Automated Feedback in Flow for Accelerated Reaction Screening, Optimization, and Kinetic Parameter Estimation

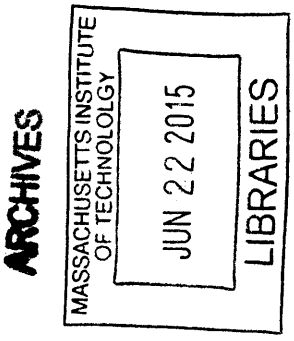
by

Brandon Jacob Reizman

B.S. Chemical Engineering, University of Illinois at Urbana-Champaign (2009)
M.S. Chemical Engineering Practice, Massachusetts Institute of Technology (2011)

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemical Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY



May 2015 [June 2015]

Signature of Author..... Signature redacted

Department of Chemical Engineering
May 15, 2015

Certified by ... Signature redacted .....

Klavs F. Jensen
Warren K. Lewis Professor of Chemical Engineering
Professor of Materials Science and Engineering
Thesis Supervisor

Accepted by ..... Signature redacted .....

Richard D. Braatz
Edwin R. Gilliland Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Automated Feedback in Flow for Accelerated Reaction Screening, Optimization, and Kinetic Parameter Estimation

Brandon Jacob Reizman

ABSTRACT

With the cost to discover and develop a drug now estimated to exceed $2 billion, the pharmaceutical industry is in search of innovative and cost-effective ways to reduce process footprint, minimize lead times, and accelerate scale-up. One path to achieving these goals is in the adoption of continuous processing. Among the many advantages offered by the use of continuous flow systems is the ease of integration of automation and online analytics for real-time monitoring of reactions. The further incorporation of feedback into automated systems invents an even greater possibility: the use of algorithms to intelligently manipulate different continuous variables—for instance temperature, time, and concentration—until an optimal synthesis is achieved. This thesis opens by reviewing the most recent applications of feedback optimization in flow. The same methodology is then applied to the estimation of reaction kinetics in a series-parallel S_NAr reaction network.

Unfortunately, the most challenging aspect of reaction development tends not to necessarily be the continuous variables, but rather the enumerate combinations of discrete variables—e.g. catalysts, ligands, and solvents—that, when paired with the continuous variables, give rise to changes in the reaction mechanism or kinetics. To address this problem, this thesis introduces a more general approach to reaction optimization with the construction of an automated segmented flow system, wherein reactants are confined to sub-20 μL slugs flowing through a heated Teflon tube microreactor and analyzed online by LC/MS. The system allows for manipulation of both discrete and continuous variables, making it possible to simultaneously screen reagents while optimizing the reaction. A sequential adaptive response surface methodology for optimizing both discrete and continuous variables is presented. The algorithm employs optimal design of experiments in feedback to greatly accelerate convergence of the mixed integer nonlinear programming (MINLP). Examples of real-time simultaneous screening and optimization are explored, including optimal solvent selection in a selective alkylation reaction and optimal palladacycle-ligand precatalyst selection for Suzuki-Miyaura cross-coupling reactions. We conclude by showing how the automated system can be utilized to gain further understanding of reaction mechanisms and kinetics and by demonstrating that the optimal results can be scaled to larger chemical syntheses.

Thesis Supervisor: Klavs F. Jensen
Department Head, Chemical Engineering
Warren K. Lewis Professor of Chemical Engineering
Professor of Materials Science and Engineering

To my dad, who thought I should be an accountant,
and to my mom, who thought otherwise.

ACKNOWLEDGEMENTS

Undoubtedly the person to acknowledge first and foremost for this thesis is my advisor, Prof. Klavs Jensen. He has been a great mentor and supporter of my work, and most of the innovations shared below would have not been possible without either his foresight or the culture of community and knowledge he has fostered in his research group. I still remember sitting in his office in fall 2010, with my thesis proposal looming, wondering what would make for a captivating project to embark on for the rest of my thesis. When Prof. Jensen pitched the idea of simultaneous catalyst screening and reaction optimization, I told him first hand that it sounded like an amazing idea, but that I had no idea how to do it. His response: "that's what makes it a good thesis project."

Beyond Prof. Jensen, I have benefited greatly—both academically and personally—from the mentorship of many others in the MIT community. The faculty in particular have always challenged me to look at problems in new ways and encouraged me to go far beyond what I would have first considered to be a lofty goal. I have received insightful guidance on this research in particular from my thesis committee Prof. Steve Buchwald, Prof. Richard Braatz, and Prof. Paul Barton, as well as another member of the MIT-Novartis collaboration, Prof. Tim Jamison, and the continuous manufacturing group at Novartis led by Dr. Berthold Shenkel. Funding from the Novartis-MIT Center for Continuous Manufacturing made this research possible. As a teaching assistant, I was fortunate to benefit from the mentoring of Prof. Michael Strano and Prof. Hadley Sikes, and much of my ability to deduce problems and effectively communicate solutions can be attributed to the support I received from Prof. Claude Lupis and Prof. Bob Hanlon at Practice School. Through work in the classroom, the rest of the chemical engineering faculty and the MIT faculty on the whole has opened my eyes to new ideas and innovative strategies for problem solving.

I have been privileged to work in research with a group of equally talented and thoughtful colleagues. It is hard to imagine this particular thesis being feasible without the innovativeness of Dr. Jon McMullen and Dr. Jason Moore, who were first to demonstrate the incredible value of automated optimization in flow and were my educators in software development, online instrumentation, and assembling high-throughput systems. Recently, I have benefited immensely from discussions about instrumentation and software from new members of the automation team, Isaac Roes, Kosi Aroh, Connor Coley, and Dr. Milad Abolhasani. Many of the physical

components of the system I demonstrate herein were contributions of the thought, time, and effort of Dr. Andrea Adamo, Dr. Baris Ünal, Dr. Everett O'Neal, and Dr. Patrick Heider, along with UROP Michaelann Rodriguez. The micromixer and silicon microreactor used in Chapter 2 were contributions from Dr. Nick Zaborenko and Dr. Lei Gu, respectively. In terms of organic synthesis, my growth as a chemist is forever endeared to Dr. Chris Smith, who tolerated my endless nights of "amateur hour" in hopes that someday I would run a column better than the driver of Jose's Taco Truck (no offense intended to Jose's). I have subsequently benefited from the teachings of Dr. Stephen Born, Prof. Steve Newman, Dr. Saurabh Shahane, and Dr. Antony Fernandes. The impact of my thesis would be far less without the guidance of those in the Buchwald and Jamison groups, in particular Dr. Yiming Wang, Dr. Nick Bruno, and Andy McTeague. Along the way, I have received a great amount of feedback on my research not just from those listed above, but from the rest of the Jensen chemical synthesis subgroup and the Jensen group on the whole.

Above all, a huge thank you must be extended to those around the MIT community who have supported my growth for the past several years. Thanks in particular go to the support staff in chemical engineering who have worked tirelessly to make my experience at MIT smooth and always enjoyable: Alina Haverty, Suzanne Maguire, Joel Dashnaw, Katie Lewis, Fran Miles, Beth Tuths, Steve Wetzel, and Brian Smith. In my time spent at MIT, I have been a member of some excellent organizations, most notably the MIT Energy Initiative, the Course X Graduate Student Council, the ChemE Graduate Student Advisory Board, the Thirsty Ear Executive Committee, and a host of intramural sports teams. I have benefited immensely from the friendships built through those groups and with other graduate students in chemical engineering, most notably Dr. Caleb Class, Dr. Rathi Srinivas, Dr. Jon Harding, Dr. Dave Borrelli, Dr. Nisarg Shah, Dr. Nigel Reuel, Tim Politano, and Su Zhu. I would like to thank my family, both my sister Caitlyn and my mom and dad for the support and encouragement they have given me not just now but in my entire academic career. And I would like to thank the love and support I have received from a very newly awarded doctor, Dr. Irene Brockman, who has worked harder than anyone to help me reach my fullest potential. Irene, we made it!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

1. Feedback Systems for the Acceleration of Reaction Development

With the cost to discover and develop a drug now estimated to exceed $2 billion,[1] the pharmaceutical industry is in search of innovative and cost-effective ways to reduce process footprint, minimize lead times, and accelerate scale-up. One path to achieving these goals that has received great attention lately is the adoption of continuous flow technology.[2-4] In the past few years, the pharmaceutical industry has begun to incorporate continuous processing into more and more small molecule syntheses, with applications ranging from the replacement of individual batch unit operations for safer,[5-8] greener,[9-12] and/or more aggressive flow reaction steps,[13-16] to the design and implementation of an end-to-end pilot plant encompassing continuous drug manufacture.[17] It is widely accepted that the continued introduction of efficient and rapidly scalable technologies for the acceleration of reaction development into manufacturing will help greatly in delivering drugs in less time and at lower cost to the patient.

A breadth of technologies and methodologies fall under the heading of "acceleration of reaction development." Much of our lab's focus has been in the development of microreaction technology,[18-21] namely the use of sub-millimeter scale reactors to achieve highly controlled chemical syntheses that can then be scaled to larger flow systems.[22-25] This is made possible by the excellent rates of heat and mass transfer[26] and minimal dispersion[27] in microscale systems, which enable easier access to the intrinsic kinetics of the reaction.[23,28-31] The reduced volumes of microreactors further allow profiling of reactions at conditions that would be too hazardous or simply impossible to achieve in batch, a few examples of these being reactions of azides,[5,32] fluorinations,[33-36] nitrations,[37,38] DIBAL-H reductions,[39-41] lithiations,[39,42-45] Grignard reactions,[39,42,46,47] and reactions in supercritical media.[48-50] Data collection from continuous flow systems is accelerated by the incorporation of online analytics such as HPLC,[22,23,30,51] MS,[52,53] GC,[28] UV-Vis,[54,55] FTIR,[28,56-59] Raman,[60] and NMR.[61] These instruments allow the experimenter near infinite access to the inner workings of the chemistry, enhancing understanding of reaction mechanisms, formation rates of intermediates and byproducts, and responses to perturbations in process conditions.

Of course, to some "acceleration of reaction development" does not only imply easier scalability, but actually running more experiments in a reduced period of time. This is the essence of high-throughput experimentation (HTE),[62-66] whereby automation and robotics are

used to rapidly conduct and analyze many experiments in parallel with little to no intervention required on the part of the scientist. These tools enable fast understanding of the reaction performance as a function of many different variables, with full reaction maps being assembled in time spans of days down to a few hours.[67,68] By minimizing the volume of each reaction sample, reaction profiling can be completed with milligram-scale quantities of expensive pharmaceutical precursers[63] or nanogram-scale quantities of expensive catalysts or ligands. Yet techniques imported from high-throughput drug discovery are often limited in scope, ranging from the limited scalability of batch results to limitations in the ability to modulate key factors such as reaction time or temperature. To alleviate these concerns, efforts have been made to employ rapid automated experimentation in flow, merging the scalability of microreaction technology with the sheer speed of HTE.[22,34,37]

Unfortunately, the direct assimilation of HTE and automation approaches into chemical synthesis introduces a new problem—the curse of dimensionality. To illustrate, consider a case study where an experimenter is interested in studying the combined effects of 10 catalysts and 10 ligands. Teasing out all of the catalyst-ligand interactions would require 100 experiments, which can be easily executed in a single parallelized screen. Now consider the addition of 10 solvents and 10 bases to the study, giving 10,000 total experiments to be run. While such a system is still solvable, the cost of these experiments is substantially more. What if the kinetics of the reaction are important, such that the experimenter needs to also study 10 temperatures and 10 reaction times? Alternatively, how much money and time are lost in screening if one of these variables is found later to have no effect on the reaction? Clearly simplifications are needed and often employed, the most well-known of these being the one-factor-at-a-time approach.[69] Yet by manipulating only a single variable at a time, then optimizing, then moving to a new variable, all of the information is lost from the system with regard to how multiple variables interact with one another, which can be critical when constructing reaction mechanisms or identifying an optimal process envelope. It is also easy to see that such an approach can lead to identification of less-than-optimal process conditions, leading to reduced yields in scale-up, or the worst-case possibility of not being able to manufacture the drug altogether.

A better approach that addresses the concerns listed above is to use feedback to engineer reactions to more optimal conditions. Unlike straight HTE, all experimental conditions are not screened upfront; rather one or multiple initial experiments are chosen as an initialization, then

an optimization routine selects the next best experiment or set of experiments to run that moves the system toward finding an optimum (Figure 1.1). To accomplish this task, the optimization routine must first assess the fitness of previous experimental data points against an objective function. For deterministic routines, the data are then extrapolated to a model, which can be as simple as a piecewise response surface or as complex as to fully describe the kinetics and transport phenomena taking place within the system. Based upon this model, a more optimal experiment is selected, observed, and then the model is updated in iteration until convergence is reached.

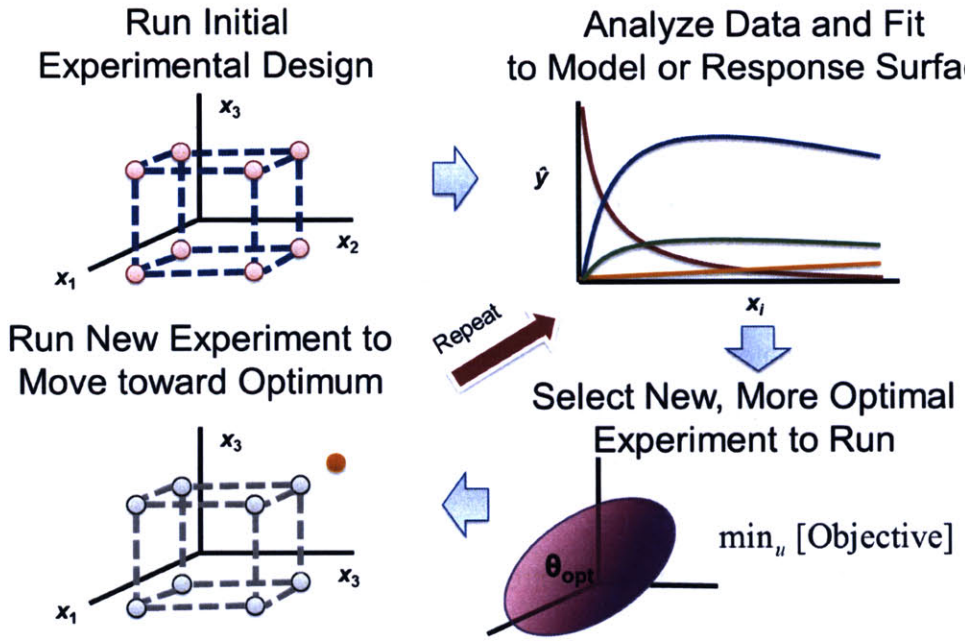


Figure 1.1 Generalized feedback loop for (deterministic) automated reaction optimization.

Naturally such an approach would be rather tedious in small-scale reaction development, hence the reason feedback optimal design has generally been reserved for instances where experimental data are expensive to collect.[70-74] With automation, however, the amount of user interface required to collect a data point becomes considerably less. With the use of continuous flow and online analytics, smart automated systems can be constructed that use experimental data collected in real-time as feedback for rapid reaction development. In the simplest case, an automated feedback system can be used to identify the optimal yield for a reaction, but algorithms can also be constructed that identify best-fit reaction kinetics and mechanisms using information theory. In more complex reaction cases, questions arise as to the best way to

elucidate reaction kinetics in a single flow system, and how to optimize reactions in which discrete variables—such as catalysts, ligands, or solvents—interact differently with continuous variables—such as temperature, reaction time, or concentration. This thesis aims to answer both of these questions, and in so doing looks for a faster, more versatile, and more economical way to accelerate reaction development.

1.1. The Tools for Feedback Optimization in Flow

The general tools for flow chemistry have been reviewed recently.[21] These comprise reagent delivery, reaction, separation, analysis, and pressure control. Figure 1.2 illustrates how all of these components can be integrated into a single continuous flow feedback system. Both reagent delivery and reaction control (temperatures and flow rates) are controlled with a central automation software. Automation software and hardware can also be used to interface with online analytical devices that may identify changes in the process (for instance disturbances in temperature or pressure) or be used to extract quantitative data from the reaction—most often yields and conversion. These data are then passed to an optimization algorithm that interprets the data—perhaps also in the context of prior experiments—and use the information to project a new set of inputs that will move the system toward the optimum of a user-defined objective. These inputs are passed to the central control system, which makes the necessary manipulations to the experiment.

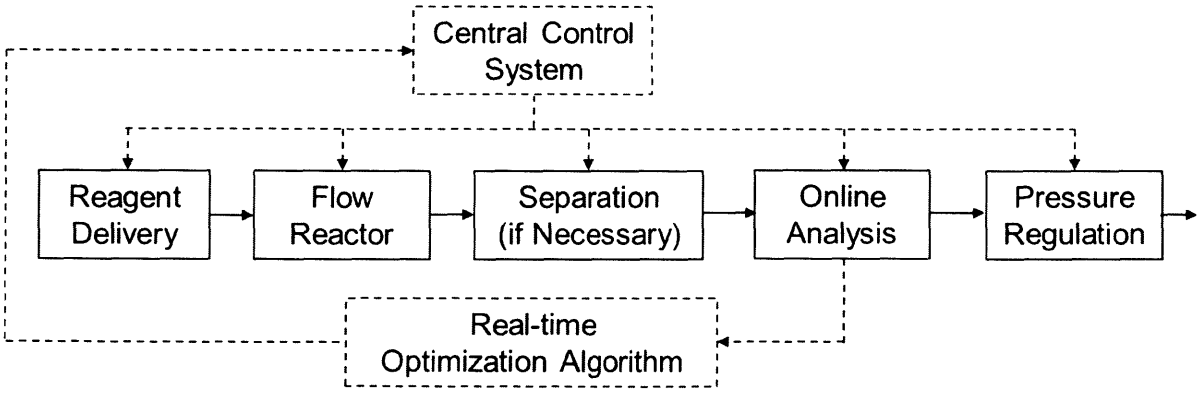


Figure 1.2. Block diagram for automated feedback optimization in flow systems.

For reagent delivery in optimization systems, our lab has had the most success using syringe pumps. These tend to produce the greatest accuracy in the flow rate range of 1 μL/min-250 μL/min, which depending on the number of pumps and size of reactor in use may translate to

residence times of 30 s to 30 min. At or below 30 s, the rate of mixing can become a key confounding variable in the optimization.[27] Reaction times of 30 min or more are usually undesirable for rapid feedback because the time to reach steady-state in continuous, non-droplet, systems is greater than or equal to triple the residence time. Long reaction times may be more suitable for flow rate ramps, where each segment of fluid is treated as a separate batch reactor.[29] For accurate optimization results, the choice of syringe material can matter greatly. Despite their convenience and low cost, plastic syringes are almost never desirable because of the tendency of the plastic plungers and barrels to buckle and produce inconsistent flow rates under even the mildest of pressures. Glass syringes are excellent in terms of accuracy and chemical compatibility; however these leak and therefore cannot be used at pressures above ~14 bar. Even prolonged use at ~7 bar will after several days require replacement of the syringe because of leaking. At high pressures, the only suitable syringe material option is stainless steel; however reagents which corrode stainless steel must be avoided.

Flow optimization systems require special treatment in reactor selection, as the nature of the experiment requires the reaction to proceed uninterrupted under a diverse and often sub-optimal range of experimental conditions. Flow reactors that are slow to reach temperature or cool down—most notably stainless steel or Teflon tubes submerged in an oil bath—tend to be undesirable for optimizations where the temperature must be manipulated regularly. Of course chemical compatibility with harsh organic reagents is of utmost concern, hence the avoidance of PDMS in flow chemistry. Our lab's preferred reactor material has been silicon because of its excellent heat transfer properties and well-known microfabrication procedures, with the only notable limitation being its incompatibility with strong bases such as NaOH or KOH. For greater chemical compatibility, glass[75,76] or silicon carbide[77] reactors are available, though these are more difficult to machine and are expensive to replace in the event of a clog. In this thesis, a new reactor is introduced which comprises a Teflon tube heated in an aluminum chuck. The advantages of this design are extensive chemical compatibility, easy replacement of the Teflon tube in the event of temperature-induced degradation or clogging, and faster equilibration to the set point temperature compared to heating in an oil bath.

The available tools for online analysis were discussed earlier, and any and all could be applied to flow optimization systems. Our lab has explored predominantly online HPLC—because of its ability to resolve a diverse spectrum of reaction products—and FTIR—because of the speed at

which data can be collected and fed back to the process. The time needed to collect data by HPLC makes LC analysis the limiting step in many instances of feedback optimization. Consequently, work has been done in recent years to improve the speed of HPLC analysis for high-throughput applications.[78,79] As is always the concern, the greater the specificity in designing an optimal LC method for a given reaction, the less generalizable the platform can be for broader scopes of substrates and products.

The design of an optimization system with automation software is dependent upon recognizing which variables are to be controlled and which observables (responses) are to be measured. As straightforward as it may seem, model inputs that cannot be controlled cannot be manipulated to achieve an optimum. Likewise outputs that cannot be measured cannot be optimized, unless there is a model relating another measured output to the desired response. Inputs can be controlled by local controllers (for instance most syringe pumps deliver at a controlled flow rate once a set point is received) or by the central controller. The choice of when to use either is usually a matter of how much control needs to be exercised and what local controllers, if any, are available (for example if the pump flow rate needs to change as a function of pressure but the pump has no pressure sensor, then this loop is built into the central controller). Sometimes sensory measurements trigger decisions by the control system, such as the decision to switch valves and flush out a reactor on account of an observed increase in pressure.[80] Constraints with regard to maximum and minimum values of the controlled inputs are almost always incorporated into the optimization and control routine and depend upon the physics of the system. For sake of clarity in the subsequent sections and chapters, the responses are the measurements that factors into the objective of the optimization—for instance product yield, which is estimated based on a calibration model of the measured analytical signal. In all of the accounts presented herein, LabView (National Instruments Corporation) was used as the central control software for manipulating inputs. Responses were sometimes interpreted directly by LabView, or worked up analyses were passed to LabView from software such as ChemStation (Agilent Technologies) or iC IR (Mettler Toledo International Inc.). Optimization routines executed in MATLAB (The MathWorks Inc.) or LabView provided new set points for inputs.

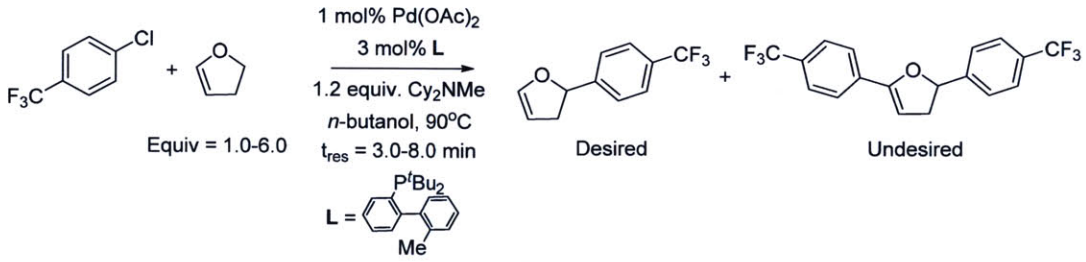
1.2. Reaction Optimization "from Scratch"

With no prior information or models to which to resort, a plausible optimization strategy must accept input factors known to influence the desired response and interpret relationships among these variables which lead to improvement—and hopefully optimality—in an objective. Such a model-free strategy is commonly referred to as "black box" optimization. An important aspect of black box optimization is that, from an engineering standpoint, little to no modeling information is gained from the search. Consequently, optimal results have no guarantee to transfer across scales. This is of course the major concern of using black box approaches for the application of reaction development, where the ultimate goal is scale-up. However, if the feedback system is engineered to detect intrinsic reaction rates, the results of black box optimization routines can still have great utility, especially when little to no *a priori* information is known about the chemistry. If optimization of the flow system is desired simply to increase production rate or yield in the same flow reactor, black box strategies are excellent simple tools to identify improved or perhaps optimal reaction conditions.

Some of the earliest methods for incorporating black box optimization algorithms into automated microreactor experiments were developed by Krishnadasan *et al.*[81] and by McMullen and Jensen.[22,51] Krishnadasan *et al.* employed a Stable Noisy Optimization by Branch and Fit (SNOBFIT)[82] algorithm to optimize the automated synthesis of CdSe quantum dots. Nanoparticles were monitored inline by fluorescent emission measurements from a CCD spectrometer, and a feedback loop allowed for tuning of the microreactor temperature and flow rates in order to maximize the quantum dot emission at a specified wavelength. McMullen and Jensen compared the SNOBFIT algorithm with two local-search black box optimization algorithms—Nelder-Mead Simplex and steepest descent—in studying the Knoevenagel condensation reaction of *p*-anisaldehyde and malononitrile. Species concentrations were monitored online by HPLC. The two-dimensional optimization of reactor residence time and temperature consistently identified the maximum objective function value as occurring at the maximum allowable reaction temperature of 100°C and the minimum residence time of 30 min, regardless of the optimization method employed. The fastest convergence upon the optimum was achieved with the gradient-based steepest descent method.

Though only a local search strategy, the Simplex method[83] has been used extensively in black box flow optimization systems. In its simplest implementation, Simplex selects $k+1$ points for a

k-dimensional optimization, then after identifying the least optimal point from that experimental set, reflects the least optimal point across the polygon defined by the other *k* points to define a new Simplex. When no improvement in the optimization is achieved, the size of the Simplex is reduced and the method repeats. This strategy is both simple to implement and does not require approximation of a gradient, which caters well to expensive experimentation. In one application, McMullen and Jensen demonstrated use of the Simplex method in a four-dimensional optimization of benzaldehyde production, increasing the yield from 21% to 80%.[51] The Simplex method was further applied in optimizing the number of alkene equivalents and reaction time in a selective Heck reaction (Scheme 1.1).[22] In both cases, syringe pumps were connected to a temperature-controlled silicon microreactor, with analysis performed by online HPLC (Figure 1.3). The feedback optimization algorithm was executed in MATLAB and the system was controlled with LabView software. For the Heck reaction example, nine sets of reaction conditions were then scaled 50 times to a 7 mL Corning Advanced-Flow glass reactor module. The scaled optimum was found to be consistent with the microscale optimum and resulted in the synthesis of 26.9 g desired product at 80% yield.



Scheme 1.1. Heck reaction optimization studied by McMullen *et al.*[22]

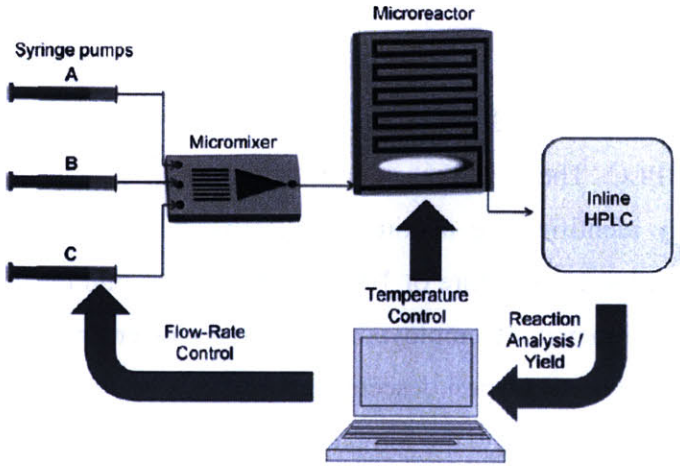
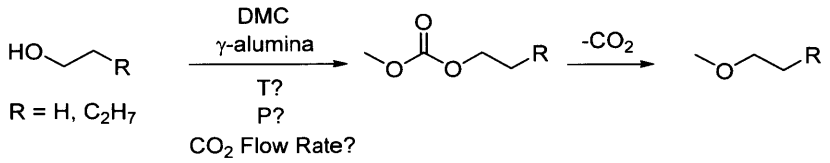


Figure 1.3. Automated feedback loop used by McMullen *et al.* for the optimization in Scheme 1.1.[22]

Poliakoff and coworkers[84-88] have been interested in the use of feedback for online optimization of the methylation of primary alcohols by dimethyl carbonate (DMC) in supercritical CO_2 (Scheme 1.2). Here the advantage of the flow system was both in rapid optimization and in the accessibility of supercritical reaction conditions. In the original system, Parrott *et al.*[84] flowed the methylation reactants through a packed bed of γ-alumina in the presence of supercritical CO_2 and monitored the reaction yield by online GC. The Super Modified Simplex algorithm[89] was used to optimize the reaction yield with respect to temperature, pressure, and the flow rate of CO_2. The Super Modified Simplex algorithm functions similarly to Nelder-Mead Simplex but additionally determines if the size of the Simplex can be expanded in regions of the experimental space where little change in the objective is observed—hence leading to faster convergence. Even with the improved convergence rate, approximately 35 hr were required for the optimizations in the study. Improvements to the system were later made by Bourne *et al.*[85] (who included equivalents of methylating agent as a variable in the optimization), Jumbam *et al.*[86] (who demonstrated optimization over several objective functions including yield, space-time yield, and E-factor), and Skilton *et al.*[87] (who accelerated data collection with use of an FTIR and compared the performance of Simplex to SNOBFIT). Most recently the group demonstrated the potential of automated chemical reaction systems by allowing a researcher to remotely study and optimize the UK-based system from Brazil.[88]



Scheme 1.2. The methylation of primary alcohols by DMC in supercritical CO_2 studied by Parrott *et al.*[84]

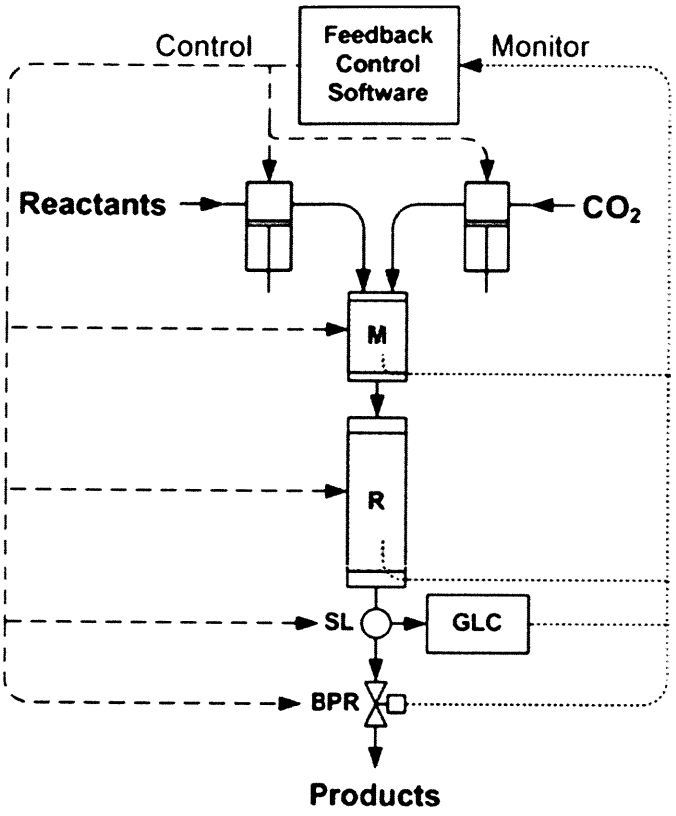
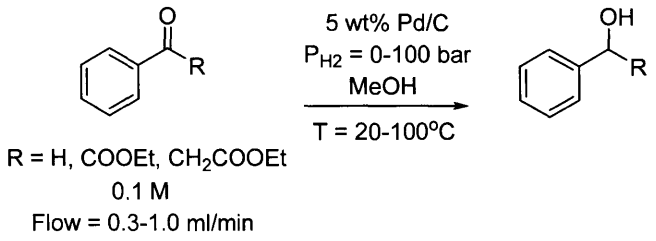


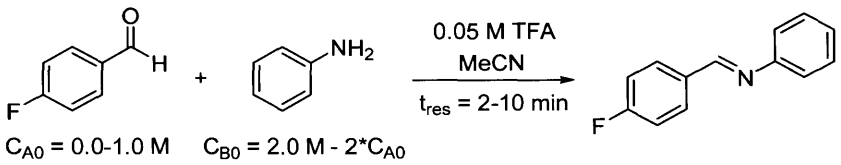
Figure 1.4. Automated system for optimization of the methylation of alcohols introduced by Parrott *et al.*[84] M is a static mixed and R is the reactor packed with catalyst.

Recent applications of automated optimization systems have demonstrated incorporation of more advanced synthesis and analytical systems into the Simplex algorithm framework. In increasing the scope of chemistries available for feedback control, the utility of smart optimization systems has increased substantially. As an example, Fabry *et al.*[90] coupled LabView software to a commercial flow hydrogenation system, the H-cube, and demonstrated optimization of Scheme 1.3 with respect to hydrogen pressure, temperature, and flow rate. Inline FTIR was used for analysis. The Simplex optimization required 24 hr to complete 17 experiments but identified optimal conditions for 99% conversion to the alcohol product. Feedback optimization with an inline NMR, introduced by Sans *et al.*,[61] offers the researcher an opportunity to study kinetics of intermediate formation and optimize chemistries that would be otherwise impossible to observe by optical or chromatographic techniques. In an application of Nelder-Mead Simplex, Sans *et al.* optimized imine formation in the reaction of 4-fluorobenzaldehyde and aniline (Scheme 1.4) by monitoring reaction yield and conversion with ^{1}H NMR. The optimization resulted in a reduction of reaction time to 2 min, while the reaction

yield was maintained above 70%. It is inevitable that to study more complex chemistries inline, multiple analytical instruments such as IRs, LCs, GCs, and/or NMRs will be incorporated into flow systems in series or in parallel to maximize the information gained on a per experiment basis.



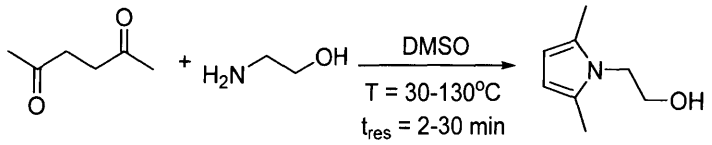
Scheme 1.3. Hydrogenation optimization studied by Fabry *et al.*[90]



Scheme 1.4. Optimization of imine formation by inline NMR from Sans *et al.*[61]

Despite the popularity of the method, the Simplex optimization routine can struggle to converge efficiently on account of the algorithm chosen for Simplex contraction, the dimensionality of the optimization, and the initial guess. Gradient-based optimization strategies tend to offer much faster convergence rates, at the expense of the extra experiments involved in gradient calculation. With a slow HPLC method, for instance, the time required for a gradient estimation can be quite limiting, but with the emergence of technology such as inline FTIR the use of gradient-based optimization methods should become more widely accepted. As an example of the speed at which feedback optimization can be accomplished with a gradient search, Moore and Jensen[58] demonstrated optimization of space-time yield for a Paul-Knorr reaction (Scheme 1.5) with three different gradient-based searches: steepest descent, conjugate gradient, and conjugate gradient with an Armijo step size. As illustrated in Figure 1.5, the rate of convergence effectively doubled when the conjugate gradient method was used compared to the steepest descent method, because of the additional gradient history incorporated into the selection of the conjugate gradient search direction. With a smarter selection of step size (following the Armijo rule[91]), the rate of convergence of the same optimization was accelerated to greater than three times the rate of convergence of the original steepest descent method. Though these methods are still black box local optimization searches, their abilities to rapidly

map out the trajectory to an optimum allow experimenters a quantitative interpretation of the response surface curvature that could be of service to kinetic investigations.



Scheme 1.5. Paul-Knorr reaction optimization studied by Moore and Jensen.[58]

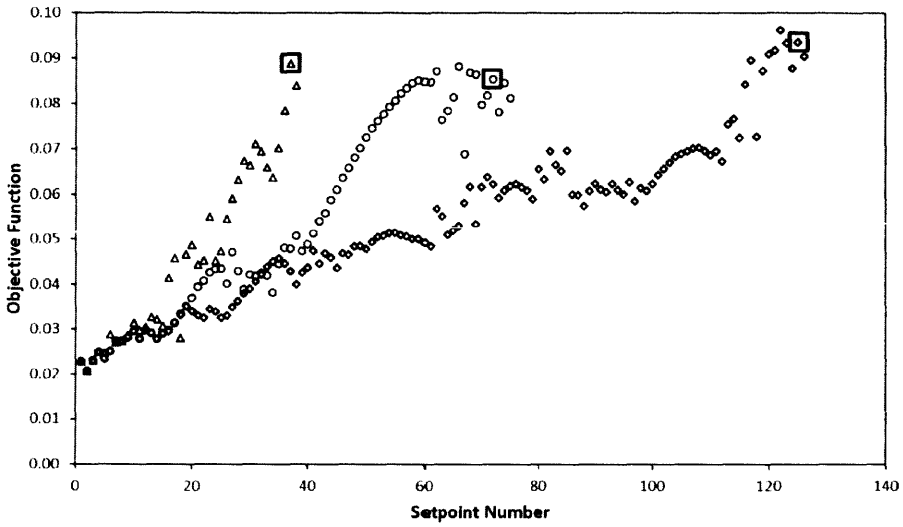


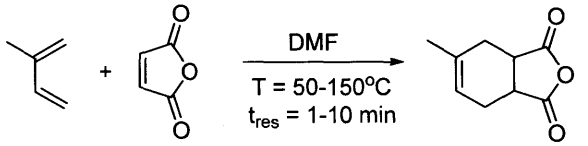
Figure 1.5. Convergence of the Paul-Knorr reaction from the automated system of Moore and Jensen.[58] Diamond—steepest descent algorithm. Circle—conjugate gradient algorithm with fixed step size. Triangle—conjugate gradient algorithm with Armijo step size.

1.3. KINETICS IN FLOW: A ROUTE TO FASTER SCALE-UP

Though black box strategies are valuable tools for identifying improved reaction conditions in less time than combinatorial or one-factor-at-a-time screening, predictable scalability of results can only come from a complete understanding of the reaction mechanism and kinetics. Identification of reaction kinetics in flow offers many of the advantages already discussed in previous sections in terms of fast heat and mass transfer rates and more precise control of reaction conditions. Additionally, with feedback an automated system can determine which kinetic experiments are most valuable to run in order to select optimal rate parameters and an optimal rate law. This offers an invaluable tool to the experimenter looking to discriminate among many possible reaction mechanisms.

As an example of using feedback in the determination of reaction kinetics, McMullen and Jensen[23] used a silicon microreactor flow system with online HPLC sampling to study the Diels-

Alder reaction of isoprene and maleic anhydride (Scheme 1.6). The experimenters initiated the algorithm by proposing four different kinetic rate laws—only one of which represented the true second order reaction rate of the system. Following the Shannon's entropy approach presented by Box and Hill,[92] the automated system selected a series of experiments that attempted to maximize the discrepancy between the predictions of each rate model. This allowed the system to discriminate between models based on the reaction's performance at each set of experimental conditions. After identification of the correct kinetic model, a D-optimal experimental design approach,[93,94] whereby experiments are chosen to minimize the collective uncertainty in fitted model parameters, was employed to estimate kinetic parameters for the model. The results were found to be in good agreement with literature and with transition-state theory calculations.



Scheme 1.6. Diels-Alder reaction used in the kinetic study by McMullen and Jensen.[23]

Demonstration of the scalability of the kinetic model required use of the Corning Advanced-Flow reactor system (Corning Incorporated). In this case, 9 glass reactors were used in series to produce a 500-fold scale-up from the silicon microreactor. It is important to note here that immediate scale-up of the microreactor optimum did not result in the same yield as observed during the optimization; this was on account of the losses of product yield to dispersion and less efficient heat removal in the scaled-up system. However, by accurately measuring the residence time distribution in the Advanced-Flow reactor and incorporating the heat of reaction into reactor temperature models along with the reaction kinetics, the optimal yield from the microreactor system was, in fact, found to scale predictably. The success of this scale-up emphasizes the superiority of model-based approaches over black box approaches when translating optimal operating conditions across scale.

1.4. Bringing Discrete Variables into the Optimization

Droplet-based flow systems have gained much interest recently as tools for continuous HTE.[95-106] By performing reactions, crystallizations, or biological assays in the confines of isolated droplets, reagent compositions can be controlled accurately, and the rates of mixing and heat transfer are consistent with batch reaction kinetics. Most importantly for optimization,

segmented-flow systems allow experimenters the opportunity to manipulate discrete variables such as catalysts or solvents in addition to the standard continuous variables of temperature, reaction time, and concentration, greatly expanding the design space for the synthesis.[97] The need for such systems in process scale-up has inspired the development and application of a commercial system[107-112] that can screen 20-900 μL volumes of reactions with different reagents, temperatures, and flow rates, though as was discussed earlier the problem of enumeration in any of these systems becomes limiting if many discrete variables and reaction conditions are incorporated into the screen without the use of feedback.

The incorporation of a global optimization algorithm into an automated, high-throughput catalyst screening process was demonstrated by Kreutz *et al.*[98] The feedback loop is shown in Figure 1.6. For the oxidation of methane by oxygen, multiple catalysts, cocatalysts, and ligands were screened in a segmented two-phase system, with the extent of each individual reaction assessed through the monitoring of an *in situ* colorimetric indicator. After 48 segmented flow experiments in quadruplicate for a given generation were completed, the "fitness" of each reaction (corresponding to the color change in an indicator solution) was supplied to a computer, which applied a genetic algorithm[113] to identify a new generation of catalyst, cocatalyst, and ligand combinations to be screened. After four generations of experiments, the genetic algorithm had sufficiently sampled the possible solution space, and the migration operator was deactivated in the algorithm in favor of attempting to select the fittest catalyst/cocatalyst/ligand combinations. The optimal active catalyst system was found after a total of eight generations to be Pt and POM-V2, in agreement with catalyst systems already reported in literature.

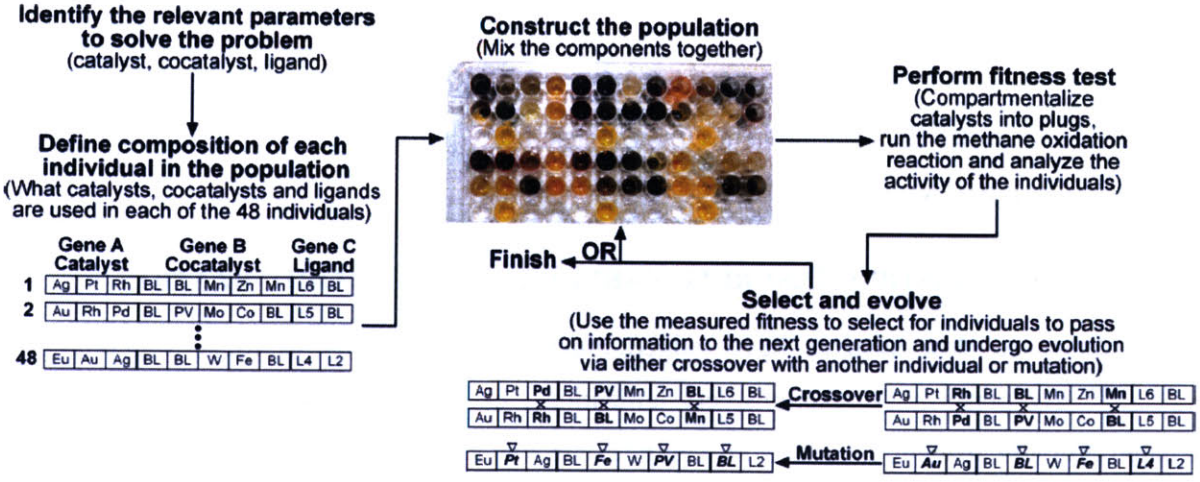


Figure 1.6. Feedback loop for the optimization of methane oxidation by Kreutz *et al.*[98]

The ability to reduce over 8 million catalyst/cocatalyst/ligand combinations to fewer than 400 experiments illustrates the tremendous time and cost savings that can be achieved with the application of feedback in discrete variable screening. The inclusion of continuous variables in the optimization has the potential to lead to even more improvements in efficiency, particularly if the performance of each catalyst differs at different temperatures, reaction times, or reagent concentration.

1.5. Thesis Overview and Goals

In a very short time, the field of continuous feedback optimization has grown from novelty to necessity. As syntheses of interest increase in complexity, the number of factors that can affect a response increase, and the level of complexity of the instrumentation needed to observe the response (or responses) increases accordingly. Along with that growth comes the requirement that smarter approaches continue to develop for solving more difficult and more integrated problems, and that technologies continue to develop in order to produce more rapidly scalable results.

This thesis begins in Chapter 2 by establishing an approach to evaluating reaction kinetics when the reaction pathway is more complex and an automated system alone cannot fully solve the kinetic model. Though a suitable approach is found by breaking the reaction pathway into smaller reaction steps and then applying the automated feedback approach, the question emerges as to how a smart system could be developed to increase the number of variables at the experimenter's disposal and hence greatly simplify the task of resolving a complex system. Specifically we ask how discrete variables can be brought into the optimization, and Chapter 3 is dedicated to the construction of a fully automated segmented flow system that allows for on-demand screening of both discrete and continuous variables. The optimization of both classes of variables simultaneously is a challenge within itself, with the introduction of a sequential adaptive response surface methodology aimed at solving this problem presented in Chapter 4. Chapters 5 and 6 then delve specifically into the application of the new method in the optimization of complex systems where both continuous variables and discrete variables strongly interact. In Chapter 5, the effect of solvent upon a selective alkylation reaction is explored. In Chapter 6, the method is applied to several case studies of Suzuki-Miyaura cross-coupling reactions in an effort to identify the best catalyst-ligand system and corresponding optimal

reaction conditions. Chapter 7 summarizes the significant contributions of this thesis and outlines future directions for feedback optimization in highly multivariate systems.

By the conclusion of this thesis, we aim to address the following goals:

- Expansion of the scope of automated kinetic parameter estimation
- Design and implementation of an automated segmented flow system capable of screening a diverse range of liquid-phase reactions
- Implementation of an algorithm that solves in real-time the mixed integer nonlinear programming (MINLP) of simultaneous discrete variable screening and continuous variable screening optimization
- Application of the screening system to simultaneous ligand, catalyst, and solvent selection and reaction optimization
- Demonstration of scale-up of optimized reaction conditions

In route to accomplishing these goals, we present a methodology that accelerates reaction development in a novel yet versatile way, transforming small amounts of material into information that can be used to optimize reactions and gain increased understanding of the underlying chemistry. We present new levels of understanding, not just of algorithms and chemistries, but of how to characterize reactions in a versatile, simple, and accurate manner. In continuation of research into smart feedback systems, it is these motives of versatility, simplicity, and accuracy that should be expected to drive innovation, helping foster greater continuity from reaction discovery all the way through drug manufacture.

2. AN AUTOMATED CONTINUOUS-FLOW PLATFORM FOR THE ESTIMATION OF MULTI-STEP REACTION KINETICS

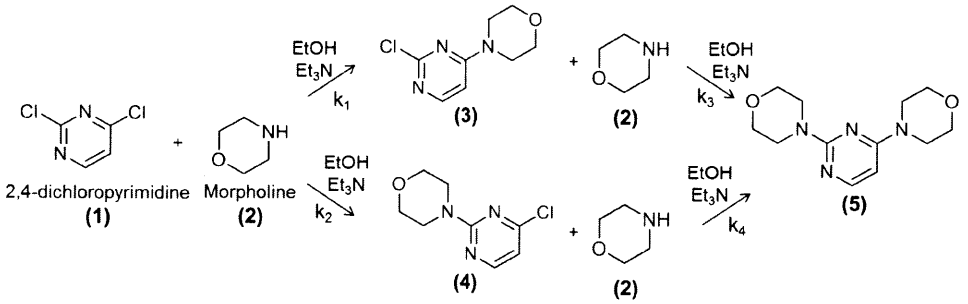
2.1. INTRODUCTION

A primary concern in pharmaceutical process chemistry is scale-up of a reaction from bench to production levels. Commonly, conditions found to be optimal on the bench scale end up non-optimal at a larger scale, due to changes in mass and heat transfer properties between reactor volumes. These changes in transport properties can lead to the formation of byproducts at the larger scale that were not accounted for in the preliminary optimization. In contrast to small-scale optimization followed by scale-up, it is traditionally preferred to model and parameterize a synthesis in terms of its kinetics at the small scale. The reaction kinetics are then coupled with knowledge of the effects of heat and mass transfer in the larger scale reactor in order to optimize the reaction.

As summarized in Chapter 1, our group has demonstrated the application of automated, continuous-flow microreactor systems for the purposes of reaction screening,[54] optimization,[22,51,58] kinetic model discrimination,[23] and kinetic parameter estimation.[23,29,31] Such systems offered the advantages of precise control of reaction conditions while minimizing both reagent consumption and user intervention. A challenge still to be considered in all of these cases was the extraction of reaction kinetic information from more complex reaction networks—those which proceed in either series or parallel and have the potential to form one or more unwanted byproducts.

In this study, we demonstrated use of an automated, continuous-flow system capable of both accurately estimating kinetic parameters for a series-parallel reaction network and optimizing the yield of a desired mono-substituted product. As is often the case in complex networks, however, we found that the kinetic parameters estimated by the automated system included a high degree of uncertainty—as great as 20% for some parameters—which had not been observed in automated studies of more simplified reaction networks.[23] To ensure scalability of our final results, we applied the same automated system to determine the kinetics of isolated steps in the reaction network, greatly reducing parameter uncertainties to more acceptable values of less than 4%. This yielded precise kinetic estimates for all steps in the reaction pathway, including those which accounted for only a small amount of byproduct formation.

The model reaction network studied was the nucleophilic aromatic substitution (S_NAr) reaction of 2,4-dichloropyrimidine (**1**) and morpholine (**2**) in ethanol to form a desired 2-substituted aminopyrimidine (**4**) and the less-desired 4-substituted (**3**) and 2,4-substituted (**5**) byproducts. The reaction network is shown in Scheme 2.1. 2- and 4-substituted aminopyrimidines have generated considerable pharmaceutical interest as inhibitors of kinases such as Cdks, p38, Aurora, KDR, and Gsk3.[114-117] The observed inhibitory effect has been attributed to hydrogen bonding interactions between the 1-nitrogen and the 2-amino group on the pyrimidine molecule and the hinge amino acid of the kinase.[118] Synthesis of 2-aminopyrimidines is complicated by the preference of the amine nucleophile to substitute at the 4-carbon position of the substrate.[119] More aggressive reaction conditions are generally required in order to promote the second nucleophilic substitution and thereby generate the 2-amino derivative.[120,121] Using a silicon microreactor for this synthesis, we were able to safely pressurize the flow system and carry out the reaction above the atmospheric boiling point of the solvent—a traditional limitation of batch experimentation. The rapid heat transfer rate of silicon additionally improved the likelihood of obtaining intrinsic reaction kinetics during experimentation without being limited by reaction exothermicity.



Scheme 2.1. Multi-step reaction network for conversion of 2,4-dichloropyrimidine to 4,4'-(2,4-pyrimidinediyl)bis-morpholine.

2.2. METHOD

A similar optimal experimental design procedure was followed to that used for kinetic estimation by McMullen and Jensen.[23] The procedure was iterative and is illustrated schematically in Figure 2.1. Experimentation began with an initial factorial design. Based upon data collected through online analysis, a regression-fitting algorithm optimized the values of parameters specified in a user-defined model in order to best agree with experimental data. Sensitivity coefficients were then calculated based upon the optimal parameter estimates for the

experiments performed and for each candidate posterior experiment to be tested for optimality. The sensitivity coefficients were stored in the Fischer information matrix, the determinant of which gave the objective function to be minimized in the selection of a D-optimal posterior experiment. The optimal experiment was subsequently identified, and the prior experimental data were augmented by the results of the D-optimal posterior experiment. This procedure iterated until the system was terminated by a user.

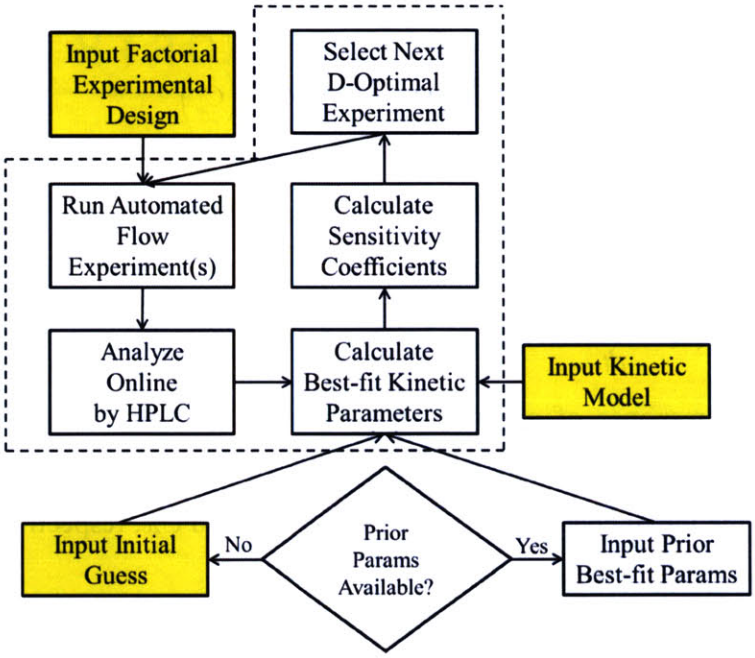


Figure 2.1. Logic flow diagram for automated kinetic parameter estimation in continuous flow.

2.2.1. Kinetic Model

We assumed for our kinetic model that all four reactions in Scheme 2.1 followed second order, bimolecular reaction kinetics and that the reaction system could be modeled as an ideal plug flow reactor (PFR). The assumption of second-order, bimolecular reaction kinetics agreed with the mechanisms previously established in literature for S_NAr reactions.[122-124] For 400 μm-reactor channels and liquid phase species diffusivities of greater than or equal to 1×10^{-9} m^2s^{-1}, the flow reactor used in these experiments could be modeled as an ideal plug flow reactor for residence times exceeding 2 min. Only small deviations from plug flow were expected for shorter reactor residence times extending down to 30 s.[27,125] We chose the minimum reaction time for our experiments to be 30 s so as not to have to de-convolute the effect of dispersion in our online kinetic analysis.

Considering our assumptions for the kinetics of the reactions in Scheme 2.1, we proposed the following rate laws governing species generation and consumption:

$$f_1 = \frac{d\left(\hat{C}_u\right)_1}{dt} = -k_1\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 - k_2\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 \tag{2.1}$$

$$f_2 = \frac{d\left(\hat{C}_u\right)_2}{dt} = -k_1\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 - k_2\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 - k_3\left(\hat{C}_u\right)_3\left(\hat{C}_u\right)_2 - k_4\left(\hat{C}_u\right)_4\left(\hat{C}_u\right)_2 \tag{2.2}$$

$$f_3 = \frac{d\left(\hat{C}_u\right)_3}{dt} = k_1\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 - k_3\left(\hat{C}_u\right)_3\left(\hat{C}_u\right)_2 \tag{2.3}$$

$$f_4 = \frac{d\left(\hat{C}_u\right)_4}{dt} = k_2\left(\hat{C}_u\right)_1\left(\hat{C}_u\right)_2 - k_4\left(\hat{C}_u\right)_4\left(\hat{C}_u\right)_2 \tag{2.4}$$

$$f_5 = \frac{d\left(\hat{C}_u\right)_5}{dt} = k_3\left(\hat{C}_u\right)_3\left(\hat{C}_u\right)_2 + k_4\left(\hat{C}_u\right)_4\left(\hat{C}_u\right)_2 \tag{2.5}$$

The model-predicted response of species i was specified in Equations 2.1-2.5 as $(\hat{C}u)_i$, the predicted concentration of i in experiment u as a function of the reaction time t, the reaction temperature T, and the initial concentrations of **1** and **2**, C_{10} and C_{20}, respectively. k_r was the rate constant for reaction r, expressed as:

$$k_r = A_r \exp\left(-\frac{E_{Ar}}{RT}\right) \quad r = 1,...,4 \tag{2.6}$$

where A_r and E_{Ar} were the pre-exponential factor and activation energy associated with k_r, respectively, and R was the gas constant. To achieve better convergence to an optimal set of kinetic parameters, we defined scaled parameters θ_p such that:

$$\theta_{p=2r-1} = \ln\left(A_r\right) \tag{2.7}$$

$$\theta_{p=2r} = \frac{E_{Ar}}{RT^*} \tag{2.8}$$

Equation 2.6 was then rewritten as:

$$k_r = \exp\left(\theta_{2r-1} - \frac{T^*}{T}\theta_{2r}\right) \tag{2.9}$$

where the pre-exponential factor and the activation energy were scaled comparably. We chose for our case T^* = 343 K to represent an average value for the reaction temperature in our experiments, ensuring that $T^*/T \sim 1$ over the range of reaction temperatures studied.

2.2.2. Approach to Parameter Estimation

An optimal set of kinetic parameters was obtained by fitting the kinetic model of Equations 2.1-2.5 to the calibrated HPLC responses of **1**, **3**, **4**, and **5** as functions of T, C_{10}, C_{20}, and the residence time t_{res}. Both maximum likelihood estimation (MLE) and maximum *a posteriori* (MAP) estimation were applied in obtaining optimal least squares regression estimates of the kinetic parameters.

For experiments in which no prior estimates for optimal kinetic parameters were available, a set of optimal kinetic parameters was found by MLE. The non-linear programming for MLE was formulated as:

$$\min_{\boldsymbol{\theta}} \sum_{u=1}^{N_{expts}} \left[\mathbf{C}_u - \hat{\mathbf{C}}_u(\boldsymbol{\theta})\right]' \mathbf{W}_u \left[\mathbf{C}_u - \hat{\mathbf{C}}_u(\boldsymbol{\theta})\right] \tag{2.10}$$

where $\mathbf{C}_u$ was the $N_{resp} \times 1$ vector of measured responses for experiment u and $\hat{\mathbf{C}}_u(\boldsymbol{\theta})$ was the $N_{resp} \times 1$ vector of model-predicted responses for experiment u with model parameters $\boldsymbol{\theta}$. $\mathbf{W}_u$ was a weighting matrix for the residuals which we chose to be:

$$\mathbf{W}_u = \mathbf{V}_B^{-1} \quad u = 1, \ldots, N_{expts} \tag{2.11}$$

$\mathbf{V}_B$ was the response-covariance matrix, defined for species i and species j as:[126]

$$(\mathbf{V}_B)_{ij} = s_{ij}^2 = \sum_{u=1}^{N_{expts}} \frac{\left[(\mathbf{C}_u)_i - (\hat{\mathbf{C}}_u)_i\right]' \left[(\mathbf{C}_u)_j - (\hat{\mathbf{C}}_u)_j\right]}{N_{expts} - N_{params}} \tag{2.12}$$

where the difference in the number of experiments and the number of optimized parameters, $N_{expts} - N_{params}$, was strictly greater than zero. Because the objective function required $\mathbf{V}_B$ as an input, we used $\mathbf{V}_B$ from the previous experiment as an input to the updated MLE optimization. $\mathbf{V}_B$ was initialized as the identity matrix prior to the first parameter optimization and was found experimentally to converge to a consistent set of values after only 1 to 2 posterior experiments. The optimization in Equation 2.10 was performed as a constrained sequential quadratic programming (SQP) optimization in MATLAB. The lower and upper bounds on the optimization

were found by computing the 98% confidence intervals on the prior optimal parameters. In order to limit online computational time, a maximum of 500 SQP iterations was allowed for the optimization.

Uncertainties in parameter values were evaluated based upon the parameter covariance matrix $\mathbf{V}_\theta$ following the treatment of Beck and Arnold.[126] To calculate $\mathbf{V}_\theta$, we first defined the sensitivity coefficient, $(\mathbf{X}_u)_{ip}$, for response i with respect to parameter θ_p in experiment u:

$$(\mathbf{X}_u)_{ip} = \left.\frac{\partial\left(\hat{\mathbf{C}}_u\right)_i}{\partial\theta_p}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{opt}} \tag{2.13}$$

Here $\boldsymbol{\theta}_{opt}$ denoted the optimal set of MLE or MAP parameters found by SQP. Given the kinetic rate laws f_i in Equations 2.1-2.5, we were able to analytically evaluate Equation 2.13 in the form of an ordinary differential equation:[127]

$$\frac{d(\mathbf{X}_u)_{ip}}{dt} = \frac{\partial f_i}{\partial\theta_p} + \sum_{k=1}^{N_{resp}} \frac{\partial f_i}{\partial\hat{C}_k}(\mathbf{X}_u)_{kp} \tag{2.14}$$

From the matrix of sensitivity coefficients, we calculated the Fisher information matrix, $\mathbf{Z}$, which equaled the inverse of the parameter covariance matrix for the case of MLE:

$$\mathbf{Z} = \sum_{u=1}^{N_{expts}} \mathbf{X}_u' \mathbf{V}_B^{-1} \mathbf{X}_u \tag{2.15}$$

$$\mathbf{Z} = \mathbf{V}_\theta^{-1} \text{ for MLE} \tag{2.16}$$

Estimation of the parameter covariance matrix allowed for the uncertainties of our kinetic parameter estimates to be calculated. For a single parameter θ_p, a 1-dimensional confidence interval was calculated from the expression:[128]

$$\left(\boldsymbol{\theta}_{opt}\right)_p - \left[\left(\mathbf{V}_\theta\right)_{pp}\right]^{1/2} t_{\alpha/2,\nu=N_{expts}-N_{params}} \le \theta_p \le \left(\boldsymbol{\theta}_{opt}\right)_p + \left[\left(\mathbf{V}_\theta\right)_{pp}\right]^{1/2} t_{\alpha/2,\nu=N_{expts}-N_{params}} \tag{2.17}$$

where $t_{\alpha/2,\nu=N_{expts}-N_{param}}$ was the Student's t-distribution value for $\alpha/2$ confidence and N_{expts} - N_{params} degrees of freedom.

The approach to MAP estimation was similar to the approach for MLE, though MAP estimation considered *a priori* estimates and uncertainties for the vector of model parameters $\boldsymbol{\theta}$. The quadratic program for MAP estimation was given as:

$$\min_{\theta}\left[\boldsymbol{\mu}-\boldsymbol{\theta}\right]^{T}\mathbf{V}_{\mu}^{-1}\left[\boldsymbol{\mu}-\boldsymbol{\theta}\right]+\sum_{u=1}^{N_{expts}}\left[\mathbf{C}_{u}-\hat{\mathbf{C}}_{u}(\boldsymbol{\theta})\right]'\mathbf{V}_{B}^{-1}\left[\mathbf{C}_{u}-\hat{\mathbf{C}}_{u}(\boldsymbol{\theta})\right] \tag{2.18}$$

$\boldsymbol{\mu}$ was the *a priori* vector of optimal model parameters and $\mathbf{V}_{\mu}^{-1}$ was a weighting matrix for the difference between the prior model parameters and the optimal *a posteriori* model parameters. $\mathbf{V}_{\mu}$ was identified as the *a priori* parameter covariance matrix, which could be calculated as in Equations 2.15 and 2.16 for all prior experiments. The *a posteriori* parameter covariance matrix was then given by:

$$\mathbf{V}_{MAP}^{-1}=\mathbf{V}_{\mu}^{-1}+\mathbf{V}_{\theta}^{-1} \tag{2.19}$$

and was substituted into Equation 2.17 for $\mathbf{V}_{\theta}^{-1}$ in order to obtain posterior confidence intervals. The number of degrees of freedom in this case was N_{prior} + N_{expts} - N_{params}, where N_{prior} was the number of prior experiments already conducted. As in MLE, the MAP optimization was evaluated in MATLAB with the constrained SQP optimization algorithm and limited to a maximum of 500 iterations.

2.2.3. Approach to Optimal Experimental Design

Our objective in parameter estimation was to minimize the total uncertainty and joint-uncertainty in the MLE and MAP optimal parameters. Experiments were selected based upon the D-optimality criterion, introduced by Box and coworkers:[93,94]

$$D=\min_{\mathbf{u}}\left|\left(\mathbf{V}_{\theta}^{-1}+\left(\mathbf{X}_{u}\right)'\mathbf{V}_{B}^{-1}\mathbf{X}_{u}\right)^{-1}\right| \tag{2.20}$$

s.t. $\mathbf{u}\in$ Experimental design space

The optimal D corresponded to the choice of conditions for the next experiment in the design of experiments for which the predicted volume of the parameter covariance matrix was minimized. By minimizing the volume of the parameter covariance matrix, the total joint uncertainty among all parameters in the model was minimized. For MLE, we incorporated Equation 2.20, as written, into the parameter estimation program. For MAP estimation, we substituted $\mathbf{V}_{MAP}$ for $\mathbf{V}_{\theta}$ as was done for estimation of parameter confidence intervals.

2.3. EXPERIMENTAL

2.3.1. Automated Parameter Estimation System

A diagram of the automated parameter estimation system is shown in Figure 2.2. A 0.30 M solution of **1** (98%, Sigma-Aldrich) was delivered with an internal standard, 1,2-dimethoxybenzene (> 99%, TCI) in ethanol to a silicon microreactor. A solution of 0.92 M **2** (≥ 99.0%, Sigma-Aldrich) and 0.92 M triethylamine (Et_3N, ≥ 99.0%, TCI) in ethanol was delivered to a T-junction, diluted with ethanol, and delivered to the second inlet port of the microreactor. The reaction product was quenched in the quenching zone of the microreactor by a 1.6 M solution of trifluoroacetic acid (TFA, 99%, Sigma-Aldrich) in ethanol. This quenched product was then further diluted to a 3:5 ratio in a micromixer by a second stream of ethanol and injected into an HPLC for online analysis. In a separate set of experiments, solutions of **3** and **4** were each reacted with a solution of 0.36 M **2** and 0.36 M Et_3N to produce **5**. In the case of using **3** or **4** as a starting material, a 0.16 M solution of **3** dissolved with the internal standard in ethanol or a 0.08 M solution of **4** dissolved with the internal standard in ethanol, respectively, was substituted into the system in place of **1**.

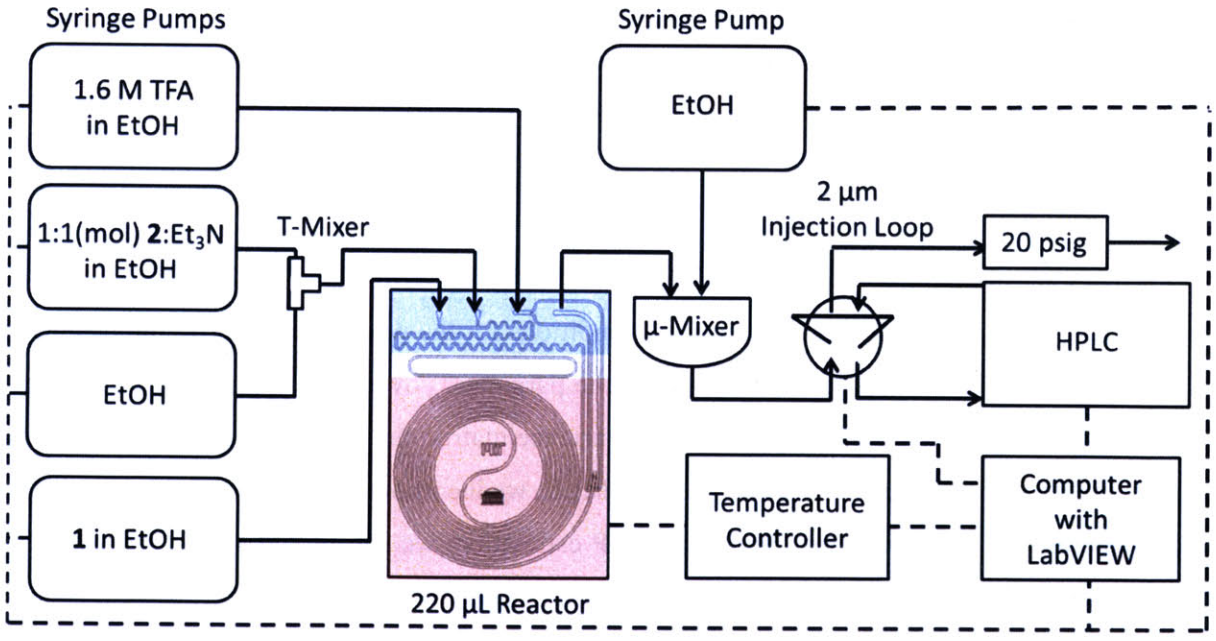


Figure 2.2. Diagram of the automated continuous-flow parameter estimation system

Pumping of fluids through the microreactor and micromixer system was accomplished using PhD 2200 syringe pumps (Harvard Apparatus) under the control of LabVIEW v8.6. Manipulation of pump flow rates allowed for a range of residence times and reactant concentrations to be explored. Connections downstream of the microreactor were made using 218 μm (0.0086 in.) internal diameter PFA tubing (Upchurch Scientific, IDEX Health & Science LLC) to minimize dead volume. Check valves were installed on all feed streams and a 1.4 bar (20 psi) backpressure regulator was installed downstream of the HPLC injection valve in order to dampen flow oscillations and to increase the boiling temperature of the primary solvent, ethanol.

The silicon microreactor employed in this experiment was fabricated following standard photolithography and deep reactive ion etching techniques.[6] The channel cross-sectional dimensions were 500 μm (width) x 400 μm (height). A halo etch of the microreactor enabled temperature control in two different zones of the reactor. At the entrance of reactants to the reactor, a 20-μL mixing zone allowed for mixing of both reactant streams at ambient temperature. The 220-μL spiral reaction zone of the reactor was then heated to a uniform temperature by a cartridge heater controlled by an OMEGA temperature controller (OMEGA Engineering Inc.) to ± 0.4°C. The silicon micromixer design has been described previously[129] and allows for rapid mixing or dilution in a 4.1-μL volume. Both the microreactor and the micromixer were compression packaged to enable continuous fluid transfer throughout the system. The total volume of the system, including the microreactor, micromixer, and transfer tubing downstream of the microreactor was approximately 280 μL. To allow adequate time for the system to reach a steady state after equilibration of the reactor temperature, 1 mL of reactants were infused into the system prior to online analysis. To allow sufficient time for the syringe pumps to equilibrate, the system was additionally required to have run for a minimum of 3 min at the same temperature and flow rates prior to online analysis.

Based on the system design, constraints were placed *a priori* upon the experimental design space. Residence times were constrained to within the range of accuracy for the syringe pumps exerting force on 5 μL glass syringes dispensing through the system under 1.4 bar backpressure. We estimated this range to be from flow rates of approximately 1 μL min^{-1} to 250 μL min^{-1} for each syringe pump. We reasoned the minimum temperature to be that at which the rate of reaction in the mixing zone of the microreactor (held at room temperature) was insignificant in comparison to the reaction rate in the reaction zone. Under this condition, the complications of

the reaction mechanism in the mixing zone could be excluded from the kinetic model. A maximum temperature of 100°C was specified so as not to exceed the boiling point of ethanol in the presence of 1.4 bar of backpressure. In future studies, a change to a higher boiling point solvent such as *n*-butanol would enable a more extensive range of temperatures to be explored. The initial concentration of **1** was limited by the solubility of **1** in ethanol, and the range of equivalents **2** and Et_3N added was chosen to extend from 0.5 to 5.0. The number of discretizations of the experimental space (1600) was chosen to achieve an extensive range of internal points at which to evaluate the D-optimality condition, while at the same time limiting the time required online to exhaust all possible combinations of the four input variables.

Analyte concentrations were measured online by HPLC. Analysis by HPLC was advantageous in allowing for a quantitative separation of reaction components and demonstrated the potential for this method to be applied to more complex reaction networks. Species were measured using a Waters HPLC (Waters Corporation) with 1525 binary pumps; a Nova-Pak C18 4μm, 3.9 x 150 mm column; a 2996 PDA detector; and Empower software. A 2-μL volume of diluted reaction product was automatically injected into the HPLC for analysis. A gradient method of water and acetonitrile was employed in order to separate **1**, **3**, **4**, **5**, and the internal standard. MATLAB code was written to integrate peak areas and determine species concentrations based upon previous calibrations with the internal standard. Analysis was completed in 9.25 min, at which point either reaction conditions were manipulated in preparation for collecting the next experimental data point or data were passed to the parameter estimation program in MATLAB in order to identify the next D-optimal experiment to conduct.

2.3.2. Experimental Design

The design of experiments proceeded in three stages: an initial set of experiments aimed at determining the eight parameters in Equations 2.1-2.6 simultaneously; a second set of experiments aimed at parameterizing each step of the reaction pathway in isolation; and a final set of experiments aimed at determining the eight kinetic parameters simultaneously using *a priori* estimates from the prior sets of experiments. Up to four factors were manipulated for each experiment: the reaction residence time (t_{res}), the reaction temperature (T), the initial substrate concentration (C_{i0}), and the equivalents of **2** fed to the reactor. Each set of experiments began with an initial factorial design, which characterized the effect of manipulating multiple factors

upon product yield. Following the initial factorial design, experiments were chosen sequentially from the solution of Equation 2.20 over 1600 candidate experimental points found from enumeration of:

$$\begin{aligned} t_{res} &\in \{0.5 \text{ min}, 1 \text{ min}, 1.5 \text{ min}, ..., 20 \text{ min}\} \\ T &\in \{40°C, 60°C, 80°C, 100°C\} \\ C_{10} &\in \{0.075 \text{ M}, 0.150 \text{ M}\} \text{ or } C_{30} \in \{0.025 \text{ M}, 0.050 \text{ M}\} \text{ or } C_{40} \in \{0.015 \text{ M}, 0.030 \text{ M}\} \\ \text{Equiv. } \mathbf{2} &\in \{0.5,\ 1.0, 1.5, 2.0, 2.5\} \end{aligned} \tag{2.21}$$

A complete list of experimental conditions tested can be found in Appendix A.

The program initiated with a set of factorial experiments designed to estimate all kinetic parameters in Equations 2.1-2.6 simultaneously. The initial number of experiments needed before an initial guess of the reaction kinetics could be calculated had to exceed eight, the number of parameters in the model. To lessen parameter uncertainty before the selection of the first D-optimal experiment, a 12-experiment factorial design was chosen which sampled from several different regions of the experimental design space. The factorial design included all combinations of 3 different residence times (30 s, 60 s, and 5 min), 2 different reaction temperatures (40°C and 80°C), 2 different initial concentrations of morpholine (0.15 M and 0.30 M), and an initial concentration of **1** of 0.15 M. Experiments were sequenced randomly, and no replicates were specified initially.

The MLE estimates for the kinetic parameters after the initial factorial design were found from initial guesses of $\log_{10}(A_r / \text{M}^{-1}\text{s}^{-1}) = 0$ and $E_{Ar} = 14.2$ kJ mol^{-1} for all four rate constants. To initiate the method, $\mathbf{V}_B$ was chosen to be the identity matrix (assuming that measurement uncertainty was the same for the starting material and all three products), and lower and upper bounds on the pre-exponential factors and activation energies were specified as:

$$-10.9 \le \log_{10}\left(A_r / \text{M}^{-1}\text{s}^{-1}\right) \le 10.9 \tag{2.22}$$

$$0 \text{ kJ mol}^{-1} \le E_{Ar} \le 285 \text{ kJ mol}^{-1} \tag{2.23}$$

Setting feasible upper and lower bounds which did not extend toward $\pm\infty$ improved the speed of convergence for the online estimation and allowed for a more accurate fit of the rate parameters to be achieved in a limited number of MLE iterations.

Following simultaneous parameter estimation, the kinetics of isolated reactions of **1** and **2**, **3** and **2**, and **4** and **2** were assessed. For the reaction of **1** and **2**, a 2×2×2 factorial design was selected which, based upon the prior parameter estimates obtained from estimating all

parameters simultaneously, was expected to most greatly minimize uncertainty in the estimates of A_1, E_{A1}, A_2, and E_{A2}. 0.5 min and 1.0 min residence times, 40°C and 100°C reaction temperatures, and initial concentrations of **2** of 0.150 M and 0.300 M were studied along with an initial concentration of **1** of 0.150 M. The initial guess for the MAP estimation corresponded to the prior optimal parameter estimates for A_1, E_{A1}, A_2, and E_{A2}. $\mathbf{V}_B$ was supplied as the final response-covariance matrix obtained from the simultaneous estimation of all kinetic parameters.

The isolated product **3** was reacted with **2** and Et_3N to produce **5**. Following from the kinetic model, we expected this reaction to obey second order kinetics and to be governed by the parameters A_3 and E_{A3}. As only two parameters were to be estimated, we limited the initial factorial design to a 2×2 set of four experiments at 80°C and 100°C with 10 min and 20 min residence times, an initial concentration of **3** of 0.050 M, and 2.5 equivalents of **2** and Et_3N. This factorial design was chosen to maximize the sensitivity coefficients for the final concentration of **5** with respect to the parameters A_3 and E_{A3}. These sensitivity coefficients were observed to increase with increasing reaction time, temperature, and starting material concentrations. After the initial factorial design, the MAP parameter estimation procedure was initiated with the prior estimates and uncertainties listed for A_3 and E_{A3} in Table 1. $\mathbf{V}_B$ was specified as the matrix of covariances corresponding to the measurements of **3** and **5** in the simultaneous parameter estimation experiment. The set of allowable initial concentrations for **3** was limited in the combinatorial optimization to 0.025 M and 0.050 M in order to conserve reagents. The allowable initial concentrations of **2** were scaled to range from 0.025 M to 0.125 M in 0.025 M increments. The allowable reaction temperatures and residence times were the same as given for the simultaneous experiment.

To determine the kinetic parameters for k_4 in isolation, **4** was reacted with **2** in the presence of Et_3N to produce **5**. Because less than 0.15 g of **4** were synthesized and purified, we limited the initial concentration of **4** in all experiments to 0.030 M. All other inputs to the initial 2×2 factorial design were the same as for Scheme 3: temperatures of 80°C and 100°C, residence times of 10 min and 20 min, and 2.5 molar equivalents of **2** and Et_3N. For the eight D-optimal experiments, the D-optimal design program was allowed to optimize with initial concentrations of **4** of 0.015 M and 0.030 M and initial concentrations of **2** and Et_3N ranging from 0.015 M to 0.075 M in increments of 0.015 M. As in the estimation of A_3 and E_{A3}, the *a priori* estimates and

uncertainties for A_4 and E_{A4} were specified as the results of the simultaneous parameter estimation experiment found in Table 1.

In a final simultaneous experiment using the optimal parameters from the set of isolated experiments, a 12-experiment factorial design was initially conducted with reaction times of 0.5 min, 10 min, and 20 min; temperatures of 40°C and 100°C; an initial concentration of **1** of 0.150 M; and 1.0 and 2.5 equivalents of **2** and Et_3N. This experimental design was then supplemented with six D-optimal experiments. $\mathbf{V}_B$ was specified as the final response covariance matrix from the simultaneous parameter estimation experiments. To ensure that $\mathbf{V}_\mu$ would be positive-definite, the blocks along the diagonal of the initial $\mathbf{V}_\mu$ were specified as the optimal parameter covariance matrices from the prior isolated experiments, with all other terms in $\mathbf{V}_\mu$ set equal to 0.

2.3.3. Synthesis and Isolation of Products

Products **3** and **4** were synthesized in the automated system described and shown in Figure 2.2. 0.150 M **1** (corresponding to 2.0 g starting material) and 2.5 equivalents of **2** and Et_3N were reacted in ethanol for 49 s at 100°C. The reaction product was quenched online by TFA. An aqueous extraction was performed offline to remove any salts formed from the quench, and the organic product was dried in Na_2SO_4. After filtration, liquid solvents were removed under vacuum to yield a white, crystalline product. This product was separated by dry loading onto 50 g of silica gel and eluting with a 4:1 solution of hexane:ethyl acetate to yield **4** in > 95% purity and eluting with a 1:2 solution of hexane:ethyl acetate to yield **3** in > 95% purity. HPLC yields of compounds **3** and **4** were 79.5% and 16.1%, respectively. Following workup, the isolated yield of **3** was 69.9% and the isolated yield of **4** was 14.1%, based upon the moles of **1** reacted. The isolated compound **3** was confirmed by HPLC, IR, ^1H and ^{13}C NMR, and GCMS. **4** was confirmed by HPLC, IR, ^1H NMR, and GCMS. Though ^1H NMR and GCMS cannot be used to distinguish the structures of **3** and **4**, we inferred from literature[19] that **3** was the compound produced in the greatest quantity and the compound that gave greater selectivity at low temperatures than at higher temperatures. **4** was also identified by HPLC as being notably less polar than **3**, which is consistent with the symmetric positioning of the electron-donating nitrogen atoms in the structure of **4**.

4-(2-chloro-4-pyrimidinyl)-morpholine (3): ^{1}H NMR(400 MHz, $CDCl_3$) δ 8.08 (1 H, d, J=6.0 Hz), 6.38 (1 H, d, J=6.0 Hz), 3.78 (4 H, t, J = 5.0Hz), 3.65 (4 H, broad); ^{13}C NMR(100 MHz, $CDCl_3$) δ 162.8, 159.7, 155.9, 101.3, 66.5, 44.6; IR ν_{max} 2361, 1653, 1586, 1559, 1540, 1355, 1265, 1234, 1165, 1117, 979, 801 cm^{-1}; GCMS m/z 52.0, 79.0, 114.0, 142.0, 167.9, 199.0; HPLC elution time 275-280 s, λ_{meas} = 328 nm.

4-(4-chloro-2-pyrimidinyl)-morpholine (4): ^{1}H NMR(400 MHz, $CDCl_3$) δ 8.27 (1 H, d, J=5.5 Hz), 6.64 (1 H, d, J=5.0 Hz), 3.90-3.75 (8 H, m); IR ν_{max} 2865, 2361, 2341, 1617, 1580, 1506, 1448, 1336, 1269, 1202, 1159, 1116, 983, 962, 780 cm^{-1}; GCMS m/z 51.9, 78.9, 113.9, 141.9, 167.9, 199.0; HPLC elution time 405-410 s, λ_{meas} = 284 nm.

Product **5** was synthesized neat in batch in an effort to achieve a high yield in a short period of time. Such an approach was considered acceptable over a flow chemistry approach for two primary reasons. First, as the proposed kinetic mechanism suggested, the yield of **5** was maximized for arbitrarily high species concentrations and infinitely long residence times. In a case such as this, the precise control of reaction conditions afforded by a regulated microreactor system was of minimal benefit unless the reaction presented a concern to safety under uncontrolled conditions. Secondly, as the production of **5** under neat conditions progressed, a viscous slurry of reactants and products developed which would have been difficult to transport in our flow system without risking unsafe pressure accumulation and/or clogging of the microchannel.

Our batch synthesis of **5** began with 2.0 g of **1** reacted with 2 mL morpholine and 3 mL Et_3N and reacted to yield 3.3 g of **5** at > 99% purity after 48 hr at room temperature. The isolated product was confirmed by HPLC, IR, ^{1}H and ^{13}C NMR, and GCMS.

4,4'-(2,4-pyrimidinediyl)bis-morpholine (5): ^{1}H NMR(300 MHz, $CDCl_3$) δ 7.96 (1 H, d, J=6.0 Hz), 5.86 (1 H, d, J=6.0 Hz), 3.80-3.70 (12 H, m), 3.54 (4 H, t, J=5.0 Hz); ^{13}C NMR(75 MHz, $CDCl_3$) δ 162.8, 161.8, 157.0, 93.3, 67.1, 66.8, 44.5, 44.3; IR ν_{max} 2852, 2361, 2341, 1582, 1558, 1472, 1438, 1263, 1237, 1001 cm^{-1}; GCMS m/z 67.1, 106.9, 134.9, 161.9, 192.8, 218.9, 249.8; HPLC elution time 180-210 s, λ_{meas} = 284 nm.

2.3.4. Automated Calibration of Analyzed Compounds

To conserve materials and minimize the amount of manual work invested in experimental preparation, HPLC calibration curves for **1**, **3**, **4**, and **5** in relation to the internal standard were developed in an automated procedure. In the case of each calibration, a sample of the isolated reactant or product was dissolved with the internal standard into a 10-mL solution of ethanol. Each solution was then set up to be delivered via syringe pump to the micromixer, where it was to be diluted with pure ethanol and injected online into the HPLC. The automated system proceeded by manipulating flow rates of both the analyte and ethanol streams so as to generate a correlation between the HPLC absorbance signal of the reactant or product and the absorbance of the internal standard at various concentrations. We found this procedure to be effective in eliminating the effect of flow rate oscillations in the final calibration curves.

2.4. RESULTS

2.4.1. Simultaneous Estimation of Kinetic Parameters

Table 2.1 lists the best-fit parameter estimates found by MLE at the conclusion of the initial simultaneous parameter estimation factorial design. Along with each parameter estimate, the calculated uncertainty is presented as ±1 standard deviation. While estimates on the parameters relating to k_1 and k_2 showed reasonable precision, the infinite uncertainties in the parameters for k_3 and k_4 implied that very little information on these parameters had been gathered from the results of the initial experimental design. Figure 2.3(a-d) shows the fit of the factorial design data by the initial rate parameter estimates. That the fit passes visual inspection is a testament more to the accuracy of temperature and residence time control in the microreactor system than it is the choice of kinetic parameters. As the standard errors in Table 2.1 indicate, an extensive range of parameters could have been found for k_3 and k_4 which would have acceptably fit the data shown in Figure 2.3.

Table 2.1. Optimal kinetic parameter estimates and uncertainties* from simultaneous estimation approach.

Number of Experiments	$\log_{10}(A_1)$	E_{A1}	$\log_{10}(A_2)$	E_{A2}	$\log_{10}(A_3)$	E_{A3}	$\log_{10}(A_4)$	E_{A4}
Initial	0.0	14.2	0.0	14.2	0.0	14.2	0.0	14.2
12	3.4 ± 0.6	26.6 ± 3.2	3.3 ± 0.6	31.0 ± 3.8	-6 ± Inf†	52 ± Inf†	-2 ± Inf†	27 ± Inf†
13	3.3 ± 0.5	26.4 ± 2.7	3.3 ± 0.5	31.1 ± 2.9	11 ± 6	102 ± 43	11 ± 29	100 ± 210
14	3.2 ± 0.4	25.5 ± 2.5	3.1 ± 0.4	30.1 ± 2.8	6.2 ± 1.2	67.9 ± 8.4	3.0 ± 3.0	44 ± 22
15	3.3 ± 0.4	26.2 ± 2.4	3.2 ± 0.4	30.8 ± 2.6	6.2 ± 1.1	68.3 ± 8.2	2.1 ± 2.1	37 ± 15
16	3.3 ± 0.4	26.1 ± 2.2	3.2 ± 0.4	30.7 ± 2.4	5.8 ± 0.8	65.4 ± 5.8	2.5 ± 1.7	40 ± 12
17	3.3 ± 0.4	26.1 ± 2.1	3.2 ± 0.4	30.6 ± 2.3	5.8 ± 0.8	65.5 ± 5.5	2.4 ± 1.5	39 ± 10
18	3.3 ± 0.3	26.3 ± 1.7	3.3 ± 0.3	31.1 ± 1.9	5.9 ± 0.7	66.5 ± 5.5	2.3 ± 1.4	38.5 ± 9.6
19	3.3 ± 0.3	26.2 ± 1.6	3.2 ± 0.3	30.8 ± 1.8	6.0 ± 0.7	67.2 ± 4.9	1.8 ± 1.0	35.1 ± 6.9
20	3.3 ± 0.3	26.4 ± 1.7	3.3 ± 0.3	31.2 ± 1.9	5.8 ± 0.7	65.1 ± 5.0	2.4 ± 1.0	39.2 ± 7.0
21	3.3 ± 0.3	26.3 ± 1.7	3.3 ± 0.3	31.1 ± 1.9	5.8 ± 0.6	65.8 ± 4.0	2.4 ± 0.9	39.1 ± 6.1
22	3.3 ± 0.3	26.2 ± 1.7	3.3 ± 0.3	31.0 ± 1.9	6.0 ± 0.6	66.7 ± 4.0	2.2 ± 0.8	37.4 ± 5.8
23	3.3 ± 0.2	26.5 ± 1.5	3.3 ± 0.3	31.4 ± 1.7	6.1 ± 0.6	67.4 ± 3.9	2.1 ± 0.7	36.8 ± 5.3
24	3.3 ± 0.2	26.5 ± 1.5	3.3 ± 0.2	31.4 ± 1.6	6.2 ± 0.6	68.3 ± 4.1	2.0 ± 0.8	36.0 ± 5.3

*-Uncertainties given as ±1 standard deviation. A_r is in $M^{-1}s^{-1}$ and E_{Ar} is in kJ mol^{-1}.
†-Inf denotes an undefined uncertainty.

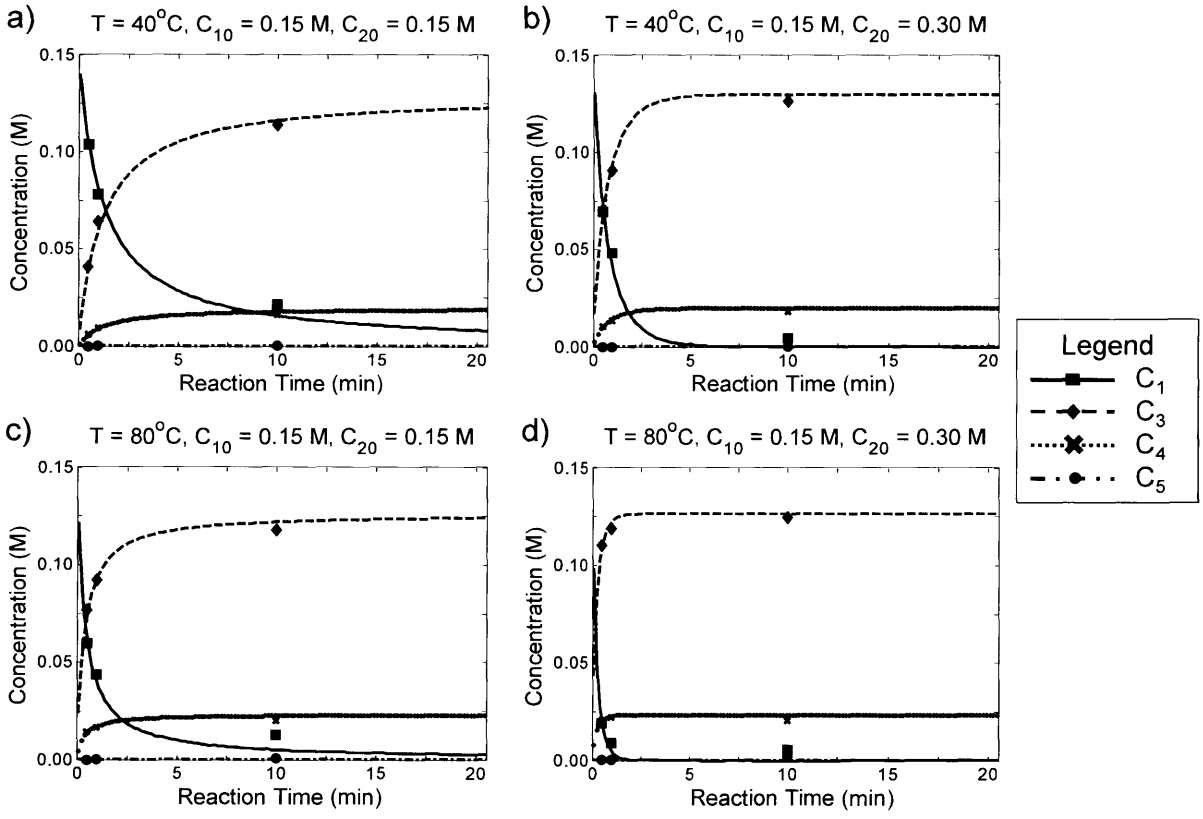


Figure 2.3(a-d). Experimental and model-predicted reactant and product concentration profiles after initial factorial design (12 automated experiments). Markers identify experimental data points. Lines indicate model prediction.

In order to minimize the uncertainty in the parameter estimates, the automated system selected the next D-optimal experiment to be performed at the maximum allowable residence time, temperature, concentrations of **1**, and equivalents of **2**. After the collection of this data point, parameter estimates were again calculated, this time using the parameter estimates and upper and lower bounds reported in Table 2.1 as inputs. $\mathbf{V}_B$ was also updated to agree with the covariance of the measured concentrations of the starting material and three products after the first 12 experiments. As Table 2.1 shows, significant improvements in the confidence of the estimates for all eight parameters were achieved after completing this first D-optimal experiment. In particular, uncertainties in the estimates of the pre-exponential factors and activation energies pertaining to k_3 and k_4 were all quantifiable, albeit reflective of greater than 50% error in the optimal parameter estimates.

The procedure for selecting and performing D-optimal experiments was repeated in an automated manner a total of 12 times (giving 24 experiments in total) before a user-specified termination. After each experiment, the initial guess for the parameter values, the bounds on the parameter values, and $\mathbf{V}_B$ were updated to agree with the results of the previous parameter estimation. It is notable from Table 2.1 that the uncertainties in parameter estimates improved greatly after experiments 13-18, but that the uncertainties improved only modestly from experiments 19-24. The optimal parameter values also changed little for all four rate constants from experiment 18 onward. Reasoning that further experimentation would only lead to modest improvements in parameter estimates and confidence intervals, we chose to terminate the method after experiment 24 and pursue a different approach to minimizing parameter uncertainty.

Figure 2.4(a-d) illustrates the agreement between the final best-fit model parameters and experimental data after 24 experiments. The model-fit and experimental data agree well across the range of temperatures tested and at short residence times. At long residence times, the model accurately fits the yield of **5**, but the conversion of **1** is overestimated and the yields of **3** and **4** are underestimated. We believe this to be a consequence of the peak resolution between **1** and **3** as measured by HPLC. A high conversion generally resulted in a strong **3** signal, which by broadening overlapped the weak **1** signal and reduced the accuracy of detecting and quantifying **1** at low concentrations. This claim is supported by the observation that at conditions of high conversion (temperatures at or above 80°C, excess initial concentrations of **2** and Et_3N), the measured conversion of **1** reproducibly reaches a maximum of 96%-97% regardless of residence

time. Providing that the SNAr reaction is irreversible, it is most likely that this replicated error results from the repeated bias introduced in detecting the weak **1** signal in close proximity to the strong **3** signal by HPLC.

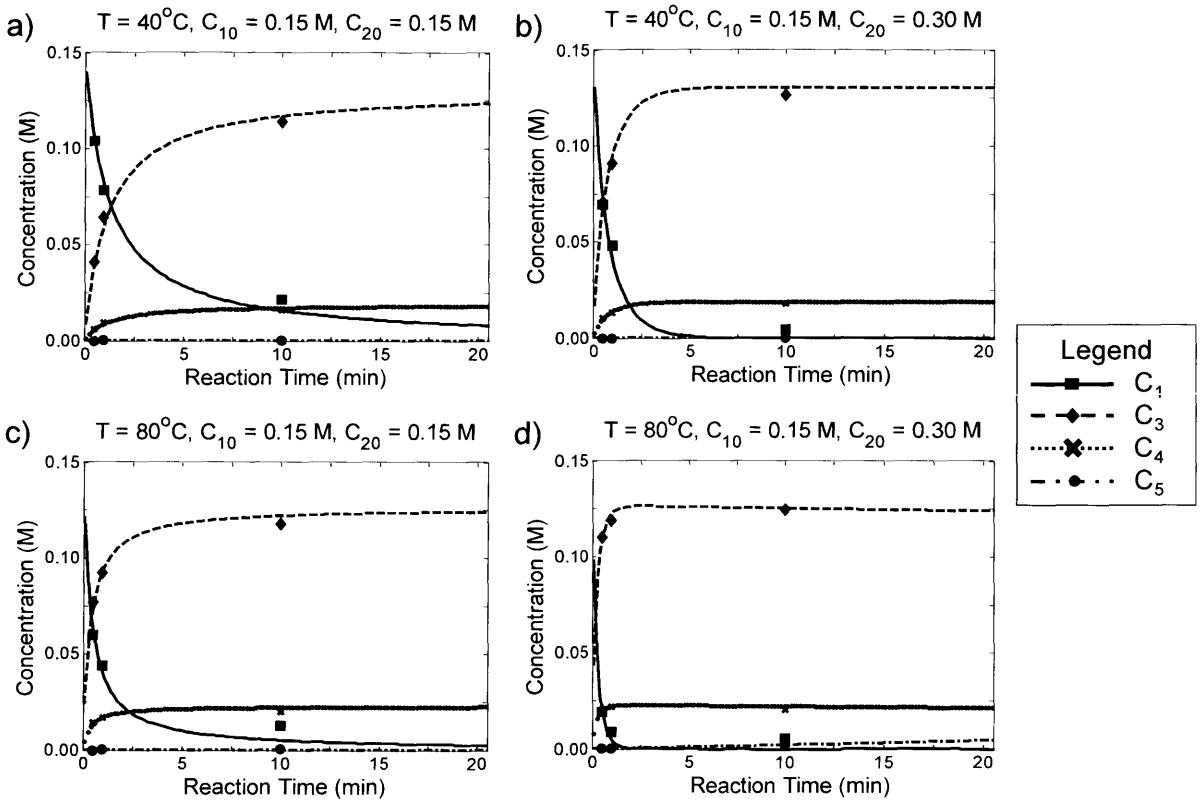
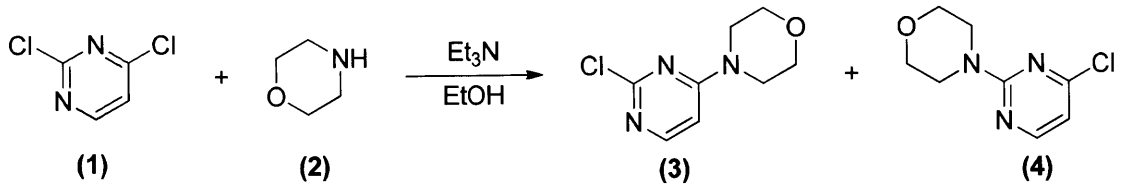


Figure 2.4(a-d). Experimental and model-predicted reactant and product concentration profiles after 24 automated experiments. Markers identify experimental data points. Lines indicate model prediction.

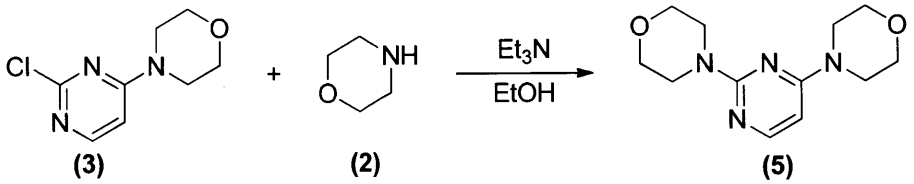
2.4.2. Estimation of Kinetic Parameters from Isolated Reactions

We proposed that parameter uncertainty could be reduced by decomposing the reactions in Scheme 2.1 into a sequence of isolated reactions. These isolated reaction steps are presented in Schemes 2.2, 2.3, and 2.4. The approach of isolating reaction steps is not uncommon to kinetic parameter estimation and may be beneficial in cases where the path through intermediates taken by the starting materials to reach the final product is unknown or ambiguous. In the case of the synthesis of **5**, we hypothesized that the large parameter uncertainties observed when attempting to estimate all eight kinetic parameters simultaneously implied an ambiguity in being able to identify whether the route from **1** to **5** went predominantly through the intermediate **3** or the intermediate **4**.



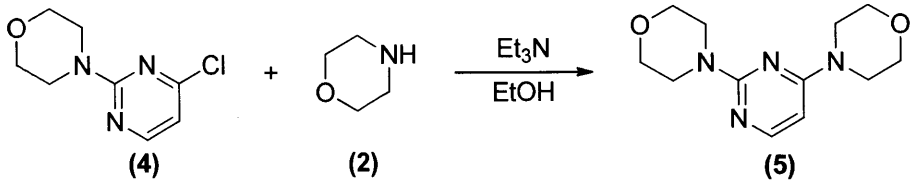
(1) (2) (3) (4)

Scheme 2.2. Reaction of 2,4-dichloropyrimidine and morpholine.



(3) (2) (5)

Scheme 2.3. Reaction of 4-(2-chloro-4-pyrimidinyl)-morpholine and morpholine.



(4) (2) (5)

Scheme 2.4. Reaction of 4-(4-chloro-2-pyrimidinyl)-morpholine and morpholine.

We sought first to optimize kinetic parameter estimates corresponding to k_1 and k_2 in Scheme 2.2. MAP estimation was employed for parameter estimation, with k_3 and k_4 constrained to their prior optimal values. Table 2.2 demonstrates that the approach of only estimating the first four kinetic parameters resulted in rapid convergence of the parameter values and substantial reductions in uncertainties. Though the optimal kinetic parameters presented in Table 2.2 shifted modestly from their prior MLE values, it is important to note that the posterior optimal values in all four cases are within the prior one standard deviation confidence intervals for each parameter found by the simultaneous parameter estimation approach. This observation would suggest that the posterior estimates of the four kinetic parameters are consistent with the results of the simultaneous parameter estimation experiment, though the posterior results have increased the likelihood that the optimal kinetic parameters are within closer proximity to the true parameter values.

Table 2.2. Optimal kinetic parameter estimates and uncertainties* from isolated estimation of parameters A_1, E_{A1}, A_2, and E_{A2}.

Number of Experiments	$\log_{10}(A_1)$	E_{A1}	$\log_{10}(A_2)$	E_{A2}
Prior	**3.3 ± 0.2**	**26.5 ± 1.5**	**3.3 ± 0.2**	**31.4 ± 1.6**
8	3.3 ± 0.2	25.9 ± 0.8	3.3 ± 0.2	31.1 ± 0.9
9	3.3 ± 0.1	26.0 ± 0.7	3.3 ± 0.1	31.3 ± 0.7
10	3.4 ± 0.1	26.6 ± 0.8	3.4 ± 0.1	31.9 ± 0.8
11	3.4 ± 0.1	27.0 ± 0.8	3.5 ± 0.1	32.2 ± 0.8
12	**3.4 ± 0.1**	**27.0 ± 0.7**	**3.5 ± 0.1**	**32.3 ± 0.7**

*-Uncertainties given as ±1 standard deviation. A_r is in $M^{-1}s^{-1}$ and E_{Ar} is in kJ mol^{-1}. Constrained values for parameters were $\log_{10}(A_3/M^{-1}s^{-1}) = 6.2$, $E_{A3} = 68.3$ kJ mol^{-1}, $\log_{10}(A_4/M^{-1}s^{-1}) = 2.0$, $E_{A4} = 36.0$ kJ mol^{-1}.

A challenge often presented in complex reaction networks is the isolation of intermediate products, such as **3** and **4** in Scheme 2.1. We already introduced the importance of selecting for 2-substituted pyrimidines as kinase inhibitors. In our kinetic investigation, it was similarly important to optimize for the synthesis of both the 2-substituted and the 4-substituted pyrimidines in order to isolate starting materials for the estimation of parameters in Schemes 2.3 and 2.4. Because **4** was known to be produced less favorably than **3**, we designed a synthesis which would maximize the yield of **4** at the maximum initial concentrations of **1** and **2.** Additionally, we required that our conversion of **1** exceed 99% to ensure that the starting material would not be present to complicate the isolation of products **3** and **4**:

$$\max_{t_{res},T,C_{10},C_{20}} \frac{\hat{C}_4}{C_{10}} \tag{2.24}$$

$$\text{s.t. } 0.5 \text{ min} \le t_{res} \le 20 \text{ min}$$

$$40°C \le T \le 100°C$$

$$C_{10} = 0.150 \text{ M}$$

$$C_{20} = 0.375 \text{ M}$$

$$\hat{C}_1 \le 0.01\left(C_{10}\right)$$

It can be derived from the proposed kinetic model that C_{10} affects the absolute concentrations of **1**, **3**, **4**, and **5** in the reaction but not the final product yields and selectivities. It followed that a greater initial concentration of **1** would allow for shorter reaction times with no adverse effect on the yield of **4**. We also found that a ridge of solution values existed for the yield optimization when C_{20} and t_{res} were allowed to vary independently. Figure 2.5a illustrates this ridge of optimal solutions at a temperature contour of 100°C, where the maximum yield is 17.1%. We reasoned

from this ridge of optimality that an optimum yield of **4** could be obtained in a minimum reaction time by specifying C_{20} at its upper bound of 0.375 M.

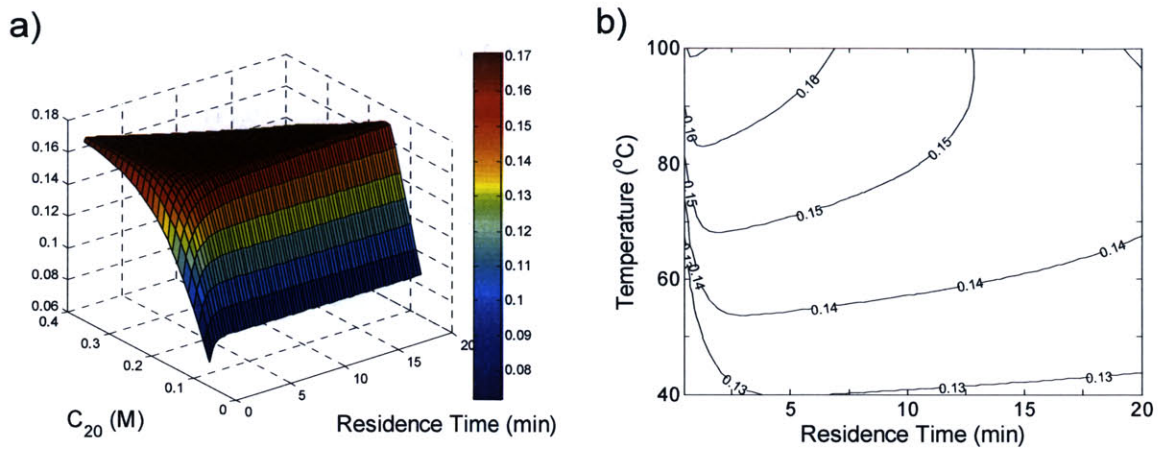


Figure 2.5. (a) Model predicted-yield of 4 with initial concentration C_{10} = 0.150 M and T = 100°C based upon optimal model parameters for k_1 and k_2 from Table 2.2 and for k_3 and k_4 from Table 2.1. The ridge of maximum yield is at 17.1%. (b) Model predicted-yield of 4 with initial concentrations C_{10} = 0.150 M and C_{20} = 0.375 M based upon optimal model parameters for k_1 and k_2 from Table 2.2 and for k_3 and k_4 from Table 2.1. The maximum predicted yield is 17.1% at t_{res} = 49 s and T = 100°C.

The formulation in Equation 2.24 generated a model-predicted optimal yield of **4** at a residence time of 49 s and a temperature of 100°C. This optimum can be seen visually from the contour plot shown in Figure 2.5b. To test the predictive capability of our model, we reacted 0.745 g of **1** at the optimized reaction conditions and analyzed reactant and product concentrations by online HPLC. The experimental yields and conversion (in mass units) based upon 0.745 g of **1** are compared in Table 2.3 to the predicted yields and conversion for the optimal set of experimental conditions. It was observed that our model predictions for the yields of **4** and **5** were fairly accurate, but that our model overestimated both the conversion of **1** and the yield of **3**. Following synthesis, **3** and **4** were each isolated by column chromatography, with isolated yields reported in Table 2.3 in comparison to the model predictions and online HPLC analysis. Isolated yields of both compounds from workup only were between 84% and 85% and, consequently, the relative selectivity of **3** to **4** was the same for the isolated yields as was measured by online analysis.

Table 2.3. Model-predicted, HPLC, and isolated yields for 1, 3, 4, and 5 for t_{res} = 49 s, T = 100°C, C_{10} = 0.150 M, and C_{20} = 0.375 M.

Product Distribution	Model-predicted†	Measured by HPLC	Isolated after Workup
C_{10} = 0.745 g			
Output **1** (g)	0.001	0.028	--‡
Output **3** (g)	0.821	0.794	0.698
Output **4** (g)	0.171	0.161	0.141
Output **5** (g)	0.006	0.008	--‡
% Conversion **1**	99.8	96.2	--‡
mol/mol% Yield **3**	82.2	79.5	69.9
mol/mol% Yield **4**	17.1	16.1	14.1
mol/mol% Yield **5**	0.5	0.6	--‡

†-Model predictions were calculated using optimal model parameters for k_1 and k_2 from Table 2.2 and for k_3 and k_4 from Table 2.1.

‡-Workup of the product was not attempted.

Using the isolated product **3**, we carried out the isolated reaction in Scheme 2.3 to estimate the parameters for A_3 and E_{A3}. Table 2.4 shows that greatly improved convergence of the parameter confidence intervals for $\log_{10}(A_3)$ and E_{A3} was observed for the isolated conversion of **3** and **2** to **5** and that the standard errors on the estimates of the two parameters were reduced by more than 50% in the eight experiments following the four-experiment factorial design. By comparison, the final eight D-optimal experiments in the simultaneous estimation experiment yielded an improvement in the uncertainties of the estimates of $\log_{10}(A_3)$ and E_{A3} of less than 30%. It was also observed that the optimal parameter estimates for A_3 and E_{A3} laid outside of the 2-standard deviation *a priori* confidence interval for the individual parameter estimates. Though this would indicate an inconsistency between the simultaneous and isolated experiment data sets, we believe the reported results to be acceptable on the basis that the path taken to a new set of optimal parameters remained within the 95% prior confidence intervals for A_3 and E_{A3} through the first eight experiments.

We next conducted the reaction of **4** with **2** (Scheme 2.4) for the estimation of A_4 and E_{A4}. Table 2.4 shows the convergence of the parameter estimates and single parameter confidence intervals for $\log_{10}(A_4)$ and E_{A4} over the course of 12 experiments. Both parameter standard errors improved by more than 67% from the *a priori* uncertainty values. The optimal parameter estimates for A_4 and E_{A4} were both found to be within the *a priori* 2-standard deviation single-parameter confidence intervals.

Table 2.4. Optimal kinetic parameter estimates from isolated estimation and uncertainties* of parameters A_3 and E_{A3} and parameters A_4 and E_{A4}.

Number of Experiments	$\log_{10}(A_3)$	E_{A3}	$\log_{10}(A_4)$	E_{A4}
Prior	**6.2 ± 0.6**	**68.3 ± 4.1**	**2.0 ± 0.8**	**36.0 ± 5.3**
4	6.2 ± 0.5	68.3 ± 3.7	2.1 ± 0.5	37.2 ± 3.1
5	6.0 ± 0.5	67.2 ± 3.5	2.2 ± 0.5	38.8 ± 3.2
6	5.8 ± 0.4	65.8 ± 2.9	2.7 ± 0.4	42.4 ± 2.5
7	5.6 ± 0.4	64.3 ± 2.9	2.7 ± 0.3	42.9 ± 2.3
8	5.3 ± 0.3	62.4 ± 2.4	2.9 ± 0.3	43.9 ± 1.9
9	5.1 ± 0.3	61.0 ± 2.3	2.8 ± 0.3	43.2 ± 2.2
10	4.9 ± 0.3	59.7 ± 1.8	2.8 ± 0.3	43.7 ± 1.9
11	4.7 ± 0.3	58.0 ± 1.9	2.8 ± 0.3	43.5 ± 1.8
12	**4.8 ± 0.2**	**59.0 ± 1.7**	**3.0 ± 0.2**	**44.7 ± 1.7**

*-Uncertainties given as ±1 standard deviation. A_3 is in $M^{-1}s^{-1}$ and E_{A3} is in kJ mol^{-1}.

We sought upon completion of the isolated experiments to reconcile our updated set of parameters and uncertainties in a final set of MAP estimation experiments. Such experiments were necessary to account for any interaction effects between species in the reaction network and to correct for the uncertainty introduced from using lower purity chemicals in the isolated reactions of **3** and **4** with **2**. Optimizing over the same set of parameters but with the prior estimates and uncertainties found from conducting the set of isolated experiments, we were able to obtain rapid convergence of our parameter estimates with uncertainties greatly reduced over the initial simultaneous estimation approach. The convergence of the parameter estimates from experiments 12 through 18 of the culminating experiment are shown in Table 2.5. Optimal log-pre-exponential terms were estimated to a standard deviation of ±0.1 $M^{-1}s^{-1}$ for k_1 and k_2 and to ±0.2 $M^{-1}s^{-1}$ for k_3 and k_4. Likewise, standard errors for the activation energy terms were estimated as ±0.6 kJ mol^{-1} for rate constants 1 and 2 and ±1.6-1.7 kJ mol^{-1} for rate constants 3 and 4. Figure 2.6(a-f) compares the predicted concentration profiles based upon the optimal model parameters in comparison to the data for all experiments. The model-fit values appear to agree with the experimental data.

Table 2.5. Optimal kinetic parameter estimates and uncertainties from final simultaneous experiment in isolated approach.

Number of Experiments	$\log_{10}(A_1)$	E_{A1}	$\log_{10}(A_2)$	E_{A2}	$\log_{10}(A_3)$	E_{A3}	$\log_{10}(A_4)$	E_{A4}
Prior	**3.4 ± 0.1**	**27.0 ± 0.7**	**3.5 ± 0.1**	**32.3 ± 0.7**	**4.8 ± 0.2**	**59.0 ± 1.7**	**3.0 ± 0.2**	**44.7 ± 1.7**
12	3.5 ± 0.1	27.3 ± 0.7	3.5 ± 0.1	32.1 ± 0.7	5.0 ± 0.2	60.4 ± 1.7	3.2 ± 0.2	46.3 ± 1.8
13	3.5 ± 0.1	27.1 ± 0.7	3.5 ± 0.1	32.2 ± 0.6	4.8 ± 0.2	59.0 ± 1.7	3.0 ± 0.2	45.0 ± 1.7
14	3.4 ± 0.1	27.0 ± 0.6	3.5 ± 0.1	32.2 ± 0.6	4.8 ± 0.2	59.0 ± 1.7	3.0 ± 0.2	45.0 ± 1.8
15	3.4 ± 0.1	27.0 ± 0.6	3.5 ± 0.1	32.1 ± 0.6	4.8 ± 0.2	58.9 ± 1.7	3.0 ± 0.2	45.0 ± 1.8
16	3.4 ± 0.1	27.0 ± 0.6	3.5 ± 0.1	32.1 ± 0.6	4.8 ± 0.2	58.7 ± 1.7	3.0 ± 0.2	45.0 ± 1.8
17	3.4 ± 0.1	27.0 ± 0.6	3.5 ± 0.1	32.1 ± 0.6	4.9 ± 0.2	59.4 ± 1.6	3.0 ± 0.2	45.0 ± 1.8
18	**3.4 ± 0.1**	**27.0 ± 0.6**	**3.5 ± 0.1**	**32.1 ± 0.6**	**4.9 ± 0.2**	**60.0 ± 1.6**	**3.0 ± 0.2**	**45.0 ± 1.7**

*-Uncertainties given as ±1 standard deviation. A_r is in $M^{-1}s^{-1}$ and E_{Ar} is in kJ mol^{-1}.

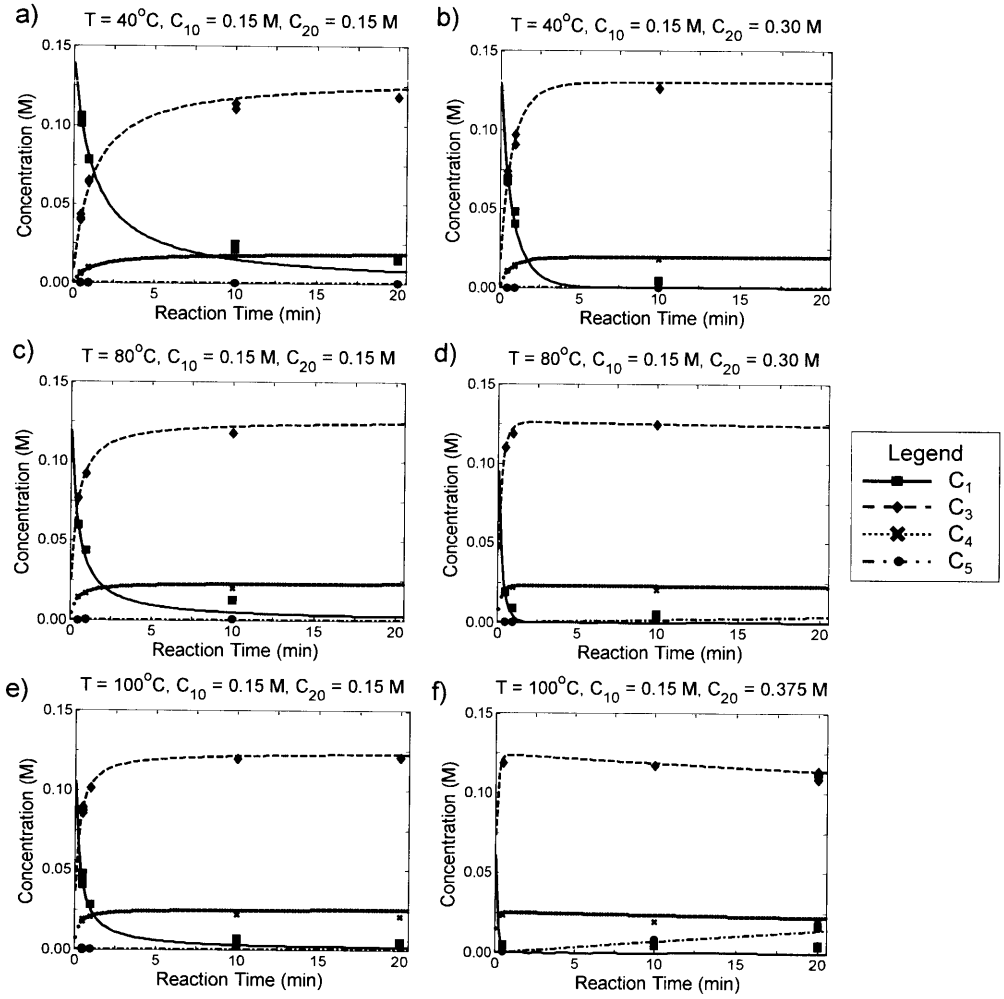


Figure 2.6(a-f). Experimental and model-predicted reactant and product concentration profiles after completion of all experiments (including simultaneous and isolated approaches). Markers identify experimental data points. Solid lines indicate model prediction.

In total, the procedure for first simultaneously estimating kinetic parameters then isolating products and refining the kinetic parameters in isolated experimentation required 78 automated experiments to complete (with 54 experiments dedicated to estimating each kinetic parameter) and required 7 days of cumulative time. Less than 5 g of **1** were consumed for all experiments and the synthesis and isolation of **3** and **4.** A more streamlined workup and isolation routine in the future would reduce the experiment time further, as workup and stock sample preparation were the only steps carried out manually. The duration of the automated experiments was primarily determined by the time required to reach steady-state for the longest residence time experiments and the rate of convergence of the system, which was a function of both the sensitivity of the experiments to the model parameters and the accuracy and precision of the continuous flow method.

2.5. DISCUSSION

The complexity of many pharmaceutical syntheses dictates that reactions be parameterized and optimized accurately and with minimal uncertainty when scaled to a production level. A large degree of uncertainty introduces the potential for inaccurate reaction scale-up, leading to lesser yields and/or increased formation of detrimental byproducts. In the initial simultaneous parameter estimation experiments, we demonstrated an automated approach that, albeit accurate, failed to reasonably minimize parameter uncertainty. The results suggested that although the model predictions in Figures 2.3 and 2.4 appeared reasonable, there in truth may have been many parameters within an error of as large as 20% yielding an acceptable fit of the experimental data. Optimizing or predictably scaling-up a system while considering these large uncertainties in kinetic parameters is infeasible.

We hypothesized from our initial results that the large uncertainties in the parameter estimates derived from the correlation in the model parameters. From examination of Equations 2.15 and 2.16, we identified two key factors which we expected to contribute to the correlated uncertainties between parameters. The first factor considered was the calculated response covariance $\mathbf{V}_B$, which indicates the variability in both the experiment and in the measurement of the data. As continuous flow systems are excellent in their control of reaction conditions and residence time, we proposed that the variability in the experimental setup was not the major factor contributing to the large uncertainties. Alternatively, we considered the large uncertainties

to be primarily a result of low parameter sensitivity across the range of experimental conditions. To address this low sensitivity, we restructured the parameter estimation so as to estimate the eight kinetic parameters from the isolated reactions in Schemes 2.2 through 2.4. From this restructuring, we were able to achieve significantly reduced parameter uncertainties compared to those in the simultaneous approach and confirm our notion that the originally high uncertainties resulted from low parameter sensitivity.

In models with strong correlations among parameters, it is of interest to calculate multi-parameter confidence regions, which show the dependence of the estimation of one parameter upon another. These $1 - \alpha$ joint confidence regions can be found for two or more parameters from the inequality:[126]

$$\left(\boldsymbol{\theta}-\boldsymbol{\theta}_{opt}\right)' \mathbf{V}_{\theta}^{-1}\left(\boldsymbol{\theta}-\boldsymbol{\theta}_{opt}\right) \leq N_{params} F_{1-\alpha,\nu_1=N_{params},\nu_2=N_{expts}-N_{params}} \tag{2.25}$$

Joint-confidence regions for this study were found by calculating the probability value associated with a multi-dimensional perturbation in $\boldsymbol{\theta}$ away from $\boldsymbol{\theta}_{opt}$. The probability distribution followed an F cumulative distribution function with $1 - \alpha$ confidence and N_{params} and N_{expts} - N_{params} degrees of freedom. As in the case of posterior confidence intervals, to find posterior joint confidence intervals $\mathbf{V}_{MAP}^{-1}$ had to be substituted for $\mathbf{V}_{\theta}^{-1}$.

Figure 2.7(a-d) shows the correlated uncertainty between pre-exponential factors and activation energies after the simultaneous parameter estimation experiment. An elongated elliptical confidence region identifies two strongly correlated parameters and suggests that a change in the optimal value of one parameter will be reflected in a change in the optimal value of the other parameter. This is often the case in determining best-fit pre-exponential factors and activation energies, which are coupled by respective rate constants. Figure 2.7(a-d) demonstrates that, indeed, the pre-exponential terms and activation energies were strongly correlated for each of the four rate constants estimated for the SNAr reaction. The joint confidence regions for A_3 and E_{A3} and for A_4 and E_{A4} in particular indicated that the optimal parameter values existed beyond what would have been expected had we only to considered the single-parameter uncertainties.

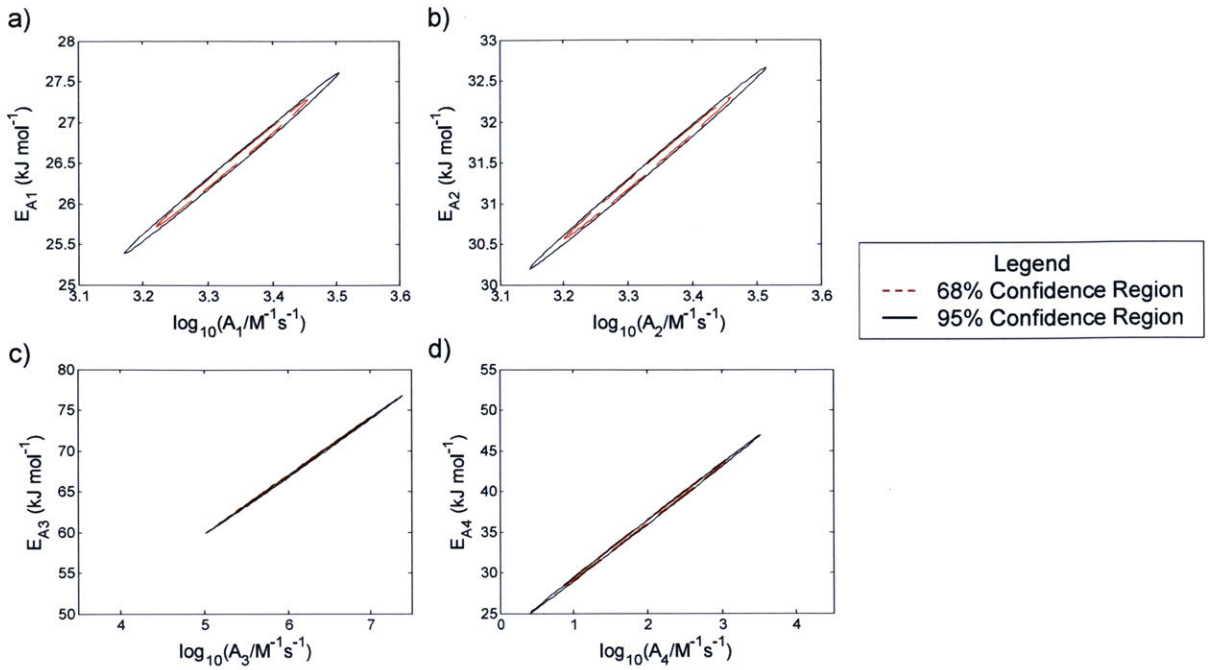


Figure 2.7(a-d). 68% and 95% joint confidence regions for estimated parameters after 24 automated experiments.

Figure 2.8(a-d) shows the final set of converged joint confidence regions corresponding to k_1 through k_4. The activation energies and pre-exponential factors, especially for k_3 and k_4, still show strong correlation, although comparison to Figure 2.7(a-d) demonstrates that the isolated experiment approach contributed greatly to minimizing the joint uncertainties between rate constant parameters. Drawing upon our reasoning that the parameter uncertainties arise from the inherent coupling of the system, we would expect to be able to reduce the size and aspect ratio of the joint confidence ellipses by exploring more sensitive regions of the experimental space for the reactions of **1**, **3**, and **4** with **2** than those considered for this automated study.

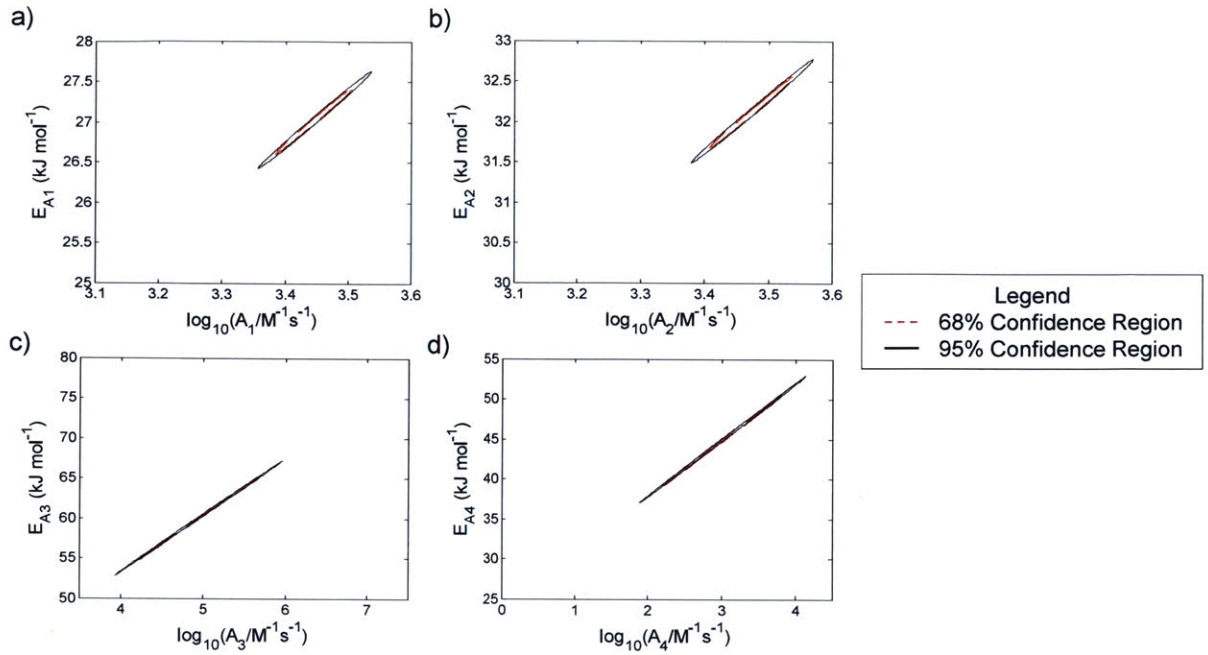


Figure 2.8(a-d). 68% and 95% joint confidence regions for estimated parameters after all simultaneous and isolated automated experiments.

The parameter correlation matrix provided a useful measure of the strength of correlation between model parameters. Each element of the matrix was defined as:

$$\boldsymbol{\rho}_{ij}(\mathbf{V}_\theta) = \frac{(\mathbf{V}_\theta)_{ij}}{\sqrt{(\mathbf{V}_\theta)_{ii}(\mathbf{V}_\theta)_{jj}}} \tag{2.26}$$

Entries of $\boldsymbol{\rho}$ with magnitude close to 1 indicated a strong correlation between the respective parameters.

After the initial set of 24 simultaneous experiments, we calculated $\boldsymbol{\rho}$ as:

$$\boldsymbol{\rho}(\mathbf{V}_\theta) = \begin{array}{c} \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \\ \theta_8 \end{array} \begin{array}{c} \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & \theta_7 & \theta_8 \end{bmatrix} \\ \begin{bmatrix} 1.000 & 0.998 & 0.977 & 0.972 & 0.005 & 0.010 & 0.060 & 0.052 \\ 0.998 & 1.000 & 0.977 & 0.976 & -0.005 & 0.000 & 0.069 & 0.061 \\ 0.977 & 0.977 & 1.000 & 0.998 & -0.024 & -0.016 & 0.126 & 0.113 \\ 0.972 & 0.976 & 0.998 & 1.000 & -0.041 & -0.034 & 0.142 & 0.130 \\ 0.005 & -0.005 & -0.024 & -0.041 & 1.000 & 1.000 & -0.840 & -0.850 \\ 0.010 & 0.000 & -0.016 & -0.034 & 1.000 & 1.000 & -0.835 & -0.846 \\ 0.060 & 0.069 & 0.126 & 0.142 & -0.840 & -0.835 & 1.000 & 0.999 \\ 0.052 & 0.061 & 0.113 & 0.130 & -0.850 & -0.846 & 0.999 & 1.000 \end{bmatrix} \end{array} \tag{2.27}$$

From the structure of the Arrhenius equation, it is typical to expect a strong correlation between the activation energy and the pre-exponential factor for the same rate constant. This high correlation was indeed observed for the initial set of simultaneous experiments. In addition, the correlation matrix for the first 24 experiments indicated strong correlation in the model between k_1 and k_2 and between k_3 and k_4. To justify this observation, we considered first that a change in either k_1 or k_2 would result in a necessary change in k_2 or k_1 to maintain the correct predicted selectivity of **3** to **4** to support the experimental data. Furthermore, k_3 and k_4 had to offset one another in order to describe the yield of **5**.

The effect of performing isolated parameter estimation experiments upon correlations between model parameters for the SNAr case was measured quantitatively in the final correlation matrix:

$$\boldsymbol{\rho}\left(\mathbf{V}_\theta\right)=\begin{array}{c}\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & \theta_7 & \theta_8 \end{bmatrix} \\ \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \\ \theta_8 \end{bmatrix}\begin{bmatrix} 1.000 & 0.997 & 0.976 & 0.968 & -0.002 & 0.003 & 0.000 & 0.000 \\ 0.997 & 1.000 & 0.976 & 0.973 & -0.003 & 0.003 & 0.000 & 0.000 \\ 0.976 & 0.976 & 1.000 & 0.997 & -0.006 & -0.006 & 0.000 & 0.000 \\ 0.968 & 0.973 & 0.997 & 1.000 & -0.007 & -0.007 & 0.000 & 0.000 \\ -0.002 & -0.003 & -0.006 & -0.007 & 1.000 & 1.000 & -0.008 & -0.008 \\ -0.003 & -0.003 & -0.006 & -0.007 & 1.000 & 1.000 & -0.008 & -0.008 \\ 0.000 & 0.000 & 0.000 & 0.000 & -0.008 & -0.008 & 1.000 & 1.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & -0.008 & -0.008 & 1.000 & 1.000 \end{bmatrix}\end{array} \tag{2.28}$$

Upon calculation of the final parameter correlation matrix, we observed that the isolated reaction approach greatly reduced the correlation between k_3 and k_4 and that minimal correlation remained between the first two rate constants and the last two rate constants. Strong correlations still remained between all activation energies, their respective pre-exponential factors, and between k_1 and k_2. To isolate the pre-exponential factors and activation energies, we would have needed to conduct experiments at largely different reaction temperatures so as to increase the model sensitivity to the activation energy parameters. Discerning better between k_1 and k_2 would likely have required much higher temperatures than were achievable with ethanol as the solvent, or instead the ability to run the reaction at much shorter residence times.

We suspected that the remaining parameter uncertainty after the set of posterior experiments was a combination of lower sensitivity in discriminating between the activation energy and the pre-exponential factor for each rate constant and of the inherent error in the system, estimated as

$\mathbf{V}_B$. From the final calculated $\mathbf{V}_B$, the error in measuring **1** was found to be ±0.0064 M (±4.3% based on C_{10}), with the errors in measuring **3**, **4**, and **5** calculating to ±0.0026 M, ±0.0019 M, and ±0.0011 M, respectively. These errors can be interpreted as the limiting precision of the kinetic model, given uncertainties in flow rates, starting material purity, temperature control, and online analysis. To achieve this limiting precision, one would need improve the sensitivity of the model to each activation energy and pre-exponential factor. This could be accomplished either by conducting experiments at more extreme temperatures, by increasing the feed concentrations of starting reagents (to increase conversion to **5**), or by incorporating quantum calculated pre-exponential factors into the MAP estimation.

It is important to recognize that although the parameter estimation improved substantially by analyzing isolated reactions under conditions of greater sensitivity, the method we have employed still relies upon obtaining reasonable estimates of the rate parameters in the simultaneous approach. By incorporating MAP estimation into the method, we demonstrated that the information gained from the simultaneous approach could be incorporated as *a priori* information in the isolated reaction approach to provide initial parameter estimates and to further reduce parameter uncertainty. Additionally, the optimal parameters found in the simultaneous experiment and for the isolated reaction in Scheme 2.2 proved to be necessary in finding conditions at which an optimal yield of **4** could be obtained.

2.6. CONCLUSIONS

The advancement of continuous flow technology with online feedback has enabled the development of automated systems capable of parameterizing and optimizing chemical syntheses with little *a priori* reaction information. For appropriate reactions and conditions, such automated systems have the potential to minimize consumption of valuable reagents while providing the requisite information for reaction scale-up. Though these systems are quite interesting for demonstration purposes, it is trivial to find cases in which the chemistry under study is too complicated to be parameterized in a handful of flow experiments. Here we demonstrated an automated platform and procedure that were both efficient in conserving reagents and effective at parameterizing a complex reaction network.

As we reviewed the performance of this automated system, we identified several limitations in the extension of this methodology to more challenging chemical syntheses. In terms of

simplicity of use, the complexity of the physical system (specifically the number of pumps) grew with the number of substrates or additives to be tested in the optimization, unless the user manually changed reagents as in the procedure outlined in this chapter. Though low amounts of material were consumed by mass, the amount of reagent consumption per data point (at least 1 mL reaction volume per 2 μL HPLC sample) would still be considered wasteful in pharmaceutical development. In terms of versatility, the limited scope of optimizations or kinetic parameter studies that could be covered over a range of 60°C temperature and $\leq$ 1 order of magnitude reaction time and reagent concentrations greatly detracted from the utility of the method. In identifying reaction kinetics or characterizing product formation, one would ideally prefer to have as much flexibility as possible to manipulate experimental conditions in order to maximize sensitivity. One way to greatly expand the versatility of such automated flow platforms would be to incorporate discrete variables along with continuous variables into the optimization. This is the essence of the work presented in the remaining chapters of this thesis.

3. A Segmented Flow System for On-Demand Screening of Discrete and Continuous Variables

3.1. Introduction

It is generally the case in chemical synthesis that the interplay of multiple variables factors significantly into the rate of product formation and the scalability of the reaction. To tease out these relationships becomes a labor-intensive and time-consuming task—first sample preparation, then reaction screening, then workup and analysis, then review of the data and model fitting, then finally feedback to explore a new region of the experimental design space. With the growth in popularity of HTE,[62-66] many of these steps can now be automated and streamlined, at the expense being able to manipulate key factors such as reaction time and temperature—factors which are essential to the extraction of reaction kinetics for scale-up. Though adaptations of 96-, 384-, and 1536-well plate systems work well for discovery, they also become costly in substrate and catalyst usage when many unnecessary experiments are run in parallel—in comparison to the careful selection of experiments designed to maximize information as was demonstrated in Chapter 2.

Reaction screening "on-demand" using an automated segmented flow system that mixes reagents, reacts, and analyzes, would stand to offer substantial material savings over batch HTE. The efficient mixing and heat transfer in microfluidic devices would further offer better scalability of results. Though rich in potential, the development of a flow screening system with the universality of batch HTE has in the past presented a number of challenges—the predominant problems being dispersion in different mobile phases, carryover from one sample to the next, and versatility in reagents and materials of construction. We sought to address these challenges with development of an automated flow screening system versatile enough to screen a wide range of organic syntheses at dynamic temperature, composition, and time conditions. The system was to be fully automated and integrated into online analysis equipment, allowing chemists a "set-it-and-forget-it" method for accurately characterizing reaction performance across scales.

Automated flow microreactor systems for HTE can be generally classified into parallel, homogeneous isolated reagent, or segmented flow designs. Automated parallel microreactor systems provide a rather intuitive means for screening various solid-phase heterogeneous catalysts. Examples of the applications of automated parallel systems have included the

enantioselective hydrogenation of methyl-2-acetamido acrylate,[130] the partial oxidation of 1,3-butadiene to furan,[131] and the selective catalytic reduction of NO_x by propane.[132] Shi *et al.* have additionally presented a parallel capillary-based system for homogeneous catalyst screening.[133] The homogeneous system was capable of screening 20 Stille cross-coupling reactions simultaneously with data analysis conducted by online GC. The demonstration succeeded in identifying an optimal palladium precatalyst and arsine ligand that closely resembled that already employed in literature. Library generation has been established in parallel systems for the syntheses of pyrazoles,[134] ciprofloxacin analogs,[135] and sulfonamides.[136] Wang *et al.* have presented a networked system of microfluidic channels for the parallel screening of 32 *in situ* click chemistry reactions.[137] A library of 20 azides was screened in the automated system, with only 120 nL of each azide consumed per reaction. An integrated system such as this has the potential for great economic upside when cost of reagents is a limiting factor in library compilation.

Homogeneous isolated reagent and segmented flow methods of reaction and catalyst screening have the potential for much higher throughput and smaller reaction volumes than parallel microreactor schemes. These sequential reaction methods further allow for greater flexibility in selecting reactive components. An example of the sequential reaction approach using homogeneous isolated reagents for library development was demonstrated by Goodell *et al.*.[54] Reactants in this system were withdrawn from a reagent block by a liquid handler and injected as isolated segments into a homogeneous solvent phase flowing through a microreactor. In total, over 1000 reactions were screened and monitored by UPLC in order to generate a library of functionalized bicycle[3.2.1]octanoid scaffold transformations. The most notable drawback of the homogeneous flow approach is the axial dispersion introduced by laminar flow, which presents the risk of cross-contamination between segmented reagents.

In large, segmented flow systems comprise the most versatile systems for continuous HTE. Segmented flow, or Taylor flow, is the description given to two-phase flow in a microreactor in which the continuous and dispersed phases form distinct, segmented droplets (slugs). The formation of slugs is dependent upon the ratio of viscous stress to the interfacial tension stresses between the two immiscible phases, a ratio given by the capillary number (Ca):[138,139]

$$\mathrm{Ca} = \frac{U\eta}{\sigma} \tag{3.1}$$

U in Equation 3.1 represents the fluid velocity, η is the fluid viscosity, and σ is the interfacial tension between the two phases. Slugs form at low values of Ca, generally < 0.01. Much of the interest in segmented microreactor flows over the last few years can be attributed to the enhanced mixing within either phase of the two-phase system. This mixing is the result of recirculating flows induced by the shear forces exerted on the slug against the microchannel wall.[140] The relative motion of the internally recirculating slugs at turns in winding microchannels even more greatly enhances chaotic advection within the dispersed phase and allows for extremely short mixing timescales within the slugs.[141]

The narrow residence time distributions achievable in segmented flow have stimulated interest in the two-phase flow technique as a means of continuously executing compartmentalized reactions or crystallizations requiring precise residence time control. The applications of segmented flow have in turn been expanded to HTE under the premise that different reactions can be run in separate slugs flowing in series through the microreactor.[95,96,100] Several different on-chip and off-chip techniques have been implemented for generating nanoliter-scale slugs of reagents, the simplest on-chip method being *via* the injection of the dispersed and continuous phases through two separate branches of a T-junction.[142,143] Slugs of the dispersed phase solvent are produced as a result of the shear and interfacial forces experienced at the interface of the two fluids at the T. Flow focusing, in which two outer continuous phase channels merge with an inner dispersed phase channel at a narrow orifice, has additionally been employed to generate slugs on-chip.[101,144,145] In on-chip techniques, slugs with different reagent compositions are prepared by merging solutions at different flow rates at a junction with the dispersed phase.[146,147]

As an alternative to passive techniques, active techniques for slug generation are available for control of slug size and composition independent of the fluid properties. Examples of active techniques include the use of valves, off-chip or on-chip, to create slugs of nanoliter precision[148,149] or to digitize slugs into smaller volumes that can be merged with different quantities of reagents.[103] The screening of many different substrates has required the preparation of individual sample droplets off-line before on-chip introduction of other reactants or catalysts.[150] Autosampling devices which extract different volumes of prepared reagent or catalyst solutions and then inject these samples as immiscible slugs into a constantly-flowing continuous phase greatly expedite this process. The automated system demonstrated by Garcia-

Egido *et al.* in the synthesis of a library of pyrazoles illustrates the versatility of the autosampling technique.[151]

For reaction screening applications in particular, the ability to controllably add reagents into slugs is essential to precise reaction time control. A simple approach to controlled reagent addition online was proposed by Hatakeyama *et al.*, who infused reagents *via* syringe pump though a T-junction directly into a premade slug.[96] With on-chip designs, the approach of adding reagents at a T-junction can be enumerated several times to generate combinatorial screens of reagents and slug volumes.[103] As an alternative to the continuous T-junction approach, multiple reactant slugs may be generated and merged together on-demand, as was demonstrated by Niu *et al.*[152] In the example, a series of pillars were designed into the microfluidic device to enable entrainment of a slug by surface tension, followed by merging with a second slug, and then release of the newly combined slug because of the change in hydraulic pressure in the system. This technology has further enabled controlled droplet dilution for a DNA-binding assay.[153]

Analysis of individual slugs has been performed either by observing each slug by spectroscopy[154] or by direct analysis of the slugs by LC, GC, or MS.[96,151,155] Regardless of the type of instrument employed for analysis, the instrument must have the capacity to handle small volumes of fluid. For real-time screening and optimization, rapid feedback is also desirable. Indexing of slugs has been achieved through the generation of pairs of slugs, with one slug carrying the reactants and the other containing an indicator identifying the composition of its partner.[156] In addition to serving as a reference guide during analysis, indicator slugs have found further application as sensors for possible inter-slug cross-contamination, which can be detected during analysis of the index.[98]

The enhanced mixing achievable in segmented flow greatly improves the accuracy and versatility of continuous HTE, particularly in chemical synthesis applications. One example segmented flow system was developed by Fang *et al.* for the purpose of screening acid catalysts in a Friedel-Crafts reaction.[155] The system incorporated a capillary microreactor, and online analysis was performed by direct injection into a UPLC. Clausell-Tormos *et al.* applied segmented flow to develop a library of inhibitors of the enzyme β-galactosidase.[99] Samples were originally prepared in a 96-well plate and then transferred *via* an autosampler into slugs to be reacted in a microchannel. The developed system had the rather novel capability of allowing for both the merging and the splitting of slugs, which respectively allowed for more robust reactant

combinations and replicate slugs to be tested. Florescence measurements were made online and even allowed for reaction monitoring throughout the length of the reactor, potentially facilitating future kinetics studies. Theberge *et al.* employed a segmented flow system for merging of three different reactants in an Ugi-type multicomponent reaction, generating a library of small molecules with the potential for thrombin inhibitory activity.[100] Li *et al.* presented an automated segmented flow system for simultaneous reaction screening and optimization.[97] Volumes of reaction slugs in this study were nominally 140 nL, with 20 different reagent combinations screened. Computer control of the system furthermore allowed for residence times and initial reactant concentrations within each slug to be manipulated such that 1000 different reactant, concentration, and residence time combinations were screened in total. Kreutz *et al.*[98] demonstrated coupling of a genetic optimization algorithm to the screening of catalysts, cocatalysts, and ligands for the oxidation of methane in a segmented flow system.

Commercial on-demand segmented flow screening systems have been introduced by the Accendo Corporation. These systems allow users to screen up to 40 reagents sequentially in slug flow through a tubular reactor with online detection by LC-MS. Slug volumes range on the order of 20-900 μL, and the system throughput is maximized at the analysis of one slug every three minutes. The systems operate by generating a three-phase segmented flow system comprising the reaction volume, an immiscible phase of perfluoromethyldecalin, and a transport phase. Screening modes, where each slug is analyzed independently, and prep modes, where the same conditions can be repeated to generate greater amounts of material, are available. Though several successful library syntheses have been reported with these commercial systems,[107-112] recent scrutiny has emerged regarding the accuracy of the use of these systems for scale-up. Accendo has reported online the phenomenon of "incubation acceleration," stating that slugs move faster than predicted through the flow system as temperature increases.[157] A recent study by Hawbaker *et al.* also illustrated the diffusion of reagents from the reaction phase along the wall contacting the fluorinated phase and into the transport phase, leading to a loss in conversion as a function of flow rate and tubing material.[158] These factors, along with a need to be able to modulate temperature and flow rate and control reaction initiation and termination, motivated us to design our own in-house system for automated on-demand reaction screening and optimization.

3.2. METHOD

Our efforts in developing an on-demand microfluidic screening system focused on the use of biphasic reaction segments (or slugs) for a more accurate representation of the batch reaction upon scale-up without the complication of the effect of dispersion on the reaction yield. For ease of reagent handling, we concentrated our efforts on screening liquid phase reactions—and hence homogenous catalysts—only. Reagents which could be dissolved in the reaction solvent or a suitable co-solvent were acceptable. We targeted optimizing with reaction volumes less than 20 μL. In general, the smaller the reaction volume, the better the mixing within the slug[159] and the greater the ratio of material sampled to material used for analysis. With smaller dimensions, however, came greater challenges to the precision of reagent handling and slug detection. Many of these challenges will be elaborated upon in the following sections.

Figure 3.1 overviews schematically the system constructed for automated and on-demand slug preparation, reaction, analysis, and feedback. Compared to previous examples of screening systems in literature, the system was found to be unique in its versatility for organic reactions and its utility in accurately representing larger batch-scale chemistry. The integration of sensory equipment and control software made the system simple to operate and fully automated, allowing the experimenter to provide reagents, propose experiments or a range of conditions over which to optimize, start the system, and return several hours later to retrieve reaction data. Slugs were prepared individually by a liquid handling robot, introduced through a sample loop, and transported by an inert carrier phase to a tubular reactor. An additional reagent could be added to the slug online to initiate the reaction. After reaction, the slugs were quenched and sampled by an online HPLC or LC/MS. With the assistance of an MINLP feedback algorithm (discussed in Chapter 4), new reagents and reaction conditions were proposed for subsequent slugs based on the yields of prior experiments. A full system diagram is presented in Figure 3.2, with photographs of main system components including pumps, liquid handling, analysis, automation, reaction, and sampling shown in Figure 3.3(a-b).

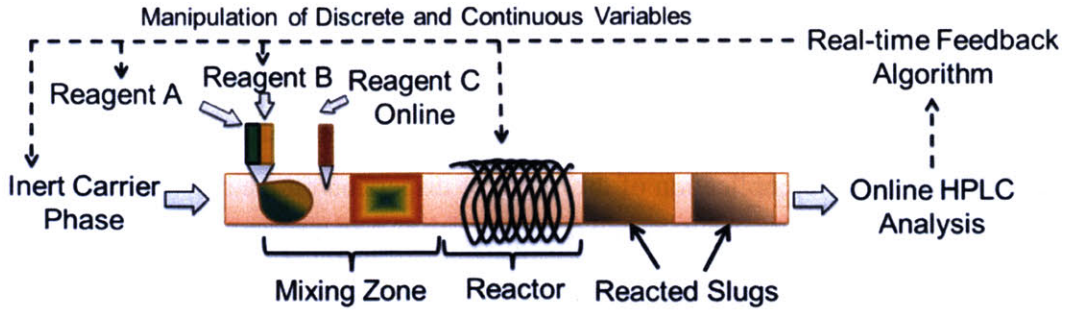


Figure 3.1. Concept diagram for on-demand preparation, reaction, analysis, and feedback in an automated reaction flow screening system.

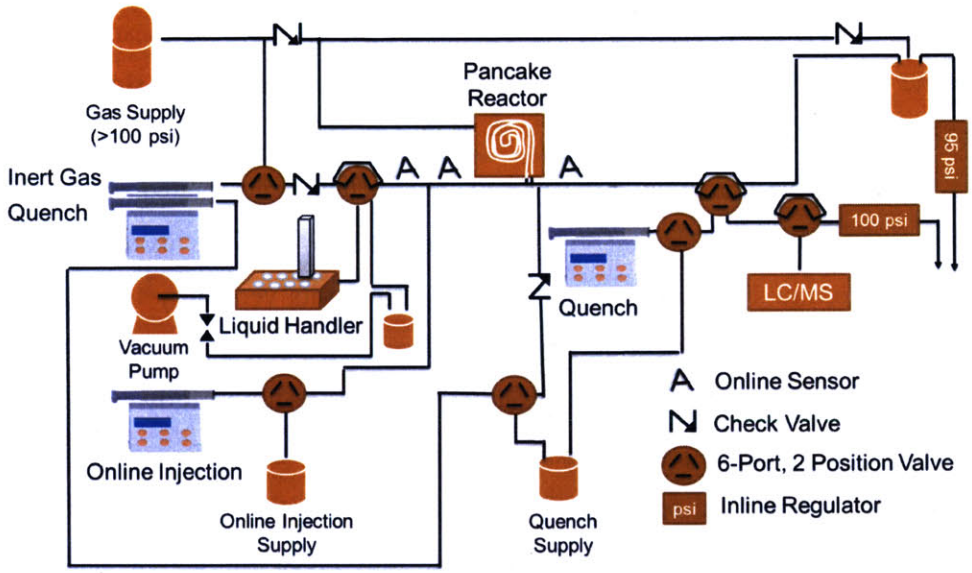


Figure 3.2. Schematic of automated flow system for alkylation reaction optimization.

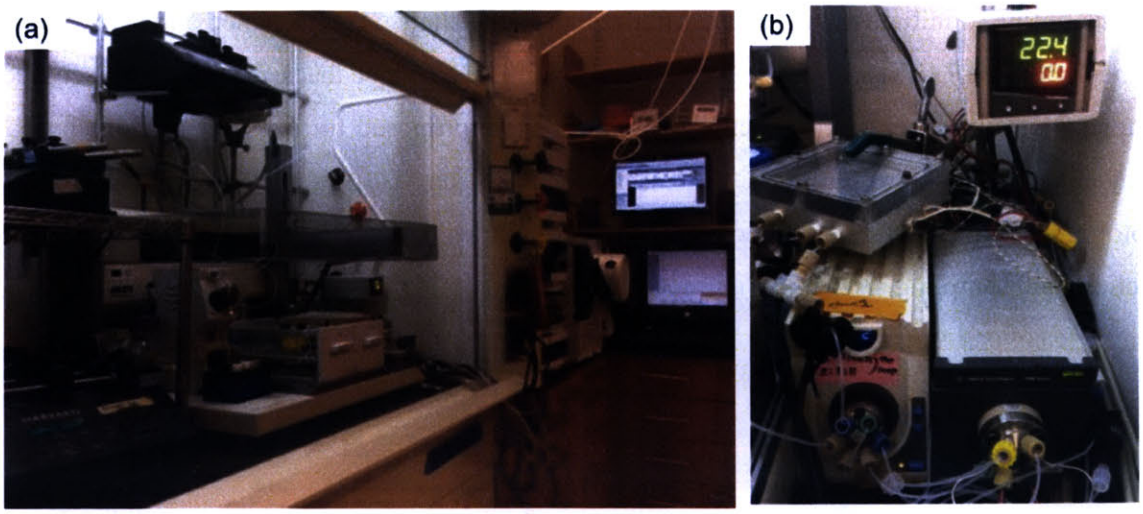


Figure 3.3. Automated system hardware including (a) pumps, automated liquid handler, LC/MS, and automation and (b) reactor and online sampling.

3.2.1. Automated Reagent Handling

On-demand reagent sampling was accomplished using an automated liquid handling robot (Gilson GX-271, Gilson, Inc.) controlled with LabView software. A syringe pump (Harvard PhD 2000) with a 100 μL glass syringe was connected to the liquid handler probe (needle) by approximately 50 cm of 500 μm PFA tubing filled with a selected transport fluid. To prepare a slug, the liquid handler aspirated first a 30 μL volume of inert gas, followed by aliquots of the chosen solvent, desired reactants, and the chosen solvent again. To minimize carryover during this process, the liquid handler probe was dipped in a wash solution before each reagent aspiration (the choice of wash solution depended upon the chemistry; both *i*PrOH and THF were used in subsequent studies). Approximately 35 μL total liquid volume was aspirated. Following sample aspiration, the sample was "stirred" three times in the probe under inert conditions by pulling and pushing with the syringe pump 30 μL volume. All reagents were then transferred into a 6 port-2 way injection valve (Cheminert 10S-0503H, Valco Instruments Co. Inc.) containing a 14-μL sample loop. Switching of the sample loop to the inject position created a 14-μL slug.

As reagent carryover from one slug to the next had the potential to significantly and adversely affect the accuracy of the optimization, we included in the sampling procedure the preparation of three blank slugs before every on-demand slug preparation. These served to clean the probe, injection valve, sample loop, and reaction system of any residue from prior experiments. The compositions of the three blank slugs changed depending upon the chemistry. In all cases, water was used as the first blank slug to dissolve any inorganic material; subsequent slugs comprised either THF and DMF or acetone and THF (in cases where the reaction solvent was fixed, use of that solvent as the final blank slug was advantageous in case any blank material became deposited on the reactor walls). In sequence, the liquid handler aspirated 20 μL inert gas, followed by 60 μL of each wash solvent with injection. The sample injection valve and sample loop were cleared following every injection by pulling a vacuum for 3 seconds on the outlet of the valve. This was accomplished by switching on and off a solenoid 3-way valve (P/N 01540-11, Cole-Parmer Instrument Company, LLC) connecting the house vacuum to a trap connected to the outlet of the valve. Residual liquid drained from the trap at the end of the optimization. Additionally, a 6-port, 2-way valve (Rheodyne MXP7960-000, IDEX Health & Science LLC) installed on the transfer line between the syringe pump and the probe was used to refill the transfer line with 80 μL fresh transfer fluid after every on-demand slug preparation. The line was

connected through an inline degasser (Agilent G1379B μ-degasser) to a supply tank of transfer fluid. This was found to both help minimize reagent carryover (from the material transferred to the transfer fluid during aspiration and stirring) and reduce the frequency of gas bubble formation in the transfer line. Gas bubbles in the transfer line were often problematic to ensuring good sampling accuracy. To correct for any gas bubble formation resulting from the probe sitting idle, the transfer line purge was repeated three times at the start of any optimization.

Accurate sampling also required knowledge of the physics of the liquid handling system. Because this particular liquid handler was an air displacement model, the amount of volume aspirated by the liquid handler changed as a function of reagent density. Ordinarily this would not have presented a problem in sampling only one solution at a time or even multiple solutions in the same solvent. However, in the studies presented herein it was common to sample multiple reagents with very different densities, which required a correction factor be introduced on a per reagent basis. Assuming that the gas volume was much greater than the liquid volume, we estimated that the change in gas bubble volume and sample volume upon upward movement of the transfer fluid could be expressed through a pressure balance as:

$$\rho_{tf} g_c \frac{\Delta V_{tf}}{A} = \rho_g g_c \frac{\Delta V_g}{A} + \rho_1 g_c \frac{\Delta V_1}{A} \tag{3.2}$$

where ρ_{tf}, ρ_g, and ρ_1 were the densities of the transfer fluid, gas, and sample 1, ΔV_{tf}, ΔV_g, and ΔV_1 were the respective volume changes upon aspiration, and A and g_c were the cross-sectional area and the gravitational constant, respectively. With ρ_{tf} and $\rho_1 >> \rho_g$, we arrived at the relationship:

$$\frac{\Delta V_1}{\Delta V_{tf}} = \frac{\rho_{tf}}{\rho_1} \tag{3.3}$$

which is well-known for the calibration of micropipettes.[160] By the same analysis, the relative aspirated volumes of samples 1 and 2 was expressed as:

$$\frac{\Delta V_1}{\Delta V_2} = \frac{\rho_2}{\rho_1} \tag{3.4}$$

Hence the relative densities of the two reagents were needed to accurately sample reagent 1 relative to reagent 2.

The minimum volume of a prepared slug was restricted both by the sampling accuracy and the dead volume of the sample injection valve. At approximately 0.1% of the syringe volume, the minimum sample accuracy was roughly 100 nL; thus reagent sampling of less than 1 μL implied

greater than 10% error in the sample. Generally it was undesirable for quantitation to aspirate reagent samples of less than 3 μL, though given that the target range for optimizations was generally a factor of 5 (for instance 0.5% to 2.5% catalyst) sometimes as little as 2 μL reagent was sampled with the acceptance that at the low end of the optimization this implied 5% error. The total sample volume aspirated needed to fill both the dead volume of the injection valve (estimated as 15-20 μL) and the 14-μL sample loop; hence nominally 35 μL of sample was aspirated. As many as six reagents were mixed in the liquid handler for a single slug.

The ability to keep samples under inert atmosphere and relatively free of evaporation was an important aspect to our optimization system, particularly for the case of catalyst screening. With Patrick Heider, we designed a 3D-printed manifold that allowed for an inert gas blanket to be maintained over air-sensitive reagents. The manifold is depicted in illustration and in application in Figure 3.4(a-b), with SOLIDWORKS (Dassault Systèmes) design specifications included in Appendix B. Screw threads on the underside of the manifold allowed vials containing reagents in solution to tighten against 15 mm PTFE-lined red rubber septa (W240594SP, Wheaton). A void space was left open above the vials through which inert gas (nitrogen or argon) at low pressure was supplied. 15 mm PTFE-lined red rubber septa were then inserted above the void space to seal the top of the manifold. We found much greater success with the rubber septa than with PTFE-lined silicone, as once a silicone septum was punctured, repeated exposure to organic solvents would cause degradation of the silicone and introduce the risk of clogging the probe. The red rubber septa were reusable for well over 100 injections before requiring replacement. The manifold itself was printed by Solid Concepts, Inc., from PEEK HP3, which has excellent chemical compatibility with nearly all organic reagents. For reactions in THF, the evaporation of THF with such a device was of concern to the accuracy of our method. We found that reducing the flow of inert gas with a bleed valve helped greatly to reduce evaporation, as did filling the vial closest to the inert gas inlet with neat THF (such that the atmosphere inside of the manifold would be THF saturated). Nonetheless, we were only able to run experiments for ~48 hr before observing significant losses in accuracy on account of evaporation.

We did not test the long-term stability of air-sensitive catalysts in this design. Though we observed stable performance for Suzuki-Miyaura cross-coupling reactions over a few days, the catalysts used in the study were known to be very stable in solution already and stable in air before activation. The manifold technique assisted greatly in facilitating engineering and

troubleshooting of the rest of the system. However, it would be anticipated that now that a system design has been established, a more generalizable solution may be to store the liquid handler and reagents in an enclosed glove box or glove bag. This may also help streamline catalyst preparation, without the need to degas individual solution vials.

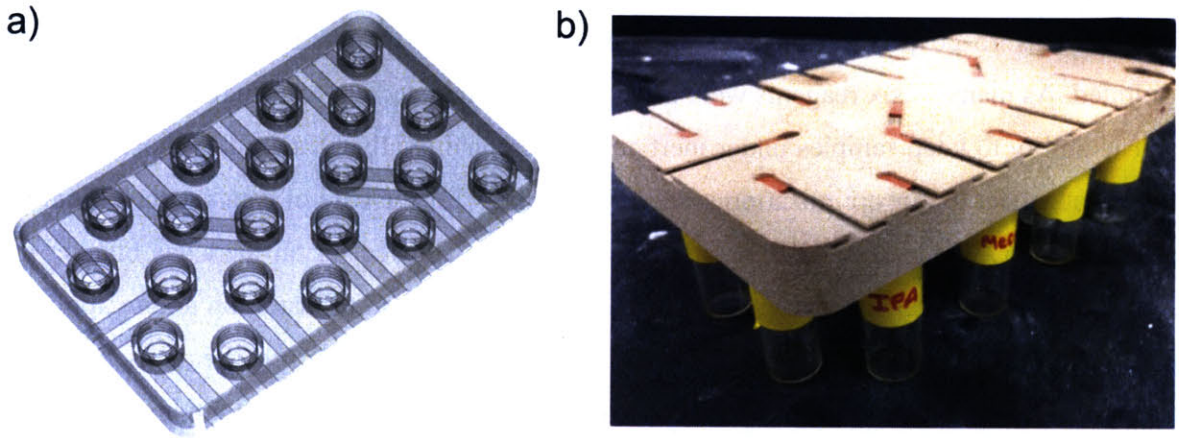


Figure 3.4. Septum-sealed inert gas manifold for reagent storage under inert gas atmosphere, (a) SOLIDWORKS rendering and (b) photograph of 3D-printed device.

3.2.2. Slug Transport and Reaction

We considered several strategies and combinations of materials for slug transport, including methodologies for organic slugs dispersed in an aqueous carrier phase, organic or aqueous slugs dispersed in a fluorinated carrier phase, and organic or aqueous slugs transported by a gas phase. Though the former two strategies targeted minimization of reagent carryover based on the preferential wetting of the carrier phase relative to the reaction slugs and allowed for more straightforward implementation because of the incompressibility of the transport phase, we concluded in both cases that the miscibility of reagents in either liquid-liquid system presented a problem to accurate representation of batch kinetics. With an aqueous-organic system, it was well understood that the partitioning of reagents between both phases would lead to inaccurate and generally unpredictable reagent compositions within the organic slugs. This was thought to be much less significant in fluorinated oil-organic systems, but recent studies have shown that mixing of nonpolar organic reagents such as toluene and THF into fluorinated oils can be appreciable even at room temperature,[161,162] with the miscibility increasing with temperature.[163] The problems associated with reagent solubility in fluorinated oils have led to more in depth investigations into the accuracy of the Accendo Conjure system for screening Suzuki-Miyaura

cross-coupling reactions.[158] Empirically, we observed 14 μL THF slugs, for example, to completely dissolve in FC-70 and FC-3283 at temperatures above 80°C. Though most limiting, miscibility was not the only negative factor brought about by the use of fluoroinated oils in the screening of organic reactions. Fluorinated compounds are also excellent solvents of oxygen, meaning considerable degassing was required for their use (not to mention polymers such as Teflon AF are incompatible with perfluorinated reagents), and the immiscibility of these compounds in aqueous environments caused nearly immediate damage to HPLC performance, requiring backflushing of the column with THF for removal. Too fast an acceleration of flow rate also caused "budding," or the breaking off of pieces of slugs into the fluorinated oil (as was also observed by Hawbaker *et al.*[158]). The density difference of most fluorinated oils relative to organic solvents limited the maximum diameter of tubing that could be used to 500 μm before gravity forces became comparable to surface tension forces.

The use of inert gases such as nitrogen or argon avoided all of the limitations listed above, though wetting, gas compressibility, and gas permeability all became new problems to address. We observed that in the uses of stainless steel, PEEK, and even PFA tubing led to significant wetting for solvents such as THF, which in turn led to degradation of the slug along the tubing wall. PDMS was not considered because of its incompatibility with organic reagents. Considerable improvement was observed with 750 μm inner diameter FEP as the tubing material. Slugs were transported through the FEP tubing by an 8 μL stainless steel syringe (Harvard Apparatus) containing 6.9 bar (100 psi) gas driven by syringe pump (Harvard PhD 2000). The compressibility of the gas mandated that for steady flow rates to be achieved, constrictions in the flow path had to be limited to no less than 500 μm and the pressure be maintained at or above 6.9 bar. A check valve was installed upstream of slug preparation and injection to dampen the effect of pressure oscillations further. Sample loops and unions upstream of the reactor were made from Teflon to ensure as little carryover as possible in the system. More consistent gas-liquid flow was observed by use of a 1 mm inner diameter T-junction at the reaction quench. Pressure in the system was controlled at 6.9 bar with an inert gas-regulated Parr bomb, approximately 40 mL in volume. The bomb was drained during refill of the gas and quench syringes by automatically opening a 6-port, 2 way valve (Rheodyne MXT715-000). To minimize gas loss during regular system operation, 6.6 bar of backpressure (5.2 bar and 1.4 bar

backpressure regulators, Upchurch Scientific) was applied to the gas vent of the Parr bomb during regular system operation.

For reaction, the FEP tubing was inserted into a "pancake" reactor housing (Figure 3.5(a-b)), comprising of an aluminum chuck with a 1.6 mm groove for the tubing, a raised lip with an O-ring, and a sheet of polycarbonate which compressed against the O-ring to allow for pressurization of the reactor to 6.9 bar. The custom reactor was designed in collaboration with Andrea Adamo, Baris Ünal, and Everett O'Neal. SOLIDWORKS design specifications are provided in Appendix B. With this device, we were able to rapidly heat and cool the reactor tubing between 30°C and 120°C and neutralize gas permeation out of the reactor—a factor which accounted for up to a 20% difference in residence time at high temperature. Residence times in the reactor were maintained between 1 min and 10 min (we empirically observed mixing to take place on the order of 1-10 s) at gas flow rates of 15-250 μL/min. A thermocouple was introduced through the gas supply line of the reactor and held in place on the aluminum surface by a thin sheet of polycarbonate and thermal paste. The reactor was heated with four 50 W cartridge heaters (McMaster-Carr Supply Company, two pairs spaced equally on opposite sides of the device). A PID temperature controller (OMEGA CN9412) controlled the reactor temperature. Reaction slugs were not introduced into the system unless the reactor temperature was within 1°C of the reaction set point temperature. Because acceleration of slugs was observed to occur when trailing slugs entered the heated reactor (a consequence of surface-tension driven thermocapillary flow[164]), blank slugs were not prepared and introduced until a reaction slug had traversed a full reactor volume in the system (240 μL).

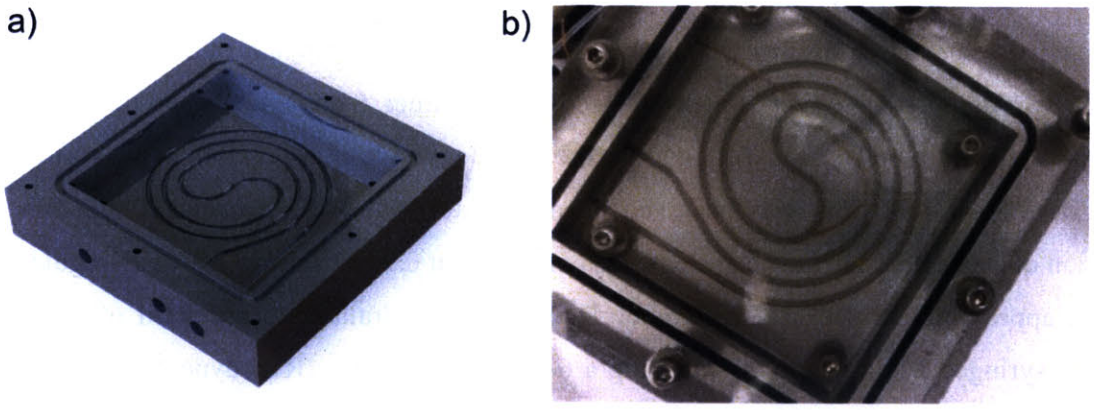


Figure 3.5. Pressure-sealed "pancake" reactor comprising a Teflon tube in an aluminum housing, (a) SOLIDWORKS rendering and (b) photograph of packaged device with polycarbonate cover, FEP tubing, and thermocouple.

3.2.3. Online Injection

Considerable effort was dedicated to enabling online injection of one or more activating reagents into flowing slugs. Having the capacity for online injection was essential to having accurate reaction time control in the optimizations, as few chemistries are completely inactive at room temperature. Though we adopted an approach similar that used Hatakeyama *et al.*,[96] where the online reagent was introduced via syringe pump into a slug through a T-junction, we found the repeatability of the online addition technique to be difficult to master without the proper combination of hardware and materials. Any compressibility—from gas or from leaks—in the online injection stream led to backflow from the main system into the normally stagnant injection line, causing inconsistent injections.

The online injection reagent was stored under an inert atmosphere and sampled by a 100 μL or a 250 μL glass syringe (Gastight, Hamilton Company) syringe driven by a syringe pump (Harvard PhD 2000). A 6-port, 2-way valve (Rheodyne MXP7960-000) was installed on the line to allow switching between refill of the syringe and online injection. Under normal continuous flow conditions, the valve was exclusively in the online injection position. Approximately 1 m of 250 μm inner diameter PFA tubing connected the valve to a T-junction (500 μm ID Teflon, Upchurch Scientific), which intersected the main system 6 cm before the reactor inlet. As a slug passed through the T-junction, the syringe pump infused a volume of 2-10 μL online injection reagent into the slug. The flow rate of the injection was chosen such that the volume was infused while 80% of the slug passed through the T-junction. Refractive index sensors (EE-SPX613, Omron Corporation) were attached to the Teflon tubing before and after the T-junction to correctly time the online reagent injection and to verify that the slug volume was within an acceptable tolerance (±4.0 μL) following the online injection. This was intended as a verification that no gas bubbles were introduced into the slug and that the slug had not broken apart upstream.

Leaks and gas bubbles were consistent inhibitors of reliable online injection. To ensure accuracy to 100-250 nL, the only chemically-suitable commercial options available for reagent dosing were glass syringes, which were found to leak over the course of a few optimizations with exposure to 6.9 bar pressure. The connection of the glass syringe to Teflon tubing was also often a source of leaks. We found the use of female Luer connectors sold by Upchurch Scientific to be most leak resistant, in comparison to the use of ¼-28-Luer adapters. Recently Hamilton began

selling ¼-28 male threaded syringes which may be even more leak free, but that was near the tail end of this thesis. Use of the 6-port, 2-way valve for refilling was mandatory compared to a 3-way solenoid valve because of the smaller dead volume and better pressure resistance. To avoid the introduction of gas bubbles into the injection line, the entire refill line was purged when new stock solutions were introduced by detaching and refilling the injection syringe. Naturally it was desirable to make the connection between the syringe and the T-junction as short as possible. By using a Teflon T-junction, less sticking and carryover of reagents were observed. A 500 μm T-junction was also required to prevent injected material or segments of slugs from becoming trapped in the swept volume of the T.

3.2.4. Quenching and Online Analysis

Downstream of the reactor, slugs were quenched at room temperature with a continuously flowing solution delivered *via* syringe pump (Harvard Apparatus PhD 2000 with 8 mL Harvard stainless steel syringe) through a T-junction (1 mm ID Teflon, Upchurch Scientific). A third refractive index sensor was used downstream of the quench to time the HPLC sampling accurately. Following sampling with a 30 μL sample loop in a 6-port, 2-way valve (Rheodyne MXP7960-000), the sample was transported *via* syringe pump (Harvard Apparatus PhD 2000 with 1 mL Hamilton Gastight syringe) containing the quench solution to a second 6-port, 2-way valve (Agilent G1158A) with a 1 μL PPEK sample loop. An LC/MS (Agilent G1312B binary pump, G1329B ALS, G1316A column compartment, G1365C multi-wavelength detector, 6120 quadrupole MS) method was remotely started with LabView software. The LC flow rate increased from 0.5 mL/min (standby mode) to 3.5 mL/min (required for method), and the sample was injected into the HPLC after 15 s. The sample passed through a 0.5 μm filter (Upchurch Scientific), then was heated to 40°C, passed through a T-junction (250 μm stainless steel Valco), and split by pressure difference between a 1.8 μm particle diameter column (Agilent Zorbax SB-C18 2.1 x 50 mm) and a 4.6 μm particle diameter column (Agilent Zorbax SB-C18 2.1 x 50 mm). The sample from the 1.8 μm particle diameter column was detected by UV and passed to the MS. A suitable HPLC method was found to be 9 min, which included a gradient ramp from 95/5 water/acetonitrile + 1% formic acid to 0/100 acetonitrile + 1% formic acid to 95/5 water/acetonitrile + 1% formic acid. Following UV analysis, the spectral baseline found by

subtraction of a reference and ChemStation outputted an Excel data file which was retrieved by LabView (ver. 8.6). The product yield was calculated in MATLAB (ver. R2011a).

3.2.5. Automation

Valve manipulation, HPLC method initiation, and analog input communication with the refractive index sensors were accomplished using a Compact FieldPoint controller from National Instruments (cFP-2020, cFP-RLY-425, cFP-AI-110). The entire system including pumps, liquid handler, temperature control, refractive index monitoring, valving, remote triggering of the HPLC gradient, and MATLAB optimization was controlled with LabView. With the exception of drivers, which were obtained from suppliers or from National Instruments, the LabView software was written completely in-house for the purpose of on-demand reaction screening. The LabView routine comprised a central VI ("Master VI.vi") that executed simultaneous loops for flow rate manipulation, reaction preparation, temperature control, online monitoring, HPLC sampling, HPLC analysis, and optimization. Data for individual slug experiments were recorded in a single matrix in MATLAB ("slug_tracker"). These data were recorded every 20 s in the form of text files documenting information about each slug experiment (concentrations, reagents, set temperature, and set reaction time), status of the experiments (time spent by slug in system and analysis and objective function value), system information (current flow and temperature conditions), and analysis data. Indices of LabView and MATLAB functions are provided in Appendix B.

3.3. EXPERIMENTAL

3.3.1. Reagent Stirring and Carryover

To test reagent stirring, a solution of 210 mg (0.20 M) 3-bromoquinoline in 5 mL DMF containing 80 mg (0.10 M) biphenyl and a solution of 80 mg (0.10 M) biphenyl in 5 mL DMF were stored in test tubes in the liquid handler. The liquid handler sampled from the biphenyl solution, the bromoquinoline/biphenyl solution, and the biphenyl solution again in differing relative amounts summing to 30 μL, with 30 μL air bubbles before and after. In one experiment the sample was injected immediately into the system as a slug, collected, and analyzed by HPLC. In a second experiment, the sample was stirred in the liquid handler probe by pushing and

pulling the liquid handler syringe via syringe pump multiple times, then the sample was injected into the system, collected, and analyzed by HPLC. The resulting calibrations were compared.

Multiple tests of reagent carryover were conducted. In a test of the effect of pinched tubing on reaction performance with a fluorinated carrier phase, the reaction in Scheme 2.1 was carried out in slug flow. Isolated solutions were prepared of 225 mg **1** (0.30 M) and 16 mg veratrole in 5 mL ethanol, 650 mg (1.50 M) **2** in 5 mL ethanol, and 750 mg (1.50 M) Et_3N in 5 mL ethanol. The solutions were stored in test tubes in the liquid handler along with a test tube of neat ethanol and sampled as described in Section 3.2 to produce 10 μL (stirred) slugs. The slugs were flowed at 14.3 μL/min through 244 μL unheated Teflon tubing, followed by 99 μL of tubing at 40°C, and finally 60 μL of unheated tubing. The slugs were collected at the outlet in a bath of isopropanol and analyzed by HPLC. Two DMF blank slugs were introduced after each analytical slug. 0.04" PEEK ¼-28 connectors (two T-junctions, one union) with flangeless ferrules were used for connections. Following the first set of experiments, the ends of the segments of Teflon tubing were cut, and the fittings were replaced with ¼-28 nuts with superflangeless ferrules. The sampling procedure was repeated. A list of slug reaction conditions for both sets of experiments is given in Table 3.1.

In a test of carryover in the gas flow system, neat 4-methoxybenzyl chloride and THF were stored under nitrogen in the liquid handler. A slug comprising 1 M 4-methoxybenzyl chloride in THF was prepared following the procedure described in Section 3.2, routed through the screening system (without online injection) at 30°C and a residence time of 5 min, and analyzed by LC/MS. Two subsequent slugs were then prepared with only THF as the reagent and routed through the system at the same flow rate and temperature and analyzed by LC/MS. The procedure was repeated 9 times.

Table 3.1. Slug experimental conditions for 2,4-dichloropyrimidine-morpholine reaction test of carryover.

Experiment	Conc. **1** in Slug (M)	Conc. **2** and Et_3N in Slug (M)
Flangeless Ferrules		
1	0.15	0.30
2	0.15	0.30
3	0.15	0.30
4	0.15	0.30
5	0.15	0.30
6	0.15	0.30
7	0.15	0.30
8	0.15	0.375
9	0.15	0.375
10	0.15	0.375
11	0.15	0.375
12	0.15	0.375
13	0.15	0.375
Superflangeless Ferrules		
1	0.15	0.0
2	0.15	0.0
3	0.15	0.15
4	0.15	0.15
5	0.15	0.375
6	0.15	0.375
7	0.15	0.375
8	0.15	0.30
9	0.15	0.30
10	0.15	0.30

3.3.2. "Pancake" Reactor Modeling

Considering the novelty of the "pancake" reactor design, we were interested in several factors affecting the heating of slugs within the Teflon tube reactor, namely the evenness of heating across the slug, the time for the slug temperature to equilibrate, and the temperature gradient across the aluminum chuck. To assess the heating of a slug, we constructed a 2D simulation of a microchannel cross section in COMSOL Multiphysics (Comsol Inc.) and simulated slug heating in time. The full cross section geometry is shown in Figure 3.6a, with a zoomed in profile of the microchannel shown in Figure 3.6b. The microchannel was assumed to be made of 750 μm FEP tubing carrying (stagnant) THF. To simulate entrance of the slug, a thin thermal boundary layer was assumed along the inner tubing wall with zero thermal conductivity before slug entrance and the thermal conductivity of FEP after slug entrance. The tube was situated in a small pocket of nitrogen against the aluminum chuck. A layer of polycarbonate was placed above the chuck, with ambient nitrogen and another layer of polycarbonate placed above the reactor surface. To model

ambient conditions, a pocket of air was added to the model above the outer polycarbonate surface. An isothermal 373.15 K boundary condition was applied to all sides of the aluminum surface, with the left and right edges of the polycarbonate, nitrogen, and air treated as periodic boundary conditions. A heat flux of 10 W/m^2 K was assumed at the outermost air boundary. Meshing was performed automatically by COMSOL. Up to 60 s were simulated in time at 0.2 s time steps.

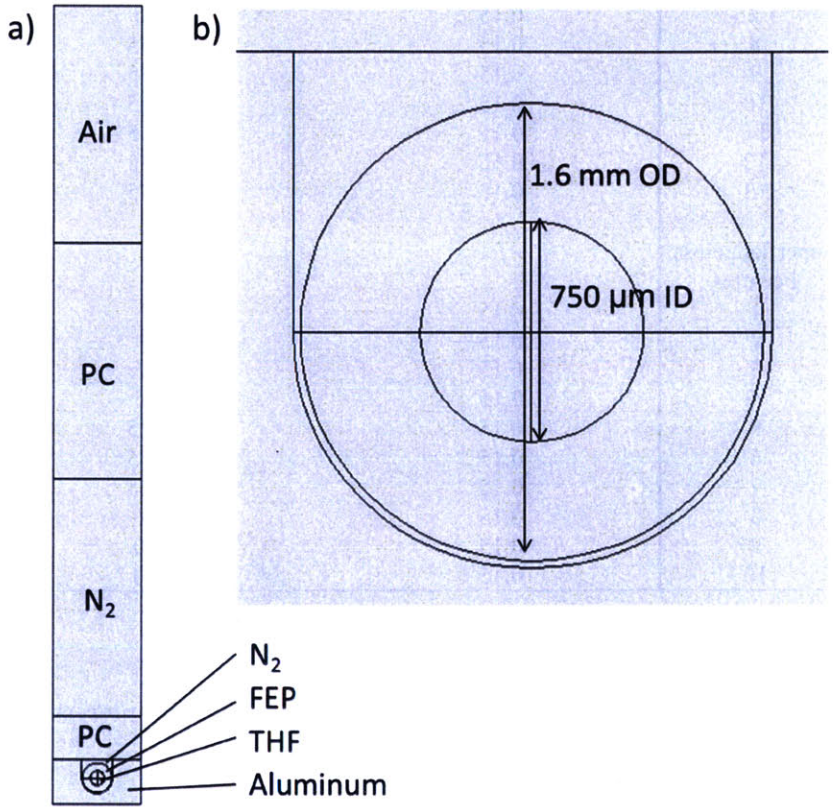
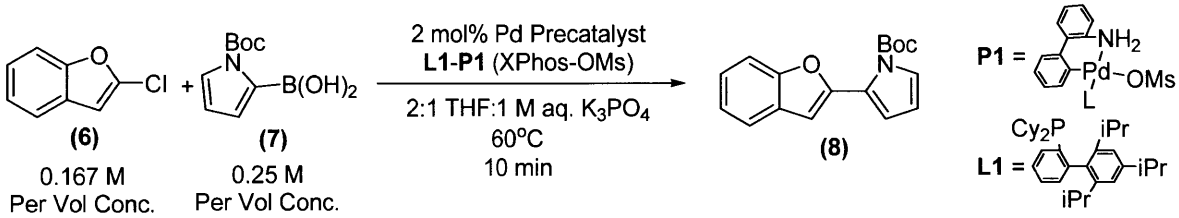


Figure 3.6. (a) Geometry studied for COMSOL simulations of the pancake reactor and (b) zoomed in view of the Teflon tube cross section. PC = polycarbonate.

Additionally, a 3D steady-state COMSOL simulation was constructed to model heat distribution throughout the aluminum reactor block. Cartridge heaters with a power input of 6 W were modeled in the simulation, which comprised a 10.3 cm x 10.3 cm x 2 cm aluminum block with raised lip and two layers of polycarbonate separated by a headspace of nitrogen. A heat flux of 10 W/m^2 K was assumed for all boundaries, aside from the insulated cartridge heaters, and radiation with a surface emissivity of 0.09 and ambient temperature of 293.15 K was assumed. A normal sized mesh was generated automatically by COMSOL.

3.3.3. Comparison of Reaction Yield in Batch and in Slugs

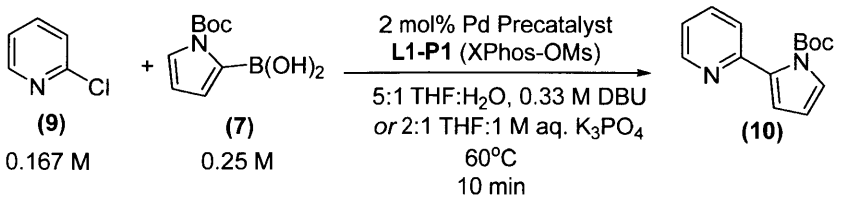
For validation of the scalability of our method, we compared the results of Suzuki-Miyaura cross-coupling reactions in batch to results from the segmented flow system. All reagents were used as received. We began in batch with the cross-coupling in Scheme 3.1. A 2 mL solution containing 187.1 mg 2-chlorobenzoxazole (**6**, 99%, Sigma-Aldrich), 378.5 mg 1-boc-2-pyrroleboronic acid (**7**, 98%, Frontier Scientific), 23.4 mg XPhos-OMs precatalyst (**L1-P1,** synthesized according to published procedure[165]), 25.7 mg naphthalene (99%, Sigma-Aldrich) as the internal standard, and THF (anhydrous, ≥99.9%, inhibitor-free, Sigma-Aldrich) was prepared under argon. To a 7 mL vial was charged 400 μL of the pre-made reagent solution, 600 μL THF, and 500 μL 1 M aq. K_3PO_4 (≥98%, Sigma-Aldrich), which had previously been degassed by sonication under vacuum and stored under argon. The reaction was stirred under argon at 60°C for 10 min, quenched with a 1:1 solution of acetonitrile and water, and analyzed by LC/MS. The product **8** was isolated by column chromatography.



(6) 0.167 M Per Vol Conc. **(7)** 0.25 M Per Vol Conc. **(8)**

Scheme 3.1. Suzuki-Miyaura cross-coupling of 2-chlorobenzoxazole and 1-boc-2-pyrroleboronic acid catalyzed by XPhos-OMs precatalyst.

Similarly, we studied in batch the single-phase cross-coupling in Scheme 3.2. A 2 mL solution containing 142.0 mg 2-chloropyridine (**9**, 99%, Sigma-Aldrich), 430.0 mg **7**, 22.6 mg **L1-P1**, 24.4 mg naphthalene as the internal standard, and THF was prepared under argon. To a 7 mL vial was charged 400 μL of the pre-made reagent solution, 600 μL THF, and 500 μL base. Two base solutions were tested: 1 M 1,8-diazabicycloundec-7-ene (DBU, 98%, Sigma-Aldrich) in 1:1 THF:H_2O and 1 M aq. K_3PO_4. In both cases the reaction was stirred under argon at 60°C for 10 min, quenched with a 1:1 solution of acetonitrile and water, and analyzed by LC/MS. The product **10** was isolated by column chromatography after using K_3PO_4 as base.



Scheme 3.2. Suzuki-Miyaura cross-coupling of 2-chloropyridine and 1-boc-2-pyrroleboronic acid catalyzed by XPhos-OMs precatalyst.

The reactions in Scheme 3.1 and 3.2 were repeated in the flow reaction system. A stock solution containing the aryl halide, boronic acid, precatalyst, and internal standard in THF was stored under argon in the liquid handler along with a vial of neat THF and a vial of degassed water. For Scheme 3.1, a 2 mL reagent solution was prepared containing 203.4 mg **6**, 388.4 mg **7**, 23.8 mg **L1-P1**, and 21.7 mg naphthalene in THF. 14 μL slugs were prepared by sampling the reagent solution and THF in a 1:1.5 ratio. For Scheme 3.2, a 5 mL reagent solution was prepared containing 135.5 mg **9**, 419.8 mg **7**, 94.0 mg **L1-P1**, and 47.3 mg naphthalene in THF. 14 μL slugs were prepared by sampling the reagent solution, THF, and water in a 1:0.875:0.625 ratio. Base solutions (1 M aq. K_3PO_4 or 1 M DBU in THF) were introduced online in a 1:2 ratio with the slugs, forming 21 μL reaction segments. The segments reacted at 60°C for 10 min with argon carrier gas and were analyzed by LC/MS.

3.4. Results and Discussion

3.4.1. Reagent Stirring and Carryover

Examining the results of reagent calibrations with and without reagent stirring (Figure 3.7), we observed substantial improvement in the system's ability to quantitatively dose reagents with the use of reagent stirring. Though all reagent concentration levels benefited in reproducibility and accuracy from stirring, in particular the accuracy improved the most with the most 3-bromoquinoline introduced in the sample. In general, the more stirring repeats, the more improved the accuracy; however no significant improvement was seen after the use of three back-and-forth stirring actions. This was reasoned optimal in an effort to conserve preparation time. Though faster flow rates also led to faster mixing, flow rates greater than 250 μL/min occasionally caused parts of the slug to break off in the transfer line, leading to cross-contamination and a worse calibration performance. Hence 250 μL/min was reasoned to be the optimal flow rate.

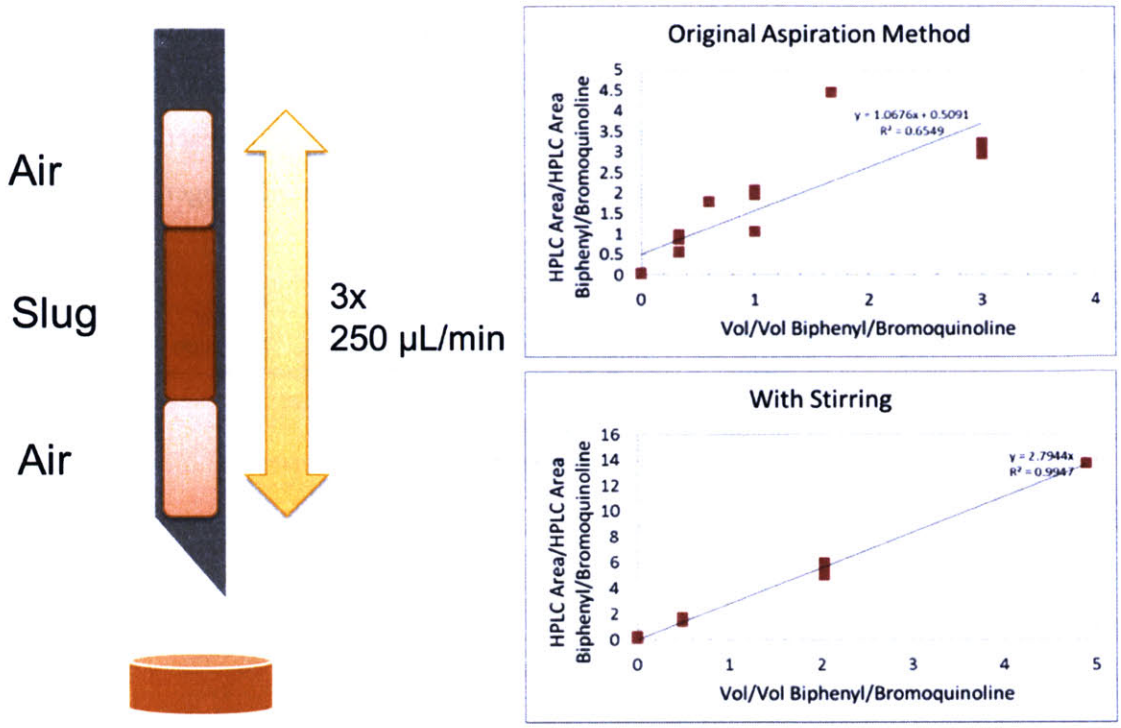


Figure 3.7. Schematic of reagent sampling and stirring and effect upon calibration reproducibility.

The effectiveness of stirring relates to the mixing rates of slugs, discussed in Section 3.1. By creating air or inert gas bubbles on either side of the sample slug, we were able to enhance mixing in the sample using the same convective mixing strategy that is employed in slug flow. This suggests an opportunity in HTE that to our knowledge has not been pursued—the controlled mixing of reactants in a needle, followed by direct sampling by HPLC. Though this presents a design challenge in terms of heating and pressure control, the ability to handle small volumes accurately may make this concept an idea worth exploring further.

Carryover of reagents was of great concern in design of this system and was tested intermittently at every stage of development. A key aspect of reaction carryover that we found not to be accounted for clearly in literature was the limitation in reaction kinetic accuracy that results from the use of blank slugs. Clearly blank slugs are beneficial in screening, as they prevent reagents or catalysts from diffusing backwards into upstream experimental slugs, potentially creating a false positive. However, the use of blank slugs alone for quantitative accuracy in HTE systems is insufficient for the reason illustrated in Figure 3.8a. Note that if a loss of material from an experimental slug is a regular enough occurrence to mandate the use of blank slugs, the loss of material from a blank slug into a following experimental slug should

happen approximately as often. This implies in the simplest case that a volume of reacting material within the experimental slug is replaced with a volume of blank material from an upstream slug, hence diluting the experimental slug. If obtaining accurate reaction kinetics or an accurate yield were the goal of the experiment, the diluted slug is no longer a model for the scaled-up batch reactor.

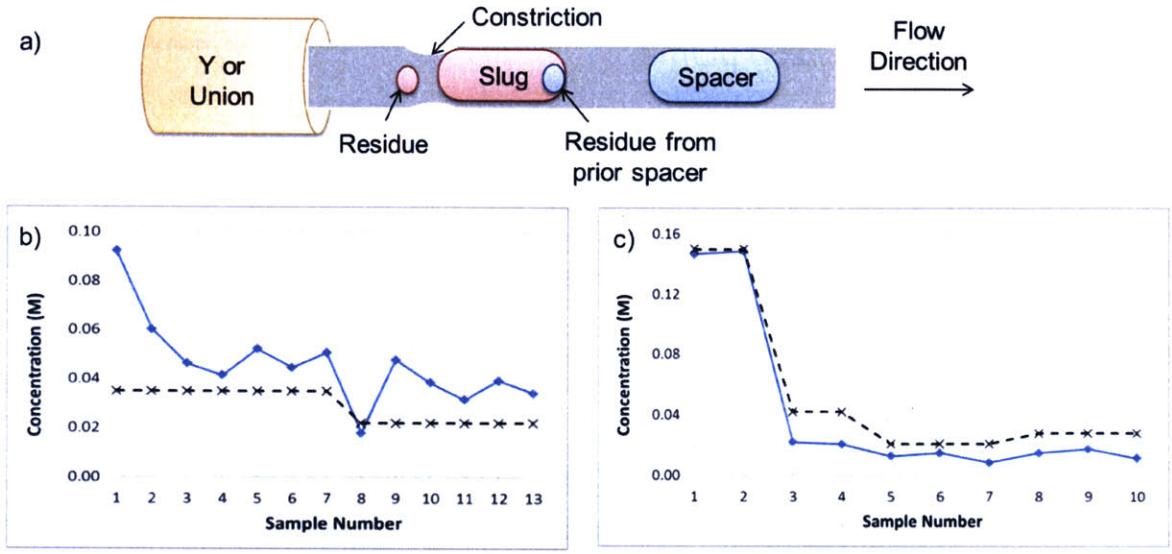


Figure 3.8. (a) Illustration of the effect of pinched tubing upon slug dilution upstream and the effect of (b) pinched and (c) new, un-pinched tubing upon reaction concentration and reproducibility. Black dotted line represents predicted starting material concentration and blue solid line represents observed. The excess conversion in (c) was attributed to not having a quench on the reactor outlet.

To address this problem, it was necessary to identify the source of reagent carryover in the automated system. Though factors such as tubing wetting, material left on the outside of the liquid handler probe or injection valve, and residual material in the HPLC sampling valve all contributed some to carryover (hence many of the system design features discussed in Section 3.2), we found the largest source of carryover to be in constrictions of the flow path. As shown in Figure 3.8a, our hypothesis was that slugs passing through flow constrictions were particularly susceptible to "budding" because of the pressure oscillation introduced across the length of the slug, and this resulted in small residues of slug material being left at tubing orifices. These residues were then exchanged with subsequent slugs. Two strategies were employed to minimize constrictions: first, connectors were chosen to most accurately match the inner diameter of the tubing; and second, flangeless fittings (which crimp onto the tubing) were replaced with superflangeless fittings (which have a metal ring that crimps onto the ferrule, leaving the tubing generally undamaged).

As evidenced by the data in Figure 3.8b and Figure 3.8 c, the reproducibility of results in switching from flangeless to superflangeless fittings was greatly improved. Additionally, the reaction conversion increased when superflangeless fittings were used. The data were compared against the kinetics for the reaction of **1** evaluated in Chapter 2. We observed increased conversion in all reaction cases with superflangeless fittings because of the lack of a quench and online injection control in the test system. Later experiments with the gas phase system demonstrated better the accuracy of the system with controlled reaction initiation/termination.

Finally, we considered in the gas phase system the amount of carryover with THF slugs. On average, the LC area for 4-methoxybenzyl chloride was 6578 in the 1 M "experimental" slugs, 216 in the first subsequent THF slugs, and 116 in the second subsequent THF slugs. The percent carryover was 3.3% ± 0.6% in the first THF slugs and 1.8% ± 0.6% in the second THF slugs with a maximum carryover in the first THF slugs of 3.9% (twice). With added flushing of the HPLC sample loop after this carryover experiment was performed, we believe the final system carryover may now be even less than the 3.3% reported.

3.4.2. COMSOL Simulations of Heat Transfer in the Pancake Reactor

The thermal profile within a reactor channel and across the reactor as a whole were modeled with COMSOL simulations. We generally observed the reactor temperature to rapidly equilibrate, with the formation of a modest temperature gradient across the Teflon tubing. The gradient formed as a result of the absence of a heat-conductive material above the Teflon tubing. Figure 3.9a illustrates that in the first 10 s of heating, this gradient was approximately 10°C from bottom to top of the tubing at 100°C; however fluid inside of the tubing only experienced a 3°C temperature gradient from top to bottom. The evolution of the temperature gradient within the microfluidic channel is shown in Figure 3.9b at 5 s, 10 s, 20 s, and 30 s time points. After 30 s, the fluid inside the channel was found to be uniformly within 2°C of the set reactor temperature. Thermal equilibration of the inner fluid was notably faster when the tubular reactor temperature had equilibrated to the set point temperature (Figure 3.9(c-d)). Hence, experimentally the reactor was required to equilibrate for a minimum of 30 s before introduction of a slug, at which point the time for the slug to heat to within 1°C of the reactor set point temperature was less than 10 s.

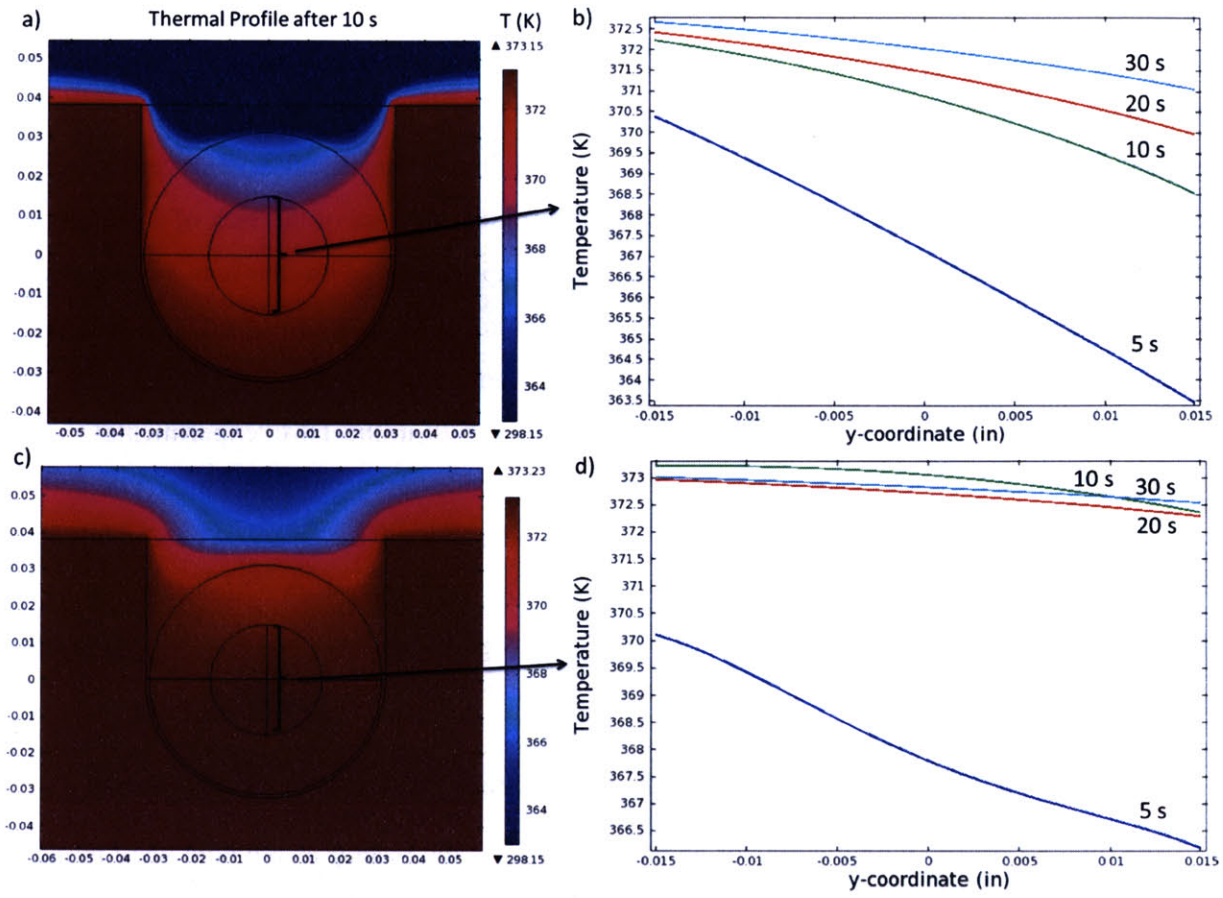


Figure 3.9. (a) Temperature profile for pancake reactor cross-section for 373.15 K aluminum block and 750 μm ID FEP tubing after 10 s. (b) Temperature of the center of the channel as a function of vertical position at 5 s (blue), 10 s (green), 20 s (red) and 30 s (aqua). (c) Temperature profile for pancake reactor cross-section for 373.15 K aluminum block and 750 μm ID FEP tubing after 10 s after allowing system to equilibrate for 30 s. (d) Temperature of the center of the channel as a function of vertical position at 5 s (blue), 10 s (green), 20 s (red) and 30 s (aqua) after 30 s of equilibration.

Steady-state COMSOL simulations of the reactor module (Figure 3.10(a-b)) allowed for validation of the assumption of an isothermal reactor surface and identification of the optimal position of the 50 W maximum cartridge heaters. We found that four cartridge heaters supplying heat in parallel were needed to distribute heat evenly across the reactor surface. Shown in Figure 3.10a, the placement of two cartridge heaters with 2.3 cm spacing on each opposite side of the reactor block allowed for the reaction area to be heated to within 1°C of the desired set point temperature. Illustrated in Figure 3.10b, the gradient across the center reactor cross section was maintained within 0.5°C. Note that although the reactor heated the thin layer of polycarbonate affixed to the reaction surface to within 1°C of set point, both the gas above the reaction surface and the outer polycarbonate layer remained much cooler—as cool as 44°C on the outer edge of

the device. We propose that further optimization of this design could allow for stackable reactors operating in parallel, as temperatures in each reactor can be controlled independently.

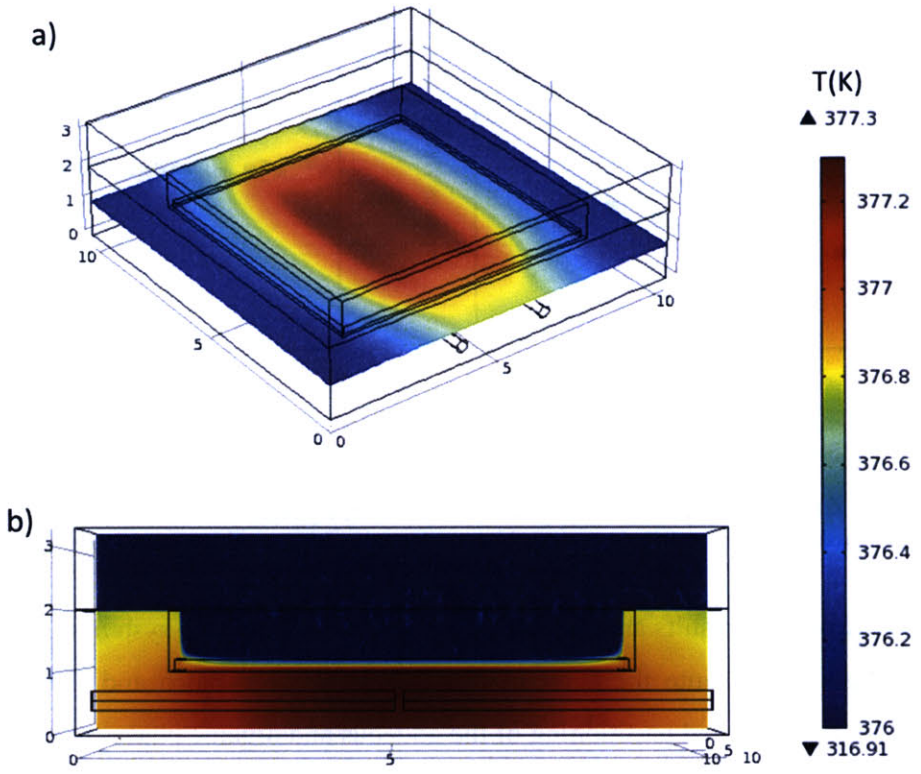


Figure 3.10. Pancake reactor temperature profile at 373.15 K for (a) surface of the aluminum chuck and (b) cross-section of the chuck with cartridge heaters and polycarbonate covers.

3.4.3. Comparison of Reaction Yield in Batch and in Slugs

The batch synthesis in Scheme 3.1 yielded 81% of the product **8** by HPLC with agreeable isolated yield. Attempts to make the product **8** at the same yield in flow were less successful, as evidenced by Figure 3.11a. Slug flow yields for the synthesis of **8** varied in the range of 30-50% despite the equivalent reaction time and temperature. A similar observation was made for the synthesis of **10** with K_3PO_4 as the base. Under batch conditions, 64% yield of **10** was observed by HPLC; however the reaction yield was less than half in the segmented flow system. By transitioning to single-phase reaction conditions with the use of DBU as the base, we were able to synthesize **10** at 57% yield, similar to the results found with K_3PO_4 as the base. Unlike in the use of an inorganic base, the use of the organic base DBU resulted in excellent agreement between batch and slug flow yields (Figure 3.11b) and much improved consistency from one

experiment to the next. Water was still found to be necessary for obtaining acceptable yields in the DBU reaction as a source of hydroxide.

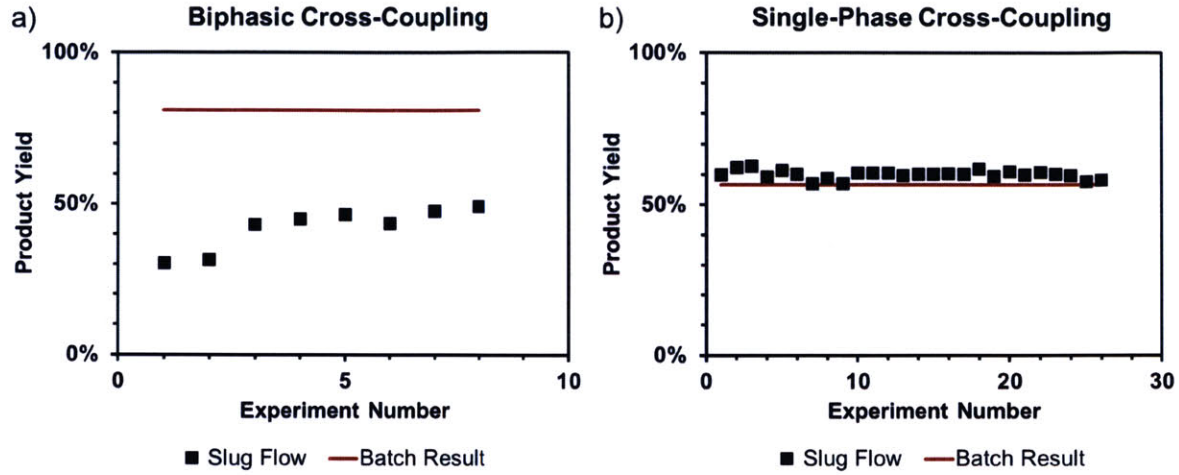


Figure 3.11. Comparison of slug flow yields to batch yield for (a) reaction of 2-chlorobenzoxazole and 1-boc-2-pyrroleboronic acid with aq. K_3PO_4 base and (b) reaction of 2-chloropyridine and 1-boc-2-pyrroleboronic acid with DBU base.

We reasoned from our results and from the physics of the slug flow system that mass transfer rates factored significantly into the agreement of yields between batch and flow syntheses. In biphasic systems, we observed partitioning of the aqueous and organic phases to opposite ends of the slug, creating only a single interface at which the reaction could occur. This occurred regardless of the injection rate of the online base stream. Although the surface tension forces contributed to mixing of the separate phases in these biphasic slugs, no mechanism was available to disrupt and regenerate the aqueous-organic interface. This interface regeneration occurs regularly in a stirred batch reactor or a flow system with static mixing elements, hence reducing the mass transfer limitation. We found static mixing elements to not be an option in our segmented flow system on account of the carryover introduced from breaking apart the slugs. The partitioning of the two phases at the initial point of online injection led to the inconsistency in yields in Figure 3.11a. For slugs where the aqueous and organic phases were more evenly distributed and in smaller segments before phase separation, the yield improved relative to having a single phase front-loaded or back-loaded in the slug.

By making the reaction medium homogeneous, the mixing across the slug was sufficient to match the yield observed in batch. This particular DBU system was found to be comparable to the more common THF/K_3PO_4 system in terms of yield and could be beneficial to continuous

flow systems in general because of the absence of mass transfer limitations. We did not have success with K^{t}BuOH as the base or with the DBU/water system with **7** (in the latter case **7** was observed to hydrolyze rapidly in the presence of water). Thus the DBU system was less general but still allowed for several Suzuki-Miyaura case studies to be explored in Chapter 6.

To improve mass transfer in the biphasic reaction segments, Abolhasani *et al.* successfully demonstrated the application of an oscillatory flow reactor (OFR).[166] In OFR screening, a single slug was moved forward and backward in a Teflon tube reactor by pushing and pulling on a gas-filled syringe, causing the aqueous phase to pass through the organic phase and hence improve mixing. This technology was applied to both Suzuki-Miyaura cross-coupling reactions and Buchwald-Hartwig amination reactions, with agreeable yields found with batch. It is expected in future work that this technology will be incorporated into the automated screening system presented in this chapter to expand the generality of the on-demand screening method from homogenous single-phase reactions to all liquid-phase reactions.

3.5. CONCLUSIONS

The literature of droplet and segmented flow screening systems is diverse in both application and design. However, the subset of systems designed for truly on-demand reaction screening is considerably more limited. The primary challenges presented in on-demand reaction preparation are versatility, both in the scope of chemistries to be studied and in the range of variables that can be accessed in a screen or optimization, and accuracy, which we found to be of particular concern in the use of liquid-liquid screening systems for reaction modeling. The packaging of maximal versatility and accuracy in a simple, easy to access automated platform presented a challenge within itself.

While undoubtedly leaving room for improvement, the system presented in this chapter makes great strides in reaction versatility and accuracy, all while allowing users a "set-it-and-forget-it" approach to reaction characterization. The scope of chemistries available to the system is to our knowledge unlimited, provided that reagents can be prepared into a single homogenous solution. In forthcoming chapters, we illustrate that the system performs equally well screening across a diverse array of organic solvents and generating accurate and reproducible results for palladium-catalyzed cross-coupling reactions. Reagent volumes are confined to 15-20 μL slugs, and the

amount of reaction material consumed per data point is reduced to 35:1 from the minimum 500:1 presented the continuous flow system in Chapter 2.

In continued work, it is expected that the incorporation of oscillatory slug flow into the system will enable enhanced mixing in biphasic liquid-liquid reactions, allowing these to be modeled as equally well as homogeneous chemistries. Better parallelization of reactor modules and streamlined HPLC injections will accelerate the reaction cycle time well beyond the current limitation of 10 min per experiment. Overall, on-demand reaction characterization will ultimately allow users access to reaction libraries, kinetics, and optimized conditions at the press of a button. Of course, all of these will require smarter computer software that compiles information about the current chemistry in the context of other related chemical systems, as a process chemist would draw upon in his or her reaction development work. To this eventual end, Chapter 4 presents our development of a black box methodology for optimizing reactions for both discrete variables and continuous variable conditions, offering a first step at solving such a complex experimental design problem.

4. AN ADAPTIVE RESPONSE SURFACE METHODOLOGY FOR OPTIMIZATION OF DISCRETE AND CONTINUOUS VARIABLE CHEMICAL SYSTEMS

4.1. INTRODUCTION

The optimization of chemical systems depends on manipulation of two classes of factors: continuous variables (such as temperature, reaction time, or concentration) and discrete or decision variables (such as choice of catalyst or solvent). Though often treated independently, it is difficult in most synthetic chemistry applications—and certainly in optimization—to separate discrete and continuous variable interactions. The best catalyst may depend on the loading at which it is most active, or the best solvent may depend on molecular interactions which change as a function of temperature or reagent composition.[167] The interplay of both discrete variables and continuous variables therefore represents an NP-hard MINLP which, in the most complex of cases, must be solved from kinetic models constructed for each distinct combination of discrete variable factors. In this chapter we do not venture to solve such a daunting problem; rather we ask whether a simple approach can be constructed for simultaneous optimization of discrete and continuous variables in the case where no mechanistic shift occurs, yet the substitution of discrete variables still has a significant effect upon reaction rate. Most importantly, no *a priori* model is made available for the reaction kinetics, as is common in early-stage reaction discovery and development, and the time and material spent optimizing for each discrete variable has to be minimized without compromising our method's ability to find the optimum.

In the absence of a model, the approximation of continuous variable interdependencies using response surface methodology (RSM) is a popular strategy, first introduced by Box and Hunter[168] and later the focus of many books and reviews.[169-174] In the simplest case, fractional or full factorial design experiments are used to tease out relationships among independent factors. More optimal designs can be constructed for estimation of interaction and quadratic effects of variables upon a response—the most common of these designs being central composite designs[175] and Box Behnken designs.[176] Though abundant in academia and in industry, these standardized designs have no built-in feedback methodology to increase resolution of an optimum and become less and less efficient as the complexity of the experimental and/or computational system increases.

To rapidly optimize more complex black box systems, Box and Draper introduced evolutionary operation (now more popularly referred to as sequential RSM),[177,178] for which the goal was to construct a response surface around a proposed optimum, test the optimum and model experimentally, and update the response surface in the subsequent iteration. Sequential RSM has emerged as an effective strategy in application to black box optimization problems.[171,173,179,180] To solve a complex optimization problem, sequential RSM is generally coupled with adaptive RSM (ARSM). In ARSM, the response surface is broken into subspaces, with regional optima compared against a common threshold to determine whether regions of the experimental space can be disregarded in the optimization.[181] Original ARSM algorithms formulated by merging central composite designs and Latin hypercube designs were proposed by Wang and coworkers.[182,183] To accelerate convergence, optimal design of experiments criteria such as minimization of the volume of the Fisher information matrix (D-optimality), minimization of the diagonal elements of the Fisher information matrix (A-optimality), or minimization of the variance on a specific point in the experimental design space (G-optimality)[184] can be incorporated into the sequential ARSM algorithm.

When integrated with a global search strategy such as branch-and-bound (B&B), sequential ARSM lends itself well to optimization over black box MINLPs.[185,186] In contrast to genetic and evolutionary algorithms,[113,187,188] ARSM is faster,[189,190] deterministic, and—as is advantageous in reaction kinetics—provides quantitative insight into the interaction of the response with each factor or combination of factors.[191] Much of the research in MINLP ARSM has focused on optimization of surrogate models, which are response surface models fitted in place of expensive computer simulations.[192-195] Several MINLP ARSM strategies have been published recently. For example, Holmström *et al.* developed an algorithm which treated all variables in an MINLP, including integer ones, as continuous variables and solved using a surrogate model with feasibility checks at each iteration.[196] Rashid *et al.* explored optimization using radial basis functions in MINLPs with nonlinear constraints.[197] Müller *et al.* further introduced SO-MI, which is a surrogate algorithm for global optimization for constrained MINLPs.[190] Studies that apply these methods to experimental situations data have been presented as well. In an application to groundwater management, Hemker *et al.* explored the advantages of B&B with surrogate models in comparison to genetic algorithms or implicit filtering.[189] For noisier problems, Davis and Ierapetritou introduced a kriging method with a B&B search that obtained

predictions and approximated variances at each iteration of the search. The authors illustrated application of the algorithm to chemical production in a continuous and binary variable optimization problem.[198]

To our knowledge, little experimental work has been conducted in the application of these strategies to experiments in chemical reaction kinetics, where interdependencies of discrete and continuous variables may often be too complex or simply too expensive to explore *via* construction of a complete reaction model. In this chapter, we demonstrate use of an adaptive response surface approach in conjunction with feedback optimization by G-optimality to identify the best discrete variable and reaction conditions, given the ability to conduct on-demand experimentation as outlined in Chapter 3. We then validate the generality of our optimization method, first through simulation, and then through experimental optimization of discrete (solvent, catalyst, and ligand) and continuous (temperature, reaction time, and concentration) variable selection in Chapters 5 and 6. We examine the ability of our method to optimize with a constraint placed on the reaction yield, a nonlinear function of the reaction turnover number (TON). Through simulation, we additionally demonstrate that in a high frequency of occurrence the method correctly identifies the optimum in a limited number of iterations regardless of a perturbation in the nominal reaction mechanism, making it suitable for cases of limited to no *a priori* reaction knowledge.

4.2. METHOD

4.2.1. Approach to Real-Time Discrete and Continuous Variable Optimization

We began with interest in optimizing the generalized chemical reaction A + B → R over N_{cv} continuous variables and N_{dv} discrete variables, which we proposed to be candidate catalysts for the synthesis of R. A general formulation of the MINLP for optimization of the chemical system was:

$$\begin{aligned} &\max_{\{\mathbf{x},\mathbf{y}\}} f(\mathbf{x},\mathbf{y}) \\ &\text{s.t. } g(\mathbf{x},\mathbf{y}) \le 0 \\ &\sum y_i = 1 \\ &x_j \in [-1,1] \quad \text{for } j = 1,\ldots,N_{cv} \\ &y_i \in \{0,1\} \quad \text{for } i = 1,\ldots,N_{dv} \end{aligned} \tag{4.1}$$

where $f(\mathbf{x},\mathbf{y})$ was the TON of the reaction—defined as moles product per mole of catalyst—and the constraint $g(\mathbf{x},\mathbf{y})$ was a constraint on the minimum yield (C_R/C_{A0}) at the optimum:

$$g(\mathbf{x},\mathbf{y}) = \gamma Y(\mathbf{x}',\mathbf{y}') - C_R / C_{A0}, \text{ where } Y(\mathbf{x}',\mathbf{y}') = \max_{\{\mathbf{x},\mathbf{y}\}} C_R / C_{A0} \quad (4.2)$$

$$\text{s.t.} \sum y_i = 1$$
$$x_j \in [-1,1] \quad \text{for } j = 1,\ldots,N_{cv}$$
$$y_i \in \{0,1\} \quad \text{for } i = 1,\ldots,N_{dv}$$

The parameter γ was adjustable in the range [0,1], with the choice of $\gamma = 1$ implying maximization with respect to yield and $\gamma = 0$ implying unconstrained maximization of TON.

The response surface procedure employed in this study comprised both an initialization phase and a feedback optimization phase. The decision diagram is shown in Figure 4.1. For initialization, two sets of fractional factorial design experiments were conducted with all candidate discrete variables. Each discrete variable was assigned a fractional factorial design that spanned the continuous variable experimental space. The sum of all fractional factorial designs minimally equaled a single-discrete-variable 2^{Ncv} full factorial design. In this regard, completion of the first fractional factorial designs allowed for calculation of a linear response surface with respect to $\mathbf{x}$ and $\mathbf{y}$ and identification of preliminary optima $J_i^* = \max[f(\mathbf{x},y_i)]$. A second set of fractional factorial designs were then chosen in the quadrants of the experimental space corresponding to the projected optimum for each discrete variable. These experiments allowed for estimation of the quadratic dependencies among continuous variables by virtue of sampling midpoints in the experimental design space.

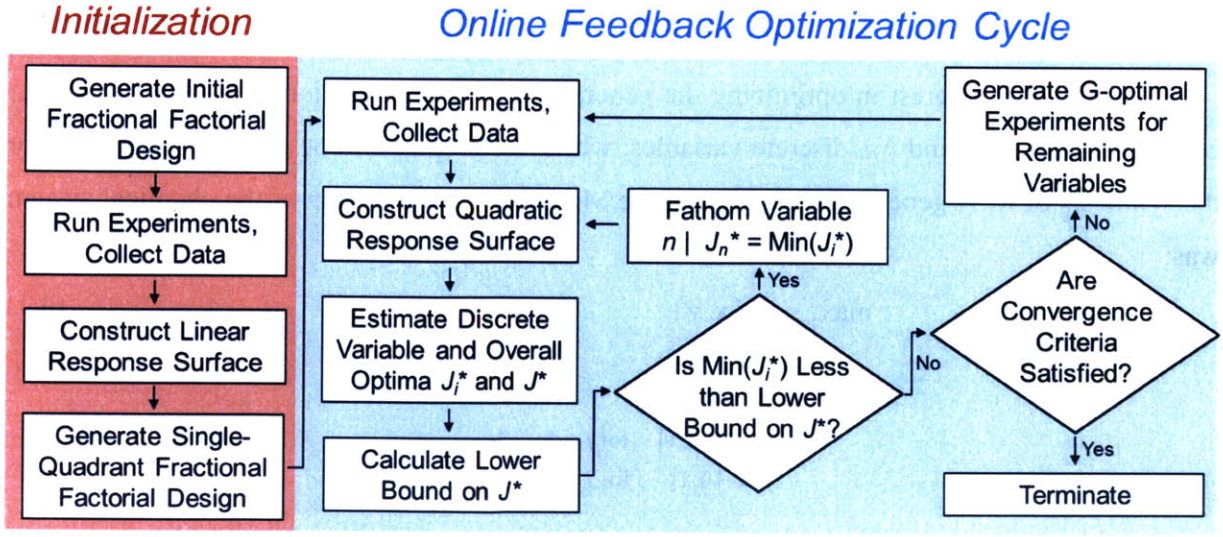


Figure 4.1. Real-time discrete and continuous variable optimization decision diagram.

Following initialization, response surfaces were iteratively developed for candidate discrete variables and discrete-variable-specific optima were predicted based on the response surfaces. Using a procedure akin to B&B, discrete variables whose performance was worse than the lower bound on the maximum of the leading discrete variable were fathomed from that iteration of the optimization, and response surfaces were recalculated using only experimental data from the remaining candidate discrete variables. Once a candidate set of discrete variables was determined, new experiments were chosen by G-optimality[173] and the procedure repeated until convergence to the final optimum.

4.2.2. Construction of Discrete Variable-Specific Response Surface Models

For the generalized bimolecular reaction A + B → R with constant $k_R = k_R$(catalyst type, C_{cat}, T), the production rate C_R/t_{res} was assumed to scale on the order of:

$$\frac{C_R}{t_{res}} \propto k_R\left(\text{catalyst}, C_{cat}, T\right) C_{A0}^m C_{B0}^n \tag{4.3}$$

For simplicity, it was then assumed that k_R(catalyst type, C_{cat}, T) could be separated into Arrhenius and catalyst-specific terms:

$$k_R\left(\text{catalyst}, C_{cat}, T\right) \propto \left(C_{cat}^p A_i e^{-E_{Ai}/RT}\right)\left(A_R e^{-E_{AR}/RT}\right) \tag{4.4}$$

giving an assumed scaling for C_R of:

$$C_R \propto A_i A_R e^{-(E_{Ai}+E_{AR})/RT} C_{A0}^m C_{B0}^n C_{cat}^p t_{res} \tag{4.5}$$

Taking the logarithm of all factors produced the linear relation:

$$\ln\left(C_R\right) \propto \ln\left(A_i\right) + \ln\left(A_R\right) - \frac{E_{Ai}}{R}\left(\frac{1}{T}\right) - \frac{E_{AR}}{R}\left(\frac{1}{T}\right) + m\ln\left(C_{A0}\right) + n\ln\left(C_{B0}\right) + p\ln\left(C_{cat}\right) + \ln\left(t_{res}\right) \tag{4.6}$$

This introduced a set of continuous factors to vary for each discrete variable: T^{-1}, $\ln(C_{i0})$, $\ln(C_{cat})$, and $\ln(t_{res})$. Naturally the assumptions leading to Equation 4.6 ignored the possibility of more complex kinetics, such as a Langmuir-Hinshelwood mechanism,[199] or the change in rate as the starting reagents A and B were consumed. To correct for inaccuracies in our assumed scaling, additional coefficients were introduced to weigh the $\ln(t_{res})$ term and account for interactions and quadratic functionality among all continuous variables. The final response surface model to fit was of the form:

$$\hat{b} = \sum_{i=1}^{N_{cat}} \left(c_i y_i + a_{i1} y_i x_1 \right) + \sum_{j=2}^{N_{cv}} c'_j x_j + \sum_{j=1}^{N_{cv}} \sum_{j'=j}^{N_{cv}} a'_{jj'} x_j x_{j'} \tag{4.7}$$

where $\hat{b}$ was the response value and a_{i1}, $a'_{jj'}$, c_i, and c'_j, were coefficients to fit.

Optimal coefficients $\boldsymbol{\theta}$ for the response surface model were found by weighted least-squares regression of the scaled experiments $\mathbf{X}$ (including the linear, interaction, and quadratic terms of discrete and continuous variables in Equation 4.7) and the vector of measured responses $\mathbf{b}$. We found empirically that the use of a weighting matrix $\mathbf{W}$ which biased the regression to most closely fit the response surface at points where the yield of R was greatest gave the most accurate optimization results. A convenient choice of $\mathbf{W}$ was thus the yield itself; giving:

$$\boldsymbol{\theta} = \left(\mathbf{X}^T \mathbf{W} \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{W} \mathbf{b}; \; \mathbf{W} = \begin{bmatrix} C_{R1}/C_{A01} & 0 & \cdots & 0 \\ 0 & C_{R2}/C_{A02} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & C_{RN_{expts}}/C_{A0N_{expts}} \end{bmatrix} \tag{4.8}$$

Incidentally the choice of TON as the weighting also gave good performance, but the rate of convergence in test cases was found to be slower than with the choice of yield as the weighting.

4.2.3. Optimization of Response Surface Models and Discrete Variable Fathoming

Optima for the each discrete variable were identified by converting Equation 4.7 to quadratic form and populating the matrices $\mathbf{A}$ and $\mathbf{c}$ with the optimal parameters $\boldsymbol{\theta}$:

$$\begin{aligned} J_i^* = \max_{[\mathbf{x} \;\; \mathbf{y}]} \hat{b} = & \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} \mathbf{A} \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix}^T + \mathbf{c} \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix}^T \\ & s.t. \; \ln \gamma + \ln \left[Y(\mathbf{x}', \mathbf{y}') \right] - \left[\hat{b} + \ln \left(C_{cat}/C_{A0} \right) \right] \le 0 \\ & x_j \in [-1, 1] \quad \text{for} \, j = 1, \ldots, N_{cv} \\ & y_i = 1 \\ & y_k = 0 \quad \text{if} \, k \ne i \end{aligned} \tag{4.9}$$

Calculation of $\ln[Y(\mathbf{x}',\mathbf{y}')]$ was performed using the same quadratic model for $\hat{b}$, with linear rescaling with respect to $\ln(C_{cat}/C_{A0})$. The overall maximum, J^*, corresponded to the maximum over all J_i^*.

To calculate the uncertainty on J^*, the prediction covariance $V_{\hat{B}}$ was estimated as:[126]

$$V_{\hat{B}} = \begin{bmatrix} \mathbf{x}^* & \mathbf{y}^* \end{bmatrix} \left(\mathbf{X}^T V_B^{-1} \mathbf{X} \right)^{-1} \begin{bmatrix} \mathbf{x}^* & \mathbf{y}^* \end{bmatrix}^T \tag{4.10}$$

With many experiments at or near the optimum, an estimate for the scalar response covariance V_B could have been obtained from the squared sum of residuals, as in Equation 2.12 in Chapter 2.

However with very few experiments, we observed significant bias in a squared sum of residuals estimate of V_B. To reduce the amount of bias introduced by the manner in which experiments were being selected, a jackknife resampling strategy was employed to generate the scalars V_{Bu}, which were the response covariance values estimated with experiment u removed from the data set:[200]

$$V_{B_u} = \frac{\left(\boldsymbol{\theta}_u{}^T \mathbf{X}'_u - \mathbf{b}_u\right)^T \left(\boldsymbol{\theta}_u{}^T \mathbf{X}'_u - \mathbf{b}_u\right)}{N_{expts} - N_{params} - 1} \tag{4.11}$$

$\mathbf{X}'_u$, $\mathbf{b}_u$, and $\boldsymbol{\theta}_u$ were the matrix of scaled experimental conditions, vector of responses, and best-fit response surface parameters calculated excluding experiment u, respectively. An overall estimate of V_B was then found by:

$$V_B = \frac{N_{expts}}{N_{expts} - 1} \sum_{u=1}^{N_{expts}} \left[V_{B_u}^{1/2} - \frac{1}{N_{\text{expts}}} \sum_{u=1}^{N_{expts}} V_{B_u}^{1/2} \right]^2 \tag{4.12}$$

and assumed to be uniform across all response surfaces. (Though this was clearly not the case, as the algorithm advanced and conducted more experiments closer to the predicted optima, V_B became more representative of the covariance near [**x*** **y***].)

Given an estimate for the response covariance, a lower bound on J^* was found from a Student's t-distribution and $V_{\hat{B}}$ evaluated at the optimum:[126]

$$J_-^* = J^* - \left(V_{\hat{B}}\right)^{1/2} \left(t_{1-\alpha, \nu = N_{expts} - N_{params}}\right) \tag{4.13}$$

α was chosen before experimentation as 0.05, corresponding to a 95% one-sided confidence level on the lower bound of J^*. For the least optimal discrete variable, a paired 2-sample t-test at 95% confidence revealed whether J_i^* was significantly less than the overall optimum J^*:

$$\begin{aligned}
&\text{Null Hypothesis: } H_0 = J^* - \min\left(J_i^*\right) = 0 \\
&H_a = J^* - \min\left(J_i^*\right) > 0 \\
&t_{stat} = \frac{J^* - \min\left(J_i^*\right)}{V_{\hat{B}}^{1/2}} > t_{crit} = t_{1-\alpha, \nu = N_{expts} - N_{params}} \\
&J^* - \left(V_{\hat{B}}^{1/2}\right)\left(t_{1-\alpha_{conf}, \nu = N_{expts} - N_{params}}\right) > \min\left(J_i^*\right) \\
&J_-^* > \min\left(J_i^*\right)
\end{aligned} \tag{4.14}$$

This assumed a constant $V_{\hat{B}}$ for both discrete variable optima. Rejection of the null hypothesis resulted in fathoming of discrete variable i from the current optimization step, and response surfaces were recalculated excluding y_i and any data points associated with variable i from the model (though in instances where the number of candidate experiments was less than $N_{params} + 1$, it was possible to leave variable i and associated data points in the model to advance the optimization). This process proceeded in a loop until all values of J_i* exceeded J_-*. These remaining discrete variables comprised the experimental set for which a new iteration of experiments was generated.

To accelerate reduction of the discrete variable space and simultaneously maximize the continuous variable information gained per experiment, new experiments were generated using a modified G-optimality criterion:

$$G_i = \min_{[\mathbf{x}\ \mathbf{y}]} \left[\mathbf{x}_i^*\ \ \mathbf{y}_i^*\right]\left(\mathbf{X}_1^T V_B^{-1} \mathbf{X}_1\right)^{-1}\left[\mathbf{x}_i^*\ \ \mathbf{y}_i^*\right]^T + \left[\mathbf{x}_i'\ \ \mathbf{y}_i'\right]\left(\mathbf{X}_1^T V_B^{-1} \mathbf{X}_1\right)^{-1}\left[\mathbf{x}_i'\ \ \mathbf{y}_i'\right]^T \qquad (4.15)$$

$$s.t.\ x_i \in [-1, 1]$$

$$y_i = 1$$

$$y_k = 0 \text{ if } i \neq k$$

$\mathbf{X}_1$ was the matrix $\mathbf{X}$ augmented to include the candidate experiment [$\mathbf{x}$ $\mathbf{y}$]. Equal weighting was assigned to minimizing the error in the yield optimum and to minimizing the error in the constrained TON optimum. [$\mathbf{x}_i$* $\mathbf{y}_i$*] was supplied as an initial guess to an SQP optimizer to generate G_i and the new G-optimal experimental conditions. These experiments were then executed, and new response surfaces and estimates for J_i* were estimated accordingly.

To achieve convergence, we specified that linear improvement had to be observed both in the predicted optimal TON (exp(J*)) and in the lower bound on the optimal TON (exp(J_-*)) to within 2% of the optimal value:

$$\text{let } J_q^* \text{ be the the optimum through iteration } q \qquad (4.16)$$

$$J_{pred,q}^* = \frac{N_{expts} \text{ between } q-2 \text{ and } q}{N_{expts} \text{ between } q-2 \text{ and } q-1}\left(J_{q-2}^* - J_{q-1}^*\right) + J_{q-2}^*$$

$$\text{if } \frac{\left|J_{pred,q}^* - J_q^*\right|}{J_q^*} \text{ and } \frac{\left|J_{pred,q-1}^* - J_{q-1}^*\right|}{J_{q-1}^*} < 0.02$$

$$\text{and } \frac{\left|J_{-pred,q}^* - J_{-q}^*\right|}{J_{-q}^*} \text{ and } \frac{\left|J_{-pred,q-1}^* - J_{-q-1}^*\right|}{J_{-q-1}^*} < 0.02, \text{ terminate}$$

Importantly, this criterion was independent of the number of remaining unfathomed solvents (meaning multiple optima could be obtained within the convergence tolerance) and independent of the scaling of f.

4.2.4. Real-Time Experimental Considerations

Some subtle changes were introduced to the algorithm in order to maximize the efficiency of conducting real-time experiments. During the sequential ARSM procedure, new optimal experimental conditions were calculated one experiment before the complete data set for a given iteration was collected. This enabled new experimentation to begin during analysis of the last experiment in an iteration, greatly minimizing experimental downtime. It was also a very realistic possibility in practice that zero product yield and TON would be observed, which in this algorithm's logarithmic coordinate system would have produced an undefined objective value. Through simulations, we found that assigning a yield of 0.1% to cases where the yield was in truth zero imposed sufficient penalty on the optimization method, yet did not interfere with prediction of the maximum. The detection limit by HPLC was also chosen to be within an order of magnitude of 0.1% of the maximum in experiments. Optimal experiments were grouped by temperature and randomized at each optimization iteration to minimize both experimental bias and the time required for temperature re-equilibration.

4.3. RESULTS

Using optimization functions written in MATLAB and provided for reference in Appendix C, we simulated the optimization of TON for batch reaction cases where catalyst selection perturbed the reaction activation barrier. The stoichiometry was A + B → R with the rate law:

$$\frac{dC_R}{dt} = k_R C_A C_B \tag{4.17}$$

$$k_R = C_{cat}^{0.5} A_0 e^{-(E_{A0}+E_{Ai})/RT} \tag{4.18}$$

$$A_R = 3.1\times10^7 \text{ L}^{0.5} \text{ mol}^{-1.5} \text{ s}^{-1} \tag{4.19}$$

$$E_{AR} = 55 \text{ kJ mol}^{-1} \tag{4.20}$$

We examined the algorithm's ability to discriminate among eight catalysts with respective values of E_{Ai} reported in Table 4.1. We assessed the optimization with respect to the continuous variables temperature (T = 30°C-110°C), reaction time (t_{res} = 1 min-10 min), and catalyst loading

(C_{cat} = 0.835 mM-4.175 mM), while C_{A0} and C_{B0} were maintained constant at 0.167 M and 0.250 M, respectively. An $8*2^{3-2}$ fractional factorial design was selected for the initialization, and a subsequent $8*2^{3-2}$ fractional factorial design was constructed based on the initial linear model. As the choice of a set of fractional factorial design experiments had the potential to bias the results of a single simulation, we generated 10 randomized designs and recorded the mean and standard deviation for each case study. Individual test case results have also been included in Appendix C.

Table 4.1. Simulated catalyst-specific activation energy correction factors (E_{Ai}) in kJ mol^{-1}.

Catalyst	Test Case 1	Test Case 2	Test Case 3	Test Case 4
1	0.0	0.0	0.0	-5 if T < 80°C -5 + 0.3*(T-80) if T ≥ 80°C
2	0.3	0.0	0.3	0.7
3	0.3	0.3	0.3	0.7
4	0.7	0.7	0.7	0.7
5	0.7	0.7	0.7	0.7
6	2.2	2.2	2.2	2.2
7	3.8	3.8	3.8	3.8
8	7.3	7.3	7.3	7.3

4.3.1. Test Case 1: The Effect of γ upon Convergence

We investigated the effect of manipulation of γ upon the accuracy and convergence speed of the MINLP method. Cases where the optima lay at the lower bound of C_{cat} (γ = 0.90) and inside the C_{cat} constraint (γ = 0.95 and 0.98) were considered. Overall, we observed excellent agreement in our method with the true optima, both in the selection of catalyst 1 as the optimal catalyst and in identification of the continuous variable optimal conditions. All simulations identified catalyst 1 as optimal. Optimization with γ = 0.90 (Table 4.2) identified the correct continuous variable conditions as optimal in all instances. For γ = 0.95 (Table 4.3), we observed one instance where the algorithm converged to a sub-optimal loading of catalyst 1 of 1.6 mM, whereas in all other instances the optimum was found between 1.305 mM and 1.357 mM (compared to the true value of 1.309 mM). With γ = 0.98 (Table 4.4), the optimum was found within 2% accuracy in 8 of 10 simulations, with one case of sub-optimal convergence resulting in over-estimation of the optimal catalyst loading and the other producing an underestimation. Convergence of the algorithm was fast (fewer than 100 simulated data points were required for all but two simulations) and improved slightly for the cases of γ = 0.95 and γ = 0.98 in comparison to γ = 0.90.

Maximal objective function values for the non-optimal discrete variables consistently under-estimated the true maximum for the discrete variable. In the majority of cases the error between

the true optimum and that estimated by the algorithm grew as the catalyst became less kinetically favorable. This was a consequence of the assumption that response surfaces simply shifted up or down for different discrete variable reaction rates, which was invalid when multiple reactions proceeded to 100% conversion. Consequently, the estimated response surfaces were also conservative in estimation of the yield constraint. Despite this over-simplification, our method qualitatively succeeded in ranking the catalysts in agreement with reaction rate, suggesting that the predicted optima for non-optimal discrete variables may still be useful in identifying common trends among classes of variables.

Table 4.2. Optimization results for $\gamma = 0.90$. $N_{expts} = 83 \pm 23$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	110.0	0.835	180.7	
2	10.0	110.0	0.835	177.5	
3	10.0	110.0	0.835	177.5	
4	10.0	110.0	1.065	139.0	
5	10.0	110.0	1.065	139.0	
6	10.0	110.0	2.732	54.2	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	110.0	0.835	180.9	0.3
2	10.0	110.0	0.935	161.0	12.5
3	10.0	110.0	0.938	160.2	9.2
4	10.0	110.0	1.265	120.6	13.0
5	10.0	110.0	1.406	109.6	16.7
6	10.0	110.0	3.591	41.9	8.6
7	10.0	110.0	4.041	27.7	4.5
8	9.8	110.0	4.175	13.6	3.4

Table 4.3. Optimization results for $\gamma = 0.95$. $N_{expts} = 66 \pm 6$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	110.0	1.309	119.5	
2	10.0	110.0	1.580	98.9	
3	10.0	110.0	1.580	98.9	
4	10.0	110.0	2.031	77.0	
5	10.0	110.0	2.031	77.0	
6	10.0	110.0	4.175	36.9	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	110.0	1.353	116.2	7.0
2	10.0	110.0	1.865	86.6	13.1
3	10.0	110.0	1.832	87.0	9.2
4	10.0	110.0	2.614	63.6	14.0
5	10.0	110.0	2.686	61.5	12.7
6	10.0	110.0	4.175	33.6	2.0
7	10.0	110.0	4.146	28.0	3.9
8	10.0	110.0	4.077	16.3	4.7

Table 4.4. Optimization results for $\gamma = 0.98$. $N_{expts} = 74 \pm 12$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	110.0	2.285	70.6	
2	10.0	110.0	2.759	58.5	
3	10.0	110.0	2.759	58.5	
4	10.0	110.0	3.546	45.5	
5	10.0	110.0	3.546	45.5	
6	10.0	110.0	4.175	36.9	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	110.0	2.262	71.7	5.4
2	10.0	110.0	2.970	55.1	6.1
3	10.0	110.0	3.149	52.2	6.8
4	10.0	110.0	4.175	38.6	0.2
5	10.0	110.0	4.120	39.2	1.9
6	10.0	110.0	4.175	35.1	2.2
7	10.0	110.0	4.140	31.7	2.1
8	10.0	110.0	4.065	17.1	4.5

Illustrated in Figure 4.2(a-c), the trajectory followed by the optimization agreed with general intuition—namely the system explored all corners of the experimental space to identify the overall optimal yield, then optimized along the axis of catalyst loading to resolve the yield constraint. For γ = 0.90 (Figure 4.2a), experiments were concentrated at the minimum value of C_{cat}; for γ = 0.95 (Figure 4.2b) and γ = 0.98 (Figure 4.2c) experiments progressed along the C_{cat} axis. As desired for the B&B style search, as the optimization progressed fewer experiments were run with less kinetically favorable catalysts, with experiments near the end of the sequential optimization run with exclusively catalyst 1. The elimination of fathomed discrete variables from the response surface model coincided with a more accurate representation of the performance of catalyst 1, hence improving resolution of the maximum yield and TON.

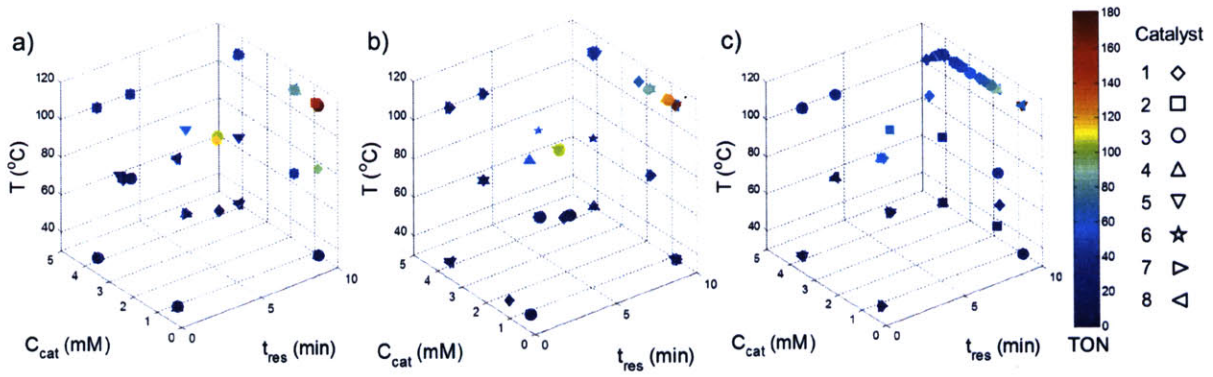


Figure 4.2. Optimization trajectory for (a) γ = 0.90, (b) γ = 0.95, and (c) γ = 0.98. Catalyst 1 is optimal in all cases with T = 110°C, t_{res} = 10 min, and C_{cat} = (a) 0.835 mM, (b) 1.311 mM, and (c) 2.507 mM.

4.3.2. Test Case 2: Optimization with Multiple Discrete Variable Optima

In kinetic problems where the rate is insensitive to the choice of discrete variable (for instance as a consequence of a change in the rate-limiting step), a suitable algorithm would have to identify both discrete variables as optimal. This test case hence presented a challenge to our B&B-inspired algorithm, as too fast of a discrimination among like discrete variables would imply a loss of information regarding the optimality of both catalysts 1 and 2. We considered for this case situations where the optimum for C_{cat} lay both on and off the constraint (γ = 0.90 and 0.95).

Shown in Table 4.5, we observed 100% success in our algorithm's ability to identify both 1 and 2 as co-optimal for γ = 0.90. The optimal TON found over 10 simulations agreed with the true value of 180.7 within a minimum standard deviation of 0.1 TON. In comparison to the first

test case with only a single optimal catalyst, the convergence rate for this case was considerably slower at 116 ± 39 iterations. This was perhaps counterintuitive, because with more discrete variables and associated experiments remaining in the experimental set, the number of degrees of freedom in the system increased, implying a lesser estimate of the variance V_B. In truth, however, the variance was generally greater for the optimization with two similar discrete variables because the optimal response surface had to capture the behaviors of twice the original number discrete variables using only an additional two model parameters.

For the case of $\gamma = 0.95$ (Table 4.6), our method was able to identify catalysts 1 and 2 as co-optimal in 9 of 10 cases. In all of the successful cases the optimal TON for both catalysts was within 4% of the true value and was more often underestimated than overestimated. In the lone unsuccessful case, catalyst 2 was identified with a maximum TON of 111, whereas the maximum TON for catalyst 1 was only estimated as 100. As in earlier test cases, the difficulty in resolving the constraint coincided with an over-approximation of the maximum reaction yield. Convergence was much faster for $\gamma = 0.95$, with the optimum found within 76 iterations on average. The simulated experimental trajectories shown in Figure 4.3(a-b) show that, as in the first test case, most experiments were conducted with the high-probability catalysts at the maximum temperature and time. For $\gamma = 0.90$ (Figure 4.3a), these experiments clustered more at the minimum catalyst loading, whereas for $\gamma = 0.95$ (Figure 4.3b) the simulated experiments were run with C_{cat} predominantly between 0.835 mM and 2.5 mM.

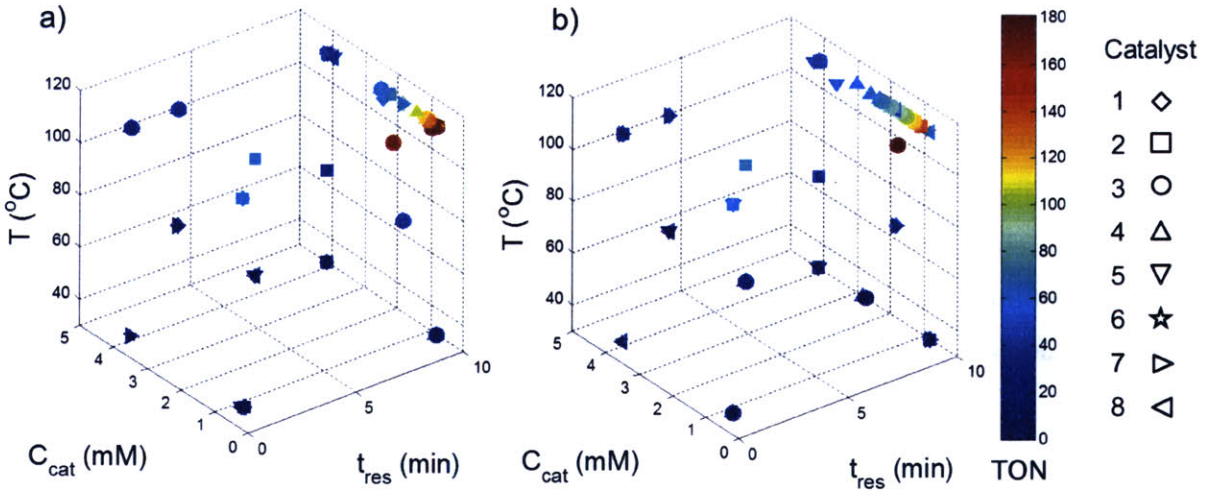


Figure 4.3. Optimization trajectory for the case of two co-optimal catalysts. (a) $\gamma = 0.90$: catalysts 1 and 2 are co-optimal at $T = 110$°C, $t_{res} = 10.0$ min, and $C_{cat} = 0.835$ mM. (b) $\gamma = 0.95$: catalysts 1 and 2 are co-optimal at $T = 110$°C, $t_{res} = 10.0$ min, and $C_{cat} = 1.320$-1.326 mM.

Table 4.5. Optimization results for $\gamma = 0.90$. $N_{expts} = 116 \pm 39$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	110.0	0.835	180.7	
2	10.0	110.0	0.835	180.7	
3	10.0	110.0	0.835	177.5	
4	10.0	110.0	1.065	139.0	
5	10.0	110.0	1.065	139.0	
6	10.0	110.0	2.732	54.2	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	110.0	0.835	180.8	0.2
2	10.0	110.0	0.835	180.8	0.1
3	10.0	110.0	0.875	170.4	5.0
4	10.0	110.0	1.240	123.0	17.1
5	10.0	110.0	1.217	125.7	18.3
6	9.9	110.0	3.825	38.5	9.8
7	9.8	110.0	4.079	30.3	4.2
8	9.8	110.0	3.946	18.6	4.0

Table 4.6. Optimization results for $\gamma = 0.95$. $N_{expts} = 76 \pm 13$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	110.0	1.309	119.5	
2	10.0	110.0	1.309	119.5	
3	10.0	110.0	1.580	98.9	
4	10.0	110.0	2.031	77.0	
5	10.0	110.0	2.031	77.0	
6	10.0	110.0	4.175	36.9	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	110.0	1.349	116.4	6.0
2	10.0	110.0	1.338	117.2	2.9
3	10.0	110.0	1.741	90.7	6.8
4	10.0	110.0	2.504	65.1	12.6
5	10.0	110.0	2.639	62.6	13.9
6	10.0	110.0	4.175	35.4	1.2
7	9.8	110.0	4.175	26.6	4.1
8	9.7	110.0	4.175	15.5	5.3

4.3.3. Test Case 3: Perturbations to the Reaction Pathway

In addition to the reaction of A + B → R with kinetics in Equations 4.17-4.20, we considered cases where species B or C may be consumed by alternative pathways over the course of the reaction. As a first case study, we introduced degradation of B as a parallel reaction step. The degradation pathway was represented as B → S_1 with kinetics:

$$\frac{dC_{S_1}}{dt} = k_{S1}C_B \tag{4.21}$$

$$k_{S1} = A_{S1}e^{-E_{AS1}/RT} \tag{4.22}$$

$$A_{S1} = 1.0\times10^{12}\ \text{s}^{-1} \tag{4.23}$$

$$E_{AS1} = 100\ \text{kJ mol}^{-1} \tag{4.24}$$

This pathway was intended to mimic the effect of protodeboronation in Suzuki-Miyaura cross-coupling reactions, which often leads to significant degradation of the starting boronic acid at high temperatures. In this case we selected $\gamma = 0.90$.

From the results found in Table 4.7, we observed our algorithm to successfully identify both the optimal catalyst 1 and the internal temperature and catalyst loading optima in all cases. The optimization routine converged rapidly in an average of 66 simulated experiments. In 4 of the 10 cases, convergence was achieved in fewer than 64 iterations, meaning the algorithm found an internal constrained optimum in fewer experiments than would be required to run a 2^3 full factorial design for all 8 discrete variables. Good agreement was observed between the optimum found in the black box routine (33.5 ± 1.0 TON) and the true optimum (34.9 TON) given the limited number of experiments.

Likewise we considered the consumption of R in a series reaction, which was expected to produce an internal optimum for reaction time instead of temperature. In addition to A + B → R, the reaction B + R → S_2 was introduced into the model with kinetics:

$$\frac{dC_{S_2}}{dt} = k_{S2}C_BC_R \tag{4.25}$$

$$k_{S2} = A_{S2}e^{-E_{AS2}/RT} \tag{4.26}$$

$$A_{S2} = 3.1\times10^{5}\ \text{L mol}^{-1}\ \text{s}^{-1} \tag{4.27}$$

$$E_{AS2} = 50\ \text{kJ mol}^{-1} \tag{4.28}$$

Again γ was chosen as 0.90. Figure 4.4 illustrates the response surface for this case for catalyst 1 at the optimal catalyst concentration of 2.66 mM. This example is challenging from a response-surface-based optimization perspective because of the ridge of near-optimal values that forms in the temperature-residence time plane, much akin to that which was observed for the series reaction in Chapter 2.

Table 4.7. Optimization results for parallel reactions A + B → R and B → S_1. N_{expts} = 66 ± 5.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	81.8	2.662	34.9	
2	10.0	81.9	3.262	28.4	
3	10.0	81.9	3.262	28.4	
4	10.0	82.0	4.175	22.1	
5	10.0	82.0	4.175	22.1	
6	10.0	83.0	4.175	16.7	
7	10.0	83.8	4.175	11.6	
8	10.0	85.9	4.175	4.4	
			Response Surface Optimum		
1	10.0	81.1	2.756	33.5	1.0
2	10.0	81.4	3.777	24.4	2.3
3	10.0	81.5	3.672	25.3	2.6
4	10.0	81.6	4.175	20.4	0.7
5	10.0	82.4	4.175	20.1	0.9
6	10.0	84.8	4.175	13.5	1.0
7	10.0	86.1	4.175	8.8	0.7
8	10.0	89.9	4.175	3.0	0.2

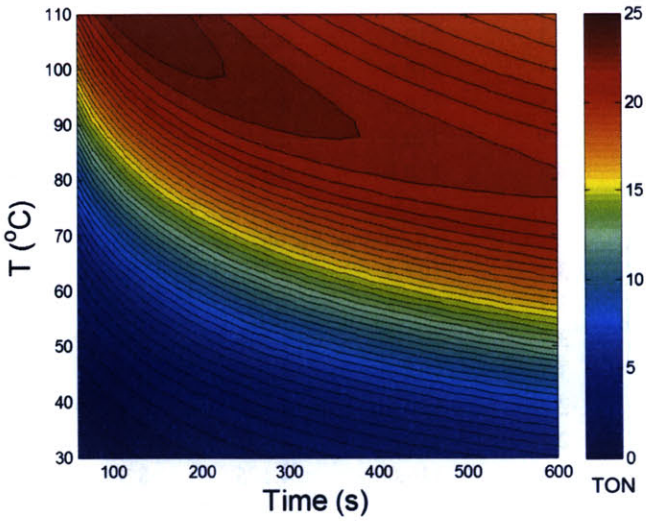


Figure 4.4. TON response surface for the optimal conditions of catalyst 1 at 2.66 mM for the series reactions A + B → R and B + R → S_2.

The optimization results presented in Table 4.8 are representative of the ridge in the response surface solution. Statistically, the mean optimal reaction time of 3.1 min is substantially different than the true optimum for this case at 2.1 min, but the mean optimal catalyst concentration and objective function value are comparable (23.9 for the true objective value as compared to 23.1 ± 2.5 for the ARSM algorithm). The optimal reaction time deviates significantly from the known optimum on account of two different local optima being identified by the algorithm: one optimum at t_{res} = 2.2-2.5 min and T = 110°C, and the second optimum at t_{res} = 10 min and T = 84.7°C. The best-fit response surfaces for either case are shown in Figure 4.5(a-b). Convergence to the T = 110°C optimum occurred in 9 of 10 cases, with an example response surface presented in Figure 4.5a, evidently in good agreement with the true response surface. In the case of convergence to the t_{res} = 10 min optimum, the best-fit response surface was in fact bimodal, with a flatter ridge of optimality than observed in Figures 4.4 and 4.5a. Because catalyst 1 was still identified as optimal in this case, we suspect that this limitation is a reflection of the limited robustness of the G-optimality criterion. More global sampling techniques will need to be incorporated into the algorithm in future adaptations in order to minimize the probability of local convergence.

Table 4.8. Optimization results for series reactions A + B → R and B + R → S_2. N_{expts} = 65 ± 6.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	2.1	110.0	2.665	23.9	
2	2.1	110.0	3.217	19.8	
3	2.1	110.0	3.217	19.8	
4	2.1	110.0	4.136	15.4	
5	2.1	110.0	4.136	15.4	
6	2.5	110.0	4.175	11.9	
7	2.9	110.0	4.175	8.8	
8	3.8	110.0	4.175	4.0	
			Response Surface Optimum		
1	3.1	107.4	2.789	23.1	2.5
2	3.2	107.5	3.263	19.3	1.1
3	3.2	107.5	3.319	19.0	1.2
4	3.1	107.5	4.143	14.8	0.5
5	2.3	110.0	4.109	15.0	0.9
6	2.3	110.0	4.175	11.3	0.4
7	2.4	110.0	4.175	8.3	0.2
8	2.4	110.0	4.175	3.7	0.3

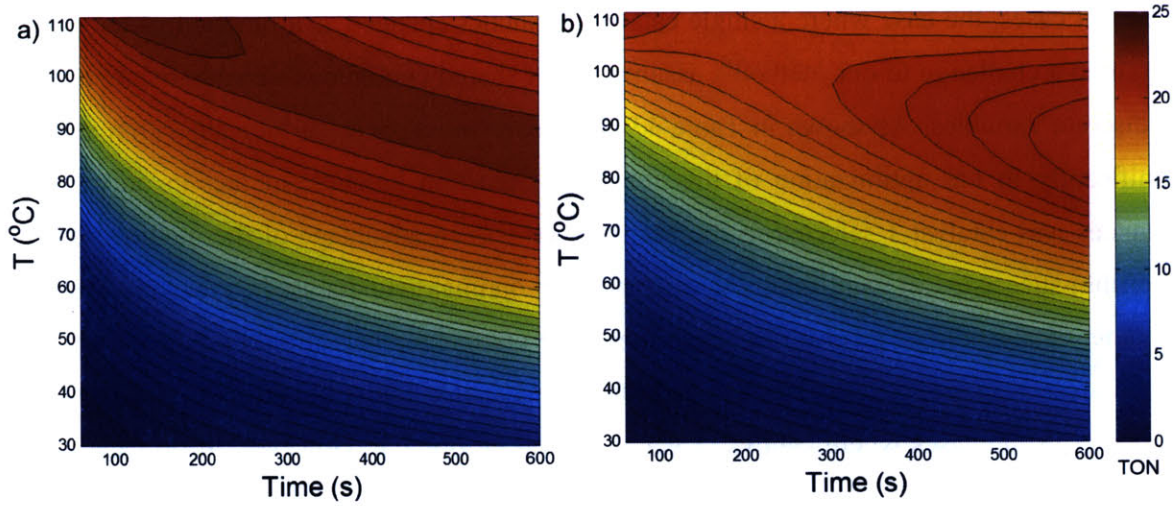


Figure 4.5. TON response surface for the optimal conditions of catalyst 1 at 2.66 mM for the series reactions A + B → R and B + R → S_2 for cases of (a) convergence to the global optimum and (b) convergence to a sub-optimal combination of temperature and reaction time.

Though limited in select test cases by local convergence, both of these examples illustrate the strength of this optimization approach in resolving optima when the reaction mechanism is perturbed. As seen in the case of series reactions, the response surfaces generated by the algorithm are locally consistent with the true response surface, despite the few assumptions made about the reaction model. In cases of process scale-up where it is be essential to know the effect on production of a perturbations to one or more process variables, this sequential ARSM algorithm provides insight in a limited number of experiments (on average 65-66 experiments for the case studies above). Local rate of reaction information can also be extracted by taking the derivative in time of the proposed response surface:

$$\frac{dC_R}{dt} = \sum_j \frac{\partial C_R}{\partial x_j} \frac{dx_j}{dt} \tag{4.29}$$

$$\ln\left(\frac{C_R}{C_{cat}}\right) = \mathbf{x}^T \mathbf{A}\mathbf{x} + \mathbf{c}^T \mathbf{x} \tag{4.30}$$

The correct choice of manipulated inputs would presumably allow mechanisms to be posed based on similarity to known rate law behavior.

4.3.4. Test Case 4: Catalyst 1 Deactivation at High Temperature

Catalyst deactivation can factor significantly into organometallic reactions, particularly in cases where the catalyst-ligand complex breaks down or changes ligated state. Unlike the first

three test cases, the case where a single catalyst's kinetics changed as a function of temperature presented a challenge to our method's assumption of shared continuous variable behavior across all discrete variables. As shown in Table 4.9, the adaptive response surface algorithm struggled significantly in the optimization of this set of test reaction kinetics; in fact in only 2 of 10 instances was catalyst 1 selected as the best catalyst. As an advantage of the methodology, the algorithm succeeded in identifying the distinction in optimal temperature between catalyst 1 and the other catalysts in all case studies examined. The optimal temperature found by the algorithm for catalyst 1 ranged from 71°C-81°C, with the most accurate estimates of 79°C and 81°C occurring in the cases where catalyst 1 was found to be optimal. However, in no case was the optimal C_{cat} for catalyst 1 found to lie at the lower bound of 0.835 mM, as the maximal yield in the two optimal cases was estimated as greater than 100%.

Table 4.9. Optimization results for case of catalyst 1 deactivation at T > 80°C. N_{expts} = 151 ± 83.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	80.0	0.835	187.6	
2	10.0	110.0	1.161	128.7	
3	10.0	110.0	1.161	128.7	
4	10.0	110.0	1.161	128.7	
5	10.0	110.0	1.161	128.7	
6	10.0	110.0	2.976	50.2	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	76.6	2.228	75.4	29.2
2	10.0	110.0	1.091	142.3	25.7
3	10.0	110.0	1.100	142.3	27.8
4	10.0	110.0	1.102	142.7	26.4
5	10.0	110.0	1.100	141.5	26.9
6	10.0	110.0	3.761	42.5	15.8
7	9.9	110.0	4.096	31.5	3.9
8	9.8	110.0	4.132	16.3	5.0

We suspected that the poor accuracy of the algorithm in this case resulted from the assumption of a constant $V_{\hat{\beta}}$ across the experimental design space. A limitation of the G-optimal design objective is that the uncertainty of predictions away from the optimum is untested; hence no test was in place to assess the quality of the prediction of the optimum of catalyst 1 given the results for the other seven (sub-optimal) catalysts. We addressed this problem by proposing a

trust region for the prediction covariance. Rather than assuming a universal $V_{\hat{B}}$, a discrete-variable specific $V_{\hat{B}i}$ was calculated by incorporating into the matrix **X** in Equation 4.10 only experiments conducted with discrete variable *i* or within specified continuous variable bounds. Only the columns of **X** corresponding to continuous variable-continuous variable interactions and continuous variable-discrete variable *i* interactions were considered. Discrete variables satisfying the requisite number of degrees of freedom were then fathomed sequentially based on failure of an unpaired two-sample *t*-test at 95% confidence:

$$\text{Null Hypothesis: } H_0 = J_i^* - J_i = 0 \tag{4.31}$$

$$H_a = J_i^* - J_i > 0$$

$$t_{stat} = \frac{J_i^* - J_i}{\sqrt{V_{\hat{B}} + V_{\hat{B}i}}} > t_{crit} = t_{1-\alpha,\nu}$$

$$\nu = \frac{\left(V_{\hat{B}} + V_{\hat{B}i}\right)^2}{\dfrac{V_{\hat{B}}^2}{N_{expts} - N_{params}} + \dfrac{V_{\hat{B}i}^2}{\left(N_{expts} - N_{params}\right)_{\text{Reduced Model}}}}$$

We guessed initially trust region bounds of ±10% on the continuous variable factors in the estimation of $V_{\hat{B}i}$. As shown in Table 4.10, a 10% tolerance did in fact improve the probability of convergence to the optimal catalyst 1, albeit not as well as we had hoped. Though 7 of the 10 case studies identified catalyst 1 as optimal, only one identified an optimal TON of greater than 180. By reducing the volume of the trust region by a factor of 64 to ±2.5% in each continuous variable (Table 4.11), we found that we could increase the frequency of identification of catalyst 1 as optimal to 80%, with half of these instances identifying the optimum C_{cat} at 0.835 mM and an optimal TON greater than 180. Coupling this algorithm to a more robust continuous variable local search algorithm would likely increase the probability of finding the correct continuous variable optimum once all other discrete variables have been fathomed. Interestingly, we learned that the use of a trust-region based $V_{\hat{B}i}$ did nothing to slow the overall convergence rate of our algorithm; as a matter of fact on average the number of experiments required to converge to an optimum decreased from 151 in the non-trust region case to 138 in the trust region case, though substantial variability was observed from one simulation to the next.

Table 4.10. Optimization results for case of catalyst 1 deactivation at T > 80°C, assuming a trust region on the calculation of the prediction covariance of *i* of ±0.9 min t_{res}, ±8°C *T*, and ±0.33 mM C_{cat}. N_{expts} = 140 ± 73.

Catalyst	t_{res} (min)	*T* (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	80.0	0.835	187.6	
2	10.0	110.0	1.161	128.7	
3	10.0	110.0	1.161	128.7	
4	10.0	110.0	1.161	128.7	
5	10.0	110.0	1.161	128.7	
6	10.0	110.0	2.976	50.2	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	77.2	1.367	116.7	26.9
2	10.0	110.0	1.632	104.9	36.7
3	10.0	110.0	1.636	105.2	37.7
4	10.0	110.0	1.579	107.7	35.0
5	10.0	110.0	1.481	110.3	31.3
6	10.0	110.0	3.918	39.8	10.2
7	10.0	110.0	4.175	31.3	0.9
8	10.0	110.0	4.175	18.5	1.5

Table 4.11. Optimization results for case of catalyst 1 deactivation at T > 80°C, assuming a trust region on the calculation of the prediction covariance of *i* of ±0.22 min t_{res}, ±2°C *T*, and ±0.08 mM C_{cat}. N_{expts} = 136 ± 33.

Catalyst	t_{res} (min)	*T* (°C)	C_{cat} (mM)	TON	TON Std.
			True Optimum		
1	10.0	80.0	0.835	187.6	
2	10.0	110.0	1.161	128.7	
3	10.0	110.0	1.161	128.7	
4	10.0	110.0	1.161	128.7	
5	10.0	110.0	1.161	128.7	
6	10.0	110.0	2.976	50.2	
7	10.0	110.0	4.175	33.2	
8	10.0	110.0	4.175	21.3	
			Response Surface Optimum		
1	10.0	77.9	1.187	139.4	39.5
2	10.0	110.0	1.629	101.8	31.7
3	10.0	110.0	1.713	97.7	32.7
4	10.0	110.0	1.674	99.5	32.2
5	10.0	110.0	1.528	107.7	30.3
6	10.0	110.0	4.163	36.2	0.4
7	10.0	110.0	4.175	32.0	0.8
8	10.0	110.0	4.175	18.4	1.0

These results suggest that algorithms which test the variance over a broader region of the experimental design space—for example I-optimality-based design of experiments algorithms[201]—may be better suited for solving problems where changes in the discrete variable effect significant change in the curvature of the response surface. Heuristics and rules must be developed, however, for determination of the appropriate size of region over which the variance can be safely computed. If we examine in Figure 4.6 the true response surface for the catalyst deactivation example, we observe that the objective function value decreases sharply close to the optimum, making the choice of an acceptable prediction covariance tolerance region all the more complicated. It is likely that adaptive algorithms will be needed in the future to test both the optimum and the size of the variance trust region, which can pose a challenge to the goal of minimizing the number of real-time experiments needed to optimize the process.

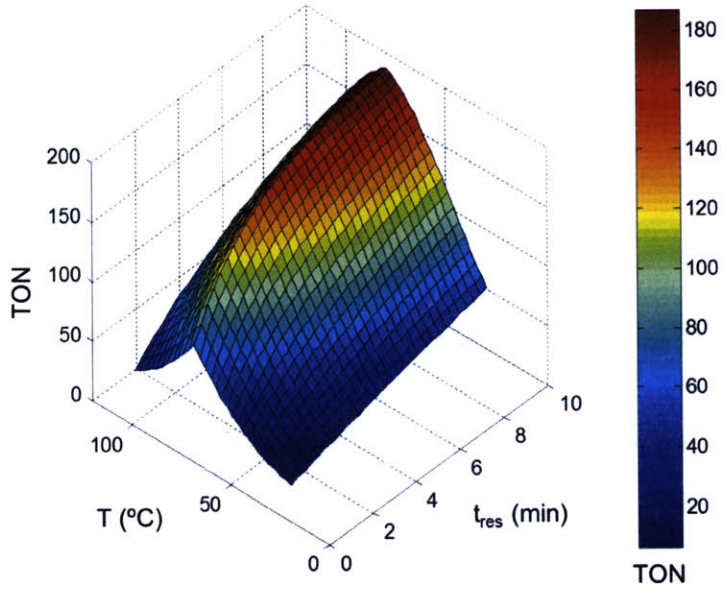


Figure 4.6. True response surface for catalyst 1 the case of catalyst 1 deactivation at T > 80°C at the minimum C_{cat} = 0.835 mM. The optimum is at T = 80°C and t_{res} = 10 min.

4.4. CONCLUSIONS

The optimization of chemical reaction networks without prior knowledge of the reaction mechanism is generally slow on account of the interdependence and often inconvenient scaling of independent and dependent continuous variables in the problem. Discrete factors such as catalysts, ligands, and solvents only complicate the problem more by changing perhaps the rate limiting step of the reaction or the mechanism altogether. In these instances, it may be truly unfavorable to propose a model for all discrete variable cases, particularly when the number of

model parameters is large, requiring that many expensive experiments be run to minimize uncertainty in the proposed model or optimum.

An adaptive response surface technique, on the other hand, allows for correlations to be identified rapidly among variables. These correlations can be used in some instances to identify relationships that may not be evident without proposing the perfect mechanism or by simply optimizing a single factor at a time. The generality of the method makes for a more versatile approach to reaction optimization when only the reaction stoichiometry is known. Still, caution must be taken in ARSM to test predictions rather than to simply propose numerical correlations. With sequential ARSM, best-fit response surfaces can be tested iteratively to refine, for instance, the maximum of an objective function or nonlinear constraints, and B&B strategies can be imposed to preferentially refine more optimal response surfaces compared to other sub-optimal surfaces.

The coupling of sequential ARSM to discrete and continuous variable optimization offers a simple approach to preliminary reaction optimization, when the effects of many variables on a synthesis are unknown. It is not a universal modeling tool, and as we have illustrated is not immune to the pitfalls of traditional local black box optimization. However, considering the complexity of most synthetic reaction models, the ability to optimize over an expansive set of discrete variables and range of continuous variables with nonlinear constraints imposed—in many cases in fewer than 200 experiments, easily examinable by today's HTE equipment—is an accomplishment that the discovery and development community should hope to embrace. Once the optimal set of variables in a synthesis is chosen, more complete kinetic models can be constructed for reaction scale-up as demonstrated in Chapter 2. As an unforeseen advantage of this method, our simulation studies showed that qualitative relationships could be accurately constructed among discrete variables. For example, the relative order of discrete variables in terms of optimality was consistent with pre-defined kinetics in almost all cases, and even when the method failed to produce the true optimum (particularly in the example of catalyst deactivation), the algorithm results highlighted the unique property of a particular discrete variable that were qualitatively consistent with the reaction kinetics, leaving open the potential to probe these differences specifically in further experiments. These qualitative outcomes will be explored extensively in Chapters 5 and 6, as they prove to provide insight into similar positive attributes of discrete variables that will ultimately lead to improved mechanistic understanding.

5. Simultaneous Solvent Screening and Reaction Optimization for the Alkylation of 1,2-Diaminocyclohexane

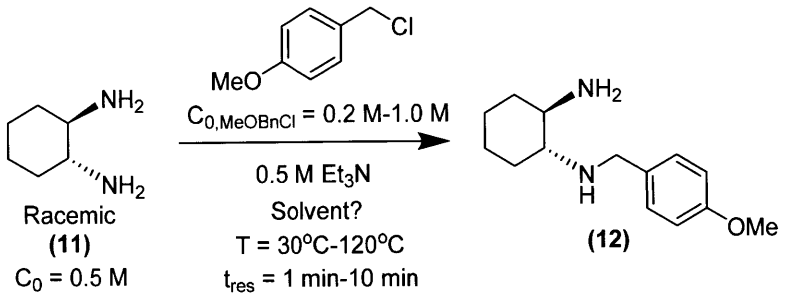
5.1. Introduction

Among the discrete variable factors that profoundly affect reaction activation and kinetics, medium effects stand out as perhaps the most universal factor to examine, yet remain perhaps the most perplexing effects to rationalize and quantify. The reason for this complication is that the reaction solvent has an integral role in every aspect of the reaction, from the diffusion barrier that two molecules must traverse to react, to the formation and stability of a solvent cage around the two molecules, to the stabilization of the reaction transition state.[202] Aspects such as the polarity of the solvent are known to play a critical role in activation-limited reactions,[203-205] but specific attributes of the solvent that promote reactivity or selectivity are impossible to isolate because of the numerous macroscopic and microscopic properties that change in switching a reaction from one solvent to another.[206] A change in solvent could imply a change in reaction mechanism (such as a change from S_N1 to S_N2) or a change in the rate-limiting step of more complex reaction pathways.[206]

For nucleophilic substitution reactions, it has been widely established that the reaction rate depends both on the solvent dielectric constant and the hydrogen-bond donating capacity of the medium. Hughes and Ingold[203] were the first to identify that the solvent electronics work to stabilize the transition state in the case of formation of a charged intermediate, hence the acceleration of S_N1 reactions in polar media. For an uncharged transition state, as is the case in S_N2, polar protic and polar aprotic solvents produce opposite trends with increasing dipole strength, with aprotic solvents accelerating the reaction and protic solvents, by virtue of sharing hydrogen atoms with the nucleophile, inhibiting the reaction as polarity increases.[207-209] Though easy to generalize, the properties that define polarity differ from case study to case study. The dielectric constant (ε) may be the most referenced metric associated with polarity, but this property alone often fails to describe the observed trends in reactivity in solution. Other metrics such as dipole moment (μ_d), molecular size, surface tension, the electrostatic factor ($\varepsilon\mu_d$), and hydrogen-bond donor ability have been used independently or collectively to try to rationalize the observed medium effects of the reaction, with empirical correlations fit retroactively to justify proposed reaction mechanisms or rate-limiting steps.[206]

It would be ideal to have an intelligent system that could assess quantitatively the relative reactivity of reagents in different media and relate qualitative insights back to the reaction mechanism. Naturally this would be an advantageous application of feedback reaction screening in flow. The scientific literature abounds in examples of high-throughput reaction screening with dispersed phase reactions carried by inert fluorinated oils in PDMS devices.[97,99,101,103,106,210] Though applicable to biological, crystal growth, and simplified chemical reaction studies, these systems—through miscibility and lack of chemical compatibility[211]—are unable to withstand the reagents and solvents most common in organic synthesis,[161,162] particularly at high temperatures.[163] The optimization of organic solvent selection in segmented flow has therefore only been studied in limited examples and independently of manipulation of other reaction conditions.[111,112] In Chapter 3, we introduced a system which, through the use of an inert gas carrier phase and FEP tubing, provided scalable results for reactions in common organic solvents. Additionally, Chapter 4 developed an optimization method which could be employed with the automated screening system to optimize for reaction conditions simultaneously with solvent selection. By employing the system and algorithm together, we proposed that the rate acceleration caused by a given solvent could be not only optimized, but rationalized in terms of continuous factors such as temperature and concentration.

In this chapter, we demonstrate the utility of the automated system in solvent optimization for the synthesis of an asymmetric-catalyst building block. Shown in Scheme 5.1, the addition of 4-methoxybenzyl chloride to *trans*-1,2-diaminocyclohexane (**11**) yields the mono-alkylated product **12**, which can be consumed by over-alkylation at either amine position. The rate of alkylation depends on the relative orientation of the two nitrogen atoms in solution, which is controlled both by solvent electronics and by the strain on the cyclohexyl ring. Selective alkylation of polyamines has traditionally been done very slowly in batch (at 0-23°C with limiting equivalents of the alkyl-halide)[212,213] or inefficiently with protection-deprotection of one or both primary amines.[214-217] Our study herein illustrates that through optimal selection of solvent and tuning of the continuous variables temperature, reaction time, and equivalents 4-methoxybenzyl chloride, the rate of formation of **12** can be controlled to generate high yield in a sub-10 minute reaction time.



Scheme 5.1. Optimization conditions for the mono-alkylation of *trans*-1,2-diaminocyclohexane.

5.2. METHOD

Following the optimization method outlined in Chapter 4, we began with a three-step sequential optimization routine comprising an initial fractional factorial design, a fractional factorial design targeted in the experimental design space where the optimum was predicted, and a sequential RSM optimization algorithm, which minimized the uncertainty on the estimated maxima for each solvent by G-optimality. No trust region was specified for the prediction covariance. In this example, we also did not impose a weighting on the least-squares regression, nor did we unbias the response covariance by jackknifing or eliminate the experiments associated with fathomed variables from the data set. Hence, to rectify errors in fitting the final response surface, a quasi-Newton gradient and line search around the most likely optimum (optima) was performed for the remaining solvent(s) following termination of the RSM algorithm. Response surfaces were constructed for each solvent with the independent variables $\ln(t_{res})$, T^{-1}, and $\ln(C_{0,MeOBnCl})$ and dependent variable ln(yield) (implying $\gamma = 1$).

The quasi-Newton search comprised a gradient estimation and a back-stepping line search, the search direction for which was calculated using the estimated gradient and a BFGS approximation for the Hessian. The search initialized at $\mathbf{x}_i^*$. To estimate a gradient, a randomly chosen 2^{3-1} fractional factorial design was executed around $\mathbf{x}_i^*$ at vertex points given by $\mathbf{x}_i^* \pm \Delta\mathbf{x}$, where $\Delta\mathbf{x}$ was a user-defined tolerance on the desired optimization accuracy. For continuous variables where $(x_i^*)_j \pm \Delta x_j$ exceeded a constraint, the new x_j was set equal to the constraint. $\mathbf{x}_i^*$ was also examined experimentally, and a gradient was calculated from all experiments run throughout the optimization within $\mathbf{x}_i^* \pm 1.05^*\Delta\mathbf{x}$. A new estimate of J_i^* was calculated based on the linearized response surface within $\mathbf{x}_i^* \pm \Delta\mathbf{x}$, with uncertainty ΔJ_i^*:

$$\Delta J_i^* = \pm\left(V_{J^*}^{1/2}\right)\left(t_{1-\alpha,\nu=N_{expts}-N_{params}}\right) \tag{5.1}$$

V_{J^*} was the prediction covariance estimated over only $\mathbf{x}_i^* \pm 1.05^*\Delta\mathbf{x}$. The confidence parameter α was chosen to be 0.16, corresponding to a one-sided single-standard deviation confidence level. The (uphill) search direction $\mathbf{p}$ was evaluated from the gradient $\mathbf{g}$ and approximate Hessian $\mathbf{H}$:

$$\mathbf{p}_q = \mathbf{H}_q^{-1}\mathbf{g}_q \tag{5.2}$$

$\mathbf{H}^{-1}$ was initialized as the inverse of Hessian estimated from the optimal response surface, and an update for $\mathbf{H}^{-1}$ was calculated following the BFGS approximation:[218]

$$\mathbf{H}_{q+1}^{-1} = \mathbf{H}_q^{-1} + \frac{\left[\mathbf{s}_q^T\mathbf{y}_q + \mathbf{y}_q{}^T\mathbf{H}_q^{-1}\mathbf{y}_q\right]\mathbf{s}_q^T\mathbf{s}_q}{\left(\mathbf{s}_q^T\mathbf{y}_q\right)^2} - \frac{\mathbf{H}_q^{-1}\mathbf{y}_q\mathbf{s}_q^T + \mathbf{s}_q\mathbf{y}_q^T\mathbf{H}_q^{-1}}{\mathbf{s}_q^T\mathbf{y}_q} \tag{5.3}$$

$$\mathbf{s}_q = -\delta\mathbf{p}_q \tag{5.4}$$

$$\mathbf{y}_q = \mathbf{g}_{q+1} - \mathbf{g}_q \tag{5.5}$$

New experiments along a line search were selected with the scaling factor δ according to:

$$\mathbf{x}_i = \mathbf{x}_i^* + \delta\mathbf{p}_q \tag{5.6}$$

For $(x_i)_j$ in violation of a constraint, $(x_i)_j$ was chosen equal to the constraint. δ was halved iteratively until the subsequent experiment would not be a replicate of the constrained experiment.

The objective function value found at $\mathbf{x}_i$, b_u, was evaluated statistically against the uncertainty in J_i^* to determine if the line search experiment was a candidate optimum. A paired t-test was employed with the criterion:

$$J_i^* - \Delta J_i^* > b_u \tag{5.7}$$

Satisfaction of this criterion implied that $\mathbf{x}_i$ was not a candidate optimum, and hence δ was reduced to $\delta/2$ and a new line search experiment commenced. In the event that $\mathbf{x}_i$ could not be disregarded as an optimum, a gradient was calculated around $\mathbf{x}_i$ with a randomly-selected 2^{3-1} fractional factorial design. A new candidate optimum, J_i', was established in the region $\mathbf{x}_i \pm \Delta\mathbf{x}$ with variance $V_{J'}$. J_i' was compared to J_i^* with an unpaired 2-sample t-test:

$$\text{Null Hypothesis: } H_0 = J_i' - J_i^* = 0 \tag{5.8}$$

$$H_a = J_i' - J_i^* > 0$$

$$t_{stat} = \frac{J_i' - J_i^*}{\sqrt{V_{J'} + V_{J^*}}} > t_{crit} = t_{1-\alpha,\nu}$$

$$\nu = \frac{\left(V_{J'} + V_{J^*}\right)^2}{\frac{V_{J'}^2}{N_{expts}' - N_{cv}} + \frac{V_{J^*}^2}{N_{expts}^* - N_{cv}}}$$

Rejection of the null hypothesis established $\mathbf{x}_i$ as the new optimum $\mathbf{x}_i$*, and a new search direction was calculated following Equations 5.2-5.5. Failure to reject the null hypothesis resulted in a continuation of the line search with δ was reduced to $\delta/2$. When the step size was small enough such that $\delta\mathbf{p}_q < \Delta\mathbf{x}$ for all $(x_i)_j$, the optimization terminated and J_i* was recorded as the optimum for solvent i with uncertainty ΔJ_i*.

5.3. EXPERIMENTAL

5.3.1. Procedure for On-Demand Solvent Screening

A tank of nitrogen (≥99.998%, Airgas) supplied both the main process flow and provided an inert gas blanket for the reagents. Ten organic solvents from generally distinct solvent classes were considered for the optimization: acetonitrile (MeCN), dichloroethane (DCE), DMC, DMF, DMSO, dimethoxyethane (DME), *i*PrOH, pyridine, THF, and toluene. All solvents were purchased anhydrous and used as received from Sigma-Aldrich, purged with nitrogen, and stored in 7 mL vials under the septum-sealed manifold. 4-methoxybenzyl chloride (98% containing K_2CO_3 as stabilizer, Sigma-Aldrich) was purged with nitrogen and stored in a 7 mL vial under the manifold. A 5 mL solution of 2.22 g **11** (98%, Sigma-Aldrich), 1.97 g Et_3N (≥99%, Sigma-Aldrich), and 135 mg naphthalene (99%, Sigma-Aldrich) was prepared and stored under nitrogen. The online addition line was purged and filled with the solution of **11**.

To prepare a slug, the liquid handler aspirated first a 30 µL volume of nitrogen from an empty vial under the nitrogen manifold, followed by aliquots of a chosen solvent, 4-methoxybenzyl chloride, and the chosen solvent again. To minimize carryover during this process, the liquid handler probe was dipped in a wash solution of *i*PrOH before each reagent aspiration. 35 µL total liquid volume was nominally aspirated, although the relative volume of 4-methoxybenzyl

chloride to solvent was corrected by the density ratio between the two reagents. Following sample aspiration, the sample was "stirred" three times in the probe under nitrogen by pulling and pushing with the syringe pump 30 μL volume. The reaction samples were introduced as 14-μL slugs into the microfluidic system. Each slug was transported by 6.9 bar nitrogen gas supplied from a syringe pump and was mixed online with 2.1 μL of the solution of **11**, Et_3N, and naphthalene, before being carried to the reactor.

Downstream of the reactor, slugs were quenched at room temperature with a continuously flowing solution of 10% acetic acid in acetonitrile. Following sampling with a 30 μL sample loop, the sample was transported *via* syringe pump in a solution of 10% acetic acid in acetonitrile to a 2 μL sample loop. The sample was observed by reversed-phase LC/MS with split 4.6 μm particle diameter and 1.8 μm particle diameter columns. The product **12** was detected by UV at 270 nm. A complete list of experiments and results can be found in Appendix D.

5.3.2. Preparation of (N-4-methoxybenzyl)-(1*R*,2*R*)-(-)-diaminocyclohexane (**1*R*,2*R*-(-)-12**)

All reagents were used as received. (1*R*,2*R*)-(-)-1,2-diaminocyclohexane (**1R,2R-(-)-11**) (410.9 mg, 98% purity, 99% ee, Sigma-Aldrich), Et_3N (358.6 mg, ≥99%, Sigma-Aldrich), naphthalene (29.1 mg, Sigma-Aldrich), and DMSO (5 mL, anhydrous ≥99.9% Sigma-Aldrich) were stirred in a 25 mL round-bottomed flask. To the flask was added 4-methoxybenzyl chloride (1064.6 mg, 6.80 mmol, 2 equiv, 98% w/ K_2CO_3 as stabilizer Sigma-Aldrich). The flask was heated to 78°C and stirred for 7.5 min. The reaction was quenched with 4 M aq. NaOH (5 mL) at room temperature, and analysis was taken by HPLC showing 61% yield of **12**. The resulting solution was extracted 5 times with 40 mL ethyl acetate. The collected organic product was then extracted once with 50 mL 4 M NaOH, and the subsequent aqueous product was extracted twice with 25 mL ethyl acetate. The cumulative extracted organic product was then extracted again with 50 mL 4 M NaOH, and the subsequent aqueous product was extracted again twice with 25 mL ethyl acetate. The cumulative organic product was dried with Na_2SO_4, filtered, and the solvent was removed by rotary evaporation to yield an orange oil. This oil was purified using column chromatography with silica gel. The column mobile phase was increased from DCM/0.1% Et_3N to DCM/8% MeOH/0.1% Et_3N to yield an isolated sample of the product **1*R*,2*R*-(-)-12**. The solvents were removed by rotary evaporation, and the resulting product was

washed 5 times with 5 mL ethyl acetate followed by rotary evaporation to remove excess Et_3N. The product was then dissolved in DCM, precipitated with hexane, filtered, and dried under vacuum, yielding the product **1*R*,2*R*-(-)-12** (500.9 mg, 2.14 mmol, 59% yield) as a white powder. ^1H and ^{13}C NMR spectra are included in Appendix D.

Product characterization: ^1H NMR (400 MHz, $CDCl_3$): δ = 7.36 (d, J = 8.6 Hz, 2H), 6.88 (d, J = 8.7 Hz, 2H), 4.23 (s, 2H), 4.11 – 3.74 (dd, J = 115.5, J = 12.9, 2H), 3.79 (s, 3H), 2.81 (td, J = 11.2, 4.1 Hz, 1H), 2.53 (td, J = 10.9, 3.9 Hz, 1H), 2.18 (dd, J = 12.1 Hz, 2H), 1.75 (dd, J = 25.0, 11.7 Hz, 2H), 1.54 – 1.14 (m, 5H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 159.28, 130.12, 129.52, 114.15, 59.67, 55.37, 54.76, 49.57, 31.65, 30.32, 24.68, 24.36. HRMS (ESI) *m/z* 235.1805 (calculated for $C_{14}H_{22}N_2O$ 235.1805 $[M+H]^+$).

5.3.3. Automated Reagent Calibration

A solution of 0.542 M **12** and 0.0433 M naphthalene in DMSO and a solution of 0.0474 M naphthalene in DMSO were stored under nitrogen in the liquid handler. Two replicates each of slugs containing 0 M, 0.13 M, 0.26 M, 0.39 M, and 0.52 M **12** were prepared following the same procedure as in the optimization, routed through the screening system (without online injection) at 30°C and a residence time of 5 min, and analyzed by LC/MS. A calibration was constructed based on integrated peak absorbance measurements of the desired product and naphthalene at 270 nm. The calibrated slope was $C_{prod} = 6.98 * C_{naphthalene} * A_{prod}/A_{naphthalene}$ with $R^2 = 0.987$.

5.4. RESULTS

Optimizing in the range of conditions given in Scheme 5.1, two initial sets of 20 fractional factorial design slugs (four slugs for each of ten solvents) identified seven candidate solvents to continue exploring in the first iteration of the sequential RSM algorithm. DMSO was predicted to be the most favorable solvent for producing **12** in high yield, with MeCN, DME, and DMC considered the three least favorable solvents for the reaction. These three solvents were fathomed in the next optimization iteration. A comparison of predicted optima and observed optima through 40 total experiments is shown in Table 5.1. The large uncertainty in the predicted maximum yields was attributed to the inefficiency of a combinatorial factorial design approach in refining a complex response surface. However, by assuming similar trends among the continuous variables, our method was able to predict higher yields for solvents in other regions

of the experimental space that were not considered during the fractional factorial design of experiments. For that reason, THF, with a maximum observed yield of only 36% in four slugs, ranked above DMC, with a maximum observed yield of 57%.

Table 5.1. Observed maxima and maxima predicted by a linear response surface model through 40 fractional factorial design experiments. Dashed border represents solvents below minimum yield tolerance.

	Observed Maximum				Predicted Maximum			
Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
DMSO	600	69	1.00	63%	144	64	1.00	197%± 121%
DMF	600	69	1.00	56%	159	76	1.00	106%
DCE	190	120	1.00	53%	176	89	1.00	85%
Pyridine	190	120	1.00	51%	168	84	1.00	82%
*i*PrOH	600	120	0.44	43%	216	120	1.00	77%
THF	600	120	0.44	36%	216	120	1.00	77%
Toluene	190	120	0.44	44%	160	77	1.00	77%
DMC	600	120	1.00	57%	216	120	1.00	67%
DME	190	120	0.44	42%	170	85	1.00	65%
MeCN	600	120	1.00	47%	199	107	1.00	62%

Using the sequential RSM algorithm, our system subsequently ran through several iterations of response surface refinement and discrete variable elimination until linear convergence of the uncertainty in the maximum was achieved. This refinement required only an additional 27 experiments but was able to reduce the number of solvents under consideration from seven down to one, the optimal solvent being DMSO. Figures 5.1a and 5.1b illustrate the steady improvement of the observed yield over the course of the experimental design and the convergence of the predicted yields for all ten solvents, respectively. As the search progressed, a greater number of experiments were conducted with the polar aprotic solvents DMSO, DMF, and pyridine, improving the response surfaces and refining the estimated maxima for these solvents in preference to less favorable discrete variables. This resulted in rapid reduction in the uncertainty on the optimum yield from 121% to 10%. Shown in Figure 5.1c, experiments clustered at the corners of the continuous variable space (owing to the factorial design initialization) and in the interior of the space near the eventual temperature and reaction time optima for DMSO, DMF, and pyridine.

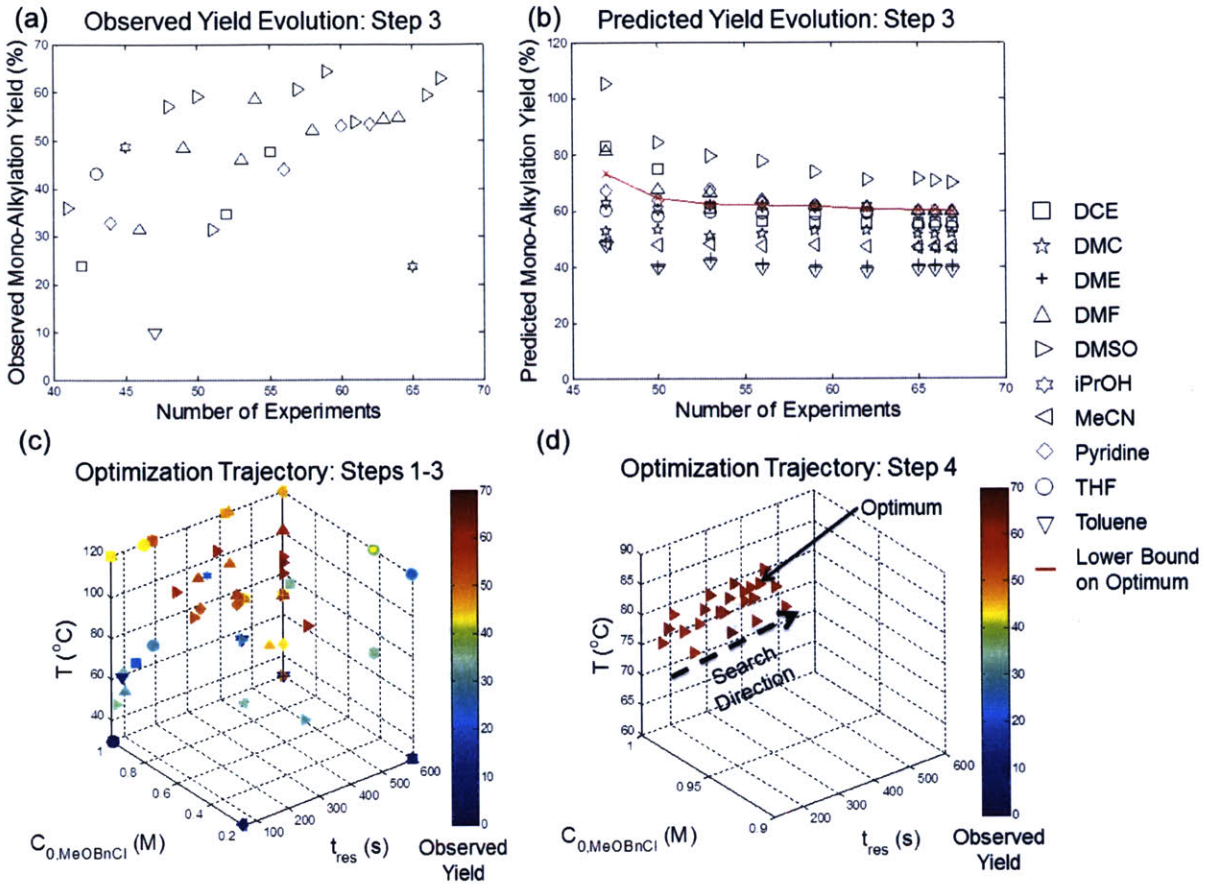
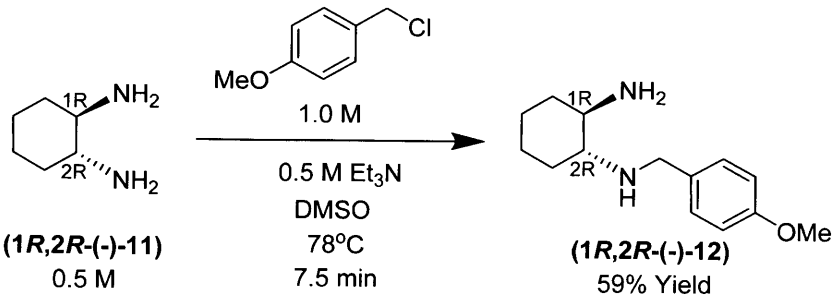


Figure 5.1. (a) Evolution in observed yield as a function of solvent during a 27-experiment sequential RSM optimization. (b) Evolution in predicted yield as a function of solvent during the same RSM optimization. (c) Optimization trajectory and observed yield for fractional factorial design and RSM experiments. (d) Optimization trajectory and observed yield for quasi-Newton gradient search with DMSO.

Review of the predicted optima for each solvent following the sequential RSM optimization (Table 5.2) revealed much improved agreement with experimentally observed optima. The optimal reaction conditions were predicted for DMSO at a moderate temperature and reaction time (78°C and 444 s) with the maximum equivalents of 4-methoxybenzyl chloride. These conditions provided a starting point for a quasi-Newton optimization of the yield of the reaction in DMSO with respect to the three continuous variables. As shown in Figure 5.1d, this optimization resulted in re-convergence to the sequential RSM optimum within the statistical noise of the system. The final optimized yield was found to be 62.1 ± 0.2% with optimal conditions t_{res} = 444 ± 15 s, T = 78 ± 2°C, $C_{0,MeOBnCl}$ = 1.00 ± 0.02 M, and DMSO as the solvent. 93 slug experiments were required collectively to find the discrete and continuous variable optimum. To validate the scalability of the method, we synthesized 0.5 g of the enantiospecific

product **1*R*,2*R*-(-)-12** in batch at the aforementioned optimal conditions (Scheme 5.2). The scale-up yielded 61% product by HPLC, with an isolated yield of 59% and 99% ee.



Scheme 5.2. Mono-alkylation of (*1R,2R*)-(-)-1,2-diaminocyclohexane.

Table 5.2. Observed maxima and maxima predicted by a linear response surface model through 67 fractional factorial and sequential RSM experiments. Dashed border represents solvents below minimum yield tolerance.

	Observed Maximum				Predicted Maximum			
Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
DMSO	600	85	1.00	64%	444	78	1.00	70% ± 10%
DMF	600	101	1.00	59%	404	95	1.00	60%
Pyridine	340	78	1.00	53%	449	76	1.00	59%
DCE	190	120	1.00	53%	349	120	0.91	56%
THF	161	120	1.00	43%	349	120	0.91	55%
DMC	600	120	1.00	57%	349	120	0.91	52%
MeCN	600	120	1.00	47%	380	104	0.96	47%
*i*PrOH	161	120	1.00	49%	349	120	0.91	47%
DME	190	120	0.44	42%	349	120	0.91	41%
Toluene	190	120	0.44	44%	368	110	0.95	39%

5.5. Discussion

Based on the predicted optima in Table 5.2 and the generated response surfaces (illustrated at the respective optimal temperatures for each solvent in Figure 5.2), we observed that a higher yield of the mono-alkylated product correlated strongly with the aprotic solvent polarity, with DMSO, DMF, and pyridine outperforming the other seven solvents. In the cases of all three favorable solvents and MeCN, increasing temperature led to over-alkylation and reduced selectivity. We reasoned this enhanced rate of alkylation derived from better stabilization of the dipolar transition state in polar aprotic solvents, as compared to other solvent classes. While in batch experimentation or combinatorial screening careful tuning and control of the reaction temperature and time would present a challenge to identifying conditions where good product

selectivity could be achieved in polar aprotic solvents, our system was able to rapidly identify conditions where mono-alkylation was most favorable, preventing the need for use of slow-reacting nonpolar solvents in the synthesis of **12**.

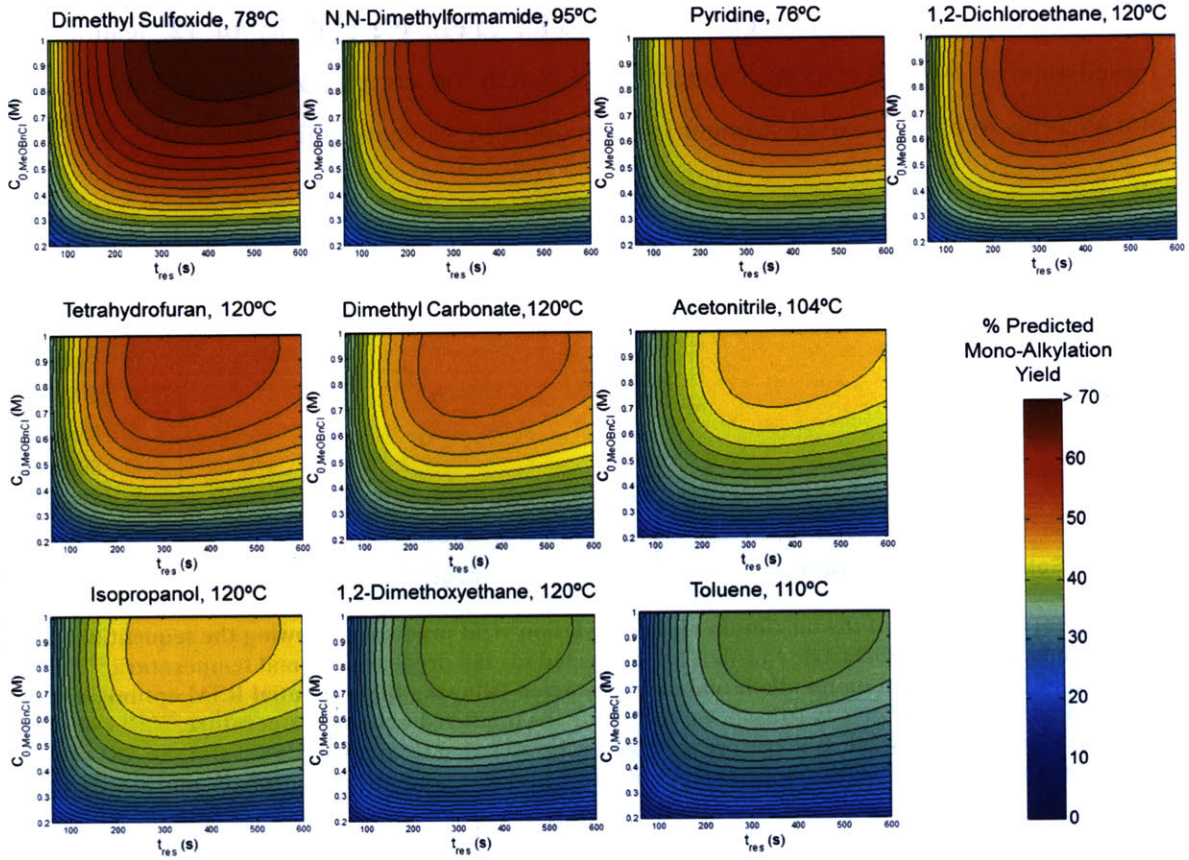


Figure 5.2. Quadratic response surfaces for the predicted mono-alkylation yield of product 12. Response surfaces were calculated at the optimal temperature for each solvent at termination of the sequential RSM optimization.

We compared the optimal predicted yields found from the sequential RSM optimization to two common metrics of solvent polarity: the solvent dielectric constant (Figure 5.3a) and the solvent hydrogen bond basicity (Figure 5.3b). Good qualitative agreement between the predicted maximum yield and the solvent dielectric constant was observed for the solvents DMSO, DMF, *i*PrOH, DME, and toluene; however the remaining five solvents deviated greatly from the reaction performance predicted by the dielectric constant alone. Seeking better rationalization for the solvent performances in our system, we considered a recent study by Lebleu *et al.* which demonstrated that selective mono-methylation of primary amines could be achieved with a good H-bond donating solvent.[219] Comparing our optimization results to the hydrogen-bond basicities

of the solvents in our study, we as well observed good correlation in the predicted reaction yield with the H-bond donating capacity of our solvent (DCE being a lone outlier). Knowing that chiral diamines like the product **12** were themselves good H-bond donors,[220] we hypothesized that the basicity of the solvent worked to counteract the nucleophilicity of **12**, leading to a decreased rate of over-alkylation. Though based strictly on correlation, this hypothesis could inspire attempts to optimize over stronger H-bond donating solvents in future studies.

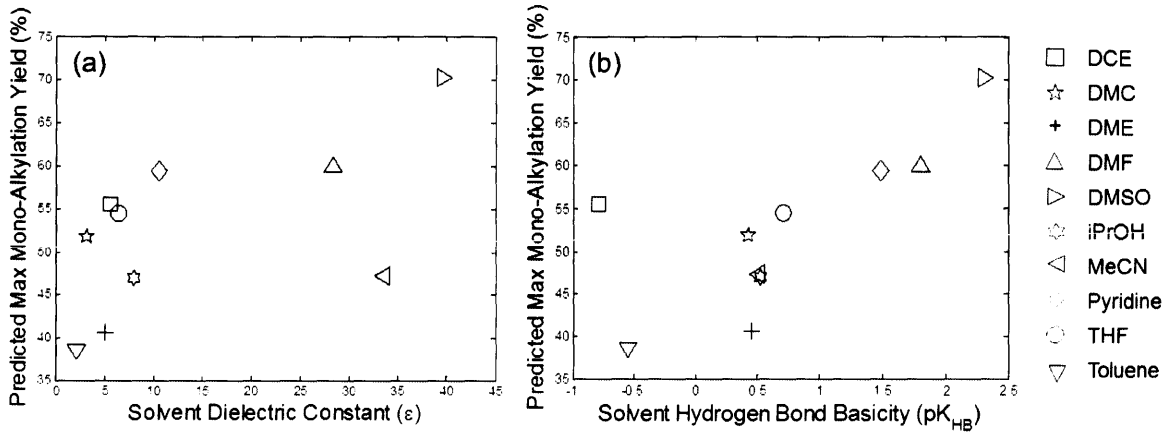


Figure 5.3. (a) Correlation of the maximum mono-alkylation yield predicted following the sequential RSM optimization to the solvent dielectric constant (ε), corrected for the predicted optimal temperature.[221-223] (b) Correlation of the maximum mono-alkylation yield predicted following the sequential RSM optimization to the solvent hydrogen bond basicity (pK_{HB}), corrected for the predicted optimal temperature.[224-232] For cases where ΔS° was not available in literature, $pK_{HB}(T)$ was estimated using $pK_{HB}(25^{\circ}C)$ and the ΔS° of a comparable molecule: for *i*PrOH, 1-propanol;[233] for DCE and DME, 1,3-dichloropropane and 1,4-dioxane, respectively.[224]

5.6. CONCLUSIONS

The ability to screen discrete variables while simultaneously optimizing continuous variables offers new possibilities that have, to a large extent, been unutilized in batch HTE. The use of a smart search algorithm, for instance, has been shown to greatly reduce the number of experiments and amount of uncertainty on the optimum relative to what would be achievable for a traditional design of experiments in the confines of a 96-well-plate. From the response surfaces found through feedback, quantitative information was gained with regard to the change in reaction yield as a function of a perturbation in one or more process variables. The accuracy of a greater than 300-fold scale-up from 16 μL slugs to a 5 mL batch synthesis demonstrated the utility of microscale reaction modeling.

Overall, this case study illustrates how intelligent experimentation, along with the introduction of a greater diversity of model variables, may ultimately allow greater insight into

the chemistry than a vast library of high-throughput experimental studies. The significant relationships affecting yield or reaction kinetics are easiest to identify when a model relating all variables is repeatedly proposed, tested, and refined. By refining the solvent response surface model in this chapter, we were able to identify relationships between discrete and continuous variables that were consistent with scientific literature but non-trivial to quantify prior to experimentation. This is the first step in the direction of an automated platform capable of identifying the relationships needed to build insight into reaction mechanisms and to identify structural or reactive properties shared by optimal discrete variables. In Chapter 6, we demonstrate that similar understanding can be built from analysis of ligand and palladium precursor effects in Suzuki-Miyaura cross-coupling reactions.

6. Optimization and Kinetic Investigations of Suzuki-Miyaura Cross-Coupling Reactions

6.1. Introduction

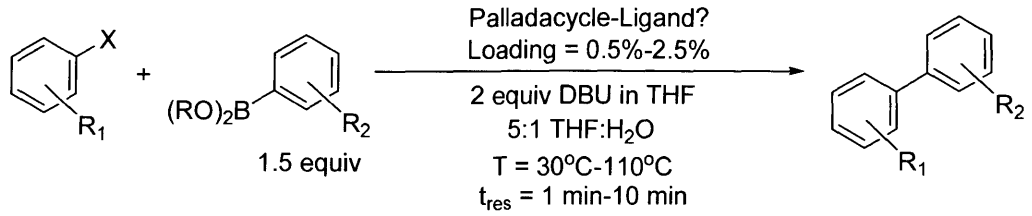
The Suzuki-Miyaura cross-coupling reaction has enjoyed much fanfare in recent decades on account of the reaction's mild conditions, high functional group tolerance, and general accessibility through the use of relatively stable and easy-to-handle organoboron reagents.[234-237] The coupling of an aryl halide to a boronic acid to form a new C-C bond holds widespread applicability in the pharmaceutical and fine chemical industries[238] and has drawn significant interest as a tool for fast scale-up from process chemistry to production in the realm of continuous manufacturing.[239] Yet as ubiquitous as the Suzuki-Miyaura cross-coupling may be in organic synthesis, a somewhat fundamental mystery has always remained unsolved: how to choose the right catalyst/ligand system for a given pair of coupling partners. Though generations of catalyst precursors and ligands have been developed and iteratively improved to produce more generalizable results, higher yields, and better turnover numbers (TONs),[240-246] it has always been to the experimentalist to instinctively or methodically identify the optimal catalytic system, while at the same time evaluating dependences on factors such as temperature, reaction time, and concentration.

The emergence of HTE[66] has to an extent simplified the chore of optimization, enabling as many as 1532 cross-coupling reactions to be examined in a single automated experiment in a matter of hours.[63] While this high-throughput approach is impressive in its enumerative power, the amount of information that can be gleaned from such a strategy is still limited by virtue of having to run all experiments at a fixed temperature and time, the reliance upon diffusion for mixing, and the high probability that experiments are not "intelligently" chosen to be run where the optimum is most likely to be found (which, in fairness to the experimentalist, cannot be expected to be known before the experiment begins). Those who have invested strictly in the HTE approach also must consider the colossal magnitude of numbers involved the screening of transition-metal catalyzed reactions. Murray *et al.*[247] estimated that even in a simplified case there may be greater than 50 million unique combinations of discrete and continuous variable factors to consider in a typical Suzuki-Miyaura cross-coupling reaction. If each of these sets of

conditions were screened with 1 mg of substrate, more than 50 kg of material would be needed to complete an optimization.

To combat the curse of dimensionality, strategies for optimal catalyst and reaction condition selection have focused upon the use of principle component analysis (PCA) and design of experiments (DoE) to select the most representative discrete and continuous variables to choose in more selective cross-coupling experiments.[248,249] PCA allows for quantification of the relative effects of ligands[250-252] based upon shared physical characteristics such as cone or bite angle,[253-255] allowing researchers to identify representative ligands that can be sampled to estimate the effect of the principle component on the reaction output. This approach comes with the disadvantage of having to characterize and map the ligands *a priori*, and as is the case with all discrete to continuous variable transformations, can easily fail by virtue of oversimplifying the properties essential to good catalytic performance with a given substrate or within a specific medium. Such properties can be factored into consideration using a feedback approach to select the optimal catalyst-ligand system for a reaction, as was accomplished for solvent selection in Chapter 5. With slug flow microfluidics we can also simultaneously minimize the amounts of expensive starting materials, transition metals, and ligands consumed during a given study. Based upon the success of our method in Chapter 5, we were also interested to explore whether our automated system could be used to tease out important mechanistic relationships that could contribute to better understanding of the Suzuki-Miyaura cross-coupling reaction.

We had particular interest in studying the coupling of aryl halides and boronic acids in the presence of the organic base DBU. This reaction has compelling implications for flow chemistry, as reaction yields can be achieved without mass transfer limitation that are comparable to the use of inorganic bases such as K_3PO_4 (see Section 3.4.3). The general optimization is presented in Scheme 6.1. We considered a fixed ratio of 5:1 THF to water (the water being necessary for the generation of OH^- ions) and temperatures and residence times ranging from 30-110°C and from 1-10 min, respectively. We sought to optimize the catalyst loading by way of maximizing the TON for the reaction. However, simply maximizing TON presented a problem in that a high TON could be achieved with relatively little yield; to avoid this, we stipulated in the optimization routine that the maximum TON be identified subject to the yield being within 90% of the overall maximum.



Scheme 6.1. Optimization scheme for Suzuki-Miyaura cross-couplings in the presence of DBU and THF/Water.

We considered for optimization the family of palladacycle-ligand precatalyst systems developed by Buchwald and coworkers (Figure 6.1).[245,256] These precatalysts have been found to be both air and moisture stable, with precatalysts of the form **P1** stable for more extensive periods of time in solution and more conducive to the incorporation of bulky ligands.[245] Upon exposure to base, **P1** and **P2** rapidly undergo deprotonation and reductive elimination. This rapid reduction to the active Pd^0-ligand complex, accelerated by the strongly electron withdrawing methane sulfonate and chloride groups, makes this family of precatalysts highly effective in Suzuki-Miyaura cross-coupling reactions, particularly in the presence of boronic acids which are prone to rapid protodeboronation.[256,257] The monoligated complex that forms upon activation of **P1** or **P2** is particularly favorable for oxidative addition of aryl chlorides.[258] Methylated and phenylated variants of **P1** have also been synthesized to prevent the formation of NH_2-aminobiphenyls during catalyst activation.[165]

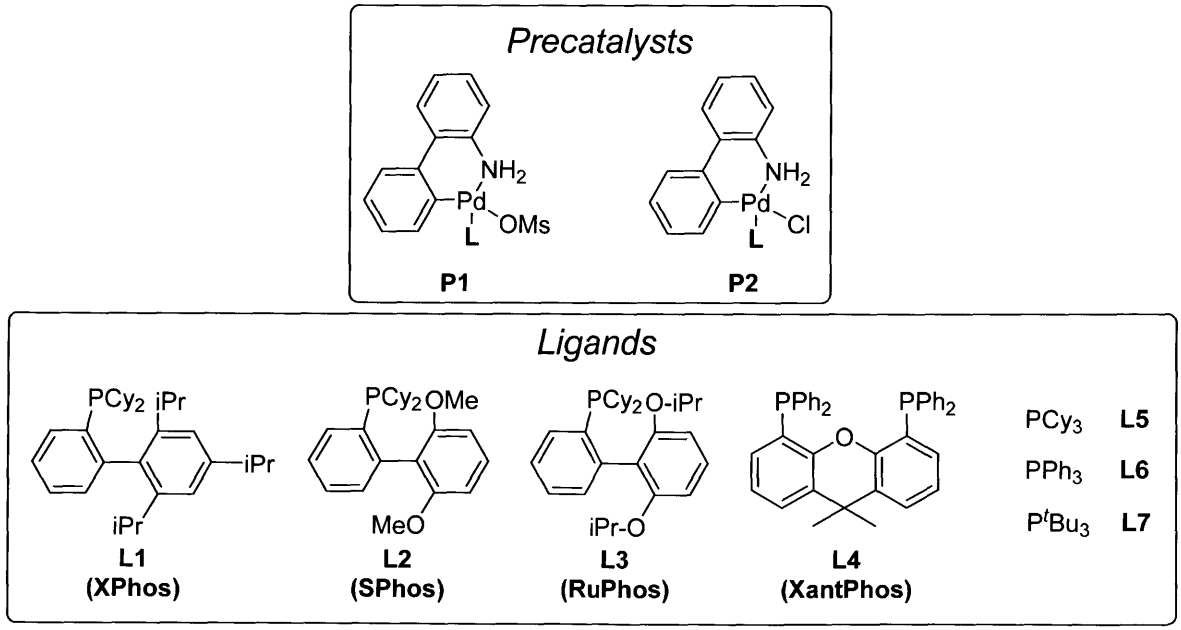


Figure 6.1. Precatalysts and ligands for Suzuki-Miyaura reaction optimization.

6.2. METHOD

Following the optimization method described in Chapter 4, the reactions were optimized for product yield with respect to total amount of ligand, which for all studies with 1:1 palladacycle-ligand solutions equaled the catalytic TON. The constraint parameter γ was chosen as 0.90. The reaction TON was chosen as the weighting factor in the least squares regression for all examples except for the optimization of 2-chloropyridine and 1-boc-2-pyrroleboronic acid. In the latter case the reaction yield was chosen as the weighting factor. No trust region was specified for the prediction covariance in these examples.

Full catalyst-ligand optimization studies commenced with a randomized 16-experiment fractional factorial design, followed by a second refined 16-experiment fractional factorial design. The ligand equivalent optimization study commenced with a randomized 12-experiment fractional factorial design, followed by a second refined 12-experiment fractional factorial design. To prevent a loss in accuracy from solvent evaporation, optimization studies were terminated at a maximum of 96 experiments, regardless of whether the termination criteria presented in Chapter 4 had been satisfied. Optimization routines were executed in MATLAB.

6.3. EXPERIMENTAL

6.3.1. General Solution Preparation Procedure

The precatalyst-ligand complexes used in this study were synthesized and isolated by Yiming Wang following the procedure published by Bruno *et al.*[245] All other reagents were used as received. Reagent solutions were prepared under ambient conditions but transferred to nitrogen-backfilled vials and stored under argon in the inert gas manifold. These solutions were prepared fresh for each optimization or kinetic parameter ramp. A 5 mL aryl halide and (the internal standard) naphthalene solution was prepared by diluting with THF to 1.4 M aryl halide and 0.4 M naphthalene and transferring the solution to a 7 mL vial. A 5 mL boronic acid or boronic pinacol ester solution was prepared by diluting with THF to 1.0 M and transferring the solution to a 7 mL vial. Eight palladacycle-ligand combinations were considered: **P1-L1**, **P2-L1**, **P1-L2**, **P1-L3**, **P1-L4**, **P1-L5**, **P1-L6**, and **P1-L7**. Individual 2 mL precatalyst solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF to yield a 0.018 M solution. A solution of makeup THF was transferred to a 7 mL vial. A solution of water,

degassed with sonication and backfilled under nitrogen, was transferred to a 7 mL vial. A 10 mL DBU solution was prepared by diluting 2.5 g DBU to 1.66 M in THF and transferring the solution to a 20 mL scintillation vial. The online injection line and syringe were purged with this DBU solution. For the ligand equivalent optimization only, individual ligand solutions containing **L1** and **L5**-HBF_4 were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF to yield a 0.05 M solution. A solution of **L7**-HBF_4 was prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL degassed water to yield a 0.05 M solution.

6.3.2. Automated Reaction Optimization and Screening

A tank of argon (≥99.997%, Airgas) supplied both the main process flow and provided an inert gas blanket for the reagents. To prepare a slug, the liquid handler aspirated first a 30 μL volume of argon from an empty vial under the argon manifold, followed by aliquots of THF, aryl halide, excess ligand (if needed), precatalyst, boronic acid or boronic acid pinacol ester, water, and THF again. To minimize carryover during this process, the liquid handler probe was dipped in a wash solution of THF before each reagent aspiration. 35 μL total liquid volume was nominally aspirated. Following sample aspiration, the sample was "stirred" three times in the probe under argon by pulling and pushing with the syringe pump 30 μL volume. The reaction samples were introduced as 14-μL slugs into the microfluidic system. The slugs were transported at 6.9 bar by argon gas driven by a syringe pump and mixed online with 3.5 μL DBU in THF before reaction in a heated FEP tube reactor. Reacting slugs comprised 0.167 M aryl halide, 0.250 M boronic acid, 0.333 M DBU, and 0.5-2.5 mol% precatalyst. Temperatures ranged 30-110°C and reaction times ranged 1-10 min. Each slug was quenched with a 1:1 solution of water and acetone, with 1 μL sampled for analysis by LC/MS. The sample was split in the HPLC between a 1.8 μm analytical column and a 4.6 μm pressure resistance column. Product quantitation was performed by UV at pre-determined wavelengths of 230, 254, 270, 285, 300, or 340 nm. Before each experiment, blank slugs of water, acetone, and THF were prepared and injected into the system. A complete list of experiments, measured product yields and TONs, and sample preparation procedures can be found in Appendix E.

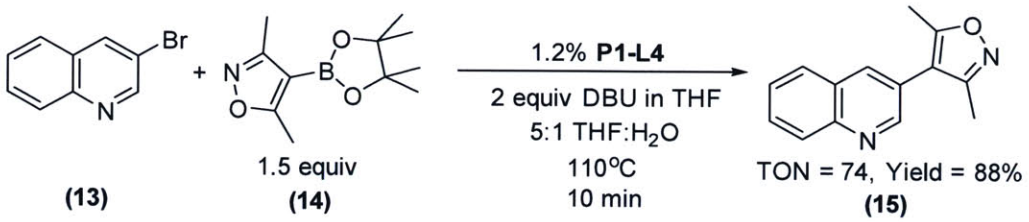
6.3.3. Automated Boronic Acid and Ester Degradation Studies

The system was set up as described above. To prepare a slug, the liquid handler aspirated first a 30 μL volume of argon from an empty vial under the argon manifold, followed by aliquots of THF, naphthalene, boronic acid or boronic acid pinacol ester, water, and THF again. The sample was rinsed, stirred, and injected into the system as previously described. The slugs were mixed online with 3.5 μL DBU in THF before reaction at 110°C. Each slug was quenched with acetonitrile (to prevent additional water-induced protodeboronation), with 1 μL sampled for analysis by LC/MS. The HPLC method was as previously described. The boronic acid concentration was quantified by UV at 300 nm and compared to the absorbance found after flowing the boronic acid and naphthalene through the system in the absence of water or DBU for 5 min at 30°C.

6.4. RESULTS

6.4.1. Optimization of TON in Suzuki-Miyaura Cross-Coupling Systems

We first considered with our system the coupling of 3-bromoquinoline (**13**) with 3,5-dimethylisoxazole-4-boronic acid pinacol ester (**14**) (Scheme 6.2). The system began by searching the extremes of the experimental space for temperature, reaction time, and catalyst loading before moving to interior of the experimental space and disregarding **L6** as a candidate optimal ligand. Subsequent experiments led to further reduction in the number of possible optimal palladacycle-ligand combinations until only **P1-L4** remained under consideration after 78 experimental slugs. Further experiments were employed to reduce the uncertainty on the optimal yield (98% **15** at the maximum temperature of 110°C, residence time of 10 min, and precatalyst loading of 2.5%) and the optimal TON (74 at 110°C, 10 min, and 1.2% precatalyst). The experimental trajectory progressed much in similarity to the simulations in Section 4.3.1. As observed in Figure 6.2, most of the automated system's effort focused on resolving the cutoff at which 90% of the maximum yield could be achieved with a reduced loading of **P1-L4**, just as a traditional process chemist would have identified the best temperature, reaction time, and catalyst and tuned the catalyst loading to improve TON.



Scheme 6.2. Optimum for the Suzuki-Miyaura cross-coupling of 3-bromoquinoline and 3,5-dimethylisoxazole-4-boronic acid pinacol ester.

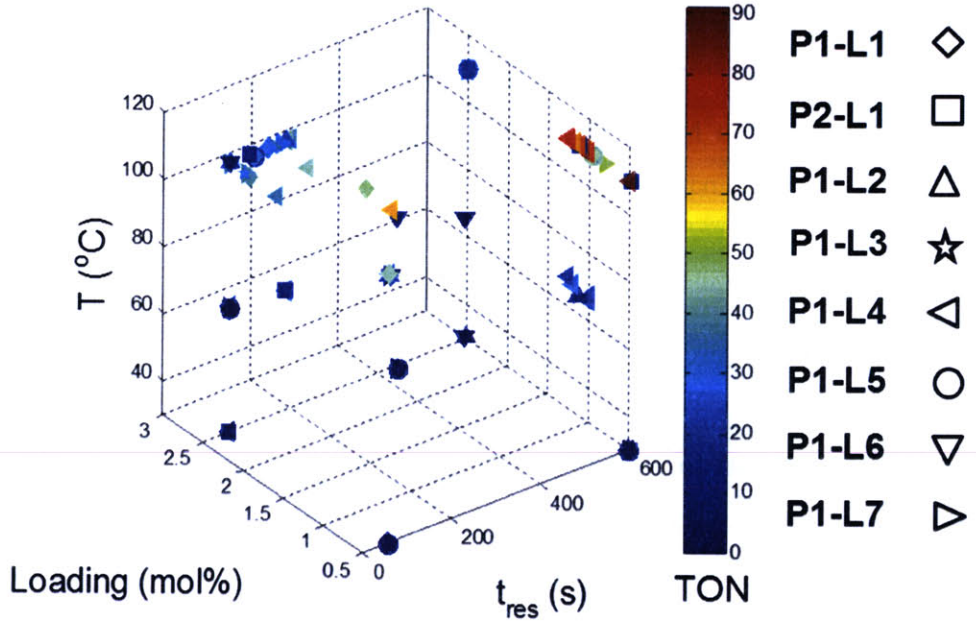
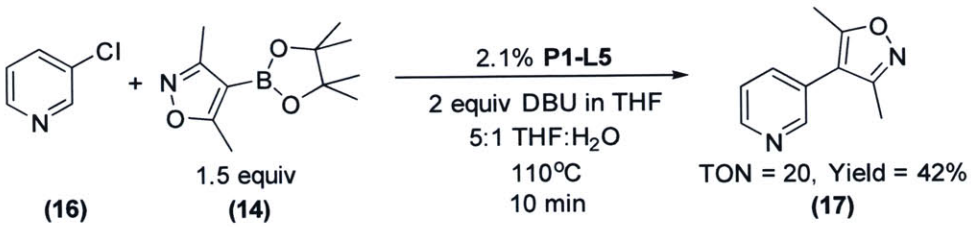


Figure 6.2. Automated optimization trajectory for the synthesis of 15.

Confident in our system's ability to rapidly identify the best catalytic system and experimental variables, we challenged the system by switching from the aryl bromide to 3-chloropyridine (**16**) (Scheme 6.3). The reaction of **16** with **14** was considerably slower and gave poor yields when **P1-L4** was selected as the precatalyst. Among the eight candidate palladacycle-ligand pairs, the system identified **P1-L5** as a suitable catalyst system, generating 42% yield of **17** in 10 min at 110°C and 2.1% Pd loading. By contrast, selection of **P1-L5** was only found to produce a maximum yield of 3%. We were also surprised to see the trialkylphosphine ligand **L5** outperform dialkylbiarylphosphine ligands **L1-L3**. Further analysis of the optimization results revealed that although the use of **P1-L1** at reduced temperatures gave the product **17** in moderate yield, **P1-L5** once activated at 110°C appreciably outperformed the maximum yield achievable with **P1-L1** at all temperatures.



Scheme 6.3. Optimum for the Suzuki-Miyaura cross-coupling of 3-chloropyridine and 3,5-dimethylisoxazole-4-boronic acid pinacol ester.

Alternatively, we considered the cross-coupling of **16** with benzofuran-2-boronic acid (**18**) (Scheme 6.4). Unlike in the two prior cases, the maximum TON for the production of **19** was found to occur at a shortened residence time of 3.9 min, with 1.2% **P1-L1** catalyst at 110°C. Review of the optimization trajectory (Figure 6.3) showed that for this case the automated system conducted a sweep of moderate residence times and catalyst loadings at 110°C before converging upon an optimum. Experiments were also conducted at 10 min residence times with **P1-L1** over a range of catalyst loadings, but no improvements in TON were observed beyond that which could be achieved in a sub-4 minute reaction. The dialkylbiarylphosphine ligands **L1-L3** faired particularly well in this example, all generating maximum yields of greater than 75%. Catalyst systems of the other four ligands, including **L4** and **L5**, were fathomed from the optimization early and found to be unlikely to produce yields in excess of 40%.

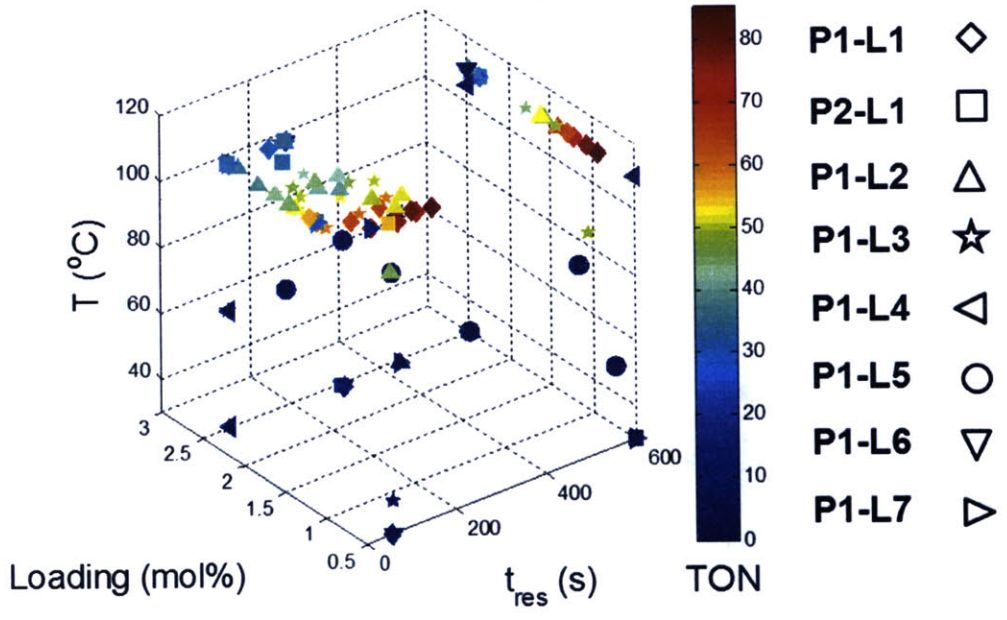
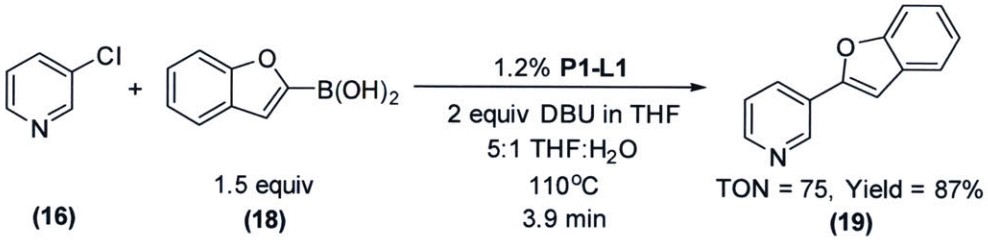


Figure 6.3. Automated optimization trajectory for the synthesis of 19.



Scheme 6.4. Optimum for the Suzuki-Miyaura cross-coupling of 3-chloropyridine and benzofuran-2-boronic acid.

Intrigued by the shift in the continuous variable optimum from substitution of a boronic acid pinacol ester for a boronic acid, we attempted what was anticipated to be an even faster and more unstable reaction of 2-chloropyridine (**9**) and 1-boc-2-pyrroleboronic acid (**7**) (Scheme 6.5). As was reasonable from conversion from the *meta*-substituted to the *ortho*-substituted pyridine, the reaction completed to 90% yield in less than 5 min, with an optimum of 1.0% **P1-L1** at 97°C. Like in the case of synthesizing **19**, the dialkylbiarylphosphine ligands **L1-L3** were distinguished by the optimization system as providing significantly higher yields (greater than 88%) and TONs in comparison to other considered ligands. Illustrated in Figure 6.4, the experimental trajectory in this case extensively searched the experimental design space before convergence to the interior reaction time-temperature-catalyst loading optimum.

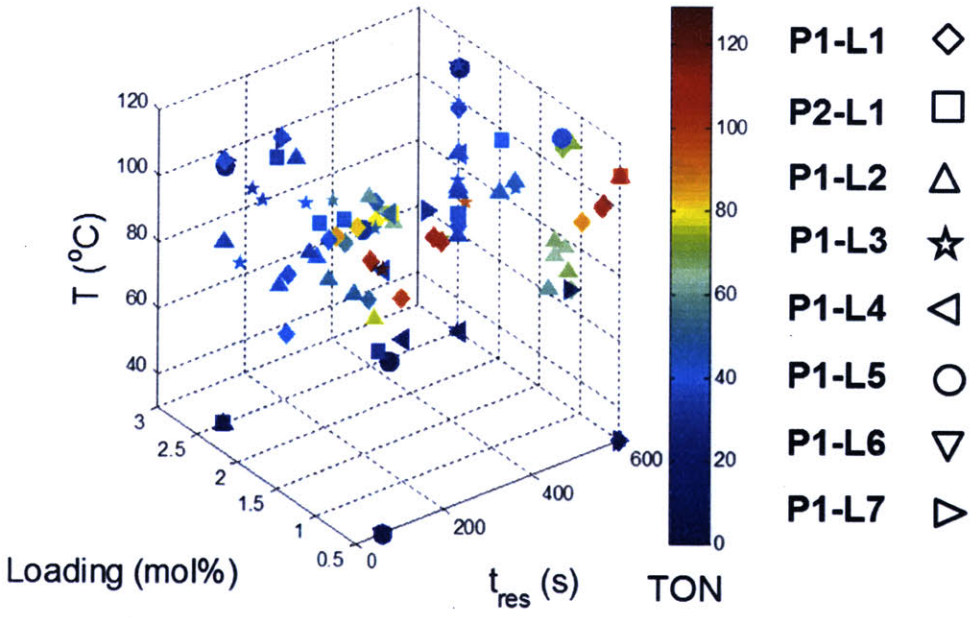
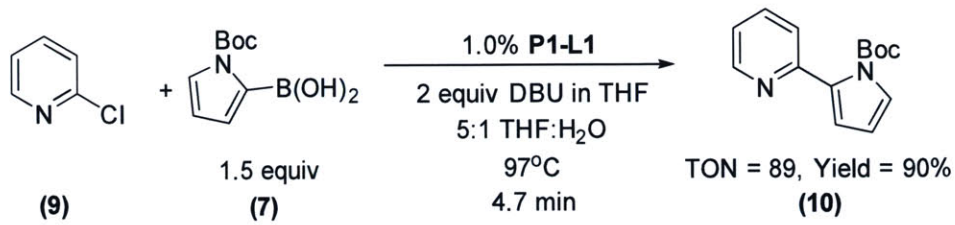


Figure 6.4. Automated optimization trajectory for the synthesis of 10.



Scheme 6.5. Optimum for the Suzuki-Miyaura cross-coupling of 2-chloropyridine and 1-boc-2-pyrroleboronic acid.

Given that no upper or lower bound constraint was active upon convergence to the optimum in Scheme 6.5, we questioned the validity of the proposed optimum and proceeded to conduct a series of automated experiments to synthesize **10** in slug flow at different reaction times and temperatures at 1.0% loading of **P1-L1**. Figure 6.5a shows the response surface estimated by the optimization algorithm, and Figure 6.5b overlays the results of the **P1-L1** screening experiments upon the predicted yields from the response surface model at 80°C, 97°C, and 110°C. The response surface predictions agree closely with the screening results near the optimum of 4.7 min and 97°C and capture the reduced TON at both long residence times and high temperature and short residence times and low temperature. A consequence of the adaptive response surface algorithm is that farther away from the optimum (for instance at 80°C) the approximated response surface does not capture accurately the slower activation time of the reaction.

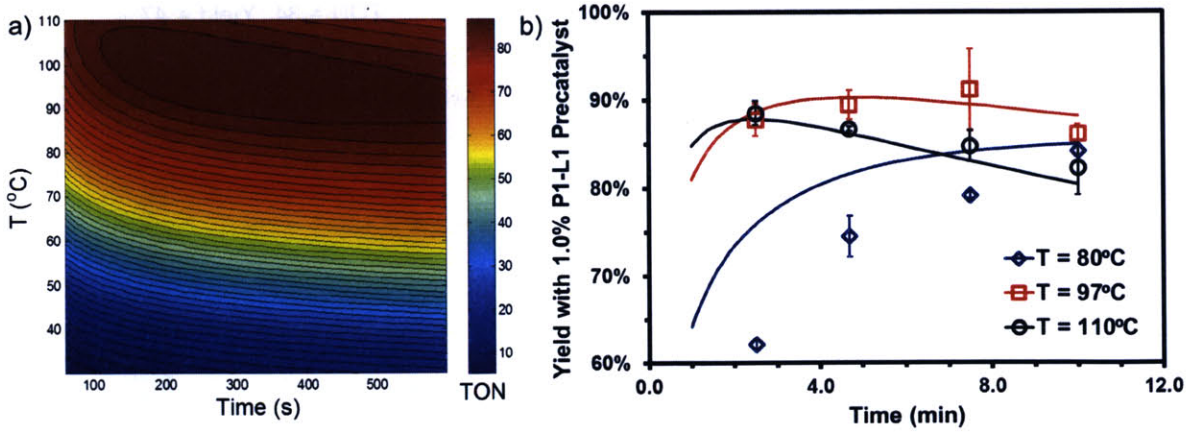
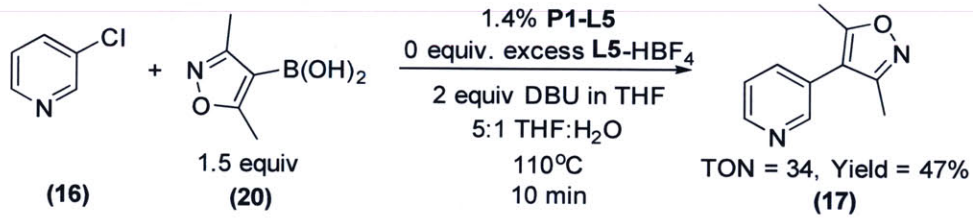


Figure 6.5. (a) Predicted response surface for the synthesis of 10 with 1.0% P1-L1 precatalyst. (b) Comparison of automated screening experiments (markers) on the predicted yield based on the best-fit response surface (solid line) for the synthesis of 10 at same coniditons.

6.4.2. Optimization of Ligand Equivalents

We postulated that the reactions of **16** may be accelerated by adding excess ligand to solution. To study the effect of adding excess ligand, we employed the same automated system and characterized the reaction of **16** with 3,5-dimethylisoxazoleboronic acid (**20**) (Scheme 6.6). As a simplification, we considered only the precatalyst **P1** and the ligands **L1**, **L5**, and **L7**, and examined the effect of manipulating temperature and excess equivalents ligand (from 0.0 to 2.0) at 10 min reaction time and 1.4% precatalyst-ligand loading. As in the cross-coupling of **16** with the boronic acid pinacol ester, the optimization algorithm rapidly identified **P1-L5** as the best catalytic system at 110°C. Shown in Figure 6.6(a-b), the reaction yield improved overall in the range of 0.2-0.8 excess ligand equivalents but decreased significantly with the use of 2 excess ligand equivalents for all precatalysts. On a per-ligand basis, it was found to be non-optimal to introduce excess **L5**-HBF_4 to the **P1-L5** system. For both **L1** and **L7** the automated system postulated only 0.3 excess ligand equivalents as optimal both on a per ligand basis and for the overall reaction yield.



Scheme 6.6. Optimum for the Suzuki-Miyaura cross-coupling of 3-chloropyridine and 3,5-dimethylisoxazole-4-boronic acid.

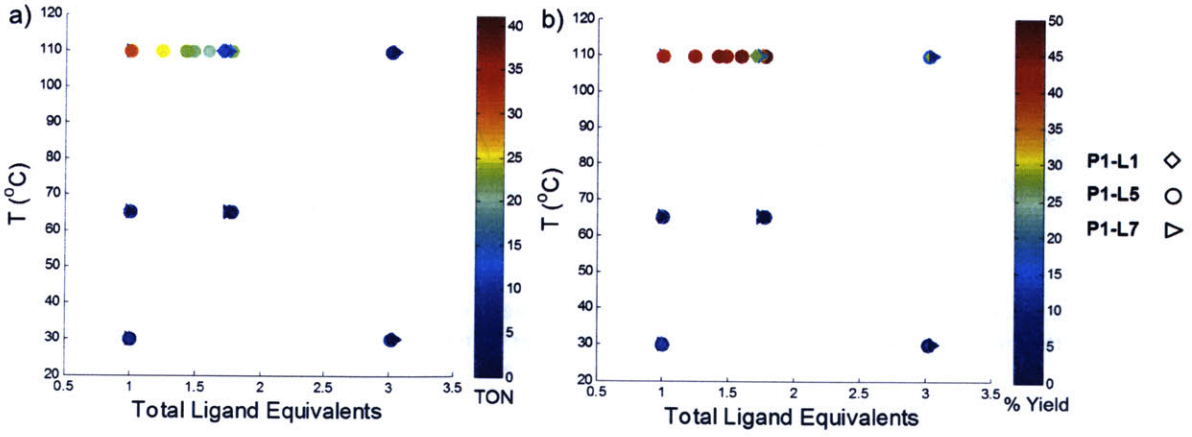


Figure 6.6. Automated optimization trajectory for the synthesis of 17 by reaction of 16 and 20. (a) TON optimization profile with respect to ligand equivalents. (b) Yield optimization profile.

6.5. DISCUSSION

6.5.1. Mechanistic Insights

The Suzuki-Miyaura reaction mechanism has been studied extensively,[258-261] and though the role of the base in the transmetallation step is still controversial,[262-265] the catalytic cycle generally proceeds as illustrated in Figure 6.7. The Pd^0-ligand complex is generated from the activation of the Pd^{II} precatalyst and undergoes oxidative addition of the aryl halide. This is followed by transmetallation and then by reductive elimination to generate the product and regenerate Pd^0. Faster uptake of the aryl halide into the catalytic cycle makes the transmetallation step rate-limiting. Substrates which are poor oxidants, such as aryl chlorides or unactivated aryl bromides, limit the oxidative addition step of the cycle and can therefore change the rate limiting step from transmetallation to oxidative addition.[258] For sterically bulky ligands and aryl chlorides, the monoligated palladium species is preferential for oxidative addition to the palladium, with inhibitory reaction kinetics observed with increasing ligand concentration.[266,267]

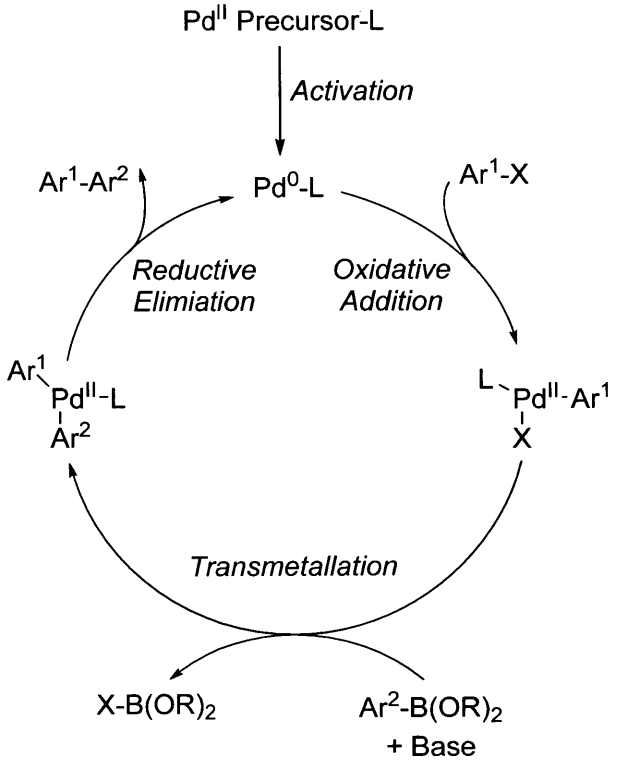


Figure 6.7. Generalized catalytic cycle for the Suzuki-Miyaura cross-coupling of an aryl halide and an aryl boronic acid.

The results of our optimization study suggest such a transition in rate-limiting step when converting from the aryl bromide **13** to the aryl chloride **16**. From Table 6.1, which summarizes the maximum yield and TON predicted by the automated system for the Suzuki-Miyaura reaction case studies, we observed a significant decline in turnover rate with **L4** using the chloride substrate, whereas the optimal performance for the other ligands in the study only declined slightly or improved, as in the case of **L5**. This suggested that **L4** accelerated the rate of transmetallation upon fast uptake of the bromide into the catalytic cycle, but with the chloride the slow rate of oxidative addition offset the effectiveness of the bidentate ligand. By contrast, trialkylphosphine ligands **L5** and **L7** and dialkylbiarylphosphine ligands **L1-L3**, by nature of being good electron donors, continued to promote oxidative addition even in the presence of the chloride substrate. Though it was speculated that the coordination of multiple ligands to the Pd center may contribute to the enhanced performance of **L1**, **L5**, and **L7**, optimization of the synthesis of **17** with the boronic acid substrate showed no improved reaction rate per unit ligand. This would seem to support the optimality of a monoligated palladium species. The decrease in reaction yield observed with a large excess of ligand suggests that the excess ligand competes with the aryl halide to bind to the palladium center, further retarding the reaction.

Table 6.1 Optimal yield and TON found by automated optimization of Suzuki-Miyaura case studies. Yields for syntheses of 10, 15, and 19 are based on conversion of the aryl halide.

Precat-Ligand	13 + 14 → 15		16 + 14 → 17		16 + 18 → 19		9 + 7 → 10	
	Max Yield	Max TON	Max Yield	Max TON	Max Yield	Max TON	Max Yield	Max TON
P1-L1	85%	39	45%	18	**97%**	**75**	**99%**	**89**
P2-L1	22%	9	8%	3	78%	41	88%	42
P1-L2	50%	20	30%	12	81%	44	95%	65
P1-L3	71%	28	35%	14	87%	52	90%	62
P1-L4	**98%**	**74**	3%	1	4%	2	73%	29
P1-L5	52%	21	**47%**	**20**	6%	2	54%	32
P1-L6	9%	4	1%	1	29%	11	34%	19
P1-L7	81%	32	25%	10	40%	20	27%	16

6.5.2. Unstable Reactants and Products, and the Correlation to Ligand Selection

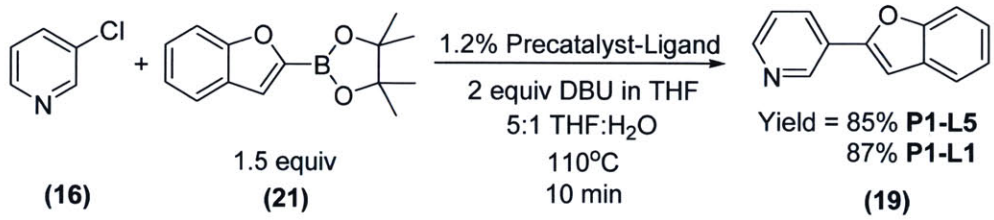
In the shift from the boronic pinacol ester **14** to the boronic acids **18** and **10**, a clear transition was observed in the preference of dialkylbiarylphosphine ligands **L1-L3** over other ligands considered in the study. In both boronic acid cases we also observed optimality in the reaction residence time at less than 5 min (as compared to the maximum of 10 min for the boronic pinacol ester **14**), and in the case of the synthesis of **10** the optimal yield and TON were both achieved at

less than 100°C with **L1**. For this last case, the optimal reaction conditions for all palladacycle-ligand combinations listed in Table 6.2 also showed a distinct segregation in the optimality of ligands **L1-L3** in comparison to the other ligands in the study. Whereas catalysts bound to less sterically-hindered alkyl or aryl ligands were observed most optimal at the maximum temperature and short reaction times, **L1-L3** were found to be optimal in the range of 85-97°C and at moderately longer reaction times of 4-6 min.

Table 6.2. Optimal TON conditions for the reaction of 9 and 7 to produce 10. P1-L1 was found to be optimal in 97 experiments.

Precat-Ligand	t_{res} (min)	T (°C)	Cat. Loading (mol%)	TON
P1-L1	4.7	97.2	1.012	88.7
P2-L1	6.1	88.2	2.088	42.2
P1-L2	6.2	85.9	1.381	65.0
P1-L3	4.3	95.5	1.442	61.7
P1-L4	10.0	79.6	2.500	29.0
P1-L5	1.5	110.0	1.707	31.7
P1-L6	1.7	110.0	1.790	18.8
P1-L7	2.4	110.0	1.760	15.6

We attributed the preferences for dialkylbiarylphosphine ligands in these cases to the competing rates of catalyst activation and reagent degradation. Both **18** and **10** were expected to undergo rapid protodeboronation upon exposure to base at high temperature. Using the automated system, we studied the kinetics of the degradation of **18** by co-flowing slugs of 0.250 M boronic acid with 0.333 M DBU in 5:1 THF-water. As illustrated in Figure 6.8, the loss of **18** to protodeboronation at 110°C was indeed significant, with only 20% of the original boronic acid remaining after 5 min—hence the requirement for fast activation kinetics of the catalyst. We then questioned whether controlled release of the boronic acid—by virtue of starting with the boronic pinacol ester—would neutralize the advantage of using ligands **L1-L3** in this case study. In collaboration with Yiming Wang, the boronic acid pinacol ester of **18** was synthesized and studied kinetically in comparison to **18**. Sure enough, the amount of **18** present starting with the boronic acid pinacol ester **21** was observed to remain between 50% and 90% over the course of 10 min. Reaction of **21** with **16** in the presence of 1.2% **P1-L5** (Scheme 6.8) was observed to give a much-improved yield of 85% in 10 min at 110°C, whereas the yield found from use of **P1-L1** with the ester remained consistent with the previous optimization at 87%.



Scheme 6.7. Suzuki-Miyaura cross-coupling of 3-chloropyridine and benzofuran-2-boronic acid pinacol ester.

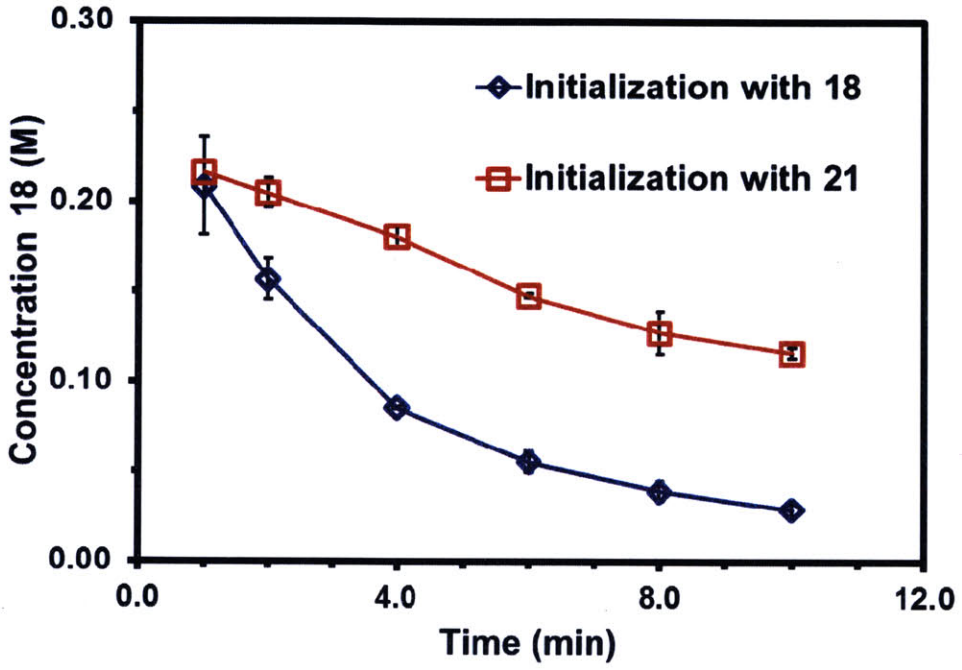


Figure 6.8. Observed HPLC concentration of 18 at 110°C starting with benzofuran-2-boronic acid (18) and benzofuran-2-boronic acid pinacol ester (21).

The optimal synthesis of **10** involved a three-component competition between activation of the catalyst complex, the protodeboronation of **7**, and the instability of the product. Based on the similarity of the optimal response surface (Figure 6.5a) to the simulation results in Figure 4.4, it was evident in this situation that the primary competition was between catalyst activation and product instability. Here the true advantage of a feedback flow optimization system was realized, as ligands were discriminated based on catalyst activity while the system simultaneously pinpointed a controlled reaction time and temperature for maximum product yield. As in the prior boronic acid case, the dialkylbiarylphosphine ligands were found to be more efficient than **L4** and trialkylphosphine ligands **L5-L7**. **P1** also outperformed **P2** in this and all other cases, consistent with the observations of Bruno *et al.*[245]

6.6. CONCLUSIONS

With the integration of automation and online analytics into cross-coupling reactions in flow, we have demonstrated a systematic methodology for both optimizing reactions and extracting key mechanistic insights. In 1-2 days of automated experiments, a single Suzuki-Miyaura cross-coupling reaction can now be assessed, optimized, and debugged, with precious resources such as the substrate, palladium, and ligand consumed in milligram quantities over the span of the full optimization. The reproducible results generated by the system further allow for the opportunity to test hypotheses and draw unbiased conclusions about rate-limiting reaction steps and process variables.

Our studies herein have shown the importance of dialkylbiarylphosphine ligands in promoting high catalytic turnover in the presence of aryl chloride substrates and unstable boronic acids and products. With boronic acid pinacol esters, where the rate of hydrolysis to the boronic acid can be limiting, the advantage of using the dialkylbiarylphosphine ligand may be neutralized. The same observation applies for the substitution of aryl bromides for aryl chlorides, where the reaction can transition to being oxidative addition-limited to transmetallation-limited. Heuristically, the data of these case studies show that the choice of the **P1-L1** precatalyst is favorable in most cases, but optimality can certainly not be guaranteed without a comprehensive search of all palladium precursor-ligand combinations.

Given the abundance of reaction insights that can be gleaned from the simultaneous study of only a few discrete and continuous variables, it can only be anticipated that future studies incorporating more variables—for instance candidate bases, solvents, and further manipulation of reagent equivalents—will allow even more extensive insight into design of the catalytic system for improved performance. These investigations will require both more intelligent ways of exploring discrete variables (compared to enumeration) and ways to incorporate learning from one optimization to the next. It was limiting in these optimization studies to repeatedly start with a design of experiments initialization and no prior preferentiality among different precatalysts and ligands. Future revitalizations of this work will in turn bring together chemoinformatics, PCA, and Bayesian statistics to prioritize discrete and continuous variables given the chemical properties of the substrates to greatly expedite the optimization of library syntheses and the resolution of more complex reaction mechanisms. It is clear that one day these automated tools will be working alongside the experimentalist solve chemistry's most challenging problems.

7. Conclusions and Future Research Directions

7.1. Introduction

The increased role of automation in chemical discovery and development has aroused the interests and sparked the curiosities of academics and industrials alike. In 2013, Elvira *et al.* posed seven goals of microfluidic research, and in doing so placed specific emphasis upon the advancement of automation.[268] With regard to automated reaction optimization, the authors challenged the microfluidic community to be able use integrated flow devices to map reactions and products finely over an exhaustive range of experimental conditions. This challenge resonated with us, not because the goal was insurmountable, but because the overarching goal of automated flow systems in this thesis has always been to find a way to make process development and scale-up faster, not bigger. An essential part of intelligence is the ability to maximize the use of resources and solve problems as directly as possible; likewise automation should be a tool that thinks and assists in that process. Put differently: if hiring a researcher to work in a chemistry lab, would we hire the one who can run the most experiments in the shortest time, or the one who draws upon past knowledge and inferences to make the smartest decision on what experiment to run next? It would only seem logical to have the same expectations for automated systems.

In another article in 2014, Peplow posed the idea of the robo-chemist,[269] a machine that could one day synthesize any compound automatically at the touch of a button. Current research is already underway build such a machine for different classes of reactions.[270-272] For discovery chemists, an automated synthesis tool would open the door to billions upon billions of yet-unsynthesized molecules. Chemists and engineers concerned with scale-up, however, should recognize that making the molecules may be a much simpler task than making the molecules in a cost-effective, scalable manner. The incorporation of flow chemistry offers one very viable option to facilitating scalability of small scale syntheses, for all of the reasons presented in Chapter 1. The need for smarter algorithms and analytical tools that can mine through past data and collect new data is, however, a recurrent theme that will need to be addressed to make automated synthesis a reality.

By bringing together discrete and continuous variables in automated optimization, we have illustrated a tool with the potential both to screen a diverse array of chemistries—as required for

automated synthesis—and learn about the chemistry—as required for accelerated reaction characterization and scale-up. The challenges posed by the Elvira *et al.* and Peplow articles, in truth, are not for a system that maps the reaction space fully or makes every conceivable molecule; rather they are challenges for chemists and engineers to look critically at organic chemistry and question what radical changes can be made that will ultimately reshape the way materials are made and drugs are designed. The key takeaways from both articles are that synthetic chemistry is constantly searching for generalizable methods that access a more diverse scope of molecules and conditions in higher throughput. We have shown through examples how feedback optimization of complex chemical systems offers one contribution in advancement this ongoing goal.

7.2. SUMMARY OF THESIS CONTRIBUTIONS

Several goals were introduced at the beginning of this thesis, which constituted both physical and conceptual advances to the field of feedback automation:

- Expansion of the scope of automated kinetic parameter estimation
- Design and implementation of an automated segmented flow system capable of screening a diverse range of liquid-phase reactions
- Implementation of an algorithm that solves in real-time the MINLP of simultaneous discrete variable screening and continuous variable screening optimization
- Application of the screening system to simultaneous ligand, catalyst, and solvent selection and reaction optimization
- Demonstration of scale-up of optimized reaction conditions

In large, the goal was to develop tools that would simplify and accelerate reaction development for a wide range of flow chemistry applications.

We began by exploring estimation of reaction kinetics in a series-parallel SNAr reaction pathway, which in determining four sets of pre-exponential factors and activation energies pushed the limits of what could be achieved with continuous variable feedback optimization. We demonstrated in this example the impressive accuracy of flow systems; however it was also clear that the sensitivity of the continuous flow system design to changes in the model parameters was insufficient to estimate all kinetic parameters with low uncertainty. Rather than estimating all model parameters simultaneously, we found that decoupling the system by studying the kinetics

of isolated reaction steps greatly reduced model uncertainty, with the limitation that intermediates had to be synthesized and isolated. Using the preliminary kinetics from the flow system, we were able to optimize the production of the unfavorable intermediate and use both reaction intermediates as starting materials in later parameter estimation studies.

The limitation posed by optimization over exclusively continuous variables motivated design and construction of a new system that would allow for optimization over both discrete variables and continuous variables. To accomplish this goal, we constructed a segmented flow microfluidic system that with the service of a liquid handling robot enabled different reactants, catalysts, and solvents to be paired together, mixed and reacted in a microfluidic slug, and analyzed online with feedback. We overcame numerous design challenges in producing a versatile system. Reagents were sampled and mixed accurately by aspirating and withdrawing samples repeatedly under an inert atmosphere. Transport of the slug was accomplished most effectively and with minimal carryover with the use of a compressed inert gas flowing through in an FEP tube. A custom-designed aluminum chuck with a groove for the FEP tube that sealed with polycarbonate to 6.9 bar prevented gas permeation from the system and promoted rapid heating and cooling. Automated syringe pumps connected to the flow system at T-junctions allowed reagents to be added to the slugs online for reaction initiation and quenching. Splitting of the slug into a 1 μL sample that was further partitioned in an online LC/MS enabled on-demand reaction analysis. The control to execute reactions on-demand was achieved with LabView. Studies showed the system to be reproducible over the course of many repeated experiments and to have a slug-to-slug carryover of approximately 3%.

Though the system as designed was advantageous for HTE, we sought as well to incorporate automated feedback into the optimization of reactions. As an initial attempt at black box optimization of discrete and continuous variable reaction systems, we developed and assessed a sequential response-surface based optimization algorithm. The algorithm employed a design of experiments approach to preliminarily characterize the experimental design space, then constructed response surface models for the performance of the objective with respect to linear, interaction, and quadratic effects in the continuous variables. Discrete variable response surfaces shared these common response surface terms but were distinguished with intercept and temperature-specific terms. Using the best-fit response surfaces, a lower bound on the overall optimum was established at each optimization iteration; discrete variables which were not

predicted to outperform this lower bound were fathomed from the next optimization iteration, akin to B&B. The response surfaces were then re-adapted for the exclusion of these discrete variables. To minimize the uncertainty on the optimum the fastest, a G-optimality criterion was implemented to select new on-demand experiments. This algorithm was demonstrated to be successful in a great majority of bimolecular, parallel, and series reaction pathway examples. We found the local search aspect of the G-optimality criterion to be limiting depending on the quality of the design of experiments initialization, and the method struggled when one discrete variable reaction rate law differed from the others.

The system and numerical methodology was applied to systems for solvent optimization and catalyst-ligand optimization. We considered optimal solvent selection for the benzylation of *trans*-1,2-diaminocylohexane, which is difficult to control in batch because of the propensity to over-alkylate at either amine position. Using the segmented flow system and the developed algorithm, we identified the solvent (DMSO) and reaction conditions that would allow for controlled alkylation at a single amine position and used these conditions in scale-up to make 500 mg product in batch. Similarly, we explored the optimization of several examples of single-phase Suzuki-Miyaura cross-coupling reactions with respect to the palladacycle-ligand complex and the reaction conditions. Generally, we found our method advantageous in rapidly identifying conditions where catalyst TON could be maximized given a constraint on yield. We observed cases where ligands which performed well with one substrate failed with the next, where the optimal ligand at one temperature was not optimal at a different temperature, and where an internal temperature-reaction time-catalyst loading optimum was found.

What we did not necessarily anticipate in accomplishing these goals was the learning with regard to the chemistry that would come from each reaction study. By merging discrete and continuous variables, we were able to explore to a greater depth the similarities and differences among optimal solutions. In the solvent optimization case, our system and response surface approach in a relatively limited number of experiments identified commonalities among preferential solvents that we later correlated back to solvent polarity and, more specifically, solvent hydrogen bond basicity. Similar discoveries were made in the Suzuki-Miyaura case studies. Depending on the choice of aryl halide, we observed transitioning of the rate limiting reaction step from oxidative addition to transmetallation, as evidenced by ligand preference and optimal TON. Using our system, we were able to construct kinetic profiles for the evolution of a

boronic acid and boronic acid pinacol ester, allowing visualization of the effect of protodeboronation on reaction yield. From the optimal response surface, we were also able to diagnose the competing effects of product degradation and catalyst activation and justify without bias the advantages of using dialkylbiarylphosphine ligands in the case of an unstable boronic acid and Suzuki-Miyaura product.

Though these are the primary thesis contributions, it would be unfortunate to overlook the other contributions this research has made in the directions of flow chemistry and automation. It was only a few years ago that automation was a luxury in process development, designed for demonstration purposes in specialized cases. Now the use of automation is prolific in flow chemistry examples,[273-276] simply because it enables experiments to run longer, makes analysis easier—especially at hazardous conditions—, and allows researchers more time to focus on more challenging chemistry or process engineering tasks. By collaborating actively with process chemists, we are also finding ways to engineer systems in flow without having to manipulate the underlying chemistry, a key aspect that will have to progress forward in order for flow chemistry to gain greater acceptance in both academic and industrial labs. Essential to that goal is the increased versatility of reactors, reagent delivery, and online analytical methods, all of which have been addressed to some extent in this thesis.

7.3. Future Research Directions

Relating back to the introduction of this chapter, continued work in automated feedback reaction systems will, in the most general sense, focus upon throughput acceleration and increased versatility, all while attempting to bring greater accessibility to the chemists and engineers looking to use the technology. Whether nanomolar reagent thresholds, sub-hour optimization times, or complete online scale-up and isolation of product are needed in future years will be determined by the demands of the pharmaceutical and fine chemical industries. In all of these cases, it will be important to balance the collection of large quantities of data with the quality and scalability of the data collected.

In terms of immediate enhancements to be made to the system presented in this thesis, further exploration into the optimization of multi-phase reactions (in particular liquid-liquid) is needed. This will require use of oscillatory slug flow as presented in a recently.[166] Future investigations will as well focus on the handling of less stable starting reagents—perhaps those with half-lives

of a few hours or less or ones which are insoluble at high concentration. Use of a glove box or glove bag for inert sample preparation will be far more robust than the manifold presented in Chapter 3. Additionally, it will be advantageous to provide active cooling to reagents in the liquid handler or to the online addition reagent for extended thermal stability. Technologies also exist for gravimetric handling of solid reagents. Though these may not be able to dose accurately to the microgram scale, as needed for individual slugs, it may be possible to incorporate this technology into the formation of small on-demand reaction batches that could turn into multiple slug injections, still accelerating sample preparation. Of course the incorporation of online analysis such as FTIR could greatly accelerate reaction modeling on a case-by-case basis.

We have postulated one algorithmic approach for simultaneous discrete and continuous variable optimization. Surely many other approaches exist, including pattern search methods and evolutionary methods.[277-280] Even in the subset of optimal experimental design methodologies, it would be intriguing to explore by comparison other optimality approaches such as D-optimality or I-optimality, which has a similar objective to G-optimality but examines the collective uncertainty of multiple points.[201] Perhaps these strategies will be less subjective to convergence to local minima or more robust to variations in reaction mechanisms. Ultimately it will be important to examine the best way to optimize multiple classes of discrete variables. In the simplest case, all possible combinations of discrete variables could be tested in enumeration, as was presented with palladacycle-ligand pairs in Chapter 6. But it may well be the case that this is an inefficient approach, and an experimental design approach more along the lines of a Latin Hypercube for the pairing of different discrete variables may result in accelerated convergence.[183]

Though these are worthy explorations in themselves, the true test of this method will come in the conceptual advances it will facilitate. More facile flow mechanistic studies are now enabled with this system, given the user's ability to characterize the reaction with respect to both discrete and continuous variables. Kinetic explorations like those in Chapter 2 are certainly attainable, and time- and temperature-course data have been shown to be collected automatically in Chapter 6. In the long term, it is easy to foresee research in this field expanding into *in silico* reaction modeling and chemoinformatics, as it will be essential in time savings for the feedback algorithm to be able to discern the common attributes shared by more optimal discrete variables and identify other unscreened discrete variables which share those optimal attributes. Collaborations

between programmers and process and medicinal chemists will help immensely in bringing the software to the intelligence level of a "robo-chemist." On-demand library development and product isolation are also achievable with this system with collection and purification of individual slugs by preparatory HPLC, leading to the possibility of biological assays of flow-synthesized compounds.[281]

In summary, the outlook for automated, segmented flow reaction optimization systems is positive, given the need both in academia and industry for tools that can rapidly and accurately characterize reactions using minimal amounts of material. Attention must always be paid to the versatility of the systems in terms of reagent tolerance and the sensitivity achievable for measured responses to manipulation in the process variables. In terms of automation and simplicity, the adage "set it and forget it" is a foremost goal to have in mind. It is with the continued advancement of automated reaction characterization systems that researchers will begin to gain much faster insight into reaction mechanisms, sparking faster reaction optimization and scale-up, and truly achieving the vision of "accelerated reaction development."

8. NOMENCLATURE

Table 8.1. Table of Latin nomenclature.

Symbol	Definition
$\mathbf{A}$	Matrix of response surface parameters
A	Cross-sectional area
A_i	Pre-exponential factor for discrete variable i
A_R	Pre-exponential factor for k_R
A_{S1}	Pre-exponential factor for k_{S1}
A_{S2}	Pre-exponential factor for k_{S2}
A_r	Pre-exponential factor for rate constant k_r
a_{ij}	Response surface model parameter
a'_{ij}	Response surface model parameter
$\mathbf{b}$	Vector of measured responses
$\mathbf{b}_u$	Vector of measured responses with experiment u removed from data set
b_u	Measured objective function (response) value for experiment u
$\hat{b}$	Model-predicted response
$C_{0,MeOBnCl}$	Initial concentration of 4-methoxybenzyl chloride
C_{A0}	Initial concentration of limiting reagent A
C_{A0u}	Initial concentration of limiting reagent A in experiment u
C_{B0}	Initial concentration of reagent B
C_{cat}	Catalyst concentration
C_R	Outlet concentration of desired product R
C_{Ru}	Outlet concentration of desired product R in experiment u
C_{S1}	Outlet concentration of undesired product S_1
$\mathbf{C}_u$	Vector of concentrations C_{iu}
$\hat{C}_{iu}$	Model-predicted concentration of species **1**, **3**, **4**, and **5** in experiment u
$\hat{\mathbf{C}}_u$	Vector of predicted concentrations $\hat{C}_{iu}$
$\mathbf{c}$	Vector of response surface parameters
c_i	Response surface model parameter
c'_i	Response surface model parameter
Ca	Capillary number
E_{Ai}	Activation energy for discrete variable i
E_{AR}	Nominal activation energy for k_R
E_{Ar}	Activation energy for rate constant k_r
E_{AS1}	Activation energy for k_{S1}
E_{AS2}	Activation energy for k_{S2}
$F_{\alpha,\nu1,\nu2}$	F distribution with α confidence and ν_1 and ν_2 degrees of freedom
f	Objective function
f_i	Rate of formation for species i
G_i	G-optimal value for discrete variable i
g	Inequality constraint
g_c	Gravitational constant
$\mathbf{g}_q$	Gradient at iteration q
H_0	Null hypothesis
H_a	Alternative hypothesis
$\mathbf{H}_q$	Hessian at iteration q
J^*	Maximal maximum of f for all discrete variables
J_-^*	Lower bound on J^*
$J_{-pred,q}^*$	Predicted J_-^* through iteration q
J_{-q}^*	J_-^* through iteration q
J_i'	Optimal predicted response value at candidate optimum $\mathbf{x}_i$
J_i^*	Maximum of f for discrete variable y_i

Table 8.1. (cont.) Table of Latin nomenclature.

Symbol	Definition
J_n*	Minimal maximum of f for all discrete variables
$J_{pred,q}$*	Predicted J* through iteration q
J_q*	J* through iteration q
k_R	Rate constant for A + B → R
k_r	Rate constant for reaction r
k_{S1}	Rate constant for B → S_1
k_{S2}	Rate constant for B + R → S_2
m	Rate law power dependence of C_{A0}
N_{cv}	Number of continuous variables
N_{dv}	Number of discrete variables
N_{expts}	Number of experiments
N_{expts}'	Number of experiments used to calculate J_i'
N_{expts}*	Number of experiments used to calculate J_i*
N_{params}	Number of parameters
N_{prior}	Number of prior experiments
N_{resp}	Number of measured responses
n	Rate law power dependence of C_{B0}
p	Rate law power dependence of C_{cat}
$\mathbf{p}_q$	Search direction for line search at iteration q
q	Optimization iteration index
R	Gas constant (8.314 J/mol K)
s_{ij}^2	Covariance of responses i and j
$\mathbf{s}_q$	Scaled search direction at iteration q
T	Reaction temperature
T*	Scaling temperature for parameter optimization
t	Time (instantaneous)
t_{crit}	Critical t-statistic
t_{res}	Residence time
t_{stat}	Calculated t-statistic
$t_{\alpha,\nu}$	Student's t value for 1-α confidence and ν degrees of freedom
U	Bulk fluid velocity
$\mathbf{u}$	Vector of experimental conditions
$\mathbf{V}_{MAP}$	*a posteriori* parameter covariance matrix
$\mathbf{V}_B$	Response covariance matrix
V_B	Scalar response covariance
$V_{\hat{B}}$	Prediction covariance
$V_{\hat{B}i}$	Prediction covariance for discrete variable i
V_{Bu}	Response covariance with experiment u removed from data set
$V_{J'}$	Prediction covariance matrix at candidate optimum $\mathbf{x}_i$
V_{J*}	Prediction covariance matrix at the optimum x*
$\mathbf{V}_\mu$	*a priori* parameter covariance matrix
$\mathbf{W}$	Weighting matrix for least-squares regression
$\mathbf{W}_u$	Weighting matrix for experiment u in MLE
$\mathbf{X}$	Matrix of scaled experimental conditions (also sensitivity matrix)
$\mathbf{X}_1$	Matrix of scaled experimental conditions augmented with new candidate experiment
$\mathbf{X}_u$	Sensitivity coefficient matrix for experiment u
$\mathbf{X'}_u$	Matrix of scaled experimental conditions with experiment u removed from data set
$\mathbf{x}$	Vector of scaled continuous variables
$\mathbf{x}$*	Optimal vector of continuous variables for f
$\mathbf{x}$'	Optimal vector of continuous variables at yield optimum

Table 8.1. (cont.) Table of Latin nomenclature.

Symbol	Definition
$\mathbf{x}_i$	Vector of scaled continuous variables for solvent *i*
$(x_i)_j$	Continous variable *j* for solvent *i*
$\mathbf{x}_i'$	Optimal vector of continuous variables at yield optimum for discrete variable *i*
$\mathbf{x}_i^*$	Optimal vector of continuous variables for *f* for discrete variable *i*
$(x_i^*)_j$	Optimal continous variable *j* for solvent *i*
x_j	Continuous variable *j*
Y	Yield objective function
$\mathbf{y}$	Vector of scaled discrete variables
$\mathbf{y}^*$	Optimal vector of discrete variables for *f*
$\mathbf{y}'$	Optimal vector of discrete variables at yield optimum
$\mathbf{y}_i^*$	Optimal vector of discrete variables for *f* for discrete variable *i*
$\mathbf{y}_i'$	Optimal vector of discrete variables at yield optimum for discrete variable *i*
y_i	Discrete variable *i*
$\mathbf{y}_q$	Gradient difference at iteration *q*
$\mathbf{Z}$	Fisher information matrix

Table 8.2. Table of Greek nomenclature.

Symbol	Definition
α	Rejection confidence level specified for Student's *t*-test
γ	Fraction of maximum yield that satisfies yield constraint of turnover number optimum
ΔJ_i^*	Error on the optimal response for solvent *i*
ΔV_g	Change in gas volume
ΔV_i	Aspirated volume fluid *i*
ΔV_{tf}	Change in transfer fluid volume
$\Delta\mathbf{x}$	Tolerance on the vector of optimal continuous variables
Δx_j	Tolerance on the optimum for continuous variable *j*
δ	Line search step size
ε	Dielectric constant
η	Viscosity
$\boldsymbol{\theta}$	Vector of model parameters
$\boldsymbol{\theta}_{opt}$	Optimal vector of model parameters
$\boldsymbol{\theta}_u$	Vector of optimal model parameters with experiment *u* removed from data set
θ_p	Model parameter *p*
$\boldsymbol{\mu}$	*a priori* vector of optimal model parameters
μ_d	Dipole moment
$\boldsymbol{\rho}$	Correlation matrix
ρ_g	Density of gas
ρ_i	Density of fluid *i*
ρ_{tf}	Transfer fluid density
σ	Interfacial tension

9. REFERENCES

(1) "Tufts Center for the Study of Drug Development. Cost to Develop and Win Marketing Approval for a New Drug Is $2.6 Billion." **2014**.

(2) Poechlauer, P.; Manley, J.; Broxterman, R.; Gregertsen, B.; Ridemark, M. "Continuous Processing in the Manufacture of Active Pharmaceutical Ingredients and Finished Dosage Forms: An Industry Perspective." *Org. Process Res. Dev.* **2012**, *16*, 1586.

(3) Baxendale, I. R.; Braatz, R. D.; Hodnett, B. K.; Jensen, K. F.; Johnson, M. D.; Sharratt, P.; Sherlock, J.-P.; Florence, A. J. "Achieving Continuous Manufacturing: Technologies and Approaches for Synthesis, Workup, and Isolation of Drug Substance. May 20-21, 2014 Continuous Manufacturing Symposium." *J. Pharm. Sci.* **2015**, *104*, 781.

(4) Anderson, N. G. "Using Continuous Processes to Increase Production." *Org. Process Res. Dev.* **2012**, *16*, 852.

(5) Sahoo, H. R.; Kralj, J. G.; Jensen, K. F. "Multistep Continuous-Flow Microchemical Synthesis Involving Multiple Reactions and Separations." *Angew. Chem.-Int. Edit.* **2007**, *46*, 5704.

(6) Pellegatti, L.; Sedelmeier, J. "Synthesis of Vildagliptin Utilizing Continuous Flow and Batch Technologies." *Org. Process Res. Dev.* **2015**.

(7) Johnson, M. D.; May, S. A.; Calvin, J. R.; Remacle, J.; Stout, J. R.; Diseroad, W. D.; Zaborenko, N.; Haeberle, B. D.; Sun, W.-M.; Miller, M. T.; Brennan, J. "Development and Scale-Up of a Continuous, High-Pressure, Asymmetric Hydrogenation Reaction, Workup, and Isolation." *Org. Process Res. Dev.* **2012**, *16*, 1017.

(8) Kockmann, N.; Roberge, D. M. "Harsh Reaction Conditions in Continuous-Flow Microreactors for Pharmaceutical Production." *Chem. Eng. Technol.* **2009**, *32*, 1682.

(9) Newman, S. G.; Jensen, K. F. "The Role of Flow in Green Chemistry and Engineering." *Green Chem.* **2013**, *15*, 1456.

(10) Vaccaro, L.; Lanari, D.; Marrocchi, A.; Strappaveccia, G. "Flow Approaches towards Sustainability." *Green Chem.* **2014**, *16*, 3680.

(11) Ley, S. V. "On Being Green: Can Flow Chemistry Help?" *Chem. Rec.* **2012**, *12*, 378.

(12) Poechlauer, P.; Colberg, J.; Fisher, E.; Jansen, M.; Johnson, M. D.; Koenig, S. G.; Lawler, M.; Laporte, T.; Manley, J.; Martin, B.; O'Kearney-McMullan, A. "Pharmaceutical Roundtable Study Demonstrates the Value of Continuous Manufacturing in the Design of Greener Processes." *Org. Process Res. Dev.* **2013**, *17*, 1472.

(13) Hartman, R. L.; McMullen, J. P.; Jensen, K. F. "Deciding Whether To Go with the Flow: Evaluating the Merits of Flow Reactors for Synthesis." *Angew. Chem.-Int. Edit.* **2011**, *50*, 7502.

(14) Murphy, E. R.; Martinelli, J. R.; Zaborenko, N.; Buchwald, S. L.; Jensen, K. F. "Accelerating Reactions with Microreactors at Elevated Temperatures and Pressures: Profiling Aminocarbonylation Reactions." *Angew. Chem.-Int. Edit.* **2007**, *46*, 1734.

(15) Snead, D. R.; Jamison, T. F. "A Three-Minute Synthesis and Purification of Ibuprofen: Pushing the Limits of Continuous-Flow Processing." *Angew. Chem.-Int. Edit.* **2015**, *54*, 983.

(16) Hessel, V.; Kralisch, D.; Kockmann, N.; Noel, T.; Wang, Q. "Novel Process Windows for Enabling, Accelerating, and Uplifting Flow Chemistry." *ChemSusChem* **2013**, *6*, 746.

(17) Mascia, S.; Heider, P. L.; Zhang, H. T.; Lakerveld, R.; Benyahia, B.; Barton, P. I.; Braatz, R. D.; Cooney, C. L.; Evans, J. M. B.; Jamison, T. F.; Jensen, K. F.; Myerson, A. S.; Trout, B. L. "End-to-End Continuous Manufacturing of Pharmaceuticals: Integrated Synthesis, Purification, and Final Dosage Formation." *Angew. Chem.-Int. Edit.* **2013**, *52*, 12359.

(18) Jensen, K. F. "Microreaction Engineering - Is Small Better?" *Chem. Eng. Sci.* **2001**, *56*, 293.

(19) Jensen, K. F. "Silicon-Based Microchemical Systems: Characteristics and Applications." *MRS Bull.* **2006**, *31*, 101.

(20) Hartman, R. L.; Jensen, K. F. "Microchemical Systems for Continuous-Flow Synthesis." *Lab Chip* **2009**, *9*, 2495.

(21) Jensen, K. F.; Reizman, B. J.; Newman, S. G. "Tools for Chemical Synthesis in Microsystems." *Lab Chip* **2014**, *14*, 3206.

(22) McMullen, J. P.; Stone, M. T.; Buchwald, S. L.; Jensen, K. F. "An Integrated Microreactor System for Self-Optimization of a Heck Reaction: From Micro- to Mesoscale Flow Systems." *Angew. Chem.-Int. Edit.* **2010**, *49*, 7076.

(23) McMullen, J. P.; Jensen, K. F. "Rapid Determination of Reaction Kinetics with an Automated Microfluidic System." *Org. Process Res. Dev.* **2011**, *15*, 398.

(24) Zhang, Y. J.; Born, S. C.; Jensen, K. F. "Scale-Up Investigation of the Continuous Phase-Transfer-Catalyzed Hypochlorite Oxidation of Alcohols and Aldehydes." *Org. Process Res. Dev.* **2014**, *18*, 1476.

(25) Zaborenko, N.; Bedore, M. W.; Jamison, T. F.; Jensen, K. F. "Kinetic and Scale-Up Investigations of Epoxide Aminolysis in Microreactors at High Temperatures and Pressures." *Org. Process Res. Dev.* **2011**, *15*, 131.

(26) Woitalka, A.; Kuhn, S.; Jensen, K. F. "Scalability of Mass Transfer in Liquid-Liquid Flow." *Chem. Eng. Sci.* **2014**, *116*, 1.

(27) Nagy, K. D.; Shen, B.; Jamison, T. F.; Jensen, K. F. "Mixing and Dispersion in Small-Scale Flow Systems." *Org. Process Res. Dev.* **2012**.

(28) Keybl, J.; Jensen, K. F. "Microreactor System for High-Pressure Continuous Flow Homogeneous Catalysis Measurements." *Ind. Eng. Chem. Res.* **2011**, *50*, 11013.

(29) Moore, J. S.; Jensen, K. F. ""Batch" Kinetics in Flow: Online IR Analysis and Continuous Control." *Angew. Chem.-Int. Edit.* **2014**, *53*, 470.

(30) Reizman, B. J.; Jensen, K. F. "An Automated Continuous-Flow Platform for the Estimation of Multistep Reaction Kinetics." *Org. Process Res. Dev.* **2012**, *16*, 1770.

(31) Heublein, N.; Moore, J. S.; Smith, C. D.; Jensen, K. F. "Investigation of Petasis and Ugi Reactions in Series in an Automated Microreactor System." *RSC Adv.* **2014**, *4*, 63627.

(32) Born, S. C.; Jensen, K. F. "Risk Evaluation for the Use of Azide Reagents in Pharmaceutical Development: DPPA On-Demand." *Abstr. Pap. Am. Chem. S.* **2013**, *245*.

(33) Chambers, R. D.; Fox, M. A.; Sandford, G.; Trmcic, J.; Goeta, A. "Elemental Fluorine - Part 20. Direct Fluorination of Deactivated Aromatic Systems Using Microreactor Techniques." *J. Fluorine Chem.* **2007**, *128*, 29.

(34) Chambers, R. D.; Sandford, G.; Trmcic, J.; Okazoe, T. "Elemental Fluorine. Part 21. Direct Fluorination of Benzaldehyde Derivatives." *Org. Process Res. Dev.* **2008**, *12*, 339.

(35) Navarrini, W.; Venturini, F.; Tortelli, V.; Basak, S.; Pimparkar, K. P.; Adamo, A.; Jensen, K. F. "Direct Fluorination of Carbon Monoxide in Microreactors." *J. Fluorine Chem.* **2012**, *142*, 19.

(36) de Mas, N.; Gunther, A.; Schmidt, M. A.; Jensen, K. F. "Increasing Productivity of Microreactors for Fast Gas-Liquid Reactions: The Case of Direct Fluorination of Toluene." *Ind. Eng. Chem. Res.* **2009**, *48*, 1428.

(37) Ducry, L.; Roberge, D. M. "Controlled Autocatalytic Nitration of Phenol in a Microreactor." *Angew. Chem.-Int. Edit.* **2005**, *44*, 7972.

(38) Pelleter, J.; Renaud, F. "Facile, Fast and Safe Process Development of Nitration and Bromination Reactions Using Continuous Flow Reactors." *Org. Process Res. Dev.* **2009**, *13*, 698.

(39) Murray, P. R. D.; Browne, D. L.; Pastre, J. C.; Butters, C.; Guthrie, D.; Ley, S. V. "Continuous Flow-Processing of Organometallic Reagents Using an Advanced Peristaltic Pumping System and the Telescoped Flow Synthesis of (E/Z)-Tamoxifen." *Org. Process Res. Dev.* **2013**, *17*, 1192.

(40) Munoz, J. D.; Alcazar, J.; de la Hoz, A.; Diaz-Ortiz, A. "Application of Flow Chemistry to the Reduction of Nitriles to Aldehydes." *Tetrahedron Lett.* **2011**, *52*, 6058.

(41) Ducry, L.; Roberge, D. M. "DIBAL-H Reduction of Methyl Butyrate into Butyraldehyde Using Microreactors." *Org. Process Res. Dev.* **2008**, *12*, 163.

(42) Wu, J.; Yang, X. Q.; He, Z.; Mao, X. W.; Hatton, T. A.; Jamison, T. F. "Continuous Flow Synthesis of Ketones from Carbon Dioxide and Organolithium or Grignard Reagents." *Angew. Chem.-Int. Edit.* **2014**, *53*, 8416.

(43) Nagaki, A.; Takahashi, Y.; Yoshida, J. I. "Extremely Fast Gas/Liquid Reactions in Flow Microreactors: Carboxylation of Short-Lived Organolithiums." *Chem.-Eur. J.* **2014**, *20*, 7931.

(44) Shu, W.; Pellegatti, L.; Oberli, M. A.; Buchwald, S. L. "Continuous-Flow Synthesis of Biaryls Enabled by Multistep Solid-Handling in a Lithiation/Borylation/Suzuki-Miyaura Cross-Coupling Sequence." *Angew. Chem.-Int. Edit.* **2011**, *50*, 10665.

(45) Tomida, Y.; Nagaki, A.; Yoshida, J. "Asymmetric Carbolithiation of Conjugated Enynes: A Flow Microreactor Enables the Use of Configurationally Unstable Intermediates before They Epimerize." *J. Am. Chem. Soc.* **2011**, *133*, 3744.

(46) Deng, Q. L.; Shen, R. W.; Ding, R.; Zhang, L. X. "Generation of Ethynyl-Grignard Reagent in a Falling Film Microreactor: An Expeditious Flow Synthesis of Propargylic Alcohols and Analogues." *Adv. Synth. Catal.* **2014**, *356*, 2931.

(47) He, Z.; Jamison, T. F. "Continuous-Flow Synthesis of Functionalized Phenols by Aerobic Oxidation of Grignard Reagents." *Angew. Chem.-Int. Edit.* **2014**, *53*, 3353.

(48) Marre, S.; Park, J.; Rempel, J.; Guan, J.; Bawendi, M. G.; Jensen, K. F. "Supercritical Continuous-Microflow Synthesis of Narrow Size Distribution Quantum Dots." *Adv. Mater.* **2008**, *20*, 4830.

(49) Nishiyama, Y.; Mori, R.; Nishida, K.; Tanimoto, H.; Morimoto, T.; Kakiuchi, K. "Diastereodifferentiating 2+2 Photocycloaddition of a Chiral Cyclohexenone with Cyclopentene in Supercritical Carbon Dioxide Using a Flow Microreactor." *J. Flow Chem.* **2014**, *4*, 185.

(50) Trachsel, F.; Tidona, B.; Desportes, S.; von Rohr, P. R. "Solid Catalyzed Hydrogenation in a Si/Glass Microreactor Using Supercritical CO2 as the Reaction Solvent." *J. Supercrit. Fluid* **2009**, *48*, 146.

(51) McMullen, J. P.; Jensen, K. F. "An Automated Microfluidic System for Online Optimization in Chemical Synthesis." *Org. Process Res. Dev.* **2010**, *14*, 1169.

(52) Silva, B. V.; Violante, F. A.; Pinto, A. C.; Santos, L. S. "The Mechanism of Sandmeyer's Cyclization Reaction by Electrospray Ionization Mass Spectrometry." *Rapid Commun. Mass Sp.* **2011**, *25*, 423.

(53) Browne, D. L.; Wright, S.; Deadman, B. J.; Dunnage, S.; Baxendale, I. R.; Turner, R. M.; Ley, S. V. "Continuous Flow Reaction Monitoring Using an On-Line Miniature Mass Spectrometer." *Rapid Commun. Mass Sp.* **2012**, *26*, 1999.

(54) Goodell, J. R.; McMullen, J. P.; Zaborenko, N.; Maloney, J. R.; Ho, C. X.; Jensen, K. F.; Porco, J. A.; Beeler, A. B. "Development of an Automated Microfluidic Reaction

Platform for Multidimensional Screening: Reaction Discovery Employing Bicyclo[3.2.1]octanoid Scaffolds." *J. Org. Chem.* **2009**, *74*, 6169.

(55) Sans, V.; Glatzel, S.; Douglas, F. J.; Maclaren, D. A.; Lapkin, A.; Cronin, L. "Non-Equilibrium Dynamic Control of Gold Nanoparticle and Hyper-Branched Nanogold Assemblies." *Chem. Sci.* **2014**, *5*, 1153.

(56) Carter, C. F.; Lange, H.; Ley, S. V.; Baxendale, I. R.; Wittkamp, B.; Goode, J. G.; Gaunt, N. L. "ReactIR Flow Cell: A New Analytical Tool for Continuous Flow Chemical Processing." *Org. Process Res. Dev.* **2010**, *14*, 393.

(57) Qian, Z.; Baxendale, I. R.; Ley, S. V. "A Continuous Flow Process Using a Sequence of Microreactors with In-Line IR Analysis for the Preparation of N,N-Diethyl-4-(3-Fluorophenylpiperidin-4-ylidenemethyl)benzamide as a Potent and Highly Selective δ-Opioid Receptor Agonist." *Chem.-Eur. J.* **2010**, *16*, 12342.

(58) Moore, J. S.; Jensen, K. F. "Automated Multitrajectory Method for Reaction Optimization in a Microfluidic System using Online IR Analysis." *Org. Process Res. Dev.* **2012**, *16*, 1409.

(59) Liu, X.; Unal, B.; Jensen, K. F. "Heterogeneous Catalysis with Continuous Flow Microreactors." *Cat. Sci. Tec.* **2012**, *2*, 2134.

(60) Mozharov, S.; Nordon, A.; Littlejohn, D.; Wiles, C.; Watts, P.; Dallin, P.; Girkin, J. M. "Improved Method for Kinetic Studies in Microreactors Using Flow Manipulation and Noninvasive Raman Spectrometry." *J. Am. Chem. Soc.* **2011**, *133*, 3601.

(61) Sans, V.; Porwol, L.; Dragone, V.; Cronin, L. "A Self Optimizing Synthetic Organic Reactor System Using Real-Time In-Line NMR Spectroscopy." *Chem. Sci.* **2015**, *6*, 1258.

(62) Davies, I. W.; Welch, C. J. "Looking Forward in Pharmaceutical Process Chemistry." *Science* **2009**, *325*, 701.

(63) Santanilla, A. B.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L. C.; Schneeweis, J.; Berritt, S.; Shi, Z. C.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. "Nanomole-Scale High-Throughput Chemistry for the Synthesis of Complex Molecules." *Science* **2015**, *347*, 49.

(64) Shultz, C. S.; Krska, S. W. "Unlocking the Potential of Asymmetric Hydrogenation at Merck." *Accounts Chem. Res.* **2007**, *40*, 1320.

(65) Rubin, A. E.; Tummala, S.; Both, D. A.; Wang, C. C.; Delaney, E. J. "Emerging Technologies Supporting Chemical Process R&D and Their Increasing Impact on Productivity in the Pharmaceutical Industry." *Chem. Rev.* **2006**, *106*, 2794.

(66) Schmink, J. R.; Bellomo, A.; Berritt, S. "Scientist-Led High-Throughput Experimentation (HTE) and Its Utility in Academia and Industry." *Aldrichim. Acta* **2013**, *46*, 71.

(67) Preshlock, S. M.; Ghaffari, B.; Maligres, P. E.; Krska, S. W.; Maleczka, R. E.; Smith, M. R. "High-Throughput Optimization of Ir-Catalyzed C-H Borylation: A Tutorial for Practical Applications." *J. Am. Chem. Soc.* **2013**, *135*, 7572.

(68) Bellomo, A.; Celebi-Olcum, N.; Bu, X.; Rivera, N.; Ruck, R. T.; Welch, C. J.; Houk, K. N.; Dreher, S. D. "Rapid Catalyst Identification for the Synthesis of the Pyrimidinone Core of HIV Integrase Inhibitors." *Angew. Chem.-Int. Edit.* **2012**, *51*, 6912.

(69) Friedman, M.; Savage, L. J. *Techniques of Statistical Analysis*; McGraw-Hill: New York, 1947.

(70) Wissmann, P. J.; Grover, M. A. "Optimization of a Chemical Vapor Deposition Process Using Sequential Experimental Design." *Ind. Eng. Chem. Res.* **2010**, *49*, 5694.

(71) de Castro, A. M.; Castilho, L. R.; Freire, D. M. G. "Multivariate Optimization and Supplementation Strategies for the Simultaneous Production of Amylases, Cellulases, Xylanases, and Proteases by Aspergillus awamori Under Solid-State Fermentation Conditions." *Appl. Biochem. Biotech.* **2015**, *175*, 1588.

(72) Kaith, B. S.; Sharma, R.; Kalia, S.; Bhatti, M. S. "Response Surface Methodology and Optimized Synthesis of Guar Gum-Based Hydrogels with Enhanced Swelling Capacity." *RSC Adv.* **2014**, *4*, 40339.

(73) Casciato, M. J.; Kim, S.; Lu, J. C.; Hess, D. W.; Grover, M. A. "Optimization of a Carbon Dioxide-Assisted Nanoparticle Deposition Process Using Sequential Experimental Design with Adaptive Design Space." *Ind. Eng. Chem. Res.* **2012**, *51*, 4363.

(74) Coetzer, R. L. J.; Morgan, D. H.; Maumela, H. "Optimization of a Catalyst System through the Sequential Application of Experimental Design Techniques." *J. Appl. Stat.* **2008**, *35*, 131.

(75) Duarte, A. R. C.; Ünal, B.; Mano, J. F.; Reis, R. L.; Jensen, K. F. "Microfluidic Production of Perfluorocarbon-Alginate Core–Shell Microparticles for Ultrasound Therapeutic Applications." *Langmuir* **2014**, *30*, 12391.

(76) Lee, W.-H. Ph.D., Massachusetts Institute of Technology, 2014.

(77) Newman, S. G.; Gu, L.; Lesniak, C.; Victor, G.; Meschke, F.; Abahmaneb, L.; Jensen, K. F. "Rapid Wolff-Kishner Reductions in a Silicon Carbide Microreactor." *Green Chem.* **2014**, *16*, 176.

(78) Welch, C. J.; Gong, X. Y.; Schafer, W.; Pratt, E. C.; Brkovic, T.; Pirzada, Z.; Cuff, J. F.; Kosjek, B. "MISER Chromatography (Multiple Injections in a Single Experimental Run): The Chromatogram is the Graph." *Tetrahedron-Asymmetr.* **2010**, *21*, 1674.

(79) Siegle, A. F.; Trapp, O. "Development of a Straightforward and Robust Technique to Implement Hadamard Encoded Multiplexing to High-Performance Liquid Chromatography." *Anal. Chem.* **2014**, *86*, 10828.

(80) McMullen, J. P. Ph.D., Massachusetts Institute of Technology, 2010.

(81) Krishnadasan, S.; Brown, R. J. C.; Demello, A. J.; Demello, J. C. "Intelligent Routes to the Controlled Synthesis of Nanoparticles." *Lab Chip* **2007**, *7*, 1434.

(82) Huyer, W.; Neumaier, A. "SNOBFIT--Stable Noisy Optimization by Branch and Fit." *ACM T. Math Software* **2008**, *35*, 9.

(83) Nelder, J. A.; Mead, R. "A Simplex Method for Function Minimization." *Comput. J.* **1965**, *7*, 308.

(84) Parrott, A. J.; Bourne, R. A.; Akien, G. R.; Irvine, D. J.; Poliakoff, M. "Self-Optimizing Continuous Reactions in Supercritical Carbon Dioxide." *Angew. Chem.-Int. Edit.* **2011**, *50*, 3788.

(85) Bourne, R. A.; Skilton, R. A.; Parrott, A. J.; Irvine, D. J.; Poliakoff, M. "Adaptive Process Optimization for Continuous Methylation of Alcohols in Supercritical Carbon Dioxide." *Org. Process Res. Dev.* **2011**, *15*, 932.

(86) Jumbam, D. N.; Skilton, R. A.; Parrott, A. J.; Bourne, R. A.; Poliakoff, M. "The Effect of Self-Optimisation Targets on the Methylation of Alcohols Using Dimethyl Carbonate in Supercritical CO2." *J. Flow Chem.* **2012**, *2*, 24.

(87) Skilton, R. A.; Parrott, A. J.; George, M. W.; Poliakoff, M.; Bourne, R. A. "Real-Time Feedback Control Using Online Attenuated Total Reflection Fourier Transform Infrared (ATR FT-IR) Spectroscopy for Continuous Flow Optimization and Process Knowledge." *Appl. Spectrosc.* **2013**, *67*, 1127.

(88) Skilton, R. A.; Bourne, R. A.; Amara, Z.; Horvath, R.; Jin, J.; Scully, M. J.; Streng, E.; Tang, S. L. Y.; Summers, P. A.; Wang, J.; Perez, E.; Asfaw, N.; Aydos, G. L. P.; Dupont, J.; Comak, G.; George, M. W.; Poliakoff, M. "Remote-Controlled Experiments with Cloud Chemistry." *Nat. Chem.* **2015**, *7*, 1.

(89) Routh, M. W.; Swartz, P. A.; Denton, M. B. "Performance of the Super Modified Simplex." *Anal. Chem.* **1977**, *49*, 1422.

(90) Fabry, D. C.; Sugiono, E.; Rueping, M. "Self-Optimizing Reactor Systems: Algorithms, On-Line Analytics, Setups, and Strategies for Accelerating Continuous Flow Process Optimization." *Isr. J. Chem.* **2014**, *54*, 341.

(91) Armijo, L. "Minimization of Functions Having Lipschitz Continuous First Partial Derivatives." *Pac. J. Math.* **1966**, *16*, 1.

(92) Box, G. E. P.; Hill, W. J. "Discrimination among Mechanistic Models." *Technometrics* **1967**, *9*, 57.

(93) Box, G. E. P.; Draper, N. R. "Baysian Estimation of Common Parameters from Several Responses." *Biometrika* **1965**, *52*, 355.

(94) Box, G. E. P.; Lucas, H. L. "Design of Experiments in Non-Linear Situations." *Biometrika* **1959**, *46*, 77.

(95) Song, H.; Chen, D. L.; Ismagilov, R. F. "Reactions in Droplets in Microflulidic Channels." *Angew. Chem.-Int. Edit.* **2006**, *45*, 7336.

(96) Hatakeyama, T.; Chen, D. L.; Ismagilov, R. F. "Microgram-Scale Testing of Reaction Conditions in Solution Using Nanoliter Plugs in Microfluidics with Detection by MALDI-MS." *J. Am. Chem. Soc.* **2006**, *128*, 2518.

(97) Li, L.; Mustafi, D.; Fu, Q.; Tereshko, V.; Chen, D. L. L.; Tice, J. D.; Ismagilov, R. F. "Nanoliter Microfluidic Hybrid Method for Simultaneous Screening and Optimization Validated with Crystallization of Membrane Proteins." *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 19243.

(98) Kreutz, J. E.; Shukhaev, A.; Du, W. B.; Druskin, S.; Daugulis, O.; Ismagilov, R. F. "Evolution of Catalysts Directed by Genetic Algorithms in a Plug-Based Microfluidic Device Tested with Oxidation of Methane by Oxygen." *J. Am. Chem. Soc.* **2010**, *132*, 3128.

(99) Clausell-Tormos, J.; Griffiths, A. D.; Merten, C. A. "An Automated Two-Phase Microfluidic System for Kinetic Analyses and the Screening of Compound Libraries." *Lab Chip* **2010**, *10*, 1302.

(100) Theberge, A. B.; Mayot, E.; El Harrak, A.; Kleinschmidt, F.; Huck, W. T. S.; Griffiths, A. D. "Microfluidic Platform for Combinatorial Synthesis in Picolitre Droplets." *Lab Chip* **2012**, *12*, 1320.

(101) Kaminski, T. S.; Jakiela, S.; Czekalska, M. A.; Postek, W.; Garstecki, P. "Automated Generation of Libraries of nL Droplets." *Lab Chip* **2012**, *12*, 3995.

(102) Niu, X.; deMello, A. J. "Building Droplet-Based Microfluidic Systems for Biological Analysis." *Biochem. Soc. T.* **2012**, *40*, 615.

(103) Zec, H.; Rane, T. D.; Wang, T.-H. "Microfluidic Platform for On-Demand Generation of Spatially Indexed Combinatorial Droplets." *Lab Chip* **2012**, *12*, 3055.

(104) Dressler, O. J.; Maceiczyk, R. M.; Chang, S. I.; deMello, A. J. "Droplet-Based Microfluidics Enabling Impact on Drug Discovery." *J. Biomol. Screen.* **2014**, *19*, 483.

(105) Kuster, S. K.; Pabst, M.; Zenobi, R.; Dittrich, P. S. "Screening for Protein Phosphorylation Using Nanoscale Reactions on Microdroplet Arrays." *Angew. Chem.-Int. Edit.* **2015**, *54*, 1671.

(106) Jeong, H.-H.; Jin, S. H.; Lee, B. J.; Kim, T.; Lee, C.-S. "Microfluidic Static Droplet Array for Analyzing Microbial Communication on a Population Gradient." *Lab Chip* **2015**, *15*, 889.

(107) Bogdan, A. R.; Sach, N. W. "The Use of Copper Flow Reactor Technology for the Continuous Synthesis of 1,4-Disubstituted 1,2,3-Triazoles." *Adv. Synth. Catal.* **2009**, *351*, 849.

(108) Bogdan, A. R.; James, K. "Efficient Access to New Chemical Space Through Flow-Construction of Druglike Macrocycles Through Copper-Surface-Catalyzed Azide-Alkyne Cycloaddition Reactions." *Chem.-Eur. J.* **2010**, *16*, 14506.

(109) Bogdan, A. R.; James, K. "Synthesis of 5-Iodo-1,2,3-Triazole-Containing Macrocycles Using Copper Flow Reactor Technology." *Org. Lett.* **2011**, *13*, 4060.

(110) Hochlowski, J. E.; Searle, P. A.; Tu, N. P.; Pan, J. Y.; Spanton, S. G.; Djuric, S. W. "An Integrated Synthesis-Purification System to Accelerate the Generation of Compounds in Pharmaceutical Discovery." *J. Flow Chem.* **2011**, *1*, 56.

(111) Lange, P. P.; Bogdan, A. R.; James, K. "A New Flow Methodology for the Expedient Synthesis of Drug-Like 3-Aminoindolizines." *Adv. Synth. Catal.* **2012**, *354*, 2373.

(112) Lange, P. P.; James, K. "Rapid Access to Compound Libraries through Flow Technology: Fully Automated Synthesis of a 3-Aminoindolizine Library via Orthogonal Diversification." *ACS Comb. Sci.* **2012**, *14*, 570.

(113) Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley Pub. Co.: Reading, MA, 1989.

(114) Zhang, Q.; Liu, Y.; Gao, F.; Ding, Q.; Cho, C.; Hur, W.; Jin, Y. H.; Uno, T.; Joazeiro, C. A. P.; Gray, N. "Discovery of EGFR Selective 4,6-Disubstituted Pyrimidines from a Combinatorial Kinase-Directed Heterocycle Library." *J. Am. Chem. Soc.* **2006**, *128*, 2182.

(115) Anderson, M.; Beattie, J. F.; Breault, G. A.; Breed, J.; Byth, K. F.; Culshaw, J. D.; Ellston, R. P. A.; Green, S.; Minshull, C. A.; Norman, R. A.; Pauptit, R. A.; Stanway, J.; Thomas, A. P.; Jewsbury, P. J. "Imidazo[1,2-a]pyridines: A Potent and Selective Class of Cyclin-Dependent Kinase Inhibitors Identified through Structure-Based Hybridisation." *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3021.

(116) Cumming, J. G.; McKenzie, C. L.; Bowden, S. G.; Campbell, D.; Masters, D. J.; Breed, J.; Jewsbury, P. J. "Novel, Potent and Selective Anilinoquinazoline and Anilinopyrimidine Inhibitors of p38 MAP Kinase." *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5389.

(117) Tavares, F. X.; Boucheron, J. A.; Dickerson, S. H.; Griffin, R. J.; Preugschat, F.; Thomson, S. A.; Wang, T. Y.; Zhou, H. Q. "N-Phenyl-4-Pyrazolo[1,5-b]Pyridazin-3-Ylpyrimidin-2-Amines as Potent and Selective Inhibitors of Glycogen Synthase Kinase 3 with Good Cellular Efficacy." *J. Med. Chem.* **2004**, *47*, 4716.

(118) Wang, S. D.; Meades, C.; Wood, G.; Osnowski, A.; Anderson, S.; Yuill, R.; Thomas, M.; Mezna, M.; Jackson, W.; Midgley, C.; Griffiths, G.; Fleming, I.; Green, S.; McNae, I.;

Wu, S. Y.; McInnes, C.; Zheleva, D.; Walkinshaw, M. D.; Fischer, P. M. "2-Anilino-4-(Thiazol-5-yl)pyrimidine CDK Inhibitors: Synthesis, SAR Analysis, X-Ray Crystallography, and Biological Activity." *J. Med. Chem.* **2004**, *47*, 1662.

(119) Zagulyaeva, O. A.; Bukhatkina, N. V.; Mamaev, V. P. "Relative Reactivity of Chlorine Atoms in 2,4-Dichloropyrimidine during Reactions with Ammonia and Amines in Isooctane and Ethanol." *Zhurnal Org. Khimii* **1978**, *14*, 409.

(120) Liu, M.; Wang, S. Y.; Clampit, J. E.; Gum, R. J.; Haasch, D. L.; Rondinone, C. M.; Trevillyan, J. M.; Abad-Zapatero, C.; Fry, E. H.; Sham, H. L.; Liu, G. "Discovery of a New Class of 4-Anilinopyrimidines as Potent c-Jun N-terminal Kinase Inhibitors: Synthesis and SAR Studies." *Bioorg. Med. Chem. Lett.* **2007**, *17*, 668.

(121) Delcorona, L.; Signorelli, G.; Manzardo, S.; Pinzetta, A.; Coppi, G. "Synthesis and Invitro Study of Platelet Antiaggregant Activity of 2(4)-Inidazol-1-yl-4(2)-Cycloalkylamino Pyrimidines." *Eur. J. Med. Chem.* **1991**, *26*, 729.

(122) Melander, L. "On the Mechanism of Electrophilic Aromatic Substitution - an Investigation by Means of the Effect of Isotopic Mass on Reaction Velocity." *Ark. Kemi* **1950**, *2*, 211.

(123) Bunnett, J. F.; Zahler, R. E. "Aromatic Nucleophilic Substitution Reactions." *Chem. Rev.* **1951**, *49*, 273.

(124) Bunnett, J. F. "Mechanism and Reactivity in Aromatic Nucleophilic Substitution Reactions." *Q. Rev. Chem. Soc.* **1958**, *12*, 1.

(125) Levenspiel, O. *Chemical Reaction Engineering*; 3rd ed.; Wiley: New York, 1999.

(126) Beck, J. V.; Arnold, K. J. *Parameter Estimation in Engineering and Science*; Wiley: New York, 1977.

(127) Steinfeld, J. I., Francisco, J. S., and Hase W. L. *Chemical Kinetics and Dynamics*; 2nd ed.; Prentice Hall: Upper Saddle River, N.J., 1999.

(128) Draper, N. R.; Smith, H. *Applied Regression Analysis*; Wiley: New York, 1981.

(129) Zaborenko, N.; Murphy, E. R.; Kralj, J. G.; Jensen, K. F. "Synthesis and Kinetics of Highly Energetic Intermediates by Micromixers: Direct Multistep Synthesis of Sodium Nitrotetrazolate." *Ind. Eng. Chem. Res.* **2010**, *49*, 4132.

(130) Allwardt, A.; Holzmuller-Laue, S.; Wendler, C.; Stoll, N. "A High Parallel Reaction System for Efficient Catalyst Research." *Catal. Today* **2008**, *137*, 11.

(131) Mills, P. L.; Nicole, J. F. "A Novel Reactor for High-Throughput Screening of Gas-Solid Catalyzed Reactions." *Chem. Eng. Sci.* **2004**, *59*, 5345.

(132) Richter, A.; Langpape, M.; Kolf, S.; Grubert, G.; Eckelt, R.; Radnik, J.; Schneider, A.; Pohl, M. M.; Fricke, R. "Combinatorial Preparation and High-Throughput Catalytic Tests of Multi-Component deNO(x) Catalysts." *Appl. Catal. B-Environ.* **2002**, *36*, 261.

(133) Shi, G. Y.; Hong, F.; Liang, Q. S.; Fang, H.; Nelson, S.; Weber, S. G. "Capillary-Based, Serial-Loading, Parallel Microreactor for Catalyst Screening." *Anal. Chem.* **2006**, *78*, 1972.

(134) Fernandez-Suarez, M.; Wong, S. Y. F.; Warrington, B. H. "Synthesis of a Three-Member Array of Cycloadducts in a Glass Microchip under Pressure Driven Flow." *Lab Chip* **2002**, *2*, 170.

(135) Schwalbe, T.; Kadzimirsz, D.; Jas, G. "Synthesis of a Library of Ciprofloxacin Analogues by Means of Sequential Organic Synthesis in Microreactors." *QSAR Comb. Sci.* **2005**, *24*, 758.

(136) Griffiths-Jones, C. M.; Hopkin, M. D.; Jonsson, D.; Ley, S. V.; Tapolczay, D. J.; Vickerstaffe, E.; Ladlow, M. "Fully Automated Flow-Through Synthesis of Secondary Sulfonamides in a Binary Reactor System." *J. Comb. Chem.* **2007**, *9*, 422.

(137) Wang, J.; Sui, G.; Mocharla, V. P.; Lin, R. J.; Phelps, M. E.; Kolb, H. C.; Tseng, H. R. "Integrated Microfluidics for Parallel Screening of an In Situ Click Chemistry Library." *Angew. Chem.-Int. Edit.* **2006**, *45*, 5276.

(138) Gunther, A.; Khan, S. A.; Thalmann, M.; Trachsel, F.; Jensen, K. F. "Transport and Reaction in Microscale Segmented Gas-Liquid Flow." *Lab Chip* **2004**, *4*, 278.

(139) Tice, J. D.; Lyon, A. D.; Ismagilov, R. F. "Effects of Viscosity on Droplet Formation and Mixing in Microfluidic Channels." *Anal. Chim. Acta* **2004**, *507*, 73.

(140) Song, H.; Bringer, M. R.; Tice, J. D.; Gerdts, C. J.; Ismagilov, R. F. "Experimental Test of Scaling of Mixing by Chaotic Advection in Droplets Moving through Microfluidic Channels." *Appl. Phys. Lett.* **2003**, *83*, 4664.

(141) Gunther, A.; Jhunjhunwala, M.; Thalmann, M.; Schmidt, M. A.; Jensen, K. F. "Micromixing of Miscible Liquids in Segmented Gas-Liquid Flow." *Langmuir* **2005**, *21*, 1547.

(142) Thorsen, T.; Roberts, R. W.; Arnold, F. H.; Quake, S. R. "Dynamic Pattern Formation in a Vesicle-Generating Microfluidic Device." *Phys. Rev. Lett.* **2001**, *86*, 4163.

(143) Nisisako, T.; Torii, T.; Higuchi, T. "Droplet Formation in a Microchannel Network." *Lab Chip* **2002**, *2*, 24.

(144) Anna, S. L.; Bontoux, N.; Stone, H. A. "Formation of Dispersions Using "Flow Focusing" in Microchannels." *Appl. Phys. Lett.* **2003**, *82*, 364.

(145) Xu, Q. Y.; Nakajima, M. "The Generation of Highly Monodisperse Droplets through the Breakup of Hydrodynamically Focused Microthread in a Microfluidic Device." *Appl. Phys. Lett.* **2004**, *85*, 3726.

(146) Zheng, B.; Roach, L. S.; Ismagilov, R. F. "Screening of Protein Crystallization Conditions on a Microfluidic Chip Using Nanoliter-Size Droplets." *J. Am. Chem. Soc.* **2003**, *125*, 11170.

(147) Shestopalov, I.; Tice, J. D.; Ismagilov, R. F. "Multi-Step Synthesis of Nanoparticles Performed on Millisecond Time Scale in a Microfluidic Droplet-Based System." *Lab Chip* **2004**, *4*, 316.

(148) Churski, K.; Korczyk, P.; Garstecki, P. "High-Throughput Automated Droplet Microfluidic System for Screening of Reaction Conditions." *Lab Chip* **2010**, *10*, 816.

(149) Churski, K.; Nowacki, M.; Korczyk, P. M.; Garstecki, P. "Simple Modular Systems for Generation of Droplets on Demand." *Lab Chip* **2013**, *13*, 3689.

(150) Li, L.; Boedicker, J. Q.; Ismagilov, R. F. "Using a Multijunction Microfluidic Device to Inject Substrate into an Array of Preformed Plugs without Cross-Contamination: Comparing Theory and Experiments." *Anal. Chem.* **2007**, *79*, 2756.

(151) Garcia-Egido, E.; Spikmans, V.; Wong, S. Y. F.; Warrington, B. H. "Synthesis and Analysis of Combinatorial Libraries Performed in an Automated Micro Reactor System." *Lab Chip* **2003**, *3*, 73.

(152) Niu, X.; Gulati, S.; Edel, J. B.; deMello, A. J. "Pillar-Induced Droplet Merging in Microfluidic Circuits." *Lab Chip* **2008**, *8*, 1837.

(153) Niu, X. Z.; Gielen, F.; Edel, J. B.; deMello, A. J. "A Microdroplet Dilutor for High-Throughput Screening." *Nat. Chem.* **2011**, *3*, 437.

(154) Song, H.; Ismagilov, R. F. "Millisecond Kinetics on a Microfluidic Chip Using Nanoliters of Reagents." *J. Am. Chem. Soc.* **2003**, *125*, 14613.

(155) Fang, H.; Xiao, Q.; Wu, F.; Floreancig, P. E.; Weber, S. G. "Rapid Catalyst Screening by a Continuous-Flow Microreactor Interfaced with Ultra-High-Pressure Liquid Chromatography." *J. Org. Chem.* **2010**, *75*, 5619.

(156) Zheng, B.; Tice, J. D.; Ismagilov, R. F. "Formation of Droplets of in Microfluidic Channels Alternating Composition and Applications to Indexing of Concentrations in Droplet-Based Assays." *Anal. Chem.* **2004**, *76*, 4977.

(157) "Accendo Corporation Segment Insights: Product Incubation Acceleration." **2010**.

(158) Hawbaker, N.; Wittgrove, E.; Christensen, B.; Sach, N.; Blackmond, D. G. "Dispersion in Compartmentalized Flow Systems: Influence of Flow Patterns on Reactivity." *Org. Process Res. Dev.* **2015**.

(159) El-Ali, J.; Gaudet, S.; Gunther, A.; Sorger, P. K.; Jensen, K. F. "Cell Stimulus and Lysis in a Microfluidic Device with Segmented Gas-Liquid Flow." *Anal. Chem.* **2005**, *77*, 3629.

(160) "Gilson Guide to Pipetting." **2005**.

(161) Lee, T. G.; Kim, S. Y.; Song, K. H.; Choe, J.; Kim, J. H. "Liquid-Liquid Equilibria for the Ternary Systems of Perfluorohexane or Perfluamine plus Hydrofluoroether plus Tetrahydrofuran at 298.15 K or 273.15 K." *J. Chem. & Eng. Data* **2013**, *58*, 2035.

(162) Eum, K. W.; Gu, H.; Lee, T. G.; Choe, J.; Lee, K.; Song, K. H. "Liquid-Liquid Equilibria for the Ternary Systems of Perfluorohexane plus Methyl Nonafluorobutyl Ether plus Toluene,+1,4-Dioxane, or plus Dimethylformamide at 298.15 K." *J. Chem. & Eng. Data* **2013**, *58*, 915.

(163) Gladysz, J. A.; Curran, D. P.; Horváth, I. T. *Handbook of Fluorous Chemistry*; Weinheim, Wiley-VCH, 2004.

(164) Burns, M. A.; Mastrangelo, C. H.; Sammarco, T. S.; Man, F. P.; Webster, J. R.; Johnsons, B. N.; Foerster, B.; Jones, D.; Fields, Y.; Kaiser, A. R.; Burke, D. T. "Microfabricated Structures for Integrated DNA Analysis." *Proc. Natl. Acad. Sci.* **1996**, *93*, 5556.

(165) Bruno, N. C.; Niljianskul, N.; Buchwald, S. L. "N-Substituted 2-Aminobiphenylpalladium Methanesulfonate Precatalysts and Their Use in C–C and C–N Cross-Couplings." *J. Org. Chem.* **2014**, *79*, 4161.

(166) Abolhasani, M.; Bruno, N. C.; Jensen, K. F. "Oscillatroy Three-Phase Flow Reactor for Studies of Bi-Phasic Catalytic Reactions." *Chem. Commun.* **2015**, Accepted.

(167) Afagh, N. A.; Yudin, A. K. "Chemoselectivity and the Curious Reactivity Preferences of Functional Groups." *Angew. Chem.-Int. Edit.* **2010**, *49*, 262.

(168) Box, G. E. P.; Hunter, J. S. "Multi-Factor Experimental Designs for Exploring Response Surfaces." *Ann. Math. Stat.* **1957**, *28*, 195.

(169) Myers, R. H. *Response Surface Methodology*; Allyn and Bacon: Boston, 1971.

(170) Myers, R. H.; Montgomery, D. C.; Vining, G. G.; Borror, C. M.; Kowalski, S. M. "Response Surface Methodology: A Retrospective and Literature Survey." *J. Qual. Technol.* **2004**, *36*, 53.

(171) Myers, R. H.; Montgomery, D. C. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*; Wiley: New York, 1995.

(172) Box, G. E. P.; Draper, N. R. *Response Surfaces, Mixtures, and Ridge Analyses*; Wiley-Interscience: Hoboken, N.J., 2007.

(173) Atkinson, A. C.; Donev, A. N. *Optimum Experimental Design*; Clarendon Press; Oxford University Press: Oxford, New York, 1992.

(174) Bezerra, M. A.; Santelli, R. E.; Oliveira, E. P.; Villar, L. S.; Escaleira, L. A. "Response Surface Methodology (RSM) as a Tool for Optimization in Analytical Chemistry." *Talanta* **2008**, *76*, 965.

(175) Box, G. E. P.; Wilson, K. B. "On the Experimental Attainment of Optimum Conditions." *J. Roy. Stat. Soc. B* **1951**, *13*, 1.

(176) Box, G. E.; Behnken, D. W. "Some New Three Level Designs for the Study of Quantitative Variables." *Technometrics* **1960**, *2*, 455.

(177) Box, G. E. P.; Draper, N. R. *Evolutionary Operation: A Statistical Method for Process Improvement*; Wiley: New York, 1969.

(178) Box, G. E. "Evolutionary Operation: A Method for Increasing Industrial Productivity." *Appl. Stat.* **1957**, 81.

(179) Box, G. E.; Draper, N. R. *Empirical Model-Building and Response Surfaces*; John Wiley & Sons, 1987.

(180) Khuri, A. I.; Cornell, J. A. *Response Surfaces: Designs and Analyses*; M. Dekker: New York, 1987.

(181) Alaeddini, A.; Murat, A.; Yang, K.; Ankenman, B. "An Efficient Adaptive Sequential Methodology for Expensive Response Surface Optimization." *Qual. Reliab. Eng. Int.* **2013**, *29*, 799.

(182) Wang, G. G.; Dong, Z.; Aitchison, P. "Adaptive Response Surface Method--A Global Optimization Scheme for Approximation-Based Design Problems." *Eng. Optimiz.* **2001**, *33*, 707.

(183) Wang, G. G. "Adaptive Response Surface Method Using Inherited Latin Hypercube Design Points." *J. Mech. Design* **2003**, *125*, 210.

(184) Steinberg, D. M.; Hunter, W. G. "Experimental-Design - Review and Comment." *Technometrics* **1984**, *26*, 71.

(185) Belotti, P.; Kirches, C.; Leyffer, S.; Linderoth, J.; Luedtke, J.; Mahajan, A. "Mixed-Integer Nonlinear Optimization." *Acta Numer.* **2013**, *22*, 1.

(186) Floudas, C. A.; Gounaris, C. E. "A Review of Recent Advances in Global Optimization." *J. Global Optim.* **2009**, *45*, 3.

(187) Holland, J. H. "Genetic Algorithms." *Sci.Am.* **1992**, *267*, 66.

(188) Geem, Z. W.; Kim, J. H.; Loganathan, G. V. "A New Heuristic Optimization Algorithm: Harmony Search." *Simulation* **2001**, *76*, 60.

(189) Hemker, T.; Fowler, K. R.; Farthing, M. W.; von Stryk, O. "A Mixed-Integer Simulation-Based Optimization Approach with Surrogate Functions in Water Resources Management." *Optim Eng* **2008**, *9*, 341.

(190) Muller, J.; Shoemaker, C. A.; Piche, R. "SO-MI: A Surrogate Model Algorithm for Computationally Expensive Nonlinear Mixed-Integer Black-Box Global Optimization Problems." *Comput. Oper. Res.* **2013**, *40*, 1383.

(191) Jones, D. R.; Schonlau, M.; Welch, W. J. "Efficient Global Optimization of Expensive Black-Box Functions." *J. Global Optim.* **1998**, *13*, 455.

(192) Sacks, J.; Schiller, S. B.; Welch, W. J. "Designs for Computer Experiments." *Technometrics* **1989**, *31*, 41.

(193) Jones, D. R. "A Taxonomy of Global Optimization Methods Based on Response Surfaces." *J. Global Optim.* **2001**, *21*, 345.

(194) Forrester, A. I. J.; Keane, A. J. "Recent Advances in Surrogate-Based Optimization." *Prof. Aerosp. Sci.* **2009**, *45*, 50.

(195) Queipo, N. V.; Haftka, R. T.; Shyy, W.; Goel, T.; Vaidyanathan, R.; Tucker, P. K. "Surrogate-Based Analysis and Optimization." *Prof. Aerosp. Sci.* **2005**, *41*, 1.

(196) Holmström, K.; Quttineh, N.-H.; Edvall, M. "An Adaptive Radial Basis Algorithm (ARBF) for Expensive Black-Box Mixed-Integer Constrained Global Optimization." *Optim Eng* **2008**, *9*, 311.

(197) Rashid, K.; Ambani, S.; Cetinkaya, E. "An Adaptive Multiquadric Radial Basis Function Method for Expensive Black-Box Mixed-Integer Nonlinear Constrained Optimization." *Eng. Optimiz.* **2013**, *45*, 185.

(198) Davis, E.; Ierapetritou, M. "A Kriging Based Method for the Solution of Mixed-Integer Nonlinear Programs Containing Black-Box Functions." *J. Global Optim.* **2009**, *43*, 191.

(199) Dumesic, J. A.; Rudd, D. F.; Aparicio, L. M.; Rekoske, J. E.; Trevino, A. A. *The Microkinetics of Heterogeneous Catalysis*; American Chemical Society: Washington, D.C., 1993.

(200) Quenouille, M. H. "Approximate Tests of Correlation in Time-Series." *J. Roy. Stat. Soc. B* **1949**, *11*, 68.

(201) Jones, B.; Goos, P. "I-Optimal Versus D-Optimal Split-Plot Response Surface Designs." *J. Qual. Technol.* **2012**, *44*, 85.

(202) Pilling, M. J.; Seakins, P. W. *Reaction Kinetics*; Oxford University Press: New York, 1995.

(203) Hughes, E. D.; Ingold, C. K. "Mechanism of Substitution at a Saturated Carbon Atom. Part IV. A Discussion of Constitutional and Solvent Effects on the Mechanism, Kinetics, Velocity, and Orientation of Substitution." *J. Chem. Soc.* **1935**, 244.

(204) Reichardt, C. "Empirical Parameters of Polarity of Solvents." *Angew. Chem.-Int. Edit.* **1965**, *4*, 29.

(205) Miller, J.; Parker, A. J. "Dipolar Aprotic Solvents in Bimolecular Aromatic Nucleophilic Substitution Reactions." *J. Am. Chem. Soc.* **1961**, *83*, 117.

(206) Connors, K. A. *Chemical Kinetics: The Study of Reaction Rates in Solution*; VCH: New York, 1990.

(207) Parker, A. J. "Protic-Dipolar Aprotic Solvent Effects on Rates of Bimolecular Reactions." *Chem. Rev.* **1969**, *69*, 1.

(208) Parker, A. J.; Mayer, U.; Schmid, R.; Gutmann, V. "Correlation of Solvent Effects on Rates of Solvolysis and SN2 Reactions." *J. Org. Chem.* **1978**, *43*, 1843.

(209) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*; 4th ed.; Wiley-VCH: Weinheim, 2010.

(210) Theberge, A. B.; Whyte, G.; Frenzel, M.; Fidalgo, L. M.; Wootton, R. C. R.; Huck, W. T. S. "Suzuki-Miyaura Coupling Reactions in Aqueous Microdroplets with Catalytically Active Fluorous Interfaces." *Chem. Commun.* **2009**, 6225.

(211) Lee, J. N.; Park, C.; Whitesides, G. M. "Solvent Compatibility of Poly(dimethylsiloxane)-Based Microfluidic Devices." *Anal. Chem.* **2003**, *75*, 6544.

(212) Salvatore, R. N.; Nagle, A. S.; Jung, K. W. "Cesium Effect: High Chemoselectivity in Direct N-Alkylation of Amines." *J. Org. Chem.* **2002**, *67*, 674.

(213) Fu, X. K.; Gong, C. B.; Ma, X. B.; Wen, S. Y. "A New Simple Route to N-Substituted 2-Aminoethylphosphonic Acids." *Synthetic Commun.* **1998**, *28*, 2659.

(214) Mitchell, J. M.; Finney, N. S. "An Efficient Method for the Preparation of N,N-Disubstituted 1,2-Diamines." *Tetrahedron Lett.* **2000**, *41*, 8431.

(215) Kaik, M.; Gawronski, J. "Facile Monoprotection of trans-1,2-Diaminocyclohexane." *Tetrahedron-Asymmetr.* **2003**, *14*, 1559.

(216) Suez, G.; Bloch, V.; Nisnevich, G.; Gandelman, M. "Design and Development of Bioinspired Guanine-Based Organic Catalyst for Asymmetric Catalysis." *Eur. J. Org. Chem.* **2012**, 2118.

(217) Xu, B.; Li, L.; Gou, S. "A chiral Primary-Tertiary-1,2-Diamine as an Efficient Catalyst in Asymmetric Aldehyde–Ketone or Ketone–Ketone Aldol Reactions." *Tetrahedron-Asymmetr.* **2013**, *24*, 1556.

(218) Beers, K. J. *Numerical Methods for Chemical Engineering: Applications in Matlab*; Cambridge University Press: New York, 2007.

(219) Lebleu, T.; Ma, X. L.; Maddaluno, J.; Legros, J. "Selective Monomethylation of Primary Amines with Simple Electrophiles." *Chem. Commun.* **2014**, *50*, 1836.

(220) Doyle, A. G.; Jacobsen, E. N. "Small-Molecule H-Bond Donors in Asymmetric Catalysis." *Chem. Rev.* **2007**, *107*, 5713.

(221) *CRC Handbook of Chemistry and Physics*; 95th Edition ed.; CRC Press: Cleveland, OH, 2015.

(222) Rivas, M. A.; Iglesias, T. P. "On Permittivity and Density of the Systems {Triglyme plus (Dimethyl or Diethyl Carbonate)} and Formulation of Delta Epsilon in Terms of Volume or Mole Fraction." *J. Chem. Thermodynamics* **2008**, *40*, 1120.

(223) Marcheselli, L.; Pistoni, G.; Tagliazucchi, M.; Tassi, L.; Tosi, G. "N,N-Dimethylformamide + 1,2-Dimethoxyethane Binary Mixtures. The Static Dielectric Constant from 40 to 80.degree.C." *J. Chem. & Eng. Data* **1993**, *38*, 204.

(224) Laurence, C.; Gal, J.-F. *Lewis Basicity and Affinity Scales: Data and Measurement*; Hoboken, N.J., Wiley, 2009.

(225) Laurence, C.; Berthelot, M. "Observations on the Strength of Hydrogen Bonding." *Perspect. Drug Discov.* **2000**, *18*, 39.

(226) Lequestel, J. Y.; Laurence, C.; Lachkar, A.; Helbert, M.; Berthelot, M. "Hydrogen-Bond Basicity of Secondary and Tertiary Amides, Carbamates, Ureas and Lactams." *J. Chem. Soc.-Perkin Transactions 2* **1992**, 2091.

(227) Berthelot, M.; Laurence, C.; Safar, M.; Besseau, F. "Hydrogen-Bond Basicity pK(HB) Scale of Six-Membered Aromatic N-Heterocycles." *J. Chem. Soc.-Perkin Transactions 2* **1998**, 283.

(228) Ouvrard, C.; Berthelot, M.; Laurence, C. "The First Basicity Scale of Fluoro-, Chloro-, Bromo- and Iodo-Alkanes: Some Cross-Comparisons with Simple Alkyl Derivatives of Other Elements." *J. Chem. Soc.-Perkin Transactions 2* **1999**, 1357.

(229) Berthelot, M.; Besseau, F.; Laurence, C. "The Hydrogen-Bond Basicity pK(HB) Scale of Peroxides and Ethers." *Eur. J. Org. Chem.* **1998**, 925.

(230) Besseau, F.; Laurence, C.; Berthelot, M. "Hydrogen-Bond Basicity of Esters, Lactones and Carbonates." *J. Chem. Soc.-Perkin Transactions 2* **1994**, 485.

(231) Laurence, C.; Berthelot, M.; Helbert, M.; Sraidi, K. "1st Measurement of the Hydrogen-Bond Basicity of Monomeric Water, Phenols, and Weakly Basic Alcohols." *J. Phys. Chem.* **1989**, *93*, 3799.

(232) Besseau, F.; Laurence, C.; Berthelot, M. "The pK(HB) Scale of Pi Bases." *B. Soc. Chim. Fr.* **1996**, *133*, 381.

(233) van der Spoel, D.; van Maaren, P. J.; Larsson, P.; Tîmneanu, N. "Thermodynamics of Hydrogen Bonding in Hydrophilic and Hydrophobic Media." *J. Phys. Chem. B* **2006**, *110*, 4393.

(234) Miyaura, N.; Suzuki, A. "Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds." *Chem. Rev.* **1995**, *95*, 2457.

(235) Bellina, F.; Carpita, A.; Rossi, R. "Palladium Catalysts for the Suzuki Cross-Coupling Reaction: An Overview of Recent Advances." *Synthesis-Stuttgart* **2004**, 2419.

(236) Miyaura, N. In *Metal-Catalyzed Cross-Coupling Reactions*; Wiley-VCH Verlag GmbH: 2008, p 41.

(237) Hartwig, J. F. "Carbon-Heteroatom Bond Formation Catalysed by Organometallic Complexes." *Nature* **2008**, *455*, 314.

(238) Seechurn, C.; Kitching, M. O.; Colacot, T. J.; Snieckus, V. "Palladium-Catalyzed Cross-Coupling: A Historical Contextual Perspective to the 2010 Nobel Prize." *Angew. Chem.-Int. Edit.* **2012**, *51*, 5062.

(239) Noel, T.; Buchwald, S. L. "Cross-Coupling in Flow." *Chem. Soc. Rev.* **2011**, *40*, 5010.

(240) Marion, N.; Nolan, S. P. "Well-Defined N-Heterocyclic Carbenes-Palladium(II) Precatalysts for Cross-Coupling Reactions." *Accounts Chem. Res.* **2008**, *41*, 1440.

(241) Kataoka, N.; Shelby, Q.; Stambuli, J. P.; Hartwig, J. F. "Air Stable, Sterically Hindered Ferrocenyl Dialkylphosphines for Palladium-Catalyzed C-C, C-N, and C-O Bond-Forming Cross-Couplings." *J. Org. Chem.* **2002**, *67*, 5553.

(242) Martin, R.; Buchwald, S. L. "Palladium-Catalyzed Suzuki-Miyaura Cross-Coupling Reactions Employing Dialkylbiaryl Phosphine Ligands." *Accounts Chem. Res.* **2008**, *41*, 1461.

(243) Barder, T. E.; Walker, S. D.; Martinelli, J. R.; Buchwald, S. L. "Catalysts for Suzuki-Miyaura Coupling Processes: Scope and Studies of the Effect of Ligand Structure." *J. Am. Chem. Soc.* **2005**, *127*, 4685.

(244) Biscoe, M. R.; Fors, B. P.; Buchwald, S. L. "A New Class of Easily Activated Palladium Precatalysts for Facile C-N Cross-Coupling Reactions and the Low Temperature Oxidative Addition of Aryl Chlorides." *J. Am. Chem. Soc.* **2008**, *130*, 6686.

(245) Bruno, N. C.; Tudge, M. T.; Buchwald, S. L. "Design and Preparation of New Palladium Precatalysts for C-C and C-N Cross-Coupling Reactions." *Chem. Sci.* **2013**, *4*, 916.

(246) Fu, G. C. "The Development of Versatile Methods for Palladium-Catalyzed Coupling Reactions of Aryl Electrophiles through the Use of P(t-Bu)(3) and PCy3 as Ligands." *Accounts Chem. Res.* **2008**, *41*, 1555.

(247) Murray, P. M.; Tyler, S. N. G.; Moseley, J. D. "Beyond the Numbers: Charting Chemical Reaction Space." *Org. Process Res. Dev.* **2013**, *17*, 40.

(248) Carlson, R.; Carlson, J. E. "Principal Properties and Designs for Discrete Variations." *Org. Process Res. Dev.* **2005**, *9*, 680.

(249) Moseley, J. D.; Murray, P. M. "Ligand and Solvent Selection in Challenging Catalytic Reactions." *J. Chem. Technol. Biot.* **2014**, *89*, 623.

(250) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. "Expansion of the Ligand Knowledge Base for Chelating P,P-Donor Ligands (LKB-PP)." *Organometallics* **2012**, *31*, 5302.

(251) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. "Expansion of the Ligand Knowledge Base for Monodentate P-Donor Ligands (LKB-P)." *Organometallics* **2010**, *29*, 6245.

(252) Fey, N.; Haddow, M. F.; Harvey, J. N.; McMullin, C. L.; Orpen, A. G. "A Ligand Knowledge Base for Carbenes (LKB-C): Maps of Ligand Space." *Dalton T.* **2009**, 8183.

(253) Tolman, C. A. "Steric Effects of Phosphorus Ligands in Organometallic Chemistry and Homogeneous Catalysis." *Chem. Rev.* **1977**, *77*, 313.

(254) Bunten, K. A.; Chen, L. Z.; Fernandez, A. L.; Poe, A. J. "Cone Angles: Tolman's and Plato's." *Coordin. Chem. Rev.* **2002**, *233*, 41.

(255) Freixa, Z.; van Leeuwen, P. "Bite Angle Effects in Diphosphine Metal Catalysts: Steric or Electronic?" *Dalton T.* **2003**, 1890.

(256) Kinzel, T.; Zhang, Y.; Buchwald, S. L. "A New Palladium Precatalyst Allows for the Fast Suzuki-Miyaura Coupling Reactions of Unstable Polyfluorophenyl and 2-Heteroaryl Boronic Acids." *J. Am. Chem. Soc.* **2010**, *132*, 14073.

(257) Bruno, N. C.; Buchwald, S. L. "Synthesis and Application of Palladium Precatalysts that Accommodate Extremely Bulky Di-tert-butylphosphino Biaryl Ligands." *Org. Lett.* **2013**, *15*, 2876.

(258) Christmann, U.; Vilar, R. "Monoligated Palladium Species as Catalysts in Cross-Coupling Reactions." *Angew. Chem.-Int. Edit.* **2005**, *44*, 366.

(259) Barrios-Landeros, F.; Hartwig, J. F. "Distinct Mechanisms for the Oxidative Addition of Chloro-, Bromo-, and Iodoarenes to a Bisphosphine Palladium(0) Complex with Hindered Ligands." *J. Am. Chem. Soc.* **2005**, *127*, 6944.

(260) Garcia-Melchor, M.; Braga, A. A. C.; Lledos, A.; Ujaque, G.; Maseras, F. "Computational Perspective on Pd-Catalyzed C-C Cross-Coupling Reaction Mechanisms." *Accounts Chem. Res.* **2013**, *46*, 2626.

(261) Braga, A. A. C.; Ujaque, G.; Maseras, F. "A DFT Study of the Full Catalytic Cycle of the Suzuki-Miyaura Cross-Coupling on a Model System." *Organometallics* **2006**, *25*, 3647.

(262) Dufert, M. A.; Billingsley, K. L.; Buchwald, S. L. "Suzuki-Miyaura Cross-Coupling of Unprotected, Nitrogen-Rich Heterocycles: Substrate Scope and Mechanistic Investigation." *J. Am. Chem. Soc.* **2013**, *135*, 12877.

(263) Braga, A. A. C.; Morgon, N. H.; Ujaque, G.; Maseras, F. "Computational Characterization of the Role of the Base in the Suzuki-Miyaura Cross-Coupling Reaction." *J. Am. Chem. Soc.* **2005**, *127*, 9298.

(264) Amatore, C.; Jutand, A.; Le Duc, G. "Kinetic Data for the Transmetalation/Reductive Elimination in Palladium-Catalyzed Suzuki-Miyaura Reactions: Unexpected Triple Role of Hydroxide Ions Used as Base." *Chem.-Eur. J.* **2011**, *17*, 2492.

(265) Lennox, A. J. J.; Lloyd-Jones, G. C. "Transmetalation in the Suzuki-Miyaura Coupling: The Fork in the Trail." *Angew. Chem.-Int. Edit.* **2013**, *52*, 7362.

(266) Barrios-Landeros, F.; Carrow, B. P.; Hartwig, J. F. "Effect of Ligand Steric Properties and Halide Identity on the Mechanism for Oxidative Addition of Haloarenes to Trialkylphosphine Pd(0) Complexes." *J. Am. Chem. Soc.* **2009**, *131*, 8141.

(267) Schoenebeck, F.; Houk, K. N. "Ligand-Controlled Regioselectivity in Palladium-Catalyzed Cross Coupling Reactions." *J. Am. Chem. Soc.* **2010**, *132*, 2496.

(268) Elvira, K. S.; Solvas, X. C. I.; Wootton, R. C. R.; deMello, A. J. "The Past, Present and Potential for Microfluidic Reactor Technology in Chemical Synthesis." *Nat. Chem.* **2013**, *5*, 905.

(269) Peplow, M. "The Robo-Chemist." *Nature* **2014**, *512*, 20.

(270) Li, J. Q.; Ballmer, S. G.; Gillis, E. P.; Fujii, S.; Schmidt, M. J.; Palazzolo, A. M. E.; Lehmann, J. W.; Morehouse, G. F.; Burke, M. D. "Synthesis of Many Different Types of Organic Small Molecules Using One Automated Process." *Science* **2015**, *347*, 1221.

(271) Seeberger, P. H. "Automated Oligosaccharide Synthesis." *Chem. Soc. Rev.* **2008**, *37*, 19.

(272) Coles, S. J.; Frey, J. G.; Bird, C. L.; Whitby, R. J.; Day, A. E. "First Steps towards Semantic Descriptions of Electronic Laboratory Notebook Records." *J. Cheminform.* **2013**, *5*.

(273) Heider, P. L.; Born, S. C.; Basak, S.; Benyahia, B.; Lakerveld, R.; Zhang, H. T.; Hogan, R.; Buchbinder, L.; Wolfe, A.; Mascia, S.; Evans, J. M. B.; Jamison, T. F.; Jensen, K. F. "Development of a Multi-Step Synthesis and Workup Sequence for an Integrated, Continuous Manufacturing Process of a Pharmaceutical." *Org. Process Res. Dev.* **2014**, *18*, 402.

(274) O'Neal, E. J.; Jensen, K. F. "Continuous Nanofiltration and Recycle of a Metathesis Catalyst in a Microflow System." *ChemCatChem* **2014**, *6*, 3004.

(275) Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X. Y.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y. K.; Liao, X. L.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. "Rapid Flow- Based Peptide Synthesis." *ChemBioChem* **2014**, *15*, 713.

(276) Ley, S. V.; Fitzpatrick, D. E.; Ingham, R. J.; Myers, R. M. "Organic Synthesis: March of the Machines." *Angew. Chem.-Int. Edit.* **2015**, *54*, 3449.

(277) Griffin, J. D.; Fowler, K. R.; Gray, G. A.; Hemker, T.; Parno, M. D. "Derivative-Free Optimization via Evolutionary Algorithms Guiding Local Search (EAGLS) for MINLP." *Pac. J. Optim.* **2011**, *7*, 425.

(278) Pickett, S. D.; Green, D. V. S.; Hunt, D. L.; Pardoe, D. A.; Hughes, I. "Automated Lead Optimization of MMP-12 Inhibitors Using a Genetic Algorithm." *ACS Med. Chem. Lett.* **2011**, *2*, 28.

(279) Singh, J.; Ator, M. A.; Jaeger, E. P.; Allen, M. P.; Whipple, D. A.; Soloweij, J. E.; Chowdhary, S.; Treasurywala, A. M. "Application of Genetic Algorithms to Combinatorial Synthesis: A Computational Approach to Lead Identification and Lead Optimization." *J. Am. Chem. Soc.* **1996**, *118*, 1669.

(280) Wolf, D.; Buyevskaya, O. V.; Baerns, M. "An Evolutionary Approach in the Combinatorial Selection and Optimization of Catalytic Materials." *Appl. Catal. A-Gen.* **2000**, *200*, 63.

(281) Werner, M.; Kuratli, C.; Martin, R. E.; Hochstrasser, R.; Wechsler, D.; Enderle, T.; Alanine, A. I.; Vogel, H. "Seamless Integration of Dose-Response Screening and Flow Chemistry: Efficient Generation of Structure-Activity Relationship Data of beta-Secretase (BACE1) Inhibitors." *Angew. Chem.-Int. Edit.* **2014**, *53*, 1704.

APPENDIX A. CHAPTER 2 SUPPORTING INFORMATION

A.1. EXPERIMENTAL DATA

Table A1. List of experimental conditions and measured outlet concentrations for initial simultaneous parameter estimation.

Expt	t_{res} (min)	T (°C)	C_{10} (M)	Equiv. **2**	C_1 (M)	C_3 (M)	C_4 (M)	C_5 (M)
1	1.0	80	0.150	1.0	0.0437	0.0922	0.0168	0.0001
2	10	80	0.150	2.0	0.0056	0.1247	0.0205	0.0022
3	1.0	40	0.150	1.0	0.0783	0.0639	0.0096	0.0001
4	0.5	40	0.150	1.0	0.1040	0.0407	0.0060	0.0000
5	1.0	40	0.150	2.0	0.0480	0.0907	0.0134	0.0000
6	10	40	0.150	1.0	0.0216	0.1139	0.0160	0.0002
7	0.5	40	0.150	2.0	0.0696	0.0706	0.0105	0.0000
8	0.5	80	0.150	2.0	0.0192	0.1101	0.0203	0.0001
9	0.5	80	0.150	1.0	0.0598	0.0766	0.0141	0.0000
10	1.0	80	0.150	2.0	0.0090	0.1189	0.0219	0.0002
11	10	40	0.150	2.0	0.0048	0.1265	0.0181	0.0003
12	10	80	0.150	1.0	0.0129	0.1178	0.0201	0.0006
13	20	100	0.150	2.5	0.0042	0.1113	0.0191	0.0185
14	20	80	0.150	2.5	0.0049	0.1195	0.0198	0.0067
15	20	100	0.150	2.5	0.0041	0.1133	0.0195	0.0177
16	20	80	0.150	2.5	0.0050	0.1208	0.0201	0.0068
17	20	100	0.150	2.5	0.0051	0.1106	0.0192	0.0181
18	0.5	100	0.150	1.5	0.0194	0.1078	0.0220	0.0003
19	20	60	0.150	2.5	0.0059	0.1239	0.0196	0.0026
20	20	40	0.150	2.5	0.0000	0.1256	0.0183	0.0005
21	20	80	0.150	2.5	0.0047	0.1184	0.0198	0.0066
22	20	100	0.150	2.5	0.0048	0.1093	0.0198	0.0183
23	0.5	100	0.150	1.5	0.0192	0.1077	0.0222	0.0003
24	20	100	0.150	2.5	0.0047	0.1086	0.0195	0.0187

Table A2. List of experimental conditions and measured outlet concentrations for isolated estimation of A_1, E_{A1}, A_2, and E_{A2}.

Expt	t_{res} (min)	T (°C)	C_{10} (M)	Equiv. **2**	C_1 (M)	C_3 (M)	C_4 (M)	C_5 (M)
1	0.5	100	0.150	1.0	0.0425	0.0895	0.0193	0.0000
2	0.5	100	0.150	2.0	0.0069	0.1166	0.0247	0.0003
3	1.0	100	0.150	2.0	0.0056	0.1172	0.0242	0.0013
4	0.5	40	0.150	1.0	0.1012	0.0435	0.0065	0.0000
5	1.0	100	0.150	1.0	0.0283	0.1012	0.0208	0.0004
6	1.0	40	0.150	1.0	0.0784	0.0651	0.0098	0.0000
7	0.5	40	0.150	2.0	0.0675	0.0738	0.0108	0.0000
8	1.0	40	0.150	2.0	0.0405	0.0970	0.0144	0.0000
9	0.5	100	0.150	1.5	0.0182	0.1100	0.0227	0.0002
10	1.0	40	0.150	2.5	0.0305	0.1054	0.0159	0.0000
11	1.0	40	0.150	2.5	0.0281	0.1071	0.0161	0.0002
12	0.5	100	0.150	1.5	0.0186	0.1085	0.0226	0.0002

Table A3. List of experimental conditions and measured outlet concentrations for isolated estimation of A_3 and E_{A3}.

Expt	t_{res} (min)	T (°C)	C_{10} (M)	Equiv. **2**	C_1 (M)	C_3 (M)	C_4 (M)	C_5 (M)
1	20	80	0.050	2.5		0.0489		0.0009
2	10	100	0.050	2.5		0.0482		0.0011
3	10	80	0.050	2.5		0.0489		0.0005
4	20	100	0.050	2.5		0.0472		0.0025
5	20	100	0.050	2.5		0.0476		0.0025
6	20	80	0.050	2.5		0.0489		0.0009
7	20	100	0.050	2.5		0.0473		0.0025
8	20	80	0.050	2.5		0.0486		0.0009
9	20	100	0.050	2.5		0.0470		0.0024
10	20	80	0.050	2.5		0.0488		0.0009
11	20	100	0.050	2.5		0.0477		0.0024
12	20	80	0.050	2.5		0.0490		0.0009

Table A4. List of experimental conditions and measured outlet concentrations for isolated estimation of A_4 and E_{A4}.

Expt	t_{res} (min)	T (°C)	C_{10} (M)	Equiv. **2**	C_1 (M)	C_3 (M)	C_4 (M)	C_5 (M)
1	20	100	0.030	2.5			0.0287	0.0014
2	10	100	0.030	2.5			0.0298	0.0007
3	20	80	0.030	2.5			0.0299	0.0006
4	10	80	0.030	2.5			0.0299	0.0004
5	20	100	0.030	2.5			0.0287	0.0013
6	20	80	0.030	2.5			0.0299	0.0006
7	20	100	0.030	2.5			0.0294	0.0014
8	20	80	0.030	2.5			0.0298	0.0007
9	20	100	0.030	2.5			0.0293	0.0013
10	20	80	0.030	2.5			0.0297	0.0006
11	20	100	0.030	2.5			0.0291	0.0014
12	20	80	0.030	2.5			0.0302	0.0006

Table A5. List of experimental conditions and measured outlet concentrations for final simultaneous parameter estimation.

Expt	t_{res} (min)	T (°C)	C_{10} (M)	Equiv. **2**	C_1 (M)	C_3 (M)	C_4 (M)	C_5 (M)
1	0.5	40	0.150	1.0	0.1062	0.0397	0.0057	0.0003
2	10	40	0.150	1.0	0.0248	0.1104	0.0162	0.0003
3	20	40	0.150	2.5	0.0043	0.1261	0.0179	0.0009
4	20	100	0.150	2.5	0.0043	0.1119	0.0179	0.0162
5	0.5	100	0.150	1.0	0.0410	0.0887	0.0177	0.0009
6	20	40	0.150	1.0	0.0145	0.1182	0.0170	0.0004
7	10	40	0.150	2.5	0.0057	0.1261	0.0181	0.0006
8	10	100	0.150	1.0	0.0072	0.1199	0.0220	0.0020
9	10	100	0.150	2.5	0.0050	0.1171	0.0195	0.0084
10	0.5	40	0.150	2.5	0.0539	0.0842	0.0126	0.0006
11	0.5	100	0.150	2.5	0.0049	0.1186	0.0237	0.0009
12	20	100	0.150	1.0	0.0052	0.1209	0.0209	0.0031
13	0.5	100	0.150	1.0	0.0471	0.0874	0.0175	0.0002
14	0.5	100	0.150	1.0	0.0475	0.0854	0.0176	0.0003
15	0.5	100	0.150	1.0	0.0455	0.0870	0.0177	0.0002
16	20	100	0.150	2.5	0.0046	0.1115	0.0176	0.0165
17	20	80	0.150	2.5	0.0054	0.1223	0.0191	0.0062
18	20	80	0.150	2.5	0.0044	0.1199	0.0187	0.0061

APPENDIX B. CHAPTER 3 SUPPORTING INFORMATION

B.1. AUTOMATED SCREENING SYSTEM STANDARD OPERATING PROCEDURE

B.1.1. HPLC Initialization

Power on all modules of the Agilent LC/MS. Ensure that the water and acetonitrile liquid volumes are adequate and that ~96 psi nitrogen is flowing to the MS. Open ChemStation. If the most recent method in the ChemStation queue was one which received samples from the Agilent automated liquid handler (rather than manual injections), follow steps 1-6 below:

1. Load the method BJR_SLUGFLOW_3_MANUAL.M.
2. Configure ChemStation to run a single sample injection
3. Under Instrument → Select Injection Source, change the injection source from manual to ALS.
4. Start a single run. It is okay to stop the run any time after time = 0.
5. Under Instrument → Select Injection Source, change the injection source back from ALS to manual. Save the method.
6. Start a single run. It is okay to stop the run any time after time = 0.

Load the method BJR_SLUGFLOW_3_MANUAL.M and sequence BJR_SLUGFLOW.S. Under Sequence → Sequence Parameters, verify that the subdirectory is consistent with the path for HPLC chromatograms in the LabView Master VI.vi. Assign the sequence a prefix that is alphanumerically after all previous stored files in the subdirectory (for instance with the time stamp YYMMDD). Exit this window and configure ChemStation to run a sequence injection. Change the instrument, including the MS, from standby mode to on by pressing the green power button beneath the MS icon (Note: pressing only the power button under the MWD icon will leave the MS in standby mode and not initiate the method). Press the button on the ChemStation interface to start the sequence. You will be automatically prompted to approve that the files are being written to a previously existing subdirectory. Choose yes. You will then be prompted to specify a blank run. Select by file and choose BLANK_SLUGFLOW_3_MANUAL.D. (If this blank becomes outdated or a new method is used, simply in single sample mode execute a new blank run and point the blank subtraction macro to the data file generated by that blank). The LC/MS should now be ready for use.

B.1.2. System Initialization

Ensure that power is being supplied to all units. The LabView Real-time Controller should be receiving 15-20 V power, relays should be receiving 4 V power, and solenoid valves should be receiving 12 V power. Ensure that the transfer line refill tank and quench refill tank are adequately filled. Replace and fill the rinse solutions in positions 1-4 of the liquid handler. If necessary, empty the contents of the trap at the outlet of the injection valve. Prepare a fresh solution of the online injection fluid. Confirming that the system is depressurized or that the online injection valve is switched to the refill position (by convention, position 2 on all valves is denoted as 'into the system', so 'refill' is position 1), remove the glass syringe from the online injection line, and purge and refill the syringe with the online injection fluid. Replace the syringe in the line and verify that the online injection valve is back to remote control, as should be all other valves. Take the remaining stock solution and submerge the refill tubing from the online injection valve into the online injection fluid. Insert the inert gas purge needle into the stock solution above the liquid level. Pressurize the system by opening the tank of inert gas supplied to the system. The pressure gauge on the bomb at the outlet of the system should rapidly equilibrate to 100 psi; if not, verify that the gas pressure out of the inert gas regulator is sufficient at >100 psi and that there are no leaks in the system. Pressurize the manifold by opening the tank of inert gas supplied to the manifold (the pressure should be maintained at 3 psi). Open the vacuum line supplied to the solenoid valve connected to the injection valve trap (note that the solenoid valve should be configured that the vacuum is only applied when the solenoid valve switches to the open position).

Prepare screening solutions as desired and store in the liquid handler, preferably beneath the inert gas manifold. These solutions must be prepared concentrated enough such that it is feasible to dilute all solutions to make slugs of the desired compositions. On the main computer, open the file Reagent Table.xlsx and input relevant data about the starting reagents and expected products on the Reagent List sheet. For the starting reagents, molar masses, concentrations (in g/L), internal standard concentrations (in g/L), densities, and sample locations must be supplied. HPLC retention times, wavelengths, and calibrations may optionally be supplied. For products, HPLC retention times, wavelengths, and calibrations must be supplied. The desired product for optimization must be denoted as type R in the reagent type column (there should only be one R). Other important types are I for internal standard (there should only be one I) and X for the online

injection solution (there should only be one X). Other classes of reagents and the calibration column have been incorporated for future development. On the Initial Slugs sheet, specify any specific experiments to run by reagent numbers, concentrations, reaction time, and temperature. This is only required if not running an optimization. Save any changes to the file Reagent Table.xlsx. Ensure that the reagent table path on Master VI.vi points to this file. For optimizations, open slug_optimization_phase_1.m and program the continuous and discrete variable inputs to read into the optimization. The entries for each row of the continuous variable matrix cv are [(min) (max) (default value) (± tolerance for prediction covariance) (optimize? 1 for yes, 0 for no)]. Assumable values for discrete variables are listed as a row in the matrix dv. The rows of the matrix cor are discrete variables that correlate with given optimized discrete variables. The columns of cor align 1:1 with the columns of dv (ie dv(1) will always be paired with the variables in the first column of cor), with the last column of cor specifying the row of dv to correlate to. I_cv and I_dv are the interaction terms to be considered between continuous and continuous variables and discrete and continuous variables, respectively. The rows/columns of I_cv correspond 1:1 to the optimized variables cv. The rows of I_dv correspond 1:1 to the optimized discrete variables and the columns of I_dv correspond to the optimized continuous variables. Gamma is as defined in Chapter 4, the minimum fraction of the maximum yield for a feasible TON. Save any changes to this file.

Prime the system by opening Automatic Refill-Carrier and Base (Flush Base Line).vi and inputting the syringe volumes, the volume of solutions to push in to waste, and the current fill volumes of the syringes. Run this vi. Once complete, open the file Master VI.vi. Update all paths for output data so that these do not overwrite old data. Mark the Reset Syringe Volumes box as yes and input new carrier and base syringe fill volumes. Check the box for Optimize as yes if optimization is desired and as no otherwise. All other values should nominally stay as default, but may be manipulated for a new reactor volume, longer HPLC method time, etc. Run the VI to begin. After the set of experiments is complete, the temperature set point should automatically return to zero. The pumps and Master VI.vi will automatically shut off after an optimization, but both will remain on in the case of on-demand screening in case more experiments are desired. The vacuum should be turned off manually, and the LC/MS should be shut down by stopping the current sequence and pressing the power off button beneath the MS icon. The valve on the gas supply to the system should be closed.

B.1.3. List of LabView Files

Master VI

- This is the main vi for optimization and screening. The vi calls all MATLAB functions and executes multiple parallel loops which update slug status with time, update flow rate and manage syringe refills, prepare slugs, inject reagent online, interpret online slug sensor measurements, sample slugs by HPLC, interpret the HPLC data files, control temperature, record data, and optimize.
- Inputs:
    - Reset Syringe Fill Volume (yes for new fill volumes, no for read in old fill volumes from Path for Syringe Fill Volumes)
    - Optimize? (yes for optimization, no for only automated screening)
    - Distance to Online Base (system volume from injection valve to online injection point; default 69.0 μL)
    - Distance to Base Phase Sensor (system volume from injection valve to online injection sensor; default 50.0 μL)
    - Init Carrier Fill Volume (if Reset Syringe Fill Volume is yes, new fill volume for carrier syringe)
    - Init Base Fill Volume (if Reset Syringe Fill Volume is yes, new fill volume for base syringe)
    - Distance to Pre-Reactor Phase Sensor (system volume from injection valve to pre-reactor sensor; default 85.0 μL)
    - Gilson Control VISA (COM port for Gilson GX-271; default is COM1)
    - Entrance Volume (system volume from injection valve to reactor inlet; default 100.0 μL)
    - HPLC Warmup Time (time from HPLC triggering to sample injection into HPLC; default 20.0 s)
    - Analysis Time (time from sample injection into HPLC to end of HPLC method; default 8.75 min)
    - Withdraw Flow Rate (flow rate for sample aspiration in liquid handler; default 100.0 μL/min)

- Injection Flow Rate (flow rate for injection of sample into liquid handler injection valve; default 250.0 µL/min)
- Reactor Volume (volume of reactor; default 240.0 µL)
- Transfer Fluid Density (density of fluid in liquid handler transfer line; default 889.0 g/L)
- Rinse Volume (volume of rinse solvent to sample for each blank slug; default 60.0 µL)
- Quench Volume (system volume from reactor outlet to sampling loop; default 88.0 µL)
- LC Valve to Valve Volume (volume from sampling loop to HPLC sample loop; default 60.0 µL)
- Injection Volume (volume of liquid handler injection valve sample loop; default 14.0 µL)
- Buffer Volume (volume before reactor at which point the flow rate must be set to the reactor flow rate; default 70.0 µL)
- Slug Volume Tolerance (maximum error in the slug volume registered by the pre-reactor phase sensor for the slug to be considered acceptable; default 4.0 µL)
- Wait Time for Report (time for HPLC report generation after method completion; default 20.0 s)
- Distance to HPLC sensor (system volume from injection valve to pre-HPLC phase sensor; default 390.0 µL)
- Slug Volume Threshold (minimum volume considered acceptable for a slug to be recognized by a phase sensor; default 2.0 µL)
- Time Between HPLC Injections (time after HPLC method is complete to wait before a new HPLC method can initiate; default 60.0 s)
- Slug Matching Tolerance ((±) maximum system volume considered acceptable for a slug to be recognized be a phase sensor; default 15.0 µL)
- Slug Volume Threshold with Quench (minimum volume considered acceptable for a slug to be recognized by the pre-HPLC phase sensor; default 15.0 µL)

- HPLC Slug Matching Tolerance ((±) maximum system volume considered acceptable for a slug to be recognized by the pre-HPLC phase sensor; default 25.0 μL)
- Maximum Carrier Flow (maximum slug flow rate; default 250.0 μL/min)
- Minimum Carrier Flow (minimum slug flow rate; default 15.0 μL/min)
- Temperature Tolerance (maximum temperature deviation from set point allowed for slug injection into system; default 1.0°C)
- T Control VISA (COM port for temperature controller; default COM9)
- Reagent Table Path (path to Excel reagent table file)
- Slug Info File (path for slug composition data)
- Slug Status File (path for slug status in system and objective function data)
- System Status File (path for system conditions (flow rate, temperature) data)
- Data Analysis File (path for worked up HPLC data and output reaction concentrations)
- Path for HPLC Chromatograms (path for ChemStation data files)
- Path for Syringe Fill Volumes (path for stored syringe fill volumes)
- Matlab Data File (path for storing MATLAB workspace)

Automatic Refill-Carrier and Base (Flush Base Line)

- Refills both the carrier and online injection line. Purges online injection line into the system. Uses phase sensors to check that system is primed. Creates three blank slugs to clean system after purging with online injection stream. Opens and closes valve on pressure bomb to drain collected liquid to waste.
- Inputs:
  - Carrier Syringe Vol (total volume of carrier syringe; default 8.000 mL)
  - Carrier Push-in Vol (volume to push to waste of carrier line before refill; default 1.0 mL)
  - Carrier Fill Vol (current fill volume of the carrier syringe)
  - Base Syringe Vol (total volume of online injection syringe; default 0.250 mL)
  - Base Push-in Vol (volume to push to waste of online injection line before refill; default 0.200 mL)
  - Base Fill Vol (current fill volume of the online injection syringe)

- Distance to Online Base (system volume from injection valve to online injection point; default is 69.0 μL)

Agilent Start Method

- Sends contact closure command to Agilent HPLC to begin method

Automatic Refill Carrier and Base

- Refills carrier and online injection line but doesn't purge online injection line or introduce blank rinse slugs. Uses phase sensors to check that system is primed. Opens and closes valve on pressure bomb to drain collected liquid to waste.

Automatic Refill LC

- Purges and refills transfer quench line between system sampling loop downstream of reactor and HPLC injection valve. Primes syringe.

Base Phase Sensor Detect Prime

- Checks that online injection line is primed.

Calibration v2

- Structured like Master VI.vi, reads reagent table, identifies reagents to be calibrated, and builds a list of slug experiments with the number of calibration discretizations and replicates specified by user. Outputs least-squares regression. Note: this file was last updated Dec. 2014 and doesn't include some of the most recent corrections to Master VI.vi.

GGSIOC

- Gilson GSIOC communication library.

Gilson 271 Driver

- Driver for GX-271 liquid handler.

Gilson Move to Drain and Dispense Sample

- Moves liquid handler probe to waste vial. Injects into waste vial.

Gilson Dip Needle

- Dips liquid handler probe into a vial, pauses, and removes probe.

Gilson Home Device

- Homes liquid handler.

Gilson Injection Valve

- Switches liquid handler injection valve from load position to inject position.

Gilson LC Valve

- Switches post-reactor sampling valve from sampling position to bypass position.

Gilson Load into Needle

- Reads in reagents to sample in a slug. Creates blank slugs and injects. Samples all reagents sequentially (with gas bubbles and rinses where appropriate). Stirs sample. Injects into injection valve.

Gilson Move to Injection Valve and Dispense Sample

- Moves probe to injection valve. Injects into injection valve.

Gilson Move to Well and Inject Probe

- Moves probe to target vial. Dips probe into vial.

Gilson Move to Well and Withdraw Sample

- Moves probe to target vial. Dips probe into vial and aspirates. Removes probe from vial.

Gilson Static Rinse and Drain Loop Only

- Moves probe to wash solution. Dips probe into wash and aspirates. Removes probe and injects wash solution into injection valve. Switches valve to inject to create blank slug.

Gilson Stir

- Stirs sample in probe by withdrawing and infusing with liquid handler syringe.

Gilson Withdraw into Needle and Inject Hold Needle

- Samples regents and stirs. Injects sample into injection valve. Holds probe in injection valve.

Harvard Change

- Changes syringe pump flow rate.

Harvard Infuse

- Infuses syringe pump to target volume at target flow rate.

Harvard Initialize Injection Base

- Sets syringe diameter, flow rate, and target volume for online injection.

Harvard Inject Base

- Infuses online injection syringe pump to target volume.

Harvard Refill

- Withdraws syringe pump to target volume at target flow rate.

Harvard Reset Volume

- Sets syringe pump target volume.

Harvard Stop

- Stops syringe pump.

Master VI from Restart with Matlab

- Functions identically to Master VI.vi. Reads in old MATLAB file and initializes system with same variable values. Useful in the event of a crash.

Omega T Controller No Loop 3

- Sends the temperature set point to the OMEGA PID controller. Returns the read value and the measured temperature.

Phase Sensor Detect Equilibration

- Uses pre-HPLC phase sensor to check if both quench and gas streams are flowing and thus system is primed.

Prime LH Pump

- Moves liquid handler probe to injection valve. Refills liquid handler syringe and purges line though the injection valve to waste.

pump22~1

- Harvard PhD pump library

Vacuum On

- Opens solenoid valve for a few seconds to pull vacuum on the injection valve waste line. Closes valve.

Valve Switch

- Switches 6-port/2-way valve from refill to infuse.

B.1.4. List of MATLAB Files

(Thank you to Connor Coley for helping compile this list)

Start_slug_tracker

- Inputs the reagent table file and initializes the slug tracker matrix and index
- Beginning of the file includes columns of all variables to track during operation (in *slug_tracker*)

Start_reagent_table

- Inputs the same reagent table file, but outputs *reagent_table* and *reagent_table_index*, which contains all of the reagent information in the excel file
- Identifies the internal standard and decides which reagents to run a calibration with

Slug_analysis

- Updates the *slug_tracker* matrix based on the time spent in the HPLC

Slug_analysis_complete

- Identifies whether the HPLC analysis is done

Slug_analysis_write

- Records the HPLC data for each slug to the specified file *ana_path*

Slug_area

- Calculates the peak areas and concentrations for each slug based on the HPLC output
- Rejects slugs if it cannot retrieve the HPLC file
- Rejects slugs if the internal standard area is not >50% of that expected from previous calibration

Slug_build_prior_opt

- Compiles the final optimization results into a single row with experimental conditions

Slug_calib_ana

- Updates *slug_tracker* for the calibration VI to include HPLC data from *slug_calib_area()* and calculate a calibration coefficient, either assuming a linear or quadratic relationship

Slug_calib_area

- Calculates the concentration and area of a slug for the calibration reagents

Slug_calib_inject

- Determines whether a slug can be injected into the system based on *slug_tracker* during calibration

Slug_calibration

- Equivalent of *start_slug_tracker()* for the calibration VI
- Creates a *slug_tracker* matrix with specific columns to record relevant calibration params

Slug_complete

- Checks whether analysis for a slug is complete, and marks it complete it *slug_tracker* if so

Slug_distance

- Figures out the distance (in μL) how far the slug has traveled
- Takes into account that the flow rate is temperature-dependent (esp. in the reactor)

Slug_exp_cond

- For the optimization, condenses *slug_tracker* information into a format suitable for an input file for the optimization procedure
- Outputs *exp_cond* (list of experimental conditions), and *B* (list of objective function values)

Slug_flowrate

- Takes in *slug_tracker* and figures out what the flow rate should be, both in and out of the reactor
- Decides if the flow rate needs to be changed (and signals so pump can be changed)
- Inputs the sizes of each zone of the system, the buffer size (to equilibrate)
- Before the reactor, the flow rate is halfway between the max and min allowed rate to allow time for mixing; in the reactor, the flow rate is specified by the residence time; after the reactor, the flow rate is halfway between the max and min allowed
- When no slugs are in system, the flow rate is the minimum flow rate to minimize waste

Slug_in_prep

- Figures out if there is a slug to be made, and outputs a matrix *comp* that lists the volumes and indices that the liquid handler must use
- *Prep_slug* is the row number in *slug_tracker* that is in-progress
- First calculates what the volume of the slug must be taking into account online reagent addition, then finds the volume of each reagent, then finds the needed make-up volume
- Tracks the total concentration of the internal standard

Slug_inj_check

- Determines whether a slug is within 1.5**match_tol* of the online injection point distance so that the syringe pump can be prepared for injection
- Outputs 0 to LabView if nothing needs to be injected, >0 otherwise

Slug_inj_fail

- Determines if the slug has gone past the sensor distance + tolerance but has not been seen, the slug is flagged as failed

Slug_inj_online

- Calculates the amount of online-reagent to add and updates *slug_tracker* to note that online-reagents have been added
- Also corrects all upstream slugs' distances by the volume added

Slug_inject_lc

- Determines if the slug is at a distance where HPLC injection should begin (incl. warmup time)

Slug_inject_sys

- Goes through each criterion to see if a slug should be held in the injection valve for longer
- Criteria are (i) hold on system placed by HPLC (ii) hold for refilling syringes (iii) temperature out of tolerance (iv) flowrates exceeding maximum or below minimum (v) slugs wouldn't have enough time in reactor or before HPLC
- Outputs 1 if the slug is ready for injection

Slug_injected

- Flags a slug in *slug_tracker* to note that it has been injected into the system

Slug_injected_lc

- Flags a slug in *slug_tracker* to note that it has been injected into the HPLC

Slug_load

- Interprets the *comp* matrix and reformats for the Gilson Load Needle sub-VI

Slug_match

- Called when the phase sensor sees a slug, and checks to see if any slugs are the *match_tol* of the sensor; if so, updates that slug's distance in *slug_tracker* to the distance of the sensor

Slug_match_lc

- Similar to *slug_match()*, checks that the slug volume is acceptable before injection into HPLC

Slug_objective

- Calculates the function objective value based on HPLC peak areas and concentrations
- Contains the objective function (hard-coded)

Slug_optimization_phase_1_suzuki

- See Appendix C

Slug_optimization_phase_2_suzuki

- See Appendix C

Slug_optimization_phase_3_suzuki

- See Appendix C

Slug_optimization_phase_4

- Conducts gradient-based linesearch initialized with Hessian from phase 3. Determines if convergence criteria are satisfied

Slug_optimization_phase_update

- Determines if the minimum number of experiments for the current phase are complete; if so, increments the phase number

Slug_predictor

- Outputs *slug_event*, which lists the important future times at which events will happen (e.g., flow rate change or slug injection into HPLC)

Slug_read_syringe_vol

- Reads the syringe volumes out of *fill_vol_path*

Slug_refill

- Checks whether a refill is needed for any of the syringes
- Updates the syringe volumes after refill
- Will queue up a hold on the system if there are slugs currently in the system (that is activated once the last slug leaves)

Slug_start_prep

- Determines whether it is ok to start making a slug
- Ensures there won't be multiple slugs in the reactor (to avoid thermocapillary flow problem) and if the system is being held for refill

Slug_temp

- Finds the set point temperature based on *slug_tracker*

Slug_tracker_augment

- Adds a fully-populated row (or rows) to *slug_tracker*
- Also calculates the minimum number of experiments until the next optimization experiment, used by *slug_optimization_phase_update()*

Slug_tracker_rebuild

- Takes in data files from previous runs and recreates the *slug_tracker* matrix
- Note: could also just reload the .mat file saved by LabView instead of using this function

Slug_volume

- Calculate the volume of the slug measured by the phase sensor, based on flow rate
- *Slug_vol1* reports the most recent slug

Slug_volume_ps

- Determines whether the volume found by *slug_volume()* is above the minimum acceptable volume for a slug; if so, reports that the slug has been matched

Slug_volume_react

- Matches the slug and checks that it is large enough
- If it is above the threshold but outside the acceptable tolerance, it flags this as a failed slug

Slug_write

- Records the data for all of the slugs (currently every 20 seconds)
- Writes to files *slug_path*, data regarding the slug composition and setpoints, *status_path*, yes/no indicators and distance, and *system_path*, current temperature and flow rate at various times

Slug_write_init

- Does the same thing as *slug_write()* but also writes the headers for each column

Slug_write_syringe

- Records the volumes of each syringe

Slug_x_linear_dependence

- Checks whether there is linear dependence in experimental conditions
- Not called by LabView, just an auxiliary function for analysis

Slug_x_matrix

- See Appendix C

B.2. DEVICES

B.2.1. Vial Manifold

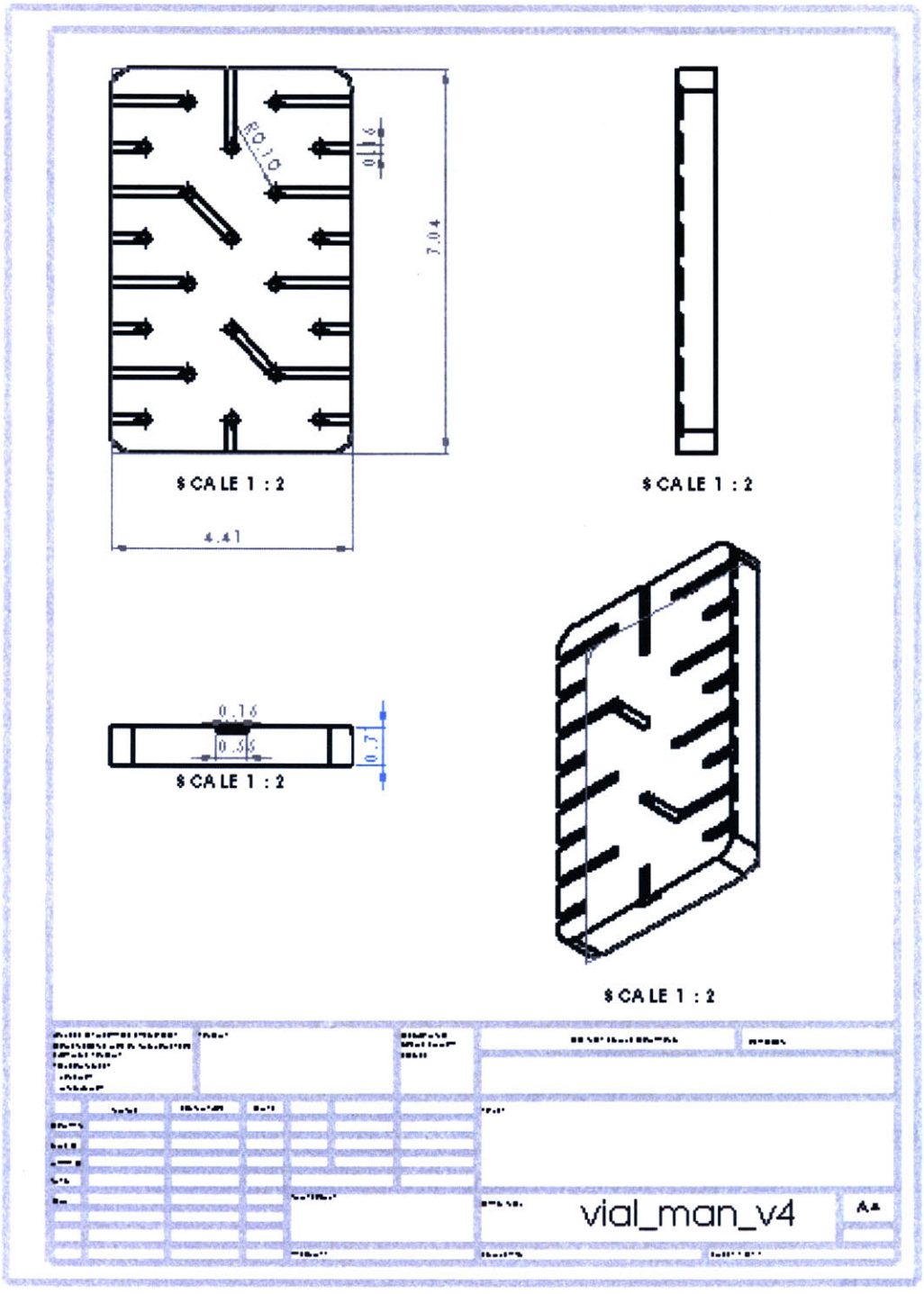


Figure B.1. SOLIDWORKS drawing of vial manifold. SOLIDWORKS file available on KFJSERVER.

B.2.2. Pancake Reactor

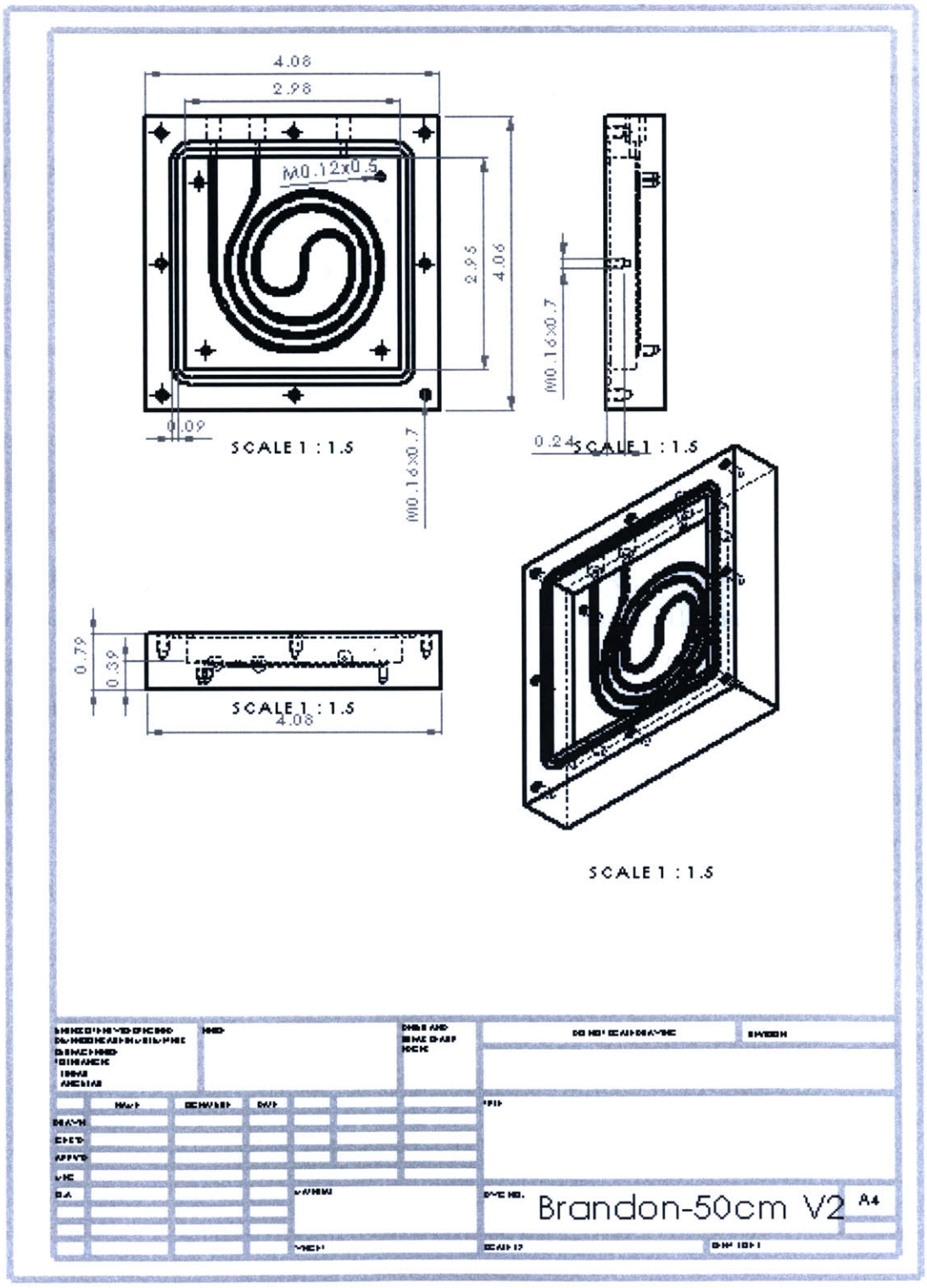


Figure B.2. SOLIDWORKS drawing of pancake reactor. SOLIDWORKS file available on KFJSERVER.

APPENDIX C. CHAPTER 4 SUPPORTING INFORMATION

C.1. OPTIMIZATION SCRIPTS

Included below, the three main optimization scripts along with slug_X_matrix.m progress through the three stages of the adaptive response surface method. Slug_optimization_phase_1_suzuki.m requires the user to input all discrete and continuous variables in the optimization and specify over what continuous range the optimization takes place. The function then constructs an initial fractional factorial design consistent with the number of interaction terms in the model. Slug_optimization_phase_2_suzuki.m receives the matrices of variables passed from phase 1 of the optimization, along with experimental conditions and objective function values from the initial fractional factorial design. This function then estimates the optimum for all discrete variables assuming a linear model and constructs a new set of fractional factorial design experiments in the quadrant of the continuous variable space where the optimum is predicted. Slug_optimization_phase_3_suzuki.m similarly receives the matrices of variables, prior experiments, objective function values, and prior estimates of the optimum and uncertainty. The phase 3 function generates a quadratic response surface model that is optimized to solve for the optimum yield and constrained optimum TON. Discrete variables are fathomed iteratively by a *t*-test, and a new model is constructed after each variable elimination. Finally, new G-optimal experiments are proposed to minimize uncertainty on the remaining candidate discrete variable optima. The function outputs the final optimum TON and yield and associated conditions if termination criteria are satisfied. In phase 2 and phase 3, slug_X_matrix.m is called to construct the matrix of scaled variables **X** from the matrix of experimental conditions.

C.1.1. Optimization Phase 1

```
function
[cv,cv_scale,N_cv,index_cv,dv,N_dv,index_dv,non_opt,index_non_opt,cor,index_cor,I_cv,I_dv,gamma,
exp_cond1] = slug_optimization_phase_1_suzuki()
%SLUG_OPTIMIZATION_PHASE_1_SUZUKI Compiles list of optimized and non-optimized
%continuous and discrete variables. Generates a set of fractional factorial
%design experiments to be optimized spanning full continuous variable space.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Submitted by Brandon Reizman
% May 15, 2015
```

```
% Inputs:
%               None
% Outputs:
%               cv is the matrix of continuous variables
%               cv_scale is the matrix of scaled continuous variables
%               N_cv is the number of continuous variables to optimize
%               index_cv is the identifier assigned to each continuous variable
%               dv is the matrix of discrete variables
%               N_dv is the number of discrete variables to optimize
%               index_dv is the identifier assigned to each discrete variable
%               non_opt is the matrix of continuous variables not to be optimized
%               index_non_opt is the identifier assigned to non-optimized continuous variables
%               cor is the matrix of non-optimized discrete variables
%               index_cor is the identifier assigned to each non-optimized discrete variable
%               I_cv is the matrix of interaction terms between continuous variables
%               I_dv is the matrix of interaction terms between discrete and continuous variables
%               gamma is the fractional yield threshold for optimal TON
%               exp_cond1 is the matrix of experiments to run
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Start by reading in all continuous variables in form
% [(min) (max) (default) (+/- prediction covariance tolerance) (optimize? 1 yes or 0 no)]
% Res Time
cv(1,:) = [60 600 180 540 1];
index_cv(1) ={'Res Time'};
% Temperature
cv(2,:) = [30 110 60 80 1];
index_cv(2) ={'Temp'};
% CB0
cv(3,:) = [0.167*0.005 0.167*0.025 0.167*0.02 0.167*0.02 1];
index_cv(3) ={'Reag 2 Conc'};
% Base Conc
cv(4,:) = [0.333 0.333 0.333 0.01 0];
index_cv(4) ={'Base Conc'};
% CA0
cv(5,:) = [0.167 0.167 0.167 0.01 0];
index_cv(5) ={'Reag 1 Conc'};
% CC0
cv(6,:) = [0.25 0.25 0.25 0.01 0];
index_cv(6) ={'Reag 3 Conc'};
% CD0
cv(7,:) = [9.24 9.24 9.24 0.01 0];
index_cv(7) ={'Reag 4 Conc'};
% CE0
cv(8,:) = [0.0 0.0 0.0 0.01 0];
index_cv(8) ={'Reag 5 Conc'};
% Precat Conc
```

```
cv(9,:) = [0.0 0.0 0.0 0.01 0];
index_cv(9) ={'Precat Conc'};
% Slug Vol
cv(10,:) = [35 35 35 1 0];
index_cv(10) ={'Slug Vol'};

% Read in discrete variables [{list of possible values}]
% Reagent 2
dv(1,:) = [1 2 3 4 5 6 7 8];
index_dv(1) ={'Reagent 2'};

% Read in correlated variables [{list of possible values} {corresponding discrete variable
number}]
cor(1,:) = [49*ones(size(dv(1,:))) 1];
index_cor(1) ={'Reagent 1'};
cor(2,:) = [55*ones(size(dv(1,:))) 1];
index_cor(2) ={'Reagent 3'};
cor(3,:) = [30*ones(size(dv(1,:))) 1];
index_cor(3) ={'Reagent 4'};
cor(4,:) = [30*ones(size(dv(1,:))) 1];
index_cor(4) ={'Reagent 5'};
cor(5,:) = [19*ones(size(dv(1,:))) 1];
index_cor(5) ={'Makeup'};

% Read in continuous variable interaction terms to consider (columns are
% continous variables)
I_cv = [1 1 1; 0 1 1; 0 0 1];

% Read in discrete variable interaction terms to consider (columns are
% continous variables)
I_dv = [0 1 0];

% Cutoff for yield in catalyst optimization
gamma = 0.90;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Sort cv
[cv,sort_cv] = sortrows(-cv,length(cv(1,:)));
cv = -cv;
index_cv = index_cv(sort_cv);

% Count continuous variables
N_cv = sum(cv(:,end));

% Randomize continuous variables
new_stream = RandStream.create('mt19937ar','seed',sum(100*clock));
RandStream.setDefaultStream(new_stream);
```

```
r = rand(N_cv,1);
[sorted,order] = sort(r);
cv(1:N_cv,:) = cv(order,:);
index_cv(1,1:N_cv) = index_cv(1,order');
I_dv = I_dv(1,order);
for i = 1:N_cv
    for j = i:N_cv
        I_cv_order(order(i),order(j)) = I_cv(i,j);
        if I_cv_order(order(i),order(j)) == 1
            if order(i) > order(j)
                I_cv_order(order(i),order(j)) = 0;
                I_cv_order(order(j),order(i)) = 1;
            end
        end
    end
end
I_cv = I_cv_order;

% Create scaled continuous variable matrix
cv_scale = cv;

% Apply proper scaling
for i = 1:length(cv(:,1))
    if strcmp(index_cv{i},'Res Time') == 1
        cv_scale(i,1:3) = log(cv(i,1:3));
    elseif strcmp(index_cv{i},'Temp') == 1
        cv_scale(i,1:3) = (cv(i,1:3) + 273.15).^-1;
    elseif strcmp(index_cv{i},'Base Conc') == 1
        cv_scale(i,1:3) = log(cv(i,1:3));
    elseif strncmp(index_cv{i},'Reag',4) == 1
        cv_scale(i,1:3) = log(cv(i,1:3));
    end
end

% Count discrete variables
N_dv = length(dv(1,:));

% Randomize discrete variables
for i = 1:length(dv(:,1))
    r = rand(size(dv(i,:)));
    [sorted,order] = sort(r);
    dv(i,:) = dv(i,order);
    for j = 1:length(cor(:,1))
        if cor(j,end) == i
            cor(j,1:end-1) = cor(j,order);
        end
    end
end
```

```
% Generate initial factorial design matrix
min_design1 = N_cv + N_dv + 1;
min_design2 = (N_cv - sum(I_dv)) + N_dv + N_dv*sum(I_dv) + sum(sum(I_cv)) + 1;
% Size of first fractional factorial design
design1 = N_dv;
% Size of first and second fractional factorial designs
design2 = 2*design1;
fraction = N_cv;
while design1 <= min_design1 || design2 <= min_design2
    % Increase fractional factorial size
    fraction = fraction - 1;
    design1 = 2^(N_cv - fraction)*N_dv;
    design2 = 2*design1;
end

% Build generator string
gen_str = [sprintf('')];
for i = 1:N_cv + N_dv - fraction
    gen_str = [gen_str char(96+i) ' '];
end
for j = 1:fraction
    if fraction >= 2
        for i = j:N_cv - fraction + j
            gen_str = [gen_str char(96+i)];
        end
        for i = j + N_cv:j + N_cv + floor(N_dv/2 - 1) - 1
            gen_str = [gen_str char(96+i)];
        end
    else
        for i = j:N_cv - fraction + j
            gen_str = [gen_str char(96+i)];
        end
    end
    gen_str = [gen_str ' '];
end

% Create fractional factorial design
exp_cond_scale1 = fracfact(gen_str);
exp_cond_scale_row = 0;
for i = 1:length(exp_cond_scale1(:,1))
    if sum(exp_cond_scale1(i,N_cv+1:end)) == -N_dv + 2
        exp_cond_scale_row = exp_cond_scale_row + 1;
        index = find(exp_cond_scale1(i,N_cv + 1:N_cv + N_dv) == 1);
        exp_cond_scale(exp_cond_scale_row,:) = [0.5*exp_cond_scale1(i,1:N_cv)+1.5 index];
    end
end
```

```
% Block randomize experiments
exp_cond_scale(:,end + 1) = rand(length(exp_cond_scale(:,1)),1);
exp_cond_scale = sortrows(exp_cond_scale,length(exp_cond_scale(1,:)));
exp_cond_scale(:,end) = [ ];
% Group by temperature
col_T = find(strcmp(deblank(index_cv),'Temp') == 1);
for i = 1:length(exp_cond_scale(:,1))-1
    swap_row = find(exp_cond_scale(i+1:end,col_T) == exp_cond_scale(i,col_T),1,'first');
    if isempty(swap_row) == 0
        exp_cond_scale_swap = exp_cond_scale;
        exp_cond_scale(i+1,:) = exp_cond_scale_swap(i+swap_row,:);
        exp_cond_scale(i+swap_row,:) = exp_cond_scale_swap(i+1,:);
    end
end

x_opt_ub = ones(1,N_cv);
x_opt_lb = -1*ones(1,N_cv);

% Generate initial matrix of conditions
for i = 1:length(exp_cond_scale(1,:))
    if i <= N_cv
        for row = 1:length(exp_cond_scale(:,1))
            if exp_cond_scale(row,i) == 1
exp_cond_scale(row,i) = x_opt_lb(1,i)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
            else
exp_cond_scale(row,i) = x_opt_ub(1,i)*(cv_scale(i,2) - cv_scale(i,1))/2 + mean([cv_scale(i,1)
cv_scale(i,2)]);
            end
        end
    else
        for row = 1:length(exp_cond_scale(:,1))
            exp_cond_scale(row,i) = dv(i - N_cv,exp_cond_scale(row,i));
        end
    end
end

% Apply proper scaling
exp_cond1 = exp_cond_scale;
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Res Time') == 1
        exp_cond1(:,i) = exp(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Temp') == 1
        exp_cond1(:,i) = exp_cond_scale(:,i).^-1 - 273.15;
    elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale(:,i));
    elseif strncmp(deblank(index_cv(i)),'Reag',4) == 1
        exp_cond1(:,i) = exp(exp_cond_scale(:,i));
```

```
    end
end

% Include non-optimized continuous variables
non_opt = find(cv(:,end) == 0);
index_non_opt = index_cv(non_opt);
for i = length(non_opt):-1:1
    index_cv(non_opt(i)) = [ ];
    exp_cond1(:,N_cv + length(dv(:,1)) + i) = cv(non_opt(i),3);
end

% Include correlated discrete variables
for i = 1:length(cor(:,1))
    dv_row = cor(i,end);
    for row = 1:length(exp_cond1(:,1))
        j = find(dv(dv_row,:) == exp_cond1(row,N_cv + dv_row),1);
        exp_cond1(row,N_cv + length(dv(:,1)) + length(non_opt) + i) = cor(i,j);
    end
end

end
```

C.1.2. Optimization Phase 2

```
function
[X,col_index,theta,V_B,V_J,x_maxJ,J_maxJ,x_maxY,Y_maxY,J_opt,J_opt_lb,dJ,dv_fathom,exp_cond1] =
slug_optimization_phase_2_suzuki(cv,cv_scale,N_cv,index_cv,dv,N_dv,non_opt,index_non_opt,cor,
I_cv,I_dv,gamma,exp_cond,B)
%SLUG_OPTIMIZATION_PHASE_2_SUZUKI Generates a linearized model and an updated
%set of targeted fractional factorial design experiments
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Submitted by Brandon Reizman
% May 15, 2015
% Inputs:
%               cv is the matrix of continuous variables
%               cv_scale is the matrix of scaled continuous variables
%               N_cv is the number of continuous variables to optimize
%               index_cv is the identifier assigned to each continuous variable
%               dv is the matrix of discrete variables
%               N_dv is the number of discrete variables to optimize
%               non_opt is the matrix of continuous variables not to be optimized
%               index_non_opt is the identifier assigned to non-optimized continuous variables
%               cor is the matrix of non-optimized discrete variables
%               I_cv is the matrix of interaction terms between continuous variables
%               I_dv is the matrix of interaction terms between discrete and continuous variables
%               gamma is the fractional yield threshold for optimal TON
%               exp_cond is the matrix of experiments previously run
```

```
%               B is the vector of objective function values for exp_cond
% Outputs:
%               X is the matrix of scaled experimental conditions
%               col_index is the matrix identifying the columns of X
%               theta is the optimal parameter vector
%               V_B is the estimated response covariance
%               V_J is the prediction covariance at the optimum
%               x_maxJ is the optimal scaled experimental conditions for each discrete variable
%               J_maxJ is the optimal objective function value for each discrete variable
%               x_maxY is the optimal scaled experimental conditions for yield for each variable
%               Y_maxY is the optimal objective function value for yield for each variable
%               J_opt is the matrix of optimal values
%               J_opt_lb is the matrix of 95% confidence lower bounds on the optimal cost
%               dJ is the matrix of errors in J_opt
%               dv_fathom is the list of fathomed discrete variables
%               exp_cond1 is the matrix of experiments to run
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Scale exp_cond
exp_cond_scale = exp_cond;
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Res Time') == 1
        exp_cond_scale(:,i) = log(exp_cond(:,i));
    elseif strcmp(deblank(index_cv(i)),'Temp') == 1
        exp_cond_scale(:,i) = (exp_cond(:,i) + 273.15).^-1;
    elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 1 Conc') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    end
end

% Build X matrix
[X,col_index] =
slug_X_matrix(cv_scale,N_cv,dv,N_dv,eye(length(I_cv(:,1))),zeros(size(I_dv)),exp_cond_scale,'off'
);

% Find starting material index
for i = 1:size(non_opt,1)
    if strcmp(deblank(index_non_opt(i)),'Reag 1 Conc') == 1
        sm_index = i + N_cv;
        sm_conc = cv(sm_index,3);
    end
end
```

```
% Find loading index
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
        loading_index = i;
    end
end

% log catalyst loading
ln_loading = exp_cond_scale(:,loading_index);

% Log(yield)
ln_yield = B + ln_loading - log(sm_conc);

W1 = eye(length(B));
for i = 1:length(B)
    W1(i,i) = exp(ln_yield(i))/sum(exp(ln_yield));
    %W1(i,i) = exp(B(i))/sum(exp(B));
end

% Find best fit parameters
theta = (X'*W1*X)\(X'*W1*B);

% Weighting matrix
W = eye(length(B));

% Residuals
e = B - X*theta;

% Jackknife estimation of response covariance
for i = 1:length(B)
    Xi = X([1:i-1,i+1:end],:);
    Bi = B([1:i-1,i+1:end]);
    Wi = W1([1:i-1,i+1:end],[1:i-1,i+1:end]);
    thetai = (Xi'*Wi*Xi)\(Xi'*Wi*Bi);
    ei = Bi - Xi*thetai;
    V_Bi(i,1) = ei'*ei/(length(Bi) - length(thetai));
end
V_B_jack = mean(sqrt(V_Bi))^2;
V_B = (length(B) - 1)/length(B)*sum((sqrt(V_Bi) - sqrt(V_B_jack)).^2);

% This is true if using same model for yield and TON
V_yield = V_B;

% Find current minimum
options = optimset('TolFun',1e-10,'TolX',1e-10,'TolCon',1e-10,'Display','off','Algorithm','SQP');
Aeq = zeros(N_dv,N_cv + N_dv);
for i = 1:N_dv
    Aeq(i,N_cv + i) = 1;
```

```
end

% Build A and c (with current values)
c = zeros(N_cv + N_dv,1);
A = zeros(N_cv + N_dv,N_cv + N_dv);
for i = 1:length(theta)
    [row,col] = find(col_index == i);
    if isempty(row) == 0
        if row == col
            c(row,1) = theta(i);
        else
            A(row,col) = 0.5*theta(i);
            A(col,row) = 0.5*theta(i);
        end
    else
        col = find(col_index_square == i);
        A(col,col) = theta(i);
    end
end

% Yield optimization
for i = 1:N_dv
    % Constrain discrete variables
    Beq = zeros(N_dv,1);
    Beq(i,1) = 1;
    % Maximize response surface function
    [x,Y,exitflag,output,lambda,gradient,hessian] =
fmincon(@maxY,[zeros(N_cv,1);Beq],[],[],Aeq,Beq,[-1*ones(N_cv,1);zeros(N_dv,1)],ones(N_cv +
N_dv,1),[],options,loading_index,A,c,cv_scale,sm_conc);
    x_maxY(i,:) = x;
    Y_maxY(i,1) = -Y;
end

% Find maximum for all cases
[Y_opt,dv_id] = max(Y_maxY);
yield_opt = x_maxY(dv_id,:);

% Optimum prediction covariance
for i = 1:length(theta)
    [row,col] = find(col_index == i);
    if isempty(col) == 1
        row = find(col_index_square == i);
        X_prime(1,i) = yield_opt(row)^2;
    else
        if row == col
            X_prime(1,i) = yield_opt(row);
        else
            X_prime(1,i) = yield_opt(row)*yield_opt(col);
```

```
        end
    end
end
V_Y = X_prime*(X'*X)^-1*X_prime'*V_yield;

% TON optimization
for i = 1:N_dv
    % Constrain discrete variables
    Beq = zeros(N_dv,1);
    Beq(i,1) = 1;
    % Maximize TON subject to yield <= gamma*max(yield)
[x,J,exitflag,output,lambda,gradient,hessian] = fmincon(@maxJ,x_maxY(i,:)',[],[],Aeq,Beq,[-
1*ones(N_cv,1);zeros(N_dv,1)],ones(N_cv +
N_dv,1),@maxJ_nlcon,options,loading_index,A,c,gamma,max(Y_maxY(:,1)),cv_scale,sm_conc);
    if exitflag <= 0
        x = x_maxY(i,:);
        J = maxJ(x',loading_index,A,c,gamma,x_maxY(i,:),cv_scale,sm_conc);
    end
    x_maxJ(i,:) = x;
    J_maxJ(i,1) = -J;
    H_inv(N_cv*(i-1)+1:N_cv*i,1:N_cv) = hessian(1:N_cv,1:N_cv)^-1;
end

% Find maximum for all cases
[J_max,dv_id] = max(J_maxJ);
x_opt = x_maxJ(dv_id,:)';
J_opt = [length(exp_cond(:,1)) J_max];

% Optimum prediction covariance
for i = 1:length(theta)
    [row,col] = find(col_index == i);
    if isempty(col) == 1
        row = find(col_index_square == i);
        X_prime(1,i) = x_opt(row)^2;
    else
        if row == col
            X_prime(1,i) = x_opt(row);
        else
            X_prime(1,i) = x_opt(row)*x_opt(col);
        end
    end
end
% Prediction covariance
V_J = X_prime*(X'*X)^-1*X_prime'*V_B;

% Optimum lower bound
J_max_lb = J_opt(end,2) - sqrt(V_J)*tinv(1-0.05,length(B) - length(theta));
J_opt_lb = [length(exp_cond(:,1)) J_max_lb];
```

```
dJ = [length(exp_cond(:,1)) (J_opt(end,2) - J_opt_lb(end,2))];

% Check of any discrete variable can be fathomed (if J_int < J_opt_lb)
dv_fathom_row = 0;
dv_fathom = [ ];

% New centered design
x_opt_ub = zeros(N_dv,length(theta));
x_opt_lb = zeros(N_dv,length(theta));
for dv_id = 1:N_dv
    for i = 1:N_cv
        if x_maxJ(dv_id,i) >= 0
            x_opt_ub(dv_id,col_index(i,i)) = 1;
            x_opt_lb(dv_id,col_index(i,i)) = 0;
        else
            x_opt_ub(dv_id,col_index(i,i)) = 0;
            x_opt_lb(dv_id,col_index(i,i)) = -1;
        end
    end
    x_opt_ub(dv_id,col_index(N_cv + dv_id,N_cv + dv_id)) = 1;
    x_opt_lb(dv_id,col_index(N_cv + dv_id,N_cv + dv_id)) = 1;
end

% Generate new factorial design matrix
fraction = ceil(N_cv - log(length(exp_cond(:,1))/N_dv)/log(2));

% Build generator string
gen_str = [sprintf('')];
for i = 1:N_cv + N_dv - fraction
    gen_str = [gen_str char(96 + i) ' '];
end
for j = 1:fraction
    gen_str = [gen_str '-'];
    if fraction >= 2
        for i = j:N_cv - fraction + j
            gen_str = [gen_str char(96+i)];
        end
        for i = j + N_cv:j + N_cv + floor(N_dv/2 - 2) - 1
            gen_str = [gen_str char(96+i)];
        end
    else
        for i = j:N_cv - fraction + j
            gen_str = [gen_str char(96+i)];
        end
    end
    gen_str = [gen_str ' '];
end
```

```
% Create fractional factorial design
exp_cond_scale1 = fracfact(gen_str);
exp_cond_scale_row = 0;
for i = 1:length(exp_cond_scale1(:,1))
    if sum(exp_cond_scale1(i,N_cv+1:end)) == -N_dv + 2
        exp_cond_scale_row = exp_cond_scale_row + 1;
        index = find(exp_cond_scale1(i,N_cv + 1:N_cv + N_dv) == 1);
        exp_cond_scale2(exp_cond_scale_row,:) = [0.5*exp_cond_scale1(i,1:N_cv)+1.5 index];
    end
end

% Generate matrix of conditions
for i = 1:length(exp_cond_scale2(1,:))
    if i <= N_cv
        col = col_index(i,i);
        for row = 1:length(exp_cond_scale2(:,1))
            if exp_cond_scale2(row,i) == 1
                dv_row = exp_cond_scale2(row,N_cv + 1);
exp_cond_scale2(row,i) = x_opt_lb(dv_row,col)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
            else
                dv_row = exp_cond_scale2(row,N_cv + 1);
exp_cond_scale2(row,i) = x_opt_ub(dv_row,col)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
            end
        end
    else
        for row = 1:length(exp_cond_scale2(:,1))
            exp_cond_scale2(row,i) = dv(i - N_cv,exp_cond_scale2(row,i));
        end
    end
end

% Block randomize experiments
new_stream = RandStream.create('mt19937ar','seed',sum(100*clock));
RandStream.setGlobalStream(new_stream);
exp_cond_scale2(:,end + 1) = rand(length(exp_cond_scale2(:,1)),1);
exp_cond_scale2 = sortrows(exp_cond_scale2,length(exp_cond_scale2(1,:)));
exp_cond_scale2(:,end) = [ ];
% Group by temperature
col_T = find(strcmp(deblank(index_cv),'Temp') == 1);
for i = 1:length(exp_cond_scale2(:,1))-1
    swap_row = find(exp_cond_scale2(i+1:end,col_T) == exp_cond_scale2(i,col_T),1,'first');
    if isempty(swap_row) == 0
        exp_cond_scale_swap = exp_cond_scale2;
        exp_cond_scale2(i+1,:) = exp_cond_scale_swap(i+swap_row,:);
        exp_cond_scale2(i+swap_row,:) = exp_cond_scale_swap(i+1,:);
    end
```

```
end

% Apply proper scaling
exp_cond1 = exp_cond_scale2;
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Res Time') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    elseif strcmp(deblank(index_cv(i)),'Temp') == 1
        exp_cond1(:,i) = exp_cond_scale2(:,i).^-1 - 273.15;
    elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 1 Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    end
end

% Include non-optimized continuous variables
for i = 1:length(non_opt)
    exp_cond1(:,N_cv + length(dv(:,1)) + i) = cv(non_opt(i),3);
end

% Include correlated discrete variables
for i = 1:length(cor(:,1))
    dv_row = cor(i,end);
    for row = 1:length(exp_cond1(:,1))
        j = find(dv(dv_row,:) == exp_cond1(row,N_cv + dv_row),1);
        exp_cond1(row,N_cv + length(dv(:,1)) + length(non_opt) + i) = cor(i,j);
    end
end

end

function [Cineq,Ceq] = maxJ_nlcon(x,loading_index,A,c,gamma,Yi,cv_scale,sm_conc)
% Constraints for TON maximization

% Response surface prediction
Y = -maxY(x,loading_index,A,c,cv_scale,sm_conc);

% Inequality
Cineq = -Y + Yi + log(gamma);

% Equality
Ceq = [ ];

end
```

```
function [J,g] = maxJ(x,loading_index,A,c,gamma,Yi,cv_scale,sm_conc)
% Objective function for TON maximization

% Objective Function
J = -x'*A*x - c'*x;

% Gradient
if nargout > 1
    g = (-2*x'*A - c')';
end

end

function [Y,g] = maxY(x,loading_index,A,c,cv_scale,sm_conc)
% Objective function for yield maximization

% log catalyst loading
ln_loading = (x(loading_index) + 1)*(cv_scale(loading_index,2) - cv_scale(loading_index,1))/2 +
cv_scale(loading_index,1);

% Objective Function
Y = -x'*A*x - c'*x - ln_loading + log(sm_conc);

% Gradient
if nargout > 1
    g = (-2*x'*A - c')';
    g(loading_index) = g(loading_index) - (cv_scale(loading_index,2) -
cv_scale(loading_index,1))/2;
end

end
```

C.1.3. Optimization Phase 3

```
function
[X,col_index,col_index_square,theta,V_B,V_J,H_inv,x_maxJ,J_maxJ,x_maxY,Y_maxY,J_opt,J_opt_lb,dJ,
x_mindJ,dJ_mindJ,dv_fathom,exp_cond1,optim_phase,final_opt] =
slug_optimization_phase_3_suzuki(cv,cv_scale,N_cv,index_cv,dv,N_dv,non_opt,index_non_opt,cor,
I_cv,I_dv,gamma,J_opt_prior,J_opt_lb_prior,dJ_prior,exp_cond,B)
%SLUG_OPTIMIZATION_PHASE_3_SUZUKI Generates a quadratic response surface
%model, optimizes, determines which discrete variables to fathom, and finds
%the next g-optimal experiment to run for all non-fathomed discrete variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Submitted by Brandon Reizman
% May 15, 2015
% Inputs:
%               cv is the matrix of continuous variables
```

```
%               cv_scale is the matrix of scaled continuous variables
%               N_cv is the number of continuous variables to optimize
%               index_cv is the identifier assigned to each continuous variable
%               dv is the matrix of discrete variables
%               N_dv is the number of discrete variables to optimize
%               non_opt is the matrix of continuous variables not to be optimized
%               cor is the matrix of non-optimized discrete variables
%               I_cv is the matrix of interaction terms between continuous variables
%               I_dv is the matrix of interaction terms between discrete and continuous variables
%               gamma is the tolerance on maximum yield for TON
%               J_opt_prior is the prior optimum values
%               J_opt_lb_prior is the prior optimum lower bound values
%               dJ_prior is the prior error in J_opt
%               exp_cond is the matrix of experiments previously run
%               B is the vector of objective function values for exp_cond
% Outputs:
%               X is the matrix of scaled experimental conditions
%               col_index is the matrix identifying the columns of X
%               col_index_square is the matrix identifying the squared term columns of X
%               theta is the optimal parameter vector
%               V_B is the estimated response covariance
%               V_J is the prediction covariance at the optimum
%               H_inv is the inverse of the objective function Hessian at the optimum
%               x_maxJ is the optimal scaled experimental conditions for each discrete variable
%               J_maxJ is the optimal objective function value for each discrete variable
%               x_maxY is the optimal scaled experimental conditions for yield
%               Y_maxY is the optimal objective function value for yield
%               J_opt is the matrix of optimal values
%               J_opt_lb is the matrix of 95% confidence lower bound on the optimum
%               dJ is the matrix of errors in J_opt
%               x_mindJ is the experimental conditions which minimize uncertainty in J for each
discrete variable
%               dJ_mindJ is the minimum uncertainty for each x_mindJ
%               dv_fathom is the list of fathomed discrete variables
%               exp_cond1 is the matrix of experiments to run
%               optim_phase is the phase of the optimization (3 if repeat, 4 if move on)
%               final_opt is the list of optimal conditions (both yield and TON)
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Scale exp_cond
exp_cond_scale = exp_cond;
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Res Time') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Temp') == 1
        exp_cond_scale(:,i) = (exp_cond_scale(:,i) + 273.15).^-1;
    elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
```

```
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 1 Conc') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
        exp_cond_scale(:,i) = log(exp_cond_scale(:,i));
    end
end

% Find starting material index
for i = 1:size(non_opt,1)
    if strcmp(deblank(index_non_opt(i)),'Reag 1 Conc') == 1
        sm_index = i + N_cv;
        sm_conc = cv(sm_index,3);
    end
end

elim_row = 0;
dv_update = dv;
N_dv_update = N_dv;
exp_cond_scale_update = exp_cond_scale;
exp_cond_update = exp_cond;
B_update = B;
dv_fathom = [ ];
dv_fathom_change = 1;
exp_cond2 = [ ];

% Loop until no more fathomed discrete variables are in reduced model
while dv_fathom_change == 1

    x_maxY = zeros(N_dv_update,N_cv + N_dv_update);
    Y_maxY = zeros(N_dv_update,1);
    x_maxJ = zeros(N_dv_update,N_cv + N_dv_update);
    J_maxJ = zeros(N_dv_update,1);
    X_prime = [ ];
    dv_fathom_change = 0;

    % Build X matrix
    [X,col_index,col_index_square] =
slug_X_matrix(cv_scale,N_cv,dv_update,N_dv_update,I_cv,I_dv,exp_cond_scale_update,'on');

    % Find loading index
    for i = 1:N_cv
        if strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
            loading_index = i;
        end
    end

    % log catalyst loading
```

```
    ln_loading = exp_cond_scale_update(:,loading_index);

    % log yield
    ln_yield = B_update + ln_loading - log(sm_conc);

    W1 = eye(length(B_update));
    for i = 1:length(B_update)
        W1(i,i) = exp(ln_yield(i))/sum(exp(ln_yield));
        %W1(i,i) = exp(B_update(i))/sum(exp(B_update));
    end

    % Find best fit parameters
    theta = (X'*W1*X)\(X'*W1*B_update);

    % Residuals
    e = B_update - X*theta;

    % Jackknife estimation of response covariance
    Xi = [ ];
    Bi = [ ];
    Wi = [ ];
    thetai = [ ];
    ei = [ ];
    V_Bi = [ ];
    for i = 1:length(B_update)
        Xi = X([1:i-1,i+1:end],:);
        Bi = B_update([1:i-1,i+1:end]);
        Wi = W1([1:i-1,i+1:end],[1:i-1,i+1:end]);
        thetai = (Xi'*Wi*Xi)\(Xi'*Wi*Bi);
        ei = Bi - Xi*thetai;
        V_Bi(i,1) = ei'*ei/(length(Bi) - length(thetai));
    end
    V_B_jack = mean(sqrt(V_Bi))^2;
    V_B = (length(B_update) - 1)/length(B_update)*sum((sqrt(V_Bi) - sqrt(V_B_jack)).^2);

    % This is true if using same model for yield and TON
    V_yield = V_B;

    % Find current minimum
    options = optimset('TolFun',1e-10,'TolX',1e-10,'TolCon',1e-
10,'Display','off','Algorithm','SQP');
    Aeq = zeros(N_dv_update,N_cv + N_dv_update);
    for i = 1:N_dv_update
        Aeq(i,N_cv + i) = 1;
    end

    % Build A and c (with current values)
    c = zeros(N_cv + N_dv_update,1);
```

```
    A = zeros(N_cv + N_dv_update,N_cv + N_dv_update);
    for i = 1:length(theta)
        [row,col] = find(col_index == i);
        if isempty(row) == 0
            if row == col
                c(row,1) = theta(i);
            else
                A(row,col) = 0.5*theta(i);
                A(col,row) = 0.5*theta(i);
            end
        else
            col = find(col_index_square == i);
            A(col,col) = theta(i);
        end
    end

    % Yield optimization
    for i = 1:N_dv_update
        % Constrain discrete variables
        Beq = zeros(N_dv_update,1);
        Beq(i,1) = 1;
        % Maximize response surface function
        [x,Y,exitflag,output,lambda,gradient,hessian] =
fmincon(@maxY,[zeros(N_cv,1);Beq],[],[],Aeq,Beq,[-1*ones(N_cv,1);zeros(N_dv_update,1)],ones(N_cv
+ N_dv_update,1),[],options,loading_index,A,c,cv_scale,sm_conc);
        x_maxY(i,:) = x;
        Y_maxY(i,1) = -Y;
    end

    % Find maximum for all cases
    [Y_opt,dv_id] = max(Y_maxY);
    yield_opt = x_maxY(dv_id,:);

    % Optimum prediction covariance
    for i = 1:length(theta)
        [row,col] = find(col_index == i);
        if isempty(col) == 1
            row = find(col_index_square == i);
            X_prime(1,i) = yield_opt(row)^2;
        else
            if row == col
                X_prime(1,i) = yield_opt(row);
            else
                X_prime(1,i) = yield_opt(row)*yield_opt(col);
            end
        end
    end
    V_Y = X_prime*(X'*V_yield^-1*X)^-1*X_prime';
```

```
    % TON optimization
    for i = 1:N_dv_update
        % Constrain discrete variables
        Beq = zeros(N_dv_update,1);
        Beq(i,1) = 1;
        % Maximize TON subject to yield <= gamma*max(yield)
        [x,J,exitflag,output,lambda,gradient,hessian] =
fmincon(@maxJ,x_maxY(i,:)',[],[],Aeq,Beq,[-1*ones(N_cv,1);zeros(N_dv_update,1)],ones(N_cv +
N_dv_update,1),@maxJ_nlcon,options,loading_index,A,c,gamma,max(Y_maxY(:,1)),cv_scale,sm_conc);
        if exitflag <= 0
           x = x_maxY(i,:);
           J = maxJ(x',loading_index,A,c,gamma,x_maxY(i,:),cv_scale,sm_conc);
        end
        x_maxJ(i,:) = x;
        J_maxJ(i,1) = -J;
        H_inv(N_cv*(i-1)+1:N_cv*i,1:N_cv) = hessian(1:N_cv,1:N_cv)^-1;
    end

    % Find maximum for all cases
    [J_max,J_opt_dv_id] = max(J_maxJ);
    x_opt = x_maxJ(J_opt_dv_id,:);
    J_opt = [J_opt_prior; length(exp_cond(:,1)) J_max];

    clear V_J_mat N_J
    % Estimation of prediction covariance for each discrete variable
    for j = 1:N_dv_update

        % Scale optimum for discrete variable
        X_keep = [ ];
        col_keep = [ ];
        x_opt = x_maxJ(j,:);
        for i = 1:length(theta)
            [row,col] = find(col_index == i);
            if isempty(col) == 1
                row = find(col_index_square == i);
                X_prime(1,i) = x_opt(row)^2;
                col_keep = [col_keep; i];
            else
                if max([row col]) <= N_cv
                    col_keep = [col_keep; i];
                elseif col == N_cv + j
                    col_keep = [col_keep; i];
                end
                if row == col
                    X_prime(1,i) = x_opt(row);
                else
                    X_prime(1,i) = x_opt(row)*x_opt(col);
```

```
            end
        end
    end

    if J_maxJ(j,1) == max(J_maxJ)

        % Overall prediction covariance
        V_J_mat(j,1) = X_prime*(X'*V_B^-1*X)^-1*X_prime';
        N_J(j,1) = length(B);

    else

        % Trust-region prediction covariance
        for i = 1:N_cv
            exp_cond_scale_j(1,i) = x_maxJ(j,i)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
        end

        % Apply proper scaling
        exp_cond_j = exp_cond_scale_j;
        for i = 1:N_cv
            if strcmp(deblank(index_cv(i)),'Res Time') == 1
                exp_cond_j(:,i) = exp(exp_cond_scale_j(:,i));
            elseif strcmp(deblank(index_cv(i)),'Temp') == 1
                exp_cond_j(:,i) = exp_cond_scale_j(:,i).^-1 - 273.15;
            elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
                exp_cond_j(:,i) = exp(exp_cond_scale_j(:,i));
            elseif strcmp(deblank(index_cv(i)),'Reag 1 Conc') == 1
                exp_cond_j(:,i) = exp(exp_cond_scale_j(:,i));
            elseif strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
                exp_cond_j(:,i) = exp(exp_cond_scale_j(:,i));
            end
        end

        % Find experiments run with discrete variable or within cv tolerance
        for i = 1:length(X(:,1))
            if abs(exp_cond_update(i,1:N_cv) - exp_cond_j) < cv(1:N_cv,4)'
                X_keep = [X_keep; i];
            elseif exp_cond_update(i,N_cv + 1) == dv_update(j)
                X_keep = [X_keep; i];
            end
        end
        if length(X_keep) > length(col_keep)
            V_J_mat(j,1) = X_prime(col_keep)*(X(X_keep,col_keep)'*V_B^-
1*X(X_keep,col_keep))^-1*X_prime(col_keep)';
        else
            V_J_mat(j,1) = NaN;
        end
```

```
            N_J(j,1) = length(X_keep);
        end
    end
    V_J = V_J_mat(J_opt_dv_id,1);

    % Optimum lower bound
    J_max_lb = J_opt(end,2) - sqrt(V_J)*tinv(1-0.05,length(B_update) - length(theta));
    J_max_ub = J_opt(end,2) + sqrt(V_J)*tinv(1-0.05,length(B_update) - length(theta));
    J_opt_lb = [J_opt_lb_prior; length(exp_cond(:,1)) J_max_lb];
    dJ = [dJ_prior; length(exp_cond(:,1)) (J_opt(end,2) - J_opt_lb(end,2))];

    dv_fathom_row = length(dv_fathom);
    clear J_maxJ_rows
    J_maxJ_rows = find(N_J > length(col_keep));
    [J_maxJ_min_sort,sort_order] = sort(J_maxJ(J_maxJ_rows));
    sort_order_row = 0;

    % Scan through eligible discrete variables, see if any can be fathomed
    while dv_fathom_change == 0 && sort_order_row ~= length(sort_order)

        sort_order_row = sort_order_row + 1;
        i = J_maxJ_rows(sort_order(sort_order_row));

        % Unpaired, 2 sample t-test
        dof = (V_J_mat(i,1) + V_J_mat(J_opt_dv_id,1))^2/(V_J_mat(i,1)^2/(N_J(i) -
length(col_keep)) + V_J_mat(J_opt_dv_id,1)^2/(length(B) - length(theta)));
        t_stat = (J_maxJ(J_opt_dv_id,1) - J_maxJ(i,1))/sqrt(V_J_mat(J_opt_dv_id,1) +
V_J_mat(i,1));
        if t_stat > tinv(1-0.05,dof)
            dv_fathom_row = dv_fathom_row + 1;
            dv_fathom_change = 1;
            dv_fathom(dv_fathom_row,1) = dv_update(1,i);
            dv_fathom_J = J_maxJ(i,1);
        end

    end

    % Exit loop if max number of discrete variables has been fathomed
    if isempty(elim_row) == 1
        break
    end

    elim_row = [ ];
    exp_cond_scale_update2 = exp_cond_scale_update;
    B_update2 = B_update;

    % New dv vector is dv_update
    dv_update_keep = [ ];
```

```
    for dv_id = 1:N_dv_update
        if ismember(dv_update(dv_id),dv_fathom) == 0
            dv_update_keep = [dv_update_keep dv_id];
        end
    end
    dv_update_prior = dv_update;
    dv_update = dv_update(dv_update_keep);
    N_dv_update = length(dv_update);

    % Remove fathomed rows from exp_cond
    for i = 1:length(exp_cond_scale_update2(:,1))

        if ismember(exp_cond_scale_update2(i,N_cv + 1),dv_fathom) == 1
            dv_id = find(dv == exp_cond_scale_update2(i,N_cv + 1));
            dv_update_id = find(dv_update_prior == exp_cond_scale_update2(i,N_cv + 1));
            if isempty(dv_update_id) == 0
                x_maxJ_final(dv_id,:) = x_maxJ(dv_update_id,1:N_cv);
                J_maxJ_final(dv_id,:) = J_maxJ(dv_update_id,:);
                x_maxY_final(dv_id,:) = x_maxY(dv_update_id,1:N_cv);
                Y_maxY_final(dv_id,:) = Y_maxY(dv_update_id,:);
            end

            % Propose new X matrix with variable eliminated
            exp_cond_scale_update = [ ];
            B_update = [ ];
            for j = 1:length(exp_cond_scale_update2(:,1))
                if ismember(j,elim_row) ~= 1 & j ~= i
                    exp_cond_scale_update = [exp_cond_scale_update; exp_cond_scale_update2(j,:)];
                    B_update = [B_update; B_update2(j,:)];
                end
            end
            [X_update,col_index_update,col_index_square_update] =
slug_X_matrix(cv_scale,N_cv,dv_update,N_dv_update,I_cv,I_dv,exp_cond_scale_update,'on');

            W1 = eye(length(B_update));
            for k = 1:length(B_update)
                W1(k,k) = exp(ln_yield(k))/sum(exp(ln_yield));
                %W1(k,k) = exp(B_update(k))/sum(exp(B_update));
            end

            if abs(det(X_update'*X_update)) > 1 && size(X_update,1) - 1 > size(X_update,2)
                % Remove row from X
                elim_row = [elim_row;i];
            end
        end
    end

    exp_cond_scale_update2(elim_row,:) = [ ];
```

```
    B_update2(elim_row,:) = [ ];
    exp_cond_scale_update = exp_cond_scale_update2;
    B_update = B_update2;
    X = slug_X_matrix(cv_scale,N_cv,dv_update,N_dv_update,I_cv,I_dv,exp_cond_scale_update,'on');
    exp_cond_update(elim_row,:) = [ ];

end

% New optimization finds experimental conditions which best test the
% current optimum for all unfathomed discrete variables
% Find current minimum
Aeq = zeros(N_dv_update,N_cv + N_dv_update);
for i = 1:N_dv_update
    Aeq(i,N_cv + i) = 1;
end

options = optimset('TolFun',1e-10,'TolX',1e-10,'TolCon',1e-10,'Display','off','Algorithm','SQP');

for i = 1:N_dv_update
    if isempty(find(dv_fathom == dv_update(1,i))) == 1
        % Constrain discrete variables
        Beq = zeros(N_dv_update,1);
        Beq(i,1) = 1;
        % Find x which maximizes change in J
        [x,min_dJ] = fmincon(@mindJ,x_maxJ(i,:)',[],[],Aeq,Beq,[-
1*ones(N_cv,1);zeros(N_dv_update,1)],ones(N_cv +
N_dv_update,1),[],options,x_maxJ(i,:)',x_maxY(i,:)',X,V_B,theta,V_yield,theta,col_index,col_index
_square);
        x_mindJ(i,:) = x';
        dJ_mindJ(i,1) = min_dJ;
    else
        x_mindJ(i,:) = zeros(1,N_cv + N_dv_update);
        dJ_mindJ(i,1) = 0;
    end
end

% Check for linearly improving exp(J_opt)
if length(J_opt(:,1)) >= 4
    J_opt_pred1 = (exp(J_opt(end-2,2)) - exp(J_opt(end-1,2)))/(exp(J_opt(end-2,1)) -
exp(J_opt(end-1,1)))*(exp(J_opt(end,1)) - exp(J_opt(end-2,1))) + exp(J_opt(end-2,2));
    J_opt_error1 = abs((J_opt_pred1 - exp(J_opt(end,2)))/exp(J_opt(end,2)));
    J_opt_pred2 = (exp(J_opt(end-3,2)) - exp(J_opt(end-2,2)))/(exp(J_opt(end-3,1)) -
exp(J_opt(end-2,1)))*(exp(J_opt(end-1,1)) - exp(J_opt(end-3,1))) + exp(J_opt(end-3,2));
    J_opt_error2 = abs((J_opt_pred2 - exp(J_opt(end-1,2)))/exp(J_opt(end-1,2)));
else
    J_opt_pred1 = [ ];
    J_opt_error1 = [ ];
    J_opt_pred2 = [ ];
```

```
    J_opt_error2 = [ ];
end

% Check for linearly improving exp(J_opt_lb)
if length(J_opt_lb(:,1)) >= 4
    J_opt_lb_pred1 = (exp(J_opt_lb(end-2,2)) - exp(J_opt_lb(end-1,2)))/(exp(J_opt_lb(end-2,1)) -
exp(J_opt_lb(end-1,1)))*(exp(J_opt_lb(end,1)) - exp(J_opt_lb(end-2,1))) + exp(J_opt_lb(end-2,2));
    J_opt_lb_error1 = abs((J_opt_lb_pred1 - exp(J_opt_lb(end,2)))/exp(J_opt_lb(end,2)));
    J_opt_lb_pred2 = (exp(J_opt_lb(end-3,2)) - exp(J_opt_lb(end-2,2)))/(exp(J_opt_lb(end-3,1)) -
exp(J_opt_lb(end-2,1)))*(exp(J_opt_lb(end-1,1)) - exp(J_opt_lb(end-3,1))) + exp(J_opt_lb(end-
3,2));
    J_opt_lb_error2 = abs((J_opt_lb_pred2 - exp(J_opt_lb(end-1,2)))/exp(J_opt_lb(end-1,2)));
else
    J_opt_lb_pred1 = [ ];
    J_opt_lb_error1 = [ ];
    J_opt_lb_pred2 = [ ];
    J_opt_lb_error2 = [ ];
end

% If uncertainty is improving linearly, terminate
if (J_opt_lb_error1 < 0.02 & J_opt_lb_error2 < 0.02 & J_opt_error1 < 0.02 & J_opt_error2 < 0.02)

    optim_phase = 4;

    for j = 1:length(dv_update)
        for i = 1:N_cv + 1
            if i <= N_cv
                exp_cond_scale2(2*j-1,i) = x_maxJ(j,i)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
                exp_cond_scale2(2*j,i) = x_maxY(j,i)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
            else
                exp_cond_scale2(2*j-1:2*j,i) = dv_update(1,j);
            end
        end
    end

else

    % Generate matrix of g-optimal conditions
    for i = 1:length(x_mindJ(1,:))
        if i <= N_cv
            exp_cond_scale2(:,i) = x_mindJ(:,i)*(cv_scale(i,2) - cv_scale(i,1))/2 +
mean([cv_scale(i,1) cv_scale(i,2)]);
        else
            for row = 1:length(x_mindJ(:,1))
                if abs(x_mindJ(row,i) - 1) < 1e-6
                    exp_cond_scale2(row,N_cv + 1) = dv_update(1,i - N_cv);
```

```
                end
            end
        end
    end

    % Don't update optim_phase
    optim_phase = 3;

end

for row = length(exp_cond_scale2(:,1)):-1:1
    if exp_cond_scale2(row,N_cv + 1) < 1
        exp_cond_scale2(row,:) = [ ];
    end
end

if isempty(exp_cond2) ~= 1
    exp_cond_scale2 = [exp_cond_scale2; exp_cond2(:,1:N_cv + 1)];
end

% Block randomize experiments
if optim_phase ~= 4
    new_stream = RandStream.create('mt19937ar','seed',sum(100*clock));
    RandStream.setGlobalStream(new_stream);
    exp_cond_scale2(:,end + 1) = rand(length(exp_cond_scale2(:,1)),1);
    exp_cond_scale2 = sortrows(exp_cond_scale2,length(exp_cond_scale2(1,:)));
    exp_cond_scale2(:,end) = [ ];
    % Group by temperature
    col_T = find(strcmp(deblank(index_cv),'Temp') == 1);
    for i = 1:length(exp_cond_scale2(:,1))-1
        swap_row = find(exp_cond_scale2(i+1:end,col_T) == exp_cond_scale2(i,col_T),1,'first');
        if isempty(swap_row) == 0
            exp_cond_scale_swap = exp_cond_scale2;
            exp_cond_scale2(i+1,:) = exp_cond_scale_swap(i+swap_row,:);
            exp_cond_scale2(i+swap_row,:) = exp_cond_scale_swap(i+1,:);
        end
    end
end

% Apply proper scaling
exp_cond1 = exp_cond_scale2;
for i = 1:N_cv
    if strcmp(deblank(index_cv(i)),'Res Time') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    elseif strcmp(deblank(index_cv(i)),'Temp') == 1
        exp_cond1(:,i) = exp_cond_scale2(:,i).^-1 - 273.15;
    elseif strcmp(deblank(index_cv(i)),'Base Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
```

```
    elseif strcmp(deblank(index_cv(i)),'Reag 1 Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    elseif strcmp(deblank(index_cv(i)),'Reag 2 Conc') == 1
        exp_cond1(:,i) = exp(exp_cond_scale2(:,i));
    end
end

% Include non-optimized continuous variables
for i = 1:length(non_opt)
    exp_cond1(:,N_cv + length(dv(:,1)) + i) = cv(non_opt(i),3);
end

% Include correlated discrete variables
for i = 1:length(cor(:,1))
    dv_row = cor(i,end);
    for row = 1:length(exp_cond1(:,1))
        j = find(dv(dv_row,:) == exp_cond1(row,N_cv + dv_row));
        exp_cond1(row,N_cv + length(dv(:,1)) + length(non_opt) + i) = cor(i,j);
    end
end

% Termination criterion
if optim_phase == 4
    final_opt = slug_build_prior_opt(N_cv,exp_cond1);
else
    final_opt = [ ];
end

for dv_update_id = 1:length(dv_update)
    dv_id = find(dv_update(dv_update_id) == dv);
    x_maxJ_final(dv_id,:) = x_maxJ(dv_update_id,1:N_cv);
    J_maxJ_final(dv_id,:) = J_maxJ(dv_update_id,:);
    x_maxY_final(dv_id,:) = x_maxY(dv_update_id,1:N_cv);
    Y_maxY_final(dv_id,:) = Y_maxY(dv_update_id,:);
end

x_maxJ = x_maxJ_final;
J_maxJ = J_maxJ_final;
x_maxY = x_maxY_final;
Y_maxY = Y_maxY_final;

end

function [Cineq,Ceq] = maxJ_nlcon(x,loading_index,A,c,gamma,Yi,cv_scale,sm_conc)
% Constraints for TON maximization

% Response surface prediction
Y = -maxY(x,loading_index,A,c,cv_scale,sm_conc);
```

```
% Inequality
Cineq = -Y + Yi + log(gamma);

% Equality
Ceq = [ ];

end

function [J,g] = maxJ(x,loading_index,A,c,gamma,Yi,cv_scale,sm_conc)
% Objective function for TON maximization

% Objective Function
J = -x'*A*x - c'*x;

% Gradient
if nargout > 1
    g = (-2*x'*A - c')';
end

end

function [Y,g] = maxY(x,loading_index,A,c,cv_scale,sm_conc)
% Objective function for yield maximization

% log catalyst loading
ln_loading = (x(loading_index) + 1)*(cv_scale(loading_index,2) - cv_scale(loading_index,1))/2 +
cv_scale(loading_index,1);

% Objective Function
Y = -x'*A*x - c'*x - ln_loading + log(sm_conc);

% Gradient
if nargout > 1
    g = (-2*x'*A - c')';
    g(loading_index) = g(loading_index) - (cv_scale(loading_index,2) -
cv_scale(loading_index,1))/2;
end

end

function [min_dJ] =
mindJ(x,x_opt,yield_opt,X,V_B,theta,V_yield,theta_yield,col_index,col_index_square)
% G-optimal experimental design criterion

% Augment X matrix and rewrite x_opt and yield_opt
x1 = zeros(1,length(theta));
x1_opt = zeros(1,length(theta));
```

```
yield1_opt = zeros(1,length(theta_yield));
for i = 1:length(theta)
    [row,col] = find(col_index == i);
    if isempty(row) == 0
        if row == col
            x1(1,i) = x(row,1);
            x1_opt(1,i) = x_opt(row,1);
            yield1_opt(1,i) = yield_opt(row,1);
        else
            x1(1,i) = x(row,1)*x(col,1);
            x1_opt(1,i) = x_opt(row,1)*x_opt(col,1);
            yield1_opt(1,i) = yield_opt(row,1)*yield_opt(col,1);
        end
    else
        col = find(col_index_square == i);
        x1(1,i) = x(col,1)^2;
        x1_opt(1,i) = x_opt(col,1)^2;
        yield1_opt(1,i) = yield_opt(col,1)^2;
    end
end
X1 = [X; x1];

% G-optimal objective
min_dJ = x1_opt*(X1'*V_B^-1*X1)^-1*x1_opt' + yield1_opt*(X1'*V_yield^-1*X1)^-1*yield1_opt';

end
```

C.1.4. Construction of **X** Matrix

```
function [X,col_index,col_index_square] =
slug_X_matrix(cv_scale,N_cv,dv,N_dv,I_cv,I_dv,exp_cond_scale,squared_terms)
%SLUG_X_MATRIX builds the matrix X from a matrix of experimental conditions.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Submitted by Brandon Reizman
% May 15, 2015
% Inputs:
%               cv_scale is the matrix of scaled continuous variables
%               N_cv is the number of continuous variables to optimize
%               dv is the matrix of discrete variables
%               N_dv is the number of discrete variables to optimize
%               I_cv is the matrix of interaction terms between continuous variables
%               I_dv is the matrix of interaction terms between discrete and continuous variables
%               exp_cond_scale is the matrix of scaled experimental conditions
%               squared_terms is 'on' if quadratic terms are included in the model
% Outputs:
%               X is the matrix of scaled experimental conditions
%               col_index is the matrix identifying the columns of X
```

```
%               col_index is the matrix identifying the squared term columns of X
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Build X matrix
X_length = 0;
dv_id = length(dv(:,1));
for i = 1:N_cv + N_dv
    if i <= N_cv
        for row = 1:length(exp_cond_scale(:,1))
            % (Value - Avg. Value)/(Max Value - Min Value)
            X(row,i) = 2*(exp_cond_scale(row,i) - mean([cv_scale(i,1)
cv_scale(i,2)]))/(cv_scale(i,2) - cv_scale(i,1));
        end
        X_length = X_length + 1;
        col_index(i,i) = X_length;
    else
        X_length = X_length + 1;
        for row = 1:length(exp_cond_scale(:,1))
            if exp_cond_scale(row,N_cv + dv_id) == dv(dv_id,i - N_cv);
                X(row,X_length) = 1;
            else
                X(row,X_length) = 0;
            end
        end
        col_index(i,i) = X_length;
    end
end

% Continuous variable interaction terms
for i = 1:N_cv
    for j = i + 1:N_cv
        if I_cv(i,j) == 1
           X_length = X_length + 1;
           col_index(i,j) = X_length;
           X(:,X_length) = X(:,i).*X(:,j);
        end
    end
end

% Discrete variable interaction terms
for i = 1:length(dv(:,1))
    for j = 1:N_cv
        if I_dv(i,j) == 1
            for k = N_cv + 1:N_cv + N_dv
                X_length = X_length + 1;
                col_index(j,k) = X_length;
                X(:,X_length) = X(:,j).*X(:,k);
```

```
            end
        end
    end
end

% Squared variable terms
if strcmp(squared_terms,'on') == 1
    for i = 1:N_cv
        if I_cv(i,i) == 1
            X_length = X_length + 1;
            col_index_square(i,1) = X_length;
            X(:,X_length) = X(:,i).^2;
        else
            col_index_square(i,1) = 0;
        end
    end
end

% This is to prevent linear dependence. Any variable with interaction in
% discrete variable will have to be removed from X
X_elim = [];
for j = 1:N_cv
    for i = 1:length(dv(:,1))
        if I_dv(i,j) == 1
            X_elim = [X_elim j];
            col_store = col_index(j,j);
            col_index(j,j) = 0;
            for k = 1:N_cv + N_dv
                for m = k:N_cv + N_dv
                    if col_index(k,m) > col_store
                        col_index(k,m) = col_index(k,m) - 1;
                    end
                end
            end
            if strcmp(squared_terms,'on') == 1
                for k = 1:N_cv
                    if col_index_square(k,1) > j
                        col_index_square(k,1) = col_index_square(k,1) - 1;
                    end
                end
            end
        end
    end
end
X(:,X_elim) = [];

% This prevents linear dependence in squared terms
if strcmp(squared_terms,'on') == 1
```

```
    for j = 1:N_cv
        if col_index_square(j) > 0
            if abs(X(:,col_index_square(j)) - ones(length(X(:,1)),1)) <= 0.1
                X(:,col_index_square(j)) = [];
                for k = j+1:N_cv
                    if col_index_square(k,1) > col_index_square(j,1)
                        col_index_square(k,1) = col_index_square(k,1) - 1;
                    end
                end
                col_index_square(j,1) = 0;
            end
        end
    end
end

% For calculating intercept
% Don't calculate intercept if there are discrete variables. The column for
% each discrete variable is effectively the intercept for that variable
if N_dv == 0
    X = [ones(size(X(:,1))) X];
    for i = 1:N_cv + N_dv
        for j = i:N_cv + N_dv
            if col_index(i,j) ~= 0
                col_index(i,j) = col_index(i,j) + 1;
            end
        end
        if strcmp(squared_terms,'on') ~= 1
            if col_index_square(i) ~= 0
                col_index_square(i) = col_index_square(i) + 1;
            end
        end
    end
end

if strcmp(squared_terms,'on') ~= 1
    col_index_square = [ ];
end

end
```

C.2. Simulation Data

Table C.1. Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.949	157.9
3	10.0	110.0	0.926	161.7
4	10.0	110.0	1.156	130.0
5	10.0	110.0	1.321	113.9
6	10.0	110.0	2.827	53.3
7	10.0	110.0	4.175	30.7
8	10.0	110.0	4.175	16.0
Test 2				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.848	175.1
3	10.0	110.0	0.974	154.2
4	10.0	110.0	1.219	123.8
5	10.0	110.0	1.560	96.8
6	10.0	110.0	2.788	54.3
7	10.0	110.0	4.175	22.7
8	10.0	110.0	4.175	15.9
Test 3				
1	10.0	110.0	0.835	180.9
2	10.0	110.0	1.035	145.7
3	10.0	110.0	0.987	152.3
4	10.0	110.0	1.543	98.8
5	10.0	110.0	1.619	94.3
6	10.0	110.0	3.623	42.2
7	9.8	110.0	4.175	23.0
8	8.8	110.0	4.175	8.8
Test 4				
1	10.0	110.0	0.835	181.7
2	10.0	110.0	0.851	175.5
3	10.0	110.0	0.876	170.6
4	10.0	110.0	1.059	141.4
5	10.0	110.0	1.164	128.9
6	10.0	110.0	4.175	30.7
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	9.2

Table C.1. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 5				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.919	162.8
3	10.0	110.0	0.873	170.6
4	10.0	110.0	1.373	110.0
5	10.0	110.0	1.222	123.1
6	10.0	110.0	3.995	37.9
7	10.0	110.0	4.175	31.6
8	10.0	110.0	4.175	16.1
Test 6				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	1.082	138.9
3	10.0	110.0	0.935	160.8
4	10.0	110.0	1.455	103.8
5	10.0	110.0	1.438	104.9
6	10.0	110.0	4.175	35.3
7	10.0	110.0	4.175	22.3
8	10.0	110.0	4.175	12.0
Test 7				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.893	167.0
3	10.0	110.0	0.871	171.1
4	10.0	110.0	1.214	123.9
5	10.0	110.0	1.139	131.7
6	10.0	110.0	4.175	34.3
7	10.0	110.0	4.175	28.2
8	10.0	110.0	4.175	13.1
Test 8				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.861	173.1
3	10.0	110.0	1.010	149.0
4	10.0	110.0	1.202	125.8
5	10.0	110.0	1.703	89.0
6	10.0	110.0	2.949	46.2
7	10.0	110.0	2.834	31.7
8	10.0	110.0	4.175	18.7

Table C.1. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 9				
1	10.0	110.0	0.835	181.2
2	10.0	110.0	0.920	162.2
3	10.0	110.0	0.907	164.8
4	10.0	110.0	1.153	129.9
5	10.0	110.0	1.208	124.2
6	10.0	110.0	4.175	34.4
7	10.0	110.0	4.175	22.8
8	9.5	110.0	4.175	10.2
Test 10				
1	10.0	110.0	0.835	180.6
2	10.0	110.0	0.987	152.2
3	10.0	110.0	1.023	147.1
4	10.0	110.0	1.276	118.6
5	10.0	110.0	1.687	89.6
6	10.0	110.0	3.031	50.0
7	10.0	110.0	4.175	30.9
8	10.0	110.0	4.175	15.8

Table C.2. Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	110.0	1.312	119.3
2	10.0	110.0	1.696	92.6
3	10.0	110.0	1.780	88.5
4	10.0	110.0	2.182	72.3
5	10.0	110.0	2.097	75.1
6	10.0	110.0	4.175	32.9
7	10.0	110.0	4.175	23.2
8	10.0	110.0	4.175	10.4
Test 2				
1	10.0	110.0	1.357	115.5
2	10.0	110.0	1.809	87.1
3	10.0	110.0	1.694	92.8
4	10.0	110.0	3.287	47.9
5	10.0	110.0	2.955	53.3
6	10.0	110.0	4.175	34.7
7	10.0	110.0	4.175	28.3
8	10.0	110.0	4.175	20.2

Table C.2. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 3				
1	10.0	110.0	1.308	119.5
2	10.0	110.0	2.155	73.6
3	10.0	110.0	1.980	79.9
4	10.0	110.0	2.568	61.8
5	10.0	110.0	3.104	51.3
6	10.0	110.0	4.175	36.8
7	10.0	110.0	4.175	27.7
8	10.0	110.0	3.931	20.0
Test 4				
1	10.0	110.0	1.305	119.8
2	10.0	110.0	1.971	80.2
3	10.0	110.0	2.188	72.5
4	10.0	110.0	4.175	37.7
5	10.0	110.0	2.667	59.6
6	10.0	110.0	4.175	34.7
7	10.0	110.0	3.890	22.6
8	10.0	110.0	3.434	18.8
Test 5				
1	10.0	110.0	1.310	119.4
2	10.0	110.0	1.561	100.5
3	10.0	110.0	1.686	93.1
4	10.0	110.0	2.217	71.0
5	10.0	110.0	2.157	72.9
6	10.0	110.0	4.175	34.7
7	10.0	110.0	4.175	31.8
8	10.0	110.0	4.175	10.7
Test 6				
1	10.0	110.0	1.631	97.1
2	10.0	110.0	2.687	59.3
3	10.0	110.0	2.185	72.5
4	10.0	110.0	2.550	62.2
5	10.0	110.0	3.922	40.8
6	10.0	110.0	4.175	30.7
7	10.0	110.0	4.175	28.0
8	10.0	110.0	4.175	9.2

Table C.2. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 7				
1	10.0	110.0	1.311	119.3
2	10.0	110.0	1.774	88.8
3	10.0	110.0	1.852	85.2
4	10.0	110.0	2.174	72.6
5	10.0	110.0	2.324	68.1
6	10.0	110.0	4.175	33.0
7	10.0	110.0	4.175	33.0
8	10.0	110.0	4.175	15.6
Test 8				
1	10.0	110.0	1.308	119.5
2	10.0	110.0	1.546	101.4
3	10.0	110.0	1.606	97.6
4	10.0	110.0	2.117	74.4
5	10.0	110.0	2.031	77.4
6	10.0	110.0	4.175	33.8
7	10.0	110.0	4.175	30.6
8	10.0	110.0	4.175	21.2
Test 9				
1	10.0	110.0	1.382	113.6
2	10.0	110.0	1.608	97.8
3	10.0	110.0	1.653	95.2
4	10.0	110.0	1.886	83.3
5	10.0	110.0	2.283	68.9
6	10.0	110.0	4.175	34.3
7	10.0	110.0	4.175	31.8
8	10.0	110.0	4.175	21.2
Test 10				
1	10.0	110.0	1.311	119.2
2	10.0	110.0	1.848	85.2
3	10.0	110.0	1.692	92.9
4	10.0	110.0	2.989	52.8
5	10.0	110.0	3.320	47.5
6	10.0	110.0	4.175	30.0
7	9.7	110.0	4.175	22.9
8	9.9	110.0	4.175	15.6

Table C.3. Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.98$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	110.0	2.285	70.6
2	10.0	110.0	3.283	49.4
3	10.0	110.0	3.247	49.9
4	10.0	110.0	4.175	38.3
5	10.0	110.0	4.175	38.7
6	10.0	110.0	4.175	36.6
7	10.0	110.0	4.175	31.6
8	10.0	110.0	4.175	21.0
Test 2				
1	10.0	110.0	2.255	71.7
2	10.0	110.0	2.733	59.2
3	10.0	110.0	2.630	61.4
4	10.0	110.0	4.175	38.5
5	10.0	110.0	3.622	44.7
6	10.0	110.0	4.175	36.0
7	10.0	110.0	4.175	29.4
8	10.0	110.0	4.175	21.2
Test 3				
1	10.0	110.0	2.286	70.6
2	10.0	110.0	2.782	58.1
3	10.0	110.0	2.741	58.9
4	10.0	110.0	4.175	38.6
5	10.0	110.0	4.175	38.6
6	10.0	110.0	4.175	34.9
7	10.0	110.0	4.017	33.5
8	10.0	110.0	3.861	18.5
Test 4				
1	10.0	110.0	2.285	70.6
2	10.0	110.0	2.592	62.3
3	10.0	110.0	2.750	58.8
4	10.0	110.0	4.175	38.5
5	10.0	110.0	4.175	38.6
6	10.0	110.0	4.175	36.5
7	10.0	110.0	3.986	33.6
8	10.0	110.0	3.740	22.8

Table C.3. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.98$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 5				
1	10.0	110.0	2.282	70.7
2	10.0	110.0	2.715	59.5
3	10.0	110.0	3.540	45.8
4	10.0	110.0	4.175	38.6
5	10.0	110.0	4.175	38.6
6	10.0	110.0	4.175	29.9
7	10.0	110.0	4.175	33.0
8	9.8	110.0	4.175	15.6
Test 6				
1	10.0	110.0	2.224	72.5
2	10.0	110.0	2.599	62.1
3	10.0	110.0	3.883	41.8
4	10.0	110.0	4.175	38.8
5	10.0	110.0	4.175	38.7
6	10.0	110.0	4.175	36.8
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	9.3
Test 7				
1	10.0	110.0	2.285	70.6
2	10.0	110.0	3.095	52.3
3	10.0	110.0	2.954	54.8
4	10.0	110.0	4.175	38.6
5	10.0	110.0	4.175	38.8
6	10.0	110.0	4.175	36.1
7	10.0	110.0	4.175	29.5
8	10.0	110.0	4.175	15.7
Test 8				
1	10.0	110.0	2.342	68.9
2	10.0	110.0	3.572	45.5
3	10.0	110.0	3.535	45.9
4	10.0	110.0	4.175	38.5
5	10.0	110.0	4.175	38.5
6	10.0	110.0	4.175	34.2
7	10.0	110.0	4.175	32.3
8	10.0	110.0	4.175	10.7

Table C.3. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 1 with $\gamma = 0.98$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 9				
1	10.0	110.0	1.871	85.7
2	10.0	110.0	2.906	55.6
3	10.0	110.0	2.828	57.1
4	10.0	110.0	4.175	38.8
5	10.0	110.0	4.175	38.2
6	10.0	110.0	4.175	36.7
7	10.0	110.0	4.175	27.6
8	10.0	110.0	3.821	20.3
Test 10				
1	10.0	110.0	2.507	64.8
2	10.0	110.0	3.419	47.4
3	10.0	110.0	3.380	47.9
4	10.0	110.0	4.175	38.5
5	10.0	110.0	4.175	38.8
6	10.0	110.0	4.175	33.6
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	16.2

Table C.4. Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	110.0	0.835	180.8
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.864	172.1
4	10.0	110.0	1.184	126.4
5	10.0	110.0	1.088	137.1
6	10.0	110.0	4.175	34.6
7	10.0	110.0	4.175	31.6
8	10.0	110.0	4.175	21.2
Test 2				
1	10.0	110.0	0.835	180.8
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.885	168.5
4	10.0	110.0	1.172	128.0
5	10.0	110.0	1.186	126.3
6	10.0	110.0	2.684	56.0
7	10.0	110.0	4.175	28.6
8	10.0	110.0	3.172	20.2

Table C.4. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 3				
1	10.0	110.0	0.835	180.9
2	10.0	110.0	0.835	180.8
3	10.0	110.0	0.838	177.0
4	10.0	110.0	0.926	160.5
5	10.0	110.0	0.953	156.2
6	9.4	110.0	4.175	29.5
7	8.0	110.0	4.175	22.1
8	7.7	110.0	4.175	15.8
Test 4				
1	10.0	110.0	0.835	181.2
2	10.0	110.0	0.835	180.9
3	10.0	110.0	0.904	165.4
4	10.0	110.0	1.506	100.0
5	10.0	110.0	1.400	107.6
6	10.0	110.0	4.175	33.6
7	10.0	110.0	4.175	32.4
8	10.0	110.0	4.175	16.8
Test 5				
1	10.0	110.0	0.835	180.8
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.903	165.3
4	10.0	110.0	1.230	121.8
5	10.0	110.0	1.255	119.7
6	10.0	110.0	4.175	35.3
7	10.0	110.0	4.175	28.7
8	10.0	110.0	4.175	21.2
Test 6				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.846	175.6
4	10.0	110.0	1.441	103.8
5	10.0	110.0	1.000	148.9
6	10.0	110.0	4.175	29.9
7	10.0	110.0	3.450	36.9
8	10.0	110.0	2.886	22.6

Table C.4. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.90$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 7				
1	10.0	110.0	0.835	181.0
2	10.0	110.0	0.835	180.9
3	10.0	110.0	0.874	170.6
4	10.0	110.0	1.227	122.1
5	10.0	110.0	1.275	117.9
6	10.0	110.0	2.653	56.7
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	8.9
Test 8				
1	10.0	110.0	0.835	180.7
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.910	164.5
4	10.0	110.0	1.363	111.0
5	10.0	110.0	1.545	98.3
6	10.0	110.0	4.175	34.9
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	19.6
Test 9				
1	10.0	110.0	0.835	180.9
2	10.0	110.0	0.835	180.7
3	10.0	110.0	0.890	167.8
4	10.0	110.0	1.108	135.3
5	10.0	110.0	1.343	112.2
6	10.0	110.0	3.692	38.5
7	10.0	110.0	3.936	25.7
8	10.0	110.0	4.175	18.3
Test 10				
1	10.0	110.0	0.835	180.6
2	10.0	110.0	0.835	180.8
3	10.0	110.0	0.835	177.6
4	10.0	110.0	1.242	120.9
5	10.0	110.0	1.130	132.7
6	10.0	110.0	4.175	35.9
7	10.0	110.0	4.175	30.9
8	10.0	110.0	4.175	21.2

Table C.5. Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	110.0	1.577	100.2
2	10.0	110.0	1.409	111.9
3	10.0	110.0	2.010	78.9
4	10.0	110.0	2.606	60.7
5	10.0	110.0	2.917	54.3
6	10.0	110.0	4.175	35.5
7	10.0	110.0	4.175	23.6
8	10.0	110.0	4.175	10.5
Test 2				
1	10.0	110.0	1.338	117.0
2	10.0	110.0	1.306	119.9
3	10.0	110.0	1.748	89.5
4	10.0	110.0	2.334	67.2
5	10.0	110.0	4.175	36.6
6	10.0	110.0	4.175	36.7
7	10.0	110.0	4.175	27.4
8	10.0	110.0	4.175	8.8
Test 3				
1	10.0	110.0	1.300	120.2
2	10.0	110.0	1.313	119.0
3	10.0	110.0	1.588	98.6
4	10.0	110.0	2.691	58.4
5	10.0	110.0	2.079	75.4
6	10.0	110.0	4.175	36.8
7	9.1	110.0	4.175	22.2
8	7.9	110.0	4.175	9.9
Test 4				
1	10.0	110.0	1.343	116.7
2	10.0	110.0	1.360	115.2
3	10.0	110.0	1.885	83.2
4	10.0	110.0	1.936	81.1
5	10.0	110.0	3.024	52.0
6	10.0	110.0	4.175	34.2
7	10.0	110.0	4.175	30.6
8	10.0	110.0	4.175	15.5

Table C.5. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 5				
1	10.0	110.0	1.325	118.1
2	10.0	110.0	1.306	119.8
3	10.0	110.0	1.755	89.4
4	10.0	110.0	2.663	58.9
5	10.0	110.0	2.826	55.7
6	10.0	110.0	4.175	36.8
7	9.4	110.0	4.175	22.5
8	9.7	110.0	4.175	21.2
Test 6				
1	10.0	110.0	1.286	121.6
2	10.0	110.0	1.307	119.7
3	10.0	110.0	1.696	92.5
4	10.0	110.0	3.187	49.5
5	10.0	110.0	2.265	69.5
6	10.0	110.0	4.175	33.6
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	21.2
Test 7				
1	10.0	110.0	1.359	115.3
2	10.0	110.0	1.388	113.0
3	10.0	110.0	1.622	96.8
4	10.0	110.0	2.378	66.2
5	10.0	110.0	2.280	69.0
6	10.0	110.0	4.175	34.2
7	10.0	110.0	4.175	28.5
8	10.0	110.0	4.175	21.2
Test 8				
1	10.0	110.0	1.305	119.9
2	10.0	110.0	1.322	118.3
3	10.0	110.0	1.576	99.5
4	10.0	110.0	1.796	87.4
5	10.0	110.0	1.872	83.9
6	10.0	110.0	4.175	34.6
7	10.0	110.0	4.175	24.8
8	9.7	110.0	4.175	10.1

Table C.5. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 2 with $\gamma = 0.95$.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 9				
1	10.0	110.0	1.334	117.4
2	10.0	110.0	1.344	116.3
3	10.0	110.0	1.854	84.8
4	10.0	110.0	3.278	48.7
5	10.0	110.0	2.174	72.6
6	10.0	110.0	4.175	36.2
7	10.0	110.0	4.175	31.3
8	10.0	110.0	4.175	20.6
Test 10				
1	10.0	110.0	1.326	118.0
2	10.0	110.0	1.320	118.6
3	10.0	110.0	1.681	93.5
4	10.0	110.0	2.171	72.5
5	10.0	110.0	2.776	56.9
6	10.0	110.0	4.175	35.5
7	9.3	110.0	4.175	22.2
8	9.5	110.0	4.175	16.0

Table C.6. Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B → S_1.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	81.7	2.807	32.8
2	10.0	81.7	3.542	26.2
3	10.0	82.3	3.970	23.4
4	10.0	82.1	4.175	20.7
5	10.0	83.3	4.175	20.6
6	10.0	82.6	4.175	13.8
7	10.0	83.8	4.175	9.4
8	10.0	90.3	4.175	2.9
Test 2				
1	10.0	80.7	2.595	34.9
2	10.0	81.8	4.175	20.2
3	10.0	80.8	3.865	23.5
4	10.0	81.1	4.175	19.9
5	10.0	84.2	4.175	18.4
6	10.0	86.7	4.175	12.2
7	10.0	86.3	4.175	7.6
8	10.0	93.5	4.175	2.8

Table C.6. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B → S_1.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 3				
1	10.0	79.8	2.592	35.2
2	10.0	80.9	3.903	23.5
3	10.0	81.1	3.291	27.8
4	10.0	81.4	4.175	20.2
5	10.0	82.0	4.175	19.6
6	10.0	85.0	4.175	13.1
7	10.0	85.2	4.175	7.7
8	10.0	90.0	4.175	3.1
Test 4				
1	10.0	81.0	2.741	33.6
2	10.0	80.7	3.780	24.5
3	10.0	79.6	3.308	27.9
4	10.0	81.6	4.175	20.1
5	10.0	82.6	4.175	21.3
6	10.0	85.5	4.175	14.5
7	10.0	87.4	4.175	9.3
8	10.0	90.2	4.175	3.0
Test 5				
1	10.0	81.9	2.903	32.3
2	10.0	81.5	3.494	26.5
3	10.0	83.6	3.497	26.5
4	10.0	80.4	4.175	21.5
5	10.0	82.9	4.175	19.8
6	10.0	86.7	4.175	14.9
7	10.0	88.1	4.175	8.1
8	10.0	87.4	4.175	3.3
Test 6				
1	10.0	82.4	2.860	32.6
2	10.0	81.3	3.506	26.3
3	10.0	81.4	3.742	24.8
4	10.0	80.4	4.175	19.8
5	10.0	80.8	4.175	20.3
6	10.0	83.0	4.175	13.0
7	10.0	87.1	4.175	9.2
8	10.0	89.6	4.175	3.0

Table C.6. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B → S_1.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 7				
1	10.0	81.5	2.805	32.9
2	10.0	82.0	4.175	22.1
3	10.0	82.1	3.695	25.1
4	10.0	82.0	4.175	20.1
5	10.0	81.9	4.175	20.1
6	10.0	84.0	4.175	15.2
7	10.0	86.3	4.175	9.5
8	10.0	87.4	4.175	3.2
Test 8				
1	10.0	81.4	2.806	32.8
2	10.0	81.4	3.538	25.8
3	10.0	82.1	3.061	29.6
4	10.0	80.8	4.175	20.9
5	10.0	81.9	4.175	19.9
6	10.0	85.3	4.175	13.1
7	10.0	87.9	4.175	9.4
8	10.0	91.8	4.175	3.2
Test 9				
1	10.0	80.4	2.754	33.3
2	10.0	81.5	3.486	26.5
3	10.0	80.7	4.117	22.5
4	10.0	83.7	4.175	21.6
5	10.0	82.9	4.175	21.4
6	10.0	83.4	4.175	13.0
7	10.0	85.5	4.175	9.3
8	10.0	90.3	4.175	3.0
Test 10				
1	10.0	80.7	2.696	34.3
2	10.0	81.5	4.175	22.0
3	10.0	81.1	4.175	21.8
4	10.0	82.2	4.175	19.5
5	10.0	81.3	4.175	19.3
6	10.0	86.1	4.175	12.5
7	10.0	83.9	4.175	8.9
8	10.0	88.3	4.175	2.7

Table C.7. Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B + R → S_2.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	2.3	109.7	2.582	24.3
2	2.3	110.0	3.159	19.8
3	2.3	110.0	3.041	20.7
4	2.3	110.0	3.940	15.9
5	2.3	110.0	3.608	17.4
6	2.3	110.0	4.175	11.5
7	2.3	110.0	4.175	8.5
8	2.3	110.0	4.175	3.8
Test 2				
1	2.5	110.0	2.701	23.3
2	2.4	110.0	3.461	18.2
3	2.5	110.0	3.426	18.4
4	2.3	110.0	4.175	15.0
5	2.3	110.0	4.084	15.5
6	2.3	110.0	4.175	11.6
7	2.3	110.0	4.175	8.5
8	2.3	110.0	4.175	3.7
Test 3				
1	2.3	110.0	2.778	22.8
2	2.3	110.0	3.170	20.0
3	2.3	110.0	3.496	18.1
4	2.2	110.0	4.086	15.5
5	2.2	110.0	4.175	14.9
6	2.4	110.0	4.175	11.4
7	2.4	110.0	4.175	8.6
8	2.4	110.0	4.175	3.7
Test 4				
1	2.6	110.0	2.534	24.7
2	2.6	110.0	3.107	20.3
3	2.6	110.0	3.213	19.7
4	2.5	110.0	4.175	14.2
5	2.5	110.0	4.175	14.9
6	2.5	110.0	4.175	11.2
7	2.6	110.0	4.175	8.3
8	2.6	110.0	4.175	4.0

Table C.7. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B + R → S_2.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 5				
1	2.4	110.0	2.569	24.5
2	2.4	110.0	3.085	20.4
3	2.4	110.0	3.272	19.3
4	2.3	110.0	4.175	14.5
5	2.3	110.0	4.175	15.1
6	2.3	110.0	4.175	11.2
7	2.4	110.0	4.175	8.2
8	2.4	110.0	4.175	3.1
Test 6				
1	2.2	110.0	4.175	16.9
2	2.6	110.0	3.624	17.4
3	2.5	110.0	3.743	16.9
4	2.2	110.0	4.175	14.6
5	2.3	110.0	4.175	14.7
6	2.5	110.0	4.175	11.4
7	2.5	110.0	4.175	8.2
8	2.5	110.0	4.175	3.3
Test 7				
1	2.3	110.0	2.504	25.1
2	2.3	110.0	3.110	20.2
3	2.3	110.0	3.490	18.0
4	2.2	110.0	4.175	14.9
5	2.2	110.0	4.175	14.8
6	2.3	110.0	4.175	11.6
7	2.4	110.0	4.175	8.2
8	2.4	110.0	4.175	4.2
Test 8				
1	2.4	110.0	2.580	24.4
2	2.4	110.0	3.434	18.4
3	2.4	110.0	3.218	19.5
4	2.4	110.0	4.175	14.5
5	2.4	110.0	4.175	14.9
6	2.5	110.0	4.175	11.4
7	2.5	110.0	4.175	8.2
8	2.6	110.0	4.175	4.0

Table C.7. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 3 with competing reaction B + R → S_2.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 9				
1	10.0	84.8	2.886	21.0
2	10.0	85.1	3.311	18.4
3	10.0	85.3	3.221	18.9
4	10.0	85.4	4.175	14.4
5	1.8	110.0	4.175	13.8
6	2.0	110.0	4.175	11.6
7	2.4	110.0	4.175	8.2
8	2.4	110.0	4.175	3.6
Test 10				
1	2.3	110.0	2.579	24.5
2	2.4	110.0	3.164	19.9
3	2.4	110.0	3.071	20.5
4	2.3	110.0	4.175	14.7
5	2.3	110.0	4.175	14.3
6	2.3	110.0	4.175	10.3
7	2.4	110.0	4.175	8.1
8	2.5	110.0	4.175	3.5

Table C.8. Predicted optimal conditions and TON for kinetics of Case Study 4 with no prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	78.6	1.085	142.5
2	10.0	110.0	1.514	102.1
3	10.0	110.0	1.690	91.0
4	10.0	110.0	1.224	126.3
5	10.0	110.0	1.789	86.0
6	10.0	110.0	4.175	35.4
7	10.0	110.0	4.175	30.3
8	10.0	110.0	4.175	9.3
Test 2				
1	10.0	71.8	2.154	67.9
2	10.0	110.0	0.939	155.7
3	10.0	110.0	0.938	155.9
4	10.0	110.0	0.938	155.8
5	10.0	110.0	0.939	155.6
6	10.0	110.0	4.175	34.5
7	10.0	110.0	4.175	32.5
8	10.0	110.0	4.175	20.2

Table C.8. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with no prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 3				
1	10.0	76.3	1.969	79.6
2	10.0	110.0	0.936	156.2
3	10.0	110.0	0.940	155.5
4	10.0	110.0	0.941	155.4
5	10.0	110.0	0.940	155.6
6	10.0	110.0	4.175	35.4
7	10.0	110.0	4.175	32.3
8	10.0	110.0	4.175	16.1
Test 4				
1	10.0	78.5	1.858	82.4
2	10.0	110.0	1.019	144.4
3	10.0	110.0	0.941	155.4
4	10.0	110.0	1.077	137.4
5	10.0	110.0	1.030	143.1
6	10.0	110.0	4.175	36.5
7	10.0	110.0	4.175	30.6
8	10.0	110.0	4.175	17.5
Test 5				
1	10.0	76.7	2.937	52.2
2	10.0	110.0	0.944	155.0
3	10.0	110.0	0.944	155.0
4	10.0	110.0	0.944	155.0
5	10.0	110.0	0.939	155.8
6	10.0	110.0	3.501	43.6
7	10.0	110.0	4.175	28.9
8	10.0	110.0	4.175	21.2
Test 6				
1	10.0	81.1	1.485	104.3
2	10.0	110.0	1.803	87.1
3	10.0	110.0	1.781	88.1
4	10.0	110.0	2.145	73.3
5	10.0	110.0	1.604	96.7
6	10.0	110.0	3.352	44.1
7	10.0	110.0	3.388	38.5
8	10.0	110.0	3.749	22.5

Table C.8. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with no prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 7				
1	10.0	78.0	3.097	48.6
2	10.0	110.0	0.948	154.4
3	10.0	110.0	0.951	154.0
4	10.0	110.0	0.944	155.1
5	10.0	110.0	0.951	154.0
6	10.0	110.0	4.175	35.9
7	8.6	110.0	4.175	23.1
8	8.3	110.0	4.175	9.1
Test 8				
1	10.0	72.1	2.226	67.4
2	10.0	110.0	0.932	156.8
3	10.0	110.0	0.934	156.5
4	10.0	110.0	0.934	156.5
5	10.0	110.0	0.932	156.8
6	10.0	110.0	4.011	37.6
7	10.0	110.0	4.175	32.3
8	10.0	110.0	4.175	16.7
Test 9				
1	10.0	76.1	2.526	58.3
2	10.0	110.0	0.937	156.0
3	10.0	110.0	0.937	155.9
4	10.0	110.0	0.936	156.2
5	10.0	110.0	0.936	156.2
6	10.0	110.0	1.693	86.4
7	10.0	110.0	4.175	33.1
8	10.0	110.0	4.175	19.4
Test 10				
1	10.0	77.0	2.943	51.2
2	10.0	110.0	0.939	155.7
3	10.0	110.0	0.942	155.3
4	10.0	110.0	0.940	155.6
5	10.0	110.0	0.940	155.6
6	10.0	110.0	4.175	35.5
7	10.0	110.0	4.175	33.2
8	10.0	110.0	4.175	10.6

Table C.9. Predicted optimal conditions and TON for kinetics of Case Study 4 with 10% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	79.9	0.835	182.1
2	10.0	110.0	2.746	57.3
3	10.0	110.0	2.901	54.3
4	10.0	110.0	2.940	53.9
5	10.0	110.0	1.906	81.9
6	10.0	110.0	4.175	36.4
7	10.0	110.0	4.175	31.4
8	10.0	110.0	4.175	18.1
Test 2				
1	10.0	73.2	1.617	92.8
2	10.0	110.0	0.949	154.4
3	10.0	110.0	0.921	158.9
4	10.0	110.0	0.945	155.0
5	10.0	110.0	0.954	153.5
6	10.0	110.0	2.145	68.4
7	10.0	110.0	4.175	33.2
8	10.0	110.0	4.175	17.1
Test 3				
1	10.0	75.2	1.803	85.9
2	10.0	110.0	0.949	154.3
3	10.0	110.0	0.937	156.3
4	10.0	110.0	0.954	153.5
5	10.0	110.0	0.954	153.5
6	10.0	110.0	4.175	36.1
7	10.0	110.0	4.175	31.1
8	10.0	110.0	4.175	19.6
Test 4				
1	10.0	79.4	1.351	114.2
2	10.0	110.0	1.841	84.4
3	10.0	110.0	1.907	81.0
4	9.6	110.0	1.741	89.2
5	10.0	110.0	1.958	79.3
6	10.0	110.0	4.175	36.4
7	10.0	110.0	4.175	31.7
8	10.0	110.0	4.175	16.8

Table C.9. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with 10% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 5				
1	10.0	80.4	1.278	119.5
2	10.0	110.0	1.585	96.8
3	10.0	110.0	1.875	83.1
4	10.0	110.0	1.729	89.6
5	9.7	110.0	1.485	103.0
6	10.0	110.0	4.175	36.3
7	10.0	110.0	4.175	30.8
8	10.0	110.0	4.175	17.5
Test 6				
1	10.0	73.1	1.622	92.4
2	10.0	110.0	0.939	155.7
3	10.0	110.0	0.937	156.1
4	10.0	110.0	0.939	155.6
5	10.0	110.0	0.939	155.7
6	10.0	110.0	3.635	41.3
7	10.0	110.0	4.175	31.2
8	10.0	110.0	4.175	20.5
Test 7				
1	10.0	77.6	1.269	122.1
2	10.0	110.0	1.468	104.3
3	10.0	110.0	1.683	91.1
4	10.0	110.0	1.500	102.5
5	10.0	110.0	1.510	102.0
6	10.0	110.0	4.175	35.9
7	10.0	110.0	4.175	29.9
8	10.0	110.0	4.175	18.1
Test 8				
1	10.0	78.0	1.274	121.1
2	10.0	110.0	2.024	77.2
3	10.0	110.0	1.624	95.3
4	10.0	110.0	1.559	99.4
5	10.0	110.0	1.755	88.4
6	10.0	110.0	4.175	34.5
7	10.0	110.0	4.175	30.8
8	10.0	110.0	4.175	21.2

Table C.9. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with 10% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 9				
1	10.0	77.7	1.362	114.3
2	10.0	110.0	1.727	89.3
3	10.0	110.0	1.606	96.1
4	10.0	110.0	1.677	91.7
5	10.0	110.0	1.740	88.7
6	10.0	110.0	4.175	36.8
7	10.0	110.0	4.175	32.2
8	10.0	110.0	4.175	17.8
Test 10				
1	10.0	77.6	1.262	122.5
2	10.0	110.0	2.086	75.3
3	10.0	110.0	1.969	79.8
4	10.0	110.0	1.804	86.6
5	10.0	110.0	1.607	96.5
6	10.0	110.0	4.175	36.3
7	10.0	110.0	4.175	30.8
8	10.0	110.0	4.175	18.7

Table C.10. Predicted optimal conditions and TON for kinetics of Case Study 4 with 2.5% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 1				
1	10.0	79.1	1.261	121.9
2	10.0	110.0	1.657	93.2
3	10.0	110.0	1.731	89.4
4	10.0	110.0	1.496	102.8
5	10.0	110.0	1.417	108.5
6	10.0	110.0	4.175	35.3
7	10.0	110.0	4.175	31.8
8	10.0	110.0	4.175	19.5
Test 2				
1	10.0	79.3	0.835	181.0
2	10.0	110.0	1.437	106.8
3	10.0	110.0	2.104	74.1
4	10.0	110.0	1.985	78.4
5	10.0	110.0	2.029	76.7
6	10.0	110.0	4.175	36.3
7	10.0	110.0	4.175	32.7
8	10.0	110.0	4.175	17.2

Table C.10. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with 2.5% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 3				
1	10.0	78.3	1.110	141.2
2	10.0	110.0	2.099	74.3
3	10.0	110.0	2.055	75.9
4	10.0	110.0	1.899	81.2
5	10.0	110.0	1.845	83.6
6	10.0	110.0	4.175	36.4
7	10.0	110.0	4.175	32.6
8	10.0	110.0	4.175	19.5
Test 4				
1	10.0	78.8	0.835	182.6
2	10.0	110.0	2.163	72.4
3	10.0	110.0	2.114	73.9
4	10.0	110.0	1.912	81.4
5	10.0	110.0	1.334	114.1
6	10.0	110.0	4.175	36.3
7	10.0	110.0	4.175	32.5
8	10.0	110.0	4.175	18.3
Test 5				
1	10.0	74.9	1.638	93.0
2	10.0	110.0	0.941	155.5
3	10.0	110.0	0.939	155.8
4	10.0	110.0	0.932	157.0
5	10.0	110.0	0.940	155.5
6	10.0	110.0	4.175	36.5
7	10.0	110.0	4.175	33.3
8	10.0	110.0	4.175	19.2
Test 6				
1	10.0	80.9	1.385	112.2
2	10.0	110.0	1.565	98.4
3	10.0	110.0	1.744	88.9
4	10.0	110.0	1.796	86.4
5	10.0	110.0	1.466	104.9
6	10.0	110.0	4.058	36.6
7	10.0	110.0	4.175	32.2
8	10.0	110.0	4.175	19.4

Table C.10. (cont.) Predicted optimal conditions and TON for kinetics of Case Study 4 with 2.5% prediction covariance trust region.

Catalyst	t_{res} (min)	T (°C)	C_{cat} (mM)	TON
Test 7				
1	10.0	77.8	0.835	183.0
2	10.0	110.0	2.016	76.8
3	10.0	110.0	1.559	98.3
4	10.0	110.0	1.648	93.1
5	10.0	110.0	1.554	98.5
6	10.0	110.0	4.175	36.4
7	10.0	110.0	4.175	30.8
8	10.0	110.0	4.175	17.1
Test 8				
1	10.0	79.5	1.412	109.3
2	10.0	110.0	2.087	74.8
3	10.0	110.0	2.384	66.0
4	10.0	110.0	2.553	61.7
5	10.0	110.0	2.426	64.8
6	10.0	110.0	4.175	35.6
7	10.0	110.0	4.175	30.8
8	10.0	110.0	4.175	16.8
Test 9				
1	10.0	77.6	0.835	181.7
2	10.0	110.0	1.395	109.0
3	10.0	110.0	1.569	97.4
4	10.0	110.0	1.587	96.6
5	10.0	110.0	1.335	113.7
6	10.0	110.0	4.175	36.6
7	10.0	110.0	4.175	31.4
8	10.0	110.0	4.175	18.3
Test 10				
1	10.0	73.0	1.729	87.9
2	10.0	110.0	0.932	156.8
3	10.0	110.0	0.931	156.9
4	10.0	110.0	0.932	156.8
5	10.0	110.0	0.931	157.0
6	10.0	110.0	4.175	36.3
7	10.0	110.0	4.175	32.1
8	10.0	110.0	4.175	18.3

APPENDIX D. CHAPTER 5 SUPPORTING INFORMATION

D.1. EXPERIMENTAL DATA

Table D.1. Observed yields for conditions screened during first fractional factorial design.

Experiment	Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
1	*i*PrOH	60	30.0	0.999	2.4%
2	THF	60	30.0	0.999	1.8%
3	Toluene	600	30.0	0.998	9.7%
4	MeCN	60	30.0	0.206	2.0%
5	DMF	600	30.0	0.206	6.1%
6	DMSO	600	30.0	0.999	52.2%
7	Pyridine	60	30.0	0.206	8.5%
8	DMC	60	30.0	0.206	1.0%
9	DME	600	30.0	1.000	5.4%
10	DCE	600	30.0	0.206	2.7%
11	DMSO	60	120.0	0.206	28.4%
12	*i*PrOH	600	120.0	0.206	23.5%
13	MeCN	600	120.0	1.000	46.6%
14	Pyridine	600	120.0	0.999	47.2%
15	DMC	600	120.0	0.999	56.8%
16	THF	600	120.0	0.206	27.7%
17	Toluene	60	120.0	0.206	9.3%
18	DMF	60	120.0	0.999	36.4%
19	DME	60	120.0	0.206	8.7%
20	DCE	60	120.0	1.000	44.6%

Table D.2. Observed yields for conditions screened during second fractional factorial design.

Experiment	Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
21	DMC	600	120.0	0.999	45.5%
22	DCE	190	120.0	1.000	53.4%
23	*i*PrOH	600	120.0	0.444	43.2%
24	MeCN	600	120.0	1.000	36.1%
25	THF	600	120.0	0.444	35.7%
26	DME	190	120.0	0.445	42.3%
27	Pyridine	190	120.0	0.999	51.5%
28	DMSO	190	120.0	0.445	52.2%
29	Toluene	190	120.0	0.444	43.7%
30	DMF	190	120.0	0.444	48.9%
31	DMF	600	69.2	0.999	56.2%
32	DCE	600	69.2	0.444	35.9%
33	THF	190	69.2	0.999	29.9%
34	MeCN	190	69.2	0.444	34.2%
35	Pyridine	600	69.2	0.444	30.5%
36	DME	600	69.2	1.000	41.5%
37	DMC	190	69.2	0.444	16.2%
38	Toluene	600	69.2	0.998	41.9%
39	*i*PrOH	190	69.2	0.999	33.6%
40	DMSO	600	69.2	0.999	62.9%

Table D.3. Observed yields for conditions screened during response surface optimization with G-optimal design of experiments criterion.

Experiment	Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
41	DMSO	77	46.4	0.999	36.0%
42	DCE	137	63.6	1.000	24.0%
43	THF	161	120.0	0.999	43.2%
44	Pyridine	94	61.1	0.999	33.0%
45	*i*PrOH	161	120.0	0.999	48.7%
46	DMF	100	52.1	0.999	31.4%
47	Toluene	91	59.4	0.998	9.9%
48	DMSO	265	90.6	0.999	57.2%
49	DMF	429	120.0	0.999	48.7%
50	DMSO	391	102.8	0.999	59.3%
51	DMSO	323	58.8	0.333	31.4%
52	DCE	326	120.0	0.429	34.8%
53	DMF	229	98.9	0.364	46.1%
54	DMF	600	101.0	0.999	58.6%
55	DCE	419	120.0	1.000	47.7%
56	Pyridine	600	45.0	0.999	44.0%
57	DMSO	600	87.8	0.999	60.7%
58	DMF	431	94.3	0.999	52.3%
59	DMSO	600	85.2	0.999	64.3%
60	Pyridine	455	73.2	0.999	53.1%
61	DMSO	319	74.9	0.999	53.9%
62	Pyridine	340	77.7	0.999	53.5%
63	DMF	332	93.1	0.999	54.4%
64	DMF	340	96.0	0.778	54.9%
65	*i*PrOH	188	120.0	0.667	23.7%
66	DMSO	514	75.6	0.682	59.4%
67	DMSO	600	79.8	0.999	62.9%
68	DMSO	516	81.1	0.999	62.6%

Table D.4. Observed yields for conditions screened during quasi-Newton gradient-based search.

Experiment	Solvent	t_{res} (s)	T (°C)	$C_{0,MeOBnCl}$ (M)	Yield
69	DMSO	444	78.1	0.999	61.1%
70	DMSO	459	76.1	0.983	60.7%
71	DMSO	429	76.1	0.999	62.2%
72	DMSO	429	80.1	0.983	61.5%
73	DMSO	459	80.1	0.999	60.4%
74	DMSO	189	76.2	0.999	62.8%
75	DMSO	204	74.2	0.983	56.7%
76	DMSO	174	74.2	0.999	60.6%
77	DMSO	204	78.2	0.999	59.2%
78	DMSO	189	76.2	0.999	54.8%
79	DMSO	174	78.2	0.983	59.8%
80	DMSO	290	77.2	0.999	64.1%
81	DMSO	305	75.2	0.983	65.8%
82	DMSO	275	75.2	0.999	57.1%
83	DMSO	305	79.2	0.999	62.4%
84	DMSO	275	79.2	0.983	59.3%
85	DMSO	290	77.2	0.999	58.3%
86	DMSO	359	77.6	0.999	63.2%
87	DMSO	374	75.6	0.983	61.3%
88	DMSO	344	75.6	0.999	62.5%
89	DMSO	359	77.6	0.999	60.9%
90	DMSO	344	79.6	0.983	60.3%
91	DMSO	374	79.6	0.999	63.8%
92	DMSO	399	77.8	0.999	60.3%
93	DMSO	421	78.0	0.999	61.5%

D.2. NMR SPECTRA

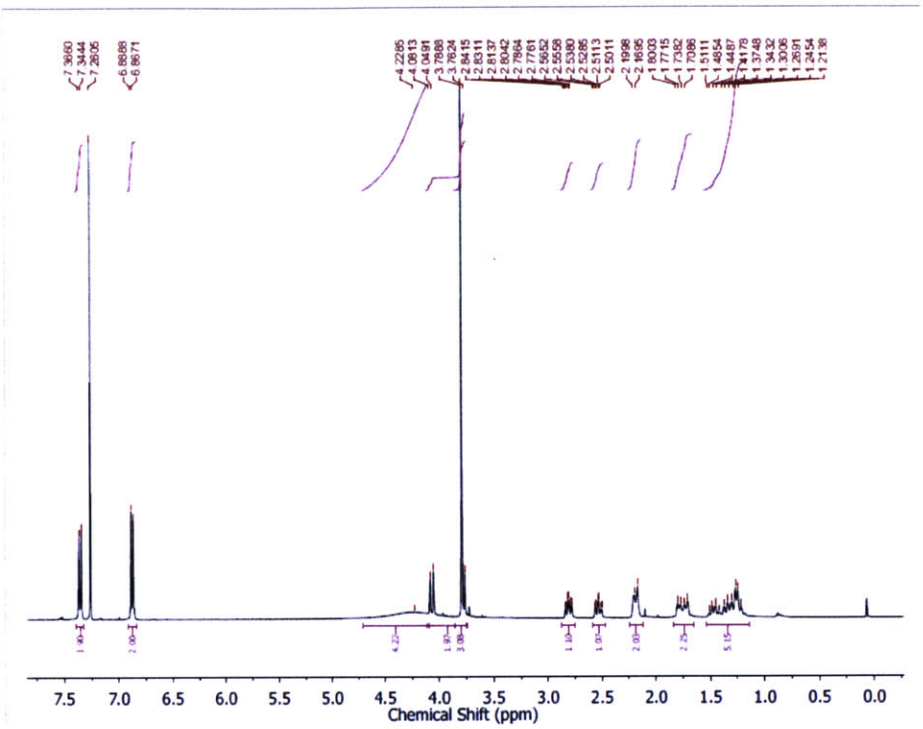


Figure D.1. (N-4-methoxybenzyl)-(1*R*,2*R*)-(-)-diaminocyclohexane ^{1}H NMR (400 MHz, $CDCl_3$)

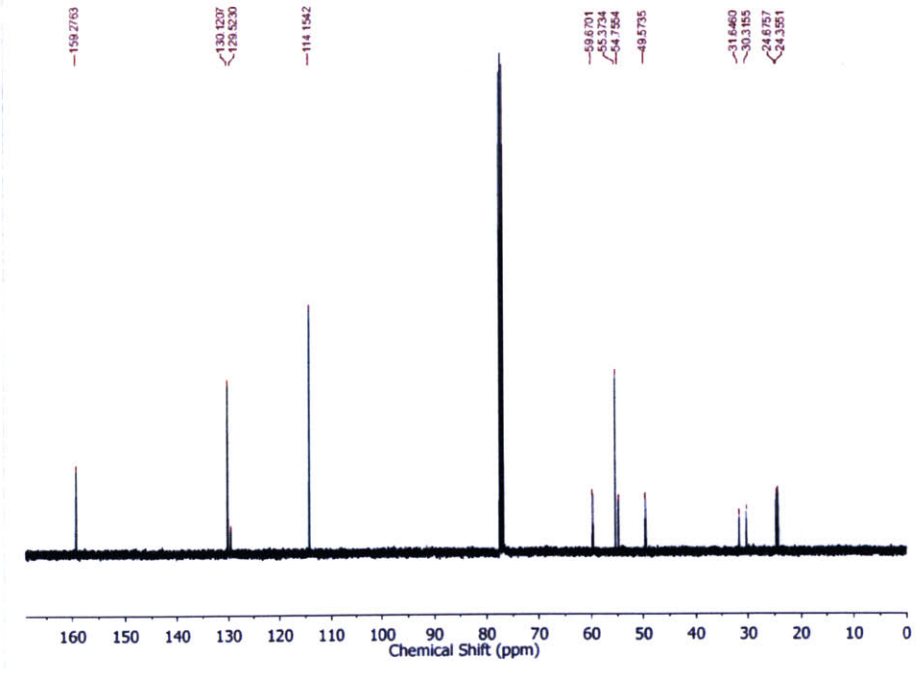


Figure D.2. (N-4-methoxybenzyl)-(1R,2R)-(-)-diaminocyclohexane ^{13}C NMR (101 MHz, $CDCl_3$)

APPENDIX E. CHAPTER 6 SUPPORTING INFORMATION

E.1. EXPERIMENTAL DATA

E.1.1. Reaction of **13** and **14**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 1450.2 mg **13** and 340.3 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid pinacol ester solution was prepared by diluting 1152.1 mg **14** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 29.637 mg **P1-L1**, 28.163 mg **P2-L1**, 28.591 mg **P1-L2**, 30.022 mg **P1-L3**, 33.195 mg **P1-L4**, 25.235 mg **P1-L5**, 21.491 mg **P1-L6**, and 20.850 mg **P1-L7**. A 10 mL volumetric DBU solution was prepared by diluting 2.5203 g DBU in THF and transferring the solution to a 20 mL scintillation vial. Solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M boronic acid pinacol ester, 0.333 M DBU, 0.000835-0.004175 M precatalyst-ligand, and a 5:1 THF-water ratio in the reacting slugs. The product **15** was detected by UV at 340 nm. Reaction data are presented in Table E.1. Optimization results are presented in Table E.2.

Table E.1. Experimental data for reaction optimization of 13 and 14. Yields based on conversion of 13.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
1	**P1-L3**	600.0	30.0	0.498	1.2	0.6
2	**P1-L6**	600.0	30.0	2.515	0.2	0.6
3	**P1-L4**	60.0	30.0	2.508	0.2	0.6
4	**P1-L1**	60.0	30.0	0.513	1.2	0.6
5	**P1-L2**	600.0	30.0	2.513	0.2	0.6
6	**P1-L5**	60.0	30.0	0.508	1.2	0.6
7	**P1-L7**	600.0	30.0	0.506	1.2	0.6
8	**P2-L1**	60.0	30.0	2.509	0.2	0.6
9	**P2-L1**	600.0	110.0	0.496	9.2	4.6
10	**P1-L4**	600.0	110.0	0.512	91.2	46.7
11	**P1-L6**	60.0	110.0	0.498	1.2	0.6
12	**P1-L1**	600.0	110.0	2.509	25.8	64.8
13	**P1-L5**	600.0	110.0	2.512	18.0	45.2
14	**P1-L7**	60.0	110.0	2.499	36.4	91.1
15	**P1-L2**	60.0	110.0	0.508	18.2	9.2
16	**P1-L3**	60.0	110.0	2.489	23.5	58.6
17	**P1-L7**	189.7	65.3	1.123	0.5	0.6
18	**P1-L1**	189.7	65.3	1.106	24.2	26.8
19	**P1-L6**	600.0	65.3	2.515	0.2	0.6

Table E.1. (cont.) Experimental data for reaction optimization of 13 and 14. Yields based on conversion of 13.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
20	**P2-L1**	189.7	65.3	2.509	9.0	22.7
21	**P1-L4**	189.7	65.3	2.508	22.8	57.2
22	**P1-L5**	189.7	65.3	1.106	0.5	0.6
23	**P1-L3**	600.0	65.3	1.106	5.5	6.0
24	**P1-L2**	600.0	65.3	1.129	7.2	8.1
25	**P1-L4**	600.0	110.0	1.106	88.9	98.3
26	**P1-L1**	600.0	110.0	2.509	29.5	73.9
27	**P1-L2**	189.7	110.0	2.513	22.7	57.0
28	**P1-L6**	189.7	110.0	1.127	7.1	8.0
29	**P1-L7**	600.0	110.0	2.499	32.3	80.8
30	**P1-L3**	189.7	110.0	2.489	23.3	57.9
31	**P1-L5**	600.0	110.0	2.512	20.4	51.3
32	**P2-L1**	600.0	110.0	1.131	16.2	18.3
33	**P1-L4**	600.0	30.0	0.512	1.2	0.6
34	**P1-L5**	600.0	30.0	0.508	1.2	0.6
35	**P1-L1**	600.0	30.0	0.513	1.2	0.6
36	**P2-L1**	600.0	30.0	0.496	1.2	0.6
37	**P1-L2**	60.0	110.0	2.513	17.9	45.1
38	**P1-L3**	60.0	110.0	2.489	33.6	83.5
39	**P1-L6**	60.0	110.0	2.515	3.7	9.2
40	**P1-L7**	60.0	110.0	2.499	36.8	92.0
41	**P1-L4**	60.0	67.5	2.508	14.4	36.0
42	**P1-L5**	60.0	66.7	2.512	0.2	0.6
43	**P1-L3**	60.0	67.0	2.489	10.6	26.3
44	**P1-L7**	60.0	66.8	2.499	1.5	3.9
45	**P1-L1**	60.0	66.3	2.509	5.3	13.4
46	**P2-L1**	60.0	67.0	2.509	4.8	12.1
47	**P1-L7**	155.6	110.0	2.499	35.0	87.5
48	**P1-L5**	109.3	110.0	2.482	17.0	42.1
49	**P2-L1**	104.5	110.0	2.509	12.2	30.6
50	**P1-L1**	109.2	110.0	2.482	32.1	79.8
51	**P1-L3**	166.5	110.0	2.489	26.4	65.6
52	**P1-L4**	60.0	110.0	0.512	21.6	11.1
53	**P1-L7**	60.0	110.0	0.506	1.2	0.6
54	**P1-L3**	60.0	110.0	0.498	47.1	23.4
55	**P1-L1**	60.0	110.0	0.513	43.1	22.1
56	**P1-L4**	60.0	110.0	0.512	18.3	9.4
57	**P1-L3**	600.0	110.0	0.968	26.4	25.5
58	**P1-L1**	600.0	110.0	0.971	45.9	44.6
59	**P1-L5**	600.0	110.0	0.957	43.7	41.8
60	**P1-L4**	600.0	110.0	1.268	78.0	98.9
61	**P1-L7**	600.0	110.0	0.814	50.2	40.9
62	**P1-L4**	161.7	110.0	2.104	43.0	90.5
63	**P1-L1**	146.1	110.0	2.509	29.8	74.9

Table E.1. (cont.) Experimental data for reaction optimization of 13 and 14. Yields based on conversion of 13.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
64	**P1-L3**	185.5	110.0	2.489	27.1	67.6
65	**P1-L7**	176.5	110.0	2.499	29.6	73.9
66	**P1-L1**	60.0	110.0	2.266	39.9	90.4
67	**P1-L4**	60.0	110.0	1.915	37.0	70.8
68	**P1-L7**	60.0	110.0	2.303	37.7	86.9
69	**P1-L3**	60.0	110.0	2.323	29.8	69.3
70	**P1-L4**	600.0	110.0	2.508	33.2	88.4
71	**P1-L4**	600.0	110.0	2.508	34.4	86.2
72	**P1-L4**	600.0	67.3	1.214	26.3	31.9
73	**P1-L4**	600.0	68.7	1.268	17.3	22.0
74	**P1-L4**	199.3	110.0	1.241	62.2	77.1
75	**P1-L1**	202.8	110.0	1.592	48.0	76.3
76	**P1-L4**	600.0	110.0	1.160	70.1	81.3
77	**P1-L4**	600.0	110.0	1.160	76.9	89.1
78	**P1-L4**	600.0	110.0	1.106	70.3	77.7
79	**P1-L4**	600.0	110.0	1.106	90.3	99.9
80	**P1-L4**	600.0	110.0	1.187	65.4	77.6
81	**P1-L4**	600.0	110.0	1.106	78.6	86.9
82	**P1-L4**	600.0	66.3	0.998	21.7	21.6
83	**P1-L4**	600.0	67.6	0.998	19.4	19.3
84	**P1-L4**	600.0	110.0	1.241	75.9	94.2
85	**P1-L4**	600.0	110.0	1.268	68.1	86.4
86	**P1-L4**	600.0	110.0	1.187	70.9	84.1
87	**P1-L4**	189.1	110.0	2.508	42.4	106.3
88	**P1-L4**	600.0	110.0	1.268	70.6	89.5
89	**P1-L4**	199.8	110.0	2.508	36.1	90.4
90	**P1-L4**	199.6	110.0	2.508	35.7	89.6
91	**P1-L4**	600.0	110.0	1.241	66.8	82.8
92	**P1-L4**	600.0	110.0	1.025	78.9	80.8
93	**P1-L4**	600.0	110.0	1.079	87.4	94.3
94	**P1-L4**	600.0	110.0	1.133	72.1	81.7
95	**P1-L4**	600.0	110.0	1.052	70.4	74.0
96	**P1-L4**	600.0	110.0	1.106	76.7	84.8

Table E.2. Optimal yield and TON conditions for optimization of 13 and 14. Yields based on conversion of 13.

	Yield Maximum				TON Maximum			
Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	Yield (%)	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON
P1-L1	242.8	110.0	2.152	84.7	242.8	110.0	2.152	39.3
P2-L1	96.0	110.0	2.500	21.9	96.0	110.0	2.500	8.7
P1-L2	82.0	110.0	2.500	49.7	82.0	110.0	2.500	19.9
P1-L3	60.0	110.0	2.500	71.0	60.0	110.0	2.500	28.4
P1-L4	206.9	110.0	2.500	98.1	600.0	110.0	1.195	73.9
P1-L5	60.0	110.0	2.500	52.2	60.0	110.0	2.500	20.9
P1-L6	101.5	110.0	2.500	9.1	101.5	110.0	2.500	3.6
P1-L7	94.8	110.0	2.500	81.0	94.8	110.0	2.500	32.4

E.1.2. Calibration of **17**

The product **17** was synthesized in batch and isolated by column chromatography. A 5 ml volumetric solution was prepared by diluting 182.2 mg **17** and 35.9 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 ml internal standard solution was prepared by diluting 32.8 mg naphthalene with THF and transferring the solution to a 7 mL vial. Both solutions were stored under argon in the liquid handler. Using the automated sample preparation procedure as described previously, two replicates each of slugs containing 0, 0.033, 0.067, 0.100, 0.133, and 0.167 M **17** were prepared and introduced into the system. The slugs were transported by compressed argon at 30°C and 5 min residence time through the FEP reactor and sampled by LC/MS. A calibration was constructed based on integrated peak absorbance measurements of **17** and naphthalene at 270 nm. The calibrated slope was $C_{prod} = 1.26054 * C_{naphthalene} * A_{prod}/A_{naphthalene}$ with $R^2 = 0.996$.

E.1.3. Reaction of **16** and **14**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 815.2 mg **16** and 336.4 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid pinacol ester solution was prepared by diluting 1176.5 mg **14** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 31.057 mg **P1-L1**, 27.263 mg **P2-L1**, 29.287 mg **P1-L2**, 29.242 mg **P1-L3**, 34.030 mg **P1-L4**, 25.335 mg **P1-L5**, 22.442 mg **P1-L6**, and 21.001 mg **P1-L7**. A 10 mL volumetric

DBU solution was prepared by diluting 2.5343 g DBU in THF and transferring the solution to a 20 mL scintillation vial. Solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M boronic acid pinacol ester, 0.333 M DBU, 0.000835-0.004175 M precatalyst-ligand, and a 5:1 THF-water ratio in the reacting slugs. The product **17** was detected by UV at 270 nm. Reaction data are presented in Table E.3. Optimization results are presented in Table E.4.

Table E.3. Experimental data for reaction optimization of 16 and 14.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
1	**P1-L6**	600.0	110.0	2.490	0.1	0.1
2	**P1-L5**	60.0	110.0	0.510	0.3	0.1
3	**P1-L4**	60.0	110.0	0.498	0.3	0.1
4	**P1-L7**	600.0	110.0	2.489	2.7	6.7
5	**P1-L1**	60.0	110.0	2.516	3.0	7.6
6	**P1-L2**	60.0	110.0	2.516	5.2	13.0
7	**P1-L3**	600.0	110.0	0.512	0.3	0.1
8	**P2-L1**	600.0	110.0	0.507	0.3	0.1
9	**P1-L5**	600.0	30.0	2.492	0.1	0.1
10	**P1-L4**	600.0	30.0	2.516	0.2	0.6
11	**P1-L7**	60.0	30.0	0.509	0.3	0.1
12	**P1-L6**	60.0	30.0	0.492	0.3	0.1
13	**P1-L1**	600.0	30.0	0.509	0.3	0.1
14	**P1-L3**	60.0	30.0	2.505	0.1	0.1
15	**P2-L1**	60.0	30.0	2.509	0.1	0.1
16	**P1-L2**	600.0	30.0	0.492	0.3	0.1
17	**P2-L1**	60.0	65.3	1.121	0.4	0.4
18	**P1-L4**	189.7	65.3	2.516	0.4	1.2
19	**P1-L5**	60.0	65.3	2.492	0.1	0.1
20	**P1-L1**	189.7	65.3	1.131	6.1	6.9
21	**P1-L6**	60.0	65.3	1.122	0.1	0.1
22	**P1-L2**	189.7	65.3	1.128	9.3	10.5
23	**P1-L7**	60.0	65.3	1.131	0.1	0.1
24	**P1-L3**	60.0	65.3	2.505	5.5	13.9
25	**P1-L6**	189.7	110.0	2.490	0.1	0.1
26	**P1-L5**	189.7	110.0	1.111	8.6	9.5
27	**P1-L2**	60.0	110.0	2.516	5.4	13.7
28	**P1-L3**	189.7	110.0	1.131	11.0	12.4
29	**P2-L1**	189.7	110.0	2.509	2.5	6.1
30	**P1-L1**	60.0	110.0	2.516	6.4	16.0
31	**P1-L4**	60.0	110.0	1.106	0.4	0.4
32	**P1-L7**	189.7	110.0	2.489	5.5	13.7
33	**P1-L2**	339.6	110.0	2.516	7.6	19.2
34	**P1-L4**	351.3	110.0	2.516	1.1	2.9
35	**P1-L5**	370.2	110.0	2.492	13.2	32.9
36	**P1-L1**	340.3	110.0	2.516	7.7	19.3

Table E.3. (cont.) Experimental data for reaction optimization of 16 and 14.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
37	**P1-L7**	600.0	33.8	2.489	0.1	0.1
38	**P2-L1**	600.0	33.4	2.509	0.6	1.4
39	**P1-L3**	600.0	30.0	2.047	0.1	0.1
40	**P1-L6**	600.0	30.0	2.490	0.1	0.1
41	**P1-L7**	149.0	110.0	2.489	13.2	32.9
42	**P1-L5**	156.7	110.0	2.492	7.7	19.2
43	**P1-L3**	162.3	110.0	2.505	7.9	20.0
44	**P1-L1**	172.3	107.8	2.516	9.4	23.6
45	**P1-L2**	177.6	95.1	2.516	6.1	15.4
46	**P1-L1**	222.1	110.0	0.792	12.9	10.2
47	**P1-L2**	192.5	110.0	2.169	6.4	13.8
48	**P1-L7**	155.6	110.0	1.923	3.3	6.4
49	**P1-L5**	184.7	110.0	2.011	10.8	21.7
50	**P1-L3**	262.1	56.1	0.754	8.4	6.3
51	**P1-L3**	203.4	81.9	2.505	8.5	21.2
52	**P1-L2**	190.0	93.8	2.516	6.7	16.8
53	**P1-L5**	184.5	110.0	2.492	8.6	21.3
54	**P1-L1**	213.2	110.0	2.516	11.6	29.1
55	**P1-L7**	157.7	110.0	2.489	3.2	8.0
56	**P1-L1**	261.0	110.0	2.516	12.3	30.9
57	**P1-L5**	206.5	110.0	2.492	11.5	28.8
58	**P1-L7**	166.9	110.0	2.489	5.1	12.7
59	**P1-L3**	198.4	100.7	2.505	8.2	20.5
60	**P1-L2**	600.0	110.0	1.938	9.9	19.1
61	**P1-L3**	600.0	110.0	1.993	11.1	22.1
62	**P1-L1**	600.0	110.0	1.781	15.8	28.2
63	**P1-L5**	600.0	110.0	1.951	18.3	35.8
64	**P1-L5**	600.0	110.0	1.951	18.1	35.4
65	**P1-L1**	600.0	110.0	1.385	17.4	24.1
66	**P1-L3**	375.4	97.2	1.940	10.5	20.3
67	**P1-L3**	600.0	103.3	2.505	10.5	26.4
68	**P1-L1**	600.0	110.0	2.516	15.4	38.7
69	**P1-L5**	600.0	110.0	2.011	23.4	47.0
70	**P1-L5**	600.0	110.0	2.492	19.8	49.3
71	**P1-L1**	60.0	30.0	2.516	0.1	0.1
72	**P1-L1**	60.0	30.0	2.516	0.2	0.5
73	**P1-L5**	600.0	110.0	2.492	22.4	55.9
74	**P1-L1**	600.0	47.1	2.516	19.7	49.7
75	**P1-L5**	600.0	71.0	2.492	0.1	0.1
76	**P1-L5**	600.0	110.0	2.492	19.8	49.4
77	**P1-L2**	600.0	39.8	2.516	0.5	1.3
78	**P1-L3**	600.0	38.1	2.505	2.1	5.4
79	**P1-L1**	600.0	61.0	2.516	6.7	16.8
80	**P1-L3**	600.0	69.4	2.505	11.5	28.9
81	**P1-L1**	600.0	63.0	2.516	7.4	18.7
82	**P1-L1**	600.0	110.0	2.516	6.6	16.6

Table E.3. (cont.) Experimental data for reaction optimization of 16 and 14.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
83	**P1-L5**	600.0	110.0	2.492	21.4	53.4
84	**P1-L5**	600.0	110.0	2.492	18.1	45.1
85	**P1-L3**	600.0	110.0	2.505	11.8	29.6
86	**P1-L5**	600.0	110.0	1.711	17.9	30.6
87	**P1-L1**	600.0	110.0	2.516	8.1	20.4
88	**P1-L5**	600.0	110.0	2.492	19.3	48.3
89	**P1-L1**	600.0	110.0	2.516	8.1	20.4
90	**P1-L1**	600.0	110.0	2.516	8.3	20.9
91	**P1-L5**	600.0	110.0	2.492	21.3	53.1
92	**P1-L5**	600.0	110.0	1.771	22.9	40.7
93	**P1-L1**	600.0	30.0	2.488	1.2	2.9
94	**P1-L1**	600.0	30.0	2.488	1.5	3.7
95	**P1-L5**	600.0	110.0	2.492	20.7	51.7
96	**P1-L5**	600.0	110.0	1.921	24.1	46.3

Table E.4. Optimal yield and TON conditions for optimization of 16 and 14.

	Yield Maximum				TON Maximum			
Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	Yield (%)	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON
P1-L1	600.0	78.5	2.500	44.4	600.0	78.5	2.500	17.8
P2-L1	600.0	96.4	2.500	7.2	600.0	96.4	2.500	2.9
P1-L2	600.0	87.6	2.500	29.9	600.0	87.6	2.500	12.0
P1-L3	600.0	85.9	2.500	34.4	600.0	85.9	2.500	13.7
P1-L4	600.0	90.9	2.500	3.2	600.0	90.9	2.500	1.3
P1-L5	600.0	110.0	2.500	46.5	600.0	110.0	2.061	20.3
P1-L6	600.0	36.4	2.500	1.4	600.0	36.4	2.500	0.6
P1-L7	600.0	110.0	2.500	24.6	600.0	110.0	2.500	9.9

E.1.4. Reaction of **16** and **18**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 829.4 mg **16** and 325.2 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid solution was prepared by diluting 863.5 mg **18** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 31.128 mg **P1-L1**, 28.858 mg **P2-L1**, 28.706 mg **P1-L2**, 30.302 mg **P1-L3**, 35.144 mg **P1-L4**, 25.230 mg **P1-L5**, 22.442 mg **P1-L6**, and 20.658 mg **P1-L7**. A 10 mL volumetric DBU solution was prepared by diluting 2.5373 g DBU in THF and transferring the solution to a

20 mL scintillation vial. Solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M boronic acid, 0.333 M DBU, 0.000835-0.004175 M precatalyst-ligand, and a 5:1 THF-water ratio in the reacting slugs. The product **19** was detected by UV at 340 nm. Reaction data are presented in Table E.5. Optimization results are presented in Table E.6.

Table E.5. Experimental data for reaction optimization of 16 and 18. Yields based on conversion of 16.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
1	**P1-L4**	60.0	30.0	2.513	0.1	0.2
2	**P1-L2**	600.0	30.0	2.494	0.1	0.2
3	**P1-L1**	60.0	30.0	0.510	0.3	0.2
4	**P1-L5**	600.0	30.0	2.511	0.1	0.2
5	**P1-L6**	60.0	30.0	0.499	0.3	0.2
6	**P1-L7**	600.0	30.0	0.501	0.3	0.2
7	**P1-L3**	60.0	30.0	2.512	0.1	0.2
8	**P2-L1**	600.0	30.0	0.509	0.3	0.2
9	**P2-L1**	60.0	110.0	2.515	32.7	82.3
10	**P1-L5**	60.0	110.0	0.508	0.3	0.2
11	**P1-L4**	600.0	110.0	0.514	0.3	0.2
12	**P1-L6**	600.0	110.0	2.496	0.1	0.2
13	**P1-L3**	600.0	110.0	0.503	50.4	25.3
14	**P1-L7**	60.0	110.0	2.504	16.8	42.1
15	**P1-L2**	60.0	110.0	0.510	46.2	23.5
16	**P1-L1**	600.0	110.0	2.493	36.5	91.0
17	**P1-L3**	189.7	110.0	1.117	77.2	86.2
18	**P2-L1**	189.7	110.0	2.515	34.2	86.1
19	**P1-L6**	189.7	110.0	2.496	0.1	0.2
20	**P1-L5**	60.0	110.0	1.106	4.0	4.4
21	**P1-L1**	189.7	110.0	2.493	31.5	78.6
22	**P1-L4**	189.7	110.0	1.114	2.6	2.9
23	**P1-L2**	60.0	110.0	2.494	34.6	86.4
24	**P1-L6**	60.0	65.3	1.109	0.1	0.2
25	**P1-L7**	60.0	110.0	2.504	19.8	49.7
26	**P1-L2**	189.7	65.3	1.106	3.9	4.3
27	**P1-L5**	189.7	65.3	2.511	2.0	5.1
28	**P1-L1**	60.0	65.3	1.105	6.9	7.6
29	**P1-L3**	60.0	65.3	2.512	8.8	22.0
30	**P1-L7**	189.7	65.3	1.113	0.7	0.8
31	**P1-L4**	60.0	65.3	2.513	0.3	0.7
32	**P2-L1**	60.0	65.3	1.130	6.1	6.9
33	**P1-L4**	600.0	105.1	2.513	1.6	4.1
34	**P1-L1**	600.0	110.0	2.352	32.5	76.4
35	**P1-L2**	600.0	110.0	2.438	32.9	80.1
36	**P1-L3**	600.0	110.0	2.317	37.6	87.1
37	**P2-L1**	600.0	110.0	2.345	34.2	80.1
38	**P1-L7**	600.0	110.0	2.393	20.8	49.7
39	**P1-L5**	600.0	71.7	1.196	0.1	0.2

Table E.5. (cont.) Experimental data for reaction optimization of 16 and 18. Yields based on conversion of 16.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
40	**P1-L5**	600.0	48.3	0.748	0.2	0.2
41	**P1-L7**	124.3	110.0	1.113	9.8	10.9
42	**P1-L2**	202.9	110.0	1.219	51.2	62.4
43	**P2-L1**	160.9	110.0	1.074	55.4	59.5
44	**P1-L1**	123.2	110.0	1.105	77.4	85.5
45	**P1-L3**	600.0	83.8	1.061	48.5	51.5
46	**P2-L1**	60.0	110.0	1.413	31.0	43.8
47	**P1-L7**	60.0	110.0	1.419	24.7	35.0
48	**P1-L2**	60.0	110.0	1.672	52.9	88.5
49	**P1-L3**	60.0	110.0	1.312	59.6	78.2
50	**P1-L1**	60.0	110.0	1.417	61.6	87.2
51	**P1-L2**	186.8	110.0	2.494	30.2	75.2
52	**P1-L1**	149.9	110.0	2.493	31.9	79.6
53	**P2-L1**	186.8	103.6	2.515	33.1	83.1
54	**P1-L3**	179.9	110.0	2.512	34.3	86.2
55	**P1-L1**	60.0	110.0	1.502	56.4	84.7
56	**P1-L2**	60.0	110.0	1.927	43.5	83.9
57	**P1-L3**	137.9	110.0	1.312	60.4	79.2
58	**P1-L3**	60.0	110.0	1.647	49.3	81.2
59	**P1-L1**	177.6	110.0	1.048	77.2	80.9
60	**P1-L2**	188.8	110.0	1.446	49.3	71.3
61	**P1-L3**	211.9	110.0	1.368	59.7	81.7
62	**P1-L2**	239.7	110.0	1.332	51.0	68.0
63	**P1-L1**	253.3	110.0	1.048	82.7	86.7
64	**P1-L1**	108.9	110.0	1.275	65.1	83.0
65	**P1-L3**	92.2	110.0	1.787	50.3	89.9
66	**P1-L2**	66.6	110.0	2.409	35.7	86.1
67	**P1-L3**	192.1	110.0	1.731	48.0	83.0
68	**P1-L1**	221.3	110.0	1.077	75.7	81.5
69	**P1-L2**	192.7	110.0	1.871	41.4	77.5
70	**P1-L1**	180.1	110.0	1.303	64.9	84.6
71	**P1-L2**	144.0	110.0	1.899	44.4	84.3
72	**P1-L3**	146.6	110.0	2.038	42.9	87.4
73	**P1-L1**	60.0	110.0	2.493	34.3	85.6
74	**P1-L3**	101.5	110.0	1.926	46.8	90.1
75	**P1-L2**	60.0	110.0	2.126	38.2	81.2
76	**P1-L1**	218.5	110.0	1.105	80.6	89.1
77	**P1-L2**	135.7	110.0	1.814	40.4	73.3
78	**P1-L1**	600.0	110.0	0.935	85.3	79.7
79	**P1-L3**	147.9	110.0	1.591	51.9	82.6
80	**P1-L2**	162.9	110.0	1.701	39.9	67.8
81	**P1-L3**	162.1	110.0	1.675	53.6	89.8
82	**P1-L1**	600.0	110.0	1.218	74.1	90.3
83	**P1-L3**	600.0	110.0	1.424	49.1	69.9
84	**P1-L2**	600.0	110.0	1.417	66.2	93.8
85	**P1-L3**	60.0	110.0	1.926	45.3	87.3

Table E.5. (cont.) Experimental data for reaction optimization of 16 and 18. Yields based on conversion of 16.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
87	**P1-L2**	60.0	110.0	1.757	44.2	77.7
88	**P1-L3**	600.0	110.0	1.787	45.2	80.8
89	**P1-L2**	600.0	110.0	1.616	52.1	84.2
90	**P1-L1**	600.0	110.0	1.303	64.2	83.7
91	**P1-L1**	600.0	110.0	1.218	65.2	79.4
92	**P1-L2**	600.0	110.0	1.587	47.1	74.8
93	**P1-L3**	600.0	110.0	1.787	43.5	77.8
94	**P1-L3**	60.0	40.2	0.503	0.3	0.2
95	**P1-L3**	60.0	40.1	0.503	0.3	0.2
96	**P1-L3**	229.2	110.0	1.619	49.4	79.9

Table E.6. Optimal yield and TON conditions for optimization of 16 and 18. Yields based on conversion of 16.

	Yield Maximum				TON Maximum			
Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	Yield (%)	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON
P1-L1	198.1	110.0	1.697	97.3	233.2	110.0	1.170	74.9
P2-L1	186.2	110.0	1.909	77.9	186.2	110.0	1.909	40.8
P1-L2	60.0	110.0	1.853	81.4	60.0	110.0	1.853	43.9
P1-L3	198.1	110.0	1.697	87.4	198.1	110.0	1.697	51.5
P1-L4	600.0	89.1	2.330	3.5	600.0	89.1	2.330	1.5
P1-L5	600.0	86.7	2.383	5.8	600.0	86.7	2.383	2.4
P1-L6	600.0	30.0	2.500	28.6	600.0	30.0	2.500	11.5
P1-L7	165.6	110.0	2.009	40.2	165.6	110.0	2.009	20.0

E.1.5. Reaction of **9** and **7**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 816.4 mg **9** and 327.1 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid solution was prepared by diluting 1079.4 mg **7** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 29.622 mg **P1-L1**, 27.872 mg **P2-L1**, 28.706 mg **P1-L2**, 29.369 mg **P1-L3**, 34.043 mg **P1-L4**, 25.128 mg **P1-L5**, 23.247 mg **P1-L6**, and 20.128 mg **P1-L7**. A 10 mL volumetric DBU solution was prepared by diluting 2.5024 g DBU in THF and transferring the solution to a 20 mL scintillation vial. Solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M boronic acid, 0.333 M DBU, 0.000835-0.004175 M precatalyst-ligand, and a 5:1

THF-water ratio in the reacting slugs. The product **10** was detected by UV at 340 nm. Reaction data are presented in Table E.7. Optimization results are presented in Table E.8. Screening results at 80°C, 97°C, and 110°C are presented in Table E.9.

Table E.7. Experimental data for reaction optimization of 9 and 7. Yields based on conversion of 9.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
1	**P1-L6**	600.0	110.0	2.504	11.7	29.4
2	**P1-L5**	600.0	110.0	2.499	21.4	53.6
3	**P1-L7**	60.0	110.0	2.491	9.5	23.7
4	**P1-L3**	60.0	110.0	0.502	129.2	64.8
5	**P2-L1**	600.0	110.0	0.506	99.3	50.3
6	**P1-L2**	600.0	110.0	0.511	107.8	55.1
7	**P1-L4**	60.0	110.0	0.489	32.4	15.8
8	**P1-L1**	60.0	110.0	2.501	37.6	94.1
9	**P1-L1**	600.0	30.0	0.500	2.9	1.5
10	**P1-L6**	60.0	30.0	0.501	0.2	0.1
11	**P2-L1**	60.0	30.0	2.508	1.6	4.0
12	**P1-L7**	600.0	30.0	0.503	0.2	0.1
13	**P1-L5**	60.0	30.0	0.500	0.2	0.1
14	**P1-L4**	600.0	30.0	2.492	4.6	11.5
15	**P1-L2**	60.0	30.0	2.507	0.0	0.1
16	**P1-L3**	600.0	30.0	2.510	2.7	6.9
17	**P2-L1**	600.0	65.3	2.508	38.1	95.5
18	**P1-L1**	600.0	65.3	1.120	65.3	73.1
19	**P1-L6**	189.7	65.3	1.127	1.3	1.4
20	**P1-L4**	600.0	65.3	2.492	29.7	74.1
21	**P1-L2**	189.7	65.3	2.507	36.4	91.3
22	**P1-L7**	600.0	65.3	1.126	0.7	0.8
23	**P1-L3**	189.7	65.3	1.123	60.8	68.3
24	**P1-L5**	189.7	65.3	1.131	1.3	1.5
25	**P1-L1**	189.7	110.0	2.501	34.9	87.2
26	**P1-L4**	189.7	110.0	1.124	50.6	56.9
27	**P1-L6**	600.0	110.0	2.504	16.1	40.4
28	**P1-L3**	600.0	110.0	2.510	32.8	82.2
29	**P1-L7**	189.7	110.0	2.491	11.6	28.8
30	**P1-L2**	600.0	110.0	1.120	72.2	80.9
31	**P1-L5**	600.0	110.0	2.499	25.7	64.3
32	**P2-L1**	189.7	110.0	1.109	80.6	89.4
33	**P1-L5**	114.7	110.0	1.079	31.4	33.9
34	**P1-L6**	114.7	110.0	1.077	21.4	23.0
35	**P1-L7**	250.1	110.0	0.982	26.8	26.3
36	**P2-L1**	150.3	71.6	1.037	27.7	28.7
37	**P1-L2**	145.8	81.2	1.071	74.9	80.2
38	**P1-L4**	226.1	69.5	1.148	8.1	9.2
39	**P1-L3**	326.4	109.2	0.932	96.1	89.6
40	**P1-L1**	168.9	89.8	0.858	100.7	86.4
41	**P1-L1**	600.0	61.9	2.501	37.8	94.4

Table E.7. (cont.) Experimental data for reaction optimization of 9 and 7. Yields based on conversion of 9.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
42	**P1-L2**	227.6	101.6	2.507	33.6	84.1
43	**P1-L4**	600.0	83.7	2.492	35.9	89.4
44	**P1-L5**	60.0	108.5	2.499	19.4	48.4
45	**P1-L3**	600.0	75.7	2.510	36.1	90.5
46	**P2-L1**	182.6	104.0	2.508	34.1	85.5
47	**P1-L2**	600.0	61.5	1.387	59.5	82.6
48	**P1-L1**	600.0	106.8	1.215	72.5	88.1
49	**P1-L3**	600.0	84.7	1.793	48.3	86.5
50	**P2-L1**	600.0	96.5	1.977	41.3	81.8
51	**P1-L5**	600.0	110.0	1.236	34.8	43.0
52	**P1-L2**	600.0	73.7	2.507	35.1	88.0
53	**P1-L3**	134.4	110.0	1.530	55.8	85.3
54	**P2-L1**	196.1	91.7	2.050	42.1	86.2
55	**P1-L1**	227.9	106.5	0.786	108.5	85.3
56	**P2-L1**	260.4	88.2	2.098	40.0	83.9
57	**P1-L2**	600.0	80.8	1.996	47.6	95.0
58	**P1-L3**	169.1	98.5	2.056	39.4	81.0
59	**P1-L1**	236.9	105.3	0.739	106.8	78.9
60	**P1-L1**	60.0	70.4	1.715	39.1	67.0
61	**P1-L3**	60.0	82.1	2.295	40.0	91.8
62	**P1-L2**	60.0	85.6	2.507	34.5	86.6
63	**P1-L1**	600.0	97.7	0.715	115.4	82.5
64	**P1-L2**	600.0	72.8	1.314	66.2	87.0
65	**P1-L3**	319.6	94.9	1.171	71.4	83.6
66	**P1-L1**	299.9	73.5	2.501	39.9	99.8
67	**P1-L3**	237.1	95.6	1.554	51.4	79.9
68	**P1-L2**	600.0	87.0	1.801	46.9	84.6
69	**P1-L2**	196.0	110.0	1.412	59.9	84.5
70	**P1-L1**	202.8	108.8	1.358	49.5	67.3
71	**P1-L2**	205.1	76.3	1.655	50.0	82.7
72	**P1-L1**	222.9	74.7	1.572	51.9	81.6
73	**P1-L1**	167.6	73.2	2.287	38.4	87.9
74	**P1-L2**	200.4	75.8	1.947	50.7	98.7
75	**P1-L2**	600.0	84.2	2.507	35.5	88.9
76	**P1-L3**	600.0	98.2	2.510	33.0	82.8
77	**P1-L1**	600.0	97.7	2.501	33.5	83.8
78	**P1-L2**	600.0	70.7	1.144	69.3	79.3
79	**P1-L3**	149.9	93.0	2.510	33.6	84.2
80	**P1-L1**	600.0	96.7	0.739	105.3	77.8
81	**P1-L3**	216.2	105.7	1.386	61.7	85.6
82	**P1-L1**	600.0	97.3	0.715	118.5	84.7
83	**P1-L1**	600.0	88.6	0.977	87.6	85.6
84	**P1-L2**	600.0	77.4	1.193	66.2	78.9
85	**P1-L2**	600.0	77.2	1.314	67.3	88.4
86	**P1-L3**	126.4	97.6	2.510	28.9	72.6
87	**P1-L2**	276.4	69.3	2.507	39.7	99.5

Table E.7. (cont.) Experimental data for reaction optimization of 9 and 7. Yields based on conversion of 9.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
88	**P1-L1**	60.0	110.0	0.643	129.4	83.2
89	**P1-L1**	60.0	110.0	0.667	105.1	70.1
90	**P1-L2**	258.2	71.6	2.507	33.7	84.5
91	**P1-L2**	211.7	104.7	1.193	65.2	77.8
92	**P1-L1**	113.3	110.0	1.120	81.4	91.1
93	**P1-L2**	600.0	58.7	2.507	34.4	86.3
94	**P1-L1**	60.0	110.0	1.096	86.3	94.5
95	**P1-L1**	154.1	110.0	1.096	80.9	88.6
96	**P1-L2**	600.0	72.4	2.507	33.7	84.4
97	**P1-L1**	60.0	110.0	0.977	55.5	54.2

Table E.8. Optimal yield and TON conditions for optimization of 9 and 7. Yields based on conversion of 9.

	Yield Maximum				TON Maximum			
Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	Yield (%)	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON
P1-L1	212.5	85.1	2.401	99.7	282.5	97.2	1.012	88.7
P2-L1	366.8	88.2	2.088	88.1	366.8	88.2	2.088	42.2
P1-L2	307.5	78.0	2.500	95.8	373.9	85.9	1.381	65.0
P1-L3	235.8	92.1	1.983	90.4	256.0	95.5	1.442	61.7
P1-L4	600.0	79.6	2.500	72.6	600.0	79.6	2.500	29.0
P1-L5	91.6	110.0	1.707	54.1	91.6	110.0	1.707	31.7
P1-L6	100.9	110.0	1.790	33.7	100.9	110.0	1.790	18.8
P1-L7	143.4	110.0	1.760	27.4	143.4	110.0	1.760	15.6

Table E.9. Experimental data for screening of 9 and 7. Yields based on conversion of 9.

Experiment	Precat-Ligand	t_{res} (s)	T (°C)	Cat. Loading (mol%)	TON	Yield (%)
1	**P1-L1**	150.0	80.0	1.018	61.5	62.6
2	**P1-L1**	600.0	80.0	1.018	81.7	83.2
3	**P1-L1**	282.0	80.0	1.018	74.8	76.2
4	**P1-L1**	600.0	80.0	1.018	83.5	85.0
5	**P1-L1**	282.0	97.0	1.018	86.7	88.3
6	**P1-L1**	282.0	97.0	1.018	89.0	90.6
7	**P1-L1**	450.0	97.0	1.018	86.3	87.8
8	**P1-L1**	600.0	97.0	1.018	83.7	85.2
9	**P1-L1**	600.0	110.0	1.018	82.9	84.4
10	**P1-L1**	282.0	110.0	1.018	85.0	86.5
11	**P1-L1**	150.0	110.0	1.018	87.9	89.5
12	**P1-L1**	450.0	110.0	1.018	82.0	83.5
13	**P1-L1**	450.0	80.0	1.018	78.0	79.4
14	**P1-L1**	450.0	80.0	1.018	77.5	78.9
15	**P1-L1**	282.0	80.0	1.018	71.6	72.9
16	**P1-L1**	150.0	80.0	1.018	60.6	61.7
17	**P1-L1**	600.0	97.0	1.018	85.3	86.8
18	**P1-L1**	450.0	97.0	1.018	92.8	94.4
19	**P1-L1**	150.0	97.0	1.018	84.9	86.4
20	**P1-L1**	150.0	97.0	1.018	87.5	89.0
21	**P1-L1**	282.0	110.0	1.018	85.4	86.9
22	**P1-L1**	150.0	110.0	1.018	86.0	87.5
23	**P1-L1**	600.0	110.0	1.018	78.6	80.0
24	**P1-L1**	450.0	110.0	1.018	84.5	86.0

E.1.6. Reaction of **16** and **20**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 816.0 mg **16** and 326.9 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid solution was prepared by diluting 726.7 mg **20** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 29.516 mg **P1-L1**, 23.738 mg **P1-L5**, and 20.394 mg **P1-L7**. Individual 2 mL ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF or water. Ligand masses were 49.629 mg **L1** (in THF), 39.050 mg **L5**-HBF_4 (in THF), and 29.177 mg **L7**-HBF_4 (in water). A 10 mL volumetric DBU solution was prepared by diluting 2.5126 g DBU in THF and transferring the solution to a 20 mL scintillation vial. Solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M boronic acid,

0.333 M DBU, 0.0023 M precatalyst-ligand, 0.0000-0.0046 M excess ligand, and a 5:1 THF-water ratio in the reacting slugs. The product **17** was detected by UV at 270 nm. Reaction data are presented in Table E.10. Optimization results are presented in Table E.11.

Table E.10. Experimental data for reaction optimization of 16 and 20.

Experiment	Precat-Ligand	*T* (°C)	Excess Ligand Equiv.	TON	Yield (%)
1	**P1-L5**	110.0	2.0	3.8	16.1
2	**P1-L1**	110.0	2.0	6.9	29.0
3	**P1-L5**	110.0	0.0	25.4	35.7
4	**P1-L7**	110.0	2.1	2.0	8.1
5	**P1-L1**	110.0	0.0	17.5	24.5
6	**P1-L7**	110.0	0.0	10.1	14.1
7	**P1-L5**	30.0	2.0	0.0	0.1
8	**P1-L1**	30.0	2.0	2.8	11.7
9	**P1-L7**	30.0	2.1	0.0	0.1
10	**P1-L5**	30.0	0.0	0.1	0.1
11	**P1-L7**	30.0	0.0	0.1	0.1
12	**P1-L1**	30.0	0.0	8.6	12.0
13	**P1-L7**	65.3	0.8	0.1	0.1
14	**P1-L1**	65.3	0.8	9.6	23.6
15	**P1-L5**	65.3	0.0	0.1	0.1
16	**P1-L1**	65.3	0.0	14.1	19.6
17	**P1-L7**	65.3	0.0	0.1	0.1
18	**P1-L5**	65.3	0.8	0.1	0.1
19	**P1-L5**	110.0	0.8	21.9	54.0
20	**P1-L5**	110.0	0.0	28.8	40.5
21	**P1-L1**	110.0	0.8	14.9	36.4
22	**P1-L7**	110.0	0.8	8.4	20.4
23	**P1-L1**	110.0	0.0	21.7	30.4
24	**P1-L7**	110.0	0.0	9.7	13.6
25	**P1-L5**	110.0	0.6	20.9	46.5
26	**P1-L1**	110.0	0.7	11.4	27.0
27	**P1-L5**	110.0	0.5	23.8	49.2
28	**P1-L5**	110.0	0.5	21.6	44.4
29	**P1-L5**	110.0	0.4	22.9	45.4
30	**P1-L5**	110.0	0.2	24.5	42.4
31	**P1-L5**	110.0	0.0	33.0	46.4
32	**P1-L5**	110.0	0.0	35.9	50.4
33	**P1-L5**	110.0	0.0	37.8	53.1
34	**P1-L5**	110.0	0.0	41.1	57.9
35	**P1-L5**	110.0	0.0	34.5	48.5
36	**P1-L5**	110.0	0.0	39.8	56.0
37	**P1-L5**	110.0	0.0	37.0	52.0
38	**P1-L5**	110.0	0.0	37.2	52.4
39	**P1-L5**	110.0	0.0	32.4	45.6
40	**P1-L5**	110.0	0.0	32.7	46.0
41	**P1-L5**	110.0	0.0	30.9	43.4

Table E.11. Optimal yield and TON conditions for optimization of 16 and 20.

	Yield Maximum			TON Maximum (Per Ligand Equiv)		
Precat-Ligand	*T* (°C)	Excess Ligand Equivalents	Yield (%)	*T* (°C)	Excess Ligand Equivalents	TON
P1-L1	110.0	0.276	31.4	110.0	0.276	17.9
P1-L5	110.0	0.225	50.0	110.0	0.000	34.4
P1-L7	110.0	0.328	15.9	110.0	0.328	8.7

E.1.7. Time-Course Evolution of **18** and Reaction of **16** and **21**

A 5 mL volumetric aryl halide and naphthalene solution was prepared by diluting 809.2 mg **16** and 347.4 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric naphthalene solution was prepared by diluting 329.3 mg naphthalene with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid solution was prepared by diluting 848.3 mg **18** with THF and transferring the solution to a 7 mL vial. A 5 mL volumetric boronic acid pinacol ester solution was prepared by diluting 1215.5 mg **21** with THF and transferring the solution to a 7 mL vial. Individual 2 mL precatalyst-ligand solutions were prepared by charging the solid to a tapered 2 mL vial, then dosing with 2 mL THF. Catalyst masses were 30.409 mg **P1-L1** and 25.430 mg **P1-L5**. A 10 mL volumetric DBU solution was prepared by diluting 2.5235 g DBU in THF and transferring the solution to a 20 mL scintillation vial. For reaction of **16** and **21**, solution volumes were automatically sampled to achieve 0.167 M aryl halide, 0.250 M **21**, 0.333 M DBU, 0.0020 M precatalyst-ligand, and a 5:1 THF-water ratio. The product **19** was detected by UV at 340 nm. The yield with **P1-L1** was 87%. The yield with **P1-L5** was 85%. For the time-course study of the boronic acid, solution volumes were automatically sampled to achieve 0.250 M **18** or **21**, 0.333 M DBU, and a 5:1 THF-water ratio. The slugs were transported at 110°C, with the boronic acid **18** detected by UV at 300 nm. Solution volumes were also automatically sampled to achieve 0.050 M naphthalene, 0.250 M **18** or **21**, 0.333 M DBU, and a 5:1 THF-water ratio. The slugs were transported at 110°C, with the boronic acid **18** detected by UV at 300 nm. The results were compared to slugs comprising 0.050 M naphthalene, 0.250 M **18** or **21**, and 100% THF flowed at 30°C.

E.2. NMR SPECTRA

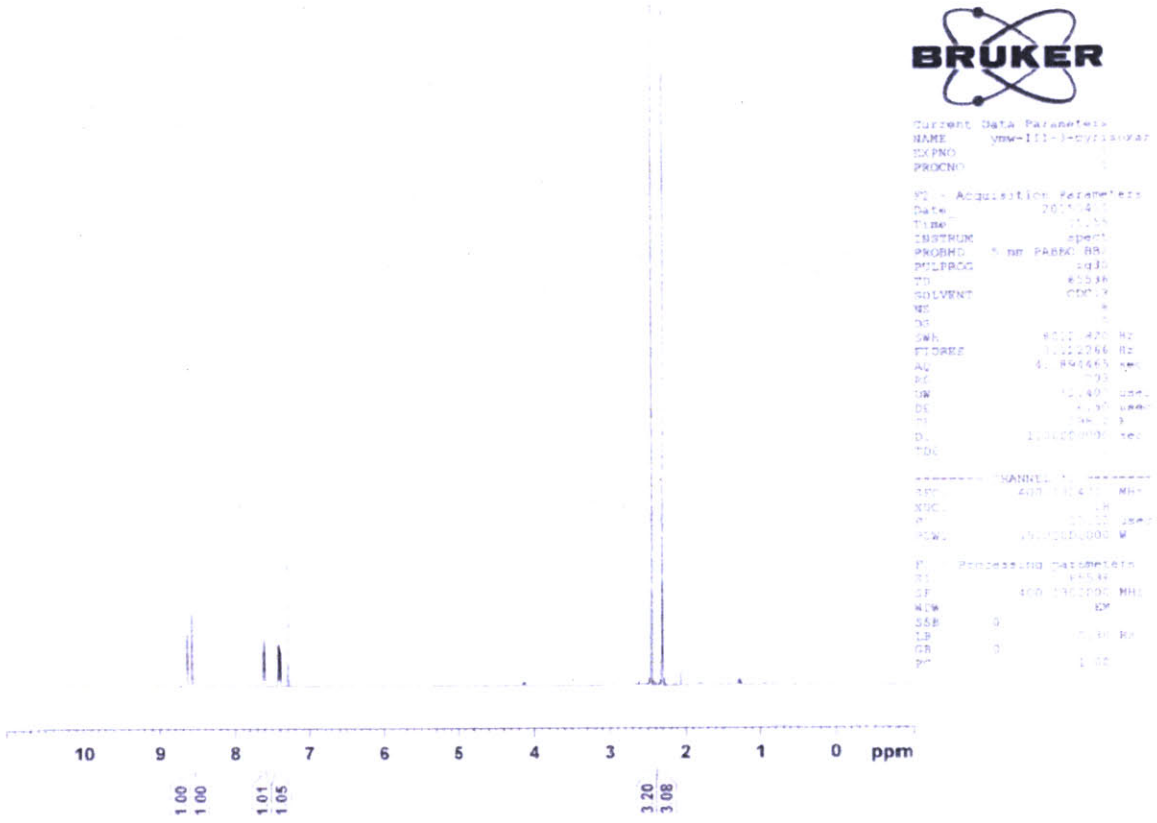


Figure E.1. 3-(3,5-dimethyl-4-isoxazolyl)-pyridine (17) ^{1}H NMR (400 MHz, $CDCl_3$)