



Guide to Database Curation and Annotation

[ChemSpider](#) is a free chemistry search engine. It has been built to aggregate and index chemical structures and their associated information into a single searchable repository and to make it available to everybody, at no charge. In order to curate data, upload structures or add associated information, and to download search results and use our embedding tools, you need to be a registered user.

What is Curation?

Curation of the [ChemSpider](#) database refers to the manual annotation and correction of data, either structural information or the nomenclature of chemical entities and the links to publications from which the substances are derived.

There are two ways to help curate data on [ChemSpider](#), these are:

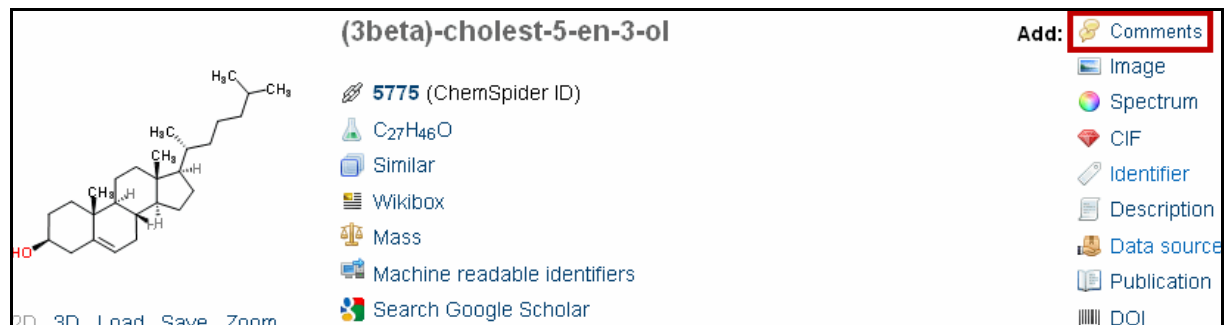
1. Post comments on a record; a Master curator will then investigate your concern(s) and take appropriate action.
2. As a registered user who has requested curation rights; where you participate directly in marking data for master curation or removing erroneous data.

By adding information you make it easier to find records and enrich their content.

By removing errors you can improve the results of [ChemSpider](#) searches. This can benefit both you and all the other users of [ChemSpider](#) by eliminating results that are plainly wrong.

Posting Comments on ChemSpider

Anyone can Post Comments regarding erroneous data. This could be an incorrect name or a structure that is incorrectly drawn. If you believe that you have found an error in a record, you can submit feedback by clicking on the **Comments** button in the top-right corner of the page.



(3beta)-cholest-5-en-3-ol

5775 (ChemSpider ID)

$C_{27}H_{46}O$

Similar

Wikibox

Mass

Machine readable identifiers

Search Google Scholar

Add: Comments

Image

Spectrum

CIF

Identifier

Description

Data source

Publication

DOI

A feedback form is then displayed on top of the record. Simply fill in the text box describing the error (and provide any suggestions of how this should be corrected)

It is important that you supply an e-mail address in order that we may respond but we do respect your privacy and our privacy policy can be found here: <http://www.chemspider.com/Privacy.aspx>

You may wish to assign a higher priority for the error (depending on the severity).

Finally, you need to complete the CAPTCHA and select **Submit**.

Please leave your feedback here... ✕

Thank you for your assistance in curating the data on ChemSpider. We welcome your feedback. Please add any comments you have regarding observed errors in any of the associated text, properties or chemical structures.

I believe that the registry number indicated in this record does not relate to this compound. Please investigate

E-mail: Status: ▼

gamered the

Type the two words:

gamered the

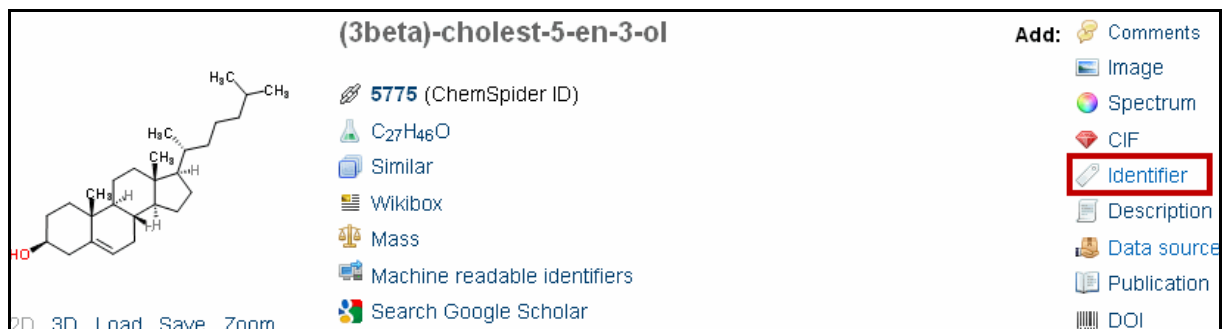
reCAPTCHA™
stop spam.
read books.

Curating Identifiers on ChemSpider

To be able to curate a [ChemSpider](#) record you need to have Curator privileges and be logged in to your account.

Tip: If you need to request Curator privileges or are unsure that you have them you can check/update your profile by clicking on **My ChemSpider** (this is displayed at the top-right hand side of all pages on [ChemSpider](#) when you are logged in).

To change any information in the Identifiers section you need to open the editing dialog. This can be accessed by clicking on the **Identifier** button at the top of the record.



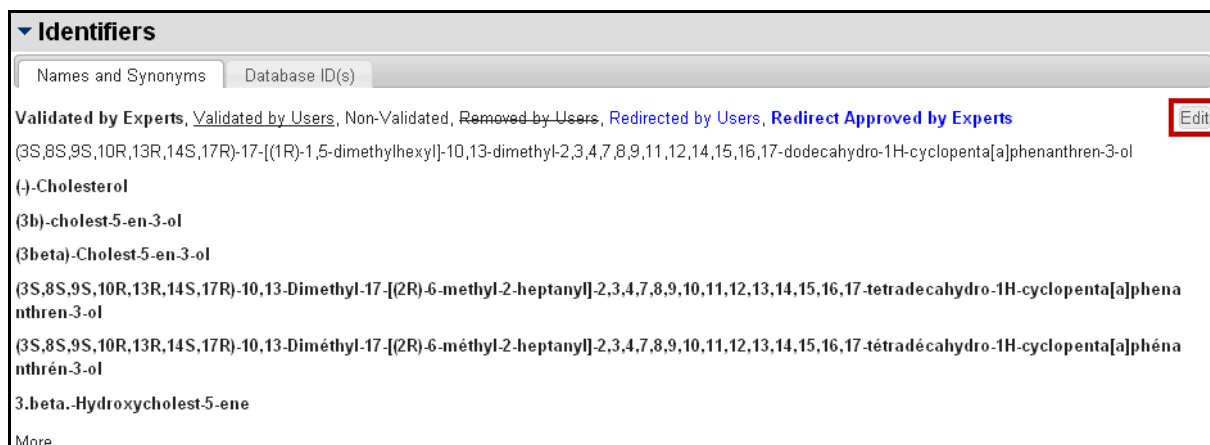
(3beta)-cholest-5-en-3-ol

Add: Comments Image Spectrum CIF **Identifier** Description Data source Publication DOI

5775 (ChemSpider ID) C₂₇H₄₆O Similar Wikibox Mass Machine readable identifiers Search Google Scholar

2D 3D Load Save Zoom

Alternatively, if the record already has some identifiers you can scroll down to the Identifiers section and click on the **Edit** button.



Identifiers

Names and Synonyms Database ID(s)

Validated by Experts, Validated by Users, Non-Validated, Removed by Users, Redirected by Users, **Redirect Approved by Experts** **Edit**

(3S,8S,9S,10R,13R,14S,17R)-17-[(1R)-1,5-dimethylhexyl]-10,13-dimethyl-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-3-ol

(-)-Cholesterol

(3b)-cholest-5-en-3-ol

(3beta)-Cholest-5-en-3-ol

(3S,8S,9S,10R,13R,14S,17R)-10,13-Dimethyl-17-[(2R)-6-methyl-2-heptanyl]-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-ol

(3S,8S,9S,10R,13R,14S,17R)-10,13-Diméthyl-17-[(2R)-6-méthyl-2-heptanyl]-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tétradécahydro-1H-cyclopenta[a]phénanthrén-3-ol

3.beta.-Hydroxycholest-5-ene

More

Overview of the Identifiers editing dialog

The editing dialog displays a list of the identifiers in the record together with a set of check boxes that allow you to select the identifiers that you want to change. As a Curator it is not possible to make changes to validated names (indicated in **bold face**).

At the top of the dialog you can see:

- Selection tools – These allow you to; Select All, Deselect All or Invert the Current Selection.
- Update button – This is used to choose how the state of the selected identifiers should be changed Giving you 4 options: Reject, Normal, Confirm and Redirect
- Add button – This is used to add new identifiers

Names and Synonyms

Selection Tools: Select all, Deselect all, Invert selection, Update, Add

Validated by Experts, Validated by Users, Non-Validated, Removed by Users, Redirected by Users, Redirect Approved by Experts

Validated names cannot be altered by Curators

(-)-Cholesterol
 (3b)-cholest-5-en-3-ol
 (3beta)-Cholest-5-en-3-ol
 (3bêta)-Cholest-5-én-3-ol
 3.beta.-Hydroxycholest-5-ene
 3b-Hydroxy-5-cholestene
 5:6-Cholesten-3.beta.-ol
 5:6-Cholesten-3b-ol
 57-88-5[RN]
 200-353-2[EINECS/ELINCS]
 Cholesterine[Wiki]
 Cholesterol[Wiki]

Save Cancel

Below these there is a key which reminds you of the different styles of text formatting that are used to indicate the current status of an identifier.

In the main body of the dialog you can see:

The Identifiers are grouped with Validated identifiers appearing at the top of the list, followed by Normal identifiers and Rejected identifiers appearing at the bottom of the list. Within these groupings the individual identifiers are listed in alphabetical order.

At the very bottom of the dialog you can see:

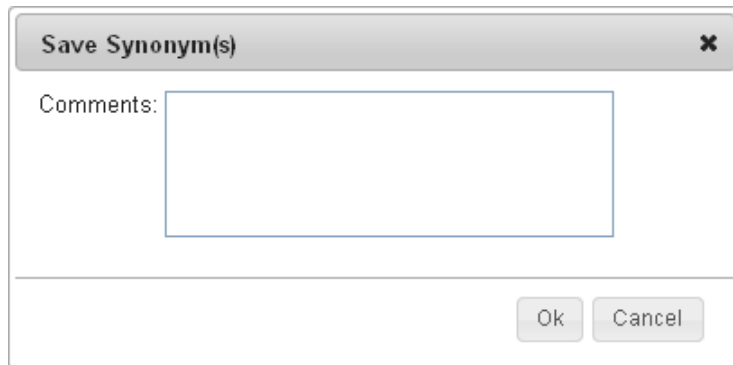
Buttons to **Save** your changes or **Cancel** them.

Adding an Identifier

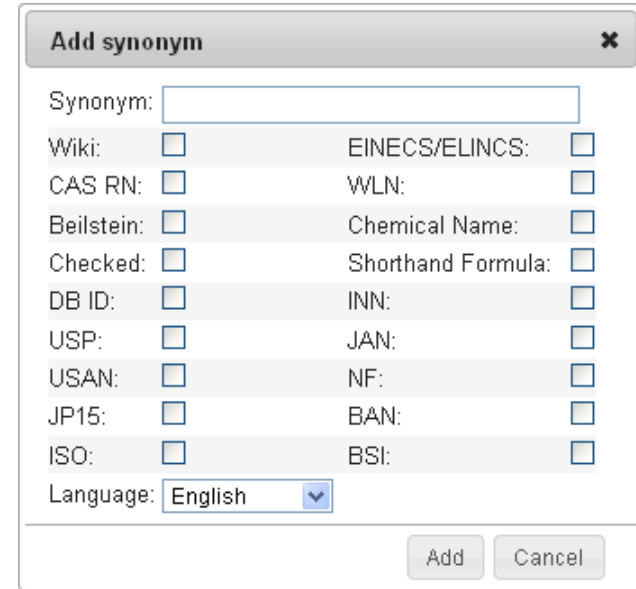
- To add a new identifier to a record open the Identifiers editing dialog (described in the previous section), and click on the **Add** button.
- The identifier entry box will appear, please type (or paste) the identifier into the Synonym field and select any appropriate check boxes that further define the nature of the identifier.

Note: Certain characters are will not be displayed correctly and should not be used (for instance, Greek characters generated by using the Symbol font face).

- When you have finished, click on the **Add** button, this will return you to the identifiers editing dialog. You should now be able to see the identifier that you added.
- To save your changes, click on the **Save** button, which will close the identifiers editing dialog.
- A comment box will pop up for the further addition of comments which would be helpful to a master curator when reviewing the suggested identifiers.
- Click **Ok** when complete, this will return you to the main record and will send an e-mail to the Master Curators for review and approval.



The 'Save Synonym(s)' dialog box features a title bar with a close button (X). Below the title bar is a text area labeled 'Comments:' with a large empty box for input. At the bottom of the dialog are two buttons: 'Ok' and 'Cancel'.



The 'Add synonym' dialog box has a title bar with a close button (X). It contains a text input field for 'Synonym:'. Below this are two columns of checkboxes for various identifiers: Wiki, CAS RN, Beilstein, Checked, DB ID, USP, USAN, JP15, ISO, EINECS/ELINCS, WLN, Chemical Name, Shorthand Formula, INN, JAN, NF, BAN, and BSI. At the bottom left is a 'Language:' dropdown menu currently set to 'English'. At the bottom right are 'Add' and 'Cancel' buttons.

Guide to the options for adding synonyms

- Wiki: This name is on Wikipedia and when linked in this way will open up Wikipedia (please check the name IS on Wikipedia as typed)
- EINES/ELINCS number. (<http://en.wikipedia.org/wiki/EINECS>)
- Chemical Abstracts registry number. (http://en.wikipedia.org/wiki/CAS_number)
- Beilstein indicates a Beilstein number
- Chemical Name is an invisible field but is useful for internal purposes
- DB ID is a database identifier and is displayed in a separate area of the record. Examples include; NCI numbers, PubChem IDs, etc.
- If the name entered is in a foreign language then the drop-down menu can be used to select the language

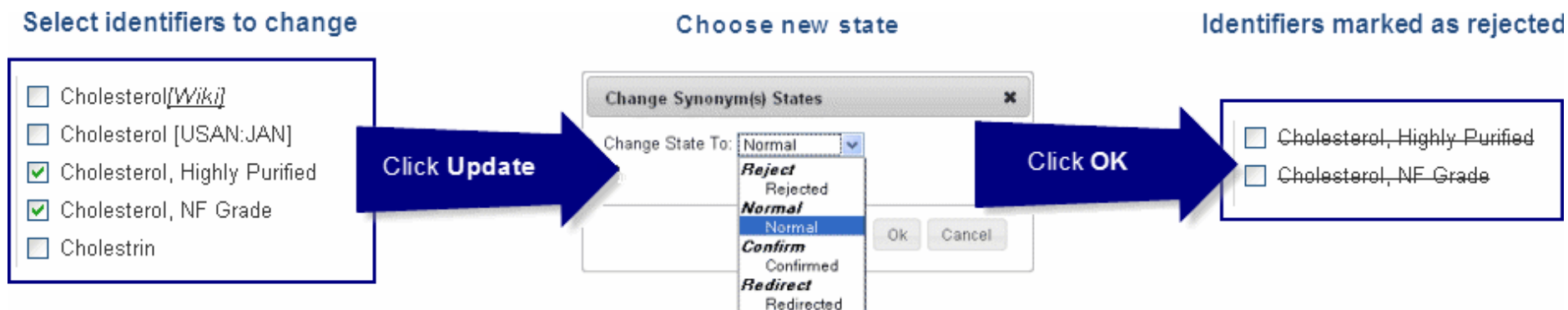
For a complete list see the ChemSpider curation manual.

Rejecting and Approving Identifiers

The process of rejecting or approving identifiers involves the selection of one or more identifiers and then specifying what the new state of these identifiers should be. A common selection is to reject an identifier since it is incorrectly associated with the structure.

There are four states:

- Rejected – displays the identifier with a strike out line to indicate that it is not match the compound
- Normal – returns a name to a normal state
- Confirm – confirms that the identifier is appropriately matched with the compound
- Redirect – offers the user to associate the compound with another ChemSpider ID (for example, when there are two tautomers or isomers that the users would like to connect)



Rejecting or approving an identifier is performed in similar way to adding an Identifier:

- Open the Identifiers editing dialog.
- Select the identifier(s) that you wish to change.
- Click on the **Update** button, this brings up a dialog box which allows you to select the state that you want to apply to the selected identifiers.
- Then click **Ok**. This returns you to the Identifiers editing dialog - the altered identifiers will have changed their position in the list of identifiers (approved go to the top of the list, rejected to the bottom) and will be formatted to display their new state.
- Click **Save** to apply the changes. You will be prompted to supply comments that explain your changes to the Master Curator when reviewing your curation.

State changes can be done on groups of identifiers at one time.

But it is necessary to separately approve or reject in separate operations.

It is not necessary to save the state changes between these operations.

Guidelines for Removal and Approval of Identifiers

What we are trying to achieve with the actions of approval or rejection of identifiers is state changes which will assist the master curators in speeding up the process of database cleansing.

Master Curators have the responsibility of moving curated identifiers to a final approval state of Confirmed or Deleted identifiers based on further research work, reversing the changes or leaving in their present state.

The intention is to remove the associations between structures and identifiers that cause confusion, mislead chemists in their understanding of the chemical structure and provide clarification.

There are various confusions requiring clarification. Specifically:

- All systematic names should match the structure as drawn. All stereochemistry in the name must be represented in the structure shown.
- Any systematic name should be adequate enough to unambiguously convert the name to the matching structure.
- CAS Registry numbers must be for the compound as shown. If the compound shown is the neutral base compound then the registry numbers should not be for the sodium salt or the chloride salt for example.
- Identifiers are not meant to be descriptors per se. For example, an identifier of "One of a series of hexamethylcyclohexanes" is not a good identifier.
- Duplicates can be subtly different but do need to be curated.