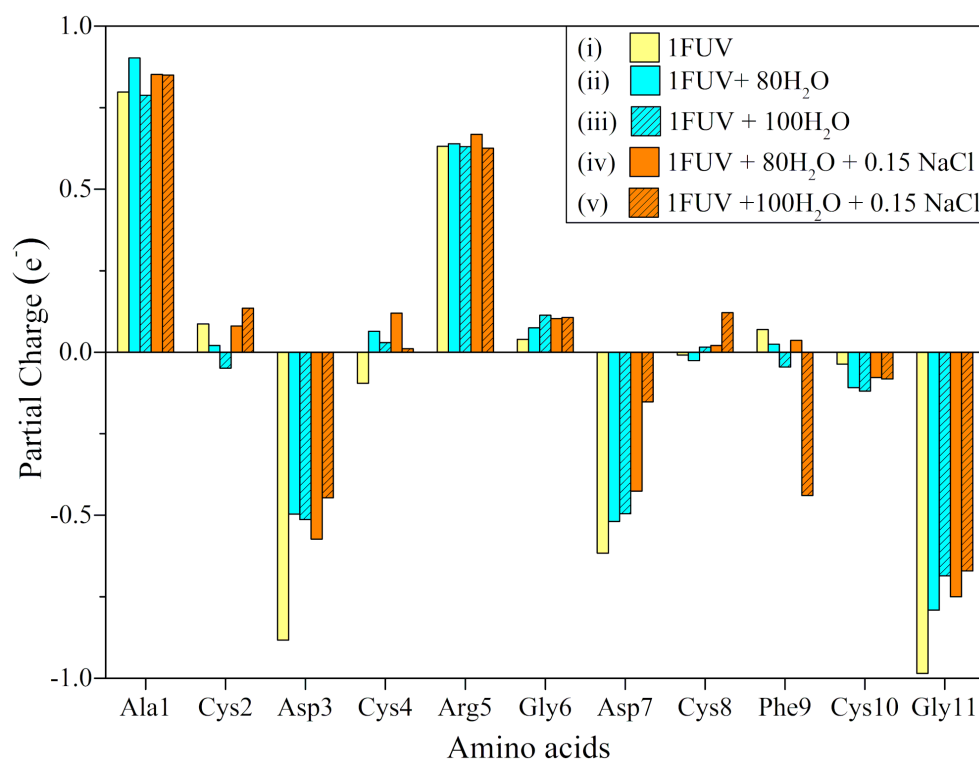
The ab initio quantum mechanical calculation allows for a very detailed analysis of the interatomic bonding in dry as well as solvated 1FUV, directly affecting the electronic

portion of the dielectric response of proteins. Quantitative evaluation of the HB in solvated model is extremely important and universally agreed by all researchers in biomolecular science, yet few have gone deep enough in the search for quantitative details, mostly characterizing the HBs qualitatively based on the separations between H and O or N without quantitative value for their bond strength, especially those using classical molecular dynamic methodology.

### 3.3. Role of Partial Charge

In addition to the dielectric spectra, the partial charge (PC) distribution is also crucial for determining and elucidating the electrostatic Coulomb interactions that play a major role in catalysis [71], drug engineering [72], etc. For different models studied here, the PC values are calculated based on the reliable orthogonalized linear combination of atomic orbitals (OLCAO) methodology [35] to understand the electrostatic effects but also the impact due to aqueous environments. We discuss here the comparison of PC distributions in the five 1FUV models under different local environment.

The calculated PCs on each of the 11 AAs in the five 1FUV models ((i) dry, (ii) solvated with 80 H<sub>2</sub>O (iii) 100 H<sub>2</sub>O, (iv) salted model of (ii), and (v) salted model of (iii)) are displayed in Figure 4. It can be observed that the variations of PC in five 1FUV models are minor for all amino acids with the exception of the negatively charged AAs (Asp3, Asp7) and Gly11, largely affected by the presence of the aqueous solution (water and ions). In particular, Phe9 residue in salted models with 100 H<sub>2</sub>O or pure 100 H<sub>2</sub>O (models (v) and (iii)) differs from other models in which its PC gains a small charge. This behavior is even more pronounced in model (v) when the PC of Phe9 is flipped from a negligible 0.07 e<sup>-</sup> in the dry case to a negative value of -0.44 e<sup>-</sup>. A close examination of such behavior reveals that the oxygen atom in the backbone of Phe9 can form a HB with one water molecule in all four solvent models, but its separation distance varies. In model (v), this HB has a much shorter distance of 1.65 Å when compared to 1.83, 1.78, and 2.52 Å for models (ii), (iii), and (iv), respectively.



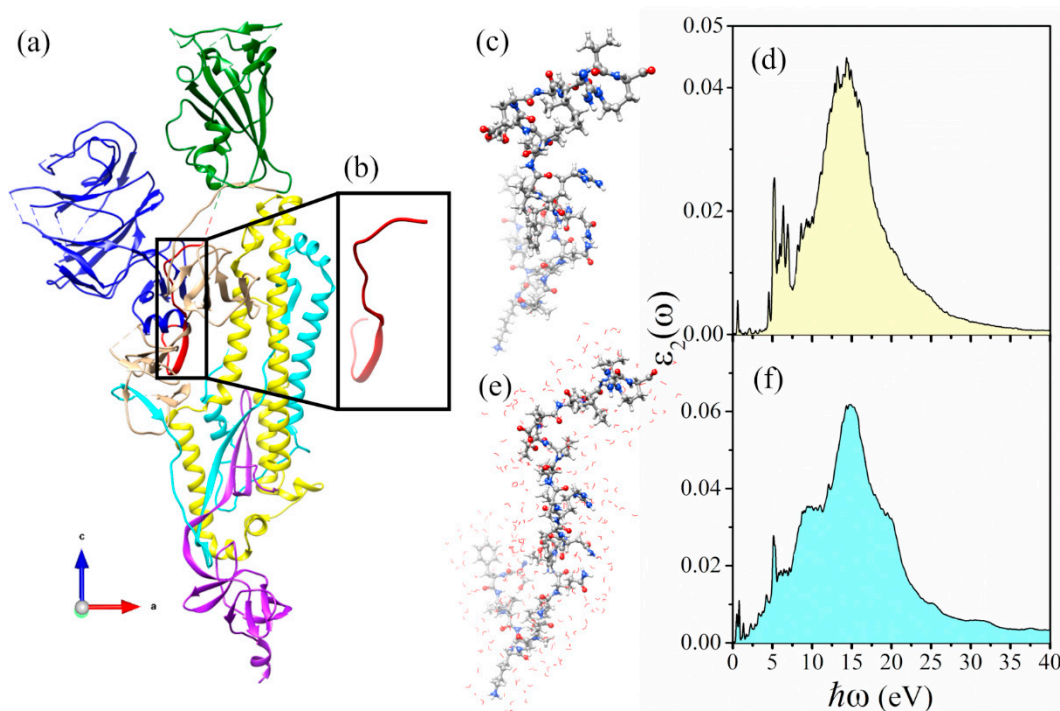
**Figure 4.** Partial charge (PC) distributions in terms of AAs for five 1FUV models (i)–(v) in various environments as shown in legends.

This finding clearly indicates that local environments have a direct effect on PC distributions. It is noteworthy that positive and negative PCs in Ala1 and the Gly11 at both terminals of 1FUV result from the  $\text{NH}_2^+$  in N-terminal and  $\text{COO}^-$  in C-terminal, respectively, and these opposite PCs neutralize each other.

#### 4. SD1 of SARS-CoV-2 Spike-Protein

The last example for the calculation of the dielectric function in proteins is a part of the spike protein in SARS-CoV-2 virus. The spike protein of SARS-CoV-2 is an ideal target for vaccine development and other therapeutic treatment due to its essential role in the virus life cycle [73,74]. The structure of spike protein is very complex as it exists in a trimeric form with each protomer consisting of two functional subunits, S1 and S2. The S1 subunit comprises a signal sequence (SS) at the N-terminal end, followed by NTD and RBD and two structurally conserved subdomains (SD1 and SD2). The S2 subunit is subdivided into a fusion peptide (FP), two heptad repeats (HR1 and 2), and a central helix (CH), a connector domain (CD), a transmembrane domain (TM), and the cytoplasmic tail (CT). Between the S1 and S2 subunits, there are two protease cleavage sites (S1/S2 and S2') [31].

The ab initio calculation of the electronic structure and interatomic bonding of this large biomolecular system has been already discussed [23]. SD1 is the smallest of the seven structural domains, having 24 AAs and 391 atoms in the dry environment. Unlike 1FUV, it has an elongated structure mimicking a very irregularly shaped mini-protein as shown in Figure 5a–c. It is therefore highly instructive to apply the QMRPA method to investigate its dielectric spectra of SD1. Other information on the electronic structure and intra- and inter-atomic bonding of SD1 can be found in Ref. [23].



**Figure 5.** (a) Chain A of spike protein (6VSB); (b) ribbon for SD1; (c) ball and stick model of SD1; (d)  $\epsilon_2(\omega)$  for dry SD1 model; (e) ball and stick model and (f)  $\epsilon_2(\omega)$  of SD1 with 300  $\text{H}_2\text{O}$  molecules.

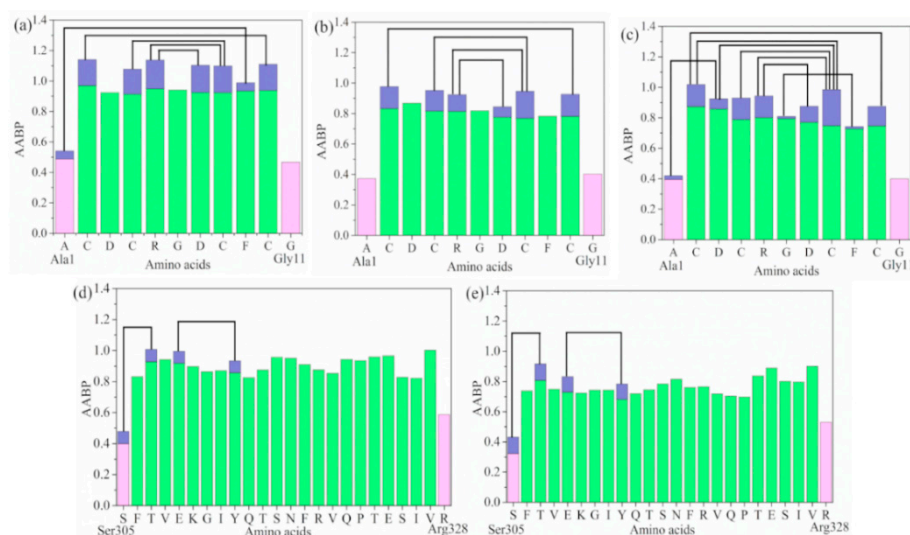
The solvation effects in SD1 have been addressed by adding 300 vicinal water molecules, as illustrated in Figure 5e. The solvated structure is then fully optimized by using VASP in the same way as the solvated model of 1FUV. The calculated dielectric absorption spectra for the dry and solvated SD1 are shown in Figure 5d,f. By comparing the dielectric spectra in the case of dry and solvated 1FUV models, shown in Figure 1d,j, we note the occurrence

of a broad peak at 14.4–15.2 eV originating from covalent interatomic bonding. In addition, we observe a peak at 5.2–5.7 eV, in cases of dry and solvated 1FUV and SD1 models.

### 5. Amino Acid Bond Pairing—AABP

All polypeptides are composed of an AA primary sequence. Because of the unique 3D structure of folded proteins, the AAs interact not only between nearest neighbors (NN) in the primary sequence, but also with other non-NN or off-diagonal or spatially not vicinal, i.e., non-local AAs. These non-local interactions between AAs along the primary sequence allow for the formation of AA-AA bond pairs (AABPs) (see Equation (3) in Method Section 2.3) [23,40].

The distribution of AABP and their characteristics in the seven domains of the spike protein in SARS-CoV-2 has been described in detail [23]. The same type of analysis is also performed for the 1FUV protein models and for the SD1 small protein models. Figure 6 shows the comparison of the calculated AABP distribution and the 3D bonding network in dry and solvated 1FUV and SD1. They differ substantially, mainly because of the differences in their structures. SD1 has a long, ribbon-like structure and naturally exhibits less off-diagonal AABP contribution (4 out of 24 AAs or 0.16%), while 1FUV is a rather compact protein with only 11 AAs, displaying a substantial contribution from AABP (8 out of 11 or 73%).



**Figure 6.** Distribution of AABP in (a) 1FUV; (b) 1FUV in 80 H<sub>2</sub>O molecules; (c) 1FUV in 100 H<sub>2</sub>O molecules; (d) SD1, and (e) SD1 in 300 H<sub>2</sub>O molecules. The color bar represents the following cases, light pink: sum AABP of AA with single nearest neighbor, green: sum AABP of AA with two nearest neighbors, light blue: AA with off-diagonal AABP. The black curve lines represent off-diagonal bonding between two AAs.

In Figure 6, we display and compare the results for the AABP values in dry and solvated modes for 1FUV and SD1. The main observations can be summarized as follows:

1. The AAs in 1FUV models have far more off-diagonal contributions to AABP than the AAs in SD1, as expected from the very different 3D structure of these two proteins.
2. The solvated models exhibit a reduced total AABP, mainly from the reduced NN interactions.
3. The five off-diagonal pairs in dry 1FUV are reduced to four pairs when solvated by 80 H<sub>2</sub>O molecules. The missing pair is between Ala1-Phe9. Surprisingly, when solvated by 100 H<sub>2</sub>O molecules, the number of pairs increases again to eight. The new pairs are Ala1-Asp3, Cys2-Cys8, Asp3-Cys8, Gly6-Phe9.
4. For SD1, the two off-diagonal pairs remain the same when solvated by 300 H<sub>2</sub>O molecules, but the total AABP values for all AAs are decreased by about 13% on average.

5. The decrease in total AABP in solvated models in comparison to the dry model is due to the interactions between AAs modified by the presence of H<sub>2</sub>O molecules.
6. It can then be concluded that the total AABP values can be changed in a rather complicated fashion depending on the nature of the protein and the amount of water molecules surrounding them.

## 6. Discussion

There is a growing interest in determining the details of the dielectric properties of proteins to better understand and estimate the molecular interactions involving proteins and polypeptides. However, the details of the dielectric response function depend on various factors such as the nature, size, and structure of the protein. In the present study, we report on the implementation of the QMRPA scheme to calculate the dielectric spectra of various biomolecular systems such as 1FUJ in five different environments and SD1 of SARS-CoV-2 spike protein in dry and solvated cases.

The present approach using *ab initio* calculations for the dielectric spectra in proteins can provide input spectral information for the calculation of the dispersion van der Waals interactions, while the bond order and the partial charge calculations can yield parametrizations for the electrostatic Coulomb interactions. Problems related to protein folding and stability, conformational change, mutation, glycosylation, etc., could be elucidated on a more solid basis involving molecular and atomic interactions that can be quantified by the *ab initio* calculations. However, the application of QMRPA methodology to larger proteins, while challenging due to computational limitations, is certainly not impossible. It is much more effective to use a QMRPA method to a selected number of specific larger macromolecules or semi-macromolecules to provide more accurate parameter-free data than just relying on standard atomic potential parametrizations.

Recently, we have succeeded in the largest *ab initio* quantum chemical computation to date on the S-protein by using a *divide and conquer* strategy for all subdomains of the spike protein available in 6VSB [23], while here the electronic dielectric properties of SD1, that act as a hinge point for the RBD in the down to up transitions [75] in dry and solvated environments have been investigated. The same type of calculations for other much larger subdomains are also possible to gain additional insights on the electrostatic interactions in the spike protein, and the results could be very useful for COVID-19 researchers working on detection techniques.

For example, a biosensor based on imaging ellipsometry was used to detect two neutralizing monoclonal antibodies and serial serum samples from SARS-CoV patients [76,77]. Ellipsometry is an optical technique closely related to the dielectric properties [77]. We speculate that the calculated dielectric spectra of the spike protein can be helpful in their design and implementation. On other hand, the electrostatic interactions have been demonstrated to play a major role in enhancing the binding affinity of RBD of SARS-CoV-2 to ACE2 as compared to SARS-CoV [78–81] and the polar solvation free energy of RBD-ACE2 complex is in fact quite sensitive to the dielectric spectra.

Our work on the RBD-ACE2 interface complex is in progress and will be reported elsewhere. We have high expectations that the calculated dielectric spectra based on a more realistic *ab initio* approach could be used to achieve a reasonable value for overall electrostatic interactions (Coulombic and polar solvation energies) as compared to the experimental data.

## 7. Conclusions

In conclusion, we have presented the dielectric spectra for two small proteins, 1FUJ and SD1, where our results indicate that the dielectric absorption spectra show distinct peaks at around 5 and 15 eV. The peaks at 15 eV are a more generic peak from the covalent interatomic bonding, whereas peaks at around 5 eV are more unique but are observed in both biomolecules. A succinct summary of the work presented above is as follows:



1. We introduced the QMRPA method for dielectric spectra for small proteins such as RGD (1FUUV) peptide and the SD1 subdomain of the spike protein of SARS-CoV-2 virus.
2. We pointed out the possible connections between atomic scale partial charges of AAs in proteins and their specific role in the electrostatic interaction.
3. We described the role of non-local AA-AA interactions via AABP values in the 3D structure of the protein comparing dry and solvated models.
4. We laid out the roadmap to use QMRPA method for applications to electrostatic interactions in spike protein and other biomolecular systems in general.

**Author Contributions:** W.-Y.C. and R.P. conceived the project. P.A. and W.-Y.C. performed the calculations, P.A. made most of the figures. W.-Y.C. and R.P. drafted the paper with inputs from P.A., B.J. All authors participated in the discussion and interpretation of the results. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project is funded by the National Science Foundation of USA: RAPID DMR/CMMT-2028803.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This research used the resources of the National Energy Research Scientific Computing Center supported by DOE under Contract No. DE-AC03-76SF00098 and the Research Computing Support Services (RCSS) of the University of Missouri System. This project is funded by the National Science Foundation of USA: RAPID DMR/CMMT-2028803. RP acknowledges funding from the Key project #12034019 of the National Natural Science Foundation of China.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Schutz, C.N.; Warshel, A. What are the dielectric “constants” of proteins and how to validate electrostatic models? *Proteins Struct. Funct. Bioinform.* **2001**, *44*, 400–417. [[CrossRef](#)] [[PubMed](#)]
2. Zhou, H.-X.; Pang, X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chem. Rev.* **2018**, *118*, 1691–1741. [[CrossRef](#)]
3. Woods, L.; Dalvit, D.A.R.; Tkatchenko, A.; Rodriguez-Lopez, P.; Rodriguez, A.W.; Podgornik, R. Materials perspective on Casimir and van der Waals interactions. *Rev. Mod. Phys.* **2016**, *88*, 045003. [[CrossRef](#)]
4. Dryden, D.M.; Hopkins, J.C.; Denoyer, L.K.; Poudel, L.; Steinmetz, N.F.; Ching, W.-Y.; Podgornik, R.; Parsegian, A.; French, R.H. van der Waals interactions on the mesoscale: Open-science implementation, anisotropy, retardation, and solvent effects. *Langmuir* **2015**, *31*, 10145–10153. [[CrossRef](#)]
5. Bai, C.; Warshel, A. Critical Differences Between the Binding Features of the Spike Proteins of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. B* **2020**, *124*, 5907–5912. [[CrossRef](#)]
6. Javidpour, L.; Bozic, A.; Naji, A.; Podgornik, R. Electrostatic interaction between SARS-CoV-2 virus and charged electret fibre. *arXiv* **2020**, arXiv:2012.07160.
7. McLaughlin, S. The electrostatic properties of membranes. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 113–136. [[CrossRef](#)] [[PubMed](#)]
8. Warshel, A.; Sharma, P.K.; Kato, M.; Parson, W.W. Modeling electrostatic effects in proteins. *Biochim. Biophys. Acta BBA-Proteins Proteom.* **2006**, *1764*, 1647–1676. [[CrossRef](#)]
9. Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the dielectric “constant” of proteins: Smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136. [[CrossRef](#)]
10. García-Moreno, B.E.; Dwyer, J.J.; Gittis, A.G.; Lattman, E.E.; Spencer, D.S.; Stites, W.E. Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophys. Chem.* **1997**, *64*, 211–224. [[CrossRef](#)]
11. Li, C.; Jia, Z.; Chakravorty, A.; Pahari, S.; Peng, Y.; Basu, S.; Koirala, M.; Panday, S.K.; Petukh, M.; Li, L. DelPhi suite: New developments and review of functionalities. *J. Comput. Chem.* **2019**, *40*, 2502–2508. [[CrossRef](#)]
12. Koehl, P. Electrostatics calculations: Latest methodological advances. *Curr. Opin. Struct. Biol.* **2006**, *16*, 142–151. [[CrossRef](#)] [[PubMed](#)]
13. Simonson, T. Electrostatics and dynamics of proteins. *Rep. Prog. Phys.* **2003**, *66*, 737. [[CrossRef](#)]
14. Gudarzi, M.M.; Aboutalebi, S.H. Self-consistent dielectric functions of materials: Toward accurate computation of Casimir–van der Waals forces. *Sci. Adv.* **2021**, *7*, eabg2272. [[CrossRef](#)]
15. Parsegian, V.; Ninham, B. Application of the Lifshitz theory to the calculation of van der Waals forces across thin lipid films. *Nature* **1969**, *224*, 1197–1198. [[CrossRef](#)] [[PubMed](#)]

16. Dill, K.A. Dominant forces in protein folding. *Biochemistry* **1990**, *29*, 7133–7155. [CrossRef] [PubMed]
17. Stöhr, M.; Tkatchenko, A. Quantum mechanics of proteins in explicit water: The role of plasmon-like solute-solvent interactions. *Sci. Adv.* **2019**, *5*, eaax0024. [CrossRef]
18. Velichko, E.; Baranov, M.; Mostepanenko, V. Change of sign in the Casimir interaction of peptide films deposited on a dielectric substrate. *Mod. Phys. Lett. A* **2020**, *35*, 2040020. [CrossRef]
19. Baranov, M.; Klimchitskaya, G.; Mostepanenko, V.; Velichko, E. Fluctuation-induced free energy of thin peptide films. *Phys. Rev. E* **2019**, *99*, 022410. [CrossRef]
20. Adhikari, P.; Wen, A.M.; French, R.H.; Parsegian, V.A.; Steinmetz, N.F.; Podgornik, R.; Ching, W.-Y. Electronic structure, dielectric response, and surface charge distribution of RGD (1FUUV) peptide. *Sci. Rep.* **2014**, *4*, 5605. [CrossRef]
21. Bloch, F. Über die quantenmechanik der elektronen in kristallgittern. *Z. Phys.* **1929**, *52*, 555–600. [CrossRef]
22. Ehrenreich, H.; Cohen, M.H. Self-consistent field approach to the many-electron problem. *Phys. Rev.* **1959**, *115*, 786–790. [CrossRef]
23. Ching, W.-Y.; Adhikari, P.; Jawad, B.; Podgornik, R. Ultra-Large-Scale Ab Initio Quantum Chemical Computation of Bio-Molecular Systems: The Case of Spike Protein of SARS-CoV-2 Virus. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1288–1301. [CrossRef]
24. Millefiori, S.; Alparone, A.; Millefiori, A.; Vanella, A. Electronic and vibrational polarizabilities of the twenty naturally occurring amino acids. *Biophys. Chem.* **2008**, *132*, 139–147. [CrossRef] [PubMed]
25. Swart, M.; Snijders, J.G.; van Duijnen, P.T. Polarizabilities of amino acid residues. *J. Comput. Methods Sci. Eng.* **2004**, *4*, 419–425. [CrossRef]
26. Abeyrathne, C.D.; Halgamuge, M.N.; Farrell, P.M.; Skafidas, E. An ab-initio computational method to determine dielectric properties of biological materials. *Sci. Rep.* **2013**, *3*, 1796. [CrossRef] [PubMed]
27. Stöhr, M.; Van Voorhis, T.; Tkatchenko, A. Theory and practice of modeling van der Waals interactions in electronic-structure calculations. *Chem. Soc. Rev.* **2019**, *48*, 4118–4154. [CrossRef] [PubMed]
28. 1FUUV Solution Structure of an RGD Peptide ISOMER-A. Available online: <https://www.rcsb.org/structure/1FUUV> (accessed on 20 March 2021).
29. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
30. Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham III, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41. [CrossRef]
31. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.-L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [CrossRef]
32. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
33. VASP—Vienna Ab initio Simulation Package. Available online: <https://www.vasp.at/> (accessed on 20 March 2021).
34. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865. [CrossRef]
35. Ching, W.-Y.; Rulis, P. *Electronic Structure Methods for Complex Materials: The Orthogonalized Linear Combination of Atomic Orbitals*; Oxford University Press: Oxford, UK, 2012.
36. Poudel, L.; Steinmetz, N.F.; French, R.H.; Parsegian, V.A.; Podgornik, R.; Ching, W.-Y. Implication of the solvent effect, metal ions and topology in the electronic structure and hydrogen bonding of human telomeric G-quadruplex DNA. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21573–21585. [CrossRef]
37. Poudel, L.; Twarock, R.; Steinmetz, N.F.; Podgornik, R.; Ching, W.-Y. Impact of hydrogen bonding in the binding site between capsid protein and MS2 bacteriophage ssRNA. *J. Phys. Chem. B* **2017**, *121*, 6321–6330. [CrossRef]
38. Eifler, J.; Podgornik, R.; Steinmetz, N.F.; French, R.H.; Parsegian, V.A.; Ching, W.Y. Charge distribution and hydrogen bonding of a collagen  $\alpha$ 2-chain in vacuum, hydrated, neutral, and charged structural models. *Int. J. Quantum Chem.* **2016**, *116*, 681–691. [CrossRef]
39. Adhikari, P.; Li, N.; Shin, M.; Steinmetz, N.F.; Twarock, R.; Podgornik, R.; Ching, W.-Y. Intra- and intermolecular atomic-scale interactions in the receptor binding domain of SARS-CoV-2 spike protein: Implication for ACE2 receptor binding. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18272–18283. [CrossRef]
40. Adhikari, P.; Ching, W.-Y. Amino acid interacting network in the receptor-binding domain of SARS-CoV-2 spike protein. *RSC Adv.* **2020**, *10*, 39831–39841. [CrossRef]
41. Mulliken, R.S. Electronic population analysis on LCAO–MO molecular wave functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840. [CrossRef]
42. Mulliken, R. Electronic population analysis on LCAO–MO molecular wave functions. II. Overlap populations, bond orders, and covalent bond energies. *J. Chem. Phys.* **1955**, *23*, 1841–1846. [CrossRef]
43. Dereka, B.; Yu, Q.; Lewis, N.H.; Carpenter, W.B.; Bowman, J.M.; Tokmakoff, A. Crossover from hydrogen to chemical bonding. *Science* **2021**, *371*, 160–164. [CrossRef] [PubMed]
44. Martin, P.C. Sum rules, Kramers-Kronig relations, and transport coefficients in charged systems. *Phys. Rev.* **1967**, *161*, 143. [CrossRef]

45. Ching, W.; Xu, Y.; Zhao, G.-L.; Wong, K.; Zandiehnam, F. Electronic structure and excitonic-enhanced superconducting mechanism in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ . *Phys. Rev. Lett.* **1987**, *59*, 1333. [[CrossRef](#)]
46. Zhao, G.-L.; Xu, Y.; Ching, W.; Wong, K. Theoretical calculation of optical properties of Y-Ba-Cu-O superconductors. *Phys. Rev. B* **1987**, *36*, 7203. [[CrossRef](#)]
47. Ching, W.; Huang, M.-Z.; Xu, Y.-N.; Harter, W.; Chan, F. First-principles calculation of optical properties of C 60 in the fcc lattice. *Phys. Rev. Lett.* **1991**, *67*, 2045. [[CrossRef](#)] [[PubMed](#)]
48. Xu, Y.-N.; Huang, M.-Z.; Ching, W. Optical properties of superconducting K 3 C 60 and insulating K 6 C 60. *Phys. Rev. B* **1991**, *44*, 13171. [[CrossRef](#)]
49. Baral, K.; Adhikari, P.; Ching, W.Y. Ab initio Modeling of the Electronic Structures and Physical Properties of  $\text{a-Si}_{1-x}\text{Ge}_x\text{O}_2$  Glass ( $x = 0$  to 1). *J. Am. Ceram. Soc.* **2016**, *99*, 3677–3684. [[CrossRef](#)]
50. Adhikari, P.; Xiong, M.; Li, N.; Zhao, X.; Rulis, P.; Ching, W.-Y. Structure and electronic properties of a continuous random network model of an amorphous zeolitic imidazolate framework (a-ZIF). *J. Phys. Chem. C* **2016**, *120*, 15362–15368. [[CrossRef](#)]
51. Adhikari, P.; Li, N.; Rulis, P.; Ching, W.-Y. Deformation behavior of an amorphous zeolitic imidazolate framework—from a supersoft material to a complex organometallic alloy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 29001–29011. [[CrossRef](#)]
52. Ouyang, L.; Randaccio, L.; Rulis, P.; Kurmaev, E.; Moewes, A.; Ching, W. Electronic structure and bonding in vitamin B12, cyanocobalamin. *J. Mol. Struct. THEOCHEM* **2003**, *622*, 221–227. [[CrossRef](#)]
53. Ouyang, L.; Rulis, P.; Ching, W.; Nardin, G.; Randaccio, L. Accurate redetermination of the X-ray structure and electronic bonding in adenocobalamin. *Inorg. Chem.* **2004**, *43*, 1235–1241. [[CrossRef](#)] [[PubMed](#)]
54. Rulis, P.; Ouyang, L.; Ching, W. Electronic structure and bonding in calcium apatite crystals: Hydroxyapatite, fluorapatite, chlorapatite, and bromapatite. *Phys. Rev. B* **2004**, *70*, 155104. [[CrossRef](#)]
55. Rulis, P.; Yao, H.; Ouyang, L.; Ching, W. Electronic structure, bonding, charge distribution, and x-ray absorption spectra of the (001) surfaces of fluorapatite and hydroxyapatite from first principles. *Phys. Rev. B* **2007**, *76*, 245410. [[CrossRef](#)]
56. Kahr, B.; Freudenthal, J.; Phillips, S.; Kaminsky, W. Herapathite. *Science* **2009**, *324*, 1407. [[CrossRef](#)] [[PubMed](#)]
57. Liang, L.; Rulis, P.; Kahr, B.; Ching, W. Theoretical study of the large linear dichroism of herapathite. *Phys. Rev. B* **2009**, *80*, 235132. [[CrossRef](#)]
58. French, R.H.; Parsegian, V.A.; Podgornik, R.; Rajter, R.F.; Jagota, A.; Luo, J.; Asthagiri, D.; Chaudhury, M.K.; Chiang, Y.-M.; Granick, S.; et al. Long range interactions in nanoscale science. *Rev. Mod. Phys.* **2010**, *82*, 1887. [[CrossRef](#)]
59. Hopkins, J.C.; Dryden, D.M.; Ching, W.-Y.; French, R.H.; Parsegian, V.A.; Podgornik, R. Dielectric response variation and the strength of van der Waals interactions. *J. Colloid Interface Sci.* **2014**, *417*, 278–284. [[CrossRef](#)]
60. Schimelman, J.B.; Dryden, D.M.; Poudel, L.; Krawiec, K.E.; Ma, Y.; Podgornik, R.; Parsegian, V.A.; Denoyer, L.K.; Ching, W.-Y.; Steinmetz, N.F.; et al. Optical properties and electronic transitions of DNA oligonucleotides as a function of composition and stacking sequence. *Phys. Chem. Chem. Phys.* **2015**, *17*, 4589–4599. [[CrossRef](#)]
61. Gong, Y.; Adhikari, P.; Liu, Q.; Wang, T.; Gong, M.; Chan, W.-L.; Ching, W.-Y.; Wu, J. Designing the interface of carbon nanotube/biomaterials for high-performance ultra-broadband photodetection. *ACS Appl. Mater. Interfaces* **2017**, *9*, 11016–11024. [[CrossRef](#)]
62. Gong, M.; Adhikari, P.; Gong, Y.; Wang, T.; Liu, Q.; Kattel, B.; Ching, W.Y.; Chan, W.L.; Wu, J.Z. Polarity-Controlled Attachment of Cytochrome C for High-Performance Cytochrome C/Graphene van der Waals Heterojunction Photodetectors. *Adv. Funct. Mater.* **2018**, *28*, 1704797. [[CrossRef](#)]
63. Donev, R. *Protein and Peptide Nanoparticles for Drug Delivery*; Academic Press: Waltham, MA, USA, 2015.
64. Wang, F.; Li, Y.; Shen, Y.; Wang, A.; Wang, S.; Xie, T. The functions and applications of RGD in tumor therapy and tissue engineering. *Int. J. Mol. Sci.* **2013**, *14*, 13447–13462. [[CrossRef](#)]
65. Avraamides, C.J.; Garmy-Susini, B.; Varner, J.A. Integrins in angiogenesis and lymphangiogenesis. *Nat. Rev. Cancer* **2008**, *8*, 604–617. [[CrossRef](#)]
66. Garanger, E.; Boturyn, D.; Dumy, P. Tumor targeting with RGD peptide ligands—design of new molecular conjugates for imaging and therapy of cancers. *Anti-Cancer Agents Med. Chem.* **2007**, *7*, 552–558. [[CrossRef](#)]
67. Metcalfe, A.D.; Ferguson, M.W. Tissue engineering of replacement skin: The crossroads of biomaterials, wound healing, embryonic development, stem cells and regeneration. *J. R. Soc. Interface* **2007**, *4*, 413–437. [[CrossRef](#)]
68. Zhou, Y.; Chakraborty, S.; Liu, S. Radiolabeled cyclic RGD peptides as radiotracers for imaging tumors and thrombosis by SPECT. *Theranostics* **2011**, *1*, 58. [[CrossRef](#)]
69. Meyers, S.R.; Grinstaff, M.W. Biocompatible and bioactive surface modifications for prolonged in vivo efficacy. *Chem. Rev.* **2012**, *112*, 1615–1632. [[CrossRef](#)]
70. Bellis, S.L. Advantages of RGD peptides for directing cell association with biomaterials. *Biomaterials* **2011**, *32*, 4205–4210. [[CrossRef](#)]
71. Zheng, H.; Yang, S.-J.; Zheng, Y.-C.; Cui, Y.; Zhang, Z.; Zhong, J.-Y.; Zhou, J. Electrostatic Effect of Functional Surfaces on the Activity of Adsorbed Enzymes: Simulations and Experiments. *ACS Appl. Mater. Interfaces* **2020**, *12*, 35676–35687. [[CrossRef](#)] [[PubMed](#)]
72. Ionescu, C.-M.; Sehnal, D.; Falginella, F.L.; Pant, P.; Pravda, L.; Bouchal, T.; Vařeková, R.S.; Geidl, S.; Koča, J. AtomicCharge-Calculator: Interactive web-based calculation of atomic charges in large biomolecular complexes and drug-like molecules. *J. Cheminformatics* **2015**, *7*, 50. [[CrossRef](#)] [[PubMed](#)]

73. Gao, Q.; Bao, L.; Mao, H.; Wang, L.; Xu, K.; Yang, M.; Li, Y.; Zhu, L.; Wang, N.; Lv, Z. Development of an inactivated vaccine candidate for SARS-CoV-2. *Science* **2020**, *369*, 77–81. [[CrossRef](#)]
74. Chi, X.; Yan, R.; Zhang, J.; Zhang, G.; Zhang, Y.; Hao, M.; Zhang, Z.; Fan, P.; Dong, Y.; Yang, Y.; et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **2020**, *369*, 650–655. [[CrossRef](#)] [[PubMed](#)]
75. Henderson, R.; Edwards, R.J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S.M.; Kopp, M.; Li, D.; Parks, R.; Hsu, A.L.; et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925–933. [[CrossRef](#)] [[PubMed](#)]
76. Qi, C.; Duan, J.-Z.; Wang, Z.-H.; Chen, Y.-Y.; Zhang, P.-H.; Zhan, L.; Yan, X.-Y.; Cao, W.-C.; Jin, G. Investigation of interaction between two neutralizing monoclonal antibodies and SARS virus using biosensor based on imaging ellipsometry. *Biomed. Microdevices* **2006**, *8*, 247–253. [[CrossRef](#)] [[PubMed](#)]
77. Orooji, Y.; Sohrabi, H.; Hemmat, N.; Oroojalian, F.; Baradaran, B.; Mokhtarzadeh, A.; Mohaghegh, M.; Karimi-Maleh, H. An overview on SARS-CoV-2 (COVID-19) and other human coronaviruses and their detection capability via amplification assay, chemical sensing, biosensing, immunosensing, and clinical assays. *Nano-Micro Lett.* **2021**, *13*, 18. [[CrossRef](#)] [[PubMed](#)]
78. Spinello, A.; Saltalamacchia, A.; Magistrato, A. Is the rigidity of SARS-CoV-2 spike receptor-binding motif the hallmark for its enhanced infectivity? Insights from all-atom simulations. *J. Phys. Chem. Lett.* **2020**, *11*, 4785–4790. [[CrossRef](#)] [[PubMed](#)]
79. Xie, Y.; Karki, C.B.; Du, D.; Li, H.; Wang, J.; Sobitan, A.; Teng, S.; Tang, Q.; Li, L. Spike proteins of SARS-CoV and SARS-CoV-2 utilize different mechanisms to bind with human ACE2. *Front. Mol. Biosci.* **2020**, *7*, 392. [[CrossRef](#)]
80. Peng, C.; Zhu, Z.; Shi, Y.; Wang, X.; Mu, K.; Yang, Y.; Zhang, X.; Xu, Z.; Zhu, W. Computational Insights into the Conformational Accessibility and Binding Strength of SARS-CoV-2 Spike Protein to Human Angiotensin-Converting Enzyme 2. *J. Phys. Chem. Lett.* **2020**, *11*, 10482–10488. [[CrossRef](#)]
81. Amin, M.; Sorour, M.K.; Kasry, A. Comparing the binding interactions in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. Lett.* **2020**, *11*, 4897–4900. [[CrossRef](#)]