



PROTEIN METRICS

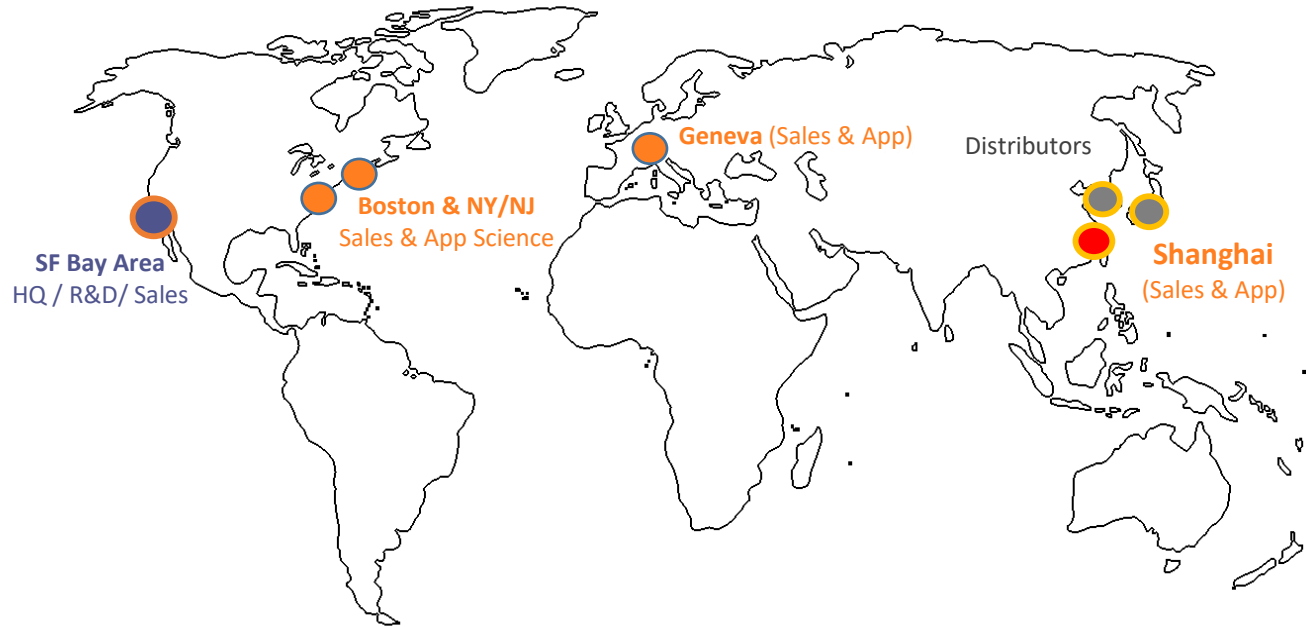
Software for Oligonucleotide Analysis by Mass Spectrometry

Marshall Bern

October 10th 2019

SUMS, Stanford CA

Protein Metrics Inc. - Overview

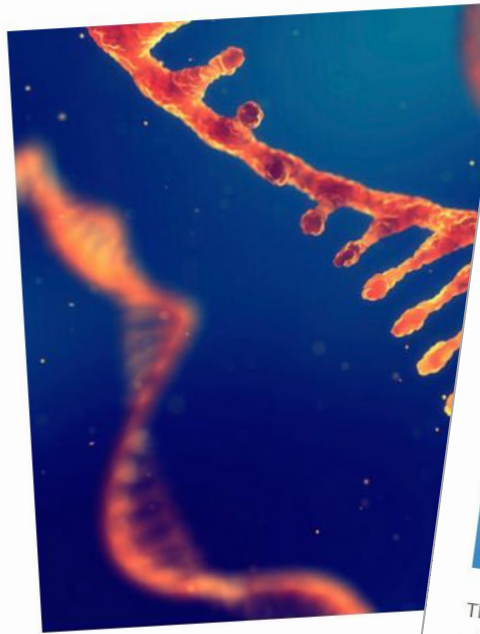


- Software company founded in 2011 (Cupertino, same address as Apple circa 1978)
- In almost all major biopharma companies
- > 100 academic laboratories
- 9 SBIR / STTR grants from NIH NIGMS



Synthetic oligonucleotides are **hot** !

"2018 the year of RNA", says Genetic Engineering and Biotechnology News.

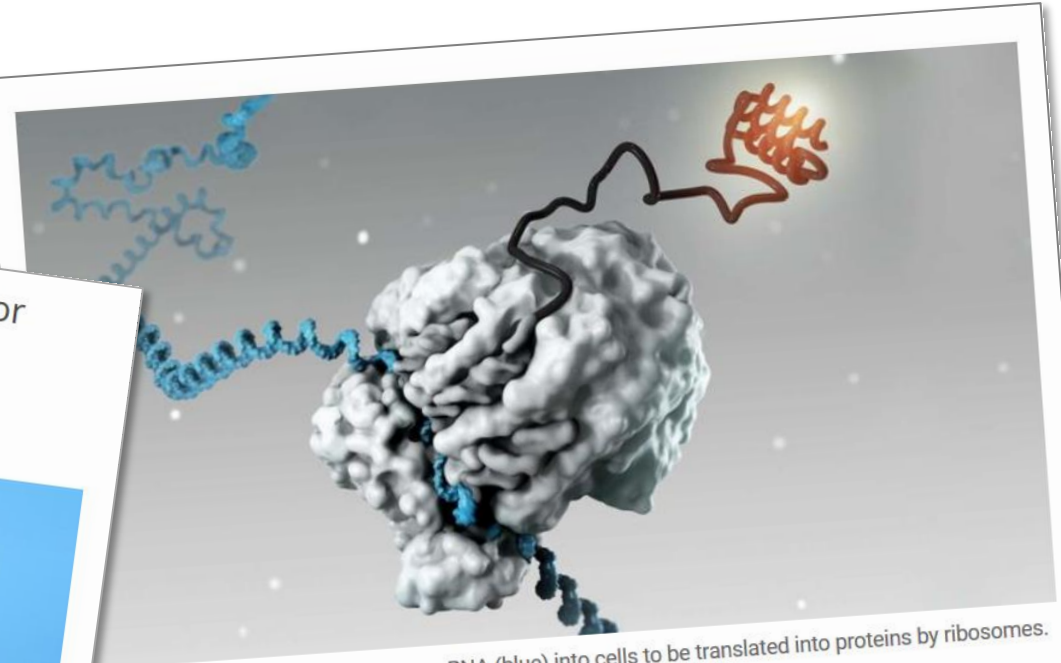


Moderna Therapeutics Sets Record for Biggest Biotech IPO

Published: Dec 07, 2018 | By Mark Terry



The long-awaited and massively hyped initial public offering (IPO) of **Moderna Therapeutics** hit the market yesterday. The company sold approximately 26.3 million shares priced at \$23 a share. This exceeded the revised goal of \$600 million by about \$4.3 million. Shares are trading on the Nasdaq under the "MRNA" ticker symbol. The raise values the company at about \$7.5 billion.



biotech Moderna delivers messenger RNA (blue) into cells to be translated into proteins by ribosomes.
LTOUNIAN/SCIENCE

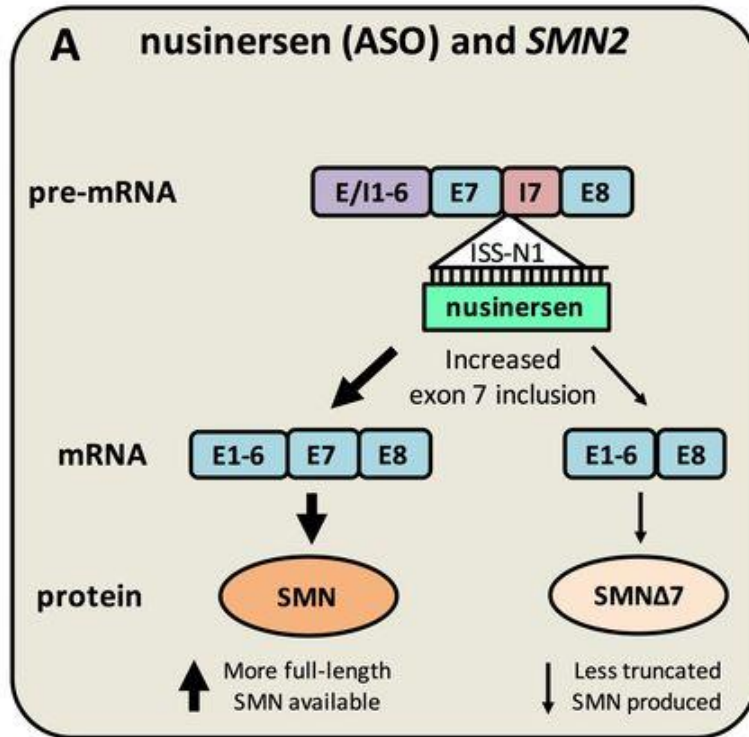
This mysterious \$2 billion biotech is revealing the secrets behind its new drugs and vaccines

By Kelly Servick | Feb. 1, 2017, 2:30 PM

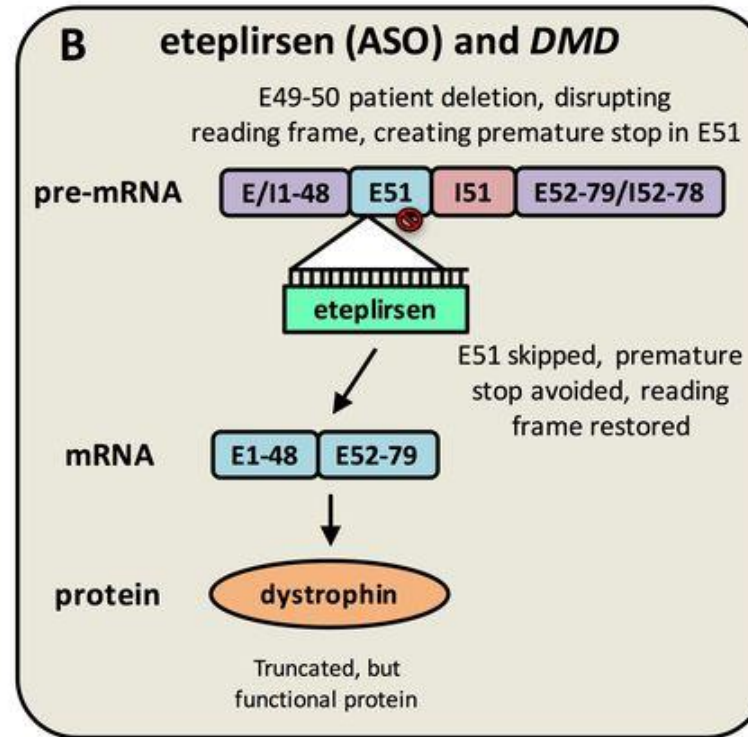
Drug Mechanisms

Single-stranded antisense oligos (ASO) can repair bad genes

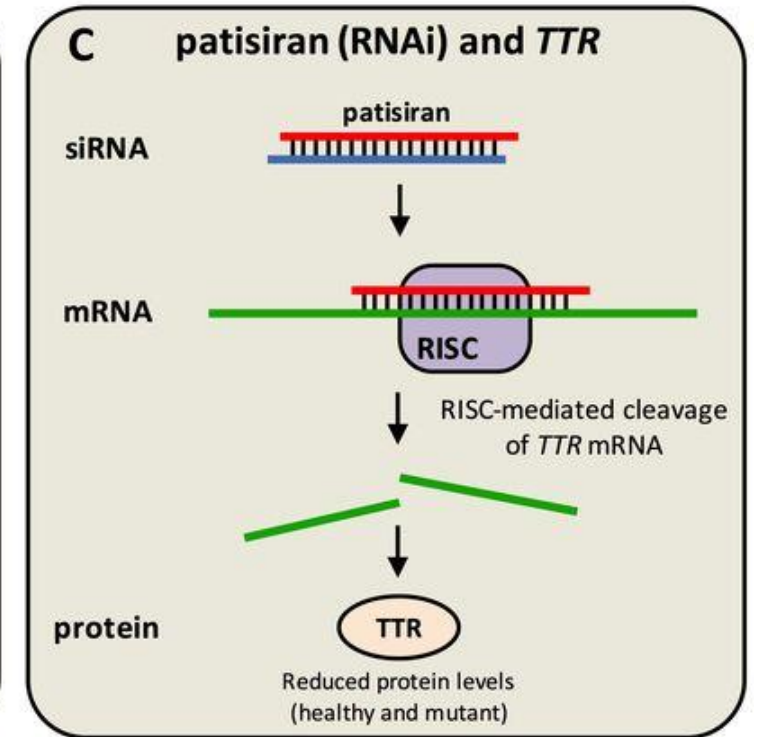
Double-stranded RNAi can kill unwanted mRNA's



spinal muscular atrophy

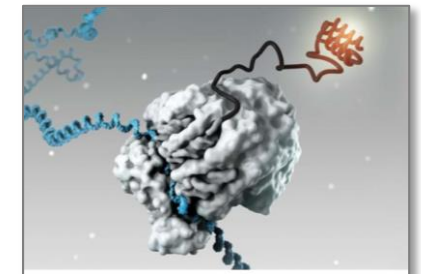


Duchenne muscular dystrophy



familial amyloid polyneuropathy

... and someday mRNAs will make new proteins!



Eteplirsen

From Wikipedia, the free encyclopedia

Eteplirsen (brand name **Exondys 51**) is a medication to treat, but not cure, some types of **Duchenne muscular dystrophy** (DMD), caused by a specific mutation. Eteplirsen only targets specific mutations and is useful in about 1.5% of cases.^{[1][2]} **DMD ≈ 1 in 3500 male births**

Eteplirsen was designed and developed by **Sarepta Therapeutics**. After a controversial debate surrounding the efficacy of the drug, during which two FDA review panel members resigned in protest, eteplirsen received accelerated approval from the US Food and Drug administration in late 2016.^{[3][4]} A year's worth of treatment is expected to cost approximately \$300,000.^[5]

CTCCAACATCAAGGAAGATGGCATTCTAG

NDA 206488

David B. Hawver, Ph.D.

Chemical Name

RNA, [*P*-deoxy-*P*-(dimethylamino)] (2',3'-dideoxy-2',3'-imino-2',3'-seco) (2'a→ 5') (C-m⁵U-C-C-A-A-C-A-m⁵U-C-A-A-G-G-A-A-G-A-m⁵U-G-G-C-A-m⁵U-m⁵U-m⁵U-C-m⁵U-A-G), 5'-[*P*-[4-[[2-[[2-(2-hydroxyethoxy)ethoxy]ethoxy]carbonyl]-1-piperazinyl]-*N,N*-dimethylphosphoramidate]

Molecular Formula

C₃₆₄H₅₆₉N₁₇₇O₁₂₂P₃₀

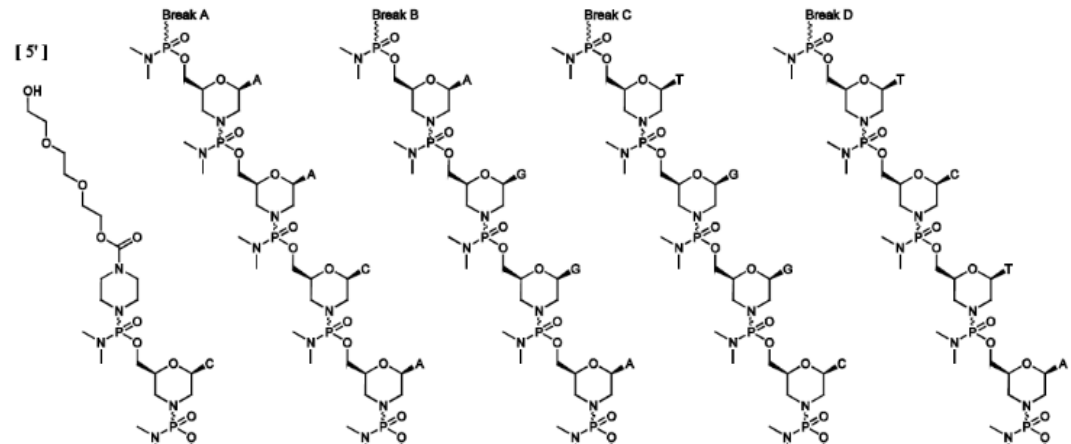
Molecular Weight

10305.7

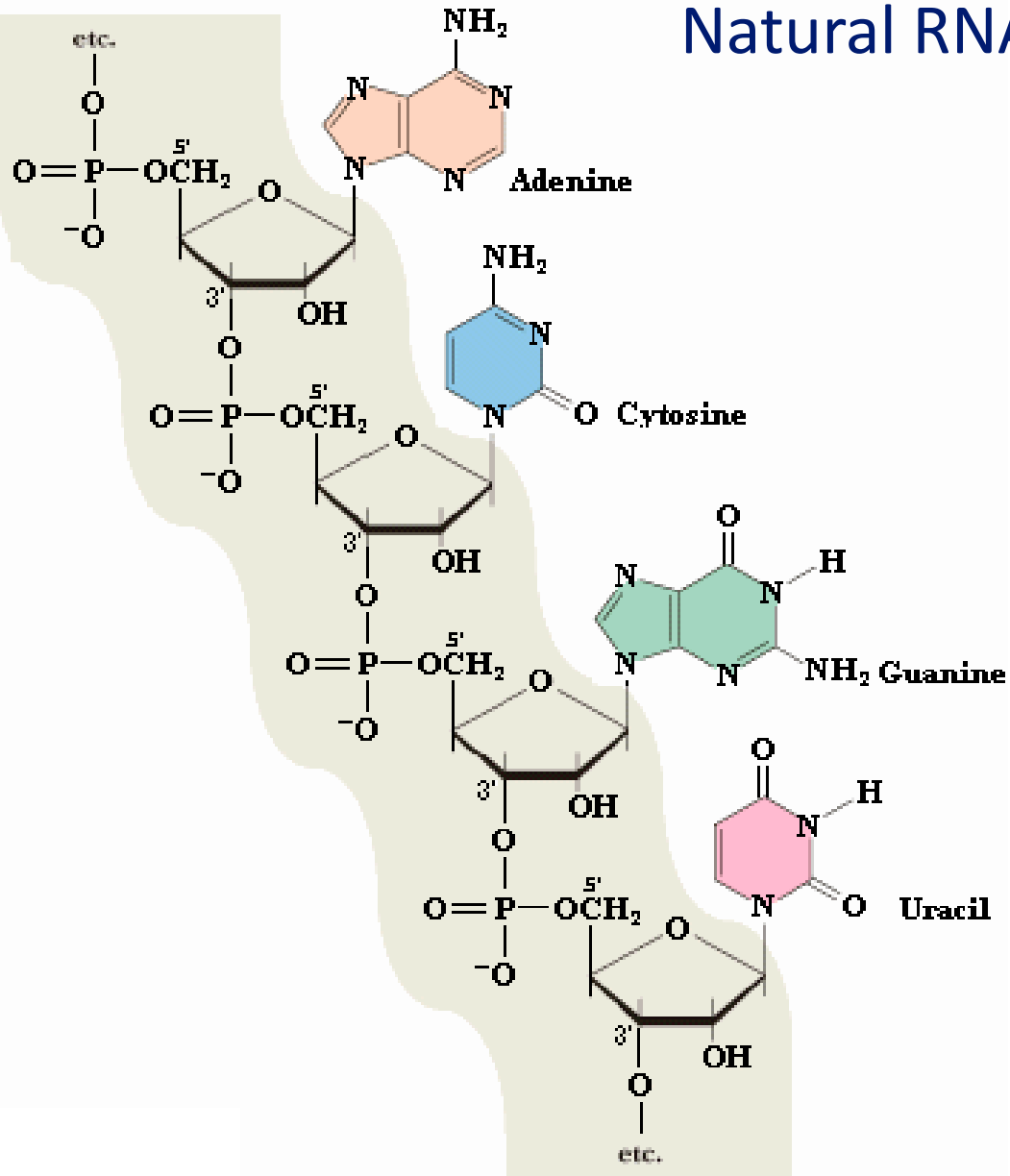
What's this?

Biochemical Description and Structure

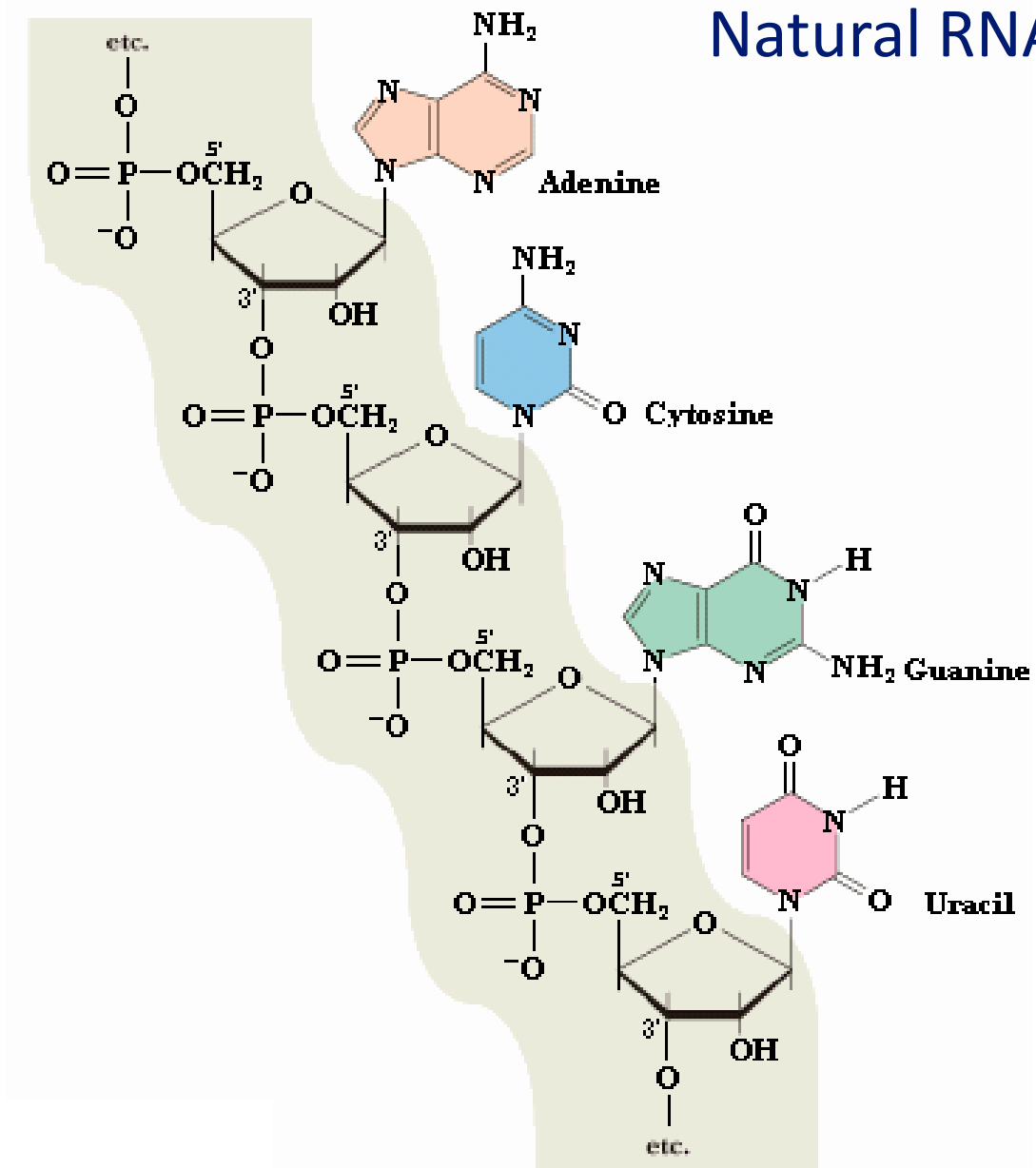
Eteplirsen is a charge-neutral phosphorodiamidate morpholine oligomer (PMO) consisting of a sequence of 30 of the following four nucleobases: 5-methyluracil (m⁵U; thymine; T), adenine (A), cytosine (C), and guanine (G).



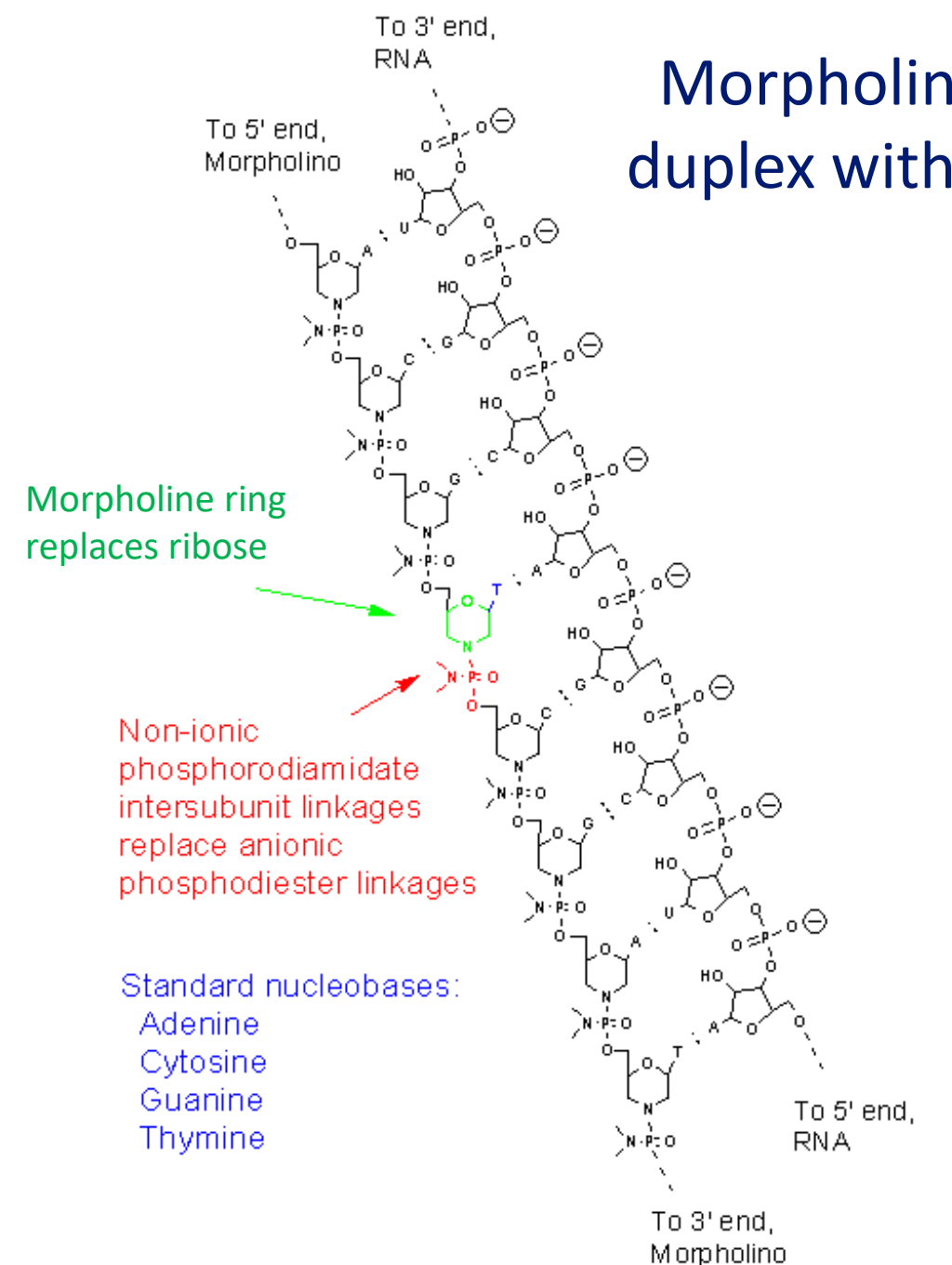
Natural RNA



Natural RNA



Morpholino in duplex with RNA



Patisiran



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

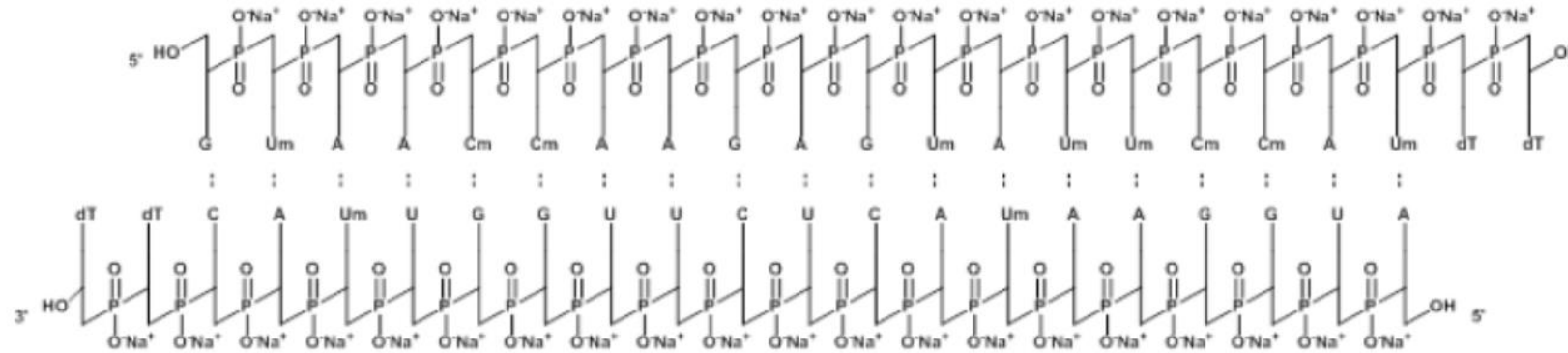
26 July 2018
EMA/554262/2018
Committee for Medicinal Products for Human Use (CHMP)

Assessment report

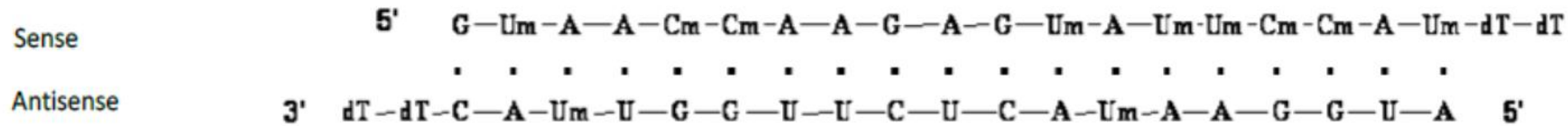
Onpattro

International non-proprietary name: patisiran

Sense Strand



Antisense Strand



The chemical structures and physical properties of patisiran and the single strand intermediates were elucidated and confirmed by a combination of **LC-MS, MS-MS sequence confirmation**, thermal dependent UV absorbance, SE-HPLC UV, FAAS, UV absorption, UV spectroscopy, ¹H-NMR, Imino-¹H-NMR, ¹³C-NMR, ³¹P-NMR, FTIR, circular dichroism, differential scanning calorimetry and thermogravimetric analysis.

LC-MS Analysis

Intact Mass™

PROTEIN METRICS INC.



Intact

Identification, characterization, and relative quantification of large molecules

New Project

PMI-Intact - Protein Metrics Inc. - Registered (v3.6-0 x64) - Days left: 120.4

File Edit Window Help

Project (double click to dock / undock) **Elution Peaks** Add Peaks... Delete Peaks Merge Peaks Split Peak Text filter Chromatogram (double click to dock / undock)

New Reference project

Sample input Protein input Sample-protein input Deconvolution Mass matching Advanced Log

Table Text filter

Import Items	Sample name	File	Sample type	Time unit	Start time of interest	End time of interest	Alignment max time	Trace channel
<input checked="" type="checkbox"/>	<Enter name for referen... --		Reference	Minutes	<enter value>	<enter value>	1.00	
<input checked="" type="checkbox"/>		<Enter filename... Unknown		Minutes				

PMI_Share_Data > Data2018 >

Name	Date modified	Type	Size
20180628_watersOST_1a.raw	9/4/2018 3:58 PM	File folder	
intact_tmp_files	8/21/2019 10:29 A...	File folder	
PMI Mass 20180628_watersOST_1a -Ilker.ntms	2/21/2019 1:14 AM	Protein Metrics Int...	12,332 KB
PMI Mass 20180628_watersOST_1a.ntms	10/3/2019 9:02 AM	Protein Metrics Int...	107,696 KB

Add sample Insert data to sample Remove Item(s) Peak construction options: UV... TIC... BPL...

Save Preset... Load Preset... Create... Cancel

Deconvolved Mass spectrum (double click to dock / undock) Protein sequence (double click to dock / undock)

Define one-letter code for oligo sequences

- Presets for natural nucleotides
- Type in atomic formula for morpholinos, “locked” nucleic acids, thiophosphate, etc.

The screenshot shows the 'New Reference project (read only)' window in Protein Metrics. The 'Protein input' tab is selected and highlighted with an orange box. Below the tabs, the 'Residues' section contains a table with columns for 'Letter', 'Chemical formula', and 'Average mass'. The 'Terminus' field is set to 'H P (-1) O (-2)'. The 'Chemical Elements' table is also visible. On the right, the 'Common modified residues' and 'Biopolymer type (pilot)' sections are shown, with the 'DNA', 'RNA', and 'Protein (default)' buttons highlighted by an orange box.

Letter	Chemical formula	Average mass
A	C ₁₀ H ₁₂ O ₅ N ₅ P	313.207578
B		
C	C ₉ H ₁₂ O ₆ N ₃ P	289.182778
D		
E		
F		
G	C ₁₀ H ₁₂ O ₆ N ₅ P	329.206948
H		
I		
J		
K		
L		
M		
N		
O		
P		
Q		
R		
S		
T	C ₁₀ H ₁₃ O ₇ N ₂ P	304.194216
U		

Element	Average mass
C	12.01079
H	1.007968
N	14.00669
O	15.99937
S	32.0639
P	30.973762
Se	78.971

Input sequences for mass matching

New Reference project (read only)

Sample input Protein input Sample-protein input Deconvolution Mass matching Advanced

Proteins Building blocks

Chains

Id	Name	Sequence/mass
A	6-nt	TTTTTT
B	7-nt	TTTTTTT
C	8-nt	TTTTTTTT
D	9-nt	TTTTTTTTT
E	10-nt	TTTTTTTTTT
F	11-nt	TTTTTTTTTTT
G	12-nt	TTTTTTTTTTTT
H	13-nt	TTTTTTTTTTTTT
I	14-nt	TTTTTTTTTTTTTT

Select from FASTA file... Add Remove selected

Proteins/protein complexes

Mirror Chains table

File	Name	Alias	Composition	Disulfides	Average mass
	6-nt		A(1)	All possible dis...	1763.20
	7-nt		B(1)	All possible dis...	2067.39
	8-nt		C(1)	All possible dis...	2371.59
	9-nt		D(1)	All possible dis...	2675.78
	10-nt		E(1)	All possible dis...	2979.98
	11-nt		F(1)	All possible dis...	3284.17
	12-nt		G(1)	All possible dis...	3588.37
	13-nt		H(1)	All possible dis...	3892.56
	14-nt		I(1)	All possible dis...	4196.75

Add Remove selected

Consider clipped species

Mass computation options

- Change N-terminal Q to pyroGlu
- Clip off C-terminal K
- N-glycans removed by PNGase F (N-X-S/T --> D)

Save Preset... Load Preset... Close

Define delta masses for impurities

New Reference project (read only)

Sample input Protein input Sample-protein input Deconvolution Mass matching Advanced

Delta masses

Text filter

Enable auto assign	Mass name	Mass (Dalton)
<input type="checkbox"/>	Man5	1217.2
<input type="checkbox"/>	G0F-GlcNAc	1242.2
<input type="checkbox"/>	G0	1299.3
<input type="checkbox"/>	G0F-Lys	1317.2
<input type="checkbox"/>	Man5F	1363.3
<input type="checkbox"/>	G0F	1445.4
<input type="checkbox"/>	G0F+Lys	1573.6
<input type="checkbox"/>	G1F	1607.6
<input type="checkbox"/>	G2F	1769.7
<input type="checkbox"/>	G1F+NeuAc	1898.8
<input type="checkbox"/>	G2F+NeuAc	2061
<input type="checkbox"/>	G2F+2NeuAc	2352.2
<input type="checkbox"/>	G0F/G1F+NeuAc	3344.2
<input type="checkbox"/>	G1F/G1F+NeuAc	3506.4
<input type="checkbox"/>	G1F/G2F+NeuAc	3668.5
<input type="checkbox"/>	G1F/G2F+2NeuAc	3959.8
<input type="checkbox"/>	G2F/G2F+2NeuAc	4121.9
<input type="checkbox"/>	G2F/G2F+3NeuAc	4413.2
<input type="checkbox"/>	Hexose	162
<input checked="" type="checkbox"/>	+Phosphate	80
<input checked="" type="checkbox"/>	+HFIP	168

Match tolerance: 1 Dalton

Import... Add row Remove rows

Mass assignments

No charge deconvolution
 Auto charge deconvolution
 Auto charge deconvolution and mass assignments

Mass area and relative intensity options

Mass area

Compute areas of mass peaks Width: 500 Da

Relative mass intensity

Report intensities relative to local base peak

Window for local base peak: 20%

Minimum % of local base peak: 10%

Save Preset... Load Preset... Close

Automatic elution peak definition and mass assignment

Project (double click to dock / undock)

R project

Peak #	Start time	End time	Apex time	Retention time	Normed area %	Mass	Intensity
> 13 (5)	3.32	3.58	3.42	1.8e+4	0.95 %	1463.87 ...	3.140e+...
> 14 (5)	3.77	4.17	3.88	3.5e+5	19.08 %	4500.76 ...	9.013e+...
> 15 (5)	4.30	4.57	4.42	8.3e+3	0.45 %	1463.87 ...	2.204e+...

Masses

Mass	Intensity	Name	Delta mass from most intense	Mass monoisotope
4500.76	7.67e+5	15-nt, _	0.00	4498.7930
4522.75	8.01e+4	4523	21.99	4520.7926
4540.74	1.79e+4	4541	39.98	4536.7006
4668.76	6.29e+4	15-nt, +HFIP	168.00	4666.7974

Chromatogram (double click to dock / undock)

MS1 (double click to dock / undock)

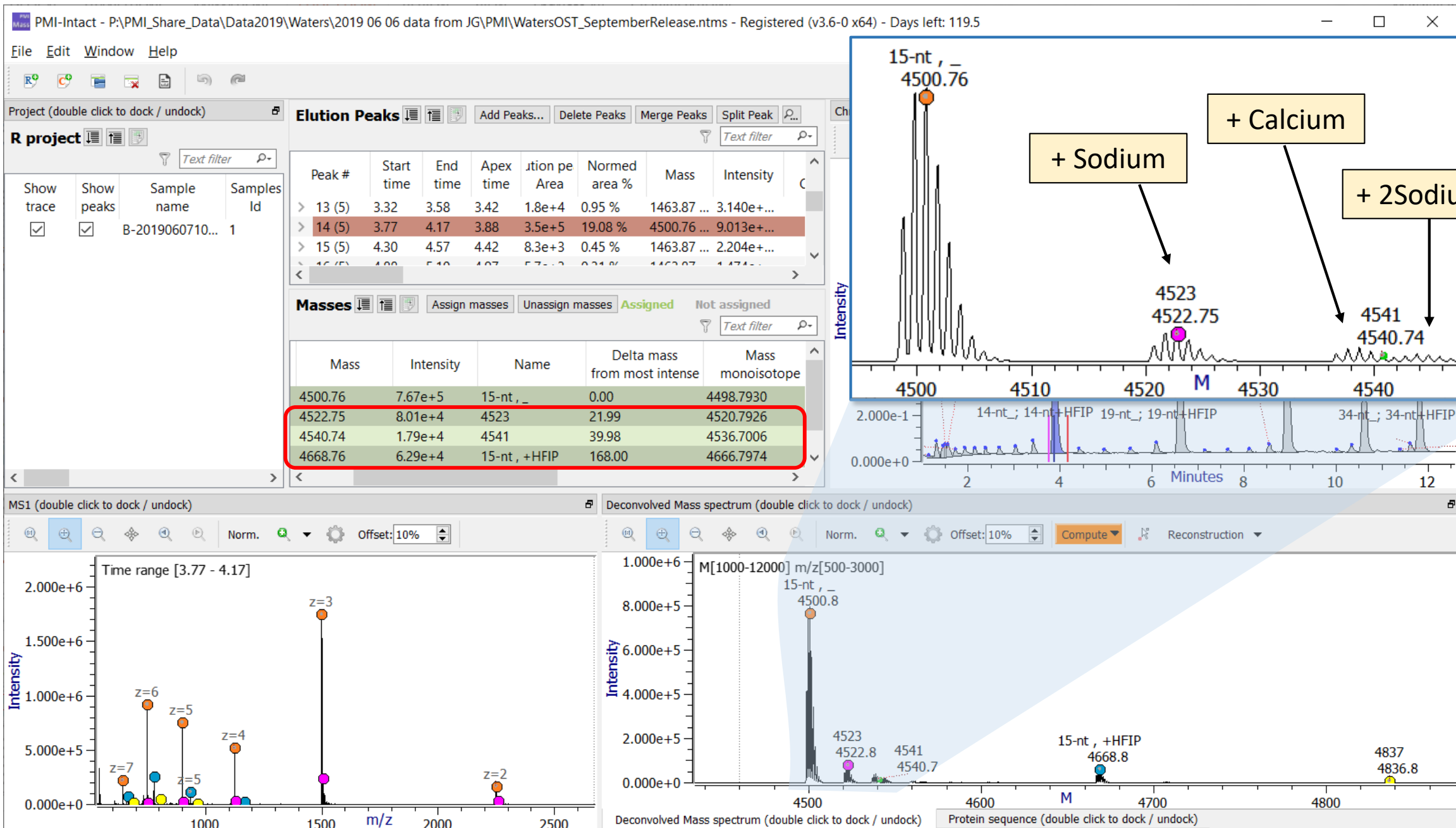
Time range [3.77 - 4.17]

Deconvolved Mass spectrum (double click to dock / undock)

HFIP = hexafluoroisopropanol

Summed MS1

Deconvolved masses



The screenshot shows a software interface with a 'File' menu open. The 'Export' option is selected, and a sub-menu is visible with the following options:

- Report...
- Generate Multi-document Report...
- Generate MS path template CSV...

The background interface includes a 'Peak Peaks' table and a 'Masses' table.

#	Start time	End time	Apex time	Retention time	Area	Normed area %
1	3.11	3.12	3.115	3.115	1.5e+5	0.15

Assigned	Mass Id	Mass	Intensity	...
<input checked="" type="checkbox"/>	226	7543.30	2.83e+5	25...

One-Click Reporting

PMI-Intact - P:\PMI_Share_Data\Data2019\Waters\2019 06 06 data from JG\PMI\WatersOST_Sep

File Edit Window Help

- New Reference Project... Ctrl+N
- New Comparison Project... Ctrl+Shift+N
- Open Project... Ctrl+O
- Recent Projects
- Close Project Ctrl+W
- Export**
 - Report...**
 - Generate Multi-document Report...
 - Generate MS path template CSV...
- Save a Copy As... Ctrl+S
- Exit Ctrl+Q

Peaks

#	Start time	End time	Apex time	Retention time	Area	Normalized area %
226					7543.30	2.83e+5

Masses

Assigned	Mass Id	Mass	Intensity
<input checked="" type="checkbox"/>	226	7543.30	2.83e+5

One-Click Reporting

Intact_Mass_Pivot_Report_withPlots (ReferenceDefault).rptc

File Edit Tabs

Show configuration Copy pivot table to clipboard Find: Enter text

Summary * Expected Mass - Intensity * Expected Mass - Relative Intensity (%) * All Mass - Intensity

Protein name

Name

Sample name ← B-20190607101208 (%)

Name ↓	Sample name ←	B-20190607101208 (%)
35-nt, _		5.77
35-nt, -HFIP		1.89
34-nt, _		0.26
34-nt, -HFIP		0.06
33-nt, _		0.08
30-nt, _		7.65
30-nt, -HFIP		1.84
29-nt, _		0.41
29-nt, -HFIP		0.10
28-nt, _		0.15
28-nt, -HFIP		0.05
25-nt, _		11.59
25-nt, -HFIP		2.07
24-nt, _		0.51
24-nt, -HFIP		0.14
23-nt, _		0.23
22-nt, _		0.22
21-nt, _		0.22
20-nt, _		18.32
20-nt, -HFIP		2.40
19-nt, _		1.10
19-nt, -HFIP		0.15
18-nt, -HFIP		0.08
17-nt, _		0.50
17-nt, -HFIP		0.06
16-nt, _		0.62
16-nt, -HFIP		0.09
15-nt, _		31.40
15-nt, -HFIP		2.57
14-nt, _		1.61
14-nt, -Phosphate		0.12
14-nt, -HFIP		0.19

_cand_src_t
 _pks_c_id
 _pks_id
 _plt_id
 _sr_idx
 Apex time
 Apex time original
 Delta mass from calc.
 Delta Mass ppm
 Delta name
 Elution peak Area
 End time
 End time Original
 Expected mass
 Expected type
 Glycans
 Intensity
 Local Rel. Int.

File Edit Window Help

- New Reference Project... Ctrl+N
- New Comparison Project... Ctrl+Shift+N
- Open Project... Ctrl+O
- Recent Projects
- Close Project Ctrl+W
- Export**
- Save a Copy As... Ctrl+S
- Exit Ctrl+Q

Peak Peaks Add Peaks... Delete Peaks

#	Start time	End time	Apex time	Integration Area	Normalized area %
1	3.0	3.1	3.05	100	100

- Report...**
- Generate Multi-document Report...
- Generate MS path template CSV...

Masses Assign masses Unassign masses

Assigned	Mass Id
<input checked="" type="checkbox"/>	226

File Edit Tabs

Show configuration Copy pivot table to clipboard Find: Enter text

Summary * Expected Mass - Intensity * Expected Mass - Relative Intensity (%) * All Mass - Intensity

Protein name

Name

_cand_src_t

_pks_c_id

_pks_id

Name ↓	Sample name ←	B-20190607101208 (%)
35-nt, _		5.77
35-nt, +HFIP		1.89
34-nt, _		0.26
34-nt, +HFIP		0.06
33-nt, _		0.08
30-nt, _		7.65
30-nt, +HFIP		1.84
29-nt, _		0.41
29-nt, +HFIP		0.10
28-nt, _		0.15
28-nt, +HFIP		0.05
25-nt, _		11.59
25-nt, +HFIP		2.07
24-nt, _		0.51
24-nt, +HFIP		0.14
23-nt, _		0.23

Name ↓	Sample name ←	B-20190607101208 (%)
35-nt, _		5.77
35-nt, +HFIP		1.89
34-nt, _		0.26
34-nt, +HFIP		0.06
33-nt, _		0.08
30-nt, _		7.65
30-nt, +HFIP		1.84
29-nt, _		0.41
29-nt, +HFIP		0.10
28-nt, _		0.15
28-nt, +HFIP		0.05
25-nt, _		11.59
25-nt, +HFIP		2.07
24-nt, _		0.51
24-nt, +HFIP		0.14
23-nt, _		0.23
22-nt, _		0.22
21-nt, _		0.22
20-nt, _		18.32
20-nt, +HFIP		2.40
19-nt, _		1.10
19-nt, +HFIP		0.15
18-nt, +HFIP		0.08
17-nt, _		0.50
17-nt, +HFIP		0.06
16-nt, _		0.62
16-nt, +HFIP		0.09
15-nt, _		31.40
15-nt, +HFIP		2.57
14-nt, _		1.61
14-nt, +Phosphate		0.12
14-nt, +HFIP		0.19

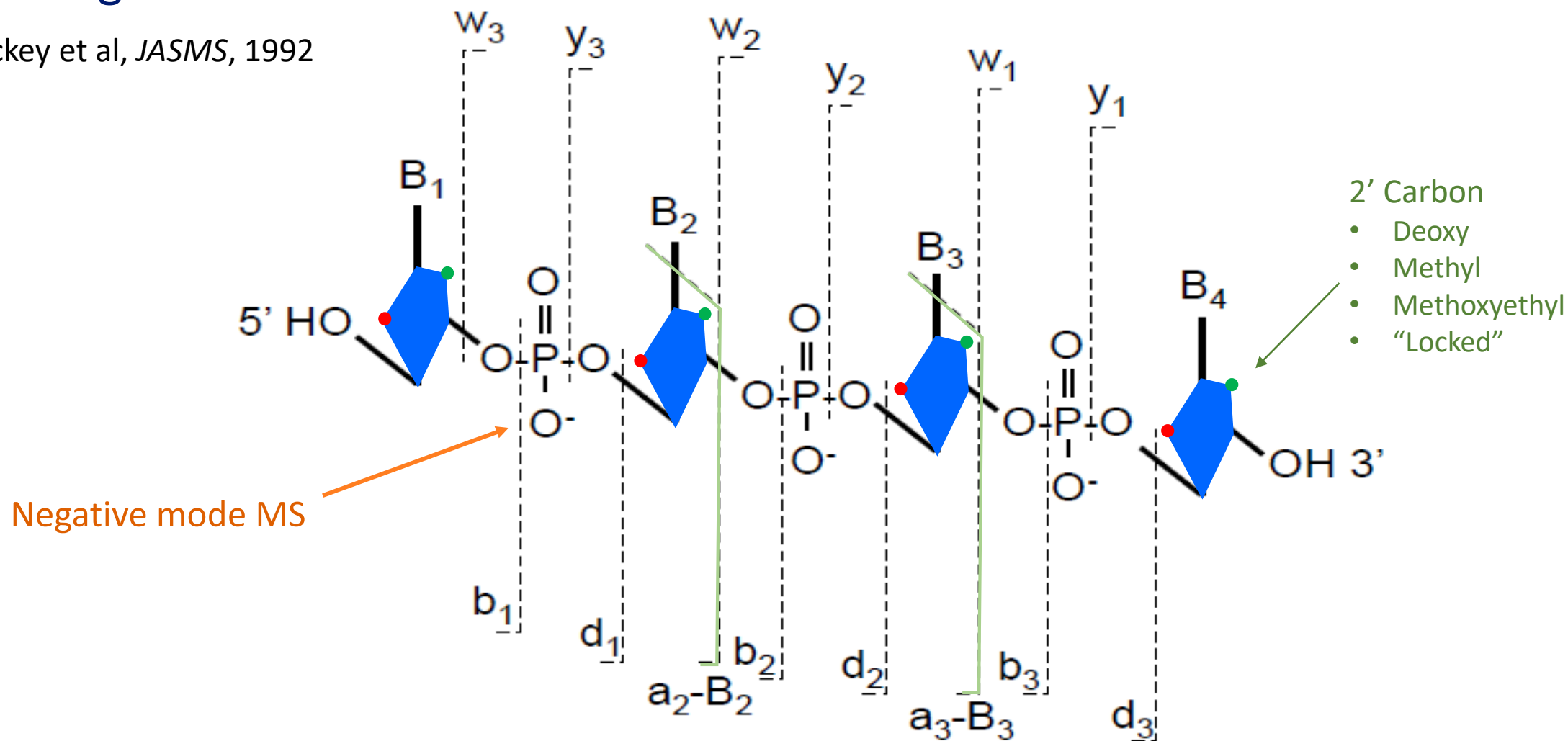
One-Click Reporting

Sequence Confirmation with MS2



Oligo Fragmentation

McLucky et al, *JASMS*, 1992



Byonic Parameters

Text commands for
experimental features

PMi-Byonic - Protein Metrics Inc. - Registered (v3.6.0) - Days left: 120.3

File Edit Help

Input files

Select MS/MS data file: P:\PMI_Share_Data\Data2019\U_Texas\Ines_Oligos\Sample1_native_Wcleanup_MS2_5-_HCD18.raw

Select Protein database file: P:\PMI_Share_Data\Data2019\U_Texas\Ines_Oligos\InesRNA.fasta

Protein database options: Add decoys: Add common contaminants:

Output folder

Output folder...: C:\data_results Results folder name: [sp...]

Options

Digestion and Instrument Parameters Modifications Glycans S-S, Xlink Inclusion Advanced Pro...

Fixed and Variable modifications

Enter/edit... Total common max: 1 Total rare max: 1

% Custom modification text below
cleavage_flags=0
machine_type=10

Wildcard search: Disabled
Minimum mass: -160
Maximum mass: 100
Restrict to residues:

Load parameters... Save parameters... Reset parameters

Run

Multicore options
Multicore: Normal

Byonic™ by Protein Metrics Inc.
© 2011-2019
www.proteinmetrics.com

PROTEIN METRICS

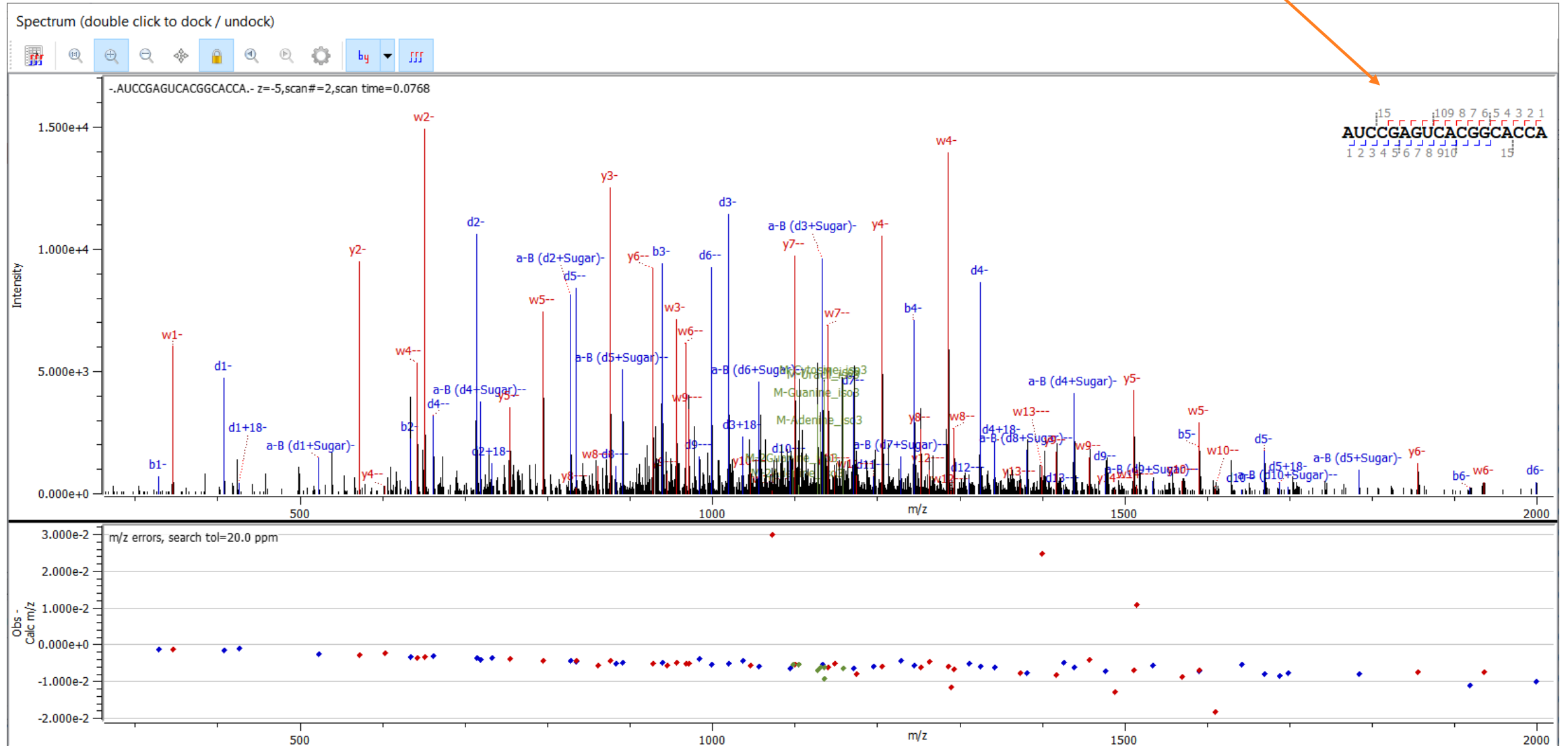
HSS3.new.fasta x InesRNA.fasta x InesRNA.fasta x

```
1  
2 >Ines RNA with phospho 5' terminus  
3 AUCCGAGUCACGGCACCA
```

Synthetic RNA with HCD fragmentation

(Brodbelt Lab, U. Texas)

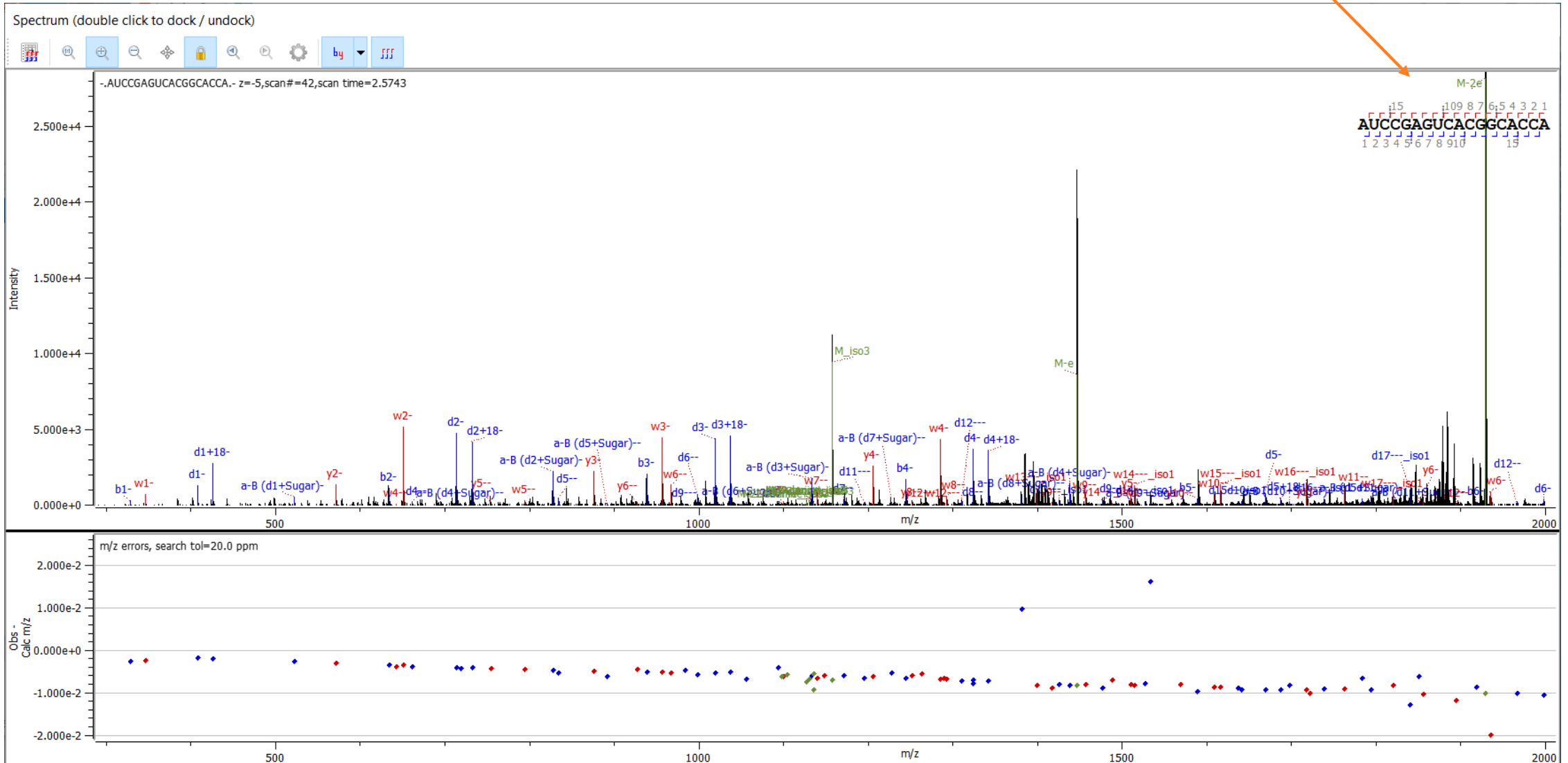
100% Sequence Coverage



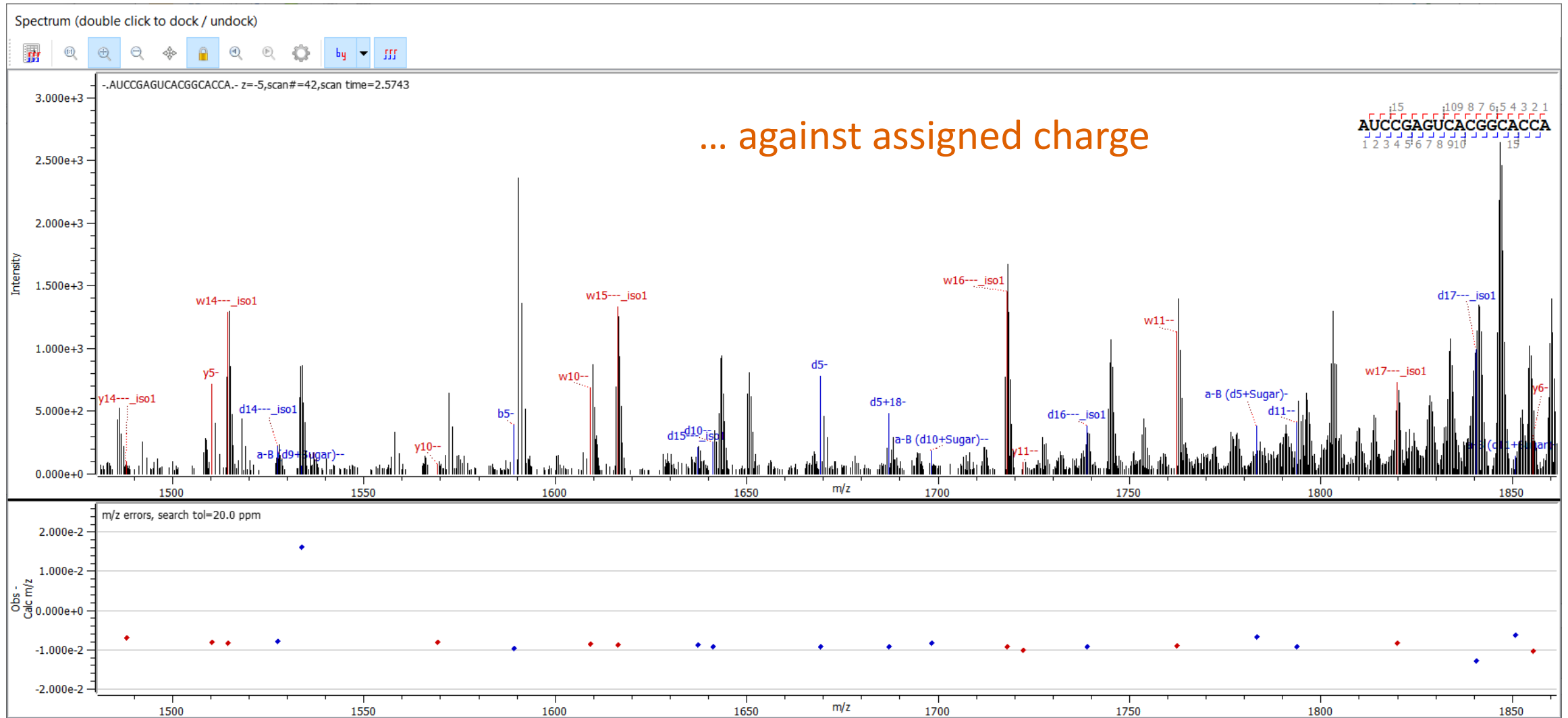
UVPD fragmentation

(Brodbelt Lab, U. Texas)

100% from both ends !




Byonic annotations for high-res MS2 check isotope peak spacing ...



Future Work

- Oligo-centric GUI
- Oligo-centric reporting
- Support for IDT codes
- Mixed protein / oligo (Crosslinked?!)

 INTEGRATED DNA TECHNOLOGIES custom oligos • qPCR • next generation sequencing • RNAi • genes & gene fragments • CRISPR genome editing 		
INSTRUCTIONS		
Scale of Synthesis	Lengths	Base Notations
25 nmole	15-60 Bases	Phosphorothioated DNA = A*, G*, C*, T* RNA = rA, rG, rC, rU Phosphorothioated RNA = rA*, rG*, rC*, rU* 2'O-Methyl RNA = mA, mG, mC, mU Phosphorothioated 2'O-Methyl RNA = mA*, mG*, mC*, mU* Locked Nucleic Acid (LNA) = +A, +G, +C, +T (available on dual Mixed Bases = Please enter bases in UPPERCASE
100 nmole	10-90 Bases	
250 nmole	5-100 Bases	
1 µmole	5-100 Bases	
5 µmole	5-100 Bases	
10 µmole	5-100 Bases	
4 nmole Ultramer	45-200 Bases	
20 nmole Ultramer	45-200 Bases	
PAGE Ultramer	60-200 Bases	
Purification Description		
Standard Desalt		R= A,G
PAGE		Y= C,T
HPLC		M= A,C
IE HPLC		K= G,T
RNase Free HPLC		S= C,G
*PAGE, HPLC, IE-HPLC, and RNase-Free HPLC available on the 100 nm scale and higher		W= A,T
Standard Desalt available on all scales		H= A,C,T
Available Modifications		B= C,G,T
Click here for Modification Codes		V= A,C,G
Please insert modifications within sequence		D= A,G,T
Tube Templates		N= A,C,G,T
		Custom Mixed Base Instructions
		Enter the desired percentage of each base (Integers Only, Totaling 100%).



Thank You !



PROTEIN METRICS

For a **live demo** contact info@proteinmetrics.com

We're hiring!