

Published in final edited form as:

Arch Biochem Biophys. 2009 December ; 492(1-2): 82–92. doi:10.1016/j.abb.2009.08.010.

An Investigation of the Catalytic Mechanism of S-adenosylmethionine Synthetase by QM/MM Calculations†

George D. Markham^{‡,*}, Fusao Takusagawa[§], Anthony M. DiJulio[‡], and Charles W. Bock^{‡,||}

[‡] Institute for Cancer Research, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111

[§] Department of Molecular Biosciences, University of Kansas, 1200 Sunnyside Avenue, Lawrence, KS 66045-7534

^{||} Department of Chemistry and Biochemistry, School of Science and Health, Philadelphia University, School House Lane and Henry Avenue, Philadelphia, PA 19144

Abstract

Catalysis by S-adenosylmethionine synthetase has been investigated by quantum mechanical/molecular mechanical calculations, exploiting structures of the active crystalline enzyme. The transition state energy of +19.1 kcal/mol computed for a nucleophilic attack of the methionyl sulfur on carbon-5' of the nucleotide was indistinguishable from the experimental (solution) value when the QM residues were an uncharged histidine that hydrogen bonds to the leaving oxygen-5' and an aspartate that chelates a Mg²⁺ ion, and was similar (+18.8 kcal/mol) when the QM region also included the active site arginine and lysines. The computed energy difference between reactant and product was also consistent with their equimolar abundance in co-crystals. The calculated geometrical changes support catalysis of a S_N2 reaction through hydrogen bonding of the liberated oxygen-5' to the histidine, charge neutralization by the 2 Mg²⁺ ions, and stabilization of the product sulfonium cation through a close, non-bonded, contact between the sulfur and the ribose 4'-oxygen.

Keywords

S-adenosylmethionine; methionine adenosyltransferase; enzyme mechanism; substrate-assisted-catalysis; QM/MM

Introduction

S-adenosylmethionine, AdoMet, and its metabolites play a vast number of roles in cellular life [1]. AdoMet is one of the few sulfonium ions found in nature, and the cationic center endows it with a chemical versatility matched by few other biological entities, enabling it to act as an alkylating agent and free radical precursor, as well as a regulatory agent [2-4]. Methyl transfer from AdoMet is perhaps its most widely recognized role, participating in intermediary metabolism and in the modification of nucleic acids and proteins. DNA methylation forms a

© 2009 Elsevier Inc. All rights reserved.

* Address correspondence to George D. Markham, Ph.D., Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111. Tel: 215-728-2439. Fax: 215-728-2412; gd_markham@fccc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

basis for the burgeoning field of epigenetics [5], while aberrant DNA methylation is common in cancers wherein errors are associated with alterations in DNA replication and transcription [6,7]. In a different family of pathways, decarboxylation of AdoMet followed by transfer of the propylamine moiety leads to the polyamines spermine and spermidine which are utilized in the regulation of cell proliferation [8,9]. In a distinct role, increasing numbers of AdoMet dependent enzymatic reactions are being recognized as having 5'-deoxyadenosyl free radicals as transient intermediates formed by homolytic C5'-S bond cleavage [4,10].

The only known biosynthetic route to S-adenosylmethionine is catalyzed by S-adenosylmethionine synthetase, (ATP:L-methionine S-adenosyltransferase, often abbreviated as MAT) [11,12]. The two-step reaction catalyzed by MAT has a number of features that are unique in biology, as it encompasses displacement of the entire tripolyphosphate chain from ATP by the sulfur of methionine, followed by hydrolysis of the resultant tripolyphosphate (PPP_i) moiety to PP_i and P_i before product release; P_i originates from the γ -phosphoryl group of ATP and incorporates an oxygen atom from a water molecule [12,13]. Thus the enzyme has a bifunctional active site that catalyzes both AdoMet formation and PPP_i hydrolysis, the latter step being required to remove a kinetic and thermodynamic trap that arises due to the high affinity of PPP_i for the enzyme [12,14]. MAT sequences from eucarya and bacteria are highly conserved, with typically greater than 80% sequence homology, and the polar active site residues are retained in all of the hundreds of known sequences [15]. MATs exist as dimers or tetramers in nature, and vary substantially in kinetic behavior (*e.g.* k_{cat} and K_m values, cooperativity) [11,15]. MATs from archaea have distinct sequences that are highly conserved within that kingdom, and representatives of the archaeal class are found in a few bacteria [15,16].

The MAT from *Escherichia coli* (denoted cMAT) is the best characterized family member in terms of catalytic mechanism [13,14,17-22]. The crystal structure of cMAT provided the first insight into the architecture of a MAT, and structures have been reported for the ligand free (apo) enzyme, as well as for several complexes [17,23,24]. cMAT is composed of 383 residue subunits and is typically found as a tetramer [13]. The four active sites are located in ~ 30 Å deep cavities between subunits and have contributions from residues from two subunits; cMAT does not show any cooperativity in kinetics [13]. Some mutants of cMAT are also active as dimers, supporting this state as the minimal functional unit (20). Additional crystal structures have shown that rat liver MAT and the human non-hepatic MAT have the same topology as cMAT [25,26], with an rmsd for the main chain carbons of less than 1.3 Å between any two of the structures. AdoMet formation is postulated to occur in a single chemical step *via* a direct S_N2 attack of the methionine on C5' of ATP, based on the observed inversion of stereochemical configuration at C5' during the reaction [27] and the magnitudes of primary and secondary kinetic isotope effects [19]. The free energy profile for the steps in the conversion of enzyme-bound substrates to products shows that in the active site AdoMet formation is energetically favorable while the subsequent PPP_i hydrolysis step has an equilibrium constant near unity [14,22]. MATs utilize two divalent cations (M²⁺) per subunit for both catalytic activities, and certain monovalent cations stimulate these reaction rates by nearly three orders of magnitude [13]; *in vivo*, these cations are presumably Mg²⁺ and K⁺. Crystallographic and spectroscopic data show that one of the divalent metal ions binds to all 3 phosphoryl groups while the second is ligated to the α and γ phosphoryl groups [18,24]. The monovalent cation activator also binds at the active site and appears to organize the active site structure rather than directly partake in catalysis [24]. Coordination of two divalent metal ions to the α phosphoryl group is anticipated to facilitate C5'-O5' bond cleavage by polarizing the C5—O5' bond in the reactant and neutralizing the negative charge that develops during AdoMet formation; however there are no data to support this notion. Furthermore the roles of the protein itself in the catalysis of AdoMet formation remain elusive, and the results of side-direct mutagenesis studies are largely ambiguous [20-22,28-30].

Crystal structures of the cMAT have provided a foundation for understanding the means by which the surrounding protein facilitates the formation of AdoMet. Most notably are the structures of the catalytically active crystalline protein in complexes formed by incubation of the protein crystals with methionine, the alternate substrate adenylylimidodiphosphate (AMPPNP), and the activators Mg^{2+} and K^+ [24]. Utilization of AMPPNP stops the reaction sequence after formation of AdoMet and PPNP, the non-hydrolyzable analog of the normal tripolyphosphate intermediate [13]. These crystals contained approximately equal abundances of the reactant complex, $E \cdot 2Mg^{2+} \cdot AMPPNP \cdot methionine \cdot K^+$ and the intermediate analog complex, $E \cdot 2Mg^{2+} \cdot AdoMet \cdot PPNP \cdot K^+$ [24]. These crystal structures provided the starting points for our computational studies which were directed toward understanding the various molecular interactions involved in catalysis.

Computer simulations of enzymatic reactions are capable of providing tests of postulated reaction mechanisms, as well as insight into the catalytic contributions of individual residues and cofactors at a resolution currently unattainable by experimental methods [31-37]. This is an important convergence for MAT; despite decades of effort, cMAT crystals have had limiting resolution of 2.5 Å [17,23,24] while the crystal structures of rat liver MAT (for which some mechanistic data are also available) have been reported at resolutions in the range of 2.5 – 3.0 Å (cf. [15,26]). Recently the 1.2 Å resolution structure of the non-hepatic human MAT, for which there have been no mechanistic studies, has been deposited by a structural genomics project (pdb file 2P02) [25]. Comparison of the highest resolution structures of each of these three MATs reveals an rmsd for α carbons of < 1 Å, and < 1.4 Å for all atoms in conserved residues, indicating that the cMAT structure provides a sound framework for further mechanistic studies despite the modest resolution of the structural model. Furthermore the initial steps in computational investigations include addition of the appropriate hydrogens (which are nearly X-ray transparent) and geometry optimization, which can mitigate some experimental uncertainty in atomic positions.

The combined quantum mechanics (QM)/molecular mechanics (MM) methodology (QM/MM) enables the rigor of QM methods to be applied to the reactive center while the remainder of the system is described by computationally more economical MM methods. The MM region exerts its influence *via* the combination of an anisotropic electrostatic environment and a framework within which the reaction site is constrained. The use of density functional theory (DFT) in the QM region allows some effects of electron correlation to be incorporated into the calculations while retaining sufficient computational efficiency to allow application to larger QM regions; the 229 atom QM region in our calculations of the Large_QM system is, to our knowledge, among the largest active site models that have been studied by DFT using these methods. The nearly identical crystal structures of the reactant and product complexes of cMAT show that large-scale protein conformational alterations need not be included in these calculations, reinforcing the suitability of a QM/MM approach utilizing a single protein conformation [24]. In the present study, QM/MM calculations of the MAT reaction in the architecture of the cMAT crystal structure have been used to investigate the molecular interactions that catalyze AdoMet formation and to evaluate whether an S_N2 reaction with the experimentally determined rate is consistent with the available crystal structures

Computational Methods

The protein structural model used as the starting point for the QM/MM calculations was derived from the experimental 2.5 Å resolution crystal structure reported in the pdb files 1P7L and 1RG9, which contain the reactants L-methionine and AMPPNP [24] or products AdoMet and PPNP. Extensive interactions between two subunits of the protein are evident in the crystal structures, consistent with a dimer (with the subunits denoted by A and B in the pdb file) being

the minimal functional unit [20]. Thus a dimer was chosen as the protein framework for our studies.

The protein structure was prepared by removing all ligands, followed by addition of hydrogens as appropriate for ionization states at neutral pH, positioning the polar side chains in tautomers and rotamers to maximize hydrogen-bonding and ion-pairing interactions. Computational software from Schrodinger L.L.C. (New York, NY) was used throughout this study. The protein was then subjected to a restrained MM minimization using the OPLS-AA force field (2001) in the program IMPACT; during this minimization all heavy atoms were constrained by a modified harmonic potential with an energy penalty $\Delta E = C*(r-r_0)^2$ for $(r - r_0) > 0.3 \text{ \AA}$ (with $C = 25 \text{ kcal}/(\text{mol}\cdot\text{\AA}^2)$) and $\Delta E = 0$ for $(r - r_0) \leq 0.3 \text{ \AA}$ so that there was no energy penalty for movement within 0.3 \AA of the crystallographic positions, The tolerance of 0.3 \AA was chosen based on an analysis of the experimental uncertainty in atomic coordinates according to Luzatti [38]. The final prepared structure included a total of 11921 atoms and had an rms deviation of 0.17 \AA for all heavy atoms relative to the un-minimized protein structure, with a maximal deviation of 0.33 \AA . The system was not “solvated” by the addition of explicit surface water molecules because the active site is buried deep within the protein structure and the surface atoms were frozen throughout the calculations. There is little quantitative information available in the literature concerning the importance of explicit solvation of the protein surface in studies such as ours, however in a QM/MM study of cytochrome P450(cam) the addition of a solvent layer around the protein had minimal influence on the relative energies of the species along the calculated reaction path [39]. Upon completion of the MAT protein preparation, the ligands were returned to the MAT structure file at their original coordinates. Ligand structures were manually prepared by the addition of appropriate hydrogen atoms and bonds to the structures in the pdb file.

The region of the reactant complex to be treated by DFT (the “QM region”) was chosen to include one active site, see Fig. 1A. This QM region (the “Small_QM region”) contains one equivalent each of AMPPNP, methionine, K^+ , two Mg^{2+} ions, as well as five water molecules that were reported in the pdb file [24]; two of these water molecules interact with the adenine moiety and the other three are in the vicinity of the terminal γ -phosphoryl group of AMPPNP. Two amino acid side chains were included in the Small_QM region, *i.e.* histidine-14, which forms a hydrogen bond to $O5'$ in the scissile bond of AMPPNP, and aspartate-16, which is coordinated to the Mg^{2+} ion denoted as MgA.

The water molecules needed to complete the presumed octahedral coordination sphere of the two Mg^{2+} ions [35,40-42] were not identified in the X-ray structure as a result of crystallographic disorder. However, EPR studies with Mn^{2+} in place of Mg^{2+} at the corresponding sites clearly demonstrated that Mn^{2+} is octahedrally coordinated in all the analogous complexes [18]; based on this experimental information for Mn^{2+} , and the intrinsic coordination preferences of Mg^{2+} , we assumed that both Mg^{2+} were octahedrally coordinated throughout [40,42]. This consideration required the addition of two explicit water molecules to the inner coordination shell of MgA and three explicit water molecules to the inner coordination shell of MgB (see Figs. 1A and 1B). The crystal structure of MAT shows that the active site K^+ is coordinated by the carboxylates of glutamate-42 and aspartate-238, the carbonyl oxygen of cysteine-239, and is $\sim 2.6 \text{ \AA}$ from the closest substrate atom, a non-bridging oxygen on the β phosphoryl group of the nucleotide; the K^+ ion is more than 6 \AA from the $C5'$ of AMPPNP that is the center of the reaction [24]. In the absence of experimental data on the hydration of the active site K^+ ion, no waters were added to this ion. The net charge on the Small_QM region was zero.

The QM/MM program QSite (v. 4.0) from Schrödinger, L.L.C. (New York, NY, 2005) was used for simulations of a postulated reaction path [43] (see below). QSite uses a frozen bond

approach to divide the QM and MM regions of the protein [43]; the bonds between the C_{α} and C_{β} of histidine-14 and aspartate-16 were chosen as sites to partition the side chains into the QM region, leaving their main chain atoms to be treated by MM. In simulations using a larger QM region, denoted as the “Large_QM region”, the side chains of the additional charged active site residues, arginine-244 and lysines-165, -245 and -265, were added to the QM region (Fig. 1B); this QM region had a net charge of +4. The B3LYP hybrid density functional was employed in all calculations [44,45]. Geometry optimizations for the Small_QM region (133 atoms) used the LACVP+** basis set [46], which includes polarization functions on all atoms and diffuse functions on the heavy atoms; optimizations for the Large_QM region (229 atoms) required use of the smaller LACVP* basis set due to computational limitations [46]; in the Large_QM case, single-point calculations at the larger B3LYP/LACVP+** computational level using the B3LYP/LACVP* optimized geometry were subsequently performed to allow some comparison with the energy changes calculated for the Small_QM region. The LACVP basis set employs the Pople split-valence 6-31G basis set for all the atoms except potassium which is modeled by a Los Alamos pseudopotential [46]. All of the ligand atoms in a single active site of the dimer were incorporated into the QM region; the other active site of the dimer was included in the frozen MM region. An envelope comprised of the entirety of any residue with an atom lying within 10 Å of any atom in the QM region was treated as mobile, with energetics described by the OPLSAA (2001) MM force field [47]. Charged amino acids in proximity to the ligands were included in the MM region without any geometrical constraints; these were residues 165, 244, 245 in chain A, and residues 118, 265, 271 in chain B. Additional MM atoms in the 10 Å envelope were constrained to the positions obtained in the pre-optimization procedure, r_0 , by a harmonic potential, with $\Delta E = 25*(r-r_0)^2$ (see Methods) which we previously found to be satisfactory in other MM applications [48]; these constrained residues were: Chain A residues 8-11, 163, 164, 166, 186-188, 227-243, 247, 248; and Chain B residues 40-42, 55, 98-102, 117, 119, 120, 259-264, 266-270 and 302. The positions of the remainder of the atoms were frozen.

For calculations using the Large_QM region, the sidechains of His-14 (Chain A), Asp-16 (Chain A), Lys-165 (Chain A), Lys-245 (Chain A), Arg-244 (Chain A) and Lys-265 (Chain B) were in the QM region; the residues in the mobile, unconstrained MM region were: the backbone atoms of residues in the QM region, and also residues 118 and 271 from Chain B. In the mobile MM region, but constrained by a restoring force of 25 (kcal/mol-Å²), were: Chain A residues 8-11, 163, 164, 166, 186-188, 227-243, 247, 248; and Chain B residues 40-42, 55, 98-102, 117, 119, 120, 259-264, 266-270 and 302. The positions of the remainder of the atoms were frozen. The constrained MM region included 858 atoms for calculations using the Small_QM representation, and 803 atoms for calculations using the Large_QM representation and their total charges were +2 and -2, respectively. The convergence criteria were a maximal energy change of 0.1 kcal/mol and a gradient maximum of 0.2 kcal/mol-Å. Non-bonded electrostatic cutoffs were not employed to avoid difficulties arising from the movement of atoms across the cutoff distance during the optimization.

After addition of waters around the two Mg^{2+} ions to complete their expected octahedral inner hydration shell [40,42] and preparation of the protein as described above, the reactant geometry was optimized by QM/MM at the B3LYP/LACVP+** level without constraints in the QM region, see Fig. 1. The rmsd for the non-hydrogen atoms was 0.38 Å between the optimized structure and the crystal structure (the maximum rmsd was 0.98 Å for a water oxygen atom) showing the lack of substantial geometrical changes upon this optimization. This structure formed the basis for further QM/MM calculations. Because computational limitations required a smaller basis set to be used in geometry optimizations of the Large_QM region, as a reference we examined the structure of the Small_QM region after optimization at both the B3LYP/LACVP+** and B3LYP/LACVP* levels for the reactant, transition state and product. Comparisons of the effect of basis set on the calculated geometry of each of these three

structures revealed only modest differences, with an rmsd of maximally 0.14 Å with a maximal deviation of 0.40 Å for a water atom.

Since AdoMet synthesis involves breaking the C5'-O5' bond and forming the C5'-S bond, we choose $\xi = (R(C5'-O5') - R(C5'-S))$ as the approximate reaction coordinate; practically, this was implemented by coordinate driving, *i.e.* incrementally decreasing the distance between the nucleophilic sulfur of methionine and the C5' of ATP, with optimization of the remainder of the QM and mobile MM regions. A variety of tests showed that an initial step size of ~0.1 Å, with propagation of the wave function at each step, allowed a relatively smooth transformation of the geometry (as well as the total and MM energies) from reactant to product. Once the C5'-S distance reached the length of a typical C-S bond (1.8 Å), the geometry of the product was optimized without any constraints in either the QM or mobile MM regions. After an initial path from reactants to products was completed, additional intermediate points along the reaction coordinate were obtained from calculations in which the C-S distance was increased, starting at the product geometry.

Charges from Natural Population Analyses (NPA) and bond orders from Wiberg bond indices in the Natural Atomic Orbital (NAO) basis were calculated using the NBO module in Jaguar v. 7.0 and NBO v. 5.0 [49]. These single-point calculations were carried out at the geometries of the reactants, TS, and product in the Large_QM representation, using the highest level practical with our available computational resources, LACVP+*. In this analysis, the QM region was extracted from the surrounding protein by cutting the C α -C β bond for each amino acid and adding hydrogens on the C β at the positions previously held by the C α ; and the geometry of solely the new C-H bonds were then regularized with a MM minimization, with the remainder of the structure held fixed [50]. The energies of these resulting structures were calculated at the LACVP+* level; the difference between the energies in the presence and absence of the bulk protein reflects to some extent the influence of the overall protein environment on the reaction energetics.

Relative thermal corrections for the reactant, TS and product structures were estimated from frequency analyses calculated for truncated versions of the Small_QM region because of the large computational demands of frequency analyses. The QM region in these calculations included His-14 and Asp-16 (both truncated and hydrogen capped at C β as above), AMPPNP, methionine, the two Mg²⁺ with the total of five waters in their first coordination spheres, and the K⁺ ion; the LACVP* basis set was used for frequency analyses, the highest level practical with our computational resources. A similar method of truncation in order to allow the computationally intensive frequency analysis has been reported previously [50].

Results

Choice of Starting Structure

The starting geometry for our series of calculations reported in this paper was taken from the crystal structure for the reactant complex containing AMPPNP and methionine (pdb file 1P7L) [24]. In the crystal structures the primary differences between the reactant and product complexes corresponded to movements of those atoms involved in the substitution at C5'; there were no substantial differences in the locations of the metal ions or of the active site amino acid residues in the reactant and product structures. Thus, the crystallographic data provided confidence that this enzymatic reaction was well suited to studies employing computational methods which model changes in bonding in the active site but do not take into account larger protein structural changes such as movement of the loop that gates access to the active site.

Modeling of the Formation of AdoMet

Fig. 2 illustrates the reaction coordinate diagram for the conversion of methionine and AMPPNP to AdoMet and PPNP in the Small_QM representation. In these calculations all species are “enzyme-bound”, i.e. the periphery of the overall catalytic reaction sequence that includes substrate binding and product release (*i.e.* enzyme turnover) was not considered. The calculated approximate activation energy of +19.1 kcal/mol, see Table 1, is consistent with the experimental value of ~+19 kcal/mol (a rate of 0.06 s^{-1} at 298K) [13]. Estimates of thermal corrections to 298 K were obtained on a reduced QM region using the LACVP* basis set (see Computational Methods). These thermal corrections differed by less than 2 kcal/mol among the reactant, TS and product structures; given the approximations involved with these calculations, we chose not to include these thermal corrections in our further analyses.

The maximal energy in the reaction coordinate occurred when ξ was $\sim -0.41 \text{ \AA}$, reflecting a C5'-S distance of 2.40 \AA and a C5'-O5' distance of 1.99 \AA ; the S-C5'-O5' and C5'-S-C_{methyl} angles in this structure were $\sim 160^\circ$ and 105° , respectively. These calculated structural parameters are consistent with a classical S_N2 displacement of the leaving O5' as the sulfur of methionine approaches. The calculated TS structure is comparable to that previously described based upon vibrational analysis of kinetic isotope effect data, from which it was deduced that the transition state occurred near a value of $\xi = \sim -0.24 \text{ \AA}$ (C5'-S and C5'-O5' distances of 1.96 \AA and 1.72 \AA , respectively), with a C5'-S-C_{methyl} angle of 102° [19]. The QM/MM calculated TS has less bond formation than that deduced from the kinetic isotope experiments; the calculated Wiberg index (bond-order [51]) for C5' was calculated to be of 3.78 at the TS (B3LYP/LACVP+* level), which is slightly less than that estimated from the kinetic isotope experiments, 3.96 [19]. In light of the approximations in both the current study and in the analysis of the experimental kinetic isotope effects, these deduced structural parameters are in reasonable agreement. An analogous finding of a looser transition state in QM/MM calculations than that deduced from kinetic isotope effect data was noted in simulations of the transfer of a methyl group from AdoMet to a primary amine [52]. The energy of the product was calculated to be -1.8 kcal/mol relative to the reactant; product formation in the crystal is nearly thermoneutral [24].

Structural Changes Across the Reaction Coordinate

Figs. 3A and 3B present a superposition of the calculated active site structures for the reactant, TS, and product in the Small_QM and Large_QM representations, respectively. The structural changes upon product formation are qualitatively the same in both models. In the Small_QM model the major motion calculated upon going from the reactant to the TS is the movement of C5' by 0.48 \AA toward the electron rich sulfur; the calculated total displacement of C5' in the product is 0.94 \AA , comparable to the displacement of 1.3 \AA deduced from the crystal structure [24, 53]. The translation of C5' is accompanied by the approach of the sulfur toward C5', moving by 0.45 \AA at the TS and by a total of 0.58 \AA at the product; the displaced O5' is calculated to move by 0.04 \AA at the TS then by an additional 0.30 \AA at the product. The reaction is accompanied by a change in the ribose ring pucker due to movements of both C3' and O4', each by $\sim 0.19 \text{ \AA}$ between the reactant and the TS, with total movements of 0.44 \AA and 0.48 \AA between reactant and product, respectively; this calculated change in the ribose puckering is in accord with that seen in the crystal structure and also deduced from NMR measurements [24, 53]. Thus the reaction path largely reflects the motion of the ribose methylene group and to a lesser extent a movement of the thioether of methionine.

To examine changes in interactions along the reaction path, we monitored the Mulliken charges on selected atoms at the B3LYP/LACVP+*** level [54]; additional charges were obtained from Natural Population Analyses (NPA) for the reactants, TS, and products in the Large_QM region (see the Computational Methods section and below). As anticipated, product formation is

accompanied by accumulation of negative charge on the oxygen atom that is liberated upon C5'-O5' bond cleavage, whereas the calculated positive charge increases on the nucleophilic sulfur atom (see Fig. 4A). The charge on the leaving oxygen is calculated to become $\sim 0.3e$ more negative, with the rest of the charge formally on the oxygen being distributed among the Mg^{2+} ions and their ligands. The calculations predict few other structural changes in the active site as the reaction proceeds, consistent with crystallographic results (illustrated in Figs. 3A and 3B); it is important to recall that none of the displayed atoms were constrained during the QM/MM optimizations.

The crystal structures show that the α phosphoryl group, which is the leaving moiety in the reaction, is encompassed by a number of hydrogen bond donors as well as the two Mg^{2+} ions [24]. In comparing the calculated reactant and product structures, the hydrogen bond distance from O5' to the $H_{\epsilon 2}$ of His-14 ($O5' \cdots H-N_{\epsilon}(\text{His-14})$) is 0.23 Å smaller in the product (see Fig. 4B), suggesting an increase in the strength of this bond as the reaction proceeds; the majority of this movement is due to changes in the AMPPNP (PPNP) positions, with the histidine side chain remaining virtually stationary in relation to its position in the reactant (see Fig. 3). The hydrogen bond angle for $O5' \cdots H-N_{\delta}(\text{His-14})$ is 168° in the reactant and changed by less than 1° throughout the reaction coordinate. Possible interactions of $N_{\delta 1}$ of His-14 with the amide hydrogens of Asp-16 and Lys-17 are reflected in $H \cdots N_{\delta 1}$ distances of 2.50 Å ($\angle N-H \cdots N$ 142°) and 2.81 Å ($\angle N-H \cdots N$ 144°) respectively in the reactant; these two $N \cdots H$ distances change by less than 0.02 Å across the reaction coordinate while the $N-H \cdots N$ angles vary by less than 2°. Apparently the interactions between His-14- N_{δ} and the amide hydrogens of Asp-16 or Lys-17 do not differ significantly along the reaction coordinate.

The length of the hydrogen bond from a H_{ζ} of Lys-165 to O5' is calculated to decrease by ~ 0.25 Å as the reaction proceeds (see Fig. 4B). It is noteworthy that the calculated length of the hydrogen bond from the cationic Lys-165 to O5' is consistently ca. 0.22 Å shorter than the hydrogen bond from the formally uncharged His-14 to the same oxygen. The hydrogen bonds from the H_{ζ} of Lys-245 to the non-bridging O_{α} are calculated to decrease in length by ~ 0.18 Å and ~ 0.24 Å as the reaction proceeds (Fig. 4C). The $H \cdots O$ distance for the O_{α} that is coordinated to MgA is ca. 0.18 - 0.21 Å shorter than the $H \cdots O$ distance for the O_{α} that is coordinated to MgB which has two fewer anionic ligands; the respective $N-H \cdots O$ bond angles are 133° and 136° in the reactants and both increase by ca. 5° in the products, the geometries suggesting consistently sub-optimal hydrogen bonding interactions. The length of the hydrogen bond between H_{ζ} of Lys-265 and the sole terminal phosphoryl group oxygen (O_{γ}) that is not coordinated to a Mg^{2+} ranges from ~ 2.4 Å in the reactants to ~ 2.3 Å in the products, ~ 0.55 to ~ 0.70 Å longer than that involving Lys-165 and O5' at the other end of the polyphosphate chain.

Fig. 4D shows that the $O5'-P_{\alpha}$ (denoted $O_{\alpha}-P_{\alpha}$ in the product) distance decreases by 0.11 Å across the reaction coordinate, while the calculated $P_{\alpha}-O_{\alpha\beta}$ distance ($O_{\alpha\beta}$ is the bridging oxygen in the $P_{\alpha}-O_{\alpha\beta}-P_{\beta}$ moiety) increases. The $P_{\alpha}-O_{\alpha\beta}$ separation is linearly related to the C5'-O5' distance, with an increase of 0.05 Å in $P_{\alpha}-O_{\alpha\beta}$ bond length for each Ångstrom increase in C5'-O5' distance (not shown). The lengths of the bonds from P_{α} to its two non-bridging O_{α} are also computed to increase upon product formation (by a total of 0.02 Å and 0.01 Å for the oxygens coordinated to MgA and MgB, respectively). These distance changes suggest that as additional negative charge accrues on the α -phosphoryl group upon C5'-O5' bond breaking, the hydrogen bonding interactions between the α -phosphoryl group oxygens and cationic active site amino acids strengthen which would stabilize the TS, and subsequently the reaction product.

Figs. 3A and 3B illustrate that the positions of each Mg^{2+} minimally alter upon product formation, and throughout the reaction the Mg^{2+} are calculated to retain their octahedral coordination geometries. Although the Mg^{2+} ions each moved slightly closer to O5' as the

reaction proceeded across the LACVP+** calculated reaction coordinate (by 0.05 Å for MgA and 0.10 Å for MgB), neither approached O5' closer than 4.39 Å, thus neither directly bound to the leaving O5' atom. The total displacements of MgA and MgB were 0.09 Å and 0.17 Å between reactant and product, leading to a decrease in the MgA---MgB distance of ~0.03 Å in the product, and a ~0.03 Å increase in both of the Mg²⁺---K⁺ distances. Fig. 4E illustrates the variation across the reaction coordinate of the distance between each Mg²⁺ ion and the oxygen of the α -phosphoryl group to which it is coordinated. The bond distance decreased in both cases, by 0.06 Å for MgA and by 0.11 Å for MgB; Mulliken charges on the Mg²⁺ ions also decreased as the reaction proceeded. The Mulliken charge changes by -0.06e for MgA which is coordinated to all three phosphoryl groups and Asp-16, and by -0.08e for MgB for which the anionic ligands are the α and γ phosphoryl groups. These calculated charge alterations substantiate the anticipated roles of the Mg²⁺ as sinks for the negative charge that must develop as the C5'-O5' bond breaks, consistent with charge dissipation as a factor in catalysis [18],

The position of the K⁺ ion is calculated to change by 0.40 Å between reactant and product, moving 0.26 Å away from O₈₂ of Asp-238 in the product (to a distance of 3.30 Å), while remaining at least 2.59 Å from the β -phosphoryl group oxygens. The alteration in the position of the monovalent cation, appearing as a blurring of its sphere in Fig. 3, did not appear to have any direct functional significance.

Fig. 4F presents the variation in the separation of the nucleophilic sulfur and the O4' of the ribose ring across the reaction coordinate. When the reaction coordinate has a value less than ~-1.2 Å, the S---O4' distance becomes less than the sum of their van der Waals radii (~3.3 Å [55]), and the separation remains near to, or less than, 3.34 Å for the remainder of the reaction coordinate. A stabilizing close, non-bonded interaction between sulfur atoms and ribose oxygens has been reported previously in studies of various compounds, including nucleosides and AdoMet [48,55,56]. MAT appears to utilize this interaction to stabilize the TS and subsequently the product, at the active site.

Alterations in the Quantum Mechanical Region

In order to assess the sensitivity of this model mechanism to our computational methodology, we performed additional calculations with alterations in the composition of the QM region, paying particular attention to the influence of His-14 and the charge in the region.

Expansion of the QM Region

We examined the effect of expanding the QM region to include additional cationic residues at the active site, an arginine and 3 lysines, on the computed reaction energies. The crystal structure showed that these residues all interact with the polyphosphate chain (see Fig. 1B). The structures of the reactant, product and the approximate TS were optimized for this Large_QM region at the B3LYP/LACVP* level, without constraints in the QM region for both the reactant and the product, but with the C5'-S distance fixed at 2.40 Å for the approximate TS (see above). The computed C5'-O5' distance for this TS remained 1.99 Å, as was present in the starting Small_QM region (B3LYP/LACVP+** level) geometry. The rmsd between the constituents of the Small_QM region and the same atoms in the optimized Large_QM region was 0.14 Å, with a maximal deviation of 0.40 Å. Single-point energy calculations at the B3LYP/LACVP+** level, using the B3LYP/LACVP* optimized geometries, reported that the energy of the reactant was +1.0 kcal/mol higher than the product (+3.0 kcal/mol at the LACVP* level), while the energy of the approximate TS was +18.8 kcal/mol above the reactant energy (+18.0 kcal/mol at the LACVP* level).

In the Large_QM active site representation, the changes in hydrogen bond lengths among reactants, TS, and products qualitatively echo those observed for the Small_QM region. For

example, comparing the optimized structures of the reactant and product reveals that the length of the hydrogen bond from His-14 to O5' is calculated to decrease by 0.12 Å (to 1.93 Å in the product) in the Large_QM representation, compared to a ~0.23 Å decrease (to 1.81 Å in the product) in the Small_QM description. The H•••O5' distance in the hydrogen bond from the ammonium group of Lys-165 to O5' was calculated to decrease by 0.29 Å in the Large_QM model, to 1.61 Å, as product formation ensued, similar to the computed decrease from ~1.86 Å to ~1.61 Å in the Small_QM representation. There were minimal changes elsewhere in the vicinity of the active site (see Figs. 3A and 3B). Apparently the contributions of these 4 cationic residues are reasonably well represented by the mobile MM description of the Small_QM system, with the result that their addition to the QM region did not substantially alter the calculated energy changes, although it greatly increased the computational cost. These results are consistent with QM/MM studies on other enzymes which have reported that relative energies among reactant, TS, and products are not dramatically sensitive to the geometries optimized at modestly different computational levels [35].

Contraction of the QM Region

The effect of contracting the QM region was examined using single-point calculations at the B3LYP/LACVP+** computational level, in order to further assess the effects of altering the net charge on the QM region on the calculated reaction energetics. The reduction was accomplished by moving the side chains of His-14 and Asp-16 from the Small_QM region to the MM region, both individually and together, and of placing only the side chains of His-14 and Lys-165 in the QM region. This allowed the sensitivity of the calculated reaction energetics to the charge of the QM region to be assessed as the charge varied from 0 to +2. Table 1 shows that the calculated relative energies varied by only ~2.2 kcal/mol for the TS and by ~4.8 kcal/mol for the product compared to the reactant. The variations are remarkably small given the complexity of this system.

The role of the protein may be most simply seen by comparison of the calculated reaction energetics when the ligands alone are in the QM region and the protein is either all in the MM region or is absent entirely (Table 1). The presence of the protein solely in the MM region provided a calculated TS energy (relative to the reactant) of +21.0 kcal/mol, and the relative energy of the product was +1.0 kcal/mol both at the LACVP+** level. This beneficial effect on the calculated reaction energetics of including the protein solely in the MM region shows the significant influence of interactions that are reasonably well represented by MM.

Completely removing the protein while maintaining the relative positions of the reactants, TS and products resulted in substantial increases in the calculated energies to +37.0 kcal/mol for the TS and +33.2 kcal/mol for the product, both relative to the reactant at the LACVP+** level. The results show that much of the catalytic influence of the protein stems from contributions in addition to its organization of the active site since the relative positions of the atoms were maintained in these calculations.

Mutation of the Active Site Histidine

Site directed mutagenesis studies have shown that the side chain having the largest influence on the rate of AdoMet formation is that of His-14; mutation of this evolutionarily conserved residue to Asn-14 reduced the maximal rate of AdoMet formation by 10⁴-fold, while changing substrate K_M values by less than 2-fold and decreasing the maximal rate of the subsequent PPP₁ hydrolysis step by ca. 30-fold [21]. We examined the effect of an *in silico* H14N mutation by altering the His-14 side chain in the geometry optimized Large_QM region, while maintaining the same rotamer. The position of the resultant Asn-14 is such that it cannot form hydrogen bonds to the substrates as is illustrated in Fig. 5. This same mutation protocol was

carried out on the reactant, TS and product using the structures from the geometry optimized Large_QM region.

The structures of the Large_QM region of the H14N mutant were geometry optimized at the B3LYP/LACVP* level for the reactant, approximate TS and product. There were only small changes in the structure of the QM region compared to the wild type enzyme, with rms deviations of 0.02 Å, 0.08 Å and 0.11 Å for reactant, TS (C5'-S distance fixed at 2.40 Å, ξ was computed to be -0.41 Å) and product, respectively (the rmsd was calculated considering all non-hydrogen atoms in the QM region with the exception of the side chain of residue 14). The largest individual displacement between the mutant and wild type enzymes was 0.50 Å for C5' in the TS structure; the maximal change in position of Asn-14 atoms among the three structures was 0.33 Å for N δ . Single-point energies were then calculated at the B3LYP/LACVP+** level for a more consistent comparison with the wild type enzyme (Table 1). The calculated energies of the TS and product were +23.2 and +12.1 kcal/mol above the reactants; this corresponds to an approximately 5 kcal/mol increase in the TS energy over that for the same model of the wild type enzyme, and a ~13 kcal/mol increase in the energy for product formation, changing the reaction from modestly exothermic (as experimentally observed for the native enzyme) to substantially endothermic. The 10⁴-fold reduction in catalytic rate experimentally observed for the H14N variant corresponds to a ~6 kcal/mol increase in activation energy [21]. While there are no experimental data available for the change in the equilibrium constant upon mutation, the computational results support a substantial contribution of the histidine to the enzyme's prowess even though it is uncharged.

Discussion

The results of this study substantiate that the active-site structure found in crystallographic studies of *E. coli* MAT is well-suited to catalyze its biological function of S-adenosylmethionine formation. The calculations reported herein mimic a single turnover of AdoMet formation, which was the experimental technique employed in the measurements to which the calculations are compared. Multiple enzyme turnovers require either the hydrolysis of the tripolyphosphate (or its analog) formed in conjunction with AdoMet, or the ~10⁴-fold slower dissociation of these reaction intermediates. In either case, multiple turnovers require protein conformational alterations including movement of the loop that gates access to the active site, events that are not required for a single turnover, which made the AdoMet formation step well suited to QM/MM studies. The energetic differences between the reactant and product structures, as well as between the reactant and the approximate TS, were calculated using a QM/MM methodology (B3LYP/(OPLS-AA (2001)) with the active site amino acids in ionized states with the exception of the neutral His-14. The computed TS energy was +18.8 kcal/mol above the reactant (at the B3LYP/LACVP+** level), and the relative energy of the product was +1.0 kcal/mol at the same level. The calculated values are in reassuringly good agreement with the experimental values of ~+19 kcal/mol and ~0 kcal/mol, respectively [13,24]. Thus the calculations support the ability of the highly polar active site to provide a suitable steric and electronic framework for catalysis of AdoMet formation. It is remarkable that experimental studies show that the residue most intimately involved in the reaction is a neutral histidine, in the N ϵ protonated tautomer, which crystallographic data show forms a hydrogen bond to the O5' atom that is displaced from the nucleotide [24]. The experimental result that showed a ~10⁴ fold reduction in catalytic activity when the evolutionarily conserved His-14 is replaced *in vitro* by asparagine revealed its critical role in catalysis [21]. Our calculations illustrate the approach of H ϵ of His-14 to O5' as the reaction proceeds suggesting a strengthening in the hydrogen bond. An analogous electrophilic role for a neutral histidine has been demonstrated for triose phosphate isomerase (TIM) [57]. The catalytic histidine-95 of TIM forms a strong hydrogen bond with the enediol reaction intermediate; the histidine's pKa was shown to be reduced by at least 2 units by the electrostatic effect of its location at the N-terminus of an α -

helix [58]. This reduced pKa was deduced to provide an improved match with the pKa of the enediol reaction intermediate, thus optimizing hydrogen bond strength and enabling transient formation of an imidazolate anion upon proton transfer [58]. In MAT, His-14 is also located at the N-terminus of an α -helix, and an abnormally low pK would be consistent with the invariance of k_{cat} over the examined pH range of 6 to 9 [13]. The calculated decrease in the length of the hydrogen bond from His-14 to the leaving O5' as the reaction proceeds implies that the proton interacts more strongly with the TS and product than with the reactant, consistent with the increasingly negative charge on the oxygen (Fig. 4A). The appropriate pKa of a di-magnesium complex of imido-tripolyphosphate (or a related polyphosphate) could not be identified in the literature, but might be expected to be well below that available to an imidazole, suggesting that a match of pKa values to optimize hydrogen bond strength has not been attained. Nevertheless, the use of a neutral imidazole as an electrophile obviates the need for the ligand-free protein to maintain an amino acid in an unusual protonation state, such as a protonated carboxylic acid, in preparation for catalysis.

The interaction of O4' of the ribose and the nucleophilic sulfur appears to reflect an example of substrate assisted catalysis. Fig. 4A illustrates that the sulfur atom gains positive charge as the reaction proceeds, and Fig. 4F shows that the distance from the sulfur to O4' concomitantly decreases, becoming less than the sum of the van der Waals radii of sulfur and oxygen ($\sim 3.3 \text{ \AA}$) at $\xi \sim -1.0 \text{ \AA}$. The minimum S-O4' separation occurs approximately at the TS, then the distance increases but remains near to, or below, the sum of the S-O4' van der Waals radii for the remainder of the reaction. The precise energetic magnitude of this close, non-bonded interaction is unclear in the context of the active site, but the interaction would be anticipated to provide a stabilizing influence for both the TS and product [48,55,56]. The computational results rationalize our previously puzzling finding that replacing Oxygen-4' by a CH₂ group (yielding 4'-deoxy-ATP, aristeromycin-TP) resulted in a complete loss in substrate activity ($< 0.1 \%$ of ATP). 4'-dATP has good affinity for the enzyme, being a competitive inhibitor with respect to ATP with a K_i of $8 \mu\text{M}$, 13-fold less than the K_m for ATP (unpublished results).

The present studies also address the significance of the active site structure reported in the 2.9 \AA resolution crystal structure of the rat hepatic MAT (rMAT) [25,26]. The rMAT active site contained ATP, methionine, 3 Mg^{2+} ions, K^+ and 2 PO_4 moieties [25,26]. The overall cMAT and rMAT structures are nearly identical, sharing the same topology with an rmsd of 1.3 \AA for C α . While the triphosphate chain of the nucleotide is located in a comparable position in the two protein structures, the methionine substrate in rMAT is stacked upon phenylalanine-251, whereas the adenine group stacks upon the analogous F230 in cMAT. Furthermore in rMAT the adenosine moiety is effectively positioned at the opposite end of the triphosphate chain from its location in cMAT, as discussed in a recent review [25,26]. The mechanistic significance of the rMAT structure is further diminished because of the bound PO_4 which is an inhibitory product of the reaction. His-30 of rMAT, the analog of His-14 in cMAT, is in proximity to the Py of ATP, and not in position to facilitate AdoMet formation. Thus the current interpretation is that the rMAT structure may reflect a complex on the route to product release rather than the enzymatically active species [25,26]

Various experimental and computational studies have now combined to show that the catalytic devices employed by MAT are a combination of binding the reactants in proper orientations to facilitate the reaction (perhaps conforming to a near-attack-conformation model [31,59, 60] and the pre-positioning of the polar active site functional groups that form hydrogen bonds to the ligands [61]. Comparison of the crystal structures of the active site of the fully liganded enzyme (pdb codes 1P7L and 1RG9) and the crystal structure of the ligand-free enzyme (pdb code 1FUG) shows that the active site residues are largely in place in the apo enzyme, with the exception of the mobile loop which forms the active site lid and is not directly involved in catalysis [17,23,24]. The Mg^{2+} ions used in catalysis of AdoMet formation help dissipate the

negative charge that would otherwise accumulate on the departing oxygen atom. The Mg^{2+} ions can also play a role in establishing the unusual bent conformation of the polyphosphate chain of AMPPNP (and PPNP) at the active site [24]. The binding of one Mg^{2+} to both ends of the polyphosphate chain (forming an 8-membered ring), and the other Mg^{2+} to all three phosphoryl groups, forming two six-membered ring systems, appears to be a unique coordination scheme among known enzymes, consistent with its exceptional task of catalyzing reactions at both ends of the polyphosphate chain. Other proteins are known to employ two Mg^{2+} coordinated to the polyphosphate chain of ATP to catalyze transfer of the γ -phosphoryl group; these include protein kinase A and pyruvate kinase [62,63]. In protein kinase A, both Mg^{2+} are coordinated to the γ phosphoryl group; in addition one is coordinated to an α phosphoryl group oxygen whereas the other binds to a β phosphoryl oxygen atom and approaches the β,γ bridge oxygen [62]. In pyruvate kinase, both Mg^{2+} are coordinated to the γ phosphoryl group of ATP; one is also coordinated to the α and β phosphoryl groups while the second does not have additional interactions with the polyphosphate chain [63]. The active site of the Mn^{2+} dependent enzyme phospho(enol)pyruvate carboxykinase contains both Mn^{2+} and Mg^{2+} ; a Mg^{2+} ion is coordinated to the β,γ phosphoryl groups while a Mn^{2+} binds to an oxygen of γ phosphoryl group [64].

The only other enzyme that is known to catalyze a reaction at the C5' position of ATP is the Coenzyme B12 adenosyltransferase, in which the nucleophile is a Co(I) embedded in a corrin framework [65]. The crystal structures of representatives of two distinct classes of this enzyme, i.e. from *Salmonella typhimurium* and from *Lactobacillus reuteri*, have been reported [66, 67]. In both cases a single Mg^{2+} ion is bound to ATP in the active site; in the *S. typhimurium* enzyme it is coordinated to the α and β phosphoryl groups [66], whereas in the *L. reuteri* enzyme all three phosphoryl groups are ligands to the Mg^{2+} [67]; in the latter case an arginine residue also interacts with the O5' from the leaving adenosyl group, whereas no comparable interacting cationic protein groups were noted in the former structure. Thus the detailed mechanisms of corrin adenosylation are likely to be significantly different from MAT.

The reverse of the physiological MAT catalyzed reaction is alkylation of the di- Mg^{2+} complex of tripolyphosphate, which invites comparison to other alkylation reactions utilizing AdoMet. Studies of methyl transfer from AdoMet to different acceptors, have led to the general conclusion that the reactions proceed by S_N2 mechanism [52,68-72], although for lysine- N_ϵ methylation a partially dissociative transition state with reduced bond order for the methyl carbon has been implicated by QM/MM studies [52]. The associative transition state of the MAT reaction is apparently similar in general terms to other alkylations. Our ongoing QM/MM studies are directed toward providing further understanding of the roles of the protein and metal ions in the two catalytic events at MAT's active site.

Acknowledgments

The work relied upon the FCCC computer cluster supported by the High Performance Workstation Facility, and the aid of Joseph Anlage from that Facility. FCCC Secretarial Services, in particular Marie Estes, contributed significantly to this manuscript. We thank Dr. Maxim Pimkin for preparing Fig. 1.

G.D.M. would like to thank the NIH (GM31186, CA006927) and NCI for financial support of this work, which was also supported by an appropriation from the Commonwealth of Pennsylvania.

1 Abbreviations

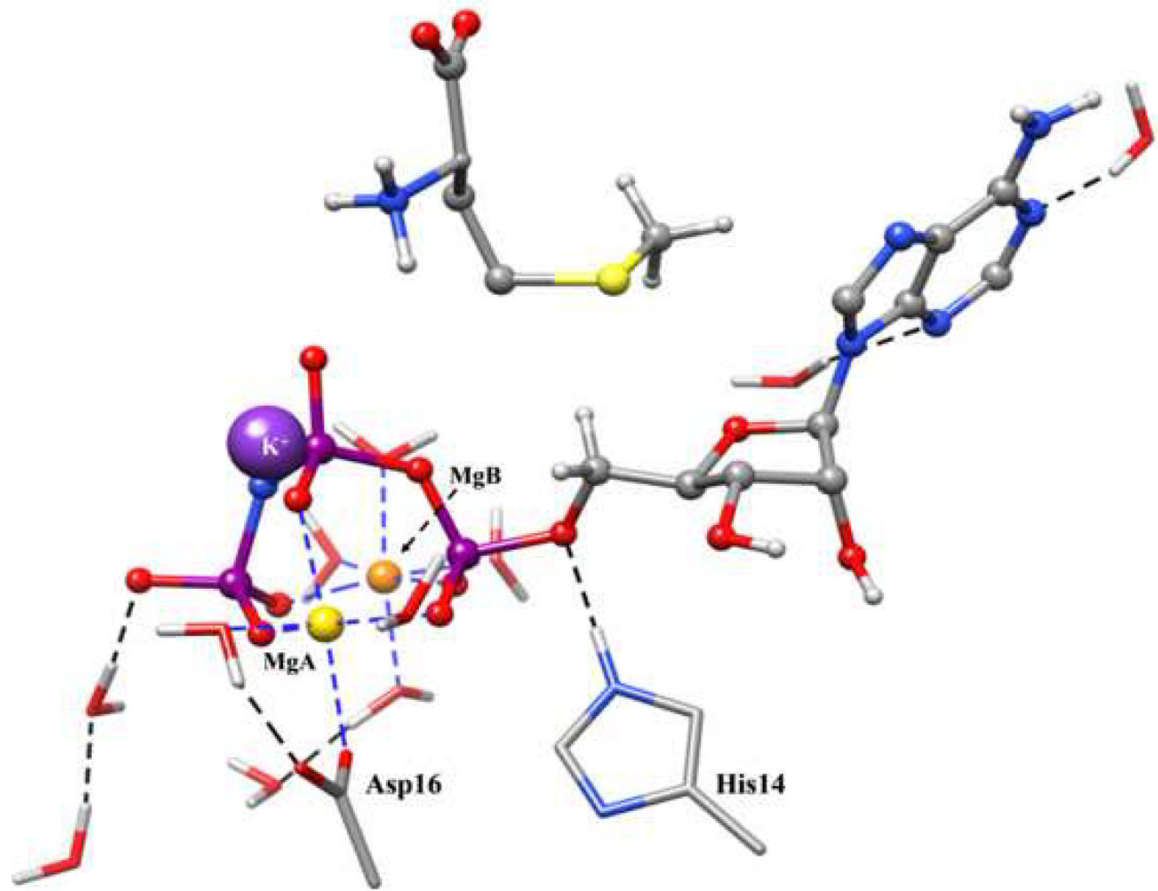
AdoMet, S-adenosylmethionine; AMPPNP, adenylylimidodiphosphate; cMAT, S-adenosylmethionine synthetase from *Escherichia coli*; DFT, density functional theory; MAT, S-adenosylmethionine synthetase (methionine adenosyltransferase); PPNP,

imidotriphosphate, $O_3P-NH-PO_2-O-PO_3$; QM/MM, quantum mechanical/molecular mechanical; TS, transition state.

References

1. Cantoni GL. *Annu. Rev. Biochem* 1975;44:435–451. [PubMed: 1094914]
2. Lu SC. *Internatl. J. Biochem. Cell Biol* 2000;32:391–395.
3. Markham GD. *Encyclopedia of Life Sciences* 2002. 2002http://www.els.net/ [doi:10.1038/npg.els.0000662]
4. Wang SC, Frey PA. *Trends Biochem. Sci* 2007;32:101–110. [PubMed: 17291766]
5. Bender J. *Annu. Rev. Plant Biol* 2004;55:41–68. [PubMed: 15725056]
6. Esteller M. *N. Engl. J. Med* 2008;358:1148–1159. [PubMed: 18337604]
7. Lu SC, Mato JM. *J. Gastroenterol. Hepatol* 2008;23(Suppl 1):S73–77. [PubMed: 18336669]
8. Tabor CW, Tabor H. *Annu. Rev. Biochem* 1984;53:749–790. [PubMed: 6206782]
9. Pegg AE. *Cancer Res* 1988;48:759–774. [PubMed: 3123052]
10. Cheek J, Broderick JB. *J. Biol. Inorg. Chem* 2001;6:209–226. [PubMed: 11315557]
11. Markham GD, Pajares MA. *Cell Mol. Life Sci* 2009;66:636–648.
12. Mudd, SH. *The Enzymes*. Vol. 3rd Edition. Academic Press; New York: 1973. p. 21-154.
13. Markham GD, Hafner EW, Tabor CW, Tabor H. *J. Biol. Chem* 1980;255:9082–9092. [PubMed: 6251075]
14. McQueney MS, Anderson KS, Markham GD. *Biochemistry* 2000;39:4443–4454. [PubMed: 10757994]
15. Sanchez-Perez GF, Bautista JM, Pajares MA. *J. Mol. Biol* 2004;335:693–706. [PubMed: 14687567]
16. Graham DE, Bock CL, Schalk-Hihi C, Lu ZJ, Markham GD. *J. Biol. Chem* 2000;275:4055–4059. [PubMed: 10660563]
17. Fu Z, Hu Y, Markham GD, Takusagawa F. *J. Biomol. Struct. Dyn* 1996;13:727–739. [PubMed: 8723769]
18. Markham GD. *J. Biol. Chem* 1981;256:1903–1909. [PubMed: 6257692]
19. Markham GD, Parkin DW, Mentch F, Schramm VL. *J. Biol. Chem* 1987;262:5609–5615. [PubMed: 3553181]
20. Reczkowski RS, Markham GD. *J. Biol. Chem* 1995;270:18484–18490. [PubMed: 7629176]
21. Taylor JC, Markham GD. *J. Biol. Chem* 2000;275:4060–4065. [PubMed: 10660564]
22. Taylor JC, Takusagawa F, Markham GD. *Biochemistry* 2002;41:9358–9369. [PubMed: 12135357]
23. Takusagawa F, Kamitori S, Markham GD. *Biochemistry* 1996;35:2586–2596. [PubMed: 8611562]
24. Komoto J, Yamada T, Takata Y, Markham GD, Takusagawa F. *Biochemistry* 2004;43:1821–1831. [PubMed: 14967023]
25. Papagrorgiou E, Shafqat N, Rojkova A, Niessen F, Kavanagh K, von Delft F, Gorrec F, Ugochukwu E, Arrowsmith C, Edwards A, Weigelt J, Sundstrom M, Oppermann U. Crystal structure of the alpha subunit of human Sadenosylmethionine synthetase 2002. 200620072018http://www.ebi.ac.uk/pdbsum/2002P2002
26. Gonzalez B, Pajares MA, Hermoso JA, Guillerm D, Guillerm G, Sanz-Aparicio J. *J. Mol. Biol* 2003;331:407–416. [PubMed: 12888348]
27. Parry RJ, Minta A. *J. Am. Chem. Soc* 1982;104:871–872.
28. McQueney MS, Markham GD. *J. Biol. Chem* 1995;270:18277–18284. [PubMed: 7629147]
29. Reczkowski RS, Taylor JC, Markham GD. *Biochemistry* 1998;37:13499–13506. [PubMed: 9753435]
30. Taylor JC, Markham GD. *J. Biol. Chem* 1999;274:32909–32914. [PubMed: 10551856]
31. Bruice TC. *Acc. Chem. Res* 2002;35:139–148. [PubMed: 11900517]
32. Friesner RA, Guallar V. *Annu. Rev. Phys. Chem* 2005;56:389–427. [PubMed: 15796706]
33. Lin H, Truhlar DG. *Theor. Chem. Accts* 2007;117:185–199.
34. Senn HM, Thiel W. *Curr. Opin. Chem. Biol* 2007;11:182–187. [PubMed: 17307018]
35. Siegbahn PE. *Q Rev. Biophys* 2003;36:91–145. [PubMed: 12643044]

36. Warshel A. *Annu. Rev. Biophys. Biomol. Struct* 2003;32:425–443. [PubMed: 12574064]
37. Zhang X, DeChancie J, Gunaydin H, Chowdry AB, Clemente FR, Smith AJ, Handel TM, Houk KN. *J. Org. Chem* 2008;73:889–899. [PubMed: 18179229]
38. Luzatti PV. *Acta Crystallogr* 1952;B 5:802–810.
39. Altun A, Shaik S, Thiel W. *J. Comput. Chem* 2006;27:1324–1337. [PubMed: 16788908]
40. Bock CW, Katz AK, Markham GD, Glusker JP. *J. Am. Chem. Soc* 1999;121:7360–7372.
41. Markham GD, Glusker JP, Bock CL, Trachtman M, Bock CW. *J. Phys. Chem* 1996;100:3488–3497.
42. Pavlov M, Siegbahn PEM, Sandstrom M. *J. Phys. Chem* 1998;A 102:219–228.
43. Philipp DM, Friesner RA. *J. Comput. Chem* 1999;20:1468–1494.
44. Lee CT, Yang WT, Parr RG. *Phys. Rev* 1988;B 37:785–789.
45. Becke AD. *J. Chem. Phys* 1993;98:5648–5652.
46. Hay PJ, Wadt WR. *J. Chem. Phys* 1985;82:299–310.
47. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. *J Phys Chem* 2001;B 105:6474–6487.
48. Markham GD, Norrby PO, Bock CW. *Biochemistry* 2002;41:7636–7646. [PubMed: 12056895]
49. Glendening ED, JKB, Reed AE, Carpenter JABJE, Morales CM, Weinhold F, et al. NBO 5.0. 2001
50. Rinaldo D, Philipp DM, Lippard SJ, Friesner RA. *J. Am. Chem. Soc* 2007;129:3135–3147. [PubMed: 17326634]
51. Wiberg KB. *Tetrahedron* 1968;24:1083–1096.
52. Hu P, Zhang Y. *J. Am. Chem. Soc* 2006;128:1272–1278. [PubMed: 16433545]
53. Schalk-Hihi C, Markham GD. *Biochemistry* 1999;38:2542–2550. [PubMed: 10029549]
54. Jensen, F. *Introduction to Computational Chemistry*. John Wiley; 1999.
55. Burling FT, Goldstein BM. *Acta Crystallogr* 1993;B 49(Pt 4):738–744.
56. Burling FT, Goldstein BM. *J. Am. Chem. Soc* 1992;114:2313–2320.
57. Lodi PJ, Knowles JR. *Biochemistry* 1991;30:6948–6956. [PubMed: 2069953]
58. Lodi PJ, Knowles JR. *Biochemistry* 1993;32:4338–4343. [PubMed: 8476863]
59. Bruice TC, Benkovic SJ. *Biochemistry* 2000;39:6267–6274. [PubMed: 10828939]
60. Shurki A, Strajbl M, Villa J, Warshel A. *J. Am. Chem. Soc* 2002;124:4097–4107. [PubMed: 11942849]
61. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MH. *Chem. Rev* 2006;106:3210–3235. [PubMed: 16895325]
62. Valiev M, Yang J, Adams JA, Taylor SS, Weare JH. *J Phys Chem* 2007;B 111:13455–13464.
63. Larsen TM, Benning MM, Rayment I, Reed GH. *Biochemistry* 1998;37:6247–6255. [PubMed: 9572839]
64. Matte A, Tari LW, Goldie H, Delbaere LT. *J. Biol. Chem* 1997;272:8105–8108. [PubMed: 9139042]
65. Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. *Nat. Prod. Rep* 2002;19:390–412. [PubMed: 12195810]
66. Bauer CB, Fonseca MV, Holden HM, Thoden JB, Thompson TB, Escalante-Semerena JC, Rayment I. *Biochemistry* 2001;40:361–374. [PubMed: 11148030]
67. Maurice M, Mera PE, Taranto MP, Sesma F, Escalante-Semerena JC, Rayment I. *J. Biol. Chem* 2007;282:2596–2605. [PubMed: 17121823]
68. Soriano A, Castillo R, Christov C, Andres J, Moliner V, Tunon I. *Biochemistry* 2006;45:14917–14925. [PubMed: 17154529]
69. Zhang X, Bruice TC. *Proc. Natl. Acad. Sci. U.S.A* 2006;103:6148–6153. [PubMed: 16606828]
70. Zhang X, Bruice TC. *Proc. Natl. Acad. Sci. U.S.A* 2006;103:16141–16146. [PubMed: 17053070]
71. Zhang X, Bruice TC. *Biochemistry* 2007;46:14838–14844. [PubMed: 18044969]
72. Zhang X, Bruice TC. *Biochemistry* 2008;47:2743–2748. [PubMed: 18260647]
73. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J. Comput. Chem* 2004;25:1605–1612. [PubMed: 15264254]



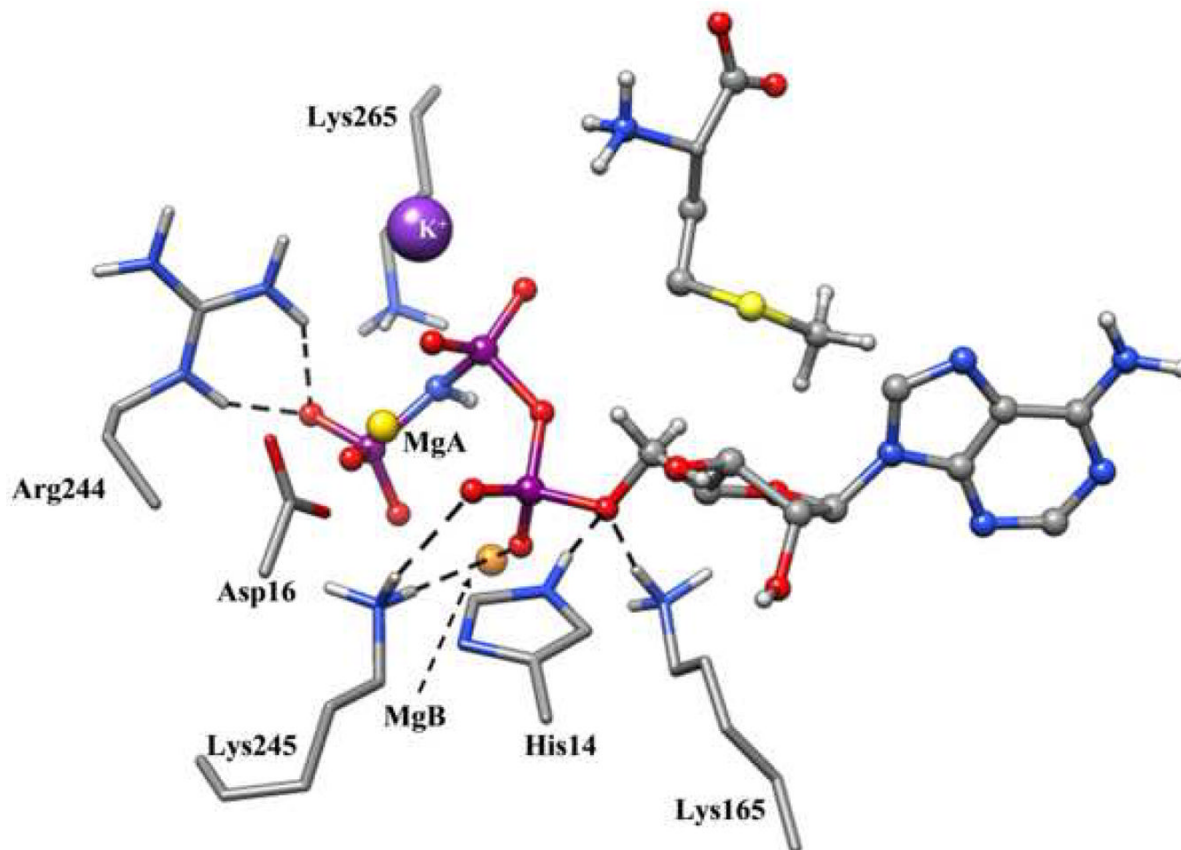


Fig. 1. The segments of the S-adenosylmethionine synthetase active site present in the QM regions of the QM/MM calculations (taken from the crystal structure of E•2Mg²⁺•AMPPNP•methionine•K⁺ complex (pdb file 1P7L) [24]). **Part A** shows the QM/MM energy minimized reactant structure of the Small_QM region, while **Part B** shows the energy minimized structure of the reactant in the Large_QM region. In both parts the protein segments are shown as sticks, and the reactants as ball-and-stick. In Part A water molecules are in wireframe; for clarity water molecules are not shown in Part B, but they occupied all of the same positions in the structure. Metal ions are represented as spheres. Hydrogens attached to carbons are not shown, except those attached to C5' of AMPPNP and the methionine methyl group. In Part A the first coordination sphere interactions of the Mg²⁺ are shown in dashed blue lines. Selected hydrogen bonds are indicated in dashed black lines. This figure was prepared using UCSF Chimera [73].

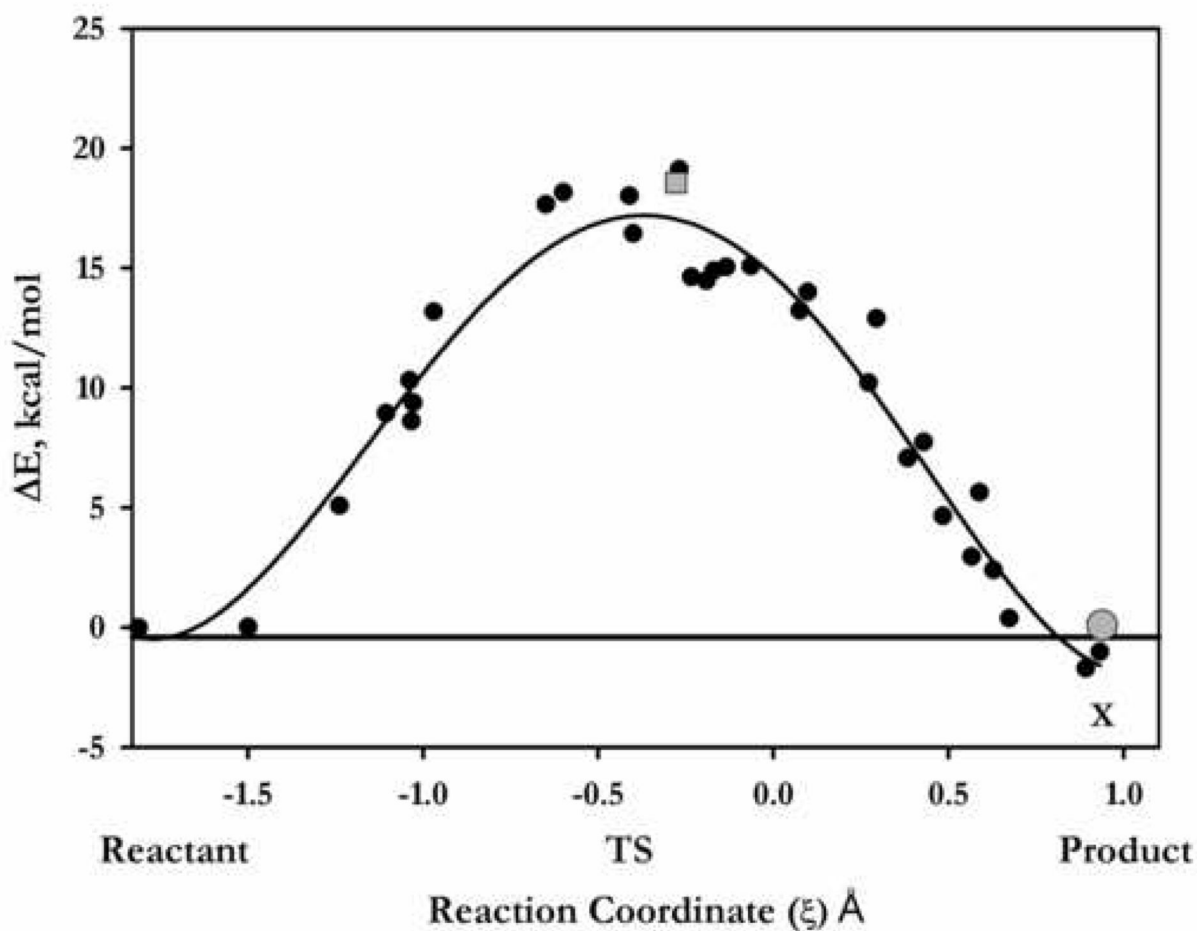
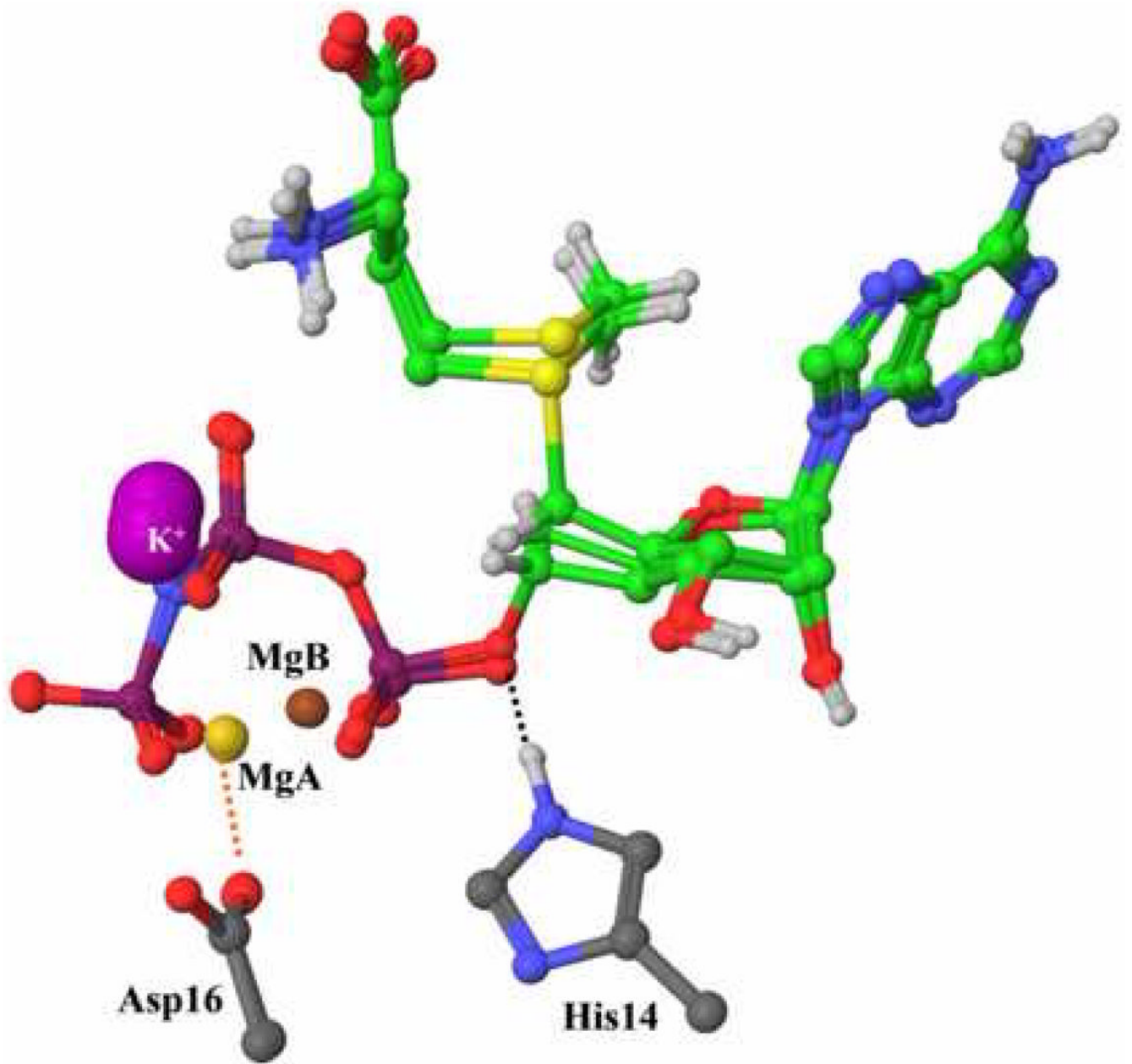


Fig. 2. QM/MM Energy as a function of the reaction coordinate ξ (defined as ((Distance C5' – O5') - (Distance C5' – S))) in Ångströms. The curve derives from a polynomial fit to the data points. It is noteworthy that structures with different combinations of C5' – O5' and C5' – S distances can yield the same value of ξ . The experimental activation energy for the reaction in solution is shown by a ■ located at the reaction coordinate for the transition state that was deduced from kinetic isotope measurements, $\xi = -0.24 \text{ \AA}$ [19]. The experimentally determined relative energies of the product complex in solution and in the crystal are shown by X- and • respectively.



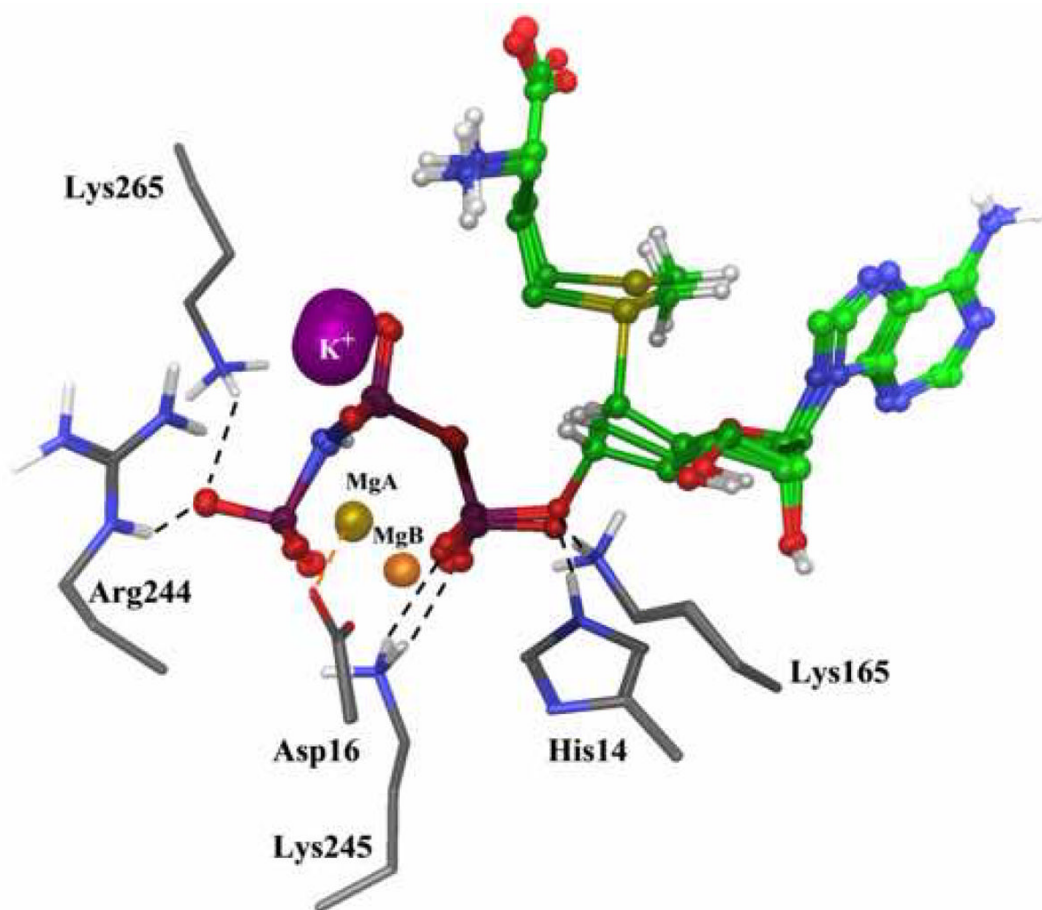
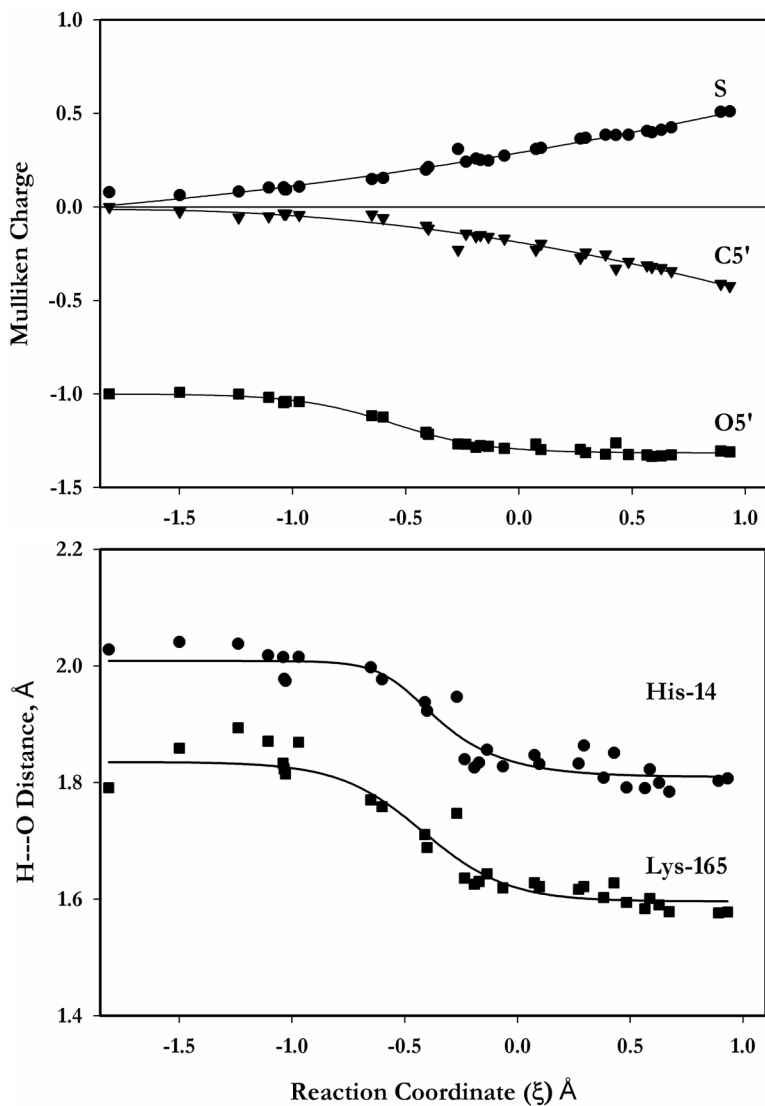
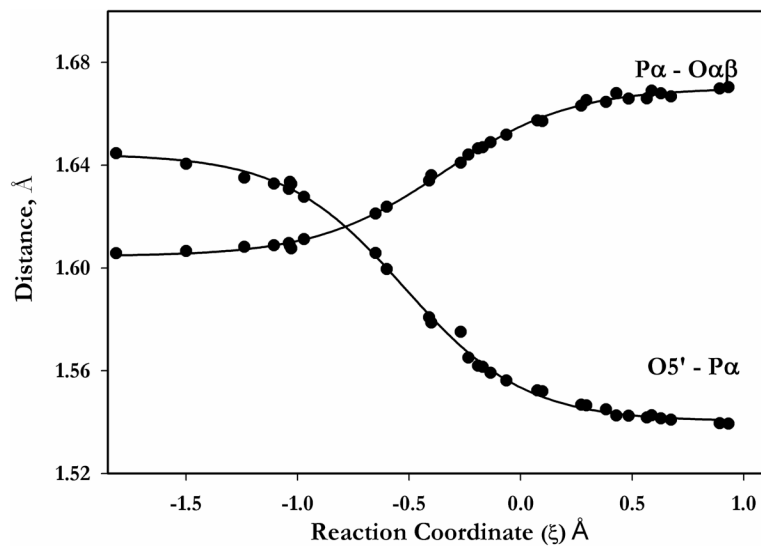
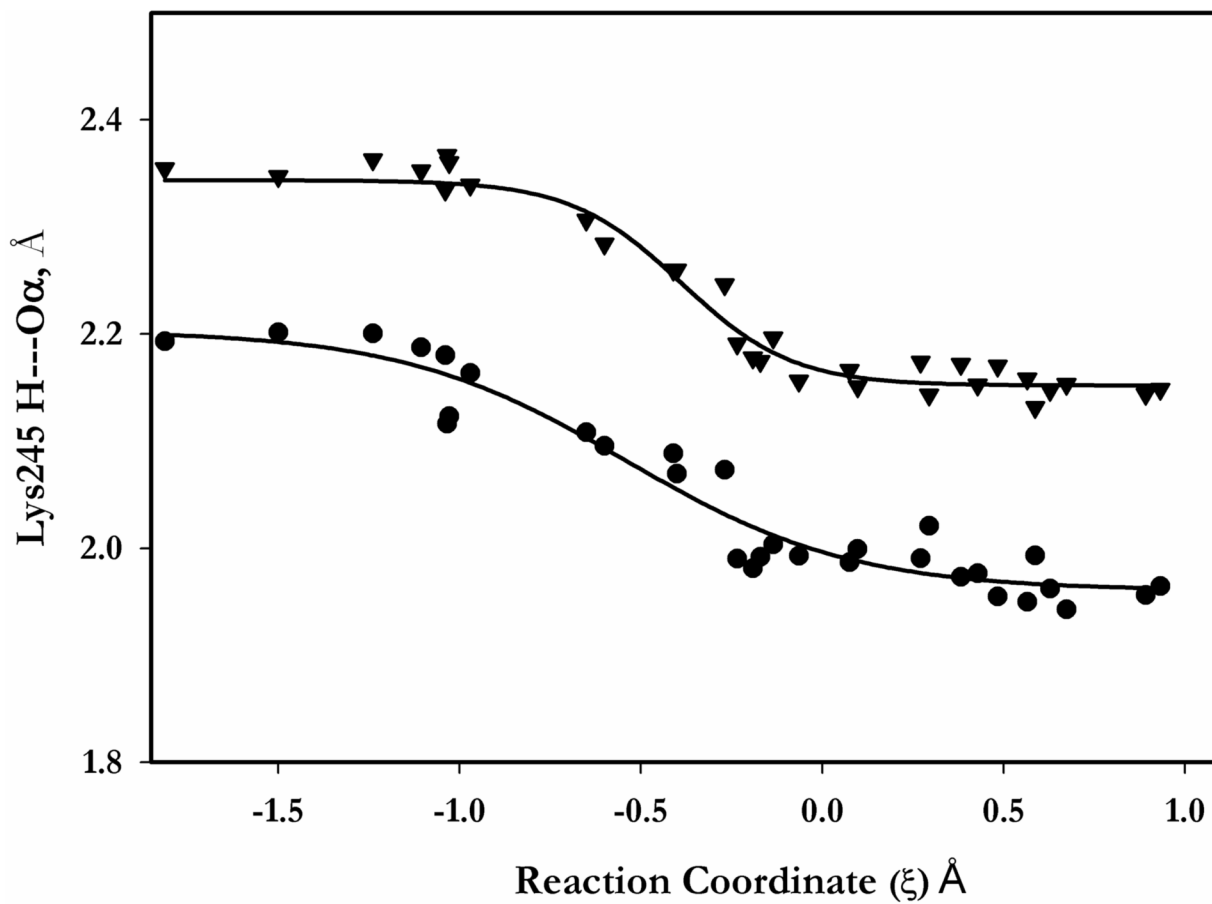


Fig. 3. Superposition of the energy minimized structures of the Small_QM region in **Part A** and the Large_QM region in **Part B**. The reactants and products are shown in ball-and-stick representation and the TS structure in tubes. Protein components are shown in ball-and-stick in Part A and in sticks in Part B. Carbons for the reactants, TS, and products are green while carbons of the protein are grey. Water molecules are not shown, nor are hydrogens attached to carbons, except for H5' and H5'' at the reaction center and the methionine methyl group.





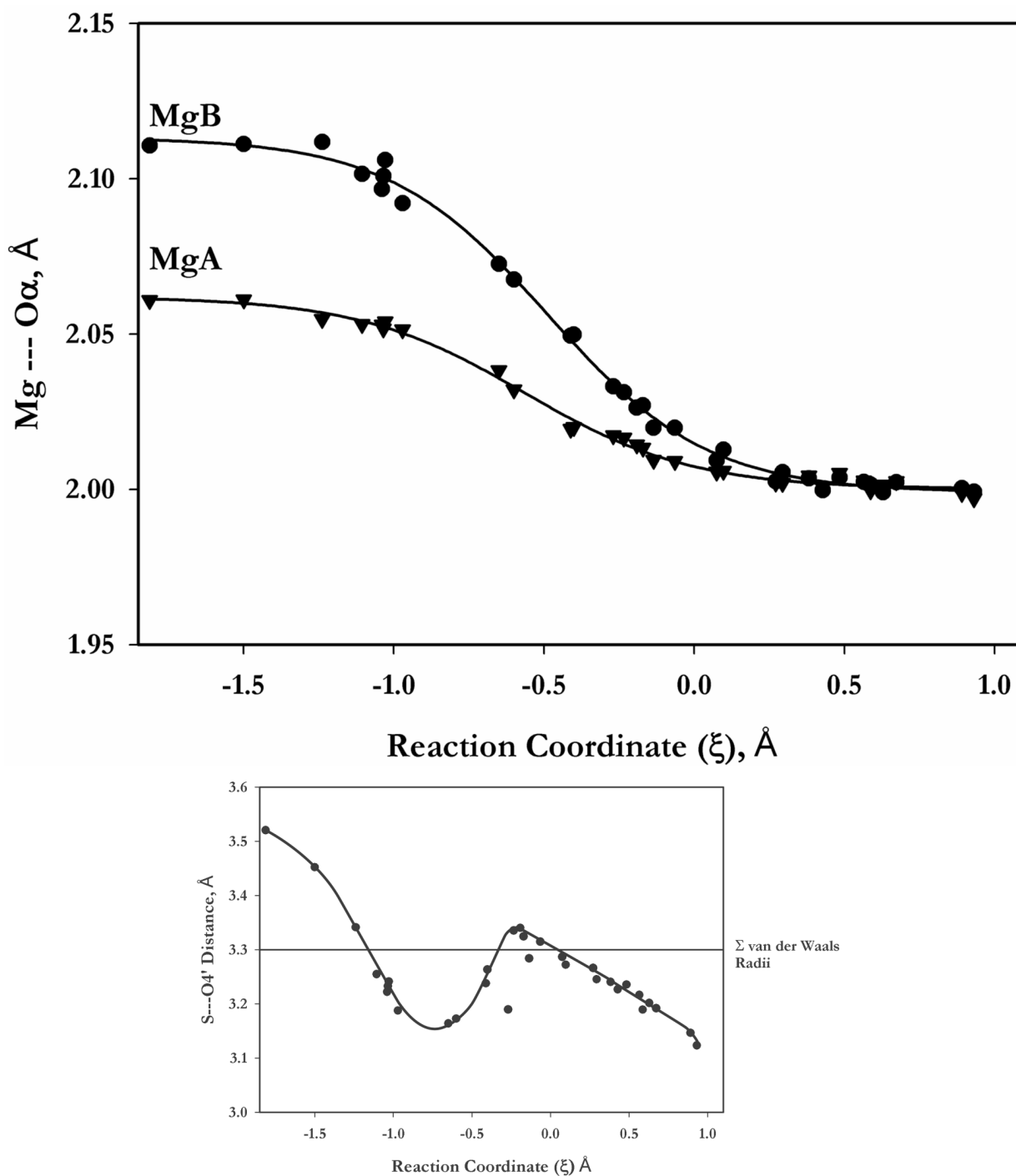


Fig. 4. Changes in geometrical parameters in the Small_QM representation across the reaction coordinate. **Part A** shows the variation in Mulliken charges on the nucleophilic sulfur atom, the central C5' atom, and the departing O5' atom with reaction coordinate. The data for O5' were fit to a sigmoidal curve (the half-maximal change occurs at $\xi = -0.55$ Å); the curves for S and C5' are drawn as visual aids. **Part B** shows the variation in the hydrogen bond lengths from O5' to H ϵ of His-14 (O5'•••H-N ϵ (His-14)), and to H ζ of Lys-165 (O5'•••H-N ζ (Lys-165)).

The curves were obtained from fits of the points to the equation for a sigmoidal curve; the half maximal changes occur at $\xi = -0.39 \text{ \AA}$ and -0.41 \AA , respectively. **Part C** shows the changes in the lengths of the hydrogen bonds between H_{ζ} of Lys-245 to the O_{α} , the two non-bridge oxygens on the α -phosphoryl group ($O_{\alpha} \cdots H-N_{\zeta}$ (Lys-245)). The curves were obtained from fitting the points to the equation for a sigmoidal curve; the half maximal changes occurred at $\xi = -0.54 \text{ \AA}$ (-●-) and -0.39 \AA (-▼-), respectively. **Part D** shows the variation in the lengths of the $O5'-P_{\alpha}$ and $P_{\alpha}-O_{\alpha\beta}$ bonds along the reaction coordinate. The data were fit to a sigmoidal equation; the half maximal changes occur at $\xi = -0.52 \text{ \AA}$ and -0.37 \AA , respectively. **Part E** shows the decrease in the lengths across the reaction coordinate of the bond between each Mg^{2+} ion and a non-bridging oxygen on the α -phosphoryl group. The curves derive from fitting the points to a sigmoidal equation; the half maximal changes occurred for MgA (-●-) at $\xi = -0.56 \text{ \AA}$ and at $\xi = -0.48 \text{ \AA}$ for MgB (-■-). Part F illustrates the variation in the distance between the nucleophilic sulfur atom and $O4'$ in the ribose ring across the reaction coordinate. The sum of the van der Waals radii for sulfur and oxygen atoms is $\sim 3.3 \text{ \AA}$ [55]. The curve is a trend line.

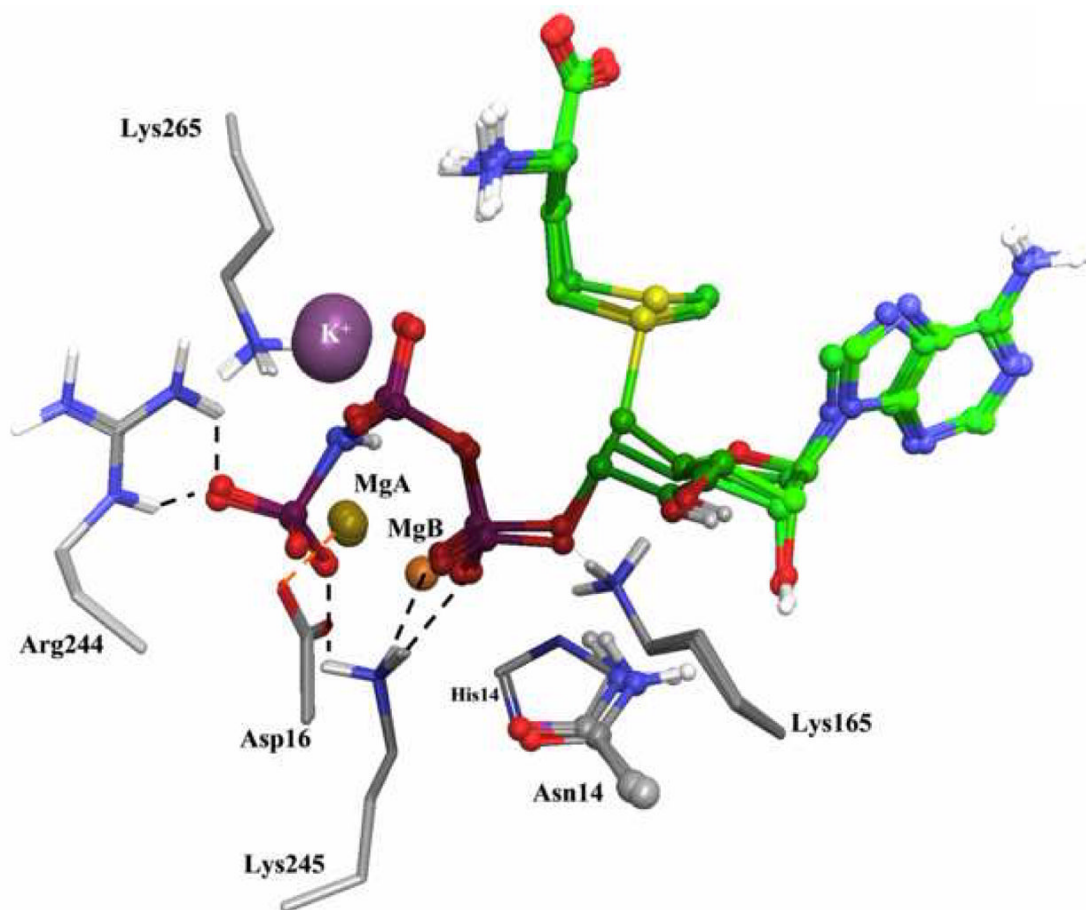


Fig. 5. Superposition of the energy minimized structures of the Large_QM region for the reactant, TS and product of the H14N mutant. The reactants and products are shown in ball-and-stick representation and the TS structure as tubes. Histidine-14 of the wild type enzyme is superimposed in wireframe for comparison; for each reaction coordinate value the rest of the structures are indistinguishable from each other and from the wild type protein. Carbons for the reactants are green, while carbons of the protein are grey. Water molecules are not shown, nor are hydrogen atoms attached to carbons, except for H5 and H5' of AMPPNP and the methionine methyl group.

Table 1

Relative QM/MM Energies (kcal/mol) of Various Structures

Protein in the QM Region ^b	Relative Energy (kcal/mol)			Basis Set ^d		
	Reactant	TS	Product	Charge	Geometry	Energy
Small_QM ^c	0.0	+19.1	-1.8	0	LACVP+**	LACVP+**
Large_QM ^d	0.0	+18.0	+3.0	+4	LACVP*	LACVP*
Large_QM ^d	0.0	+18.8	+1.0	+4	LACVP*	LACVP+**
His-14	0.0	+19.6	-2.2	+1	LACVP+**	LACVP+**
ASP-16	0.0	+20.5	+1.9	0	LACVP+**	LACVP+**
His-14, Lys-165	0.0	+21.0	+2.6	+2	LACVP+**	LACVP+**
No QM Protein Residues ^e	0.0	+21.0	+1.0	+1	LACVP+**	LACVP+**
Large_QM, H14N mutation	0.0	+31.4	+13.2	+4	LACVP*	LACVP*
Large_QM, H14N mutation	0.0	+23.2	+12.1	+4	LACVP*	LACVP+**
Ligands Only ^g	0.0	+40.1	+39.6	+1	LACVP*	LACVP*
Ligands Only ^g	0.0	+37.0	+33.2	+1	LACVP*	LACVP+**

^aThe LACVP basis set uses the Pople 6-31G basis set for atoms except potassium for which a Los Alamos pseudopotential is used [46].

^bThe QM regions in all calculations contained AMPPNP and methionine (or AdoMet and PPNP), 2 Mg²⁺, K⁺, and 11 water molecules.

^cThe Small_QM region had the side chains of Histidine-14 and Aspartate-16 in the QM region.

^dThe Large_QM region had the side chains of Histidine-14, Aspartate-16, Lysine-165, Arginine-244, Lysine-245 and Lysine-265 in the QM region.

^eAll protein components were in the MM region.

^fThe mutated Large_QM region contained the side chains of Asparagine-14, Aspartate-16, Lysine-165, Arginine-244, Lysine-245 and Lysine-265.

^gAll protein components were deleted. Atomic positions were maintained at the LACVP* optimized geometry of the Large_QM region.