

# SARS-CoV-2 virion stabilization by Zn binding

S. Morante<sup>a)b)\*</sup>, G. La Penna<sup>c)b)</sup>, G.C. Rossi<sup>a)b)d)</sup>, F. Stellato<sup>a)b)</sup>

<sup>a)</sup> Dipartimento di Fisica, Università di Roma “*Tor Vergata*”, Via della Ricerca Scientifica - 00133 Roma, Italy

<sup>b)</sup> INFN, Sezione di Roma 2, Via della Ricerca Scientifica - 00133 Roma, Italy

<sup>c)</sup> CNR, Institute for Chemistry of Organometallic Compounds - 50019 Sesto Fiorentino, Italy

<sup>d)</sup> Centro Fermi - Museo Storico della Fisica e Centro Studi e Ricerche “Enrico Fermi”, Piazza del Viminale - 00184 Roma, Italy

## Abstract

Zinc plays a crucial role in the process of virion maturation inside the host cell. The accessory Cys-rich proteins expressed in SARS-CoV-2 by genes ORF7a and ORF8 are likely involved in zinc binding and in interactions with cellular antigens activated by extensive disulfide bonds. In this report we provide a proof of concept for the feasibility of a structural study of orf7a and orf8 proteins. A conceivable hypothesis is that lack of cellular zinc, or substitution thereof, might lead to a significant slowing down of viral maturation.

---

\*Corresponding author email: [silvia.morante@roma2.infn.it](mailto:silvia.morante@roma2.infn.it)

## 1 Introduction

The SARS-CoV-2 open reading frames ORF7a and ORF8 code for virion non-structural, called accessory<sup>1</sup>, proteins of yet unknown function (orf7a and orf8, respectively, hereafter). The protein orf7a is common to all SARS-CoV type coronaviruses and highly conserved<sup>2</sup>, while orf8 is remarkably different from proteins coded by genes ORF8 and ORF8b of human SARS-CoV<sup>3</sup>.

The orf7a protein of SARS-CoV-1 has been shown to interact with several host proteins<sup>4</sup>. An analogous situation occurs in the case of the very similar SARS-CoV-2 orf7a protein<sup>5</sup>. The most supported hypothesis proposed for orf7a protein function is the interference with virion budding tethering<sup>6</sup> operated by cellular antigens<sup>7</sup>. Indeed, orf7a is expressed in the host cell to inhibit the intracellular (at endoplasmic membrane) process of virion immobilization before and after virion vesiculation. On the basis of structural similarities<sup>8</sup> we argue that also protein orf8 can be involved in the same inhibition process, strengthening the inhibition of virion immobilization. This conjecture is in agreement with Ref. 5 where it is suggested that orf8 plays a role in vesicle trafficking and in endoplasmic reticulum protein quality control, thus favouring the reconfiguration of ER/Golgi trafficking during coronavirus infection.

Virion tethering is mainly due to proteins of the tetherin family, also known as bone marrow stromal antigen 2 (BST2) or cluster of differentiation 317 (CD317). BST2 is expressed in many cells in the interferon-dependent antiviral response pathway. The mechanism of tethering involves tetherin protein dimerization via formation of extended disulfide bonds within the coiled coil region<sup>9</sup>. This step is known to be strongly influenced by divalent cations involved in Cys binding. Among these ions, the most available in cells is Zn<sup>2+</sup>.

A timely computational search for therapeutic targets of SARS-CoV-2 found the orf7a-BST2 complex as a potential target to be addressed with structural studies<sup>10</sup>.

The orf7a and orf8 protein sequences (both 121 amino acids long) hint at a high Zn binding propensity, as they display 6 Cys/3 His and 7 Cys/4 His side-chains, respectively, with motifs that are able to bind Zn, thus forming multiple zinc finger (ZF) domains. The relevance of Zn in the viral replication process has been widely investigated<sup>11,12</sup> and, indeed, two ZF domains have been discovered in the nucleocapsid protein of HIV-1 (Ncp7)<sup>13,14</sup>. Two such similar Zn domains have been also identified<sup>15</sup> along the highly conserved, within the SARS-CoV family<sup>3</sup>, nsp12 polymerase.

## 2 Method and Results

The pivotal role of Zn in SARS-CoV replication has been demonstrated by inhibition of RNA polymerase activity both *in vitro* and in cell culture<sup>16</sup>.

### Zinc Finger Domains

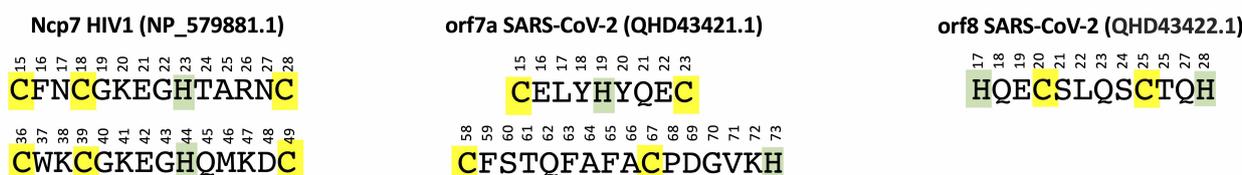


Figure 1: ZF domains in NP\_579881 (Ncp7 of HIV-1) (left), QHD43421.1 (orf7a of SARS-CoV-2) (center) and QHD43422.1 (orf8 of SARS-CoV-2) (right). Cys are highlighted in yellow and His in green.

Comparing with the well assessed ZF domains of Ncp7 CX<sub>2</sub>CX<sub>4</sub>HX<sub>4</sub>C (located from aa 15 to aa 28 and from aa 36 to aa 49), one identifies the similar CX<sub>3</sub>HX<sub>3</sub>C and CX<sub>8</sub>CX<sub>5</sub>H motifs in orf7a (located from aa

15 to aa 23 and from aa 56 to aa 73, respectively) and the  $\text{HX}_2\text{CX}_4\text{CX}_2\text{H}$  motif in orf8 (located from aa 17 to aa 28). The patterns of these motifs are shown in Fig 1.

UNIPROT ([https://www.uniprot.org/help/zn\\_fing](https://www.uniprot.org/help/zn_fing)) reports a large variety of ZF domains among which also the Cys-rich patterns of orf7a and orf8 can be accommodated.

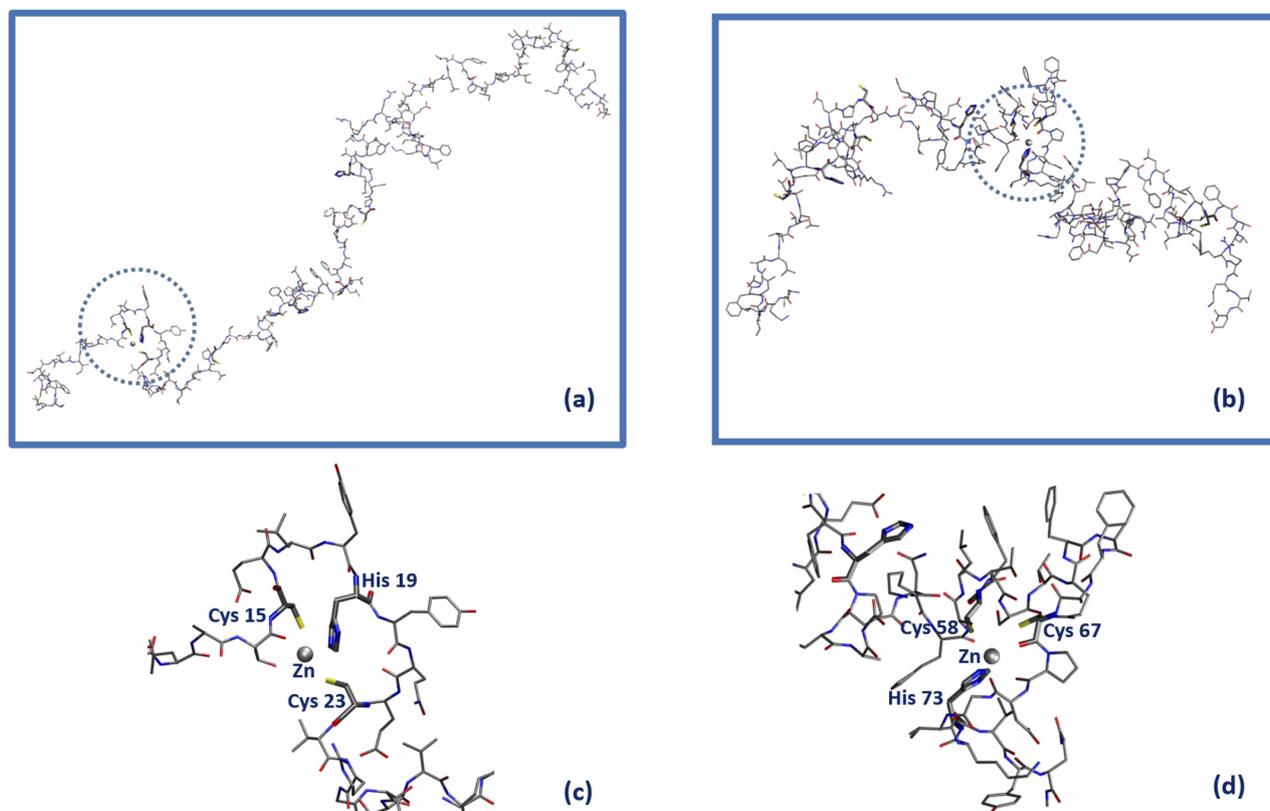


Figure 2: Sketch of the orf7a domains capable of hosting a  $\text{Zn}^{2+}$  ion (gray sphere) within the **CELYHYQEC** 15-23 segment (Zn-CHC, left) and the **CFSTQFAFACPDGVKH** 58-73 segment (Zn-CCH, right). The amino acid residues belonging to the first Zn coordination shell are highlighted in bold along the sequence. Panels (c) and (d) represent the blow-up of the local  $\text{Zn}^{2+}$  binding sites Cys(15)-His(19)-Cys(23) circled in panel (a) and Cys(58)-Cys(67)-His(73) (circled in panel (b)), respectively.

From a structural point of view it is not difficult to come up with a plausible protein structure capable of hosting a  $\text{Zn}^{2+}$  ion. Focusing on orf7a as an example, one can proceed to construct such structures by first building an atomistic model of the protein in a random coil configuration. After simulating self-avoiding random walks<sup>17,18</sup> of the chain in vacuum by randomly changing the dihedral angles (not involving H atoms) of all residues (except  $\Phi$  in Pro), for a total of 445 dihedral angles, we proceeded by monitoring along the trajectory the distance  $d$  between pairs of  $\text{S}\gamma(\text{Cys})$ - $\text{S}\gamma(\text{Cys})$  atoms,  $d(\text{CC})$ , and  $\text{S}\gamma(\text{Cys})$ - $\text{N}\epsilon(\text{His})$  atoms,  $d(\text{CH})$ . All His side-chains are neutral and protonated at  $\text{N}\delta$ . We collected a trajectory of 400 random configurations, with consecutive configurations separated by 500 attempted torsional moves. Acceptance ratio was about 0.5. Two chain configurations where three among all of the possible  $d(\text{CC})$  and  $d(\text{CH})$  distances are smaller than 1.2 nm were selected. This condition was found to be satisfied by the triplets Cys(15)...His(19)...Cys(23) in one case and Cys(58)...Cys(67)...His(73) in a second case.  $\text{Zn}^{2+}$  ions were then inserted at the geometric center of each of the two selected triplets of atoms. A cationic dummy atom model for  $\text{Zn}^{19}$  and the PARM14 Amber force-field<sup>20</sup> for the protein were used. We took a distance cut-off for non-bonding interactions of 0.5 nm, with Coulomb interactions shifted and damped to achieve local neutralization<sup>21</sup>. Starting from the selected configurations, we relaxed the systems by classical molecular dynamics (MD). At the beginning, harmonic forces with  $k = 10 \text{ kcal/mol/\AA}^2$  were added to bring the

distances between Zn and the three atoms [ $S\gamma(\text{Cys}(15))$ ,  $N\epsilon(\text{His}(19))$ ,  $S\gamma(\text{Cys}(23))$ ] in the first case (site Zn-CHC) and [ $S\gamma(\text{Cys}(58))$ ,  $S\gamma(\text{Cys}(67))$ ,  $N\epsilon(\text{His}(73))$ ] in the second (site Zn-CCH) from the initial values down to 2.5 Å in 10 ps. The final configurations of sites Zn-CHC (segment S1: 15-23) and Zn-CCH (segment S2: 58-73) are displayed in Figs. 2, left and right panels, respectively. All calculations were performed with the LAMMPS code<sup>22</sup>.

We interpret the fact that we have been able to get Zn-CHC and Zn-CCH stable configurations already with very simple classical simulation techniques, as a strong indication of the high propensity to give rise to ZF domains as soon as the occurrence frequency of Cys and His along the protein (or in its environment) is sufficiently high.

## 2.1 Refining the Zn binding sites

In the previous section by means of empirical non-polarizable force fields we have built candidates Zn binding sites for the orf7a protein of SARS-CoV-2. A more realistic description of Zn–protein interactions, especially when Zn-S bonds occur, requires polarizable or reactive force fields<sup>23,24</sup>. Semiempirical quantum mechanics methods, like those driven by the density-functional tight-binding approximated (DFTB) hamiltonian<sup>25</sup>, include polarization and charge transfer in the calculation of atomic forces and allow energy minimization of systems of thousands of atoms within periodic boundary conditions<sup>26</sup>. In this subsection we explore the energy profile around the previously selected configurations, using a DFTB hamiltonian.

Since the candidate Zn binding sites displayed in Fig. 2 involve a limited protein region, we built model configurations of the systems of interest by truncating the protein around the ZF domains involved in Zn binding. We choose for the Cys(15)-His(19)-Cys(23) and Cys(58)-Cys(67)-His(73) binding sites the Ala<sub>13</sub>...Gly<sub>26</sub> and the Asp<sub>51</sub>...Val<sub>74</sub> segments, called for short S1 and S2, respectively, in the following. The N- and C-termini of the segments were capped with acetyl and N-methyl groups, respectively. Cys binding side-chains were assumed as deprotonated according to NMR measurements of ZF domains<sup>27</sup>.

For each one of the two segments we constructed a simulation box by inserting the corresponding amino acids together with Zn into an orthorhombic cell whose volume is minimized by rotating the protein segments. The cell was then filled of TIP3P water molecules<sup>28</sup> with the help of the “solvate” utility of VMD<sup>29</sup>. At this point the energy of the system was first minimized for as long as 2000 conjugate gradient (CG) steps by moving water molecules and capping groups only. The same force field as the one we used in the previous classical self-avoiding random walk simulation was employed here. The final configurations were then used as the initial configurations in the CG energy minimization performed with the DFTB hamiltonian. The DFTB+ code<sup>26</sup> was used with the znorg-0-1 parametrization<sup>30</sup>. The energy is assumed to be in a minimum when the maximal force component was lower than 0.01 Ha/Bohr. We allowed for a maximum of 150 CG steps.

In order to estimate the formation energy of the Zn binding sites in the minimized configurations we obtained, we compare the former with two new systems that are obtained by exchanging the position of the Zn ion with that of the O of a water molecule lying far from the protein side-chains, upon transferring the entire water molecule where the Zn was located. After the exchange, the Cys side-chains turn out to be protonated. The exchange of Zn with a water molecule is represented by the chemical reaction



where  $\text{PH}_2$  is the protein segment with protonated Cys side-chains (when not bound to Zn),  $\text{ZnP}$  represents the complex formed by Zn and the deprotonated protein.

The charge  $q$  of the protein is -3 and -2 for the S1 and S2 segments, respectively. The energy of each complex was therefore corrected for the Makov–Payne term<sup>31</sup>, to take into account the difference in the cell net charge from the initial to the final state of the reaction (1).

The uncorrected energy of reactants is computed as the lowest energy of the system with Zn exchanged with a bulk water molecule. The energy of the product in reaction (1),  $\text{ZnP}$ , is that of the complex in the

Table 1: Geometry of the first Zn coordination shell of the orf7a S1 segment at the end of the DFTB minimization. We report atoms lying within 3 Å from the metal.

Residue	Atom	Distance [Å]
His 19	N	1.97
Cys 23	S	2.29
Wat	O	1.99
Wat	O	2.04

Table 2: Geometry of the first Zn coordination shell of the orf7a S2 segment at the end of the DFTB minimization. We report atoms lying within 3 Å from the metal.

Residue	Atom	Distance [Å]
Phe 54	O	2.30
Cys 58	S	2.78
Cys 67	S	2.38
Pro 68	O	2.32
Asp 69	O	2.23
His 73	N	2.00

box. The energy of hydrated proton is, as usual, taken from experiments<sup>32</sup>. After correcting for the Makov–Payne term<sup>31</sup> one gets the numbers reported in Table 3. We stress that both numbers are negative, meaning that Zn is bound to the ZF domains of the segments S1 and S2.

The lowest energy configurations obtained in the DFTB approximation are displayed in Fig 3, panels (a) and (b). The coordination geometry around the Zn binding sites at the end of the DFTB minimization are reported in Tables 1 and 2 where the distances of the atoms lying within 3 Å from the metal are reported. It is worth noticing that approximately the arrangement of the first shell Zn ligands is tetrahedral in the case of S1 and octahedral in the case of S2.

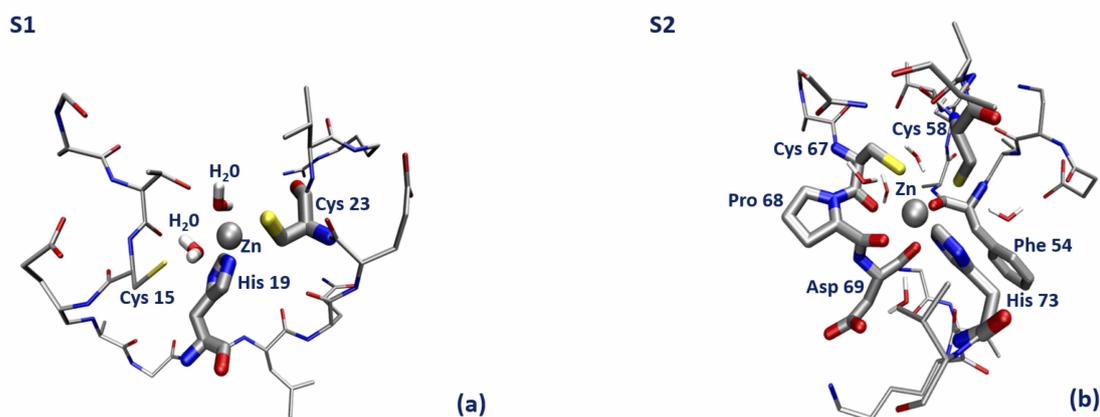


Figure 3: Sketch of the structure of the orf7a S1 (panel (a)) and S2 (panel (b)) segments at the end of the DFTB minimization.

We end with a few remarks. First of all, we observe that the mechanical constraints acting within the bent protein regions are not sufficient to break the Zn bond. Secondly, we see that, while in the S1 Zn binding site  $S_{\gamma}$ (Cys 15) is replaced by a water molecule (see Fig. 3 panel (a)), the structure of the S2 Zn binding site is remarkably stable, as a consequence of the small probability for a water molecule to enter the first Zn coordination shell. The structural stability of the Zn bonds is ensured by presence of 2-3 carbonyl groups that keep Zn bound to the two  $S_{\gamma}$  atoms on the opposite side, with the His side-chain unperturbed by the protein small relaxation. Finally, we stress that the formation energy of the Zn bound complexes is

dominated by the negative empirical contribution of the extraction of two protons into bulk water.

Table 3: Energy variation ( $\Delta E$ , kJ/mol) between the initial and final minimized configurations in the reaction 1 for our models of the binding sites S1: Cys(15)-His(19)-Cys(23) and S2: Cys(58)-Cys(67)-His(73).

Segment	$\Delta E$
S1	-376
S2	-1242

### 3 Discussion

With the help of classical MD simulations complemented with semiempirical quantum mechanics methods we have provided convincing evidence for the key role played by  $Zn^{2+}$  in stabilizing the orf7a (and possibly also) orf8 protein. Our working assumption is that orf/a (and/or orf8) can make  $Zn^{2+}$  available to BST2. The stability of Zn-S bonds, with S belonging to reduced sulfide groups, can in turn favour the breaking of some of the disulphide bridges of the folded BST2 tetherin, allowing the formation of inactivated orf7a/BST2 (and/or orf8/BST2) complexes with inhibition of the BST2 antiviral activity.

Based on the above considerations, we would like to argue that, although the affinity of Zn for the ZF domains is large (and larger than for other divalent cations of similar size available in cells, like  $Mg^{2+}$ ), one might think of displacing  $Zn^{2+}$  upon altering concentration in the host cells by temporary Zn deprivation and Mg augmentation. Indeed, situations have been identified in which, depending on the specific amino acids forming the metal coordination site, proteins can preferentially bind  $Mg^{2+}$  over  $Zn^{2+}$ <sup>33</sup>.

Thus, cellular Zn deprivation, with Zn replacement by Mg and/or specific drugs based on Ag(I) or Au(I)<sup>34</sup>, might result in a significant slowing down of viral replication, owing to the inhibition of orf7a/BST2 and orf8/BST2 complexes formation. Circumstantial support for this hypothesis comes from the well-known, important role played by Zn in inflammation<sup>35</sup>.

According to the scenario we have described, we are engaged in the long-term project of producing atomistic models of the orf7a/BST2 and orf8/BST2 systems, that can serve as reliable templates in the analysis of the forthcoming experiments aimed at unveiling the formation pathways and the detailed structure of these complexes.

The comparison to known related structures, *i.e.* the N-terminal ectodomain of orf7a of SARS-CoV-2 in the absence of Zn (PDB 6W37) and Ncp7 of HIV-1 in the presence of  $Zn^{13}$  (PDB 1ESK), will be the basis to model BST2 disulfide bond plasticity induced by modulation of orf7a/orf8 Zn binding.

The resulting predicted structures will also be useful for the forthcoming experimental studies aimed at validating (or disproving) the function of orf7a/orf8 proteins as diversion tools of the interferon-dependent antiviral response.

### References

- [1] *Retroviruses*, ed. J. M. Coffin, S. H. Hughes and V. H. E., Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 1997.
- [2] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng and T. Jiang, *Cell Host Microbe*, 2020, **27**, 325–328.
- [3] J. Xu, S. Zhao, T. Teng, A. E. Abdalla, W. Zhu, L. Xie, Y. Wang and X. Guo, *Viruses*, 2020, **12**, 244.
- [4] N. Vasilenko, I. Moshynskyy and A. Zakhartchouk, *Viol. J.*, 2010, **7**, 31–36.

- [5] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney *et al.*, *Nature*, 2020, 1–13.
- [6] J. S. Bonifacino and B. S. Glick, *Cell*, 2004, **116**, 153–166.
- [7] J. K. Taylor, C. M. Coleman, S. Postel, J. M. Sisk, J. G. Bernbaum, T. Venkataraman, E. J. Sundberg and M. B. Frieman, *J. Virol.*, 2015, **89**, 11820–11833.
- [8] M. Swiecki, N. S. Omattage and T. J. Brett, *Molecular immunology*, 2013, **54**, 132–139.
- [9] A. Le Tortorec, S. Willey and S. J. D. Neil, *Viruses*, 2011, **3**, 520–540.
- [10] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen and H. Li, *Acta Pharm. Sin. B*, 2020.
- [11] U. C. Chaturvedi and R. Shrivastava, *FEMS Immun. Med. Microbiol.*, 2005, **43**, 105–114.
- [12] C. T. Chasapis, *J. Inorg. Biochem.*, 2018, **186**, 157–161.
- [13] N. Morellet, H. Déméné, V. Teilleux, T. Huynh-Dinh, H. de Rocquigny, M.-C. Fournié-Zaluski and B. P. Roques, *J. Mol. Biol.*, 1998, **283**, 419–434.
- [14] J. Guo, T. Wu, J. Anderson, B. F. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson and J. G. Levin, *J. Virol.*, 2000, **74**, 8980–8988.
- [15] R. N. Kirchdoerfer and A. B. Ward, *Nature Commun.*, 2019, **10**, 2342.
- [16] A. J. W. te Velhuis, S. H. E. van den Worm, A. C. Sims, R. S. Baric, E. J. Snijder and M. J. van Hemert, *PLoS Pathog.*, 2010, **6**, e1001176.
- [17] G. La Penna, S. Morante, A. Perico and G. C. Rossi, *J. Chem. Phys.*, 2004, **121**, 10725–10741.
- [18] S. Furlan, G. La Penna and A. Perico, *Macromolecules*, 2008, **41**, 2938–2948.
- [19] Y.-P. Pang, *Proteins: Struct., Funct., Bioinf.*, 2001, **45**, 183–189.
- [20] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- [21] D. Wolf, P. Keblinski, S. R. Phillpot and J. Eggebrecht, *J. Chem. Phys.*, 1999, **110**, 8254–8282.
- [22] S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- [23] P. Li and K. M. Merz, *Chem. Rev.*, 2017, **117**, 1564–1686.
- [24] J. Zhang, W. Yang, J.-P. Piquemal and P. Ren, *J. Chem. Theory Comput.*, 2012, **8**, 1314–1324.
- [25] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys. Rev. B*, 1998, **58**, 7260–7268.
- [26] B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshayé, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W.-z. Yu and T. Frauenheim, *J. Chem. Phys.*, 2020, **152**, 124101–20.

- [27] N. Tochio, T. Umehara, K. Nakabayashi, , M. Yoneyama, K. Tsuda, M. Shirouzu, S. Koshiba, S. Watanabe, T. Kigawa, T. Sasazuki, S. Shirasawa and S. Yokoyama, *J. Struct. Funct. Genom.*, 2015, **16**, 55–65.
- [28] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. J. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- [29] W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33–38.
- [30] N. H. Moreira, G. Dolgonos, B. Aradi, A. L. da Rosa and T. Frauenheim, *J. Chem. Theory Comput.*, 2009, **5**, 605–614.
- [31] G. Makov and M. C. Payne, *Phys. Rev. B*, 1995, **51**, 4014.
- [32] D. F. Raffa, G. A. Rickard and A. Rauk, *J. Biol. Inorg. Chem.*, 2007, **12**, 147–164.
- [33] C. Gati, A. Stetsenko, D. J. Slotboom, S. H. Scheres and A. Guskov, *Nature communications*, 2017, **8**, 1–8.
- [34] K. Kluska, M. D. Peris-Díaz, D. Płonka, A. Moysa, M. Dadlez, A. Deniaud, W. Bal and A. Krezel, *Chem. Commun.*, 2020, **56**, 1329–1332.
- [35] N. Z. Gammoh and L. Rink, *Nutrients*, 2017, **9**, 624–637.