

Size-extensivity-corrected multireference configuration interaction schemes to accurately predict bond dissociation energies of oxygenated hydrocarbons

Victor B. Oyeyemi,¹ David B. Krisiloff,² John A. Keith,³ Florian Libisch,³ Michele Pavone,⁴ and Emily A. Carter^{3,5,6,a)}

¹Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, USA

²Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA

³Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08544, USA

⁴Department of Chemical Sciences, University of Napoli Federico II, Napoli 80120, Italy

⁵Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544, USA

⁶Andlinger Center for Energy and the Environment, Princeton University, Princeton, New Jersey 08544, USA

(Received 10 November 2013; accepted 1 January 2014; published online 29 January 2014)

Oxygenated hydrocarbons play important roles in combustion science as renewable fuels and additives, but many details about their combustion chemistry remain poorly understood. Although many methods exist for computing accurate electronic energies of molecules at equilibrium geometries, a consistent description of entire combustion reaction potential energy surfaces (PESs) requires multireference correlated wavefunction theories. Here we use bond dissociation energies (BDEs) as a foundational metric to benchmark methods based on multireference configuration interaction (MRCI) for several classes of oxygenated compounds (alcohols, aldehydes, carboxylic acids, and methyl esters). We compare results from multireference singles and doubles configuration interaction to those utilizing *a posteriori* and *a priori* size-extensivity corrections, benchmarked against experiment and coupled cluster theory. We demonstrate that size-extensivity corrections are necessary for chemically accurate BDE predictions even in relatively small molecules and furnish examples of unphysical BDE predictions resulting from using too-small orbital active spaces. We also outline the specific challenges in using MRCI methods for carbonyl-containing compounds. The resulting complete basis set extrapolated, size-extensivity-corrected MRCI scheme produces BDEs generally accurate to within 1 kcal/mol, laying the foundation for this scheme's use on larger molecules and for more complex regions of combustion PESs. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4862159>]

I. INTRODUCTION

There is growing interest in understanding the combustion of oxygenated hydrocarbons due to their roles as renewable additives and fuels.¹⁻³ Oxygenated fuels burn more cleanly than conventional hydrocarbon fuels. However, fundamental details about their combustion remain to be quantified in order to optimize their energy conversion efficiency. Current experimental and theoretical research efforts aim to elucidate their combustion properties, e.g., features that explain auto-ignition and species profiles at various stages of combustion (see Ref. 3 for additional discussion). Kinetic modeling of combustion can help us understand combustion processes, but such modeling requires accurate thermochemistry (bond dissociation energies (BDEs), reaction enthalpies, etc.) and kinetics (temperature- and pressure-dependent rate constants) as input, both of which

are extractable in part from portions of potential energy surfaces (PESs) of the relevant reactions.

Multireference *ab initio* methods can evaluate accurately and consistently PESs that span different reaction intermediates, transition states, and all points in between. The question arises as to how to benchmark such methods against experiment. PESs for polyatomic molecules are not directly measurable; only properties related to critical points on the PES are extractable from experiment and then only in simple cases where the observable definitively corresponds to, e.g., a particular bond-breaking event. Even measurements of a specific elementary rate constant are not directly comparable to a PES barrier height because of complexities in the analysis of pre-exponential factors, tunneling corrections, etc. BDEs are definitive and measurable observables that are relatable to PES critical points (differences in PES minima). Thus, at a minimum, one must ensure that MR methods are able to accurately reproduce measured BDEs before using them to determine other points on the PES such as barrier heights. In this work, we test the ability of techniques based on multireference single and double excitation configuration

^{a)} Author to whom correspondence should be addressed. Electronic mail: eac@princeton.edu. Fax: +1 609 258 5877.

interaction (MRSDCI) and its size-extensive variants to accurately predict BDEs in a variety of combustion-related oxygenated molecules for which reliable experimental and theoretical reference data exist. We are not aware of any other systematic comparison of the various size-extensive variants of MRSDCI for BDEs. Of course, methods that accurately predict BDEs do not guarantee the same accuracy for an entire PES – the breakdown of conventional, single-reference coupled cluster (CC) theory during bond breaking is an obvious example (single reference CC theory describes minima well and so can accurately predict BDEs, except in cases in which static correlation demands an MR treatment, *vide infra*). However, determining the appropriate MR method for BDEs sets the level of theory required for minima on a given PES, which in turn provides a consistent level of theory to evaluate all other points on the PES.

Many quantum-chemistry-based schemes to obtain BDEs exist.⁴ Most use electronic energies from a single reference (i.e., a single determinant zeroth-order wavefunction). Schemes fitting this classification may use variants of Kohn-Sham density functional theory (KS-DFT)^{5,6} or coupled cluster theories⁷ (e.g., CCSD(T)).⁸ Other approaches include composite schemes like the semi-empirical G - n methods (e.g., $G2$,⁹ $G3$ ^{10,11}), complete basis set (CBS- n) extrapolation (e.g., CBS-QB3¹²), the HEAT method,¹³ the focal point method,^{14,15} or the W_n methods.^{16–18} Many of these single reference methods provide a robust treatment of dynamic electron correlation that affords accurate BDEs. However, these methods can be inadequate in cases where one or more equilibrium species have significant multiconfigurational character, i.e., cases where a single determinant does not physically describe the molecule's electronic structure. Radicals derived from oxygenated molecules such as aldehydes, alcohols, and methyl esters with electronegative oxygen atoms, lone pair electrons, and π electrons are examples relevant to combustion that may require an MR description due to, e.g., resonance effects. Computationally efficient and robust MR approaches that capably predict accurate BDEs as well as consistently describe the entire PES are needed for consistent and accurate modeling of any point on the PES that would be accessible during combustion.

Many approaches have emerged over the years that generalize conventional single reference correlated wavefunction theories to handle multireference problems. A few examples are multireference coupled cluster approaches¹⁹ such as the spin-flip,²⁰ equation-of-motion,²¹ and completely renormalized coupled-cluster methods.²¹ Further assessments of these significantly more expensive methods (compared to single reference variants), which are in varying stages of development and applications, are beyond the scope of the present work.

This study addresses several issues relevant to quantum chemical modeling of biodiesel and related oxygenated fuel combustion. We do not contest the accuracy of highly mature, single-reference schemes commonly used in molecular property predictions, including BDEs. The present work rather focuses on (1) ascertaining the degree that MR character affects BDEs in oxygenated molecules, (2) benchmarking multireference configuration interaction (MRCI) approaches

to show which is most accurate for these applications, and lastly (3) outlining challenges faced when using MR methods to determine BDEs. This work aims to furnish understanding of how to systematically model combustion reaction dynamics with MR methods.

As a first approximation, MR effects can be described by the complete active space self-consistent field (CASSCF) method.²² CASSCF describes multiconfigurational systems by systematically including in the ground state wavefunction all possible near-degenerate electronic configurations in the orbital active space (typically those valence orbitals that are changing in the process under consideration). CASSCF is a type of multi-configuration SCF, so the exclusion of dynamic correlation involving excitations to virtual orbitals not included in the CAS limits its accuracy. CASPT2²³ introduces dynamic correlation through the use of second-order many-body perturbation theory on the CASSCF wavefunction. Dynamic correlation can also be included on top of CASSCF via configuration interaction (CI)²⁴ or even coupled cluster expansions.^{25,26} Another MR scheme is the MR correlation-consistent composite approach (MR-ccCA),²⁷ which uses CASPT2 electronic energies as well as other energy contributions including core-valence correlations, spin-orbit coupling, and scalar relativistic effects for improved accuracy. MR-ccCA methods are computationally very expensive (mainly due to poor $O(N^5)$ scaling in CASPT2 and use of up to a quadruple zeta basis set) and have mainly been applied to molecular systems consisting of at most three^{27,28} heavy (non-hydrogen) atoms.

We recently developed our own composite MR scheme that uses CBS-MRSDCI²⁹ to provide the dynamic correlation on top of a CASSCF wavefunction. For simple hydrocarbons, our method predicts chemically accurate BDEs (within ~ 1 kcal/mol relative to reliable experimental values). Advantageously, conventional CI codes using this scheme can be replaced with reduced scaling versions^{30–33} that dramatically improve MRSDCI's $O(N^6)$ scaling to $O(N^3)$ or better, thereby opening the possibility for future applications on molecular systems of up to at least 50 heavy atoms. However, the computational cost of our method (as well as its accuracy) will be largely dependent on the active space used in the initial CASSCF calculation, which can vary depending on the molecular systems involved.

Here, we extend our MR calculation scheme to evaluate oxygenated molecules, which have more complex electronic structures than hydrocarbons. For several reasons (*vide infra*) they require special care in their modeling. Specifically, we demonstrate how to successfully calculate bond energies in oxygenated molecules with varying number of oxygen atoms and the presence of double bonds (C=O) using MR methods. We begin by considering alcohols as the simplest oxygenated species. We then follow by studying aldehydes, acids, and esters, i.e., molecules with C=O bonds. C–H and C–C bonds beta (β) to the C=O warrant special attention in the aldehydes, acids, and esters. Finally, we investigate bond breaking that produces the hydroxyformyl and formyloxyl radicals from formic acid. We assess each bond breaking process using two steps. First, we determine the multiconfigurational (static correlation) CASSCF active space sizes needed to obtain

accurate energies in each class of oxygenated bonds. Second, we compare different CI-based electronic structure methods and determine the relative strengths or weaknesses of each. Finally, we discuss the relative importance of accurate geometries and vibrational frequencies. This study provides a roadmap for addressing technical issues arising in future studies that rigorously treat the multiconfigurational character of alcohol, aldehyde, and biodiesel ester combustion.

II. METHOD: MRSDCI/MRACPF-BASED COMPOSITE APPROACHES

Our previously reported *ab initio* MR scheme²⁹ carried a benefit over single reference BDE calculation schemes by incorporating an *ab initio* multi-configurational correlated wavefunction evaluation of electronic energies, making it able to consistently treat the entire chemical bond-breaking event. The particular correlated wavefunctions we use are MRSDCI^{34,35} and MR averaged coupled pair functional (MRACPF)^{36,37} theories. Our scheme is also extendable to large molecules, such as biodiesel fuels, through reduced scaling MRSDCI/MRACPF methods.³⁰⁻³³ In the following sections, we describe the details of our approach.

A. Geometries and vibrational frequencies

KS-DFT is a single determinant method and as such cannot be relied upon to accurately describe entire PESs. However, it is appropriate for predicting equilibrium geometries and corresponding frequencies, even if is not quantitative for energy differences like bond dissociation energies.³⁸⁻⁴⁰ Unrestricted KS-DFT with the B3LYP^{41,42} exchange-correlation (XC) functional and the 6-311G(2d,p) basis set⁴³ was therefore used to calculate both geometries and vibrational frequencies. The absence of imaginary frequencies confirmed that a true minimum energy structure in each case. Zero point vibrational energies (ZPE) and thermal corrections ($H_{0 \rightarrow 298}$) were then calculated for the optimized structures from scaled harmonic vibrational frequencies⁴⁴ using the ideal gas, rigid rotor, and harmonic oscillator approximations.⁴⁵ We define D_e (electronic energy change upon dissociation), D_0 (enthalpy of dissociation at 0 K), and D_{298} (enthalpy of dissociation at 298 K) for a generic bond A–B as

$$D_e = \Delta E_{\text{MRSDCI/MRACPF}(2)}, \quad (1)$$

$$D_0 = D_e + \Delta \text{ZPE}, \quad (2)$$

$$D_{298} = D_0 + \Delta H_{0 \rightarrow 298}, \quad (3)$$

where ΔZPE and $\Delta H_{0 \rightarrow 298}$ are the zero point energy and thermal correction differences between dissociated fragments A• and B• and the undissociated molecule AB, all calculated at their optimized equilibrium geometries.

B3LYP has received substantial use for nearly two decades as a relatively inexpensive, popular, and respectably accurate hybrid XC functional,³⁸⁻⁴⁰ which nonetheless fails in some applications, e.g., yielding unphysical charge

delocalization.⁴⁶⁻⁴⁸ We will evaluate its performance for geometries and frequencies against results from higher-level calculations for some of the more complex radicals studied here.

B. Multireference electronic structure calculations

CASSCF provides a zeroth-order MR wavefunction, which accounts for static correlation, in the calculation of D_e . However, the CAS must be manually adjusted according to each bond being studied. All calculations reported here use at least a CAS(2e,2o), an active space containing two electrons and two orbitals (bonding and anti-bonding), for each breaking bond. The nature of the bonding (σ) and anti-bonding (σ^*) orbitals changes as a bond dissociates from equilibrium to the dissociated fragment limit. At the dissociation limit, the orbitals are singly occupied, e.g., χ_1 and χ_2 on each fragment, respectively. At the equilibrium geometry, they are $\sigma = \chi_1 + \chi_2$ and $\sigma^* = \chi_1 - \chi_2$. More electrons and orbitals may be included to describe additional static correlation effects. Dynamic correlation is then obtained from an MRSDCI or related calculation (*vide infra*), allowing single and double excitations from the dominant reference CASSCF configurations, which are defined as those having CASSCF wavefunction coefficients greater than 0.05. The MRSDCI method is the most commonly used CI-based MR method, also used in our earlier applications^{29,46} on hydrocarbon BDEs. Admittedly, MRSDCI is limited in that it is not size-extensive; the resulting error increases superlinearly with system size. Furthermore, it is not clear how this size-extensivity error affects energy differences like BDEs even for small, oxygenated molecules, so we investigate these effects here.

Several different schemes already exist that correct for MRSDCI's size-extensivity errors.^{36,37,49-55} Here we test their accuracies when used to predict BDEs of oxygenated molecules. We focus specifically on Davidson-Silver-corrected MRSDCI (MRSDCI-DS)^{53,54} and two variants of the multireference averaged coupled pair functional, MRACPF³⁶ and MRACPF2.³⁷

Out of several *a posteriori* schemes available, MRSDCI-DS was shown to give a simple and correctly scaling approximation.³³ MRSDCI-DS is obtained from the MRSDCI energy and wavefunction:

$$E_{\text{MRSDCI-DS}} = E_{\text{MRSDCI}} + (E_{\text{MRSDCI}} - E_{\text{REF}}) \times \left[\frac{1 - \sum_i (c_i^{\text{REF}})^2}{2 \sum_i (c_i^{\text{REF}})^2 - 1} \right]. \quad (4)$$

Here E_{MRSDCI} is the MRSDCI total energy and the c_i^{REF} s are the coefficients of the reference configurations in the MRSDCI wavefunction. The reference energy (E_{REF}) is the energy of the corresponding reference wavefunction. Note that when the reference configurations are a subset of the total CAS, the reference energy is *not* the CASSCF energy, but instead the CI energy of a variationally optimized linear combination of the CASSCF dominant configurations that comprise the reference configurations.

We achieve *a priori* size-extensivity corrections with MRACPF-like functionals. They are obtained by

renormalizing the denominator of the MRSDCI energy functional with the introduction of so-called g -values to maintain size-extensivity:^{36,37}

$$F_{MRACPF} = \frac{\langle \Psi_{REF} + \Psi_S + \Psi_D | \hat{H} - E_{REF} | \Psi_{REF} + \Psi_S + \Psi_D \rangle}{1 + g_S \langle \Psi_S | \Psi_S \rangle + g_D \langle \Psi_D | \Psi_D \rangle}. \quad (5)$$

Here, the CI wavefunction has been separated into the (multi-configuration) reference wavefunction (Ψ_{REF}) and portions obtained from single (Ψ_S) and double (Ψ_D) excitations from the reference wavefunction. Various flavors of MRACPF are available depending on the g -values. The trivial g -value of 1 gives the MRSDCI energy. MRACPF uses $g_S = g_D = \frac{2}{N}$, where N is the number of correlated electrons.³⁶ MRACPF2 uses the same g -value as MRACPF for double excitations ($g_D = \frac{2}{N}$), but it uses a damped g -value for single excitations ($g_S = \frac{4}{N} [1 - \frac{1}{2(N-1)}]$).³⁷ Unlike MRSDCI, MRACPF and MRACPF2 energies scale approximately linearly with system size (i.e., are approximately size-extensive) and do not unphysically diverge from the thermodynamic limit.³³ One might expect MRSDCI to produce accurate BDEs for small molecules (since those would have small extensivity errors) and less accurate BDEs for larger molecules (large extensivity errors). In fact, MRSDCI BDEs are not necessarily accurate or consistent with MRACPF and MRACPF2 BDEs even for small, oxygenated molecules, as we shall show.

MRACPF2 only differs from MRACPF by using damped g -values for single excitations to improve stability.³⁷ However, our BDE calculations sometimes encountered instabilities in the MRACPF (and, more rarely, in the MRACPF2) wavefunction related to already documented MRACPF instabilities that motivated the creation of MRACPF2.^{37,56,57} This point and how to circumvent it is discussed at length in a separate paper,⁵⁸ while the present work focuses on other technical aspects that need to be addressed to obtain reliable *ab initio* MRCI-based BDEs.

Note that a related concept to size-extensivity is size-consistency. A size-consistent method for two non-interacting systems A and B produces energies such that Energy(A) + Energy(B) = Energy(AB). Neither MRSDCI nor MRACPF(2) are size-consistent. To maintain a size-consistent calculation, we employ the supermolecule approach where both fragments formed by bond dissociation are included in the same calculation separated by a large distance (10 Å). This keeps the level of excitations used for the equilibrium geometry and fragment calculations equivalent.²⁴ The supermolecule and the equilibrium geometries employ the same set of reference configurations in order to obtain consistent energies. We do this by selecting a superset of dominant configurations for the two geometries, which would also be the correct consistent set of configurations for any other point on the bond dissociation PES. For example, in a CAS(2e,2o) calculation (where the active orbitals are σ and σ^*), the equilibrium structure has two dominant references: the “20” (read as “two-zero”) configuration represents the electron configuration with two electrons in the σ orbital and zero electrons in the σ^* orbital. The “02” configuration represents the electron configuration with zero electrons in the σ orbital and two

electrons in the σ^* orbital. The supermolecule has only one dominant reference, the 11 configuration of the two singly occupied radical orbitals, so the superset of reference configurations we use for both geometries consists of the 20, 02, and 11 configurations.

MRSDCI, MRSDCI-DS, MRACPF, and MRACPF2 energies converge slowly with basis set size, making the choice of basis set non-trivial (see, e.g., Ref. 59). Consider methane as an example: the experimental D_{298} for a methane C–H bond is 104.99 ± 0.03 kcal/mol. However, MRSDCI D_{298} s from the cc-pVDZ and cc-pVTZ correlation-consistent basis sets⁶⁰ are 99.6 kcal/mol and 103.1 kcal/mol, respectively.²⁹ Calculations using cc-pVQZ basis sets are prohibitively expensive for the larger molecules that are our ultimate goal, so our CBS extrapolations instead involve the cc-pVDZ and cc-pVTZ basis sets feasible for use with larger molecules in future investigations. Many different CBS extrapolation schemes exist;^{61–71} we use the two-point set extrapolation scheme from Truhlar *et al.*^{65,67}

$$E_X^{ref} = E_\infty^{ref} + A^{ref} X^{-\alpha}, \quad (6)$$

$$E_X^{cor} = E_\infty^{cor} + A^{cor} X^{-\beta}. \quad (7)$$

X in these equations is identified with indices of the cc-pVDZ and cc-pVTZ basis sets and take values equal to 2 and 3, respectively. The reference and correlation energy extrapolations are done separately because (CAS)SCF energies converge faster than dynamic correlation energies. Thus, the extrapolation parameters α and β are different, and they have been obtained by Truhlar⁶⁵ as 3.4 and 2.4, respectively, by fitting Hartree-Fock (HF; for α) and CCSD(T) (for β) energies to CBS energy estimates from explicitly correlated wavefunction (R12)^{72,73} calculations. Even though α was fit to HF, we expect similar basis set convergence in our calculations given the small number of references used (3 for CAS(2e,2o) calculations and 8 for CAS(8e,6o) in acyloxy-related calculations).²⁹ Using $\beta = 2.4$ from CCSD(T) should be appropriate for the CI-based methods we employ because they include primarily single and double excitations, as well as selected higher order excitations due to the multiple references. We tested another extrapolation procedure⁷⁴ of the form: $E_X = E_\infty + A(X + 1)^{-4}$ and found it gives BDEs similar to the ones reported in this work using the extrapolation scheme described above.⁷⁵

III. COMPUTATIONAL DETAILS

CASSCF calculations were done with MOLCAS 7.8.⁷⁶ Both TigerCI^{30–33} and MOLCAS were used for the MRSDCI and MRACPF calculations. MRACPF2 energies were computed within TigerCI. Mulliken charge populations⁷⁷ calculated by MOLCAS aided in explaining some of the BDE results. Geometry optimizations and vibrational frequency calculations were performed at the DFT-B3LYP/6-311G(2d,p) level with GAMESS-US⁴³ using the program’s default convergence parameters. As a single-reference comparison to CI energies, we also calculated CCSD(T) BDEs and the CCSD(T) T1 diagnostic,⁷⁸ a metric frequently used to measure MR character, using MOLPRO.⁷⁹ CCSD(T) BDEs

were computed from R-CCSD(T) energies for closed shell molecules and U-CCSD(T) for open shell dissociation fragments using the same basis set extrapolation as used for the MRSDCI/MRACPF methods.

IV. RESULTS

We now discuss BDEs for several classes of oxygenated molecules including alcohols, aldehydes, carboxylic acids, and methyl esters. We determine which MRCI methods accurately calculate BDEs of the different bonds. We investigate two main elements of the MRCI methodology: (i) the active space sizes needed to model accurately static correlation effects and (ii) the performance of different MRCI methods. We compare against experimental BDEs and reference CCSD(T) calculations. In the absence of MR effects, CCSD(T) is more accurate than CI based methods because it is rigorously size extensive and includes larger sets of (disconnected) higher order excitations. We discuss the BDE results by molecule or bond type in order of increasing degree of MR complexity. In Sec. IV A we evaluate BDEs in methanol and ethanol. In Sec. IV B we report the effect of π electron delocalization when bonds beta to the carbonyl groups of aldehydes, acids, and methyl esters are broken. In Sec. IV C we focus our discussion on the HOC(=O)-H bond. Lastly, in Sec. IV D we report geometries, frequencies, and the active space convergence of BDEs in formylxyl radical formed upon HC(=O)O-H dissociation.

A. Alcohols: CAS(2e,2o) model and a comparison of CI methods

This section discusses bond breaking in methanol and ethanol. Oxygen lone pair electrons may delocalize in radical fragments of alcohols, creating multiconfigurational character. In this case, an active space larger than CAS(2e,2o) may

be required. We tested the influence of higher active spaces on the C-H bond scission in methanol because upon dissociation the oxygen $2p_\pi$ lone pair and the singly occupied carbon $2p_\pi$ orbital may exhibit some π -resonance (hyperconjugation) in the resulting H₂COH radical fragment. We therefore added the O $2p_\pi$ orbital to the CAS(2e,2o) orbitals to form a CAS(4e,3o) active space. This orbital rotated out of the CAS(4e,3o) active space, indicating that it does not require additional static correlation. Thus, in the case of alcohol BDEs, the CAS(2e,2o) calculations can be considered to be converged with respect to active space, and the static correlation is sufficiently described at the CAS(2e,2o) level for this bond. We also computed the CCSD(T) T1 diagnostics⁷⁸ of the molecules and all dissociation fragments in methanol and ethanol. All T1 diagnostic values were below the 0.02 threshold suggested as indicative of MR character.⁷⁸ We therefore expect minimal MR effects in alcohol bond dissociations and that these bonds can be properly described using either a MR method at the CAS(2e,2o) active space level or with an accurate single reference method such as CCSD(T).

To elucidate the relative accuracies of the different methods, we compare CAS(2e,2o)-based CI D_0 s with CCSD(T) D_0 s and reference D_0 s computed from enthalpies of formation available in the Active Thermochemical Tables (ATcT).⁸⁰⁻⁸² CCSD(T) and the size-extensivity-corrected MRSDCI methods agree well with the reference ATcT energies (see Table I) and have average mean deviations between 0.6 and 0.9 kcal/mol. MRSDCI BDEs are significantly lower in accuracy than the size-extensivity-corrected methods (average deviation of 1.7 kcal/mol; see Table I). Even in these small molecules, size-extensivity errors render less-accurate BDEs. Of the four size-extensive or size-extensivity corrected methods, CCSD(T) and MRACPF2 give the lowest average (0.6 and 0.7 kcal/mol, respectively) and maximum deviations (1.2 kcal/mol) from the reference data. Thus, CCSD(T) and MRACPF2 appear to be the most accurate of the methods tested.

TABLE I. Alcohol D_0 BDEs from experiment, single reference CCSD(T), and CAS(2e,2o)-based size-extensive multireference CI methods. The hyphen indicates the dissociating bond; the same convention is used in all tables.

Dissociating bond	ATcT D_0^a (kcal/mol)	CCSD(T) D_0 (kcal/mol)	CAS(2e,2o) D_0 (kcal/mol)			
			MRSDCI	MRSDCI-DS	MRACPF	MRACPF2
CH ₃ O-H	103.98 ± 0.09	104.5	102.7	105.0	103.3	104.1
CH ₃ CH ₂ O-H	103.93 ± 0.13	104.1	101.6	105.0	102.6	103.6
HOCH ₂ -H	94.66 ± 0.11	95.8	94.7	96.2	94.9	95.7
HOCH ₂ CH ₂ -H	100.46 ± 0.15	101.1	98.5	101.2	100.5	100.8
CH ₃ <u>CH</u> ₂ OH ^b	93.46 ± 0.15	94.7	93.1	95.3	93.9	94.7
CH ₃ -OH	90.10 ± 0.05	89.7	87.2	89.6	89.2	89.0
CH ₃ CH ₂ -OH	92.05 ± 0.09	92.3	89.1	92.2	91.6	91.6
CH ₃ -CH ₂ OH	85.31 ± 0.11	85.1	83.2	85.3	83.2	84.3
Statistics of absolute deviations from expt. D_0 s						
	Average	0.6	1.7	0.9	0.8	0.7
	Maximum	1.3	2.9	1.9	2.1	1.2
	Std dev.	0.4	1.1	0.6	0.7	0.4

^aReference D_0 s were computed from enthalpies of formation in the Active Thermochemical Tables.⁸²

^bOne of the underlined Hs in CH₃CH₂OH is dissociating.

B. Beta (β) C–H and C–C bonds in carbonyl-containing molecules: CAS(4e,4o) model

Dissociation of single bonds β to the C=O π -bond in aldehydes, acids, and esters produce resonance-stabilized allyl-like radicals with multiconfigurational electronic structures. We computed BDEs for β C–H and C–C bonds in the aldehydes acetaldehyde, propanal, and butanal, as well as those in acetic acid, methyl acetate, and methyl propanoate. D_{298} s for both the CAS(2e,2o) and the CAS(4e,4o) active spaces are reported in Table II. The CAS(4e,4o) is composed of the π and π^* orbitals of the carbonyl in addition to the σ and σ^* of the dissociating bond. Since enthalpies of formation for species involving these bonds are not included in the Active Thermochemical Tables, experimental D_{298} s from Luo⁸³ are used as a benchmark. Again, CCSD(T) D_{298} s are included as a single reference theory to compare to our MR values.

We first note T1 values for species as an indicator of MR character. T1 diagnostic values for the closed shell aldehyde, acid, and ester molecules lie between 0.012 and 0.015. These values are higher than those for alcohols (0.007), but they are still lower than the 0.02 threshold typically used to denote significant MR character. Non-hydrogen and non-methyl radical products of acid and ester decompositions have diagnostic values between 0.016 and 0.018, while the T1 value for the $\underline{\text{C}}\text{H}_2\text{OH}$ radical from alcohol decomposition is 0.019 (underlined C indicates location of the radical electron). By comparison, the radicals from aldehyde decomposition tend to have higher T1 values: 0.019 for $\text{CH}_3\text{CH}_2\underline{\text{C}}\text{HC}(=\text{O})\text{H}$, 0.021 for $\text{CH}_3\underline{\text{C}}\text{HC}(=\text{O})\text{H}$, and 0.024 for $\underline{\text{C}}\text{H}_2\underline{\text{C}}(=\text{O})\text{H}$. Thus the aldehydic radicals appear to be more multiconfigurational than the corresponding products of bond breaking in alcohols, acids, and esters.

Regarding BDE accuracies, CAS(2e,2o)-MRACPF fails badly for beta bonds in aldehydes. MRACPF with this active space can produce unreasonably low BDEs (deviations from experiment as large as 36 kcal/mol).⁷⁵ On the other hand, MRSDCI, MRSDCI-DS, and MRACPF2 all give physically reasonable energies with the CAS(2e,2o) active space, with all BDEs within 3.7 kcal/mol of experiment. To elucidate the origin of this MRACPF anomaly, we analyze Mulliken charges⁷⁷ (Fig. 1) for propanal in both its equilibrium geometry and the relaxed supermolecule formed upon breaking the C–C bond β to the C=O bond. Unlike BDEs for β bonds, CAS(2e,2o)-based MRACPF BDEs for α bonds yield reasonable MRACPF results.⁷⁵ We therefore compare a case with an anomalous BDE to one with a physically reasonable BDE by also performing the same Mulliken charge analysis for α C–C bond breaking in propanal. All four theories show qualitatively and quantitatively similar charge distributions for α C–C bond breaking (see left column in Fig. 1). On the other hand, the charge distribution from MRACPF substantially differs from the other three theories for the β C–C bond supermolecule (Fig. 1, middle right panel). MRACPF shows excessive charge withdrawal from the carbonyl oxygen to the β carbon (C_β) and, to a smaller extent, the α carbon (C_α). MRACPF essentially overcompensates for the resonance effect expected in the β bond breaking. No such overcompensation appears for the α bond breaking where no electron delocalization is expected. According to our recent analysis of MRACPF instability,⁵⁸ such an artifact of MRACPF should vanish with a larger active space. Indeed with CAS(4e,4o), where π electron delocalization is accounted for, all four levels of theory predict reasonable BDEs that are within 3.1 kcal/mol of experiments (Table II, rightmost three columns). Mean absolute deviations from

TABLE II. β C–H and C–C bond D_{298} BDEs (kcal/mol) in aldehydes, acids, and esters from experiment, single-reference CCSD(T) calculations, and CAS(2e,2o)- and CAS(4e,4o)-based multireference CI and ACPF calculations.

Breaking bond	Expt. D_{298} ^a (kcal/mol)	CCSD(T) D_{298} (kcal/mol)	D_{298} (kcal/mol) at CAS(2e,2o)			D_{298} (kcal/mol) at CAS(4e,4o) ^b		
			MRSDCI	MRSDCI-DS	MRACPF2	MRSDCI	MRSDCI-DS	MRACPF2
Aldehydes								
HC(=O)CH ₂ –H	94.3 ± 2.2	96.3	95.8	96.9	94.7	93.0	94.5	94.4
HC(=O)CH ₂ CH ₃ ³	91.7 ± na	91.5	90.9	92.6	90.3	88.6	90.1	89.8
HC(=O)CH ₂ CH ₂ CH ₃	...	91.9	91.1	93.3	90.7	88.7	90.7	90.1
CH ₃ –CH ₂ C(=O)H	82.0 ± 2.4	83.5	82.4	84.1	79.4	79.7	81.4	81.4
CH ₃ CH ₂ –CH ₂ C(=O)H	79.8 ± 2.5	82.5	81.1	83.5	78.1	78.4	80.6	80.3
		1.6	Average absolute deviations from expt. D_{298} s			2.1	0.8	0.8
			1.0	2.3	1.5			
Acids and esters								
HOC(=O)CH ₂ –H	95.3 ± 2.9	99.5	97.0	99.9	98.9	95.3	97.3	96.7
CH ₃ OC(=O)CH ₂ –H	97.1 ± 2.5	101.2	96.6	100.0	98.9	95.9	98.4	97.3
CH ₃ OC(=O)CH ₂ CH ₃	...	95.4	92.7	96.2	95.1	91.5	93.8	92.4
CH ₃ OC(=O)CH ₂ –CH ₃	...	87.2	83.6	87.5	85.7	82.3	84.8	83.2
		4.1	Average absolute deviations from expt. D_{298} s			0.6	1.7	0.8
			1.1	3.7	2.7			

^aExperimental D_{298} s from Luo.⁸³

^bSuperset CAS(4e,4o) reference configurations used are $\sigma\sigma^*\pi\pi^* = 2020, 2002, \text{ and } 0220$ from the equilibrium geometry and $(2p)(X)\pi\pi^* = 1120, 1111, 1102, \text{ and } 2101$ from the supermolecule, where 2p is the singly occupied radical orbital on the C=O-containing allyl-like radical and X stands for either the hydrogen 1s or the 2p radical orbital of $\underline{\text{C}}\text{H}_3$ or $\underline{\text{C}}\text{H}_2\text{CH}_3$.

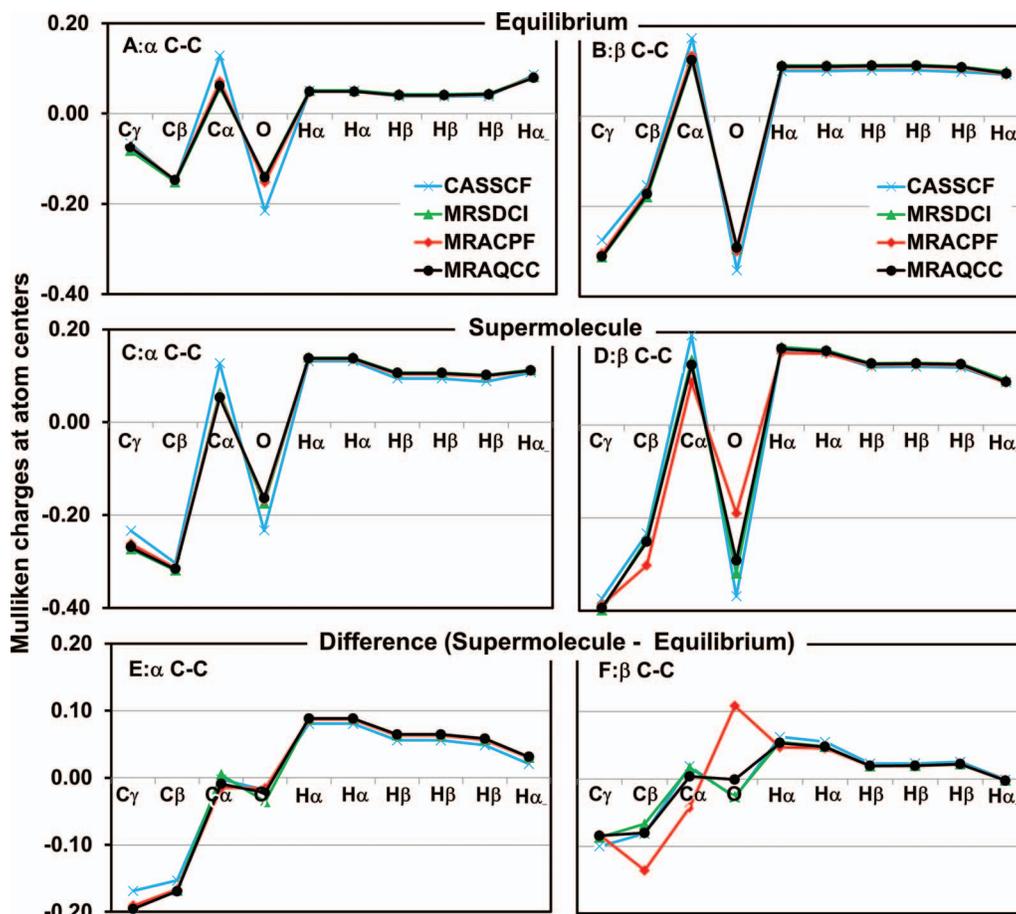


FIG. 1. MRACPF gives incorrect charge distributions for β bonds of aldehydes (propanal used here as illustration).⁸⁴ We consider the α C-C ($\text{CH}_3\text{CH}_2\text{-C(=O)H}$, left) and the β C-C ($\text{CH}_3\text{-CH}_2\text{C(=O)H}$, right) bond dissociation for different theories (see insets in (a) and (b)). Plots show propanal Mulliken charges at atomic centers for the equilibrium structure (a) and (b) and the supermolecule (c) and (d), as well as their differences (e) and (f).

experimental $D_{298\text{s}}$ of CAS(4e,4o)-MRACPF2 (0.8 kcal/mol) are furthermore half as large as those from CAS(2e,2o)-MRACPF2 BDEs (1.5 kcal/mol) and similar to those from alcohol BDEs (0.7 kcal/mol). We conclude from these results that the π resonance effects in aldehydes must be treated at least at the CAS(4e,4o) level and that MRACPF2 again gives the most accurate energies. While average deviations from reference BDEs for BDEs computed with CCSD(T) are larger than those from CAS(4e,4o)-MRACPF2 (1.6 kcal/mol vs. 0.8 kcal/mol, respectively), one cannot say much regarding the relative performance of the methods given the large uncertainty in the reference experimental energies because most energies fall within this error range. However, the key point here is that for these β bonds, MR calculations are more accurate when done at the CAS(4e,4o) level than at the CAS(2e,2o) level. Unlike the bonds in aldehydes, we achieve reasonable BDEs for β C-H and C-C bonds in acids and esters using CAS(2e,2o)-MRACPF. The most obvious reason for the different behavior is the additional oxygen atom in the acids and esters, which localizes electrons near the carbonyl group, drawing charge away from the radical site that forms upon C-C and C-H bond breaking. This results in less resonance stabilization in the allylic radical of C-C and C-H bond breaking, and explains the lower T1 diagnostic values for acid and ester radicals. As we discuss later in

Sec. IV D, the other β bonds involving C-O or O-H bond breaking in acids and esters feature pronounced multiconfigurational character. Overall, we find significantly larger average deviations from the mean experimental values for CAS(2e,2o) than for CAS(4e,4o) BDEs (see Table II), further justifying the larger active space. Of course, as mentioned earlier, the large error bars on the experimental values means that these conclusions are subject to change when measurements that are more accurate become available. We also note that the MRACPF instability mentioned above provides another tool to test convergence of the active space in BDE applications.⁵⁸

C. HOC(=O) formation from C-H bond dissociation in formic acid

The HOC(=O) radical is an important precursor to CO_2 formation in combustion and atmospheric chemistry, and it has therefore been the subject of numerous prior studies (e.g., Refs. 85–87). As the simplest of ROC(=O) radicals, we study the MR requirements of HOC(=O) to elucidate how to rigorously treat the larger radicals in our studies of methyl esters. The HOC(=O) radical has two conformers, *trans*-HOC(=O) and *cis*-HOC(=O), with similar energies. We found that the *trans*-HOC(=O) conformer is more stable; DFT-B3LYP electronic energies favor *trans*-HOC(=O)

TABLE III. DFT-B3LYP vibrational frequencies and zero point energies (ZPE) for HOC(=O) radicals.

Mode	Vibrational frequencies (cm ⁻¹)				
	<i>cis</i> -HOC(=O)		<i>trans</i> -HOC(=O)		Expt. ^c
	CcCR-VPT ^a	DFT-B3LYP ^b	CcCR-VPT ^a	DFT-B3LYP ^b	
a' O ₁ -H stretch	3451	3511	3641	3777	3636
a' C=O ₂ stretch	1823	1847	1862	1881	1853
a' H-O ₁ -C bend	1284	1301	1214	1229	...
a' C-O ₁ stretch	1046	1063	1053	1068	...
a' O ₁ -C-O ₂ bend	602	600	617	614	...
a'' torsional	567	596	498	542	...
ZPE (kcal/mol)	12.7	12.9	12.9	13.0	...

^aCcCR-VPT: CCSD(T) with corrections for core correlation and relativistic effects; quadratic force field. Data taken from Refs. 85 and 86.

^bDFT-B3LYP/6-311G(2d,p) harmonic vibrational frequencies from this work.

^cExperimental O-H stretch frequency from Ref. 88 and C=O stretch frequency from Ref. 89.

over *cis*-HOC(=O) by 1.3 kcal/mol. Gas phase experimental vibrational frequencies are not available for these two radicals except for the O-H and C=O stretch modes in *trans*-HOC(=O).^{88,89} The DFT-B3LYP/6-311G(2d,p) frequencies (Table III) differ from the available experimental frequencies by 141 cm⁻¹ (0.40 kcal/mol) for the H-O stretch mode and 29 cm⁻¹ for the C=O stretch. We also compare the DFT-B3LYP frequencies to recent high level *ab initio* anharmonic frequencies from CCSD(T)-based quadratic force field calculations of Fortenberry *et al.*^{85,86} All the DFT-B3LYP frequencies for the two radicals agree with the CCSD(T) frequencies to within 50 cm⁻¹ (0.14 kcal/mol). The DFT-B3LYP/6-311G(2d,p) harmonic frequencies are thus in very good agreement with higher level theory and the limited experimental data. This provides additional validation of the general quality of DFT-B3LYP/6-311G(2d,p) geometries and frequencies, even though we identified some earlier failures in alkynyl radicals.⁴⁶ Once the ZPE and thermal correction obtained from the DFT-B3LYP frequencies are added to the HOC(=O) DFT-B3LYP energies, an energy difference of 0.9 kcal/mol separates the two conformers, with *trans*-HOC(=O) remaining the more stable of the two. We study both conformers further with MR wavefunction methods.

Table IV shows MRSDCI, MRSDCI-DS, MRACPF, and MRACPF2 D_0 s for HOC(=O)-H, when considering the two conformers of HOC(=O) as dissociation products. The experimental D_0 s for this bond are 98.77 ± 0.16 and 97.00 ± 0.14 kcal/mol, respectively, for the *cis* and *trans* HOC(=O) conformers.⁸¹ At the CAS(2e,2o) level, computed bond energies appear fairly accurate (where values averaged over MRSDCI, MRSDCI-DS, MRACPF, and MRACPF2 are 98.9 kcal/mol and 98.4 kcal/mol for the *trans*-HOC(=O) and *cis*-HOC(=O) products, respectively). The CCSD(T) T1 diagnostics for *trans*-HOC(=O) and *cis*-HOC(=O) are 0.025 and 0.024, respectively, suggesting significant MR character requiring a larger active space than the CAS(2e,2o). We first considered the substantially larger CAS(8e,8o) active space composed of C-H σ , C-H σ^* , O-C σ , O-C σ^* ,

TABLE IV. D_0 s for HOC(=O)-H bond breaking to form either the *cis*- or *trans*-HOC(=O) radical, comparing size-extensivity-corrected multi-reference CI methods against CCSD(T) and experiment.

Method	D_0 (HOC(=O)-H) (kcal/mol)			
	<i>cis</i> -HOC(=O)		<i>trans</i> -HOC(=O)	
	CAS(2e,2o)	CAS(10e,9o)	CAS(2e,2o)	CAS(10e,9o)
MRSDCI	98.6	100.8	98.9	99.3
MRSDCI-DS	99.0	100.3	99.3	98.2
MRACPF	98.0	101.0	96.6	98.9
MRACPF2	100.1	101.3	98.7	99.3
CCSD(T)	99.8		97.9	
ATcT ^a	98.77 ± 0.16		97.00 ± 0.14	

^aReference D_0 s were computed from enthalpies of formation from the Active Thermochemical Tables.⁸²

O=C σ , O=C σ^* , O=C π , and O=C π^* orbitals. However, the carbonyl O lone pair rotated into the CAS(8e,8o) of the supermolecule and was therefore added to the active space to form a CAS(10e,9o). Average computed D_0 s at the CAS(10e,9o) level are 100.4 kcal/mol and 98.9 kcal/mol for *trans*-HOC(=O) and *cis*-HOC(=O), respectively. Calculations with the CAS(10e,9o) active space show slightly larger deviations from experiment than calculations with the CAS(2e,2o) active space and single reference CCSD(T). Overall, the lower BDE predicted when forming the *trans*-HOC(=O) conformer means that this conformer is the ground state isomer for HOC(=O) for all MR approaches we consider, in agreement with previous theoretical studies.

D. Formyloxyl radical (HC(=O)O) formation in acids and esters: CAS(8e,6o) active space

1. Geometries and vibrational frequencies

Dissociation of RC(=O)O-R' bonds leads to acyloxyl radicals. Multiple near-degenerate orbitals in these radicals complicate their electronic structure.⁹⁰⁻⁹⁸ For example, the formyloxyl radical, HC(=O)O, has four near-degenerate orbitals (see right-hand side of Fig. 2 for the σ (b_2 and a_1) and π (b_1 and a_2) orbitals of HC(=O)O). There are six unique electronic states for HC(=O)O, each with a unique geometry: three σ radicals (2B_2 , 2A_1 , ${}^2A'$) and three π radicals (2B_1 , 2A_2 , ${}^2A''$). Mixing between the 2B_2 and 2A_1 C_{2v} states or the 2B_1 and 2A_2 C_{2v} states produce the ${}^2A'$ and ${}^2A''$ states, respectively. The π states are higher in energy than the σ states and are thus relatively unimportant in combustion, except perhaps at very high temperatures ($k_B T = 3.0$ kcal/mol at $T = 1500$ K while *ab initio* theories (CASSCF, CASPT2, MRCI) predict any of the π radicals are > 5 kcal/mol less stable than the 2B_2 σ radical).⁹⁰

Numerous *ab initio* methods qualitatively disagree on the energetic ordering of states of HC(=O)O. Rauk *et al.*^{90,91} reported that HF yields an energy ordering of the states as ${}^2A' < {}^2B_2 < {}^2A_1$. CASSCF(13e,11o) finds the states to be ${}^2A' < {}^2B_2$ but with the 2A_1 state being a transient leading to decomposition into H and CO₂. Single reference MP2 predicts three minima with ${}^2A_1 < {}^2B_2 < {}^2A'$ and, similarly, CASPT2

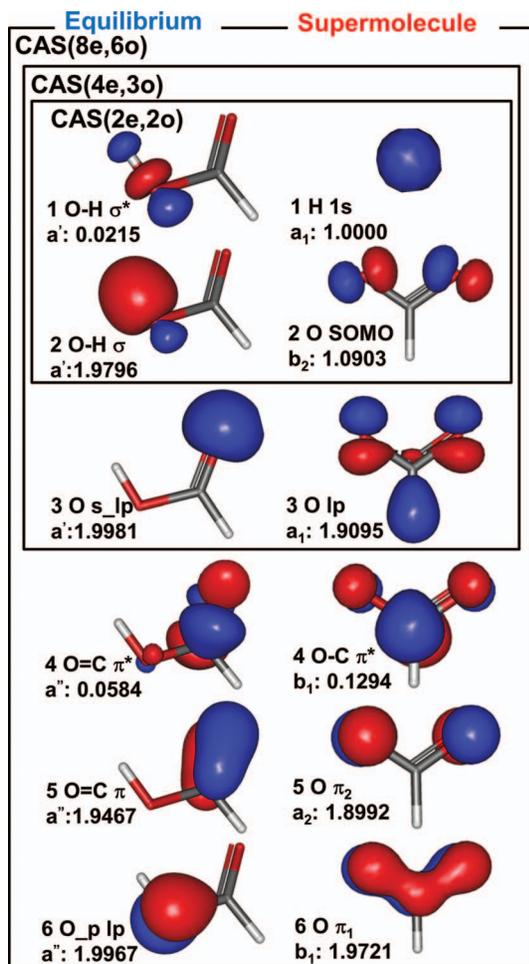


FIG. 2. Formylxoyl orbitals are strongly correlated. Orbitals and occupation numbers for CAS(2e,2o), CAS(4e,3o), and CAS(8e,6o) active spaces for HC(=O)O–H bond breaking (see Sec. IV D 2 for explanation). On the left are active orbitals for formic acid at its equilibrium geometry and on the right are the active orbitals for H atom and formylxoyl radical in the supermolecule geometry.

finds ${}^2A_1 < {}^2B_2$, but with the ${}^2A'$ being a transition state for isomerization from HC(=O)O (2B_2) \rightarrow HOC(=O) (2A_1). Both single reference and MR CI find that the ground state structure is 2B_2 . Vibronic coupling between the low-lying formylxoyl states complicates the experimental identification of the true ground state.^{96,99} Recent studies by Garand *et al.*⁹⁶ using photoelectron velocity mapping spectroscopy aided by CCSD(T)-based *ab initio* calculations determined that the 2A_1 state is the global minimum structure and the 2B_2 state is slightly higher in energy by 0.9 kcal/mol. While it is not clear that this work provides the definitive resolution to the disagreements between *ab initio* methods, it appears that the true ground state geometry of HC(=O)O has C_{2v} symmetry and that both 2B_2 and 2A_1 will be important at combustion temperatures, which are typically above 600 K. Both states are considered in what follows.

Up until now, our validations have focused on electronic energies assuming DFT-B3LYP geometries and frequencies are accurate. However, in an earlier study,⁴⁶ we reported that DFT functionals with exact exchange below a threshold amount (e.g., less than 25% HF exchange in alkynyl radical

cases) can produce unphysical structures of alkynyl radicals. In particular, DFT-B3LYP produced an unphysically bent geometry for alkynyl radical. Here we investigate whether DFT-B3LYP finds correct ground state geometries for oxygenated radicals. To avoid the computational expense of post-HF methods, Kieninger *et al.*⁹⁵ sought to ascertain the performances of various DFT approximations for formylxoyl radical geometries. They found that the pure GGA XC functional BLYP^{41,42} and the hybrid XC-functionals B3LYP^{41,42} and B3PW91^{41,100} (each with 20% HF exchange) all predict qualitatively similar state orderings as CASPT2 (${}^2A_1 < {}^2B_2 < {}^2A'$). We tested several DFT XC functionals ourselves and found that the amount of exact exchange included definitely affects the geometries obtained. We performed a scan of the HC(=O)O PES along the OCO bond angle, i.e., here the OCO angle is constrained at each point of the search while the other degrees of freedom are allowed to relax. XC functionals with more HF exchange (>50%) locate only the ${}^2A'$ state as a stable structure while functionals with less HF exchange (<27%) find the 2B_2 and 2A_1 states are stable structures with nearly the same energies (Fig. 3). Here, it appears that less exact exchange allows the molecule to (fortuitously) form the experimentally observed C_{2v} geometry, while the combination of higher amounts of exact exchange and no explicit treatment of near-degeneracies via a MR theory favors the more ionic, incorrect Cs geometry. Notably, DFT-B3LYP finds the correct minimum geometries and transition states.

We further compare DFT-B3LYP frequencies to those from higher level CASPT2 theory reported in the literature (see Table V). Most DFT-B3LYP frequencies agree well with CASPT2, again confirming that DFT-B3LYP produces reasonably accurate frequencies. An exception is the H in-plane wag of the 2A_1 conformer (B3LYP: 103 cm^{-1} , CASPT2: 642 cm^{-1}). ZPEs agree well for 2B_2 (within 0.1 kcal/mol) but disagree by 0.7 kcal/mol for 2A_1 . Note also that the DFT-B3LYP ZPE for 2B_2 (12.1 kcal/mol) is higher than the 2A_1 ZPE (9.8 kcal/mol) by 2.3 kcal/mol, again consistent with a

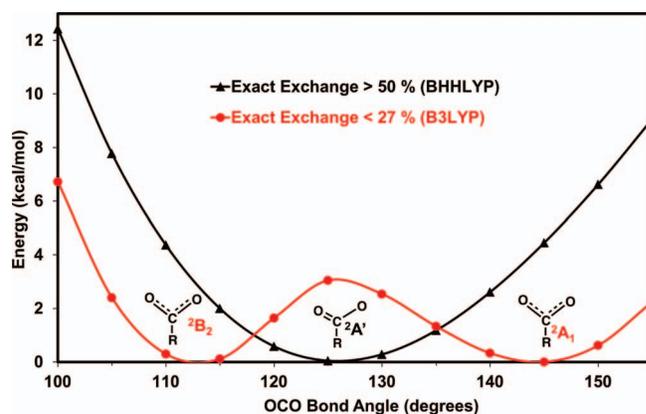


FIG. 3. DFT using exchange-correlation functionals with exact exchange (X) greater than 50% (including Hartree-Fock theory) locate only the ${}^2A'$ state, while those with lower exact exchange correctly locate all three states (2B_2 , 2A_1 , and ${}^2A'$). The lowest point of the potential energy curves for BHHLYP and B3LYP are referenced to zero on the vertical axis. Several functionals or methods (X \leq 27%: U-M06, UBLYP, UB3LYP; X \geq 50%: UBHHLYP, UM06-2X, UHF, ROHF) were tested. All exhibit behavior characterized by either B3LYP or BHHLYP, respectively.

TABLE V. Vibrational frequencies and zero point energies (ZPE) of HC(=O)O conformers from DFT-B3LYP/6-311G(2d,p) and CASPT2/ANO (Ref. 90).

Mode	Vibrational frequencies (cm ⁻¹)			
	² B ₂		² A ₁	
	DFT-B3LYP	CASPT2	DFT-B3LYP	CASPT2
CO ₂ bend	646	624	664	653
CO ₂ asym. stretch	977	1008	1643	1669
H out of plane wag	1017	1150	841	830
H in-plane wag	1275	1287	103	642
CO ₂ symm. stretch	1453	1437	1200	1153
C-H stretch	3008	3053	2297	2392
ZPE (kcal/mol)	12.1	12.2	9.8	10.5

²A₁ ground state. Overall, the DFT-B3LYP frequencies and ZPEs are satisfactory.

2. Multiconfigurational nature of the formyloxyl radical: Active space convergence

Breaking the weakest C–O bond in methyl esters (the CH₃–O bond) produces acyloxyl radicals.^{101,102} These bonds and the resulting radicals are therefore likely to be critically important in biodiesel combustion chemistry. We elucidate the multiconfigurational character of the HC(=O)O electronic structure to guide our follow-on studies of acyloxyl radicals in biodiesel methyl esters. Although energies from calculations with the CAS(2e,2o) active space appear reasonable for MRACPF, the large errors from the other CI methods suggests that the active space for this radical is actually not converged until the CAS(8e,6o) active space (Fig. 2) is used (Fig. 4). In particular, MRACPF2 results for CAS(8e,6o) agree very well with experimental and ATcT values (MRACPF2 D_{298} : 111.5 kcal/mol; Expt. D_{298} : 112 ± 3 kcal/mol from Luo;⁸³ ATcT D_{298} : 113.9 + 0.2 kcal/mol from ATcT obtained by adding ATcT D_0 = 112.53 kcal/mol and

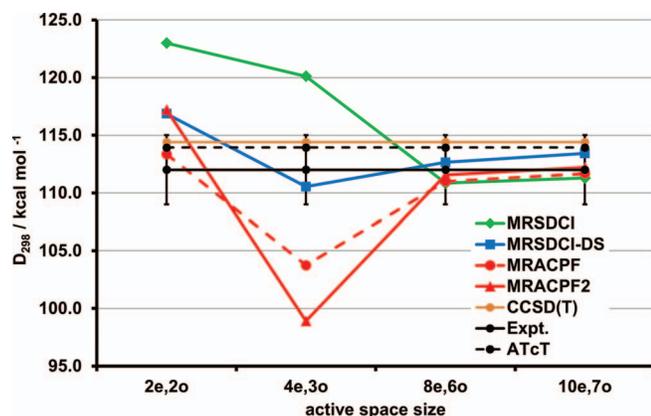


FIG. 4. D_{298} (HC(=O)O–H) as a function of active space size. The BDEs converge at CAS(8e,6o) for all CI methods. CCSD(T) D_{298} is included for comparison, in addition to the experimental D_{298} from Luo (“Expt.”)⁸³ and ATcT. The ATcT D_{298} was obtained by adding D_0 computed from ATcT enthalpies of formation to the DFT-B3LYP $\Delta H_{0 \rightarrow 298}$.

B3LYP/6-311G(2d,p) $\Delta H_{0 \rightarrow 298}$ = 1.4 kcal/mol). MRCI calculations with intermediate CAS sizes may feature larger errors, however. As discussed in detail below, including only orbitals participating in σ -resonance as in the CAS(4e,3o) leads to large errors while including the π -resonance is critical (as in the CAS(8e,6o)). Further increasing the active space by adding the carbonyl O 2p_σ lone pair electrons and orbital in the CAS(10e,7o) (not shown in Fig. 2) has little effect on the results. MRCI BDEs from CAS(8e,6o) calculations are converged with respect to active space size and therefore much more trustworthy than CAS(2e,2o) or CAS(4e,4o) results. These findings further demonstrate the sensitivity of the MR approaches to active space sizes and the importance of identifying a converged active space; while the latter is well known, it never hurts to re-emphasize this point with more data. CCSD(T)’s accuracy suggests that MR effects may be less important than including enough higher level excitations in the wavefunction. Indeed, the CAS(8e,6o) includes many more excitations than the CAS(2e,2o). We note that the MRS-DCI and MRACPF results agree very well with Luo’s tabulated experimental value while CCSD(T) and MRSDCI-DS agree more closely with ATcT-derived values.

Having determined that converged and accurate BDEs for the formyloxyl producing bond require a CAS(8e,6o), we now explain how these eight electrons are correlated within the six orbitals. First, in the supermolecule geometry (Fig. 2, right column), the CAS(8e,6o) wavefunction correlates σ orbitals a_1 and b_2 by allowing electrons to excite from doubly occupied a_1 to the singly occupied b_2 . Since the 1s orbital of hydrogen is 10 Å away, it does not interact with the formyloxyl a_1 and b_2 orbitals and retains an exact single electron occupation. Electronic occupations in the supermolecule show that the doubly occupied π orbitals b_1 and a_2 are correlated by the b_1 π^* orbital. Likewise, the σ orbitals and the electrons residing in them are separately correlated from the π orbitals and their electrons. This is because the s-like (σ) and p-like (π) orbitals have different symmetries. Similarly, in the formic acid equilibrium geometry (Fig. 2, left column), the π O=C bonding orbital is correlated with the O=C π^* . The fractional occupations indicate that the oxygen 2p lone pair of the non-carbonyl oxygen and the O=C π^* are also correlated. In the formic acid σ orbital set, the O–H σ and the O–H σ^* are correlated. As alluded to earlier, an oxygen lone pair orbital is included in the equilibrium active space to be consistent with the supermolecule active space. For consistency with the latter, we included the oxygen 2p lone pair in our initial guess, but the CAS(8e,6o) optimization rotates the oxygen 2s lone pair into the active space instead. Including both oxygen 2s- and 2p- lone pairs gives rise to the CAS(10e,7o) active space mentioned earlier, which only changes the MRS-DCI and MRSDCI-DS/MRACPF/MRACPF2 D_0 by 0.4 and 0.7 kcal/mol, respectively. Thus, the CAS(8e,6o)-derived MRSDCI/MRACPF calculations were performed using eight dominant references and give chemically accurate BDEs.⁷⁵

V. CONCLUSIONS

Our analyses establish the MRCI-based methods and active space levels needed to accurately determine BDEs in

various bonds of oxygenated compounds. We benchmark our results against experimental or ATcT data and CCSD(T) calculations for BDEs. MRSDCI is consistently less accurate than MRSDCI-DS, MRACPF, and MRACPF2 for calculating BDEs in alcohols, aldehydes, acids, and esters due to size-extensivity errors, even in small molecules. In many cases, CAS(2e,2o) active spaces are suitable to calculate accurate BDEs, especially those of alcohols. In situations involving more complicated electronic structure, larger active spaces are needed to obtain reasonably accurate BDEs. To properly describe π -resonance in the resulting radical, a CAS(4e,4o) active space is necessary for C–H and C–C bonds β to C=O and a CAS(8e,6o) active space is needed for breaking bonds producing acyloxy radicals. Based on the accuracy of CCSD(T), the larger active space may not always be needed to capture multiconfigurational character but rather to recover contributions from (disconnected) higher order excitations.

We have also discussed how MRACPF is highly sensitive to inadequacies in active space size and in fact may be used to verify the sufficiency of active space size due to this sensitivity. MRACPF2 is the most accurate of all CI methods tested, achieving chemical accuracies in BDE predictions, while correctly accounting for MR character encountered during bond-breaking processes. Thus, for computations of entire PESs we recommend using MRACPF2 calculations, of course with adequately sized active spaces. Our results establish the toolset ultimately needed for future investigations of the entire bond breaking/forming process and of other combustion-related chemical reactions.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Stephen J. Klippenstein for helpful discussions. This material is based upon work supported as part of the Combustion Energy Frontier Research Center, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award No. De-SC0001198. The authors are pleased to acknowledge computational support by the TIGRESS high performance computer center at Princeton University, which is jointly supported by the Princeton Institute for Computational Science and Engineering and the Princeton University Office of Information Technology.

¹A. K. Agarwal, *Prog. Energy Combust. Sci.* **33**(3), 233–271 (2007).

²C. K. Westbrook, W. J. Pitz, and H. J. Curran, *J. Phys. Chem. A* **110**(21), 6912–6922 (2006).

³L. S. Tran, B. Sirjean, P. A. Glaude, R. Fournet, and F. Battin-Leclerc, *Energy* **43**(1), 4–18 (2012).

⁴K. Peterson, D. Feller, and D. Dixon, *Theor. Chem. Acc.* **131**(1), 1079 (2012).

⁵P. Hohenberg and W. Kohn, *Phys. Rev.* **136**(3B), B864–B871 (1964).

⁶W. Kohn and L. J. Sham, *Phys. Rev.* **140**(4A), A1133–A1138 (1965).

⁷R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**(1), 291–352 (2007).

⁸R. J. Bartlett and G. D. Purvis, *Int. J. Quantum Chem.* **14**(5), 561–581 (1978).

⁹L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **109**(1), 42–55 (1998).

¹⁰L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **109**(18), 7764–7776 (1998).

¹¹L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **114**(1), 108–117 (2001).

¹²J. A. Montgomery, M. J. Frisch, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **110**(6), 2822–2827 (1999).

¹³A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**(23), 11599–11613 (2004).

¹⁴A. L. East and W. D. Allen, *J. Chem. Phys.* **99**(6), 4638–4650 (1993).

¹⁵J. P. Kenny, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **118**(16), 7353–7365 (2003).

¹⁶J. M. L. Martin and G. D. Oliveira, *J. Chem. Phys.* **111**(5), 1843–1856 (1999).

¹⁷A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, *J. Chem. Phys.* **120**(9), 4129–4141 (2004).

¹⁸A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, *J. Chem. Phys.* **125**(14), 144108 (2006).

¹⁹R. J. Bartlett, *Int. J. Mol. Sci.* **3**(6), 579–603 (2002).

²⁰A. I. Krylov, *Acc. Chem. Res.* **39**(2), 83–91 (2006).

²¹P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, *Int. Rev. Phys. Chem.* **21**(4), 527–655 (2002).

²²B. O. Roos, “The complete active space self-consistent field method and its applications in electronic structure calculations,” in *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry - Part 2*, edited by K. P. Lawley (John Wiley & Sons, New York, 1987), Vol. 69, pp. 399–445.

²³K. Andersson, P.-A. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**(2), 1218–1226 (1992).

²⁴A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (McGraw-Hill, New York, 1989), p. 480.

²⁵X. Li and J. Paldus, *J. Phys. Chem. A* **111**(44), 11189–11197 (2007).

²⁶M. Kállay, P. G. Szalay, and P. R. Surjan, *J. Chem. Phys.* **117**(3), 980–990 (2002).

²⁷G. A. Oyedepo and A. K. Wilson, *J. Phys. Chem. A* **114**(33), 8806–8816 (2010).

²⁸G. A. Oyedepo, C. Peterson, and A. K. Wilson, *J. Chem. Phys.* **135**(9), 094103 (2011).

²⁹V. B. Oyeyemi, M. Pavone, and E. A. Carter, *ChemPhysChem* **12**, 3354–3364 (2011).

³⁰D. Walter, A. Venkatnathan, and E. A. Carter, *J. Chem. Phys.* **118**(18), 8127–8139 (2003).

³¹T. S. Chwee, A. B. Szilva, R. Lindh, and E. A. Carter, *J. Chem. Phys.* **128**(22), 224106 (2008).

³²T. S. Chwee and E. A. Carter, *J. Chem. Phys.* **132**(7), 074104 (2010).

³³D. Krisiloff and E. A. Carter, *Phys. Chem. Chem. Phys.* **14**, 7710–7717 (2012).

³⁴V. R. Saunders and J. H. van Lenthe, *Mol. Phys.* **48**(5), 923–954 (1983).

³⁵P. E. M. Siegbahn, *Int. J. Quantum Chem.* **18**(5), 1229–1242 (1980).

³⁶R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.* **143**(5), 413–420 (1988).

³⁷R. J. Gdanitz, *Int. J. Quantum Chem.* **85**(4–5), 281–300 (2001).

³⁸S. F. Sousa, P. A. Fernandes, and M. J. Ramos, *J. Phys. Chem. A* **111**(42), 10439–10452 (2007).

³⁹M. Swart and J. G. Snijders, *Theor. Chem. Acc.* **110**(1), 34–41 (2003).

⁴⁰C. Puzzarini, M. Biczysko, and V. Barone, *J. Chem. Theory Comput.* **6**(3), 828–838 (2010).

⁴¹A. D. Becke, *J. Chem. Phys.* **98**(7), 5648–5652 (1993).

⁴²C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**(2), 785–789 (1988).

⁴³M. S. Gordon and M. W. Schmidt, “Advances in electronic structure theory: GAMESS a decade later,” in *Theory and Applications of Computational Chemistry, the First Forty Years*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.

⁴⁴A frequency scaling factor of 0.99 is proposed for use in the CBS-QB3 method with DFT-B3LYP/6-311G(d,p) for 1st row elements and DFT-B3LYP/6-311G(2d,p) for 2nd row elements. This factor is so close to unity that frequency scaling makes negligible difference in energies.

⁴⁵D. McQuarrie and J. Simon, *Molecular Thermodynamics* (University Science Books, California, 1999).

⁴⁶V. B. Oyeyemi, J. A. Keith, M. Pavone, and E. A. Carter, *J. Phys. Chem. Lett.* **3**(3), 289–293 (2012).

⁴⁷S. Grimme, *Org. Lett.* **12**(20), 4670–4673 (2010).

⁴⁸S. N. Steinmann, C. Piemontesi, A. Delachat, and C. Corminboeuf, *J. Chem. Theory Comput.* **8**(5), 1629–1640 (2012).

⁴⁹L. Meissner, *Chem. Phys. Lett.* **146**(3–4), 204–210 (1988).

⁵⁰J. A. Pople, R. Seeger, and R. Krishnan, *Int. J. Quantum Chem.* **12**(S11), 149–163 (1977).

⁵¹K. A. Brueckner, *Phys. Rev.* **100**(1), 36–45 (1955).

- ⁵²W. L. Luken, *Chem. Phys. Lett.* **58**(3), 421–424 (1978).
- ⁵³E. R. Davidson and D. W. Silver, *Chem. Phys. Lett.* **52**(3), 403–406 (1977).
- ⁵⁴P. E. M. Siegbahn, *Chem. Phys. Lett.* **55**(2), 386–394 (1978).
- ⁵⁵W. Duch and G. H. F. Diercks, *J. Chem. Phys.* **101**(4), 3018–3030 (1994).
- ⁵⁶C. W. Bauschlicher, S. R. Langhoff, and A. Komornicki, *Theor. Chim. Acta* **77**(4), 263–279 (1990).
- ⁵⁷W. Cardoen and R. J. Gdanitz, *Chem. Phys. Lett.* **364**(1–2), 39–43 (2002).
- ⁵⁸D. B. Krisiloff, V. B. Oyeyemi, F. Libisch, and E. A. Carter, *J. Chem. Phys.* **140**, 024102 (2014).
- ⁵⁹A. K. Wilson, T. van Mourik, and T. H. Dunning, *J. Mol. Struct.: THEOCHEM* **388**, 339–349 (1996).
- ⁶⁰T. Dunning, *J. Chem. Phys.* **90**(2), 1007–1024 (1989).
- ⁶¹M. R. Nyden and G. A. Petersson, *J. Chem. Phys.* **75**(4), 1843–1862 (1981).
- ⁶²J. M. L. Martin, *Chem. Phys. Lett.* **259**(5–6), 669–678 (1996).
- ⁶³J. M. L. Martin, *J. Mol. Struct.: THEOCHEM* **398–399**, 135–144 (1997).
- ⁶⁴A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**(3–4), 243–252 (1998).
- ⁶⁵D. G. Truhlar, *Chem. Phys. Lett.* **294**(1–3), 45–48 (1998).
- ⁶⁶A. G. Császár, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **108**(23), 9751–9764 (1998).
- ⁶⁷P. L. Fast, M. L. Sanchez, and D. G. Truhlar, *J. Chem. Phys.* **111**(7), 2921–2926 (1999).
- ⁶⁸D. Feller and D. A. Dixon, *J. Phys. Chem. A* **104**(13), 3048–3056 (2000).
- ⁶⁹A. J. C. Varandas, *J. Chem. Phys.* **113**(20), 8880–8887 (2000).
- ⁷⁰T. Helgaker, T. A. Ruden, P. Jørgensen, J. Olsen, and W. Klopper, *J. Phys. Org. Chem.* **17**(11), 913–933 (2004).
- ⁷¹S. E. Wheeler, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **121**(18), 8800–8813 (2004).
- ⁷²J. Noga, W. Kutzelnigg, and W. Klopper, *Chem. Phys. Lett.* **199**(5), 497–504 (1992).
- ⁷³J. Noga, D. Tunega, W. Klopper, and W. Kutzelnigg, *J. Chem. Phys.* **103**(1), 309–320 (1995).
- ⁷⁴J. M. L. Martin and O. Uzan, *Chem. Phys. Lett.* **282**(1), 16–24 (1998).
- ⁷⁵See supplementary material at <http://dx.doi.org/10.1063/1.4862159> for Table S1 comparing BDEs from the two extrapolation schemes, Table S2 showing unphysical BDEs due to MRACPF instability, and Figure S1 showing the full set of references used for all three active spaces.
- ⁷⁶F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, and R. Lindh, *J. Comput. Chem.* **31**(1), 224–247 (2010).
- ⁷⁷R. S. Mulliken, *J. Chem. Phys.* **23**(10), 1833–1840 (1955).
- ⁷⁸T. J. Lee and P. R. Taylor, *Int. J. Quantum Chem.* **36**(S23), 199–207 (1989).
- ⁷⁹H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, *WIREs Comput. Mol. Sci.* **2**(2), 242–253 (2012).
- ⁸⁰B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszewski, S. J. Bittner, S. G. Nijssure, K. A. Amin, M. Minkoff, and A. F. Wagner, *J. Phys. Chem. A* **108**(45), 9979–9997 (2004).
- ⁸¹B. Ruscic, R. E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoy, and A. F. Wagner, *J. Phys.: Conf. Ser.* **16**(1), 561–570 (2005).
- ⁸²S. J. Klippenstein, L. B. Harding, and B. Ruscic, “*Ab initio* computations and active thermochemical tables hand in hand: Heats of formation of core combustion species,” *J. Phys. Chem. A* (unpublished).
- ⁸³Y.-R. Luo, *Comprehensive Handbook of Chemical Bond Energies* (CRC Press, Boca Raton, 2007).
- ⁸⁴MRAQCC is used here in place of MRACPF2 because the Mulliken charges for MRACPF2 are not available in MOLCAS. MRAQCC and MRACPF2 use same *g*-values for single excitations.
- ⁸⁵R. C. Fortenberry, X. Huang, J. S. Francisco, T. D. Crawford, and T. J. Lee, *J. Chem. Phys.* **135**(21), 214303 (2011).
- ⁸⁶R. C. Fortenberry, X. Huang, J. S. Francisco, T. D. Crawford, and T. J. Lee, *J. Chem. Phys.* **135**(13), 134301 (2011).
- ⁸⁷J. S. Francisco, J. T. Muckerman, and H.-G. Yu, *Acc. Chem. Res.* **43**(12), 1519–1526 (2010).
- ⁸⁸J. T. Petty and C. B. Moore, *J. Mol. Spectrosc.* **161**(1), 149–156 (1993).
- ⁸⁹T. J. Sears, W. M. Fawzy, and P. M. Johnson, *J. Chem. Phys.* **97**(6), 3996–4007 (1992).
- ⁹⁰A. Rauk, D. Yu, P. Borowski, and B. Roos, *Chem. Phys.* **197**(1), 73–80 (1995).
- ⁹¹A. Rauk, D. Yu, and D. A. Armstrong, *J. Am. Chem. Soc.* **116**(18), 8222–8228 (1994).
- ⁹²S. D. Peyerimhoff, P. S. Skell, D. D. May, and R. J. Buenker, *J. Am. Chem. Soc.* **104**(17), 4515–4520 (1982).
- ⁹³A. D. McLean, B. H. Lengsfeld III, J. Pacansky, and Y. Ellinger, *J. Chem. Phys.* **83**(7), 3567–3576 (1985).
- ⁹⁴D. Feller, E. S. Huyser, W. T. Borden, and E. R. Davidson, *J. Am. Chem. Soc.* **105**(6), 1459–1466 (1983).
- ⁹⁵M. Kieninger, O. N. Ventura, and S. Suhai, *Int. J. Quantum Chem.* **70**(2), 253–267 (1998).
- ⁹⁶E. Garand, K. Klein, J. F. Stanton, J. Zhou, T. I. Yacovitch, and D. M. Neumark, *J. Phys. Chem. A* **114**(3), 1374–1383 (2010).
- ⁹⁷T. D. Crawford and J. F. Stanton, *J. Chem. Phys.* **112**(18), 7873–7879 (2000).
- ⁹⁸P. Y. Ayala and H. B. Schlegel, *J. Chem. Phys.* **108**(18), 7560–7567 (1998).
- ⁹⁹K. Klein, E. Garand, T. Ichino, D. Neumark, J. Gauss, and J. Stanton, *Theor. Chem. Acc.* **129**(3–5), 527–543 (2011).
- ¹⁰⁰J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B* **46**(11), 6671–6687 (1992).
- ¹⁰¹A. M. El-Nahas, M. V. Navarro, J. M. Simmie, J. W. Bozzelli, H. J. Curran, S. Dooley, and W. Metcalfe, *J. Phys. Chem. A* **111**(19), 3727–3739 (2007).
- ¹⁰²V. B. Oyeyemi, J. A. Keith, and E. A. Carter, “Accurate bond energies of biodiesel methyl esters from multireference averaged coupled-pair functional calculations,” *J. Phys. Chem. A* (submitted).