Organic Process Research & Development **2008**, *12*, 1287–1292

Full Papers

Selection of the Proper Calorimetric Test Strategy in Reactive Chemicals Hazard Evaluation

David J. Frurip

The Dow Chemical Company, Environmental Technology Center/Analytical Support/Thermal Laboratory, Building 1897F, Midland, Michigan 48667, U.S.A.

Abstract:

The job of a reactive chemicals testing subject matter expert (a so-called SME) is to effectively and accurately evaluate a chemical process for known or potential reactive chemical hazards such as thermal stability, runaway potential, and flammability−dust explosion hazards. This evaluation allows designers to ensure the process operates within a safe operating envelope, with minimum risk of a "significant" reactive chemicals event/incident. This paper will explore the decision making process for developing a successful testing strategy based on scale (i.e., the size of the process) and energy release potential. Examples will be given which demonstrate the application of this strategy to modern chemical production. We will focus on the decision process with regard to thermal hazards which in turn primarily use calorimetric test methods. The main calorimetric methods discussed are differential scanning calorimetry (DSC), accelerating rate calorimetry (ARC), heat of mix calorimetry (HOM), and reaction calorimetry.

1. Introduction

A reactive chemicals hazard assessment is the identification and quantification of the hazardous energy-release scenarios of a process. This is typically accomplished by calculations and experimental testing.[1−8] To carry out an efficient, accurate, and meaningful reactive chemicals hazard assessment, the SME needs to devise an equally efficient, yet accurate, calculational and experimental strategy. First the SME will evaluate the known properties of the materials, then perform various calculations (such as heats of reaction), and finally he/she may recommend testing. The emphasis on the word "may" arises from the fact that the decision to perform or not perform testing is based on a number of factors discussed below. It is certainly true that some processes may require no (or minimal) testing. This will be discussed in more detail later in this paper.

Table 1. **Scales of operations**

small scale:	<5 L (typically found in an R&D laboratory or very small pilot plant)
moderate scale:	5−200 L (typical pilot-plant scale)
large/production scale:	>200 L

At many companies, a risk/consequence-based approach is successfully applied to the testing strategy decision process. This approach balances the scale of operations (e.g., a one liter in a R&D laboratory versus a 10,000 L reactor in a plant) with the overall energy release potential.

As an example, consider the mixing/blending of brine (salt water) with a nonvolatile, high flash point polyglycol at ambient conditions at a large scale, say 5000 L.

This process may not require any testing due to the known properties and known nonreactivity of these materials.

Despite the fact that this is at a large scale and may have never been practiced in the plant before, the risk of not performing any testing is low enough to be deemed acceptable. On the other hand, a 5-L glassware R&D synthesis on a new substituted nitrobenzene would probably require testing due to known energy release potential of the nitro moiety and the unacceptable risk (of injury/fatality) if the new process leads to a catastrophic energy release event.

The typical testing strategy applied is to screen first, then apply more sophisticated testing if necessary. By the term "screen" we mean to apply less expensive, quicker turnaround, smaller scale, and less experimentally complicated tests. In the thermal stability area, DSC is the commonly used screening technique. By "more sophisticated" testing, we infer the application of more expensive, slower turnaround, larger scale, and more experimentally complicated tests. In the thermal

---

(1) *Guidelines for Chemical Reactivity Evaluation and Application to Process Design*; American Institute of Chemical Engineers: New York, 1995; 0-8169-0479-0.

(2) Crowl, D. A.; Louvar, J. F. *Chemical Process Safety: Fundamentals with Applications*; Prentice Hall: New York, 1990.

(3) Yoshida, T. *Safety of Reactive Systems*; Elsevier: Dordrecht, The Netherlands, 1987.

(4) Grewer, T. *Thermal Hazards of Chemical Reactions*; Elsevier: Dordrecht, The Netherlands, 1994.

(5) Frurip, D. J.; Britton, L.; Fenlon, W.; Going, J.; Harrison, B. K.; Niemeier, J.; Ural, E. The Role of ASTM E27 Methods in Hazard Assessment, Proceedings of the 38th AIChE Loss Prevention Symposium, New Orleans, LA, 2004. This paper was divided into two parts and subsequently published in *Process Saf. Prog.* **2004**, *23* (4), 266 and **2005**, *24* (1), 12.

(6) Frurip, D. J.; Elwell, T. Effective use of Differential Scanning Calorimetry in Reactive Chemicals Hazard Evaluation. *Process Saf. Prog.* **2007**, *26* (1), 51.

(7) Rowe, S. Thermal Stability: A Review of Methods and Interpretation of Data. *Org. Process Res. Dev.* **2002**, *6*, 877–883.

(8) Dale, D. J. A Comparison of Reaction Screening Techniques. *Org. Process Res. Dev.* **2002**, *6*, 933–937.

10.1021/op800121x CCC: $40.75

Table 2. **Energy release potential**

low energy release potential:	no chemistry *per se* (e.g., blending or compounding) with low heats (less exothermic than −100 J/g), no high process temperature degradation issues
moderate energy release potential:	chemistry with heats less exothermic than −200 J/g, run dilute (<25 wt %) with moderate kinetics (>4 h for reaction to complete) at moderate temperatures (<100 °C). TMRs from worst case *T* for any reaction or degradation processes >8 h
significant energy release potential:	chemistry with heats more exothermic than −200 J/g, run more concentrated (>25 wt %) with faster kinetics (<4 h for reaction to complete) at elevated temperatures (>100 °C). TMRs from worst case *T* for any reaction or degradation processes <8 h

stability area, this might be an ARC test. Generally the more complicated tests provide more quantitative and more accurate information.

This strategy allows us to be efficient and economically optimized in this important area of process design. Overtesting (i.e., performing unneeded tests) can add significantly to the cost of scale-up and commercialization and may be totally unnecessary. Thus, by testing as needed for the scale-up process, we can avoid unnecessary costs and time delays.

2. Designing an Appropriate Reactive Chemicals Evaluation Scheme

In order to illustrate in more detail the aforementioned process of designing an experimental and calculational reactive chemicals evaluation we first propose to arbitrarily segregate three scales of operation, three scales of energy release potential, and three types of testing and calculations for a typical chemical industry. Bear in mind that there are no hard and fast rules to this categorization;, in fact, as will be seen later, there are many "fuzzy" boundaries. However, based on our experience, we think that dividing up the work process in this fashion is a helpful way to approach the problem.

Scale (Size of Operations) Categories. As seen in Table 1, we define the small, moderate, and large/production scales of operations.

We believe these categories are typical of any large chemical company and represent the laboratory, pilot pant, and production scales. Certainly at a "world class" operation for a commodity chemical (like ethylene oxide, for example), the scales are even 10−100 times the largest cutoff point in the table. We also realize that in pharmaceutical production, the largest scale may be of more moderate size. Clearly, by the time a reaction vessel gets to, say, the 200-L scale, the heat losses are diminished significantly so that a runaway reaction is more likely when cooling is lost.

Energy Release Categories. It is important to recognize that energy release potential really combines two dimensions: thermodynamics (how much energy can be released) and kinetics (how fast can the energy be released). A rusting iron pipe is very "hot" oxidation chemistry (∼−400 kJ/mol or −3600J/g)) but proceeds slowly enough not to present a hazard. So the energy release potential must always contain elements of both thermodynamics and kinetics. The kinetics part of the risk also involves how "close" the process operates to a condition which would be hazardous. This concept has been explored in great detail by Stoessel.[9] We incorporate this concept into the discussion through the adiabatic time to maximum rate (TMR). TMRs may be estimated from experimental data (such as ARC) or DSC.[10] Shorter TMRs (typically calculated from the highest process upset temperature) are more hazardous than longer ones. Our experience guides us to three logical categories of energy release potential (Table 2).

The numerical energy release cutoff points are expressed in J/g and are intended to represent the energy release from a reaction mass (including solvent) or from exothermic degradation of any process stream, raw material, or product. Clearly, if a reaction is especially energetic on a *molar* basis (say, more exothermic than −200 kJ/mol), then the energy release classification should account for any credible misoperations that might allow for a more *concentrated* reaction mass and therefore a larger energy release potential on a per gram basis.

We note that not all the individual criteria may be met for each category. It is up to the SME to choose the appropriate category (i.e., one that will lead to adequate calculations and testing to ensure an accurate result). Also, the 8 h for a trigger point in TMR was chosen since it represents a single shift (time spent at work by a typical worker).

With these categorizations, we can turn our attention to how we may approach the problem of designing a testing scheme based on these definitions. Our aim is to devise a testing strategy that manages risk in a way that is accurate and efficient and results in acceptable risk.

In other words, we cannot eliminate risk entirely but our evaluation of the process including exactly what we test (which raw materials and process streams) and how we test it (what instruments we employ) must result in a clear understanding of the safe operating limits of the process that are sufficiently accurate enough for us to accept and tolerate the risk. In practical terms, this means that we require a much deeper understanding of a chemical process when that process has an inherently greater propensity for an undesired energy release event than a more benign process with lower potential energy release,

---

(9) Stoessel, F.; Ubrich, O. Safety Assessment and Optimization of Semibatch Reactions by Calorimetry. *J. Therm. Anal. Calorim.* **2001**, (1), 64.

(10) Keller, A.; Stark, D.; Fierz, H.; Heinzle, E.; Hungerbühler, K. Estimation of the Time to Maximum Rate using Dynamic DSC Experiments. *J. Loss Prev. Process Ind.* **1997**, *10*, 31.

Table 3. **Detailed chart of testing strategies based on energy release potential and process scale**

	Low energy release potential No chemistry per se (e.g. blending or compounding) with low heats (less exothermic than -100 J/g), no high process temperature degradation issues	**Moderate energy release potential** Chemistry or mixing with heats less exothermic than -200 J/g, run dilute (<25 wt. %) with moderate kinetics (> 4 hours for reaction to complete) at moderate temperatures (<100 °C). TMR's from worst case T for any degradation processes > 8 hrs.	**Significant energy release potential** Chemistry with heats more exothermic than -200 J/g, run more concentrated (> 25 wt. %) with faster kinetics (< 4 hours for reaction to complete) at elevated temperatures (>100 °C). TMR's from worst case T for any degradation processes < 8 hrs.
Large/Production scale > 200 liter	"Desktop" methods and some screening tests of selected process streams (including credible misoperations), raw materials, more quantitative testing such as ARC if deemed warranted, testing of final product	"Desktop" methods and screening tests of many process streams/raw materials, wastes, (including credible misoperations), more quantitative testing such as ARC, and Scalable tests such as Reaction Calorimetry if deemed warranted, VSP if reactive relief deemed credible, TMR, TNR calculations, testing of final product	"Desktop" methods , screening tests of most process streams/raw materials, wastes, (including credible misoperations like accidental mixing), ARC, and Scalable tests such as Reaction Calorimetry if deemed warranted, VSP if reactive relief deemed credible, possible kinetic modeling, TMR, TNR calculations, testing of final product
Moderate scale 5-200 liter (typical pilot plant scale)	"Desktop" methods and some screening tests of selected process streams/raw materials	"Desktop" methods and some screening tests of selected process streams (including credible misoperations), raw materials, more quantitative testing such as ARC if deemed warranted, possible TMR calculations	"Desktop" methods and screening tests of most process streams/raw materials, (including credible misoperations), more quantitative testing such as ARC, and Scalable tests such as Reaction Calorimetry if deemed warranted, TMR, TNR calculations
Small scale <5 liter (typically found in an R&D laboratory or very small pilot plant)	"Desktop" methods and screening tests only if materials are new and properties unknown	"Desktop" methods and some screening tests of selected process streams/raw materials	"Desktop" methods and some screening tests of selected process streams (including credible misoperations), raw materials, more quantitative testing such as ARC if deemed warranted

especially if the event results in a significant injury potential to the workers and negative effects on the environment and local community.

3. Testing and Calculational Methods

Another dimension to this discussion is choosing the type of thermal stability testing (typically calorimetry-based tests) and calculations/estimations to be performed. Along the lines of what we have proposed for the scale and energy categories above, one may divide the testing/calculational methodologies into three categories: screening, quantitative, and scalable. The boundaries here, like our classifications before, are somewhat arbitrary but arise from experience; furthermore, there are clear overlaps between the methods. As mentioned earlier, screening tests are typically easier and less expensive to run. Typically data acquisition is fast ($<$4 h), and the equipment utilizes small samples ($<$1 mL). These tests and the corresponding instrumentation are also not very expensive. However, while valuable in the screening sense, these tests may not yield scalable data that can be used for process engineering design. That being said, in many cases the absence of thermal activity in the test (up to the maximum test temperature) is all we need to know. The testing listed below has been described in great detail in the literature (see refs 1−6 for example). It is assumed that the reader has some familiarity with these methods. For example, the widely used CHETAH program[11] is useful not only for estimation of the reaction heat (needed for adiabatic temperature rise estimates) but also for prediction of the total/maximum energy release potential of a compound or mixture. Screening type literature reviews and calculations/estimations are, again, easy to do and do not take a lot of time to complete. Sometimes these are called "desktop" methods.

The screening tests that we perceive as being widely employed in industry are listed below:

Screening Tests.

- heat of mix calorimetry, (HOM, isothermal)
- differential scanning calorimetry (DSC - typically

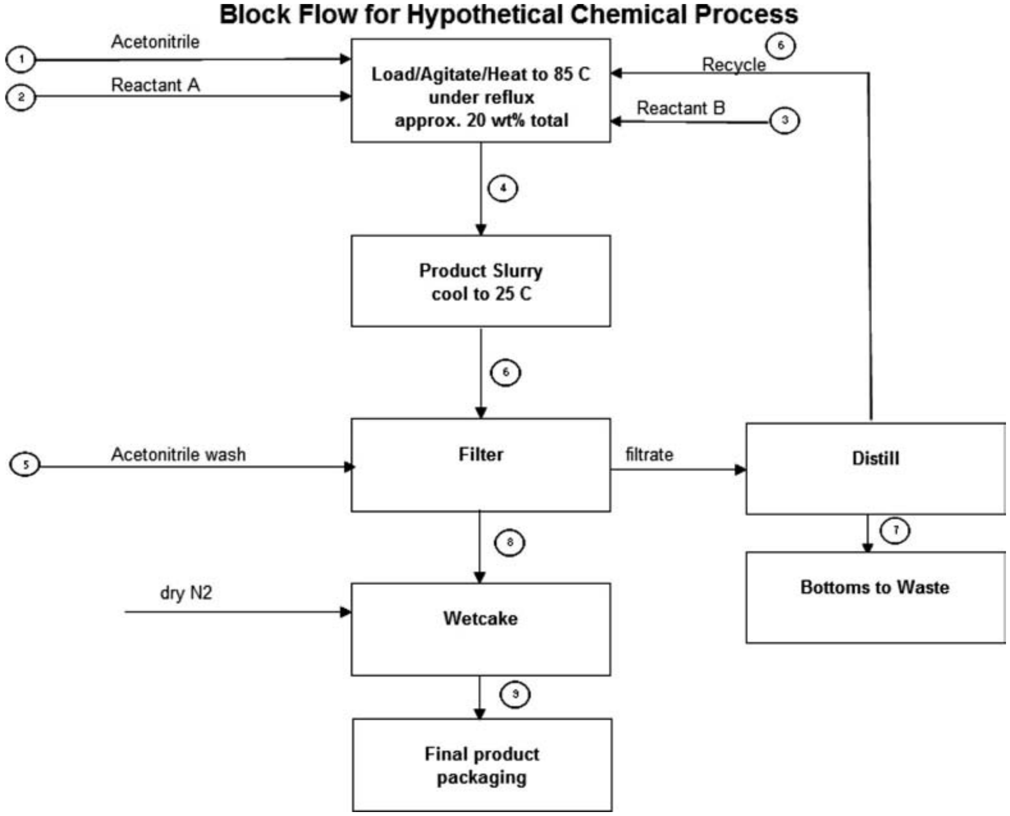


Figure 1. **Simplified block flow diagram for a hypothetical chemical process to illustrate the concepts discussed in this work.**

requires a robust sample holder to accommodate elevated pressure, see ref 6 and references therein)
- "desktop" methods such as heat of reaction estimation, CHETAH evaluation, adiabatic temperature rise estimates, review of the literature (Bretherick,[12] MSDS, safe handling guidelines, etc.)
- estimates of TMR from DSC[10]
- estimates of adiabatic temperature rise drop weight (not a calorimetry test but it belongs on the screening list, tests for potential impact sensitivity)

In the "quantitative" testing/calculation category list below, we include testing which is typically considered more "involved" from an experimental standpoint. Typically the testing requires more sample (4−50 mL) and takes longer for the experiment to be completed (1−3 days). The instrumentation is more sensitive (sometimes by one or more orders of magnitude) and has additional capability such as pressure measurement.

However, the added effort can yield data that are inherently more valuable to the process chemists and engineers by helping to make accurate engineering decisions about the hazards of a process and, more importantly, how to manage them.

It needs to be emphasized that, in this proposed work process, the results are not known a priori, for if they were, there would be no need to test at all.

Thus, in this category of testing and in the next, the results of any test may cause the strategy to be modified to run a more quantitative test. A good example is the DSC/ARC connection, where a DSC of a material may show significant and perhaps unexpected energy release detected at a low enough temperature to fully justify a subsequent ARC run.

Quantitative Tests.

- ARC (yields good thermokinetics with pressure data)
- isothermal calorimetry (can confirm ARC kinetic extrapolation to lower *T*)
- card gap testing (noncalorimetry, tests for detonability)
- estimates of TMR from ARC (more accurate than TMR from DSC)

Scalable tests are even more involved and more expensive but yield results which are directly and quantitatively applicable to engineering process design. For example the vent sizing package (VSP) test while similar to an ARC test, is run in such a way that thermokinetic and pressure data required for an accurate pressure relief design are obtained. While less calorimetrically sensitive that the ARC, its ability to run under so-called low ϕ conditions, closer to real plant conditions, makes VSP a good choice for vent design work. Reaction calorimetry is another example of this type of scalable test which closely mimics the process reaction conditions. It has the capability of measuring heat release in a dynamic manner and allows dosing of the reactants under controlled conditions. It may be thought of as a mini-pilot plant in that sense and is widely used for reaction engineering and process optimization. This method is particularly good at investigating scenarios where reactant accumulation (and therefore energy release potential) may become hazardous. In some companies, in particular those involved mainly in the manufacture of pharmaceuticals, reaction calorimetry has found greater use in the routine, day-to-day screening for reaction hazards. In the calculation area,

Table 4. **Probable reactive chemicals testing/calculational strategy for the hypothetical chemistry shown in section 5 based on the three arbitrary process scale categories presented in this paper**

	scale of operation (see definitions in Table 3 and discussion above)		
raw material or process stream	testing recommended at small scale	testing recommended at moderate scale	testing recommended at large/production scale
reactant A	DSC	DSC and ARC if DSC shows substantial energy release, especially if detected near process temperature	DSC and ARC if DSC shows substantial energy release, especially if detected near process temperature
reactant B	DSC	“	“
product	DSC	“	DSC, ARC
solvent (acetonitrile)	known properties/stable	known properties/stable	known properties/stable
reaction mixture	heat of reaction estimate only, adiabatic temperature rise estimate	heat of reaction estimate, DSC, possible heat of mix, possible ARC	heat of reaction estimate, DSC, heat of mix, ARC, VSP if reactive relief is deemed credible, reaction calorimetry if issues detected in other testing warrant
post reaction mixture	None if DSC of wetcake is benign	DSC and ARC if DSC shows substantial energy release, especially if detected near process temperature	DSC, ARC if DSC shows substantial energy release
distillation overheads	none (note: this step may not be applicable at this scale)	none if overheads analyzed as neat solvent	DSC
distillation bottoms	none (note: this step may not be practised at this scale)	DSC and ARC if DSC shows substantial energy release, especially if detected near process temperature	DSC and ARC if DSC shows substantial energy release, especially if detected near process temperature, VSP if reactive relief is deemed credible
product wetcake	DSC	“	“
dry product	none if DSC of wetcake is benign	“	ARC
	heat of reaction estimate[16]	heat of reaction estimate, Adiabatic temperature rise, TMR estimate from DSC of reaction mix	heat of reaction estimate, Adiabatic temperature rise, TMR from ARC

development of detailed thermokinetic models fits in this category. This model may be used to determine the so-called temperature of no return, or TNR, where heat input from the process exactly balances the heat removal capabilities of the system.

Examples of the application of TNR may be found in.[13,14] Beyond the TNR, the system will run away.

Scalable Testing/Modeling.

- VSP
- reaction calorimetry
- development of detailed kinetic models for the process chemistry (such as those needed for use in ASPEN[15] process simulation program)
- heat gain, heat loss evaluation for determination of TNR

4. Evaluation Scheme

Combining all the thoughts above, we present in Table 3 how one may approach a reactive chemicals evaluation of a process based on the three dimensions of scale, energy release, and testing types. The lower left-hand part of Table 3 is the relatively benign area, requiring little or even no testing. Conversely, at the upper right-hand part of the table (large scale and high energy release), much more testing and calculations would be needed to reduce the risk to an acceptable level. Of course, at intermediate scales and energy, intermediate approaches may be used.

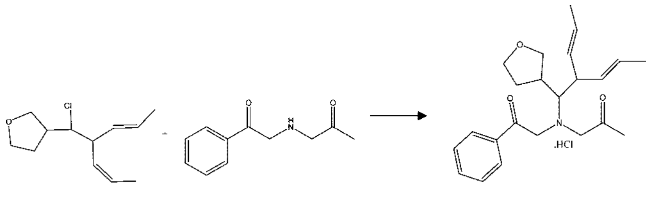
Note the expected symmetry in Table 3. For example a process with low energy release potential and run at a moderate scale would most likely be approached the same from a testing strategy perspective as a process with a moderate energy release potential and run at a small scale. Also note that as the testing scheme becomes more involved, the number of process streams tested does too.

In the less hazardous region of the chart in Table 3, only selected raw materials and process streams may be tested, whereas in the higher risk areas, many more streams may be tested, and those tests may be in the quantitative or scalable categories. Finally, for responsible “product stewardship”, we believe that it is important to fully understand the reactive chemicals hazards of the final product since this is the material that will be used by external customers. Thus, we may perform a test like an ARC on a final product, even though we are confident there are no hazards.

To fully vet this scheme, and to help explain this further, it is best to proceed with a worked hypothetical example.

5. Hypothetical Example

Consider the (hypothetical) process chemistry illustrated below.



Reactant A Reactant B Product (salt)

According to the block flow diagram in Figure 1, the recipe calls for the two starting materials to be loaded into the inerted reaction vessel after a solvent (acetonitrile) has been added. The total concentration of reactants is on the order of 20 wt %. After agitation to dissolve the starting materials, the contents are slowly heated to 85 °C and held for 6 h under reflux. The product, a salt, remains in solution until after the reaction, the reactor is cooled, and the material precipitates out of solution. The product is filtered off and washed with fresh acetonitrile, and the mother liquor is recycled back to the reactor after distillation. The distillation bottoms are sent to a common waste pack used in the facility. The filtered and washed product is dried with warm (80 °C) nitrogen and packaged off. The R&D chemist who performs this chemistry at 100 mL scale sees minimal heat release, but a CHETAH[11] calculation shows the reaction enthalpy is on the order of -130 kJ/mol (~ -120 J/g for the process conditions), so the experimental observation is probably explained by the slow kinetics and heat losses at the small scale.

On the basis of the criteria we have delineated above, this process probably falls into the moderate energy release category. In Table 4, we list the recommended testing and calculations for all of the three scales of operation based on guidance from Table 3. (Our focus here is thermal stability testing; we will assume the flammability issues have also been managed through inerting and other means).

Table 4 is self-explanatory and clearly shows the concepts discussed above. Certainly the testing strategies are always dynamic in the sense that our testing may show an energy release which may be hazardous in the process. That may in turn lead to further testing with different mixtures or testing using some of the more quantitative or scalable calorimetric methods.

A common testing strategy used here is to run the first test on the desired reaction composition but, depending on what is observed in that test, run another test on a mixture that is "worst case", mimicking credible misoperation of say, 2× loading of one of the raw materials or a catalyst.

6. Conclusions

We have presented here a systemized way to approach the task of designing a calculational and experimental plan for reactive chemicals evaluation of a chemical process. This approach is based on the two dimensions of process scale/size and energy release potential. The latter category, in itself, inherently contains thermodynamics (how much energy can be released) and kinetics (how fast can it be released). Using this approach, and without paying too close attention to the somewhat arbitrarily chosen boundaries of the categories, can help the reactive chemicals SME make efficient and accurate decisions about testing and calculations. The process is flexible enough to ensure that if any unacceptably large energy release events are identified in smaller scale, so-called screening tests, the testing strategy dictates that further testing with more quantitative methods be applied. Ultimately, the result of the testing must be a solid definition of the safe operating boundaries (envelope) of the process.

Acknowledgment

The author would like to express sincere appreciation to Tim Elwell for helpful discussions in framing the ideas in this paper. The author also acknowledges the historical contributions of many members in our Reactive Chemicals Testing group who have collectively developed the methodologies presented here including Larry Whiting, Bruce Powers, Ashok Chakrabarti, Harold Johnstone and Tom Hofelich. Finally, the author thanks the organizing Committee of the "22nd Reaction Calorimetry Conference for Safety and Scale-up of Chemical Processes", San Donato (Milan, Italy), December 4, 2007, where this paper was presented, for their support. In particular the author expresses his appreciation to Paolo Cardillo (SSC), Angelo Lunghi (SSC), Lucia Gigante (SSC), and finally Marco Dellavedova (SSC).

Received for review May 22, 2008.

OP800121X

---

(11) Harrison, B. K. *CHETAH*, v 8.0; available from www.ASTM.org; see http://www.chetah.usouthal.edu/ for a complete citation list for CHETAH; see also: Frurip, D. Computational Thermochemistry In Support of Calorimetry for Reactive Chemicals Assessment *Chim. Ind.* **2006**, *88*, 3, 140.

(12) Urben, P. *Bretherick's Handbook of Reactive Chemical Hazards*; Bretherick, L., Ed.; Butterworth-Heinemann: Woburn, MA, 1999.

(13) Ferguson, H. D.; Townsend, D. I.; Hofelich, T. C.; Russell, P. M. Reactive Chemicals Hazard Evaluation: Impact of Thermal Characteristics of Transportation/Storage Vessels. *J. Hazard. Mater.* **1994**, *37*, 285.

(14) Hofelich, T. C.; Frurip, D. J.; Powers, J. B. The Determination of Compatibility via Thermal Analysis and Mathematical Modeling. *Process Saf. Prog.* **1994**, *13* (4), 227.

(15) http://www.aspentech.com.

(16) Weisenburger, G. A.; Barnhart, R. W.; Clark, J. D.; Dale, D. J.; Hawksworth, M.; Higginson, P. D.; Kang, Y.; Knoechel, D. J.; Moon, B. S.; Shaw, S. M.; Taber, G. P.; Tickner, D. L. *Org. Process Res. Dev.* **2007**, *11*, 1112–1125 and references therein.