

I. Créer un compte SciFinder.

II. En 10 minutes par outil, comparer Scifinder-n, Reaxys et ChempSpider d'après :

- leurs contenus (types d'informations, quantité...)
- l'accessibilité de l'outil et des données
- les éventuels points forts et points faibles
- essayez d'évaluer l'utilité pour un projet de type big data/machine learning

Si intéressés, ajoutez PubChem à la liste.

III. Chercher des informations sur la cetirizine dans les 3 outils (et d'autres si vous voulez) : structure, propriétés, publications qui mentionnent cette molécule, réactions...

IV. Chercher une réaction ayant comme produit l'hydroxychloroquine ; proposer une méthode de préparation de cette molécule

V. Chercher les molécules synthétisées dans le laboratoire du prof. Nicolai Cramer qui contiennent du soufre et de l'azote

VI. Identifier une molécule ayant les propriétés suivantes, et proposer un plan de synthèse à partir de molécules suffisamment simples :

- Point de fusion entre 98 et 100 degrés Celsius
- Masse moléculaire entre 300 et 305
- Structure contenant le fragment 1-phenylpiperidine, la pipéridine ne devant porter aucune autre substitution

VII. Chercher un brevet en rapport avec l'hydroxychloroquine

VIII. Dans Reaxys ou dans Scifinder, chercher des réactions réduisant la triple liaison d'un groupe nitrile lié à au moins un carbone C-CN en double liaison C-C=N. Combien en trouvez-vous?

IX. On autorise maintenant des substitutions quelconques sur le fragment pipéridine de l'ex. VI. Dans combien de publications liées à ces molécules est-il question de photocatalyse?

X. Installer Zotero (+ éventuellement Firefox) et le plug-in pour MS Word ou LibreOffice, récupérer quelques références de votre choix et les insérer dans un document. Formater la bibliographie au standard du journal Nature.

II. ChemSpider :

- **Can I download the complete data set?**

You can assemble a database of 5000 structures or less, and their associated properties, from our database without our permission. You can download up to 1000 structures per day from the website. If you want to do more than this, please contact us for help – we don't make the entire set available for free download. Don't forget we also have an extensive set of web services to connect to the ChemSpider database.

- **I'd like to scrape content from ChemSpider, can I do that?**

We do not allow the scraping of ChemSpider pages for a number of reasons, primarily because scraping places undue strain on our systems and reduces the responsiveness and functionality of the site for everybody else. In addition, some data is kindly made available by third parties who are happy to permit users to access and make use of data on an individual basis but do not allow us to make large sets of data available. If a few users seek to scrape this content it may be that we will be required to remove some data to the detriment of all users.

http://www.chemspider.com/FAQ.aspx#Download_Data

Scifinder-n: up to 500 substances in SDF format

https://scifinder-n.cas.org/help/#t=Working_with_Search_Results/Substances/Substances_Screen_-_Download_Answers.htm

Reaxys:

Limits – after logging in (otherwise not possible): The number of *exported* items is 5000, the number of *exports* per 24 hours is now limited to 25.

https://service.elsevier.com/app/answers/detail/a_id/11635/c/10546/supporthub/reaxys/kw/export/

Batch Queries https://service.elsevier.com/app/answers/detail/a_id/26151/supporthub/reaxys/kw/batch+queries/

PubChem:

Bulk download possible <https://pubchemdocs.ncbi.nlm.nih.gov/downloads>

III. <http://www.chemspider.com/Chemical-Structure.2577.html> (par ex.)

IV. Scifinder-n : hydroxychloroquine => 118-42-3 => 258 reactions => Product 195 réactions (or directly « Synthesize » tab after clicking on the substance in the search results). Scifinder-n : retrosynthèse relativement rapide (net progrès par rapport à l'an passé) ; on voit le prix des réactants, intéressant à comparer avec le produit fini. Reaxys search for name => substance => 93 réactions et 21 préparations. Autoplan => 5 plans de synthèse proposés

V. Scifinder-n Advanced Reference Search/Author Name Cramer, Nicolai => 264 references (250 dans CAPLUS) => 8107 substances but no contextual filter at this point anymore ; « product » could be by some other authors, 7814 => Search within results S + N substructure 818

VI. DANS REAXYS avec 2H sur chaque carbone de la piperidine => 2 résultats (**sinon 19**)

Chemical Name:

[2-Cyano-3-(1-Piperidinyl)Phenylamino]Oxoacetic Acid Ethyl Ester

Reaxys Registry Number: 5123203

CAS Registry Number: 63365-15-1

Type of Substance: heterocyclic

Molecular Formula: C₁₆H₁₉N₃O₃

Linear Structure Formula: C₁₆H₁₉N₃O₃

Molecular Weight: 301.345
InChI Key: HFJOOHKASRVNQJ-UHFFFAOYSA-N

Chemical Name:
6,9-dimethyl-7-piperidin-1-yl-2,3,4,4a-tetrahydro-1H-benzo[4,5]oxazolo[3,2-a]pyridin-8-ol

Reaxys Registry Number: 1592975
CAS Registry Number: 18735-58-5
Type of Substance: heterocyclic
Molecular Formula: C₁₈H₂₆N₂O₂
Linear Structure Formula: C₁₈H₂₆N₂O₂
Molecular Weight: 302.417
InChI Key: JDESLOKBKFKHWM-UHFFFAOYSA-N

Dans Scifinder-n : on peut quantifier le poids moléculaire mais on ne peut que vérifier l'existence du point de fusion, on n'arrive pas à descendre en-dessous de 56 substances à vérifier à la main ?

VII. Par exemple Scifinder-n -> Substance Identifier hydroxychloroquine -> Get References 12'151 publiés (8596 en mai 2022) -> Document Type filter Patent 1482 brevets. Google Patents : > 100'000 résultats ! Qui a raison ? Limité à publication 2019 64 vs. 7127

Un qui n'a pas le mot dans le titre :

[https://patents.google.com/patent/CA3077089A1/en?](https://patents.google.com/patent/CA3077089A1/en?q=(hydroxychloroquine)&before=publication:20191231&after=publication:20190101)

[q=\(hydroxychloroquine\)&before=publication:20191231&after=publication:20190101](https://patents.google.com/patent/CA3077089A1/en?q=(hydroxychloroquine)&before=publication:20191231&after=publication:20190101) pas dans les résultats, pourquoi ? Dans Scifinder, lié à hydroxychloroquine sulfate ! => en fait 117 substances au lieu de 2 mais on n'arrive qu'à 1654 brevets (71 en 2019)

[https://patents.google.com/patent/CN109311967A/en?](https://patents.google.com/patent/CN109311967A/en?q=(hydroxychloroquine)&country=CN&before=publication:20191231&after=publication:20190101&num=100)

[q=\(hydroxychloroquine\)&country=CN&before=publication:20191231&after=publication:20190101&num=100](https://patents.google.com/patent/CN109311967A/en?q=(hydroxychloroquine)&country=CN&before=publication:20191231&after=publication:20190101&num=100) est dans SFS mais pas associé à l'hydroxychloroquine. 336 pages en chinois avec une traduction automatique dans Google, difficile d'évaluer le niveau d'importance et d'exactitude

VIII. Reaxys: 1,175,469 réactions SLOOOOOOW, can freeze at 98+ % -> 763'047 avec mapping (!?! résultats encore très bruités; un peu plus concluant en excluant les tautomères 141,896) (~5 minutes par recherche)

Scifinder-n: 4,733,875 -> 278,890

IX. Scifinder-n (pas aidés par le filtre qualitatif du point de fusion, laissons tomber): 447 substances 6438 refs => 4 avec photocatalysis

Reaxys [2-methoxy-4-\[4-\(4-methylpiperazin-1-yl\)-piperidin-1-yl\]-phenylamine](#)

1211284-93-3 (trier par nombre de docs, 36 documents dans Reaxys) refine docs par Index term « hydrogenation » => 4 publiés

Idées à explorer si le temps le permet :

- Créer des recherches Reaxys en json
- Approfondir les recherches par mots-clés