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



Raul Garcia-Sanchez¹, Paul Mahaffy², Prabhakar Misra¹ ¹Howard University, ²NASA Goddard Space Flight Center Code 699

Introduction

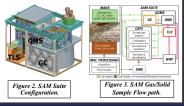
HOWARD UNIVERSITY

The primary purpose of the research is to develop an organic contami nants 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 Gas Chromatography Mass Spectral (GCMS) data

The MSL is NASA's Mars Exploration Program newest rover. It contains the most advanced suite of instruments to date. It is the biggest rover developed and will implement an innovative landing system



Sample, Analysis at Mars (SAM) is the largest payload of the MSL rover. It contains equipment for chemica analysis such as a Gas Chromatograph (GC), mass Spectrometer (MS) and Tunable Laser Spectrometer (TLS). With these instruments, it will analyze the composition of soil and rock on Mars with the intention of finding organic compounds that are associated with life. The QMS detects gases sampled from the atmosphere or those released from solid samples by heating. The GC can separate out individual gases from a complex mixture into molecular components for QMS and GC stand alone analysis. The TLS implements a sensitive search for methane and makes precision measurements of oxygen and carbon isotope ratios in carbon dioxide

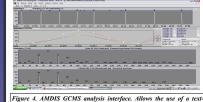


Research Questions

· 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?

Software

1. Automated Mass Spectral Deconvolution and Identification System (AMDIS): Used to extract spectra from the GCMS data we obtain through heating tests and creating a target compounds library.



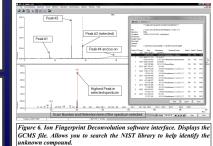
based target compound library used for deconvolution of the GCMS data.

National Institute for Standards and Technology (NIST) MS Search: Used for searching organic contaminant libraries for spectra to match the unknown compounds in our GCMS data runs

Compound to	identify fror		Program	
	identify fror			
	,	ters (encount it was 14 / 17 / 14 mini Mart	
son 100' (L. 🏦 100'	((Jars)		
				d++674 a 05/670/a
		45.50	CORT MODELE NAME TO DE TOOL	2 8
	1	(25a) 28		
80	· * *	7	CORR.	1 10 10
	\sim	91 Dia 20	a wanaj samaj 22158 2	29114
BACK 1	2010 10			
21 40	NO 180 TOO 120 \$40 \$88			ampound
(TecTik) Conso	entilessen 21 (10 214 mill) 30e	avidss supril (a for searched o	ompound
1, Suffer Clark	MAN CONTRACTOR	and the cele	cted nossible co	hound
		and the sele	cteu possible co	impound.
				18
-				
2.2		10 100 MD 100 100 10	e en en en en en en	210 200 200 200
- Constraints	or 21/02H mit Hot + 1	alvera		0001
	any series Years	//		10,100,000
	·	/	Nodelice	
	. /	/ <u>[</u>	-994	
an i	1~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~ 🎼	Firs 1904 RIEGE BRA HOOLD IN	E BESS FOR
	μ.			
	44 C	11000	THERE ADDRESS	
				2 700
	a	Second Second	4	
	10 50 10			
	the second rate and	the lots and loss 106am		
		4		2
	or we de			
en linee 1675				
possible identity	of the com	pounds sorte	d by probability	
		$\begin{array}{c c} \hline \\ \hline $	and the selection of the compounds sorter	The second se

rigure 5. 1951 mass spectral Search Interface. Compares between the unshown composition and loss thrute composition and further (IFD): Chromatograph. It is also used to create a database of

compounds to search for after deconvolution



Methodology

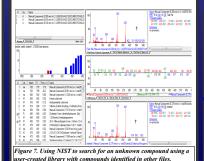
1. Utilize IFD to identify the TIC peaks, then select the most significant ones.

- 2. Run a similarity (based for highest match/rmatch ratio) and an identity search (based on probability) using IFD to see if different searches yield similar results. If they are different, compare all related
- 3. Export the TIC peak spectra to NIST.
- 4. Use NIST to identify the spectra and test if it agrees with IFD results
- 5. Insert entry into contaminants spreadsheet.

6. Add the identified (or unidentified) spectrum into a user-created NIST Library. Also add to a target compound library using IFD or AMDIS.

File path and name	Sample Name	Sample Type	Method			Ramp Temp Rate	Final Temp
Finnicari/CDSpyrel/SF120900CT15hmanb		1.1	COSpra 1	SGE GC Capitary Column HTS		200°C/min	600°C
Finnicani CDSpyral SF 1280 bcocktail			COSpra.	SGE GC Capitary Calumn HTS		200°C/min	600°C
Finnicari/CDSpuni/SF129(2)htcscktal 000			COSpira 1	SGE GC Capitary Calumn HTS		200°C/mn	60010
Finnicari (CDSpiral SF 129(24)			COSpira	SGE GC Capitary Calumn HTS		200°C/mn	60010
Fineigeni CD Sourg/SF 1204/h			COSpira	SGE GC Capitary Column HTS		200°C/min	600°C
Ennigeridurako tatza/talnk@1	bink1	Blank	TESLC	SGE GC Capillary Column HTS		20°Cinin	
Envioariduralco taballalink82 0708301567	bink2	Blank	TOSLC	SGE GC Capillary Column HTS		20°Chrin	
Ennigeriduseloo taba/talixit3	bink3	Elank	TESLC	SGE GC Capitary Column HTS		20°Chrin	
integaridutalco tatta/Dutalco208d1	200 C Duralco 4468	Unkrawn	TESLC	SGE GC Capitary Column HTS		29'Ohin	200°C
Fineiganiduralco tefze/Duralco208d2	208 C Duratico 4460 d2	Unkrown	TD6Lc	SGE GC Capitary Column HTS		29'Olmin	200°C
Fineigeniduratico tefze/Duratics208d3	200 C Duraico 4460 e3	Unknown	TOSLC	SGE GC Capitary Column HTS	260°C	20'Ohin	20010
Fineigeniduratico tetzel/duratico4525d1	sangle (d)	Unknown	TOSLC	SGE GC Capitary Calumn HTS	2000	28°Omn	21010
Finnicaniduralco tetzel/duralce4525d2	sample 1/2	Unknown	TOSLO	SGE GC Capitary Calumn HTS	20/C	28°O/min	21010
Finnicaniduralco tatzal/duralco4525d3	sample 1d3	Unknown	TOSLO	SGE GC Capillary Calumn HTS	250°C	20°Cimin	21010
Fineiganiduralco tatzal/duralco452564	sample 104	Unknown	TESLO	SGE GC Capillary Column HTS		20°Citrin	21010
Enniganiduralco talbal postduralco4525d1	pestsampla1d1	Unknown	TOSLO	SGE GC Capillary Column HTS		20°Citrin	21010
Ennicariduralco tatzal postdaralco452562	postsampia1d2	Unknown	TESLC	SGE GC Capitary Column HTS		20°Chrin	21810
Firmiganiduralco taltal postduralco4525d3	ocitiandia1d3	Unknown	TD6Lc	SGE GC Capitary Column HTS	250°C	20'Ohin	21010





Discussion

· 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

· 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.

· Development of Target Compound Library: As seen in Figure 7, 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 It can be seen that our unknown spectrum is similar to another spectrum we had found in a previous file and had placed in the user-defined library. These library compounds have the format of: <Compound Name at RT = x.xxx in Filepath> and is how we are able to reference them in the previous database spreadsheets mentioned. 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

Future Work

- · Continue analyzing more test data files as well as real-time data as it becomes available
- · Develop a database program prompt that will ease the retrieval of the file information instead of an indexed spreadsheet
- · Analyze Coldfinger (CF) data

Reference

SAM MSL GSFC Atmospheric Experiments Laboratory Website:

spectra and interpret.

Results

· Development of File Database

Compound name in Library	Retention Time	Compound Name (most probable)	Formula	Probability	Mash	1.10
timeperil: Sweeblendard.D						
Nephthalane, 1-fueros- in Finalgen/1, SomeIntendent D at \$7 = 8 208		Nephthelene, 1-Europ-	CJU .	75.8	948	96
Unknown compound in Fispigan/1, Semebland and at 87 + 15.435	15.416	c-Phenylene benzeneboronate	0,8,80,	342	665	75
Descentione in Decease/1_Semicistendent D at 87 + 18 945	18.945	tissteribere	CORMOD	37	928	92
SuteBOI/SuiteB0 0028.0						
Alkane Hedrocerbon in SuiteBO\SuiteB0 (6535.D at 87 - 5.546)	5,816	Undersee, 3-methol-	C11HQ6	9.42	748	80
Alkane Hedrocarbon in SubeBOLSuite80 0828.D at 87 - 6.068	6.058	Decene	C10H22	617	766	87
Alkane Hydrocarbon in Suite601/Suite60 8828.D at RT = 6.397	6,207	Dodecene, 2.6.10-trimethyl-	C15H02	22.1	768	- 60
Alkane Hydrocarbon in SuiteBO\SuiteB0_0828.D at RT = 6.351	6.351	Tetradecene	C14H00	6.27	725	45
Alkane Hydrocarbon in SuiteBO\SuiteB0_0028.D at RT = 6.405	8.408	Decase 2-methyl-	C11HQ4	11.3	778	86
Alkane Medrocarbon in SuffetOOLS at e60, 0525, D at 87 = 6.454	6.484	regions 5-styl-2.2.3-prinethyl-	C11HQ6	2.44	726	84
Alkane Hedrocarbon in SuiteBOLSuiteB0 0828.D at 87 - 6.545	6.545	Dodecare, 2.6.13-trimethyl-	C15H02	8.29	784	83
Alkane Hydrocarbon in SuiteBO\SuiteBd_BR28.D at RT = 6.722	6.722	Decens, 2.5.5.8-tetramethyl-	C14H00	5.89	765	- 60
Alkane Hydrocarbon in SuiteBO\SuiteB0_8828.D at RT = 6.829	6.839	Tetradecene	C14H00	7.27	782	- 84
Alkane Hydrocarbon in SutteBO\SuiteB0_0028.D at RT = 6.905	6.538	Endstane, 4-weifigt-	C12HQ8	15.2	795	84
Alizane Hydrocarbon in SuffeetOl Suiteb0_0525.D at 87 = 6.565	6.888	Undecase, 4-methyl-	C11HQ6	28.9	728	86
Unknown compound in SustemplySuiteen_DE28.D at RT = 8.385		Formal colchicine	623425407	31	534	54
Unknown compound in SuiteRO/SuiteR0_0828.0 at RT = 9.471		15.4 Dimethyl-5 avo 2.5-dihetro LP	C27H41M00252	11.7	490	50
Unknown compound in SuiteRO/SuiteR0_0828.D at RT = 18.228		Pentannia exid, 2,2,4 planethyl 3 sP	C36H0004	9-01	564	-80
Unknown compound in SuffeBO/SuiteB0_0028.D at RT = 18.292		Acres and (1.2.3.4.5.6.7.8-min/or*	CS6HQ802	4.57	538	59
Unknown compound in SufjetOO/SaiteBO_0528.D at #7 = 19.421		Beropic acid, 2-amino-3-hydroxy-, *		11.1	531	21
Unknown compound in SubeBO/SuiteB0_B828.D at R7 = 33.648		D4-Cyclopropal3.4[berg[1,2-4]acs/#		54	554	- 59
Unknown compound in Suite90/Suite80_BR28.D at RT = 11.23		5 Brono 5 ethel 5 hydroxy 1.5 d M	02246554902543		556	59