Electronic structure benchmark calculations of CO_2 fixing elementary chemical steps in RuBisCO using the projector-based embedding approach[†]

Oscar A. Douglas-Gallardo,[‡] Ian Shepherd,[¶] Simon J. Bennie,[¶] Kara E. Ranaghan,[¶] Adrian J. Mulholland,[∗,¶] and Esteban Vöhringer-Martinez[∗,‡]

[‡]*Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Chile.*

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

E-mail: adrian.mulholland@bristol.ac.uk; evohringer@udec.cl

Abstract

Ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO) is the main enzyme involved in atmospheric carbon dioxide (CO_2) fixation in the biosphere. This enzyme catalyses a set of five chemical steps that take place in the same active-site within magnesium (II) coordination sphere. Here, a set of electronic structure benchmark calculations have been carried out on a reaction path proposed by Gready *et al.* by means of the projector-based embedding approach. Activation and reaction energies for all main steps catalyzed by RuBisCO have been calculated at the MP2, SCS-MP2, CCSD and CCSD(T)/aug-cc-pVDZ and cc-pVDZ levels of theory. The treatment of the magnesium cation with post-HF methods is explored to determine the nature of its

---

[†]A footnote for the title

involvement in the mechanism. With the high-level ab initio values as a reference, we tested the performance of a set of density functional theory (DFT) exchange-correlation (xc) functionals in reproducing the reaction energetics of RuBisCO carboxylase activity on a set of model fragments. Different DFT xc-functionals show large variation in activation and reaction energies. Activation and reaction energies computed at the B3LYP level are close to the reference SCS-MP2 results for carboxylation, hydration and protonation reactions. However, for the carbon-carbon bond dissociation reaction, B3LYP and other functionals give results that differ significantly from the ab initio reference values. The results show the applicability of the projector-based embedding approach to metalloenzymes. This technique removes the uncertainty associated with the selection of different DFT xc-functionals and so can overcome some of inherent limitations of DFT calculations, complementing and potentially adding to modelling of enzyme reaction mechanisms with DFT methods.

Introduction

Ribulose 1,5 bisphosphate carboxylase-oxygenase (RuBisCO) is a magnesium metalloenzyme found in the chloroplasts of the majority of photosynthetic organisms.[1–7] This enzyme is involved in the first main step of the Calvin-Benson-Bassham (CBB) cycle associated with atmospheric carbon dioxide (CO_2) assimilation.[1,2,4–9] Within the enzyme's catalytic pocket, a set of five chemical reactions convert the initial substrate, ribulose 1,5-bisphosphate (RuBP), to the final product, two-molecules of 3-phosphoglycerate (3-PGA).[4–6,10,11] The catalyzed chemical reactions involve CO_2 and H_2O addition to an activated RuBP intermediate by enolization, which subsequently undergoes a bond dissociation reaction to produce two-molecules of 3-PGA through stereospecific protonation.[5,6,10–12] All these chemical transformations (enolization, carboxylation, hydration, bond cleavage and stereospecific protonation) occur in a highly charged chemical environment centered on the first magnesium (II) coordination sphere (see Figure 1A).[4,5]

In the last five decades, much effort has been devoted to studying the mechanisms of RuBisCO carboxylase activity.[5,6,10–14] Computational and experimental studies have identified the role of different amino-acid residues within each chemical step.[4–7,10–24] Specifically, Gready et al.[5] have proposed a complete sequence of chemical species that connect the initial RuBP with the two-molecules of 3-PGA from calculations on a set of model fragments representing the active-site and substrate. These models represent the first magnesium coordination sphere including a reduced version of the substrate and the protonated residues Kcx201 (carbamylated lysine), Asp203 and Glu204 together with the second magnesium coordination sphere consisting of residues His294, Lys334, Lys177 and Lys175, schematically shown in Figure 1A. The numbering of amino-acid residues here used corresponds to *spinach* RuBisCO (PDB ID: 8RUC).[4]

In the proposed mechanism, amino-acid residues within the first and second magnesium coordination sphere were identified as promoting each main chemical step. In general, a proton transfer event was identified for each one of the main steps from the geometry optimization process carried out on each model fragment.[5] Gready et al. carried out electronic structure calculations using density functional theory (DFT). Geometry optimizations and energy calculations were carried out for each chemical step in order to obtain activation and reaction energies. Mechanistic details are still debated and the focus of ongoing research on RuBisCO.[6,10–14] The proposed reaction path by Gready et al. still represents the largest and most complete model reported so far to describe all chemical events that take place within enzyme active-site. However, a systematic study of the effect of the electronic structure method used to describe the energetics associated with each chemical step (activation and reaction energy) is still missing. We have showed that DFT methods provide accurate geometries but may present considerable errors in predicting carboxylation reaction electronic energies compared to complete basis set CCSD(T) calculations.[25] High-level post-HF ab initio electronic structure methods can provide accurate electronic energies to predict energetics of each chemical step involved in RuBisCO carboxylase activity. The inclusion of

electronic correlation energy in a systematic way helps to explore the role played by the magnesium cation and its first and second coordination sphere in the reaction. The validation of a computational method is crucial for the future development of more complex RuBisCO models taking into account the whole protein and its dynamics in multi-scale quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations which represents, nowadays, a powerful tool to explore enzymes.[26]

Here, a computational study of the carboxylase activity of RuBisCO enzyme assesses different ab initio and DFT electronic structure methods for the reaction path proposed by Gready *et al.*[5] The projector-based embedding approach is used: this can eliminate the variability often found when different DFT exchange-correlation functionals (DFT xc-functional) are selected.[27–30] This embedding approach has not previously been applied to a metalloenzyme, to our knowledge. This methodology allows us to obtain accurate activation and reaction energies, by describing the most important part of the reactions with a high-level ab initio post-HF electronic structure method, within a DFT treatment of the active-site model. The projector-based embedding approach tests and overcomes some inherent limitations of DFT methods by applying correlated ab initio methods. Similar to the projector-based embedding technique, other high-level approaches are emerging for the calculation of accurate activation barriers for enzyme-catalysed reactions.[31]

Computational Methods

For HF, DFT and MP2 calculations, the ORCA (version 4.0)[32,33] quantum chemistry package with the aug-cc-pVDZ basis set[34–36] was used, taking all electrons into account. A very dense DFT integration grid (Grid5 and Final Grid6 keywords in the ORCA input file) was chosen to reduce the effect of the grid size in these calculations. For DFTB,[37] the DFTB+ package[38] was used along with DFTB parameter set 3ob[39–41] which were specifically parameterized for organic and biological systems. In the case of projector-based embedding calculations,

the electronic energy was computed using the MOLPRO (Molpro 2019.1)[42] software. A specific active zone was selected to describe the most reactive part with a high-level wave-function (WF) method and the rest of the model fragment with different DFT xc-functionals (WF-in-DFT). The aug-cc-pVDZ[34–36] basis set was employed for both the high-level and low-level subregions. To accurately describe the chemically reactive region in the projector-based embedding calculations, different sets of active atoms were selected for each chemical step including or not the magnesium cation and applying a basis set truncation technique (truncation threshold parameter: 0.0001)[43] to reduce the computational cost associated with the embedding calculations. All model fragments contain 77 atoms with a total of 316 electrons. To compute the electronic energy for each model fragment, we have used the same optimized structures proposed by Gready *et al.* which have been reported in the Supporting Information of the original report.[5] However, we have detected that two of these structures showed a mismatch with the energetics and structural features reported by Gready *et al.* Both structures have been obtained for us by means of a optimization process by using previous transition state structures as initial guess (see Supporting Information).

Results and Discussion

The main goal in the present study is to obtain reliable activation and reaction energies involved in the RuBisCO carboxylase activity with a set of electronic structure methods of different nature using the projector-based embedding technique to enable high-level post-HF method on systems of several hundreds of electrons with reduced computational resources. The chemical reactions here considered (carboxylation, hydration, bond dissociation and protonation reactions) are schematically shown in Figure 1B where the main chemical change on the substrate molecule are highlighted in color red.

It is important to stress that model fragments proposed by Gready *et al.* only contains a reduced version of RuBP with four carbon atoms and just one phosphate (R-PO$_4^{-2}$) group
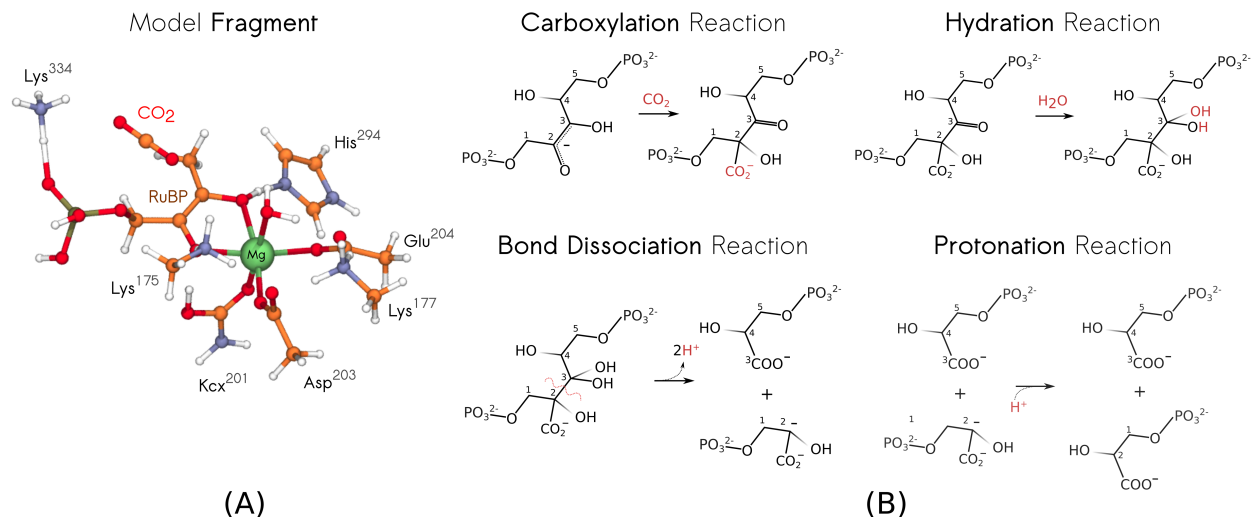
Figure 1: Schematic representation of (A) a model fragment of the enzymatic active site proposed by Gready *et al.*[5] to describe carboxylation step. (B) The main chemical reactions considered for RuBisCO carboxylase activity: carboxylation, hydration, bond dissociation and protonation reactions. The color red highlights the main chemical event on the substrate molecule, ribulose 1,5-bisphosphate (RuBP). It is important to stress that the model fragment shown in (A) only contains a reduced version of the substrate with four carbon atoms and just one phosphate group.

which has been neutralized with two protons to prevent artifactual energies. In the original paper, the electronic energy was computed at DFT level using the B3LYP xc-functional and the 6-31G(d,p) basis set. To reproduce Gready's electronic energy values, a set of single-point calculations at the same level of theory (B3LYP, 6-31G(d,p)) was carried out on each model fragment and obtained electronic changes are shown in Figure 2 with the label B3LYP* and in Table S1 in the Supporting Information. The obtained values were $\Delta E^{\ddagger}= 5.9$ and $\Delta E_0= $ -7.7 kcal mol^{-1} (carboxylation), $\Delta E^{\ddagger}= 21.6$ and $\Delta E_0= 0.3$ kcal mol^{-1} (hydration), $\Delta E^{\ddagger}= 27.4$ and $\Delta E_0= 11.3$ kcal mol^{-1} (bond dissociation) and $\Delta E^{\ddagger}= 13.5$ and $\Delta E_0= $ -18.1 kcal mol^{-1}(bond dissociation), respectively. In order to assess the effect of different electronic structure methods on activation and reaction energies, a set of single-point calculations was computed on the same optimized structures proposed by Gready *et al.* (Cartesian coordinates are provided in the Supporting Information) and the obtained energetic changes are also shown in Figure 2.

6

DFT Results

First, the performance of a set of DFT xc-functionals of different nature compared to a set of high and low-level electronic structure methods (HF, DFTB, MP2, SCS-MP2) was analyzed. Figure 2 shows the activation (blue) and reaction (orange) energies obtained for each chemical step within the sequence of reactions catalyzed by RuBisCO enzyme (carboxylation, hydration, bond dissociation and protonation). In biological systems that have an electronic state that is well described by a single electronic configuration the performance of SCS-MP2 is often comparable to other high-level electronic structure methods like CCSD(T),[27–29,44] therefore, SCS-MP2 was selected as the reference in this particular selection of methods and its results are shown as dotted line in Figure 2.

As can be seen in Figure 2, the calculated barriers and reaction energies are notably sensitive to the DFT xc-functional employed. This shows the need to be cautious in drawing mechanistic conclusions from DFT calculations. Of the WF methods (last three columns in Figure 2), HF and MP2 show the largest and smallest activation and reaction energy values in all steps. The inclusion of the spin-component-scaling within the MP2 energy calculations (SCS-MP2) increases its activation and reaction energy values, placing it between HF and MP2. The HF results are only included as a comparison and we do not of course expect them to be useful in predicting reaction energetics. Of DFT results, the smallest and largest activation and reaction energy values were obtained with PBE and BH&HLYP, respectively, which is a general trend of GGA functionals underestimating barriers and with inclusion of HF exchange increasing them. In between, a range of different values for different DFT xc-functionals (B3LYP, M06-L, M06-2x, ωB97x and PBE0) is observed. For this set of reactions, the semiempirical DFTB method turned out to be very close to B3LYP particularly, for the activation energy values. DFTB is widely used in QM/MM molecular dynamics simulations.[45,46] This indicates that DFTB is a reasonable choice for modelling this RuBisCO activity.

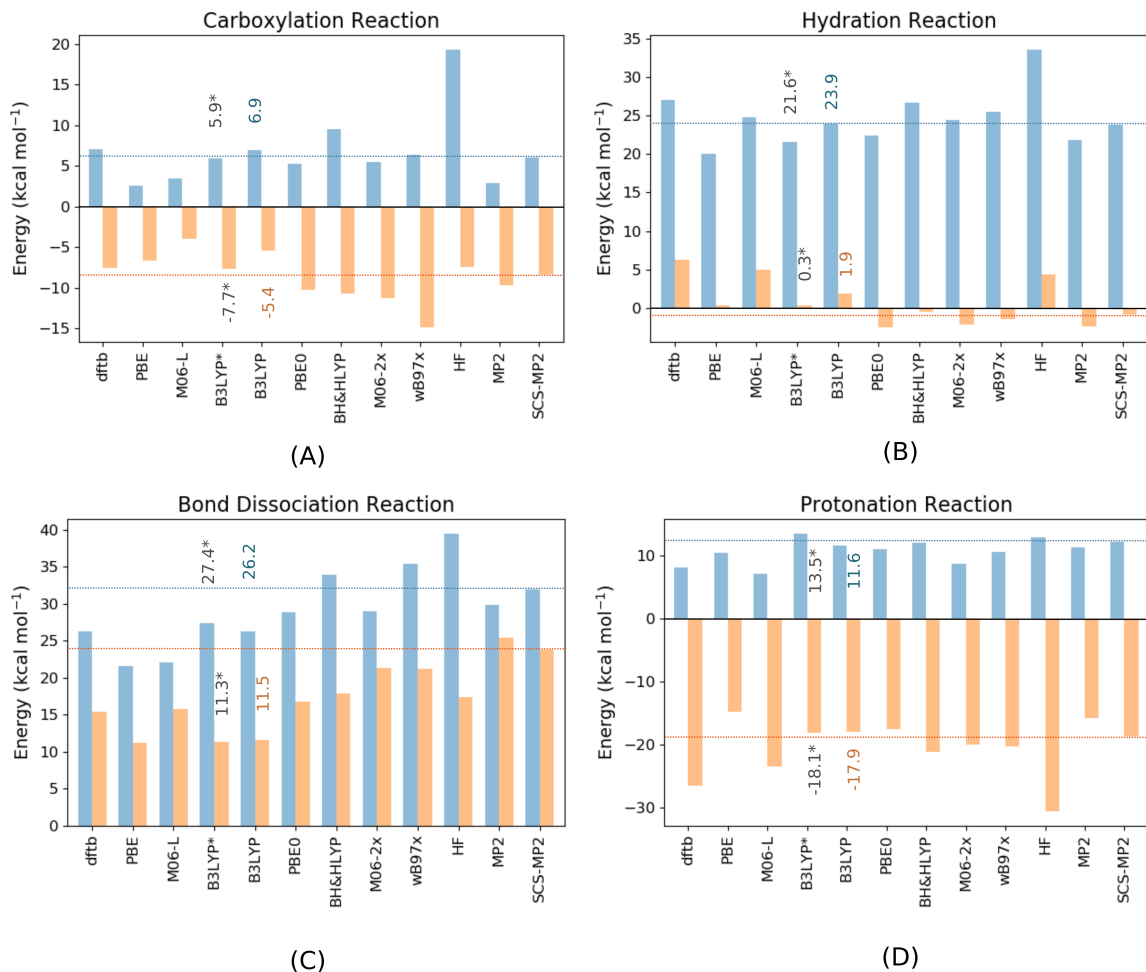Among the electronic structure methods, SCS-MP2 is the highest level method consid-

Figure 2: Activation (bar in color blue) and reaction (bar in color orange) energies calculated for the (A) carboxylation , (B) hydration, (C) bond cleavage and (D) protonation reactions. In all cases the electronic energy was determined by using a Dunning basis set, aug-cc-pVDZ. The electronic energy obtained at the B3LYP/6-31G(d,p) level was included as comparison and highlighted with an asterisk.

ering the 77 atoms and 316 electrons of the whole system and provides the reference, which in general gives better results than MP2 and is very close to CCSD(T) in some cases.[27–29,44]

Projector-based Embedding Results

A set of projector-based embedding calculations was carried out on each RuBisCO model fragment determined by Gready $et\ al.$[5] The projector-based embedding approach allows high-level post-HF calculations by reducing the computational resource costs (e.g. memory). Here

we calculated the activation and reaction energies of each chemical step. In the embedding technique, a high-level region is defined within a larger QM region; the high-level region is typically treated by a correlated ab initio method, and the larger QM region is typically treated by DFT. The embedding approach has been shown to remove the dependence of enzyme reaction barriers on density functional.[29] To our knowledge, the calculations here are the first application of this approach to a metalloenzyme.

A suitable high-level region was selected for each RuBisCO model including all active atoms close to the reactive center. The effects of including the magnesium cation within the high-level region was also explored to study its role in the reaction. We selected the carboxylation reaction to study the effect of the high-level region size on the computed activation and reaction energies. Three different sizes of high-level regions were explored (shadowed atoms on reactive species, Figure 3), with an increasing number of active atoms (first column in Figure 3).

The smallest high-level region (region I, 39 high-level orbitals) only contains the CO_2 molecule and the majority of substrate atoms (up to two bonds from C2 carbon atom). Here, it is important to stress that carboxylation takes place on the C2 carbon atom and in Gready et $al.$'s model, this step is assisted by His294 as a base. Part of this residue was considered in the high-level region in the calculations using region II (51 high-level orbitals) as shown in Figure 3. Thus, the proton transfer process that takes places between the RuBP O3 oxygen atom and His294 (ϵN) nitrogen atom was included in the high-level region, along with the carboxylation reaction. Finally, in the largest high-level region (region III, 68 high-level orbitals) a part of both amino-acid residues Kcx201 and Lys175 were also included due to their proximity to the reactive centers located on the C2 carbon atom.

For all these systems, the electronic energy was computed with the projector-based embedding approach in SCS-MP2-in-DFT scheme using different DFT xc-functionals (B3LYP, PBE, BH&HLYP, and M06-2x) to describe the chemical environment and to compare it with the full SCS-MP2 result. The computed reaction and activation energy for each system
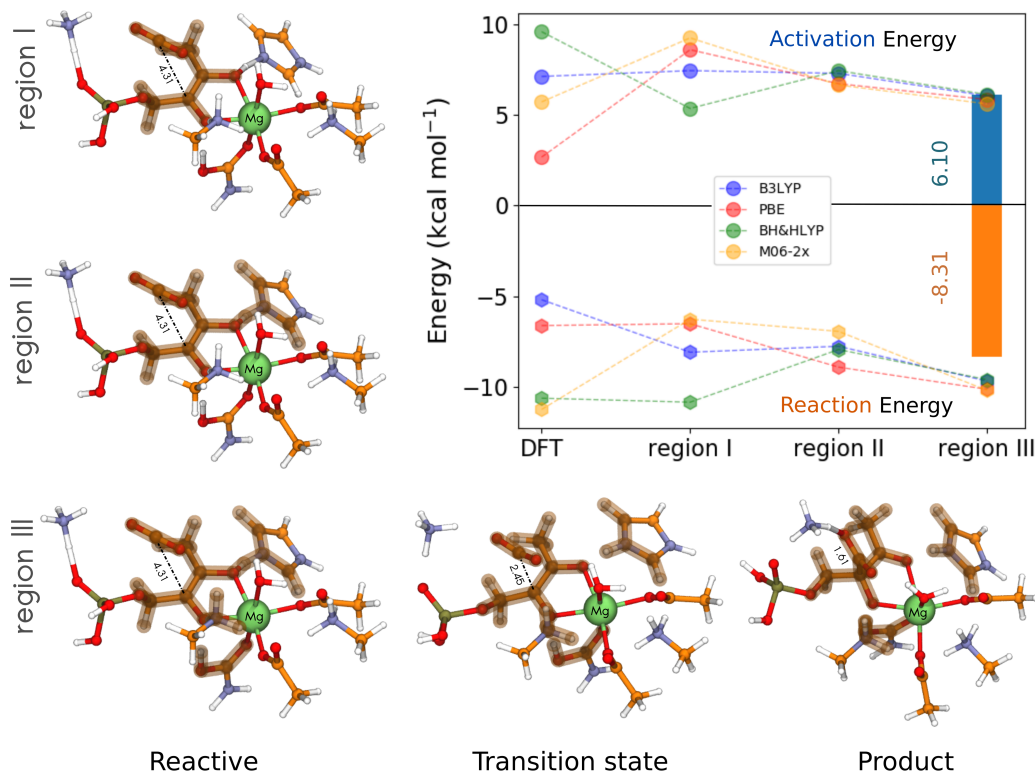
Figure 3: On the left, three different sizes of the high-level region are shown: (A) small : region I, (B) medium: region II and (C) large : region III. The active atoms selected in each high-level region are highlighted with a shadow in each model fragment. On the bottom the reactant, transition state and product structure of the carboxylation reaction are shown using the largest high-level region. On the right top the activation and reaction energy with different high-level region size employed for projector-based embedding calculations in a SCS-MP2-in-DFT scheme are shown. For reference, full-DFT (first point, labeled DFT) and full-SCS-MP2 (bar graph) results of the whole model fragment are shown.

including the magnesium cation in the high-level region are shown in Figure 3. The full SCS-MP2 electronic energy value for the whole model fragment was included as a bar graph at the right as a reference value. The full SCS-MP2 activation and reaction energy for the carboxylation reaction are $\Delta E^{\ddagger}= 6.10$ and $\Delta E_0= $ -8.31 kcal mol^{-1}. The full DFT result for each DFT xc-functional is also shown in the same Figure with the label DFT. As can be seen in the inset of Figure 3, when more atoms near to the reactive center are included within the high-level region, the variability initially found with different DFT xc-functional is reduced and the energy values converge to the full SCS-MP2 result. Therefore, it is possible to ob-

tain energy values very close to the full SCS-MP2 result with the projector-based embedding approach, with different DFT xc-functionals, reducing the computational resources (memory) considerably, showing that this is a practical and accurate approach to high accuracy calculations of reaction energies for metalloenzymes.

Using high-level regions of comparable size as region III, we extended the analysis to other chemical reactions of RuBisCO. Table 1 shows the error due to embedding of the activation and reaction energies obtained at SCS-MP2-in-B3LYP or MP2-in-B3LYP/aug-cc-pVDZ level compared to the full SCS-MP2 and MP2 calculations for each reaction step. It can be seen that the errors introduced by embedding are very small: the results from embedded calculations are very similar to the full ab initio results. The active atoms considered for each step are shown in Figure S1 in Supplementary Information. For each reaction a different set of atoms was selected including all atoms separated by up to two bonds from the reactive center.

Table 1: Reaction energies and activation barriers for chemical steps, calculated with MP2 and SCS-MP2 methods. Errors from emdedding are calculated by comparing these values to projector-based embedding results from a WF-in-B3LYP/aug-cc-pVDZ scheme. All energy values are expressed in kcal mol^{-1}. The numbers of orbitals included in the high-level region are 68, 64, 80 and 69, respectively.

Method	Carboxylation		Hydration		Bond Dissociation		Protonation	
	$\Delta E^{\ddagger}$	ΔE_0	$\Delta E^{\ddagger}$	ΔE_0	$\Delta E^{\ddagger}$	ΔE_0	$\Delta E^{\ddagger}$	ΔE_0
Full MP2	2.89	−9.65	21.85	−2.45	29.88	25.40	11.33	−15.82
Full SCS-MP2	6.10	−8.31	23.91	−0.90	31.99	23.78	12.16	−18.61
Errors from embedding:								
MP2-in-B3LYP	0.06	−0.57	−0.40	−2.63	2.59	0.16	3.01	2.47
SCS-MP2-in-B3LYP	−0.06	−1.35	−0.68	−2.26	2.34	−0.10	2.78	1.90

The error from using the projector-based embedding technique is smaller than 3 kcal mol^{-1} for all reactions and almost half of the values are within 1 kcal mol^{-1}. This shows that the embedding method provides accurate results for this metalloenzyme, as has been found previously for reactions catalyzed by enzymes that do not use metal ions in catalysis.[27–30] Considering that this method enables high-level post-HF calculations on electronic systems

of several hundreds of electrons, its direct application along the reaction path obtained from QM/MM calculations becomes possible. From the potential energy profiles, high-level post-HF activation barriers will be derived, which will consider the electrostatic effect of the surrounding enzyme, contributing to the understanding of enzymatic reaction mechanisms.

To investigate the accuracy of the SCS-MP2 calculations in the WF-in-DFT schemes, a smaller high-level region with a reduced basis set (cc-pVDZ) was selected to obtain CCSD and CCSD(T) results. Table S2 in the Supporting Information shows the results obtained at MP2, SCS-MP2, CCSD and CCSD(T)-in-B3LYP/cc-pVDZ level for this system. In this Table, we also address the effect of including the magnesium cation in the active (high level) zone (Mg) or describing it at the DFT level. The results show that the inclusion of the magnesium atom in the high-level region has a negligible effect on the computed energies for all the reaction steps. The energy differences are below 0.5 kcal mol^{-1} for all the reactions. This embedding result indicates that DFT is sufficient to describe the Mg^{2+} cation in this system. The SCS-MP2 and CCSD(T) results are very close to each other, showing that SCS-MP2 provides an adequate reference. The energy differences are often below 2 kcal mol^{-1}. Finally, we combine the obtained results to define the energetics of the different steps and compare them to the values obtained by Gready *et al.* at the B3LYP/6-31G(d,p) level. The calculated energy profiles are compared to B3LYP results in Figure 4.

The (full, canonical) SCS-MP2/aug-cc-pVDZ method was selected as our reference as they provide values close to CCSD(T) as discussed above (they are also very close to SCS-MP2-in-DFT results). It is important to stress that it was not feasible to perform CCSD(T) calculations with larger basis sets with our available computing resources. The largest energy difference in the reaction energetics (DFT vs SCS-MP2) is in the reaction energy for the dissociation reaction, which is less exothermic. This chemical step involves bond breaking, which needs a proper description of the electronic correlation energy as we have shown previously.[25] A notable difference is that SCS-MP2 calculations predict the transition state of the proton transfer reaction as the highest energy point along the whole path from the
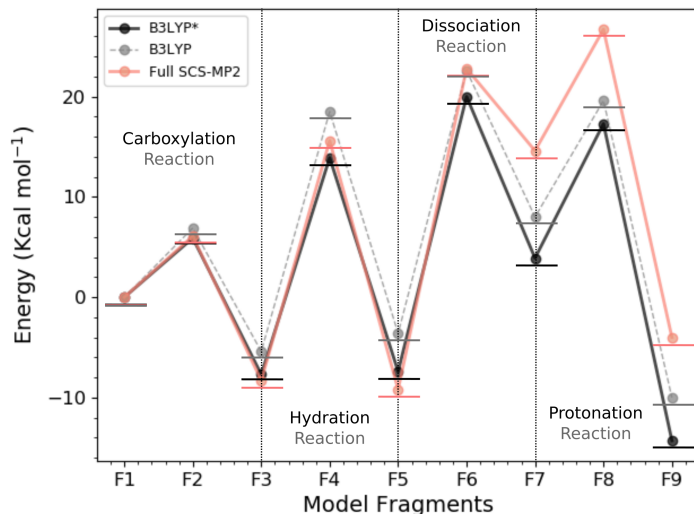
Figure 4: Energy profiles for all chemical steps calculated using different electronic structure methods. B3LYP*:(B3LYP/6-31G(d,p)), B3LYP: (B3LYP/aug-cc-pVDZ) and Full SCS-MP2: (SCS-MP2/aug-cc-pVDZ). In this case, a full SCS-MP2 calculation was selected to compare the energetic changes instead of the SCS-MP2-in-DFT embedding scheme. As it can be seen in Table 1, embedding results provide electronic energy values very close to full SCS-MP2 if a high-level region large enough is selected.

reactants to the product.

Considering the small errors incurred in the projector-based embedding approach and its reduced computational cost, this study shows that DFT potential energy profiles can be improved with more accurate post-HF methods in embedding calculations, removing uncertainly due to choice of DFT functional, to address reaction mechanisms in enzymes, including metalloenzymes.

Conclusions

Our results show the projector-based embedding technique is applicable to a metalloenzyme, using post-HF calculations which include the metal cation within the high-level region. The SCS-MP2-in-B3LYP method with the aug-cc-pVDZ basis set correctly describes the energetics of the chemical steps associated with the CO_2 fixation mechanism of RuBisCO resulting

13

in errors below 3 kcal mol^{-1} (and almost half of the values are within 1 kcal mol^{-1}) compared to our reference calculations.

Acknowledgement

Supporting Information Available

Figure S1 shows the largest high-level regions selected for each chemical step. DFT single-point electronic energy calculation on optimized geometries proposed by Gready *et al.* at the B3LYP/6-31G(d,p) level are shown in Table S1. These electronic energy values are referred to electronic energy associated with the first fragment model. In Table S2 activation and reaction energies calculated with projector-based embedding approach at SCS-MP2, MP2, CCSD and CCSD(T)-in-B3LYP level with the cc-pVDZ basis set are shown. We left available the new Cartesian coordinates for F3 and F9 which have been recalculated for us.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Calvin, M. Angew. Chem. Int. Ed. **1962**, 1, 65–75.

(2) Erb, T. J. Appl. Environ. Microbiol. **2011**, 77, 8466–8477.

(3) Vitillo, J. G. RSC Adv. **2015**, 5, 36192–36239.

(4) Andersson, I. J. Mol. Biol. **1996**, 259, 160 – 174.

(5) Kannappan, B.; Gready, J. E. J. Am. Chem. Soc. **2008**, 130, 15063–15080.

(6) Cleland, W. W.; Andrews, T. J.; Gutteridge, S.; Hartman, F. C.; Lorimer, G. H. Chem. Rev. **1998**, 98, 549–562.

(7) Erb, T. J.; Zarzycki, J. Curr. Opin. Biotechnol. **2018**, 49, 100 – 107.

(8) Portis, A. R.; Parry, M. A. J. Photosynth. Res. **2007**, 94, 121–143.

(9) Appel, A. M. et al. Chem. Rev. **2013**, 113, 6621–6658.

(10) Cummins, P. L.; Kannappan, B.; Gready, J. E. J. Comput. Chem. **2018**, 39, 1656–1665.

(11) Cummins, P. L.; Kannappan, B.; Gready, J. E. J. Phys. Chem. B **2019**, 123, 2679–2686.

(12) Tcherkez, G. Plant, Cell & Environment **2013**, 36, 1586–1596.

(13) Flamholz, A. I.; Prywes, N.; Moran, U.; Davidi, D.; Bar-On, Y. M.; Oltrogge, L. M.; Alves, R.; Savage, D.; Milo, R. Biochemistry **2019**, 58, 3365–3376.

(14) Tcherkez, G. G. B.; Bathellier, C.; Stuart-Williams, H.; Whitney, S.; Gout, E.; Bligny, R.; Badger, M.; Farquhar, G. D. Biochemistry **2013**, 52, 869–877.

(15) King, W. A.; Gready, J. E.; Andrews, T. J. Biochemistry **1998**, 37, 15414–15422.

(16) Mauser, H.; King, W. A.; Gready, J. E.; Andrews, T. J. J. Am. Chem. Soc. **2001**, 123, 10821–10829.

(17) Oliva, M.; Safont, V. S.; Andrs, J.; Tapia, O. J. Phys. Chem. A **2001**, 105, 4726–4736.

(18) El-Hendawy, M. M.; English, N. J.; Mooney, D. A. J. Mol. Model. **2013**, 19, 2329–2334.

(19) Andersson, I.; Taylor, T. C. Arch. Biochem. Biophys. **2003**, 414, 130 – 140.

(20) Andersson, I. J. Exp. Bot. **2008**, 59, 1555–1568.

(21) Andersson, I.; Backlund, A. Plant Physiol. Bioch **2008**, 46, 275 – 291.

(22) van Lun, M.; Hub, J. S.; van der Spoel, D.; Andersson, I. J. Am. Chem. Soc. **2014**, 136, 3165–3171.

(23) Tcherkez, G. G. B.; Farquhar, G. D.; Andrews, T. J. Proc. Natl. Acad. Sci. **2006**, 103, 7246–7251.

(24) Stec, B. Proc. Natl. Acad. Sci. **2012**, 109, 18785–18790.

(25) Douglas-Gallardo, O. A.; Saez, D. A.; Vogt-Geisse, S.; Vöhringer-Martinez, E. J. Comput. Chem. **2019**, 40, 1401–1413.

(26) Amaro, R. E.; Mulholland, A. J. Nat. Rev. Chem. **2018**, 98, 148.

(27) Xinglong, Z.; Bennie, S.; der Kamp, M. V.; Glowacki, D.; Manby, F.; Mulholland, A. J. R. Soc. open. sci. **2018**, 5, 171390.

(28) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. J. Chem. Theory Comput. **2016**, 12, 2689–2697.

(29) Ranaghan, K. E.; Shchepanovska, D.; Bennie, S. J.; Lawan, N.; Macrae, S. J.; Zurek, J.; Manby, F. R.; Mulholland, A. J. J. Chem. Inf. Model. **2019**, 59, 2063–2078.

(30) Ranaghan, K. E.; Morris, W. G.; Masgrau, L.; Senthilkumar, K.; Johannissen, L. O.; Scrutton, N. S.; Harvey, J. N.; Manby, F. R.; Mulholland, A. J. J. Phys. Chem. B **2017**, 121, 9785–9798.

(31) Bistoni, G.; Polyak, I.; Sparta, M.; Thiel, W.; Neese, F. Journal of Chemical Theory and Computation **2018**, 14, 3524–3531, PMID: 29883118.

(32) Neese, F. Wiley Interdiscip. Rev. Comput. Mol. Sci. **2012**, 2, 73–78.

(33) Neese, F. Wiley Interdiscip. Rev. Comput. Mol. Sci. **2017**, 1–6.

(34) Dunning, T. H. J. Chem. Phys. **1989**, 90, 1007–1023.

(35) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. J. Chem. Phys. **1992**, 96, 6796–6806.

(36) Woon, D. E.; Dunning, T. H. J. Chem. Phys. **1993**, 98, 1358–1371.

(37) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Phys. Rev. B **1998**, 58, 7260–7268.

(38) Aradi, B.; Hourahine, B.; Frauenheim, T. J. Phys. Chem. A **2007**, 111, 5678–5684.

(39) Gaus, M.; Goez, A.; Elstner, M. J. Chem. Theory Comput. **2013**, 9, 338–354, PMID: 26589037.

(40) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. J. Chem. Theory Comput. **2014**, 10, 1518–1537.

(41) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. J. Phys. Chem. B **2015**, 119, 1062–1082.

(42) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schtz, M. Wiley Interdiscip. Rev. Comput. Mol. Sci. **2012**, 2, 242–253.

(43) Bennie, S. J.; Stella, M.; Miller, T. F.; Manby, F. R. J. Chem. Phys. **2015**, 143, 024105.

(44) Sirirak, J.; Lawan, N.; van der Kamp, M. W.; Harvey, J. N.; Mulholland, A. J. PeerJ **2020**, in press.

(45) Cui, Q.; Elstner, M. Phys. Chem. Chem. Phys. **2014**, 16, 14368–14377.

(46) Christensen, A. S.; Kuba, T.; Cui, Q.; Elstner, M. Chemical Reviews **2016**, 116, 5301–5337, PMID: 27074247.

Graphical TOC Entry