

CAS STNEXT<sup>®</sup>

# Overview on current issues in Derwent Markush Resource Database

Last update: June 2024

This document summarizes all known issues that can affect the results of structure searches in the Derwent Markush Resource (DWPIM) database.

If you need support concerning your structure query or regarding DWPIM in general, please contact [EMEAhelp@cas.org](mailto:EMEAhelp@cas.org).

# 1. Use M (any metal) generic node for Metallacycles

Derwent metal generic nodes A35, ACT, AMX, LAN and TRM do not work correctly within ring systems and lead to 0 hits.

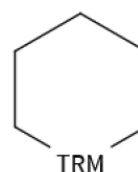
Possible Workaround: Use generic node M (any metal) within rings.

Variables	
<input type="checkbox"/> X	Any halogen
<input checked="" type="checkbox"/> M	Any metal
<input type="checkbox"/> A	Any atom except H
<input type="checkbox"/> Q	Any atom except C or H
<input type="checkbox"/> Ak	Any carbon chain
<input type="checkbox"/> Cy	Any cycle
<input type="checkbox"/> Cb	Any carbocycle
<input type="checkbox"/> Hy	Any heterocycle
<input type="checkbox"/> Id	ID generic node

Derwent (DWPIM/DCR) generic nodes	
▼ Metals	
<input type="checkbox"/> A35	Group IIIA-VA metals
<input type="checkbox"/> ACT	Actinides
<input type="checkbox"/> AMX	Alkali and alkaline earth metals
<input type="checkbox"/> LAN	Lanthanides
<input type="checkbox"/> TRM	Transition metals
► Carbon Chains	
► Carbocycles	
► Heterocycles	
► Miscellaneous (DWPIM only)	

Do not use metal  
superatoms in rings,  
e.g.

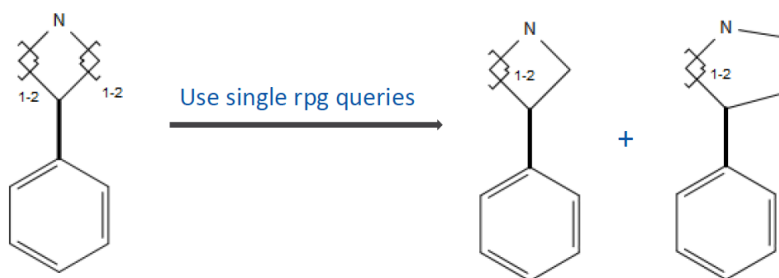


## 2. Within rings only one repeating group (rpg) per ring is allowed

In case of two or more repeating groups within the same ring incorrect results are obtained.

There is no workaround other than drawing out all possible combinations for every additional repeating group.

Example:



### 3.DWPIM Hitstructures in WPIX/WPINDEX after Cross-Over

DWPIM hitstructures are only preserved in WPIX/WPINDEX answer sets within the same session (including log off hold). Otherwise hitstructures are lost and the full Markush record is displayed. This limitation may also affect results from SDI's and batch searches.

Recommendation: hitstructure information within DWPIM is preserved. Therefore all answer sets should always be also saved in DWPIM.

Records from DWPIM and WPIX/WPINDEX can be mutually allocated by Markush number (AN in DWPIM corresponds to MCN in WPIX/WPINDEX).

## 4.STNext: Adjust bond value for Carboxylic acid derivates to exact/normalized (e/n)

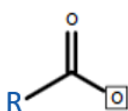
Carboxylic acid derivates as well as corresponding phosphoric, sulfonic and selenic acid derivates are defined with normalized bonds in the STNext structure editor (as shortcuts or drawn, see table below). In order to achieve correct matching in DWPIIM those bonds need to be set to exact/normalized. This can be done either by drawing a respective group with locked heteroatom (bonds are automatically set to e/n in the editor) or by changing the normalized bonds of the Hydrogen containing group manually.

The bond value of the shortcuts cannot be changed and therefore those shortcuts should not be used at all!

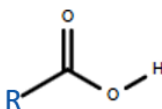
The following groups are affected:

-COOH	-OPO3H2	-SO3H	-SeO3H
-CO2H	-OSO3H	-SO2H	-SeO2H
-COSH	-PO3H2		
-CSSH	-PO2H		
-CS2H			

How to draw those groups correctly:



e/n (default)



Switch from n (default) to e/n

## 5. Usage of generic ring and chain nodes via variable attachment points (VPA) can lead to incorrect results

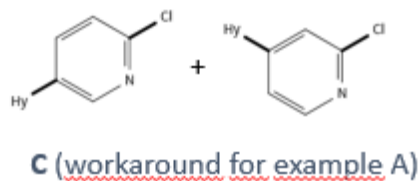
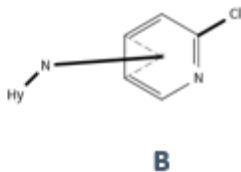
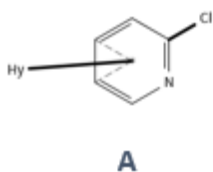
The problem only applies to directly VPA connected generic ring or chain nodes (example A). If there is at least one atom in the VPA between the generic ring or chain node and the main molecule, the problem does not exist (example B).

The following nodes are affected:

- Cy, Cb, Hy, CYC, ARY, HEA, HET, HEF
- Ak, CHK, CHE, CHK

The only practicable workaround for this issue is to use the respective individual structures encoded in the VPA query (example C).

Examples:



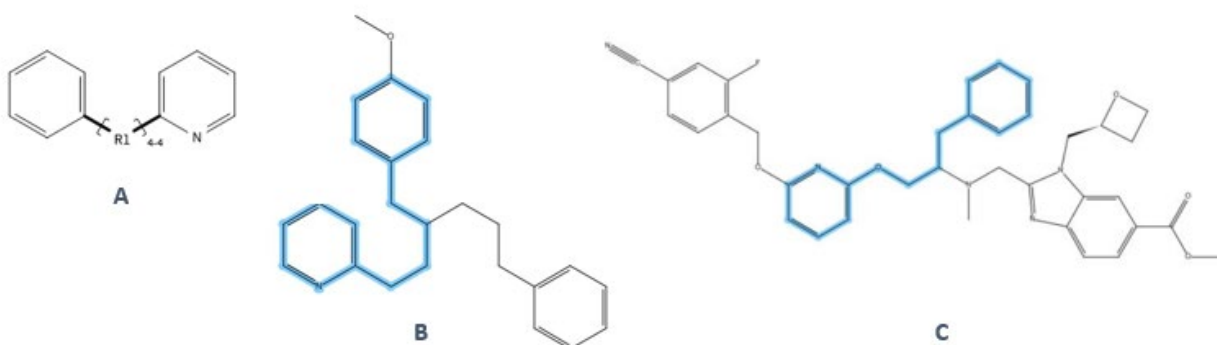
## 6.R-groups within repeating groups may lead to incomplete results

The variants within R groups within repeating groups do not provide mixed results but only results of one of the occurring variants.

Examples:

Query structure A with R1=C,O does only yield results with either C's (example B) or O's but not mixed results (example C).

The only practicable workaround is to search for structures with mixed variants separately as individual structures.



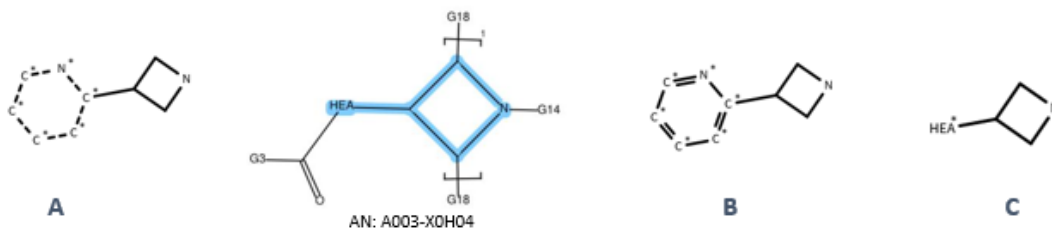
## 7. Rings with unspecified bonds do not match on aromatic generic ring nodes (ARY, HEA)

Examples:

Query structure A does not match on hit records containing HEA (e.g. record A003-X0H04 ).

The only practicable workaround is to use query structures with alternating double and single bonds (example B, Match level Class) or to use the respective generic aromatic ring node HEA (example C, Match level Class).

The same applies for carbocycles. In these cases hit records with ARY are missed. For the workaround the respective benzene ring or the ARY node have to be used.





Solution images are included for illustrative purposes only. Your experience may vary based on recent enhancements or product license.



CAS is a leader in scientific information solutions, partnering with innovators around the world to accelerate scientific breakthroughs. CAS employs over 1,400 experts who curate, connect, and analyze scientific knowledge to reveal unseen connections. For over 100 years, scientists, patent professionals, and business leaders have relied on CAS solutions and expertise to provide the hindsight, insight, and foresight they need so they can build upon the learnings of the past to discover a better future. CAS is a division of the American Chemical Society.

**Connect with us at [cas.org](https://cas.org)**

