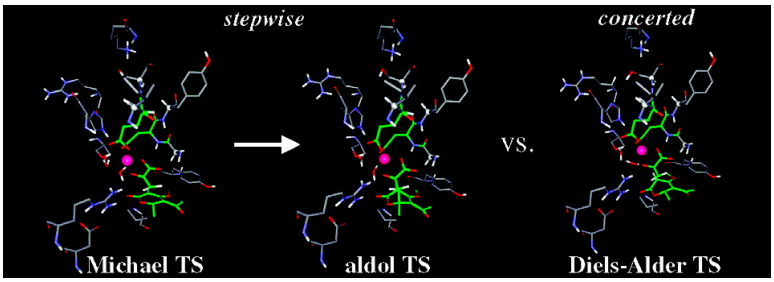
Article

Macrophomate Synthase: QM/MM Simulations Address the Diels–Alder versus Michael–Aldol Reaction Mechanism

Cristiano Ruch Werneck Guimares, Marina Udier-Blagovi, and William L. Jorgensen

J. Am. Chem. Soc., **2005**, 127 (10), 3577-3588• DOI: 10.1021/ja043905b • Publication Date (Web): 16 February 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 14 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 02/16/2005

Macrophomate Synthase: QM/MM Simulations Address the Diels−Alder versus Michael−Aldol Reaction Mechanism

Cristiano Ruch Werneck Guimarães, Marina Udier-Blagović, and William L. Jorgensen*

Contribution from the Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06520-8107

Received October 6, 2004; E-mail: william.jorgensen@yale.edu

Abstract: Macrophomate synthase (MPS) of the phytopathogenic fungus *Macrophoma commelinae* catalyzes the transformation of 2-pyrone derivatives to the corresponding benzoate analogues. The transformation proceeds through three separate chemical reactions, including decarboxylation of oxalacetate to produce pyruvate enolate, two C−C bond formations between 2-pyrone and pyruvate enolate that form a bicyclic intermediate, and final decarboxylation with concomitant dehydration. Although some evidence suggests that the second step of the reaction catalyzed by MPS is a Diels−Alder reaction, definite proof that the C−C bond formations occur via a concerted mechanism is still required. An alternative route for formation of the C−C bonds is a stepwise Michael−aldol reaction. In this work, mixed quantum and molecular mechanics (QM/MM) combined with Monte Carlo simulations and free-energy perturbation (FEP) calculations were performed to investigate the relative stabilities of the transition states (TS) for both reaction mechanisms. The key results are that the Diels−Alder TS model is 17.7 and 12.1 kcal/mol less stable than the Michael and aldol TSs in the stepwise route, respectively. Therefore, this work indicates that the Michael−aldol mechanism is the route used by MPS to catalyze the second step of the overall transformation, and that the enzyme is not a natural Diels−Alderase, as claimed by Ose and co-workers (*Nature* **2003**, *422*, 185−189; *Acta Crystallogr.* **2004**, *D60*, 1187−1197). A modified link-atom treatment for the bonds at the QM/MM interface is also presented.

Introduction

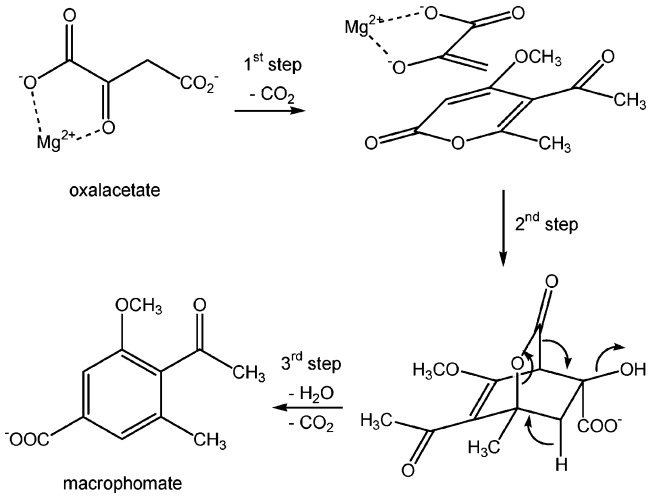
The Diels−Alder reaction between 1,3-dienes and alkenes is a cycloaddition that forms cyclohexenes in a concerted manner. Due to its versatility and extraordinary stereoselectivity (two new C−C bonds and up to four new stereocenters are created), the Diels−Alder reaction is considered one of the most important reactions in organic synthesis.[1] This pericyclic reaction has been applied extensively in the synthesis of complex pharmaceutical and biologically active compounds.[2]

The involvement of enzymes in the biosynthesis of secondary metabolites through Diels−Alder reactions has been postulated for over 100 natural products.[3] Nevertheless, to date, only three purified or partially purified enzymes have been reported as naturally occurring Diels−Alderases.[4,5] Solanapyrone synthase of the fungus *Alternaria solani* catalyzes the conversion of prosolanapyrone II to solanapyrones A and D via oxidation and subsequent intramolecular Diels−Alder reaction.[6−8] Fungal lovastatin nonaketide synthase has been shown to catalyze the [4 + 2] cycloaddition of an intermediate hexaketide triene to generate the decalin system of lovastatin.[9] The last example of natural Diels−Alderases is macrophomate synthase (MPS) of the phytopathogenic fungus *Macrophoma commelinae.* The enzyme catalyzes the transformation of 2-pyrone derivatives into the corresponding benzoate analogues.[10−12]

As Diels−Alder cycloadditions are still rare in nature, catalytic antibodies have provided an alternate source of detailed information on protein-catalyzed cycloadditions.[5] The immune system solves the problem of molecular recognition by generating structurally distinct antibodies and amplifying those with the requisite binding affinity and specificity.[13] When the immune system is challenged with transition-state analogues, the resultant antibodies can show catalytic activity for the respective reaction. Several Diels−Alderase catalytic antibodies have been elicited

(1) Houk, K. N.; Gonzalez, J.; Li, Y. *Acc. Chem. Res.* **1995**, *28*, 81−90.
(2) Desimoni, G.; Tacconi, G.; Barco, A.; Pollini, G. P. In *Natural Products Synthesis Through Pericyclic Reactions*; American Chemical Society: Washington, D.C., 1983.
(3) Ichihara, A.; Oikawa, H. *Curr. Org. Chem.* **1998**, *2*, 365−394.
(4) Pohnert, G. *Chem. Biochem.* **2003**, *4*, 713−715.
(5) Pohnert, G. *Chem. Biochem.* **2001**, *2*, 873−875.
(6) Oikawa, H.; Katayama, K.; Suzuki, Y.; Ichihara, A. *J. Chem. Soc., Chem. Commun.* **1995**, 1321−1322.
(7) Katayama, K.; Kobayashi, T.; Oikawa, H.; Honma, M.; Ichihara, A. *Biochim. Biophys. Acta* **1998**, *1384*, 387−395.
(8) Oikawa, H.; Kobayashi, T.; Katayama, K.; Suzuki, Y.; Ichihara, A. *J. Org. Chem.* **1998**, *63*, 8748−8756.
(9) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 11519−11520.
(10) Sakurai, I.; Miyajima, H.; Akiyama, K.; Shimizu, S.; Yamamoto, Y. *Chem. Pharm. Bull.* **1988**, *36*, 2003−2011.
(11) Watanabe, K.; Mie, T.; Ichihara, A.; Oikawa, H.; Honma, M. *J. Biol. Chem.* **2000**, *49*, 38393−38401.
(12) Ose, T.; Watanabe, K.; Mie, T.; Honma, M.; Watanabe, H.; Yao, M.; Oikawa, H.; Tanaka, I. *Nature* **2003**, *422*, 185−189. (b) Ose, T.; Watanabe, K.; Yao, M.; Honma, M.; Oikawa, H.; Tanaka, I. *Acta Crystallogr.* **2004**, *D60*, 1187−1197.
(13) Romesberg, F. E.; Spiller, B.; Schultz, P. G.; Stevens, R. C. *Science* **1998**, *279*, 1929−1933.

Scheme 1. Chemical Reactions Catalyzed by MPS

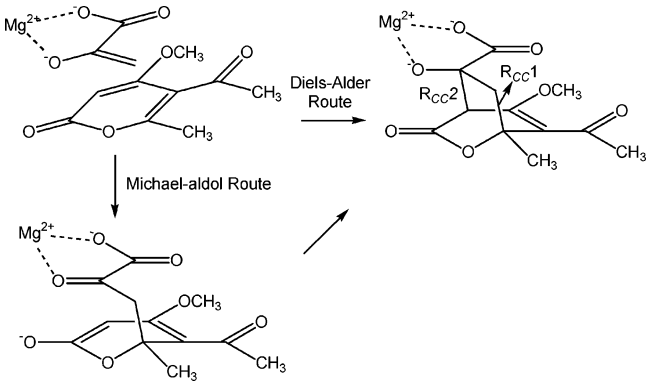


using this strategy,[14−18] and crystal structures for three of them are available.[13,19,20] Modified RNA has also been shown to be an effective Diels−Alderase with its activity entirely dependent on the nature of the base modification and presence of cupric ion.[21]

According to the mechanistic studies of Watanabe and co-workers,[11] MPS catalyzes the transformation of 2-pyrone derivatives into the corresponding benzoate analogues through three separate chemical reactions, including decarboxylation of oxalacetate to produce pyruvate enolate, C−C bond formations between 2-pyrone and pyruvate enolate to form a bicyclic intermediate, and final decarboxylation with concomitant dehydration to produce benzoate analogues, like macrophomate (Scheme 1). Examination of the kinetic parameters indicates that the third step is rate-determining.[11] In addition, Watanabe and co-workers suggest that the second step is an intermolecular Diels−Alder reaction.[11,12] This combination of an electron-rich dienophile, pyruvate enolate, and an electron-deficient diene, the substituted 2-pyrone, is categorized as an inverse-electron demand Diels−Alder reaction.[22]

The crystal structure of MPS complexed to pyruvate enolate and Mg^{2+} was obtained recently, making MPS the first potential Diels−Alderase with its structure determined.[12] MPS is a Mg^{2+}-dependent enzyme consisting of 339 residues. Mg^{2+}, which is essential for the decarboxylation of oxalacetate to produce the reactive enolate (see Scheme 1),[11] is coordinated in an octahedral geometry. Two of the ligands of Mg^{2+} are side-chain carboxyl oxygens from residues Glu185 and Asp211. Two additional coordination sites are filled with water molecules, and the last two coordination sites are occupied by the C2-carbonyl and C1-carboxyl oxygens of pyruvate enolate. The X-ray crystal structure of MPS contains only residues 1−299 for each of the monomers belonging to a hexameric structure. The coordinates of the C-terminal 40 residues from 300 to 339 were not determined due to poor electron density in this region. However, this region is not important for the catalytic activity since the mutant with these 40 residues deleted showed the same MPS-specific activity as that of the native enzyme.[12]

Scheme 2. Diels−Alder versus Michael−Aldol Reaction Mechanism



Although some evidence suggests that the second step of the reaction catalyzed by MPS is a Diels−Alder reaction, proof that the C−C bond formations occur via a concerted mechanism is still required. The other plausible route is a tandem Michael−aldol reaction (Scheme 2).[11] This alternative is initiated with attack of the reactive enolate to form the first C−C bond with the negative charge highly delocalized in a 1,7-dioxoheptatrienyl substructure, followed by formation of the second C−C bond to generate the bicyclic intermediate.

In this work, mixed quantum and molecular mechanics (QM/MM)[23] combined with Monte Carlo (MC)[24] simulations and free-energy perturbation (FEP)[25−28] calculations were performed to determine the preferred route for the second step of the transformation catalyzed by MPS. More specifically, potential of mean force (PMF)[26] calculations were carried out to investigate the relative stabilities of the transition states (TS) identified for both the concerted Diels−Alder pathway and for the stepwise Michael−aldol reaction. On the technical side, a new treatment for the bonds at the QM/MM interface is also presented.

The QM/MM Model. A major complication in hybrid QM/MM methods is the treatment of the interface between the quantum part, describing the reactive region, and the classical part, describing the environment. The most popular approaches are the link atom[29−31] and the local self-consistent field formalisms.[32−34] A modified link-atom method has been utilized

(14) Hilvert, D.; Hill, K. W.; Nared, K. D.; Auditor, M.-T. *J. Am. Chem. Soc.* **1989**, *111*, 9261−9262.
(15) Braisted, A. C.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 7430−7431.
(16) Suckling, C. J.; Tedford, M. C.; Bence, L. M.; Irvine, J. I.; Stimson, W. H. *J. Chem. Soc., Perkin Trans. 1* **1993**, 1925−1929.
(17) Gouverneur, V. E.; Houk, K. N.; Pascual-Teresa, B.; Beno, B.; Janda, K. D.; Lerner, R. A. *Science* **1993**, *262*, 204−208.
(18) Ylikauhaluoma, J. T.; Ashley, J. A.; Lo, C. H.; Tucker, L.; Wolfe, M. M.; Janda, K. D. *J. Am. Chem. Soc.* **1995**, *117*, 7041−7047.
(19) Heine, A.; Stura, E. A.; Ylikauhaluoma, J. T.; Gao, C.; Deng, Q.; Beno, B. R.; Houk, K. N.; Janda, K. D.; Wilson, I. A. *Science* **1998**, *279*, 1934−1940.
(20) Xu, J.; Deng, Q.; Chen, J.; Houk, K. N.; Bartek, J.; Hilvert, D.; Wilson, I. A. *Science* **1999**, *286*, 2345−2348.
(21) Tarasow, T. M.; Eaton, B. E. *Cell. Mol. Life Sci.* **1999**, *55*, 1463−1472.
(22) Afarinkia, K.; Vinader, V.; Nelson, T. D.; Posner, G. H. *Tetrahedron* **1992**, *48*, 9111−9171.
(23) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(24) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon Press: Oxford, U.K., 1987.
(25) Zwanzig, R. *J. Chem. Phys.* **1954**, *22* (8), 1420−1426.
(26) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Biochem.* **1989**, *18*, 431−492.
(27) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184−189.
(28) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417.
(29) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.
(30) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(31) Eurenius, K. P.; Chatfield, D. C.; Brooks, B. R. *Int. J. Quantum Chem.* **1996**, *60*, 1189−1200.
(32) Théry, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. *J. Comput. Chem.* **1994**, *15*, 269−282.

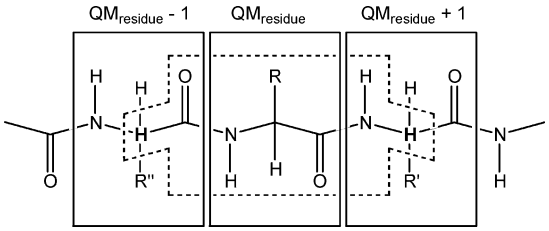


Figure 1. QM/MM partition scheme.

here. In our QM/MM implementation, nonbonded interactions between the QM part and the environment are treated classically, as described previously for uncatalyzed reactions in solution[35−37] and for an enzymatic reaction where no substrate−protein covalent linkages are formed or broken.[38] Specifically, the nonbonded interactions are given by eq 1 as the sum of all interactions between the QM and MM atoms, where q_i and q_j are the partial atomic charges, and σ_{ij} and ϵ_{ij} are Lennard-Jones (LJ) parameters.

$$E_{\mathrm{QM/MM}} = \sum_{i}^{\text{on QM}} \sum_{j}^{\text{on MM}} [q_i q_j e^2/r_{ij} + 4\epsilon_{ij}(\sigma_{ij}^{12}/r_{ij}^{12} - \sigma_{ij}^{6}/r_{ij}^{6})] \quad (1)$$

The interfacial scheme for the current QM/MM implementation involves the backbone of the protein. Important amino acid residues ($QM_{residue}$) that are involved in the enzymatic reaction are always fully included in the QM region. The QM/MM partition is then performed at two regions in the peptide chain using three bonds at each QM/MM interface: N−C_α, H−C_α, and R″−C_α in $QM_{residue} - 1$, and C_α−C, C_α−H, and C_α−R′ in $QM_{residue} + 1$ (Figure 1). The QM atoms are represented by the region within the dashed lines in Figure 1. The link atoms are formally hydrogens, but in contrast to the standard link-atom formalism, their bonds to the QM region are elongated by ca. 0.5 Å to allow spatial coincidence between the hydrogens and the C_α atoms. In this manner, the link atom and C_α form a single atom and collectively interact with the system.

The interactions of the overlapped atoms with the QM region are described through the link atoms as hydrogens, included in the SCF calculation. The connection of the QM and MM regions requires the inclusion of the classical bond stretching, angle bending, and torsion terms if any MM atom is involved in the interaction. These bonded interactions are described through the C_α carbon atoms. Nonbonded interactions (eq 1) are also included between the QM and MM atoms; the link atoms in this case have the charge determined from the QM calculation and normal Lennard-Jones parameters for a C_α atom. However, to avoid redundancies with the MM bond stretching, angle bending, and torsion terms, any nonbonded terms between QM and MM atoms that have 1,2 or 1,3 or 1,4 spatial relationship are excluded.

The elongation of bonds between the link atoms and the QM region is obviously artificial. To minimize potential errors, the elongated bonds are kept as far from the reaction center as possible. Also, these bond lengths are kept fixed during the MC simulations. As fixed bonds, their energy contributions to the entire system are roughly constant and should not affect the Metropolis test.[24]

In view of the MC sampling, the need for very rapid QM energy evaluation is clear. Therefore, in the present QM/MM implementation,[35] the QM choice is a semiempirical method, such as AM1,[39] PM3,[40,41] or PDDG/PM3.[42,43] The choice of overlapping the C_α atom with the hydrogen link atom is supported by both Mulliken and CM[44,45] charges from the semiempirical methods. Specifically, the charges for a hydrogen atom connected to either a carbonyl or a nitrogen of a peptide bond are similar to the charge for a C_α atom in the OPLS-AA force field.[46] In other words, the semiempirical QM charges obtained for the hydrogen link atoms can represent well the electrostatic interactions of the C_α atoms. Furthermore, the atoms in the elongated bonds have almost the same QM charges as those for the normal bond length, and the elongation does not significantly alter the charges of other QM atoms. In general, there is good accord between the OPLS-AA charges in the MM region and the semiempirical QM charges for corresponding QM atoms. Consequently, use of the OPLS-AA Lennard-Jones parameters for both the QM and MM atoms in eq 1 has been appropriate.[35−38]

Computational Details

PM3 Validation. As parameters for Mg^{2+} have not been obtained for the PDDG methods, the present QM/MM calculations used PM3 for the QM part. However, before applying the PM3 method to the full enzymatic reaction, a reduced QM model with 28 heavy atoms comprising Mg^{2+}, two water molecules, two formates to represent the side chains of Glu185 and Asp211, pyruvate enolate, and the 2-pyrone derivative shown in Scheme 2 was investigated. The energetics and structures for species along the two possible routes to form the bicyclic intermediate were obtained with PM3 and compared to results of DFT calculations at the B3LYP/6-31+G(d)//3-21+G level using Gaussian 98.[47] Stationary points for the Michael and bicyclic intermediates were obtained for both PM3 and DFT. Models for the Michael and aldol TSs, with a single negative eigenvalue corresponding to the appropriate C−C bond formation, were also identified with both PM3 and DFT.

(33) Monard, G.; Loos, M.; Théry, V.; Baka, K.; Rivail, J.-L. *Int. J. Quantum Chem.* **1996**, *58*, 153−159.
(34) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(35) Kaminski, G. A.; Jorgensen, W. L. *J. Phys. Chem. B* **1998**, *102*, 1787−1796.
(36) Chandrasekhar, J.; Shariffskul, S.; Jorgensen, W. L. *J. Phys. Chem. B* **2002**, *106*, 8078−8085.
(37) Repasky, M. P.; Guimarães, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6663−6672.
(38) Guimarães, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6892−6899.
(39) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(40) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.
(41) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.
(42) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601−1622.
(43) Tubert-Brohman, I.; Guimarães, C. R. W.; Repasky, M. P.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 138−150.
(44) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87−109.
(45) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291−1304.
(46) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9, Gaussian, Inc.: Pittsburgh, PA, 1998.

However, no Diels−Alder TS could be located at either level. To obtain a Diels−Alder TS model, the number of degrees of freedom was reduced by removing the two water molecules and the two formates, but this procedure worked only for the PM3 method; a single negative eigenvalue corresponding to the appropriate reaction coordinate, the symmetric combination of the two forming C−C bonds (ξ), was found. The PM3 Diels−Alder TS model was finally obtained by "resaturating" Mg^{2+} with the water molecules and formates and by energy minimizing the final structure while keeping the forming C−C bond lengths fixed at the value found for the system with fewer degrees of freedom. B3LYP calculations using a larger basis set [6-31+G(d)//6-31+G(d)] were performed in a further simplified QM model consisting of 2-pyrone and the enolate of acetaldehyde, and again, no Diels−Alder TS could be obtained.[48] The quantitative accord between the PM3 and DFT results is good, as described in the Results.

QM Region. The QM region, which consists of Mg^{2+}, Glu185, Asp211, neighboring peptide bonds, link atoms, two water molecules, pyruvate enolate, and the 2-pyrone derivative, has a total of 77 atoms. The QM/MM partition was done for Glu185 and Asp211, as shown in Figure 1, using a total of four link atoms. The initial coordinates for the heavy atoms of Glu185, Asp211, neighboring peptide bonds, the two water molecules, and the Mg^{2+} were taken from the crystal structure, whereas initial coordinates for the water hydrogens, pyruvate enolate, and the 2-pyrone derivative were taken from the reduced PM3 models for both the Michael and Diels−Alder TSs, generating two preliminary TS models for the QM/MM simulations. Coordinates for the hydrogens of Glu185 and Asp211 were obtained using standard bond lengths, bond angles, and dihedral angles. As described above, the bonds between the link atoms and the QM region were elongated by 0.5 Å to allow for effective spatial overlap between the hydrogens and the C_α atoms.

QM/MM System. Cartesian coordinates for the 1.70 Å *Macrophoma commelinae* MPS crystal structure complexed to pyruvate enolate and Mg^{2+} (Brookhaven Protein Data Bank code: 1IZC) were employed.[12] Of the six active sites for MPS identified in the 1IZC structure, only the first was retained and taken as the center of the reacting system (Figure 2). The selected active site contains residues from the chains A and C of the enzyme. To insert the QM region into the protein environment, the Mg^{2+} and the Glu185 and Asp211 heavy atoms corresponding to the preliminary Michael and Diels−Alder TS models were optimally overlaid by matching the coordinates of the corresponding atoms. Subsequently, all atoms from the crystal structure that were present in the QM model (except for the C_α atoms overlapped with the link atoms) were deleted. Residues with any atom within 15 Å from the center of the reacting system were retained in the simulations (residues in yellow and magenta), and any clipped residues were capped with acetyl and/or *N*-methylamine groups. The enzymatic system then has 2085 atoms, including 123 amino acid residues.

Degrees of freedom for the protein backbone atoms were not sampled in the QM/MM calculations. Only side chains of residues with any atom within 10 Å from the center of the solute were varied (residues in yellow). As the QM region has a charge of −2 e, charge neutrality was imposed by having a total MM charge of +2 e; charged residues near the active site were assigned normal protonation states at physiological pH, and the adjustments for neutrality were made to the most distant residues. The entire system was solvated with a 22 Å radius water cap consisting of 794 molecules, and a half-harmonic potential with a force constant of 1.5 kcal/mol·Å^2 was applied to water molecules

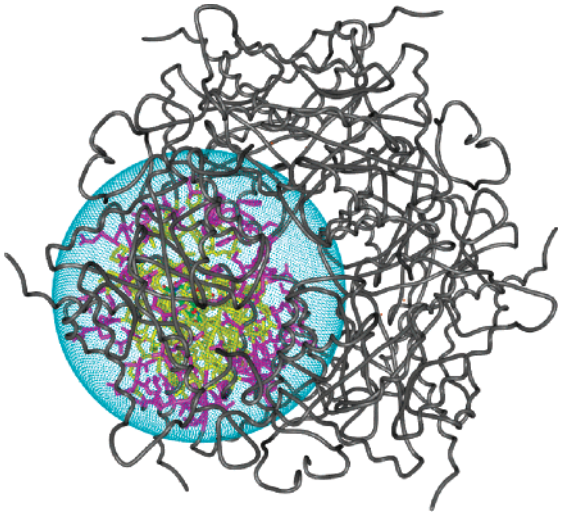


Figure 2. MPS active site model surrounded by a 22 Å water cap (coordinates from 1IZC).

at distances greater than 22 Å from the center of the solute to inhibit evaporation. The preparation of the enzymatic system was much facilitated by use of the Chop delegate and MidasPlus 2.1.[49] As the CM charge models[44,45] have not been parametrized for systems containing Mg, partial atomic charges for the QM region were obtained from the Mulliken procedure, while standard Lennard-Jones parameters were assigned for QM atoms (for C, $\sigma = 3.550$ Å, $\epsilon = 0.070$ kcal/mol; for H on C, $\sigma = 2.460$ Å, $\epsilon = 0.030$ kcal/mol; for N, $\sigma = 3.250$ Å, $\epsilon = 0.170$ kcal/mol; for O, $\sigma = 3.000$ Å, $\epsilon = 0.170$ kcal/mol; for Mg, $\sigma = 3.400$ Å, $\epsilon = 0.050$ kcal/mol; for H on a heteroatom, $\sigma = \epsilon = 0.000$).[46]

PMF Calculations. In this work, PMF calculations were used to investigate the relative stabilities for the TSs of both reaction mechanisms and, consequently, to identify the most viable route. Figure 3 represents a *hypothetical* free-energy surface for the two C−C bond formations between the 2-pyrone derivative and pyruvate enolate, generating the bicyclic intermediate in the protein via stepwise and concerted reaction mechanisms. The aldol and Michael TSs are characterized by ensembles of structures near free-energy maxima for elongating $R_{CC}2$ and then $R_{CC}1$ in the product (see Scheme 2), while the Diels−Alder TS is characterized by a collection of structures displaying the free-energy maximum for ξ, the symmetric combination of $R_{CC}1$ and $R_{CC}2$.

In theory, two-dimensional (2D) PMF calculations could be performed to characterize all species in the free-energy surface. However, the computational demands to obtain the full free-energy surface in the protein complex are very high, especially because of the large size of the QM region. Thus, less demanding 1D PMF calculations were performed to locate the Michael and aldol TSs in the protein complex, while 2D PMF calculations on a reduced region of the full free-energy surface were performed to locate the Diels−Alder TS. Specifically, $R_{CC}1$ was followed from the preliminary Michael TS model to the free-energy maximum for $R_{CC}1$, corresponding to the optimized Michael TS, and to the free-energy minimum for $R_{CC}1$, corresponding to the Michael intermediate. The same procedure was done for the second step on the Michael−aldol route. $R_{CC}2$ was followed from the value of 4.00 Å to the free-energy minimum for $R_{CC}2$, corresponding to the Michael intermediate, and to the free-energy maximum for $R_{CC}2$, corresponding to the optimized aldol TS. The initial value for $R_{CC}2$ is long enough to guarantee the location of both the Michael intermediate and the aldol TS.

(48) B3LYP/6-31+G(d) calculations were performed to study the reaction between 2-pyrone and the enolate of acetaldehyde. No Diels−Alder TS could be obtained using a larger basis set for this additional, simplified model. Initial configurations for the Diels−Alder TS were optimized to the Michael TS, with a single negative eigenvalue corresponding to the appropriate C−C bond-forming mechanism. The bicyclic intermediate for this model is no longer a stationary point at this level of calculation. Initial configurations for the bicyclic intermediate optimized to the Michael intermediate. The potential energy difference between the Michael TS and intermediate is 13.5 kcal/mol.

(49) Tirado-Rives, J. *Chop*; Yale University: New Haven, CT, 2002. (b) Huang, C.; Pettersen, E.; Couch, G.; Ferrin, T. *MidasPlus 2.1*; University of California: San Francisco, CA, 1994.

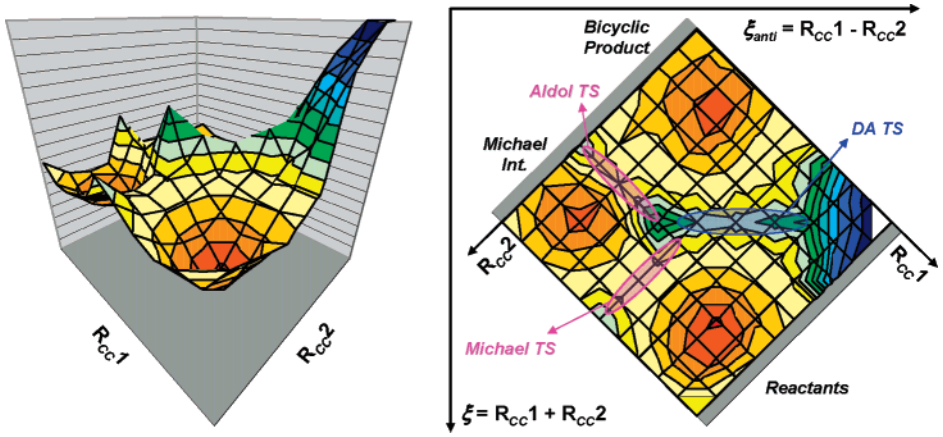


Figure 3. Hypothetical free-energy surface for the reaction between the 2-pyrone derivative and pyruvate enolate via stepwise and concerted reaction mechanisms.

It should be noted that the Michael intermediate is a collection of structures displaying a range of values for both $R_{CC}1$ and $R_{CC}2$; the free-energy minima for $R_{CC}1$ and $R_{CC}2$ correspond to subsets of the entire population. Thus, a small correction for the ΔG between the minimum and maximum for $R_{CC}1$ and $R_{CC}2$ was considered and calculated by eq 2:

$$\Delta G_{\text{all}\rightarrow\text{subset}} = -RT\ln P_{\text{subset}} \qquad (2)$$

Equation 2 represents the free energy associated with the conversion of all Michael intermediate structures into the subset displaying the free-energy minimum geometry for $R_{CC}1$ (or $R_{CC}2$), and P_{subset} is the probability of sampling the particular subset. $\Delta G_{\text{all}\rightarrow\text{subset}}$ is then added to the ΔG values between the minimum and maximum for $R_{CC}1$ and $R_{CC}2$ to yield the free-energy differences between the Michael intermediate and the Michael and aldol TSs, respectively.

Two-dimensional PMF calculations mapping $R_{CC}1$ and $R_{CC}2$ were performed in a region surrounding the preliminary Diels–Alder TS model. As the free-energy difference between the Michael and aldol TSs may be obtained as detailed above, only the free-energy difference between the Michael and Diels–Alder TSs was considered. Once the true Diels–Alder TS is located, its stability can be compared to that of the two TSs on the Michael–aldol route by performing a series of PMF calculations connecting the Michael and the Diels–Alder TSs.

To calculate P_{subset}, the radial distribution function for a general reaction coordinate, R ($g(R)$), was obtained from the respective PMF curve ($G(R)$), which was computed using increments of 0.02 Å. The probabilities of sampling R ($P(R)$ or P_{subset}) were obtained as the product of the respective $g(R)$ and the volume element of the configuration space corresponding to the coordinate R. The volume element can be interpreted as a quantity proportional to the probability of sampling R with the potential function set to zero.[26] In the present case, the volume element is $4\pi R^2$. This procedure was detailed previously.[37,38] The 2D PMF map was also obtained using increments of 0.02 Å.

MC Simulations Protocol. PMF curves were obtained via MC statistical mechanics at 25 °C using FEP calculations. To ensure convergence, each MC/FEP simulation was extensive. Initial reorganization of the solvent was performed for 5×10^6 configurations. This was followed by 30×10^6 configurations of full equilibration and 30×10^6 configurations of averaging for each window in each MC/FEP calculation using double-wide sampling. Subsequently, for further analyses, energies for the QM and MM regions and the interaction energies between QM and MM components were averaged via MC simulations at 25 °C for the TSs found for the two reaction mechanisms. In these cases, initial reorganization of the solvent was performed for 5×10^6 configurations. This was followed by 200×10^6 configurations of full equilibration and 200×10^6 configurations of averaging for each TS model. Computation of the QM energy and atomic charges is required for every MC move of the QM region, which was attempted every 400 configurations. Specifically, in an FEP calculation using double-wide sampling, three QM calculations are performed for each accepted move of the QM region corresponding to the value for the reaction coordinate and its two perturbed values. A total of nearly 200 million QM calculations were performed in this work. The combined QM/MM method as implemented in MCPRO 2.1[50] was used to perform all calculations, and established procedures including Metropolis and preferential sampling were employed.[24] Statistical uncertainties were obtained from the batch means procedure with batch sizes of 1×10^6 configurations.[24] The MM part of the protein was represented with the OPLS-AA force field;[46] the TIP4P model was used for water,[51] and residue-based cutoffs of 10 Å were employed for all nonbonded solute–solute and solvent–solute interactions.

Results and Discussion

PM3 versus B3LYP. The energetics and structures for the species along the two possible routes to form the bicyclic intermediate were obtained from PM3 and DFT calculations on a reduced model of the system. As mentioned above, the reduced QM model consists of the Mg^{2+} ion, two water molecules, two formates representing the side chains of Glu185 and Asp211, pyruvate enolate, and the 2-pyrone derivative shown in Scheme 1. Stationary points at both levels for the Michael and aldol TSs and the Michael and bicyclic intermediates are depicted in Figure 4. The PM3 Diels–Alder TS model obtained as described in the Computational Details is also displayed. The $R_{CC}1$ and $R_{CC}2$ distances are 2.02 and 2.38 Å, respectively. As mentioned above, no Diels–Alder TS could be obtained at the B3LYP level. The $R_{CC}1$ distance for the Michael TS is 2.40 Å with PM3 and 2.50 Å with DFT, whereas for the Michael intermediate, it is 1.56 and 1.59 Å, respectively. For the aldol TS, the $R_{CC}2$ distance is 2.04 Å with PM3 and 2.03 Å with DFT. Finally, the $R_{CC}1$ and $R_{CC}2$ distances are 1.54 and 1.65 Å, respectively, for the PM3 bicyclic intermediate, whereas the DFT calculations give values of 1.55 and 1.65 Å. Although both levels of theory provide almost identical values for $R_{CC}1$ and $R_{CC}2$, the PM3 and DFT structures for the Michael and bicyclic intermediates and aldol TS are slightly different; one of the water molecules coordinated to Mg^{2+} donates a hydrogen bond to the C2-carbonyl oxygen of the pyrone moiety

(50) Jorgensen, W. L.; Tirado-Rives, J. *MCPRO*, version 2.1; Yale University: New Haven, CT, 2004.
(51) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

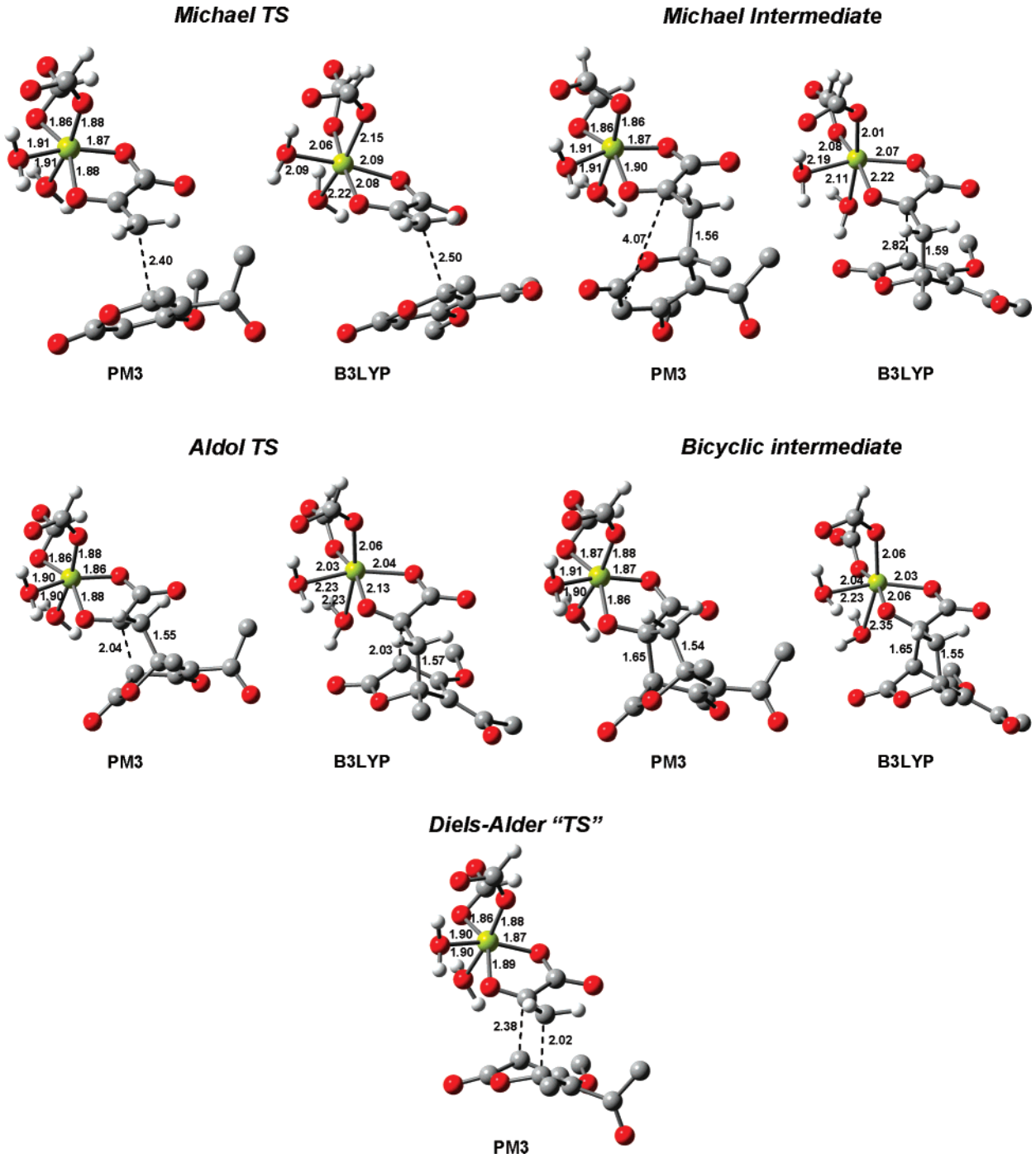


Figure 4. Structures for species along the two possible routes to form the bicyclic intermediate obtained from PM3 and DFT calculations.

with DFT, but not with PM3. It can also be seen that for both PM3 and DFT, the overall coordination geometry to Mg^{2+}, which remains fairly unaltered for all species along the reaction paths, is octahedral and very similar to the coordination geometry observed in the crystal structure (Figure 5). However, the distances from PM3 are ca. 0.2−0.3 Å shorter than those from the DFT calculations and the crystal structure.

Table 1 compares the relative energies for these stationary points from the PM3 and B3LYP/6-31+G(d)//3-21+G calculations. The Michael intermediate is 26.6 kcal/mol more stable than the Michael TS with DFT and 32.5 kcal/mol with PM3. The aldol TS is 20.3 and 7.9 kcal/mol more stable than the Michael TS with DFT and PM3, respectively. The bicyclic intermediate is 21.6 kcal/mol more stable than the Michael TS with DFT, while the corresponding value computed from PM3 is only 10.1 kcal/mol. However, the energy differences between

Table 1. Relative Energies (kcal/mol) from the PM3 and DFT Calculations on the Model Reaction

	PM3	B3LYP/6-31+G(d)//3-21+G
Michael TS	0.0	0.0
Michael intermediate	−32.5	−26.6
aldol TS	−7.9	−20.3
bicyclic intermediate	−10.1	−21.6
Diels−Alder TS[a]	11.4	n.a.

[a] See Computational Details for location procedure.

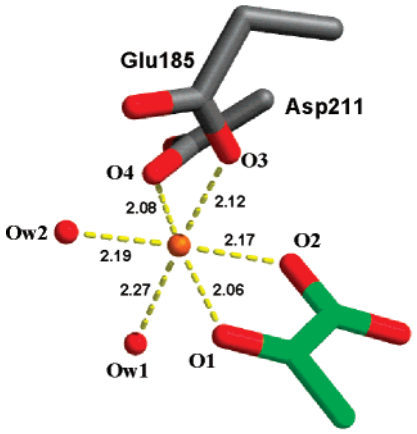


Figure 5. Coordination geometry observed for the crystal structure.

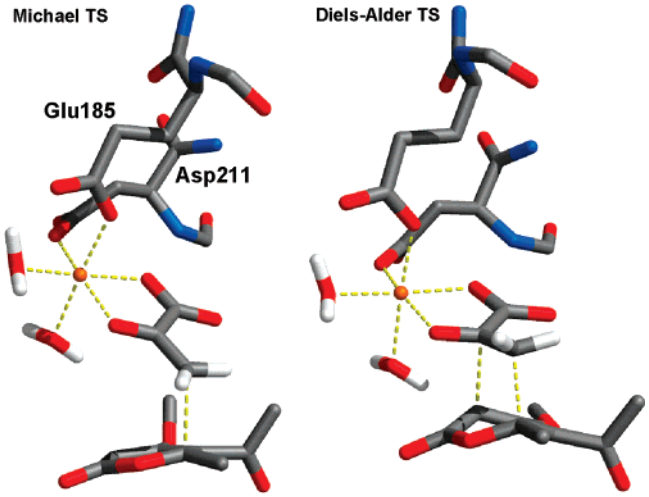


Figure 6. Initial configurations of the QM region (77 atoms) for the QM/MM simulations of both the Michael and Diels−Alder TSs.

the aldol TS and the bicyclic intermediate given by both levels of theory are in close agreement. As for the Diels−Alder TS, it is 11.4 kcal/mol less stable than the Michael TS at the PM3 level, but the TS model for the concerted route is not a true stationary point; the PM3 Diels−Alder TS was generated by removing the two water molecules and the two formates, "resaturating" the Mg^{2+}, and energy minimizing this structure with the forming of C−C bond lengths fixed. Overall, both QM levels find the Michael intermediate as the most stable species, followed by the bicyclic intermediate, the aldol TS, and the Michael TS, while with PM3, the Michael TS is more stable than the Diels−Alder TS. The general agreement between the PM3 and DFT results for the energetics and structures supports the use of the semiempirical method in the QM/MM simulations.

The Michael−Aldol Route. Figure 6 shows the QM region of the unoptimized Michael and Diels−Alder TS models used as starting configurations in the PMF calculations. Figure 7 plots the PMF curves following the reaction coordinates $R_{CC}1$ and $R_{CC}2$ in the stepwise path to generate the bicyclic intermediate and the probability of sampling $R_{CC}1$ and $R_{CC}2$ for the Michael intermediate. The PMF curve for the first step on the Michael−aldol route (Figure 7a) was zeroed at $R_{CC}1 = 2.40$ Å, which is the value for the preliminary (unoptimized for the protein environment) Michael TS. The free-energy minimum ($R_{CC}1 = 1.60$ Å) and maximum ($R_{CC}1 = 2.08$ Å), corresponding to the optimized Michael TS, were identified from the PMF curve, and the ΔG between them amounts to 11.5 kcal/mol. The probability of sampling the subset displaying the free-energy minimum geometry for $R_{CC}1$ from all Michael intermediate structures was computed to be 23.6%, giving a $\Delta G_{all\rightarrow subset}$ value of 0.8 kcal/mol. Thus, the final value for the free-energy difference between the Michael intermediate and the Michael TS on the stepwise route is 12.3 ± 0.1 kcal/mol.

Figure 7b shows the PMF curve for the second step on the Michael−aldol route. The free-energy minimum occurs at 3.48 Å, and the maximum, corresponding to the aldol TS, is at 2.20 Å; the ΔG between them amounts to 16.6 kcal/mol. The probability of sampling the subset displaying the free-energy minimum geometry for $R_{CC}2$ from all Michael intermediate structures was computed to be 10.2%, giving a $\Delta G_{all\rightarrow subset}$ value of 1.3 kcal/mol. Thus, the final value for the free-energy difference between the Michael intermediate and the aldol TS is 17.9 ± 0.2 kcal/mol. Consequently, the aldol TS is 5.6 kcal/mol less stable than the Michael TS.

The Diels−Alder Route. The PM3 and DFT calculations for the reduced QM model of the reacting system did not identify a Diels−Alder TS. As discussed above, the PM3 Diels−Alder TS, which was used to build the QM region of the preliminary QM/MM Diels−Alder TS model ($R_{CC}1 = 2.02$ Å, $R_{CC}2 = 2.38$ Å), is not a true stationary point. To locate the concerted TS in the protein environment, the 2D PMF calculations should cover the region surrounding the hypothetical Diels−Alder TS in Figure 3. More specifically, a free-energy map was built for the vicinity of the preliminary Diels−Alder TS model, a good starting point for the Diels−Alder TS search in the enzyme active site. Figure 8 shows that no Diels−Alder TS is again apparent. Instead, the free-energy map points to a stepwise mechanism, with formation of the first C−C bond ($R_{CC}1$) generating the stable Michael intermediate, followed by formation of the second C−C bond ($R_{CC}2$).

Diels−Alder versus Michael−Aldol. At this point, it is very clear that the Diels−Alder TS for a reduced QM model in the gas phase or in the protein environment either does not exist or is very unstable, at least at the levels of calculation used in this work. Even without locating the true Diels−Alder TS in the free-energy map, it is worthwhile to compare the stabilities of the TSs identified for the Michael−aldol route with those of the preliminary Diels−Alder TS. As the free-energy difference between the Michael and aldol TSs in the stepwise route has already been calculated (5.6 kcal/mol favoring the Michael TS), only the free-energy difference between the Michael TS and the unoptimized Diels−Alder TS was computed. The conversion of the Diels−Alder TS model to Michael TS was performed through a series of PMF calculations and found to be favorable by 17.7 kcal/mol. As a consequence, the unoptimized Diels−Alder TS is 12.1 kcal/mol less stable than the aldol TS.

Energy Contributions to the Free-Energy Differences between the TSs. Long MC simulations (200×10^6 configurations of equilibration followed by 200×10^6 configurations of averaging) were performed for the Michael and aldol TSs identified by the PMF calculations and the preliminary Diels−Alder TS model. Interatomic distances,[52] extracted from 200 structures saved every 1×10^6 configurations, were averaged

(52) Average values for the coordination distances were obtained with the Mok program (www.perlmol.org/pod/mok.html).

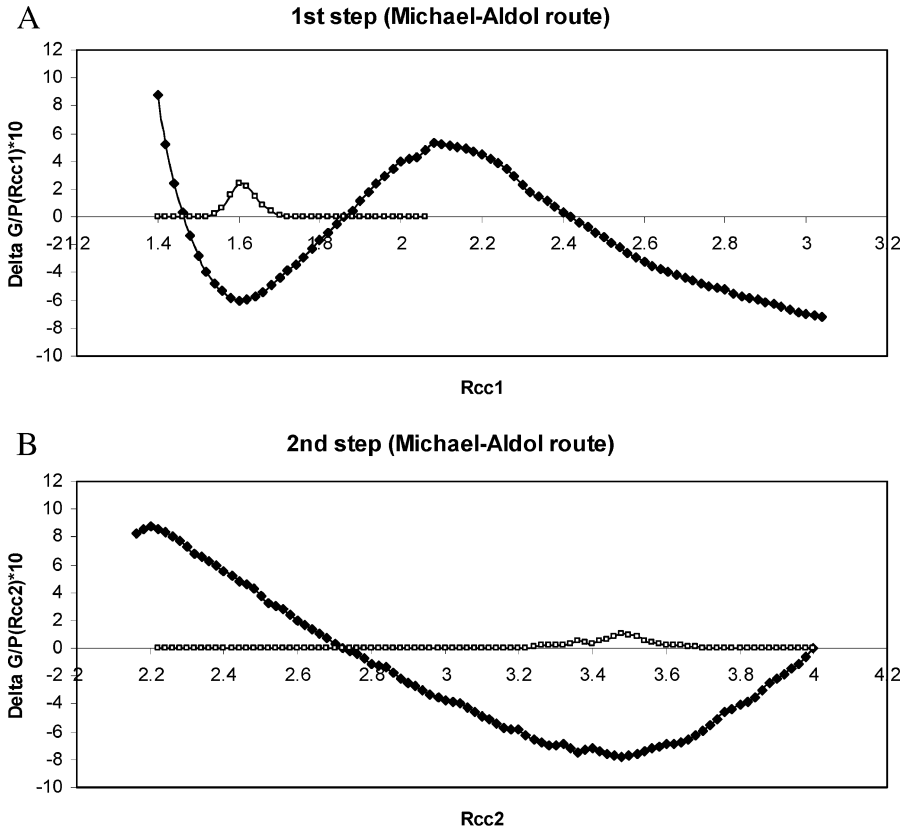


Figure 7. PMF curves following the reaction coordinates $R_{CC}1$ (a) and $R_{CC}2$ (b) in the stepwise path, and the probability of sampling $R_{CC}1$ and $R_{CC}2$ for the Michael intermediate.

Table 2. Average Coordination Distances Obtained for All TS Models[a]

distance	Diels–Alder TS	Michael TS	aldol TS	crystal structure[b]
Mg−O1	1.88 ± 0.03	1.86 ± 0.02	1.87 ± 0.03	2.06
Mg−O2	1.86 ± 0.02	1.86 ± 0.03	1.86 ± 0.03	2.17
Mg−O3	1.84 ± 0.03	1.83 ± 0.03	1.85 ± 0.03	2.12
Mg−O4	1.85 ± 0.03	2.46 ± 0.04	1.85 ± 0.03	2.08
Mg−Ow1	2.42 ± 0.04	2.40 ± 0.05	2.42 ± 0.05	2.27
Mg−Ow2	2.44 ± 0.04	2.39 ± 0.03	2.41 ± 0.04	2.19

[a] Obtained with the Mok program.[52] [b] For pyruvate enolate bound to MPS.[12]

to analyze the coordination between the Mg^{2+} and the side-chain carboxyl oxygens from residues Glu185 and Asp211, the two water molecules, and the C2-carbonyl and C1-carboxyl oxygens of pyruvate enolate, all part of the QM region. Table 2 (see also Figure 5) shows that the average coordination distances are mostly the same for all TSs, and they compare fairly well to the crystal structure for bound pyruvate enolate. For the Diels−Alder and aldol TSs, the distances for the coordination between Mg^{2+} and the carbonyl and carboxyl oxygens are ca. 0.2−0.3 Å shorter than those for the crystal structure, whereas the coordination to the water molecules are ca. 0.2 Å longer. However, for the Michael TS, the average coordination distance between Mg^{2+} and the carboxyl oxygen from residue Asp211 (oxygen O4) is ca. 0.6 Å longer than that for the other TSs and 0.4 Å longer than that for the crystal structure.

Energies for the QM (E_{QM}) and MM (E_{MM}) regions and the interaction energies between them were also averaged for all TSs. To facilitate the analysis, $E_{QM/MM}$ was broken into the interactions between the QM region and the solvent ($E_{QM/Solv}$) and the QM region and the protein ($E_{QM/Ptn}$). It should be noted that long MC simulations for all TSs were carried out to obtain reasonably converged results. Figure 9 shows that the convergence rate for all energy components is slow, so the energy differences in Table 3 between the TSs have significant statistical uncertainties. Nevertheless, the results indicate that the aldol TS has the most favorable E_{QM}, followed by the Michael and Diels−Alder TSs; E_{QM} for the aldol and Michael TSs are 22.6 and 10.8 kcal/mol, respectively, more stable than E_{QM} for the Diels−Alder TS. These results agree with the PM3 calculations in the gas phase for the reduced QM model (Table 1), which point to aldol and Michael TSs that are 19.3 and 11.4 kcal/mol, respectively, more stable than those of the Diels−Alder TS.

Figure 10 illustrates the interactions between the QM region and the nearest residues for all TS models. The average structures that emerged feature an extensive hydrogen-bonded network between the QM region and the protein that is present for every TS. More specifically, the QM water molecule (W2) is hydrogen bonded to the side chains of the MM residues His125C and Ser148A, which also donate hydrogen bonds to the carboxyl oxygen of Asp211C. Additional hydrogen bonds are formed between the hydroxyl oxygen of Ser148A and the QM water (W1) and between the backbone nitrogen of His125C and the carboxyl oxygen of Glu185C. The C2-carbonyl oxygens of pyruvate enolate and the 2-pyrone derivative accept hydrogen bonds from the guanidinium nitrogens of Arg101C, which is in turn hydrogen bonded to Asp70C. In contrast to suggestions by Ose and co-workers,[12] the MC simulations showed no hydrogen bonds between the side chain of Tyr169A and the

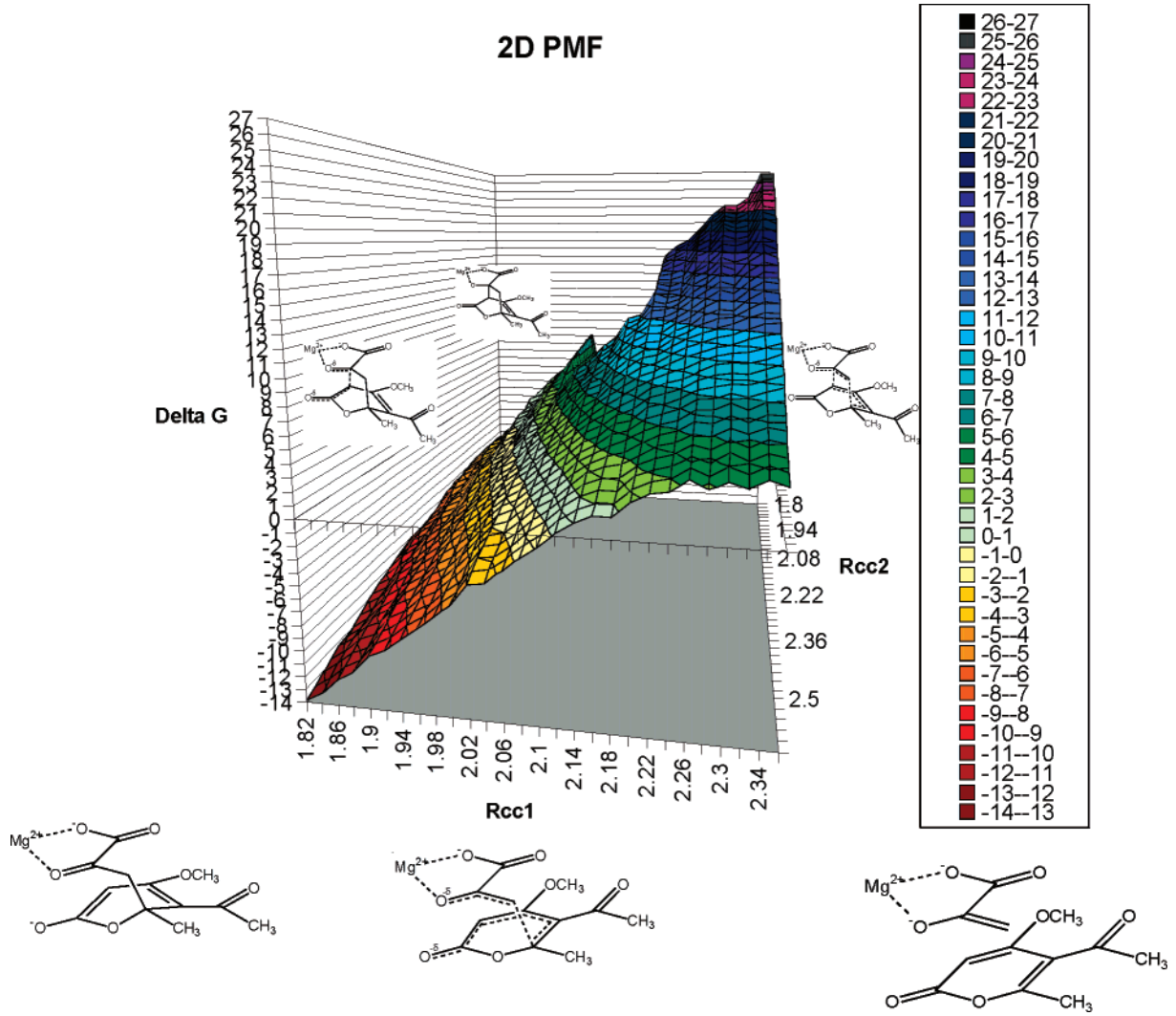


Figure 8. Free-energy map near the preliminary Diels−Alder TS model, obtained from 2D PMF calculations.

Table 3. Average Intramolecular and Intermolecular Energies (kcal/mol), and Energy Differences Obtained from Long MC Simulations for All TS Models[a]

	Diels–Alder TS	Michael TS	aldol TS
E_{QM}	−575.8 ± 0.2	−586.6 ± 0.2	−598.4 ± 0.2
$E_{QM/Solv}$	−128.3 ± 0.4	−117.1 ± 0.5	−128.8 ± 0.5
$E_{QM/Ptn}$	−454.6 ± 0.3	−458.5 ± 0.3	−451.1 ± 0.5
E_{MM}	−9138.8 ± 2.3	−9157.8 ± 2.4	−9144.9 ± 2.0
E_{Total}	−10297.5 ± 2.4	−10320.0 ± 2.5	−10323.2 ± 2.1
ΔE_{QM}	0	−10.8 ± 0.3	−22.6 ± 0.2
$\Delta E_{QM/Solv}$	0	11.2 ± 0.6	−0.5 ± 0.6
$\Delta E_{QM/Ptn}$	0	−3.9 ± 0.4	3.5 ± 0.6
ΔE_{MM}	0	−19.0 ± 3.4	−6.1 ± 3.0
ΔE_{Total}	0	−22.5 ± 3.5	−25.7 ± 3.2

[a] MC simulations, consisting of 200 × 10^6 configurations of equilibration followed by 200 × 10^6 configurations of averaging, were performed with the reaction coordinates for the Michael ($R_{CC}1$) and aldol ($R_{CC}2$) TSs fixed at the values identified by the PMF calculations. As the true Diels−Alder TS in the protein environment was not identified by the 2D PMF map, $R_{CC}1$ and $R_{CC}2$ were fixed at the values for the Diels−Alder TS model.

2-pyrone; the C5-acyl−carbonyl and C4-methoxy oxygens of the 2-pyrone are hydrogen bonded to water molecules. It is also worth mentioning the hydrogen bond interactions described by the PM3 Hamiltonian within the QM region. The C1-carboxyl oxygens of pyruvate enolate are hydrogen bonded to the backbone nitrogen of Asp211C; W2 donates a hydrogen bond to the carboxyl oxygen of Glu185C, and W1 donates a hydrogen bond to the C2-carbonyl oxygen of the pyrone moiety; as mentioned above, this interaction is absent in the reduced PM3 models for the TSs in the gas phase but occurs with DFT. Although the hydrogen-bonded network exists for all TSs, Table 3 shows that $E_{QM/Ptn}$ for the Michael TS is the most favorable, followed by the Diels−Alder and aldol TSs.

To investigate the more favorable interactions between the QM region and the protein for the Michael TS, per-residue interaction energies were calculated during the MC simulations. Figure 11 plots the per-residue interaction energy differences between the Michael and aldol TSs (Figure 11a) and Michael and Diels−Alder TSs (Figure 11b). Very similar profiles were obtained in both cases, indicating that the QM regions of both the aldol and Diels−Alder TSs interact similarly with the protein. This should be expected since the transition structures differ primarily by the aldol TS being just tighter than the Diels−Alder TS, with $R_{CC}1$ and $R_{CC}2$ values of ca. 1.60 and 2.20 Å compared to 2.02 and 2.38 Å.

The QM region of the Michael TS interacts more favorably with the residues Arg146A, Ser148A, Asp70C, His125C, and Lys188C (Figure 11). As for the aldol and Diels−Alder TSs, their QM regions interact better with Pro151A, Arg101C, Gly210C, Asp215C, and Met236C. The peak for residue

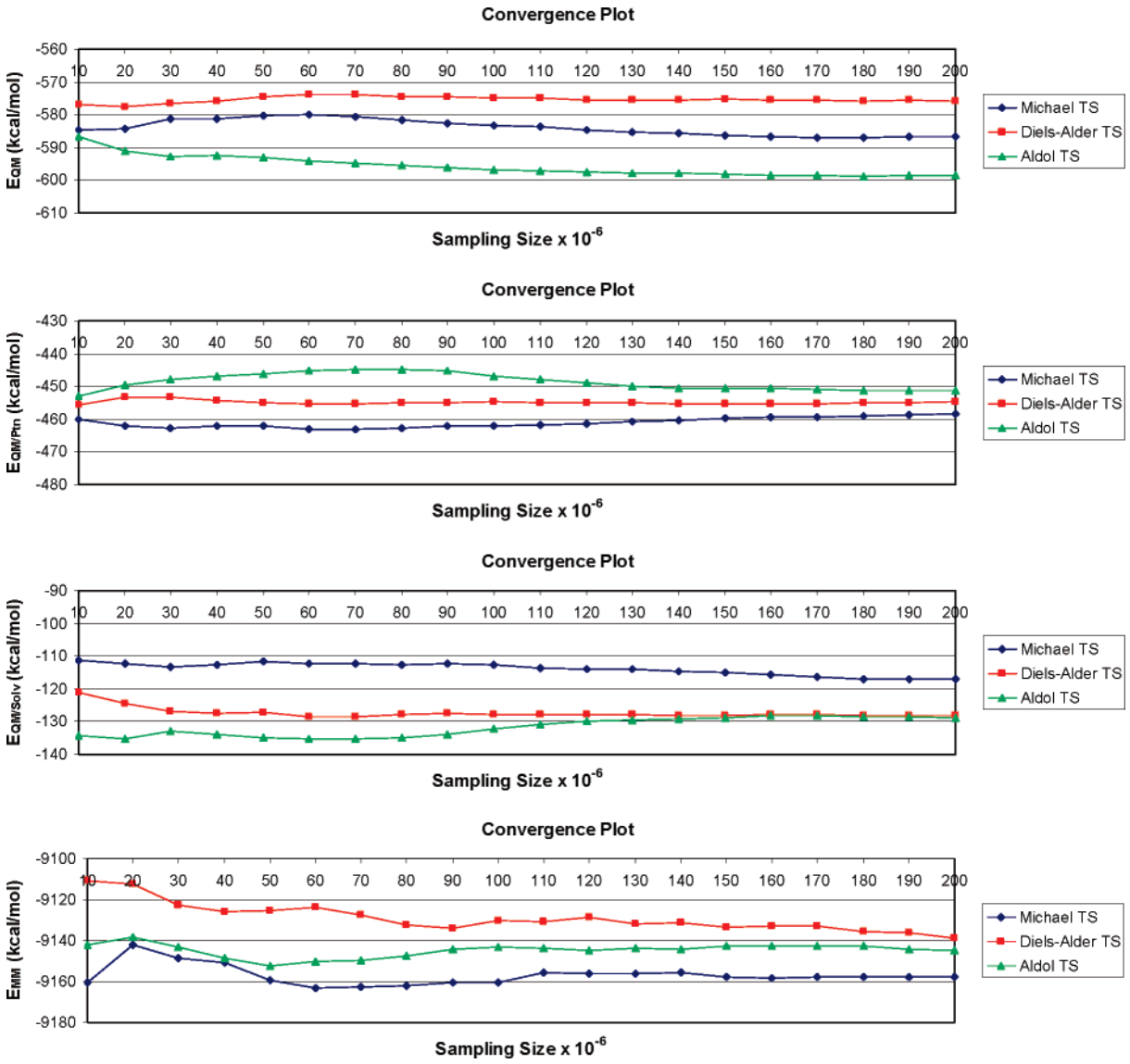


Figure 9. Convergence profiles for the energy components considered for the TS models.

Tyr169A in Figure 11b and its absence in Figure 11a indicates that the QM region of the Diels−Alder TS interacts more strongly with this residue than for the Michael and aldol TSs. Figure 11a,b also shows that residue Gln183C interacts more favorably with the QM region of the aldol TS, followed by the Michael and Diels−Alder TSs. It can be seen that part of the difference in interactions between the protein and the QM region of each TS is related to the more distant residues, such as Arg146A, Asp70C, Gln183C, Lys188C, Asp215C, and Met236C (some are not shown in Figure 10). This should be expected as the QM region has a charge of −2. Overall, the more favorable $E_{QM/Ptn}$ value for the Michael TS is probably related to the fact that it is more flexible than the aldol and Diels−Alder TSs. In other words, the Michael TS is more easily deformed in the MPS active site to optimize intermolecular interactions with the protein.

Counts of hydrogen bonds from each MC simulation show that the QM region of the Michael TS accepts, on average, 6.0 hydrogen bonds from the water molecules, while for the Diels−Alder and aldol TSs, this number is increased to 7.2 and 7.6, respectively. No hydrogen bonds are donated to the solvent for all TS models, which reflects the −2 charge. As the Michael TS is more open than both the Diels−Alder and aldol TSs, it occupies a larger space in the active site, reducing the space available for the water molecules. Consequently, the intermolecular interactions between the QM region and the solvent are more favorable for the aldol and Diels−Alder TSs than those for Michael TS, as shown in Table 3.

Table 3 also shows that the Michael TS has the most favorable E_{MM}, followed by the aldol and Diels−Alder TS. In other words, the environment, particularly, the enzyme, is more adapted to accommodate the TSs of the stepwise path than the Diels−Alder TS. The final energy differences indicate that the unoptimized Diels−Alder TS is 22.5 and 25.7 kcal/mol less stable than the first and second TSs in the stepwise route. These numbers compare to the computed free-energy differences of 17.7 and 12.1 kcal/mol. The differences cannot just be assigned to entropic factors in view of the statistical uncertainties for the energies. However, the results for both the energy and free-energy differences point to more stable TSs for the stepwise path than for the concerted one. The results indicate that the Michael−aldol mechanism is the route used by MPS to catalyze

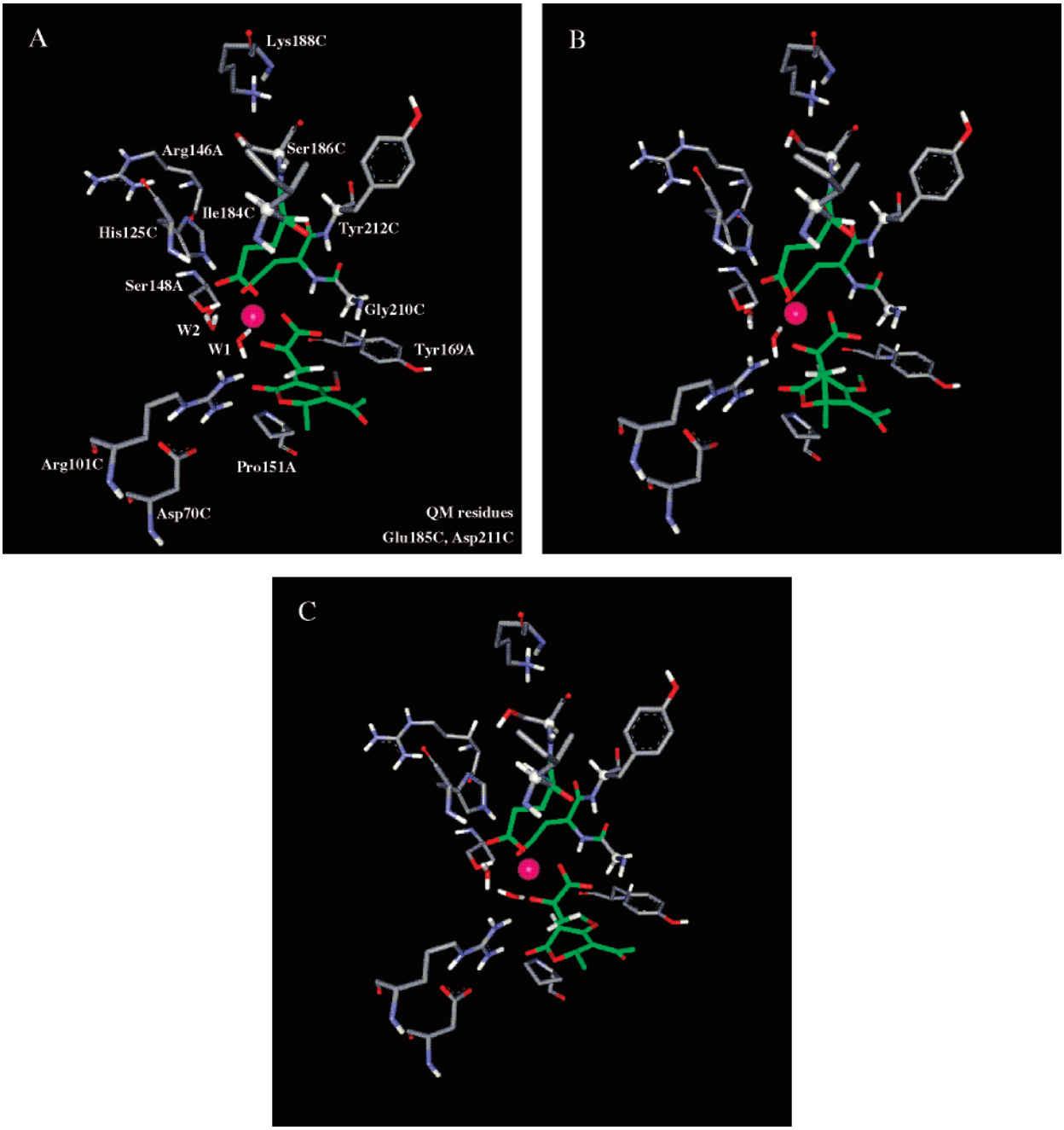


Figure 10. Snapshots illustrating the interactions between the QM region and selected residues for the Michael TS (a), the aldol TS (b), and the Diels–Alder TS (c). White dots represent the link atoms. The carbon atoms of the QM region are colored green.

the second step of the transformation of 2-pyrone derivatives into the corresponding benzoate analogues, and that the enzyme is not a Diels–Alderase.

Conclusions

In this work, mixed quantum and molecular mechanics (QM/MM), combined with Monte Carlo (MC) simulations and free-energy perturbation (FEP) calculations, were performed to investigate the most viable route for the second step of the transformation catalyzed by macrophomate synthase (MPS) of the phytopathogenic fungus *Macrophoma commelinae*. The QM/MM FEP calculations found that the Diels–Alder TS model is 17.7 and 12.1 kcal/mol higher in free energy than the Michael and aldol TSs in the stepwise route, respectively. The intervening Michael intermediate is computed to be 12.3 and 17.9 kcal/mol lower in free energy than the Michael and aldol TSs. Energies for the QM (E_{QM}) and MM (E_{MM}) regions and the interaction energies between them ($E_{QM/MM}$) were also averaged during MC simulations for the TSs. Although the statistical uncertainties are larger, the final results indicate that the Diels–Alder TS model is more than 20 kcal/mol higher in energy than that of the Michael and aldol TSs for the stepwise route. Therefore, the present computations consistently indicate that the Michael–aldol mechanism is the route used by MPS to catalyze the second step of the overall transformation in Scheme 1.

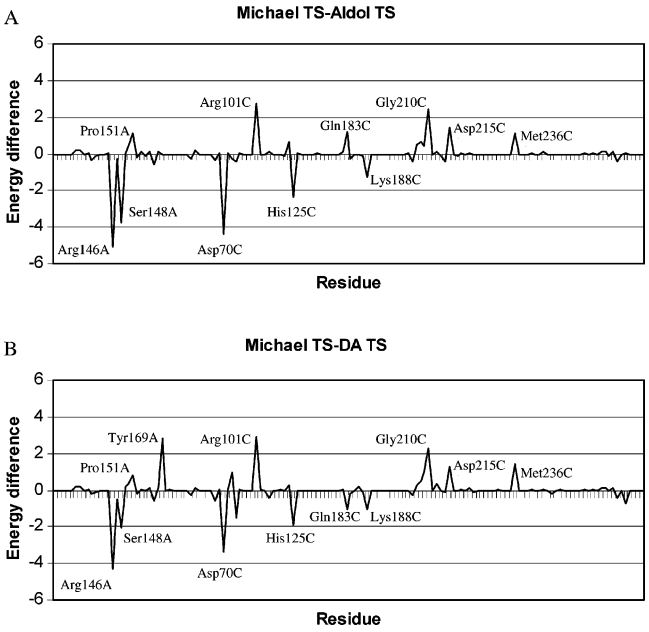


Figure 11. Per-residue $E_{QM/Ptn}$ differences between the Michael and aldol TSs (a) and between the Michael and Diels−Alder TSs (b).

The calculations in the gas phase also indicate that the Michael intermediate is more stable than the bicyclic intermediate, and that the retroaldol cleavage of the latter is facile. As the third step (decarboxylation with concomitant dehydration of the bicyclic intermediate) is rate-determining for the overall transformation, the present results indicate that experiments directed at trapping the Michael intermediate might be fruitful. The stepwise nature of the reaction mechanism could also possibly be supported by studies of kinetic isotope effects.

Acknowledgment. Gratitude is expressed to the National Institutes of Health (GM032136) for support, and to Dr. Matthew P. Repasky, Ivan Tubert-Brohman, Dr. Yukio Tominaga, Dr. J. Chandrasekhar, and Dr. Dmitry Klimenko for assistance. C.R.W.G. also acknowledges CNPq/Brazil (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for partial financial support.

JA043905B