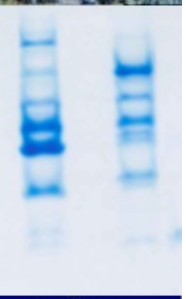
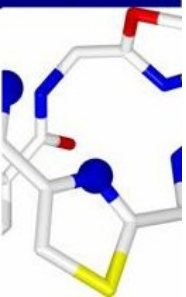




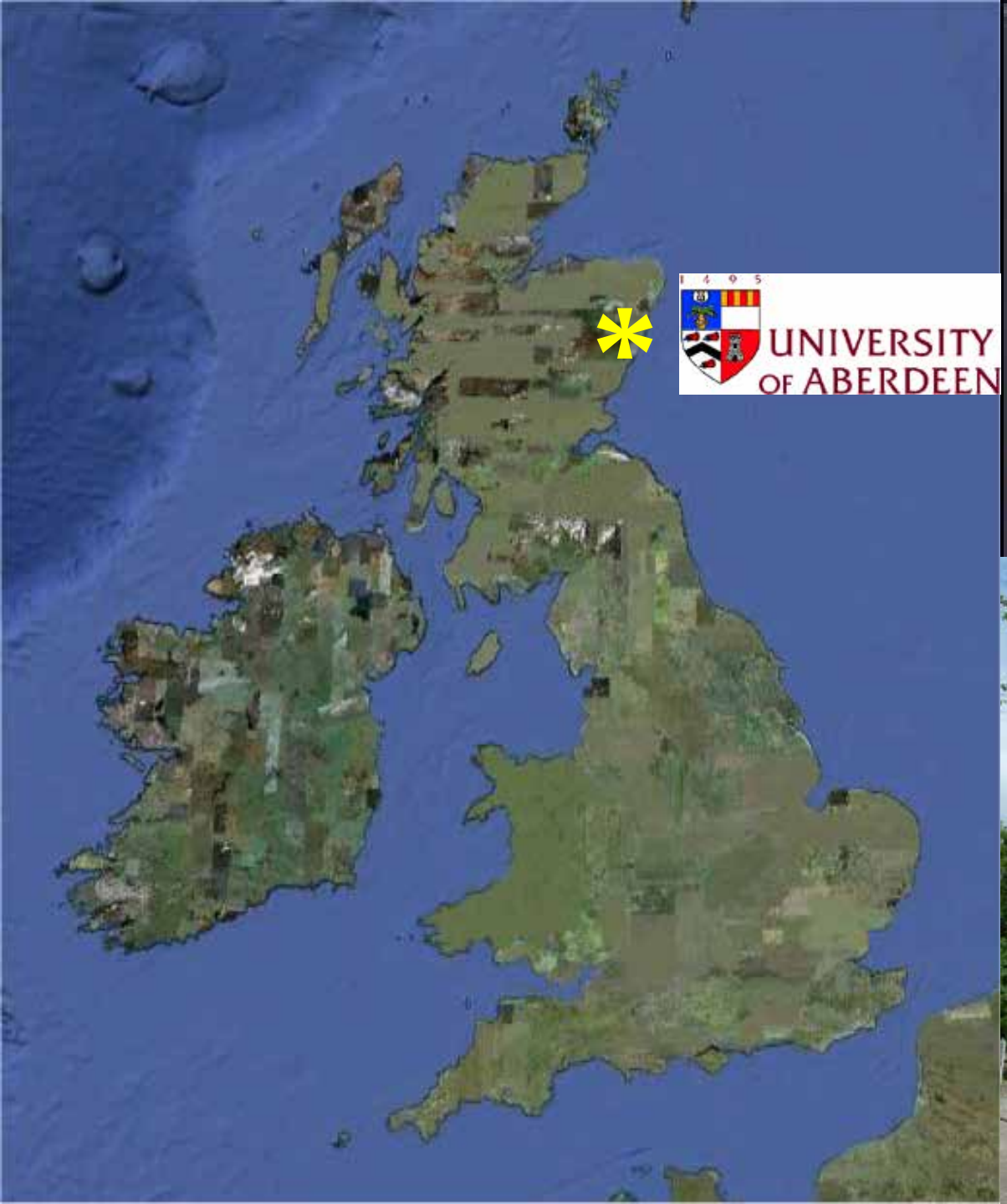


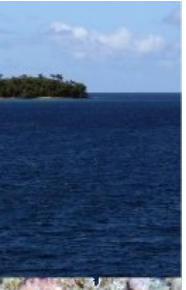
marine biodiscovery centre

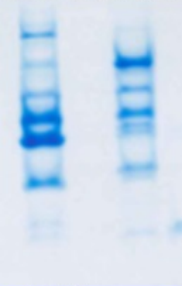
UNIVERSITY OF ABERDEEN

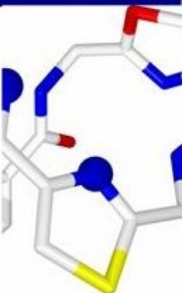
New Approaches for the Dereplication, Isolation and Structural Elucidation of New Bioactive Marine Natural Products.

Marcel Jaspars


1495
UNIVERSITY
OF ABERDEEN







marine biodiscovery centre

The Marine Biodiscovery Centre

UNIVERSITY OF ABERDEEN

Aim: The discovery of novel pharmaceutical candidates and research tools from extreme environments.

Tools

Chromatography
Spectroscopy
Molecular biology
Synthesis
Molecular modelling
New software

Marine invertebrates
Marine microorganisms
& Symbionts

Functions

Chemical ecology
Symbiosis
Biosynthesis
Bioinorganic chemistry

Applications:

Pharmaceutical
Cancer
Infection
Inflammation
Parasites
Epilepsy
Alzheimer's

People

Prof Marcel Jaspars	Natural Product Chemistry
Dr Rainer Ebel	Pharmacy & Pharmacognosy
Dr Hai Deng	Chemical Biology
Dr Wael Houssen	Synthetic Biology
Prof Jörg Feldmann	Analytical Methods
Dr Eva Krupp	Analytical Methods
Dr Laurent Trembleau	Organic Synthesis
Dr Andrea Raab	Mass spectrometry
Mr Russell Gray	Nuclear Magnetic Resonance

marine biodiscovery centre

The Biodiscovery Pipeline

UNIVERSITY OF ABERDEEN

Bottleneck

Sampling → Access to resources (Physical and legal)

Curation → Taxonomic identification; Storage with associated metadata

Biomass → Whole organism collection; Microbial cultivation/cryptic pathways

Extraction → Extraction process; **Early Dereplication**

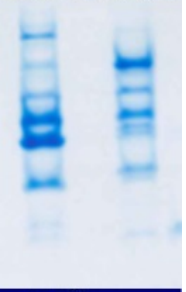
Assay → Assay throughput *vs* information obtained; False positives/negatives

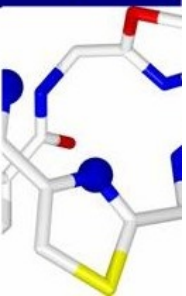
Purification → **Purification process**; Removal of nuisance compounds

Active NCE → **Structure determination**; **Late Dereplication**

Development → **SUPPLY**; **Scale-up/process intensification**





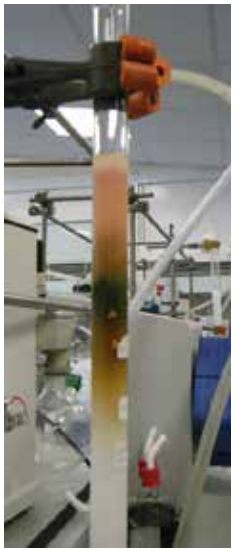


marine biodiscovery centre

UNIVERSITY OF ABERDEEN

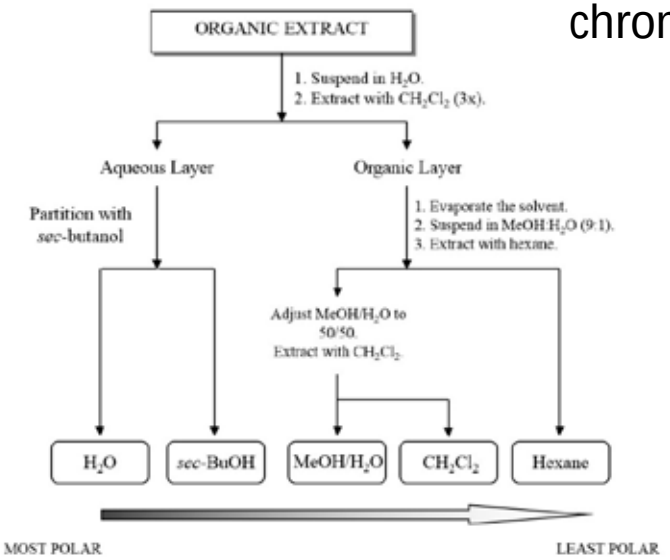
Compound Isolation

Solvent-solvent partition



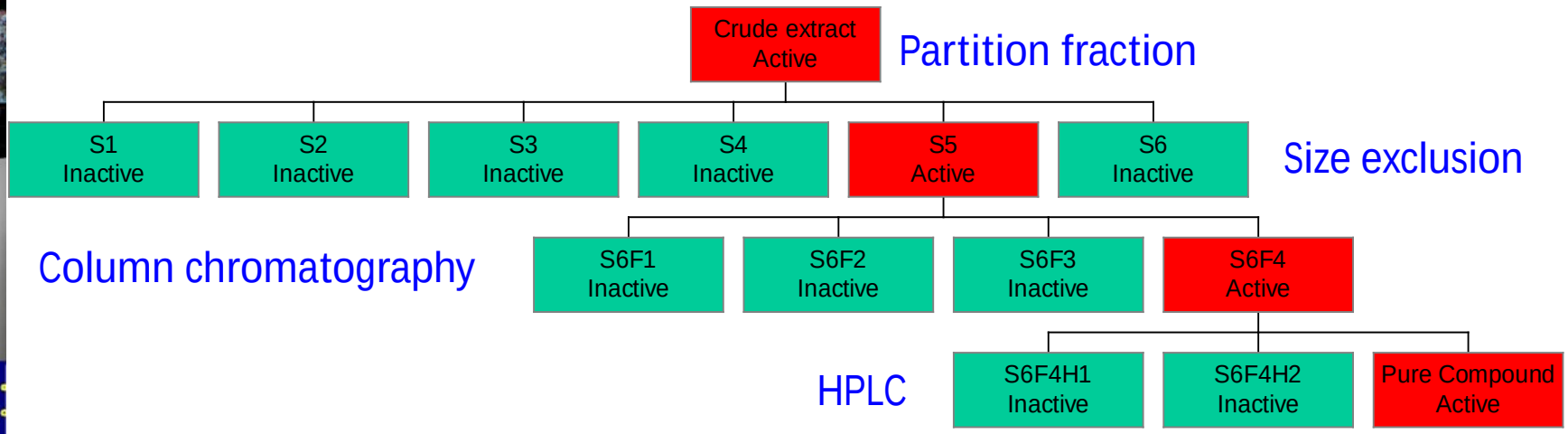
Size-exclusion chromatography

Medium pressure liquid chromatography



High Pressure Liquid Chromatography

Bioassay Guided Isolation



marine biodiscovery centre

Problems with bioassay guided isolation

UNIVERSITY OF ABERDEEN

- Poor assay reproducibility using crude/semi-purified extracts
- False positives/negatives
- Long times between cycles of purification/bioassay are common
- Repeated rediscovery of known compounds (bioactives)

marine biodiscovery centre

UNIVERSITY OF ABERDEEN

One Solution – Compound Libraries

Collection/Taxonomy
Curation/Informatics

Library generation

SCREEN (in-house/partner)

Crude extracts

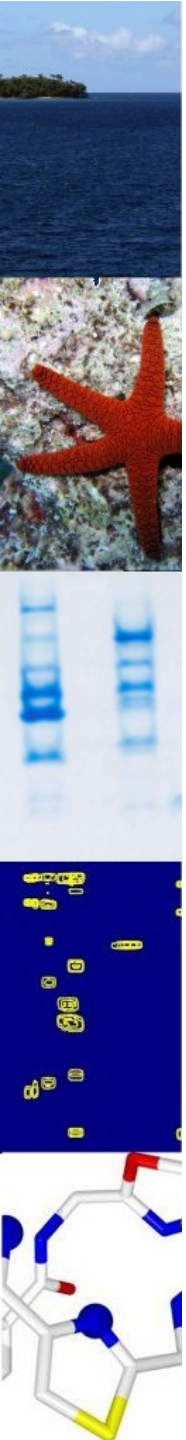
Purified extracts

Pure compounds

Pure compounds with structure

Library size

Library value

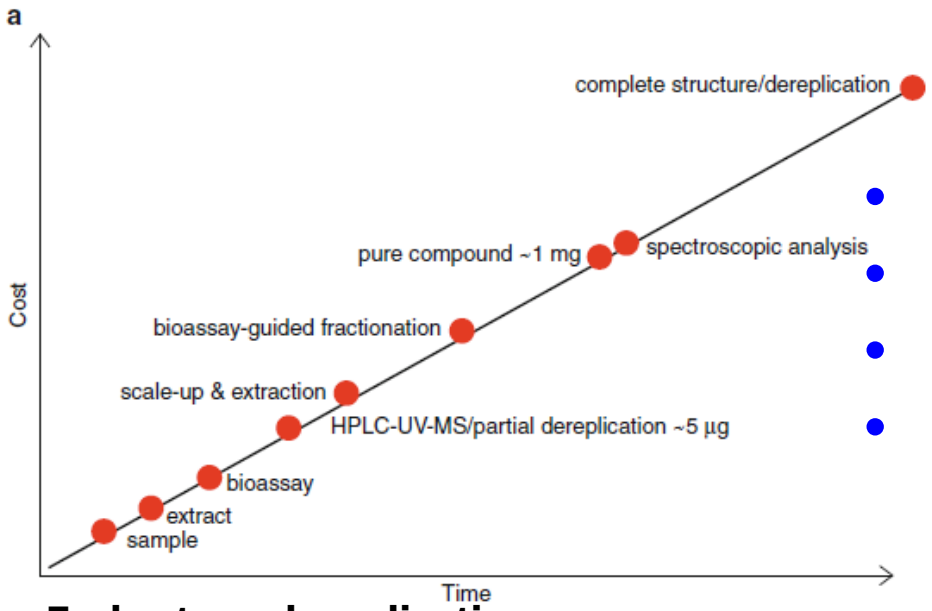


marine biodiscovery centre

Dereplication

Has your compound been reported before and how do you find out?

Isolation of known compounds is time consuming and costly



Available databases:

- 30,000 - MarinLit (RSC)
- 36,000 - Antibase
- 50,000 – AntiMarin
- 160,000 – Dictionary of natural products

Early stage dereplication

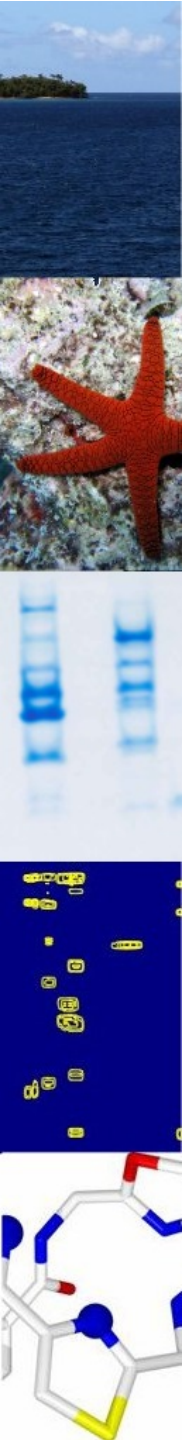
Identifying known compounds in extracts

Prioritising samples for further work

Late stage dereplication

After compounds have been purified.

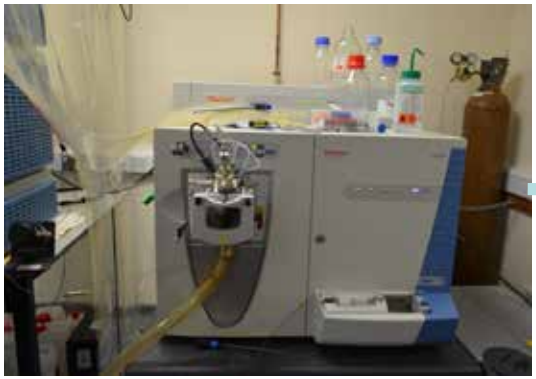
John Blunt et al. in Handbook of Marine Natural Products, 2012



marine biodiscovery centre

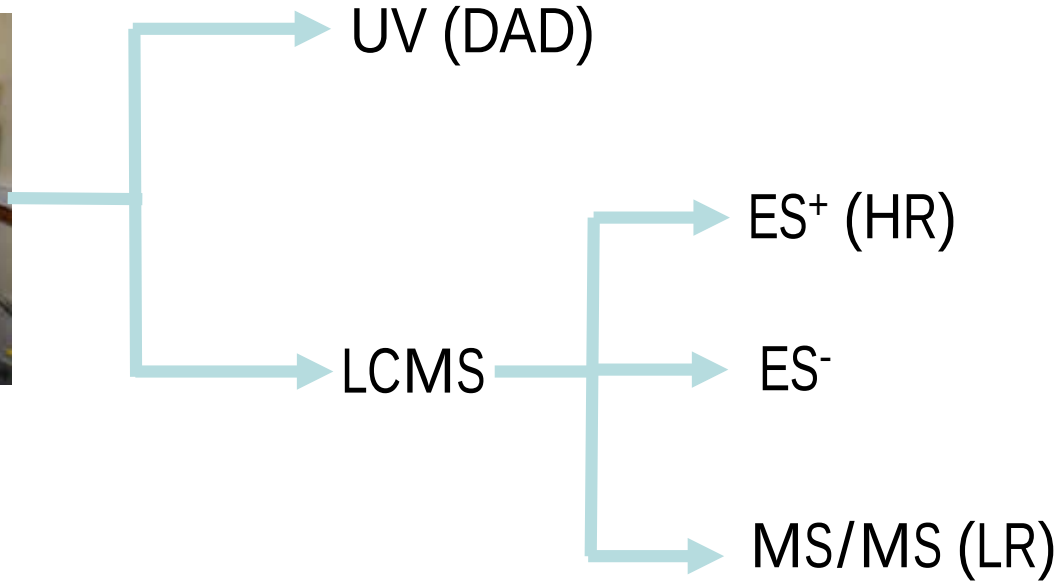
Hyphenated Data

UNIVERSITY OF ABERDEEN

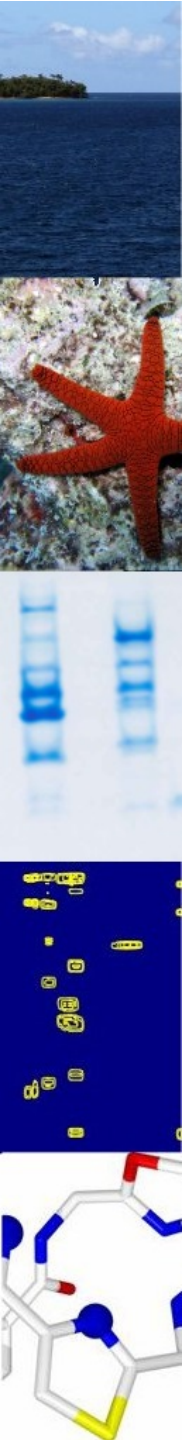


Sample submission

[Crude] = 0.5 mg/ml
[Pure] = 0.05 mg/ml



- Coupling of LC and MS increases resolution of both LC and MS methods
  - Can monitor selected ions
  - Can monitor selected reactions
- Using high resolution, molecular formula can be determined and searched in database.
- Useful for early dereplication – at extract stage

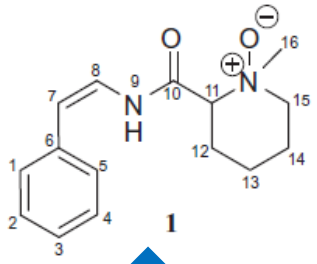


marine biodiscovery centre

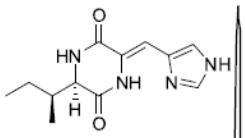
Targeted Dereplication

Searching for Every Database Compound in the LC-MS

UNIVERSITY OF ABERDEEN

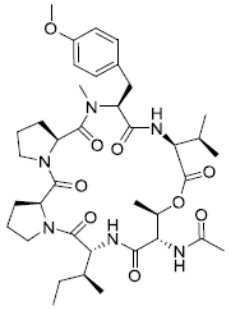


1

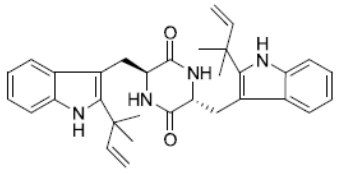


JBIR-74

Unidentified Peak



Aspergillicin A

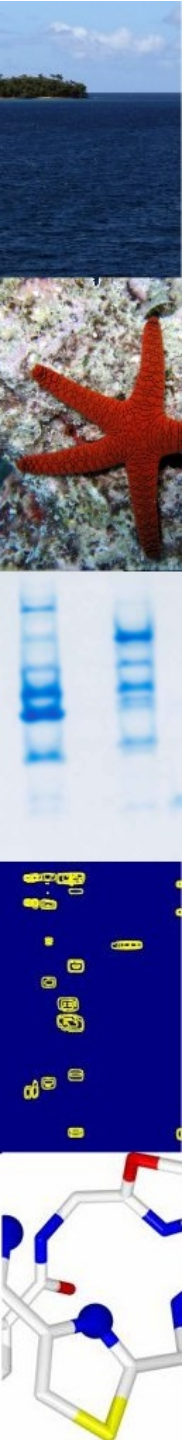


Fellutanine C

***Aspergillus indologenus* extract**

Use database of all known *Aspergillus* and *Penicillium* compounds and retention times to identify known peaks and locate unknowns

Tetrahedron Lett. 2015, **56** 1847

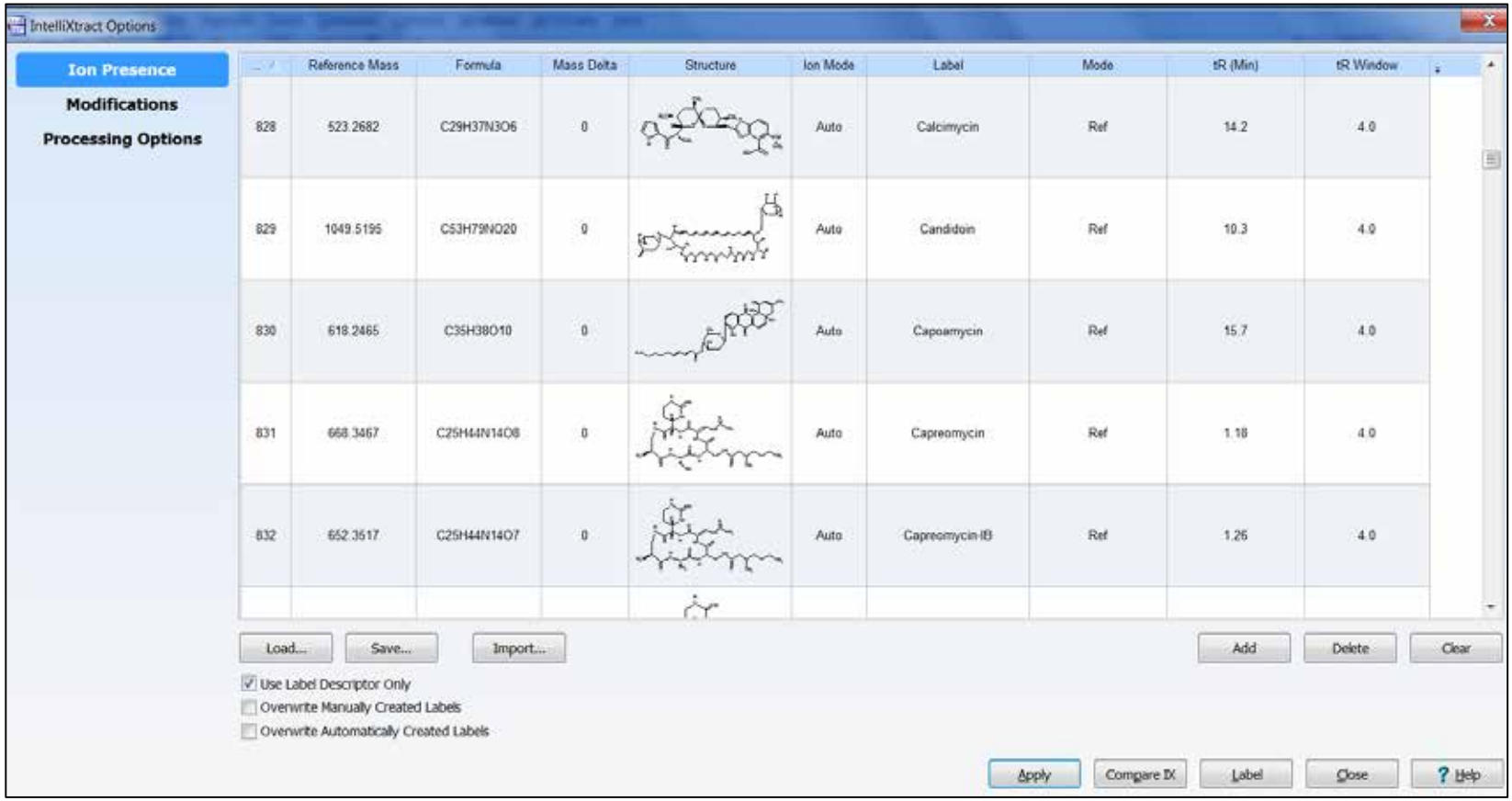


marine biodiscovery centre

MBC *Streptomyces* Database
5070 *Streptomyces* compounds

UNIVERSITY OF ABERDEEN

ACD/Labs



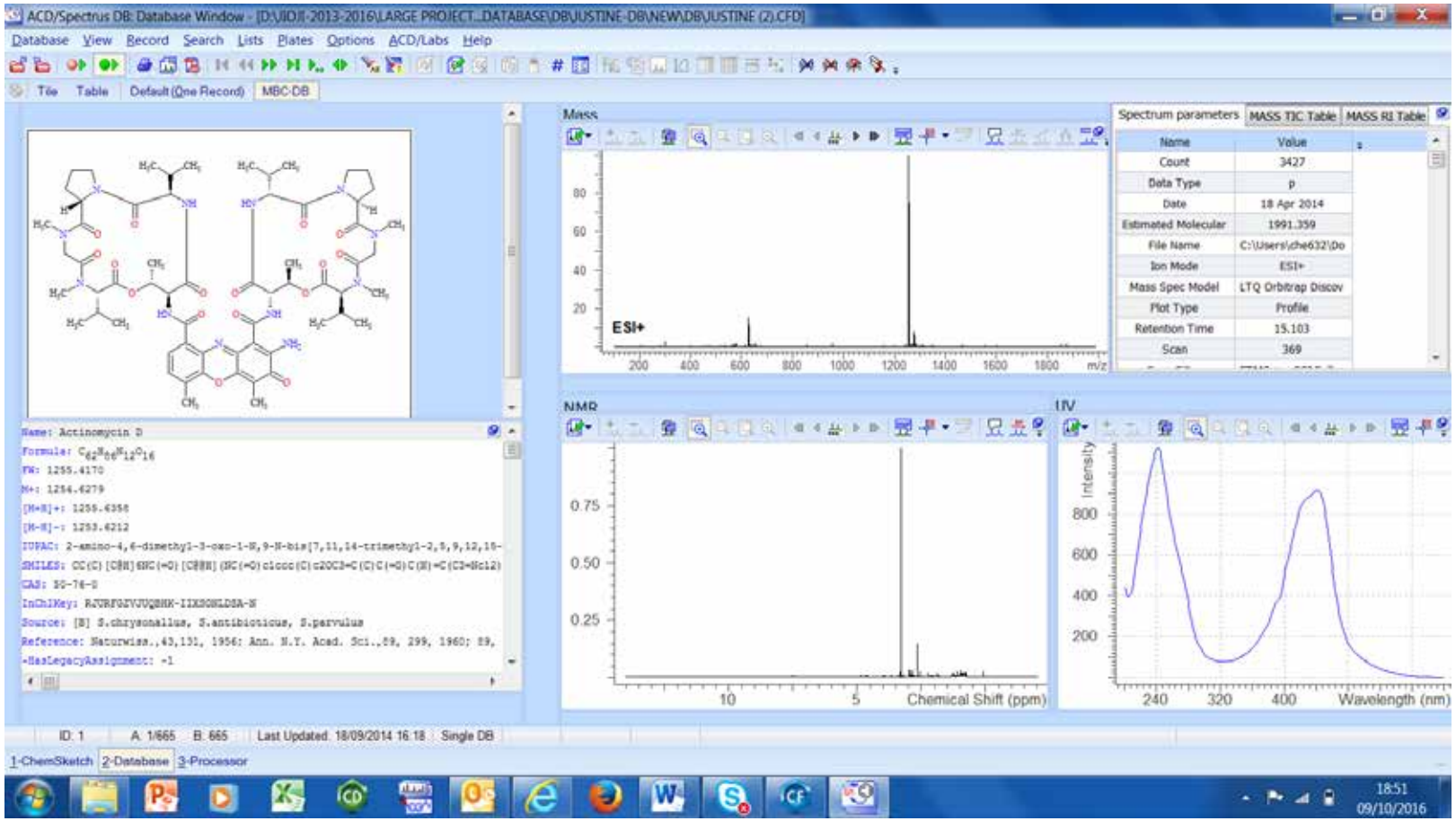
Jioji Tabudravu

marine biodiscovery centre

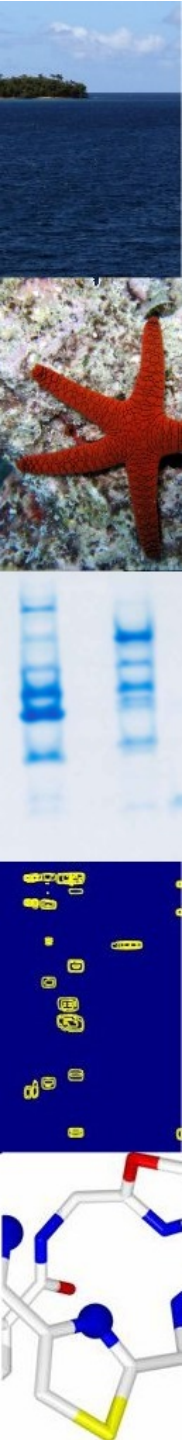
MBC Spectroscopic Database

UNIVERSITY OF ABERDEEN

ACD/Labs



665 natural products from marine and terrestrial sources representing several classes of natural products including peptides, alkaloids, terpenes and others



marine biodiscovery centre

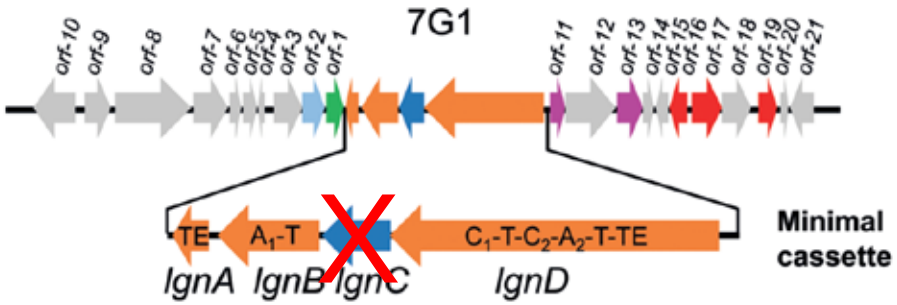
UNIVERSITY OF ABERDEEN

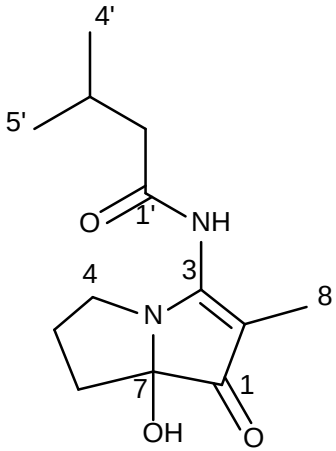
Case Study

Dereplication of *Streptomyces albus*, Δ*lgnC*

Tamarindus indica, Legon, Ghana

Streptomyces albus





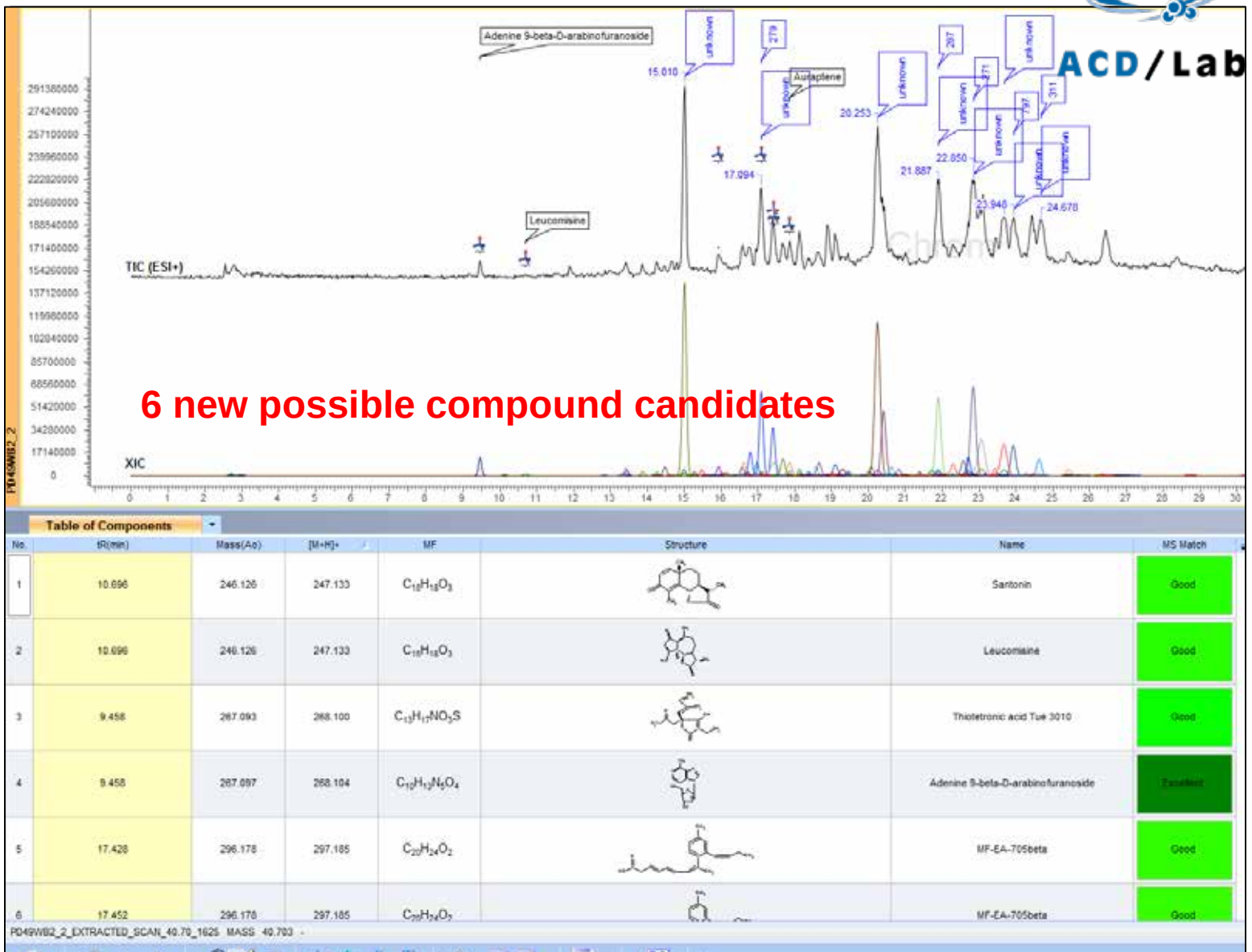
Angew. Chem. Int. Ed. 2015, **54**, 12697

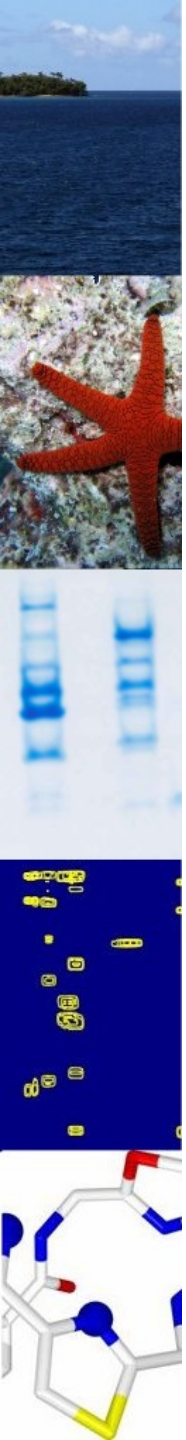
Legonmycin A

marine biodiscovery centre

Knowns and Unknowns

UNIVERSITY OF ABERDEEN

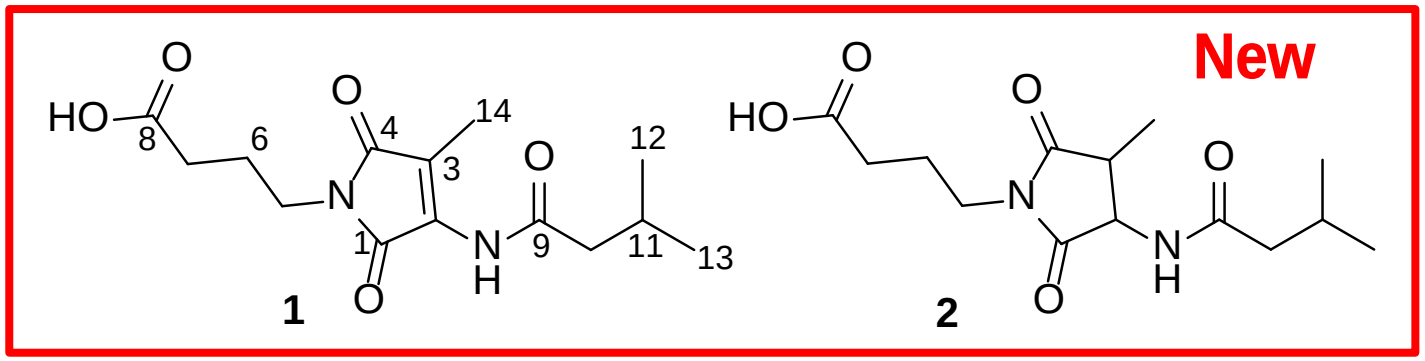


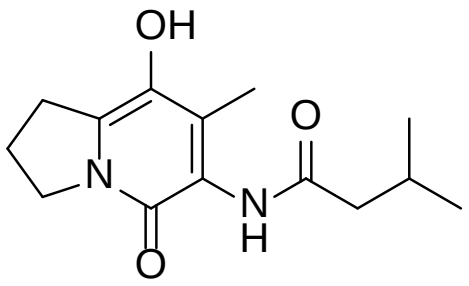


marine biodiscovery centre

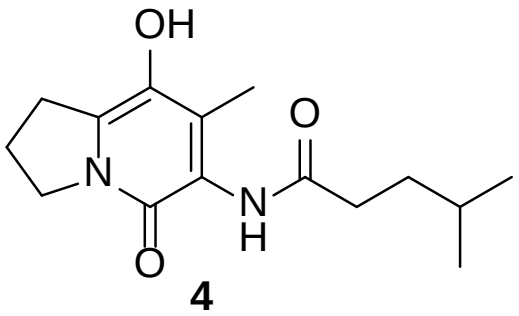
UNIVERSITY OF ABERDEEN

New Compounds

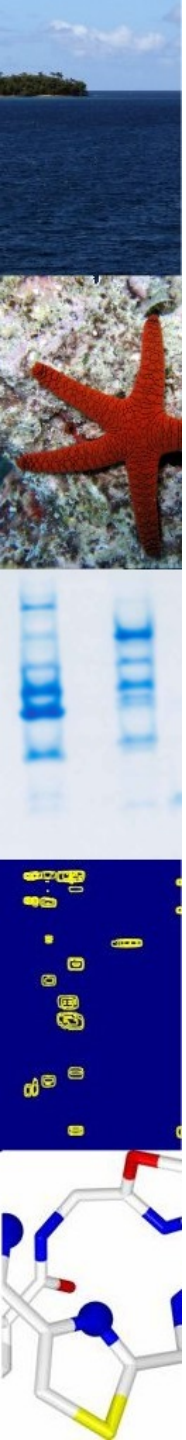




Known



Known

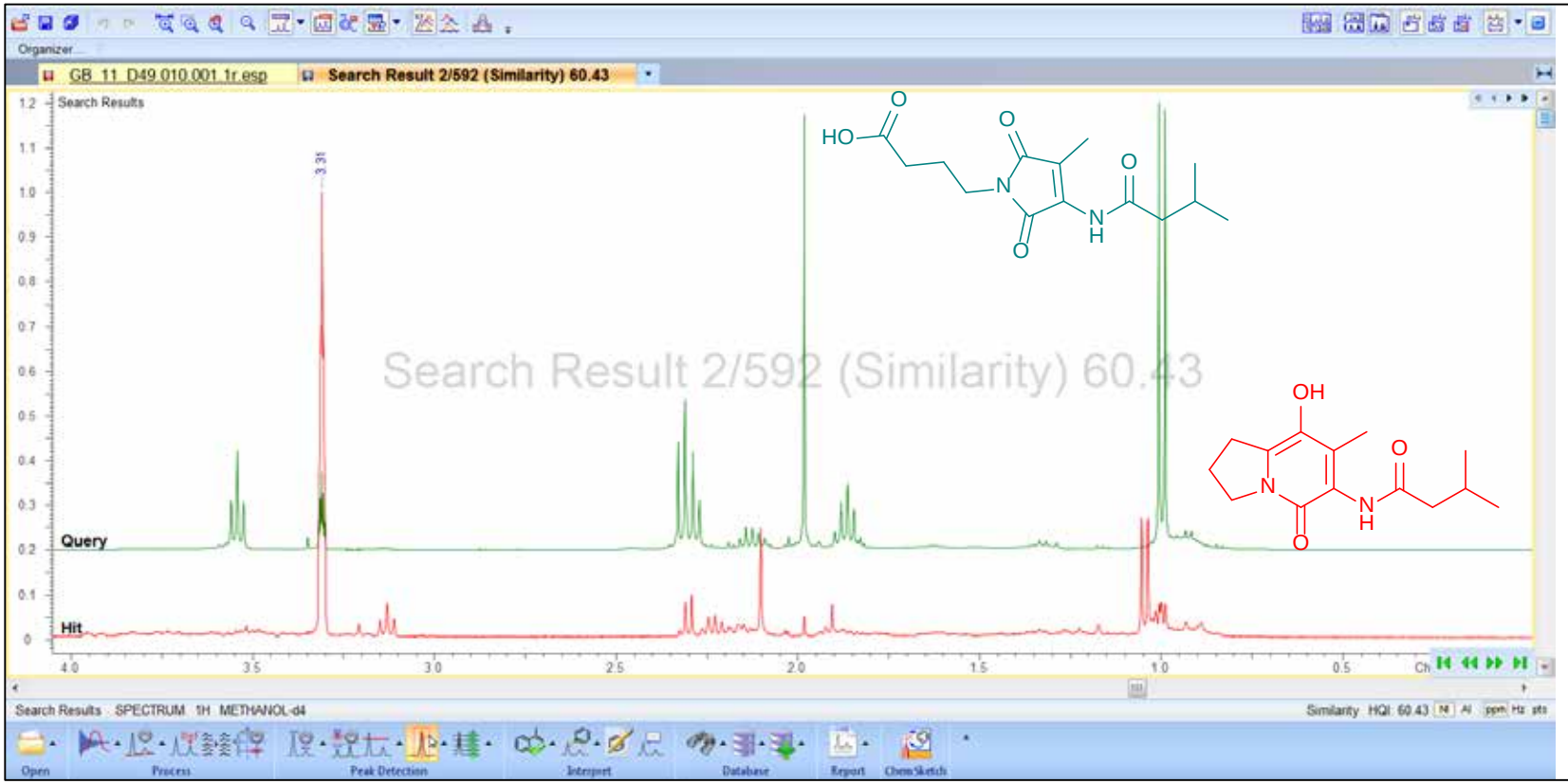


marine biodiscovery centre

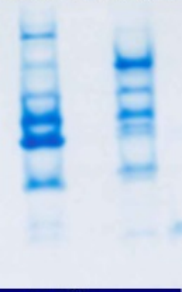
MBC Spectroscopic Database Similarity Match

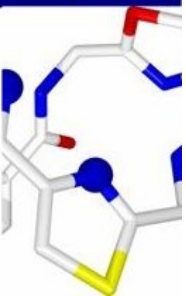
UNIVERSITY OF ABERDEEN

ACD/Labs





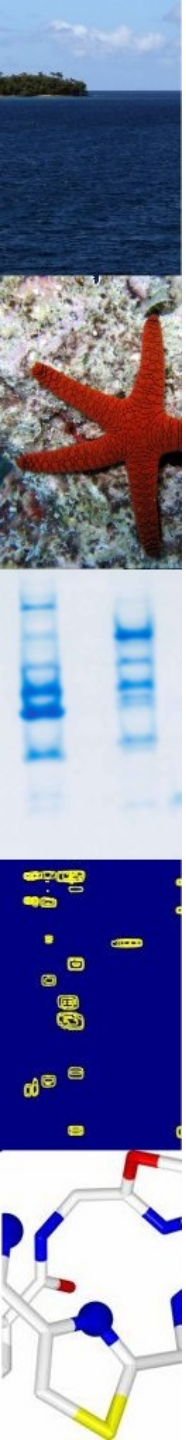




marine biodiscovery centre

UNIVERSITY OF ABERDEEN

Ross Sea *Pseudomonas* sp. BTN1



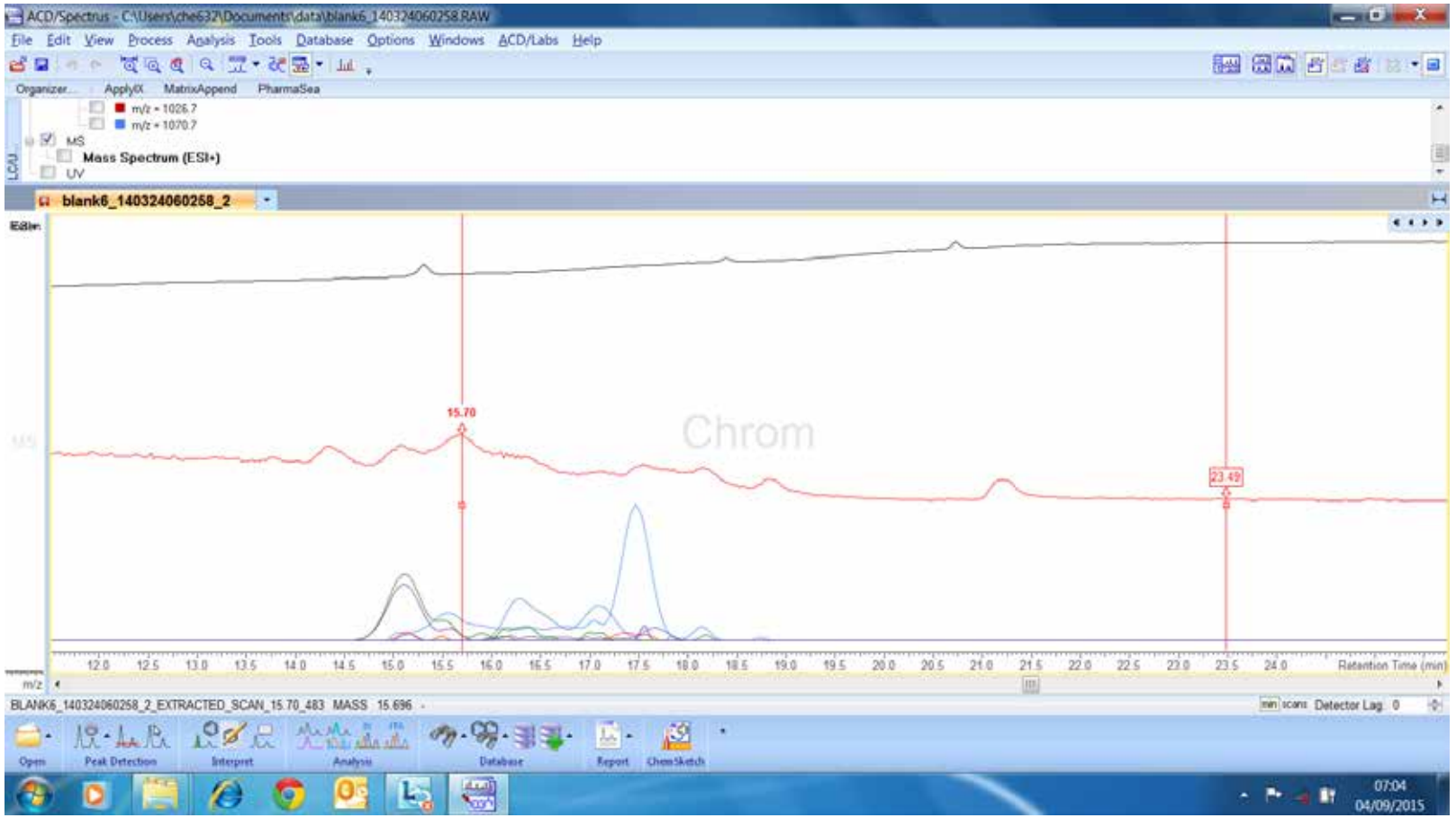
marine biodiscovery centre

Untargeted Dereplication

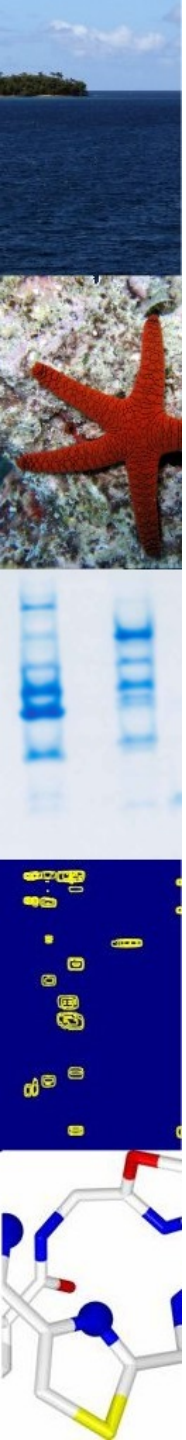
UNIVERSITY OF ABERDEEN

Searching Every LC-MS peak in the Database

Analysis of the Blank



Jioji Tabudravu

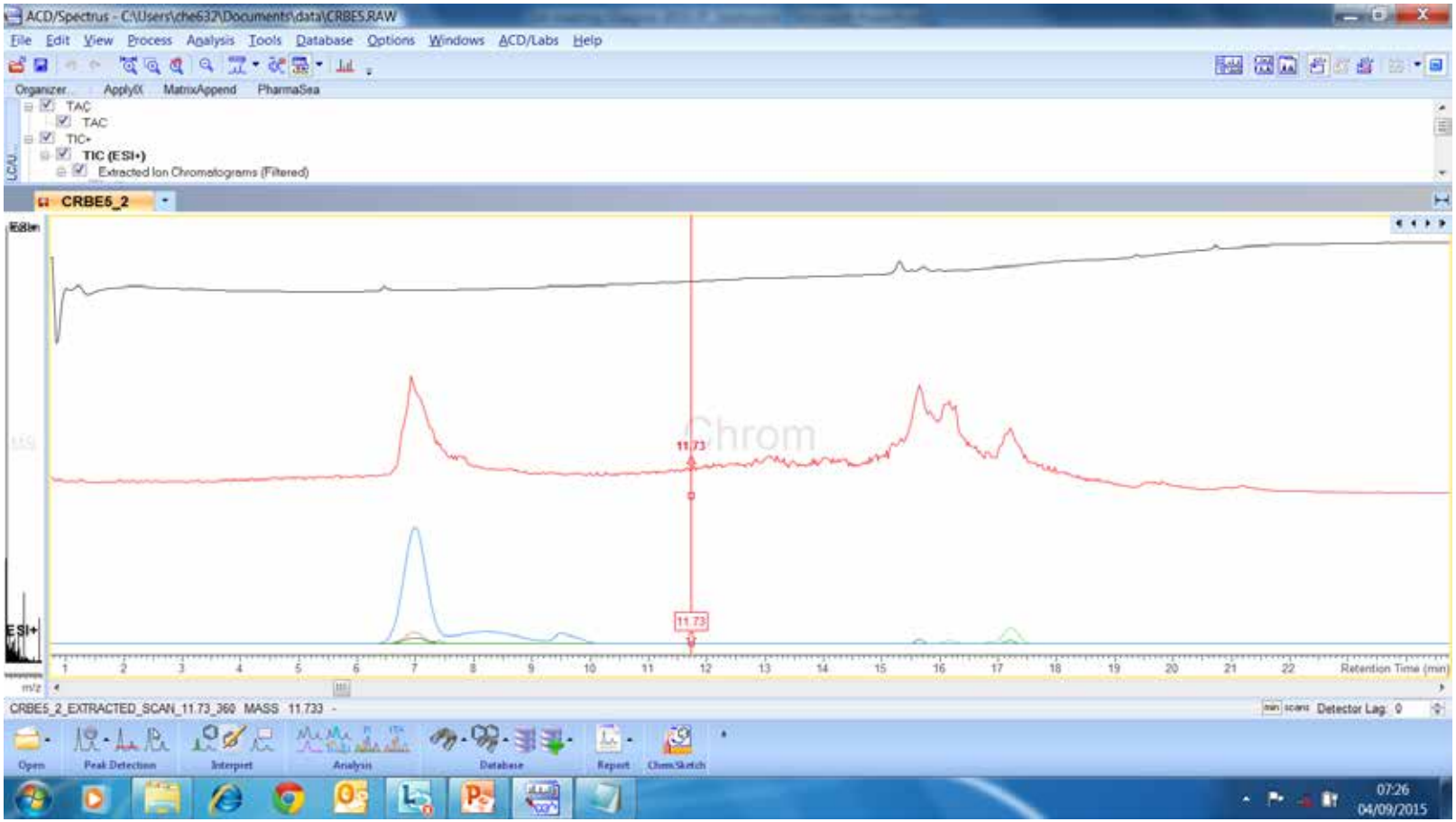


marine biodiscovery centre

Untargeted Dereplication

Blank and Medium Removed

UNIVERSITY OF ABERDEEN

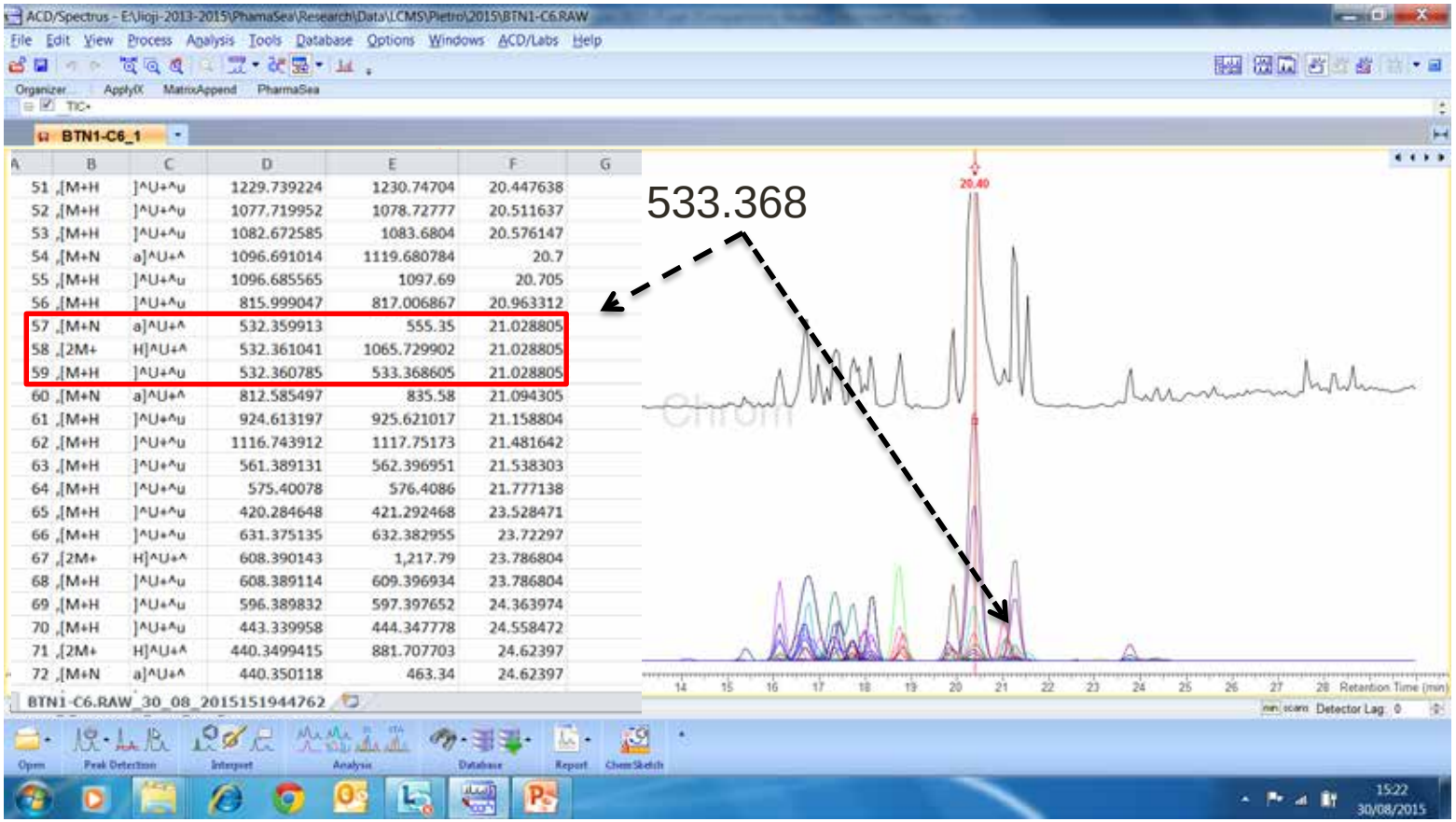


marine biodiscovery centre

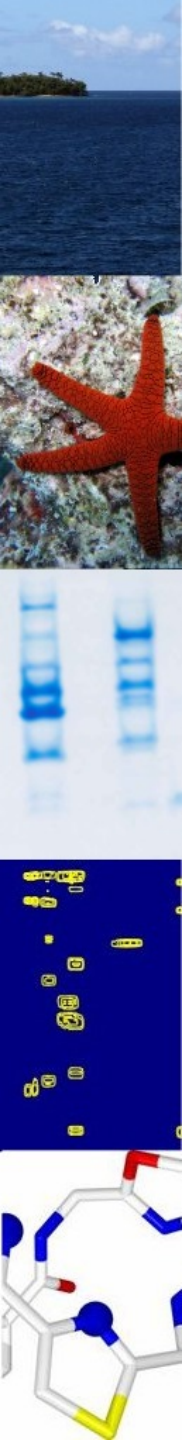
Untargeted Dereplication

UNIVERSITY OF ABERDEEN

Blank, Medium and PharmaSea Database (28,000 compounds) removed



	A	B	C	D	E	F	G
51		[M+H	]^U+^u	1229.739224	1230.74704	20.447638	
52		[M+H	]^U+^u	1077.719952	1078.72777	20.511637	
53		[M+H	]^U+^u	1082.672585	1083.6804	20.576147	
54		[M+N	a]^U+^	1096.691014	1119.680784	20.7	
55		[M+H	]^U+^u	1096.685565	1097.69	20.705	
56		[M+H	]^U+^u	815.999047	817.006867	20.963312	
57		[M+N	a]^U+^	532.359913	555.35	21.028805	
58		[2M+	H]^U+^	532.361041	1065.729902	21.028805	
59		[M+H	]^U+^u	532.360785	533.368605	21.028805	
60		[M+N	a]^U+^	812.585497	835.58	21.094305	
61		[M+H	]^U+^u	924.613197	925.621017	21.158804	
62		[M+H	]^U+^u	1116.743912	1117.75173	21.481642	
63		[M+H	]^U+^u	561.389131	562.396951	21.538303	
64		[M+H	]^U+^u	575.40078	576.4086	21.777138	
65		[M+H	]^U+^u	420.284648	421.292468	23.528471	
66		[M+H	]^U+^u	631.375135	632.382955	23.72297	
67		[2M+	H]^U+^	608.390143	1,217.79	23.786804	
68		[M+H	]^U+^u	608.389114	609.396934	23.786804	
69		[M+H	]^U+^u	596.389832	597.397652	24.363974	
70		[M+H	]^U+^u	443.339958	444.347778	24.558472	
71		[2M+	H]^U+^	440.3499415	881.707703	24.62397	
72		[M+N	a]^U+^	440.350118	463.34	24.62397	

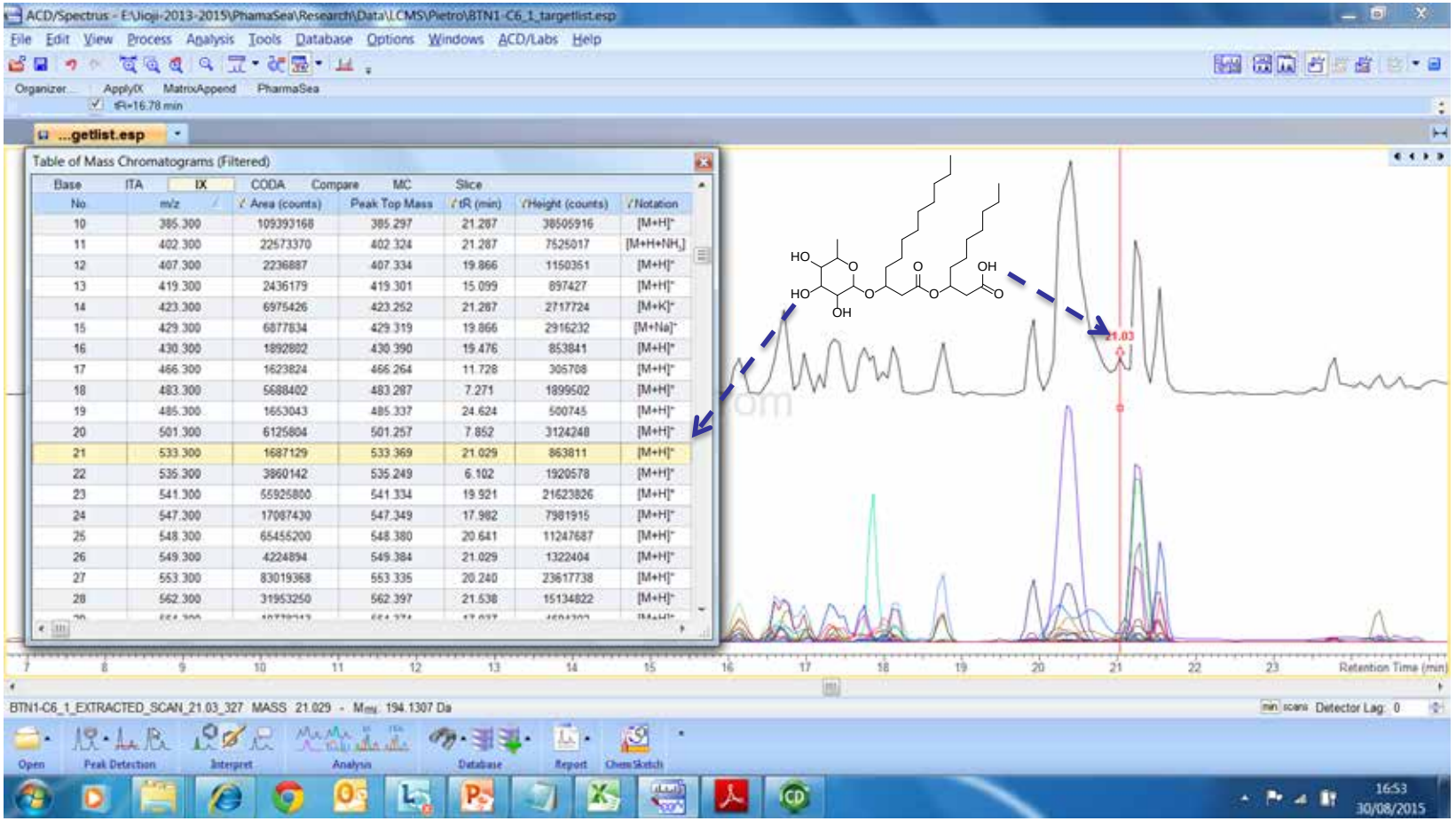


marine biodiscovery centre

Untargeted Dereplication

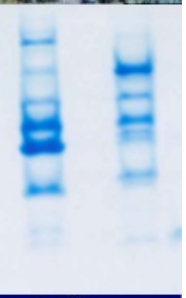
UNIVERSITY OF ABERDEEN

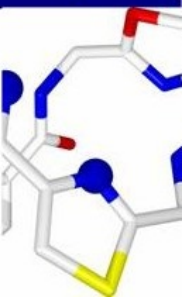
After Removal of Blank/Medium/Knowns, Remainder are New Compounds



Marine Drugs, 2016, **14**, 83



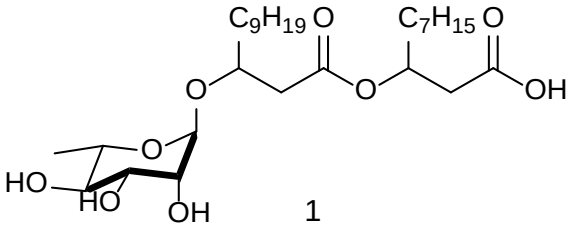




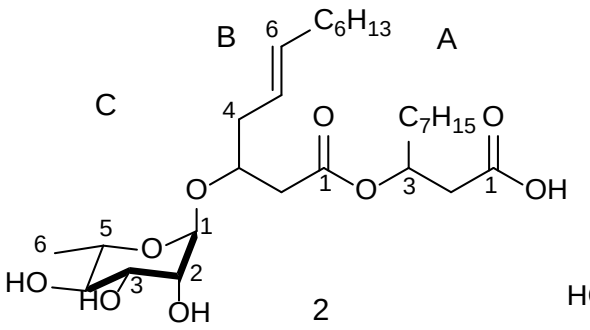
marine biodiscovery centre

Monorhamnolipids

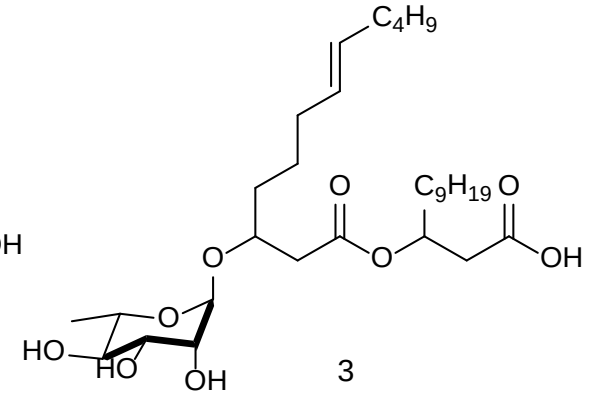
Activity Against *Burkholderia*



1



2

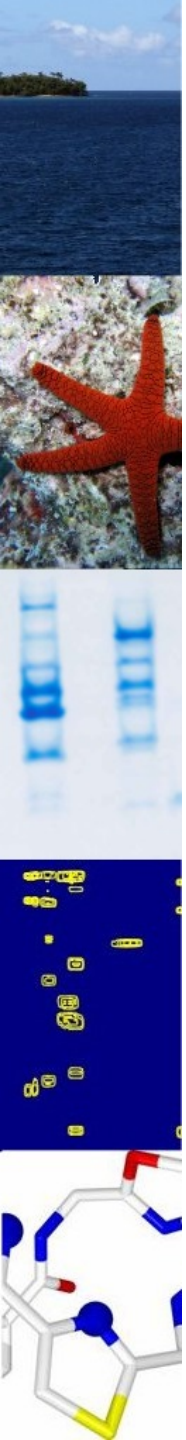


3

Antimicrobial Activity (μg/mL)

	B. cenocepacia LMG 16656		*B. metallica* LMG 24068		*B. seminalis* LMG 24067		*B. diffusa* LMG 24065		*B. latens* LMG 24064		*S. aureus* 6538P	
	MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC
C1	3.12	3.12	50	50	12.5	12.5	>200	>200	12.5	12.5	1.56	1.56
C2	3.12	3.12	25	25	3.12	3.12	200	200	12.5	12.5	3.12	3.12
C3	200	200	>200	>200	>200	>200	>200	>200	>200	>200	100	100

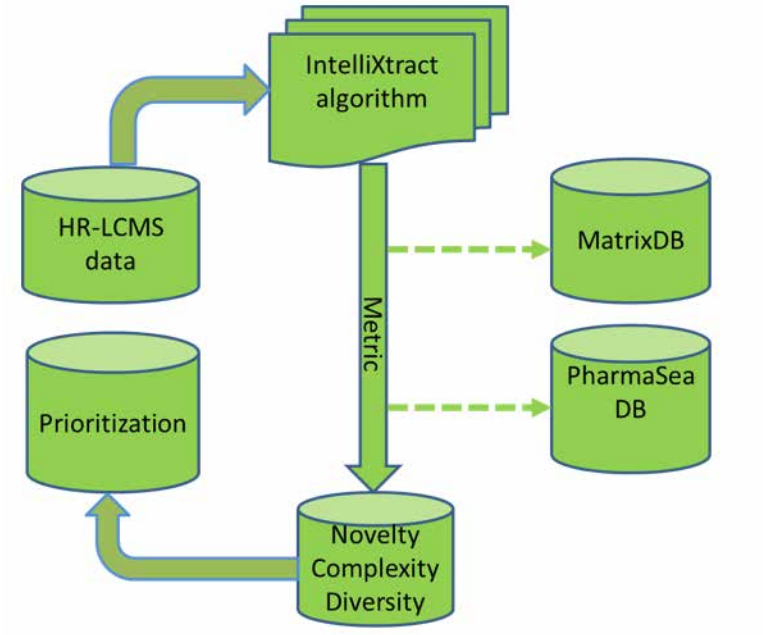
Mar. Drugs **2016**, *14*, 83



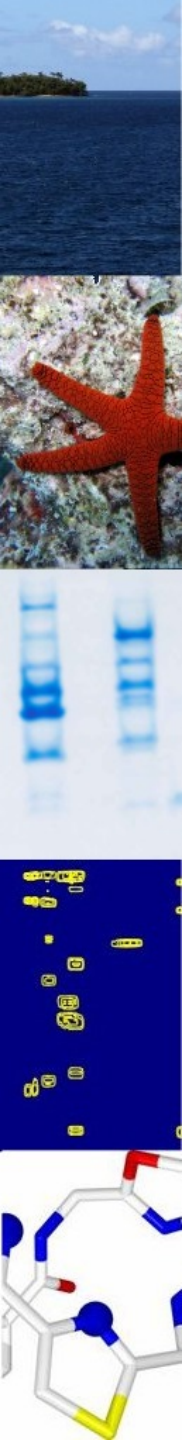
marine biodiscovery centre

Novelty, Complexity and Diversity Metrics

UNIVERSITY OF ABERDEEN



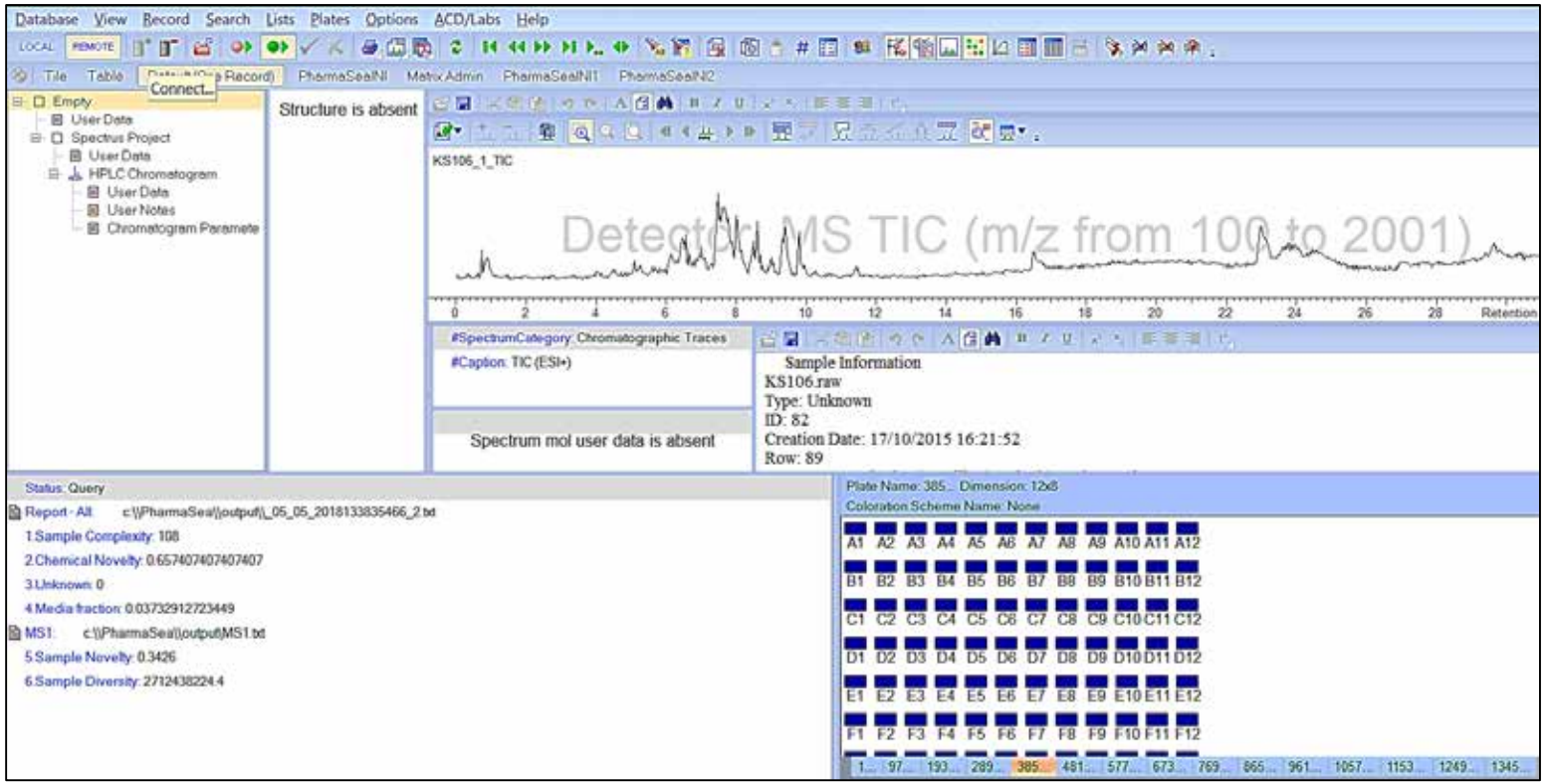
Jioji Tabudravu

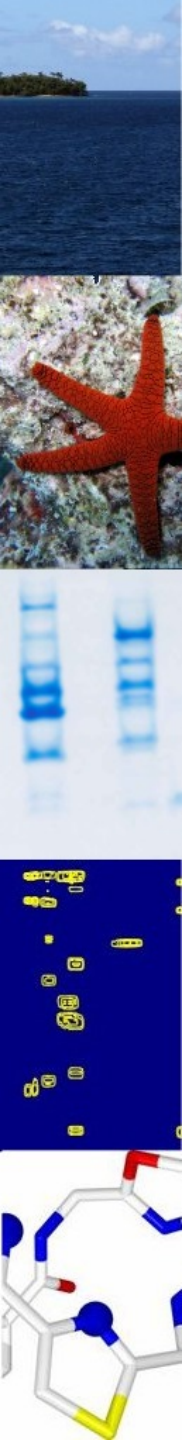


marine biodiscovery centre

Novelty, Complexity and Diversity Metrics

UNIVERSITY OF ABERDEEN

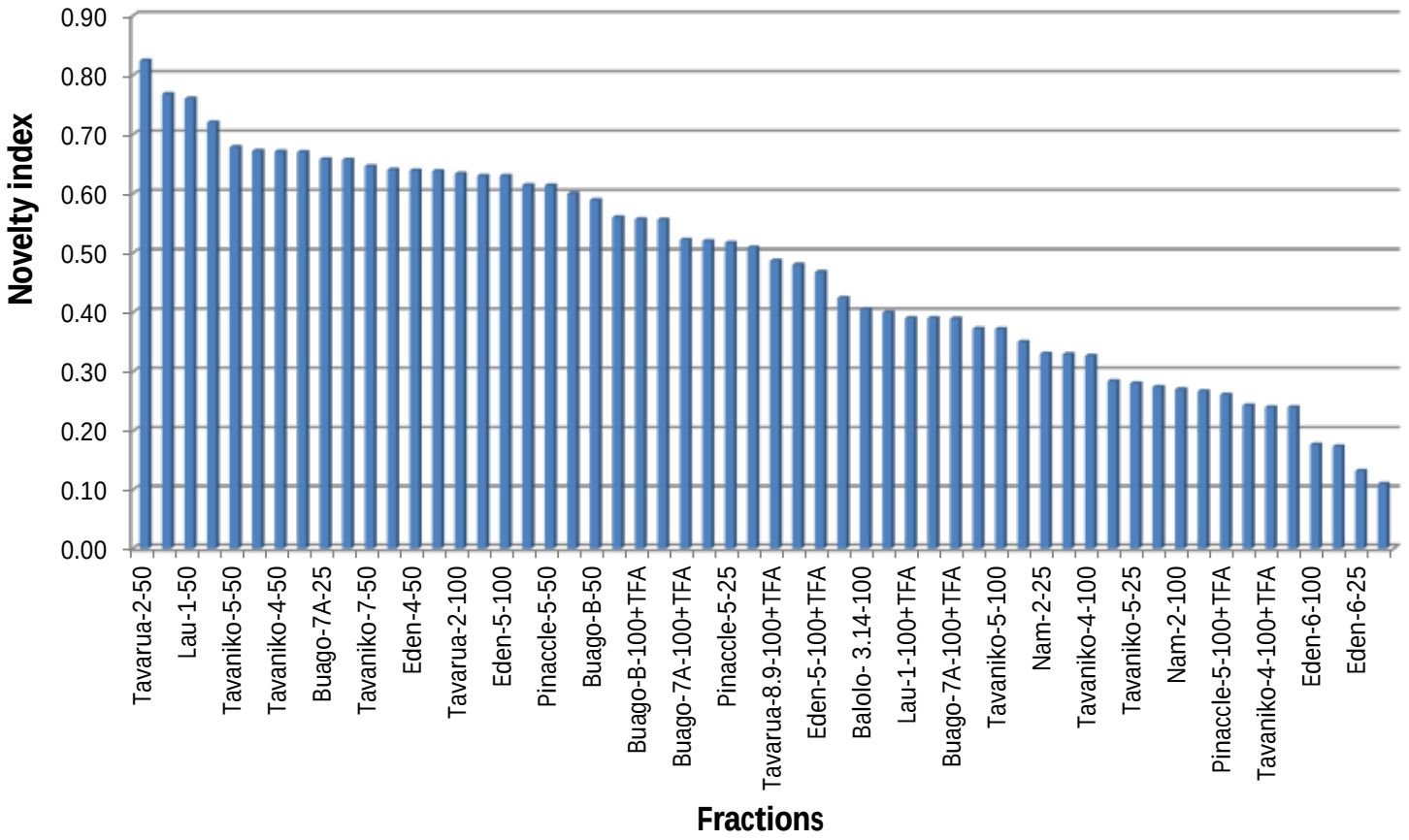




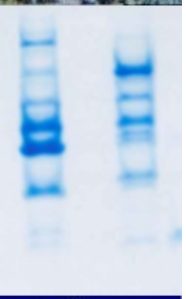
marine biodiscovery centre

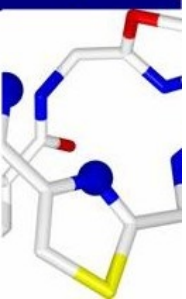
Novelty for Fijian Invertebrate Extracts

UNIVERSITY OF ABERDEEN









marine biodiscovery centre

UNIVERSITY OF ABERDEEN

New Compound Discovered

No	Sample code	SPE fract	Sample complexity	Chemical novelty	Sample Media c	Sample novelty	Sample diversity
1	Tavarua-2	50	74	0.82	3.30E-02	0.17	2.52E+09
2	Tavarua-8.9	100	69	0.77	2.87E-02	0.23	2.53E+09
3	Lau-1	50	79	0.76	3.20E-02	0.24	9.08E+09

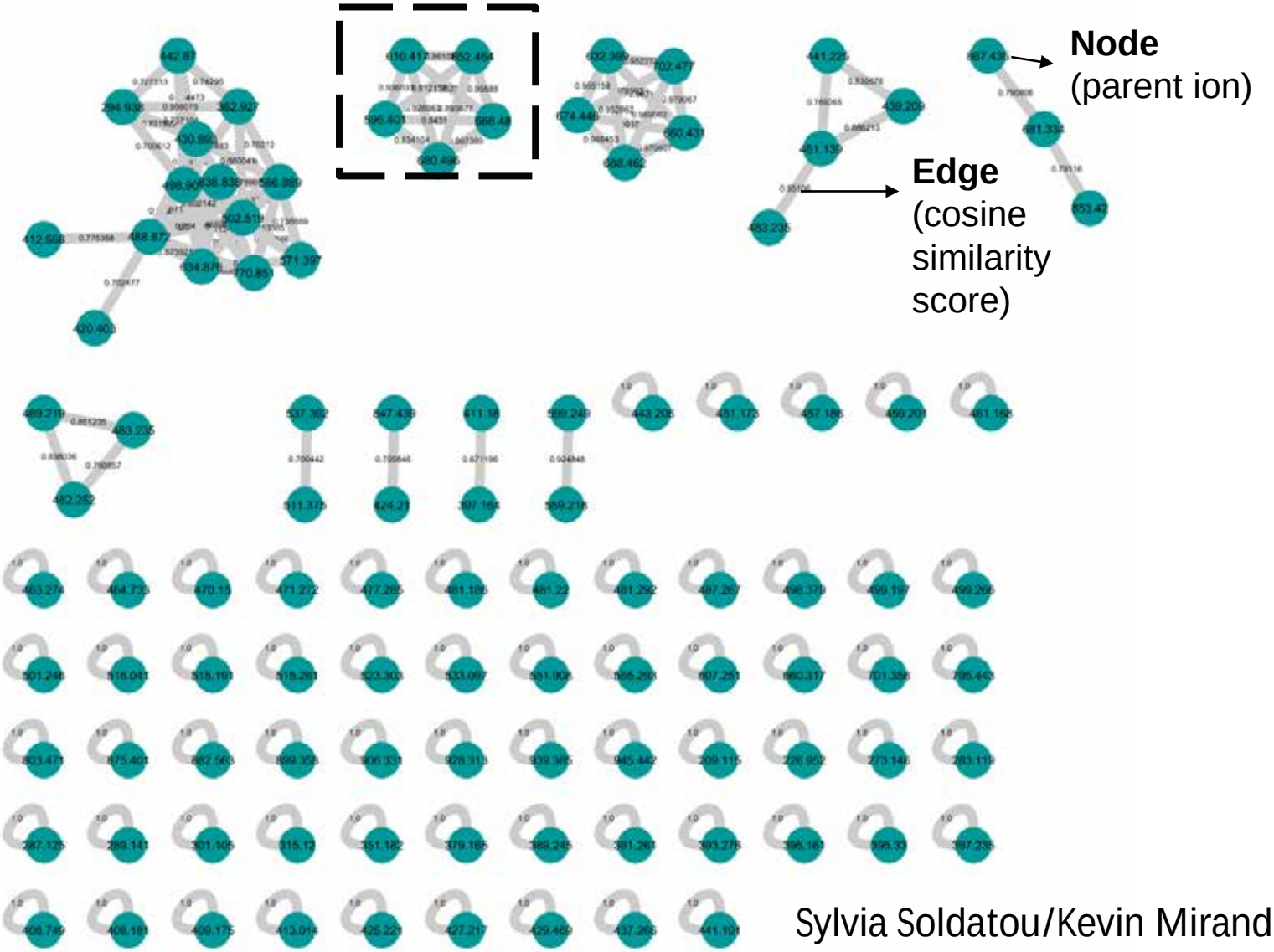
- Analysis MS/MS data is time consuming
- No publicly available dereplication databases

GNPS

- Compares fragmentation patterns
- Clusters compounds with similar fragmentation patterns based on similarity
- Dereplication by comparing newly generated MS/MS data with deposited data

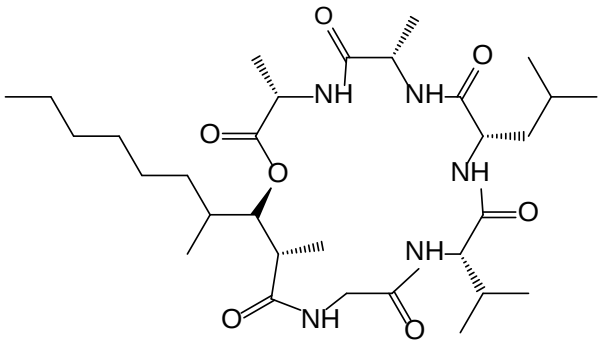
Sylvia Soldatou/Kevin Miranda

Molecular Network of Marine Endophytic ***Aspergillus*** sp. Extract



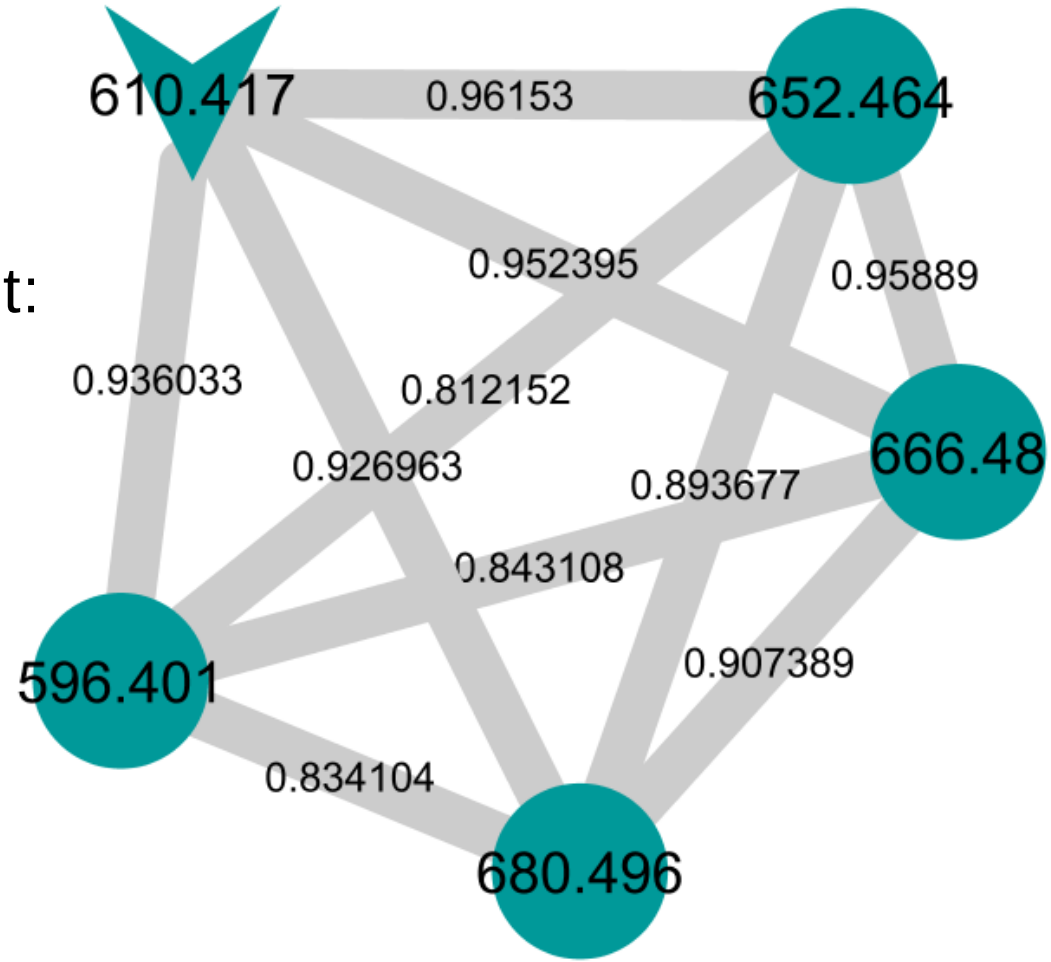
Sylvia Soldatou/Kevin Miranda

Molecular Network of Marine Endophytic ***Aspergillus*** sp. Extract



GNPS spectral library hit:
Emericellamide A
Confirmed by MS-MS Fragmentation

Circles: potentially new analogues

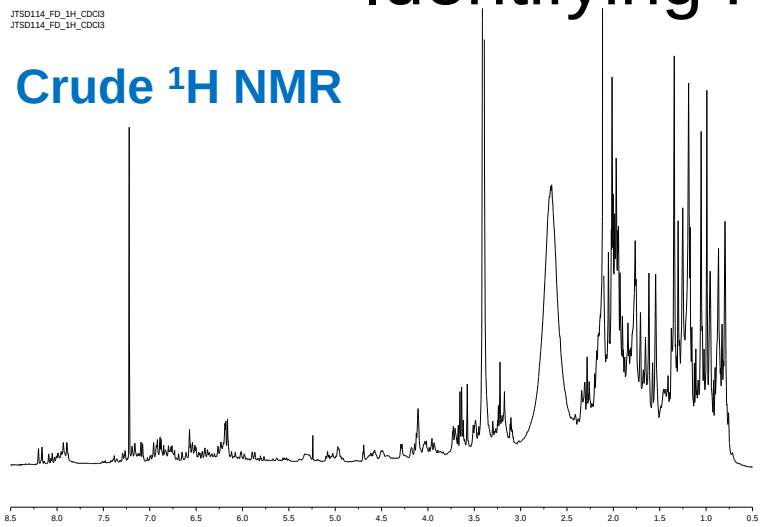


Sylvia Soldatou/Kevin Miranda

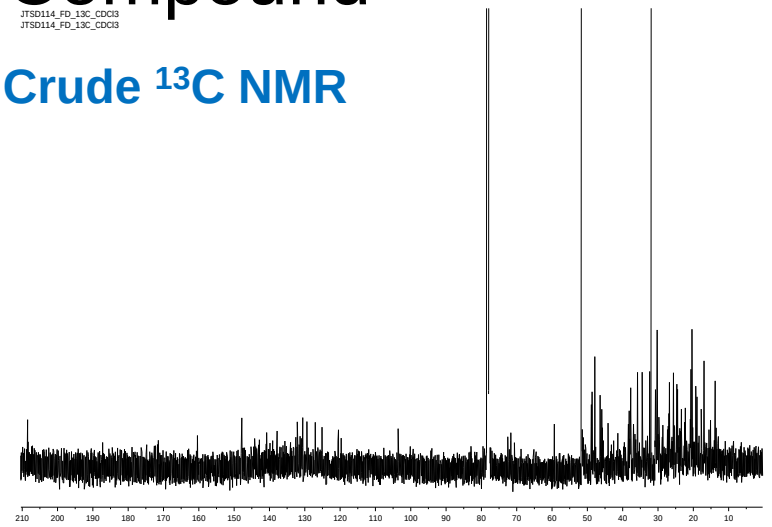
Late Stage Dereplication

From Prioritising Crude Extract to Identifying Pure Compound

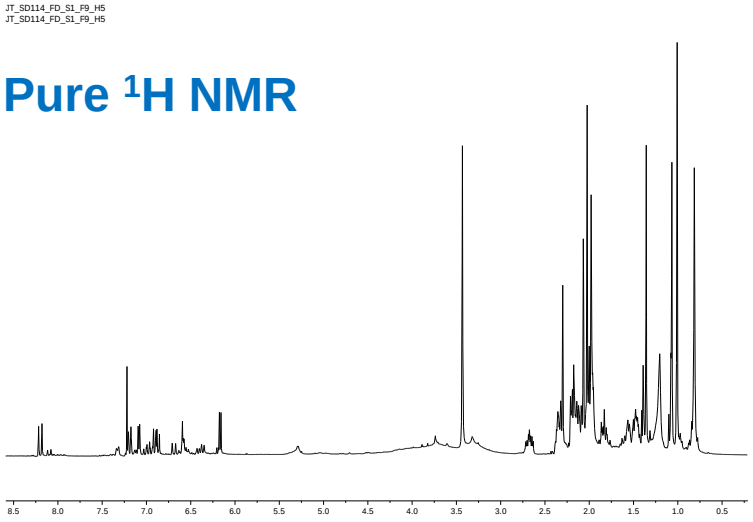
Crude ^{1}H NMR



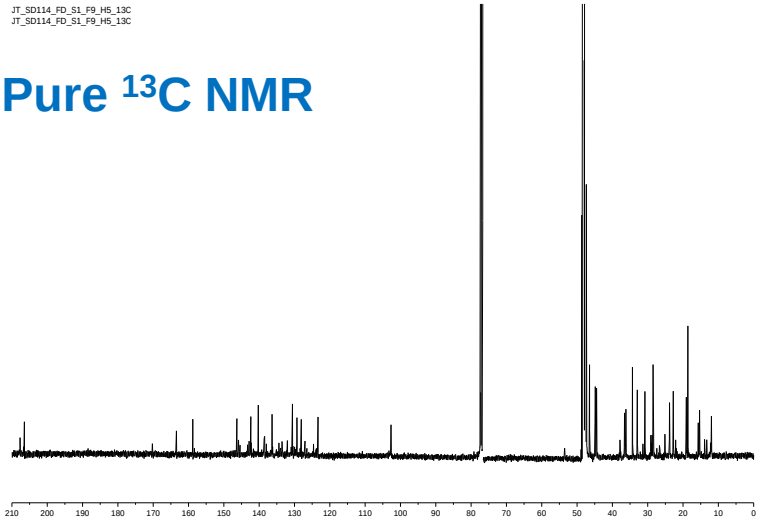
Crude ^{13}C NMR



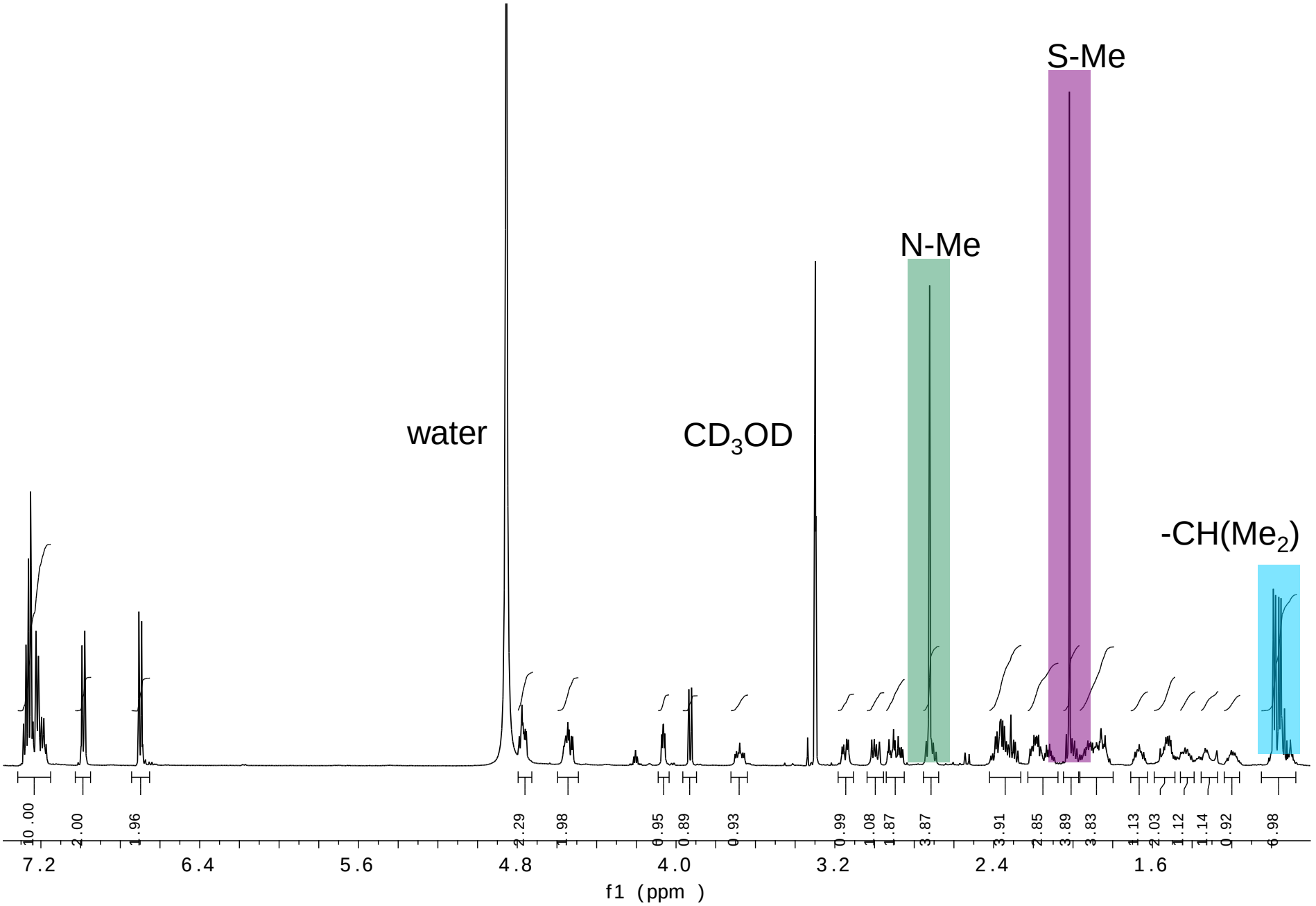
Pure ^{1}H NMR



Pure ^{13}C NMR

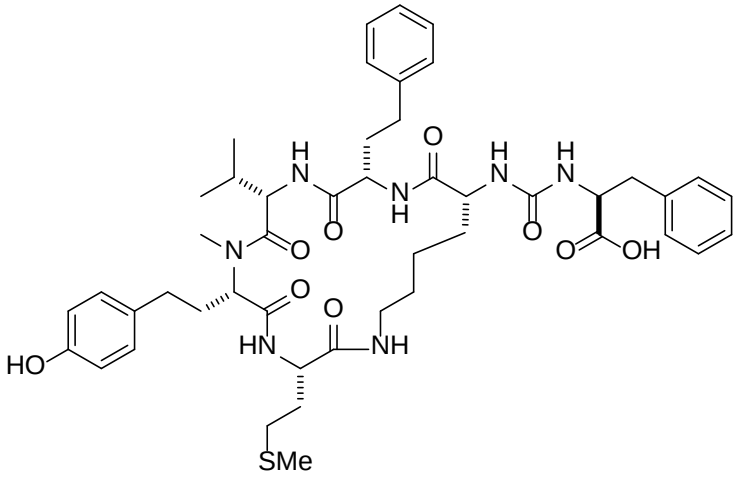


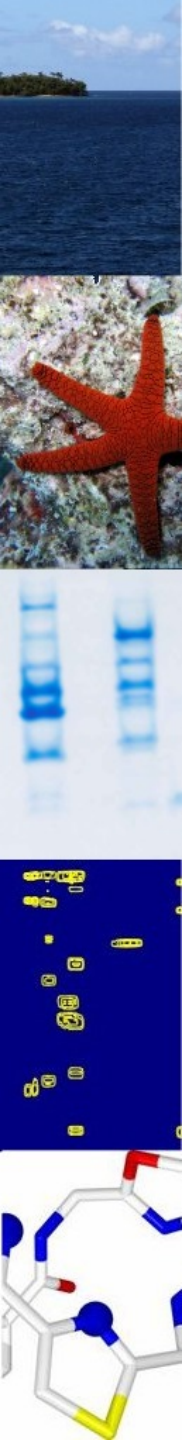
Using Simple Features




PHARMASEA
Simple search
Structure search
Advanced search
Structure
Monoisotopic Mass
Functional groups
Clear
Functional group search guidelines
Insert number or range (e.g. 1-3) of each type of functional group. To exclude a functional group, set the number to 0. Fields left blank will be included in matches. See Help for further details.
Carbon skeleton
sp3 carbons
Total methyl groups (includes methoxyl, N-methyl, S-methyl and acetyl)
Singlet methyl (R_4C-CH_3)
Doublet methyl (R_2HC-CH_3)
2
Triplet methyl (RH_2C-CH_3)
Aromatic methyl
Vinyl methyl ($C=CH-CH_3$)
Acetyl (including acetyl ester and acetyl amide)
Methoxy ($-OCH_3$)
N-methyl ($-NCH_3$)
1
S-methyl ($-SCH_3$)
1


PHARMASEA
Results 2
Structure Molecular Formula Molecular Weight Monoisotopic Mass
69493
902.110
901.440796
69497
918.109
917.435730
Displaying 1 to 2 of 2 items


HO
HN
NH
SMe
OH



marine biodiscovery centre

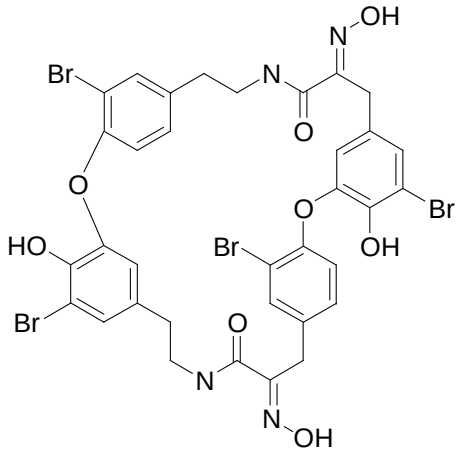
Chemical Structure

2D Framework

1 Carbon skeleton

2 Functional groups

3 Location of functional groups



OK

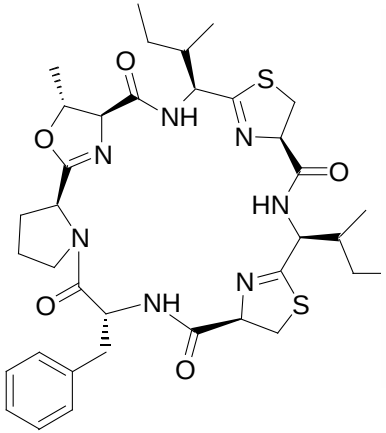
But what about:

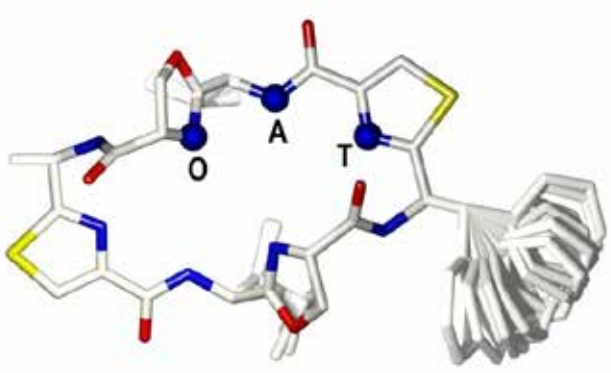
3D Framework

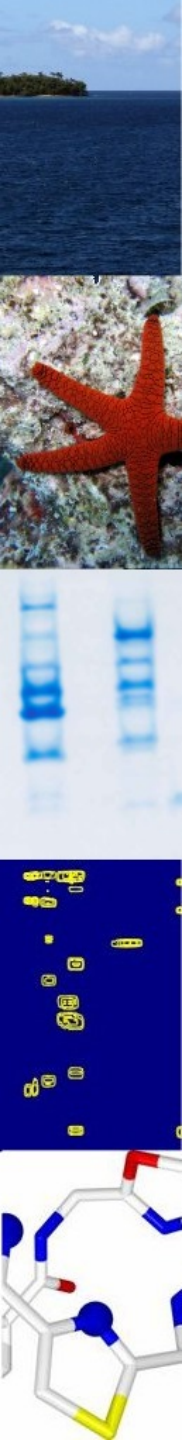
1 Relative stereo

2 Absolute stereo

3 Conformational features



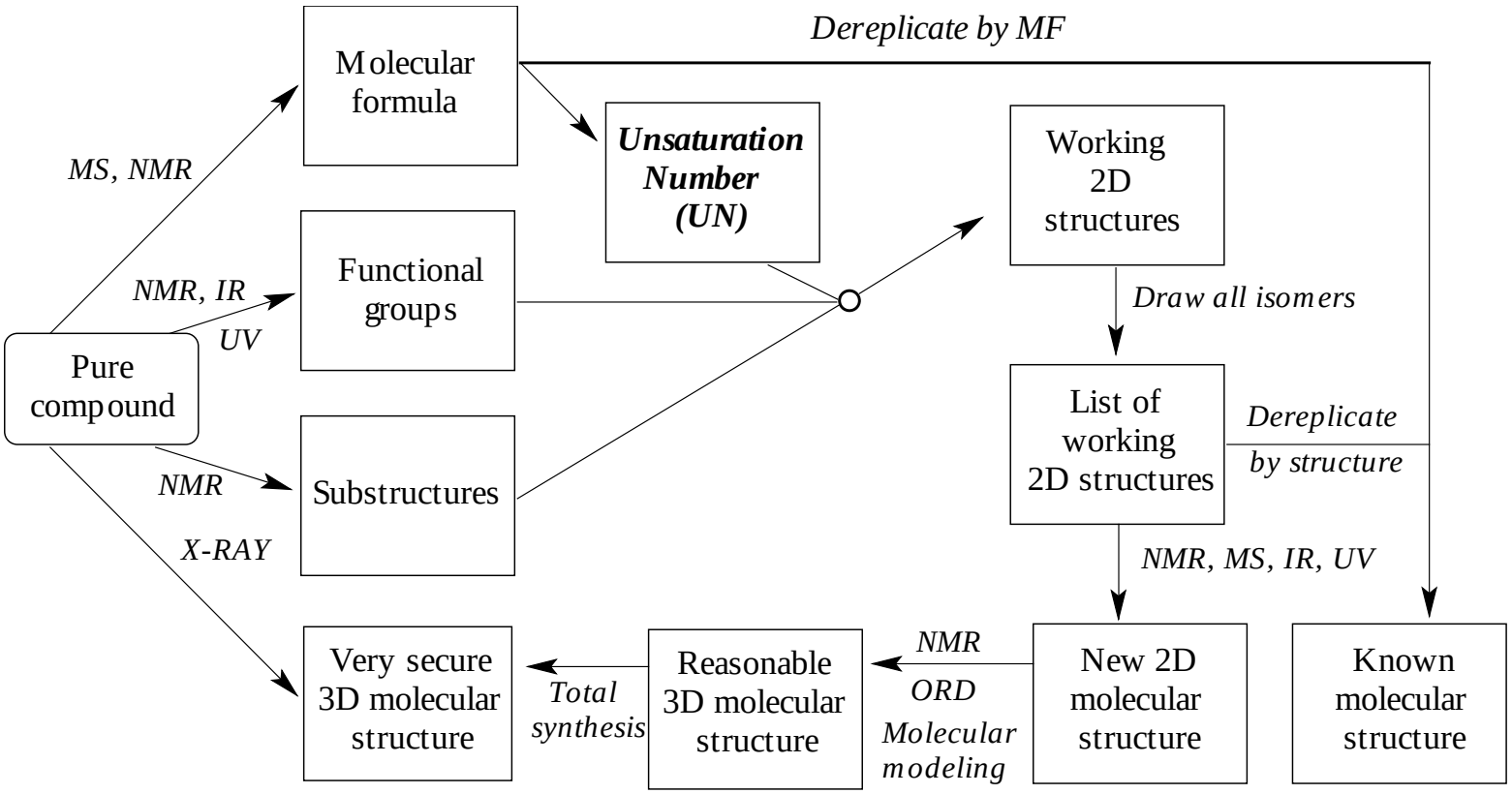




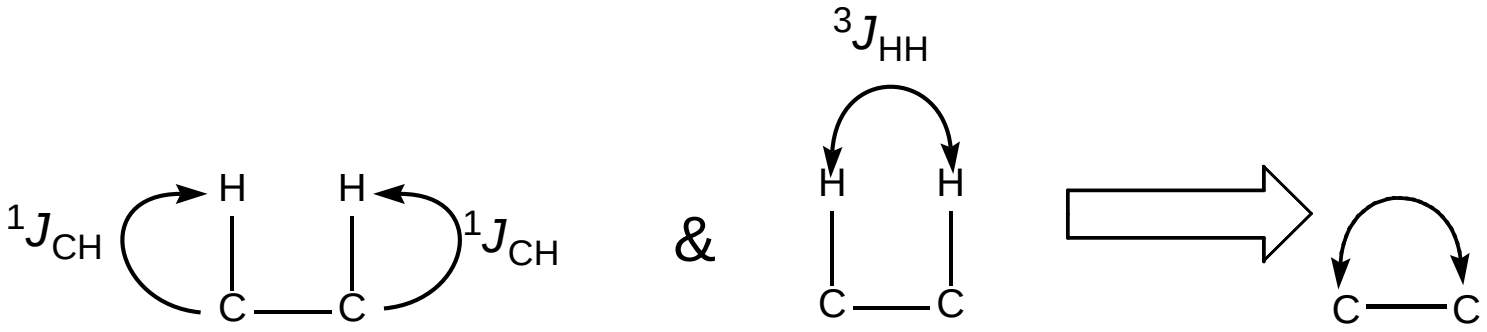
marine biodiscovery centre

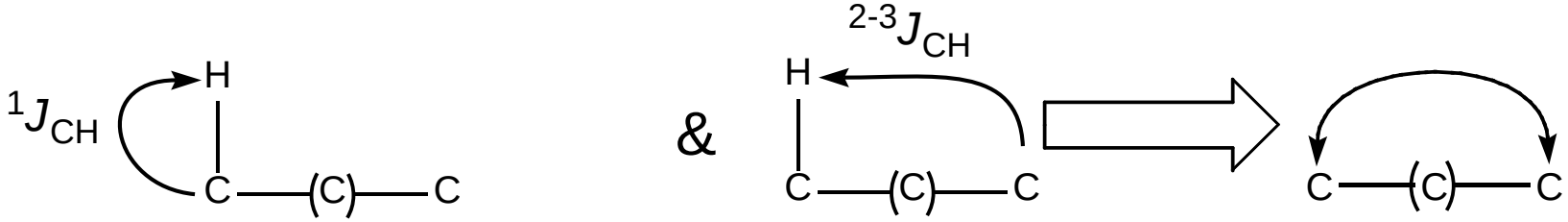
Structure Determination Steps

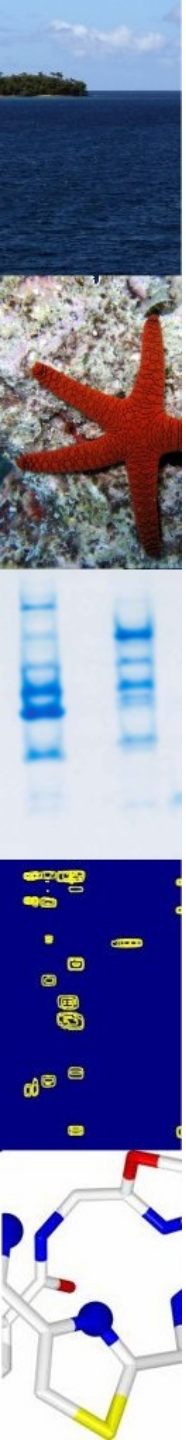
UNIVERSITY OF ABERDEEN



Strategy Based on C-H Connectivity







marine biodiscovery centre

Structure Determination Steps

MF = $C_{10}H_{18}O$

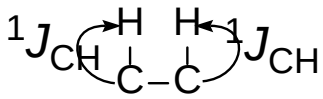
sp^3: Me Me Me

sp^3: CH_2 CH 2 CH 2

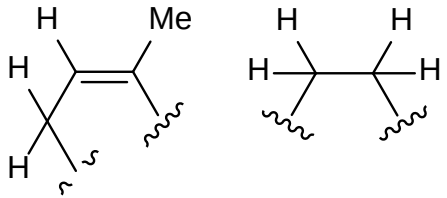
sp^3: CH; sp^2: CH

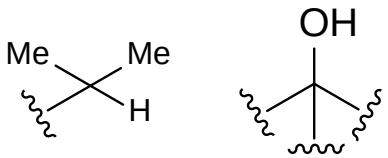
sp^3: C-OH; sp^2: C

1.) Obtain C/H pairs from HSQC

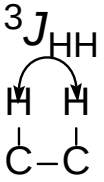


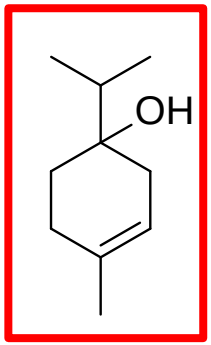
2.) Determine functional groups (NMR, IR etc)

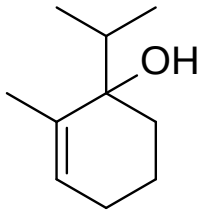


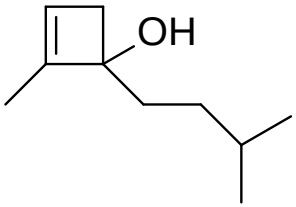


3.) Generate substructures

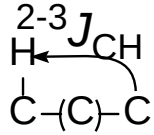






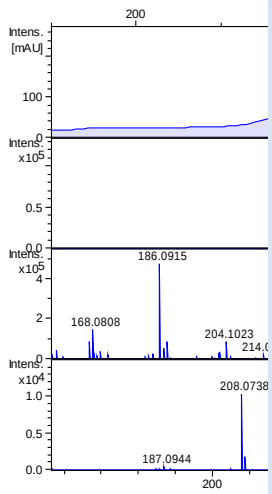


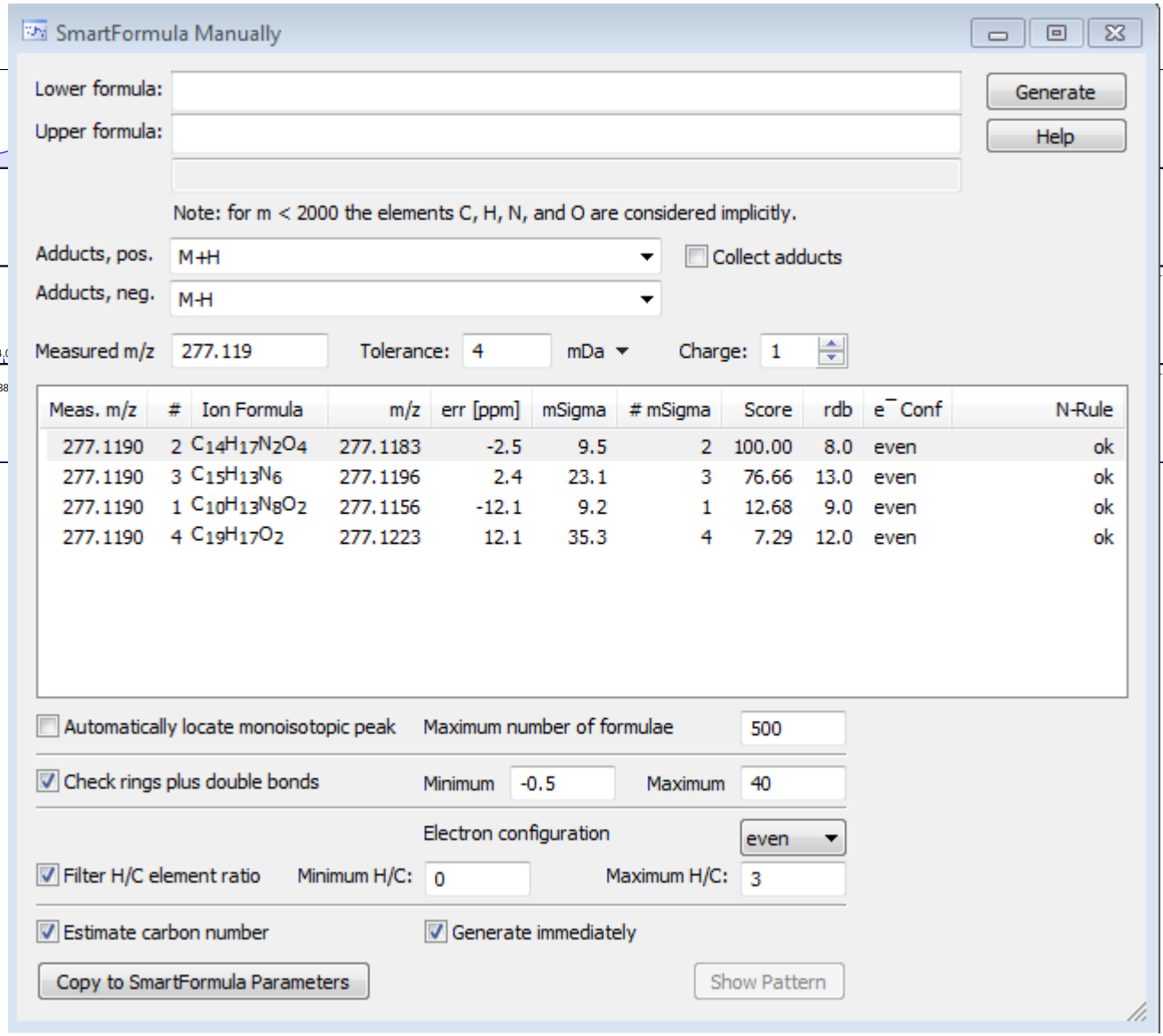
4.) Assemble working structures

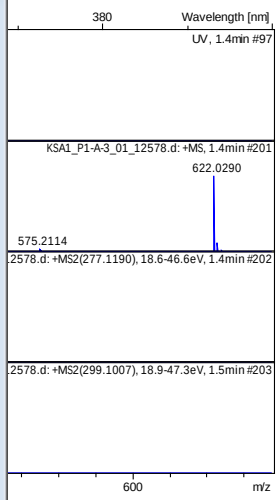


5.) Make structural proposal

Determining the MF



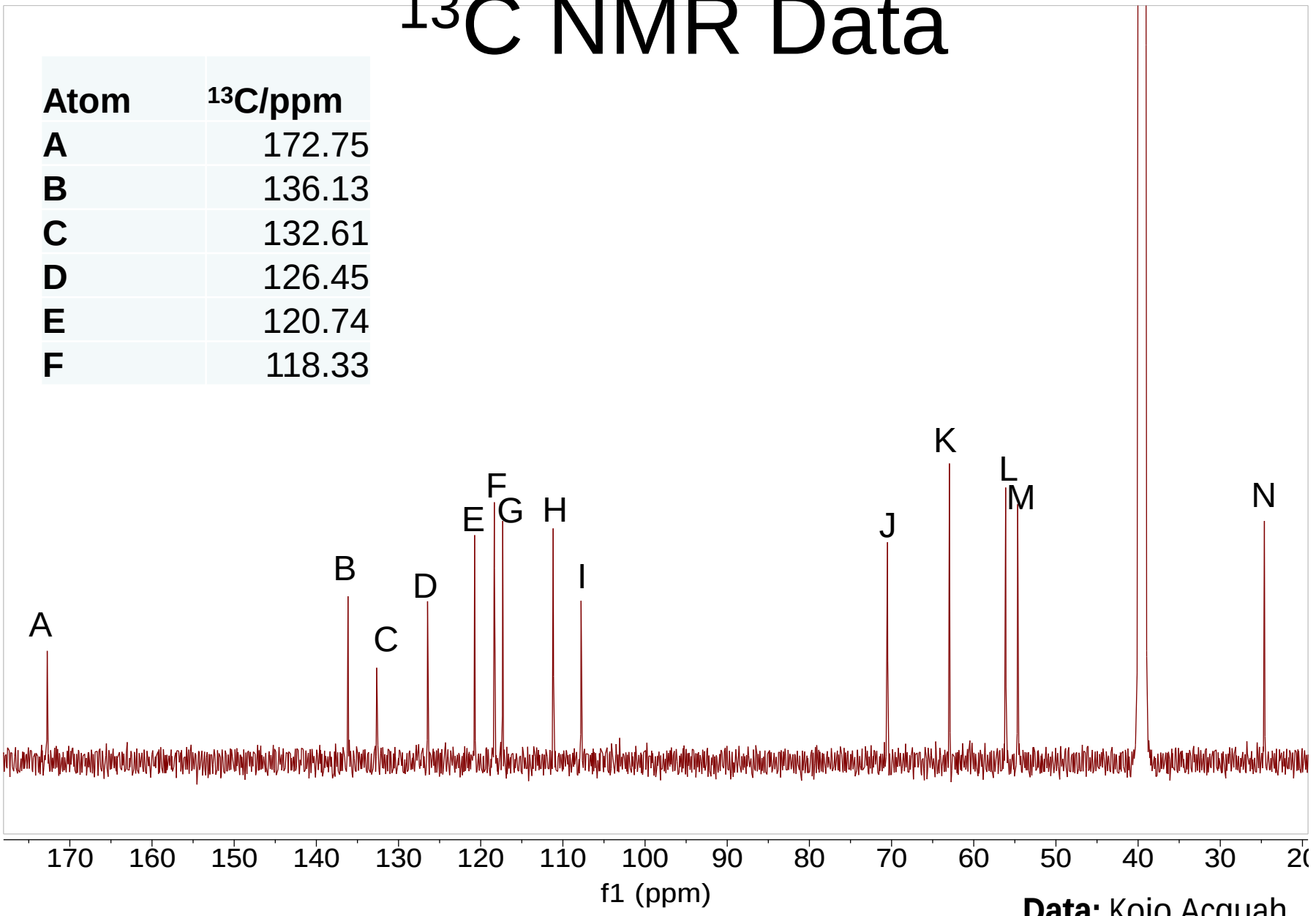




Data:
Kojo Acquah

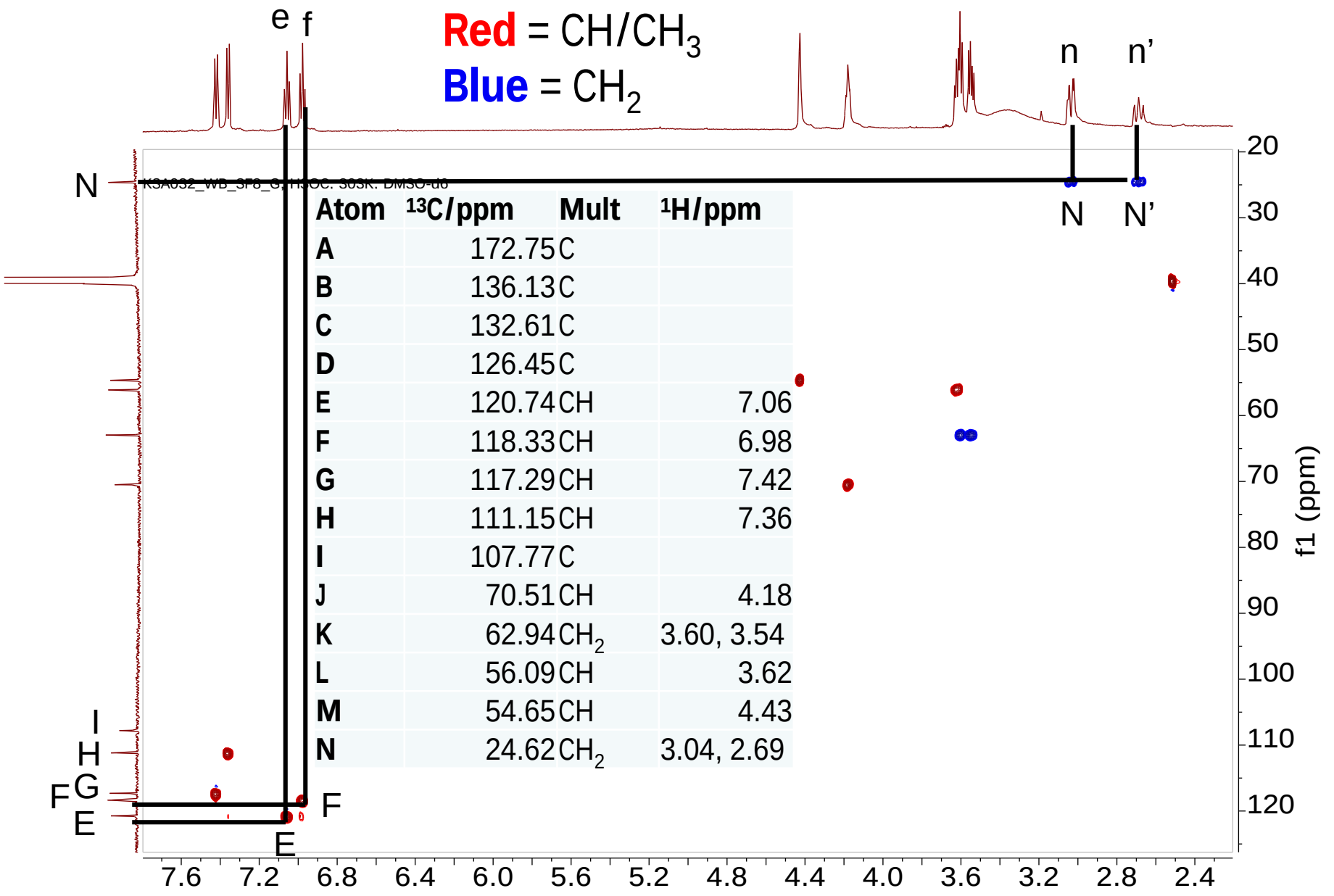
^{13}C NMR Data

Atom	^{13}C/ppm
A	172.75
B	136.13
C	132.61
D	126.45
E	120.74
F	118.33



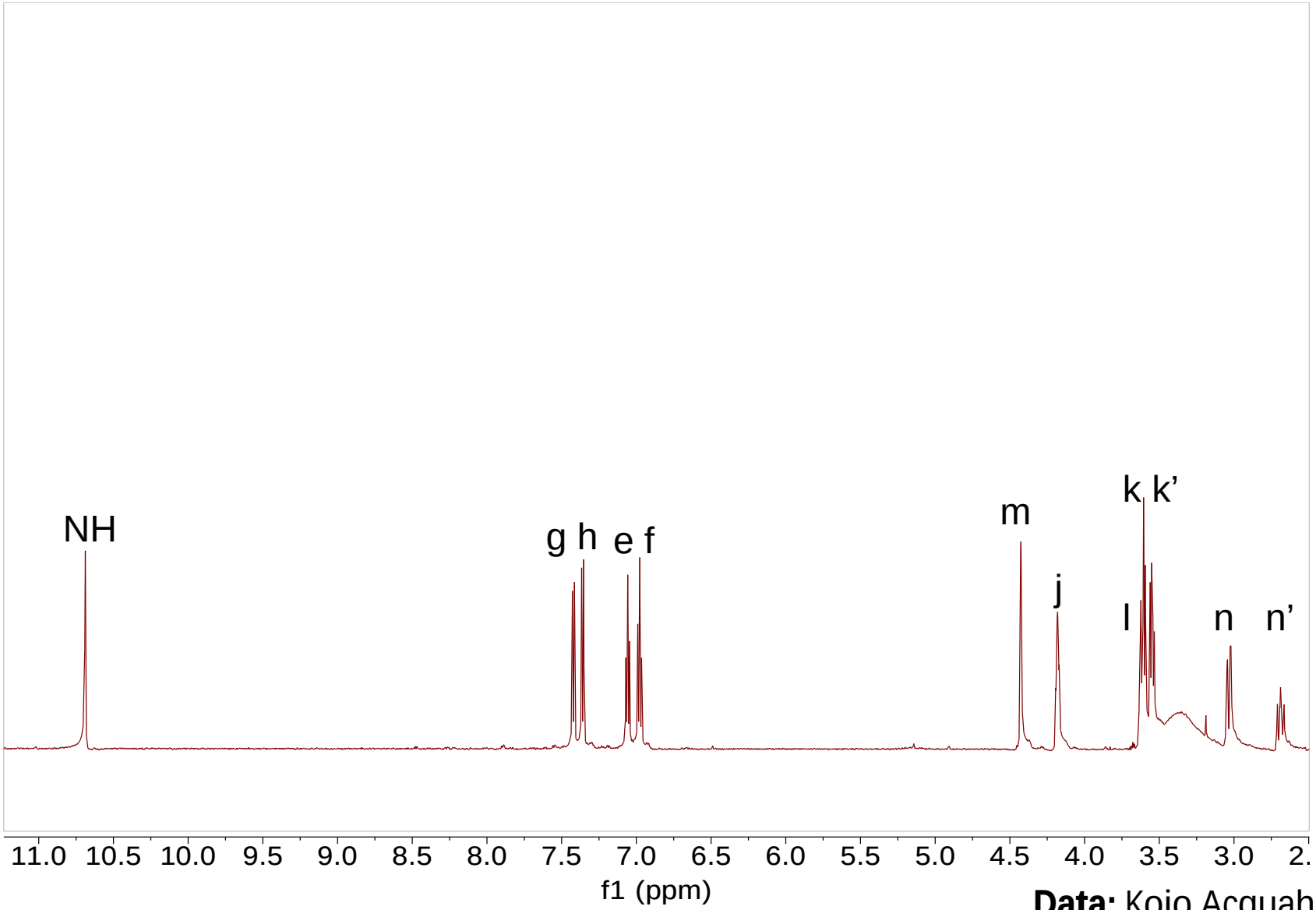
Data: Kojo Acquah

HSQC Data - C-H (1 bond)



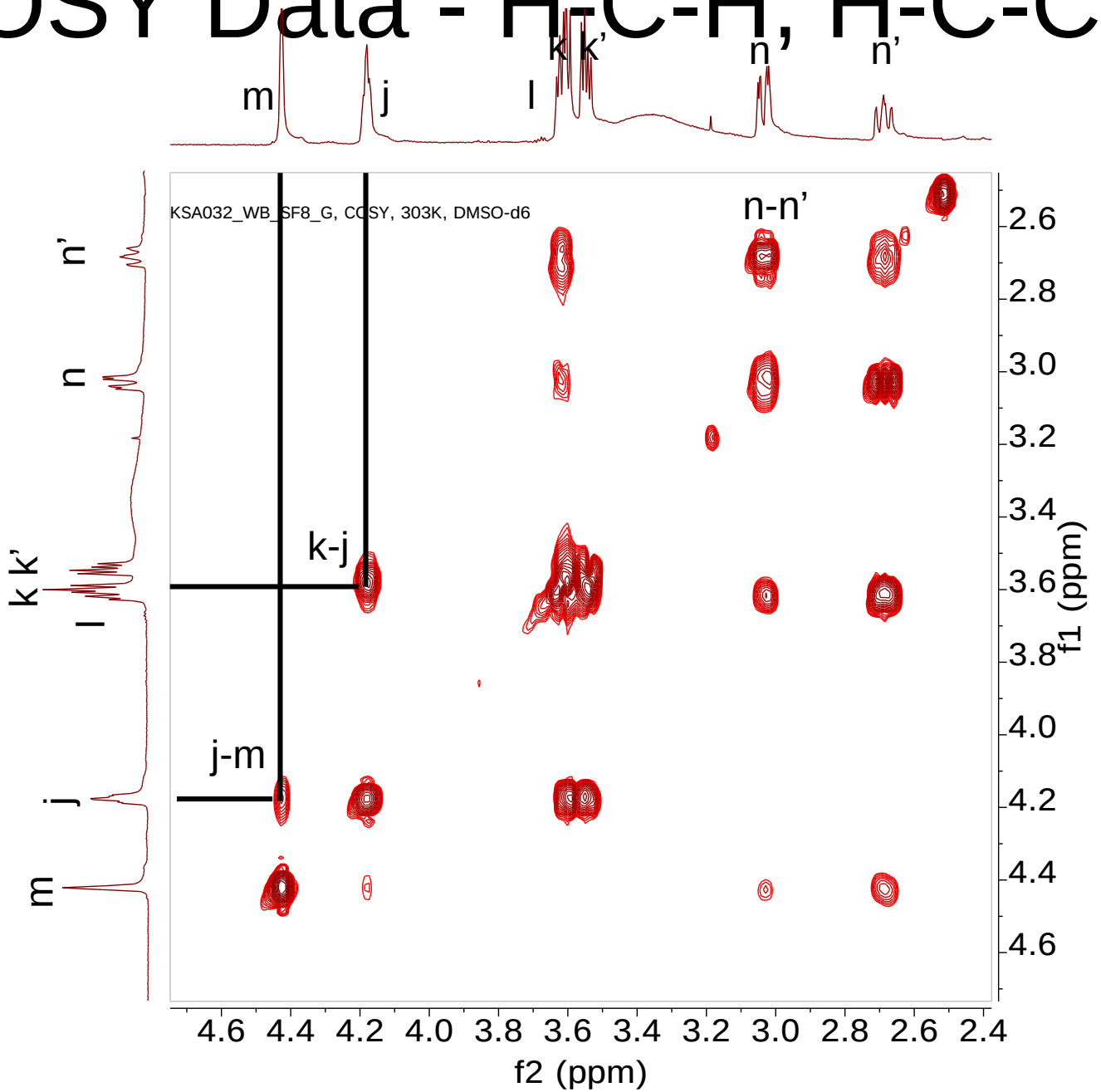
Atom	^{13}C/ppm	Mult	^{1}H/ppm
A	172.75	C	
B	136.13	C	
C	132.61	C	
D	126.45	C	
E	120.74	CH	7.06
F	118.33	CH	6.98
G	117.29	CH	7.42
H	111.15	CH	7.36
I	107.77	C	
J	70.51	CH	4.18
K	62.94	CH_2	3.60, 3.54
L	56.09	CH	3.62
M	54.65	CH	4.43
N	24.62	CH_2	3.04, 2.69

^{1}H NMR Data



Data: Kojo Acquah

COSY Data - H-C-H, H-C-C-H

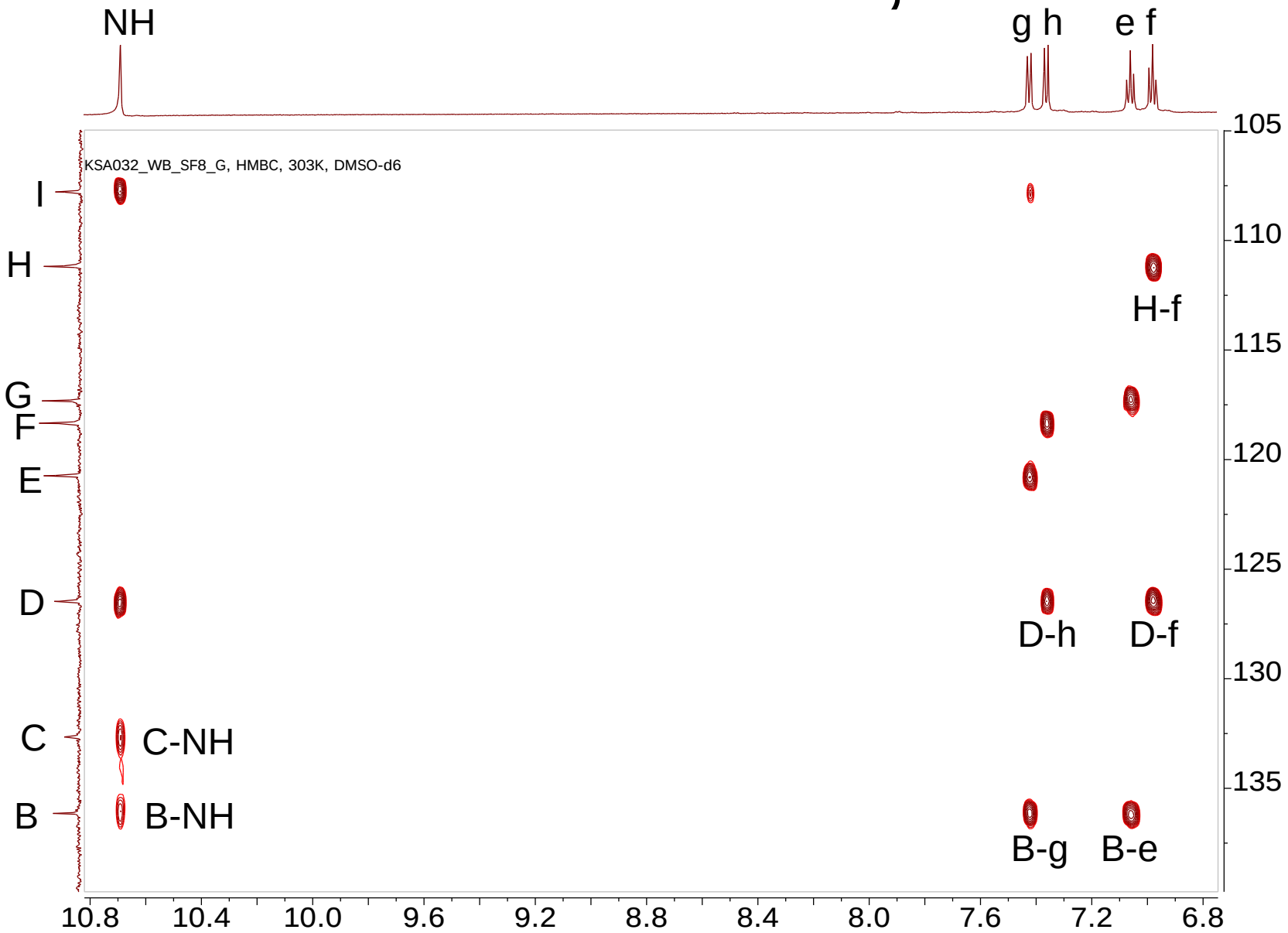


COSY Data - H-C-H, H-C-C-H

Atom	^{13}C/ppm	Mult	^{1}H	COSY
A	172.75	C		
B	136.13	C		
C	132.61	C		
D	126.45	C		
E	120.74	CH	7.06	f, h
F	118.33	CH	6.98	e, g
G	117.29	CH	7.42	f
H	111.15	CH	7.36	e
I	107.77	C		
J	70.51	CH	4.18	m, k
K	62.94	CH_2	3.60, 3.54	j
L	56.09	CH	3.62	n
M	54.65	CH	4.43	j, n
N	24.62	CH_2	3.04, 2.69	l, m

Data: Kojo Acquah

HMBC Data - C-C-H, C-C-C-H



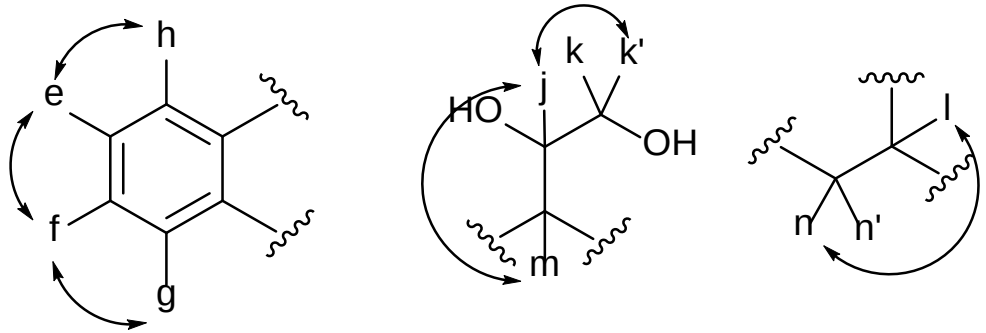
HMBC Data - C-C-H, C-C-C-H

Atom	^{13}C/ppm	Mult	^{1}H	COSY	HMBC
A	172.75	C			l, n
B	136.13	C			HN, g, e
C	132.61	C			HN, m, j, n
D	126.45	C			HN, h, f, n
E	120.74	CH	7.06	f, h	g, f
F	118.33	CH	6.98	e, g	h
G	117.29	CH	7.42	f	e,f
H	111.15	CH	7.36	e	f, g, e
I	107.77	C			HN, g, m, n
J	70.51	CH	4.18	m, k	m, k
K	62.94	CH_2	3.60, 3.54	j	j, m
L	56.09	CH	3.62	n	n
M	54.65	CH	4.43	j, n	l, k
N	24.62	CH_2	3.04, 2.69	l, m	l

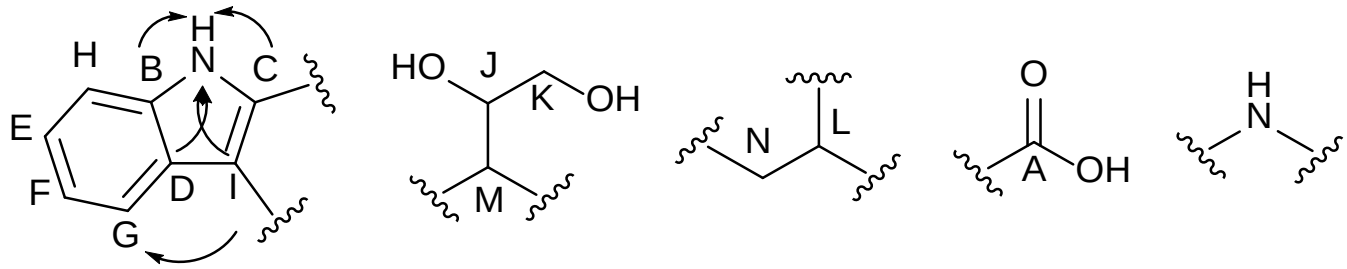
Data: Kojo Acquah

Substructures

COSY Data

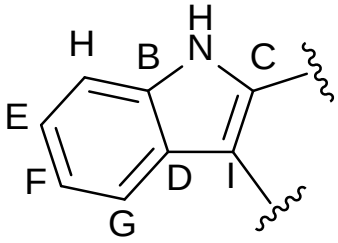


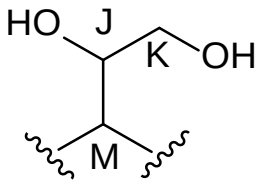
HMBC Data

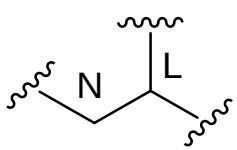


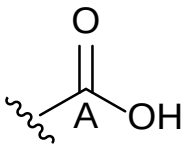
$C_{14}H_{16}N_2O_4$

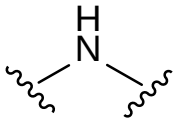
Combining the Pieces



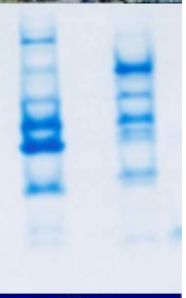


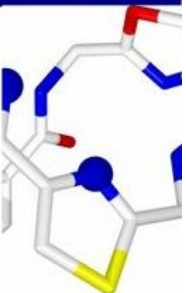












marine biodiscovery centre

UNIVERSITY OF ABERDEEN

Computer Aided Structure Elucidation

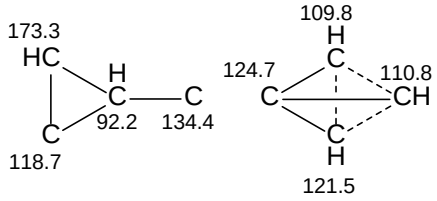
HSQC analysis - assign C-H

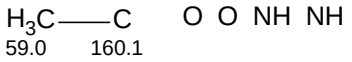
Tabulate data

#	d_C	d_H
1	173.3	9.27
2	121.5	6.87
3	110.8	6.09
⋮	⋮	⋮

COSY/HMBC analysis

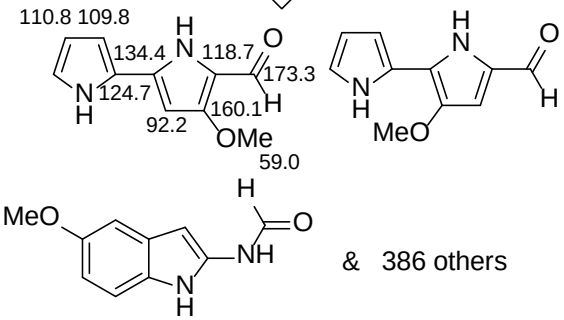
Generate connectivities





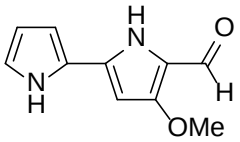
= MF ($C_{10}H_{10}N_2O_2$)

Structure generation

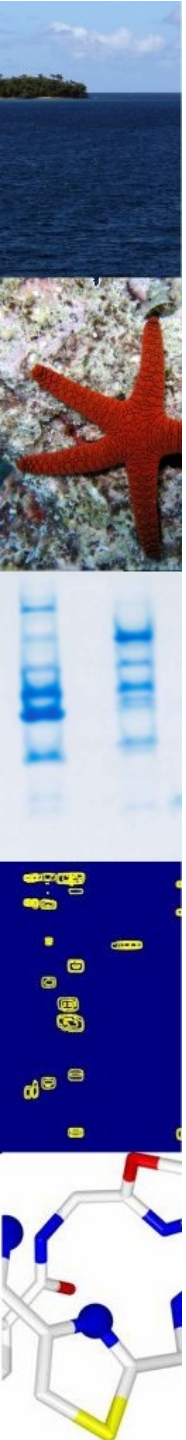


& 386 others

Calculate match with ^{13}C shifts



Correct structure



marine biodiscovery centre

The Determinator

UNIVERSITY OF ABERDEEN

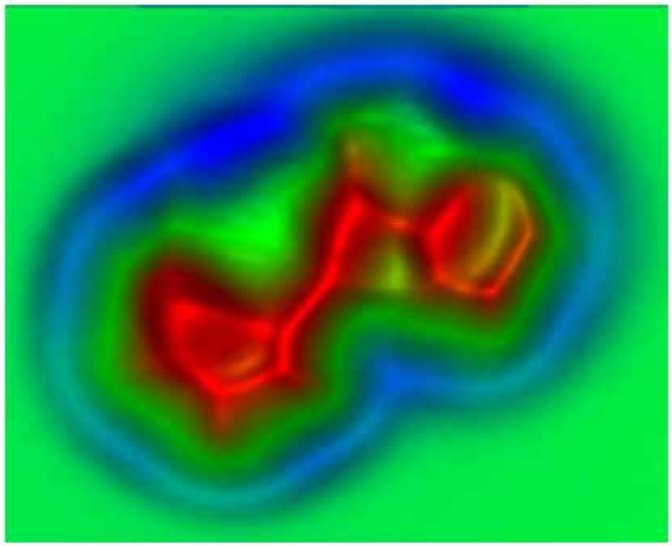
- Fast
- Easy to use
- Generally applicable
- Reliable
- Inexpensive

marine biodiscovery centre

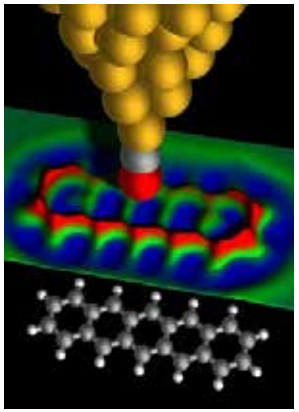
Solving Structures Using Atomic Force Microscopy

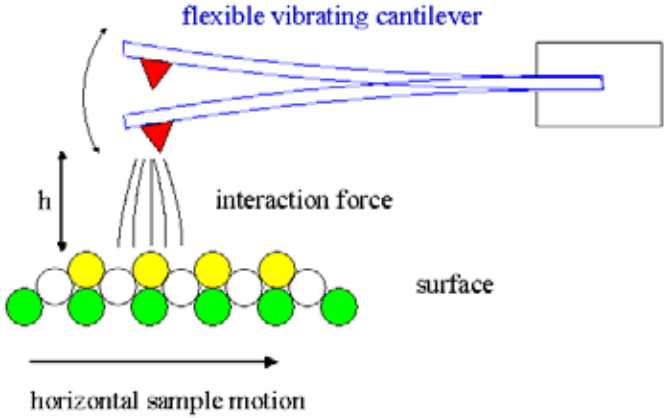
UNIVERSITY OF ABERDEEN

IBM



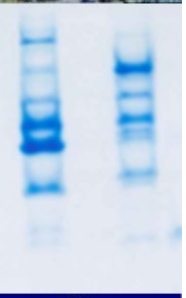
Const. height AFM, Δf channel, 3D representation

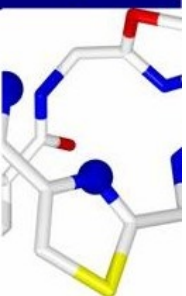




Nat. Chem. 2010

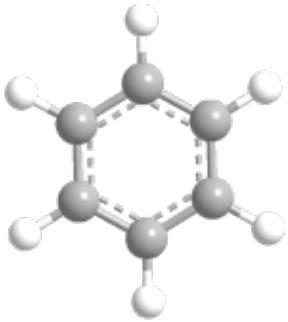


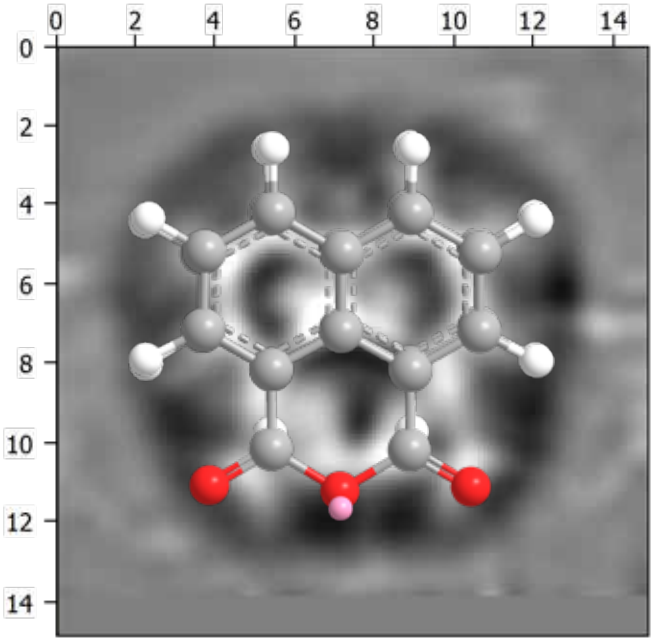


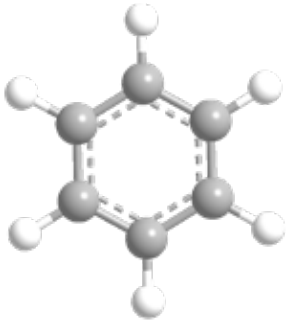


marine biodiscovery centre

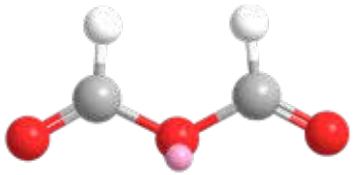
The Determinator

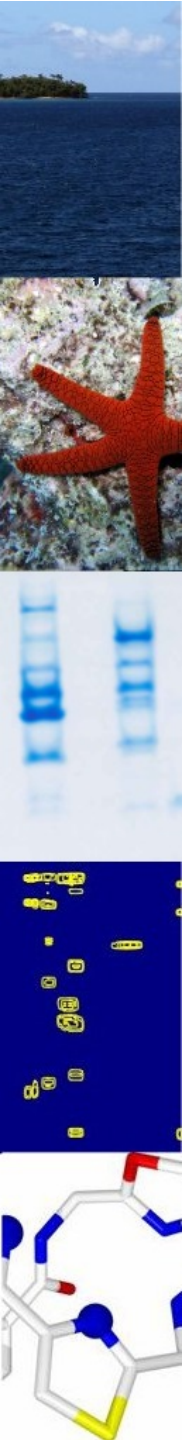






AFM of Unknown





marine biodiscovery centre

Applying the Determinator to Breitfussin A

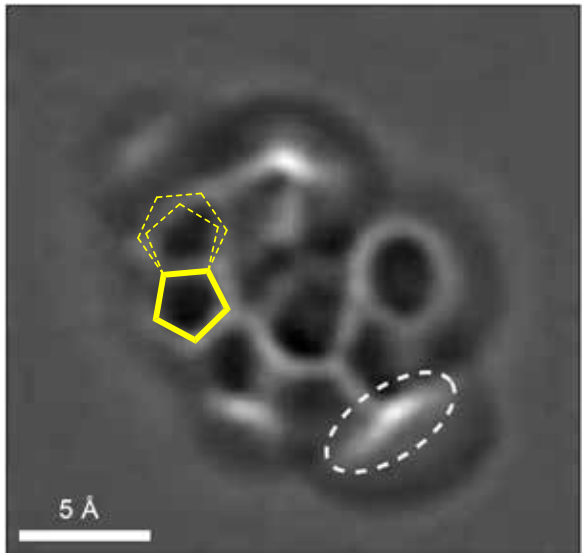
UNIVERSITY OF ABERDEEN

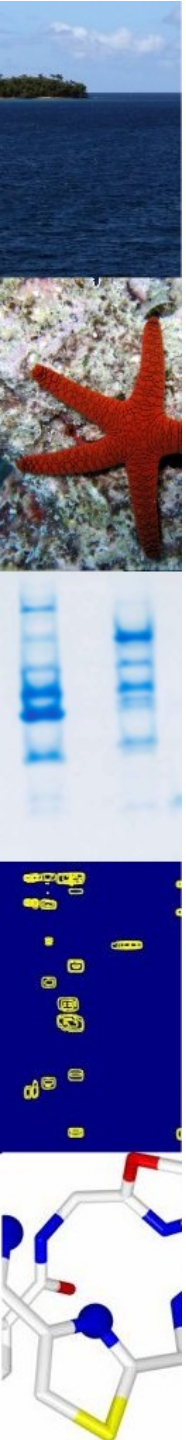
IBM

Prerequisite

$C_{15}H_8N_3O_1$

—Br —I —OMe



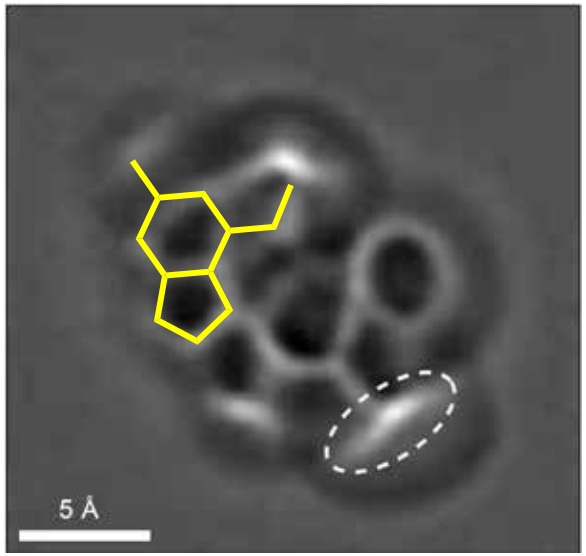


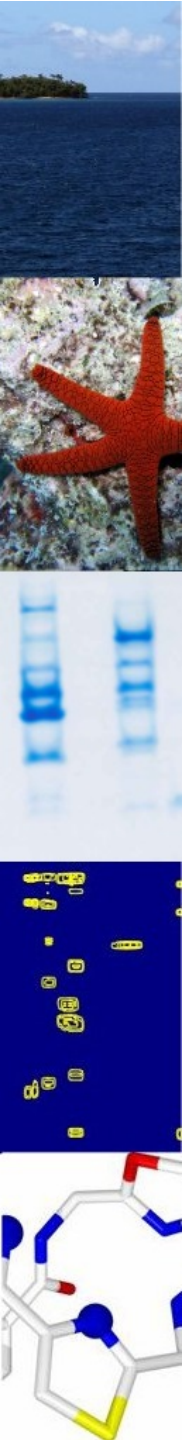
marine biodiscovery centre

Applying the Determinator to Breitfussin A

UNIVERSITY OF ABERDEEN

IBM



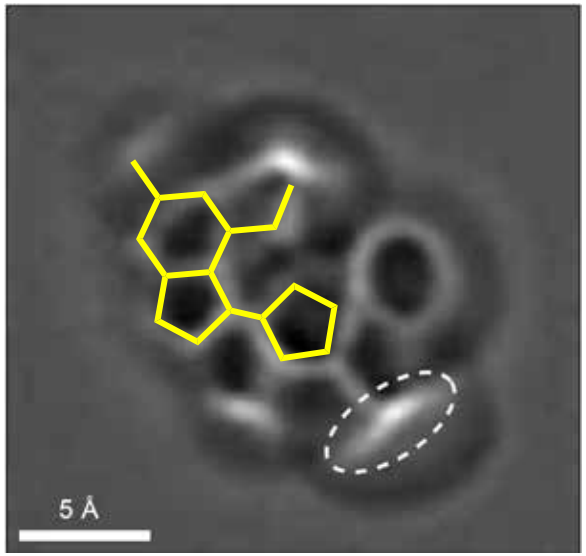


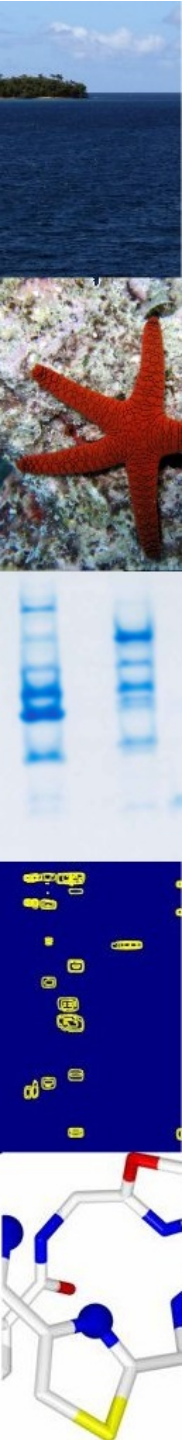
marine biodiscovery centre

Applying the Determinator to Breitfussin A

UNIVERSITY OF ABERDEEN

IBM



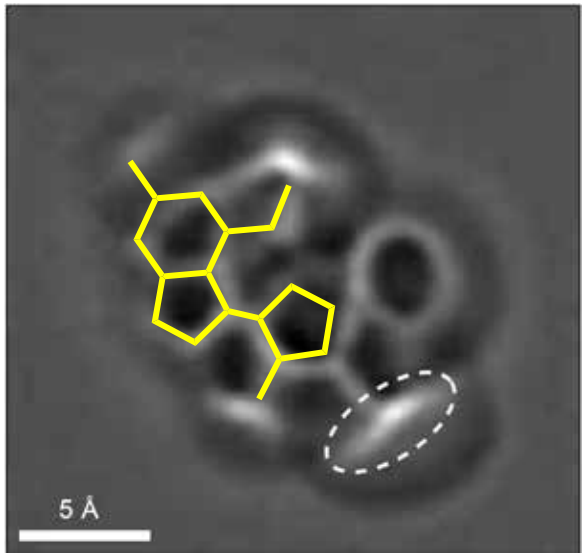


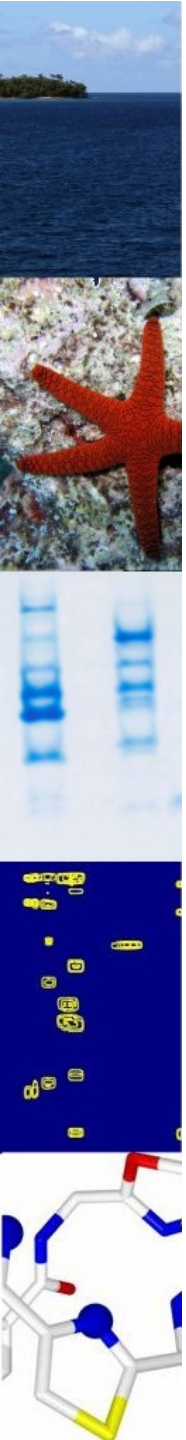
marine biodiscovery centre

Applying the Determinator to Breitfussin A

UNIVERSITY OF ABERDEEN

IBM



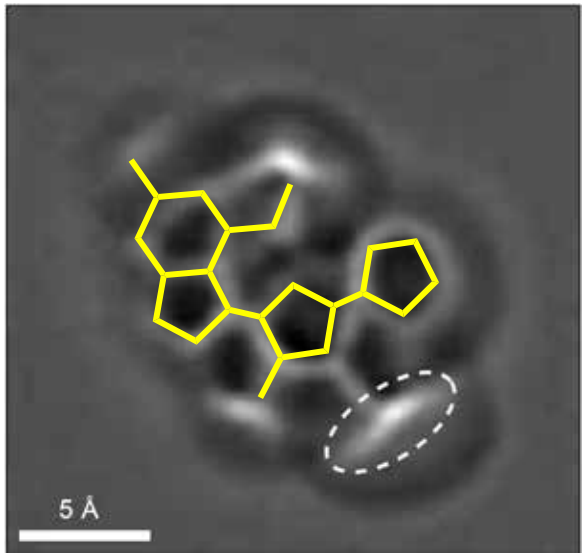


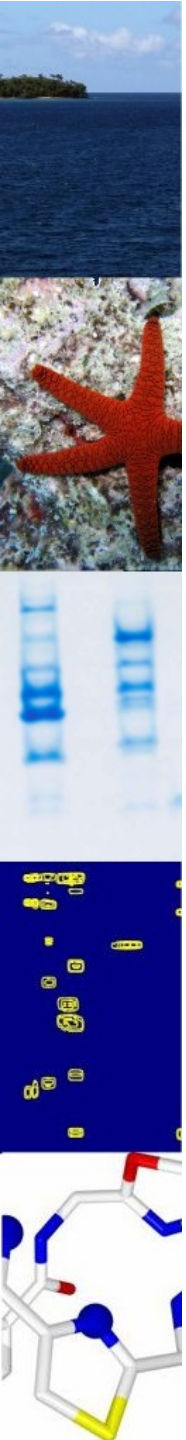
marine biodiscovery centre

Applying the Determinator to Breitfussin A

UNIVERSITY OF ABERDEEN

IBM



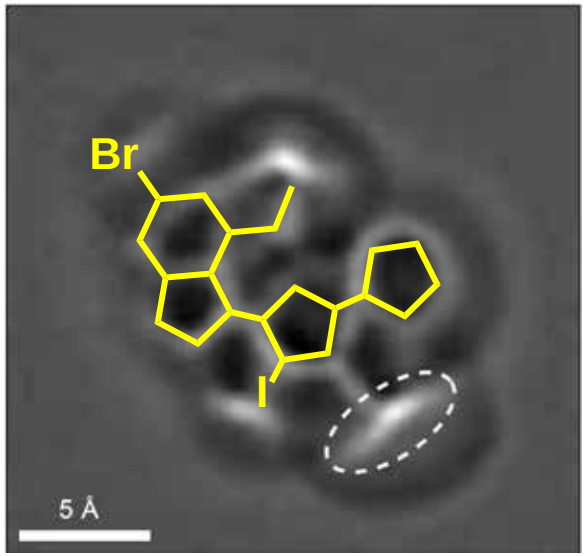


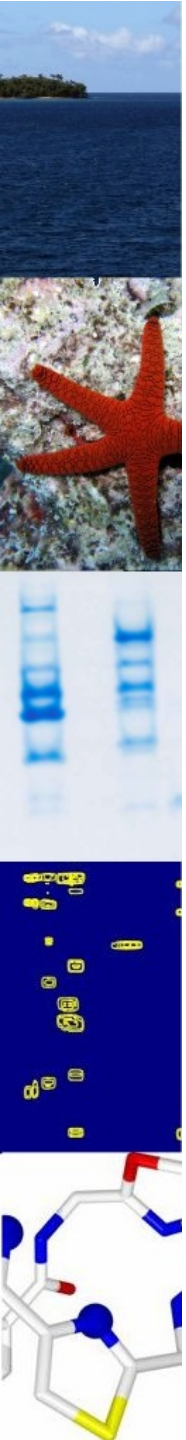
marine biodiscovery centre

Applying the Determinator to Breitfussin A

UNIVERSITY OF ABERDEEN

IBM





marine biodiscovery centre

UNIVERSITY OF ABERDEEN

IBM

"If you have a strange substance and you want to know what it is, you go through a long and complicated process of chemical analysis. ..... It would be very easy to make an analysis of any complicated chemical substance; all one would have to do would be to look at it and see where the atoms are."

Plenty of Room at the Bottom
Richard P. Feynman
December 1959