Thermochemical Properties and Group Values for Nitrogen-Containing Molecules

Robert W. Ashcraft and William H. Green*

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received: April 29, 2008; Revised Manuscript Received: July 10, 2008

Gas-phase thermochemical group additivity values were derived from CBS-QB3 computational chemistry calculations for 105 noncyclic C/H/O/N molecules. The molecules contain nitrile, nitro, nitroso, nitrite, nitrate, amine, imino, and azo functional groups. The enthalpy of formation, entropy, and heat capacity values for 49 atom-centered groups were derived. The effect of hindered internal rotations was included via rotor potential energy scans and solution of the one-dimensional Schrödinger equation. The average 95% confidence intervals across all derived groups are 1.4 kcal mol^{-1} for the enthalpy, 1.3 cal mol^{-1} K^{-1} for the entropy, and 1.0 cal/mol K for the heat capacity. The presented group values will be useful when employing automatic reaction mechanism generation tools to examine the role of fuel-bound or molecular nitrogen in energy-related or atmospheric processes.

Introduction

The desire to construct large, detailed chemical kinetic mechanisms for gas-phase processes has resulted in the development of several computational tools to automatically generate such mechanisms.[1−4] To construct these large mechanisms, thermochemical and rate information is needed to determine the importance of each potential species or reaction to the overall system; lack of this type of information for chemical species is often the most limiting aspect of these tools. The group additivity method, pioneered by Benson,[5−7] is one way an arbitrary molecule's thermochemical parameters can be estimated. Since Benson's original work, many researchers have derived group additivity values for unsubstituted, oxygenated, and halogenated hydrocarbons,[7−21] but much less effort has been put toward deriving nitrogen-containing groups.[22−24] Extension of these tools to include the detailed chemistry of gas-phase nitrogen species will allow new problems to be addressed, particularly as it relates to pollutants from combustion processes and atmospheric chemistry. The chemistry of fuel-bound nitrogen may also play a greater role as the use of fuel additives becomes increasingly necessary to meet both vehicle performance targets and emission regulations. Another potential application is in the modeling of the gasification, pyrolysis, and combustion of biomass feed stocks, which may have an inherently higher percentage of bound nitrogen than traditional fossil fuels. This work presents new nitrogen-containing group additivity values based on computational chemistry calculations

Methodology

The Gaussian03[25] suite of programs was used to perform the ab initio calculations. Gas-phase geometry optimization and frequency calculations were performed at the B3LYP/6-311G(2d,d,p) (B3LYP/CBSB7) level of theory as implemented in the CBS-QB3 method. The gas-phase enthalpies of formation are based on the energies derived using the CBS-QB3 method, which is a compound method that seeks to approximate the complete basis set limit.[26] They are estimated through the commonly used atomization method, which includes spin−orbit and bond additivity corrections (BACs) from Petersson et al.[27] In addition, BACs derived for N/O and N/N bond types from a previous work are included.[28] Results are also presented excluding the bond corrections to provide a more complete picture.

The entropy at 298 K and the constant-pressure heat capacity values at 300, 400, 500, 600, 800, 1000, and 1500 K were calculated using statistical mechanics, including the effect of hindered internal rotations. The effect of internal rotors was estimated using the method outlined by Sumathi et al.,[29] a brief description of which is given here. The hindered rotor potential was obtained by scanning the dihedral angle ϕ for 360° in increments of 30°, allowing for the relaxation of all other degrees of freedom. This potential energy scan was performed at the B3LYP/6-31G(d) level of theory, and the corresponding values of the potential, $V(\phi)$, were fit to a fifth-order Fourier series, eq 1, where a_0, a_m, and b_m are adjustable parameters.

$$V(\varphi)=\frac{a_0}{2}+\sum_{m=1}^{5}a_m\cos(m\varphi)+b_m\sin(m\varphi) \qquad (1)$$

$$\frac{-h^2}{8\pi^2 I}\frac{\partial^2\psi}{\partial\varphi^2}+V(\varphi)\psi=E\psi \qquad (2)$$

The one-dimensional Schrödinger equation, shown in eq 2, was solved in a sine/cosine basis; the basis size was increased until the computed partition function converged. The reduced moment of inertia was computed at the equilibrium geometry of the molecule, which is $I^{2,3}$ in East's notation.[30] The thermodynamic quantities of interest were estimated by a combination of the standard RRHO approximation methodology (translational, rotational, and vibrational contributions with the vibration corresponding to the hindered rotor removed) and the contributions from the hindered rotor as described here. The contributions of the hindered rotor to the heat capacity and entropy were calculated using ensemble average energies, $\langle E\rangle$ and $\langle E^2\rangle$, as appropriate.

The ability of amine groups to easily invert via the umbrella motion caused minor problems when attempting to calculate the rotational potential for some rotors. The detailed

---

* To whom correspondence should be addressed. Tel: +1-617-253-4580. Fax: +1-617-324-0066. E-mail: whgreen@mit.edu.

10.1021/jp803759m CCC: $40.75

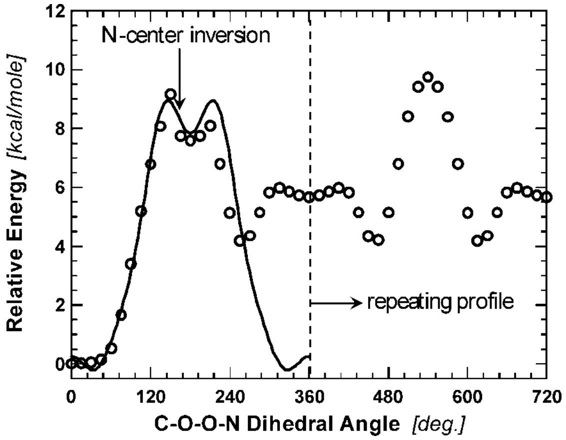

Figure 1. Calculated potential energy profile for the COON rotor in $CH_3OON(CH_3)_2$ for two full rotations (○) and the corrected Fourier series potential used when calculating rotor energy levels (solid line).

interactions of rotational/vibrational motions are very complex,[31] and we do not attempt such an undertaking here. During the course of the dihedral angle scan, an inversion typically results in a discontinuity in the rotor potential. The inversion, although favorable at a given dihedral angle, perturbs the molecule and prevents it from returning to the globally optimum energy at a dihedral angle of 360°. An example potential energy surface for two full rotations of such a rotor is shown in Figure 1, for the COON dihedral in $CH_3OON(CH_3)_2$. The ab initio potential is shown by the open circles, while the Fourier series used to calculate the energy levels is given by the solid line. It is clear that the inversion, which occurs near 150°, creates a situation in which the molecule cannot return to its globally minimum energy. Continued rotation does not reverse the inversion, and a stable, higher-energy rotation is obtained as shown between 360 and 720° in the figure. The differences in the peaks near 180 and 540° are due to a 60° difference in the $O-CH_3$ group rotor as well as the umbrella inversion. In cases where an inversion occurred, we either copied the values from 0−90° to the 120−210° and 240−330° regions for rotors that had a symmetry number of 3 or mirrored the values from 0−180° for rotors that had an energy peak near 180° (which were symmetric or near-symmetric rotors about $\phi = 180°$). The molecules that required corrections are noted in Table 1.

The group additivity values were determined through linear regression of the computed molecular data, using many of the previously established group values from Benson[7] and Bozzelli.[10,11,15] As is customary when deriving group values, the entropy used in the fitting procedure was the so-called intrinsic entropy, which ignores external and rotor symmetry and treats each atom in the molecule as distinguishable.[8] The intrinsic entropy is computed by adding $R\ln(\sigma)$ to the molecular entropy. Optical isomers and the potential umbrella motion of amine groups were not considered when reporting the molecular or intrinsic entropy values. The 95% confidence intervals for the new group value estimates were also computed to give a general indication of the uncertainty.

The confidence intervals were determined using a combination of Monte Carlo and traditional linear model confidence interval techniques. Monte Carlo was used in an attempt to propagate the error present in the CBS-QB3 estimates into to the final confidence intervals in the group values. The estimated molecular thermochemistry was assumed to have a Gaussian distribution with mean equal to the CBS-QB3 estimate and a standard deviation of 0.5 kcal mol^{-1} for enthalpy or 0.5 cal mol^{-1} K^{-1} for entropy and heat capacity. A value for each molecule was taken randomly from this distribution, and the vector of values was used in the regression algorithm, resulting in a unique set of best-fit group values and confidence intervals for each molecular property vector. This process was repeated 1000 times to yield a distribution of group values and confidence intervals. The confidence interval reported here was taken to be the mean value of the distribution of confidence intervals, attempting to capture some of the uncertainty in the molecular data.

As will be shown in the discussion, not all atom-centered groups could be uniquely determined because certain ones always appear together. The specific groups that were affected by this will be discussed later.

Results and Discussion

The thermochemical properties of 105 molecules were computed, which contain 66 groups to be determined and 14 groups with previously derived values. The computed structures for all of the molecules are included in the Supporting Information. Of the 66 unknown groups, 17 had to be assigned values or lumped in with other groups (equivalent to assigning a value of zero) to create a tractable linear regression problem.

Estimates of the thermochemical parameters for the molecules are given in Table 1 along with the limited experimental enthalpy data; not enough entropy or heat capacity data are available to make a meaningful comparison. The error (relative to the CBS-QB3 estimates) from the group additivity method is also presented with and without the bond corrections recommended in the literature.[27,28] The values without bond additivity corrections appear in parentheses. The bond-corrected enthalpy estimates agree significantly better with the limited experimental data than their uncorrected counterparts, giving credence to their inclusion. The mean absolute deviations relative to the 19 experimental data points were 0.7 and 1.7 kcal mol^{-1}, with and without bond corrections, respectively. The overall fit to enthalpies calculated using bond corrections was also better, with a mean absolute deviation of 0.40 kcal mol^{-1}, as compared to an MAD of 0.44 without corrections. The fitting errors in the entropy estimates generally ranged from −1.5 to +1.5 cal mol^{-1} K^{-1}, with a MAD of 0.37 cal mol^{-1} K^{-1}. Note that the entropy listed in the table is the molecular entropy, not the intrinsic entropy. The errors in the heat capacity data are not presented here for succinctness, but the average MAD across all temperatures was 0.21 cal mol^{-1} K^{-1}.

Derivation of new group values required the molecular data, known group values, and assignment of arbitrary values to certain unknown groups. The known and assigned group values used in this work are presented in Table 2, where 14 groups have been previously derived, three groups ($N/C_d/H_2$, $C/H_3/C{=}N$, and $N/H_2/O$) were assigned values by analogy to previously defined groups, three groups ($N{=}N/C$, $N{=}C/H$, and $N{=}C/C$) were assigned special values to keep them consistent with Benson's values, and 11 groups (CN/C, CN/C_d, CN/O, CN/N, NO/C, NO/O, NO/N, NO_2/C, NO_2/C_d, NO_2/O, and NO_2/N) were assigned zeros for each parameter. The $N{=}N/C$, $N{=}C/H$, and $N{=}C/C$ groups were all assigned Benson's values for the enthalpy of formation; however, no entropy and heat capacity values were available. The entropy and heat capacity data for $N{=}N/C$ were derived by taking the $N{=}N/H$ values and adding

TABLE 1: CBS-QB3 Molecular Thermochemistry Estimates with (without) Bond Corrections[a]

molecule		ΔH_f (298 K) (kcal mol^{-1})			S (298 K) (cal mol^{-1} K^{-1})		C_P values (cal mol^{-1} K^{-1})							symmetry
		CBS-QB3	expt.	GAV error	CBS-QB3	GAV error	300 K	400 K	500 K	600 K	800 K	1000 K	1500 K	σ
1	C_2H_3CN	45.5 (47.1)	43.0^{b}	−0.7 (−0.4)	65.3	−0.1	15.2	18.3	20.8	23.0	26.3	28.8	32.5	1
2	C_2H_3ONO	8.5 (7.1)		1.3 (1.8)	73.2	0.1	21.5	25.1	27.7	29.6	32.5	34.7	38.1	1
3	$C_2H_3ONO_2$	−3.6 (−6.0)		0.2 (0.5)	77.8	0.0	21.6	26.5	30.4	33.3	37.4	40.0	43.5	2
4	$CH_3CHC(CH_3)NO_2$	−9.3 (−9.9)		0.0 (−0.2)	87.2	−0.1	28.4	34.7	40.4	45.3	52.8	58.2	66.2	18
5	$CH_3CHCHNO_2$	−0.1 (−1.3)	2.4^{c}	0.6 (0.3)	79.1	0.0	23.2	28.3	32.8	36.6	42.3	46.4	52.2	6
6	$(C_2H_5)(CH_3)_2CNO$	−10.6 (−9.2)		−0.3 (−0.5)	94.7	0.6	36.7	44.1	50.7	56.4	65.5	72.4	83.2	27
7	$(C_2H_5)(CH_3)CHNO$	−1.1 (−0.2)		−0.4 (−0.7)	88.9	0.1	29.0	35.3	41.0	46.0	53.9	59.8	68.9	9
8	$(C_2H_5)(CH_3)NNO$	5.7 (7.9)		0.1 (−0.1)	84.9	0.6	27.6	33.5	39.0	43.7	51.2	56.7	65.2	9
9	C_2H_5CN	13.1 (15.1)	12.3^{b}	−0.2 (0.0)	68.2	−0.1	17.5	21.3	24.7	27.6	32.2	35.6	41.1	3
10	$(C_2H_5)(CH_3)NNO_2$	−10.2 (−9.1)		0.1 (−0.2)	90.2	0.5	29.7	36.5	42.7	48.0	56.2	62.2	71.0	18
11	$(C_2H_5)NHNO$	8.5 (10.7)		0.0 (−0.2)	78.7	1.0	21.3	26.3	30.9	34.8	40.8	45.1	51.4	3
12	$(C_2H_5)NHNO_2$	−8.3 (−7.3)		0.0 (−0.2)	81.6	0.3	24.8	30.5	35.5	39.5	45.6	50.0	56.5	6
13	C_2H_5NO	11.2 (11.1)		0.1 (0.4)	72.3	0.3	18.3	21.6	24.7	27.4	31.9	35.3	40.5	3
14	$C_2H_5NO_2$	−24.7 (−26.0)	$-24.5^{d,e}$	0.0 (0.2)	75.8	0.0	19.3	23.7	27.7	31.1	36.3	40.1	45.9	6
15	C_2H_5OCN	−8.6 (−7.6)		0.5 (0.2)	76.7	0.4	20.5	24.7	28.4	31.5	36.5	40.2	45.8	3
16	$C_2H_5ON(CH_3)_2$	−21.2 (−20.1)		0.2 (0.2)	88.4	0.6	31.5	39.3	46.5	52.6	62.1	69.0	79.4	27
17	$C_2H_5ONH(CH_3)$	−17.2 (−16.1)		0.1 (0.1)	81.8	−0.2	25.6	32.0	37.9	42.8	50.4	55.9	64.3	9
18	C_2H_5ONO	−22.8 (−23.7)		−0.2 (−0.2)	77.6	0.4	22.0	26.0	29.6	32.7	37.7	41.4	46.9	3
19	$C_2H_5ONO_2$	−37.3 (−39.3)	-37.0^{f}	0.0 (0.0)	81.5	0.4	23.6	28.7	33.2	36.9	42.6	46.5	52.0	6
20	C_2H_5OONO	−8.7 (−9.7)		0.2 (−0.1)	86.9	0.4	26.4	31.1	35.1	38.3	43.3	46.9	52.3	3
21	$(CH_3)_2CHNO$	3.0 (3.4)		0.4 (0.7)	79.7	−0.1	23.7	28.5	32.8	36.6	42.8	47.4	54.6	9
22	$(CH_3)_2CHNO_2$	−33.6 (−34.3)	$-33.2^{g,e}$	−0.3 (−0.1)	82.9	−0.4	25.3	31.2	36.4	40.8	47.6	52.6	60.2	18
23	$(CH_3)_2NCN$	33.5 (35.4)		−0.3 (0.0)	77.9	−0.4	20.4	24.7	28.9	32.7	38.7	43.2	50.1	9
24	$(CH_3)_2NNO$	12.7 (14.4)		−0.1 (0.1)	76.3	−0.6	21.2	25.7	29.9	33.7	39.6	43.9	50.6	9
25	$(CH_3)_2NON(CH_3)_2$	18.3 (19.4)		0.1 (0.1)	94.0	−1.1	37.0	45.9	53.8	60.4	70.6	77.9	88.9	162
26	$(CH_3)_2NONO$	17.7 (16.8)		−0.5 (−0.6)	84.5	0.1	26.8	32.2	36.8	40.8	46.8	51.2	57.5	9
27	$^{q}C_3H_7NO$	6.5 (6.9)		−0.1 (−0.4)	82.3	−0.3	23.1	28.2	32.8	36.8	43.1	47.7	54.8	3
28	$C_3H_7NO_2$	−29.8 (−30.5)	$-29.6^{d,e}$	0.0 (−0.2)	85.2	0.0	24.4	30.5	35.9	40.4	47.4	52.5	60.2	6
29	$(CH_3)_3CCN$	−1.6 (1.5)		−0.1 (0.2)	79.8	−0.2	29.8	36.8	42.7	47.6	55.3	60.9	69.9	81
30	$(CH_3)_3CNO$	−6.3 (−5.3)		0.3 (0.5)	84.2	−0.6	30.2	36.4	41.9	46.6	54.2	59.9	68.9	81
31	$(CH_3)_3CNO_2$	−42.6 (−42.8)	-42.3^{h}	−0.2 (0.1)	86.8	−0.4	31.6	39.1	45.6	51.0	59.2	65.2	74.5	162
32	CH_3CN	17.6 (19.2)	17.7^{i}	0.0 (0.0)	58.0	0.0	12.5	14.6	16.6	18.3	21.2	23.4	26.9	3
33	$CH_3NC(CH_3)_2{}^{q}$	1.3 (3.0)		−0.4 (−0.4)	80.5	−0.2	23.6	29.5	35.2	40.1	48.2	54.2	63.6	27
34	$CH_3NCCHCH_3$	40.4 (41.7)		−0.3 (−0.3)	78.1	−0.3	22.9	27.7	32.2	36.2	42.7	47.5	55.1	9
35	CH_3NCH_2	19.4 (20.1)		0.6 (0.6)	62.4	0.2	14.2	17.4	20.4	23.0	27.3	30.6	35.7	3
36	CH_3NCHCH_3	8.6 (9.8)		0.1 (0.1)	71.1	0.3	19.9	24.1	28.3	31.9	38.0	42.6	49.8	9
37	CH_3NHCN	34.4 (36.3)		−0.3 (0.0)	68.9	0.0	16.3	19.3	22.1	24.5	28.4	31.4	36.1	3
38	$CH_3NHN(CH_3)_2$	17.7 (21.4)		−0.3 (−0.3)	78.7	1.2	27.5	34.8	41.1	46.3	54.2	60.0	69.1	27
39	$CH_3NHN(CH_3)(C_2H_3)$	38.9 (42.6)		0.7 (0.9)	86.3	1.2	31.4	37.9	44.1	49.5	58.2	64.7	74.6	9
40	$CH_3NHNH(C_2H_3)$	41.0 (44.7)		−1.1 (−0.9)	78.3	0.9	26.0	32.1	37.2	41.5	48.1	52.9	60.4	3
41	CH_3NHNO	15.3 (16.9)		0.0 (0.2)	70.9	−1.0	17.1	20.0	23.0	25.5	29.6	32.6	36.9	3
42	$CH_3NHON(CH_3)_2$	22.6 (23.7)		−0.3 (−0.3)	86.9	0.0	31.5	38.8	45.4	51.0	59.7	66.1	75.4	27
43	CH_3NHONO	21.5 (20.5)		−0.3 (−0.4)	76.8	0.4	22.2	26.1	29.2	31.8	35.6	38.4	42.7	3
44	$CH_3NNC_2H_5$	29.5 (32.6)		0.0 (0.0)	80.0	0.0	24.4	29.6	34.5	38.8	45.8	51.0	59.2	9
45	CH_3NNCH_3	36.1 (38.6)	35.5^{c}	0.0 (0.0)	69.0	0.0	18.9	22.6	26.3	29.5	34.8	38.8	45.0	18
46	CH_3OCN	0.5 (1.1)		−0.5 (−0.2)	67.7	−0.4	15.1	17.5	19.9	21.9	25.3	27.7	31.6	3
47	$CH_3OON(CH_3)_2{}^{q}$	−4.3 (−3.7)		0.4 (0.6)	85.9	1.5	31.1	37.8	44.1	49.7	58.2	63.9	71.6	27
48	$CH_3OONH(CH_3)$	−0.2 (0.3)		0.3 (0.4)	81.3	−1.3	24.6	29.7	34.5	38.7	45.2	49.7	56.1	9
49	CH_3OONO	−0.2 (−1.6)		−0.2 (0.1)	77.9	−0.4	20.8	23.9	26.5	28.7	32.0	34.4	38.0	3
50	$HNCCH(CH_3)$	40.4 (41.7)		0.3 (0.3)	69.0	0.3	17.8	21.6	24.9	27.8	32.3	35.7	41.1	3
51	$HNCH(CH_3)$	10.4 (11.6)		−1.1 (−1.1)	63.0	−0.2	14.1	17.2	20.2	22.9	27.1	30.4	35.5	3
52	$NH_2N(CH_3)_2$	17.7 (21.4)		0.3 (0.3)	72.3	−1.2	21.8	27.5	32.4	36.5	42.8	47.3	54.5	9
53	$CH_2C(CH_3)(CN)$	36.4 (38.6)		0.1 (0.4)	73.4	0.3	20.9	25.2	28.9	32.1	37.1	40.8	46.6	3
54	$CH_2C(CH_3)(NO_2)$	−1.3 (−2.5)		0.0 (0.2)	78.4	0.1	23.6	28.9	33.4	37.0	42.6	46.4	52.1	6
55	$C_2H_3N(CH_3)(NH_2)^{q}$	39.1 (42.9)		1.1 (1.2)	80.3	−1.5	25.6	30.9	35.9	40.2	47.1	52.2	59.9	3
56	$C_2H_3NH(NH_2)$	39.4 (43.1)		1.1 (1.3)	71.6	−1.2	19.7	24.6	28.9	32.4	37.6	41.2	46.5	1

TABLE 1: Countinued

molecule		ΔH_f (298 K) (kcal mol^{-1})			S (298 K) (cal mol^{-1} K^{-1})		C_P values (cal mol^{-1} K^{-1})							symmetry
		CBS-QB3	expt.	GAV error	CBS-QB3	GAV error	300 K	400 K	500 K	600 K	800 K	1000 K	1500 K	σ
57	$C_2H_3NO_2$	8.9 (7.2)	8.0[c]	−0.6 (−0.3)	70.6	0.0	17.6	21.6	25.0	27.8	31.7	34.3	38.1	2
58	$(C_2H_5)(CH_3)_2CCN$	−6.7 (−3.1)		0.1 (−0.2)	90.9	0.2	35.6	44.0	51.2	57.1	66.5	73.4	84.2	27
59	$(C_2H_5)(CH_3)_2CNO_2$	−47.8 (−47.5)		0.2 (−0.1)	97.5	0.4	38.1	46.9	54.5	60.9	70.7	77.9	88.9	54
60	$(C_2H_5)(CH_3)CHNO_2$	−39.2 (−39.4)	−39.1[e]	0.3 (0.1)	91.6	0.4	32.2	39.2	45.5	50.9	59.2	65.2	74.6	18
61	$(C_2H_5)(CH_3)NCN$[q]	26.2 (28.6)		0.3 (0.0)	86.9	0.4	26.8	32.6	38.0	42.7	50.2	55.8	64.6	9
62	$C_2H_5NC(CH_3)_2$[q]	−6.1 (−3.9)		−0.1 (−0.1)	88.9	0.0	29.3	36.9	43.9	50.1	59.8	67.1	78.3	27
63	$C_2H_5NCH_2$	12.8 (14.0)		0.1 (0.1)	71.4	0.0	19.8	24.7	29.1	32.8	38.7	43.1	50.0	3
64	C_2H_5NHCN	27.1 (29.6)		0.3 (0.0)	78.7	0.0	21.9	26.6	30.7	34.2	39.7	43.9	50.4	3
65	$C_2H_5NNC_2H_5$	23.0 (26.5)		0.0 (0.0)	88.2	0.0	29.9	36.6	42.8	48.2	56.8	63.3	73.4	18
66	C_3H_7CN	7.7 (10.3)	7.5[j,k]	0.2 (0.0)	77.4	0.1	22.9	28.2	33.0	37.0	43.3	48.0	55.4	3
67	C_3H_7ONO	−27.9 (−28.3)	−28.4[k]	−0.1 (−0.6)	87.2	0.3	27.0	32.6	37.6	42.0	48.7	53.7	61.2	3
68	$C_3H_7ONO_2$	−42.3 (−43.8)	−41.6[l,m]	0.1 (−0.4)	91.6	−0.2	28.8	35.4	41.3	46.2	53.6	58.8	66.3	6
69	$CH_2C(CH_3)(ONO)$[q]	−0.4 (−1.2)		−0.4 (−0.4)	82.7	0.3	25.7	30.6	34.6	37.9	43.0	46.7	52.1	3
70	$CH_2C(CH_3)(ONO_2)$	−13.7 (−15.7)		−0.3 (−0.5)	86.9	0.6	28.1	33.9	38.7	42.5	48.0	51.8	57.4	6
71	$(CH_3)CHC(CH_3)(CN)$	28.8 (31.4)		−0.1 (−0.4)	82.6	−0.3	25.3	30.8	35.9	40.3	47.3	52.5	60.6	9
72	$(CH_3)CHC(CH_3)(ONO)$	−7.3 (−7.6)		−1.4 (−1.9)	92.2	−0.6	30.4	36.0	41.1	45.6	52.8	58.0	65.9	9
73	$(CH_3)CHCHCN$	36.3 (38.4)		0.7 (0.4)	73.7	0.1	20.8	25.0	28.7	31.9	37.0	40.8	46.6	3
74	$(CH_3)CHCHN(CH_3)(NH_2)$[q]	32.8 (37.0)		−0.5 (−0.8)	87.0	0.4	31.5	38.7	45.2	50.7	59.2	65.3	74.6	9
75	$(CH_3)CHCHNH(NH_2)$	33.5 (37.7)		−0.9 (−1.2)	78.7	0.3	26.0	32.1	37.5	42.0	48.8	53.5	60.6	3
76	$(CH_3)CHCHONO$	1.4 (0.6)		0.5 (0.5)	81.7	0.2	26.7	31.3	34.9	38.0	42.9	46.6	52.3	3
77	$(CH_3)CHCHONO_2$	−11.4 (−13.3)		0.1 (−0.1)	87.0	−0.6	27.2	33.0	37.9	41.9	47.9	51.9	57.5	6
78	CH_3ONO	−15.1 (−16.6)	−15.6[n]	0.3 (0.8)	68.9	−0.7	16.9	19.4	21.6	23.6	26.6	29.0	32.4	3
79	CH_3ONO_2	−29.0 (−31.6)	−29.2[o]	−0.1 (0.5)	72.2	−0.1	18.2	21.7	24.7	27.3	31.2	33.9	37.5	6
80	CH_3NHNO_2	−1.6 (−1.0)		0.0 (0.2)	72.4	−0.3	19.0	23.1	26.8	29.8	34.3	37.5	42.2	6
81	$(CH_3)_2NNO_2$	−3.3 (−2.7)	−1.2[c]	−0.1 (0.2)	81.3	−0.5	23.3	28.6	33.6	37.9	44.6	49.3	56.3	18
82	$CH_2C(CH_3)N(CH_3)(NH_2)$[q]	30.7 (34.9)		0.9 (1.0)	85.9	0.0	31.9	39.3	45.7	51.1	59.4	65.4	74.6	9
83	$CH_2C(CH_3)NH(NH_2)$	29.5 (33.7)		2.3 (2.5)	78.6	−1.1	24.6	31.0	36.7	41.4	48.5	53.5	60.8	3
84	$CH_3CHC(CH_3)N(CH_3)(NH_2)$[q]	26.0 (30.8)		−2.4 (−2.7)	95.1	−0.6	36.2	44.6	52.4	59.0	69.3	76.8	88.4	27
85	$CH_3CHC(CH_3)NH(NH_2)$	25.0 (29.7)		−1.1 (−1.4)	86.1	0.0	30.0	37.9	45.2	51.2	59.9	66.0	74.9	9
86	$CH_3NC(CH_3)C(CH_3)_3$	−17.5 (−14.2)		−0.8 (−0.8)	101.0	0.1	41.5	52.0	61.3	69.3	82.0	91.5	106.4	729
87	$CH_3NC(CH_3)CH_2(CH_3)$[q]	−3.9 (−1.7)		−0.1 (−0.1)	90.5	−0.1	29.3	36.5	43.3	49.4	59.1	66.4	77.8	27
88	$CH_3NC(CH_3)CH(CH_3)_2$	−10.4 (−7.6)		−0.1 (−0.1)	97.3	0.1	34.8	43.9	52.2	59.3	70.6	79.0	92.1	81
89	$CH_3NCHC(CH_3)_3$	−11.3 (−8.6)		0.8 (0.8)	92.2	−0.1	37.3	46.2	54.2	61.0	71.8	79.9	92.6	243
90	$CH_3NCHCH_2(CH_3)$	3.8 (5.5)		0.1 (0.1)	81.4	0.1	25.2	30.9	36.3	41.0	48.8	54.7	63.9	9
91	$CH_3NCHCH(CH_3)_2$	−2.9 (−0.6)		0.1 (0.1)	88.5	−0.1	31.1	38.4	45.1	50.9	60.2	67.2	78.2	27
92	$CH_3NHN(CH_3)C(CH_3)CH_2$[q]	30.7 (35.0)		0.3 (0.4)	94.1	0.5	37.5	45.1	52.3	58.6	68.8	76.6	88.6	27
93	$CH_3NHNHC(CH_3)CH_2$	31.7 (35.9)		−0.4 (−0.3)	85.1	1.1	32.1	39.4	45.8	51.1	59.3	65.4	74.8	9
94	$HNC(CH_3)_2$	0.2 (1.8)		1.5 (1.5)	71.7	0.1	19.4	23.9	28.1	31.8	37.8	42.4	49.5	9
95	$HNCH_2$	21.5 (22.1)	21.8[p]	−0.8 (−0.8)	54.3	−0.2	9.1	10.7	12.3	13.9	16.4	18.3	21.4	1
96	$NH_2OC_2H_5$	−15.5 (−14.5)		−0.4 (−0.3)	74.7	−0.3	21.0	25.7	30.0	33.6	39.1	43.2	49.5	3
97	NH_2ONHCH_3	27.2 (28.2)		0.2 (0.3)	71.2	1.1	21.8	26.9	30.8	33.8	38.1	41.2	45.9	3
98	NH_2ONO	21.4 (20.4)		0.9 (1.0)	70.5	−0.5	15.9	18.4	20.4	22.1	24.4	25.9	28.1	1
99	$NH_2OOC_2H_5$	−6.3 (−5.3)		−0.7 (−1.0)	82.8	−0.2	24.4	30.0	35.0	39.1	45.3	49.5	55.7	3
100	$CH_2C(CH_3)N(CH_3)_2$[q]	6.1 (8.1)		−0.6 (−0.6)	86.8	0.5	30.5	38.3	45.3	51.4	60.9	67.9	78.7	27
101	$CH_2C(CH_3)NH_2$	4.1 (6.0)		0.5 (0.5)	71.8	−0.3	20.9	25.8	30.1	33.7	39.3	43.5	50.2	3
102	$CH_2C(CH_3)NH(CH_3)$	4.5 (6.4)		0.5 (0.5)	78.7	−0.1	26.5	32.9	38.4	43.1	50.5	56.0	64.6	9
103	$CH_2CHN(CH_3)_2$[q]	13.7 (15.1)		0.6 (0.6)	80.7	−0.5	24.7	30.5	36.1	41.1	49.1	55.1	64.3	9
104	CH_2CHNH_2	13.8 (15.2)		−0.5 (−0.5)	64.1	0.3	15.4	18.8	21.8	24.3	28.3	31.3	36.0	1
105	$CH_2CHNH(CH_3)$	14.2 (15.6)		−0.5 (−0.5)	71.6	0.1	22.2	26.6	30.4	33.8	39.3	43.5	50.2	3

[a] Group additivity value (GAV) errors are relative to the CBS-QB3 estimates. Values in parentheses are enthalpy estimates and errors without employing gas-phase BACs. Some gas-phase enthalpies were estimated from the liquid-phase enthalpy and heat of vaporization data. [b] Data from ref 32. [c] Ref 33. [d] Ref 34. [e] Ref 35. [f] Ref 36. [g] Ref 37. [h] Ref 38. [i] Ref 39. [j] Ref 40. [k] Ref 41. [l] Ref 42. [m] Ref 43. [n] Ref 44. [o] Ref 45. [p] Ref 46. [q] Molecules that had at least one rotor potential corrected due to inversion of an amine group.

TABLE 2: Known and Assigned Group Values Used in the Derivation[a]

			C_P values (cal mol^{-1} K^{-1})						
group	ΔH_f (298 K) (kcal mol^{-1})	S (298 K) (cal mol^{-1} K^{-1})	300 K	400 K	500 K	600 K	800 K	1000 K	1500 K
C/C/H3[b]	−10.0	30.4	6.2	7.8	9.4	10.8	13.0	14.8	17.6
C/C2/H2[b]	−5.0	9.4	5.5	7.0	8.3	9.4	11.1	12.3	14.3
C/N/H3	−10.1	30.4	6.2	7.8	9.4	10.8	13.0	14.8	17.6
C/N/C/H2	−6.6	9.8	5.3	6.9	8.3	9.4	11.1	12.3	[14.8]
C/C/H2/O [b]	−8.1	9.8	5.0	6.9	8.3	9.4	11.1	12.3	[14.7]
C/H3/O [b]	−10.1	30.4	6.2	7.8	9.4	10.8	13.0	14.8	17.6
C_d/C/H [b]	8.6	8.0	4.2	5.0	5.8	6.5	7.7	8.5	9.6
C_d/H2 [c]	6.3	27.6	5.1	6.4	7.5	8.5	10.1	11.3	13.1
C/C_d/H3 [b]	−10.0	30.4	6.2	7.8	9.4	10.8	13.0	14.8	17.6
C_d/H/O[b]	8.6	6.2	4.8	6.5	7.6	8.4	9.1	9.6	10.5
C_d/C/O [d]	8.2	−12.3	3.6	4.6	5.0	5.3	5.8	6.1	6.2
O/C/O [e,f]	−5.5	8.5	3.2	3.4	3.6	3.8	4.1	4.2	4.2
N/C/H/N	20.9	9.6	4.8	5.8	6.5	7.0	7.8	8.3	9.0
N/H2/N	11.4	29.1	6.1	7.4	8.4	9.3	10.5	11.5	13.2
N/C_d/H2 [g]	4.8	29.7	5.7	6.5	7.3	8.1	9.4	10.5	12.3
C/H3/C=N [h]	−10.2	30.4	6.2	7.8	9.4	10.8	13.0	14.8	17.6
N/H2/O [i]	11.4	29.1	6.1	7.4	8.4	9.3	10.5	11.5	13.2
N=N/C [j]	27.0	7.2	3.4	3.6	3.7	3.9	4.3	4.6	4.9
N=C/H [k]	16.3	13.3	3.0	3.5	3.9	4.3	5.0	5.5	6.4
N=C/C [l]	21.3	−6.3	2.0	2.2	2.2	2.3	2.5	2.7	2.9

[a] Data from Benson[7] unless otherwise noted. Values in brackets were estimated in the present work. Groups with assigned values of zero for every parameter: CN/C, CN/C_d, CN/O, CN/N, NO/C, NO/O, NO/N, NO_2/C, NO_2/C_d, NO_2/O, and NO_2/N. [b] Enthalpy values from ref 18. [c] From ref 11. [d] From ref 15. [e] Enthalpy from ref 10. [f] Entropy and C_P from ref 15. [g] Assigned Benson's values for N/C_B/H2. [h] Assigned values for C/C_d/H3. [i] Assigned values for N/H2/N. [j] Enthalpy, assigned Benson's value; entropy and C_P, N=N/H + (C=C/C/H − C=C/H2). [k] Enthalpy, assigned Benson's value; entropy and C_P, N=N/H − (C=O/C/H − C=C/C/H)/2. [l] Enthalpy, assigned Benson's value; entropy and C_P, N=C/H + (C=C/C/H − C=C/H2).

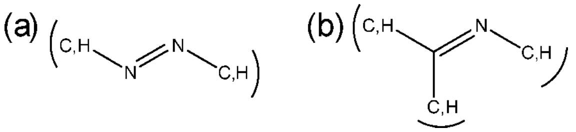

Figure 2. Generalized azo and imino functional groups.

the difference between the C=C/C/H and C=C/H2 groups in an attempt to mimic the effect of changing a neighbor from a hydrogen to a carbon. The N=C/R groups were more uncertain because Benson did not provide any entropy or heat capacity values. The N=C/H data were derived using the following relationship: N=N/H − (C=O/C/H − C=C/C/H)/2. We attempt to capture the effect of changing an N=N bond to an N=C bond by analogy to the C=C to C=O change. This is a very rough approximation and is meant to keep fitted values in line with Benson's[7] and Cohen's[18] previously published values. The values for N=C/C were assigned using the following relationship: N=C/H + (C=C/C/H − C=C/H2).

The reason that 17 groups were assigned values lies in the fact that they always appear with another subset of groups, resulting in an underspecified system. This is a well-known problem with group additivity methods.[5,6] A simple illustration of this is the N=N/C group, which will always appear with some member of the C/R/R′/R″/N=N group that is also present. The system will always be numerically rank-deficient because of linear combinations. More generally, any atom-centered group of the form R=N−R′ will be redundant and cannot be determined independently from the R-centered and R′-centered groups on either side, assuming that the R′ atom-centered group is defined as R′/N=R and not simply as R′/N. The choice to include the specific nitrogen double bond in the group definition of the neighbor makes defining the nitrogen-centered group unnecessary because no new information is included. In the general case, the thermochemical contribution of this section of the molecule can be arbitrarily divided up between the R-centered, N-centered, and R′-centered groups.

Consider the generalized azo and imino functional groups shown in Figure 2. For simplicity, consider only the case where the atoms on the edge of the group are either H or sp3 carbon. Enumeration of all combinations of atom-centered groups for the azo group results in 15 distinct combinations and six groups: N=N/H, N=N/C, C/H3/N=N, C/C/H2/N=N, C/C2/H/N=N, and C/C3/N=N. All other groups outside of the azo subgroup are assumed to be known and would not appear in the design matrix. It seems that 15 equations and six unknowns should make for a simple linear regression, but the absolute interdependence of some groups results in a rank of 5 (i.e., only five of the 15 equations are linearly independent). Therefore, an arbitrary value must be assigned to one group to create a tractable system, and this assignment can be made for any group except N=N/H. In this work, we chose to assign a value to N=N/C as discussed earlier. The imino functional group suffers from similar problems. In this case, the enumeration of possibilities, under similar assumptions as the azo example, results in 75 distinct combinations and 13 groups, with a numerical rank of only 10. Thus, three groups must be assigned values, and there are 129 acceptable, unique combinations that result in a tractable system of rank 10 (i.e., some combinations of three specified groups will still result in an underspecified system). Specifying more than one group out of each of the following sets will always result in a rank-deficient system: [C/H3/C=N, C/C/H2/C=N, C/C2/H/C=N, C/C3/C=N] and [C/H3/N=C, C/C/H2/N=C, C/C2/H/N=C, C/C3/N=C]. The complete set of feasible selections is complicated, and we do not attempt to present it here; however, selecting one group from each of the above sets and the final group from [C=N/H2, C=N/C/H, C=N/C2, N=C/H, N=C/C] will ensure that the system is of full rank. However, we chose to assign values to both N-centered groups and the C/H3/C=N group due to availability of Benson's data, which also resulted in a system of full rank. This type of analysis can also be carried out for other subsystems to determine the groups that need to be

TABLE 3: Group Additivity Values Derived from CBS-QB3 Molecular Data with Bond Corrections

			C_P values (cal mol^{-1} K^{-1})							confidence interval (95%)		
group	ΔH_f (298 K) (kcal mol^{-1})	S (298 K) (cal mol^{-1} K^{-1})	300 K	400 K	500 K	600 K	800 K	1000 K	1500 K	ΔH_f	S	C_P (avg)
C/H3/CN	17.7	60.2	12.5	14.6	16.6	18.3	21.2	23.4	26.9	1.9	1.7	1.3
C/C/H2/CN	22.9	39.8	11.3	13.5	15.3	16.8	19.2	20.9	23.5	1.3	1.2	1.0
[a]	22.5	40.2	11.1	13.4	15.5	17.2	19.7	21.3				
C/C3/CN	28.3	−3.0	11.4	13.4	14.6	15.3	16.3	16.7	17.2	1.3	1.2	1.0
[a]	29.0	−2.8										
C/C3/NO	24.1	1.2	12.2	13.3	14.0	14.5	15.3	15.7	16.2	1.3	1.2	1.0
C/C2/H/NO	23.4	23.1	11.2	12.7	14.0	15.1	16.8	17.9	19.5	1.3	1.2	1.0
C/C/H2/NO	21.4	44.3	11.8	13.6	15.2	16.7	18.9	20.5	23.0	1.3	1.2	1.0
C/C3/NO_2	−12.7	5.2	13.5	16.0	17.8	18.9	20.3	21.1	21.9	1.3	1.2	1.0
C/C2/H/NO_2	−13.9	27.5	13.6	16.1	18.1	19.6	21.8	23.2	25.1	1.3	1.2	1.0
[a]	−15.8	26.9										
C/C/H2/NO_2	−14.8	48.9	12.9	15.8	18.3	20.3	23.3	25.4	28.3	1.3	1.2	1.0
[a]	−15.1	48.4										
C/C/H2/C=N	−5.1	10.1	5.5	6.9	8.1	9.2	10.9	12.2	14.1	1.7	1.6	1.2
C/C2/H/C=N	−1.6	−11.2	5.0	6.5	7.5	8.2	9.3	9.9	10.9	1.7	1.6	1.2
C/C3/C=N	0.6	−33.5	5.3	6.6	7.3	7.5	7.8	7.8	7.7	1.7	1.6	1.2
C/H3/N=C	−5.7	30.4	6.0	7.7	9.3	10.7	13.1	14.8	17.7	1.3	1.2	1.0
C/C/H2/N=C	−2.9	8.6	5.3	7.2	8.7	9.8	11.6	12.8	14.7	1.7	1.6	1.2
C/H3/N=N	−9.0	30.2	6.0	7.8	9.4	10.8	13.1	14.8	17.6	0.8	0.8	0.6
C/C/H2/N=N	−5.5	9.4	5.3	6.9	8.3	9.4	11.1	12.3	14.2	0.8	0.8	0.6
[a]	−6.0											
C_d/C/N	3.5	−14.1	3.8	5.0	5.9	6.4	6.9	7.1	7.2	1.4	1.3	1.0
C_d/H/N	2.2	7.1	4.7	6.0	7.0	7.7	8.8	9.5	10.6	1.4	1.3	1.0
C_d/C/CN	40.2	17.9	9.2	10.6	11.7	12.5	13.8	14.7	15.9	1.3	1.2	1.0
C_d/H/CN	38.5	37.6	10.3	12.0	13.4	14.6	16.3	17.5	19.4	1.3	1.2	1.0
[a]	37.4	36.58	9.8	11.7	13.3	14.5	16.3	17.3				
C_d/C/NO_2	2.3	24.0	12.1	14.3	16.1	17.5	19.3	20.3	21.4	1.3	1.2	1.0
C_d/H/NO_2	2.0	44.3	12.7	15.4	17.6	19.3	21.7	23.1	25.0	1.3	1.2	1.0
[a]		44.4	12.3	15.1	17.4	19.2	21.6	23.2	25.3			
C=N/C2	5.7	2.0	3.5	4.2	5.0	5.6	6.6	7.2	7.9	1.2	1.1	0.9
C=N/C/H	3.3	21.2	5.5	6.3	7.2	8.0	9.3	10.2	11.6	1.3	1.2	0.9
C=N/H2	4.4	40.8	6.2	7.4	8.7	9.8	11.5	12.9	15.0	1.4	1.3	1.0
C_{dd}/C/N	25.9	19.7	4.5	5.1	5.6	6.0	6.5	6.9	7.4	1.5	1.4	1.1
O/C/CN	10.0	39.1	9.1	9.8	10.6	11.2	12.3	13.0	13.8	1.3	1.2	1.0
O/C/N	−9.2	7.2	3.5	3.6	4.0	4.3	4.7	4.8	4.2	1.3	1.2	1.0
O/C/NO	−4.8	40.0	10.6	11.3	11.9	12.6	13.6	14.3	14.8	1.1	1.0	0.8
[a]	−5.9	41.9	9.1	10.3	11.2	12	13.3	13.9	14.5			
O/C_d/NO	−5.3	39.5	11.1	11.7	12.2	12.7	13.5	14.1	14.9	0.9	0.9	0.7
O/O/NO	15.2	40.7	11.7	12.9	13.6	14.2	15.0	15.5	16.0	1.3	1.2	1.0
O/N/NO	10.8	40.8	10.2	11.5	12.4	13.0	13.9	14.3	14.8	1.3	1.2	1.0
O/C/NO_2	−19.1	45.3	12.2	13.9	15.4	16.6	18.4	19.3	19.9	1.1	1.0	0.8
[a]	−19.4	48.5										
O/C_d/NO_2	−18.4	45.4	12.4	14.2	15.7	16.9	18.5	19.3	20.1	1.1	1.0	0.8
O/O/N	5.3	6.9	4.3	4.9	5.6	6.3	7.0	7.1	6.5	1.3	1.2	1.0
O/N2	5.7	6.8	3.8	4.6	5.1	5.2	5.2	4.9	4.3	2.2	2.1	1.6
N/C/H/CN	44.1	40.7	10.3	11.6	12.8	13.9	15.5	16.7	18.3	1.3	1.2	1.0
N/C2/CN	53.3	21.0	8.6	9.6	10.5	11.4	12.9	13.8	14.8	1.3	1.2	1.0
N/C/H/O	20.4	8.1	5.2	6.2	7.0	7.7	8.7	9.4	10.5	1.4	1.3	1.0
N/C2/O	26.6	−12.7	4.3	5.1	5.7	6.2	7.0	7.3	7.5	1.2	1.1	0.8
N/C2/N	26.8	−14.5	3.7	4.9	5.8	6.3	6.8	6.8	6.7	1.3	1.2	1.0
[a]	29.2	−13.8										
N/C_d/C2	25.9	−11.0	2.8	2.9	3.3	3.7	4.6	5.0	5.5	1.9	1.7	1.3
N/C_d/C/H	15.3	8.7	5.8	6.1	6.4	6.7	7.5	8.1	9.1	1.9	1.7	1.3
[a]	15.4											
N/C_d/C/N	30.3	−13.2	4.2	4.2	4.4	4.8	5.4	5.7	6.0	1.5	1.4	1.1
[a]	30											
N/C_d/H/N	20.5	6.6	4.5	5.4	6.5	7.3	8.5	9.1	9.9	1.5	1.4	1.1
[a]	21.5											
N/C2/NO	32.6	19.3	9.4	10.5	11.5	12.4	13.8	14.6	15.3	1.3	1.2	1.0
N/C/H/NO	25.2	41.7	10.4	11.9	13.4	14.7	16.6	17.9	19.2	1.3	1.2	1.0
N/C/H/NO_2	8.4	45.3	13.1	15.5	17.6	19.2	21.4	22.8	24.4	1.3	1.2	1.0
N/C2/NO_2	16.7	25.8	11.5	13.4	15.2	16.7	18.8	20.0	21.1	1.3	1.2	1.0

[a] Values reported by Benson[7] are listed for comparison.

specified. The main point to take away is that one cannot independently determine all group values, and the choice of groups that are assigned values can be rather arbitrary.

The linear least-squares fitting procedure yielded estimates for 49 groups containing nitrogen. The values from fitting the bond-corrected molecular data for the enthalpy of

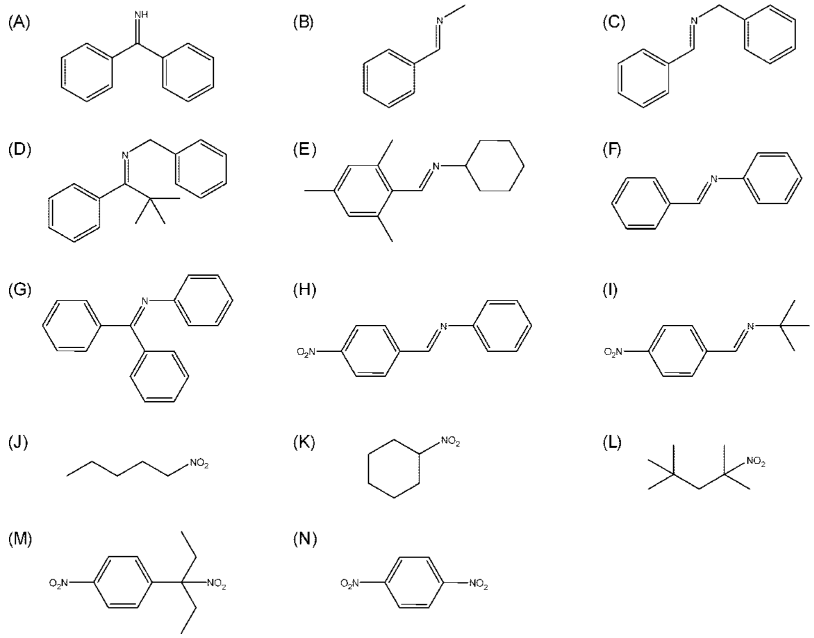

Figure 3. Set of molecules used to test group value estimates. Structures A−H are from ref 47, structure I is from ref 48, and structures J−N are from ref 24.

TABLE 4: Group Value Enthalpy Estimates at 298 K using Imperfect Group Assignments (kcal mol^{-1})[a]

molecule	GAV estimate	experiment	deviation
A	66.0	59.4 ± 0.3[a,b]	6.6
B	40.9	37.5 ± 0.2[a,b]	3.4
C	65.7	64.2 ± 0.3[a,b]	1.5
D	38.0	37.4 ± 0.8[a,b]	0.6
E	−6.3	−5.5 ± 0.5[a,b]	−0.8
F	62.6	66.8 ± 0.4[a,b]	−4.2
G	61.6	63.0 ± 0.4[a,b]	−1.4
H	87.0	89.0 ± 0.5[a,b]	−2.0
I	12.4	11.8 ± 0.9[a,c]	0.6
J	−39.8	−39.5 ± 0.4[d]	−0.3
K	−38.2	−38.1 ± 0.2[d]	−0.1
L	−67.8	−59.6 ± 0.3[d,e]	−8.2
M	−21.7	−21.8 ± 0.3[a,d]	0.1
N	17.8	13.3 ± 0.2[a,d]	4.5

[a] The precise groups needed to build the molecules were not available, requiring the use of similar groups. The group substitutions are discussed in the text. [b] Ref 47. [c] Ref 24. [d] Ref 48. [e] Molecular strain[24] of 8.3 kcal mol^{-1} not included in the estimate.

TABLE 5: Substitutions for Unavailable Groups

unavailable	substitute
C_b/C=N	C_b/C
C=N/C_b2	C=N/C2
C=N/C_b/H	C=N/C/H
C/C3/C_b	C/C4
C/C_b/H2/N=C	C/C/H2/N=C
C/C2/H/N=C	C/C2/H/N
C_b/N=C	C_b/N
N=C/C_b	N=C/C
C_b/NO_2	C_d/C/NO_2
C/C3/N=C	C/C3/N

formation are presented in Table 3. We chose not to provide group values for the uncorrected enthalpy data because of their relatively larger errors with respect to experimental data. Confidence limits (95%) in the parameter estimates are given for the enthalpy, entropy, and average for heat capacity across all temperatures. It is worth noting that only 105 data were used to determine 49 parameters; obviously, more data would be preferred to yield higher confidence in the parameter values. Despite this apparent small data/parameter ratio, the confidence intervals in the parameters are relatively tight, typically less than 1.5 kcal mol^{-1} for group enthalpy, 1.5 cal mol^{-1} K^{-1} for entropy, and 1.0 cal mol^{-1} K^{-1} for heat capacity. As mentioned previously, the confidence intervals reported in Table 3 are the results of a Monte Carlo approach and are larger than if the molecular data were assumed to be precisely correct.

Partial values for 13 of the groups were also available from Benson's previous work,[7] and these data are presented in Table 3 below the values derived here. Three groups had nearly complete data provided by Benson (C/C/H2/CN, C_d/H/CN, and O/C/NO); however, we chose not to treat these as known group values in the fitting procedure because we were fitting similar groups where full data were not available and wanted to keep the derived estimates more internally consistent. These also serve as a good metric to which to compare our fitted values. The enthalpy and heat capacity data derived here agree well with Benson's previous estimates. We do not hope to match exactly because the new values should be an improvement over the previously published values; however, it does serve as a good check to ensure that the fitted values and molecular data are reasonable. The MAD relative to Benson for the group enthalpy values was 0.84 kcal mol^{-1}, with a signed average error of 0.14 kcal mol^{-1}. Comparison of the group entropy values yields a MAD of 0.94 cal mol^{-1} K^{-1} and a signed error of −0.48 cal mol^{-1} K^{-1}. The heat capacity values agree very well. Extensive lists of previously derived group values not listed here are available from Cohen[18] and Benson,[7] among other sources cited previously.

There is a limited amount of experimental data available to test the group values derived in this work.[24,47,48] The molecules in the test set are shown in Figure 3 and represent imine and nitro type functional groups. Only gas-phase enthalpy of formation estimates were compared with experimental data. The main problem associated with these comparisons is the lack of correct groups needed to construct the molecules. Many of the groups adjacent to aromatic rings were not available, requiring substitute groups to be used. The exact groups used to construct each molecule and the associated group values are included in the Supporting Information. The group values used to calculate the enthalpies in Table 4 were taken from, in order of preference, the groups values derived here, Cohen's revised values,[18] and Benson's original values.[7] Table 4 lists the group additivity estimates for the enthalpy of formation at 298 K, the experimental values, and the deviation between the two. The average signed error for the data is 0.02 kcal mol^{-1}, and the mean absolute deviation is 2.45 kcal mol^{-1}. Much of this error is likely due to the imperfect group usage, requiring substitute groups to be used to make the estimates possible. Additionally, the uncertainties for group values derived in the present work are on the order of 1.5 kcal mol^{-1}, so the observed deviations are to be expected. Steric effects not captured in group values may also play a role. The group substitutes used in the calculations are shown in Table 5. Most molecules in the test set required at least one substitute group, and those that did not (J and K) have relatively small errors and fall within the experimental error bars. Molecule L is a unique case because it has a large strain enthalpy of 8.3 kcal mol^{-1},[24] and the large error is likely due to the group values not capturing this effect. If the strain enthalpy is included, then the group value estimate is in line with the experimental data. Overall, the predictions are as expected, with significant errors in many cases when substitute groups were used and generally good agreement when all groups were available for a given molecule.

Conclusion

The CBS-QB3 compound method was employed to calculate the enthalpy of formation, entropy, and heat capacity for a set of 105 nitrogen-containing molecules. The atomization procedure used to calculate the enthalpy of formation has been shown to be quite accurate by Petersson et al.[27] The entropy and heat capacity calculations employed traditional gas-phase statistical thermodynamics, including corrections for hindered internal rotations. The 49 derived, and 17 assigned, group additivity values appear reasonable and fit the calculated molecular data well. The uncertainty of the group values was estimated using a Monte Carlo method to propagate errors in the molecular data into the fit. The overall uncertainty is a combination of the error in the molecular data and fitting procedure and is difficult to estimate accurately, but the Monte Carlo method yields a reasonable first estimate that is more rigorous than assuming error-free CBS-QB3 thermochemistry estimates. These new group values will prove useful in group-based automatic mechanism generation tools when examining the role of nitrogen in high-temperature processes involving organics. Accurate experimental data for gas-phase species of the types examined here would allow for a more definitive estimate of the uncertainty and potentially allow for refined group value estimates.

Acknowledgment. Financial support is gratefully acknowledged from NSF Grant CHE-0535604, Collaborative Research in Process Informatics for Chemical Reaction Systems. This work is also supported by the Division of Chemical Sciences, Geosciences, and Biosciences, the Office of Basic Energy Sciences, the U.S. Department of Energy, in part through Grant DE-FG02-98ER14914.

Supporting Information Available: Group construction of the experimental test set molecules; 2D connectivity diagrams of the molecular structures and group definitions; B3LYP/6-311G(2d,d,p) level optimized geometries, vibrational frequencies, and rotational frequencies for all molecules examined in the study; and the group assignment matrix detailing the groups that make up each molecule. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Green, W. H. *Adv. Chem. Eng.* **2007**, *32*, 1.
(2) Broadbelt, L. J.; Stark, S. M.; Klein, M. T. *Ind. Eng. Chem. Res.* **1994**, *33*, 790.
(3) Warnatz, J.; Muharam, Y. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4218.
(4) Blurock, E. S. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 607.
(5) Benson, S. W.; Buss, J. H. *J. Chem. Phys.* **1958**, *29*, 546.
(6) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. *Chem. Rev.* **1969**, *69*, 279.
(7) Benson, S. W. *Thermochemical Kinetics*, 2nd ed.; John Wiley & Sons Inc.: New York, 1976.
(8) Sumathi, R.; Green, W. H. *J. Phys. Chem. A* **2002**, *106*, 7937.
(9) Sumathi, R.; Green, W. H. *J. Phys. Chem. A* **2002**, *106*, 11141.
(10) Lay, T. H.; Bozzelli, J. W. *J. Phys. Chem. A* **1997**, *101*, 9505.
(11) Chen, C.-J.; Wong, D.; Bozzelli, J. W. *J. Phys. Chem. A* **1998**, *102*, 4551.
(12) Zhu, L.; Chen, C.-J.; Bozzelli, J. W. *J. Phys. Chem. A* **2000**, *104*, 9197.
(13) Sun, H.; Bozzelli, J. W. *J. Phys. Chem. A* **2002**, *106*, 3947.
(14) Chen, C.-C.; Bozzelli, J. W. *J. Phys. Chem. A* **2003**, *107*, 4531.
(15) Sebbar, N.; Bozzelli, J. W.; Bockhorn, H. *J. Phys. Chem. A* **2004**, *108*, 8353.
(16) Sebbar, N.; Bockhorn, H.; Bozzelli, J. W. *J. Phys. Chem. A* **2005**, *109*, 2233.
(17) Fishtik, I.; Datta, R. *J. Phys. Chem. A* **2003**, *107*, 6698.
(18) Cohen, N. *J. Phys. Chem. Ref. Data* **1996**, *25*, 1411.
(19) Yu, J.; Sumathi, R.; Green, W. H. *J. Phys. Chem. A* **2006**, *110*, 6971.
(20) Yu, J.; Sumathi, R.; Green, W. H. *J. Am. Chem. Soc.* **2004**, *126*, 12685.
(21) Salmon, A.; Dalmazzone, D. *J. Phys. Chem. Ref. Data* **2007**, *36*, 19.
(22) Verevkin, S. P. *J. Therm. Anal. Calorim.* **2000**, *60*, 437.
(23) Rakus, K.; Verevkin, S. P.; Beckhaus, H.-D.; Ruchardt, C. *Chem. Ber.* **1994**, *127*, 2225.
(24) Verevkin, S. P. *Thermochim. Acta* **1997**, *307*, 17.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02; Gaussian Inc.: Wallingford, CT, 2004.
(26) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532.
(27) Petersson, G. A.; Malick, D.; Wilson, W.; Ochterski, J. W., Jr.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 10570.
(28) Ashcraft, R. W.; Raman, S.; Green, W. H. *J. Phys. Chem. B* **2007**, *111*, 11968.
(29) Sumathi, R.; Carstensen, H.-H.; Green, W. H. *J. Phys. Chem.* **2001**, *105*, 6910.
(30) East, A. L. L.; Radom, L. *J. Chem. Phys.* **1997**, *106*, 6655.
(31) Johnson, P. M.; Sears, T. J. *J. Chem. Phys.* **1999**, *111*, 9222.
(32) Hall, H. K.; Baldt, J. H. *J. Am. Chem. Soc.* **1971**, *93*, 140.
(33) Burcat, A.; Ruscic, B. Ideal Gas Thermochemical Database with Updates from Active Thermochemical Tables; ftp://ftp.technion.ac.il/pub/supported/aetdd/thermodynamics, March 2008.
(34) Lebedeva, N. D.; Ryadenko, V. L. R. *Russ. J. Phys. Chem.* **1973**, *47*, 1382.
(35) Holcomb, D. E.; Dorsey, C. L. *Ind. Eng. Chem. Res.* **1949**, *41*, 2788.
(36) Gray, P.; Pratt, M. W. T.; Larkin, M. J. *J. Chem. Soc.* **1956**, 210.

(37) Cass, R. C.; Fletcher, S. E.; Mortimer, C. T.; Quincey, P. G.; Springall, H. D. *J. Chem. Soc.* **1958**, 958.

(38) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*; Chapman and Hall: New York, 1986.

(39) An, X.; Mansson, M. *J. Chem. Thermodyn.* **1983**, *15*, 287.

(40) Evans, F. W.; Skinner, H. A. *Trans. Faraday Soc.* **1959**, *55*, 255.

(41) Batt, L.; Christie, K.; Milne, R. T.; Summers, A. J. *Int. J. Chem. Kinet.* **1974**, *6*, 877.

(42) Fairbrother, D. M.; Skinner, H. A.; Evans, F. W. *Trans. Faraday Soc.* **1957**, *53*, 779.

(43) Gray, P.; Pratt, M. W. T. *J. Chem. Soc.* **1957**, 2163.

(44) Ray, J. D.; Gershon, A. A. *J. Phys. Chem.* **1962**, *66*, 1750.

(45) Ray, J. D.; Ogg, R. A. *J. Phys. Chem.* **1959**, *63*, 1522.

(46) Dean, A. M.; Bozzelli, J. W. Combustion chemistry of nitrogen. In *Gas Phase Combustion Chemistry*; Gardiner, W. C., Ed.; Springer: New York, 2000.

(47) Verevkin, S. P.; Morgenthaler, J.; Ruchardt, C. *J. Chem. Thermodyn.* **1997**, *29*, 1175.

(48) Acree, W. E.; Kirchner, J. J.; Tucker, S. A.; Pilcher, G.; Ribeiro da Silva, M. D. M. C. *J. Chem. Thermodyn.* **1989**, *21*, 443.

JP803759M