

Supplementary information

Electric Fields and Fast Protein Dynamics in Enzymes

Ioanna Zoi, Dimitri Antoniou, and Steven D. Schwartz*

*Department of Chemistry and Biochemistry, University of Arizona, 1306 East University Blvd.,
Tucson, AZ 85721*

E-mail: sschwartz@email.arizona.edu

Phone: (520) 621-6363

All simulations were performed using the CHARMM^{1,2} molecular dynamics package. The starting point for the simulations was crystal structure 1OH0 from the Protein Data Bank. To build our substrate, we modified the 6,8-didehydroestrone (equilenin) molecule to 5-androstene-3,7-dione in Accelrys Discovery Studio and we obtained parameters for the CHARMM force field from www.paramchem.org. Asp103 was protonated (cite) and all crystallographic waters were retained. The system was partitioned into quantum and molecular mechanics regions where the QM region consisted of 95 atoms including the substrate, the catalytic base Asp 40, the protonated Asp 103,³ and three tyrosines (Tyr16, Tyr32, Tyr57) as they form necessary hydrogen bonds. The QM region was modeled using the PM3 method. There are various opinion on the best way to model the QM region when one is interested in electrostatic effects^{4,5} but the agreement we obtained between experimental and calculated values of the electric field validates our choice for the system we examined. The remaining atoms of the protein, ions, and solvent water molecules were in the MM region and modeled using the CHARMM27 force field. The generalized hybrid orbital (GHO) method⁶ was used to couple the two regions through the CB atoms of Tyr16/Tyr32/Tyr57 and Asp40/Asp103.

The protein was solvated in a sphere of waters of 60 Å radius, using explicit TIP3 water molecules and then the system was neutralized by placing 13 potassium ions. Then the protein was minimized using 50 steps of steepest descent followed by three stages of adopted basis Newton-Raphson minimization (ABNR), 2000 steps each. In the first ABNR stage we applied harmonic constraints to the protein backbone that were released until we reached the third stage of ABNR minimization where we switched the QM/MM on. The heating was conducted slowly from 0 to 300 K for 45 ps and this temperature was held for the rest of the simulations. Any harmonic constraints applied to the system were gradually released. After the heating phase, the system was dynamically equilibrated with velocities assigned from a Gaussian distribution every 100 steps at 300 K for 150 ps. The final structure was used as a starting configuration for the TPS study. TPS was used to harvest trajectories of the chemical event. The reaction happens in two steps involving a proton transfer from C2 to the OD2 oxygen of the catalytic base, which in turn transfers the

same proton to C10. We defined the order parameters using the second step's bond-breaking and bond-forming distances so that the reactant region contained all configurations with the OD2-H bond length $< 1.35 \text{ \AA}$ and the H-C10 bond length $> 1.35 \text{ \AA}$, while the product region contained all configurations with OD2-H1 bond length $> 1.35 \text{ \AA}$ and H1-C10 bond length $< 1.35 \text{ \AA}$. The initial reactive trajectory was generated by applying a harmonic constraint with force constant $55 \text{ (kcal/mol)/\AA}^2$ on the bond-forming distance followed by propagation for 250 fs. Using this biased first reactive trajectory as a seed, we perturb the momenta of a randomly chosen slice of this reactive trajectory and propagated using the new momenta until we generate a new reactive trajectory. This process was iterated until we generated 150 reactive trajectories of 500 fs length.

Next, we performed committor analysis to locate the transition state, defined as the time slice with the property that new trajectories initiated from that slice with random momenta have probability 0.5 to reach reactants or products. The set of these isocommittor structures, is the transition state ensemble (TSE). This TSE was used as starting points for the constrained walk that leads to the identification of the reaction coordinate following these steps: first, we make a guess for the reaction coordinate and starting from a TS structure we evolve the system while restraining the degrees of freedom of the guess for the reaction coordinate, then at various time slices along this constrained trajectory we initiate trajectories assigning random initial velocities. Finally, we calculated the ratios of the trajectories of the previous step that end up either in reactants or products: if the guess for the reaction coordinate were correct, the commitment probabilities would be 0.5 for either reactants or products. Trajectories and structures for PNP and hLDH enzymes were obtained from our previous studies.

As mentioned in the text, constraining only the QM atoms of KSI during the algorithm for the identification of the reaction coordinate, leads to commitment histograms peaked at 0.5 (Figure 5 of paper), showing that protein motions are not part of the reaction coordinate in KSI.

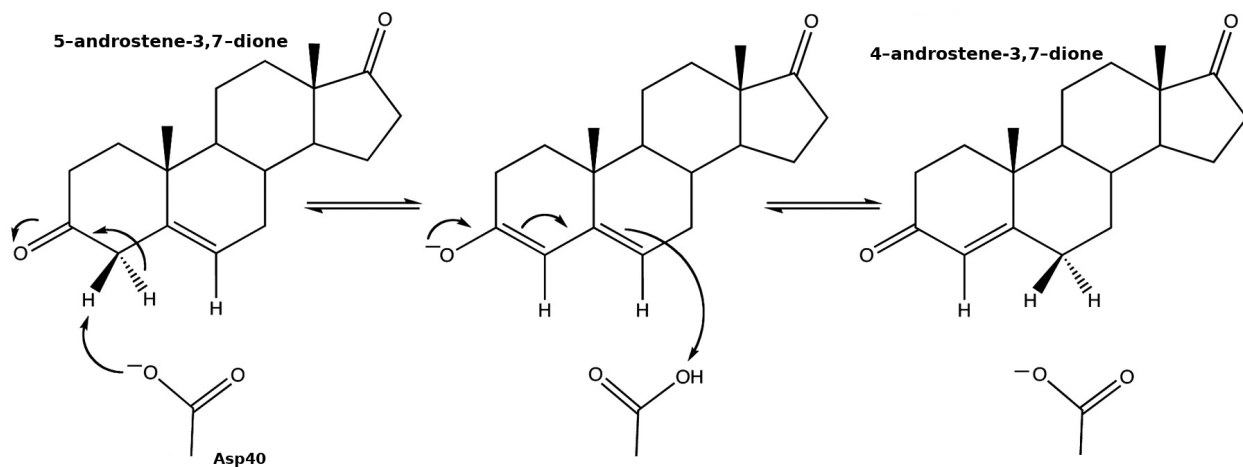


Figure S1: Mechanism of the isomerization reaction as catalysed by KSI.

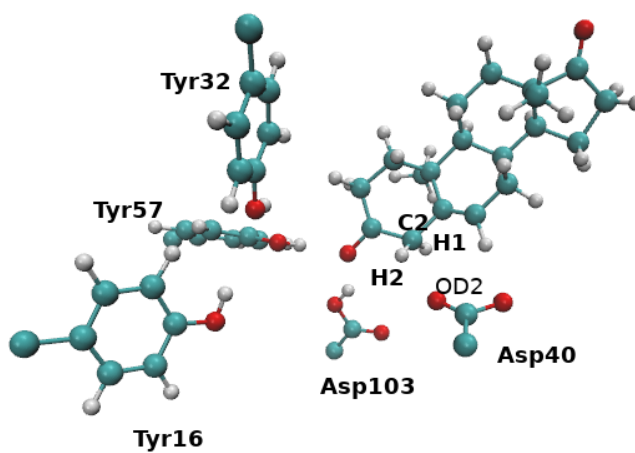


Figure S2: QM region of KSI in our simulations.

Table S1: Electric field calculation of the substrate in water in various equilibration times.

El.field on C=O dipole (MV/cm)	
10 ps	317
50 ps	379
100 ps	348
150 ps	348
200 ps	351

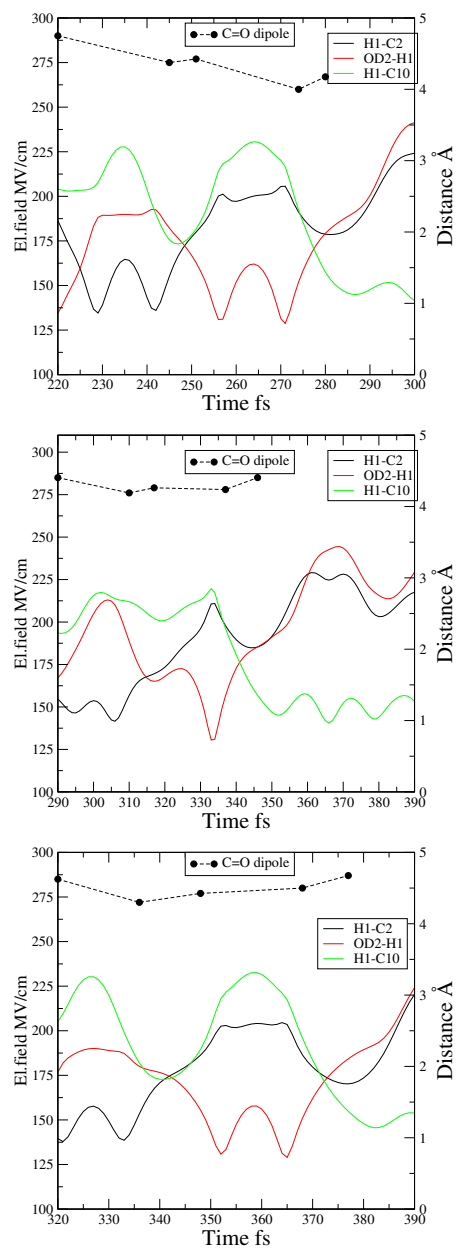


Figure S3: Electric field calculation on the second C=O dipole of the substrate for the same three representative trajectories in KSI. The five points are: in a reactant configuration, bond breaking of first step, TS of the first step, bond breaking of second step, TS of the second step.

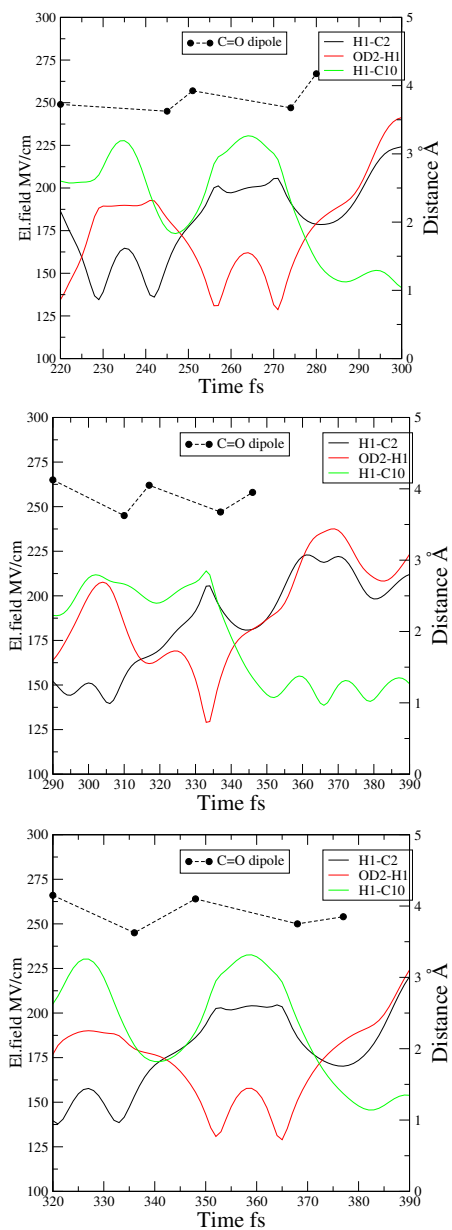


Figure S4: Electric field calculation on the C=O dipole of Asp 103 for the same three representative trajectories in KSI. The five points are: in a reactant configuration, bond breaking of first step, TS of the first step, bond breaking of second step, TS of the second step.

References

- (1) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: the biomolecular simulation program. *J. Comp. Chem.* **2009**, *30*, 1545–1614.
- (2) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (3) Thornburg, L.; Hénot, F.; Bash, D.; Hawkinson, D.; Bartel, S.; Pollack, R. Electrophilic assistance by asp-99 of 3-oxo- Δ^5 -steroid isomerase. *Biochemistry* **1998**, *37*, 10499–10506.
- (4) Bhomwick, A.; Sharma, S.; Head-Gordon, T. The importance of the scaffold for de novo enzymes: a case study with Kemp eliminase. *J. Am. Chem. Soc.* **2017**, *139*, 5793–5800.
- (5) Wang, X.; He, X.; Zhang, J. Accurate calculation of electric fields inside enzymes. *Methods in Enzymology* **2016**, *578*, 1473–1476.
- (6) Gao, J.; Amara, P.; Alhambra, C.; Field, M. A generalized hybrid orbital (GHO) approach for the treatment of link-atoms using combined QM/MM potentials. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.