SUPPORTING INFORMATION

Connecting the Mechanistic Steps of Cyclic Dipeptide Formation by Proton-Transfer Network: pH, Temperature, Pressure, and Nuclear Quantum Effects

Pimjai Pimbaotham[1], John K. Villanueva[2], Siriporn Jungsuttiwong[1], Masanori Tachikawa[3], and Robert K. Szilagyi[2*]

[1]*Department of Chemistry, Faculty of Science, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand*

[2]*Department of Chemistry, University of British Columbia Okanagan, 1177 Research Rd., Kelowna, BC V1V 1V7, Canada.*

[3]*Quantum Chemistry Division, Yokohama City University, Yokohama 236-0027, Japan;*

** Correspondence should be addressed: Robert.Szilagyi@UBC.CA*

Table of Contents:

Computational log files, XYZ coordinates, examples for MC-DFT calculations, IRC scan results can be accessed free of charge at DOI: 10.5281/zenodo.18264551.

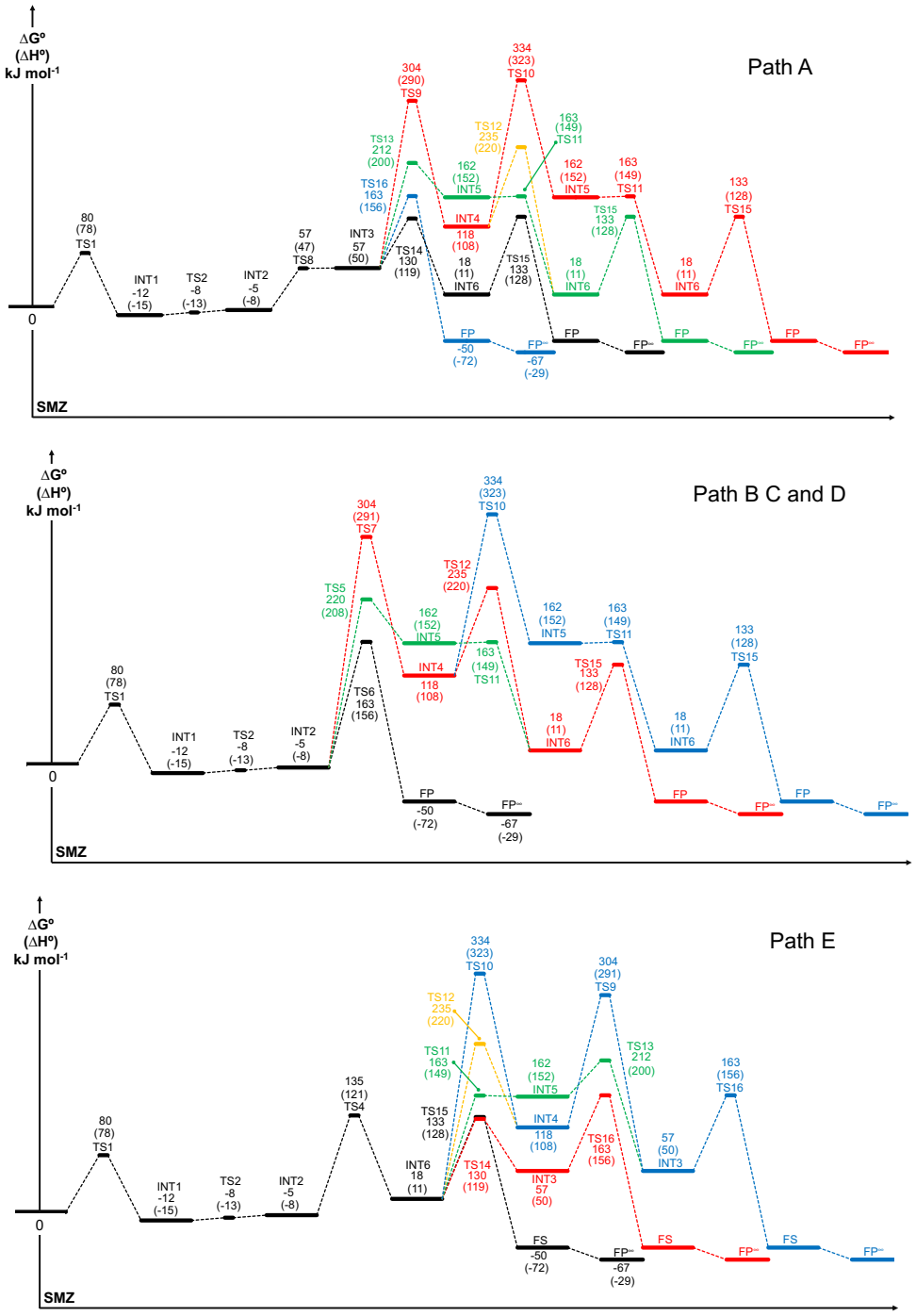

Figure S1: Free energy profile of cyclic peptide formation pathways A-E (enthalpies in parentheses, in kJ mol^{-1}) at the MN15/def2TZVPP|SMD level of theory calculated for *Scenario III* at standard state. The reference state (zwitterionic starting material, SMZ) is defined by implicitly solvated zwitterionic trans isomer of linear diglycine peptide. The intermediates and transition states are defined in **Scheme 1** and the proton-transfer network is shown in **Scheme 2**.

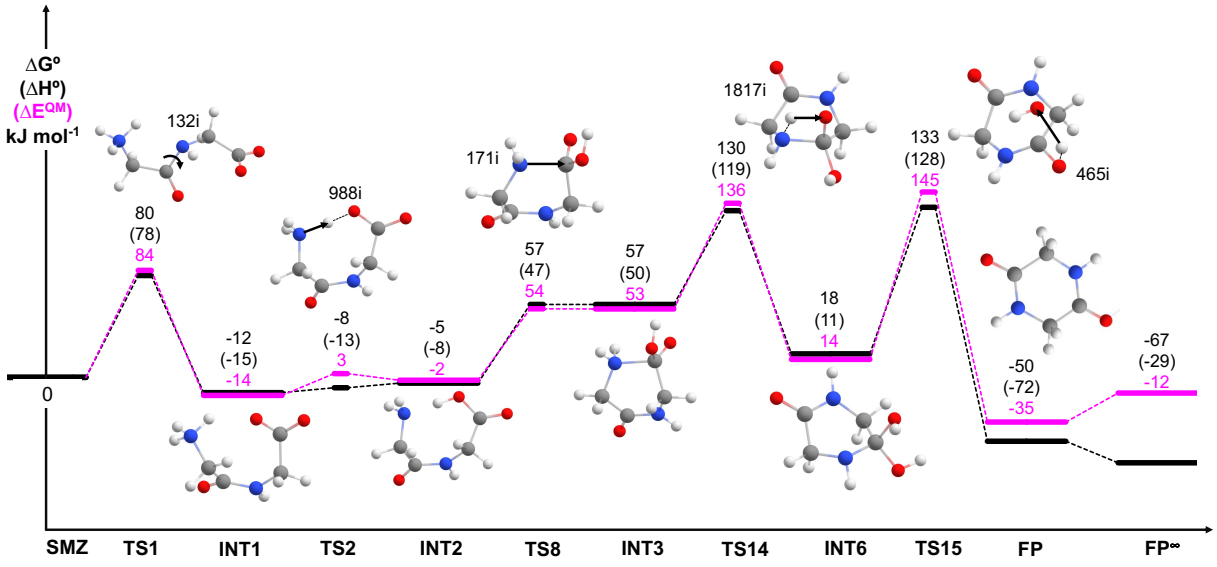

Figure S2: Comparison of the free-energy profile (ΔG^{o}, black trace) and the potential energy plot (ΔE^{QM}, pink trace) calculated at MN15/def2TZVPP|SMD level of theory in order to support the existence of TS2 proton-transfer and TS8 C–N bond formation transition states that are distorted by excessive entropic terms in thermochemistry calculations that are based on the ideal gas approximation at standard state.

Extended discussion of the nuclear quantum effects (NQE):

Table 1 summarized the changes in the activation energies relative to the respective pre-TS intermediates in the reaction direction of amide bond formation. In addition to the reference level of theory utilized for all four scenarios (MN15/def2TZVPP), we varied the basis set (def2TZVPP Ahlrichs[9] → cc-pVTZ Dunning[10] → 6-31G(d) Pople[2, 3]) and the density functional (MN15[12] → B3LYP[1, 5]). The highlighted levels of theory are the commonly employed methods in MC-DFT calculations.[8, 11] The introduction of the Dunning or Pople basis sets is due to the technical limitations of transition state search in MC-DFT calculations when using the Ahlrichs basis set. The first two data rows in **Table 1** are from classical DFT. They show that the thermochemistry between the two triple-ζ quality basis sets is practically identical within a few kJ mol^{-1} for the initial proton-transfer (TS2) and the elimination step (TS15). However, there is a significant difference for the *gem*-diol formation due to the large change in positions of the explicit solvent molecules. Two nuclear quantum protons impact the transition states by up to 20 kJ mol^{-1} through lowering of the activation potential energies of each species. Switching to a commonly used hybrid functional, B3LYP maintains the difference for TS2 neutralization step, but significantly varies in opposite directions for TS14 and TS15. The difference in tunneling can be related the electronic structure changes between INT3 and INT6 being more N-centered as the amide bond develops in TS14, while TS15 connects INT6 and FP with O-centered proton transfer for the final elimination step. The overestimation of the reaction thermochemistry by the B3LYP functional has also been seen for the linear peptide mechanism.[4] This also impacted the integrated complete-basis set extrapolation (CBS-QB3[6, 7]) calculations, where B3LYP geometries are used for describing the equilibrium structures. Reduction of the basis set from triple- to double-ζ quality reduces the barrier height deviations from the reference level of theory, but also further diminishes the impact of NQE. The geometry changes when employing a quantum protonic wave function for the transition states and the two intermediates they connect to are summarized on the inset structures of **Figure S3**, which can be up to a modest 0.1 Å elongation of the X–H bond. In general, we observe that the MC-DFT calculations can change the energetics barrier; however, when the protonic wavefunctions are adjacent to each other (< 1.6 Å) a significant stabilization can take place of -7 kJ mol^{-1}.

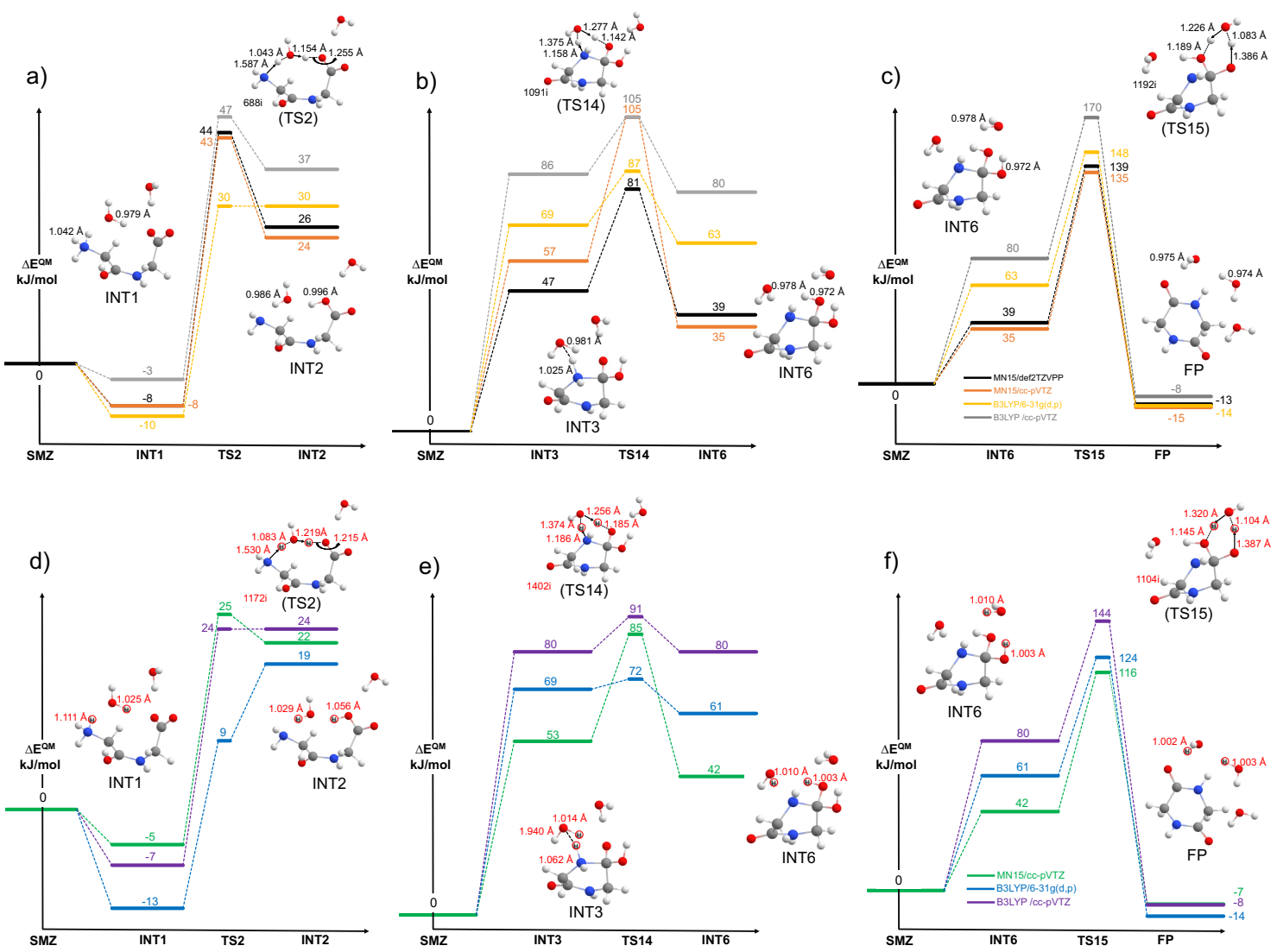

Figure S3: Comparison of potential energy profiles (ΔE^{QM}) at various levels of theory from classical DFT (Panels a-c) to multi-component DFT (Panels d-f). The proton-transfer transition state TS2 in the neutralization step is shown in Panels a and d, TS14 to form the gem-diol intermediate INT6 are in Panels b and e, followed by the water elimination step TS15 to FP in Panels c and f. All potential energy values are given relative to the zwitterionic trans-linear peptide structure (SMZ) from *Scenario IV* at standard state. Black, brown, orange, and gray traces are MN15/def2TZVPP, MN15/cc-pVTZ, B3LYP/6-31G(d,p), and B3LYP/cc-pVTZ levels, respectively. The green, blue, and purple traces are MC-MN15/ccpVTZ, MC-B3LYP/6-31G(d), and MC-B3LYP/cc-pVTZ levels, respectively with protonic wave functions on centers involved in the TS displacements. All models were embedded into SMD polarizable continuum model.

A deeper dive into the impact of NQE on the transition state from Intrinsic Reaction Coordinate calculations reveals significant changes in the potential energy surface curvature when the transferring proton is described as a wave function. **Figure S4** summarizes the changes in the IRC path in both forward and reverse directions. In Panel a), both pre- and post-TS structure intermediates get destabilized by the protonic wave function treatment, while it is mixed in TS14, where right hand side (INT6) becomes less stable compared to the left-hand side (INT3). Lastly, there is no major differences among INT6 and FP connected by TS15. As can be seen from the inset structures in each panel, the nuclear wave function modestly (<0.1 Å) increases the bond lengths due to the delocalized, cloud-like quantum treatment of the proton. There is one significant change which is the position of the explicit water solvent relative to the secondary ammonium cation and the negatively charged, deprotonated group of the gem-diol. The distance shortens from 2.65 Å to 1.94 Å as the H becomes a shared proton between N and O of the newly formed C–N bond. This shifts the potential energy surface to a large degree as it contributes the increase in the barrier in comparison to TS2 and TS15. The H-bond network differs between the classical DFT and MC-DFT optimized TS structure. Therefore, we assessed the impact of switching in between the optimized structures and levels of theory without observing more than a kJ mol^{-1} effect on TS energies. The overall TS geometry remains the same; however, the NQE imposes an elongation on the intermolecular H-bonding distances and even more the intramolecular H–O/N distances. A small angle adjustment is to follow; however, the energetics differences will be in the order of 7 kJ mol^{-1}, at most. The potential energy surface profile is the most delicately impacted that can result in challenges of localizing transition states when using MC-DFT, since for some of the examples highlighted, the flatness of the PES is rather visible in **Figure S4.**

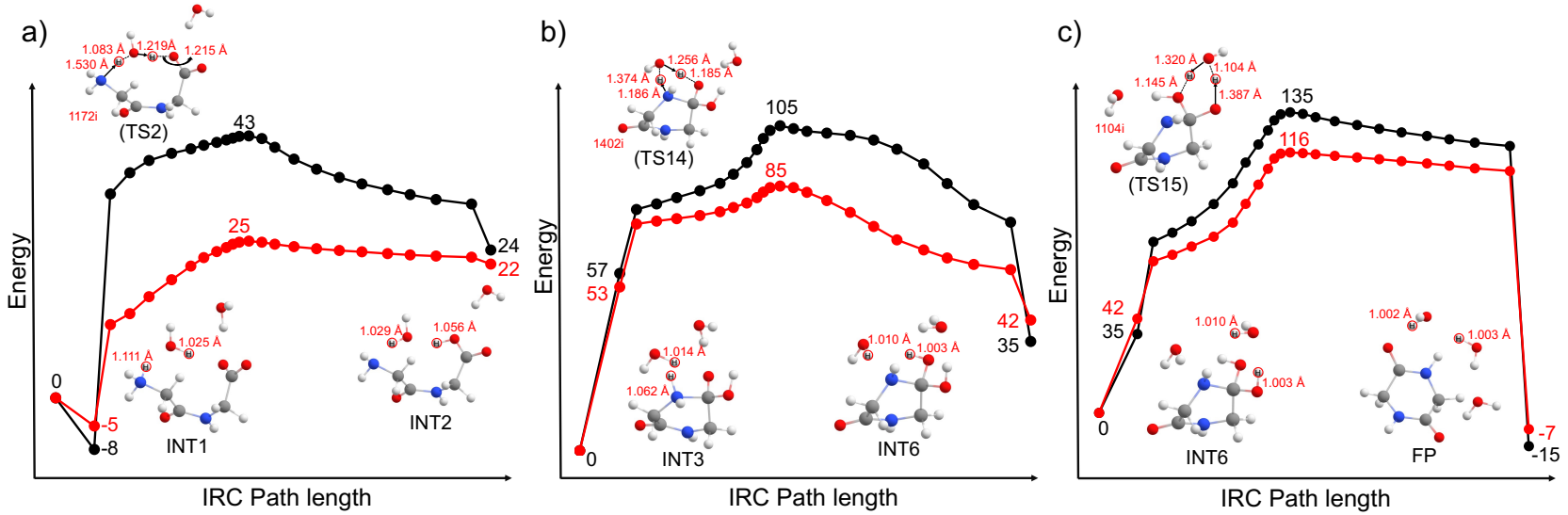

Figure S4: IRC forward and backward pathways represented by potential energy values, (ΔE^{QM}) from proton-transfer transition state TS2 in the neutralization step (Panel a), TS14 to form the gem-diol intermediate INT6 (Panel b) followed by the water elimination step TS15 to FP (Panel c). Black and red traces are MN15/cc-pVTZ|SMD and MC-MN15/cc-pVTZ|SMD levels of theory, respectively.

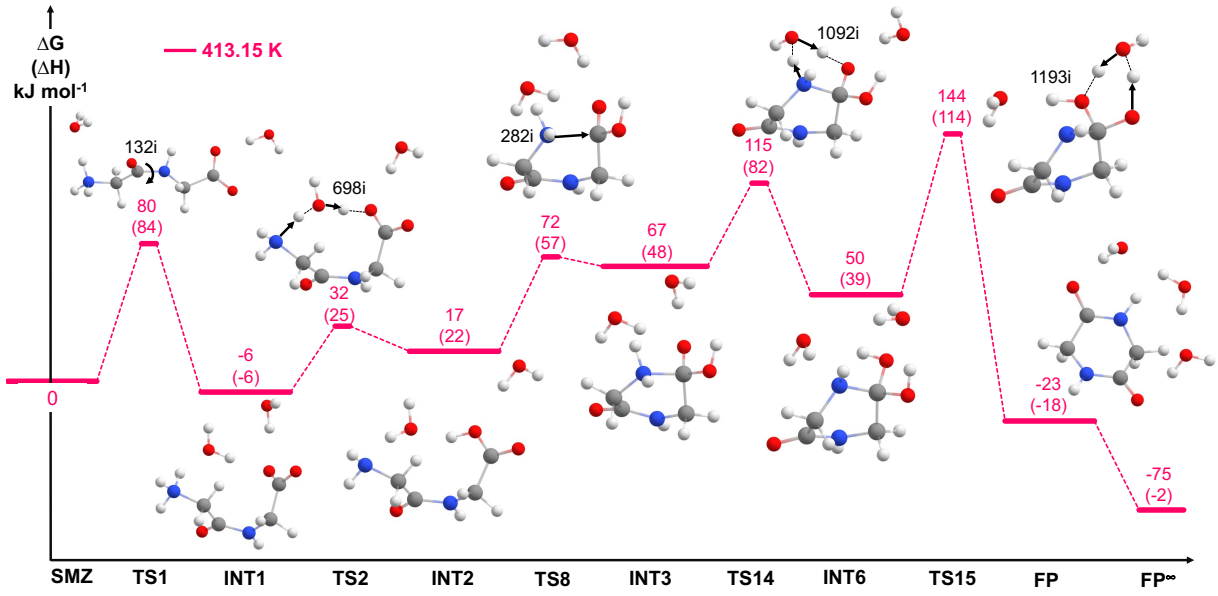

Figure S5: Free energy profile for the cyclic peptide condensation reaction at 140 ºC and 1 atm mimicking experimental hydrothermal conditions as described for Scenario IV. Equilibrium structures and transition states were obtained from MN15/def2TZVPP|SMD calculations.

Table S1: Summary of energy components, changes in enthalpy and free-energy corrections to the SCF electronic energy (E^{QM}, au.) as a function of hydrothermal non-standard conditions, T = 140 ºC and P = 1,748 atm)

Intermediate/TS (description)	T, K	P, atm	E^{QM}, a.u	Thermal Correction to Enthalpy, a.u.	Thermal Correction to Free Energy, a.u.
SMZ (trans dipeptide)	298	1	-644.958573	0.202567	0.142461
	413	1		0.213566	0.117454
	413	1748		0.213566	0.127223
TS1 (amide rotation)	298	1	-644.922905	0.199161	0.137716
	413	1		0.209758	0.112259
	413	1748		0.209758	0.122027
INT1 (cis dipeptide)	298	1	-644.961446	0.202915	0.142908
	413	1		0.213988	0.117926
	413	1748		0.213988	0.127695
TS2 (N^{+}-H $\rightarrow$ O^{-}–C=O)	298	1	-644.941689	0.195461	0.137071
	413	1		0.206137	0.112784
	413	1748		0.206137	0.122553
INT2 (neutralized ends)	298	1	-644.948792	0.200897	0.139776
	413	1		0.212226	0.114319
	413	1748		0.212226	0.124088
TS8 (C–N formation)	298	1	-644.935003	0.201026	0.144884
	413	1		0.211859	0.121438
	413	1748		0.211859	0.131206
INT3 (zwitterionic C–N)	298	1	-644.940513	0.203037	0.147889
	413	1		0.213889	0.124825
	413	1748		0.213889	0.134594
TS14 (N^{+}-H $\rightarrow$ O^{-}–C–OH)	298	1	-644.917337	0.193308	0.141662
	413	1		0.203374	0.120089
	413	1748		0.203374	0.129857
INT6 (gem diol)	298	1	-644.943824	0.202534	0.145639
	413	1		0.213619	0.121860
	413	1748		0.213619	0.131629
TS15 (elimination step)	298	1	-644.905817	0.193866	0.141412
	413	1		0.204330	0.119459
	413	1748		0.204330	0.129228
FP (DKP product)	298	1	-644.963664	0.200790	0.139372
	413	1		0.212367	0.113758
	413	1748		0.212367	0.123527

References used in the supporting information

1. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652. 10.1063/1.464913
2. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654-3665. 10.1063/1.444267
3. W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257-+. 10.1063/1.1677527
4. T. Herriman and R. K. Szilagyi, *Org. Biomol. Chem.*, 2023, **21**, 5953-5963. 10.1039/d3ob00410d
5. C. T. Lee, W. T. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789. 10.1103/PhysRevB.37.785
6. J. A. Montgomery, M. J. Frisch, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 1999, **110**, 2822-2827. 10.1063/1.477924
7. J. A. Montgomery, M. J. Frisch, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 2000, **112**, 6532-6542. 10.1063/1.481224
8. Y. Shimohata, H. Sakagami, Y. Kanematsu, D. S. R. Rocabado, M. Tachikawa and T. Ishimoto, *J. Phys. Chem. C*, 2023, **127**, 24316-24323. 10.1021/acs.jpcc.3c04610
9. F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297-3305. 10.1039/b508541a
10. D. Woon and T. Dunning, *J. Chem. Phys.*, 1993, **98**, 1358-1371. 10.1063/1.464303
11. N. Yodsin, H. Sakagami, T. Udagawa, T. Ishimoto, S. Jungsuttiwong and M. Tachikawa, *Mol. Catal.*, 2021, **504**, 10. 10.1016/j.mcat.2021.111486
12. H. S. Yu, X. He, S. L. Li and D. G. Truhlar, *Chemical Science*, 2016, **7**, 5032-5051. 10.1039/C6SC00705H