accelrys®

JDRAW

Rgroup Query User Guide

Copyright Notice

Trademarks

The registered trademarks or trademarks of Accelrys Software, Inc. include but are not limited to:

ACCELRYS®
ACCELRYS & Logo®
JDRAW

All other trademarks are the property of their respective owners.

Restrictions on Government Use

This is a “commercial” product. Use, release, duplication, or disclosure by United States Government agencies is subject to restrictions set forth in DFARS 252.227-7013 or FAR 52.227-19, as applicable, and any successor rules and regulations.

Acknowledgments and References

To print photographs or files of computational results (figures and/or data) obtained using Accelrys software, acknowledge the source in an appropriate format. For example:

> Imaging results obtained using software programs from Accelrys Software, Inc. Data management and analysis performed with the Pipeline Pilot Imaging collection. Graphical displays generated with the Discovery Studio Visualizer.

To reference an Accelrys Software, Inc. publication in another publication, Accelrys Software, Inc. is the author and the publisher. For example:

> Accelrys Software, Inc., *Chemistry Collection: Basic Chemistry User Guide, Pipeline Pilot, San Diego*: Accelrys Software, Inc., 2010.

Request for Permission to Reprint and Acknowledgment

Accelrys may grant permission to republish or reprint its copyrighted materials. Requests should be submitted to Accelrys Scientific and Technical Support, either through email to support@accelrys.com or in writing to:

Accelrys Scientific and Technical Support
10188 Telesis Court
Suite 100
San Diego, CA 92121

Please include an acknowledgment “Reprinted with permission from Accelrys Software, Inc., [*Document name*], [*Month Year*], Accelrys Software, Inc., San Diego.” For example:

> Reprinted with permission from Accelrys Software, Inc., Rgroup Query User Guide *JDraw*: *Rgroup Query* User Guide, *December*, *2010*, Accelrys Software, Inc., San Diego.

Contents

Chapter 1:
Introduction

About Accelrys JDraw

The Accelrys JDraw Java applet gives application developers a lightweight, no-fee structure editor for the development of custom web and rich client applications that require chemical structure queries and/or chemical structure rendering functionality.

In addition to accessing the no-fee Accelrys Draw desktop application, students, teachers and researchers working in a non-commercial setting can now access capabilities for rapidly drawing chemical structures and chemically intelligent queries in Web applications–capabilities built on industry-leading Accelrys chemistry representation technology.

Accelrys JDraw applet includes the following key features:

- Loads quickly in web based applications.
- Provides access to Accelrys' best in class chemical representation.
- Includes components that can use Accelrys' query features to refine hit sets.
- Can be used as a full editor and renderer, or in a render only mode.
- Allows editing of structure in the web-page; you do not have to wait for a separate editing application to start up.
- Supports Clipboard operations with major chemical structure drawing applications, including Accelrys Draw, ISIS/Draw, and ChemDraw.
- Has a familiar look-and-feel that allows Accelrys Draw and ISIS/Draw users to immediately use the functionality without dedicated training.
- Does not require installation; deployment on web server as JAR files that a customer-developer can include in a Java applet or Java application. The Accelrys JDraw components might be used for various purposes, such as allowing end-user chemists to draw query structures, or allowing a result set of structures to display in a grid.
- Provides examples that show:
  - How to use JavaScript when including the Accelrys JDraw editor in an applet, and how an applet can render multiple structures in a grid.
  - How a Java application can include the Accelrys JDraw editor, and how a Java application can render multiple structures in a grid.
- Includes an API Reference (javadoc).
- Enhancements to support Sgroups (polymers, mixtures, and formulations) are planned for a future release.

Additional Information

For more information about Accelrys JDraw and other Accelrys software products, visit http://www.accelrys.com.

Chapter 2:
Topics in this Guide

Rgroup Topics

This guide includes many chapters, each providing comprehensive information and details pertaining to the following Rgroup topics:

- Rgroup Structure

An **Rgroup structure** is a structure that consists of one or more members that can be substituted, or prohibited, at labeled sites on the root member or on other Rgroup members. See Chapter 3: Rgroup Structure.

- Rgroup Attachment Point

An **Rgroup attachment** point is the atom in a structural fragment that attaches the fragment to the rest of an Rgroup structure. See Chapter 4: Rgroup Attachment Point.

- Rgroup Member

Rgroup members are the chemical fragments that may be substituted at defined attachment points on the root structure. See Chapter 5: Rgroup Member.

- Rgroup Logic

Accelrys JDraw allows you to add **logic** when using Rgroups in your structures. See Chapter 6: Rgroup Logic.

Related Topics

- Markush Structure

A **Markush structure** (sometimes called a **generic** structure) is a chemical representation of a set of unique compounds by using the most common structural feature, the root structure, and representing substitution or change points in or on the root structures. See Chapter 7: Markush Structure.

- Markush (Rgroup) Query Tool, Root Structure

The **root structure** of a Markush query contains the portion of the query that must be present in all of the structures retrieved. See Chapter 8: Markush (Rgroup) Query Tool, Root Structure.

- Markush (Rgroup) Query Tool

This topic provides details about Rgroups, an example, and a detailed procedure how to draw a Markush query. See Chapter 9: Markush (Rgroup) Query Tool.

- Markush (Rgroup) Query Tool, Example

This topic provides an example and illustration of a Markush query. See Chapter 10: Markush (Rgroup) Query Tool, Example.

- Markush Query Tool

This topic provides simple steps on how to draw a Markush query for users of Accelrys Direct or ISIS/Host Databases. See Chapter 11: Markush Query Tool.

Chapter 3:
Rgroup Structure

Definition of an Rgroup Structure

An **Rgroup structure** is a structure that consists of one or more **members** that can be substituted, or prohibited, at labeled sites on the root member or on other Rgroup members.

There are three categories of Rgroup structure

- Markush structures (see Definition of a Markush Structure), which can be registered to an Accelrys Direct Generic library[1] database.
- Markush queries (see Example and Illustration of a Markush Query), which can be used in substructure search.
- Building blocks[2], which can be registered to a building blocks database.

1 A **library** is a set of chemical structures that are related through a common Markush (Rgroup) structure.

[2] A **building block** is a structural fragment in a building blocks database with one or two attachment points that specify the connectivity of the fragment in a Markush structure.

A **building block database** is a database that contains collections of building blocks. These reusable collections are useful for rapidly specifying Markush structures.

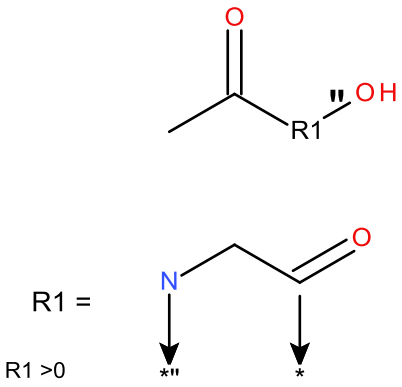
Chapter 4:
Rgroup Attachment Point

Definition of an Rgroup Attachment Point

An **Rgroup attachment point** is the atom in a structural fragment that attaches the fragment to the rest of an Rgroup structure (see Chapter 3: Rgroup Structure).

If an Rgroup member contains two attachment points, you can specify one as a **primary attachment point** and the other as a **secondary attachment point**. If there are two attachment points on an Rgroup member, there must be two corresponding attachments (bonds) to the Rgroup atom that has the same Rgroup number. The bond from the Rgroup atom to the primary attachment point is marked with a quote (') and the bond to the secondary attachment point is marked with a double quote (").

The following figure shows a Markush structure (see Chapter 7: Markush Structure) with an Rgroup atom that has two attachments. The double quote (") on the bond from R1 to the hydroxy group is the secondary attachment point:



Rgroup attachment point

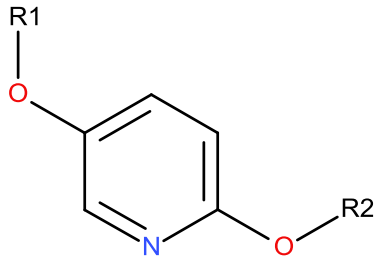
Chapter 5:
Rgroup Member

Definition of an Rgroup Member

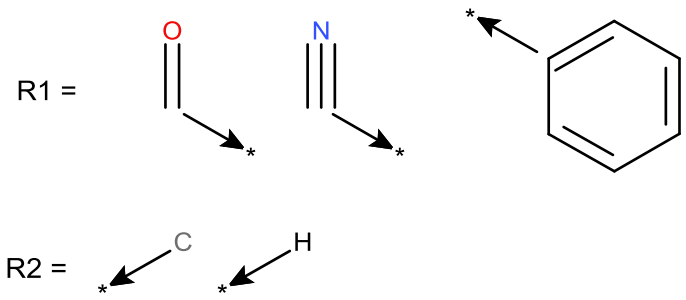
Rgroup members are the chemical fragments that may be substituted at defined attachment points on the root structure (see Chapter 8: Markush (Rgroup) Query Tool, Root Structure). You define the Rgroup members that you will allow to be found at specified positions in the molecules retrieved. Rgroup members can be any structural fragment, including functional groups and single atoms or atom lists.

In the example below, the Rgroup members **CO**, **CN**, or **Ph** are allowed at position 5 (R1) on the root structure. The Rgroup members **H** or **CH_3** are allowed at position 2 (R2) on the root structure.

Root Structure



Rgroup Members



Rgroup member

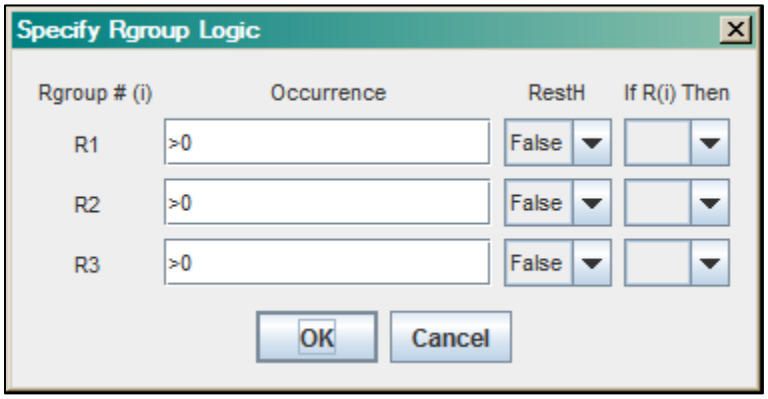
Chapter 6:
Rgroup Logic

Working with Rgroup Logic

JDraw allows you to add logic when using Rgroups in your structures. Once you have drawn your structure and identified the location of each Rgroup, then you can choose the logic you want.

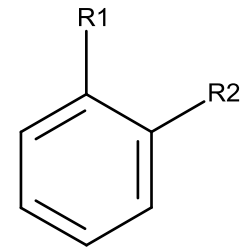
To access the advanced Rgroup logic feature

1. Draw your structure with the All Purpose Drawing tool.
2. Define the Rgroups (see Definition of a Root Structure).
3. Right-click the R1= (or R2=, Rn=, etc.) on the canvas.
4. Select **Advanced Rgroup Logic**.
   - You see the Rgroup Logic dialog.

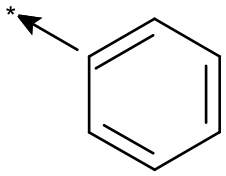


Rgroup Logic dialog

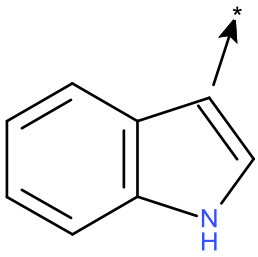
5. Edit **Occurrence/RestH** group box defines how many of a Rgroup occurs and assigns a value of **True** or **False**.
6. Edit **If R(i) Then** logic group box controls specify logical conditions. Use the Rgroup condition **If R(i) Then** to specify whether the presence of one Rgroup is dependent on the presence of another Rgroup. For example:



R1 =



R2 =



IF R1 >0
  THEN R2 >0 (RestH)
IF R2 >0 (RestH)
  THEN R1 >0

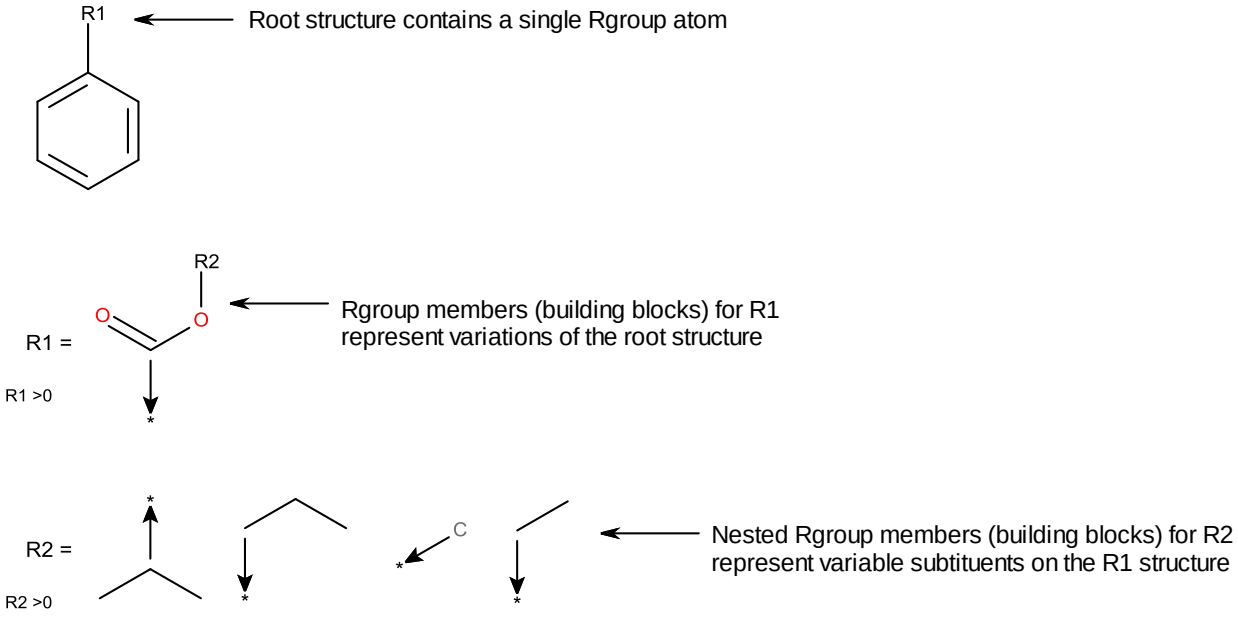
Rgroup logic

Chapter 7:
Markush Structure

Definition of a Markush Structure

A **Markush structure** (sometimes called a **generic** structure) is a chemical representation of a set of unique compounds by using the most common structural feature – the root structure[3] – and representing substitution or change points in or on the root structures. For each change point there is a list of the options, called Rgroup members. Markush structures can be stored in Accelrys Direct Generic databases.

The following figure shows a Markush structure:



Markush structure

Although superficially similar to Markush queries, Markush structures differ in important ways:

- The root structure of a Markush query **must** contain at least one atom that is **not** an Rgroup atom. The root of a Markush structure, however, can consist solely of Rgroup atoms. For example, the root of the Markush structure in the preceding figure contains only one Rgroup atom, and **cannot** be used as a Markush query.
- A Markush query can contain only **one level of nesting**, but a Markush structure can contain Rgroups that are nested to **any number** of levels.
- Unlike a Markush query, a Markush structure **cannot** contain Occurrence, RestH, or If-Then logic.

3 The **root structure** is the portion of a Markush query or Markush structure that contains the portion of the structure that is always present. For a Markush query, the root contains the portion of the query structure that must be present in all structures that you retrieve. For a Markush structure, the root contains the structure that is present in all compounds in the library.

- Each Rgroup member of a Markush query must contain either one or two attachment points, whereas members (building blocks[4]) of a Markush structure can contain zero, one, or two attachment points.
- Markush queries **cannot** be registered to a database. In contrast, generic structures can be registered to library databases, and building block structures can be registered to building blocks **databases**.
- Many Accelrys applications that use Markush structures can create and manipulate the root structure separately from the associated Rgroup members or building blocks. For example, individual Rgroup members within a Markush structure can be registered to a building blocks database.

Tip: To avoid confusion between the different types of Rgroup structures, JDraw supplies a separate palette of tools for drawing generic structures and building blocks. Use the Rgroup Atom tool to add Rgroup atoms to the root of a Markush structure.

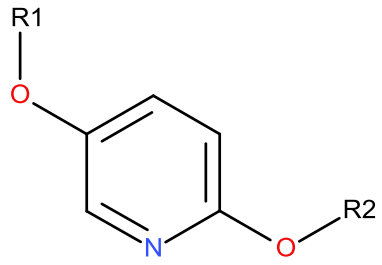
4 A **building block** is a structural fragment in a building blocks database with one or two attachment points that specify the connectivity of the fragment in a Markush structure.

A **building block database** is a database that contains collections of building blocks. These reusable collections are useful for rapidly specifying Markush structures.

Chapter 8:
Markush (Rgroup) Query Tool, Root Structure

Definition of a Root Structure

The **root structure** of a Markush query contains the portion of the query that must be present in all of the structures retrieved. For example, you might want to find 2, 5-dihydroxypyridine in which the substituent at position 2 can be H or CH3, and the substituent at position 5 can be C=O, CN, or a phenyl group. The root structure contains 2, 5-dihydroxypyridine with bonds to **Rgroups** (such as R1 and R2) that specify substituents:



Markush root structure

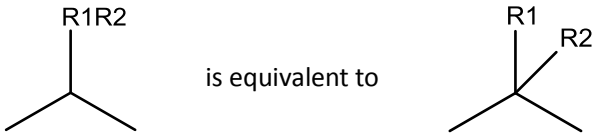
Rgroup Atoms

On the root structure, **Rgroup atoms** (R1 through R32) are placeholders that designate where members of the lists of functional groups and atoms can attach to the root structure. In the example above, the Rgroup R1 designates that the members of the list R1 (the Rgroup substituents) are allowed at position 5 on the root structure (CO, CN, or Ph). The Rgroup R2 designates that the members of the list R2 (the Rgroup substituents) are allowed at position 2 on the root structure (H or CH3).

Multiple Rgroup Atoms at One Position

You can specify **multiple Rgroup atoms** at a single site on the root structure.

For example:

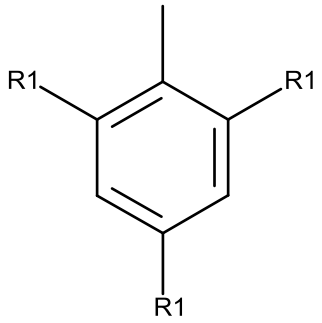


Multiple Rgroup atoms, one position

One Rgroup Atom at Multiple Positions

You can specify a single Rgroup atom (such as R1) at **multiple positions** on the root structure.

For example:



One Rgroup atom, multiple positions

If the root structure contains more than one instance of an Rgroup atom, you can specify the number of times that any member of this Rgroup is to appear in retrieved structures. For additional information, see Working with Rgroup Logic.

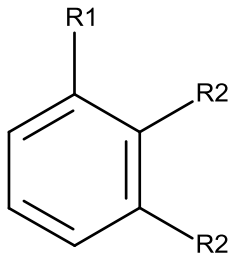
Chapter 9:
Markush (Rgroup) Query Tool

Working with the Rgroup Query Tool

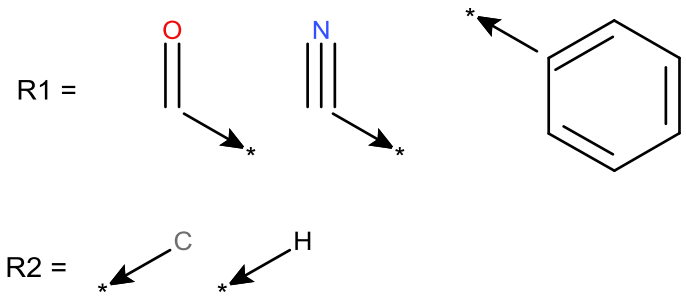
Example and Illustration of a Markush Query

For the Rgroup members, an asterisk and arrow denote an "attachment point", an atom or bond in the Rgroup member that attaches the member to the root structure at the Rgroup atom.

Root Structure



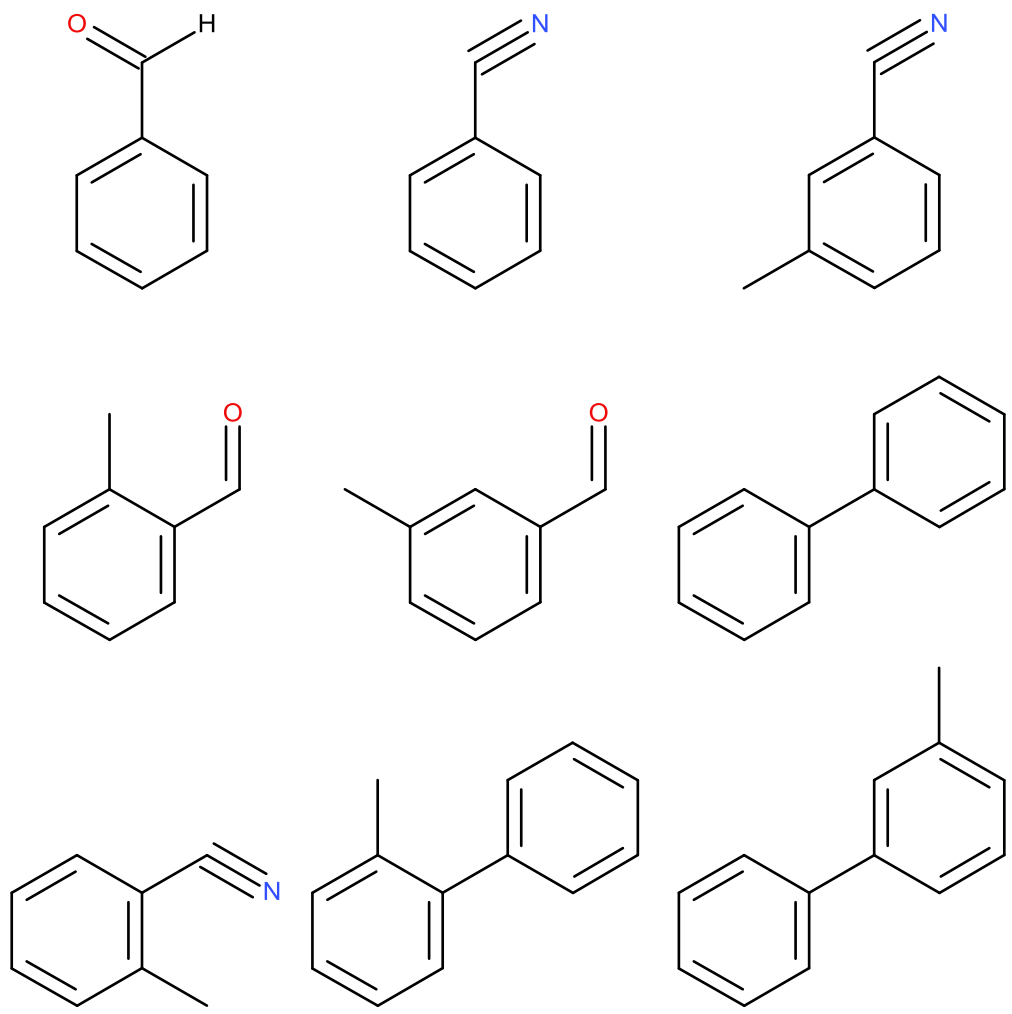
Rgroup Members



For R2, Occurrence = 1

Rgroup attachment point

Some structures that you retrieve with this query:



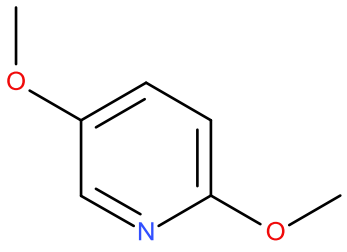
Rgroup attachment point queries

To draw a Markush query, perform the following steps in order

Note: In Rn, n = Rgroup number = 1, 2, etc. A selected item appears in a rectangle with a blue border.

1. **Draw the root structure**. The "root structure" is the part of the query that does not vary. Use the chemical drawing tools to draw the root structure, as you would draw any structural query.

   Example of a root structure:

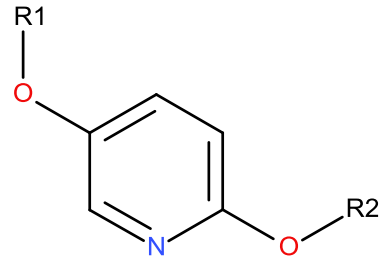


Draw the Rgroup root structure

2. **Add Rgroup atoms to the root structure**. You need to specify where in the query you want to allow Rgroups, and which one to allow at a given site. To add Rgroup atoms (R1, R2, etc.) to the root structure:

- Click the Rgroup Tool .
- Click the site on the root structure where you want the substituent (Rgroup member) to bond.
- Click the Rgroup atom that you want from the palette of values.
- Repeat this step for each Rgroup atom.

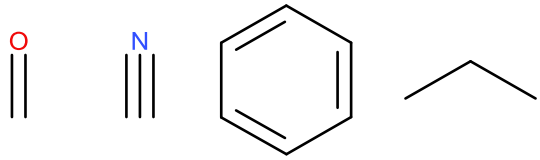
Example of a query in which Rgroups 1 and 2 are allowed at specific sites:



Add Rgroup atoms to the root structure

3. **Draw the Rgroup members** (possible choices for substituents or fragments). The structure editor allows only one possibility of the set of members for a given Rgroup, to bond at a given site on the root structure. Use the chemical drawing tools to draw the Rgroup members, as you would draw any structural query.

   Example of fragments that will become a set of Rgroup members, such as for Rgroup 1:



Rgroup members

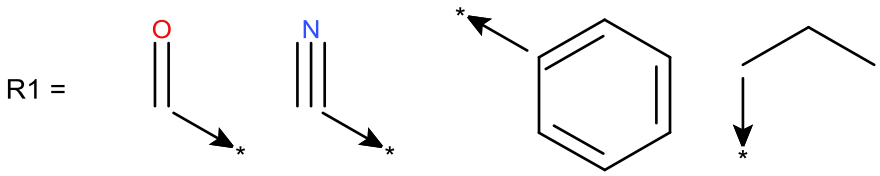
4. **Add each substituent or fragment as a member of the Rgroup**. To associate a fragment with an Rgroup as a member:
   - Click the Markush Query Tool .
   - Click the fragment.
   - Drag and drop the fragment onto the Rgroup definition (Rn =).
   - Repeat this step for each fragment that you want to associate with an Rgroup.

   Example of dragging and dropping a fragment onto an Rgroup definition:

Add Rgroup member onto Rgroup definition

When a fragment becomes a member of an Rgroup, you see an asterisk and an arrow appear at an atom, the "attachment point". The attachment point is the atom of the Rgroup member that bonds to the root structure at the site of the Rgroup atom (**Step 2**). Attachment points are added automatically. You can have up to two attachment points for an Rgroup.

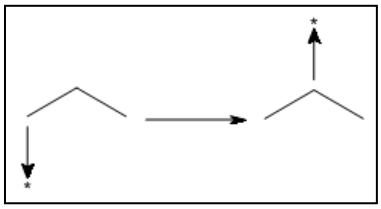
Example of fragments that have become a set of Rgroup members:



Defined Rgroup members, noted by an asterisk and an arrow

5. **(Optional) Change attachment points**. To change an attachment point:
   - Click the asterisk to select it.
   - Drag and drop it onto the other atom that you want, in the Rgroup member.
   - Repeat this step for each attachment point that you want to change. Only the asterisk of the attachment point must be selected before you can change the attachment point.

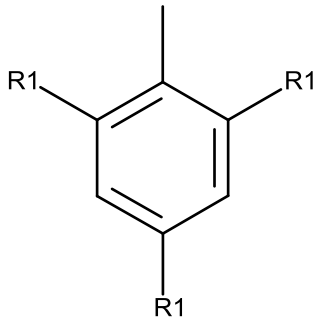
   Example of changing an attachment point:



Define a different Rgroup attachment point

6. **(Optional) Change the frequency**. If an Rgroup atom appears at more than one instance (or place) in the root structure, you see "R1 = n (where n is defined as the number of occurences), R2 >0, etc." appear automatically next to the Rgroup definition (Rn =). For each such Rgroup, you need to specify the frequency (occurrence), the number of times that a member of this Rgroup must appear in retrieved structures. To change the frequency:
   - Select the Rgroup Query Tool.
   - Click the Rgroup definition, Rn= and click **Advanced Rgroup Logic…**
   - Change **occurrence** field of that Rgroup
   - Select a number from the dialog box that appears.
   - Click **OK** to accept your selection. The frequency definition is updated with your selection.

   Example that uses the default value of frequency for Rgroup R1:



Rgroup frequency

Additional Commands and Capabilities

For all of these commands and capabilities, a highlighted item such as an Rgroup member or an Rgroup definition appears in a blue rectangle.

To delete an Rgroup member

1. Click the Rgroup tool.
2. Click the highlighted member.
3. Select Delete.

To move an Rgroup member to a different Rgroup

1. Click the highlighted member.
2. Select Move to different Rgroup.
3. Click the Rgroup number.

To delete an Rgroup

1. Click the Rgroup definition (e.g., "R1=").
2. Select Remove Rgroup Definition.

Using the Rgroup tool, you can click a highlighted Rgroup member to do the following operations:

- Remove from Rgroup
- Move to a different Rgroup
- Delete
- Duplicate
- Select

Using the Rgroup tool, you can click the Rgroup definition (Rn=) to do the following operations:

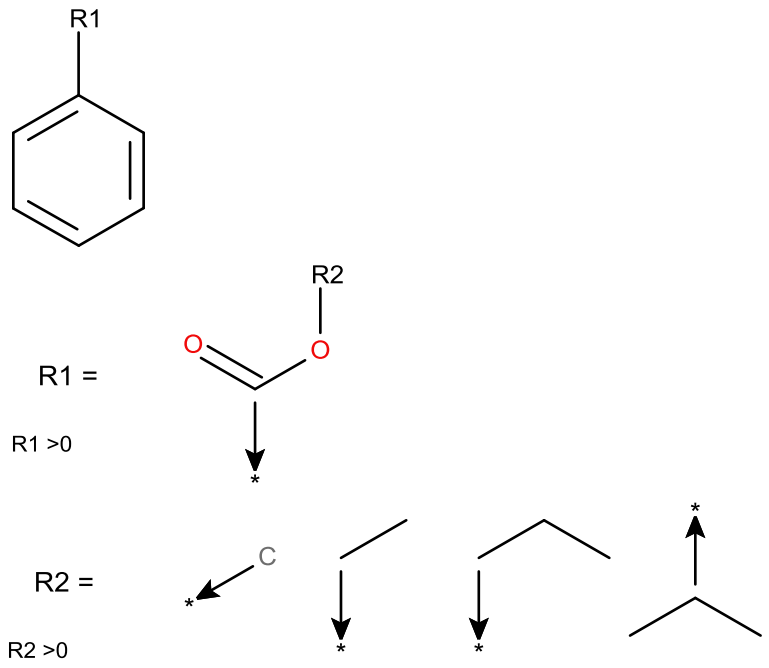
- Select members
- Exclude members
- Advanced Rgroup Logic
- Remove Rgroup definition

Nested Markush queries

A **nested Markush query** is a structural query in which one Rgroup is contained within another. Nested Markush queries give you an additional capability to create more advanced, precise queries. You can nest Rgroups to a maximum of 2 levels.

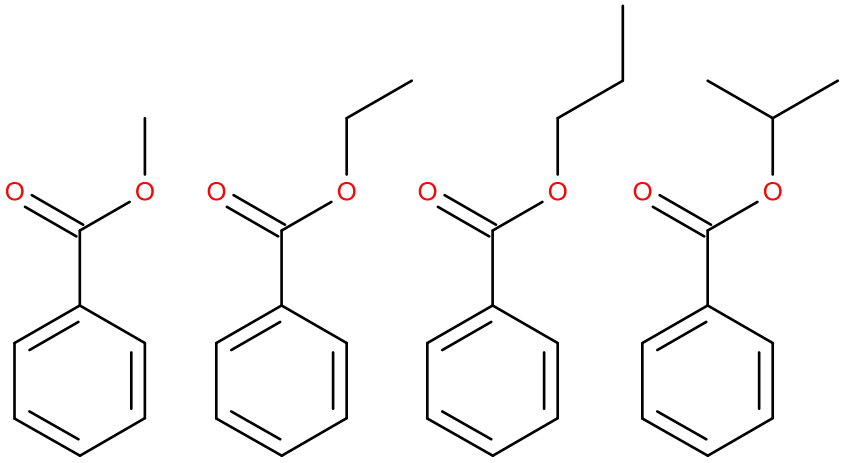
Example: Find all 1-, 2-, and 3-carbon esters of benzoic, benzenesulfonic, and benzene-phosphonic acids.

Query:



Nested Markush query

Sample results:

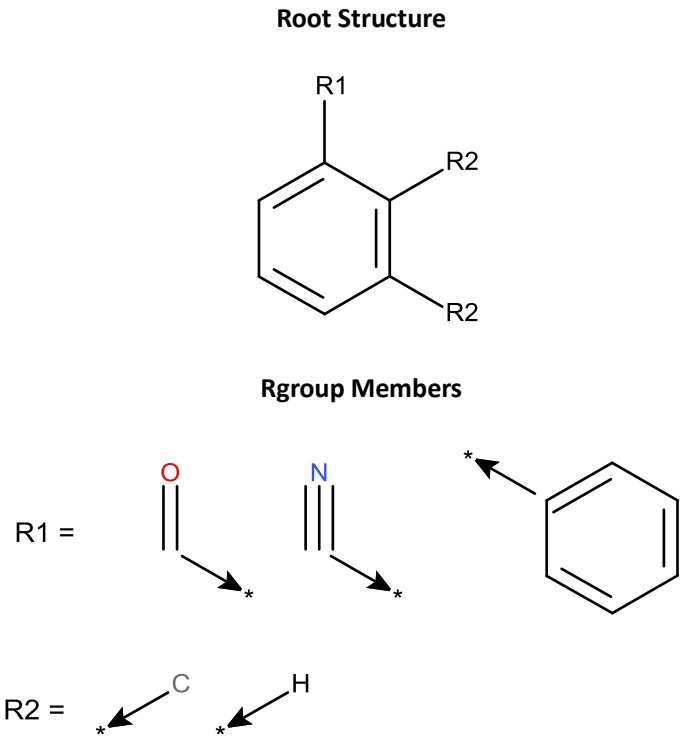


Nested Markush query results

Chapter 10:
Markush (Rgroup) Query Tool, Example

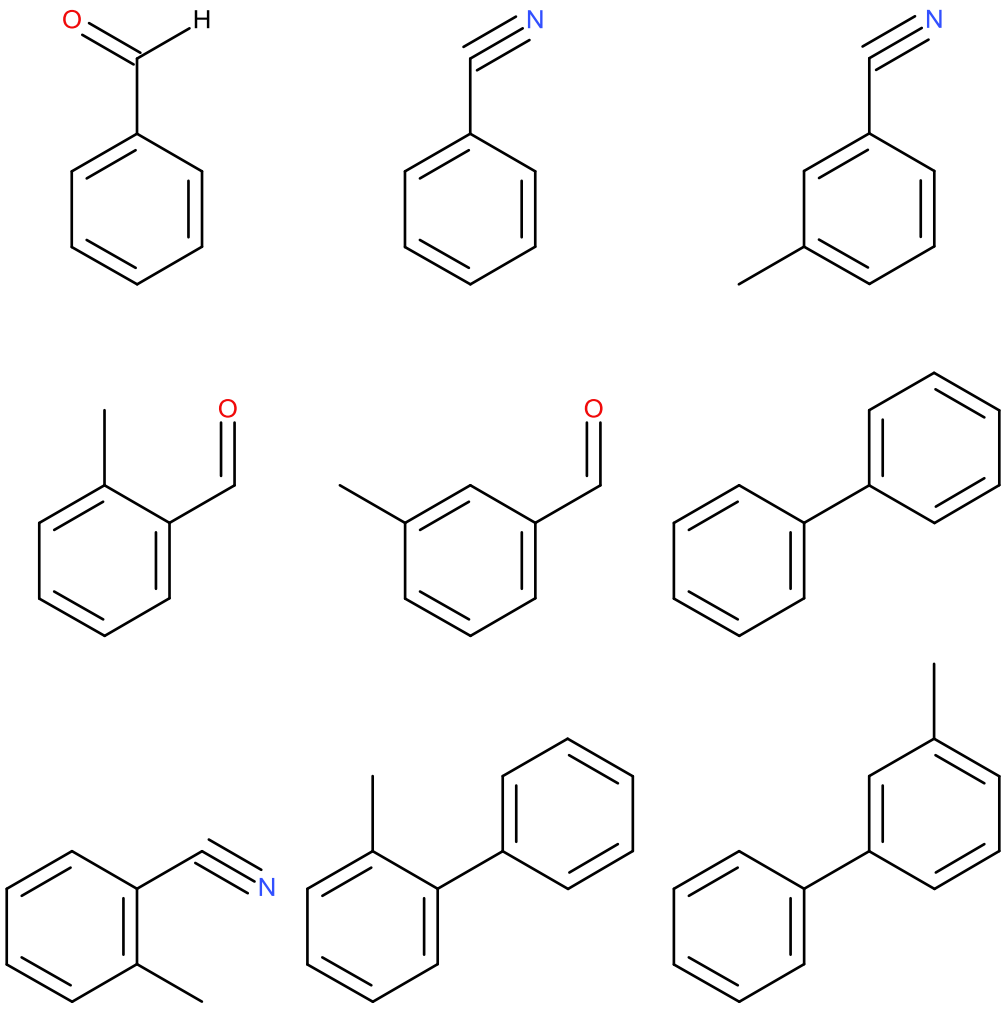
Example of an Rgroup Query

Example and illustration of a Markush query:



Example of a Markush query

Some structures that you retrieve with this query:



Markush query results

Chapter 11:
Markush Query Tool

Working with the Rgroup Query Tool

See Example of an Rgroup Query for more details about Markush queries, an example, and a detailed drawing procedure.

To draw a Markush query

1. Draw the root structure. Use the other drawing tools.
2. Add Rgroup atoms to the root structure.
   - Click the Markush Query tool.
   - Click the atom that you want to replace.
   - Select an Rgroup atom from the palette.
3. Draw the Rgroup members. Use the other drawing tools.
4. Add Rgroup members.
   - Click the Markush Query tool.
   - Click the fragment that you want to add.
   - Drag and drop the fragment onto the Rgroup definition (Rn=).
5. **(Optional)** Move attachment points.
   - Click the Markush Query tool.
   - Click the asterisk of the attachment point.
   - Drag and drop the asterisk onto the atom that you want.
6. **(Optional)** Exclude all members of an Rgroup.
   - Click the Rgroup definition (Rn=).
   - Select **Exclude All Members** from the menu that appears.
7. **(Optional)** Change the occurrence.
   - Click the Markush Query tool.
   - Click the Rgroup definition, Rn= and click **Advanced Rgroup Logic…**
   - Change **occurrence** field of that Rgroup
   - Select a number from the dialog box that appears.
8. **(Optional)** Specify H at unoccupied Rgroup sites (RestH).
   - Click the Markush Query tool.
   - Click the Rgroup definition, **Rn =**.
   - Click the Rgroup definition, Rn= and click **Advanced Rgroup Logic…**
   - Change **RestH** field of that Rgroup