

**Catalytic Mechanism of Mevalonate Kinase Revisited, a  
QM/MM Study**

Journal:	<i>Organic &amp; Biomolecular Chemistry</i>
Manuscript ID	OB-ART-12-2018-003197
Article Type:	Paper
Date Submitted by the Author:	25-Dec-2018
Complete List of Authors:	McClory, James; Queen's University Belfast School of Chemistry and Chemical Engineering Lin, Jun-Tang; Xinxiang Medical University Zhang, Jian; Shanghai Jiaotong university, Timson, David; University of Brighton, School of Pharmacy and Biomolecular Sciences Huang, M; Queen's University Belfast,



## Catalytic Mechanism of Mevalonate Kinase Revisited, a QM/MM Study

James McClory,<sup>a</sup> Jun-Tang Lin,<sup>b</sup> David Timson,<sup>c</sup> Jian Zhang<sup>d</sup> and Meilan Huang<sup>\*a</sup>

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Mevalonate Kinase (MVK) catalyses the ATP-Mg<sup>2+</sup> mediated phosphate transfer of mevalonate to produce mevalonate 5-phosphate and is a key kinase in the mevalonate pathway in the biosynthesis of isopentenyl diphosphate, the precursor of isoprenoid-based biofuels. However, the crystal structure in complex with the native substrate mevalonate, ATP and Mg<sup>2+</sup> has not been resolved, which has limited the understanding of its reaction mechanism and therefore its application in the production of isoprenoid-based biofuels. Here using molecular docking, molecular dynamic (MD) simulations and a hybrid QM/MM study, we revisited the location of Mg<sup>2+</sup> resolved in the crystal structure of MVK and determined a catalytically competent MVK structure in complex with the native substrate mevalonate and ATP.

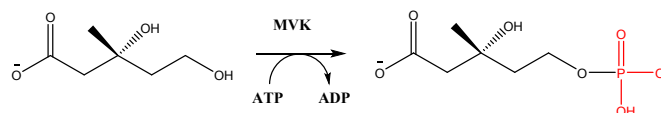
We demonstrated that significant conformational change on a flexible loop connecting the α6 and α7 helix is induced by the substrate binding. Further, we found that Asp204 is coordinated to the Mg<sup>2+</sup> ion. Arg241 plays a crucial role in organizing the triphosphoryl tail of ATP for in-line phosphate transfer and stabilizing the negative charge that accumulates at the β,γ-bridging oxygen of ATP upon bond cleavage. Remarkably, we revealed that the phosphorylation of mevalonate catalyzed by MVK occurs via a direct phosphorylation mechanism, instead of the conventionally postulated catalytic base mechanism. The catalytically competent complex structure of MVK as well as the mechanism of reaction will pave the way for the rational engineering of MVK to exploit its applications in the production of biofuels.

**Keywords:** GHMP kinase; reaction mechanism; biocatalysis; phosphorylation; isoprenoid biosynthesis; QM/MM

### 1 Introduction

Mevalonate kinase (MVK; EC 2.7.1.36) is a key kinase in the mevalonate (MVA) pathway. It catalyses the Mg<sup>2+</sup>-ATP dependent phosphorylation of mevalonate to mevalonate-5-phosphate (Figure 1), a crucial step in the mevalonate pathway for the production of isopentenyl diphosphate (IPP) and its stereoisomer dimethylallyl diphosphate (DMAPP). IPP is used in production of isoprenoids, the important precursors to five-carbon alcohols and as the initial precursors to longer chain terpenes, therefore, MVK has become an attractive biocatalyst for the production of environmentally benign biofuels<sup>1-2</sup>. On the other hand, mutations of MVK causes decreased prenylation of

associated with a range of metabolic autoinflammatory diseases such as mevalonic aciduria (MA, MIM# 610377)<sup>6</sup> and hyper-immunoglobulin D syndrome (HIDS, MIM# 260920)<sup>7-8</sup>. MKD is a rare, orphan and neglected disease and so far, there is no effective treatment for it. Understanding the catalytic mechanism of the MVK would be helpful for exploiting the potential biotechnological applications and developing possible pharmacological interventions for the MKD-associated disease. However, lack of complete kinase structure in complex with the substrate mevalonate and unambiguous position/ coordination of Mg<sup>2+</sup> ion has limited the understanding of the phosphorylation mechanism and broader biotechnological applications of the kinase.



**Fig. 1** Phosphorylation of mevalonate to mevalonate-5-phosphate catalysed by MVK.

MVK belongs to GHMP superfamily, which is named after the four originally identified members, Galactokinase (GALK; EC

<sup>a</sup> School of Chemistry and Chemical Engineering, Queen's University Belfast, David Keir Building, Stranmillis Road, Belfast, BT9 5AG, Northern Ireland, United Kingdom

<sup>b</sup> Stem Cells and Biotherapy Engineering Research Centre of Henan, College of Biomedical Engineering, Xinxiang Medical University, Xinxiang 453003, China

<sup>c</sup> School of Pharmacy and Biomolecular Sciences, The University of Brighton, Huxley Building, Lewes Road, Brighton, BN2 4GJ, United Kingdom

<sup>d</sup> Shanghai Jiaotong University, 280 Chongqing Road, Shanghai 200025, China

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

GTPases<sup>3</sup> and mevalonate kinase deficiency (MKD)<sup>4-5</sup>, which is

2.7.1.6), Homoserine kinase (HSK; EC 2.7.1.39), MVK and Phosphomevalonate kinase (PMK; EC 2.7.4.2)<sup>9</sup> and has since been expanded to include additional members such as N-acetylgalactosamine kinase (GALK2; EC 2.7.1.157)<sup>10-11</sup>, galacturonic acid kinase (GalAK; EC 2.7.1.43)<sup>12</sup>, arabinose kinase (EC 2.7.1.46)<sup>13</sup> and pantoate kinase (EC 2.7.1.169)<sup>14</sup>. The family is characterised by three highly conserved motifs (I-III). Motif I is involved in substrate binding, motif II is a glycine-rich loop <sup>138</sup>PPGAGLGSSA<sup>147</sup> engaged in nucleotide binding and motif III is involved in the phosphorylation reaction<sup>9</sup>.

Two mechanisms have been proposed for the GHMP kinases. The first is the catalytic base mechanism proposed for MVK, GALK and MDD whereby phosphate transfer was proposed to be preceded by the abstraction of proton from the substrate hydroxyl group by a negatively charged residue that functions as the catalytic base<sup>15-19</sup>. The second plausible mechanism was proposed for HSK in which phosphorylation of the substrate occurs in the absence of a potential catalytic base thus the substrate hydroxyl group directly attacks the  $\gamma$ -phosphate of ATP<sup>20</sup>. Moreover, even if there is a potential catalytic base in the active site, the kinase may not necessarily undergo a base-assisted phosphorylation. *e.g.* it was suggested that Asp190 in the active site of GALK2 is too far from the substrate to function as a catalytic base<sup>21</sup>. Recently, QM/MM studies revealed that the reactions catalyzed by GALK and PMK occur via direct phosphorylation mechanism<sup>22-23</sup>.

Although several different crystal structures of mevalonate kinase have been resolved, the position of Mg<sup>2+</sup> remains elusive, along with the location of its native substrate mevalonate. This results in great uncertainty surrounding the reaction mechanism of MVK. Here we report a catalytically competent structure of the ATP-Mg<sup>2+</sup>-MVK-MVA complex and reveal the reaction mechanism of MVK by using molecular docking, molecular dynamic simulations and a combined QM/MM study. These findings would provide the basis for the rational engineering of MVK for biosynthesis of biofuels and provide valuable insights for development of potential therapies for MKD-related diseases.

## 2 Theoretical Methods

### 2.1 Protein Preparation

The crystal structure of *Rattus norvegicus* MVK in complex with ATP (PDB code: 1KVK)<sup>24</sup> with a resolution of 2.4 Å was used to build the complex model of MVK. Magnesium together with three coordinating waters was placed beside the  $\gamma$ -phosphate of ATP referring to the magnesium coordination in *Streptococcus pneumoniae* PMK, a homologous kinase in GHMP family, for which the location of Mg<sup>2+</sup> ion has been explicitly resolved<sup>25</sup>. The missing loop region between residues 73-88 was rebuilt using Modeller v9.17<sup>26</sup>. The protonation states of the titratable residues were decided

using the H++ webserver<sup>27</sup> and by visual inspection (Table S1). Asn/Gln/His flips were checked using Reduce encoded in AmberTools<sup>28</sup>.

### 2.2 Molecular Docking

The mevalonate substrate was docked in the MVK structure using AutoDock 4.2<sup>29</sup>. Polar hydrogen atoms were added. Kollman charges were applied to the protein while RESP charges were utilized for the ligand. A grid box was set to cover the catalytic residues proposed in previous kinetic studies. The Lamarckian genetic algorithm (LGA) was applied with a GA population size of 100 and a maximum number of 2,500,000 energy evaluations. The docked poses were selected for the subsequent MD simulations based on the docking score and visual inspection of the binding mode.

### 2.3 Molecular Dynamics Simulations

The AMBER 14 (University of California, USA) package was used for Molecular Dynamic Simulations<sup>30</sup>. All MD simulations were performed with the FF14SB forcefield<sup>31</sup>. Single point charge calculations were run for the substrate and nucleotide using the Hartree-Fock (HF) method in conjunction with the 6-31G(d) basis set implemented in Gaussian 09 package<sup>32</sup> to obtain the RESP charges of the ligands (Table S2). The GAFF forcefield<sup>33</sup> was used for the ligands. The ATP-Mg<sup>2+</sup>-MVK-MVA complex was then soaked in a TIP3P water box within a minimum distance of 10 Å to the boundary of the protein. A total of 11,739 waters were added to the complex and Na<sup>+</sup> counterions were used to maintain the neutrality of the system.

An initial minimization was performed exclusively for the solute involving 1,250 steps of steepest descent minimization followed by 1,250 steps of conjugate gradient minimization, where the nucleotide, substrate, the magnesium ion and the three structural waters were restrained with a force constant of 50 kcal mol<sup>-1</sup>Å<sup>-2</sup>. A subsequent minimization was conducted with no restraint force applied to the entire protein system.

The system was then heated slowly from 0 to 300 K for 250 ps using the Langevin dynamics with a collision frequency of 1.0 ps<sup>-1</sup>. An equilibration simulation was run with a NVT ensemble and a periodic boundary condition for a further 50 ps at 300 K. A cutoff distance of 8 Å was set for non-bonded Van der Waals force while the Particle Mesh Ewald method was used to account for the long-range electrostatic interactions. The SHAKE method was utilized to fix the covalent bonds associated with the hydrogen atoms within the system<sup>34</sup>. A 100-ns production simulation with a time step of 1 fs was performed for the ATP-Mg<sup>2+</sup>-MVK-MVA complex utilizing the NPT ensemble at 300 K and a pressure of 1 atm. Four replicas of MD simulations were run for the wild-type kinase and the variants with randomly assigned initial velocities. Root mean square deviation (RMSD) of the C $\alpha$  atoms was calculated (Figure S1) using the initially minimized crystal structure as the reference.

Cluster analysis was carried out using the AMBER tool CPPTRAJ to attain representative structures from the simulation trajectory.

#### 2.4 QM/MM Calculations

The representative structures of the ATP-Mg<sup>2+</sup>-MVK-MVA complex were obtained by cluster analysis of the MD simulation trajectory and then were used for the subsequent QM/MM study. All QM/MM calculations were performed using the ONIOM protocol implemented in Gaussian09. The Tao package was used to prepare the structures for QM/MM calculation<sup>35</sup>. The system was divided into two layers, the QM region, which was treated with the b97d functional<sup>36</sup> accounting for Grimme's dispersion effect in conjunction with a 6-31g(d) basis set and the MM region, which was treated with Amber Parm99 force field. An electronic embedding scheme was incorporated to integrate the point charge of the MM layer into the QM layer<sup>37</sup>. The quadratically convergent Hartree-Fock (QC-SCF) method was used to speed up convergence. The residues within a radius of 6 Å of the active site were allowed to move freely while the rest residues were kept frozen.

The QM region contains the divalent metal Mg<sup>2+</sup> ion, the side chains of Lys13, Ser146, Asp204 and Arg241, the truncated triphosphoryl tail of ATP along with the connecting methyl group and a total of seven water molecules. It was composed of a total of 45 heavy atoms with a net charge of -2. Charges used for mevalonate and ATP in the MD simulations were employed for the QM/MM calculations.

Representative structures from the MD trajectory were subjected to an initial geometry optimization followed by potential energy scan (PES) with a decrement of 0.1 Å for the distance between the  $\gamma$ -phosphorous atom of ATP and the C5-hydroxyl oxygen of mevalonate. The transition state (TS) structure attained from the PES was optimized as TS and validated using frequency calculation. The energies of the stationary points were corrected with the 6-31++g(d,p) basis set (Table S3).

## 3 Results and Discussions

### 3.1 The catalytically competent model of ATP-Mg<sup>2+</sup>-MVK-MVA

Among the GHMP kinases that are involved in the mevalonate pathway, four kinases have their crystal structures be resolved so far, namely MVK, MDD, M3K and PMK. However, except for PMK, the location of Mg<sup>2+</sup> has not been defined in other three kinases. Although several crystal structures of MVK have been reported, none of these structures were fully resolved with the native substrate mevalonate, the nucleotide ATP and the divalent metal Mg<sup>2+</sup> ion thus leaving ambiguity surrounding the reaction mechanism of MVK. In order to obtain the adopted conformation of the substrate MVA in presence of Mg<sup>2+</sup> and ATP, MVA was docked into the MVK structure using Autodock 4.2. An initial catalytically competent conformation was chosen based on the docking scores, as well as the feasibility of the phosphorylation in view of the reaction coordinates, *i.e.* the

distance between the  $\gamma$ -phosphate of ATP and the C5-hydroxyl of mevalonate (Figure S2).

#### 3.1.1 Mg<sup>2+</sup> coordination in MVK

In the crystal structure of *Rattus norvegicus* MVK (PDB Code: 1KVK)<sup>24</sup>, a Mg<sup>2+</sup> ion is coordinated with the  $\beta$ - and  $\gamma$ - phosphates of ATP. However, it should be noted that the distance between the metal ion and phosphate oxygen of ATP is longer than a standard coordinate bond, and it is actually analogous to that of a hydrogen bond with a water molecule. Glu193 was postulated to be coordinated to the Mg<sup>2+</sup> ion<sup>18,24</sup>. However, the distance between Glu193 and the Mg<sup>2+</sup> ion (2.8 Å) is also larger than a standard coordinate bond. Therefore, in order to elucidate the phosphorylation mechanism of the kinase, it is necessary to revisit the location/coordination of Mg<sup>2+</sup> to decide a catalytically competent complex of MVK.

We first attempted to build a MVK model with Mg<sup>2+</sup> coordinated to the  $\beta$ - and  $\gamma$ -phosphate oxygen atoms of ATP, which was analogous to the suspicious Mg<sup>2+</sup> coordination in the crystal structure of *Rattus norvegicus* MVK (PDB Code: 1KVK)<sup>24</sup>. However, MD simulations revealed that with time evolution, the substrate and ATP moved considerably far away from each other, resulting in a non-catalytic model of MVK (Figure S3), indicating the crystal structure represents a catalytically inactive structure.

In order to determine the catalytically competent configuration of MVK complex, a second MVK model was built where a Mg<sup>2+</sup> ion was placed to be coordinated to the  $\gamma$ -phosphate of ATP, analogous to the Mg<sup>2+</sup> coordination in *S. pneumoniae* PMK (PDB Code: 3GON)<sup>25</sup>, a GHMP kinase homologous to MVK. Further, three coordinating water molecules taken from *S. pneumoniae* PMK were retained to coordinate with the Mg<sup>2+</sup> in MVK. Further, the Mg<sup>2+</sup> ion was positioned to be coordinated with the hydroxyl group of the MVA substrate, by analogy to PMK, where the metal ion is coordinated to the hydroxyl group of the substrate phosphomevalonate (Note the coordination was replaced by a fourth coordinating water during the subsequent QM/MM optimization). In PMK Mg<sup>2+</sup> is coordinated to the conserved aspartate Asp297<sup>25</sup>, which was previously proposed to function as a catalytic base<sup>18</sup>. Interestingly, in the resulting ATP-Mg<sup>2+</sup>-MVK-MVA complex model, the hexa-coordination of Mg<sup>2+</sup> ion is also saturated by the conserved Asp204.

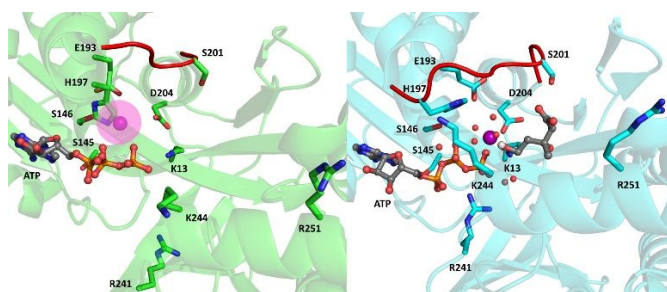
The MD simulations showed that the distance between the  $\gamma$ -phosphate of ATP and the hydroxyl group of MVA is maintained between 3.0 Å to 4.0 Å (Figure S4), indicating that the simulated structures represent the catalytically competent configurations with the mevalonate substrate positioned in a reasonable binding pose in relation to ATP in favor of phosphate transfer. Asp204 forms a salt bridge interaction with Lys13, which in turn interacts with the  $\gamma$ -phosphate of ATP (Figure 2 and Figure S5). This is consistent with the configuration observed in the crystal structure of *Rattus norvegicus* MVK<sup>24</sup>. It should be noted that in the crystal structure of *Rattus norvegicus* MVK<sup>24</sup>, Asp204 is located 3.4 Å away from the postulated Mg<sup>2+</sup> ion, obviously longer than a standard coordinate bond.

## ARTICLE

## Organic &amp; Biomolecular Chemistry

However, the coordinate bond between Asp204 and  $Mg^{2+}$  was retained in the ATP- $Mg^{2+}$ -MVK-MVA complex throughout the MD simulations (Figure S6).

Remarkably, the MD simulations of the complete ATP- $Mg^{2+}$ -MVK-MVA complex disclosed that Glu193 is not directly involved in  $Mg^{2+}$  coordination, instead, it forms H-bond interactions with two coordinating water molecules (WAT441 and WAT461). Thus the Hydrogen bonds between Glu193 and the two structural water molecules assist in maintaining an ordered octahedral geometry centred on the metal ion (Figure S7). Notably, in one of the replicas of the MD simulations (Replica 1), Glu193 forms a hydrogen bond with His197 on the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix (Figure 2b and S8); the backbone of His197 is in turn engaged in a H-bond interaction with Lys244. Thus, His197 may assist to organize the active site via the Hydrogen bond with Glu193, rather than help to order the nucleotide as previously suggested<sup>24</sup>.



**Fig. 2** (a) Elusive location of  $Mg^{2+}$  in the crystal structure of *Rattus norvegicus* MVK in absence of mevalonate (pdb code: 1KVK), (b) Representative MD cluster structure of ATP- $Mg^{2+}$ -MVK-MVA complex.

### 3.1.2 Ser201 involved in substrate binding

Kinetic and site-directed mutagenesis studies of the invariant serine/threonine residue in over twenty MVK sequences disclosed that a S201A mutation in human MVK resulted in a 100-fold increase in  $K_m$  while the turnover rate was relatively unaffected. Thus this residue was implicated in substrate binding rather than catalytic activity<sup>38</sup>. Interestingly, the MD simulations of ATP- $Mg^{2+}$ -MVK-MVA complex reveals that as time evolves the hydroxyl side chain of Ser201 forms a stable hydrogen bond with the carboxylate group of MVA (Figure 2 & S9a). Thus Ser201 is actually involved in substrate binding, in accordance with previous kinetic study<sup>38</sup>.

### 3.1.3 The invariant tandem serine residues on P-loop rigidifies ATP

Ser145 and Ser146 are two tandem serine residues highly conserved among GHMP kinases (Figure 2). Ser146 was suggested to help to orientate ATP, based on large effect on  $k_{cat}$  as well as reduced binding affinity associated with S146A mutation<sup>38</sup>. In the crystal structure of *R. norvegicus* MVK, the hydroxyl side chains of both Ser145 and Ser146 form hydrogen bond with the  $\beta$ -phosphate of ATP. With time evolution of the MD simulations, Ser146 maintains a hydrogen bond with the  $\beta$ -phosphate of ATP (Figure 2 & S9b). Ser145 also maintains

a hydrogen bond with the  $\beta$ -phosphate of ATP (Figure 2 & S9c). Thus the two tandem serine residues Ser145 and Ser146 help to rigidify the flexible tail of ATP, facilitating the inline phosphate transfer.

The conserved serine residues Ser145 and Ser146 in MVK are located on a glycine-rich loop comprised of the sequence PXXGLXSSA (Figure S10), which is the characteristic feature of GHMP kinases<sup>39</sup>. Actually the interactions of these tandem serine residues on the flexible P-loop with triphosphoryl tail of ATP have been also reported for other GHMP kinases, e.g. *Thermoplasma acidophilum* M3K (Ser107 and Ser108), *Staphylococcus epidermidis* MDD (Ser106 and Ser107) and *Streptococcus pneumoniae* PMK (Ser106 and Ser107). Thus, it is evident that these conserved serine residues are crucial in arranging the mobile phosphate tail of ATP to facilitate the phosphate transfer in GHMP kinases.

### 3.1.4 Arg241 organizes ATP for in-line phosphate transfer

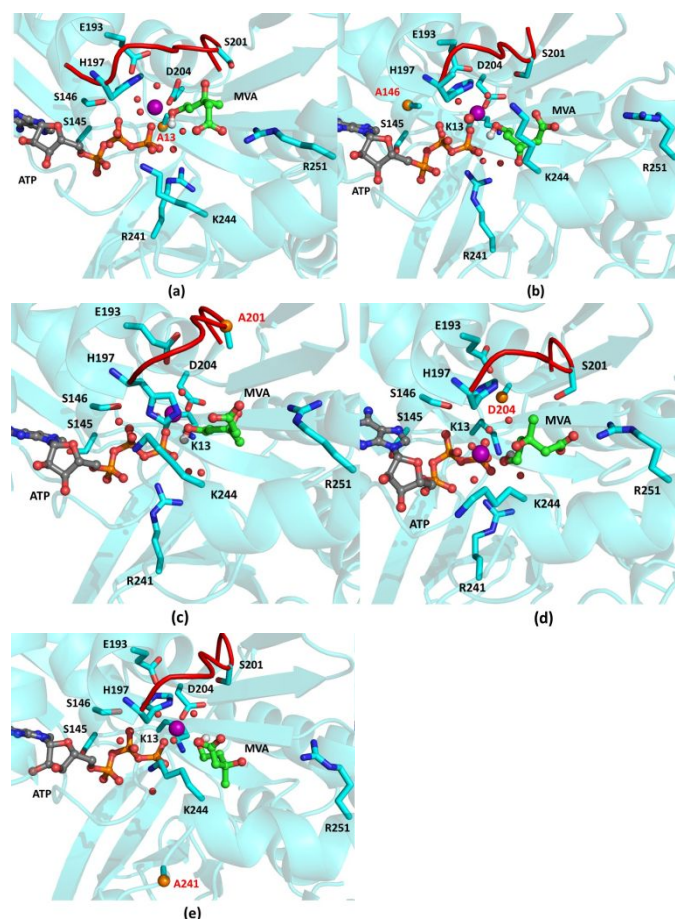
In the crystal structure of *Rattus norvegicus* MVK, Arg241 located at the loop (238-246) preceding the  $\alpha 8$  helix is away from the active site. In the ATP- $Mg^{2+}$ -MVK-MVA complex model, the carboxylate group of the docked MVA was positioned in line with the guanidinium group of Arg241 establishing an ionic interaction with the substrate. Interestingly, with time evolution of the MD simulation, Arg241 moved gradually towards the highly negatively charged triphosphate tail of ATP, forming an ionic interaction with the  $\alpha$ - and  $\gamma$ -phosphate of ATP (Figure S11), whereas the carboxylate group of MVA forms hydrogen bond interaction with Ser201 (Figure 2). Thus Arg241 plays a predominant role in organizing the nucleotide for an in-line phosphate transfer.

### 3.1.5 Effect of mutation of the key residues in MVK

In order to validate the role of the key residues in the active site of MVK identified from MD simulations, these residues were mutated using the mutation feature in Accelrys Discovery studio and the resulted mutants were subjected to MD simulations.

In the wild-type enzyme (Figure 2b), Lys13 forms an ionic interaction with Asp204. MD simulations of the K13A mutant (Figure 3a) reveal that lack of this interaction in the mutant abolished the hydrogen bond interaction between Ser201 and the carboxylate group of the substrate. As a result, mevalonate adopts a significantly different 'bent' conformation, with its carboxylate group of mevalonate forming an ionic interaction with Arg251. Interestingly, a notable conformational change of the loop preceding the  $\alpha 8$  helix is observed causing the displacement of the  $\alpha 8$  helix, as a result, the catalytic residue Arg241 moves away from the catalytic centre whereas Lys244 move in the proximity of the  $\beta$ -,  $\gamma$ -phosphate of ATP, nested in a similar location to where Arg241 is in the WT kinase. Thus this loop may play an auxiliary role in orientating the catalytic Arg241 around the ATP phosphate tail so as to facilitate the phosphate transfer.





**Fig 3.** MD simulated structures of MVK mutants (a) K13A, (b) S146A, (c) S201A, (d) D204A, (e) R241A.

Previous kinetic study showed a S146A mutation resulted in significantly reduced  $k_{\text{cat}}$ <sup>38</sup>. Our MD simulations showed that mutating of Ser146 on the glycine-rich loop into alanine causes conformational change around the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix, and therefore the hydrogen bond interaction between the carboxylate group of mevalonate and Ser201 is abolished (Figure 3b). The loss of the interaction between the substrate mevalonate and S201 mutation due to the S146A mutation may account for the largely reduced activity observed in the previous kinetic study<sup>38</sup>. It also indicates that Ser146 on the glycine-rich loop is critical in maintaining the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix to hold the mevalonate substrate in the catalytic site, so as to facilitate the catalysis of the phosphorylation.

MD simulations of S201A caused the loop linking the  $\alpha 6$  and  $\alpha 7$  helix to undergo a remarkable conformational shift such that the hydrogen bond formed between His197 on this loop and Glu193 is lost, and His197 forms a H-bond with the carboxylate of mevalonate instead. The loss of the interaction between the substrate carboxylate and Ser201 due to the S201A mutation is compensated by an ionic interaction with Arg251 (Figure 3c). Thus, compared to the WT kinase, the mutant displays a less organized catalytic site analogous to the configuration of MVK in absence of the substrate (Figure 2a).

In D204A, the coordination between  $\text{Mg}^{2+}$  and of the aspartate is lost due to lack of the carboxylate side chain, such that the magnesium ion adopts an alternative position at an opposite plane of the ATP triphosphoryl tail and the coordination with the  $\beta$ - and  $\gamma$ - phosphates of ATP is established. The movement of the metal ion caused significant displacement of mevalonate, resulting in an unfavourable orientation for phosphate transfer to occur (Figure 3d). Furthermore, the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix also undergoes notable change such that the interaction between mevalonate and Ser201 is abolished and also replaced by the ionic interaction with Arg251.

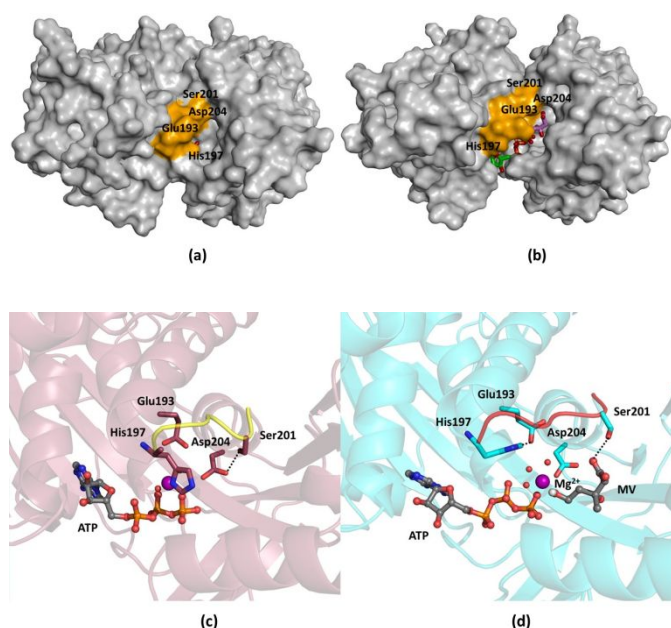
MD simulations disclosed the R241A mutation induced conformational change of the loop proceeding the  $\alpha 8$  helix and the displacement of the  $\alpha 8$  helix so that Lys244 to move toward the active site and nest at an almost identical position to Arg241 in the wild-type MVK, helping to hold the triphosphoryl moiety of ATP in the catalytic site (Figure 3e). Significant conformational change is also observed at the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix such that the interaction between the carboxylate group of the mevalonate and Ser201 is lost. As a result, the substrate adopts a catalytically unfavourable orientation in relation to ATP.

In summary, MD simulations of the MVK variants validated the role of the key residues in the catalytic reaction; The loop connecting  $\alpha 6$  and  $\alpha 7$  helix is crucial in positioning mevalonate substrate in favourable orientation in relation to ATP, and the loop proceeding the  $\alpha 8$  helix plays an auxiliary role in orientating the Arg251 toward the phosphate tail of ATP, facilitating the phosphorylation reaction to happen.

### 3.2 Conformational change during substrate binding

Molecular dynamics simulations were conducted for ATP- $\text{Mg}^{2+}$ -MVK to elucidate the conformational changes in MVK induced by substrate binding. In the absence of the substrate mevalonate, a concave opening in the surface of the protein created an entrance to the active site of ATP- $\text{Mg}^{2+}$ -MVK, forming an 'open' conformation (Figure 4a). In contrast, while the substrate mevalonate is bound, this concave opening is pinched together forming a narrow active site such that the substrate is locked into this catalytically active state in the 'closed' conformation (Figure 4b).

The significant conformational change of the active site is mainly mediated by the loop connecting the  $\alpha 6$  and  $\alpha 7$  helix. In absence of the substrate, Ser201 forms an interaction with Asp204 (Figure 4c). In the substrate bound complex, Ser201 on the loop forms a hydrogen bond interaction with the carboxylate group of mevalonate, which results in a remarkable conformational change in the loop (Figure 4d). In addition, in the ATP- $\text{Mg}^{2+}$ -MVK structure, His197 is hydrogen-bonded with ATP (Figure 4c), whereas in presence of the substrate, His197 undergoes a notable change, forming hydrogen bond interactions with Glu193 and the water molecules coordinated to the metal ion.



**Fig. 4** Open and Closed conformation of MVK during substrate binding. (a) Surface representation of MD simulated ATP-Mg<sup>2+</sup>-MVK complex, involved residues coloured in orange; (b) Surface representation of MD simulated ATP-Mg<sup>2+</sup>-MVK-MVA complex, involved residues coloured in orange; (c) Secondary Structure of MD simulated ATP-Mg<sup>2+</sup>-MVK complex with loop coloured yellow; (d) Secondary Structure of MD simulated ATP-Mg<sup>2+</sup>-MVK-MVA complex with loop coloured red.

*S. pneumoniae* MVK was reported to be able to catalyze the phosphorylation of mevalonate diphosphate, the substrate of its GHMP fellow mevalonate pathway enzyme, mevalonate diphosphate decarboxylase (MDD)<sup>41</sup>. Thus the binding mode of mevalonate in MVK elucidated here would provide crucial information on the tailoring of the enzyme to expand substrate scope to exploit its broader biotechnological applications.

### 3.3 Phosphorylation mechanism of MVK

Mutation of Asp204 in human MVK resulted in a >40,000-fold diminution in the catalytic rate constant,  $k_{cat}$ . Thus, it was hypothesized that this aspartate residue may function as a catalytic base to deprotonate the C5-OH of mevalonate, which then undergoes a nucleophilic attack on the  $\gamma$ -phosphate of ATP<sup>18,24</sup>. However, the detection of activity, albeit substantially reduced compared to the wild-type enzyme, indicates either that Asp204 does not have a catalytic role or that an alternative catalytic pathway is possible.

Recently, QM/MM studies on GHMP kinases galactokinase (GALK)<sup>22</sup> and phosphomevalonate kinase (PMK)<sup>23</sup> revealed that these enzymes actually function through a direct phosphorylation mechanism. Thus it is necessary to resolve the uncertainty around the mechanism of MVK, a representative GHMP kinase that was conventionally deemed to catalyse the phosphorylation via the catalytic base mechanism.

Cluster analysis was conducted to obtain the representative structures from the MD simulated trajectory of the ATP-Mg<sup>2+</sup>-MVK-MVA complex. Two key interatomic distances were monitored to choose a suitable structure for QM/MM. *i.e.* the distance between the  $\gamma$ -phosphate of ATP and the C5-hydroxyl oxygen of mevalonate and the distance between Asp204 and the C5-hydroxyl hydrogen of mevalonate (Table S4). From the ten representative structures of the last 10 ns of the MD trajectory (Figure S12), it was evident that no interatomic interaction exists between the aspartate residue and the C5-OH hydrogen; Moreover, the proton on the hydroxyl group points away from Asp204 in all representative structures, thus it is unlikely that the phosphorylation would occur via a catalytic base mechanism involving Asp204.

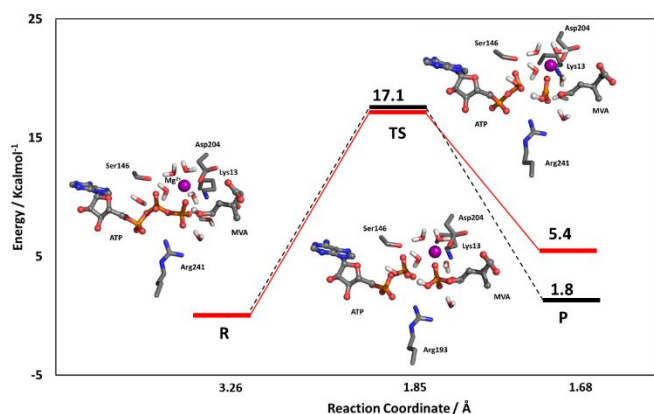
Here QM/MM studies were conducted to elucidate the mechanism of phosphorylation reaction catalyzed by MVK. The QM region was composed of the side chains of key residues Lys13, Ser146, Asp204 and Arg241, the Mg<sup>2+</sup> ion, the three water molecules coordinated to magnesium, mevalonate and the triphosphoryl tail of ATP. In total 45 heavy atoms were included. In addition, four structural water molecules identified from MD simulations were included in the QM region, two of which formed hydrogen bonds with the transferring  $\gamma$ -phosphate of ATP and another located at the  $\beta$ -phosphate of ATP. Strikingly, in the QM/MM optimized geometry, the hydroxyl group of the substrate approached the  $\gamma$ -phosphate of ATP (the distance between the  $\gamma$ -phosphate of ATP and the C5-OH group of mevalonate was 3.26 Å), in the meantime, the coordinate bond between the metal ion and the hydroxyl oxygen of mevalonate was broken. As a result, a fourth water molecule approached the active site to coordinate with Mg<sup>2+</sup>, replacing the coordination with the substrate observed in the MD simulated structure such that the favoured octahedral geometry of the metal ion was maintained.

In the optimized stationary points corresponding to the reactant, TS and product, Arg241 forms an ionic interaction with the  $\alpha$ -,  $\beta$ - and  $\gamma$ -phosphate groups of ATP. Furthermore, Arg241 forms an important ionic interaction with the  $\beta$ -, $\gamma$ -bridging oxygen of ATP, thus the negative charge that accumulates at the  $\beta$ -, $\gamma$ -bridging oxygen during bond cleavage in the phosphate transfer process is balanced (Figure 6).

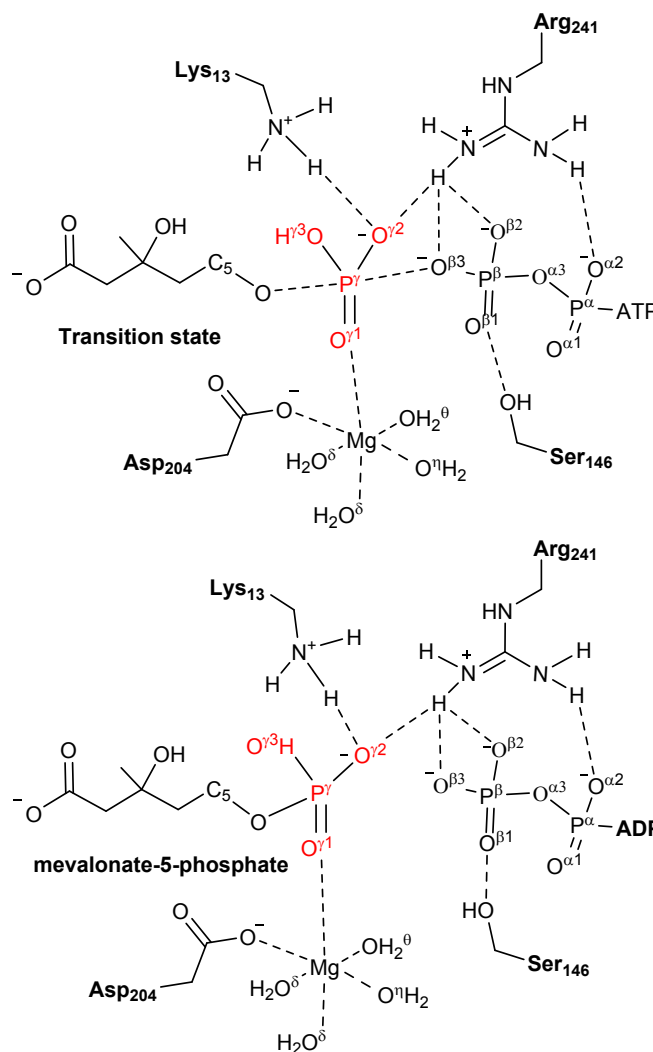
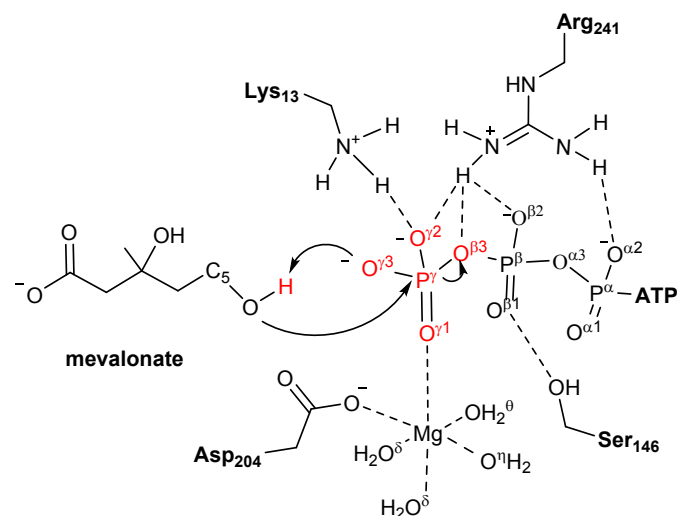
The calculated reaction barrier is 17.1 kcal/mol (Figure 5), which is in close agreement with the experimental barrier of 15.8 kcal/mol<sup>18</sup>. The transition state was validated by a unique imaginary frequency of -166.3 cm<sup>-1</sup>, which corresponds to the vibration of the phosphate transfer. From the QM/MM study we find that MVK follows a direct phosphorylation mechanism, despite the presence of Asp204 in the active site, which was conventionally proposed to be a catalytic base in previous literature<sup>18</sup>.

Pauling's formula<sup>42</sup>  $D(n) = D(l) - 0.6 \log n$  was used to quantify the associative or dissociative nature of the TS. In the formula the term  $D(l)$  is the average length of a P-O bond valued at 1.73 and  $D(n)$  is the

average of  $r_{\text{O-P}\gamma}$  and  $r_{\text{P}\gamma\text{-O}\beta 3}$  in the transition state. In the optimized transition state of MVK  $r_{\text{O-P}\gamma}$  was 1.86 Å and  $r_{\text{P}\gamma\text{-O}\beta 3}$  was 1.83 Å. The estimated fractional bond number ( $n$ ) is 0.61, representing 39% dissociative character. This is in conjunction with other GHMP kinases such as GALK<sup>22</sup> and PMK<sup>23</sup>, which were found to also follow a direct phosphorylation mechanism with a dissociative nature, as well as Isopentenyl phosphate kinase (IPK) in the mevalonate pathway that belongs to a distinct kinase family, amino acid kinase (AAK) family<sup>43</sup>.



**Fig. 5** Reaction profile of mevalonate kinase obtained by QM/MM calculations. The red line is the energy profile calculated using B97D/6-31g(d) method, and the dashed back line is the corrected energy profile using same DFT functional and 6-311++g(d,p) basis set.



**Fig. 6** 2D sketch of the reactant, TS and the product in the phosphorylation of mevalonate catalysed by MVK. The reaction follows a phosphorylation mechanism, *i.e.* the proton of the substrate hydroxyl group is abstracted by the  $\gamma$ -phosphate oxygen of ATP while the hydroxyl oxygen of the substrate undergoes nucleophilic attack on the  $\gamma$ -phosphate phosphore of ATP.

## Conclusions

MVK is a key enzyme in the biosynthesis of IPP, the precursor for isoprenoids. The lack of a complete kinase structure containing the substrate mevalonate and uncertainty of the  $\text{Mg}^{2+}$  location has limited the understanding of the catalytic mechanism of the kinase and therefore its biotechnological applications of MVK. Here we determined a catalytically competent structure of MVK-ATP- $\text{Mg}^{2+}$ -MVA based molecular docking, MD simulations and QM/MM studies. Our results show that Glu193 is not directly coordinated to  $\text{Mg}^{2+}$ , instead, it forms hydrogen bond interactions with the coordinating water molecules. Strikingly, Asp204 which was previously implicated as a catalytic base, forms a coordinate bond with  $\text{Mg}^{2+}$ . Furthermore, we found that Ser146 on the glycine-rich loop is actually involved in hydrogen bond interactions with the highly mobile triphosphoryl tail of ATP, instead of being coordinated to  $\text{Mg}^{2+}$ , as hypothesized previously. Notably, our MD simulations revealed that Arg241 moves



## ARTICLE

## Organic &amp; Biomolecular Chemistry

into the active site to form salt bridge interactions with the  $\alpha$ - and  $\gamma$ -phosphate of ATP, indicating it is pivotal in organizing the highly flexible nucleotide tail for in-line phosphate transfer, and neutralising the negative charge developed during phosphate transfer. This ordering of the triphosphate tail would result in a decrease in the entropy of the system which is, most likely, balanced by favourable enthalpy changes resulting from the non-covalent interactions between the enzymes and its substrates.

We also investigated the conformational change induced by substrate binding in MVK was by MD simulations. We found that in the absence of the substrate, there is an opening at the surface of MVK for the entry of a substrate molecule, which is mediated by the hydrogen bond interaction between Ser201 and Asp204; In the presence of the substrate, significant conformational change was observed in the loop connecting the  $\alpha_6$  and  $\alpha_7$  helix and closed conformation is observed as a result of Ser201 forming a hydrogen bond with the carboxylate group of the substrate. It is possible that similar conformational changes occur in other GHMP kinases which share a similar overall structure with active sites nested in the cleft between the two main motifs.

Furthermore, our QM/MM studies reveal that MVK follows a direct phosphorylation reaction mechanism. This argues the conventionally postulated catalytic base mechanism. Arg241 plays a pivotal role in organizing the highly flexible triphosphate tail of ATP, as well as neutralizing the negative charge accumulated at the  $\beta$ - $\gamma$ -bridging oxygen of ATP during phosphate transfer. Thus far, MVK, GALK<sup>22</sup> and PMK<sup>23</sup> have all been shown through QM/MM studies to have an in-line, direct phosphate transfer mechanism instead of a catalytic base mechanism. Homoserine kinase, which lacks such potential catalytic base residue in the active site would also function through direct phosphorylation. Taken together, we suggest that direct phosphorylation may be the predominant mechanism in most GHMP kinases. Such a conclusion challenges the generally held view that the majority of these enzymes function via the catalytic base mechanism. The results presented here would provide the basis for the rational design and engineering of MVK, a key enzyme in the mevalonate pathway for biosynthesis of isoprenoid-based biofuels. It may also provide valuable insights for development of potential pharmacological chaperone therapies for the stabilisation of the variant proteins in MKD-related diseases.

### Conflicts of interest

There are no conflicts to declare

### Acknowledgements

J.M. acknowledges the financial support from the Department of Education and Learning (DEL), Northern Ireland. We are grateful for the computing resources from QUB high performance computing centre.

### References

- 1 K. W. George, J. Alonso-Gutierrez, J. D. Keasling and T. S. Lee, *Adv. Biochem. Eng. Biotechnol.*, 2015, **148**, 355–389.
- 2 P. Gupta and S. C. Phulara, *J. Appl. Microbiol.*, 2015, **119**, 605–619.
- 3 J. Jurczyk, M. A. Munoz, O. P. Skinner, R. C. Chai, N. Ali, U. Palendira, J. M. Quinn, A. Preston, S. G. Tangye, A. J. Brown, E. Argent, J. B. Ziegler, S. Mehr and M. J. Rogers, *Immunol. Cell Biol.*, 2016, **94**, 994–999.
- 4 C. Browne and D. J. Timson, *Ann. Hum. Genet.*, 2015, **79**, 451–459.
- 5 S. H. L. Mandey, M. S. Schneiders, J. Koster and H. R. Waterham, *Human Mutation*, 2006, **27**, 796–802.
- 6 R. Berger, G. P. A. Smit, H. Schierbeek, K. Bijsterveld and R. le Coulter, *Clinica Chimica Acta*, 1985, **152**, 219–222.
- 7 J. M. Van Der Meer, J. Radl, C. L. M. Meyer, J. Vossen, J. Van Nieuwkoop, S. Lobatto and R. Van Furth, *The Lancet*, 1984, **323**, 1087–1090.
- 8 A. M. Prieur and C. Griscelli, *Sem Hop*, 1984, **60**, 163–167.
- 9 P. Bork, C. Sander and A. Valencia, *Protein Sci.*, 1993, **2**, 31–40.
- 10 Pastuszak, J. O'Donnell and A. D. Elbein, *J. Biol. Chem.*, 1996, **271**, 23653–23656.
- 11 I. Pastuszak, R. Drake and A. D. Elbein, *J. Biol. Chem.*, 1996, **271**, 20776–20782.
- 12 T. Yang, L. Bar-Peled, L. Gebhart, S. G. Lee and M. Bar-Peled, *J. Biol. Chem.*, 2009, **284**, 21526–21535.
- 13 S. Sherson, I. Gy, J. Medd, R. Schmidt, C. Dean, M. Kreis, A. Lecharny and C. Cobbett, *Plant Mol. Biol.*, 1999, **39**, 1003–1012.
- 14 H. Tomita, Y. Yokooji, T. Ishibashi, T. Imanaka and H. Atomi, *J. Bacteriol.*, 2012, **194**, 5434–5443.
- 15 L. A. Reinhardt, J. B. Thoden, G. S. Peters, H. M. Holden and W. W. Cleland, *FEBS Lett.*, 2013, **587**, 2876–2881.
- 16 X. Chu, N. Li, X. Liu and D. Li, *J. Biotechnol.*, 2009, **141**, 142–146.
- 17 C. F. Megarity, M. Huang, C. Warnock and D. J. Timson, *Bioorg. Chem.*, 2011, **39**, 120–126.
- 18 D. Potter and H. M. Miziorko, *J. Biol. Chem.*, 1997, **272**, 25449–25454.
- 19 D. Krepiak and H. M. Miziorko, *Protein Sci*, 2004, **13**, 1875–1881.
- 20 S. S. Krishna, T. Zhou, M. Daugherty, A. Osterman and H. Zhang, *Biochemistry*, 2001, **40**, 10810–10818.
- 21 J. B. Thoden and H. M. Holden, *J. Biol. Chem.*, 2005, **280**, 32784–32791.
- 22 M. Huang, X. Li, J.-W. Zou and D. J. Timson, *Biochemistry*, 2013, **52**, 4858–4868.
- 23 M. Huang, K. Wei, X. Li, J. McClory, G. Hu, J.-W. Zou and D. Timson, *J. Phys. Chem. B*, 2016, **120**, 10714–10722.
- 24 Z. Fu, M. Wang, D. Potter, H. M. Miziorko and J. J. P. Kim, *Journal of Biological Chemistry*, 2002, **277**, 18134–18142.
- 25 J. L. Andreassi, M. W. Vetting, P. W. Bilder, S. L. Roderick and T. S. Leyh, *Biochemistry*, 2009, **48**, 6461–6468.
- 26 A. Sali and T. L. Blundell, *J. Mol. Biol.*, 1993, **234**, 779–815.
- 27 J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath and A. Onufriev, *Nucleic Acids Res.*, 2005, **33**, W368–371.
- 28 J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *J. Mol. Biol.*, 1999, **285**, 1735–1747.
- 29 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J Comput Chem*, 2009, **30**, 2785–2791.
- 30 D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. SalomonFerrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and

- P. A. Kollman, *AMBER 14*, University of California, San Francisco, 2014.
- 31 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- 32 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2009.
- 33 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J Comput Chem*, 2004, **25**, 1157–1174.
- 34 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *Journal of Computational Physics*, 1977, **23**, 327–341.
- 35 P. Tao and H. B. Schlegel, *J. Comput. Chem.*, 2010, **31**, 2363–2369.
- 36 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 37 T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, *J. Chem. Theory Comput.*, 2006, **2**, 815–826.
- 38 Y.-K. Cho, S. E. Ríos, J.-J. P. Kim and H. M. Miziorko, *J. Biol. Chem.*, 2001, **276**, 12573–12578.
- 39 T. Zhou, M. Daugherty, N.V. Grishin, A.L. Osterman, H. Zhang. *Structure*. 2000, **8**, 1247-57.
- 40 F. Sievers and D. G. Higgins, in *Multiple Sequence Alignment Methods*, Humana Press, Totowa, NJ, 2014, pp. 105–116.
- 41 J. L. Andreassi, P. W. Bilder, M. W. Vetting, S. L. Roderick and T. S. Leyh, *Protein Sci.*, 2007, **16**, 983–989.
- 42 A. S. Mildvan, *Proteins*, 1997, **29**, 401–416.
- 43 J. McClory, D. J. Timson, W. Singh, J. Zhang and M. Huang, *J Phys Chem B*, 2017, **121**, 11062–11071.