SUPPPORTING INFORMATION

Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 Mpro from QM/MM Computational Methods.

Katarzyna Świderek, Vicent Moliner

Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain.

✉ swiderek@uji.es; moliner@uji.es

COMPUTATIONAL METHODS

Computational details of the molecular models set up. The coordinates of atoms of SARS-CoV-2 Mpro were adapted from the X-Ray structure of its complex with the N3 inhibitor, as available in Protein Data Bank (PDB ID 6LU7).[1] Biological assemble (homodimer) was build using Discovery Studio Visualizer 19. Inhibitor N3 was replaced by the polypeptide Ac-Val-Lys-Leu-Gln-ACC (ACC is the fluorescent tag 7-amino-4-carbamoylmethylcoumarin) covalently bound to Cys145 through Gln, the substrate used in the kinetic studies.[2] Missing force field parameters for ACC residue, the part of substrate were generated using the Antechamber[3] program available in AmberTools package. Generated parameters are provided in Table S1. The protonation state of titratable residues was determined at pH 7 by estimating pKa shifts generated by the local environment on titratable groups using the empirical program PropKa ver. 3.0.3.[4,5] All residues were found in their standard protonation stage as shown in Table S2. After a detailed inspection of surrounding of each histidine residue, it was concluded that all should be neutral. However catalytically important His41 and His80 should be protonated in Nδ position, while remaining His64, His163, His164, His172, and His246 on Nε. Presence of S-S linkage between Cys residues was not detected. According to determine pKa values all Cys residues should be protonated.

After adding missing hydrogen atoms, the system was neutralized by addition of 7 sodium counter ions and placed in the box of 8.374 × 8.087 × 10.786 nm^3 of TIP3P[6] water molecules. Such a prepared model was further optimized, heated up to 310 Kelvins and equilibrated using Molecular Dynamic (MD) simulations under NPT conditions with AMBER force field,[7] as implemented in NAMD software.[8] Subsequently, 10 ns NVT MD simulations were performed in the peptide-protein covalently bound complex. The constant temperature during MD simulation was controlled using the Langevin thermostat.[9] Then we move backwards (though exploration QM/MM potential energy surfaces, as described below) to the reactants **E:S** non-covalent complex, and this structure was used for additional 42 ns of non-biased MD with time step of 1fs to observe the behavior of hydrogen bonds. It is important to point out that the calculations (and the system set up) were initiated from the conformation of the product of the acylation step (**E-I2** in Scheme 1 of the manuscript) in order to preliminarily avoid the debate on the protonation state of the catalytic dyad Cys145/His41 in the reactant state, **E:S**.

Robustness of MD simulations was examined by monitoring time dependence of RMSD (Fig. S1). Analysis was done using cpptraj software.[10] The structure of the MD simulation with the smallest deviation from average RMSD value was used to start the study of the molecular mechanism of the proteolysis. After optimization of the system, those residues located 20 Å beyond any atoms of the Gln5 of substrate were kept frozen in the remaining calculations. Potential energy surfaces, free energy surfaces and spline corrections have been performed using fDynamo[11] library together with implemented AMBER force field.[12] A Cut-off for nonbonding interactions was applied using a smooth switching function between 14.5 to 16 Å.

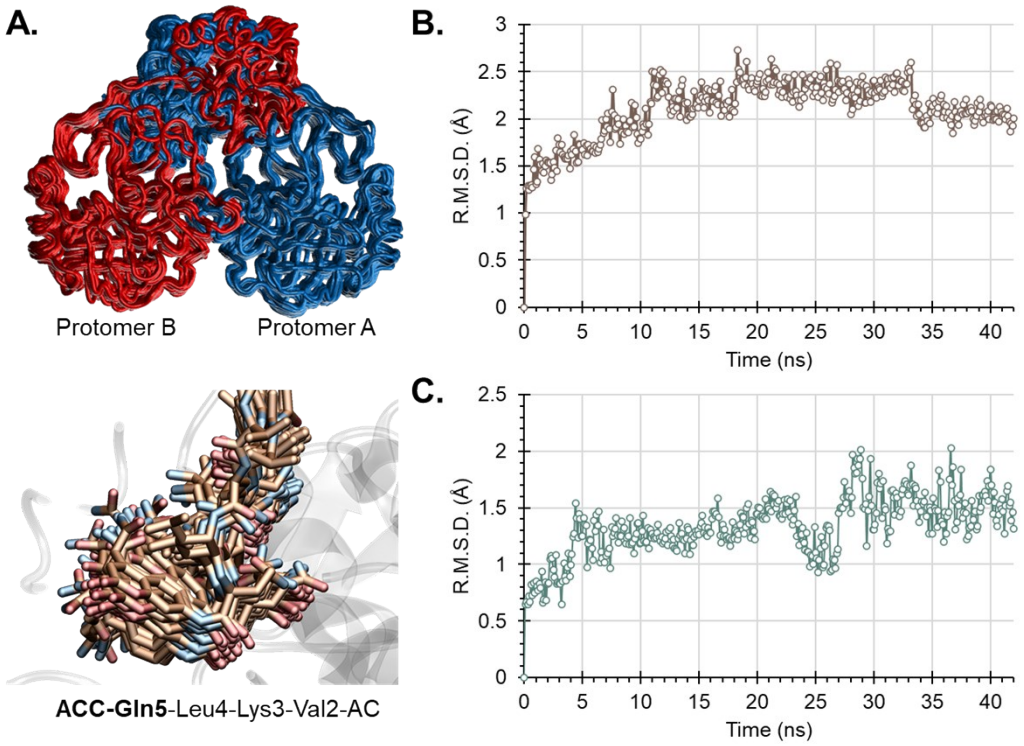

Figure S1. Analysis of classical non-accelerated 42 ns of NVT-MD simulations (T = 310 K) explored using AMBER force field implemented in NAMD software. A. Displacement of Backbone of the protein (top) along MD simulations and of the part of the substrate and residues involve in the catalysis (bottom). B. Time evolution of the Root Mean Square Deviation (RMSD) of C-Cα-N atoms of protein backbone C. Time evolution of the RMSD for all atoms of substrate.

Computational details of the QM/MM simulations. In this work, an additive hybrid QM/MM scheme was employed for the construction of the total Hamiltonian where the total energy is obtained from the sum of each contribution to the energy.

$$E_{QM/MM} = E_{QM} + E_{QM/MM}^{elect} + E_{QM/MM}^{vdW} + E_{MM} \quad \text{(S1)}$$

Here, E_{QM} describes the atoms in the QM part, $E_{QM/MM}$ defines the interaction between the QM and MM region and E_{MM} describes the rest of the MM part. As shown in Figure S2, the QM subset of atoms includes ACC and Gln5 residues of the substrate, together with Cys145 and His41 residues of the protein. Four link atoms[13] were inserted where the QM/MM boundary intersected covalent bonds in the positions indicated on Figure S2. It is important to point out that after acylation step, when ACC group is departing from the active site the definition of QM region slightly changes. Now ACC group realized after the acylation step is substituted by four water molecules, one of which was defined in QM subset. Thus, for acylation step QM part consisted of 77 atoms, while for diacylation step of 54. The Austin Model 1 (AM1)[14] semiempirical and the Minnesota Functional M06-2X[15] with the standard 6-31+G(d,p) basis set were used to treat the QM sub-set of atoms, as implemented in Gaussian09 program.[16]

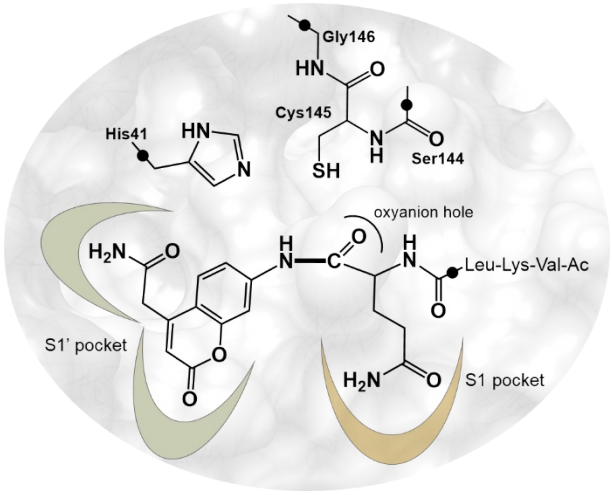

Figure S2. Schematic representation of the active site of SARS-CoV-2 M^{pro} with atoms included in the QM region in QM/MM calculations. Four link atoms are indicated as black dots.

Contribution of each residue of the protein to the interaction energy with defined part of substrate was computer using the following expression:

$$E_{QM/MM}^{Int} = \sum \left\langle \Psi \left| \frac{q_{MM}}{r_{e,MM}} \right| \Psi \right\rangle + \sum\sum \frac{Z_{QM} q_{MM}}{r_{QM,MM}} + E_{QM/MM}^{vdW} \quad \text{(S2)}$$

This interaction energy can be exactly decomposed in a sum over residues provided that the polarized wave function (Ψ) is employed to evaluate this energy contribution. The global polarization effect can be obtained from the gas phase energy difference between the polarized, Ψ, and non-polarized, Ψ_0, wave functions.

Potential Energy Surfaces. Potential energy surfaces (PESs) were obtained by grid scanning of distinguish reaction coordinates most suitable for describing each of the chemical step. A harmonic constraint was used to maintain the proper interatomic distances along the reaction coordinate, and a series of conjugate gradient optimizations and L-BFGS-B optimization algorithms were applied to obtain the final potential energy of the minimized constrained geometry. A micro-macro iteration optimization algorithm[17,18] was used to localize, optimize, and characterize the key TS structures using a Hessian matrix containing all the coordinates of the QM subsystem at M06-2X/MM level of theory with the standard 6-31+G(d,p) basis set. The gradient norm of the remaining movable atoms was maintained less than 0.01 kcal·mol^{-1}·Å^{-1}. Intrinsic reaction coordinates (IRCs) were traced down from located TSs to the valleys of the reactants, intermediates and products in mass-weighted Cartesian coordinates.

Free Energy Surfaces. FESs were obtained, in terms of Potentials of Mean Force (2D-PMF), for every step of the reaction using the Umbrella Sampling approach[19] combined with the Weighted Histogram Analysis Method (WHAM).[20] Series of MD simulations were performed adding a constraint along the selected reaction coordinates with an umbrella force constant of 2500 kJ·mol^{-1}·Å^{-2}. In every window QM/MM MD simulations were performed with a total of 5 ps of equilibration and 20 ps of production at 310 K with a time step of 1 fs. Structures obtained in previously computed PESs were used as starting points for the MD simulations in every window. The resulting Free Energy Surfaces are shown in Figures S3, S4 and S5 while key interatomic distances of every state involved in the reaction, optimized at M06-2X/MM are listed in Table S3.

Spline Corrections

The generation of PMFs requires a large number of structures that are generated from QM/MM MD calculations. Inevitably, we are restricted to the use of a semiempirical Hamiltonian, AM1 in this work. Then, in order to improve the quality of our free energy surfaces, based on the work of Truhlar and co-workers for reactions in solution[21,22,23] a spline under tension[24],[25] is used to interpolate this correction term at any value of the reaction coordinate, ξ_1 and ξ_2 in the case of two dimensional PMFs, selected to generate the free energy surfaces.[26,27] In this way we obtain a continuous function in a new energy function to obtain corrected PMFs:

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi_1,\xi_2)] \qquad \text{(S3)}$$

where S refers to the spline function, and its argument $\Delta E_{LL}^{HL}(\xi_1,\xi_2)$ is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. Herein the hybrid M06-2X functional with the 6-31+G(d,p) basis set, as suggested by Truhlar and co-workers,[28,29] has been employed as the HL method. As our simulations demonstrate, the corrections in the energy applied to any single chemical step, can be significant thus justifying our computational protocol, while the geometries generated from the sampling with LL calculations can be considered as representative of the HL geometries, thus giving credit to the selected technique. These calculations were carried out using the Gaussian09 program.[30]

It is important to point out that this method can have some limitations when the LL is not capable of describing the electronic structures of the system subject of study.[31] On the other side, even when the topology of the PESs are significantly different, spline corrections at high level have shown good results.[32] In the present study, structures optimized at the selected LL/MM combination of potentials (AM1/AMBER) were qualitatively equivalents to those later optimized at HL/MM level (M06-2X/AMBER).

Table S1. Atom types, charges and parameters obtained for ACC residue of the substrate inhibitor generated using antenchamber package included in AmberTools.

Atom name	Atom type	Charge
N	nh	-0.79555
H1	hn	0.437896
C6	ca	0.230216
C5	ca	-0.21304
C4	ca	-0.03882
C3	ca	-0.20339
C1	ca	0.21284
C2	ca	-0.22725
H4	ha	0.170609
O1	os	-0.30981
C9	c	0.760157
C8	cc	-0.24827
C7	cd	0.049076
C10	c3	-0.12423
C11	c	0.694125
N2	n	-0.63345
H8	hn	0.333319
H9	hn	0.335425
O2	o	-0.57579
H6	hc	0.09236
H7	hc	0.102892
H5	ha	0.183246
O3	o	-0.53166
H3	ha	0.149546
H2	ha	0.149546

Parameters

NONBON

nh	1.8240	0.1700
hn	0.6000	0.0157
ca	1.9080	0.0860
ha	1.4590	0.0150
os	1.6837	0.1700
c	1.9080	0.0860
cc	1.9080	0.0860
cd	1.9080	0.0860
c3	1.9080	0.1094
n	1.8240	0.1700
o	1.6612	0.2100
hc	1.4870	0.0157

BONDS:

nh-hn	401.20	1.014
nh-ca	449.00	1.364
ca-ca	478.40	1.387
ca-ha	344.30	1.087
ca-cd	411.70	1.434
ca-os	372.40	1.373
os-c	411.30	1.343
c -cc	377.40	1.462
c -o	648.00	1.214
cc-cd	504.00	1.371
cc-ha	347.20	1.085
cd-c3	337.30	1.499
c3-c	328.30	1.508
c3-hc	337.30	1.092
c -n	478.20	1.345
n -hn	410.20	1.009

ANGLES

nh-ca-ca	69.340	120.130
hn-nh-ca	49.080	116.130
ca-ca-ca	67.180	119.970
ca-ca-ha	48.460	120.010
ca-ca-cd	65.990	120.100
ca-ca-os	69.790	119.200
ca-cd-cc	68.230	113.510
ca-cd-c3	64.050	117.810
ca-os-c	63.750	120.870
os-c -cc	70.500	112.300
os-c -o	75.930	123.330
c -cc-cd	65.230	121.510
c -cc-ha	46.990	117.020
cc-c -o	68.910	125.710
cc-cd-c3	64.810	119.450
cd-cc-ha	48.350	122.890
cd-c3-c	63.810	113.170
cd-c3-hc	47.200	110.860
c3-c -n	67.860	115.150
c3-c -o	68.030	123.110
c -c3-hc	47.200	109.680
c -n -hn	49.210	118.460
n -c -o	75.830	122.030
hn-n -hn	39.730	117.850
hc-c3-hc	39.430	108.350

Parameters

DIHEDRALS:

nh-ca-ca-ca	1	3.625	180.000	2.000
nh-ca-ca-ha	1	3.625	180.000	2.000
hn-nh-ca-ca	1	1.050	180.000	2.000
ca-ca-ca-ca	1	3.625	180.000	2.000
ca-ca-ca-ha	1	3.625	180.000	2.000
ca-ca-ca-os	1	3.625	180.000	2.000
ca-ca-ca-cd	1	3.625	180.000	2.000
ca-ca-cd-cc	1	2.550	180.000	2.000
ca-ca-cd-c3	1	2.550	180.000	2.000
ca-ca-os-c	1	0.900	180.000	2.000
ca-cd-cc-c	1	4.000	180.000	2.000
ca-cd-cc-ha	1	4.000	180.000	2.000
ca-cd-c3-c	1	0.000	0.000	3.000
ca-cd-c3-hc	1	0.000	0.000	3.000
ca-os-c -cc	1	2.700	180.000	2.000
ca-os-c -o	1	2.700	180.000	2.000
ha-ca-ca-os	1	3.625	180.000	2.000
os-ca-ca-cd	1	3.625	180.000	2.000

IMPROPERS:

ca-ca-ca-nh	1.1	180.0	2.0
ca-ca-ca-ha	1.1	180.0	2.0
ca-ca-ca-cd	1.1	180.0	2.0
ca-ca-ca-os	1.1	180.0	2.0
cc-o -c -os	10.5	180.0	2.0
c -cd-cc-ha	1.1	180.0	2.0
c3-ca-cd-cc	1.1	180.0	2.0
c3-n -c -o	10.5	180.0	2.0
c -hn-n -hn	1.1	180.0	2.0

os-c -cc-cd	1	2.875	180.000	2.000
os-c -cc-ha	1	2.875	180.000	2.000
c -cc-cd-c3	1	4.000	180.000	2.000
cc-cd-c3-c	1	0.000	0.000	3.000
cc-cd-c3-hc	1	0.000	0.000	3.000
cd-ca-ca-ha	1	3.625	180.000	2.000
cd-cc-c -o	1	2.875	180.000	2.000
cd-c3-c -n	1	0.000	180.000	2.000
cd-c3-c -o	1	0.000	180.000	2.000
c3-cd-cc-ha	1	4.000	180.000	2.000
c3-c -n -hn	1	2.500	180.000	2.000
n -c -c3-hc	1	0.000	180.000	2.000
hn-n -c -o	1	2.500	180.000	-2.000
hn-n -c -o	1	2.000	0.000	1.000
o -c -c3-hc	1	0.800	0.000	-1.000
o -c -c3-hc	1	0.080	180.000	3.000
ha-cc-c -o	1	2.875	180.000	2.000
ha-ca-ca-ha	1	3.625	180.000	2.000

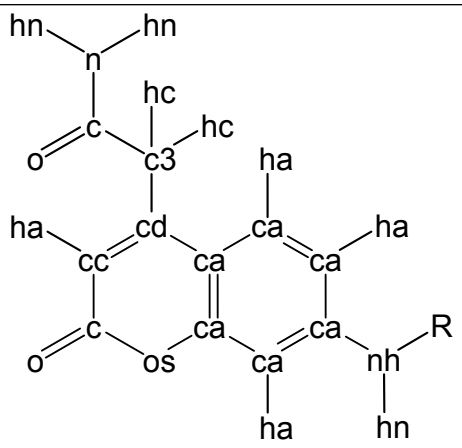

Table S2. pKa values predicted for titratable residues present in Protomer A and B of SARS-CoV-2 M^{pro}.

Residue		pKa^{calc}	pKa^{stand}	Residue		pKa^{calc}	pKa^{stand}	Residue		pKa^{calc}	pKa^{stand}
ASP	33	3.81	3.8	GLU	14	6.81	4.5	LYS	5	9.85	10.5
ASP	34	3.45	3.8	GLU	47	4.64	4.5	LYS	12	9.52	10.5
ASP	48	3.71	3.8	GLU	55	4.76	4.5	LYS	61	10.85	10.5
ASP	56	3.77	3.8	GLU	166	3.17	4.5	LYS	88	10.13	10.5
ASP	92	3.53	3.8	GLU	178	4.89	4.5	LYS	90	10.56	10.5
ASP	153	3.81	3.8	GLU	240	6.01	4.5	LYS	97	10.29	10.5
ASP	155	3.26	3.8	GLU	270	4.82	4.5	LYS	100	10.81	10.5
ASP	176	3.74	3.8	GLU	288	5.2	4.5	LYS	102	10.68	10.5
ASP	187	2.9	3.8	GLU	290	2.15	4.5	LYS	137	9.64	10.5
ASP	197	3.81	3.8	HIS	41	3.04	6.5	LYS	236	10.62	10.5
ASP	216	4.29	3.8	HIS	64	6.22	6.5	LYS	269	9.97	10.5
ASP	229	3.61	3.8	HIS	80	6.58	6.5	ARG	4	15.07	12.5
ASP	245	3.92	3.8	HIS	163	-0.11	6.5	ARG	40	14.28	12.5
ASP	248	4.15	3.8	HIS	164	4.62	6.5	ARG	60	12.62	12.5
ASP	263	3.68	3.8	HIS	172	4.04	6.5	ARG	76	12.82	12.5
ASP	289	4.08	3.8	HIS	246	5.43	6.5	ARG	105	12.98	12.5
ASP	295	4.46	3.8	TYR	37	11.92	10	ARG	131	15.58	12.5
CYS	16	11.72	9	TYR	54	15.43	10	ARG	188	12.28	12.5
CYS	22	10.3	9	TYR	101	12.69	10	ARG	217	12.23	12.5
CYS	38	12.87	9	TYR	118	13.08	10	ARG	222	12.29	12.5
CYS	44	11.8	9	TYR	126	15.03	10	ARG	279	12.02	12.5
CYS	85	11.93	9	TYR	154	10.12	10	ARG	298	10.79	12.5
CYS	117	12.98	9	TYR	161	15.62	10				
CYS	128	11.77	9	TYR	182	13.82	10				
CYS	145	16.67	9	TYR	209	12.75	10				
CYS	156	9.25	9	TYR	237	10.13	10				
CYS	160	12.98	9	TYR	239	12.75	10				
CYS	265	12.28	9								
CYS	300	11.81	9								

COMPUTATIONAL RESULTS

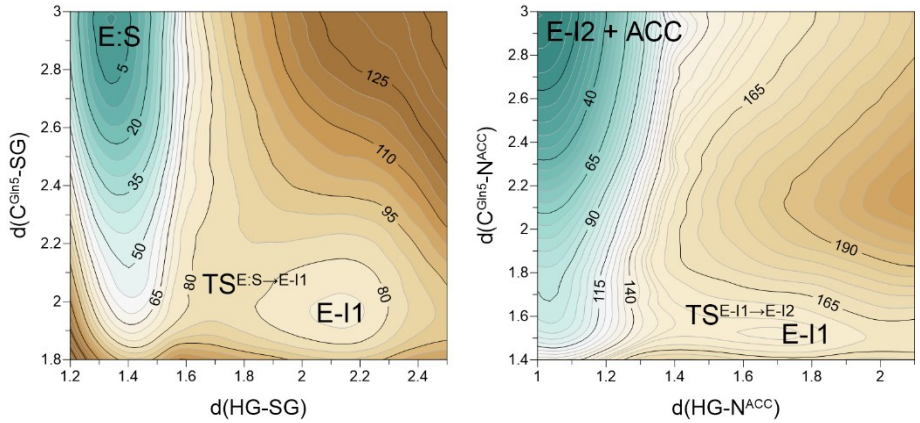

Figure S3. Free Energy Surfaces computed at M06-2X:AM1/MM level for acylation process catalyzed by SARS-CoV-2 M^{pro}. The values of energies are in kJ·mol^{-1} and the distances defined on axes are given in Å.

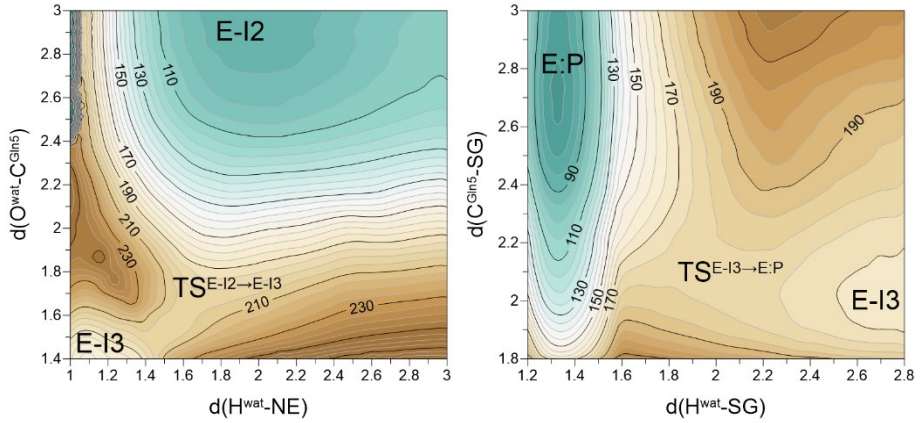

Figure S4. Free Energy Surfaces computed at M06-2X:AM1/MM level for deacylation process catalyzed by SARS-CoV-2 M^{pro}. The values of energies are in kJ·mol^{-1} and the distances defined on axes are given in Å.

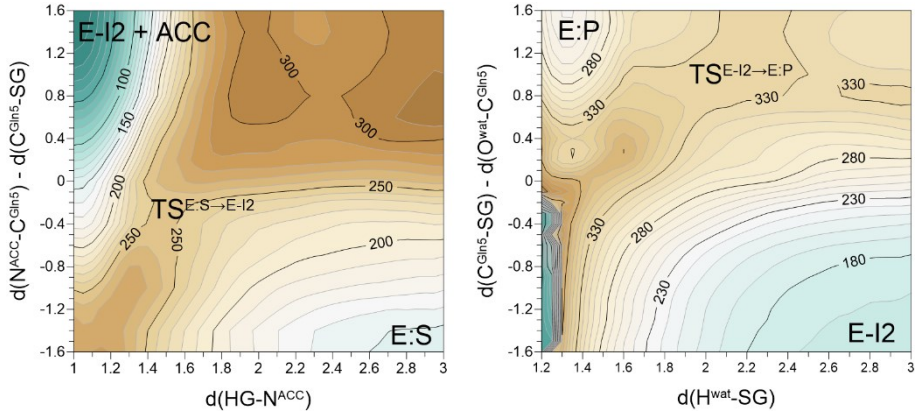

Figure S5. Free Energy Surfaces computed at M06-2X:AM1/MM level for concerted process of acylation and deacylation process catalyzed by SARS-CoV-2 M^{pro}. The values of energies are in kJ·mol^{-1} and the distances defined on axes are given in Å.

Table S3. Key distances (in Å) of structures of **E:S**, rate limiting step of the acylation reaction, $\mathbf{TS^{E:S \rightarrow E\text{-}I1}}$, intermediate **E-I2**, rate limiting step of the deacylation reaction, $\mathbf{TS^{E\text{-}I2 \rightarrow E\text{-}I3}}$, and **E:P**, optimized at M06-2X/AMBER leve310 K l of theory.

Distance	**E:S**	$\mathbf{TS^{E:S \rightarrow E\text{-}I1}}$	**E-I2**	$\mathbf{TS^{E\text{-}I2 \rightarrow E\text{-}I3}}$	**E:P**
SG^{Cys145}-HG^{Cys145}	1.36	1.89	-	-	-
NE^{His41}-HG^{Cys145}	1.91	1.13	-	-	-
SG^{Cys145}-C^{Gln5}	3.10	2.35	1.76	1.88	2.99
HG^{Cys145}-N^{ACC}	3.30	2.98	-	-	-
N^{ACC}-C^{Gln5}	1.38	1.46	-	-	-
C^{Gln5}-O^{Gln5}	1.23	1.25	1.22	1.26	1.22
SG-H^{wat}	-	-	3.07	2.95	1.35
NE^{His41}-H^{wat}	-	-	1.99	1.20	2.09
H^{wat}-O^{wat}	-	-	0.98	1.38	2.91
C^{Gln5}-O^{wat}	-	-	2.81	1.71	1.34
Oxyanion hole:					
$O^{Gln5}\cdots H^{Cys145}$	2.04	2.03	2.20	2.25	2.37
$O^{Gln5}\cdots H^{Ser144}$	2.81	2.88	3.25	3.52	3.01
$O^{Gln5}\cdots H^{Gly143}$	1.83	1.76	1.89	2.02	1.82
S1'-pocket:					
$HG^{Ser46}\cdots{=}Ox^{ACC}$	1.66	1.64	-	-	-
$HE3^{Met49}\cdots{=}Ox^{ACC}$	3.64	3.68	-	-	-
$HE1^{Met49}\cdots Oxx^{ACC}$	3.15	2.74	-	-	-
S1-pocket:					
$HE2^{His163}\cdots OE1^{Gln5}$	1.88	1.89	1.84	1.82	1.81
$HB3^{Phe140}\cdots OE1^{Gln5}$	2.84	2.87	3.07	3.12	2.98
$O^{Phe140}\cdots HE22^{Gln5}$	2.59	2.69	3.24	3.33	3.15
$O^{Phe140}\cdots HE21^{Gln5}$	4.10	3.45	4.46	4.51	4.33
$OE1^{Glu166}\cdots HE22^{Gln5}$	3.45	3.45	1.82	1.90	1.81
$OE1^{Glu166}\cdots HE21^{Gln5}$	3.65	3.80	3.28	5.41	3.23
$ND2^{Asn142}\cdots HA^{Gln5}$	2.25	2.41	3.74	3.68	3.21
$OD1^{Asn142}\cdots HE21^{Gln5}$	4.10	3.95	2.87	2.79	3.02
$O^{His164}\cdots H^{Gln5(backbone)}$	2.24	2.11	1.91	1.95	2.04

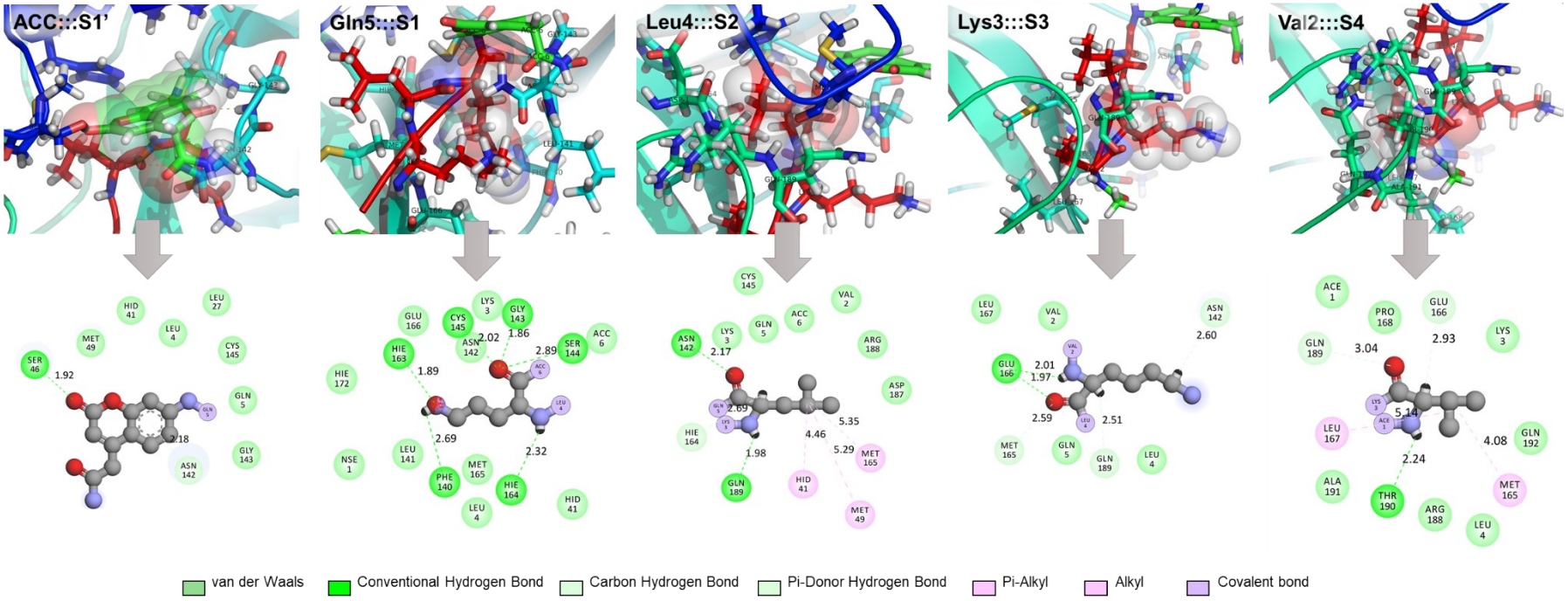

Figure S6. Structural representation of each binding pocket for the **E:S** structure optimized at M06-2X/MM level (top) and schematic representation of the specific interactions (bottom).

Table S4. Non-bonding interaction energies (E_{int}) between substrate and its binding pockets, decomposed in the contribution from electrostatic (E_{elec}) and Lennard-Jones (E_{LJ}) interactions, obtained as an average over 1000 structures from the AM1/MM MD simulation. Only those residues interacting with the substrate with E_{int} higher than 3 kcal·mol^{-1} are reported. All values in kcal·mol^{-1}.

Residue #	$\mathbf{E_{int}}$	$\mathbf{E_{elec}}$	$\mathbf{E_{LJ}}$	Residue #	$\mathbf{E_{int}}$	$\mathbf{E_{elec}}$	$\mathbf{E_{LJ}}$
ACC:::**S1'**				Leu4:::**S2**			
41	3.34 ± 0.48	4.12	-0.78	142	-5.60 ± 2.22	-4.99	-0.62
46	-3.94 ± 2.85	-2.72	-1.22	165	-6.54 ± 0.81	-4.25	-2.29
49	-4.30 ± 0.83	-2.63	-1.67	166	15.14 ± 1.69	15.55	-0.41
142	-19.09 ± 1.95	-14.72	-4.38	188	-4.68 ± 0.62	-3.13	-1.54
143	-5.79 ± 0.78	-4.80	-1.00				
144	-3.93 ± 0.73	-3.66	-0.27	Lys3:::**S3**			
145	-5.63 ± 2.11	-4.81	-0.82	165	-3.87 ± 0.65	-2.96	-0.9
164	4.57 ± 0.73	4.71	-0.14	166	-11.45 ± 1.96	-9.38	-2.1
166	7.44 ± 0.68	7.46	-0.03	189	-8.50 ± 1.17	-6.31	-2.2
187	4.70 ± 0.64	4.72	-0.02				
				Val2:::**S4**			
Gln5:::**S1**				166	8.59 ± 1.57	9.74	-1.1
27	-3.01 ± 0.26	-2.67	-0.34	167	-4.94 ± 0.86	-3.00	-1.9
41	4.49 ± 0.87	5.02	-0.52	168	-4.27 ± 0.55	-3.13	-1.1
140	-5.70 ± 0.90	-3.94	-1.77	188	-5.60 ± 0.76	-4.96	-0.6
142	-16.83 ± 1.49	-13.44	-3.39	189	-3.46 ± 1.46	-1.01	-2.5
143	-10.31 ± 0.89	-9.52	-0.79	190	6.47 ± 1.43	7.67	-1.2
144	-8.17 ± 1.78	-6.98	-1.19				
145	-5.47 ± 2.17	-5.25	-0.22				
163	-12.08 ± 1.12	-11.46	-0.62				
164	4.87 ± 0.80	6.00	-1.13				
165	-3.25 ± 1.37	-1.68	-1.57				
166	4.50 ± 1.97	6.71	-2.21				
172	-4.16 ± 0.61	-3.59	-0.57				
187	6.52 ± 0.96	6.55	-0.03				
189	3.84 ± 0.49	3.95	-0.12				

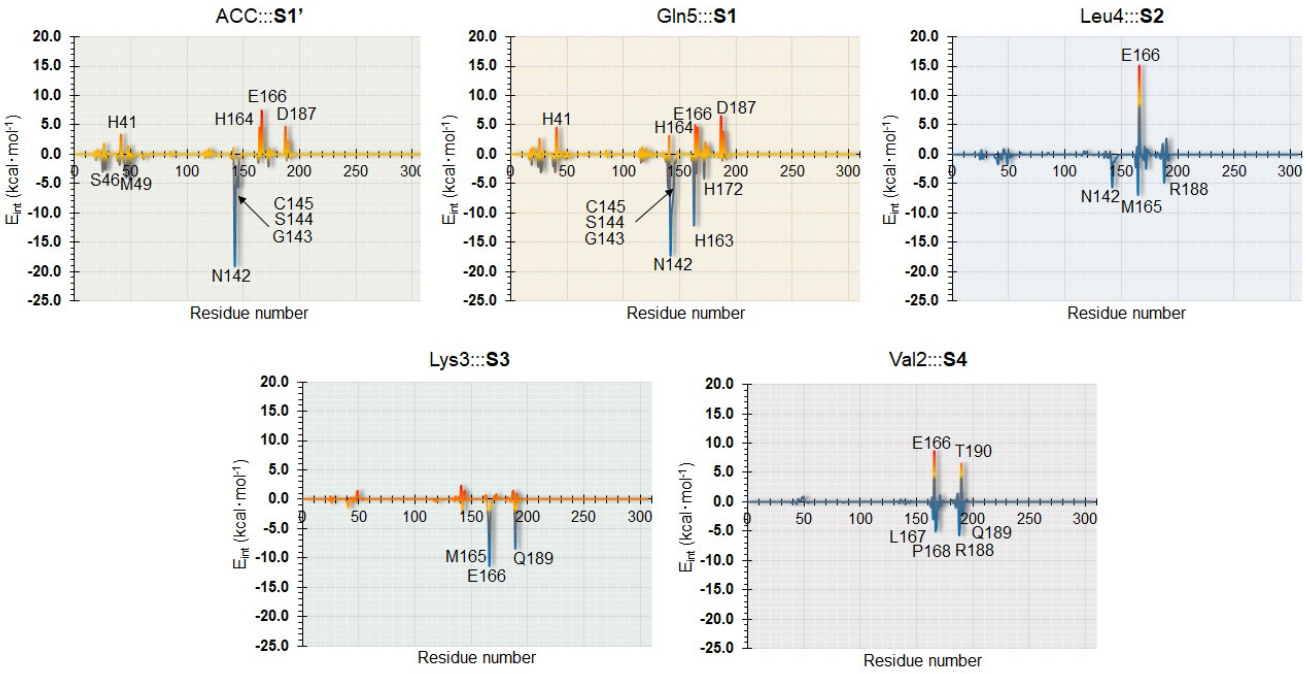

Figure S7. Non-bonding interaction energy between residues of Chain-A and each fragment of substrate, obtained as an average over 1000 structures of the AM1/MM MD simulations.

COORDINATES OF TS STRUCTURES

Table S5. Cartesian coordinates (in Å) of QM atoms for Transition State structures optimized at M06-2X/MM level for the rate limiting step of the acylation and deacylation reactions.

$TS^{E:S \rightarrow E\text{-}I1}$ (ν_i = -111.3 cm^{-1})				$TS^{E\text{-}I2 \rightarrow E\text{-}I3}$ (ν_i = -695.4 cm^{-1})			
Atoms	x	y	z	Atoms	x	y	z
C	4.038663	7.015889	25.21455	C	25.62189	7.18704	-3.24642
H	3.331241	6.453941	25.83161	H	25.9585	6.652549	-2.35282
H	5.033165	6.776299	25.60086	H	26.31448	6.891638	-4.04083
C	3.9481	6.559066	23.78876	C	24.24477	6.688861	-3.65688
N	2.825861	5.90821	23.29606	N	23.42177	5.984296	-2.79043
H	1.978411	5.695413	23.81195	H	23.59963	5.804492	-1.80886
C	3.063007	5.502348	22.0409	C	22.35365	5.514602	-3.45697
H	2.387003	4.945744	21.40984	H	21.55552	4.917151	-3.04048
N	4.281811	5.876163	21.70156	N	22.43157	5.882131	-4.72665
H	4.644013	5.571002	20.67908	H	21.60726	5.547813	-5.53707
C	4.858921	6.530869	22.76555	C	23.6026	6.609346	-4.87065
H	5.871283	6.904469	22.72479	H	23.91567	7.008036	-5.82231
C	4.635634	4.883631	14.24618	C	14.6899	5.117649	-4.51793
O	3.492918	4.928667	13.78504	O	14.26985	5.108881	-3.35792
N	4.835124	4.997355	15.57558	N	15.99875	5.361805	-4.74329
H	5.772389	4.93461	15.96156	H	16.37804	5.414149	-5.6834
C	3.798969	5.510856	16.45986	C	16.81662	5.857854	-3.65462
H	3.044071	4.743376	16.66178	H	16.81396	5.12517	-2.84231
C	4.446504	5.97909	17.77229	C	18.24672	6.092391	-4.11988
H	3.83122	6.775049	18.19671	H	18.7593	6.719555	-3.38872
H	5.425942	6.413936	17.5356	H	18.26311	6.595204	-5.09332
S	4.62024	4.679687	19.01415	S	19.15236	4.546278	-4.25003
C	3.150582	6.756443	15.84324	C	16.28503	7.194046	-3.12899
O	3.892842	7.656108	15.42515	O	15.87867	8.049782	-3.92444
N	1.82634	6.826331	15.89386	N	16.38593	7.357536	-1.81421
H	1.31276	6.043972	16.29258	H	16.69938	6.564237	-1.26065
C	1.062872	7.933632	15.28129	C	15.74783	8.494739	-1.12452
H	1.683473	8.824475	15.35657	H	15.87048	9.36594	-1.76722
H	0.178117	8.090237	15.90134	H	16.31164	8.672597	-0.20664
C	6.706643	1.954986	21.43024	C	21.71333	2.113192	-6.47977
O	7.929051	1.846056	21.56081	O	21.79872	2.250067	-7.70894
N	6.148741	2.535559	20.35125	N	20.62428	2.424839	-5.77877
H	5.13253	2.5077	20.24784	H	20.63405	2.304942	-4.76415
C	6.944356	2.808227	19.15798	C	19.42548	2.887435	-6.4414
H	7.986481	2.618416	19.45567	H	19.63447	2.838892	-7.51432
C	6.561301	1.848087	18.0263	C	18.23986	1.967716	-6.09073
H	5.504499	1.989589	17.77634	H	18.03967	2.009992	-5.01558
H	7.122048	2.125517	17.12803	H	17.34156	2.35152	-6.58892

C	6.834442	0.378575	18.37025	C	18.51535	0.520709	-6.52008
H	7.897394	0.215419	18.58695	H	18.66976	0.445099	-7.60563
C	6.39362	-0.52955	17.22947	C	17.45794	-0.47625	-6.06488
N	7.119079	-1.65448	17.01995	N	17.23091	-1.519	-6.89248
H	6.791665	-2.27297	16.28867	H	16.70354	-2.32851	-6.5506
H	7.751611	-2.05749	17.70388	H	17.72879	-1.60377	-7.76786
O	5.440089	-0.24952	16.50573	O	16.87946	-0.37315	-4.98299
H	6.280482	0.067659	19.26767	H	19.44871	0.169338	-6.05686
C	6.909562	4.279786	18.68721	C	19.17256	4.371687	-6.12648
N	7.41898	5.260393	19.64072	O	18.28878	4.969156	-6.7902
H	7.368011	6.14312	19.13324	O	20.68269	5.095624	-6.4587
C	8.774958	5.092503	20.10136	H	20.38865	5.78843	-7.06781
C	9.901314	5.302202	19.27964	H	22.45687	1.729908	-5.93182
C	11.18516	5.157512	19.79906	H	14.78744	8.328946	-0.90055
C	11.39715	4.814802	21.1481	H	14.07455	4.977584	-5.29364
C	10.25583	4.692223	21.95434	H	25.77187	8.16235	-3.08433
C	8.96577	4.795372	21.44231				
H	8.119727	4.648796	22.10275				
O	10.33684	4.486814	23.29261				
C	11.53322	4.403305	23.932				
C	12.7343	4.36493	23.10857				
C	12.69539	4.559755	21.77185				
C	13.94579	4.399962	20.94691				
H	14.82414	4.67954	21.53802				
C	14.15272	2.924671	20.58899				
N	14.39212	2.635359	19.29898				
H	14.36432	3.324577	18.55507				
H	14.66676	1.687352	19.07343				
O	14.12273	2.05524	21.45743				
H	13.95581	5.040906	20.06409				
H	13.66187	4.169027	23.63567				
O	11.49709	4.369492	25.14598				
H	12.03599	5.32914	19.14722				
H	9.772526	5.615352	18.24748				
O	7.068583	4.525456	17.47088				
H	6.092275	1.607575	22.13866				
H	0.78787	7.77105	14.33369				
H	5.43296	4.82554	13.64543				
H	3.856218	7.983317	25.39004				

Table S6. Cartesian coordinates (in Å) of QM atoms for Transition State structures of the concerted acylation and deacylation reactions optimized at M06-2X/MM level.

$TS^{E:S \rightarrow E\text{-}I2}$ (ν_i = -1503.9 cm^{-1})				$TS^{E\text{-}I2 \rightarrow E:P}$ (ν_i = -1453.9 cm^{-1})			
Atoms	x	y	z	Atoms	x	y	z
C	3.995381	7.050899	25.39507	C	25.80598	7.272366	-3.0572
H	3.154854	6.513588	25.84346	H	26.18773	6.772918	-2.16311
H	4.87253	6.796992	25.99765	H	26.46166	6.970456	-3.87888
C	4.23526	6.61029	23.98727	C	24.41075	6.811036	-3.36732
N	3.294112	5.934879	23.24173	N	23.61184	6.131248	-2.47207
H	2.368722	5.66151	23.54511	H	23.84119	5.884264	-1.51763
C	3.869554	5.592973	22.05846	C	22.46498	5.785872	-3.11958
H	3.348576	5.034873	21.29383	H	21.67419	5.225425	-2.64198
N	5.113516	6.015335	21.99026	N	22.46795	6.203987	-4.36549
H	5.898964	5.175614	19.92482	H	20.40664	5.172113	-5.57074
C	5.351005	6.648087	23.18808	C	23.67678	6.839905	-4.52853
H	6.31149	7.087575	23.42203	H	23.9716	7.278382	-5.47091
C	4.616913	4.915418	14.24536	C	14.7577	5.100952	-4.53093
O	3.479142	4.971294	13.78019	O	14.35337	5.066764	-3.36828
N	4.817173	5.010714	15.58197	N	16.06971	5.318901	-4.78328
H	5.76443	4.964686	15.93917	H	16.37531	5.426557	-5.74368
C	3.79607	5.532697	16.4848	C	16.93462	5.839136	-3.73638
H	3.033582	4.772301	16.68844	H	17.02845	5.099657	-2.93502
C	4.467234	5.979076	17.79458	C	18.31131	6.160956	-4.31363
H	3.85283	6.745014	18.2721	H	18.84262	6.833087	-3.63705
H	5.433276	6.440058	17.55458	H	18.21201	6.658509	-5.28554
S	4.66197	4.635136	19.00092	S	19.33772	4.68266	-4.45666
C	3.151969	6.788448	15.86656	C	16.36698	7.150242	-3.16696
O	3.910882	7.676483	15.45355	O	15.94987	8.007872	-3.95656
N	1.82766	6.8556	15.88024	N	16.43872	7.288558	-1.84983
H	1.30657	6.083114	16.28914	H	16.77795	6.500198	-1.30113
C	1.070964	7.960973	15.24711	C	15.80834	8.429009	-1.14786
H	1.697123	8.847934	15.31799	H	15.94553	9.307182	-1.77841
H	0.188042	8.132336	15.86625	H	16.3737	8.583468	-0.22675
C	6.705042	1.92804	21.41821	C	21.65289	2.143518	-6.52697
O	7.928196	1.855926	21.55028	O	21.77806	2.306974	-7.74641
N	6.127205	2.441367	20.30715	N	20.51214	2.40936	-5.87459
H	5.108422	2.392621	20.22666	H	20.4792	2.263892	-4.86218
C	6.905048	2.652398	19.09272	C	19.29997	2.737735	-6.59654
H	7.954481	2.477072	19.37787	H	19.52585	2.562092	-7.65658
C	6.498815	1.677835	17.98451	C	18.14355	1.840329	-6.15204
H	5.436036	1.819803	17.75809	H	17.96531	1.964861	-5.07771
H	7.040703	1.947962	17.07245	H	17.22992	2.161114	-6.66518
C	6.774577	0.210058	18.32914	C	18.46292	0.380999	-6.49404
H	7.836319	0.051458	18.55435	H	18.65865	0.279851	-7.57106

C	6.354897	-0.67229	17.15886
N	7.082955	-1.78705	16.9246
H	6.779373	-2.36707	16.1516
H	7.706312	-2.21501	17.60405
O	5.416486	-0.3643	16.42568
H	6.210658	-0.10954	19.21743
C	6.863598	4.094523	18.55682
N	7.219383	5.166752	19.56866
H	7.207312	6.021934	19.00539
C	8.556856	5.08041	20.14674
C	9.669498	5.283419	19.31107
C	10.95346	5.170239	19.82705
C	11.1692	4.867467	21.18346
C	10.03476	4.761114	22.00621
C	8.730877	4.838028	21.49977
H	7.881354	4.717762	22.16409
O	10.14683	4.589761	23.34814
C	11.36569	4.527422	23.96617
C	12.54901	4.47238	23.12072
C	12.4846	4.646899	21.78539
C	13.7314	4.517721	20.9503
H	14.59054	4.901877	21.5126
C	14.05278	3.044183	20.67657
N	14.42332	2.724597	19.42625
H	14.41478	3.382569	18.65516
H	14.7896	1.794539	19.26657
O	13.99963	2.213058	21.57988
H	13.69265	5.101756	20.03017
H	13.48745	4.29701	23.63636
O	11.37518	4.532437	25.17833
H	11.79899	5.340288	19.16916
H	9.531526	5.5592	18.26992
O	7.131735	4.346159	17.37846
H	6.095214	1.591237	22.13562
H	0.795223	7.791622	14.30092
H	5.415477	4.852096	13.64679
H	3.815456	8.028248	25.5065
C	17.39036	-0.59918	-6.05125
N	17.16603	-1.63946	-6.87848
H	16.63413	-2.44761	-6.53713
H	17.69389	-1.74499	-7.73347
O	16.78949	-0.47395	-4.98389
H	19.3824	0.06662	-5.97968
C	18.99031	4.23112	-6.50523
O	18.00365	4.757963	-7.02145
O	20.32361	4.976686	-6.81818
H	20.05404	5.791046	-7.27701
H	22.38715	1.773365	-5.95789
H	14.84456	8.281252	-0.92585
H	14.13481	5.006112	-5.30746
H	25.91966	8.257568	-2.92892

REFERENCES

(1) Z. Jin et al. Nature 2020 https://doi.org/10.1038/s41586-020-2223-y.
(2) W. Rut, K. Groborz, L. Zhang, X. Sun, M. Zmudzinski, R. Hilgenfeld, M. Drag bioRxiv, 2020, preprint, DOI: https://doi.org/10.1101/2020.03.07.981928.
(3) J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* 2020, **25**, 247260.
(4) M. H. M.Olsson, C. R. SØndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput*. 2011, **7**, 525–537.
(5) C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *J Chem Theory Comput*. 2011, **7**, 2284–2295.
(6) W. L.Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* 1983, ***79***, 926–935.
(7) W. E. I. Zhang, *J. Comput. Chem.* 2003, **24**, 1999–2012.
(8) J. C. Phillips, *J. Comput. Chem.* 2005, **26**, 1781–1802.
(9) G. S. Grest, K. Kremer, *Phys. Rev. A*, 1986, **33**, 3628-3631.
(10) D. R. Roe, T. E. Cheatham, III *J. Chem. Theory Comput*. 2013, **7**, 3084-3095.
(11) M. J.Field, M. Albe, C. Bret, M. F. Proust-De Martin, A. Thomas, *J. Comput. Chem.* 2000, **21**, 1088–1100.
(12) A. Krzeminska, P. Paneth, V. Moliner, K. Świderek, *J. Phys. Chem. B*. 2015, **119**, 917-927.
(13) M. J. Field, P. A. Bash, M. Karplus, M. *J. Comput. Chem*. 1990, **11**, 700−733.
(14) M.J.S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* 1985, **107**, 3902–3909.
(15) Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc*. 2008, **120**, 215–241.
(16) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.Gaussian 09, Gaussian, Inc.: Wallingford, CT, 2009.
(17) A. Turner, V. Moliner, I. H. Williams, *Phys. Chem. Chem. Phys*. 1999, **1**, 1323−1331.
(18) S. Martí, V. Moliner, I. Tuñón, I. *J. Chem. Theory Comput*. 2005, **1**, 1008−1016.
(19) G. M. Torrie, J. P. Valleau, *J. Comput. Phys*. 1977, **23**, 187−199.
(20) S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *J. Comput. Chem*. 1992, **13**, 1011−1021.
(21) K. A. Nguyen, I. Rossi, D. G. Truhlar *J. Chem. Phys.* 1995, **103**, 5522-5530.
(22) J. C. Corchado, E. L. Coitiño, Y. Chuang, P. L. Fast, D. G. Truhlar *J. Phys. Chem. A* 1998, **102**, 2424-2438.
(23) Y. Y. Chuang, J. C. Corchado, D. G. Truhlar *J. Phys. Chem. A* 1999, **103**, 1140−1149.
(24) R. J. Renka, *SIAM J. Stat. Comput.* 1987, **8**, 393-415.
(25) R. J. Renka, *ACM Trans. Math. Software* 1993, **19**, 81-94.
(26) J. J. Ruiz-Pernia, E. Silla, I. Tuñón, S. Martí, V. Moliner *J. Phys. Chem. B* 2004, **108**, 8427−8433.
(27) J. J. Ruiz-Pernia, E. Silla, I. Tuñón, S. Martí, *J. Phys. Chem. B*. 2006, **110**, 17663-17670.
(28) Y. Zhao, D. G. Truhlar *Theor. Chem. Acc.* 2008, **120**, 215−241.
(29) B. J. Lynch, Y. Zhao, D. G. Truhlar *J. Phys. Chem. A* 2003, **107**, 1384−1388.
(30)Gaussian 09, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K.

Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
(31) S. Martí, V. Moliner, K. Świderek *Theor. Chem. Acc.* 2019*,* **138**:120.
(32) K. Świderek, I. Tuñón, S. Martí, V. Moliner *ACS Catal.*,2015, **5**, 1172-1185.