



รายงานวิจัยฉบับสมบูรณ์

การจำลองทางควอนตัมแบบหลายระดับความละเอียดกับการเร่ง
ปฏิกิริยาทางชีวภาพ: เครื่องมือสำหรับทำนายสมบัติความว่องไว
และความจำเพาะต่อปฏิกิริยาของเอนไซม์

โดย รองศาสตราจารย์ ดร.จิตรยุทธ จิตอ่อนนุ่ม

เมษายน 2565

สัญญาเลขที่ RSA6280104

รายงานวิจัยฉบับสมบูรณ์

การจำลองทางควอนตัมแบบหลายระดับความละเอียดกับการเร่ง
ปฏิกิริยาทางชีวภาพ: เครื่องมือสำหรับทำนายสมบัติความว่องไว
และความจำเพาะต่อปฏิกิริยาของเอนไซม์

รองศาสตราจารย์ ดร.จิตรยุทธ์ จิตอ่อนนุ่ม
คณะวิทยาศาสตร์ มหาวิทยาลัยพะเยา

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย
และมหาวิทยาลัยพะเยา

บทคัดย่อ

รหัสโครงการ: RSA6280104

ชื่อโครงการ: การจำลองทางควอนตัมแบบหลายระดับความละเอียดกับการเร่งปฏิกิริยาทางชีวภาพ: เครื่องมือสำหรับทำนายสมบัติความว่องไวและความจำเพาะต่อปฏิกิริยาของเอนไซม์

ชื่อนักวิจัย: รองศาสตราจารย์ ดร.จิตรยุทธ์ จิตอ่อนนุ่ม มหาวิทยาลัยพะเยา

E-mail: jitrayut.018@gmail.com, jitrayut.ji@up.ac.th

ระยะเวลา: 30 เมษายน พ.ศ. 2562 ถึง 29 เมษายน พ.ศ. 2565

ปัจจุบันถึงแม้ว่าเอนไซม์ในทางอุตสาหกรรมจะนำมาใช้เพื่อการสังเคราะห์มากขึ้น แต่การพัฒนาเอนไซม์ยังต้องอาศัยกลไกการเกิดปฏิกิริยาที่แน่ชัด โดยเฉพาะข้อมูลตำแหน่งกรดอะมิโนที่มีผลต่อกลไกการสังเคราะห์ ดังนั้นในโครงการวิจัยนี้ ผู้วิจัยจึงประยุกต์ใช้ระเบียบวิธีการจำลองแบบหลายระดับความละเอียด อาทิ QM cluster QM/MM ONIOM-QM/QM' และ QM/MM molecular dynamics simulation (QM/MM MD) เพื่อศึกษาข้อมูลระดับโมเลกุลของปฏิกิริยาที่เกิดขึ้นในเอนไซม์ ทำนายสมบัติความว่องไวและความจำเพาะของเอนไซม์ การระบุสถานะโปรโตเนชันและหน้าที่ของกรดอะมิโน และการศึกษาผลของการกลายพันธุ์ของเอนไซม์ สำหรับเอนไซม์ที่รายงานในครั้งนี้ ประกอบด้วยเอนไซม์อะราบินานเนส GH43 เอนไซม์เบต้ากลูโคซิเดส GH116 เอนไซม์ไดไฮโดรพิริมิดีเนส และเอนไซม์ครีเอติเนส ผลการทดสอบและพัฒนาเทคนิค SCC-DFTB/MM ให้มีความน่าเชื่อถือด้วยวิธีการปรับปรุงค่าพลังงาน (Energy correction) พบว่าเป็นระเบียบวิธีที่ให้ผลการคำนวณที่รวดเร็ว และถูกต้อง สามารถอธิบายความยืดหยุ่นของโครงสร้างน้ำตาลวงห้าเหลี่ยม และวงหกเหลี่ยมในสารละลายได้ อธิบายพฤติกรรมการแตกออกของพันธะไกลโคสิติกได้ และเป็นวิธีการที่ให้ผลการทำนายค่าพลังงานอิสระของปฏิกิริยาได้อย่างมีประสิทธิภาพ เมื่อนำมาประยุกต์ใช้ร่วมกับเทคนิคการจำลองพลวัตแบบ umbrella sampling แต่มีข้อพึงระวังเมื่อแปลผลข้อมูลการถ่ายโอนของโปรตอน องค์ความรู้ใหม่ต่อวงการวิชาการ ได้แก่ (1) การทำนายกลไกการเร่งปฏิกิริยาย่อยสลายพันธะไกลโคสิติก การเปลี่ยนแปลงโครงสร้างน้ำตาลของซับสเตรต และบทบาทของกรดอะมิโนของเอนไซม์อะราบินานเนส GH43 (2) ได้แบบจำลองควอนตัมที่สามารถทำนายความจำเพาะทางสเตอริโอของเอนไซม์ไดไฮโดรพิริมิดีเนส (3) ค้นพบตำแหน่งกรดอะมิโนของเอนไซม์เบต้ากลูโคซิเดส GH116 ที่มีประโยชน์ในการพัฒนาสารยับยั้งต้านโรคทางพันธุกรรม และ (4) ค้นพบบทบาทของพันธะโคออร์ดิเนชันของลิแกนด์บริเวณโลหะสังกะสีตำแหน่งอัลฟาที่มีผลต่อวิถีของการเร่งปฏิกิริยาของเอนไซม์ครีเอติเนส ข้อมูลเหล่านี้สามารถนำไปศึกษาต่อทางด้านวิศวกรรมเอนไซม์เพื่อปรับปรุงสมบัติของเอนไซม์ในทางอุตสาหกรรมให้ดียิ่งขึ้น

คำสำคัญ: การจำลองแบบหลายระดับความละเอียด การเร่งปฏิกิริยาทางชีวภาพ กลไกการเร่งปฏิกิริยาของเอนไซม์ เคมีคอมพิวเตอร์ สมบัติความจำเพาะทางสเตอริโอเคมี

Abstract

Project code: RSA6280104
Title: Multiscale Quantum Modeling in Biocatalysis: A Tool for Predicting Enzymatic Reactivity and Selectivity
Investigator : Associate Professor Dr. Jitrayut Jitonnorn, University of Phayao
E-mail: jitrayut.018@gmail.com, jitrayut.ji@up.ac.th
Project period: 30 April 2019 to 29 April 2022

Although industrial enzymes are nowadays increasingly being used for synthesis purpose, the design of enzymes still requires a precise reaction mechanism, especially the molecular details on the “hot-spot” amino acids affecting the rate of enzyme catalysis. Accordingly, in this research project, we adopted several multiscale quantum modeling methods, such as QM cluster, QM/MM, ONIOM-QM/QM', and QM/MM molecular dynamics simulation (QM/MM MD), to obtain important insights into the mechanism, reactivity and selectivity, protonation state and the roles of amino acids, and the mutation effect. The studied enzymes are GH43 endo-arabinanase, GH116 beta-glucosidase, dihydropyrimidinase and creatininase. After the energy correction scheme, the accurate results of SCC-DFTB/MM technique can be obtained with less computational effort: it provides good descriptions for the structural flexibility of the five-membered and six-membered ring sugar in solution, the cleavage of glycosidic bonds and the efficient prediction of the free energy of a reaction, when combined with umbrella sampling MD. Proton affinity is the only issue when interpreting the results of this technique. The significant findings for scientific community are, for example, (i) reaction mechanism of glycosidic bond cleavage, conformational itinerary, and the roles of amino acids in GH43 endo-arabinanase, (ii) a quantum cluster model for stereospecific reaction in dihydropyrimidinase, (iii) analysis of amino acid binding residues in the protein-ligand interaction of GH116 beta-glucosidase for GBA2 inhibitor design, and (iv) role of ligand coordination at Zn α and its significance on creatininase reaction pathway. These results will be useful for protein engineering and rational design of industrial enzymes with desired properties.

Keywords: multiscale modeling, biocatalysis, enzyme mechanism, computational chemistry, stereoselectivity

Research Output (in Thai)

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ หรือผลงานตามที่ได้คาดไว้ในสัญญาโครงการ

1.1 บทความวิจัยระดับนานาชาติที่ได้รับการตีพิมพ์ในฐานข้อมูล SCOPUS/ISI และมีค่า **impact factor** จำนวนทั้งสิ้น 7 เรื่อง (และกำลังรอส่งตีพิมพ์อีก 2 เรื่อง)

ก. บทความหลักที่คาดไว้ในสัญญาโครงการ จำนวน 3 เรื่อง

1. Meelua, W., J. Olah and **J. Jitonnorn** (2022). "Role of water coordination at zinc binding site and its catalytic pathway of dizinc creatininase: insights from quantum cluster approach." **J. Comput. Aided Mol. Des.:** *accepted*. (Q1, IF 2020 =3.686, Cited by 0).
2. Huang, M., S. Pengthaisong, R. Charoenwattanasatien, N. Thinkumrob, **J. Jitonnorn** and J. R. Ketudat Cairns (2022). "Systematic Functional and Computational Analysis of Glucose-Binding Residues in Glycoside Hydrolase Family GH116." **Catalysts** 12(3): 343. (Q2, IF 2020 =4.146, Cited by 0).
3. Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. I. Mujika, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn** (2021). "Active site dynamics and catalytic mechanism in arabinan hydrolysis catalyzed by GH43 endo-arabinanase from QM/MM molecular dynamics simulation and potential energy surface." **J. Biomol. Struct. Dyn.:** 1-11. (Q3, IF 2020 =3.310, Cited by 3).

ข. บทความอื่นๆ ที่ต่อยอดจากงานวิจัยหลักของโครงการ จำนวน 4 เรื่อง

4. Castrosanto, M. A., K. B. Alviar and **J. Jitonnorn** (2022). "Virtual Evaluation of Insecticidal Potential of Fused Vip3Aa and Garlic Lectin against Fall Armyworm (Spodoptera frugiperda)." **Chiang Mai J. Sci** 49(4): 1-12. (Q4, IF 2020 =0.523, Cited by 0).
5. Hemmati, S. A., N. Karam Kiani, J. E. Serrão and **J. Jitonnorn** (2021). "Inhibitory Potential of a Designed Peptide Inhibitor Based on Zymogen Structure of Trypsin from Spodoptera frugiperda: In Silico Insights." *Int. J. Pept. Res. Ther.* 27(3): 1677-1687. (Q3, IF 2020 =1.931, Cited by 1).

6. **Jitonnorn, J.** and W. Meelua (2020). "DFT study of lactide ring-opening polymerizations by aluminium trialkoxides: Understanding the effects of monomer, alkoxide substituent, solvent and metal." **Chem. Phys. Lett.** 750: 137482. (Q2, IF 2020 =2.328, Cited by 2).
7. Meelua, W., N. Keawkla, J. Oláh and **J. Jitonnorn** (2020). "DFT study of formation and properties of dinuclear zirconocene cations: Effects of ligand structure, solvent, and metal on the dimerization process." **J. Organomet. Chem.** 905: 121024. (Q2, IF 2020 =2.368, Cited by 3).

ค. บทความที่กำลังรอส่งตีพิมพ์ จำนวน 2 เรื่อง

1. Meelua, W., T. Wanjai, N. Thinkumrob, and **J. Jitonnorn** (2022). "QM/MM Modeling of Glycosylation Reaction in GH32 Cell-Wall Invertase, A Key Enzyme in Sucrose Metabolism in Plant." **Mol. Simul.**: *submitted*. (Q2, IF 2020 =2.178).
2. Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn**, "Reaction mechanism and stereospecificity of catalytic reaction in DHP from *Saccharomyces kluyveri* and *Sinorhizobium meliloti* CECT4114: A systematic quantum chemical investigations", *to be submitted*.

2. การนำผลงานวิจัยไปใช้ประโยชน์

- เชิงวิชาการ (มีการพัฒนาการเรียนการสอน/สร้างนักวิจัยใหม่)

ผลการดำเนินงานโครงการ “การจำลองทางควอนตัมแบบหลายระดับความละเอียดกับการเร่งปฏิกิริยาทางชีวภาพ: เครื่องมือสำหรับทำนายสมบัติความว่องไวและความจำเพาะต่อปฏิกิริยาของเอนไซม์” ในช่วงระยะเวลา 3 ปี นั้น สามารถตอบโจทย์งานวิจัยแนวหน้าของประเทศ (Frontier Research) ผ่านการสร้างองค์ความรู้ใหม่และการตีพิมพ์ โดยสามารถตีพิมพ์บทความวิจัยในระดับนานาชาติที่มีค่า impact factor ได้ทั้งสิ้น จำนวน 7 เรื่อง (โดยมีค่าการอ้างอิงรวม (Total citations) เท่ากับ 9, 2020-2022, ณ 23 มีนาคม 2565) และยังอยู่ระหว่างรอส่งเพื่อตีพิมพ์อีก 2 เรื่อง นอกจากนี้ระเบียบวิธีการวิจัยที่ได้ศึกษาทำให้เกิดการนำไปต่อยอดเพื่อตอบโจทย์ปัญหาอื่น ๆ นอกเหนือจากโจทย์วิจัยหลักของโครงการได้อย่างมีนัย โดยหากสรุปการนำผลงานวิจัยไปใช้ประโยชน์ สามารถแบ่งผลลัพธ์ของโครงการวิจัยในด้านต่าง ๆ ดังต่อไปนี้

ในด้านองค์ความรู้ โครงการวิจัยนี้ทำให้เกิดการสร้างองค์ความรู้ใหม่ต่อวงการวิชาการ สำหรับการค้นพบสำคัญภายใต้โครงการวิจัย มีดังต่อไปนี้

☑ ค้นพบกลไกการเร่งปฏิกิริยาย่อยสลายพันธะไกลโคสิดิก การเปลี่ยนแปลงโครงสร้างน้ำตาลของซัสเตรตเพื่อประโยชน์ในการค้นหาสารยับยั้งในอนาคต รวมถึงการค้นพบตำแหน่งของกรดอะมิโนที่สามารถนำไปศึกษาวิศวกรรมเอนไซม์เพื่อปรับปรุงสมบัติของเอนไซม์ให้ดียิ่งขึ้น นอกจากนี้ยังค้นพบวิธีการทำแผนที่การกระจายตัวของโครงสร้างน้ำตาลในระหว่างการจำลอง (ดูรายละเอียดเพิ่มเติมได้ใน Chapter I ของรายงานวิจัยฉบับสมบูรณ์)

☑ ค้นพบแบบจำลอง QM และ QM/MM ที่สามารถใช้ทำนายความจำเพาะทางสเตอริโอของเอนไซม์ไดไฮโดรพิริมิดีเนส (DHP) ได้ครั้งแรก ซึ่ง DHP เป็นเอนไซม์ในทางอุตสาหกรรมการผลิตและสังเคราะห์สารเบต้ากรดอะมิโน (Chapter II ของรายงานวิจัยฉบับสมบูรณ์)

☑ ค้นพบตำแหน่งกรดอะมิโนของเอนไซม์เบต้ากลูโคซิเดส GH116 ที่มีประโยชน์ในการพัฒนาสารยับยั้ง GBA2 ด้านโรคทางพันธุกรรม และการศึกษาวิศวกรรมเอนไซม์เพื่อการสังเคราะห์ นอกจากนี้กระบวนการวิเคราะห์ค่าพลังงานอันตรกิริยาของกรดอะมิโน (โดยใช้ ONIOM QM/QM') ยังเป็นประโยชน์ในการศึกษาครั้งต่อไป (Chapter III ของรายงานวิจัยฉบับสมบูรณ์)

☑ ค้นพบกลไกการเร่งปฏิกิริยาของเอนไซม์ครีเอทีเนส (creatininase) ที่เป็นไปได้หลายกลไก รวมถึงพฤติกรรมการปรับเปลี่ยนพันธะโคออร์ดิเนชันของลิแกนด์บริเวณโลหะสังกะสีตำแหน่งอัลฟา (Chapter IV ของรายงานวิจัยฉบับสมบูรณ์)

☑ ได้แนวทางการศึกษาและระเบียบวิธีการคำนวณทางควอนตัมเพื่อการหาโครงสร้างและพลังงานของปฏิกิริยาที่ใช้สารประกอบแทรนซิชัน (transition metal compound) เป็นตัวเร่ง (Chapter V ของรายงานวิจัยฉบับสมบูรณ์)

☑ ได้แบบจำลองที่ช่วยอธิบายอันตรกิริยาระหว่างโปรตีนกับโปรตีน และโปรตีนกับลิแกนด์ เพื่อการค้นหายาและสารยับยั้งในอนาคต (Chapter V ของรายงานวิจัยฉบับสมบูรณ์)

ในด้านการเรียนการสอน มีการนำองค์ความรู้มาจัดการเรียนการสอนในรายวิชา 242413 เคมีคำนวณเบื้องต้น และ 242411 งานวิจัยแนวใหม่ทางเคมีเชิงฟิสิกส์ และ 242311 เคมีเชิงฟิสิกส์ 1 และ 242312 เคมีเชิงฟิสิกส์ 2 ของคณะวิทยาศาสตร์ มหาวิทยาลัยพะเยา และรายวิชาโครงการวิทยาศาสตร์ ของหลักสูตรรวมว. โรงเรียนสาธิตมหาวิทยาลัยพะเยา

ในด้านพัฒนากลุ่มวิจัย เพราะ สกว. สนับสนุนทุนวิจัยอย่างต่อเนื่องทำให้สามารถพัฒนางานวิจัยให้มีลักษณะโดดเด่นเฉพาะทาง จนสามารถต่อยอดไปสู่การเสนอขอเพื่อจัดตั้งหน่วยวิจัยความเป็นเลิศด้านวิทยาการโมเลกุลและการเร่งปฏิกิริยา (Unit of Excellence in Computational Molecular Science and Catalysis) ในปีการศึกษา พ.ศ. 2565

ในด้านพัฒนากำลังคน งานวิจัยและความรู้ที่เกิดขึ้น ได้นำไปถ่ายทอดแก่นิสิตระดับปริญญาตรี สาขาเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยพะเยา จำนวน 2 คน ได้แก่ นางสาวธัญชนก วันใจ และ นางสาวเนตรชนก ถิ่นคำรพ รวมถึงนักเรียนระดับมัธยมศึกษาตอนปลายในโครงการ รวมว. โรงเรียนสาธิตมหาวิทยาลัยพะเยา มากกว่า 10 คน ซึ่งจะเติบโตเป็นนักวิจัยในอนาคตต่อไป

ในด้านการเผยแพร่และประชาสัมพันธ์โครงการวิจัยและทุนวิจัย สกว. โดยมี
การนำผลงานวิจัยไปเผยแพร่ในงานประชุมวิชาการระดับนานาชาติ งานสัมมนา
จำนวน 4 ครั้ง (ดูข้อ 3.1)

**3. อื่น ๆ (เช่น ผลงานตีพิมพ์ในวารสารวิชาการในประเทศ การเสนอผลงานในที่ประชุม
วิชาการ หนังสือ การจดสิทธิบัตร)**

3.1 การเสนอผลงาน/เข้าร่วมในที่ประชุมวิชาการ

- วิทยากรบรรยายพิเศษ "Multiscale modeling in enzyme catalysis" ในระหว่าง
วันที่ 1-30 เมษายน 2565 ณ ภาควิชาเคมีและวิทยาศาสตร์ชีวการแพทย์
มหาวิทยาลัยลินเนียส เมืองคาลมาร์ ประเทศสวีเดน
- วิทยากรบรรยายพิเศษ "Computational enzymology: methods and
applications" งานประชุมระดับนานาชาติ ANSCSE ครั้งที่ 23 จัดโดยสมาคม
วิทยาการและวิศวกรรมเชิงคอมพิวเตอร์ ระหว่างวันที่ 27-29 มิถุนายน 2562 ณ
คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่
- วิทยากรบรรยายพิเศษ "Active Site Dynamics and Mechanism of Arabinan
Hydrolysis by Inverting GH43 Arabinanases: A QM/MM Study" งานประชุม
ระดับนานาชาติ ANSCSE ครั้งที่ 24 จัดโดยสมาคมวิทยาการและวิศวกรรมเชิง
คอมพิวเตอร์ ณ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
วิทยาเขตกำแพงแสน (รูปแบบออนไลน์) ระหว่างวันที่ 28-30 เมษายน 2564
- ได้รับคัดเลือกให้เข้าร่วมการอบรมโปรแกรม CHARMM-GUI และเผยแพร่
ประชาสัมพันธ์งานวิจัยที่ได้รับการสนับสนุนจาก สกว-มพ. 62 ระหว่างวันที่ 23-25
กันยายน 2562 ณ สถาบันการศึกษาขั้นสูงแห่งประเทศเกาหลีใต้ (Korea Institute
for Advanced Study, KIAS) กรุงโซล ประเทศเกาหลี

หมายเหตุ จำนวนการเสนอผลงานในที่ประชุมอาจน้อยกว่าปกติ เนื่องจากสถานการณ์การแพร่ระบาดของ
เชื้อไวรัสโควิด 19

Research Output 2020-2022 (in English)

Total publications: 7

Total manuscripts: 2

Total citations: 9¹

Paper I:

Meelua, W., J. Olah and **J. Jitonnorn** (2022). "Role of water coordination at zinc binding site and its catalytic pathway of dizinc creatininase: insights from quantum cluster approach." **J. Comput. Aided Mol. Des.**, *accepted*. (Q1, IF 2020 =3.686, Cited by 0).

Paper II:

Huang, M., S. Pengthaisong, R. Charoenwattanasatien, N. Thinkumrob, **J. Jitonnorn** and J. R. Ketudat Cairns (2022). "Systematic Functional and Computational Analysis of Glucose-Binding Residues in Glycoside Hydrolase Family GH116." **Catalysts** 12(3): 343. (Q2, IF 2020 =4.146, Cited by 0).

Paper III:

Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. I. Mujika, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn** (2021). "Active site dynamics and catalytic mechanism in arabinan hydrolysis catalyzed by GH43 endo-arabinanase from QM/MM molecular dynamics simulation and potential energy surface." **J. Biomol. Struct. Dyn.**: 1-11. (Q3, IF 2020 =3.310, Cited by 3).

Paper IV:

Castrosanto, M. A., K. B. Alviar and **J. Jitonnorn** (2022). "Virtual Evaluation of Insecticidal Potential of Fused Vip3Aa and Garlic Lectin against Fall Armyworm (*Spodoptera frugiperda*)." **Chiang Mai J. Sci** 49(4): 1-12. (Q4, IF 2020 =0.523, Cited by 0).

Paper V:

Hemmati, S. A., N. Karam Kiani, J. E. Serrão and **J. Jitonnorn** (2021). "Inhibitory Potential of a Designed Peptide Inhibitor Based on Zymogen

¹ Accessed on 23/03/2565

Structure of Trypsin from *Spodoptera frugiperda*: In Silico Insights." **Int. J. Pept. Res. Ther.** 27(3): 1677-1687. (Q3, IF 2020 =1.931, Cited by 1).

Paper VI:

Jitonnorn, J. and W. Meelua (2020). "DFT study of lactide ring-opening polymerizations by aluminium trialkoxides: Understanding the effects of monomer, alkoxide substituent, solvent and metal." **Chem. Phys. Lett.** 750: 137482. (Q2, IF 2020 =2.328, Cited by 2).

Paper VII:

Meelua, W., N. Keawkla, J. Oláh and **J. Jitonnorn** (2020). "DFT study of formation and properties of dinuclear zirconocene cations: Effects of ligand structure, solvent, and metal on the dimerization process." **J. Organomet. Chem.** 905: 121024. (Q2, IF 2020 =2.368, Cited by 3).

Manuscript I:

Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn**, Reaction mechanism and stereospecificity of catalytic reaction in DHP from *Saccharomyces kluyveri* and *Sinorhizobium meliloti* CECT4114: A systematic quantum chemical investigations, *to be submitted*.

Manuscript II:

Meelua, W., T. Wanjai, N. Thinkumrob, and **J. Jitonnorn** (2022). "QM/MM Modeling of Glycosylation Reaction in GH32 Cell-Wall Invertase, A Key Enzyme in Sucrose Metabolism in Plant." **Mol. Simul.**: *submitted*. (Q2, IF 2020 =2.178).

Table of Content

	Pages
Chapter I: Multiscale modeling approaches for understanding reaction mechanism in carbohydrate-active enzymes.....	1
Chapter II: Quantum chemical cluster approach as a tool for understanding enzyme mechanism and selectivity in cascade biocatalysis for synthesis of beta-amino acid.....	17
Chapter III: Application of ONIOM QM/QM' method for understanding protein-ligand interaction in carbohydrate-active enzyme for drug discovery.....	50
Chapter IV: Quantum chemistry method for mechanistic and structural investigation of zinc metalloenzyme.....	62
References:	81
Appendix: Additional papers supported by Thailand Research Fund.....	91

Chapter I

Multiscale Modeling Approaches for Understanding Reaction Mechanism in Carbohydrate-Active Enzymes

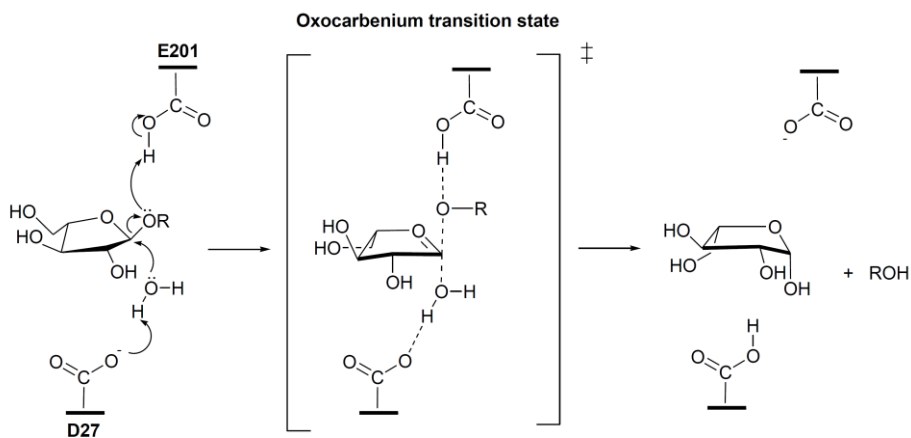
Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. I. Mujika, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn** (2021). "Active site dynamics and catalytic mechanism in arabinan hydrolysis catalyzed by GH43 endo-arabinanase from QM/MM molecular dynamics simulation and potential energy surface." **J. Biomol. Struct. Dyn.**: 1-11. (Q3, IF 2020 =3.310, Cited by 3).

Meelua, W., T. Wanjai, N. Thinkumrob, and **J. Jitonnorn** (2022). "QM/MM Modeling of Glycosylation Reaction in GH32 Cell-Wall Invertase, A Key Enzyme in Sucrose Metabolism in Plant." **Mol. Simul.**: *submitted*. (Q2, IF 2020 =2.178).

ACTIVE SITE DYNAMICS AND CATALYTIC MECHANISM IN ARABINAN HYDROLYSIS CATALYZED BY GH43 ENDO-ARABINANASE FROM QM/MM MOLECULAR DYNAMICS SIMULATION AND POTENTIAL ENERGY SURFACE

INTRODUCTION

Arabinan is a plant structural polysaccharide degraded by two enzymes; α -l-arabinofuranosidase (E.C. 3.2.1.55) and endo-1,5- α -l-arabinanase (E.C. 3.2.1.99). The latter, belonging to glycoside hydrolase (GH) family 43 (Cantarel, Coutinho et al. 2009), are of great industrial interest. Due to their ability to hydrolyze α -1,5-arabinofuranosidic bonds in arabinose-containing polysaccharides, they are applied in food technology, organic synthesis as well as in biofuel production. Furthermore, there is a current interest in the development of arabinanase inhibitors, such as iminosugar analogues of arabinobiose (Goddard-Borger, Carapito et al. 2011). Understanding the catalytic mechanism and active-site characteristics of the arabinanase enzyme requires knowledge of the protonation states of active site residues and of the hydrogen-bond (H-bond) network in atomic detail. This information at a later stage will be a the key for efficient enzyme engineering and drug discovery (Wan, Parks et al. 2015). Most glycoside hydrolysis reactions involve several proton transfer steps among the enzyme, substrate, and water at physiological pH. Thus, glycosidase catalysis cannot be fully understood without accurate mapping of hydrogen atom positions in these biological catalysts. Furthermore, direct information on the location of hydrogen atoms is difficult to obtain by X-ray crystallography, but biomolecular simulation can yield insight into the atomic details of the species participating in the catalytic cycle of enzymes, e.g. of the Michaelis complex, an initial stage of the glycoside hydrolysis reaction catalyzed by GH43 arabinanase enzyme.



Scheme 1. Proposed reaction mechanism for ABN

Arabinanases (ABNs, EC 3.2.1.-) comprise *endo*- and *exo- α* -L-arabinanases, which differ in the *endo*- and *exo* modes of action towards arabinan or arabinooligosaccharides, depending on both the overall structures of the catalytic clefts and the structure of the glycon site (Proctor, Taylor et al. 2005, Fujimoto, Ichinose et al. 2010, Park, Jang et al. 2012). The crystal structures of two *exo*-ABNs from *C. japonicus* (Nurizzo, Turkenburg et al. 2002) and *S. avermitilis* (Fujimoto, Ichinose et al. 2010), and three *endo*-ABNs from *B. subtilis* (Proctor, Taylor et al. 2005), *B. thermodenitrificans* (Yamaguchi, Tada et al. 2005) and *G. stearothermophilus* (Yamaguchi, Tada et al. 2005), are available, showing a catalytic domain consisting of a five-bladed β -propeller fold. Hydrolysis of arabinan catalyzed by ABN (Scheme 1) is suggested to occur via acid/base catalysis that requires two critical residues, E201 (a proton donor) and D27 (a nucleophilic base), giving rise to the inversion of the anomeric configuration (Alhassid, Ben-David et al. 2009). However, the roles of the residues surrounding the catalytic pocket in changing the pKa of these acid/base catalysts are poorly understood. We have recently demonstrated, using DFT-based cluster approach, that catalytic activity of ABNs relies on the interplay between three active site residues S164, E201 and Y229 (Jitonnom and Hannongbua 2018). Two different orientations of E201 with respect to the neighboring pKa modulator (S164 and Y229) were modeled, and only the system containing the H-bond network pattern S164...E201...Y229 was found to catalyze the reaction in an energetically feasible way. However, enzyme engineering of ABNs for improved activity requires more detailed understanding of the (1) dynamic behavior of the micro-environment relation to the interaction pattern of the pKa modulators (2) the conformational space of the furanose ring of the arabinan in solution, and the (3) detailed information on the key amino acid residues important to the catalysis of ABN is still missing which limits the enzyme engineering.

To address these questions we performed MD simulations, potential energy surface and free energy (PMF) calculations within the QM/MM framework. This computational study firstly uncovers the active site dynamics of ABNs towards the arabinan substrate, and then analyzes ring-puckering distribution of the furanose ring of the substrate (to address the interaction pattern of the pKa modulators and to provide a Michaelis complex for the reaction). The inverting mechanism of the enzyme as proposed in the literature is modeled using QM/MM. Finally, electrostatic contributions of amino acid residues to the reaction were analyzed to provide useful information for mutational studies.

OBJECTIVES

- (1) To address the interaction pattern of the pKa modulators of ABNs
- (2) To provide a Michaelis complex for the reaction of ABNs
- (3) To provide useful information for mutational studies

METHOD

2.1 Model preparation

All calculations were performed using the CHARMM program (Brooks, Bruccoleri et al. 1983). The initial structure for molecular simulation was obtained from the Protein Data Bank (PDB)

code 3D5Z (Alhassid, Ben-David et al. 2009), which is the X-ray structure of the mutant (E201A) arabinanase GH43 from *Geobacillus stearothermophilus* complexed with its natural substrate, arabinotriose (Figure 1). Since the wildtype and mutant structures are structurally similar (Alhassid, Ben-David et al. 2009), the wildtype was recovered by manually mutating Ala201 to Glu201 (A201E). All crystallographic water molecules were kept. Hydrogen atoms were added using the HBUILD subroutine in CHARMM and titratable residues in the enzyme were assigned their protonation status based on the pKa estimated by PROPKA 3.1 at pH 7.0 (<http://propka.ki.ku.dk>) (Olsson, Sondergaard et al. 2011). As suggested previously (Alhassid, Ben-David et al. 2009, Jitonnom and Hannongbua 2018), the general acid Glu201 was protonated and the general base Asp27 and the transition state stabilizer Asp147 was deprotonated. Other aspartate and glutamate residues were deprotonated. Histidine residues were modeled in their neutral states, with their tautomeric state assigned on the basis of the hydrogen-bonding network by WHAT-IF (<http://swift.cmbi.ru.nl>) (Vriend 1990). Here, according to the CHARMM format, HSD26, HSE41, HSE82, HSE136, HSD271, and HSE286 (HSD and HSE are protonated on ND1 and NE2, respectively) were modeled in their neutral states, while HSP5 (HSP is doubly protonated on ND1 and NE2) were fully protonated. This reactant (**React**) complex model provides a good starting point for subsequent simulations (Figure 1).

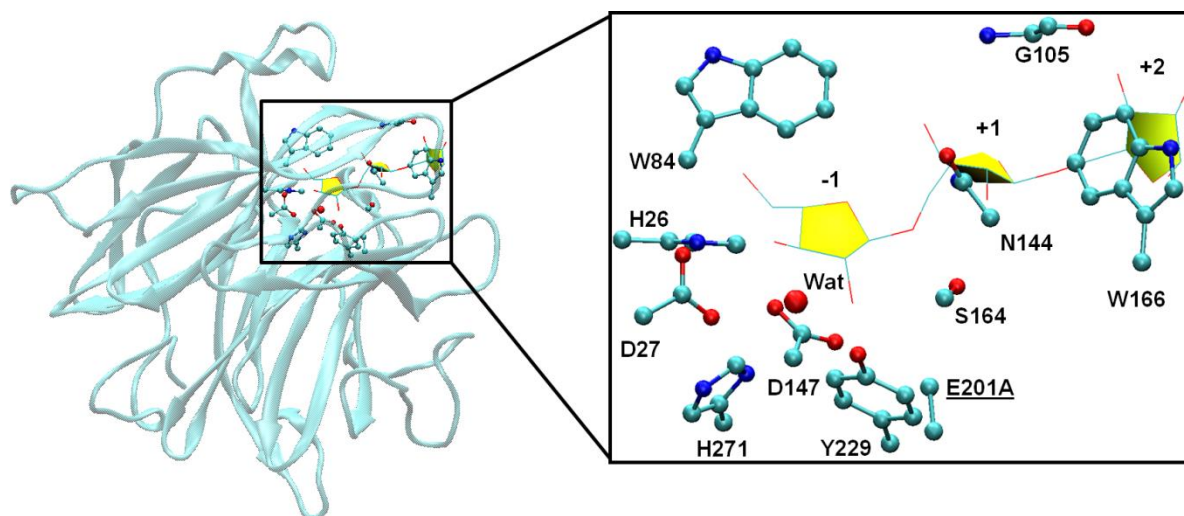


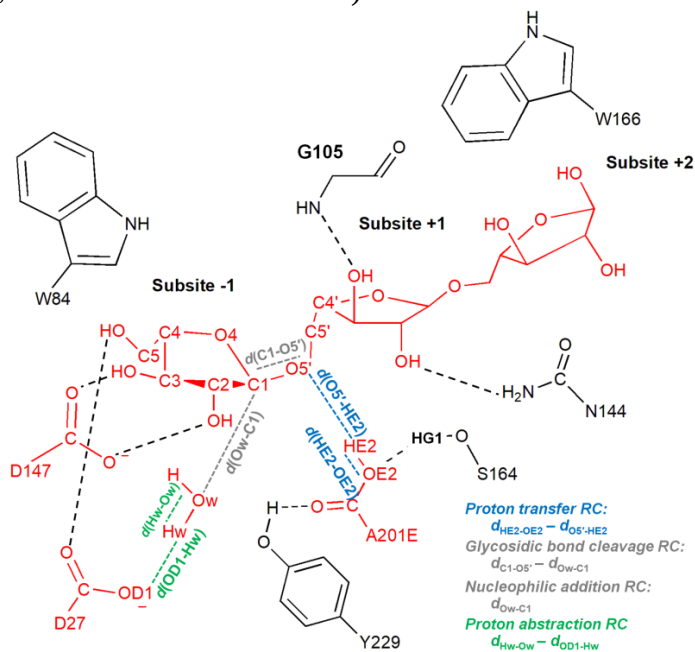
Figure 1. X-ray structure of active site in the GH43 endo-arabinanase from *Geobacillus stearothermophilus* complexed with its natural substrate (shown in yellow shade), arabinotriose (PDB code: 3D5Z).

2.2 QM/MM molecular dynamics simulation

To set the QM/MM molecular dynamics (MD) simulations up, the **React** system must be partitioned into two regions: QM and MM. Here, the QM region consists of a water molecule, the arabinotriose substrate and the side chains of three catalytic residues, Asp27, Asp147 and Glu201 (truncated at the C_β, C_β and C_γ atoms, respectively). The resulting QM region comprises 79 atoms and has a net charge of -2 , corresponding to the negative charge of Asp27 and Asp147. The QM region was treated at the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) (Eltner, Porezag et al. 1998), while the MM region was treated with the

CHARMM27 all-atom force field (MacKerell, Bashford et al. 1998) for protein, carbohydrate, and solvent. The link atom approach (Field, Bash et al. 1990) was used to couple the QM and MM regions, where the H-link atoms were placed between C_α and C_β of the Asp27 and Asp147 side chains and between C_β and C_γ of Glu201 side chain. Thus, the QM/MM MD simulations were performed at SCC-DFTB/CHARMM27 level. The SCC-DFTB method has been extensively tested and applied to investigate structural dynamics and mechanism of several GHs (Barnett and Naidoo 2010, Jitonnorn, Limb et al. 2014, Jitonnorn, Ketudat-Cairns et al. 2018).

After QM/MM partition, a 25 Å radius sphere of pre-equilibrated TIP3P waters (Jorgensen, Chandrasekhar et al. 1983) was placed on the center of the reactant complex, i.e., the anomeric carbon atom (C1) of the substrate (Scheme 1). A spherical deformable boundary potential (Brooks III and Karplus 1983) with a 25 Å radius was used to prevent the water from diffusing away from the system. All atoms outside the 25 Å sphere centered on C1 were deleted, while protein heavy atoms in the buffer zone (21–25 Å) were subject to Langevin dynamics with positional restraints using force constants scaled to increase from the inside to the outside of the buffer. All atoms within a 21 Å sphere of the reaction zone were treated by Newtonian dynamics with no positional restraints. The reactant complex was thermalized in the NVT ensemble at 310 K with 2 ns of stochastic boundary QM/MM MD simulation. An integration time-step of 1 fs was used, and all bonds involving hydrogen atoms were constrained using the SHAKE algorithm (Ryckaert, Ciccotti et al. 1977). The overall protocol has been described in more detail in previous publications (Jitonnorn, Lee et al. 2011, Jitonnorn, Limb et al. 2014, Jitonnorn, Mujika et al. 2017, Jitonnorn, Ketudat-Cairns et al. 2018).



Scheme 2. The QM region (shown in red) used in the QM/MM study of the ABN reaction. Definition of the reaction coordinates (RC) for the proton transfer, glycosidic bond cleavage, nucleophilic addition, and proton abstraction are also indicated as well as the atomic definitions used in the main text.

2.3 Potential energy surface calculations

The starting points for the two-dimensional QM/MM potential energy surface (2D-PES) calculations were taken from the final snapshot of the QM/MM/MD simulation. Prior to the 2D-PES calculations, the snapshots were subject to QM/MM energy minimization. The QM/MM calculations were carried out using the SCC-DFTB/CHARMM27 adiabatic mapping approach (Cui, Elstner et al. 2001, Woodcock, Hodošček et al. 2007, Jitonnorn, Lee et al. 2011), were performed to explore the reaction mechanism of ABN in Scheme 1. Reaction coordinates (RC) were defined by linear combinations of interatomic distances (see Scheme 2 for details). For the first reaction step, the RCs are defined as $d_{\text{C1-O5}}-d_{\text{Ow-C1}}$ for glycosidic bond cleavage and $d_{\text{HE2-OE2}}-d_{\text{O5'-HE2}}$ for proton transfer from E201 to O5' atom of the scissile linkage. In the second reaction step, the RCs are defined as $d_{\text{Ow-C1}}$ for nucleophilic addition performed by the catalytic water and $d_{\text{Ow-Hw}}-d_{\text{Hw-OD1}}$ for proton abstraction by the general base D27. These RC values were gradually increased (or decreased) by 0.1 Å for each reaction step using a force constant of 5000 kcal/mol·Å² to drive the reaction. Energy minimizations at each RC value were performed to a gradient tolerance of 0.01 kcal/mol·Å. The final energies were computed by a single-point calculation, removing the energy contributions due to the coordinate restraints. High-level energy corrections (at the B3LYP/6-31+G(d) DFT level) to the SCC-DFTB/CHARMM27 PES were also performed to obtain more accurate reaction energies and barriers. The corrected PESs were obtained by subtracting the SCC-DFTB energy of the isolated QM region from the total QM/MM energy, and adding the B3LYP energy. This energy correction method has been previously applied successfully to other enzyme-catalyzed reactions (Hermann, Hensen et al. 2005, Jitonnorn, Limb et al. 2014, Lence, van der Kamp et al. 2018).

2.4 Potential of mean force calculations

The QM/MM minimized geometries along the minimum energy path on the PES were used as putative RCs for the free-energy calculations or potential of mean force (PMF). This type of simulation allows more efficient conformational sampling of the reaction pathways, and accounts for thermal fluctuations of the solvent and enzyme environment (Lu, Fang et al. 2016). In the PMF calculations, the reaction free energies were computed on each RC value with the QM/MM umbrella sampling MD simulations, which consist of a series of “window” simulations that carried out with a fixed force constant of 200 kcal/mol·Å² as biasing potential. Other variables of the PMF simulations were the same as used for the QM/MM MD simulations described above. For each window-biased simulation, 60 ps simulation was first performed to equilibrate the system, followed by another 40 ps for sampling dynamics. The free energy profiles were plotted by combining the statistical results from all of the biased simulations for each RC by the Weighted Histogram Analysis Method (WHAM) (Kumar, Rosenberg et al. 1992).

2.5 Point-charge deletion analysis

To test the contribution of side-chains surrounding the active site in enhancing the catalytic activity of ABN, electrostatic interaction energies of the 38 residues were investigated at the SCC-DFTB/CHARMM27 level using point charge deletion analysis (also called differential transition state analysis (Jitonnorn, Lee et al. 2011, Grácz, Szimler et al. 2016, Jitonnorn, Ketudat-Cairns et al. 2018, Zhang, Shi et al. 2018, Zhu, Tang et al. 2018, Song, Yue et al. 2020). The electrostatic contribution of amino acid *i* can be described as:

$$\Delta\Delta E_{\text{React} \rightarrow \text{X}} = \Delta E^{i-0}(\text{X}) - \Delta E^{i-0}(\text{React}) \quad \text{when X = TS, Prod}$$

$$\Delta E^{i-0} = \Delta E^i - \Delta E^0$$

Where, ΔE^{i-0} is the change of the barrier, ΔE^i is the potential barrier by eliminate the charge on the individual residue, and ΔE^0 is the original value of the potential barrier. The electrostatic influence of residues in the QM region (Asp27, Glu201, Asp147) was excluded from this analysis. The geometries of all the reactant and transition states were kept unchanged during the process. Positive and negative values of ΔE indicate that the i th residue stabilizes or destabilizes the reacting system, respectively.

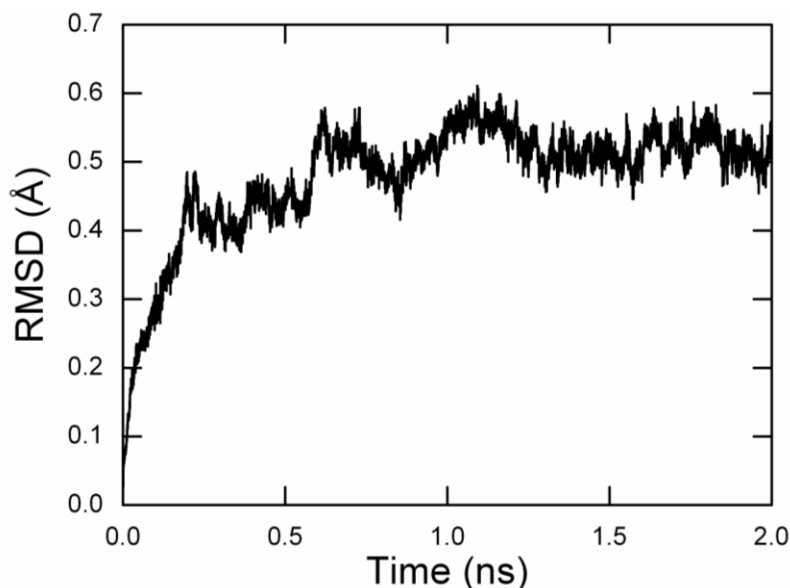


Figure 2. Time evolution of the RMSD of protein heavy atoms with respect to the initial structure

RESULTS AND DISCUSSION

3.1 QM/MM MD simulation

(a) Michaelis complex: In order to obtain the thermally equilibrated Michaelis complex structure for modeling ABN-catalyzed hydrolysis, we have conducted a QM/MM MD simulation of the reactant complex. The root mean square deviations (RMSDs) for the protein heavy atoms of the simulation indicated that the Michaelis complex is very stable and reached equilibrium after ~0.5 ns of simulation (Figure 2), with averaged heavy-atom RMSDs of 0.52 ± 0.03 Å. Throughout the simulations, arabinan is tightly bound within the enzyme's active site (occupying subsites -1 to +2) via electrostatic and hydrophobic interactions. For example, the carboxylate oxygen atoms (OD1 and OD2) of Asp147 form strong, stable hydrogen bonds with two hydroxyl groups, C2-OH and C3-OH, of the furanosyl moiety at subsite -1, as can be seen from the distances C2(-1)@OH...OD2@D147 and C3(-1)@OH...OD2@D147 (Figure S1). The sugar moiety at

subsite +1 is stabilized by several hydrogen bonds formed by the OH group of Ser164 (2-OH...HO@Ser164 = 2.04 ± 0.39 Å), the NH₂ group of Asn144 (2-OH...NH₂@N144 = 2.22 ± 0.22 Å), and the NH backbone atom of Gly105 (3-OH...NH@G105 = 2.24 ± 0.32 Å). The remaining subsite +2 is loosely bound as it is outside the binding cleft allowing it to be more exposed to the solvent (data not shown). The general acid Glu201 forms a hydrogen bond with the glycosidic bond oxygen O5' (O5'-HE2@E201 = 1.88 ± 0.14 Å), whereas the nucleophilic water is oriented towards the anomeric carbon (C1-O_w = 3.96 ± 0.18 Å) with the help of Asp27 via hydrogen bonding (Hw-OD1@D27 = 1.98 ± 0.27 Å, Figure S2). As pointed out by Zechel and Withers (Zechel and Withers 2000), the distance between two conserved carboxylate groups of Glu201 and Asp27 should be about 10.5 Å for an inverting glycosidase. In our simulation, this distance (between the C α atoms of D27 and E201) was measured to be 11.56 ± 0.21 Å, which is comparable to the experimental value (10.99 Å) (Alhassid, Ben-David et al. 2009). W84 and W166 also form pi-stacking interaction at subsites -1 and +2, respectively. Overall, these interactions are maintained over the course of the simulation, indicating that the Michaelis complex is thermally equilibrated and is ready for the reaction to take place.

(b) Substrate distortion along the glycosidic bond. The SCC-DFTB/MM simulation predicted a near coplanarity along the scissile glycosidic bond linkage (between subsite -1 and +1), with average torsion angle (C1-O5'-C5-C4) of $171.5 \pm 6.5^\circ$ (Figure S3). The second linkage between subsite +1 and +2 has a more distorted conformation with higher fluctuated ($161.7 \pm 16.7^\circ$) originating from the weak binding of the subsite +2 furanose ring. This calculated torsion angles differ to some extent from those observed crystallographically in other ABNs: a more distorted geometry of this bond was found at 133.8° (-1 and +1) and 178.3° (+1 and +2) in GsABN43B (Alhassid, Ben-David et al. 2009) and -156.8° (-1 and +1) and -175.1° (+1 and +2) in CjABN43A (Nurizzo, Turkenburg et al. 2002). This discrepancy for the description of glycosidic linkage geometry may reflect the stabilization of the oligosaccharide conformation by crystal packing in the X-ray structure, whereas the relaxed MD structure mimics the more dynamic solution structure, allowing a stretching of the glycosidic linkage.

(c) Dynamics of the pKa modulators (Tyr229 and Ser164) and their H-bond networks: As aforementioned, two active site residues (Y229 and S164) in BsABN43 are thought to increase the pKa of E201 (Alhassid, Ben-David et al. 2009) enabling it to act as a catalytic acid; the functional effect of its neighboring residues (Y229 and S164) as the pKa modulators on the catalytic power has recently been studied theoretically based on a static X-ray structure (Jitonnom and Hannongbua 2018). Here, we further investigated the dynamics and interaction patterns of these residues using the QM/MM MD simulation. The key snapshots representing the flipping of E201 side chain towards Y229 (active site configuration **I** $\rightarrow$ **II**) are visualized in Figure 3a and some relevant distances are plotted in Figure 3b. After 530 ps, the acid residue (E201) starts to rotate its side chain towards Y229, due to the loss of a H-bond with the HG1 atom of S164 side chain (**I** $\rightarrow$ **II**, Figure 3a). This can be seen from the decrease of the S164@HG1-OE2@E201 distance from 3.48 ± 0.24 Å in **I** to 1.94 ± 0.20 Å in **II**. Similar situation was also observed for S164@HG1-OE1@E201. In the configuration **II**, two different configurations (**IIa** and **IIb**) can be envisioned upon networks of H-bond (Wat...Y229...E201) and their dynamics, i.e., whether the OH group of Y229 makes a stable hydrogen bond with the catalytic water (**IIa**; Y229@OH-O_w@Wat = 1.97 ± 0.17 Å; H-bond cut-off = 2.5 Å) or the E201 OE1 atom (**IIb**; Y229@OH-OE1@E201 = 1.91 ± 0.20 Å; H-bond cut-off = 2.5 Å). During the

simulation, the appearance of **IIa** is more obvious than the case of **IIb**. These results support the experiment that Y229, together with S164, serve as pKa modulators of the acid E201 and help in maintaining a productive configuration for catalysis through the H-bond network.

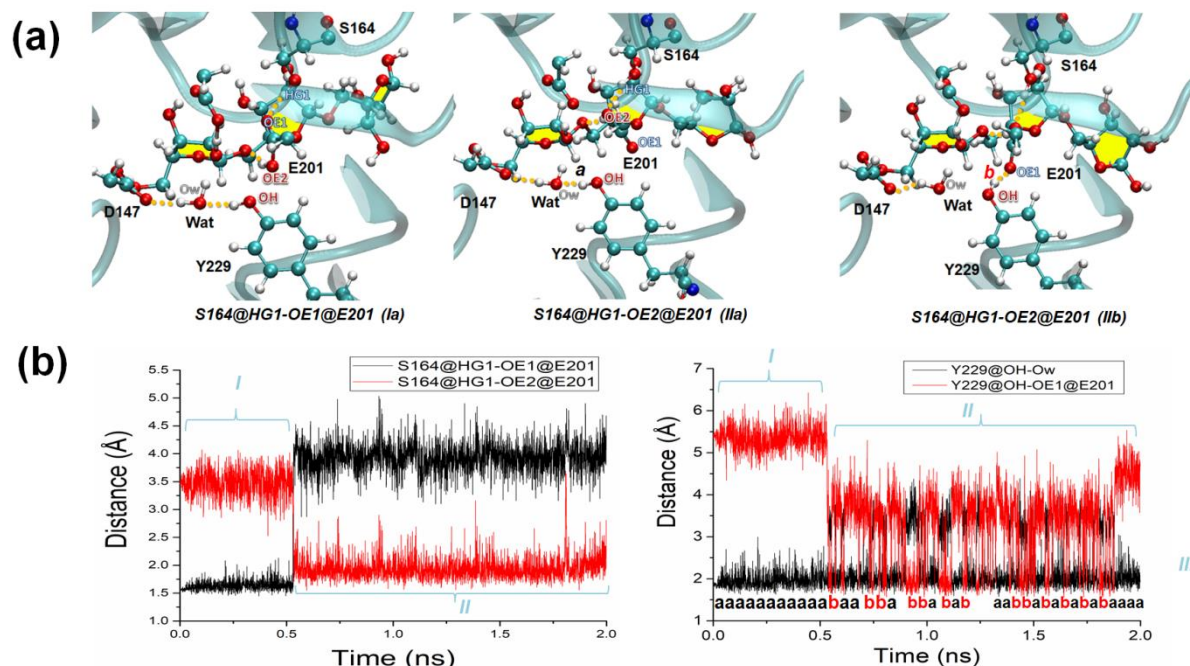


Figure 3. (a) Observed flipping of E201 sidechain towards Y229 and (b) time-evolution of the distances, S164@HG1-OE1/OE2@E201, Y229@OH-Ow and Y229@OH-OE1@E201. The latter two reflect the role of Y229 in stabilizing the catalytic water, Wat, and the E201 sidechain. The furanosyl units of the arabinotriose substrate are highlighted by yellow filling.

(d) Conformational sampling of the furanose ring. It is known that five-membered rings such as furanose can be classified into envelope (E) and twists (T). The ring pucker of these forms can be described using a triangular decomposition system that involves three angles, C3-O4-C2 (a reference plane), C3-C4-O4 (θ_0) and O4-C1-C2 (θ_1), as defined by Barnett and Naidoo (Barnett and Naidoo 2010) (see Scheme 2 for atomic labels). We then analyzed the ring pucker distribution for the furanose moiety occupying at subsites -1, +1 and +2 during the simulation of Michaelis complex, as shown in Figure 4. The plot clearly shows that two mixtures of the envelope (E) and twist (T) conformations were observed for each furanosyl subsite. This would be expected since both conformations have similar energy, and the barrier to their interconversions is rather small (Agirre 2017). For subsite -1 where the glycosidic cleavage taking place, a mixture of two mainly $^0E/^0T_1$ conformations appears along the diagonal (Q3 zone) of the origin ($\theta_1 = 0^\circ$ and $\theta_0 = 0^\circ$). Despite the majority of $^0E/^0T_1$ distribution at Q3, other minor conformers such as $^2E/E_2$ and $^2T_1/^1T_2$, as seen at Q2/Q4, are predicted as a global minimum in conformational landscape of furanose ring (Stortz and Sarotti 2019). This might be the reason why the two envelope conformations $^0E/^2E$ were found in our previous QM cluster study (Jitonnorn and Hannongbua 2018). For subsite +1, the conformational distribution appears mainly at Q1 and a minor at Q4, covering a wide variety of the $E_0/^4T_0/^4E$ conformations. The remaining subsite at +2 exhibits a wide distribution covering all possible furanose conformations.

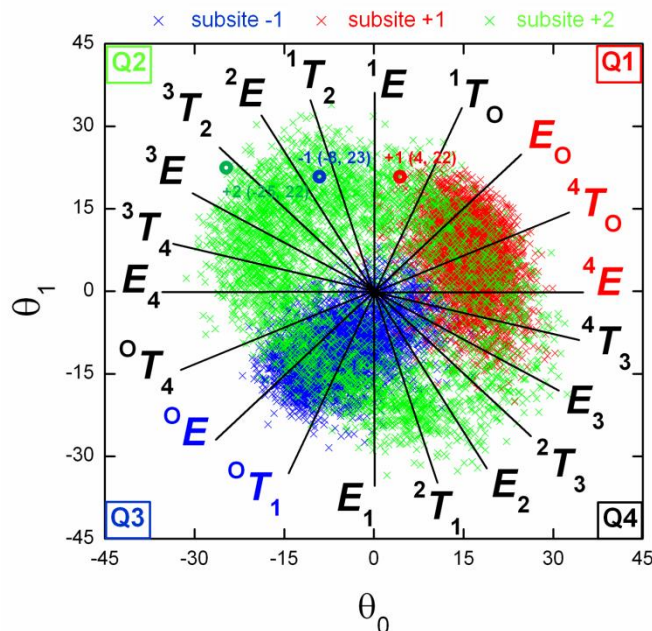


Figure 4. Conformational sampling analysis for the furanose ring at subsites -1 (blue), $+1$ (red) and $+2$ (green), as observed from the SCC-DFTB/CHARMM27 QM/MM simulation. Triangular decomposition pucker space is mapped out onto the plot for comparison. Angles obtained from the X-ray data (PDB code 3D5Z) for the furanose rings at subsites -1 , $+1$ and $+2$ were also included in the plot as shown in blue, red, green circle, respectively.

3.2 2D-PES and PMF

Starting from the snapshots of the equilibrium state of the 2-ns QM/MM MD simulation on the reaction systems, we first explored the 2D potential energy surfaces, using QM/MM adiabatic mapping, and then identified minimum energy path along the 2D-PES. Two B3LYP/6-31+G(d)//SCC-DFTB/CHARMM27 PESs for the arabinotriose hydrolysis are shown in Figure 5. Consistent with our previous QM cluster study (Jitonnom and Hannongbua 2018), a saddle point that separates the reactant complex (**React**) and the reaction product (**Prod**) was located on the first PES (at RC = ~ 0.0 Å and RC = ~ 0.7 Å, Figure 5a), suggesting a concerted process between the proton transfer and the glycosidic bond cleavage. After the TS, the minimum energy path goes approximately diagonally without any barrier into the stable product **Prod** (at RC = ~ 1.5 Å and RC = ~ 0.7 Å, Figure 5b), releasing the energy of 8 kcal/mol. The second PES also indicates that the proton abstraction by Asp27 is concerted with the nucleophilic attacking of water onto the anomer carbon C1. Details of important structures on these PESs are described below.

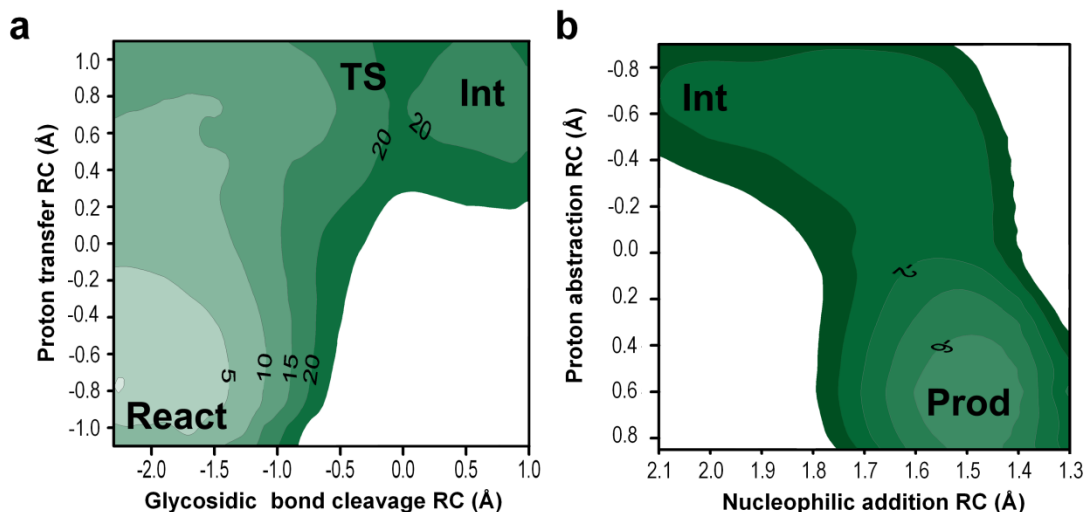


Figure 5. **a** 2D potential energy surface for the first reaction step as a function of the glycosidic bond cleavage ($d_{C1-O5'}-d_{Ow-C1}$) and the proton transfer ($d_{HE2-OE2}-d_{O5'-HE2}$). **b** 2D potential energy surface for the second reaction step as a function of the nucleophilic addition (d_{Ow-C1}) and the proton abstraction ($d_{Ow-Hw}-d_{Hw-OD1}$). Both PESs were obtained by B3LYP/6-31+G(d)//SCC-DFTB/CHARMM27 QM/MM calculations. The energies given are the total QM/MM energies in kcal/mol relative to **React**. RC, reaction coordinate.

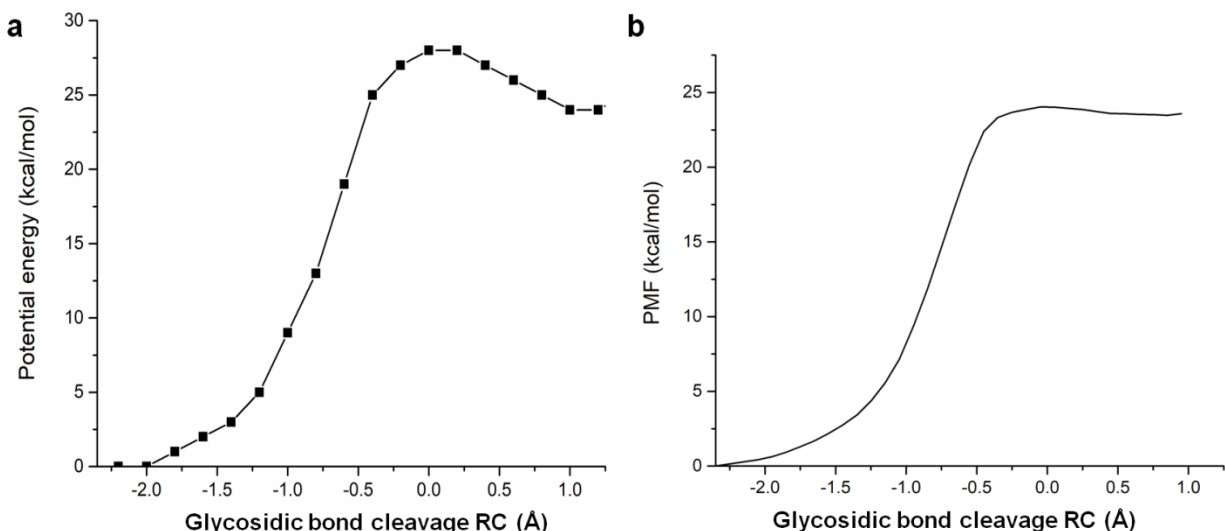


Figure 6. Potential energy profile (a) and potential of mean force (PMF) (b) for the first reaction step as a function of the glycosidic bond cleavage ($d_{C1-O5'}-d_{Ow-C1}$).

The experimental rate constant was measured to be 191.6 s^{-1} (de Sanctis, Inacio et al. 2010) which can be converted to a barrier of $\sim 14 \text{ kcal/mol}$ using classical transition-state theory. The barrier (**React**→**TS**) is calculated to be $\sim 21 \text{ kcal/mol}$, which is overestimated by $\sim 7 \text{ kcal/mol}$ compared to the experimental one. The overall reaction is exothermic by 8 kcal/mol , indicating

that the release of aglycon leaving group in the product complex is thermodynamically driven. To obtain the reaction free energy profile or PMF in aqueous solution, we additionally performed the PMF simulations using QM/MM(SCC-DFTB/CHARMM27) MD umbrella sampling simulations (Seabra, Walker et al. 2007, Jitonnorn, Limb et al. 2014, Jitonnorn, Mujika et al. 2017) and WHAM method (Kumar, Rosenberg et al. 1992). The calculated potential energy (adiabatic) map and the PMF profile as a function of the glycosidic bond cleavage ($d_{C1-O5'}-d_{C1-Ow}$) is shown in Figure 6. The energy barriers obtained from both adiabatic mapping and the PMF for the reaction are computed to be around 27 and 23 kcal/mol, respectively [at the SCC-DFTB/CHARMM27 QM/MM level], which are in-line with the 2D-PES results in Figure 5.

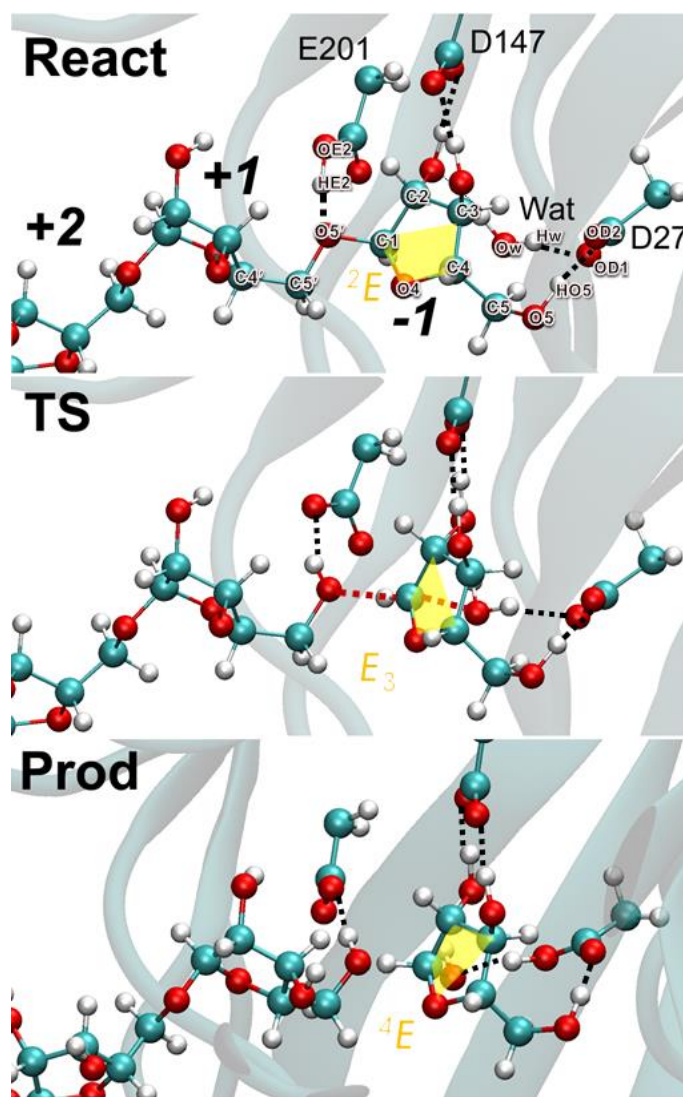


Figure 7. QM/MM structures of the reactant (**React**), transition state (**TS**) and product (**Prod**) along the minimum energy path of potential energy surface. Only QM atoms (shown in ball and stick) are included in the figure. The envelope conformation of the -1 sugar ring with its four atoms lie in a plane is also indicated in yellow.

The QM/MM minimized structures resulting from the PES calculations are shown in Figure 7 and their geometric parameters are listed in Table 1. The mechanism described here is generally similar to previous studies (Alhassid, Ben-David et al. 2009, Jitonnom and Hannongbua 2018), which consists of protonation on the glycosidic bond of substrate assisted by a proton transfer from Glu201, followed by the activation of the catalytic water to promote a single displacement attack on the anomeric carbon. Our calculations indicate that these steps occur through a transition state without any fructosyl–enzyme intermediate being formed, as in other retaining GHs. To activate the fission of the α -1,5-arabinofuranosidic bond, Glu201 must donate its proton HE2 onto the glycosidic bond oxygen (O5'). At **React**, the Glu201 sidechain is hydrogen bonding with the O5' atom (O5'–HE2@E201 = 1.80 Å) and the catalytic water is 3.78 Å distant from C1. At **TS**, the proton transfer between the E201 oxygen and the glycosidic oxygen is complete (O5'–HE2@E201 = 1.00 Å) and the glycosidic bond elongates to 2.30 Å, and the nucleophilic water molecule now approaches to the C1 atom (2.19 Å). At this stage, the TS stabilizer (D147) tightens its hydrogen bonds with the C2-OH and C3-OH groups of the fructosyl residue in subsite –1 from 1.68 and 1.71 Å at **React** to 1.51 and 1.70 Å at **TS**, respectively. This proton transfer process is well stabilized by the enzyme, as evidenced by the positive stabilization energies (Figure S4). The oxocarbenium species is observed, with a partial double character of the C1-O4 bond (shortening from 1.45 Å at **React** to 1.30 Å at **TS**). Furthermore, the C1 atom becomes sp^2 -hybridized, forcing the atoms attached to C1 and O4 into the same plane and has an E_3 conformation (van Rijssel, van Delft et al. 2014) (see umbrella sampling analysis in Figure S5). After the TS, the product is formed, where the former glycosidic oxygen O5' atom is now 3.59 Å away from the anomeric carbon C1, and the nucleophilic water molecule forms a covalent bond of 1.41 Å with C1, transferring the extra proton to the proton acceptor (D27) to complete the reaction. The C1 atom recovers sp^3 hybridization with a C1-O4 distance of 1.47 Å. During the reaction, the dihedral angle (C1'-O5-C5-C4) along the glycosidic bond linkage is slightly distorted away from its plane (177.8°) at **React** to some extent at **Prod** (162.8°). The final product has an 4E conformation in the –1 sugar, as also observed in our previous DFT study (Jitonnom and Hannongbua 2018). Thus, the substrate conformation follows the $^2E \leftrightarrow E_3 \leftrightarrow ^4E$ itinerary.

3.3 Individual residue contribution

To probe the roles of the residues surrounding the active site of ABN, electrostatic interaction energies of the 38 residues were investigated at the SCC-DFTB/CHARMM27 level. Although residues beyond the reactive center do not participate directly in the reaction, they still exert their effect via electrostatic interactions, which is considered as the basis of the catalytic effect of enzymes by Warshel and coworkers (Warshel and Levitt 1976, Warshel, Sharma et al. 2006).

The electrostatic contributions of the investigated residues on the stabilization of the QM region during the reaction were analyzed for the **React**, **TS** and **Prod** states (Figure S6 and Table S1). Nine of them show a significant contribution with $\Delta\Delta E > 1$ kcal/mol and the structures of individual residues relative to the reactive region are visualized in Figures 8 and S7. These residues are His26, Trp84, Asp87, Lys106, Phe162, Glu196, Arg223, His271 and Tyr275.

Table 1. Some relevant geometric parameters for the stationary structures along the potential energy surface calculated at B3LYP/6-31+G(d)//SCC-DFTB/CHARMM27 level

Distance (Å), angle (°)	Expt.	React	TS	Prod
Substrate				
C1–O _w	3.64	3.78	2.19	1.41
C1–O5'	1.46	1.48	2.30	3.59
C1–O4	1.45	1.45	1.30	1.47
HO2(+1)–OG@S164	2.94 ^a	1.91	1.82	1.93
∠C1(–)O5'(+)C5(+1)	113.9	111.1	113.5	102.1
∠C1(–)O5'(+)C5(+1)C4(+1)	133.8	177.8	166.6	162.8
∠C2(–)C1(–)O4(–)C4(–)	–28.3	–13.0	–7.4	–26.4
∠C2(–)C3(–)C4(–)O4(–)	13.3	–1.4	6.9	–5.0
∠C1(+1)C2(+1)C3(+1)C4(+1)	–6.0	–4.7	–3.8	0.8
∠C2(+2)C3(+2)C4(+2)O4(+2)	38.2	2.1	0.0	–2.1
General acid-base residues				
O5'–HE2@E201	–	1.80	1.00	1.01
OE2@E201–HE2@E201	–	1.00	1.79	1.73
Hw–OD1@D27	2.76 ^a	1.83	1.91	1.05
Hw–Ow	–	0.98	0.99	1.95
pK_a modulator				
OE2@E201–HG1@S164	–	1.82	1.70	1.74
OE1@E201–HH@Y229	–	1.77	1.64	1.63
TS stabilizer				
C2(–)@OH–OD2@D147	2.51 ^a	1.68	1.51	1.52
C3(–)@OH–OD2@D147	2.63 ^a	1.71	1.70	1.72
C5(–)@OH–OD2@D27	3.21 ^a	1.56	1.54	1.81

^a O–O distances and the experimental values are obtained from PDB code 3D5Z

It is shown that Tyr275, which located distant from the QM region (Figure S7), makes the largest contribution to stabilization ($\Delta\Delta E > 6.5$ kcal/mol). This result to some extent agrees with previous suggestions that distant residues can have a significant impact upon the catalytic reaction (Labas, Szabo et al. 2013). Other residues, Glu196, His26, Phe162, also show a stabilizing effect but to a lesser extent. It is found that Phe162 effectively stabilize the reaction product ($\Delta\Delta E_{\text{React} \rightarrow \text{Prod}} = \sim 7$ kcal/mol). Meanwhile, our results showed that the negative values at TS of residues Asp87, Lys106, and His271 were –0.6, –1.9, –1.0, respectively. As a result, residues Asp87, Lys106, His271 were highlighted as potential targets for future mutations to improve enzymatic efficiency.

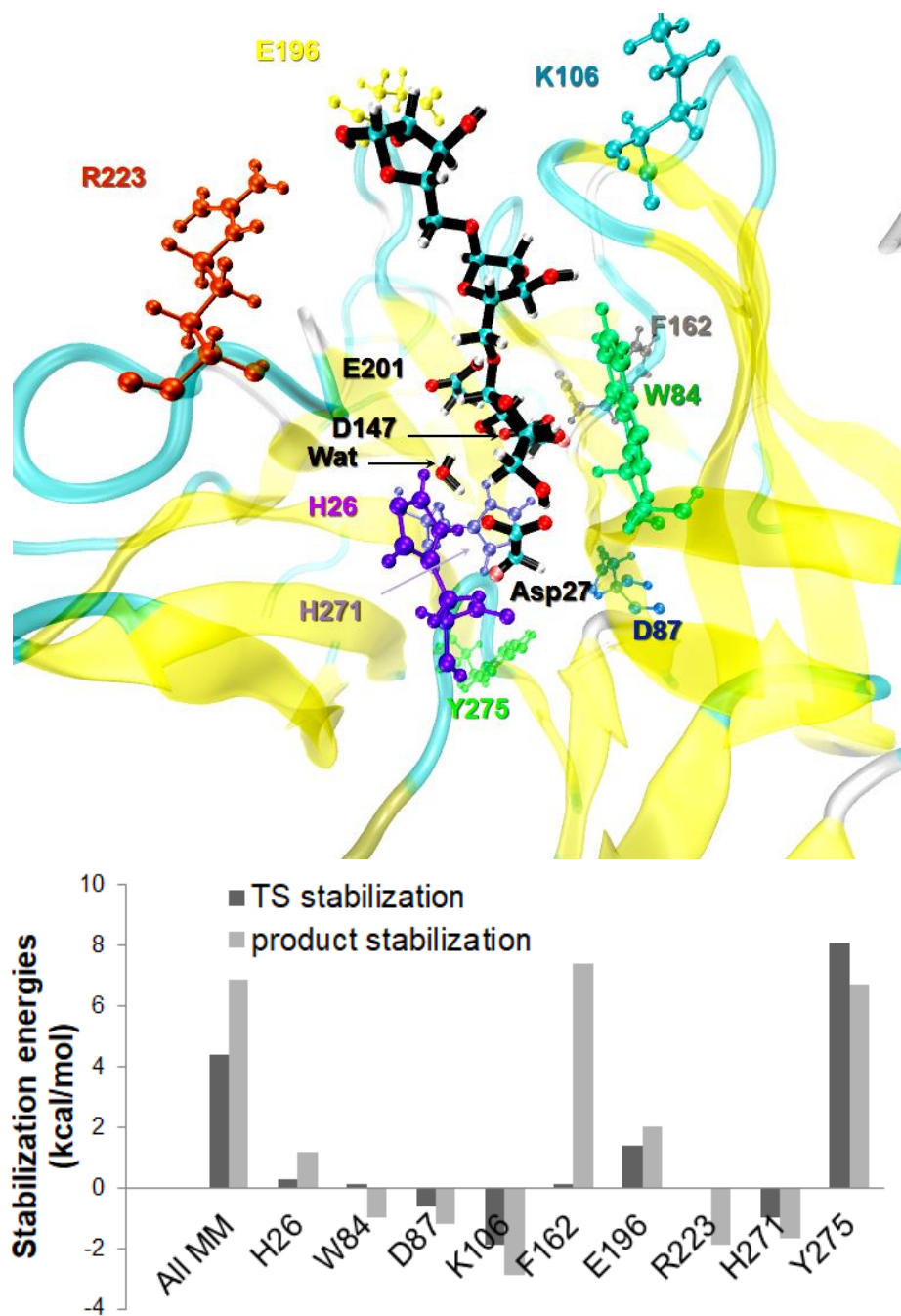


Figure 8. Key residues in TS and product stabilization (top) and their contributions (bottom) in stabilizing/destabilizing the QM region, relative to the reactant in the ABN reaction.

CONCLUSIONS

In the present work, we investigated in solution the glycosidic bond hydrolysis of arabinotriose catalyzed by ABN enzyme using MD simulations, potential energy surface, and PMF free energy calculations within the QM/MM framework. The QM/MM MD simulation yielded insights into the nature of the H-bond network around the pKa modulators in the active site, as required for efficient catalytic performance. Using the QM/MM/MD structures as the Michaelis complex of the reaction, the combined QM/MM(SCC-DFTB/CHARMM27) potential and free energy calculations were performed to explore the glycosidic bond hydrolysis of ABN toward the arabinotriose. With two-sets of RCs, the two-dimensional potential energy surface (2D-PES) was determined and the mechanistic and key geometric (active site) results were obtained to provide generally support the inverting mechanism of the GH43 ABN, as reported in the literature. The proton transfer to the glycosidic bond is found to be the rate-determining step of the acid/base catalysis, with the barriers of 21 and 23 kcal/mol for 2D-PES and PMF methods, respectively. Conformational sampling analysis of the -1 sugar ring at the Michaelis complex reveals a mixture of envelope/twist conformations. A conformational itinerary ($^2E \leftrightarrow E_3 \leftrightarrow ^4E$) in the -1 fructosyl ring is captured with significant oxocarbenium TS character. Residue electrostatic analysis on the key step (Figure 8) shows that residues Asp87, Lys106, His271 make negative contributions to the reaction, highlighting them as potential targets for experimental mutation studies.

SUGGESTIONS

The detailed mechanistic understanding developed here might be extended to other members of the GH family and may facilitate future rational design of biocatalyst

Chapter II

Quantum Chemical Cluster Approach as a Tool for Understanding Enzyme Mechanism and Selectivity in the Cascade Biocatalysis for Synthesis of Beta-Amino Acid

Manuscript I:

Meelua, W., T. Wanjai, N. Thinkumrob, J. Olah, J. R. Ketudat-Cairns, S. Hannongbua and **J. Jitonnorn**, Reaction mechanism and stereospecificity of catalytic reaction in DHP from *Saccharomyces kluyveri* and *Sinorhizobium meliloti* CECT4114: A systematic quantum chemical investigations, *to be submitted*.

REACTION MECHANISM AND STEREOSPECIFICITY OF CATALYTIC REACTION IN DIHYDROPYRIMIDINASE FROM *SACCHAROMYCES KLUYVERI* AND *SINORHIZOBIUM MELILOTI* CECT4114: A SYSTEMATIC QUANTUM CHEMICAL INVESTIGATIONS

INTRODUCTION

Dihydropyrimidinase (DHPase, EC 3.5.2.2) is involved in the degradation of pyrimidine nucleotide and is found in diverse organisms such as bacteria, yeast, animals, and plants (Gojkovic, Rislund et al. 2003, Ashihara, Ludwig et al. 2020). It plays an integral part of the pyrimidine catabolic pathway (Schnackerz and Dobritsch 2008), which is responsible for the regulation of the pyrimidine pool size available for nucleic acid synthesis and for supplying the cell with β -alanine. Deficiency of this enzyme may cause a risk from developing severe 5-fluorouracil (5FU)-associated toxicity (van Kuilenburg, Meinsma et al. 2003, Tsuchiya, Akiyama et al. 2019) and thus elucidating its role and inhibitory mechanism is the focus of current research in anticancer therapy (Huang 2015, Huang, Ning et al. 2019, Huang, Lien et al. 2020). The enzyme is also capable of detoxifying xenobiotics, making it attractive for antibiotic development and drug design (Huang 2015). Recently, this enzyme finds biotechnological applications for the industrial production of β -amino acids (Martínez-Gómez, Clemente-Jiménez et al. 2012, Clemente-Jiménez, Martínez-Rodríguez et al. 2014, Slomka, Zhong et al. 2015). Therefore, understanding of the mechanism of DHPase activity at atomic detail is of critical importance for both medical and biotechnological applications.

DHPases catalyze the reversible hydrolysis of dihydropyrimidines, 5,6-dihydrouracil (DHU) and 5,6-dihydrothymine (DHT), to the corresponding *N*-carbamyl- β -amino alanine (NC β A) and *N*-carbamyl- β -amino isobutyrate (NC β I), respectively, (see **Scheme 1**), in the second step of reductive pyrimidine degradation. This enzyme also hydrolyses a variety of other 5,6-dihydropyrimidines (Martínez-Rodríguez, Martínez-Gómez et al. 2010), as well as hydantoins and succinimides (Dudley, Butler et al. 1974). DHPases belong to the cyclic amidohydrolases superfamily (Seibert and Raushel 2005) that includes allantoinase, dihydroorotase, hydantoinase, and imidase. This group of enzymes shares their functional and structural similarity to each other (Holm and Sander 1997, Gojkovic, Rislund et al. 2003, Nam, Park et al. 2005) and has similar active site architecture, containing four histidines, one aspartate, and one post-carboxylated lysine residue, which are required for metal binding and catalytic activity (Huang, Hsu et al. 2009). Previous studies indicated that cyclic amidohydrolases may use a nearly identical mechanism, mainly based on the dihydroorotase (DHOase)-like scenario (Thoden, Phillips et al. 2001, Liao, Yu et al. 2008). However, depending on the nature of substrate and the enzyme

active site, a minor difference in detailed mechanism could be observed such as creatininase, a cyclic amidohydrolase enzyme acting on creatinine substrate (Jitonnorn, Mujika et al. 2017).

Previous mutagenesis and structural studies showed that the substrate specificity of dihydropyrimidinase (DHPase), hydantoinase (HYDase), allantoinase (ALNase), and dihydroorotase (DHOase) highly differ, despite structural similarities among their substrates (**Figure 1**). For example, the yeast (*Saccharomyces kluyveri*) DHPase (*SkDHPase*) and the slime mold (*Dictyostelium discoideum*) do not hydrolyze hydantoin (Gojkovic, Rislund et al. 2003) but the *Sinorhizobium meliloti* CECT4114 (*SmelDHPase*) does (Martínez-Rodríguez, Martínez-Gómez et al. 2010). Some bacterial HYDases are still named and identified as DHPase because of the catalytic activity toward the natural substrates DHU and DHT, including DHPase from *Pseudomonas aeruginosa* and *Thermus sp.* (Abendroth, Niefind et al. 2002, Huang 2015). Furthermore, DHOase does not hydrolyze dihydropyrimidine (DHU/DHT), hydantoin (HYD), and allantoin (ALN) (Peng and Huang 2014). These examples have raised the question whether the substrate of each enzyme competitively inhibits other enzymes in this family (Huang 2015). In the current study, these substrates were modeled in the active sites of *SkDHPase* and *SmelDHPase*, and their binding affinity and chemical reactivity were analyzed.

In recent years, several crystal structures of DHPases from different organisms have been solved (Abendroth, Niefind et al. 2002, Lohkamp, Andersen et al. 2006, Tzeng, Huang et al. 2016, Cheng, Huang et al. 2018, Huang, Ning et al. 2019), including the structure of *SkDHPase* in complex with the substrate DHU and product NCβI (Lohkamp, Andersen et al. 2006). An apo crystal structure for *SmelDHPase* has been reported (Martínez-Rodríguez, Martínez-Gómez et al. 2010) and its active site shares a similar geometries to *SkDHPase*. Furthermore, the structures of vertebrate DHPase from recombinant bacterial (Hsieh, Chen et al. 2013) have also been determined. The X-ray crystal structure of *SkDHPase* (Lohkamp, Andersen et al. 2006) shows that a unique hydrophobic pocket at the 5- and 6- positions of substrate is responsible for the stereospecificity of the enzyme. Within the active site of this enzyme, four histidine residues (His, His, His, and His), a carboxylated lysine 167 (Kcx167) and a nucleophilic water were crucial for the assembly of the binuclear metal center while other residues Tyr172, Ser331, and Asn392 are directly responsible for the strict substrate recognition (see **Figure 2**), as a part of three stereo-gate-loops (SGL1-3) (Cheon, Kim et al. 2002). Residues in these SGL regions are thought to be crucial for substrate binding prior to catalysis of enzymes from the amidohydrolase family: Ser331 is spatially conserved in ALNase (PDB 3E74, Ser288), imidazolonepropionase (PDB 2BB0, Ser329) and D-aminoacylase (PDB 1M7J, Ser289). Cys360 is at the van der Waals distance (3.6–3.8 Å) to N1 and C6 of DHU.

Tyr172 is of particular interest as its role in substrate binding and/or catalysis has been a subject of controversy by structural and mutagenesis studies (Cheon, Park et al. 2003, Lohkamp, Andersen et al. 2006). Since the Tyr conformation is high mobility, only one chain (of four other chains A-D) could be observed in the X-ray structure with its side-chain point towards the substrate (Lohkamp, Andersen et al. 2006). The structure of *SkDHPase* also points out that Tyr172 is far from interacting with the substrate DHU (PDB 2FVK), most probably due to the absence of clear electron densities for this residue, though it is at binding distance from the amine group of the reaction product (PDB 2FVM). Conversely, mutation of the analogous tyrosine (Tyr155) in *Bacillus stearothermophilus* D1 enzyme (Cheon, Park et al. 2003) and that

from *Sme*/DHPase, results in inactivation of the enzymes. This Tyr residue is also suggested to support correct positioning of the substrate, and also polarize the scissile amide bond together with Zn β , facilitating the attack of an activated hydroxyl group. Mutagenesis study also showed that this Tyr residue is involved in the inhibitor binding (Huang, Lien et al. 2020). These experimental observations clearly provide a clue to the importance of this active site conserved tyrosine residue. However, further investigations are still necessary to confirm this hypothesis whether it plays a role in substrate binding and/or catalysis.

Since DHPase enzymes are gaining more attractive as biocatalyst for the synthesis of industrially important non-natural amino acids (Syldatk, May et al. 1999, Engel, Syldatk et al. 2012, Heras-Vazquez, Clemente-Jimenez et al. 2012, Martínez-Gómez, Clemente-Jiménez et al. 2012, Clemente-Jiménez, Martínez-Rodríguez et al. 2014, Slomka, Zhong et al. 2015, Aganyants, Weigel et al. 2020), it is essential to know the reaction mechanism, the substrate specificity, and the stereospecificity of the enzyme towards different substrates. In particular, understanding the mechanism and origin of stereospecificity in different DHPase species towards natural and non-natural hydantoids/dihydropyrimidines is of great importance for biocatalytic application. Currently, to the best of our knowledge, the detailed information at atomic level for the hydrolysis of substrates of cyclic amidohydrolase family enzymes (**Figure 1**) is very limited. To extend the current knowledge of mechanism and stereospecificity of the enzyme, we have investigated the hydrolysis of the cyclic amidohydrolase substrates in the active site pocket of *Sk*DHPase. In particular, both natural substrates, dihydrouracil (DHU) and dihydrothymine (DHT), and non-natural substrates, hydantoid (HYD) and dihydroorotate (DHO), were considered as shown in **Figure 1**. On the basis of an available X-ray structure of the enzyme–substrate (ES) complex, a quantum cluster approach (Himo 2006, Siegbahn and Himo 2009, Siegbahn and Himo 2011, Jitonnorn, Sattayanon et al. 2015, Himo 2017) is used to (1) model enzyme reaction mechanism and (2) to probe the electronic and electrostatic factors governing the activity and stereospecificity of DHPase towards different substrates (**Figure 1**). This approach has previously been successfully applied to elucidate reaction mechanism and stereochemistry of various enzymes, including limonene epoxide hydrolase (Hopmann, Hallberg et al. 2005, Lind and Himo 2013), ω -transaminases (Cassimjee, Manta et al. 2015), human glyoxalase I (Jafari, Ryde et al. 2016), arylmalonate decarboxylase (Lind and Himo 2014). Our theoretical calculations suggest that DHPase may utilize a slightly different mechanism to the previously proposed DHOase-like mechanism (**Scheme 2**) and explain why dihydrouracil and dihydrothymine is the preferential substrates of this family of DHPase, not the hydantoid or other related substrates. Moreover, the role of the conserved tyrosine (Tyr172) and leucine (Leu) residues in catalysis is also investigated using an *in silico* point mutation approach. The ability of dihydropyrimidinase to hydrolyze DHU bearing methyl substituents at C5 and C6 position (5,6-dimethyl substituted DHU) was investigated. Our study provides a rationale for the experimentally observed activity and stereospecificity of the enzyme and details of the reaction mechanism, stereospecificity and mutation effects at the atomic scale may further aid the development of rationally designed enzyme variants.

OBJECTIVES

- (4) To model enzyme reaction mechanism of DHPase towards different substrates

- (5) To probe the electronic and electrostatic factors governing the activity and stereospecificity of DHPase towards different substrates (**Figure 1**).

METHOD

A model of the enzyme active site was constructed on the basis of the crystal structure of the native *SkDHPase* in complex with the substrate dihydrouracil (PDB 2FVK, **Figure 2**) (Lohkamp, Andersen et al. 2006). All crystallographic water molecules are removed. The hydroxide (OH⁻) was added between the two zinc ions and the ligand DHU was displaced with its oxygen atom positioned within hydrogen-bonding distance to Tyr172. Truncations were made for all the amino acids at the β -carbons, except three amino acids (Leu72, Tyr172, and Cys360) at α -carbons, and were kept fixed to their crystallographic positions, which was used to avoid unrealistic movements of the groups during the geometry optimizations (Siegbahn and Himo 2011, Jitonnorn, Sattayanon et al. 2015). These fixed atoms are marked with asterisks in the figures below. Hydrogen atoms were added manually and bonds that were truncated were saturated by hydrogen atoms. The geometries were optimized at the B3LYP/6-31G(d,p) level of theory. Based on these geometries, more accurate energies were obtained by performing single-point calculations with a larger basis set 6-311+G(2d,2p). Solvation effects were evaluated at the large basis set by performing single-point calculations on the optimized geometries using the CPCM solvation model (Barone and Cossi 1998, Cossi, Rega et al. 2003). The CPCM calculations used UFF atomic radii and default water solvent parameters, but the dielectric constant (ϵ) was set to 4, as recommended by Himo et al. (Himo 2006, Siegbahn and Himo 2011, Blomberg, Borowski et al. 2014). Frequency calculations were performed at the same theory level as the geometry optimizations to obtain zero-point energies (ZPE) and to verify the nature of the stationary points (minima or transition states). The final energies reported in the text are the large basis set energies corrected for both solvation and zero-point vibrational effects. All theoretical calculations were performed using Gaussian 09 program package (M. J. Frisch 2009).

RESULTS AND DISCUSSION

Generally, it is not well understood how the model size and the dielectric constant value chosen for the QM cluster calculations will influence the calculated properties. Therefore, in the present study, we start the mechanistic studies with a relatively small model and then increase the size, which we designed below as **Model I** and **Model II**, respectively.

3.1 Model I

For comparison purpose, we first design a small model (hereafter **Model I**) of the DHPase active site, containing only substrate (dihydropyrimidines DHU/DHT, hydantoid HYD or dihydroorotate DHO), the two zinc ions, and their ligands (namely, His62, His64, His199, His255, the carboxylated lysine Kcx167, and the general base/acid residue Asp358). The resulting models consist of 81-83 atoms, and have a total charge of +1, except for DHO which has zero net charge, owing to the negative charge of the COO⁻ group. This model essentially lacks the important substrate binding residues of the SGLs and, therefore, it is not expected to reproduce or rationalize the substrate specificity and stereospecificity. Nevertheless, it still

provides important information on the reaction energetics, mechanism, and the transition states, which will be used for comparison purpose when discussed on the impact of SGLs by **Model II** below. We first examined the sensitivity of the solvent dielectric on the results of **Model I** for DHU hydrolysis, data not shown. The results show that the solvent affects significantly on the shape of energy profile: it higher the activation energies and reaction energies, especially for the ring-opening step (**INT1** $\rightarrow$ **EP**) involving the protonation of cyclic amide ring of the substrate, assisted by Asp358. In particular, the activation barrier for the DHU hydrolysis is 21.9 kcal mol⁻¹, which is too high compared to the experiment (17.9 kcal mol⁻¹). The calculated energy profiles for the hydrolysis of all substrates using **Model I** and CPCM model are computed (data not shown), which clearly indicates that the enzyme can hydrolyze all substrates, as evident by the similarity in relative energies and activation barriers (17.5 – 22.3 kcal mol⁻¹). These data suggest that this model is too small to rationalize the stereospecificity of DHPase. Similar results have also been reported when the small model might not account well enough for the different electronic and flexibility of the surrounding environment (Jafari, Ryde et al. 2016). Therefore, more realistic representation of the active site is required, which will be further discussed by **Model II** below.

3.2 Model II

In the large model (hereafter **Model II**), some of the SGL residues (see **Figure 2**), which are closely related to substrate/product binding, were included in addition to the groups of **Model I**. These additional groups are Leu72 (in SGL-1), Tyr172 (in SGL-2), and Ser331 and Pro332 (in SGL-3). Also, the residues (Cys360, Asn392, and Gly393) that are not located in SGLs but are hydrogen bonding with the substrate or the SGL loops were also considered. A total size of the large model ranges from 176 to 182 atoms (176 atoms in HYD, 179 atoms in DHU, 182 atoms in DHT, and 181 and 182 atoms in charged and neutral DHO, respectively). All model systems have a net charge of +1, except DHO, which was modeled in both charged and neutral forms, leading to a net charge of 0 and +1, respectively. This model size is considered to be large compared to the earlier work of Liao et al. on DHOase (Liao, Yu et al. 2008). In the Liao's proposed models, the conserved residues (Ser331, Pro332 and Asn392) of the cyclic amidohydrolase family of the enzyme, does not include in their models. Besides, the analogous position of Tyr172 in the DHPase enzyme is absent in the DHOase enzyme, which is expected to cause different mechanistic detail between the two enzymes due to different electrostatic environment. In order to test the sensitivity of the large model (**Model II**) with the solvent dielectric, the energy profile of DHU hydrolysis was computed using the CPCM calculations with different dielectric constants ($\epsilon = 1, 4.0$ and 78.39 for gas-phase, protein and water, respectively). The results (data not shown) show that **Model II** is very insensitive to the higher dielectric constant (i.e., the computed energy profiles at $\epsilon = 4.0$ and $\epsilon = 78.39$ are almost identical). Thus, we chose **Model II** to describe the catalytic mechanism and the origin of stereospecificity of DHPase in the next section.

3.3 Reaction mechanism and substrate specificity in SkDHPase

In order to understand the mechanism and origin of stereospecificity of SkDHPase, both natural substrates, DHU and DHT, and non-natural substrates, HYD and DHO, were considered (**Figure 1**). The electronic and electrostatic factors governing the substrate specificity and stereospecificity of the enzyme, acting on these distinct substrates were obtained using DFT QM-cluster calculations in **Model II**. The binding affinity and specificity of these compounds within the

active site of *SkDHPase* were analyzed and the corresponding energetic properties (binding energies, activation energies, strained energies) are summarized in **Table 1**. We also modeled the hydrolysis in *SmeDHPase* to compare the activity differences between the two species and the results were discussed where it is necessary. Details of *SmeDHPase* model and calculations were provided in **ESI**. We discuss the results for each of these substrates in separate subsections.

(a) Dihydrouracil. The potential energy profile for the hydrolysis of DHU by *SkDHPase* is shown in **Figure 3** and the corresponding stationary points are shown in **Figure 4** for ES complex and **Figure 5** for transition states and intermediates. In the ES complex, the substrate tightly binds within the active site through several hydrogen bonds with the backbone atoms of Ser331, Pro332, Gly393 and the sidechain OH group of Tyr172. These interactions are maintained during the catalysis and also help in positioning the substrate for the subsequent nucleophilic attack (see details below). Note that this ES complex is not observed in the X-ray structure (PDB 2FVK) (Lohkamp, Andersen et al. 2006) where the hydroxide/water molecule was replaced by the 4-oxo group of DHU (see **Figures 2 and 4**).

The plot in **Figure 3** clearly indicates that the hydrolysis of DHU consists of two elementary steps, i.e., nucleophilic attack and ring opening. In the first step of the reaction, the bridging hydroxide performs a nucleophilic attack on the C4 atom of the substrate, leading to the formation of tetrahedral intermediate (**ES** $\rightarrow$ **INT**). The energy of **INT** is calculated to be as equal as the **ES** (0.1 and -0.7 kcal mol⁻¹ with and without solvation, respectively). In **ES**, the substrate initially binds at Zn β with a bond length of 2.12 Å (2.12 Å in **Model I**), as its fifth ligand, and its carbonyl group also forms a hydrogen bond (2.17 Å) with sidechain of Tyr172 (**Figure 4**). This observation is in contrast to the previous finding for the same Zn β -O distance (2.87 Å) in DHOase (Liao, Yu et al. 2008), which is probably due to the missing of Tyr residue at the analogous position. The distance between the bridging hydroxide and the carbon atom (C4) of the substrate is 2.60 Å. The bridging hydroxide symmetrically bonds to the two zinc ions (1.94 and 2.02 Å to Zn α and Zn β , respectively) and is hydrogen bonding (1.84 Å) with the catalytic residue Asp358.

From **ES** $\rightarrow$ **INT**, the critical C4-O_w distance is 1.92 Å at **TS1** and is fully bonded at **INT** with a bond distance of 1.50 Å (**Figure 5**). The barrier of this step is calculated to be 6.1 kcal mol⁻¹ (4.0 kcal mol⁻¹ without solvation). The resulting oxyanion of the carbonyl group binds to Zn β with a bond length of 1.90 Å, which is shortening by 0.2 Å from **ES**. This demonstrates that Zn β provides electrostatic stabilization for the transition state and intermediate, thereby lowering the barrier of this step. Also, the shortening of the two hydrogen-bonds between Asp358 and the bridging hydroxide (from 1.84 Å in **ES** to 1.66 Å in **INT**) and between the hydroxyl group of Tyr172 and the O1 atom of substrate (from 2.17 Å in **ES** to 1.78 Å in **INT**) implies that Asp358, Tyr172, and Zn β all anticipate in stabilizing the tetrahedral intermediate. During this step, the coordination number of Zn β changes from 5 to 4, resulting from the detachment of O_w from Zn β (2.02 Å at **ES** to 2.70 Å at **INT**).

The next step is the opening of ring nitrogen upon the protonation assisted by Asp358. From **INT** $\rightarrow$ **EP**, Asp358 abstracts a proton from the hydroxide and delivers it to the nitrogen leaving group. This step was found to occur simultaneously through one transition state (**TS2**). At **TS2**, the scissile C-N bond is 1.59 Å and the transferring proton is in between the carboxylate Asp358 and the ring nitrogen (1.32 Å and 1.41 Å, **Figure 5**). The barrier of this step

was 14.9 kcal mol⁻¹ (9.0 kcal mol⁻¹ without solvation correction) indicating that this process is the rate-determining step of the whole reaction. No experimental rate constant (k_{cat}) has been measured for SkDHPase. However, k_{cat} value for DHU hydrolysis catalyzed by *Smel*DHPase has been reported to be 2 s⁻¹ (Martínez-Rodríguez, Martínez-Gómez et al. 2010), corresponding to energy barriers of ~17.9 kcal mol⁻¹. Finally, the product carbamoyl aspartate is obtained and the overall reaction is exothermic by -6.6 kcal mol⁻¹ (-9.7 kcal mol⁻¹ without solvation correction). The reverse reaction has the barrier of 21.5 kcal mol⁻¹, which is 6.6 kcal mol⁻¹ higher than that of the forward reaction. This might be the reason why only the product NCβA was identified in the co-crystallization experiment with DHU (Hsieh, Chen et al. 2013).

(b) Dihydrothymine. DHT can bind in the DHPase active site in two different forms, namely *L*- and *D*-configuration (hereafter *L*-DHT and *D*-DHT, respectively). Previous crystallographic studies suggested that the *L*-conformation of the substrate is more likely based on the X-ray observation of the NCβI product (Lohkamp, Andersen et al. 2006, Hsieh, Chen et al. 2013). To confirm the stereopreference of DHPase enzyme toward this substrate, we modeled the DHT substrate in both forms (*L*-DHT and *D*-DHT) and their subtle energetic difference in the hydrolytic reaction was evaluated, as shown in **Figure 6**. The geometries of the ES complex for *L*-DHT and *D*-DHT are shown in **Figure 7**, whereas the corresponding stationary points are depicted in **Figure 1**, respectively.

The reaction mechanism for hydrolysis of DHT is essentially similar to the case of DHU, i.e., the nucleophilic substitution by the hydroxide ion, followed by the abstraction of its proton by Asp358, and then the ring-opening of the substrate assisted by the same Asp. The main difference is that a barrier is observed during the proton abstraction from a hydroxide ion by Asp358 (**INT1** → **INT2**, **Figure 6**), and thus the overall reaction consists of three elementary steps through **TS1**, **TS2** and **TS3**. In case of *L*-DHT, the substrate initially binds to Znβ with a bond length of 2.13 Å and its position is placed through four H-bonds with Tyr172, Ser331, Asn392, and Gly393, supporting their critical role in substrate binding (Lohkamp, Andersen et al. 2006, Huang 2015). These H-bond networks maintain throughout the catalysis. On the contrary, for *D*-DHT, these interactions, especially Ser331 and Tyr172, are loosely bound to the substrate (1.93 vs 3.28 Å for DHT@O1...HO@Tyr172 distance and 1.86 vs 2.52 Å for DHT@N1-H...O@Ser331 distance for *L*-DHT vs *D*-DHT, respectively; **Figure 7**). The poor affinity of *D*-DHT towards the enzyme is also reflected by the lowest binding interaction energies (-5.8 kcal mol⁻¹, **Table 1**). This unfavorable binding of *D*-DHT leads to a higher barrier (~10 kcal mol⁻¹) in the nucleophilic attack step in comparison with that of *L*-DHT.

In the first step, the C4-O_w distance is shortening from 2.62 Å at **ES** to 1.89 Å at **TS1** and finally 1.51 Å at **INT1** for *L*-DHT (corresponding to 2.88, 1.87 and 1.50 Å for *D*-DHT, respectively). The bridging hydroxide symmetrically bonds to the dizinc center at **ES** (~1.9 and ~2.0 Å to Znα and Znβ, respectively for both *L*-DHT and *D*-DHT) and then become asymmetric with lengthening of the O_w-Znβ distance (2.01 Å at **ES** to 2.27 Å at **TS1** and 2.66 Å at **INT1**). During this step, for *L*-DHT, Asp358 forms a stronger hydrogen bonding to the hydroxide in ranges of 1.67-1.84 Å in comparison with the values for *D*-DHT (1.73-1.86 Å). In the second step, the proton transfer from the hydroxide to the Asp358 residue occurs via **TS2** with the overall barriers of 9.2 and 21.0 kcal mol⁻¹ and the reaction are endothermic by 7.2 and 17.6 kcal mol⁻¹ relative the **ES** energy for *L*-DHT and *D*-DHT, respectively. The transient bonds, H_w-O_w and

H_w-O@Asp358, are 1.35 and 1.29 Å at **TS2**, respectively. The last step is the ring-opening of cyclic substrate, where Asp358 donates its received proton to the amide nitrogen of DHT to yield the opened product NCβI with the C-N bond being fully breakdown (2.85 Å at both **EP-II-L** and **EP-II-D**). The scissile C-N bonds are 1.61 and 1.59 Å at **TS3** for *L*-DHT and *D*-DHT, respectively, and the transient bonds, Asp358@O...H_w and H_w...N3@DHT, are 1.32 and 1.42 Å for *L*-DHT (1.29 and 1.39 Å for *D*-DHT, respectively). This process is the rate-determining step with the highest barrier (21.1 kcal mol⁻¹ for *L*-DHT and 26.8 kcal mol⁻¹ for *D*-DHT; **TS3** in **Figure 6**).

Overall, the reaction with *D*-DHT exhibits a higher energy profile in comparison with the case of *L*-DHT. This can be realized from the unfavorable interaction between the C5-methyl substituent of DHT and the dizinc center at **ES**. The steric hindrance at C5 position also results in a distorted conformation of *D*-DHT and a loose binding to the SGL-2/SGL-3 residues (Tyr172, Ser331 and Pro332) (**Figures 7 and 9**). The *k*_{cat} value for the DHT hydrolysis catalyzed by *Sme*I DHPase has been reported to be 1 s⁻¹ (Martínez-Rodríguez, Martínez-Gómez et al. 2010), which is corresponding to an energy barrier of ~18.4 kcal mol⁻¹. Thus, the calculated overall barrier for DHT hydrolysis of 21.1 kcal mol⁻¹ (**Figure 6**), which is 3.2 kcal mol⁻¹ higher than the corresponding value for the DHU hydrolysis, is still in-line with the kinetic observed trend (i.e., DHU is hydrolyzed faster than DHT). The reverse reaction has the barriers of 25.9 kcal mol⁻¹ for *L*-DHT and 23.8 kcal mol⁻¹ for *D*-DHT, which are higher than its forward reaction. This finding could explain why only the product of DHT substrate was observable in the X-ray structure of vertebrate DHPase (Hsieh, Chen et al. 2013).

(c) Hydantoid. The reaction energy profile calculated for HYD hydrolysis is shown in **Figure 10** and the geometries of all stationary points on the energy profile are computed. The mechanistic details of this five-membered ring substrate (HYD) are somewhat more complex than the case of the six-membered ring substrates (DHU and DHT). As shown, the reaction consists of four elementary steps, proceeding via **TS1**, **TS2**, **TS3**, and **TS4**. The substrate is first coordinated with Znβ (2.25 Å) and its orientation is placed via van der Waals and H-bond interactions. The nucleophilic attack of hydroxide on the C4 of substrate that leads to the formation of tetrahedral intermediate occurs, as can be seen from the shortening of C4-O_w distance from 2.52 Å at **ES**, 1.99 Å at **TS1** and 1.48 Å at **INT1**. This step has a barrier of 5.9 kcal mol⁻¹ and is slightly exothermic (-1.0 kcal mol⁻¹ at **INT1**). Next, Asp358 functions as a base in abstracting a hydroxide proton, which requires the barrier of 4.3 kcal mol⁻¹ relative to **ES**. The transferred proton is in the middle of the hydroxide oxygen and the Asp358 carboxylate (~1.26 Å at **TS2**). The third and fourth steps involve the rotation of Asp358 with respect to the HYD amide group and its proton transfer to the amide nitrogen of the substrate, which subsequently breaks down the hydantoin ring, via C-N bond cleavage, and forms the *N*-carbamoyl-α-amino acid product. This proton transfer was found to take place in a separate step via **TS3** and **TS4**, which is not observed in other six-membered ring substrates. At **TS3**, the transferred proton is at 1.92 and 2.17 Å distant from the bridging oxygen and the substrate amide nitrogen, respectively. At **TS4**, the transferring proton is in between the carboxylate Asp358 and the ring nitrogen (1.30 and 1.39 Å). The ring opening is completed as evident by the elongation of the scissile C-N bond from 1.51 Å at **INT3** to 2.81 Å at **EP**. The last step is the rate-determining step, with the barrier of 11.3 kcal mol⁻¹, which is only 3 kcal mol⁻¹ lower than the reverse reaction (14.4 kcal mol⁻¹). The close in the barriers also implies that the product could transform back to the reactant, and this

can explain why the ring-opened product of HYD was not observed in the X-ray structure (Hsieh, Chen et al. 2013).

The k_{cat} value for the HYD hydrolysis catalyzed by *Sme*/DHPase has been reported to be 12 s^{-1} ¹², which corresponds to an energy barrier of $\sim 16.8 \text{ kcal mol}^{-1}$. The difference in both calculated and experimental barriers (11.3 vs $16.8 \text{ kcal mol}^{-1}$) could be realized from the different specificity between the two enzymes, *Sk*DHPase and *Sme*/DHPase, toward hydantoin, (i.e., the latter does hydrolyze HYD¹² but the former does not (Gojkovic, Rislund et al. 2003)). See section 3.4 for more details. This interpretation is also supported by the relatively low binding affinity of HYD towards the *Sk*DHPase active site in comparison with the case of *Sme*/DHPase (BE = $-6.0 \text{ kcal mol}^{-1}$ for *Sk*DHPase and $-11.3 \text{ kcal mol}^{-1}$ for *Sme*/DHPase, **Table 1**), which is also observed from structural comparison (Lohkamp, Andersen et al. 2006).

(d) Dihydroorotate. The potential energy surface of the reaction with DHO is shown in **Figure 11**. The optimized geometries of the stationary points on the surface are computed. Starting from **ES**, the hydroxide performs a nucleophilic attack at the C4 atom of substrate and a tetrahedral intermediate (**INT**) is formed with the C4-O_w bond length of $1.51 \text{ \AA}$. The barrier of this step is $8.1 \text{ kcal mol}^{-1}$ at **TS1**. During this step, the coordination number of Zn β changes from 5 to 4, due to the elongation of O_w-Zn β distance ($1.98 \text{ \AA}$ at **ES** to $2.64 \text{ \AA}$ at **INT**). The shortening of the two distances between the substrate carbonyl oxygen to Zn β and Tyr172 (see O1-Zn β and O1-HO@Tyr172) indicates that these molecules play a role in transition state stabilization. The next step is the opening of the DHO ring and Asp358 plays a dual role in shuttle a proton from the hydroxide of **INT** to the nitrogen leaving group. This step was found to occur simultaneously through **TS2** and is the rate-determining step with the highest barrier ($17.9 \text{ kcal mol}^{-1}$). The scissile C-N bond is $1.59 \text{ \AA}$ and the transferring proton is in between the carboxylate Asp and the ring nitrogen ($1.26 \text{ \AA}$ and $1.40 \text{ \AA}$). After **TS2**, the C-N bond is fully breakdown with a distance of $2.73 \text{ \AA}$, and the product carbamoyl aspartate is obtained. The activation barrier for the hydrolysis of DHO in both charged and neutral forms are calculated to be 17.9 - $18.4 \text{ kcal mol}^{-1}$ ($14.4 \text{ kcal mol}^{-1}$ without solvation), which is $\sim 3 \text{ kcal mol}^{-1}$ higher than the corresponding barrier ($14.9 \text{ kcal mol}^{-1}$) in the case of DHU. Similar binding energies between the DHU and DHO were found. Values of BE of DHU and DHO are -10.8 and $-10.4 \text{ kcal mol}^{-1}$. This data could explain from the energetic point of view why DHU is a substrate of DHPase, not DHO, and, *vice versa*, why the X-ray crystal structure of DHOase in the presence of DHU was not observed (Lee, Maher et al. 2007). These data further confirms the substrate specificity of the enzyme.

(e) 5FU and 5FUH2. It is also interesting to gain further insights into how 5FU and 5FUH2 were metabolized by the enzyme DHPase (see **Scheme 3**). Here, two compounds 5FU and 5FUH2 in both *L* and *D* forms (namely, *L*-5FUH2 and *D*-5FUH2, respectively) were modeled in the active site of *Sk*DHPase and their possible degradation pathways were theoretically explored and characterized at B3LYP/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d) level. The models were also superimposed with the crystal structure of the 5-FU-dihydropyrimidinase complex (PDB entry 6KLK) (data not shown). Values of BE of 5FU and 5FUH2 are calculated to be $-11.8 \text{ kcal mol}^{-1}$ for 5FU and -16.1 and $-7.7 \text{ kcal mol}^{-1}$ for *L*-5FUH2 and *D*-5FUH2, respectively. This data clearly suggests that the enzyme selectively bind with 5FUH2 over 5FU, and the *L* form is the favorable one. The optimized structures at **ES** and **TS3** for both compounds

are shown in **Figure 12**. All stationary geometries are provided in ESI. Considering the energy profiles shown in **Figures 13-14**, the degradation of 5FU follows a 4-step mechanism, while the hydrolysis of *L*-5FUH2 and *D*-5FUH2 proceed in three steps as in the case of **DHT**.

(f) Relationship between ring strain energy and activation barrier. In this section, we further calculated the ring strain energy, which is defined to be the reaction energy of a balanced chemical reaction in which the reactant and product differ by the presence of a ring. As shown in **Table 1**, using **Model I**, the relative strain energies of DHU, *D*-DHT, HYD, *L*-HYD, *D*-HYD, DHO are calculated to be 5.4, 3.9, 7.0, 3.6, 4.8, 8.5 kcal mol⁻¹ relative to the lowest RSE of *L*-DHT (-32.2 kcal mol⁻¹) as a reference molecule. The overall activation barriers of both **Model I** and **II** (namely, E_a -I and E_a -II) and the ring strain energies (RSE) for the reaction with the natural substrates and non-natural substrates in this study are summarized in **Table 1**. Plot of activation barrier E_a -I against the ring strain energy RSE is shown in **Figure 15**, which indicates a good linear correlation between the two values ($R^2 = 0.84$). Considering E_a -I and E_a -II, it is seen that the SGL residues of the enzyme mostly stabilizes the reaction with dihydropyrimidine and hydantoin by lowering the barrier of the ring-opening step, as evidenced by the negative ΔE_a values. On contrary, the SGL residues show a destabilizing effect, with positive ΔE_a , for the non-natural substrates (DHO in neutral, PTA, and ALN), the 5FU compounds, and the case of unfavorable substrate *D*-DHT.

To predict the enzyme-substrate specificity, both the substrate binding and the chemical steps (i.e., activation energies) should be considered (Tian, Wallrapp et al. 2013). Accordingly, we further evaluated the binding affinities (BE) of the substrates of these family enzymes (**Figure 1**) within the active site of *SkDHPase*, as obtained by deleting the substrate from the ES structure (complex) using equation: $BE = E(\text{complex}) - E(\text{receptor}) - E(\text{substrate})$. This QM-based approach (Ryde and Söderhjelm 2016, Yilmazer and Korth 2016) has been successfully applied in protein-ligand interaction studies (Ahumado, Drosos et al. 2014). The results of B3LYP and M06-2X are tabulated in **Tables 1** and **Figure 16**. Both methods show a linear correlation ($R^2 = 0.94$), which suggest that the dispersion interaction does not affect the relative binding of the substrates studied. The binding energies of the natural substrates DHU and *L*-DHT are estimated to be -10.8 and -15.0 kcal mol⁻¹, which are 2-3 order more stable than the cases of non-natural substrates (-5.8, -6.0, -6.8 kcal mol⁻¹ for *D*-DHT, HYD, *L*-HYD, respectively). Thus, our BE calculations do reflect the experimentally observed stereospecificity of the enzyme and this further confirms the fact that the substrate specificity of the enzyme is mainly dictated in the enzyme-substrate binding.

3.4 Probing the stereospecific difference between *SkDHPase* and *SmeIDHPase* towards hydantoin and dihydrouracil derivatives

Previous studies (Martínez-Rodríguez, Martínez-Gómez et al. 2010) have reported that *SmeIDHPase* can be used for the hydrolytic opening of non-natural 6-monosubstituted dihydrouracils (e.g., 6-MeDHU). However, the *SkDHPase* activities towards 6-MeDHU have not been described experimentally. To investigate stereospecific behavior of *SkDHPase* and *SmeIDHPase* towards hydantoin and dihydrouracil derivatives bearing mono- or di-methyl substituents, we have theoretically modeled the hydrolysis of the 6-MeDHU (both *R* and *S* configurations at C6 position) and the 5,6-MeDHU (both *SS* and *SR* configurations at C5

and C6 positions, respectively). See details of chemical structures of these derivatives in **Figure 17**. Energy profiles for the hydrolysis of (5*R*)-methylhydantoid (denoted as (5*R*)-MeHYD) and (5*S*)-methyldihydrouracil (denoted as (5*S*)-MeDHU), 6-methyldihydrouracil (*S* and *R* isomers, denoted as (6*S*)-MeDHU and (6*R*)-MeDHU) and 5,6-dimethyldihydrouracil (*SS* and *RR* isomers, denoted as (5*S*, 6*S*)-Me₂DHU and (5*R*, 6*R*)-Me₂DHU) are plotted in **Figure 18** and the corresponding activation barriers and reaction energies are tabulated in **Table 2**.

Results in **Figure 18** and **Table 2** clearly showed that the two DHPases from yeast and bacteria have different preferences in the stereospecificity towards 5-monosubstituted HYD/DHU substrates. For *SkDHPase*, the activation barrier and reaction energy are calculated to be 18.4 and +3.6 kcal mol⁻¹ for (5*R*)-MeHYD, and 15.8 and -6.1 kcal mol⁻¹ for (5*S*)-MeDHU, respectively. Notably, no kinetic experiment has been done for *SkDHPase*, but only those *SmelDHPase* were previously reported for a range of hydantoid and dihydrouracil derivatives. For *SmelDHPase*, a good agreement is found between the experimental and calculated barriers toward (5*R*)-MeHYD and (5*S*)-MeDHU, i.e., 16.7 vs 15.1 kcal mol⁻¹ and 20.3 vs 18.4 kcal mol⁻¹, respectively. The overall reaction of the six-membered ring substrate (5*S*)-MeDHU is slightly exothermic for both enzymes (~3-6 kcal mol⁻¹). These results indicate that *SkDHPase* has stereospecific preference toward (5*S*)-MeDHU over (5*R*)-MeHYD, while *SmelDHPase* shows opposite trend as observed by Martínez-Rodríguez et al. (Martínez-Rodríguez, Martínez-Gómez et al. 2010)

For 5,6-disubstituted DHU, *SkDHPase* catalyzed the hydrolysis of (5*S*,6*S*)-MeDHU with the $\Delta E_{TS}^\ddagger$ and ΔE_R values of 21.1 and +4.4 kcal mol⁻¹, while the reaction barrier and energies were lower by 5-9 kcal mol⁻¹ (16.0 and -4.1 kcal mol⁻¹) when reaction with (5*S*,6*R*)-MeDHU. On the other hand, the *SmelDHPase* specifically hydrolyzed the (5*S*,6*S*)-MeDHU over the case of (5*S*,6*R*)-MeDHU (16.5 vs 19.5 kcal mol⁻¹, respectively). The overall reaction with (5*S*,6*R*)-MeDHU has an exothermic character similar to the cases of (5*S*)-MeDHU and (6*R*)-MeDHU, whereas a small endothermicity is found in (5*S*,6*S*)-MeDHU (~1-4 kcal mol⁻¹), as also seen in (6*S*)-MeDHU. Such the ΔE_R values seem to depend on the stereochemistry at C6 position. Overall, the results clearly indicate that *SkDHPase* specifically hydrolyzed (5*S*,6*R*)-MeDHU over (5*S*,6*S*)-MeDHU, while *SmelDHPase* favors in (5*S*,6*S*)-MeDHU.

3.6 Expanding the substrate scope with *in silico* point mutations

To further quantify the impact of stereo-gate-loop (SGL) residue(s) (Cheon, Kim et al. 2002) in governing the enzyme substrate specificity and reactivity, we have performed point mutations in *SkDHPase* by replacing a mutated SGL residue with a hydrophobic residue such as phenylalanine using *in silico* method. Here, two mutations, Leu72Phe and Tyr172Phe, located at SGL-1 and SGL-2, respectively, were considered and their mutation effects on the energy profiles towards DHU were computed as shown in **Figure 19**, with the result of WT included for comparison. Detailed information on the mechanistic results is provided in Supplementary. Clearly, both mutants largely increase the energy profile in comparison with the WT enzyme. This data also support why the single substitution of this Tyr residue inactivates the observed activities in Y155F and Y155E mutants from *Bacillus stearothermophilus D1* (Cheon, Park et al. 2003) and the Y152A *SmelDHPase* mutants.

Furthermore, we also estimated the binding (interaction) energies of these mutants towards a range of substrate, with the corresponding values included in **Table 4**. In this study, six substrates (DHU, *L*-DHT, *D*-DHT, HYD, *L*-HYD, and *D*-HYD) were chosen for the binding energy calculations. All of them are bound in very similar ways via hydrogen bonds involving atoms N1, O2, and N3 of the substrates with backbone atoms (S331 N and O and N392 O for *SkDHPase*). The mutant results for DHO were excluded here due to the converse problem. It is found that despite the observed lower rate of DHU hydrolysis in both mutants, the calculated binding affinities seems to be improved for most substrates, with the BE values in ranges of -9.2 to -13.6 kcal mol⁻¹ for Tyr172Phe and -6.7 to -11.8 kcal mol⁻¹ for Leu72Phe in comparison with the values for WT (-5.8 to -15.0 kcal mol⁻¹). For example, the poor binding affinities in WT for *D*-DHT, HYD and *L*-HYD were increased by about two-order in the Tyr172Phe. The results of DHO cannot be expected to be directly comparable to the other substrates due to the DHO charge difference with respect to the carboxylate group. It can be seen that the highest binding energies in DHO in charged form. Overall, these data point out that (1) the SGL residues (Tyr172 and Leu72 as example) have a direct impact to the reactivity and stereospecificity of the enzyme and (2) it is possible from energetic point of view to improve the substrate binding affinity via the hydrophobic point mutations. However, the activation energies of the reaction should be optimized as well.

CONCLUSIONS

In the present work, the mechanism and stereospecificity of DHPase-catalyzed reaction towards distinct substrates have been investigated by DFT quantum chemical approach. Two models of the active site were designed based on the available X-ray crystal structure of enzyme-substrate complex. The reaction mechanism proposed on the basis of the calculations is shown in **Figure 1**, and the obtained energetic values regarding substrate binding ability and substrate specificity are given in **Table 1**. Following a previously proposed DHOase-like mechanism, the overall reaction consists of three elementary steps: (1) nucleophilic attack by the bridging hydroxide is facilitated by abstraction of its proton by the side chain carboxylate of Asp358, (2) collapse of the tetrahedral intermediate and (3) cleavage of the carbon-nitrogen bond occurs with donation of a proton from the protonated form of Asp358.

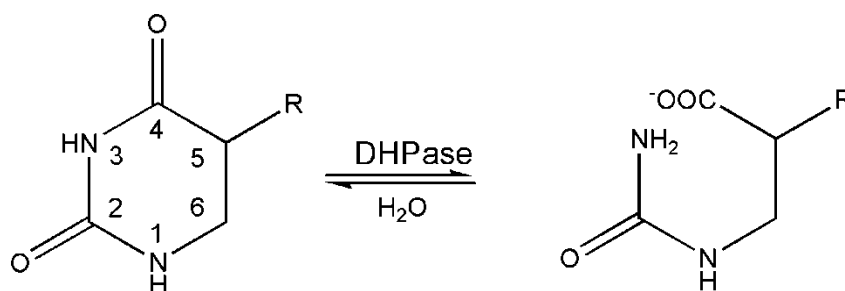
By comparing the energetics of the reaction with *L*-DHT and *D*-DHT, a preference of the *SkDHPase* for the *L*-conformer was observed, consistent to the structural studies reported previously. The calculations clearly indicated that the configuration of the C5 carbon plays a crucial role in the binding affinity of the DHT substrate and contributes to the stereospecificity of the catalysis.

In silico mutation with QM calculations have confirmed the mutational and structural evidence that Tyr172 plays a critical role in catalysis of the enzyme, i.e. stabilizing the transition state and positioning the substrate toward Zn^β. Both mutants (Tyr172Phe and Leu72Phe) exhibited the higher energy profiles for DHU hydrolysis in comparison with the wildtype, with a slightly improvement in the substrate binding. The results also demonstrate that the replacement of Leu72 and Tyr172 with phenylalanine leads to an increase in the activation barriers towards DHU.

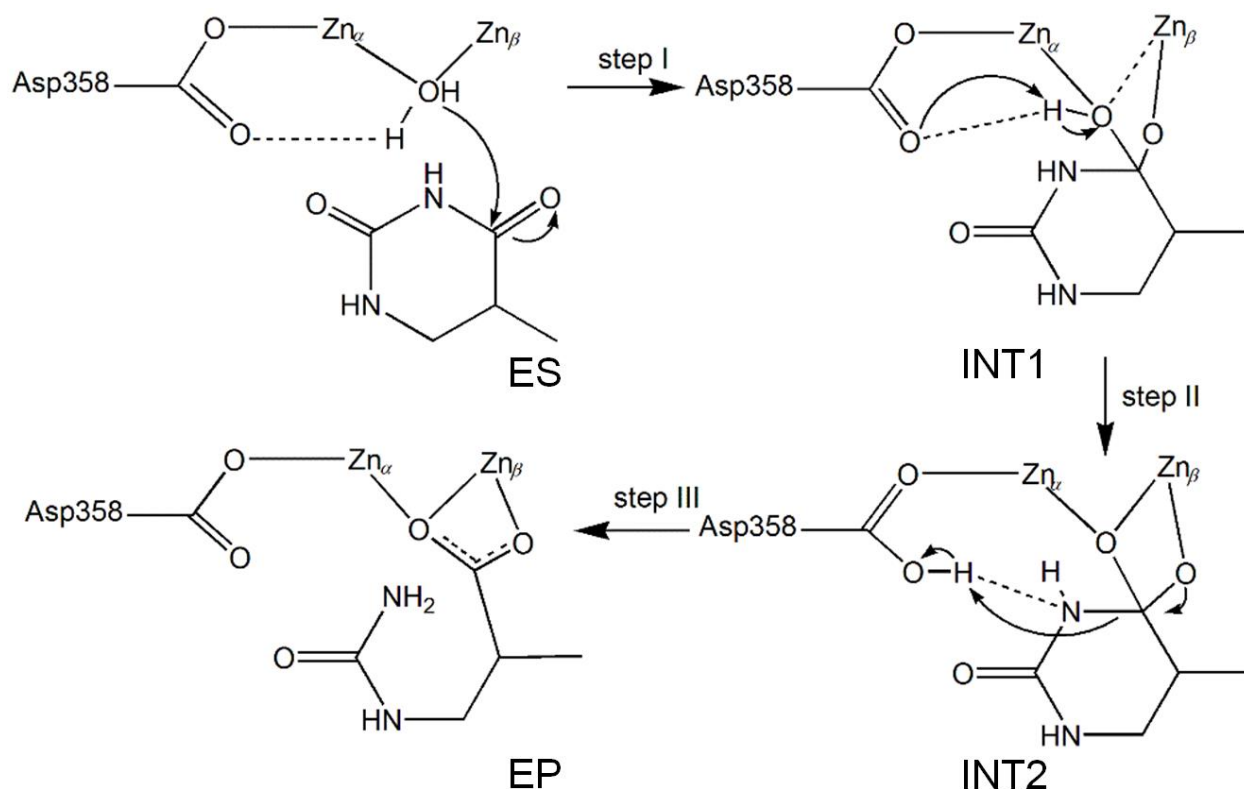
The stereospecificity of the enzyme was also investigated in this current study. The calculations could explain the experimental observation, the preferred L-conformation of the product, and also provide a rationale for this observation. We also observed different preferences in the substrate specificity of the DHPase from yeast and bacteria, and a good agreement is found in *Smel*DHPase between activation barriers observed experimentally and those predicted using computational methods.

SUGGESTIONS

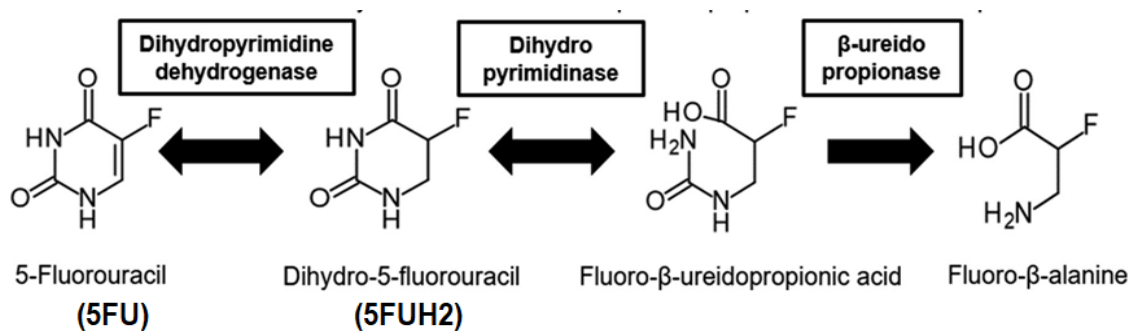
The detailed mechanistic understanding developed here might be extended to other members of the cyclic amidohydrolase family and may facilitate future rational design of biocatalyst for the production of β -amino acids.



Scheme 1. Reaction catalyzed by DHPase



Scheme 2. Mechanism of the ring-opening reaction catalyzed by DHPase, which is based on the DHOase-like mechanism (Liao, Yu et al. 2008).



Scheme 3. Degradation pathway of 5FU into fluoro- β -alanine catalyzed by dihydropyrimidine dehydrogenase, dihydropyrimidinase, and β -ureidopropionase, respectively.

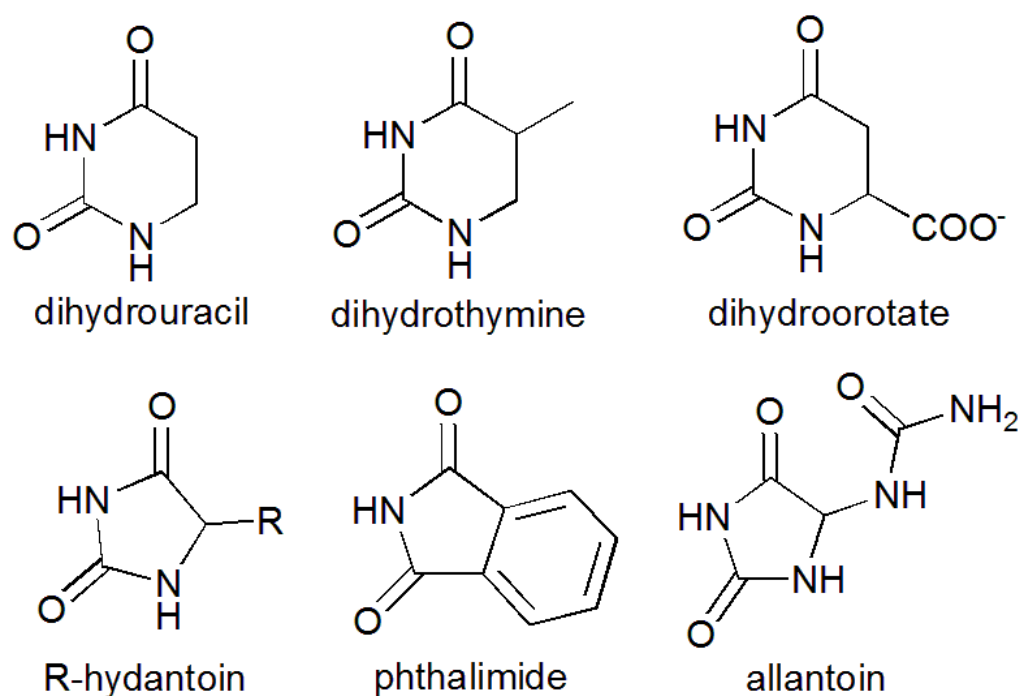


Figure 1. Substrates of dihydropyrimidinase, hydantoinase, imidase, allantoinase, and dihydroorotase.

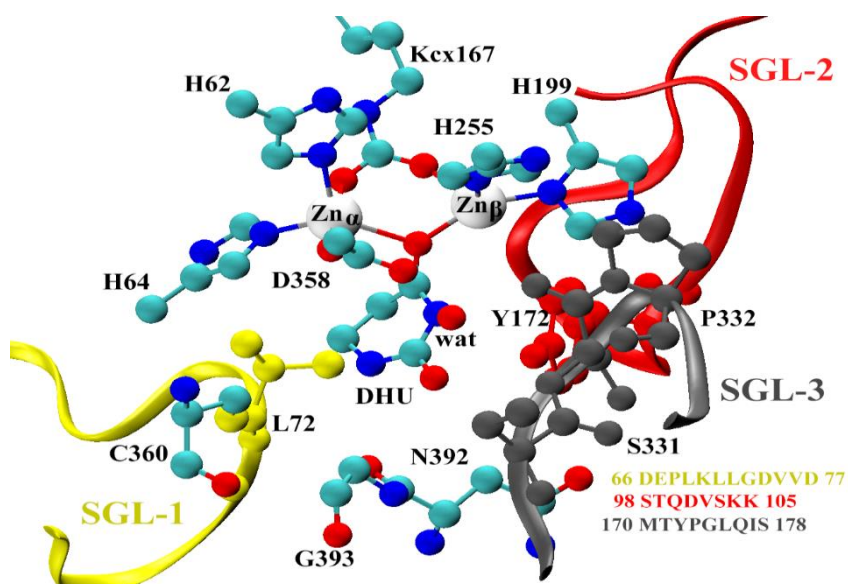


Figure 2. X-ray crystal structure of the active site of *S. kluyveri* DHPase in complex with substrate dihydrouracil (labeled as DHU) (coordinates taken from PDB 2FVK (Lohkamp,

Andersen et al. 2006)). Ribbons indicate the three stereo-gate-loops (SGL1-3) shown in yellow, red and grey, respectively. Conserved amino acids line in each SGLs are also indicated.

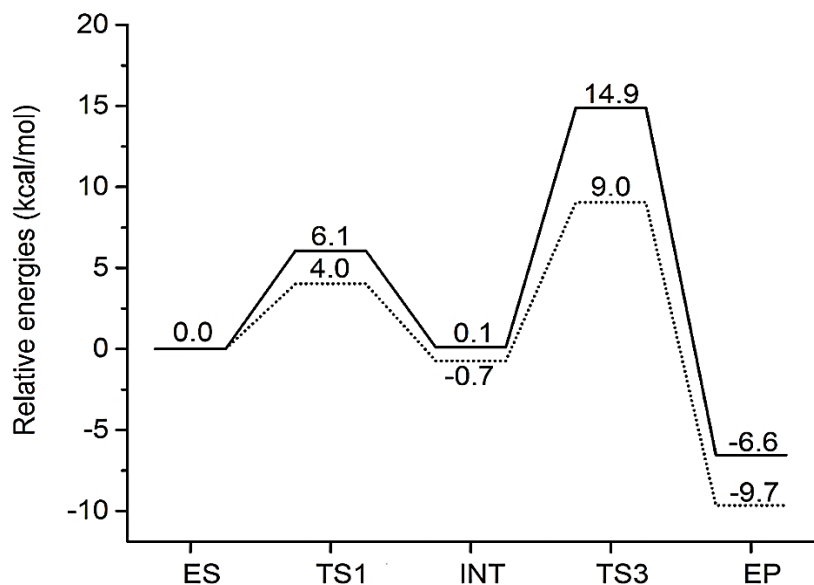
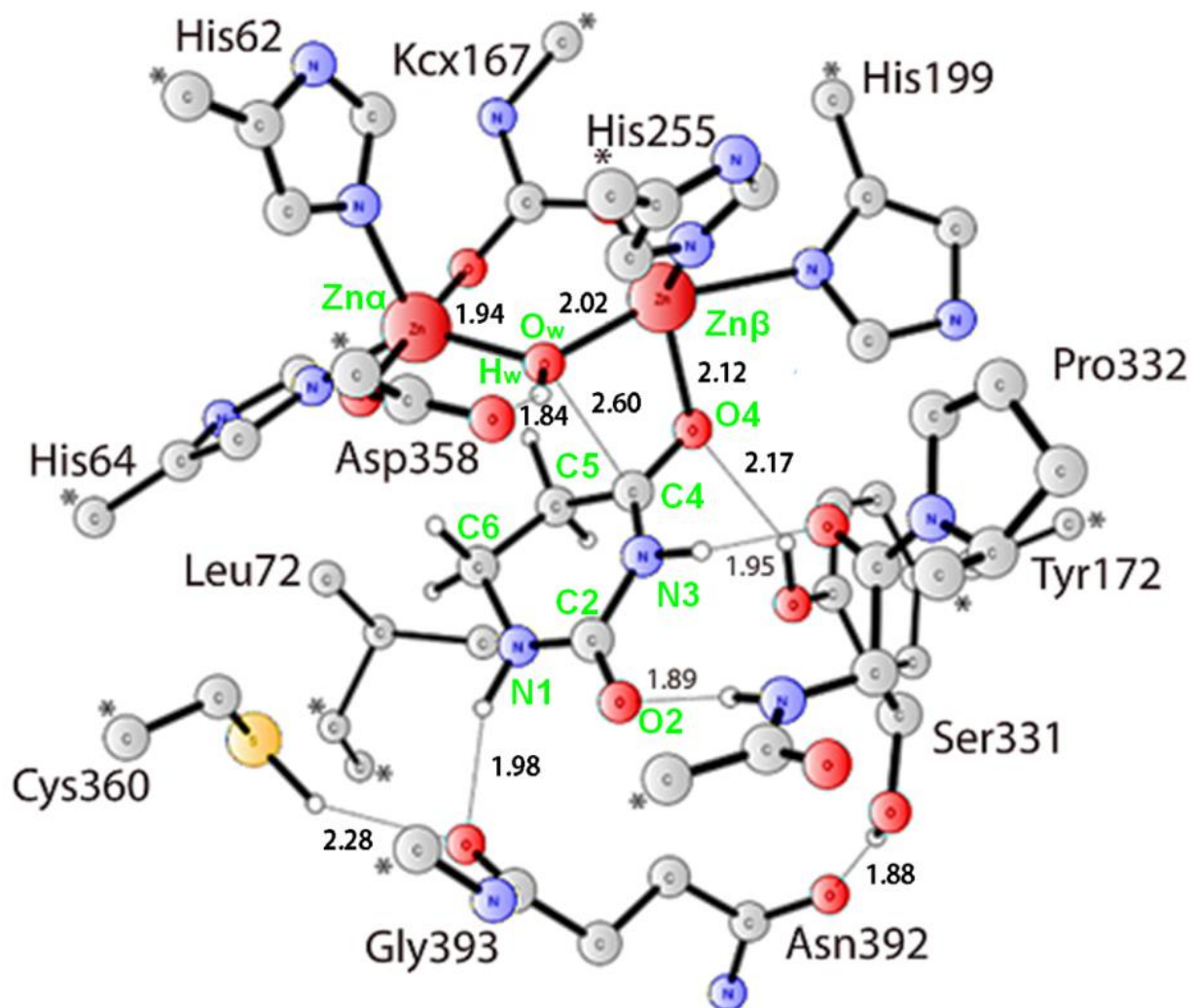


Figure 3. Calculated energy profiles for hydrolysis of substrate dihydrouracil (DHU) by SkDHPase in **Model II**. cluster + CPCM (—) and cluster (····). Energies were obtained at B3LYP/6-311+G(2d,2p) + CPCM($\epsilon=4$) level of theory



ES-II

Figure 4. Optimized structure for the large model of *SkDHPase* active-site with substrate DHU bound (**ES-II**). For clarity, most hydrogen atoms have been omitted in the figure. Fixed atoms are shown by an asterisk.

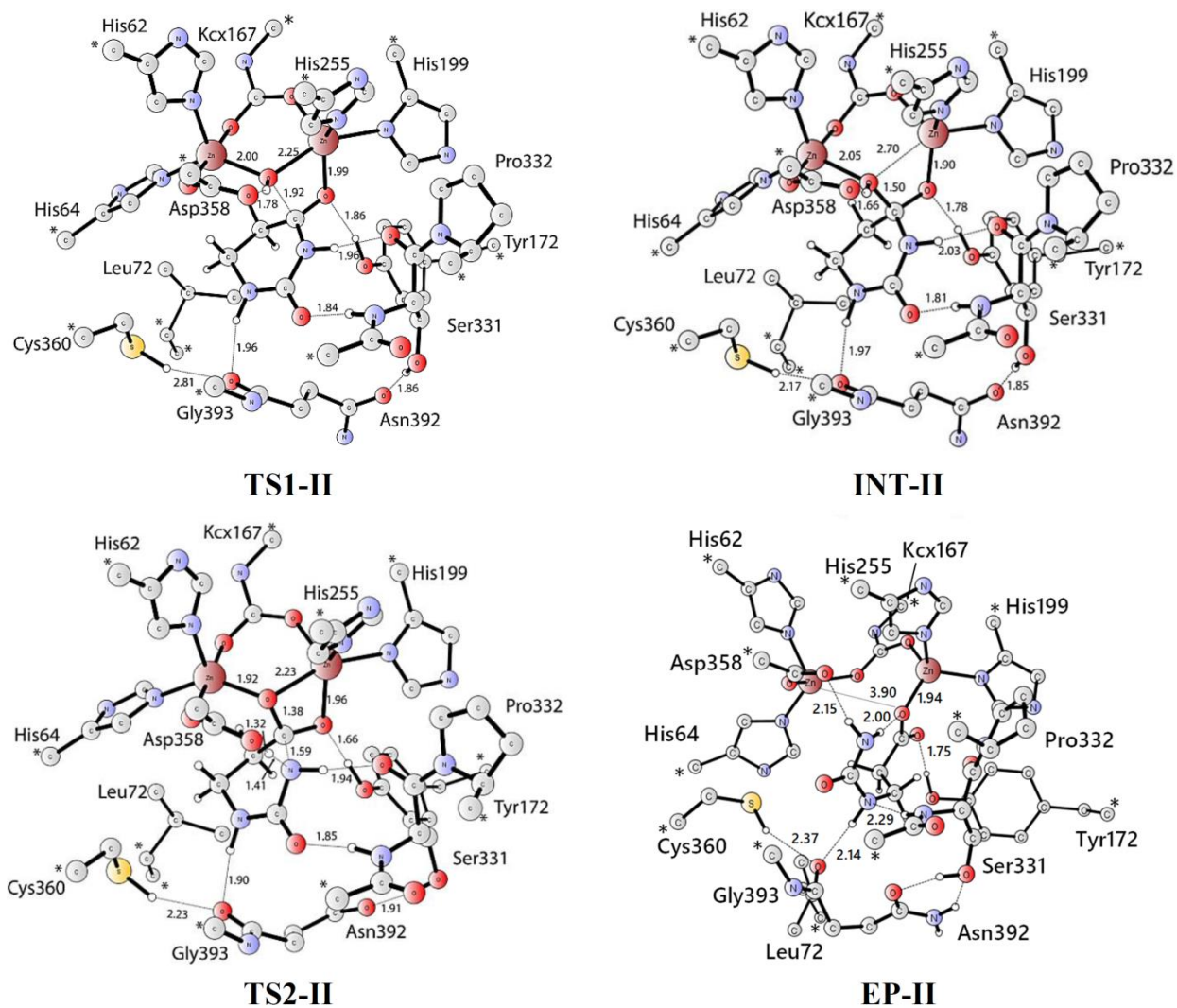


Figure 5. Optimized stationary points (**TS1**, **INT**, **TS2**, and **EP**) for the hydrolysis of DHU in **Model II**. Distances are in Å. Fixed atoms are shown by an asterisk. For clarity, only selected hydrogen atoms are shown. The product NCβA (**EP-II**) is not the same as 2FVM

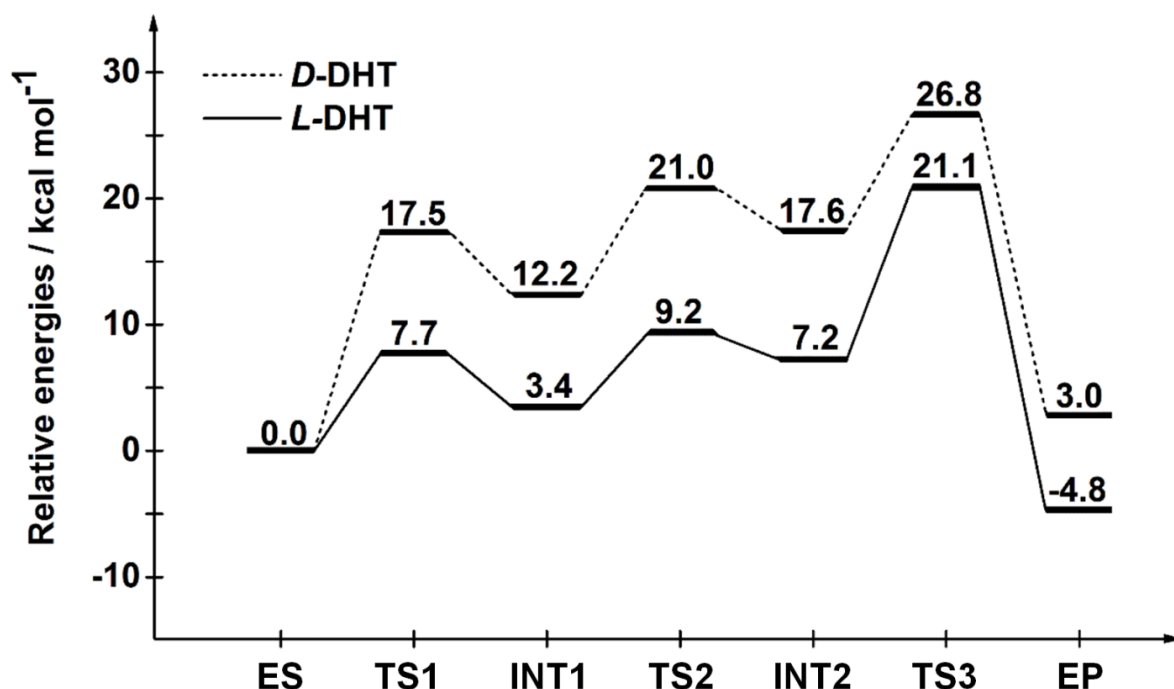


Figure 6. Calculated energy profile for the hydrolysis of DHT in either *L*- or *D*- configurations (*L*-DHT and *D*-DHT) in **Model II**. Energies were obtained at B3LYP/6-311+G(2d,2p) + CPCM($\epsilon=4$) level of theory

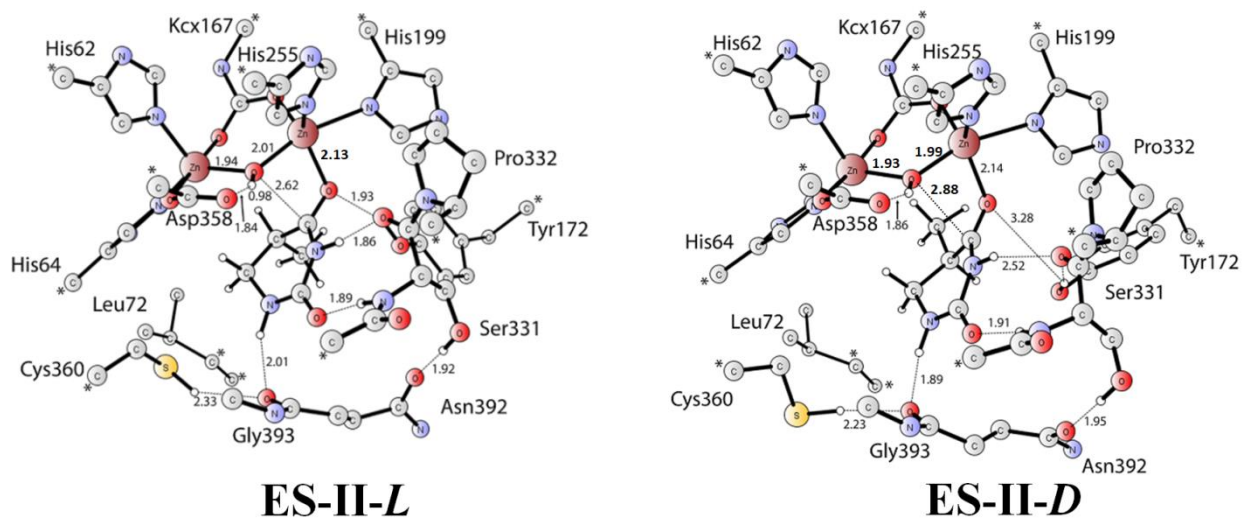


Figure 7. Optimized structures of the *SkDHPase*-DHT complexes (**ES**) in *L*- and *D*-configuration in **Model II**. Distances are given in Å. Fixed atoms are shown by an asterisk.

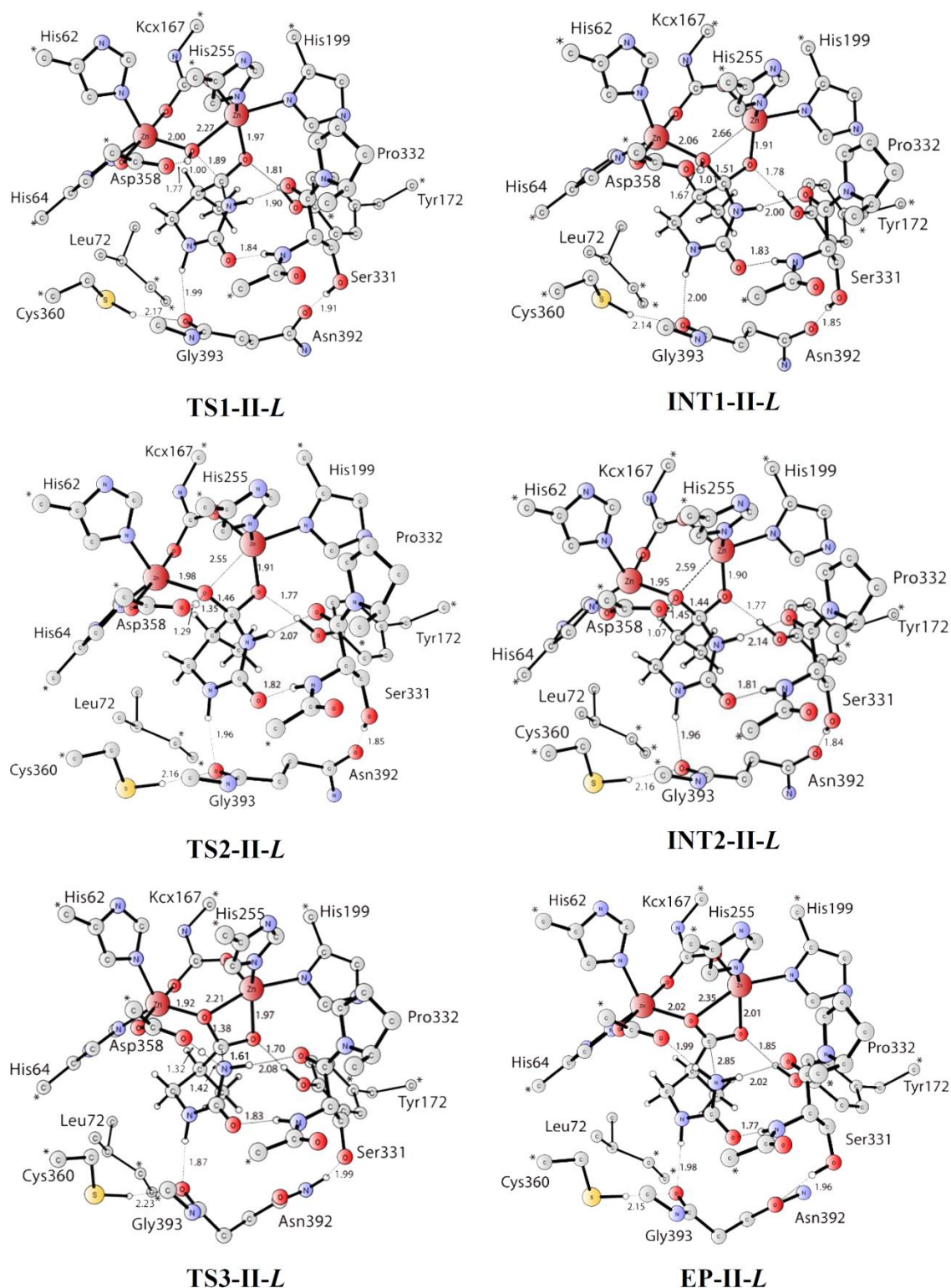


Figure 8. Optimized stationary points for the hydrolysis of *L*-DHT in **Model II**. Distances are in Å. Fixed atoms are shown by an asterisk.

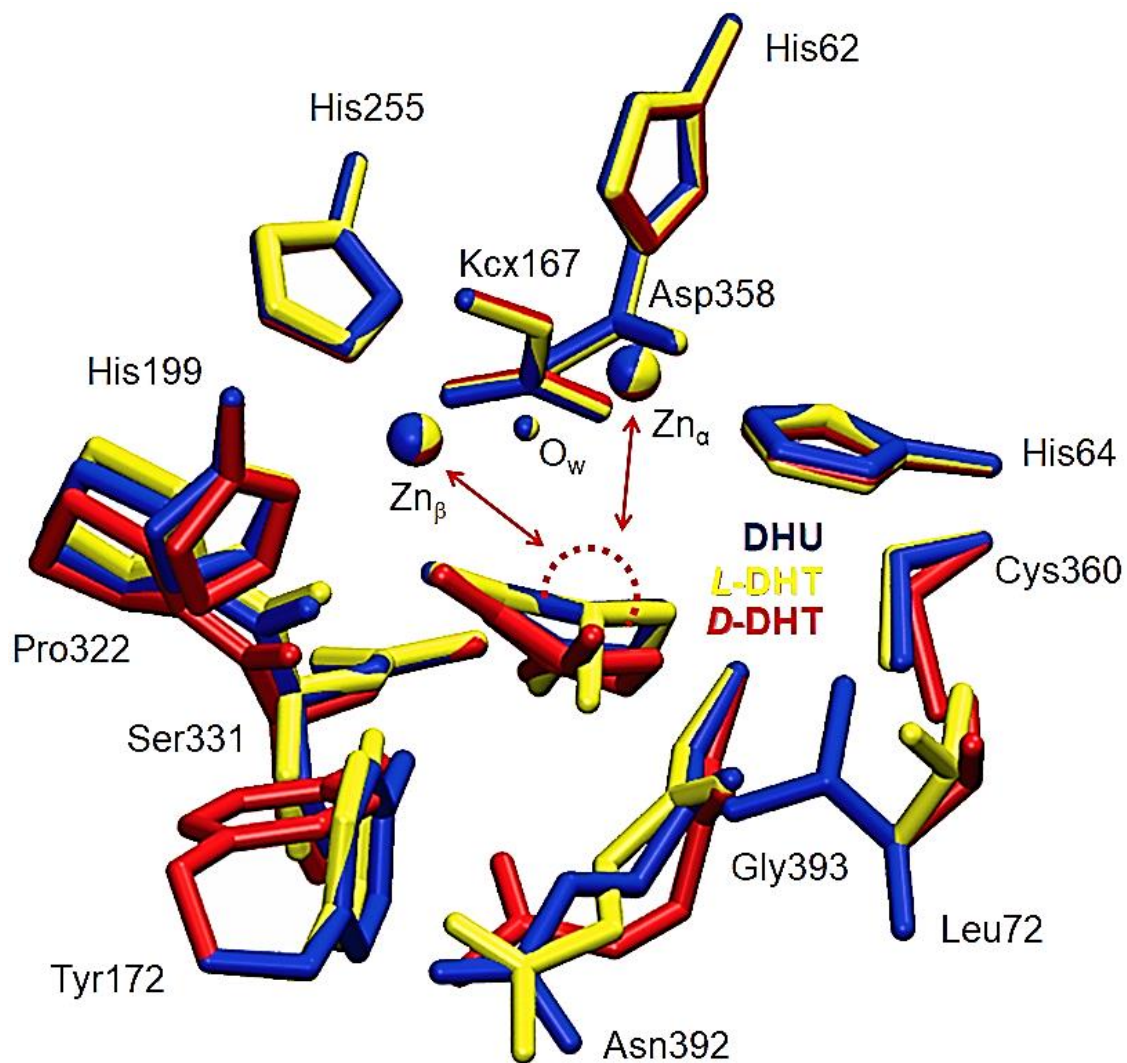
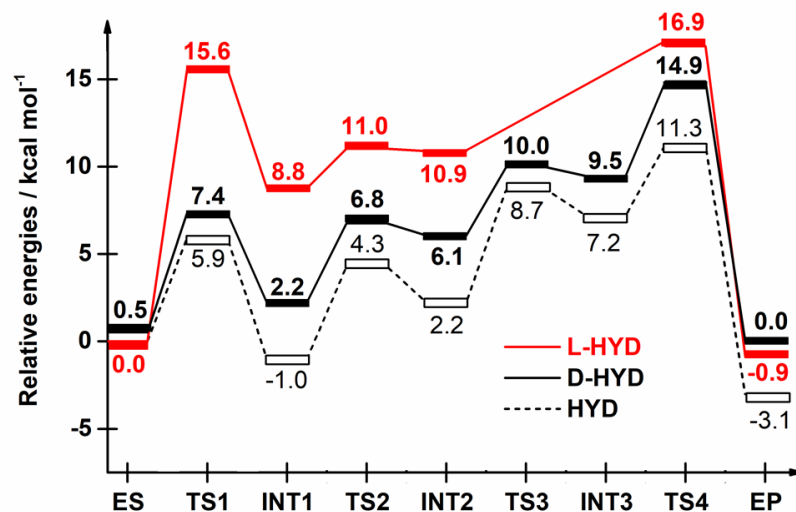
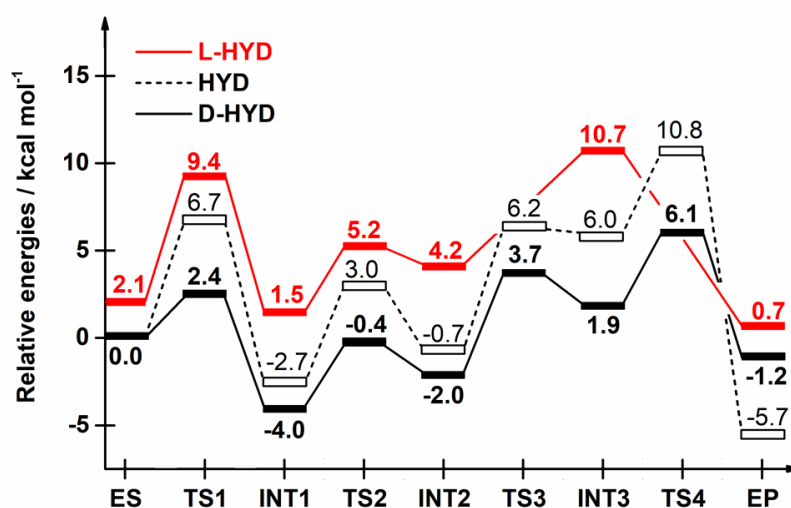


Figure 9. Overlay of the optimized structures of the **ES-II** with DHU (blue), *L*-DHT (yellow) and *D*-DHT (red) as substrates.

SkDHPase catalyzed hydrolysis of hydantoid and methyl hydantoid



(a) Without Edisp



(b) With Edisp

Figure 10. Calculated energy profile for the hydrolysis of HYD by *SkDHPase* in **Model II**. Values (in $\text{kcal}\cdot\text{mol}^{-1}$) are obtained at B3LYP/6-311+G(2d,2p)//B3LYP/6-31G(d) + ZPE correction + CPCM($\epsilon=4$)

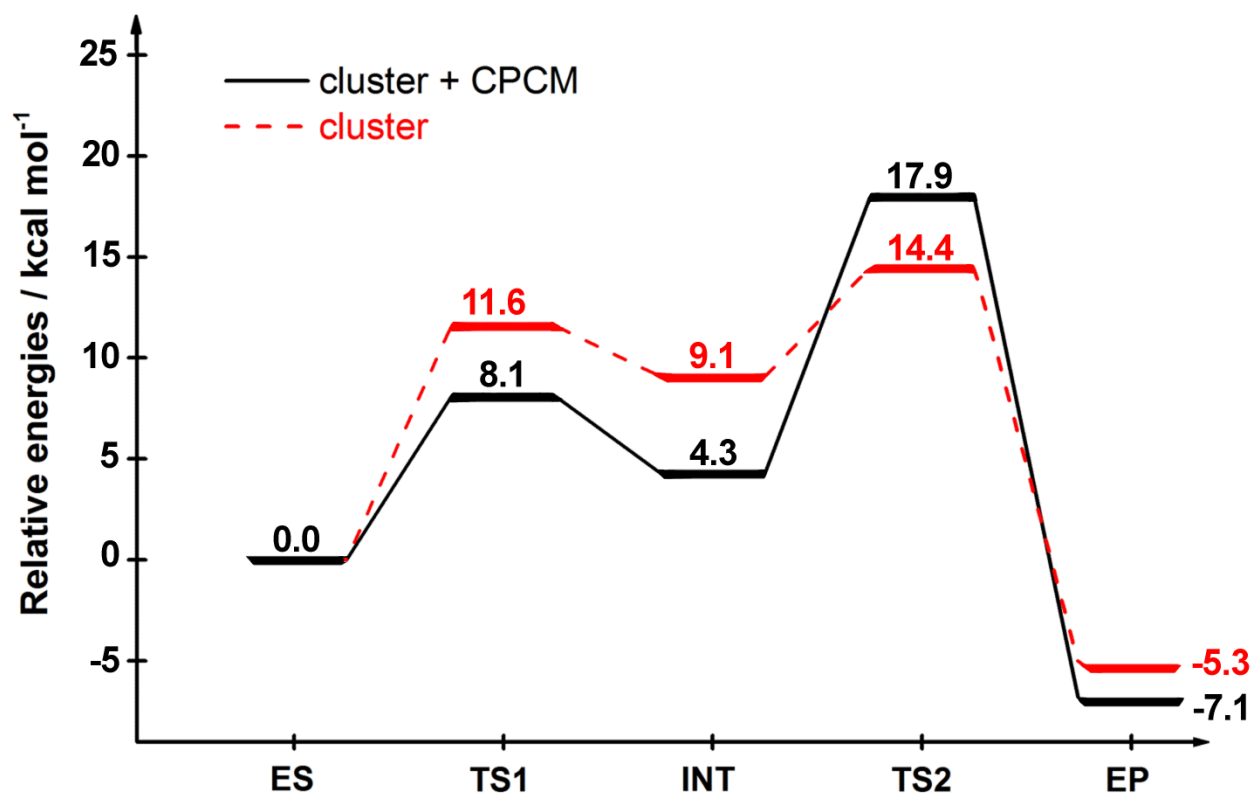


Figure 11. Calculated energy profile for the hydrolysis of DHO by *SkDHPase* in **Model II**. The energies were obtained at B3LYP/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d) level.

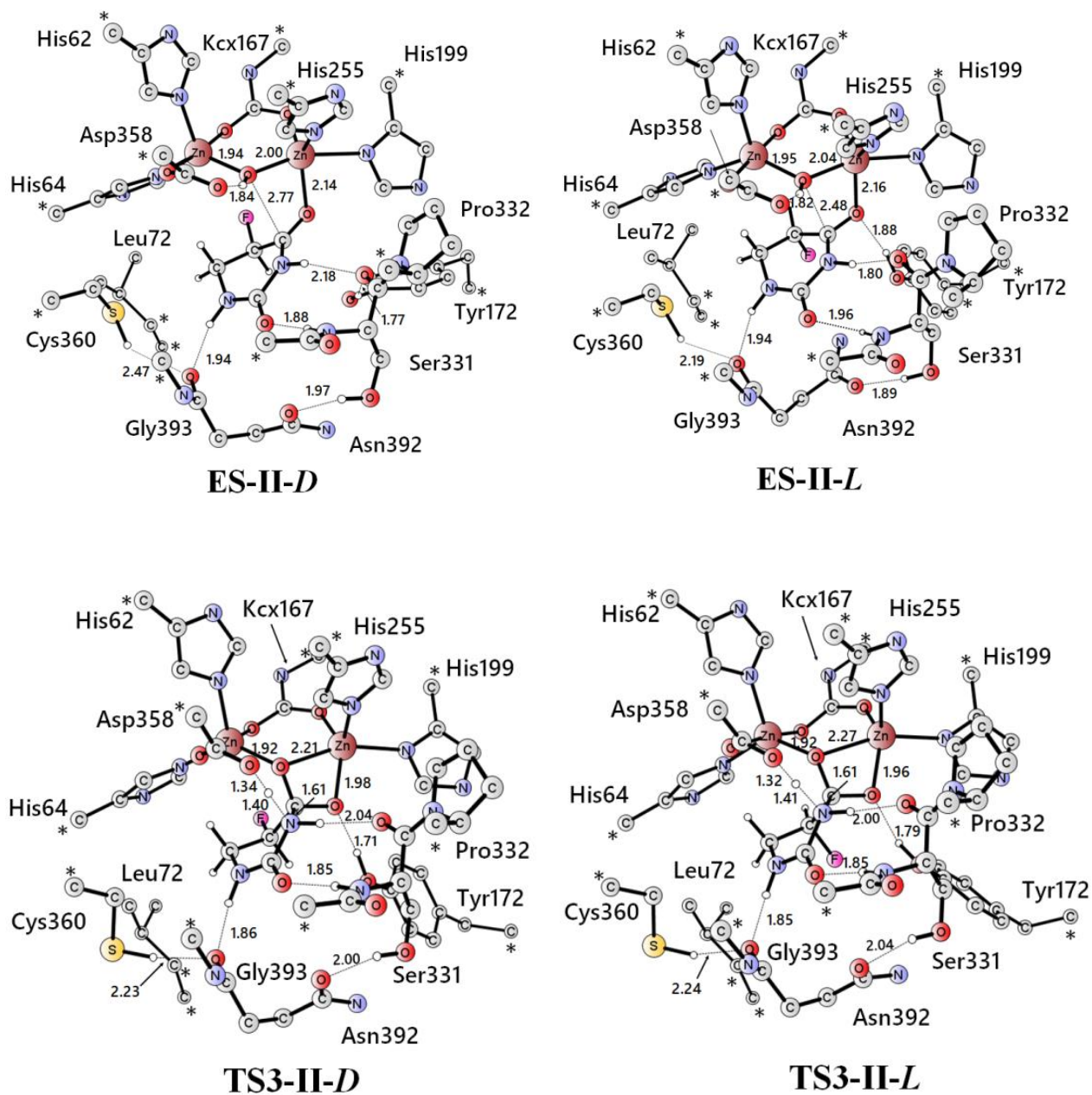


Figure 12. Optimized structures of the SkDHPase-5FUH2 complexes at ES and TS3 in *L*- and *D*-configurations in **Model II**. Distances are given in Å. Fixed atoms are shown by an asterisk.

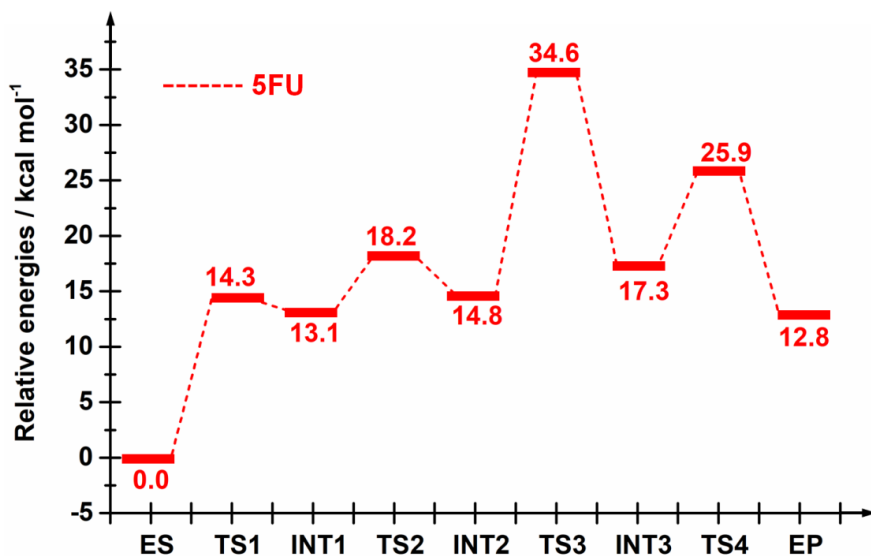


Figure 13. Calculated energy profile for the hydrolysis of 5-fluorouracil (5FU) in **Model II**. Relative energies were obtained at B3LYP/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d) level.

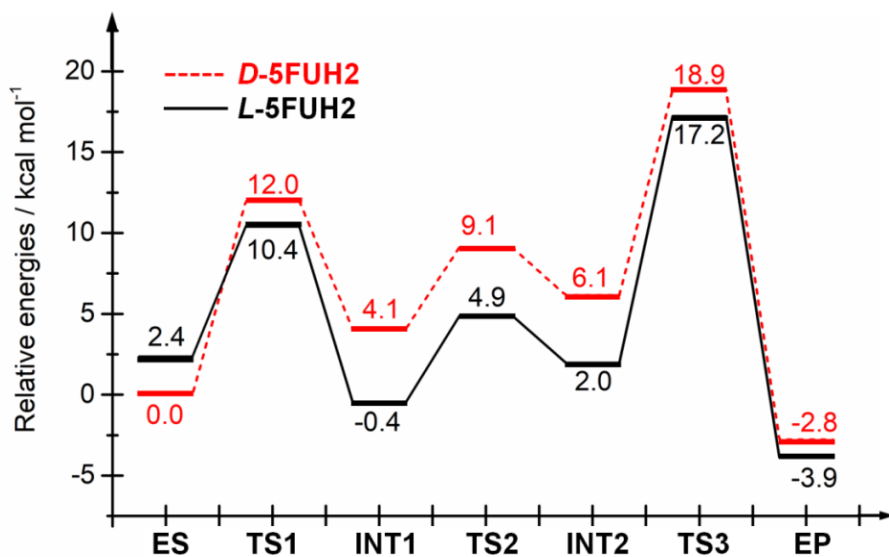


Figure 14. Calculated energy profile for the hydrolysis of dihydro-5-fluorouracil in either *L*- or *D*- configurations (*L*-5FUH2 and *D*-5FUH2) in **Model II**. Relative energies were obtained at B3LYP/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d) level.

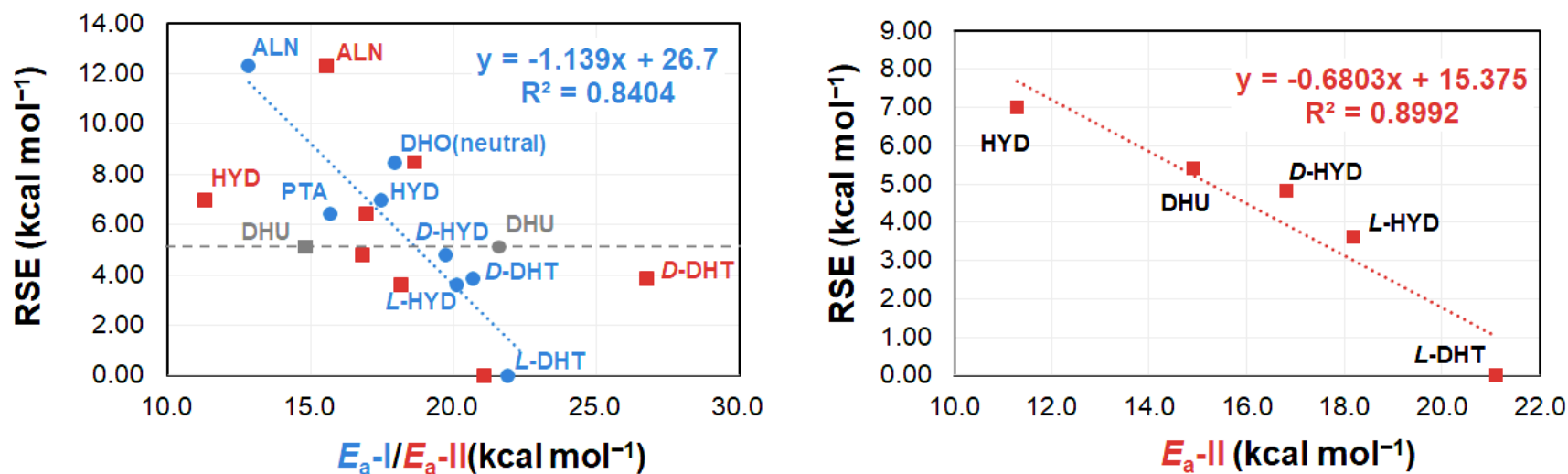


Figure 15. Relationship between the activation barriers ($E_a\text{-I}$ and $E_a\text{-II}$) and the ring-strain energies (RSE) as calculated for different substrates shown in Figure 1. The negatively charged DHO is excluded from the plot due to the large deviation (RSE = -20.0 kcal mol⁻¹ and $\Delta = -12.2$ kcal mol⁻¹).

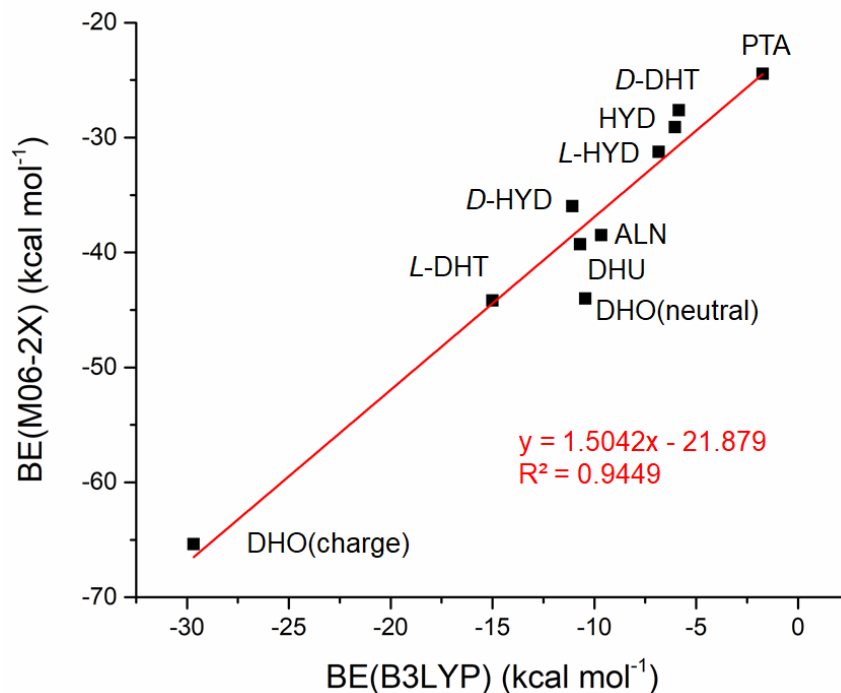


Figure 16. Correlation between the B3LYP and M06-2X binding energies (BE, kcal mol⁻¹) for different enzyme-substrate structures of **Model-II**. These BE values were obtained from the single-point energy computed on the optimized geometries [at B3LYP/6-31G(d,p) level] using B3LYP and M06-2X functional and 6-311+G(2d,2p) basis set, plus CPCM solvation model ($\epsilon = 4$).

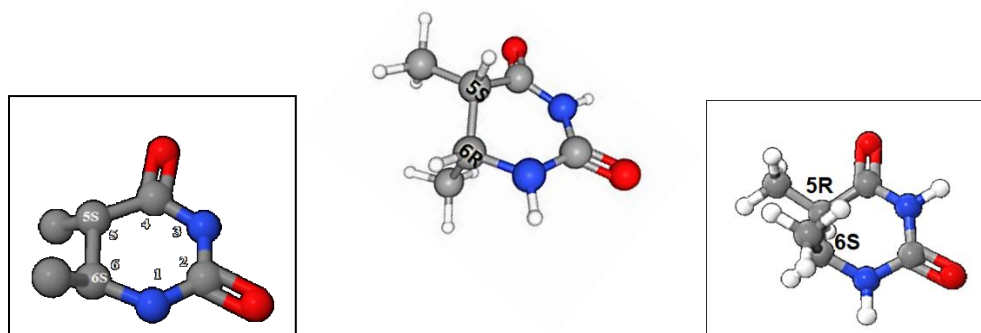


Figure 17. Three possible configurations (SS, SR, RS) for the 5,6-dimethyl dihydrouracil substrate.

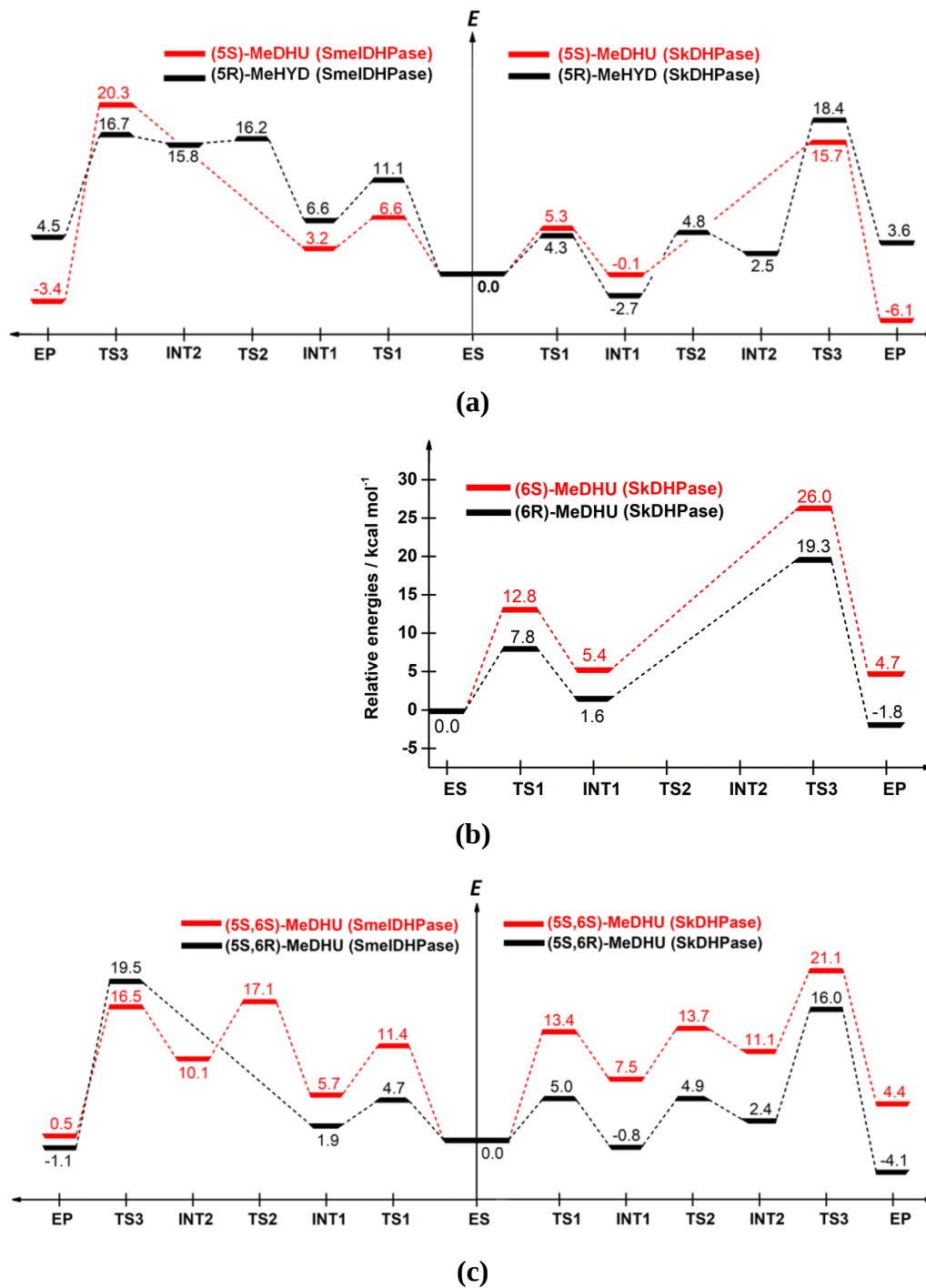


Figure 18. Calculated energy profiles for stereospecific hydrolysis of (5R)-methylhydantoid and (5S)-methyldihydrouacil **(a)**, 6-methyldihydrouacil (*S* and *R* isomers) **(b)** and 5,6-dimethyldihydrouacil (*SS* and *RR* isomers) **(c)** catalyzed by *SkDHPase* and *SmeIDHPase*, as obtained at B3LYP-D3/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d,p) level.

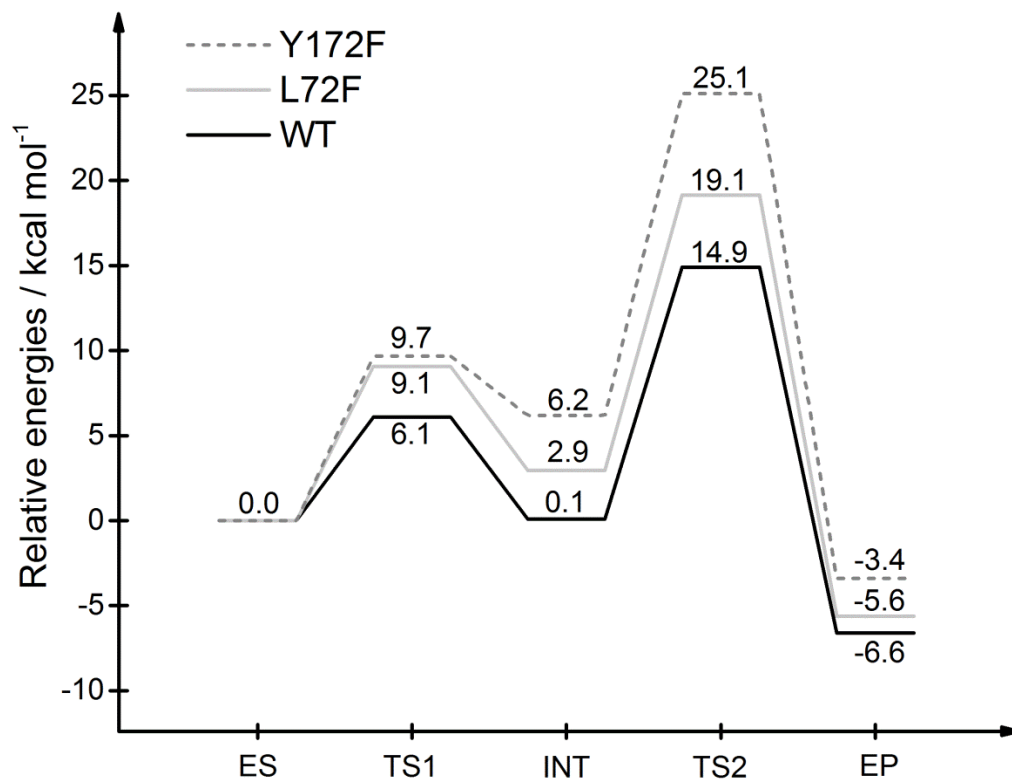


Figure 19. Calculated energy profile for the hydrolysis of DHU by WT, Leu72Phe and Tyr172Phe *SkDHPase* in **Model II**. Energies were obtained at the B3LYP/6-311+G(2d,2p) + CPCM($\epsilon=4$) level of theory.

Table 1 Values ^a of binding interaction energies (BE) ^b, activation energies (E_a) of WT models (I and II) and ring strain energies (RSE) ^c for different substrates

Substrate	BE			E_a			RSE		
	WT	Tyr172Phe	Leu72Phe	I	II	ΔE_a	$\varepsilon = 1$	$\varepsilon = 4$	Δ^d
DHU	-10.8	-11.9	-11.0	21.7	14.9	-6.8	-63.7	-37.6	5.4
<i>L</i> -DHT	-15.0	-13.6	-11.8	21.9	21.1	-0.8	-58.9	-32.2	0.0
<i>D</i> -DHT	-5.8	-11.9	-9.8	20.7	26.8	+6.1	-62.1	-36.1	3.9
HYD	-6.0	-10.8	-11.5	17.5	11.3	-6.2	-65.2	-39.2	7.0
<i>L</i> -HYD	-6.8	-9.2	-6.7	20.1	16.9	-3.2	-62.2	-35.8	3.6
<i>D</i> -HYD	-11.1	-11.0	-8.9	19.8	14.4	-5.4	-63.5	-37.0	4.8
DHO (neutral)	-10.4	n/a	n/a	18.0	18.4	+0.4	—	-40.7	8.5
DHO (charged)	-29.7	n/a	n/a	22.3	17.9	-4.4	—	-20.0	-12.2

^a Values (in kcal mol⁻¹) were obtained at B3LYP/6-311+G(2d,2p) + CPCM(ε =4) level of theory

^b Obtained from the interaction energies according to the equation: BE = $E(\text{complex}) - E(\text{receptor}) - E(\text{substrate})$

^c Obtained from the energy difference between the linear product and the cyclic substrate according to the equation: cyclic substrate + OH⁻ → linear product + RSE

^d Relative ring strain energies as calculated from the difference of all RSE values with $\varepsilon = 4$ relative to the lowest RSE (-32.2 kcal mol⁻¹ in *L*-DHT)

Table 2. Activation barriers and reaction energies ($\Delta E_{\text{TS}}^{\ddagger}$ and ΔE_{R} , in kcal mol⁻¹)^a for hydrolysis of (5*R*)-methylhydantoid and (5*S*)-methyldihydouracil, 6-methyldihydouracil (*S* and *R* isomers), and 5,6-dimethyldihydouracil (*SS* and *RR* isomers) by *SkDHPase* and *SmeIDHPase*, obtained using the B3LYP-D3 and B3LYP (in parenthesis) methods.

Substrate	<i>SkDHPase</i>		<i>SmeIDHPase</i>		Expt. ^b	Error
	$\Delta E_{\text{TS}}^{\ddagger}$	ΔE_{R}	$\Delta E_{\text{TS}}^{\ddagger}$	ΔE_{R}		
HYD	10.8(11.3)	-5.7(-3.1)	19.1(20.7)	+5.8(+2.9)	16.8	2.3
DHU	- (14.9)	- (-6.6)	n/a	n/a	17.9	-
(5 <i>R</i>)-MeHYD	18.4(24.4)	+3.6(+5.4)	16.7(23.1)	+4.5(+6.0)	15.1	1.6
(5 <i>S</i>)-MeDHU	15.8(20.5)	-6.1(-4.9)	20.3(24.1)	-3.4(-2.7)	18.4	1.9
(6 <i>S</i>)-MeDHU	26.0(33.9)	+4.7(+9.0)	11.4	+5.2	17.1	-
(6 <i>R</i>)-MeDHU	19.3(27.0)	-1.8(+2.2)	4.6	-2.5		-
(5 <i>S</i> ,6 <i>S</i>)-MeDHU	21.1(27.5)	+4.4(+11.1)	17.1(18.5)	+0.5(+2.0)	n/a	n/a
(5 <i>S</i> ,6 <i>R</i>)-MeDHU	16.0(28.1)	-4.1(+4.2)	19.5(24.8)	-1.1(-4.1)	n/a	n/a

^a All energies were calculated at the B3LYP-D3/6-311+G(2d,2p)(CPCM, $\epsilon=4$)/B3LYP/6-31G(d,p) level. Energies without dispersion correction were also included in parenthesis

^b Experimental barriers were estimated using the k_{cat} values reported in Ref. 12 and the transition state theory.

Table 3 Binding interaction energies (BE, kcal/mol) ^a for different substrates in the active site pocket of *SkDHPase* and *SmelDHPase* as obtained using B3LYP and B3LYP-D3 methods

Substrate	B3LYP			B3LYP-D3		
	<i>SkDHPase</i>	<i>SmelDHPase</i>	Δ	<i>SkDHPase</i>	<i>SmelDHPase</i>	Δ
HYD	−6.0	−11.3	5.4	−30.2	−39.1	8.9
DHU	−10.8	n/a		−40.6	n/a	
(5 <i>R</i>)-MeHYD	+10.9	−7.8	18.7	−16.4	−37.4	21.0
(5 <i>S</i>)-MeDHU	+8.5	−6.2	14.7	−22.4	−40.8	18.4
(6 <i>S</i>)-MeDHU	+12.3	−5.9	18.2	−15.0	−39.2	24.2
(6 <i>R</i>)-MeDHU	+11.8	−4.5	16.3	−15.3	−39.0	23.7
(5 <i>S</i> ,6 <i>S</i>)-MeDHU	−11.1	−7.3	3.8	−39.6	−43.9	4.3
(5 <i>S</i> ,6 <i>R</i>)-MeDHU	−13.1	−8.5	4.6	−46.5	−45.2	1.3

Table 4. Effects of Tyr172Phe and Leu72Phe *SkDHPase* mutants on the binding energies (BE, kcal/mol) and activation energies E_a (model II). Energies were obtained at B3LYP/6-311+G(2d,2p) + CPCM(ϵ =4) level of theory

Substrate	BE			E_a -II		
	WT	Tyr172Phe	Leu72Phe	WT	Tyr172Phe	Leu72Phe
DHU	−10.8	−11.9	−11.0	14.9	25.1	19.1
<i>L</i> -DHT	−15.0	−13.6	−11.8	21.1	33.1	27.2
<i>D</i> -DHT	−5.8	−11.9	−9.8	26.8	34.4	33.1
HYD	−6.0	−10.8	−11.5	11.3	25.5	16.2
<i>L</i> -MEHYD	−6.8	−9.2	−6.7	16.9	26.7	21.8
<i>D</i> -MEHYD (D-selective)	−11.1	−11.0	−8.9	14.4	23.3	17.2

Chapter III

Application of ONIOM QM/QM' Method for Understanding Protein-Ligand Interaction in Carbohydrate-Active Enzyme for Drug Discovery

Huang, M., S. Pengthaisong, R. Charoenwattanasatien, N. Thinkumrob, **J. Jitonnorn*** and J. R. Ketudat Cairns* (2022). "Systematic Functional and Computational Analysis of Glucose-Binding Residues in Glycoside Hydrolase Family GH116." **Catalysts** 12(3): 343. (Q2, IF 2020 =4.146, *Co-corresponding author).

COMPUTATIONAL ANALYSIS OF GLUCOSE-BINDING RESIDUES IN GLYCOSIDE HYDROLASE FAMILY GH116

INTRODUCTION

Enzymes are particularly malleable catalysts due to their ability to change specific functional groups and the catalytic environment by the genetic engineering of the amino acid sequence. An enzyme's catalytic reaction often occurs in a buried groove or pocket-like active site, where the functional groups of the surrounding amino acid residues build a unique environment that promotes the highly specific and efficient catalysis of reactions. In the active site and the surrounding region outside this site, a relatively small number of specific amino acid residues are involved in substrate binding, and few of these act directly in the catalytic reaction (Holliday, Bartlett et al. 2005). The roles of amino acid residues during catalysis have been delegated into seven basic classes—stabilization, steric roles, activation, proton shuttling, hydrogen shuttling, electron shuttling residues, and covalent catalysis (Holliday, Mitchell et al. 2009).

β -Glucosidases (β -D-glucopyranoside glucohydrolases, E.C. 3.2.1.21) are enzymes that release the nonreducing terminal β -D-glucosyl residue from glucoconjugates, including glucosides, 1-O-glucosyl esters, and oligosaccharides, by glycosidic bond hydrolysis (Ketudat Cairns, Mahong et al. 2015). β -Glucosidases have been categorized into the protein sequence-based glycoside hydrolase families GH1, GH2, GH3, GH5, GH16, GH30, GH39, and GH116 (<http://www.cazy.org>, accessed on 14 February, 2022) (Lombard, Golaconda Ramulu et al. 2014), and the structure and catalytic mechanism are generally conserved within each family. Their specificity toward different substrates varies, depending on the enzyme and its biological function.

Hydrolysis of the glycosidic bond can occur with either inversion or retention of the anomeric configuration, and this property is generally conserved within one enzyme family (McCarter and Stephen Withers 1994, Zechel and Withers 2000). Most characterized β -glucosidases catalyze hydrolysis via a two-step retaining mechanism, in which the first step entails the departure of the aglycone and the formation of a covalent intermediate with the enzymatic nucleophile residue, while the second step is a reversal of this process, with water serving as the enzyme-displacing nucleophile (Burmeister, Cottaz et al. 1997, McCarthy, Uzelac et al. 2004, Kurakata, Uechi et al. 2008). A catalytic acid/base residue serves as an acid to protonate the glycosidic oxygen as the aglycone departs in the enzyme glycosylation step and as a base to extract a proton from water in the enzyme deglycosylation step. In each step, the glucose is thought to pass through an oxocarbenium cation at or near the transition state, which requires distortion to a conformation that allows sp^2 hybridization at C1 and O5, typically a 4H_3 half-chair or closely related 4E envelope in β -glucosidases (Vocadlo and Davies 2008). In most β -glucosidases, the formation of the transition state shape appears to be primed by the distortion of the pyranose ring to a 1S_3 conformation. This induced fit of the substrate implies that the surrounding residues, which generally form tight hydrogen bonds with the hydroxyl groups and aromatic-sugar stacking interactions with the ring hydrogens of the β -D-glucopyranose, are essential to the transition state formation in catalysis, rather than simply serving to bind the substrate. It has been noted that the electrostatic stabilization of the oxocarbenium ion-like transition state by nearby residues also

contributes to catalysis (Chuenchor, Pengthaisong et al. 2011, Geronimo, Payne et al. 2018). Indeed, some previous studies have shown that the contributions of enzyme active-site residues not directly involved in bond making and breaking could be as essential to the enzyme's function as catalytic residues (Dopitová, Mazura et al. 2008, Pengthaisong, Chen et al. 2012, Zhang, Wang et al. 2015). Nonetheless, systematic kinetic analysis of the contributions of the active-site residues of β -glucosidases is rarely carried out, due to the low activities of such mutants (Charoenwattanasatien, Pengthaisong et al. 2016).

The enzyme used in this work, *Thermoanaerobacterium xylanolyticum* GH116 (TxGH116), is a thermostable β -glucosidase which consists of an N-terminal domain formed by a two-sheet β -sandwich and a C-terminal $(\alpha/\alpha)_6$ solenoid domain containing the slot-like active site [17]. As the first family GH116 enzyme to have its X-ray crystal structure determined, TxGH116 has helped to elucidate the molecular basis of pathogenic mutations in the human GH116 member, GBA2 glucosylceramidase, that lead to hereditary ataxias and paraplegias. Aside from its mutations resulting in these rare genetic disorders, the over- and under-production of GBA2 have been shown to affect cellular sphingolipid balance and the depletion of GBA2 has also been shown to ameliorate some of the effects of loss of lysosomal glucosylcerebrosidase (GBA) activity, suggesting the inhibitors of GBA2 may have therapeutic applications. High-resolution crystal structures of TxGH116 in complex with glucose and inhibitors allow for the identification of the active-site residues involved in substrate binding and catalysis, which are completely conserved with human GBA2 (Charoenwattanasatien, Pengthaisong et al. 2016). Aside from its role as a model structure for human GBA2, mutants of the TxGH116 enzyme have been used to efficiently generate glucosylazides for the production of α - and β -glucosidase inhibitors, which have potential applications in treating human diseases.

Our previous work verified that TxGH116 utilizes a retaining mechanism by identifying the initial hydrolysis product as β -D-glucose by monitoring the reaction time course by NMR spectroscopy (Charoenwattanasatien, Pengthaisong et al. 2016). The identities of E441 as the catalytic nucleophile and D593 as the catalytic acid/base were proven by the mutation of these residues to alanine, followed by chemical rescue using small nucleophiles. TxGH116 is thought to catalyze hydrolysis via the usual β -glucosidase conformational itinerary from 4C_1 to 1S_3 to 4H_3 to 4C_1 in the glycosylation step, but no distortion of the pyranose ring has been observed for glucose and glucose-like inhibitors in the active site complexes. Glucoimidazole binds in the transition state-like 4H_3 conformation, but this is the low energy conformation for this inhibitor, so it does not provide strong evidence for distortion toward this transition state conformation during catalysis (Charoenwattanasatien, Pengthaisong et al. 2016). The yet-to-be-verified transition state and the unusual orientation of the catalytic acid/base compared to β -glucosidases from other families make the roles of GH116 active-site residues particularly worthy of investigation.

Although the TxGH116 catalytic acid/base and nucleophile residues were analyzed (Charoenwattanasatien, Pengthaisong et al. 2016), the importance and roles of other residues within the TxGH116 and GH116 family active-site pocket remain to be clarified. In this work, the energetic effects of their mutations were calculated to further elucidate their roles. The insight into the relative importance of active site residues in substrate and catalysis facilitates the modification of TxGH116 β -glucosidase for improved application, the design of human GBA2 inhibitors, and a better understanding of human disorders caused by mutations in GBA2.

OBJECTIVES

- (1) To elucidate the mutation-induced effect on glucose-binding residues
- (2) To identify the important residues that contributed to the binding of glucose

METHOD

Our own two-layered N-layered integrated molecular orbital and molecular mechanics (ONIOM) (Chung, Sameera et al. 2015) calculations (ONIOM2(B3LYP/6-31G(d,p):PM3)) were performed to examine the mutation-induced effects on the binding affinity and stability of glucose ligand in the TxGH116 active site. Here, eleven protein–ligand complexes for TxGH116 wild type and mutants (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K) were modeled and optimized using the X-ray structure of the TxGH116-glucose complex (PDB: 5BX5 (Charoenwattanasatien, Pengthaisong et al. 2016)) because of their potential hydrogen bond interactions. The ONIOM2 calculations were set up in a similar manner as in the previously published papers (Saen-oon, Kuno et al. 2005, Boonsri, Kuno et al. 2011), and were performed on the TxGH116 active site residues in a sphere of 5 Å around the glucose. In this complex, the high-level region (region A in **Figure 1**) comprises the atoms in the region of the complex, including the glucose and the residues E441 (the catalytic nucleophile), H507, D452, D508, Y523, T591, D593 (the catalytic acid/base), E777, R786, and R792, and were treated to a high-level calculation, in this case, B3LYP/6-31G(d,p). These residues are, from the crystallographic point of view, involved mainly in hydrogen-bonding interactions with the glucose (Charoenwattanasatien, Pengthaisong et al. 2016). The low-level region (region B in **Figure 1**) includes all the atoms not previously selected in the high-level region, which were treated at the semi-empirical level PM3. The ONIOM2 optimized structures, as well as their binding energies, BE, were obtained from the extrapolated energy of the ONIOM2 approach, which is defined as follows:

$$\begin{aligned} \text{BE(ONIOM2)} &= E[\text{Cpx}] - E[\text{P}] - E[\text{L}] \\ &= \Delta E (\text{high, A}) + [\Delta E (\text{low, AB}) - \Delta E (\text{low, A})] \end{aligned}$$

where $E[\text{Cpx}]$ is the total energy of the protein–ligand complex, $E[\text{P}]$ is the total energy of the protein pocket, $E[\text{L}]$ is the total energy of the ligand (i.e., glucose), $\Delta E (\text{high, A})$ is the energy of the region A calculated using a high level of calculations, $\Delta E (\text{low, AB})$ is the energy of the whole model system (regions A and B in **Figure 1**) calculated using a low level of calculations, and $\Delta E (\text{low, A})$ is the energy of the region A calculated using a low level of calculations.

To further identify the important contribution of individual residues in the TxGH116 active site, we have performed a particular interaction analysis on those mutated amino acids by calculating the interaction energies, $\text{IE}(\text{L} + \text{X}_i)$, between ligand (glucose) and individual residues, X_i , at the B3LYP/6-31G(d,p) level of theory using the ONIOM geometry described above. Counterpoise correction for the basis set superposition error (BSSE-CP) was also calculated to correct the interaction energy. The total interaction energy, IE, can be expressed as follows:

$$\text{IE}(\text{L} + \text{X}_i) = E(\text{L} + \text{X}_i) - E(\text{L}) - E(\text{X}_i)$$

where $E(L)$ and $E(X_i)$ are the energies of the ligand (i.e., glucose) and each individual residue, respectively. All computations were performed using the Gaussian09 program (Frisch, Trucks et al. 2009).

RESULTS AND DISCUSSION

To elucidate the mutation-induced effect on glucose-binding residues, particular interaction analysis, which describes the contribution of each residue in the total binding energy, was conducted for wild type and 10 mutant models (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, R786K), based on the ONIOM method [36–38]. The ONIOM-optimized structure of wild type TxGH116-glucose complex and the particular interaction energy (IE) between glucose and the TxGH116 binding site for the wild type and mutants are illustrated in **Figures 1** and **2**. The BE values of all complexes are plotted in **Figure 3**, which are in the ranges of -94.8 to -131.7 kcal mol⁻¹ for the mutants and -134.1 kcal mol⁻¹ for the wild type. This result clearly shows that all mutations decreased the binding affinity of glucose to the TxGH116 active site, suggesting the important roles of the TxGH116 binding residues.

The particular interaction energies were calculated to identify the important residues that contributed to the binding of glucose. They can also be used as a simple and qualitative indicator for understanding the mutation effect on the electronic structure and protein–ligand interaction. The interaction energies for the individual residues of both the wild type and mutants are shown in **Figure 2**. It was calculated that six residues, E441 (nucleophile), D452, H507, D593 (acid/base), E777, and R786, are major (electrostatic) contributors for glucose binding to TxGH116, as evidenced by the large negative IE values ($|IE| > 15$ kcal mol⁻¹). Further protein–ligand analysis illustrated in **Figure 4** also showed that the hydrogen bond and van der Waals interactions were the dominant stabilization factors of the non-bonded interactions between the glucose ligand and the TxGH116 active site.

The geometries of the wild type and the mutated residues (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K) and their relevant H-bond interactions with the C2OH, C3OH, C4OH, or C6OH of the glucose are visualized in **Figures 5–7**. Noticeably, the mutations caused significant changes in the binding energy contribution, for which TxGH116 D452 and E777 mutations reduced the favorable contribution of the residues E441, D593, E777, and R786, and increased the favorable contributions of D452 and H507 (**Figures 1** and **2**). Furthermore, the mutations also resulted in a reduction in the total ONIOM binding energy compared to the wild type, as seen in **Figure 3**, with D452N, E777Q, and R786K exhibiting a small effect (~ 2 – 7 kcal mol⁻¹) compared to the alanine mutations (~ 10 – 40 kcal mol⁻¹).

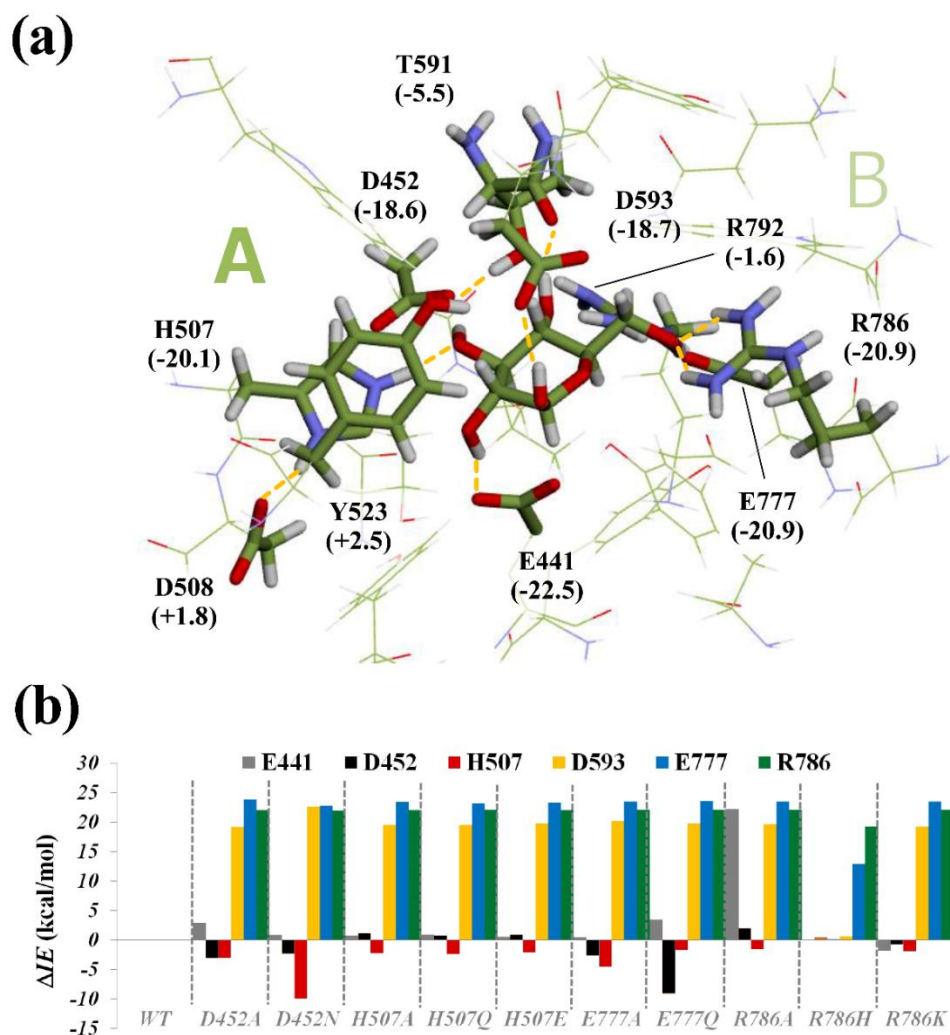


Figure 1. ONIOM2-optimized structure of wild type (WT) TxGH116-glucose complex (a) and change in particular interaction energy (ΔIE) between glucose and six important residues for different mutations (b). The residues analyzed were those with $|IE| > 15 \text{ kcal mol}^{-1}$ in **Figure 3**, and the mutations analyzed were D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K) (b). Geometries were obtained at ONIOM2(B3LYP/6-31G(d,p):PM3) level of theory. All energies in the bar plot for the mutants are relative to the IE values of WT (shown in parentheses in (a) in kcal mol^{-1}). The energies were calculated at the B3LYP/6-31G(d,p) level with BSSE-CP. Negative and positive ΔIE values mean the particular residue stabilizes and destabilizes the glucose binding, respectively. The hydrogen bonds are displayed as orange dash lines.

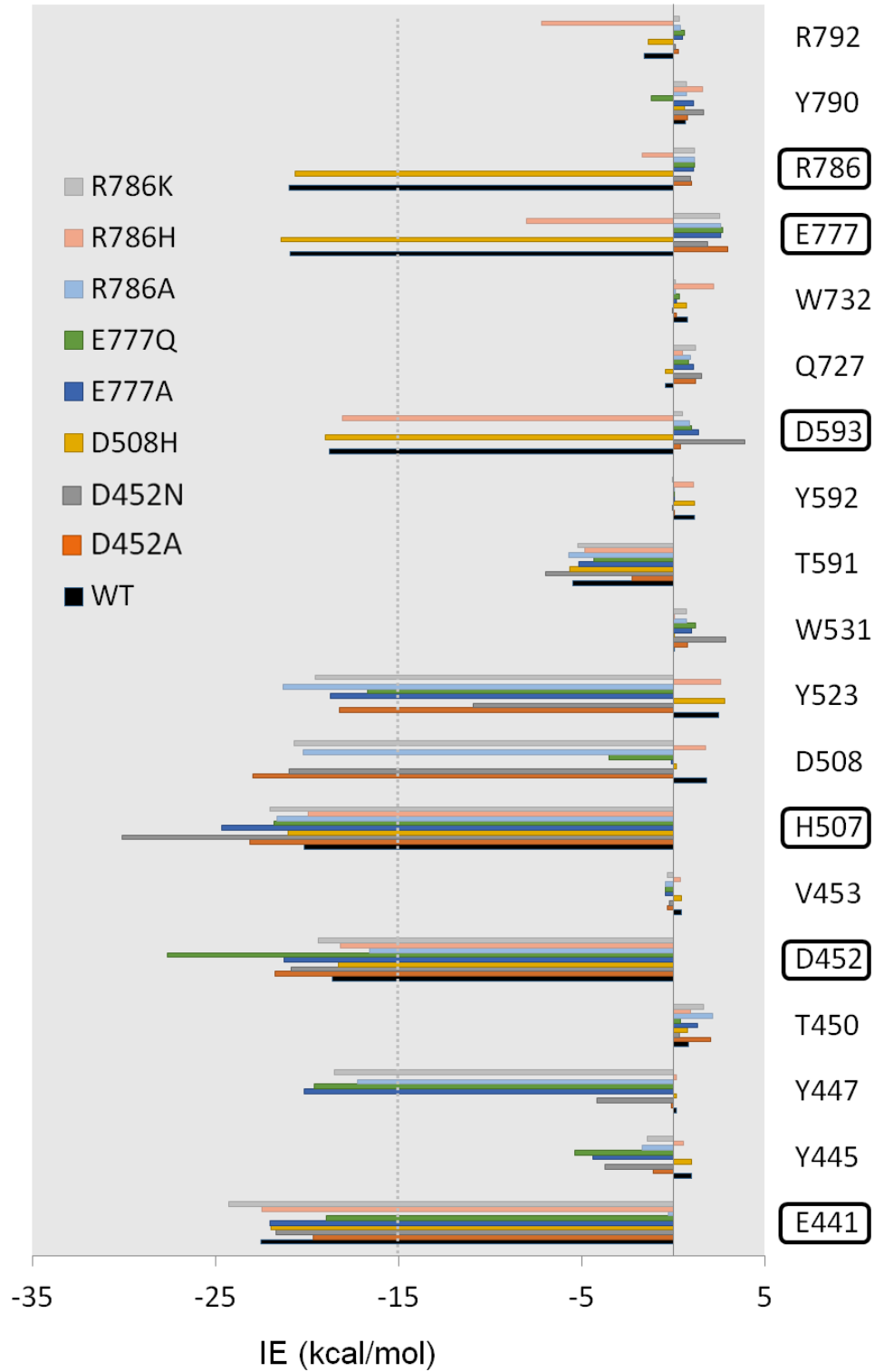


Figure 2. Particular interaction energy (*IE*, kcal/mol) between the glucose and individual residues for wildtype (WT) and mutant *TxGH116* models (D452A, D452N, D508H, E777A, E777Q, R786A, R786H, R786K). Six residues (E441, D452, H507, D593, E777, R786) were identified as strong electrostatic contributors for the *TxGH116*-glucose binding, with $|IE| > 15$ kcal/mol (shown in rectangular line).

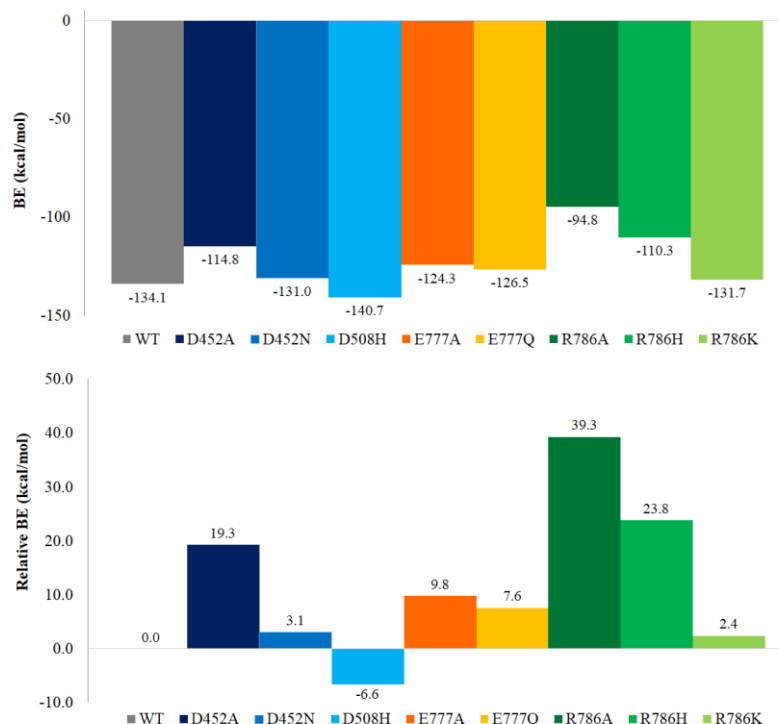


Figure 3. Binding interaction energies (BE, kcal/mol) of glucose for WT and mutant models (D452A, D452N, D508H, E777A, E777Q, R786A, R786H, R786K) (Top). The corresponding binding energies relative to the BE of WT were also included (Bottom). Positive and negative values of the relative BE (ΔBE) for particular mutant models indicate the unfavorable and favorable binding of the glucose ligand in the TxGH116 pocket, respectively.

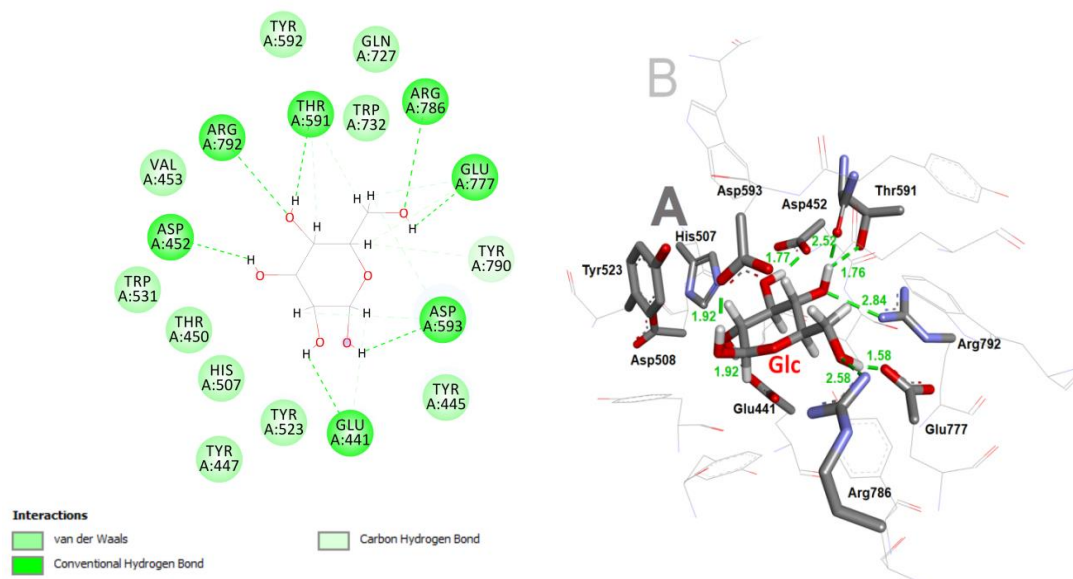


Figure 4. 2D protein-ligand interaction map (left) and 3D (optimized) structure (right) of WT TxGH116-glucose (Glc) complex obtained at ONIOM2 (B3LYP/6-31G(d,p):PM3) level. Regions A (shown in tube model) and B (shown in wire model) used in the ONIOM calculations. Green dash lines indicate H-bond distances (in Å) between glucose, Glc, and TxGH116 active site. Important interactions are also indicated, which mainly involves the hydrogen bond and van der Waals interactions.

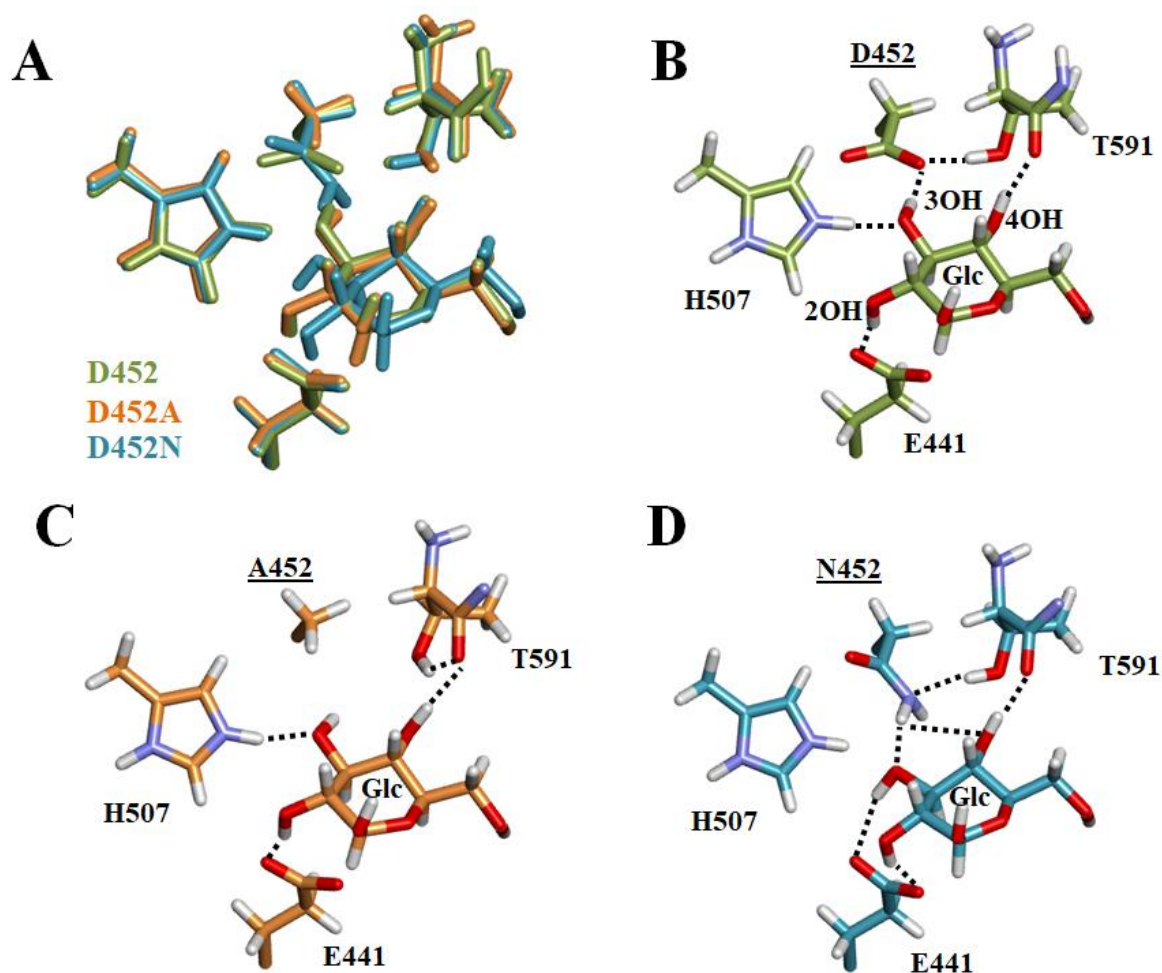


Figure 5. Structural overlap of the TxGH116-glucose complex models (obtained at the ONIOM2 [B3LYP/6-31G(d,p):PM3] level) between the WT (green) and the mutants, D452A (orange) and D452N (cyan) (A) and their hydrogen bond (H-bond) interactions involved; (B) WT, (C) D452A, (D) D452N. For clarity, only the residues making hydrogen bonds at C2OH, C3OH and C4OH of glucose are shown. The mutated residue 452 is also indicated with underlined labels.

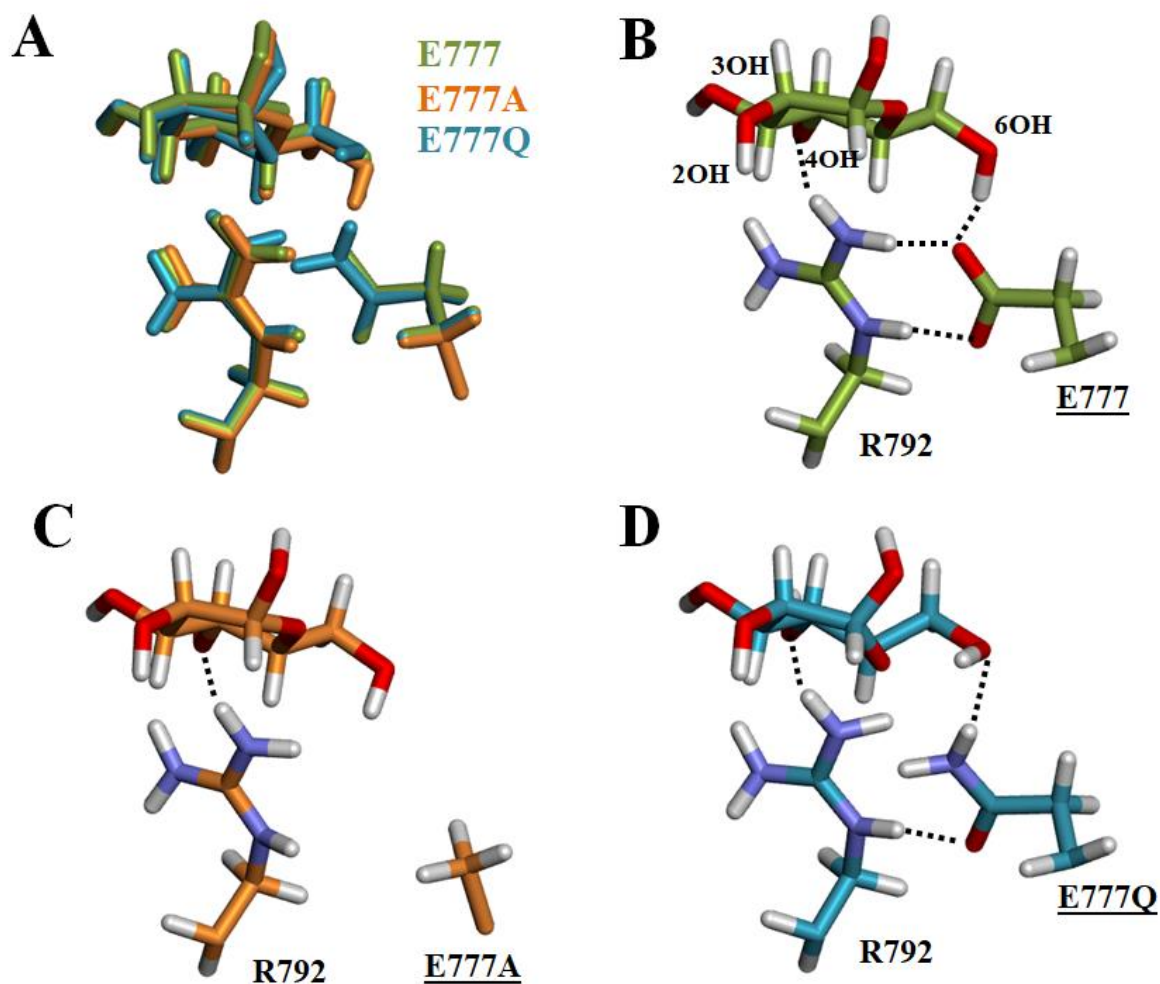


Figure 6. Structural overlap of the TxGH116-glucose complex models (obtained at the ONIOM2 (B3LYP/6-31G(d,p):PM3) level) between the WT (green) and the mutants, E777A (orange) and E777Q (cyan) (A) and their hydrogen bond (H-bond) interactions involved; (B) WT, (C) E777A, (D) E777Q. For clarity, only the residues R792 and E777 (E777Q/E777A) were shown. Mutated position at 777 is also indicated with underline.

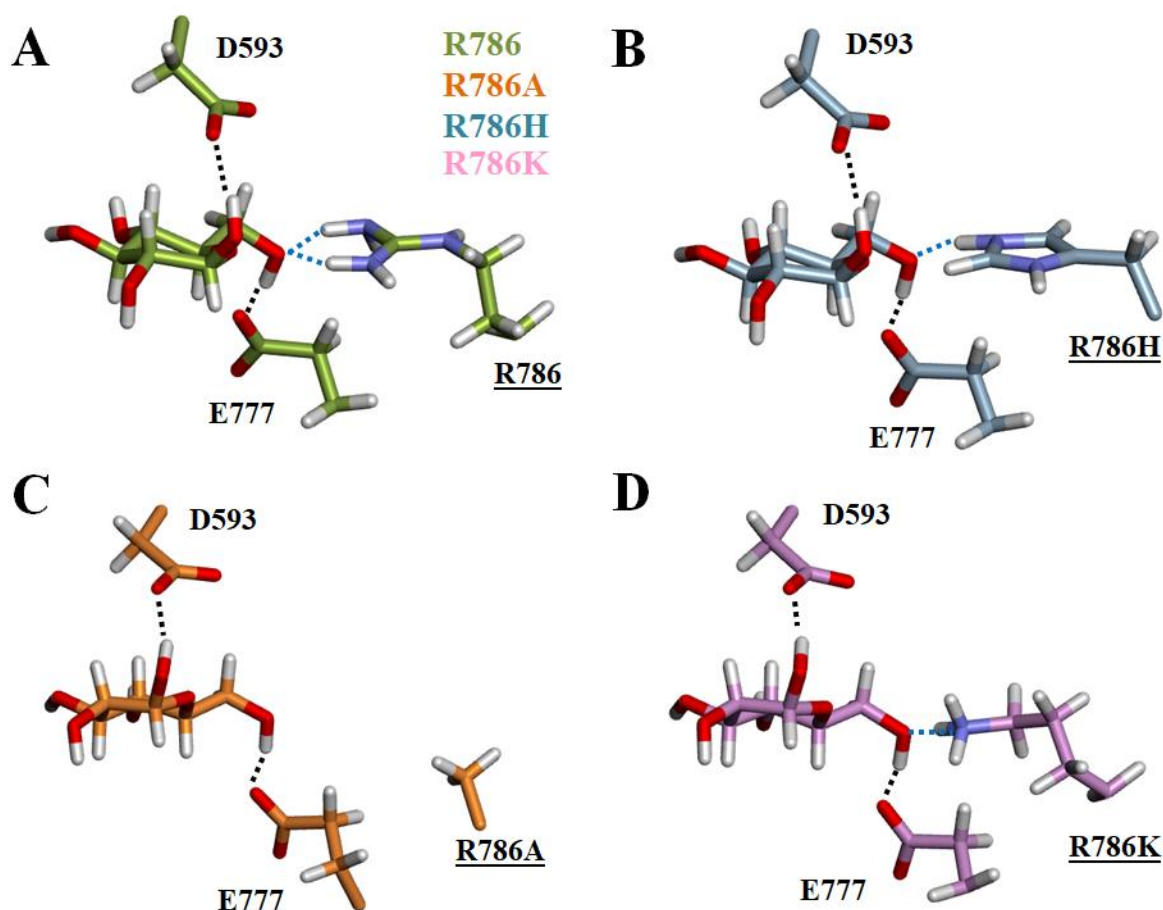


Figure 7. H-bond interactions in the TxGH116-glucose complex models (obtained at the ONIOM2 (B3LYP/6-31G(d,p):PM3) level) for (A) WT (green), (B) R786H (cyan), (C) E786A (orange) and (D) R786K (pink). For clarity, only the residues D593, E777 and R786 (R786Q/R786A) were shown. Mutated position at 786 is also indicated with underline.

CONCLUSIONS

We endeavored to assess the importance of glucose-binding residues in GH family 116 (GH116) β -glucosidases, which include human β -glucosylceramidase 2 (GBA2), by ONIOM calculations on *Thermoanaerobacterium xylanolyticum* TxGH116, the structural model for GH116 enzymes. The ONIOM binding energy calculations identified D452, H507, E777, and R786, along with the catalytic residues E441 and D593, as strong electrostatic contributors to glucose binding with predicted interaction energies $> 15 \text{ kcal mol}^{-1}$, consistent with the effects of the D452, H507, E777 and R786 mutations on enzyme kinetics. The relative importance of GH116 active site residues in substrate binding and catalysis identified in this work improves the prospects for the design of inhibitors for GBA2 and the engineering of GH116 enzymes for hydrolytic and synthetic applications.

SUGGESTIONS

- (1) Accurate energies on the ONIOM model can also be improved with *Ab initio* method such as MP2, CCST.
- (2) Increasing the size of the model with thermal fluctuation can generate more realistic models and results.

Chapter IV

Quantum Chemistry Method for Mechanistic and Structural Investigation of Zinc Metalloenzyme

Meelua, W., J. Olah and **J. Jitonnorn** (2022). "Role of water coordination at zinc binding site and its catalytic pathway of dizinc creatininase: insights from quantum cluster approach." **J. Comput. Aided Mol. Des.**: DOI 10.1007/s10822-022-00451-8. (Q1, IF 2020 =3.686, Cited by 0).

ROLE OF WATER COORDINATION AT ZINC BINDING SITE AND ITS CATALYTIC PATHWAY OF DIZINC CREATININASE: INSIGHTS FROM QUANTUM CLUSTER APPROACH

INTRODUCTION

Creatininase (creatinine amidohydrolase; EC 3.5.2.10), is a member of the urease-related amidohydrolase superfamily. It is a creatinine-metabolizing enzyme that catalyzes hydrolysis of creatinine to creatine, which is a breakdown product of creatine metabolism in mammals. This enzyme plays a key role in the bacterial degradation of creatinine (Szulmajster 1958) and participates in arginine and proline metabolism. In medicine, the creatininase has a great potential for diagnostic application, including its use as biosensors for accurate measurement of renal function (Wyss and Kaddurah-Daouk 2000). Creatininase from *Pseudomonas putida* has been extensively studied (Kaplan and Szabo 1974, Tsuru, Oka et al. 1976, Yoshimoto, Oka et al. 1976, Rikitake, Oka et al. 1979, Schumann, Böhm et al. 1993, Yamamoto and Oka 1995, Tang and Wen 2000). Pioneering works on the structure and function of the creatininase enzyme were described by Beuth et al. and Yoshimoto et al. (Beuth, Niefind et al. 2002, Ito, Kanada et al. 2002, Beuth, Niefind et al. 2003, Yoshimoto, Tanaka et al. 2004, Yamashita, Nakajima et al. 2010). The X-ray crystal structures of creatininase in its apo and product complex forms revealed that it is a typical zinc enzyme with a binuclear metal center as a metal cofactor at the active site of each subunit (Mn^{2+}/Zn^{2+} or Zn^{2+}/Zn^{2+}) (Beuth, Niefind et al. 2003, Yoshimoto, Tanaka et al. 2004). The binuclear metal center is bridged by an aspartate (Asp45) and a zinc-bound water or a hydroxide ion (Wat1) (see **Scheme 1**). Zn1/Mn1 is liganded by three protein residues (Glu34, Asp45, and His120), a carbonyl oxygen of the creatinine substrate, and the bridging hydroxide Wat1, while Zn2 is liganded by three protein residues (His36, Asp45, and Glu183) and Wat1. In the active site, there is a second water molecule (Wat2) situated very close to Zn1 and the Glu122 residue, which belongs to the second-shell. Mutational and computational studies (Yamashita, Nakajima et al. 2010, Jitonnorn, Mujika et al. 2017) suggest that the lack of Glu122-Wat2 interaction in the Glu122Asn mutant accounts for the reduced activity.

Similarly to other dinuclear zinc enzymes ((Seibert and Raushel 2005, Rong-Zhen Liao, Jian-Guo Yu et al. 2008, Liao, Yu et al. 2009, Liao, Himo et al. 2010, Krivitskaya, Khrenova et al. 2021, Reidl, Mascarenhas et al. 2021), the reaction mechanism of creatininase occurs via nucleophilic attack and ring opening. The first step is quite common to in these zinc metalloenzymes, i.e., the bridging water performs a nucleophilic attack on the substrate carbonyl carbon, leading to the formation of tetrahedral intermediate. However, in the ring-opening step, the precise mechanism and the significant roles of the proton donor remain unclear. Several experimental and computational studies have been carried out to identify the exact proton donor of the catalytic reaction in creatininase, which was suggested to be either the His178 residue or a water molecule, Wat2, in the enzyme active site (hereafter named path A and path B, respectively; Scheme 1). In a proposed mechanism

suggested by Beuth et al. (Beuth, Niefind et al. 2003) (path A in **Scheme 1**), the bridging hydroxide Wat1 acts as the nucleophile in the first nucleophilic attack step, whereas the His178 serves as a proton donor in the ring-opening step. The H178A mutation abolished the activity of the enzymes calling the attention to the crucial role of His178 in catalysis (Yamashita, Nakajima et al. 2010). Contrary this suggestion, Yoshimoto et al. (Yoshimoto, Tanaka et al. 2004, Yamashita, Nakajima et al. 2010) (path B2 in **Scheme 1**), proposed that instead of His178 a second water molecule Wat2 acts as a proton donor in the ring-opening step. However, our recent QM/MM free-energy simulations (Jitonnom, Mujika et al. 2017) suggest that this role of Wat2 is unlikely. In particular, our computer simulations clearly show that the most likely ring-opening step occurs through one concerted proton transfer step with His178 playing a central role as a proton shuttle. Still, several details remain unclear, such as the possible dual role of both His178 and Wat2 and the zinc coordination mode during catalysis. Alternatively, apart from path A and B2 where the proton donor is His178 or Wat2, they can participate together and performs a couple role, i.e., the His178 is firstly protonated and then Wat2 serves as a proton donor in the ring opening (see path B1 in **Scheme 1**). This pathway B1 was not considered in the previous QM/MM study. Particularly, we wondered whether this water molecule (Wat2) would act as a proton donor once the His178 is protonated through a proton transfer from the gem-diolate proton. Here, we considered this second water molecule and its catalytic role are elucidated. Overall, this information clearly indicates that clarifying the possible dual role of both His178 and Wat2 is worthy of investigation.

In this work, we have used a combination of the density functional theory (DFT) method and the quantum cluster approach to investigate the reaction mechanism of creatininase concerning possible pathways regarding the chemistry of zinc coordination and protonation state of active site environment. A model of the active site was designed on the basis of the crystal structure, and the stationary structures and energetics of the reaction pathways (**Scheme 1**) were calculated. This DFT cluster approach has previously been successfully applied to study reaction mechanism of different enzymes (Rong-Zhen Liao, Jian-Guo Yu et al. 2008, Liao, Yu et al. 2009, Liao, Himo et al. 2010, Liao, Yu et al. 2010, Zhu, Barman et al. 2012, Manta, Raushel et al. 2014, Jitonnom, Sattayanon et al. 2015, Jitonnom and Hannongbua 2018). Our cluster model calculations can discriminate different mechanisms proposed earlier (**Scheme 1**) and provide a computational evidence for the dual role of His178 and Wat2, which have not yet been well described. A detailed mechanistic understanding of the enzyme will enhance current knowledge and application of an enzymatic biosensor for kidney function.

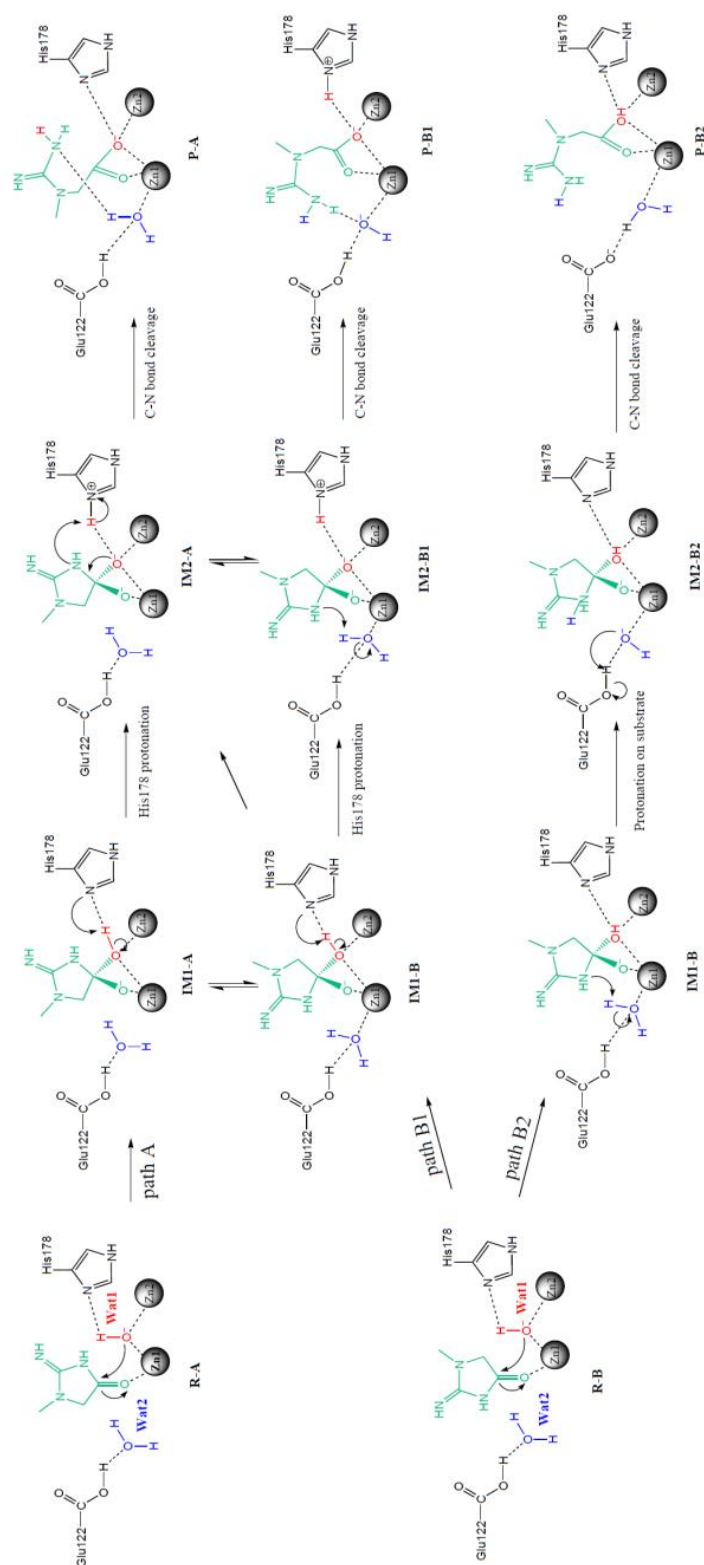
OBJECTIVES

- (1) To elucidate the possible reaction mechanism of creatininase
- (2) To describe the dual role of His178 and Wat2 during the creatininase reaction

METHOD

The model system for the active site dinuclear zinc cluster of the creatininase was constructed on the basis of a high resolution X-ray crystal structure of the enzyme-product (creatine) complex of the Mn-activated creatininase (PDB entry 1V7Z) (Yoshimoto, Tanaka et al. 2004)), a typical Zn^{2+} enzyme with one zinc ion replaced

with Mn^{2+} . To build a dizinc enzyme cluster model (Zn-Zn enzyme), the creatinine substrate was built from the X-ray structure by manually adjusting the carbon-nitrogen bond to form the five-membered ring. One oxygen atom of creatine that is bridged with two metal centers was replaced with a hydroxide anion (OH^-), a highly reactive and unstable species (Fernandes, Teixeira et al. 2019). The manganese ion in the active site was replaced with a zinc ion to build the Zn-Zn enzyme cluster model. Hydrogen atoms were added manually, and the amino acids were truncated so that in principle only side chains were kept in the model (see **Figures 1-2**). Truncated bonds were saturated with hydrogen atoms. Thus, the model consists of the two zinc ions along with their ligands His36, His120, His178, Glu34, Glu122, Glu183 and the bridging carboxylated Asp45 and the hydroxide ion (Wat1). With an exception of Glu122, all Glu/Asp residues were negatively charged, while His residues were neutral. For Glu122, we assigned this residue as the neutral form



Ring-opening

Nucleophilic attack

Scheme 1. Possible reaction pathways (paths A, B1 and B2) considered in this study. Path A refers to the His178-assisted mechanism (via protonated His178) suggested by Beuth et al. [11] and Jitonnorn et al. [15], while path B involved the water-assisted mechanism proposed by Yoshimoto et al. [13,14] where the active site water, Wat2, nearby Glu122 and Zn1 acts as a proton donor of the ring-opening step if His178 is protonated or unprotonated (referred to paths B1 and B2, respectively).

according to our previous study (Jitonnom, Lee et al. 2011). Meanwhile, according to the electronic structures the possibility of the ionized Glu122 at the initial stage of the reaction is unlikely as this would rule out a proton transfer from Wat2 to the cyclic amide nitrogen (path B in **Scheme 1**), considering its potential hydrogen bonds with Wat2 and Glu34 (data not shown). Additionally, the second water molecule (Wat2), which is thought to be important for catalysis (Yamashita, Nakajima et al. 2010), were also included in the model. The resulting active-site model is thus composed of 87 atoms and has a zero net charge.

Following the cluster approach (Siegbahn and Himo 2009, Siegbahn and Himo 2011, Himo 2017, Ahmadi, Barrios Herrera et al. 2018), truncation atoms (C_{β} atoms) were fixed to their corresponding positions from the X-ray structure. The model system was then optimized at the B3LYP/6-31G(d) level, with no initial symmetry restrictions and assuming C1 point group. Frequency calculations were performed on all optimized structures to obtain zero-point energy corrections to the electronic energies, and to identify all the stationary points as minima (zero imaginary frequency) or transition states (one imaginary frequency). The atomic charge distributions were also computed at the same level of theory using Mulliken population analysis (Mulliken 1955). More accurate B3LYP energies were also obtained by performing single-point energy calculations on the optimized geometries using a larger 6-31+G(d,p) basis set for all elements. Solvation effect was taken into account using the conductor-like polarized continuum model (CPCM) (Cossi, Rega et al. 2003) using a dielectric constant (ϵ) of 4, as suggested by Himo et al. (Siegbahn and Himo 2011, Himo 2017). It was shown that inclusion of solvent effects significantly lowers the barriers and intermediate energies compared to gas phase for path A, while they increase the barriers and relative energies for path B (data not shown). Altering this dielectric constant to higher or lower value does not change significantly the overall shape of the potential energy profile ($\epsilon = 1, 4, 80$; data not shown). Dispersion effects were performed with DFT-D3 program as developed by Grimme et al. (Grimme, Antony et al. 2010, Grimme, Ehrlich et al. 2011). It is shown that both implicit solvation and dispersion have a significant and important effect in lowering the energy barrier (Lonsdale, Harvey et al. 2010, Lonsdale, Harvey et al. 2012, Brás, Fernandes et al. 2014, Ahmadi, Barrios Herrera et al. 2018). Final reported energies are the electronic energies with the large basis set plus corrections for zero-point energy, CPCM solvation, and dispersion in kcal/mol. All calculations were performed using Gaussian 09 program (Frisch, Trucks et al. 2009).

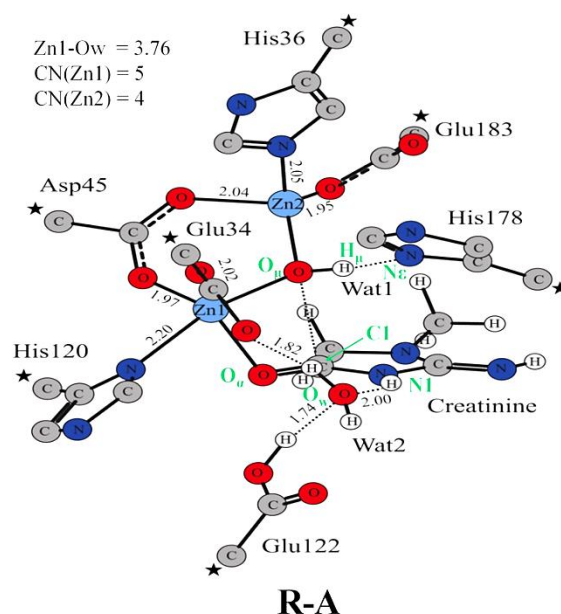


Figure 1. DFT optimized structure of reactant for path A. For clarity, some hydrogen atoms are omitted in the figures. The fixed atoms are marked with asterisks. Distances are in angstrom ($\text{\AA}$).

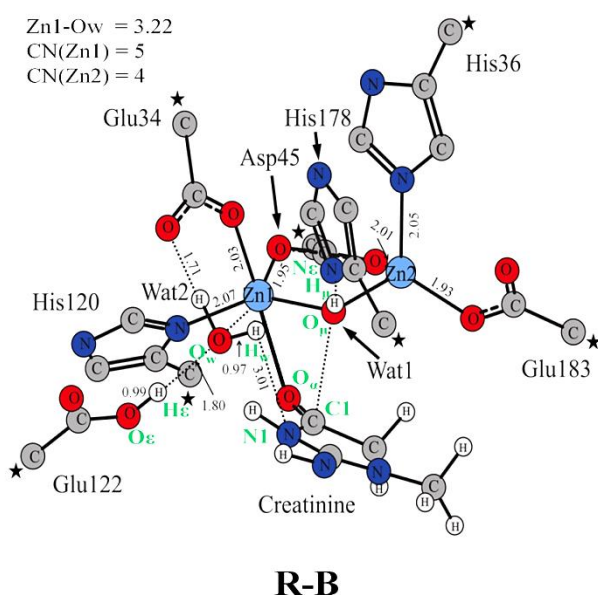


Figure 2. DFT optimized structure of reactant for path B. For clarity, some hydrogen atoms are omitted in the figures. The fixed atoms are marked with asterisks. Distances are in angstrom ($\text{\AA}$).

RESULTS AND DISCUSSION

To further scrutinize these previously suggested mechanisms in **Scheme 1** with QM cluster calculations, we first optimized the structures for the reactant complex and the structures of the proposed intermediates and their corresponding energies were evaluated for paths A, B1 and B2. The results of the QM calculations have identified seven stationary points (**R**, **TS1**, **IM1**, **TS2**, **IM2**, **TS3** and **P**) in all studied pathways.

The main geometric results are summarized in **Table 1** and visualized in **Figures 1-5**. All coordinates for all important structures are given in Supporting Information. Relative energies calculated for paths A, B1 and B2 using cluster model are depicted in **Figure 6**. The criteria for the ligation ability of Wat2 at Zn1 were measured by the Zn1-O_w distance, having a cut-off of < 3.00 Å for bound geometry (*Wat2_bound*) and > 3.00 Å for unbound geometry (*Wat2_unbound*).

3.1 Reactant

Initially, several structures at the reactant (**R**) and intermediate (**IM**) in both bound and unbound modes of Wat2 with respect to Zn1 (hereafter, *Wat2_bound* and *Wat2_unbound*, respectively) were built and optimized. Both geometries were successfully obtained at the intermediate (**IM1-A** and **IM1-B**, see **Figures 3-4**) but only *unbound* geometries were obtained at the reactant (see **R-A** and **R-B** in **Figures 1-2**, respectively). Attempts to generate a stable structure of *bound* reactant from the *bound* intermediate geometry (**IM1-B**) are not satisfied during the relaxed scan along the C-O distance (data not shown). Therefore, we obtained one *unbound* for the reactant, while both *bound* and *unbound* for the intermediate.

In the reactant, due to hydrogen bonding possibility of Wat2 with respect to Glu34, Glu122 and the substrate amide bond, this water could form, via its oxygen, a direct hydrogen bond to the substrate nitrogen atom (N1) (**R-A** in **Figure 1**) or situate above the substrate prepared for a proton transfer as suggested by Yoshimoto et al. (**R-B** in **Figure 2**). The optimized structure of the reactant (**R-A**) for path A is shown in **Figure 1**. This structure represent the initial stage prior to the reaction to take place, which is not seen in the X-ray structure (PDB 1V7Z) (Yoshimoto, Tanaka et al. 2004) where the hydroxide/water was replaced by the carboxylate terminus of the creatine product. In particular, the bridging hydroxide has symmetric bonds to the two zinc ions (Zn1-O_μ = 2.03 Å and Zn2-O_μ = 1.94 Å) and forms a hydrogen bond with the His178 residue (Nε-H_μ = 1.87 Å).

The reactant structure (**R-B**) of path B is first optimized and it is shown, in **Figure 2**, that the computed R-B structure is quite similar to the one calculated for path A. The minor structural difference is the orientation of Wat2 with respect to the N1 atom of substrate and the ligand coordination geometry of Zn1 (Zn1-O_w = 3.22 Å; Zn1-O_α = 2.39 Å for path B and Zn1-O_w = 3.76 Å; Zn1-O_α = 2.10 Å for path A). This Wat2 is completely unliganded to Zn1. Therefore, the coordination number is 5. With the C1-O_μ distance of 2.49 Å, the bridging hydroxide is, however, possible for in-line nucleophilic attack, stabilized by His178 (Nε-H_μ = 1.99 Å at **R-B**). It should be mentioned that the relative energy of the reactant structure of path B (**R-B**) is by 6.5 kcal/mol more stable than that of path A (**R-A**). It is found that for both R-A and R-B, Zn1 is more electropositive than Zn2 (0.90e/0.88e on Zn1/Zn2 for path A and 0.92e/0.90e on Zn1/Zn2 for path B). Similar results were also previously observed in a dizinc metallopeptidase (Zhu, Barman et al. 2012).

3.2 Nucleophilic attack and ring opening

The optimized structures of transition states, intermediate and product are shown in **Figures 3-5** and the potential energies of these stationary points calculated for all pathways are depicted in **Figure 6**, with respect to the energy of the reactant **R-B**.

3.2.1 Path A The first step of the reaction (**Scheme 1**) is a nucleophilic attack by the bridging hydroxide on the C1 atom of the substrate, leading to the formation of

tetrahedral gem-diolate intermediate (**R-A**→**IM1-A**). This step is common to other hydrolytic dinuclear zinc enzymes (Seibert and Raushel 2005). Analysis of the same step using the water H₂O (Wat1) as a nucleophile was also carried out, which suggests this mechanistic scenario to be energetically demanding (data not shown). **Figure 3** shows the optimized structures of transition states, intermediates and product obtained for this path. The critical C1-O_μ distance is 1.91 Å at **TS1-A** and is fully bonded at **IM1-A** with a bond distance of 1.51 Å. The corresponding **TS1-A** has an imaginary frequency of 190i cm⁻¹. The barrier of this step is calculated to be 6.3 kcal/mol (7.2 kcal/mol without solvation) and the energy of **IM1-A** is calculated to be 5.3 kcal/mol (7.5 kcal/mol without solvation) relative to the **R-A**. The hydroxyl group of the gem-diolate forms a hydrogen bond to His178 with a Nε-H_μ distance of 1.63 Å, setting a stage for further proton transfer assisted by His178. The shortening of the Nε-H_μ distance (from 1.87 Å in **R-A** to 1.63 Å in **IM1-A**), and the Zn1-O_α distance (from 2.10 Å in **R-A** to 1.87 Å in **IM1-A**) implies that His178 and Zn1 anticipate in stabilizing the transition state and tetrahedral intermediate, thereby lowering the barrier for the nucleophilic attack. During this step, the coordination number (CN) of Zn1 changes from 5 to 4, resulting from the detachment of O_μ from Zn1 (2.03 Å at **R-A** to 2.96 Å at **IM1-A**), while the tetracoordination geometry of Zn2 remains unaltered (CN=4). This also explains the laboratory experiment why Zn1 can be replaced by other divalent cations such as Mn²⁺ or Co²⁺ (Kim and Lipscomb 1993).

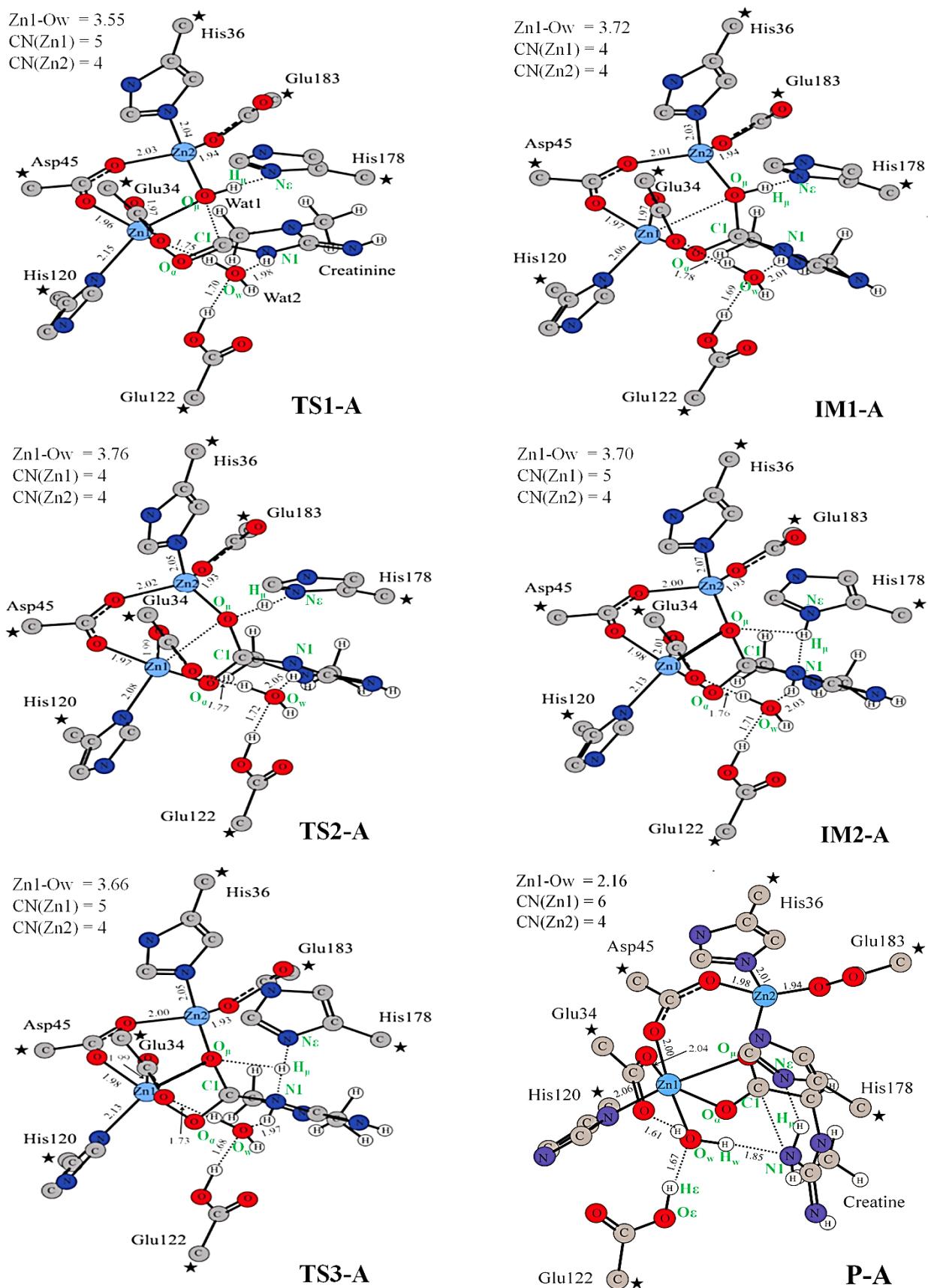


Figure 3. DFT optimized structures of transition states, intermediates and product for path A. The fixed atoms are marked with asterisks. Distances are in angstrom (Å).

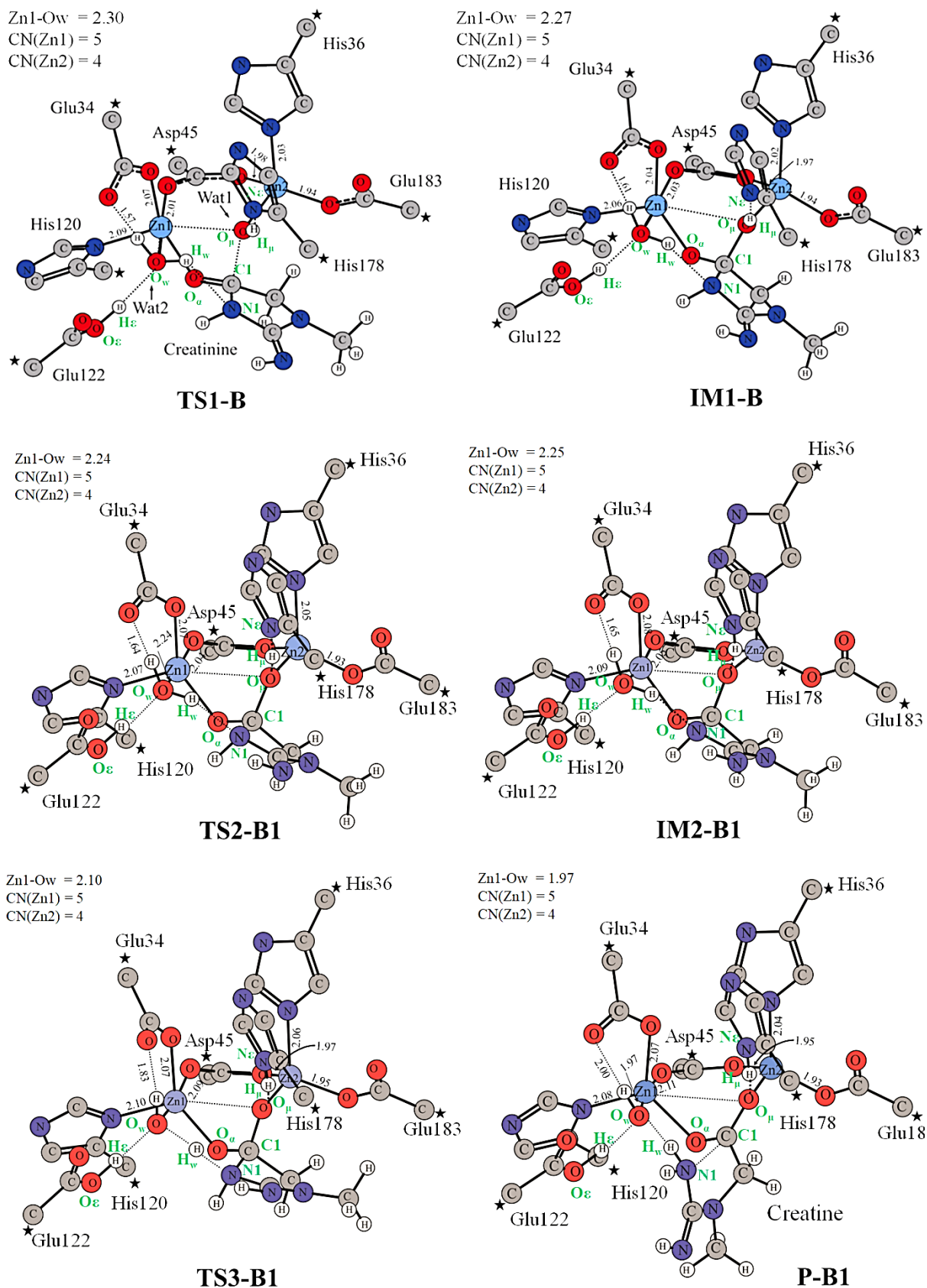


Figure 4. DFT optimized structures of transition states, intermediates and product for path B1. The fixed atoms are marked with asterisks. Distances are in angstrom (Å).

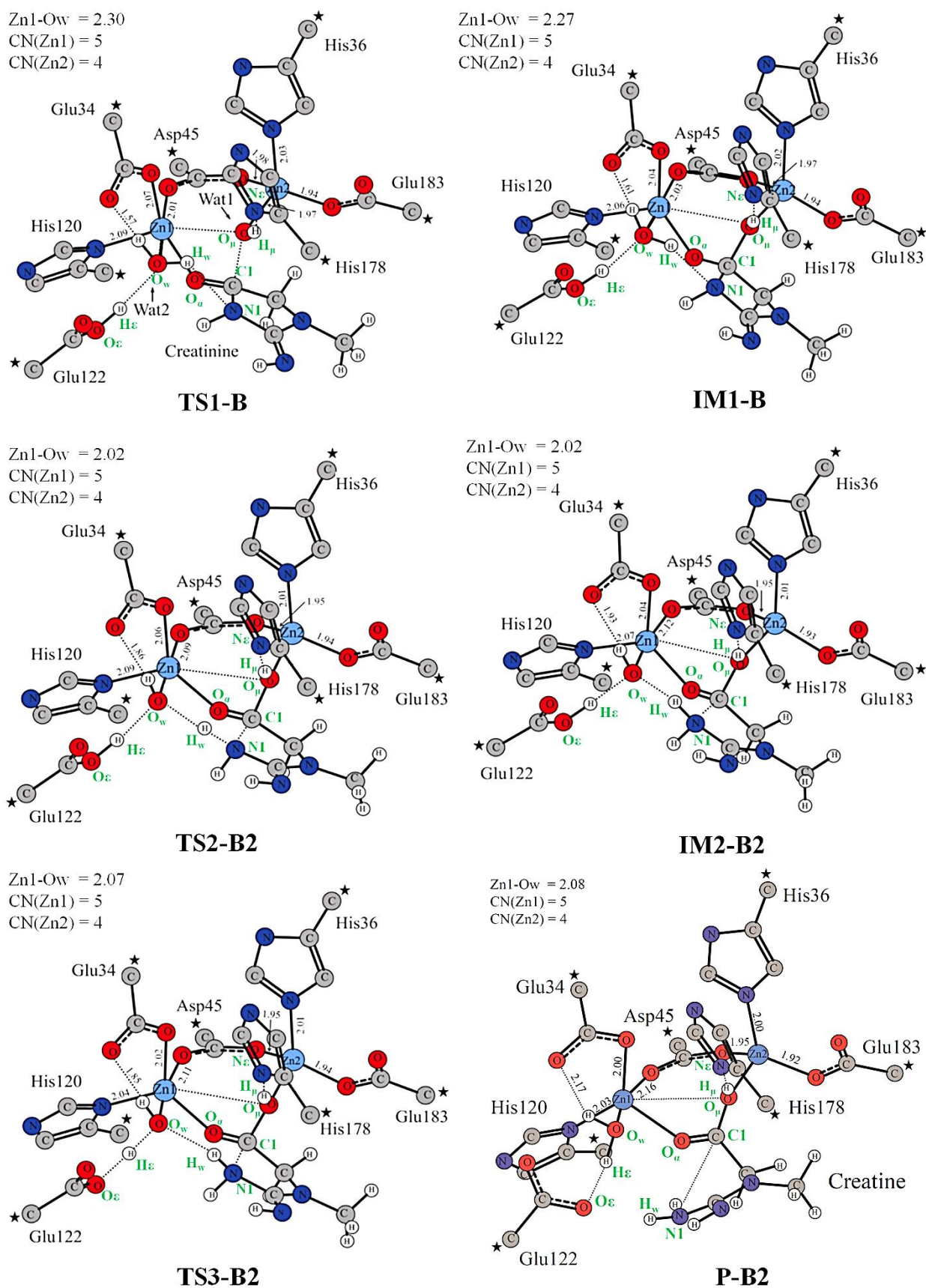


Figure 5. DFT optimized structures of transition states, intermediates and product for path B2. The fixed atoms are marked with asterisks. Distances are in angstrom (Å).

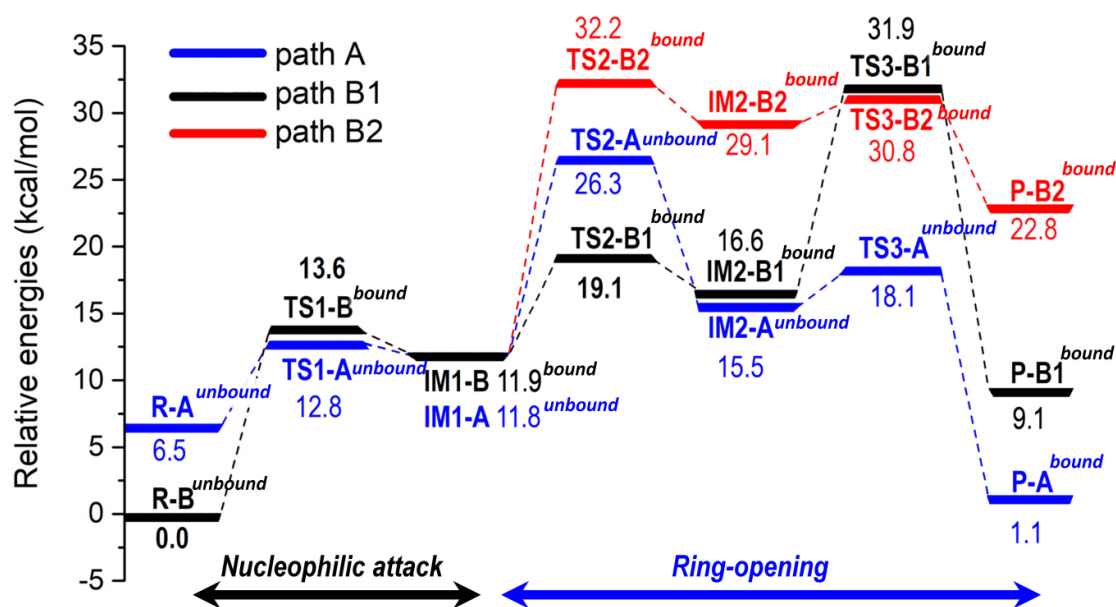


Figure 6. Relative potential energy profiles of the creatinine hydrolysis for paths A, B1, and B2. The *bound* and *unbound* modes of Wat2 at Zn1 are also indicated.

The next step (**IM1-A**→**P-A**) is the ring opening which involved the proton transfer and the C1-N1 bond cleavage, assisted by the His178 residue. As suggested by previous studies (Beuth, Niefind et al. 2003, Jitonnorn, Mujika et al. 2017), His178 now plays a central role by shuttling a proton from the hydroxide of gem-diolate intermediate to the nitrogen leaving group of the creatinine ring. From the QM-cluster calculations, this process appears to occur as two separate steps, consisting of His178 protonation and subsequent C1-N1 bond fission (via two transition states **TS2-A** and **TS3-A**). At **TS2-A**, the transferring proton (H_{μ}) is in between the N_{ϵ} atom of His178 imidazole ring and the gem-diolate O_{μ} atom ($N_{\epsilon}-H_{\mu} = 1.43 \text{ \AA}$, $O_{\mu}-H_{\mu} = 1.48 \text{ \AA}$; **TS2-A** in Figure 3) and this corresponds to the stretching of the $N_{\epsilon}-H_{\mu}$ and $O_{\mu}-H_{\mu}$ bonds as evident by the imaginary value of $1691i \text{ cm}^{-1}$. After **TS2-A**, His178 is now fully protonated ($N_{\epsilon}-H_{\mu} = 1.08 \text{ \AA}$) and, by orienting its imidazole ring, forms a hydrogen bond to the nitrogen leaving group ($N1-H_{\mu} = 1.63 \text{ \AA}$ at **IM2-A**). At **IM2-A**, due to the loss of one proton to His178, the bridging hydroxyl is now ionized and unstable, which is then stabilized by coordinating with Zn1 as a fifth ligand ($CN=5$ at Zn1). This Zn1 remains pentacoordinated until the formation of the product state, where Wat2 is ligated as a sixth ligand in a hexacoordination sphere. The **TS3-A** structure, as confirmed by one imaginary value of $1355i \text{ cm}^{-1}$, shows that the transferring proton is approximately in half way between the N_{ϵ} (His178) and the nitrogen leaving group ($N_{\epsilon}-H_{\mu} = 1.31 \text{ \AA}$; $N1-H_{\mu} = 1.39 \text{ \AA}$). Downhill from **TS3-A**, the H_{μ} proton is completely transferred to the C1-N1 amide bond ($N1-H_{\mu} = 1.03 \text{ \AA}$), which is subsequently breaks ($C1-N1 = 2.06 \text{ \AA}$) to yield a product creatine with C-terminal carboxylate bound in the dizinc site in a bidentate fashion. The barriers of this ring-opening step were 19.8 kcal/mol (24.0 kcal/mol without solvation) for the His178 protonation and 11.6 kcal/mol (14.0 kcal/mol without solvation) for the C1-N1 bond cleavage, indicating the first proton transfer by His178 is the rate-determining step. The overall reaction is slightly exothermic by -5.4 kcal/mol (-4.9 kcal/mol without solvation). Note that the role of the proton shuttle His178 was also seen in other amino acid residues of both mono- and dizinc metalloenzymes (e.g., Asp250 in

dihydroorotase (Rong-Zhen Liao, Jian-Guo Yu et al. 2008), Asp288 in human renal peptidase (Liao, Himo et al. 2010), Glu384 of human angiotensin-converting enzyme (Brás, Fernandes et al. 2014), His216 of the β -carbonic anhydrase (Rowlett, Tu et al. 2002)).

3.2.2 Path B1

The proposed mechanism in path A suggests that the His178 residue plays the key roles during catalysis: it serves as an electrostatic stabilizer in the nucleophilic attack and a proton shuttle in the ring-opening step. However, a water molecule that situates very close to Zn1 and hydrogen bonds with Glu34 and Glu122 (Glu34-Wat2-Glu122 partner) might serve as a proton donor in the ring-opening, as suggested by Yoshimoto et al. (Yoshimoto, Tanaka et al. 2004, Yamashita, Nakajima et al. 2010). Figure 4 shows the optimized transition states, intermediates and product for the nucleophilic attack and ring-opening of path B2. As in the first step of path A, the nucleophilic attack readily takes place from a bridging position, and the C1-O $_{\mu}$ distance decreases from 2.49 Å at **R-B** to 1.79 Å at **TS1-B**, and further to 1.49 Å at **IM1-B**. This **TS1-B** exhibits an imaginary value of $245i\text{ cm}^{-1}$. The resulting intermediate **IM1-B** lies 11.9 kcal/mol higher than **R-B**. Despite the fact that the transition state structure **TS1-B** is quite similar to that of **TS1-A**, however, the barrier is calculated to be 13.6 kcal/mol (11.6 kcal/mol without solvation), which is two-fold higher than the corresponding barrier (6.3 kcal/mol) of path A. The reason for the higher barrier at **TS1-B** may be explained by the higher coordination number (CN) at Zn1 in path B (CN=5), compared with that of path A (CN=4). The lower **TS1-A** barrier compared to the corresponding **TS1-B** might be explained from the better nucleophile and electrophile of O $_{\mu}$ H $_{\mu}^{-}$ and C1 atom, as evident from the more negative and positive charge on them (O $_{\mu}$ = -0.81e to -0.95e and C1 = +0.62e for path A vs O $_{\mu}$ = -0.78e to -0.96e and C1 = +0.60e for path B). From **TS1-B**→**IM1-B**, Wat2 is now fully coordinated to Zn1 (Zn1-O $_{\alpha}$ = 2.30 and 2.27 Å, respectively), which is not present in path A (Zn1-O $_{\alpha}$ > 3.5 Å) and the previous free-energy simulations (Jitonnom, Mujika et al. 2017). At **IM1-B**, after Zn1 coordination rearrangement, Wat2 now serves as the fifth ligand of Zn1 and, as stabilized by Glu34 and Glu122 via two hydrogen bonds (1.61 and 1.68 Å, respectively), is allowed to orient its position to form a hydrogen bond (2.02 Å at **IM1-B**) to the N1 atom of creatinine, which is distorted away from the planar ring (see **IM1-B**, Figure 4). Conversely, this amide bond is directly hydrogen bonding with the oxygen atom (O $_w$) of Wat2 in path A. This process is required for next proton transfer step.

The ring-opening step in this pathway basically followed the same mechanism as in path A, i.e., the protonation of His178 and the breaking of the scissile bond. In this scenario, it is not His178 that plays this role as a central catalyst. On the other hand, it is also envisioned that both His178 and Wat2 might play a catalytic role as base and acid, respectively. In this path B1, the protonated His178 intermediate (**IM1-B**→**IM2-B1**) is formed, as evident from the shortening of the N $_e$ -H $_{\mu}$ distance from 1.76 Å to 1.00 Å, respectively. At **TS2-B**, the critical bonds, O $_{\mu}$ -H $_{\mu}$ and N1-H $_{\mu}$, are 1.40 and 1.34 Å, respectively, with one imaginary frequency ($1546i\text{ cm}^{-1}$). During this stage, the substrate amide bond is more distorted due to the shortened H $_w$ -N1 distance from 2.02 Å at **IM1-B** to 1.80 Å at **IM2-B2**. This structural arrangement prompts further proton transfer in the protonation of substrate amide bond that leads to the ring-opening. In **IM2-B**→**P-B1**, the protonation of the substrate by Wat2 results in the elongation of C1-N1 bond (1.50 Å at **IM2-B1** to 3.43 Å at **P-B1**) and the breakdown of creatinine ring into creatine product. The two critical distances, O $_w$ -H $_w$ and N1-H $_w$,

of this process are 1.42 and 1.35 Å at **TS3-B**, yielding one imaginary number ($1488i\text{ cm}^{-1}$) corresponding to the stretching of these bonds. At the end of this process, a penta-coordinated complex is obtained.

3.2.3 Path B2 The two proposed mechanisms above suggest that the reaction proceeds via the protonated His178 before the C1-N1 bond cleavage and ring-opening assisted by a proton donor (His178 in path A or Wat2 in path B1). In fact, a water molecule that situated very close to Zn1 might play this role during the course of the reaction, as suggested by Yoshimoto et al. (Yamashita, Nakajima et al. 2010), without the His178 protonation as described earlier. In particular, we also tested another pathway where Wat2 serves as the catalytic acid and His178 maintains its neutral form during the catalysis. Figure 5 shows the optimized transition states, intermediates and product for this path. Since the nucleophilic attack follows the same pathway as in the case of path B1, therefore we will discuss only the second ring-opening step.

In the ring-opening step, the tetrahedral intermediate is subsequently collapsed by the donation of a proton (H_w) from Wat2 to the amide of the substrate (**IM1-B**→**IM2-B2**), followed by the transfer of a proton (H_e) from Glu122 to Wat2 to regenerate a water molecule (**IM2-B2**→**P-B2**). The critical bond distances for the first proton transfer are calculated at 1.60 (O_w-H_w) and 1.22 ($N1-H_w$) Å at **TS2-B2**. Meanwhile, the transient bonds for the second proton transfer by Glu122 are 1.39 (O_e-H_e) and 1.28 (O_w-H_e) Å at **TS3-B2**. Values of imaginary frequency of these two transition states (**TS2-B2** and **TS3-B2**) are $-827i$ and $-1351i\text{ cm}^{-1}$, respectively. The barriers of these proton transfers are 32.2 at **TS2-B2** and 30.8 kcal/mol at **TS3-B2**, which is much higher than the other pathways. This clearly indicates that the proton transfer from Wat2 to the scissile bond of substrate is very unfeasible when His178 is neutral and this renders ineffective catalysis. However, His178 still provides stabilization to the reaction via hydrogen bonding to the gem-diolate hydroxide, as can be seen from the decreasing of N_e-H_μ distances (1.95, 1.76, 1.69, 1.68, 1.68 and 1.41 Å at **TS1-B**, **IM1-B**, **TS2-B2**, **IM2-B2**, **TS3-B2**, **P-B2**, respectively). The C1–N1 distance of 3.15 Å at **P-B2** clearly indicates that the amide bond of the creatinine ring is fully broken, yielding the creatine product with the amine and carboxylic groups. These stepwise proton transfers lead to the cleavage of the amide bond of the substrate releasing a creatine molecule.

3.2.3 Summary

In general, the first nucleophilic attack for all pathways (A, B1 and B2) is the same, i.e. the attack of hydroxide on the substrate carbonyl that leads to the formation of tetrahedral gem-diolate intermediate (see **IM1-A** in **Figure 2** and **IM1-B** in **Figures 4-5**). However, in the second ring-opening step, the proton donor can be either His178 or Wat2 (see **Scheme 1**) or both, which delivers its proton to the substrate for opening the creatinine ring to yield the creatine product. In the ring opening of path A, His178 serves as a proton shuttle by accepting one proton of the -OH group of the gem-diolate intermediate (and His178 is now protonated; see *His178 protonation* step in **Scheme 1**) and transferring it to the substrate. During this step, Wat2 is unbound at Zn1 and does not participate during catalysis. In the ring opening of path B1, His178 is protonated via the same pathway as in path A but the cleavage of the substrate ring is now activated by the proton transfer from Wat2, which turn into a hydroxide at the final product. Alternatively, in path B2, His178 is unprotonated and the proton transfer from Wat2 to the substrate is assisted by Glu122.

Comparison of energetic profiles for all possible pathways in **Figure 1** also shows that the enzyme likely favors the pathway involved in the protonation of His178 (paths A and B1) prior to the ring cleavage. In path A, His178 acts as a catalytic base/acid, and in path B1 His178 and Wat2 can serve as an equal foot as the catalytic base and the catalytic acid, respectively. Alternatively, Wat2 might function as the catalytic acid in assisting ring-opening pathway (path B2). However, due to the high-energy pathway, this mechanistic scenario is not very likely.

3.3 Can Wat2 ligand exchange at Zn1 lead to different catalytic route?

From **Figure 6**, it is clear that path A gives quite feasible barrier (19.8 kcal/mol) for the creatinine hydrolysis. The calculated energetics of path A described above agrees well with the available experimental kinetics ($k_{\text{cat}} = 252 \text{ s}^{-1}$ (Yamashita, Nakajima et al. 2010) which can be converted to a barrier of ~ 14.1 kcal/mol according to transition state theory). It is also in line with the recent QM/MM study (Jitonnom, Mujika et al. 2017). However, our QM study suggests one more step is needed in the ring-opening step starting from **IM1-A**.

It can be seen from **Figure 6** that, for path A, the highest barrier of 26.3 kcal/mol (19.8 kcal/mol relative to **R-A**) is found at **TS2-A** for His178 protonation (**IM1-A** $\rightarrow$ **IM2-A**) while the proton transfer from His178 to substrate (**IM2-B1** $\rightarrow$ **P-B1**) is the rate-limiting step for path B1 (**TS3-B1** = 31.9 and 25.4 kcal/mol, relative to **R-B** and **R-A**, respectively). Both path A and B1 share the same intermediate species at **IM2**, which involves a protonated His178 residue. The two species, **IM2-A** and **IM2-B**, have similar energies on the potential energy surface (15.5 vs 16.6 kcal/mol, **Figure 6**). Analysis of relative energies of these pathways using different DFT functional and a bigger basis set (6-311+G(2d,2p)) was also carried out and it shows that these two geometries have a close energy (18.5 ± 2.7 kcal/mol for **IM2-A** and 20.0 ± 2.6 kcal/mol **IM2-B1**). The main difference between **IM2-A** and **IM2-B1** is the geometrical rearrangement of Wat1 and Wat2 with respect to Zn1, via O_{μ} and O_{ω} atoms, respectively ($\text{Zn1-O}_{\mu} = 2.30 \text{ \AA}$, $\text{Zn1-O}_{\omega} = 3.70 \text{ \AA}$ for **IM2-A** and $\text{Zn1-O}_{\mu} = 2.81 \text{ \AA}$, $\text{Zn1-O}_{\omega} = 2.26 \text{ \AA}$ for **IM2-B1**), as shown in **Figure 7**. These data also reflect that this arrangement is necessary to keep Zn1 in a penta-coordination fashion. Hydrogen bond interactions also play an important role and they are found in both structures where Glu122 and Glu34 participate in stabilizing the Wat2. His178, on the other hand, shows different H-bond interactions between the NH- group of His178 and the substrate.

Considering the close energy for two different Zn1 coordination geometries at **IM2-B1** and **IM2-A**, it could be envisioned that another alternative pathway could also be possible, i.e., bound (**R-B** $\rightarrow$ **TS1-B** $\rightarrow$ **IM1-B** $\rightarrow$ **TS2-B1**) $\rightarrow$ unbound (**IM2-A** $\rightarrow$ **TS3-A**) $\rightarrow$ bound (**P-A**). This alternative pathway has similar barriers of 19.1 and 18.1 kcal/mol at **TS2-B1**^{bound} and **TS3-A**^{unbound}, respectively, which indicates that the two mechanistic steps in the ring-opening, occurs sequentially with comparable rates. Such a route was not suggested by our previous QM/MM simulations, possibly due to the weak coordination of Wat2 at Zn1 during the course of reaction (Jitonnom, Mujika et al. 2017).

Now, we look further into the change of coordination number at Zn1 and Zn2. As depicted in **Figure 8**, during the ring-opening step of path B1 and B2, the 5-fold coordination is maintained at Zn1, while the 4-fold coordination in Zn2. However, for

the same step in path A, there is an increasing of Zn1 coordination number from 4 to 6 at **IM1-A**, **IM2-A** and **P-A**, respectively. There is a good correlation between the Zn1 coordination with oxygen atom of Wat2 (O_{Wat2}) and oxygen atom (O_{μ}) of substrate in path A. Meanwhile, path B1 and B2 show a good correlation between the O_w and oxygen atom (O_{μ}) of hydroxide (Wat1). These results also implied that the fast ligand exchange at Zn1 can lead to different coordination modes that yield different several energetically favorable pathways, as described by Brás et al. (Brás, Fernandes et al. 2014).

Overall, this DFT cluster approach not only allows us to examine reaction pathways of this enzyme but also provide important insights into key intermediate species of the ring-opening step, which was found to occur via a stepwise mechanism assisted by His178 (a proton transfer takes places from the bridging hydroxyl group to His178, couple with a rotation of the latter to form a hydrogen bond to the substrate amide nitrogen, which becomes ready for protonation). In other word, the protonation of His178 and the C1-N1 bond cleavage occurs via two transition states (via **TS2** and **TS3**). In the previous QM/MM study, the proton transfer from the bridging hydroxyl group to the amide nitrogen proceeds in a single concerted step.

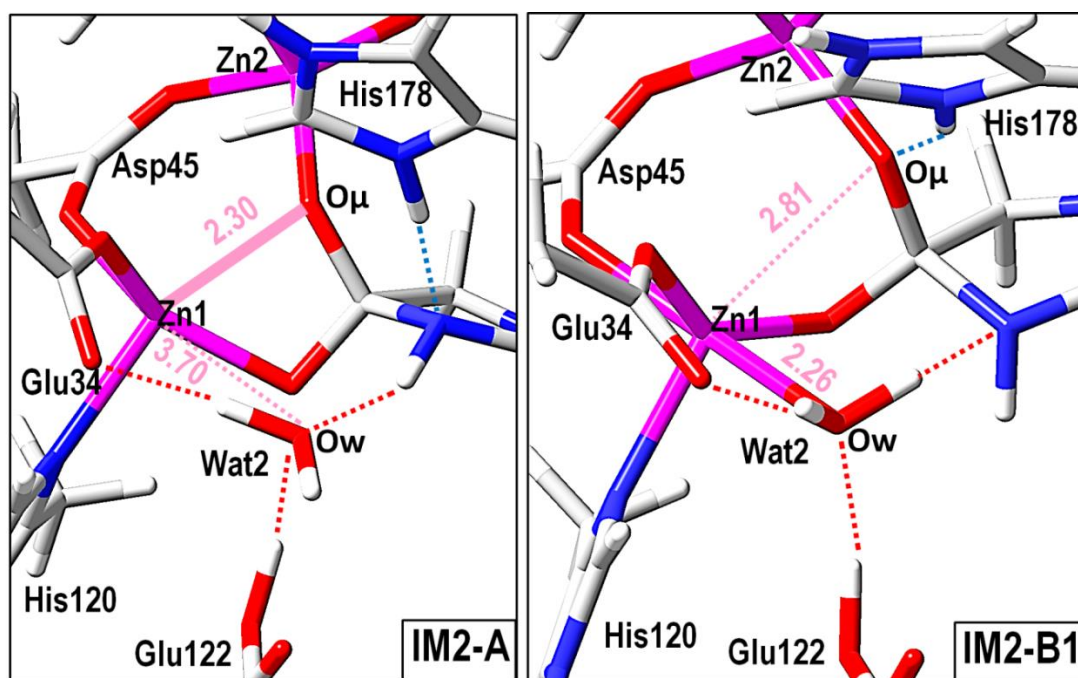


Figure 7. Comparison of structures IM2-A and IM2-B1. Hydrogen bonds originated from Wat2 and His178 were depicted as dotted lines with red and blue colors, respectively. Zn1 coordination distances with O_{μ} (Wat1) and O_w atoms (Wat2) were also indicated. Distances are in angstrom (Å).

3.4. The influence of different functional and basis set

In order to assess the influence of functional on the reaction of creatininase, several different functional (B3LYP, B3LYP-D3, X3LYP, B97-D, B97-2, B1B95, mPWB1K, M06-L, M06-2X, M05-2) were performed using the single-point energy calculations with the larger basis set (6-311+G(2d,2p) (data not shown). As results, the lowest

activation barriers of 22.1 ± 2.4 kcal/mol were obtained for path A, while paths B1 and B2 yielded very large barriers at **TS3** (37.2 ± 2.8 and 37.6 ± 2.3 kcal/mol). These results clearly indicate that path A is the likely mechanism of creatininase, with the proximate energy value to experimental value. Therefore, using different functional and a larger basis set (6-311+G(2d,2p)) do not change significantly the results drawn above and the calculations discussed in the main text are still valid.

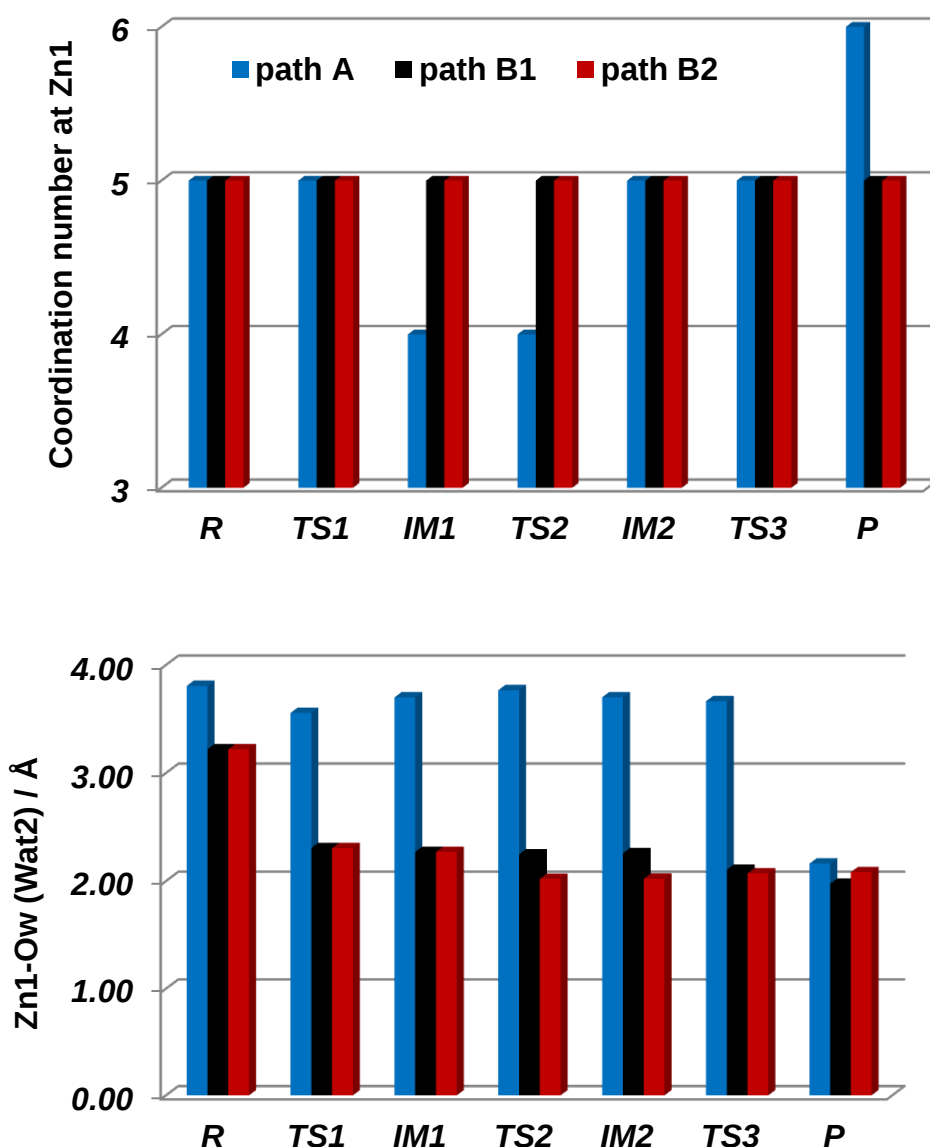


Figure 8. Bar graph showing the changes of coordination number at Zn1 (top) and Zn1-O_w(Wat2) distance (bottom) for paths A, B1, and B2, shown in blue, black and red, respectively. The criteria for the *bound* and *unbound* geometries of Wat2 at Zn1 were measured by the Zn1-O_w distance, having a cut-off of < 3.00 Å and > 3.00 Å, respectively.

CONCLUSIONS

In the present work, the catalytic mechanism of creatininase-catalyzed reaction has been investigated by DFT cluster approach. Three reaction pathways (paths A, B1 and

B2, **Scheme 1**) were modeled using the X-ray crystal structure of the product-bound enzyme as a starting point of the current investigation. It is shown that the reaction mechanism shown in **Scheme 1** (corresponding to path A) is energetically more feasible and is in agreement with available experimental kinetic (14.1 kcal/mol derived from k_{cat} of 252 s^{-1} (Yamashita, Nakajima et al. 2010)) and our recent QM/MM study (Jitonnom, Mujika et al. 2017). Owing to fast exchange of Wat2 ligand at Zn1, an alternative/competitive ring-opening pathway to path A (via **TS2-B1** and **TS3-A**) can occur with His178 and Wat2 serving as a dual role as the catalytic base and the catalytic acid, respectively. The overall mechanism of the creatininase is proposed to follow three elementary steps: nucleophilic attack by the bridging hydroxide. Collapse of the tetrahedral intermediate and cleavage of the C1-N1 bond occur via a sequence of two proton transfer steps with the assistance of His178 or Wat2: the abstraction of a proton from the gem-diolate tetrahedral intermediate by His178 followed by the proton transfer from this protonated histidine to the amide of substrate, resulting in the formation of creatine product. The first proton transfer is the rate-limiting step with the computed barrier of $\sim 19 \text{ kcal/mol}$. Calculations have also confirmed that His178 plays a crucial role in stabilizing the transition state and intermediate during catalysis. The detailed mechanistic understanding in this study might be extended to other members of the cyclic amidohydrolase family and may facilitate future rational design of enzymes for biosensor application.

SUGGESTIONS

- (1) Flexibility of Zn1 will be further investigated by QM/MM MD.
- (2) Replacement of Zn^{2+} at site 1 with Mn^{2+} can be used to evaluate the role of Zn1 on the activity of the enzyme.

References

- Abendroth, J., K. Niefind and D. Schomburg (2002). "X-ray structure of a dihydropyrimidinase from *Thermus* sp. at 1.3 Å resolution." *J. Mol. Biol.* **320**(1): 143-156.
- Aganyants, H., P. Weigel, Y. Hovhannisyan, M. Lecocq, H. Koloyan, A. Hambardzumyan, A. Hovsepyan, J. N. Hallet and V. Sakanyan (2020). "Rational engineering of the substrate specificity of a thermostable d-hydantoinase (dihydropyrimidinase)." *High Throughput* **9**(1).
- Agirre, J. (2017). "Strategies for carbohydrate model building, refinement and validation." *Acta Crystallogr. D Struct. Biol.* **73**(Pt 2): 171-186.
- Ahmadi, S., L. Barrios Herrera, M. Chehelamirani, J. Hostaš, S. Jalife and D. R. Salahub (2018). "Multiscale modeling of enzymes: QM-cluster, QM/MM, and QM/MM/MD: A tutorial review." *International Journal of Quantum Chemistry* **118**(9): e25558.
- Ahumedo, M., J. C. Drosos and R. Vivas-Reyes (2014). "Application of molecular docking and ONIOM methods for the description of interactions between anti-quorum sensing active (AHL) analogues and the *Pseudomonas aeruginosa* LasR binding site." *Mol. BioSyst.* **10**(5): 1162-1171.
- Alhassid, A., A. Ben-David, O. Tabachnikov, D. Libster, E. Naveh, G. Zolotnitsky, Y. Shoham and G. Shoham (2009). "Crystal structure of an inverting GH 43 1,5- α -L-arabinanase from *Geobacillus stearothermophilus* complexed with its substrate." *Biochem. J.* **422**(1): 73-82.
- Ashihara, H., I. A. Ludwig and A. Crozier (2020). Degradation of pyrimidine nucleotides. *Plant nucleotide metabolism - biosynthesis, degradation, and alkaloid formation*, John Wiley & Sons Ltd.: 165-171.
- Barnett, C. B. and K. J. Naidoo (2010). "Ring puckering: a metric for evaluating the accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB carbohydrate QM/MM simulations." *J. Phys. Chem. B* **114**(51): 17142-17154.
- Barone, V. and M. Cossi (1998). "Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model." *J. Phys. Chem. A* **102**(11): 1995-2001.
- Beuth, B., K. Niefind and D. Schomburg (2002). "Crystallization and preliminary crystallographic analysis of creatininase from *Pseudomonas putida*." *Acta Crystallogr D Biol Crystallogr* **58**(8): 1356-1358.
- Beuth, B., K. Niefind and D. Schomburg (2003). "Crystal structure of creatininase from *Pseudomonas putida*: a novel fold and a case of convergent evolution." *J. Mol. Biol.* **332**(1): 287-301.
- Blomberg, M. R. A., T. Borowski, F. Himo, R.-Z. Liao and P. E. M. Siegbahn (2014). "Quantum chemical studies of mechanisms for metalloenzymes." *Chem. Rev.* **114**(7): 3601-3658.
- Boonsri, P., M. Kuno and S. Hannongbua (2011). "Key interactions of the mutant HIV-1 reverse transcriptase/efavirenz: an evidence obtained from ONIOM method." *MedChemComm* **2**(12): 1181-1187.
- Brás, N. F., P. A. Fernandes and M. J. Ramos (2014). "QM/MM study and MD simulations on the hypertension regulator angiotensin-converting enzyme." *ACS Catal.* **4**(8): 2587-2597.

- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus (1983). "CHARMM: a program for macromolecular energy, minimization, and dynamics calculations." J. Comp. Chem. **4**(2): 187-217.
- Brooks III, C. L. and M. Karplus (1983). "Deformable stochastic boundaries in molecular dynamics." J. Chem. Phys. **79**(12): 6312-6325.
- Burmeister, W. P., S. Cottaz, H. Driguez, R. Iori, S. Palmieri and B. Henrissat (1997). "The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase." Structure **5**(5): 663-675.
- Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard and B. Henrissat (2009). "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics." Nucleic Acids Res. **37**(Database issue): D233-238.
- Cassimjee, K. E., B. Manta and F. Himo (2015). "A quantum chemical study of the omega-transaminase reaction mechanism." Org. Biomol. Chem. **13**(31): 8453-8464.
- Charoenwattanasatien, R., S. Pengthaisong, I. Breen, R. Mutoh, S. Sansenya, Y. Hua, A. Tankrathok, L. Wu, C. Songsiriritthigul, H. Tanaka, S. J. Williams, G. J. Davies, G. Kurisu and J. R. Cairns (2016). "Bacterial β -Glucosidase Reveals the Structural and Functional Basis of Genetic Defects in Human Glucocerebrosidase 2 (GBA2)." ACS Chem Biol **11**(7): 1891-1900.
- Cheng, J.-H., C.-C. Huang, Y.-H. Huang and C.-Y. Huang (2018). "Structural basis for pH-dependent oligomerization of dihydropyrimidinase from *Pseudomonas aeruginosa* PAO1." Bioinorg. Chem. Appl. **2018**: 9564391.
- Cheon, Y.-H., H.-S. Kim, K.-H. Han, J. Abendroth, K. Niefind, D. Schomburg, J. Wang and Y. Kim (2002). "Crystal structure of D-hydantoinase from *Bacillus stearothermophilus*: insight into the stereochemistry of enantioselectivity." Biochemistry **41**(30): 9410-9417.
- Cheon, Y.-H., H.-S. Park, S.-C. Lee, D.-E. Lee and H.-S. Kim (2003). "Structure-based mutational analysis of the active site residues of d-hydantoinase." J. Mol. Catal. B: Enzym. **26**(3-6): 217-222.
- Chuenchor, W., S. Pengthaisong, R. C. Robinson, J. Yuvaniyama, J. Svasti and J. R. K. Cairns (2011). "The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase." Journal of Structural Biology **173**(1): 169-179.
- Chung, L. W., W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding and K. Morokuma (2015). "The ONIOM Method and Its Applications." Chemical Reviews **115**(12): 5678-5796.
- Clemente-Jiménez, J. M., S. Martínez-Rodríguez, F. Rodríguez-Vico and F. J. Las Heras-Vázquez (2014). Synergies of chemistry and biochemistry for the production of β -amino acids. Cascade Biocatalysis, Wiley-VCH Verlag GmbH & Co. KGaA: 161-178.
- Cossi, M., N. Rega, G. Scalmani and V. Barone (2003). "Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model." J. Comput. Chem. **24**(6): 669-681.
- Cossi, M., N. Rega, G. Scalmani and V. Barone (2003). "Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model." J. Comput. Chem. **24**(6): 669-681.
- Cui, Q., M. Elstner, E. Kaxiras, T. Frauenheim and M. Karplus (2001). "A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method." J. Phys. Chem. B **105**(2): 569-585.

- de Sanctis, D., J. M. Inacio, P. F. Lindley, I. de Sa-Nogueira and I. Bento (2010). "New evidence for the role of calcium in the glycosidase reaction of GH43 arabinanases." FEBS J. **277**(21): 4562-4574.
- Dopitová, R., P. Mazura, L. Janda, R. Chaloupková, P. Jerábek, J. Damborský, T. Filipi, N. S. Kiran and B. Brzobohatý (2008). "Functional analysis of the aglycone-binding site of the maize beta-glucosidase Zm-p60.1." Febs j **275**(24): 6123-6135.
- Dudley, K. H., T. C. Butler and D. L. Bius (1974). "The role of dihydropyrimidinase in the metabolism of some hydantoin and succinimide drugs." Drug Metab. Dispos. **2**(2): 103-112.
- Elstner, M., D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert (1998). "Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties." Phys. Rev. B **58**(11): 7260-7268.
- Engel, U., C. Syltatk and J. Rudat (2012). "Stereoselective hydrolysis of aryl-substituted dihydropyrimidines by hydantoinases." Appl. Microbiol. Biotechnol. **94**(5): 1221-1231.
- Fernandes, H. S., C. S. S. Teixeira, S. F. Sousa and N. M. F. S. A. Cerqueira (2019). "Formation of unstable and very reactive chemical species catalyzed by metalloenzymes: a mechanistic overview." Molecules **24**(13).
- Field, M. J., P. A. Bash and M. Karplus (1990). "A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations." J. Comput. Chem. **11**(6): 700-733.
- Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox (2009). Gaussian 09 Rev. A.02. Wallingford, CT.
- Fujimoto, Z., H. Ichinose, T. Maehara, M. Honda, M. Kitaoka and S. Kaneko (2010). "Crystal structure of an exo-1,5- α -l-arabinofuranosidase from *Streptomyces avermitilis* provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43." J. Biol. Chem. **285**(44): 34134-34143.
- Geronimo, I., C. M. Payne and M. Sandgren (2018). "The role of catalytic residue pKa on the hydrolysis/transglycosylation partition in family 3 β -glucosidases." Organic & Biomolecular Chemistry **16**(2): 316-324.
- Goddard-Borger, E. D., R. Carapito, J. M. Jeltsch, V. Phalip, R. V. Stick and A. Varrot (2011). " α -L-arabinofuranosylated pyrrolidines as arabinanase inhibitors." Chem. Comm. **47**(34): 9684-9686.
- Gojkovic, Z., L. Rislund, B. Andersen, M. P. Sandrini, P. F. Cook, K. D. Schnackerz and J. Piskur (2003). "Dihydropyrimidine amidohydrolases and dihydroorotases share the same origin and several enzymatic properties." Nucleic Acids Res. **31**(6): 1683-1692.
- Grácz, É., T. Szimler, A. Garamszegi, P. V. Konarev, A. Lábás, J. Oláh, A. Palló, D. I. Svergun, A. Merli, P. Závodszy, M. S. Weiss and M. Vas (2016). "Dual role of

- the active site residues of *Thermus thermophilus* 3-isopropylmalate dehydrogenase: chemical catalysis and domain closure." *Biochemistry* **55**(3): 560-574.
- Grimme, S., J. Antony, S. Ehrlich and H. Krieg (2010). "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu." *J Chem Phys* **132**(15): 154104.
- Grimme, S., S. Ehrlich and L. Goerigk (2011). "Effect of the damping function in dispersion corrected density functional theory." *J Comput Chem* **32**(7): 1456-1465.
- Heras-Vazquez, F. J., J. M. Clemente-Jimenez, S. Martinez-Rodriguez and F. Rodriguez-Vico (2012). "Engineering cyclic amidases for non-natural amino acid synthesis." *Methods Mol. Biol.* **794**: 87-104.
- Hermann, J. C., C. Hensen, L. Ridder, A. J. Mulholland and H.-D. Höltje (2005). "Mechanisms of antibiotic resistance: QM/MM modeling of the acylation reaction of a class A beta-lactamase with benzylpenicillin." *J. Am. Chem. Soc.* **127**(12): 4454-4465.
- Himo, F. (2006). "Quantum chemical modeling of enzyme active sites and reaction mechanisms." *Theor. Chem. Acc.* **116**(1): 232-240.
- Himo, F. (2017). "Recent trends in quantum chemical modeling of enzymatic reactions." *J. Am. Chem. Soc.* **139**(20): 6780-6786.
- Himo, F. (2017). "Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions." *Journal of the American Chemical Society* **139**(20): 6780-6786.
- Holliday, G. L., G. J. Bartlett, D. E. Almonacid, N. M. O'Boyle, P. Murray-Rust, J. M. Thornton and J. B. O. Mitchell (2005). "MACiE: a database of enzyme reaction mechanisms." *Bioinformatics* **21**(23): 4315-4316.
- Holliday, G. L., J. B. Mitchell and J. M. Thornton (2009). "Understanding the functional roles of amino acid residues in enzyme catalysis." *J Mol Biol* **390**(3): 560-577.
- Holm, L. and C. Sander (1997). "An evolutionary treasure: unification of a broad set of amidohydrolases related to urease." *Proteins* **28**(1): 72-82.
- Hopmann, K. H., B. M. Hallberg and F. Himo (2005). "Catalytic mechanism of limonene epoxide hydrolase, a theoretical study." *J Am Chem Soc* **127**(41): 14339-14347.
- Hsieh, Y.-C., M.-C. Chen, C.-C. Hsu, S. I. Chan, Y.-S. Yang and C.-J. Chen (2013). "Crystal structures of vertebrate dihydropyrimidinase and complexes from *Tetraodon nigroviridis* with lysine carbamylation: metal and structural requirements for post-translational modification and function. ." *J. Biol. Chem.* **288**(42): 30645-30658.
- Huang, C.-Y. (2015). "Inhibition of a putative dihydropyrimidinase from *Pseudomonas aeruginosa* PAO1 by flavonoids and substrates of cyclic amidohydrolases." *PLoS One* **10**(5): e0127634.
- Huang, C. Y., C. C. Hsu, M. C. Chen and Y. S. Yang (2009). "Effect of metal binding and posttranslational lysine carboxylation on the activity of recombinant hydantoinase." *J. Biol. Inorg. Chem.* **14**(1): 111-121.
- Huang, Y.-H., Y. Lien, J.-H. Chen, E.-S. Lin and C.-Y. Huang (2020). "Identification and characterization of dihydropyrimidinase inhibited by plumbagin isolated from *Nepenthes miranda* extract." *Biochimie* **171-172**: 124-135.
- Huang, Y.-H., Z.-J. Ning and C.-Y. Huang (2019). "Crystal structure of dihydropyrimidinase in complex with anticancer drug 5-fluorouracil." *Biochem. Biophys. Res. Commun.* **519**(1): 160-165.
- Ito, K., N. Kanada, T. Inoue, K. Furukawa, K. Yamashita, N. Tanaka, K. T. Nakamura, Y. Nishiya, A. Sogabe and T. Yoshimoto (2002). "Preliminary crystallographic studies of the creatinine amidohydrolase from *Pseudomonas putida*." *Acta Crystallogr. Sect. D.* **58**: 2180-2181.

- Jafari, S., U. Ryde and M. Irani (2016). "Catalytic mechanism of human glyoxalase I studied by quantum-mechanical cluster calculations." J. Mol. Catal. B: Enzym. **131**: 18-30.
- Jitonnorn, J. and S. Hannongbua (2018). "Theoretical study of the arabinan hydrolysis by an inverting GH43 arabinanase." Mol. Simul. **44**(8): 631-637.
- Jitonnorn, J., J. R. Ketudat-Cairns and S. Hannongbua (2018). "QM/MM modeling of the hydrolysis and transfructosylation reactions of fructosyltransferase from *Aspergillus japonicus*, an enzyme that produces prebiotic fructooligosaccharide." J. Mol. Graph. Model. **79**: 175-184.
- Jitonnorn, J., V. S. Lee, P. Nimmanpipug, H. A. Rowlands and A. J. Mulholland (2011). "Quantum mechanics/molecular mechanics modeling of substrate-assisted catalysis in family 18 chitinases: conformational changes and the role of Asp142 in catalysis in ChiB." Biochemistry **50**(21): 4697-4711.
- Jitonnorn, J., V. S. Lee, P. Nimmanpipug, H. A. Rowlands and A. J. Mulholland (2011). "Quantum mechanics/molecular mechanics modeling of substrate-assisted catalysis in family 18 chitinases: conformational changes and the role of Asp142 in catalysis in ChiB." Biochemistry **50**(21): 4697-4711.
- Jitonnorn, J., M. A. Limb and A. J. Mulholland (2014). "QM/MM free-energy simulations of reaction in *Serratia marcescens* Chitinase B reveal the protonation state of Asp142 and the critical role of Tyr214." J. Phys. Chem. B **118**(18): 4771-4783.
- Jitonnorn, J., J. I. Mujika, M. W. van der Kamp and A. J. Mulholland (2017). "Quantum Mechanics/Molecular Mechanics Simulations Identify the Ring-Opening Mechanism of Creatininase." Biochemistry **56**(48): 6377-6388.
- Jitonnorn, J., C. Sattayanon, N. Kungwan and S. Hannongbua (2015). "A DFT study of the unusual substrate-assisted mechanism of *Serratia marcescens* chitinase B reveals the role of solvent and mutational effect on catalysis." J. Mol. Graph. Model. **56**: 53-59.
- Jitonnorn, J., C. Sattayanon, N. Kungwan and S. Hannongbua (2015). "A DFT study of the unusual substrate-assisted mechanism of *Serratia marcescens* chitinase B reveals the role of solvent and mutational effect on catalysis." J. Mol. Graph. Model. **56**: 53-59.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein (1983). "Comparison of simple potential functions for simulating liquid water." J. Chem. Phys. **79**(2): 926.
- Kaplan, A. and L. L. Szabo (1974). "Creatinine hydrolase and creatine amidinohydrolase. II. Partial purification and properties." Mol. Cell. Biochem. **3**: 17-25.
- Ketudat Cairns, J. R., B. Mahong, S. Baiya and J. S. Jeon (2015). "β-Glucosidases: Multitasking, moonlighting or simply misunderstood?" Plant Sci **241**: 246-259.
- Kim, H. and W. N. Lipscomb (1993). "Differentiation and identification of the two catalytic metal binding sites in bovine lens leucine aminopeptidase by x-ray crystallography." Proc. Natl. Acad. Sci. USA **90**(11): 5006-5010.
- Krivitskaya, A. V., M. G. Khrenova and A. V. Nemukhin (2021). "Two Sides of Quantum-Based Modeling of Enzyme-Catalyzed Reactions: Mechanistic and Electronic Structure Aspects of the Hydrolysis by Glutamate Carboxypeptidase." Molecules **26**(20): 6280.
- Kumar, S., J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman (1992). "THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method." J. Comput. Chem. **13**(8): 1011-1021.
- Kurakata, Y., A. Uechi, H. Yoshida, S. Kamitori, Y. Sakano, A. Nishikawa and T. Tonozuka (2008). "Structural insights into the substrate specificity and function of

- Escherichia coli K12 YgjK, a glucosidase belonging to the glycoside hydrolase family 63." *J Mol Biol* **381**(1): 116-128.
- Labas, A., E. Szabo, L. Mones and M. Fuxreiter (2013). "Optimization of reorganization energy drives evolution of the designed Kemp eliminase KE07." *Biochim. Biophys. Acta Proteins Proteom.* **1834**(5): 908-917.
- Lee, M., M. J. Maher, R. I. Christopherson and J. M. Guss (2007). "Kinetic and structural analysis of mutant Escherichia coli dihydroorotases: a flexible loop stabilizes the transition state." *Biochemistry* **46**(37): 10538-10550.
- Lence, E., M. W. van der Kamp, C. González-Bello and A. J. Mulholland (2018). "QM/MM simulations identify the determinants of catalytic activity differences between type II dehydroquinase enzymes." *Org. Biomol. Chem.* **16**(24): 4443-4455.
- Liao, R.-Z., F. Himo, J.-G. Yu and R.-Z. Liu (2010). "Dipeptide hydrolysis by the dinuclear zinc enzyme human renal dipeptidase: Mechanistic insights from DFT calculations." *Journal of Inorganic Biochemistry* **104**(1): 37-46.
- Liao, R.-Z., J.-G. Yu and F. Himo (2009). "Reaction Mechanism of the Dinuclear Zinc Enzyme N-Acyl-L-homoserine Lactone Hydrolase: A Quantum Chemical Study." *Inorganic Chemistry* **48**(4): 1442-1448.
- Liao, R.-Z., J.-G. Yu and F. Himo (2010). "Reaction Mechanism of the Trinuclear Zinc Enzyme Phospholipase C: A Density Functional Theory Study." *The Journal of Physical Chemistry B* **114**(7): 2533-2540.
- Liao, R. Z., J. G. Yu, F. M. Raushel and F. Himo (2008). "Theoretical investigation of the reaction mechanism of the dinuclear zinc enzyme dihydroorotase." *Chem. Eur. J.* **14**(14): 4287-4292.
- Lind, M. E. S. and F. Himo (2013). "Quantum chemistry as a tool in asymmetric biocatalysis: limonene epoxide hydrolase test case." *Angew. Chem. Int. Ed.* **52**(17): 4563-4567.
- Lind, M. E. S. and F. Himo (2014). "Theoretical study of reaction mechanism and stereoselectivity of arylmalonate decarboxylase." *ACS Catal.* **4**(11): 4153-4160.
- Lohkamp, B., B. Andersen, J. Piskur and D. Dobritzsch (2006). "The crystal structures of dihydropyrimidinases reaffirm the close relationship between cyclic amidohydrolases and explain their substrate specificity." *J. Biol. Chem.* **281**(19): 13762-13776.
- Lombard, V., H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat (2014). "The carbohydrate-active enzymes database (CAZy) in 2013." *Nucleic Acids Res* **42**(Database issue): D490-495.
- Lonsdale, R., J. N. Harvey and A. J. Mulholland (2010). "Inclusion of dispersion effects significantly improves accuracy of calculated reaction barriers for cytochrome P450 catalyzed reactions." *J. Phys. Chem. Lett.* **1**(21): 3232-3237.
- Lonsdale, R., J. N. Harvey and A. J. Mulholland (2012). "Effects of dispersion in density functional based quantum mechanical/molecular mechanical calculations on cytochrome P450 catalyzed reactions." *J. Chem. Theory Comput.* **8**(11): 4637-4645.
- Lu, X., D. Fang, S. Ito, Y. Okamoto, V. Ovchinnikov and Q. Cui (2016). "QM/MM free energy simulations: recent progress and challenges." *Mol. Simul.* **42**(13): 1056-1078.
- M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J.

- Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox (2009). Gaussian 09, Revision A.02; Gaussian, Inc.: Wallingford, CT.
- MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus (1998). "All-atom empirical potential for molecular modeling and dynamics studies of proteins." J. Phys. Chem. B **102**(18): 3586-3616.
- Manta, B., F. M. Raushel and F. Himo (2014). "Reaction mechanism of zinc-dependent cytosine deaminase from *Escherichia coli*: a quantum-chemical study." J Phys Chem B **118**(21): 5644-5652.
- Martínez-Gómez, A. I., J. M. Clemente-Jiménez, F. Rodríguez-Vico, L. T. Kanerva, X.-G. Li, F. J. L. Heras-Vázquez and S. Martínez-Rodríguez (2012). "New biocatalytic route for the production of enantioenriched β -alanine derivatives starting from 5- and 6-monosubstituted dihydouracils." Process Biochem. **47**(12): 2090-2096.
- Martínez-Rodríguez, S., A. I. Martínez-Gómez, J. M. Clemente-Jiménez, F. Rodríguez-Vico, J. M. García-Ruiz, F. J. Las Heras-Vázquez and J. A. Gavira (2010). "Structure of dihydropyrimidinase from *Sinorhizobium meliloti* CECT4114: New features in an amidohydrolase family member." J. Struct. Biol. **169**(2): 200-208.
- McCarter, J. D. and G. Stephen Withers (1994). "Mechanisms of enzymatic glycoside hydrolysis." Current Opinion in Structural Biology **4**(6): 885-892.
- McCarthy, J. K., A. Uzelac, D. F. Davis and D. E. Eveleigh (2004). "Improved catalytic efficiency and active site modification of 1,4-beta-D-glucan glucohydrolase A from *Thermotoga neapolitana* by directed evolution." J Biol Chem **279**(12): 11495-11502.
- Mulliken, R. S. (1955). "Electronic population analysis on LCAO-MO molecular wave functions. I." J. Chem. Phys. **23**(10): 1833-1840.
- Nam, S. H., H. S. Park and H. S. Kim (2005). "Evolutionary relationship and application of a superfamily of cyclic amidohydrolase enzymes." Chem. Rec. **5**(5): 298-307.
- Nurizzo, D., J. P. Turkenburg, S. J. Charnock, S. M. Roberts, E. J. Dodson, V. A. McKie, E. J. Taylor, H. J. Gilbert and G. J. Davies (2002). "Cellvibrio japonicus alpha-L-arabinanase 43A has a novel five-blade beta-propeller fold." Nat. Struct. Biol. **9**(9): 665-668.
- Olsson, M. H., C. R. Sondergaard, M. Rostkowski and J. H. Jensen (2011). "PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions." J. Chem. Theory Comput. **7**(2): 525-537.
- Park, J.-M., M.-U. Jang, J.-H. Kang, M.-J. Kim, S.-W. Lee, Y. B. Song, C.-S. Shin, N. S. Han and T.-J. Kim (2012). "Detailed modes of action and biochemical characterization of endo-arabinanase from *Bacillus licheniformis* DSM13." J. Microbiol. **50**(6): 1041-1046.
- Peng, W.-F. and C.-Y. Huang (2014). "Allantoinase and dihydroorotase binding and inhibition by flavonols and the substrates of cyclic amidohydrolases." Biochimie **101**: 113-122.

- Pengthaisong, S., C.-F. Chen, S. G. Withers, B. Kuaprasert and J. R. Ketudat Cairns (2012). "Rice BGlu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition." Carbohydrate Research **352**: 51-59.
- Proctor, M. R., E. J. Taylor, D. Nurizzo, J. P. Turkenburg, R. M. Lloyd, M. Vardakou, G. J. Davies and H. J. Gilbert (2005). "Tailored catalysts for plant cell-wall degradation: redesigning the exo/endo preference of *Cellvibrio japonicus* arabinanase 43A." Proc. Natl. Acad. Sci. USA **102**(8): 2697-2702.
- Reidl, C. T., R. Mascarenhas, T. S. H. Mohammad, M. R. Lutz, Jr., P. W. Thomas, W. Fast, D. Liu and D. P. Becker (2021). "Cyclobutanone Inhibitor of Cobalt-Functionalized Metallo- γ -Lactonase AiiA with Cyclobutanone Ring Opening in the Active Site." ACS omega **6**(21): 13567-13578.
- Rikitake, K., I. Oka, M. Ando, T. Yoshimoto and D. Tsuru (1979). "Creatinine amidohydrolase (creatininase) from *Pseudomonas putida*. Purification and some properties." J. Biochem. **86**: 1109-1117.
- Rong-Zhen Liao, Jian-Guo Yu, Frank M. Raushel and F. Himo (2008). "Theoretical investigation of the reaction mechanism of the dinuclear zinc enzyme dihydroorotase." Chem. Eur. J. **14**: 4287 – 4292.
- Rowlett, R. S., C. Tu, M. M. McKay, J. R. Preiss, R. J. Loomis, K. A. Hicks, R. J. Marchione, J. A. Strong, G. S. Donovan, Jr. and J. E. Chamberlin (2002). "Kinetic characterization of wild-type and proton transfer-impaired variants of beta-carbonic anhydrase from *Arabidopsis thaliana*." Arch. Biochem. Biophys. **404**(2): 197-209.
- Ryckaert, J.-P., G. Ciccotti and H. J. C. Berendsen (1977). "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes." J. Comp. Phys. **23**(3): 327-341.
- Ryde, U. and P. Söderhjelm (2016). "Ligand-binding affinity estimates supported by quantum-mechanical methods." Chem. Rev. **116**(9): 5520-5566.
- Saen-oon, S., M. Kuno and S. Hannongbua (2005). "Binding energy analysis for wild-type and Y181C mutant HIV-1 RT/8-Cl TIBO complex structures: quantum chemical calculations based on the ONIOM method." Proteins **61**(4): 859-869.
- Schnackerz, K. D. and D. Dobritzsch (2008). "Amidohydrolases of the reductive pyrimidine catabolic pathway purification, characterization, structure, reaction mechanisms and enzyme deficiency." Biochim. Biophys. Acta **1784**(3): 431-444.
- Schumann, J., G. Böhm, G. Schumacher, R. Rudolph and R. Jaenicke (1993). "Stabilization of creatinase from *Pseudomonas putida* by random mutagenesis." Protein Sci **2**(10): 1612-1620.
- Seabra, G. d. M., R. C. Walker, M. Elstner, D. A. Case and A. E. Roitberg (2007). "Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package." J. Phys. Chem. A **111**(26): 5655-5664.
- Seibert, C. M. and F. M. Raushel (2005). "Structural and catalytic diversity within the amidohydrolase superfamily." Biochemistry **44**(17): 6383-6391.
- Siegbahn, P. E. and F. Himo (2009). "Recent developments of the quantum chemical cluster approach for modeling enzyme reactions." J Biol Inorg Chem **14**(5): 643-651.
- Siegbahn, P. E. M. and F. Himo (2009). "Recent developments of the quantum chemical cluster approach for modeling enzyme reactions." J. Biol. Inorg. Chem. **14**(5): 643-651.
- Siegbahn, P. E. M. and F. Himo (2011). "The quantum chemical cluster approach for modeling enzyme reactions." Wiley Interdisciplinary Reviews: Computational Molecular Science **1**(3): 323-336.
- Siegbahn, P. E. M. and F. Himo (2011). "The quantum chemical cluster approach for modeling enzyme reactions." Wiley Interdiscip. Rev. Comput. Mol. Sci. **1**(3): 323-336.

- Slomka, C., S. Zhong, A. Fellingner, U. Engel, C. Syltatk, S. Brase and J. Rudat (2015). "Chemical synthesis and enzymatic, stereoselective hydrolysis of a functionalized dihydropyrimidine for the synthesis of beta-amino acids." AMB Express **5**(1): 85.
- Song, Z., Y. Yue, S. Feng, H. Sun, Y. Li, F. Xu, Q. Zhang and W. Wang (2020). "Cysteine dioxygenase catalyzed CF bond cleavage: An in silico approach." Chem. Phys. Lett **750**: 137449.
- Stortz, C. A. and A. M. Sarotti (2019). "Exhaustive exploration of the conformational landscape of mono- and disubstituted five-membered rings by DFT and MP2 calculations." RSC Adv. **9**(42): 24134-24145.
- Syltatk, C., O. May, J. Altenbuchner, R. Mattes and M. Siemann (1999). "Microbial hydantoinases--industrial enzymes from the origin of life?" Appl. Microbiol. Biotechnol. **51**(3): 293-309.
- Szulmajster, J. (1958). "Bacterial degradation of creatinine. II. Creatinine desimidase." Biochim Biophys Acta **30**(1): 154-163.
- Tang, T. Y. and C. J. Wen (2000). "Expression of the creatininase gene from *Pseudomonas putida* RS65 in *Escherichia coli*." J. Ind. Microbiol. **24**: 2-6.
- Thoden, J. B., G. N. Phillips, Jr., T. M. Neal, F. M. Raushel and H. M. Holden (2001). "Molecular structure of dihydroorotase: a paradigm for catalysis through the use of a binuclear metal center." Biochemistry **40**(24): 6989-6997.
- Tian, B., F. Wallrapp, C. Kalyanaraman, S. Zhao, L. A. Eriksson and M. P. Jacobson (2013). "Predicting enzyme-substrate specificity with QM/MM methods: a case study of the stereospecificity of (D)-glucarate dehydratase." Biochemistry **52**(33): 5511-5513.
- Tsuchiya, H., T. Akiyama, T. Kuhara, Y. Nakajima, M. Ohse, H. Kurahashi, T. Kato, Y. Maeda, H. Yoshinaga and K. Kobayashi (2019). "A case of dihydropyrimidinase deficiency incidentally detected by urine metabolome analysis." Brain Dev. **41**(3): 280-284.
- Tsuru, D., I. Oka and T. Yoshimoto (1976). "Creatinine decomposing enzymes in *Pseudomonas putida*." Agric. Biol. Chem. **40**: 1011-1018.
- Tzeng, C.-T., Y.-H. Huang and C.-Y. Huang (2016). "Crystal structure of dihydropyrimidinase from *Pseudomonas aeruginosa* PAO1: Insights into the molecular basis of formation of a dimer." Biochem. Biophys. Res. Commun. **478**(3): 1449-1455.
- van Kuilenburg, A. B. P., R. Meinsma, B. A. Zonnenberg, L. Zoetekouw, F. Baas, K. Matsuda, N. Tamaki and A. H. van Gennip (2003). "Dihydropyrimidinase deficiency and severe 5-fluorouracil toxicity." Clin. Cancer Res. **9**(12): 4363-4367.
- van Rijssel, E. R., P. van Delft, G. Lodder, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov and J. D. C. Codée (2014). "Furanosyl oxocarbenium ion stability and stereoselectivity." Angew. Chem. Int. Ed. **53**(39): 10381-10385.
- Vocadlo, D. J. and G. J. Davies (2008). "Mechanistic insights into glycosidase chemistry." Curr Opin Chem Biol **12**(5): 539-555.
- Vriend, G. (1990). "WHAT IF: A molecular modeling and drug design program." J. Mol. Graph. **8**(1): 52-56.
- Wan, Q., J. M. Parks, B. L. Hanson, S. Z. Fisher, A. Ostermann, T. E. Schrader, D. E. Graham, L. Coates, P. Langan and A. Kovalevsky (2015). "Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography." Proc. Natl. Acad. Sci. U.S.A. **112**(40): 12384.
- Warshel, A. and M. Levitt (1976). "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme." J. Mol. Biol. **103**(2): 227-249.

- Warshel, A., P. K. Sharma, M. Kato, Y. Xiang, H. Liu and M. H. Olsson (2006). "Electrostatic basis for enzyme catalysis." Chem. Rev. **106**(8): 3210-3235.
- Woodcock, H. L., M. Hodošček and B. R. Brooks (2007). "Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms." J. Phys. Chem. A **111**(26): 5720-5728.
- Wyss, M. and R. Kaddurah-Daouk (2000). "Creatine and creatinine metabolism." Physiol Rev **80**(3): 1107-1213.
- Yamaguchi, A., T. Tada, K. Wada, T. Nakaniwa, T. Kitatani, Y. Sogabe, M. Takao, T. Sakai and K. Nishimura (2005). "Structural basis for thermostability of endo-1,5-alpha-L-arabinanase from *Bacillus thermodenitrificans* TS-3." J. Biochem. **137**(5): 587-592.
- Yamamoto, K. and M. Oka (1995). "Cloning of the creatinine amidohydrolase gene from *Pseudomonas* sp. PS-7." Biosci. Biotechnol. Biochem. **59**: 1331-1332.
- Yamashita, K., Y. Nakajima, H. Matsushita, Y. Nishiya, R. Yamazawa, Y. F. Wu, F. Matsubara, H. Oyama, K. Ito and T. Yoshimoto (2010). "Substitution of Glu122 by glutamine revealed the function of the second water molecule as a proton donor in the binuclear metal enzyme creatininase." J. Mol. Biol. **396**(4): 1081-1096.
- Yilmazer, N. D. and M. Korth (2016). "Recent progress in treating Protein-Ligand interactions with quantum-mechanical methods." Int. J. Mol. Sci. **17**(5).
- Yoshimoto, T., I. Oka and D. Tsuru (1976). "Creatine amidinohydrolase of *Pseudomonas putida*: Crystallization and some properties." Arch. Biochem. Biophys **177**(2): 508-515.
- Yoshimoto, T., N. Tanaka, N. Kanada, T. Inoue, Y. Nakajima, M. Haratake, K. T. Nakamura, Y. Xu and K. Ito (2004). "Crystal structures of creatininase reveal the substrate binding site and provide an insight into the catalytic mechanism." J. Mol. Biol. **337**(2): 399-416.
- Zechel, D. L. and S. G. Withers (2000). "Glycosidase mechanisms: anatomy of a finely tuned catalyst." Acc. Chem. Res. **33**(1): 11-18.
- Zechel, D. L. and S. G. Withers (2000). "Glycosidase mechanisms: anatomy of a finely tuned catalyst." Acc Chem Res **33**(1): 11-18.
- Zhang, R., X. Shi, Y. Sun, Q. Zhang and W. Wang (2018). "Insights into the catalytic mechanism of dehydrogenase BphB: A quantum mechanics/molecular mechanics study." Chemosphere **208**: 69-76.
- Zhang, X., S. Wang, X. Wu, S. Liu, D. Li, H. Xu, P. Gao, G. Chen and L. Wang (2015). "Subsite-specific contributions of different aromatic residues in the active site architecture of glycoside hydrolase family 12." Scientific Reports **5**(1): 18357.
- Zhu, L., X. Tang, Y. Li, R. Zhang, J. Wang, Q. Zhang and W. Wang (2018). "QM/MM study of the reaction mechanism of Cl-cis,cis-muconate with muconate lactonizing enzyme." Bioorg. Chem. **80**: 453-460.
- Zhu, X., A. Barman, M. Ozbil, T. Zhang, S. Li and R. Prabhakar (2012). "Mechanism of peptide hydrolysis by co-catalytic metal centers containing leucine aminopeptidase enzyme: a DFT approach." J Biol Inorg Chem **17**(2): 209-222.
- Zhu, X., A. Barman, M. Ozbil, T. Zhang, S. Li and R. Prabhakar (2012). "Mechanism of peptide hydrolysis by co-catalytic metal centers containing leucine aminopeptidase enzyme: a DFT approach." J. Biol. Inorg. Chem. **17**(2): 209-222.

Appendix

**Additional Papers Supported by
Thailand Research Fund
(RSA6280104)**
