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

Supporting Information

Impacts of QM Region Size and Conformation Number on Modelling Enzyme Reactions: A Case Study of Polyethylene Terephthalate Hydrolase

Mingna Zheng[1], Yanwei Li[1]*, Qingzhu Zhang[1], Wenxing Wang[1]

[1]Environment Research Institute, Shandong University, Qingdao, 266237, PR China

*Corresponding author.

Yanwei Li: E-mail: lyw@sdu.edu.cn

Number of pages: 22 (S1-S22)

Number of Figure: 12 (Figure S1-S12)

Number of Table: 3 (Table S1 to S3)

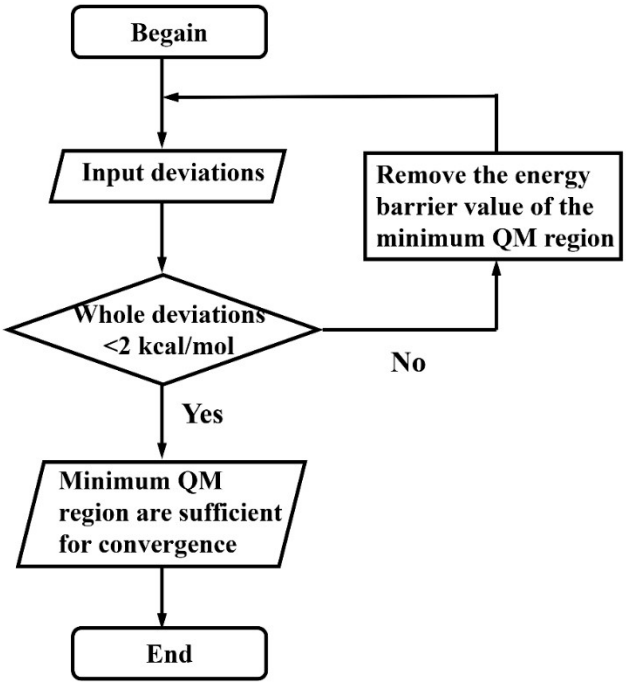
Contents

Theoretical details

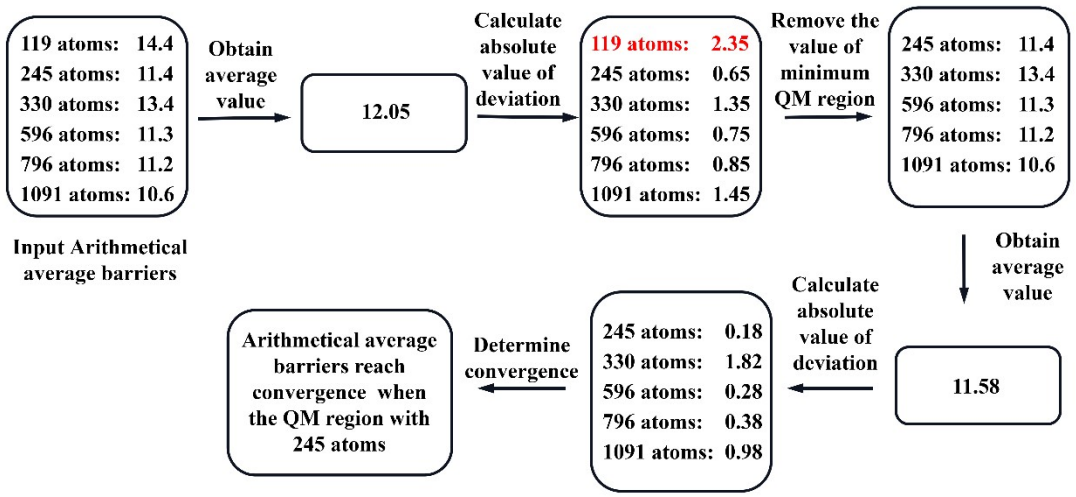
1. The criterion of determine the convergence behavior of energy barrier

In this study, the calculated way of relative energy barrier of convergence behavior is as follows:

The deviation of barrier of six different QM regions was calculated based on their average barrier, respectively. If the six deviation values are all less than 2 kcal/mol, we think that atom number in the minimum QM region are sufficient for convergence. However, when there is a value greater than 2 kcal/mol among the six deviations, we will remove the energy barrier value of the minimum QM region and repeat the above process.



Taking the arithmetical average barrier of step i as an example:



2. The reasons for selecting the BP86/6-31G(d) for geometry optimizations and M06-2X/6-311G(d,p) for single-point energy calculations are as follows:

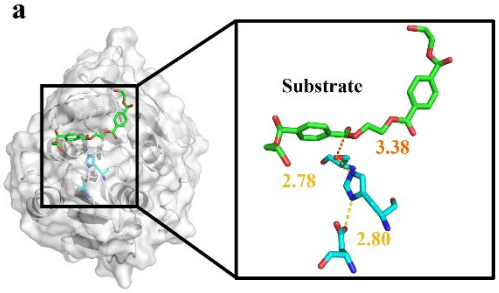
(1) This study mainly focuses on investigating the influence of conformation number and QM region size on the reaction energy barriers calculated at DFT level.

(2) We tested the optimization with high-level functional M06-2X, the results show that optimization of four elementary steps took at least one month for relatively smaller QM region (~300 atoms) with routinely used servers (e.g. 120 cores). Therefore, due to the fact that extremely big QM region (100-1000 atoms), six enzyme conformations, and four elementary steps were considered, extremely long time will be needed if M06-2X/6-31G(d) level were used during the optimization.

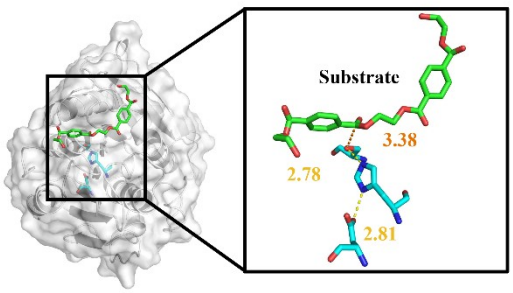
(3) The density functional BP86 was widely used in the study of large QM systems to speed up the calculations.[1-3] In order to reduce the computational cost, the functional BP86 was selected for the optimization in this study. We tested the calculation time for different functionals (BP86 and M06-2X). The results show that BP86/6-31G(d) is about 3.5 times faster than M06-2X/6-31G(d) method for the QM region size of 245 QM atoms. In addition, we compared the fluctuation of key structures of six enzyme conformations in R1 (100 QM region atoms) optimized by BP86/6-31G(d) and M06-2X/6-31G(d). As shown in **Table S1**, the results show that the fluctuation range of key structures is similar, indicating that the functional BP86 is reasonable to investigate the influence of conformation number and QM region size on the reaction energy barriers.

(4) To keep consistency with our previous study,[4] M06-2X/6-311G(d,p) was selected for single-point energy calculations in this study.

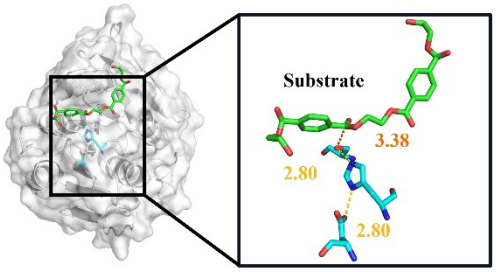
a



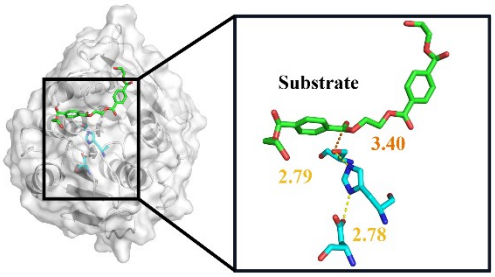
Conformation 1: 5.13 ns



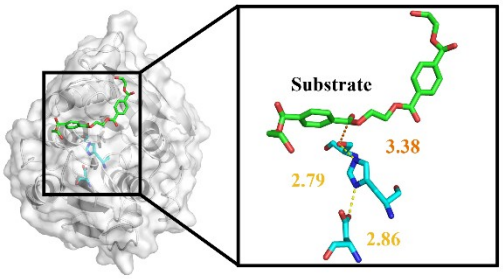
Conformation 2: 6.73 ns



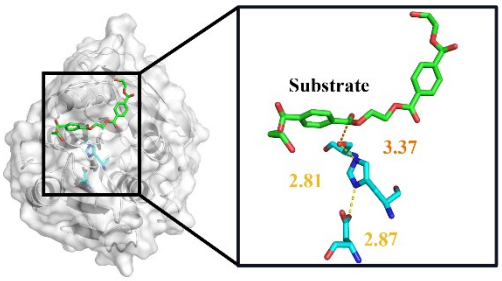
Conformation 3: 10.25 ns



Conformation 4: 18.44 ns



Conformation 5: 25.56 ns



Conformation 6: 27.24 ns

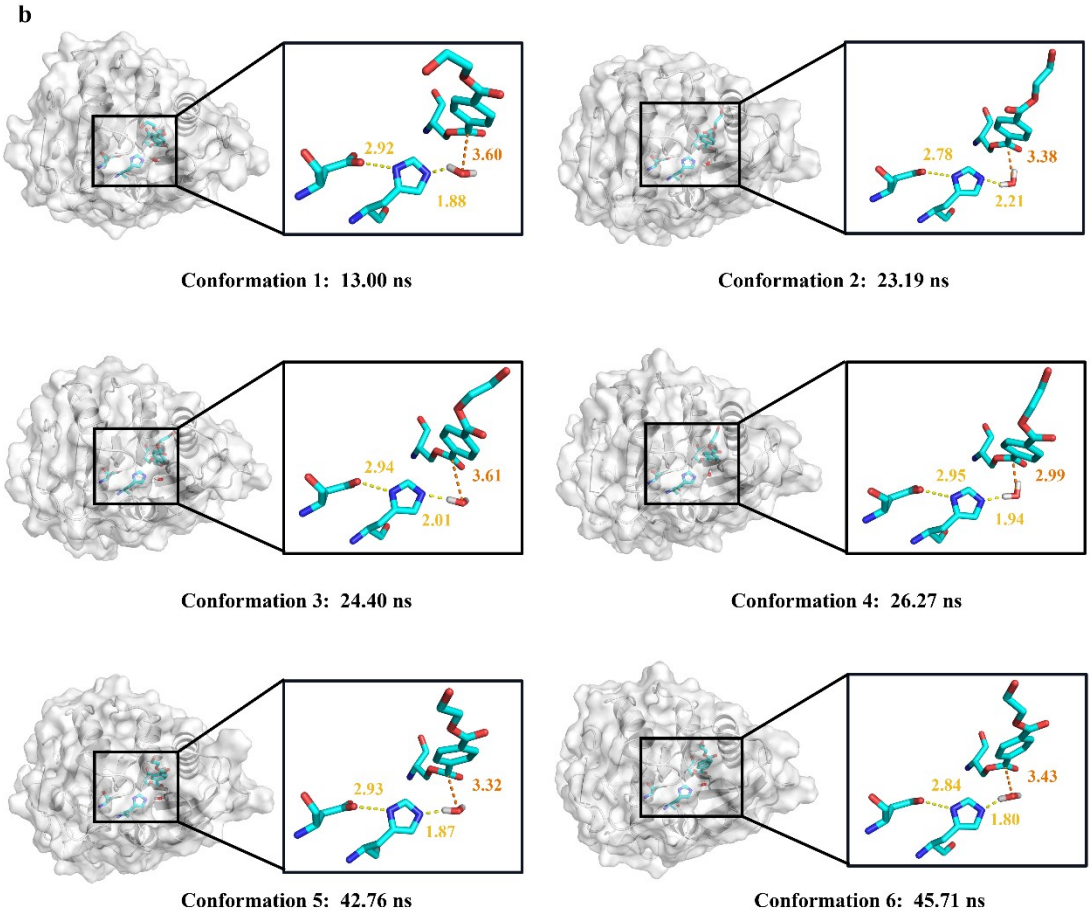


Fig. S1. Six independent enzyme conformations for acylation (a) and deacylation (b) processes.

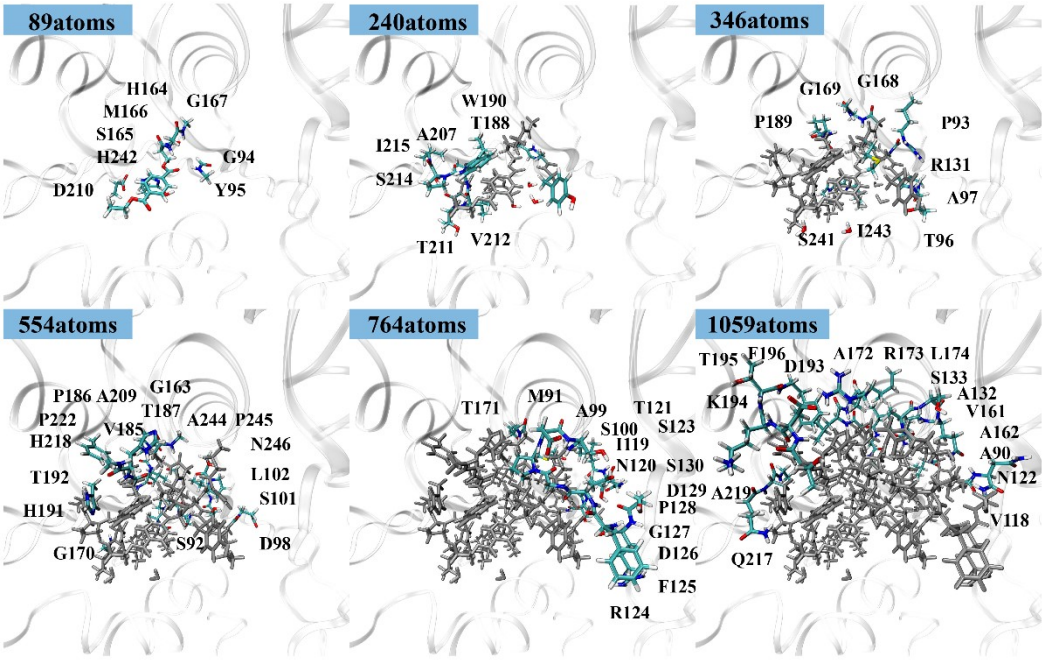


Fig. S2. Structures of six QM region sizes for deacylation process. The added group was shown in blue stick.

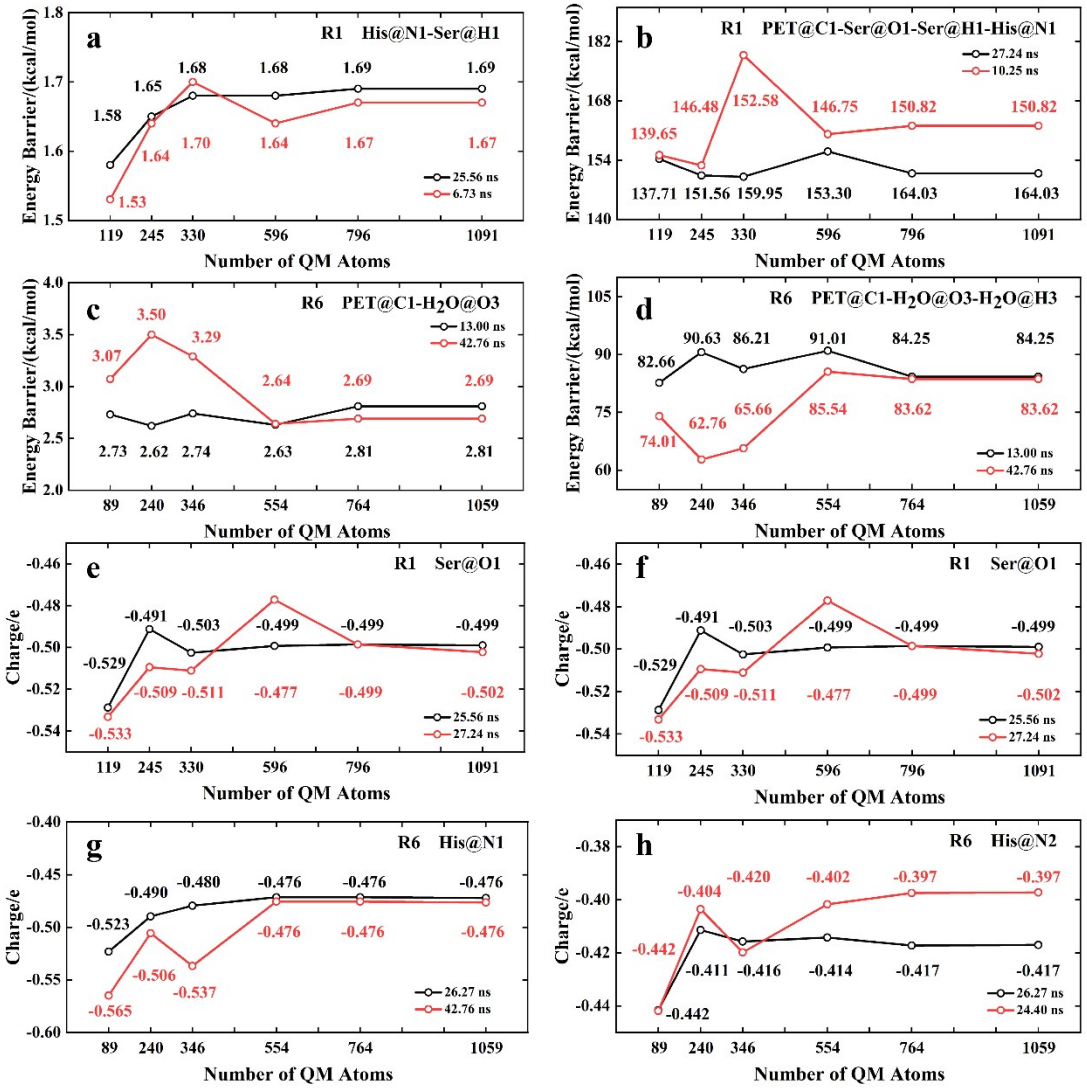


Fig. S3. Variation of key structural and charge parameters with expanding QM region for a single conformation. **a-d** Variations of key structural parameters with increasing QM region. **d-f** Variations of key charge parameters with increasing QM region.

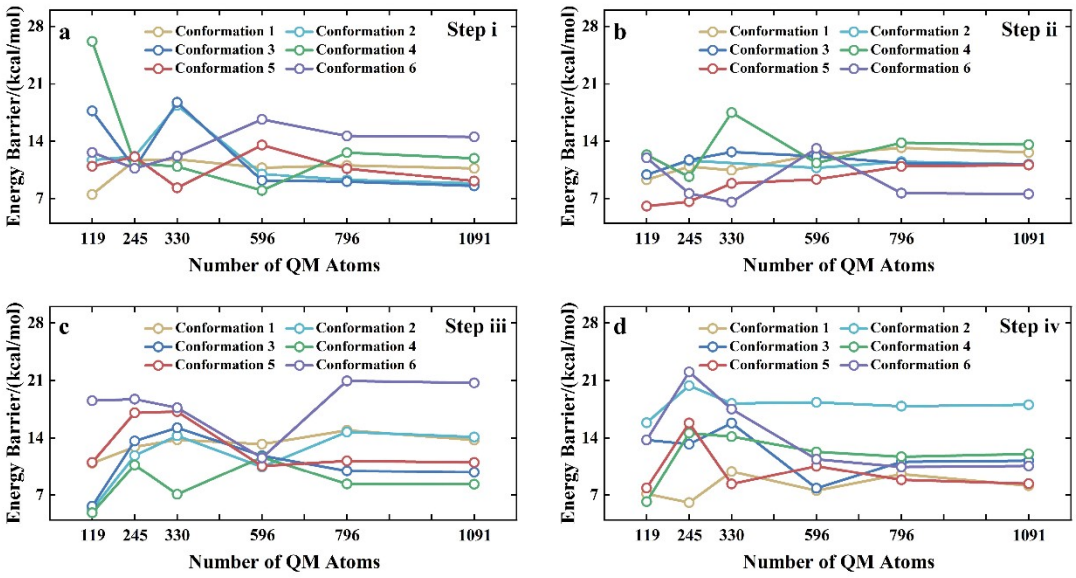


Fig. S4. Variation of energy barriers for six independent enzyme conformations. **a-d** represents elementary steps i to iv, respectively.

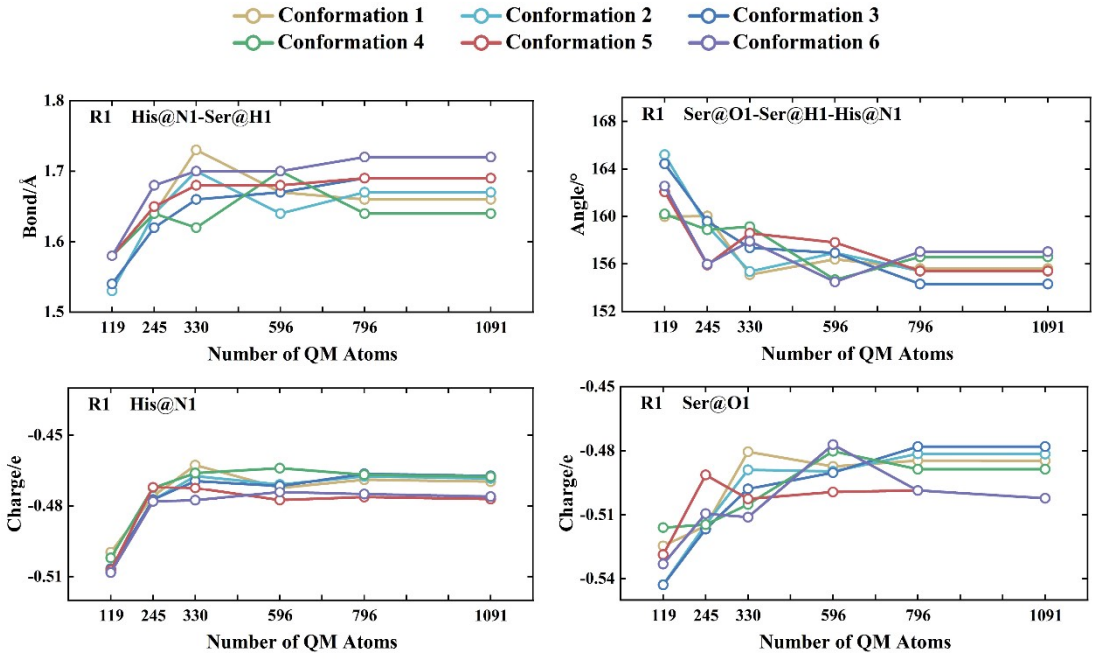


Fig. S5. Variation of key structural and charge parameters for six independent enzyme conformations with respect to QM region in R1.

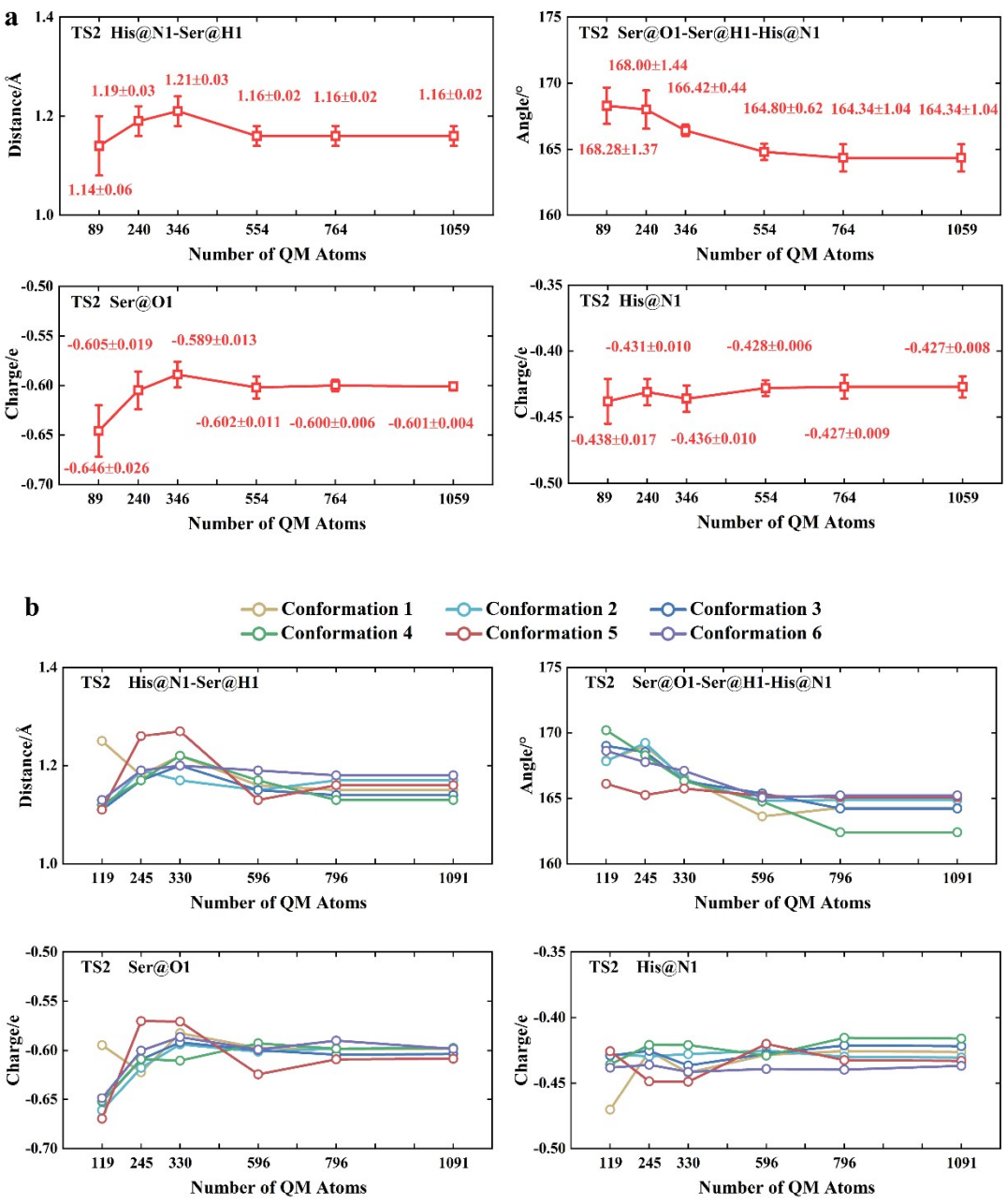


Fig. S6. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in TS2.

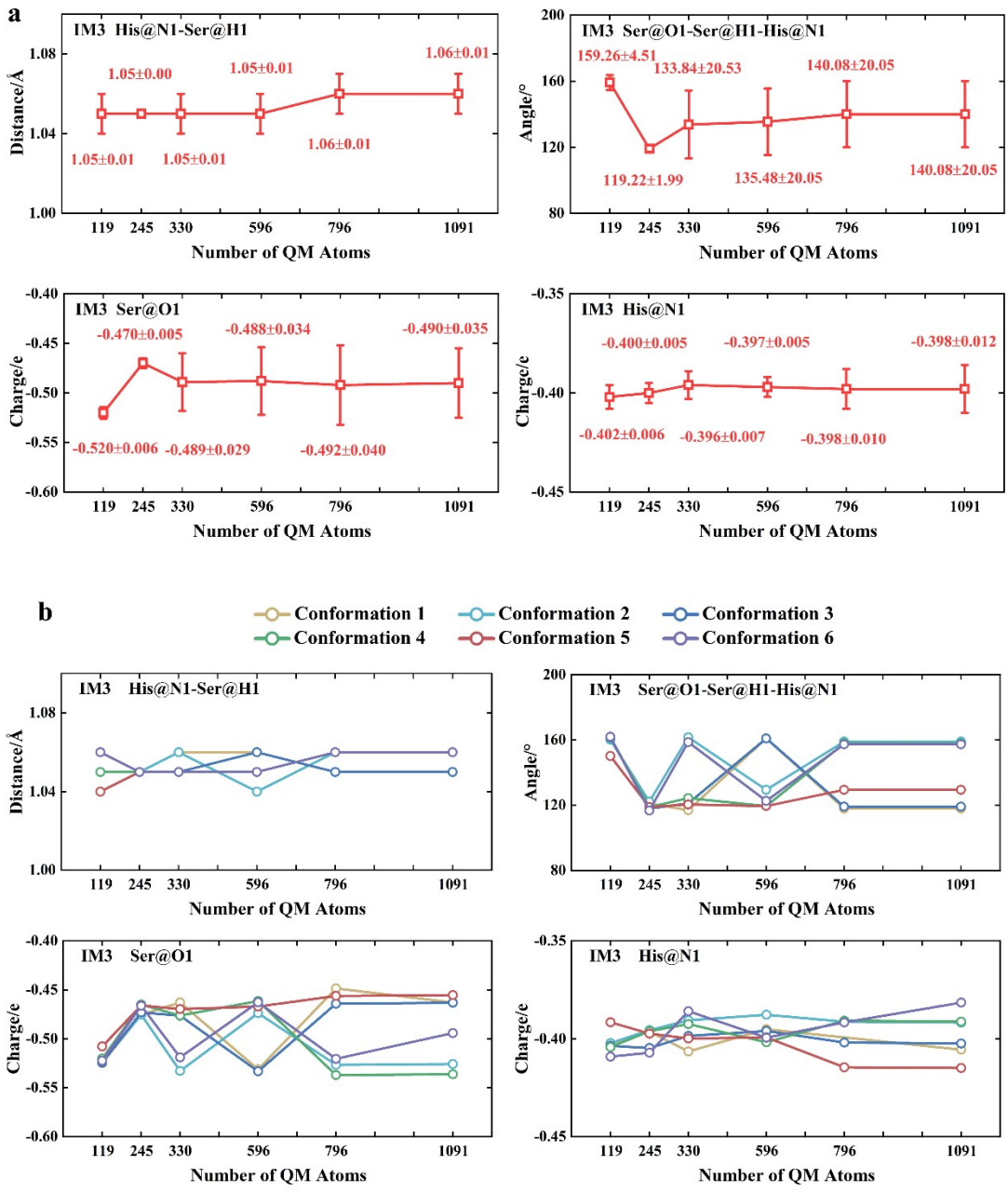


Fig. S7. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in IM3.

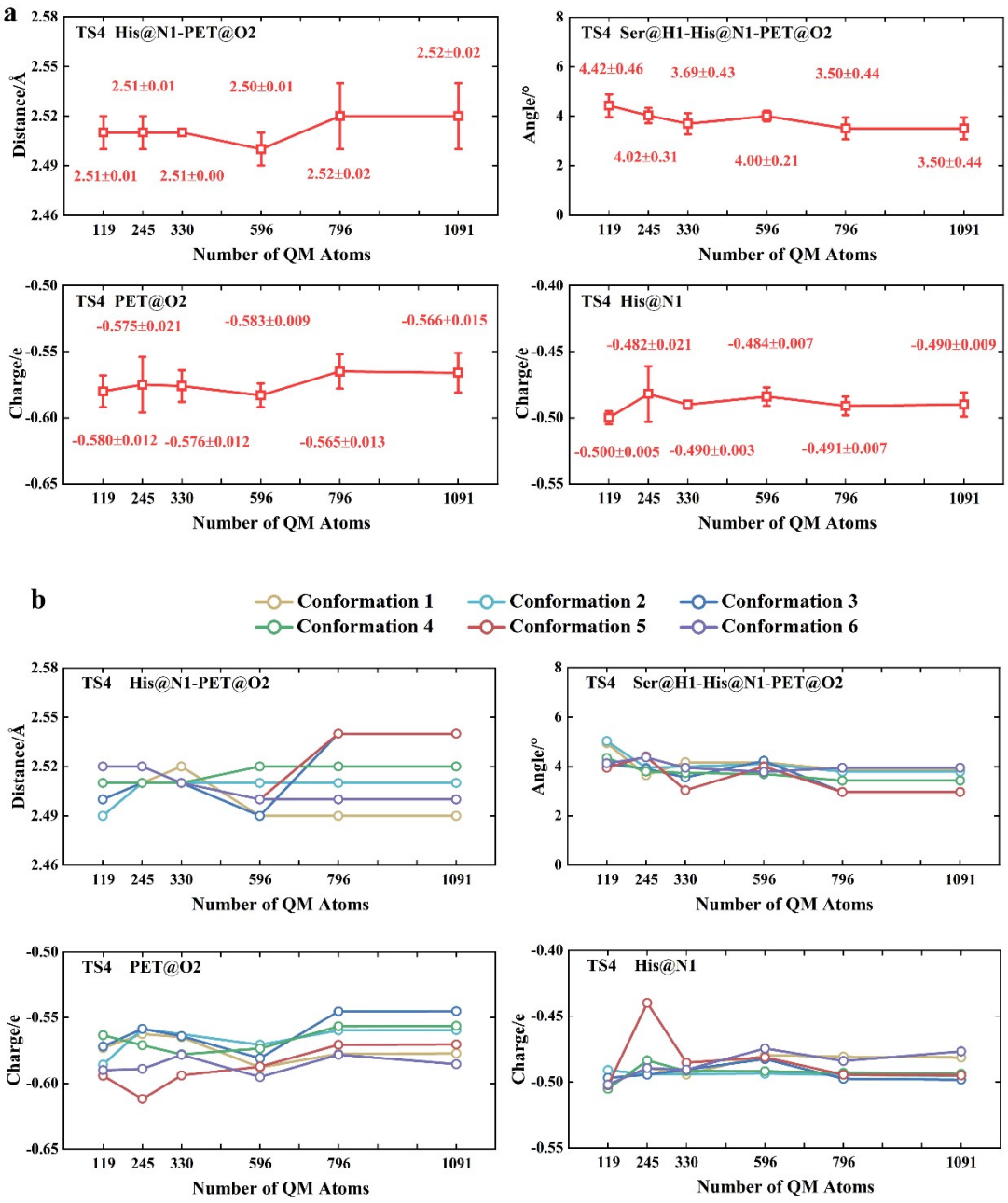


Fig. S8. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in TS4.

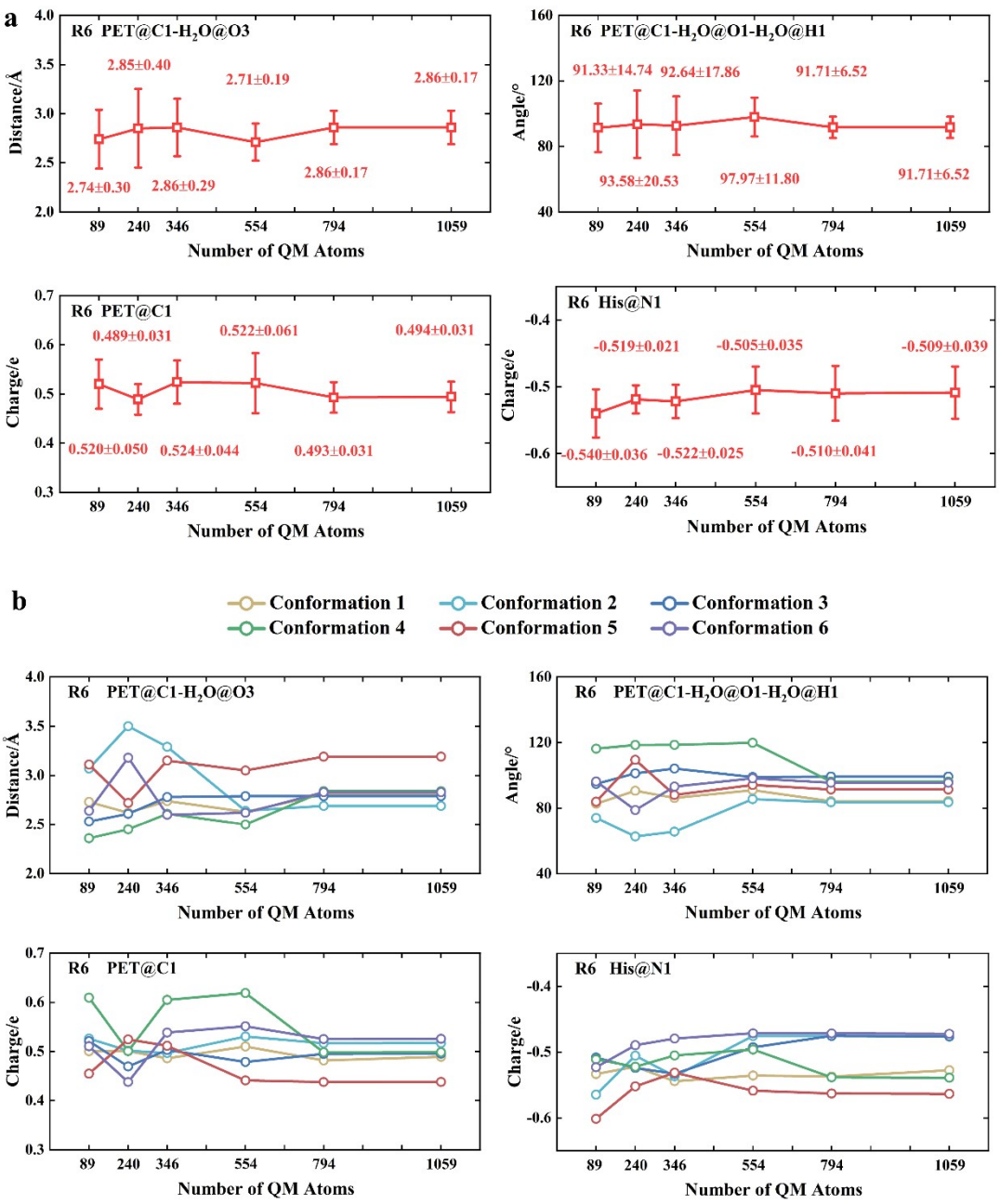


Fig. S9. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in R6.

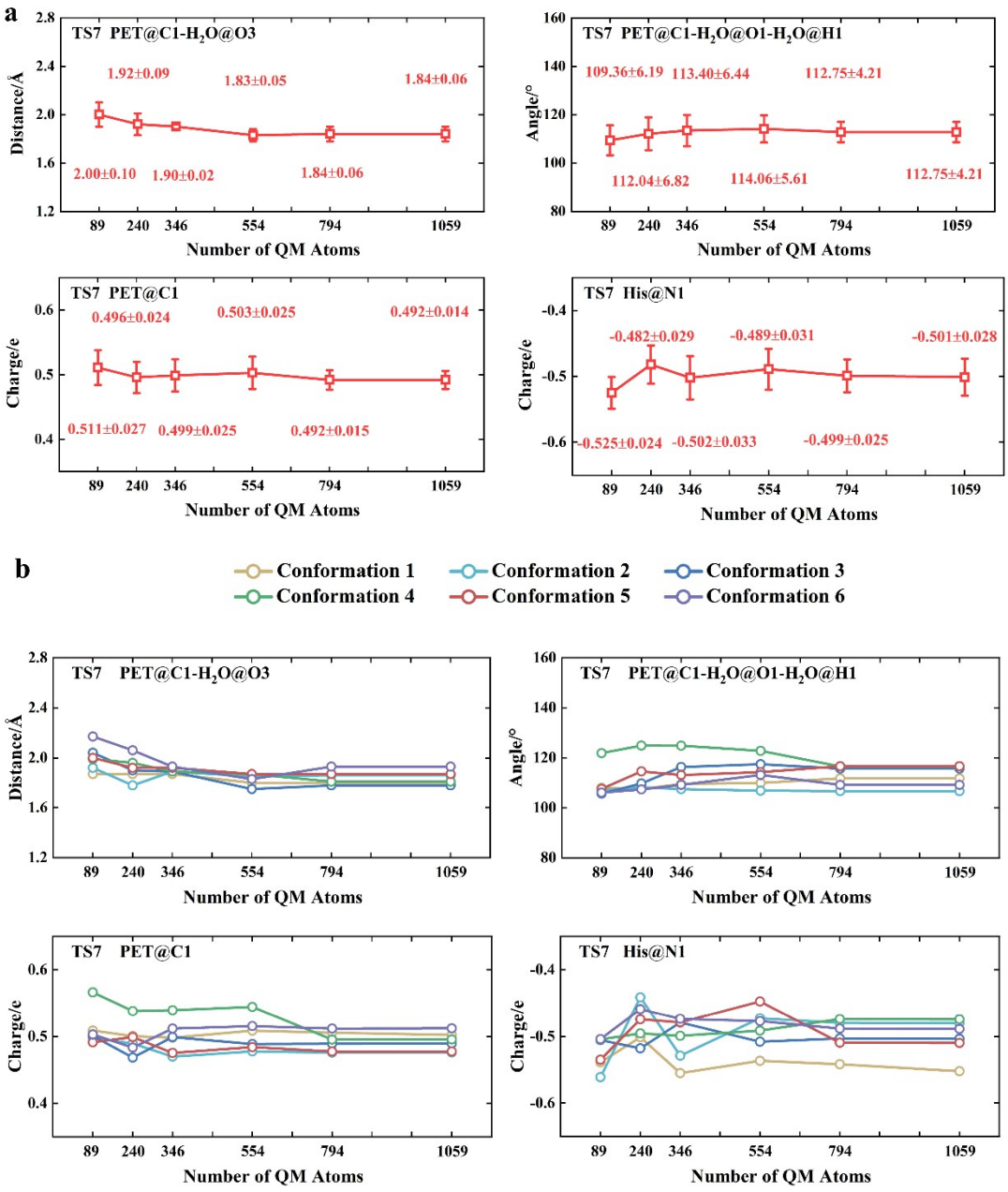


Fig. S10. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in TS7.

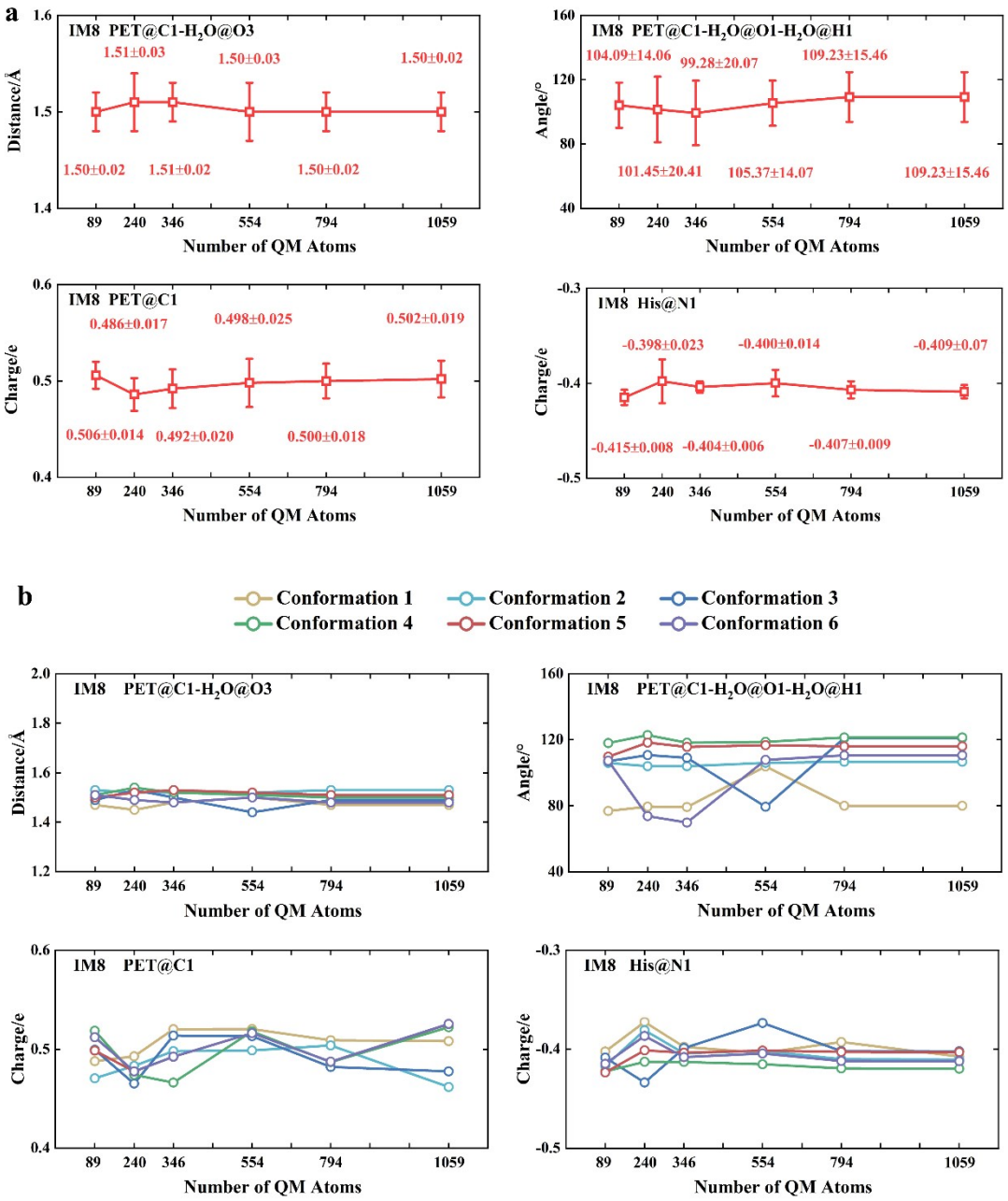


Fig. S11. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in IM8.

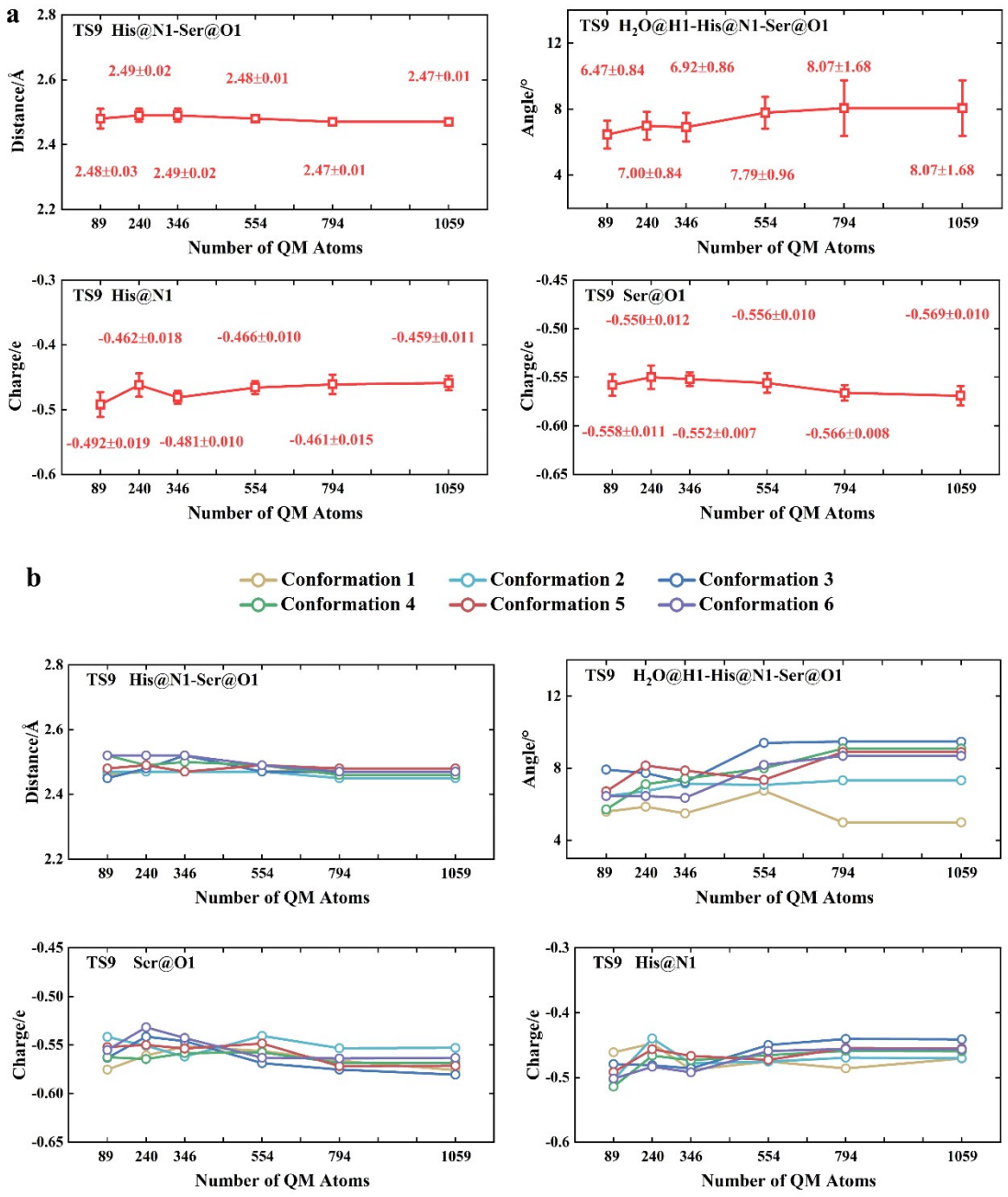


Fig. S12. Variation of key structural and charge parameters for average values (a) and six independent enzyme conformations (b) with respect to QM region in TS9.

Table S1. The fluctuation range of key structures in R1 (100 QM region atoms) calculated based on the six enzyme conformations, which is optimized by BP86/6-31G(d) and M06-2X/6-31G(d), respectively. The unit is Å.

	PET@C1-Ser@O1	His@N1-Ser@H1	His@H4-Asp@O5	PET@O4-Tyr@H6	PET@O4-Met@H7
BP86/6-31G(d)	3.30±0.09	1.57±0.02	1.69±0.01	3.33±0.40	1.86±0.04
M06-2X/6-31G(d)	3.27±0.12	1.60±0.03	1.69±0.01	3.40±0.14	1.89±0.03

Table S2. Summary of charges, QM atoms and residues in each of the studied systems.

System	Charge	QM atoms	Residues	System	Charge	QM atoms	Residues
1	0	119	GLY94, TYR95, ARG124, SER165, MET166, ASP210, HSD242, PET	1	-1	89	GLY94, TYR95, HSD164, SER165, MET166, GLY167, ASP210, HSD242, MHET, TIP3
2	-1	245	+THR96, ALA97, ASP98, ALA209, THR211, VAI212, ALA213, SER241, TIP3, TIP3, TIP3	2	-1	240	+THR188, TRP190, ALA207, THR211, VAL212, ALA213, PRO214, VAL215, TIP3, TIP3, TIP3, TIP3
3	-1	330	+SER101, HSD164, PRO214, VAL215, ILE243, ASN246	3	0	346	+PRO93, THR96, ALA97, ARG131, GLY168, GLY169, PRO189, SER241, ILE243
4	0	596	+SER92, PRO93, ARG131, GLY163, MET166, GLY167, GLY168, GLY169, GLY170, VAL185, PRO186, LEU187, THR188, PRO189, TRP190, HSD191, THR192, HSD218, PHE222, ALA244, PRO245, ASN246	4	-1	554	+SER92, SER101, LEU102, ALA244, PRO245, ASN246, ASP98, GLY163, GLY170, VAL185, PRO186, THR187, HSD191, THR192, ALA209, HSD218, PHE222
5	-2	796	+MET91, ALA99, SER100, ILE119, ASN120, THR121, SER123, PHE125, ASP126, GLY127, PRO128, ASP129, SER130, THR171	5	-2	764	+MET91, ALA99, SER100, ILE119, ASN120, THR121, SER123, ARG124, PHE125, ASP126, GLY127, PRO128, ASP129, SER130, THR171
6	-1	1091	+ALA90, VAL118, ASN122, ALA132, SER133, VAL161, ALA162, LEU172, ARG173, ILE174, ASP193, LYS194, THR195, PHE196, GLN217, ALA219	6	-1	1059	+ALA90, VAL118, ASN122, ALA132, SER133, VAL161, ALA162, LEU172, ARG173, ILE174, ASP193, LYS194, THR195, PHE196, GLN217, ALA219

Table S3. Comparison of standard deviations between single conformation and six conformations with respect to the QM region. Feature 1~6 are PET@O4-Met@H7, Ser@O1-Ser@H1-His@N1, PET@C1-Ser@O1-Ser@H1-His@N1, PET@C1-H_2O@O3, PET@C1-H_2O@O3-H_2O@H3, PET@C1-H_2O@O3-H_2O@H3-His@N1, respectively.

	R1				R6		
	Feature 1	**Feature 2**	**Feature 3**		**Feature 4**	**Feature 5**	**Feature 6**
Conformation 1	0.4	2.3	4.0	**Conformation 1**	0.1	3.5	12.6
Conformation 2	0.5	3.9	4.7	**Conformation 2**	0.4	9.9	8.2
Conformation 3	0.5	3.8	4.7	**Conformation 3**	0.1	3.1	19.2
Conformation 4	0.5	2.1	3.4	**Conformation 4**	0.2	11.5	12.4
Conformation 5	0.5	2.6	9.3	**Conformation 5**	0.2	8.7	24.9
Conformation 6	0.4	2.7	10.0	**Conformation 6**	0.2	7.1	13.4
Arithmetical	0.2	2.6	5.6	Arithmetical	0.1	2.5	9.3

Reference

1. S. Sumner, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput.*, 2013, **9**, 4205-4214.
2. A. Fouda and U. Ryde, *J. Chem. Theory Comput.*, 2016, **12**, 5667-5679.
3. L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput.*, 2013, **9**, 640-649.
4. M. Zheng, Y. Li, W. Dong, S. Feng, Q. Zhang and W. Wang, *J. Hazard. Mater.*, 2022, **423**, 127017.