

King-Smith Group Working Guide

In this document you will find the useful links, guides, and procedures for the King-Smith Group. As we are an interdisciplinary group, you will be expected to understand synthetic chemistry and machine learning, thus the length of this document. This packet is meant to serve as both a starting up guide and something you can refer back to as you progress in your PhD, PGT, BSc, and 5P. As the group grows, this guide will likely change and update. Those changes will be communicated to the group and a live version of this document will be made continually available for reference.

Remember that learning is a continual process – there is no shame in asking for more clarification or feeling confused. We are all here to support you!

Table of Contents

Group Conduct.....	pg. 3
General Resources.....	pg. 7
Chemistry Resources.....	pg. 9
Computational Resources.....	pg. 11
Machine Learning Resources.....	pg. 13
Searching Reaxys & SciFinder.....	pg. 14
Curating Data from Reaxys.....	pg. 20

Group Conduct

The King-Smith Group is a welcoming space for all of its members, irrespective of their length of tenure in the group. We all come from different parts of the world, cultures, and backgrounds. While this can occasionally lead to miscommunications, our diversity is something to be celebrated. The King-Smith Group is one that is respectful and professional; no discrimination or harassment will be tolerated. If you feel that someone is displaying bullying behavior towards yourself or another member, and you feel you are unable to resolve it, you must inform me as soon as possible. This statement holds even if that person is not a member of our group.

What can You Expect from Me:

My working hours are 8:00 - 18:30. I aim to be in the office (Rm 241) for the bulk of that time, so please feel free to knock on my door. If I am not in the office for whatever reason, please do not hesitate to reach out *via* email (emma.king-smith@ed.ac.uk) or by WhatsApp (XXXXXXXXXX). Unfortunately, Rm 241 is a communal office, so if you wish to have a private and/or confidential conversation with me, please let me know so that we can book a suitable location.

I will respond to communication within 1 business day or if urgent, within 24 hours. If you haven't heard back from me after that time, please re-send your email or text. Remember that you are not wasting my time – it is my job to support you and your research.

What I Expect from You:

I expect the members of my group to be diligent in their research, problem solving where they can and reaching out for assistance when they need it. Every member and I will have biweekly meetings that will be ~45 min in length, on average. In our biweekly meetings, I do not expect a cohesive narrative, but I do want to see that you are actively engaging with your research and have a plan for the upcoming weeks. I expect you to show me bad results and for the research to be a team effort.

Evening and weekend work is not an expectation, so long as there are no immovable deadlines and you are managing your time well. That being said, research often benefits from a "little extra". For example, checking that your code is still running or working up a critical reaction over the weekend. The metric that you should use for a successful time in this group is "one big thing completed a day". That may be one reaction (setup, run, worked up, purified, characterized) or it may be a presentation (figures drawn and in the slides in the desired order) or it may be a piece of code (planned, written, debugged).

Interdisciplinary Research Good Practices:

We all come to the group with different strengths, and hopefully your project will help you hone those skills. However, no matter what, you will be learning something new: chemistry or machine learning. In addition to keeping up with both sets of literature, I recommend that chemistry-aficionados do daily coding either with Sololearn's Python Developer Course (<https://www.sololearn.com/en/>) and/or LeetCode's problem sets (<https://leetcode.com/problemset/>). Those comfortable with coding and machine

learning but not so much with chemistry should be doing mechanism practice daily. I find *Strategic Applications of Named Reactions in Organic Synthesis* by Kürti and Czakó (https://www.dirzon.com/Doc/Details/telegram:strategic_applications_of_organic.pdf) and Chemistry by Design (<https://chemistrybydesign.oia.arizona.edu/app/>) helpful. There are also flash cards developed by Prof. Chris Newton at Georgia Tech that offer an alternative way to practice (<https://www.newton-lab.org/flashcards>).

Large Language Model / Generative AI Statement: Group members are **NOT** allowed to use generative AI to create or synthesize any aspect of their research. This includes writing code, reports, or generating figures. You are allowed to use LLMs for debugging, editing, finding references, and when appropriate, "inspiration" (e.g., mockups for slide design). *You are expected to understand and explain every sentence and every line of code you write.*

Group Meetings:

We have two sets of group meetings: the first is presentation-based and the second is mechanism-based. Presentation-based group meetings occur once every 2 weeks on Mondays 16:00 - 17:30. A maximum of two people will present per session. A description of each presentation type is given below.

Presentation Type	Length	Notes
Data-Driven Literature	30 min presentation + 15 Q&A	A comprehensive walk through of a data-driven chemistry paper or related set of papers. They do not need to be ASAPs. All fundamental ML concepts should be explained in the presentation.
Synthesis Literature	30 min presentation + 15 Q&A	A comprehensive walk through of a chemical synthesis paper or related set of papers. They do not need to be ASAPs nor do they need to be total syntheses. Any chemistry that seems unfamiliar or worth discussion should be highlighted in the presentation.
Ideation to Code	45 min presentation	Your experimentation of taking an idea and implementing it as code. Feel free to showcase multiple iterations of your code. Typically, I will give you the first task. The goal is to have 1) you become better at coding and ML; 2) the group become used to reading other people's code on the fly; 3) improving the code by troubleshooting any (un)foreseen bugs.
Named Reaction / Reagent	30 min presentation + 15 Q&A	A detailed examination of one named reaction or named reagent. This can include a family of related reactions. A non-exhaustive list of things to include are: the developers of the reaction, prior work in the field, canonical examples in synthesis, selectivity of the reaction, limitations, and future directions of the reaction / reagent.
Research Update	30 min presentation + 15 Q&A	A research update talk with narrative. This should include an introduction and motivation as to why you are pursuing your project and should not be solely a collection of results. Expect to give a research update once every 6 months. This slot can be used as a practice presentation for upcoming conferences, oral exams, etc.

Topical Lecture	60 min presentation + 30 min Q&A	A comprehensive overview of a field that can be used as a reference for both members of the group and those within our field. Consider this to be a mini-review.
-----------------	----------------------------------	--

You may find it helpful to refer to [Baran](#), [Renata](#), and [Myers Lab](#) presentations to get a sense of the depth I expect for topical lectures. Named reactions are shorter in length and therefore cover less ground, but the goal is to have everyone (including me!) walk away with a deeper understanding of the chemistry.

Mechanism-based group meetings are held every 2 weeks (on presentation-GM off weeks) on Thursdays from 16:30 - 17:30. They are an opportunity to practice fundamental organic chemistry mechanisms at the board - great for interview prep and managing "whiteboard fever".

At the beginning of meetings, there will be an opportunity to discuss any issues that have arisen over the past week. These may include problems with accessing the high performance computing facility, near misses in the lab, or equipment failures. We should all be on time for group meetings and your presentations should be rehearsed at least once prior to your presentation.

How to Exit JBB After Hours:

As our group meetings run until the end of business day, you may find yourself in the situation where you cannot get through the turnstiles and our lovely servitor is not present. In that case, please make your way through the east exit, past G.069 and to the door that looks like a fire alarmed door, but actually isn't.

Lab Safety:

Like all UK labs, we implement Control of Substances Hazardous to Health (CoSHH) forms prior to any new lab work. All CoSHH forms must be approved by me before you begin your experiment and the form must be displayed on your fumehood. If you are at all ever unsure about a procedure, particularly if it involves pyrophoric agents, **ask someone**. All chemists can tell you about a horrible lab accident that they have personally witnessed – do not become another anecdote! Also remember that increasing the scale of your reaction also comes with potentially unforeseen risks. You must submit another CoSHH form if the scale of your reaction exceeds the previously approved size. Our CoSHH portal can be found here:

<https://uoe.sharepoint.com/sites/chemcollab-coshh/SitePages/King-Smith%20Group.aspx?csf=1&web=1&e=Ff2yxc&CID=2abdbdfa-fb0d-48a0-aed8-c7b64668b643>

Note that you must be on UoE internet to access sharepoint sites.

Standard working hours for lab work are 8:30 - 17:30. If you have not been approved for lone work or out-of-hours work you must stop all wet lab work at 17:30, although you may continue to do dry work and admin after hours. Additionally, members who are not PDRAs or PhDs will not be given permission for out-of-hours wet lab work.

Sick Leave:

If you are ill, please stay home. We have members of staff who are immunocompromised and will be greatly affected if infected. In the same vein, do not work when you are sick – you will only delay your recovery. Edinburgh does have an extended sick leave policy, which if, after consultation with me, we feel is the best course of action, we will implement.

This goes as well for mental health, please take the time you need to recover and come back to research fresh and motivated.

Annual Leave, Holiday Break, & Lab Cleanup:

Every member of staff is given 32 of paid time off in addition to the holiday break (8 days). Note that these 32 days include bank holidays. Prior to you going on leave, please get my approval for the dates of your absence. I very rarely refuse the time off. The university is closed over Christmas and New Years, so no work can be conducted during that time.

Once a year, we will have a lab cleanup and inventory prior to the holiday lab shut down. I expect all members to contribute to this effort, even if your work is computational. Details will be made known at a date closer to the end of the year. We will also have a festive celebration afterwards, but attendance for this is optional.

General Resources

Internet:

The most reliable internet is *via* eduroam. You can access it by following these steps:

<https://information-services.ed.ac.uk/computing/desktop-personal/wifi-networking/configure-device>

Edinburgh Shuttle & Lothian Buses:

A convenient way to move from the central campus to Kings Buildings during term time is the Edinburgh Shuttle. It is free to everyone who has a UoE badge and the schedule can be found here:

<https://transport.ed.ac.uk/public-transport/shuttle-bus>

Alternatively, the #9, #24, and #38 Lothian buses stop close to Kings Buildings. Live information can be found on the Lothian Bus & Tam App:

<https://www.lothianbuses.com/app/>

Chemistry Stores:

This is the storefront for purchasing general chemicals, consumables, PPE, and stationary. Note that stationery and cleaning items are free of charge.

<https://uoesharepoint.com/sites/chem-stores>

Purchasing:

Items not available through Chemistry Stores can be purchased through People & Money (P&M) under Procurement > Purchase Requisitions.

<https://www.ed.ac.uk/staff/services-support/hr-and-finance/people-and-money-system>

Anything that cannot be found through the default search bar must go through additional approval.

WARNING: The default codes in P&M purchasing are NOT our codes! To avoid spending on unapproved funds please use the codes provided.

Group Website:

Our website is <https://www.kingsmithgroup.com/>. If you are comfortable, please send me a headshot and a short ~100-word blurb about yourself. And of course keep in touch after you depart!

Dilbert & PhD Comics:

What's life without some simple pleasures? Great comics to indulge. 'Nuff said.

<https://dilbert-viewer.herokuapp.com>

<https://phdcomics.com/>

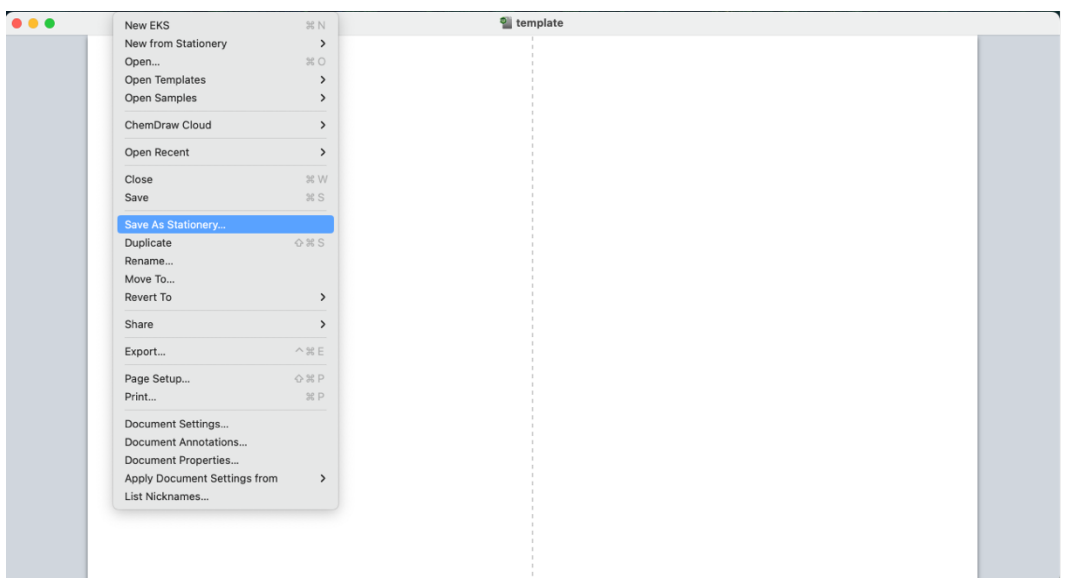
Chemistry Resources

ChemDraw:

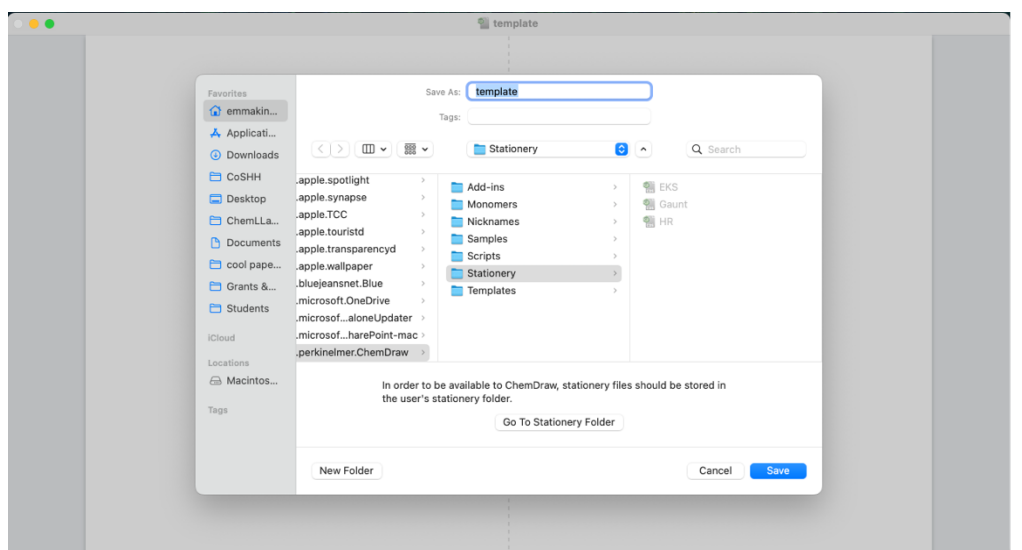
ChemDraw can be downloaded from the University of Edinburgh link:

<https://uoe.sharepoint.com/sites/chem-it/SitePages/School-Provided-Software.aspx#chemdraw>

The ChemDraw template that I use (and happen to find very aesthetically pleasing) was sent along with your introductory email, but I will make it available on our website as well. To set it as your stationary (e.g. default) open the template in ChemDraw, go to File > Save as Stationary.



You may name it whatever you like. Opening new ChemDraws will default to this style.



MNova:

NMR and Mass Spec analysis tool. This can be downloaded from the University of Edinburgh link:

<https://uoesharepoint.com/sites/chem-it/SitePages/School-Provided-Software.aspx#mnova>

NMR Training:

NMR training must be done prior to running your first experiment. Details can be found here:

<http://nmr-server.chem.ed.ac.uk/lowfield/automation-training.html>

Master Organic Chemistry:

A website that concisely explains fundamental organic chemistry concepts. For those who are dipping their toes into chemistry for the first time, this is a great first stop.

<https://www.masterorganicchemistry.com/>

Chemistry By Design:

A great resource to practice your synthesis – 4,000 syntheses and counting!

<https://chemistrybydesign.oia.arizona.edu/app/>

Fukuyama/Yokoshima Mechanisms:

Difficult to very difficult mechanism practice covering a variety of chemistry. Rest assured they come with answer sheets.

https://www.ps.nagoya-u.ac.jp/lab_pages/natural_products/problem-e.html

Computational Resources

Git & GitHub Access:

We will be using GitHub to store our code which requires git. A guide to git installation and basic git commands can be found here:

https://www.w3schools.com/git/git_workflow.asp?remote=github

Prior to publication of papers, all code should be on a private repository. This means that all members will need an SSH key to access other members' code. All code should be committed at the end of the day to prevent catastrophic losses. Approximately once a week (depending on how frequently coding is being debugged / created) you should push to your remote (GitHub repository).

A free GitHub account can be made by following this guide:

<https://docs.github.com/en/get-started/start-your-journey/creating-an-account-on-github>

SSH Key access can be created by following these steps:

<https://docs.github.com/en/authentication/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent>

<https://docs.github.com/en/authentication/connecting-to-github-with-ssh/adding-a-new-ssh-key-to-your-github-account>

Make sure you remember your password!

EDDIE Access:

EDDIE is Edinburgh's free-to-access computing service for all staff. It is a decent place to work through your code and test small bits of it, but it isn't suitable for the deep learning we will do later on. Documentation for setting up your EDDIE login can be found below and it was also be linked on our group site:

<https://uoee.sharepoint.com/sites/DRSFacilitation/Shared%20Documents/Forms/AllItems.aspx?id=%2Fsites%2FDRSFacilitation%2FShared%20Documents%2FGettingStartedOnEddie%2Epdf&parent=%2Fsites%2FDRSFacilitation%2FShared%20Documents&p=true&ga=1>

This is an EDDIE wiki that will come in handy later:

<https://www.wiki.ed.ac.uk/pages/viewpage.action?spaceKey=ResearchServices&title=Eddie>

EIDF GPU Access:

The deep learning will be done through the Edinburgh International Data Facility's (EIDF) GPU service. You will need an SSH key for remote desktop access but you can re-use the key you made for your GitHub account. You will need to create a EPCC SAFE account and request a Machine Account. Instructions can be found here:

<https://epcced.github.io/safe-docs/safe-for-users/>

The service uses kubernetes to manage the GPUs and computational jobs. They offer a nice guide on getting started with kubernetes:

https://docs.eidf.ac.uk/services/gpuservice/training/L1_getting_started/

Note that kubernetes uses Docker images. Docker can be downloaded by following the instructions here:

<https://docs.docker.com/desktop/>

Miniconda:

Running code can be a messy business – new packages have instabilities with others, updates to certain functions may break whole systems, versions of Python have major impacts on the types of packages you can install. All of this means we want to run things in virtual environments: clean spaces that we can setup just the way we need and tear it down if we need to. Miniconda is what we will be using and can be downloaded and installed here:

<https://www.anaconda.com/docs/getting-started/miniconda/install>

Pycharm & Other Interactive Development Environments (IDEs):

IDEs are a great way to get going quickly on the coding portion of your project. They offer nice features like debugging, running your code, and autofilling function names, parentheses, indents, etc. Personally, I use PyCharm, but you may use something else (just know that my assistance will be limited to my knowledge of your IDE).

<https://www.jetbrains.com/pycharm/>

For any IDE, you will need to specify the python interpreter / virtual environment. I find it's easiest to manage these environments by first creating a conda environment in the command line, then loading that environment into the IDE.

IDEs are, however, optional. You can do the same thing with just a text editor and a Linux terminal.

Jupyter notebooks are a quick and easy way to do things, but they will not have the memory needed to run most deep learning models, so I don't advise relying on them.

Machine Learning Resources

Python:

The language we will be using for the majority of our applications is Python. In the beginning, coding is a harsh mistress and will give you many many many unintelligible error messages. This is normal! Error messages can be Googled (look for StackExchange answers) or put into the LLM of your choice (GPT, Claude, DeepSeek, etc.).

For an introduction to coding, the university offers a crash course (Data-Driven Chemistry) which can be accessed here:

https://git.ecdf.ed.ac.uk/data-driven-chemistry/data-driven-chemistry-student/-/blob/master/Session1_introduction/Session_1.1.ipynb?ref_type=heads

GeeksforGeeks is a good resource as well that offers both introductory and step-by-step guides for specific python functions.

<https://www.geeksforgeeks.org/python/python-programming-language-tutorial/>

Remember that the best way to learn a new language (including a computer language) is exposure! Courses can help improve your fundamentals but ultimately your learning will begin when you start trying to code on your project.

A Math-Free Introduction:

I put together a course for a math-free introduction to machine learning. Admittedly, I am a little biased, but I think it serves as a gentle guide to the fundamental concepts of classic machine learning.

https://github.com/emmaking-smith/ML_for_Chem/tree/master

Machine Learning with Math:

Given how dense math symbolism can be, I find that it's easier to learn the mathematical aspects of machine / deep learning through lectures rather than text.

Two creators I think do an excellent job of this are StatQuest and 3Blue1Brown.

https://www.youtube.com/watch?v=Gv9_4yMHFhI&list=PLblh5JKOoLUICTaGLRoHQDuF_7q2GfuJF

<https://www.youtube.com/watch?v=aircAruvnKk&list=PLcCe-ymWq77ow42k4-ZrLzIM3F7Ha7smT>

https://www.youtube.com/watch?v=fNk_zzaMoSs

I highly recommend watching through the latter's deep learning and linear algebra playlists (linked above). They are usually 10 - 20 min long and very well put together.

Searching Reaxys & SciFinder

One of the best resources for a chemist is Reaxys and SciFinder. Remember that these are two different databases that are run by two different companies, so they will not have identical information. It's always worthwhile checking both.

Starting with SciFinder, click the "Draw" button.

The screenshot shows the SciFinder homepage. At the top, there's a navigation bar with the CAS logo and 'SciFinder' text. Below it, a user greeting 'Good Morning, Emma' is displayed. A search bar is present with the text 'Search by CAS Reaction Number, Substance Name, CAS RN, Patent Number, PubMed ID, AN, CAN, and/or DOI.' To the right of the search bar, a 'Draw' button is highlighted with a red rectangle. Below the search bar, there are four featured search options: 'Prior Art Discovery', 'Patent Markush', 'Retrosynthetic Analysis', and 'Search CAS Lexicon'. Each option has a brief description and an icon.

A popup window will appear where you can enter your reaction of interest.

The screenshot shows the 'CAS Draw' popup window. It has a title bar 'CAS Draw' and a close button. Below the title bar, there's a toolbar with various drawing tools. A text input field says 'Enter a CAS Registry Number, SMILES, or InChI...'. Below this, a blue box says 'Draw or change atoms or bonds.' The main area shows a chemical reaction: a benzene ring with a bromine atom (labeled 'reactant') reacts with a cyclopentane ring with a bromine atom (labeled 'reactant') to form a biphenyl derivative (labeled 'product'). Below the reaction, the molecular formula is given: 'Molecular Formula: C₆H₅Br (157.01) . C₅H₉Br (149.03) . C₁₁H₁₄ (146.23)'. At the bottom, there's a 'Zoom: 100%' slider and 'Cancel' and 'OK' buttons.

Once it is drawn, click "OK". You will be returned to the home page that now has your reaction of interest underneath the edit button. Press "Search".

Good Morning, Emma

Search by CAS Reaction Number, Substance Name, CAS RN, Patent Number, PubMed ID, AN, CAN, and/or DOI.

Featured Search

- Prior Art Discovery**
Discover prior art in patents and non-patent literature using AI-enhanced search technologies.
- Patent Markush**
Search Patent Markush by structure and view associated references.
- Retrosynthetic Analysis**
Make reaction plans with conditions, yields, catalysts, and experimental procedures.

Edit Drawing taxonomy.

The results of the search will be displayed. You can filter for exact matches, substructure matches, or similar reactions. Scrolling through the filters, you can show results that are from specific journals, use specific reagents, or have specific yields.

Reactions search for drawn structure

View Related Results

Filter Results

Structure Match

- As Drawn (3)
- Substructure (1,913)
- Similarity (18K)

Behavior

Filter by Exclude

3 Results

Group: By Scheme Sort: Relevance View: Expanded

Scheme 1 (3 Reactions)

Steps: 1 Yield: 63-72%

31-056-CAS-15724583 Steps: 1 Yield: 72%

Nickel-Catalyzed Reductive Cross-Coupling of Aryl Bromides with Alkyl Bromides: Et₃N as the Terminal Reductant

By: Duan, Zhengli; et al

Additionally, you can quickly go to the primary source by clicking "Full Text" and selecting either "findit@edinburgh" or "DOI".

Reaxys employs a very similar search engine. From the homepage, press the "Draw" button.

A popup screen will appear where you can draw your reaction of interest. On the right-hand side, you can search for "as drawn" matches (exact matches), "as substructure" matches, or "similar" matches. "As substructure" has the option of matching all atoms or only heteroatoms. You can also select whether or not to include tautomers, stereochemistry, etc. in your matching.

Structure editor selected: ☒ MarvinJS ☐ ChemDrawJS

Insert structure from name

Search this structure as:

- ☐ As drawn
- ☒ As substructure
- ☒ On all atoms
- ☐ On heteroatoms
- ☐ Similar
- ☒ Tautomers
- ☒ Stereo
- ☒ Stereo
- ☒ Additional ring closures
- ☐ Related Markush
- ☒ Salts
- ☒ Mixtures
- ☒ Isotopes
- ☒ Charges
- ☒ Radicals
- [+ More Options](#)

Clear

When done, click "Transfer to query" which will bring you back to the homepage.

Discover a more intuitive way to search — with Reaxys AI Search Beta.

Search for

Import [...](#)

Search Reaxys

Substance Molecular Formula, e.g. Pt(PPh₃)₃

AND

On all atoms [...](#)

Content Overview | Latest update: 18. August 2025

350M Substances 71M Reactions 122M Documents 47M Patents 49M Bioactivities

Then click "Find".

Results for

191 Reactions

Reaction Query : substructure; included: tautomers, only absolute stereo, additional ring closures allowed, salts, mixtures, isotopes, charges, radicals

Edit in Query Builder Create Alert

[View Results](#)

Remote access Terms and Conditions Privacy policy About content Accessibility Contact support

We use cookies to help provide and enhance our service and tailor content. [Cookie Settings](#)

All content on this site: Copyright © 2025 Elsevier B.V., its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the relevant licensing terms apply.

Feedback

We see that 191 reactions match our reaction pattern either exactly or as a substructure. To view results, click "View Results". Then we can perform further filtering by yield, journal, reagent, etc.

191 Reactions out of 76 Documents, containing 327 Substances, 63 Targets

Sort by Reaxys Ranking

0 selected Limit To Exclude Export Hide Conditions

1

2 Conditions Find Similar Reaction ID: 28432258

Conditions

With hydrogen; palladium 10% on activated carbon in methanol at 25°C; for 24h; Experimental Procedure

With palladium 10% on activated carbon; hydrogen in methanol at 25°C; for 24h; Inert atmosphere;

Yield	Reference
76%	Current Patent Assignee: COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK - WO2009/38731, 2009, A2 Location in patent: Page/Page column 357; 4/37; 24/37 Full Text Details Abstract
76%	Snyder, Scott A.; Breszozzo, Steven P.; Ross, Audrey G.; Lin, Yunqing; Zografos, Alexandros L. [Journal of the American Chemical Society, 2009, vol. 131, p. 1753 - 1765] Full Text Details Abstract

Feedback

You can also easily navigate to the primary source from both. Click on "Full Text" which will redirect you to the article or patent in question.

www.reaxys.com/#/results/reactions/8/RX001__370961287080 80% Search

Getting Started Dinner List SmartR Dashboard SmartR Email AR Caramelized Shallot... Renegade Platinum Renegade Platinum ... Pantry Pasta! Pasta ... What the heck is my...

Reaxys Quick search Query builder Results Retrosynthesis History Alerts + Reaxys AI Search Beta

191 Reactions out of 76 Documents, containing 327 Substances, 63 Targets

0 selected Limit To Exclude Export Hide Conditions

Sort by Reaxys Ranking

Filters

Limit to > Exclude >

By Structure >

Yield >

Reagent/Catalyst >

Solvent >

Catalyst Classes >

Solvent Classes >

Product Availability >

Reactant Availability >

Reaction Classes >

Document Type >

Publication Year >

☐ Single step reactions only

☐ Experimental procedure only

1

2

Conditions Find Similar > Reaction ID: 28432258

Conditions

With hydrogen; palladium 10% on activated carbon in methanol at 25°C; for 24h; Experimental Procedure >

With palladium 10% on activated carbon; hydrogen in methanol at 25°C; for 24h; Inert atmosphere;

Yield Reference

76% Current Patent Assignee: COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK - WO2009/38731, 2009, A2 Location in patent: Page/Page column 357; 4/37; 24/37 Full Text > Details > Abstract >

76% Snyder, Scott A.; Breazzano, Steven P.; Ross, Audrey G.; Lin, Yunqing; Zografos, Alexandros I. Journal of the American Chemical Society, 2009, vol. 131, p. 1753 - 1765 Full Text > Details > Abstract >

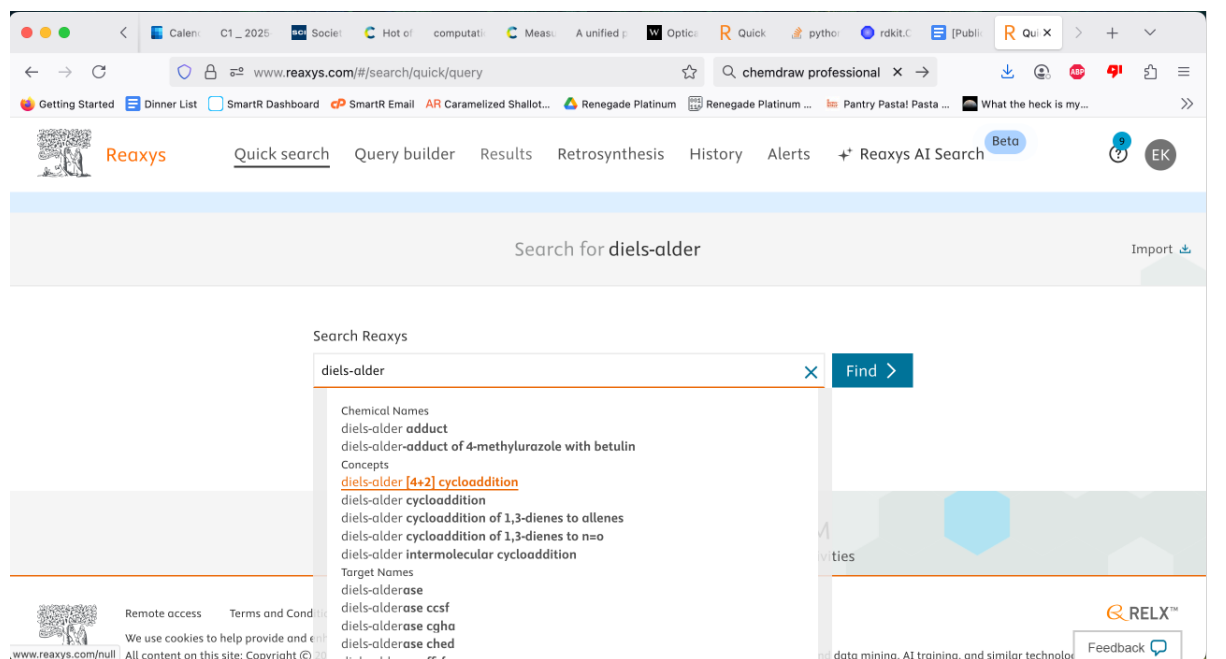
Feedback

TIP: If you don't get any hits for your reaction in SciFinder or Reaxys, try making it more general: remove rings or functional groups that do not play a part in the reaction. If you get too many irrelevant hits, try adding explicit hydrogens or use the atom mapping functions.

Curating Data from Reaxys

We will often need to augment our own experimental datasets with information from Reaxys. SciFinder is a great resource for synthesis, but their download feature is not as good as Reaxys'. There are a couple of ways to gather information. The first is through the Quick Search feature.

Here we see the potential Quick Search options for "Diels-Alder". When searching for reactions that fit a particular mechanism, you want to select options in "concepts". Click on the concept of interest then click "Find".



Search for diels-alder

Search Reaxys

diels-alder

Find

Chemical Names

- diels-alder adduct
- diels-alder-adduct of 4-methylurazole with betulin

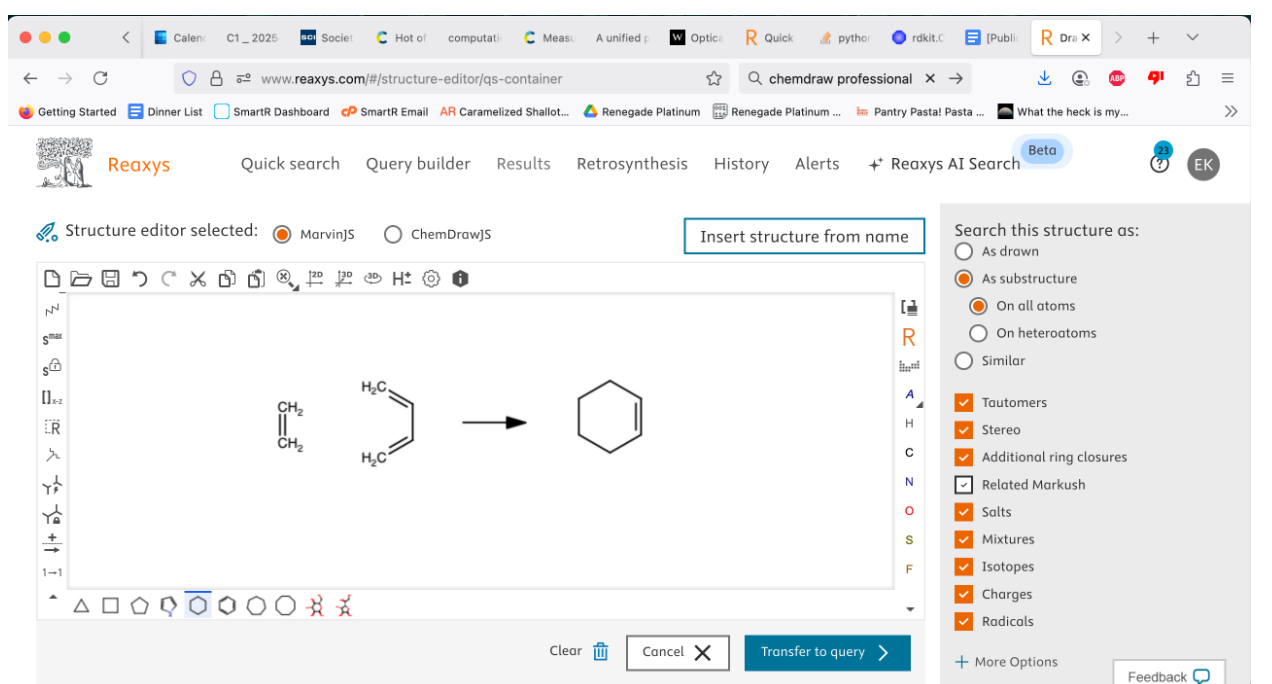
Concepts

- diels-alder [4+2] cycloaddition**
- diels-alder cycloaddition
- diels-alder cycloaddition of 1,3-dienes to allenes
- diels-alder cycloaddition of 1,3-dienes to n=0
- diels-alder intermolecular cycloaddition

Target Names

- diels-alderase
- diels-alderase ccsf
- diels-alderase cgha
- diels-alderase ched

Alternatively, you can draw the reaction of interest and hit "Transfer to query", then "Find".



Structure editor selected: ☒ MarvinJS ☐ ChemDrawJS

Insert structure from name

Search this structure as:

- ☐ As drawn
- ☒ As substructure
- ☐ On all atoms
- ☐ On heteroatoms
- ☐ Similar

- ☒ Tautomers
- ☒ Stereo
- ☒ Additional ring closures
- ☒ Related Markush
- ☒ Salts
- ☒ Mixtures
- ☒ Isotopes
- ☒ Charges
- ☒ Radicals

More Options

Feedback

Clear Cancel Transfer to query

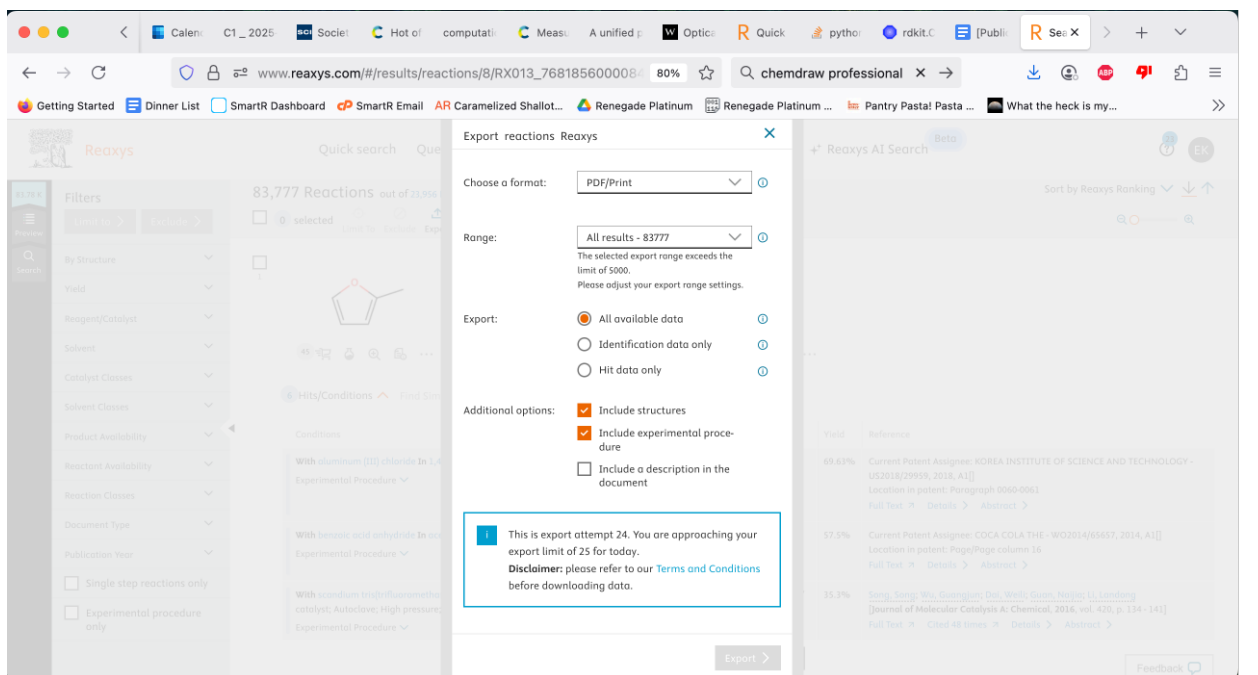
In either case, "Reactions" are what we are interested in. Click "View Results".

The screenshot shows the Reaxys search results page for the query "diels-alder cycloaddition". The page displays two main categories: Reactions (83,777) and Documents (16,826). The Reactions section is highlighted with a blue icon of a flask. The Documents section is highlighted with a blue icon of a document. Both sections have a "View Results" button. The page also includes a navigation bar with links like "Quick search", "Query builder", "Results", "Retrosynthesis", "History", "Alerts", and "Reaxys AI Search". At the bottom, there is a footer with "Remote access", "Terms and Conditions", "Privacy policy", "About content", "Accessibility", "Contact support", and a "Feedback" button.

All the reactions that Reaxys believes match your query will appear in list form. To export the reactions, click on "Export". A box will pop up.

The screenshot shows the Reaxys results page for the query "diels-alder cycloaddition". The page displays 83,777 Reactions out of 23,956 Documents, containing 147,273 Substances, 3,119 Targets. The "Export" button is highlighted with a red box. The page also includes a navigation bar with links like "Quick search", "Query builder", "Results", "Retrosynthesis", "History", "Alerts", and "Reaxys AI Search". At the bottom, there is a footer with "Remote access", "Terms and Conditions", "Privacy policy", "About content", "Accessibility", "Contact support", and a "Feedback" button.

Conditions	Yield	Reference
With aluminum (III) chloride In 1,4-dioxane at 250°C; under 22502.3 Torr; for 24h; Reagent/catalyst; Solvent; Autoclave; Experimental Procedure	69.63%	Current Patent Assignee: KOREA INSTITUTE OF SCIENCE AND TECHNOLOGY - US2018/29959, 2018, A1[] Location in patent: Paragraph 0060-0061 Full Text Details Abstract
With benzoic acid anhydride In acetic acid at 240°C; under 22502.3 Torr; for 8h; Reagent/catalyst; Temperature; Experimental Procedure	57.5%	Current Patent Assignee: COCA COLA THE - WO2014/65657, 2014, A1[] Location in patent: Page/Page column 16 Full Text Details Abstract
With scandium tris(trifluoromethanesulfonate) In 1,4-dioxane at 199.84°C; under 26252.6 Torr; for 24h; Kinetics; Reagent/catalyst; Autoclave; High pressure; Inert atmosphere; Experimental Procedure	35.3%	Song, Song; Wu, Guangjun; Dai, Wei; Guan, Nalija; Li, Landong [Journal of Molecular Catalysis A: Chemical, 2016, vol. 420, p. 134 - 141] Full Text Cited 48 times Details Abstract



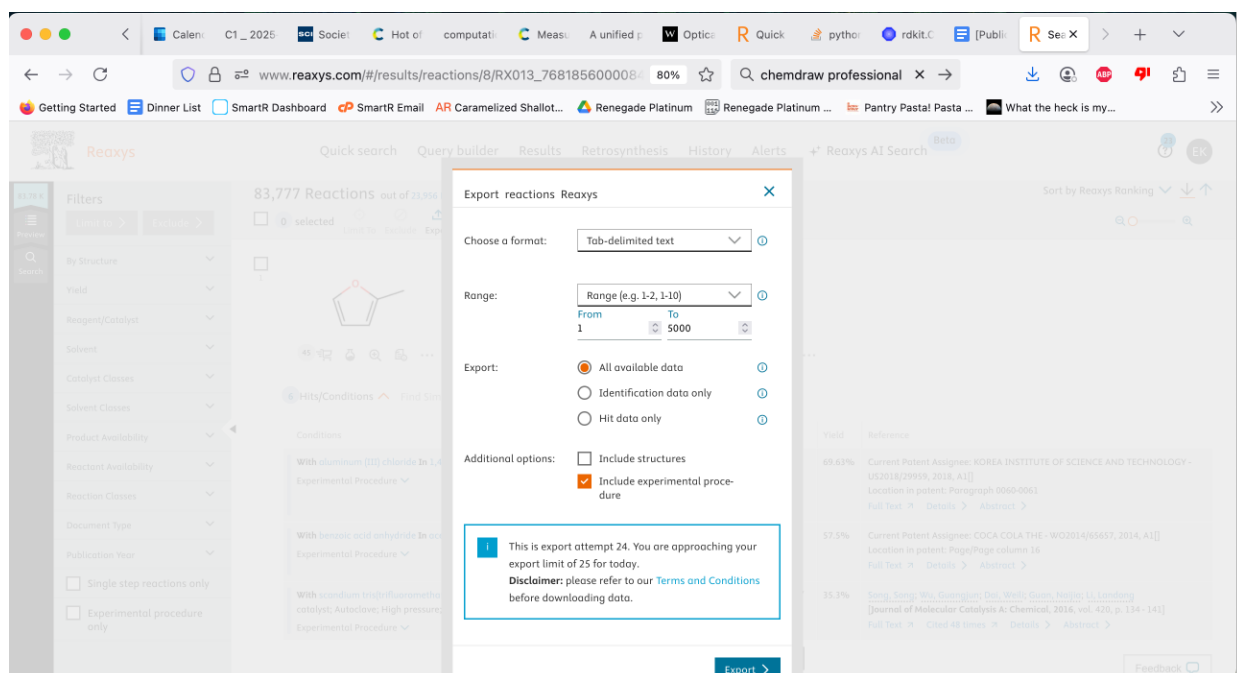
For format, typically we want to export to a tab-delimited text.

For range, we can only export 5,000 entries at a time, so set the range accordingly. It is normal to do 5K entries at a time until you get all of the reported reactions. **NOTE:** You only get 25 downloads a day, so at most you can download 125,000 reactions in a 24-hour period.

For export, we will typically want to get all available data.

We typically include experimental procedures but do not include structures.

Once you have entered all the information, hit "Export".



Your download will start and will let you know when you've finished as a pop up in the bottom right.

The screenshot displays the Reaxys web interface. The top navigation bar includes the Reaxys logo and various search and analysis tools. The main content area shows search results for a specific reaction, with 83,777 reactions found. A chemical reaction scheme is shown, involving a furan derivative and a benzene ring. Below the reaction scheme, a table lists conditions and references for the reaction. An 'Exports' pop-up window is visible in the bottom right corner, indicating that the export is ready for download.

Reaxys

Quick search Query builder Results Retrosynthesis History Alerts + Reaxys AI Search Beta

83,777 Reactions out of 23,956 Documents, containing 147,273 Substances, 3,119 Targets

Sort by Reaxys Ranking

6 selected Limit To Exclude Export Hide Conditions

By Structure Yield Reagent/Catalyst Solvent Catalyst Classes Solvent Classes Product Availability Reactant Availability Reaction Classes Document Type Publication Year

Single step reactions only Experimental procedure only

Reaction ID: 35171620

Conditions	Yield	Reference
With aluminum (III) chloride In 1,4-dioxane at 250°C; under 22502.3 Torr; for 24h; Reagent/catalyst; Solvent; Autoclave; Experimental Procedure	69.63%	Current Patent Assignee: KOREA INSTITUTE OF SCIENCE AND TECHNOLOGY - US2018/29959, 2018, A1[] Location in patent: Paragraph 0060-0061 Full Text Details Abstract
With benzoic acid anhydride In acetic acid at 240°C; under 22502.3 Torr; for 8h; Reagent/catalyst; Temperature; Experimental Procedure	57.5%	Current Patent Assignee: COCA COLA THE - WO2014/65657, 2014, A1[] Full Text
With scandium tris(trifluoromethanesulfonate) In 1,4-dioxane at 199.84°C; under 26252.6 Torr; for 24h; Kinetics; Reagent/catalyst; Autoclave; High pressure; Inert atmosphere; Experimental Procedure	35.3%	Song, S. Journal Full Text

+ Show all hits

Exports Feedback

Exports

Your export is ready

Download Cancel