Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Insights into the reaction mechanism of 3-*O*-Sulfotransferase through QM/MM calculations

*Rui P. Sousa, Pedro A. Fernandes, Maria J. Ramos and Natércia F. Brás**

REQUIMTE, UCIBIO/Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal.

* corresponding author, NFB (nbras@fc.up.pt)

SUPPORTING INFORMATION

Table S1 - List of the atoms included in the high-layer of the QM/MM calculations.

Atom Number	Atom Type	Residue	Residue Number	X	Y	Z
1	CE	LYS	162	36.328	32.266	43.287
2	HE2	LYS	162	36.774	32.806	44.123
3	HE3	LYS	162	35.284	32.03	43.495
4	NZ	LYS	162	36.995	30.979	43.147
5	HZ1	LYS	162	37.992	31.03	43.301
6	HZ2	LYS	162	36.747	30.304	43.857
7	HZ3	LYS	162	36.805	30.441	42.313
8	CB	GLU	184	39.457	25.649	32.708
9	HB2	GLU	184	40.498	25.946	32.841
10	HB3	GLU	184	39.022	26.352	31.998
11	CG	GLU	184	38.806	25.992	34.008
12	HG2	GLU	184	37.748	25.729	34.01
13	HG3	GLU	184	39.2	25.291	34.744
14	CD	GLU	184	38.954	27.449	34.581
15	OE1	GLU	184	38.019	27.829	35.383
16	OE2	GLU	184	39.916	28.182	34.258

17	CB	HIP	186	41.196	28.022	28.231
18	HB2	HIP	186	41.978	27.265	28.293
19	HB3	HIP	186	41.573	28.819	27.59
20	CG	HIP	186	40.887	28.615	29.603
21	ND1	HIP	186	40.244	29.781	29.739
22	HD1	HIP	186	39.985	30.444	29.023
23	CE1	HIP	186	39.963	29.912	31.059
24	HE1	HIP	186	39.358	30.68	31.519
25	NE2	HIP	186	40.398	28.8	31.703
26	HE2	HIP	186	40.43	28.736	32.711
27	CD2	HIP	186	41.024	27.972	30.78
28	HD2	HIP	186	41.586	27.069	30.967
29	CA	ASP	189	38.236	32.129	24.378
30	HA	ASP	189	37.822	32.969	23.82
31	CB	ASP	189	37.758	32.172	25.851
32	HB2	ASP	189	38.074	33.158	26.193
33	HB3	ASP	189	36.675	32.102	25.944
34	CG	ASP	189	38.51	31.175	26.689

35	OD1	ASP	189	38.711	30.043	26.199
36	OD2	ASP	189	38.803	31.468	27.892
37	CE	LYS	215	35.331	26.063	36.601
38	HE2	LYS	215	34.629	25.765	37.38
39	HE3	LYS	215	36.09	25.284	36.673
40	NZ	LYS	215	35.875	27.435	36.915
41	HZ1	LYS	215	36.421	27.375	37.762
42	HZ2	LYS	215	36.494	27.715	36.168
43	HZ3	LYS	215	35.135	28.122	36.875
44	CE	LYS	368	37.328	25.003	43.194
45	HE2	LYS	368	36.555	24.364	42.767
46	HE3	LYS	368	37.279	25.908	42.589
47	NZ	LYS	368	38.592	24.397	42.872
48	HZ1	LYS	368	38.702	24.061	41.926
49	HZ2	LYS	368	39.244	25.14	43.08
50	HZ3	LYS	368	38.726	23.631	43.516
51	C10	SUG		40.9	29.765	37.213
52	H7	SUG		40.832	28.901	36.547

53	C9	SUG	39.722	30.695	36.919
54	O9	SUG	38.486	29.942	37.047
55	H6	SUG	38.539	29.325	36.295
56	H16	SUG	39.748	31.518	37.638
57	C8	SUG	39.855	31.417	35.605
58	N1	SUG	38.829	32.527	35.493
59	S2	SUG	38.819	33.363	34.143
60	O13	SUG	39.188	32.405	33.095
61	O14	SUG	39.703	34.378	34.469
62	O15	SUG	37.428	33.645	34.164
63	H11	SUG	37.897	32.119	35.482
64	H10	SUG	39.579	30.695	34.832
65	C7	SUG	41.267	31.988	35.32
66	H9	SUG	41.366	32.311	34.273
67	O19	SUG	41.653	32.996	36.207
68	H14	SUG	41.056	33.683	35.86
69	O11	SUG	42.223	30.898	35.529
70	C11	SUG	42.288	30.406	36.897

71	H8	SUG	42.519	31.18	37.634
72	O11	PAP	35.483	30.132	39.564
73	S1	PAP	35.854	28.989	40.303
74	O12	PAP	35.567	27.745	39.637
75	O13	PAP	37.121	29.062	41.003
76	O6	PAP	34.871	29.071	41.477
77	P2	PAP	33.415	28.553	41.218
78	O4	PAP	32.714	29.117	40.085
79	O5	PAP	32.948	27.161	41.531
80	O7	PAP	32.73	29.42	42.389
81	C1	PAP	33.16	29.126	43.774
82	H9	PAP	33.259	28.045	43.901
83	H10	PAP	34.119	29.626	43.932

Table S2 - Activation and reaction energies obtained for the catalytic mechanism of the 3-OST enzyme (conformation extracted after 3 ns) with the QM region treated with different density functionals (with dispersion corrections where applicable), using the 6-311+G(2d,2p) basis set and the electrostatic embedding scheme.

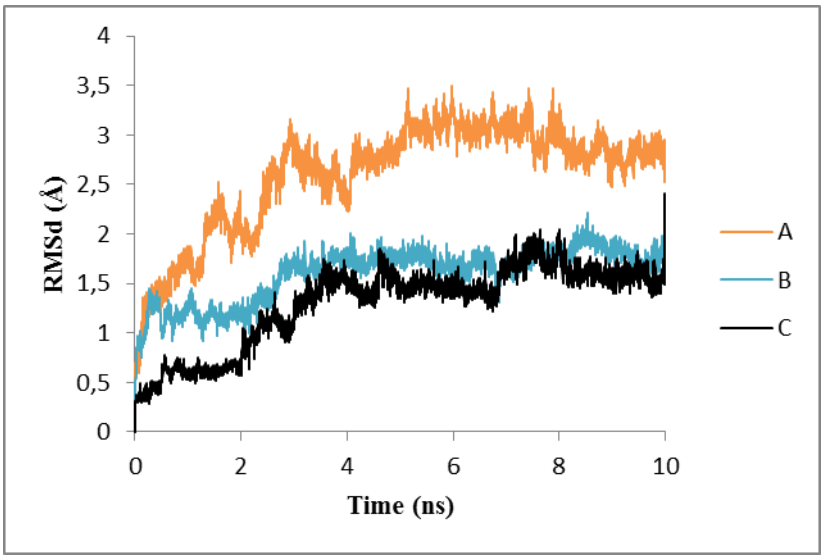
Functional	Functional Type	$\Delta G^{\ddagger}$	$\Delta G^{\ddagger}$ (D3)	ΔG_R	ΔG_R (D3)
B3LYP	h-GGA	19.0	16.9 (-2.1)	-36.2	-38.0 (-1.6)
BMK	hm-GGA	24.3	22.4 (-1.8)	-34.0	-35.6 (-1.6)
B1B95	hm-GGA	24.3	22.7 (-1.6)	-31.2	-33.0 (-1.8)
M06	hm-GGA	22.6	22.2 (-0.4)	-35.8	-36.1 (-0.2)
M062X	hm-GGA	21.6	20.9 (-0.7)	-36.8	-36.7 (-0.1)
CAM-B3LYP	h-GGA	20.8	19.5 (-1.3)	-36.2	-37.2 (-1.1)
B98	h-GGA	20.1	-	-35.5	-
wB97XD	h-GGA	20.9	-	-35.5	-
PBE1PBE	h-GGA	20.6	19.3 (-1.4)	-36.1	-37.1 (-1.0)
mPW1PBE	hm-GGA	20.7	-	-36.1	-

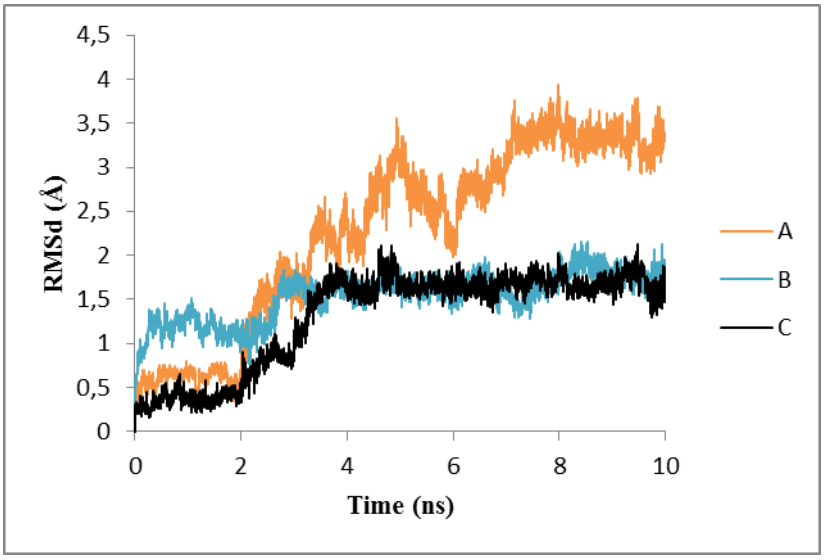
Table S3 - Activation and reaction energies obtained for the catalytic mechanism of the 3-OST enzyme (conformation extracted after 6 ns) with the QM region treated with different density functionals (with dispersion corrections where applicable), using the 6-311+G(2d,2p) basis set and the electrostatic embedding scheme.

Functional	Functional Type	$\Delta G^{\ddagger}$	$\Delta G^{\ddagger}$ (D3)	ΔG_R	ΔG_R (D3)
B3LYP	h-GGA	13.6	8.9 (-4.4)	-32.8	-36.0 (-2.8)
BMK	hm-GGA	14.9	10.2 (-4.7)	-35.8	-39.0 (-3.1)
M06	hm-GGA	10.9	10.5 (-0.4)	-37.8	-38.4 (-0.6)
M062X	hm-GGA	10.6	10.2 (-0.4)	-38.0	-38.2 (-0.2)
CAM-B3LYP	h-GGA	14.8	11.9 (-2.9)	-34.4	-36.6 (-2.2)
B98	h-GGA	13.6	-	-33.2	-
wB97XD	h-GGA	13.2	-	-35.5	-
PBE1PBE	h-GGA	14.4	11.8 (-2.6)	-33.0	-35.0 (-2.0)
mPW1PBE	hm-GGA	15.2	-	-32.7	-

Table S4 - Representation of the breakdown of the energy values, in kcal/mol, obtained for the different stationary points from the QM/MM calculations performed with the structures extracted from the MD simulations, at 3, 6 and 8 ns.

		R	**TS**	**P**
3 ns	ONIOM Total	0.00	22.16	-33.13
	High Model	0.00	40.34	28.71
	Low Real	0.00	-18.18	-61.84
6 ns	ONIOM Total	0.00	10.90	-33.76
	High Model	0.00	36.36	-5.41
	Low Real	0.00	-24.83	-27.72
8 ns	ONIOM Total	0.00	14.26	-17.25
	High Model	0.00	16.49	-1.49
	Low Real	0.00	-2.24	-15.75


RMSd (Å)
Time (ns)
A
B
C


RMSd (Å)
Time (ns)
A
B
C

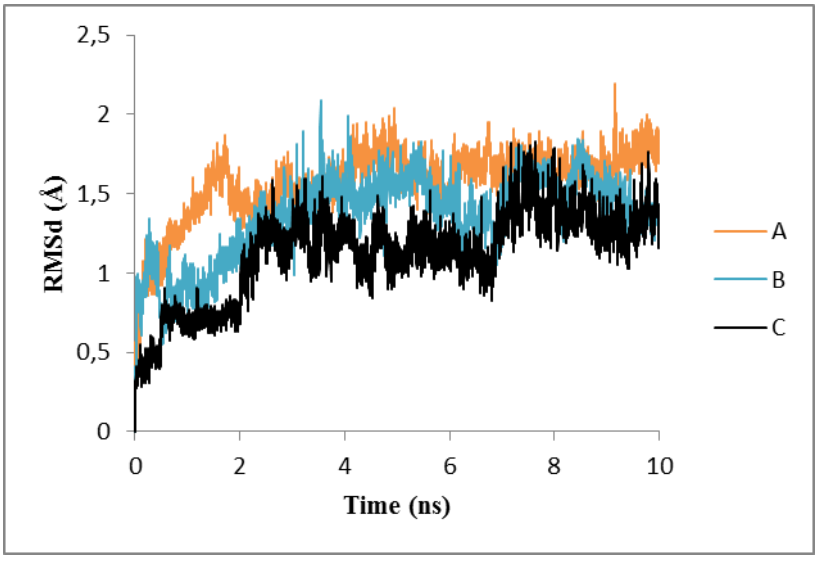

RMSd (Å)
Time (ns)
A
B
C

Figure S1 - RMSd values for the active site residues and substrate molecules (top), the substrate molecules (middle) and active site residues (bottom), throughout the MD simulations, for the three models studied.

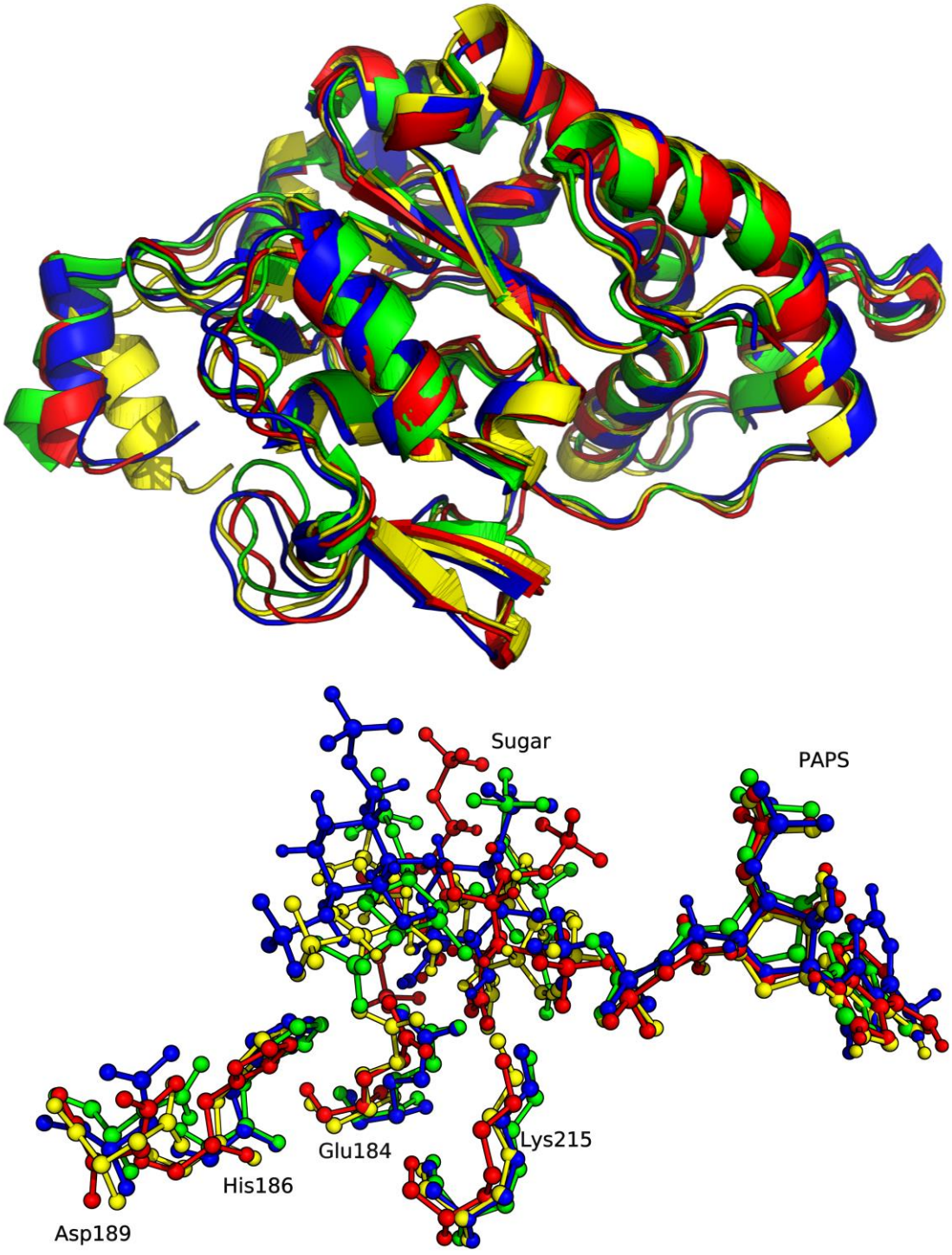
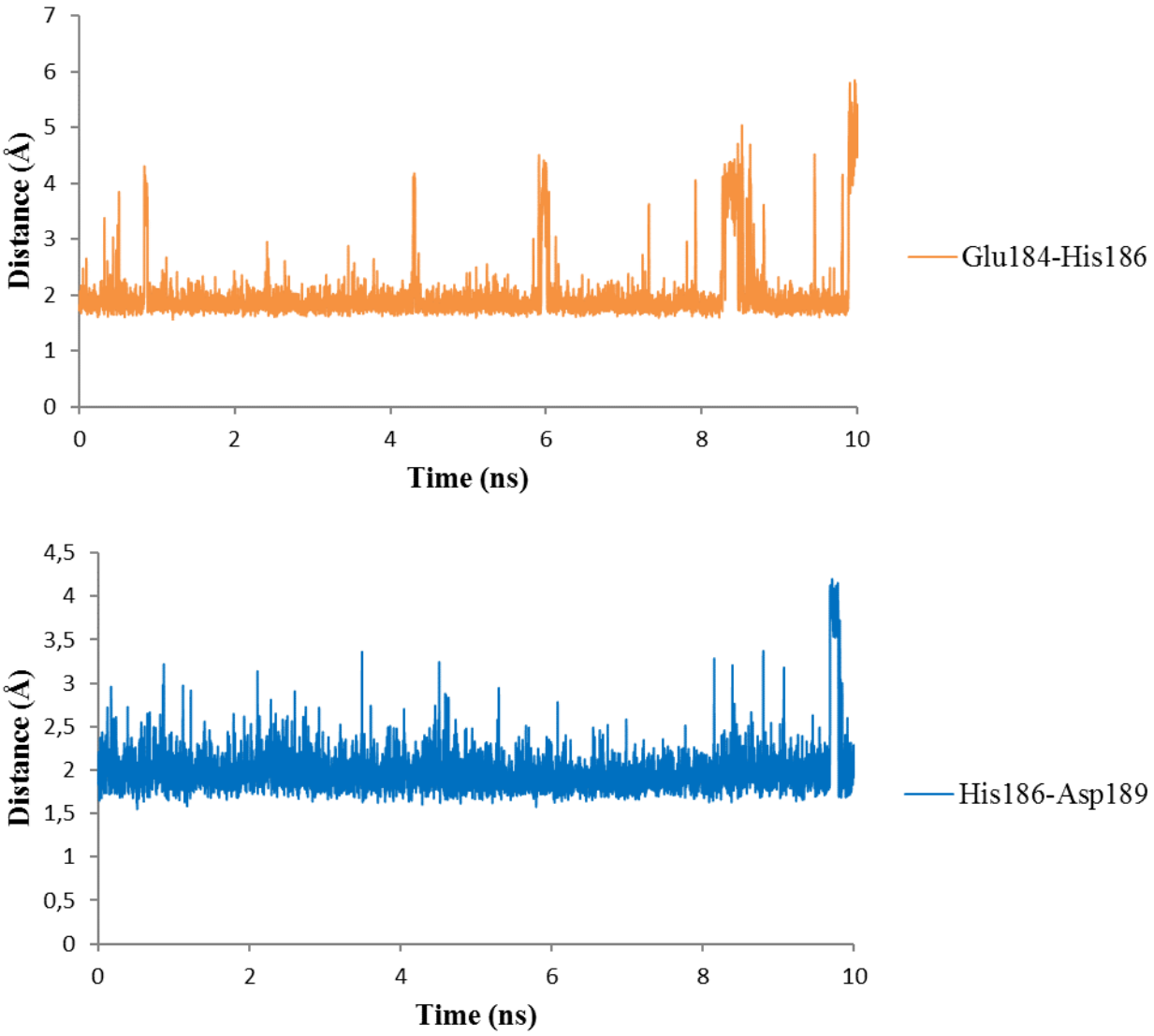


Figure S2 - Representation of the superimposition of the resulting structures extracted from the end of the MD simulations (after 10 ns) (red - model A; blue – model B; yellow - model C) and the crystallographic structure (green), with a cartoon representation for the whole protein (top) and a CPK representation of the key active site residues and substrate molecules (bottom).



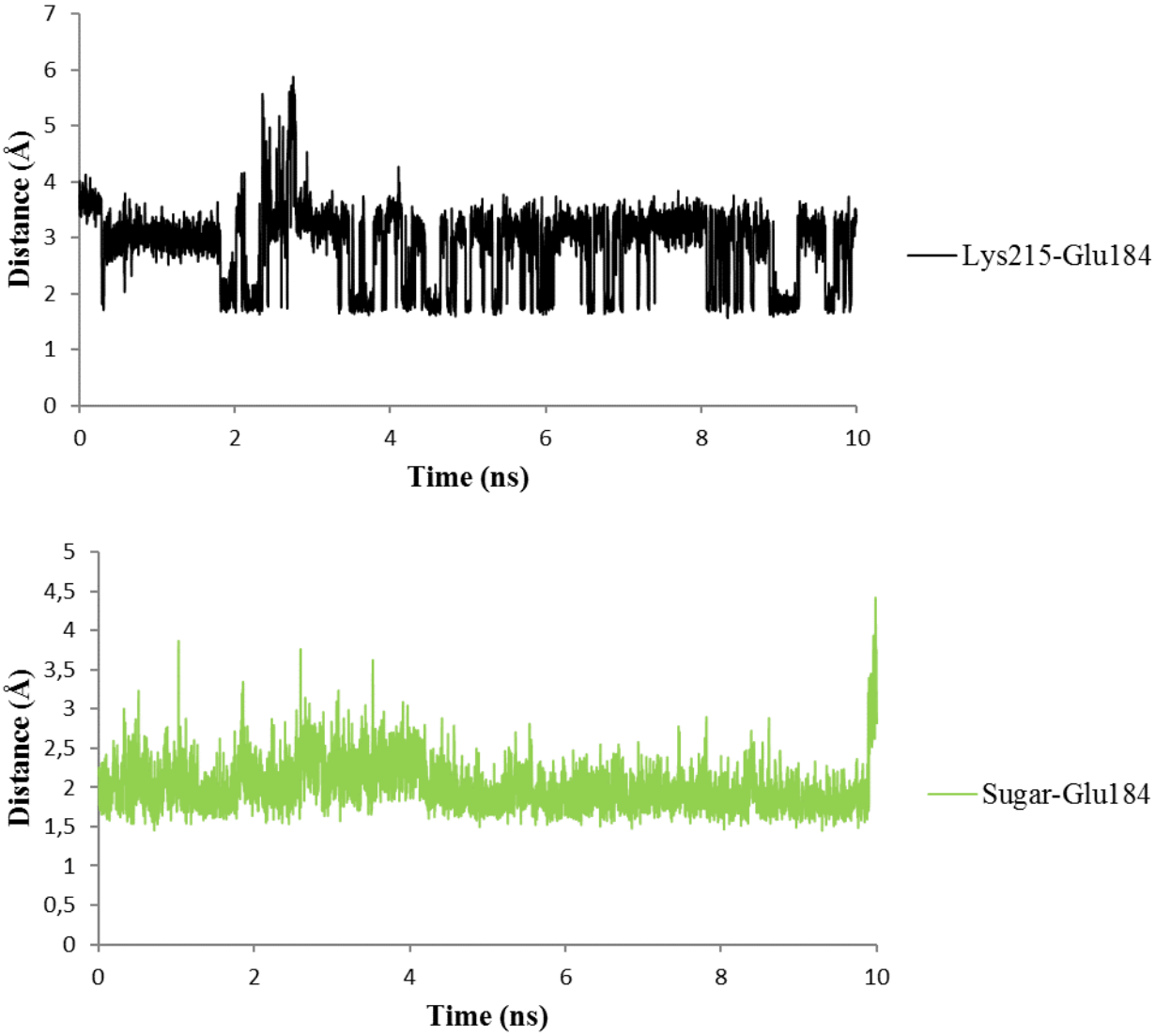


Figure S3 – Representation of the key distances studied throughout the MD simulation of model C.

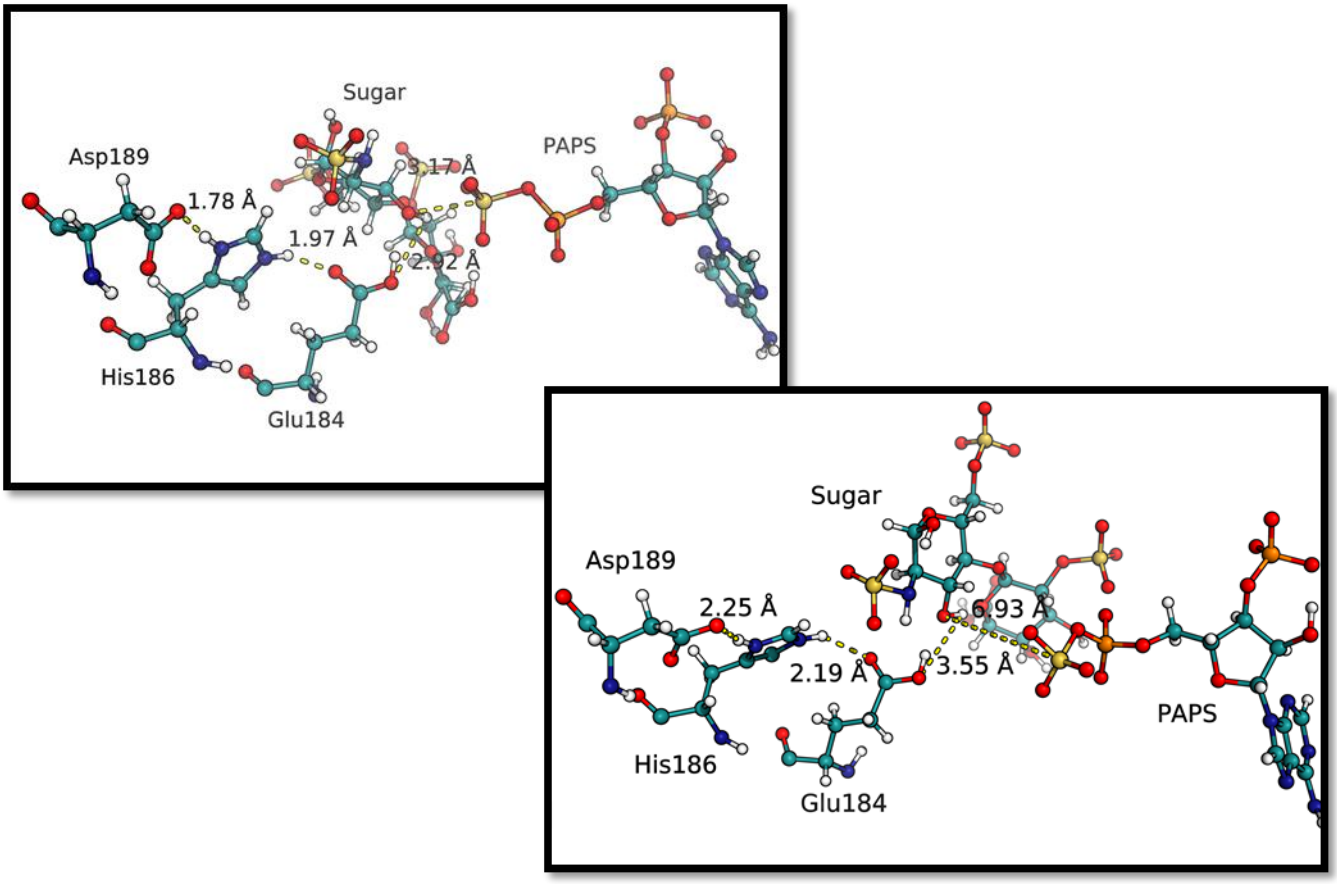


Figure S4 - Representation of the structures at the beginning (left) and end (right) of the MD simulation with doubly protonated His186, negatively charged Asp189, neutral Glu184 and Lys215.

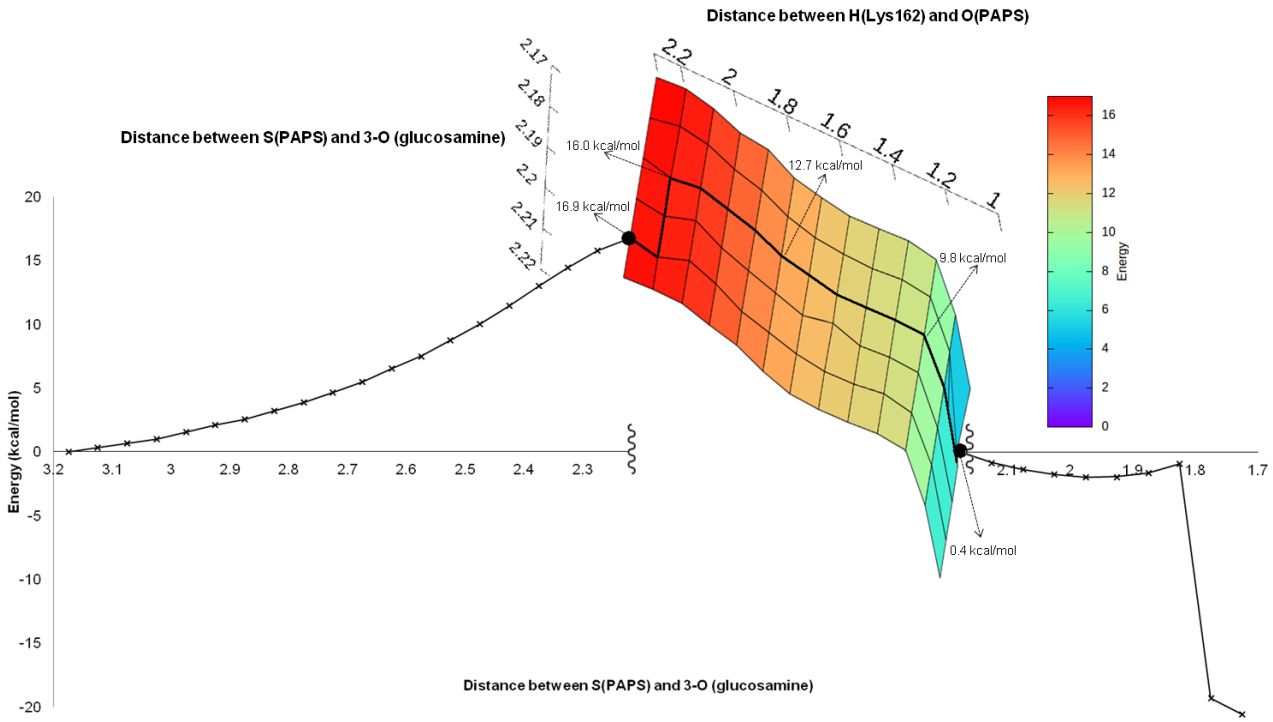


Figure S5 – Potential Energy Surface of the QM/MM calculations, along the chosen reaction coordinate. Near the TS structure, a bidimensional PES scan was performed, with the shortening of the distance between the proton from Lys162 and the oxygen atom from the phosphate of the PAPS molecule (represented horizontally in the 3D plot) and the distance between the sulfur atom of the PAPS molecule and the 3-O position of the glucosamine (represented vertically in the 3D plot) as reaction coordinates. The colors in the 3D plot represent the energy in kcal/mol. All distances are in Angstrom.

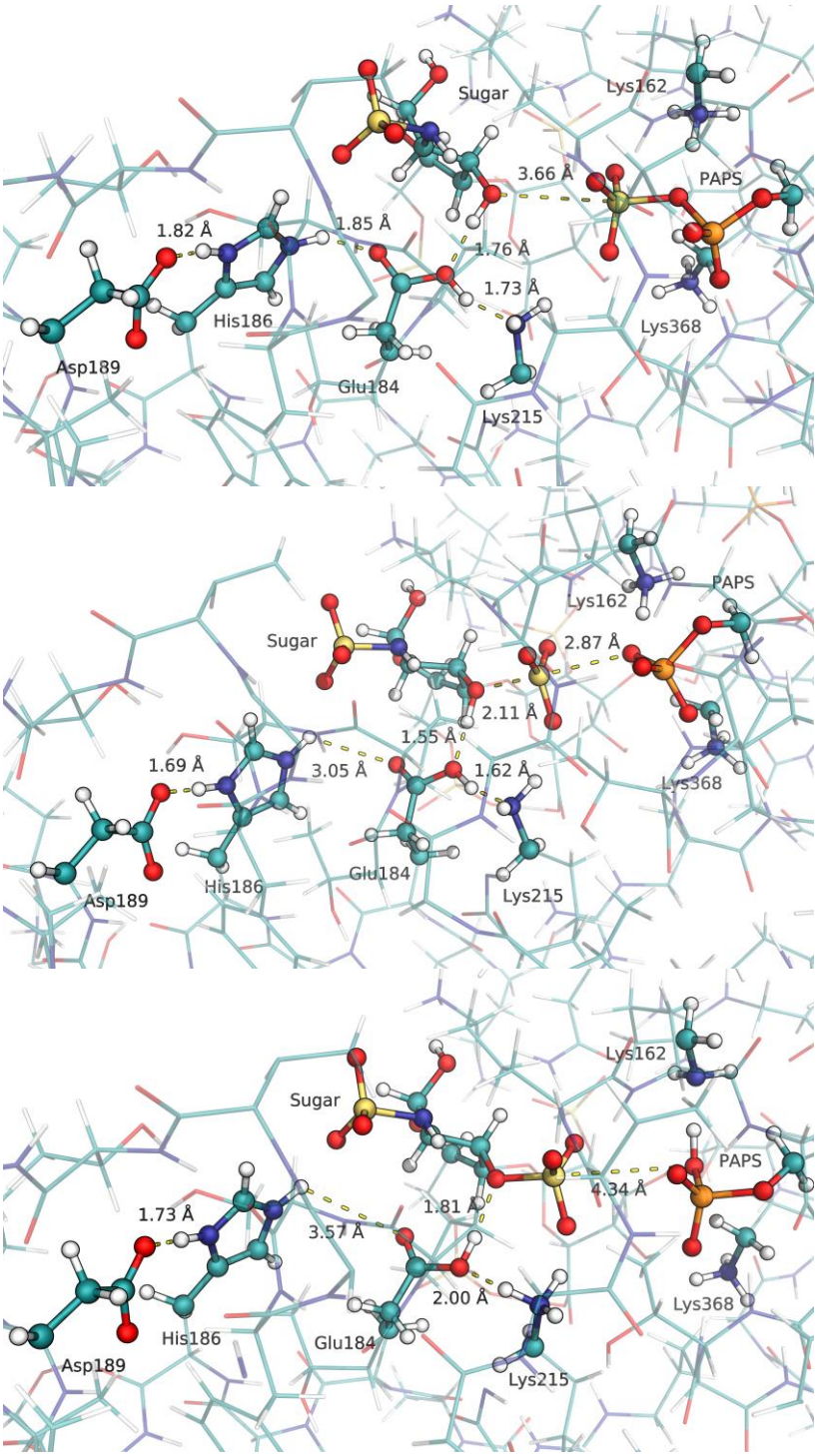


Figure S6 – Representation of the reactants (top), TS (middle) and products (bottom) structures obtained from the QM/MM calculations performed on the structure extracted at 3 ns from the MD simulations.

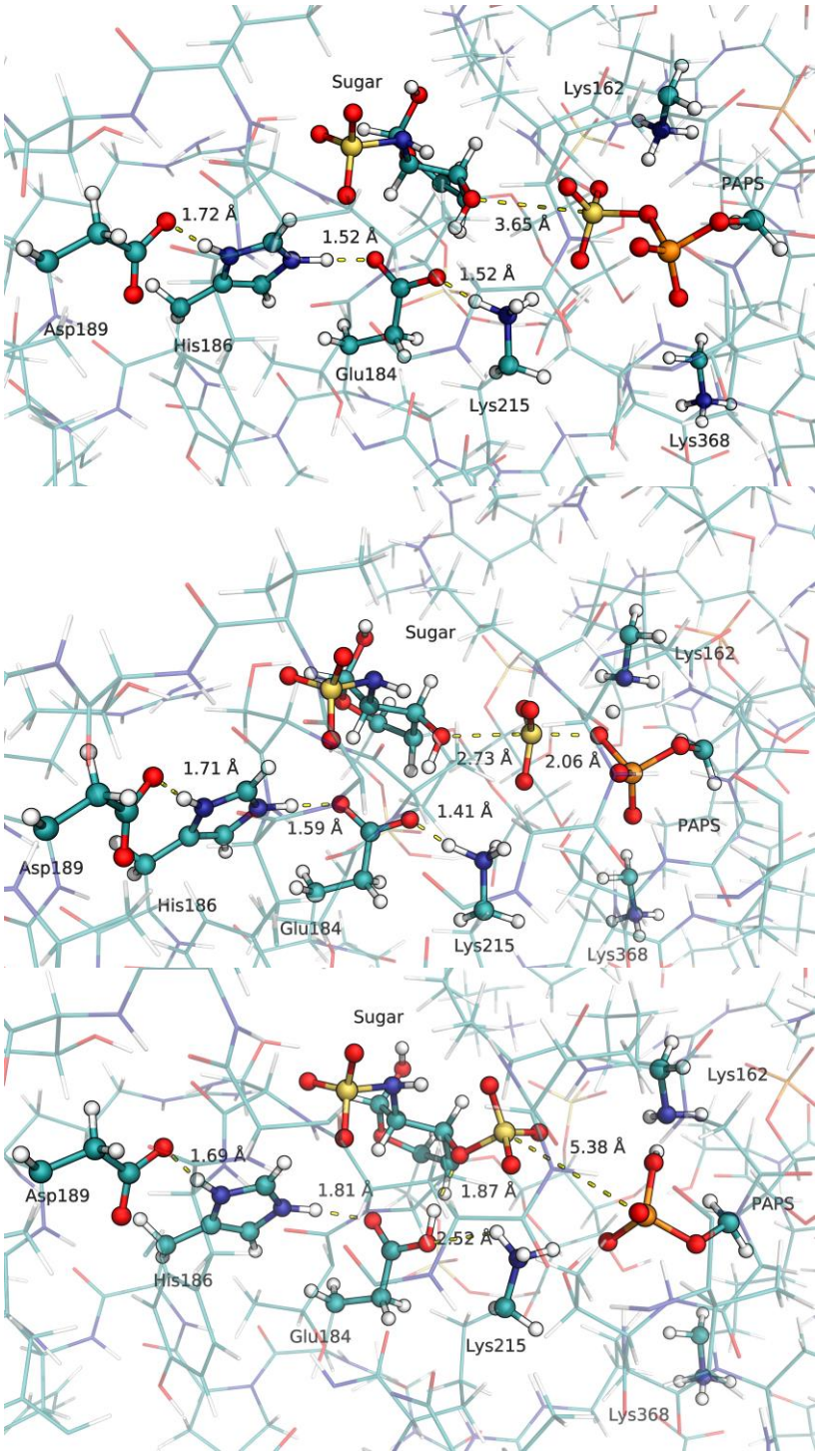


Figure S7 - Representation of the reactants (top), TS (middle) and products (bottom) structures obtained from the QM/MM calculations performed on the structure extracted at 6 ns from the MD simulations.