Target Organic Contaminant Library Development in Support of Sample Analysis at Mars (SAM)

Raul Garcia-Sanchez
Research Investigator: Dr. Paul R. Mahaffy
Code 699, NASA Goddard Space Flight Center
Research Mentor: Dr. Prabhakar Misra
Department of Physics & Astronomy, Howard University

Outline

- Summary
- Research Purpose
- Research Questions
- Challenges
- Tools
- Methodology
- File Database
- Contaminants Database
- Project Timeline
- Issues

Summary

- The primary purpose of the research is to develop an organic contaminants database for the SAM project that supports the Mars Science Laboratory (MSL).

- Our work involves the development of an organic contaminants database that will allow us to determine what compounds are found here on Earth and would be inadvertently detected in the Mars soil and gaseous samples as impurities.

- In order to develop a comprehensive target database, we utilize the National Institute of Standards & Technology (NIST) Automated Mass Spectral Deconvolution and Identification System (AMDIS) and Ion Fingerprint Deconvolution (IFD) software to analyze the GC/MS data.

Research Purpose

- Analyze data from the rover environment using Gas Chromatography Mass Spectroscopy (GC/MS).
- Develop an organic contaminants database for the Sample Analysis at Mars utilizing GC/MS analysis tools.
- Utilize this database to identify compounds from actual mission samples.

Research Questions

- What organic compounds can be found in the rover structure? Where in the rover can they be found?
- How accurately can we determine the compounds from tests to those on the actual rover?
- How will identifying these compounds help when the time comes for the actual mission analysis?

Challenges

- Identifying the correct contaminant from very similar family of compounds.
  - i.e., Hexamethyl Cyclotrisiloxane vs
  - Octamethyl Cyclotetrasiloxane
- Obtaining the information required to determine the accuracy of our results.
- Learning to interpret the spectra efficiently.

Tools

- Software
  - Automated Mass Spectral Deconvolution and Identification System (AMDIS)
  - National Institute for Standards and Technology (NIST) MS Search
  - Ion Signature Quantitative Deconvolution (IFD)
- Hardware
  - GC/MS device

Methodology

- Utilize IFD to identify the TIC peaks, then select the most significant ones.
- Run a similarity (based for highest match/rmatch ratio) and an identity search (based on probability) using IFD.
- Interpret the results between these searches.
- Export the TIC peak spectra to NIST.
- Use NIST to identify the spectra and test if it agrees with IFD results.
- Insert entry into contaminants spreadsheet.

Automated Mass Spectral Deconvolution and Identification System (AMDIS)

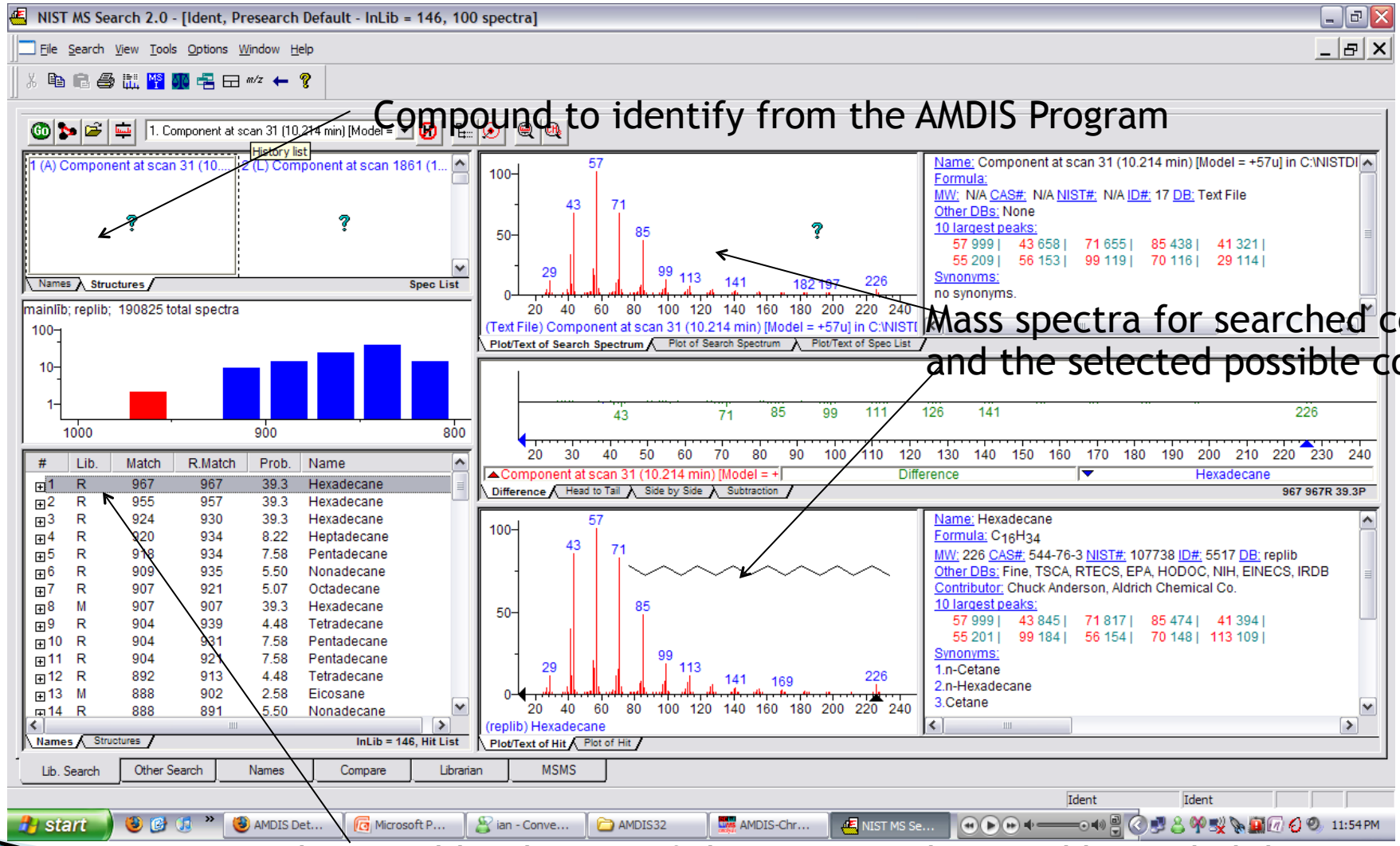
- AMDIS is a software used to extract clean spectra from complex GC-MS analysis.
- It allows you to build your own target compound library.
- It is useful when you have noisy Total Ion Current (TIC) spectral files.
- It deconvolutes the data, matching the compounds and finds target compounds from the selected library.
- It accepts large data files in different formats (.cdf, .ms, .d, etc.)

National Institute of Standards and Technology (NIST) MS Search

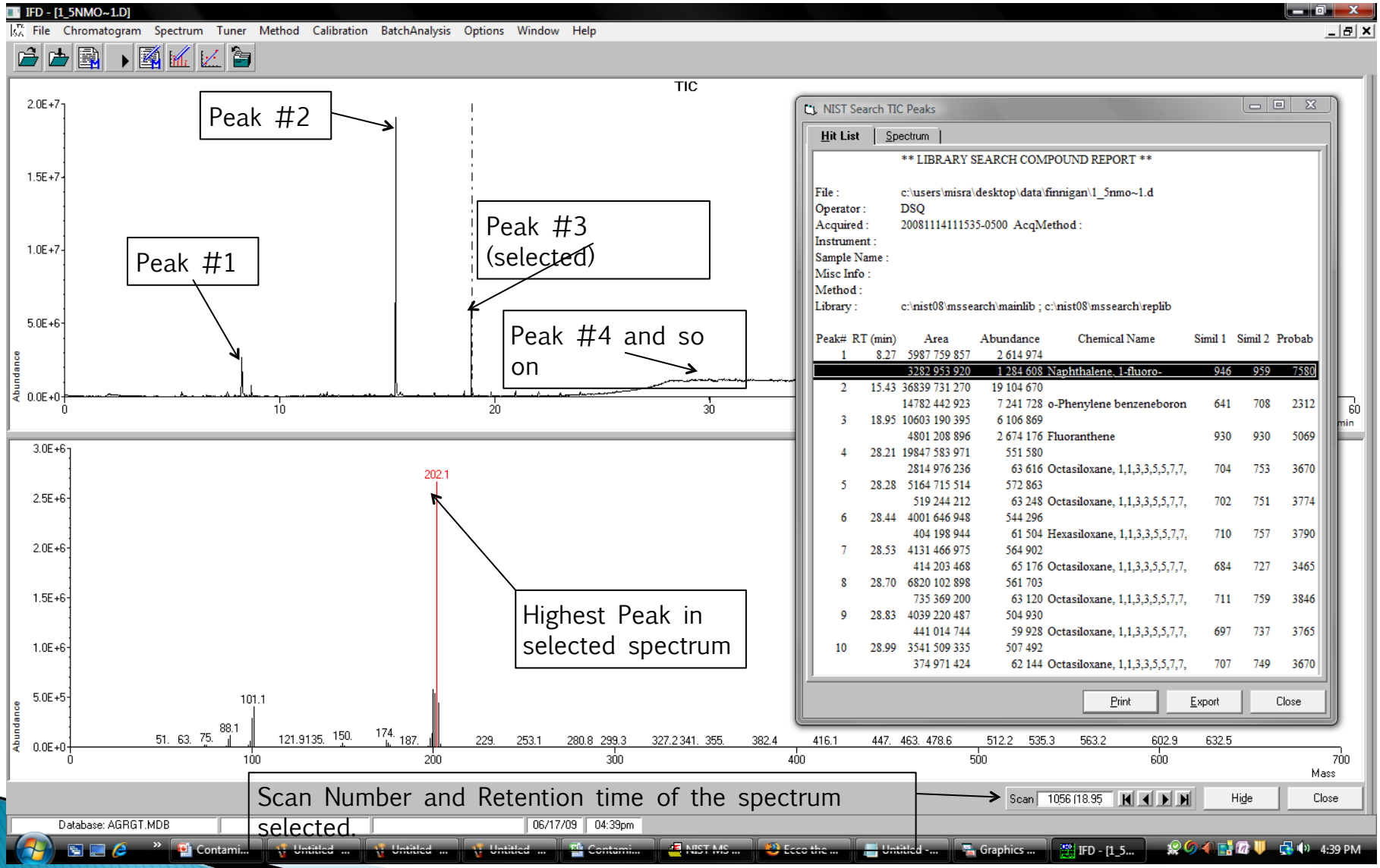
- When you have the MS for a compound at a certain retention time in the GC, you can run a search command to see if the compound matches an entry in the NIST library.

- The NIST library search program returns the compounds that are most likely to be the compound you are searching for.

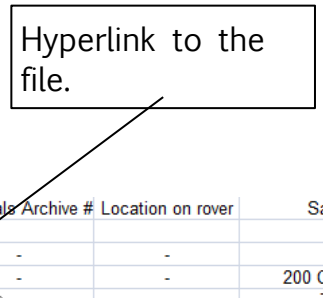
NIST MS Search Interface

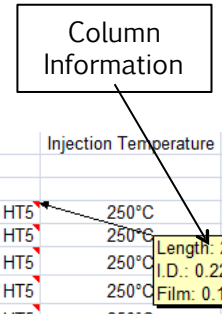


Ion Signature Quantitative Deconvolution: TIC Peak Identification



File Database





	File Name & Path	Materials Archive #	Location on rover	Sample Name	GC Column Type	Injection Temperature	Ramp Temp Rate	Final Temperature
2								
3	Finnigan\CDSpyro\CaCO30320	-	-	-			200°C/min	600°C
4	Finnigan\duralco_tefzel\Duralco200d1	-	-	200 C Duralco 4460	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C
5	Finnigan\duralco_tefzel\tefzelwire	-	-	Tefzel wire	SGE Forte GC Capillary Column HT5	250°C	20°C/min	50°C
6	Finnigan\Insulation\HCtrap\306bla	-	-	Chemglaze306blackd1	SGE Forte GC Capillary Column HT5	250°C	20°C/min	100°C
7	Finnigan\Insulation\HCtrap\Cham	-	-	chamber 200Cd2	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C
8	Finnigan\Insulation\HCtrap\Chothermt412d1	-	-	chothermd1	SGE Forte GC Capillary Column HT5	250°C	20°C/min	100°C
9	Finnigan\Insulation\Hitemppoly0407\KetronPEEKd1	-	-	Ketron Peek 1070 at 250 C	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
10	Finnigan\Insulation\Hitemppoly0407\Labair0407d1	-	-	Lab air one week	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
11	Finnigan\Insulation\Hitemppoly0407\torlon4203d1	-	-	Torlon 4203 at 250C	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
12	Finnigan\Insulation\Rampbond\MSmoldreleased1	-	-	m-s mold release	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
13	Finnigan\Kingston\306black\306black01	-	-	blank1	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
14	Finnigan\Kingston\Bunan\BunaN	-	-	Buna N o-rings	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
15	Finnigan\Kingston\Dacrond\Dacrond1	-	-	Gudebrod Braided dacron	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
16	Finnigan\Kingston\Delrind\Delrind1	-	-	Delrin reextract after 2 weeks	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
17	Finnigan\Kingston\JPLstuff7\kaptfilmd1	-	-	Kapton film adhesive 60C	SGE Forte GC Capillary Column HT5	250°C	20°C/min	60°C
18	Finnigan\Kingston\Nomex_Dacron\dacrond1	-	-	sample2d1	SGE Forte GC Capillary Column HT5	250°C	20°C/min	150°C
19	Finnigan\Kingston\Nusil\Nusil2646d1	-	-	Nusil CV2646	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
20	Finnigan\Kingston\Nylon\nylond1	-	-	Gudebrod braided nylon	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
21	Finnigan\Kingston\sauereisen_parofluor\parofluord2	-	-	sample2d2	SGE Forte GC Capillary Column HT5	250°C	20°C/min	225°C
22	Finnigan\O-ring_parafluor\O-ring_parafluor\SamOringparafluor01d1	-	-	Oringparafluor sample 1 225C	SGE Forte GC Capillary Column HT5	250°C	20°C/min	225°C
23	Finnigan\O-ring_parafluor\Polymerics\Sixday\poly_carboxen01	-	-	Blank	SGE Forte GC Capillary Column HT5	250°C	20°C/min	
24	Finnigan\O-ring_parafluor\SAMGC\Airseal22 1hr	-	-	Airseal22 1 hr	SGE Forte GC Capillary Column HT5	250°C	20°C/min	50°C
25	Finnigan\Wire\Dry lubed wire\dry_lubed1	-	-	Dry lube JPL	SGE Forte GC Capillary Column HT5	250°C	20°C/min	

Contaminants Database

	Compound name in Library	Retention Time	Compound Name (most probable)	Formula	Probability	Match	R. Match	Library
2								
3	Finnigan\1_5nmolstandard.D							
4								
5	Naphthalene, 1-fluoro- in Finnigan/1_5nmolstandard.D at RT = 8.268	8.268	Naphthalene, 1-fluoro-	$C_{10}H_7F$	75.8	948	964	Mainlib
6	Unknown compound in Finnigan/1_5nmolstandard at RT = 15.416	15.416	o-Phenylene benzeneboronate	$C_{12}H_9BO_2$	34.2	669	733	Mainlib
7	Fluoranthene in Finnigan/1_5nmolstandard.D at RT = 18.945	18.945	Flouranthene	C16H10	57	926	926	Mainlib
8								
9	SuiteBO\SuiteB0_0828.D							
10	Alkane Hydrocarbon in SuiteBO\	5.946	Undecane, 3-methyl-	C12H26	9.42	748	803	Mainlib
11	Alkane Hydrocarbon in SuiteBO\	6.068	Decane	C10H22	6.17	766	873	Mainlib
12	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.207	6.207	Dodecane, 2,6,10-trimethyl-	C15H32	22.1	768	881	Replib
13	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.351	6.351	Tetradecane	C14H30	6.27	725	859	Replib
14	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.406	6.406	Decane, 2-methyl-	C11H24	11.2	776	868	Mainlib
15	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.484	6.484	Heptane, 5-ethyl-2,2,3-trimethyl-	C12H26	7.44	726	847	Mainlib
16	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.545	6.545	Dodecane, 2,6,11-trimethyl-	C15H32	8.38	784	838	Mainlib
17	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.722	6.722	Decane, 2,3,5,8-tetramethyl-	C14H30	5.89	765	802	Replib
18	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.839	6.839	Tetradecane	C14H30	7.27	782	849	Mainlib
19	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.938	6.938	Undecane, 4-methyl-	C12H26	15.2	751	844	Mainlib
20	Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.988	6.988	Undecane, 4-methyl-	C12H26	19.9	729	861	Mainlib
21	Unknown compound in SuiteBO\SuiteB0_0828.D at RT = 8.385	8.385	Formyl colchicine	C23H25NO7	31	524	545	Mainlib
22	Unknown compound in SuiteBO\SuiteB0_0828.D at RT = 9.471	9.471	(3,4-Dimethyl-5-oxo-2,5-dihydro-1H-p	C27H41N3O2S2	11.7	490	508	Mainlib
23	Unknown compound in SuiteBO\SuiteB0_0828.D at RT = 10.236	10.236	Pentanoic acid, 2,2,4-trimethyl-3-carb	C16H30O4	9.01	564	809	Mainlib

file:///C:\Users\Misra\Desktop\Fileinformation.xls - Sheet1!A849 - Click once to follow. Click and hold to select this cell.

Link to file database.

Unknown compounds have very low probability, do not meet user criteria and have a low match ratio.

Project Timeline: Current Work

- Retrieve and tabulate information regarding the data files and their run parameters, rover location, etc.
- Retrieve the temperatures for the data files based on the method used.
- Consult with other researchers to ensure that the results returned by our analysis matches with their expectations of what contaminants should be present.
- Integrate the new data into the target library and monitor how effective it is in successfully identifying compounds found in the rover's environment to ensure correct contaminant detection.

Project Timeline: Future Work

- Continue analyzing more test data files as well as real-time data as it becomes available.
- Wrap-up the contaminants data base for the Finnigan data.
- Develop a program prompt that will ease the retrieval of the file information.
- Analyze Coldfinger (CF) data.

Issue #1: Method Inconsistencies

- Is the method obtained by Doris Jallice and Dr. John Canham the same?

TDSLc.meth

```
                                    jcanham          Oven                    Method
     Initial    Temperature     (C):                  50
     Initial    Time    (min):                        2.00
     Number    of    Ramps:                           1
     Rate    #1    (deg/min):                        20.0
     Final    Temperature     #1    (C):             330
     Hold    Time    #1    (min):                    5.00
     Post    Run    Temperature:                     Off
     Enable    Cryogenics:                          Off
     Maximum    Temperature     (C):                450
     Prep    Run    Timeout    (min):                360.00
     Equilibration     Time    (min):                0.00
-------------------------------------------------------------------
Right    SSL    Method
     Base    Temperature:                           On
     Base    Temperature     (C):                   250
     Mode:                                          Splitless
     Split    Flow:                                 On
     Split    Flow    Flow    (ml/min):             20
     Splitless     Time    (min):                   1.00
     Surge    Pressure:                             Off
     Surge    Pressure     (psi):                   0.44
     Surge    Duration     (min):                   0.00
     Constant    Purge:                             On
```

Ramp Temp Rate Entry in File database.

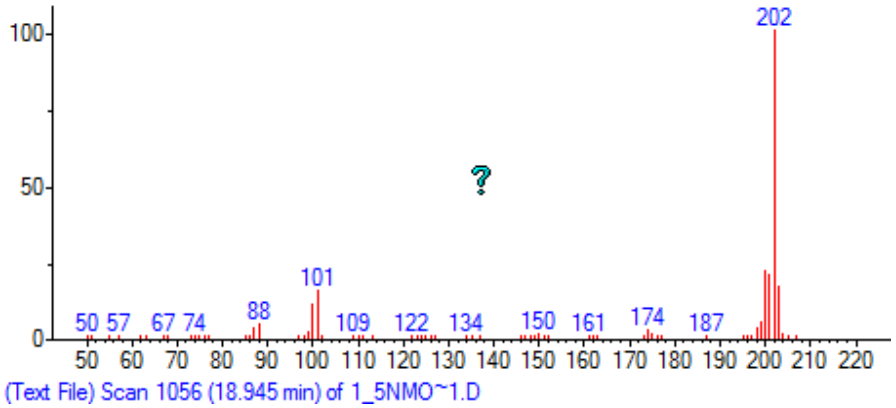
Final Temperature Entry in File database***.

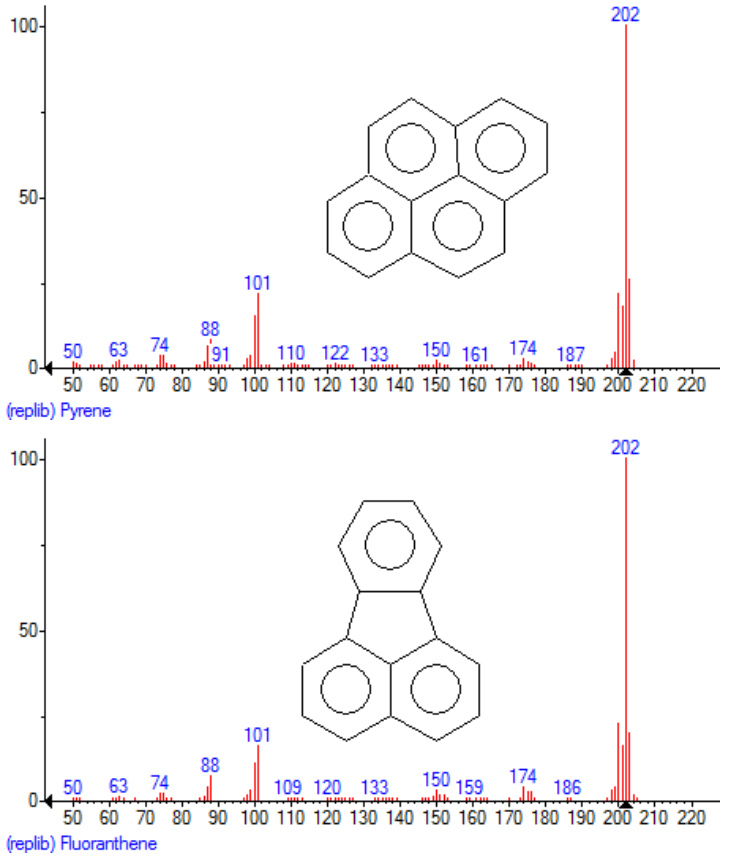
Injection Temperature Entry in File database.

Issue #2: Compound similarity

- Search 1_5nmolstandard.D at RT = 18.945.
  - Fluoranthene ($C_{16}H_{10}$) @ 57.0%, 926/926 Match
  - Pyrene ($C_{16}H_{10}$) @ 35.7%, 909/909 Match

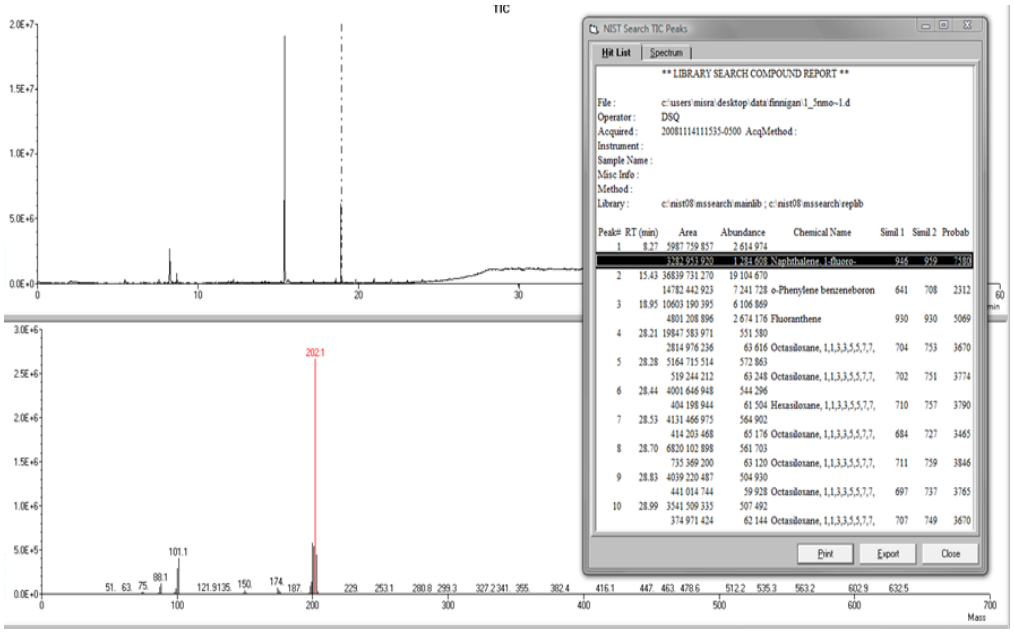
- Which compound is correct?

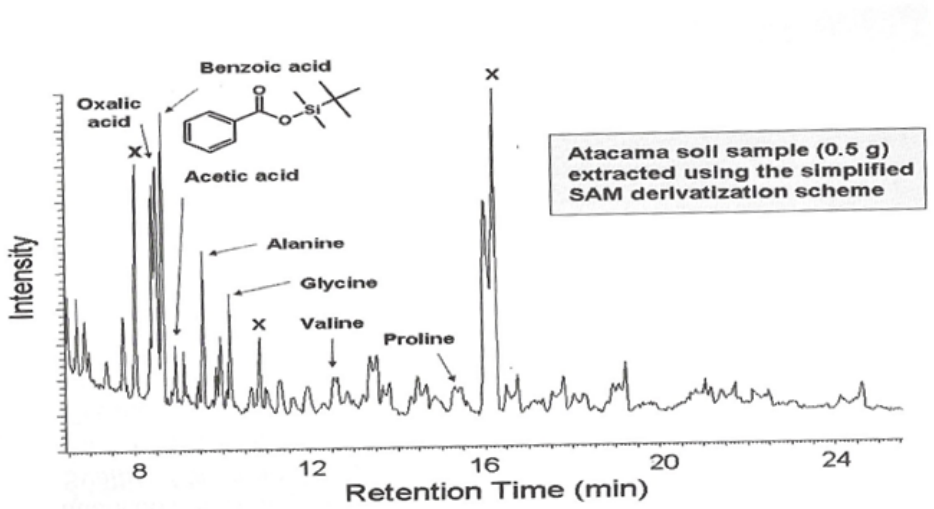




Issue #3: TIC Peak criteria

- What are acceptable criteria to determine that a TIC peak is significant?





GCMS Analysis of Soil Sample, SAM CDR 6-feb-2007, p.46, D. McLennan

Issue #4: File Spreadsheet Information

- Missing and uncertain information on files.

What format is the archive number and where is the information located?

Unable to find information regarding location on rover of materials analyzed.

What is the nomenclature for naming samples?

Why are all sample types listed as unknown?

File path and name	Materials Archive Num.	Location on rover	Sample Name	Sample Type	Method	GC Column Type	Injection Temp	Ramp Temp Rate	Final Temp	Acquisition D
Finnigan\CDSpyro\SF120t24h.D		-	-	-	CDSpyro	-		200°C/min	600°C	-
Finnigan\CDSpyro\SF120t47h.D	-	-	-	-	CDSpyro	-		200°C/min	600°C	-
Finnigan\duralco_tefzel\blnk01.D	-	-	blnk1	Blank	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min		08/30/07 03:1(
nigan\duralco_tefzel\blnk02_070830160749.D	-	-	blnk2	Blank	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min		08/30/07 04:07
Finnigan\duralco_tefzel\blnk03.D	-	-	blnk3	Blank	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min		08/31/07 04:5(
Finnigan\duralco_tefzel\Duralco200d1.D	-	-	200 C Duralco 4460	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/23/06 03:2(
Finnigan\duralco_tefzel\Duralco200d2.D	-	-	200 C Duralco 4460 d2	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/24/06 04:27
Finnigan\duralco_tefzel\Duralco200d3.D	-	-	200 C Duralco 4460 d3	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/24/06 09:1(
Finnigan\duralco_tefzel\duralco4525d1.D	-	-	sample1d1	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 05:32
Finnigan\duralco_tefzel\duralco4525d2.D	-	-	sample1d2	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 07:04
Finnigan\duralco_tefzel\duralco4525d3.D	-	-	sample1d3	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 08:37
Finnigan\duralco_tefzel\duralco4525d4.D	-	-	sample1d4	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 10:0(
innigan\duralco_tefzel\postduralco4525d1.D	-	-	postsample1d1	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 06:1(
innigan\duralco_tefzel\postduralco4525d2.D	-	-	postsample1d2	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 07:5(
innigan\duralco_tefzel\postduralco4525d3.D	-	-	postsample1d3	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 09:23
innigan\duralco_tefzel\postduralco4525d4.D	-	-	postsample1d4	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	210°C	08/30/07 10:55
innigan\duralco_tefzel\posttefzelm22759d1.D	-	-	postsample2d1	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 01:0(
innigan\duralco_tefzel\posttefzelm22759d2.D	-	-	postsample2d2	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 02:33
innigan\duralco_tefzel\posttefzelm22759d3.D	-	-	postsample3d3	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 04:0(
Finnigan\duralco_tefzel\tefzelm22759d1.D	-	-	sample2d1	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 12:14
Finnigan\duralco_tefzel\tefzelm22759d2.D	-	-	sample2d2	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 01:46
Finnigan\duralco_tefzel\tefzelm22759d3.D	-	-	sample3d3	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	200°C	08/31/07 03:1(
Finnigan\duralco_tefzel\tefzelwire.D	-	-	Tefzel wire	Unknown	TDSLc	SGE Forte GC Capillary Column HT5	250°C	20°C/min	50°C	08/08/06 02:47