

A Guide to Adding Screens in STN[®] Structure Searching

December 2010

Copyright © 2010 American Chemical Society
All rights reserved

Table of Contents

Preface	3
Screen Numbers as Search Terms in STN.....	5
Automatic Screen Generation.....	9
The Screen Dictionary.....	13
Search Strategies in STN.....	25
Adding Screens Search Examples	28
Searches Illustrating Addition of Screens That Are Not Generated Automatically by STN.....	29
Searches Illustrating Addition of Screens That Are Conditionally Generated by STN	39
Search Illustrating Addition of Screens That Are Usually Generated Automatically by STN	49
Searches in Which Some of the Requirements Are More Conveniently Specified in Terms of Screens Than in Terms of One or More Structures.....	55
Summary.....	63

A Guide to Adding Screens in STN[®] Structure Searching

Preface

Introduction

Chemical structure searching is available in several STN databases. You may build your structure query (strategy) using STN Express[®], STN[®] on the WebSM, or the online STRUCTURE command.

STN automatically dissects your structure query and derives appropriate “screen numbers” (numeric codes also called filters) that represent the various structural features required by the structure query. The search proceeds automatically in two steps.

Two-Step Structure Search

Step 1 (screening) uses the numeric codes to reduce the number of possible answers for your search by “screening out”, or eliminating from further consideration, those substances in the structure file being searched that do not contain your required structural features and therefore cannot be valid answers. All remaining substances are candidate answers. This step proceeds very quickly.

Step 2 (iteration) then follows screening with an atom-by-atom, bond-by-bond comparison (or iterative search) on the candidate answers that passed the screening step, i.e., those that remain as possibilities after the initial screen out. This two-step procedure is highly efficient in searching very large databases.

Structure Search Limits

Each of the STN structure searchable databases has specific search limits. In most searches, the input structure query usually will cause enough screen numbers to be automatically generated that the search will run to completion within the system limits. Enter HELP SLIMITS at an arrow prompt (=>) to see the limits for your current database.

Adding Screens Manually

You can enhance your search strategy by *manually adding screen numbers* to further define the requirements or increase the efficiency of the “screen out” so that your search will complete within system limits and that the proper answers are retrieved.

This strategy may be necessary if the query structures are so general or so small that they do not automatically eliminate enough candidate answers to allow the search to run to completion within the database limits.

Another reason to manually add screens is if your search question has requirements that cannot be completely defined in terms of one or more drawn structures.

For example, when searching for polymers:

- There is no way to indicate in the structure that the desired substances are to be restricted to polymers.
- There is a screen number that you can manually add to the search strategy to require that all substances retrieved be polymers.

Purpose of This Guide

The purpose of this guide is to explain:

- What types of screens are available
- How you can locate the screen numbers you need in the *STN Screen Dictionary for Structure Searching*¹
- How you can incorporate screen numbers into various search strategies
- How to decide which screen(s) to add to your strategy
- When and where screens cannot be added

Prerequisites

You should already be:

- An experienced STN user
 - Familiar with how to create a structure query
 - Familiar with the basics of STN structure searching
-

¹ Hereafter referred to as the screen dictionary, available at www.cas.org.

Screen Numbers as Search Terms in STN

Introduction

Screen numbers (“screens” or “filters”) in STN are numeric codes that have been assigned to various structural features, e.g., the occurrence of a specified element, group of atoms, sequence of bonds, unusual mass, unusual valence, unusual charge, class of compounds, number and type of rings, etc.

The list of screens with their definitions is found in the screen dictionary.

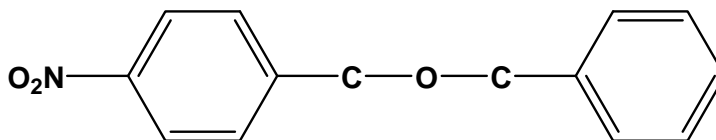
Screen Number Search

When your query is uploaded into STN or you exit STRUCTURE command, a list of screen numbers is automatically generated, representing the structural features of your query. All answers to your query must include the set of screen numbers required by your query, thus eliminating all substances that do not include the screen number set.

In most cases, this step will bring the number of candidate answers within the system limits of the database. This set of candidate answers is then passed to the iteration step.

Example of Screens

For example, if you are searching for all substances containing the substructure



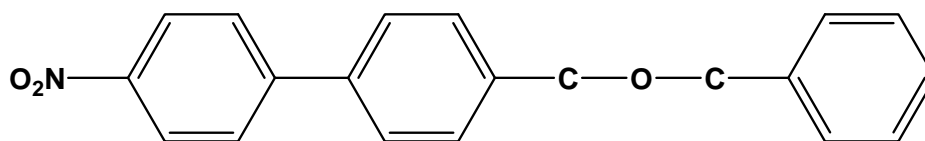
screens would be required for the presence of:

- 14 (or more) carbon atoms
- 1 (or more) nitrogen atom(s)
- 3 (or more) oxygen atoms
- 2 (or more) rings
- 1 (or more) occurrence(s) of a sequence of C C O C C atoms connected by chain bonds.

Additional screens would also be generated by the algorithm.

Iterative Search Substances that pass the screen search are automatically subjected to “iterative search”. This second step is the atom-by-atom, bond-by-bond comparison of the candidate answers from the initial screen search with your input structure. This “iteration” step vastly improves the precision of the search.

Every substance record that passes the initial screen search step will contain, somewhere in the structure, all of the features that could be described by one or more screens. However, they may not have the exact relationship that you desire. As a result, the undesired structure



could also be represented by the screens noted in the example above. (The structure may not have passed the screen search based on the exhaustive list of screens generated automatically by STN, but it is used here as an illustration.) However, this structure would not pass the atom-by-atom, bond-by-bond comparison because it does not meet all of the requirements expressed in the input structure.

Dropped Screens Screens with a higher frequency of occurrence within the database are removed automatically from the list, since they are less effective in eliminating invalid candidates. This makes the process more efficient.

Structure Search Process The steps you must take to carry out a structure search are explained in the table below.

Step	Action
1	Build a query using STN Express or STN on the Web.
2	Upload the query to an STN structure database.
3	Initiate a SAMPLE search, if available.
4	Examine results and full-file projections.
5	Add screens, if necessary, using software or the SCREEN command.
6	Initiate the full-file search (the screen search followed by the iterative search).
7	Review your results.

Types of Screens

STN uses 12 types of screens to represent structural features:

Screen Type	Function
(AA) Augmented Atom	Describes a central atom (any non-hydrogen atom) and its environment, i.e., the non-hydrogen atoms that are directly attached to the central atom.
(HA) Hydrogen Augmented Atom	A special type of augmented atom screen describing a central atom and its attachments, including the exact number of hydrogen atoms.
(TW) Twin Augmented Atoms	Analogous to HA screens but are more specific in which they allow the exact number of hydrogens on two adjacent atoms to be specified.
(AS) Atom Sequence	Represents a sequence of four to six atoms within the structure and may include the type of bonds in that sequence.
(BS) Bond Sequence	Represents the exact bond type and bond value in a sequence of four to six atoms. The atoms in the sequence are undefined (the symbol "A" is used for all atoms).
(CS) Connectivity Sequence	Represents the exact number of non-hydrogen atoms attached to central atoms in a sequence of four to six atoms. Bond types between the atoms in the sequence may be specified.
(AC) Atom Count	Describes the minimum number of non-hydrogen atoms in a substance.
(DC) Degree of Connectivity	Represents the minimum number of non-hydrogen attachments to a central atom.
(RC) Ring Count	Describes the minimum number of ring systems present.
(TR) Type of Ring	Describes the nature of rings atoms, i.e., whether they are fused (attached to more than two other ring atoms) or unfused (attached to only two other ring atoms).
(EC) Element Count	Specifies the minimum number of occurrences of an element in a substance.
(GM) Graph Modifier	Represents unusual structural features (attributes) or classes of compounds.

Screen Number Frequency

The screen dictionary includes all these various types of screens and their frequency of occurrence in the CAS REGISTRYSM database as of May 2005.

Screen List Is Comprehensive

The list of screens associated with the record of each substance in STN is comprehensive. This means that there is often some “overlap” in screens.

For example, a six-atom sequence described by AS, BS, or CS screens may also be described in part by AS, BS, or CS screens for various four- and five-atom sequences that are embedded within that six-atom sequence.

Records for substances with AC 10 (or more) atoms also contain screens for AC 9 (or more), AC 8 (or more), etc. Similarly substance records that contain EC 2 (or more) oxygen or RC 2 (or more) also contain EC 1 (or more) oxygen and RC 1 (or more).

Overlapping Screens

Multiple AA screens may exist for a central atom and given set of attached atoms. The AA screens may exist with both bond type and bond value specified, with only bond type specified, or with neither bond type nor bond value specified. Sometimes an AA screen with bond types and bond values specified, but not describing all of the attached atoms, will have a lower frequency of occurrence than another AA screen that describes all of the attached atoms but does not specify bond types and values.

Lower Frequency Screens Preferred

In general, it is preferable to add a screen with lower frequency of occurrence to the search strategy because it will be more effective in eliminating undesirable candidates. Remember this when manually adding screens to your strategy.

Further Information About Screens

Additional information about these screens will be given in the context of the examples throughout this document. The screen type abbreviations shown in parentheses in the list on the previous page will be used throughout the remainder of this text to refer to a particular type of screen.

Automatic Screen Generation

Introduction

STN automatically derives appropriate screens from your structure. This automatic generation of screens minimizes the number of situations in which you will need to add screens manually to the search strategy. The following chart illustrates when the various types of screens are generated depending on the type of search – SUBSTRUCTURE, EXACT, or FAMILY – specified by the user in the SEARCH command.

Automatic Screen Generation Table

	Screen Type	Substructure Search	EXACT or FAMILY Search
(AA)	Augmented Atoms	Yes	Yes
(HA)	Hydrogen Augmented Atoms	Conditional	Conditional
(TW)	Twin Augmented Atoms	No	No
(AS)	Atom Sequence	Yes	Yes
(BS)	Bond Sequence	Yes	Yes
(CS)	Connectivity Sequence	No	Yes
(AC)	Atom Count	Yes	Yes
(DC)	Degree of Connectivity	Yes	Yes
(RC)	Ring Count	Yes	Yes
(TR)	Type of Ring	Conditional	Yes
(EC)	Element Count	Conditional	Conditional
(GM)	Graph Modifier	No	No

Automatically Generated Screens

As shown in the table, the majority of screens are automatically generated by STN. As a result, you will seldom incorporate these screens, i.e., AA, AS, BS, AC, DC, RC, into your search strategies.

Manually Added Screens

The AA, AS, BS, AC, DC, and RC screens can still be useful as manually added search terms in very broad (generic) searches, e.g., for structures containing multiple variable atoms. This is especially true if the generated screen was dropped because of higher frequency of occurrence.

Variable Atoms and Bonds **Variable atoms** (either system-defined or user-defined) **are ignored** during automatic screen generation. In small structures, this could result in sequences of three or fewer fully defined atoms for which no AS or BS screens could be generated. Similarly, the AA screens that are generated would be those with fewer attached atoms and higher frequencies of occurrence.

Automatic generation of AA, AS, and BS screens can be similarly affected when bonds are left unspecified or specified as ring or chain.

Conditionally Generated Screens Automatic generation of the HA, TR, and EC screens is conditional. That is, they may or may not be generated depending on how you have defined the structure to be searched.

HA Screens When you have specified *exact hydrogen counts* and *exact bonding*, appropriate HA screens will be generated. When you do not specify hydrogen counts, no HA screens are generated.

EC Screens EC screens are generated automatically for all atoms in the input structure except:

- H
- System variables A, Q, M, X, Id
- User-defined variables G1-G20
- Single-atom fragments (fragments of multicomponent structures that consist of a single atom).

Hydrogen atoms are not connection table atoms in STN, so that the Br in the HBr component of a salt is a single-atom fragment (SAF).

TR Screens In a substructure search, TR screens are generated only when you have specified that the ring system(s) is (are) isolated. When you allow for embedment or fusion of the input structure in larger ring systems, no TR screens are generated. However, TR screens are automatically generated in EXACT or FAMILY searches, since by definition they do not allow further substitution.

**Manual
Addition of TR
Screens**

Manual addition of these conditionally generated screens may be useful in some search strategies, for example, a search in which some rings may be embedded within larger ring systems while others must be isolated could benefit from the manual addition of TR screens.

**Screens Not
Generated
Automatically**

CS screens are not generated in substructure searches, since STN considers all sites open for additional substitution.

In EXACT or FAMILY searches, no further substitution of the input structure is allowed by definition and CS screens are generated.

Manual addition of CS screens may be useful in substructure searches in which you desire to prevent (or require) substitution at certain locations.

TW Screens

The few TW screens that exist describe the exact number of hydrogen atoms in CH₃, OH, SH, and NH₂ groups as well as the exact number of hydrogens on the carbon atom to which these groups are attached. These TW screens are never automatically generated by STN. Manual addition of TW screens is only rarely useful.

GM Screens

GM screens that are usually not generated automatically describe various special structural features (often called attributes) such as:

- Mass, valence, and charge
- Stereochemical data
- Numbers of components in multicomponent substances
- Classes of substances such as alloys, coordination compounds, and polymers

The GM screens for numbers of components are automatically generated in EXACT searches for multicomponent substances.

Manual addition of GM screens to search strategies is often very useful.

When To Add Screens to Your Search Strategy

Since STN generates an exhaustive list of screens for you, it is not necessary for you to routinely add screen numbers to your search strategy. There are three common situations in which you might consider manual addition of screens:

- Add screens as part of the initial search strategy because not all of the search requirements can be explicitly expressed in terms of one or more structures (e.g., only alkyl derivatives of a given structure).
- Add screens after a SAMPLE search in response to a full-file projection that indicates that the system threshold will be exceeded and the search will not run to completion.
- Add screens to improve the precision of the search (e.g., you are retrieving too many undesired answers, e.g., polymers, when you desire only non-polymeric substances).

Further information on when to add screens to your search strategy is found in the examples cited later in this guide.

The Screen Dictionary

How To Find Screens

To find the screen numbers to add to the search strategy, you need to consult the screen dictionary. The introduction of the screen dictionary provides a detailed description of its content and use.

The various types of screens used in the examples in the next section of this guide are also described as they appear in the screen dictionary.

Screen Dictionary Arrangement

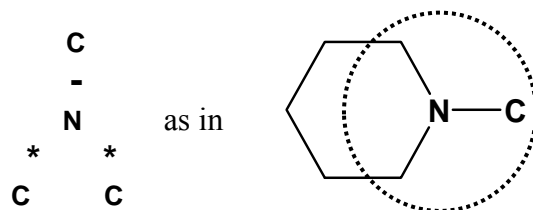
The screen dictionary is an ordered listing of the screens that are used in STN and it consists of:

- An alphabetically ordered section arranged by types of screens in the order AA, HA, TW, AS, BS, CS, AC, DC, RC, TR, EC, and GM
 - A numerically ordered section arranged in order of increasing screen numbers
-

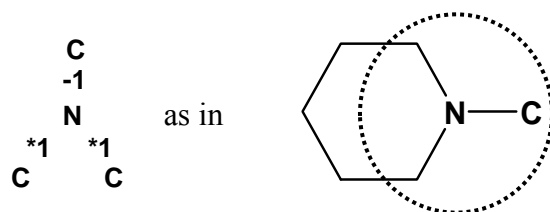
**Screen
Dictionary
Format**

The general format of a page from the alphabetically arranged section of the screen dictionary is illustrated in the following pages using a page from the Augmented Atom (AA) section. The horizontal lines drawn after every fifth screen number are for visual clarity.

- The first (left-hand) column shows the screen number. There are 2127 screen numbers.
- The second column indicates the type of screen using a two-letter abbreviation. In this example, AA indicates Augmented Atom.
- The third column indicates the minimum number of occurrences of the fragment in the structure. If no number appears in this column, the number of occurrences is one (or more). If a two appears, as in the row for screen 1603, the fragment must occur in the structure two or more times.
- The fourth column describes the fragment using linear notations and ordering of symbols. For example, screen 1607 (see ==> in the example) represents a structural fragment in which a nitrogen atom is bonded to a carbon atom by a ring bond (*), to a second carbon atom by a ring bond, and to a third carbon atom by a chain bond (-), i.e.,



Screen 1608 (see the example) represents the same structural fragment with the additional specification that the bonds are single ring (*1) or single chain (-1) bonds,



In the fragment definition for AA screens, the central atom is always listed first followed by the attached atoms in alphabetic order (see screens 1614 and 1615 for an example).

**Fragments
Sharing Screen
Numbers**

A number (or pound) sign (#) in the fifth column indicates that two or more fragments share a screen number. Sharing of screen numbers allows STN to use 2127 screens to describe over 5500 structural features. Structural features that share screens sometimes occur adjacent to one another in the alphabetically ordered section of the screen dictionary [see screen 1613 (***) for an example].

**Special Atom
Symbols Used
in the Screen
Dictionary**

Atoms in the screen dictionary are represented by the standard element symbols, e.g., C for carbon, N for nitrogen, Br for bromine, etc. Special atom symbols used in the screen dictionary are:

Special Atom Symbols

X = Halogen
A = Any Atom
M = Metal

Note that only three variable atom symbols are used in the screen dictionary.

**Bond Symbols
Used in the
Screen
Dictionary**

The bond symbols specify both bond type (* for ring bonds, - for chain bonds) and bond value (e.g., **1** for a single bond, **2** for a double bond, etc.). Thus, ***3** represents a triple ring bond and **-2** indicates an exact double chain bond. The normalized (***4** or **-4**) bonds represent aromatic, tautomeric, or delocalized bonds, i.e., resonating bonds in rings or chains, respectively. A blank represents any type of bond with no value specified.

Bond Symbols

* = Any Ring Bond
*1 = Single Exact Ring Bond
*2 = Double Exact Ring Bond
*3 = Triple Ring Bond
*4 = Normalized Ring Bond

- = Any Chain Bond
-1 = Single Exact Chain Bond
-2 = Double Exact Chain Bond
-3 = Triple Chain Bond
-4 = Normalized Chain Bond

“blank” = Any Bond

From the Augmented Atom Section of the Screen Dictionary

Multiple Use of Screen Number

Screen Number	Screen Type	Number of Occurrences (if >1)	Fragment Definition						Frequency (%)
1602	AA		N	C	C	C			31.96
1603	AA	2	N	C	C	C			10.54
1604	AA	3	N	C	C	C			2.71
1605	AA		N	* C	* C	* C			3.71
1606	AA	2	N	* C	* C	* C			0.51
==>1607	AA		N	* C	* C	- C			22.55
**1608	AA		N	*1 C	*1 C	-1 C			21.59
1609	AA		N	- C	- C	- C			8.91
1610	AA	2	N	- C	- C	- C			1.24
1611	AA		N	C	C	C	C		1.17
1612	AA	2	N	C	C	C	C		0.16
***1613	AA		N	* C	* C	* C	* C	#	0.35
***1613	AA		N	* C	* C	* C	- C	#	0.35
***1613	AA		N	* C	* C	- C	- C	#	0.35
1614	AA		N	- C	- C	- C	- C		0.83
1615	AA		N	C	C	C	N	#	0.08
1615	AA		N	C	C	C	O	#	0.08
1615	AA		N	C	C	C	S	#	0.08
1616	AA		N	C	C	N			5.31
1617	AA		N	* C	* C	* N		#	0.97

General Approach

The most general approach to determine what structural features share a particular screen number is to consult the Screen Number Order section of the screen dictionary. For example, screen 1617 is shared by the following structural fragments.

From the Screen Number Order Section of the Screen Dictionary

Screen Number	Screen Type	Fragment Definition	Freq. %
1617	AA	N * C * C * N	0.97
	AA	N * C * C * O	
	AA	N * C * C * S	

Frequency of Occurrence

Finally, the frequency of occurrence of the screen number in CAS REGISTRY is based on the approximately 23.5 million structures in the database in May 2005. It is shown in the right-hand column in both the alphabetically and numerically ordered sections of the screen dictionary. This number gives an indication of relative frequency of occurrence of structural features in CAS REGISTRY. The frequency in other structure databases may differ.

Screen Sequence in the Alphabetically Ordered Section

Some screens, i.e., AC, DC, RC, TR, EC, GM, are listed in the screen dictionary in increasing numerical or alphabetic order of the screen definition.

From the Atom Count Section of the Screen Dictionary

Screen Number	Fragment Definition	Freq. %
1899 AC 1	} increasing numerical order	98.15
1900 AC 6		92.64
1901 AC 8		91.95
1902 AC 10		90.59
1903 AC 12		88.21
1904 AC 14		84.50
1905 AC 16		79.75

From the Type of Ring Section of the Screen Dictionary

Screen Number	Fragment Definition	Freq. %
1848 TR DDD	} alphabetical order	2.06
1849 TR 2 DDD		0.21
1850 TR DDT		0.26
1851 TR DTT		# 2.46
1851 TR TTT		# 2.46
1852 TR DDDD		0.97
1853 TR DDDT		2.17

CS Screens

CS screens are first listed in increasing numerical order (considering individual digits in turn from left to right). Thus, in the example shown below, screen 548 is listed before screen 575 since at the first point of difference (from the left) 1 - 3 * 3 * 2 * 3 - 4 comes before 1 - 3 * 3 - 3.

When two or more screens have the same numerical order, then they are arranged by bond type in the order unspecified before ring before chain, i.e., (blank) before * before -. Thus in the example that follows, screen 575 (see ==>) occurs before screen 576 (see **) because * is ordered before - at the first point of difference from the left.

From the Connectivity Sequence Section of the Screen Dictionary

Screen Number		Fragment Definition		Freq. %
573	CS	1 - 3 * 3 * 2 * 3 - 2		6.02
574	CS	1 - 3 * 3 * 2 * 3 - 3		5.24
548	CS	1 - 3 * 3 * 2 * 3 - 4	#	3.16
==>575	CS	1 - 3 * 3 - 3		15.42
**576	CS	1 - 3 - 3 * 3		18.41
577	CS	1 - 3 - 3 - 3		9.33
578	CS	1 - 3 * 3 * 3 - 1		9.40

Hierarchical Arrangement

AA, HA, TW, AS, and BS screens are listed in a hierarchical arrangement:

- Increasing alphabetic order of element symbols (considering each element symbol individually from left to right)
- Bond type [unspecified (blank) before * before -]
- Bond type and value [unspecified (blank) before * before *1 before *2 before *3 before *4 before - before -1 before -2 before -3 before -4]

Hierarchical Ordering**ALPHABETIC**

```
C C C C
C C C N|
C C N
C C N N
```

expanded →

BOND TYPE

```
C C C N
C * C * C * N
C * C * C - N
C * C - C * N|
C - C - C - N
```

expanded →

BOND TYPE & VALUE

```
C *1 C -1 C *1 N
C *1 C -1 C *2 N
C *1 C -2 C *1 N
C *2 C -1 C *1 N
C *4 C *4 C *1 N
• C *4 C *4 C *4 N
• C *4 C *4 C -1 N
C *4 C -1 C *4 N
```

Hierarchical Ordering

In this example, AA screens would first be alphabetized (left to right) as shown in the left-hand column. Considering only screens with the definition C C C N, additional screens with this definition would next be arranged by bond type as shown in the center column. Finally, additional screens with the definition C*C-C*N would be arranged by bond type and bond value as shown in the right-hand column.

Note that, because the final ordering is by both bond type and bond value, screens that contain a sequence other than C*C-C*N may occur (see ●) between screens that have that sequence. If several screens have the same fragment definition, they are arranged by increasing number of occurrences.

**Double Posting
of AS, BS, and
CS Screens**

When the fragment definitions for AS, BS, and CS screens are not symmetrical, the fragment definitions are “double-posted”. That is, both the “forward” and the “reverse” definitions appear in the screen dictionary. This is done as an aid to the user.

From the Atom Sequence Section of the Screen Dictionary

Screen Number		Fragment Definition	Freq. %
432	AS	N * C * N * N	3.24
433	AS	N * C - N - N	0.55
434	AS	N - C * N * N	1.27
435	AS	N - C - N - N	0.73
	•		
	•		
	•		
432	AS	N * N * C * N	3.24
434	AS	N * N * C - N	1.27
433	AS	N - N - C * N	0.55
435	AS	N - N - C - N	0.73

The SCREEN Command

The SCREEN command is used to create sets of screen numbers that you wish to use to describe a desired substructure, attribute, class of substances, or other search requirement. The SCREEN command provides for the input of either a single screen number or a series of screen numbers connected by a single type of Boolean operator (AND or OR) in an unnested expression, i.e., parentheses are not allowed. The NOT operator cannot be used in the SCREEN command.

QUERY and SEARCH Commands

The QUERY or SEARCH commands may be used to combine screen set L-numbers and structure L-numbers using any of the Boolean operators. If desired, you may exclude a structural feature represented by a screen using the NOT operator in the QUERY or SEARCH command. For example, the query logic may include

Structure query AND screen set 1 NOT screen set 2

where screen set 1 and screen set 2 may contain multiple screen numbers combined with AND or OR operators.

Note: Always use the NOT operator with caution and discretion since NOT logic is absolute. There is a possibility that some relevant answers may be inadvertently lost because the structural feature represented by the screen excluded with NOT logic occurred in an unexpected location in an otherwise desirable substance. Also be aware that screen numbers may be shared and the use of screens with NOT logic may eliminate a shared structural feature you may not have considered. Remember to check the numeric screen number list to see if a screen number is shared.

**Structure and
Screen Queries**

In QUERY or SEARCH statements that contain both structures and screens, the structures and screens are interpreted separately by STN. For example, the logic statement

Structure A OR structure B AND screen C NOT screen D

is interpreted, in effect, in three steps:

- (1) Structure A OR Structure B
- (2) Screen C NOT Screen D
- (3) (1) AND (2)

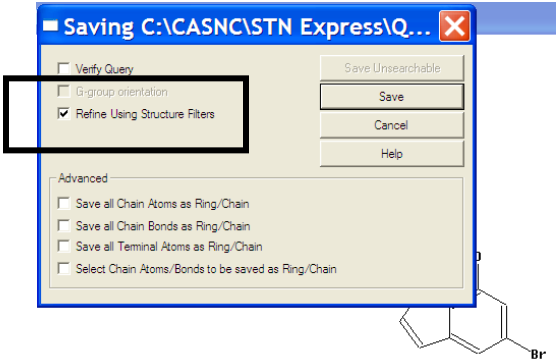
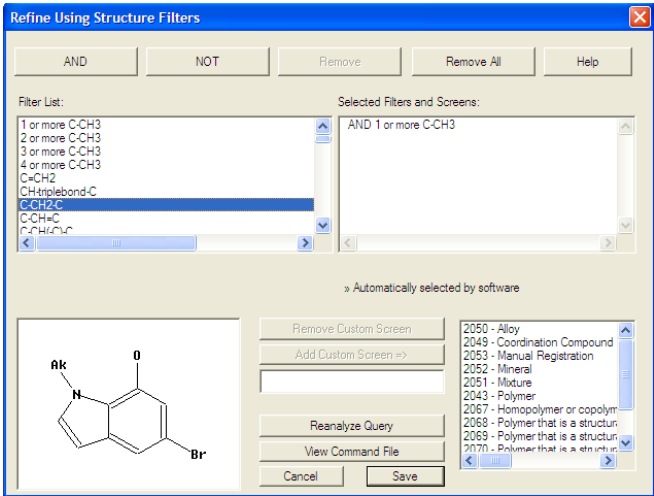
Although it is not required in formulating such search strategies, it may be helpful for you to use parentheses to explicitly separate the structures from the screens, i.e.,

(Structure A OR structure B) AND (screen C NOT screen D)

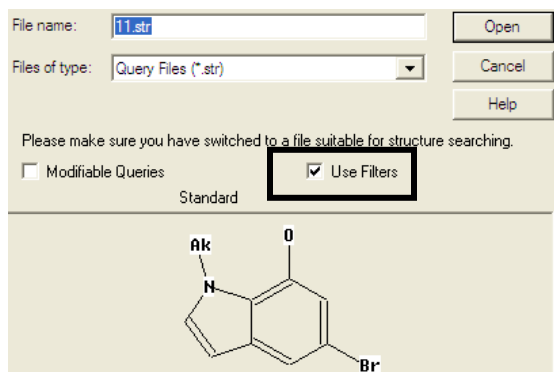
Since structures and screens are interpreted separately, a screen (or set of screens) does not apply to just one structure if a query contains several structures. The screens *must* apply to all structures in the query.

Adding Screens: STN Express or STN on the Web

STN Express and STN on the Web support the addition of screens with the *Refine Using Structure Filters* feature. This feature may be invoked as part of the *Save* process as follows:

Step	Action
1	<p>Save the query, and choose <i>Refine Using Structure Filters</i>.</p> 
2	<p>Select a screen from the <i>Filter List</i>, and then click AND or NOT. The chosen filter will appear in the <i>Selected Filters and Screens</i> window.</p> 
3	<p>Repeat the process for all screens you wish to add. Click <i>OK</i>. For more information consult the <i>STN Express User Guide</i>.</p>

Note: The selected screens will be automatically uploaded with your query if you choose *Use Filters* during the upload process.



A search query combining the drawn query and the selected screens will be automatically generated.

```
=> ....Testing the current file.... screen  
ENTER SCREEN EXPRESSION OR (END):end
```

```
=> screen 963 AND 1398 AND 1945 AND 2004  
L1      SCREEN CREATED
```

```
=> screen 2050 OR 2049  
L2      SCREEN CREATED
```

```
=>  
Uploading C:\Program Files\STN  
Express\Queries\11.str  
L3      STRUCTURE UPLOADED
```

```
=> QUE L3 AND L1 NOT L2  
L4      QUE L3 AND L1 NOT L2
```


Search Strategies in STN

Introduction

The examples in this section illustrate various search requirements in which the manual addition of screen numbers can be used to refine a search strategy initially based on structures alone.

Structure Search Process

Earlier we noted that a structure search in STN is a two-step process. First, the system conducts a “screen search” based on the screen numbers automatically generated from all of the structures in the search query. The candidate answers from this preliminary screening then undergo an “iterative search,” performing an atom-by-atom, bond-by-bond comparison with the structure(s) you required.

Search Limits and Adding Screens

If a large number of candidate answers pass the screen search, then STN must perform a large number of iterations. This search becomes less efficient. In some cases the input structure is so general that the number of iterations that must be performed exceeds the system threshold limits. As a result, the search will not run to completion in the full file. This will be shown in the “Full File Projections” obtained from the SAMPLE search.

The efficiency of such searches can often be improved by adding to the search strategy those higher frequency screens that are likely to be dropped prior to screening or those that are not generated automatically.

For details on structure search limits, enter HELP SLIMITS at an arrow prompt (=>) in the structure search file.

Special Considerations

In some searches, for example when looking for a class of substances, not all of your requirements can be defined in terms of one or more structures. The addition of screen numbers may be the only way to specify your requirements.

Some search strategies may appear to require that many structures be built. It may be possible to minimize the building of multiple structures by manual addition of appropriate screens. This approach may involve sacrificing some specificity of search requirements as fewer structures are built and may result in a decrease of the precision of the search.

**Search
Precision**

The precision of some searches may be improved by more complex strategies involving the addition of multiple screens. These applications are useful to experienced and advanced searchers. You may find it useful to add screens to your search strategy when screens are necessary to:

- Complete the specification of the search requirements
 - Revise the strategy after a SAMPLE search has indicated that system thresholds will be exceeded in a full-file search
-

**When To Add
Screens
Manually**

Some of the common situations in which you may want to consider the manual addition of screens to your search strategy are:

- When your structure is “small”, especially if it contains only commonly occurring atoms (carbon, hydrogen, oxygen, and nitrogen)
 - When your structure contains variable atoms and/or undefined bonds because you cannot explicitly define the structure
 - When your search requirements will not allow you to isolate all ring systems
 - When you are searching for a class of substances
 - When a single structure contains multiple fragments
-

**Low-Frequency
Screens**

Addition of one or two lower frequency screens representing a required structural feature is usually very effective in increasing the efficiency of the initial screen search and the effectiveness of the overall search.

Many searchers prefer to add lower frequency screens for required structural features rather than eliminating numerous screens representing undesired features with NOT logic. (Review the caution regarding the absolute nature of NOT logic.) However, in some cases, such as searches for hydrocarbons or alkyl esters, it is necessary to adopt the approach of elimination of screen sets containing numerous screens. Both approaches are illustrated in the examples that follow.

**Ways To Refine
Search Strategy**

The purpose of the discussions presented here is to illustrate ways to refine search strategies by adding screen numbers when any of the above-mentioned limitations of “structure-only” searches are encountered.

We have attempted to keep the examples straightforward by illustrating only one type of screen in each of the early examples. In later examples, the combined use of multiple screens is illustrated.

Experienced STN searchers may devise alternate approaches to meet the search requirements of some of the examples discussed here. These examples are not necessarily the best, or only, solutions to the stated search requirements. They do, however, allow us to illustrate the use of most of the 12 types of screens in one or more searches.

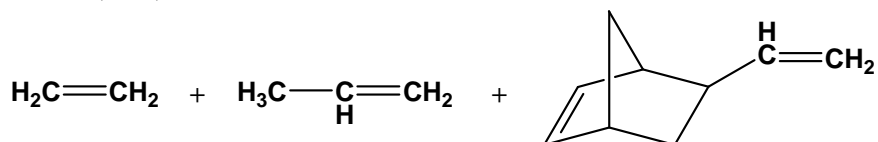
Adding Screens Search Examples

Example	Page
• SEARCHES ILLUSTRATING ADDITION OF SCREENS THAT ARE NOT GENERATED AUTOMATICALLY BY STN	
1. Copolymers of some defined and some undefined monomers	29
2. Homopolymers of a substructure	32
3. Isotopically labeled substances	34
4. Substances with restricted sites for substitution	35
• SEARCHES ILLUSTRATING ADDITION OF SCREENS THAT ARE CONDITIONALLY GENERATED BY STN	
5. Substances containing rings that may be either isolated or embedded	39
6. Substances containing isolated and embedded rings	43
7. Substances containing embedded ring systems	44
8. Small structures containing multiple system-defined variable atoms	46
• SEARCHES ILLUSTRATING ADDITION OF SCREENS THAT ARE USUALLY GENERATED AUTOMATICALLY BY STN	
9. Alkyl, alkenyl, and alkynyl esters	49
10. Substances with variable locations for attachment of desired atoms	51
• SEARCHES IN WHICH SOME OF THE REQUIREMENTS ARE MORE CONVENIENTLY SPECIFIED IN TERMS OF SCREENS THAN IN TERMS OF ONE OR MORE STRUCTURES	
11. Linear hydrocarbons containing single, double, or triple bonds	55
12. Locating substances corresponding to known analytical data	57
13. Coordination compounds with variable central metal atoms	60
14. Substances with variable bonding and restricted positions for substitution	61

Searches Illustrating Addition of Screens That Are Not Generated Automatically by STN

Example 1 Copolymers with Some Defined Monomers and Some Undefined Monomers

Find all copolymers containing ethylene, propylene, and 5-vinyl-2-norbornenes, i.e.,



with no restriction on the maximum number of components. The norbornene ring is not isolated.

Search Strategy Polymer searching, in particular, illustrates the power of combining structures with screen numbers to enhance search results in STN. Classes of substances, such as polymers, are described by Graph Modifier (GM) screens that are not generated automatically by STN.

GM: Graph Modifier Screens

The table below lists the GM screens that are most useful in polymer searching. A complete listing of GM screens may be found in the screen dictionary.

Polymer Screens	
Screen Number	Definition
2043	Polymer (general definition)
2067	Homopolymers and copolymers (A) _x , (A.B) _x , etc.
2068	SRUs; polymers defined as structural repeating units
2069	SRU with end groups X-(-Y-) _n -Z
2070	SRU without end groups (-Y-) _n

Multicomponent Substance Screens	
Screen Number	Definition
2127	two or more components
2077	three or more components
2078	four or more components

Search Strategy The search is conducted in the following steps:

Step	Action
1	<p>Build separate structures for each of the required monomers. Upload the three query structures in a single step.</p> <pre> Uploading c:\casnc\stn express\queries\vinylbornene.str L1 STRUCTURE UPLOADED Uploading c:\casnc\stn express\queries\ethylene.str L2 STRUCTURE UPLOADED Uploading c:\casnc\stn express\queries\propylene.str L3 STRUCTURE UPLOADED </pre>
2	<p>Create a screen set consisting of the screen number (2067) for homopolymer or copolymer and that for three or more components (2077). This screen restricts the search to polymeric substances defined in terms of the component monomers. [Screen 2043 (polymers - general category) could have been used, but 2067 is more precise.]</p> <pre> => SCREEN 2067 AND 2077 L4 SCREEN CREATED </pre>
3	<p>Run a substructure SAMPLE search (no fee) using the strategy combining the query L-numbers with AND logic.</p> <pre> => S L1 AND L2 AND L3 AND L4 </pre> <p>Examine the FULL FILE PROJECTIONS and the results. In this case, no answers were retrieved by the sample search, but the search will run within system limits.</p>
4	<p>Run the FULL search.</p> <pre> => S L1 AND L2 AND L3 and L4 FUL FULL SEARCH INITIATED 09:19:20 FULL SCREEN SEARCH COMPLETED - 22580 TO ITERATE 100.0% PROCESSED 22580 ITERATIONS 50 ANSWERS SEARCH TIME: 00.00.01 L6 62 SEA SSS FUL L1 AND L2 AND L3 AND L4 </pre>
5	<p>Display results using D SCAN (no fee) or other display formats. See the CAS REGISTRY Database Summary Sheet at www.cas.org.</p>

Search Results => D 50

L6 ANSWER 50 OF 50 REGISTRY COPYRIGHT 2008 ACS on STN
RN 27155-45-9 REGISTRY
ED Entered STN: 16 Nov 1984
CN Bicyclo[2.2.1]hept-2-ene, 5-ethenyl-, polymer with ethene
and 1-propene (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN 1-Propene, polymer with ethene and 5-ethenylbicyclo[2.2.1]hept-
2-ene (9CI)
CN 2-Norbornene, 5-vinyl-, polymer with ethylene and propene (8CI)
CN Ethene, polymer with 5-ethenylbicyclo[2.2.1]hept-2-ene and 1-
propene (9CI)
CN Ethylene, polymer with propene and 5-vinyl-2-norbornene (8CI)
CN Propene, polymer with ethylene and 5-vinyl-2-norbornene (8CI)

OTHER NAMES:

CN Ethylene-propene-vinylnorbornene copolymer
CN Ethylene-propylene-5-vinyl-2-norbornene copolymer
CN Ethylene-propylene-5-vinyl-2-norbornene polymer
CN Ethylene-propylene-5-vinylnorbornene copolymer
CN Ethylene-propylene-vinylnorbornene copolymer
MF (C9 H12 . C3 H6 . C2 H4)x

CI PMS

PCT Polyolefin, Polyvinyl

LC STN Files: CA, CAPLUS, CHEMLIST, IFICDB, IFIPAT,
IFIUDB, TOXCENTER, USPAT2, USPATFULL

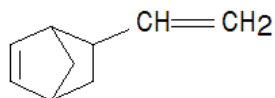
Other Sources: TSCA**

(**Enter CHEMLIST File for up-to-date regulatory information)

CM 1

CRN 3048-64-4

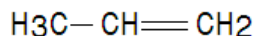
CMF C9 H12



CM 2

CRN 115-07-1

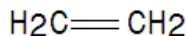
CMF C3 H6



CM 3

CRN 74-85-1

CMF C2 H4



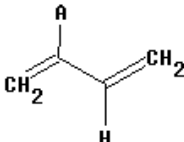
143 REFERENCES IN FILE CA (1907 TO DATE)
30 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
143 REFERENCES IN FILE CAPLUS (1907 TO DATE)

Note: Polymers are displayed as the structures of the component monomers. The CAS Registry Number and molecular formula of each component monomer are also given.

Example 2 Homopolymers of a Substructure

Find all homopolymers of 1,3-butadiene with any substitution at the 2-position except hydrogen.

Search Strategy The strategy for this search is as follows:

Step	Action
1	<p>Build the substructure of interest with the 2-position open for substitution.</p> 
2	<p>Create separate screen sets for homopolymer or copolymer (2067) and two or more components (2127).</p> <p>=> SCR 2067 L1 SCREEN CREATED</p> <p>=> SCR 2127 L2 SCREEN CREATED</p> <p>Note: Since the two screen numbers will be combined using NOT logic, they must be created in separate L-numbers.</p> <p style="text-align: right;"><i>(Continued on next page)</i></p>

3	<p>Upload the structure query (L3), and run the sample search with the screen sets.</p> <pre> Uploading C:\CASNC\STN Express\Queries\butadiene.str L3 STRUCTURE UPLOADED => S L3 AND L1 NOT L2 SAMPLE SEARCH INITIATED 10:51:13 SAMPLE SCREEN SEARCH COMPLETED - 100.0% PROCESSED 148 ITERATIONS SEARCH TIME: 00.00.01 FULL FILE PROJECTIONS: ONLINE **COMPLETE** BATCH **COMPLETE** PROJECTED ITERATIONS: 2231 TO 3689 PROJECTED ANSWERS: 22 TO 418 L4 5 SEA SSS SAM L3 AND L1 NOT L2 The FULL FILE PROJECTIONS indicate the full search will run to completion. Display one or more answers using the no-cost D SCAN feature.</pre>
4	<p>Run the search, and review results.</p> <pre> => S L4 FUL FULL SEARCH INITIATED 10:51:41 FULL SCREEN SEARCH COMPLETED - 3156 TO ITERATE 100.0% PROCESSED 3156 ITERATIONS 145 ANSWERS SEARCH TIME: 00.00.01 L5 83 SEA SSS FUL L3 AND L1 NOT L2</pre>

Note: No parentheses can be used to enclose a structure L-number and a screen L-number.

Additional Comments

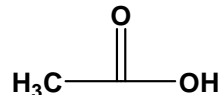
The same results could have been achieved using a structure with no substituent at the 2-position and an H-COUNT of exactly 0 at that position. The "A" substituent is preferred, since it is visible when viewing the query structure.

If the requirement had been for homopolymers of a specific monomer rather than of a substructure as in this example, the search could have been run as an EXACT search and screen 2127 (two or more components) would not have to be added manually, since STN automatically excludes substances whose record contains screen 2127 when an EXACT search is specified for a structure containing a single component.

Screen 2067 (homopolymer) should be retained in the search strategy even when an EXACT search is specified unless you also want to retrieve the monomer, any of its stereoisomers, or any of its isotopically labeled, ionic, or radical counterparts.

Example 3**Isotopically Labeled Compounds**

Find all isotopically labeled acetic acids, i.e.,



in which the isotopically labeled atoms could be H, C, and/or O.

Search Strategy

The strategy for this search is to build the structure for acetic acid (L1) and combine it with GM screen 2039 (requires the presence of 1 or more atoms with abnormal mass) (L2) using AND logic. The search can be run at lower cost as an EXACT search, since EXACT searches allow for isotopically labeled equivalents.

```
=> S L1 AND L2 EXACT FULL
FULL SEARCH INITIATED 12:59:59
FULL SCREEN SEARCH COMPLETED - 368 TO ITERATE
100.0% PROCESSED 368 ITERATIONS 70 ANSWERS
SEARCH TIME: 00.00.01
```

```
L3 72 SEA EXA FUL L1 AND L2
```

The answers include labeled acetate ions as well acetic acids.

**Additional
Comments**

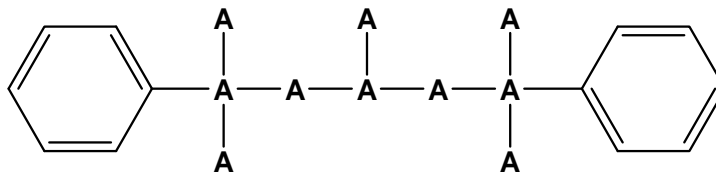
The labeled ionic species could have been excluded from the answer set by excluding GM screen 2040 (charge - fixed, tautomeric, or delocalized) and GM screen 2054 (radical ion) from the search strategy, viz.,

```
Structure AND (screen 2039 NOT (screen 2040 OR screen 2054))
```

Note that it is permissible to “nest” the screen numbers, when combining several screens with more than one type of Boolean operator. The best approach when excluding several screen numbers is to create a single screen set containing those screen numbers (use the OR operator) and then exclude that entire screen set (use the NOT operator).

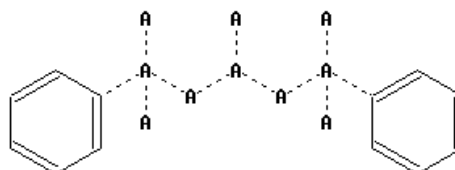
Example 4 Substances with Restricted Sites for Substitution

Find substances having the substructure



where A is any element except hydrogen, any chain bonds (shown above as single) are acceptable between A atoms, and substitution and/or ring fusion are allowed on the benzene rings. No further substitution on the A atoms is allowed.

Search Strategy The structure query (L1, shown below) includes Non-H attachments on all the "A" atoms to limit substitution and all the chain bonds are unspecified.



A SAMPLE search projected the search would complete in CAS REGISTRY, but the same is not true in REAXYSFILE (lower system limits). The addition of CS screens will make the search more efficient in CAS REGISTRY (no ITERATION INCOMPLETES) and allow it to run to completion in REAXYSFILE.

CS Screens

The CS screens represent the exact number of non-hydrogen atoms attached to central atoms in a sequence of 3-6 atoms. For example, screen 588 (see ==> following) represents a three-atom sequence in which the first atom is connected to exactly one (1) other non-hydrogen atom, the second atom is connected to exactly four (4) other non-hydrogen atoms, and the third is connected to exactly one (1) other non-hydrogen atom. No bond types or bond values are included in the fragment definition for screen 588.

However, bond types (* for ring, - for chain), but not bond values, may be specified (see screen 684).

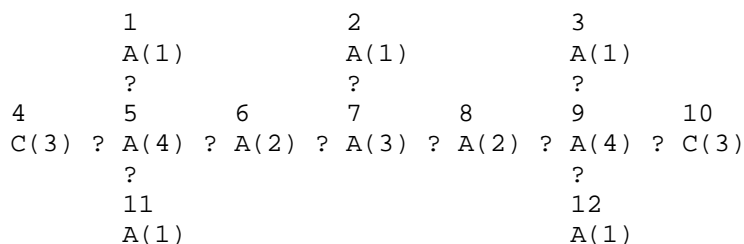
From the Connectivity Sequence Section of the Screen Dictionary

Screen Number	CS	Fragment Definition	Freq. %
563	CS	1 3 2 4	6.81
564	CS	1 3 2 1 1	4.90
565	CS	1 - 3 * 3 - 1	9.71
566	CS	1 - 3 - 3 - 1	9.12
		•	
		•	
		•	
587	CS	1 3 4	8.58
==>588	CS	1 4 1	27.09
589	CS	1 - 4 * 2	7.98
		•	
		•	
		•	
592	CS	2 3 2 2 3 4 1	7.50
647	CS	2 3 2 4	11.56
567	CS	2 - 3 * 3 - 1	19.66
		•	
		•	
		•	
683	CS	2 - 4 * 3	2.54
==>684	CS	2 - 4 - 3	6.98
516	CS	3 - 2 - 2 - 2 - 1	5.04
		•	
		•	
		•	
720	CS	4 2 3 2 3 4	# 3.12
721	CS	4 2 3 2 4	# 5.18
705	CS	4 2 3 3	9.47

CS Screens

CS screens can be very useful in describing substances where additional substitution is not allowed. In this example, a series of CS screens describing the various desired possibilities will be included in the search strategy.

The A atoms and the two ring carbon atoms have been numbered as shown for clarity in the following discussion. The connectivity value of each atom, i.e., the number of non-hydrogen atoms to which it is attached, is shown in parentheses. The “?” is used to represent the unspecified bonds.



Various connectivity sequences that should be considered in defining this structure are listed in the following table:

Atom Sequence	Connectivity Sequence	Screen
A2-A7-A8-A9-A3	1 3 2 4 1	564
A2-A7-A8-A9-A12	1 3 2 4 1	564
A2-A7-A6-A5-A1	1 3 2 4 1	564
A2-A7-A6-A5-A11	1 3 2 4 1	564
A1-A5-A11	1 4 1	588
A3-A9-A12	1 4 1	588
A6-A7-A8-A9	2 3 2 4	647
A8-A7-A6-A5	2 3 2 4	647
A6-A5-c4	2 - 4 - 3	684
A8-A9-c10	2 - 4 - 3	684
A5-A6-A7-A8-A9	4 2 3 2 4	721

Note

Screen 721 is a shared screen (note the # in the screen dictionary record). To determine what connectivity sequences are shared in the definition of screen 721, consult the Screen Number Order section of the screen dictionary.

From the Screen Number Order Section of the Screen Dictionary

Screen Number		Fragment Definition					Freq. %
721	CS	4	2	2	2	4	5.18
		4	2	2	3	4	
		4	2	3	2	4	
		4	2	3	3	4	
		4	3	2	2	4	
		4	3	3	2	4	
		4	3	3	3	4	

It is unlikely that any other connectivity sequence shared in the definition of screen 721 will adversely affect the use of this screen in this example.

CS Screens

All of the CS screens listed in the above table will be added to the search strategy using AND logic since all of these connectivity sequences must be present in valid answers.

Search Results

```
L1      STRUCTURE UPLOADED

=> SCR 564 AND 588 AND 647 AND 684 AND 721
L2      SCREEN CREATED

=> S L1 AND L2 FUL
FULL SEARCH INITIATED 08:03:29
FULL SCREEN SEARCH COMPLETED - 6837 TO ITERATE

100.0% PROCESSED 6837 ITERATIONS 222 ANSWERS
SEARCH TIME: 00.00.01


L3      252 SEA SSS FUL L1 AND L2
```

Additional Comments

This example illustrates the use of CS screens to *block* positions to prevent further substitution. CS screens may also be used to *require* substitution at certain locations. In both instances, the function of the added CS screens is the same, i.e., determining the pattern of substitution.

Searches Illustrating Addition of Screens That Are Conditionally Generated by STN

Example 5 Substances Containing Rings That May Be Isolated or Embedded

Find substances containing  where the ring may be either isolated as shown or embedded within a larger ring system and any substitution is allowed.

Search Strategy In this search, the ring system cannot be specified as “isolated” in the structure query, since we want both isolated and embedded ring systems. This means that no TR screens will be generated automatically by STN.

The screens that are automatically generated in this example describe commonly occurring structural features, e.g., substances containing three or more carbon atoms (>90% of the file), one or more nitrogen atoms (>62% of the file), one or more rings (>81% of the file), and an atom sequence of C C C N (>56% of the file). As a result, the query likely will not run to completion due to too many candidate answers.

SAMPLE Search The query is uploaded to STN, a SAMPLE search is run in CAS REGISTRY, and the FULL FILE PROJECTIONS examined.

```
L4          STRUCTURE UPLOADED

=> S L4
SAMPLE SEARCH INITIATED 08:26:00
SAMPLE SCREEN SEARCH COMPLETED - 167003 TO ITERATE
  1.2% PROCESSED      2000 ITERATIONS      0 ANSWERS
INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)
SEARCH TIME: 00.00.01
FULL FILE PROJECTIONS:  ONLINE  **INCOMPLETE**
                        BATCH  **INCOMPLETE**
PROJECTED ITERATIONS:   3316034 TO 3364086
PROJECTED ANSWERS:     0 TO      0
L5          0 SEA SSS SAM L4
```

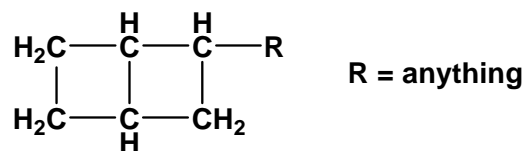
The SAMPLE search confirms that the query will not run to completion. TR screens will need to be added. By manually adding TR screens for all possible four-membered rings (isolated and embedded), the search requirements can be specified in a search that will not exceed the system thresholds.

**How To Find
TR Screens**

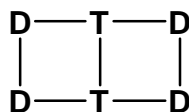
Since the ring in Example 5 is defined as a single ring, another ring system will be used in the following discussion. TR screens describe the nature and size of the ring in terms of whether the ring atoms are fused or not fused. Note that TR screens are the only screens that specify the size of a ring. All non-hydrogen atoms are described by the notation “D” or “T”:

- “D” ring atom is one that is not fused, i.e., it is connected to only two other ring atoms (it is part of only one ring).
- “T” ring atom is one that is fused, i.e., it is connected to three or more other ring atoms (it is part of two or more rings).

For example, the ring system



would be coded (using the screen dictionary notation) as



(Note that other attached atoms that are not part of the ring system being described, such as the R group, are ignored in determining whether a ring atom is a “D” or “T” atom.)

TR Screens

The fragment definitions for TR screens are ordered by first citing the maximum number of adjacent D's followed by the remaining T's and D's in the order in which they are encountered around the ring system.

The rings represented by TR screens are also governed by a rule that considers only the "smallest number of smallest rings". This means that the above ring system is represented by two occurrences of the DDTT ring, but the larger, or enveloping, DDTDDT ring is not represented in the STN record for this substance even though the DDTDDT screen *is valid* for other substances.

From the Type of Ring Section of the Screen Dictionary

Screen Number		Fragment Definition	Freq. %
1848	TR	DDD	2.06
1849	TR	2 DDD	0.21
1850	TR	DDT	0.26
1851	TR	DTT	# 2.46
1851	TR	TTT	# 2.46
==>1852	TR	DDDD	0.97
**1853	TR	DDDT	# 2.17
**1853	TR	DDTT	# 2.17
1854	TR	2 DDTT	0.04
**1853	TR	DTDT	# 2.06
**1853	TR	DTTT	# 2.06
**1853	TR	TTTT	# 2.06
1855	TR	DDDDD	21.27

Addition of screens 1852 (==>) or 1853 (**) to the search strategy would increase the efficiency of the search. These two relatively low frequency screens would allow for one or more occurrences of DDDD, DDDT, DTDT, DTTT, and/or TTTT rings because screen 1853 is a shared screen. These two screens not only cover all of the possible types of four-atom rings, but also restrict the search to substances that contain one or more four-atom rings.

Search Results

A new SAMPLE search shows the search will run to completion. The full file search results are shown.

```
=> SCR 1852 OR 1853
L6      SCREEN CREATED
```

```
=> S L4 AND L6 FUL
FULL SEARCH INITIATED 08:27:38
FULL SCREEN SEARCH COMPLETED - 83842 TO ITERATE
100.0% PROCESSED      83842 ITERATIONS 182 ANSWERS
SEARCH TIME: 00.00.01
L8              198 SEA SSS FUL L4 AND L6
```

**Additional
Comments**

In this example, it was not necessary to consider screens for rings smaller than four members since it was chemically unlikely that rings such as



would occur. However, if the initial requirement had been for a search for all substances containing the



ring regardless of where it occurred, then it would be necessary to allow for three-membered rings (screen 1851) in order to retrieve substances such as (CAS Registry Numbers[®] given)



19540-05-7



34743-96-9



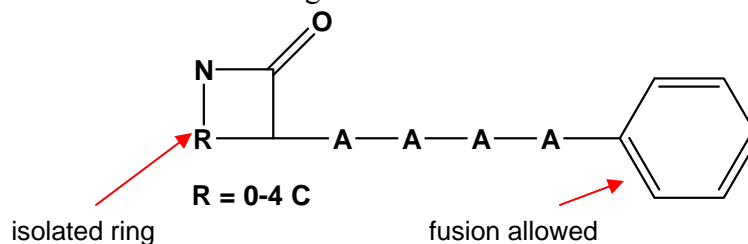
76152-46-0

or any of their derivatives. These substances would not have been retrieved by the original strategy (even if the bonds in the structure had been changed to unspecified) since STN, as mentioned above, does not consider the larger enveloping rings when assigning TR screens to substance records.

Note: The general rule for adding TR screens for a ring of size “n” is to consider *all* possibilities for rings equal to and smaller than “n”.

Example 6 Substances Containing One Isolated Ring and One Embedded Ring

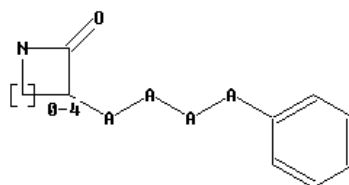
Find substances containing the substructure



in which the N-containing ring varies in size from 3-7 atoms and is isolated but the phenyl ring may be isolated (as shown) or embedded in a larger ring system, and any atom except hydrogen (A) is allowed in the four-atom chain which joins the two ring systems.

Search Strategy In this example, the SAMPLE search shows the search will not run to completion. The addition of TR screens may bring the search within system limits. The repeating group in the query allows an isolated N-containing ring from 3-7 atoms. This will require the addition of five TR screens for the five isolated rings: DDD, DDDD, DDDDD, DDDDDD, DDDDDDD. The screen numbers are 1848, 1852, 1855, 1868, and 1893.

Search Results The search query (L1)



with the screens resulted in a complete search.

=> SCR 1848 OR 1852 OR 1855 OR 1868 OR 1893

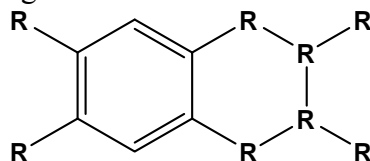
L3 SCREEN CREATED

=> S L1 AND L3 FUL

L5 11809 SEA SSS FUL L1 AND L3

Example 7 Substances Containing Embedded Rings

Find substances containing the substructure

**R = Any ring node**

where the R nodes are any ring atoms and the C-R bonds are ring bonds.

Search Strategy The SAMPLE search projects far more candidate answers than STN can handle.

```
=> S L1
SAMPLE SEARCH INITIATED 11:03:27
SAMPLE SCREEN SEARCH COMPLETED - 779853 TO ITERATE
  0.3% PROCESSED      2000 ITERATIONS      6 ANSWERS
INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)
SEARCH TIME: 00.00.01
FULL FILE PROJECTIONS:  ONLINE  **INCOMPLETE**
                        BATCH   **INCOMPLETE**
PROJECTED ITERATIONS:   15549256 TO 15644864
PROJECTED ANSWERS:      43890 TO 49692
L4                      6 SEA SSS SAM L3
```

In this search requiring the substructure to be embedded in a larger ring system, all of the requirements cannot be specified in the structure since STN only allows ring specifications of “isolated” or “isolated or embedded”, but does not allow a ring specification of “embedded”. Manual addition of a TR screen with DTTDTT specification should help. The addition of a ring count screen requiring four or more rings may also be useful.

From the Type of Ring Section of the Screen Dictionary

Screen Number	Fragment Definition	Freq. %
1889 TR	DTTDTT	3.09
1890 TR 2	DTTDTT	0.66

Note that screen 1890 requires two or more occurrences of DTTDTT rings.

**Ring Count
Screens**

Addition of the RC screen, requiring four (or more) rings in the structure (screen 1841), would also be beneficial.

From the Ring Count Section of the Screen Dictionary

Screen Number		Fragment Definition	Freq. %
1840	RC	3	53.05
1841	RC	4	33.05
1842	RC	5	17.59

Search Results

Addition of the TR screen 1890 and RC screen 1841 permit this search to run to completion with good results.

=> **SCR 1890 AND 1841**

L3 SCREEN CREATED

=> **S L1 AND L3**

FULL FILE PROJECTIONS: ONLINE **COMPLETE**
BATCH **COMPLETE**

PROJECTED ITERATIONS: 164750 TO 175810

PROJECTED ANSWERS: 38732 TO 44194

L4 50 SEA SSS SAM L1 AND L3

=> **S L4 FUL**

FULL SEARCH INITIATED 12:34:15

FULL SCREEN SEARCH COMPLETED - 169333 TO ITERATE

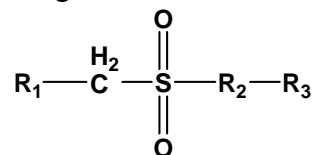
100.0% PROCESSED 169333 ITERATIONS 41317 ANSWERS

SEARCH TIME: 00.00.02

L5 41317 SEA SSS FUL L1 AND L3

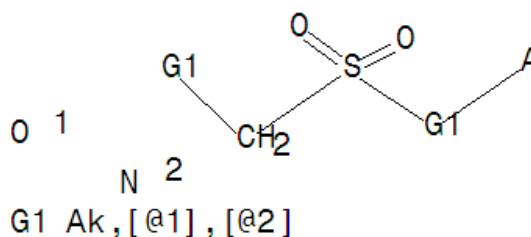
Example 8 Small Structures Containing Several Variable Atoms

Find substances containing the substructure



R1 = R2 = N, O, any carbon chain
R3 = any element except H

Search Strategy The SAMPLE search resulted in a full file projection of incomplete. The query



The node specification for the O, N, A is ring/chain.

AA Screens AA screens describe an atom and the non-hydrogen atoms to which that atom is attached and, optionally, the bond types and bond values. AA screens are automatically generated for structures in which all nodes are defined as specific atoms.

Remember that user-defined variable atom symbols are ignored during automatic screen generation. When generating AA screens, STN does not automatically “expand” user-defined variable atom symbols to include all possibilities. (A limited number of AA screens, including the system-defined variable atom symbols A, M, and X, are found in the screen dictionary and may be automatically generated.)

**Example 8
Requirements**

This means that, for the structure in this example, AA screens indicating the attachment of either the N or O or C to the C or to the S atom are not automatically generated. Manual addition of these AA screens will greatly improve the efficiency of the initial screen search. For example:

- If G1 = N, then the substructure is N-C-S AA screen 1492
- If G1 = O, then the substructure is O-C-S AA screen 1532
- If G1 = C, then the substructure is C-C-S AA screen 1163

At least one of these AA screens must be present in the STN record for the desired substances.

AA Screens

The AA screens for the sulfur (1788 S-C-N-O-O or 1796 S C O O O or 1782 S-C-C-O-O) describe the various possibilities of environment.

AS Screens

No AS screens would be automatically generated for this structure. Recall that variable atoms symbols, either user-defined or system-defined, are ignored during automatic screen generation. This results in atom sequences of three (or fewer) defined atoms in the structure, and a minimum of four is necessary by definition for AS screens.

The frequency of occurrence of the atom sequences N C S O (G1 at the left node is nitrogen) and O C S O (G1 at the left node is oxygen) is very low – see screens 446 and 511, respectively. While C-C-S-O (G1 at the left node is C) has a much higher frequency, it is still useful. Manual addition of these AS screens will greatly improve the efficiency of the initial screen search and allow the search to run to completion.

From the Atom Sequence Section of the Screen Dictionary

Screen Number		Fragment Definition		Freq. %
445	AS	N - C - S - N	#	0.06
446	AS	N C S O	#	0.46
446	AS	N C S S	#	0.46
	•			
	•			
	•			
468	AS	O C S N	#	0.59
511	AS	O C S O	#	0.48
511	AS	O C S S	#	0.48
	•			
	•			
	•			
200	AS	C * C - S - N - C	#	2.34
2085	AS	O C S O	#	26.07
201	AS	C * C - S - O	#	6.11

Search Results The full search with screens added gave good results. Note that each group of screens is a separate L-number.

```
L3          SCR 1492 OR 1532 OR 1163
L4          SCR 1788 OR 1796 OR 1782
L5          SCR 446 OR 511 OR 2085
L6          50 S L1 AND L3 AND L4 AND L5
```

=> **S L6 FULL**

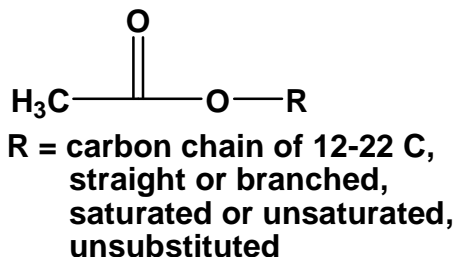
```
L7          4563 SEA SSS FUL L1 AND L3 AND L4 AND L5
```

EC Screens EC screens are also not generated automatically by STN for user-defined or system-defined variables in the structure. It was not necessary to add EC screens to the final strategy since the presence of the N or O is implicit in the definition of the low-frequency AA and AS screens that were included in the strategy.

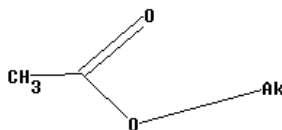
Search Illustrating Addition of Screens That Are Usually Generated Automatically by STN

Example 9 Alkyl, Alkenyl, and Alkynyl Esters

Find all alkyl, alkenyl, and alkynyl acetates where the hydrocarbon chain contains 14-22 carbon atoms, is straight or branched, may be saturated or unsaturated, and is unsubstituted.



Search Strategy A SAMPLE search of the query



with element count of 14-24 C and non-H attachment exactly one on the Ak, projected more than 7 million substances would pass the screen search.

The strategy is modified by use of NOT logic:

- Exclude substances with 25 or more carbon atoms (EC screen 1951), since the search requirements allow for a maximum of 24 carbon atoms.
- Exclude “28 or more total non-hydrogen atoms” (AC screen 1911), since a maximum of 26 non-hydrogen atoms is desired.

Note: This strategy would eliminate copolymers, mixtures, and other “dot-disconnected” multicomponent substances containing >25 carbon atoms and >28 total non-hydrogen atoms. This was considered acceptable for the purposes of this search. However, if you are not willing to exclude such substances, you should not use this technique to limit the range of carbon atoms in a hydrocarbon chain.

A second SAMPLE search was run that projected over 3 million substances will pass the screen search. Still more screens need to be added.

**Further
Modification**

Use of the no-fee D SCAN feature detected answers that were not desired.
The following screens will be added to eliminate these undesired substances:

Undesired Feature	Screen Type	Screen Number
three or more oxygen atoms	EC	2006
one or more halogen atoms	GM	1929
one or more sulfur atoms	EC	2021
one or more nitrogen atoms	EC	1992
one or more rings	RC	1838

Search Results

This time the SAMPLE search projected a full search to completion.

```
=> SCR 1951 OR 1911 OR 2006 OR 1929 OR 2021 OR 1992 OR 1838  
L3      SCREEN CREATED
```

```
=> S L1 NOT L3
```

```
FULL FILE PROJECTIONS:  ONLINE  **COMPLETE**  
                        BATCH   **COMPLETE**  
PROJECTED ITERATIONS:   61679 TO   68521  
PROJECTED ANSWERS:     1059 TO   2129
```

```
L4      50 SEA SSS SAM L1 NOT L3
```

```
=> S L4 FUL
```

```
L5      7516 SEA SSS FUL L1 NOT L3
```

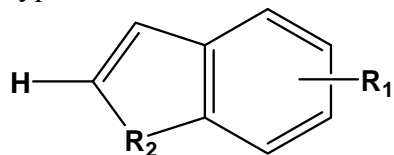
Note

When several screen numbers are to be excluded from the search profile, these screen numbers are first combined with OR logic in one screen set L-number and then the entire screen set is excluded using NOT logic.

Example 10

Substances with Variable Locations for Attachment of Desired Atoms

Find substances of the type

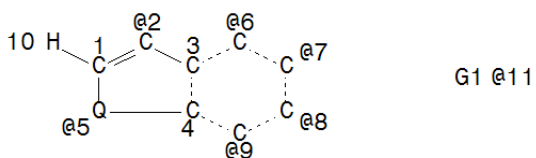


$R_1 = \text{NH}_2, \text{OH}, \text{SH}, \text{Br}, \text{Cl}, \text{F}, \text{or I}$
 $R_2 = \text{any element except C,H}$

The R substituent may be attached to any open ring position. The ring is isolated, but any other substitution is allowed.

Search Strategy

The SAMPLE search of the query projects too many candidate answers for iteration with the following query (STRucture command version shown).



```
VAR G1=NH2/OH/SH/X
VPA 11-2/5/6/7/8/9 U
NODE ATTRIBUTES:
DEFAULT ECLEVEL IS LIMITED
GRAPH ATTRIBUTES:
RSPEC I
NUMBER OF NODES IS 11
```

It will be necessary to add some screens for the search to run to completion.

Addition of AA Screens

AA screens can be added to the search strategy to require that the G1 atoms be attached to a ring carbon atom (but not necessarily one in the 6-5 aromatic ring system):

- When G1 is nitrogen, screen 1097 applies.
- When G1 is oxygen, screen 1135 applies.
- When G1 is sulfur, screen 1158 applies.
- When G1 is halogen, screen 1077 applies.

A screen set containing these alternative screens (OR logic) will be added to the search strategy using AND logic.

From the Augmented Atom Section of the Screen Dictionary

Screen Number			Fragment Definition	Freq. %
1096	AA	5	C * C * C * N	0.82
1097	AA		C * C * C - N	28.97
1098	AA	2	C * C * C - N	10.75
		•		
		•		
		•		
1134	AA	2	C * C * C * O	4.34
1135	AA		C * C * C - O	35.34
1136	AA	2	C * C * C - O	17.07
		•		
		•		
		•		
1157	AA	2	C * C * C * S	0.72
1158	AA		C * C * C - S	7.91
1159	AA	2	C * C * C - S	1.93
		•		
		•		
		•		
1087	AA		C -1C -1C -1S -1X	# 7.76
1077	AA		C * C * C - X	18.59
1079	AA		C - C - C - X	1.51

SAMPLE Search Results

The addition of the AA screens (1097 or 1135 or 1158 or 1077) allows the search to run to completion but only as a BATCH (overnight) search.

FULL FILE PROJECTIONS: ONLINE **INCOMPLETE**
BATCH **COMPLETE**

Addition of TW Screens

An alternative approach would be to add TW screens to the search profile in order to specify the presence of the desired substituents, i.e., the N, O, or S atoms.

TW screens describe the exact number of hydrogens attached to two adjacent atoms and the bond types and bond values involved. The following TW screens can be used for this example:

Group	Screen Number
NH ₂ group on a carbon in the five-membered ring	1827
NH ₂ group on a carbon in the six-membered ring	1828
OH group on a carbon in the five-membered ring	1831
OH group on a carbon in the six-membered ring	1832
SH group on a carbon in the five-membered ring	1831
SH group on a carbon in the six-membered ring	1832

Note

Two screens must be considered for each of the desired substituents, since the bonds in the five-membered ring (one double and the others single) are different from those in the six-membered ring (aromatic, i.e. normalized).

From the Twin Augmented Atom Section of the Screen Dictionary

Screen Number			Fragment Definition					Freq. %
1827	TW		C	*1C	*2C	-1N H2	#	0.63
1828	TW		C	*4C	*4C	-1N H2		1.65
		•						
		•						
		•						
1831	TW		C	*1C	*2C	-1O H	#	5.45
1832	TW		C	*4C	*4C	-1O H	#	6.53
		•						
		•						
		•						
1831	TW		C	*1C	*2C	-1S H	#	5.45
1832	TW		C	*4C	*4C	-1S H	#	6.53

Since any halogen is also an acceptable substituent, AA screen number 1077 (describing the attachment of a halogen to a ring carbon) is also included in the screen set using OR logic for the acceptable alternative substituents. A screen set containing these alternative screens would be added to the search strategy using AND logic.

**Final Search
Results**

The addition of the second set of TW screens allows the search to complete.
(L1 is the structure query, SAMPLE searches not shown.)

=> **SCR 1097 OR 1135 OR 1158 OR 1077**

L3 SCREEN CREATED

=> **SCR 1827 OR 1828 OR 1831 OR 1832**

L5 SCREEN CREATED

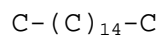
=> **S L1 AND L3 AND L5 FUL**

L7 11868 SEA SSS FUL L1 AND L3 AND L5

Searches in Which Some of the Requirements Are More Conveniently Specified in Terms of Screens Than in Terms of One or More Structures

Example 11 Linear Hydrocarbons Containing Single, Double, or Triple Bonds

Find all non-ring C₁₈₋₂₃ hydrocarbons containing the linear structure



In which single, double, and/or triple bonds are allowed.

Search Strategy In this example, it is impractical to create multiple structures (scores of structures would be required). It is possible to specify the basic linear C₁₆ chain in a single structure (with all bonds left unspecified to allow for single, double, and triple bonds and with non-H attachments of exactly 2 for all chain members), and the remainder of the search requirements with a combination of screens.

- Screen used with AND logic
 - EC screen 1949 C18 requires 18 or more C atoms
- Screen set used with NOT logic
 - AC screen 1909 24 requires 24 or more atoms
 - RC screen 1838 1 requires one or more rings
 - GM screen 1918 M requires one or more metals (all elements except Ar, As, At, B, Br, C, Cl, F, H, He, I, Kr, N, Ne, O, P, Rn, S, Se, Si, Te, and Xe)
 - GM screen 1924 0 requires one or more group 0 element
 - GM screen 1929 VIIb requires one or more halogen
 - GM screen 2003 O,S requires one or more O or S
 - GM screen 2019 Po, Se, Te requires one or more Po, Se, or Te
 - GM screen 1925 Vb requires one or more group Vb element
 - EC screen 1992 N requires one or more N atom
 - EC screen 2026 Si requires one or more Si atom
 - EC screen 1932 B requires one or more B atom

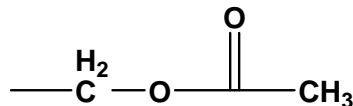
GM and EC Screens

The screen set of the GM and EC screens above is useful when you want to restrict your answers to hydrocarbons. You may wish to build this set and store it with your STN login ID for later use.

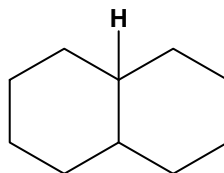
Example 12**Locating Substances Corresponding to Known Analytical Data**

Find substances in CAS REGISTRY that have the following characteristics:

- Combustion analysis and mass spectrometry indicate that the molecular formula is $C_{11}H_{12}O_5$.
- NMR (proton) indicates the presence of six acetate hydrogen atoms and four acetoxymethyl hydrogen atoms; that is, two occurrences of the structure



- NMR also shows the presence of bridgehead methine hydrogen atoms (a hydrogen occurring at the point of fusion of two ring systems), i.e.,



- IR shows the presence of another type of carbonyl group in addition to the acetate carbonyls.
-

Search Strategy

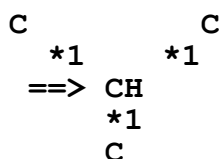
In this search, a multifragment structure requiring the presence of a carbonyl group (node specification on the carbon must be “ring or chain” and the bond should be “double” rather than “double exact”) and two occurrences of the acetoxymethyl group anywhere within the substance is combined with screen numbers to describe the bridgehead methine hydrogen (HA screen), the number of carbon and oxygen atoms required by the molecular formula (EC screens), and the maximum number of non-hydrogen atoms in the structure (AC screen).

Hydrogen Augmented Atom (HA) screens describe the exact number of hydrogen atoms attached to a central atom as well as other non-hydrogen atoms to which that central atom may be attached. Bond types and bond values must also be specified.

From the Hydrogen Augmented Atoms Section of the Screen Dictionary

Screen Number	HA		Fragment Definition	Freq. %
1016	HA	2	C H -1c -2C	5.64
1017	HA	3	C H -1C -2C	1.54
1018	HA	4	C H -1C -2C	0.99
1028	HA		C H *1C *1C *1C	6.09
1034	HA		C H *1C *1C -1C	10.75
1051	HA		C H -1C -1C -1C	10.06
1052	HA	3	C H -1C -1C -1C	3.43

For example, screen 1028 describes a carbon atom (==> in the diagram below) attached to exactly one hydrogen atom and to three other carbon atoms by single ring (*1) bonds. This corresponds to the bridgehead methine hydrogen required by the NMR spectrum of the substance being analyzed.



Some searchers may choose to build another fragment in the structure to require the presence of the bridgehead methine hydrogen. If you do this, remember to indicate that the node specification of the carbon atoms is “ring”, and specify the hydrogen using an H-count of exactly 1. This approach was not used in this example so that we could illustrate the use of HA screens.

Molecular Formula Screens

The formula requires the presence of 11 carbon atoms, five oxygen atoms, and a total of 16 non-hydrogen atoms. EC screen 1945 requires the presence of 11 or more carbon atoms and EC screen 2008 requires the presence of five or more oxygen atoms. AC screen 1906 requires the presence of 18 or more non-hydrogen atoms. These screens can be combined in the logic expression

Screen 1945 AND screen 2008 NOT screen 1906

to approximate the required formula. Note that this logic expression will allow for one additional non-hydrogen atom beyond that required by the molecular formula. This results since the more precise (for this example) AC screen for 17 or more non-hydrogen atoms does not exist in the screen dictionary.

Note: There are no EC screens for hydrogen atoms.

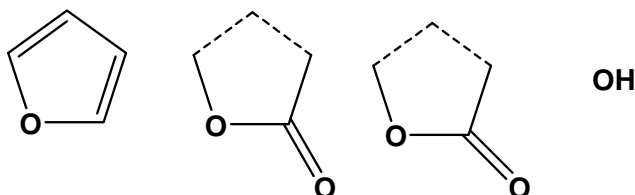
Search Results

L1 STRUCTURE UPLOADED
=> **SCR 1028**
L2 SCREEN CREATED
=> **SCR 1945**
L3 SCREEN CREATED
=> **SCR 2008**
L4 SCREEN CREATED
=> **SCR 1906**
L5 SCREEN CREATED
=> **S L1 AND L2 AND L3 AND L4 NOT L5 FUL**
L6 1 SEA SSS FUL L1 AND L2 AND L3 AND L4 NOT L5

**Additional
Comments**

In this example, only one candidate substance was retrieved. In other searches of this type more than one candidate substance may be retrieved.

For example, mass spectrometry and combustion analysis of another substance indicated a molecular formula of $C_{20}H_{22}O_7$. The infrared spectrum of this substance was consistent with the occurrence of a furan ring, two γ -lactones, and one hydroxyl group. A full-file search for this substance based on a four fragment strategy with screens:



The dotted bonds are unspecified. In addition the query includes:

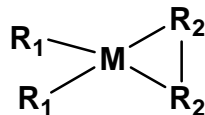
- TR screens 1855-1866 (all isolated or embedded 5-membered rings)
- EC screen 1950 for 20 or more C atoms
- EC screen 2010 for seven or more O atoms
- AC screen 1911 for 28 or more non-hydrogen atoms

and resulted in 24 answers having the correct number of carbon and oxygen atoms. Of these, 12 had the required number of hydrogens.

Additional information would be needed to determine which, if any, of the remaining 12 substances corresponds to the substance being analyzed, a reasonable number of possibilities has been identified.

Example 13 Coordination Compounds with Variable Central Metal Atoms

Find coordination compounds containing the substructure



R1 = any atom except C, H

R2 = N, O, S

where M = Fe, Co, Ni, Ru, Rh, Pd, Os, Ir, and Pt (the Group VIII metals) and any other ligands may be coordinated with the central metal atom.

Search Strategy In this example, many separate structures could be created and combined with OR logic.

An alternative approach would be to build one structure using the variable atom symbol “M” to indicate any metal, and then specify the desired metals by use of one or more screen numbers. In addition to EC screens for the individual elements, generic EC and GM screens are available for various groups and series of elements. In this example, EC/GM screen 1965 can be used to describe all of the Group VIII metals.

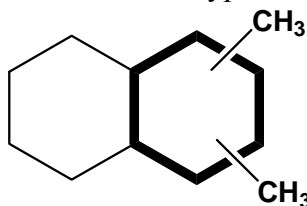
Since the R₁ atoms may occur in bidentate ligands (thus creating additional cyclic structures containing the M atom), no ring specification can be made. However, TR screens 1848 (DDD ring) and 1850 (DDT ring) can be added to increase the efficiency of the search. The node specification for the two R₁ atoms must be “ring or chain” and the bonds between the M and R₁ atoms may be designated as “ring or chain unspecified”, unless you decide that your requirements allow you to be more specific, e.g., “ring or chain single exact”.

Search Results

```
=> D HIS
L1          STRUCTURE UPLOADED
L2          SCR 1965
L3          SCR 1848 OR 1850
L4          971 S L1 AND L2 AND L3 FUL
```

Example 14 Substances with Variable Bonding and Restricted Positions for Substitution

Find all of the dimethyl derivatives of the type



where any degree of saturation or unsaturation of the rings is allowed, both methyl groups must be on the same ring (dark bonds) as illustrated, and no other substitution is desired.

Search Strategy Our structure query will consist of a 10-membered naphthalene type ring with two methyl groups both bonded (VPA) at various positions of the same ring. All ring bonds are unspecified and since no further substitution is allowed, we will run a closed substructure search (CSS).

Since no further substitution is allowed, the desired substances all contain 12 non-hydrogen atoms (all carbon atoms). All other substances could be eliminated by excluding those substances having an atom count of 13 or more non-hydrogen atoms. Unfortunately, the AC screen for 13 or more non-hydrogen atoms does not exist in the screen dictionary. However, the screen for 14 or more atoms does exist.

Adding AC Screens

Excluding all substances whose CAS REGISTRY records are associated with the AC 14 screen (screen number 1904) would still allow undesired substances to be retrieved, e.g., those substituted by a single halogen, oxygen, nitrogen, or sulfur atom, or an additional carbon atom, e.g., the trimethyl or isopropyl derivatives. However, using the CSS search prevents any extra answers.

Search Results

The CSS SAMPLE search shows that the screen is needed to allow a complete full-file search.

=> **S L1 CSS**

FULL FILE PROJECTIONS: ONLINE **INCOMPLETE**
 BATCH **INCOMPLETE**
PROJECTED ITERATIONS: 2562071 TO 2604529
PROJECTED ANSWERS: 0 TO 0

L2 0 SEA CSS SAM L1

=> **SCR 1904**

L3 SCREEN CREATED

=> **S L1 NOT L3 CSS**

FULL FILE PROJECTIONS: ONLINE **COMPLETE**
 BATCH **COMPLETE**
PROJECTED ITERATIONS: 4215 TO 6145
PROJECTED ANSWERS: 33 TO 447

L4 12 SEA CSS SAM L1 NOT L3

=> **S L1 NOT L3 CSS FUL**

FULL SEARCH INITIATED 12:44:44

FULL SCREEN SEARCH COMPLETED - 5087 TO ITERATE

100.0% PROCESSED 5087 ITERATIONS 215 ANSWERS
SEARCH TIME: 00.00.01

L5 215 SEA CSS FUL L1 NOT L3

Summary

Screen Numbers

Screen numbers are intrinsically involved in STN structure searching. They are automatically generated by STN from the structure(s) included in your search strategy. Usually these automatically generated screens are all that is necessary for your search to run efficiently to completion and give you the desired results.

As we have seen in this document, there are situations in which it becomes necessary to manually add screen numbers to the search strategy to define an efficient search. Situations in which you may want to consider adding screens include, but are not necessarily limited to, searches for:

- Structures for which the SAMPLE search gives a FULL FILE PROJECTION of INCOMPLETE
 - Classes of substances, such as polymers, for which the addition of screens is the most efficient way to restrict the search
 - Very general structures, i.e., those containing multiple variable atoms and/or numerous undefined bonds and/or several nodes defined as either “ring” or “chain”
 - “Small” structures, especially those with variable atoms, fewer than four fully defined atoms in a row, or only commonly occurring elements such as carbon, oxygen, and nitrogen
 - Substances for which all rings cannot be specified as “isolated”
 - Cyclic or linear hydrocarbons
 - Structures in which you desire to specify acceptable points of attachment
 - Combinations of various factors, none of which by themselves would result in an incomplete search
-

Boolean Logic

- Addition (AND logic) of a few low-frequency screens representing **required** structural features is usually very effective in increasing the efficiency of the search.
- Addition of screen sets containing alternative screens (OR logic) is also useful.

You may also exclude undesired features using screens and NOT logic. ***Always use the NOT operator with care and discretion since NOT logic is absolute and “unforgiving.”*** There is a possibility that some relevant answers may be inadvertently lost because the structural feature represented by the screen excluded with NOT logic occurred in an unexpected location in an otherwise desirable substance.

You should also recall that screen numbers may be shared, and the use of screens with NOT logic may eliminate a shared structural feature that you did not consider. It is always good practice to examine the screen number(s) you intend to use in the number ordered section of the screen dictionary to see what fragments are represented by the screen.

Structures and Screens

When your search is run, STN processes the structure(s) in your search profile separately from any screens that you include. This means that, if you add a single screen number, screen number must apply to all structures in your search requirements. If you add multiple screens in one or more screen sets, the entire screen expression must apply to all structures in the profile.

Proficiency in the use of screens to increase the efficiency, precision, or effectiveness of your search comes with practice and experience, that is, it is no different – nor necessarily harder – than developing any other skill. If you encounter difficulties with screen searching that this booklet does not resolved for you, please contact your STN Help Desk.
