

Supporting Information

Leveraging GPT-4 to transform chemistry from paper to practice

Wenyu Zhang,^a Mason A. Guy,^a Jerrica Yang,^a Lucy Hao,^a Junliang Liu,^b Joel M. Hawkins,^c Jason Mustakis,^c Sebastian Monfette,^{*c} and Jason E. Hein^{*abde}

*Corresponding author

- a) Department of Chemistry, The University of British Columbia, Vancouver, BC V6T 1Z1, Canada
- b) Telescope Innovations Corp., Vancouver, BC, Canada
- c) Pfizer Worldwide Chemical Research and Development, Pfizer Inc., Groton, Connecticut 06340, USA
- d) Department of Chemistry, University of Bergen, Norway
- e) Acceleration Consortium, University of Toronto, Toronto, ON, Canada

Contents

S1.	General Information.....	2
S1.1.	Chemicals and Material	2
S1.2.	Analytical equipment and methods	2
S1.3.	Automated Sampling Platforms and Parameters	3
S2.	Experimental Detail	4
S2.1.	General procedure.....	4
S2.2.	Case study 1: S _N Ar aminolysis.....	5
S2.3.	Case study 2: synthesis of hydrazone compound	5
S2.4.	Case study 3: Curtius rearrangement.....	5
S3.	Prompt detail and outputs	6
S3.1.	Prompt full templates.....	6
S3.2.	Prompt development.....	9
S3.3.	GPT-4 output	10
S3.4.	Comparison to ChemIDE and RXN	40
S3.5.	Example errors and solutions	41

S3.6.	Prompt reproducibility.....	47
S4.	Workflow and Tutorial	48
S4.1.	iC software installation.....	48
S4.2.	Configuration.....	48
S4.3.	Save .XML from GPT	50
S5.	NMR spectra	51
S5.1.	Case 1: Synthesis of 1-methyl-4-(4-nitrophenyl)piperazine	51
S5.2.	Case 2: Synthesis of (4-fluorobenzylidene)hydrazine.....	52
S5.3.	Case 3: synthesis of carbamate after Curtius rearrangement.....	53
Reference	54

S1. General Information

S1.1. Chemicals and Material

Potassium carbonate was purchased from Anachemia. Biphenyl, 4-fluoronitrobenzene, 1-methylpiperazine, 4-fluorobenzaldehyde, hydrazine hydrate, triethylamine (TEA), 4-fluorobenzoic acid, were purchased from Sigma-Aldrich. Diphenylphosphoryl azide (DPPA) and hexafluoroisopropanol (HFIPA) were purchased from Oakwood Chemicals. Toluene and other HPLC solvents were from Fisher Scientific. Solvents and reagents were used as received from commercial suppliers unless otherwise stated. No special consideration was given regarding reaction atmosphere and no additional steps were taken to ensure moisture control of the reagents or solvents used.

S1.2. Analytical equipment and methods

NMR measurements were performed with a Bruker AV-400 MHz spectrometer. Spectra were collected at 298 K; chemical shifts are reported in ppm and were calibrated for the (residual) NMR solvent signal. ¹H NMR multiplicity was reported using the following abbreviations: br = broad, s = singlet, d = doublet, t = triplet, q = quartet, quint. = quintet, sext. = sextuplet, sept = septuplet, and m = multiplet. The spectra were processed with Mestre Nova 14.3.0.

HPLC-UV/Vis (MS) analysis was conducted using an Agilent 1290 infinity HPLC equipped with a: G4220A binary pump unit, G4226A multi-sampler, G1316C thermal column compartment, G4212A diode array detector, and G6140A single quad mass spectrometer. Data processing and analysis were carried out using both Chem Station (Agilent) and another proprietary third-party software. Analysis on both systems was conducted using the method outlined in **Table S1**. Reactions were performed in and temperature controlled using a Mettler Toledo EasyMax 102 Thermostat system.

Table S1. HPLC-UV/Vis method.

Column:	Poroshell 120 EC-C18 2.7 μm , 2.1 x 50 mm	
Column Temperature:	35.0 $^{\circ}\text{C}$	
Flow Rate:	0.650 mL/min	
Acquisition Time:	3.50 min	
Mobile Phase:	Solvent A: Water (0.1 % formic acid) Solvent B: Acetonitrile	
Mobile Phase Program (case 1 and 2):	Time (min)	%B
	0.00	5
	0.10	20
	2.30	100
Mobile Phase Program (case 3):	Time (min)	%B
	0.00	5
	0.50	100
Compound Name:	Retention time (detection wavelength):	
Case Study 1		
(4-fluorobenzylidene) hydrazine	0.99 min (210 nm)	
4-fluorobenzaldehyde	1.40 min (270 nm)	
1-(dimethoxymethyl)-4-fluorobenzene	1.69 min (210 nm)	
Case Study 2		
1-methyl-4-(4-nitrophenyl) piperazine	0.62 min (270 nm)	
4-fluoronitrobenzene	1.38 min (230 nm)	
biphenyl	1.97 min (210 nm)	
Case Study 3		
biphenyl	2.090 min (270 nm)	
DPPA	1.713 min (270 nm)	
4-fluorobenzoic acid	1.519 min (230 nm)	
4-fluorophenyl isocyanate	1.746 min (270 nm)	
4-fluorobenzoyl azide	1.630 min (254 nm)	

S1.3. Automated Sampling Platforms and Parameters

Time course data of reactions was collected using a DirectInject-LCTM platform developed by Telescope Innovations described below (**Figure S1**).

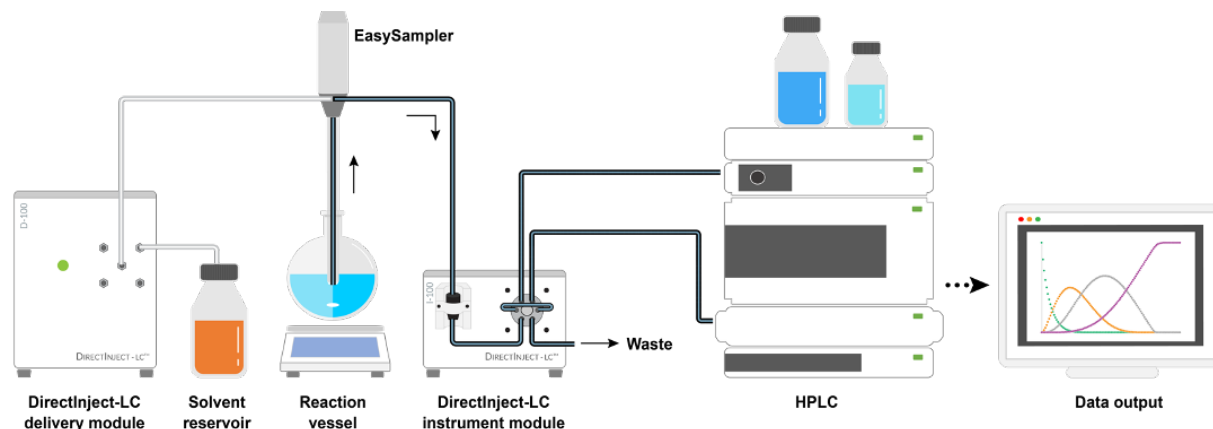


Figure S1. DirectInject-LC™ automated sampling system coupled with online HPLC.

DirectInject-LC™ with online HPLC. To commence sampling, the EasySampler was actuated to extend the sampling pocket into the reaction mixture, collecting a 20 μ L aliquot.

S1.3.1. Case study 1 and 2 sampling conditions

The system was filled with solvent (MeCN, 2.5 mL, 5.0 mL/min). The pocket was retracted, and MeCN (1.8 mL, 5.0 mL/min) was pushed to position the aliquot immediately before the injection valve and sample loop. During delivery, the sample was passed through an in-line mixer. The injection valve was switched to bring the sample loop in-line with the EasySampler, and MeCN (0.1 mL, 0.5 mL/min) was pushed to load the reaction aliquot onto the sample loop. The injection valve was switched again to align the sample loop containing the reaction aliquot with the HPLC mobile phase and column for online analysis. The system was then flushed with MeCN (2.5 mL, 5.0 mL/min) to clear the system prior to further sampling.

S1.3.2. Case 3 sampling condition

The system was filled with solvent (toluene) before pocket extend. The pocket was extended, and MeTHF (4 mL, 2.5 mL/min) was pushed through the line, then pocket was retracted, MeTHF (2.3 mL, 2.5 mL/min) was pushed to position the aliquot immediately before the injection valve and sample loop. During delivery, the sample was passed through an in-line mixer. The injection valve was switched to bring the sample loop in-line with the EasySampler, and MeTHF (0.1 mL, 0.5 mL/min) was pushed to load the reaction aliquot onto the sample loop. The injection valve was switched again to align the sample loop containing the reaction aliquot with the HPLC mobile phase and column for online analysis. The system was then flushed with toluene (5 mL, 2.5 mL/min) to clear the system prior to further sampling.

S2. Experimental Detail

S2.1. General procedure

An EasyMax 102 device, equipped with a 100 mL reaction vessel, a 5-port lid, a temperature probe, a mechanized stirrer, and a SP-50 dosing unit. Before starting the experiment, the solvent of the dosing unit was changed to the solvent chemical prompted in the imported procedure. Once the iControl experiments started, the required chemicals were weighed according to the

prompted steps manually using a balance, then added to the reaction vessel. The remaining procedure was not interrupted until the end of the reaction.

S2.2. Case study 1: S_NAr aminolysis

The solvent of DU-1 was changed to acetonitrile. Upon the start of the iControl procedure, 9.55 g of potassium carbonate, 65.3 mg of biphenyl and 6.5 g of 4-fluoronitrobenzene were measured and added manually to the 100 ml reactor on a EasyMax 102 workstation. 32.51 ml of acetonitrile was pumped to the reactor according to the iControl operation at 5.0 ml/min rate. The stir rate was set to 400 rpm. The temperature was ramped to 60 °C over 15 minutes. The reaction was monitored by sampling every 10 minutes using the DirectInject-LCTM for 8 hours. Upon iControl's prompt after reaching to 60 °C, 6.69 g of 1-methylpiperazine was added manually to the reactor at once.

S2.3. Case study 2: synthesis of hydrazone compound

The solvent of dosing unit 1 (DU-1) was changed to methanol. Upon the start of the iControl procedure, 2.482g of 4-fluorobenzaldehyde was added to the 100 ml reactor on a EasyMax 102 workstation. 20 ml of methanol was pumped to the reactor according to the iControl operation with the DU-1 at 10 ml/min rate. The stir rate was set to 500 rpm with temperature set to 25 °C. The reaction was monitored every 5 minutes using the DirectInject-LCTM for 1 hour. 1.2 ml of hydrazine hydrate was added at 17 minutes after the start of the reaction.

S2.4. Case study 3: Curtius rearrangement

The solvent of dosing unit 1 (DU-1) was changed to toluene. Upon the start of the iControl procedure, 10 mL of toluene was dosed to a 100 ml reactor on a EasyMax 102 workstation. 1.1 mL of TEA and 1.036 g of 4-fluorobenzoic acid was manually added to the reactor at 25 minutes. The stir rate was set to 400 rpm. After 42.5 minutes, 1.8 mL of DPPA was added manually to the reactor. The reaction was continued for 127 minutes and changed temperature to 80 °C. Gases release was observed immediately after ramping the temperature. After the consumption of nitrene intermediate slowed down at 30.7 hours, HFIPA was added to the reactor for another 3 hours. The reaction was sampled every 10 minutes using the DirectInject-LCTM for the 33.78 hours.

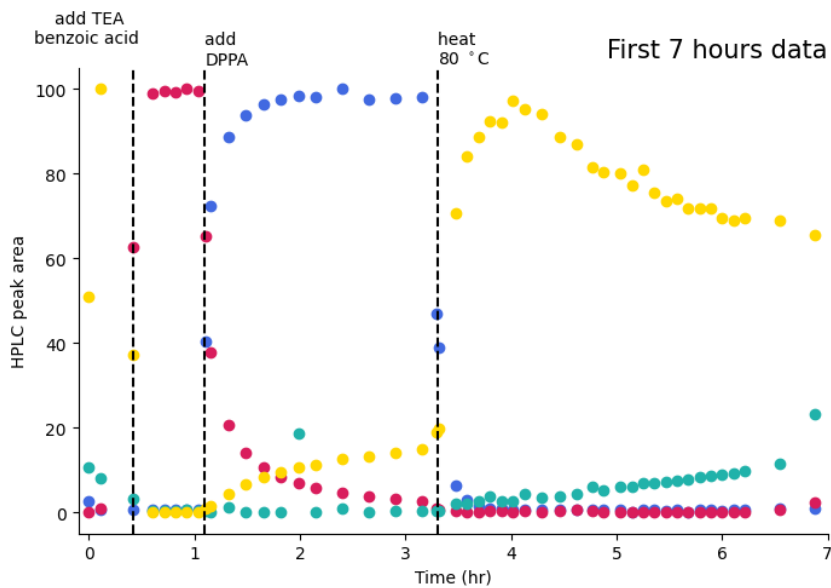
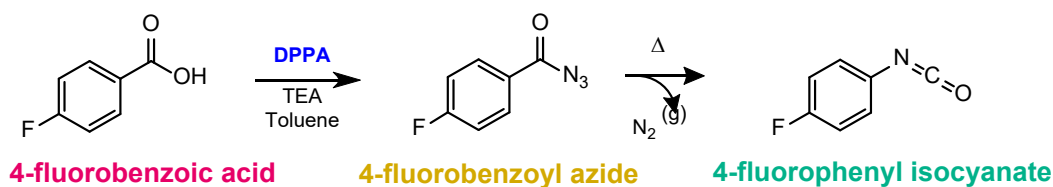


Figure S2. Online HPLC monitoring of 4-fluorobenzoic acid, 4-fluorobenzoyl azide, 4-fluorophenyl isocyanate and reactive nitrene immediate with DPPA addition at 42.5 min, and heat at 169.5 min. (zoomed 0 to 7 hours data of Fig. 8)

S3. Prompt detail and outputs

S3.1. Prompt full templates

Table S2. Prompt template

Prompt 1
Reaction: <i>"Paste your experimental protocol here"</i>
Instruction: Help me understand this reaction, and break it into phases if this is really long. After you identified the step, if the mass or volume is not given, please help me with calculating the mass/volume used for each compound. You can look up the molecular weight too. help me rewrite this procedure to a detailed stepwise procedure with the actual mass/volume to be added. Each step should only have one chemical being added.
Prompt 2
Now write a complete procedure including all the steps using the template below,

add reagents with **AddAtOnceOperation**;

add solvent if there is any using **DoseAtRateOperation**, ignore this if there is no solvent;

set temperature with **HeatCoolOperation**;

set stir rate with **StirOperation**;

Note that the duration in Heat and Stir are the ramping duration (default is 10, must > 0), so set reaction duration with **WaitOperation** and end the experiment with **EndExperimentOperation**.

Make sure to match the operation method and generate a uuid using uuid4() for each TrackingId. At the end, generate the correct <Chemicals> for all solid and liquid that are used in the procedure according to their phases. Make sure to include all elements. You can use the default value from the template if you can't find any information.

```
```xml
```

```
<Experiment
 UniqueElnId="Exp name"
 User="username"
 Project="Project name"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance" >
 <Process ProcessType="Synthesis" TrackingId="####">
 <Stages>
 <Stage StageType="Reaction" TrackingId="####">
 <Phases>
 <Phase TrackingId="####">
 <OperationSequences>
 <OperationSequence>
 <!-- add reagent liquid and solid -->
 <AddAtOnceOperation TrackingId="####">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="####" />
 <Amount Value="1" Unit="ml" /> <!-- unit is g or mg for solid -->
 </DesignValues>
 </AddAtOnceOperation>
 <!-- only use this to add solvent, rate < 10 -->
 <DoseAtRateOperation TrackingId="####" >
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="####" />
 <AutomaticParameterSet>
 <Amount Value="10" Unit="ml" /> <!-- only use ml here -->
 <Rate Value="10" Unit="ml/min" />
 </AutomaticParameterSet>
 </DesignValues>
 </DoseAtRateOperation>
 <!-- stir control, should be before heatcool-->
 <StirOperation TrackingId="####">
 <DesignValues>
 <Duration Value="10" Unit="sec" /> <!-- this is ramping duration -->
 <EndValue Value="400" Unit="rpm" />
 </DesignValues>
 </StirOperation>
 </OperationSequence>
 </Phase>
 </Phases>
 </Stage>
 </Stages>
 </Process>
</Experiment>
```

```

 </DesignValues>
 </StirOperation>
 <!-- temperature control, always use Tr for control mode-->
 <HeatCoolOperation TrackingId="####">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="20" Unit="°C" />
 <Duration Value="30" Unit="sec" /> <!-- this is ramping duration -->
 </RampByDurationParameterSet>
 </DesignValues>
 </HeatCoolOperation>
 <!-- set reaction duration -->
 <WaitOperation TrackingId="####">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="30" Unit="sec" />
 </DurationParameterSet>
 </DesignValues>
 </WaitOperation>
</OperationSequence>
</OperationSequences>
</Phase>
<!-- add second phase if need -->
<Phase TrackingId="####">
 <OperationSequences>
 <OperationSequence>
 <!-- end experiment, turn off devices -->
 <EndExperimentOperation TrackingId="####">
 <DesignValues>
 <EndConditions>
 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValues>
 </EndExperimentOperation>
 </OperationSequence>
 </OperationSequences>
</Phase>
</Phases>
</Stage>
</Stages>
</Process>
<Chemicals>
 <!--solid chemicals-->
 <Chemical Name="Chemical Solid" LotNumber="Lot1" TrackingId="####">
 <MassAmount Value="1000" Unit="g">
 <Density Value="0.500" Unit="g/ml" />
 </MassAmount>
 </Chemical>
 <!--Liquid chemicals-->
 <Chemical Name="Chemical liquid" LotNumber="Lot2" TrackingId="####">
 <VolumeAmount Value="1000" Unit="ml">
 <Density Value="0.500" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
</Chemicals>

```



```
<Substances />
</Experiment>
'''
```

## S3.2. Prompt development

### S3.2.1. Prompt 1 – additional request

The prompt design is flexible, allowing more detailed or specific instructions. For example, while the current GPT-4 model doesn't require role assignment for improved performance, assigning a specific role - like 'expert digital assistant in organic synthesis chemistry' - may enhance its output. To streamline documentation, one can request a markdown table listing all chemicals and solvents with details such as stoichiometry, molecular weight, and density. This not only aids in organizing the chemical inventory but also ensures accurate calculations. In addition, including a statement like 'no additional steps should be added to the procedure' may prevent model creativity without fine-tuning or changing model temperature. Depending on vessel capacity, incorporating a multiplying factor in this prompt could be optional. However, it is worth noting that the XML template plays a more crucial role for guiding the XML format. The instructions in prompt 1 can be tailored according to the reaction complexity to save token usage.

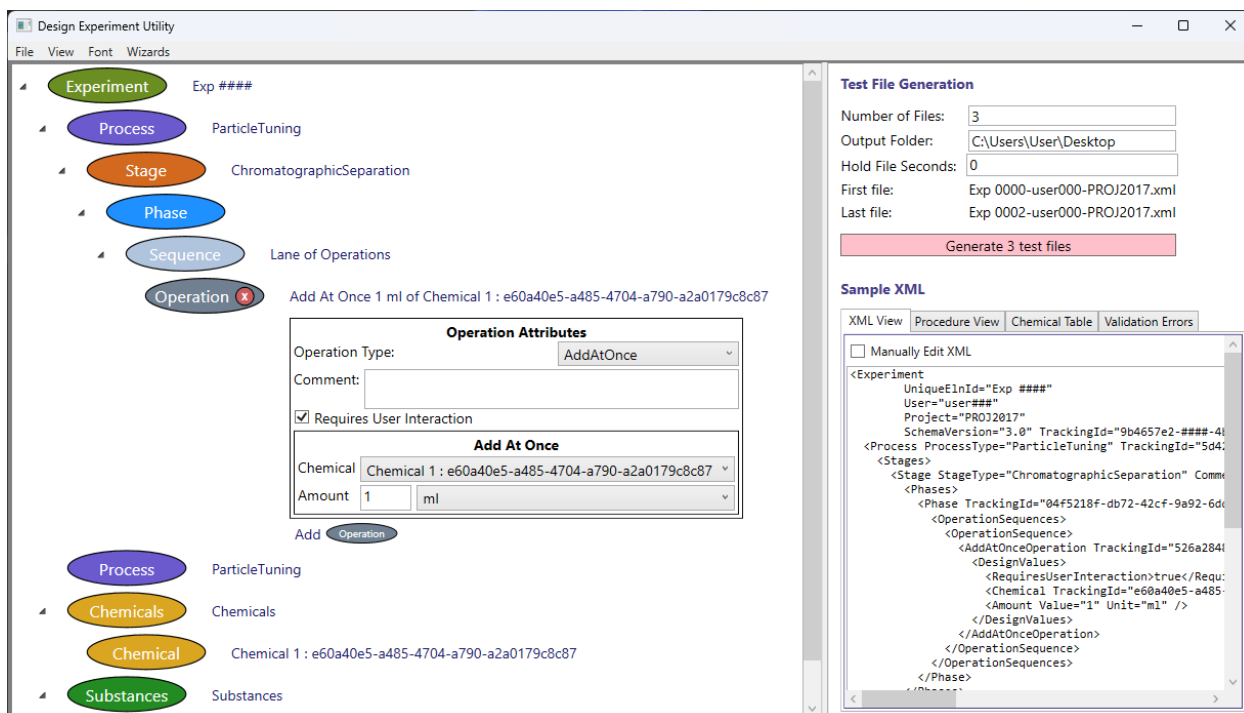
### S3.2.2. Prompt 1 – additional request example

You are an expert digital assistant in organic synthesis chemistry. Help me write a chemical materials table for the reaction below. Always look up whether a chemical is a solid or a liquid, molecular weights and densities. Always output a table in markdown format with the following columns from left to right: entry number, chemical name, phase (liquid or solid which you just looked up), molecular weight, stoichiometry, moles, weight, density (which you just looked up), volume (if a liquid), and duration of addition. The limiting reagent has a stoichiometry of 1. Calculate solvent volumes (mL) for indicated concentrations (e.g. molar = moles/L). When reagents are given as solutions, calculate the volume to add. Always show your calculations before providing the requested output because accuracy is very important.

Then, write a chemical procedure with numbered steps. The procedure should be at the level of detail of procedures in Journal of Organic Chemistry, but written as numbered steps. Do not add details to the procedure which I do not specify and use 25 Celsius as room temperature. In the procedure give quantities as mL (g, equivalents) for liquids, and g (equivalents) for solids. The equivalents in the procedure should always match the stoichiometry in the table.

### S3.2.3. Prompt 2 development

To construct and fine-tune the XML template in the second prompt, Mettler Toledo's Experiment Design Utility software<sup>1</sup> allowed us to manually draft the experimental design. Coupled with the XML Mapping Guide,<sup>2</sup> the drafted XML was used as a fundamental reference for modifying the prompt template. The Experiment Design Utility application also served as a preliminary syntax validator if importing XML. To allow autonomous scripting of iControl operations, an example operation block can be stacked to the <OperationSequence> to serve as a template. Taking <AddAtOnce> as an example in **Figure S3**, a xml script for this operation shown in the **XML View** panel on the right. Similarly, the xml syntax template of other operations or attributes can be mapped by adding to the XML draft.



**Figure S3.** Screenshot of adding operation in Mettler Toledo Design Experiment Utility software.

### S3.3. GPT-4 output

#### S3.3.1. Case study 1: SNAr aminolysis

**Table S3.** Prompt input and GPT-4 output of detailed procedure for SnAr reaction

Input
<p>Reaction: “To a 50 mL vessel fitted with condenser, nitrogen inlet, internal temperature probe, and overhead agitator were added potassium carbonate (9.55 g, 69.10mmol), biphenyl (65.3 mg, 0.4 mmol), and 4-fluoronitrobenzene 2(6.50 g, 46.10 mmol). Acetonitrile (32.5 mL, 620 mmol) was then added, and the contents of the vessel were agitated at 400 rpm. The contents of the vessel were heated to 60°C over a period of 15 min. 1-Methylpiperazine 3 (6.69 g, 7.43 mL, 66.80 mmol) was then added to the vessel in a single portion by syringe addition. The reaction mixture was sampled immediately after addition of the 1-methylpiperazine and then at further time points over the course of 8 h at 60°C to generate a reaction profile. The reaction mixture was then cooled to 20°C.”</p> <p>Instruction: Help me understand this reaction, and break it into phases if this is really long. After you identified the step, if the mass or volume is not given, please help me with calculating the mass/volume used for each compound. You can look up the molecular weight too. help me rewrite this procedure to a detailed stepwise procedure with the actual mass/volume to be added. Each step should only have one chemical being added.</p>

## Output

This chemistry procedure describes a reaction setup and execution for a chemical synthesis involving potassium carbonate, biphenyl, 4-fluoronitrobenzene, acetonitrile, and 1-methylpiperazine. To help you better understand and follow this procedure, we'll break it down into distinct phases, including setting up the reaction apparatus, combining and mixing reactants, and managing the reaction conditions over time. We'll also calculate and verify the necessary mass or volume for each compound where not provided.

### Detailed Stepwise Procedure

#### *Phase 1: Setup*

##### 1. Assemble the reaction apparatus:

- Equip a 50 mL reaction vessel with:
  - A condenser (for cooling vapors back to liquid)
  - A nitrogen inlet (to maintain an inert atmosphere)
  - An internal temperature probe (to monitor the reaction temperature)
  - An overhead agitator (for mixing)

#### *Phase 2: Adding Reactants*

##### 2. Add Potassium Carbonate:

- **Amount:** 9.55 g
- **Moles:** 69.10 mmol
- Add potassium carbonate to the reaction vessel.

##### 3. Add Biphenyl:

- **Amount:** 65.3 mg
- **Moles:** 0.4 mmol
- Add biphenyl to the reaction vessel.

##### 4. Add 4-Fluoronitrobenzene:

- **Amount:** 6.50 g
- **Moles:** 46.10 mmol
- Add 4-fluoronitrobenzene to the reaction vessel.

##### 5. Add Acetonitrile:

- **Volume:** 32.5 mL (calculated from density, assuming acetonitrile's density is approximately 0.786 g/mL at 20°C, giving a mass around 25.54 g)
- **Moles:** 620 mmol (calculated based on the molecular weight of 41.05 g/mol)
- Add acetonitrile to the reaction vessel and start the overhead agitator at 400 rpm.

#### *Phase 3: Reaction Conditions and Monitoring*

##### 6. Heat the mixture:

- Gradually heat the contents of the vessel to 60°C over a period of 15 minutes.

##### 7. Add 1-Methylpiperazine:

- **Amount:** 6.69 g
- **Volume:** 7.43 mL
- **Moles:** 66.80 mmol
- Using a syringe, add 1-methylpiperazine to the vessel in a single portion.

*Phase 4: Sampling and Cooling*

**8. Monitor the reaction:**

- Sample the reaction mixture immediately after the addition of 1-methylpiperazine.
- Continue to sample at further time points over the course of 8 hours at 60°C to generate a reaction profile.

**9. Cool the reaction mixture:**

- After completion of the reaction monitoring, cool the mixture to 20°C.

*Additional Calculations and Verifications*

For a thorough understanding and verification, you would typically check the densities and molecular weights for accurate volume calculations. However, for this setup, I've made assumptions based on typical values. Let me know if you need precise calculations using specific densities or any additional modifications to the procedure.

**Input**

see Table S2 – Prompt 2

**Output**

Creating a comprehensive XML-based procedure for the described experiment involves integrating the various steps and chemical reagents with specified operations and unique identifiers. Below is a structured XML template tailored to your provided experiment, incorporating the operations as outlined in your requirements.

```
<Experiment
 UniqueElnId="Synthesis of Biphenyl Derivative"
 User="chemuser"
 Project="Organic Synthesis Project"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance" >
 <Process ProcessType="Synthesis" TrackingId="d01b4a48-c768-4d09-a042-5fa3e2bb9b93">
 <Stages>
 <Stage StageType="Reaction" TrackingId="a4138cf2-92d8-496b-8618-1f5b8e5f23ff">
 <Phases>
 <Phase TrackingId="31769f76-f97f-4b97-b1c0-d6f998607f89">
 <OperationSequences>
 <OperationSequence>
 <AddAtOnceOperation TrackingId="f4c5c1f5-7a00-498d-b8e9-2afaff6c9eeb">
```

```

<DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="2de81288-3062-41c9-947c-3aefa47deb3b" />
 <Amount Value="9.55" Unit="g" />
</DesignValues>
</AddAtOnceOperation>
<AddAtOnceOperation TrackingId="0ff421af-d528-45b0-864e-90bb7c8795f7">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="7dd2bf64-0c5e-4d33-a7b5-370b6b14b6fa" />
 <Amount Value="65.3" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<AddAtOnceOperation TrackingId="73d56b70-6757-479f-a631-9285ab271532">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="f7c97774-8427-4f97-a179-3b5da29e7834" />
 <Amount Value="6.50" Unit="g" />
 </DesignValues>
</AddAtOnceOperation>
<DoseAtRateOperation TrackingId="4296aa01-9ee3-4ee6-af47-d2fbf55140f5">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="eb8c5a90-3f07-4d0a-88d2-204deefb6c9b" />
 <AutomaticParameterSet>
 <Amount Value="32.5" Unit="ml" />
 <Rate Value="5" Unit="ml/min" />
 </AutomaticParameterSet>
 </DesignValues>
</DoseAtRateOperation>
<StirOperation TrackingId="2b330f6e-58cc-43f1-81e2-993e6e88f14e">
 <DesignValues>
 <Duration Value="10" Unit="sec" />
 <EndValue Value="400" Unit="rpm" />
 </DesignValues>
</StirOperation>
<HeatCoolOperation TrackingId="fe1f0b7a-15d7-4643-833b-d91db06c8e47">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="60" Unit="°C" />
 <Duration Value="15" Unit="min" />
 </RampByDurationParameterSet>
 </DesignValues>
</HeatCoolOperation>
<AddAtOnceOperation TrackingId="4be6f14c-6c94-4606-97ec-bf1df7a08d92">

```

```

 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="9be6f1d8-58fa-40aa-812f-77c6ff267aae" />
 <Amount Value="6.69" Unit="g" />
 </DesignValues>
 </AddAtOnceOperation>
 <WaitOperation TrackingId="5b7c34d6-ccb7-4d86-9041-374abcfe728c">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="8" Unit="h" />
 </DurationParameterSet>
 </DesignValues>
 </WaitOperation>
</OperationSequence>
</OperationSequences>
</Phase>
<Phase TrackingId="a9874df1-6cfc-456c-9b8e-9e95571a2fbc">
 <OperationSequences>
 <OperationSequence>
 <HeatCoolOperation TrackingId="6a2e400e-47a1-43b9-b807-5b76b5a5a17e">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="20" Unit="°C" />
 <Duration Value="30" Unit="min" />
 </RampByDurationParameterSet>
 </DesignValues>
 </HeatCoolOperation>
 <EndExperimentOperation TrackingId="98965acd-3815-4579-8b13-fb8631f7c1cf">
 <DesignValues>
 <EndConditions>
 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValues>
 </EndExperimentOperation>
 </OperationSequence>
 </OperationSequences>
</Phase>
</Phases>
</Stage>
</Stages>
</Process>
<Chemicals>
 <Chemical Name="Potassium Carbonate" LotNumber="Lot K2CO3" TrackingId="2de81288-3062-41c9-947c-3aefa47deb3b">

```

```

 <MassAmount Value="9.55" Unit="g" />
 </Chemical>
 <Chemical Name="Biphenyl" LotNumber="Lot Biphenyl" TrackingId="7dd2bf64-0c5e-4d33-a7b5-370b6b14b6fa">
 <MassAmount Value="65.3" Unit="mg" />
 </Chemical>
 <Chemical Name="4-Fluoronitrobenzene" LotNumber="Lot 4FNB" TrackingId="f7c97774-8427-4f97-a179-3b5da29e7834">
 <MassAmount Value="6.50" Unit="g" />
 </Chemical>
 <Chemical Name="Acetonitrile" LotNumber="Lot MeCN" TrackingId="eb8c5a90-3f07-4d0a-88d2-204deefb6c9b">
 <VolumeAmount Value="32.5" Unit="ml" />
 </Chemical>
 <Chemical Name="1-Methylpiperazine" LotNumber="Lot 1MP" TrackingId="9be6f1d8-58fa-40aa-812f-77c6ff267aae">
 <MassAmount Value="6.69" Unit="g" />
 </Chemical>
</Chemicals>
<Substances />
</Experiment>

```

### S3.3.2. Case 2: synthesis of hydrazone

**Table S4.** Prompt input and GPT-4 output of detailed procedure for synthesis of hydrazone

Input
<p>Reaction: <i>"To a solution of 4-fluorobenzaldehyde (20.0 mmol, 1.0 equiv.) in methanol (20 mL) was added hydrazine hydrate (1.2 mL, 24 mmol, 1.2 equiv.) and the mixture stirred at r.t. for 1 h."</i></p> <p>Instruction: Help me understand this reaction, and break it into phases if this is really long. After you identified the step, if the mass or volume is not given, please help me with calculating the mass/volume used for each compound. You can look up the molecular weight too. help me rewrite this procedure to a detailed stepwise procedure with the actual mass/volume to be added. Each step should only have one chemical being added.</p>
Output
<p>This reaction involves a chemical synthesis step where 4-fluorobenzaldehyde is reacted with hydrazine hydrate to presumably form a hydrazone derivative. Here's a breakdown of the reaction and step-by-step procedure with more detailed calculations.</p> <p><b>Reaction Understanding:</b></p> <ol style="list-style-type: none"> <li><b>Reactants:</b> <ul style="list-style-type: none"> <li><b>4-fluorobenzaldehyde</b> - An aromatic aldehyde where a fluorine atom is attached to the benzene ring opposite to the aldehyde group.</li> </ul> </li> </ol>

- **Hydrazine hydrate** - Used as a nucleophilic reagent that reacts with the carbonyl group of the aldehyde to form a hydrazone linkage.
2. **Solvent:**
    - **Methanol** - Acts as the solvent in which the reaction takes place. It is chosen likely for its ability to dissolve both the reactant and the product while providing a medium for the reaction to occur.
  3. **Reaction Type:**
    - This is a nucleophilic addition where hydrazine attacks the carbon of the aldehyde group of 4-fluorobenzaldehyde, followed by a dehydration step to form the hydrazone.

### Calculations:

1. **4-Fluorobenzaldehyde:**

- Molecular weight (MW) = 124.11 g/mol
- Moles = 20.0 mmol
- Mass = Moles  $\times$  MW = 20.0 mmol  $\times$  124.11 g/mol = 2482.2 mg

2. **Hydrazine hydrate:**

- Hydrazine hydrate often contains approximately 64% hydrazine by weight.
- MW of Hydrazine (N<sub>2</sub>H<sub>4</sub>) = 32.05 g/mol
- Amount of pure hydrazine in 1.2 mL hydrazine hydrate:
  - Density of hydrazine hydrate  $\approx$  1 g/mL (approximate for calculation)
  - Mass of hydrazine hydrate in 1.2 mL = 1.2 g (assuming density is 1 g/mL)
  - Amount of pure hydrazine = 1.2 g  $\times$  64% = 0.768 g
  - Moles of hydrazine = Mass / MW = 0.768 g / 32.05 g/mol  $\approx$  0.024 mol (24 mmol)

3. **Methanol:**

- Volume given = 20 mL (no need for density calculations as the volume is already appropriate for the solvent role)

### Detailed Stepwise Procedure:

1. **Weighing the Aldehyde:**

- Weigh out 2482.2 mg of 4-fluorobenzaldehyde.

2. **Preparing the Solvent:**

- Measure 20 mL of methanol into a suitable reaction flask.

3. **Adding the Aldehyde:**

- Add the weighed 4-fluorobenzaldehyde to the methanol in the flask.

4. **Measuring Hydrazine Hydrate:**

- Measure 1.2 mL of hydrazine hydrate using a graduated syringe or pipette.



### 5. Adding Hydrazine Hydrate:

- Add the measured hydrazine hydrate to the flask containing the methanol and aldehyde solution.

### 6. Stirring the Reaction Mixture:

- Stir the mixture at room temperature (r.t.) for 1 hour. Ensure that the mixture is well mixed and that the aldehyde is fully reacted.

### 7. Monitoring the Reaction:

- After the reaction is complete, the product can be isolated by usual work-up procedures involving extraction and purification techniques, such as recrystallization or column chromatography, depending on the solubility and stability of the formed hydrazone.

This detailed procedure lays out each step explicitly with the exact amounts of each reagent and solvent used, enhancing clarity and reproducibility in the lab.

## Input

see Table S2 – Prompt 2

## Output

To create a detailed experimental procedure based on the XML template provided, here are the steps incorporating the specific operations and tracking details. The steps align with the earlier chemistry procedure and now include the technical operations required by your self-driving lab system:

```
<Experiment
 UniqueElnId="Synthesis of Hydrazone"
 User="chem_user"
 Project="Hydrazone Synthesis Project"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance" >
<Process ProcessType="Synthesis" TrackingId="7dc4f8b3-c4f2-489a-b0f7-bdfad7a2c1d3">
 <Stages>
 <Stage StageType="Reaction" TrackingId="47ec77f2-e182-4560-b2a2-d2b32ba6c173">
 <Phases>
 <Phase TrackingId="4c1cc1c9-58b2-40b5-965d-b8a67e01177c">
 <OperationSequences>
 <OperationSequence>
 <DoseAtRateOperation TrackingId="07ad67a5-14b2-43e7-b85a-0ff4a7d6782a">
 <DesignValues>
 <RequiresUserInteraction>false</RequiresUserInteraction>
 <Chemical TrackingId="d1f8fa36-4e2a-4816-bb50-32f2234e0392"/>
 </DesignValues>
 </DoseAtRateOperation>
 </OperationSequence>
 </OperationSequences>
 </Phase>
 </Phases>
 </Stage>
 </Stages>
</Process>
```

```

 <Amount Value="20" Unit="ml" />
 <Rate Value="10" Unit="ml/min" />
 </AutomaticParameterSet>
</DesignValues>
</DoseAtRateOperation>
<AddAtOnceOperation TrackingId="b94eb787-6a8f-4b24-b0d0-cc5576b23617">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="1feea5c0-9cb3-4736-a234-3b249ce0f57d" />
 <Amount Value="2482.2" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<AddAtOnceOperation TrackingId="11c00324-cf52-4b97-a975-9b0ee6c96ba5">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="e41f4859-031b-4ad8-8075-d15dd96a4321" />
 <Amount Value="1.2" Unit="ml" />
 </DesignValues>
</AddAtOnceOperation>
<StirOperation TrackingId="c3b2d507-3f92-42f7-9295-5f2d2c405dd8">
 <DesignValues>
 <Duration Value="10" Unit="sec" />
 <EndValue Value="400" Unit="rpm" />
 </DesignValues>
</StirOperation>
<HeatCoolOperation TrackingId="6b3f9c6f-f9d7-4620-97b0-cf0e0c5f8cbe">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="20" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
</HeatCoolOperation>
<WaitOperation TrackingId="8f3923db-f72b-490d-aba7-9c083ddf5e97">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="3600" Unit="sec" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
</OperationSequence>
</OperationSequences>
</Phase>
<Phase TrackingId="d2e1b2c3-eb7a-4d47-aed7-a0c6e98129f4">
 <OperationSequences>

```

```

 <OperationSequence>
 <EndExperimentOperation TrackingId="dbf32f87-45d2-4874-a4d6-acaf98092b29">
 <DesignValues>
 <EndConditions>
 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValues>
 </EndExperimentOperation>
 </OperationSequence>
 </OperationSequences>
</Phase>
</Phases>
</Stage>
</Stages>
</Process>
<Chemicals>
 <Chemical Name="4-Fluorobenzaldehyde" LotNumber="FB001" TrackingId="1feea5c0-9cb3-4736-
a234-3b249ce0f57d">
 <MassAmount Value="2482.2" Unit="mg">
 <Density Value="1.278" Unit="g/ml" /> <!-- Approximate density -->
 </MassAmount>
 </Chemical>
 <Chemical Name="Hydrazine hydrate" LotNumber="HH001" TrackingId="e41f4859-031b-4ad8-
8075-d15dd96a4321">
 <VolumeAmount Value="1.2" Unit="ml">
 <Density Value="1.000" Unit="g/ml" /> <!-- Approximate density for calculations -->
 </VolumeAmount>
 </Chemical>
 <Chemical Name="Methanol" LotNumber="M001" TrackingId="d1f8fa36-4e2a-4816-bb50-
32f2234e0392">
 <VolumeAmount Value="20" Unit="ml">
 <Density Value="0.792" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
</Chemicals>
<Substances />
</Experiment>

```

### S3.3.3. Case study 3: Curtius rearrangement

Table S5. Prompt input and GPT-4 output of detailed procedure for Curtius rearrangement

Input
<p>Reaction: “Once the system was configured, the system was tested with the Curtius rearrangement process previously highlighted and with conditions analogously described in SI section 2.d. Therefore, 10.00 mL of toluene was added to a scintillation vial equipped with a magnetic stir bar and septum filled cap (with venting needle). With the stopped-flow system pumping fresh toluene (collecting into waste), the ReactIR was blanked on toluene and these ‘in’ and ‘out’ lines were connected to the scintillation vial (adding 1.89 mL toluene to overall volume). Next, TEA (1.1 mL, 1.0 eq.) was added to the vial and allowed to circulate. With the ‘in’ line pulled out of solution, 7 (1.0361 g, 7.32 mmol, 1 eq.) was added. Once the solution appeared homogeneous, the ‘in’ line was again placed well below the liquid level. After, the NMR spectrometer and Python script were configured with experimental parameters, the stoppedflow script was started and IR spectra began to be collected every minute using the iC IR software (Mettler Toledo). Once the 5th 19F NMR spectrum started acquisition, DPPA (1.8 mL, 1.1 eq.) was added. Then, once the 15th spectrum started to be acquired, the reaction solution was lowered to a pre-heated oil bath (80 °C) as significant conversion of 7 to the corresponding 10 was observed. 12 (0.84 mL, 1.1 eq.) was added to the reaction solution once the 135th spectrum started to be acquired as significant conversion from the 10 to the corresponding isocyanate (11) was observed. The reaction was then monitored for the remaining time points as denoted with the script. For this trial, 19F NMR spectra were acquired with four transients, zero dummy scans, 90° pulse, 44 dB rg, 61 s d1, -90.0 ppm o1p, 219.6 ppm SW, and 2.79 s AQ (34,816 complex data points). In the Python script, 170 timepoints were collected every 510 s, with a 15 s pause time, 1.78 mL/min calibrated pump rate, and an additional t=0 s spectrum collected.”</p> <p>Instruction: Help me understand this reaction, and break it into phases if this is really long. After you identified the step, if the mass or volume is not given, please help me with calculating the mass/volume used for each compound. You can look up the molecular weight too. help me rewrite this procedure to a detailed stepwise procedure with the actual mass/volume to be added. Each step should only have one chemical being added.</p>
Output
<p>Let's break down the Curtius rearrangement process described and rewrite it into a detailed, step-by-step procedure with clear instructions for adding reagents, stirring, heating, and other operations.</p> <p><b>Reagents and Conditions:</b></p> <ul style="list-style-type: none"><li>• <b>Toluene</b> (C<sub>7</sub>H<sub>8</sub>)</li><li>• <b>Triethylamine (TEA)</b> (C<sub>6</sub>H<sub>15</sub>N), MW = 101.19 g/mol, density = 0.726 g/mL</li><li>• <b>Compound 7</b> (MW provided)</li></ul>

- **DPPA (Diphenylphosphoryl azide)** (C<sub>12</sub>H<sub>10</sub>N<sub>3</sub>O, MW = 239.24 g/mol, density = 1.31 g/mL)
- **Compound 12** (MW provided)
- **NMR Spectrometer**
- **Stopped-flow system**
- **ReactIR with iC IR software**

### Detailed Step-by-Step Procedure:

#### *Step 1: Preparation*

##### 1. System Configuration:

- Ensure the stopped-flow system is configured correctly.
- Prepare the NMR spectrometer and Python script with the required experimental parameters.

#### *Step 2: Setup of Initial Reaction Solution*

##### 2. Add Toluene:

- Measure and add 10.00 mL of toluene to a scintillation vial equipped with a magnetic stir bar and a septum-filled cap with a venting needle.

##### 3. ReactIR Blanking:

- Use the stopped-flow system to pump fresh toluene (which will be collected into waste).
- Blank the ReactIR on toluene.
- Connect the 'in' and 'out' lines of the ReactIR to the scintillation vial, adding 1.89 mL of toluene to the overall volume (final total: 11.89 mL).

#### *Step 3: Addition of Triethylamine (TEA)*

##### 4. Add TEA:

- Measure 1.1 mL of triethylamine (TEA) (0.726 g/mL \* 1.1 mL = 0.7986 g) and add it to the vial.
- Allow the solution to circulate.

#### *Step 4: Addition of Compound 7*

##### 5. Add Compound 7:

- Remove the 'in' line from the solution.
- Measure 1.0361 g of Compound 7 (7.32 mmol, MW = 141.56 g/mol).
- Add Compound 7 to the vial.
- Ensure the solution appears homogeneous.
- Place the 'in' line back below the liquid level.

*Step 5: Start Data Collection and Addition of DPPA*

**6. Start Stopped-Flow Script:**

- Begin the stopped-flow script.
- Collect IR spectra every minute using the iC IR software.

**7. Add DPPA:**

- After the 5th  $^{19}\text{F}$  NMR spectrum starts acquisition, measure and add 1.8 mL of DPPA ( $1.8 \text{ mL} * 1.31 \text{ g/mL} = 2.358 \text{ g}$ ).

*Step 6: Heating and Further Additions*

**8. Heat Reaction Solution:**

- After the 15th spectrum starts acquisition, lower the reaction solution into a pre-heated oil bath set at  $80 \text{ }^\circ\text{C}$ .
- Observe significant conversion of Compound 7 to Compound 10.

**9. Add Compound 12:**

- Once the 135th spectrum starts to be acquired, measure and add 0.84 mL of Compound 12 (1.1 eq.).

*Step 7: Monitoring*

**10. Monitor Reaction:**

- Continue monitoring the reaction for the remaining time points as per the script.

**NMR Spectra Acquisition Parameters:**

- **Number of Transients:** 4
- **Dummy Scans:** 0
- **Pulse Angle:**  $90^\circ$
- **Receiver Gain (rg):** 44 dB
- **Delay Time (d1):** 61 s
- **Carrier Frequency (o1p):** -90.0 ppm
- **Sweep Width (SW):** 219.6 ppm
- **Acquisition Time (AQ):** 2.79 s (34,816 complex data points)
- **Python Script Parameters:**
  - Collect 170 time points every 510 s.
  - Pause time: 15 s
  - Calibrated pump rate: 1.78 mL/min
  - Additional spectrum at  $t=0$  s.

This procedure outlines each step clearly, ensuring all necessary details are included for accurate execution of the reaction.

**Input (additional request)**

can you estimate the reaction time, according to the number of NMR spectrum, and the sampling interval, and update the procedure.

## Output (additional request)

Certainly! Let's estimate the reaction time based on the NMR spectra collected and the sampling interval.

### Parameters given:

- $^{19}\text{F}$  NMR spectra collected: 170 timepoints
- Sampling interval: 510 s (8.5 minutes)

### Calculation:

- Total reaction time = Number of spectra  $\times$  Sampling interval
- Total reaction time = 170 spectra  $\times$  510 s/spectrum = 86,700 s

Convert the total time to hours:

- Total reaction time in hours = 86,700 s / 3600 s/hr = 24.08 hours

Now, let's update the procedure to include this estimated reaction time.

## Stepwise Procedure with Reaction Time

### Phase 1: Setup and Initial Solution Preparation

1. Configure the system as described in SI section 2.d. Ensure the system is ready for the Curtius rearrangement process.
2. Add 10.00 mL of toluene to a scintillation vial.
3. Equip the vial with a magnetic stir bar and a septum-filled cap with a venting needle.
4. With the stopped-flow system pumping fresh toluene (collecting into waste), blank the ReactIR on toluene.
5. Connect the 'in' and 'out' lines of the ReactIR to the scintillation vial, adding an additional 1.89 mL of toluene (total volume now 11.89 mL).

### Phase 2: Reagent Addition and Circulation

6. Add 1.1 mL of TEA (Triethylamine) to the vial and allow it to circulate.
7. Pull the 'in' line out of the solution.
8. Add 1.0361 g of compound 7 (7.32 mmol) to the vial.
9. Once the solution appears homogeneous, place the 'in' line back well below the liquid level.

### Phase 3: Data Collection and Reaction Monitoring

10. Configure the NMR spectrometer and Python script with the experimental parameters.

11. Start the stopped-flow script and begin collecting IR spectra every minute using the iC IR software.

#### Phase 4: Sequential Reagent Addition During Monitoring

12. Once the 5th  $^{19}\text{F}$  NMR spectrum starts acquisition (approximately 42.5 minutes from the start), add 1.8 mL of DPPA (1.1 eq.).
13. Once the 15th spectrum starts to be acquired (approximately 2.13 hours from the start), lower the reaction solution to a pre-heated oil bath at 80 °C, as significant conversion of compound 7 to compound 10 is observed.
14. Once the 135th spectrum starts to be acquired (approximately 19.13 hours from the start), add 0.84 mL of compound 12 (1.1 eq.), as significant conversion from compound 10 to the corresponding isocyanate (compound 11) is observed.
15. Continue to monitor the reaction for the remaining time points as denoted in the script.

#### Estimated Total Reaction Time:

- The reaction is monitored over a total period of approximately 24.08 hours.

#### NMR Data Acquisition Parameters:

- $^{19}\text{F}$  NMR spectra were acquired with four transients, zero dummy scans, 90° pulse, 44 dB rg, 61 s d1, -90.0 ppm o1p, 219.6 ppm SW, and 2.79 s AQ (34,816 complex data points).
- 170 timepoints were collected every 510 s, with a 15 s pause time, 1.78 mL/min calibrated pump rate, and an additional t=0 s spectrum collected.

This procedure should now provide a clear, step-by-step guide including the estimated reaction time based on the NMR sampling intervals.

#### Input

see Table S2 – Prompt 2

#### Output

```
<Experiment
 UniqueElnId="Curtius Rearrangement"
 User="Ivory"
 Project="Self-Driving Lab Synthesis"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance">
<Process ProcessType="Synthesis" TrackingId="f685c72e-14d2-4638-ba12-fbe2466b6bb0">
 <Stages>
 <Stage StageType="Reaction" TrackingId="8819c96b-5bfc-4979-92f8-0189e902f4c3">
 <Phases>
```



```

<Phase TrackingId="0c58e7f4-b887-4278-bf09-7f34cbb1ac56">
 <OperationSequences>
 <OperationSequence>
 <!-- Adding Toluene as solvent -->
 <DoseAtRateOperation TrackingId="0451fdce-5de2-4c59-95dc-4f576ef99323">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="dc8b2a8d-1d9a-4bc5-aae4-81e1a4a55c91" />
 <AutomaticParameterSet>
 <Amount Value="11.89" Unit="ml" />
 <Rate Value="5" Unit="ml/min" />
 </AutomaticParameterSet>
 </DesignValues>
 </DoseAtRateOperation>
 <!-- Adding TEA -->
 <AddAtOnceOperation TrackingId="cc3320fa-c662-4cdf-a654-a9b3e3b398aa">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="9fe2b996-4e28-4c21-97e3-5e0c4ec9253a" />
 <Amount Value="1.1" Unit="ml" />
 </DesignValues>
 </AddAtOnceOperation>
 <!-- Adding compound 7 -->
 <AddAtOnceOperation TrackingId="f5c2e6bc-9078-4d0e-9246-d55fcbd1b3ec">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="aef9487-632c-47f4-aa9b-4cefa8f5b491" />
 <Amount Value="1.0361" Unit="g" />
 </DesignValues>
 </AddAtOnceOperation>
 <!-- Stirring at 400 rpm -->
 <StirOperation TrackingId="6588c5dd-5b8a-4b2d-937e-a789c6c46694">
 <DesignValues>
 <Duration Value="10" Unit="sec" />
 <EndValue Value="400" Unit="rpm" />
 </DesignValues>
 </StirOperation>
 <!-- Heating to 80°C -->
 <HeatCoolOperation TrackingId="af4dbf2f-f4f6-4608-b8e5-cc7a1794499f">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="80" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
 </HeatCoolOperation>
 </OperationSequence>
 </OperationSequences>
</Phase>

```

```

</HeatCoolOperation>
<!-- Waiting for reaction to proceed to the next addition -->
<WaitOperation TrackingId="1f790d80-b6d6-43af-b0ea-82776a711dc5">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="127.5" Unit="min" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<!-- Adding DPPA after observing significant conversion -->
<AddAtOnceOperation TrackingId="cb34f8e4-1d34-4e1b-b9b3-5fa5e55a6e94">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="b3f0d86a-78e9-47b8-85d6-d993bea56e49" />
 <Amount Value="1.8" Unit="ml" />
 </DesignValues>
</AddAtOnceOperation>
<!-- Waiting for reaction to complete before adding Compound 12 -->
<WaitOperation TrackingId="61d0e094-1f78-447b-8b82-a5c9146ebfcd">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="17.5" Unit="h" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<!-- Adding Compound 12 -->
<AddAtOnceOperation TrackingId="4e212d4d-b25b-4a07-845a-df8dcd6113b2">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="d930dbf3-a3ed-4703-9c8a-3433155d542b" />
 <Amount Value="0.84" Unit="ml" />
 </DesignValues>
</AddAtOnceOperation>
<!-- Waiting for the final reaction to complete -->
<WaitOperation TrackingId="ef8f2303-2023-4d69-b8cb-2e42b0d4e413">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="4.5" Unit="h" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<!-- Ending the experiment -->
<EndExperimentOperation TrackingId="5b839654-dc17-47d0-b073-9c6c8aabc291">
 <DesignValues>
 <EndConditions>

```

```

 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValues>
 </EndExperimentOperation>
 </OperationSequence>
 </OperationSequences>
 </Phase>
 </Phases>
 </Stage>
 </Stages>
 </Process>
 <Chemicals>
 <!-- Toluene as liquid -->
 <Chemical Name="Toluene" LotNumber="Tol-001" TrackingId="dc8b2a8d-1d9a-4bc5-aae4-81e1a4a55c91">
 <VolumeAmount Value="11.89" Unit="ml">
 <Density Value="0.866" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
 <!-- TEA as liquid -->
 <Chemical Name="Triethylamine" LotNumber="TEA-102" TrackingId="9fe2b996-4e28-4c21-97e3-5e0c4ec9253a">
 <VolumeAmount Value="1.1" Unit="ml">
 <Density Value="0.726" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
 <!-- Compound 7 as solid -->
 <Chemical Name="Compound 7" LotNumber="C7-019" TrackingId="aeef9487-632c-47f4-aa9b-4cefa8f5b491">
 <MassAmount Value="1.0361" Unit="g">
 <Density Value="1.2" Unit="g/ml" />
 </MassAmount>
 </Chemical>
 <!-- DPPA as liquid -->
 <Chemical Name="Diphenylphosphoryl azide" LotNumber="DPPA-301" TrackingId="b3f0d86a-78e9-47b8-85d6-d993bea56e49">
 <VolumeAmount Value="1.8" Unit="ml">
 <Density Value="1.204" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
 <!-- Compound 12 as liquid -->
 <Chemical Name="Compound 12" LotNumber="C12-045" TrackingId="d930dbf3-a3ed-4703-9c8a-3433155d542b">
 <VolumeAmount Value="0.84" Unit="ml">

```

```

 <Density Value="1.000" Unit="g/ml" /> <!-- Assumed density -->
 </VolumeAmount>
</Chemical>
</Chemicals>
<Substances />
</Experiment>

```

### S3.3.4. Length limit study: C-N cross-coupling

**Table S6.** Prompt input and GPT-4 output of detailed procedure for C-N cross-coupling

Input
<p>Reaction: “To a screw-cap vial equipped with a stir bar were added benzylamine (4a, 0.65 mmol, 1.3 equiv), anhydrous 1,4-Dioxane (2.0 mL), 5,6-dichloropyrazine-2,3-dicarbonitrile (1, 0.60 mmol, 1.2 equiv, C<sub>6</sub>Cl<sub>2</sub>N<sub>4</sub>), and anhydrous K<sub>3</sub>PO<sub>4</sub> powder (2.5 mmol, 5.0 equiv). The reaction mixture was heated at 50 °C for 2 h during which time the amine and 5,6-dichloropyrazine-2,3-dicarbonitrile were converted to the S<sub>N</sub>Ar adduct. Next, DMSO (4.0 ml) and phenol (3a, 0.50 mmol, 1.0 equiv) were added, and the resulting mixture was heated at 100 °C for 30 minutes to induce the Smiles rearrangement. The reaction mixture was cooled to room temperature and treated with AcOH (6.0 mL) and zinc (5.00 mmol, 10.0 equiv). The resulting mixture was heated at 80 °C for 30 minutes to ensure full reductive cleavage of the rearranged intermediate.”</p> <p>Instruction: Help me understand this reaction, and break it into phases if this is really long. After you identified the step, if the mass or volume is not given, please help me with calculating the mass/volume used for each compound. You can look up the molecular weight too. help me rewrite this procedure to a detailed stepwise procedure with the actual mass/volume to be added. Each step should only have one chemical being added.</p>
Output
<p>To help you understand and rewrite this reaction procedure, I will break it down into distinct phases, calculate the required masses and volumes based on molar equivalents and molecular weights, and rewrite the steps accordingly.</p> <p><b>Phase 1: Formation of S<sub>N</sub>Ar Adduct</b></p> <p><b>Chemicals Involved:</b></p> <ul style="list-style-type: none"> <li>• <b>Benzylamine</b> (molecular weight: 107.16 g/mol)</li> <li>• <b>5,6-dichloropyrazine-2,3-dicarbonitrile</b> (molecular weight: 230.02 g/mol)</li> <li>• <b>Anhydrous K<sub>3</sub>PO<sub>4</sub></b> (molecular weight: 212.27 g/mol)</li> <li>• <b>1,4-Dioxane</b></li> </ul> <p><b>Reaction Steps:</b></p> <ol style="list-style-type: none"> <li>1. Prepare a screw-cap vial and equip it with a magnetic stir bar.</li> <li>2. Calculate and add 0.65 mmol of benzylamine.</li> <li>3. Add 2.0 mL of anhydrous 1,4-dioxane.</li> <li>4. Calculate and add 0.60 mmol of 5,6-dichloropyrazine-2,3-dicarbonitrile.</li> </ol>

5. Calculate and add 2.5 mmol of anhydrous K<sub>3</sub>PO<sub>4</sub>.
6. Heat the reaction mixture to 50 °C and stir for 2 hours.

### Phase 2: Smiles Rearrangement

#### Chemicals Involved:

- **DMSO**
- **Phenol** (molecular weight: 94.11 g/mol)

**Reaction Steps:** 7. Add 4.0 mL of DMSO. 8. Calculate and add 0.50 mmol of phenol. 9. Heat the reaction mixture to 100 °C and stir for 30 minutes.

### Phase 3: Reductive Cleavage

#### Chemicals Involved:

- **AcOH (Acetic Acid)**
- **Zinc** (molecular weight: 65.38 g/mol)

**Reaction Steps:** 10. Cool the reaction mixture to room temperature. 11. Add 6.0 mL of AcOH. 12. Calculate and add 5.00 mmol of zinc. 13. Heat the reaction mixture to 80 °C and stir for 30 minutes.

### Calculations for Masses and Volumes

For clarity and to rewrite the procedure with the actual mass/volume to be added, I will calculate the necessary amounts based on the molar amounts and molecular weights provided:

1. **Benzylamine:**  $0.65 \text{ mmol} \times 107.16 \text{ g/mol} = 69.65 \text{ mg}$
2. **5,6-dichloropyrazine-2,3-dicarbonitrile:**  $0.60 \text{ mmol} \times 230.02 \text{ g/mol} = 138.01 \text{ mg}$
3. **Anhydrous K<sub>3</sub>PO<sub>4</sub>:**  $2.5 \text{ mmol} \times 212.27 \text{ g/mol} = 530.67 \text{ mg}$
4. **Phenol:**  $0.50 \text{ mmol} \times 94.11 \text{ g/mol} = 47.05 \text{ mg}$
5. **Zinc:**  $5.00 \text{ mmol} \times 65.38 \text{ g/mol} = 326.90 \text{ mg}$

### Rewritten Detailed Procedure

1. Equip a screw-cap vial with a magnetic stir bar.
2. Add 69.65 mg of benzylamine to the vial.
3. Add 2.0 mL of anhydrous 1,4-dioxane.
4. Add 138.01 mg of 5,6-dichloropyrazine-2,3-dicarbonitrile.
5. Add 530.67 mg of anhydrous K<sub>3</sub>PO<sub>4</sub>.
6. Heat the vial to 50 °C and stir for 2 hours.
7. Add 4.0 mL of DMSO.
8. Add 47.05 mg of phenol.
9. Heat the vial to 100 °C and stir for 30 minutes.
10. Cool the vial to room temperature.
11. Add 6.0 mL of AcOH.
12. Add 326.90 mg of zinc.

13. Heat the vial to 80 °C and stir for 30 minutes.

This detailed stepwise procedure ensures each chemical addition is clear and the amounts are precise, aiding in the reproducibility of the experiment.

### Input

see Table S2 – Prompt 2

### Output

```
<Experiment
 UniqueElnId="Synthesis of SNAr Adduct and Smiles Rearrangement"
 User="Ivory"
 Project="Synthetic Chemistry Research"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance" >
<Process ProcessType="Synthesis" TrackingId="b1ed56af-75a2-489e-829a-4223dbde73a3">
 <Stages>
 <Stage StageType="Reaction" TrackingId="06f292cf-4e57-498c-bc85-7228c37f181c">
 <Phases>
 <Phase TrackingId="10834a42-c6d7-4ea3-b45e-ebf53698d176">
 <OperationSequences>
 <OperationSequence>
 <!-- Initial reagents addition -->
 <AddAtOnceOperation TrackingId="fa1806f2-0470-4644-bf3f-d8e3ceea923f">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="2c0f77ef-d197-43ab-a056-4b7625f9ddec" />
 <Amount Value="69.65" Unit="mg" />
 </DesignValues>
 </AddAtOnceOperation>
 <DoseAtRateOperation TrackingId="a17e5bda-365d-4a16-9ec1-09775239cafe">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="e3a47b34-6c7e-4a02-9468-cedf78a899f1" />
 <AutomaticParameterSet>
 <Amount Value="2.0" Unit="ml" />
 <Rate Value="5" Unit="ml/min" />
 </AutomaticParameterSet>
 </DesignValues>
 </DoseAtRateOperation>
 <AddAtOnceOperation TrackingId="c2dd1c2d-f65d-4d8e-a806-3926143ecfe1">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
```

```

 <Chemical TrackingId="591ad507-87f7-42b2-ba9c-5c0932c1fb5d" />
 <Amount Value="138.01" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<AddAtOnceOperation TrackingId="d1e0f693-075c-49ec-b7e5-1d2b8729ba51">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="6dd9ef3b-03de-48b7-9f04-84bd38ea90c8" />
 <Amount Value="530.67" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<StirOperation TrackingId="9f01ecb1-ba7f-420f-9d08-92cf6bda1a22">
 <DesignValues>
 <Duration Value="10" Unit="sec" />
 <EndValue Value="400" Unit="rpm" />
 </DesignValues>
</StirOperation>
<HeatCoolOperation TrackingId="716aadb6-f820-4c8f-a7e7-64c6f2ba9c1d">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="50" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
</HeatCoolOperation>
<WaitOperation TrackingId="2e7257d8-ff08-42c9-9480-cdc0b31cd93b">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="7200" Unit="sec" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<!-- Smiles Rearrangement -->
<AddAtOnceOperation TrackingId="c6d1d012-768e-4047-b9e4-3e5e15827d80">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="7f22e2cf-f718-4379-b8db-77e5e0c6c758" />
 <Amount Value="47.05" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<DoseAtRateOperation TrackingId="06d2043a-729f-4b7b-873f-2e585666c7c7">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="08f1ac5a-ee07-44cc-bded-ef8eae8def4c" />
 <AutomaticParameterSet>

```

```

 <Amount Value="4.0" Unit="ml" />
 <Rate Value="5" Unit="ml/min" />
 </AutomaticParameterSet>
</DesignValues>
</DoseAtRateOperation>
<HeatCoolOperation TrackingId="a5806f77-6ef4-49c5-8f54-eeb0b1a36e09">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="100" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
</HeatCoolOperation>
<WaitOperation TrackingId="5d4872fb-ffdb-4985-bbfe-0aabb3d29488">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="1800" Unit="sec" />
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<HeatCoolOperation TrackingId="bef42ec3-3ae1-43e1-a995-df6c0128e333">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="25" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
</HeatCoolOperation>
<WaitOperation TrackingId="c0b9e3b4-771f-4b7a-a8ee-5ce3213a6747">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="600" Unit="sec" /> <!-- Allow 10 minutes for cooling
-->
 </DurationParameterSet>
 </DesignValues>
</WaitOperation>
<!-- Reductive Cleavage -->
<AddAtOnceOperation TrackingId="9bea3f7c-d86e-4f29-bc56-97c33f5b9345">
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="0ec8efee-8825-4688-bce6-d2986d0f8c34" />
 <Amount Value="326.90" Unit="mg" />
 </DesignValues>
</AddAtOnceOperation>
<DoseAtRateOperation TrackingId="b46feca4-e1ed-409d-a5b6-9c3c1e3c3e76">

```



```

 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="d18fd9be-bdfd-4b8e-a6a4-44df2b9639fd" />
 <AutomaticParameterSet>
 <Amount Value="6.0" Unit="ml" />
 <Rate Value="5" Unit="ml/min" />
 </AutomaticParameterSet>
 </DesignValues>
 </DoseAtRateOperation>
 <HeatCoolOperation TrackingId="aa4d821a-3522-4aeb-afcc-fc692c641a1c">
 <DesignValues>
 <RampByDurationParameterSet TemperatureControlMode="Tr">
 <EndValue Value="80" Unit="°C" />
 <Duration Value="10" Unit="sec" />
 </RampByDurationParameterSet>
 </DesignValues>
 </HeatCoolOperation>
 <WaitOperation TrackingId="4c1bc292-6897-4b5f-835f-53c9926a1b9d">
 <DesignValues>
 <DurationParameterSet>
 <Duration Value="1800" Unit="sec" />
 </DurationParameterSet>
 </DesignValues>
 </WaitOperation>
 <!-- End of experiment -->
 <EndExperimentOperation TrackingId="734af91d-1c5b-4771-abe5-c9760e67c540">
 <DesignValues>
 <EndConditions>
 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValues>
 </EndExperimentOperation>
</OperationSequence>
</OperationSequences>
</Phase>
</Phases>
</Stage>
</Stages>
</Process>
<Chemicals>
 <Chemical Name="Benzylamine" LotNumber="LotBZA" TrackingId="2c0f77ef-d197-43ab-a056-4b7625f9ddec">
 <MassAmount Value="69.65" Unit="mg">
 <Density Value="1.084" Unit="g/ml" />

```

```
</MassAmount>
</Chemical>
<Chemical Name="1,4-Dioxane" LotNumber="LotDIO" TrackingId="e3a47b34-6c7e-4a02-9468-
cedf78a899f1">
 <VolumeAmount Value="2.0" Unit="ml">
 <Density Value="1.033" Unit="g/ml" />
 </VolumeAmount>
</Chemical>
<Chemical Name="5,6-dichloropyrazine-2,3-dicarbonitrile" LotNumber="LotPYZ"
TrackingId="591ad507-87f7-42b2-ba9c-5c0932c1fb5d">
 <MassAmount Value="138.01" Unit="mg">
 <Density Value="1.6" Unit="g/ml" />
 </MassAmount>
</Chemical>
<Chemical Name="Anhydrous K3P04" LotNumber="LotK3P" TrackingId="6dd9ef3b-03de-48b7-9f04-
84bd38ea90c8">
 <MassAmount Value="530.67" Unit="mg">
 <Density Value="2.338" Unit="g/ml" />
 </MassAmount>
</Chemical>
<Chemical Name="Phenol" LotNumber="LotPHN" TrackingId="7f22e2cf-f718-4379-b8db-
77e5e0c6c758">
 <MassAmount Value="47.05" Unit="mg">
 <Density Value="1.071" Unit="g/ml" />
 </MassAmount>
</Chemical>
<Chemical Name="DMSO" LotNumber="LotDMS" TrackingId="08f1ac5a-ee07-44cc-bded-
ef8eae8def4c">
 <VolumeAmount Value="4.0" Unit="ml">
 <Density Value="1.100" Unit="g/ml" />
 </VolumeAmount>
</Chemical>
<Chemical Name="Zinc" LotNumber="LotZn" TrackingId="0ec8efee-8825-4688-bce6-
d2986d0f8c34">
 <MassAmount Value="326.90" Unit="mg">
 <Density Value="7.14" Unit="g/ml" />
 </MassAmount>
</Chemical>
<Chemical Name="AcOH" LotNumber="LotAcOH" TrackingId="d18fd9be-bdfd-4b8e-a6a4-
44df2b9639fd">
 <VolumeAmount Value="6.0" Unit="ml">
 <Density Value="1.049" Unit="g/ml" />
 </VolumeAmount>
</Chemical>
</Chemicals>
```

```
<Substances />
</Experiment>
```

### S3.3.5. Prompt 2 – Python version

**Table S7.** Prompt input and GPT-4 output of complete procedure for synthesis of hydrazone in Python using self driving solubility SDL.

Input
<p>Now, following the steps above strictly, write a complete procedure using the template below,</p> <pre>class CrystalizationDeck:     def add_reagent(self, chemical: str, amount: float, units: str):     def heat_stir(self, temperature: float, stir_rate: int, duration_in_miniute: int):     def add_solvent(self, solvent: str):  if __name__ == "__main__":     deck = CrystalizationDeck()     # script action here     deck.add_reagent("methanol")</pre> <p>add reagents with <code>add_reagent(self, chemical: str, amount: float, units: str);</code> add solvent if there is any using <code>add_solvent(self, solvent: str, volume_in_ml:float)</code>, ignore this if there is no solvent; set temperature and stir rate with <code>heat_stir(self, temperature: float, stir_rate: int, duration_in_miniute: int)</code></p>
Output
<p>Certainly! Below is a Python script that simulates the described chemical synthesis procedure using a class named <code>CrystalizationDeck</code>. The class is designed with methods to add reagents, add solvent, and control the reaction conditions such as temperature and stirring. This script assumes the <code>CrystalizationDeck</code> class methods are properly implemented elsewhere in your project.</p> <pre>if __name__ == "__main__":     deck = CrystalizationDeck()      # Add solvent to the reaction flask     deck.add_solvent("methanol", 20)      # Add 4-fluorobenzaldehyde     deck.add_reagent("4-fluorobenzaldehyde", 2802.2, "mg")</pre>

```

Add hydrazine hydrate
deck.add_reagent("hydrazine hydrate", 1.2, "mL")

Set the reaction conditions
deck.heat_stir(temperature=25, stir_rate=300, duration_in_minutes=60)

```

This script does the following:

1. **Adds solvent:** Methanol is added to the reaction flask.
2. **Adds reagents:** Both 4-fluorobenzaldehyde and hydrazine hydrate are added sequentially to the methanol in the specified amounts and units.
3. **Controls reaction conditions:** The reaction mixture is stirred at room temperature (assumed to be around 25°C) with a stir rate of 300 RPM for 60 minutes.

This setup will effectively simulate the chemical synthesis procedure you outlined, assuming you have the operational details and hardware to handle such tasks in your setup.

**Table S8.** Prompt input and GPT-4 output of complete procedure for synthesis of hydrazone in python using Kinova 3 SDL.

Input
<p>Now, following the steps above strictly, write a complete procedure using the template below,</p> <pre> kinova_shaker = Shaker(name='kinova shaker') europa = EuropaHandling(name='europa')  acetone = Liquid(name='acetone', formula='C3H6O') salt = Solid(name='salt', formula=NaCl)  # vial tray vials vial_1 = Vial(name='vial 1') vial_2 = Vial(name='vial 2')  # solvent tray vials acetone_solvent_vial = Vial(name='acetone solvent', chemical=acetone)  # script action here # only new liquids, solids and solvent tray vials can be added  add_solvent(europa, acetone_solvent_vial, vial_1, 10) </pre>

add solid reagents with the function `add_reagent(reagent: Solid, vial_to_dose_to: Vial, amount: float, units: str);`

add solvent with the function `add_solvent(handler: EuropaHandling, solvent_vial: Vial, vial_to_dispense_to: Vial, volume_in_ml: float),` ignore this if there is no solvent;

start to heat and stirring with the function `heat_and_stir(shaker: Shaker, vial: Vial, duration: int, temperature: float, stir_rate: int);`

## Output

Here is the complete procedure rewritten using the provided template:

```
kinova_shaker = Shaker(name='kinova shaker')
europa = EuropaHandling(name='europa')

methanol = Liquid(name='methanol', formula='CH3OH')
fluorobenzaldehyde = Solid(name='4-fluorobenzaldehyde', formula='C7H5FO')
hydrazine_hydrate = Liquid(name='hydrazine hydrate', formula='N2H4.H2O')

vial tray vials
vial_1 = Vial(name='vial 1')
vial_2 = Vial(name='vial 2')

solvent tray vials
methanol_solvent_vial = Vial(name='methanol solvent', chemical=methanol)
hydrazine_solvent_vial = Vial(name='hydrazine solvent',
chemical=hydrazine_hydrate)

script action here

Step 1: Add methanol solvent to vial 1
add_solvent(europa, methanol_solvent_vial, vial_1, 20)

Step 2: Add 4-fluorobenzaldehyde to the methanol in vial 1
add_reagent(fluorobenzaldehyde, vial_1, 2.4822, 'g')

Step 3: Add hydrazine hydrate to the reaction mixture in vial 1
add_solvent(europa, hydrazine_solvent_vial, vial_1, 1.2)

Step 4: Stir the reaction mixture at room temperature for 1 hour
heat_and_stir(kinova_shaker, vial_1, duration=60, temperature=25,
stir_rate=300)
```

In this script:

- `add_solvent` is used to add methanol and hydrazine hydrate to the vial.
- `add_reagent` is used to add 4-fluorobenzaldehyde to the vial.
- `heat_and_stir` is used to stir the reaction mixture for 1 hour at room temperature.

You can adjust the `stir_rate` parameter according to your specific requirements for stirring speed.

**Table S9.** Prompt input and GPT-4 output of complete procedure for synthesis of hydrazone in Python using XML generator script.

### Input

Now, following the steps above strictly, write a complete procedure using the template below,

```
class ExampleXmlGen:
 def dose_at_rate(self, chemical_formula: str, amount_in_ml: float,
rate_ml_per_min: float = 5):
 def add_at_once(self, chemical_formula: str, amount: float, unit: str =
"ml"):
 def stir(self, rate: str = 400, ramping_duration: float = 10):
 def heat_cool(self, temperature:float, ramping_duration_in_second: float =
10):
 def wait(self, duration_in_minute: float):

if __name__ == "__main__":
 gen = ExampleXmlGen()
 # script action here
 gen.dose_at_rate("methanol", 10)
```

add reagents with `add_at_once(self, chemical_formula: str, amount: float, unit: str = "ml")`

add solvent if there is any using `dose_at_rate(self, chemical_formula: str, amount_in_ml: float, rate_ml_per_min: float = 5)`, ignore this if there is no solvent;

set temperature with `heat_cool(self, temperature:float, ramping_duration_in_second: float = 10)` and stir rate with `stir(self, rate: str = 400, ramping_duration: float = 10)`

set reaction wait time using `wait(self, duration_in_minute: float)`

## Output

Here is the complete procedure using the provided template:

```
class ExampleXmlGen:
 def dose_at_rate(self, chemical_formula: str, amount_in_ml: float,
rate_ml_per_min: float = 5):
 pass # Add implementation here

 def add_at_once(self, chemical_formula: str, amount: float, unit: str =
"ml"):
 pass # Add implementation here

 def stir(self, rate: str = 400, ramping_duration: float = 10):
 pass # Add implementation here

 def heat_cool(self, temperature: float, ramping_duration_in_second: float =
10):
 pass # Add implementation here

 def wait(self, duration_in_minute: float):
 pass # Add implementation here

if __name__ == "__main__":
 gen = ExampleXmlGen()

 # Add solvent (methanol)
 gen.dose_at_rate("methanol", 20, rate_ml_per_min=5)

 # Add 4-fluorobenzaldehyde
 gen.add_at_once("4-fluorobenzaldehyde", 2.4822, unit="g")

 # Stir to dissolve
 gen.stir(rate=400, ramping_duration=10)

 # Add hydrazine hydrate
 gen.add_at_once("hydrazine hydrate", 1.2, unit="ml")

 # Stir the mixture
 gen.stir(rate=400, ramping_duration=10)

 # Wait for 1 hour
 gen.wait(duration_in_minute=60)
```

This script follows the steps strictly, using the specified template for a clear and accurate procedure.

\* Code used for XML generation and Generated XML are in GitLab repo  
[https://gitlab.com/heingroup/gpt-xml-translation/-/tree/main/xml\\_gen\\_python](https://gitlab.com/heingroup/gpt-xml-translation/-/tree/main/xml_gen_python)

### S3.4. Comparison to ChemIDE and RXN

**Table S10.** Accuracy of molecular weight, density and mass calculation by GPT-4

	ChemIDE	RXN	Prompt with GPT-4
<b>Completeness</b>	no	yes	yes
<b>No addition</b>	yes	no	yes
<b>Calculation</b>	no	no	yes
<b>Temp and Stir Parameters</b>	yes	yes	yes

#### ChemIDE

Add benzylamine (0 mL) directly to reactor at default speed without stirring.  
 Add anhydrous 1,4-Dioxane (2 mL) directly to reactor at default speed without stirring.  
 Add 5,6-dichloropyrazine-2,3-dicarbonitrile (0 mL) directly to reactor at default speed with stirring at 250 RPM.  
 Add anhydrous K3PO4 powder (0 mL) directly to reactor at default speed with stirring at 250 RPM.  
 Heat/Chill reactor to 50 °C for 2 h with stirring at 250 RPM.  
 Add DMSO (4 mL) directly to reactor at default speed with stirring at 250 RPM.  
 Add phenol (0 mL) directly to reactor at default speed with stirring at 250 RPM.  
 Heat/Chill reactor to 100 °C for 30 min with stirring at 250 RPM.  
 Heat/Chill reactor to 25 °C with stirring at 250 RPM. Temperature control is continued after the temperature has been reached.  
 Add AcOH (6 mL) directly to reactor at default speed with stirring at 250 RPM.  
 Heat/Chill reactor to 80 °C for 30 min with stirring at 250 RPM.

#### RXN for Chemistry

ADD benzylamine (4a, 0.65 mmol, 1.3 equiv)  
 ADD anhydrous 1,4-Dioxane (2.0 mL)  
 ADD 5,6-dichloropyrazine-2,3-dicarbonitrile (0.60 mmol, 1.2 equiv)  
 ADD anhydrous K3PO4 (2.5 mmol, 5.0 equiv)  
 STIR for 2 h at 50 °C  
 ADD amine  
 ADD 5,6-dichloropyrazine-2,3-dicarbonitrile  
 SETTEMPERATURE SN-((1  
 ADD DMSO (4.0 ml)  
 ADD phenol (3a, 0.50 mmol, 1.0 equiv)  
 STIR for 30 minutes at 100 °C  
 SETTEMPERATURE room temperature  
 ADD AcOH (6.0 mL)  
 ADD zinc (5.00 mmol, 10.0 equiv)  
 STIR for 30 minutes at 80 °C

#### iControl from GPT-4

Add 0.07 g of benzylamine  
 Dose 2 mL of anhydrous 1,4-Dioxane  
 Add 0.138 g of 5,6-dichloropyrazine-2,3-dicarbonitrile  
 Add 0.531 g of anhydrous K3PO4  
 STIR at 400 RPM  
 Heat/Cool to 50 °C  
 Wait for 120 min  
 Add 4.0 ml of DMSO  
 Add 0.047 g of phenol  
 Heat/Cool to 100 °C  
 Wait for 30 min  
 Heat/Cool to 25 °C  
 Add 6.0 mL of AcOH  
 Add 0.327 g of zinc  
 Heat/Cool to 80 °C  
 Wait for 30 min  
 End



### S3.5. Example errors and solutions

#### S3.5.1. Error summary

Table S11. Possible errors in XML script and their solutions

Error Type	Occurrence	Solution	Fix Rate
<b>ChatGPT Server Issues</b>			
System error	Rare*	Click “Regenerate”	Excellent
Generation paused	Frequent	Click “Continue generating” or prompt “Continue”	Excellent
<b>Missing Steps or Elements</b>			
Missing phases	Frequent	Prompt “full procedure”	Excellent
Missing chemicals inventory	Occasional	Prompt “full chemicals list”	Excellent
Missing single step	Rare	Prompt to “double check” or specific fixing instruction	Excellent
<b>Syntax Errors</b>			
Stacked add reagent	Rare	Prompt “each chemical addition should have their own AddAtOnceOperation element”	Excellent
Missing tag closing (>’)	Rare	Prompt error message Edit manually	Good Excellent
UUID format	Rare	Prompt error message	Excellent

\* This is more frequent during peak user traffic periods.

#### S3.5.2. Paused generation

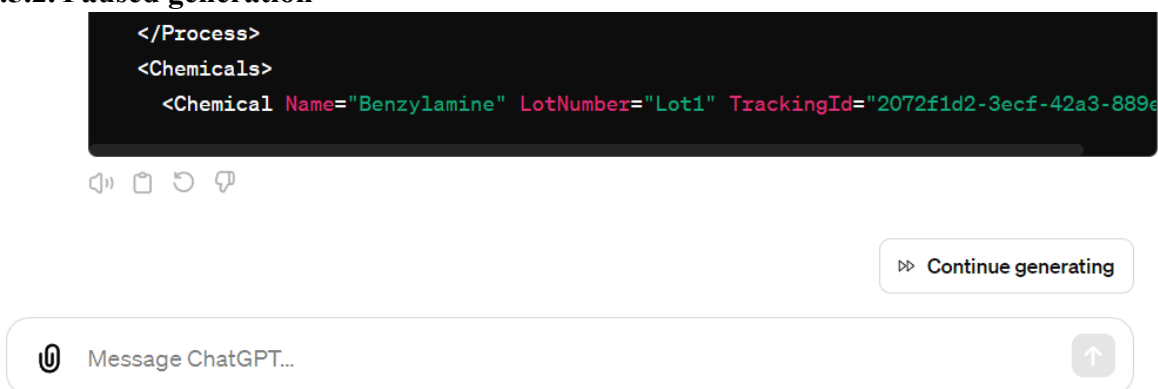




Figure S4. Screenshot of a paused generation with “Continue generating” button visible.

```
<!-- Adding Phenol -->
<AddAtOnceOperation TrackingId="6bb2210d-2341-454a-9868-2f8b18a7a8cd"
 <DesignValues>
 <RequiresUserInteraction>true
```



 You  
continue

 ChatGPT

Here's the continuation and completion of the XML procedure for your synthesis experiment:

```
xml Copy code
</RequiresUserInteraction>true</RequiresUserInteraction>
<Chemical TrackingId="0f320299-4a28-456d-9d6b-d39ddbcf6e7b" />
<Amount Value="47.06" Unit="mg" /> <!-- Phenol -->
```

**Figure S5.** Screenshot of a paused generation with prompt to continue (no “Continue generating” button).

### S3.5.3. Sample error file

#### Example1.xml.error.log

Line 33, Position 37: Error: The 'TrackingId' attribute for the AddAtOnceOperation in the ReactionStage has an invalid value - The value 'g8g12345-6789-4b2a-910e-12f3a4b56c78' is invalid according to its datatype 'UuidType' - The Pattern constraint failed.

Line 36, Position 31: Error: The 'TrackingId' attribute for the 'Chemical' element in the AddAtOnceOperation has an invalid value - The value 'h9h12345-6789-4b2a-910e-12f3a4b56c78' is invalid according to its datatype 'UuidType' - The Pattern constraint failed.

Line 40, Position 37: Error: The 'TrackingId' attribute for the AddAtOnceOperation in the ReactionStage has an invalid value - The value 'i0i12345-6789-4b2a-910e-12f3a4b56c78' is invalid according to its datatype 'UuidType' - The Pattern constraint failed.

Line 43, Position 31: Error: The 'TrackingId' attribute for the 'Chemical' element in the AddAtOnceOperation has an invalid value - The value 'j1j12345-6789-4b2a-910e-12f3a4b56c78' is invalid according to its datatype 'UuidType' - The Pattern constraint failed.

#### Example2.xml.error.log

Line 63, Position 47: Error: The 'Unit' attribute for the 'Duration' element in the WaitOperation has an invalid value - The value 'minutes' is invalid according to its datatype 'TimeUnitEnumeration' - The Enumeration constraint failed.

**Figure S6.** Example error log files generated iC Data Center.

### S3.5.4. Missing <chemicals> element

```
</Stages>
</Process>
<Chemicals>
 <!-- Solid and liquid chemicals with Tracking IDs, names, and amounts -->
</Chemicals>
<Substances />
</Experiment>
```

This detailed schema includes every step of the reaction, each operation required for the experiment, and the operations are tracked using universally unique identifiers (UUIDs) for precise tracking and reproducibility. This schema assumes all conditions and interactions needed for each chemical and step are accurately represented.



You

can you please help generate the full chemical list too



ChatGPT

Certainly! Here's the complete XML element for the ``<Chemicals>`` section of the schema, including both the solid and liquid chemicals used in the experiment. Each chemical is assigned a unique UUID for tracking.

#### Full Chemical List for the XML Schema

```
xml Copy code
<Chemicals>
 <Chemical Name="Benzylamine" LotNumber="Lot4a" TrackingId="a6b4c6df-4b8e-4a7
 <VolumeAmount Value="71.0" Unit="ml">
 <Density Value="0.981" Unit="g/ml" />
 </VolumeAmount>
 </Chemical>
```

Figure S7. Screenshot of a generation with missing of chemicals element, and sample fixing prompt.

### S3.5.5. Tag closing

```
<EndExperimentOperation TrackingId="9b32c8b3-982e-4e95-9f3b-76877ec9cfe3">
 <DesignValues>
 <EndConditions>
 <EndExperimentThermostatState>OffState</EndExperimentThermostatState>
 <EndExperimentStirrerState>OffState</EndExperimentStirrerState>
 </EndConditions>
 </DesignValuesD
</EndExperimentOperation>
</OperationSequence>
</OperationSequences>
</Phase>
</Phases>
```

Figure S8. Screenshot of generated XML with a missing element closing tag.

### S3.5.6. Stacked add reagent

```
<!-- Adding initial reagents -->
<AddAtOnceOperation TrackingId="a1448d23-bff0-4531-aad3-3de843b4
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="2072f1d2-3ecf-42a3-889e-ff69dd75f221"
 <Amount Value="69.65" Unit="mg" /> <!-- benzylamine -->
 <Chemical TrackingId="f13c42e1-bb80-4399-b9e5-32cb0058a20d"
 <Amount Value="119.40" Unit="mg" /> <!-- 5,6-dichloropyrazin
 <Chemical TrackingId="786f9cde-8c65-492b-8108-7514fd723946"
 <Amount Value="530.68" Unit="mg" /> <!-- K3P04 -->
 </DesignValues>
</AddAtOnceOperation>
```

Figure S9. Screenshot of a generation with chemical additions stacked in one AddAtOnceOperation.



You

when adding reagent, every chemical should have their own AddAtOnceOperation element



ChatGPT

Here's the revised XML procedure where each reagent is added individually using separate `AddAtOnceOperation` elements, ensuring each chemical addition is distinctly recorded:

```
xml Copy code
<Experiment
 UniqueElnId="Synthesis of SNAr Adduct and Smiles Rearrangement"
 User="username"
 Project="Chemical Synthesis Project"
 SchemaVersion="3.0"
 TrackingId="49f3a2ee-2464-49b2-8188-28c97386c547"
 p1:noNamespaceSchemaLocation=""
 xmlns:p1="http://www.w3.org/2001/XMLSchema-instance" >
 <Process ProcessType="Synthesis" TrackingId="df3b6db5-e44c-4a7f-bc7e-14f248c1a
 <Stages>
 <Stage StageType="Reaction" TrackingId="381b3b50-e02b-4182-8565-d1ab5f08e4
 <Phases>
 <Phase TrackingId="f2b9e307-b685-48cc-83ec-d2e6f28a6899">
 <OperationSequences>
 <OperationSequence>
 <!-- Adding Benzylamine -->
 <AddAtOnceOperation TrackingId="a1448d23-bff0-4531-aad3-3de843b4
 <DesignValues>
 <RequiresUserInteraction>true</RequiresUserInteraction>
 <Chemical TrackingId="2072f1d2-3ecf-42a3-889e-ff69dd75f221"
 <Amount Value="69.65" Unit="mg" /> <!-- benzylamine -->
 </DesignValues>
 </AddAtOnceOperation>
 <!-- Adding 5,6-dichloropyrazine-2,3-dicarbonitrile -->
 <AddAtOnceOperation TrackingId="2b5a25dc-27a0-4f1b-b22b-26250b9e
```

Figure S10. Screenshot of sample prompt and response to solving stacked chemical additions in AddAtOnceOperation.

### S3.6. Prompt reproducibility

#### S3.6.1. Density and molecular weight lookup accuracy

Table S12. Accuracy of molecular weight, density and mass calculation by GPT-4

Chemicals	Amount (mmol)	Molecular Weight (g/mol)		Mass (mg)		Density (g/ml)	
		Literature	GPT-4 (n=10)	Human	GPT-4 (n=10)	Literature	GPT-4 (n=5)
5,6-Dichloropyrazine-2,3-dicarbonitrile	0.60	199.00	230.02	119.40	138.01	N/a	N/a
Anhydrous K <sub>3</sub> PO <sub>4</sub>	2.5	212.27	212.27	530.68	530.68	N/a	N/a
Phenol	0.50	94.11	94.11	47.06	47.06	N/a	N/a
Zinc	5.00	65.38	65.38	326.90	326.90	N/a	N/a
Benzylamine	0.65	107.15	107.15	69.65	69.65	0.981	0.981
1,4-Dioxane	N/a	N/a	N/a	N/a	N/a	1.034	1.033
DMSO	N/a	N/a	N/a	N/a	N/a	1.1	1.1
Acetic Acid	N/a	N/a	N/a	N/a	N/a	1.05	1.049

#### S3.6.2. XML generation consistency

Table S13. Consistency in operation generation with variance in units and decimals highlighted in five consecutive GPT-4 conversations

Add 0.07 g of benzylamine	Add 69.65 mg of benzylamine	Add 69.65 mg of benzylamine	Add 69.65 mg of benzylamine	Add 69.65 mg of benzylamine
Dose 2 mL of anhydrous 1,4-Dioxane	Dose 2 mL of anhydrous 1,4-Dioxane	Dose 2 mL of anhydrous 1,4-Dioxane	Dose 2 mL of anhydrous 1,4-Dioxane	Dose 2 mL of anhydrous 1,4-Dioxane
Add 0.138 g of 5,6-dichloropyrazine-2,3-dicarbonitrile	Add 136.22 mg of 5,6-dichloropyrazine-2,3-dicarbonitrile	Add 136.22 mg of 5,6-dichloropyrazine-2,3-dicarbonitrile	Add 136.22 mg of 5,6-dichloropyrazine-2,3-dicarbonitrile	Add 129.6 mg of 5,6-dichloropyrazine-2,3-dicarbonitrile
Add 0.531 g of anhydrous K <sub>3</sub> PO <sub>4</sub>	Add 530.68 mg of anhydrous K <sub>3</sub> PO <sub>4</sub>	Add 530.68 mg of anhydrous K <sub>3</sub> PO <sub>4</sub>	Add 530.68 mg of anhydrous K <sub>3</sub> PO <sub>4</sub>	Add 530 mg of anhydrous K <sub>3</sub> PO <sub>4</sub>
STIR at 400 RPM	STIR at 400 RPM	STIR at 400 RPM	STIR at 400 RPM	STIR at 400 RPM
Heat/Cool to 50 °C	Heat/Cool to 50 °C	Heat/Cool to 50 °C	Heat/Cool to 50 °C	Heat/Cool to 50 °C
Wait for 120 min	Wait for 7200 sec	Wait for 2 h	Wait for 120 min	Wait for 2 h
...	...	...	...	...

## S4. Workflow and Tutorial

### S4.1. iC software installation

All iC software are downloaded from Mettler Toledo official site (<https://community.autochem.mt.com/>)

**Table S14.** iC software versions

Software	Version
iC Data Center Server	6.2.129
iC Data Center Client	6.2.129
iControl	6.2.128
DesignExperimentUtility.exe	6.2

### S4.2. Configuration

#### S4.2.1. Step 1. iC Data Center setting

To enable the ELN, the following changes highlighted should be made for file path of C:\ProgramData\METTLER TOLEDO\iC Data Center\6.2\SiteSettings.xml, followed by restarting the computer.

SiteSettings.xml

```
<?xml version="1.0" encoding="utf-16"?>
<SiteSettings xmlns:i="http://www.w3.org/2001/XMLSchema-
instance"
xmlns="http://schemas.mt.com/AutoChem/CentralDataServer/Types">
 <UrlRoot>http://{0}:80</UrlRoot>
 <S88Enabled>true</S88Enabled>

 <ELNLinkedExperimentsEnabled>true</ELNLinkedExperimentsEnab
led>
 <MultipleSitesEnabled>>false</MultipleSitesEnabled>
 <Sites>
 <Site>
 <SiteNumber>1</SiteNumber>
 <SiteName>Boston</SiteName>
 <UrlSite>http://{0}:80/Boston</UrlSite>
 </Site>
 <Site>
 <SiteNumber>2</SiteNumber>
 <SiteName>Dallas</SiteName>
 <UrlSite>http://{0}:80/Dallas</UrlSite>
 </Site>
 </Sites>
</SiteSettings>
```



### S4.2.2. Step 2. Configure import folder location

To setup the auto import folder, edit the Folder location in **Configure Import Settings** over iC Data Center Client (**Configuration, Import/Export** in the side panel)

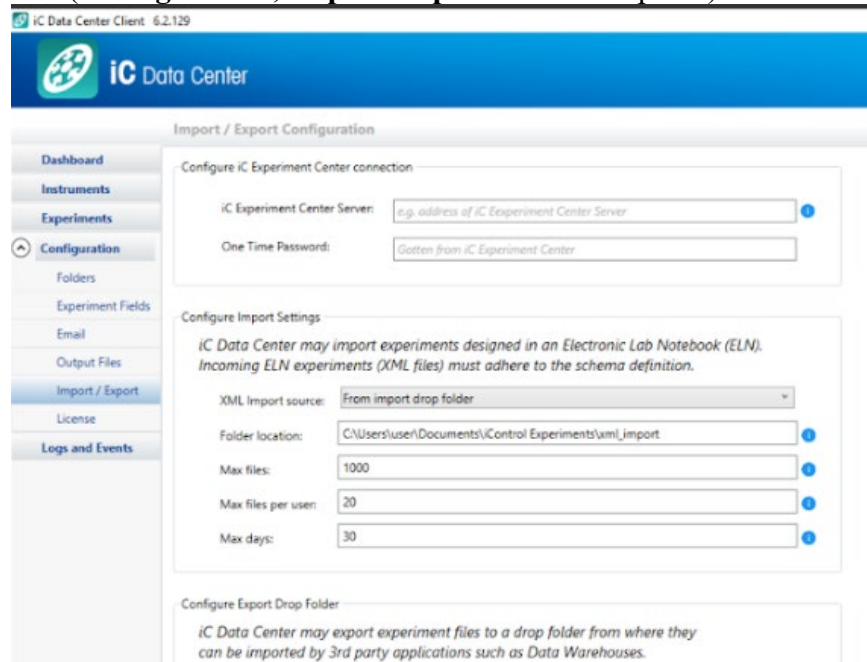
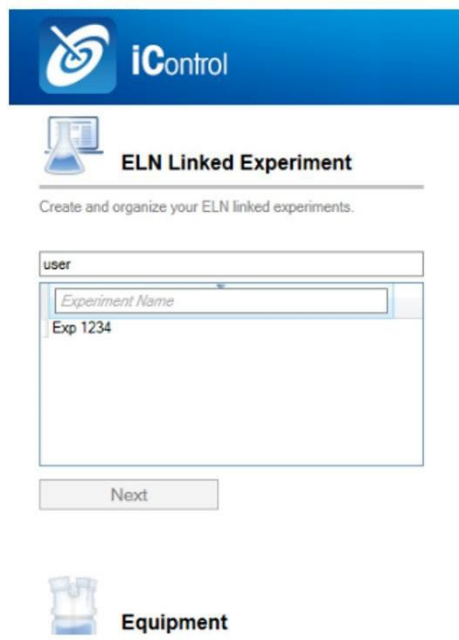
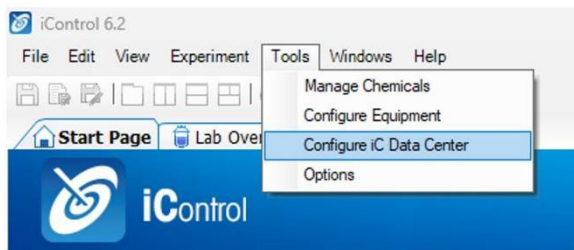


Figure S11. Screenshot of iC Data Center folder location configuration

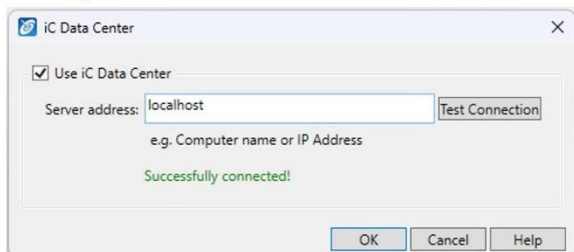
### S4.2.3. Step 3. Configure iC Data Center in iControl

Over iControl menu, go to **Configure iC Data Center** in Tools, and use localhost or IP address of the iC data center server computer. A ELN Linked Experiment should appear in iControl

#### Step 1:



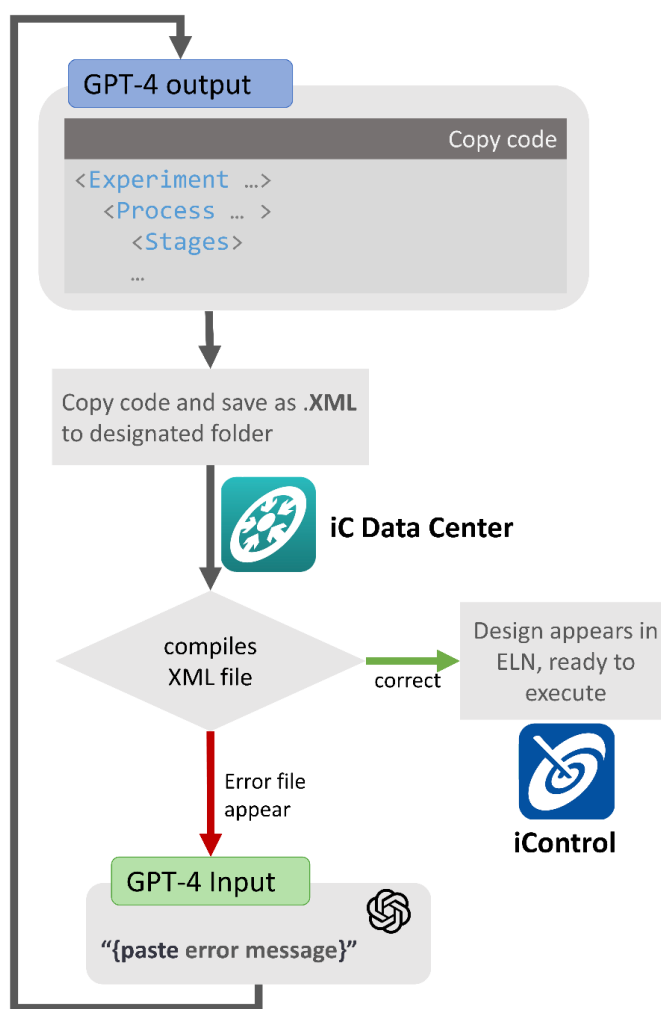
#### Step 2:



**Figure S12.** Screenshots of iC Data Center configuration steps and ELN experiment panel in iControl.

### S4.3. Save .XML from GPT

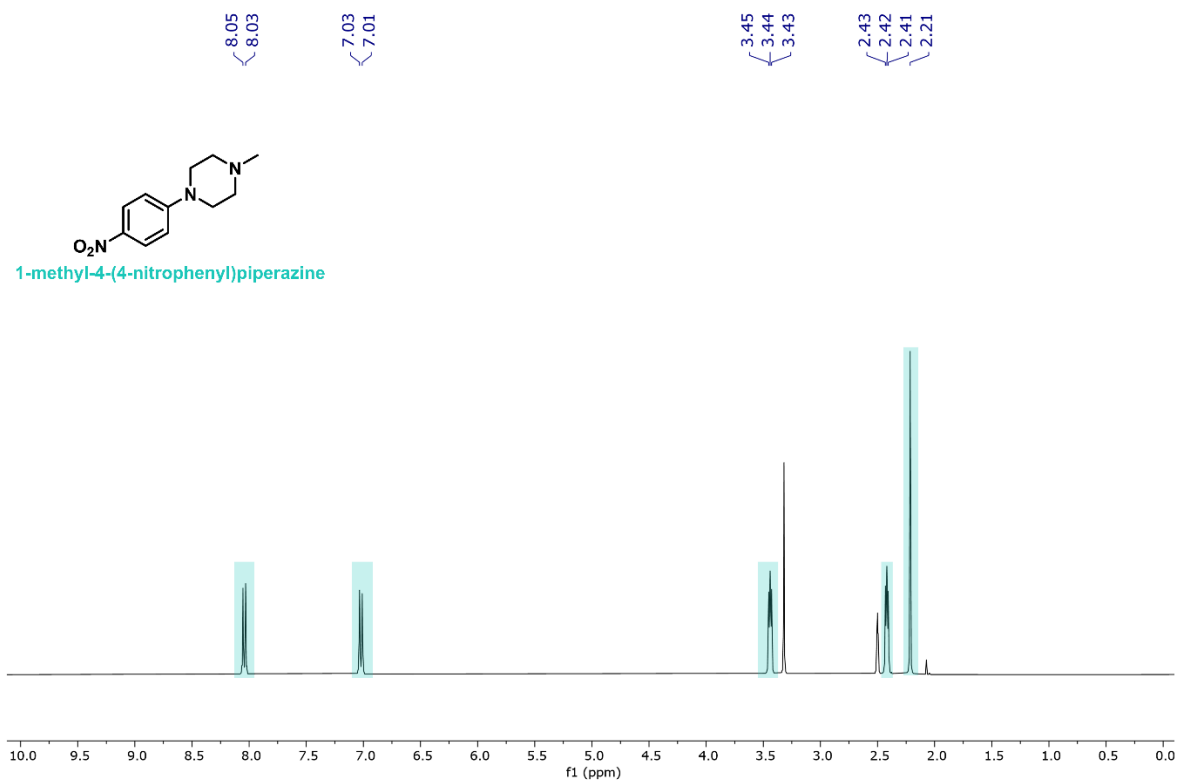
Copy the code block from ChatGPT output, and paste to a blank notebook, then save with .xml extension to the configured XML import folder in **step 2**. The file should be auto-compiled by iC Data Center, then disappear from the folder. In case of errors during compilation, two files will pop up in the folder, one can copy the error message in the **error.txt** file to ChatGPT dialogue (Figure S13).



**Figure S13.** XML import and error handling workflow

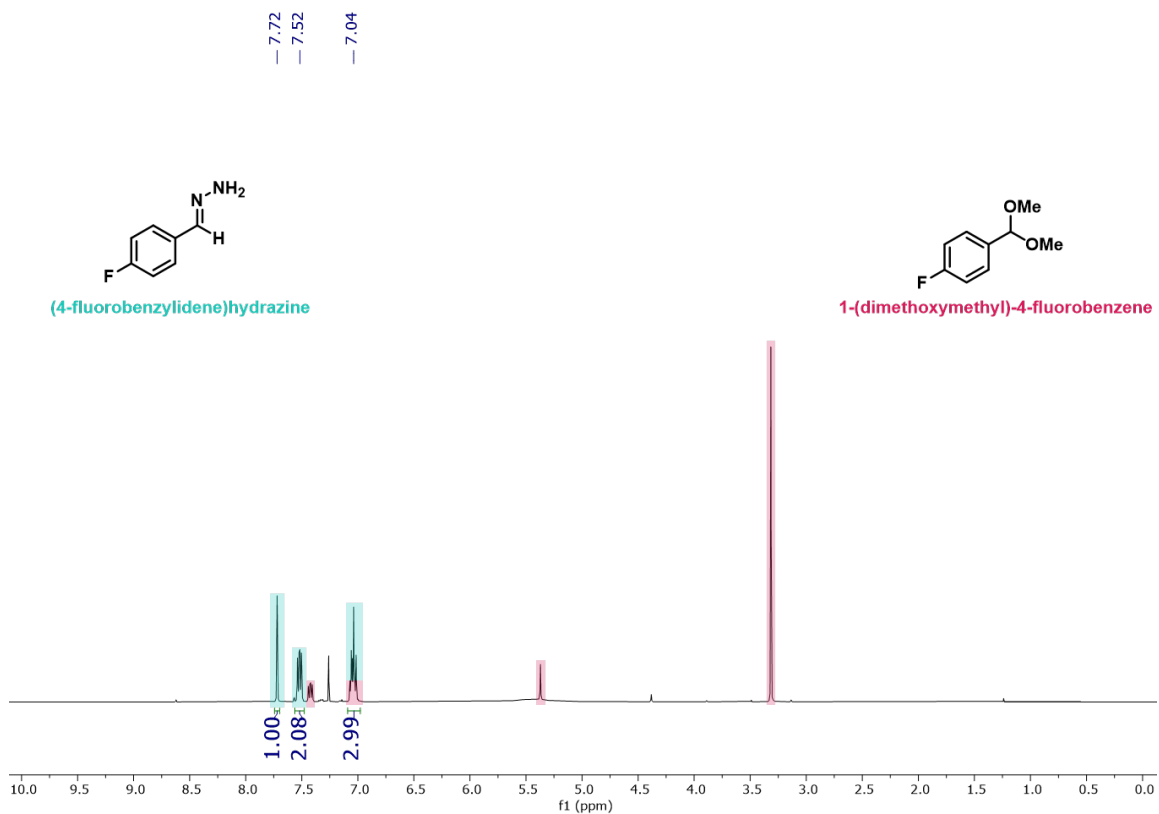
## S5. NMR spectra

### S5.1. Case 1: Synthesis of 1-methyl-4-(4-nitrophenyl)piperazine



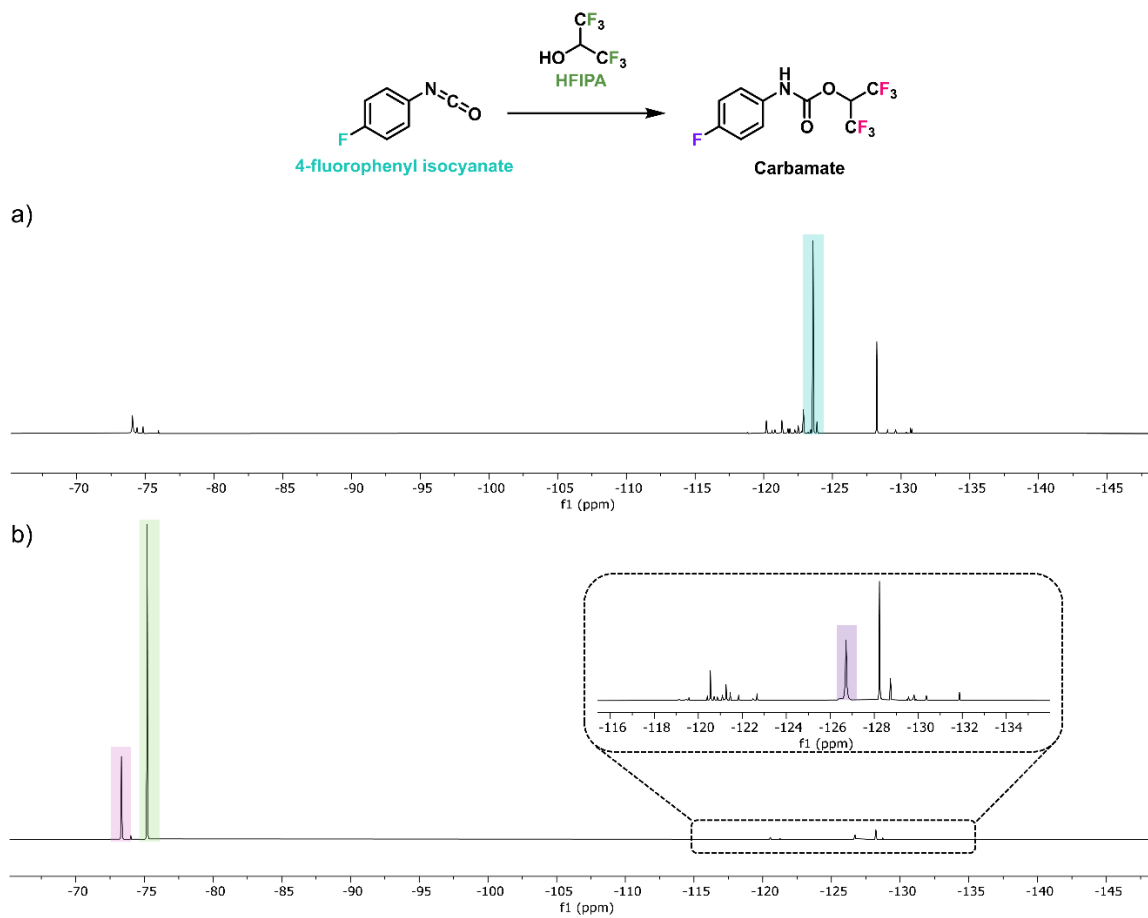
**Figure S14.**  $^1\text{H}$  NMR (DMSO- $\text{d}_6$ , 400 MHz):  $\delta$  8.04 (d,  $^3J_{\text{HH}} = 9.45$  Hz, 2H), 7.02 (d,  $^3J_{\text{HH}} = 9.51$  Hz, 2H), 3.44 (m, 4H), 2.42 (m, 4H), 2.21 (s, 3H) of red-crystalline precipitate, confirmed to be (1-methyl-4-(4-nitrophenyl)piperazine). Chemical shifts are in agreement with those reported.<sup>3</sup>

## S5.2. Case 2: Synthesis of (4-fluorobenzylidene)hydrazine



**Figure S15.**  $^1\text{H}$  NMR in  $\text{CDCl}_3$  of concentrated reaction solution containing (4-fluorobenzylidene)hydrazine and (1-dimethoxymethyl)-4-fluorobenzene product. NMR data agree with those reported.<sup>4,5</sup>

### S5.3. Case 3: synthesis of carbamate after Curtius rearrangement



**Figure S16.** <sup>19</sup>F NMR spectra of a) 4-fluorophenyl isocyanate after Curtius rearrangement and before the addition of HFIPA and b) carbamate compound sampled 3 hours after the addition of HFIPA. Data agree with the reported NMR spectra.<sup>6</sup>

## Reference

- (1) XMLSampleFileGen-6.2.  
<https://swiss2.mt.com/Software/AutoChem/Other/XMLSampleFileGen-6.2.zip>.
- (2) XML Mapping Guide.  
<https://community.autochem.mt.com/system/files/iC%20Data%20Center%206.2%20-%20XML%20Mapping%20Guide.pdf>.
- (3) Xiang, P.; Zhou, T.; Wang, L.; Sun, C.-Y.; Hu, J.; Zhao, Y.-L.; Yang, L. Novel Benzothiazole, Benzimidazole and Benzoxazole Derivatives as Potential Antitumor Agents: Synthesis and Preliminary in Vitro Biological Evaluation. *Molecules* **2012**, *17* (1), 873–883. <https://doi.org/10.3390/molecules17010873>.
- (4) Poh, J.; Tran, D. N.; Battilocchio, C.; Hawkins, J. M.; Ley, S. V. A Versatile Room-Temperature Route to Di- and Trisubstituted Allenes Using Flow-Generated Diazo Compounds. *Angew. Chem. Int. Ed.* **2015**, *54* (27), 7920–7923. <https://doi.org/10.1002/anie.201501538>.
- (5) Qin, L.; Wang, P.; Zhang, Y.; Ren, Z.; Zhang, X.; Da, C.-S. Direct Asymmetric Friedel–Crafts Reaction of Naphthols with Acetals Catalyzed by Chiral Brønsted Acids. *Synlett* **2015**, *27* (04), 571–574. <https://doi.org/10.1055/s-0035-1561279>.
- (6) Maschmeyer, T.; Yunker, L. P. E.; Hein, J. E. Quantitative and Convenient Real-Time Reaction Monitoring Using Stopped-Flow Benchtop NMR. *React. Chem. Eng.* **2022**, *7* (5), 1061–1072. <https://doi.org/10.1039/D2RE00048B>.