

# Multi and single-reference methods for the analysis of multi-state peroxidation of enolates

P. Ortega,<sup>1</sup> S. Gil-Guerrero,<sup>1</sup> A. Veselinova,<sup>1</sup> A. Zanchet,<sup>2</sup> L. González-Sánchez,<sup>1</sup> P. G. Jambrina,<sup>1</sup> and C. Sanz-Sanz<sup>3, a)</sup>

<sup>1)</sup> *Departamento de Química-Física. Universidad de Salamanca, Salamanca, 37008 (Spain)*

<sup>2)</sup> *Instituto de Física Fundamental (CSIC), Madrid, 28006 (Spain)*

<sup>3)</sup> *Departamento de Química Física Aplicada. Universidad Autónoma de Madrid, Madrid, 28049 (Spain)*

(Dated: 17 March 2021)

In spite of being spin-forbidden, some enzymes are capable of catalyzing the incorporation of  $O_2(^3\Sigma_g^-)$  to organic substrates without needing any cofactor. It has been established that the process followed by these enzymes starts with the deprotonation of the substrate forming an enolate. In a second stage, the peroxidation of the enolate formation occurs, a process in which the system changes its spin multiplicity from a triplet state to a singlet state. In this article, we study the addition of  $O_2$  to enolates using state-of-the-art multi-reference and single-reference methods. Our results confirm that intersystem crossing is promoted by stabilization of the singlet state along the reaction path. When multi-reference methods are used, large active spaces are required, and in this situation, Semistochastic Heat-Bath Configuration Interaction (SHCI) emerges as a powerful method to study these multi-configurational systems and is in good agreement with LCCSD(T) when the system is well-represented by a single-configuration.

## I. INTRODUCTION

Every computational chemistry study starts with the selection of a suitable electronic structure method. This selection is made based on a trade-off between accuracy and computational cost, and the method is finally chosen depending on the size of the system under study. If the system can be reasonably described by one configuration, single-reference methods such as MP2,<sup>1</sup> and specially DFT,<sup>2</sup> are often used. CCSD(T),<sup>3</sup> considered as the “gold standard” in quantum chemistry, is the best choice for small systems for which the calculation is affordable. However, when the ground state is strongly coupled to excited states, the aforementioned are not reliable, and an accurate description of the physics of these systems requires the use of multi-reference methods.

The most widely used multi-configuration method is the CASSCF method, developed more than 30 years ago.<sup>4</sup> In CASSCF, the total orbital space is divided into three sets: inactive orbitals that are doubly occupied in all configurations, virtual orbitals that are always unoccupied, and the active orbitals where all possible occupations are considered. Based on these configurations, the multi-reference wavefunction is built, and subsequently, the coefficients of all configurations and orbitals are optimized. This allows to account for the static electronic correlation and describe properly electronic states with a strongly mixed character. Dynamic correlation can be introduced via perturbation theory (CASPT2) or variational multi-reference configuration interaction (MRCI) method. The number of configurations (and the computational cost) rises exponentially with the number of

active orbitals, which usually limits the number of active space orbitals to 18 electrons in 18 orbitals (CAS(18,18)) considering current computational resources, with the most recent and efficient implementations,<sup>5</sup> which could be increased up to 20 electrons for 20 orbitals in large parallel machines.<sup>6</sup> Larger active spaces can be used using restricted active space methods (RAS), which limits the number of possible occupations arising from the active space orbitals.<sup>7</sup> Not only the limitation in the number of active space orbitals hampers the use of CASSCF-like methods. The active space should be robust throughout the reaction path and include all the configurations of interest along this path. Although some strategies have been developed to automatically detect rotations in the active space orbitals,<sup>8–10</sup> the selection of a robust and coherent active space can still be very cumbersome.

In the last decade, we have witnessed the development of a plethora of new multi-configuration methods.<sup>11</sup> On one hand, single-reference approaches such as the spin-flip family of methods (see Ref. 12 for a recent review) and on the other hand stochastic multi-reference methods using Monte Carlo methods,<sup>13,14</sup> that allows extending the number of active space to around 100 orbitals.<sup>15–17</sup> Among these methods, density matrix renormalization group (DMRG) method<sup>18,19</sup> and the Semistochastic Heat-Bath Configuration Interaction (SHCI) have been widely used.<sup>20–22</sup> These methods alleviate the problems arising from the selection of the active space, and using SHCI, active spaces of (30,108)<sup>23</sup> and (32,204)<sup>24</sup> have been reported. In this article, we use the SHCI method to describe the addition of molecular oxygen to enolates, intrinsically multi-configuration systems.

From an electronic structure perspective, reactions between  $O_2$  and most organic molecules are in principle spin-forbidden. This is because  $O_2$  in its ground state is a triplet ( $^3\Sigma_g^-$ ) while most of the organic molecules, includ-

<sup>a)</sup> Electronic mail: cristina.sanz@uam.es

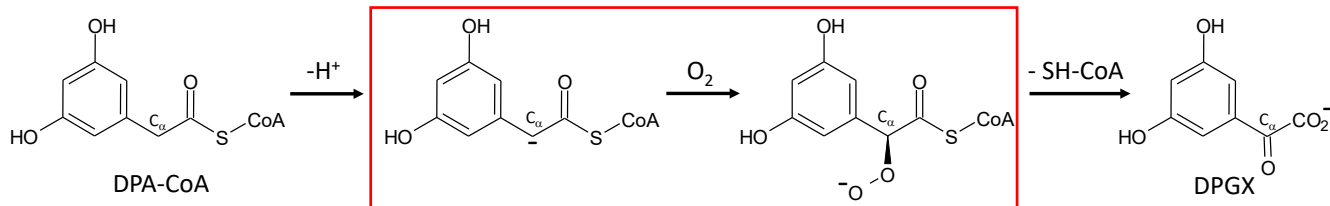


FIG. 1: Scheme of the proposed mechanism for the oxidation of DPA-CoA and synthesis of DPGX. The step highlighted in red is that studied in the present manuscript (reaction between  $O_2$  and the enolate).

ing the peroxide formed by the addition of  $O_2$  to double bonds, are singlets. At very high temperatures, radicals are formed, which react with  $O_2$  via a spin-allowed process. Since working at such temperatures is not an option for living organisms, enzymes devised other strategies to catalyze a reaction with  $O_2$ . Most enzymes rely on transition metals, particularly Fe with or without a heme group.<sup>25–31</sup> Crystallographic structures have shown unequivocally that other enzymes use organic cofactors, such as flavin,<sup>32,33</sup> while others carry out their catalytic activity without the presence of any cofactors.<sup>34–36</sup> The mechanism followed by these enzymes has been thoroughly studied (see for example Refs. 37–42), and it has been established that the first stage is deprotonation of the substrate prior to the diffusion of  $O_2$  to the active site,<sup>43</sup> forming an enolate that will subsequently react with  $O_2$  forming a peroxide.

3,5-dihydroxyphenylacetyl-CoA 1,2-dioxygenase (DpgC) is a protein which plays a key role in the biosynthesis of antibiotics of last resort such as vancomycin or teicoplanin and was the first protein for which a crystal structure showing a substrate analog and  $O_2$  bound to the enzyme was resolved.<sup>44,45</sup> DpgC catalyzes the cofactor-independent oxidation of 3,5-dihydroxyphenylacetyl-CoA (DPA-CoA) to 3-5-dihydroxyphenyl-glyoxylate (DPGX), and a scheme of the proposed mechanism is depicted in Fig. 1. Previous calculations in our group<sup>42</sup> combining QM/MM (DFT), MRCI calculations on a model system with a reduced active space, have shown that, once the enolate is formed, the reaction proceeds via transference of two electrons from the enolate to  $O_2$  and without requiring activation of  $O_2$ . The intersystem crossing was found to be concomitant with the first electron transfer, and the minimum energy crossing point, which plays the role of an effective barrier,<sup>46</sup> is stabilized by the formation of a stable intermolecular complex between DPA-CoA and  $O_2$ .

The fact that at least two electronic states of different spin multiplicities are necessary to describe the system implies that multi-configuration methods should be used to understand the physics of these systems. Ideally, to treat this system using a CAS-like method, the active space should include all the  $\pi$  orbitals of the aromatic system and  $O_2$ . However, since the hybridization of  $C_\alpha$

changes from  $sp^2$  to  $sp^3$ , the aromaticity is broken along the reaction path and the related orbitals are generally expelled from the active space leading to an artificial discontinuity in the electronic wavefunction. In Ref. 42, the problem of the stability was circumvented by using a small active space including only the frontier orbitals. Our results allowed us to demonstrate the multi-surface nature of the peroxidation reaction. The drawback of this approach is that only a small amount of correlation was gathered. The use of stochastic multi-reference methods permits to use a different approach, including all valence orbitals and electrons in the active space.

In this work, we will calculate the reaction path for this reaction using SHCI on three models of different size, and we will compare these reaction paths with those obtained using single-reference methods such as DFT and CCSD(T), the latter in the form of the explicitly correlated open-shell pair natural orbital local coupled-cluster methods, PNO-LCCSD(T)-F12.<sup>47</sup> Using these calculations, we will shed light on the mechanism of the peroxidation reaction of enolates, and also assess if single-reference methods as CCSD(T) and stochastic multi-reference methods (SHCI) are capable of describing the process. Moreover, we will evaluate the performance of DFT method, which is widely used for the study of these reactions in biological conditions.

The manuscript is organized as follows. In the methods section, we present the three models used to describe this process, along with a short description of the main features of the electronic structure methods used. In the results section, we will show the energy profiles for the peroxidation of the enolates. Finally, a summary and the conclusions of this work are presented.

## II. METHODS

There are two reasons why the reaction between DPA-CoA and  $O_2$  is an excellent system for the study of cofactorless addition of  $O_2$  to enolates. On one hand, no basic residues are present in the active site, and the protein plays a passive role in promoting intersystem crossing,<sup>42</sup> by restraining the carbon atoms backbone close to a planar conformation and guiding the oxygen

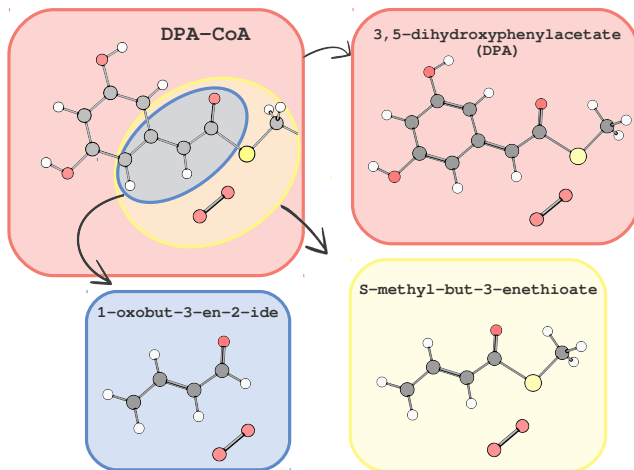


FIG. 2: Scheme of the three model systems used to describe the reaction between DPA-CoA and  $O_2$ . In the top-left panel,  $O_2$  and the DPA moiety of DPA-CoA are depicted. The three model systems are depicted in the bottom-left and right panels, where the colors show the relationship between the models and the system.

to the active site. On the other hand, peroxidation takes place out from the aromatic ring, making possible to propose smaller systems that may reproduce the main features of the overall system. Following this idea, in this manuscript we built three models to study the reaction between DPA and  $O_2$ . The three models, and their analogy with DPA-CoA +  $O_2$  system, are depicted in Fig. 2. In the largest system the CoA moiety is replaced by a methyl group. For the middle size system, the DPA moiety is also simplified to S-methyl-but-3-enethioate (hereinafter S-methyl-butenthioate), and in the smallest system to 1-oxobut-3-en-2-ide (hereinafter oxobutenide). To maintain the effect of the chemical bond, hydrogen atoms were added along the “broken” bonds.

For each of the models, minimum energy path (MEP) along the distance between  $C_\alpha$  and  $O_2$  ( $r_{C_\alpha O}$ ) were obtained at a B3LYP/aug-cc-pVDZ calculation level of theory as implemented in Gaussian16.<sup>48</sup> Following the same procedure of Ref. 42, constraints were applied to keep the planarity of all the C atoms, and better reproduce the effect of the protein scaffold on the reaction. These constraints are only important at small  $r_{C_\alpha O}$  distances, as for larger  $r_{C_\alpha O}$  the molecule is intrinsically planar. No constraints were applied to the H atom bonded to  $C_\alpha$ . Starting from the optimized geometries, the following methods were used to characterize the reaction paths: broken-symmetry DFT, SHCI, and PNO-RCCSD(T)-F12.

### A. Broken-symmetry DFT

Broken-Symmetry DFT method (BS-DFT)<sup>49–52</sup> is an affordable method to describe open-shell singlets within a monodeterminantal scenario. This method, widely used

for the treatment of diradical systems,<sup>39,53–58</sup> introduces the breakdown of the spatial and spin symmetry in decoupled alpha and beta spin-orbital contributions. Then, the splitting of the spin contribution and the frontier orbital mixing introduces the correction of the static correlation. The main advantage of this method against multi-reference methods is that it does not require the selection of any active space, and it is considerably less expensive, making possible the study of significantly larger systems.

BS-DFT methodology may suffer from the spin contamination which arises intrinsically from the method splitting procedure. It is possible, however, to correct it using Yamaguchi’s formula.<sup>59–61</sup> It consists on a spin projection (SP) method, in which the spin symmetry for the open-shell singlet is recovered. In a recent paper, Visitsatthawong *et al.*<sup>39</sup> introduced this correction for the study of the oxygen activation in a flavin-dependent monooxygenase by means of Eqs. (1) y (2),<sup>59</sup>

$$E_{SP}^{\text{singlet}} = E^{\text{singlet}} + C_{SC} [E^{\text{singlet}} - E^{\text{triplet}}] \quad (1)$$

$$C_{SC} = \frac{\langle S^2 \rangle^{\text{singlet}}}{\langle S^2 \rangle^{\text{triplet}} - \langle S^2 \rangle^{\text{singlet}}} \quad (2)$$

where,  $E_{SP}^{\text{singlet}}$  is the corrected open-shell singlet energy,  $E^{\text{singlet}}$ , the open-shell singlet electronic energy obtained with BS-DFT and  $E^{\text{triplet}}$ , the triplet electronic energy.

### B. Semistochastic Heat-Bath Configuration Interaction

SHCI method involves two stages. First, the wavefunction is optimized selecting iteratively a set of determinants until meeting convergence. These determinants, which are pre-sorted according to their energy, are filtered out below a threshold ( $\epsilon_1$ ). Then, the Hamiltonian is diagonalized only in the subspace of these determinants and the total energy is calculated. In the second stage, the second-order correction to the variational energy,  $E_p$ , is computed using multi-reference Epstein-Nesbet perturbation theory. The value of  $\epsilon_1$  determines the accuracy and the computational cost of the calculations. Values of  $\epsilon_1$  below  $10^{-8} E_H$  could yield highly accurate results, while results for  $\epsilon_1 > 10^{-2} E_H$  are probably meaningless. Following the procedure described in Ref. 24, we carried out calculations for different values of  $\epsilon_1$ , and finally extrapolated to  $\epsilon_1 \rightarrow 0$ . The lowest value of  $\epsilon_1$  in this work is  $5 \cdot 10^{-3} E_H$  for orbital optimization (only variational energy) and  $5 \cdot 10^{-4} E_H$  for FCI extrapolation.

SHCI calculations have been carried out in PySCF software,<sup>62</sup> interfaced to Dice and Arrow.<sup>63,64</sup> The active space selected included all the valence orbitals, i.e., the 2s and 2p orbitals for the C and O, the 1s of the H, and the 3s and 3p of the S. For the smallest system, oxobutenide +  $O_2$ , this active space spans 40 electrons in 33 orbitals, (40,33). For the middle-size system, S-methyl-butenthioate +  $O_2$ , it results in a (52,43) active

space. For the largest system, DPA + O<sub>2</sub>, it involves an (82,69) active space which is, to the best of our knowledge, one of the largest active space calculations up to date, spanning a 10<sup>38</sup> Slater determinants space. To better achieve convergence, calculations were carried out in three steps: a first step in which only active space orbitals were optimized using  $\epsilon_1 = 5 \cdot 10^{-3} E_H$ , a second step where the full set of orbitals were reoptimized at the same value of  $\epsilon_1$  (vHCISCF), and a third step in which starting from the fully optimized set of orbitals, calculations are repeated for  $\epsilon_1 = 1 \cdot 10^{-3}$ , and  $5 \cdot 10^{-4} E_H$  and finally extrapolated towards  $\epsilon_1 \rightarrow 0$ .

### C. PNO-LCCSD(T)-F12

Coupled cluster methods, and in particular CCSD(T), are among the most accurate single-reference methods.<sup>47,65,66</sup> However, its computational cost impairs its use for medium-size systems. To reduce the computational burden of the calculations, some approximations have been proposed. One of these approximations consists in restricting the electron correlation within the whole system to local contributions. This is performed by means of the Pair Natural Orbitals (PNOs)<sup>47,67,68</sup> which allow the reduction of the virtual space to an independent set of virtual orbitals for each occupied orbital pair. For systems for which CCSD(T) calculations are feasible, both methods are in very good agreement, with relative energies differing in around 0.2 kcal/mol. In this manuscript, we will use PNO-LCCSD(T) method as the reference single-reference method. PNO-LCCSD(T) calculations were carried out using Molpro 2020<sup>69</sup> with a cc-pVDZ-F12 basis set.

### D. Electron-density descriptors

Electron density descriptors are a useful tool to understand how electron density rearranges along a reaction pathway. In these methods, the total electron density of a system splits into atomic contributions by means of several fragmentation methods.<sup>70,71</sup> In this frame, the electron sharing between a set of atomic centers can be calculated with the so-called electron delocalization indices (DIs).<sup>72–74</sup> These indices can be tightly related with some chemical concepts such as bond orders, atomic charges and aromaticity. Herein DIs were calculated as implemented in the NDELOC code.<sup>75</sup>

Here, the fragmentation procedure chosen according to the size of the system was the Mulliken scheme.<sup>71</sup> To avoid a wrong definition of the Hilbert space, calculations were performed without a diffuse basis set. Also, during this work, we will focus on the analysis of two-order delocalization indices which measure the number of electron

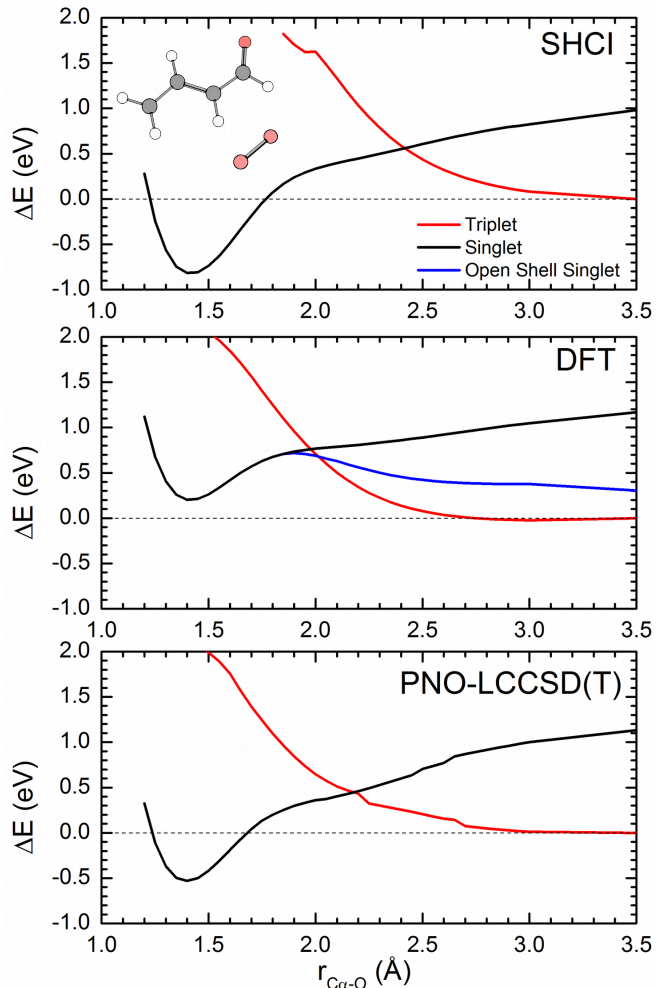


FIG. 3: Potential energy curves for the addition of O<sub>2</sub> to oxobutenide anion at SHCI/aug-cc-pVDZ (Top panel), DFT B3LYP/aug-cc-pVDZ (Medium panel) and PNO-LCCSD(T)-F12/cc-pVDZ-F12 level (Bottom panel). The energy of the triplet state for the O<sub>2</sub> + oxobutenide asymptote is taken as zero reference energy.

between two regions as follows,

$$\delta_{AB} = \sum_{\alpha \in A} \sum_{\beta \in B} (PS)_{\alpha\beta} (PS)_{\beta\alpha} \quad (3)$$

where, A and B are the atomic centers,  $\alpha$  and  $\beta$  the basis functions, and  $S$  and  $P$  the overlap and density matrices, respectively.

## III. RESULTS

We will start by analyzing the results obtained for the smallest system, in which DPA-CoA is replaced by oxobutenide (see Fig. 2). The reaction paths for the peroxidation of oxobutenide as a function of  $r_{C\alpha O}$  are depicted in Fig. 3 for the three different methods used:

SHCI, DFT (using the B3LYP functional) and PNO-LCCSD(T).

Regardless of the method considered, all the sets of potential energy curves (PECs) share the same main features. At large  $r_{C_\alpha O}$  distances (where  $O_2$  and oxobutenide are far from each other) the ground state is a triplet, while at short  $r_{C_\alpha O}$ , where the peroxide is formed, the triplet is a repulsive state, and the system is only stable as a singlet. With decreasing  $r_{C_\alpha O}$  the energy of the triplet PEC increases and, simultaneously, there is a stabilization of the singlet state.

The PEC for the triplet state at SHCI level of theory is more repulsive than for the single-reference methods, causing that the distance at which the singlet and triplet curves cross is significantly larger at SHCI level ( $r_{C_\alpha O}=2.4$  Å) than for PNO-LCCSD(T) and DFT (2.2 Å, and 2.0 Å, respectively). The energy of the crossing point is significantly larger for DFT (0.69 eV) than for PNO-CCSD(T) (0.43 eV), and SHCI (0.57 eV). We should remark that due to the lack of analytical gradients for SHCI method, we did not attempt to optimize the minimum energy crossing point (MECP), which in any case should be lower than the crossing points of Fig. 2. In Ref. 42 the MECP was calculated for the reaction between  $O_2$  and S-methyl-but-3-enethioate and it was located at  $r_{C_\alpha O}=2.1$  Å, and its energy was 0.7 eV above the triplet asymptote.

Regarding the singlet PECs, SHCI and PNO-LCCSD(T) singlet curves are similar, and mainly differ on the depth of the peroxide well, being the SHCI singlet the most stable. For both methods, the energy of the singlet states monotonically decreases with decreasing  $r_{C_\alpha O}$ . As it was noticed in the methods section, to account for the possible effect of the protein on the enolate, constraints were applied to keep the planarity of the C scaffold of the molecule<sup>42</sup> so, if relaxed, the singlet state would be significantly more stable regardless of the method. As no constraints are applied to the H bonded to  $C_\alpha$ , it moves outside the plane of the molecule for  $r_{C_\alpha O} < 2.2$  Å. It is a consequence of the change of hybridization from a  $sp^2$  to a  $sp^3$ .

We calculated two different singlet states using DFT: a restricted singlet state, and the open-shell singlet using the Broken-Symmetry framework. At close  $r_{C_\alpha O}$  distance, where the peroxide is stable, the wavefunction is well approximated by just one configuration and, accordingly, the restricted and open-shell singlet states converge to the same solution. The peroxide minimum is less stable for DFT than for SHCI or PNO-LCCSD(T). For  $r_{C_\alpha O} > 1.9$  Å, the broken-symmetry solution does not longer overlaps the restricted singlet PEC and at large  $r_{C_\alpha O}$  the difference between these states is about 0.8 eV. The open-shell character of  $O_2$  ( $\pi_g^*$ )<sup>2</sup> of the first singlet state is the actual cause of this divergence, as the restricted solution does not longer represents correctly the system. In the broken-symmetry framework, an unrestricted wavefunction is constructed in which the mixture

of the frontier orbitals is allowed (an analog to a multi-reference wavefunction with 2 electrons in 2 orbitals). According to this, one would expect that the open-singlet PEC would reproduce the behavior of the SHCI or PNO-LCCSD(T) singlet PECs at large  $r_{C_\alpha O}$  better than its restricted counterpart. However, that is not the case, and the energy gap between the open-singlet and triplet PEC at large  $r_{C_\alpha O}$  is significantly smaller than for any other method. Incidentally, it creates a barrier in the singlet PEC which is not present in any other method. For the three model systems studied, curves obtained using MP2 showed discontinuities and were excluded from the comparison. Similar errors were obtained when we attempt to converge CASPT2 calculations. We believe that this may be caused by the appearance of intruder states, even in MP2.

In the absence of a barrier for the singlet state, the energy gap between singlet and triplet states may have a strong influence on the reactivity. At very large  $r_{C_\alpha O}$  distances, we can assume that there is no interaction between oxobutenide and  $O_2$ . Under these circumstances, the negative charge should be located in the enolate, and the difference between the singlet and triplet PECs should correspond to that for isolated  $O_2$ . Experimentally, the energy difference between the first triplet ( $^3\Sigma_g^-$ ) and singlet ( $^1\Delta_g$ ) states of  $O_2$  is 0.981 eV.<sup>76</sup> The energy gap obtained with SHCI (1.20 eV), PNO-LCCSD(T) (1.19 eV), and standard DFT (1.19eV) is in a reasonable agreement with the experimental datum, but for Broken-Symmetry DFT the energy difference is just 0.36 eV, clearly underestimating the experimental energy gap. It is worth mentioning that spin-contamination was found to be very small and correction using Yamaguchi's formula (see methods) barely affects the result.

To discard the hypothesis that the differences in the energy gap were not intrinsic of the electronic structure method but caused by some peculiarities of the system, we calculated the energy difference between triplet and singlet  $O_2$  with the methods employed in this study applying a small set of basis sets, and additionally other widely used DFT functionals. The energy gaps, collected in Tables I and II, are somewhat insensitive of the basis set but strongly depend on the electronic structure method used. Regardless of the DFT functional used (even for M06-HF that includes a 100 % Hartree-Fock exchange) there is a clear underestimation of the energy gap when the Broken-Symmetry framework is used, and a clear overestimation of the energy gap when it is not used. Our results agree with those by Kitagawa and co-workers<sup>77</sup> who reported that the energy-gap for carbene ( $CH_2$ ) calculated using broken symmetry methods along with hybrid functionals are clearly underestimated. It should be noticed, however, that the Broken-Symmetry in combination with HF yields values of the energy-gap close to the experimental ones.

According to Tables I and II, the triplet-singlet energy gap is best reproduced by SHCI. However, CCSD(T) and PNO-LCCSD(T) also provide reasonable values for the

TABLE I: Energy gap between (closed-shell) singlet and triplet  $O_2$  for a series of methods and basis sets. Energies are given in eV. For PNO-LCCSD(T), the basis set used was cc-pVDZ-F12.

$\Delta E_{ST}$	B3LYP	B3PW91	M062X	M06HF	CAM-B3LYP	HF	MP2	CCSD(T)	PNO-LCCSD(T)
3-21G	1.712	1.759	1.638	1.172	1.738	2.229	1.257	1.167	
6-31+G(d,p)	1.674	1.729	1.614	1.405	1.698	2.316	1.534	1.335	
aug-cc-pVDZ	1.675	1.728	1.613	1.405	1.699	2.317	1.540	1.336	*1.259
aug-cc-pVTZ	1.667	1.723	1.608	1.401	1.689	2.321	1.477	1.295	

TABLE II: Energy gap between (open-shell) singlet and triplet  $O_2$ . Energies in eV. For DFT, HF, and MP2 methods, open-shell singlet was calculated using the broken-symmetry framework, and the energy was corrected using Yamaguchi's formula. For SHCI, we used a (12,8) active space including the 2s and 2p electrons and orbitals.

$\Delta E_{ST}$	B3LYP	B3PW91	M062X	M06HF	CAM-B3LYP	HF	MP2	SHCI
3-21G	0.469	0.472	0.595	0.264	0.487	0.858	0.827	0.987
6-31+G(d,p)	0.459	0.463	0.619	0.558	0.476	0.865	0.526	0.824
aug-cc-pVDZ	0.454	0.458	0.623	0.557	0.471	0.850	0.464	0.966
aug-cc-pVTZ	0.445	0.450	0.637	0.571	0.463	0.836	0.341	0.943

energy gap. In Ref. 42, the energy gap was also calculated using MRCI and a small active space (4 electrons in 3 orbitals) and the value obtained was 1.13 eV. It should be mentioned that, for DFT calculations, there are some discrepancies between the energy-gap obtained for  $O_2$ , and for our model at large  $r_{C_{\alpha}O}$ . As we will discuss later, this discrepancy arises from the charge of the system, which in DFT is not fully located on the enolate.

For strong-correlated systems single-configuration methods may fail due to the lack of static correlation. To assess the reliability of the PNO-LCCSD(T) calculations, we calculated the diagnostic  $T_1$ .<sup>78</sup>  $T_1$  is the most widely used diagnostic tool to verify if non-dynamical correlation effects are important enough to compromise CCSD calculations. It is assumed that calculations for which  $T_1 < 0.02$  are reasonably described using a CCSD approach. For open-shell systems, it has been suggested that this threshold value may be closer to 0.044.<sup>79</sup> When we calculated the  $T_1$  along the reaction path, we observed that both the singlet minimum and the triplet asymptote are reasonably described with the PNO-LCCSD(T) method as they have associated values under the 0.02 threshold ( $T_1 \simeq 0.015$ -0.017). However,  $T_1$  values rise in the crossing region, especially for the triplet PEC whose typical  $T_1$  is larger than 0.03.

In Fig. 4 we show the PECs corresponding to the addition of  $O_2$  to S-methyl-butenthioate (see Fig. 2). S-methyl-butenthioate was the system used in Ref. 42 to simulate the addition of  $O_2$  to DPA-CoA using MRCI calculations with a small active-space (4 electrons in 3 orbitals). Using this small active-space it was possible to determine the reaction paths corresponding to all the 9 states that arise from that active space, showing that the peroxide well correlates with the  $O_2^{2-}$  asymptote, thus requiring a double electron transfer. In Ref. 42 SHCI calculations were carried out using an active space of (18,23), significantly smaller than the active space used here (52,43). Although the curves using (18,23) and (52,43) are similar, the peroxide well is three times deeper

for the larger active space.

The curves calculated for peroxidation of S-methyl-butenthioate using SHCI, DFT, PNO-LCCSD(T) are in a good qualitative agreement and are also similar to those obtained for the smallest system. The singlet PECs are less attractive for peroxidation of S-methyl-butenthioate than for peroxidation of oxobutenide, and as a consequence, the energies of the crossing points are slightly larger: 0.73 eV for SHCI, 0.77 eV for DFT, and 0.60 eV at PNO-LCCSD(T).

Results for the largest model (DPA, see Fig. 2) are shown in Fig. 5. Again there is a qualitative agreement between all methods, and they are also similar to those obtained for smaller systems. For this system, the singlet PEC is less attractive, leading to a shallower peroxide minimum, in particular using DFT. The crossing points are also higher in energy than for the smaller systems, in particular for SHCI method, as the energy gap between the triplet and singlet PEC is larger than that obtained for  $O_2$  (see Table II). The discrepancy in the singlet-triplet energy gap comes from the selection of the active space, which is selected at the smallest  $r_{C_{\alpha}O}$  and then the wavefunction is propagated to larger  $r_{C_{\alpha}O}$ . Using this procedure the active space is robust, but it originates a larger energy gap at the largest  $r_{C_{\alpha}O}$ .

Regarding the PNO-LCCSD(T) calculations, values of  $T_1$  diagnosis are similar for the three systems at large  $r_{C_{\alpha}O}$  ( $T_1 \simeq 0.015$ -0.017) while around the crossing region, although presenting larger values, they decrease with the size of the system. For  $DPA^- + O_2$ , on the singlet PEC they are typically below 0.02 and only around  $r_{C_{\alpha}O} = 1.8 \text{ \AA}$ , they go up to 0.022. On the triplet PEC,  $T_1$  values are also smaller than for the smallest models and are always below 0.03. These results suggest that even for this intrinsically multi-surface reaction, it is possible to rely on the CCSD(T) treatment. This conclusion is reinforced by the good agreement between the PNO-LCCSD(T) and SHCI PECs.

To get more insight about the reaction mechanism, in



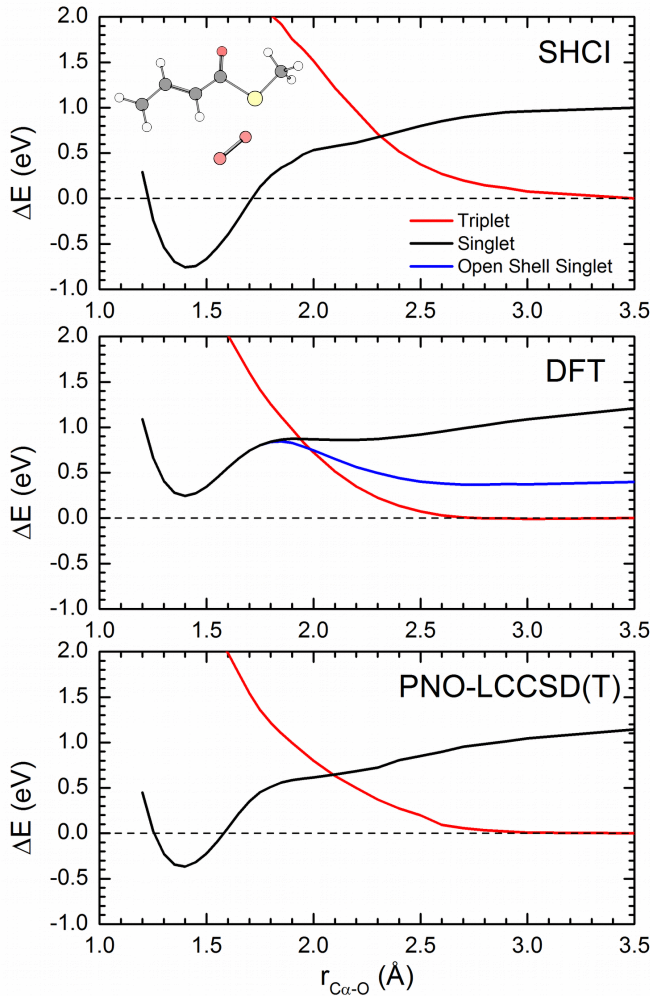


FIG. 4: Potential energy curves of the S-methyl-butenthioate anion reaction with  $O_2$ , at SHCI/aug-cc-pVDZ (Top panel) B3LYP/aug-cc-pVDZ (Medium panel) and PNO-LCCSD(T)-F12/cc-pVDZ-F12 level (Bottom panel). The energy of the triplet state for the  $O_2 +$  S-methyl-butenthioate is taken as zero reference energy.

Fig. 6 we show the two-center electron delocalization indices along the reaction path for the DFT singlet and triplet PECs of the  $DPA^- + O_2$  reaction. They represent the bond-order, we calculated these indexes for the two covalent bonds that are expected to change throughout the reaction: i) that between the two O atoms ( $\delta_{OO}$ ) and ii) that between  $C_\alpha$  and O ( $\delta_{C_\alpha O}$ ). At large  $r_{C_\alpha O}$  the values of  $\delta_{OO}$  and  $\delta_{C_\alpha O}$  are almost constant, the latter close to zero, showing that there is almost no interaction between the enolate and  $O_2$ . When the two reactants approach, in particular for  $r_{C_\alpha O} < 2.2$  Å,  $\delta_{OO}$  shrinks while  $\delta_{C_\alpha O}$  rises. This indicates that the peroxidation occurs via a concerted mechanism in which electron density is transferred from the O-O to the  $C_\alpha$ -O bond. At  $r_{C_\alpha O} = 1.4$  Å, at the equilibrium distance of the peroxide, both indices are similar, showing “bond-orders” that

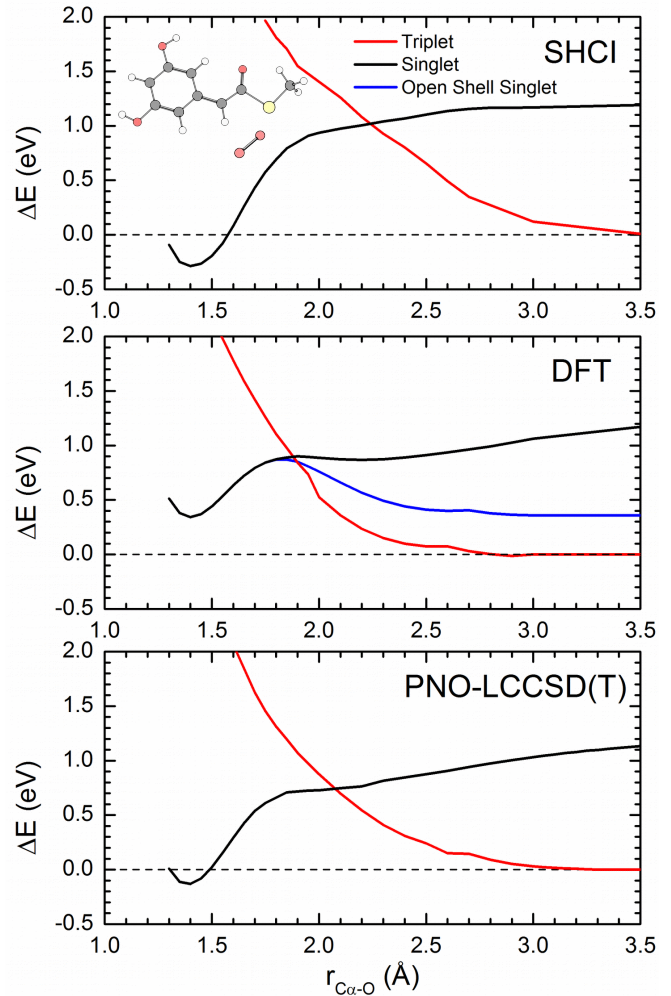


FIG. 5: Potential energy curves for peroxidation of  $DPA^-$ . Calculations are carried out at SHCI/aug-cc-pVDZ B3LYP/aug-cc-pVDZ, and PNO-LCCSD(T)-F12/cc-pVDZ-F12 level. The energy of the triplet state for the  $O_2 + DPA^-$  is taken as zero reference energy.

are slightly below those for standard single-bonds (such as the C-O bond in ether). There are two aspects that deserve further discussion: first, regardless of  $r_{C_\alpha O}$ ,  $\delta_{OO}$  is larger for the triplet than for the singlet PEC while the opposite occurs for  $\delta_{C_\alpha O}$ . Hence, intersystem crossing intrinsically involves electron transfer from the O-O to the  $C_\alpha$ -O bond. This result agrees with predictions based on the Mulliken charges, that showed that intersystem crossing was associated with some electron density transfer from  $O_2$  to the enolate<sup>42</sup>. Second, it is also interesting that  $\delta_{OO}$  never approaches the value obtained for isolated  $O_2$ . This discrepancy arises from an excess of electronic charge on  $O_2$ , that is not zero even at very large  $r_{C_\alpha O}$ . This is the reason why the triplet-singlet energy gap for DFT calculations is lower for our systems than for isolated  $O_2$ . To ascertain this statement, the singlet-triplet energy gap and delocalization indices were recalculated

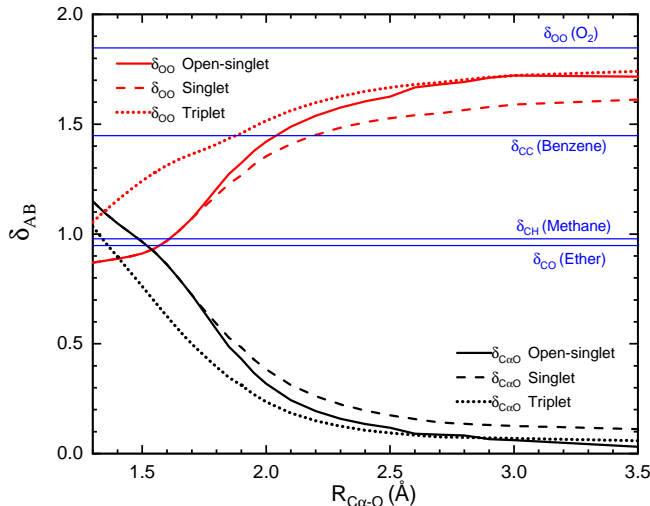


FIG. 6: Delocalization indices evolution along the addition of  $O_2$  to  $DPA^-$  as a function of the  $C_\alpha$ -O distance calculated at B3LYP/cc-pVDZ (top panel) level of theory. Reference values are shown as horizontal solid blue lines.

for an uncharged model system, in which the charge was removed by adding a second hydrogen to  $C_\alpha$ , that now shows an  $sp^3$  hybridization. On this system, the triplet-singlet energy gap is coincident with that obtained for  $O_2$ , and that is also the case of  $\delta_{OO}$ .

Using the IBO/IAO formalism,<sup>10,80</sup> we represented the orbitals whose nature changed significantly along the addition of  $O_2$  to  $DPA^-$ . The results (shown in Fig. S1) reveal that for both singlet and triplet states, they are related to the formation of the bond between O and  $C_\alpha$ . On the singlet state, they correspond to the mixing between one  $\pi$  and one  $\pi^*$  of  $O_2$  and  $p_z$  orbital of  $C_\alpha$ . On the triplet state, however, mixing between those orbitals is not significant, explaining why lack of stability of the triplet peroxide.

#### IV. CONCLUSIONS

Throughout this article, we have studied the addition of  $O_2$  to enolates using two recently developed theoretical approaches such as SHCI and PNO-LCCSD(T), and compared them to DFT methods, widely used to study these reactions on biological media. On top of the fundamental importance of these intrinsically multi-configuration systems, the addition of  $O_2$  to enolates is the key step for the enzymatically catalyzed cofactorless incorporation of  $O_2$  into organic substrates. To shed light on the reaction mechanism, we calculated the delocalization indices along the reaction path. Our results show that the reaction follows a concerted mechanism in which hopping from the triplet to the singlet potential energy surface is associated with an electron density

transfer from the O-O to the C-O bond.

To study this reaction we used three different model systems of different sizes. The potential energy curves (PECs) obtained using these three methodologies and the three models are qualitatively similar. Regardless of the model used, the singlet PEC is less attractive for the largest systems, leading to crossing points of larger energies. DFT calculations cast some doubts due to inaccuracies in the singlet-triplet gap at large  $r_{C_\alpha O}$  where  $O_2$  and enolates are not interacting, and a comparison with experimental values are possible. When the energy of the singlet state is corrected using the broken-symmetry framework, the energy gap is in turn too small, causing a barrier on the singlet PEC, which is not predicted for any other method.

The singlet PECs obtained using SHCI and PNO-LCCSD(T) are very similar to each other, while the triplet PEC is more repulsive for SHCI. The  $T_1$  diagnosis was evaluated showing that PNO-LCCSD(T) calculations are reliable both when the peroxide is formed (small  $r_{C_\alpha O}$ ) and in the triplet asymptote. In the crossing region,  $T_1$  diagnosis is larger, especially for the smallest enolate, while for the largest one, its highest value is 0.022 for the singlet and 0.028 for the triplet, slightly above the usually accepted empirical threshold value of 0.02 for closed-shell systems and well below the proposed threshold of 0.044 for open-shell systems. It suggests that PNO-LCCSD(T) results may be accurate for these systems, opening the gate to use this method for the study of this kind of reactions.

The study of these reactions using multi-reference methods is challenging, requiring a very large active space. Using SHCI we could carry out calculations using an active space that includes all the valence orbitals and electrons, 82 electrons in 69 orbitals, far beyond the possibility of the classical CAS-like methods. SHCI PECs are in a reasonable agreement with PNO-LCCSD(T) in the regions where PNO-LCCSD(T) is expected to reproduce faithfully the behavior of the system and shows energy crossing points of higher energy. Altogether, our results suggest that SHCI is a good approach to deal with intrinsic multisurface problems, as it permits to include oversized active spaces which makes easier to select a robust active space along the reaction path, and also to recover most of the correlation energy, which makes the results more accurate.

#### V. DEDICATION

Throughout our scientific careers we have worked along with many outstanding female Physical Chemists (and learn from them). We would like to dedicate this article to Prof. Carmen Barrientos, Dr. Béatrice Bussery-Honvault, Dr. Nadja Doslic, Dr. Shirin Faraji, Dr. Susana Gómez-Carrasco, Dr. Mari-Luz Hernández, Dr. Marta Menéndez, Dr. Ángeles Peña-Gallego, Dr. Edina Rosta, and Prof. Carmela Valdemoro.



## VI. SUPPLEMENTARY MATERIAL

See the supplementary material for the geometries of oxobutenide, S-methyl-butenthioate and DPA anions + O<sub>2</sub>, the raw data for the energy profiles of oxobutenide, S-methyl-butenthioate and DPA anions and the SHCI extrapolation raw data for the three mentioned anions in their singlet and triplet states. It is as well included a representation of the orbitals whose nature changes more significantly along the addition of O<sub>2</sub> to DPA<sup>-</sup> using IBO/IAO formalism.

## VII. ACKNOWLEDGMENTS

Funding by the Spanish Ministry of Science and Innovation (grant FIS2017-83473-C2-P, PID2019-107115GB-C21, and PGC2018-096444-B-I00) is also acknowledged. P.O. acknowledges grant EDU/601/2020 (Junta de Castilla y León and European Social Fund) and thank James E. T. Smith for his practical hints and helpful comments on PySCF code. The authors acknowledge funding by Fundación Salamanca City of Culture and Knowledge (program for attracting scientific talent to Salamanca). The authors thank Dr. M. Mandado from University of Vigo for kindly providing the NDELOC code used to calculate the electron density descriptors.

## VIII. DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## IX. BIBLIOGRAPHY

- <sup>1</sup>D. Cremer, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2011, **1**, 509–530.
- <sup>2</sup>R. G. Parr and W. Yang, in *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, Oxford, 1989.
- <sup>3</sup>K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479–483.
- <sup>4</sup>B. O. Roos, P. R. Taylor and P. E. Sigbahn, *Chem. Phys.*, 1980, **48**, 157–173.
- <sup>5</sup>J. Townsend, J. K. Kirkland and K. D. Vogiatzis, in *Mathematical Physics in Theoretical Chemistry*, Elsevier, 2019, pp. 63–117.
- <sup>6</sup>K. D. Vogiatzis, D. Ma, J. Olsen, L. Gagliardi and W. A. de Jong, *J. Chem. Phys.*, 2017, **147**, 184111.
- <sup>7</sup>V. Sauri, L. Serrano-Andres, A. R. M. Shahi, L. Gagliardi, S. Vancoillie and K. Pierloot, *J. Chem. Theory Comput.*, 2011, **7**, 153–168.
- <sup>8</sup>G. Cárdenas and J. J. Nogueira, *Int J Quantum Chem.*, 2021, <https://doi.org/10.1002/qua.26533>.
- <sup>9</sup>J. J. Bao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2019, **15**, 5308–5318.
- <sup>10</sup>G. Knizia and J. E. M. N. Klein, *Angew. Chem. Int. Ed.*, **54**, 5518–5522.
- <sup>11</sup>S. Mai and L. González, *Angew. Chem. Int. Ed.*, 2020, **59**, 16832–16846.
- <sup>12</sup>D. Casanova and A. I. Krylov, *Phys. Chem. Chem. Phys.*, 2020, **22**, 4326–4342.
- <sup>13</sup>G. H. Booth, A. J. W. Thom and A. Alavi, *J. Chem. Phys.*, 2009, **131**, 054106.
- <sup>14</sup>G. H. Booth, D. Cleland, A. J. W. Thom and A. Alavi, *J. Chem. Phys.*, 2011, **135**, 084104.
- <sup>15</sup>N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon and K. B. Whaley, *J. Chem. Phys.*, 2016, **145**, 044112.
- <sup>16</sup>W. Liu and M. R. Hoffmann, *Theor. Chim. Acta*, 2014, **133**, 141.
- <sup>17</sup>G. K.-L. Chan and S. Sharma, *Ann. Rev. Phys. Chem.*, 2011, **62**, 465–481.
- <sup>18</sup>K. H. Marti and M. Reiher, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6750–6759.
- <sup>19</sup>S. Sharma and G. K.-L. Chan, *J. Chem. Phys.*, 2012, **136**, 124121.
- <sup>20</sup>S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi and C. J. Umrigar, *J. Chem. Theory Comput.*, 2017, **13**, 1595–1604.
- <sup>21</sup>A. A. Holmes, N. M. Tubman and C. J. Umrigar, *J. Chem. Theory Comput.*, 2016, **12**, 3674–3680.
- <sup>22</sup>J. E. T. Smith, B. Mussard, A. A. Holmes and S. Sharma, *J. Chem. Theory Comput.*, 2017, **13**, 5468–5478.
- <sup>23</sup>J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar and J. Gauss, *J. Phys. Chem. Lett.*, 2020, **11**, 8922–8929.
- <sup>24</sup>A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma and P. M. Zimmerman, *J. Phys. Chem. A*, 2018, **122**, 2714–2722.
- <sup>25</sup>T. D. Bugg, *Tetrahedron*, 2003, **59**, 7075 – 7101.
- <sup>26</sup>S. Sahu and D. P. Goldberg, *J. Am. Chem. Soc.*, 2016, **138**, 11410–11428.
- <sup>27</sup>A. J. Jasiewicz and L. Que, *Chem. Res.*, 2018, **118**, 2554–2592.
- <sup>28</sup>M. Wikstrom, K. Krab and V. Sharma, *Chem. Res.*, 2018, **118**, 2469–2490.
- <sup>29</sup>S. S. Chaturvedi, R. Ramanan, N. Lehnert, C. J. Schofield, T. G. Karabencheva-Christova and C. Z. Christov, *ACS Catal.*, 2020, **10**, 1195–1209.
- <sup>30</sup>R. Suardiaz, P. G. Jambrina, L. Masgrau, A. González-Lafont, E. Rosta and J. M. Lluch, *J. Chem. Theory Comput.*, 2016, **4**, 2079.
- <sup>31</sup>K. P. Jensen and U. Ryde, *J. Biol. Chem.*, 2004, **279**, 14561–14569.
- <sup>32</sup>P. Chaiyen, M. W. Fraaije and A. Mattevi, *Trends Biochem Sci.*, 2012, **37**, 373 – 380.
- <sup>33</sup>E. Romero, J. R. Gomez Castellanos, G. Gadda, M. W. Fraaije and A. Mattevi, *Chem. Res.*, 2018, **118**, 1742–1769.
- <sup>34</sup>S. Bui and R. A. Steiner, *Curr Opin Struct Biol.*, 2016, **41**, 109 – 118.
- <sup>35</sup>X. Fretzner and R. A. Steiner, *Appl Microbiol Biotechnol.*, 2010, **4186**, 791–804.
- <sup>36</sup>S. Bui, D. von Stetten, P. G. Jambrina, T. Prange, N. Colloc'h, D. de Sanctis, A. Royant, E. Rosta and R. A. Steiner, *Angew. Chem. Int. Ed.*, 2014, **53**, 13710.
- <sup>37</sup>P. J. Silva, *PeerJ*, 2016, **4**, e2805.
- <sup>38</sup>F. G. Cantu Reinhard, J. L. DuBois and S. P. de Visser, *J. Phys. Chem. B*, 2018, **122**, 10841–10854.
- <sup>39</sup>S. Visitsatthawong, P. Chenprakhon, P. Chaiyen and P. Surawatanawong, *J. Am. Chem. Soc.*, 2015, **137**, 9363–9374.
- <sup>40</sup>B. Minaev, *Chem. Phys.*, 2019, **521**, 61 – 68.
- <sup>41</sup>D. Wei, X. Huang, Y. Qiao, J. Rao, L. Wang, F. Liao and C.-G. Zhan, *ACS Catal.*, 2017, **7**, 4623–4636.
- <sup>42</sup>P. Ortega, A. Zanchet, C. Sanz-Sanz, S. Gómez-Carrasco, L. González-Sánchez and P. G. Jambrina, *Chem. Eur. J.*, 2021, **21**, 1700–1712.
- <sup>43</sup>C. C. Tseng, F. H. Vaillancourt, S. D. Bruner and C. T. Walsh, *Chem. Biol.*, 2004, **11**, 1195 – 1203.

- <sup>44</sup>P. F. Widboom, E. N. Fielding, Y. Liu and S. D. Bruner, *Nature*, 2007, **447**, 342–345.
- <sup>45</sup>K. Li, E. N. Fielding, H. L. Condurso and S. D. Bruner, *Acta Cryst.*, 2017, **D73**, 573–580.
- <sup>46</sup>J. N. Harvey, *WIREs Comput Mol Sci*, 2014, **4**, 1–14.
- <sup>47</sup>Q. Ma and H.-J. Werner, *Journal of Chemical Theory and Computation*, 0, **0**, null.
- <sup>48</sup>M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb and J. R. Cheeseman et al., *Gaussian-16 Revision A.03*, 2016, Gaussian Inc. Wallingford CT.
- <sup>49</sup>H. Shull and P. Löwdin, *J. Chem. Phys.*, 1956, **25**, 1035–1040.
- <sup>50</sup>L. Noodleman and E. R. Davidson, *Chem. Phys.*, 1986, **109**, 131–143.
- <sup>51</sup>J.-M. Mouesca, *Density Functional Theory–Broken Symmetry (DFT–BS) Methodology Applied to Electronic and Magnetic Properties of Bioinorganic Prosthetic Groups*, Humana Press, 2014, pp. 269–296.
- <sup>52</sup>L. Noodleman, *J. Chem. Phys.*, 1981, **74**, 5737–5743.
- <sup>53</sup>T. Sadhukhan, D. Beckett, B. Thapa and K. Raghavachari, *J. Chem. Theory Comput.*, 2019, **15**, 5998–6009.
- <sup>54</sup>G. Trinquier and J.-P. Malrieu, *Chem. Eur. J.*, 2014, **21**, 814–828.
- <sup>55</sup>E. Louis, E. San-Fabián, G. Chiappe and J. Vergés, *Carbon*, 2019, **154**, 211–218.
- <sup>56</sup>J.-P. Malrieu and G. Trinquier, *J. Chem. Phys.*, 2016, **144**, 211101.
- <sup>57</sup>S. Gil-Guerrero, M. Melle-Franco, Á. Peña-Gallego and M. Mandado, *Chem. Eur. J.*, 2020, **26**, 16138–1614.
- <sup>58</sup>N. Ramos-Berdullas, S. Gil-Guerrero, A. Peña-Gallego and M. Mandado, *Phys. Chem. Chem. Phys.*, 2021, **23**, 4777–4783.
- <sup>59</sup>H. Isobe, S. Yamanaka, S. Kuramitsu and K. Yamaguchi, *J. Am. Chem. Soc.*, 2008, **130**, 132–149.
- <sup>60</sup>K. Yamaguchi, S. Yamanaka, M. Nishino, Y. Takano, Y. Kitagawa, H. Nagao and Y. Yoshioka, *Theor. Chem. Acc.*, 1999, **102**, 328–345.
- <sup>61</sup>Y. Kitagawa, T. Saito and K. Yamaguchi, in *Symmetry (Group Theory) and Mathematical Treatment in Chemistry*, InTech, 2018.
- <sup>62</sup>Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters and G. K.-L. Chan, *PySCF: the Python-based simulations of chemistry framework*, 2017, <https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1340>.
- <sup>63</sup>A. A. Holmes, N. M. Tubman and C. J. Umrigar, *J. Chem. Theory Comput.*, 2016, **12**, 3674–3680.
- <sup>64</sup>J. Li, M. Otten, A. A. Holmes, S. Sharma and C. J. Umrigar, *J. Chem. Phys.*, 2018, **149**, 214110.
- <sup>65</sup>F. Neese, A. Hansen and D. G. Liakos, *J. Chem. Phys.*, 2009, **131**, 064103.
- <sup>66</sup>G. Schmitz and C. Hättig, *J. Chem. Theory Comput.*, 2017, **13**, 2623–2633.
- <sup>67</sup>W. Meyer, *Int. J. Quant. Chem.*, 2009, **5**, 341–348.
- <sup>68</sup>R. Ahlrichs, H. Lischka, V. Staemmler and W. Kutzelnigg, *J. Chem. Phys.*, 1975, **62**, 1225–1234.
- <sup>69</sup>H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz et al., *MOLPRO, version 2020.1, a package of ab initio programs*, 2019, see.
- <sup>70</sup>R. Bader, *Atoms in molecules : a quantum theory*, Clarendon Press, Oxford, 1990.
- <sup>71</sup>R. S. Mulliken, *J. Chem. Phys.*, 1962, **36**, 3428–3439.
- <sup>72</sup>M. Giambiagi, M. S. de Giambiagi and K. C. Mundim, *Struct. Chem.*, 1990, **1**, 423–427.
- <sup>73</sup>P. Bultinck, R. Ponc and S. V. Damme, *J. Phys. Org. Chem.*, 2005, **18**, 706–718.
- <sup>74</sup>K. C. Mundim, M. Giambiagi and M. S. de Giambiagi, *J. Phys. Chem.*, 1994, **98**, 6118–6119.
- <sup>75</sup>M. Mandado, M. J. González-Moa and R. A. Mosquera, *J. Comput. Chem.*, 2006, **28**, 127–136.
- <sup>76</sup>G. Herzberg, *Spectra of Diatomic Molecules*, Van Nostrand Reinhold, New York, 1950.
- <sup>77</sup>Y. Kitagawa, T. Saito, Y. Nakanishi, Y. Kataoka, T. Matsui, T. Kawakami, M. Okumura and K. Yamaguchi, *J. Phys. Chem. A*, 2009, **113**, 15041–15046.
- <sup>78</sup>T. J. Lee and P. R. Taylor, *Int. J. Quantum Chem.*, 1989, **36**, 199–207.
- <sup>79</sup>J. C. Rienstra-Kiracofe, W. D. Allen and H. F. Schaefer, *J. Phys. Chem. A*, 2000, **104**, 9823–9840.
- <sup>80</sup>G. Knizia, *J. Chem. Theory Comput.*, 2013, **9**, 4834–4843.