

How to Predict Heats of Formation of Molecules?

Jun Li and Chan Kyung Kim*

Department of Chemistry and Chemical Engineering and Center for Design and Applications of Molecular Catalysts, Inha University, Incheon 402-751, Korea.

In the area of propellants, explosives, pyrotechnics and high energy density molecules (HEDMs), including military and commercial aims, two factors are of great interest: One is performance and the other is sensitivity.¹ Performance depends theoretically on both energy contents and the crystal packing of the molecules. The determining factors are, therefore, molecular heats of formation and solid densities.² To design or test new HEDMs, the two factors should be considered ahead of synthetic work. These properties can be obtained from experimental work. But for not-yet-synthesized molecules, there is no way to obtain these properties. In the past few years, numerous theoretical methods have been proposed for the prediction of the former with excellent accuracy in the gas phase as well as in solid.³ Although the methods are originated from different progress and theoretical background, they can be categorized broadly to three groups: 1. Group and atom additive method. 2. Quantitative Structure–Property Relationship (QSPR). 3. Quantum mechanics.

1. Group and Atom Additive Method

In the group and atom additive method, the thermochemical properties, for example, heats of formation or heat capacities can be estimated additively from the group or atom types consisting of the molecules. Benson and Buss firstly proposed the additive concepts in 1958⁴, and showed that the thermochemical properties can be estimated within 3 kcal/mol error. Following this pioneering work, Shaw,⁵ Bourasseau,⁶ Rouse,⁷ Mader,⁸ Kamlet–Jacobs,⁹ Bureš,¹⁰ and Cohen¹¹ improved the additive method for specific group of molecules.

Taking nitro benzene as an example, the heat of formation can be calculated as the formula as below:

$$\Delta H_f(\text{PhNO}_2) = 1(\text{C}_B\text{-NO}_2) + 5(\text{C}_B\text{-H}) = 16.9 \text{ kcal/mol} \quad (1)$$

where the $\text{C}_B\text{-NO}_2$ and $\text{C}_B\text{-H}$ are the number of carbon attached to the nitro group and hydrogen atom, respectively, and their contributions are known to be 3.3 and 0.4 kcal/mol, respectively. Once the parameters are derived, the group and atom additive method is very

*Corresponding author: kckyung@inha.ac.kr

simple and convenient to apply, and affords accepted values because it does not require much computing power.¹² But there is also disadvantages for this type of method. In order to arrive very accurate results, it often requires “further correction” or more additive items¹³ which in turn, decreases the efficiency of the application. In the parameter set from Pedley,¹⁴ there are 20 parameters for alkane only. Apparently, it is not very practical to apply.

There is equilibrium between accuracy and efficiency for such group of methods. To increase the accuracy, more types of additive groups or atoms should be defined and introduced, which inversely will decrease the efficiency. Based on the scheme from Laidler,¹⁵ Leal collected 200 molecules,¹² and divided them into 10 groups including alkanes, alkenes, alkynes, allenes, dienes/polyenes, diynes/alkyne-enes, cycloalkanes/cycloalkenes, alkyl radicals, benzene derivatives, and biphenyls/polyaromatic hydrocarbons, for which altogether 103 additive terms were defined. For multi-parameter regression analysis, the Excel 2000 was used to derive these parameters. The accuracy from the set is well accepted with an average difference of 1.28 kJ/mol for the gas phase enthalpy of formation. Compared with parameter set from Pedley for alkane,¹⁴ this set minimized the number of the additive items to 14, which increased the efficiency. Compared with parameter set from Tatevskii¹⁶, derived only for alkanes, the additive parameters from Leal are more versatile and comprehensive.

2. Quantitative Structure–Property Relationship (QSPR) method

Group and atom additive method usually requires a large number of molecules to derive the parameters. There are often too many additive terms that lowers the speed of the applications. Thus QSPR method that can derive optimal linear regression models between physicochemical properties and molecular descriptors became important. In 1994, Sukhachev and Pivina developed an efficient modelling of molecular activity (EMMA) model to predict the heat of formation.¹⁷ In their model, the heats of formation can be calculated using eq (2):

$$\Delta H_f^0 = -98.86 + 1.14SBE + 49.77 \times 4_{kr} - 437.78c_{mid} + \frac{61.76V_{mid}^2}{N_{at}} + 71.39Fr_1 - 195.44Fr_2 - 2933.25 \ln \left(\frac{Fr_3}{N_{at}} \right) \quad (2)$$

where, SBE is the enthalpy of the molecule without steric hindrance, 4_{kr} the Randić's index, c_{mid} the mean value of the diagonal elements in the inverse matrix of the weighed connectivity matrix G , V_{mid} mean value of the components of vector V , where V is the solution of the linear system $GV = C$, and C is the vector of corresponding vertex degrees in the structure graph, N_{at} number of atoms in the molecule. Fr_1 sum of the minimum charges at atoms in fragments of the type $-C-C-NO_2$, Fr_2 the minimum of all the minimum charges

at atoms in fragments of the type $At-N=O$, where At is an arbitrary atom, and

F_{r3} the sum of maximum charges at atoms in all five-atom linear chains with single bonds in the structure.

In 2007, Ali Vatani and coworkers utilized a training set of 1115 molecules and proposed a five-descriptor equation to estimate the standard heats of formation.¹⁸ The 5 descriptors utilized in this method are number of non-H atoms (nSK), sum of conventional bond orders ($SCBO$), number of oxygen atoms (nO), number of fluorine atoms (nF), and number of heavy atoms (nHM).

$$\Delta H_f^0 = 50.1688 - 80.52012nSK + 5364546SCBO - 169.21889nO - 174.75477nF - 266.57659nHM \quad (3)$$

For this method, the correlation coefficient (R^2) and the standard deviation were 0.9830 and 58.541 kJ/mol, respectively, which justified the accuracy of the method. Many other correlation equations for specific types of molecules were also proposed.¹⁹⁻²⁵ Usually, a QSPR study involves the process shown in Figure 1.

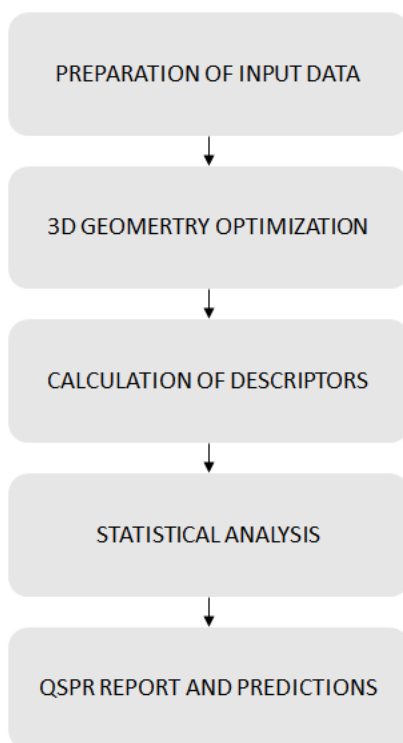


Figure 1: Flow chart for a QSPR study

3. Quantum Mechanics.

With the help of newly developed algorithms and continuous speed-up in computer hardware, QM method was applied to use complicated procedure but to obtain quite accurate results. Semi-empirical methods as MNDO, MINDO/3, AM1, and PM3 are practical and accurate enough for the estimation of heats of formation although are not as accurate as *ab initio* methods. In this area, there were some benchmark studies by many researchers²⁶⁻³⁵. Disch and Schulman performed *ab initio* calculations on the [n]prismanes.³⁶ Duan and coworkers utilized combined Hartree–Fock/density functional theory calculations with linear regression correction, and proposed an equation containing 5 descriptors. With B3LYP method and 6-3111G(d,p) basis set, the mean absolute deviation was 2.4 kcal/mol.³⁷ Bhattacharya converted *ab initio* energies to enthalpies for free radicals species.³⁸ Nicolaidis and coworkers transferred G2, G2(MP2), and G2(MP2,SVP) total energies to heats of formation.³⁹ Byrd and Rice combined QM calculation and derived additive parameters to predict heats of formation.⁴⁰ Shafagh derived a three-parameter equation based on DFT and *ab initio* thermal energies.⁴¹

These is an important trend in recent development: more and more methods are combined methods. For example, QM or QSPR method is combined with group additive method. As one of the direct methods cannot obtain good results, the second step treatment of the initial value from first step can improve the accuracy efficiently.^{42, 43}

In summary, we examined recent studies on estimation of heats of formation, including group additive, QSPR and QM methods. Group additive method required a large number of molecules to set up various terms and usually needed many additional types. In QSPR method, proper selection of molecular descriptors is very crucial for the accuracy and efficiency of the estimated values. In the QM method, *ab initio* and DFT calculations require more time but can obtain more accurate values. In terms of speed, semi-empirical method could be an alternative for screening a large database. To reach acceptable results with less time, combined method is a potential one to develop.

References

1. J. P. Agrawal, "Recent trends in high-energy materials," *Prog. Energy Combust. Sci.*, 24 [1] 1-30 (1998).
2. P. A. Persson, R. Holmberg, J. Lee, *Rock blasting and explosives engineering*, CRC Press 1993.
3. A. K. Sikder, G. Maddala, J. P. Agrawal, H. Singh, "Important aspects of behaviour of organic energetic compounds: a review," *J. Hazard. Mater.*, 84 [1] 1-26 (2001).
4. S. W. Benson, J. H. Buss, "Additivity rules for the estimation of molecular properties. Thermodynamic properties," *J. Chem. Phys.*, 29 [3] 546-572 (1958).
5. R. Shaw, "Heats of formation of nitroaromatics. Group additivity for solids," *J. Phys. Chem.*, 75 [26] 4047-4049 (1971).
6. S. Bourasseau, "A systematic procedure for estimating the standard heats of formation in the

- condensed state of non aromatic polynitro-compounds," *J. Energ. Mater.*, 8 [4] 266-291 (1990).
7. Jr, P. E. Rouse, "Enthalpies of formation and calculated detonation properties of some thermally stable explosives," *J. Chem. Eng. Data*, 21 [1] 16-20 (1976).
 8. C. L. Mader, Fortran BKW: A Code for Computing the Detonation Properties of Explosives (No. LA--3704). *Los Alamos Scientific Lab.*, N. Mex. 1967.
 9. M. J. Kamlet, S. J. Jacobs, "Chemistry of detonations. I. A simple method for calculating detonation properties of C-H-N-O explosives," *J. Chem. Phys.*, 48 [1] 23-35 (1968).
 10. M. Bureš, V. Majer, M. Záborský, "Modification of Benson method for estimation of ideal-gas heat capacities," *Chem. Eng. Sci.*, 36 [3] 529-537 (1981).
 11. N. Cohen, S. W. Benson, "Estimation of heats of formation of organic compounds by additivity methods," *Chem. Rev.*, 93 [7] 2419-2438 (1993).
 12. J. P. Leal, "Additive methods for prediction of thermochemical properties. The Laidler method revisited. 1. Hydrocarbons," *J. Phys. Chem. Ref. Data*, 35 [1] 55-76 (2006).
 13. D. Habibollahzadeh, M. E. Grice, M. G. Concha, J. S. Murray, P. Politzer, "Nonlocal density functional calculation of gas phase heats of formation," *J. Comput. Chem.*, 16 [5] 654-658 (1995).
 14. J. B. Pedley, *Thermochemical Data and Structures of Organic Compounds*, Vol. I Thermodynamic Research Center, College Station, TX, 1994.
 15. K. J. Laidler, "A system of molecular thermochemistry for organic gases and liquids," *Can. J. Chem.*, 34 [5] 626-648 (1956).
 16. V. M. Tatevskii, V. A. Benderskii, S. S. Yarovoi, *Rules and Methods for Calculating the Physicochemical Properties of Paraffinic Hydrocarbons*, Pergamon, Oxford, 1961.
 17. D. V. Sukhachev, T. S. Pivina, F. S. Volk, "Estimation and prediction of the heats of formation for non-aromatic polynitro compounds on the basis of the QSPR approach," *Propell. Explos. Pyrot.*, 19 [4] 159-164 (1994).
 18. A. Vatani, M. Mehrpooya, F. Gharagheizi, "Prediction of standard enthalpy of formation by a QSPR model," *Int. J. Mol. Sci.*, 8 [5] 407-432 (2007).
 19. M. Garbalena, W. C. Herndon, "Optimum graph-theoretical models for enthalpic properties of alkanes," *J. Chem. Inf. Comp. Sci.*, 32 [1] 37-42 (1992).
 20. M. Kuanar, S. K. Kuanar, B. K. Mishra, I. Gutman, "Correlation of line graph parameters with physicochemical properties of octane isomers," *Indian. J. Chem. A*, 38, 525-528 (1999).
 21. S. Liu, S. Cai, C. Cao, Z. Li, "Molecular electronegative distance vector (MEDV) related to 15 properties of alkanes," *J. Chem. Inf. Comp. Sci.*, 40 [6] 1337-1348 (2000).
 22. P. Thanikaivelan, V. Subramanian, J. R. Rao, B. U. Nair, "Application of quantum chemical descriptor in quantitative structure activity and structure property relationship," *Chem. Phys. Lett.*, 323 [1] 59-70 (2000).
 23. A. Mercader, E. A. Castro, A. A. Toropov, "Maximum topological distances based indices as molecular descriptors for QSPR. 4. Modeling the enthalpy of formation of hydrocarbons from elements," *Int. J. Mol. Sci.*, 2 [2] 121-132 (2001).
 24. G. G. Cash, "Heats of Formation of Polyhex Polycyclic Aromatic Hydrocarbons from Their Adjacency Matrixes," *J. Chem. Inf. Comp. Sci.*, 35 [5] 815-818 (1995).
 25. I. B. Golovanov, S. M. Zhenodarova, O. A. Smolyaninova, "Quantitative structure-property relationship: XIII. Properties of aliphatic alcohols," *Russ. J. Gen. Chem.*, 73 [4] 519-524 (2003).

26. L. P. Davis, D. Storch, R. M. Guidry, "MINDO/3, MNDO and AM1 calculations for nitro compounds," *J. Energ. Mater.*, 5 [2] 89-142 (1987).
27. J. J. Stewart, "Optimization of parameters for semiempirical methods I. Method," *J. Comput. Chem.*, 10 [2] 209-220 (1989).
28. D. R. Hwang, M. Tamura, T. Yoshida, N. Tanaka, F. Hosoya, "Estimation of Δf of nitro derivatives of benzene and toluene using AM1 and DSC," *J. Energ. Mater.*, 8 [1-2] 85-98 (1990).
29. Y. Akutsu, S. Y. Tahara, M. Tamura, T. Yoshida, "Calculations of heats of formation for nitro compounds by semi-empirical MO methods and molecular mechanics," *J. Energ. Mater.*, 9 [3] 161-171 (1991).
30. M. Sana, G. Leroy, D. Peeters, C. Wilante, "The theoretical study of the heats of formation of organic compounds containing the substituents CH 3, CF 3, NH 2, NF 2, NO 2, OH and F," *J. Mol. Struc-THEOCHEM*, 164 [3] 249-274 (1988).
31. Y. Akutsu, J. I. Takayama, M. Tamura, T. Yoshida, "Calculations of heats of formation for nitroaromatics with PM3 and MM2," *J. Energ. Mater.* 10 [2-3] 173-182 (1992).
32. M. J. Dewar, W. Thiel, "Ground states of molecules. 38. The MNDO method. Approximations and parameters," *JACS*, 99 [15] 4899-4907 (1977).
33. T. S. Pivina, E. A. Arnautova, A. V. Dzyabehenko, Computer modeling of possible polymorphic transformations in HNIW, in: Proceedings of the ICT, Vol. 39, pp. 1-12, 25-28 June, Karlsruhe, Germany, 1996.
34. K. Jayasurya, Design of new explosives molecules, pp. 12-13, GC-TR-1853, US Army ARDEC, Picatinny Arsenal, NJ, 1993.
35. Y. Akutsu, R. Che, M. Tamura, "Calculations of heats of formation for nitramines and alkyl nitrates with PM3 and Mm2," *J. Energ. Mater.*, 11 [3] 195-203 (1993).
36. R. L. Disch, J. M. Schulman, "Ab initio heats of formation of medium-sized hydrocarbons. 7. The [n] prismanes," *JACS*, 110 [7] 2102-2105 (1988).
37. X. M. Duan, G. L. Song, Z. H. Li, X. J. Wang, G. H. Chen, K. N. Fan, "Accurate prediction of heat of formation by combining Hartree-Fock/density functional theory calculation with linear regression correction approach," *The J. Chem. Phys.*, 121 [15] 7086-7095 (2004).
38. A. Bhattacharya, "Converting ab initio energies to enthalpies of formation of free radicals. I. New atom equivalents for alkyl radicals," *AIChE J.*, 58 [2] 600-609 (2012).
39. A. Nicolaidis, A. Rauk, M. N. Glukhovtsev, L. Radom, "Heats of formation from G2, G2 (MP2), and G2 (MP2, SVP) total energies," *J. Phys. Chem.*, 100 [44] 17460-17464 (1996).
40. E. F. Byrd, B. M. Rice, "Improved prediction of heats of formation of energetic materials using quantum mechanical calculations," *J. Phys. Chem. A*, 110 [3] 1005-1013 (2006).
41. I. Shafagh, K. J. Hughes, M. Pourkashanian, "Modified enthalpies of formation for hydrocarbons from DFT and ab initio thermal energies," *Comp. Theor. Chem.*, 964 [1] 100-107 (2011).
42. D. Habibollahzadeh, M. E. Grice, M. C. Concha, J. S. Murray, P. Politzer, *J. Comput. Chem.* 1995, 16, 654.
43. J. Li, C. K. Kim, *Bull. Korean Chem. Soc.* 2015, 36 1536.