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

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

Overview

- 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

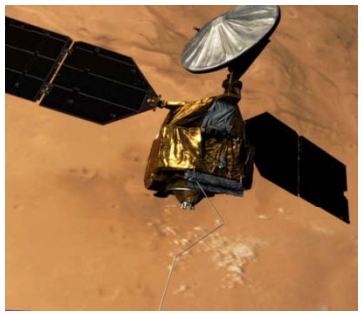
- Analyze data from the rover environment using Gas Chromatography-Mass Spectroscopy (GC-MS).
- Develop an organic contaminants database for the Sample Analysis at Mars (SAM) utilizing GC-MS analysis tools.
- Utilize this database to identify compounds from actual mission samples.

Research Questions

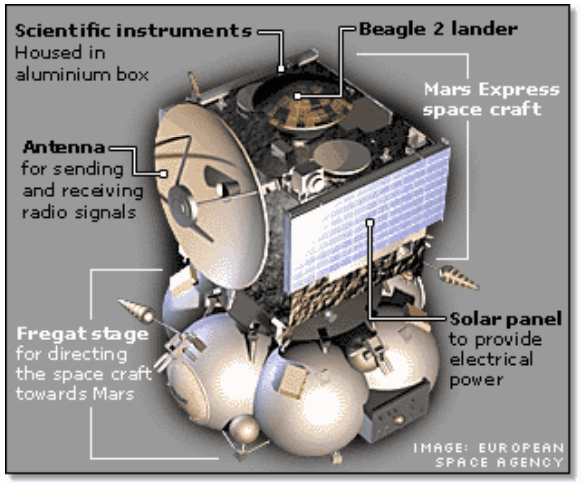
- What organic compounds can be found in the rover structure? Where in the rover can they be found?
- How accurately can we determine the compounds from tests to those on the actual rover?
- How will identifying these compounds help when the time comes for the actual mission analysis?

NASA Mars Exploration Timeline

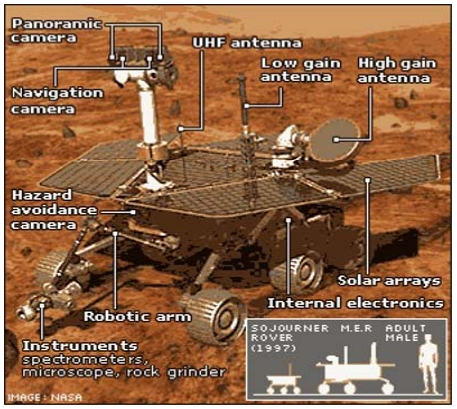
- **2001:**
  - 2001 Mars Odyssey - 7 April 2001 - Mars Orbiter



    - Still in orbit as of March 2008, and the ending date was extended to September 2008.
    - Odyssey's Gamma Ray Spectrometer detected significant amounts of hydrogen, which was thought to be contained in large deposits of water ice.
- **2003:**
  - Mars Express - 2 June 2003 - Mars Orbiter and Lander
    - Lander failed during descent and was declared lost in early February 2004.
    - In early 2004 the Planetary Fourier Spectrometer team announced it had detected methane in the Martian atmosphere. ESA announced in June 2006 the discovery of aurorae on Mars.



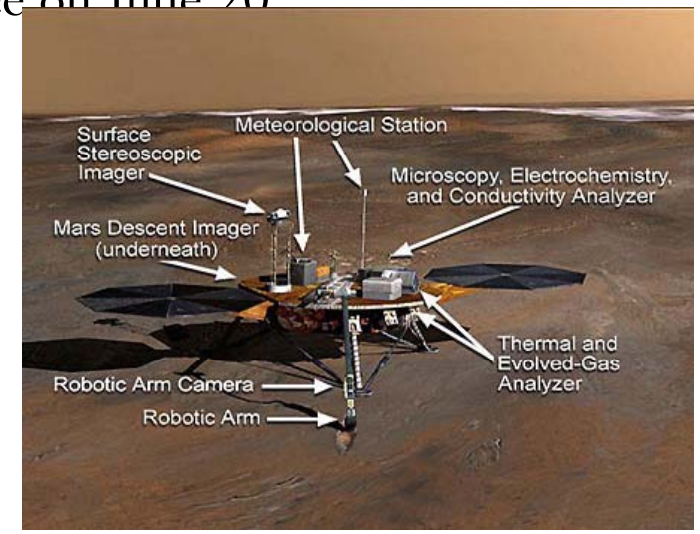
NASA Mars Exploration Timeline



- **2003:**
  - Spirit (MER-A) - 10 June 2003 - Mars Rover
  - Opportunity (MER-B) - 7 July 2003 - Mars Rover
    - Both missions landed successfully in January 2004 and
    - have met or exceeded all their targets.
    - Among the most significant scientific returns has been conclusive evidence that liquid water existed at some time in the past at both landing sites.
    - Martian dust devils and windstorms have occasionally cleaned both rovers' solar panels, and thus increased their lifespan.
- **2005:**
  - Mars Reconnaissance Orbiter - 10 August 2005 - Mars Orbiter
    - The Mars Reconnaissance Orbiter snapped the first image of a series of active avalanches near the planet's north pole, scientists said March 3, 2008.

NASA Mars Exploration Timeline

- **2007**
  - Phoenix - 04 August 2007 - Mars Scout Lander
    - The Lander has a robotic arm with a
    - 2.5 m reach and capable of digging a
    - meter into the Martian soil.
    - It also has a microscopic camera capable of resolving to one-thousandth the width of a human hair, and discovered a substance at its landing site on June 15, 2008, which was confirmed to be water ice on June 20



Mars Science Laboratory (MSL)

- "Curiosity" is the newest rover in NASA's Mars Exploration Program expected to launch in 2011.
  - It is larger and can travel farther than previous rovers such as Spirit and Opportunity.

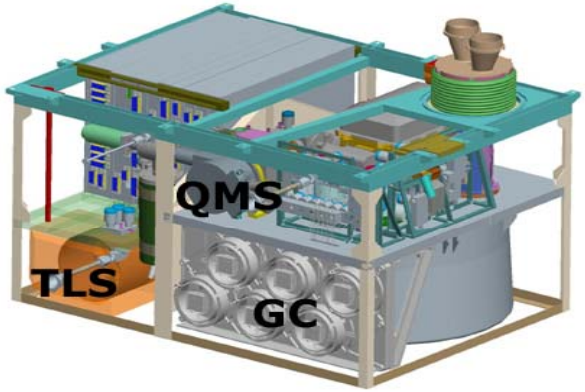
  - Contains the most advanced suite of instruments (such as SAM) to date. It will also make use of innovative technologies for landing.

MSL Goals

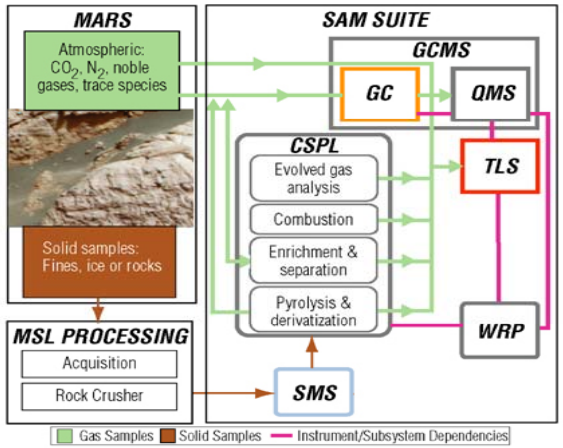
- Assess the biological potential by determining the nature and inventory of organic carbon compounds, etc.
- Investigate past habitability on Mars by determining information about carbon dioxide and water.
- Characterize the geology of the landing region and interpreting the processes that have formed rocks and soils.
- Characterize the broad spectrum of surface radiation.

Sample Analysis at Mars (SAM)

- It consists mainly of three devices for chemical analysis:
  - Gas Chromatograph
  - Mass Spectrometer
  - Tunable Laser Spectrometer



- It will analyze the composition of soil and rock on Mars with the intention of finding compounds that are associated with life.



SAM Process

- The mass spectrometer will divide the sample and its elements by mass.
- The gas chromatograph will heat the sample until they vaporize, and will separate the gases so the compounds arrive at different intervals.
- The laser spectrometer will measure the abundance of organic isotopes such as carbon, hydrogen and oxygen in the compounds.

Tools

- Software
  - Automated Mass Spectral Deconvolution and Identification System (AMDIS)
  - National Institute for Standards and Technology (NIST) MS Search
  - Ion Fingerprint Deconvolution (IFD)

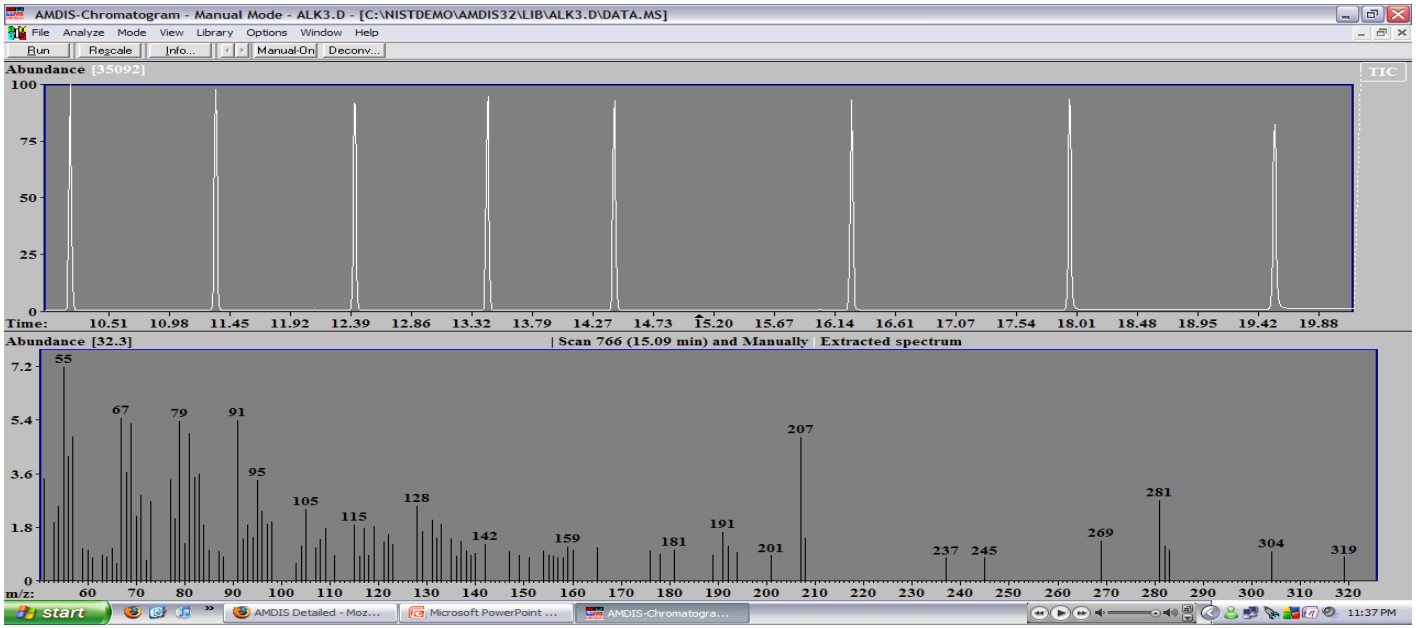
Automated Mass Spectral Deconvolution and Identification System (AMDIS)

- AMDIS is a free software used to extract clean spectra from complex GC-MS analysis.
- It allows you to build your own target compound library. This allows you to analyze the samples and match them with the components you want.
- It is useful when you have noisy Total Ion Current (TIC) spectra files.
- It allows you to search components in the NIST library through the NIST MS Search program.
- It deconvolutes the data, matching the compounds and finding target compounds from the selected library.
- It accepts large data files in different formats.

AMDIS Procedure

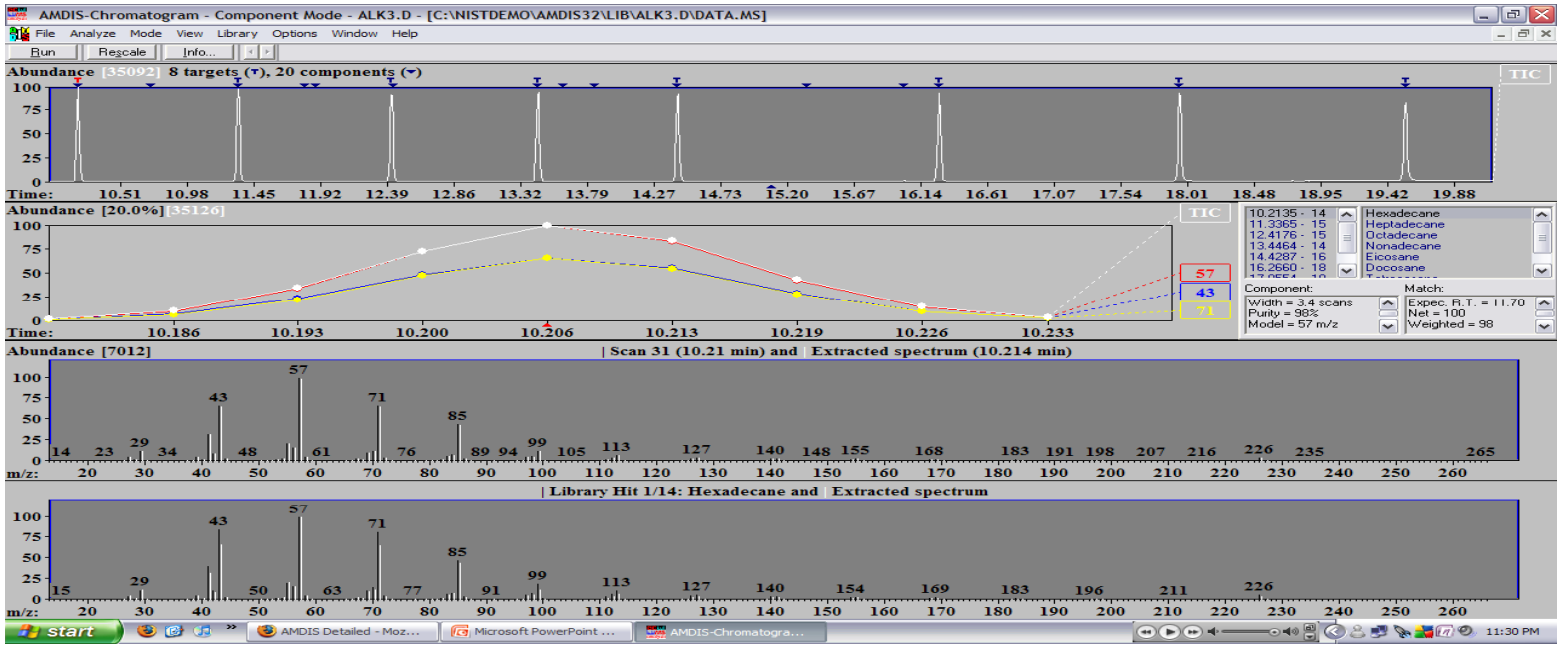
- AMDIS analyzes the background and calculates a noise level for later processing.
- It then analyzes the data for an increase of a special ion trace.
- If there is maxima for other traces at the same time, it assumes there is a peak and shapes a model peak.
- From there, it calculates clean spectra for each peak.
- Lastly, it identifies the compound via a library search.

AMDIS Interface



- The confirm window in manual mode: Above is the GC data and below the MS data for the selected GC retention time.

AMDIS Interface: Deconvolution Analysis



Confirm window after deconvolution analysis: The upper window is the GC data. The graph depicts the TIC over a period of time for the chosen component. The analysis identifies with an arrow the compounds found in the GC data, and with a T those compounds it found in the search library. The lower windows are the compound spectra.

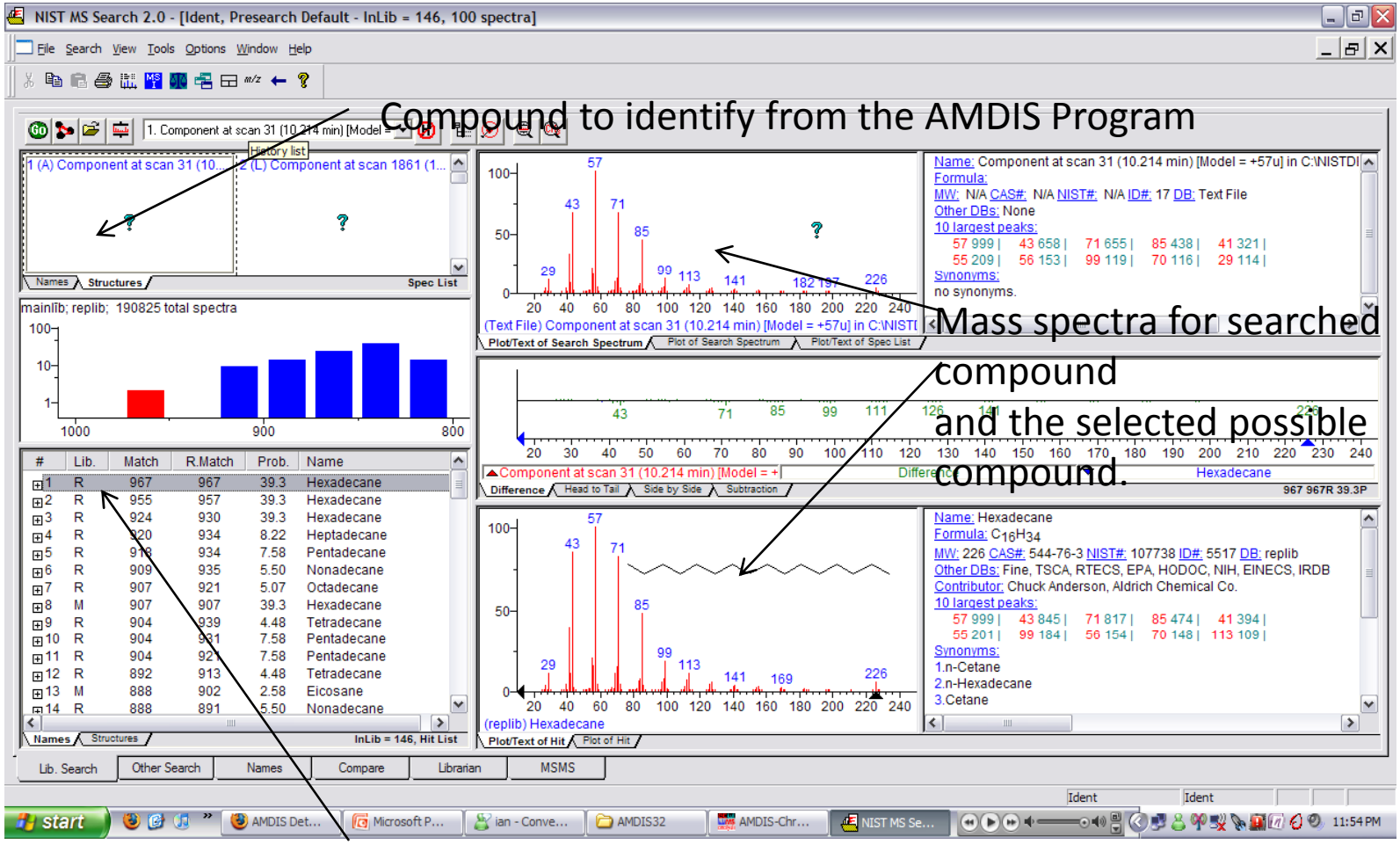
National Institute for Standards and Technology (NIST) MS Search

- 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 the identity of the compound you are searching for.

- It can search by structure or spectra.

NIST MS Search Interface



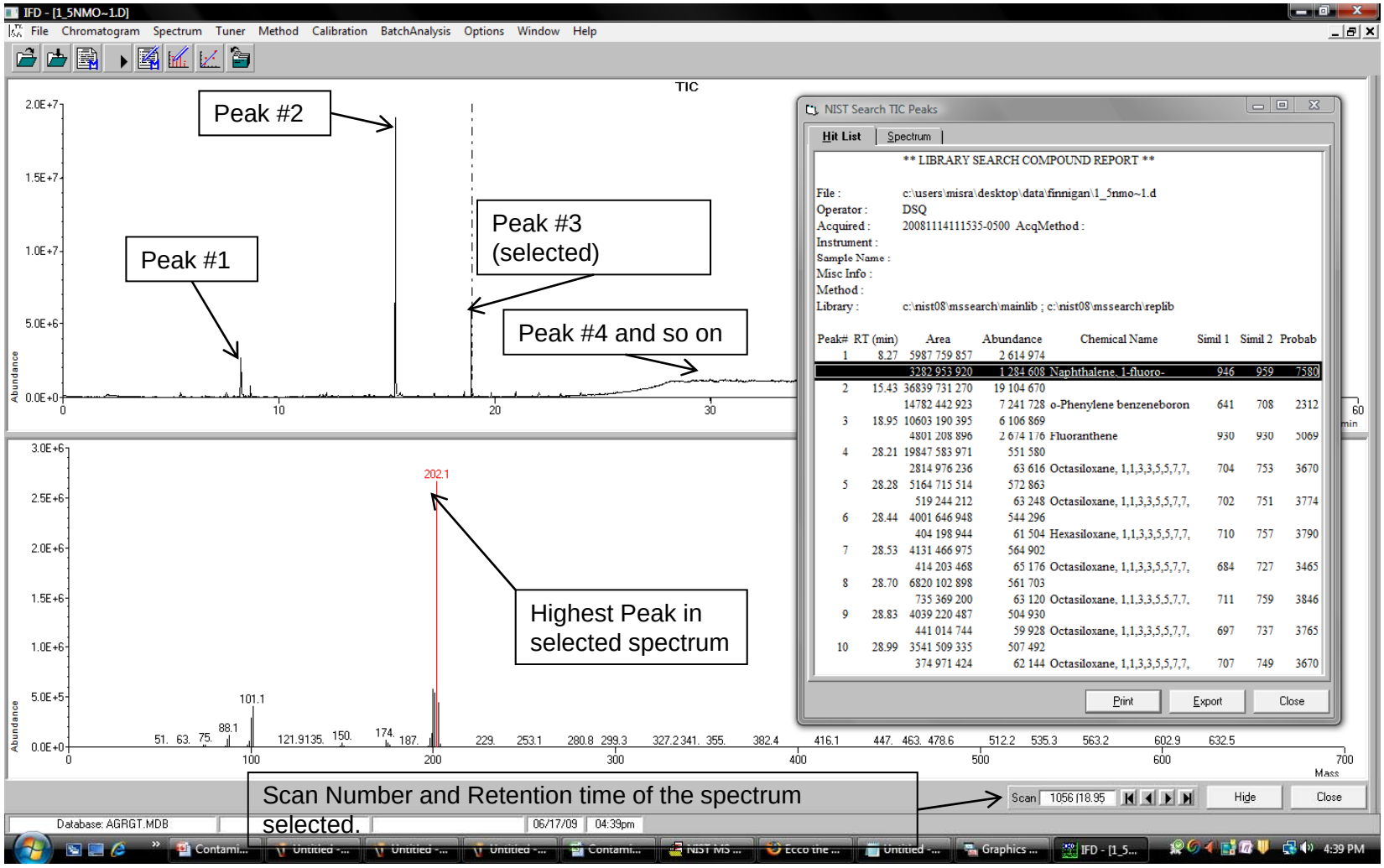
Ion Signature Quantitative Deconvolution (ISQD) Software

- Features a set of advanced deconvolution algorithms that identify target compounds present even in difficult and complex mixtures.

- The software extracts each compound's characteristic signature, untangling it from the surrounding sample matrix.

- Will be able to identify compounds “ultra-fast”.

Ion Signature Quantitative Deconvolution: TIC Peak Identification



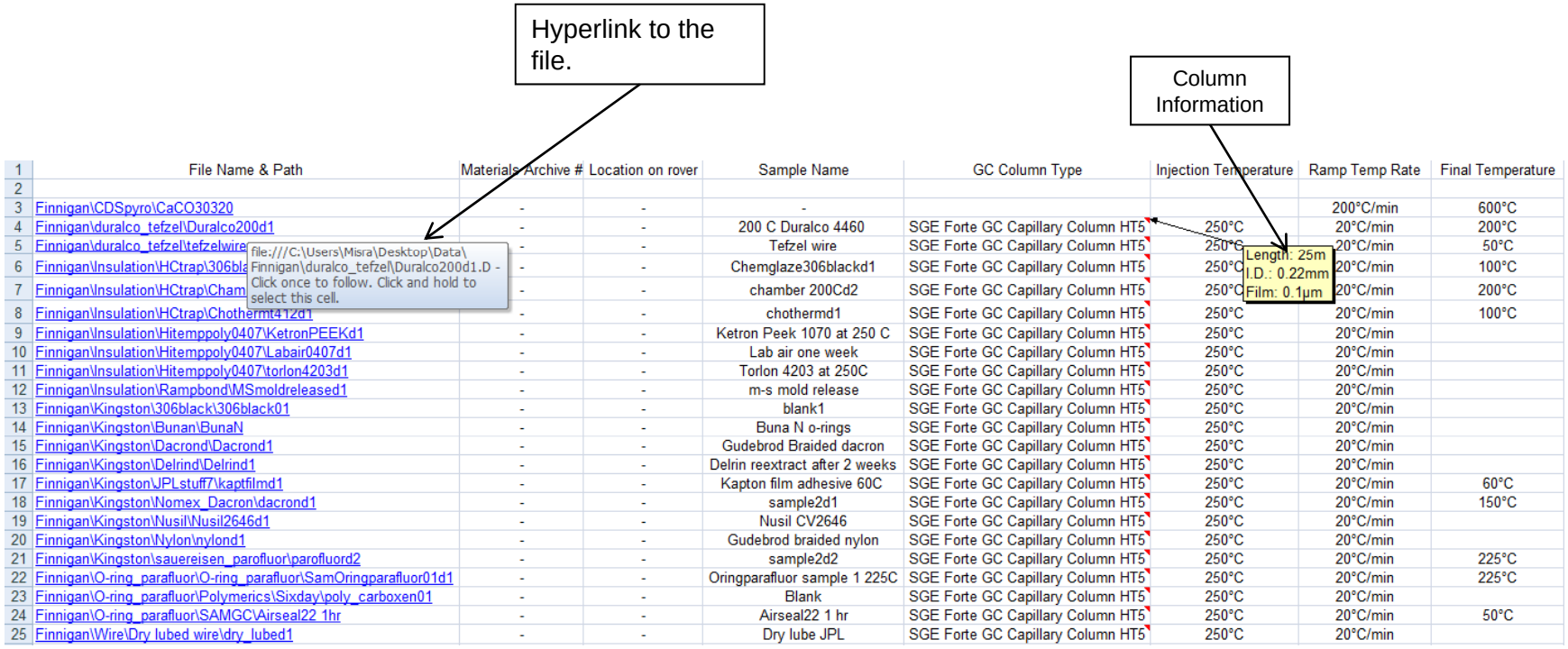
Methodology

- 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.

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.

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	

Development of File Database

- The file database is a spreadsheet file that contains file locations with information on the runtime parameters and environment in which the tests were taken.

- This database contains parameters such as: location of the file, the number by which it is identified in the material archive, location on rover, sample name and Type, the GC Column Type used, temperatures and the date the run was done.

- So far we have been able to obtain most of these parameters for over 1000 Finnigan Themo files. We are awaiting input from JPL to determine the material archive number and the location on the rover.

Contaminants Database

	Compound name in Library	Retention Time	Compound Name (most probable)	Formula	Probability	Match	R. Match	Library
1								
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.

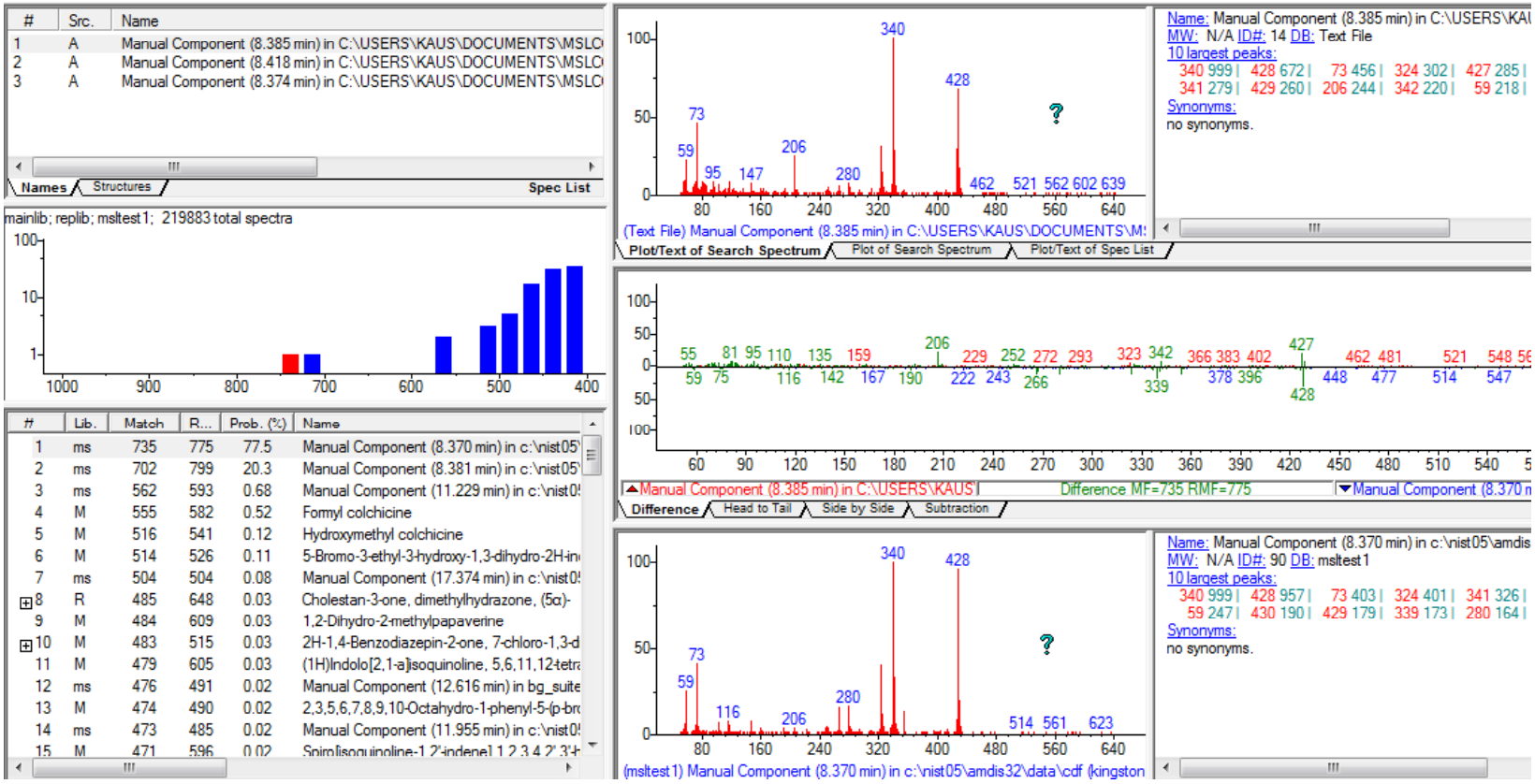
Development of Contaminants Database

- The contaminants database is a reference spreadsheet file that contains the compounds found in each file and how they can be identified by their library name.

- It lists the name and formula of the most likely compound according to the NIST search that matches an unknown compound. It also lists how similar the two compounds are in terms of their peaks (match/rmatch) and their abundance (percentage).

- Based on how similar they are, they are appropriately named in the library.

Target Compound Library



#	Lib.	Match	R...	Prob. (%)	Name
1	ms	735	775	77.5	Manual Component (8.370 min) in c:\nist05
2	ms	702	799	20.3	Manual Component (8.381 min) in c:\nist05
3	ms	562	593	0.68	Manual Component (11.229 min) in c:\nist0
4	M	555	582	0.52	Formyl colchicine
5	M	516	541	0.12	Hydroxymethyl colchicine
6	M	514	526	0.11	5-Bromo-3-ethyl-3-hydroxy-1,3-dihydro-2H-in
7	ms	504	504	0.08	Manual Component (17.374 min) in c:\nist0
8	R	485	648	0.03	Cholestan-3-one, dimethylhydrazone, (5α)-
9	M	484	609	0.03	1,2-Dihydro-2-methylpapaverine
10	M	483	515	0.03	2H-1,4-Benzodiazepin-2-one, 7-chloro-1,3-d
11	M	479	605	0.03	(1H)Indolo[2,1-a]isoquinoline, 5,6,11,12-tetra
12	ms	476	491	0.02	Manual Component (12.616 min) in bg_suite
13	M	474	490	0.02	2,3,5,6,7,8,9,10-Octahydro-1-phenyl-5-(p-br
14	ms	473	485	0.02	Manual Component (11.955 min) in c:\nist0
15	M	471	596	0.02	Spiro[isoquinoline-1,2'-indene]-1,2,3,4,2',3'-h

Development of Target Compound Library

- On the upper side of the figure we have the name, spectral and other information regarding our unknown spectrum, in the middle portion is a peak comparison and the lower part is the compound that most likely matches your unknown compound.

- We see that the spectra match very closely to each other and the peaks are similar as well in terms of abundance. After this, we go to our contaminants database, look for the component in the library by following the file path. Then we can determine what the known compound was identified as and how probable it is.

Project Timeline: Future Work

- Continue analyzing more test data files as well as real-time data as it becomes available.
- Continue updating the contaminants database for the Finnigan data.
- Develop a program prompt that will ease the retrieval of the file information.

Acknowledgements

- We would like to acknowledge with appreciation the financial support from NASA GSFC and the DC Space Grant Consortium (DCSGC) that allowed Raul Garcia-Sanchez to spend the 2009 summer in the Atmospheric Experiments Laboratory (Code 699 GSFC) and the award of a Faculty Summer Fellowship to Dr. Prabhakar Misra.
- We would like to gratefully acknowledge the support and guidance of Dr. Paul Mahaffy and his research group in this project.