On Calculating a Polymer's Enthalpy of Formation with Quantum Chemical Methods

Edward N. Brothers,* Artur F. Izmaylov, Alexander A. Rusakov, and Gustavo E. Scuseria

Department of Chemistry, Mail Stop 60, Rice University, Houston, Texas 77005-1892

Received: August 30, 2007; In Final Form: October 16, 2007

Calculating the enthalpy of formation of a polymer with ab initio methods requires two choices. The first decision is whether to use oligomeric extrapolation or periodic boundary conditions to model the extended system, and the second choice is between formation reactions to be modeled, for example, formation from atoms, formation from standard states, or formation from some set of molecular systems. Utilizing *trans*-polyacetylene and polyethylene as examples, the oligomeric and periodic techniques are contrasted, leading to a discussion of the larger than minimal unit cell required when frequency calculations only include in-phase vibrations, that is, only the $k = 0$ frequencies, in an enthalpy of formation calculation. The accuracy of calculating the enthalpy of formation, in light of density functional theory's increased error with larger systems and with respect to various reference states, is also discussed. The calculation of the enthalpy of formation for a polymer is most accurate when the reference states are chosen carefully and most efficient when using periodic boundary conditions.

The ab initio study of polymers and solids is typically accomplished in one of two ways. In the oligomeric approach, a molecule consisting of one or more units that repeat in the bulk material is modeled using standard molecular techniques, and then additional units are included in the model system until the value of the property of interest converges. The bulk system is thus modeled through the extrapolation of molecular results. In the periodic approach, periodic boundary conditions (PBC) are imposed, with a single repeating unit modeled as interacting with effectively infinite copies of itself, and the property of interest is taken as the value calculated for the central unit. The usefulness of PBC cannot be overstated because the results are free from edge effects and PBC calculations are computationally more efficient than any oligomeric approach.[1,2] The manifestation of the latter point is especially evident when investigating properties that necessarily converge slowly in oligomers, such as polarizability.[3,4]

Because of recent advances in quantum chemistry codes, it is now possible to calculate almost any property for a polymer or solid while using PBC that is calculable for molecules. One property commonly calculated for molecules is the enthalpy of formation (ΔH_f), which has become a litmus test for new density functionals and other quantum chemical methods. For example, density functionals are often rated on their performance versus the G3/99 set, which is mostly enthalpies of formation.[5] Calculating this property requires the calculation of both the energy and the vibrational spectrum. Techniques for determining vibrational frequencies and thus ΔH_f for molecules (and by extension oligomers) are well known,[6] and oligomeric extrapolation provides a route to polymer ΔH_f. There are, however, several subtleties in calculating the ΔH_f for a polymer with PBC that arise from calculating the vibrational spectra for a periodic system.[7,8] The first half of this letter is thus devoted to discussing the periodic approach, especially in comparison to the oligomeric approach.

All vibrations in periodic systems can be divided in two categories: in-phase and out-of-phase. In-phase vibrations consist of the motions of atoms within the unit cell, which are simultaneously repeated by atoms in other cells; hence, there is no difference in phase between different cells ($k = 0$). Out-of-phase vibrations, in contrast, have a phase difference between motions in different cells ($k \neq 0$). To model the thermodynamic properties of a solid, the full vibrational spectrum is needed. However, the current work uses an implementation of harmonic frequencies in PBC where only the in-phase part can be evaluated.[1,7,8] This does not prevent the calculation of the entire vibrational spectrum because the out-of-phase portion can be included through the so-called frozen phonon technique.[9] The idea behind the frozen phonon approach is taking multiple unit cells as a new reference super-cell, which results in vibrational frequencies becoming available as in-phase vibrations within the new reference super-cell. These new in-phase vibrations are out-of-phase motions when considered in the framework of the minimum translationally invariant reference cell. Thus, using M minimal cells to build the reference super-cell effectively obtains all frequencies for $k = 2\pi m/(a_c M)$ (where $m = 1, ..., M$, and a_c is the lattice parameter). With this information, one can potentially generate uniformly discretized phonon dispersion curves $\nu_i = \nu_i(k)$, where ν_i is a vibrational frequency. However, in the case of enthalpy of formation, this is an unnecessary intermediate construction because one can use obtained frequencies $\nu_i(k_j)$ directly in the discrete sum over states' equation for calculating the vibrational contribution

$$E_{\text{vib}}^{\text{dis}} = \sum_{i,j} \nu_i(k_j) \left(\frac{1}{2} + \frac{1}{e^{\nu_i(k_j)/T} - 1} \right) \qquad (1)$$

at temperature T. Note that $E_{\text{vib}}^{\text{dis}}$ has the same continuous limit

$$E_{\text{vib}} = \int_0^\infty \nu g(\nu) \left(\frac{1}{2} + \frac{1}{e^{\nu/T} - 1} \right) d\nu \qquad (2)$$

* Corresponding author. E-mail: enb@rice.edu.

TABLE 1: Enthalpies of Formation of Per Unit Cell of *trans*-Polyacetylene and Polyethylene in kcal/mol[a]

	polyacetylene		polyethylene	
N	oligomer	PBC	oligomer	PBC
2	20.5	18.1	−8.6	−8.4
3	19.0	17.8	−8.5	−8.5
4	18.4	17.8	−8.5	−8.5
5	18.1	17.7		
6	18.0	17.7		
7	17.9			
8	17.9			
9	17.8			
10	17.8			
11	17.8			
12	17.7			

[a] These are calculated as $\Delta H_f(N) - \Delta H_f(N-1)$, where N is the number of units, at the TPSS/3-21G level of theory.

as the more common approach based on the generation of discretized density of state function $g(\nu)$ from phonon dispersion curves.[10−12] The implications of this for calculating ΔH_f for polymers is most easily seen by an example.

The ΔH_f per minimum unit cell of *trans*-polyacetylene and polyethylene can be trivially calculated through an oligomeric extrapolation, and this has been done here with the density functional of Tao, Perdew, Staroverov, and Scuseria (TPSS)[13] and a small basis set (3-21G)[14] for illustration purposes. In these specific calculations and throughout this letter, computations were performed with the GAUSSIAN suite of programs[15] using high-precision settings.[16] Molecular frequencies were calculated analytically,[6] while periodic system frequencies were calculated through the numerical differentiation of atomic displacement forces for each atom in the unit (or super) cell, as per previous studies.[1] Please note that all results were obtained at geometries that were optimized with the method/basis set of interest. In comparing oligomers to PBC results, the reaction being modeled is the commonly used formation from atoms in the gas phase. The ΔH_f obtained from these calculations is not expected to be accurate, but rather to converge to a single value in both molecular and PBC calculations. Enthalpies of formation per unit cell were calculated for molecules H−[HC=CH]$_N$−H and H−[H$_2$CCH$_2$]$_N$−H for $N = 1−12$, as can be seen in Table 1. Note that these values were calculated as $\Delta H_f(N) - \Delta H_f(N-1)$ rather than $\Delta H_f(N)/N$ as we found the subtraction method to converge more rapidly for both PBC and oligomers. The ΔH_f per minimum repeating unit was found to converge fairly rapidly, especially for polyethylene.

A similar extrapolation was done with PBC. As stated previously, the ΔH_f calculation required a non-minimum unit cell to be accurate because only $k = 0$ frequencies are calculated. (Please note that this issue does not arise in energy-only calculations because doubling the unit cell exactly doubles the energy.) Using only the minimum unit cell resulted in an imprecision of ∼1 kcal/mol arising from incomplete vibrational spectra. It is interesting to note how much faster the super-cell approach converged in the PBC case for *trans*-polyacetylene, requiring 6 units to reach its final value, which can be favorably compared with the more than 12 units required in the molecular oligomer calculation. (For polyethylene, both periodic and oligomeric calculations converge almost immediately, after just three unit cells.) This demonstrates that correct ΔH_f calculation for periodic systems requires the out-of-phase ($k \neq 0$) portion of the vibrational spectra. Using a super-cell when calculating vibrations is thus a technique to implicitly provide information equivalent to that from discretized phonon dispersion curves in the ΔH_f vibrational contribution. Because of the fast conver-

gence with PBC, the example also shows that although a minimal unit cell is insufficient a PBC calculation of polymeric ΔH_f is still the optimal route to this property. For higher dimensional systems, edge effects would be expected to be even larger and thus very large structures would be needed for oligomers, making periodic calculations even more advantageous.

Now that the method for obtaining the enthalpy of the polymer is understood, it is necessary to consider the formation reaction that is to be modeled, in order to have a complete route to ΔH_f. Formally, ΔH_f for a molecule or polymer is the change in enthalpy between products and reactants of the formation reaction that constructs the molecule/polymer of interest from the constituent elements in their standard states. For example, the ΔH_f of *trans*-polyacetylene can be defined as the enthalpy of reaction

$$H_2^{gas} + 2C^{graphite} \rightarrow [-HC=CH-] \tag{3}$$

Attempting to model the formation reaction from standard states is fundamentally impractical, however, because many standard states are not amenable to computation. For example, modeling graphite requires that the method of interest contains accurate van der Waals forces to model the interlayer interactions adequately, which many methods fail to do, even those that produce otherwise excellent thermochemistry.[17] Thus, the modeling of other formation reactions is required for computational practicality.

Continuing with the formation of *trans*-polyacetylene as an example, the formation reaction most commonly modeled in quantum chemical prediction of thermochemistry involves gas phase atoms as reactants

$$2H + 2C \rightarrow [-HC=CH-] \tag{4}$$

In a typical calculation, the atomic energies and the energy (with vibrational corrections) of the compound of interest are used to form an enthalpy of reaction, and this is combined with the ΔH_f of the atoms to determine the ΔH_f of the molecule. There are several issues with this, most notably that many density functional theory (DFT) based methods have errors in atomic energy calculations, which leads to larger errors in larger molecules, that is, size-dependent systematic errors, in the ΔH_f.[18] One option is to replace the atomic energies with parameters, which reduces the errors but requires a re-parametrization for each method/basis set combination of interest.[19] This increases the accuracy dramatically and can even remove qualitative errors but is fraught with all of the peril of parametrization, such as the large computational effort of the parametrization, test-set dependence, and so forth.

Although there are effectively an infinite number of formation reactions that could be considered, there are several that are especially interesting for polymeric hydrocarbons such as *trans*-polyacetylene. The most obvious reactions to model are the simple polymerization reactions

$$HC\equiv CH \rightarrow [-HC=CH-] \tag{5}$$

$$H_2C=CH_2 \rightarrow [-HC=CH-] + H_2 \tag{6}$$

$$H_3CCH_3 \rightarrow [-HC=CH-] + 2H_2 \tag{7}$$
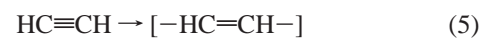
Two other reactions that must be considered have a great deal of bond-type similarity with the two polymers under consideration here, namely formation reactions from benzene and propane. Please note that in the case where a theoretical method

TABLE 2: *trans*-Polyacetylene's Enthalpy of Formation Per Unit Cell in kcal/mol Using Various Density Functionals and the 3-21G and 6-31G(d,p)[27] Basis Sets[a]

	atomic reference states		molecular reference states				
method	standard	parametrized	acetylene	ethylene	ethane	propane	benzene
LSDA[20]/3-21G	−54.7	11.0	−6.9	6.6	20.3	21.1	9.9
PBE[21]/3-21G	0.2	11.5	3.4	9.9	16.2	16.0	10.9
PBEh[22]/3-21G	14.8	13.4	3.5	12.6	22.0	21.6	12.9
TPSS[13]/3-21G	17.6	11.7	6.7	10.7	13.8	13.4	11.5
TPSSh[23]/3-21G	21.5	12.6	6.7	11.9	16.6	16.1	12.4
HF[24]/3-21G	147.9	15.5	17.3	19.9	24.0	22.1	15.4
LSDA/6-31G(d,p)	−62.0		−10.3	4.6	17.0	17.7	9.5
PBE/6-31G(d,p)	−7.0		0.6	8.0	12.6	12.2	10.5
PBEh/6-31G(d,p)	5.0		0.8	10.8	18.5	17.9	12.5
TPSS/6-31G(d,p)	6.9		4.00	9.1	10.8	10.2	11.2
TPSSh/6-31G(d,p)	10.1		3.9	10.3	13.5	12.8	12.0
HF/6-31G(d,p)	134.8		18.6	23.4	26.5	24.7	20.1

[a] The labels "standard" and "parametrized" refer to the type of atomic energy used in an atomic formation reaction, with parametrized energies taken from ref 19. (Blanks arise from the lack of parameters for that method/basis combination.) The columns labeled with molecular names refer to formation reactions given in the text. The target value, from G3MP2, is 12.4 kcal/mol.

TABLE 3: Polyethylene's Enthalpy of Formation Per Unit Cell in kcal/mol Using Various Density Functionals and the 3-21G and 6-31G(d,p) Basis Sets[a]

	polyethylene						
	atomic reference states		molecular reference states				
method	standard	parametrized	acetylene	ethylene	ethane	propane	benzene
LSDA/3-21G	−92.4	−13.9	−39.2	−25.8	−12.1	−11.3	−22.4
PBE/3-21G	−21.3	−11.5	−21.9	−15.3	−9.1	−9.3	−14.4
PBEh/3-21G	−10.6	−11.6	−26.8	−17.7	−8.3	−8.7	−17.4
TPSS/3-21G	−8.4	−10.2	−15.5	−11.5	−8.4	−8.8	−10.7
TPSSh/3-21G	−6.4	−10.4	−18.0	−12.7	−8.1	−8.6	−12.3
HF/3-21G	148.2	−7.9	−10.4	−7.8	−3.6	−5.6	−12.3
LSDA/6-31G(d,p)	−96.8		−39.1	−24.3	−11.8	−11.2	−19.4
PBE/6-31G(d,p)	−25.2		−20.6	−13.2	−8.6	−9.0	−10.7
PBEh/6-31G(d,p)	−17.7		−25.7	−15.6	−7.9	−8.5	−13.9
TPSS/6-31G(d,p)	−16.6		−14.6	−9.5	−7.8	−8.5	−7.4
TPSSh/6-31G(d,p)	−15.4		−17.2	−10.7	−7.5	−8.3	−9.0
HF/6-31G(d,p)	129.4		−12.1	−7.3	−4.3	−6.0	−10.6

[a] The labels "standard" and "parametrized" refer to the type of atomic energy used in an atomic formation reaction, with parametrized energies taken from ref 19. (Blanks arise from the lack of parameters for that method/basis combination.) The columns labeled with molecular names refer to formation reactions given in the text. The target values are −9.9 kcal/mol from extrapolated experimental results or −10.2 from G3MP2 calculations.

predicts little to no bond length alternation in *trans*-polyacetylene a formation from benzene is effectively an isodesmic reaction, that is, a reaction where products and reactants have the same number and type of bonds.

The ΔH_f per unit cell of *trans*-polyacetylene and polyethylene has been calculated by several electronic structure methods using each of these reactions, and the results are summarized in Tables 2 and 3. As stated above, all properties of the molecule/polymer under consideration including optimum geometry, energy, and vibrational frequencies were determined with the basis set and method being tested. The polymer results were obtained by performing calculations on a super-cell made by combining six minimum periodic unit cells, which converged the results in the worst case to 0.01−0.02 kcal/mol. Note that these results are thus calculated as $\Delta H_f^{N=6}/6$ rather than $\Delta H_f^{N=6} - \Delta H_f^{N=5}$ as was the case in Table 1. This was done in order to keep computational costs down.

Before discussing the results of using these reactions, it is necessary to mention two minor points. First, there are currently no experimental values for the ΔH_f for a single strand of *trans*-polyacetylene. It can, however, be easily extrapolated from high-accuracy/high-expense thermochemistry methods like G3MP2.[25] We performed these calculations and used an oligomeric extrapolation to produce a ΔH_f per unit cell of 12.4 kcal/mol. Second, the ΔH_f per unit cell for polyethylene can be extrapolated easily from experimental results for n-alkanes,[26] resulting

in a value of −9.9 kcal/mol, as well as from G3MP2 oligomeric calculations, which gives a result of −10.2 kcal/mol.

There are several points to be taken from the data in Tables 2 and 3. First, there are common qualities in both the *trans*-polyacetylene and polyethylene results. Unsurprisingly, there are large errors when the standard formation reaction from gas phase atoms is used, which was expected because of DFT's systematic errors with large systems.[19] Also expected was the success of parametrized atomic energies, which results from the method being explicitly developed to calculate ΔH_f accurately. In addition, both polymers have larger errors with the larger 6-31G(d,p) basis set for all results, which probably arises from cancellation of errors with the small basis. Finally, the ΔH_f for both polymers is modeled with increased accuracy when using reference states other than the gas-phase atoms. This can be seen most easily by considering the two methods whose performance with regular atomic energies are the worst, HF and LSDA. Using any reference state other than gas-phase atoms removes the qualitative failure in ΔH_f prediction, which was present universally in tests of these methods. This unequivocally demonstrates the importance of the reaction being modeled.

Another example of the importance of the reference reaction is the failure of formation from acetylene, which has errors large enough to be considered qualitative failures for both polymers with many of the methods/basis sets tested. This can be attributed to both polymers having different carbon hybridiza-

tions than the carbon in acetylene; that is, acetylene is not a good reference for either of these polymers.

Overall, *trans*-polyacetylene was modeled more accurately than polyethylene. If ethane is used as a reactant, then the *trans*-polyacetylene ΔH_f has large errors in general and the propane reference state results in errors nearly as large. In the case of the ethylene and benzene formation reaction, good results are obtained with most of the functionals, with benzene producing the best results. This can also be attributed in part to the preservation of hybridization of the carbon atoms in the formation reaction because these formation reactions are close isodesmic reactions, especially the formation from benzene. The variety of results indicates that formation reactions need to be tested carefully before use; for example, it is not immediately obvious that benzene would be significantly better than ethylene. For this specific polymer, benzene is a very good reference state because of its small size (and high symmetry) contributing to computational efficiency in addition to the accuracy of the results.

This preservation of hybridization effect can be seen again when modeling polyethylene. The formation reaction from ethane, a nearly isodesmic reaction, results in fairly good accuracy, whereas the other formation reactions lose accuracy as compared to their performance on *trans*-polyacetylene. The formation from the system closest in atom type to the infinite alkane, namely from propane, has the best accuracy for this polymer. Although accuracy like that found for *trans*-poly-acetylene is not achieved with the reactions considered here for polyethylene, a significant improvement is seen when atoms are no longer used as a reference state. For example, errors on the order of 100% of the target value are seen with 6-31G(d,p) and are greatly reduced by modeling the formation from propane reaction.

It has therefore been shown that periodic calculations, even though they require non-minimal unit cells, converge faster than molecular oligomers for the calculation of ΔH_f per unit cell of a polymer, and that these values can be calculated accurately using a well-selected formation reaction, even with small basis sets, with great improvements in accuracy occurring when atomic reference states are avoided. It is hoped that this study demonstrates both the ease and the potential pitfalls of calculating the ΔH_f of periodic systems.

References and Notes

(1) Kudin, K. N.; Scuseria, G. E. *Phys. Rev. B* **2000**, *61*, 16440.

(2) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 104103.

(3) Kudin, K. N.; Carr, R.; Resta, R. *J. Chem. Phys.* **2005**, *122*, 134907.

(4) Izmaylov, A. F.; Brothers, E. N.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 224105.

(5) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.

(6) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1979**, *13*, 225.

(7) Hirata, S.; Iwata, S. *J. Mol. Struct.: THEOCHEM* **1998**, *451*, 121.

(8) Izmaylov, A. F.; Scuseria, G. E. *J. Chem. Phys.* **2007**, *127*, 144106.

(9) Baroni, S.; de Gironcoli, S.; Dal Corso, A.; Giannozzi, P. *Rev. Mod. Phys.* **2001**, *73*, 515.

(10) Gilat, G.; Raubenheimer, L. J. *Phys. Rev.* **1966**, *144*, 390.

(11) Raubenheimer, L. J.; Gilat, G. *Phys. Rev.* **1967**, *157*, 586.

(12) Kam, Z.; Gilat, G. *Phys. Rev.* **1968**, *175*, 1156.

(13) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(14) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Scalmani, K. N.; Kudin, K. N.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; 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.; Li, X.; Hratchian, H. P.; Peralta, J. E.; Izmaylov, A. F.; Brothers, E.; Staroverov, V.; Kobayashi, R.; Normand, J.; Burant, J. C.; Millam, J. M.; Klene, M.; Knox, J. E.; 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.; Chen, W.; Wong, M. W.; Pople, J. A. *Gaussian Development Version*, revision E.05, Gaussian, Inc.: Wallingford, CT, 2006.

(16) SCF and geometry optimization criteria were set to "tight," which requests root-mean-squared values of 1×10^{-8} and 1×10^{-5} for density matrix change and atomic forces, and DFT quadrature used "ultrafine" grids, a pruned (99,590) grids.

(17) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463 and references therein.

(18) Csonka, G. I.; Ruzsinszky, A.; Tao, J.; Perdew, J. P. *Int. J. Quantum Chem.* **2005**, *101*, 506.

(19) Brothers, E. N.; Scuseria, G. E. *J. Chem. Theory Comput.* **2006**, *2*, 1045, and references therein.

(20) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200. We used the fifth proposed formula in this paper for local correlation.

(21) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865. Errata: Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

(22) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(23) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129.

(24) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(25) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703.

(26) *NIST Thermochemistry Webbook*, available online at *webbook.nist.gov*.

(27) Hariharan, P. C.; Pople, J. A. *Theoret. Chimica Acta* **1973**, *28*, 213.