Energetics of the OH radical H-abstraction reactions from simple aldehydes and their geminal diol forms

Zoi Salta[1], Thomas Schaefer[2], Nicola Tasinato[1], Martina Kieninger[3], Aline Katz[3], Hartmut Herrmann[2], Oscar N. Ventura[3]*

1 Scuola Normale Superiore, piazza dei Cavalieri 7, 56126 Pisa, Italy
2 Atmospheric Chemistry Department (ACD), Leibniz Institute for Tropospheric Research (TROPOS), Permoserstr. 15, 04318 Leipzig, Germany
3 Computational Chemistry and Biology Group, CCBG, DETEMA, Facultad de Química, Universidad de la República, 11400 Montevideo, Uruguay

Submitted to

Journal of Molecular Modelling, April 2024[1]

*Corresponding author: Prof. Oscar N. Ventura, Computational Chemistry and Biology Group, CCBG, DETEMA, Facultad de Química, Universidad de la República, 11400 Montevideo, Uruguay, E-mail: Oscar.N.Ventura@gmail.com

ORCID

Zoi Salta	0000-0002-7826-0182
Thomas Schaefer	0000-0001-7995-4285
Nicola Tasinato	0000-0003-1755-7238
Martina Kieninger	0000-0002-5255-7148
Aline Katz	0000-0001-5134-8400
Hartmut Herrmann	0000-0001-7044-2101
Oscar N. Ventura	0000-0001-5474-0061

[1] This manuscript has been submitted as part of the "Festschrift" issue in honor of Alejandro Toro-Labbé, on the occasion of his 70th birthday.

Abstract:

Context: Carbonyl compounds, especially aldehydes, emitted to the atmosphere, may suffer hydration in aerosols or water droplets in clouds. At the same time, they can react with hydroxyl radicals which may add or abstract hydrogen atoms from these species. The interplay between hydration and hydrogen abstraction is studied using density functional and quantum composite theoretical methods, both in the gas phase and in simulated bulk water. The H-abstraction from the aldehydic and geminal diol forms of formaldehyde, acetaldehyde, glycolaldehyde, glyoxal, methylglyoxal and acrolein are studied to determine whether the substituent has any noticeable effect in the preference for the abstraction of one form or another. It is found that abstraction of the H-atom adjacent to the carbonyl group gives a more stable radical than same abstraction from the geminal diol in the case of formaldehyde, acetaldehyde and glycolaldehyde. The presence of a delocalizing group in the C_α (a carbonyl group in glyoxal and methylglyoxal, and a vinyl group in acrolein), reverts this trend and now the abstraction of the H-atom from the geminal diol gives more stable radicals. A further study was conducted abstracting hydrogen atoms from the other different positions in the species considered, both in the aldehydic and geminal diol forms. Only in the case of glycolaldehyde, the radical formed by H-abstraction from the $–CH_2OH$ group is more stable than any of the other radical species. Abstraction of the hydrogen atom in one of the hydroxyl groups in the geminal diol is equivalent to the addition of the $^{\bullet}OH$ radical to the aldehyde. It leads, in some cases, to decomposition into a smaller radical and a neutral molecule. In these cases, some interesting theoretical differences are observed between the results in gas-phase and (simulated) bulk solvent, as well as with respect to the method of calculation chosen.

Methods: DFT (M06-2X, B2PLYP, PW6B95), CCSD(T) and composite (CBS-QB3, jun-ChS, SCVECV-f12) methods using Dunning basis sets and extrapolation to the CBS limit were used to study the energetics of closed shell aldehydes in their keto and geminal-diol forms, as well as the radical derived from them by hydrogen abstraction. Both gas-phase and simulated bulk solvent calculations were performed, in the last case using the Polarizable Continuum Model.

Keywords: Glyoxal, Methylglyoxal, Hydration, Hydrogen abstraction, Atmospheric chemistry

1. INTRODUCTION

Atmospheric chemistry is a subject of continuous and growing interest in the scientific community. Older work was mainly focused on gas-phase reactivity, aiming principally to determine reaction mechanisms, participant species, and the thermochemistry and kinetics of said mechanisms. In more recent years, the importance of water mediated chemistry in the atmosphere has been highlighted, in relation to phenomena like aerosol particle growth, cloud formation, and contamination [1–5]. Over the past decade, it has become clear that aqueous chemical processes occurring in cloud droplets and wet atmospheric particles are an important source of organic atmospheric particulate matter. Reactions of water-soluble volatile (or semivolatile) organic gases (VOCs or SVOCs) in these aqueous media lead to the formation of highly oxidized organic particulate matter (secondary organic aerosol; SOA) and key tracer species, such as organosulfates [4].

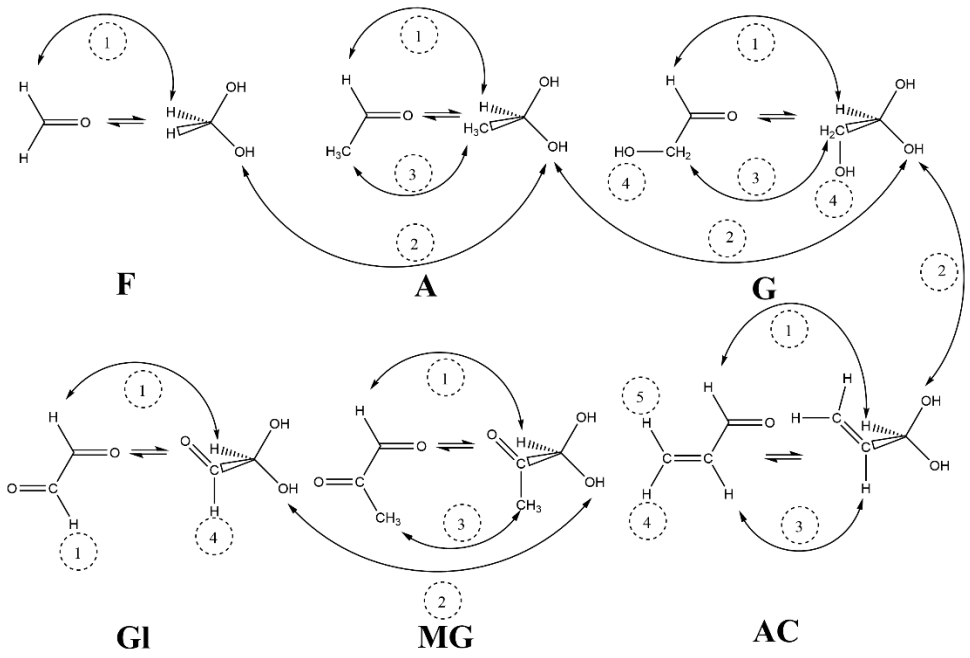
Atmospheric aerosols play an important role in the Earth's radiative balance [6,7], which, however, is not fully understood yet. Particularly, the formation of Secondary Organic Aerosols (SOA) remains uncertain, as a large SOA source seems to be missing e.g. small volatile organic compounds (VOC's) from current atmospheric models.

An increasing amount of research has been focused on carbonyl chemistry in water droplets in the atmosphere, especially that of acids [8] and α-oxoaldehydes [9-11]. Organic acids, like glutaric, adipic [8], pyruvic, and oxalic acids, have been found in SOA, and there are studies suggesting that they originate from the oxidation of aldehydes [2,12-14]. Among the latter, glyoxal, glycolaldehyde, and methylglyoxal attracted ample attention. Aldehydes in water may suffer hydration rather easily, depending on the substituent attached to the aldehydic carbon atom. The best-known case is formaldehyde, where the equilibrium of the $H_2C{=}O + H_2O \rightarrow H_2C(OH)_2$ reaction is displaced to the right. Substitution of one of the hydrogen atoms by a methyl group, as in acetaldehyde, diminishes markedly the preference towards the diol species, but it is kept if the C_α is another carbonyl like in α-oxo aldehydes. In this context, glyoxal [9,10,15] and methylglyoxal [10,12,13-21] have been frequently studied in recent years.

Oxidation in the atmosphere occurs via photochemical reactions or by the action of radicals like $^{\bullet}OH$, $NO_3^{\bullet}$ or $^{\bullet}Cl$. In the latter case, the most frequent reactions are H-abstraction and addition. Although hydrogen abstraction is more readily observed from alkyl CH_n groups, other hydrogens may be susceptible to abstraction in some conditions [22]. The hydration of aldehydes, *i.e.*, the nucleophilic addition of water to the carbonyl group to form hydrated geminal species [23], dramatically affects their volatility and Henry's law solubility [24]. Since carbonyl species are an important component of SOA, it is then pertinent to ask whether the aldehydic or geminal diol forms are more readily attacked by the hydroxyl radical in atmospheric systems. Thus, the central question we seek to answer in this study is whether H-abstraction is easier to occur from the oxo or the diol forms, depending on the substituent attached to the aldehydic carbon.

2. COMPUTATIONAL METHODS

The species selected for this work were formaldehyde (methanal), acetaldehyde (ethanal), glycolaldehyde (hydroxyethanal), glyoxal (ethanedial), methylglyoxal (2-oxopropanal), and acrolein (vinylacetaldehyde). For reasons of economy, we will use the abbreviations **F**, **A**, **G**, **Gl**, **MG**, and **AC** respectively, to identify the oxo species, followed by *diol* to identify the corresponding *geminal* diol, and *rad* to identify the radical species. If H-abstraction from different sites may lead to different radicals, we identify them by ordinal numbers, so that *1* always identifies the abstraction from the C atom bearing the aldehydic or *gem* diol groups, *2* identifies the radical obtained by abstraction from the most exposed OH substituent in the *gem* diol group, *3* is reserved for the H-abstraction from the CH_3 group in **A** or **MG**, CH_2 in **G**, and the proximal H in the vinyl group of **AC**, *4* is used in the case of **G** to identify the radical obtained abstracting the hydrogen from the OH in the $–CH_2OH$ substituent group of **G** or the cis H in the CH_2 of the vinyl group of **AC**, while *5* is used to label the radical resulting from abstraction of the trans H in the CH_2 of the vinyl group of **AC**. This numbering is shown explicitly in Scheme 1.



Scheme 1. Numbering of the hydrogen abstraction sites to obtain the radicals studied in this work. Notice that only mono-hydration at the more reactive carbonyl has been considered when there is more than one possibility.

According to the calculations performed by Elrod et al. [24], the $log_{10}K_H$ values for **F**, **A**, **G**, **Gl** and **MG** are 3.4, 0.5, 1.6, 3.3 and 2.8 respectively, thus providing the order **F** > **Gl** > **MG** > **G** > **A** for the extension of hydration (no values are provided for acrolein). These data will have to be later combined with the values obtained in this study to assess the atmospheric implications of the combined hydration/H-abstraction pathways.

In this study, we used density functional (DFT) [25–28], coupled-cluster theory using singles, double and perturbative triple excitations CCSD(T) [29], and wavefunction-based composite methods [30–32] to obtain optimum geometries, frequencies, energies, and thermochemical properties of the different species. The M06-2X [33], PW6B95 [34] and B2PLYP [35] DFT methods were employed. Grimme's empirical dispersion [36] was included for the first two methods, while no empirical dispersion was included for B2PLYP. Dunning's aug-cc-pVTZ [37] basis set was used within the M06-2X and B2PLYP calculations, while Truhlar's jun–cc–pVTZ was employed for PW6B95 [38]. CCSD(T) calculations were performed on

top of the B2PLYP [34] optimum geometries with the aug-cc-pVTZ basis set [37] without further optimization. Thermochemical properties were obtained using the thermal contributions obtained at the B2PLYP level. A comparison with the thermal contributions obtained at the CCSD(T) level was performed in the case of **F** and **A** and shown to be nearly equal. The simple CBS-QB3 [39, 40] composite protocol, which relies on the B3LYP [41] DFT method to calculate optimum geometries and CCSD(T) [29], MP4SDQ [42], and MP2 [43] calculations with different basis sets to calculate the energies, was one of the procedures employed for the fast determination of all the species. Two more sophisticated wavefunction-based composite methods were also employed, namely SVECV-f12 [44–46] and jun-Cheap (jChS) [47–49], which in both cases imply the calculation of optimum geometries at the DFT level, calculation of the complete basis set (CBS) [50] energy on this fixed geometry by CCSD(T)-F12 [51] methods in the first case, and a combination of MP2 and CCSD(T) in the second case, and adding core-valence correlation energy employing MP2 calculations using the cc-pwCVTZ basis set [52]. A simulated bulk water environment was introduced in the calculation employing the self-consistent reaction field developed by Tomasi and coworkers, called Polarizable Continuum Model (PCM) [53]. As usual, water was simulated by introducing its experimental dielectric constant ($\varepsilon = 78.355$). DFT, MP2, and CBS-QB3 calculations were performed using Gaussian 16, Revision C.01 [54] while the F12 calculations were done using the Molpro2020 package [55–57].

3. RESULTS AND DISCUSSION

3.1 Aldehyde vs. geminal diol H-abstraction

The most significant results in this study from the point of view of atmospheric chemistry are the stability of the radicals formed from the oxo and geminal diol forms of the species considered. The aldehydes and their geminal diol forms, as well as the radicals obtained from H-abstraction from the carbonyl or gem-diol C atoms are shown in Figure 1. The numerical results obtained for each of the molecules are shown in Tables SI1-SI6 in the Supplementary Information (SI). A summary of the most important results is contained in Table 1, where the free energies (ΔG) at 298.15 K for the reaction $X + HO^{\bullet} \rightarrow X^{\bullet} + H_2O$, with

X either the aldehyde or the geminal diol, are collected for the two most accurate methods used in this study, jChS and SVECV-f12. In the first case, results are shown both in the gas-phase and in bulk, but in the second case only gas phase results are shown because the PCM model cannot be applied to the CCSD(T)-F12 components of the SVECV-f12 composite method.

The radical generated from the aldehyde is more stable with respect to the parent structure than the radical obtained from the geminal diol in the case of **F**, **A**, and **G**. The opposite is true for **Gl**, **MG** and **AC**. This conclusion is valid both for the gas phase and in the bulk solvent, and, except marginally for the case of **AC**, the bulk solvent makes the radicals more stable with respect to the parent species when they are in the gas-phase, although the effect is relatively small. However, it should be stressed that PCM, lacking a description of hydrogen bonded interactions, is not the best approach to model water solvation. It is probable that explicit treatment of water molecules offers a different picture and a more accurate description of local effects.

Table 1. ΔG (298.15 K) for the reaction X + $HO^{\bullet}$ → $X^{\bullet}$ + H_2O, where X is either the aldehyde or the geminal diol. Units are kJ mol^{-1}.

Species	Form	Gas-phase		Bulk water
		jChS	SVECV-f12	jChS
Formaldehyde	Oxo	-133.9	-130.3	-133.1
	Gem-diol	*-98.4*	*-97.6*	*-103.1*
Acetaldehyde	Oxo	-130.0	-128.8	-136.6
	Gem-diol	*-111.8*	*-110.0*	*-111.3*
Glycolaldehyde	Oxo	-121.8	-120.0	-123.5
	Gem-diol	*-112.9*	*-111.7*	*-112.7*
Glyoxal	Oxo	-133.2	-133.6	-140.1
	Gem-diol	***-183.3***	***-183.9***	***-190.5***
Methylglyoxal	Oxo	-117.9	-116.3	-128.4
	Gem-diol	***-173.4***	***-173.1***	***-187.7***
Acrolein	Oxo	-122.7	-122.3	-122.7
	Gem-diol	***-179.5***	***-175.4***	***-169.1***

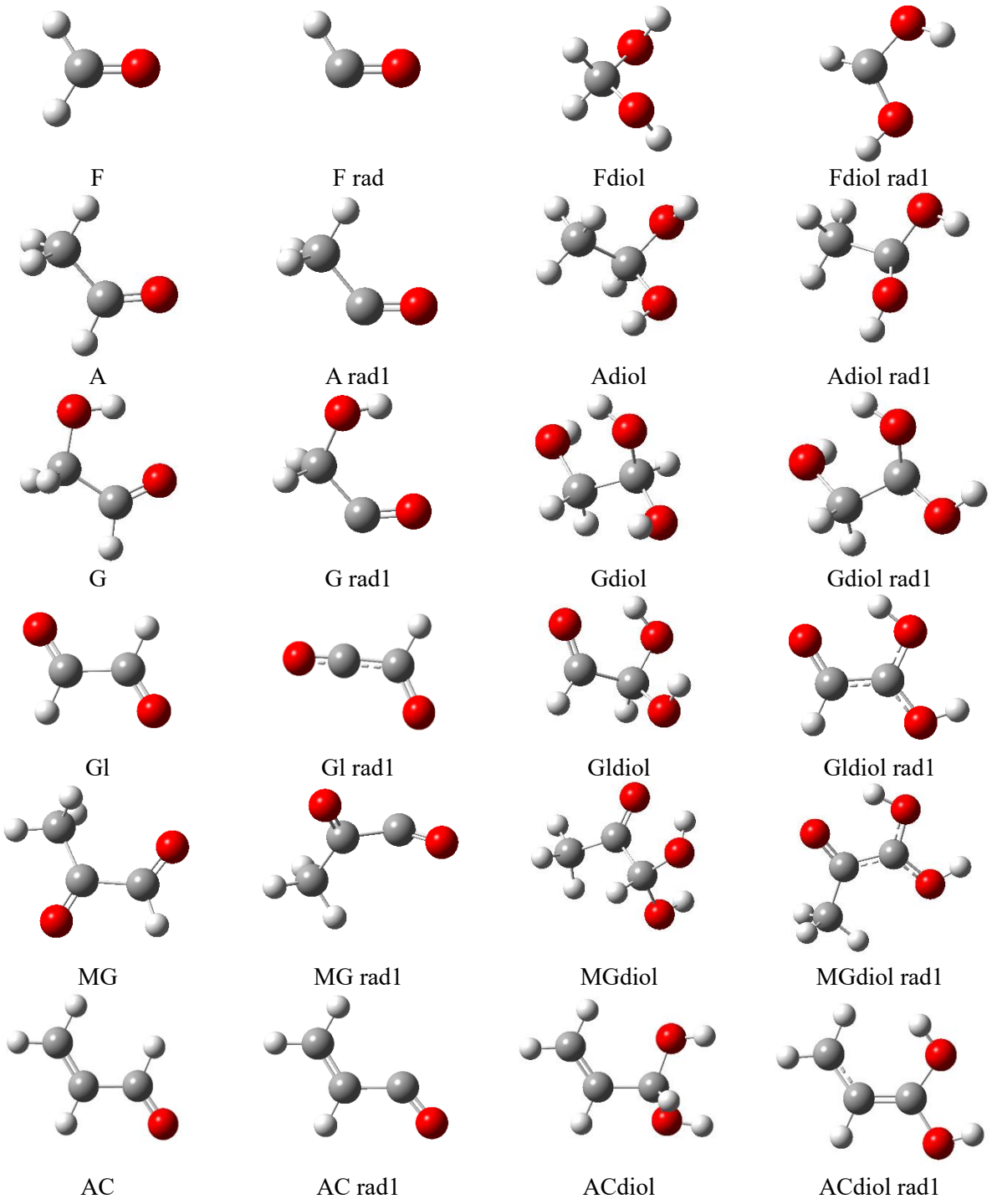


Figure 1. Structure of the aldehydes and the geminal diols considered in this study, together with that of the radicals formed by removal of a hydrogen atom from the carbon bearing the carbonyl and gem diol functions respectively.

The observation of the previous paragraph can be easily rationalized in terms of the delocalization provided by the double bond of the substituent to the unpaired electron generated over the carbon atom when the hydrogen is abstracted. This is particularly noticeable in the structures of the radicals of the gem-diols of **Gl**, **MG** and **AC**, which exhibit a planar arrangement of atoms around the carbon atom bearing the unpaired electron. The whole process has two components, hydration, and hydrogen abstraction. Whether the hydrogen will be preferentially abstracted from the aldehyde or, after hydration, from the geminal diol, depends both on thermochemical and kinetic considerations. We have collected in Table 2 the free energies for direct abstraction from the aldehyde, hydration, and hydrogen abstraction from the geminal diol with respect to the aldehyde plus a hydroxyl radical. We show the numbers only at one of the best levels of theory used (jChS), and in the gas phase, since all methods give approximately the same results.

Table 2. Abstraction free energies (ΔG) from the C1 carbon atom in the aldehyde and the gem diol radicals with respect to their parent structures. Hydration energies of the aldehydes with respect to the aldehyde plus the hydroxyl radical, and gem diol radical stability with respect to the aldehyde and the hydroxyl radical. All free energies calculated at 298.15 K are given in kJ mol^{-1}.

Species	R–CHO→ R–C˙O	R–CHO→ R–$CH(OH)_2$	R–$CH(OH)_2$→ R–C˙$(OH)_2$	Gem diol radical stabilization
Formaldehyde	-133.9	1.5	-96.8	-95.3
Acetaldehyde	-130.0	26.4	-85.3	-58.9
Glycolaldehyde	-121.8	24.8	-88.2	-63.4
Glyoxal	-133.2	3.5	-179.8	-176.3
Methylglyoxal	-122.3	4.9	-168.5	-163.6
Acrolein	-122.7	40.3	-139.2	-98.9

F, **Gl** and **MG** aldehydic and gem-diol forms are in equilibrium, while the aldehydic form is largely favored for **A**, **G** and especially **AC**. Since abstraction from the gem diol form is favored for **Gl**, **MG** and **AC**, it is interesting to see if the whole process of hydration + hydrogen abstraction is also more favorable in these three species. As shown in the last column of Table 2, this is true for **Gl** and **MG**, but not for **AC**, since the gem diol suffers a larger destabilization with respect to the aldehydic form.

It is interesting then to point out this singularity of glyoxal and methylglyoxal which are the most abundant α-dicarbonyls in the atmosphere [12,58-60]. In their gas-phase study of the water-mediated equilibrium

between methylglyoxal and its geminal diol, Axson et al. [17] found that the gas-phase hydration of methylglyoxal is possible at low relative humidity (RH% < 5%) and the amount of geminal diol should be larger than previously assumed. These results apply equally well to glyoxal so that, as shown in Table 2, hydrogen abstraction from C1 will occur preferentially from the geminal diol and not from the aldehyde.

3.2 Abstraction vs addition

Besides the abstraction from the geminal diol carbon atom, a hydrogen atom can be abstracted also from one of the hydroxyl groups. The resulting radical is the same product that would be obtained by ˙OH addition to the corresponding aldehyde. The six structures of the respective radicals are shown in Figure 2. Since both addition and H-abstraction are processes observed in atmospheric chemistry, it is interesting to compare their energies for the aldehydes we considered in this work. We collected in Table 3 those energies in the gas phase with the most accurate methods used in this study.

Table 3. Stability of the radicals obtained after OH addition to the aldehyde and H-abstraction from the aldehyde respectively. All free energies calculated at 298.15 K are given in kJ mol^{-1}.

Species	Addition			Abstraction		
	jChS	SVECV-f12	Average	jChS	SVECV-f12	Average
Formaldehyde	-59.1	-57.7	**-58.4**	-96.8	-95.5	**-96.2**
Acetaldehyde	-39.7	-38.1	**-38.9**	-85.3	-84.5	**-84.9**
Glycolaldehyde	-29.6	-31.3	**-30.4**	-88.2	-87.3	**-87.7**
Glyoxal	-58.2	-60.0	**-59.1**	-179.8	-182.1	**-180.9**
Methylglyoxal	-56.6	-61.7	**-59.1**	-168.5	-171.8	**-170.2**
Acrolein	-26.2	-25.9	**-26.0**	-139.2	-135.1	**-137.2**

It is immediately obvious from the table that the radical resulting by H-abstraction from the carbon bearing the geminal diol group is much more stable than the radical produced by the addition of ˙OH to the aldehyde species. As we commented before, this is the effect of the delocalization of the unpaired electron when it is located over the carbon atom, an effect that is not present in the diols *rad2* structures, where the unpaired electron is located over the oxygen formerly present in the carbonyl group.

It is also observed that the methyl group in **MG** does not alter the energy of the addition product with respect to that of **Gl**. These two, together with **F**, produce the most stable addition products and, simultaneously, the most stable abstraction products. Acrolein, on the other hand, gives rise to an addition product almost as stable as that of glycolaldehyde, notwithstanding the very different structure of the radicals. The abstraction radicals, however, show a very different stability. Since the *ACdiol rad1* is highly delocalized, as in the case for **Gl** and **MG**, the radical is much more stable than the corresponding G radical, where there is a fair amount of H-bonding but no delocalization. *ACdiol rad2* is, however, less stable than *Gldiol rad2* or *MGdiol rad2* because there is a stabilizing hydrogen bond between the hydrogen in the hydroxyl group and the oxygen in the carbonyl group, which is not present in *ACdiol rad2*.

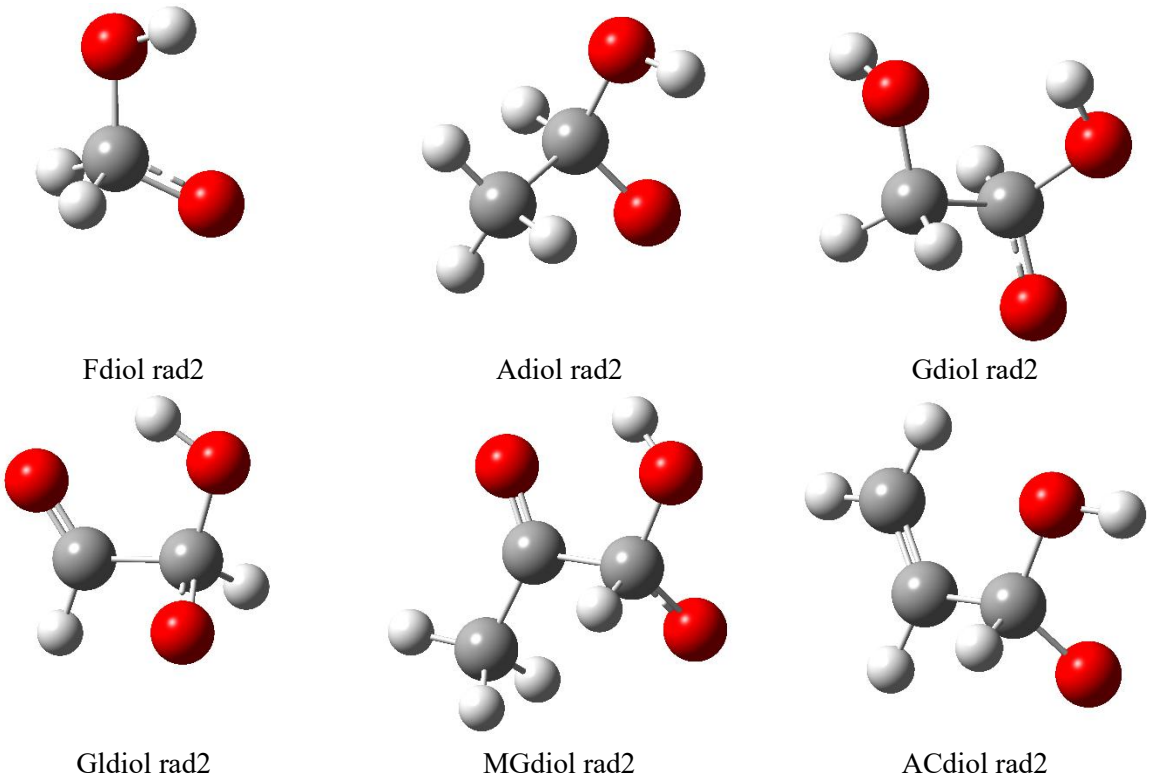


Figure 2. Structure of the radicals obtained from the gem diol forms of the aldehydes by removal of the hydrogen atom from one of the OH groups. They are the same structures that are obtained by addition of the ˙OH radical to the aldehydes.

3.3 Other H-abstractions

Besides the abstraction from C1 (the carbon atom bearing the carbonyl or the gem diol functionalities), or one of the hydroxyl groups in the diol (which gives rise to the addition product), other hydrogen abstractions are possible: in A from the methyl group, in **G**, from the $–CH_2OH$ group carbon or oxygen atoms, in **MG** from the methyl group, and in **AC** from the two carbons in the vinyl substituent. These structures are shown in Figure 3. The radicals from the aldehyde forms are collected in the upper rows and those from the diols in the lower ones.

Several of the alternative radicals from the aldehydes exhibit resonant structures, but only the radical from *MGdiol* does have this resonance. In principle, one could guess that those radicals would be more stable, but, as we will see in the next section, this is mostly not true. The only case in which an alternative radical is more stable than that obtained from hydrogen removal of the aldehyde group is *G rad3*, which is more stable than *G rad1*. This fact does not change any conclusion, because already *G rad1* was more stable than any of the *Gdiol* radicals.

An interesting observation on the structure of these radicals is that some of them might be dissociated to a closed shell molecule and a simpler radical. For instance, the *Gdiol rad 3* for glycolaldehyde can dissociate into formaldehyde and *Fdiol rad1*, *Gdiol rad4* into CO + *Fdiol rad1* (same happens to the isomers *Gldiol rad3* and *Gldiol rad4*), etc. We did not investigate fully these decompositions but faced some spontaneous transformations depending on the method employed, as will be described later. It must be mentioned in this context, however, that Shannon et al. [61] recently performed a theoretical study of the decomposition of *Gl rad* which sheds some light on the relative importance of these reactions vis-à-vis the more common path of reaction with O_2.

3.4 Considerations on the accuracy of the methodologies

We have collected in Table 4 and 5 the free energies with the DFT, CCSD(T), and composite methods for all the radicals studied in this work, obtained from both the aldehyde and geminal diol forms of the species.

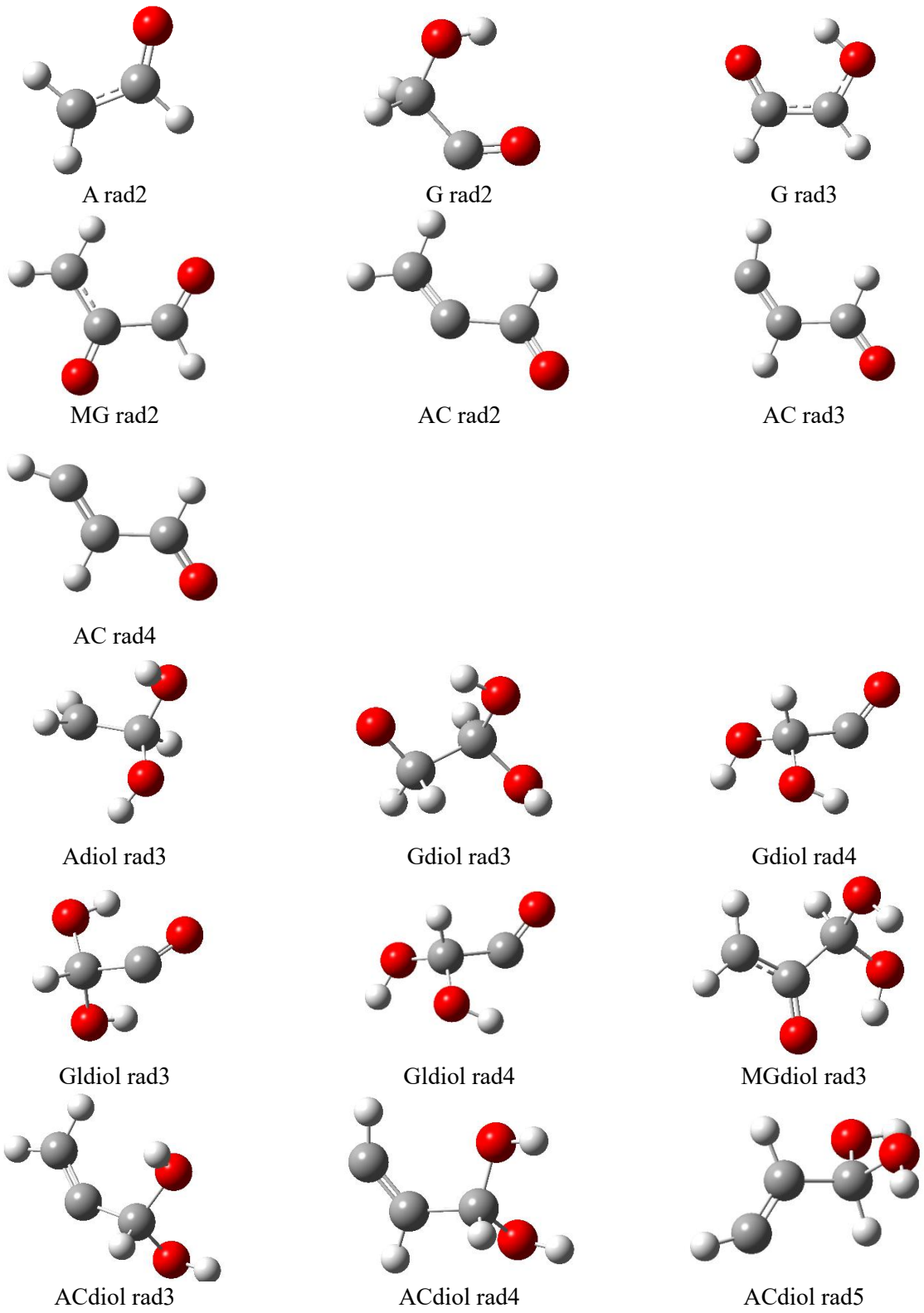


Figure 3. Structure of the radicals obtained by other hydrogen abstractions from **A** to **AC**, and from *Adiol* to *ACdiol*.

When performing theoretical calculations, it is crucial to control the expected accuracy of the results. Both from the theoretical and practical points of view, composite methods, like jChS and SVECV-f12, are the most accurate feasible procedures applicable from medium to large molecules. Their accuracy is expected to be better than 4 kJ mol^{-1} which is normally more than appropriate to study thermochemical properties like the ones investigated in this study.

Composite methods depend on DFT for the geometry optimization and thermochemical components of the final energy, and on post-Hartree-Fock procedures, like MP2 and CCSD(T), to obtain the correlation energy. An essential component is the extrapolation to the complete basis set, which means that the amount of correlation energy included is larger than that in the simpler procedures. Finally, the use of theoretically different addition procedures in jChS and SVECV-f12 implies that agreement between their results assures a reasonable concordance with the experimental results, when and if they are obtained.

Considering this reasoning is that we built Tables 4 and 5 where the free energy of stabilization of the radicals with respect to their parent species is compared for all methods, first in the gas phase and then in simulated water solution. Highlighted in grey in these tables are the most stable hydrogen abstractions both from the aldehyde and the geminal diol, as we discussed before. What is important in this section is that all methods applied afford the same picture concerning which is the most stable radical product. Some blanks in the tables correspond to difficult cases that will be described in the next section.

Since both jChS and SVECV-f12 are the best methodologies and their results are pretty close in gas phase, as shown by their absolute difference in Table 4 (one-before-last column), we used the average of their values as the most reliable estimate. Notice in the last two lines of Table 4 that the average absolute error between both methods is just 1.9 kJ mol^{-1} with a spread (given by the standard deviation of the errors) of 1.6 kJ mol^{-1}. This implies that the differences between the methods are not systematic, probably due to the different paths taken by the two composite methods for obtaining the final energy. Compared to these average values, the other methods exhibit different situations.

Table 4. Relative free energies in the gas phase of the radicals obtained from the aldehyde and geminal diol forms with respect to their parent molecules for all methods used in this study. Shadowed cells indicate the most stable radicals for each species and underlined values are those differing by more than 4.2 kJ mol^{-1} from the average jChS and SVECV-f12 results. Blank cells are those for which the species could not be calculated at that level (see text). Units of kJ mol^{-1}.

		M06-2X-D3		PW6B95-D3		B2PLYP		CCSD(T)		jChS	SVECV-f12	Absolute	Average
Species	Form	Value	Difference	Value	Difference	Value	Difference	Value	Difference			Diff.SV-jC	
F	OXO CH	-126.7	5.5	-129.5	2.6	-132.0	0.1	-114.2	17.9	-133.9	-130.3	3.6	-132.1
	DIOL CH	-97.5	0.5	-99.3	1.3	-98.9	0.9	-79.5	18.6	-98.4	-97.6	0.7	-98.0
	DIOL OH	-61.0	0.8	-74.4	14.2	-63.7	3.5	-45.8	14.4	-60.6	-59.8	0.8	-60.2
A	OXO CH	-125.2	4.2	-126.6	2.8	-129.4	0.0	-111.7	17.7	-130.0	-128.8	1.2	-129.4
	OXO CH3	-99.7	1.2	-98.4	0.2	-97.0	1.6	-81.1	17.5	-97.7	-99.5	1.8	-98.6
	DIOL CH	-110.5	0.4	-114.5	3.6	-114.3	3.4	-93.0	18.0	-111.8	-110.0	1.7	-110.9
	DIOL OH	-64.3	5.8	-79.1	8.9	-70.4	0.3	-54.1	16.0	-70.9	-69.3	1.6	-70.1
	DIOL CH3	-68.6	3.7	-64.8	0.1	-70.4	5.4	-52.6	12.3	-66.2	-63.7	2.5	-64.9
G	OXO CH	-116.3	4.6	-117.6	3.3	-121.3	0.4	-103.4	17.5	-121.8	-120.0	1.8	-120.9
	OXO OH	-45.1	0.1	-56.3	11.3	-49.3	4.3	-32.5	12.4	-44.7	-45.2	0.5	-45.0
	OXO CH2	-168.7	3.9	-174.4	9.6	-168.6	3.8	-144.6	20.2	-164.5	-165.2	0.7	-164.8
	DIOL CH	-112.5	0.1	-116.7	4.3	-115.1	2.8	-93.3	19.1	-112.9	-111.7	1.2	-112.3
	DIOL OH g	-55.7	0.7	-70.2	15.2	-59.7	4.7	-41.2	13.8	-54.3	-55.7	1.3	-55.0
	DIOL OH s	-58.4	4.3	-68.4	14.2	-60.3	6.2	-42.8	11.3	-50.9	-57.4	6.5	-54.1
	DIOL CH2	-107.5	0.6	-110.7	2.6	-109.5	1.5	-89.5	18.5	-108.8	-107.3	1.5	-108.0
Gl	OXO CH	-145.4	12.1	-152.5	19.1	-142.8	9.4	-113.7	19.7	-133.2	-133.6	0.4	-133.4
	DIOL CH	-195.1	11.5	-203.4	19.8	-192.0	8.4	-161.3	22.3	-183.3	-183.9	0.6	-183.6
	DIOL OH	-61.7	0.1	-81.6	19.8					-61.7	-61.9	0.2	-61.8
	DIOL CHO	-113.0	3.0	-113.4	2.7	-117.4	1.4	-99.3	16.7	-116.6	-115.5	1.0	-116.0
	DIOL CHO best	-116.8	2.9	-118.6	1.0	-121.7	2.1	-102.0	17.6	-120.4	-119.0	1.4	-119.7
MG	OXO CH	-133.0	9.9	-140.1	17.0	-137.5	14.4	-113.8	9.3	-122.3	-124.0	1.7	-123.1
	OXO CH3	-101.2	4.2	-98.7	1.7	-96.7	0.3	-79.5	17.6	-94.4	-99.7	5.3	-97.1
	DIOL CH	-182.5	9.3	-191.4	18.1	-181.3	8.1	-151.4	21.8	-173.4	-173.1	0.4	-173.2
	DIOL OH	-63.5	2.2	-79.8	18.6					-61.5	-61.0	0.5	-61.3
	DIOL CH3	-96.9	2.0	-94.7	0.2	-94.7	0.2	-77.7	17.2	-94.3	-95.5	1.1	-94.9

AC	OXO CHO	-118.9	3.6	-120.3	2.1	-122.9	0.5	-104.7	17.7	-122.7	-122.3	0.4	-122.5
	OXO CH	-32.7	2.8	-31.1	1.3	-29.5	0.4	-11.9	17.9	-28.2	-31.4	3.2	-29.8
	OXO CH2 cis	-33.9	3.3	-28.1	2.4	-28.5	2.0	-13.5	17.0	-28.7	-32.4	3.7	-30.5
	OXO CH2 trans	-34.7	3.7	-28.3	2.7	-29.1	1.8	-14.1	16.8	-29.1	-32.8	3.7	-30.9
	DIOL CH(OH)2	-182.3	4.8	-194.1	16.6	-180.9	3.4	-154.0	23.5	-179.5	-175.4	4.1	-177.5
	DIOL OH	-67.6	1.3	-79.5	13.2	-69.2	2.9	-52.0	14.3	-66.5	-66.2	0.3	-66.3
	Diol CH	-40.4	2.7	-36.6	1.1	-37.3	0.5	-21.1	16.6	-37.3	-38.2	0.9	-37.7
	Diol CH2 cis	-35.3	3.3	-28.4	3.6	-30.4	1.6	-16.7	15.3	-32.6	-31.4	1.1	-32.0
	Diol CH2 trans	-35.1	0.3	-34.2	1.1	-36.7	1.4	-22.1	13.2	-38.3	-32.4	6.0	-35.3
	Average Error		3.5		7.5		3.0		16.9			1.9	
	Std. Dev. Error		3.1		7.0		3.2		3.1			1.6	

B2PLYP, the most accurate of the DFT methods used, exhibits an error larger than 4.2 kJ mol^{-1} (a threshold normally called chemical accuracy[2]) only in a few cases. The other two DFT methods, M06-2X-D3 and PW6B95-D3, are over this threshold on more occasions. This shows that, while the DFT methods here used to obtain the equilibrium geometry of the stable bound species on top of which the other calculations are performed are equivalent (provided there are no multi-conformation or degeneracy problems, as we showed in previous publications [44,63]) not all of them are equally useful to obtain relative energies, unless further corrections are carried on. The size of the average error of the three DFT procedures, B2PLYP (3.0 kJ mol^{-1}), M06-2X-D3 (3.5 kJ mol^{-1}), and PW6B95-D3 (7.5 kJ mol^{-1}), shows that these methods are reasonably appropriate for this type of study on larger molecules for which the coupled-cluster calculations cannot be performed, although B2PLYP should be the method of choice, resources allowing.

Interestingly, CCSD(T)/aug-cc-pVTZ calculations always show an error much larger than the threshold chosen of 4.2 kJ mol^{-1}. This situation might be caused by five types of drawbacks, namely i) the B2PLYP optimum equilibrium geometry used ; ii) the larger amount of correlation included in the composite methods because of the F12 component; iii) the size of the basis set; iv) the use of the DFT thermochemical corrections; or v) the lack of the core-valence correlation energy. This last item is not a reason for the failure since the comparison of free energies of the composite methods with and without such corrections exhibits the same trend. For the fourth error, we can compare the results using the DFT corrections or calculating them at the CCSD(T) level, a much more resource-intensive procedure. We did this comparison for formaldehyde and acetaldehyde and the results are shown in Table 6. A small average difference of 1.5 kJ mol^{-1} was found, with no significant difference between gas phase and bulk solvent, implying that this is not the cause of the large error found in gas phase. Similarly, the geometry used is unlikely to be the cause of the discrepancy, since its use in the composite method calculations did not introduce any large error.

[2] Chemical accuracy is actually an abbreviation for thermochemical accuracy, a term that was pushed forward by John A. Pople in the development of his Gn methods. The justification comes from the fact that a 1.4 kcal mol^{-1} error in the free energy produces approximately a factor of 10 error in equilibrium constants and rates coefficients. Thus, 4.1 kJ mol^{-1}, which is roughly equivalent to 1 kcal mol^{-1}, allows for an agreement of experimental and theoretical thermochemistry values within one order of magnitude.

Notice also that the energies obtained at the B2PLYP level by itself (i.e. without further correction) show much smaller deviations from our best values. Furthermore, the accuracy of CCSD(T), as well as of the other DFT approximations used in this work, for equilibrium geometries is well established [64-66]. Therefore, it is possible to conclude that the error arises from the inadequacy of the basis set size at this level of calculation and/or the lack of valence correlation energy (which is included in the CCSD(T)-F12 methods and higher orders of CCSD, CSDTQ5 for instance). This agrees with previous findings in the investigation of the formation pathway of ethanimine, where the CCSD(T)/aug-cc-pVTZ computations showed an average error of 14.5 kJ mol^{-1} from reference values considering both CBS and core-valence contributions [67].

Table 5. Relative free energies in bulk solvent of the radicals obtained from the aldehyde and geminal diol forms with respect to their parent molecules for all methods used in this study. Shadowed cells indicate the most stable radicals for each species and underlined values are those differing by more than 9.4 kJ mol^{-1} from the jChS results. Units of kJ mol^{-1}.

		M06-2X-D3		PW6B95-D3		B2PLYP		CCSD(T)		jChS
Species	Form	Value	Difference							
F	OXO CH	-126.4	6.7	-129.3	3.7	-131.2	1.8	-136.4	3.3	-133.1
	DIOL CH	-103.1	0.0	-104.6	1.5	-103.9	0.8	-107.5	4.4	-103.1
	DIOL OH	-64.7	1.1	-77.0	13.3	-69.6	5.9	-53.1	10.6	-63.7
A	OXO CH	-126.0	10.6	-127.6	9.0	-129.8	6.8	-135.0	1.6	-136.6
	OXO CH3	-106.7	3.7	-104.2	6.1	-103.9	6.5	-110.7	0.3	-110.4
	DIOL CH	-112.8	1.5	-113.9	2.6	-114.6	3.3	-116.5	5.2	-111.3
	DIOL OH	-64.0	11.5	-76.3	0.9	-67.0	8.5	-82.7	7.2	-75.5
	DIOL CH3	-75.5	12.1	-68.9	5.4	-75.8	12.4	-72.3	8.8	-63.4
G	OXO CH	-118.3	5.2	-120.3	3.2	-123.9	0.3	-128.4	4.8	-123.5
	OXO OH	-48.8	0.9	-60.6	12.7	-52.9	5.0	-58.6	10.7	-47.9
	OXO CH2	-172.2	4.8	-177.6	10.1	-172.1	4.7	-170.5	3.1	-167.4
	DIOL CH	-113.3	0.6	-116.4	3.7	-114.9	2.2	-116.2	3.5	-112.7
	DIOL OH g	-65.6	5.7	-81.5	21.6	-69.6	9.7	-71.5	11.6	-59.9
	DIOL OH s	-58.8	1.2	-69.9	12.3	-61.1	3.5	-66.4	8.8	-57.6
	DIOL CH2	-96.6	13.6	-111.8	1.7	-111.1	1.0	-114.1	3.9	-110.2
Gl	OXO CH	-153.0	12.9	-159.5	19.4	-150.5	10.4	-143.4	3.3	-140.1
	DIOL CH	-203.1	12.6	-210.9	20.4	-199.9	9.4	-191.5	1.0	-190.5
	DIOL OH	-67.1	2.5	-89.1	24.5					-64.6
	DIOL CHO	-115.5	3.6	-116.7	2.5	-119.9	0.7	-124.3	5.1	-119.2
	DIOL CHO bst	-122.3	3.3	-124.5	1.0	-127.4	1.8	-130.5	4.9	-125.6
MG	OXO CH	-139.6	11.2	-140.9	12.5	-140.9	12.5	-139.6	11.2	-128.4

	OXO CH3	-108.2	20.1	-104.9	23.5	-103.8	24.5	-109.0	19.4	-128.4
	DIOL CH	-189.5	1.8	-206.8	19.1	-187.1	0.6	-180.0	7.7	-187.7
	DIOL OH									
	DIOL CH3	-104.2	5.3	-111.1	1.5	-101.5	8.1	-107.3	2.2	-109.6
AC	OXO CHO	-119.4	3.3	-120.9	1.8	-122.9	0.2	-127.9	5.2	-122.7
	OXO CH	-39.0	3.5	-37.2	1.7	-33.2	2.4	-40.1	4.6	-35.5
	OXO CH2 cis	-38.8	4.3	-32.6	1.9	-33.2	1.3	-42.4	7.9	-34.5
	OXO CH2 trans	-39.6	4.8	-29.1	5.7	-33.8	1.0	-42.9	8.1	-34.8
	DIOL CH(OH)2	-178.8	9.6	-184.4	15.3	-176.3	7.2	-172.7	3.6	-169.1
	DIOL OH	-62.7	0.0	-75.3	12.6	-65.3	2.5	-71.9	9.1	-62.8
	Diol CH	-46.9	2.5	-44.1	0.2	-47.2	2.9	-54.0	9.7	-44.3
	Diol CH2 cis	-41.9	2.8	-35.0	4.1	-36.9	2.2	-46.9	7.8	-39.1
	Diol CH2 trans	-42.3	1.8	-36.6	3.9	-41.7	1.1	-50.8	10.3	-40.5
	Average Error		5.6		8.5		5.0		6.5	
	Std. Dev. Error		4.9		7.5		5.1		3.9	

The fact that the average error of CCSD(T) is large (16.9 kJ mol^{-1}) but the spread is even smaller than for the DFT methods (3.1 kJ mol^{-1}) suggests that the origin of the error is systematic, supporting the previous conclusion. Whatever the case, it shows that the simple CCSD(T)/aug-cc-pVTZ calculations are not accurate enough to reproduce quantitatively the energy differences between the radicals and their parent species, although qualitatively exhibits the same behavior than the other methods.

Table 6. Relative free energies of some of the radicals of formaldehyde and acetaldehyde with respect to their parent species using the DFT and CCSD(T) thermochemical corrections respectively. The shadowed cells correspond to the most stable radicals. Units are kJ mol^{-1}.

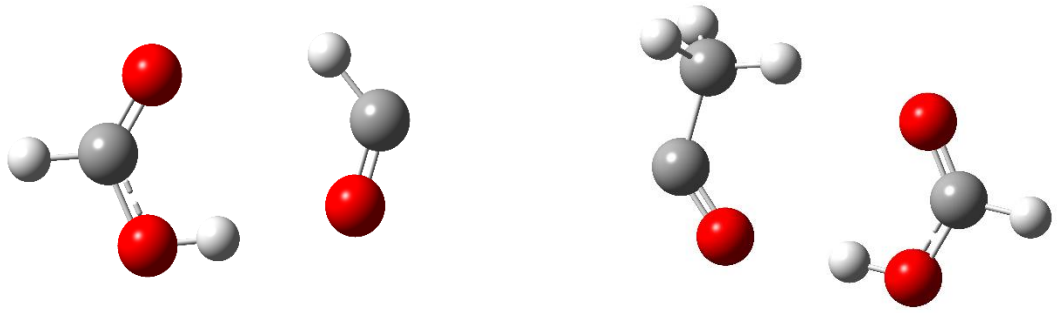
		Gas-phase		Bulk solvent	
Species	Form	B2PLYP	CCSD(T)	B2PLYP	CCSD(T)
F	OXO CH	-115.7	-114.2	-137.9	-136.4
	DIOL CH	-79.4	-79.5	-107.4	-107.5
	DIOL OH	-45.6	-45.8	-52.0	-53.1
A	OXO CH	-107.6	-111.7	-131.5	-135.0
	OXO CH3	-79.5	-81.1	-109.6	-110.7

In the case of the bulk solvent results, we were not able to calculate them at the SVECV-f12 level, because of the lack of solvent corrections to the F12 calculations in the programs we used. Nonetheless, given the

same accuracy delivered by SVECV-f12 and jChS calculations, the latter can be considered as the reference values for comparison of the other results. This information is collected in Table 5, and we see that in this case the DFT methods exhibit slightly larger average errors than in the gas phase, with the B2PLYP results still being the closer ones to jChS. However, in bulk solvent we obtain the somehow surprising result that the CCSD(T)/aug-cc-pVTZ calculations are much more accurate than in gas phase. Since neither the basis set or any of the other procedures have changed, we must conclude that the explanation should lie in the way that the electronic density of the radicals is treated. When the field of the solvent is introduced through the polarizable continuum approach, the unpaired electron is finding a better way to distribute its density by polarizing the charge elements on the shell surrounding the molecule, a core fact of the methodology. This result is basically the same that can be obtained by increasing the basis set, a conclusion that reinforces our previous discussion of the results in gas phase. Finally, one can observe that qualitatively the results in the bulk do not differ much from the results in gas phase. This statement, however, should be taken with care from the quantitative point of view, because, as previously pointed out, the PCM neglects the atomistic nature of hydrogen-bond interactions, which can be significant in ruling the reactivity mediated by water in the condensed phases [68]. It would be untrue to conclude from here that the solvent has no important impact on the calculations, since for that we should perform micro-solvated or even dynamical studies, both of which are work in progress in our groups.

3.5 Peculiarities

As commented before, we confronted difficulties with some of the species and methods because of spontaneous dissociation. This happened specifically for **Gl** and **MG** when we attempted the abstraction of a hydrogen from one of the hydroxyl groups in the diols. The resulting structures, which show an interaction between formic acid and either *F rad* or *A rad 1* are gathered in Figure 4.



Dissociacion from *Gldiol* Dissociation from *MGdiol*

Figure 4. Dissociated structures obtained from *Gldiol* and *MGdiol* when attempting hydrogen abstraction from one of the hydroxyl groups.

Some particularities must be mentioned. First, the presence of the second carbonyl seems to be indispensable to obtain this dissociation. It was not observed either for *Adiol*, *Gdiol*, or *ACdiol* hydrogen abstractions. In the second place, obtaining these structures was dependent on the species and method used. In the case of *Gldiol*, all methods except B2PLYP afforded the same type of non-dissociated structure, while B2PLYP, both in the gas phase and bulk solvent, resulted in the structure shown in Figure 4. Consequently, the CCSD(T) calculations could not be performed either. For *MGdiol*, instead, the initial structure at the exploratory CBS-QB3 level, as well as the B2PLYP structure like in *Gldiol*, ended up dissociated in the gas phase, while the other two DFT methods afforded non-dissociated structures. In bulk solvent, however, none of the methods afforded a non-dissociated structure, with all the optimizations ending in the structure shown in Figure 4. Since this was not the aim of this study, we didn't explore in detail what would be the result if other DFT or high-level post Hartree-Fock geometry optimizations were performed, but these results again point out the singularity of the α-oxo aldehydes in terms of hydration and hydrogen abstraction, a fact that might be a factor to be considered in the study of SOA in the atmosphere. It may be worthwhile to point out here to the study of Shannon et al. [62] where they explored the dissociation of *Gl rad* into *F rad* and CO

$$^{\bullet}OH + HC(O)C(O)H \rightarrow HC(O)C^{\bullet}O + H_2O \quad (1)$$

$$HC(O)C^{\bullet}O \rightarrow HCO + CO \quad (2)$$

This study shows that the water molecule in reaction (1) forms with far more energy than the HC(O)C$^{\bullet}$O co-fragment and that the energy partition between both co-fragments determine whether decomposition will take place like in reaction (2) or further reaction with an oxygen molecule will take place. In our case, the analogue situation is produced for *Gldiol rad2* which seems to be able to decompose into formic acid (instead of CO) and HC$^{\bullet}$O like in reaction (2). Further research is needed to find out whether these decompositions are an artificial issue of the methods employed or a real feature analogous to the phenomenon studied by Shannon et al. [62]

4. CONCLUSIONS

A study of the abstraction of all possible hydrogen atoms from both the aldehyde and geminal diol forms of six aldehydes (formaldehyde, acetaldehyde, glycolaldehyde, glyoxal, methylglyoxal and acrolein) was performed theoretically simulating conditions in the gas phase and water environments. The main purpose was to obtain reasonably accurate values of the stabilization energies of the radicals with respect to the parent species and to decide whether H-abstraction would be preferred to occur from the aldehyde or geminal diol forms.

For this purpose, we used three DFT methods, namely M06-2X-D3, PW6B95-D3, and B2PLYP, and one post-Hartree-Fock method, CCSD(T). Additionally, we applied composite quantum chemical methods, CBS-QB3 (applied only for exploratory purposes), jChS, and SVECV-f12, the two later ones developed in our groups, which main characteristic is to include correlation energy at advanced levels and to calculate the energy at the complete basis set level. The purposes of applying this panoply of methods were, on one side, to obtain precise and accurate values for the specific cases, and on the other, to evaluate the accuracy of the DFT methods so that they can be applied with confidence for larger molecules, for which the more accurate procedures are extremely difficult or impossible to apply.

The results showed different behaviors for the six species. When there is a conjugate group adjunct to the carbonyl group (**Gl**, **MG**, and **AC**) the abstraction of the hydrogen occurs preferentially from the geminal diol form. It is the opposite when there is no such conjugation (**F**, **A**, and **G**). If the energy necessary for hydration is considered, the conclusion is not modified for the latter, but for the former the conclusion is only valid for **Gl** and **MG**, not for **AC**. In this last case, the hydration energy needed is much larger and the full process, i.e. hydration + H-abstraction, results in a less stable product than the one obtained by direct H-abstraction from the aldehyde form. These facts then demonstrate that α-oxo aldehydes exhibit a distinct behavior than other aldehydes, notwithstanding whether the latter show conjugation (**AC**) or not (**F**, **A**, and **G**).

The theoretical results shown in this study agree mostly well with the experimental findings. According to the calculations performed by Elrod et al. [24], the $\log_{10}K_{hyd}$ values for **F**, **A**, **G**, **Gl** and **MG** are 3.4, 0.5, 1.6, 3.3 and 2.8 respectively, thus providing the order **F** > **Gl** > **MG** > **G** > **A** for the extension of hydration (no values are provided for acrolein). The experimental data summarized by Doussin and Monod [69] and Tilgner et al. [70] provided a recommended value of 2000 for the K_{hyd} of **F**, 1.2 as a recommendation for **A**, the value of 16.0 ± 1.3 for **G** as given by Zhao et al. [71] (compared to 10 obtained by Sorensen [72], 17.5 by Yaylayan et al. [73] and 15.7 Doussin and Monod [69]). MG exhibits two carbonyl groups, a ketonic C=O with K_{hyd} of 0.5 [74] and an aldehyde C=O with K_{hyd} of 1515 [70]. In the case of **Gl** the Khyd is 207 and 20000 for the second aldehyde group [69, 70]. The K_{hyd} for acrolein (**AC**) is not reported in the literature, but a value of methacrolein of 5×10^{-3} by Zhao et al. [71] can be assumed for the hydration of acrolein. These data provide the order for simple aldehydes as **F** > **G** > **A** > methacrolein, and for the α-oxo-aldehydes as **Gl** > **MG**. The theoretical results given in Table 2 provide the order **F** > **G** ≈ **A** > **AC** for simple aldehydes and **Gl** > **MG** for α-oxo-aldehydes.

It seems plausible that **F**, **Gl** and **MG** would be almost fully hydrated in the aqueous phase. However, this fact does not invalidate our previous discussion. Globally, four reactions will be competing at the interface between the gas and aqueous phases, namely:

$R–C(=O)H + H_2O \rightarrow R–C(OH)_2H$	Hydration	(3)
$R–C(=O)H + HO^{\bullet} \rightarrow R–C^{\bullet}(=O) + H_2O$	H-abstraction in gas phase	(4)
$R–C(OH)_2H + HO^{\bullet} \rightarrow R–C^{\bullet}(OH)_2 + H_2O$	H-abstraction in liquid phase	(5)
$R–C^{\bullet}(=O) + H_2O \rightarrow R–C^{\bullet}(OH)_2$	Hydration of gas-phase radical	(6)

Although we have not calculated reaction rate coefficients in this study, it is probable that reaction (3) is the slowest one in certain conditions (specifically if no ions are present). Then the two paths

$$R–C(=O)H + HO^{\bullet} \rightarrow R–C^{\bullet}(=O) + H_2O \xrightarrow{-[O_2]} R–C(=O)OO^{\bullet}\ \textit{(I1)} + H_2O \quad (A)$$

$$R–C(=O)H + H_2O \rightarrow R–C(OH)_2H \xrightarrow{-[HO^{\bullet}]} R–C^{\bullet}(OH)_2 + H_2O \xrightarrow{-[O_2]} R–COO^{\bullet}(OH)_2\ \textit{(I2)} + H_2O \quad (B)$$

will compete depending on the speed with which the aldehyde hydrates, and not on the extent of hydration. Thus, whether *(I1)* or *(I2)* are the intermediate radicals that will be formed depends on the specific characteristics of the system, an aspect that we have not investigated in the current study.

From the methodological point of view, it was shown that all methods afforded the same qualitative conclusion, although not so quantitatively. While the accurate composite methods, jChS and SVECV-f12, show very close results for all the species studied (with an average error of 1.9 kJ mol^{-1} and a similar spread of 1.6 kJ mol^{-1}, implying random errors), the DFT methods B2PLYP, M06-2X-D3, and PW6B95-D3 exhibit larger discrepancies, but still reasonably low on average and uniformly spread around the mean. In the case of CCSD(T) the situation is different. While in the gas phase the CCSD(T) results exhibit a large discrepancy with the average jChS and SVECV-f12 results, and the spread of these discrepancies around

their mean shows that there is some systematic error, the results in simulated solvent are much nearer to the best ones and, in fact, CCSD(T) behaves then as well as the DFT methods. We explored the possible reasons of this behavior and concluded that the polarization of charges on the shell around the molecules, which is the essence of the PCM method, allows a better polarization of the electronic density due to the unpaired electron. This characteristic results in the same stabilization that would have been produced by an extension of the basis set to the CBS limit and then brings the CCSD(T) results in line with the other values.

Finally, it was observed that both **Gl** and **MG** distinguish themselves from the rest because of the conflicting results that different theoretical methods produce when hydrogen abstraction is attempted from one of the hydroxyl groups in the gem diol species. Depending on the method and on whether the simulated solvent is introduced or not, either a dissociated species – a complex between formic acid and *F rad* and *A rad1* respectively – or a non-dissociated species were found. It was not studied whether these discrepancies reflect a real chemical fact or are just artifacts of the methods of calculation.

Extension of this work, both in the form of the study of transition states for kinetic studies, and in the form of a better representation of the solvent, by micro-solvation with water molecules and dynamic studies, are presently under way in our groups and will be reported in future studies.

5. ACKNOWLEDGMENTS

The CCBG participants thank the ANII, CSIC, and Pedeciba financial Uruguayan organizations for their continuous support of our research. ZS and NT thank the STARK group for the high-performance computing facilities and Scuola Normale Superiore for financial support through the project "Computational Modeling for Environmental Chemistry and Sustainability: from atmospheric monitoring to photocatalysis."

6. BIBLIOGRAPHY

1. Graedel TE, Weschler CJ (1981) Chemistry within aqueous atmospheric aerosols and raindrops. Rev Geophys Space Phys 19(4):505–539. https://doi.org/10.1029/RG019i004p00505

2. Ervens B, Turpin BJ, Weber RJ (2011) Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field, and model studies. Atm Chem Phys 11(21):11069–11102. https://doi.org/10.5194/acp-11-11069-2011
3. Herrmann H, Schaefer T, Tilgner A, Styler SA, Weller C, Teich M, Otto T (2015) Tropospheric Aqueous-Phase Chemistry: Kinetics, Mechanisms, and Its Coupling to a Changing Gas Phase. Chem Rev 115(10):4259−4334. https://doi.org/10.1021/cr500447k
4. McNeill VF (2015) Aqueous Organic Chemistry in the Atmosphere: Sources and Chemical Processing of Organic Aerosols. Environ Sci Technol 49(3):1237–1234. https://doi.org/10.1021/es5043707
5. Sun W, Zhang G, Guo Z, Fu Y, Peng X, Yang Y, Hu X, Lin J, Jiang F, Jiang B, Liao Y, Chen D, Chen J, Ou J, Wang X, Peng P, Bi X (2024) Formation of In-Cloud Aqueous-Phase Secondary Organic Matter and Related Characteristic Molecules. J Geophys Res: Atmos 129: e2023JD040355. https://doi.org/10.1029/2023JD040355
6. IPCC, 2007: Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A.(eds.)]. IPCC, Geneva, Switzerland, 104 pp. https://www.ipcc.ch/site/assets/uploads/2018/02/ar4_syr_full_report.pdf
7. Pennington EA, Seltzer KM, Murphy BN, Qin M, Seinfeld JH, Pye HOT (2021) Modeling secondary organic aerosol formation from volatile chemical products. Atmos Chem Phys 21:18247–18261. https://doi.org/10.5194/acp-21-18247-2021
8. Wen L, Schaefer T, He L, Zhang Y, Sun X, Ventura ON, Herrmann H (2021) T- and pH-Dependent Kinetics of the Reactions of ·OH(aq) with Glutaric and Adipic Acid for Atmospheric Aqueous-Phase Chemistry. ACS Earth Space Chem 5(8):1854–1864. https://doi.org/10.1021/acsearthspacechem.1c00163

9. Ervens B, Volkamer R (2010) Glyoxal processing by aerosol multiphase chemistry: towards a kinetic modeling framework of secondary organic aerosol formation in aqueous particles. Atm Chem Phys 10(17):8219-8244, https://doi.org/10.5194/acp-10-8219-2010
10. Zhang Y, He L, Sun X, Ventura ON, Herrmann H (2022) Theoretical Investigation on the Oligomerization of Methylglyoxal and Glyoxal in Aqueous Atmospheric Aerosol Particles, ACS Earth Space Chem 6(4):1031–1043. https://doi.org/10.1021/acsearthspacechem.1c00422
11. Mekic M, Schaefer T, Hoffmann EH, Aiyuk MBE, Tigner A, Herrmann H (2023) Temperature-Dependent Oxidation of Hydroxylated Aldehydes by $^{\bullet}OH$, $SO_4^{\bullet-}$, and $NO_3^{\bullet}$ Radicals in the Atmospheric Aqueous Phase. J Phys Chem A 127(31):6495–6508. https://doi.org/10.1021/acs.jpca.3c00700
12. Altieri KE, Seitzinger SR, Carlton AG, Turpin BJ, Klein GC, Marshall AG (2008) Oligomers formed through in-cloud methylglyoxal reactions: Chemical composition, properties, and mechanisms investigated by ultra-high resolution FT-ICR mass spectrometry. Atmos Environ 42(7):1476–1490. https://doi.org/10.1016/j.atmosenv.2007.11.015
13. Carlton AG, Turpin BJ, Altieri KE, Seitzinger S, Reff A. Lim HJ, Ervens B (2007) Atmospheric oxalic acid and SOA production from glyoxal: Results of aqueous photooxidation experiments. Atmos Environ 41(35):7588–7602. https://doi.org/10.1016/j.atmosenv.2007.05.035
14. Warneck P (2003) In-cloud chemistry opens pathway to the formation of oxalic acid in the marine atmosphere. Atmos Environ 37(17):2423–2427. https://doi.org/10.1016/S1352-2310(03)00136-5
15. Wei B, Zhang R, Sit PH, He M, Chan CK (2023) Theoretical study on the aqueous phase oxidation of Glyoxal. Environ Sci: Atmos 3(9):1296–1305. https://doi.org/10.1039/D3EA00049D
16. Toumi I, Dalbouha S, Al-Mogren MM, Yazidi O, Jaïdane NE, Carvajal M, Senent ML (2022) Theoretical Spectroscopic Study of Two Ketones of Atmospheric Interest: Methyl Glyoxal (CH_3COCHO) and Methyl Vinyl Ketone ($CH_3COCH{=}CH_2$). J Phys Chem A 126(40):7230–7241, https://doi.org/10.1021/acs.jpca.2c05653

17. Axson JL, Takahashi K, De Haan DO, Vaida V (2010) Gas-phase water-mediated equilibrium between methylglyoxal and its geminal diol. Proc Nat Acad Sci USA 107(15):6687–6692. https://doi.org/10.1073/pnas.091212110
18. Krizner HE, De Haan DO, Kua J (2009) Thermodynamics and kinetics of methylglyoxal dimer formation: A computational study. J Phys Chem A 113(25):6994–7001. https://doi.org/10.1021/jp903213k
19. Betterton EA, Hoffmann MR (1987) Kinetics, mechanism, and thermodynamics of the reversible reaction of methylglyoxal (CH_3COCHO) with S(IV). J Phys Chem 91(11):3011–3020. https://doi.org/10.1021/j100295a074
20. Nemet I, Vikic-Topic D, Varga-Defterdarovic L (2004) Spectroscopic studies of methylglyoxal in water and dimethylsulfoxide. Bioorg Chem 32(6):560–570. https://doi.org/10.1016/j.bioorg.2004.05.008
21. Loeffler KW, Koehler CA, Paul NM, De Haan DO (2006) Oligomer formation in evaporating aqueous glyoxal and methyl glyoxal solutions. Environ Sci Technol 40(20):6318–6323. https://doi.org/10.1021/es060810w
22. Asmus KD, Moeckel H, Henglein A (1973) Pulse radiolytic study of the site of hydroxyl radical attack on aliphatic alcohols in aqueous solution. J Phys Chem 77(10):1218–1221. https://doi.org/10.1021/j100629a007
23. Carey F, Giuliano R (2013) Organic Chemistry. McGraw Hill Education: New York.
24. Elrod MJ, Sedlak JA, Ren H (2021) Accurate Computational Model for the Hydration Extent of Atmospherically Relevant Carbonyls on Aqueous Atmospheric Particles, ACS Earth Space Chem 5(2):348−355. https://doi.org/10.1021/acsearthspacechem.0c00322
25. Bursch M, Mewes JM, Hansen A, Grimme S (2022) Best-Practice DFT Protocols for Basic Molecular Computational Chemistry. Angew Chem 61:e202205735. https://doi.org/10.1002/anie.202205735

26. Teale AM, Helgaker T, Savin A, Adamo C, Aradi B, Arbuznikov AV, Ayers PW, Baerends EJ, Barone V, Calaminici P, Cancès E, Carter EA, Chattaraj PK, Chermette H, Ciofini I, Crawford TD, De Proft F, Dobson JF, Draxl C, Frauenheim T, Fromager E, Fuentealba P, Gagliardi L, Galli, G, Gao J, Geerlings P, Gidopoulos N, Gill PMW, Gori-Giorgi P, Görling A, Gould T, Grimme S, Gritsenko O, Jensen HJA, Johnson ER, Jones RO, Kaupp M, Köster AM, Kronik L, Krylov AI, Kvaal S, Laestadius A, Lewin M, Liu S, Loos PF, Maitra NT, Neese F, Perdew JP, Pernal K, Pernot P, Piecuch P, Rebolini E, Reining L, Romaniello P, Ruzsinszky A, Salahub DR, Scheffler M, Schwerdtfeger P, Staroverov VN, Sun J, Tellgren E, Tozer DJ, Trickey SB, Ullrich CA, Vela A, Vignale G, Wesolowski TA, Xu X, Yang W (2022) DFT exchange: sharing perspectives on the workhorse of quantum chemistry and materials science. Phys. Chem. Chem. Phys. 24, 28700-28781, https://doi.org/10.1039/D2CP02827A
27. Morgante P, Peverati R (2020) The devil in the details: A tutorial review on some undervalued aspects of density functional theory calculations. Int J Quantum Chem 18:e26332, https://doi.org/10.1002/qua.26332
28. Verma P, Truhlar DG (2020) Status and Challenges of Density Functional Theory. Trends Chem. 2(4):302–318. https://doi.org/10.1016/j.trechm.2020.02.005
29. Bartlett RJ, Purvis GD (1978) Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem. Int J Quantum Chem 14(5):561–568. https://doi.org/10.1002/qua.560140504
30. Helgaker T, Klopper W, Tew DP (2008) Quantitative quantum chemistry. Mol Phys 106(16-18):2107-2143. https://doi.org/10.1080/00268970802258591
31. Karton A (2016) A computational chemist's guide to accurate thermochemistry for organic molecules. Wires Comp Mol Sci 6(3):292–310. https://doi.org/10.1002/wcms.1249
32. Karton A (2022) Quantum mechanical thermochemical predictions 100 years after the Schrödinger equation. Ann Rep Comp Chem 18:123–166. https://doi.org/10.1016/bs.arcc.2022.09.003

33. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc 120:215–241. https://doi.org/10.1007/s00214-007-0310-x
34. Zhao Y, Truhlar DG (2005) Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. J Phys Chem A 109(25):5656–5667. https://doi.org/10.1021/jp050536c
35. Grimme S (2006) Semiempirical hybrid density functional with perturbative second-order correlation. J Chem Phys 124(3):034108. https://doi.org/10.1063/1.2148954
36. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132(15):154104. https://doi.org/10.1063/1.3382344
37. Kendall RA, Dunning Jr TH, Harrison RJ (1992) Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J Chem Phys 96:6796–6806. https://doi.org/10.1063/1.462569
38. Papajak E, Zheng J, Xu X, Leverentz HR, Truhlar DG (2011) Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. J Chem Theory Comput 7(10):3027–3034. https://doi.org/10.1021/ct200106a
39. Montgomery Jr JA, Frisch MJ, Ochterski JW, Petersson GA (1999) A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. J Chem Phys 110(6):2822–2827. https://doi.org/10.1063/1.477924
40. Montgomery Jr JA, Frisch MJ, Ochterski JW, Petersson GA (2000) A complete basis set model chemistry. VII. Use of the minimum population localization method. J Chem Phys 112(15):6532–6542. https://doi.org/10.1063/1.481224
41. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98(7):5648–5652. https://doi.org/10.1063/1.464913

42. Krishnan R, Pople JA (1978) Approximate fourth-order perturbation theory of the electron correlation energy. Int J Quantum Chem 14(1):91–100. https://doi.org/10.1002/qua.560140109
43. Møller C, Plesset MS (1934) Phys Rev 46(7):618–622. https://doi.org/10.1103/PhysRev.46.618
44. Ventura ON, Kieninger M, Katz A, Vega-Teijido M, Segovia M, Irving K (2021) SVECV-f12: Benchmark of a composite scheme for accurate and cost-effective evaluation of reaction barriers. Int J Quantum Chem 121(18):e26745. https://doi.org/10.1002/qua.26745
45. Kieninger M, Ventura ON (2023) SVECV-f12: A composite scheme for accurate and cost-effective evaluation of reaction barriers. II. Benchmarking using Karton's BH28 barrier heights database. Int J Quantum Chem 123(24):e27069. https://doi.org/10.1002/qua.27069
46. Vega-Teijido M, Segovia M, Kieninger M, Ventura ON (2024) Accuracy of enthalpies of formation of hydrocarbons using the SVECV-f12 protocol and comparison to other composite methods. J Chem Therm 189:107197. https://doi.org/10.1016/j.jct.2023.107197
47. Alessandrini S, Barone V, Puzzarini C (2020) Extension of the “Cheap” Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. J Chem Theory Comput 16(2):988–1006. https://doi.org/10.1021/acs.jctc.9b01037
48. Lupi J, Alessandrini S, Puzzarini C, Barone V (2021) junChS and junChS-F12 Models: Parameter-free Efficient yet Accurate Composite Schemes for Energies and Structures of Noncovalent Complexes. J Chem Theory Comput 17(11):6974–6992. https://doi.org/10.1021/acs.jctc.1c00869
49. Barone V, Lupi J, Salta Z, Tasinato N (2021) Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates. J Chem Theory Comput 17(8):4913–4928. https://doi.org/10.1021/acs.jctc.1c00406
50. Helgaker T, Klopper W, Koch H, Noga J (1997) Basis-set convergence of correlated calculations on water. J Chem Phys 106(23):9639–9646. https://doi.org/10.1063/1.473863
51. Knizia G, Adler TB, Werner HJ (2009) Simplified CCSD(T)-F12 methods: Theory and benchmarks. Chem Phys 130:054104. https://doi.org/10.1063/1.3054300

52. Pritchard BP, Altarawy D, Didier B, Gibson TD, Windus TL (2019) A New Basis Set Exchange: An Open, Up-to-date Resource for the Molecular Sciences Community. J Chem Inf Model 59(11):4814–4820. https://doi.org/10.1021/acs.jcim.9b00725.

53. Cossi M, Barone V, Cammi R, Tomasi J (1996) Ab initio study of solvated molecules: a new implementation of the polarizable continuum model. Chem Phys Lett 255(4-6), 327–335. https://doi.org/10.1016/0009-2614(96)00349-1

54. Gaussian 16, Revision C.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian, Inc., Wallingford CT.

55. Werner HJ, Knowles PJ, Knizia G, Manby FR, Schütz M (2012) Molpro: a general-purpose quantum chemistry program package. WIREs Comput Mol Sci 2(2):242–253. https://doi.org/10.1002/wcms.82

56. Werner HJ, Knowles PJ, Manby FR, Black JA, Doll K, Heßelmann A, Kats D, Köhn A, Korona T, Kreplin DA, Ma Q, Miller TF III, Mitrushchenkov A, Peterson KA, Polyak I, Rauhut G, Sibaev M (2020) The Molpro quantum chemistry package. J Chem Phys 152(14):144107. https://doi.org/10.1063/5.0005081

57. MOLPRO, version 19.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, and others, see https://www.molpro.net. 2020.

58. Fu TM, Jacob DJ, Wittrock F, Burrows, JP, Vrekoussis M, Henze DK (2008) Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols. J Geophys Res: Atmos 113(D15):D15303. https://doi.org/10.1029/2007JD009505

59. Jang MS, Kamens RM (2001) Characterization of secondary aerosol from the photooxidation of toluene in the presence of NOx and 1-propene. Environ Sci Technol 35(18):3626–3639. https://doi.org/10.1021/es010676+

60. Tuazon EC, Macleod H, Atkinson R, Carter WPL (1986) Alpha-dicarbonyl yields from the NOX-air photooxidations of a series of aromatic-hydrocarbons in air. Environ Sci Technol 20(4):383–387. https://doi.org/10.1021/es00146a010

61. Shannon R, Blitz MA, Seakins PW (2024) Solving the OH + Glyoxal Problem: A Complete Theoretical Description of Post-Transition-State Energy Deposition in Activated Systems. J Phys Chem A 128(8):1501–1510. https://doi.org/10.1021/acs.jpca.3c07823

62. Pople JA (1999) Nobel Lecture: Quantum chemical models. Rev Mod Phys 71(5):1267–1274. https://doi.org/10.1002/(SICI)1521-3773(19990712)38:13/14<1894::AID-ANIE1894>3.0.CO;2-H

63. Salta Z, Ventura ON, Rais N, Tasinato N, Barone V (2024) A new chapter in the never-ending story of cycloadditions: the puzzling case of SO_2 and acetylene. J. Comput. Chem., in press. https://doi.org/10.1002/jcc.27350

64. Boussessi R, Ceselin G, Tasinato N, Barone V (2020) DFT meets the segmented polarization consistent basis sets: Performances in the computation of molecular structures, rotational and vibrational spectroscopic properties. J Mol. Struct 1208:127886. https://doi.org/10.1016/j.molstruc.2020.127886

65. Barone V, Ceselin G, Fusè M, Tasinato N (2020) Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. Front Chem 8. https://doi.org/10.3389/fchem.2020.584203

66. Tasinato N, Grimme S (2015) Unveiling the non-covalent interactions of molecular homodimers by dispersion-corrected DFT calculations and collision-induced broadening of ro-vibrational transitions: application to $(CH_2F_2)_2$ and $(SO_2)_2$. Phys Chem Chem Phys 17:5659–5669. https://doi.org/10.1039/C4CP05680A

67. Baiano C, Lupi J, Tasinato N, Puzzarini C, Barone V (2020) The Role of State-of-the-Art Quantum-Chemical Calculations in Astrochemistry: Formation Route and Spectroscopy of Ethanimine as a Paradigmatic Case. Molecules 25(12):2873. https://doi.org/10.3390/molecules25122873

68. Baiano C, Lupi J, Barone V, Tasinato N (2022) Gliding on ice in search of accurate and cost-effective computational methods for astrochemistry on grains: The puzzling case of the HCN isomerization. 18:3111–3121. https://doi.org/10.1021/acs.jctc.1c01252

69. Doussin JF, Monod A (2013) Structure–activity relationship for the estimation of OH-oxidation rate constants of carbonyl compounds in the aqueous phase. Atmos Chem Phys 13:11625–11641. https://doi.org/10.5194/acp-13-11625-2013

70. Tilgner A, Schaefer T, Alexander B, Barth M, Collett Jr JL, Fahey KM, Nenes A, Pye HOT, Herrmann H, McNeill VF (2021) Atm Chem Phys 21(17):13483–13536. https://doi.org/10.5194/acp-21-13483-2021

71. Zhao R, Lee AKY, Soong R, Simpson AJ, Abbatt JPD (2013) Formation of aqueous-phase α-hydroxyhydroperoxides (α-HHP): potential atmospheric impacts. Atmos Chem Phys 13:5857–5872. https://doi.org/10.5194/acp-13-5857-2013

72. Sorensen P (1972) The reversible addition of water to glycolaldehyde in aqueous solution. Acta Chem Scand 26:3357–3365. https://doi.org/10.3891/acta.chem.scand.26-3357

73. Yaylayan V, Harty-Majors S, Ismail A (1998) Investigation of the mechanism of dissociation of glycolaldehyde dimer (2,5-dihydroxy-1,4-dioxane) by FTIR spectroscopy. Carbohydr Res 309(1):31–38. https://doi.org/10.1016/S0008-6215(98)00129-3

74. Creighton DJ, Migliorini M, Pourmotabbed T, Guha MK (1988) Optimization of efficiency in the glyoxalase pathway. Biochem 27:7376-7384. https://doi.org/10.1021/bi00419a031

Statements and Declarations

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript other than the general funds mentioned in the Acknowledgments section.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

All authors contributed to the study conception and design. Calculations and data collection were performed by Zoi Salta, Aline Katz, Martina Kieninger and Oscar N. Ventura. Analysis was performed by Thomas Schaefer, Hartmut Herrmann, Nicola Tasinato and Oscar N. Ventura. The first draft of the manuscript was written by Oscar N. Ventura and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability

The datasets generated during and/or analyzed during the current study are provided in the Supplementary Information section. Other data pertaining to the study are available from the corresponding author on reasonable request.