JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

High-Level QM/MM Calculations Support the Concerted Mechanism for Michael Addition and Covalent Complex Formation in Thymidylate Synthase

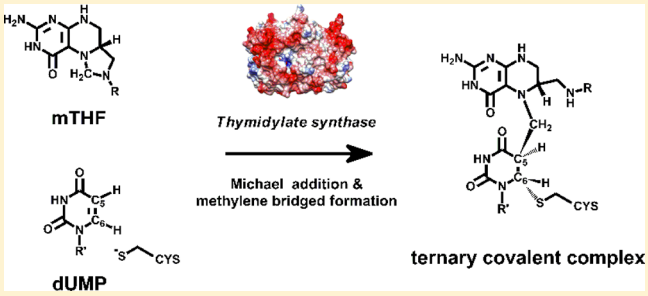
Nopporn Kaiyawet,[†] Richard Lonsdale,[‡,§] Thanyada Rungrotmongkol,[∥] Adrian J. Mulholland,*[,§] and Supot Hannongbua*[,†]

[†]Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, and [∥]Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand

[§]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom

Supporting Information

ABSTRACT: Thymidylate synthase (TS) is a promising cancer target, due to its crucial function in thymine synthesis. It performs the reductive methylation of 2′-deoxyuridine-5′-phosphate (dUMP) to thymidine-5′-phosphate (dTMP), using *N*-5,10-methylene-5,6,7,8-tetrahydrofolate (mTHF) as a cofactor. After the formation of the dUMP/mTHF/TS noncovalent complex, and subsequent conformational activation, this complex has been proposed to react via nucleophilic attack (Michael addition) by Cys146, followed by methylene-bridge formation to generate the ternary covalent intermediate. Herein, QM/MM (B3LYP-D/6-31+G(d)-CHARMM27) methods are used to model the formation of the ternary covalent intermediate. A two-dimensional potential energy surface reveals that the methylene-bridged intermediate is formed via a concerted mechanism, as indicated by a single transition state on the minimum energy pathway and the absence of a stable enolate intermediate. A range of different QM methods (B3LYP, MP2 and SCS-MP2, and different basis sets) are tested for the calculation of the activation energy barrier for the formation of the methylene-bridged intermediate. We test convergence of the QM/MM results with respect to size of the QM region. Inclusion of Arg166, which interacts with the nucleophilic thiolate, in the QM region is important for reliable results; the MM model apparently does not reproduce energies for distortion of the guanidinium side chain correctly. The spin component scaled-Møller−Plessett perturbation theory (SCS-MP2) approach was shown to be in best agreement (within 1.1 kcal/mol) while the results obtained with MP2 and B3LYP also yielded acceptable values (deviating by less than 3 kcal/mol) compared with the barrier derived from experiment. Our results indicate that using a dispersion-corrected DFT method, or a QM method with an accurate treatment of electron correlation, increases the agreement between the calculated and experimental activation energy barriers, compared with the semiempirical AM1 method. These calculations provide important insight into the reaction mechanism of TS and may be useful in the design of new TS inhibitors.



1. INTRODUCTION

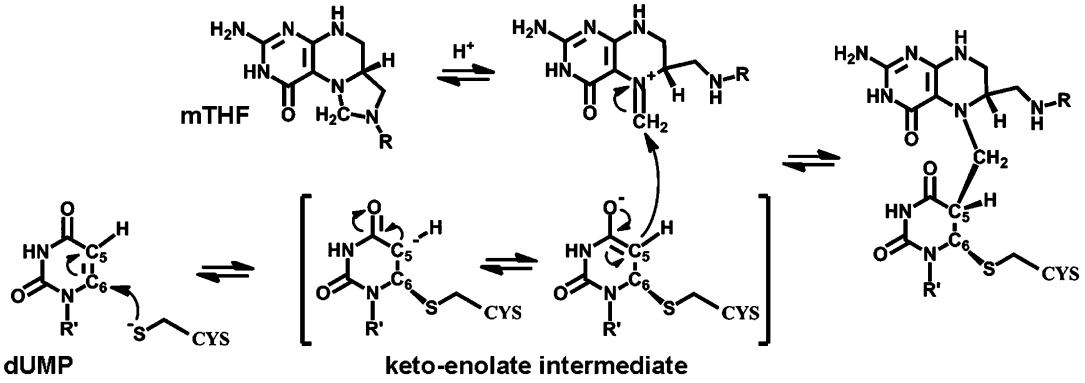
The DNA double helix is constructed by the sequential combination of nucleotide monomers, which are classified into four different types (adenine, cytosine, guanine, and thymine), depending on their nitrogenous bases. Thymine starvation directly leads to the apoptosis event called the "thymineless death phenomenon";[1−3] therefore thymine synthesis is a potential target for cancer chemotherapy. Thymidylate synthase[4] (TS, EC 2.1.1.45) catalyzes the reductive methylation of 2′-deoxyuridine-5′-phosphate (dUMP) to thymidine-5′-phosphate (dTMP, thymine monomer) using the cofactor *N*-5,10-methylene-5,6,7,8-tetrahydrofolate (mTHF) as a methylene donor and reductant.[5−9] Mechanistic details of the TS catalytic reaction have been proposed by Carreras and Santi.[5] While the main features of this mechanism are widely accepted, some details remain open to discussion.

The Michael addition and covalent complex formation steps have been proposed to occur either through a concerted or stepwise process by different authors.[5,10−13] The initiation of the Michael addition step proceeds by nucleophilic attack at the C-6 position of dUMP by the thiolate sulfur of Cys146 to form a thio-ether linkage. This step increases the nucleophilicity of C-5, which reacts with the methylene of mTHF to form a methylene bridged intermediate. In the stepwise mechanism (Scheme I), it is proposed that an enolate intermediate is formed prior to the ternary covalent intermediate, whereas in the concerted mechanism (Scheme II) the enolate intermediate is not observed. Hence, the existence of an enolate intermediate is still uncertain at present.

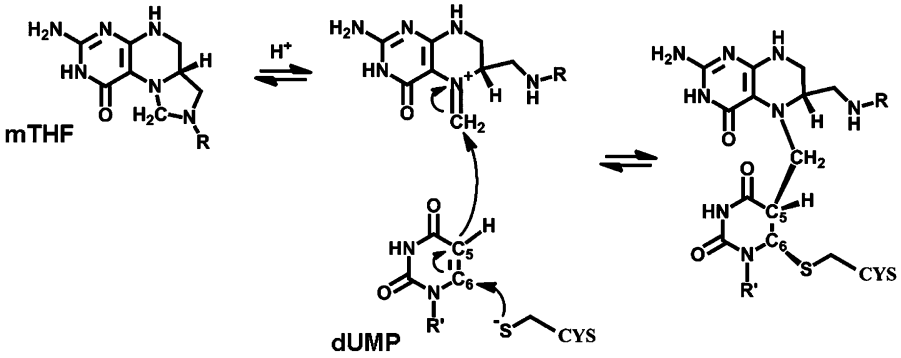
Received: June 10, 2014
Published: January 20, 2015

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

Scheme I. Stepwise Mechanism



Scheme II. Concerted Mechanism



5-Fluoro-dUMP (FdUMP), a metabolite form of 5-fluorouracil, is a structural mimic of dUMP, which can react with mTHF to form a ternary covalent intermediate in competition with the native dUMP substrate.[14,15] However, the subsequent β-elimination reaction step cannot proceed in the FdUMP complex, due to the strength of the C−F bond and poor leaving group behavior of fluorine, leading to the trapping of the ternary covalent intermediate, and subsequent irreversible enzyme inhibition. Therefore, the ternary covalent intermediate formation between dUMP and mTHF is an important reaction relevant to the mechanistic inhibition by the 5-fluorouracil anticancer drug. Although the methylene bridged intermediate between dUMP and mTHF has been isolated and confirmed in experimental investigations,[11,16] the mechanism for their formation has not been conclusively determined, due in part to the fast formation.[8,17,18]

Previous QM/MM (AM1/MM) calculations of the TS reaction pathway suggested that the reaction proceeds via the concerted mechanism, with an activation energy barrier of 12.4 kcal/mol.[13] Although semiempirical QM/MM calculations have provided useful mechanistic insight for many enzymes,[19−21] energies calculated with these approximate QM methods can often be significantly in error. Density functional theory based QM/MM methods offer a significant improvement, but DFT methods also have limitations in calculations of energies of transition states and intermediates in some cases.[22,23] Only at high levels of correlated ab intio QM treatment can near-quantitative agreement with experiment be expected,[19] and such calculations provide an important "gold-standard" for QM/MM calculations of potential energy surfaces for enzyme-catalyzed reactions. It is important to test different levels of QM/MM treatment, particularly for enzymes where two (or more) mechanisms may be close in energy, and to test conclusions regarding the nature of the energy surface for reaction.

In the present work, a realistic model of the dUMP/mTHF/TS ternary complex has been prepared and used to study methylene addition to THF with density functional theory and ab initio based QM/MM calculations. We use QM/MM methods which have been applied successfully to several enzymes[19−21] to investigate the mechanism of formation of the covalent intermediate in thymidylate synthase. Clarification of the reaction mechanism may help in the future design of new mechanism-based inhibitors, by answering fundamental questions, such as the following: does the enolate intermediate exist? and should functionalization of dUMP be aimed at enhancing the stability (or electrophilicity) of the postulated enolate?

2. COMPUTATIONAL DETAILS

The model system was constructed from an X-ray crystal structure of TS from *Escherichia coli* containing dUMP and THF (PDB code: 1KZI[24]). The topologies and parameters for ligands were generated by analogy to similar standard amino acids and nucleic acids in the CHARMM27 force field (see the Supporting Information for more details).[25,26] The protonation states of titratable protein residues were calculated at pH 7.0 using the PROPKA3 program.[27,28] All of the missing hydrogen atoms were added using the HBUILD module in the CHARMM software package[29] (version 30b2). The system was solvated using the TIP3P water model of CHARMM[30] and subsequently truncated to a sphere of radius 25 Å, centered on the iminium nitrogen of mTHF. All added water molecules located within 2.8 Å of any heavy atom were removed. A

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

deformable boundary potential, within a sphere of radius 20.75 Å to the iminium nitrogen of mTHF, was applied to water molecules during the MD simulation, to prevent evaporation from the solvent surface. All covalent bonds involving hydrogen atoms were constrained during MD using the SHAKE algorithm.[31] A 13 Å cutoff was used in the calculation of nonbonded interactions in MD. The system was optimized by MM with 2000 steps of steepest descent (SD) and adopted basis Newton−Raphson (ABNR) methods, under the molecular restrained condition on both substrates and water molecules. All water molecules were unrestrained and then equilibrated for 100 ps under stochastic boundary conditions at 310 K while the restraint was still applied to the two substrates. Afterward, the whole system was simulated for 5 ns with the CHARMM27 force field.

The five structures obtained at 1 ns intervals of MD simulation were further minimized without any reaction coordinate constraints using a QM/MM approach. The effect of QM region size was examined using three different QM regions: (i) a minimal QM model, consisting of the Cys146 side chain, the uracil and deoxyribose fragments of dUMP and the pteridine fragment of mTHF (referred to herein as QM1—see Figure 1); (ii) a larger QM region consisting of QM1 and

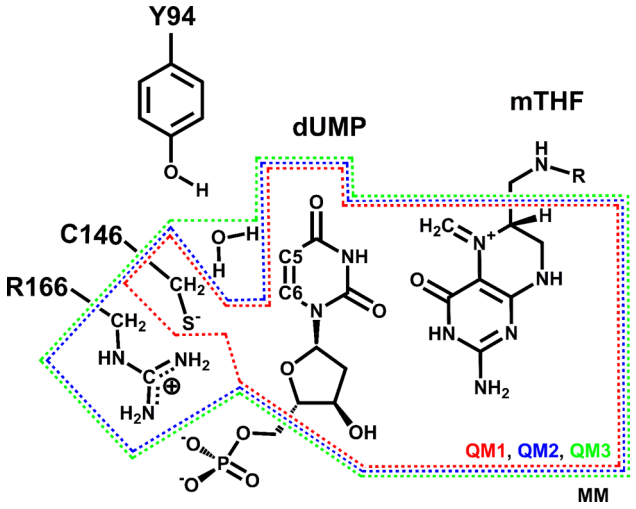


Figure 1. Three different QM regions (QM1, QM2, and QM3) used in QM/MM calculations for investigation of the Michael addition and ternary covalent complex formation reactions.

the Arg166 side chain up to the beta carbon atom (QM2); (iii) QM2 and a water molecule that is observed to interact with Cys146 (QM3). Link atoms[32] (treated as QM hydrogen atoms) were introduced at the QM/MM boundary. Larger QM regions were also tested for comparison in single point QM/MM calculations (see below). QM/MM calculations were performed using the B3LYP[33−35] density functional with the 6-31+G(d) basis set using Jaguar 5.5.[36] An empirical dispersion correction was applied to the B3LYP energy and gradient calculations;[37] inclusion of dispersion has been found to be important in DFT QM and QM/MM calculations on enzyme-catalyzed reactions.[38,39] The rest of the system was treated with the CHARMM27 force field in the Tinker program.[40] QM/MM optimization was performed using the QoMMMa program.[41] Corrections to the QM part were calculated by performing single point energy calculations for the QM region of the B3LYP-D/6-31+G(d)-CHARMM27 optimized geometries using the MP2 and SCS-MP2 methods.[42] The SCS-MP2 method has been found to give results close to CCSD(T) for other enzyme-catalyzed reations.[22] For these calculations, the MOLPRO program was used,[43] with the cc-pVTZ and aug-cc-pVTZ basis sets.[44−46]

As mentioned above, the ternary covalent intermediate formation involves nucleophilic attack of the Cys146 thiolate at the C-6 position of dUMP, and methylene linkage formation between the C-5 atom of dUMP and the reactive methylene of the mTHF cofactor. Based on the reaction coordinate driving approach,[47,48] the potential energy surface associated with these chemical reactions was calculated as a function of two reaction coordinates: RC1 = $d_{(S,Cys146-C6,dUMP)}$ and RC2 = $d_{(C5,dUMP-C5M,mTHF)}$, as shown in Figure 2. Snapshot 1 from the MD simulations was used as the initial structure. Harmonic restraints with a force constant of 3000 kcal/mol·Å^2 were used to drive RC1 and RC2 from 3.5 to 1.6 and 3.4 to 1.4 Å, respectively, with an interval step of 0.1 Å. These calculations yielded a two-dimensional potential energy surface. Additionally, a one-dimensional potential energy surface for the reaction was calculated over all of the five selected snapshots using a reaction coordinate defined by the combination of RC1 and RC2 (RC1 + RC2, ranging from 6.4 to 3.2 Å), in order to calculate a less constrained pathway. Single point QM calculations with higher levels of QM theory (i.e., larger basis sets and SCS-MP2) were performed on the structures along the reaction pathway optimized with the latter method.

3. RESULTS AND DISCUSSION

3.1. Effect of QM Region Size. The optimum relationship between the size of the QM region and the accuracy of results is an important consideration for QM/MM calculations, because while a large QM region leads to high computational expense, misleading results can be obtained if the chosen QM region is too small. From the last snapshot of MD simulation, the thiolate anion of Cys146 is stabilized by Arg166 and a water molecule as shown in Figure 3A (see Figure S4 in Supporting Information for the superposition of all five MD snapshots). Therefore, the effect of including the latter two fragments in the QM region was examined (QM regions QM2 and QM3), as described in the Computational Details section. The QM/MM total energy (the combination of both QM and MM energy parts), together with its separate QM, QM without dispersion and MM components along the reaction pathway with each QM model is shown in Figure 3B.

In the energy profiles calculated with the different QM models, the transition states and products (covalent intermediate) are found at nearly identical values of the reaction coordinate, (the black lines shown in Figure 3B indicate the total QM/MM energy, relative to the substrate complex). In addition, the figure reveals that the dispersion-corrected QM energies (the red lines) are lower in magnitude than the corresponding energies excluding dispersion by ∼5 kcal/mol for all model systems. This information suggests that the effect of the dispersion correction is significant, and its inclusion may therefore provide a more accurate reaction barrier in comparison with semiempirical methods or DFT methods where dispersion is not accounted for.

The MM energy, calculated for the minimal QM model (QM1), decreases by ∼15 kcal/mol as the reaction progresses, whereas it slightly increases (by 5 kcal/mol) for the other two QM models (QM2 and QM3). Prior to the reaction, the terminal hydrogen atom of the guanidinium group of Arg166

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

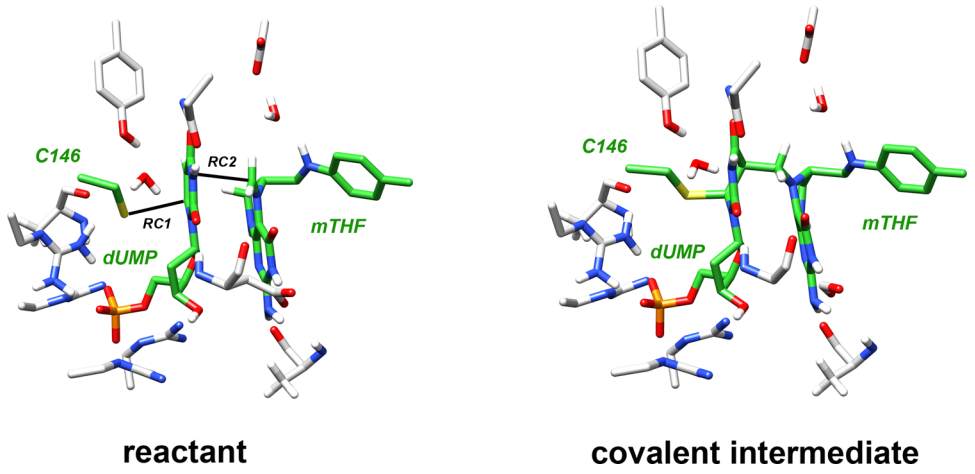


Figure 2. Reaction coordinates used for modeling the Michael addition (RC1) and ternary covalent complex formation (RC2) reactions.

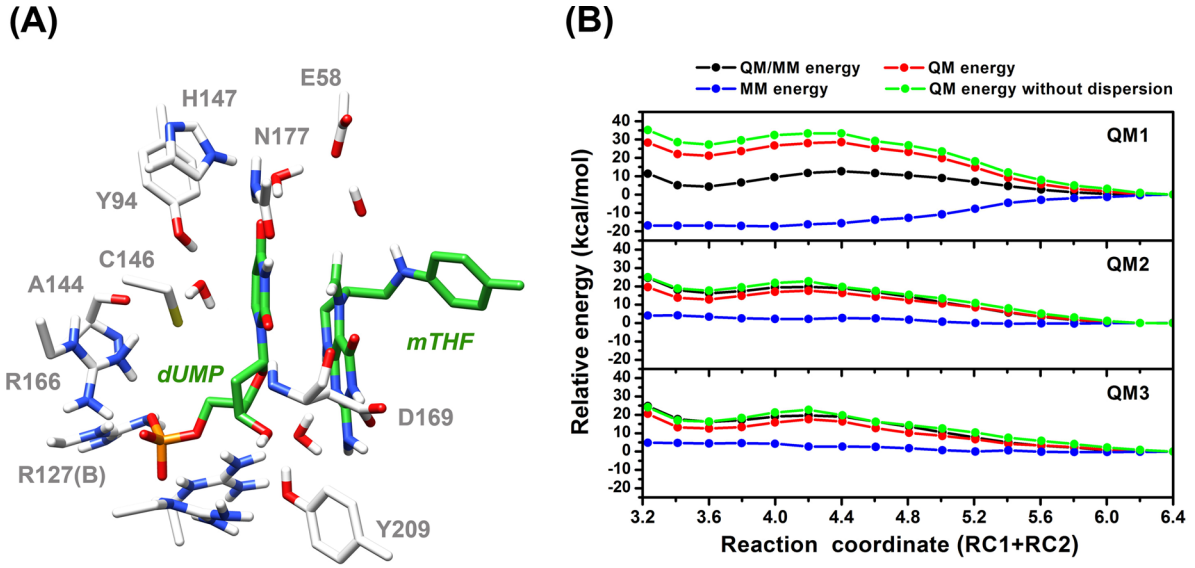


Figure 3. (A) Residues surrounding the thiolate anion of C146. (B) QM/MM (at B3LYP-D/6-31+G(d)-CHARMM27 level), QM, QM without dispersion, and MM energies for the three different QM models (QM1, QM2, and QM3).

forms a strong electrostatic interaction with the thiolate of Cys146, resulting in the former adopting a slightly bent configuration at the usually sp^2 hybridized nitrogen atom. In addition, the bending behavior around this nitrogen is observed, regardless of whether the Arg166 side chain is treated by MM or QM methods (Figure S5 in the Supporting Information). After nucleophilic attack of the thiolate on the dUMP substrate, the guanidinium nitrogen returns to a sp^2 planar conformation, forming a new H-bond to the backbone carbonyl oxygen of Ala144 (see the Structural Analysis section, Figure 7, for details).

In the QM1 model, the Arg166 side chain is treated by MM, while in the QM2 and QM3 models, it is treated by QM. Differences between the results for QM1 and the other systems could be due either to differences between the QM and MM energetics for the internal structural changes of the side chain, or for its interaction with the rest of the system. There is a strong electrostatic interaction between the Arg166 and Cys146 side chains in the reactant complex structure. As mentioned above, the guanidinium nitrogen of Arg166 which directly interacts with the thiolate side chain of Cys146 will return to the sp^2 planar configuration and form a new H-bonding interaction with the carbonyl oxygen of the Ala144 backbone. This leads to a decrease in the MM energy during the reaction when the Arg166 is treated at the MM level in the QM1 model. Conversely, the Arg166-Cys146 and Arg166-Ala141 interactions are treated by QM−QM and QM−MM interactions in QM2 and QM3 models, which is found to reduce the QM energy by ∼10 kcal/mol (the red lines in Figure 3B) and simultaneously increase the MM energy.

To approximately quantify the pronounced effect of Arg166 over the other active site residues in the stabilization of the thiolate anion of Cys146, single-point QM/MM energy calculations were performed with other nearby residues included in the QM1 region (Arg166, Ala144, Arg21, Tyr94, and Arg127(B)). For each of these individual residues, a similar reaction profile was calculated to that of the original QM1 model, suggesting that the stabilization effect of these residues on the thiolate anion of Cys146 is much smaller in magnitude than that of Arg166 (see Figure S1 in the Supporting Information). Although the Arg21 and Arg127(B) residues are placed near to the phosphate moiety of the dUMP substrate, in a similar manner to Arg166, only the Arg166 side chain is included in the QM region because it undergoes a

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

significant conformational change during the reaction, while the other two residues maintain their strong interactions with the phosphate group (see the H-bonding interactions in Figure S3). It is clear from these findings that an MM treatment of Arg166 is insufficient for modeling this enzyme since the observed QM/MM energy for the QM1 model is contaminated by the high change in the MM energy component, which occurs outside of the QM specified region.

Inclusion of the water molecule that stabilizes Cys146 in the QM region (QM3) does not significantly affect the energy profile, as indicated by the very similar energy profiles calculated for the QM2 and QM3 models (Figure 3B). These results indicate that this water molecule, and its interaction with Cys146, are treated similarly by QM and QM/MM, and thus an MM treatment is acceptable. For computational efficiency and accuracy, this suggests that the optimum QM region is QM2 (fragments of the dUMP and mTHF substrates, and the side chains of Cys146 and Arg166). For this reason, all results below correspond to calculations performed using the QM2 model.

3.2. Effect of Basis Set Size. The basis set for a DFT or ab initio calculation is also an important consideration for QM/MM modeling.[22] To test the effect of the basis set on the energy profile for the Michael addition and methylene bridge formation, single point B3LYP-D QM/MM energy calculations were calculated on the B3LYP-D/6-31+G(d)-CHARMM27 optimized geometries (with the QM2 model) using different numbers of diffuse and polarization functions, as well as a triple-ζ basis set (Figure 4).

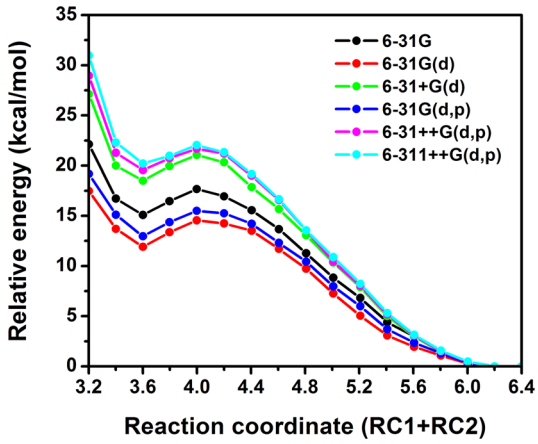


Figure 4. QM/MM energy profiles for the reaction using the combined reaction coordinate approach (RC1 + RC2) using different basis sets.

Although the energy profiles are similar in shape, with stationary points located at the same values of the reaction coordinate, the reaction energy barriers and energies of the methylene bridged intermediate have significantly different values. The inclusion of additional polarization functions (i.e., 6-31G(d) and 6-31G(d,p) versus 6-31G) for the QM model lowers the energies of the transition state and methylene bridged intermediate by ~3−4 kcal/mol, relative to the substrate complex. However, the combined inclusion of polarization and diffuse functions (6-31+G(d) and 6-31++G(d,p)) raises the energies of the transition state and covalent intermediate by ~5 kcal/mol, relative to the values calculated with 6-31G. Polarization and diffuse functions have a significant effect on the QM energies and should therefore be included in calculations. The B3LYP energy profiles calculated with the 6-31+G(d), 6-31++G(d,p), and 6-311++G(d,p) basis sets, on the other hand, were all very similar (Figure 4); therefore, the 6-31+G(d) basis set represents the optimum choice for geometry optimization in terms of accuracy and computational expense.

3.3. Potential Energy Surface for Reaction. To investigate whether the formation of the ternary covalent intermediate in TS occurs in a synchronous manner, a two-dimensional potential energy surface (2D-PES) was calculated. The 2D-PES for the ternary covalent intermediate formation is plotted relative to the reactant complex in Figure 5A. As seen from the 2D-PES, two local energy minima corresponding to the (I) reactant and (III) intermediate were detected with a single approximate transition state (II) along the minimum energy pathway. The energy barrier is 20.8 kcal/mol (Figure 5B). The minimum energy pathway crossing the middle of the 2D-PES suggests that the Michael addition and covalent complex formation steps occur in a synchronous manner.

3.4. Conformational Sampling and ab initio Energy Correction. The effects of conformational variability are another important factor to consider in the QM/MM modeling of enzyme-catalyzed reactions.[22] The calculated barrier may vary significantly, even with small changes in starting structure. In order to sample different conformations of the enzyme, five structures (taken at 1 ns intervals) were selected from the 5 ns unrestrained molecular dynamics simulation. These structures were used to model the Michael addition and methylene bridge formation reactions using the combined reaction coordinate (RC1+RC2, described in the Computational Details section) with the B3LYP-D/6-31+G(d)-CHARMM27 method.

The average activation energy barrier for the formation of the methylene bridged intermediate, calculated from the five different starting structures, is 20.7 kcal/mol. The highest and lowest energy barriers were calculated to be 21.8 and 20.1 kcal/mol, respectively. The Boltzmann-weighted average, which places additional weighting on the lowest energy barriers and is more representative of the true enzyme reactivity, is 20.5 kcal/mol. The similar values obtained using the two different averaging methods reflects a small degree of variation between the barriers, as is confirmed by the small value of the standard deviation (~0.7 kcal/mol, see Table 1).

To confirm that the five calculated pathways are consistent with the expected chemical reaction mechanism, all five reactant and transition state structures (obtained from the combined reaction coordinate profiles) were superimposed on top of the QM/MM optimized structures calculated from the last snapshot of the MD simulation (Figure S4). All five of the optimized reactant structures display similar conformations of the six membered rings of the dUMP and mTHF substrates. The thiolate side chain of Cys146 and methylene group of mTHF move closer to the dUMP molecule as the reaction proceeds. Noticeably, the transition state geometries are very similar in structure. This is consistent with the very similar barriers found for the different starting structures.

Ab initio approaches, such as spin-component scaled Møller−Plesset perturbation theory (SCS-MP2) provide a more accurate treatment of electron correlation and therefore provide greater accuracy than B3LYP calculations.[22] Single point energies were calculated for the QM region of the B3LYP-D/6-31+G(d)-CHARMM27 optimized structures along the reaction coordinate using MP2 and SCS-MP2 with the aug-cc-pVTZ basis set. The results obtained from

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

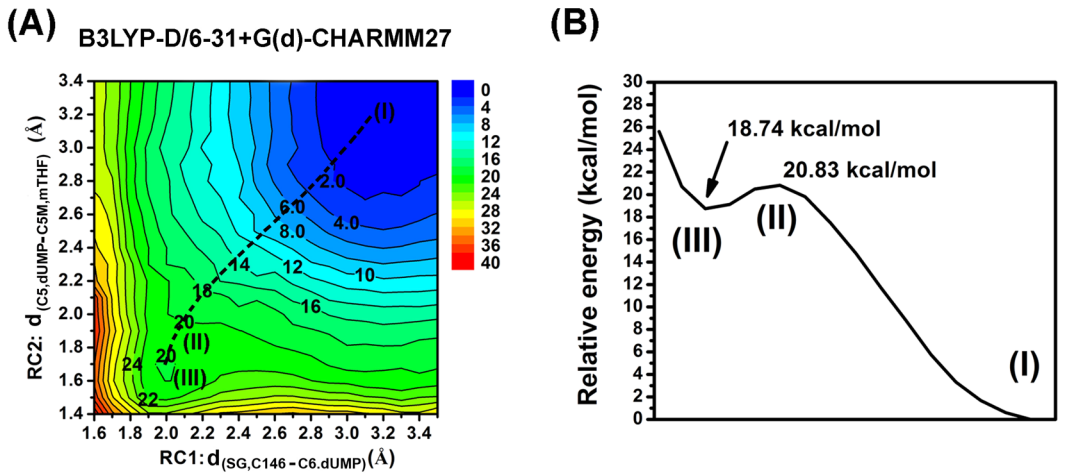


Figure 5. (A) Two-dimensional QM/MM (B3LYP-D/6-31+G(d)-CHARMM27) potential energy surface for Michael addition and covalent complex formation. (B) Simplified overall energy profile along the reaction pathway (the dash line of part A). In both parts, (I), (II), and (III) correspond to the reactant, transition state, and intermediate, respectively, whose structures are depicted in Figure 7.

Table 1. Energy Barriers (kcal/mol) for the Michael Addition and Formation of the Methylene Bridged Intermediate Obtained from the B3LYP-D/6-31+G(d), MP2/aug-cc-pVTZ, and SCS-MP2/aug-cc-pVTZ Methods, Together with the Experimental Value[a]

	activation energy						
calculation methods	snap1[b]	snap2	snap3	snap4	snap5	average[c]	SD
B3LYP-D/6-31+G(d)	21.8	20.3	20.1	20.8	20.6	20.7	0.67
MP2/aug-cc-pVTZ	18.2	16.0	15.6	17.2	16.8	16.7	1.03
SCS-MP2/aug-cc-pVTZ	21.1	19.4	18.5	19.2	18.7	19.4	1.04
multi-AM1/MM[13]						12.4	
experiment						18.3	

[a] $\Delta G^{\ddagger}$ = 16.2 kcal/mol, $\Delta E^{\ddagger}$ = 18.3 kcal/mol using the same value of entropic contribution as the FdUMP/mTHF complex for subtraction.[17] [b] The conformational details of reactant, TS, and intermediate structures are presented in Figure 7. [c] The averaged value is calculated through the arithmetic method.

performing conformational sampling and calculating MP2 and SCS-MP2 energies are shown in Figure 6 and Table 1.

The MP2 and SCS-MP2 single point energy profiles differ significantly from the B3LYP QM/MM energy profiles. At the transition state, the QM/MM energy calculated with B3LYP suggests that the position of highest activation energy corresponds to a reaction coordinate value of 4.0 Å when using the combined reaction coordinate approach, whereas it is slightly increased to 4.2 Å when using the MP2 and SCS-MP2 methods. In addition, the average activation energy of reaction from the B3LYP-D/6-31+G(d) profiles is 20.7 kcal/mol whereas the MP2 and SCS-MP2 values are 16.7 and 19.4 kcal/mol, respectively.

The experimental activation free energy, $\Delta G^{\ddagger}$, to transform the FdUMP/mTHF noncovalent to covalent intermediate, converted from the value of k_{cat} (0.586 s^{-1} at 25 °C) using transition state theory, is approximately 17.8 kcal/mol, while the activation energy, $\Delta E^{\ddagger}$, and activation enthalpy, $\Delta H^{\ddagger}$, are reported to be 19.9 and 19.3 kcal/mol, respectively.[17] In computational investigations, the $\Delta E^{\ddagger}$ parameter is typically used since it is directly calculated from the total energy difference between any two states, while the calculation of $\Delta H^{\ddagger}$

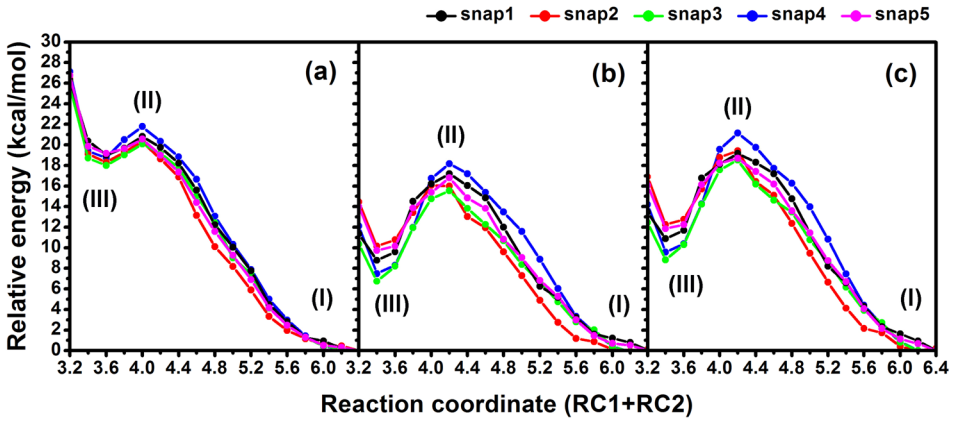


Figure 6. Reaction energy profiles for concerted Michael addition and methylene bridging formation calculated at the (a) B3LYP-D/6-31+G(d)-CHARMM27, (b) MP2/aug-cc-pVTZ-CHARMM27, and (c) SCS-MP2/aug-cc-pVTZ-CHARMM27 levels. All geometries were optimized at the B3LYP-D/6-31+G(d)-CHARMM27 level. The reactant, transition state, and covalent intermediate are labeled from right to left as (I), (II), and (III), respectively.

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

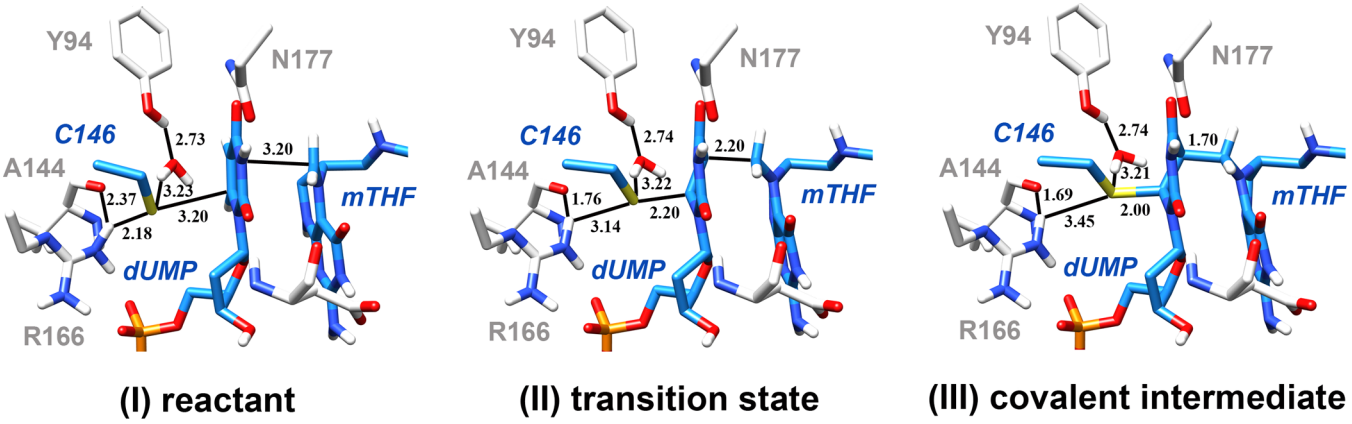


Figure 7. QM/MM (B3LYP-D/6-31+G(d)-CHARMM27) optimized structures of the (I) reactant, (II) transition state, and (III) covalent intermediate in the active site of thymidylate synthase obtained from the combined reaction coordinate profile for snapshot 1.

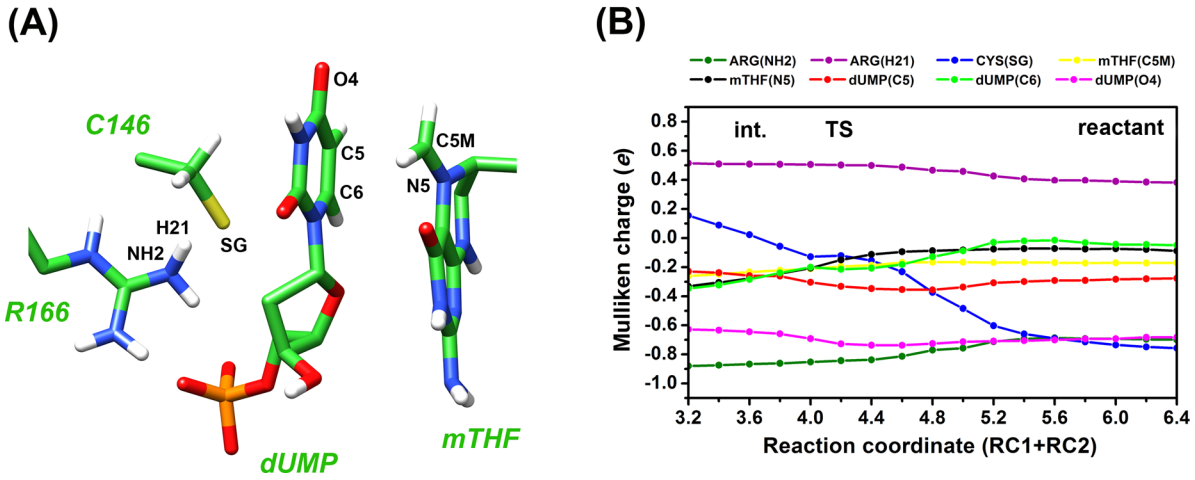


Figure 8. (A) Atomic labels over the reactant structure of combined reaction coordinate snapshot 1 and (B) Mulliken charges from B3LYP-D/6-31+G(d)-CHARMM27 optimized structures averaged over five reaction profiles.

is more complicated as it requires the calculation of the total energy, zero-point energy and enthalpic temperature effect.[22] The entropic contribution to ternary covalent formation in the FdUMP/mTHF complex is approximately −2.1 (17.8−19.9) kcal/mol. The activation free energy of this reaction in the dUMP/mTHF complex is reported to be 16.2 kcal/mol, while the activation energy of reaction is not determined.[17] The entropic contribution to the reaction of dUMP can be assumed to be of similar magnitude to that of the substrate analogue. This assumption allows us to approximate the $\Delta E^{\ddagger}$ value of the dUMP/mTHF system to be 18.3 kcal/mol (16.2 − (−2.1)), consequently.

Herein, the MP2 result underestimates the average activation energy barrier by 1.6 kcal/mol, while the SCS-MP2 and B3LYP-D overestimates the barrier by 1.1 and 2.4 kcal/mol, respectively. In a previous publication, the activation barrier of this step was observed as 12.4 kcal/mol (underestimated by 5.9 kcal/mol from experiment) by a multi-level AM1/MM study.[13] We find here that the use of higher-level QM methods provides significantly less deviation of the observed activation energy barrier from the experimental value. Although the value for the B3LYP-D/6-31+G(d) average reaction energy barrier is higher than the MP2 or SCS-MP2 methods, it is quite close to the SCS-MP2 value, indicating that the B3LYP functional performs well for this reaction. This result also reinforces that it should not generally be assumed that barriers calculated with density functional theory are too low. The main success of this method is the ability to search for the minimized geometries that correspond to the chemical reaction pathway.

3.5. Structural Analysis. The geometries of the reactant, transition state and intermediate extracted from the snapshot 1 of the combined reaction coordinate profile are depicted in Figure 7. It can be seen that the hybridization of the dUMP-C6 and C5 atoms gradually changes from sp^2 to sp^3, as expected during the reaction. The two reaction coordinate distances decrease (RC1:3.20 to 2.00 Å and RC2:3.20 to 1.70 Å). The hydrogen bonding interaction between the side chain of Arg166 and the thiolate sulfur of Cys146 is broken concomitantly with the Michael addition. However, this is compensated for by a new interaction formed between this guanidinium hydrogen and the backbone carbonyl oxygen of Ala144, $d_{(HH21,Arg166-O,Ala144)}$ decreases from 2.37 Å (reactant) to 1.76 Å (TS) and 1.69 Å (intermediate). In addition, the dramatic decrease of the $d_{(HH21,Arg166-O,Ala144)}$ distance between the reactant and transition state suggests that the H-bonding between the Ala144 and Arg166 residues is quickly formed after the breaking of the Cys146-Arg166 interaction and before the occurrence of the transition state.

Interestingly, the water molecules surrounding the Cys146 thiolate not only provide electrostatic stabilization but also

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

retain the strong hydrogen bonding interactions with the side chain of Cys146 and the Tyr94 hydroxyl group, as indicated by the consistent values of ∼1.86 and ∼2.7 Å for $d_{(SG,Cys146-HO,WAT)}$ and $d_{(O,WAT-HH,Tyr94)}$, respectively between reactant and intermediate. In addition, the number of H-bonding interactions during the reaction was also investigated (see Figure S3 in the Supporting Information). The H-bonding interactions between the two substrates and the surrounding residues are mostly observed in the same pattern and magnitude in the reactant, transition state, and covalent intermediate states. Because the conformations of the other surrounding residues and bridging waters do not change significantly during the two reactions, the PES reasonably represents the expected chemical reaction without any interference from other intrinsic changes.

3.6. Mulliken Charge Analysis. The Mulliken atomic charges, averaged over the five reaction profiles, were calculated for all points along the reaction pathway and are shown in Figure 8. While Mulliken charges have well-known limitations, they can give a useful indication of changes in electronic distribution during reactions.[49] This analysis provides information about the movements of charge during the reaction. These results are summarized in Table 2.

Table 2. Selected Atomic Charges (e^-) Obtained from Mulliken Charge Analysis for the Reactant, Transition State, and Intermediate of the Concerted Michael Addition and Methylene Bridge Formation Reactions at B3LYP-D/6-31+G(d) Level

	Mulliken charge		
relevant atom	reactant	TS	intermediate
Cys-SG	−0.75	−0.12	0.08
dUMP-C6	−0.04	−0.21	−0.32
dUMP-C5	−0.27	−0.31	−0.26
dUMP-O4	−0.67	−0.72	−0.65
mTHF-C5M	−0.16	−0.19	−0.24
mTHF-N5	−0.08	−0.15	−0.30

The Mulliken charge analysis indicates that the electronic properties of the reactive cysteine (Cys146) and two substrates (dUMP and mTHF) change during the reaction, as expected. As shown in Figure 8B, the negative charge on the Cys-SG atom (blue line) decreases, while the negative charge on the dUMP-C6 (light green) and mTHF-N5 (black) atoms increases, as the reaction proceeds. The observed changes in Mulliken charge over the Cys-SG, dUMP-C6, and mTHF-N5 atoms indicates the delocalization of electron density from the thiolate of Cys146 to the iminium of mTHF which is in accord with the proposed reaction scheme. The atomic charges of the dUMP-C5 (red) and dUMP-O4 (magenta) atoms remain approximately constant (around −0.27 and −0.67 e^-, respectively) indicating that the proposed enolate intermediate is undetectable during the reaction. This information agrees well with the single transition state observed in the 2D-PES, which clearly indicates that the Michael addition and covalent complex formation reactions between dUMP and mTHF occur through a concerted pathway, rather than by a stepwise mechanism.

4. CONCLUSIONS

Understanding the mechanism of the reaction catalyzed by thymidylate synthase is important from the point of view of fundamental biochemistry and also potentially for design of inhibitors against this cancer target enzyme. A particular question has been whether formation of the ternary covalent intermediate (between dUMP and mTHF), involving Michael addition, occurs via a concerted mechanism, or via a stable enolate. The results indicate that the reaction is concerted. Experimental information, such as (i) simultaneous occurrence of several steps (imidazolidine activation, Michael addition, and covalent bonding formation[8,17]) during the mTHF cofactor binding and (ii) the inseparable kinetic rate of Michael addition from the ternary covalent complex formation, also agrees with and supports our concerted mechanism finding indirectly. Together, these findings are potentially useful for mechanism-based drug design, because future attention should be focused on increasing Michael addition reactivity, rather than on the stabilization of the enolate.

Comparison of DFT and ab initio results shows that B3LYP provides a good description of the reaction, though it overestimates the barrier somewhat compared to the ab initio results. Inclusion of both polarization and diffuse functions was found to be important for QM/MM energy profiles. The 6-31+G(d) basis set was selected for B3LYP QM/MM geometry optimizations, as it represents an efficient trade-off between accuracy and computational expense. The average activation energy barriers, calculated from profiles from five different starting structures, obtained using the B3LYP-D/6-31+G(d), MP2/aug-cc-pVTZ, and SCS-MP2/aug-cc-pVTZ QM/MM methods, are 20.7, 16.7, and 19.4 kcal/mol, respectively. These are in good agreement with the value derived from experiment ($\Delta E^{\ddagger}$ = 18.3 kcal/mol, $\Delta G^{\ddagger}$ = 16.2 kcal/mol). The MP2 result therefore underestimates the activation energy barrier by 1.6 kcal/mol, while the SCS-MP2 overestimates the barrier by a smaller amount (1.1 kcal/mol).

The results also emphasize the need to test the size of the QM region in QM/MM calculations. The size of the QM region has been shown to have a significant effect on the calculated activation energies here. The QM2 model (containing the uracil and deoxyribose fragments of dUMP, pteridine of mTHF, Cys146 and Arg166 side chains) was selected, as it yielded results that were similar to those of the larger QM3 region. The QM1 region was found to be too small, as it was discovered that an MM representation of Arg166 gives inaccurate results for modeling this reaction, due to an unrealistic large change in MM energy during the reaction. This effect should be considered in QM/MM modeling of reactions involving arginine as a stabilizing residue and may indicate a need to examine MM parameters for arginine generally.

■ ASSOCIATED CONTENT

Supporting Information

QM/MM single point energies (and components) calculated for the reaction initiated from snapshot 1, with additional single environmental residues added to the QM1 model (as indicated). Chemical structures of mTHF and dUMP molecules with the atomic labeling scheme used in this work. Configuration of dUMP and mTHF in the TS catalytic pocket. Solid lines represent the hydrogen bonding interactions and H-bonding distances between donor−acceptor heavy atoms in

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

reactant, transition state, and ternary covalent intermediate states. Superposition of the QM/MM minimized structures of the five MD snapshots (in comparison with the final snapshot of the MD simulation, gray in the models) at the reactant and transition state. Superposition of structures between the final snapshot of the MD simulation (gray) and the five reactant structures (after QM/MM minimization). The structure of the guanidinium group of R166 is distorted from a planar sp^2 geometry. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: supot.h@chula.ac.th (S.H.).

*E-mail: adrian.mulholland@bristol.ac.uk (A.J.M.).

Present Address

‡R.L.: Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany, and Department of Chemistry, Philipps-Universität Marburg, Hans-Meerwein-Strasse, 35032 Marburg, Germany.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the overseas research grants from Asahi Glass Foundation and National Research University project, Office of Higher Education Commission (WCU-004-HR-57). N.K. thanks the Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University. T.R. thanks the Ratchadaphiseksomphot Endowment Fund for a grant for the development of new faculty staff. A.J.M. thanks EPSRC for a Leadership Fellowship (grant number: EP/G007705/01) and (together with R.L.) thanks EPSRC for support. A.J.M. and R.L. also thank EPSRC for support under the CCP-BioSim project (grant number EP/J010588/1, www.ccpbiosim.ac.uk).

REFERENCES

(1) Ahmad, S. I.; Kirk, S. H.; Eisenstark, A. Thymine metabolism and thymineless death in prokaryotes and eukaryotes. *Annu. Rev. Microbiol.* **1998**, *52*, 591−625.

(2) Seno, T.; D, A.; Shimizu, K.; Koyama, H.; Takeishi, K.; Hori, T. Thymineless death and genetic events in mammalian cells. *Basic Life Sci.* **1985**, *31*, 241−63.

(3) Rolfe, R. On the mechanism of thymineless death in Bacillus subtilis. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 114−121.

(4) Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Montfort, W. R.; Maley, G. F.; Maley, F.; Stroud, R. M. Plastic adaptation toward mutations in proteins: Structural comparison of thymidylate synthases. *Proteins* **1990**, *8*, 315−333.

(5) Carreras, C. W.; Santi, D. V. The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.* **1995**, *64*, 721−762.

(6) Danenberg, P. V. Thymidylate synthetase - a target enzyme in cancer chemotherapy. *Biochim. Biophys. Acta* **1977**, *473*, 73−92.

(7) Santi, D. V. Perspective on the design and biochemical pharmacology of inhibitors of thymidylate synthetase. *J. Med. Chem.* **1980**, *23*, 103−111.

(8) Spencer, H. T.; Villafranca, J. E.; Appleman, J. R. Kinetic scheme for thymidylate synthase from Escherichia coli: determination from measurements of ligand binding, primary and secondary isotope effects, and pre-steady-state catalysis. *Biochemistry* **1997**, *36*, 4212−4222.

(9) Stroud, R. M.; Finer-Moore, J. S. Stereochemistry of a multistep/bipartite methyl transfer reaction: thymidylate synthase. *FASEB J.* **1993**, *7*, 671−677.

(10) Rode, W.; Leś, A. Molecular mechanism of thymidylate synthase-catalyzed reaction and interaction of the enzyme with 2- and/or 4-substituted analogues of dUMP and 5-fluoro-dUMP. *Acta Biochim. Polym.* **1996**, *43*, 133−142.

(11) Barrett, J. E.; Maltby, D. A.; Santi, D. V.; Schultz, P. G. Trapping of the C5Methylene Intermediate in Thymidylate Synthase. *J. Am. Chem. Soc.* **1998**, *120*, 449−450.

(12) Phan, J.; Mahdavian, E.; Nivens, M. C.; Minor, W.; Berger, S.; Spencer, H. T.; Dunlap, R. B.; Lebioda, L. Catalytic Cysteine of Thymidylate Synthase Is Activated upon Substrate Binding. *Biochemistry* **2000**, *39*, 6969−6978.

(13) Kanaan, N.; Martí, S.; Moliner, V.; Kohen, A. A quantum mechanics/molecular mechanics study of the catalytic mechanism of the thymidylate synthase. *Biochemistry* **2007**, *46*, 3704−3713.

(14) Longley, D. B.; Harkin, D. P.; Johnston, P. G. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **2003**, *3*, 330−338.

(15) Bijnsdorp, I. V.; Comijn, E. M.; Padron, J. M.; Gmeiner, W. H.; Peters, G. J. Mechanisms of action of FdUMP[10]: Metabolite activation and thymidylate synthase inhibition. *Oncol. Rep.* **2007**, *18*, 287−291.

(16) Moore, M. A.; Ahmed, F.; Dunlap, R. B. Trapping and partial characterization of an adduct postulated to be the covalent catalytic ternary complex of thymidylate synthase. *Biochemistry* **1986**, *25*, 3311−3317.

(17) Santi, D. V.; McHenry, C. S.; Raines, R. T.; Ivanetic, K. M. Kinetics and thermodynamics of the interaction of 5-fluoro-2′-deoxyuridylate with thymidylate synthase. *Biochemistry* **1987**, *26*, 8606−8613.

(18) Finer-Moore, J. S.; Santi, D. V.; Stroud, R. M. Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. *Biochemistry* **2003**, *42*, 248−256.

(19) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schutz, M.; Thiel, S.; Thiel, W.; Werner, H. J. High-accuracy computation of reaction barriers in enzymes. *Angew. Chem., Int. Ed.* **2006**, *45*, 6856−9.

(20) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. High-level QM/MM modelling predicts an arginine as the acid in the condensation reaction catalysed by citrate synthase. *Chem. Commun.* **2008**, *16*, 1874−1876.

(21) Lonsdale, R.; Hoyle, S.; Grey, D. T.; Ridder, L.; Mulholland, A. J. Determinants of Reactivity and Selectivity in Soluble Epoxide Hydrolase from QM/MM Modeling. *Biochemistry* **2012**, *51*, 1774−1786.

(22) van der Kamp, M. W.; Zurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Testing high-level QM/MM methods for modeling enzyme reactions: acetyl-CoA deprotonation in citrate synthase. *J. Phys. Chem. B* **2010**, *114*, 11303−11314.

(23) Lawan, N.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J. Comparison of DFT and ab initio QM/MM methods for modelling reaction in chorismate synthase. *Chem. Phys. Lett.* **2014**, *608*, 380−385.

(24) Fritz, T. A.; Liu, L.; Finer-Moore, J. S.; Stroud, R. M. Tryptophan 80 and leucine 143 are critical for the hydride transfer step of thymidylate synthase by controlling active site access. *Biochemistry* **2002**, *41*, 7021−7029.

(25) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(26) Foloppe, N.; MacKerell, A. D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86−104.

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722

(27) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.
(28) Olsson, M. H. M.; Søndergard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(29) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(30) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(32) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.
(33) Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti conelation energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(34) Becke, A. D. Density functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(35) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(36) *Jaguar*, 5.5 ed.; Schrödinger, LLC: Portland, OR, 2003.
(37) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27* (15), 1787−1799.
(38) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Inclusion of Dispersion Effects Significantly Improves Accuracy of Calculated Reaction Barriers for Cytochrome P450 Catalyzed Reactions. *J. Phys. Chem. Lett.* **2010**, *1*, 3232−3237.
(39) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *J. Chem. Theory Comput.* **2012**, *8*, 4637−4645.
(40) Ponder, J. W. *TINKER: Software Tools for Molecular Design*, version 4.1; Washington University School of Medicine: St. Louis., 2003.
(41) Harvey, J. N. Spin-forbidden CO ligand recombination in myoglobin. *Faraday Discuss.* **2004**, *127*, 165−177.
(42) Gerenkamp, M.; Grimme, S. Spin-component scaled second-order Møller-Plesset perturbation theory for the calculation of molecular geometries and harmonic vibrational frequencies. *Chem. Phys. Lett.* **2004**, *392*, 229−235.
(43) Werner, H.-J., et al. *MOLPRO*, version 2006.1; University College Cardiff Consultants Ltd.: Cardiff, U.K., 2006.
(44) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(45) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(46) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358−1371.
(47) Woodcock, H. L.; Hodoscek, M.; Brooks, B. R. Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms. *J. Phys. Chem. A* **2007**, *111*, 5720−5728.
(48) Crehuet, R.; Field, M. J. A transition path sampling study of the reaction catalyzed by the enzyme chorismate mutase. *J. Phys. Chem. B* **2007**, *111*, 5708−5718.
(49) Ridder, L.; Mulholland, A. J.; Rietjens, I. M.; Vervoort, J. Combined quantum mechanical and molecular mechanical reaction pathway calculation for aromatic hydroxylation by p-hydroxybenzoate-3-hydroxylase. *J. Mol. Graphics Modell.* **1999**, *17*, 163−175.

DOI: 10.1021/ct5005033
J. Chem. Theory Comput. 2015, 11, 713−722