

# Fit for Purpose Experimental Designs and Analyses in Chemical Development

Sergio Bacchi,<sup>\*,†</sup> Mohammad Yahyah,<sup>\*,‡</sup> Antonella Carangio,<sup>§</sup> Arianna Ribecai,<sup>†</sup> and Maria Concepción Cerrato Oliveros<sup>†</sup>

Chemical Development, Synthetic Chemistry, GlaxoSmithKline Medicine Research Centre, Via Fleming 4, 37135, Verona, Italy, Statistical Science, GlaxoSmithKline, Gunnels Wood Road, Stevenage, SG1 2NY, U.K., and Chemical Development, Synthetic Chemistry, GlaxoSmithKline, Gunnels Wood Road, Stevenage, SG1 2NY, U.K.

## Abstract:

Classical experimental designs have been used as configurations of choice at GlaxoSmithKline within Chemical Development in establishing robust products and processes while achieving other goals such as reduction of cost, waste, process investigation times and most importantly improved quality. Occasionally, as in this paper, these first intent designs have been proved to be unworkable through irregularly shaped experimental regions induced through chemistry constraints. To achieve optimum efficiency, a fit for purpose computer-aided D-Optimal Design in considering the reduction of cyclic imides to amines has been investigated here, proving to be an invaluable additional tool in applying the appropriate design at the relevant time. In addition, partial least squares (PLS) analyses has been explored with a view of simultaneously modelling the responses whilst primarily identifying those factors having the largest impact across the responses of interest (quantity of reducing agent and solvent, temperature, and time). This combined with process analytical technology (PAT) through online IR monitoring has facilitated part of this cyclic process of investigations using scientific and business judgment.

## 1. Introduction

At GlaxoSmithKline within Chemical Development, experimental design methodology has been an integral and established tool in obtaining robust products and processes as well as helping to improve quality and to reduce costs, waste, and timelines of process investigations. It has been used to great effect in answering questions such as (i) is my process lean? and (ii) is my process on the brink of chaos - i.e. how reliable is it to operate? In other words it has been used to identify critical factors, optimise factor settings, minimise waste, and identify robust operating regions for our processes. Classical screening and response surface designs such as factorial, fractional factorial, and central composite designs have on the whole been used to facilitate the described process. These standard designs have assured degrees of precision, orthogonality, and other optimal properties that are important in the context for which they are used.<sup>1</sup> Over the past few years, the availability of well designed, easy to use experimental design

software such as Design Expert (DX-7) have empowered the scientist to be more proactive in setting up and analysing these designs, thereby having less reliance on statisticians to do this. On some occasions these standard designs have proved to be inappropriate and impractical. One such limiting case seen more often within Chemical Development is the impossibility of running certain combinations of the factor settings or when the experimental region is constrained or irregularly shaped. To achieve optimum efficiency, fit for purpose computer-aided designs such as alphabetical optimal designs<sup>1</sup> have proved to be an invaluable additional tool in applying the appropriate design at the relevant time using scientific judgment and common sense. In this paper we consider one such example where a D-Optimal Design has been systematically built up to investigate the reduction of imides to amines. Partial least squares projection to latent structure (PLS),<sup>2</sup> has been used to model the relationship between the input factors and the correlated responses. This method generalizes and combines features from principal components analysis as well as multiple regressions and has been shown to compare favorably with other estimation techniques such as ordinary least squares, variable subset selection, ridge regression, and principal components regression.<sup>3,4</sup>

## 2. Chemical Example and Statistical Approach

Reduction of cyclic imides to amines is a well-described process that routinely employs a selection of three reagents, Lithium aluminum hydride (LiAlH<sub>4</sub>), sodium bis(2-methoxyethoxy)aluminium hydride (Red-Al), and borane (BH<sub>3</sub>) complexes.<sup>5–12</sup> All of them have some issues in terms of scalability of the process. Both LiAlH<sub>4</sub> and Red-Al are considered not ideal for their reactivity towards other functional groups of the molecule (Scheme 1) as well as intrinsic difficulties in handling these reagents

\* To whom the correspondence should be addressed. E-mail: sergio.k.bacchi@gsk.com; mohammad.2.yahyah@gsk.com.

<sup>†</sup> Chemical Development, Synthetic Chemistry, GlaxoSmithKline Medicine Research Centre.

<sup>‡</sup> Statistical Science, GlaxoSmithKline.

<sup>§</sup> Chemical Development, Synthetic Chemistry, GlaxoSmithKline.

(1) Myers, R. H.; Montgomery, D. C. *Response Surface Methodology Process and Product Optimisation Using Designed Experiments*, 2nd ed.; Wiley: New York, 1998.

(2) Eriksson, L.; Hermens, J. L. M.; Johansson, E.; Verhaar, H. J. M.; Wold, S. *Aquat. Sci.* **1995**, *57*, 217–241.

(3) Porter, M. A. *Statistician* **1993**, *42*, 217–227.

(4) Rockhold, F. W. *Statistic. Med.* **2000**, *19*, 3211–3217.

(5) For a general review on methods of generating diborane in situ from sodium borohydride see: Souza, M V. N.; Vasconcelos, T. R. A. *Appl. Organomet. Chem.* **2006**, *20*, 798.

(6) Polonski, T.; Milewska, M. *J. Tetrahedron Lett.* **1991**, *32*, 3255–3258.

(7) Rao, V. D.; Periasamy, M. *Synthesis* **2000**, *5*, 703–706.

(8) Herbert, C. B. U.S. Patent 3,634,277, 19720111, 1972.

(9) Volkov, V.; Myakishev, K. G.; Gorbacheva, I. *Inst. Neorg. Khim. Novosibirsk* **1983**, *6*, 1442.

(10) Toft, M. A.; Leach, J. B.; Himpsl, F. L.; Shore, S. G. *J. Inorg. Chem.* **1982**, *21*, 1952–1957.

(11) Myakishev, K. G.; Gorbacheva, I.; Volkov, V. *Inst. Neorg. Khim. Novosibirsk* **1984**, *29* (4), 912–916.

(12) Burkhardt, E.; Corella, J. A. U.S. Patent 6,048,985, 2000.

### Scheme 1. Chemical transformations

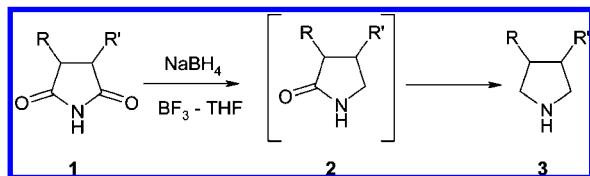


Table 1. Factors and selected factor ranges

factor	min	max	current target
NaBH <sub>4</sub> (moles per mole of starting material)	1.5	6	4
BF <sub>3</sub> ·THF (moles per mole of starting material)	2	8	5.7
total THF (volumes vs amount of starting material)	5	15	10.5
temp (°C)	20	35	25
time (h)	6	24	24

in bulk quantities. The use of borane complexes as an alternative provides a good reaction profile but has drawbacks in terms of safety issues.<sup>13</sup> In fact diborane gas, which can easily be evolved from the commercially available solution of borane in ether, has a very low auto-ignition temperature of ~38–52 °C and a wide explosive range in air (0.8–90% vol).

A process for the in situ preparation of borane was considered a valid alternative. A review of the literature<sup>5</sup> suggested that there were numerous examples describing the use of sodium borohydride (NaBH<sub>4</sub>) and a Lewis acid for the in situ generation of borane, BH<sub>3</sub>, but very little relating to the direct reduction of an imide. Even if, in such a case, the use of iodine as Lewis acid was recommended, due to the additional restriction in selected functional groups interconversion, we had to direct our attention to borontrifluoride (BF<sub>3</sub>).

**2.1. Objectives.** The main aims of this study were to:

- Identify critical factors (Table 1) impacting a number of response variables: residual starting material (res. SM) **1**, pyrrolidinone-like intermediate **2**, product **3** (Scheme 1), and impurities formation, whilst requiring tight control on scale-up. Two constraints were imposed on the region leading to an irregular design and hence to the impossibility of applying a simple factorial design:
  - the mole ratio between NaBH<sub>4</sub> and BF<sub>3</sub>·THF to be greater than 3/4; this constraint imposed through literature survey<sup>6–8</sup>
  - the total volumes of THF, relative to the unit amount of starting material, to be at least twice the moles of NaBH<sub>4</sub> - on grounds of experimental evidence in making the stirring effective.
- Investigate the possibility of reducing the amount of NaBH<sub>4</sub> used in order to facilitate the work-up of the reaction.
- Achieve high conversions with an acceptable impurity profile.
- Investigate the reduction in the time needed for the reaction to reach complete conversion (presently around 22 h).

(13) Urben, P. G. *Bretherick's Handbook of Reactive Chemical Hazards*, 6th ed.; Butterworth Heinemann: Woburn, MA, 1999; p 1937.



Figure 1. Anachem SK233 platform.

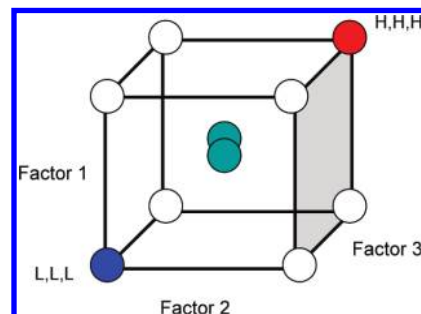


Figure 2. Scoping study runs.

Introducing experimental design at the same time as other techniques such as automated equipment at GlaxoSmithKline, has had a beneficial impact on the up-take of experimental design methodology. The use of both automated equipment and experimental design has allowed a coherent block of experiments to be planned in advance and then implemented. The SK233 autosampler platform is capable of carrying out 10 reactions simultaneously in the investigation (Figure 1). It allows automated reaction preparation, has good process control (range of temperature, concurrent reactions in different temperature zones: condensing, stirring, and nitrogen blanketing), and permits online HPLC monitoring.<sup>9</sup>

**2.2. Sequential Design Approach.** An initial set of four scoping experiments were run to test the reaction and analysis methods prior to committing time and materials to the eventual experimental design campaign as well as getting a feel for the appropriateness of the factor range settings (Table 1). Typically the four experiments that constitute a scoping study include two control reactions run close to the midpoints of each factor range, providing an estimate of the background variability in the system with two extreme sets of reaction conditions representing the mildest (L,L,L) and most forcing (H,H,H) reaction conditions (Figure 2). Note, the mildest conditions for a factor are not necessarily the lowest settings of a factor, e.g. dilution of a reaction mixture.

The ordering of the scoping study was considered important, with control experiments run at the start and end of the study to emphasise any time-related bias. The data from the study is given in Table 2 with associated output from the measured responses. To prevent a transgression of the constraints imposed on the design region, small reductions for BF<sub>3</sub>·THF were

**Table 2. Scoping design**

run	NaBH <sub>4</sub> (equiv)	BF <sub>3</sub> ·THF (equiv)	total vol THF (equiv)	temp. (°C)	time (h)	product <b>3</b> (%a/a)	intermed. <b>2</b> (%a/a)	impurities (%a/a)	res. SM <b>1</b> (%a/a)
1	4.00	3.67	10.5	27.5	15	84.9	4.6	10.5	0.0
2	1.50	2.00	5.0	20.0	6	39.6	56.4	2.0	2.0
3	6.00	8.00	15.0	35.0	24	78.4	11.6	10.0	0.0
4	4.00	3.67	10.5	27.5	15	75.9	13.8	9.8	0.4

necessary for the control runs. This change was not deemed significant for the purpose of the scoping study. Figure 3 shows clear differences in responses of the extreme conditions relative to the control experiments for residual starting material **1**, pyrrolidinone intermediate **2**, product **3**, and impurities. We believed that the extreme conditions could meet the desired specifications so that the current factor ranges were appropriate for the purpose of the investigation. However, there was some evidence of curvature in the responses, suggesting that this may need to be modelled in later studies. However, at this stage, focus was directed towards main effects and certain targeted two-factor interactions.

### 3. Results and Discussion

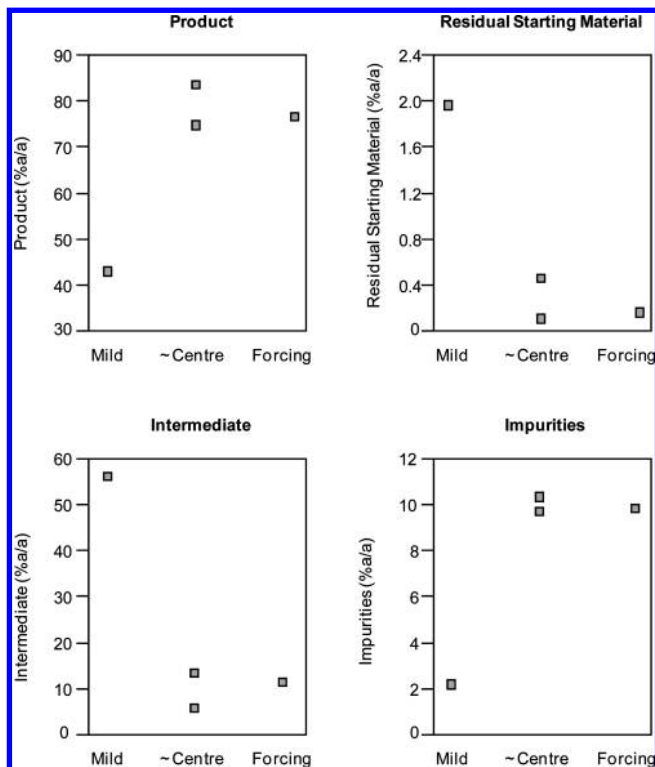
**3.1. D-Optimal Experimental Design.** Given resource limitations, the scoping experiments were then used to form part of the next experimental design campaign. This was generated using the D-Optimal Design option in DX-7 by specifying up front all the main effects and suspected two factor interactions that the subsequent design would be capable of estimating.

A 20-reactions D-Optimal Design was generated in three reaction sets (Table 2 for runs 1–4, plus Table 3 for runs 5–10 and runs 11–20) with the first set corresponding to the scoping reactions. A further consideration was made to incorporate

another control experiment in the design to investigate any time-related biases as well as assessing the reproducibility of the process (run 11). However, one of the control experiments (run 1, Table 2) was discarded from future investigations on the basis of its irregularity with the other two repeats. An evaluation of the efficiency of the resulting design was assessed against other competing D-Optimal Designs for the same resource using a number of measures including the variance inflation factors (VIF) and condition number.<sup>14</sup> In particular, the minimum VIF was 1.67 for the interaction between temperature and time and a maximum of 13.91 for the interaction between BF<sub>3</sub>·THF and the total volumes of THF (vol total THF).

**3.2. Analysis and Interpretation.** For the analysis of data, multivariate analysis was used to gain insight into the effects of the factors to the multiple responses considered. Principal component analysis (PCA) in many ways forms the basis for multivariate data analysis. PCA's main function is to reduce dimensions of the multivariate, multichannelled data to a few manageable dimensions, practically 2, 3, or 4, hence, referred to as a 2-, 3-, or 4-principle component (PCA) model. The first principal component explains the maximum amount of variation in the original data. The second principal component describes the maximum amount of remaining variation and is perpendicular to the first. Successive principal components describe decreasing amounts of variation and ultimately noise. The reduced data serves as an approximation to the original data and allows a chemist to use a scatter plot to overview the data in the reduced dimensions to study different observations and variables for their contribution and relationship to the overall variability of the data. Most of the information of the reduced data is contained in the loadings and scores plots of those components. Loadings are weights given to the original variables and are useful for locating the important variables/effects. Scores are linear combinations of the original variables whose weights are given by loadings. Scores are used in the place of the original data to look at the relationship between the observations and can be interpreted as projected data in the rotated coordinates. The coordinates of the rotation are determined to maximize the capability of capturing the information in the data.<sup>15,16</sup>

Partial least squares projection to latent structures (PLS)<sup>2,17</sup> is a proven multivariate calibration method in quantitative analysis. Similarly to PCA, its main idea is to make latent variables of the original matrix *X* (main effects and selected two-factor interactions in this case) and matrix *Y* (dependent variables). Latent variables are formed as a linear combination

**Figure 3. Summary of results from scoping study.**

(14) Shah, H. K.; Montgomery, D. C.; Carlyle, M. W. *Qual. Eng.* **2004**, *16*, 387–397.

(15) Wold, S. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52.

(16) Wold, S. *Technometrics* **1978**, *20*, 397.

(17) Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J. *J. Sci. Statistic. Comput.* **1984**, *5*, 735–743.



**Table 3.** Full experimental design, excluding the scoping design

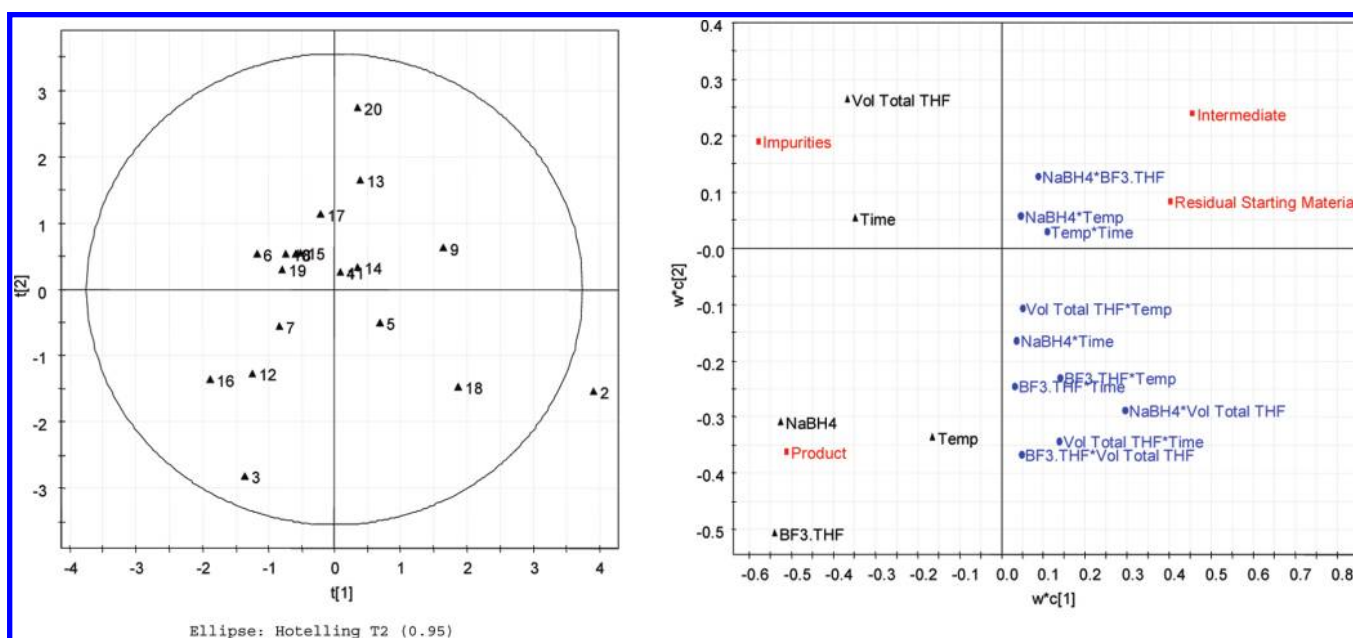
run	NaBH <sub>4</sub> (equiv)	BF <sub>3</sub> ·THF (equiv)	total vol THF (equiv)	temp. (°C)	time (h)	product 3 (%a/a)	intermed. 2 (%a/a)	impurities (%a/a)	res. SM 1 (%a/a)
5	4.25	2.00	8.5	35.0	6	64.5	24.7	9.7	1.1
6	6.00	8.00	12.0	20.0	6	69.2	18.3	11.8	0.7
7	3.75	5.00	15.0	35.0	24	83.0	4.9	11.4	0.7
8	6.00	2.00	15.0	20.0	24	65.5	23.5	10.0	1.0
9	1.50	2.00	5.0	35.0	24	48.0	41.6	9.6	0.8
10	4.25	5.67	8.5	20.0	24	77.1	7.2	15.0	0.7
11	4.00	3.67	10.5	27.5	15	75.4	12.4	11.9	0.3
12	6.00	8.00	15.0	35.0	6	81.7	5.9	11.8	0.6
13	1.50	2.00	15.0	35.0	6	48.5	41.0	9.6	0.9
14	6.00	2.00	12.0	20.0	6	59.1	33.0	6.8	1.0
15	6.00	2.00	15.0	35.0	6	63.5	26.6	9.4	0.5
16	6.00	8.00	15.0	20.0	24	76.7	8.2	14.6	0.5
17	3.75	5.00	15.0	20.0	6	59.7	28.9	11.1	0.4
18	2.50	3.33	5.0	35.0	6	70.5	24.1	5.1	0.4
19	6.00	2.00	12.0	35.0	24	64.5	17.8	17.3	0.3
20	1.50	2.00	15.0	20.0	24	45.6	42.5	11.3	0.6

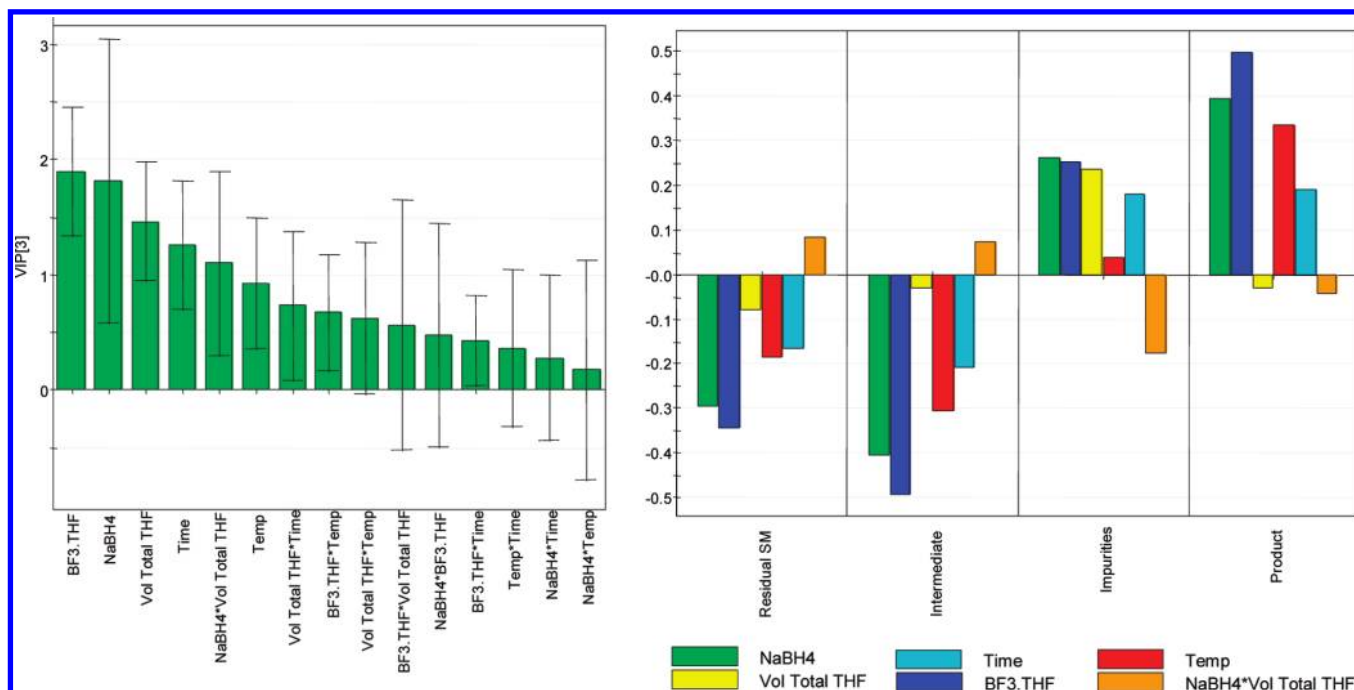
of all the original variables in  $X$  in such a way that most of the association with  $Y$  variables can be explained. The weights of the linear combination are called loadings, and the resultant linear combinations are called scores. The next main function of PLS is dimensional reduction. As in the case of PCA where dimensional reduction is achieved by explaining most of the variation in the  $X$  matrix, PLS can achieve dimensional reduction when the first few linear combinations of the  $X$  matrix can explain most of the variation in the  $Y$  matrix. The resulting model is often referred to as a 2-, 3-, or 4-component PLS model.

A 2-component PLS model was fitted to the data (one component through cross validation and one by forced fitting to allow meaningful interpretation from the scores and loadings plots) in which approximately 28% of the variation in the input data was used in modelling approximately 64% of the response variation. The PLS score plot in Figure 4 (left) is invaluable in identifying patterns in the design configuration, whereas Figure 4 (right) provides an overview of the correlations between all

variable effects for the first two components. The loadings plot shows the relationship between the main effects, two factor interactions and the four responses simultaneously. This plot can also provide an indication of how the input variables influence the response variables for the components shown. Patterns seen in this loadings plot may be brought back to the score plot to see which observations are closest to fulfilling the objectives of the study. Amongst the responses, both pyrrolidinone intermediate and the residual starting material are correlated with each other as they are situated close together in the loadings plot (Figure 4, right).

This would suggest that this pair of responses has similar profiles different from those observed with impurities and product which are both well dispersed. NaBH<sub>4</sub>, BF<sub>3</sub>·THF and temperature are positively correlated with product whereas negatively with the intermediate and the residual starting material, respectively. This would suggest that as NaBH<sub>4</sub>, BF<sub>3</sub>·THF, and temperature increase, the product is anticipated to increase, whereas the pyrrolidinone intermediate and the

**Figure 4.** PLS score plot of runs 1–20 of Tables 2 and 3 (left) and loading plot of variable effects for components 1 and 2 (right).



**Figure 5.** VIP for overall influence (left) and PLS reg. coefficients pooled over components 1 and 2 (right).

residual starting material are not: a required and desirable outcome from the objectives of the study.

The total volume of the THF is positively linked with impurities, suggesting that levels of impurities can be potentially reduced by reducing the amount of solvent used. Given that this factor has little impact on any other responses, it provides a clear indication forward for its control. A condensed review of the model interpretation is available through the variable influence on projection parameter plot (Figure 5, left). This provides a compressed summary of the overall influence of each input variable across all dimensions and response variables. Variables with VIP larger than  $\sim 1$  are considered to be the most influential for the model.  $\text{BF}_3 \cdot \text{THF}$ ,  $\text{NaBH}_4$ , the total volume of THF, time, and the interaction between  $\text{NaBH}_4$  and the total volume of THF contribute most strongly to the modelling of the five responses.

To eliminate some of the insignificant terms in the model, all effects with a  $\text{VIP} < 1$  were removed from the model with the exception of temperature which was retained for chemically practical reasons. Note, the VIP is a squared function of the PLS weights and will lead to only positive numerical values. The resulting 2-component PLS model utilized 52% of the input data for modelling approximately 60% of the response variation. By doing so, the earlier findings discussed in the previous paragraph were confirmed. Size and sign of PLS regression coefficients for scaled and centered data pooled over the two components for each response are given in Figure 5 (right). This plot indicates small or large, negative or positive influences of the input variables on the resulting responses and further aids model interpretation. A similar profile of coefficients is encountered for the pyrrolidinone intermediate and the residual starting material. As mentioned earlier, a close resemblance of these plots indicates that the intermediate and the residual starting material are strongly correlated with each other. This also substantiates the fact that these responses are situated close

to each other in the loadings plot of Figure 4. The score and loadings plots in conjunction with each other can be used to unveil observations likely to meet the objectives of the study. On first sight, the cluster of observations 3, 7, 12, and 16 (Table 3) would seem to be in a good position to meet the general demands for all the responses as they are located in the same quadrant and position in the scores plot as product, diagonally opposite to both the intermediate and the residual starting material and well distanced from impurities. This would suggest higher levels of conversion, lower levels of intermediates and residual starting material as well as acceptable levels of impurities. The remaining two conditions of reducing the time needed for the reaction to reach completion and reducing the amount of  $\text{NaBH}_4$  used for the synthesis of the product to facilitate the work-up of the reaction cannot be achieved simultaneously. Although run 12 has a reaction time of only 6 h, it has 6 equiv of  $\text{NaBH}_4$  whereas run 7 has a lower level of 3.75 equiv of  $\text{NaBH}_4$  but the reaction time is for sure longer than 6 h.

#### 4. Process Analytical Technologies (PAT) for Online Monitoring

Although Scheme 1 illustrates a conversion of functional groups that might look very simple, the use of  $\text{NaBH}_4$  in combination with  $\text{BF}_3$  was leading to a number of “unclear” transient intermediates, generically referred as boron-derivatives. Standard HPLC analysis in combination with LC/MS techniques allowed following the reduction of the starting cyclic imides **1** to the pyrrolidine derivative **3** via the lactame **2**, but did not close the gap around the understanding of the transient species. This was remedied by using online process analytical technology<sup>18,19</sup> to gain a continuous view of the reaction profile avoiding the need to take off-line samples. A 20-L vessel was equipped with a mid-infrared probe ReactIR 4000 from Mettler-Toledo,<sup>20,21</sup> and infrared spectra were acquired at 25 °C every

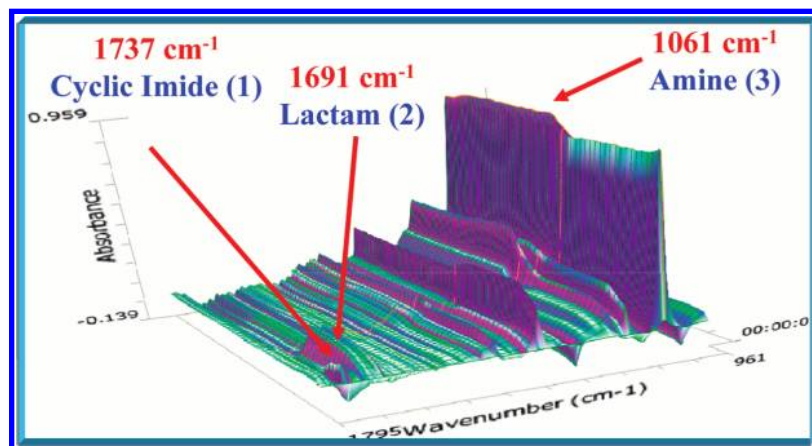


Figure 6. ReactIR three-dimensional plot.

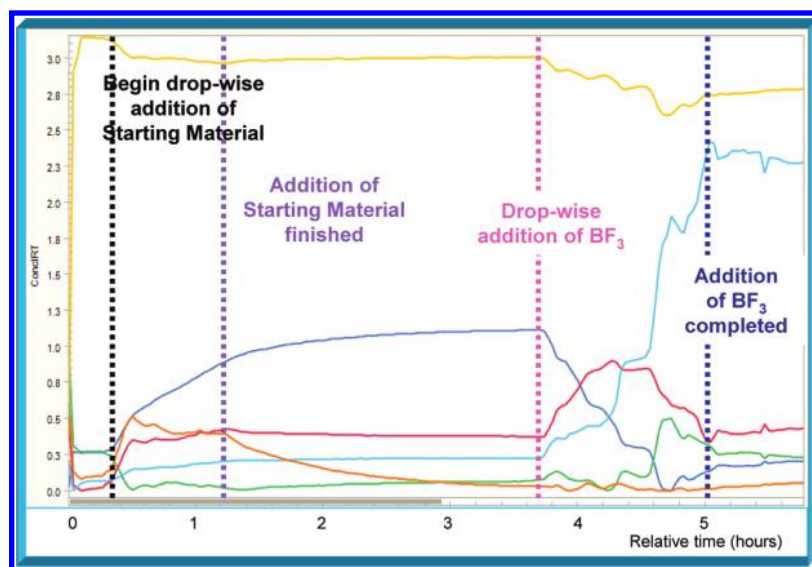


Figure 7. IR monitoring of reaction progression.

2 min for the first 7 h and every 10 min for the additional 22 h. Experimentally, the starting material was stirred with  $\text{NaBH}_4$  in THF for the first 3.5 h; then it was treated with the  $\text{BF}_3$  in THF complex over 1.5 h, and finally was left running for a further 24 h. This experimental procedure is similar to what was done in the SK233; the amounts of reagents and THF used were the ones indicated in the “current target” column of the Table 1. The spectra variations were interpreted with the ConcIRT software.

This software helps the quick deconvolution of the different components in the reaction mixture, even if it is important to bear in mind that it works through mathematical models and the results obtained from ConcIRT should be critically evaluated using chemical knowledge. Figure 6 shows the three-dimensional plot identifying the species indicated in Scheme 1.

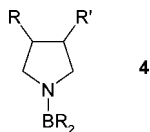
- (18) Guidance for Industry: PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance; Pharmaceutical CGMPs (Current Good Manufacturing Practice). U.S. Department of Health and Human Services, Food and Drug Administration: Washington, DC, September, 2004.
- (19) Hinz, D. C. *Process Anal. Technol.* **2004**, *1*, 16–18.
- (20) Marziano, I.; Sharp, D. C. A.; Dunn, P. J.; Hailey, P. A. *Org. Process Res. Dev.* **2000**, *4*, 357–361.
- (21) Socrates, G. *Infrared and Raman Characteristic Group Frequencies: Tables and Charts*, 3rd, ed.; John Wiley & Sons: New York, 2001.

The use of multivariate data together with a chemometric-based analysis and careful examination of the peak positioning in the IR spectra<sup>18</sup> allowed the identification of six different species during the reaction progression. Figure 7 shows the profiling of these six components during the first six hours of reaction.

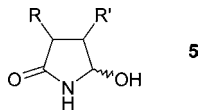
The component colored orange was identified as the starting cyclic imide **1** via bands at 1771 and 1737  $\text{cm}^{-1}$ . The light-blue component was identified as the final pyrrolidine **3** with diagnostic bands at 1000–1200  $\text{cm}^{-1}$ . The yellow component was identified via bands at 1069 and 910  $\text{cm}^{-1}$  as THF. The red component had a spectrum that resembled that of the final product **3** with an additional intense band at 1330  $\text{cm}^{-1}$  attributable to an *N*-boranyl derivative **4** (Figure 8).

The blue component was identified as an intermediate formed soon after the addition of  $\text{NaBH}_4$ . As the carbonyl stretches at 1771 and 1737  $\text{cm}^{-1}$  decreased, two bands at 1695 and 1664  $\text{cm}^{-1}$  increased. These two peaks, together with bands at 1050 and 1332  $\text{cm}^{-1}$  could be in agreement with a hydroxypyrrolidinone **5** derivative (Figure 9).

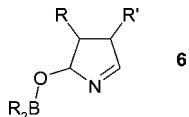
The green component appeared and then disappeared during the reaction. This transient intermediate showed intensity bands at 1175 and 1000  $\text{cm}^{-1}$  due to C–O stretch (single bond) from



**Figure 8.** *N*-Boranyl derivative **4**.



**Figure 9.** Hydroxypyrrolidinone **5**.



**Figure 10.** Azacycloborinate derivative **6**.

reduction of the carbonyl, indicating a possible C–O–B bond. An additional band in the region between 1590–1690  $\text{cm}^{-1}$  could be ascribable to an imine, suggesting this component to be an azacycloborinate derivative **6** (Figure 10). After an acidic work-up, this component should lead to the previously seen lactam derivative **2**.

## 5. Conclusion

In contrast to classical designs which are based on elegant arrangements of balance and precision, D-Optimal design techniques provide a pragmatic and fit for purpose approach on reaching a goal, this goal being to minimise the generalized variance of the estimates based on a prespecified model. They have and will always remain a useful option to use in cases where classical designs do not apply and the model must be designed for special terms. PLS has been an effective method of analysis in gaining an invaluable insight into the objectives

of the study. Whilst it has an advantage of having the ability of simultaneously modelling a number of correlated responses, the very nature of projection methods provides a latent variable model which is not always an ideal way of considering processes, especially if there are only a few response types. For that reason, it may not be considered the best tool for a more careful modelling that one may want to do. There is the notion that this method has become well used because of the widespread availability of the software, and a lack of familiarity in the scientific community of other multivariate software. A potential downside of the PLS approach could arise in situations where a factor is only effecting a single response in which case the overall analysis may mask this effect and this could be key if it is an impurity that is being reduced. In addition, PAT techniques were invaluable in providing an in-depth study of the mechanism and the process understanding of such a simple reaction, helping to allow better appreciation of the reaction profile.

## Acknowledgment

We acknowledge Martin Owen, Marion Chatfield, Ivana Damjanov, Paolo Repeto, Mick Ireson, Gillian Amphlett, Davide Tamarin, and Marsia Tampieri for the invaluable advice and encouragement they have provided us during the preparation of this paper.

## Supporting Information Available

Experimental design and process analytical technology data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review November 3, 2009.

OP900286R