



# Tips from FIZ Search Service

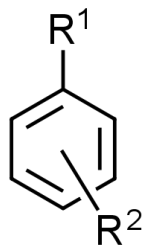
## No 2- Rethink your structure

Renate Hedderich

# Typical structure search question

- Requested is a core structure with various substituents attached

Task: find all di-substituted benzene rings with the following formula:



$R^1$  = carbon chain, opt. subst.

$R^2$  =  $C_{3-10}$  carbocycle, opt. subst.

*No further substitution* at the benzene ring

- Our task is to transfer the requested structure into an STN compatible structure query
- Usually the resulting structure query looks similar to what the customer has specified
- But different approaches can be made for this request

# Approach 1 – variable point of attachment - SSS

Structure Editor - 2020\_0212\_Structure\_11

Enter a CAS RN, SMILES or InChI

Click and drag from the substituent position to each ring position where attachment may occur.

Variables

- X Any halogen
- M Any metal
- A Any atom except H
- Q Any atom except C or H
- Ak Any carbon chain
- Cy Any cycle
- Cb** Any carbocycle
- Hy Any heterocycle
- Id ID generic node

Node Attributes

Valency  
Hydrogen Count  
Non-Hydrogen Count  
Element Count

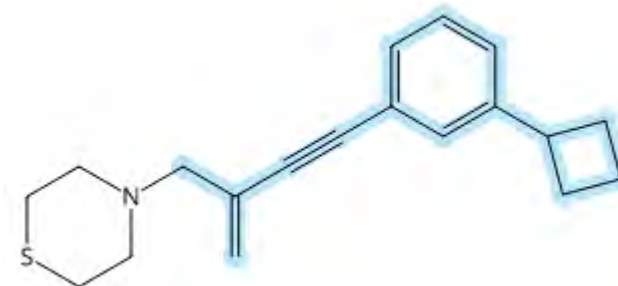
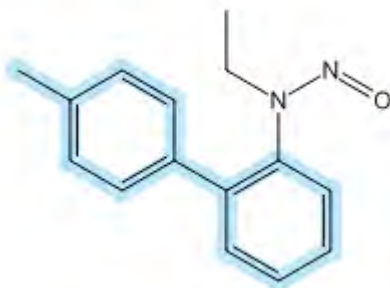
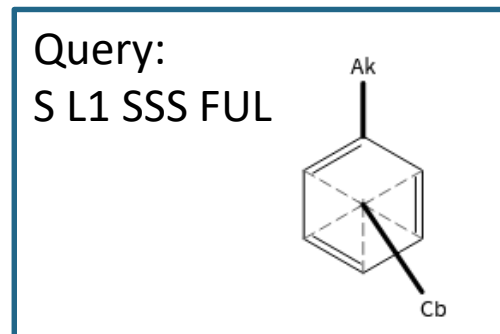
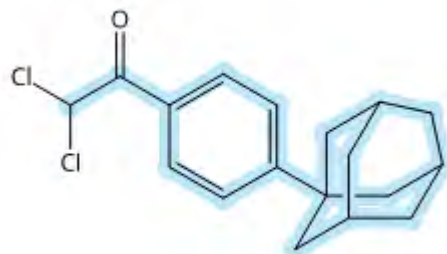
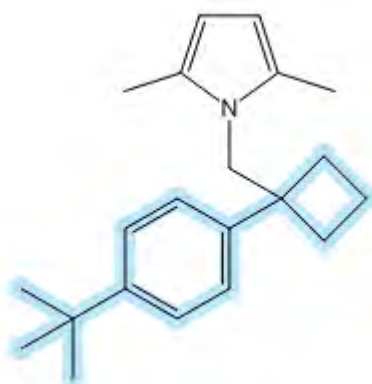
Open structure editor and draw structure using variable point of attachment.

Draw a connection between the substituent and each point of attachment

Final structure query:

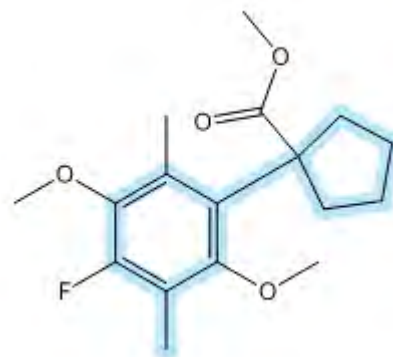
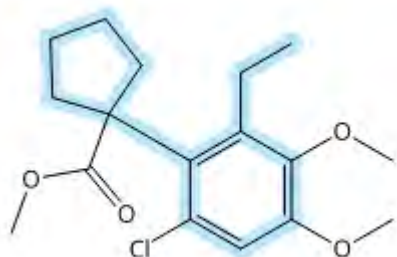
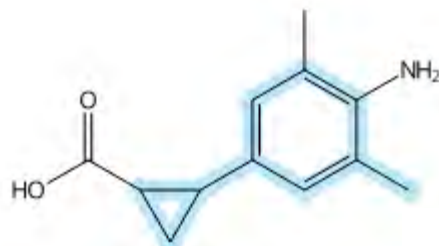
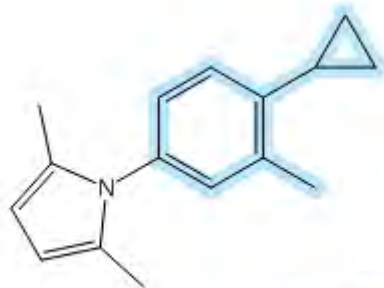
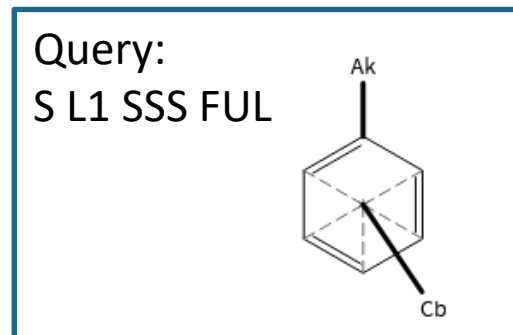
# Approach 1 – variable point of attachment - SSS

- A substructure search (SSS) will retrieve valid matches:



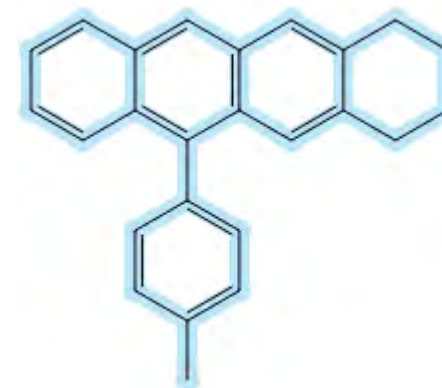
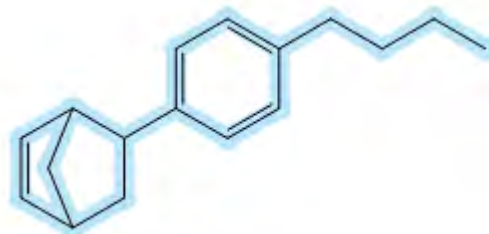
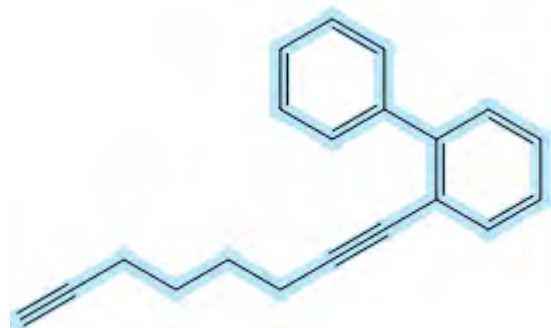
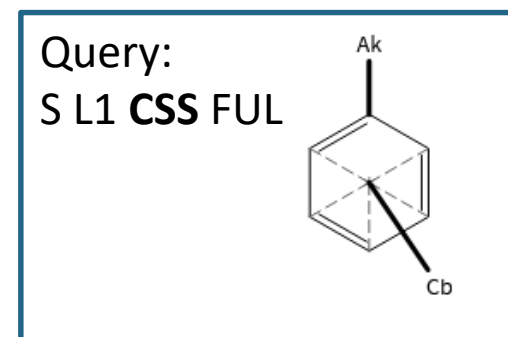
# Approach 1 – variable point of attachment - SSS

- But also unwanted hits with *more* than 2 substituents at the central benzene ring will be captured



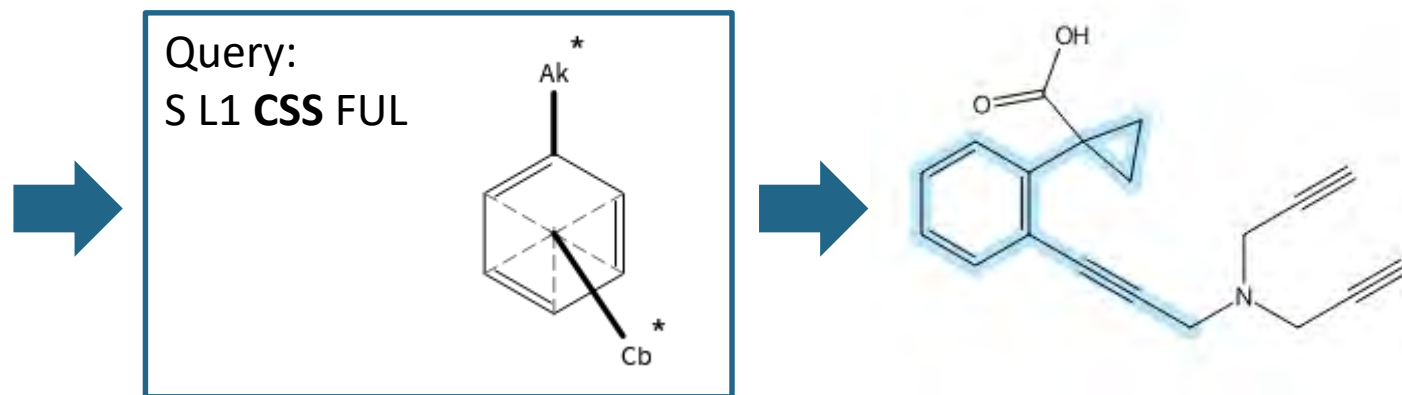
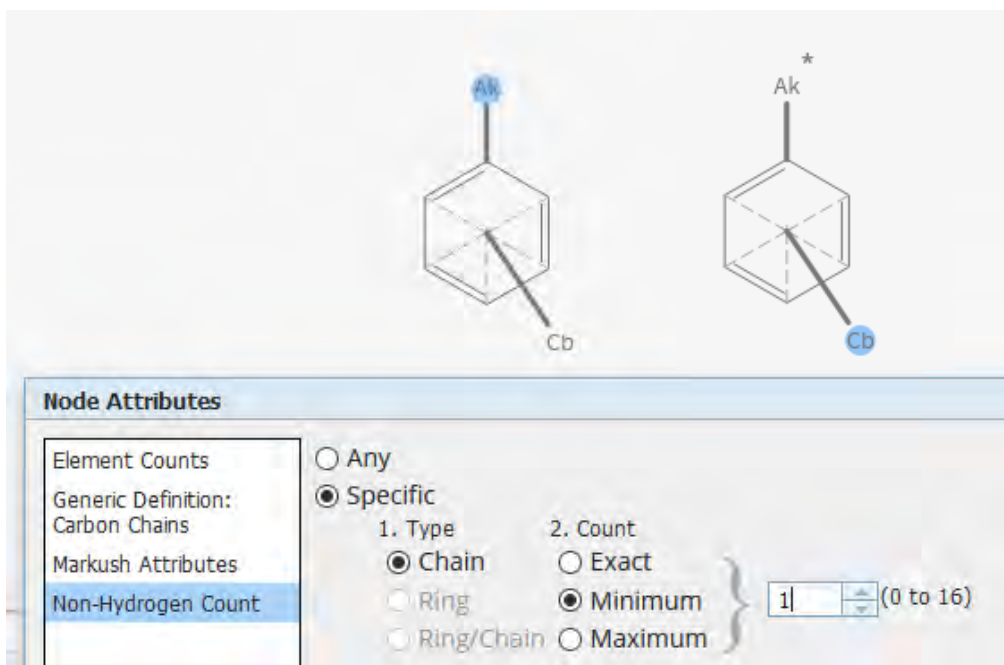
# Approach 2 – variable point of attachment - CSS

- Perform a Closed Substructure Search (CSS) using the same structure query
- Additional substitution at the benzene ring will be prevented ✓
- Also further substitution at the substituents is prevented ✗
  - „Alkyl“ will only retrieve *unsubstituted* carbon chains
  - „carbocycle“ will only retrieve *unsubstituted* rings



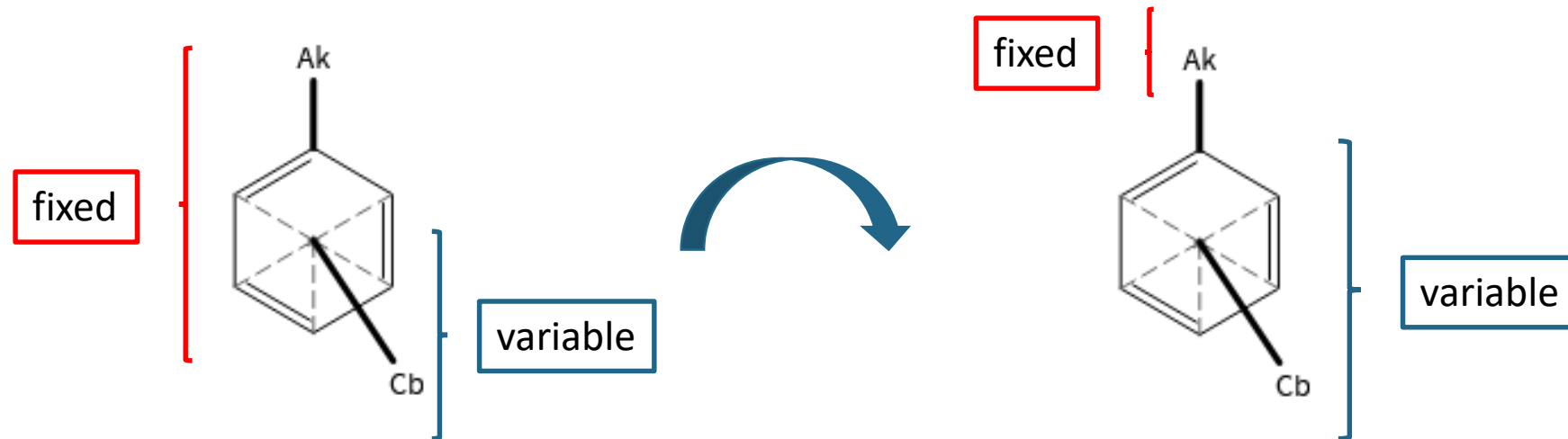
# Approach 2 – variable point of attachment - CSS

- Structure query has to be adjusted
- Allow further substitution at Ak and Cb node by defining *non-H attachments*
  - Advanced structure control tool
  - In this simple example this is a good strategy
  - For larger structures it may be more effort to care for all nodes which have to be opened




# Approach 3 – change the perspective

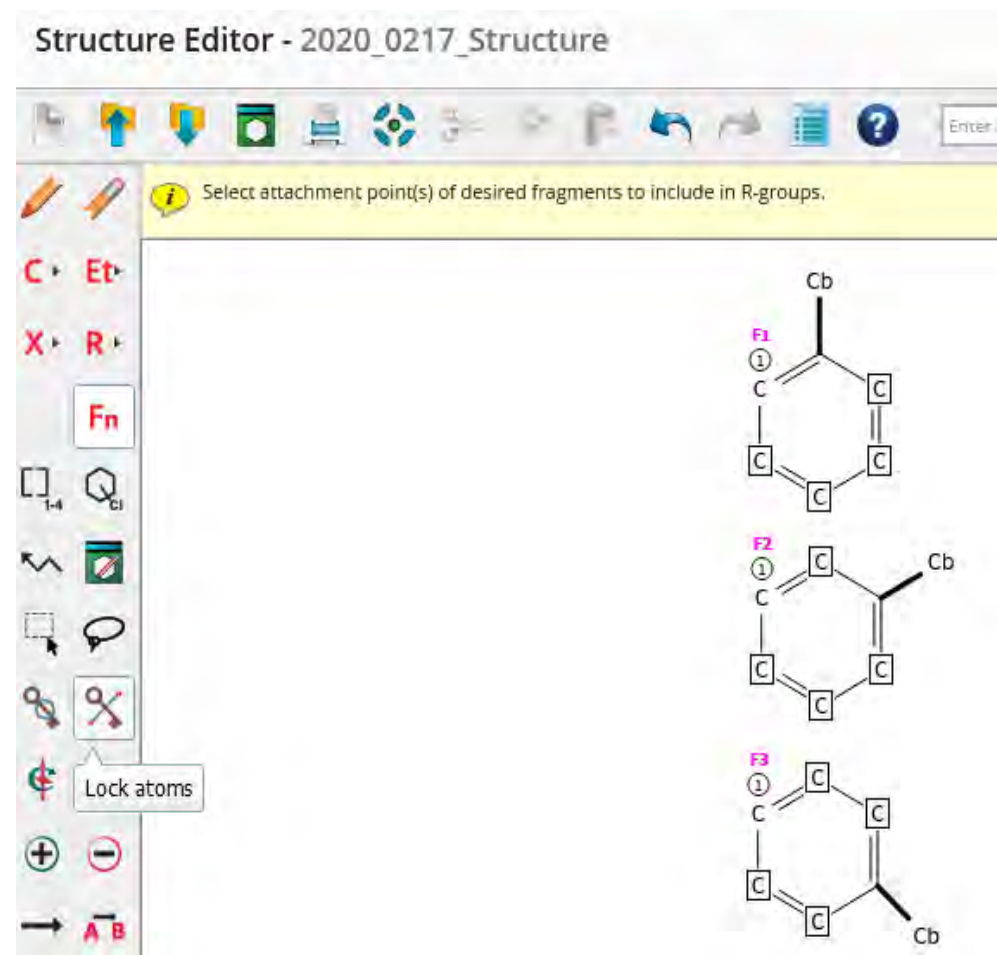
- Find a fix node in the structure and shift the „search focus“
- Move from the central ring to the substituent





# Approach 3 – change the perspective

- Use fragments for the definition of the variable part
- Define fragments **Fn** with the requested substitution pattern
  - Ortho, meta and para position
- Lock  all other positions at the benzene ring
  - Alternative: non-H attachments on exact 2 ring/chain



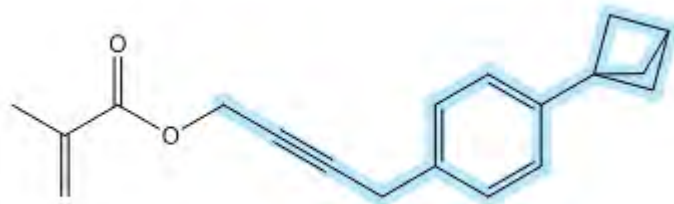
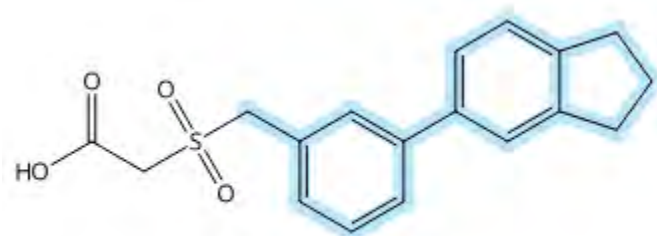
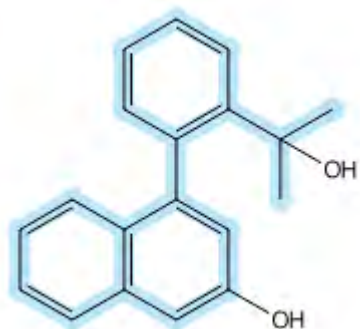
# Approach 3 – change the perspective

Group the fragments into an R-group

The screenshot shows the 'Structure Editor - 2020\_0217\_Structure' interface. On the left, a toolbar contains various drawing tools, with the 'R' group definition tool highlighted. A callout box points to this tool with the text 'Group the fragments into an R-group'. The main workspace displays a chemical structure of a substituted benzene ring with a 'Cb' group and three fragments labeled 'F1', 'F2', and 'F3'. A label 'Ak - R1' is placed near the structure. On the right, the 'R-group Definitions' panel shows a list of R-groups (R1 to R10) and a definition for R1:  $R1 = F1, F2, F3$ . Below this, the 'Fragments' section displays three separate views of the fragments F1, F2, and F3, each with its corresponding atom numbering.

# Approach 3 – change the perspective

- Only di-substituted benzene rings are retrieved
- Substitution on Ak and Cb is allowed



Query:


S L3 SSS FUL

Ak → R1



# Comparison of strategies

- Approach 1: VPA and SSS
  - Broadest strategy
  - Retrieves a lot of unwanted answers (tri- and more substituted)
- Approach 2: VPA and CSS
  - Exact matches
  - Remember to open all nodes which allow substitution
  - Be familiar with non-H attachments
- Approach 3: Fragments/R-groups and SSS
  - Exact matches
  - Change perspective
  - Be familiar with drawing fragments



Approach 2 and 3 are equivalent for the results

# FIZ Search Service Team



- Scientists with several years of search experience
- We offer all kinds of patent and literature searches on STN
- Structure searches in REGISTRY/CAPLUS, MARPAT, DCR/DWPI and DWPIM/DWPI

Thank you for your attention  
and have a nice day!

[search.service@fiz-karlsruhe.de](mailto:search.service@fiz-karlsruhe.de)

[www.fiz-karlsruhe.de/en/search-service.html](http://www.fiz-karlsruhe.de/en/search-service.html)

 **FIZ Search Service**

FIZ Karlsruhe – Leibniz Institute for Information Infrastructure