Correlations

Predictive Correlation for C_p of Organic Solids Based on Elemental Composition

V. Laštovka and J. M. Shaw*

Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta T6G 2G6, Canada

A correlation for the constant pressure heat capacity (C_p) valid from 50 K to the fusion temperature for pure organic solids is presented. The predictive correlation includes seven universal coefficients. The variables are temperature and a parameter α, which possesses a value proportional to the number of atoms in a molecule divided by molar mass. This parameter is shown to be more robust than molecular structure or composition as a descriptor for C_p. The training set comprises 72 organic solids with a wide range of molecular structures and elemental compositions including nitrogen-, oxygen-, and sulfur-substituted organic compounds. Solid−solid transition regions were excluded from the data set. The average absolute deviation for the comparison test set including 93 additional compounds with 2080 specific heat capacity values is less than 0.06 $J\cdot K^{-1}\cdot g^{-1}$. The principal applications for the correlation are estimation of heat capacities for large organic molecules, where data are normally unavailable, and for poorly defined mixed organic solids such as asphaltenes where elemental analysis but not molecular structures are available.

Introduction

There are a limited number of options for estimating the heat capacity of solid organic compounds at constant pressure. Their application is limited to narrow temperature ranges, and in most cases they are based on knowledge of molecular structure, or have only been designed for specific groups of organic compounds, or are only applicable to low molar mass compounds. For example, there is only one group contribution method available that can be used to estimate heat capacities for solids over a wide range of temperatures. It was developed by Goodman et al.[1] Other available group contribution methods are applicable only at 298 K. These include methods by Domalski and Hearing,[2,3] by Chickos et al.,[4] and by Helgeson et al.,[5] whose method was extended to heavy organic compounds by Richard and Hegleson.[6]

There is also a method based on the sum of elemental contributions[7] using a modified Kopp's rule applicable at 298 K. A low-temperature additive method for predicting specific heat capacities of alkanes, alkenes, alkanols, and alkanones from (10 to 150) K is also available.[8] This method has been modified and extended to alkyl[9] and phenyl[10] derivatives of urea.

In the petroleum industry, the physical and chemical properties such as heat capacities for large solid organic compounds in temperature ranges arising in production, transport, and refining are of growing interest in themselves. They are the focus of this contribution and are a means to detect phase transitions in the complex mixtures arising in, for example, bitumen or heavy oil segments of the industry, which will be the subject of a future contribution.

The correlation presented here is a modification of a four-parameter equation successfully used by Briard et al.[11] to correlate heat capacities of alkanes. This new predictive correlation is based on knowledge of elemental composition and can be applied to a broad range of compound classes from 50 K to the first transition temperature. We expect it to have broad application because C_p data for solid heavy organic compounds is exceedingly sparse.

Temperature Dependence of C_p

In previous studies, many researchers applied Einstein's model for the description of heat capacities as a function of temperature. Briard et al.[11] successfully established a four-parameter equation to correlate the temperature dependence of molar heat capacity for *n*-alkanes:

$$C_p = 3AR\left(\frac{\theta}{T}\right)^2 \frac{\exp\left(\frac{\theta}{T}\right)}{\left[\exp\left(\frac{\theta}{T}\right) - 1\right]^2} + cT + dT^2 \qquad (1)$$

where parameters A, θ, c, and d are functions of the number of carbons in the alkane chain. The adjustable parameters were optimized for 17 *n*-alkanes individually by fitting this equation to experimental heat capacities between 0 K and the first transition temperature. Einstein's term takes into account contributions from the group vibrations. Contributions from skeletal vibrations (stretching, bending, wagging, twisting, and rocking) are described by a quadratic expression derived from Benson's group contribution method.

Sallamie and Shaw[12] developed a predictive technique employing a combined Debye−Einstein model for computation of heat capacities from 0 K to the temperature of fusion. Their model substitutes quantum-mechanical calculation of 3N-6 skeletal vibrations[13] for the Einstein term.

More recently, we have performed Density Functional Theory (DFT) calculations for a wide range of organic compounds. We found that compounds with the same number of vibrations per mass of molecule exhibit very similar temperature dependence of specific heat capacity per unit mass [SCP/($J\cdot K^{-1}\cdot g^{-1}$)] and

---

* Corresponding author. E-mail: jmshaw@ualberta.ca.

10.1021/je700078b CCC: $37.00

that this consistency improves with increasing molar mass. A probable explanation of this thermodynamic behavior is linked to the weak dependence of SCP on the specific nature of the $3N$-6 atomistic vibrations (N is number of atoms per molecule) that dominate over a broad range of temperatures and the saturation of the six skeletal vibrations at low temperatures. We came to the conclusion that the SCP of heavy organic compounds is primarily a function of the number of vibrations per mass of molecule Φ linked directly to the number of atoms per mass of molecule, irrespective of the nature of the atoms or the molecular structures involved:

$$\Phi = \frac{f}{m} = \frac{3N_AN}{M} = \frac{3N}{m} \tag{2}$$

where f is the number of vibrations per molecule, m is the mass of a molecule, N_A is Avogadro's number, N is the number of atoms per molecule, and M is the molar mass of a molecule. This work is nearing completion. In the interim, we have applied the principle identified to the present work where we have modified eq 1 to describe the temperature dependence of SCP.

We modified the expressions for parameters A, θ, c, and d to a power series of second order and established a new parameter α. Thus, the expression of the heat capacity becomes

$$C_p = 3(A_1\alpha + A_2\alpha^2)R\left(\frac{\theta}{T}\right)^2 \frac{\exp\left(\frac{\theta}{T}\right)}{\left[\exp\left(\frac{\theta}{T}\right) - 1\right]^2} + (C_1\alpha + C_2\alpha^2)T + (D_1\alpha + D_2\alpha^2)T^2 \tag{3}$$

where R is the universal gas constant (8.314472 $\text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$), and α can be evaluated from elemental composition:

$$\alpha = \frac{N}{M} = \frac{\sum_{i=1}^{n} v_i}{\sum_{i=1}^{n} v_iM_i} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_iM_i} = \frac{\sum_{i=1}^{n} \frac{w_i}{M_i}}{\sum_{i=1}^{n} w_i} \tag{4}$$

where v_i is the stoichiometric coefficient for element i in a compound consisting of N atoms, n is the number of elements in a compound, and M_i is the molar mass of chemical element i ($\text{g}\cdot\text{mol}^{-1}$). Variable x_i is the mole fraction of element i in a compound, and w_i is the mass fraction of element i. For example, for n-alkanes with the formula C_nH_{2n-2}:

$$\alpha = \frac{\sum_{i=1}^{n} v_i}{\sum_{i=1}^{n} v_iM_i} = \frac{v_C + v_H}{v_C\cdot 12.0107 + v_H\cdot 1.00794} = \frac{v_C + (2v_C - 2)}{v_C\cdot 12.0107 + (2v_C - 2)\cdot 1.00794} \tag{5}$$

Data Treatment: Database and Correlation of SCP

A database comprising 165 organic compounds of different classes with molar masses ranging from (130 to 1100) $\text{g}\cdot\text{mol}^{-1}$ and temperatures ranging from 0 K to the temperatures of the first phase transition (T_{trs}) were culled from the literature. The thermodynamic properties of crystals at $T > T_{trs}$ were not used. The compounds, the temperature range for the SCP data, the number of SCP values, the regression quality, and the literature sources are summarized as Supporting Information. This database contains, for the most part, smoothed heat capacity values based on comprehensive analysis. Discussion of data quality is beyond the scope of this work. A comprehensive critical review of some data included in our database was published elsewhere.[14] As the experimental error for measured heat capacities exceeds the error induced by uncertainty over the temperature scales employed, we did not correct the data in the database to the latest international temperature scale (ITS-90).

Table 1. Evaluated Parameters of Equation 2 as a Function of α

parameter	expression
N	$0.013183\alpha + 0.249381a^2$
q	151.8675
c	$0.026526\alpha - 0.024942\alpha^2$
d	$0.000025\alpha - 0.000123\alpha^2$

To identify the seven parameters for eq 3, 72 compounds with a total of 2020 SCP values in the temperature range from 0 K to T_{trs} were drawn from the database and used as a training set. These compounds (identified in the Supporting Information along with the data source, numbers of SCP values, and the temperature ranges used) include linear alkanes, alcohols, esters, carboxylic acids, aromatic, and naphthenic compounds among others. Data not used in the development of the correlation were used for testing. These compounds are similarly identified in the Supporting Information. The parameters of the SCP equation regressed from the training set were obtained by minimizing the objective function S defined as

$$S = \sum_{i=1}^{m}\left[\sum_{j=1}^{n}(C_p^{exp} - C_p^{calc})_j^{\,2}/n\right] \tag{6}$$

Parameters with the superscript "exp" relate to experimental data, whereas parameters with the superscript "calc" are derived from eq 3; m denotes the number of compounds, and n represents the number of experimental SCP values for a compound.

It is important to note that SCP data arising in the region of solid−solid transitions or where a premelting effect was apparent were not included in the parameter regression calculations. This is a common approach and the one used, for example, by Briard et al.[11]

Results and Discussion

Fitting the Adjustable Parameters Using the Training Data Set. The optimized values for parameters of eq 3 and their expressions as a function of α are indicated in Table 1. The quality of the fit of the correlation for the training set is evaluated in terms of an average standard deviation of the fit σ_F, average relative deviation δ, and average absolute deviation ϵ. These three measures provide useful statistics for different parts of the temperature ranges and are defined as follows:

$$\sigma_F = \frac{1}{m}\sum_{i=1}^{m}\sqrt{\sum_{j=1}^{n}(C_p^{exp} - C_p^{calc})_j^{\,2}/(n - p)} \tag{7}$$

$$\delta = \frac{1}{m}\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}\sqrt{[(C_p^{exp} - C_p^{calc})/C_p^{exp}]_j^{\,2}/n}\right\} \tag{8}$$

$$\epsilon = \frac{1}{m}\sum_{i=1}^{m}\left[\sum_{j}\sqrt{(C_p^{exp} - C_p^{calc})_j^{\,2}/n}\right] \tag{9}$$

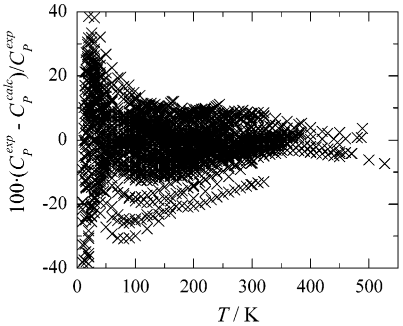

Figure 1. Relative deviation of the correlated SCP values from experimental values (training set).

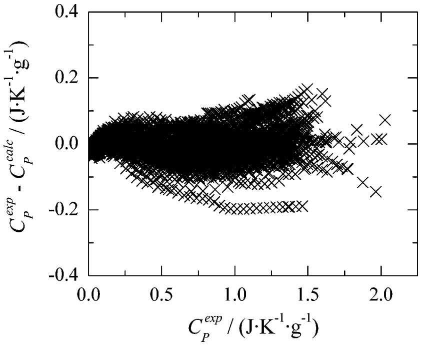

Figure 2. Absolute deviation of the correlated SCP values from experimental values (training set).

Table 2. Quality of Training Set Correlation from 50 K to First Transition Temperature

no. of compounds	72
no. of SCP values	1621
σ_F/J·K^{-1}·g^{-1}	0.059
ϵ/J·K^{-1}·g^{-1}	0.040
100·δ	5.36

where parameter p is the number of fitted parameters. The relative deviation amplifies absolute deviations at low temperatures, where the magnitude of SCP is small, while the standard deviation of the fit weights more heavily the larger absolute deviation expected at higher temperatures where SCP is larger. Since the magnitude of SCP below 50 K is small, the relative deviation might reach hundreds of percent, even if the absolute deviation is less than 0.02 J·K^{-1}·g^{-1}. Therefore, we focus on temperatures above 50 K in the discussion below. The overall quality for the regression of the training from 50 K to the first transition temperature is summarized in Table 2.

Relative deviations for the correlated SCP values from the training set are shown in Figure 1. Diphenylsulfide and 1,2-dibenzoylethylene have the highest positive deviations. 1,1,2,3,6,7,8-Hexahydropyrene and *trans*-2-cyclohexylcyclohexanol have the largest negative deviations. Correlation performance clearly degrades at low temperatures when relative deviation is the objective function. This is due to the small magnitude of SPC at low temperatures. The absolute deviation distribution of the correlated SCP values from the experimental data, shown in Figure 2, is largely independent of the SPC value.

Testing the Predictive Character of the Correlation. To test the predictive capability of the correlation, SCP values for a test set, designated test set I, containing 93 solids with 2080 values were computed. The test set I compounds and details are also identified in the Supporting Information. The statistical overview for test set I is presented in Table 3 where the average relative deviation, average absolute deviation, and average standard deviation defined as

$$\sigma = \frac{1}{m}\sum_{i=1}^{m}\sqrt{\sum_{j=1}^{n}(C_p^{\text{exp}} - C_p^{\text{calc}})_j^2/(n-1)} \qquad (10)$$

are reported for test set I as a whole. Fit statistics for individual compounds are shown in the Supporting Information. The average relative deviation for test set I in the estimation of SCP is 6.8 %, and the average standard deviation is 0.065 J·K^{-1}·g^{-1}. The fit between the experimental and predicted SCP values is shown in Figures 3 and 4.

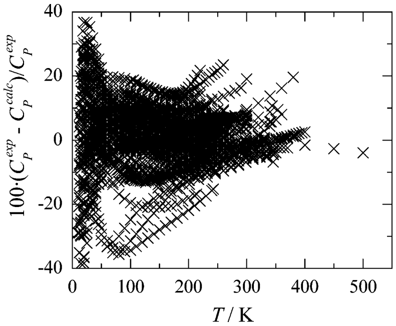

Figure 3. Relative deviation of the correlated SCP values from experimental values (test set I).

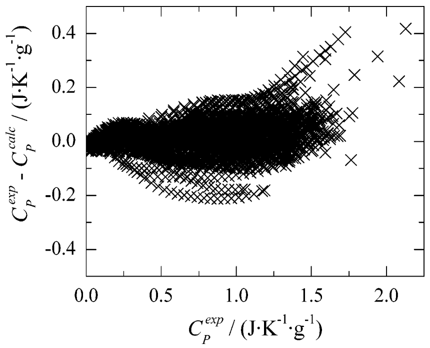

Figure 4. Absolute deviation of the correlated SCP values from experimental values (test set I).

Table 3. Statistical Overview of Test Set I, Estimates from 50 K to First Transition Temperature

no. of compounds	93
no. of SCP values	2080
σ/J·K^{-1}·g^{-1}	0.065
ϵ/J·K^{-1}·g^{-1}	0.058
100·δ	6.80

Comparison of the New Correlation with Specialized Methods Available in the Literature. The predictive capability of the new general correlation was compared with the calculated SCP values obtained using two available group contribution techniques.[1,8] These techniques provide highly accurate estimates for specific compound classes with respect to the functional groups for which the methods were developed. Goodman et al.[1] developed two techniques applicable to different temperature ranges. The power-law method utilizes an empirical temperature dependence, while the partition-function method is based on the Einstein−Debye partition-function for crystals with a modified frequency distribution function. Both methods can be used for temperatures above 50 K. However the power-law method is recommended for temperature range from (50 to 250) K whereas the partition-function method provides better predictions above 250 K.

Here, we compare the estimates only with the power-law method applicable for the temperature range from (50 to 250)

Table 4. Statistical Comparison with Available Methods

method	temp range/K	no. of compds	SCP values	σ/J·K^{-1}·g^{-1}	ϵ/J·K^{-1}·g^{-1}	100·δ
			Test Set II			
ref 1	50−250	88	1745	0.113	0.103	12.69
this work	50−250	88	1745	0.062	0.056	7.03
			Test Set III (Alkanes, Alkenes, Alkanols)			
ref 8	50−150	24	218	0.025	0.019	2.50
ref 1	50−150	24	218	0.040	0.035	4.82
this work	50−150	24	218	0.043	0.038	4.98

K. Results of the comparison are illustrated in Table 4. Because group contribution factors are unavailable for numerous compounds in the database, the comparison subset (test set II) includes only 88 compounds with 1745 SCP values.

Kabo et al.[8] established additive correlations utilizing 10 K increments from (10 to 150) K for "effective atom" and "effective bond" structure fragments. The effective atom fragments were calculated for alkanes whereas more complex effective bond fragments include alkanes, alkenes, alkanols, and alkanones. We have compared estimates from the effective bond method in the Table 4 as well. Test set III comprises 24 compounds and 218 SCP values.

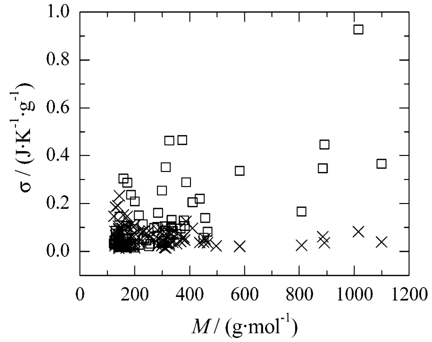

Figure 5. Distribution of the standard deviation σ as a function of molar mass: ×, this work; □, power-law method.

With respect to test set II, eq 3 exhibits an average relative deviation of 7.0 % and an average standard deviation of 0.062 J·K^{-1}·g^{-1}. The power-law method provides an average relative deviation of 13 % and an average standard deviation 0.113 J·K^{-1}·g^{-1}, which is consistent with an average expected accuracy of 13 % noted in the work of Goodman et al. In Figure 5, we show that the accuracy of the power-law method decreases with increasing molar mass for solids in the test set II database. The accuracy of eq 3 for test set I, which includes test set II, is also shown in Figure 5. It does not suffer from this deficiency. For some of the compounds with molar mass less than 200 g·mol^{-1}, the average standard deviation exceeded 0.1 J·K^{-1}·g^{-1}; therefore, we particularly recommend eq 3 for compounds possessing higher molar masses where accuracy is largely independent of molar mass.

The overall results for test set III confirm that the method of effective bonds is very accurate for a limited range of compounds. Equation 3, as expected, is not as accurate. However, eq 3 is simple, is based only on knowledge of elemental composition, and can be applied satisfactorily to estimate the SCP for a broad range of organic compounds containing elements C, H, N, S, and O over a broad range of temperatures.

Conclusions

A generalized heat capacity correlation for pure solid organic compounds based only on elemental composition valid from 50 K to the melting point temperature has been developed. The principle employed in developing the correlation is that SCP is a function of the number of fundamental vibrations per mass of molecule, which in this context is linked directly to the number of atoms per mass of molecule irrespective of the nature of the atoms. The correlation has been applied to diverse pure organic compounds containing C, H, N, O, and S, using seven universal coefficients. The range of compounds includes alcohols, *n*-alkanes, and polynuclear aromatic hydrocarbons among others. The average standard deviation achieved is 0.065 J·K^{-1}·g^{-1}. The accuracy of this simple correlation, rooted in quantum-mechanical theory, is competitive with highly specialized computational methods with restricted ranges of temperature and composition.

After comparison with available estimation techniques, we recommend this correlation in particular for prediction of heat capacities of larger organic compounds (e.g., for compounds with molar masses greater than 200 g·mol^{-1}). The principal applications for the correlation are expected to be the estimation of SCP for pure solid organic molecules, where SCP data are normally unavailable, and for SCP estimation and phase transition identification for poorly defined mixed organic solids such as asphaltenes where elemental analysis but not molecular structures is available.

Supporting Information Available:

List of organic compounds used in the training and test sets, their literature sources, temperature ranges, number of SCP values, and individual statistical results. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Goodman, B. T.; Wilding, W. V.; Oscarson, J. L.; Rowley, R. L. Use of the DIPPR database for development of quantitative structure−property relationship correlations: heat capacity of solid organic compounds. *J. Chem. Eng. Data* **2004**, *49*, 24−31.

(2) Domalski, E. S.; Hearing, E. D. Estimation of the thermodynamic properties of hydrocarbons at 298.15 K. *J. Phys. Chem. Ref. Data* **1988**, *17*, 1637−1678.

(3) Domalski, E. S.; Hearing, E. D. Estimation of the thermodynamic properties of C-H-N-O-S-halogen compounds at 298.15 K. *J. Phys. Chem. Ref. Data* **1993**, *22*, 805−1159.

(4) Chickos, J. S.; Hesse, D. G.; Liebman, J. F. A group additivity approach for the estimation of heat-capacities of organic liquids and solids at 298 K. *Struct. Chem.* **1993**, *4*, 261−269.

(5) Helgeson, H. C.; Owens, C. E.; Knox, A. M.; Richard, L. Calculation of the standard molal thermodynamic properties of crystalline, liquid, and gas organic molecules at high temperatures and pressures. *Geochim. Cosmochim. Acta* **1998**, *62*, 985−1081.

(6) Richard, L.; Helgeson, H. C. Calculation of the thermodynamic properties at elevated temperatures and pressures of saturated and aromatic high molecular weight solid and liquid hydrocarbons in kerogen, bitumen, petroleum, and other organic matter of biogeochemical interest. *Geochim. Cosmochim. Acta* **1998**, *62*, 3591−3636.

(7) Hurst, J. E.; Harrison, B. K. Estimation of liquid and solid heat-capacities using a modified Kopps rule. *Chem. Eng. Commun.* **1992**, *112*, 21−30.

(8) Kabo, G. J.; Kozyro, A. A.; Diky, V. V. Additivity of thermodynamic properties of organic compounds in the crystalline state. 1. Additive calculations for thermodynamic properties of alkanes, alkenes, alkanols, and alkanones. *J. Chem. Eng. Data* **1995**, *40*, 160−166.

(9) Kabo, G. J.; Kozyro, A. A.; Diky, V. V.; Simirsky, V. V. Additivity of thermodynamic properties of organic compounds in crystalline state. 2. Heat capacities and enthalpies of phase transition of alkyl derivatives of urea in crystalline state. *J. Chem. Eng. Data* **1995**, *40*, 371−393.

(10) Diky, V. V.; Kozyro, A. A.; Kabo, G. J. Additivity of thermodynamic properties of organic compounds in the crystalline state. 3. Heat capacities and related properties of urea phenyl derivatives. *J. Chem. Eng. Data* **2002**, *47*, 239−244.

(11) Briard, A. J.; Bouroukba, M.; Petitjean, D.; Dirand, M. Models for estimation of pure *n*-alkanes thermodynamic properties as a function of carbon chain length. *J. Chem. Eng. Data* **2003**, *48*, 1508−1516.

(12) Sallamie, N.; Shaw, J. M. Heat capacity prediction for polynuclear aromatic solids using vibration spectra. *Fluid Phase Equilib.* **2005**, *237*, 100−110.

(13) 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.

(14) Domalski, E. S.; Hearing, E. D. Heat capacities and entropies of organic compounds in the condensed phase, vol. 3. *J. Phys. Chem. Ref. Data* **1996**, *25*, 1−525.

Received for review February 8, 2007. Accepted March 22, 2007.

JE700078B