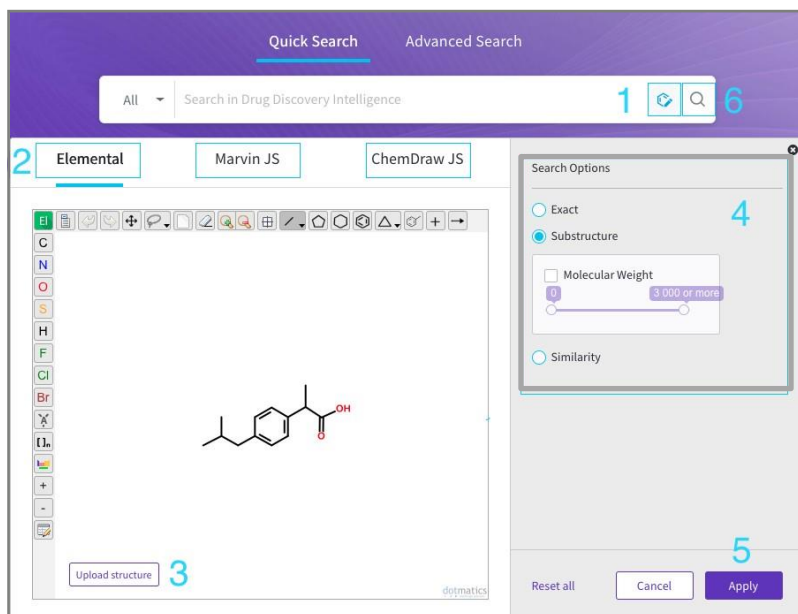# Using Structure Search to identify chemical compounds

## Cortellis Drug Discovery Intelligence

Do you need to find compounds that contain your exact structure, are structurally similar or contain a substructure from your target molecule?

Structure search enables you to identify compounds that contain your exact molecule, are similar to your target molecule or that contain a portion (substructure). Similarity searching is different from substructure searching in that it relies on molecular descriptors, fingerprints, and algorithms (more details can be found in the Appendix). It is complementary to substructure and exact structure searches because it returns answers not found with these chemical structure search strategies. Example Scenario: You are a researcher working on pain disorders and want to review the landscape of compounds obtained by similarity searching (or that contain a substructure core) on your candidate molecule within *Cortellis Drug Discovery Intelligence*.


Executing a Structure Search within Quick Search

From the **homepage**, click the **Structure Search** icon [1] within the Quick Search box. Select the preferred Structure Search editor (Elemental (default), Marvin JS, or ChemDraw JS) [2]. Draw the structure you wish to query using the structure drawing tools. Alternatively, you can upload a mol file by clicking the **Upload** structure button. [3] Select the type of structure search you want to run by selecting the **Exact, Substructure** (and optionally select a molecular weight range) or **Similarity** (and select a % range of similarity) button [4].
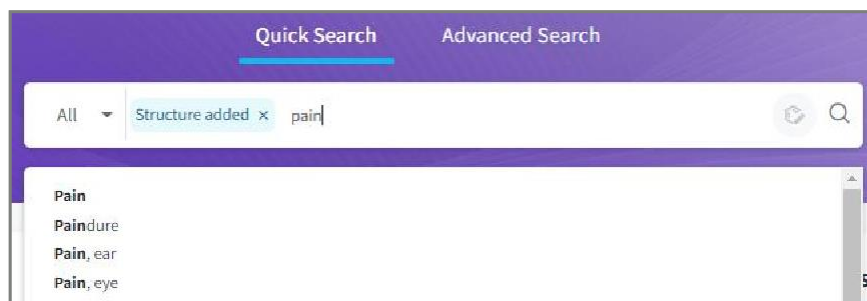
When your structure is complete in the structure editor, click **Apply** [5] to add your structure query to the Quick search box, where you can directly **run the query** [6] or combine with text-based terms.

Executing a Structure Search within Advanced Search
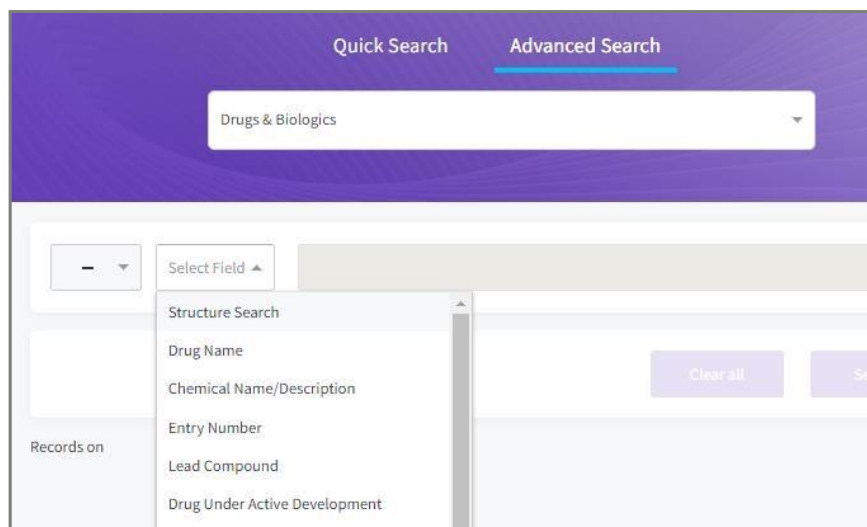
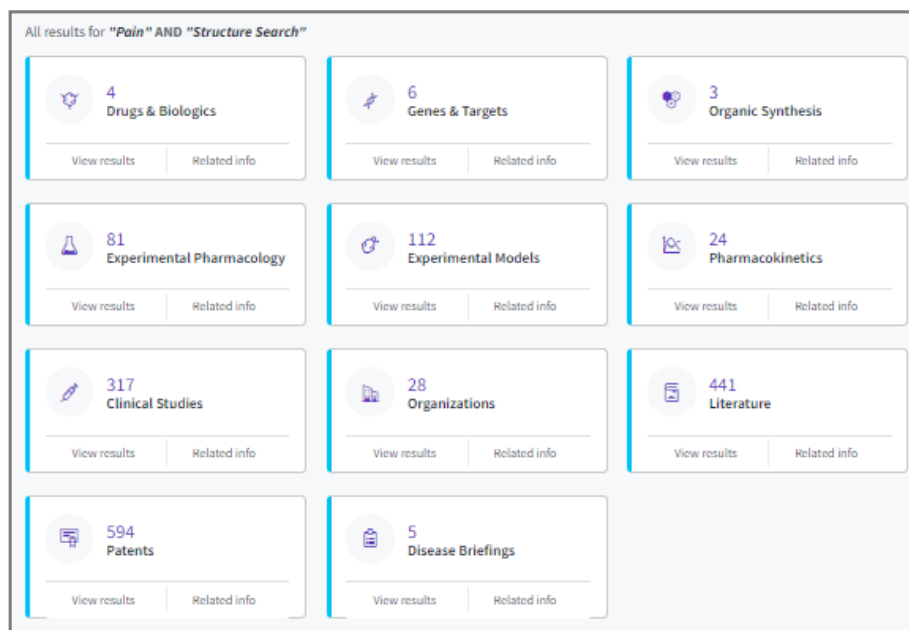From the **homepage**, click on **Advanced Search** [1].

Select the knowledge area of interest from the **drop-down menu** [2].

Select **Structure Search** [4] within the **Select Field** menu [3] to open the structure editors. Once the structure query is added (see previous page on how to do so), you may select additional fields of relevance for your **Advanced Search** and hit **Search** to run the query.
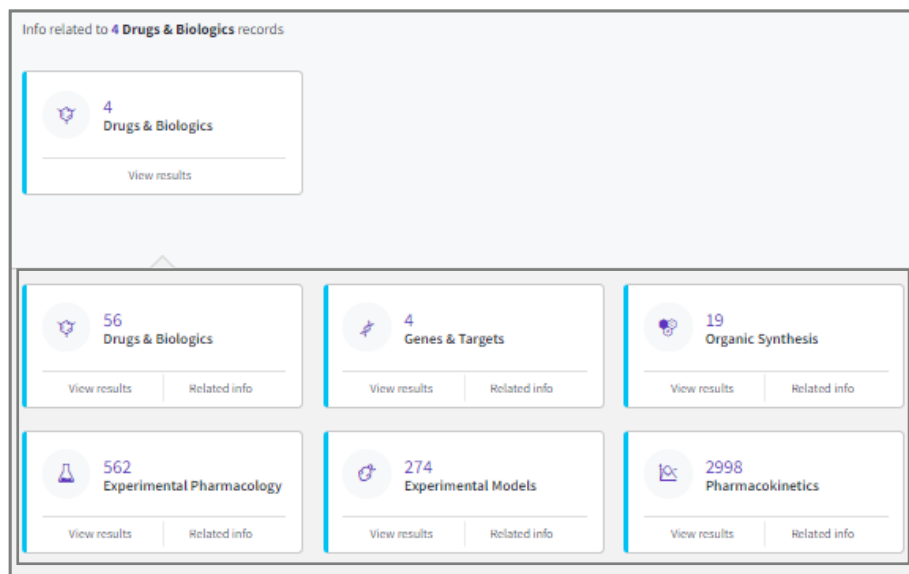
Tips -

- If you are using the structure search feature for the first time, please see the **Appendix** for further details about the structure editors. Note that ChemDraw JS will be added by the end of 2019.

- An "Exact" search will identify compounds containing the exact structure you have drawn. A "Substructure" search allows you find compounds containing the same substructure core and optionally, that fall within the selected range of molecular weight. Finally, a "Similarity" search allows you to find compounds that are similar to candidate molecule. The default similarity threshold value is 80%. i.e., retrieve structures that are at least 80% similar to the query; this can be manually adjusted by the user
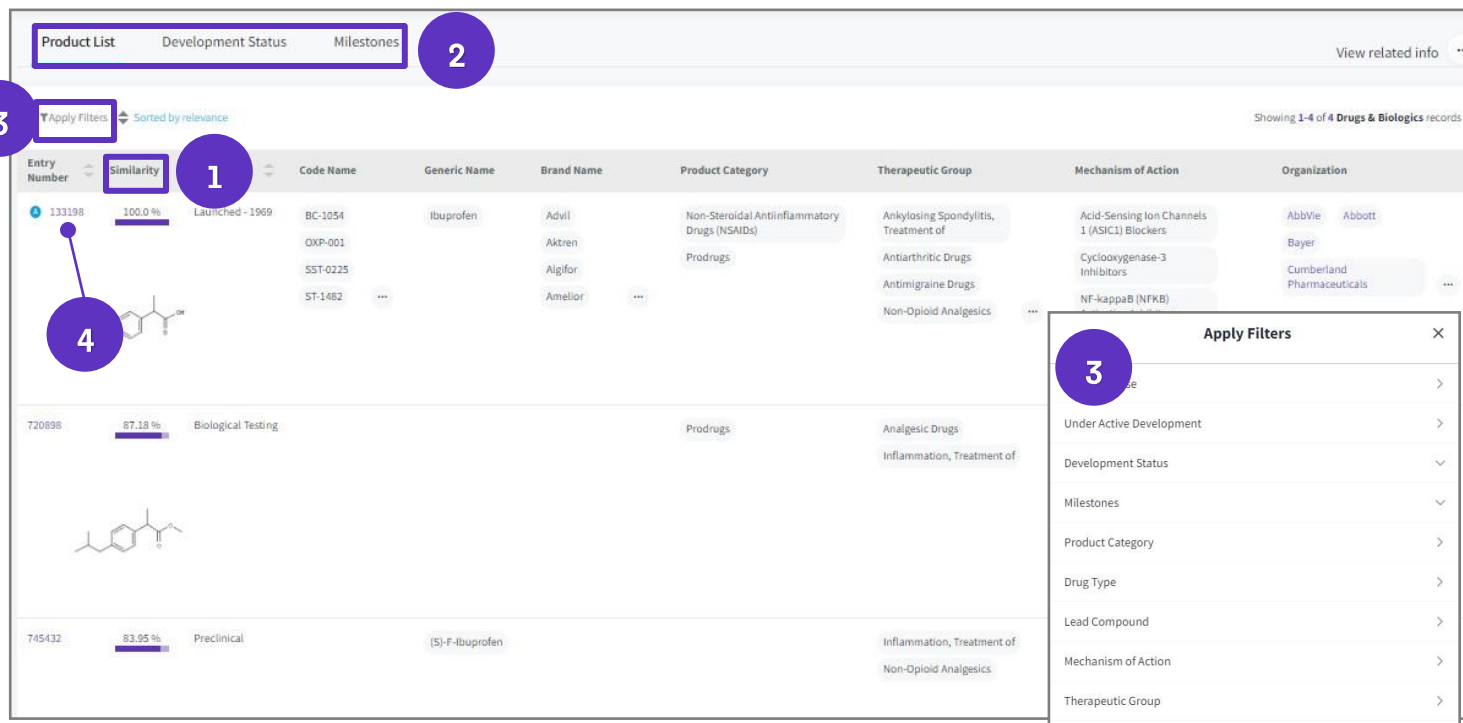
![Clarivate logo]

Working with your data

All results for *"Pain"* AND *"Structure Search"*

4 Drugs & Biologics	**6** Genes & Targets	**3** Organic Synthesis
View results / Related info	View results / Related info	View results / Related info
81 Experimental Pharmacology	**112** Experimental Models	**24** Pharmacokinetics
View results / Related info	View results / Related info	View results / Related info
317 Clinical Studies	**28** Organizations	**441** Literature
View results / Related info	View results / Related info	View results / Related info
594 Patents	**5** Disease Briefings	
View results / Related info	View results / Related info	

Having run the query (example usedis a Quick Search including a Structure similarity search (>80%) of the compound of interest AND the term **Pain**), results are found across all the knowledge areas in Cortellis Drug Discovery Intelligence.

To view the retrieved compounds, click on the **Drug & Biologics** card by either clicking the number of results found or the **View results** button.

Info related to **4 Drugs & Biologics** records

4 Drugs & Biologics — View results

56 Drugs & Biologics	**4** Genes & Targets	**19** Organic Synthesis
View results / Related info	View results / Related info	View results / Related info
562 Experimental Pharmacology	**274** Experimental Models	**2998** Pharmacokinetics
View results / Related info	View results / Related info	View results / Related info

Alternatively, to identify all related content to the compounds identified, click on the **Related Info** button on thecard.

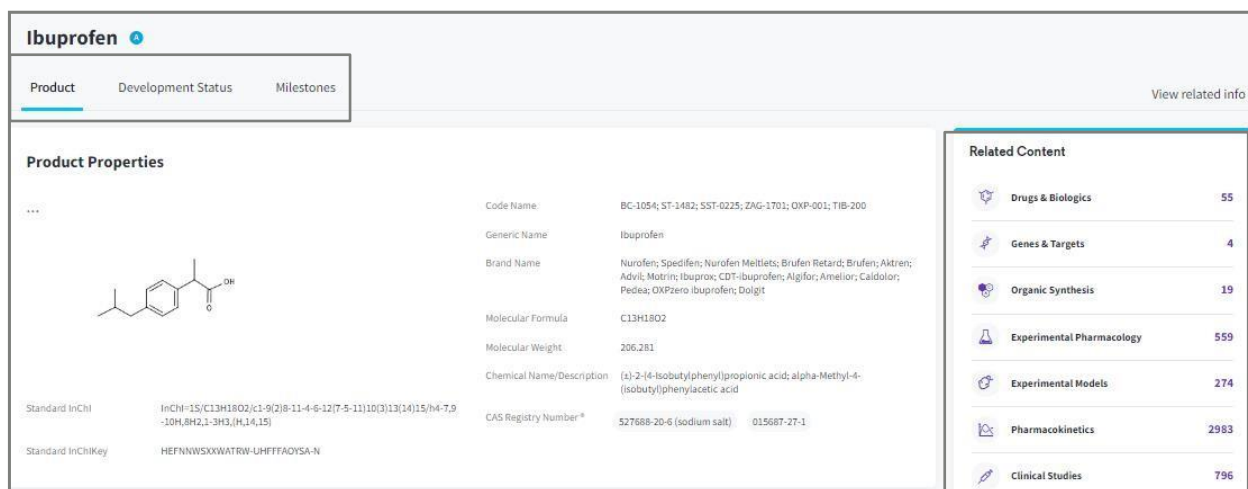The related information is immediately displayed in the bottom half of the screen.

Results are ordered by **% similarity** (for a structure similarity search) or by **relevance** to the query for an exact or substructure search [1]. To view the **Development status** information or **Milestones** data for the compounds of interest, click on their respective tabs [2]. To refine your results further, click on **Apply Filters** [3] and choose among the many options available for filtering. To know more about a specific compound, click on the **Entry Number** to open the drug entitypage [4].

The drug entity page shows information on the drug including a **product summary**, product properties, development status summary and **regulatory information.** For more detailed Development Status and Milestones on the drug, click on their respective tabs.
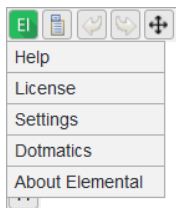
On the right side you will find easy access to the related content in other knowledge areas for your record.

Appendix

Structure editors

For further Information on the structure editors available in *Cortellis Drug Discovery Intelligence*, please refer to thefollowing information and user guides:

- Elemental: **https://www.dotmatics.com/products/elemental**

Furthermore, the Help menu within the Elemental editor opens a pdf file that shows the main drawing features.

- Marvin JS: **https://docs.chemaxon.com/display/docs/Marvin+JS+User%27s+Guide**
- ChemDraw JS: **http://media.cambridgesoft.com/support/manuals/16/ChemDrawHelp.pdf**

Similarity searching - molecular descriptors, fingerprints, and algorithms

Similarity searching finds molecules that are similar to the query structure. The calculation applies the Tanimoto coefficientthat has two arguments:

- the fingerprint (a bit string that contains structural information on the molecule) of the query structure
- the fingerprint of the molecule in the database

The Tanimoto coefficient is calculated by the following formula:

NA&B/(NA+NB-NA&B)

where NA and NB are the number of bits set in molecule A and B, respectively, NA&B is the number of bits that are set inboth.

The similarity threshold specifies a lower limit for the Tanimoto coefficients. If a Tanimoto value is greater than the threshold, then the query structure and the given database structure are considered similar.

The hits returned will be sorted in the decreasing order of the similarity between the query and the found

structure.Similarity Threshold: The value must be between 0 and 100%.

A greater threshold results in fewer hits that are more similar to the query structure.

For more information contact Customer Service at **LS Product Support**