

# Thermochemistry of Radicals and Molecules Relevant to Atmospheric Chemistry: Determination of Group Additivity Values using G3//B3LYP Theory

Shumaila S. Khan, Xinrui Yu, Jeffrey R. Wade, R. Dean Malmgren, and Linda J. Broadbelt\*

Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208-3120

Received: October 22, 2008; Revised Manuscript Received: February 24, 2009

Thermochemistry of radicals is not as extensively tabulated as that of stable molecular species, even when group additivity schemes are applied. When these radicals contain oxygen or nitrogen atoms, the availability of radical groups is even more limited. Many oxygen- and nitrogen-containing radicals and molecules are present in the atmosphere, and thermochemistry is a valuable component of the development of atmospheric models with predictive capabilities. This paper presents quantum chemical calculations using G3//B3LYP that have been performed to obtain heats of formation, entropies, and heat capacities as a function of the temperature of radicals and molecules from which group additivity values were obtained. Isodesmic and homodesmotic reactions were used to obtain improved estimates of the heats of formation. Thermodynamic property estimates were corrected to account for internal rotations. A total of 323 molecules were studied from which a total of 122 different groups, 21 gauche and cis corrections, and 5 secondary corrections were regressed.

## 1. Introduction

Ozone formation in the troposphere has been a stubborn and persistent problem in urban areas. Interactions of volatile organic compounds (VOCs) and nitrogen oxides ( $\text{NO}_x$ ) in the atmosphere lead to many radical and molecular species, and the reactivity of many of these species is unknown. To develop accurate atmospheric models that include quantitative estimates of the reactivity of many different species, rate coefficients must be specified. Structure/reactivity relationships are used to fill gaps in experimental databases for rate coefficients, and these correlations often relate thermodynamic properties to rate coefficients. Thus, thermochemical properties for all species must be available.

Group additivity has been traditionally used to obtain thermodynamic property estimates for species that have no experimental thermochemical data available.<sup>1</sup> The approach involves decomposing a molecule into its constituent groups and then summing the contributions of the group to a particular property to obtain the value for the molecule. This technique has been shown to be very accurate for heats of formation of organic compounds, oxygen-containing organic compounds,<sup>2–4</sup> and organic radicals.<sup>5,6</sup>

Recently, high-level quantum chemical calculations have found increased use because of their high accuracy in predicting thermodynamic properties.<sup>3,7–18</sup> However, it is often prohibitive to carry out an *ab initio* quantum chemical calculation for every molecule in a large molecule set because of the high computational cost. An alternative is to perform a small number of calculations for a select set of molecules and generalize these values using a group additivity scheme. This approach has been used recently by numerous researchers, including Lay and Bozzelli,<sup>3</sup> Marsi et al.,<sup>5</sup> Sebban et al.,<sup>19</sup> Sumathi and Green,<sup>10,20</sup> Swihart and Girshick,<sup>8</sup> Wong et al.,<sup>9</sup> and Sabbe et al.,<sup>11,21</sup> who have filled in missing group values for a diverse range of compounds including alkyl peroxides and trioxides, alkyl

radicals, large polyunsaturated hydrocarbons, ketenes, and silicon–hydrogen compounds.

Most atmospheric models have used lumped structure models that do not require properties of specific molecules as part of the network generation process.<sup>22–24</sup> However, explicit automated mechanism generation has proven valuable for a wide range of different chemistries,<sup>25–27</sup> and we are currently undertaking the effort to develop mechanistic models of atmospheric chemistry. Preliminary analysis of atmospheric mechanisms has revealed that the number of species that can be generated is on the order of  $10^5$  for even simple systems such as ethane–air– $\text{NO}_x$ . Other research has also shown that atmospheric models of octane can have as many as 2.5 million species.<sup>28</sup> Thus, while high-level quantum chemical calculations are becoming more tractable, calculation of properties for this number of molecules is still prohibitive, and an approach such as group additivity is still extremely valuable.

In this work, quantum chemical calculations were conducted using Gaussian-3 (G3), a high-level composite method,<sup>29</sup> with geometries based on density functional theory (DFT) using the B3LYP functional. Although G3 has been traditionally applied using MP2 for geometry optimization,<sup>30</sup> a review article by Henry and Radom recommends using DFT to obtain geometries for radicals because of its lower sensitivity to spin contamination.<sup>29</sup> On the basis of the thermodynamic properties obtained, group additivity values for 122 groups, 21 gauche and cis corrections, and 5 secondary corrections were regressed.

## 2. Computational Methods

Quantum chemical calculations were performed using Gaussian-3 theory with geometries from B3LYP with the 6-31G(d) basis set (G3//B3LYP)<sup>31</sup> (a variation of G3<sup>30</sup> using B3LYP density functional theory to perform geometry optimizations rather than second-order Møller–Plesset perturbation theory (MP2)) using the Gaussian 98 software package.<sup>32</sup> Geometry optimization to obtain a minimum-energy structure was carried out, followed by a frequency calculation to obtain thermody-

\* To whom correspondence should be addressed. Phone: (847) 491-5351. E-mail: broadbelt@northwestern.edu.

dynamic data. All G3//B3LYP calculations were tested for wave function stability. In the G3//B3LYP method, vibrational frequencies are scaled by 0.96 to obtain the energy reported. The DFT methods use scaling factors ranging from 0.96 to 0.98, and during the development of the G3//B3LYP method, a factor of 0.96 was chosen by Baboul et al.<sup>31</sup> Thus, all properties reported here, including standard entropy and heat capacities, were estimated from the vibrational frequencies scaled by the same scaling factor of 0.96.

In the calculation of thermodynamic properties, low frequencies were treated using the internal rotor (IR) model instead of the conventional harmonic oscillator (HO) model. Such treatment has been applied widely in the calculation of thermodynamic properties and kinetic parameters.<sup>4,33–41</sup> The conventional form of the partition function is shown in eq 1 as the product of electronic ( $Q_e$ ), translational ( $Q_{tr}$ ), external rotational ( $Q_r$ ), and vibrational ( $Q_v$ ) partition functions.

$$Q = Q_e Q_{tr} Q_r Q_v \quad (1)$$

In this research, the low frequencies which are mainly composed of rotational motions were separated from the vibrational partition function ( $Q_v$ ) and treated separately using the internal rotor model ( $Q_{int,rot}$ ). The remaining vibrational motions were still treated using the harmonic oscillator model ( $Q_{vib}$ ). The total partition function is revised as shown in eq 2.

$$Q = Q_e Q_{tr} Q_r Q_{vib} Q_{int,rot} \quad (2)$$

Because each low frequency may have rotational components from multiple rotation tops, it is infeasible to use the rotational partition function based on a single rotation top to substitute the harmonic oscillator contribution from a low frequency in a one-to-one correspondence. In this research, we calculated the partition function for each rotation based on the rotational energy barrier as shown in eq 3 and then substituted the product of the partition functions of low frequencies, the number of which equals the number of rotational tops.

$$Q_{ir,m} = \frac{1}{\sigma_m} \sum_i \exp\left(-\frac{\epsilon_i}{k_B T}\right) \quad (3)$$

In eq 3,  $\sigma_m$  is the symmetry number of the  $m$ th internal rotation,  $k_B$  is Boltzmann's constant ( $1.3806 \times 10^{-23} \text{ J K}^{-1}$ ), and  $\epsilon_i$  are the energy levels of the internal rotation, which were calculated by solving a one-dimensional Schrödinger equation using the Fourier Grid Hamiltonian (FGH) method<sup>42</sup>

$$-\frac{\hbar^2}{2I_{red}} \frac{d^2}{d\theta^2} \Psi + V(\theta)\Psi = \epsilon\Psi \quad (4)$$

where  $\hbar$  is Planck's constant divided by  $2\pi$ ,  $\Psi$  is the wave function,  $\theta$  is the torsional angle, and  $I_{red}$  is the reduced moment of inertia, which was defined as  $I^{2,3}$  in accordance with the systematic classification of moments of inertia by East et al.<sup>43</sup>  $V(\theta)$  was determined from relaxed potential energy scans calculated using unrestricted B3LYP/6-31G(d) using the keyword "modredundant" and intervals of  $30^\circ$ ; thus, a total 12 optimized conformations from  $0$  to  $360^\circ$  were obtained for each rotor. These one-dimensional scans were also used to ensure that the optimized geometry used in the G3//B3LYP energy

evaluation was the lowest-energy conformer. Although one-dimensional torsional scans do not guarantee that the global minimum structure is identified, they did often identify lower-energy conformers. The one-dimensional scans were then repeated until no structures of lower energy were obtained.  $V(\theta)$  was expanded using a full Fourier series as in eq 5

$$V(\theta) = \sum_{i=1}^n [a_i(1 - \cos i\theta) + b_i \sin i\theta] \quad (5)$$

where  $a_i$  and  $b_i$  are the coefficients of the expansion. Equation 5 can be written in its matrix product form, and the coefficients  $a_i$  and  $b_i$  can be determined by solving an overdetermined matrix. In order to ensure that this matrix was overdetermined, the number of coefficients was less than the number of energy points during the rotation. For a scan interval of  $30^\circ$ ,  $n$  was set equal to 3. This was demonstrated to be a sufficient number of coefficients to ensure acceptable fits for all potential energy scans.<sup>44–47</sup> The calculation of the reduced moment of inertia, solution of the Fourier coefficients and Schrödinger equation, calculation of the partition functions, and thermodynamic data correction were carried out with "calctherm" developed in our group.<sup>46,48</sup>

In some cases, Gaussian 98 did not identify the external symmetry number correctly that was used in the statistical mechanical calculation of the entropy, and therefore, these values were corrected as in eq 6

$$S^{298} = S_{\text{gaussian}}^{298} + R \ln(\sigma_{\text{ext}}^{\text{gaussian}}) - R \ln(\sigma_{\text{ext}}^{\text{correct}}) \quad (6)$$

where  $S_{\text{gaussian}}^{298}$  is the entropy reported by Gaussian 98,  $R$  is the universal gas constant,  $\sigma_{\text{ext}}^{\text{gaussian}}$  is the external symmetry number used in Gaussian 98, and  $\sigma_{\text{ext}}^{\text{correct}}$  is the correct external symmetry number. Furthermore, calculation of entropy values using group additivity includes a correction of  $-R \ln \sigma$ , in which  $\sigma$  is defined in eq 7 as

$$\sigma = \sigma_{\text{ext}}^* \sigma_{\text{int}} / n_{\text{opt}} \quad (7)$$

where  $\sigma_{\text{ext}}$  is the external symmetry number,  $\sigma_{\text{int}}$  is the internal rotational symmetry number, and  $n_{\text{opt}}$  is the number of optical isomers. Therefore,  $R \ln \sigma$  was added to  $S^{298}$  in eq 6 before fitting the group additivity values.

Enthalpies can be calculated as the sum of the electronic energies ( $E_e$ ), zero-point energies (ZPE), and the thermal corrections ( $E_v^{298}$ ,  $E_{tr}^{298}$ , and  $E_r^{298}$ ) from  $0$  to  $298 \text{ K}$  as shown in eq 8 below

$$H^{298} = E_e + \text{ZPE} + E_v^{298} + E_{tr}^{298} + E_r^{298} + \Delta PV \quad (8)$$

All of the quantities on the right-hand side of eq 8 are obtained from the quantum chemical calculations.

Three different methods were used to obtain the standard enthalpy of formation of the species. The method was based on either atomization energies, isodesmic bond separation reactions, or isodesmic reactions. The values based on atomization energies and isodesmic bond separation reactions are provided for comparison purposes only; all group additivity values were calculated using enthalpies of formation based on the more accurate method of isodesmic reactions. The method

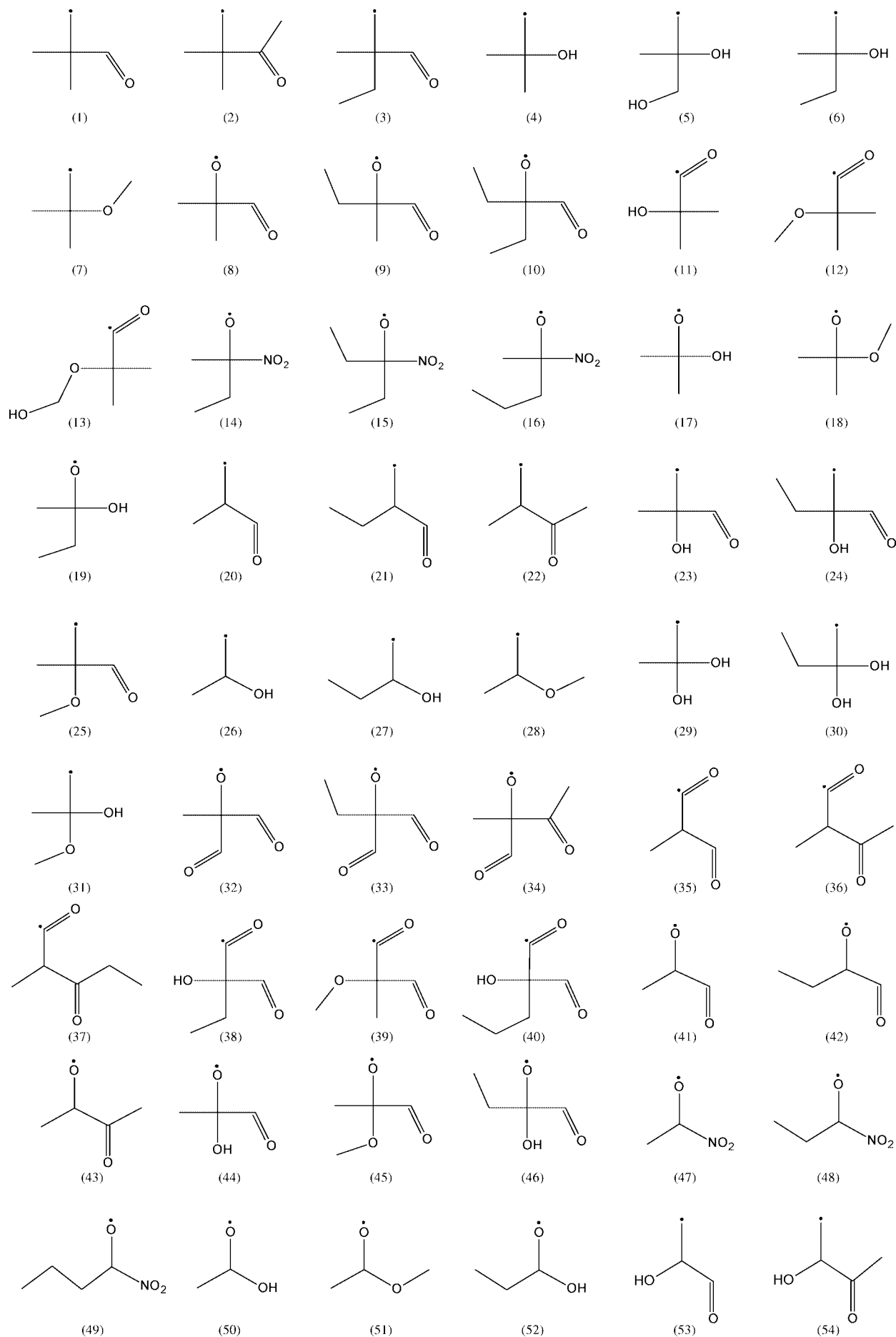


Figure 1. Part 1 of 6.

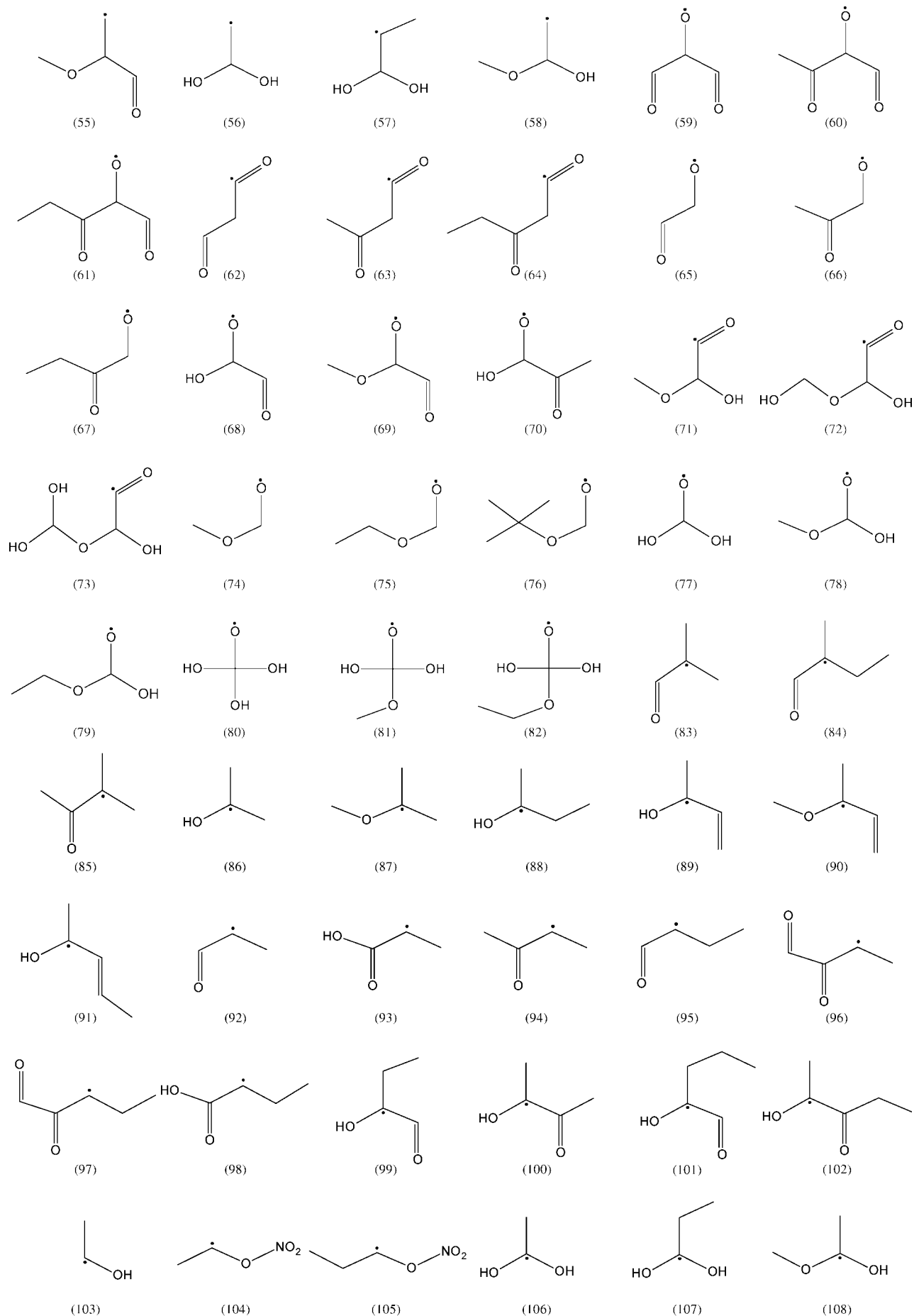


Figure 1. Part 2 of 6.

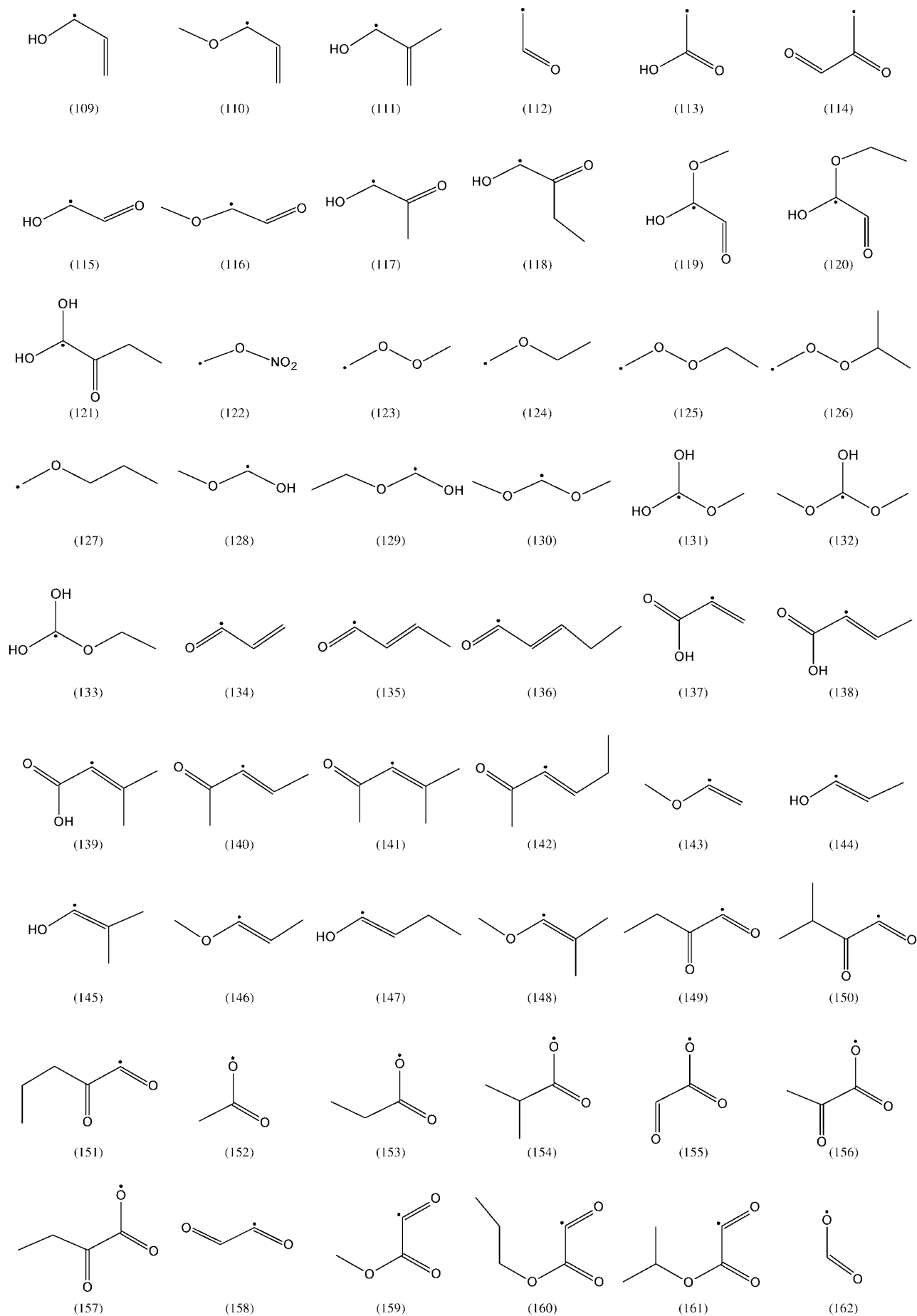


Figure 1. Part 3 of 6.

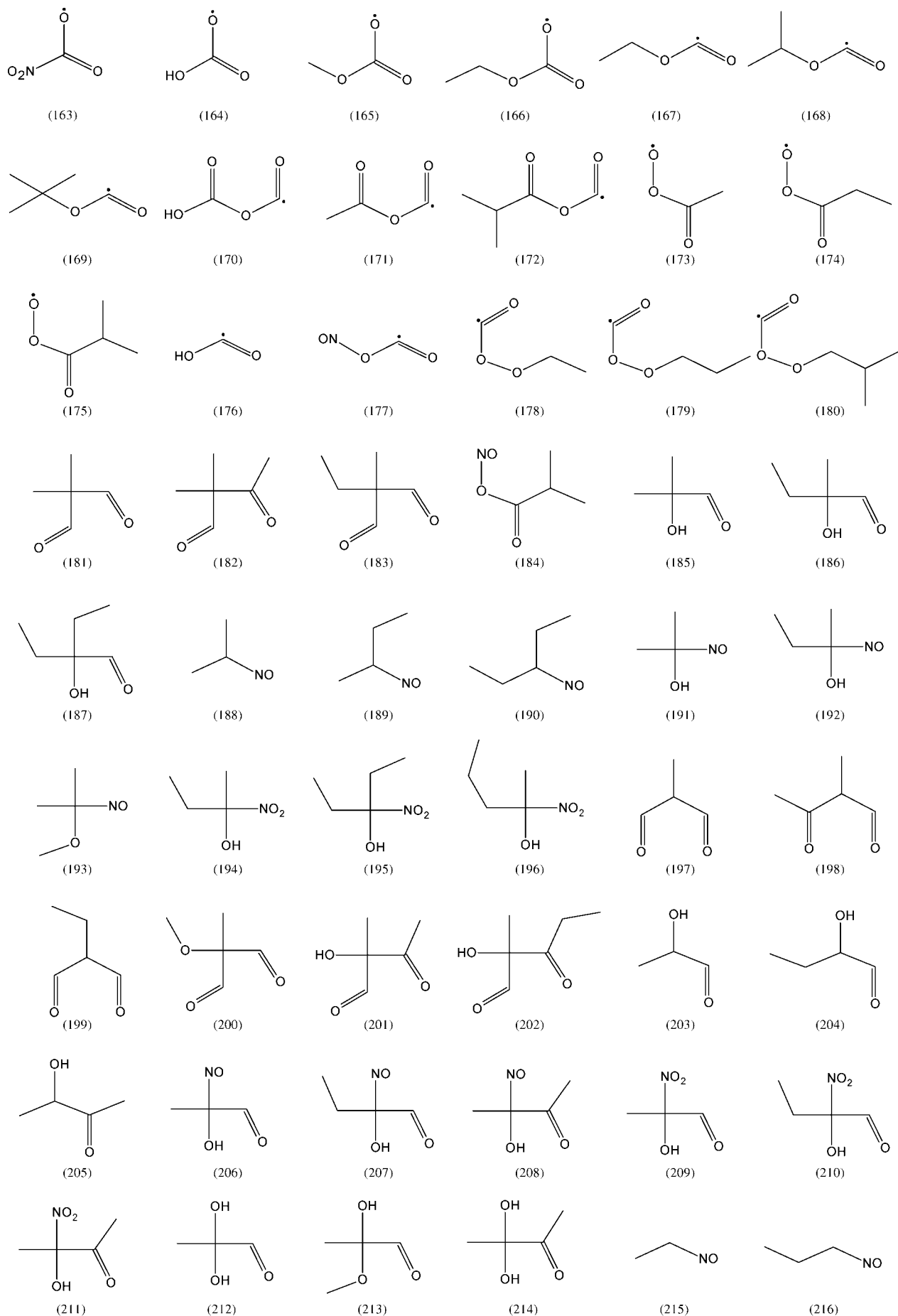


Figure 1. Part 4 of 6.

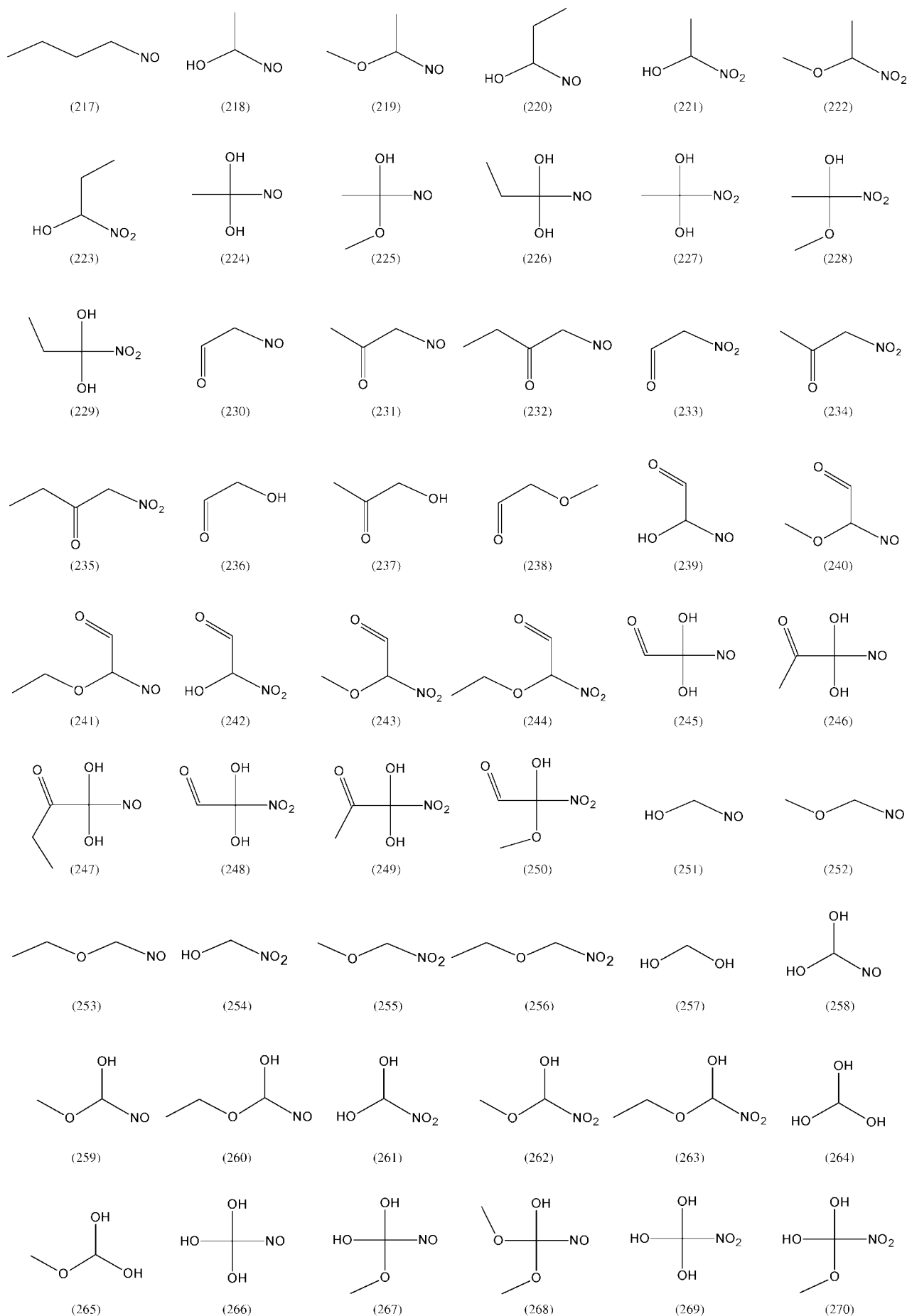
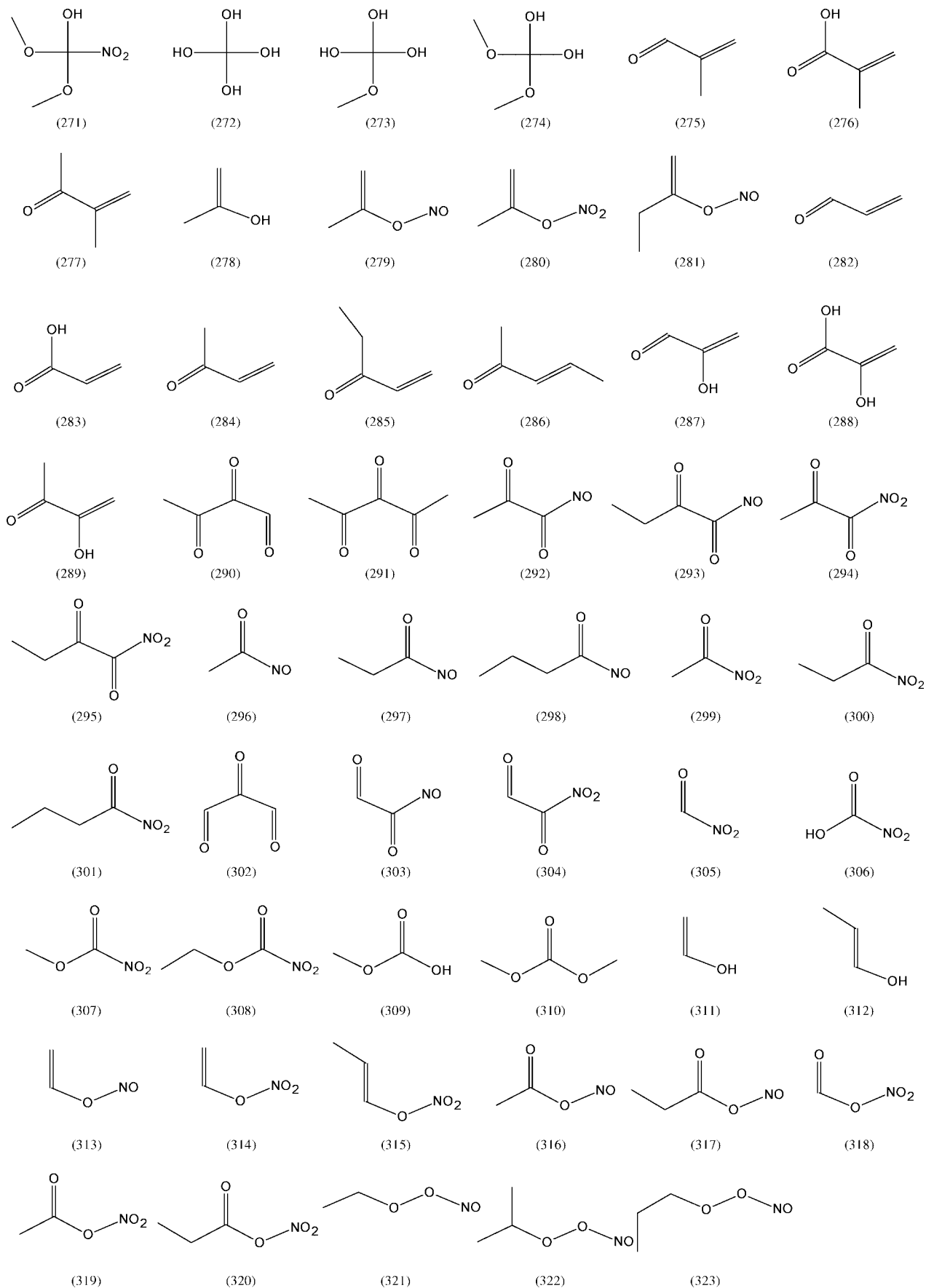


Figure 1. Part 5 of 6.



**Figure 1.** Part 6 of 6. Structures of molecules for which quantum chemical calculations were performed. The numbers associated with each molecule are used to identify molecules in the tables.



**TABLE 1: List of All Possible gauche and cis Next-Nearest-Neighbor and Secondary Interactions<sup>a</sup>**

interaction number	type of interaction	group	interaction number	type of interaction	group
Next-Nearest-Neighbor Corrections					
NN1	gauche	CH <sub>3</sub> -C-C-NO	NN42	gauche	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-Ċ(O)
NN2	gauche	CH <sub>3</sub> CH <sub>2</sub> -C-C-Ċ	NN43	gauche	Ċ-C(O)-C-CH <sub>3</sub>
NN3	gauche	CH <sub>3</sub> -C(O)-C-OH	NN44	gauche	Ċ(O)O-C(O)-C-CH <sub>3</sub>
NN4	gauche	CH <sub>3</sub> -C(O)-C-CH <sub>3</sub>	NN45	gauche	CH <sub>3</sub> -O-C-C(O)H
NN5	gauche	CH <sub>3</sub> -C(O)-C-ĊH <sub>2</sub>	NN46	gauche	CH <sub>3</sub> -O-C-ĊH <sub>2</sub>
NN6	gauche	CH <sub>3</sub> -C(O)-C-Ċ	NN47	gauche	CH <sub>3</sub> -O-C-Ċ(O)
NN7	gauche	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-OH	NN48	gauche	CH <sub>3</sub> -O-C-Ċ
NN8	gauche	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-Ċ	NN49	gauche	CH <sub>3</sub> CH <sub>2</sub> -O-C-Ċ
NN9	gauche	CH <sub>3</sub> -O-C-OH	NN50	gauche	HOCH <sub>2</sub> -O-C-Ċ(O)
NN10	gauche	CH <sub>3</sub> -O-C-OCH <sub>3</sub>	NN51	gauche	HOCH <sub>2</sub> -O-C-OH
NN11	gauche	CH <sub>3</sub> -O-C-NO <sub>2</sub>	NN52	gauche	(HO) <sub>2</sub> CH-O-C-OH
NN12	gauche	CH <sub>3</sub> CH <sub>2</sub> -O-C-OH	NN53	gauche	(HO) <sub>2</sub> CH-O-C-Ċ(O)
NN13	gauche	(CH <sub>3</sub> ) <sub>3</sub> C-O-C-Ċ	NN54	gauche	(HO)(Ċ(O))CH-O-C-OH
NN14	gauche	HOCH <sub>2</sub> -O-C-CH <sub>3</sub>	NN55	gauche	ONO-O-C-CH <sub>3</sub>
NN15	gauche	Ċ(O)-O-C-CH <sub>3</sub>	NN56	gauche	ĊH <sub>2</sub> -O-C-CH <sub>3</sub>
NN16	gauche	CH <sub>3</sub> -O-Ċ-CH <sub>3</sub>	NN57	gauche	CH <sub>3</sub> -O-Ċ-OH
NN17	cis	CH <sub>3</sub> -C(O)-C-OH	NN58	gauche	CH <sub>3</sub> -O-Ċ-OCH <sub>3</sub>
NN18	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-OH	NN59	gauche	CH <sub>3</sub> CH <sub>2</sub> -O-Ċ-OH
NN19	cis	CH <sub>3</sub> -C <sub>d</sub> =C <sub>d</sub> -OH	NN60	gauche	CH <sub>3</sub> -C-C <sub>d</sub> -ONO
NN20	cis	CH <sub>3</sub> -C(O)-Ċ-CH <sub>3</sub>	NN61	gauche	H <sub>2</sub> C <sub>d</sub> =C <sub>d</sub> -O-NO
NN21	cis	CH <sub>3</sub> -C(O)-C(O)-Ċ	NN62	gauche	CH <sub>3</sub> -C <sub>d</sub> =Ċ <sub>d</sub> -OH
NN22	gauche	CH <sub>3</sub> -C-C-ĊH <sub>2</sub>	NN63	cis	CH <sub>3</sub> -C(O)-C-NO <sub>2</sub>
NN23	gauche	CH <sub>3</sub> -C-C-C(O)H	NN64	cis	CH <sub>3</sub> O-Ċ-C <sub>d</sub> =C <sub>d</sub> H <sub>2</sub>
NN24	gauche	CH <sub>3</sub> -C-C-OH	NN65	cis	HO-Ċ-C <sub>d</sub> =C <sub>d</sub> H <sub>2</sub>
NN25	gauche	CH <sub>3</sub> -C-C-Ċ	NN66	cis	HO-Ċ-C <sub>d</sub> =C <sub>d</sub> H(CH <sub>3</sub> )
NN26	gauche	CH <sub>3</sub> -C-C-NO <sub>2</sub>	NN67	cis	CH <sub>3</sub> -C <sub>d</sub> -O-NO
NN27	gauche	CH <sub>3</sub> -C-C-C(CH <sub>3</sub> )(NO <sub>2</sub> )(Ċ)	NN68	cis	CH <sub>3</sub> -C <sub>d</sub> =Ċ <sub>d</sub> -C(O)OH
NN28	gauche	CH <sub>3</sub> -C-C-OĊH <sub>2</sub>	NN69	cis	CH <sub>3</sub> -C <sub>d</sub> =Ċ <sub>d</sub> -OCH <sub>3</sub>
NN29	gauche	CH <sub>3</sub> -C-C-O(C(O))Ċ(O))	NN70	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-NO <sub>2</sub>
NN30	gauche	CH <sub>3</sub> -C-C-OOC(O)	NN71	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-Ċ-OH
NN31	gauche	CH <sub>3</sub> CH <sub>2</sub> -C-C-C(O)H	NN72	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-Ċ-CH <sub>3</sub>
NN32	gauche	CH <sub>3</sub> CH <sub>2</sub> -C-C-NO <sub>2</sub>	NN73	cis	CH <sub>3</sub> -C(O)-C <sub>d</sub> =C <sub>d</sub> H <sub>2</sub>
NN33	gauche	CH <sub>3</sub> CH <sub>2</sub> -C-C-OH	NN74	cis	HO-C(O)-C <sub>d</sub> =C <sub>d</sub> H <sub>2</sub>
NN34	gauche	HO-C-C-OH	NN75	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C(O)-Ċ
NN35	gauche	CH <sub>3</sub> -C-Ċ-OH	NN76	cis	Ċ-C(O)-O-CH <sub>3</sub>
NN36	gauche	CH <sub>3</sub> CH <sub>2</sub> -C-Ċ-OH	NN77	cis	CH <sub>3</sub> -C(O)-O-Ċ
NN37	gauche	HO-C-Ċ-CH <sub>3</sub>	NN78	cis	CH <sub>3</sub> CH <sub>2</sub> -C(O)-O-Ċ
NN38	gauche	CH <sub>3</sub> -C(O)-C-NO	NN79	cis	(CH <sub>3</sub> ) <sub>2</sub> CH-C(O)-O-Ċ
NN39	gauche	CH <sub>3</sub> -C(O)-C-NO <sub>2</sub>	NN80	cis	CH <sub>3</sub> -O-Ċ-OH
NN40	gauche	CH <sub>3</sub> -C(O)-C-Ċ(O)	NN81	cis	CH <sub>3</sub> CH <sub>2</sub> -O-Ċ-OH
NN41	gauche	CH <sub>3</sub> CH <sub>2</sub> -C(O)-C-NO			
Secondary Corrections					
SC(a)		HO-C-C=O	SC(e)		HO-C-C-O-R
SC(b)		HO-Ċ-C=O	SC(f)		HO-C-O-R
SC(c)		HO-C-N=O	SC(g)		HO-Ċ-O-R
SC(d)		HO-C-NOO	SC(h)		HO-C-Ċ

<sup>a</sup> NN1–NN21 groups are the only ones that had statistically significant interactions upon regression out of the 81 possible types. The secondary interactions between the hydrogen atom of a hydroxyl group with (a) carbonyl oxygen bonded to carbon, (b) carbonyl oxygen bonded to the carbon radical center, (c) N(O) oxygen bonded to carbon, (d) N(O<sub>2</sub>) oxygen bonded to carbon, (e) the hydroxyl or ether group bonded to the β-carbon, (f) the hydroxyl or ether group bonded to carbon, (g) the hydroxyl or ether group bonded to the carbon radical center, and (h) the oxygen radical center bonded to carbon. Secondary corrections (a)–(e) are the only statistically significant corrections.

based on atomization energies estimates the enthalpy of formation by using experimental and theoretical results for elements in their standard states. As shown in eq 9, the heat of formation of C<sub>a</sub>H<sub>b</sub>O<sub>c</sub>N<sub>d</sub> can be calculated from the atomization energy

$$\Delta H_{f,298}^{\circ}(C_a H_b O_c N_d) = [a\Delta H_{f,298}^{\circ}(C) + b\Delta H_{f,298}^{\circ}(H) + c\Delta H_{f,298}^{\circ}(O) + d\Delta H_{f,298}^{\circ}(N)] - [aH^{298}(C) + bH^{298}(H) + cH^{298}(O) + dH^{298}(N)] + H^{298}(C_a H_b O_c N_d) \quad (9)$$

where the heats of formation of atomic carbon, hydrogen, oxygen, and nitrogen are the experimental values obtained from

the JANAF tables<sup>49</sup> ( $\Delta H_{f,298}^{\circ}(C) = 171.21 \text{ kcal mol}^{-1}$ ,  $\Delta H_{f,298}^{\circ}(H) = 52.10 \text{ kcal mol}^{-1}$ ,  $\Delta H_{f,298}^{\circ}(O) = 59.43 \text{ kcal mol}^{-1}$ ,  $\Delta H_{f,298}^{\circ}(N) = 112.97 \text{ kcal mol}^{-1}$ ), and  $H^{298}(C)$ ,  $H^{298}(H)$ ,  $H^{298}(O)$ , and  $H^{298}(N)$  are the atomization enthalpies of atomic carbon, hydrogen, oxygen, and nitrogen at 298 K, respectively, and  $H^{298}(C_a H_b O_c N_d)$  is the enthalpy of C<sub>a</sub>H<sub>b</sub>O<sub>c</sub>N<sub>d</sub> at the same temperature. These enthalpies are calculated using eq 8.

The other methods, that is, the methods based on bond separation reactions and isodesmic reactions, are known to cancel out consistent errors in a particular quantum chemical calculation and are thus more accurate than the atomization energy method. An isodesmic reaction is constructed by writing a reaction containing the species for which the heat of formation

**TABLE 2: Experimental and Calculated Enthalpies of Formation in kcal mol<sup>-1</sup> at 298 K for Reference Molecules Used in Isodesmic Reactions<sup>a</sup>**

molecule	reference	exp	calc	abs err
C <sub>2</sub> H <sub>6</sub>	Pittam and Pilcher <sup>55</sup>	-20.04 ± 0.07	-20.03	0.01
C <sub>2</sub> H <sub>4</sub>	Chase <sup>49</sup>	12.54 ± 0.07	12.30	0.24
CH <sub>4</sub>	Prosen and Rossini <sup>56</sup>	-17.89 ± 0.08	-17.89	0.00
HCHO	Fletcher and Pilcher <sup>57</sup>	-25.95 ± 0.11	-26.89	0.94
CH <sub>3</sub> OH	Hine and Arata <sup>58</sup>	-48.00	-48.23	0.23
H <sub>2</sub> O	Cox et al. <sup>59</sup>	-57.7978 ± 0.0096	-57.60	0.20
H <sub>2</sub> O <sub>2</sub>	Ruscic et al. <sup>60</sup>	-32.45 ± 0.04	-31.99	0.46
HONO	Gurvich et al. <sup>61</sup>	-18.87	-19.02	0.15
HONO <sub>2</sub>	Dorofeeva et al. <sup>62</sup>	-32.10 ± 0.12	-32.77	0.67
CH <sub>3</sub> NO	Pedley <sup>63</sup>	16.71	16.98	0.27
CH <sub>3</sub> NO <sub>2</sub>	Knobel et al. <sup>64</sup>	-19.3 ± 0.3	-18.53	0.77
C <sub>2</sub> H <sub>5</sub>	Tsang <sup>65</sup>	28.4 ± 0.5	28.49	0.09
C <sub>2</sub> H <sub>3</sub>	Pilgrim and Taatjes <sup>66,67</sup>	70.9 ± 0.3	70.29	0.61
CH <sub>3</sub>	Ruscic et al. <sup>68</sup>	35.06 ± 0.07	34.29	0.77
HC(O)	Baulch et al. <sup>69</sup>	10.3	9.44	0.86
CH <sub>3</sub> O	Ruscic et al. <sup>68</sup>	5.02 ± 0.50	4.43	0.59
HOCH <sub>2</sub>	Ruscic et al. <sup>68</sup>	-4.06 ± 0.17	-4.08	0.02
HO	Ruscic et al. <sup>60</sup>	2.94 ± 0.06	3.07	0.13
MAD				0.39
LD, MD				0.00, 0.94
C <sub>3</sub> H <sub>8</sub>	Pittam and Pilcher	-25.02 ± 0.12	-24.84	0.18
CH <sub>3</sub> CHO	Pedley et al. <sup>70</sup>	-39.70 ± 0.12	-40.00	0.30
CH <sub>3</sub> C(O)CH <sub>3</sub>	Wiberg et al. <sup>71</sup>	-52.23 ± 0.14	-52.09	0.14
C <sub>2</sub> H <sub>5</sub> OH	Green <sup>72</sup>	-56.23 ± 0.12	-56.03	0.20
CH <sub>3</sub> OCH <sub>3</sub>	Pilcher et al. <sup>73</sup>	-43.99 ± 0.12	-44.39	0.40
CH <sub>3</sub> OOH	Khursan and Martemyanov <sup>74</sup>	-31.3 ± 1.2	-30.69	0.62
CH <sub>3</sub> ONO <sub>2</sub>	Ray and Ogg <sup>75</sup>	-29.2 ± 0.3	-30.56	1.36
C <sub>3</sub> H <sub>7</sub>	Tsang <sup>76</sup>	24.0 ± 0.5	24.06	0.06
CH <sub>3</sub> C(O)	Ruscic et al. <sup>68</sup>	-2.46 ± 0.43	-2.75	0.29
C <sub>2</sub> H <sub>5</sub> O	Burcat and Ruscic <sup>77</sup>	-3.25 ± 0.96	-3.36	0.34
C <sub>2</sub> H <sub>5</sub> O	Blanksby et al. <sup>78</sup>	-6.8 ± 2.3	-5.42	1.38
MAD				0.47
LD, MD				0.11, 1.38

<sup>a</sup> Abbreviations: exp = experimental value from literature; calc = value from this study; abs err = |calc - exp|; MAD = mean absolute deviation; and LD, MD = minimum and maximum deviation.

value is to be estimated from the absolute enthalpy obtained from quantum mechanics. For example, if the enthalpy of formation of propane is sought, an isodesmic reaction can be written, as shown in eq 10



This reaction conserves the number of the same types of bonds in the reactants and products. All of the other reactants and products in the isodesmic reaction must have experimental enthalpies of formation, and the unknown heat of formation is calculated from the heat of reaction obtained from the absolute enthalpies from quantum chemistry as shown in eq 11

$$\Delta H_{f,298}^{\circ}(\text{C}_3\text{H}_8) = 2(\Delta H_{f,298}^{\circ}(\text{C}_2\text{H}_6) - H^{298}(\text{C}_2\text{H}_6)) + (H^{298}(\text{CH}_4) - \Delta H_{f,298}^{\circ}(\text{CH}_4)) + H^{298}(\text{C}_3\text{H}_8) \quad (11)$$

Bond separation reactions are a subclass of isodesmic reactions and use the smallest molecules that contain the desired bond types. Isodesmic reactions have been shown to give accurate estimates of enthalpies of formation.<sup>3,19,20,33,36,37,50,51</sup> In this work, the enthalpies of formation calculated from isodesmic reactions are used to obtain the group additivity values reported.

A group additivity scheme was developed for  $\Delta H_f^{\circ}$ ,  $S^{\circ}$ , and  $C_p$  based on the thermodynamic values obtained from the quantum chemical calculations. This method followed the development by Benson and accounts for the interactions between neighboring atoms. The molecule is assumed to consist of substructural units

called groups, which are defined as a polyvalent atom and all of its associated ligands. Each of these groups contributes a constant amount to the thermodynamic property of interest for a particular molecule. Nonbonded or next-nearest-neighbor (NN) interactions were also included to account for longer range interactions. These included gauche and cis corrections and other interactions that were identified to be present in the species in this study, as discussed in the following section.

### 3. Statistical Methods

The group values for groups, NN, and additional interaction parameters were estimated using the multiple linear regression (MLR) statistical methodology. In MLR, the simultaneous linear equations for all of the group values can be represented by matrix notation, where the solution for the group values is given by eq 12 as follows

$$\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (12)$$

where  $\beta$  represents the  $k \times 1$  vector of  $k$  unknown group values for the thermodynamic properties  $\Delta H_f^{\circ}$ ,  $S^{\circ}$ , and  $C_p$  from 300 to 1500 K,  $\mathbf{X}$  is the  $n \times k$  matrix of the  $k$  groups for the  $n$  molecules, and  $\mathbf{y}$  is the  $n \times 1$  matrix of the calculated thermodynamic properties of  $n$  molecules. The two underlying assumptions of the MLR methodology are normality and constant variance of the residuals. The first assumption of normality can be proven by comparing the cumulative distribution of the residuals to the cumulative normal distribution with the mean and variance of the residual data. The second

**TABLE 3: Comparison of Calculated (Recommended Values in This Work Based on G3//B3LYP) Enthalpies of Formation in kcal mol<sup>-1</sup> at 298 K with Literature (Experimental and Calculated) Values for Molecules Which Had Properties Reported Previously<sup>a</sup>**

molecule	reference	method	lit	calc	err
(7)	Chen and Bozzelli <sup>37</sup>	QC	-17.74 ± 1.13	<b>-18.43</b>	-0.69
(26)	Holmes et al. <sup>52</sup>	exp	-23.0 ± 3	<b>-15.53</b>	7.47
(28)	Chen and Bozzelli <sup>37</sup>	QC	-10.06 ± 0.85	<b>-10.15</b>	-0.09
(86)	Sumathi and Green <sup>15</sup>	QC	-22.97	<b>-24.68</b>	-1.71
	Holmes et al. <sup>52</sup>	exp	-25.6 ± 3		0.92
	Tsang <sup>79</sup>	exp	-24.8 ± 1.1		0.12
(87)	Sumathi and Green <sup>15</sup>	QC	-17.20	<b>-18.54</b>	-1.34
	Chen and Bozzelli <sup>37</sup>	QC	-17.33 ± 2.38		-1.21
(92)	da Silva and Bozzelli <sup>12</sup>	QC	-7.1	<b>-7.39</b>	-0.29
(95)	da Silva and Bozzelli <sup>12</sup>	QC	-11.5	<b>-11.90</b>	-0.40
(103)	Sumathi and Green <sup>15</sup>	QC	-12.72	<b>-14.03</b>	-1.31
	Holmes et al. <sup>52</sup>	exp	-14.5 ± 3		0.47
	Tsang <sup>79</sup>	exp	-15.3 ± 1		1.27
(112)	da Silva and Bozzelli <sup>12</sup>	QC	3.6	<b>3.35</b>	-0.25
	Berkowitz et al. <sup>80</sup>	exp	2.5 ± 2.2		0.85
(113)	Wenthold and Squires <sup>81</sup>	exp	-59.6 ± 2.9	<b>-56.86</b>	2.74
	Holmes et al. <sup>52</sup>	exp	-61.6 ± 3		4.74
(123)	Sumathi and Green <sup>15</sup>	QC	17.07	<b>15.44</b>	-1.63
(124)	Sumathi and Green <sup>15</sup>	QC	-7.47	<b>-7.93</b>	-0.46
	Chen and Bozzelli <sup>37</sup>	QC	-7.93 ± 1.81		0.00
	Holmes et al. <sup>52</sup>	exp	-10.6 ± 3		2.67
(152)	Finlayson-Pitts and Pitts <sup>82</sup>	exp	-49.59	<b>-45.99</b>	3.60
	Holmes et al. <sup>52</sup>	exp	-51.7 ± 3		5.71
(162)	Holmes et al. <sup>52</sup>	exp	-37.7 ± 3	<b>-30.30</b>	7.40
	Schatz et al. <sup>83</sup>	QC	-31.1		0.80
(167)	Sumathi and Green <sup>15</sup>	QC	-46.50	<b>-46.76</b>	-0.26
(168)	Sumathi and Green <sup>15</sup>	QC	-56.41	<b>-56.47</b>	-0.06
(169)	Sumathi and Green <sup>15</sup>	QC	-65.98	<b>-65.71</b>	0.27
(173)	Finlayson-Pitts and Pitts <sup>82</sup>	exp	-41.11	<b>-39.75</b>	1.36
(176)	Sumathi and Green <sup>15</sup>	QC	-41.62	<b>-44.10</b>	-2.48
	Holmes et al. <sup>52</sup>	exp	-46.0 ± 3		1.90
(275)	NIST	exp	-25.43 ± 0.48	<b>-25.61</b>	-0.18
(278)	NIST	exp	-38.2 ± 2	<b>-40.18</b>	-1.98
(282)	NIST	exp	-18.0	<b>-15.81</b>	2.19
(283)	NIST	exp	-79.0 ± 1.0	<b>-77.20</b>	1.80
(284)	NIST	exp	-27.4 ± 2.6	<b>-26.30</b>	1.10
(286)	NIST	exp	-32.6	<b>-34.95</b>	-2.35
(311)	da Silva et al. <sup>14</sup>	QC	-29.98	<b>-29.44</b>	0.54
	Turecek and Havlas <sup>84</sup>	exp	-30.6		1.16
	NIST	exp	-29.8 ± 2		0.36
	NIST	exp	-26.5 ± 2		-2.94
(312)	Turecek <sup>85</sup>	exp	-40.4	<b>-35.47</b>	4.93

<sup>a</sup> Abbreviations: lit = value from data in the literature; calc = value from this study; err = calc - lit; exp = experimental result; QC = quantum chemical calculation; NIST = National Institute of Standards and Technology database.

assumption of constant variance can be tested by plotting the residuals versus the fitted thermodynamic values. If the variability in the residuals is roughly constant over the range of fitted values, then the assumption of constant variance is met.

Once the group values are calculated, inferences can be made about whether they have a statistically significant effect on the thermodynamic properties. This can be done by testing the null hypothesis that a group value is 0 versus the alternative hypothesis that the group value is not equal to 0. If the null hypothesis is not rejected, this indicates that the corresponding group is not a useful predictor of the thermodynamic property and may be removed from the model. A *t* test is useful in testing this hypothesis, and the null hypothesis is rejected if eq 13 is satisfied

$$|t_j| = \frac{|\beta_j|}{\text{SE}(\beta_j)} > t_{n-k,\alpha/2} \quad (13)$$

where  $\beta_j$  is the group value for the *j*th group, SE is the standard error of the *j*th group value, and  $t_{n-k,\alpha/2}$  is the *t* value with *n* - *k* degrees of freedom at the  $\alpha$  confidence level. In this study, both 80 and 98% confidence levels were used to remove groups that

were deemed insignificant. A group was only removed if it failed to reject the null hypothesis for  $\Delta H_f^\circ$ ,  $S^\circ$ , and  $C_p$  at every temperature.

#### 4. Results and Discussion

G3//B3LYP calculations were performed for a total of 323 molecules shown in Figure 1 and summarized in Table S1 of the Supporting Information with their corresponding external and internal symmetry numbers, optical isomers as defined in eq 7, and constituent groups and interactions. The interactions included in this table are gauche, cis, and secondary interactions as defined in Table 1. The geometries and vibrational frequencies along with hindered internal rotation results (symmetry number, reduced moment of inertia, and Fourier series coefficients) are also provided for each molecule in the Supporting Information. There were 122 unknown group values, 16 gauche corrections, 5 cis corrections, and 5 secondary corrections that were estimated from this database of molecules.

Bond separation reactions were then formulated to estimate the enthalpies of formation for each molecule. The reference molecules used for bond separation reactions were C<sub>2</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>4</sub>, CH<sub>4</sub>, H<sub>2</sub>C(O), CH<sub>3</sub>OH, H<sub>2</sub>O, H<sub>2</sub>O<sub>2</sub>, HONO, HONO<sub>2</sub>, CH<sub>3</sub>NO,

TABLE 4: Group Values for Known Groups and gauche Correction Values

group	footnote	$\Delta H_f^{298K}$ (kcal mol <sup>-1</sup> )	$S^{298K}$ (cal (mol-K) <sup>-1</sup> )	$C_p^T$ (cal (mol-K) <sup>-1</sup> )						
				300 K	400 K	500 K	600 K	800 K	1000 K	1500 K
C-(C) <sub>3</sub> (H)	<i>a</i>	-1.65	-11.54	5.57	6.98	8.01	8.77	9.80	10.47	11.43
C-(C) <sub>3</sub> (O)	<i>b</i>	-7.03	-34.46	5.08	6.78	7.86	8.39	8.63	8.50	8.05
C-(C) <sub>2</sub> (H) <sub>2</sub>	<i>a</i>	-4.90	9.53	6.02	7.22	8.39	9.43	11.09	12.34	14.27
C-(C) <sub>2</sub> (H)(O)	<i>b</i>	-7.42	-12.38	5.03	6.74	8.05	8.97	10.07	10.69	11.39
C-(C)(C)(H) <sub>2</sub>	<i>c</i>	-4.90	9.42	5.50	6.95	8.25	9.35	11.07	12.34	14.25
C-(C)(C <sub>d</sub> )(H) <sub>2</sub>	<i>a</i>	-4.52	10.02	5.24	6.77	8.10	9.19	10.91	12.17	14.14
C-(C)(CO)(H) <sub>2</sub>	<i>d,e</i>	-5.26	9.6	6.2	7.7	8.7	9.5	11.1	12.2	14.2
C-(C)(H) <sub>3</sub>	<i>a</i>	-10.25	30.40	5.96	7.60	9.13	10.49	12.72	14.46	17.28
C-(C)(H) <sub>2</sub> (O)	<i>b</i>	-7.92	9.10	5.37	6.99	8.41	9.58	11.35	12.60	14.45
C-(C)(H) <sub>3</sub>	<i>c</i>	-10.08	30.41	6.19	7.84	9.40	10.79	13.02	14.77	17.58
C-(CO)(H) <sub>3</sub>	<i>d</i>	-10.31	30.41	6.19	7.84	9.40	10.79	13.02	14.77	17.58
C-(C <sub>d</sub> )(H) <sub>3</sub>	<i>a</i>	-10.25	30.40	5.96	7.60	9.13	10.49	12.72	14.46	17.28
C-(H) <sub>3</sub> (O)	<i>f</i>	-10.00	30.41	6.19	7.84	9.40	10.79	13.02	14.77	17.58
Ċ-(C) <sub>2</sub> (H)	<i>a</i>	42.38	13.52	4.43	5.08	5.79	6.45	7.51	8.30	9.45
Ċ-(C)(H) <sub>2</sub>	<i>a</i>	40.15	33.22	5.45	6.65	7.67	8.53	9.88	10.94	12.69
C <sub>d</sub> -(C) <sub>2</sub>	<i>a</i>	10.83	-13.36	4.68	5.19	5.53	5.80	6.20	6.46	6.78
C <sub>d</sub> -(C)(C)	<i>c</i>	10.34	-12.30	4.10	4.71	5.09	5.36	5.90	6.18	6.40
C <sub>d</sub> -(C)(H)	<i>a</i>	8.87	7.87	4.41	5.20	5.98	6.68	7.80	8.62	9.84
C <sub>d</sub> -(Ċ)(H)	<i>c</i>	8.87	7.97	4.16	5.03	5.81	6.50	7.65	8.45	9.62
C <sub>d</sub> -(H) <sub>2</sub>	<i>a</i>	6.00	27.67	4.92	6.20	7.35	8.33	9.90	11.10	13.04
C <sub>d</sub> -(H)(O)	<i>f</i>	8.60	6.2	4.8	6.5	7.6	8.4	9.1	9.6	10.5
CO-(C) <sub>2</sub>	<i>f,e</i>	-31.69	15.01	5.6	6.3	7.1	7.8	8.9	9.6	11.0
CO-(C)(C)	<i>c,e</i>	-31.4	15.01	5.6	6.3	7.1	7.8	8.9	9.6	11.0
CO-(C)(H)	<i>d,e</i>	-29.47	34.9	7	7.8	8.8	9.7	11.2	12.2	13.9
CO-(C)(O)	<i>b</i>	-34.86	15.14	5.52	6.04	6.80	7.64	8.91	9.99	11.54
CO-(CO)(H)	<i>g</i>	-25.19	21.3	7.22	8.51	9.73	10.77	12.24	13.13	14.20
CO-(H)(O)	<i>b</i>	-32.16	34.93	6.84	7.67	8.81	10.06	11.97	13.32	14.99
O-(C) <sub>2</sub>	<i>b</i>	-23.18	9.11	2.86	2.47	2.56	2.79	3.27	3.67	4.10
O-(C)(C)	<i>c,e</i>	-23.2	8.68	3.4	3.7	3.7	3.8	4.4	4.6	4.9
O-(C)(CO)	<i>b</i>	-43.43	4.93	2.68	3.24	3.77	4.12	4.76	5.09	5.10
O-(C)(H)	<i>b</i>	-37.86	29.45	4.15	4.20	4.54	4.94	5.67	6.30	7.36
O-(C)(O)	<i>h</i>	-4.54	8.50	3.20	3.40	3.60	3.80	4.10	4.20	4.20
O-(Ċ)(CO)	<i>c,e</i>	-23.2	8.68	3.4	3.7	3.7	3.8	4.4	4.6	4.9
O-(Ċ)(H)	<i>c,e</i>	-37.9	29.07	4.30	4.40	4.82	5.23	6.02	6.61	7.25
O-(CO)(H)	<i>f,e</i>	-57.79	24.5	3.8	5	5.8	6.3	7.2	7.8	9.3
alkane gauche	<i>a</i>	0.69	-0.17	-0.22	-0.26	-0.26	-0.24	-0.18	-0.13	-0.08
ether gauche	<i>c</i>	0.5								

<sup>a</sup>  $\Delta H_f^{298K}$  from Sabbe et al.<sup>11</sup> and  $S^{298K}$  and  $C_p^T$  from Sabbe et al.<sup>21</sup> <sup>b</sup>  $\Delta H_f^{298K}$ ,  $S^{298K}$ , and  $C_p^T$  from Sumathi and Green.<sup>15</sup> <sup>c</sup>  $\Delta H_f^{298K}$ ,  $S^{298K}$ , and  $C_p^T$  from Benson.<sup>1</sup> <sup>d</sup>  $\Delta H_f^{298K}$  from da Silva and Bozzelli<sup>12</sup> and  $S^{298K}$  and  $C_p^T$  from Benson.<sup>1</sup> <sup>e</sup>  $C_p^T$  at 1500 K extrapolated from fitting the 300–1000 K  $C_p$  data to a log function. <sup>f</sup>  $\Delta H_f^{298K}$  from Cohen<sup>86</sup> and  $S^{298K}$  and  $C_p^T$  from Benson.<sup>1</sup> <sup>g</sup>  $\Delta H_f^{298K}$  from Cohen,<sup>86</sup>  $S^{298K}$  from Benson,<sup>1</sup> and  $C_p^T$  estimated from the heat capacity data for HC(O)C(O)H.<sup>87</sup> <sup>h</sup>  $\Delta H_f^{298K}$  from Sumathi and Green<sup>15</sup> and  $S^{298K}$  and  $C_p^T$  from Sebban et al.<sup>19</sup>

CH<sub>3</sub>NO<sub>2</sub>, Ċ<sub>2</sub>H<sub>5</sub>, Ċ<sub>2</sub>H<sub>3</sub>, ĊH<sub>3</sub>, ḢC(O), CH<sub>3</sub>Ö, HȮCH<sub>2</sub>, and HȮO. The top section of Table 2 lists the enthalpies of formation at 298 K for these reference molecules obtained from G3//B3LYP calculations using the atomization energy method and shows a comparison of these values to experimental data. The average absolute deviation between experimental and calculated heats of formation was 0.39 kcal mol<sup>-1</sup>, with a maximum deviation of 0.94 kcal mol<sup>-1</sup> and a minimum deviation of 0.0 kcal mol<sup>-1</sup>. The bond separation reactions along with the enthalpy of reaction calculated from the G3 enthalpies for all molecules are listed in Table S2 of the Supporting Information.

The other reactions listed in Table S2 of the Supporting Information are isodesmic reactions. Two isodesmic reactions were designed for each molecule for which the heat of formation was to be estimated. The reactions in italics indicate those reactions which are homodesmotic, which means that in addition to the bond types being conserved overall between reactants and products, the bond orders of the bonds surrounding each atom are also conserved. The reference molecules used for isodesmic reactions included C<sub>3</sub>H<sub>8</sub>, CH<sub>3</sub>CHO, CH<sub>3</sub>C(O)CH<sub>3</sub>, C<sub>2</sub>H<sub>5</sub>OH, CH<sub>3</sub>OCH<sub>3</sub>, CH<sub>3</sub>OOH, CH<sub>3</sub>ONO<sub>2</sub>, C<sub>3</sub>H<sub>7</sub>, CH<sub>3</sub>Ċ(O), C<sub>2</sub>H<sub>5</sub>Ö, and C<sub>2</sub>H<sub>5</sub>OÖ in addition to all of the reference molecules used in the bond separation reactions. The bottom section of

Table 2 lists these additional molecules and their experimental and calculated heats of formation. The average absolute deviation is 0.47 kcal mol<sup>-1</sup>, with a maximum deviation of 1.38 kcal mol<sup>-1</sup> and a minimum deviation of 0.11 kcal mol<sup>-1</sup>. The highest deviations were found for CH<sub>3</sub>ONO<sub>2</sub> and C<sub>2</sub>H<sub>5</sub>OÖ.

The estimated enthalpies of formation for all of the molecules from the atomization energy, bond separation reaction, and isodesmic reaction methods are shown in Table S3 of the Supporting Information. The recommended values for the enthalpies of formation are an average of the two values obtained from the isodesmic reactions. At the bottom of each column, the mean absolute, maximum, and minimum deviations between the recommended value and the value obtained for a given method are tabulated. The mean absolute deviations for the atomization energy (AE) and bond separation (BS) reaction methods are 0.48 and 0.87 kcal mol<sup>-1</sup>, respectively. In most cases, the AE and BS methods predict very similar enthalpies of formation, but there are cases where the enthalpy predicted by both of these methods is much more than that obtained from the isodesmic reactions. Generally, isodesmic reactions are known to predict the enthalpy of formation more accurately, and therefore, those values were used in this work to regress the group additivity values.



TABLE 5: Group Values Estimated in This Work

group	$\Delta H_f^{298K}$ (kcal mol <sup>-1</sup> )	$S^{298K}$ (cal (mol-K) <sup>-1</sup> )	$C_p^T$ (cal (mol-K) <sup>-1</sup> )						
			300 K	400 K	500 K	600 K	800 K	1000 K	1500 K
C-(C) <sub>2</sub> (C)(CO)	-0.05	-33.83	8.05	9.15	9.32	9.26	8.97	8.80	8.03
C-(C) <sub>2</sub> (C)(O)	-6.16	-34.80	7.93	9.59	10.21	10.37	10.22	9.90	9.31
C-(C) <sub>2</sub> (CO)(O)	11.68	-7.75	9.64	11.82	12.85	13.35	13.55	13.51	12.91
C-(C) <sub>2</sub> (C)(O)	2.90	0.65	14.03	16.18	17.25	17.81	18.25	18.29	18.17
C-(C) <sub>2</sub> (NO <sub>2</sub> )(O)	-4.19	30.18	17.95	20.96	22.93	24.26	25.85	26.73	27.65
C-(C) <sub>2</sub> (O)(O)	-2.00	-7.29	9.49	12.04	13.38	14.02	14.35	14.20	13.71
C-(C)(C)(CO)(H)	-1.83	-13.54	7.72	8.66	9.16	9.55	10.12	10.27	10.67
C-(C)(C)(CO)(O)	-4.63	-39.40	8.93	11.54	12.34	12.37	11.70	10.97	9.63
C-(C)(C)(H)(O)	-7.19	-14.27	7.89	9.42	10.26	10.78	11.41	11.75	12.24
C-(C)(C)(O) <sub>2</sub>	-18.51	-40.12	9.04	12.66	14.00	14.14	13.17	12.01	10.37
C-(C)(CO) <sub>2</sub> (O)	14.46	-10.21	12.58	14.48	14.69	14.50	13.84	13.48	12.45
C-(C)(CO)(C)(O)(H)	8.49	21.79	15.11	16.41	17.21	17.82	18.62	19.29	19.83
C-(C)(CO)(C)(O)(O)	7.47	-5.31	16.00	18.74	19.74	20.08	19.95	19.55	18.53
C-(C)(CO)(H)(O)	11.40	13.29	9.73	11.68	12.83	13.62	14.55	15.19	15.75
C-(C)(CO)(O)(O)	-0.93	-11.19	10.39	13.44	14.95	15.65	15.78	15.39	14.12
C-(C)(H)(NO <sub>2</sub> )(O)	-2.60	51.80	17.72	20.84	23.16	24.85	27.04	28.34	30.01
C-(C)(H)(O)(O)	-2.79	13.50	9.16	11.75	13.41	14.45	15.58	16.10	16.68
C-(C)(CO)(H)(O)	-4.46	-17.62	8.54	10.99	12.12	12.64	12.86	12.79	12.48
C-(C)(H)(O) <sub>2</sub>	-18.79	-18.29	7.79	11.63	13.61	14.37	14.05	13.57	12.94
C-(CO) <sub>2</sub> (H)(O)	15.04	12.74	11.58	13.05	13.68	14.08	14.53	14.96	15.02
C-(CO)(C)(O)(H) <sub>2</sub>	5.52	44.46	14.47	16.15	17.37	18.40	19.97	21.26	22.81
C-(CO)(H) <sub>2</sub> (O)	10.97	35.92	9.07	11.49	13.18	14.43	16.10	17.28	18.71
C-(CO)(H)(O)(O)	0.86	11.47	10.06	13.27	14.93	15.78	16.38	16.58	16.56
C-(C)(O)(H)(O) <sub>2</sub>	-2.37	15.65	14.78	17.68	19.42	20.48	21.50	21.80	22.09
C-(H) <sub>2</sub> (O)(O)	-2.55	36.67	9.01	11.49	13.30	14.64	16.50	17.69	19.49
C-(H)(O) <sub>2</sub> (O)	-8.15	12.24	10.77	13.37	14.71	15.43	16.15	16.44	16.91
C-(O) <sub>3</sub> (O)	-23.31	-10.05	9.71	13.30	14.85	15.43	15.53	15.18	14.53
C-(C) <sub>2</sub> (CO)	31.48	-12.45	4.13	5.22	6.00	6.37	7.04	7.40	6.75
C-(C) <sub>2</sub> (O)	33.49	-10.03	4.98	5.91	6.34	6.57	6.77	6.81	7.08
C-(C)(C <sub>d</sub> )(O)	20.04	-12.82	6.58	7.95	8.44	8.59	8.48	8.36	8.48
C-(C)(CO)(H)	33.44	8.89	5.61	6.97	7.85	8.46	9.28	9.88	10.03
C-(C)(CO)(O)	22.89	-15.41	4.63	6.11	7.34	8.14	8.28	8.25	6.90
C-(C)(H)(O) <sup>a</sup>	33.65	11.22	4.93	6.19	7.18	7.91	8.72	9.30	10.10
C-(C)(O) <sub>2</sub>	23.55	-12.23	6.73	8.55	9.19	9.32	8.83	8.32	8.27
C-(C <sub>d</sub> )(H)(O)	21.14	8.92	6.27	7.99	9.17	9.91	10.13	10.54	11.39
C-(CO)(H) <sub>2</sub>	34.37	29.02	6.69	8.23	9.29	10.09	11.29	12.30	13.32
C-(CO)(H)(O)	25.33	5.83	5.47	7.00	8.33	9.18	9.82	10.40	10.23
C-(CO)(O) <sub>2</sub>	13.93	-18.69	5.46	7.62	9.26	10.06	9.83	9.65	8.38
C-(H) <sub>2</sub> (O) <sup>a,b</sup>	33.44	31.14	6.07	7.34	8.65	9.66	10.78	11.77	13.37
C-(H)(O) <sub>2</sub>	26.09	8.89	7.17	8.18	8.94	9.30	9.17	9.39	9.99
C-(O) <sub>3</sub>	13.72	-12.17	6.71	7.91	8.54	8.69	7.94	7.60	7.56
C <sub>d</sub> -(CO)(H)	15.03	39.59	10.97	12.73	14.17	15.35	17.08	18.27	19.92
C <sub>d</sub> -(CO)	67.03	4.13	4.15	5.43	6.07	6.29	6.17	6.00	5.63
C <sub>d</sub> -(O)	60.08	7.41	4.49	5.47	5.85	6.03	6.08	6.07	6.32
CO-(C)(C <sub>d</sub> )	-34.09	17.65	4.69	4.95	5.53	6.25	7.69	8.64	9.87
CO-(C)(CO)	-18.12	48.14	12.31	13.82	15.31	16.55	18.08	19.15	20.27
CO-(C)(O)	-35.36	40.88	8.02	9.22	10.40	11.37	12.60	13.40	13.56
CO-(C)(CO)	-32.34	24.34	5.68	6.91	8.11	9.08	10.32	10.84	11.81
CO-(C)(H)	-30.82	35.10	6.73	7.53	8.51	9.44	10.99	12.03	14.31
CO-(C)(O)	-33.06	15.06	5.78	6.40	7.31	8.32	9.82	10.73	11.15
CO-(C <sub>d</sub> )(O)	-33.06	15.06	5.78	6.40	7.31	8.32	9.82	10.73	11.15
CO-(CO)(O)	-26.17	51.55	9.92	11.01	11.88	12.55	13.42	14.10	15.11
CO-(CO)(H) <sup>c</sup>	-16.33	67.65	14.03	15.77	17.26	18.50	20.38	21.69	23.56
CO-(CO)(O)	-17.12	49.61	13.91	15.80	17.26	18.47	19.87	20.52	20.70
CO-(H)(O) <sup>c</sup>	-30.30	60.09	9.94	11.47	12.83	13.93	15.53	16.59	18.07
CO-(NO <sub>2</sub> )(O) <sup>c</sup>	-29.60	77.18	19.31	22.09	24.29	25.98	28.20	29.40	30.50
CO-(O)(O)	-30.26	41.48	9.44	10.81	11.94	12.88	13.87	14.37	14.45
O-(C)(C <sub>d</sub> )	-23.56	10.65	2.83	2.22	2.28	2.50	3.12	3.54	3.70
O-(C)(CO) <sup>a</sup>	-28.57	42.29	9.73	10.93	11.86	12.60	13.65	14.28	14.88
O-(C)(NO <sub>2</sub> )	-14.42	45.19	13.11	14.81	16.03	16.99	18.54	19.34	19.74
O-(C)(O) <sup>a</sup>	-3.13	10.08	5.92	5.81	5.45	5.23	5.40	5.48	5.68
O-(C <sub>d</sub> )(H)	-37.90	29.07	4.30	4.40	4.82	5.23	6.02	6.61	7.25
O-(CO)(CO)	-36.52	36.86	12.00	14.00	14.94	15.18	15.02	14.54	13.48
O-(CO)(O)	5.71	32.92	8.53	9.42	9.99	10.21	10.19	9.80	8.73
O-(CO)(H) <sup>a,c</sup>	-44.10	60.22	10.76	12.11	13.22	14.11	15.44	16.36	17.78
O-(CO)(NO) <sup>c</sup>	-3.25	73.42	17.24	18.82	20.06	21.06	22.51	23.44	24.62
O-(CO)(O)	-4.44	41.06	10.39	11.87	12.89	13.61	14.50	15.08	15.79
C-(C) <sub>2</sub> (CO) <sub>2</sub>	2.27	-33.96	6.99	7.93	8.09	8.15	8.04	8.08	7.67
C-(C) <sub>2</sub> (CO)(H) <sup>d</sup>	-1.67	-11.81	5.94	7.46	8.41	9.06	10.04	10.60	11.51

TABLE 5: Continued

group	$\Delta H_f^{298K}$ (kcal mol <sup>-1</sup> )	$S^{298K}$ (cal (mol-K) <sup>-1</sup> )	$C_p^T$ (cal (mol-K) <sup>-1</sup> )						
			300 K	400 K	500 K	600 K	800 K	1000 K	1500 K
	(-2.73)	(-11.54)	(4.99)	(6.18)	(7.15)	(7.85)	(8.73)	(9.33)	(-)
C-(C) <sub>2</sub> (CO)(O)	-4.58	-37.45	7.37	10.02	11.05	11.32	11.01	10.52	9.44
C-(C) <sub>2</sub> (H)(NO) <sup>e</sup>	23.08	23.68	12.31	13.71	15.00	16.09	17.67	18.69	20.10
	(23.4)	(23.1)	(11.2)	(12.7)	(14.0)	(15.1)	(16.8)	(17.9)	(19.5)
C-(C) <sub>2</sub> (NO)(O)	17.58	0.19	13.80	15.98	17.01	17.48	17.81	17.86	17.95
C-(C) <sub>2</sub> (NO <sub>2</sub> )(O)	-23.73	-2.59	15.42	18.92	21.20	22.71	24.37	25.01	25.16
C-(C)(CO) <sub>2</sub> (H)	0.66	-12.75	6.69	8.04	8.72	9.22	9.83	10.38	10.58
C-(C)(CO) <sub>2</sub> (O)	-0.70	-38.65	7.58	10.10	11.12	11.58	11.65	11.36	9.86
C-(C)(CO)(H)(O)	-4.89	-16.14	6.96	9.69	11.03	11.66	12.04	12.19	12.07
C-(C)(CO)(NO)(O)	22.44	-4.38	14.40	17.35	18.42	18.74	18.82	18.84	18.46
C-(C)(CO)(NO <sub>2</sub> )(O)	-13.58	-0.35	18.37	21.23	22.14	22.40	22.42	22.46	22.75
C-(C)(CO)(O) <sub>2</sub>	-13.52	-36.77	7.19	9.94	10.94	11.24	10.99	10.53	9.79
C-(C)(H) <sub>2</sub> (NO) <sup>e</sup>	20.67	44.56	11.61	13.67	15.53	17.09	19.45	21.11	23.47
	(21.4)	(44.3)	(11.8)	(13.6)	(15.2)	(16.7)	(18.9)	(20.5)	(23.0)
C-(C)(H)(NO)(O)	17.61	19.63	13.21	15.95	17.60	18.63	19.85	20.48	21.33
C-(C)(H)(NO <sub>2</sub> )(O)	-21.33	21.49	14.95	18.41	20.91	22.68	24.87	26.01	27.24
C-(C)(NO)(O) <sub>2</sub>	8.39	-1.09	15.84	18.70	19.19	18.77	17.76	17.28	17.46
C-(C)(NO <sub>2</sub> )(O) <sub>2</sub>	-32.24	-4.49	16.88	20.91	23.14	24.45	25.71	25.98	25.68
C-(CO)(H) <sub>2</sub> (NO)	25.14	42.51	13.93	15.42	16.75	17.94	19.72	21.12	22.76
C-(CO)(H) <sub>2</sub> (NO <sub>2</sub> )	-11.43	46.28	15.39	17.92	20.06	21.83	24.33	26.13	28.25
C-(CO)(H) <sub>2</sub> (O)	-5.49	7.36	6.69	9.35	11.02	12.11	13.29	14.00	14.97
C-(CO)(H)(NO)(O)	23.93	15.85	14.59	18.09	19.98	21.00	21.74	21.90	21.75
C-(CO)(H)(NO <sub>2</sub> )(O)	-11.22	20.52	16.12	20.29	22.62	23.95	25.24	25.93	26.69
C-(CO)(NO)(O) <sub>2</sub>	18.18	-3.25	15.67	19.22	20.15	19.97	19.06	18.48	18.05
C-(CO)(NO <sub>2</sub> )(O) <sub>2</sub>	-23.85	-5.16	15.74	20.22	22.81	24.38	25.84	26.27	25.16
C-(H) <sub>2</sub> (NO)(O)	17.51	42.55	13.91	16.48	18.28	19.61	21.48	22.65	24.42
C-(H) <sub>2</sub> (NO <sub>2</sub> )(O)	-19.98	44.44	15.08	18.76	21.51	23.55	26.30	27.95	30.17
C-(H) <sub>2</sub> (O) <sub>2</sub>	-18.01	2.87	6.96	11.43	14.20	15.60	16.38	16.37	16.47
C-(H)(NO)(O) <sub>2</sub>	11.91	21.99	15.17	17.66	18.42	18.46	18.40	18.71	20.17
C-(H)(NO <sub>2</sub> )(O) <sub>2</sub>	-28.77	20.03	16.17	19.98	22.48	24.17	26.17	27.08	28.00
C-(H)(O) <sub>3</sub>	-30.63	-15.21	7.46	10.23	11.45	12.01	12.49	12.69	13.06
C-(NO)(O) <sub>3</sub>	-1.93	-2.41	14.71	17.96	18.93	18.89	18.35	18.00	18.15
C-(NO <sub>2</sub> )(O) <sub>3</sub>	-41.57	-3.46	14.46	19.15	21.79	23.38	25.00	24.45	23.43
C-(O) <sub>4</sub>	-48.22	-45.90	7.95	13.23	15.19	15.45	14.30	12.92	11.08
C <sub>d</sub> -(C)(CO)	11.02	-15.45	4.73	6.24	7.25	7.92	8.51	8.38	6.82
C <sub>d</sub> -(C)(O) <sup>f</sup>	8.94	-13.28	4.60	5.81	6.51	6.98	7.16	7.22	7.30
	(-8.2)	(-12.3)	(3.6)	(4.6)	(5.0)	(5.3)	(5.8)	(6.1)	(6.2)
C <sub>d</sub> -(CO)(H)	10.14	4.84	5.25	6.90	8.18	9.15	10.27	10.65	10.03
C <sub>d</sub> -(CO)(O)	11.26	-16.49	3.73	5.38	6.34	7.13	8.12	8.74	8.19
CO-(C)(C <sub>d</sub> )	-32.71	15.73	5.98	6.12	6.48	6.86	7.70	8.64	10.77
CO-(C)(CO)	-29.88	0.87	6.74	7.71	8.68	9.45	10.31	10.69	10.63
CO-(C)(NO)	-1.12	48.84	13.17	14.27	15.43	16.43	17.72	18.59	19.17
CO-(C)(NO <sub>2</sub> )	-37.63	51.37	16.02	17.76	19.42	20.81	22.62	23.81	25.10
CO-(C <sub>d</sub> )(H)	-32.15	34.66	6.51	7.22	8.02	8.79	10.27	11.68	14.81
CO-(C <sub>d</sub> )(O)	-34.86	15.14	5.52	6.04	6.80	7.64	8.91	9.99	11.54
CO-(CO)(CO)	-21.95	42.12	7.39	7.51	7.35	7.25	7.56	8.08	9.48
CO-(CO)(NO)	6.66	61.76	14.66	15.67	16.31	16.80	17.56	18.21	19.02
CO-(CO)(NO <sub>2</sub> )	-28.25	65.31	16.80	18.33	19.49	20.45	21.93	23.04	24.42
CO-(H)(NO <sub>2</sub> ) <sup>c</sup>	-33.58	71.16	16.52	19.03	21.02	22.60	24.88	26.38	28.42
CO-(NO <sub>2</sub> )(O)	-32.53	51.32	15.08	17.54	19.69	21.56	24.00	25.32	25.98
CO-(O) <sub>2</sub>	-30.04	15.68	6.63	7.73	8.72	9.66	10.58	11.07	10.31
O-(C <sub>d</sub> )(H) <sup>g</sup>	-44.45	27.38	5.22	5.84	6.27	6.50	7.10	7.48	8.02
	(-44.84)	(26.15)	(5.51)	(6.77)	(7.41)	(7.72)	(7.93)	(8.03)	(8.36)
O-(C <sub>d</sub> )(NO) <sup>e</sup>	-7.09	39.56	11.23	11.93	12.40	12.73	13.61	14.19	14.72
	(-5.3)	(39.5)	(11.1)	(11.7)	(12.2)	(12.7)	(13.5)	(14.1)	(14.9)
O-(C <sub>d</sub> )(NO <sub>2</sub> ) <sup>e</sup>	-19.01	44.35	12.45	14.22	15.71	16.80	18.51	19.37	20.00
	(-18.4)	(45.4)	(12.4)	(14.2)	(15.7)	(16.9)	(18.5)	(19.3)	(20.1)
O-(CO)(NO)	-14.15	37.80	12.25	14.28	15.46	15.85	15.65	14.99	13.28
O-(CO)(NO <sub>2</sub> )	-25.61	42.90	13.38	15.99	17.79	18.43	19.17	19.13	18.64
O-(NO)(O) <sup>e</sup>	13.28	41.17	11.91	13.29	14.11	14.62	15.28	15.80	16.57
	(15.2)	(40.7)	(11.7)	(12.9)	(13.6)	(14.2)	(15.0)	(15.5)	(16.0)
Next-Nearest-Neighbor Corrections									
NN1	-0.36	-1.79	0.57	0.62	0.52	0.42	0.32	0.31	0.28
NN2	-1.30	-4.58	-0.53	0.22	0.68	0.99	1.40	1.68	1.81
NN3	-1.99	-0.54	0.84	0.69	0.61	0.55	0.51	0.38	-0.12
NN4	0.38	0.19	0.12	0.02	0.04	0.04	0.05	-0.04	-0.81
NN5	-0.75	1.38	-0.03	-0.06	-0.16	-1.37	-1.63	-1.76	-1.94
NN6	-1.28	0.17	0.13	-0.22	-0.22	-0.16	0.01	-0.01	-0.39
NN7	-2.33	-2.14	1.41	0.72	0.48	0.44	0.38	0.53	0.20

TABLE 5: Continued

group	$\Delta H_f^{298K}$ (kcal mol <sup>-1</sup> )	$S^{298K}$ (cal (mol-K) <sup>-1</sup> )	$C_p^T$ (cal (mol-K) <sup>-1</sup> )						
			300 K	400 K	500 K	600 K	800 K	1000 K	1500 K
NN8	-1.37	-0.92	0.64	0.26	0.43	0.67	0.69	0.72	0.23
NN9	0.54	-0.31	0.24	0.26	0.40	0.54	0.75	0.72	0.45
NN10	0.98	2.92	-1.23	-0.87	-0.58	-0.47	-0.51	-0.56	-0.40
NN11	-0.81	-3.11	4.36	3.97	3.24	2.58	1.68	2.20	2.69
NN12	1.16	-3.31	3.10	3.57	3.49	3.19	2.47	1.87	1.08
NN13	3.23	-1.15	1.22	1.06	0.70	0.48	0.36	0.42	0.61
NN14	2.31	1.37	0.64	-0.69	-1.86	-2.44	-2.37	-1.80	-0.83
NN15	0.32	-1.63	0.42	0.20	0.04	-0.03	0.00	0.07	0.17
NN16	1.70	0.82	-0.06	-1.02	-1.09	-1.09	-1.20	-1.05	-1.29
NN17	-1.46	-1.41	-0.13	-0.25	-0.26	-0.19	0.14	0.32	0.01
NN18	-2.42	-4.41	0.76	0.30	0.37	0.68	1.15	1.56	1.11
NN19	1.76	1.21	-1.02	-1.42	-1.49	-1.39	-1.08	-0.80	-1.43
NN20	0.79	2.07	-1.02	-1.33	-1.59	-2.24	-2.23	-2.22	-2.61
NN21	-0.33	1.18	-0.51	-0.18	-0.20	-0.32	-0.34	-0.48	-1.45
Secondary Corrections									
SC(a)	-3.32	-3.22	-0.79	-0.17	0.52	1.08	1.66	1.75	1.19
SC(b)	-6.54	-3.32	-2.02	-1.17	-0.94	-0.60	0.39	1.23	2.92
SC(c)	-3.67	-5.53	-1.57	-0.59	0.65	1.79	2.99	3.12	2.11
SC(d)	0.37	4.47	-1.17	-1.57	-1.85	-2.03	-2.18	-2.03	-1.42
SC(e)	-2.67	-5.56	-0.12	1.19	2.33	3.09	3.52	3.16	2.00

<sup>a</sup> Hydrogen bond increment value presented by Sumathi and Green.<sup>15</sup> <sup>b</sup> Hydrogen bond increment value presented by da Silva et al.<sup>14</sup> <sup>c</sup> Group additivity value obtained from one molecule. <sup>d</sup> Group additivity value presented by Sumathi and Green<sup>15</sup> shown in parentheses and italics. <sup>e</sup> Group additivity value presented by Ashcraft and Green<sup>53</sup> shown in parentheses and italics. <sup>f</sup> Group additivity value presented by Sebban and Bozzelli<sup>19</sup> shown in parentheses and italics. <sup>g</sup> Group additivity value presented by da Silva et al.<sup>14</sup> shown in parentheses and italics.

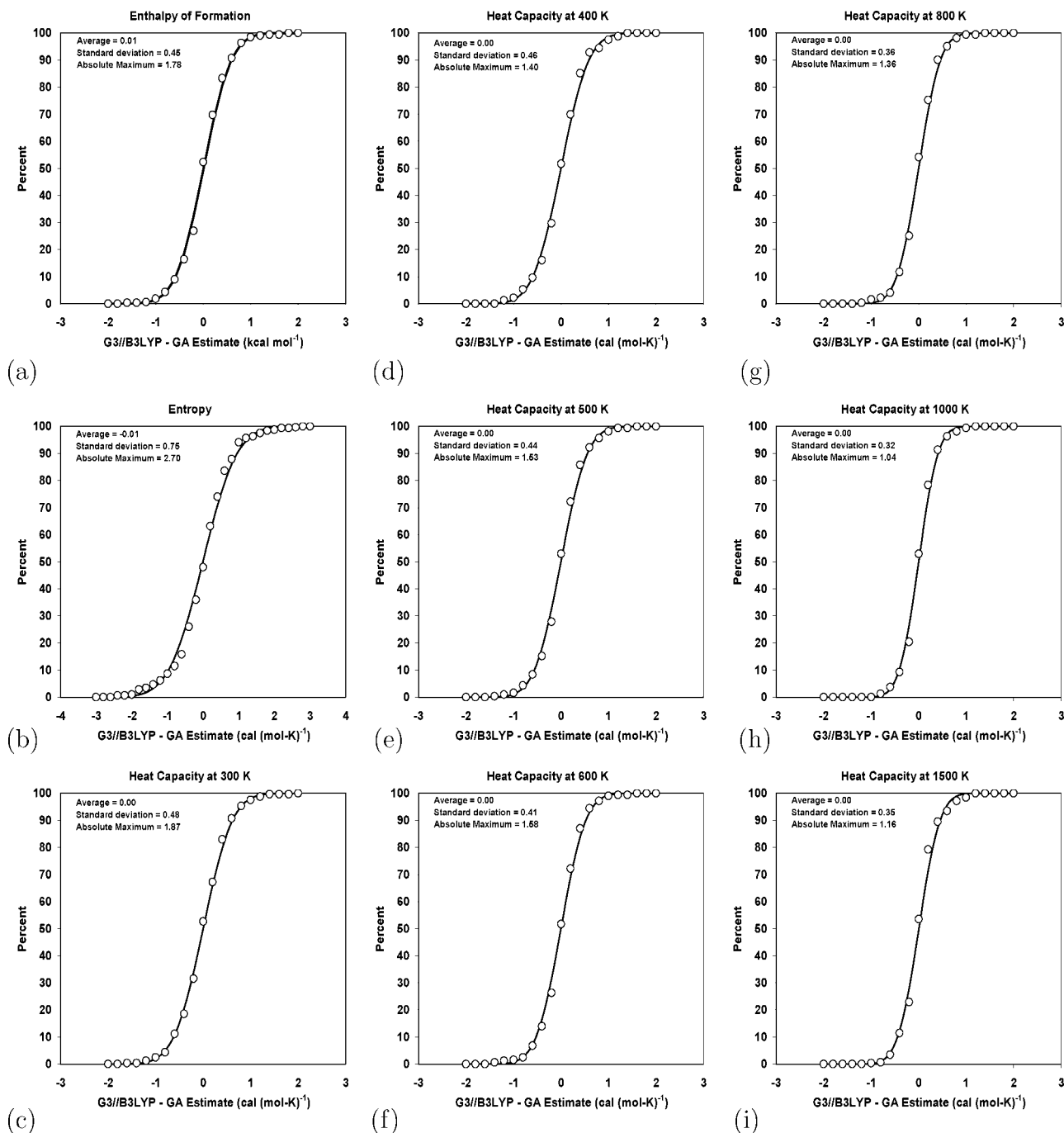
A subset of the molecules investigated had experimental or estimated quantum chemical data for the enthalpy of formation at 298 K available to which the G3//B3LYP values could be compared, as summarized in Table 3. Most molecules have errors within 2 kcal mol<sup>-1</sup> or the uncertainty indicated by the experimental data source, but in some cases, the error was higher. In particular, the values did not agree well for the 2-hydroxy-1-propyl radical (molecule 26) and 1-propenol (molecule 312). However, the 2-methoxy-1-propyl radical (molecule 28) and hydroxyethene (molecule 311) had the same unknown groups as 26 and 312, respectively, and very good agreement between the experimental and the G3//B3LYP values was observed. Furthermore, the group additivity estimates obtained upon regression of the unknown groups agree very well with the G3//B3LYP values for all four molecules. It is also noteworthy that several molecules in Table 3 had more than one experimental value available, one of which was from Holmes et al.,<sup>52</sup> the same source from which the experimental value for molecule 26 was taken. For all of these molecules, the experimental values from Holmes et al.<sup>52</sup> agreed less well with the G3//B3LYP values than the alternate experimental values. Therefore, we believe that the G3//B3LYP value for molecule 26 is a better estimate of the heat of formation than the experimental value reported by Holmes et al.<sup>52</sup> Similarly, molecule 311 had several experimental values with a range of 4.1 kcal mol<sup>-1</sup>. The value that had the highest error compared to the G3//B3LYP value was obtained by Turecek, the same author who reported an experimental value for molecule 312. Given these observations, the G3//B3LYP values are all in strikingly good agreement with the reported experimental and quantum chemical values.

Once the thermodynamic properties were obtained for all 323 molecules, the group values for the enthalpy of formation, entropy, and  $C_p$  values at 300, 400, 500, 600, 800, 1000, and 1500 K were determined through regression. Table 4 summarizes the known group values and the values for the gauche interactions for alkanes and ethers that were fixed during

regression. The unknown group values were then fit using the MLR method, and insignificant groups were eliminated at the 80% confidence level. Although the 323 molecules were selected to cover all 122 unknown groups, the rank of the matrix composed of 323 molecules as the rows and 122 unknown groups as the columns was 119, revealing that three degrees of freedom had to be specified to obtain a unique solution. To resolve this, values for three groups were set equal to values of known or unknown groups. Specifically, CO-(C<sub>d</sub>)(O) was set equal to CO-(C)(O), O-(C<sub>d</sub>)(H) was set equal to O-(C)(H), and CO-(C<sub>d</sub>)(O) was set equal to CO-(C)(O) based on their chemical similarity.

Before regressing the group values, the gauche and cis interactions were identified for every molecule. This resulted in 57 gauche interactions and 24 cis interactions to give a total of 81 interactions. These interactions are also listed in Table S1 (Supporting Information) for each molecule. With the addition of these unknown group values as variables in the model, the rank was analyzed again, and it was found that the rank of the entire matrix of 203 group values with the three constraints noted above imposed was equal to 193. This required the introduction of 10 additional constraints. On the basis of the interactions with the closest chemical similarity to the known alkane and ether gauche corrections, the following constraints were imposed: NN31 and NN32 were set equal to the alkane gauche correction, and NN51, NN53, NN54, NN56, NN59, NN61, NN79, and NN81 were constrained to equal the ether gauche correction.

The group values were regressed, and on the basis of the *t* test, it was found that only 43 of the 81 gauche and cis interactions were statistically significant. At this point, it was found that the values fit from group additivity for certain molecules had atypically high errors when compared to the G3//B3LYP values. Scrutiny of the molecules revealed that interactions due to the close proximity of a hydrogen atom on a hydroxyl group and an oxygen atom of another hydroxyl, carbonyl, alkoxy radical, or nitrate group provided additional



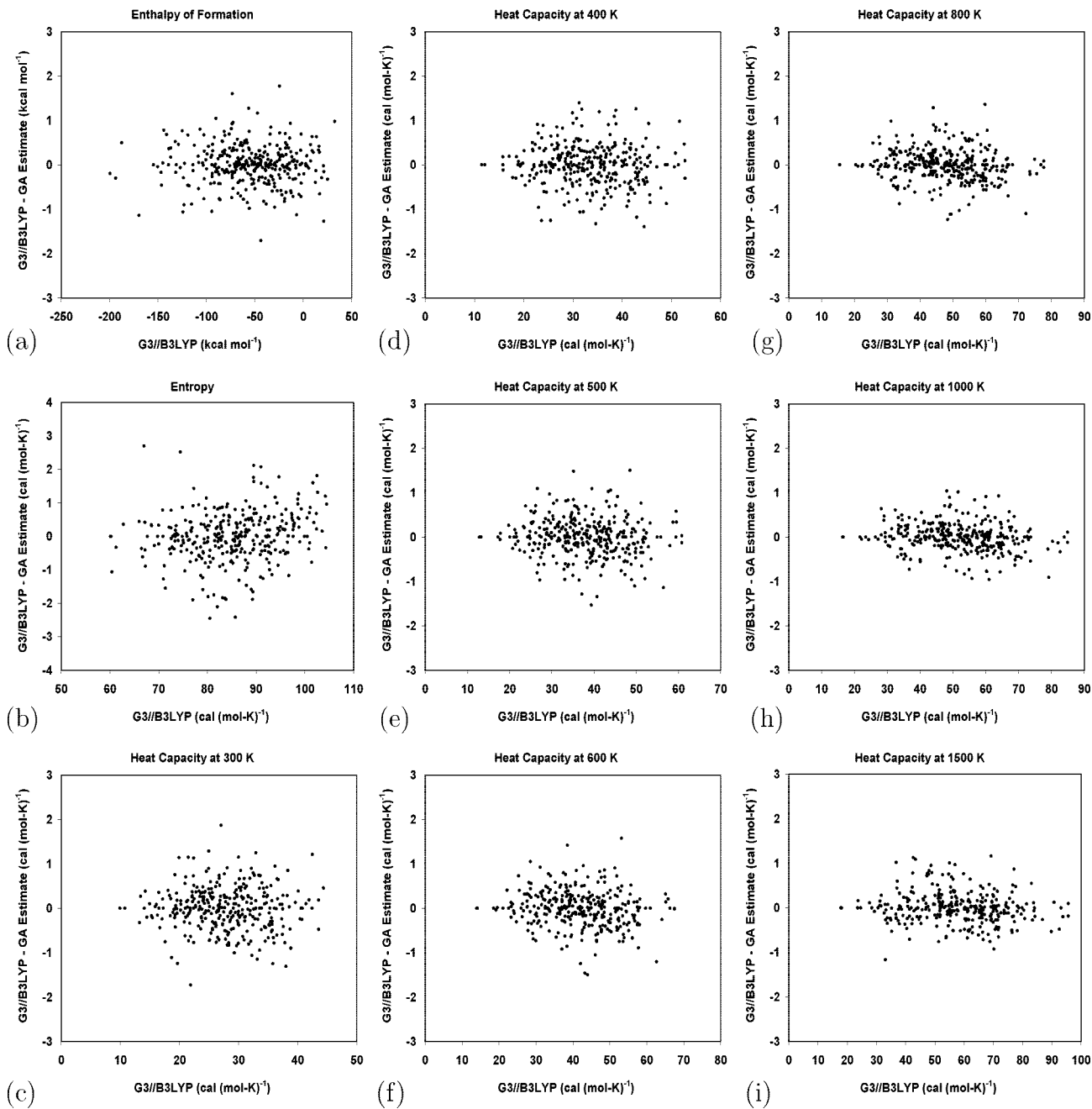
**Figure 2.** Cumulative percentage normal distribution plot of the residuals of (a) the enthalpies of formation at 298 K, (b) the entropy at 298 K, and (c–i) the heat capacity at 300, 400, 500, 600, 800, 1000, and 1500 K, respectively. The lines represent the normal distribution, and the dots represent the actual distribution of the residuals.

secondary interactions that were not taken into account in the group additivity scheme. These interactions are listed in Table 1 for a total of eight different secondary corrections. Similar corrections were introduced by Sumathi and Green<sup>20</sup> to account for eclipsing interactions of ketenes with the C=C=O group.<sup>20</sup> The addition of these eight secondary corrections resulted in 173 unknown groups. The rank of the associated matrix with no constraints imposed was 169, which therefore required four constraints, inclusion of the three constraints for the unknown groups used throughout and one gauche correction constraint, which was selected to set NN32 equal to the ether gauche correction as used previously. After regression, these corrections were also subjected to the *t* test at the 80% confidence level, and none of the eight corrections failed to disprove the null

hypothesis. However, two more of the gauche and cis corrections did now fail and were removed from the optimization.

At this point, a 98% confidence level *t* test was used to further test the remaining groups and corrections. It was found that using this more rigorous level of testing reduced the number of statistically significant corrections to 16 gauche corrections, 5 cis corrections, and 5 secondary corrections without significantly increasing the overall errors of the fit of the thermodynamic properties. These are listed first as NN1–NN21 and as SC(a)–SC(e) in Table 1. The *t* values at each optimization are also presented in the Supporting Information in Table S4. Thus, the final set of groups and corrections was 122 groups, 16 gauche, 5 cis, and 5 secondary corrections.





**Figure 3.** Plots of residuals versus the G3//B3LYP values of (a) the enthalpies of formation at 298 K, (b) the entropy at 298 K, and (c–i) the heat capacity at 300, 400, 500, 600, 800, 1000, and 1500 K, respectively.

Table 5 lists the optimized group values for all of the unknown groups. The first assumption of normality for MLR is graphically tested in Figure 2a–i for the enthalpy of formation, entropy, and heat capacity from 300 to 1500 K. The residuals, that is, the fitted thermodynamic value subtracted from the G3//B3LYP value, are plotted in the form of a normal probability plot, where the percent cumulative distribution of the residuals at each increment is plotted. The line shows the cumulative normal distribution, and the circles indicate the distribution for the actual model residuals. The figure clearly indicates that the residuals are normally distributed for all thermodynamic properties. The average, the standard deviation, and the maximum difference between the group additivity and the G3//B3LYP values for each property are shown in the inset of the figures. The second MLR assumption of constant variance is met as shown in Figure 3a–i for the enthalpy of formation, entropy, and heat capacity from 300 to 1500 K. The residuals

are plotted versus the G3//B3LYP value for the enthalpy of formation, entropy, and heat capacity. All graphs indicate that there is no significant difference in variance of the residuals over the range of fitted thermodynamic values. The error in the enthalpy of formation reaches a maximum value of 1.78 kcal mol<sup>-1</sup> for molecule 84 ( $\dot{C}(\text{CH}_3)(\text{HC}(\text{O}))(\text{C}_2\text{H}_5)$ ).

Shown in italics in Table 5 are group values that have been regressed in other studies. There are eight molecular groups, C–(C)<sub>2</sub>(CO)(H), C–(C)<sub>2</sub>(H)(NO), C–(C)(H)<sub>2</sub>(NO), C<sub>d</sub>–(C)(O), O–(C<sub>d</sub>)(H), O–(C<sub>d</sub>)(NO), O–(C<sub>d</sub>)(NO<sub>2</sub>), and O–(NO)(O), that have been regressed by Sumathi and Green,<sup>15</sup> Ashcraft and Green,<sup>53</sup> Sebbar and Bozzelli,<sup>19</sup> and da Silva et al.<sup>14</sup> Group values from their work were not used explicitly in this work but do offer a chance for comparison to our values. The values from our study compare very well with the values of these authors, thereby giving further validity to the group values obtained in this work. Two of the above studies<sup>14,15</sup> also report

radical hydrogen bond increment (HBI) values, which would enable the prediction of properties of radicals containing six of the radical group values in our study. While no direct comparison between our group values and their HBI values is possible, the groups for which HBI values are available in the literature are still noted in Table 5.

Finally, the sensitivity of the group values to the molecules included in the optimization was probed using a method similar to that employed by Mavrovouniotis, who developed a group contribution method to estimate standard Gibbs free energies of formation for biochemical compounds.<sup>54</sup> This analysis randomly eliminated 10% of the database used in optimizing the group values and compared the predictions of the Gibbs free energies for those molecules not used in optimization to the predictions obtained when they were used in optimization. In a similar manner, we randomly removed 10% of the database of molecules, or 32 molecules, and the group values were reoptimized. Although removal was random, it was ensured that the matrix of molecules and group values was still of full rank. The process was repeated four times with different sets of molecules removed during each optimization run. The difference between the heat of formation values calculated from the full optimization (323 molecules) and the heats of formation predicted when the optimization was carried out with different sets of 32 molecules eliminated was then examined for the 32 molecules that were removed. These differences are shown in the form of a histogram in Figure S1 (Supporting Information) for each of the four runs in which a different set of molecules was excluded from the optimization. In these histograms, the difference spans a range of  $\pm 1$  kcal mol<sup>-1</sup>. This is an indication that the use of this group additivity method is not compromised for molecules not included in the optimization.

## 5. Conclusion

The database of thermodynamic group values was expanded substantially by performing G3//B3LYP calculations on a set of molecules relevant to atmospheric chemistry. The thermodynamic properties were corrected for internal rotations. Values based on atomization energies, bond separation reactions, and isodesmic reactions agreed well, as the maximum mean absolute deviation among the methods for all molecules examined was 0.66 kcal mol<sup>-1</sup> for the enthalpies of formation. However, given that isodesmic reactions are known to improve estimates of the enthalpies of formation, the average of two different isodesmic reactions for each molecule was used as the recommended value. Group values were estimated using MLR, and statistically significant corrections were found by using *t* test analysis at the 80 and 98% confidence levels. The 122 new group values were regressed from the G3//B3LYP data, and 21 gauche and cis corrections and 5 secondary corrections were quantified, resulting in a total of 151 new values that can be included in existing group additivity databases. Enthalpies of formation, entropies, and heat capacities obtained from group additivity using the regressed values were in very good agreement with the values obtained from quantum chemical calculations. Furthermore, the group values obtained showed little sensitivity to the molecules not used in the optimization, as evidenced by the minimal changes in the calculated enthalpies of formation for the removed molecules. The group additivity values obtained in this study substantially augment the current database of groups for estimating thermodynamic properties.

**Acknowledgment.** This research was funded by U.S. EPA - Science to Achieve Results (STAR) Program Grant #R-

82816901-0 and the Inter-American Materials Collaboration Program of the National Science Foundation (DMR-0303435).

**Supporting Information Available:** The groups, corrections, and symmetry numbers for all molecules for which calculations have been performed are listed in Table S1. The isodesmic reactions corresponding to the molecules for which heats of formation have been calculated in this work are listed in Table S2. The estimated enthalpies of formation for all of the molecules from the atomization energy, bond separation reaction, and isodesmic reaction methods as well as the entropy and heat capacity as a function of temperature are shown in Table S3. The *t* values at each optimization of the group values are also presented in Table S4. Figure S1 has been provided to show the results from the brute-force statistical analysis. An Excel file with geometries and frequencies for every molecule studied and their hindered rotor information (symmetry number, reduced moment of inertia, and Fourier series coefficients) is also available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References and Notes

- (1) Benson, S. W. *Thermochemical Kinetics: Methods for the Estimation of Thermochemical Data and Rate Parameters*; John Wiley and Sons, Inc.: New York, 1968.
- (2) Eigenmann, H. K.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1973**, *77* (13), 1687–1691.
- (3) Lay, T. H.; Bozzelli, J. W. *J. Phys. Chem. A* **1997**, *101* (49), 9505–9510.
- (4) Zhu, L.; Chen, C.-J.; Bozzelli, J. W. *J. Phys. Chem. A* **2000**, *104* (40), 9197–9206.
- (5) Marsi, I.; Viskolcz, B.; Seres, L. *J. Phys. Chem. A* **2000**, *104* (19), 4497–4504.
- (6) Zhu, L.; Bozzelli, J. W. *J. Phys. Chem. A* **2002**, *106* (2), 345–355.
- (7) Mayer, P. M.; Parkinson, C. J.; Smith, D. M.; Radom, L. *J. Chem. Phys.* **1998**, *108* (2), 604–615.
- (8) Swihart, M. T.; Girshick, S. L. *J. Phys. Chem. B* **1999**, *103* (1), 64–76.
- (9) Wong, H.-W.; Nieto, J. C. A.; Swihart, M. T.; Broadbelt, L. J. *J. Phys. Chem. A* **2004**, *108* (5), 874–897.
- (10) Sumathi, R.; Green, W. H., Jr. *J. Phys. Chem. A* **2002**, *106* (46), 11141–11149.
- (11) Sabbe, M. K.; Saeys, M.; Reyniers, M.-F.; Marin, G. B.; Van Speybroeck, V.; Waroquier, M. *J. Phys. Chem. A* **2005**, *109* (33), 7466–7480.
- (12) da Silva, G.; Bozzelli, J. W. *J. Phys. Chem. A* **2006**, *110*, 13058–13067.
- (13) Zhu, L.; Bozzelli, J. W.; Kardos, L. M. *J. Phys. Chem. A* **2007**, *111*, 6361–6377.
- (14) da Silva, G.; Kim, C.-H.; Bozzelli, J. W. *J. Phys. Chem. A* **2006**, *110*, 7925–7934.
- (15) Sumathi, R.; Green, W. H., Jr. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3402–3417.
- (16) Menon, A. S.; Wood, G. P. F.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111* (51), 13638–13644.
- (17) Sabbe, M. K.; Vandeputte, A. G.; Reyniers, M.-F.; Van Speybroeck, V.; Waroquier, M.; Marin, G. B. *J. Phys. Chem. A* **2007**, *111* (34), 8416–8428.
- (18) Sabbe, M. K.; Reyniers, M.-F.; Van Speybroeck, V.; Waroquier, M.; Marin, G. B. *ChemPhysChem* **2008**, *9*, 124–140.
- (19) Sebban, N.; Bozzelli, J. W.; Bockhorn, H. *J. Phys. Chem. A* **2004**, *108* (40), 8353–8366.
- (20) Sumathi, R.; Green, W. H., Jr. *J. Phys. Chem. A* **2002**, *106* (34), 7937–7949.
- (21) Sabbe, M. K.; Vleschouwer, F. D.; Reyniers, M.-F.; Waroquier, M.; Marin, G. B. *J. Phys. Chem. A* **2008**, *112* (47), 12235–12251.
- (22) Carter, W. P. L. *Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment*, Technical Report; University of California: Riverside, CA, 2000; Final Report to California Air Resources Board.
- (23) Stockwell, W. R.; Kirchner, F.; Kuhn, M.; Seefeld, S. *J. Geophys. Res.* **1997**, *102*, 25847–25880.
- (24) Tonnesen, S.; Jeffries, H. E. *Atmos. Environ.* **1994**, *28* (7), 1339–1349.
- (25) Wong, H.-W.; Li, X. G.; Swihart, M. T.; Broadbelt, L. J. *J. Phys. Chem. A* **2004**, *108* (46), 10122–10132.

- (26) DeWitt, M. J.; Dooling, D. J.; Broadbelt, L. J. *Ind. Eng. Chem. Res.* **2000**, *39* (7), 2228–2237.
- (27) Pfaendtner, J.; Broadbelt, L. J. *Ind. Eng. Chem. Res.* **2008**, *47* (9), 2897–2904.
- (28) Aumont, B.; Szopa, S.; Madronich, S. *Atmos. Chem. Phys.* **2005**, *5*, 2497–2517.
- (29) Henry, D. J.; Radom, L. In *Quantum-Mechanical Prediction of Thermochemical Data*; Cioslowski, J., Ed.; Kluwer Academic Publishers: Norwell, MA, 2001; Vol. 22 of *Understanding Chemical Reactivity*, pp 161–197.
- (30) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109* (18), 7764–7776.
- (31) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110* (16), 7650–7657.
- (32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scusera, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E., Jr.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (33) Chen, C.-J.; Bozzelli, J. W. *J. Phys. Chem. A* **2000**, *104* (21), 4997–5012.
- (34) Sumathi, R.; Carstensen, H.-H.; Green, W. H., Jr. *J. Phys. Chem. A* **2001**, *105* (28), 6910–6925.
- (35) Sumathi, R.; Carstensen, H.-H.; Green, W. H., Jr. *J. Phys. Chem. A* **2001**, *105* (39), 8969–8984.
- (36) Sun, H.; Bozzelli, J. W. *J. Phys. Chem. A* **2002**, *106* (15), 3947–3956.
- (37) Chen, C.-J.; Bozzelli, J. W. *J. Phys. Chem. A* **2003**, *107* (22), 4531–4546.
- (38) Van Cauter, K.; Van Speybroeck, V.; Vansteenkiste, P.; Reyniers, M. F.; Waroquier, M. *ChemPhysChem* **2006**, *7*, 131.
- (39) Van Speybroeck, V.; Van Neck, D.; Waroquier, W. *J. Phys. Chem. A* **2000**, *104*, 10939.
- (40) Van Speybroeck, V.; Van Cauter, K.; Coussens, B.; Waroquier, M. *ChemPhysChem* **2005**, *6*, 189.
- (41) Van Speybroeck, V.; Van Neck, D.; Waroquier, M.; Wauters, S.; Saeys, M.; Martin, G. B. *J. Phys. Chem. A* **2001**, *105*, 7713.
- (42) Balint-Kurti, G. G.; Dixon, R. N.; Marston, C. C. *Int. Rev. Phys. Chem.* **1992**, *11* (2), 317.
- (43) East, A. L. L.; Radom, L. *J. Chem. Phys.* **1997**, *106* (16), 6655.
- (44) Yu, X.; Levine, S. E.; Broadbelt, L. J. *Macromolecules* **2008**, *41*, 8242.
- (45) Yu, X.; Pfaendtner, J.; Broadbelt, L. J. *J. Phys. Chem. A* **2008**, *112* (29), 6772–6782.
- (46) Pfaendtner, J.; Yu, X.; Broadbelt, L. J. *Theor. Chem. Acc.* **2007**, *118*, 881.
- (47) Pfaendtner, J.; Yu, X.; Broadbelt, L. J. *J. Phys. Chem. A* **2006**, *110*, 10863.
- (48) Pfaendtner, J.; Yu, X.; Broadbelt, L. J. calctherm, version 1.0; available for use free of charge by contacting Linda J. Broadbelt.
- (49) Chase, J. M. W. *J. Phys. Chem. Ref. Data, Monograph* **1998**, *9*, 1–1951.
- (50) Viskolcz, B.; Bérces, T. *Phys. Chem. Chem. Phys.* **2000**, *2* (23), 5430–5436.
- (51) Colson, A.-O.; Becker, D.; Eliezer, I.; Sevilla, M. D. *J. Phys. Chem. A* **1997**, *101* (47), 8935–8941.
- (52) Holmes, J. L.; Lossing, F. P.; Mayer, P. M. *J. Am. Chem. Soc.* **1991**, *113* (26), 9723–9728.
- (53) Ashcraft, R. W.; Green, W. H. *J. Phys. Chem. A* **2008**, *112* (38), 9144–9152.
- (54) Mavrouniotis, M. L. *Biotechnol. Bioeng.* **1990**, *36*, 1070–1082.
- (55) Pittam, D. A.; Pilcher, G. *J. Chem. Soc., Faraday Trans.* **1972**, *68*, 2224–2229.
- (56) Prosen, E. J.; Rossini, F. D. *J. Res. Natl. Bur. Stand. (U.S.)* **1945**, *263–267*.
- (57) Fletcher, R. A.; Pilcher, G. *Trans. Faraday Soc.* **1970**, *66*, 794–799.
- (58) Hine, J.; Arata, K. *Bull. Chem. Soc. Jpn.* **1976**, *49*, 3089–3092.
- (59) Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*; Hemisphere Publishing Corp.: New York, 1984.
- (60) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; Srinivasan, N. K.; Su, M.-C.; Sutherland, J. W.; Michael, J. V. *J. Phys. Chem. A* **2006**, *110* (21), 6592–6601.
- (61) Gurvich, L.; Veyts, I. V.; Alcock, C. B. *Thermodynamic Properties of Individual Substances*, 4th ed.; Hemisphere: New York, 1989; Vol. 2.
- (62) Dorofeeva, O. V.; Iorish, V. S.; Novikov, V. P.; Neumann, D. B. *J. Phys. Chem. Ref. Data* **2003**, *32* (2), 879–901.
- (63) Pedley, J. B. *Thermodynamic Data and Structures of Organic Compounds*, Vol. 1; Thermodynamics Research Center: The Texas A&M University System: College Station, TX, 1994.
- (64) Knobel, Y. K.; Miroshnichenko, E. A.; Lebedev, Y. A. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1971**, 425–428.
- (65) Tsang, W. *Heats of Formation of Organic Free Radicals by Kinetic Methods*; Blackie Academic and Professional: London, 1996; pp 22–58.
- (66) Pilgrim, J. S.; Taatjes, C. A. *J. Phys. Chem. A* **1997**, *101*, 8741.
- (67) Pilgrim, J. S.; Taatjes, C. A. *J. Phys. Chem. A* **1997**, *101*, 4172.
- (68) Ruscic, B.; Boggs, J. E.; Burcat, A.; Császár, A. G.; Demaison, J.; Janoschek, R.; Martin, J. M.; Morton, M. L.; Rossi, M. J.; Stanton, J. F.; Szalay, P. G.; Westmoreland, P. R.; Zabel, F.; Bérces, T. *J. Phys. Chem. Ref. Data* **2005**, *34* (2), 573.
- (69) Baulch, D. L.; Bowman, C. T.; Cobos, C. J.; Cox, R. A.; Just, T.; Kerr, J. A.; Pilling, M. J.; Stocker, D.; Troe, J.; Tsang, W.; Walker, R. W.; Warnatz, J. *J. Phys. Chem. Ref. Data* **2005**, *34* (3), 757.
- (70) Pedley, J. B.; Naylor, R. O.; Kirby, S. P. *Thermodynamic Data of Organic Compounds*, 2nd ed.; Chapman and Hall: London, 1986.
- (71) Wiberg, K. B.; Crocker, L. S.; Morgan, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 3447–3450.
- (72) Green, J. H. S. *Chem. Ind. (London)* **1960**, 1215–1216.
- (73) Pilcher, G.; Pell, A. S.; Coleman, D. J. *Trans. Faraday Soc.* **1964**, *60*, 499–505.
- (74) Khursan, S. L.; Martemyanov, V. S. *Russ. J. Phys. Chem. (Engl. Transl.)* **1991**, *65*, 321–325.
- (75) Ray, J. D.; Ogg, R. A., Jr. *J. Phys. Chem.* **1959**, *63*, 1522–1523.
- (76) Tsang, W. *J. Am. Chem. Soc.* **1985**, *107*, 2872–2880.
- (77) Burcat, A.; Ruscic, B. *Report TAE960*, Technical report; Technion Israel Institute of Technology; September 16, 2005.
- (78) Blanksby, S. J.; Ramond, T. M.; Davico, G. E.; Nimlos, M. R.; Kato, S.; Bierbaum, V. M.; Lineberger, W. C.; Ellison, G. B.; Okumura, M. *J. Am. Chem. Soc.* **2001**, *123* (39), 9585–9596.
- (79) Tsang, W. *Int. J. Chem. Kinet.* **1976**, *8* (2), 173–192.
- (80) Berkowitz, J.; Ellison, G. B.; Gutman, D. *J. Phys. Chem.* **1994**, *98* (11), 2744–2765.
- (81) Wentholt, P. G.; Squires, R. R. *J. Am. Chem. Soc.* **1994**, *116* (26), 11890–11897.
- (82) Finlayson-Pitts, B. J.; Pitts, J. N., Jr. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*; Academic Press: San Diego, CA, 1999.
- (83) Schatz, G. C.; Fitzcharles, M. S.; Harding, L. B. *Faraday Discuss. Chem. Soc.* **1987**, *84* (84), 359–369.
- (84) Turecek, F.; Havlas, V. *J. Org. Chem.* **1986**, *51*, 4066–4067.
- (85) Turecek, F. *J. Chem. Soc., Chem. Commun.* **1984**, *20*, 1374–1375.
- (86) Cohen, N. *J. Phys. Chem. Ref. Data* **1996**, *25* (6), 1411–1481.
- (87) *Selected Values of Properties of Chemical Compounds*; Thermodynamics Research Center: The Texas A&M University System: College Station, TX, 1997.