ACS Catalysis

Letter

pubs.acs.org/acscatalysis

Estrogen Formation via H-Abstraction from the O−H Bond of *gem*-Diol by Compound I in the Reaction of CYP19A1: Mechanistic Scenario Derived from Multiscale QM/MM Calculations

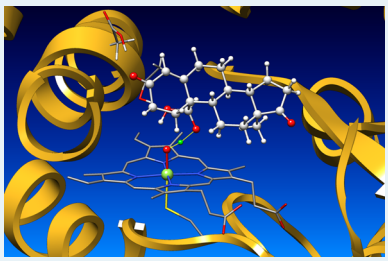
Kai Xu,[†] Yong Wang,*[,‡] and Hajime Hirao*[,†]

[†]Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371

[‡]State Key Laboratory for Oxo Synthesis and Selective Oxidation, Suzhou Research Institute of LICP, Lanzhou Institute of Chemical Physics (LICP), Chinese Academy of Sciences, Lanzhou 730000, China

Supporting Information

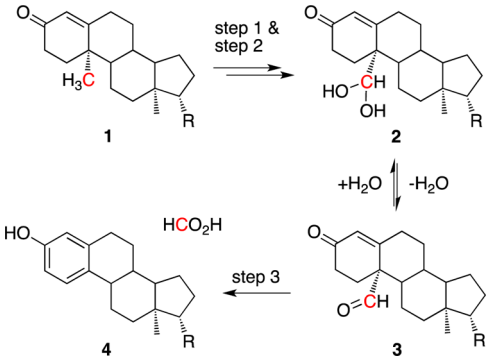
ABSTRACT: Recent experiments suggested that, contrary to traditional belief, the third step of aromatase-catalyzed estrogen formation should be effected by compound I (Cpd I), rather than by ferric peroxide. We performed QM/MM calculations to address the question of how Cpd I drives the aromatization process. Surprisingly, the calculations show that the reaction begins with hydrogen abstraction from the O−H bond of a *gem*-diol substrate, which is followed by barrierless homolytic C−C bond cleavage and then 1β-H-abstraction. Proton-coupled electron transfer enables the cleavage of the strong O−H bond. Another product, carboxylic acid, can be formed from either the *gem*-diol or aldehyde.



KEYWORDS: *aromatase, Cpd I, QM/MM, H-abstraction, PCET*

In view of their remarkable versatility and ubiquity, cytochrome P450 enzymes (P450s) are clearly among the most important metalloenzymes that exist in nature.[1−5] For us humans, xenobiotic metabolism and hormone biosynthesis are the two major functions of P450s, and a specific P450 called aromatase (CYP19A1) is responsible for the latter. As shown in Scheme 1, aromatase catalyzes the oxidative conversion of androgens such as androstenedione and testosterone (**1**) into the respective aromatized estrogens (**4**) and HCO_2H in three steps, each of which requires O_2 and NADPH.[6−8] Because only aromatase can catalyze these essential biosynthetic steps, the enzyme has been a promising drug target for the treatment of breast cancer.

Scheme 1. A Possible Reaction Sequence for CYP19A1-Catalyzed Estrogen Formation[a]



[a]The red carbon denotes C19.

In the aromatase-catalyzed reaction of androgens (Scheme 1), there is now little doubt that compound I (Cpd I)[9,10] acts as the reactive species in the first two hydroxylation steps (steps 1 and 2 in Scheme 1). By contrast, there is ongoing debate regarding the nature of the reactive species involved in step 3. Cpd I and ferric peroxide have been implicated as likely candidates. Until recently, the most convincing mechanism involved nucleophilic attack of the ferric-peroxo species on aldehyde **3**. This ferric-peroxide-driven mechanism was supported by the ^{18}O-labeling experiments performed by Akhtar et al.[11] and by Caspi et al.,[12] as well as by the theoretical calculations performed by Senn and Hackett.[13]

However, recent experimental studies using the techniques of resonance Raman spectroscopy (Mak et al.)[14] and kinetic solvent isotope effect (Kharti et al.)[15] implicated the involvement of Cpd I. More recently, Yoshimoto and Guengerich performed ^{18}O-labeling experiments with more sensitive techniques and instruments than had been used previously for such studies and demonstrated that the aromatization mechanism could be rationalized only in terms of Cpd I.[16] Surprisingly, in the latter study, no $^{18}O_2$-derived oxygen atom was found in the produced formic acid. Thus, the ferric-peroxide mechanism, in which ^{18}O is inevitably incorporated into the formic acid, could not explain this

Received: March 10, 2015
Revised: May 29, 2015
Published: June 3, 2015

DOI: 10.1021/acscatal.5b00510
ACS Catal. 2015, 5, 4175−4179

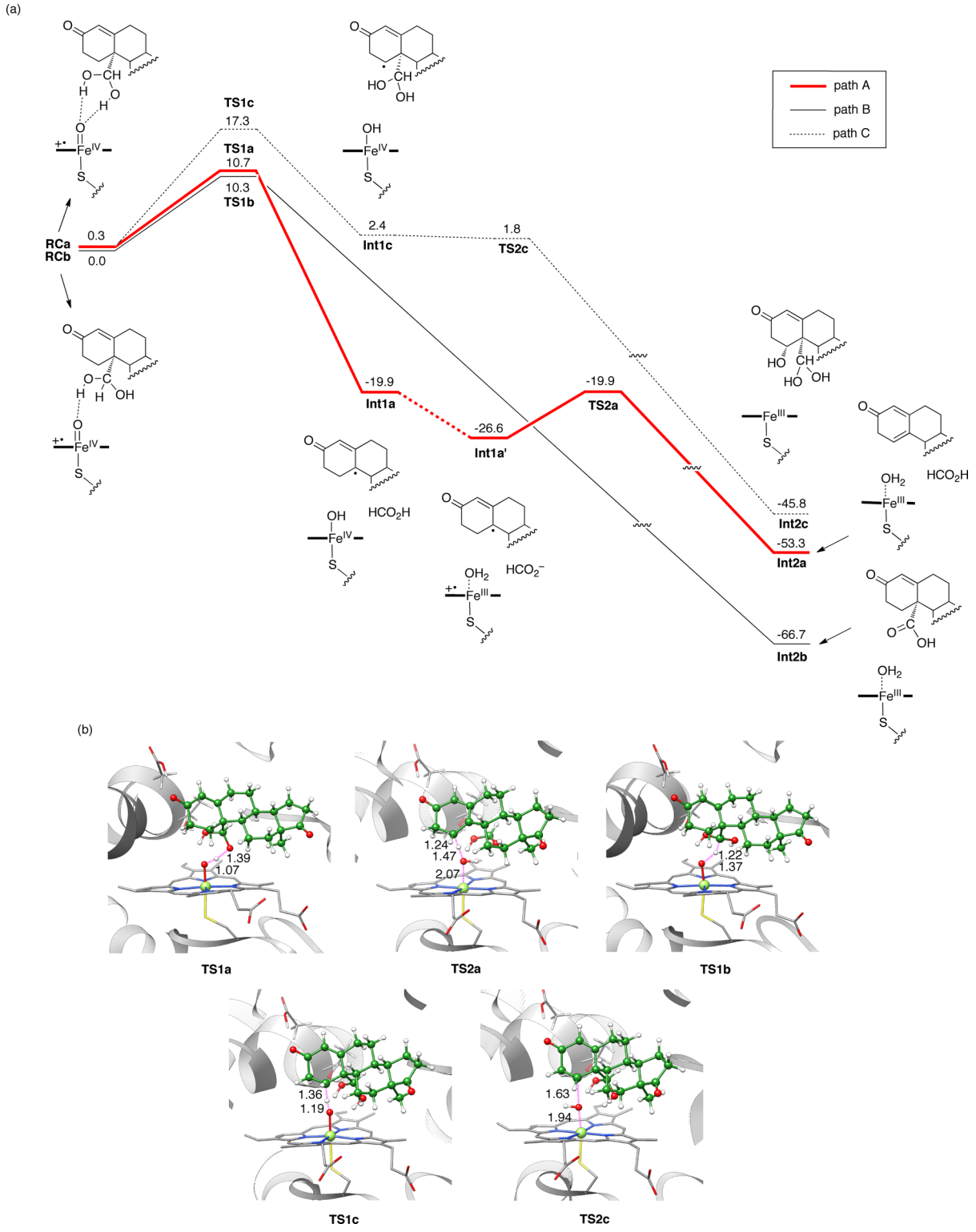


Figure 1. (a) Energy profiles (kcal/mol) for the reactions of *gem*-diol with Cpd I. (b) Transition-state geometries. Key bond distances are shown in Å.

experimental observation. These recent ground-breaking experimental outcomes and arguments raised the need for a reinvestigation of the role of Cpd I in the aromatase reaction. In particular, if Cpd I is the reactive species, how does Cpd I effect the enzymatic aromatization? We herein report the results of our hybrid quantum mechanics and molecular mechanics (QM/MM) study, which has been conducted to address this conundrum. To our delight, we could derive an unprecedented mechanism through extensive QM/MM calculations.

An X-ray crystal structure of human aromatase determined by Ghosh et al. at 2.90 Å resolution (PDB code 3EQM) was used for our QM/MM study.[17] The methyl group of androstenedione at the C19 position (see Scheme 1) in the crystal structure was slightly modified to *gem*-diol **2** and

DOI: 10.1021/acscatal.5b00510
ACS Catal. 2015, 5, 4175−4179

aldehyde **3**, and the reactions of these substrates were examined. Note that although the aldehyde is often assumed to act as the sole substrate (Scheme 1), Yoshimoto and Guengerich determined the ratio of the *gem*-diol to the aldehyde in D_2O (pH 7.8) to be 1.5:1.0, which may imply that the *gem*-diol is slightly more stable.[16] They also showed that either the *gem*-diol or the aldehyde could be the substrate. The ONIOM(QM/MM) method implemented in Gaussian 09 was used for all QM/MM calculations.[18−20] In the ONIOM calculations, B3LYP/[SDD(Fe),6-31G*(others)] was used for the QM calculation step,[21−24] whereas the AMBER03 force field was used for the MM calculations (GAFF was used for the substrate) through the "amber=softfirst" option of Gaussian.[25,26] The mechanical embedding (ME) scheme of ONIOM was used with these methods for geometry optimization and frequency calculations. Single-point energy calculations were performed by using the 6-311+G(d,p) basis set and the electronic embedding (EE) scheme of ONIOM. The doublet spin state was considered. The reported energy data were obtained from single-point energy calculations and zero-point energy corrections. UCSF Chimera was used for visualization of protein molecules,[27] and Molekel was used to plot orbitals.[28] Full details of the QM/MM calculations are given in the Supporting Information (SI).

We first studied the reaction of the *gem*-diol substrate. Three pathways (A−C) were considered. The initial steps of paths A−C are H-abstraction from the O−H bond on C19 (A), H-abstraction from the C−H bond on C19 (B), and H-abstraction from the 1β-carbon of the cyclohexenone ring (C). Figure 1 shows the reaction energy profiles and the optimized geometries of transition states obtained for these pathways. Interestingly, paths A and B are seen to have relatively low energy barriers (10.7 and 10.3 kcal/mol, respectively) for the first step; thus, both H-abstractions from the O−H bond via **TS1a** and from the C−H bond via **TS1b** should be plausible in the first step of the reaction. In path A, H-abstraction is followed by a barrierless, homolytic C−C bond cleavage to form an intermediate (**Int1a**) that involves HCO_2H, a substrate radical, and a protonated ferryl (Cpd II). As a result of an electron transfer from porphyrin to Fe and a proton transfer from formic acid to Cpd II, a somewhat more stable intermediate (**Int1a′**) containing a porphyrin-π-cation-radical-ligated Fe(III)-H_2O species and HCO_2^- can be formed, and hydrogen is then abstracted from the 1β-carbon via **TS2a** to establish a new C−C double bond within the 6-membered steroid-A ring of the substrate (**Int2a**). Note that the proton, which was transferred temporarily from HCOOH to Fe(III)-OH in **Int1a′**, is donated back to $HCOO^-$ in **Int2a**. Additional density functional theory (DFT) calculations showed that the cyclodienone moiety in **Int2a** can easily undergo keto−enol tautomerization to form phenol, especially if there are two water molecules in the vicinity of the carbonyl group (Figure S1). Therefore, it is probable that the produced cyclodienone species may be released first from the active site, and that a phenol-type product may be formed through keto−enol tautomerization in a water-rich environment outside the enzyme.

Interestingly, the H-abstraction from the C−H bond via **TS1b** in path B is followed by barrierless H-abstraction from one of the two O−H bonds on the same carbon (C19), resulting in the formation of a molecule bearing a carboxylic acid group (**Int2b**). The relatively low energy barrier for path B suggests that this route is one possible channel leading to a carboxylic acid product. Indeed, in their experiments, Yoshimoto and Guengerich observed a carboxylic acid as another product of the aromatase-catalyzed reaction of 19-oxo androstenedione (**3**).[16]

The third pathway, which starts with H-abstraction from the 1β-carbon on the cyclohexenone ring (path C) and involves **TS1c**, has a much higher barrier (17.3 kcal/mol). In this case, the H-abstraction is followed by a facile radical rebound step to form an intermediate containing a hydroxylated cyclohexenone molecule. H-abstraction from the 1β-carbon is often postulated as a viable mechanism for the third step effected by Cpd I.[29−31] However, according to our comparative analyses, path C is apparently a less favorable pathway. The reason for the higher barrier for path C should be the steric restriction imposed on the substrate within the active-site pocket.

In general, O−H bonds of alcohols are stronger than their C−H bonds.[32] Despite this, our calculations predicted that **TS1a** and **TS1b** are almost equally stable. To gain insight into this somewhat counterintuitive outcome, we performed an additional spin natural orbital (SNO) analysis.[33] Figure 2 shows

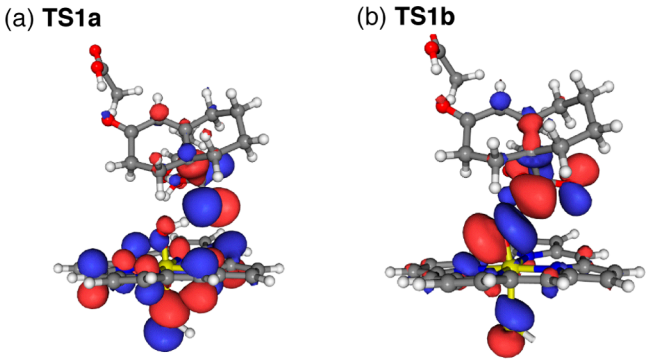


Figure 2. SNOs for (a) **TS1a** and (b) **TS1b**.

the SNOs (with an eigenvalue ∼ −1) for **TS1a** and **TS1b**. Interestingly, the SNO for **TS1a** is not localized on the breaking O−H bond, but it has a lone-pair-like distribution perpendicular to the bond. This is an orbital signature typical of proton-coupled electron transfer (PCET), and the PCET effect contributes to transition-state stabilization in the reactions of oxoiron(IV) species.[33−35] By contrast, the SNO for **TS1b** is firmly localized on the forming and breaking bonds, and in this case, H-abstraction is characterized as H atom transfer. The trend observed here is also reminiscent of the experiment results reported by Wang et al., who showed that a synthetic diiron(IV) complex cleaved the O−H bonds of alcohols, rather than the weaker C−H bonds.[36] In addition, low energy barriers for H-abstraction from an O−H bond were obtained computationally for other reactions of P450s.[37−39]

The reaction of the aldehyde substrate has also been examined (paths D and E, Figure 3a). Interestingly, the barrier for the H-abstraction from the formyl group via **TS1d** (Figure 3b) in path D was low (6.0 kcal/mol). After the H-abstraction, a carboxylic acid product (**Int2d**) was obtained without a rebound barrier. Provided that the aldehyde is slightly less stable than the *gem*-diol, path D should also constitute another channel for the formation of a carboxylic acid product. An aldehyde hydroxylation mechanism was previously discussed by Wang (one of the authors) et al. for the reaction of CYP2E1, on the basis of DFT computational results.[40,41] Yoshimoto and Guengerich suggested that H-abstraction from the aldehyde

DOI: 10.1021/acscatal.5b00510
ACS Catal. 2015, 5, 4175−4179

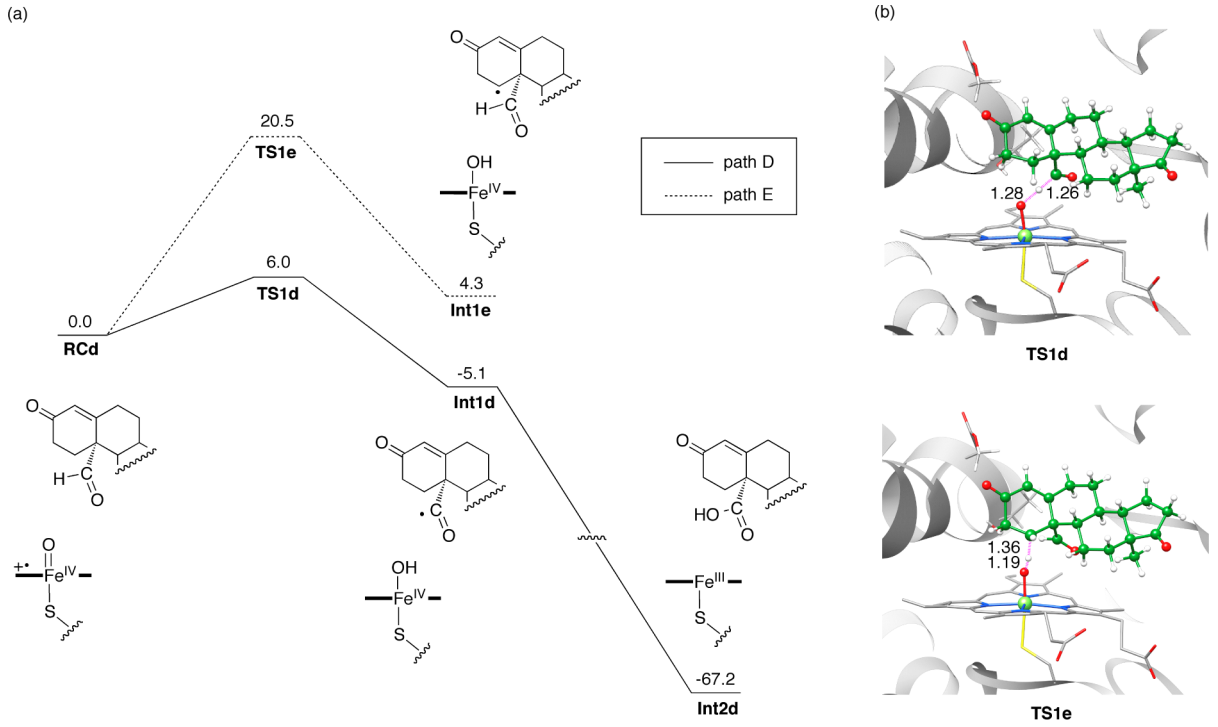


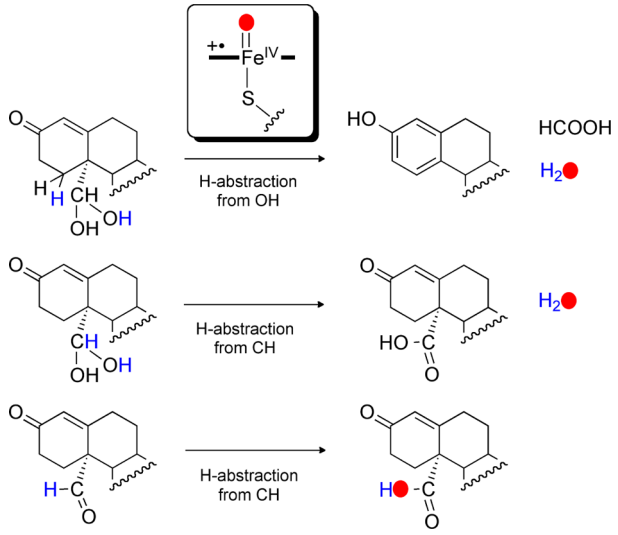
Figure 3. (a) Energy profiles (kcal/mol) for the reactions of aldehyde with Cpd I. (b) Transition-state geometries. Key bond distances are shown in Å.

may also be possible in the aromatase reaction; this mechanism explains why the amount of the carboxylic acid product decreased significantly when the aldehyde hydrogen on C19 was replaced by deuterium.[16] Our result is consistent with their argument, and the deuterium substitution may also affect the barrier height of path B. The possibility of 1β-H-abstraction from the aldehyde substrate (path E) was also examined, but similar to path C, this pathway has a high energy barrier (20.5 kcal/mol, Figure 3a).

We have seen above that the *gem*-diol substrate can be converted into an estrogen or a carboxylic acid (via path A or B, respectively), whereas the aldehyde is converted into a carboxylic acid (path D). Scheme 2 summarizes the computationally derived mechanistic scenario. An O_2-derived oxygen atom will be included in Cpd I as an oxo group (shown in red in Scheme 2) because O_2 undergoes O−O bond cleavage to form Cpd I.[42] If the *gem*-diol undergoes O−H cleavage, an aromatized product is eventually formed, and in this case, an O_2-derived oxygen atom (in Cpd I) will not be incorporated into HCO_2H. Rather, the oxygen will be found in the produced water molecule (see Figure 2). The lack of an O_2-derived oxygen in the produced HCO_2H in path A is consistent with the results of ^{18}O-labeling experiments performed by Yoshimoto and Guengerich,[16] and this agreement reinforces our argument that aromatization should occur via O−H bond cleavage. Importantly, our computational study identified the O−H cleavage pathway (path A) as the sole low-energy-barrier channel to the aromatized product.

Scheme 2 also shows that both of the *gem*-diol and aldehyde substrates can be converted into a carboxylic acid; nevertheless, the fate of the oxo oxygen in Cpd I is different in these two cases. Thus, if a *gem*-diol undergoes C−H cleavage at C19, the oxo oxygen must be incorporated into a water molecule. By contrast, in the aldehyde reaction, the oxo oxygen is incorporated into the carboxyl group. Yoshimoto and

Scheme 2. Summary of the Mechanisms Derived from Our QM/MM Study



Guengerich found that the ratio of the amount of the carboxylic acid containing ^{18}O (coming from $^{18}O_2$) to the amount of the carboxylic acid without ^{18}O was 3:1. A comparison of the experimental and computational data leads us to suggest that the pathway for the formation of the carboxylic acid containing ^{18}O (path D) should be somewhat more favorable than path B.

In conclusion, on the basis of the recent experimental results, which suggested that it should be Cpd I that actually effects step 3 of the aromatase-catalyzed aromatization of androstenedione, we have performed QM/MM calculations on this particular reaction. Our QM/MM calculations have revealed an unprecedented mechanism for the aromatization reaction that features H-abstraction from the O−H bond of the *gem*-diol.

DOI: 10.1021/acscatal.5b00510
ACS Catal. 2015, 5, 4175−4179

The O−H bond is strong, but PCET lowers the energy barrier and thereby enables this process. We also found that a carboxylic acid product can be formed via H-abstraction from the C−H bond of either the *gem*-diol or the aldehyde. Our new mechanistic scenario explains the latest experimental data for step 3 of the aromatase reaction very well.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.5b00510.

QM/MM setup, additional DFT calculations on keto−enol tautomerization, and XYZ coordinates of optimized geometries (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: wangyong@licp.cas.cn.
*E-mail: hirao@ntu.edu.sg.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

H.H. is grateful for financial support in the form of a Nanyang Assistant Professorship and also thanks the High-Performance Computing Centre at Nanyang Technological University for computer resources. W.Y. gratefully acknowledges research support from the National Natural Science Foundation of China (NFSC) (Grants 21003116 and 21173211).

■ REFERENCES

(1) *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Press: New York, 2005.
(2) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841−2887.
(3) Guengerich, F. P. *Chem. Res. Toxicol.* **2001**, *14*, 611−650.
(4) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949−1017.
(5) Poulos, T. L. *Chem. Rev.* **2014**, *114*, 3919−3962.
(6) Thompson, E. A.; Siiteri, P. K. *J. Biol. Chem.* **1974**, *249*, 5364−5372.
(7) Cole, P. A.; Robinson, C. H. *J. Med. Chem.* **1990**, *33*, 2933−2942.
(8) Simpson, E. R.; Mahendroo, M. S.; Means, G. D.; Kilgore, M. W.; Hinshelwood, M. M.; Graham-Lorence, S.; Amarneh, B.; Ito, Y.; Fisher, C. R.; Michael, M. D. *Endocr. Rev.* **1994**, *15*, 342−355.
(9) Groves, J. T. *J. Inorg. Biochem.* **2006**, *100*, 434−447.
(10) Rittle, J.; Green, M. T. *Science* **2010**, *330*, 933−937.
(11) Akhtar, M.; Calder, M. R.; Corina, D. L.; Wright, J. N. *Biochem. J.* **1982**, *201*, 569−580.
(12) Caspi, E.; Wicha, J.; Arunachalam, T.; Nelson, P.; Spiteller, G. *J. Am. Chem. Soc.* **1984**, *106*, 7282−7283.
(13) Sen, K.; Hackett, J. C. *Biochemistry* **2012**, *51*, 3039−3049.
(14) Mak, P. J.; Luthra, A.; Sligar, S. G.; Kincaid, J. R. *J. Am. Chem. Soc.* **2014**, *136*, 4825−4828.
(15) Khatri, Y.; Luthra, A.; Duggal, R.; Sligar, S. G. *FEBS Lett.* **2014**, *588*, 3117−3122.
(16) Yoshimoto, F. K.; Guengerich, F. P. *J. Am. Chem. Soc.* **2014**, *136*, 15016−15025.
(17) Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W. *Nature* **2009**, *457*, 219−223.
(18) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815−826.
(19) Chung, L. W.; Hirao, H.; Li, X.; Morokuma, K. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 327−350.
(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Rev. D.01*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(22) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(23) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(24) Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986; Vol. *67*.
(25) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Kollman, J.; Wang, J. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(26) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(27) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(28) Portmann, S.; Lüthi, H. P. *Chimia* **2000**, *54*, 766−770.
(29) Korzekwa, K. R.; Trager, W. F.; Mancewicz, J.; Osawa, Y. *J. Steroid Biochem. Mol. Biol.* **1993**, *44*, 367−373.
(30) Hackett, J. C.; Brueggemeier, R. W.; Hadad, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 5224−5237.
(31) Krámos, B.; Oláh, J. *Struct. Chem.* **2015**, *26*, 279−300.
(32) Blanksby, S. J.; Ellison, G. B. *Acc. Chem. Res.* **2003**, *36*, 255−263.
(33) Usharani, D.; Lacy, D. C.; Borovik, A. S.; Shaik, S. *J. Am. Chem. Soc.* **2013**, *135*, 17090−17104.
(34) Hirao, H.; Chuanprasit, P. *Chem. Phys. Lett.* **2015**, *621*, 188−192.
(35) Hirao, H.; Ng, W. K. H.; Moeljadi, A. M. P.; Bureekaew, S. *ACS Catal.* **2015**, 3287−3291.
(36) Wang, D.; Farquhar, E. R.; Stubna, A.; Münck, E.; Que, L., Jr. *Nat. Chem.* **2009**, *1*, 145−150.
(37) Wang, Y.; Yang, C.; Wang, H.; Han, K.; Shaik, S. *ChemBioChem* **2007**, *8*, 277−281.
(38) Schyman, P.; Lai, W.; Chen, H.; Wang, Y.; Shaik, S. *J. Am. Chem. Soc.* **2011**, *133*, 7977−7984.
(39) Hirao, H.; Thellamurege, N.; Chuanprasit, P.; Xu, K. *Int. J. Mol. Sci.* **2013**, *14*, 24692−24705.
(40) Wang, Y.; Wang, H.; Wang, Y.; Yang, C.; Yang, L.; Han, K. *J. Phys. Chem. B* **2006**, *110*, 6154−6159.
(41) Liu, X.; Wang, Y.; Han, K. *JBIC, J. Biol. Inorg. Chem.* **2007**, *12*, 1073−1081.
(42) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253−2277.

DOI: 10.1021/acscatal.5b00510
ACS Catal. 2015, 5, 4175−4179