

Jasbir Singh



Scale up of Hazardous Chemistry with thermal runaway potential

KEYWORDS: runaway reaction, reaction calorimetry, exotherm, adiabatic calorimeter, reaction hazard

Abstract Chemical reactions involving potentially unstable chemicals or indeed even quite standard reactions which are exothermic, are an accepted part of the chemical industry. When there is no alternative to the selected route, it is vital that the process is subjected to modern methods of experimental and analytical investigation so that the potential hazard is properly quantified and a protection strategy based on scalable results developed, before even pilot scale operation is anticipated.

This paper will describe, with examples, why it is virtually impossible for even experienced chemists to "spot" hazardous reactions without some measurement and how reactions that might have been safely performed for many years can lead to disaster unless the risk is understood. This is now quite simple to do with a range of experimental methods available. With the correct equipment, it is possible even to experimentally simulate a thermal runaway accident in order to fully understand the potential hazard, using less than a 100ml of sample.

INTRODUCTION

Successful scale up of exothermic chemical reactions relies critically on the data and understanding developed during process development. Too often however, chemists develop processes while focusing very narrowly on the synthesis steps they ideally need to undertake with little or no attention to deviations in operation, especially if something goes wrong. Also, severity of the exothermic hazard is often not measured, relying instead on the chemist's visual observations to decide when there might be a potential problem and hence when heat generation data might be warranted.

The problem with focusing only on the normal (often called "desired") operation is that it ignores the real world situation where things do go wrong and where the process will go into "undesired" situations.

Expecting chemists to "notice" potentially exothermic problems is also fraught with problems as in practice this is near impossible to do as small scale reaction flasks or vessels will give little clue as to the heat being generated. This is basically a scale effect where the natural heat loss from small vessels is so large that they have to be heated, even though the chemistry is exothermic, making it impossible to "notice" exotherms. This can be appreciated more easily from the information presented in table 1 (1), where the natural heat loss (or cooling) from water at 60°C, is presented for different size vessels. A 100ml beaker for example, cools down at 3°C/minute while a 5m³ reactor only cools at 0.02°C/minute. With background cooling of 3°C/minute, it will be impossible to notice the presence of reaction heat.

This shortcoming is compounded by another characteristic

Item	Volume	Cooling rate Deg C/min	Heat Loss Rate *	
			Watts	Watts/Kg
Test tube	10ml	5.5	3.08	385
Beaker	100ml	3.0	16.8	210
Flask	1000ml	0.5	28.0	35.0
Reactor	2.5m ³	0.047	6580	3.29
Reactor	5.0m ³	0.023	6440	1.61
Reactor	12.7m ³	0.0169	12019	1.18
Reactor	25m ³	0.0043	26020	0.30

*For an 80% fill level with water

Table 1. Typical cooling rates and heat loss data. Laboratory and industrial scale equipment. 60 deg. C hotter than surroundings

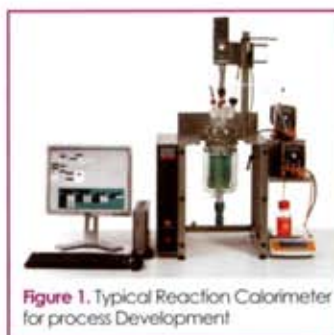


Figure 1. Typical Reaction Calorimeter for process Development

of working at small scale which is that the apparatus will proportionally absorb such a large fraction of the reaction heat being generated that the effect of any unplanned incident may not be noticed at all (or comes across as quite mild). The same event on larger scale could have devastating consequences. Therefore the secret to

successful (and fast) scale up is to put more effort into the development step, firstly to measure heat generation of the normal synthesis and secondly to consider deviations into undesired situations. The tools to do this are readily available.

UNDERSTANDING THE "DESIRED" SYNTHESIS

Basic data from a reaction calorimeter

Safety and operability of stirred reactors can be studied in a "reaction calorimeter" (2) which is a combination of a fully automated (computer controlled) reactor (see figure 1, SIMULAR as an example (3)) plus additional features that enable rate the of heat release to be tracked. This is much more of a tool for understanding the operability of stirred reactions, including their safety, and less an analytical device that provides a "number" or "answer", though it can do that too.

It is important to recognise the value of running a process in a totally controlled (and hopefully reproducible) way and having a complete record of the operating conditions, it is key to repeating the synthesis at larger scale. A common problem when scaling up a laboratory synthesis is the lack of a reliable account of the small scale operation and consequently wasting valuable resources working it all out on the pilot plant. Consider as an example, an esterification reaction involving the semi-batch addition of an anhydride to the alcohol, using equipment very much like that in figure 1, a 1litre glass reactor with temperature control and agitation control plus automated dosing of the anhydride at 2ml/minute (a total of 100g). The results from a run at 30°C are plotted in figure 2 where in addition to the dose data, the rate of heat generation (ie reaction power, measured in Watt) is plotted. The heat rate peaks at around 15W and though the feed was stopped at ~110minutes into the experiment, the heat output continues long afterwards. The heat output eventually stops at ~350minutes into the experiment, which is an on-line indication of when the reaction is finished. Also plotted on figure 2 is the cumulative amount of energy released (expressed in kilo-Joule); of the total (~68 kJ), around 33.4 kJ were released when the dosing was stopped, representing ~50 percent of the total energy.

Kinetics and reactant accumulation

Assuming the amount of heat released is proportional to the extent of reaction, this means that ~50 percent of the anhydride dosed (ie 50 g) had accumulated in the reactor, taking a further ~240 minutes to react out. The total heat release can be expressed as a reaction enthalpy ~30 kJ/mole but much more useful is the understanding of kinetics and the insight into potential hazard which is provided.

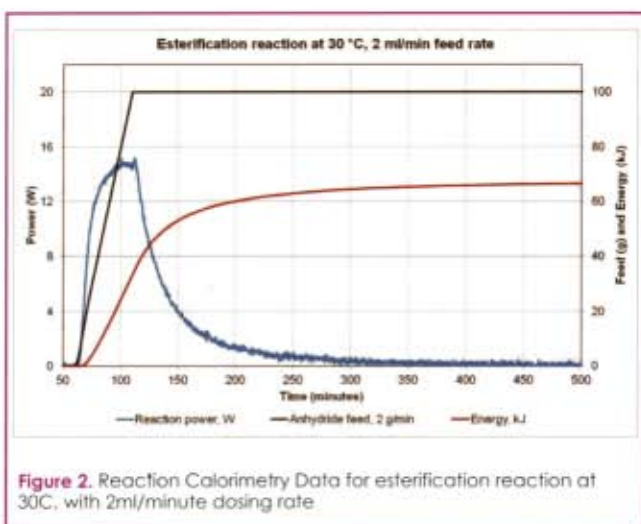


Figure 2. Reaction Calorimetry Data for esterification reaction at 30°C, with 2ml/minute dosing rate

The large accumulation implies that the kinetics is relatively slow (compared with the dose rate) and therefore not all the anhydride added is being converted. This is normally termed a "kinetic controlled" reaction since kinetics are controlling the synthesis time. The anhydride proportion that sits un-reacted will thermally runaway if at any time cooling is lost or agitation fails. The true extent of the problem during such a runaway will depend upon when the problem occurs, the worst case being just before or after the feed is stopped.

The data shows that at this "worst" point, 34.6 kJ of energy would be released (assuming no other reaction occurred to make it worse) and the temperature rise that would be caused can be readily estimated:

Temperature rise = $Q/mC_p = 106\text{ }^\circ\text{C}$ (using $Q=34,600\text{ J}$, $C_p=2\text{ J/g}^\circ\text{C}$ and $m=(100+63)\text{g}$ being the mass of anhydride and alcohol respectively).

Starting at 30°C, this means that the mixture would reach a maximum temperature of ~136°C; the time taken to reach this point (ie speed of the thermal runaway) cannot be derived from the available data nor the precise pressure that would be generated. At the maximum temperature, the alcohol would reach a pressure of ~8bar (being simply the vapour pressure of the alcohol at ~136°C), as it would be well above its boiling point.

The true worst case scenario would result if all of the reaction heat were released without any cooling; this would lead to twice the temperature rise compared to the above calculation, reaching a predicted temperature of ~242°C; for reference, the alcohol would have a pressure of 75bar (based on the vapour pressure of the alcohol at this higher temperature).

This illustrates the huge importance of running in semi-batch mode, the potential hazard being halved. The residual hazard, represented by ~50 percent accumulation of the unreacted feed in this example, still represents a significant risk and further effort should be made to reduce this.

These very preliminary conclusions based on reaction calorimetry (isothermal) data are often enough to give further direction to process development. For example it is clearly desirable to reduce the accumulation.

Moving towards a "dose controlled" process

The options open for reducing the accumulation are following:

- Increase the reaction temperature
- Add a catalyst
- Reduce the feed rate

The first two would reduce accumulation by increasing the reaction rate while the third would achieve the same objective by giving more time for conversion. A comparison of the heat generation with a feed rate of 1 and 2g/minute is shown in figure 3. Increasing the temperature, while keeping the feed rate the same is compared in figure 4.

Slowing down the feed rate reduces the accumulation to ~32 percent and increasing the temperature to 40°C at this lower feed takes it down to ~22 percent. This is still not ideal but a lot safer than the original scenario and certainly far safer than a batch synthesis.

The ideal situation would be where the dosed reagent is converted as fast as it is added, giving a zero accumulation and this would constitute a "dose controlled" reaction. In this case, the heat output would fall sharply to zero when the dosing is stopped, rather like an acid/base neutralisation. In such processes, switching off the feed (in most cases) would prevent a thermal runaway.

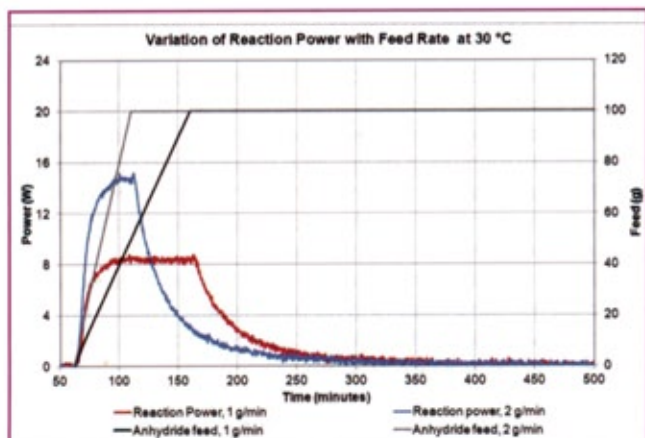


Figure 3. Esterification at 1 and 2g/minute, at 30C.

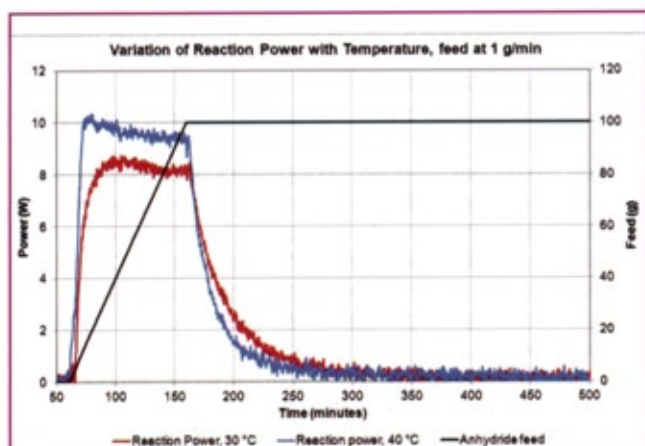


Figure 4. Esterification at 30 and 40C, with 1ml/minute dosing rate.

PROCESS SCALE UP

Matching the plant

The three situations explored on the reaction calorimeter, have different consequences on scale-up and not all will be suit a given pilot plant, which for example will often have a limited cooling capability. The key parameter to consider is the maximum heat release rate in each run; at 30°C and 2g/minute this is ~15W and ~8W at half this feed rate while at 40°C and 1g/minute heat release rate is ~10W. As a first check, these rates can be directly scaled up according to volume; for example if the pilot reactor has a capacity of 2000 litre, then to run the reaction at 40°C and 2kg/minute (instead of 1g/minute) dose rate, will require a cooling capacity of at least 20kW (instead of 10W). If this amount of cooling is not possible, then one of the other conditions might have to be considered, even though they might be worse from an accumulation viewpoint. It is possible that further laboratory trials on the reaction calorimeter will be needed to both match the plant cooling capacity and minimise accumulation. In some situations the synthesis route will have to be altered altogether, if an acceptable compromise cannot be reached.

Predicting the thermal runaway risk

In many important commercial cases, the process may have to be scaled up in spite of significant accumulation.

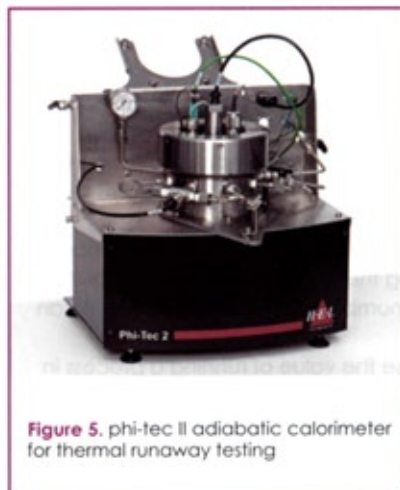


Figure 5. phi-tec II adiabatic calorimeter for thermal runaway testing

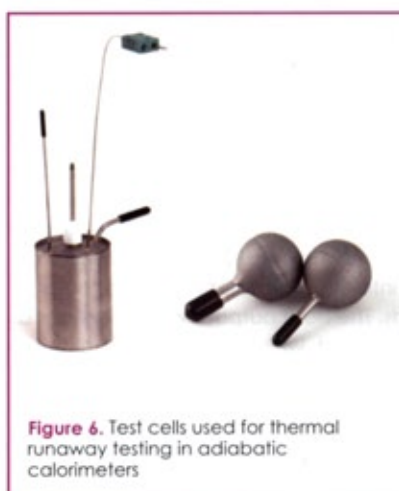


Figure 6. Test cells used for thermal runaway testing in adiabatic calorimeters

This is not a disaster; it simply requires further more detailed analysis of the risk in order that a suitable protection system can be implemented. The most un-ambiguous way to quantify the thermal runaway risk posed by the accumulation is to literally simulate the scenario on an adiabatic calorimeter (4). The adiabatic instrument, phi-tec II, used for this study is shown in figure 5 (5). This is an enhanced version of the more widely used adiabatic calorimeter, often called an "ARC" (4), and a typical test cell used in this experiment (cylindrical with stirring capability) and a classic ARC-cell which is spherical and much more robust looking, are shown in figure 6. Simulation of the thermal runaway incident for the esterification reaction involves charging the anhydride into the cylindrical test cell (~110ml in volume), warming it to the starting temperature (30 or 40°C in the current situation) and then injecting the appropriate amount of alcohol. To replicate the worst case, equimolar amounts of alcohol and anhydride would be reacted; changing this proportion would simulate the less severe cases. The results of an experiment with worst case mixture are

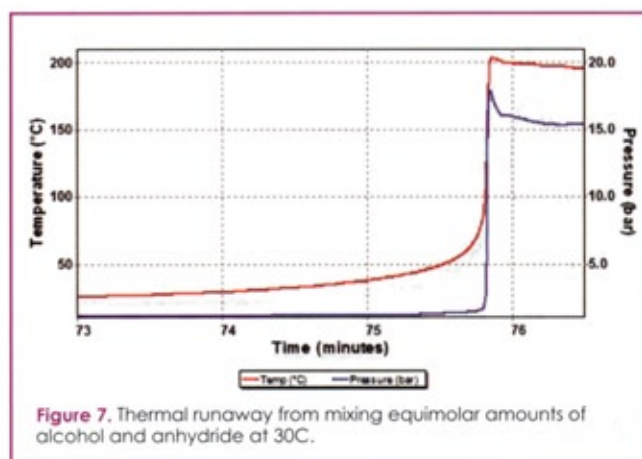


Figure 7. Thermal runaway from mixing equimolar amounts of alcohol and anhydride at 30C.

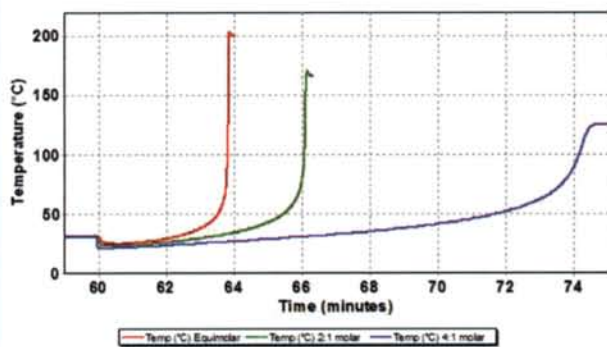


Figure 8. Temperature data resulting from different alcohol/anhydride ratios at 30°C.

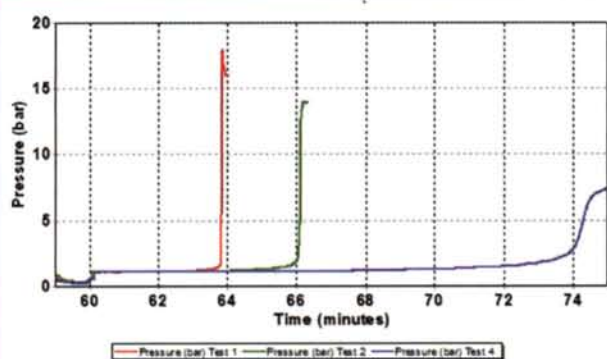


Figure 9. Pressure data resulting from different alcohol/anhydride ratios at 30°C.

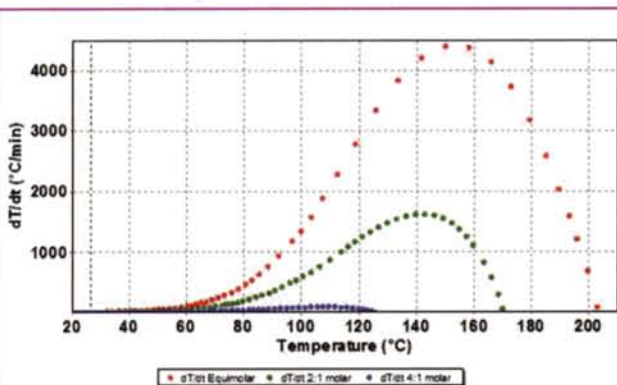


Figure 10. Rate of temperature rise data resulting from different alcohol/anhydride ratios at 30°C.

shown in figure 7, where pressure and temperature, as a function of time are plotted. The consequences of reducing the relative amount anhydride, corresponding to reduced amounts of accumulation, are compared in figures 8 and 9, where the temperature and pressure respectively, is plotted against time. Another useful way to evaluate the data is a plot of rate of temperature (dT/dt) against temperature – this is shown in figure 10, again for the 3 levels of anhydride accumulation.

Use of this type of adiabatic calorimeter provides a totally different insight into the process, complementing the isothermal data from the reaction calorimeter. For example, a runaway resulting from equimolar amounts of

anhydride and alcohol, barely takes 3 minutes (after the chemicals come together) to runaway to completion, reaching a temperature and pressure of $\sim 200^\circ\text{C}$ and $\sim 16\text{bar}$ respectively (figure 7); the rate of temperature rise peaks at over $4000^\circ\text{C}/\text{minute}$. When the relative amounts of alcohol and anhydride are reduced to 4 to 1, the incident takes around 12 minutes to reach the worst conditions and the maximum temperature and pressure are now only $\sim 125^\circ\text{C}$ and $\sim 8\text{bar}$ respectively. (Note that the time-scales in figures 7 to 9 are the experimental time, allowing for warming of the chemical, setting of initial steady conditions etc and the interesting part of the data starts when the two reagents are mixed and the exotherm begins. In figures 8 and 9, the times were shifted so that reagent mixing is at 60mins for all cases, in order to allow an easy comparison of the timescales for the 3 incidents).

This information would enable the correct combination reactor type and protection features to be selected. For example, the size of relief vent could be directly estimated from the adiabatic data for each scenario (6).

THERMAL HAZARD EVALUATION

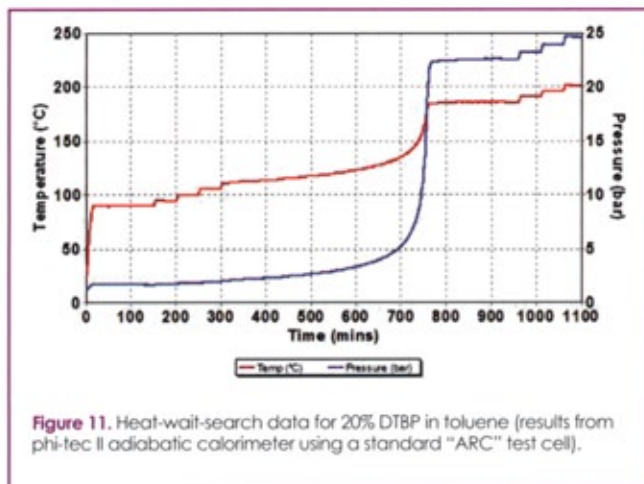
Adiabatic calorimetry

An adiabatic calorimeter ensures none of the heat generated by a chemical reaction (during thermal runaway) is lost to the surroundings, a situation that leads to the maximum temperature rise and hence a worst case scenario for the event. Unlike smaller reactors in which research is commonly done, large vessels are considered to be close to being adiabatic and little heat is actually lost. This is clear from the comparative data in table 1 discussed earlier. The reaction data obtained from an adiabatic calorimeter is therefore most appropriate in providing an understanding of the likely consequences.

The instrument used to produce the data in figures 7 to 10 has one other important feature, namely that the test cell used (into which the reagents were injected) was very light in weight, being made from thin-walled metal. Relative to the chemicals added to initiate the runaway reaction, the thermal mass of the metal was less than 10 percent and so very little of the heat produced was consumed in warming up the test cell. This is another important requirement because it is a close simulation of large scale plant and hence, the results in figures 7 to 10 can be considered to directly represent the pressure and temperature, as well as the time scale, that can be expected during a plant scale incident. No manipulation of the data is needed – in fact, it requires no knowledge of how the reaction proceeded or what the products of reaction are, the instrument mimics the possible incident.

Adiabatic calorimetry for Thermal screening

Important as the adiabatic calorimeter type (phi-tec II) used above is, it is not ideally suited to the detailed screening of liquids and powders which is very common requirement. Also, there is actually a much larger range of chemical hazard problems than accurately simulating a large-scale thermal runaway and this requires a different type of device, based around the classic "ARC" instrument. For reaction screening, a thicker walled and

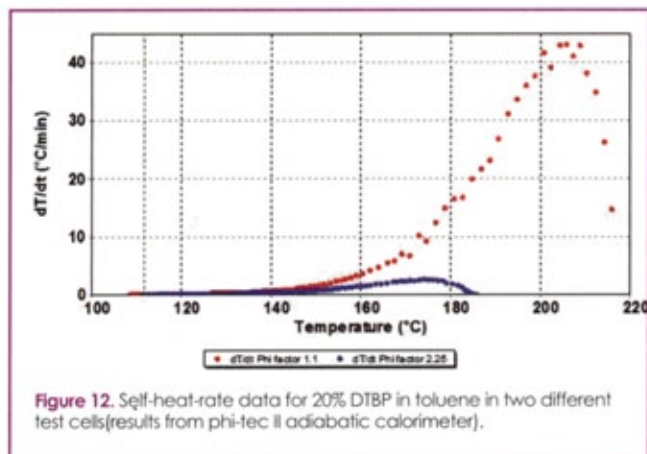


smaller (~10ml) test cell can be used - this is also shown in figure 6. This small and heavier test cells has a thermal mass that is similar to that of the reagents and hence, the temperature rise (and rate of rise) is much lower than in a real, large scale plant incident. However, much useful information can still be produced and this set up is widely used for an initial understanding of thermal hazards.

The most common way of using this type of test cell is to charge the sample (liquid, powder or mixture) and then heat it stepwise (so-called heat-wait-search method) until an exotherm is detected. An example is shown in figure 11 where 20wt percent di-tertiary butyl peroxide (DTBP) in toluene is screened for thermal instability and the resulting decomposition is tracked adiabatically. The exotherm is detected around ~115°C after several "search" steps; at the end of the exotherm (~180°C), the search was re-commenced to see if further reactivity is detected at higher temperatures, in this case nothing was found.

Though the reaction data in figure 11 is adiabatic, the rate of reaction is slowed down by the heat absorbed by the test cell and this aspect is normally expressed in terms of the phi-factor:

Phi-factor = $1 + (mC_p)_{\text{cell}} / (mC_p)_{\text{chemicals}}$
 where m is mass (g) and C_p is heat capacity (J/g °C).
 For the data in figure 11, the phi-factor was ~2.25. To replicate the behaviour of large scale equipment, this needs to be ideally ~1.05 – 1.1.



If the same test with 20 percent DTBP is performed in the larger test cell with smaller relative thermal capacity (ie phi-factor closer to 1) the reaction would be detected at a slightly lower temperature but more importantly the timescales for the runaway reaction would be two orders of magnitude shorter. This is shown very clearly in figure 12, where the rate of temperature rise is plotted against temperature. The maximum rate of temperature rise for the lower phi-factor data is ~45°C, compared with only ~3°C/minute for the high phi-factor data.

The reason for the large difference in rates of rise is the exponential dependence of the reaction rate on temperature. The lower relative thermal capacity data reports the timescales on which the event might develop in a real-life incident and hence its importance in predicting scales up. (In terms of visualisation of the effect of phi-factor on time-scale, the effect is the same as that discussed for the esterification reaction discussed earlier, where different amounts of dilution lead to a slower exotherm and then longer timescales over which the maximum temperature and pressure are reached, see figures 8 and 10 for example.) In spite of this limitation, the high phi-factor data has a wide range of uses and justifies the simpler construction of the device. It is also worth pointing out that the low phi-factor adiabatic device (phi-tec II) is able to do both types of tests simply by changing the choice of test cell.

CONCLUSIONS

The shift from batch to semi-batch processes and the subsequent development including minimisation of thermal hazard potential can be achieved with little effort using the modern version of reaction calorimeters. The ability to distinguish between kinetic and dose controlled processes provides valuable insight that enables optimisation of batch times, matching the process to available plant and a reduction in unreacted reagent accumulation all at the same time. The improvement in safety becomes an integral part of the development task.

However well a process is designed, there remains the chance that something could go wrong and additional safety features to cope with this can be designed using another type of calorimeter, based on adiabatic operation. Safer storage, transportation and proper relief sizing can all be directly undertaken on the basis of adiabatic data.

REFERENCES AND NOTES

1. Harsbook, Safety Engineering in Practice, Volume 6, 6-2. Dechema, 2002.
2. Schulz K., Reaktionskalorimetrie: Entwicklung, Anwendung und Vergleich, Wissenschaft und Technik Verl., 2002. Universität Hamburg Diss., 2002.
3. Mitchel M, et al. Organic Process Research and Development, 6, 311-311, 2002.
4. Townsend D.I., Tou J.C., Thermochemica Acta, 37, (1-30), 1980.
5. Singh J., Thermochemica Acta 226, 211-220, 1993.
6. Fisher H. G. et al, The Design Inst for Emergency Relief (DIERS) Project Manual. American Inst Chemical Engineers, 1992.