Finding Property Information in Reaxys

Learning Objectives

- Understand the definition of a property "data point" in Reaxys (excerpted information from original documents)
- Navigate the structure of Substance Records in Reaxys, which include property information
- Explore opportunities to link substances to specific property information and vice versa
- Try different ways to search for this type of information

Background

Most chemical databases extract bibliographic information, titles and abstracts from original documents. Many enhance the resulting database of bibliographic/document records with indexing terms that facilitate searching and finding information.

Some databases additionally extract substances and reactions from source documents and organize that information into records that can be searched and browsed. As part of its historical legacy, Reaxys also extracts millions of property facts from publications and organizes them into corresponding Substance Records. Because of this organization, you can easily explore the reported properties of a substance or find substances that have a specific property.
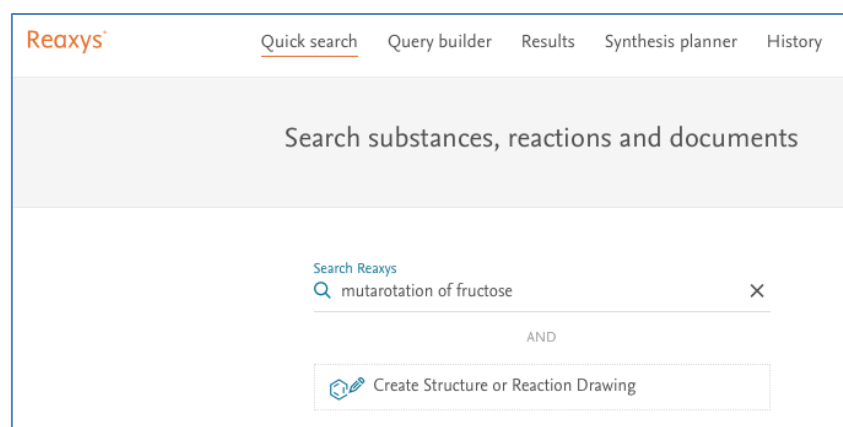
A Reaxys Data Point

Data points in Reaxys can be numeric or textual. Important is that they are information bits professionally excerpted from source documents (e.g., journal articles, patents) and organized into the information structure of the database. This makes those data points findable and usable.

Let's find a data point.

In the new Reaxys, enter into the **Search Reaxys** box:

mutarotation of fructose

and click **Search**.

A results screen appears that looks like this:

Reaxys interprets the query you performed and delivers different hit sets based on different interpretations. Take a look at the different sets. Can you explain how Reaxys found each set?

Click **View Results** for the **Substances** (likely the top hit set).

The hit set consists of different substance records. Click the arrow next to *mutarotation* under **Hit Data**. A table appears.

Eight columns display a data point in each cell. Together, they constitute information about mutarotation values reported for that substance. Most of these data points are searchable, so we can find very specific information about substances directly in Reaxys.

Searching property information

Combine structure and text queries with **Quick Search** in Reaxys

In the example above, we use the name *fructose* to indicate to Reaxys what substance we were interested in.

Suppose, however, that we need to find mutarotation data for all substances that contain the fructose substructure.

Quick search in Reaxys allows you to combine structure and text searches. Try doing the search.

When you get your answers (you should get 41) click **Export** at the top of the answer screen to order a report (include All Substances, select **Hit Data only** and **Include structures**). You can download the report and email it to colleagues.

What properties does Search Reaxys include?

Reaxys includes and recognizes over 500 properties. A search may not always be as straightforward because there is often more than one way of naming a property.

It is helpful to explore a bit what is contained in Reaxys.

Try some of these examples

- Find reported melting points of carbamazepine.
- Find information on the fluorescence spectrum of anthracene.
- What has been reported on NOESY NMR spectra for testosterone?
- What can you find on the cyclovoltammetry of tetraphenylporphyrin
- Search for cyclovoltammetry data of substances with the tetraphenylporphyrin substructure (Hint: When drawing the structure in the structure editor, take advantage of the **Create structure template from name** function and make sure to tell Reaxys to search tetraphenylporphyrin **As substructure**).
- What is the density of barium sulfate?
- Find data on the electrical conductivity of copper.
- What information does Reaxys contain on quantum chemical calculations for pyrroles (Hint: you don't need to do a substructure search for pyrroles; just do a text search).
- Find conformer equilibrium data for biphenyls.

But how smart is Reaxys for finding property information?

The **Quick search** options in Reaxys are quite good at recognizing properties. There are around 130 substance property fields in Reaxys, most of which have subfields of information that are also searchable. Some of the data in these fields and subfields are numeric and some are textual. To search the data, it helps to know what field or subfield to query.

To help you visualize this difference, use **Quick search** to find information on:

critical superconductivity temperature over +40 degrees Celsius

Simply type the above into the **Search Reaxys** box.

TRY IT NOW BEFORE YOU READ ON!

Reaxys retrieves almost 20,000 substances that have critical superconductivity temperature data. For this search, Reaxys was not able to interpret *over +40 degrees Celsius*.

Looking at the first record (hydrogen) of the Substances hit set (click **View Results**) the first Hit Data is:

We see that a reported temperature range is recorded in the subfield **Critical Superconductivity Temperature, °C** and that this subfield is subordinate of the field **Electrical Data**.

To search for specific values in this subfield, we must tell Reaxys in which field and/or subfield we are interested. This is done through **Query builder** (in the top navigation bar of Reaxys).

Building a Query

Click **Query builder**. The **Search properties** box (top right) allows you to search for relevant fields. Enter *electric* and the field **Electrical Data** appears in the list. Drag this field to the main working screen, and click **Show fields**. The screen now shows the subfields that are subordinate to the field Electrical Data:

Click the **Lookup** icon on the very right in the row **Description (Electrical Data)**. This generates a list of all terms that can appear in the Description Subfield:

All these terms are recognized by Reaxys in a text search (Note: mouse over a term to see the full name) and this is how Reaxys pulled up the set of 20,000 substances that have reported data on critical superconductivity temperature.

To get a hit set of substances that have reported critical superconductivity temperature *over +40 degrees Celsius,* we just need to enter that search criterion into the subfield that contains those data.

Under the subfield **Critical Superconductivity Temperature, °C** enter *40* and change the corresponding operator to **>=**.

Click **Search**, then **Substances**. Take a look at the retrieved substances.

Try these examples

- Find NMR spectra for substances containing ^{95}Mo
    - Hint: through **Query builder**, go from the **NMR Spectroscopy** field to the **Nucleus** subfield. Click the **Lookup** icon and enter *95Mo* in the search box.
- We are interested in substances for which various Magnetic Data have been excerpted and added to Reaxys. Locate the terms in the Description Subfield (Magnetic Data), then, in turn, choose a couple of the terms and search them. The point of this exercise is to help you become familiar with the general types of data points that Reaxys has excerpted for this property.   (The teaching point here is that we can't know all the types of information in the over 500 million data points in Reaxys – so at times we may want to perform a general search to see what is present and then, having found the information we want, construct a more targeted search accordingly).

Summary

Reaxys contains extensive property information for substances, and this module gives a simple introduction to some of the content and search opportunities.

At times, we may want property information on a substance. In this case, it may be sufficient to find the substance and look through the property fields available.

Other times, we may want to search for specific property information. In this case, we have two possible workflows:

A. Use **Quick search** – here it may help to understand the different property terms in the Description subfields
B. Use **Query builder** – here we may investigate in depth the many opportunities available through the fields and subfields, with numeric and text data points.