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 1958,\(^4\) 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 Cohen\(^11\) 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:

\[
\Delta H_f (\text{PhNO}_2) = 1 (C_\text{B}-\text{NO}_2) + 5 (C_\text{B}-\text{H}) = 16.9 \text{ kcal/mol} \tag{1}
\]

where the $C_\text{B}-\text{NO}_2$ and $C_\text{B}-\text{H}$ 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. 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 items which in turn, decreases the efficiency of the application. In the parameter set from Pedley, 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, Leal collected 200 molecules, 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, this set minimized the number of the additive items to 14, which increased the efficiency. Compared with parameter set from Tatevskii, 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. In their model, the heats of formation can be calculated using eq (2):

$$\Delta H_f^0 = -98.86 + 1.14\text{SBE} + 49.77 \times 4_{kr} - 437.78c_{mid} + \frac{61.76V_{mid}^2}{N_{at}}$$

$$+ 71.39Fr_1 - 195.44Fr_2 - 2933.25\ln\left(\frac{Fr_3}{N_{at}}\right)$$

(2)

where, SBE is the enthalpy of the molecule without steric hindrance, $4_{kr}$ the Randic’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. $F_{r1}$ sum of the minimum charges at atoms in fragments of the type –C–C–NO$_2$, $F_{r2}$ the minimum of all the minimum charges
at atoms in fragments of the type At–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. The 5 descriptors utilized in this method are number of non-H atoms ($n_{SK}$), sum of conventional bond orders ($SCBO$), number of oxygen atoms ($n_{O}$), number of fluorine atoms ($n_{F}$), and number of heavy atoms ($n_{HM}$).

\[
\Delta H_f^0 = 50.1688 - 80.52012n_{SK} + 5364546SCBO - 169.21889n_{O} \\
- 174.75477n_{F} - 266.57659n_{HM}
\] (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. 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 \textit{ab initio} methods. In this area, there were some benchmark studies by many researchers\textsuperscript{26-35}. Disch and Schulman performed \textit{ab initio} calculations on the [n]prismanes.\textsuperscript{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 6-311G(d,p) basis set, the mean absolute deviation was 2.4 kcal/mol.\textsuperscript{37} Bhattacharya converted \textit{ab initio} energies to enthalpies for free radicals species.\textsuperscript{38} Nicolaides and coworkers transferred G2, G2(MP2), and G2(MP2,SVP) total energies to heats of formation.\textsuperscript{39} Byrd and Rice combined QM calculation and derived additive parameters to predict heats of formation.\textsuperscript{40} Shafagh derived a three-parameter equation based on DFT and \textit{ab initio} thermal energies.\textsuperscript{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.\textsuperscript{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, \textit{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.

\textit{References}

6. S. Bourrasseau, “A systematic procedure for estimating the standard heats of formation in the