

Article

A New Way to Explore the Mechanism of Catalysis - the Unified Reaction Valley Approach (URVA)

Elfi Kraka ^{1*}, Wenli Zou ², Yunwen Tao ³, and Marek Freindorf ⁴

¹ Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275-0314, United States; ekraka@gmail.com

² Institute of Modern Physics, Northwest University, and Shaanxi Key Laboratory for Theoretical Physics Frontiers, Xi'an, Shaanxi, 710127, China; qcband@gmail.com

³ Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275-0314, United States; ywtao.smu@gmail.com

⁴ Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275-0314, United States; mfreindorf@gmail.edu

* Correspondence: ekraka@gmail.com

Version May 29, 2020 submitted to Catalysts

Contents

| | | |
|---|------------------------------------|---|
| 1 | QM/MM computational details | 1 |
| 2 | Summary of reaction videos | 2 |
| 3 | Browsing file examples | 3 |
| 4 | References | 5 |

1. QM/MM computational details

The IRC calculations of the Claisen rearrangement of chorismate to prephenate in explicit aqueous solution (**R3**) were performed using molecular dynamics (MD), followed by the QM/MM calculations using the B3LYP/6-31+G(d,p) level of theory for the QM part,[1,2] and AMBER force field for the MM part.[3,4] First, the chorismate molecule was surrounded by atoms of TIP3P water molecules [5] with a sphere of a radius 20 Å, and applying an external harmonic potential of a force constant 10 kcal/mol/Å². MD consisted of 2000 steps of minimisation, 100 ps heating from 0 to 300 K, followed by production dynamics for 2 ns, and followed by annealing dynamics, performed for 200 ps by heating the system to 500 K and cooling to 0 K. The simulations were performed with a constrained TS geometry of the solute with a force constant 1500 kcal/mol/Å². Based on the final geometry of the molecular system from MD, the IRC calculations of the reaction **R3** were performed without constraints using the QM/MM method with ONIOM.[6]. The QM part included the chorismate molecule and the MM part included all water molecules. The calculations of the reaction in the protein (**R4**) were performed using the QM/MM method and applying the same level of theory as the reaction in the explicit water solution (**R3**). The initial protein coordinates were taken from the crystal structure of the Arg90Cit (citrulline) mutation of the protein in complexation with a TS analogue.[7] The Arg90Cit mutation was manually replaced back to the original Arg90 structure. The molecular structure was neutralized by six Na⁺ ions and minimised at the MM level with a fixed TS geometry.

24 The following IRC QM/MM calculations were performed without geometry constraints with ONIOM.
 25 The calculations of the chemical reaction of chorismate in the protein (**R5**) were performed using the
 26 QM/MM method similar as in the reaction **R4**, but with the increased size of the QM part. The QM
 27 part in the reaction **R5** included the chorismate molecule and side chains of Arg90, Arg7 and Glu78,
 28 which make hydrogen bonds with chorismate. The reaction **R6** was composed of chorsiamte in the
 29 original Arg90Cit protein mutation, and the IRC calculations of the reaction **R6** were performed using
 30 the same computational protocol and the same level of theory as the reaction **R5**.

31 2. Summary of reaction videos

Table S1. Description of reaction videos and IRC coordinates files of the chemical reactions discussed in this study

| Reaction | File Name | | Description |
|--|--------------------|--------------------|--|
| | Movie | Coordinates | |
| Rh-catalyzed carbonylation | RhCarbonyl-st1.mpg | RhCarbonyl-st1.xyz | S1: First reaction step |
| | RhCarbonyl-st2.mpg | RhCarbonyl-st2.xyz | S2: Second reaction step |
| | RhCarbonyl-st3.mpg | RhCarbonyl-st3.xyz | S3: Third reaction step |
| | RhCarbonyl-st4.mpg | RhCarbonyl-st4.xyz | S4: Fourth reaction step |
| Ti-catalyzed Sharpless epoxidation | Sharpless.mpg | Sharpless.xyz | S5: Catalyzed reaction |
| Au(I)-catalyzed[3,3]-sigmatropic rearrangement of allyl acetate | Claisen-noncat.mpg | Claisen-noncat.xyz | S6: Non-catalyzed |
| | AuClaisen-st1.mpg | AuClaisen-st1.xyz | S7: Catalyzed, first reaction step |
| | AuClaisen-st2.mpg | AuClaisen-st2.xyz | S8: Catalyzed, second reaction step |
| <i>Bacillus subtilis</i> chorismate mutase catalyzed Claisen rearrangement | Chorism-gas.mpg | Chorism-gas.xyz | S9: Reaction R1: Gas phase |
| | Chorism-pcm.mpg | Chorism-pcm.xyz | S10: Reaction R2: Water (PCM) |
| | Chorism-tip3p .mpg | Chorism-tip3p .xyz | S11: Reaction R3: Water (TIP3P) |
| | Chorism-prot1.mpg | Chorism-prot1.xyz | S12: Reaction R4: BsCM (QM/MM) |
| | Chorism-prot2.mpg | Chorism-prot2.xyz | S13: Reaction R5: BsCM (QM _{ext} /MM), part 1 |
| | Chorism-prot3.mpg | Chorism-prot3.xyz | S14: Reaction R5: BsCM (QM _{ext} /MM), part 2 |

32 **3. Browsing file examples**

The diagram illustrates the structure of a browsing file, with annotations pointing to specific sections:

- Begin of each IRC point:** Points to the `BEGIN` keyword.
- Number of atoms:** Points to the `NZ, NSubs` line, which contains the value `51 153`.
- Atomic numbers:** Points to the `IAnZ, IZ1, IZ2, IZ3, IZ4, LBl, LAlpha, LBeta` line, which contains a 4x4 grid of values:

| | | | | | | | |
|---|---|---|---|---|----|----|----|
| 6 | 0 | 0 | 0 | 0 | 1 | 2 | 3 |
| 1 | 0 | 0 | 0 | 0 | 4 | 5 | 6 |
| 1 | 0 | 0 | 0 | 0 | 7 | 8 | 9 |
| 1 | 0 | 0 | 0 | 0 | 10 | 11 | 12 |
- Coordinates:** Points to the `CC` section, which contains a 4x4 grid of Cartesian coordinates.
- Gradient:** Points to the `FX_ZMat_Orientation` section, which contains a 4x4 grid of gradient values.
- Hessian:** Points to the `FFX_ZMat_Orientation` section, which contains a 4x4 grid of Hessian values.
- IRC point number:** Points to the `IPOCou, Energy, XXIRC` line, which contains the values `375 -0.107937661443453E+04 -11.2471818062`.
- End of each IRC point:** Points to the `END` keyword.
- Energy:** Points to the `Energy` value in the `IPOCou, Energy, XXIRC` line.
- Reaction coordinate:** Points to the `XXIRC` value in the `IPOCou, Energy, XXIRC` line.

Figure S1. Browsing file *version 1* with storage of the Hessian at one representative IRC point s.

```

Begin of each IRC point → BEGIN (no Hessian)
                          Natoms,NatomQ
                          2324  0
Number of atoms → Atomic masses needed for the decomposition
                          0.120000000000000E+02 0.100782503700000E+01 0.100782503700000E+01
Atomic masses →          0.100782503700000E+01 0.120000000000000E+02 0.120000000000000E+02
                          0.100782503700000E+01 0.100782503700000E+01 0.120000000000000E+02
                          0.100782503700000E+01 0.100782503700000E+01 0.120000000000000E+02
                          ⋮
                          CC
Coordinates →          -0.567613048769446E+01-0.100731020628619E+02 0.192929422171218E+01
                          -0.439466996809797E+01-0.111441013936450E+02 0.712776632567566E+00
                          -0.645581580739358E+01-0.113958070676757E+02 0.330949727996440E+01
                          -0.454709407929562E+01-0.871136419774327E+01 0.298898982471451E+01
                          ⋮
                          Tangent vector eta (mass-weighted)
Vector  $\eta$  →          -0.468960230865146E-01 0.453080403733721E-01-0.201959583101049E-01
                          -0.922705107789463E-02 0.151896981707737E-01-0.313019479968394E-02
                          -0.173400678506947E-01 0.102161675835089E-01-0.107718406844267E-01
                          -0.181181902629704E-01 0.116313089012388E-01 0.550779658411630E-03
                          ⋮
                          Curvature vector kappa (mass-weighted)
Vector  $\kappa$  →          0.339021615568806E-02-0.185307589203597E-02 0.263150224910690E-03
                          0.482065100420639E-03-0.689402291898235E-03-0.282019377363770E-03
                          0.118885566637580E-02-0.407529474320692E-03 0.270770350097487E-03
                          0.147283299835390E-02-0.701167238617733E-03-0.296489517426708E-03
                          ⋮
IRC point number →      IPoCou,Energy,XXIRC
                          396 -0.106870873006623E+04 -19.7987868899
                          ⋮
Normal mode frequencies → 0.375636000000000E+04 0.375819000000000E+04 0.376007000000000E+04
                          0.376049000000000E+04 0.376097000000000E+04 0.376126000000000E+04
                          0.376252000000000E+04 0.376349000000000E+04 0.376753000000000E+04
                          QM Imaginary Freq No.
                          0
End of each IRC point →  QM Imaginary Normal Modes
                          END

```

Figure S2. Browsing file *version 2* with storage of reaction path direction $\eta(s)$ and reaction path curvature $\kappa(s)$ instead of gradient and Hessian at one representative IRC point s .

33

- 34 1. Becke, A.D. Density–functional Thermochemistry. III. The role of Exact Exchange. *J. Chem. Phys.* **1993**,
35 98, 5648–5652.
- 36 2. Ditchfield, R.; Hehre, W.J.; Pople, J.A. Self–Consistent Molecular–Orbital Methods. IX. An Extended
37 Gaussian–Type Basis for Molecular–Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, 54, 724–728.
- 38 3. Case, D.A.; Ben-Shalom, I.Y.; Brozell, S.R.; Cerutti, D.S.; Cheatham, T.E.; Cruzeiro, V.W.D.; Darden, T.A.;
39 Duke, R.E.; Ghoreishi, D.; Gilson, M.K.; Gohlke, H.; Goetz, A.W.; Greene, D.; Harris, R.; Homeyer, N.;
40 Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T.S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.;
41 Mermelstein, D.J.; Merz, K.M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.;
42 Pan, F.; Qi, R.; Roe, D.R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C.L.; Smith, J.;
43 Salomon-Ferrer, R.; Swails, J.; Walker, R.C.; Wang, J.; Wei, H.; Wolf, R.M.; Wu, X.; Xiao, L.; York, D.M.;
44 Kollman, P.A. AMBER, 2018. University of California, San Francisco.
- 45 4. Nikitin, A.M.; Milchevskiy, Y.V.; Lyubartsev, A.P. A new AMBER-compatible force field parameter set for
46 alkanes. *J. Mol. Mod.* **2014**, 20, 2143.
- 47 5. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential
48 functions for simulating liquid water. *J. Chem. Phys.* **1983**, 79, 926–935.
- 49 6. Chung, L.W.; Sameera, W.M.C.; Ramozzi, R.; Page, A.J.; Hatanaka, M.; Petrova, G.P.; Harris, T.V.; Li, X.; Ke,
50 Z.; Liu, F.; Li, H.B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**,
51 115, 5678–5796.
- 52 7. Burschowsky, D.; van Eerde, A.; Ökvist, M.; Kienhöfer, A.; Kast, P.; Hilvert, D.; Krengel, U. Electrostatic
53 transition state stabilization rather than reactant destabilization provides the chemical basis for efficient
54 chorismate mutase catalysis. *PNAS* **2014**, 111, 17516.

55 © 2020 by the authors. Submitted to *Catalysts* for possible open access publication under the terms and conditions
56 of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).