

# Batch data upload

---

Batch upload wizard allows you to upload multiple records simultaneously from a specially formatted Excel file, CSV file or an SDF. This page describes the wizard steps. For file format and keyword details refer to [batch upload file format](#) section.

## Step 1 - File selection

**Instructions**

Here you have the possibility to upload data from an excel sheet. The format of this sheet 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@eadmet.com](mailto:info@eadmet.com)

Select a file to upload

Upload file

Settings

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

You can reach first step of the batch upload wizard by selecting "Database > Batch data upload" menu item. At this step you can select the file to upload.

The **Allow molecule lookup by name on PubChem** checkbox is relevant for cases, when molecules in the uploaded file are identified by names rather than by structure (SDF or SMILES). When this checkbox is selected, the OCHEM batch upload wizard will attempt to find molecule structures by the provided name in the [PubChem](#) database. The process can be rather slow and unsuitable for large amount of molecules.

When the **Allow article lookup by PMID on PubMed** checkbox is checked, the articles provided by [PubMed](#) PMID rather than OCHEM AID will be retrieved from PubMed automatically.

When the **Make the uploaded records hidden** checkbox is checked, all the uploaded data is marked as hidden (private). Note, that you can make your data hidden by defining a "HIDDEN" column in the uploaded data file (see file format section for details).

Once you have selected the data file you would like to upload, you can proceed to Step 2 by pressing the "Upload" button. If the file is corrupt or of the unrecognized format, you will be redirected back to the current step with the error message.

## Step 2 - Column remapping

LogKOC

<input checked="" type="checkbox"/> casrn	<input checked="" type="checkbox"/> name	<input checked="" type="checkbox"/> LogKoc_Tutorial	<input checked="" type="checkbox"/> molecule	<input checked="" type="checkbox"/> article
15972-60-8	alachlor	2.28	CCC1=CC=CC(CC)=C1...	A17025
23184-66-9	butachlor	2.86	CCCCOCN(C(=O)CCl)...	A17025
1918-16-7	propachlor	2.42	CC(C)N(C(=O)CCl)C...	A17025
1646-88-4	Aldicarb Sulfone	0.42	CNC(=O)O\N=C\C(C)...	A17025
2008-41-5	butylate	2.11	CCSC(=O)N(CC(C)C)...	A17025
63-25-2	carbaryl	2.4	CNC(=O)OC1=CC=CC2...	A17025
1563-66-2	Carbofuran	1.75	CNC(=O)OC1=CC=CC2...	A17025
101-21-3	Chlorpropham	2.53	CC(C)OC(=O)NC1=CC...	A17025
1134-23-2	cycloate	2.54	CCSC(=O)N(CC)C1CC...	A17025

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@eadmet.com](mailto:info@eadmet.com).

At step 2 you have the possibility to preview the uploaded file. At this point the file contents is not processed and only the column headers are analyzed. The columns that were correctly recognized will be highlighted in green. The dark-green columns will hold the recognized property or condition names, the light-green columns hold the recognized column names. At this step you can select or unselect the columns for upload. By clicking on the column headers you can remap the unrecognized columns to an appropriate valid column name or property/condition name.

### Step 3 - Value remapping

Database entities remapping page

Property: [LogKoc\\_Tutorial](#)

Values  
Unit: [Log unit](#), min value: 0.42, max value: 5.31

Article: [A17025](#)

Molecule set: [72022\\_batch-sample-sci.xls](#)  
**Notice: Molecule set does not exist and will be created.**

Total rows in sheet: 162

At step 3 you have a basic preview of the data file contents. You can see a list of properties, conditions, units, articles and baskets referenced in the file. If the units, for example, are not explicitly specified in the file, the default unit for the property will be shown. At this step you can also remap the unit, if the default value does not fit the uploaded data. You can also select an existing basket to fill with newly uploaded records.

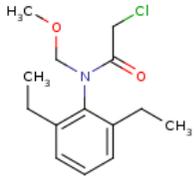
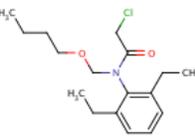
## Step 4 - Batch upload preview browser

**Batch upload preview browser**

**Summary:**  
All rows in the sheet Count: **162**  
Type: internal\_duplicate Status: preprocessed Count: **1**  
Type: valid Status: preprocessed Count: **161**

Filter by row number:  and row type:  Batch operations

1 - 5 of 162  items on page  of 33 > >>

<b>Row 2</b> <input checked="" type="radio"/> Save <input type="radio"/> Skip	 <chem>CCOC(=O)N(C)C1=CC=C(C)C=C1Cl</chem>	<b>LogKOC_Tutorial = 2.28 (in Log unit)</b> eADMET, Team Tutorial article... N: AUTO_1 Unpublished <b>2012</b> ; 11 (5) 100 - 111 15972-60-8 ; alachlor MoleculeID: M12820 RecordID: R-6590106 novserj 
<b>Row 3</b> <input checked="" type="radio"/> Save <input type="radio"/> Skip	 <chem>CCCCOC(=O)N(C)C1=CC=C(C)C=C1Cl</chem>	<b>LogKOC_Tutorial = 2.86 (in Log unit)</b> eADMET, Team Tutorial article... N: AUTO_2 Unpublished <b>2012</b> ; 11 (5) 100 - 111 23184-66-9 ; butachlor MoleculeID: M3885 RecordID: R-6590107 novserj 

At step 4 you can have a real preview of the uploaded data. No data has been uploaded to the database yet. The summary at the top of the screen shows the number of valid records contained in the uploaded file, the number of internal duplicates (i.e. duplicated within the file), the number of external duplicates (i.e. duplicates of records already existing in the database), erroneous records, etc. You can individually or through **Batch operations** button select or unselect records from further upload.

Once you are satisfied with the data that will be uploaded, click the **Proceed with upload** button. After this point the data will be introduced to the database. You will see the progress meter with time estimates.

  
Saving record 100 of 162 (less than 1 minute remaining)...  
[\[interrupt\]](#)

## Step 5 - Finish

**Batch upload results**

Batch upload is finished. You can review your [162 uploaded records](#) or look at the [detailed report of the upload](#).

**Summary:**  
All rows in the sheet Count: **162**  
Type: internal\_duplicate Status: uploaded Count: **1**  
Type: valid Status: uploaded Count: **161**

Once the data has been uploaded, the resulting report will be presented. You have the option to either download the report of the upload process in Excel format, or to directly jump to the uploaded records in the Experimental property browser.