Supporting Information

N-Substituted Valiolamine Derivatives as Potent Inhibitors of Endoplasmic Reticulum α-Glucosidases I and II with Antiviral Activity

Sharanbasappa S. Karade[1,2], Michelle L. Hill[3], J. L. Kiappes[4], Rajkumar Manne[5], Balakishan Aakula[5], Nicole Zitzmann[3], Kelly L. Warfield[6], Anthony M. Treston[6], and Roy A. Mariuzza[1,2]*

[1]University of Maryland Institute for Bioscience and Biotechnology Research, Rockville, MD 20850, USA

[2]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA

[3]Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

[4]Department of Chemistry, University College, London WC1H 0AJ, UK

[5]Sai Life Sciences Ltd., 500032 Telangana, India

[6]Emergent BioSolutions, Gaithersburg, MD 20879, USA

*Correspondence: Roy A. Mariuzza (rmariuzz@umd.edu)

Supplementary Table S1

Analytical Data

Supporting Information

Supplementary Table S1. Data Collection and Structure Refinement Statistics

	α-GluII+EB-0149	α-GluII+EB-0150	α-GluII+EB-0151
PDB accession code	7JTY	7KRY	7K9N
Data collection			
Resolution range (Å)[a]	39.9–2.21 (2.29–2.21)	47.3–2.55 (2.64–2.55)	41.8–2.07 (2.15–2.07)
Space group	P 32	P 32	P 32
Unit cell parameters	102.56 Å, 102.56 Å, 239.56 Å, 90°, 90°, 120°	102.92 Å, 102.92 Å, 240.21 Å 90°, 90°, 120°	102.78 Å, 102.78 Å, 240.90 Å 90°, 90°, 120°
Total reflections[a]	1183616 (100478)	702688 (63294)	1360283 (120629)
Unique reflections[a]	140777 (14113)	92459 (9308)	172944 (17243)
Multiplicity[a]	8.4 (7.1)	7.6 (6.8)	7.9 (7.0)
Completeness (%)[a]	99.97 (99.97	99.96 (99.89)	99.86 (99.30)
Mean $I/\sigma(I)$[a]	10.38 (1.31)	11.21 (1.11)	10.10 (1.11)
Wilson *B* factor (Å^2)	38.19	49.48	35.16
R_{merge}[a,b]	0.1091 (0.8133)	0.166 (1.236)	0.1132 (1.071)
$CC_{1/2}$[a]	0.996 (0.695)	0.989 (0.553)	0.997 (0.603)
Refinement			
Resolution range (Å)	39.9–2.21	47.3–2.55	41.8–2.07
Reflections used in refinement[a]	140769 (14113)	92448 (9307)	172904 (17242)
R_{work}[a,c]	0.1607 (0.2421)	0.1902 (0.2932)	0.1699 (0.2537)
R_{free}[a,c]	0.1890 (0.2676)	0.2154 (0.3107)	0.1952 (0.2746)
No. of protein atoms	14950	14912	15002
No. of ligand atoms	114	443	339
No. of waters	855	438	1126
Protein residues	1880	1880	1887
r.m.s.d. from ideality			
Bond lengths (Å)	0.008	0.003	0.006
Bond angles (°)	1.28	0.71	0.88
Ramachandran plot statistics			
Favored (%)	97.31	96.56	96.84
Allowed (%)	2.58	3.44	3.11
Disallowed (%)	0.11	0.00	0.05
Rotamer outliers (%)	2.36	2.43	1.55
Clashscore	3.24	4.78	3.04
Average *B* factor (Å^2)	45.55	52.66	40.65
Protein	45.24	52.49	39.98
Ligand	64.53	61.54	55.80
Waters	48.54	49.40	45.04

[a]Values in parentheses correspond to the highest resolution shell.

[b]$R_{merge} = \sum|I_j - \langle I \rangle|/\sum I_j$, where I_j is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

[c]R_{work} (R_{free}) = $\sum||F_o| - |F_c||/\sum|F_o|$; 5.0% of data were used for R_{free}.

Supplementary Table S1 (cont.). Data Collection and Structure Refinement Statistics

	α-GluII+EB-0152	α-GluII+EB-0155	α-GluII+EB-0156
PDB accession code	7K9O	7K9Q	7K9T
Data collection			
Resolution range (Å)[a]	39.9–2.30 (2.38–2.30)	42.5–2.30 (2.39–2.30)	42.4–2.10 (2.18–2.10)
Space group	P 32	P 32	P 32
Unit cell parameters	102.7 Å, 102.7 Å, 239.6 Å 90°, 90°, 120°	102.9 Å, 102.9 Å, 241.7 Å 90°, 90°, 120°	102.9 Å, 102.9 Å, 240.7 Å 90°, 90°, 120°
Total reflections[a]	898536 (69790)	1048755 (91812)	1357526 (111691)
Unique reflections[a]	125193 (12513)	126356 (12577)	165417 (16531)
Multiplicity[a]	7.2 (5.6)	8.3 (7.3)	8.2 (6.8)
Completeness (%)[a]	99.89 (99.17)	99.86 (99.31)	99.93 (99.73)
Mean $I/\sigma(I)$[a]	8.09 (1.02)	14.7 (1.0)	9.53 (1.06)
Wilson B factor (Å^2)	45.37	41.69	35.78
R_{merge}[a,b]	0.1386 (0.9136)	0.131 (1.105)	0.1298 (1.137)
$CC_{1/2}$[a]	0.992 (0.445)	0.992 (0.608)	0.996 (0.579)
Refinement			
Resolution range (Å)	39.9–2.30	42.5–2.30	42.4–2.10
Reflections used in refinement[a]	125188 (12512)	126344 (12577)	165373 (16527)
R_{work}[a,c]	0.2099 (0.2954)	0.1720 (0.2563)	0.1706 (0.2640)
R_{free}[a,c]	0.2383 (0.3050)	0.1976 (0.2866)	0.1915 (0.2999)
No. of protein atoms	15077	15086	15052
No. of ligand atoms	243	315	421
No. of waters	623	873	983
Protein residues	1882	1939	1881
r.m.s.d. from ideality			
Bond lengths (Å)	0.004	0.005	0.183
Bond angles (°)	1.07	1.12	2.88
Ramachandran plot statistics			
Favored (%)	96.62	96.52	96.78
Allowed (%)	3.38	3.32	3.22
Disallowed (%)	0.00	0.16	0.00
Rotamer outliers (%)	0.79	2.02	0.91
Clashscore	5.58	4.58	4.13
Average B factor (Å^2)	48.93	45.68	43.89
Protein	48.83	45.37	43.07
Ligand	54.49	57.14	65.54
Waters	49.16	46.84	47.16

[a]Values in parentheses correspond to the highest resolution shell.

[b]$R_{merge} = \sum|I_j - \langle I \rangle|/\sum I_j$, where I_j is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

[c]R_{work} (R_{free}) $= \sum||F_o| - |F_c||/\sum|F_o|$; 5.0% of data were used for R_{free}.

Supplementary Table S1 (cont.). Data Collection and Structure Refinement Statistics

	α-GluII+EB-0159	α-GluII+EB-0176	α-GluII+EB-0281
PDB accession code	7KAD	7KB6	7KB8
Data collection			
Resolution range (Å)[a]	42.3–2.51 (2.60–2.51)	44.6–2.20 (2.28–2.20)	40.0–2.39 (2.47–2.39)
Space group	P 32	P 32	P 32
Unit cell parameters	102.7 Å, 102.7 Å, 240.3 Å 90°, 90°, 120°	102.9 Å, 102.9 Å, 239.8 Å 90°, 90°, 120°	102.8 Å, 102.8 Å, 240.1 Å 90°, 90°, 120°
Total reflections[a]	818543 (81824)	896581 (77846)	741789 (57920)
Unique reflections[a]	97233 (9617)	140562 (13540)	112724 (11064)
Multiplicity[a]	8.4 (8.5)	6.4 (5.7)	6.6 (5.2)
Completeness (%)[a]	99.92 (100.00)	97.47 (93.75)	99.72 (97.93)
Mean $I/\sigma(I)$[a]	8.42 (1.03)	9.20 (1.13)	8.42 (0.88)
Wilson B factor (Å^2)	48.51	38.09	48.18
R_{merge}[a,b]	0.1507 (1.122)	0.1286 (0.9757)	0.1151 (0.9231)
$CC_{1/2}$[a]	0.992 (0.613)	0.993 (0.654)	0.992 (0.537)
Refinement			
Resolution range (Å)	42.3–2.51	44.6–2.20	40.0–2.39
Reflections used in refinement[a]	97204 (9619)	140521 (13536)	112697 (11059)
R_{work}[a,c]	0.1735 (0.2415)	0.1663 (0.2519)	0.1735 (0.2619)
R_{free}[a,c]	0.2033 (0.3054)	0.1932 (0.3217)	0.2103 (0.3007)
No. of protein atoms	15057	14896	15052
No. of ligand atoms	433	518	325
No. of waters	453	856	592
Protein residues	1889	1877	1893
r.m.s.d. from ideality			
Bond lengths (Å)	0.005	0.005	0.007
Bond angles (°)	1.05	0.94	1.25
Ramachandran plot statistics			
Favored (%)	96.52	96.28	96.15
Allowed (%)	3.48	3.56	3.80
Disallowed (%)	0.00	0.16	0.05
Rotamer outliers (%)	0.98	2.25	1.72
Clashscore	4.53	4.49	5.03
Average B factor (Å^2)	51.00	44.36	50.78
Protein	50.47	43.50	50.55
Ligand	71.45	66.16	63.88
Waters	49.10	46.28	49.54

[a]Values in parentheses correspond to the highest resolution shell.

[b]$R_{merge} = \sum|I_j - \langle I \rangle|/\sum I_j$, where I_j is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

[c]R_{work} (R_{free}) $= \sum||F_o| - |F_c||/\sum|F_o|$; 5.0% of data were used for R_{free}.

Supplementary Table S1 (cont.). Data Collection and Structure Refinement Statistics

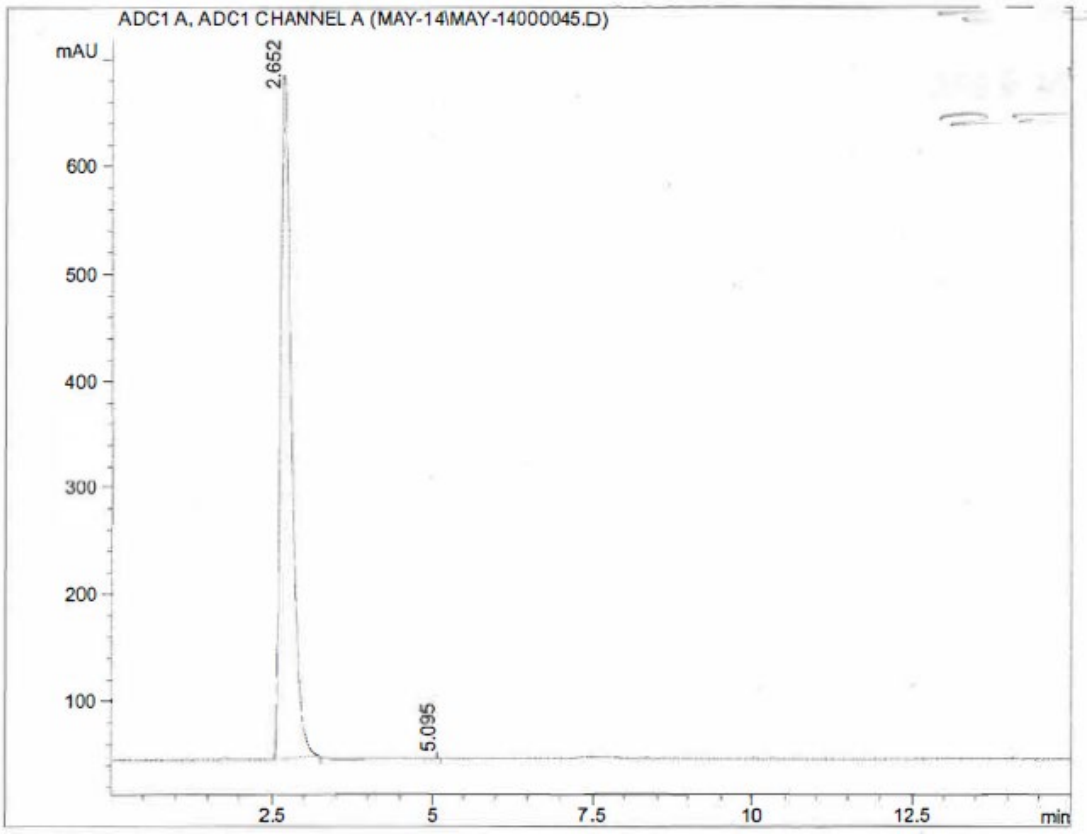
	α-GluII+EB-0283	α-GluII+EB-0288	apo α-GluII
PDB accession code	7KBJ	7KBR	7L9E
Data collection			
Resolution range (Å)[a]	42.1–2.21 (2.29–2.21)	42.3–2.09 (2.17–2.09)	44.6–2.29 (2.37–2.29)
Space group	P 32	P 32	P 32
Unit cell parameters	102.7 Å, 102.7 Å, 238.9 Å 90°, 90°, 120°	102.8 Å, 102.8 Å, 240.6 Å 90°, 90°, 120°,	103.0 Å, 103.0 Å, 240.6 Å 90°, 90°, 120°
Total reflections[a]	1266416 (53087)	1383828 (116777)	358907 (34287)
Unique reflections[a]	140010 (13256)	166814 (16182)	128181 (12699)
Multiplicity[a]	9.0 (4.0)	8.3 (7.2)	2.8 (2.7)
Completeness (%)[a]	99.29 (93.27)	99.52 (97.06)	99.85 (99.73)
Mean $I/\sigma(I)$[a]	8.33 (0.94)	10.57 (1.22)	8.2 (1.77)
Wilson *B* factor (Å^2)	37.17	32.55	44.16
R_{merge}[a,b]	0.1655 (0.8132)	0.1313 (0.9599)	0.125 (0.547)
$CC_{1/2}$[a]	0.994 (0.395)	0.993 (0.683)	0.967 (0.649)
Refinement			
Resolution range (Å)	42.1–2.21	42.3–2.09	44.6–2.29
Reflections used in refinement[a]	139911 (13176)	166559 (16180)	128171 (12703)
R_{work}[a,c]	0.1670 (0.2784)	0.1665 (0.2694)	0.1815 (0.2371)
R_{free}[a,c]	0.1997 (0.3139)	0.1889 (0.3294)	0.2080 (0.2569)
No. of protein atoms	15123	15064	14955
No. of ligand atoms	399	423	137
No. of waters	1004	1290	539
Protein residues	1880	1876	1882
r.m.s.d. from ideality			
Bond lengths (Å)	0.005	0.004	0.010
Bond angles (°)	0.85	0.75	1.14
Ramachandran plot statistics			
Favored (%)	96.67	97.20	97.58
Allowed (%)	3.33	2.80	2.31
Disallowed (%)	0.00	0.00	0.11
Rotamer outliers (%)	0.91	1.35	2.92
Clashscore	4.17	3.67	4.99
Average *B* factor (Å^2)	42.37	37.51	51.88
Protein	41.77	36.38	51.91
Ligand	59.03	58.62	61.04
Waters	44.85	43.80	48.81

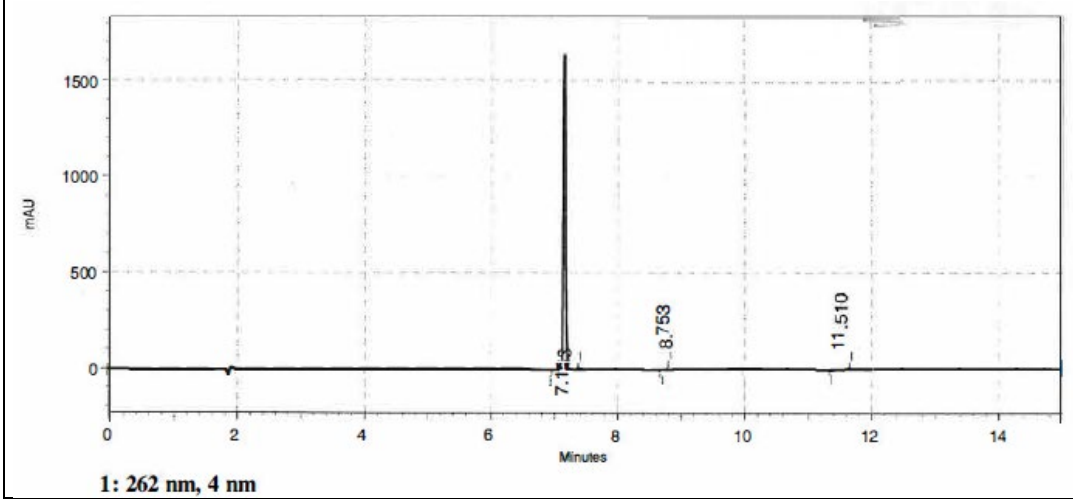
[a]Values in parentheses correspond to the highest resolution shell.

[b]$R_{merge} = \sum|I_j - \langle I \rangle|/\sum I_j$, where I_j is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

[c]R_{work} (R_{free}) $= \sum||F_o| - |F_c||/\sum|F_o|$; 5.0% of data were used for R_{free}.

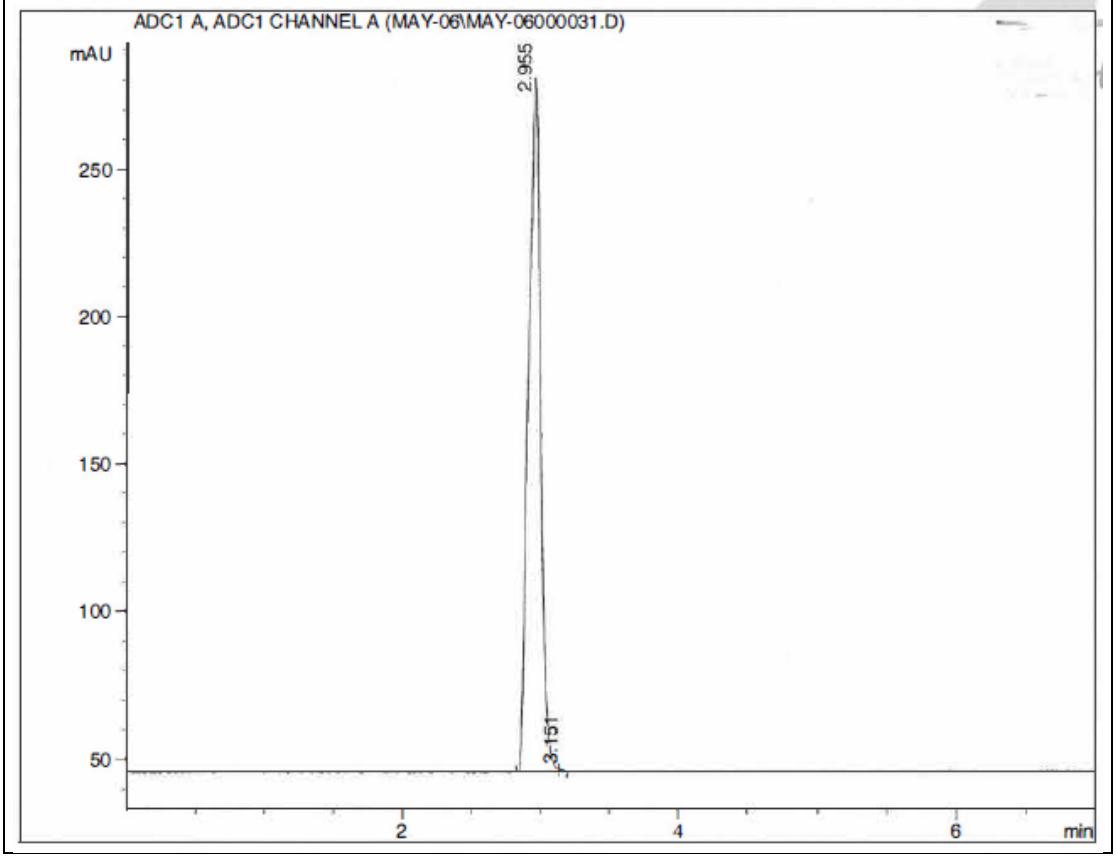
Analytical data

Calculated Mass	Observed Mass	Proton NMR (MHz, Solvent, and Peak List)
EB-0155: O[C@@H]1[C@](CO)(O)C[C@H](N)[C@H](O)[C@H]1O (chromatogram not shown for commercial substance)		
193.2	194.2(M+H)	**(500 MHz, CD3OD):** δ 3.79 (t, 1H), 3.56 (d, 1H), 3.43 (d, 1H), 3.39-3.37 (m, 1.5H), 3.36 -3.34 (m, 0.5), 3.33-3.31 (m, 1H), 2.83 (dd, 1H), 2.73 (dd, 1H)
EB-0149: O[C@@H]1[C@](CO)(O)C[C@H](NCCCC)[C@H](O)[C@H]1O		
249.3	250.2(M+H)	**(500 MHz, CD3OD):** δ 3.79-3.75 (m, 1H), 3.60-3.51 (m, 2H), 3.37-3.35 (m, 2H), 3.30-3.21 (m, 1H), 2.99-2.94 (m, 1H), 2.69-2.63 (m, 1H), 2.08-2.03 (m, 1H), 1.62-1.55 (m, 3H), 1.42-1.35 (m, 2H), 0.89 (t, 3H)
LCMS: 		

EB-0150: O[C@@H]1[C@](CO)(O)C[C@H](NCCCCCCNC2=C([N+]([O-])=O)C=C(N=[N+]=[N-])C=C2)[C@H](O)[C@H]1O		
454.4	455.4 (M+H)	**(400 MHz, CD_3OD):** δ7.80 (d, J = 2.7 Hz, 1H), 7.26 (dd, J = 9.2, 2.8 Hz, 1H), 7.09 (d, J = 9.2 Hz, 1H), 3.75-3.67 (m, 1H), 3.57-3.48 (m, 2H), 3.40 -3.37 (m, 2H), 3.36 (t, J = 1.6 Hz, 1H), 3.12-3.06 (m, 1H), 2.85-2.77 (m, 1H), 2.57 - 2.45 (m, 1H), 2.01 (dd, J = 15.0, 3.2Hz, 1H), 1.78-1.69 (m, 2H), 1.62-1.39 (m, 8H)
HPLC: 		

EB-0151: O[C@@H]1[C@](CO)(O)C[C@H](NCCCCCCCCCOC)[C@H](O)[C@H]1O		
349.4	350.3(M+H)	**(500 MHz, CD_3OD):** δ 3.75 (t, 1H), 3.59-3.51 (m, 2H), 3.39-3.35 (m, 4H), 3.34-3.31 (m, 4H), 3.10-3.05 (m,1H), 2.80-2.75 (m, 1H), 2.71-2.65 (m, 1H), 2.01 (dd, 1H), 1.49-1.41 (m, 5H), 1.39-1.31 (m, 9H)

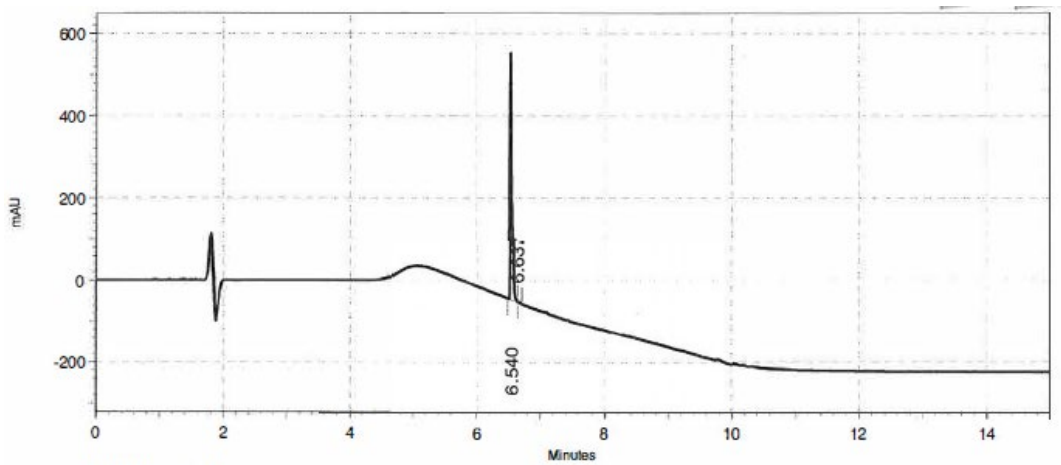
LCMS:



EB-0152: O[C@@H]1[C@](CO)(O)C[C@H](NCCCCCCNC2=C([N+]([O-])=O)C=C(N3N=NN=C3)C=C2)[C@H](O)[C@H]1O

481.5	482.4(M+H)	**(500 MHz, CD_3OD):** δ 9.75 (s, 1H), 8.62 (d, 1H), 7.99-7.90 (m, 1H), 7.25 (d, 1H), 3.79-3.75 (m, 2H), 3.59-3.49 (m, 4H), 3.39-3.35 (m, 2H), 3.10-3.05 (m,1H), 2.80-2.75 (m, 1.6H), 2.71-2.65 (m, 1.4H), 2.01 (dd, 1H), 1.82-1.75 (m, 2H), 1.61-1.43 (m, 6H)

HPLC:

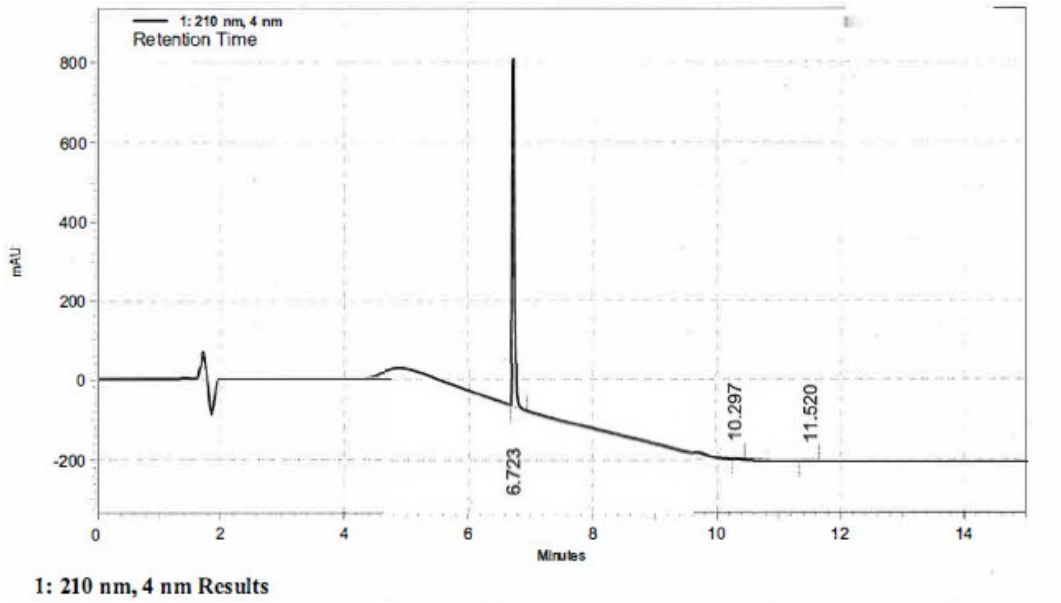


1: 210 nm, 4 nm

EB-0156: O[C@@H]1[C@](CO)(O)C[C@H](NCCCCCCNC2=C([N+]([O-])=O)C=C(N3N=CC=N3)C=C2)[C@H](O)[C@H]1O

480.5	481.4M+H)	**(500 MHz, CD_3OD):** δ 8.79 (s, 1H), 8.32-8.21 (m, 1H), 7.91 (s, 2H), 7.21 (dd, 1H), 3.75 (t, 1H), 3.59-3.51 (m, 2H), 3.39-3.35 (m, 2H), 3.34-3.31 (m, 2H), 3.10-3.05 (m, 1H), 2.85-2.80 (m, 1H), 2.59-2.51 (m 1H), 2.01 (dd, 1H), 1.82-1.75 (m, 2H), 1.61-1.42 (m, 7H)

HPLC:

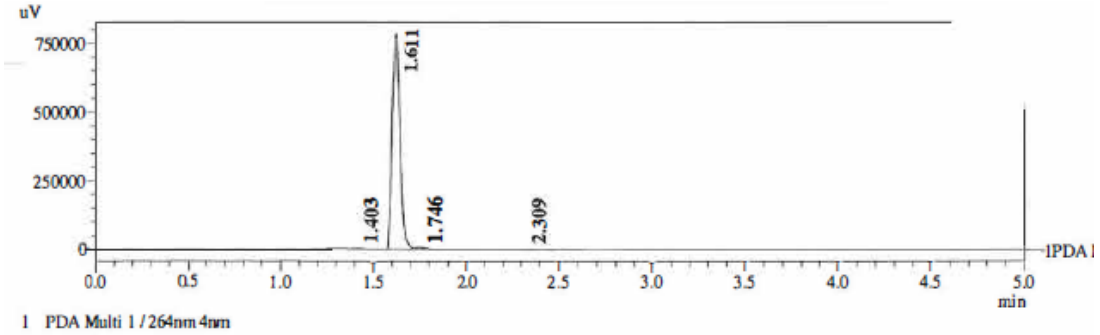


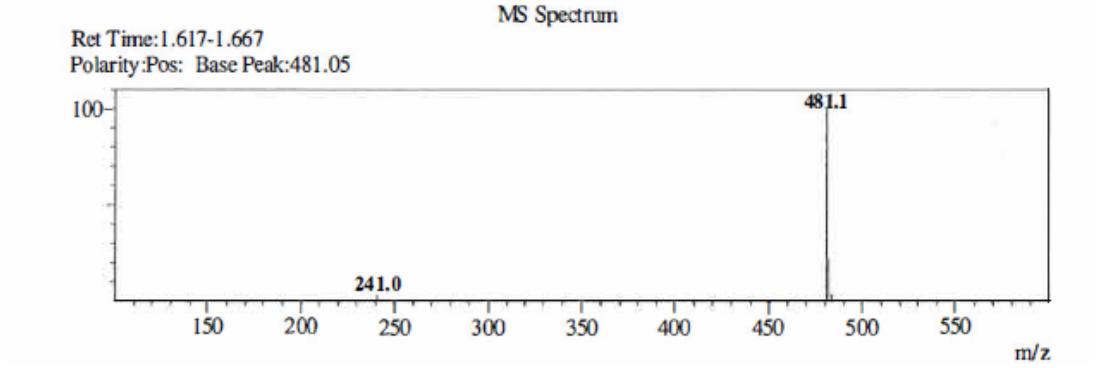
1: 210 nm, 4 nm Results

EB-0159: O[C@@H]1[C@](CO)(O)C[C@H](NCCCCCCNC2=C([N+]([O-])=O)C=C(N3N=NC=C3)C=C2)[C@H](O)[C@H]1O

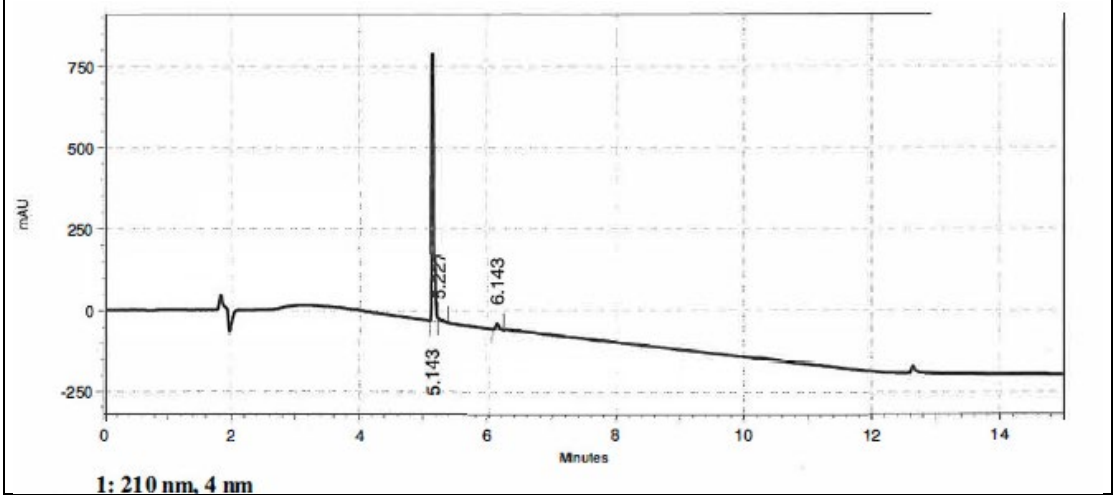
480.5	481.1(M+H)	**(400 MHz, CD_3OD):** δ 8.57 (d, *J* = 2.7 Hz, 1H), 8.49 (d, *J* = 1.0 Hz, 1H), 7.99 (dd, *J* = 9.3, 2.7 Hz, 1H), 7.88 (d, *J* = 1.1 Hz, 1H), 7.24 (d, *J* = 9.4 Hz, 1H), 3.72 (t, *J* = 9.8 Hz, 1H), 3.57-3.43 (m, 4H), 3.40-3.32 (m, 1.93H), 3.12-3.04 (m, 1H), 2.85-2.75 (m, 1.13H), 2.53-2.45 (m, 1.03H), 2.00 (dd, *J* = 14.9, 3.2 Hz, 1.02H), 1.82-1.73 (m, 1.99H), 1.60-1.42 (m, 7.12H)

LCMS:





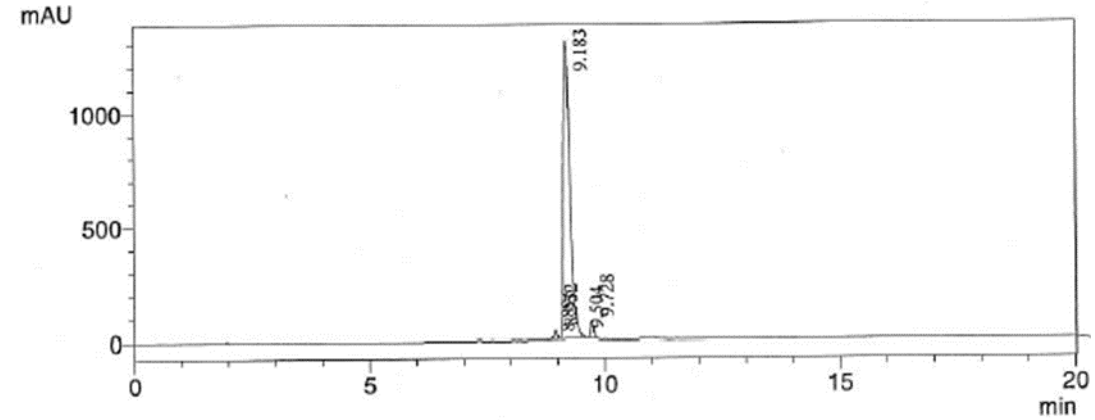
HPLC:



EB-0176: O[C@@H]1[C@@H](O)[C@H](O)[C@](CO)(O)C[C@@H]1NCCCCCCNC2=C([N+]([O-])=O)C=C(C3=NC=CC=N3)C=C2

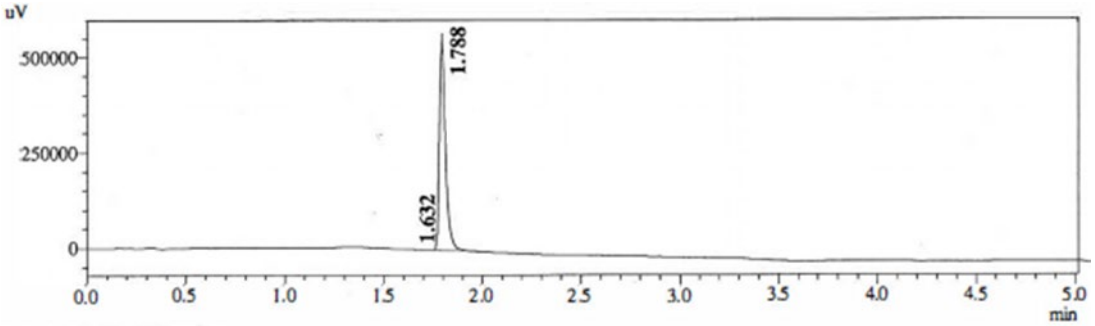
491.5	492.3(M+H)	**(400 MHz, CD_3OD)**: δ 9.23 (d, *J* = 2.1 Hz, 1H), 8.78 (d, *J* = 4.9 Hz, 2H), 8.52 (dd, *J* = 9.2, 2.1 Hz, 1H), 7.28 (t, *J* = 4.8 Hz, 1H), 7.14 (d, *J* = 9.2 Hz, 1H), 3.75-3.69 (m, 1H), 3.56-3.50 (m, 2H), 3.49-3.43 (m, 2H), 3.39-3.35 (m, 1H), 3.14-3.08 (m, 1H), 2.92-2.71 (m, 1H), 2.60-2.43 (m, 1H), 2.02 (dd, *J* = 15.0, 3.2 Hz, 1H), 1.82-1.73 (m, 2H), 1.65 - 1.38 (m, 7H)

LCMS:

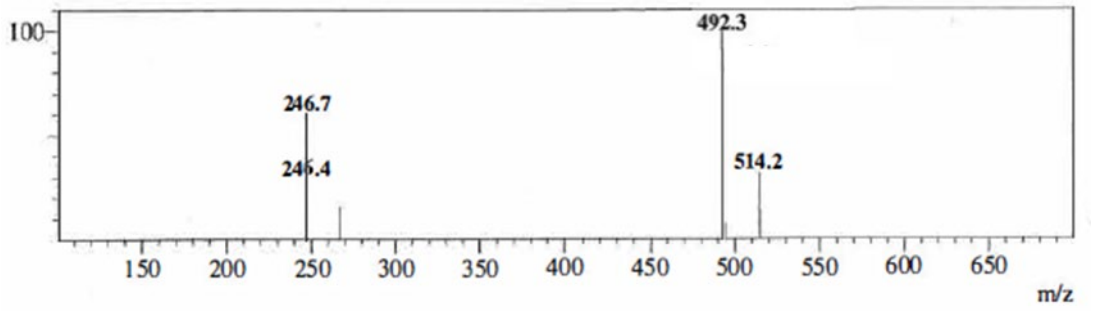


1 PDA Multi 1/200nm - 380nm 4nm

LCMS:

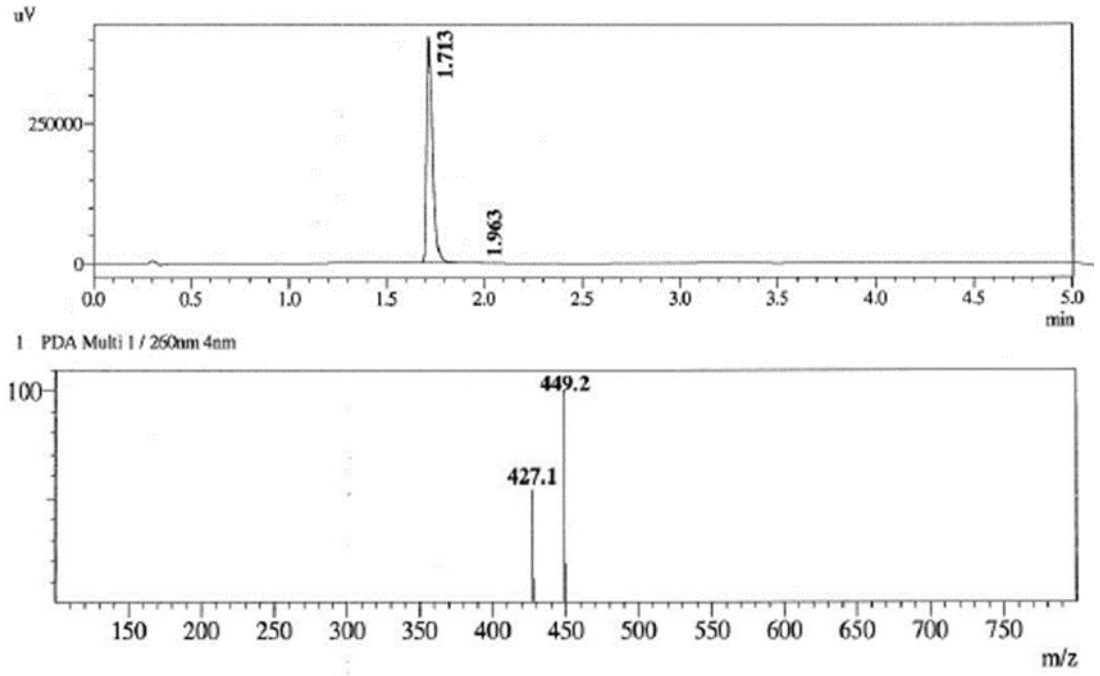


1 PDA Multi 1 / 220nm 4nm

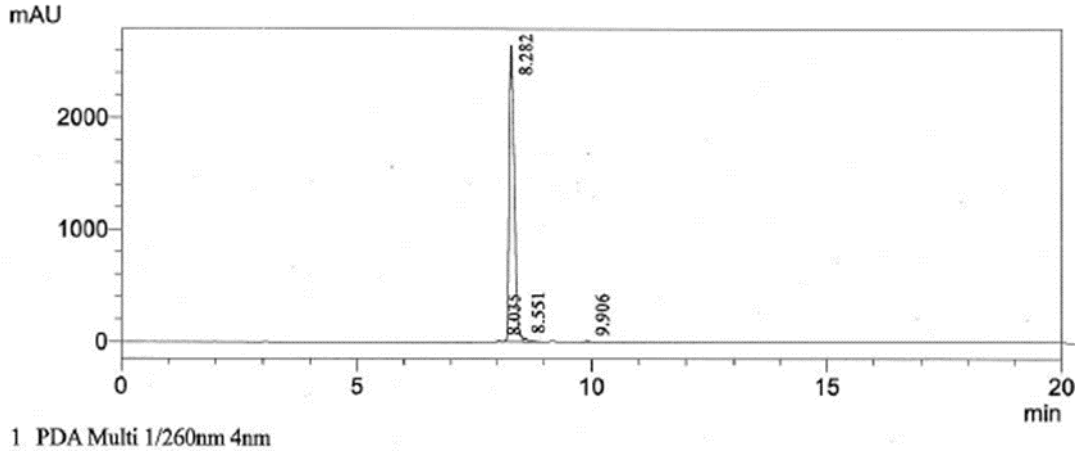


EB-0281: O[C@@H]1[C@@H](O)[C@H](O)[C@](CO)(O)C[C@@H]1NCCCCNC2=C([N+]([O-])=O)C=C(N=[N+]=[N-])C=C2		
426.4	427.1 (M+H), 449.2(M+Na)	**(400 MHz, CD_3OD):** δ 7.81 (d, J = 2.8 Hz, 1H), 7.28 (dd, J = 9.3, 2.8 Hz, 1H), 7.12 (d, J = 9.3 Hz, 1H), 3.78-3.69 (m, 1H), 3.66-3.59 (m, 1H), 3.56 (d, J = 10.9 Hz, 1H), 3.49-3.42 (m, 2H), 3.42-3.37 (m, 2H), 3.21-2.99 (m, 2H), 2.95-2.76 (m, 1H), 2.15-2.09 (m, 1H), 1.86-1.72 (m, 4H), 1.71-1.62 (m, 1H)

LCMS:



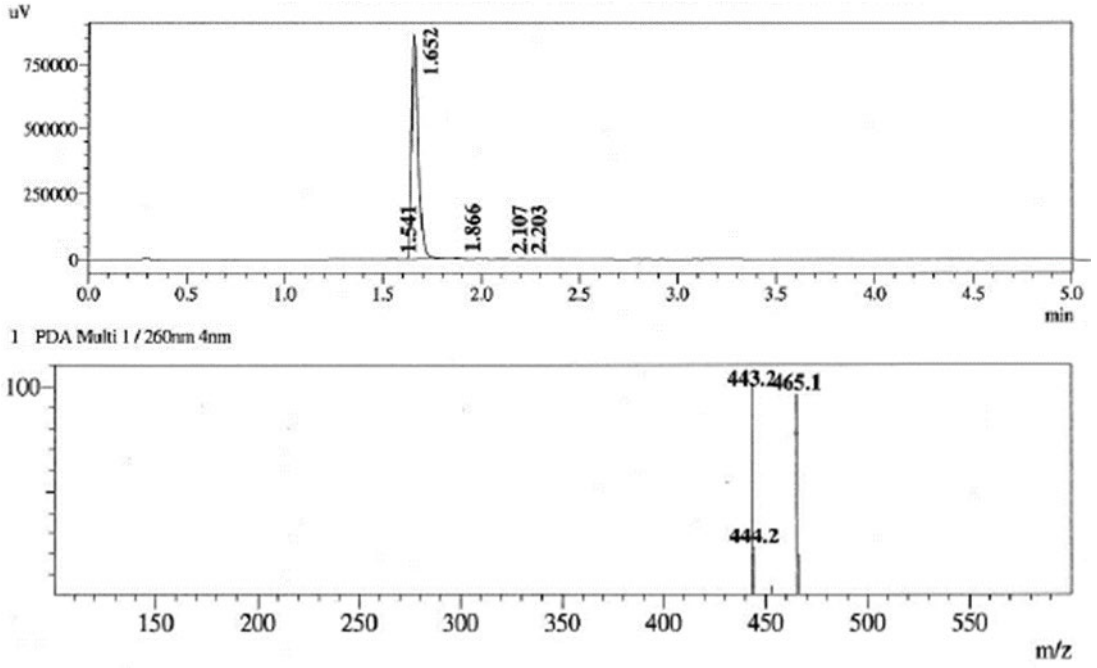
HPLC:



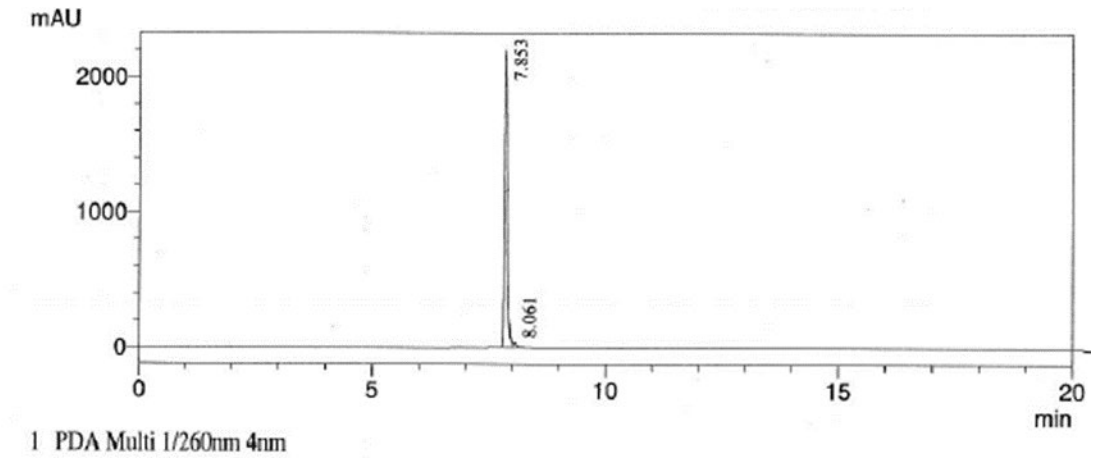
EB-0283: O[C@@H]1[C@@H](O)[C@H](O)[C@](CO)(O)C[C@@H]1NCCOCCNC2=C([N+]([O-])=O)C=C(N=[N+]=[N-])C=C2

442.4	443.2 (M+H), 465.1 (M+Na)	**(400 MHz, CD_3OD)**: δ 8.24 (br s, 1H), 7.80 (d, J = 2.6 Hz, 1H), 7.27 (dd, J = 9.3, 2.8 Hz, 1H), 7.14 (d, J = 9.3 Hz, 1H), 3.80-3.71 (m, 3H), 3.70 -3.63 (m, 1H), 3.63-3.49 (m, 5H), 3.38-3.36 (m, 1H), 3.34 (d, J = 4.3 Hz, 1H), 3.13-3.08 (m, 1H), 2.98-2.90 (m, 1H), 2.75-2.67 (m, 1H), 2.01 (dd, J = 15.1, 3.3Hz, 1H), 1.47 (dd, J = 2.7, 15.0 Hz, 1H)

LCMS



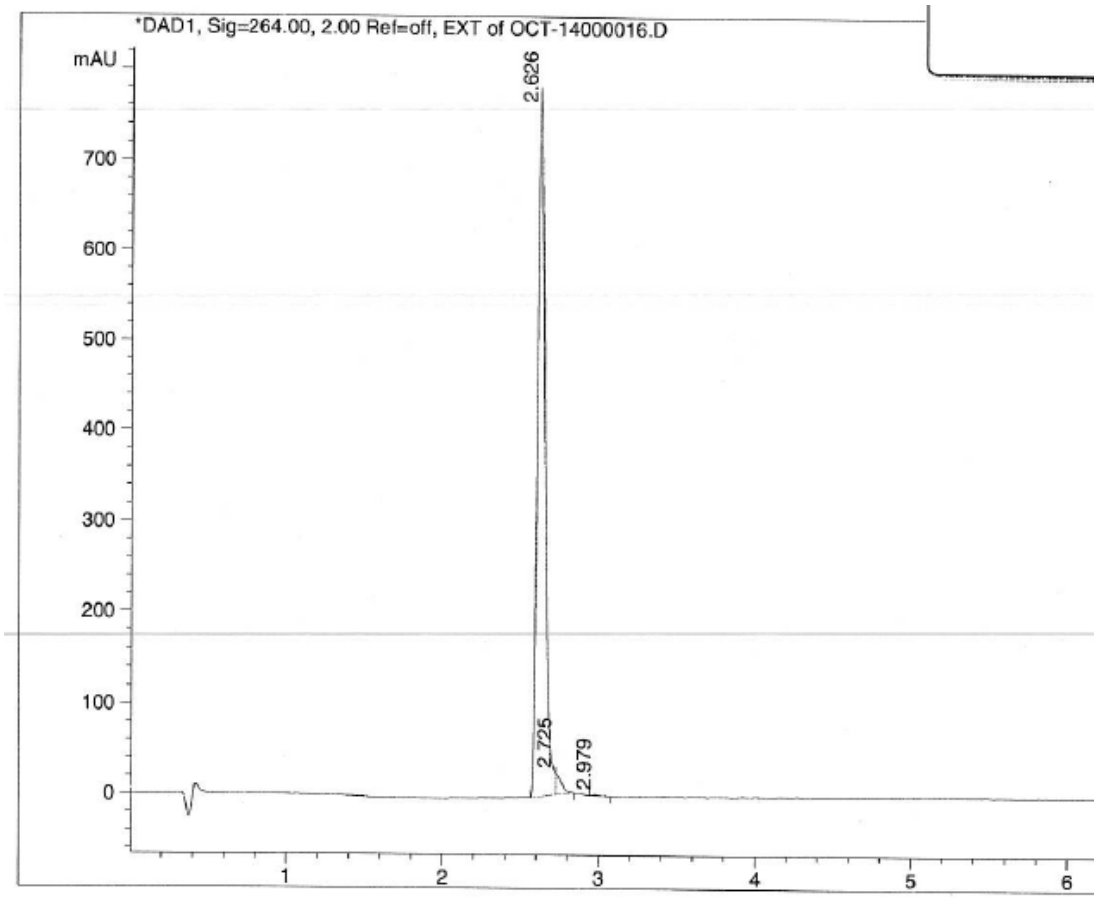
HPLC:



EB-0288: O[C@@H]1[C@@H](O)[C@H](O)[C@](CO)(O)C[C@@H]1NCCNC(OCCNC2=C([N+]([O-])=O)C=C(N=[N+]=[N-])C=C2)=O

485.4	486.4(M+H)	**(400 MHz, CD_3OD)**: δ 7.81 (d, J = 2.6 Hz, 1H), 7.28 (dd, J = 9.2, 2.7 Hz, 1H), 7.21-7.13 (m, 1H), 4.37-4.24 (m, 2H), 3.76- 3.59 (m, 3H), 3.56-3.43 (m, 2H), 3.37-3.33 (m, 2H), 3.28-3.16 (m, 2H), 3.09-3.06 (m, 1H), 2.90-2.82 (m, 1H), 2.62-2.54 (m, 1H), 2.09-1.86 (m, 1H), 1.57-1.37 (m, 1H)

LCMS:



HPLC: