Chapter 4
Intermittent Flow and Practical Considerations for Continuous Drug Substance Manufacturing

Martin D. Johnson, Scott A. May, Jennifer McClary Groh, Timothy Braden, and Richard D. Spencer

Abstract Intermittent flow enables slurry flow out of continuous stirred tank crystallizers without solids plugging or clogging. It enables semi-continuous filtration, washing, and re-dissolving downstream from continuous crystallization. It allows solvent exchange distillation with strip to dryness in rotary evaporators to be a legitimate manufacturing unit operation for small-volume continuous processes. Intermittent flow back pressure regulation and vapor–liquid separation downstream from continuous high-pressure hydrogenation reactors tolerates a small amount of solids precipitate flowing out of the reactors without clogging or plugging, and they promote efficient pressure purge stripping of excess gas reagent. Intermittent flow stirred tank reactors are a practical alternative to plug flow reactors (PFRs) for heterogeneous reactions. Eleven examples of continuous reactions are given that have been run at manufacturing scale in PFRs. Mean residence time ranges from 0.7 to 24 h in the 11 examples; therefore, it is not necessary for a reaction to be extremely fast in order to be a viable candidate for flow chemistry. This chapter gives many general guidelines on how to design and operate a continuous process, avoiding many of the common operational, equipment, analytical, and process chemistry pitfalls. The continuous process checklist serves to help prevent common oversights.

Keywords Continuous processing · Drug substance · PFR · CSTR · Intermittent flow · Recycle · Crystallization · Filtration · Distillation · Reaction

M. D. Johnson (✉) · S. A. May · J. M. Groh · T. Braden · R. D. Spencer
Eli Lilly and Company, Indianapolis, IN, USA
e-mail: johnson_martin_d@lilly.com; may_scott_a@lilly.com; groh_jennifer_mcclary@lilly.com; braden_timothy@lilly.com; spencer_richard_d@lilly.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_4

Abbreviations

API	Active pharmaceutical ingredient
CSD	Crystal size distribution
CSTR	Continuous stirred tank reactor
DARA	Direct asymmetric reductive amination
DCS	Distributed control system
FBRM	Focused Beam Reflectance Measurement
i.d.	Inside diameter
MSMPR	Mixed suspension mixed product removal
PAT	Process analytical technology
PFA	Perfluoroalkoxy
PFR	Plug flow reactor
PIDs	Process and instrumentation diagrams
PMI	Process Mass Intensity
RTD	Residence time distribution
τ	Mean residence time

4.1 Intermittent Flow

Intermittent flow enables solids processing, heterogeneous reactions, different modes of reagent addition, crystallization, filtration, more efficient solvent exchanges, back pressure regulation without restricting orifices, and complete conversion for reactions. Other researchers have used intermittent flow. In the work of Adams this is known as semi-continuous operation (Adams and Pascall 2012). Adams describes the forced cyclic process with no steady states, and demonstrates that it is possible to achieve multiple separations steps and high reaction conversion using fewer vessels than would be required in truly continuous operation.

4.1.1 Slurry Flow Out of Continuous Stirred Tank Crystallizers

Intermittent flow of slurries is especially important when average volumetric throughput is less than about 200 mL/min and internal tubing diameter is more than about 4 mm. If Reynolds number is less than about 2000, then flow is in the laminar regime, which is not sufficient for keeping the solids suspended and keeping the solids from clogging. Solids gradually settle in the piping and fittings and eventually accumulate enough to clog the process tubes. Settling and accumulation of solids especially occurs at locations of fittings, elbows, tees, valves, or any other slight constriction or expansion in the flow path. On the other hand, flowing slurry quickly and intermittently out of a stirred tank for brief time periods enables high

enough linear velocities to achieve turbulence. Intermittent flow is generally achieved by pressure differences and sequenced automated block valves, but it can also be achieved by automated intermittent mechanical pumping.

Continuous crystallization was utilized for kinetic rejection of a chiral impurity in the manufacture of a key intermediate in the synthesis of LY500307 (Johnson et al. 2012). A fully continuous process step was used to generate 144 kg of penultimate in laboratory fume hoods and a laboratory bunker. The continuous process train included high-pressure asymmetric hydrogenation, liquid–liquid extraction, solvent exchange distillation, crystallization, and filtration. Crude reaction product solution had ee in the range 92–94% throughout the continuous campaign. Crystallization was required to upgrade the advanced intermediate to greater than 99% ee. Unfortunately, crude ee was on the unfavorable side of the eutectic; therefore, thermodynamic equilibrium in the crystallizer would fail to provide the required ee upgrade. However, kinetics favored crystallization of the desired chirality product, because the undesired enantiomer crystallized out of solution more slowly. Continuous crystallization in stirred tanks in series is superior to batch for kinetic impurity rejection because the crystallizers operate in the kinetic regime, by definition, with steady-state supersaturation. Stirred tank crystallizers are often called mixed-suspension mixed-product removal (MSMPR) crystallizers. The product was in solution prior to crystallization, with toluene being the main solvent. Isopropyl alcohol anti-solvent and cooling were used to generate the supersaturation that drove the crystallization. The desired compound dissolved in toluene was continuously pumped into the first of two MSMPRs in series. Anti-solvent was continuously pumped into both MSMPRs. A sketch of the setup is shown in Fig. 4.1.

Overall average slurry flow rate was 61 mL/min from MSMPR1 to MSMPR2, and it was 119 mL/min from MSMPR2 to the dual alternating filters. Without intermittent operation, the slurry flow would have been in the laminar regime and resulted in solids clogging over time. However, because slurry flowed intermittently with about 1 m/s linear velocity, the process ran a total of more than 400 h without

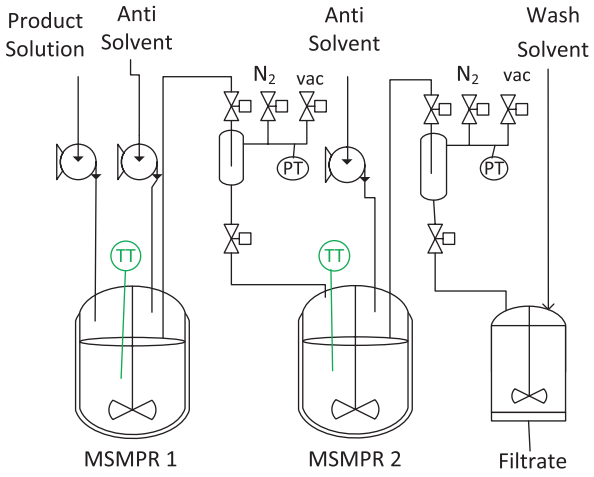

Fig. 4.1 Sketch of 2 MSMPRs in series with intermittent flow slurry transfer zones

fouling or clogging in the transfer tubes. The process was stopped each weekend. The longest individual operation time without stopping was 92 h. At the end of the 92 h there was no sign of clogging or fouling in any of the process tubing. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 5 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 30 min. Slurry slug volume between the MSMPRs was 0.3 L, and slurry slug volume between MSMPR2 and filter was 3.6 L. Please see Table 12 in the publication for more details on the intermittent slurry flow out of each MSMPR (Johnson et al. 2012). The filter cake was automatically intermittently washed after each transfer of slurry to the filter to keep the undesired enantiomer from precipitating in the filter cake. Two four-valve transfer zones designed for intermittently pumping slurries out of MSMPR crystallization vessels are shown in Fig. 4.1. Each four-valve transfer zone operated with a repeating sequence controlled by the automation system. Consider the transfer zone between MSMPR 1 and MSMPR 2 in the figure. The vacuum valve opened until pressure in the zone reached about 300 mm Hg, then it closed. The inlet valve opened to pull slurry at about 1 m/s linear velocity from MSMPR 1 into the transfer zone, then it closed. Slurry pulled out of MSMPR1 to decrease slurry level in the vessel to the dip tube, which was positioned at a height corresponding to 5.3 L remaining in the vessel. The nitrogen valve opened until the zone was pressured up to about 2 bar, then it closed. Finally, the outlet valve opened to push slurry into MSMPR 2 at about 1 m/s linear velocity, then it closed. Total time for this sequence was about 40 s. The sequence repeated automatically once every 5 min, controlled by the deltaV distributed control system (DCS). A similar automated sequence was used to transfer slurry to the filters once every 30 min. Similar to MSMPR1, slurry pulled out of MSMPR2 to empty the vessel to dip tube level, which was 8.9 L. Therefore, level sensors were not needed for automated level control in either of the crystallization vessels. A picture of the 8 L transfer zone for transferring slurry from MSMPR2 to the filter is shown in Fig. 4.2.

It is important that the automated valves have approximately the same internal diameter as the perfluoroalkoxy (PFA) tubing, minimizing any constrictions or expansions. In this example, the process tubing and the valve inside diameter were about 1 cm. Constrictions or expansions are typical locations for solids to build up and eventually clog. Automated ball valves with internal diameter as close as possible to the internal diameter of the PFA tubing work well. Figure 4.2 shows tubing and valves with internal diameter about 1 cm. It is also important to avoid elbows or any other types of fittings with sharp bends, because these are common locations for solids fouling. It is best if the PFA tubing flows straight into a two-way ball valve and straight out the other side with no bends. Valve orientation is also important because of the impact of gravity. If the valve is mounted vertically at the bottom outlet of the transfer zone, then solids can settle on the ball valve during the time that the transfer zone is filling and the slurry is accumulating in the zone. As seen in Fig. 4.2, the outlet valve from the bottom of the transfer zone was mounted horizontally and at a higher elevation than the bottom outlet. PFA tubing at the transfer zone outlet bends smoothly 270° from the bottom outlet up, around, and into the side of the outlet valve. This way, if solids settle by gravity in the transfer zone, then they

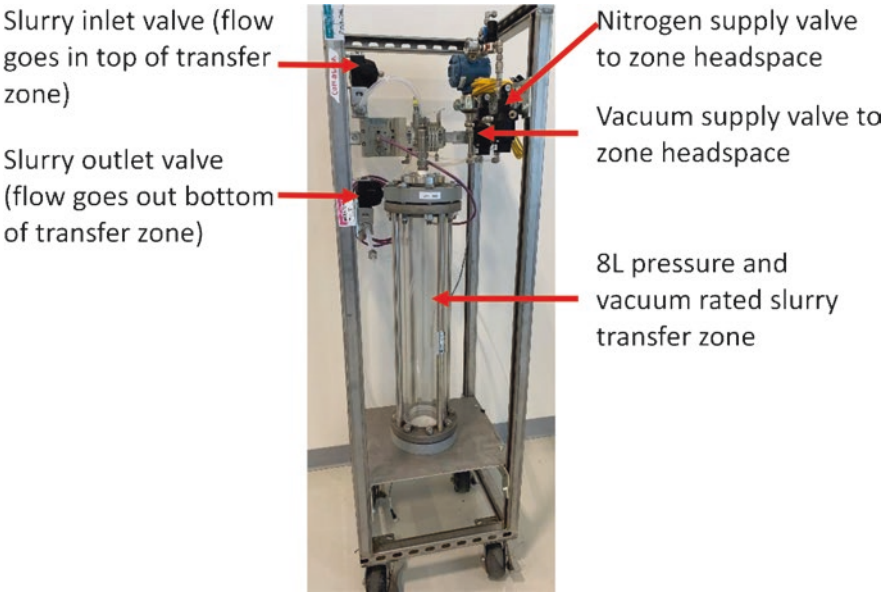

Fig. 4.2 Picture of 8 L four-valve transfer zone for transferring slurry from MSMPR2 to a filter

settle down into the smoothly curving part of the PFA tubing, so that when flows commence the solids will have less opportunity to hang up in the tubing.

Continuous crystallization was used for impurity rejection in the middle of a multistep continuous process for merestinib (Reizman et al. 2019). A picture of the 50 L glass MSMPRs with overhead stirring and custom baffle cages inserted is shown in Fig. 4.3.

One of the main benefits of running this process continuously was that the product was a highly potent compound. Small-volume continuous allowed the process to run in laboratory fume hoods at commercial manufacturing scale. Similar to the previous example, the crystallization was driven by cooling and anti-solvent. However, this was primarily a cooling crystallization, and the feed was kept heated at 50 °C to maintain solubility into the first MSMPR-in-series. Two feeds continuously flowed into the first MSMPR, the pharmaceutical compound dissolved in THF, and cyclohexane anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 15 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 15 min as well. Overall average slurry throughput was 179 mL/min, which would have been in the laminar flow regime if flows out of each MSMPR were truly continuous. However, 2.7 L slurry transferred quickly by turbulent flow once every 15 min, with linear velocity about 1 m/s, through 1 cm inside diameter PFA tubing. The end result was that the GMP manufacturing campaign ran for 17 days with no stopping and no clogging or fouling in the transfer tubes. The campaign produced more than 180 kg of API. Please see the publication for more details (Reizman et al. 2019). Like in the previous example, slurry flowed

Fig. 4.3 Picture of the 50 L MSMPRs

from MSMPR2 to the dual filters using an 8 L four-valve transfer zone, as described and shown in Figs. 4.1 and 4.2. However, in this example, slurry flowed from MSMPR1 to MSMPR2 by a different mechanism. Vacuum was temporarily applied to MSMPR2, which pulled slurry from MSMPR1. Between the MSMPRs was only a simple 1 cm *i.d.* curved PFA tube that arched up and over from one vessel to the next. During the transfer time, an automated block valve closed the vent from MSMPR2, a second automated block valve opened the vacuum supply to MSMPR2, and a third automated block valve opened an extra nitrogen supply to the headspace of MSMPR1 so that it remained at atmospheric pressure during the transfer and did not suck back from the vent bubbler. In this manner, slurry level in MSMPR1 was reduced to dip tube level once every 15 min. This method is generally preferred for cooling crystallizations, because the time for slurry to travel from MSMPR1 to MSMPR2 is only about 2 s; thus, it does not have time to cool and precipitate in between the vessels. Figure 4.4 shows a picture of the skid with automated block valves that accomplished the transfer from MSMPR1 to MSMPR2. The skid was designed and constructed to control slurry flows for up to four MSMPRs-in-series, which is why the extra valves, transmitters, and vacuum pots are seen in the picture.

A detailed description of the automated sequence used for this type of intermittent transfer, and a design sketch of the apparatus, is given on pages S9 through S11 in the supporting information section of a publication (Kopach et al. 2016).

Continuous crystallization was used for impurity rejection and isolation of a cytotoxic API in a multistep continuous process for tasisulam (White et al. 2012). One of the main benefits of continuous processing for this product was that the entire cytotoxic segment of the synthetic route was run in laboratory fume hoods in

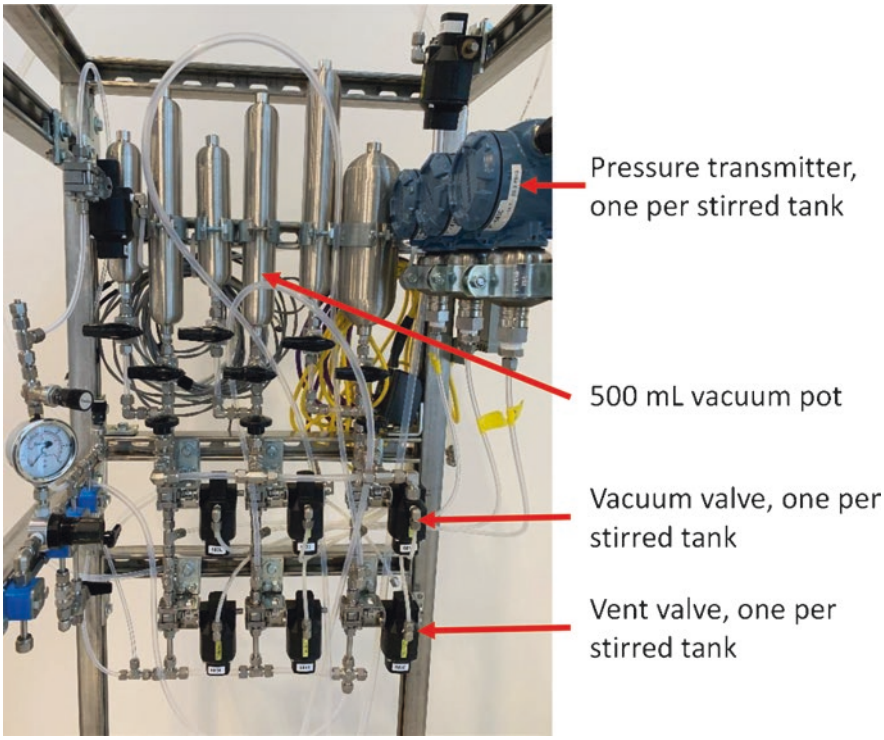

Fig. 4.4 Automated skid with sequenced block valves that transferred slurry from MSMPR1 to MSMPR2

disposable, inexpensive flow equipment. Two MSMPRs-in-series were used for the continuous crystallization. Two feeds were continuously pumped into the first MSMPR. One was the desired compound dissolved in 60/40 isopropyl acetate and isopropanol. The second was heptane anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 3 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 30 min. Before scaling up to production, the crystallization was operated at research scale. Overall average slurry flow at research scale was 0.6 mL/min. PFA tubing with 3 mm inside diameter was used for slurry flow. Therefore, the slurry flow at research scale would have had about 0.001 m/s linear velocity. Solids would have precipitated in the tubing and clogged if flow was truly continuous. However, because of intermittent flow, linear velocity in the 3 mm *i.d.* tubing was actually about 0.1 m/s in the research scale continuous crystallization experiments, which prevented solids clogging.

Continuous crystallization was used for impurity rejection in the middle of a multistep continuous process for prexasertib (Cole et al. 2017a). Crystallization was important to impurity control strategy, necessary to reject pyrazine and regio isomers downstream from a continuous S_NAr reaction. Purity of the crystallized solids was in excess of 99.8 area% throughout the campaign. There were many benefits of

continuous processing for this product, as described in the publication. One of the main benefits was containment of the entire process in laboratory fume hoods, which was important because the product was cytotoxic. Two feeds continuously flowed into the first MSMPR, the pharmaceutical compound dissolved in DMSO, and methanol anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 3 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 6 min. The next section describes the importance of the dual alternating filters for eliminating solids handling.

4.1.2 Filtration Downstream from Continuous Crystallization

In the prexasertib example, continuous crystallization flowed into semi-continuous filtration. Semi-continuous filtration eliminated manual handling of the genotoxic and cytotoxic intermediate. Therefore, the impurity rejection benefits of crystallization were realized, without the manual handling of solids that would typically be required in batch. The automated dual filters switched back and forth once every hour. The on-line filter received slurry slugs from the crystallizer, while the off-line filter was automatically rinsed with methanol, washed with MTBE, and dried with nitrogen. Then, the product was automatically dissolved off the filter in formic acid. Formic acid was the solvent and reagent for the downstream BOC deprotection continuous reaction. A simplified sketch of the automated intermittent flow dual filters in shown in Fig. 4.5.

A detailed step-by-step description of the automated dual filter process is given in the supplementary information for the publication (Cole et al. 2017a). A picture of the dual filter skid is shown in Fig. 4.6.

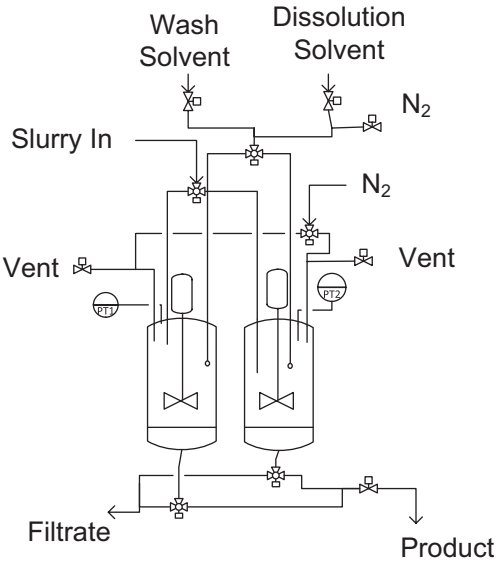

Fig. 4.5 Sketch of the automated intermittent flow dual filters

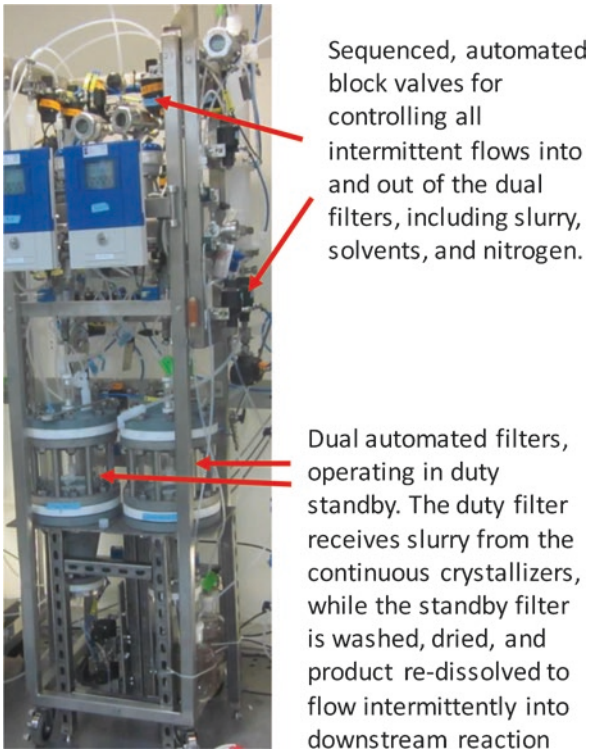

Fig. 4.6 Picture of the automated intermittent flow dual filters

4.1.3 Solvent Exchange Distillation with Strip to Dryness in Rotary Evaporators

Intermittent flow also enables efficient solvent exchanges in the middle of an otherwise fully continuous process train. Solvent exchange from high boiling point solvent to low boiling point solvent is much more efficient with stripping to dryness and then adding back the lower boiling solvent. This is not possible in a typical batch vessel in a manufacturing plant because of the need to maintain a minimum stir-able volume, but strip to dryness solvent exchange is possible using intermittent flow 20–50 L laboratory rotary evaporators, which can give commercial scale throughput for small-volume continuous processes. The automation charges product solution, then strips off the first solvent, then empties the distillate, then adds back the second solvent dissolving the product, and then empties. The DCS automated control system repeats the sequence, typically about once per hour, for the entire duration of the continuous campaign. It is not truly continuous, it is really automated repeating batch, but it is practically continuous, and it is an effective way to incorporate solvent exchange into the middle of an otherwise continuous process train in laboratory fume hoods.

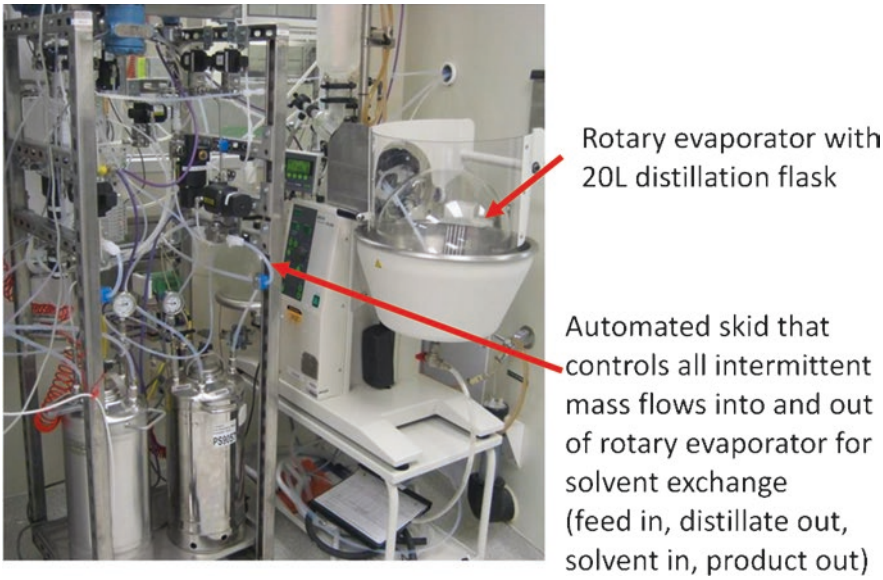

Fig. 4.7 Picture of automated intermittent solvent exchange distillation equipment

Automated intermittent solvent exchange with strip to dryness was used in the continuous process steps for tasisulam (White et al. 2012). This was a particularly challenging solvent exchange. Desired product was dissolved in a mixture of toluene, methyl THF, and residual water after a continuous Schotten-Baumann reaction and liquid–liquid extraction. The goal was to solvent exchange into a precise ratio of 60/40 isopropyl acetate and isopropanol. Residual solvent target was <0.1% water and <0.5% toluene. Furthermore, achieving the precise ratio of isopropyl acetate to isopropanol was important for the subsequent continuous crystallization. These are challenging requirements for either a batch solvent exchange distillation or a truly continuous solvent exchange distillation. However, the automated intermittent rotary evaporator with strip to dryness achieved the solvent exchange goals with ease. The solvent exchange system that was used to process 20 kg cytotoxic API at 5 kg/day throughput is shown in Fig. 4.7.

A simplified sketch of the apparatus is shown in Fig. 4.8.

Solids were present temporarily in the rotary evaporator during the solvent strip, but the solids were redissolved prior to product solution flowing out of the evaporator.

Automated intermittent solvent exchange with strip to dryness was also used in the continuous process steps for prexasertib (Cole et al. 2017a). In this example, a formate salt isolation of a cytotoxic intermediate was eliminated by using stripping to remove formic acid to less than 0.8 equivalents relative to the API. This was not feasible by either batch distillation or truly continuous distillation. Batch distillation could not accomplish the formic acid removal because the extended times required for distillation resulted in impurity formation. Truly continuous distillation was not feasible because of the solid precipitate that forms when formic acid is stripped away. The intermittent flow rotary evaporator was able to achieve formic acid removal

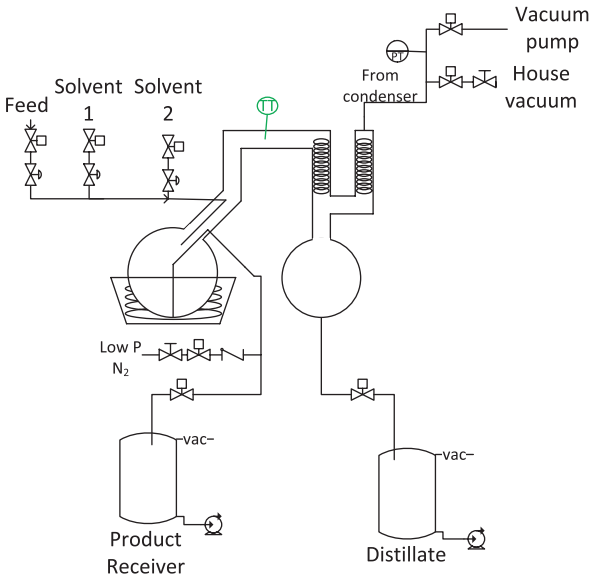

Fig. 4.8 Sketch of automated intermittent solvent exchange distillation equipment

without significant impurity formation because of the ability to strip to a dry thin film. The automated sequence charged formic acid solution of the API, then charged lactic acid, then stripped to dryness, then charged water, then stripped to dryness again, emptied the distillate, then charged water and THF, dissolved the API, and emptied the product solution. A detailed step-by-step description of the automated intermittent evaporator sequence is given in the supplementary information for the publication (Cole et al. 2017a). Figure 4.9 shows the front side of the automated skid that controlled all the flows into and out of the rotary evaporator and controlled the automated pressure profile. Figure 4.10 shows the back side of the unit. From the back side, the Coriolis mass flowmeters are seen, which were used by the DCS system to achieve accurate and precise mass charges of each intermittent flow.

4.1.4 Back Pressure Regulators and Vapor–Liquid Separators Downstream from Continuous High-Pressure Hydrogenation Reactors

Continuous reaction was used for an Ru-catalyzed direct asymmetric reductive amination (DARA) reaction for producing an API intermediate. The reaction operated with 68 bar hydrogen pressure. Continuous processing offered safety, throughput, and capital cost advantages compared to batch high-pressure hydrogenation. Intermittent flow was used at the exit of the continuous reactor to depressurize reactor contents and forward reaction product material to a surge vessel. First, the process was developed at research scale with about 0.2 kg/day throughput, then it was demonstrated at pilot scale with about 0.7 kg/day throughput, and finally it was

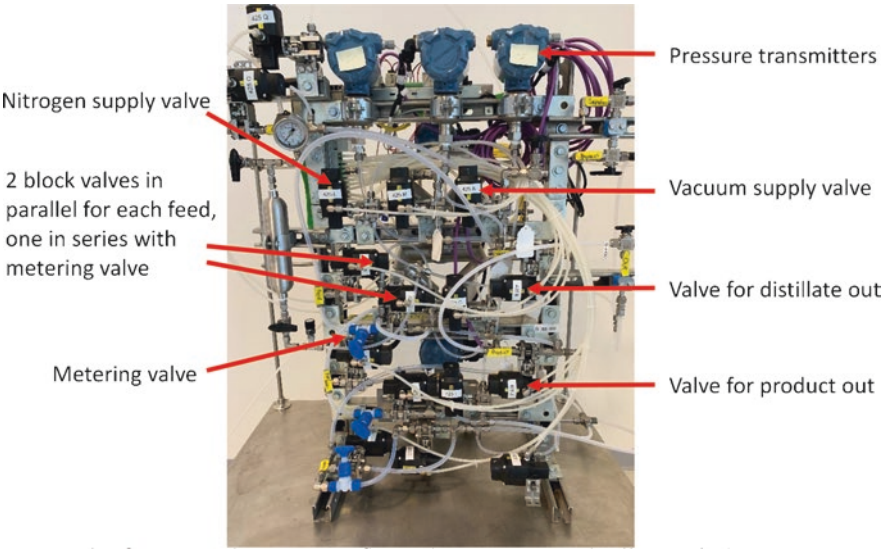

Fig. 4.9 Front side of automated intermittent flow solvent exchange distillation skid

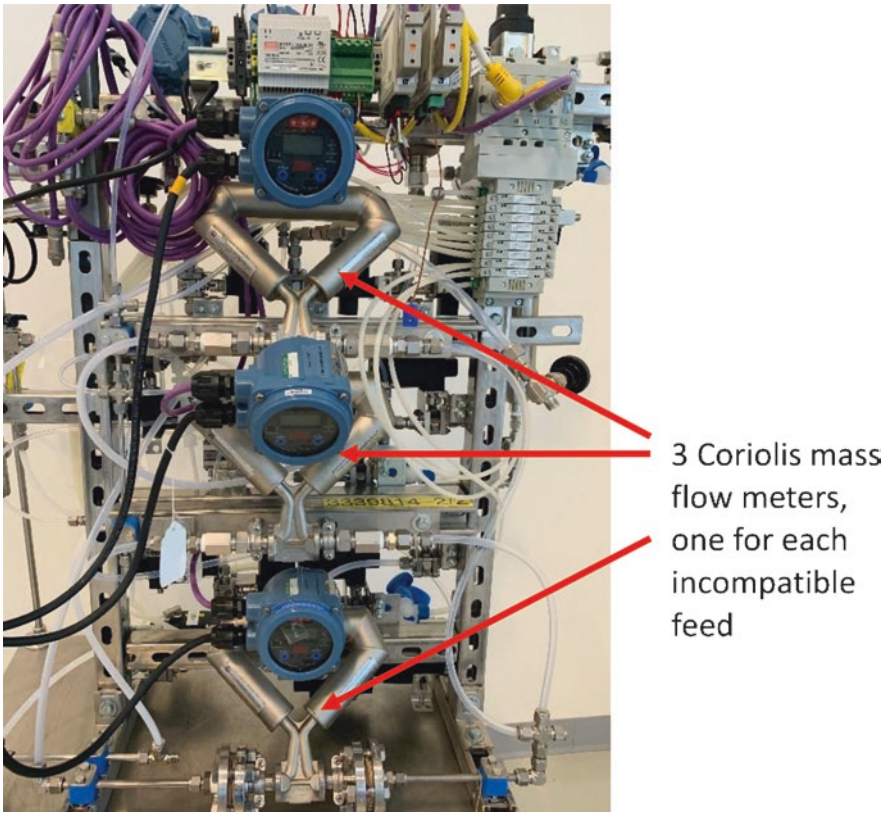

Fig. 4.10 Back side of automated intermittent flow solvent exchange distillation skid

scaled up to manufacturing for a 3000 kg GMP validation campaign with throughput about 100 kg/day (Johnson et al. 2016; Changi et al. 2017). Intermittent flow at the reactor exit was used at all three scales. The main advantage of intermittent flow at the PFR outlet was that it tolerated some amount of solids precipitate in the back pressure regulation and gas/liquid separation section, without solids clogging or fouling. A detailed description of the design, automation, and operation of the intermittent flow back pressure regulator is given on pages S14 to S25 of the supporting information to the publication (Johnson et al. 2016). Excess gas reagent and product solution temporarily pooled in the first of a series of pressurized vessels. This vessel operated at pressure equal to reactor outlet pressure. Intermittently, once every 14.5 min, the reaction product solution was transferred through expansion vessels in series by the opening and closing of sequenced automated block valves between the vessels. The pressure trends in the reactor as a result of this mode of operation are shown in the supporting information section of the publication. This type of back pressure regulation is an alternative to a standard restricting orifice back pressure regulator that would maintain truly continuous flow at the exit of the reactor. For example, a pressurized diaphragm dome style back pressure regulator, a spring-loaded back pressure regulator, or a small orifice automated metering valve could be used as an alternative. The traditional commercially available style back pressure regulator would also maintain constant pressure within the reactor, rather than the deliberate oscillating pressure swings. However, these have restricting orifices that can clog and foul with solid particles. In contrast, the intermittent flow approach using the expansion chambers in series and the automated block valves does not have any restricting orifices. When an automated block valve is opened to allow intermittent flow from one chamber to the next, the flow path has large diameter and material transports at extremely high linear velocities. At production scale, the valves have at least a 1 cm inside diameter flow path. The end result is that the intermittent flow back pressure regulation system can run for longer times and tolerate small amounts of solids without clogging or fouling. Figure 4.11 shows a research-scale intermittent flow back pressure regulation system, and Fig. 4.12 shows a pilot-scale intermittent flow back pressure regulation system.

The pressure pots and automated block valves are stainless steel. They are pressure rated to more than 100 bar, although they are typically used at pressures 70 bar or less.

The same type of intermittent flow back pressure regulator was used for the hydroformylation example given in another chapter in this book (Johnson et al. 2020a); 50/50 H_2/CO was the reagent gas mixture, and the reactor operated at 68 bar pressure. Continuous reaction had safety and capital cost advantages compared to batch. In that example, catalyst particles precipitated because the catalyst/ligand was not soluble in the mixed aldehyde product. In addition, some of the methylmethacrylate polymerized, forming solids in the reactor. The slurry exiting the reactor contained about 1 wt.% solids, much of which was sticky polymer. Nevertheless, the reactor operated continuously for the entire 314 h continuous campaign without any signs of fouling or clogging at the outlet of the reactor through the back pressure regulation section, depressurization, and gas/liquid separation.

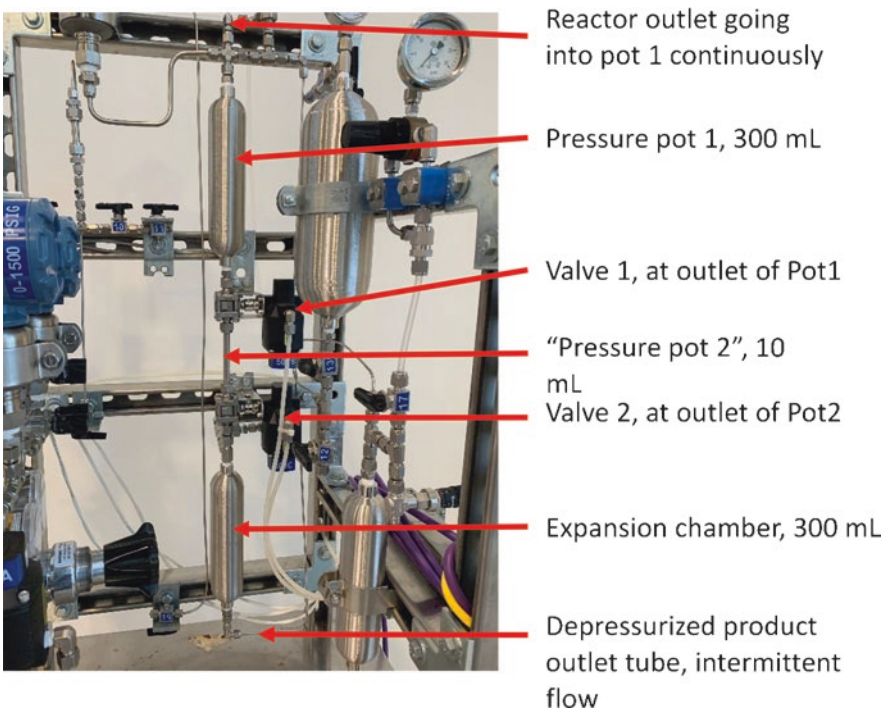

Fig. 4.11 Research-scale intermittent flow back pressure regulation system

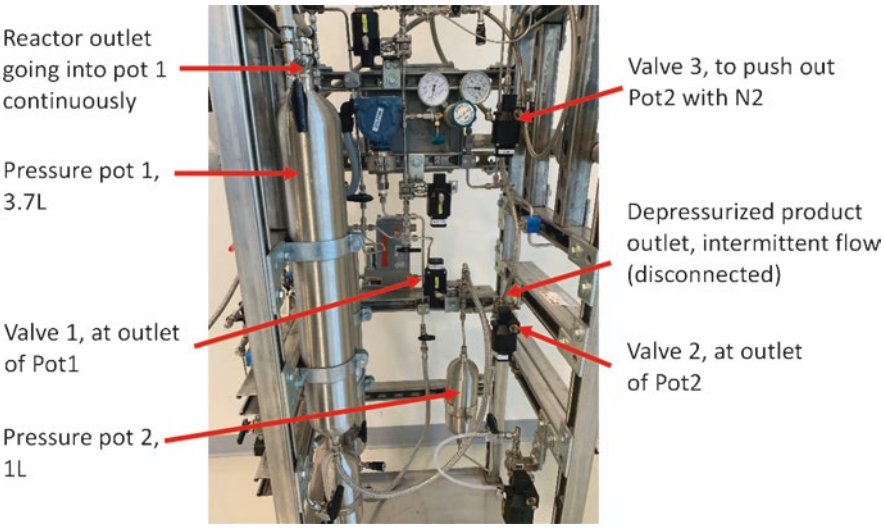

Fig. 4.12 Pilot-scale intermittent flow back pressure regulation system

Intermittent flow was used at the exit of a packed catalyst bed reactor because of catalyst fines that exit the reactor during the first several hours of start-up with each new catalyst bed. A hydrogenolysis reaction was run in the continuous packed bed reactor because it minimized formation of a desF impurity compared to batch. For details about the chemistry and the reactor design please see the publication (Zaborenko et al. 2015). If a traditional flow-restricting orifice back pressure regulation device had been used, it would have fouled because of the small catalyst particles. Filtration of the tiny catalyst particles was not an option because of the buildup of high pressure drop across the filter, restricting flow. The intermittent flow back pressure regulation at the exit of the packed catalyst bed reactor was a more reliable approach because it operated for 92 h without any signs of fouling. Material of construction was hastelloy C276 rather than stainless steel because of the HCl used in the hydrogenolysis reaction.

Intermittent flow was used downstream from a high-pressure reductive amination reactor in order to facilitate gas/liquid separation and hydrogen stripping from the liquid product solution. The manufacturing plant produced 2000 kg of product in a GMP registration stability campaign. The continuous reactor ran with 50 bar hydrogen pressure. The reasons for continuous processing, the reaction process, and automated equipment are described in detail in the publication (May et al. 2016). The main safety advantage of running this high-pressure hydrogenation reaction continuously was that the hydrogen supply, reactor, back pressure regulator, gas liquid separator, and gas stripping were all located outside the building. Reaction product solution flowing back into the building was practically free of hydrogen. This is something that is feasible continuous but is not feasible batch, because batch autoclaves must be opened for charging materials. They must be inside the building in a controlled bunker. In contrast, the continuous reactor is always sealed; therefore, it can be located outside. Detailed designs, description of automated sequences, and pressure trends are given on pages S25 through S37 in the supporting information of a different publication (Johnson et al. 2016). This gas–liquid reaction was relatively clean, meaning that solids precipitates were not expected, only vapor and liquid phases. Therefore, a dome diaphragm style restricting orifice back pressure regulator was used to maintain constant back pressure of 50 bar at the exit of the reactor. However, immediately downstream from the back pressure regulator, intermittent flow was used to separate excess hydrogen gas from the liquid product solution without misting out the vent. The intermittent flow chambers in series stripped hydrogen from the solution with nitrogen and forward material intermittently to a surge vessel downstream once every 19.5 min. Intermittent flow chambers were efficient at gas stripping because of the ability to do pressure purges.

4.1.5 Stirred Tank Reactors for Heterogeneous Reactions

A separate chapter in this book described an intermittent stirred tank reactor used for a Suzuki-Miyaura cross-coupling reaction in the merestinib continuous process (Johnson et al. 2020a; Cole et al. 2016, 2019). As described in that example, the

intermittent operation of the fill–empty reactor enabled full conversion in just 20 min, while it would have required much longer residence times in a CSTR or CSTRs-in-series. Furthermore, intermittent flow allowed the catalyst to be added last at a controlled flow rate, which improved catalyst activity. A 6 L stirred tank was used for a 100 hour continuous run to make 22 kg of product. The reactor turned over automatically about 45 times per day, therefore reactor volume was about 50–100 times smaller than what would have been required in traditional batch processing. It was not a truly continuous reaction, rather it was automated repeating batch, but it was practically continuous, and it integrated well with an otherwise fully continuous process in laboratory fume hoods.

An intermittent flow stirred tank reactor was used for a thermal cyclization reaction in which the product precipitated in the reactor (White et al. 2014). This would have clogged a plug flow tubular reactor. A truly continuous stirred tank reactor was not used because it would have required a much longer reaction time to reach the same conversion. The reactor operated under elevated pressure, therefore filling and emptying the reactor was more feasible with intermittent flow than truly continuous flow. The benefit of intermittent flow compared to batch was that the reaction ran at extremely high temperatures and pressures, beyond the capabilities of most batch reactors. The reactor turned over once every 8 min, therefore it was more than 100 times smaller than a batch reactor operating once per day. Temperature was 265 °C and pressure was 41 bar; 300 g of advanced intermediate was produced with 700 automated turnovers of a 25 mL autoclave.

An intermittent flow stirred tank reactor with catalyst recycle was used for an enantio-selective Aza-Henry reaction (Tsukanov et al. 2016). The reagents were homogeneous solutions, but the product was insoluble in the reactor. The process was more efficient and the reaction time was shorter when the catalyst and the excess reagent were recycled. A safety benefit of intermittent flow compared to batch for this reaction was that the intermittent flow reactor was smaller, reducing the amount of nitro alkane in the system. The reactor turned over once every 40 min, and 25 g product was made with 16 automated turnovers of a 250 mL reactor.

An intermittent flow stirred tank reactor was used for a nitro group reduction (Cole et al. 2017b). A trickle bed hydrogenation reactor with fixed catalyst packed into a column could have been used as a truly continuous plug flow reactor for this type of reaction. However, a trickle bed reactor requires a larger catalyst particle size than what is used batch, because small particle size catalysts create excessive pressure drop as the reaction solution and hydrogen flow through the column. Therefore, a benefit of the intermittent flow stirred tank compared to trickle bed for this chemistry was that the same catalyst that was developed for the batch reaction could be used for the intermittent flow reaction. Another benefit of the intermittent flow reaction compared to traditional batch hydrogenation was that the reactor was about 2 orders of magnitude smaller for the same overall process throughput. This is a significant safety advantage because it reduces the amount of hydrogen in the system. Moreover, the 1 L hydrogenation reactor was approved for operation in a laboratory fume hood, while a 100 L batch hydrogenation reactor surely would have needed a specially designed hydrogenation bunker, which is not available to most research labs. The small particle size heterogeneous catalyst was sequestered in the reactor for multiple turnovers. The homogeneous solution reagents and hydrogen

flowed into the reactor, and products flowed out of the reactor, but the solid catalyst particles were maintained within the vessel by filtering in situ. Overall catalyst used was less than the batch process, but the molar equivalent of catalyst in the reactor was higher compared batch because the catalyst was sequestered for about 25–50 reactor volume turnovers. Average time per reactor intermittent flow in and out of the reactor was 17 min; 1.9 kg advanced intermediate was produced with 107 automated turnovers of a 1 L autoclave reactor. Furthermore, part of the campaign was run with less feed solvent which resulted in a slurry feed. The intermittent flow stirred tank hydrogenation reactor enabled the slurry feed, which would not have been possible with a truly continuous packed catalyst bed reactor.

In general, intermittent flow stirred tank reactors (automated repeating batch) are similar to truly continuous reactors in many ways. Fast heat-up and fast cool-down is achieved in heat exchangers, while material flows in and flows out of the reactor. Intermittent reactors have higher heat transfer A/V compared to standard batch. Heat-up time, cool-down time, and time at reaction temperature is more scalable than standard batch. Intermittent reactors remain at approximately constant temperature and pressure at all times. They achieve a large number of turnovers per day. They are small compared to batch for the same throughput. If reaction time is 30 min, then the intermittent stirred tank is about 100 times smaller reactor volume compared to batch (assuming 48 h start to start cycle time for batch campaigning). Intermittent reactors are small enough so that they fit in lab hoods for 5–10 kg/day processes. The reagent feed tanks gradually empty and the product tanks gradually fill over time. These should be switched about once per day or once per shift.

In addition, there are many benefits of intermittent flow stirred tank reactors compared to PFRs. They can handle heterogeneous reactions with solids in flow, and two-phase liquid–liquid with long reaction times (τ longer than 10 min). They facilitate a much wider range of reagent addition strategies, for example, all-at-once addition, controlled addition of one or more feeds, any order of addition of multiple feeds, or co-addition, depending on which gives higher yield and/or minimizes key impurities (PFR is mainly all-at-once stoichiometric addition at the reactor inlet). They can run closer to end of reaction conditions, for example, remove and add back 10% of reactor volume each cycle, if it benefits impurity profile.

Furthermore, there are obvious benefits of intermittent flow stirred tank reactors compared to true CSTRs. They require much less reaction time for the same conversion, and thus smaller reactor volumes, for positive order reactions. They can operate with all-at-once reagent addition, controlled addition of one or more reagents, or co-addition, depending on which gives higher yield and/or minimizes key impurities (CSTR is only co-addition).

4.2 Recycle

There are many potential benefits of incorporating recycle into a continuous process. Continuous reaction with recycle of unreacted reagents after a downstream separation step can improve overall yield and selectivity of some chemical transformations. Late forming impurities can be avoided, while still maintaining high yield,

by running at incomplete conversion and recycling reagents back to the reactor inlet (White et al. 2014). For equilibrium limited reactions, recycle can be used to achieve higher overall conversion by removing a product or by-product and recycling reagents to drive the reaction forward. If there is a selectivity advantage of partial conversion, then recycle keeps the reaction at the conversion where the ratio of desired to undesired product is highest, yet maintains high yield by recycle of reagents. If there is a benefit of high stoichiometric ratio of reagents in a continuous reactor, then recycle can be used to achieve high relative stoichiometry in the reactor by separating downstream and recycling the reagent used in excess. This can be accomplished without using a large excess of reagent overall. Recycle can be used to achieve lower overall catalyst loading but higher instantaneous catalyst loading by recycling a homogeneous catalyst (Tsukanov et al. 2016). Finally, recycle is well known to achieve lower process mass intensity by recycling solvent, and achieve higher overall yield in by recycling product from filtrate. Process Mass Intensity (PMI) is defined as mass waste generated divided by mass product. Recycle could be one of the most powerful aspects of continuous processing compared to batch. Chemical engineering texts on material and energy balances often teach recycle incorporated into process flowsheets for continuous processes (Felder and Rousseau 1986). It is the key to achieving higher yield at the same purity, or the same yield with higher purity, and minimizing waste, compared to batch.

4.3 Common Misconception About Needing Fast Reactions in PFRs

Kinetics do not necessarily need to be fast in order for a reaction to be a viable candidate for continuous processing. It is common practice to speed up reaction rates by operating at elevated temperatures in continuous reactors; however, there may be a trade-off between accelerating reactions and thermal stability of reagents and products. Reaction selectivity and impurity profile often suffer at higher temperatures.

Table 4.1 lists continuous reactions that have been scaled up by Eli Lilly and Company to pilot and plant scale, most of them in GMP production. Compared to what is seen in the majority of literature on continuous chemistry, the examples listed in the table show long PFR reaction times (0.7–24 h) and large PFR reactor volumes (3–360 L). The table lists the impurity issues that become more significant when the reaction is run at higher temperatures. As seen in the last column in the table, isomeric impurities, dimers, chiral impurities, degradation of product, and other impurities may result, if reaction temperature is increased in order to reduce τ.

The imidazole cyclization with 90 min τ (May et al. 2012) and the hydrazine addition with 90 min τ (Cole et al. 2017a) were both run at moderate temperature to minimize deprotection of product (Table 4.1). The thermal EE deprotection with 40 min τ (Reizman et al. 2019; Cole et al. 2019; Frederick et al. 2015) was run at moderate temperature to avoid hydrolysis of an amide bond. The S_NAr reaction was

Table 4.1 Continuous reactions with long τ that have been scaled up in PFRs by Eli Lilly and Company to pilot- and plant-scale production

Reaction in PFR	τ	Material produced	PFR vol.	T	P	Issues at higher temperatures (and shorter τ)
Imidazole cyclization	1.5 h	29 kg GMP	7 L	140 °C	69 bar	Deprotection of product
Thermal deprotection of ee group	0.7 h	183 kg GMP	7 L	170 °C	25 bar	Hydrolysis of an amide bond
Hydrazine addition	1.5 h	26 kg GMP	1.5 L	130 °C	20 bar	Deprotection of product
S_NAr	3 h	31 kg GMP	3 L	70 °C	1 bar	Dimer
Acid deprotection of boc group	4 h	24 kg GMP	12 L	25 °C	1 bar	t-butyl amide
Hydroformylation, Rh catalyst	24 h	178 kg	32 L	55 °C	70 bar	Lower isomer selectivity
Asymmetric hydrogenation, Rh catalyst	12 h	144 kg	73 L	70 °C	70 bar	Lower ee
Reductive amination, Ir catalyst	12 h	2000 kg GMP	380 L	25 °C	55 bar	Cis isomer
Asymmetric reductive amination, Ru catalyst	5 h	3000 kg GMP	200 L	130 °C	60 bar	Dimer

run at moderate temperature and 3 h τ (Cole et al. 2017a), and a direct asymmetric reductive amination was run at moderate temperatures and 5 h τ, to minimize dimer impurities (Changi et al. 2017). The BOC deprotection was run at mild temperature and 4 h τ to minimize a t-butyl amide impurity (Cole et al. 2017a). A hydroformylation was run at moderate temperature and 24 h τ to minimize formation of the undesired linear aldehyde isomer (Johnson et al. 2020a). An asymmetric hydrogenation was run at moderate temperature and 12 h τ to minimize an undesired enantiomer (Johnson et al. 2012). A reductive amination was run at moderate temperature and 12 h τ to minimize the cis isomer (May et al. 2016). These all lead to longer reaction times, larger required reactor volumes, and less need for fast mass transfer and heat transfer rates.

The last four entries in the table are homogeneously catalyzed reactions where there are significant cost, quality, and environmental reasons for minimizing the amount of catalyst/ligand, which can represent a significant part of the overall manufacturing cost. The loading of Rh, Ir, Ru is minimized for environmental stewardship, and also for reducing levels of these metals in the product for toxicology reasons. It is vitally important to sufficiently remove metals from pharmaceutical products, and this can require additional separation steps. The best option is to put less metals in the process in the first place by using high substrate-to-catalyst ratios, which results in long (>12 h) reaction times and large reactor volumes, for example, the 360 L PFR in GMP manufacturing for homogeneously catalyzed hydrogenation (May et al. 2016).

In addition, a practical reason for not accelerating the reactions by going to higher temperatures in some circumstances is process fit. When these reactions are integrated into a fully continuous process, it is best if the reactions can tolerate flow stoppages or throughput changes. This is usually more feasible at the milder temperatures and longer target τ. This way, if the process must stop for an extended time period and then restart, risk is lower for holding material in the reactor.

Whether the reason for longer τ is to minimize impurities, improve process fit, or reduce catalyst and ligand loading, the downside is the need for larger volume PFR reactors. However, PFRs are generally low cost and the trade-off is favorable to install larger reactors.

4.4 Continuous Process Checklist

Prior to running a continuous process demonstration, the following questions should be answered.

Flows:

- Achieving and maintaining accurate and precise mass flow rates for each of your continuous feeds is one of the most critical aspects of flow chemistry. Do not start a continuous process unless you know that your mass flow rates will be accurate, consistent, and quantifiable. How do you know that your mass flow rates will be correct? Do catch and weighs. This means that you collect the liquid in a tared container for a measured amount of time, weigh the mass of liquid collected, then use the information to calculate mass flow rate. This can also be done if feed solutions and product solution are continuously collected on data logging balances, by calculating change in mass versus time. Design your system in a way that you will have redundant measures of mass or volumetric flow rates, so that you get double-checks on mass flowmeters. For example, the primary measurement is performed using a Coriolis mass flowmeter (because this is more useful for feedback control), and the secondary measurement is change in feed vessel mass or level versus time. Calculate real-time mass balances to make sure they remain 95–105%, for example, mass balance for the previous hour, by looking at change in mass of all inputs, outputs, accumulation, and generation over a given time period.
- What if one or more of the pumps gives oscillating flow rates for example, every few minutes? Does the process dampen this out sufficiently at reactor inlets to tolerate the fluctuations? You may need to include a small mixing pot or stirred vessel at the inlet to a PFR.
- If this is a reaction with gas reagent (e.g., H_2), then how do you know the real reaction gas flow rate? Verify the mass flowmeter reading by monitoring gas supply cylinder pressure versus time.
- How do you know that your check valves are going to prevent backflow? They won't. Check valves fail. Use them only as a backup line of defense.

- Have you done pump testing? For example, if you plan to use a peristaltic pump for a solution with THF solvent at 60 °C, then can you set up a pump-around recirculation loop with a small amount of the feed in an inerted box with secondary containment, and pump the feed for a long duration to make sure the peristaltic pump tubing will hold up to the hot solvent? It is better to test failure limits like this with small amounts of material in safe environments.
- How will you know if the instantaneous liquid and gas flow rates are steady? If there are oscillations in the liquid or gas flow rates, then how will you know how much they oscillated?
- Are your tubing and piping line sizes correct? If this is a technology transfer from one site to another, are any line sizes different than where the process was previously run? Oversized tubing can result in carryover and dead zones, while undersized tubing or fittings can result in high pressure drop and plugging.
- Are there any process or vent lines for which the inner diameters are too small? Look through every inch of your process. Typical examples: You want size of tubing and fittings on suction side of pumps to be larger than discharge side of pumps, preferably 3/8″ or larger for 100 L/day scale. You want process tubing and fittings to be at least ¼″ even at smallest research scale if there is any potential for solids moving through the tubing. You want the aqueous overflow from gravity decanters to be at least ½″ until after the siphon break, even at the smallest research scale, because of the impact of water surface tension.

Feed solutions:

- What is the composition of all feed solutions prepared batch? This includes weights and volumes of all components, and also the density and molarity of each of the feed solutions.
- Can you prepare feeds that are all homogeneous solutions at room temperature? This greatly simplifies the complexity of the continuous process and improves reliability. If not, then try to change solvents and concentrations so that you have homogeneous solution feeds at room temperature. A small amount of effort here to obtain solubility can save hours of effort later resolving fouling/clogging issues. The alternatives are to feed slurries or hot solutions, which are both more complex and difficult than homogeneous solutions at room temperature.
- Is there any reason why composition in feed tanks would change over time? For example, could you be losing solvent or volatile reagent to evaporation out the vent? This could cause stoichiometry to drift over time and the process to go out of spec. Make sure that you do not have a live nitrogen sweep in the headspace of the feed tank.
- Is the feed solution in the feed tank well mixed? For example, if charging multiple cans or drums of feed solution to the feed vessel, are they well blended so that concentration is uniform throughout?
- Is the feed solution stable? What is the stability of starting reagent solutions over time? The goal is at least 1 week stability at feed tank temperature.
- Is the feed solution filtered? What if in-line filters are clogging during the continuous run? It is best to filter all homogeneous solution feeds before you start

pumping them especially if you are using pumps with small clearances like gear pumps. In-line filters should still be used at the inlet of gear pumps even if the feed was already filtered before it was in the feed tank. Use dual in-line filters with valves so that you can switch to the second filter and clean the first when it starts to clog or blind, and switch back and forth.

- If in-line filters are installed, how do you know the pressure drop across them so that you can swap them out before they foul? Do you have a pressure indicator between the filter and the pump inlet?
- Do you have enough tanks, vessels, and cans for feed solutions, surge, and product solutions? Can you avoid filling and emptying a vessel at the same time? Can you avoid frequent manual refills? It is best if refilling is not more frequent than once every 12 h if you are doing it manually.
- Will you have enough materials? Are you making up enough solution volume for all feeds including excess of some feeds in case you need to change stoichiometric ratio? Which feed do you want to run out first, that is, what is the limiting reagent? Preparation of a small excess of inexpensive feeds is recommended. You don't want to run out of another feed before your limiting reagent is used up completely at the end of the campaign.
- Is a solvent feed tank ready to go, and tubing, valves, and Tee connected, in case you want to do a solvent pushout in the middle of the campaign? Is there a switching valve in line and ready for the easy feed swap?
- If you are ordering feed solutions from a vendor rather than mixing the solutions yourself, and if the solution needs to stay inerted, then does the vendor transport vessel have the needed fittings, valves, and connections so that you can transfer the solution out of the shipping container and into your plant feed vessel in an inert fashion?

Process Parameters and Data Collection:

- How will you measure to determine the real reaction temperature? Do you have redundant temperature measurements? Is temperature measured on the process side or shell/jacket side? How will you know the location of the hot spot in the reactor? Actual reaction temperature can be difficult to know and measure. Temperature measurements at the outlet of a reactor or unit operation can be inaccurate because of ambient cooling, even if they are only inches away from the jacket. Therefore, do not measure temperature at the outlet and assume that it is the same as the temperature inside the reactor. For PFRs, you may be able to use the shell side temperature and call it reaction temperature, but you must know heat transfer coefficients, reaction energetics, and kinetics so that you can calculate hot spots. What in-process temperatures do you plan to monitor? What is the number and location of in-process thermocouples (at the mixing Tees and also anticipated reactor hot spots?)
- How do you know that the actual reaction temperature at the hot spot will be the same when you scale up in a PFR?
- What will be the difference in jacket temperature (research vs. production scale) to keep reaction temperature the same when you scale up in a CSTR? For exam-

ple, if the reactor is a CSTR, then the difference between jacket temperature and reactor internal temperature will be larger when scaling up, because heat transfer surface area per to volume ratio (A/V) decreases when you scale up.

- How will you measure to determine the real reaction pressure? Do you have redundant pressure measurements? What is the pressure at the reactor inlet and the reactor outlet? Do you anticipate pressure gradually drifting over time due to fouling? Will you know if pressure drift and fouling is happening?
- How will you measure to determine the real reaction τ? Can you measure start-up transition curve and get τ from the midpoint of the F-curve? Can you calculate τ from reactor volume, % liquid filled, liquid flow rate, and thermal expansion of the liquid? What will be the impact of off-gassing in the reactor on τ? How do you know if the PFR will be completely liquid filled? How do you know if the pressure is high enough to prevent a gas phase? If the process exceeds critical temperature, then does it also exceed critical pressure? Actual reaction τ can be difficult to know and measure. τ is best measured during startup F-curve transitions from solvent to steady state.
- What is the difference between τ and V/Q for the continuous reactors? For example, for thermal deprotection in THF at 150 °C, we should say, "V/Q = 120 minutes and τ = 100 minutes because of thermal expansion." V/Q is more practical to specify for the plant operating ticket, because Q is the flow rate set point for the pumps, and the reactor volume should be known. However, τ is more important because it is the real reactor residence time. For example, if we want to compare reaction time in batch to reaction time flow, or if we want to measure reaction kinetics in the flow reactor, then we must use τ, the real mean residence time.
- What is the acceptable operating range for all of the important process parameters, for example τ, temperature, solvent volumes? Can you do anything to widen the acceptable ranges, for example, run at lower temperature if it allows you to have wider operating windows for τ because it makes product more stable to end of reaction conditions?
- How do you know the real reactor volume? This may be different than manufacturer specified, and it may be difficult to measure for a PFR because of the difficulty to get it 100% liquid filled and then completely empty.
- Does the liquid flow in the bottom and out the top of the coiled tube reactor? This is especially important for larger than 4 mm inside diameter tubing.
- How will you measure to determine the real reaction liquid flow rate? Do you have a secondary check in addition to a mass flowmeter for each stream, for example, change in mass or liquid level in feed tanks over time? Can you do catch and weighs? Do you know density of feed solutions?
- How do you know all the feeds were inerted and the reactor was inerted? Is the system leak free? Prove it with leak tests and pressure/vent or pressure/vacuum purges.
- How do you know that you will be able to calculate mass balance? Do you have enough vessels on weigh scales, vessels with level transmitters, or mass flowmeters to account for all inputs and outputs? How will you know if accumulation is happening? Is the data being collected, so that you will be able to close mass

balances for each lot? Are you manually weighing feed drums going into the process and product and waste drums coming out before and after they are filled or emptied, along with start and stop times for each drum?

- How do you know that experimental data for all important operating parameters will be collected and preserved? What is logged to the distributed control system (DCS) data historian (if available) and what is not? How will you retrieve the data? What do you need to write down hourly on checklists because it is not logged automatically? You can make a table of readings for the operator to write down at a specified frequency. How will that information be preserved?
- Is your numerical model predictive? What perturbations or step changes can be done during the continuous run to test if the model is predictive?
- What information are you planning to document after the scale-up continuous run? Are you set up to collect all of this information?
- Is there video of the entire process? If not can you get a video of the entire process during the demonstration?
- Do you expect anything to change gradually over time, like fixed catalyst bed activity? If so, do you have a planned frequency of change-out?
- Does anything need preconditioning or seasoning, like a new packed catalyst bed or metal walls of a reactor to be used for a catalytic reaction?

Scale-up

- What is the appropriate scale-up factor for a production campaign? This depends on intended duration of flow campaign. In this campaign, is it important to demonstrate longer time periods, for example, if using a packed catalyst bed reactor or running a continuous crystallization with long-term encrustation potential, or is it more important to generate product as quickly as possible?
- Are your process flowsheets and spreadsheet calculations representative for scale-up? Do you have a scale-up spreadsheet with calculations of all flow rates, volumes, masses? Or are you using a flowsheeting program?
- Have you done the engineering calculations on heat transfer rates, sizing heat exchangers, mass transfer rates, sizing mixers, line sizes, vessel sizes, agitation systems, filtration times, and settling times?
- If running continuous crystallization, how will you ensure that % supersaturation is constant with scale-up? What is the shear sensitivity of the crystals? What is the impact of agitation on particle size distribution, and impact of shear on secondary nucleation? How does it impact filtration rates and filter cake wash as a function of the attrition, and might attrition be greater at larger scale? Also, if running continuous crystallization, have the scaled-up vessels been properly designed for complete solids suspension mixing? Multiple flat baffles, vessel height to diameter ratio of 1.0 or less, and down-pumping impellers are recommended for solids suspension mixing.
- Are you planning to do a long continuous run 24 h per day with multiple unit operations simultaneously? If so, then it is best to do 12 h continuous demo runs beforehand, through small sections of the continuous train at a time, during the day shifts. This allows you to make sure everything is working correctly before

trying to start up the entire flow train for the long run. Scheduling people for 24 h per day coverage is a significant burden on resources especially if it is an R&D facility that does not normally work shifts around the clock. Make sure the 12 h daytime runs work out the bugs in the continuous train before beginning the 24 h coverage and committing larger amounts of materials.

- Do you need to run scale-up demonstrations in the lab before scaling up to the manufacturing facility? Here are 4 reasons why you might:

  1. Heat and mass transfer rates. Mass transfer changes because you cannot maintain all the same mixing parameters with scale-up. Heat transfer changes because heat transfer area/volume changes and fluid mechanics change with scale. Thus, heat transfer distances and film layers change at the boundaries. If scaling up a mixer-settler or centrifugal extractor, then what is the stage efficiency? How close does each stage operate to equilibrium? What is the relationship between τ, agitation rate, and scale on stage efficiency?
  2. Fouling/encrustation. Solids accumulating on surfaces depends on shear rate and fluid mechanics such as Reynolds number and linear velocity, which are differently scaled up, for example, in a continuous crystallization. Fouling also depends on blend time, which impacts local supersaturation levels, and mixing rate versus crystallization rate. Mixing rates and blend times change with scale-up. Another example is fouling and plugging in heat exchangers, which can be much worse scaled-up because the ΔT between jacket and process is usually higher in the scaled-up unit. ΔT must be higher when scaling up to overcome lower A/V, unless the heat exchanger maintains the same characteristic dimension as for numbering up micro-reactors.
  3. Technology development. If a reactor or unit operation is new and not tested at scale because it did not previously exist, then it may be best to test the scaled-up prototype before starting a GMP manufacturing campaign, where it is more difficult to make changes. This can often be done with short duration experiment for proof of concept and proof of sufficient heat and mass transfer rates, for example, 3–5 volume turnovers.
  4. Material production.

  Chemistry:

- Are you sure that the chemistry is going to work as expected (purity, yield) with the actual starting material lots and solvents? Have you run these exact planned conditions continuously previously? Can you run the reaction batch to confirm by use test, before starting the continuous process? Furthermore, after reagent solutions are made up in feed tanks for the continuous campaign in the plant, before starting flows, it is best to get a sample of the actual feed materials and do a batch (or small-scale flow) reaction off-line in the lab to confirm that the feed solutions are good to go.
- If the reaction uses a catalyst, then have you tested multiple lots of catalyst/ligand/support to make sure reaction results are consistent across a number of catalyst lots?

- How tight are the stoichiometry requirements if the continuous reaction mixes two reagent feeds together? What reagent is in excess, and how much excess is acceptable? It is best if there is an acceptable range on stoichiometry and the excess reagent target is set such that if one of the mass flows is off by 10% the reaction results will still be acceptable.
- Is a seasoning run warranted? Consider a brief reaction run to condition the reactor surfaces, especially if the reactor is a PFR.
- Is the flow system clean? What was used in the reactor previously, and how could it interfere with reaction? At research scale, it is usually best to install a new tube reactor if it is a new chemistry in a PFR. At production scale, it is best if the PFRs are disposable and dedicated to a specific chemistry. PFR tubes are typically inexpensive, and cost is negligible compared to the cost of repeating or delaying an experiment or production campaign.
- Was chemical reaction safety analysis completed? This includes ARC and DSC and calculations for heat removal. Were calculations done for worst-case scenario heat release for exothermic reactions? For example, if running a Grignard formation in CSTRs, did you make sure the reaction is dilute enough so that it is not possible for all the solvent to boil off in an exothermic event (latent heat of vaporization for the total solvent mass is greater than maximum heat of reaction)?
- Did you consider appropriate material compatibility, flammability, reactivity, reaction calorimetry, and thermal stability test data? Do you have appropriate materials of construction? How do you know that there will be no corrosion problems? Was coupon corrosion testing done for this reactor material of construction? All wetted parts of the reaction system including gaskets, o-rings, pressure reliefs, and gauges should be considered.
- What is the chemical stability of starting reagent solutions over time?
- Do you have data on stable hold points?
- Do you have solubility at room temperature where you need it or want it? For example, it is desirable to have solubility at room temperature for liquids flowing in/out of extractions, solvent exchanges, adsorption columns, in addition to reagent feed solutions.
- Is the product stable to end-of-reaction conditions in the reactor? If not then use a PFR, tightly control τ, and have solvent supply vessel ready to swap and valve and pump out the reactor contents in the event of a process stoppage.

Heat and Mass Transfer:

- How do you know that you will have sufficient heat transfer rates? Make sure heat transfer calculations have been done and you know heat-up and cool-down times and distances flowing through the tubes. Make sure heat exchangers are not undersized or oversized. What will be the hot spots for reaction exotherms predicted from the model of reaction rate, heat of reaction, and reactor heat transfer? Is it possible to insert thermocouples to measure the hot spots? Do you know overall heat transfer coefficients for CSTRs and can you remove the heat generated with the jacket alone, or is an internal cooling coil needed as well?

- Is sufficient heat exchange provided for process liquids flowing into and out of reactors or separators?
- If individual circulators are used instead of house cooling system, then do the circulators have enough cooling capacity to remove heat from reaction exotherms and maintain desired process temperature? What are the consequences if the circulators or cooling devices fail during the campaign? Are the safety interlocks in place for shutdown on high-temperature alarms?
- How do you know that mass transfer rates will be sufficient? Will mixing be sufficient when two streams Tee together? If a static mixer is utilized, then do you know the minimum recommended linear velocity through the static mixer for sufficiently fast mixing, given the viscosities and volumetric flow ratio of the two feeds? Are mixing zones or stirred tanks at the inlet to PFRs sufficiently large to dampen out flow fluctuations?
- If the process has separations like liquid–liquid extraction, then what is the relationship between τ, agitation rate, and scale on stage efficiency?

Residence Time Distribution (RTD):

- Is RTD known for each unit operation? CSTRs with τ more than 10 min and blend time less than 1 min typically can be modeled as ideal CSTRs, as long as inlet and outlet tube/pipe positions are sufficiently separated to prevent short-circuiting. If a PFR is used, F-curves should be measured during start-up transitions from solvent to reagents/products, to quantify RTD of substrate. Is the start-up transition time to steady state understood, and is the information used in the decision of when to start collecting product?
- How do you know what RTD will be overall for a continuous process train with multiple unit operations together in series? A numerical modeling software package may be needed.
- Will the planned data collection provide sufficient information to calculate lot genealogy? Documentation will include the time of switching reagents, switching parallel surge vessels, heels in feed and surge vessels, overall τ, and overall RTD.
- Does the numerical model of overall RTD facilitate the calculation of deviation boundaries?
- Is there a plan to quantify the potential difference between RTD of the actual product versus RTD of nonreactive tracers or solvents? These may not be the same for many reasons, including interactions with tube walls or adsorption/desorption from column packing materials.
- Are CSTRs strategically placed where process fluctuations or flow oscillations are expected? CSTRs serve to dampen out process fluctuations. If the reaction has fast kinetics and impurity profile at end of reaction conditions is favorable, then a CSTR is more forgiving to process fluctuations than a PFR. If reaction kinetics is not fast, then CSTR for initial stoichiometric mixing followed by PFR for full conversion may be a good option.
- The importance of knowing RTD is explained in another chapter in this book (Johnson et al. 2020b).

Stop/Restart:

- Is the process designed for ease of stop/restart if adjacent unit operations are having problems and need to be stopped temporarily, or in case upstream material was diverted to waste temporarily?
- Because of resource limitations, it may be preferable to run 16 h/day rather than 24 h/day in a production campaign. This is typically preferred in R&D, because it is easier on people, and it is advantageous to demonstrate that the process has sufficient stop/restart capabilities before going to manufacturing.
- If the process is stopped and restarted, then what automation interlocks might get tripped? For example, if there is an automation interlock based on low pressure, and if pumps are stopped, then does the low pressure alarm trip and do something like close an automated valve? If so, then make sure to reopen the valve before restarting.

Start-up and Shutdown

- Has the research-scale continuous process been started up with the exact same procedure and timing intended for production? This is important so that you know start-up transition will be as expected. For example, solid precipitates can form during start-up transitions and potentially clog lines, depending on timing feeds.
- When will you start collecting product, when will you switch collection tanks, and when will you stop collecting product? These should all be planned in advance. It may be advisable to switch collection vessels more frequently during start-up transition while you are deciding when to start forward processing, and then less frequently once the process reaches steady state. This approach can help minimize start-up transition waste.
- How long will it take to reach steady state? For drug substance, it can take days for a fully continuous multistep process train to reach steady state. Recycle increases the time to steady state. Also, some processes are always in transition, for example, packed catalyst bed reaction with drifting catalyst activity.
- Is it possible to achieve zero start-up transition waste for some of the sections to the continuous processing train? Will a surge vessel downstream from the reactor dampen out the impact of start-up such that the entire transition can be forward processed?
- What are the manufacturing complexity trade-offs? For example, you could start-up a countercurrent extraction train full of solvents rather than empty. This would cause start-up dilution and transition waste, but it would be much simpler than gradually filling each vessel in semi-batch fashion.
- Has a detailed plan for start-up and shutdown transition been created, and are the feed and product cans, tubing and valves set up so that you can follow the plan? For example, you may want to start the process flowing solvents only, then switch to reagent solutions at time zero. This means that you need additional feed vessels and you need switching valves installed. Also, the timing of switching each of the reagents from solvent may not be the same, depending on the time it takes

to flow from the pumps to the mixing Tee for each. The opposite order occurs at shutdown transition. You need a source of solvent ready to switch on-line to the pump inlet for solvent pushout at the end.

- Is the volume of all feed lines measured, and are the times that it takes for each feed to reach the mixing *T* at the reactor inlet known? If not simultaneous, then which reagent or catalyst feed solution do you want to reach the mixer at the inlet to the reactor first? For example, you may want the limiting reagent to reach the reactor inlet last.
- Do you know which reagent you want to run out first, and do you have excess of the other reagents?
- If you are running a PFR or continuous extractor, then do you plan to start with the process full of solvent and then switch over to reagents? Do you plan to do the opposite during shutdown for solvent pushout, meaning that you will switch from reagents to solvents and keep pumping at the same flow rates until the product is pushed out? If so, then can the downstream processing steps handle the start-up and shutdown dilution? For example, a downstream distillation may be able to normalize the concentration by stripping more solvent during start-up transition. If not, then do you need to divert reaction product solution that is not full strength? If you are planning to divert material that is not full strength, then are you using a PFR designed for low axial dispersion to minimize product loss?
- CSTRs can be started up with the reactor in semi-batch mode. Then, flows start when reaction reaches full conversion in the vessel. This method can eliminate start-up transition waste for the reactor. However, if this is the plan, then have the experiments been done to prove that semi-batch start-up mode meets product quality needs? Subsequently, semi-batch shutdown with gradual emptying of the CSTR can eliminate shutdown transition waste from the reactor. Does data exist to support this mode of operation. Also, do you have capabilities to pump down the CSTR at controlled rate during shutdown transition? If flow goes out a dip tube, then will the dip tube be gradually lowered, and if so, is it designed with the capability to gradually lower it without getting into the impellers?
- If you are running a continuous distillation, then plan to delay flowing out of the distillate bottoms until the evaporator reaches steady-state concentrations. The total mass pumped into the evaporator before starting outlet flows should be calculated ahead of time.
- If you are running a continuous crystallization in stirred tanks, one option is to start the process with a batch crystallization in the first stirred tank, then start flows when the vessel is filled and at its crystallization endpoint. Another option is to start with the first stirred tank filled with final crystallization solvent composition and fully seeded with solids closely resembling steady-state particle size distribution. How many stops and restarts do you expect for cleaning out encrustation? Do you have enough seed crystals on hand depending on your desired strategy and expected restart frequency?
- Are manual bypass valves installed where needed for start-up and shutdown transitions? For example, if you expect a small amount of solids precipitation or solids eluting from packed beds during start-up transition but not at steady state,

then do you have two three-way valves installed so that you can bypass on-line HPLC samplers, IR probe flow cells, or other sensitive or easily fouled process components until the process reaches steady state? Process analytical technology (PAT) is useful during start-up transition because it helps to determine time to steady state and help decide when to start collecting product, but if it leads to solids fouling then it may be better to temporarily bypass it.

- If you are running continuous crystallization in MSMPRs, how many system volume turnovers are required to reach steady-state concentration, and how many turnovers are required to reach steady-state crystal size distribution (CSD)? This could be determined by sampling once per turnover and analyzing with something like a Malvern, or it could be measured with an in situ Focused Beam Reflectance Measurement (FBRM). What is the target percentage supersaturation in all vessels at steady state, what was it in the research and development runs, and how will you measure supersaturation in the production run? One option is an in situ probe like IR calibrated for concentration. Another option is sampling, immediately filtering the slurry sample, diluting the sample quantitatively so that product remains in solution, and quantifying potency, then comparing to equilibrium concentration at these conditions. What is the best way to start up the process? Is the first MSMPR started batch-wise before initiating flows? Is it seeded? If start-up is done batch-wise to establish slurry in the MSMPRs before starting flows, then what is the purity of the solids by HPLC or NMR at steady state compared to batch startup? What is the polymorph form of the crystals, CSD, and crystal shape for batch start-up versus steady state?
- How will you quench excess hazardous reagents like LDA or BuLi downstream from the reactor, or activated Mg at the end of processing? Have a plan to do this in a safe and controlled manner that does not generate high concentrations of hazardous gas by-product.
- What is the plan for waste disposal?

Diverting:

- Where are the best points for diverting? This is typically immediately downstream from a PFR, where RTD is narrow, because it minimizes the amount of material that must be diverted before reestablishing the process within specifications. Is RTD known for the system? Does a plan exist for when and why to divert? For example, calculating a rolling average for impurity concentration over a time period equal to downstream fill–empty surge vessels is a reasonable approach to calculating acceptable magnitude and duration of disturbances. Do you have a decision tree for diverting decisions, and will you be set up and ready to divert with valves and catch vessels if needed?
- What are the points in the system of broad residence time distribution? For example, CSTRs, MSMPRs, and surge vessels have broad RTD. It is best to divert before the point of broad RTD, because broad RTD usually results in more material being diverted before that part of the process has returned to steady state or within range. It is better to divert at points of narrow RTD immediately after the reactors.

- Where is the divert valve positioned relative to on-line LC sample point. It is best to insert a delay loop between the on-line LC sample point and the divert valve. For example, a simple highly plug flow coiled tube with high *L/d* ratio can be used for a delay coil. The delay coil should have τ greater than LC method plus result turnaround time (e.g., 30 min). This allows you to "see into the future" with your on-line LC rather than getting results too late to respond. Of course, the on-line LC is not actually seeing into the future, but it is measuring what the composition will be at the divert valve 30 minutes after the sample is taken.

Analytical:

- How do you know that analytical will be correct and accurate? Make sure that the analytical method is up and running and that you have verified it is measuring your product and key impurities correctly before starting flows. Will the process require analytical results in real time? Make sure that on-line and off-line analyses confirm each other. Periodically take samples for off-line analysis at the same time and from the same location as the on-line PAT. Will on-line HPLC be used? Do you have a manual sampling point at the same location as your on-line HPLC, so that you can confirm on-line results with an off-line sample?
- Do you have a backup analytical plan so that the process can continue to run if the PAT is down temporarily? For example, the decision may be to take manual samples for downstream surge vessels and analyze them off-line until the PAT is back on line. More material is at risk while the PAT is down, because if a downstream surge vessels must be diverted it is probably a larger amount of material to divert. However, all material is still analyzed before it is pumped into the next step.
- What will be the first off-line analytical sample of the continuous run during start-up transition? When will it be expected to be pulled from the process? How soon will you need it analyzed in order to respond with a process adjustment if needed? How quickly can you get it analyzed and results reported from the lab? Is the analyst ready and expecting the sample? How do you know whether or not the sample represented steady state? Is this known flat-lining of PAT, or by proving the same result with multiple off-line samples pulled at some frequency?
- Is the sampling plan realistic, that is, is it possible to keep up with the number and frequency of samples that you plan to take? Where will samples be placed once pulled? How will the samples be labeled? For any continuous process, it is best to include the date and time in the sample name, because then every sample is necessarily unique and you can look back at process parameters at that time? What off-line sample analytical will be needed near real time throughout the campaign? Do you have a unique way to label and a unique place to put those urgent samples? What is required analytical frequency at start-up versus at steady state?
- How do you know that the samples will be representative? The sample position must be designed so that it receives representative flows, that is, it is not in a dead leg or dead zone. How big are the waste cuts and how big is the dead leg in the sample zone? How should I quantify the volume or mass of each waste cut to

make sure it was sufficient? Justify why you believe that the samples are representative.
- What if the sample port clogs with solids? Have a plan, and install the proper valves so that the sample port can be unclogged without disrupting the process.
- It is possible for PAT to interfere with the process in a negative way? It is best if the PAT does not obstruct the process flows, for example, restrict flows or introduce valves in line with the main process flows. If using on-line HPLC, how do you know that you will not get diluent, quench, or derivatizing reagent back into the process? You should always have more than one automated block valve separating the process from these PAT-related solutions, in case one valve fails.
- If you are using a sample port, then does the process dip tube go to the bottom of the overflow tee (via bored through fitting) for manual sampling points and for on-line dilution cart automated sampling? The overflow Tee is a good way to get samples without the potential to interrupt process flows, but you must install the dip tube such that the sample zone is not a low dead leg.
- Did you install manual bypass valves so that you can bypass the probe or auto sampler? This may be important during start-up transition if you have solids issues, for example, the outlet of a packed catalyst bed reactor where solids fines can elute from the reactor for a time when a new bed is put on line. Furthermore, this may be important at any time in the long run if you need to swap out parts or do maintenance or troubleshooting on the PAT.
- Is there a plan for how to use PAT for troubleshooting? What could go wrong in the process and what would be the PAT response to look for? If conversion or selectivity of a continuous reactor changes, a typical question is this: What reagent feed was changed 1 τ before the composition at the reactor exit changed? For example, if the reactor has a 6 h mean residence time and conversion drops at 12:00, then you will look to see if one of the reagent or catalyst feed lots was switched at 06:00.
- If the process has batch collection vessels after a continuous reactor, then have you selected the size of collection vessels that best meets your needs for analytical testing and forward processing decision? For example, you may want 24 h collection vessels, so that you only need to do analytical for forward processing once per day. The decision is based on a balance between minimizing the required analytical frequency and minimizing the amount of product at risk at a given time (putting too many eggs in one basket).

Operations:

- Do you have written start-up, shutdown, and operating procedures? Is there a plant "ticket" written to give instructions for transitions as well as steady-state operation?
- How can you work out the bugs first in a solvent only run? Are the planned solvent run conditions identical to the process conditions that you plan to use in the real chemistry campaign?
- Do you have a plan for daily checks for long continuous runs? Check all things daily that you know could be problematic. For example, check liquid level in

circulators. If the liquid level gets too low because of evaporation or other loss of heat transfer fluid, then the circulator could shut down. Be proactive and keep them topped up. Think about all the heating and cooling utilities. Also, check the level in feed vessels that are filled infrequently (e.g., once per week). Check the amount of solids buildup in continuous crystallizers. Check pressure drop across in-line filters and replace if needed. Are vent bubblers still bubbling, indicating that vent system is remaining inert with nitrogen and that vent piping and reactor headspaces are sufficiently sealed? Check to see if manual nitrogen rotameters are still operating at desired settings. Check transparent tubing at the outlet from pressure relief valves, or vessels at the outlet from relief valves, to make sure that valves have not relieved and discharged material from the process. Do secondary calculations of mass flow rates by tank outages to verify that mass flowmeters are still accurate. Check to make sure there is no buildup of material in vent lines. Are filtration rates slowing, indicating that filter pads need to be cleaned?

- Do you have an operator checklist and data table prepared? What readings do you want recorded once per hour? For example, check liquid levels in vessels, temperatures and pressures not recorded by the DCS. Check DCS trends for pressures, temperatures, and levels that are recorded, including pressure at the suction side of all pumps. Check interface levels in mixer-setters. Check slurry transfer tubes for continuous crystallizers and filters. Look for signs of pump cavitation. Confirm that no feed vessels are about to run out.
- How frequently do you plan to switch product collection vessels? Even if the product collection vessel has capacity for several days, make sure to deliberately switch your product collection vessel at least twice per day during a long multi-day run. That way, if there is a process upset you have less material at risk in any one product collection vessel. Write your instructions with flexibility so you have the ability to switch to a new product collection vessel at any time.
- How will you prevent reactant accumulation, or total mass accumulation? Make sure vessels are not gradually filling if they are supposed to be at constant level. How will you prevent overfilling vessels? This especially applies to product collection vessels.
- Do you plan to run more than one continuous unit operation simultaneously? If so then it is better to run each of them individually first to work out the bugs before putting it all together in a real production campaign.
- Is there a plan for monitoring vent knockout vessels? It is best to monitor them automatically with alarmed level detectors or weigh scales, but they can be monitored manually if frequency is proceduralized and monitoring frequency is sufficiently high that they cannot overfill.
- What will happen if you temporarily lose power during the run? Would this be a safety hazard? What are the fail-safes? What items would be difficult on the restart after power is restored?
- Is everybody ready (engineering, chemistry, operations, automation, analytical), and do all know the start date and the planned people coverage? This is generally well known for real manufacturing campaigns, but it has the tendency to be less well communicated in laboratory development runs.

- Do you have enough people? For example, suppose that you intend to run a three-step fully continuous process in manufacturing, each step with reaction, separation, and purification unit operations. The start-up transition will require the most people, and the most technical expertise. You might need two process engineers, two process chemists, one automation engineer, two analytical chemists for on-line PAT, six operators around the clock during start-up transition to get the entire train up and running. After the entire process reaches steady state, the staff can decrease. The main point is that start-up transitions are resource intensive, and require the technical experts, and a mix of R&D people who developed the process and manufacturing plant personnel collaborating on the plant floor together.
- Make sure all vessels are inerted before flammable solvents flow into the vessel, especially if they flow into the vessel through nonconducting tubing?
- Grounding and bonding is necessary when flowing from one vessel to another to prevent static charge buildup and sparking. Avoiding the use of nonconductive heat transfer fluid with nonconductive piping in the heat transfer fluid pump around loops.
- Conduct process hazard reviews (e.g., What-if or HAZOP reviews, preferably involving multiple people). Hold pre-startup safety reviews to ensure equipment and controls were installed as designed, and to communicate critical safety issues to operations.
- Select a safe location for the flow experiment or manufacturing considering the number of people affected by a worst-case scenario.
- Program DCS calculations of real-time mass and energy balances for safety critical operations.
- Prevent plugging of vent lines, and prevent process materials flowing out vents. Prevent closed block valves in vent lines. Prevent backflow into feed lines or feed vessels, and prevent backflow of liquid from vent lines such as backflow of caustic from scrubbers. Prevent undesired phase changes in process lines like boiling or freezing.

Equipment:

- Have you operated the temperature control units for extended times prior to the real chemistry run? If it will be a multiweek continuous process run, then prove beforehand that all heat transfer units will maintain jacket set points for at least the length of the planned duration. For example, cooling circulators can gradually accumulate internal ice and lose cooling capacity over time. Make sure to get constant temperature baths adjusted to your desired operating temperatures the day before, if heat-up time is significant. For example, heat-up of 200 L water or oil baths used for submerging PFRs can take time depending on the power of your circulators.
- What equipment will be used for overall reactor system back pressure regulation? Is there a potential for solids during transitions? If so, then it may be best to use an expansion chambers in series back pressure regulator rather than diaphragm or spring-loaded regulators (Johnson et al. 2016).

- Is all of the equipment ready? Has it all been tested in a solvent run?
- Do you have redundancy where you may need it? For example, if you intend to run a long campaign with multiple continuous steps operating simultaneously, do you have redundancy of key temperature and pressure measurements, and on-line PAT instruments?
- Do you have equipment backups for pumps, peristaltic pump tubing, agitators, glass vessels and heads, automated open/close block valves, and temperature control units, all of which could fail or break in the middle of the campaign? Have a plan to stop flows to swap out equipment with minimal down time.
- Will electrical and utilities be sufficient to support all the portable flow equipment that must be hooked up? In particular, think about the electrical needs for heaters and chillers that can draw high amps and make sure there are enough high amp circuits, and that you do not have too many of these running off the same circuit.
- How do you know that the equipment will run without fail for the duration of the planned campaign? Off-line long-term reliability testing on key equipment before the real continuous chemistry run is recommended.
- Do you have a sufficient supply of fittings that are GMP approved and kept in a regulated storage facility such that you have backups and can replace process lines and fittings in case of leaks or fouling?
- Get started on specifying and ordering pressure relief valves at least 6 months in advance, and order backups of all pressure reliefs. Include pressure relief devices and/or auto-shutoff pressures on the discharge of all positive displacement pumps. Provide pressure relief devices immediately downstream from back pressure regulators (lower pressure set point to protect downstream equipment). Are all pressure reliefs compatible with their process stream?
- Do you have representative and complete process and instrumentation diagrams (PIDs)? Do the PIDs match the actual equipment in the plant or lab after it is set up, including all manual block valves? Have plant personnel walked the process lines referencing the PID?
- Does the equipment have sufficient venting so that vessels do not pressure up when filling (with vent paths free of block valves). Use separate vent headers when multiple vessels have incompatible headspaces. Provide vent knockout tanks to catch process materials in the event that vessels overfill and also catch bubbler liquid in case of suck back. Be especially careful with separation of liquids and vapors when a reactor is depressurizing and the gas/liquid mixture is flowing into a vessel and gas is exiting to a vent at the same time (e.g., use hydro cyclones to prevent misting out the vent).
- Is secondary containment installed where leaks, breaks, or overfills may occur? Examples are beneath pumps, bubblers, sample valves, and CSTRs.
- Include chemical sensors to detect leaks and hazardous concentrations, for example, hydrogen detectors in the tops of hoods where H_2 is used.
- Have all gaskets been checked for compatibility with the solvents in my process? For example, Viton® o-rings and gaskets can look like Kalrez®; therefore, if Kalrez® is needed for solvents like THF then check them all carefully. If the

wrong o-ring or gasket material is installed, and if it goes unnoticed during construction and assembly, then you should at least identify the problem during the solvent run before the real chemistry run.

- Does the equipment have alarms, interlocks and/or auto-shutoffs based on temperatures, pressures, and fill levels? Does equipment have emergency stops for power and automation fail safe?
- If this is a tech transfer, what is the difference between feed, surge, and product cans between the two locations? At the new location are the surge vessels and product receivers sufficiently sized so that the tank switching or tank emptying can be done only 1 or 2 times per day for each, if the intention is to do it manually?

Planned Duration of Scale-up Demonstration

- How long should the scale-up demonstration be run before the manufacturing campaign? Run at least long enough for the entire process to reach steady state. This might only require a few hours. Run longer for special circumstances. Several examples are given below. Many of these will not need to operate 24 h per day, if they return to steady state quickly after stopping and re-starting, but the total cumulative run time without cleanout may need to be several days. If possible, design continuous reactions and separations unit operations that can easily be stopped, held for a time (e.g., 10 h), and restarted without any negative consequence. It is better for manufacturing. Therefore, in most demonstration campaigns, it is more valuable to prove that you can run the continuous process during the days and stop flows overnight, rather than operate 24 h per day, and it is much easier on people.

    1. If you are using a packed catalyst bed or CSTR with sequestered catalyst and activity decreases over time, then you need to run for the expected catalyst life, which could be several days.
    2. If you are running a process with recycle, then a long time could be required for impurities to build up to steady-state concentrations in the recycle loop, perhaps 100 h.
    3. If you are running a packed bed adsorption, then it could take more than 300 h for the mass transfer zone to take shape along the bed length and another >300 h for the mass transfer zone to move one MTZ length down the bed to see if it is spreading over time.
    4. If the process has fouling, plugging, crusting of solids over time, like continuous crystallization, cryogenic PFRs, or WFE distillation.
    5. If it is possible for material to accumulate in a rector or a unit operation, for example, precipitate of metal impurities in continuous Grignard formation reactor, if the Mg is not 100% pure.
    6. If you are running a continuous filtration and dissolve-off process, and other insoluble solids and tiny debris accumulate on the filter pads over time.
    7. If you are running a continuous extraction and you are interested in the insoluble materials that accumulate in a rag layer between the liquid phases over time.
    8. If you need to test PAT for drift over time due to crusting, fouling, instrument drift, or tube swelling.

Around the Clock Operation:

- Do you need to run 24 h per day or can you stop flows overnight and restart in the morning? Most of the reactions and separations unit operations can be operated during the day shifts in R&D. You only need to keep the process running 24 h per day if there is a negative consequence of stopping and restarting flows. For example, if it is not a stable hold point. Here are a few examples:

  1. Asymmetric hydrogenation, where PFR τ is 12 h and the product is not stable at end of reaction conditions (Johnson et al. 2012).
  2. Continuous cooling crystallization. You may not want to hold the slurries in the MSMPRs hot overnight because of stability and potential solvent loss. If you cool down the MSMPRs, then it will change particle size and CSD. Once you start the process backup, it might take 24 h to reach steady-state CSD again in all the MSMPRs. If this is the case, it is better to run 24 h per day and not stop flows until the end.
  3. Continuous crystallization with kinetic rejection of impurities (Johnson et al. 2012). In these examples, it is better to run 24 h per day and not stop flows until the very end of the campaign. When flows stop the stirred tanks will reach thermodynamic equilibrium, which is obviously not desirable for demonstrating a kinetic impurity rejection.

Automation:

- Note that experts in automation engineering are required for running fully continuous processes.
- What level of automation will be used? What will be automated by the plant-wide DCS and what unit operations will be automated by small local PLCs? Will a plant-wide control system be used, or individual unit operation controlled with surge tank decoupling?
- Are the parameters tuned for all feedback control loops?
- Are unit operations linked in series, and if so, is sufficient time allotted for automation testing before a campaign?
- Are the skids programmed the same for the manufacturing campaign as they were at research scale or pilot scale in R&D? This is especially important for sequenced operations, like intermittent flow evaporators and filters, as described previously in this chapter. If this is a technology transfer from one site to the next, and especially if the automation platform is different at the new location, did the personnel who ran the process at the previous site, or in development, consult with the automation engineers at the new site to make sure startup and shutdown transitions, ramp rates, and ratio control are programmed consistently with the previous process?
- How can the system be started up if stopped in the middle? Which stop/restart sequences will be automated, and which will be manual.
- Is there an overall control screen where multiple variables can be viewed at once?
- Allot time prior to the production campaign to make sure there are no bad channels on the cards, that solenoids are functioning, that transmitters are properly spanned, etc.

Plugging/Fouling

- What is the potential for plugging and fouling during the continuous production campaign and how will it be measured? What changes can be made to decrease the chances of fouling and plugging? For example, if you are running a continuous crystallization, then operating at lower supersaturation, using a larger number of vessels in series so that less product needs to come out of solution in each vessel, installing slurry transfer tubing subsurface, and locating the inlet away from agitator shaft, walls, or baffles may all minimize solids fouling.
- Is tube size large enough to minimize plugging if you suspect you could have solids? Can you use high-velocity intermittent flow from vessel to vessel for slurries, as described previously in this chapter?
- Can you monitor for plugging and fouling visually? This is facilitated by using transparent glass vessels and PFA tubing for slurry transfers if the reactor, crystallizer, extractor, or evaporator runs at low pressure. If nontransparent metal tubing or vessels are used for high-pressure operations, then can you monitor for plugging with pressure transmitters? Can you provide mitigations for plugging and fouling (e.g., double block and bleed valves) so that you will not need to use a wrench on a clogged line under pressure?
- Have you installed pressure gauges and transmitters for measurement of pressure versus time at pump outlets (provides indication of gradual fouling)?
- What are the most likely unit operations for solids fouling? For example, continuous crystallization, filtration, or cryogenic reaction have fouling potential, so these should be a main focus during preparations.
- Does the process have any heat traced lines because the flowing stream would not be a homogeneous solution at room temperature? If so, is every inch of the process line sufficiently heated? Small "cold" spots can plug over time or when there is a flow stoppage.
- What are the smallest flow restrictions for process materials? For example, flow restrictions may be needle valves, small fittings, restricting orifices for mass flow control or back pressure control. These will plug first.
- What are the parts of the system where you may not have solubility at lowest possible process temperatures? What are the locations of potential supersaturation?
- Does the process have a continuous liquid/vapor separator downstream from a back pressure regulation system? If so, then plugging in vent lines from misting process solutions with dissolved substrate may be a concern. In this circumstance, design the process to vent slowly, design cyclones for V/L separation, use large vent lines, and avoid vent restrictions.
- If you are running continuous crystallization, then what is the expected rate of fouling with solids over time? How does fouling and crusting change with scale-up? Where are the most likely locations for solids fouling, for example, on the agitator shaft, on the baffles, at the location of feed entry into the vessel, in the tubing between vessels? How does the level of supersaturation impact fouling rates? How long do you plan to run the process at steady state before scheduled

periodic shutdowns for solvent cleanouts? Is the plan to operate the continuous crystallization until the operations staff makes a judgment call on when it is time for a cleanout? Do the fouling solids on the walls have the same purity as the flowing crystals? What is the chemical stability of the solids on the walls, in case some break off and flow downstream? This can be tested by getting a sample of solids stuck to a wall or baffle and analyzing them for purity. Where are the inlet points, where product solution and anti-solvent flow into the stirred vessel? It is best if they are on opposite sides of the vessel. Are the inlet tubes located at a sufficient distance away from impellers, shaft, baffles, and wall? Do they mix in a static mixing device in a pump-around loop? What is the mixing rate at these addition points?

- How do you plan to unclog a clogged line? For example, use back-and-forth pressure pulses or back-and-forth pumping as a first attempt to free the clog, before using a wrench to open up process equipment.

GMP related:

- Do you intend to have any start-up and shutdown waste? Or, do you intend to start-up and shut down in a manner that all products meet specifications and thus there will not be any transition waste?
- What will be the cleaning procedures at the end of campaign and after a process upset? How will you clean a long PFR that is difficult to inspect internally? It may be better to use inexpensive tubing and dedicate the PFR to a specific chemistry, disposing it when it is no longer needed for that specific product.
- How will you calculate lot genealogy? Do you have a plan to collect sufficient information on τ and RTD to calculate lot-to-lot carryover?
- Do you have a plan for deviation management (deviation boundaries—operating space, design space)?
- If this will be a validation campaign, then what is the required number of start-ups and shutdowns, required batch size and required steady-state operation time?
- What is the required analytical sampling frequency? Is the analytical sample for controlling process parameters taken from a point of narrow RTD, and is the analytical sample for forward processing decision taken from a point of broad RTD?
- What are the process parameters that you will use for feedback control (automated PID control loops based on T, P, flow, level, versus analytical PAT)?
- Do you intend to use on-line PAT for real-time information, or real-time forward processing, or real-time feedback control, or real-time release, or some combination of these? Do you need redundancy in case the PAT instrument stops working?
- What is an acceptable degree of variability and fluctuations of process parameters?
- How will in-spec process parameter adjustments be made? In other words, if no material is out-of-spec yet, but adjustments within the acceptable processing range are being made to process parameters to adjust closer to target set points, then how do you know that each operator would make the same adjustment? Use decisions trees or automation.

- What are the points in the continuous unit operations for diverting flow in the event of a deviation?
- Do you have any deliberately non-steady-state unit operations, such as fixed catalyst bed reaction where catalyst activity changes over time (e.g., 300 h) or fixed bed adsorption where solid phase becomes exhausted over time (e.g., 600 h). What if the continuous unit operation has deactivation over time?
- Where are you building in redundancy for processing equipment?
- What if the continuous unit operation has known fouling over time? What is the plan for periodic off-line solvent cleaning and flushing to remove solids buildup?
- Will this process deliver flexible batch size or fixed batch size?
- How will you document material movements from one part of the continuous train to another?
- Does the process include recycle? If so, how will you monitor for impurities building up in recycle loops?
- Will the GMP processing instructions and master batch record be handled differently than batch?
- What will be the material hold times?
- How can you use the numerical model to determine when to divert flow, or if diverted material is unacceptable?

Safety:

- See the continuous drug substance processing safety chapter in this same book (Johnson and Niemeier 2020).

References

Adams TA, Pascall A. Semicontinuous thermal separation systems. Chem Eng Technol. 2012;35(7):1153–70.

Changi SM, Yokozawa T, Yamamoto T, Nakajima H, Embry MC, Vaid R, Luciani CV, Wong SW, Johnson MD, Moher ED. Mechanistic investigation of a Ru-catalyzed direct asymmetric reductive amination reaction for a batch or continuous process scale-up: an industrial perspective. React Chem Eng. 2017;2:720–39.

Cole KP, Campbell BM, Forst MB, McClary Groh J, Hess M, Johnson MD, Miller RD, Mitchell D, Polster CS, Reizman BJ, Rosemeyer M. An automated intermittent flow approach to continuous Suzuki coupling. Org Process Res Dev. 2016;20(4):820–30.

Cole KP, Groh JM, Johnson MD, Burcham CL, Campbell BM, Diseroad WD, et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow cGMP conditions. Science. 2017a;356:1144–50.

Cole KP, Johnson MD, Laurila ME, Stout JR. An automated repeating batch with catalyst recycle approach to nitro group hydrogenolysis. React Chem Eng. 2017b;2:288–94.

Cole KP, et al. Small-volume continuous manufacturing of merestinib. Part 1. Process development and demonstration. Org Process Res Dev. 2019;23:858–69.

Felder RM, Rousseau RW. Elementary principles of chemical processes. New York: Wiley; 1986.

Frederick MO, Calvin JR, Cope RF, LeTourneau ME, Lorenz KT, Johnson MD, et al. Development of an NH4Cl-catalyzed ethoxy ethyl deprotection in flow for the synthesis of merestinib. Org Process Res Dev. 2015;19(10):1411–7.

Johnson MD, Niemeier J. Safety guidelines for continuous processing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020. (this volume).

Johnson MD, May SA, Calvin JR, Remacle J, Stout JR, Diseroad WD, Zaborenko N, Haeberle BD, Sun WM, Miller MT, Brennan J. Development and scale-up of a continuous, high-pressure, asymmetric hydrogenation reaction, workup, and isolation. Org Process Res Dev. 2012;16(5):1017–38.

Johnson MD, May SA, Haeberle BD, Lambertus GR, Pulley SR, Stout JR. Design and comparison of tubular and pipes-in-series continuous reactors for direct asymmetric reductive amination. Org Process Res Dev. 2016;20(7):1305–20.

Johnson MD, et al. Continuous reactors for pharmaceutical manufacturing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020a. (this volume).

Johnson MD, et al. Understanding residence time, residence time distribution, and impact of surge vessels. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020b. (this volume).

Kopach ME, Cole KP, Pollock PM, Johnson MD, Braden TM, Webster LP, Groh JM. Flow Grignard and lithiation: screening tools and development of continuous processes for a benzyl alcohol starting material. Org Process Res Dev. 2016;20:1581–92.

May SA, Johnson MD, Braden TM, Calvin JR, Haeberle BD, Jines AR, et al. Rapid development and scale-up of a 1 H -4-substituted imidazole intermediate enabled by chemistry in continuous plug flow reactors. Org Process Res Dev. 2012;16(5):982–1002.

May SA, Johnson MD, Buser JY, Campbell AN, Frank SA, Haeberle BD, et al. Development and manufacturing GMP scale up of a continuous Ir-catalyzed homogeneous reductive amination reaction. Org Process Res Dev. 2016;20(11):1870–98.

Reizman BJ, et al. Small-volume continuous manufacturing of merestinib. Part 2. Technology transfer and cGMP manufacturing. Org Process Res Dev. 2019;23(5):870–81.

Tsukanov SV, Johnson MD, May SA, Rosemeyer M, Watkins MA, et al. Development of an intermittent-flow enantioselective Aza-Henry reaction using an arylnitromethane and homogeneous Bronsted acid-base catalyst with recycle. Org Process Res Dev. 2016;20:215–26.

White TD, Berglund KD, Groh JM, Johnson MD, Miller RD, Yates MH. Development of a continuous Schotten–Baumann route to an acyl sulfonamide. Org Process Res Dev. 2012;16(5):939–57.

White TD, Alt CA, Cole KP, Groh JM, Johnson MD, Miller RD. How to convert a walk-in hood into a manufacturing facility: demonstration of a continuous, high-temperature cyclization to process solids in flow. Org Process Res Dev. 2014;18:1482–91.

Zaborenko N, Linder RJ, Braden TM, Campbell BM, Hansen MM, Johnson MD. Development of pilot-scale continuous production of an LY2886721 starting material by packed-bed hydrogenolysis. Org Process Res Dev. 2015;19(9):1231–43.