



## Prediction of reactive hazards based on molecular structure

S.R. Saraf, W.J. Rogers, M.S. Mannan\*

*Mary Kay O'Connor Process Safety Center (MKOPSC), Department of Chemical Engineering,  
Texas A&M University, College Station, TX 77843-3122, USA*

Received 8 February 2002; received in revised form 25 October 2002; accepted 29 October 2002

---

### Abstract

There is considerable interest in prediction of reactive hazards based on chemical structure. Calorimetric measurements to determine reactivity can be resource consuming, so computational methods to predict reactivity hazards present an attractive option. This paper reviews some of the commonly employed theoretical hazard evaluation techniques, including the oxygen-balance method, ASTM CHETAH, and calculated adiabatic reaction temperature (CART). It also discusses the development of a study table to correlate and predict calorimetric properties of pure compounds. Quantitative structure–property relationships (QSPR) based on quantum mechanical calculations can be employed to correlate calorimetrically measured onset temperatures,  $T_o$ , and energies of reaction,  $-\Delta H$ , with molecular properties. To test the feasibility of this approach, the QSPR technique is used to correlate differential scanning calorimeter (DSC) data,  $T_o$  and  $-\Delta H$ , with molecular properties for 19 nitro compounds.

© 2003 Elsevier Science B.V. All rights reserved.

*Keywords:* Reactive hazards; QSPR; Onset temperature; CHETAH; DSC; CART

---

### 1. Introduction

A question that is frequently encountered during processing, storage, or handling of a chemical is “Does this chemical pose reactive hazards?” An analogous problem is the estimation of chemical incompatibilities. Certain compositions are unstable by nature and susceptible to violent reaction in which chemical energy is converted to heat or mechanical energy that are potential hazards.

Generally, rules of thumb based on prior experience and chemical knowledge are used for screening and estimating reactive hazards of compounds. For example, the presence of a

---

\* Corresponding author. Tel.: +1-979-862-3985; fax: +1-979-458-1493.

*E-mail address:* mannan@tamu.edu (M.S. Mannan).

'nitro' group is regarded as an indicator of potential energy, as in trinitrotoluene (TNT). Attempts have been made to develop a generalized framework for estimating reactive hazards based on molecular structure, such as the oxygen-balance method [1], chemical thermodynamic and energy release evaluation (CHETAH) [2], and calculated adiabatic reaction temperature (CART) [3]. However, these methodologies have limitations, and considerable chemical intuition and experience are required for their effective use. Also, the reliability of estimations for a range of compounds and process conditions varies significantly [4].

A reliable experimental technique for assessing reactivity is calorimetric analysis, which can be resource consuming and thus, possible only for a limited set of compounds. So, one of the challenges is extension of calorimetric tests on these limited set of compounds to a myriad of other compounds and compositions. This paper provides a brief review of computational methods that can be employed quickly to estimate reactive hazards and a description of our recent efforts to develop a method for correlating and predicting calorimetric data.

## 2. Review of available methods

This section reviews some popular methods for reactivity hazard evaluation, including their strengths and limitations, and it attempts to provide the reader with enough information to choose a method for a particular application. Some of these methods have been reviewed previously [3].

### 2.1. Rule of thumb

The presence of certain functional groups is considered an indicator of reactivity. This is the simplest possible reactivity screening method and serves as a guideline for further analysis. For example, chemicals containing the following functional groups can be considered potentially reactive:

- NO<sub>2</sub>: organic nitro compounds;
- O–O–, –O–OH : organic/inorganic peroxide and hydroperoxide compounds;
- C≡C–: triple bonded carbon atoms as in acetylene and acetylenic compounds.

A comprehensive summary of reactive groups can be found in Bretherick's handbook [5].

### 2.2. Oxygen-balance method

A quantitative correlation was shown between oxygen-balance and various measures of explosive effectiveness for several classes of more than 300 compounds organic explosives [1], and the following formula was recommended for calculating oxygen-balance for a compound:

$$\text{oxygen balance} = \frac{-1600(2X + Y/2 - Z)}{\text{MW}} \quad (1)$$

Table 1  
Oxygen-balance and hazard rank

Oxygen-balance	Hazard rank
More positive than +160	Low
+160 to +80	Medium
+80 to –120	High
–120 to –240	Medium
More negative than –240	Low

where  $X$  is the number of atoms of carbon,  $Y$  the number of atoms of hydrogen,  $Z$  the number of atoms of oxygen, and  $MW$  the molecular weight.

The above formula yields a value of zero for oxygen-balanced compounds, negative for oxygen-poor, and positive for oxygen-rich compositions. This method is a criterion used in the CHETAH program, and the classification indicated in Table 1 is recommended for estimating hazard potential based on oxygen-balance [6].

### 2.2.1. Strengths

The oxygen-balance method is useful for estimating hazards of organic nitro compounds and is universally employed in the explosive industry. In general, this method is applicable to compounds containing C, H, N, and O [6].

### 2.2.2. Weaknesses

However, it has been shown that there is no necessary connection between oxygen-balance and self-reactivity [6]. For example, water ( $H_2O$ ) has an oxygen-balance value of zero (0) and is given a ‘high’ hazard ranking by this criterion. Also, the method cannot be applied to oxygen-free but hazardous compounds such as acetylene. Application of the above oxygen-balance equation to low-oxygen content or oxygen-free compounds produces a highly negative, non-hazardous ranking regardless of the actual hazard potential.

## 2.3. CHETAH

CHETAH is a popular program by the American Society for Testing and Materials (ASTM) for prediction of reactivity hazards [2]. The software uses the ‘Benson group contribution method’ [7] to estimate heat capacity, heat of formation, and heat of combustion for a multitude of compounds. Also, the program includes a database of thermo-chemical properties for selected organic and inorganic compounds. CHETAH classifies chemicals based on their potential for violent explosion and includes the following hazard evaluation criteria:

1. Maximum heat of decomposition.
2. Oxygen-balance.

According to the first criterion, compositions with heats of reaction more negative and therefore, more exothermic than  $-2.929$  kJ/g are placed in a ‘high’ hazard category. A detailed explanation of the above evaluation criteria and three additional ones can be found in the CHETAH reference manual, and a critical review of CHETAH for predicting reactivity

hazards is available [8]. The first criterion has proved to be a reliable indicator of potential reactive hazards. The other criteria, however, are not effective for all chemicals and compositions [8].

### 2.3.1. Input to the program

The molecular formula of a compound is the only input to the program. From the included database, the thermodynamic properties are estimated and the hazard criteria are determined from these values. An energy release evaluation sheet from CHETAH 7.2 is shown in Fig. 1.

## ENERGY RELEASE EVALUATION

**Compound Name:** Hydrogen Peroxide  
**Formula:** H<sub>2</sub>O<sub>2</sub>  
**Molecular weight:** 34.015  
**Amount:** 1 Mole(s)  
**Heat of formation at 25 °C:** -32.530 kcal/mol

Plosive Hazard Classification:

**Over-all Energy Release Potential is HIGH**

**Value: -1.375**

Contributing Details:

Criterion	Value	Units	Hazard Classification
Maximum Heat of Decomposition (#2)	-0.743	kcal/g	HIGH
Fuel Value – Heat of Decomposition	0.000	kcal/g	HIGH
Oxygen Balance (#3)	47.037	percent	HIGH
CHETAH ERE Criterion 4	46.934	kcal <sup>2</sup> /mol	MEDIUM
Total Number of Peroxide Bonds	1.000		
Net Plosive Density (#4)	0.435		PLOSIVE

Warning: These ratings only apply to hazards associated with strong mechanical shock. This does not imply the absence of other hazards.

- Notes #1 This evaluation was developed to classify a composition as able or not able to decompose with violence, if subjected to the proper conditions. Information on the interpretation of hazard classification criteria used by CHETAH may be found in the CHETAH documentation. (ASTM publication DS-51A) and in J. Chem. Ed., v66, A137 (1989).
- #2 For decomposition products shown.
- #3 Experience has shown that the oxygen balance criterion is useful only for compounds composed of the elements C, H, N, and O.
- #4 Sum of auxoplosive and plosphoric weights per gram of mixture.

Decomposition Products (chosen to maximize heat of decomposition)

Moles	State	Species
0.500	ref-gas	O <sub>2</sub> Oxygen
1.000	gas	H <sub>2</sub> O Water

### Heat of Combustion Section

Fuel Value (Net Heat of Combustion)

Mass Basis	Mole Basis
-0.743 kcal/g	- 25.270 kcal/mol
-3.108 kJ/g	-105.730 kJ/mol
-1337.246 Btu/lb	

Combustion Products (Chosen for Fuel Value and Net Heat of Combustion)

Moles	State	Species
1.000	gas	H <sub>2</sub> O Water

Fig. 1. Hazard evaluation and ranking for hydrogen peroxide using CHETAH 7.2.

Based on its maximum heat of decomposition,  $\text{H}_2\text{O}_2$  is given a “high” hazard classification. However, it should be noted that the estimated product spectrum might be incorrect. Thus, one problem is the thermodynamic feasibility of a proposed stoichiometry under process conditions. Further, the program gives no indication about the sensitivity to initiation of reaction or process conditions to be avoided. It is difficult to determine conditions under which  $\text{H}_2\text{O}_2$  may pose reactive hazards based only on such an analysis. Thus, the problem of reactivity is not just a combinatorial problem (stoichiometric analysis) as implicitly suggested by this method. It is important to realize that CHETAH provides an estimate of a material hazard but not an estimate of the process risk in using the material.

### 2.3.2. Strengths

The software is user-friendly and offers the flexibility to include user-defined group values. It is computationally inexpensive and can be installed on a PC.

### 2.3.3. Weaknesses

The Benson method can fail for group values that are not available in the database or are incorrect. In the evaluation criteria, the program classifies compounds or compositions into specific hazard categories. Thus, the program places a boundary based on a threshold value and thereby blurs the distinction between hazardous and non-hazardous chemicals. Also, the program provides no insight into process conditions to be avoided or information about the sensitivity of compounds to initiation of a reaction.

## 2.4. CART

Adiabatic temperature rise due to a reaction is defined as [3]:

$$\Delta T_{\text{adiabatic}} = \frac{-\Delta H}{C_p} \quad (2)$$

where  $\Delta H$  is the heat of reaction,  $C_p$  the average heat capacity of the reacting mixture,  $\Delta T_{\text{adiabatic}}$  the adiabatic temperature rise.

The code developed for CART performs multiphase Gibbs free energy minimization and adiabatic reaction temperature calculations. Based on the calculated  $\Delta T_{\text{adiabatic}}$ , the substance is classified as [3]:

- E—can explode when unconfined;
- N—no known explosion hazard when unconfined.

An adiabatic temperature rise of 1400 K is considered a cut-off value for the above classification. Thus, substances with an adiabatic temperature rise of more than 1400 K are classified as E and lower than 1400 K as N. This value is based on the fact that most combustion reactions leading to formation of  $\text{CO}_2$  and  $\text{H}_2\text{O}$  have a threshold temperature value near 1400 K, which is the minimum temperature required for carbon monoxide to propagate a self-sustaining flame. A cut-off value of 1200 K for conservative estimates is recommended [3].

The heat of reaction can be approximated by the maximum heat of decomposition that is calculated by CHETAH. For  $\text{H}_2\text{O}_2$ , CHETAH estimates a maximum heat of decomposition

of  $-25.27$  kcal/mol and  $C_p$  as  $10.31$  cal/mol K, yielding an adiabatic temperature rise of  $2450$  K.

Based on the above discussion,  $H_2O_2$  would be classified as E. Again, this value is useful but it does not provide information about process conditions to be avoided. It is worth noting that for  $H_2O_2$ , the CHETAH criteria and the adiabatic temperature value indicate a potential reactive hazard.

#### 2.4.1. Input to the CART method

The heat of reaction, heat capacity of the reaction mixture, and other system thermochemical values are required program inputs. All of these values can be obtained from the literature or estimated using the CHETAH program.

#### 2.4.2. Strengths

The CART criterion takes into account the heat capacity of the reaction mixture and is therefore, more effective than employing only the reaction energy of the first CHETAH criterion. Higher CART values are associated with greater sensitivities to initiation and higher propagation rates. CHETAH programmers should probably include 'adiabatic temperature rise' as one of the hazard evaluation criteria.

#### 2.4.3. Weaknesses

Thermo-physical values, such as  $\Delta H$  and  $C_p$ , must be used to estimate the adiabatic rise in temperature. The heat of reaction can be approximated by the maximum heat of decomposition calculated by CHETAH, and CHETAH can also be used to estimate the average  $C_p$  value. If the CHETAH program is employed to calculate thermo-physical values, limitations similar to the group contribution method are encountered, as discussed above. The classification based on adiabatic temperature rise works well for hazard estimation of compounds undergoing combustion reactions. For compounds that do not undergo combustion type reactions, determining a threshold value of  $\Delta T_{\text{adiabatic}}$  is difficult. As pointed out in the same reference [3], the CART classification and heats of reaction values fail for hazard ranking of organic peroxides. Also, like CHETAH, this classification places a distinct boundary based on a threshold value. It is also worth noting that for a realistic calculation of  $\Delta T_{\text{adiabatic}}$ , an average value of heat capacity for the reaction mixture is required, and accurate heat capacity values are difficult to estimate.

### 3. Modeling reactivity

A reactive incident involves conversion of stored chemical energy of the components and can thus, be symbolically represented as:



It is the uncontrolled release of this stored energy that is hazardous. Information about the amount of energy released and the rate of energy released for a process chemical can be obtained by performing calorimetric tests on the chemical. There are many different calorimeters available for screening or detailed testing. Before a detailed calorimetric

analysis is performed, the chemicals are screened by employing relatively less resource consuming techniques such as a differential scanning calorimeter (DSC).

Typically in a DSC a small amount of sample (1–20 mg) is placed in a capsule and heated at a constant rate (1–10 °C/min) from room temperature to 400 °C. In a DSC, a sample and a reference are subjected to a continuously increasing temperature, and heat is added to the reference to maintain it at the same temperature as the sample. This added heat compensates for heat lost or gained as a consequence of an overall endothermic or exothermic reaction.

When the rate of heat generation ( $W$ ) in the sample exceeds a particular value, the heat supply to the sample is removed, and this additional heat gain is attributed to exothermic activity within the sample. This detection of heat generation above a threshold value depends on the sensitivity of the particular instrument. The temperature at which chemical reaction causes a detectable rise in heat generation is called the onset temperature ( $T_{\text{onset}}$ ), which is a measure of the reaction kinetics and serves as a guide for selecting process or storage temperatures. The  $T_{\text{onset}}$  is determined following a graphical procedure to detect either the first deviation from the baseline or the intersection of the baseline and a tangent to the rise in supplied heat. The energy released ( $-\Delta H$ ) during the process is calculated as the area under the heat-supplied ( $W$ ) and time curve. Thus, the onset temperature ( $T_{\text{onset}}$ ) and the energy of reaction ( $-\Delta H$ ) are two important parameters obtained from a DSC run. The onset temperature indicates the temperature at which the reaction is initiated (as detected by the calorimeter), and the energy of reaction is a measure of the magnitude of an exotherm. A sample DSC run is shown in Fig. 2.

But a reactivity assessment cannot be achieved by experimental methods alone. For example, the two hydrogen atoms in hydrogen peroxide (HO–OH) can be replaced by any of the following 10 groups: H, C1–C6 (aliphatic chains), benzene, toluene, or xylene. Thus, 100 different peroxide compounds must be evaluated to obtain the reactivity data for the possible compounds containing the –O–O– linkage. Assuming two DSC runs per compound (at least) and 2 h per calorimetric run, the analysis would take at least 200 h. But there are many

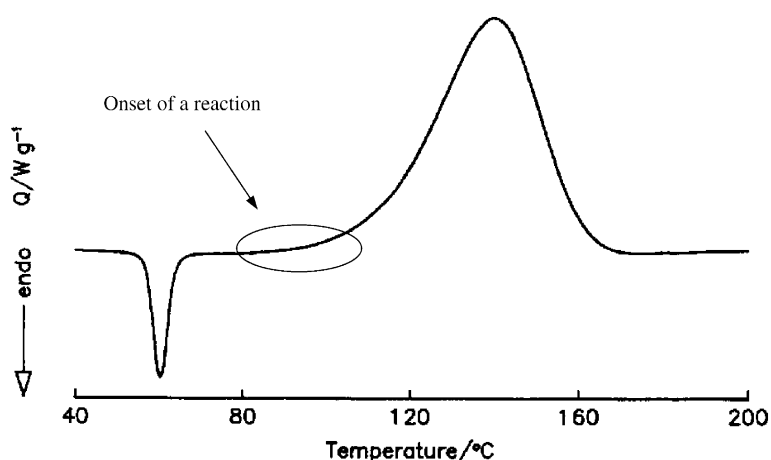


Fig. 2. DSC run [26].

different possible combinations of functional groups (e.g. alkenes, alkynes, isomers) and compositions (e.g. 10, 20, 80% H<sub>2</sub>O<sub>2</sub>). Sometimes experimentation is not feasible because of the toxic nature of chemicals and products or if the compound is not available or does not exist as in product design. Experiments with more sophisticated calorimeters such as the vent size package (VSP) (<http://www.fauske.com>), the automated pressure tracking adiabatic calorimeter (APTAC) (<http://www.adlittle.com>), or the accelerating rate calorimeter (ARC) [9] require more time, larger samples, and greater efforts.

The computational methods cannot be employed independently to estimate reactive hazards but must be implemented synergistically with experimental measurements to reduce experimental effort and to expedite reactive hazard evaluations. The primary objective is to extend and complement available experimental data with predictions for the species where no or few experimental data exist. Also, theoretical methods avoid the risks of experimental testing. The following section describes another approach to predict calorimetric data, by bridging the gap between macroscopic (calorimetric data) and microscopic (molecular) properties.

#### 4. Structure based screening tool

The exothermic behavior of a substance is influenced by the presence of functional groups [10–12], which also form a basis for reactivity classification. This approach suggests that there is an inherent structure–property relationship between the observed calorimetric properties and molecular structure. For example, as mentioned earlier, the presence of nitro group (NO<sub>2</sub>) can be considered a potential source of significant reactivity. The reactive nature is manifested in calorimetric data, but this dependence of observed behavior and molecular structure has not been quantified.

The quantitative structure–property relationship (QSPR) is a popular tool for correlating observed values based on molecular properties. QSPR techniques have been successfully employed for drug design [13] and for correlating physical properties such as boiling point [14], autoignition temperature [15], and molecular properties. In addition to providing a means of predicting properties, a QSPR study may also lead to better understanding of structural features affecting the observed data. Our objective here is to correlate and predict DSC calorimetric data, namely onset temperature and energy of reaction.

- (a) Onset temperature ( $T_{\text{onset}}$ ): The onset temperature represents detection of a chemical reaction and is a measure of potential energy and kinetics of the reaction. It is indeed a ‘detected’ onset because the value depends on the sensitivity of the instrument and experimental technique. Depending on the type of calorimeter, sample size, and scanning rate,  $T_{\text{onset}}$  can vary within 5–50 °C for the same compound.
- (b) Energy of reaction ( $-\Delta H$ ): The energy of reaction, often due to decomposition or polymerization, is the net heat released during a reaction. This energy is not the thermodynamic heat of reaction, because it includes other effects such as evaporation and enthalpy of mixing, and heat absorbed by the sample cell.

We built a quantitative property–structure relationship (QSPR) study table to develop the correlation. The first column of this table is either  $T_{\text{onset}}$  or energy of reaction obtained from



calorimetric experiments and is called the dependent variable. The remaining columns are the independent variables (characteristic of the molecules) called descriptors, which are characteristics of a molecule and account for the chemical structure of the molecule. A descriptor value can be obtained by experimental measurement or calculated based on molecular structure. We have used calculated descriptors to facilitate property predictions for unknown molecules.

#### 4.1. Data set selection

We chose compounds belonging to the organic nitro family since a better correlation of properties is expected within a family of similar compounds. This set of compounds used to develop the correlation is called a ‘training set’. The choice of data is critical for an effective correlation, and we chose data from a single reference [16] to maintain consistency in experimental procedure and calorimetric sensitivity. With a larger data set a predictive model could be developed, but we could not find a large set of consistent data in the open literature.

#### 4.2. Discussion of a few descriptors

The choice of descriptors depends on the property to be correlated. Here we chose descriptors that were expected to correlate with the detected onset temperature or determined energy of reaction.

- (a) Highest occupied molecular orbital (HOMO): The HOMO is the highest ground state energy level in the molecule that contains electrons. Molecules with high energy HOMOs are more easily able to donate their electrons and are relatively reactive compared to molecules with lower HOMO values.
- (b) Lowest unoccupied molecular orbital (LUMO): The LUMO is the lowest ground state energy level in the molecule that contains no electrons. Molecules with lower LUMO values are able to accept electrons more easily than those with higher LUMO values.
- (c,d) Highest positive charge (HPC) and highest negative charge (HNC): These descriptors are probable indicators of electrophilic or nucleophilic attacks and are expected to correlate with the  $T_{\text{onset}}$ , which is a measure of kinetics.
- (e) Weakest bond (WB): The weakest bond, here the C–NO<sub>2</sub> bond, in a molecule is a kinetic descriptor, since it represents the minimum activation energy required for initiating the reaction.
- (f) Mid-point potential ( $V_{\text{mid}}$ ): Previous work [17] has shown that the electrostatic potential produced by carbon and nitrogen charges at the C–NO<sub>2</sub> bond mid-point correlates with the impact sensitivity.

$$V_{\text{mid}} = \frac{Q_{\text{C}}}{0.5R} + \frac{Q_{\text{N}}}{0.5R} \quad (3)$$

where  $Q_{\text{C}}$  is the atomic charge on carbon,  $Q_{\text{N}}$  the atomic charge on nitrogen,  $R$  the C–NO<sub>2</sub> bond distance.

- (g) Delocalizability index (Sr): Here the descriptor (Sr) is defined as

$$Sr = \sum_{i=1}^{\text{HOMO}} \frac{1}{|\varepsilon_i|} \quad (4)$$

where  $\varepsilon_i$  is the eigenvalue for the molecular orbital.

The Sr index is similar to Fukui's superdelocalizability index [18]. In the summation, the eigenvalues for higher molecular orbitals (which are numerically smaller) will dominate, and Sr is therefore, a measure of delocalizability of electrons and a probable measure of reactivity.

- (h) Dipole moment: The degree of polarity of a molecule is expressed in terms of dipole moment. The electric dipole moment for a pair of opposite charges is defined as the magnitude of the charge times the distance between them, and the defined direction is toward the positive charge. It is useful in atoms and molecules where the effects of charge separation are measurable, but the distances between the charges are too small to be easily measured. The dipole moment of the molecule was also included as one of the descriptors.
- (i) Charge–bond strength descriptor ( $x$ ): This descriptor was calculated as follows:

$$x = \frac{\text{HPC} + \text{HNC}}{0.5\text{WB}} \quad (5)$$

and is an indicator of charge to bond strength ratio in a molecule.

Values of these descriptors were obtained using the Gaussian [19] suite of programs with the B3P86 [20] (Becke 3 Perdew-Wang 86 (B3P86)) density functional model and the cc-pVDZ [21] basis set. To develop a computationally inexpensive method, more sophisticated models were not tested at this stage of the project. Further statistical calculations were performed using the statistical analysis system (SAS) [22]. The details of the descriptors values used for developing the correlation's are available on request.

### 4.3. Correlation

The correlations were obtained by performing least square regression analysis on the training set of molecules. The developed correlation has the form:

$$Y = A_1X_1 + A_2X_2 + A_3X_3 + \cdots + A_nX_n$$

where  $Y$  is the dependent variable,  $X$  the independent variable (descriptor),  $A$  the regression constant.

#### 4.3.1. $T_{\text{onset}}$

The details of the experimental onset temperatures and predicted values used are summarized in Table 2. A statistical analysis was performed on the all descriptors and their combinations, and the chosen descriptors were the ones that maximized the  $R^2$  value and minimized the error square terms [23]. As a result of this analysis the variables that exhibited significant correlation with the dependent variable ( $T_{\text{onset}}$ ) were retained, and the statistically insignificant ones were discarded.

Table 2  
Onset temperatures, observed and predicted

S. no.	Structure	$T_{\text{onset}}$ ( $^{\circ}\text{C}$ )		Residual (Experiment–predicted)	Error (%)
		Experiment	Predicted		
1	Nitrobenzene	380	348	32	8
2	1,2-Dinitrobenzene	280	297	–17	–6
3	1,3-Dinitrobenzene	270	304	–34	–12
4	1,4-Dinitrobenzene	350	328	22	6
5	2-Nitrotoluene	290	309	–19	–7
6	3-Nitrotoluene	310	315	–5	–2
7	4-Nitrotoluene	320	311	9	3
8	2,6-Dinitrotoluene	290	280	10	3
9	3,4-Dinitrotoluene	280	261	19	7
10	2,4-Dinitrotoluene	250	260	–10	–4
11	2-Nitroaniline	280	256	24	9
12	3-Nitroaniline	300	302	–2	–1
13	4-Nitroaniline	310	279	31	10
14	2-Nitrobenzoic acid	270	281	–11	–4
15	3-Nitrobenzoic acid	300	296	4	1
16	4-Nitrobenzoic acid	310	295	15	5
17	2-Nitrophenol	250	274	–24	–10
18	3-Nitrophenol	310	316	–6	–2
19	4-Nitrophenol	270	309	–39	–14

Table 3  
Experimental energy of reaction

S. no.	Structure	Energy of reaction ( $-\Delta H$ )	
		J/g	kcal/mol
1	Nitrobenzene	2757	81.1
2	1,2-Dinitrobenzene	3310	133.0
3	1,3-Dinitrobenzene	3488	140.1
4	1,4-Dinitrobenzene	3701	148.7
5	2-Nitrotoluene	2404	78.8
6	3-Nitrotoluene	2070	67.9
7	4-Nitrotoluene	2322	76.1
8	2,6-Dinitrotoluene	3451	150.2
9	3,4-Dinitrotoluene	3574	155.6
10	2,4-Dinitrotoluene	3987	173.6
11	2-Nitroaniline	2225	73.5
12	3-Nitroaniline	2269	74.9
13	4-Nitroaniline	2026	66.9
14	2-Nitrobenzoic acid	1894	75.6
15	3-Nitrobenzoic acid	1899	75.8
16	4-Nitrobenzoic acid	1934	77.2
17	2-Nitrophenol	2481	82.5
18	3-Nitrophenol	2269	75.4
19	4-Nitrophenol	2155	71.7

We recommend the following correlation based on the training set of 19 nitro compounds:

$$T_{\text{onset}} (^{\circ}\text{C}) = 827.0 - 1036 \times \text{HPC} - 4.43 \times \text{Sr} - 5.07 \times \text{dipole} \quad (6)$$

A standard overall  $F$ -test ( $\alpha = 0.05$ ) indicates that the fitted correlation is significant and not a chance correlation. The predicted onset temperature values with an absolute average aggregate error of 6% and a bias of  $-0.5\%$  are listed in Table 2 and are plotted against the experimental values in Fig. 3. A correlation of 0.6 is obtained between the predicted and the observed values. This level of correlation is reasonable given significant variations in the experimentally determined onset temperatures due to the graphical detection procedure, as discussed above, and other associated errors.

#### 4.3.2. Energy of reaction

It is observed that the energy of reaction values correlate strongly with the count of  $-\text{NO}_2$  in the nitro compounds [24], which is consistent with  $\text{NO}_2$  as an indicator of reactivity. The experimental heat of reaction values for all the 19 compounds are noted in Table 3 and summarized for the nitro and dinitro compounds in Table 4. Thus, the correlation for the

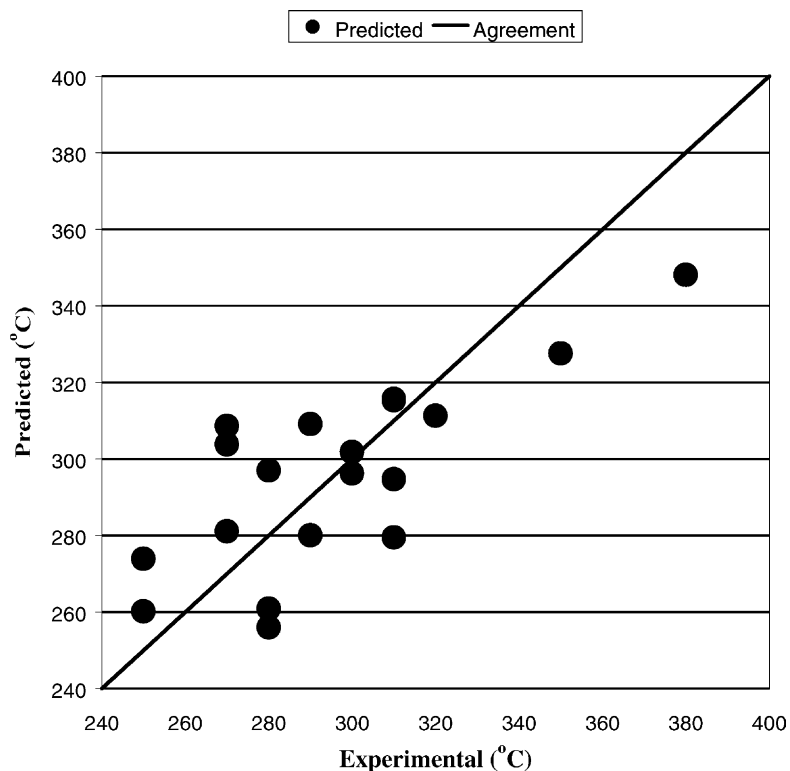


Fig. 3. Predicted onset temperatures.

Table 4  
Summary of average energy of reaction ( $-\Delta H$ ) values

Compounds	$-\Delta H$ (kcal/mol)
Mononitro <sup>a</sup>	$75 \pm 5$
Dinitro <sup>b</sup>	$150 \pm 14$

<sup>a</sup> Statistical analysis was performed on 1, 5–7, 11–19 compounds in Table 3.

<sup>b</sup> Statistical analysis was performed on 2–4, 8–10 compounds in Table 3.

energy of reaction in (kcal/mol) is:

$$-\Delta H = 75 \times \text{number of nitro groups} \quad (7)$$

Thus, for TNT (three nitro groups) the predicted energy of reaction is 225 kcal/mol, which is consistent with the experimentally determined value of 239 kcal/mol [10].

## 5. Correlation using a semi-empirical method

The descriptors for the above study were generated using the B3P86/cc-pVDZ model. Typically an optimization for an aromatic nitro molecule using this model requires about an hour of CPU time on the supercomputer. Therefore, for the descriptors to be easily calculable it is important that predictions be possible using the computationally inexpensive semi-empirical theory, such as AM1 [25].

An optimization using AM1 can be performed in few seconds. To use the correlation's generated earlier we must scale the AM1 descriptor values to the B3P86/cc-pVDZ values. We found that the Sr and dipole descriptors calculated using the AM1 model correlated with the B3P86/cc-pVDZ values to yield  $R^2$  values of 0.98 and 0.82, respectively. However, the HPC values calculated using the two models did not show a good correlation. We recommend the following equations to scale the Sr and dipole moment:

$$Sr_{B3P86} = 1.31 Sr_{AM1} - 0.11 \quad (8)$$

$$\text{dipole}_{B3P86} = 0.93 \text{dipole}_{AM1} - 0.21 \quad (9)$$

With the scaled descriptors, the following three-parameter correlation can be employed for calculating  $T_{\text{onset}}$ :

$$T_{\text{onset}} (^{\circ}\text{C}) = 828.5 - 5.80 \times Sr_{AM1} - 4.73 \times \text{dipole}_{AM1} \quad (10)$$

Using the AM1 scaled descriptors and the above two-parameter correlation, predictions for onset temperatures for the 19 compounds yielded an average absolute aggregate error of 7% and a bias of  $-1\%$ , which is in good agreement with the predictions obtained using the more expensive B3P86 descriptors.

## 6. Conclusions

Computational screening methods are needed for efficient evaluation of reactive hazards. The earlier methods attempted to classify the substances into ‘hazard’ categories. We have demonstrated that it is possible to correlate the experimental DSC thermokinetic data with molecular structure based descriptors for the nitro family of compounds. The descriptors discussed earlier are general can be used to correlate the onset temperature for any system of compounds. By employing large data sets for a variety of compounds, a predictive model could be developed. The QSPR technique appears to provide a realistic estimation of hazards and can be further improved by using a larger training set and more sophisticated quantum chemical models. However, for the approach to be practical at the industrial level, it is important that the developed correlation’s yield reasonable results with simpler models, such as the semi-empirical AM1.

In this work, the onset temperatures and energies of reaction obtained from calorimetric data were correlated with molecular descriptors, but similar methodology can be applied to develop correlation’s for other experimental parameters such as self-accelerating decomposition temperature (SADT), time to maximum rate (TMR), and kinetic rate constants. We are collaborating with Dow Chemical and Eastman Kodak to obtain larger sets of calorimetric data, and we are testing reactivity descriptors for diverse classes of chemicals. Our objective is to develop correlation’s that would yield satisfactory results with computationally inexpensive theories that can be applied by industry personnel.

## Acknowledgements

We thank David Frurip and Peter Margl from Dow Chemical for their valuable input. We are thankful to the Laboratory for Molecular Simulation (LMS, <http://www.chem.tamu.edu/LMS>), Texas A&M University, for the use of software and computer time and to Lisa M. Thomson, Manager of the LMS, for training in the use of software and for stimulating discussions. The Gaussian computations were performed on the Texas A&M supercomputer, and we acknowledge the technical support of the TAMU supercomputing centre. This research was sponsored by the Mary Kay O’Connor Process Safety Center (<http://process-safety.tamu.edu>).

## References

- [1] W.C. Lothrop, G.R. Handrick, *Chem. Rev.* 44 (1949) 419.
- [2] CHETAH<sup>TM</sup>, Version 7.2, The ASTM Computer Program for Chemical Thermodynamic and Energy Release Evaluation (NIST Special Database 16), ASTM Subcommittee E27.07, ASTM, West Conshohocken, PA, 1998.
- [3] G.A. Melhem, E.S. Shanley, *Process Safety Prog.* 15 (3) (1996) 69.
- [4] D.J. Frurip, A. Chakrabarti, J.R. Downey, H.D. Ferguson, S.K. Gupta, T.C. Hoeflich, M.S. LaBarge, A.J. Pasztor, L.M. Peerey, S.M. Eveland, R.A. Suckow, in: *Proceedings of the International Symposium on Runaway Reactions and Pressure Relief Design*, Boston, MA, AIChE, 1995, pp. 95–149.
- [5] P.G. Urben, *Bretherick’s Handbook of Reactive Chemicals*, 5th ed., Butterworth, London, UK, 1995.
- [6] G.A. Melhem, E.S. Shanley, *Process Safety Prog.* 14 (1) (1995) 29.

- [7] S.W. Benson, *Thermochemical Kinetics: Methods for the Estimation of Thermochemical Data and Rate Parameters*, 2nd ed., Wiley, New York, 1976.
- [8] G.A. Melhem, E.S. Shanley, *J. Loss Preven. Process Ind.* 8 (5) (1995) 261.
- [9] D.I. Townsend, J.C. Tou, *Thermochim. Acta* 37 (1980) 1.
- [10] T. Grewer, *Thermal Hazards of Chemical Reactions*, Elsevier, Amsterdam, Netherlands, 1994.
- [11] T. Ando, Y. Fujimato, S. Morisaki, *J. Hazard. Mat.* 28 (1991) 251.
- [12] T. Grewer, *Thermochim. Acta* 187 (1991) 133.
- [13] D.C. Young, *Computational Thermochemistry: A Practical Guide for Applying Techniques to Real World Problems*, Wiley, New York, 2001.
- [14] M.D. Wessel, P.C. Jurs, *J. Chem. Inf. Comput. Sci.* 35 (1995) 841.
- [15] B.E. Mitchell, P.C. Jurs, *J. Chem. Inf. Comput. Sci.* 37 (1997) 538.
- [16] Y. Duh, C. Lee, C. Hsu, D. Hwang, C. Kao, *J. Hazard. Mat.* 53 (1997) 183.
- [17] P. Politzer, J.S. Murray, *J. Mol. Struct.* 376 (1996) 419.
- [18] H. Hosoya, S. Iwata, *Theor. Chem. Acc.* 102 (1999) 293.
- [19] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, *Gaussian 98 (Revision A.9)*, Gaussian Inc., Pittsburgh PA, 1998.
- [20] (a) A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648;  
(b) J.P. Perdew, *Phys. Rev. B* 33 (1986) 8822.
- [21] T.H. Dunning Jr., *J. Chem. Phys.* 90 (1989) 1007.
- [22] *The SAS System for Windows*, Release 8.01, copyright(c), SAS Institute Inc., Cary, NC, USA, 1999–2000.
- [23] D.C. Montgomery, G.C. Runger, *Applied Statistics and Probability for Engineers*, 2nd ed., Wiley, 1999.
- [24] S. Chervin, G.T. Bodman, in: *Proceedings of the 29th North American Thermal Analysis Society (NATAS)*, St. Louis, MO, 24–26 September 2001.
- [25] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, *J. Am. Chem. Soc.* 107 (1985) 3902 (Austin Model 1).
- [26] D. Jones, R. Augsten, *Thermochim. Acta* 286 (1996) 366.