Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2024

Non-*Aufbau* Electronic Structure in Radical Enzymes and the Control of the Highly Reactive Intermediates

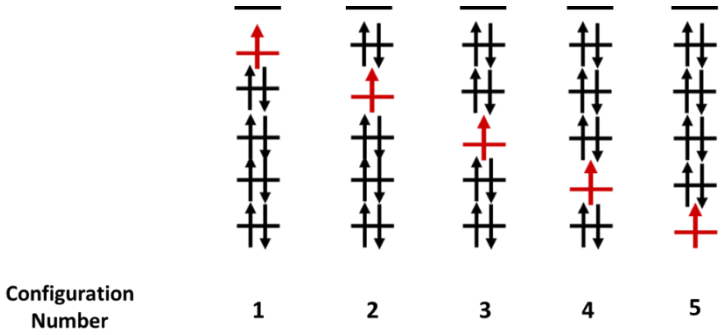
M. Hossein Khalilian, and Gino A. DiLabio*

Department of Chemistry, The University of British Columbia
3247 University Way, Kelowna, British Columbia, Canada V1V 1V7

Table of Contents

Computational Methods

CASSCF calculations: The choice of UM06-2X DFT functional for predicting electronic structures is based on the fact that this method was shown to perform well in previous related studies.[1–3] Nevertheless, we also performed our own benchmarking with high level *ab initio* multi-reference approach. This CASSCF calculations were performed on the GLM-Pre-s, GLM-Post-s, MCM-Pre-s, and MCM-Post-s models. We included five highest energy occupied MOs to active space involving those related to the glutamate residue and the radical centre, and one unoccupied MO associated with the β counterpart of the α-SOMO. Therefore the active space of our CASSCF has nine electrons and six orbitals (CASSCF(9,6)). A large basis set, 6-311+G(d,p), was used for these calculations. The major possible electronic configurations that could be occupied are given below. The contribution of these configurations to the total wave function is summarized in the table. For these CASSCF calculations GAMESS 2017 program was used.[4]



Five of the possible configurations for nine electrons and six orbitals (9,6)

The contribution of each configuration (in percentage) to the total wave function is shown in the table below. Configurations that are not reported in this table contribute a very small amount to the overall wave function. According to these results the electronic configuration predicted by the DFT calculations is the configuration that contributes the most to the CASSCF wave functions. These results, consistent with previous studies, justifies the use of the unrestricted DFT calculations with M06-2X functional for predicting the electronic structures.

GLM-Pre-s	Config. No.	Percentage (CASSCF)
	1	0%
	2	0%
	3	1%
	4	97%
	5	0%

GLM-Post-s	Config. No.	Percentage (CASSCF)
	1	91%
	2	2%
	3	4%
	4	1%
	5	0%

MCM-Pre-s	Config. No.	Percentage (CASSCF)
	1	0%
	2	14%
	3	10%
	4	74%
	5	0%

MCM-Post-s	Config. No.	Percentage (CASSCF)
	1	1%
	2	93%
	3	0%
	4	1%
	5	2%

Radical Centre Contribution (RCC): The RCC to the SOMO or HOMO is employed for the verification of the radical centre. RCC estimates the contribution of a particular atomic orbital – the carbon-centred radical p-type orbital in this case – to the MO of interest, for example the SOMO or HOMO. RCC values were determined using the following expression:

$$RCC\ (MO) = \frac{C_\lambda^2}{\sum_\mu C_\mu^2} \times 100\%$$

where C_λ is the particular atomic orbital coefficients associated with the radical center in a given MO, and $C\mu$ is the atomic orbital coefficients of all of the atoms contribute to the MO. For a particular MO, C_λ is the sum of the coefficients of the p-type orbitals on the carbon radical centre in which the unpaired radical is nominally contained, while the $C\mu$ is the atomic orbital coefficients of other atoms. The C_λ and $C\mu$ coefficients needed for the RCC calculations were obtained from the computations described previously.

***Density of State (DOS) Plots*:** In order to understand how the MOs composition derived from different chemical fragments in the systems studied, DOS plots were computed. The total DOS at energy E can be calculated by the following equation:

$$N(E) = \sum_i \delta(E - \epsilon_n)$$

where *N(E)* is the number of states, δ is the Dirac delta function, and ϵ_n is the MO energies.[5,6] The type of DOS analysis performed in this work is known as projected DOS (PDOS). In PDOS the total density is projected out into certain fragments using the following equation:

$$N_m(E) = \sum_i < \chi_\sigma | \phi_n > \delta(E - \epsilon_n)$$

The function χ_σ here shows the contribution of each fragment in the orbital φ_n.

***Barrier height calculations*:** Another set of models were prepared for the barrier height/reactivity calculations associated with the hydrogen abstraction step. These models are similar to the small models described above, but the substrate is included. In order to obtain transition states (TS) for hydrogen abstraction, the protonated version of the substrate, L-threo-3-methylaspartic acid, was used. Similarly, for obtaining MCM hydrogen abstraction's TSs the substrate, methylmalonyl-CoA, was protonated. The CoA tail of the substrate was capped after the carbon atom next to the sulfur atom. The protonation was done to remove unrealistic electrostatic interactions in small models. Given that the purpose of the TS calculations is to compare the relative reactivity of Ado$^{\bullet}$ in Pre and Post systems, protonation of the substrate is a reasonable simplification and also makes the model more realistic.

For the barrier height calculations of (s) models, the geometries of the E330, Ado$^{\bullet}$ (except the $^{\bullet}CH_2$ group) were kept fixed and the coordinates of the substrate were relaxed. This constrained optimization was performed in order to keep the geometry of Ado$^{\bullet}$ and Glu the same as in the enzyme's active site. This reduces the uncertainties associated with the use of small model structures and it better represents Ado$^{\bullet}$ in enzymes. Further, we chose TS structures in which the substrate had the same conformation in both Pre and Post models, i.e. they had the same orientation relative to the substrate (see Figure S5). Note that we are interested in the relative reactivity of the two models.

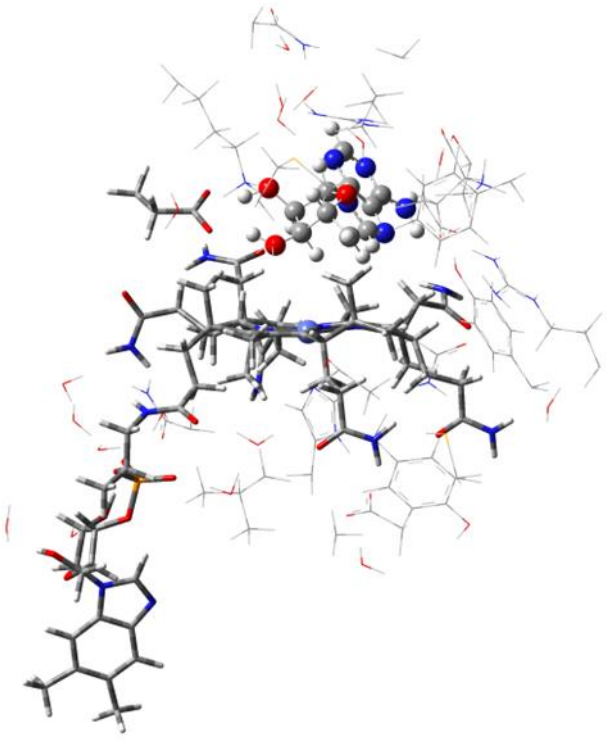
As mentioned in the main text, barrier height calculation is also performed when the side-chain of Glu^{-} residue is replaced by the classical partial charges (Figure S9 and S10). For this calculation the Mulliken partial charges were used and replaced the Glu^{-} atoms in both TS and reactants structures. These Mulliken charges were obtained from above calculations when the actual Glu^{-} side-chain presents. The single point calculations then performed to obtain the barrier height. All the calculations for obtaining the TSs and barrier heights were performed using full quantum mechanical calculations at DFT UM06-2X/6-311+G(d,p) level.

The influence of hydrogen bonding on SHI: It was shown in the main text that the magnitude of SHI in the Pre model is larger than that in the Post model. The main difference between these models is the degree of Ado$^{\bullet}$ and Glu^{-} hydrogen bonding (because of the different Ado$^{\bullet}$ geometry), suggesting that this could be the origin of the differences in the electronic structures and SHI.

The role of hydrogen bonding in mediating SHI was further explored in GLM-Post-s by rotating the $-COO^{-}$ moiety of Glu side chain (figure S3). For these calculations, we optimized the GLM-Post-s geometry and performed single-point calculations at each rotation about the plane of the hydrogen bond. As can be seen in Figure S3, at $\Phi=80^{o}$ the hydrogen bond is largely broken, and SHI is at its highest (SOMO energy is at its lowest). As the Glu^{-} is rotated towards $\Phi=180^{o}$ and the hydrogen bond is reformed, SHI diminishes. The appearance of the SHI when the hydrogen bond is lost is due to the stabilization of the SOMO and destabilization of the MOs that correspond to the Glu^{-} residue (in agreement with DOS plots in Figure 4). All of these findings point to the fact that the magnitude of SHI and the electronic structure in these systems is significantly influenced by hydrogen bond interactions between the Ado$^{\bullet}$ and Glu^{-}. We believe that the hydrogen bond interactions result in a mixing of the orbitals associated with

Glu^- and $Ado^\bullet$, which leads to the elimination of the effect. This suggests that B_{12} enzymes can make use of $Ado^\bullet$ and Glu^- hydrogen bonding interactions to control the SHI magnitude.

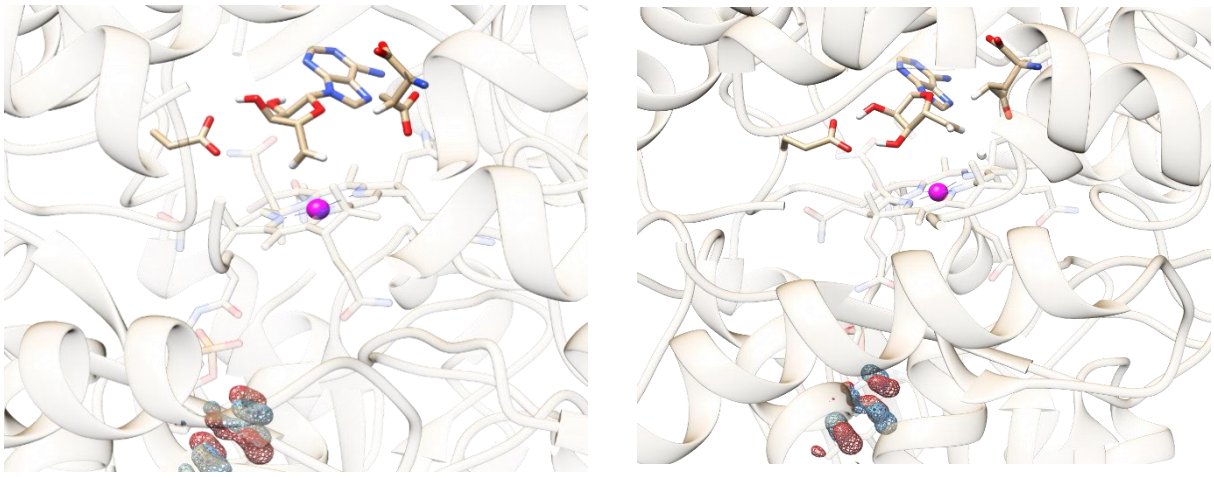
Figures and Discussions



Residues included in the QM region:

Ado•, B_{12} cofactor, substrate, Asp14, Cys15, His16, Ala17, Leu63, Tyr64, Arg66, Ala67, Asn93, Thr94, Arg100, Asn123, Tyr181, Phe216, Thr220, Met294, Lys326, Glu330, Asn340, 18 nearby crystal waters (480 atoms)

Figure S1 | QM region (charge=0) used for the GLM models with residues included. The QM residues are truncated after the alpha carbons. For clarity B_{12} and Ado radical are shown in with bigger atoms.



GLM-Pre
HOMO (-45 kJ/mol)

GLM-Post
HOMO (-45 kJ/mol)

Figure S2 | MO plots of HOMO with their associated absolute MO energies for GLM models. The electron density is distributed over the base (DMB) of B_{12} cofactor and no electron density is present on the radical centre in these MOs (RCC=0%). Absolute energy of HOMOs in both Pre and Post models are the same.

Method/6-311+G(d,p)	Model	SOMO = HOMO-X	ε_{SOMO} (kJ/mol)
M06-2X	GLM-Pre	19	-231
	GLM-Post	6	-145
BP86	GLM-Pre	23	+26
	GLM-Post	4	+118
PBE0	GLM-Pre	25	-110
	GLM-Post	6	-23

Table S1. The number and absolute energy of the SOMO associated with Ado• for Pre and Post GLM models predicted by different computational methods.

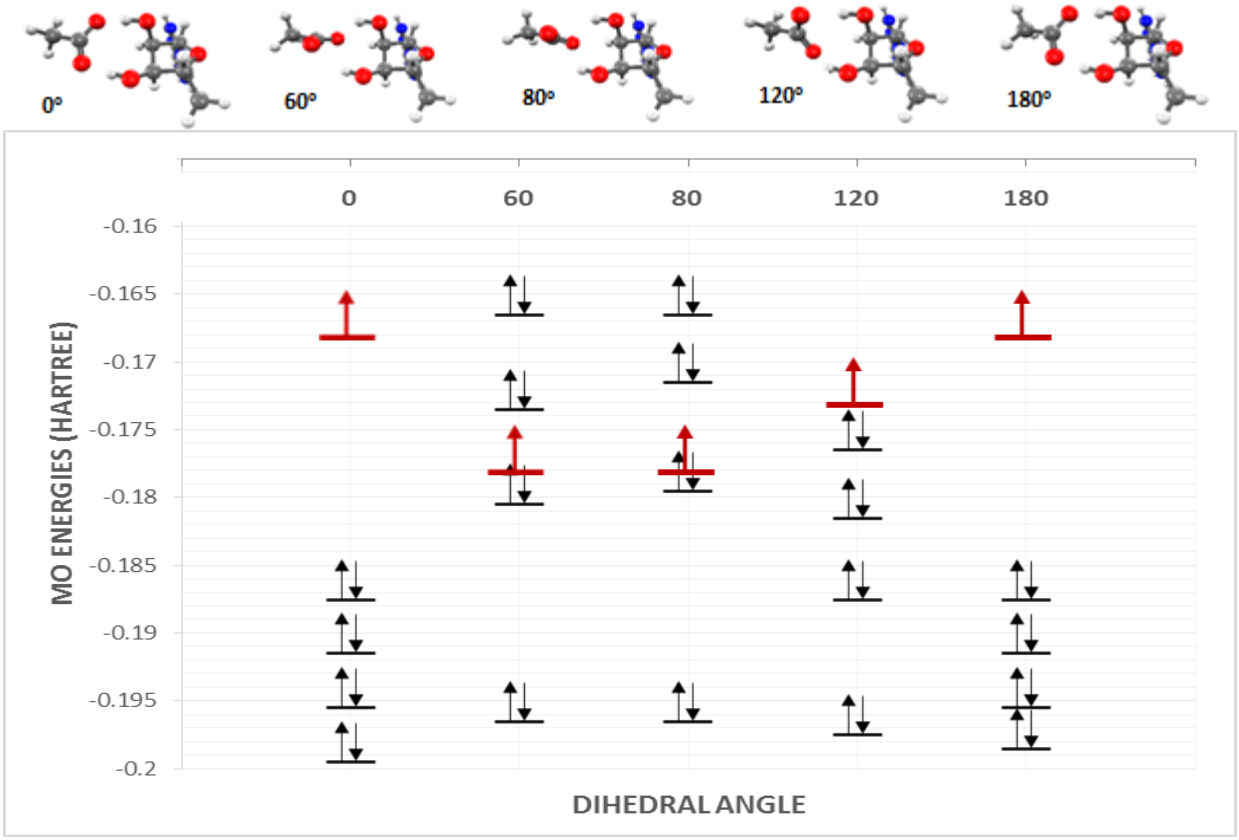


Figure S3 | The influence of the hydrogen bonding between Glu^- and Ado• on SHI. At 0° and 180° the hydrogen bonding is maximum and the SOMO is the orbital with the highest energy. When E330 is rotated by 60° or 80°, the SHI is present and the SOMO moves to lower energies while the Glu^- related MOs move to higher energies.

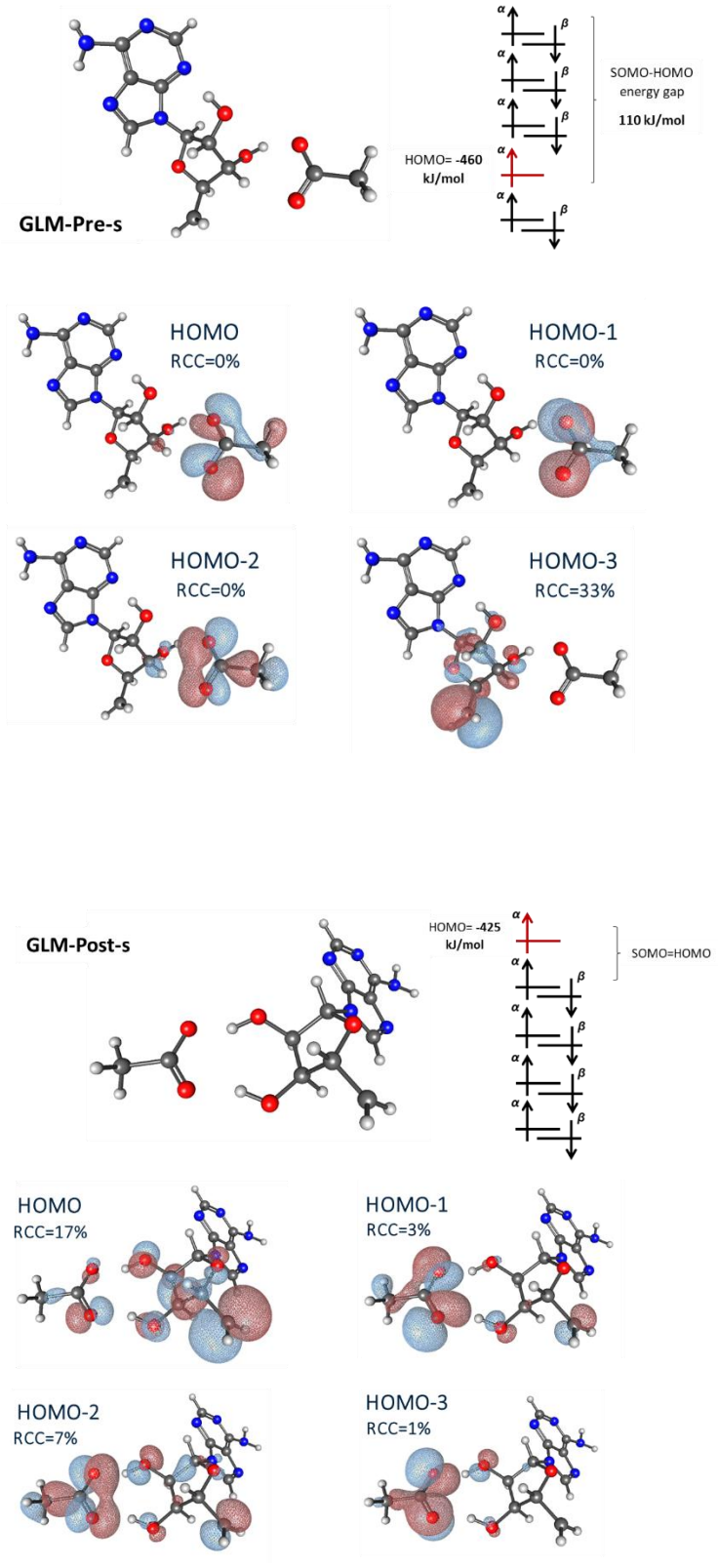


Figure S4 | Illustration of the orbital configuration, the absolute SOMO energy, SOMO-HOMO energy gap, and MO plots for four highest energy α-MOs with RCC values in GLM-Pre-s and GLM-Post-s models.

Figure S5 | The TS structures obtained from full-QM calculations.
enzyme b, MCM enzyme.
having the same conformation.

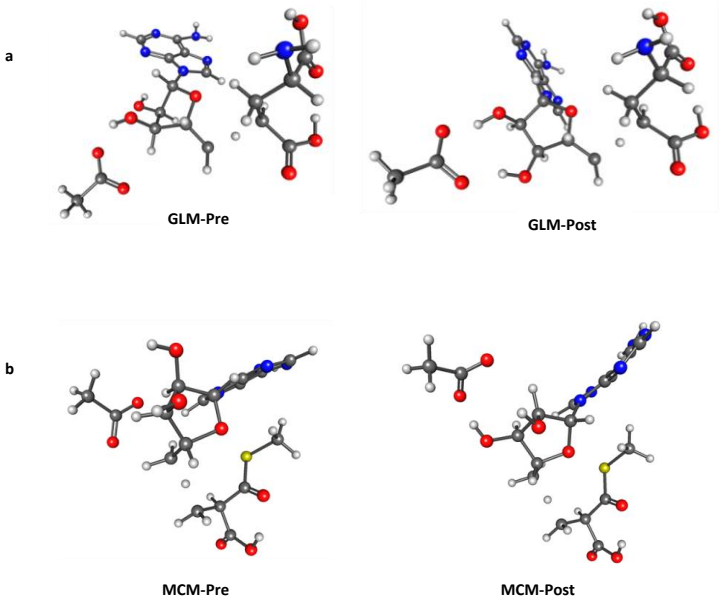
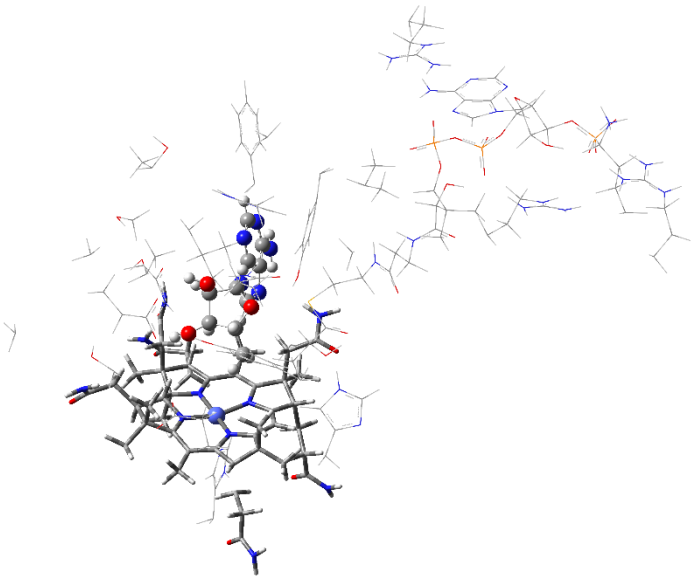


Table S2. Calculated barrier heights using the small models with comparisons to the measured tritium-exchange rate constants in GLM as presented in ref. 68. For these calculations the distance between the adenosyl 5′-carbon and the substrate hydrogen is adjusted for each mutants according to the distances reported in litrature. The overall trend of the calculated barrier heights are in agreement with the overall trend of the experimental rate constants.

Mutation	Barrier Height (kJ/mol)[a]	K_T (x 10^3 s^{-1})[b]
Wild type[a]	24.7	5000
Glu330Gln	47.7	50
Glu330Asp	54.9	20

[a]From this work. [b]From ref. 68 and 69.



Residues included in the QM region: Ado•, truncated B_{12} cofactor, substrate, Arg82, Tyr89, Ala90, Phe92, Ala116, Ala139, Val141, Tyr243, His244, Glu247, Arg283, Lys321, Arg326, Gln330, Thr331, Ser332, His364, Thr365, Asn366, Ser367, Glu370, Ala371, Ala373, Leu374, Arg383, His610 (total 551 atoms)

Figure S6 | QM region (charge=-1) used for the MCM models with residues included. The residues are truncated after the alpha carbons. For clarity B_{12} and Ado radical are shown with bigger atoms.

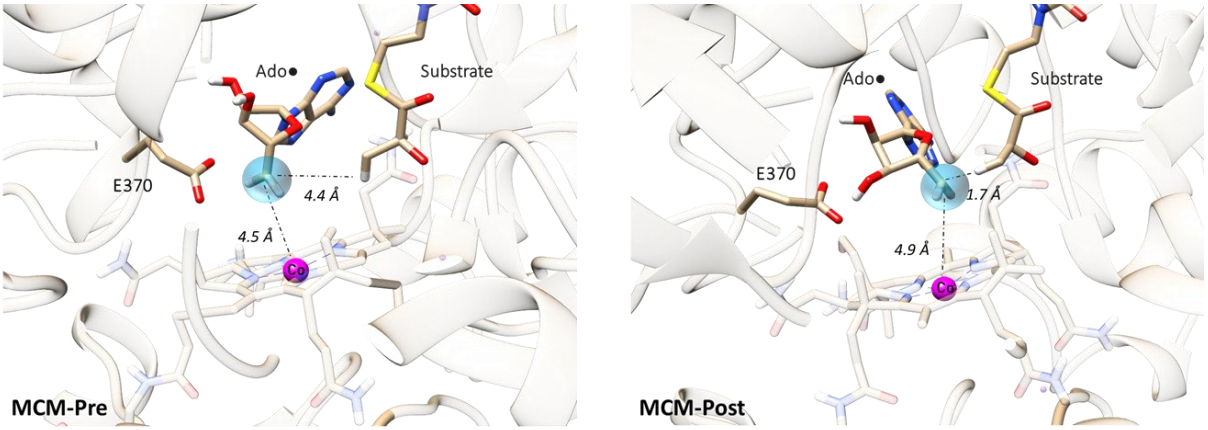


Figure S7 | Models of MCM enzyme at before and after migration of Ado radical. These models are obtained from previous study as noted in the computational methods.

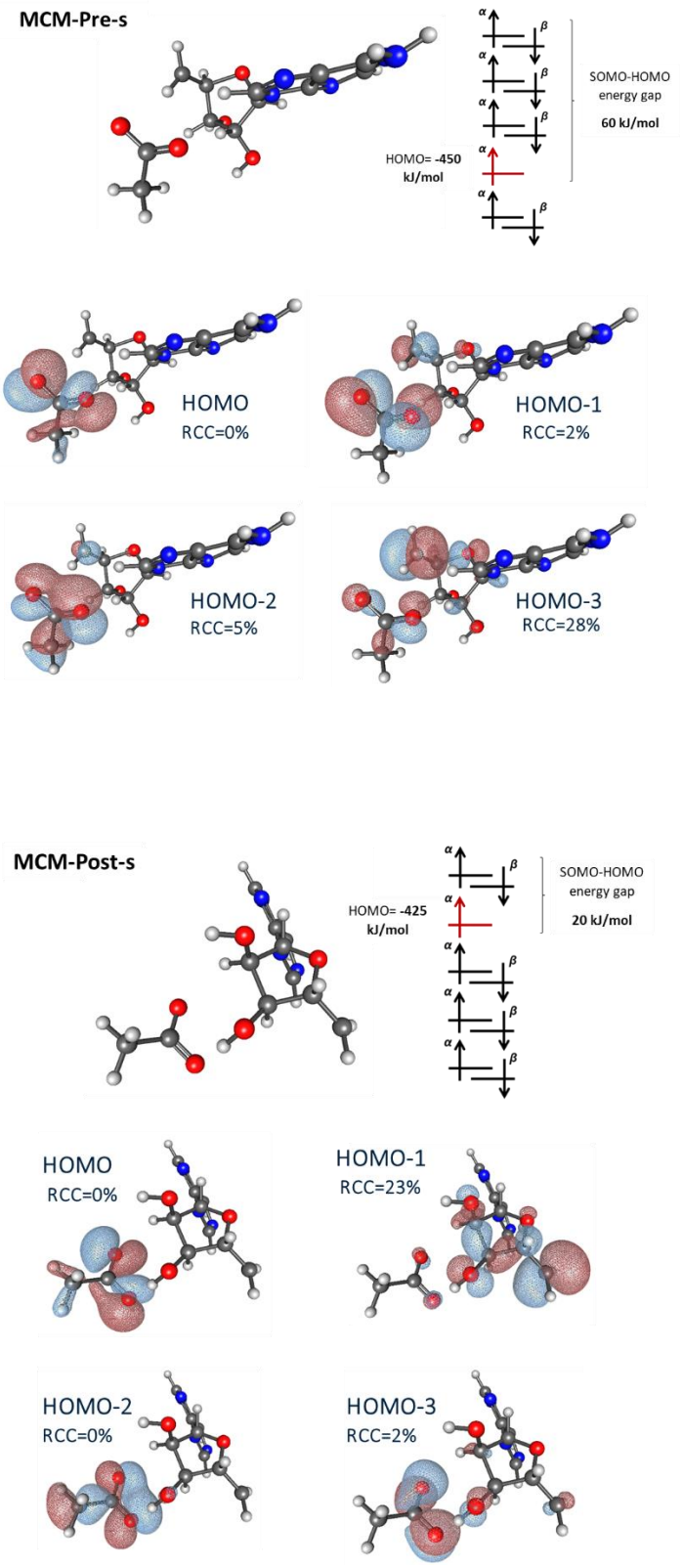
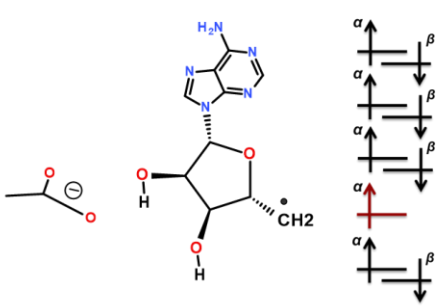
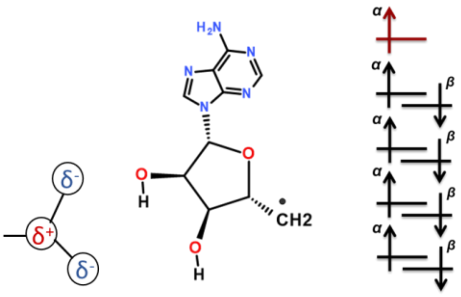


Figure S8 | Illustration of the orbital configuration, SOMO-HOMO energy gap, the absolute SOMO energy, and MO plots for four highest energy α-MOs with RCC values in MCM-Pre-s MCM-Post-s models.



Electronic structure of the GLM-Pre(s) when the glutamate side chain presents



Electronic structure of the GLM-Post(s) when the glutamate side chain is replaced with partial point charges

Figure S9 | The Illustration of the electronic structure of GLM-Pre-s model in the presence of the glutamate and a classical partial negative point charges. As can be seen in this figure we have found that the non-Aufbau electronic structure which exists in the GLM-Pre-s model disappears in the presence of the partial charges.

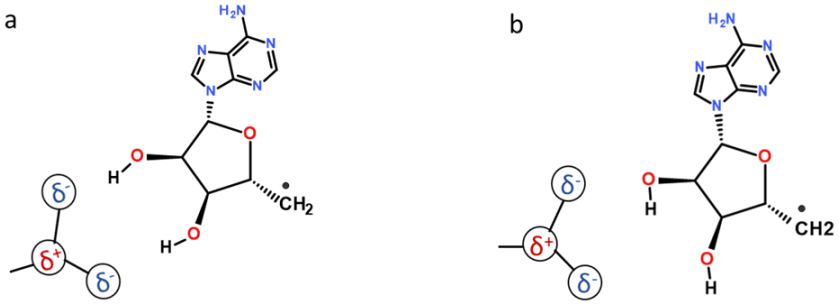


Figure S10 | a. The GLM-Pre(s) when the Glu^- is replaced with partial point charges. **b**, The GLM-Post(s) when the Glu^- side chain is replaced with partial point charges. The reactivity of **a** is calculated to be 49.1 kJ/mol while the reactivity of **b** is 49.0 kJ/mol. Thus, the reactivity of both structures is approximately the same when the glutamate residue is replaced by the partial point charges. This is consistent with the absolute SOMO energies of these structures which are approximately the same as well (ε_{Pre}=-510kJ/mol and ε_{Post}=-520kJ/mol). These results suggest that the barrier reduction (and change in SOMO energies) in the actual models are a result of a quantum effect. Note that partial point charges were obtained from previous calculation in which the Glu^- was present.

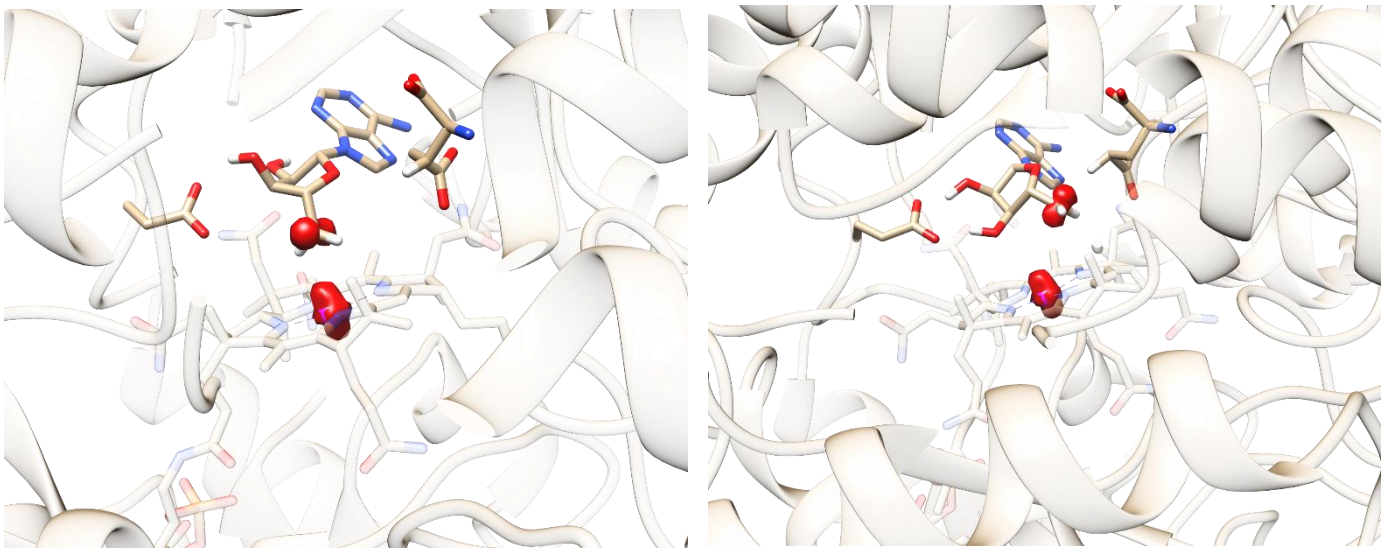


Figure S11 | Spin density plots for GLM-Pre and GLM-Post models.

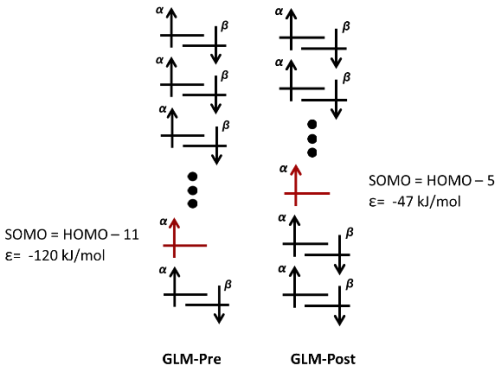


Figure S12 | The electronic structures obtained after optimization. The SOMO in the GLM-Pre structure is lower in energy compared to GLM-Post, therefore SHI is larger in GLM-Pre as expected.

References

1. Gryn'ova, G., Marshall, D. L., Blanksby, S. J. & Coote, M. L. Switching radical stability by pH-induced orbital conversion. *Nat. Chem.* **5**, 474–481 (2013).
2. Gryn'ova, G. & Coote, M. L. Origin and Scope of Long-Range Stabilizing Interactions and Associated SOMO–HOMO Conversion in Distonic Radical Anions. *J. Am. Chem. Soc.* **135**, 15392–15403 (2013).
3. Kumar, A. & Sevilla, M. D. Proton Transfer Induced SOMO-to-HOMO Level Switching in One- Electron Oxidized A - T and G - C Base Pairs: A Density Functional Theory Study. *J. Phys. Chem. B* **118**, 5453–5458 (2014).
4. Schmidt, M. W. *et al.* General atomic and molecular electronic structure system. *J. Comput. Chem.* **14**, 1347–1363 (1993).
5. *Electronic Structure and Electronic Transitions in Layered Materials*. *Electronic Structure and Electronic Transitions in Layered Materials* (Springer Netherlands, 1986). doi:10.1007/978-94-009-4542-5.
6. Rangel, T., Rignanese, G.-M. & Olevano, V. Can molecular projected density of states (PDOS) be systematically used in electronic conductance analysis? *Beilstein J. Nanotechnol.* **6**, 1247–1259 (2015).