STN®: Structure Drawing in STN

e-Seminars

Today's presenter is…

Gin-Yun Eggerichs

Peter Blasi

STN® offers several tools to customize structures for your search needs

- Chemical structure searching has been used heavily in many drug discovery studies for years
- Often times, several types of fragments are used to attach to multiple types of core structures and in any orientation
- Other times, there is a need to make specific modifications to standard variables such as an alkyl chain (Ak)
- STN has the capability to draw structures and make specific modifications to the fragments and/or standard variables to achieve your needs

Goals for this seminar

- Attach user defined variable groups (G-groups) onto a structure (substituent G-groups)
  - STN defined variables
  - User defined fragments
- Incorporate G-groups into a core structure (embedded G-groups)
  - Two points of attachment on a single fragment
  - Refine answer set by atom elimination
- Apply variable points of attachments with bridging group between two rings

Drawing features are often used to expand or narrow variability

- STN defined
  - Shortcuts
  - Variables
  - System defaults for rings, bonds, and nodes
- User defined
  - Repeating groups
  - G-groups
  - Variable point of attachment (VPA)

Features differ in the amount of user control available for these characteristics

- Shortcuts
  - Specifically defined, no modification allowed
  - Valence warning appears with improper use
- Variables X, A, Q, M represent single atoms
  - Node characteristics
  - Hydrogen attachments
  - Non-H attachments

Features differ in the amount of user control available for these characteristics

- Variables Cb, Cy, Hy, Ak represent fragments
  - Non-H attachments
  - Generic definition
    - Saturation
    - Type of chain/ring system
    - Number of carbons
    - Number of hetero atoms
  - Element count

Users can specify variables by creating G-groups

- A query may contain up to 20 G-groups
- A G-group may contain up to 20 substituents
  - Specific elements
  - Shortcuts
  - System-defined variable groups
  - Structural fragments
  - Other G-groups

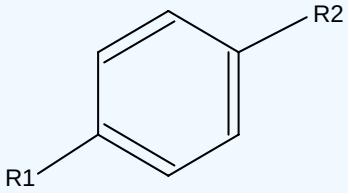
G-groups as substituents: single point of attachment

Search Question 1:

Find compounds described here:
Benzene ring with G1 and G2 variables attached at the para position.

R1 *can be either a monocyclic ring system with less than 7 carbon atoms, 1 heteroatom, or a chain of 2-5 carbons.*

R2 *needs to be a butyl ester that can be attached at either end of the fragment.*



Search Strategy

To find the cyclic compounds described...

Step 1. Draw the structures

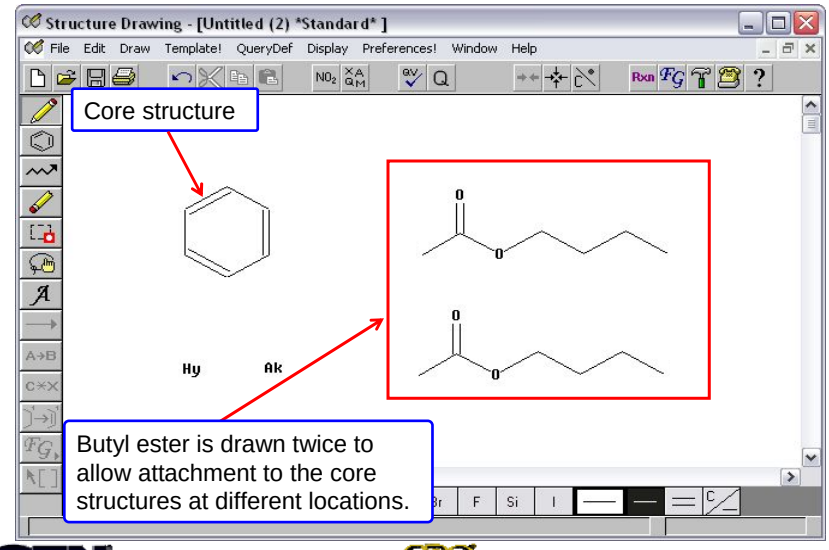
- Draw the fragments
- Assign point of attachment to the fragments
- Attach the fragments to the core structure via G-group

Step 2. Upload the query to STN

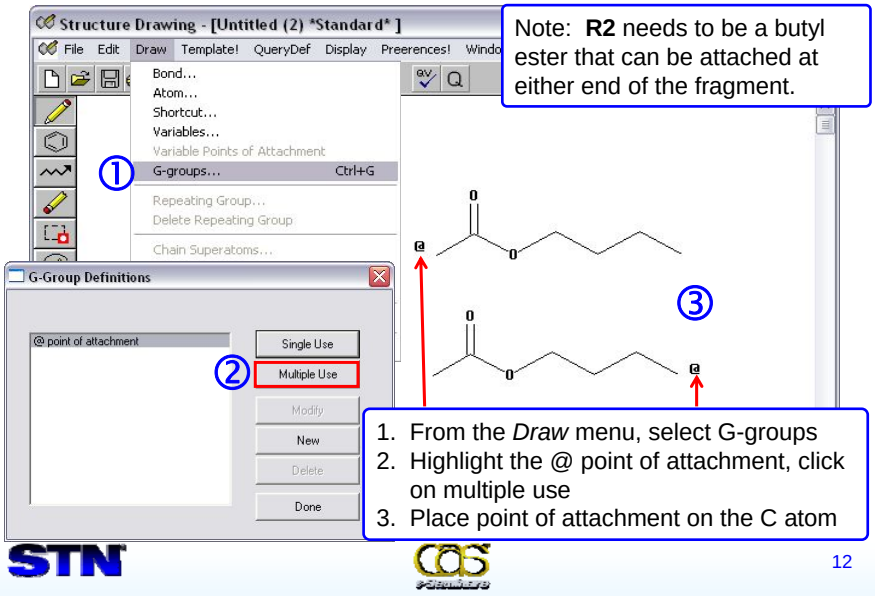
Step 3. Run the search in CAS REGISTRY℠

- Sample search
- Full search

Draw all the structures on STN

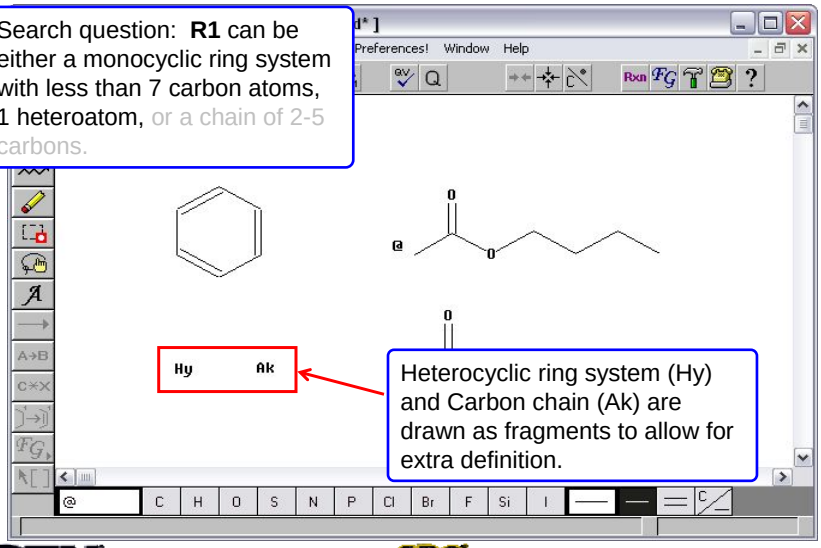


Identify points of attachment for the butyl ester fragments



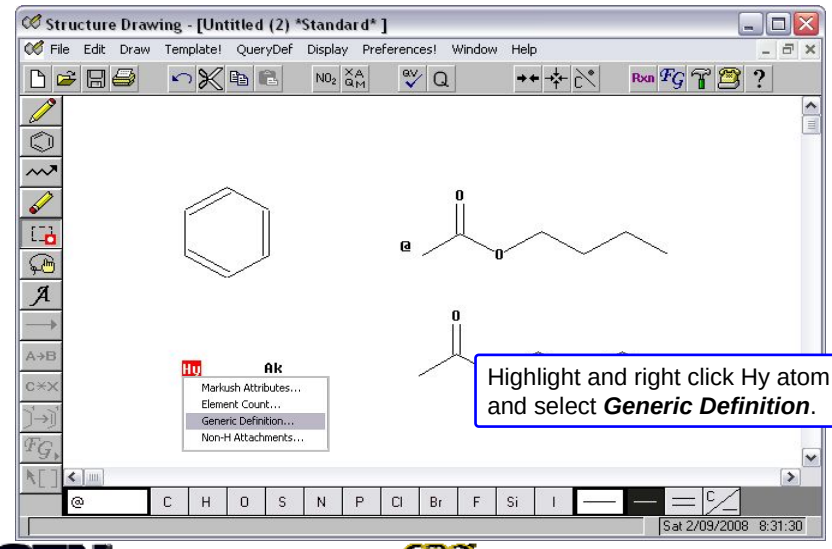
STN defined variables can be drawn as fragments

Search question: **R1** can be either a monocyclic ring system with less than 7 carbon atoms, 1 heteroatom, or a chain of 2-5 carbons.



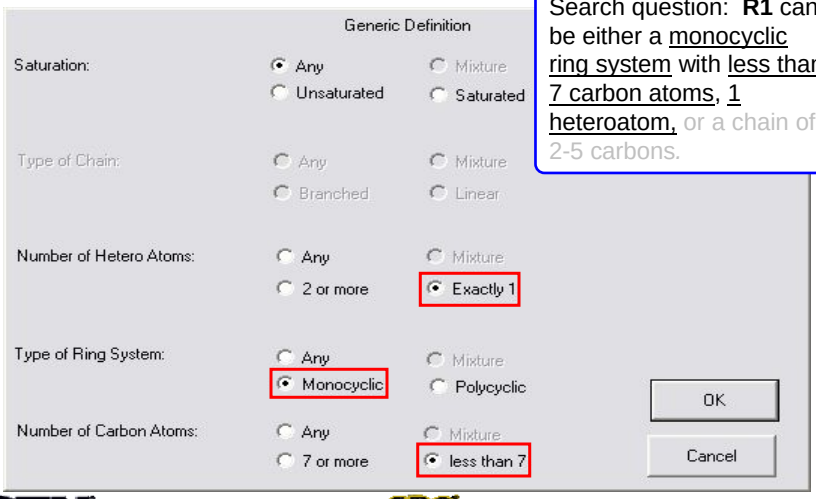
Heterocyclic ring system (Hy) and Carbon chain (Ak) are drawn as fragments to allow for extra definition.

The Hy variable can be defined to allow precise structure retrieval



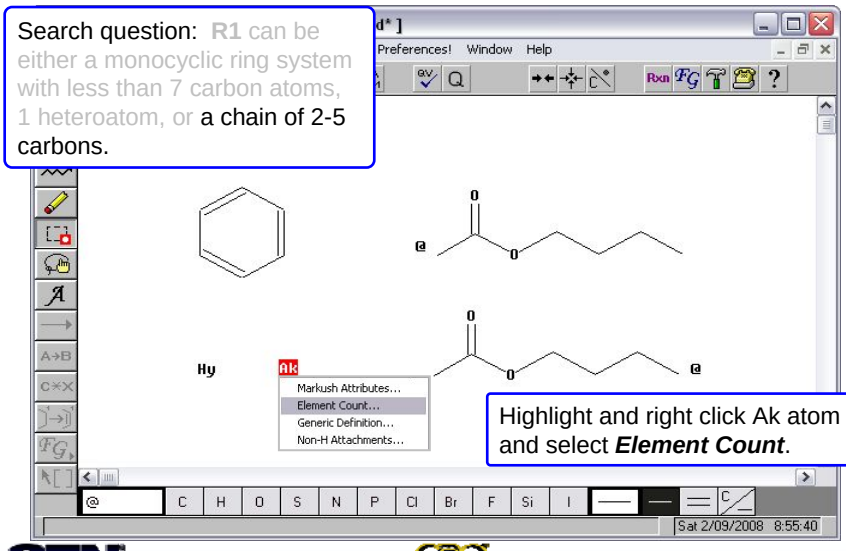
Highlight and right click Hy atom and select ***Generic Definition***.

Generic Definition options allow for precise structure retrieval



Search question: **R1** can be either a monocyclic ring system with less than 7 carbon atoms, 1 heteroatom, or a chain of 2-5 carbons.

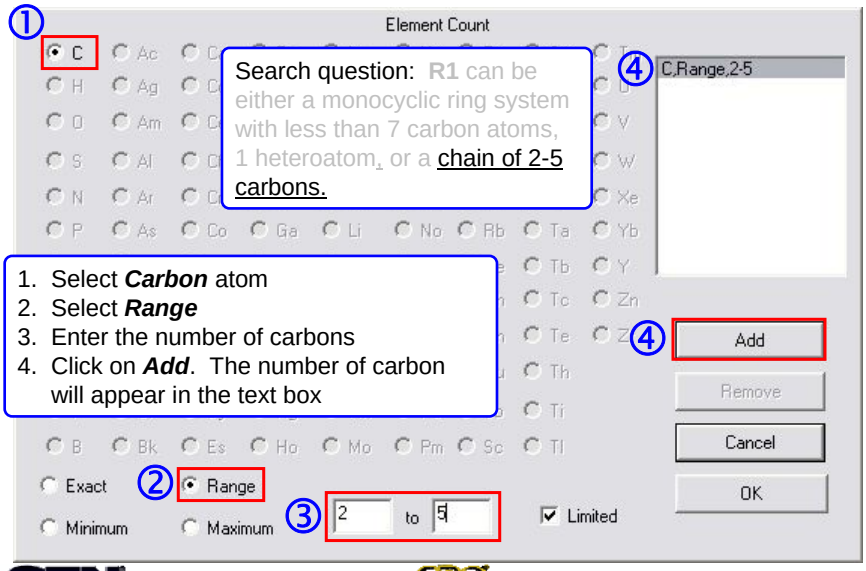
Define the Ak Element Counts



Search question: **R1** can be either a monocyclic ring system with less than 7 carbon atoms, 1 heteroatom, or a chain of 2-5 carbons.

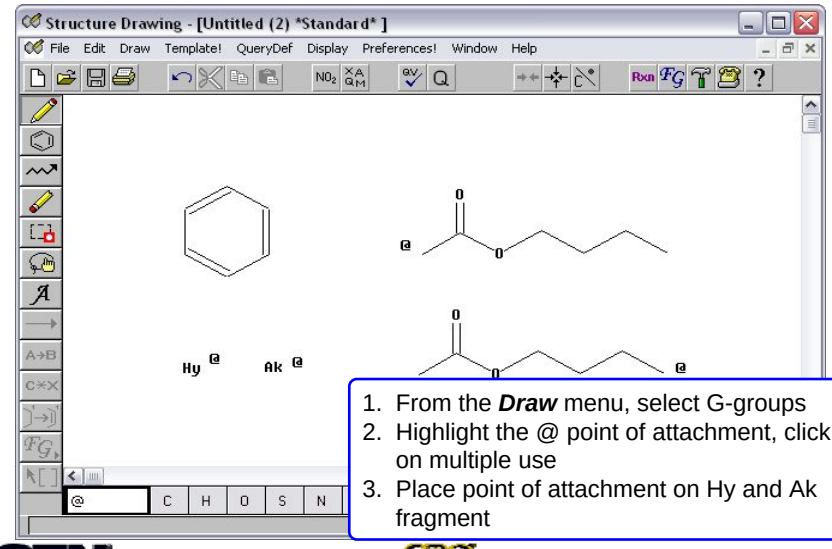
Highlight and right click Ak atom and select ***Element Count***.

Define the Ak Element Count



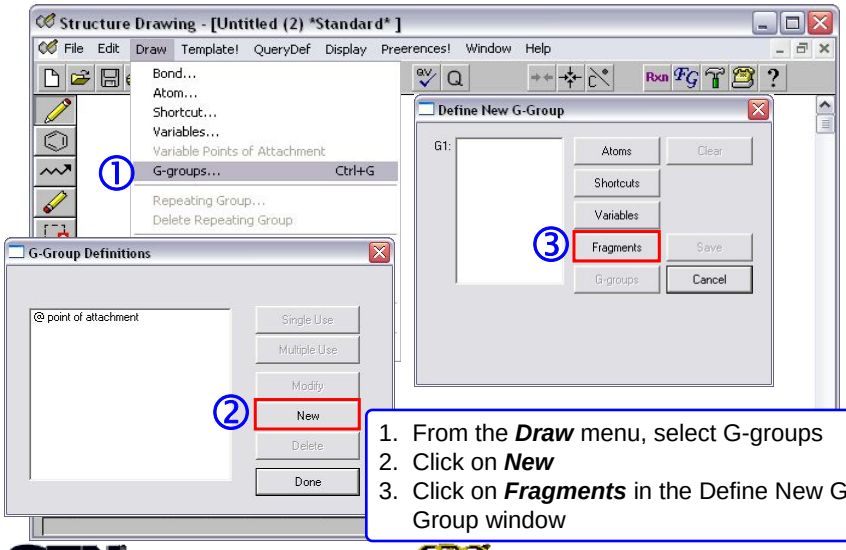
1. Select ***Carbon*** atom
2. Select ***Range***
3. Enter the number of carbons
4. Click on ***Add***. The number of carbon will appear in the text box

Assign the points of attachments for Hy and Ak



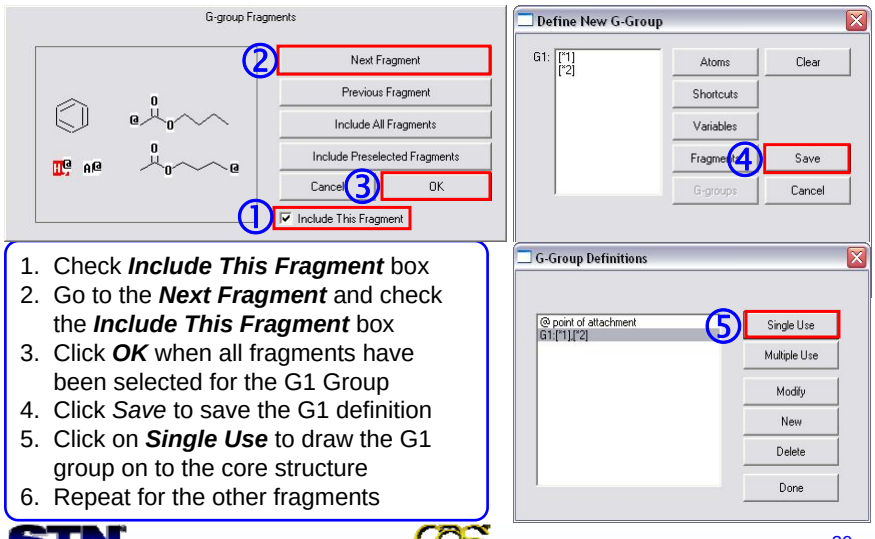
1. From the ***Draw*** menu, select G-groups
2. Highlight the @ point of attachment, click on multiple use
3. Place point of attachment on Hy and Ak fragment

Define G-groups with fragments



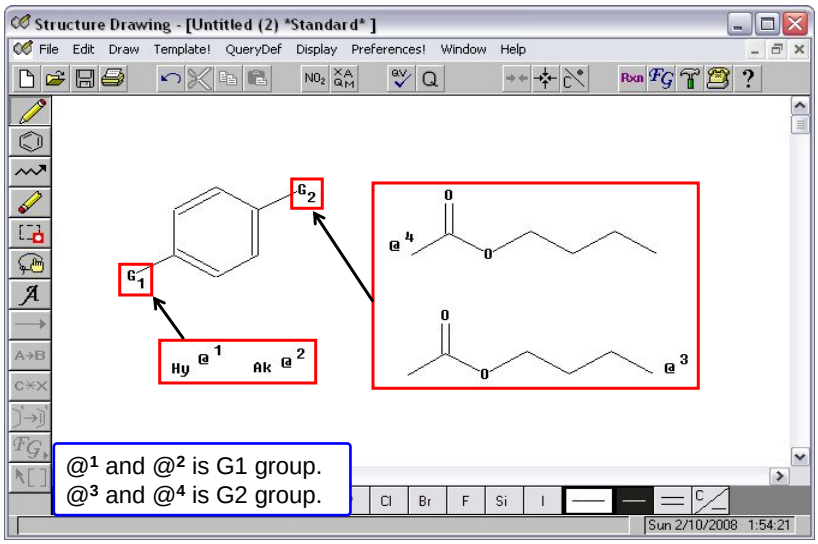
1. From the ***Draw*** menu, select G-groups
2. Click on ***New***
3. Click on ***Fragments*** in the Define New G-Group window

Define G-groups with fragments



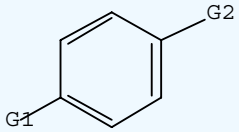
1. Check ***Include This Fragment*** box
2. Go to the ***Next Fragment*** and check the ***Include This Fragment*** box
3. Click ***OK*** when all fragments have been selected for the G1 Group
4. Click *Save* to save the G1 definition
5. Click on ***Single Use*** to draw the G1 group on to the core structure
6. Repeat for the other fragments

Place the G1 and G2 groups on to the core structure

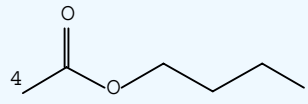


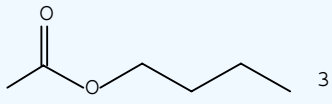
Upload structure and check query

```
=> FIL REG

=> Uploading C:\CASNC\STN Express\Queries\Structure
   Drawing E-seminar\Query 1\Query
L1      STRUCTURE UPLOADED

=> D L1 QUERY
L1              STR
```



Hy 1 Ak 2






```
  G1 [@1],[@2]
  G2 [@3],[@4]
```

Check the query before running the search.

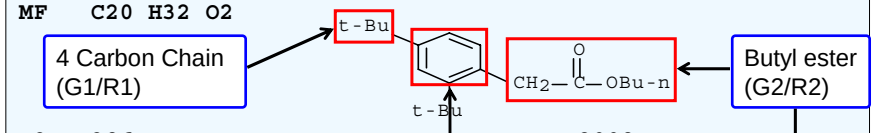
Run the substructure search

```
=> S L1 SSS SAM
SAMPLE SEARCH INITIATED 15:44:24 FILE 'REGISTRY'
SAMPLE SCREEN SEARCH COMPLETED -    10368 TO ITERATE

FULL FILE PROJECTIONS:  ONLINE  **COMPLETE**
                        BATCH   **COMPLETE**
PROJECTED ITERATIONS:       201258 TO   213462
PROJECTED ANSWERS:              75 TO      547

L2            3 SEA SSS SAM L

=> S L1 SSS FULL
FULL SEARCH INITIATED 15:42:4
FULL SCREEN SEARCH COMPLETED -   207081 TO ITERATE

100.0% PROCESSED   207081 ITERATIONS   386 ANSWERS

L3          386 SEA SSS FUL L1
```

The sample search is useful to determine if the search will run to completion within system limits before running a full structure search.

G1/G2 groups are attached to the core structure as substituents

```
=> D SCAN

L3    386 ANSWERS   REGISTRY  COPYRIGHT 2008 ACS on STN
IN    Acetic acid, (2,4-di-tert-butylphenyl)-, butyl ester
      (7CI, 8CI)
MF    C20 H32 O2
```

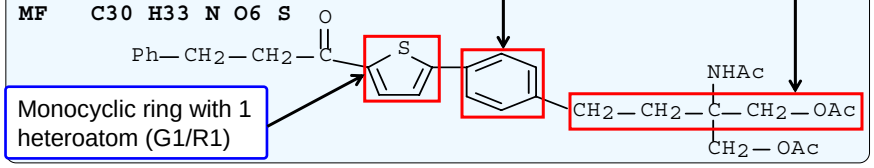


4 Carbon Chain (G1/R1)

Butyl ester (G2/R2)

```
L3    386 ANSWERS   REGISTRY  COPYRIGHT 2008 ACS on STN
IN    Acetamide, N-[1,1-bi                    methyl]-3-[4-[5-(1-
      oxo-3-phenylpropyl)-2-thienyl]phenyl]propyl]-
MF    C30 H33 N O6 S
```

Core structure

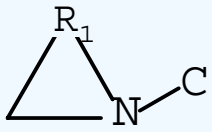


Monocyclic ring with 1 heteroatom (G1/R1)

G-groups: fragment with 2 points of attachment

Search Question 2:

Find compounds described here:
Cyclic compound containing 3-8 members with oxygen located at R_1. In case of the 5 member ring, the oxygen can attach through the carbon as well as the nitrogen. The carbon attached to the nitrogen is a ring.



Search Strategy

To find the cyclic compounds described...

Step 1. Draw the structure

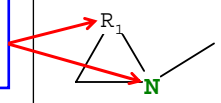
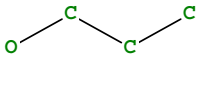
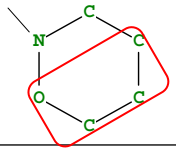
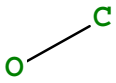
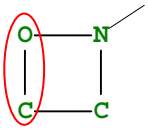
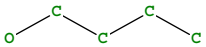
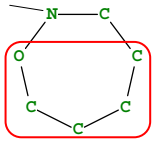
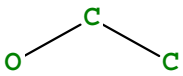
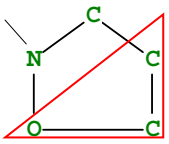
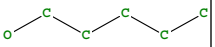
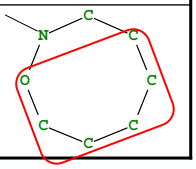
- Draw the smallest ring size desired
- Draw fragments that will account for the larger ring size
- Assign points of attachments to the fragments
- Attach the fragments to the ring via G-group
- Verify the orientation of the fragments

Step 2. Upload the query to STN

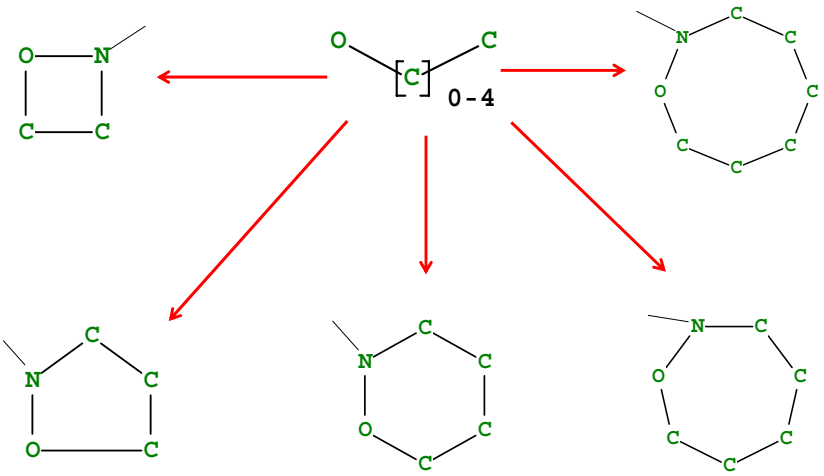
Step 3. Run the search in REGISTRY

- Sample search
- Full search

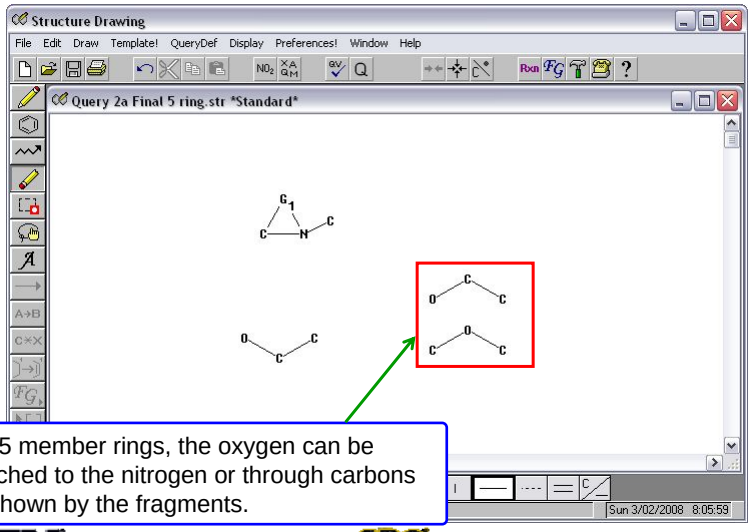
Various fragment lengths result in different ring size

Fragment (R_1)	Resulting Ring	Fragment (R_1)	Resulting Ring
Fragment attachment points			
			
			

Specifying repeating groups in a fragment will save a lot of time

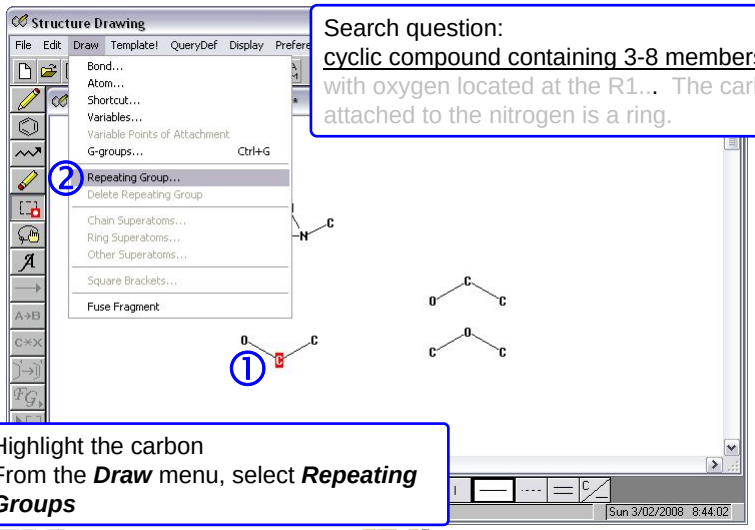


Draw the smallest ring structure and fragments



For 5 member rings, the oxygen can be attached to the nitrogen or through carbons as shown by the fragments.

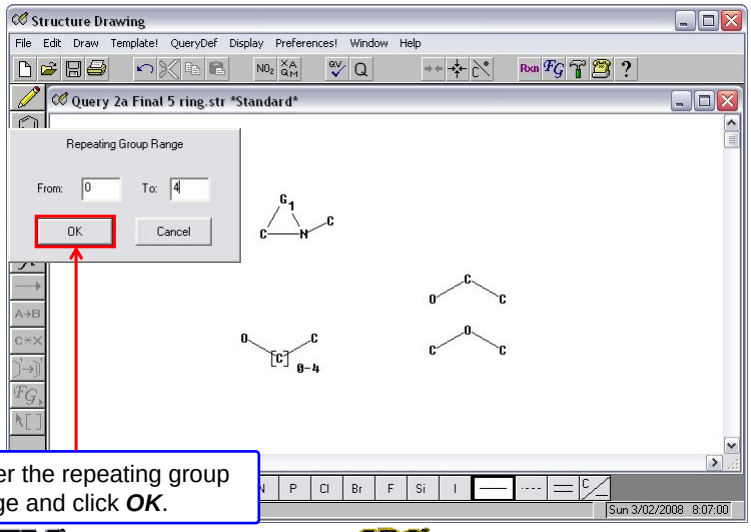
Assign repeating groups to a node



Search question:
cyclic compound containing 3-8 members with oxygen located at the R1... The carbon attached to the nitrogen is a ring.

1. Highlight the carbon
2. From the ***Draw*** menu, select ***Repeating Groups***

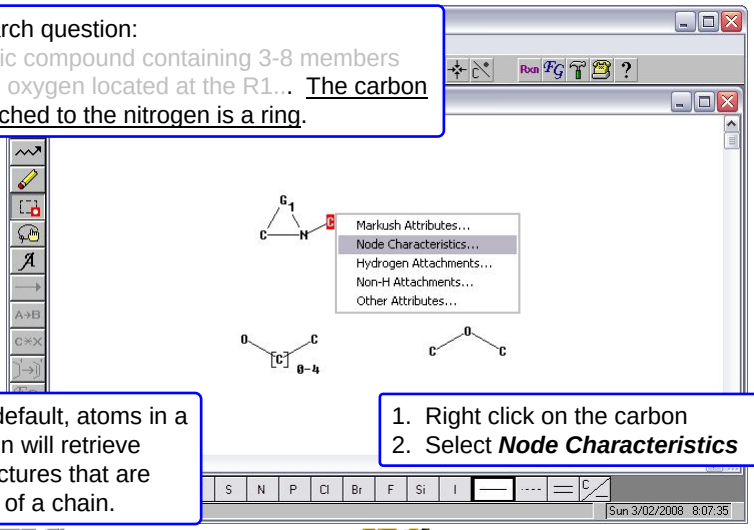
Assign repeating groups to a node



Enter the repeating group range and click ***OK***.

Modify Node Characteristics as needed

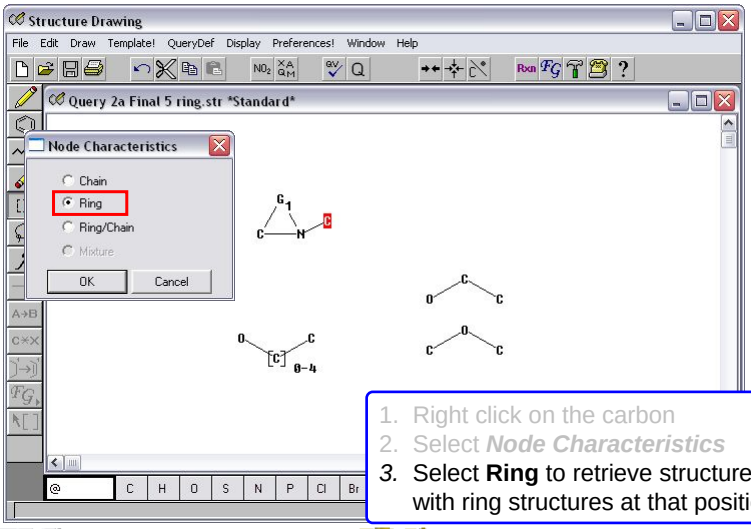
Search question:
cyclic compound containing 3-8 members with oxygen located at the R1... The carbon attached to the nitrogen is a ring.



By default, atoms in a chain will retrieve structures that are part of a chain.

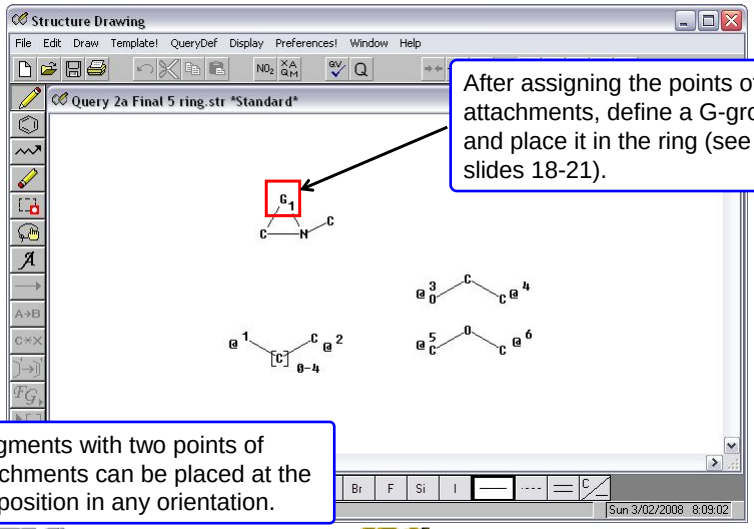
1. Right click on the carbon
2. Select ***Node Characteristics***

By default, atoms in a chain will retrieve structures that are part of a chain



1. Right click on the carbon
2. Select ***Node Characteristics***
3. Select **Ring** to retrieve structures with ring structures at that position

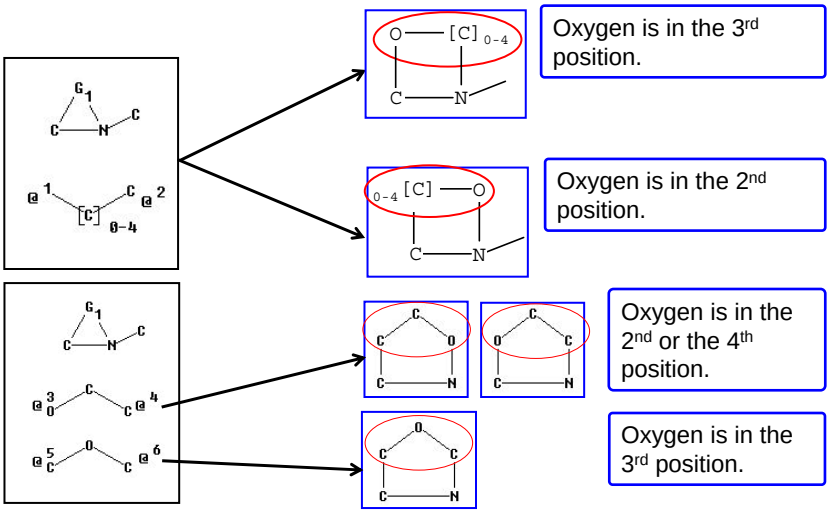
Assign TWO points of attachments to the fragment



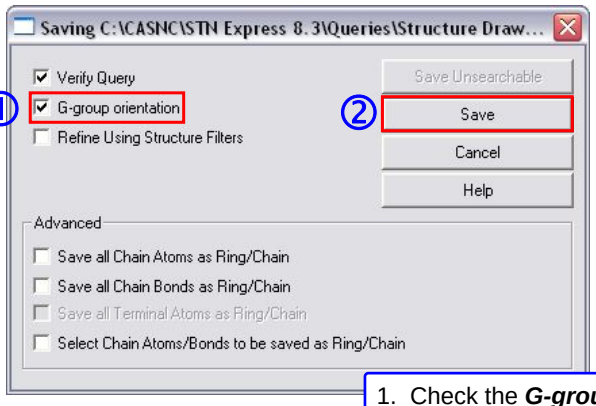
After assigning the points of attachments, define a G-group and place it in the ring (see slides 18-21).

Fragments with two points of attachments can be placed at the G1 position in any orientation.

Fragments with two points of attachments result in two placement orientations

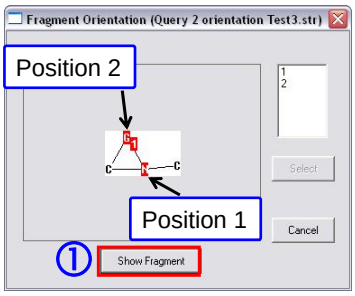


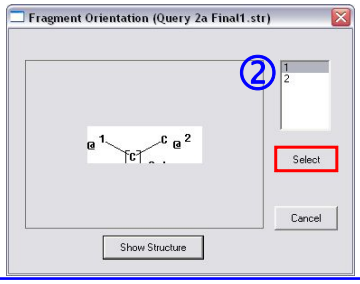
Verify the G-group orientation during a *SAVE* process



1. Check the ***G-group orientation*** box
2. Click ***Save***

Verify the G-group orientation during a *SAVE* process



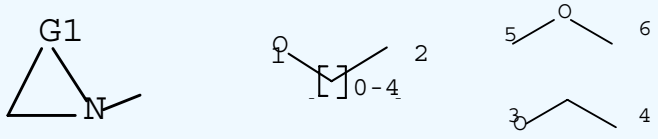


The fragment's points of attachment (@¹ and @²) corresponds to the 1 and 2 in the text box.

1. Click ***Show Fragment***
2. Highlight *1* to indicate that oxygen (@¹) connects to the nitrogen at position 1 and click ***Select***
3. Repeat with other fragments

Upload the structure in REGISTRY

```
=> FIL REG

=> Uploading C:\CASNC\STN Express 8.3\Queries\Structure
   Drawing E-seminar\Query 2\Final\Query 2a\Query 2a Final
   5 ring.str
L1      STRUCTURE UPLOADED

=> D L1 QUERY
L1                STR
```

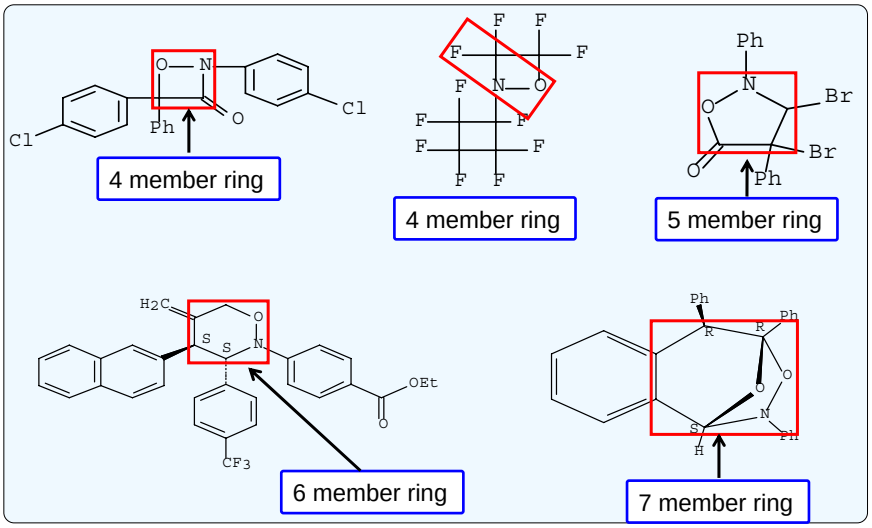



```
 G1  [@1-@2],[@3-@4],[@5-@6]
```

Run the substructure search

```
=> S L1 SSS SAM
SAMPLE SEARCH INITIATED 22:09:02 FILE 'REGISTRY'
SAMPLE SCREEN SEARCH COMPLETED -     8686 TO ITERATE
FULL FILE PROJECTIONS:  ONLINE  **COMPLETE**
                        BATCH   **COMPLETE**
PROJECTED ITERATIONS:       168134 TO   179306
PROJECTED ANSWERS:           39038 TO    44520

L2           50 SEA SSS SAM L1

=> S L1 SSS FULL
FULL SEARCH INITIATED 21:59:41 FILE 'REGISTRY'
FULL SCREEN SEARCH COMPLETED -   175293 TO ITERATE
100.0% PROCESSED   175293 ITERATIONS     42751 ANSWERS

L3        42751 SEA SSS FUL L1
```

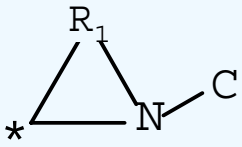

Evaluate structures using D SCAN



Modify the structure and search it as a SUBSET

Extension from Search Question 2:

From the previous answer set, find structures that do not have a carbon or any halogens attached to the carbon indicated by a asterisk.



Search Strategy

To find the cyclic compounds described...

Step 1. Modify the existing structure
- Exclude carbon and halogens as fragments
- Define G2 group

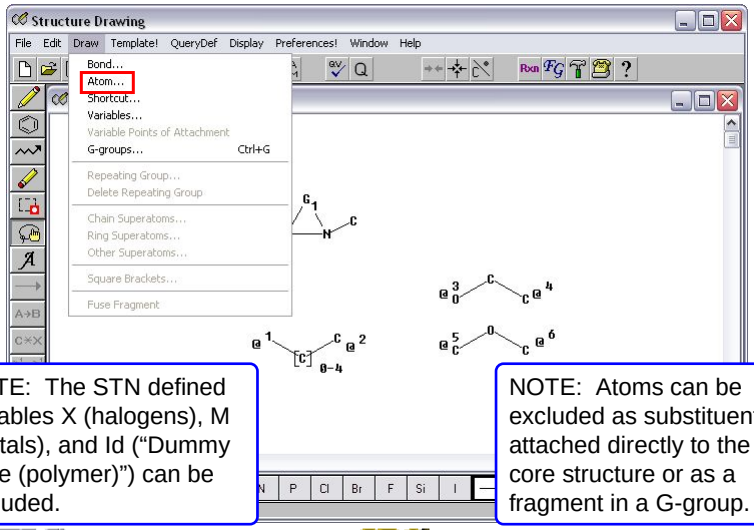
Step 2. Upload the query to STN

Step 3. Run the search in REGISTRY as a subset
- Full search

SUBSET searching is cost effective

- The first structure search resulted in many broad structures
- Uses more specific (narrow) queries that are run against the answers from the broad search
- It is a cost efficient way to add precision to a previously searched query

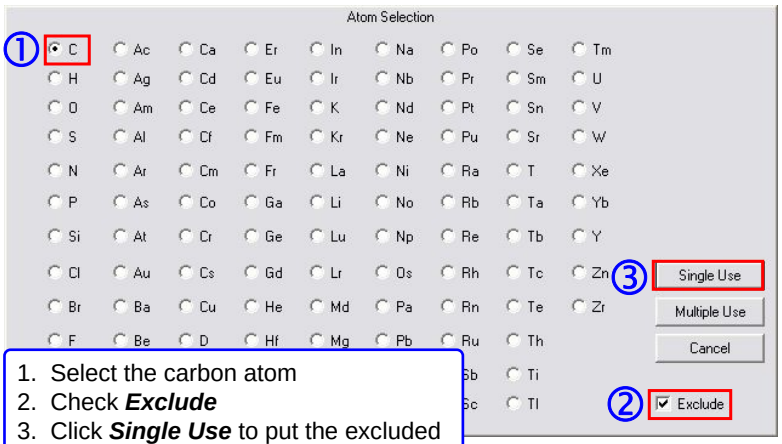
Modify the existing structure by specifying excluded atom(s)



NOTE: The STN defined variables X (halogens), M (metals), and Id ("Dummy node (polymer)") can be excluded.

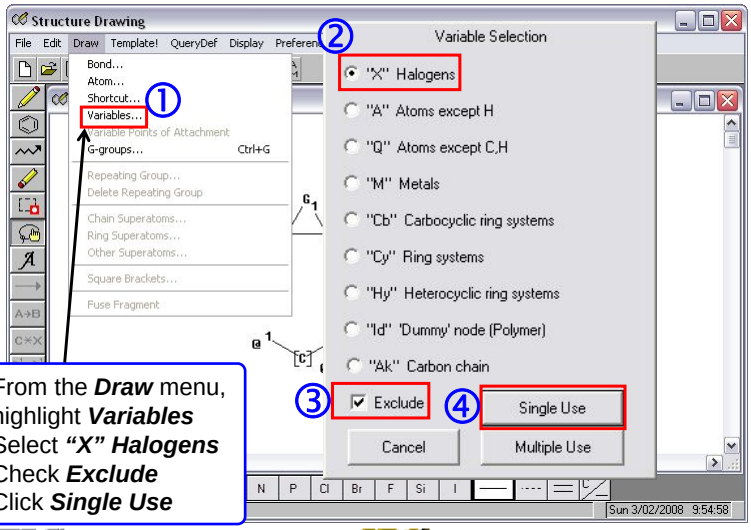
NOTE: Atoms can be excluded as substituents attached directly to the core structure or as a fragment in a G-group.

Atoms and certain variables can be excluded from the structure



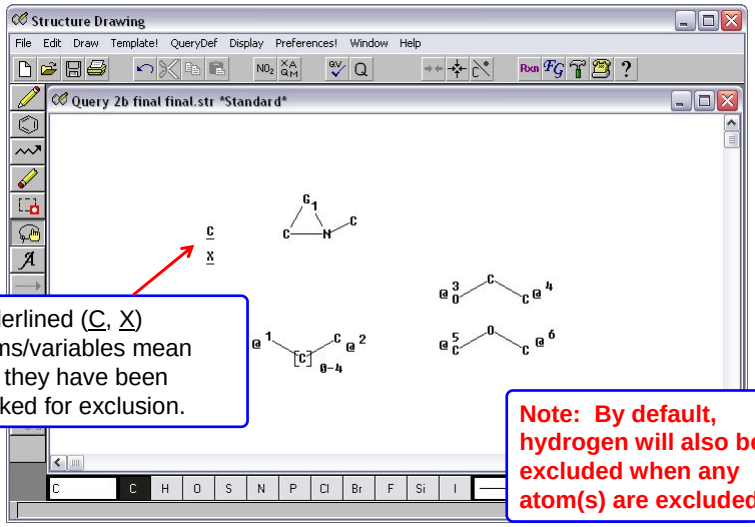
1. Select the carbon atom
2. Check ***Exclude***
3. Click ***Single Use*** to put the excluded atom on the drawing board

Modify the existing structure by specifying excluded variable(s)

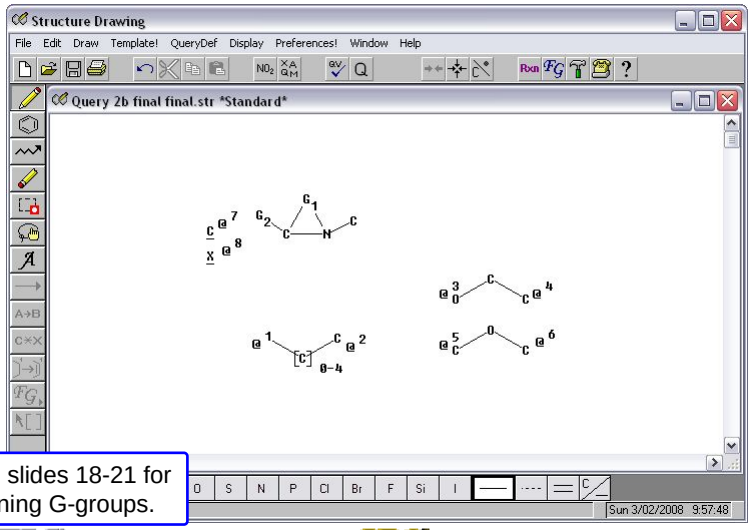


1. From the ***Draw*** menu, highlight ***Variables***
2. Select ***"X" Halogens***
3. Check ***Exclude***
4. Click ***Single Use***

Place excluded atoms/variables as fragment



Define atom and variable fragment as G-group



Upload the modified structure in REGISTRY

```
=> S L1 SSS FULL
L3          42751 SEA SSS FUL L1
```

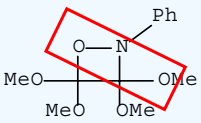
This search was the initial SSS search from slide 39.

```
=> Uploading C:\CASNC\STN Express 8.3\Queries\Structure
   Drawing E-seminar\Query 2\Final\Query 2b-exclusion\Final
   Final\Query 2b final final exclude C and X.str
L4        STRUCTURE UPLOADED

=> S L4 SSS FULL SUBSET=L3
FULL SUBSET SEARCH INITIATED 22:00:16 FILE 'REGISTRY'
FULL SUBSET SCREEN SEARCH COMPLETED -     42751 TO ITERATE
100.0% PROCESSED     42751 ITERATIONS      2926 ANSWERS

L5          2926 SEA SUB=L4 SSS FUL L3
```

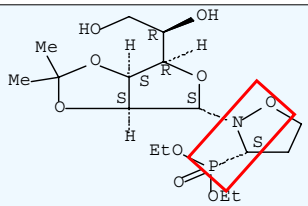
Structures can also be refined using screens/filters. Please see archived CAS e-Seminars on structure screens/filters.

SUBSET searches only in existing answer set.

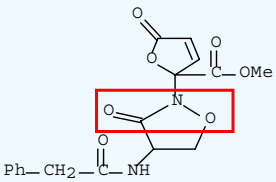
Evaluate structures with D SCAN



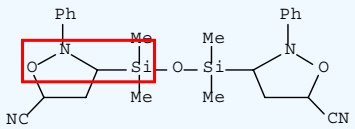
The carbon is attached to oxygen.



The carbon is attached to phosphorous.



The carbon is attached to oxygen.

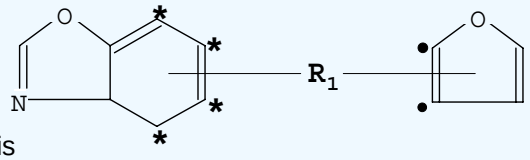


The carbon is attached to silicon.

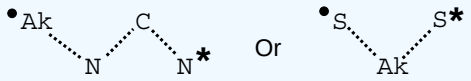
Search Question 3:

What is claimed is:

1. A compound of the formula



Wherein R_1 is



The carbon chain (Ak) can be 1-5 carbons long with various bond types between each node. The fragments are attached by matching the symbols next to the nodes.

Search Strategy

To find the cyclic compounds described...

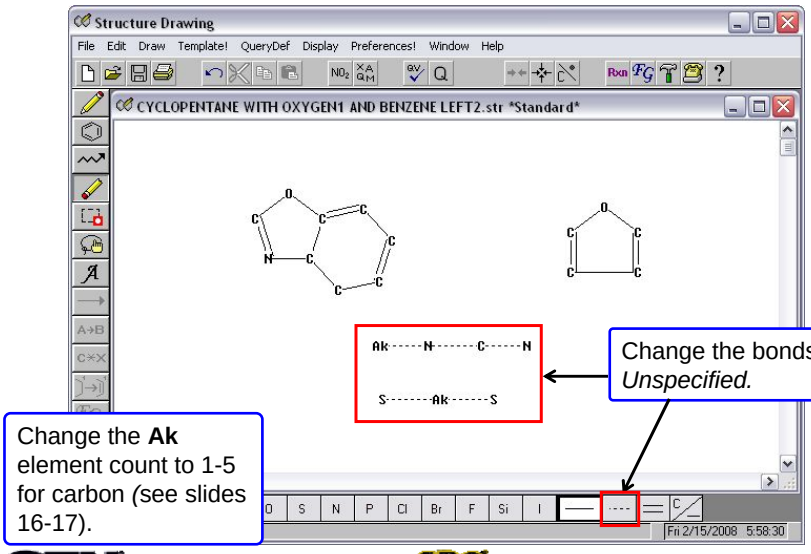
Step 1. Draw the structures

- Build first structure with Variable Points of Attachment to one ring only
- Build second structure with Variable Points of Attachment to the other ring

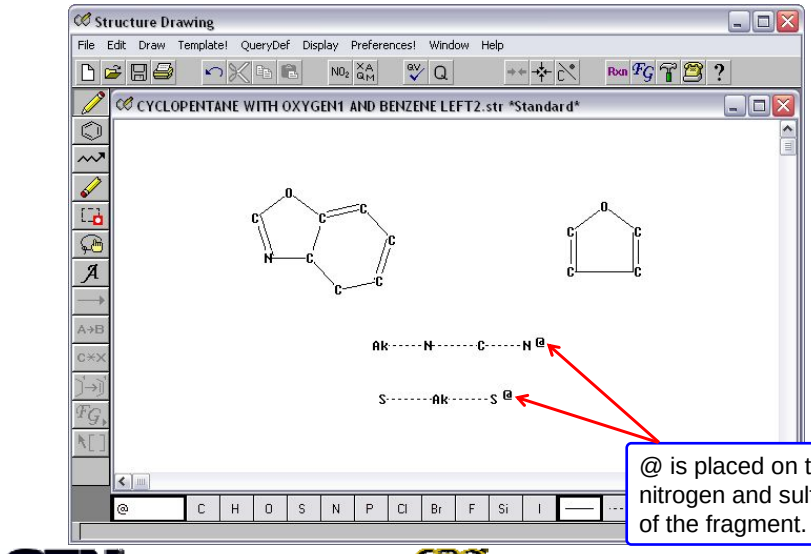
Step 2. Upload both structures to STN

Step 3. Search the first structure and the second structure at the same time using the AND operator

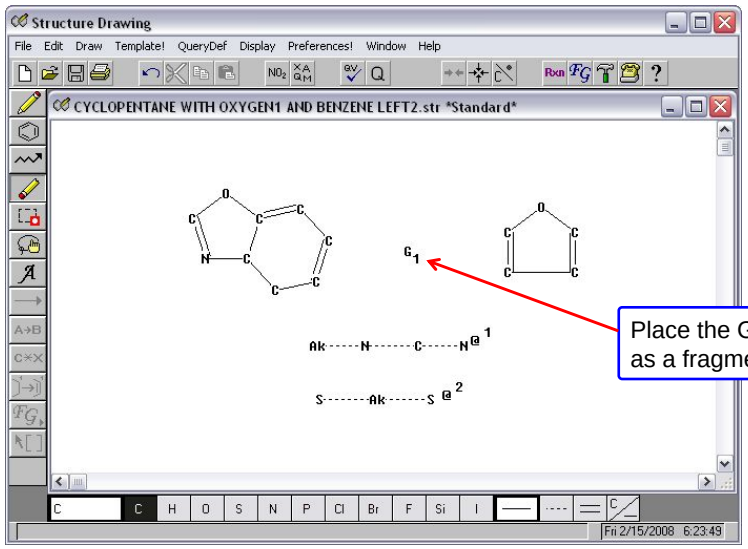
Draw the core structures and fragments



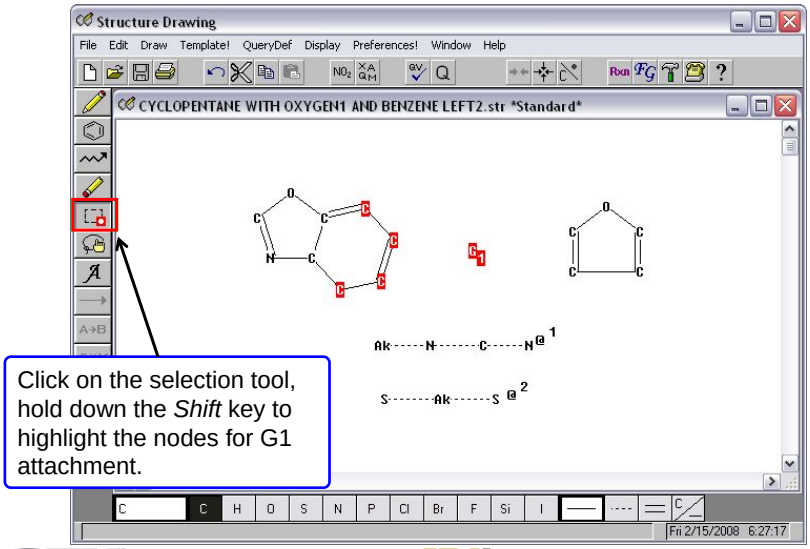
Identify the fragment sides that will attach to the first core structure



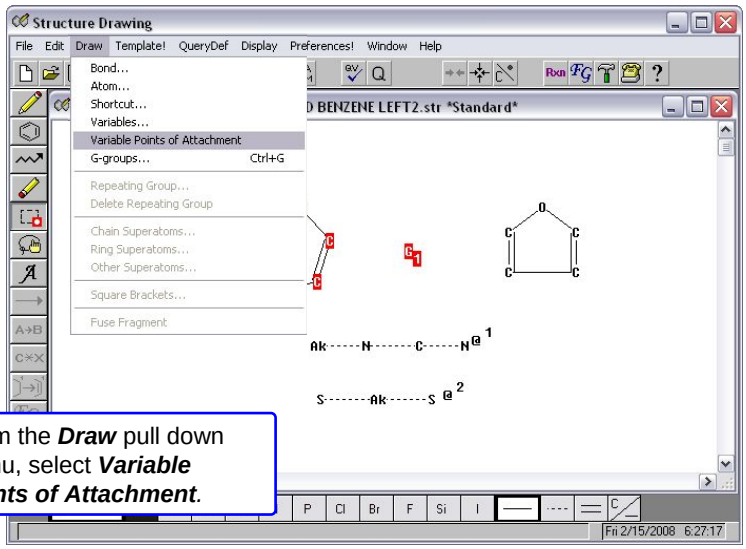
Define the G-group for the first core structure



Select the nodes for the Variable Points of Attachment

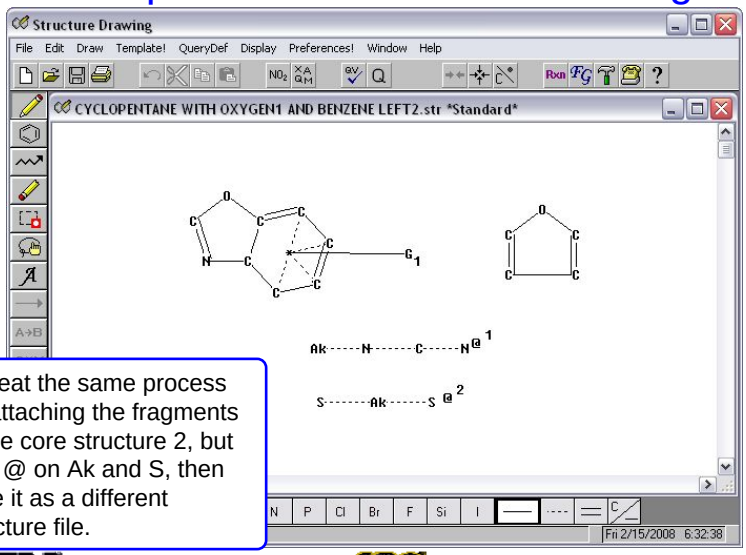


Attach G-group to the highlighted nodes



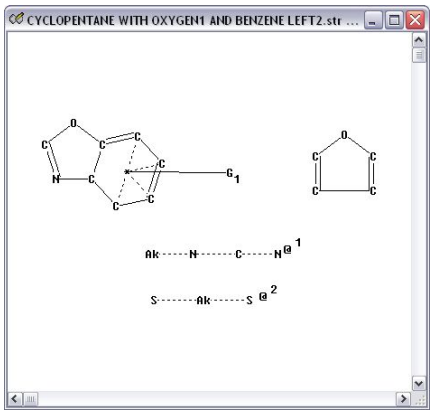
From the ***Draw*** pull down menu, select ***Variable Points of Attachment***.

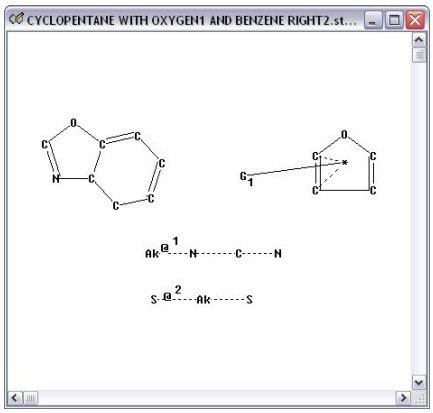
Variable Points of Attachment are illustrated with multiple bonds connect to the G-group



Repeat the same process for attaching the fragments to the core structure 2, but with @ on Ak and S, then save it as a different structure file.

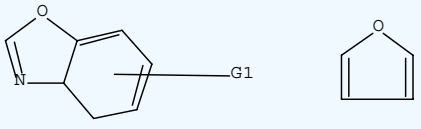
Two files of the same structures but with different fragment orientation attachment

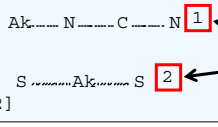




Load each structure separately in REGISTRY

```
=> FIL REG
=> Uploading C:\CASNC\STN Express 8.3\Queries\CYCLOPENTANE
   WITH OXYGEN1 AND BENZENE LEFT2.str
L1      STRUCTURE UPLOADED
=> D L1 QUERY
L1              STR
```



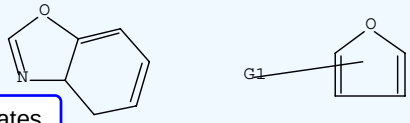



```
G1 [@1], [@2]
```

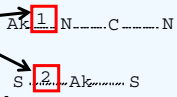
1 and 2 indicates the G1 fragments on structure 1.

Upload the 2nd structure with G-group to REGISTRY

```
=> Uploading C:\CASNC\STN Express 8.3\Queries\CYCLOPENTANE
   WITH OXYGEN1 AND BENZENE RIGHT2.str

L2      STRUCTURE UPLOADED

=> D L2 QUERY
L2              STR
```



1 and 2 indicates the G1 fragments on the second structure.



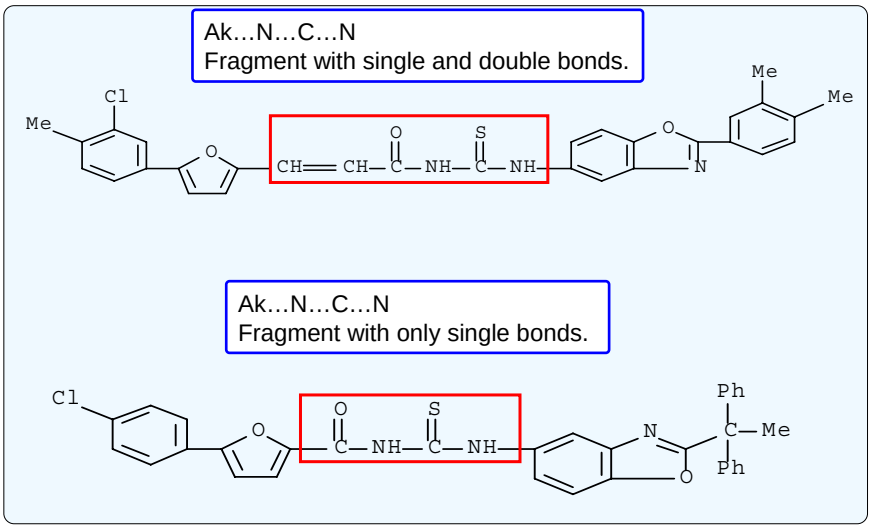

```
G1 [@1],[@2]
```


Use the AND operator to combine both substructure searches

```
=> S L1 AND L2 SSS FULL
FULL SEARCH INITIATED 15:43:27 FILE 'REGISTRY'
FULL SCREEN SEARCH COMPLETED -     7318 TO ITERATE
100.0% PROCESSED     7318 ITERATIONS       432 ANSWERS

L3          432 SEA SSS FUL L1 AND L2
```

Combining **L1** and **L2 SSS FULL** search results in 1 search cost.

Retrieved structures in REGISTRY



Summary

- Structure drawing and searching on STN offers several easy and flexible tools
- Users can modify some of the existing STN defined variables
- User can define specific fragments via G-groups
  - A query may contain up to 20 G-groups
  - A G-group may contain up to 20 substituents

Summary

- Usage of repeating groups allow users to retrieve structures of more than one ring size without having to draw them
- Variable points of attachment makes it easy to indicate fragment attachment on several positions within a ring or multiple rings
- Use subset search to refine a set of broad answers in a cost effective manner

Additional resources

- Short one hour e-Seminars
  - CAS e-Seminars (http://casevents.webex.com)
  - FIZ Karlsruhe e-Seminars (http://www.stninternational.de/training_center/seminars.html)
- STN Virtual Workshops NEW!!!
  - Live training via WebEx
  - Users can practice searching using STN login ID in a virtual setting
- Instructor-led STN workshops

Please visit www.cas.org for more CAS training information.

Please visit http://www.stn-international.com for more FIZ Karlsruhe training information.

Questions and answers...