Interaction of Hydroxychloroquine with SARS-CoV2 Functional Proteins Using All-atoms Non-equilibrium Alchemical Simulations

Piero Procacci,*,† Marina Macchiagodena,† Marco Pagliai,† Guido Guarnieri,‡ and Francesco Iannone*,‡

†*Dipartimento di Chimica "Ugo Schiff", UniversitÃǎ degli Studi di Firenze, Via della Lastruccia 3, Sesto Fiorentino, I-50019 Italy*

‡*ENEA, Portici Research Centre, DTE-ICT-HPC P.le E. Fermi, 1, Portici (NA), I-80055 Italy*

E-mail: procacci@unifi.it; francesco.iannone@enea.it

Abstract

Using a combination of enhanced sampling molecular dynamics techniques and non-equilibrium alchemical transformations with full atomistic details, we have shown that hydroxychloroquine (HCQ) may act as a mild inhibitor of important functional proteins for SARS-CoV2 replication, with potency increasing in the series PL$^{\text{pro}}$, 3CL$^{\text{pro}}$, RdRp. By analyzing the bound state configurations, we were able to improve the potency for the 3CL$^{\text{pro}}$ target, designing a novel HCQ-inspired compound, named PMP329, with predicted nanomolar activity. If confirmed *in vitro*, our results provide a molecular rationale for the use of HCQ or of strictly related derivatives in the treatment of Covid-19

On 12 March 2020, the World Health Organization declared the Covid-19 outbreak a pandemic.[1] The viral infection, believed to have a zoonotic origin in late January from the "wet markets" of the Chinese city of Wuhan, rapidly spread throughout the five continents involving, as of today, nearly 2 million of active cases and causing the death of more than 120,000. The coronavirus SARS-CoV2, responsible for Covid-19, is in strict genomic relationship[2-4] with the SARS-CoV. The latter, also with a zoonotic origin, was the cause for the SARS outbreak in 2003 in China, effectively contained in few months by the summer of the same year. The RNA of the two closely related coronaviridae SARS-CoV and SARS-CoV2 encodes two large poli-proteins, pp1a and pp1ab, that are cleaved and transformed in mature non structural proteins (nsp) by the two virus proteases, 3CL[pro] (3C-like protease) and PL[pro] (Papain Like Protease).[5] 3CL[pro], in particular, generates from pp1a the fragment nsp12 that fold by forming, along with the nsp7 and nsp8 cofactors, the RNA dependent RNA polymerase (RdRp), i.e. the RNA-synthesizing machinery in both SARS-CoV and SARS-CoV2 positive strand RNA.

In less than two month, from early February to March 2020, the X-ray structures of the two SARS-CoV2 proteases, 3CL[pro],[6] PL[pro][7] and the cryo-EM 3.1 Å resolution structure of RdRp[8] were deposited in the Protein Data Bank. These functional SARS-CoV2 proteins (PDB codes: 6LU7, 6W9C, 6M71) exhibited a sequence percent identity with their SARS-CoV homologues (PDB codes: 1UK4, 2FE8, 6NUR) exceeding 95%, with 3CL[pro] and RdRp of SARS-CoV2 differing by only 12 (out of 306)[9] and 36 (out of 932) amino acids from their SARS-CoV counterpart. This striking similarity level leads naturally to believe that inhibitors of SARS-CoV functional proteins are very likely active for the SARS-CoV2 variants as well, and that, with appropriate funding in the mid-2000s, there would probably be a drug available for Covid-19 today.[10] Unfortunately, no antiviral compound targeting 3CL[pro], PL[pro] or RdRp for SARS was ever approved in the past decades due to a sharp decline in funding. "With no clear prospect of income from investment in drugs for future epidemics, most companies choose to focus on potential blockbusters",[10] probably relying on the wrong

assumption that chance of a repetition of a new zoonotic viral transmission was extremely unlikely.

While the research on specific anti-Covid-19 drugs is frantically ongoing[11] building upon past results on SARS,[5,12] there is a growing interest in testing *approved* drugs, with hence known limited or mild side effects, for *off-label* or compassionate use against Covid-19. Among these drugs, we mention the anti-influenza virus Favipiravir (Avigan), tested in China and Japan on patients with mild symptoms, the biologic drug Tocilizumab, administered in Italy to patient in intensive care units, the adenosine-analog Remdesivir, the antiviral drug Kaletra, a combination of anti-HIV Lopinavir and antiretroviral Ritonavir, and the anti-malaria hydroxychloroquine (HCQ), approved by the FDA for the treatment of lupus and rheumatoid arthritis as well. This latter cost-effective drug is currently undergoing clinical trials for the treatment of Covid-19 in several countries, including Italy, USA and France.

Following the SARS outbreak back in the year 2005, Vincent *et al.* reported[13] that chloroquine (CQ) had strong antiviral effects on SARS-CoV infection of primate cells. In a recent study,[14] it was found that HCQ interfered with the entry mechanism of virions in Vero E6 cell cultures, by blocking the transport of SARS-CoV2 from early endosomes to endolysosomes. In a March 29 preprint on the bioRxiv database,[15] Italian researchers found evidence that the anti-SARS-CoV2 activity of a clinically achievable HCQ concentration was maximized when administered before but also *after* the infection of Vero E6 cells in agreement with previous reports on CQ.[13] These results suggest that CQ or the less toxic HCQ may be active not only by blocking the entry of the virions through the host cell membrane, but also by way an anti-replication mechanism occurring after the release of the viral genome into the host cells.

This observations lead naturally to speculate that HCQ can be an inhibitor of any of the three cited viral functional proteins 3CL$^{\text{pro}}$, PL$^{\text{pro}}$ or RdRp, that are essential in the virus replication. In order to shed light on this matter, we have performed an accurate *in silico* "measurement" of the dissociation constant of HCQ with 3CL$^{\text{pro}}$, PL$^{\text{pro}}$ and RdRp, using
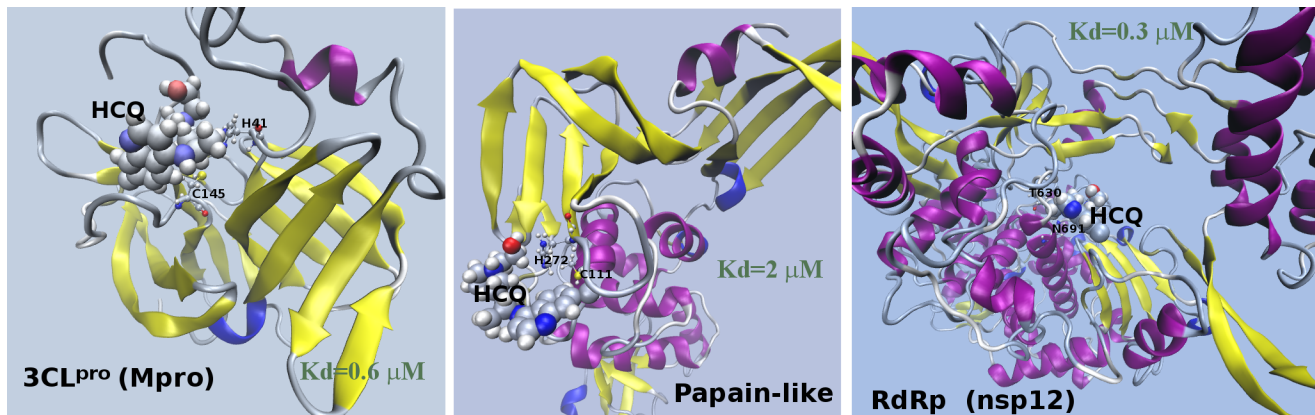
Figure 1: HCQ complexes with domain I+II of 3CL^{pro}, PL^{pro} and RdRp. HCQ is in Van der Waals representation. The active site residues are in CPK representation. Reported dissociation constants are computed from the HREM/NE determined standard dissociation free energies ΔG_0 (see text) as $K_d = e^{-\beta \Delta G}$.

a state of the art computational techniques combined with reliable modern force fields for biological systems.[16] The computational approach is based on a combination of molecular dynamics (MD) enhanced sampling techniques (Hamiltonian Replica Exchange, HREM) and non-equilibrium alchemical MD simulations,[17] encoded in an algorithm[18] amenable to massive parallelization on high performing computational facilities.[19]

We started by docking the OpenBabel[20] generated 3D structure of HCQ to the PDB structures of 3CL^{pro} (domains I+II only), PL^{pro} and RdRp (nsp12 only) using Autodock4.[21] The optimal initial docking pose was found by running 50 minimization rounds with the center of mass (COM) of the fully flexible HCQ ligand placed within a 15 Å side-length cubic box centered at the protein active sites. The latter were identified by the midpoint vectors connecting the alpha carbons of the CYS-HIS catalytic dyad in 3CL^{pro} and PL^{pro} and by the midpoint vector connecting THR680 and ASN691 in RdRp.[22] More details on the Docking approach are provided in the ESI.

The so-generated initial structures of the HCQ-3CL^{pro}, HCQ-PL^{pro} and HCQ-RdRp complexes (see Figure 1) were first equilibrated in cubic box of appropriate size, filled with TIP3P explicit water molecules, by running short simulation (100 ps) in the NPT ensemble. The resulting solvated complexes were then fed to the ORAC MD program[18] for the HREM

sampling of the bound states using a very effective torsional tempering scheme on the binding site region engaging only 8 replicas.[23] We collected, for each of the three complexes, 540 configurations sampled at regular intervals during 25 ns NPT simulation of the HREM target (unscaled) state (T=300 K and P=1 atm). From these HREM-harvested equilibrium configurations, we launched, on a single parallel job, a swarm of 540 independent alchemical non-equilibrium (NE) trajectories[24,25] where the HCQ-environment interactions were rapidly decoupled in 0.36 ns, eventually producing a ligand annihilation work distribution. During the HREM and NE simulations, HCQ was prevented to drift away from the active site using a weak harmonic restraint between the centers of mass (COM) of the ligand and the receptors.[24]

The recoupling HCQ work distribution in bulk solvent was obtained using fast growth (0.36 ps) alchemical simulations. The starting configurations in this case were generated by combining 540 HCQ solvent-decoupled conformations, sampled in a 8 ns HREM simulation of the isolated (gas-phase) molecule, with equilibrated structures of pure TIP3P water molecules in standard conditions.
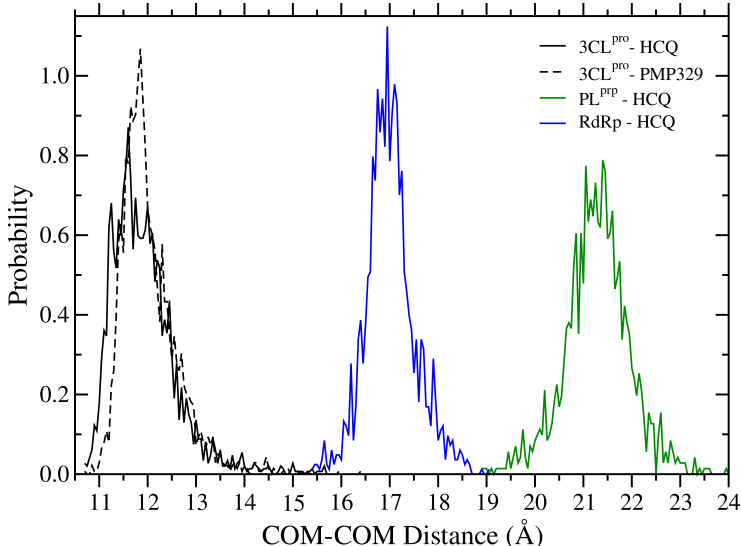

Figure 2: Probability distribution of the ligand-receptor COM-COM distance for the complex of 3CL^{pro}-HCQ, 3CL^{pro}-PMP329, PL^{pro}-HCQ, RdRp-HCQ, as obtained in 24 ns of HREM simulation.

The standard dissociation free energies, ΔG_0, were computed using the Jarzynski identity[26] evaluated on the work distribution obtained by combining the negative growth work values of HCQ in bulk with the positive decoupling work values of HCQ in the bound state, and by adding a standard state binding site volume correction.[24] The 95% confidence interval of the predicted dissociation free energies was obtained by bootstrap with resampling on the two independent sets of growth and decoupling work values. All MD calculations were performed using the program ORAC[18] on the CRESCO6 high performing computing facility located in Portici (Italy) and managed by ENEA.[19] Further details of the HREM and NE techniques are reported in the ESI.

In Figure 2 we show the probability distributions of the distance between the COM of HCQ and that of the three viral functional proteins as obtained from the HREM simulation with torsional tempering in the active site region. It can be seen that in all three cases, HCQ lingers in the active site in spite of large allowance volume (1661 $\mathring{A}^3$, see ESI) due the harmonic restraint potential. The dissociation constants $K_d = e^{-\beta \Delta G_0}$ from the HREM-NE computed standard dissociation free energy ΔG_0 are reported in the Figure 1. HCQ is predicted to have micromolar activity for all three proteins, corresponding to ΔG_0 values of 7.7 ± 0.3 kcal/mol, 8.5 ± 0.4 kcal/mol, and 9.1 ± 0.6 kcal/mol for PL$^{\text{pro}}$, 3CL$^{\text{pro}}$ and RdRP, respectively. For the HCQ-3CL$^{\text{pro}}$ complex, in particular, by analyzing the HREM sampled configuration of the bound state (see details in ESI), we noticed that in the most probable conformation HCQ adopts a compact structure characterized by the ethyl moiety insisting on the chloroquinoline planar moiety, with the hydroxy group not engaging in any stable hydrogen-bond with the surrounding protein residues. We hence decided to modify HCQ so at to further stabilize this compact structure in bulk solvent as well, in order to reduce the penalty due to the conformational entropy contribution upon binding.[27]

This was achieved by moving the OH group away to the penthyl moiety, thereby enhancing the hydrophobic character of the OH-depleted propyl moiety and its interaction with the planar group (transformed in a methyl-naphtyl moeiety) in both the bound and unbound
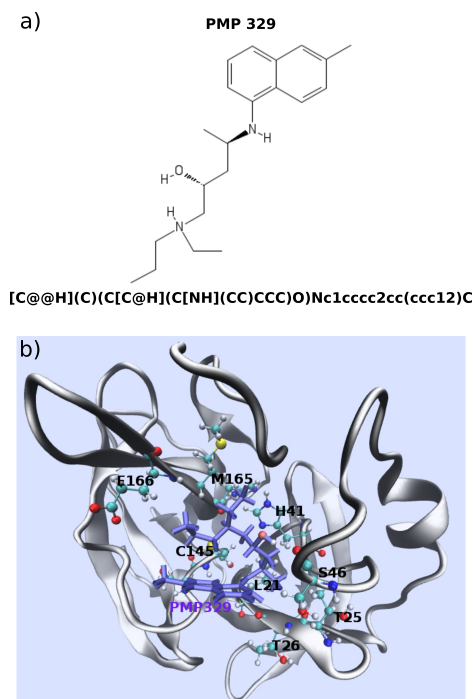
Figure 3: a): Chemical structure and SMILES code of PMP329; b) Typical PMP329 compact conformation in the binding site of 3CL$^{\text{pro}}$.

states. In the Figure 3 we show in (a) the structure and the SMILES code of the modified compound, named PMP329, and in (b) a typical compact conformation in the bound state with the hydrophobic interaction between the methyl-naphtyl moiety and the propyl group. The ligand is surrounded by mostly hydrophobic residues with a single transient H-bond between the hydroxy group of PMP329 and the carboxy group in SER46. The ΔG_0 of PMP329 for 3CL$^{\text{pro}}$, computed using the previously described HREM/NE protocol, is found to be of 9.8 ± 1.4 kcal/mol, translating in the *nanomolar activity* of $K_d \simeq 70$. PMP329 is not commercially available according to the ZINC[28] or PUBCHEM[29] databases. However its synthesis should be straightforward starting from inexpensive an readily available commercial precursors.

The results presented in this communication, obtained using advanced molecular dynamics techniques with full atomistic details, show that HCQ may act as a mild inhibitor of important functional proteins for SARS-CoV2 replication, with potency increasing in the series PL$^{\text{pro}}$, 3CL$^{\text{pro}}$, RdRp. By analyzing the HREM bound state configurations, we were

7

able to improve the potency for 3CL$^{\text{pro}}$ target, designing a novel HCQ-inspired compound, named PMP329, with predicted nanomolar activity. Our results call for a rapid *in vitro* measurements of the activity of HCQ (and possibly of PMP329) on each of the three viral protein, PL$^{\text{pro}}$, 3CL$^{\text{pro}}$ and RdRp (nsp12). If confirmed, our prediction provides a convincing molecular rationale for the use of HCQ or of strictly related derivatives in the treatment of Covid-19.

Acknowledgement

Conflicts of interest

There are no conflicts to declare.

References

(1) WHO announces COVID-19 outbreak a pandemic. http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic, 2020; Accessed 14 April 2020.

(2) 2020; Viralzone News, https://viralzone.expasy.org.

(3) 2020; The National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov.

(4) Shanker, A.; Bhanu, D.; Alluri, A. Analysis of Whole Genome Sequences and Homology Modelling of a 3C Like Peptidase and a Non-Structural Protein of the Novel Coronavirus COVID-19 Shows Protein Ligand Interaction with an Aza-Peptide and a Noncovalent Lead Inhibitor with Possible Antiviral Properties. *ChemRxiv* **2020**,

(5) Hilgenfeld, R. From SARS to MERS: Crystallographic Studies on Coronaviral Proteases Enable Antiviral Drug Design. *FEBS J.* **2014**, *281*, 4085–4096.

(6) Liu, X.; Zhang, B.; Jin, Z.; Yang, H.; Rao, Z. The Crystal Structure of 2019-nCoV Main Protease in Complex with an Inhibitor N3. RSCB PDB, 2020; pdbode: 6LU7.

(7) Osipiuk, J.; Jedrzejczak, R.; Tesar, C.; Endres, M.; Stols, L.; Babnigg, G.; Kim, Y.; Michalska, K.; Joachimiak, A. The crystal structure of papain-like protease of SARS CoV-2. RSCB PDB, 2020; pdbode: 6W9C.

(8) Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Cao, L.; Wang, T.; Wang, Q.; Lou, Z.; Rao, Z. SARS-Cov-2 RNA-dependent RNA polymerase in complex with cofactors. RSCB PDB, 2020; pdbode: 6M71.

(9) Macchiagodena, M.; Pagliai, M.; Procacci, P. Identification of potential binders of the Main Protease 3CL-pro of the COVID-19 via Structure-Based Ligand Design and Molecular Modeling. *Chem. Phys. Lett.* **2020**, Accepted.

(10) Burley, S. K. How to help the free market fight coronavirus. *Nature* **2020**, *580*, 157.

(11) Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Hilgenfeld, R.; Drag, M. Substrate specificity profiling of SARS-CoV-2 Mpro protease provides basis for anti-COVID-19 drug design. *bioRxiv* **2020**,

(12) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An Overview of Severe Acute Respiratory SyndromeâĂŞCoronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J. Med. Chem.* **2016**, *59*, 6595–6628, PMID: 26878082.

(13) Vincent, M. J.; Bergeron, E.; Benjannet, S.; Erickson, B. R.; Rollin, P. E.; Ksiazek, T. G.; Seidah, N. G.; Nichol, S. T. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virology Journal* **2005**, *2*, 69.

(14) Liu, J.; Cao, R.; Xu, M.; Wang, X.; Zhang, H.; Hu, H.; Li, Y.; Hu, Z.; Zhong, W.; Wang, M. Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. *Cell Discovery* **2020**, *6*, 16.

(15) Clementi, N.; Criscuolo, E.; Diotti, R. A.; Ferrarese, R.; Castelli, M.; Burioni, R.; Clementi, M.; Mancini, N. Combined prophylactic and therapeutic use maximizes hydroxychloroquine anti-SARS-CoV-2 effects in vitro. *bioRxiv* **2020**,

(16) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958.

(17) Procacci, P.; Guarrasi, M.; Guarnieri, G. SAMPL6 host–guest blind predictions using a non equilibrium alchemical approach. *J. Comput. Aided Mol. Des.* **2018**, *32*, 965–982.

(18) Procacci, P. Hybrid MPI/OpenMP Implementation of the ORAC Molecular Dynamics Program for Generalized Ensemble and Fast Switching Alchemical Simulations. *J. Chem. Inf. Model.* **2016**, *56*, 1117–1121.

(19) Ponti, G.; Palombi, F.; Abate, D.; Ambrosino, F.; Aprea, G.; Bastianelli, T.; Beone, F.; Bertini, R.; Bracco, G.; Caporicci, M.; Calosso, B.; Chinnici, M.; Colavincenzo, A.; Cucurullo, A.; Dangelo, P.; De Rosa, M.; De Michele, P.; Funel, A.; Furini, G.; Giammattei, D.; Giusepponi, S.; Guadagni, R.; Guarnieri, G.; Italiano, A.; Magagnino, S.; Mariano, A.; Mencuccini, G.; Mercuri, C.; Migliori, S.; Ornelli, P.; Pecoraro, S.; Perozziello, A.; Pierattini, S.; Podda, S.; Poggi, F.; Quintiliani, A.; Rocchi, A.; Scio, C.; Simoni, F.; Vita, A. *Proceeding of the International Conference on High Performance Computing & Simulation*; Institute of Electrical and Electronics Engineers (IEEE), 2014; pp 1030–1033.

(20) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3*, 33.

(21) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(22) Kirchdoerfer, R. N.; Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nature Communications* **2019**, *10*, 2342.

(23) Marsili, S.; Signorini, G. F.; Chelli, R.; Marchi, M.; Procacci, P. ORAC: A Molecular Dynamics Simulation Program to Explore Free Energy Surfaces in Biomolecular Systems at the Atomistic Level. *J. Comput. Chem.* **2010**, *31*, 1106–1116.

(24) Procacci, P. I. Dissociation free energies of drug-receptor systems via non-equilibrium alchemical simulations: a theoretical framework. *Phys. Chem. Chem. Phys.* **2016**, *18*, 14991–15004.

(25) Nerattini, F.; Chelli, R.; Procacci, P. II. Dissociation free energies in drug-receptor systems via nonequilibrium alchemical simulations: application to the FK506-related immunophilin ligands. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15005–15018.

(26) Jarzynski, C. Nonequilibrium equality for Free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.

(27) Bizzarri, M.; Marsili, S.; Procacci, P. Intraligand Hydrophobic Interactions Rationalize Drug Affinities for Peptidyl- Prolyl Cis- Trans Isomerase Protein. *J. Phys. Chem. B* **2011**, 6193–6201.

(28) Irwin, J. J.; Shoichet, B. K. ZINC-A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model* **2005**, *45*, 177–182.

(29) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound Databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213.
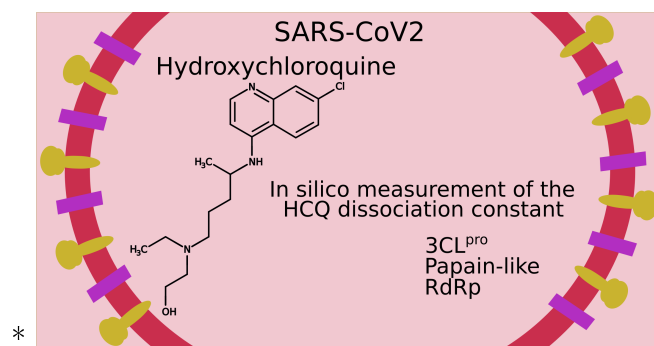
Figure 4: Graphical abstract

*