

Glucosylation Mechanism of Resveratrol through the Mutant Q345F Sucrose Phosphorylase from the organism *Bifidobacterium adolescentis*: a Computational Study

Electronic Supplementary Information

Camilo Febres-Molina,^a Leslie Sánchez,^a Xavier Prat-Resina^b and Gonzalo A. Jaña^{*c}

^a Doctorado en Físicoquímica Molecular, Facultad de Ciencias Exactas, Universidad Andres Bello, Santiago - Chile.

^b Center for Learning Innovation, University of Minnesota Rochester, Rochester, Minnesota 55904, United States.

^c Departamento de Ciencias Químicas, Facultad de Ciencias Exactas, Universidad Andres Bello, Concepción - Chile.

* Author to whom correspondence should be addressed.
gonzalo.jana@unab.cl

Contents

1	Fig. SF1 Domains in the crystallographic structure of SP.	3
2	Fig. SF2 RMSD of the <i>BaSP Q345F</i> and <i>BaSP</i> proteins.	3
3	Fig. SF3 RMSF per residue of the <i>BaSP Q345F</i> and <i>BaSP</i> proteins. . . .	4
4	Fig. SF4 B-factor of the aligned average structures <i>BaSP Q345F</i> and <i>BaSP</i> . 4	4
5	Fig. SF5 Alignment of the <i>BaSP</i> average structure with the P- and S-binding conformations.	5
6	Fig. SF6 Alignment of the <i>BaSP Q345F</i> average structure with the P- and S-binding conformations.	6
7	Fig. SF7 SASA values with respect to time of active site residues for both systems.	7
8	Fig. SF8 Radius of gyration of the whole protein for both systems.	7
9	Fig. SF9 Solvent-accessible volume of the biggest internal pocket for both systems.	8
10	Fig. SF10 Catalytic distances from unrestrained MD.	9
11	Fig. SF11 SPE calculations.	10
12	Eq. S1 Transition state theory Equation.	11

Supplementary Figures

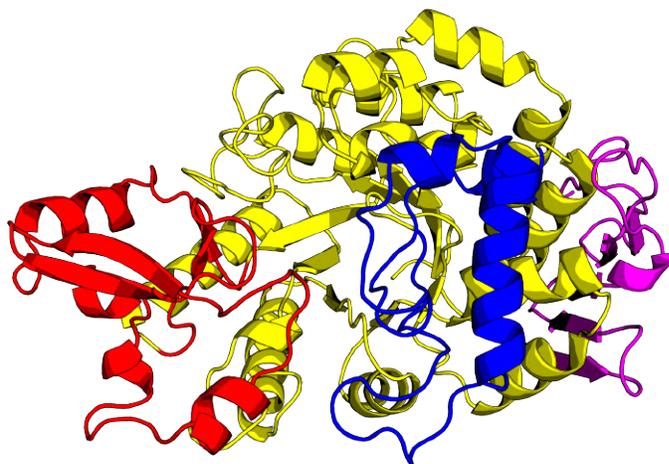


Fig. SF1 Crystallographic structure of SP from the organism *Bifidobacterium adolescentis* with PDB ID 2GDV - chain B. Domains A, B, B', and C are depicted in yellow, red, blue, and magenta, respectively.

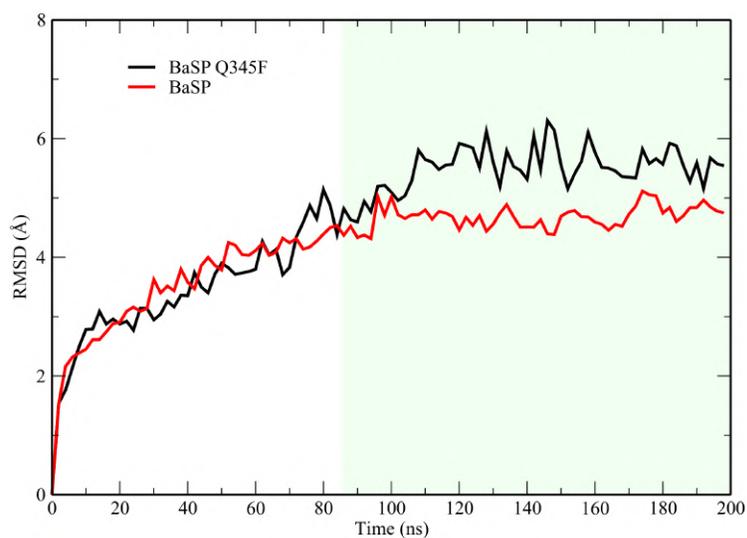


Fig. SF2 Root mean-square deviation (RMSD) of the *BaSP Q345F* and *BaSP* proteins. Both systems display stability during the last 115 ns (light green).

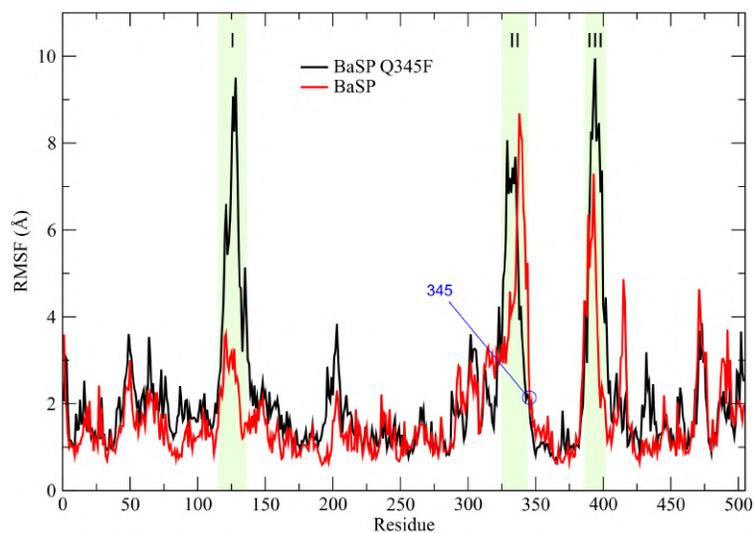


Fig. SF3 Root mean-square fluctuation (RMSF) per residue of the last 115 ns MD of the *BaSP Q345F* and *BaSP* proteins. The outer loops of the B, B', and A domains are represented in light green and marked with I, II, and III, respectively. Mutated residue 345 is shown in blue.

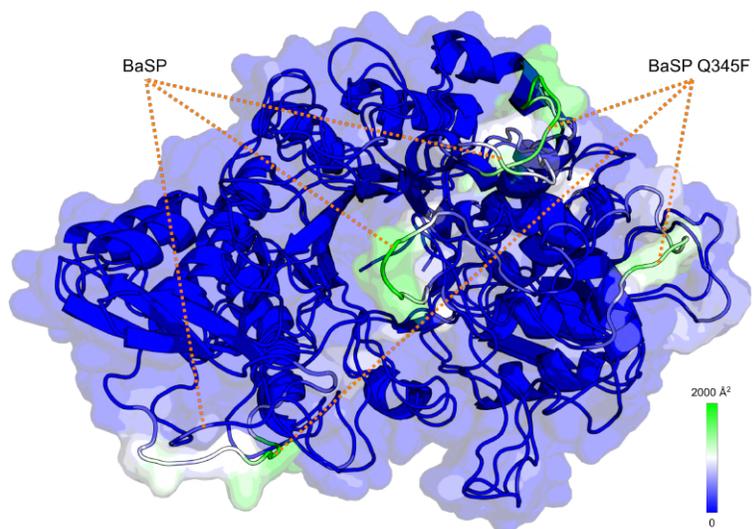


Fig. SF4 *BaSP Q345F* and *BaSP* average structures aligned with colors representing B-factor. Blue, white, and green colors indicate low, medium, and high residue fluctuations, respectively, of the last 115 ns of the MD.

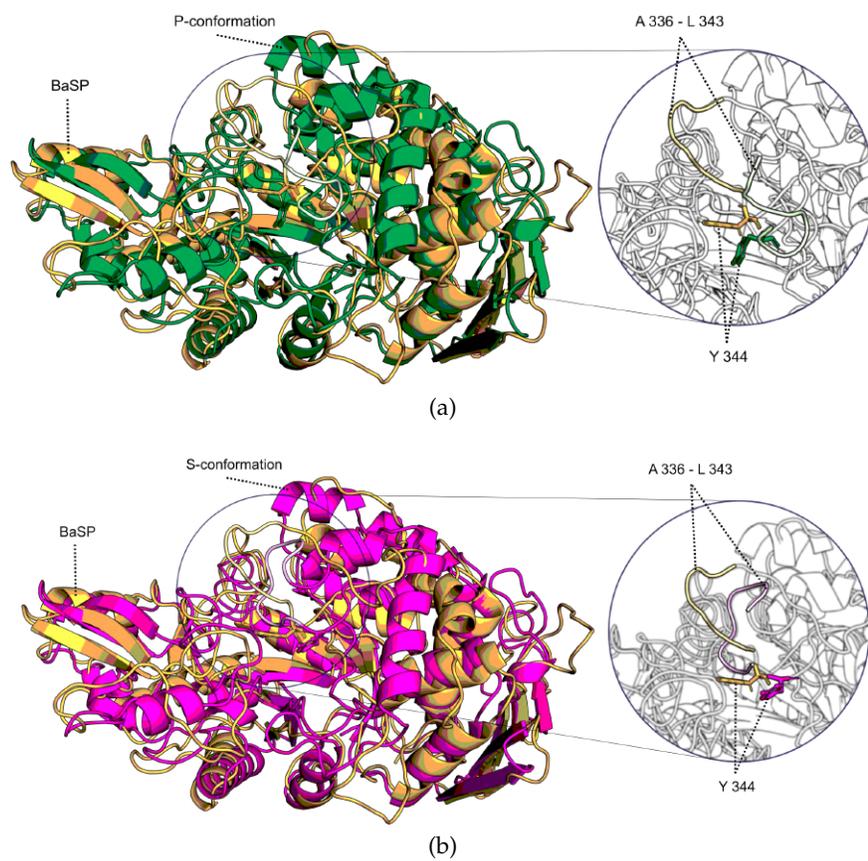


Fig. SF5 Alignment of the *BaSP* average structure with the phosphate- (a) and sucrose- (b) binding conformations with RMSD values of 2.6 and 2.8 Å, respectively. a) The zoom shows how the residues 336-343 of both structures point away from the active site, as well as both residues 344 belong to this site. b) Instead, the zoom here shows the loop of the S-conformation pointing towards the active site, while its 344 residue points away.

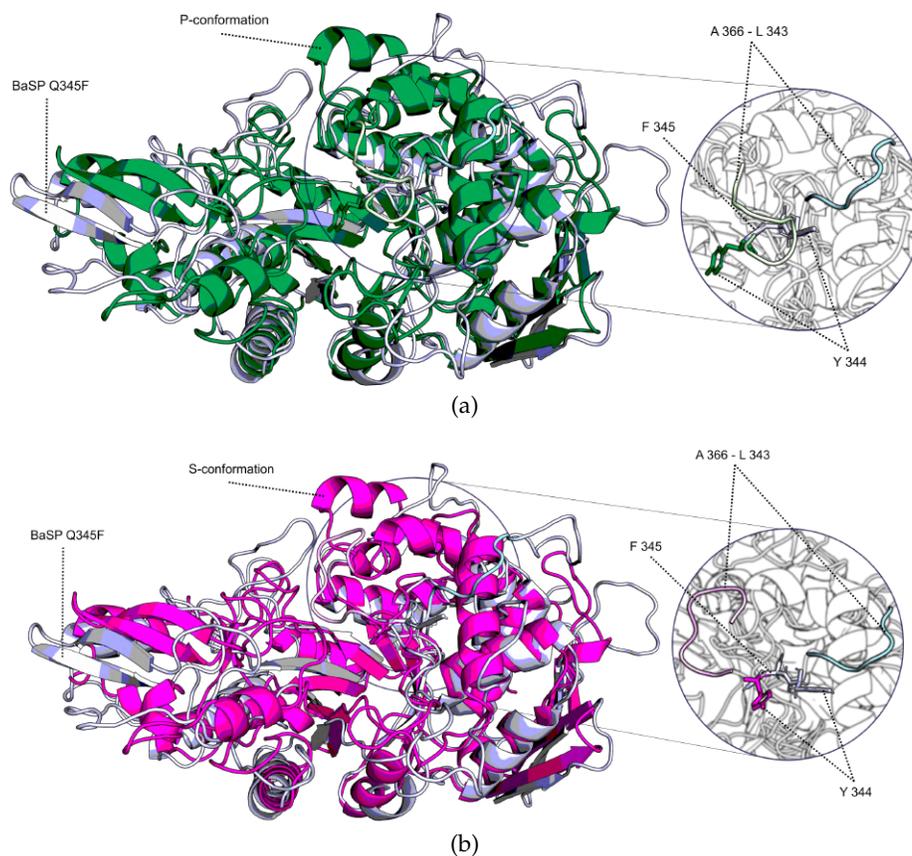


Fig. SF6 Alignment of the *BaSP Q345F* average structure with the phosphate- (a) and sucrose- (b) binding conformations with RMSD values of 2.3 and 2.4 Å, respectively. Due to steric hindrance generated by Phe³⁴⁵ residue in the mutant enzyme, its Tyr³⁴⁴ residue points outwards. a) The zoom shows how the residues 336-343 of both structures point away from the active site, but just residue 344 of the P-conformation belongs to this site. b) The zoom here shows the loop of the S-conformation pointing towards the active site, while both 344 residues point away.

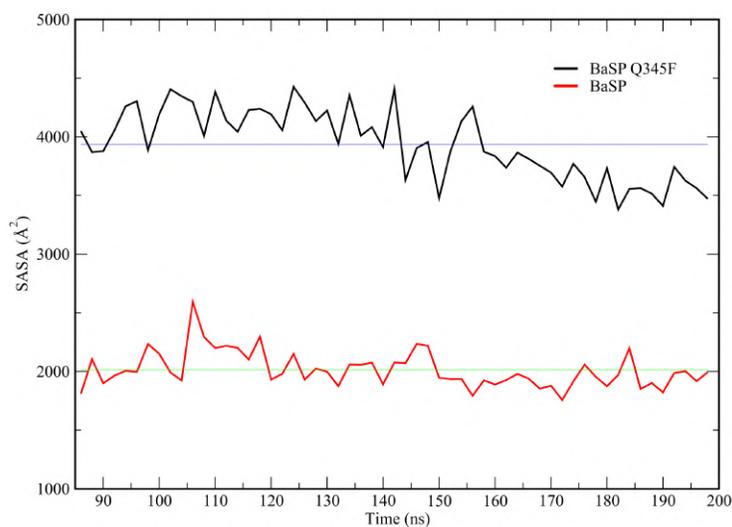


Fig. SF7 Solvent-accessible surface area (SASA) values of active site residues for both systems with respect to time during the last 115 ns. Blue and green lines represent de mean values of 3937.2 ± 297.6 and 2013.7 ± 152.4 \AA^2 for the mutant and native enzymes, respectively.

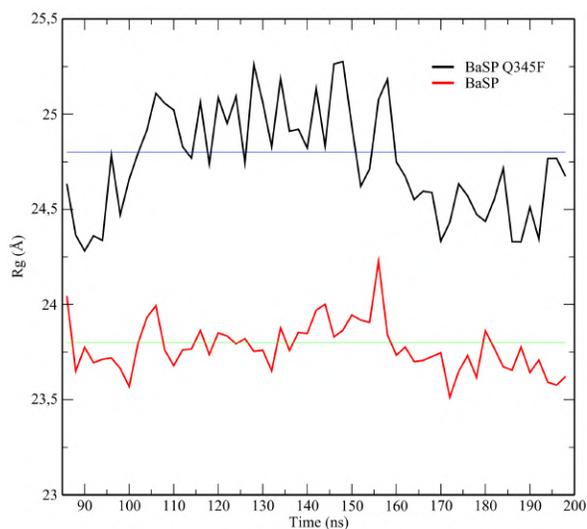
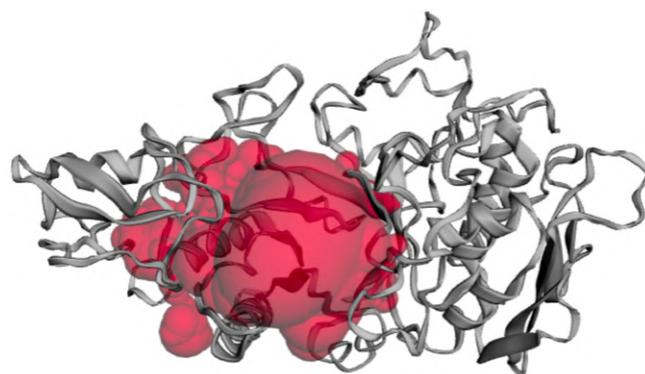
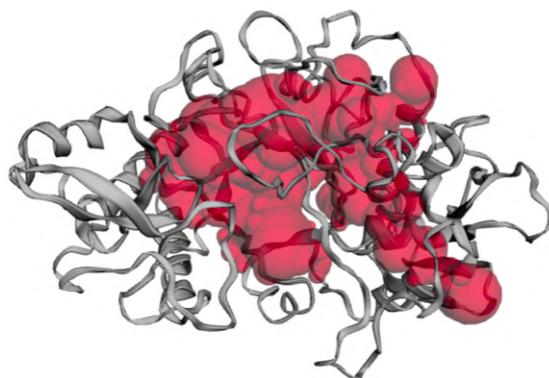


Fig. SF8 Radius of gyration (Rg) of the whole protein for both systems during the last 115 ns. Blue and green lines represent de mean values of 24.8 ± 0.3 and 23.8 ± 0.1 \AA for the mutant and native enzymes, respectively.

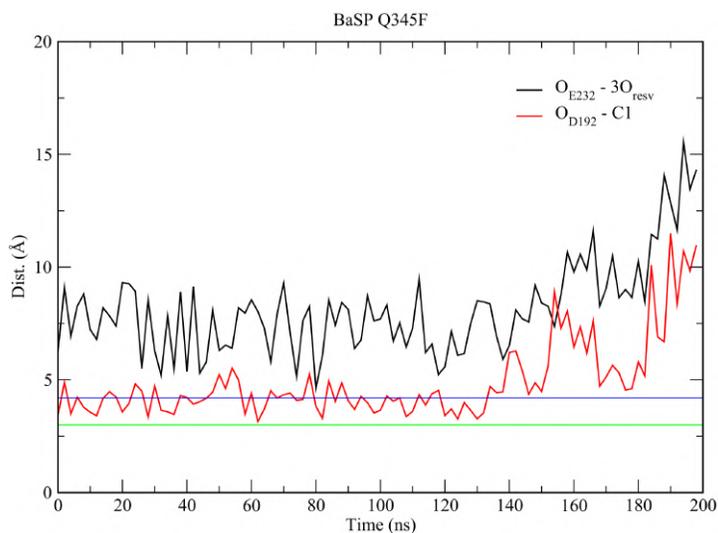


(a)

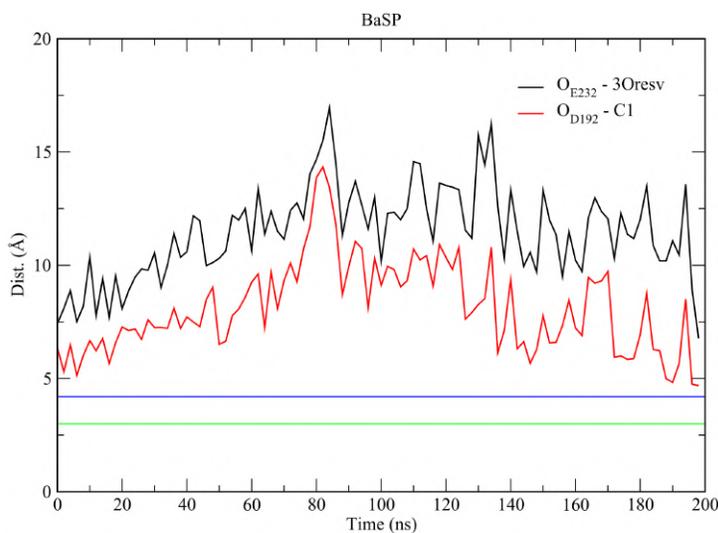


(b)

Fig. SF9 Solvent-accessible volume (SAV) represented in red spheres of the biggest internal pocket of the average proteins calculated using the CASTp webserver.¹ a) Mutant enzyme with a SAV pocket of 2991.2 Å³. b) Native enzyme with a SAV pocket of 2334.9 Å³.



(a)



(b)

Fig. SF10 Main catalytic distances along the unrestrained MD of the mutant (a) and wild-type (b) systems. Blue and green lines account for the optimal catalytic distances between $O_{E232}-3O_{resv}$ and $O_{D192}-C1$, respectively.

The goal of the MD simulations was to compare the binding between the monoglucoside resveratrol substrate with the wild-type and the mutant in the catalysis configuration. Figure SF10 shows the two main catalytic distances ($O_{E232}-3O_{resv}$ and $O_{D192}-C1$) but from the unrestrained MD. This figure tells us that the unrestrained MD of both mutant and wild-type possess structures that do not belong to the optimal catalytic configuration. Therefore, this additional unrestrained simulation did not provide additional information on the differences between the wild-type and the mutant. The reason for this could be because the nature of the interaction between a large substrate such as resveratrol and a relatively solvent-exposed active site is not rigid and some fluctuations appeared dur-

ing the unrestrained simulation that made the two systems more difficult to compare. For this reason, the MD simulations were performed under relatively mild restrain.

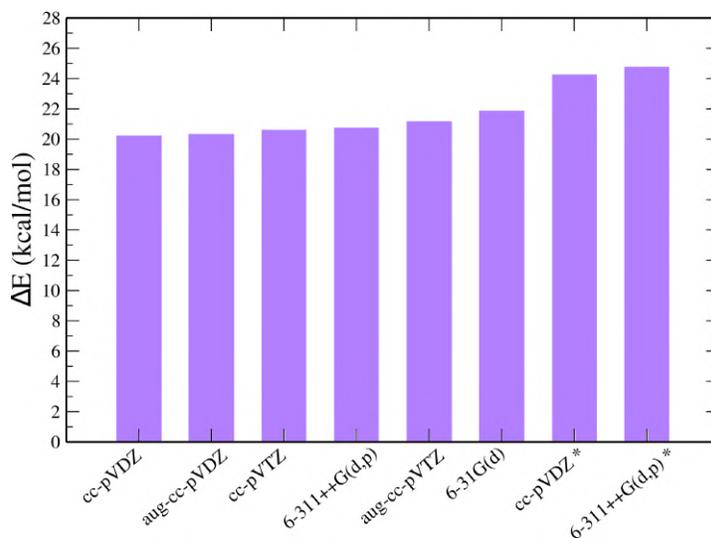


Fig. SF11 SPE calculations. The functional ω B97X-V and different basis sets were used. Asterisk (*) indicates that the D3(0) dispersion corrections were not included in those calculations.

Figure SF11 shows the SPE calculations. ΔE^\ddagger energy barriers of 20.3, 20.4, 20.6, 20.8, 21.2, and 21.9 kcal·mol⁻¹ were obtained using the functional ω B97X-V, the D3(0) dispersion corrections, and the following basis sets: cc-pVDZ, aug-cc-pVDZ, cc-pVTZ, 6-311++G(d,p), aug-cc-pVTZ, and 6-31G(d), respectively. This implies that although the size of the basis set increases and dispersion functions are added to it, there is a difference between these values of less than 2 kcal·mol⁻¹, which does not represent a significant difference in terms of theoretical-computational calculations, on the contrary, this allows us to use the computationally less expensive Dunning double- ζ basis set and still obtain very good results.

Calculations that did not have the D3(0) dispersion corrections (last two columns marked with an asterisk in Figure SF10) presented energy barrier values of 24.30 and 24.80 kcal·mol⁻¹ for the cc-pVDZ, and 6-311++G(d,p) basis sets, respectively. These last values present around 4 kcal·mol⁻¹ above the average of the previous ones, thus, showing the importance of including the aforementioned D3(0) dispersion corrections, which have empirical parameters.

Using the transition state theory equation (Eq. S1), and with the reported k_{cat} ($\sim 0.131 \text{ s}^{-1}$) and temperature ($37 \text{ }^\circ\text{C}$) for the transglucosylation reaction of *BaSP Q345F* with resveratrol it can be calculated the free energy barrier in order to have a reference energy value to compare the QM/MM results obtained.^{2,3}

$$\Delta^\ddagger G_{cat}^o \approx -RT \ln[(k_{cat}h)/(k_B T)] \quad (\text{Eq. S1})$$

Where R is the ideal gas constant, T is temperature, h is Planck's constant, and k_B is Boltzmann's constant.
The resulting value is $\sim 19.4 \text{ kcal/mol}$.

References

- [1] Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids research*. 2018;46(W1):W363–W367.
- [2] Kraus M, Grimm C, Seibel J. Switching enzyme specificity from phosphate to resveratrol glucosylation. *Chemical Communications*. 2017;53(90):12181–12184.
- [3] Sousa SF, Calixto AR, Ferreira P, Ramos MJ, Lim C, Fernandes PA. Activation free energy, substrate binding free energy, and enzyme efficiency fall in a very narrow range of values for most enzymes. *ACS Catalysis*. 2020;10(15):8444–8453.