

# Rational Screening: Parallel Experimentation and Predictive Modeling applied to Chemical Process R&D

**GÖRAN VERSPUI\***  
**MARCEL HOOGENRAAD**  
**ERIC DAMEN**  
**ERIK-JAN RAS**  
**NIELS LUCHTERS**

**Avantium Technologies BV**  
**Zekeringstraat 29**  
**NL-1014 BV Amsterdam**  
**The Netherlands**

**Tel +31 20 58 68 080**  
**Fax +31 20 58 68 085**

**goran.verspui@avantium.com**

**\* Corresponding author**

## INTRODUCTION

In the pharmaceutical industry, currently much effort is being spent to increase the efficiency in Chemical Process R&D. The general goal is to shorten the development times for production processes of new chemical entities (NCE's) to facilitate a shorter time to market and also to maximize the knowledge return from the investments in process R&D. In the related area of drug discovery, automated parallel experimentation, combinatorial synthesis, computer aided molecular design, and technologies that evolved from these innovations, have been applied to increase the efficiency of the research efforts. Indeed, huge investments have been made to increase the chances of finding suitable drug-candidates.

In process R&D however, the implementation of automated parallel experimentation and computer-aided research has not yet been taken up as promptly compared to the drug discovery arena. Despite the wish to innovate the development pathway from bench-scale medicinal chemistry to fully validated, economically attractive large-scale production, there are a few hurdles to be taken in order to lift also this area of research to a higher level. The requirements for automated parallel experimentation are more challenging and the financial benefits of efficient R&D often become apparent only after a more extended period of time (e.g., during production or even life-cycle management), which may explain the much lower investments compared to those in drug discovery.

In this paper we discuss some of the advances we have made by applying new, innovative approaches to Chemical Process R&D, such as high throughput experimentation and statistical modeling. In our opinion, the pharmaceutical industry and its outsourcing partners in the production of pharmaceutical intermediates will certainly benefit from the use of these methods.

## HIGH THROUGHPUT EXPERIMENTATION IN CHEMICAL PROCESS R&D

In principle, the speed and efficiency in the laboratory can be improved by increasing the average experimental throughput. By conducting a multitude of experiments in automated parallel reactor stations, the traditional chemist is easily outnumbered in terms of data generation. Some types of chemical reactions are unfortunately not that easy to perform in downscaled parallel reactor platforms. This could be due to difficulties with atmosphere control, mass transfer limitations or inaccurate reagent addition, etc. In other words: there are often issues with the parallelizability of experiments.

Catalytic asymmetric transformations, in particular asymmetric hydrogenations, are attractive chemical transformations for the production of enantiopure compounds (1). When performing these reactions, often air sensitive transition metal complexes are used, that require handling under inert conditions. The presence of oxygen can have a deleterious effect on the performance of the catalysts and especially at a small scale, traces of oxygen can result into a complete catalyst-deactivation.

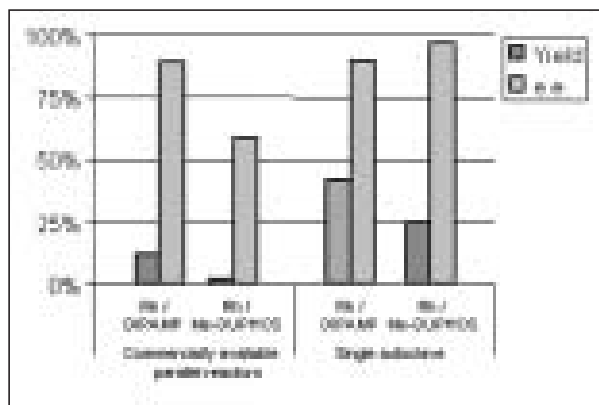


Figure 1 - Not all reactions are readily parallelizable, the results of an asymmetric hydrogenation reaction on a prochiral C=C bond, obtained in a commercially available parallel reactor station, are poor in comparison with the results from a single autoclave experiment.

We have studied the asymmetric hydrogenation of a C=C bond, catalyzed by the Rh-diphosphine catalysts Rh/DIPAMP and Rh/Me-DUPHOS.

As shown in Figure 1, the results that were obtained from an experiment at 1 ml scale in a 'state of the art' commercially available orbital shaken parallel autoclave setup, were rather poor in comparison with the results of a 25 ml scale single-autoclave experiment. After a careful analysis of the experimental procedures, we concluded that for this reaction a more sophisticated inerting and handling procedures of both the reagents and the reactor platform were required. Thus, when performing asymmetric catalytic transformations at small scale in parallel, the strict inert conditions and the exact way of performing the experiments have become critical parameters for a satisfactory result (Figure 1).

To overcome the issues with atmosphere control and mass transfer limitations, we have built a reactor station, which can host 16 reactor blocks, each block containing 12 minireactors (16 x 12 = 192 parallel reactors). This reactor arrangement allows for a large flexibility in operation, one could decide to use the full capacity of 192 reactors, or could also choose for smaller sets of experiments, e.g. 12 or 24. The reactors are magnetically stirred, and show excellent mass-transfer characteristics. Cross contamination studies have shown that even relatively volatile compounds such as n-decane and pyrrolidine do not end up in neighboring reactors.

Moreover, memory effects (from catalytically active remains from previous experiments) are eliminated by the use of new inserts for each experiment. The reactor blocks are easily opened and closed in a glove box, excluding oxygen and moist, by a simple procedure, taking only a few seconds. Outside the glove box the closed reactors are connected to the reactive gas lines, under exclusion of air (Figure 2).

Equally important as the hardware design, is that the reactor station is part of a workflow. This workflow also comprises automated solids dispensing

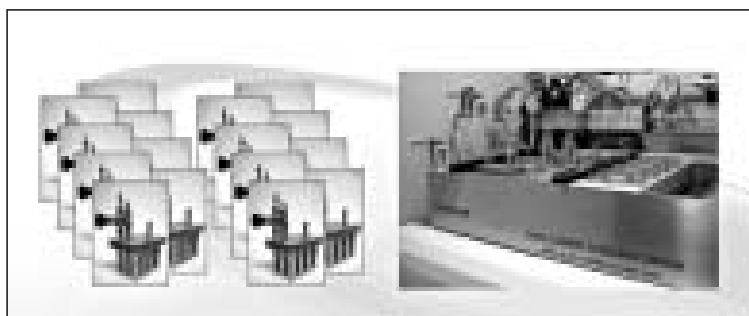


Figure 2 - Avantium's 192-parallel reactor station for studying gas-liquid batch reactions.

and automated liquids dispensing under an inert atmosphere and automated sampling and sample preparation. All components are adjusted for an efficient and reproducible execution of the experiments.

As an example, an experiment was performed using Noyori's [RuCl<sub>2</sub>((S)-BINAP)((S)-DAIPEN)] catalyst in the reduction of acetophenone to enantiomerically enriched 1-phenylethanol. When the reaction was performed at half capacity (96 parallel reactors, all with the same contents) with a catalyst loading of 0.02%, after two hours reaction time a quantitative conversion was obtained in all reactors with enantiomeric excesses (e.e.'s) ranging from 94-96%. In a subsequent smaller session with 24 parallel reactors using 0.01% of catalyst at 1 hour reaction time, across the reactors conversions ranging from 40% to 50% yield were obtained while maintaining the high enantioselectivity. These results show that our experimental setup is perfectly capable of handling oxygen-sensitive materials in a parallel fashion. For comparison, in typical screens the catalyst loadings generally vary between 0.5 and 5% (at substrate concentrations of 30-10 volumes) with reaction times between 5-24 hours. We thus conclude that our equipment and workflow is qualified in terms of performance and reproducibility for performing asymmetric hydrogenation screens. The equipment and workflow was also found to be suitable for other gas-liquid

batch reactions, such as hydroformylation (2), carbonylation (3), reductive amination (4), selective nitro reduction (5), aerobic alcohol oxidation (6) and polymerization (7) reactions (Figure 3).

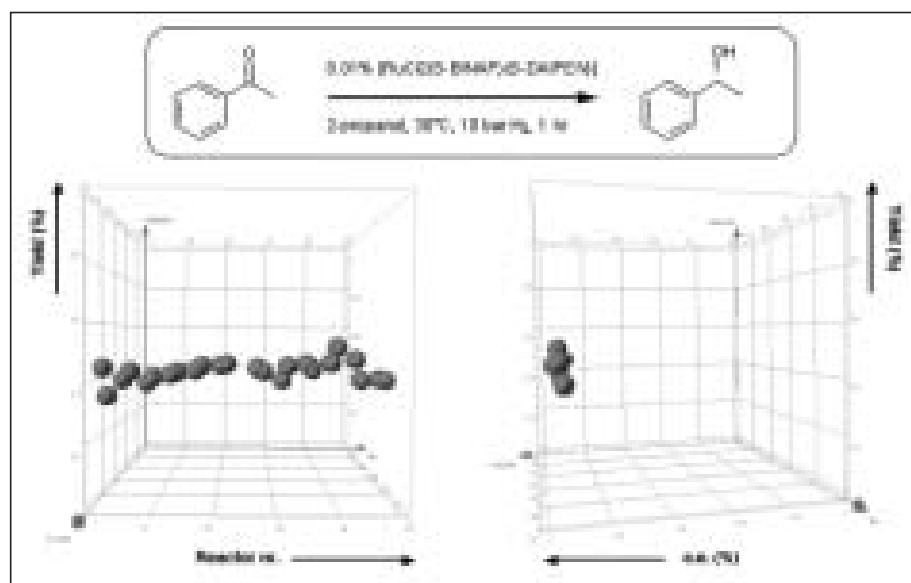


Figure 3 - An asymmetric hydrogenation experiment in 24-parallel reactors, showing a good reproducibility across the reactors.



Property Relationships (QSAR/QSSR/QSPR) using practical, fast computational methods (practically running on normal desktop computers) in combination with statistical analysis. By converting structures into descriptor information and correlating these descriptors with measured responses of a chemical reaction (e.g. enantioselectivity), we are able to predict the effect of a change in structure to the response of the system. In contrast to many other methods, this approach does not rely on knowledge of, or conjecture upon reaction mechanisms. The empirical approach makes this a practical and powerful tool for in route selection.

As an illustration, we have studied the application of QSSR modeling to oxazaborolidine mediated asymmetric reduction reactions (10). A selection of 28 chiral aminoalcohols was treated with BH<sub>3</sub> in THF in an automated orbital shaken parallel reactor station. The *in situ* formed oxazaborolidines were used in the asymmetric reduction of acetophenone. The reactions proceeded to quantitative conversion and from the duplicate, triplicate experiments we concluded that we were able to adequately and reproducibly perform this type of chemistry in a parallel fashion, affording enantioselective excesses (e.e.'s) in agreement with the literature (Figure 5).

The high quality and reproducibility of the experimental results, combined with an evenhanded spread in observed e.e.'s, made the data set attractive for modeling. As illustrated by the workflow in figure 6, we applied descriptor calculations scripts to the structures (training set, 24 structures), converting them into a large array of numbers that contain the structural information. By means of partial least squares regression (PLS) a correlation was found between the descriptor information and the e.e.'s that were observed in the automated experimentation. A prediction set of 4 aminoalcohols (these structures were not used to build a model) was

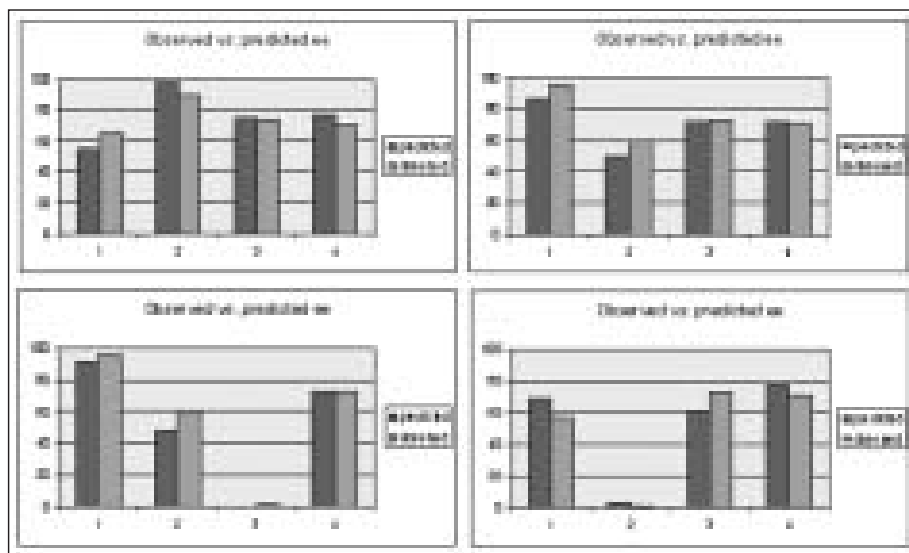


Figure 6 - QSAR modeling for e.e. prediction in oxazaborolidine mediated acetophenone reduction. The predicted vs. the measure e.e. of 4 different prediction sets (in 4 different models) is shown.

interpolated using the model and the predicted e.e.'s were in good agreement with the observed values. We have repeated the modeling with a number of different training set /prediction set combinations all of them providing a good correlation ( $r^2 \sim 0.97$ ) and predictivity ( $q^2 \sim 0.80$ ) (Figure 6).

This practical statistical approach, ignoring mechanisms and transition states and avoiding heavy computation, was found to be

applicable to many types of chemical transformations, including biocatalysis. For example, for the enantioselective hydrolytic ring opening of racemic epoxides by bacterial epoxide hydrolases, reported by Faber and coworkers (11), transition state modeling requires structure information of the active site. For bacterial epoxy hydrolases, which are used as whole cells in the hydrolytic reaction, such structural information was not available. Fortunately, a descriptor based QSSR

modeling study does not require this information (Figure 7).

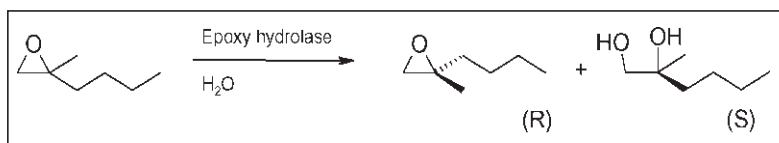


Figure 7 - The enzymatic S-specific hydrolytic ring opening of racemic epoxides, catalyzed by bacterial epoxyhydrolases.

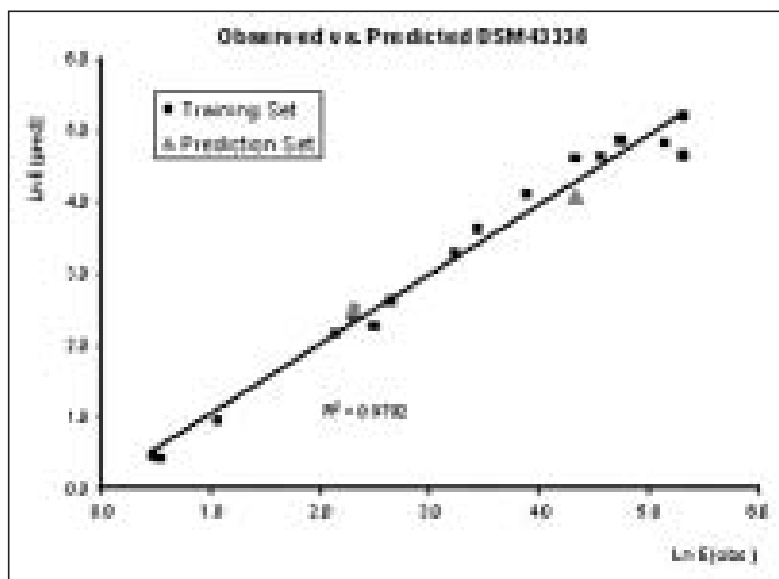


Figure 8 - A descriptor based QSSR model correlating the structures of the epoxide substrates with the measured E-value of the bacterial epoxide hydrolase DSM 43338. The data points in red represent the estimated E-value for the epoxides in the prediction set (this data was not used to construct the model)

We constructed a descriptor-based model and found a good correlation between the structural information of the epoxides (training set consists of ca. 20 epoxides) and the observed E-value of the epoxyhydrolase DSM 43338. A prediction set (2 epoxides that were not used to build the model) was fitted into to the correlation and as shown in figure 8) the predicted E-value of these independent epoxides was in good agreement with the E-value observed in the reaction (Figure 8).

The short, but effective modeling sessions are convenient in the selection of catalysts, reagents and solvents in screening programs. The approach could also be useful in the development of improved catalysts.

## CONCLUSIONS

By analogy with the automation efforts in drug discovery, it is recognized that just increasing the experimental throughput alone is not the answer for cost effective process R&D. Equipment and work flows need to be carefully designed, tested and validated before they can effectively be applied to screening sessions aimed at finding practical solutions for a production processes. Smart selection of variables and the careful use of DoE are providing maximum information from the parameter space in a reduced set of experiments. Furthermore, QSAR modeling studies offer rapid, practical and accurate predictions,

without the need for the knowledge or speculation of the reaction mechanism; especially attractive in modern organic chemistry where mechanisms are still under investigation. In conclusion, the new approaches discussed in this paper could provide for substantial efficiency increases to process R&D.

## ACKNOWLEDGEMENTS

The expertise, creative efforts and team spirit of our colleagues at Avantium Technologies is gratefully acknowledged.

## REFERENCES

- 1) *Asymmetric Catalysis on Industrial Scale*, eds. Blaser, H.U. and Schmidt, E.; Wiley-VCH, Weinheim, Germany, 2004; *Catalytic Asymmetric Synthesis, 2nd edition*, Ojima, I. (ed.), Wiley-VCH, Weinheim, Germany, 2000; *Comprehensive Asymmetric Catalysis*, eds. Jacobsen, E.N.; Pfaltz, A.; Yamamoto, H.; Jacobson, E.N.; Springer Verlag, 1999

- 2) VERSPUI, G.; HOOGENRAAD, M. *unpublished results*
- 3) VERSPUI, G. *unpublished results*
- 4) VERSPUI, G.; HOOGENRAAD, M.; VAN DER LINDEN, J.B.; DANI, P. *unpublished results*
- 5) HOOGENRAAD, M.; VAN DER LINDEN, J.B.; SMITH, A.A.; HUGHES, B.; DERRICK, A.M.; HARRIS, L.J.; HIGGINSON, P.D.; PETTMAN, A.J. *Org. Process Res. Dev.* **2004**, in press
- 6) TEN BRINK, G.J.; ARENDS, I.W.C.E.; HOOGENRAAD, M.; VERSPUI, G.; SHELDON R.A., *Adv. Synth Catal.*, **2003**, *345*, 1341-1352
- 7) MEERENDONK, W.J.; DUCHATEAU, R.; KONING, C.E.; GRUTER, G.J. *Macromol. Rapid Commun.* **2004**, *25*, 382
- 8) MCKAY, B.; HOOGENRAAD, M.; DAMEN, E.W.P.; SMITH, A.A. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 966
- 9) LIPKOWITZ, K.B.; KOZLOWSKI, M.C. *Synlett* **2003**, 1547
- 10) HOOGENRAAD, M.; KLAUS, G.M.; ELDERS, N.; HOOISCHUUR, S.M.; MCKAY, B.; SMITH, A.A.; DAMEN, E.W.P. *Tetrahedron: Asymm.* **2004**, *15*, 519
- 11) FUNAR-TIMOFEI, S.; SUZUKI, T.; PAIER, J.A.; STEINREIBER, A.; FABER, K.; FABIAN, W.M.F. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 934

