

Pharmaceutical Chemistry

QSAR through the OCHEM portal



SAPIENZA
UNIVERSITÀ DI ROMA

Introduction

J Comput Aided Mol Des (2011) 25:533–554
DOI 10.1007/s10822-011-9440-2

Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information

The Online Chemical Modeling Environment is a web-based platform that aims to automate and simplify the typical steps required for QSAR modeling. The platform consists of two major subsystems: the database of experimental measurements and the modeling framework. A user-contributed database contains a set of tools for easy input, search and modification of thousands of records. The OCHEM database is based on the wiki principle and focuses primarily on the quality and verifiability of the data. The database is tightly integrated with the modeling framework, which supports all the steps required to create a predictive model: data search, calculation and selection of a vast variety of molecular descriptors, application of machine learning methods, validation, analysis of the model and assessment of the applicability domain. As compared to other similar systems, OCHEM is not intended to re-implement the existing tools or models but rather to invite the original authors to contribute their results, make them publicly available, share them with other users and to become members of the growing research community. Our intention is to make OCHEM a widely used platform to perform the QSPR/QSAR studies online and share it with other users on the Web. The ultimate goal of OCHEM is collecting all possible chemoinformatics tools within one simple, reliable and user-friendly resource. The OCHEM is free for web users and it is available online at <http://www.ochem.eu>

Preparing the dataset

Structure-Based Design of 6-Chloro-4-aminoquinazoline-2-carboxamide Derivatives as Potent and Selective p21-Activated Kinase 4 (PAK4) Inhibitors | Journal of Medicinal Chemi...

Autenticazione | Home | Sapienz... | HomePage - Citt... | Online Chemical... | Online Chemical... | Journal of Medic... | Structure-Based... | +

https://pubs.acs.org/doi/10.1021/acs.jmedchem.7b01342

ACS Publications C&EN CAS Access provided by UNIVERSITA DI ROMA LA SAPIENZA Log In

ACS Publications Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

My Activity Publications

RETURN TO ISSUE | < PREV ARTICLE NEXT >

Structure-Based Design of 6-Chloro-4-aminoquinazoline-2-carboxamide Derivatives as Potent and Selective p21-Activated Kinase 4 (PAK4) Inhibitors

Chenzhou Hao¹, Fan Zhao¹, Hongyan Song¹, Jing Guo¹, Xiaodong Li¹, Xiaolin Jiang¹, Ran Huan¹, Shuai Song¹, Qiaoling Zhang¹, Ruifeng Wang¹, Kai Wang¹, Yu Pang¹, Tongchao Liu¹, Tianqi Lu², Wanxu Huang¹, Jian Wang¹, Bin Lin¹, Zhonggui He¹, Haitao Li¹, Feng Li², Dongmei Zhao^{1*} and Maosheng Cheng^{1*}

View Author Information

Cite this: *J. Med. Chem.* 2018, 61, 1, 265-285
Publication Date: November 30, 2017
https://doi.org/10.1021/acs.jmedchem.7b01342
Copyright © 2017 © 2017 American Chemical Society

Article Views 6322 Altmetric 4 Citations -

LEARN ABOUT THESE METRICS

Share Add to Export

Journal of Medicinal Chemistry

PDF (8 MB) Supporting Info (2)

Abstract

Herein, we report the discovery and characterization of a novel class of PAK4 inhibitors with a quinazoline scaffold. Based on the shape and chemical composition of the ATP-binding pocket of PAKs, we chose a 2,4-diaminoquinazoline series of inhibitors as a starting point. Guided by X-ray crystallography and a structure-based drug design (SBDD) approach, a series of novel 4-aminoquinazoline-2-carboxamide PAK4 inhibitors were designed and synthesized. The inhibitors' selectivity, therapeutic potency, and pharmaceutical properties were optimized. One of the best compounds, **31** (CZh226), showed remarkable PAK4 selectivity (346-fold vs PAK1) and favorable kinase selectivity profile. Moreover, this compound potently inhibited the migration and invasion of A549 tumor cells by regulating the PAK4-directed downstream signaling pathways *in vitro*. Taken together, these data support the further development of **31** as a lead compound for PAK4-targeted anticancer drug discovery and as a valuable research probe for the further biological investigation of group II PAKs.

This website uses cookies to improve your user experience. By continuing to use the site, you are accepting our use of cookies. Read the ACS privacy policy.

CONTINUE

Preparing the dataset

Structure-Based Design of 6-Chloro-4-aminoquinazoline-2-carboxamide Derivatives as Potent and Selective p21-Activated Kinase 4 (PAK4) Inhibitors | Journal of Medicinal Chem...

Autenticazione | Home | Sapientz | HomePage - Citt | Online Chemical | Online Chemical | Journal of Medic | Structure-Based

https://pubs.acs.org/doi/10.1021/acs.jmedchem.7b01342

Journal of Medicinal Chemistry

Structure-Based Design of 6-Chloro-4-aminoquinazoline-2-carboxamide Derivatives as Potent and Selective p21-Activated Kinase 4 (PAK4) Inhibitors

Introduction

ARTICLE SECTIONS

The p21-activated kinases (PAKs) are serine/threonine (Ser/Thr) protein kinases that have been identified as downstream signaling effectors of Rho-family GTPases.(1,2) The six mammalian PAK isoforms are categorized into two groups: PAK1–3 (group I) and PAK4–6 (group II), based on their structural homologies and biochemical features.(3) As key components of the Ras-Rac/Cdc42-PAK pathway, PAKs have pivotal roles in many fundamental cellular processes, including cytoskeletal reorganization, focal adhesion, cell motility, morphological changes, cell-cycle progression, etc.(2) Moreover, the overexpression, amplification, and mutational activation of PAK isoforms, in particular, PAK1 and PAK4, have been linked to many human diseases, including breast cancer, lung cancer, prostate cancer, colon cancer, and human head and neck squamous cell carcinoma.(4) Consequently, PAKs have emerged as attractive targets for new anticancer therapies and have been the subject of extensive drug discovery efforts.(5)

Although the two groups of PAK proteins are similar in overall sequence and structure, they are differentiated by their tissue expression profiles,(6) subcellular localization,(7) GTPase specificity,(7) activation mechanism,(8) and downstream substrate specificity.(2,7) Studies using knockout mice lacking one or more specific PAK isoforms revealed the role of each isoform in normal tissue development, with phenotypes that range from no apparent effect to early embryonic death.(2,9–10) Among all of the PAKs, PAK4 is the most studied group II PAK member, and it has a place at critical nodal points in multiple signaling pathways that are associated with cell growth, cytoskeletal dynamics, cell polarity, survival, and development.(11) PAK4 is particularly highly expressed in prostate, testis, lung, heart, brain, and liver.(12) It has attracted considerable interest because of its role in cancer invasion, metastasis, and proliferation of BRAF- or KRAS-driven cancers.(13) In addition to PAK4, there is emerging evidence for the roles for PAK5 and PAK6 in cancer progression.(14) Moreover, a recent study revealed that PAK2 inhibition correlates with increased acute cardiovascular toxicity, which may be enhanced by PAK1 inhibition.(15) Thus, the development of specific and potent PAK4 inhibitors is highly desirable for minimizing the risk of the potential side effects associated with the inhibition of normal function of group I PAKs and will also shed further light on its role in cancer progression.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.7b01342.

Sequence alignment of the active site residues of PAK1–6; crystallographic parameters of all X-ray structures (cocrystal structures of 10a, 30, and 31 bound to PAK4); comparison of the orientation differences of the DFG Asp (Asp458^{PAK4}/Asp407^{PAK1}); binding mode analysis of compounds 12, 17, 18, and 20; detailed kinase selectivity data of compound 31; ¹H NMR, ¹³C MMR, and HRMS spectra of compounds 10a–d, 11a–d, and 12–41; liver microsomal stability, rat plasma stability, physicochemical properties determination, CYP450, and hERG inhibition reports of compound 31 (PDF) Molecular formula strings and associated biological data (CSV)

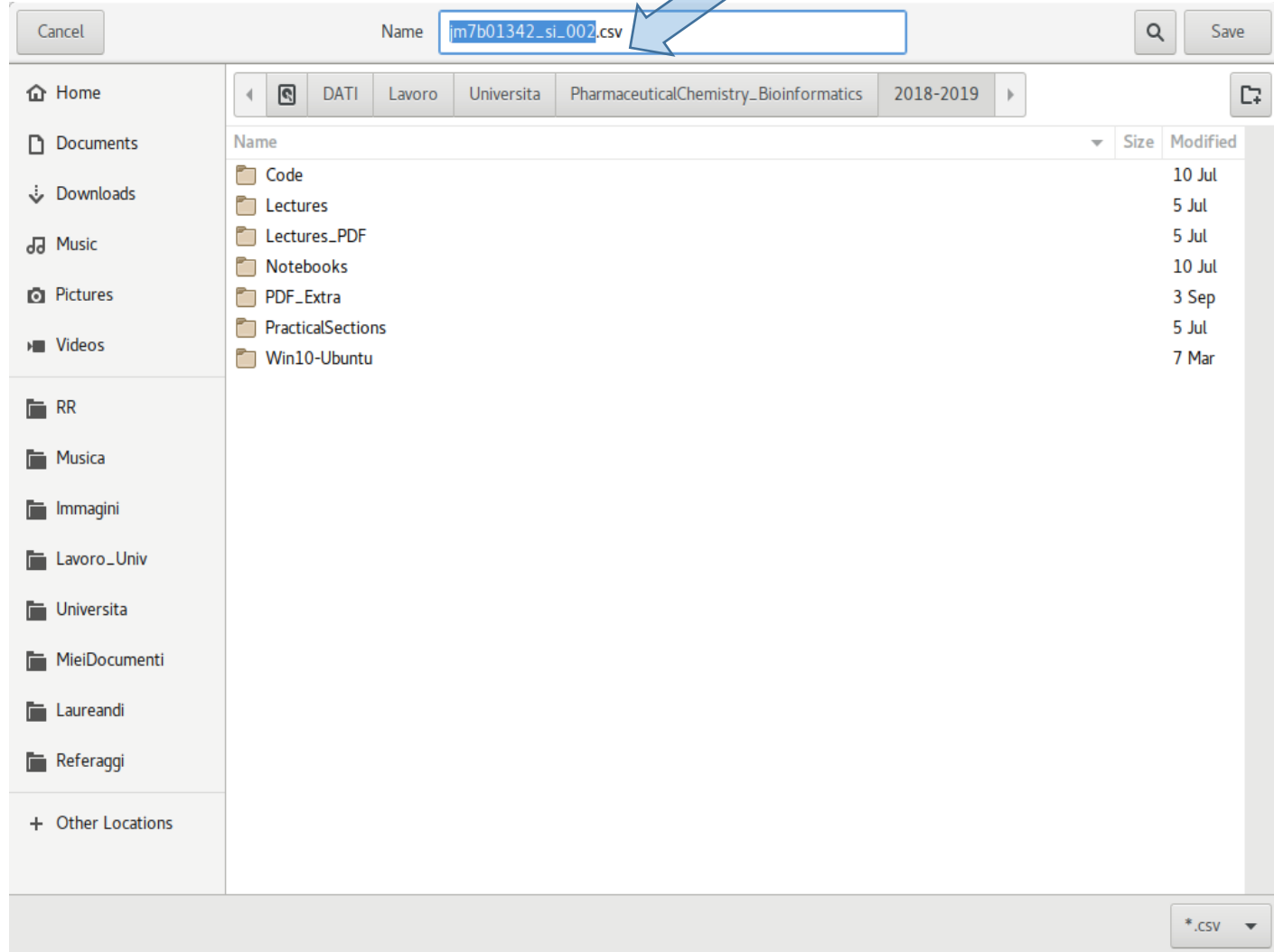
+ jm7b01342_sl002.csv (4.13 kb)

pdf

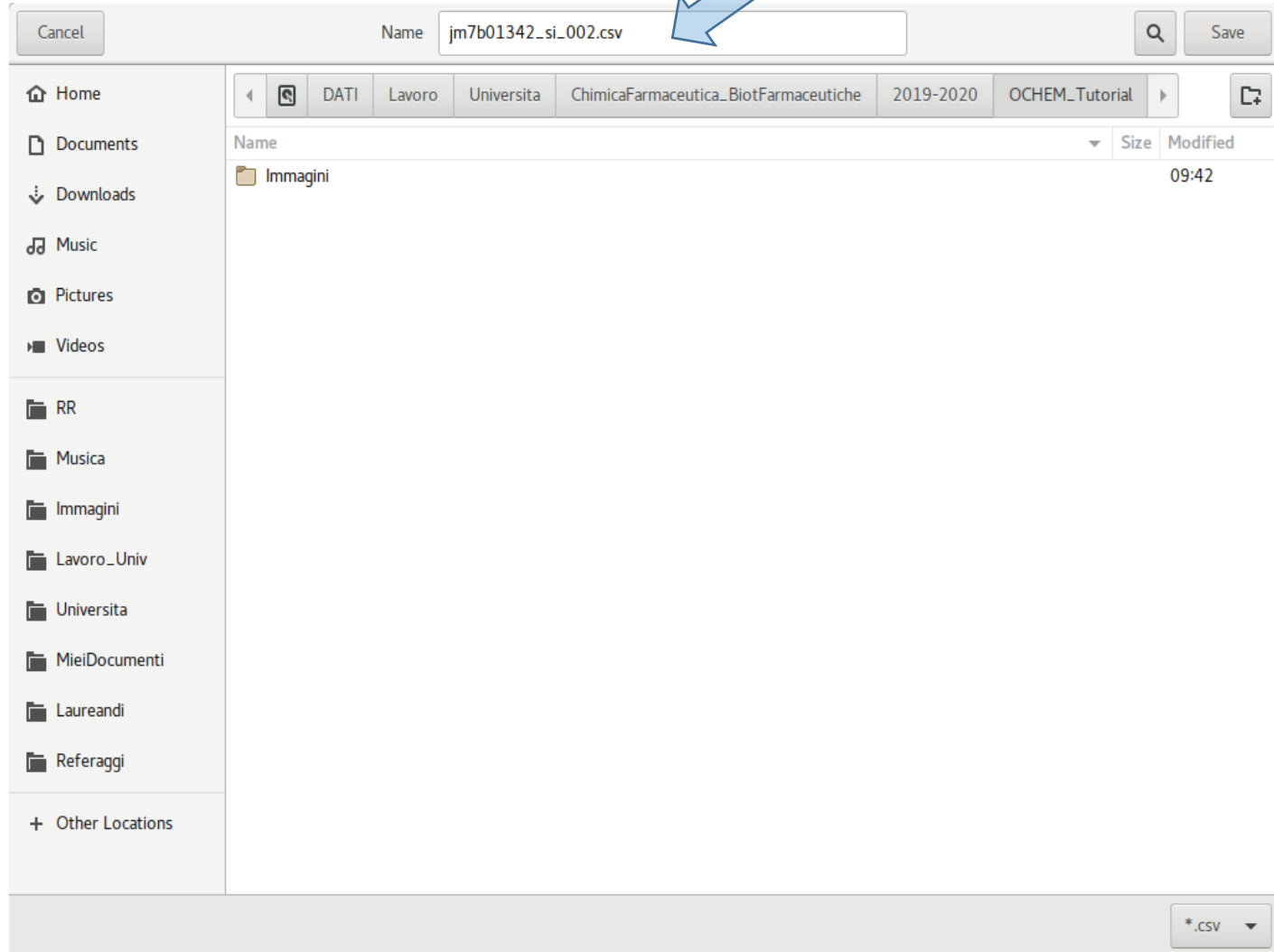
+ jm7b01342_sl001.pdf (8.69 MB)

This website uses cookies to improve your user experience. By continuing to use the site, you are accepting our use of cookies. Read the ACS privacy policy. **CONTINUE**

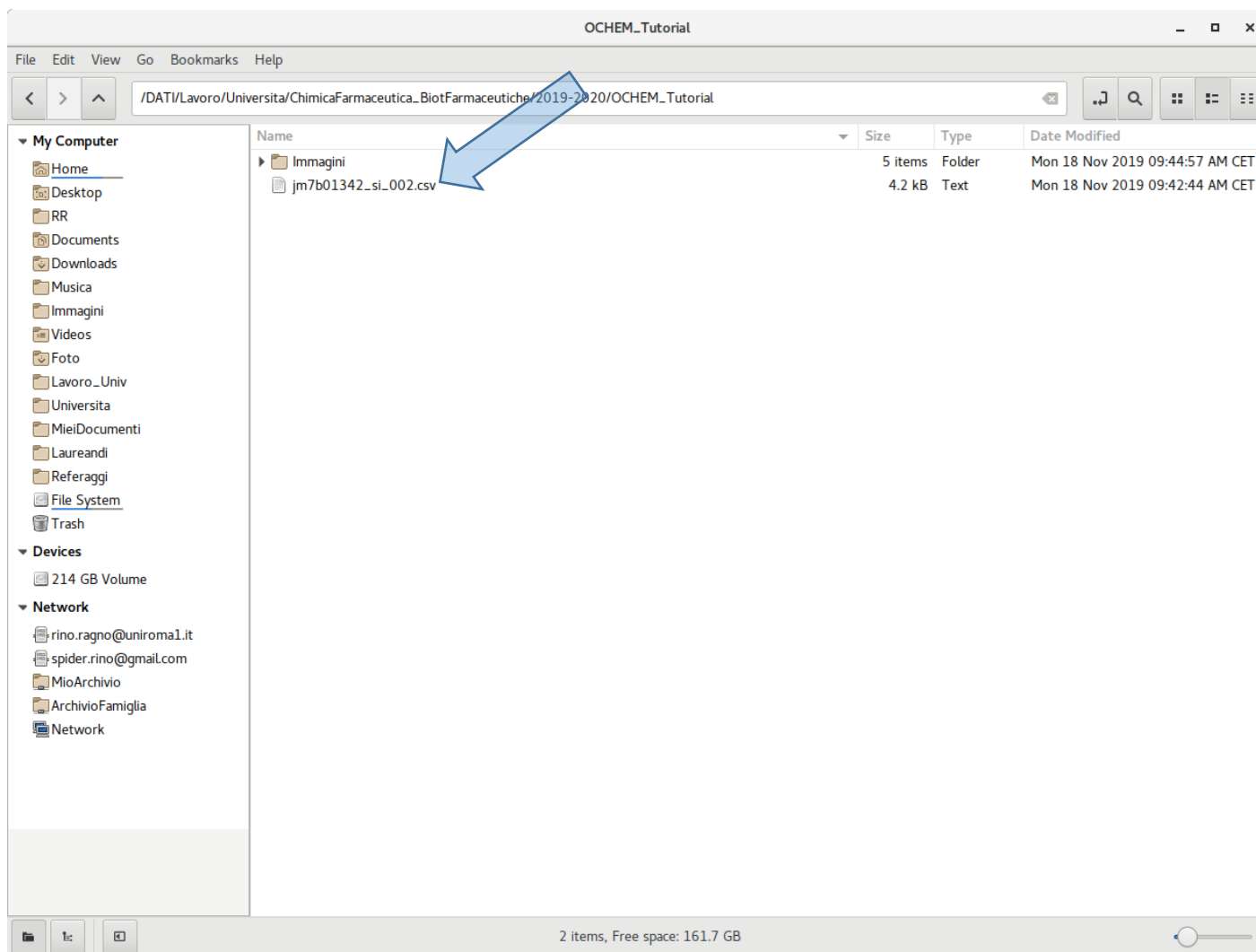
Preparing the dataset



Preparing the dataset



Preparing the dataset



Preparing the dataset

The screenshot shows a file explorer window titled "OCHEM_Tutorial" with a menu bar (File, Edit, View, Go, Bookmarks, Help) and a search bar. The address bar shows the path: "/DATI/Lavoro/Universita/ChimicaFarmaceutica..BiotFarmaceutiche/2019-2020/OCHEM_Tutorial". The file list shows a folder "Immagini" and a file "jm7b01342_si_002.csv" (4.2 kB, Text). A "Text Import" dialog box is open over the file, titled "Text Import - [jm7b01342_si_002.csv] (on rcmd-ProBook.rcmd.it)". The dialog has the following sections:

- Import:** Character set: Western Europe (ISO-8859-14); Language: English (USA); From row: 1.
- Separator Options:** Fixed width, Separated by. Under "Separated by": Tab, Comma, Semicolon, Space, Other. Merge delimiters, Trim spaces. String delimiter: " ". A blue arrow points to the "Comma" checkbox.
- Other Options:** Format quoted field as text, Detect special numbers.
- Fields:** Column type: [dropdown]. A table with 2 columns: "Standard" and "SMILES".

| | Standard | SMILES |
|---|---------------|---|
| 1 | Compound | SMILES |
| 2 | Staurosporine | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)NSC7(C)OC([H]</chem> |
| 3 | PF3758309 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)N(C2(C)C)CC3=C2N</chem> |
| 4 | 10a | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C</chem> |
| 5 | 10b | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4</chem> |
| 6 | 10c | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4</chem> |
| 7 | 10d | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5</chem> |
| 8 | 11a | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C4?)=N</chem> |

At the bottom of the dialog are buttons for "Help", "OK", and "Cancel". The status bar at the bottom of the file explorer shows: "jm7b01342_si_002.csv" selected (4.2 kB), Free space: 161.7 GB.

Preparing the dataset

The screenshot shows a spreadsheet with two columns: 'Compound' and 'SMILES'. The data is as follows:

| Compound | SMILES |
|---------------|--|
| staurosporine | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@H](NC)[C@@H]7O</chem> |
| PF3758309 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3NC4=NC(C)=NC5=C4</chem> |
| 10a | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(CI)C=C42</chem> |
| 10b | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C1</chem> |
| 10c | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1</chem> |
| 10d | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N4)C3=C1</chem> |
| 11a | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(CI)C=C42)=NN1</chem> |
| 11b | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C1</chem> |
| 11c | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1</chem> |
| 11d | <chem>CIC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C4)C3=C1</chem> |
| 12 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 13 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 14 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 15 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 16 | <chem>CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 17 | <chem>CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 18 | <chem>CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 19 | <chem>CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 20 | <chem>CIC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=C1</chem> |
| 21 | <chem>CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 22 | <chem>CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |
| 23 | <chem>CC1=CC(NC2=NC(C(N(C)C3CNCC3)=O)=NC4=CC=C(C=C42)CI)=NN1</chem> |

Preparing the dataset

jm7b01342_si_002.csv - LibreOffice Calc (on rcmd-ProBook.rcmd.it)

File Edit View Insert Format Styles Sheet Data Tools Window Help

LibreOffice Calc toolbar

fx = PAK1 Inhibitory Activity (Ki, ?M)

| | C | D | E | F | G | H | I | J |
|----|-------|-----------------------------------|-----------------------------------|--------------------|-------|------|------|-------|
| 1 | MW | PAK1 Inhibitory Activity (Ki, ?M) | PAK4 Inhibitory Activity (Ki, ?M) | Tmax-T0 (PAK4, FC) | clogP | PSA | | |
| 2 | 466 | 0.003 | 0.009 | ND | 3.7 | 69.8 | | |
| 3 | 490 | 0.052 | 0.026 | ND | 4.3 | 98.3 | | |
| 4 | 357.5 | >10 | 0.71 | 0.5 | 0.0 | 2.6 | 71.3 | |
| 5 | 383.5 | ND | 0.65 | 2.0 | 5 | 0.0 | 3 | 70.9 |
| 6 | 385.5 | ND | 0.88 | 4.0 | 8 | 0.3 | 3.2 | 70.4 |
| 7 | 419.5 | ND | 1.44 | 0.9 | 0.3 | | 3.8 | 71 |
| 8 | 357.5 | | 1.891 | 0.099 | 2.1 | 0.0 | 2.3 | 86.4 |
| 9 | 383.5 | | 0.288 | 0.016 | 4.7 | 0.3 | 2.9 | 85 |
| 10 | 385.5 | ND | | 0.077 | 2.3 | 0.3 | 3.1 | 84.8 |
| 11 | 419.5 | ND | | 0.256 | 0.7 | 0.3 | 3.5 | 85.1 |
| 12 | 385.5 | ND | | 0.354 | 0.5 | 0.0 | 1.7 | 108.9 |
| 13 | 386.5 | ND | | 3.817 | 0.1 | 0.3 | 2.1 | 105.3 |
| 14 | 399.5 | ND | | 0.674 | 0.4 | 0.3 | 2.1 | 104.3 |
| 15 | 400.5 | ND | | 3.89 | -0.1 | 0.0 | 2.9 | 101.5 |
| 16 | 399.5 | ND | | 0.672 | 0.3 | 0.3 | 2.3 | 97 |
| 17 | 371.5 | >10 | | 0.045 | 2.4 | 0.5 | 1.6 | 98.7 |
| 18 | 372.5 | ND | | 2.172 | 0.2 | 0.3 | 2.2 | 92.5 |
| 19 | 385.5 | ND | | 0.768 | ND | | 2 | 88.9 |
| 20 | 385.5 | ND | | 2.408 | ND | | 0.8 | 132.2 |
| 21 | 399.5 | ND | | 0.576 | 0.2 | 0.3 | 2.7 | 95.7 |
| 22 | 385.5 | ND | | 0.325 | 0.2 | 0.3 | 2.4 | 108.8 |
| 23 | 399.5 | ND | | 1.630 | 0.0 | 0.3 | 2.7 | 95.6 |
| 24 | 357.5 | ND | | 0.782 | 0.7 | 0.3 | 2.8 | 111.4 |
| 25 | 397.5 | | 2.75 | 0.016 | 4.1 | 0.3 | 2.2 | 97.4 |
| 26 | 385.5 | >4.52 | | 0.218 | 1.5 | 0.3 | 2.4 | 91.4 |
| 27 | 411.5 | | 0.383 | 0.026 | 3.6 | 0.3 | 3 | 90.1 |
| 28 | 385.5 | >4.52 | | 0.151 | 1.8 | 0.3 | 1.9 | 92.5 |
| 29 | 411.5 | | 2.08 | 0.017 | 3.6 | 0.3 | 2.5 | 91.1 |

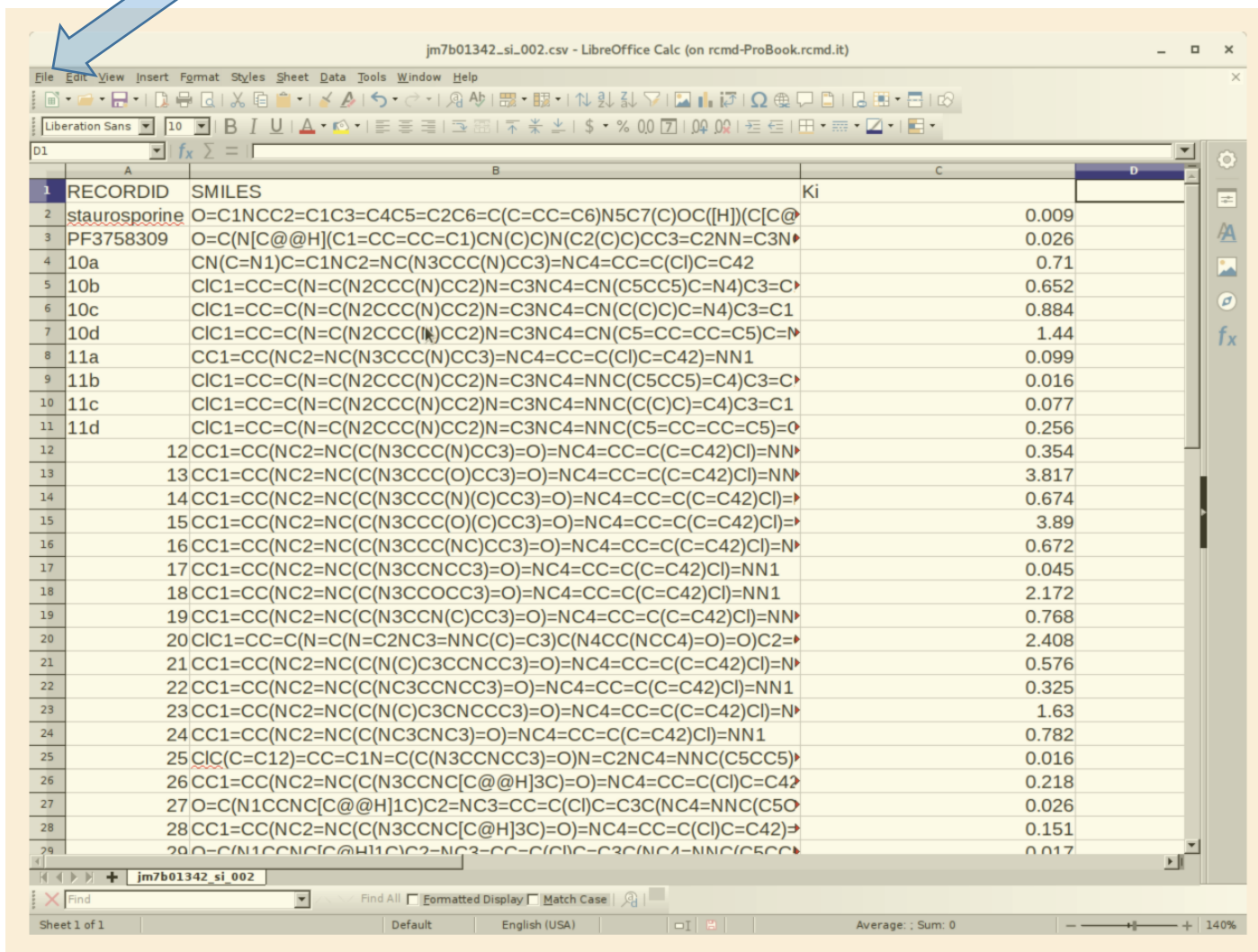
Sheet 1 of 1 | Default | English (USA) | Average: ; Sum: 0 | 140%

Preparing the dataset

The screenshot shows a LibreOffice Calc spreadsheet with the following data:

| | C | D | E | F | G | H | I | J |
|----|-------|-----------------------------------|-------|--------------------|-------|-----|-------|---|
| 1 | MW | PAK1 Inhibitory Activity (Ki, ?M) | Ki | Tmax-T0 (PAK4, FC) | clogP | PSA | | |
| 2 | 466 | | 0.003 | 0.009ND | | 3.7 | 69.8 | |
| 3 | 490 | | 0.052 | 0.026ND | | 4.3 | 98.3 | |
| 4 | 357.5 | >10 | | 0.710.5f0.0 | | 2.6 | 71.3 | |
| 5 | 383.5 | ND | | 0.6520.5f0.0 | | 3 | 70.9 | |
| 6 | 385.5 | ND | | 0.8840.8f0.3 | | 3.2 | 70.4 | |
| 7 | 419.5 | ND | | 1.440.9f0.3 | | 3.8 | 71 | |
| 8 | 357.5 | | 1.891 | 0.0992.1f0.0 | | 2.3 | 86.4 | |
| 9 | 383.5 | | 0.288 | 0.0164.7f0.3 | | 2.9 | 85 | |
| 10 | 385.5 | ND | | 0.0772.3f0.3 | | 3.1 | 84.8 | |
| 11 | 419.5 | ND | | 0.2560.7f0.3 | | 3.5 | 85.1 | |
| 12 | 385.5 | ND | | 0.3540.5f0.0 | | 1.7 | 108.9 | |
| 13 | 386.5 | ND | | 3.8170.1f0.3 | | 2.1 | 105.3 | |
| 14 | 399.5 | ND | | 0.6740.4f0.3 | | 2.1 | 104.3 | |
| 15 | 400.5 | ND | | 3.89-0.1f0.0 | | 2.9 | 101.5 | |
| 16 | 399.5 | ND | | 0.6720.3f0.3 | | 2.3 | 97 | |
| 17 | 371.5 | >10 | | 0.0452.4f0.5 | | 1.6 | 98.7 | |
| 18 | 372.5 | ND | | 2.1720.2f0.3 | | 2.2 | 92.5 | |
| 19 | 385.5 | ND | | 0.768ND | | 2 | 88.9 | |
| 20 | 385.5 | ND | | 2.408ND | | 0.8 | 132.2 | |
| 21 | 399.5 | ND | | 0.5760.2f0.3 | | 2.7 | 95.7 | |
| 22 | 385.5 | ND | | 0.3250.2f0.3 | | 2.4 | 108.8 | |
| 23 | 399.5 | ND | | 1.6300.0f0.3 | | 2.7 | 95.6 | |
| 24 | 357.5 | ND | | 0.7820.7f0.3 | | 2.8 | 111.4 | |
| 25 | 397.5 | | 2.75 | 0.0164.1f0.3 | | 2.2 | 97.4 | |
| 26 | 385.5 | >4.52 | | 0.2181.5f0.3 | | 2.4 | 91.4 | |
| 27 | 411.5 | | 0.383 | 0.0263.6f0.3 | | 3 | 90.1 | |
| 28 | 385.5 | >4.52 | | 0.1511.8f0.3 | | 1.9 | 92.5 | |
| 29 | 411.5 | | 2.08 | 0.0173.6f0.3 | | 2.5 | 91.1 | |

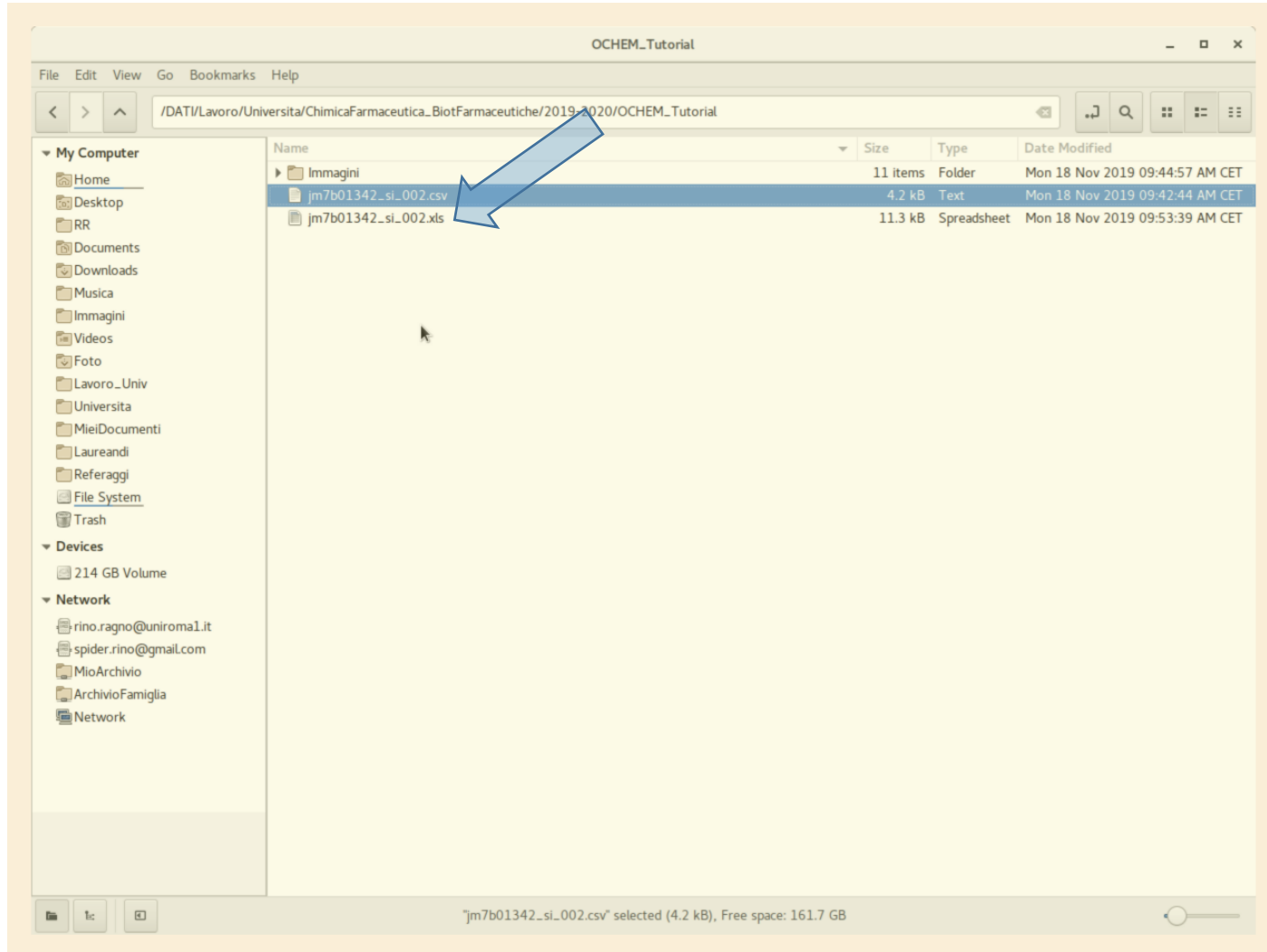
Preparing the dataset



The screenshot shows a LibreOffice Calc spreadsheet with the following data:

| RECORDID | SMILES | Ki |
|----------|---|-------|
| 1 | stauroporine | 0.009 |
| 3 | PF3758309 | 0.026 |
| 4 | 10a | 0.71 |
| 5 | 10b | 0.652 |
| 6 | 10c | 0.884 |
| 7 | 10d | 1.44 |
| 8 | 11a | 0.099 |
| 9 | 11b | 0.016 |
| 10 | 11c | 0.077 |
| 11 | 11d | 0.256 |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.354 |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 3.817 |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.674 |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 3.89 |
| 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.672 |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.045 |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 2.172 |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.768 |
| 20 | C1C=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=NN1 | 2.408 |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.576 |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.325 |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 1.63 |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.782 |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)=O | 0.016 |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.218 |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C=C3)C(NC4=NNC(C5CC5)=O)C2=O | 0.026 |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C=C4)Cl)=NN1 | 0.151 |
| 29 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C=C3)C(NC4=NNC(C5CC5)=O)C2=O | 0.017 |

Preparing the dataset



Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | Home Page - C... | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/home/show.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

v.3.0.96.1

log in create account

Home Database Models A+ a- Privacy statement

Welcome to OCHEM! Your possible actions

- Explore OCHEM data**
Search chemical and biological data: experimentally measured, published and exposed to public access by our users. You can also [upload your data](#).
- Create QSAR models**
Build QSAR models for predictions of chemical properties. The models can be based on the experimental data published in our database.
- Run predictions**
Apply one of the available models to predict property you are interested in for your set of compounds.
- Screen compounds with ToxAlerts**
Screen your compound libraries against structural alerts for such endpoints as mutagenicity, skin sensitization, aqueous toxicity, etc.
- Optimise your molecules**
Optimise different properties for your molecules (e.g., reduce their toxicity or improve their ADME properties) using the state-of-the-art MolOptimiser utility based on matched molecular pairs
- Tutorials**
<https://cbiankoni.epibio.waech/dhowtie> to know more about the

Check out the properties available on OCHEM

OCHEM contains 2854383 records for 638 properties (with at least 50 records) collected from 12973 sources

- Melting Point**
logPow logBB
LogL(water) LogD logPI(+)
- Water solubility**
LogL(blood) LogL(oil) ER
Cbrain/Cplasma IC50
- Papp(Caco-2) Papp(MDCK)**
- Oral absorption** LIC 50
Papp ratio(Caco-2)
- Plasma protein binding**
Papp ratio(MDCK-mdr1) pIC50
%Human FA Human IA
Human FA
- fraction unbound (fu)**

Latest active users

- rino.ragno: Prof. Rino Ragno seconds ago
- Xingguomeng: Miss. guomeng xing seconds ago
- vkovalishyn: Dr. Vasyi Kovalishyn seconds ago
- echmstry: Mr. Ely Setiawan seconds ago
- corde: Mr. Jose Andres Cordero Solano seconds ago
- zaira1: Mrs. Zaira Rehman about an hour ago

Latest published models

- IC50 model published by carpovpv 1 hours ago
- Ld50 mouse oral model published by Tinkov_Oleg 9 months ago
- Drug-Induced Rhabdomyolysis model published by qingshuang0501 9 months ago
- o ooo orl LD model published by

jm7b01342_sl_0...csv Show all


Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Citi | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +


https://ochem.eu/user/newuser.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

 **Online chemical database**
with modeling environment v.3.0.96.1

[log in](#) [create account](#)

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

 **User account**
Details of your personal OCHEM account

Registration Information

Login*
(min. 4 characters and max. 20 characters)

e-mail*

Password* Password can contain only letters and numbers.

Confirm password*

Personal Information

Title* Please, select your title.

First name*

Last name*

Affiliation*

Form of organization* Please, select your organization type.

City

jm7b01342_sl_0...csv

Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Cit | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/user/newuser.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

v.3.0.96.1

log in create account

Home Database Models A+ a- Privacy statement

Country

Zip

Phone

Position

Web

Terms of Service

Online License Agreement and Terms of Use

This is a legal agreement between you (the "USER"), and Helmholtz Zentrum Muenchen - Deutsches Forschungszentrum fuer Gesundheit und Umwelt ("HMGU"), covering your use of the Online Chemical Database & Modeling Environment and the accompanying Software ("OCHEM"). Be sure to read the following agreement before using OCHEM.

HMGU is willing to license the use of OCHEM upon the condition that you accept all of the terms contained in this License Agreement and any supplementary license terms included herewith ("Agreement").

By using OCHEM, you agree to be bound by the terms of this Agreement. If you do not agree to the terms of this Agreement, you should not use OCHEM.

By clicking on 'I accept' below I acknowledge that I have read and fully understand the foregoing information and agree to abide by License agreement above and the Privacy Policy.

I ACCEPT, CREATE MY ACCOUNT. I REJECT.

jm7b01342_sl_0...csv Show all


Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Citi | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/login/show.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

 **Online chemical database**
with modeling environment

v.3.0.96.1

log in create account

Home Database Models A+ a- Privacy statement

Please, login

Instant login

In order to access OCHEM, you must login. If you do not wish to register now, you can login as a guest. Guest users have access to less features than registered users.

[LOGIN AS A GUEST](#)


Already have an account?

If you already have an account, please enter you login and password below:

Login ID

Password

[LOGIN](#) [PASSWORD REMINDER](#)



Join OCHEM - register a new user!

Create a free account to upload data, create and apply QSAR models, screen chemical libraries and many more. Registered users can correct data uploaded by other registered users publish models. As a registered user, you can configure flexible access policies for your data and models.

[REGISTER A NEW USER](#)

jm7b01342_sl_0_...csv Show all

Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Citi | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/home/show.do?render-mode=full

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Welcome to OCHEM! Your possible actions

- Explore OCHEM data**
Search chemical and biological data: experimentally measured, published and exposed to public access by our users. You can also [upload your data](#).
- Create QSAR models**
Build QSAR models for predictions of chemical properties. The models can be based on the experimental data published in our database.
- Run predictions**
Apply one of the available models to predict property you are interested in for your set of compounds.
- Screen compounds with ToxAlerts**
Screen your compound libraries against structural alerts for such endpoints as mutagenicity, skin sensitization, aqueous toxicity, etc.
- Optimise your molecules**
Optimise different properties for your molecules (e.g., reduce their toxicity or improve their ADME properties) using the state-of-the art MolOptimiser utility based on matched molecular pairs
- Tutorials**
<https://fbiankoni.epibio.univ.it/show/tutorials-to-know-more-about-the>

Check out the properties available on OCHEM

OCHEM contains 2854383 records for 638 properties (with at least 50 records) collected from 12973 sources

- Melting Point**
logPow logBB
LogL(water) LogD logPI(+)
- Water solubility**
LogL(blood) LogL(oil) ER
Cbrain/Cplasma IC50
- Oral absorption** LIC 50
Papp(Caco-2) Papp(MDCK)
- Plasma protein binding**
Papp ratio(MDCK-mdr1) pIC50
%Human FA Human IA
Human FA
- fraction unbound (fu)**

Latest active users

- rino.ragno: Prof. Rino Ragno seconds ago
- vkovalishyn: Dr. Vasyl Kovalishyn seconds ago
- corde: Mr. Jose Andres Cordero Solano seconds ago
- Xingguomeng: Miss. guomeng xing 2 minutes ago
- echmstry: Mr. Ely Setiawan 8 minutes ago
- zaira1: Mrs. Zaira Rehman about 1 hours ago

Latest published models

- IC50 model published by carpovpv 1 hours ago
- Ld50 mouse oral model published by Tinkov_Oleg 9 months ago
- Drug-Induced Rhabdomyolysis model published by qingshuang0501 9 months ago
- o oaa orl LD model published by

jm7b01342_sl_0...csv Show all

Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Cit | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/home/show.do?render-mode=full

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Compound properties

- Properties
- Conditions
- Units
- Articles/Books
- Journals
- ToxAlerts
- MatchedPairs
- Baskets
- Tags
- Set area of interest...
- User-related changes
- Batch data upload**
- Trash

Check out the properties available on OCHEM

OCHEM contains 2854383 records for 638 properties (with at least 50 records) collected from 12973 sources

Melting Point
logPow logBB
LogL(water) LogD logPI(+)

Water solubility
LogL(blood) LogL(oil) ER

Cbrain/Cplasma IC50
Papp(Caco-2) Papp(MDCK)
Oral absorption LIC 50
Papp ratio(Caco-2)

Plasma protein binding
Papp ratio(MDCK-mdr1) pIC50
%Human FA Human IA
Human FA
fraction unbound (fu)

Latest active users

- rino.ragno: Prof. Rino Ragno seconds ago
- vkovalishyn: Dr. Vasyl Kovalishyn seconds ago
- corde: Mr. Jose Andres Cordero Solano seconds ago
- Xingguomeng: Miss. guomeng xing 2 minutes ago
- echmstry: Mr. Ely Setiawan 8 minutes ago
- zaira1: Mrs. Zaira Rehman about 1 hours ago

Latest published models

- IC50 model published by carpovpv 1 hours ago
- Ld50 mouse oral model published by Tinkov_Oleg 9 months ago
- Drug-Induced Rhabdomyolysis model published by qingshuang0501 9 months ago
- o oaa orl LD model published by

Screen compounds with ToxAlerts
Screen your compound libraries against structural alerts for such endpoints as mutagenicity, skin sensitization, aqueous toxicity, etc.

Optimise your molecules
Optimise different properties for your molecules (e.g., reduce their toxicity or improve their ADME properties) using the state-of-the-art MolOptimiser utility based on matched molecular pairs

Tutorials

https://ochem.eu/batchupload30/show.do

jm7b01342_sl_0_...csv

Show all

Preparing the dataset

The screenshot shows a web browser window titled "Online Chemical Modeling Environment - Google Chrome". The address bar displays the URL "https://ochem.eu/batchupload30/show.do". The page header includes the logo for "Online chemical database with modeling environment" and a user greeting: "Welcome, Dear Prof.Ragno! My account Logout". Navigation tabs for "Home", "Database", and "Models" are visible. The main content area is titled "Batch Upload 3.0 - File selection" and contains instructions for uploading data. A blue arrow points to the "Choose File" button in the file selection area. Below the instructions, there are checkboxes for "Settings" and an "Upload" button. At the bottom, a file list shows "jm7b01342_sl_0....csv" with a "Show all" button.

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Cit | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/batchupload30/show.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Batch Upload 3.0 - File selection
Select the CSV, SDF or Excel file to upload multiple records to the database.

Instructions

Here you have the possibility to upload data from an **excel sheet, sdf or csv** .
Backslash \ is used as stereochemistry in cvs, which should not contain "\uffff" characters.
The format of these data is strict, and can be viewed in [this sample](#) (scientific format) and [this sample](#) (technical format).
For more information, consider the wiki page that you can access by clicking on the wiki icon next to the title ("Batch upload browser").
If you have difficulties uploading your data, feel free to drop us an e-mail at info@ochem.eu.

Select a file to upload

Upload file
Choose File No file chosen

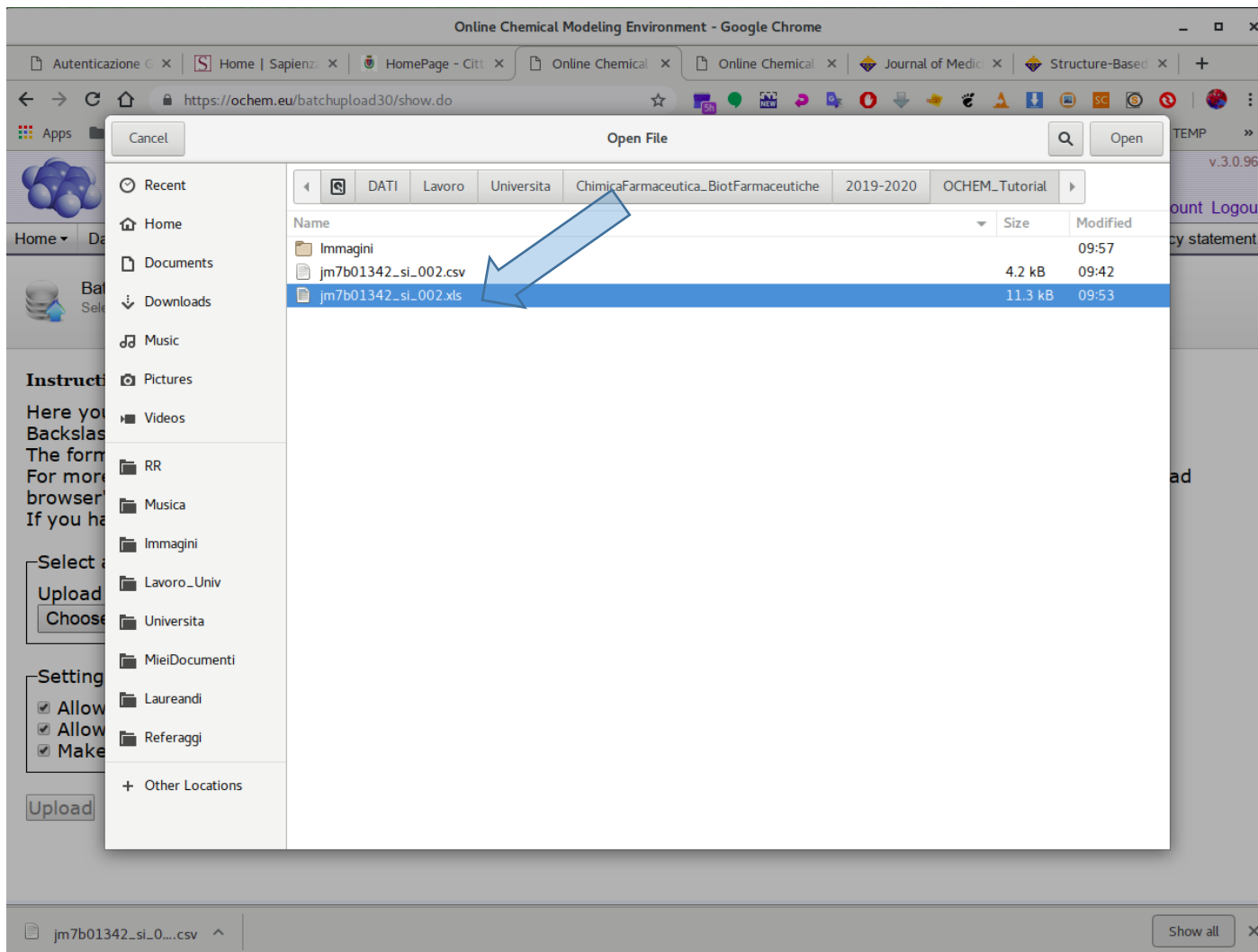
Settings

- Allow molecule lookup by name on PubChem
- Allow article lookup by PMID on PubMed
- Make the uploaded records hidden

Upload

jm7b01342_sl_0....csv Show all

Preparing the dataset



Preparing the dataset

Online Chemical Modeling Environment - Google Chrome

Autenticazione | Home | Sapienz | HomePage - Citi | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/batchupload30/show.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Batch Upload 3.0 - File selection
Select the CSV, SDF or Excel file to upload multiple records to the database.

Instructions

Here you have the possibility to upload data from an **excel sheet, sdf or csv** .
Backslash \ is used as stereochemistry in cvs, which should not contain "\uffff" characters.
The format of these data is strict, and can be viewed in [this sample](#) (scientific format) and [this sample](#) (technical format).
For more information, consider the wiki page that you can access by clicking on the wiki icon next to the title ("Batch upload browser").
If you have difficulties uploading your data, feel free to drop us an e-mail at info@ochem.eu.

Select a file to upload

Upload file
Choose File jm7b01342_si_002.xls

Settings

- Allow molecule lookup by name on PubChem
- Allow article lookup by PMID on PubMed
- Make the uploaded records hidden

Upload

jm7b01342_si_0...csv Show all

Preparing the dataset

Autenticazione x Home | Sapienz x HomePage - Citi x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/batchupload30/show.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Batch Upload 3.0 - File preview and column remapping
Preview your data, select the sheet and the columns you would like to upload

jm7b01342_si_002

| <input checked="" type="checkbox"/> RECORDID | <input checked="" type="checkbox"/> SMILES | <input checked="" type="checkbox"/> KI |
|--|---|--|
| staurosporine | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C...</chem> | 0.009000000000000001 |
| PF3758309 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C...</chem> | 0.026000000000000002 |
| 10a | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=C...</chem> | 0.71 |
| 10b | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 0.652 |
| 10c | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 0.884 |
| 10d | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 1.44 |
| 11a | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C...</chem> | 0.099 |
| 11b | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 0.016 |
| 11c | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 0.077 |
| 11d | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 0.256 |

The ARTICLE column is missing, the stub unpublished article will be assigned by default

Green titles indicate recognized columns, red titles indicate errors. Please click on the red columns and select whether the column indicates a property, condition or another column type like name, value or molecule, then select the matching entity and confirm your selection by clicking on the green button on the left.

If you have irrelevant columns in your sheet, you can leave them red and they will be ignored in the further process. If you

jm7b01342_si_0...csv Show all x

Preparing the dataset

Autenticazione x Home | Sapienz x HomePage - Citi x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/batchupload30/show.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP »

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

| <input checked="" type="checkbox"/> RECORDID | <input checked="" type="checkbox"/> SMILES | <input checked="" type="checkbox"/> Ki |
|--|--|--|
| staurosporine | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C... | 0.009000000000000001 |
| PF3758309 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C... | 0.026000000000000002 |
| 10a | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=C... | 0.71 |
| 10b | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(... | 0.652 |
| 10c | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(... | 0.884 |
| 10d | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(... | 1.44 |
| 11a | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C... | 0.099 |
| 11b | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC... | 0.016 |
| 11c | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC... | 0.077 |
| 11d | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC... | 0.256 |

The ARTICLE column is missing, the stub unpublished article will be assigned by default

Green titles indicate recognized columns, red titles indicate errors. Please click on the red columns and select whether the column indicates a property, condition or another column type like name, value or molecule, then select the matching entity and confirm your selection by clicking on the green button on the left.

If you have irrelevant columns in your sheet, you can leave them red and they will be ignored in the further process. If you need help, feel free to drop us an e-mail at info@ochem.eu.

Upload this sheet

Cancel Batch Upload Download Excel file

jm7b01342_si_0....csv Show all x

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/batchupload30/show.do>. The page title is "Online chemical database with modeling environment" and the version is v.3.0.96.1. The user is logged in as Prof. Ragno. The main content area is titled "Batch Upload 3.0 - Entity remapping" and includes a sub-section "Database entities remapping".

Property: **Ki**

Values
Unit: **-log(M)**, min value: 0.006, max value: 3.89

Article: **unpublished**

Molecule set: **default**

At the bottom, a file named "jm7b01342_si_0....csv" is visible with a "Show all" button.

Preparing the dataset

The screenshot shows the 'Online chemical database with modeling environment' website. The page title is 'Online chemical database with modeling environment' and the version is 'v.3.0.96.1'. The user is logged in as 'Prof.Ragno!'. The main navigation menu includes 'Home', 'Database', and 'Models'. The 'Batch upload (reloaded) X' and 'Select unit (Ki) X' tabs are active. The main content area displays a table of concentration units. The 'uM' unit is highlighted in green, and a blue arrow points to it. The table columns include the unit name, its type (Concentration), a 'Show records' link, a description, and the user's name.

| Unit | Type | Action | Description | User |
|------------|-----------------|--------------|------------------------------------|------------------------|
| pM | (Concentration) | Show records | 1 pM(picomolar) corresponds ... | Pankaj_OCHEM |
| ppb | (Concentration) | Show records | parts per billion Denotes one ... | bhataba / itetko |
| ppb food | (Concentration) | Show records | in food - parts per billion De ... | charochkina |
| pph | (Concentration) | Show records | parts per hundred, one molecu ... | bhataba / admin |
| ppm | (Concentration) | Show records | Parts per million ('ppm') deno ... | itetko |
| ppm food | (Concentration) | Show records | in food - Parts per million (' ... | charochkina |
| ppt | (Concentration) | Show records | Parts per trillion ('ppt') den ... | indykpol |
| ug | (Concentration) | Show records | ug/ml -- ug stands for micro g ... | Pankaj_OCHEM / enamine |
| ug/bee | (Concentration) | Show records | The quantity of substance in o ... | mojca / itetko |
| ug/g | (Concentration) | Show records | The unit to measure concentrat ... | charochkina |
| ug/kg | (Concentration) | Show records | The unit to measure concentrat ... | charochkina |
| ug/kg food | (Concentration) | Show records | The unit to measure concentrat ... | charochkina |
| ug/L | (Concentration) | Show records | ug/l -- micrograms per liter, ... | charochkina |
| ug/L/d | (Concentration) | Show records | Concentration in micrograms (u ... | charochkina |
| ug/m3 | (Concentration) | Show records | Microgram per cubic meter unit ... | bhataba / itetko |
| ug/ml | (Concentration) | Show records | Micrograms / milliliter = (10- ... | ucb |
| ul/L | (Concentration) | Show records | We assume that 1ul = 1 micro l ... | kamel / itetko |
| uM | (Concentration) | Show records | Micro mol corresponds to 10 | ucb |
| umol/L | (Concentration) | Show records | Micromol per liter. The mole i ... | amaziz / itetko |
| ug/cm^3 | (Concentration) | Show records | ug/cm^3 -- micrograms per cubi ... | tanzeem |
| ug/dcl | (Concentration) | Show records | ug/dcl -- micrograms per deci ... | kamel / tanzeem |
| ug/L | (Concentration) | Show records | ug/l -- micrograms per liter, ... | itetko |

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/batchupload30/show.do>. The page title is "Online chemical database with modeling environment" and the version is v.3.0.96.1. The user is logged in as "Prof.Ragno!". The main content area is titled "Batch Upload 3.0 - Entity remapping" and contains the following information:

- Database entities remapping**
- Property: **Ki**
- Values: Unit: **uM**, min value: 0.006, max value: 3.89
- Article: **unpublished**
- Molecule set: **default**

A blue arrow points to the "submit" button. At the bottom right, there are buttons for "Cancel Batch Upload" and "Download Excel file". The browser's taskbar at the bottom shows a file named "jm7b01342_si_0....csv" and a "Show all" button.

Preparing the dataset

Autenticazione x Home | Sapienz x HomePage - Citi x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/batchupload30/show.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Batch upload 3.0 - records preview Preview the records you are about to upload, select the desired actions

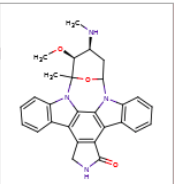
Batch upload preview browser

Summary:
All rows in the sheet Count: **40**
Status: error, Count: **40**

Filter by row number: and row type: **all**

1 - 10 of 40 items on page of 4 >>

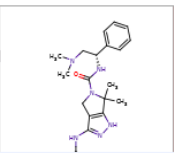
Row 1
 Save as error
 Skip



Ki = 0.009000000000000001 (in uM) = 8.05 (in -log(M))
Ragno, R
jm7b01342_si_002.xls...
N: AUTO_1
MoleculeID: M4402773
Some obligatory conditions for property Ki have not been specified: [Target] RecordID: R-1
rino.ragno Only visible to rino.ra

error: Some obligatory conditions for property Ki have not been specified: [Target]

Row 2
 Save as error
 Skip



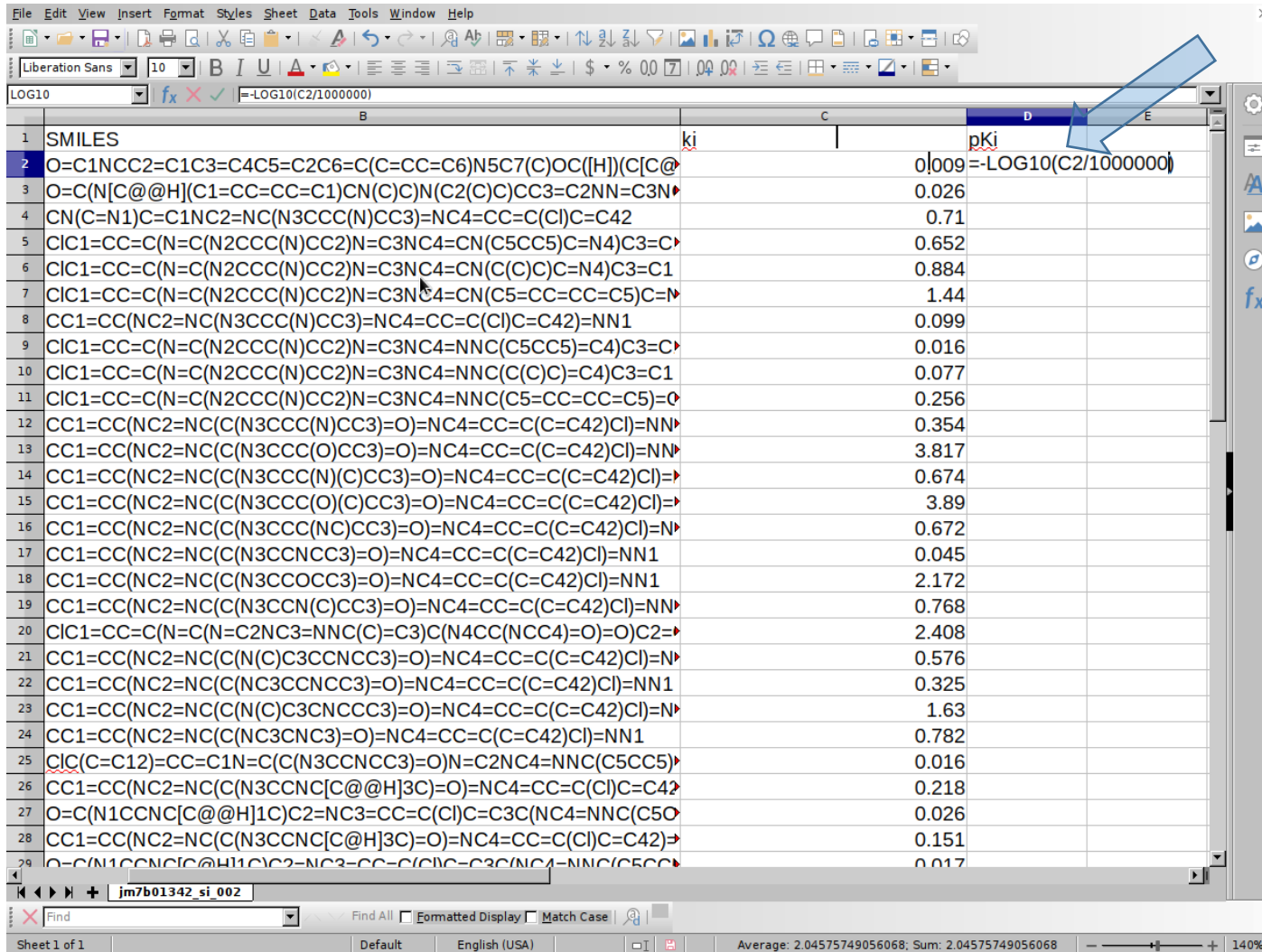
Ki = 0.026000000000000002 (in uM) = 7.59 (in -log(M))
Ragno, R
jm7b01342_si_002.xls...
N: AUTO_2
MoleculeID: M95419909
Some obligatory conditions for property Ki have not been specified: [Target] RecordID: R-2

jm7b01342_si_0...csv x

Preparing the dataset

| SMILES | | pKi |
|---|--|-------|
| <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C)C@H</chem> | | 0.009 |
| <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N</chem> | | 0.026 |
| <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C)C=C42</chem> | | 0.71 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C</chem> | | 0.652 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C)C(C)C=N4)C3=C1</chem> | | 0.884 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N</chem> | | 1.44 |
| <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C)C=C42)=NN1</chem> | | 0.099 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C</chem> | | 0.016 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C)C=C4)C3=C1</chem> | | 0.077 |
| <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C</chem> | | 0.256 |
| <chem>CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.354 |
| <chem>CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 3.817 |
| <chem>CC1=CC(NC2=NC(C(N3CCC(N)C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 0.674 |
| <chem>CC1=CC(NC2=NC(C(N3CCC(O)C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 3.89 |
| <chem>CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.672 |
| <chem>CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.045 |
| <chem>CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 2.172 |
| <chem>CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.768 |
| <chem>C1C1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=</chem> | | 2.408 |
| <chem>CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.576 |
| <chem>CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.325 |
| <chem>CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 1.63 |
| <chem>CC1=CC(NC2=NC(C(NC3CNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.782 |
| <chem>C1C(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | | 0.016 |
| <chem>CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C)C=C42</chem> | | 0.218 |
| <chem>O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5O</chem> | | 0.026 |
| <chem>CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C)C=C42)</chem> | | 0.151 |
| <chem>O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5O</chem> | | 0.017 |

Preparing the dataset



| | B | C | D | E |
|----|---|----|-------|--------------------|
| 1 | SMILES | ki | pKi | |
| 2 | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@H]1C=CC=C1)C=C1</chem> | | 0.009 | =LOG10(C2/1000000) |
| 3 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N</chem> | | 0.026 | |
| 4 | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42</chem> | | 0.71 | |
| 5 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C</chem> | | 0.652 | |
| 6 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1</chem> | | 0.884 | |
| 7 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N</chem> | | 1.44 | |
| 8 | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42)=NN1</chem> | | 0.099 | |
| 9 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C</chem> | | 0.016 | |
| 10 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1</chem> | | 0.077 | |
| 11 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C</chem> | | 0.256 | |
| 12 | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.354 | |
| 13 | <chem>CC1=CC(NC2=NC(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 3.817 | |
| 14 | <chem>CC1=CC(NC2=NC(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 0.674 | |
| 15 | <chem>CC1=CC(NC2=NC(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 3.89 | |
| 16 | <chem>CC1=CC(NC2=NC(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.672 | |
| 17 | <chem>CC1=CC(NC2=NC(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.045 | |
| 18 | <chem>CC1=CC(NC2=NC(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 2.172 | |
| 19 | <chem>CC1=CC(NC2=NC(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.768 | |
| 20 | <chem>C1C1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=</chem> | | 2.408 | |
| 21 | <chem>CC1=CC(NC2=NC(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.576 | |
| 22 | <chem>CC1=CC(NC2=NC(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.325 | |
| 23 | <chem>CC1=CC(NC2=NC(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 1.63 | |
| 24 | <chem>CC1=CC(NC2=NC(N(C)C3CNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.782 | |
| 25 | <chem>C1C(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | | 0.016 | |
| 26 | <chem>CC1=CC(NC2=NC(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | | 0.218 | |
| 27 | <chem>O=C(N1CCN1C[C@@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | | 0.026 | |
| 28 | <chem>CC1=CC(NC2=NC(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | | 0.151 | |
| 29 | <chem>O=C(N1CCN1C[C@@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | | 0.017 | |

Preparing the dataset

The screenshot shows an Excel spreadsheet with the following data:

| | B | C | D | E |
|----|---|----|-------|------------|
| 1 | SMILES | ki | pKi | |
| 2 | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@</chem> | | 0.009 | 8.04575749 |
| 3 | <chem>O=C(N[C@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N</chem> | | 0.026 | |
| 4 | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42</chem> | | 0.71 | |
| 5 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C</chem> | | 0.652 | |
| 6 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1</chem> | | 0.884 | |
| 7 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N</chem> | | 1.44 | |
| 8 | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42)=NN1</chem> | | 0.099 | |
| 9 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C</chem> | | 0.016 | |
| 10 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1</chem> | | 0.077 | |
| 11 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=</chem> | | 0.256 | |
| 12 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.354 | |
| 13 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 3.817 | |
| 14 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 0.674 | |
| 15 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | | 3.89 | |
| 16 | <chem>CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.672 | |
| 17 | <chem>CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.045 | |
| 18 | <chem>CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 2.172 | |
| 19 | <chem>CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | | 0.768 | |
| 20 | <chem>C1C1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=</chem> | | 2.408 | |
| 21 | <chem>CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 0.576 | |
| 22 | <chem>CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.325 | |
| 23 | <chem>CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | | 1.63 | |
| 24 | <chem>CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | | 0.782 | |
| 25 | <chem>C1C(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | | 0.016 | |
| 26 | <chem>CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(C1)C=C42</chem> | | 0.218 | |
| 27 | <chem>O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | | 0.026 | |
| 28 | <chem>CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(C1)C=C42)</chem> | | 0.151 | |
| 29 | <chem>O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | | 0.017 | |

Preparing the dataset

| | B | C | D | E |
|----|--|-------|------------|---|
| 17 | CC1=CC(NC2=NC(C(N3CCNCCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | 7.34678749 | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | 5.66314018 | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | 6.11463878 | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | 5.61834352 | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | 6.23957752 | |
| 22 | CC1=CC(NC2=NC(C(N3CCNCCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | 6.48811664 | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | 5.7878124 | |
| 24 | CC1=CC(NC2=NC(C(N3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | 6.10679325 | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCCC3)=O)N=C2NC4=NNC(C5CC5) | 0.016 | 7.79588002 | |
| 26 | CC1=CC(NC2=NC(C(N3CCNCC[C@@H]3C)=O)=NC4=CC=C(C)C=C42 | 0.218 | 6.66154351 | |
| 27 | O=C(N1CCNCC[C@@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5O | 0.026 | 7.58502665 | |
| 28 | CC1=CC(NC2=NC(C(N3CCNCC[C@H]3C)=O)=NC4=CC=C(C)C=C42) | 0.151 | 6.82102305 | |
| 29 | O=C(N1CCNCC[C@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5CC | 0.017 | 7.76955108 | |
| 30 | CC1=CC(NC2=NC(C(N3CCNCC[C@H](C)C3)=O)=NC4=CC=C(C)C=C42 | 0.051 | 7.29242982 | |
| 31 | O=C(N1CCNCC[C@H](C)C1)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5C | 0.009 | 8.04575749 | |
| 32 | CC1=CC(NC2=NC(C(N3CCNCC[C@@H](C)C3)=O)=NC4=CC=C(C)C=C | 0.306 | 6.51427857 | |
| 33 | O=C(N1CCNCC[C@@H](C)C1)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5 | 0.028 | 7.55284197 | |
| 34 | CC1=CC(NC2=NC(C(N3CC(C)NC(C)C3)=O)=NC4=CC=C(C)C=C42) | 0.119 | 6.92445304 | |
| 35 | O=C(N1CC(C)NC(C)C1)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5CC | 0.114 | 6.94309515 | |
| 36 | O=C(N1CCNCC[C@@H](C)C1)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5 | 0.006 | 8.22184875 | |
| 37 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=CC=C3C(NC4=NNC(C5CC | 0.017 | 7.76955108 | |
| 38 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5 | 0.007 | 8.15490196 | |
| 39 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(F)C=C3C(NC4=NNC(C5 | 0.016 | 7.79588002 | |
| 40 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(Br)C=C3C(NC4=NNC(C5 | 0.011 | 7.95860731 | |
| 41 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(OC)C=C3C(NC4=NNC(C | 0.036 | 7.4436975 | |
| 42 | | | | |
| 43 | | | | |
| 44 | | | | |
| 45 | | | | |

Preparing the dataset

The screenshot shows the LibreOffice Calc interface with a spreadsheet containing SMILES strings in column B and numerical values in columns C and D. A 'Paste Special' dialog box is open, allowing for various paste options. A blue arrow points to the 'ki' header in column C.

| | B | C | D | E |
|----|--|-------|-------------|---|
| 1 | SMILES | ki | ki | |
| 2 | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 0.009 | 8.04575749; | |
| 3 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 0.026 | 7.58502665; | |
| 4 | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C)C=C42 | 0.71 | 6.14874165; | |
| 5 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C2=C | 0.652 | 6.1857524; | |
| 6 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C2=C | 0.884 | 6.05354773; | |
| 7 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C2=C | 1.44 | 5.84163751; | |
| 8 | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C)C=C4)C=C1 | 0.099 | 7.00436481; | |
| 9 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NN(C5CC5)C=N4)C2=C | 0.016 | 7.79588002; | |
| 10 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NN(C5CC5)C=N4)C2=C | 0.077 | 7.11350927; | |
| 11 | C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NN(C5CC5)C=N4)C2=C | 0.256 | 6.59176003; | |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.354 | 6.45099674; | |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 3.817 | 5.41827784; | |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)C)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.674 | 6.1713401; | |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)C)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 3.89 | 5.4100504; | |
| 16 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.672 | 6.17263073; | |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.045 | 7.34678749; | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 2.172 | 5.66314018; | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.768 | 6.11463878; | |
| 20 | C1C1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4C)C=C4)C=C1 | 2.408 | 5.61834352; | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.576 | 6.23957752; | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)C)C=C1 | 0.325 | 6.48811664; | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)C)C=C1 | 1.63 | 5.7878124; | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)C)C=C1 | 0.782 | 6.10679325; | |
| 25 | C1C(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)C=C4 | 0.016 | 7.79588002; | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.218 | 6.66154351; | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5O)C=C4)C=C2 | 0.026 | 7.58502665; | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(C)C=C4)C=C1 | 0.151 | 6.82102305; | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(C)C=C3C(NC4=NNC(C5O)C=C4)C=C2 | 0.017 | 7.76955108; | |

Preparing the dataset

The screenshot shows a spreadsheet with the following data:

| | B | C | D | E |
|----|--|------------------|---|---|
| 1 | SMILES | pKi | | |
| 2 | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@H]1C=CC=CC=C1)CN(C)C(C)C</chem> | 8.04575749056068 | | |
| 3 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)C(C)C)CC3=C2NN=C3N</chem> | 7.58502665202918 | | |
| 4 | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42</chem> | 6.14874165128092 | | |
| 5 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C</chem> | 6.18575240426808 | | |
| 6 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C)C(C)C)C=N4)C3=C1</chem> | 6.05354773498693 | | |
| 7 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N</chem> | 5.84163750790475 | | |
| 8 | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C1)C=C42)=NN1</chem> | 7.00436480540245 | | |
| 9 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C</chem> | 7.79588001734408 | | |
| 10 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1</chem> | 7.11350927482752 | | |
| 11 | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C</chem> | 6.59176003468815 | | |
| 12 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | 6.45099673797421 | | |
| 13 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | 5.4182778400509 | | |
| 14 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | 6.17134010346468 | | |
| 15 | <chem>CC1=CC(NC2=NC(C(N3CCC(O)C)CC3)=O)=NC4=CC=C(C=C42)Cl)=</chem> | 5.41005039867429 | | |
| 16 | <chem>CC1=CC(NC2=NC(C(N3CCC(N)C)CC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | 6.17263072694618 | | |
| 17 | <chem>CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | 7.34678748622466 | | |
| 18 | <chem>CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | 5.66314017908319 | | |
| 19 | <chem>CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN</chem> | 6.11463877996849 | | |
| 20 | <chem>C1C1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=</chem> | 5.61834351741421 | | |
| 21 | <chem>CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | 6.23957751657679 | | |
| 22 | <chem>CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | 6.48811663902113 | | |
| 23 | <chem>CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N</chem> | 5.78781239559604 | | |
| 24 | <chem>CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1</chem> | 6.10679324694015 | | |
| 25 | <chem>C1C(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)</chem> | 7.79588001734408 | | |
| 26 | <chem>CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C1)C=C42</chem> | 6.6615435063954 | | |
| 27 | <chem>O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | 7.58502665202918 | | |
| 28 | <chem>CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(C1)C=C42)</chem> | 6.82102305270683 | | |
| 29 | <chem>O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(C1)C=C3C(NC4=NNC(C5O</chem> | 7.76055107862173 | | |

Preparing the dataset

The screenshot shows a web browser window with multiple tabs. The active tab is 'Online Chemical', with the URL <https://ochem.eu/batchupload30/show.do>. The page header includes the 'Online chemical database with modeling environment' logo and a welcome message for 'Prof. Ragno!'. Below the header is a navigation menu with 'Home', 'Database', and 'Models' options. The main content area is titled 'Batch Upload 3.0 - File preview and column remapping' and contains a table of data from a CSV file named 'jm7b01342_si_002'. The table has three columns: 'RECORDID', 'SMILES', and 'pKI'. The 'RECORDID' and 'SMILES' columns are highlighted in green, indicating they are recognized. The 'pKI' column is also highlighted in green. Below the table, there is a message: 'The ARTICLE column is missing, the stub unpublished article will be assigned by default'. Further instructions explain that green titles indicate recognized columns, red titles indicate errors, and irrelevant columns can be left red to be ignored.

Batch Upload 3.0 - File preview and column remapping
Preview your data, select the sheet and the columns you would like to upload

jm7b01342_si_002

| <input checked="" type="checkbox"/> RECORDID | <input checked="" type="checkbox"/> SMILES | <input checked="" type="checkbox"/> pKI |
|--|---|---|
| staurosporine | <chem>O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C...</chem> | 8.045757490560675 |
| PF3758309 | <chem>O=C(N[C@@H](C1=CC=CC=C1)CN(C)N(C2(C...</chem> | 7.585026652029182 |
| 10a | <chem>CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=C...</chem> | 6.1487416512809245 |
| 10b | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 6.185752404268079 |
| 10c | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 6.053547734986927 |
| 10d | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(...</chem> | 5.841637507904751 |
| 11a | <chem>CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(C...</chem> | 7.00436480540245 |
| 11b | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 7.795880017344075 |
| 11c | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 7.113509274827518 |
| 11d | <chem>C1C1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC...</chem> | 6.59176003468815 |

The ARTICLE column is missing, the stub unpublished article will be assigned by default

Green titles indicate recognized columns, red titles indicate errors. Please click on the red columns and select whether the column indicates a property, condition or another column type like name, value or molecule, then select the matching entity and confirm your selection by clicking on the green button on the left.
If you have irrelevant columns in your sheet, you can leave them red and they will be ignored in the further process. If you

jm7b01342_si_0...csv Show all

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/batchupload30/show.do>. The page title is "Online chemical database with modeling environment" and the version is v.3.0.96.1. The user is logged in as Prof. Ragno. The main content area is titled "Batch Upload 3.0 - Entity remapping" and includes a description: "Review and remap the properties, conditions, units, articles and baskets involved in the data upload".

Database entities remapping

Property: **pKi**

Values
Unit: **-log(mol/L)**, min value: 5.410050398674292, max value: 8.221848749616356

Article: **unpublished**

Molecule set: **default**

submit

Buttons: **Cancel Batch Upload**, **Download Excel file**

File name: **jm7b01342_si_0...csv** | **Show all**

Preparing the dataset

Autenticazione | Home | Sapienz. | HomePage - Citi. | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/batchupload30/show.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendeley | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home | Database | Models | A+ a- Privacy statement

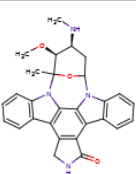
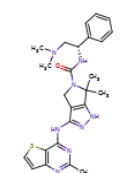
Batch upload 3.0 - records preview
Preview the records you are about to upload, select the desired actions

Batch upload preview browser

Summary:
All rows in the sheet Count: **40**
Status: valid, Count: **40**

Filter by row number: and row type: **all** Batch operations

1 - 10 of 40 10 items on page 1 of 4

| | |
|---|--|
| Row 1 <input checked="" type="radio"/> Save <input type="radio"/> Skip |  pK_i = 8.045757490560675 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls... N: AUTO_1 MoleculeID: M4402773 RecordID: R-1 rino.ragno <input checked="" type="checkbox"/> Only visible to rino.ra |
| Row 2 <input checked="" type="radio"/> Save <input type="radio"/> Skip |  pK_i = 7.585026652029182 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls... N: AUTO_2 MoleculeID: M95419909 RecordID: R-2 rino.ragno <input checked="" type="checkbox"/> Only visible to rino.ra |

jm7b01342_si_0...csv Show all ×

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

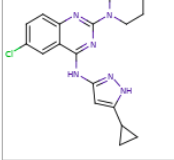
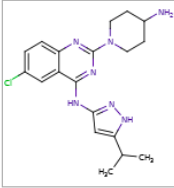
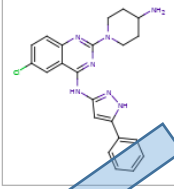
https://ochem.eu/batchupload30/show.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

| | | |
|--|---|--|
| Row 8 <input checked="" type="radio"/> Save <input type="radio"/> Skip |  | pKi = 7.793880017344073 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls... N: AUTO_8 MoleculeID: M97153552 RecordID: R-8 rino.ragno Only visible to rino.ra |
| Row 9 <input checked="" type="radio"/> Save <input type="radio"/> Skip |  | pKi = 7.113509274827518 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls... N: AUTO_9 MoleculeID: M97153553 RecordID: R-9 rino.ragno Only visible to rino.ra |
| Row 10 <input checked="" type="radio"/> Save <input type="radio"/> Skip |  | pKi = 6.59176003468815 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls... N: AUTO_10 MoleculeID: M97153554 RecordID: R-10 rino.ragno Only visible to rino.ra |

1 - 10 of 40
Proceed with upload

10 items on page 1 of 4

Cancel Batch Upload Download Excel file

jm7b01342_si_0...csv Show all x

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/batchupload30/show.do>. The page header includes the logo for "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". Navigation tabs for "Home", "Database", and "Models" are visible. The main content area displays a notification: "Batch upload 3.0 - finished" with the subtext "Your upload has been finished". Below this, a section titled "Batch upload results" contains the text: "Batch upload is finished. You can download the detailed upload report." A blue arrow points to the link "detailed upload report". A summary table is shown below:

| Summary: | |
|----------------------------|-----------|
| All rows in the sheet | Count: 40 |
| Status: valid, saved_valid | Count: 40 |

At the bottom right of the main content area, there are two buttons: "New Batch Upload" and "Download Excel file". The browser's taskbar at the bottom shows a file named "jm7b01342_si_0...csv" and a "Show all" button.

Preparing the dataset

The screenshot shows the 'Online chemical database with modeling environment' interface. The browser address bar displays 'https://ochem.eu/batchupload30/show.do'. The page header includes the site logo, the text 'Online chemical database with modeling environment', and a user greeting: 'Welcome, Dear Prof.Ragno! My account Logout'. A navigation menu at the top contains 'Home', 'Database', and 'Models'. The 'Models' dropdown menu is open, listing several options: 'Create model', 'Apply a model', 'Create multiple models', 'Create multiple models with conditions (experimental)', 'Open predictor', 'Upload a linear model', 'Upload a stub model', 'View pending tasks', 'View published tasks', 'SetCompare utility', 'MolOptimiser', 'Calculate descriptors', and 'Descriptors storage'. A blue arrow points to the 'Create model' option. Below the menu, there are buttons for 'New Batch Upload' and 'Download Excel file'. The footer shows the URL 'https://ochem.eu/modelconfigurator/choose.do' and a file upload area with a file named 'jm7b01342_si_0...csv'.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page title is "Online chemical database with modeling environment" and the version is "v.3.0.96.1". The user is logged in as "Prof. Ragno!". The navigation menu includes "Home", "Database", and "Models". The main content area is titled "Create a model" and contains the following sections:

- Select the training and validation sets:**
 - Training set (required): [...]
 - [Add a validation set](#)
- Choose the learning method:**
 - Suggested modeling methods:*
 - ASNN: ASsociative Neural Networks
 - CHEMCHAINER: Chainer Chemistry models (GPU)
 - CNF - Convolutional Neural Network Fingerprint (GPU)
 - Consensus model (based on models developed for the same set)
 - DEEPCHEM: several methods from DeepChem (GPU)
 - DNN: Deep Neural Network (GPU)
 - EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
 - FSMLR: Fast Stagewise Multiple Linear Regression
 - KNN: k - Nearest Neighbors
 - Library model (A local bias correction model based on another ASNN model)
 - LibSVM: grid-search parameter optimisation
 - LSSVMG: Least Squares Support Vector Machine (GPU)
 - MLR: Multiple Linear Regression
 - PLS: Partial Least Squares
 - RFR: Random Forest regression and classification
 - WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging
 - WEKA-RF: Random Forest, only classification
 - XGBoost: Scalable and Flexible Gradient Boosting

At the bottom of the page, there is a file upload area with the filename "jm7b01342_si_0....csv" and a "Show all" button.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page title is "Online chemical database with modeling environment" and the version is "v.3.0.96.1". The user is logged in as "Prof. Ragno!". The navigation menu includes "Home", "Database", and "Models". The "Models" dropdown is open, showing a list of machine learning models:

- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid-search parameter optimisation
- LSSVMG: Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression
- PLS: Partial Least Squares
- RFR: Random Forest regression and classification
- WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging
- WEKA-RF: Random Forest, only classification
- XGBoost: Scalable and Flexible Gradient Boosting

Below the list, there is a section "Methods under development:" and a "Model validation" section with the following options:

Validation method:

Number of folds:

Stratified cross-validation (classification only)

Consider each record as a molecule. [?](#)

You can create a model from template: [import an XML model template](#) or [use another model as a template](#)

At the bottom left, there is a "Next>>" button, which is highlighted by a blue arrow. At the bottom right, there is a "Show all" button.

Preparing the dataset

Autenticazione x Home | Sapienz: x HomePage - Citt: x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

Online chemical database with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Create a model ⓘ
Select the training and validation sets, the machine learning method and the validation protocol

Select the training and validation sets:

Training set (required): []
Add a validation set
Click to change

Choose the learning method:

Suggested modeling methods:

- ASNN: ASsociative Neural Networks
- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid-search parameter optimisation
- LSSVMG: Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression
- PLS: Partial Least Squares
- RFR: Random Forest regression and classification
- WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging
- WEKA-RF: Random Forest, only classification

javascript:void(0)

jm7b01342_sl_0....csv Show all x

Preparing the dataset

The screenshot shows the Online Chemical Database interface. The browser address bar displays `https://ochem.eu/modelconfigurator/choose.do`. The page header includes the logo and text "Online chemical database with modeling environment" and a version number "v.3.0.96.1". A navigation menu contains "Home", "Database", and "Models". A user greeting "Welcome, Dear Prof.Ragno!" and links for "My account" and "Logout" are visible. Below the navigation, there are tabs for "Model Builder X" and "Select compound set X".

The main content area is titled "Basket browser" and includes the instruction "Browse, Compare or Join molecule sets". A filter section contains a text input field, a "[Create new]" button, and a "Show public sets" checkbox. Below this, a table lists baskets:

| 1 - 2 of 2 | |
|--------------------------------------|------------|
| Selected records | 0 records |
| jm7b01342_si_002.xls | 40 records |

A blue arrow points to the selected record "jm7b01342_si_002.xls". A tooltip below the table says "Click to select this basket". At the bottom of the page, a file download bar shows "jm7b01342_si_0...csv" and a "Show all" button.

Preparing the dataset

Autenticazione x Home | Sapienz: x HomePage - Citt: x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home Database Models A+ a- Privacy statement

Create a model ⓘ
Select the training and validation sets, the machine learning method and the validation protocol

Select the training and validation sets:

Training set (required): [jm7b01342_si_002.xls](#) [details]
[Add a validation set](#)

The model will predict this property:
pKi using unit:

Choose the learning method: ⓘ

Suggested modeling methods:

- ASNN: ASsociative Neural Networks
- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid search parameter optimisation
- LSVMG - Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression
- PLS: Partial Least Squares
- RFR: Random Forest regression and classification
- WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging

jm7b01342_si_0...csv ^ Show all x

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for the Online Chemical Database and the text "with modeling environment". A welcome message reads "Welcome, Dear Prof.Ragno!" with links for "My account" and "Logout". The main content area is titled "Model configuration" and lists various machine learning models. A blue arrow points to the "Next >" button at the bottom left of the configuration area.

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home Database Models A+ a- Privacy statement

- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid-search parameter optimisation
- LSSVMG: Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression
- PLS: Partial Least Squares
- RFR: Random Forest regression and classification
- WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging
- WEKA-RF: Random Forest, only classification
- XGBoost: Scalable and Flexible Gradient Boosting

Methods under development:

Model validation

Validation method:

Number of folds:

- Stratified cross-validation (classification only)
- Consider each record as a molecule. [?](#)

You can create a model from template: [import an XML model template](#) or [use another model as a template](#)

jm7b01342_sl_0....csv

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for 'Online chemical database with modeling environment' and a version number 'v.3.0.96.1'. A navigation bar contains 'Home', 'Database', and 'Models' menus. The main content area is titled 'Model creator' and contains the instruction 'Select model template and training set'. Below this, a section titled 'Select the preferred data preprocessing options' is shown, with a sub-section 'Preprocessing of molecules (Chemaxon)'. Four checkboxes are present, all of which are checked: 'Standardization', 'Neutralize', 'Remove salts', and 'Clean structure'. At the bottom of this section are two buttons: '<<Back' and 'Next>>'. A blue arrow points to the 'Next>>' button. The browser's address bar and tabs are visible at the top, and a file upload area with a 'Show all' button is at the bottom.

Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP »

Online chemical database
with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Model creator
Select model template and training set

Select the preferred data preprocessing options

Preprocessing of molecules (Chemaxon)

- Standardization
- Neutralize
- Remove salts
- Clean structure

<<Back Next>>

jm7b01342_si_0...csv Show all x

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Model creator
Select model template and training set

Select the molecular descriptors

Recommended descriptor types

- E-state
 - E-State types:
 - Atom indices
 - Bonds indices
 - Atom counts
 - Bonds counts
- Aromatize structures: Chemaxon Basic
- ALogPS (2)
- GSFragment (1138)
- CDK 2.0 descriptors (256/3D)
- Dragon v. 7 (5270/3D)
- alvaDesc v.1.0.14 (5305/3D)
- ISIDA fragments
- 'Inductive' descriptors (54/3D)
- MERA descriptors (529/3D)
- MERSY descriptors (42/3D)
- Chemaxon descriptors (499/3D)
- QNPR
- Spectrophores (144/3D)
- Structural alerts (ToxAlerts)

Predictions by OCHEM's featured models

- Ames levenberg
- Toxicity against T. Pyriformis
- ALogPS 3.0
- CYP1A2 Estate+ALogPS
- CYP2C9 Estate+ALogPS
- CYP2C19 Estate+ALogPS
- CYP2D6 Estate+ALogPS
- CYP3A4 Estate+ALogPS
- Pyrolysis point prediction (best Estate)
- Melting Point prediction (best Estate)
- Water solubility model based on logP and Melting Point
- ALOGPS 2.1 logP
- ALOGPS 2.1 logS
- Outputs of other OCHEM models

Obsolete/Additional descriptor types

- CDK 1.4.11 descriptors (256/3D)
- OESState
- Dragon v. 5.4 (1644/3D)
- Dragon v. 5.5 (3224/3D)

jm7b01342_sl_0....csv Show all

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Model creator
Select model template and training set

Select the molecular descriptors

Recommended descriptor types

- E-state
- ALogPS (2)
- GSFfragment (1138)
- CDK 2.0 descriptors (256/3D)
- Dragon v. 7 (5270/3D)
- alvaDesc v.1.0.14 (5305/3D)
- ISIDA fragments
- 'Inductive' descriptors (54/3D)
- MERA descriptors (529/3D)
- MERSY descriptors (42/3D)
- Chemaxon descriptors (499/3D)
- QNPR
- Spectrophores (144/3D)
- Structural alerts (ToxAlerts)

Predictions by OCHEM's featured models

- Ames levenberg
- Toxicity against T. Pyriformis
- ALogPS 3.0
- CYP1A2 Estate+ALogPS
- CYP2C9 Estate+ALogPS
- CYP2C19 Estate+ALogPS
- CYP2D6 Estate+ALogPS
- CYP3A4 Estate+ALogPS
- Pyrolysis point prediction (best Estate)
- Melting Point prediction (best Estate)
- Water solubility model based on logP and Melting Point
- ALOGPS 2.1 logP
- ALOGPS 2.1 logS

Outputs of other OCHEM models

Obsolete/Additional descriptor types

- CDK 1.4.11 descriptors (256/3D)
- OESate
- Dragon v. 5.4 (1644/3D)
- Dragon v. 5.5 (3224/3D)

jm7b01342_sl_0...csv Show all

Preparing the dataset

Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

Online chemical database with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Select model template and training set

Select the molecular descriptors

Recommended descriptor types

- E-state
- ALogPS (2)
- GSF fragment (1138)
- CDK 2.0 descriptors (256/3D)
- Dragon v. 7 (5270/3D)

[select all] [select none]

| | |
|---|--|
| <input checked="" type="checkbox"/> Constitutional descriptors (47) | <input checked="" type="checkbox"/> Ring descriptors (32) |
| <input checked="" type="checkbox"/> Topological indices (75) | <input checked="" type="checkbox"/> Walk and path counts (46) |
| <input checked="" type="checkbox"/> Connectivity indices (37) | <input checked="" type="checkbox"/> Information indices (50) |
| <input checked="" type="checkbox"/> 2D matrix-based descriptors (607) | <input checked="" type="checkbox"/> 2D autocorrelations (213) |
| <input checked="" type="checkbox"/> Burden eigenvalues (96) | <input checked="" type="checkbox"/> P_VSA-like descriptors (55) |
| <input checked="" type="checkbox"/> ETA indices (23) | <input checked="" type="checkbox"/> Edge adjacency indices (324) |
| <input checked="" type="checkbox"/> Geometrical descriptors (3D, 38) | <input checked="" type="checkbox"/> 3D matrix-based descriptors (3D, 99) |
| <input checked="" type="checkbox"/> 3D autocorrelations (3D, 80) | <input checked="" type="checkbox"/> RDF descriptors (3D, 210) |
| <input checked="" type="checkbox"/> 3D-MoRSE descriptors (3D, 224) | <input checked="" type="checkbox"/> WHIM descriptors (3D, 114) |
| <input checked="" type="checkbox"/> GETAWAY descriptors (3D, 273) | <input checked="" type="checkbox"/> Randic molecular profiles (3D, 41) |
| <input checked="" type="checkbox"/> Functional group counts (3D, 154) | <input checked="" type="checkbox"/> Atom-centred fragments (115) |
| <input checked="" type="checkbox"/> Atom-type E-state indices (172) | <input checked="" type="checkbox"/> CATS 2D (150) |
| <input checked="" type="checkbox"/> 2D Atom Pairs (1596) | <input checked="" type="checkbox"/> 3D Atom Pairs (3D, 36) |
| <input checked="" type="checkbox"/> Charge descriptors (3D, 15) | <input checked="" type="checkbox"/> Molecular properties (20) |
| <input checked="" type="checkbox"/> Drug-like indices (28) | <input checked="" type="checkbox"/> CATS 3D (3D, 300) |

alvaDesc v.1.0.14 (5305/3D)

ISIDA fragments

Predictions by OCHEM's featured models

- Ames levenberg
- Toxicity against T. Pyriformis
- ALogPS 3.0
- CYP1A2 Estate+ALogPS
- CYP2C9 Estate+ALogPS
- CYP2C19 Estate+ALogPS
- CYP2D6 Estate+ALogPS
- CYP3A4 Estate+ALogPS
- Pyrolysis point prediction (best Estate)
- Melting Point prediction (best Estate)
- Water solubility model based on logP and Melting Point
- ALOGPS 2.1 logP
- ALOGPS 2.1 logS

Outputs of other OCHEM models

Obsolete/Additional descriptor types

- CDK 1.4.11 descriptors (256/3D)
- OESState
- Dragon v. 5.4 (1644/3D)
- Dragon v. 5.5 (3224/3D)
- Dragon v. 6 (4885/3D)
- MOPAC 7.1 descriptors (25/3D)

jm7b01342_sl_0....csv Show all x

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

v.3.0.96.1

Home Database Models A+ a- Privacy statement

- Chemaxon descriptors (430/3D)
- QNPR
- Spectrophores (144/3D)
- Structural alerts (ToxAlerts)

Special descriptors (scaffolds, fingerprints):

- Chemaxon Scaffolds
- Silicos-It Scaffolds
- ECFP Fingerprints *Not supported by your installation*
- MolPrint Fingerprints

Under development: can change anytime and backward compatibility is not guaranteed. Use at your own risk!

- JPligP
- CDK 2.2 descriptors (256/3D)
- RDKit descriptors (3D)
- RDKit additional descriptors (3D)
- MORDRED descriptors (1826/3D)
- CDDD
- MOPAC2016 descriptors (35/3D)
- SIRMS
- PyDescriptor descriptors (16251/3D)
- External descriptors

- Allow Merging Descriptors (experimental)

<<Back Next>>

jm7b01342_si_0....csv Show all x

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for the Online chemical database with modeling environment (v.3.0.96.1) and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". The navigation menu contains "Home", "Database", and "Models". The main content area is titled "Model creator" and "Select model template and training set". Below this, there is a section "Select a tool to optimize molecule structures" with the following options:

- No optimisation
- Optimise with Corina
- Optimise with OpenBabel
- Optimise with OBGEN (part of OpenBabel distribution)
- Optimise with BALLOON

At the bottom of the selection area, there are two buttons: "<<Back" and "Next>>". A blue arrow points to the "Next>>" button. The browser's address bar shows several tabs, including "Autenticazione", "Home | Sapienz...", "HomePage - Citt...", "Online Chemical", "Journal of Medic...", and "Structure-Based...". The browser's taskbar at the bottom shows a file named "jm7b01342_si_0...csv" and a "Show all" button.

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP »

Online chemical database
with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

Model creator
Select model template and training set

Select filters of descriptors

- Eliminate descriptors with less than unique values
- Delete descriptors that have absolute values larger than
- Delete descriptors that have variance smaller than
- Group descriptors, that have pair-wise correlations Pearson's correlation coefficient R larger than
- Use Unsupervised Forward Selection to delete variables using the above value of multiple correlation coefficient R
- After filtering, I want to select necessary descriptors myself (*advanced*)

Normalisation parameters

Descriptors normalization

Values normalization


jm7b01342_si_0...csv ^ Show all x

Preparing the dataset

Autenticazione x Home | Sapienz. x HomePage - Citt. x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

 **Online chemical database**
with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

Model creator
Select model template and training set

Select filters of descriptors

- Eliminate descriptors with less than unique values
- Delete descriptors that have absolute values larger than
- Delete descriptors that have variance smaller than
- Group descriptors, that have pair-wise correlations Pearson's correlation coefficient R larger than
- Use Unsupervised Forward Selection to delete variables using the above value of multiple correlation coefficient R
- After filtering, I want to select necessary descriptors myself (advanced)

Normalisation parameters

Descriptors normalization ▾
Values normalization ▾

<<Back

jm7b01342_si_0...csv ^ x

Preparing the dataset

The screenshot shows a web browser window titled "Online Chemical Modeling Environment - Google Chrome". The address bar displays the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for "Online chemical database with modeling environment" (version v.3.0.96.1) and a user greeting: "Welcome, Dear Prof.Ragno! My account Logout". Navigation links for "Home", "Database", and "Models" are present, along with a "Privacy statement" link.

The main content area is titled "Model creator" with the subtitle "Select model template and training set". Below this, the "Configure PLS method" section is active. It contains the following elements:

- "Number of latent variables:" followed by a text input field containing the value "0".
- A checked checkbox labeled "Optimize the number of latent variables automatically".
- An unchecked checkbox labeled "Limit predicted values to the training set range".
- Navigation buttons: "<<Back" and "Next>>".

A blue arrow points from the "Next>>" button towards the "Optimize the number of latent variables automatically" checkbox. At the bottom of the browser window, a file upload area shows a file named "jm7b01342_sl_0...csv" with a "Show all" button.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". A navigation menu contains "Home", "Database", and "Models".

The main content area is titled "Model creator" with the instruction "Select model template and training set". Below this, a section titled "Start calculation of the model" contains the text "Now we are ready to start calculation. Please provide the name for your model:" followed by a text input field containing "pKi_PLS_[Dragon7 (blocks: 1-30)] - 336922". A blue arrow points to this input field.

Below the input field, there is a checked checkbox for "Save models" and a "Task priority:" section with three radio button options: "High priority (please, use for fast tasks only)", "Normal priority" (which is selected), and "Low priority (for long tasks)". A blue arrow points to the "Normal priority" option.

At the bottom of the form, there are three buttons: "<<Back", "Start calculation>>" (highlighted in red), and "Discard". A blue arrow points to the "Start calculation>>" button.

The browser's taskbar at the bottom shows a file named "jm7b01342_sl_0...csv" and a "Show all" button.

Preparing the dataset

The screenshot shows a web browser window with multiple tabs. The active tab is 'Online Chemical' at the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the 'Online chemical database with modeling environment' logo and a version number 'v.3.0.96.1'. A user is logged in as 'Prof.Ragno!' with links for 'My account' and 'Logout'. A navigation menu contains 'Home', 'Database', and 'Models'. The main content area is titled 'Run model builder' and displays a progress indicator: 'Finished posting ... - Processing task Corina - Waiting for a free server -- 09:16'. Below this, there are links for '[cancel]' and '[fetch result later]'. At the bottom of the content area, there are '<<Back' and 'Next>>' buttons. A file manager at the bottom shows a file named 'jm7b01342_sl_0...csv' with a 'Show all' button.

Preparing the dataset

Autenticazione G x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Model creator
Select model template and training set

Save the model
Please enter your model's name:

Overview

Model name: pKi_PLS_[Dragon7 (blocks: 1-30)] - 336922 [rename] [Dragon7 (blocks: 1-30)]
Temporal Public ID: 37774824 - use this link to share the model
Correl. limit: 0.0 Variance threshold: 0.01, Maximum value: 999999,

Predicted property: **pKi** modeled in -log(mol/L)
Training method: PLS

| Data Set | # | R2 | q2 | RMSE | MAE |
|--|------------|-----------|-----------|-----------|-------------|
| Training set: jm7b01342_si_002.xls | 40 records | 0.4 ± 0.2 | 0.3 ± 0.3 | 0.7 ± 0.2 | 0.46 ± 0.08 |

scale X: STANDARDIZE0 latent variables 3D by Corina
5-fold cross-validation
3008 pre-filtered descriptors

scale X: STANDARDIZE2 latent variables
Y = -1.03 + 7.3E-6**MW* - 0.00143**AMW* + 9.56E-5**Sv* + 7.81E-5**Se* + 9.37E-5**Sp* + 7.9E-5**Si* - 0.0664**Mv* - 0.108**Me* - 0.0773**Mp* + 0.0959**Mi* + 0.0383**GD* + 8.21E-5**nAT* + 7.64E-5**nSK* - 0.00167**nTA* + 1.01E-4**nBT* + 1.66E-4**nBO* - 4.82E-4**nBM* + 3.18E-5**SCBO* + 0.00108**nBN* + 0.0675**nRBF* + 1.51E-4**nDB* - 4.65E-4**nAB* + 1.93E-4**nH* + 9.4E-5**nC* + 0.00145**nN* - 0.00104**nO* - 0.00236**nCL* - 0.00246**nHM* - 3.41E-4**nHet* - 8.88E-4**nX* + 3.52E-4**nH%* - 1.59E-4**nC%* + 1.81E-4**nN%* - 6.33E-4**nO%* - 4.98E-4**nX%* + 0.00108**nCsp3* -

LigandScout_4....dmg ^ LigandScout_4....exe ^ LigandScout....tar.gz ^ jm7b01342_si_0....csv ^ Show all x

Preparing the dataset

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home Database Models

Model c
Select m

Save

Pleas
MatchedPairs
pKi_PLS_[Dragon7 (bloc

Baskets

Tags

Set area of interest...

User-related changes

Batch data upload

Trash

Dragon7 (blocks: 1-30)) - 336922 [rename]

Correl. limit: 0.0 Variance threshold: 0.01,
Maximum value: 999999,

scale X: STANDARDIZE0 latent variables 3D by
Corina
5-fold cross-validation
-
3008 pre-filtered descriptors

scale X: STANDARDIZE2 latent variables
Y = -1.03 + 7.3E-6**MW* - 0.00143**AMW* +
9.56E-5**Sv* + 7.81E-5**Se* + 9.37E-5**Sp* + 7.9E-
5**Si* - 0.0664**Mv* - 0.108**Me* - 0.0773**Mp* +
0.0959**Mi* + 0.0383**GD* + 8.21E-5**nAT* +
7.64E-5**nSK* - 0.00167**nTA* + 1.01E-4**nBT* +
1.66E-4**nBO* - 4.82E-4**nBM* + 3.18E-5**SCBO*
+ 0.00108**RBN* + 0.0675**RBF* + 1.51E-4**nDB* -
4.65E-4**nAB* + 1.93E-4**nH* + 9.4E-5**nC* +
0.00145**nN* - 0.00104**nO* - 0.00236**nCL* -
0.00246**nHM* - 3.41E-4**nHet* - 8.88E-4**nX* +
3.52E-4**H%* - 1.59E-4**C%* + 1.81E-4**N%* -
6.33E-4**O%* - 4.98E-4**X%* + 0.00108**nCsn3* -

| Data Set | # | R2 | q2 | RMSE | MAE |
|--|------------|-----------|-----------|-----------|-------------|
| Training set: jm7b01342_si_002.xls | 40 records | 0.4 ± 0.2 | 0.3 ± 0.3 | 0.7 ± 0.2 | 0.46 ± 0.08 |

8.0

7.5

7.0

https://ochem.eu/basket/show.do

LigandScout_4....dmg ^

LigandScout_4....exe ^

LigandScout....tar.gz ^

jm7b01342_si_0....csv ^

Show all X

Preparing the dataset

The screenshot shows the Online Chemical Database (ochem.eu) interface. The browser address bar displays <https://ochem.eu/basket/show.do>. The page header includes the logo and text "Online chemical database with modeling environment" and a version number "v.3.0.96.1". A navigation menu contains "Home", "Database", and "Models". A user greeting "Welcome, Dear Prof.Ragno!" and links for "My account" and "Logout" are present. The main content area is titled "Basket browser" and includes a filter by name field, a "Create new" button, and a "Show public sets" checkbox. Below this, a table lists records:

| Record Name | Records | Pending Models |
|----------------------|------------|------------------|
| Selected records | 0 records | |
| jm7b01342_si_002.xls | 40 records | 1 pending models |

A blue arrow points to the "jm7b01342_si_002.xls" record, and a tooltip "Open basket profile" is visible over the record's icon. The bottom of the browser shows a taskbar with files like "LigandScout_4....dmg", "LigandScout_4.....exe", "LigandScout.....tar.gz", and "jm7b01342_si_0....csv".

Preparing the dataset

Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/basket/show.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP »

Online chemical database with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Molecule sets X Edit basket X

Basket editor

Add new basket or edit existing basket

Name:
(min. 2 characters)

Description (optional):

Excluded implicit records (under development):

Actions

- Create a copy of this basket
- Create a primary records basket
- Add or delete particular records
- Discretize the numerical values
- Models summary for 1 models
- Split the basket into two sets
- Transform the basket using OScript
- Export this basket into Excel, CSV or SDF

Statistics of the basket

| Properties | Records | Unique compounds | |
|------------|------------|------------------|-----------|
| pKi | 40 records | 40 compounds | Show MMPs |

https://ochem.eu/epbrowser/show.do?basket-select=189582&property=48967

LigandScout_4....dmg ^ LigandScout_4_....exe ^ LigandScout_....tar.gz ^ jm7b01342_si_0....csv ^ Show all x

Preparing the dataset

The screenshot displays the 'Online chemical database with modeling environment' interface. The browser address bar shows the URL <https://ochem.eu/basket/show.do>. The page title is 'Compounds properties browser' with a subtitle 'Search for numerical compounds properties linked to scientific articles'. The user is logged in as 'Prof. Ragno'.

The interface includes a navigation menu with 'Home', 'Database', and 'Models'. Below this, there are tabs for 'Molecule sets X', 'Edit basket X', and 'records X'. The main content area is divided into a 'FILTERS' sidebar on the left and a main list of records on the right.

Filters:

- SOURCE:** Article/Source [select]
- PROPERTY:** Activity/Property [select], with 'pKi' highlighted in yellow. There is an option to 'Hide records without property'.
- MOLECULE FILTERS:** Name / OCHEM ID / Inchi-Key

Records List:

| Chemical Structure | pKi Value | Record ID | Visibility |
|--------------------|------------------------------------|-----------|----------------------------|
| | 7.443697499232712 (in -log(mol/L)) | R38465835 | Only visible to rino.ragno |
| | 7.958607314841775 (in -log(mol/L)) | R38465834 | Only visible to rino.ragno |
| | 7.795880017344075 (in -log(mol/L)) | | |

The records list also includes the name 'Ragno, R', file names like 'jm7b01342_si_002.xls', and molecule IDs such as 'M97153584'. A blue arrow points to the right side of the records list, highlighting the visibility controls.

The bottom of the browser shows several open tabs: 'LigandScout_4....dmg', 'LigandScout_4....exe', 'LigandScout....tar.gz', and 'jm7b01342_si_0....csv'. A 'Show all' button is visible in the bottom right corner.

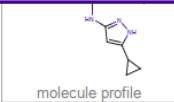
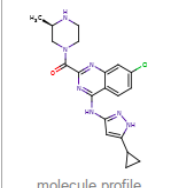
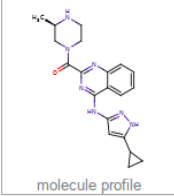
Preparing the dataset

The screenshot displays the Online Chemical Database interface. The browser address bar shows <https://ochem.eu/basket/show.do>. The page header includes the logo and text "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout".

The main content area is divided into a left sidebar with filters and a right main panel with a list of records. The sidebar includes:

- Molecular mass between [] and []
- ADVANCED MOLECULE FILTERS
- MISCELLANEOUS: Current set (jm7b01342_si_002.xls)
- Data origin and quality: Data introducers (All users), Data visibility (All data), Data from other users (All data), and checkboxes for Original records and Primary records.
- Discover issues with the data: checkboxes for Error records, Error in chies, Mismatching names, Include stereochem., Empty molecules, Show only duplicates, and No stereochemistry.
- Sort by: Creation time, with an Ascending order checkbox.
- Buttons: REFRESH, RESET

The main panel displays a list of records, each with a molecule profile image, pKi value, name, and RecordID. A blue arrow points to the "5" in the pagination control "5 items on page 1 of 8".

| Molecule profile | RecordID | Visibility |
|--|-----------|----------------------------|
|  MoleculeID: M97153582 Private record | R38465833 | Only visible to rino.ragno |
|  ● pKi = 8.154901959985743 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls N: AUTO_37 | R38465832 | Only visible to rino.ragno |
|  ● pKi = 7.769551078621726 (in -log(mol/L)) Ragno, R jm7b01342_si_002.xls N: AUTO_36 | R38465831 | Only visible to rino.ragno |

1 - 5 of 40 items on page 1 of 8

Preparing the dataset

The screenshot displays the Online Chemical Database (OChem) interface. The browser address bar shows the URL <https://ochem.eu/basket/show.do>. The page title is "Online chemical database with modeling environment". The user is logged in as "Prof. Ragno".

The main content area is titled "Compounds properties browser" and shows a list of records. A blue arrow points to the "Records" button in the toolbar, which has a tooltip that says "Select all records matching current filters".

The records are filtered by "pKi" and are displayed in a table-like format. Each record includes a chemical structure, a pKi value, and other metadata.

| RecordID | pKi | Author | File | Notes |
|-----------|-------------------|----------|----------------------|------------|
| R38465835 | 7.443697499232712 | Ragno, R | jm7b01342_si_002.xls | N: AUTO_40 |
| R38465834 | 7.958607314841775 | Ragno, R | jm7b01342_si_002.xls | N: AUTO_39 |
| R38465801 | 7.795880017344075 | Ragno, R | jm7b01342_si_002.xls | N: AUTO_38 |

The interface also includes a "FILTERS" sidebar on the left, a "Basket" icon, and a "Records" icon. The "Records" icon is highlighted with a blue arrow. The "Records" icon has a tooltip that says "Select all records matching current filters".

Preparing the dataset

The screenshot displays the 'Online chemical database with modeling environment' interface. The browser address bar shows 'https://ochem.eu/basket/show.do'. The page header includes the site logo, version 'v.3.0.96.1', and a welcome message for 'Prof. Ragno'. Navigation tabs include 'Home', 'Database', and 'Models'. A secondary navigation bar contains 'Molecule sets X', 'Edit basket X', and 'records X'. The main content area lists three molecule records, each with a chemical structure, pKi value, name, molecule ID, and record ID. Two blue arrows point to the 'Select/unselect this record' checkboxes for the second and third records.

| Chemical Structure | pKi (in -log(mol/L)) | Name | MoleculeID | RecordID | Timestamp | Visibility |
|--------------------|----------------------|---|------------|-----------|------------------|----------------------------|
| | 6.1487416512809245 | Ragno, R jm7b01342_si_002.xls N: AUTO_3 | M97153547 | R38465798 | 09:09, 18 Nov 19 | Only visible to rino.ragno |
| | 7.585026652029182 | Ragno, R jm7b01342_si_002.xls N: AUTO_2 | M95419909 | R38465797 | 09:09, 18 Nov 19 | Only visible to rino.ragno |
| | 8.045757490560675 | Ragno, R jm7b01342_si_002.xls N: AUTO_1 | M4402773 | R38465796 | 09:09, 18 Nov 19 | Only visible to rino.ragno |

1 - 40 of 40

Taskbar: LigandScout_4....dmg, LigandScout_4....exe, LigandScout....tar.gz, jm7b01342_si_0....csv, Show all X

Preparing the dataset

The screenshot displays the 'Online chemical database with modeling environment' interface. The browser address bar shows 'https://ochem.eu/basket/show.do'. The page header includes the site logo, version 'v.3.0.96.1', and a welcome message for 'Prof. Ragno'. The main navigation menu includes 'Home', 'Database', and 'Models'. A secondary menu shows 'Molecule sets', 'Edit basket', and 'records'. The 'Compounds properties browser' section is active, displaying a list of 40 records. The first three records are visible, each showing a chemical structure, a pKi value, and associated metadata. The 'pKi' property is highlighted in the filters section on the left. The taskbar at the bottom shows several open files, including 'LigandScout_4...dmg', 'LigandScout_4...exe', 'LigandScout...tar.gz', and 'jm7b01342_si_0...csv'.

Online chemical database
with modeling environment
v.3.0.96.1
Welcome, Dear Prof. Ragno! My account Logout
At a- Privacy statement

Home Database Models
Molecule sets X Edit basket X records X

Compounds properties browser
Search for numerical compounds properties linked to scientific articles
Area of your interest: no tags selected [change]

Your saved selection contains 38 records [clear]

FILTERS
SOURCE
Article/Source [select]
Page Table
PROPERTY
Activity/Property [select]
pKi
Hide records without property
CONDITIONS
MOLECULE FILTERS
Name / OCHEM ID / Inchi-Key
Similarity/substructure search
Draw a structure and search all the molecules containing it or similar to it

Basket Records Tags
1 - 40 of 40

| Chemical Structure | pKi | RecordID | Date | Author | Visibility |
|--------------------|------------------------------------|-----------|------------------|------------|----------------------------|
| | 7.443697499232712 (in -log(mol/L)) | R38465835 | 09:09, 18 Nov 19 | rino.ragno | Only visible to rino.ragno |
| | 7.958607314841775 (in -log(mol/L)) | R38465834 | 09:09, 18 Nov 19 | rino.ragno | Only visible to rino.ragno |
| | 7.795880017344075 (in -log(mol/L)) | | | | |

LigandScout_4...dmg LigandScout_4...exe LigandScout...tar.gz jm7b01342_si_0...csv Show all X

Preparing the dataset

The screenshot shows the Online Chemical Database interface. The browser address bar displays <https://ochem.eu/basket/show.do>. The page header includes the logo for the Online chemical database with modeling environment, the version number v.3.0.96.1, and a welcome message for Prof. Ragno with links for My account and Logout. The navigation menu includes Home, Database, and Models. The main content area is titled 'Basket browser' and contains a filter by name field, a 'Create new' button, and a 'Show public sets' checkbox. Below this, there are two entries in the basket:

| Entry Name | Records | Pending Models |
|----------------------|------------|------------------|
| Selected records | 38 records | |
| jm7b01342_si_002.xls | 40 records | 1 pending models |

A blue arrow points to the '38 records' value for the 'Selected records' entry. The taskbar at the bottom shows several files related to LigandScout, including a .dmg file, an .exe file, a .tar.gz file, and a .csv file.

Preparing the dataset

The screenshot shows the 'Basket editor' interface of the Online Chemical Database. The page title is 'Online chemical database with modeling environment'. The user is logged in as 'Prof. Ragno'. The interface includes a navigation menu with 'Home', 'Database', and 'Models'. The main content area is titled 'Basket editor' and contains the following elements:

- Name:** A text input field containing '38 Mols' with a note '(min. 2 characters)'. A blue arrow points to this field.
- Description (optional):** A text input field.
- Excluded implicit records (under development):** A text input field.
- Actions:** A list of actions including 'Create a copy of this basket', 'Split the basket into two sets', 'Create a primary records basket', 'Transform the basket using OScript', 'Add or delete particular records', 'Export this basket into Excel, CSV or SDF', and 'Discretize the numerical values'.
- Statistics of the basket:** A table showing the following data:

| Properties | Records | Unique compounds | |
|------------|------------|------------------|-----------|
| pKi | 38 records | 38 compounds | Show MMPs |

The bottom of the browser window shows several open files: 'LigandScout_4...dmg', 'LigandScout_4...exe', 'LigandScout...tar.gz', and 'jm7b01342_si_0...csv'. A 'Show all' button is visible in the bottom right corner.

Preparing the dataset

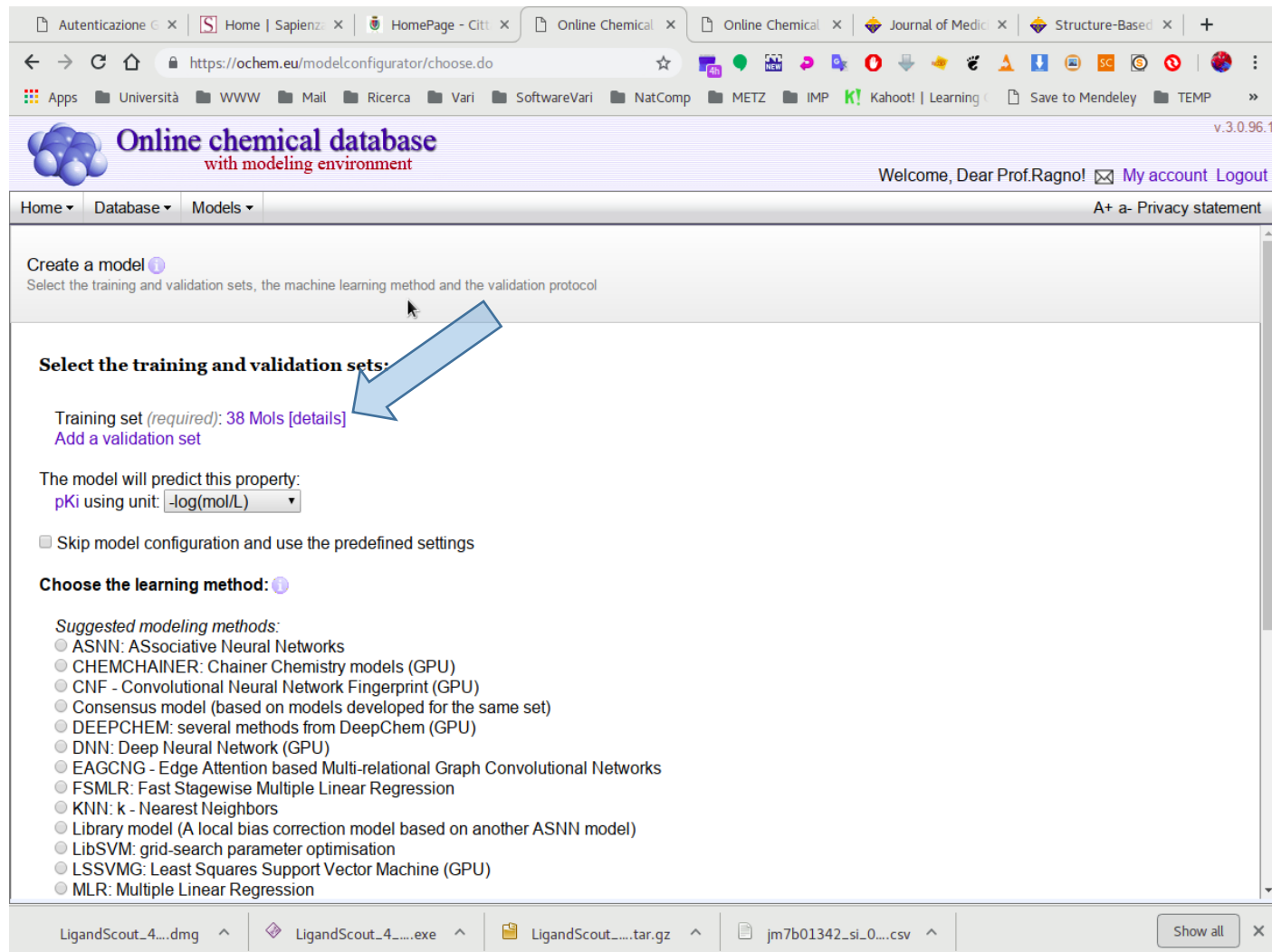
The screenshot shows the Online Chemical Database interface. The browser address bar displays <https://ochem.eu/basket/show.do>. The page header includes the logo and text "Online chemical database with modeling environment" and a version number "v.3.0.96.1". A welcome message "Welcome, Dear Prof.Ragno!" and links for "My account" and "Logout" are visible. The navigation menu contains "Home", "Database", and "Models".

The main content area is titled "Basket browser" and includes the subtext "Browse, Compare or Join molecule sets". Below this is a search filter "Filter by name:" with a text input field, a "[Create new]" button, and a "Show public sets" checkbox. The results show "1 - 2 of 2" items:

| Item Name | Records | Models |
|----------------------|------------|------------------|
| 38 Mols | 38 records | |
| jm7b01342_si_002.xls | 40 records | 1 pending models |

A blue arrow points to the "38 Mols" entry. The taskbar at the bottom shows several files related to "LigandScout_4" and a "jm7b01342_si_002.csv" file.

Preparing the dataset



Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models At a- Privacy statement

Create a model ⓘ
Select the training and validation sets, the machine learning method and the validation protocol

Select the training and validation sets:

Training set (required): 38 Mols [details]
Add a validation set

The model will predict this property:
pKi using unit: -log(mol/L)

Skip model configuration and use the predefined settings

Choose the learning method: ⓘ

Suggested modeling methods:

- ASNN: ASsociative Neural Networks
- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid-search parameter optimisation
- LSSVMG: Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression

LigandScout_4....dmg ^ | LigandScout_4....exe ^ | LigandScout....tar.gz ^ | jm7b01342_si_0....csv ^ | Show all x

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page title is "Online chemical database with modeling environment" and the version is "v.3.0.96.1". The user is logged in as "Prof.Ragno!".

The main content area displays a list of models under the "Models" tab:

- CHEMCHAINER: Chainer Chemistry models (GPU)
- CNF - Convolutional Neural Network Fingerprint (GPU)
- Consensus model (based on models developed for the same set)
- DEEPCHEM: several methods from DeepChem (GPU)
- DNN: Deep Neural Network (GPU)
- EAGCNG - Edge Attention based Multi-relational Graph Convolutional Networks
- FSMLR: Fast Stagewise Multiple Linear Regression
- KNN: k - Nearest Neighbors
- Library model (A local bias correction model based on another ASNN model)
- LibSVM: grid-search parameter optimisation
- LSSVMG: Least Squares Support Vector Machine (GPU)
- MLR: Multiple Linear Regression
- PLS: Partial Least Squares
- RFR: Random Forest regression and classification
- WEKA-J48: Weka C4.5 decision trees, only classification - use with bagging
- WEKA-RF: Random Forest, only classification
- XGBoost: Scalable and Flexible Gradient Boosting

Below the list, there is a section for "Methods under development:" and a "Model validation" section:

Model validation
Validation method:
Number of folds:
 Stratified cross-validation (classification only)
 Consider each record as a molecule. ⓘ

You can create a model from template: [import an XML model template](#) or [use another model as a template](#)

A blue arrow points to a "Next->" button at the bottom left of the configuration area.

The browser's taskbar at the bottom shows several files: "LigandScout_4_...dmg", "LigandScout_4_...exe", "LigandScout_...tar.gz", and "jm7b01342_si_0...csv".

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for 'Online chemical database with modeling environment' and a welcome message: 'Welcome, Dear Prof.Ragno! My account Logout'. A navigation bar contains 'Home', 'Database', and 'Models' menus. The main content area is titled 'Model creator' and 'Select model template and training set'. Below this, there is a section 'Select the preferred data preprocessing options' with the sub-section 'Preprocessing of molecules (Chemaxon)'. Four options are listed with checked checkboxes: 'Standardization', 'Neutralize', 'Remove salts', and 'Clean structure'. At the bottom of this section are two buttons: '<<Back' and 'Next>>'. A blue arrow points to the 'Next>>' button. The browser's taskbar at the bottom shows several files: 'LigandScout_4....dmg', 'LigandScout_4_.....exe', 'LigandScout_.....tar.gz', and 'jm7b01342_si_0....csv'.

Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! | Learning Save to Mendeley TEMP »

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Model creator
Select model template and training set

Select the preferred data preprocessing options

Preprocessing of molecules (Chemaxon)

- Standardization
- Neutralize
- Remove salts
- Clean structure

<<Back Next>>

LigandScout_4....dmg LigandScout_4_.....exe LigandScout_.....tar.gz jm7b01342_si_0....csv Show all x

Preparing the dataset

Autenticazione | Home | Sapienz | HomePage - Citt | Online Chemical | Online Chemical | Journal of Medic | Structure-Based | +

https://ochem.eu/modelconfigurator/choose.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendelej | TEMP

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

Model creator
Select model template and training set

Select the molecular descriptors ⓘ

Recommended descriptor types

- E-state
- ALogPS (2)
- GSFragment (1138)
- CDK 2.0 descriptors (256/3D)
- Dragon v. 7 (5270/3D)

[\[select all\]](#) [\[select none\]](#)

- Constitutional descriptors (47)
- Topological indices (75)
- Connectivity indices (37)
- 2D matrix-based descriptors (607)
- Burden eigenvalues (96)
- ETA indices (23)
- Geometrical descriptors (3D, 38)
- 3D autocorrelations (3D, 80)
- 3D-MoRSE descriptors (3D, 224)
- GETAWAY descriptors (3D, 273)
- Functional group counts (3D, 154)
- Atom-type E-state indices (172)
- 2D Atom Pairs (1596)
- Charge descriptors (3D, 15)
- Drug-like indices (28)
- Ring descriptors (32)
- Walk and path counts (46)
- Information indices (50)
- 2D autocorrelations (213)
- P_VSA-like descriptors (55)
- Edge adjacency indices (324)
- 3D matrix-based descriptors (3D, 99)
- RDF descriptors (3D, 210)
- WHIM descriptors (3D, 114)
- Randic molecular profiles (3D, 41)
- Atom-centred fragments (115)
- CATS 2D (150)
- 3D Atom Pairs (3D, 36)
- Molecular properties (20)
- CATS 3D (3D, 300)

Predictions by OCHEM's featured models ⓘ

- Ames levenberg
- Toxicity against T. Pyriformis
- ALogPS 3.0
- CYP1A2 Estate+ALogPS
- CYP2C9 Estate+ALogPS
- CYP2C19 Estate+ALogPS
- CYP2D6 Estate+ALogPS
- CYP3A4 Estate+ALogPS
- Pyrolysis point prediction (best Estate)
- Melting Point prediction (best Estate)
- Water solubility model based on logP and Melting Point
- ALOGPS 2.1 logP
- ALOGPS 2.1 logS

Outputs of other OCHEM models

Obsolete/Additional descriptor types

- CDK 1.4.11 descriptors (256/3D)
- OESState
- Dragon v. 5.4 (1644/3D)
- Dragon v. 5.5 (3224/3D)

LigandScout_4....dmg | LigandScout_4....exe | LigandScout....tar.gz | jm7b01342_si_0....csv | Show all

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page title is "Online chemical database with modeling environment" and the version is "v.3.0.96.1". The user is logged in as "Prof. Ragno".

The configuration page has a navigation menu with "Home", "Database", and "Models". The "Database" section is expanded, showing a list of descriptors with checkboxes:

- Chemaxon descriptors (400/3D)
- QNPR
- Spectrophores (144/3D)
- Structural alerts (ToxAlerts)

Below this, there are sections for "Special descriptors (scaffolds, fingerprints):" and "Under development: can change anytime and backward compatibility is not guaranteed. Use at your own risk!".

The "Special descriptors" section includes:

- Chemaxon Scaffolds
- Silicos-It Scaffolds
- ECFP Fingerprints *Not supported by your installation*
- MolPrint Fingerprints

The "Under development" section includes:

- JPllogP
- CDK 2.2 descriptors (256/3D)
- RDKit descriptors (3D)
- RDKit additional descriptors (3D)
- MORDRED descriptors (1826/3D)
- CDDD
- MOPAC2016 descriptors (35/3D)
- SIRMS
- PyDescriptor descriptors (16251/3D)
- External descriptors
- Allow Merging Descriptors (experimental)

At the bottom of the configuration area, there are two buttons: "<<Back" and "Next>>". A blue arrow points to the "Next>>" button, indicating the next step in the process.

The Windows taskbar at the bottom shows several files: "LigandScout_4_...dmg", "LigandScout_4_...exe", "LigandScout_...tar.gz", and "jm7b01342_si_0...csv".


Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". A navigation menu contains "Home", "Database", and "Models". The main content area is titled "Model creator" and instructs the user to "Select model template and training set". Below this, a section titled "Select a tool to optimize molecule structures" lists five options with radio buttons: "No optimisation", "Optimise with Corina", "Optimise with OpenBabel", "Optimise with OBGEN (part of OpenBabel distribution)", and "Optimise with BALLOON". A blue arrow points to the "Optimise with OpenBabel" option. At the bottom of the selection area, there are two buttons: "<<Back" and "Next>>". The browser's taskbar at the bottom shows several files, including "LigandScout_4....dmg", "LigandScout_4_....exe", "LigandScout_....tar.gz", and "jm7b01342_si_0....csv".

Autenticazione x Home | Sapienz x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

← → ↻ 🏠 <https://ochem.eu/modelconfigurator/choose.do> ☆

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP »

 **Online chemical database**
with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

Model creator
Select model template and training set

Select a tool to optimize molecule structures

- No optimisation
- Optimise with Corina
- Optimise with OpenBabel
- Optimise with OBGEN (part of OpenBabel distribution)
- Optimise with BALLOON

<<Back Next>>

LigandScout_4....dmg ^ LigandScout_4_....exe ^ LigandScout_....tar.gz ^ jm7b01342_si_0....csv ^ Show all x

Preparing the dataset

The screenshot shows the 'Model creator' interface of the Online Chemical Database. The page title is 'Online chemical database with modeling environment'. The user is logged in as 'Prof. Ragno!'. The interface includes a navigation menu with 'Home', 'Database', and 'Models'. The main content area is titled 'Model creator' and contains the following sections:

- Select filters of descriptors**
 - Eliminate descriptors with less than unique values
 - Delete descriptors that have absolute values larger than
 - Delete descriptors that have variance smaller than
 - Group descriptors, that have pair-wise correlations Pearson's correlation coefficient R larger than
 - Use Unsupervised Forward Selection to delete variables using the above value of multiple correlation coefficient R
 - After filtering, I want to select necessary descriptors myself (*advanced*)
- Normalisation parameters**
 - Descriptors normalization:
 - Values normalization:

At the bottom of the form, there are navigation buttons: '<<Back' and 'Next>>'. A blue arrow points from the 'Next>>' button towards the 'Standardize' dropdown menu. The taskbar at the bottom shows several files: 'LigandScout_4....dmg', 'LigandScout_4.....exe', 'LigandScout.....tar.gz', and 'jm7b01342_si_0....csv'.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page title is "Online chemical database with modeling environment" and the version is "v.3.0.96.1". The user is logged in as "Prof. Ragno". The navigation menu includes "Home", "Database", and "Models". The main content area is titled "Model creator" and "Select model template and training set". Under the heading "Configure PLS method", there is a form with the following elements:

- Number of latent variables:
- Optimize the number of latent variables automatically
- Limit predicted values to the training set range

At the bottom of the form, there are two buttons: "<<Back" and "Next>>". A blue arrow points to the "Next>>" button. The browser's taskbar at the bottom shows several files: "LigandScout_4....dmg", "LigandScout_4_....exe", "LigandScout_....tar.gz", and "jm7b01342_si_0....csv".

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for the Online chemical database with modeling environment, the version number v.3.0.96.1, and a welcome message for Prof. Ragno. The main content area is titled 'Model creator' and 'Start calculation of the model'. It contains a text input field with the value 'pKi_PLS_[Dragon7 (blocks: 1-30)] - 336948', a checked checkbox for 'Save models', and radio buttons for task priority: 'High priority (please, use for fast tasks only)', 'Normal priority' (selected), and 'Low priority (for long tasks)'. At the bottom, there are three buttons: '<<Back', 'Start calculation' (highlighted in red and pointed to by a blue arrow), and 'Discard'. The browser's taskbar at the bottom shows several files, including 'LigandScout_4_...dmg', 'LigandScout_4_...exe', 'LigandScout_...tar.gz', and 'jm7b01342_si_0...csv'.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". A navigation menu contains "Home", "Database", and "Models". The main content area is titled "Model creator" with the instruction "Select model template and training set". Below this, the "Run model builder" section displays a "Starting..." status with a circular progress indicator and links for "[cancel]" and "[fetch result later]". At the bottom of the main area are navigation buttons: "<<Back" and "Next>>". The browser's taskbar at the bottom shows several files: "LigandScout_4...dmg", "LigandScout_4...exe", "LigandScout...tar.gz", and "jm7b01342_si_0...csv".

Preparing the dataset

The screenshot shows a web browser window with multiple tabs. The active tab is 'Online Chemical' at the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the 'Online chemical database with modeling environment' logo and version 'v.3.0.96.1'. A user is logged in as 'Prof. Ragno!' with links for 'My account' and 'Logout'. A navigation menu contains 'Home', 'Database', and 'Models'. The main content area is titled 'Run model builder' and displays a status message: 'Finished posting ... - Processing task Corina - Waiting for a free server -- 09:37'. Below the message are links for '[cancel]' and '[fetch result later]'. At the bottom of the content area are '<<Back' and 'Next>>' buttons. The browser's taskbar at the bottom shows several files: 'LigandScout_4...dmg', 'LigandScout_4...exe', 'LigandScout...tar.gz', and 'jm7b01342_si_0...csv', along with a 'Show all' button.

Preparing the dataset

The screenshot shows a web browser window with the URL <https://ochem.eu/modelconfigurator/choose.do>. The page header includes the logo for "Online chemical database with modeling environment" and a welcome message: "Welcome, Dear Prof.Ragno! My account Logout". A navigation menu contains "Home", "Database", and "Models".

The main content area is titled "Model creator" with the instruction "Select model template and training set". Below this, the "Run model builder" section displays a progress indicator (a circular loading spinner) and the message: "Finished posting ... - Processing task Descriptors - Tasks are sent for calculations -- 09:37". There are two links below the message: "[cancel]" and "[fetch result later]". At the bottom of this section are two buttons: "<<Back" and "Next>>".

The browser's taskbar at the bottom shows several open files: "LigandScout_4....dmg", "LigandScout_4.....exe", "LigandScout.....tar.gz", and "jm7b01342_si_0....csv". A "Show all" button is visible on the right side of the taskbar.

Preparing the dataset

Autenticazione G x | Home | Sapienz: x | HomePage - Citt x | Online Chemical x | Online Chemical x | Journal of Medic x | Structure-Based x | +

https://ochem.eu/modelconfigurator/choose.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning | Save to Mendelej | TEMP >>

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

v.3.0.96.1

Home ▾ Database ▾ Models ▾ A+ a- Privacy statement

Model creator

Select model template and training set

Save the model

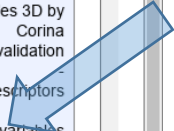
Please enter your model's name:

Overview

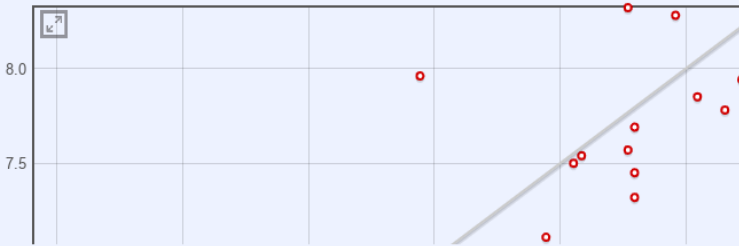
Model name: pKi_PLS_[Dragon7 (blocks: 1-30)] - 336948 [\[rename\]](#) [Dragon7 (blocks: 1-30)]
Temporal Public ID: [11990802](#) - use this link to share the model Correl. limit: 0.0 Variance threshold: 0.01,
Maximum value: 999999,

Predicted property: **pKi** modeled in -log(mol/L)
Training method: PLS scale X: STANDARDIZE0 latent variables 3D by
Corina
5-fold cross-validation

| Data Set | # | R2 | q2 | RMSE | MAE |
|-------------------------|------------|-------------|-------------|-------------|-------------|
| ● Training set: 38 Mols | 38 records | 0.79 ± 0.06 | 0.79 ± 0.06 | 0.38 ± 0.04 | 0.31 ± 0.04 |

2990 pre-filtered descriptors 

scale X: STANDARDIZE4 latent variables
Y = -0.62 + 9.86E-6**MW* - 0.00151**AMW* +
1.08E-4**Sv* + 7.59E-5**Se* + 1.51E-4**Sp* +
1.07E-4**Si* - 0.0974**Mv* - 0.346**Me* -
0.0454**Mp* + 0.133**Mi* + 0.129**GD* + 1.07E-
4**nAT* - 1.04E-4**nSK* - 0.00147**nTA* + 1.3E-
4**nBT* + 1.04E-4**nBO* - 9.95E-4**nBM* - 1.05E-
4**SCBO* - 4.56E-5**RBN* - 0.0334**RBF* - 6.43E-
4**nDB* - 7.49E-4**nAB* + 3.0E-4**nH* - 7.54E-
5**nC* + 0.0169**nN* - 0.00448**nO* - 3.04E-4**nCL*
+ 0.00136**nHM* - 2.29E-4**nHet* + 0.00449**nX* +
5.98E-4**H%* - 0.00122**C%* + 0.00224**N%* -
0.00223**O%* + 0.00189**X%* + 7.52E-4**nCsp3* -



LigandScout_4....dmg ^ | LigandScout_4.....exe ^ | LigandScout.....tar.gz ^ | jm7b01342_si_0....csv ^ | Show all x

Preparing the dataset

accounts.google.com | Il mio Drive - Google Drive | Online Chemical Modeling Environment

ochem.eu/pendingtasks/tasks.do

Online chemical database with modeling environment v.3.0.96.2

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models A+ a- Privacy statement

Pending tasks X Task results X

Please enter your model's name:

Overview

Model name: pKi_PLS_[Dragon7 (blocks: 1-30)] - 337241 [rename] [Dragon7 (blocks: 1-30)]
Temporal Public ID: 35823821 - use this link to share the model
Correl. limit: 0.0 Variance threshold: 0.01, Maximum value: 999999,

Predicted property: **pKi** modeled in -log(mol/L)
Training method: PLS

scale X: STANDARDIZE0 latent variables 3D by Corina
5-fold cross-validation
2990 pre-filtered descriptors

| Data Set | # | R2 | q2 | RMSE | MAE |
|-----------------------|------------|-------------|-------------|-------------|-------------|
| Training set: 38 Mols | 38 records | 0.79 ± 0.05 | 0.78 ± 0.06 | 0.39 ± 0.04 | 0.32 ± 0.03 |

scale X: STANDARDIZE5 latent variables
Y = -11.2 - 7.65E-6*MW - 0.00236*AMW + 8.03E-5*Sv + 5.85E-5*Se + 9.21E-5*Sp + 9.43E-5*Si - 0.093*Mv - 0.373*Me - 0.0664*Mp + 0.208*Mi + 0.115*GD + 8.62E-5*AT - 7.19E-5*nSK - 0.00171*nTA + 1.12E-4*nBT + 1.2E-4*nBO - 7.08E-4*nBM - 4.88E-5*SCBO + 1.48E-4*RBN - 0.0125*RBF - 7.32E-4*nDB - 5.02E-4*nAB + 2.37E-4*nH - 1.02E-4*nC + 0.0202*nN - 0.00489*nO - 6.76E-5*nCL - 1.06E-4*nHM + 2.88E-4*nHet + 0.00475*nX + 5.04E-4*H% - 0.00119*C% + 0.00284*N% - 0.00245*O% + 0.00216*X% + 4.81E-4*nCsp3 - 7.08E-4*nCsp2 + 0.00211*nCIC + 0.00211*nCIR + 3.84E-4*TRS + 3.84E-4*Rperim + 0.0276*MCD - 0.116*RFD - 0.0814*RCI + 0.00211*NRS + 0.0422*NNRS + 0.00286*nR03 - 4.71E-4*nR06 - 0.00301*nBnz - 0.0191*ARR + 6.28E-5*D/Dtr03 - 8.59E-5*... + 1.15E-5*... + 0.00171*... + 0.00171*...


Save Discard

Preparing the dataset

Autenticazione x Home | Sapienz: x HomePage - Citt x Online Chemical x Online Chemical x Journal of Medic x Structure-Based x +

https://ochem.eu/modelconfigurator/choose.do

Apps Università WWW Mail Ricerca Vari SoftwareVari NatComp METZ IMP Kahoot! Learning Save to Mendeley TEMP >>

 **Online chemical database**
with modeling environment v.3.0.96.1

Welcome, Dear Prof.Ragno! [My account](#) [Logout](#)

Home Database Models A+ a- Privacy statement

Model creator

Your model has been saved
Thank you for your cooperation.

Your next possible actions are:
[Apply your model](#)
[View your model's properties](#)

LigandScout_4...dmg ^ LigandScout_4_...exe ^ LigandScout_...tar.gz ^ jm7b01342_si_0...csv ^ Show all x

Preparing the dataset

accounts.google.com | Il mio Drive - Google Drive | Online Chemical Modeling Environm

ochem.eu/pendingtasks/tasks.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning G...

Online chemical database
with modeling environment
v.3.0.96.2

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models

Pending tasks X Task results X

Model profile

Statistical parameters, tables, charts - all the information related to the model.

Overview

Model name: pKi_PLS_[Dragon7 (blocks: 1-30)] - 337241 [rename]
Temporal Public ID: 35823821 - use this link to share the model

Predicted property: pKi modeled in -log(mol/L)
Training method: PLS

| Data Set | # | R2 | q2 | RMSE | MAE |
|-----------------------|------------|-------------|-------------|-------------|-------------|
| Training set: 38 Mols | 38 records | 0.79 ± 0.05 | 0.78 ± 0.06 | 0.39 ± 0.04 | 0.32 ± 0.04 |

[Dragon7 (blocks: 1-30)]
Correl. limit: 0.0 Variance threshold: 0.01,
Maximum value: 999999,
scale X: STANDARDIZED0 latent variables 3D
by Corina
5-fold cross-validation
-
2990 pre-filtered descriptors

scale X: STANDARDIZE5 latent variables
Y = -11.2 - 7.66E-6*MW - 0.00236*AMW +
8.03E-5*Sv + 5.86E-5*Se + 9.21E-5*Sp +
9.43E-5*Si - 0.093*Mv - 0.373*Me - 0.086*Mp
+ 0.208*Mi + 0.115*GD + 8.62E-5*nAT -
7.19E-5*nSK - 0.00171*nTA + 1.12E-4*nBT +
1.2E-4*nBO - 7.08E-4*nBM - 4.88E-5*SCBO +
1.48E-4*RBN - 0.0125*RBF - 7.32E-4*nDB -
5.02E-4*nAB + 2.37E-4*nH - 1.02E-4*nC +
0.0202*nN - 0.00489*nO - 6.76E-5*nCL -
1.08E-4*nHM + 2.88E-4*nHet + 0.00475*nX +
5.04E-4*H% - 0.00119*C% + 0.00284*N% -
0.00245*O% + 0.00216*X% + 4.81E-4*nCsp3
- 7.08E-4*nCsp2 + 0.00211*nClC +
0.00211*nClR + 3.84E-4*TRS + 3.84E-
4*Rperim + 0.0276*MCD - 0.116*RFD -
0.0814*RCI + 0.00211*NRS + 0.0422*NNRS +
0.00286*nR03 - 4.71E-4*nR06 - 0.00301*nBnz
- 0.0191*ARR + 6.28E-5*D/Dtr03 - 8.59E-
5*D/Dtr05 - 1.17E-5*D/Dtr06 - 9.19E-
5*D/Dtr10 + 2.18E-5*ZM1 - 4.66E-5*ZM1V -

Preparing the dataset

The screenshot displays the OCHEM web interface. At the top, there are browser tabs for 'accounts.google.com', 'Il mio Drive - Google Drive', and 'Online Chemical Modeling Environm...'. The address bar shows 'ochem.eu/pendingtasks/tasks.do'. The page header includes the OCHEM logo, the text 'Online chemical database with modeling environment', and the version 'v.3.0.96.2'. A welcome message for 'Prof. Ragno!' is visible, along with links for 'My account' and 'Logout'. Below the header, there are navigation tabs for 'Home', 'Database', and 'Models'. The main content area is divided into 'Pending tasks X' and 'Task results X'. The 'Task results X' tab is active, showing a scatter plot on the left and a list of model descriptors on the right. The scatter plot has a y-axis from 5.5 to 6.5 and an x-axis from 5.5 to 8.0. A blue arrow points to a data point at approximately (6.5, 6.5) with the label 'Measured value'. Below the plot are links for 'Download descriptors and model statistics', 'MMP-based analysis (experimental)', 'Create a copy of this model', 'View configuration XML', and 'Export configuration XML'. The list of descriptors on the right is a long mathematical expression involving various chemical descriptors and their coefficients.

0.0814*RCI + 0.00211*NRS + 0.0422*NNRS + 0.00286*nR03 - 4.71E-4*nR06 - 0.00301*nBnz - 0.0191*ARR + 8.28E-5*D/Dtr03 - 8.59E-5*D/Dtr05 - 1.17E-5*D/Dtr06 - 9.19E-5*D/Dtr10 + 2.18E-5*ZM1 - 4.68E-5*ZM1V - 3.11E-5*ZM1Kup - 2.24E-5*ZM1Mad - 2.86E-5*ZM1Per - 4.38E-5*ZM1MulPer + 2.96E-5*ZM2 - 1.46E-5*ZM2V - 1.4E-5*ZM2Kup + 1.86E-5*ZM2Mad - 1.06E-5*ZM2Per - 2.41E-5*ZM2MulPer - 0.0012*OND - 4.68E-4*ONDV - 3.78E-4*ON1 + 0.00388*ON1V + 3.01E-4*Qindex + 6.7E-5*BBI - 0.00276*DBI + 3.6E-4*SNar + 0.025*HNar + 0.031*GNar + 0.0596*Xt - 9.7E-5*Dz - 3.31E-4*Ram + 0.0613*BLI + 2.62E-5*Pol - 5.89E-6*LPRS + 0.00263*MSD - 0.001*SPI + 0.0066*PJ12 - 3.61E-5*ECC - 0.00369*AEOC - 0.0136*DECC - 1.98E-4*MDDD - 8.19E-5*UNIP + 7.34E-7*CENT - 1.58E-4*VAR - 0.0136*ICR - 4.56E-9*SMTI - 2.43E-7*SMTIV + 1.0E-7*GMTI - 2.56E-7*GMTIV - 1.29E-4*Xu - 1.08E-5*CSI + 4.48E-5*Wap - 8.11E-4*S1K - 0.00167*S2K - 0.00311*S3K - 0.00338*PHI - 0.106*PW2 - 0.288*PW3 + 0.329*PW4 + 0.0976*PW5 + 1.21E-4*MAXDN + 3.72E-4*MAXDP - 9.66E-5*DELS - 2.14E-4*TIE - 3.08E-4*Psi_i_s - 0.0141*Psi_i_A + 2.61E-4*Psi_i_0 + 4.12E-4*Psi_i_1 - 0.0141*Psi_e_A + 1.66E-5*Psi_e_0 + 6.92E-4*Psi_e_1 - 2.69E-4*BAC - 0.00767*LOC + 0.0028*MWCO1 + 0.00123*MWCO2 + 0.00441*MWCO3 + 0.00408*MWCO4 + 0.00622*MWCO5 + 0.00447*MWCO6 + 0.00614*MWCO7 - 0.00177*MWCO8 - 0.00587*MWCO9 + 0.00273*MWCO10 + 0.003*SRW02 + 0.00147*SRW03 + 0.00282*SRW04 + 0.00187*SRW05 + 0.00295*SRW06 + 0.00232*SRW07 + 0.00398*SRW08 + 0.00254*SRW09 + 0.0104*SRW10 + 0.0028*MPC01 +

Preparing the dataset

accounts.google.com | Il mio Drive - Google Drive | Online Chemical Modeling Environm

ochem.eu/pendingtasks/tasks.do

Apps | Università | WWW | Mail | Ricerca | Vari | SoftwareVari | NatComp | METZ | IMP | Kahoot! | Learning G...

Online chemical database
with modeling environment

Welcome, Dear Prof.Ragno! My account Logout

Home Database Models

Pending tasks X Task results X Download descriptors and model statistics X

Data export
Export the selected data as an Excel, CSV or SDF file

Please, select the items that you want to export:

[select all] [select unrestricted only] [select none]

- Structure (SMILES or SDF)
- CASRN
- RECORDID
- MOLECULEID
- External unique identifier
- Identifier in article (N)
- NAMES
- Introdurers of the records
- Last modifiers of the records
- Publication IDs
- Error messages
- Experimentally measured values
- Predicted values (in converted units)
- Experimentally measured values (in converted units)
- DM (distance to model) values
- Applicability Domain (FALSE if predictions are outside of the AD)
- Conditions of experiments
- DESCRIPTORSNAMES
- Comments
- Inchi-key
- Merge information for the same molecule

Select the units to which the exported values will be converted:
pKi [-log(M)]

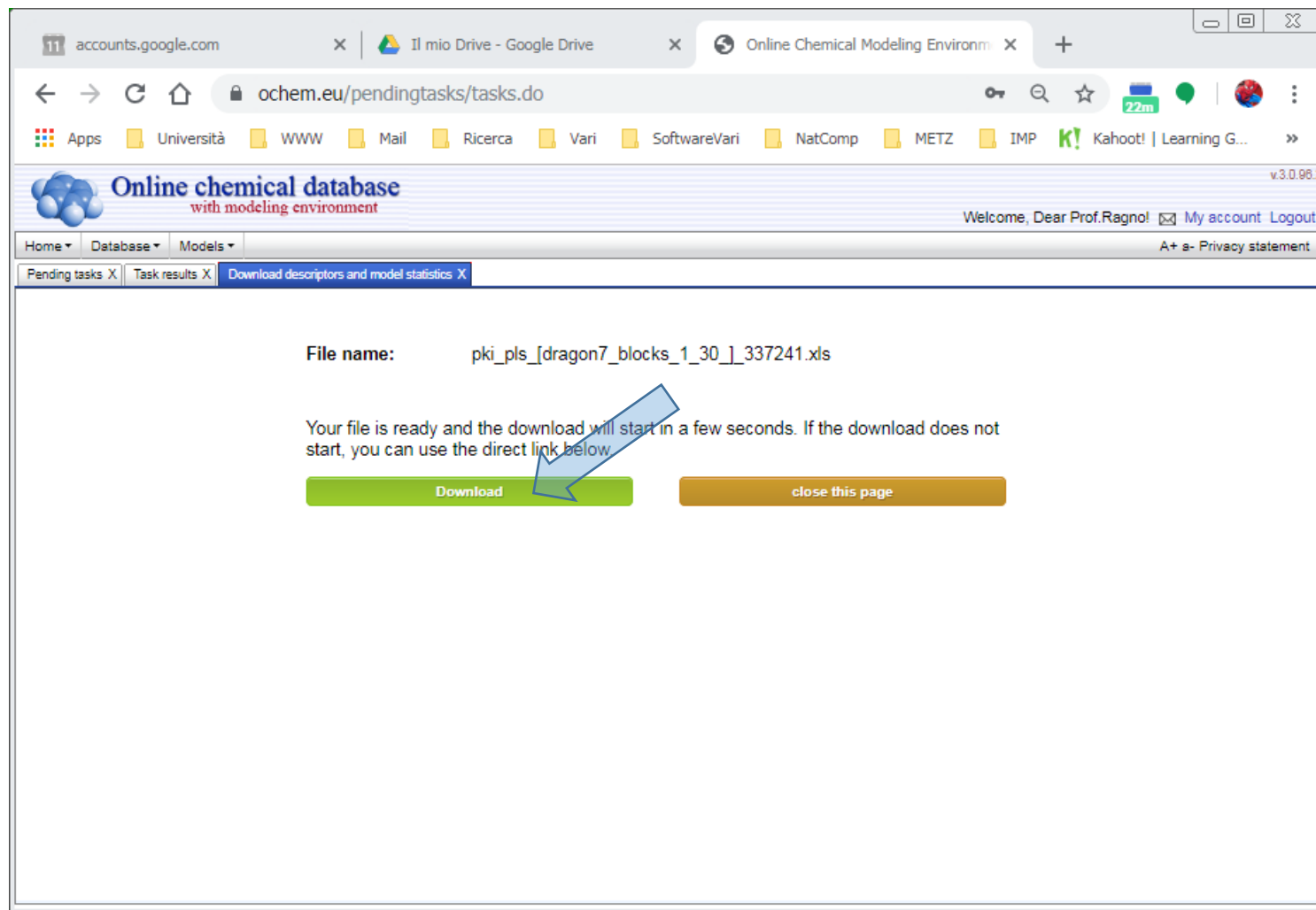
Preparing the dataset

The screenshot shows a web browser window with the URL `ochem.eu/pendingtasks/tasks.do`. The page title is "Online chemical database with modeling environment" (version 3.0.98.2). The user is logged in as "Prof.Ragno!". The main content area is titled "Download descriptors and model statistics" and contains the following elements:

- Instructions: "Please, select the items that you want to export."
- Selection options: "[select all] [select unrestricted only] [select none]"
- Checklist of exportable items (all checked):
 - Structure (SMILES or SDF)
 - CASRN
 - RECORDID
 - MOLECULEID
 - External unique identifier
 - Identifier in article (N)
 - NAMES
 - Introducers of the records
 - Last modifiers of the records
 - Publication IDs
 - Error messages
 - Experimentally measured values
 - Predicted values (in converted units)
 - Experimentally measured values (in converted units)
 - DM (distance to model) values
 - Applicability Domain (FALSE if predictions are outside of the AD)
 - Conditions of experiments
 - DESCRIPTORSNAMES
 - Comments
 - Inchi-key
 - Merge information for the same molecule
- Unit selection: "Select the units to which the exported values will be converted:" with a dropdown menu showing "pKi -log(M)".
- Export buttons: "Get Excel file", "Get CSV file", "Get SDF file", and "Get R script".

A blue arrow points from the unit selection dropdown to the "Get Excel file" button.

Preparing the dataset



The screenshot shows a web browser window with the URL `ochem.eu/pendingtasks/tasks.do`. The page header includes the logo for "Online chemical database with modeling environment" and a version number "v.3.0.98.2". A navigation bar contains links for "Home", "Database", and "Models". The main content area displays the following information:

File name: `pki_pls_[dragon7_blocks_1_30_]_337241.xls`

Your file is ready and the download will start in a few seconds. If the download does not start, you can use the direct link below.

Two buttons are visible: a green "Download" button and a brown "close this page" button. A blue arrow points to the "Download" button.