Predicting and Experimentally Validating Hot-Spot Residues at Protein-Protein Interfaces

SUPPORTING INFORMATION

Amaurys A. Ibarra,[a†] Gail J. Bartlett,[b†] Zsófia Hegedüs,[c,d] Som Dutt,[c,d] Fruzsina Hobor,[d,e] Katherine A. Horner,[c,d] Kristina Hetherington,[c,d] Kirstin Spence,[d,e] Adam Nelson,[c,d] Thomas A. Edwards,[d,e] Derek N. Woolfson,[a,b,f] Richard B. Sessions[a,f*] and Andrew J. Wilson,[c,d*]

[a]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK

[b]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[c]School of Chemistry, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

[d]Astbury Centre for Structural Molecular Biology, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

[e]School of Molecular and Cellular Biology, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

[f]BrisSynBio, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK

*To correspondence should be addressed: R.Sessions@bristol.ac.uk

A.J.Wilson@leeds.ac.uk

[†]Contributed equally to this work.

Contents

Supplementary Figures and Tables

Table S1. Summary of methods used for computational prediction. [1] Only with scripting multiple separate calls to the command line and collating data; [2] Requires user-specified input file for residues; [3] Interface residues must be manually assigned for each interface to allow for repacking

Method	Command-line?	Web-site	Interface or residue	Multiple structures	Mutations to ala / all	Interface chain assignment	Interface residue assignment	Multiple residues
BUDE	x		interface	yes	Ala	User specified	Automatic or user-specified	Yes
FoldX	x		interface	No[1]	Ala	Automatic	Automatic	No
Robetta		x	interface	No	Ala	User specified	Automatic	No
Rosetta Flex_ddG	x		interface	No[1]	All[2]	User-specified	user-specified (command-line)[3]	Yes[2]
BeAtMuSiC		x	interface	No	All	User-specified	Automatic	No
mCSM		x	residue	No	User-specified only	User-specified?	User-specified	No

Table S2. Comparison of different methods in predicting the experimental ΔΔG values from the 748 data points from the SKEMPI database. Positive: ΔΔG >= 4.184 kJ/mol; Neutral: 4.184 kJ/mol > ΔΔG < -4.184 kJ/mol; Negative: ΔΔG <= -4.184 kJ/mol. R-value is the Pearson correlation coefficient between predicted and experimental. Fraction Correct = (Tpos+Tneut+Tneg)/748

	SKEMPI ***(Positive 224; Neutral 510; Negative 4)***							
Method	***True Positive***	***True Neutral***	***True Negative***	***False Positive***	***False Neutral***	***False Negative***	***Fraction Correct***	***Pearson R value***
BUDE	130	435	0	76	107	0	0.76	0.50
FoldX*	112	268	2	111	77	178	0.51	0.28
Flex_ddG (Talaris)	113	456	0	56	123	0	0.76	0.51
Flex_ddG (ref2015)	139	416	0	86	94	13	0.74	0.49
BeAtMuSiC	145	391	0	120	92	0	0.72	0.47
mCSM	151	422	0	90	85	0	0.77	0.66
Average per mutation of all methods	137	446	0	120	92	0	0.78	0.60

Table S3. Comparison of different methods in predicting the experimental ΔΔG values from the 25 alanine mutant data points from the NOXA-B/MCL-1 (competition), SIMS/SUMO (direct titration) and GKAP/SHANK-PDZ (direct titration) experiments combined. Positive: ΔΔG >= 4.184 kJ/mol; Neutral: 4.184 kJ/mol > ΔΔG < -4.184 kJ/mol; Negative ΔΔG <= -4.184 kJ/mol. R value is the Pearson correlation coefficient between predicted and experimental values. Fraction Correct = (Tpos+Tneut+Tneg)/25. BUDE-NMR* and BUDE-MD* comprises data from NOXA-B/MCL-1, SIMS/SUMO only since GKAP/SHANK-PDZ is an X-ray structure.

	Combined data from NOXA-B/MCL-1, SIMS/SUMO and GKAP/Shank-PDZ alanine mutants							
Method	***True Positive***	***True Neutral***	***True Negative***	***False Positive***	***False Neutral***	***False Negative***	***Fraction Correct***	***Pearson R value***
BUDE	8	12	0	2	3	0	0.80	0.58
FoldX	4	13	0	0	6	2	0.68	0.52
Flex_ddG (Talaris)	7	11	0	3	4	0	0.72	0.56
Robetta	9	13	0	1	2	0	0.88	0.71
BeAtMuSiC	9	12	0	2	2	0	0.84	0.71
mCSM	8	9	0	5	3	0	0.68	0.36
MD-BUDE*	7	10	0	4	4	0	0.64	0.70
NMR-BUDE*	7	9	0	1	2	0	0.84	0.62
Average per mutation of all methods (not including MD/NMR)	9	14	0	0	2	0	0.92	0.76

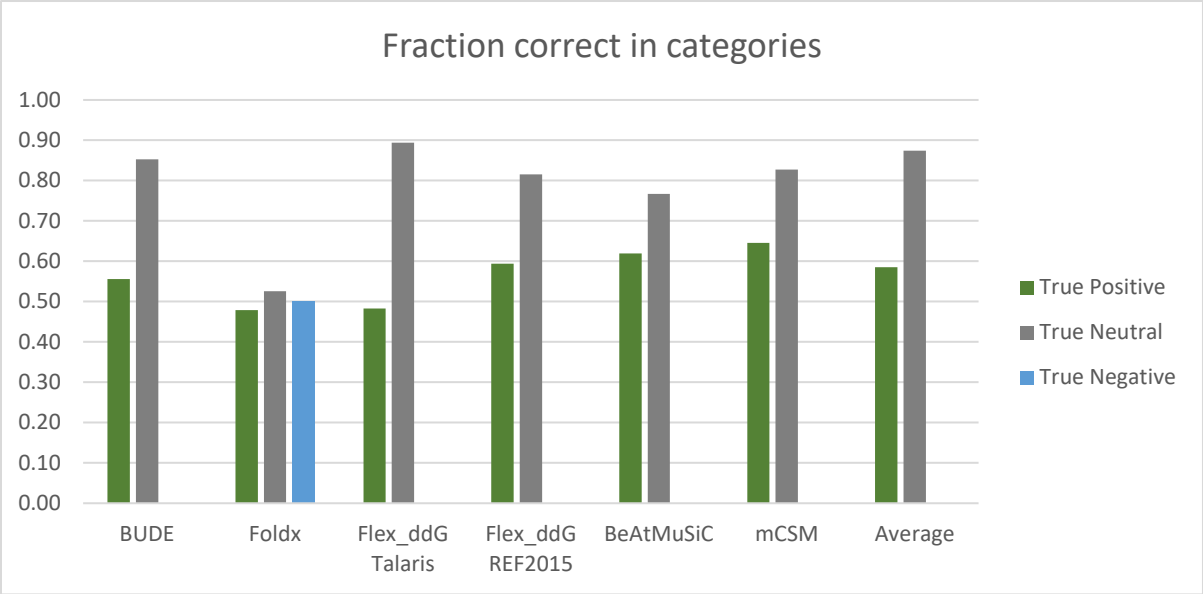


Figure S1. Experimentally determined positive (hot-spot), neutral and negative residues in the SKEMPI database versus the computational predictions using the different methods. True Positive, True Neutral and True Negative bars represent predictions that match the experimental values.

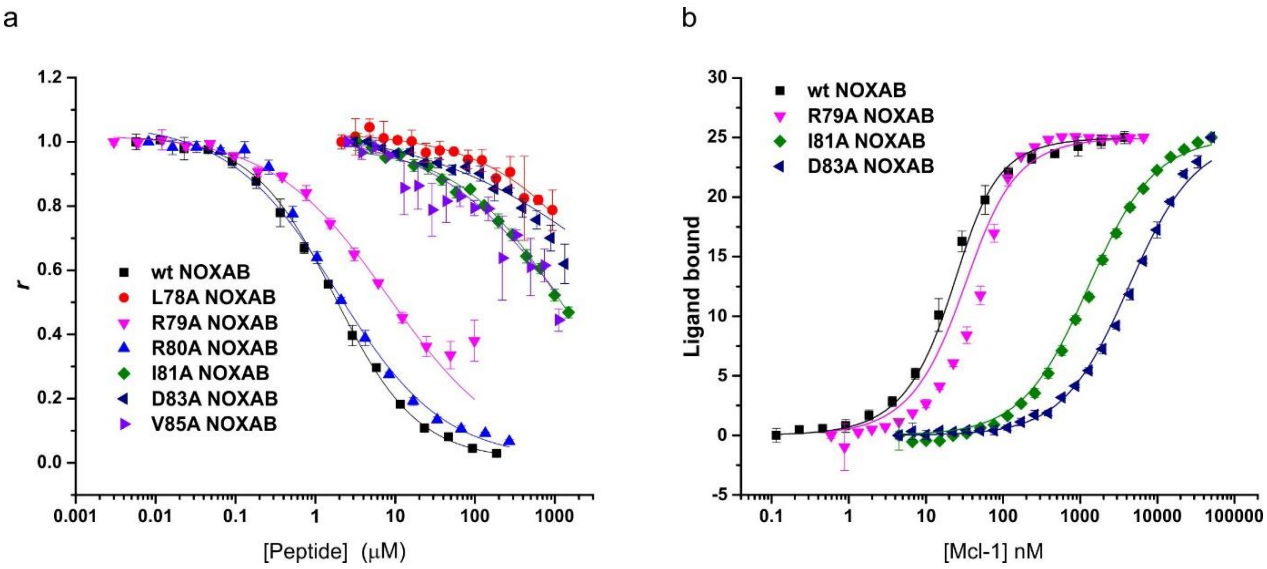


Figure S2. (a) Competition fluorescence anisotropy assay for testing hot-spot for NOXA-B/Mcl-1 interaction in 50 mM Tris, 150 mM NaCl, 0.01 % Triton-X, pH 7.4 using FITC-Ahx-NOXA-B as a tracer at 25 nM concentration and Mcl-1 at 200 nM concentration. (b) Fluorescence anisotropy direct titrations of selected FITC labelled NOXA-B variant peptides using the same conditions.

Table S4. IC_{50} values for inhibition of FITC-Ahx-NOXA-B/MCL-1 interaction by NOXA-B variant peptides determined by competition fluorescence anisotropy.

Peptide	IC_{50} (μM)
Ac-NOXA-B	1.77 ± 0.03
Ac-NOXA-B L78A	7000 ± 2000
Ac-NOXA-B R79A	8.3 ± 0.3
Ac-NOXA-B R80A	1.9 ± 0.3
Ac-NOXA-B I81A	1220 ± 70
Ac-NOXA-B D83A	13000 ± 8000
Ac-NOXA-B V85A	1200 ± 200

Table S5. K_d values of NOXA-B variant peptides binding to MCL-1 protein determined by fluorescence anisotropy.

Peptide	K_d (nM)
FITC-Ahx-NOXA-B	10 ± 1
FITC-Ga-NOXA-B R79A	19 ± 3
FITC-Ahx-NOXA-B I81A	1300 ± 50
FITC-Ahx-NOXA-B D83A	4300 ± 200

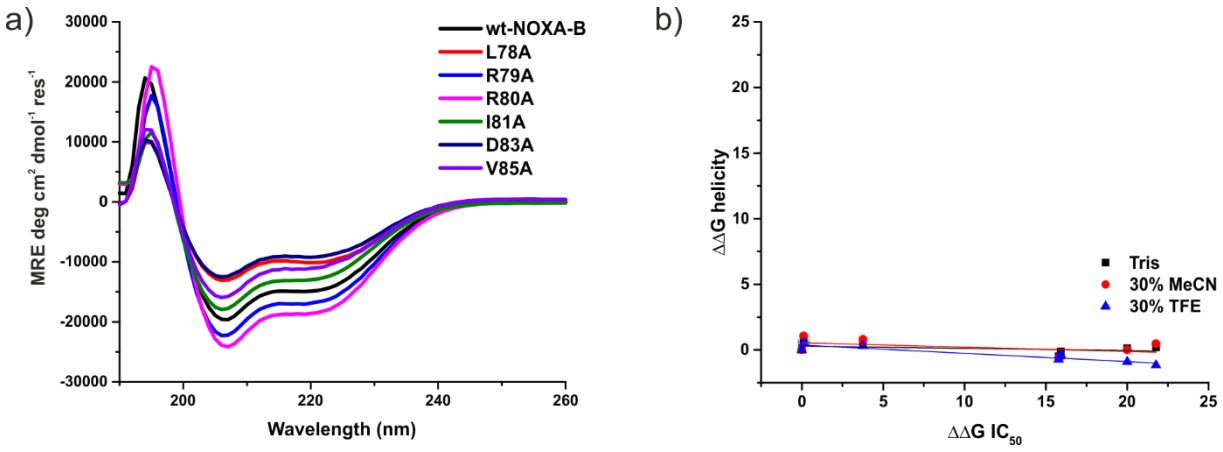


Figure S4. a) CD spectra of all NoxaB alanine variants and the wild type peptide at 20 ºC, 30% TFE in Tris buffer. b) Plot of ΔΔG helicity against ΔΔG IC_{50} of all NoxaB alanine variants. All data calculated in comparison to % helicity and IC_{50} of *wt* NoxaB peptide and calculated using $\Delta G = -RTLnK_{eq}$

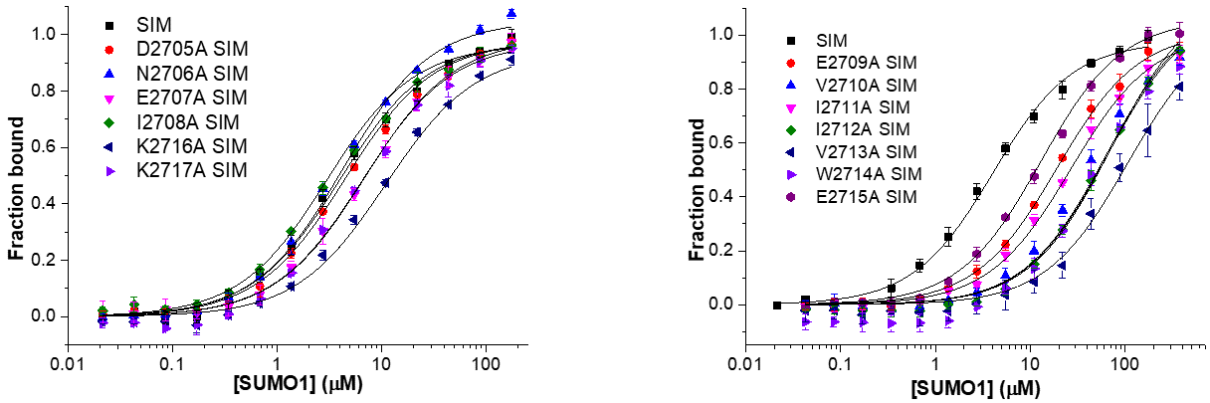


Figure S5. Fluorescence anisotropy direct titrations for testing hot-spots of SIM/SUMO-1 interaction in 20 mM Tris, 150 mM NaCl, 0.01 % Triton-X, pH 7.4 using 50 nM FAM PEG-SIM tracer.

Table S6. K_d values of FAM-PEG SIM variant peptides binding to SUMO-1 protein determined by fluorescence anisotropy.

Peptide	K_d (μM)	Peptide	K_d (μM)
FAM-SIM	3.7 ± 0.3	FAM- I2711A SIM	26.4 ± 1.5
FAM-D2705A SIM	4.2 ± 0.1	**FAM-I2712A SIM**	**64.8 ± 8.0**
FAM -N2706A SIM	4.0 ± 0.2	**FAM-V2713A SIM**	**105.9 ± 23.6**
FAM-E2707A SIM	6.6 ± 0.2	**FAM-W2714A SIM**	**63.6 ± 16.1**
FAM-I2708A SIM	2.7 ± 0.1	FAM-E2715A SIM	12.5 ± 0.7
FAM-E2709A SIM	18.9 ± 0.9	FAM-K2716A SIM	10.7 ± 0.8
FAM-V2710A SIM	**70.8 ± 8.9**	FAM-K2717A SIM	6.4 ± 0.6

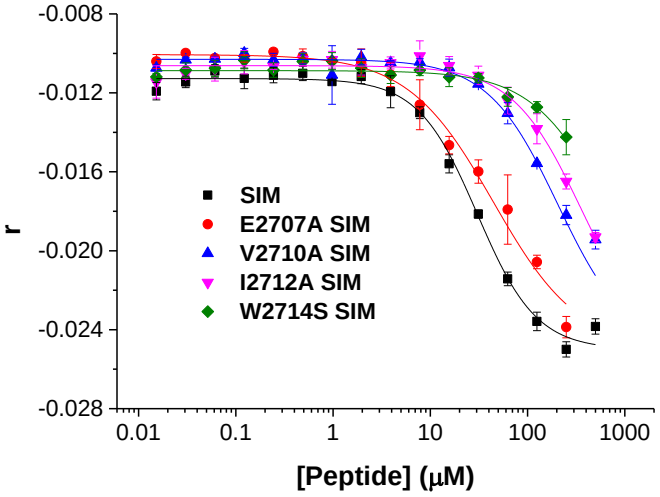


Figure S6. Fluorescence anisotropy competition titrations for testing hot-spots of SIM/SUMO-1 interaction in 50 mM Tris, 150 mM NaCl, 0.01 % Triton-X, pH 7.4 using 50 nM FAM PEG-SIM tracer and 3μM SUMO1 protein.

Table S7. IC_{50} values of acetylated SIM and Ala mutant SIM peptides determined by fluorescence competition titration assays. IC_{50} for E, V, I and W variant SIMs were extracted by fitting the curves by fixing final anisotropy at – 0.025, which is the lowest anisotropy obtained from saturation curve for Ac SIM peptide.

Peptide	IC_{50} (μM)
SIM	30 ± 1
E2707A SIM	45 ± 4
V2710A SIM	205 ± 7
I2712A SIM	351 ± 12
W2714S SIM	684 ± 264

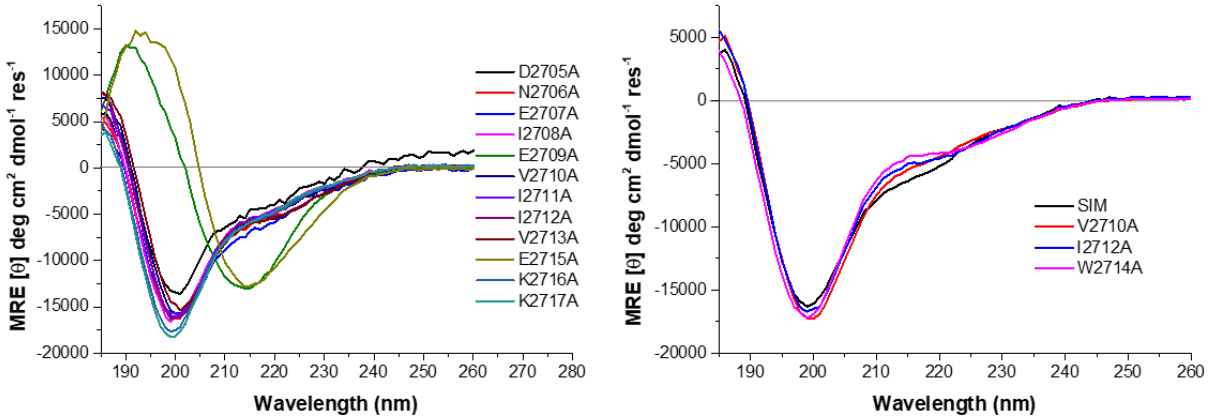


Figure S7. CD spectra of 100 μM labelled and non-labelled SIM alanine variants and the wild type peptide at 20 ºC in 10 mM phosphate buffer, pH 7.4.

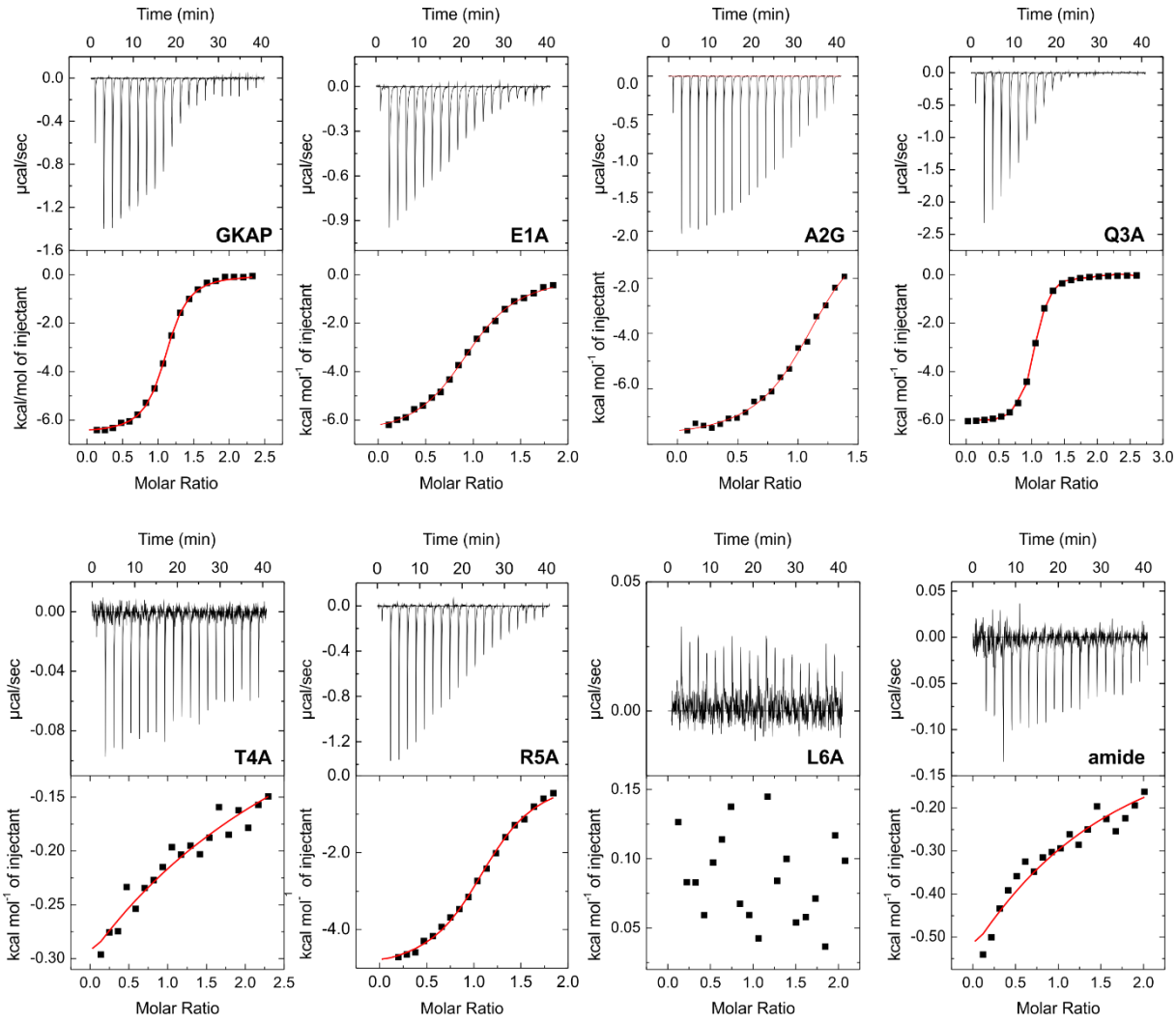


Figure S8. Raw ITC (upper) and fitted enthalpograms (lower) for wt-GKAP and all the variant peptides tested. ITC was performed in 20 mM Tris, 150 mM NaCl pH 7.5 at 25 °C.

Table S9. Fitted ITC data for GKAP/Shank1 interaction

	Sequence	n	K_d (µM)	ΔH (kcal/mol)	ΔS (cal/molK)
GKAP	Ac-EAQTRL-COOH	1.08 ± 0.004	2.77 ± 0.15	-6.54 ± 0.04	3.49
E1A	Ac-**A**AQTRL-COOH	0.979 ± 0.007	10.16 ± 0.57	-6.71 ± 0.07	0.35
A2G	Ac-E**G**QTRL-COOH	1.14 ± 0.007	6.99 ± 0.58	-7.81 ± 0.06	-2.61
Q3A	Ac-EA**A**TRA-COOH	0.989 ± 0.003	1.46 ± 0.09	-6.12 ± 0.03	6.17
T4A	Ac-EAQ**A**RL-COOH	1	1187.6 ± 276.5	-3.09 ± 0.47	3.03
R5A	Ac-EAQT**A**L-COOH	1.140 ± 0.008	9.52 ± 0.75	-5.09 ± 0.07	5.88
L6A	Ac-EAQTR**A**-COOH	no binding observed			
amide	Ac-EAQTRL-CO**NH$_2$**	1	255.7 ± 64.3	-1.57 ± 0.18	11.2

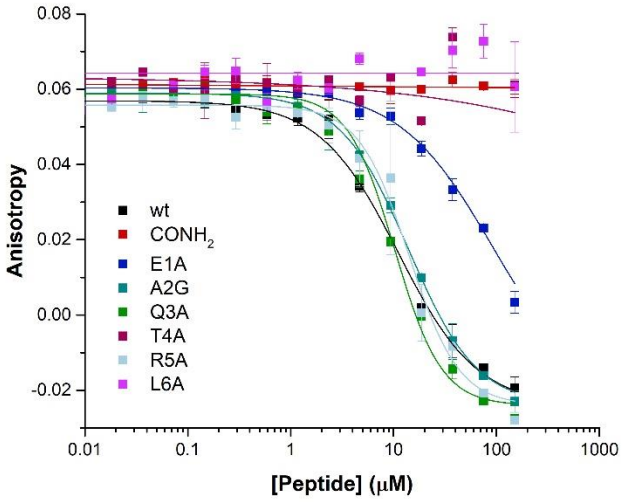


Figure S9. Competition fluorescence anisotropy assay for testing hot-spots of GKAP/Shank1 PDZ interaction in 20 mM Tris, 50 mM NaCl, pH 7.45, using FITC-Ahx-GKAP as a tracer at 50 nM and Shank1 at 1μM concentration.

Table S10. IC_{50} values of acetylated GKAP and Ala variant GKAP SIM peptides determined by fluorescence competition titration assays.

Peptide	**IC_{50} (μM)**
GKAP	11 ± 1
E1A GKAP	91 ± 5
A2G GKAP	14.0 ± 0.3
Q3A GKAP	10.3 ± 0.9
T4A GKAP	did not inhibit
R5A GKAP	14.1 ± 0.9
L6A GKAP	did not inhibit
CO_2H-$CONH_2$ GKAP	did not inhibit

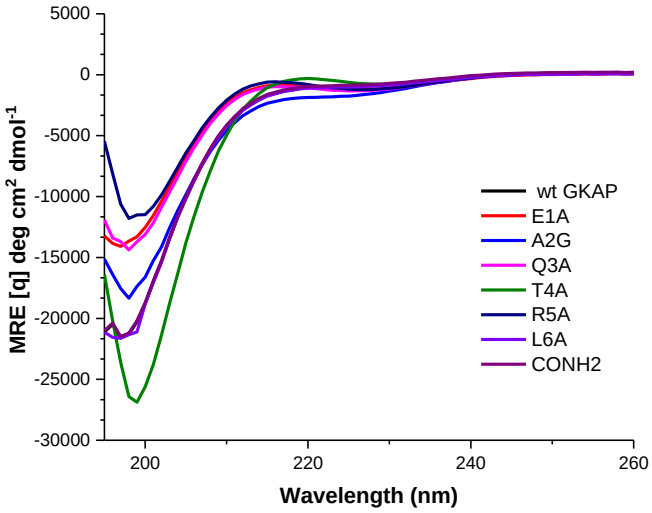


Figure S10. CD spectra of all GKAP alanine variants and the wild type peptide at 20 ºC in 20 mM Tris buffer

Table S11. IC_{50} values for inhibition of FITC-BID/BCL-x_L interaction by wild-type affimer and variants determined by competition fluorescence anisotropy.

Affimer	IC_{50} (nM)
wt	327 ± 32
W41A	no inhibition.
W44A	no inhibition
W76A	352 ± 49

Supplementary methods

Expression and purification of proteins

MCL-1 and BCL-x_L were expressed, purified and characterized following minor adaptations to methods described previously[1, 2] For SUMO and SHANK1-PDZ new protocols were elaborated. Experimental methods for preparation of these proteins are detailed in the following sections.

MCL-1

The pet28a His-SUMO Mcl-1 (172-327) construct was over-expressed in the E.coli strain Rosetta 2. 10 ml of overnight starter culture was used to inoculate 1 L 2 xYT containing 50 µg/ml Kanamycin and 50ug/ml Chloramphenicol. Cultures were grown at 37 °C plus shaking until OD600 ~ 0.6 – 0.8, the temperature was then switched to 18 °C and protein expression induced by the addition of 0.8mM IPTG.

Induced cultures were grown at 18 °C plus shaking overnight before harvesting by centrifugation. Cells were resuspended in 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole and lysed by sonication in the presence of 10 µL of 1 U.ml-1 DNase I per litre of over-expression culture and 5 mM MgCl2. The cell lysate was centrifuged (Beckman JA25.50 rotor, 17,000 rpm, 45 min, 4 °C) and the supernatant filtered (0.22 µM syringe filter) before application onto a 5 ml HisTrap that had previously been equilibrated with 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole. The cleared cell lysate was then allowed to flow through the HisTrap with the aid of a peristaltic pump. The HisTrap was then washed with 10 CV of 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole followed by 10 CV 50 mM TRIS pH 8.0, 500 mM NaCl, 50 mM imidazole. The His-SUMO-Mcl-1 fusion protein was then eluted from the HisTrap with 50 mM TRIS pH 8.0, 500 mM NaCl, 300 mM imidazole. The His-SUMO-Mcl-1 fusion protein was cleaved overnight in dialysis into 50 mM TRIS pH 8.0, 250 mM NaCl in the presence of Smt3 protease, Ulp1, overnight at 4 °C. To remove any uncleaved Mcl-1, His-SUMO and Ulp1, the sample was reapplied to a HisTrap in 50 mM TRIS pH 8.0, 250 mM NaCl and the flow through containing Mcl-1 collected. The flow through containing cleaved Mcl-1 was concentrated (Amicon Ultra centrifugal filter, MWCO 10,000) to approximately 5 ml. The sample was then filtered before being loaded onto a Superdex 75 column (GE healthcare) equilibrated in 50 mM TRIS pH 8.0, 250 mM NaCl, 0.5 mM DTT, 2.5% Glycerol. The purified was concentrated to ~ 6 mg/ml and stored at – 80 °C.

SHANK1 PDZ

Human Shank1 PDZ domain (656-762) was cloned into the pGEX-6P-2 expression vector and transformed to BL21 Gold cell line. 10 ml overnight starter culture was inoculated to 1 l 2xYT media containing 50 µg/ml chloramphenicol. Cells were grown at 37°C until OD600 0.6-0.8 and induced with 0.1 mM IPTG and incubated overnight at 18ºC. Cells were harvested and resuspended in 20 mM Tris, pH 8, 500 mM NaCl, containing protease inhibitor and 1U of DNAseI per litre of cell culture. Cells were lysed by sonication (8 cycles, 20 seconds on 40 seconds off, 10 µA) and centrifuged at 25.000 g for 45 minutes at 4ºC. The supernatant was filtered (0.22 µm membrane) and applied to glutathione beads and washed with 10 CV 20 mM Tris, pH 8, 500 mM NaCl. GST was cleaved on-column overnight at 4 ºC using Prescission protease. The eluted fractions were concentrated and purified by size-exclusion chromatography on S75 26/60 pg column in 20 mM Tris, 150 mM NaCl, 5% glycerol pH 7.5 buffer. Collected fractions were analysed by SDS-PAGE and concentrated. Pure protein was analysed by high resolution mass spectrometry: expected m/z: 12326.3 measured m/z: 12325.6. Concentration was determined using UV-VIS spectroscopy in 6M urea using the 8480 M^{-1} cm^{-1} extinction coefficient.

SUMO-1

The pet19b His-TEV-SUMO-1 (18-97) construct was over-expressed in the E.coli strain Rosetta 2. 10 ml of overnight starter culture was used to inoculate 1 L 2 xYT containing 50 μg/ml Ampicillin. Cells were grown at 37°C until OD600 0.6-0.8 and induced with 1 mM IPTG and incubated overnight at 18ºC. Cells were harvested by centrifugation and resuspended in 25 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole, containing protease inhibitor and 1U of DNAseI per litre of cell culture. Cells were lysed by sonication (6 cycles, 20 seconds on 40 seconds off, 10 μA) followed by centrifugation at 25.000 g for 45 minutes at 4ºC. Supernatant was filtered and applied onto a preequilibrated 5 ml HisTrap column. The cleared cell lysate was then allowed to flow through the HisTrap with the aid of a peristaltic pump. The HisTrap was then washed with 10 CV of 25 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole followed by 10 CV 25 mM TRIS pH 8.0, 500 mM NaCl, 50 mM imidazole. The His-TEV-SUMO-1 was then eluted from the HisTrap with 25 mM TRIS pH 8.0, 500 mM NaCl, 300 mM imidazole, and the His-tag was cleaved overnight at 4°C, in a dialysis tube in presence of TEV protease in 25 mM TRIS pH 8.0, 250 mM NaCl. The dialysed sample was reapplied to a HisTrap in 50 mM TRIS pH 8.0, 250 mM NaCl and the flow through containing SUMO-1 collected and concentrated (Amicon Ultra centrifugal filter, MWCO 10,000) to approximately 5 ml. The sample was then filtered before being loaded onto a Superdex 75 column (GE healthcare) equilibrated in 25 mM TRIS pH 8.0, 250 mM NaCl, 2.5% Glycerol. The purified protein was concentrated and stored at – 80 °C. Protein quality was assessed using SDS-PAGE, high resolution mass spectrometry and CD.

BCL-x_L

pet28a His-SUMO Bcl-x_l (Chimera with BCL-2, 1-198, missing 26-81) was expressed using a similar protocol as described above for MCL-1. Briefly bacterial cultures were grown at 37°C to a density of OD_{600} = 0.5-0.8 and protein expression then induced by addition of 0.5 mM isopropyl-β-D-thiogalactoside (IPTG). Upon induction, cultures were cooled and maintained at 18°C for protein expression overnight. Cell pellets were harvested by centrifugation for 7 minutes at 8655 xg and then resuspended in lysis buffer (25 mM Tris-HCl pH 8, 500 mM NaCl) containing (1 mg mL^-1 lysozyme, 10mg mL^-1 of DNase1 and 1 tablet per 1 L culture protease inhibitor tablets (Complete Mini, Roche) and lysed by sonication. The cell lysate was then centrifuged for 45 minutes at 23655 xg, the supernatant was collected and applied onto a Ni-Sepharose resin column (bed size 5 mL). The column was washed with 40 mL wash buffer I (25 mM Tris-HCl, 500 mM NaCl, pH 8), 40 mL wash buffer II (25 mM Tris-HCl, 15 mM Imidazole, 500 mM NaCl, pH 8) and 400mL wash buffer III (25 mM Tris-HCl, 40 mM Imidazole, 500 mM NaCl, pH 8). Bound proteins were eluted with 20 mL of elution buffer (25 mM Tris-HCl, 400 mM Imidazole, 500 mM NaCl, pH 8). The elution fraction was loaded on Superdex75 (26/600) prep grade size exclusion chromatography column, equilibrated with Gel filtration buffer (25

mM Tris-HCl, 250 mM NaCl, 0.5mM DTT, 2.5% glycerol pH 7.4). Samples of the fractionated eluent were applied to 15% SDS gels to assess sample purity and the purified protein was concentrated using centrifugal concentrators with a MWCO of 10kD and stored at -80°C in small aliquots.

BCL-x_L affimers

BCL-x_L affimer variants were produced using the Quikchange Site Directed Mutagenesis Kit (Agilent) and the wild type protein expression vector as template. The sequences of all constructs were confirmed by DNA sequencing prior to expression. Affimers were expressed as previously described.[3] Briefly: pet11a Affimer constructs were over-expressed in the *E. coli* strain Rosetta 2. 10 ml of overnight starter culture was used to inoculate 1 L 2 x YT containing 125 μg/ml Ampicillin Cultures were grown at 37 °C plus shaking until OD600 ~ 0.6 – 0.8, the temperature was then switched to 18 °C and protein expression induced by the addition of 0.5mM IPTG. Induced cultures were grown at 18 °C plus shaking overnight before harvesting by centrifugation. Cells were resuspended in 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole and lysed by sonication in the presence of 10 μL of 1 U.ml-1 DNase I per litre of over-expression culture and cell lysate was clarified (Beckman JA25.50 rotor, 17,000 rpm, 45 min, 4 °C). The supernatant was filtered (0.45 μM syringe filter) before application onto a 5 ml HisTrap that had previously been equilibrated with 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole. The cleared cell lysate was then allowed to flow through the HisTrap with the aid of a peristaltic pump. The HisTrap was then washed with 10 CV of 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole followed by 10 CV 50 mM TRIS pH 8.0, 500 mM NaCl, 50 mM imidazole and 10 CV 50 mM TRIS pH 8.0, 500 mM NaCl, 100 mM imidazole. The Affimer was then eluted from the HisTrap with 50 mM TRIS pH 8.0, 500 mM NaCl, 300 mM imidazole. Successful elution was confirmed on a gel before further purification was undertaken. The eluted Affimer was concentrated (Amicon Ultra centrifugal filter, MWCO 10,000) to approximately 5 ml. The sample was then filtered before being loaded onto a Superdex 75 column (GE healthcare) equilibrated in 50 mM TRIS pH 8.0, 250 mM NaCl, 0.5 mM DTT, 2.5% Glycerol. The protein eluted as a monomer from gel filtration. The purified was concentrated to ~ 6 mg/ml and stored at – 80 °C with the addition of 5% Glycerol.

All protein constructs were expressed in the *E. coli* strain Rosetta 2.

Computational methods and MD

Usability and limitations of CAS methods

The command-line tools of BudeAlaScan and FoldX readily process ensembles of structures, although the latter requires a script to run the program and collate the data. Likewise, Rosetta Flex_DDG can process structure ensembles with scripting but is several orders of magnitude

slower than the other methods. Robetta is accessed *via* a website and processes a single structure at a time, with the interface automatically assigned. These jobs are batch processed so the time for results to be returned depends on server load. BeAtMuSiC and mCSM are both accessed by webservers and used interactively. BeAtMuSiC has the feature of mutating the interface to all residue types, not just to alanine. mCSM requires the definition of one residue at a time and a structure upload for each calculation, but this can be scripted in contrast with servers like DrugScore PPI that requires human input for each submission.

BudeAlaScan

BudeAlaScan is command-line python application for computational alanine scanning. It employs ISAMBARD[4] for structure manipulation, a customised version of the Bristol University Docking Engine (BUDE)[5] for energy calculations and SCWRL[6] for side chain repacking (this latter feature is not used in this paper, but is required for calculating hot-regions or clusters of interfacial residues, that will be described elsewhere). Side chain flexibility in BudeAlaScan is addressed by calculating the interaction energy of a set of rotamers of the residue being replaced by alanine with the "receptor" and averaging all rotamers with a favourable energy. This feature is enabled by default for residues of the class DERKH and is designed to account for the entropic cost of freezing a residue in a salt bridge. The program was run in scan mode with default parameters. The application will be available via the BAlaS server: http://coiledcoils.chm.bris.ac.uk/balas

Molecular Dynamics Simulations

All simulations were performed using the GROMACS 5.1.4 suite and the following general protocols. Structures from the protein database (SIMS/SUMO 2LAS; GKAP/SHANK-PDZ 1Q3P; NOXA-B/MCL-1 2JM6) were processed with the GROMACS tool chain. The utility pdb2gmx was used to add hydrogen atoms consistent with pH 7, virtual-site hydrogens for mobile groups and parameterise the protein with the amber99SB-ildn forcefield. The protein was placed in an orthorhombic box 2 nm larger than the protein in each dimension and filled with TIP3P water containing 0.15 M sodium chloride ions to give a charge-neutral system overall. After 10000 steps of steepest descent minimisation, molecular dynamics was initiated with random velocities while restraining the protein backbone to its original position with a force constant of 1000 kJ/nm for 0.2 ns. Simulations were developed for a further one microsecond both with and without the backbone position restraints under periodic boundary conditions. The Particle Mesh Ewald's method was used for long range electrostatic interactions while short range

Coulombic and van de Waals energies were truncated at 1.2 nm. The temperature was maintained at 310 K using the v-rescale method and the pressure at 1 bar with the Berendsen barostat. The use of virtual-site hydrogens allowed a 5 fs time step for the leapfrog integrator. Bond constraints were implemented with the LINCS method and SETTLE used for waters. Trajectories were processed and analysed with the GROMACS tools and visualised with VMD 1.9.3.

Peptide synthesis and purification

General Remarks

All amino acids and resins were purchased from either Novabiochem (Merck) or Sigma-Aldrich. All amino acids were *N*-Fmoc protected and side chains protected with Boc (Lys); O^{t}Bu (Asp, Ser, Thr); Trt (Asn, Gln); Pbf (Arg). Synthesis of all peptides was performed using a microwave assisted automated peptide synthesiser (CEM, Liberty or Liberty Blue). Coupling of 6-aminohexanoic acid, γ-aminobutyric acid and N-terminal labelling were performed manually. DMF used in peptide synthesis was of ACS grade and from Sigma Aldrich. Peptides were synthesised on an 0.1 mmol scale and split before acetylation and fluorescent ligation. Lyophilisation was performed using a BenchTop Pro with Omnitronics^{TM} from VirTis SP Scientific. Preparative HPLC was performed on an Agilent Technologies 1260 infinity controller in conjunction with an diode array detector. Analytical HPLC was performed on an Agilent Technologies 1260 infinity controller in conjunction with an diode array detector. Mass spectrometry data were obtained on a Bruker Daltonics micrOTOF using electrospray ionisation (ES)MS instruments as appropriate.

Cycles for Automated Peptide Synthesis

Peptides were prepared on a microwave assisted Liberty Blue CEM peptide synthesiser followed this cycle:

Resin Loading

Clean reaction vessel; wash with DMF; transfer resin to reaction vessel; wash with DMF:CH_2Cl_2 (1:1); vessel draining.

Deprotection and Coupling

Add 20% piperidine in DMF (4 mL); microwave method; wash with DMF (4+4+4+4 mL); drain; add amino acid (2.5 mL); add coupling reagent (1 mL); add base (0.5 mL); microwave method; wash through manifold to waste (2 mL); drain.

For methods that ***did not*** use microwave assistance, the reaction cycle was the same, expect the microwave method for deprotection and coupling was replaced by agitation of the resin at r.t. for 15 min and 25 min respectively.

After the final residue, the resin was ejected from the reaction vessel and any further linker coupling, capping, cleavage and deprotections were performed manually using methods A to G

For the microwave methods used, the temperature and total time is shown below:

Deprotection Microwave Methods

Method	Ramp Time	Total Time	Max Temp
Standard	20-30 sec	1:05	90 ºC
75 ºC deprotection*	30 sec	0:30	~50 ºC
	30-75 sec	3:00	75 ºC
Conventional*	N/A	5:00	rt
	N/A	10:00	rt

Coupling Methods

Method	Ramp Time	Total Time	Max Temp
Standard	20-30 sec	1:05	90 ºC
50 ºC MW	N/A	2:00	rt
	30-75 sec	4:00	50 ºC
Arg* coupling	N/A	25:00	rt
	30-75 sec	2:00	75 ºC
75 ºC coupling	30 sec	0:30	~50 ºC
	30-75 sec	5:00	75 ºC

Arginine was subjected to double coupling as standard.

FMoc-PEG linker was treated as any other conventional amino acid and coupled using microwave assisted synthesis

Methods for Manual Solid Phase N-terminal Chain Elongation and Capping

Method A: Coupling of Aminohexanoic acid (Ahx) and gamma-aminobutyric acid (Ga)

Following ejection from the automated synthesiser, the resin was placed in an fritted empty SPE tube and the desired unnatural amino acid (5 equiv.), DIPEA (5 equiv.) and HCTU (5 equiv.) were dissolved in DMF (1 mL) and added to the resin, followed by agitation for 1 h. For double couplings, this step was repeated. After removal of the reagents by filtration, the resin was washed with DMF (3 × 2 mL × 2 min) and the success of the coupling determined by a negative colour test (Method C). Deprotection of the Fmoc-protected *N*-terminus then followed (Method B).

Method B: Deprotection of N-Fmoc protecting groups

N-terminal Fmoc-protecting groups were removed by the addition of 20% piperidine: DMF (v/v) (5 × 2 mL × 2 min), followed by rinsing the resin with DMF (5 × 2 mL × 2 min). Successful deprotection was determined by a positive colour test (Method C).

Method C: Kaiser Test[7]

The Kaiser Test was used for the determination of the successful coupling or deprotection for any residue coupled manually. A small number of resin beads were rinsed with CH_2Cl_2 and placed in a vial, followed by the addition of two drops of each of the three solutions below:

1) Ninhydrin (5% w/v) in ethanol
2) Phenol (80% w/v) in ethanol
3) 1 mM KCN $_{(aq.)}$ in pyridine (2% v/v)

The solution was then heated to ca. 100 °C for five minutes. A successful coupling gave no change in the colour of the beads, whereas bright blue beads demonstrated a successful deprotection.

Method D: N-terminal acetylation

Acetic anhydride (10 equiv.) and DIPEA (10 equiv.) were dissolved in DMF (1 mL) and the solution was transferred to the resin. After 2 h, the resin was drained, washed with DMF (3 × 2 mL × 2 min) and successful capping determined by a negative colour test (Method C).

Method E: N-terminal FITC labelling

Fluorescein isothiocyanate (6 equiv.) was dissolved in 12:7:5 pyridine:DMF:CH_2Cl_2 and the solution transferred to the resin in the dark. After 16 h, the resin was washed with DMF (3 × 2 mL × 2 min) ahead of cleavage and deprotection. The solvents were of anhydrous grade and distilled before use.

Method F: On resin Fluorescein (FAM) labelling

The SIM and its alanine variant peptides (X-A SIM) (1 eq., 0.1 mmol) on resin were labelled with fluorescein (FAM). To resin bound peptides (0.1 mmol, 1.0 eq) containing free N-terminal PEG linker was added pre activated solution of 5, 6-dicarboxylic fluorescein (5 eq), DIC (3eq) and HOBt (3 eq) in 4 ml DMF. The resin mixture was shaken on a vertical spinner for 16h at RT. Thereafter, the resin was filtered, washed with DMF (5x), CH_2Cl_2 (5x) and then Et_2O (3x).

Method 5: Cleavage and deprotection of Rink amide MBHA resin

After elongation and N-terminal capping was complete, the resin was washed with CH_2Cl_2 (5 × 2 mL × 2 min), Et_2O (5 × 2 mL × 2 min) and dried under vacuum for ca. 2 h. Peptides were simultaneously cleaved and side-chain deprotected using 'Reagent K' TFA:EDT:Thioanisole:Phenol:H_2O 82:3:5:5:5 (3 × 2 mL × 2 h). The solution was precipitated in ice-cold Et_2O (25 mL) and placed in a centrifuge (3000 rpm × 10 min), the supernatant removed and the precipitate resuspended in ice-cold Et_2O and placed in a centrifuge again. This process was repeated 3-4 times and the precipitate was dried under a stream of nitrogen overnight, before being dissolved in H_2O and lyophilised.

Peptide Purification

NOXA-B variant peptides were pre-purified by automated RP column chromatography on a Biotage Isolera 1.3.3., using a RediSep®Rf gold reversed phase C18 column by Teledyne Isco on an increasing gradient of acetonitrile (5-50%) in water + 0.1% TFA (v/v) at a flow rate of 12 mL min^{-1}. Crude peptides

were suspended in H_2O as concentrated as possible, fractions were checked by LCMS, concentrated *in vacuo* and lyophilised.

Peptides were purified by preparative UV- or MD- HPLC using a Jupiter Proteo or a Kinetex EVO C18 preparative column (reversed phase) on an increasing gradient of acetonitrile in water + 0.1% formic acid (v/v) at a flow rate of 10 mL min^{-1}. Crude peptides were dissolved in H_2O or DMSO at an approximate concentration. Purification runs injected a maximum of 5 mL of crude peptide solution and were allowed to run for 30 min, with acetonitrile increasing at a stated gradient. In regards to UV-HPLC, the eluent was scanned with a diode array at 220, 210 and 280 nm. In regards to MD-HPLC the mass directed chromatography software Masshunter by ChemStation (Agilent) was used to allow the collection of the desired peptide by mass, with the eluent split into an Agilent 6120 Quadropole LCMS which triggers collection of eluent at a programmed m/z. Fractions containing purified peptide were combined, concentrated *in vacuo* and lyophilised.

Biophysical assays

Isothermal titration calorimetry (ITC)

ITC experiments were carried out using Microcal ITC200i instrument (Malvern) at 25°C in 20 mM Tris, 150 mM NaCl, pH 7.5 buffer. ShankPDZ was dialysed against the buffer prior to experiment, lyophilized peptides were dissolved in the same buffer. 150 μM shankPDZ was present in the cell and titrated with 1.4-2 mM peptide solutions loaded into the syringe using 20, 2 uL injections with 120 s spacing between the injections for 20 injection. Heats of peptide dilution was subtracted from each measurement raw data. Data was analysed using Microcal Origin 8 and fitted to a one-binding site model.

Fluorescence anisotropy

Fluorescence anisotropy assays were performed in 384-well plates (Greiner Bio-one). Each experiment was run in triplicate and the fluorescence anisotropy measured using a Perkin Elmer EnVisionTM 2103 MultiLabel plate reader, with excitation at 480 nm (30 nm bandwidth), polarised dichroic mirror at 505 nm and emission at 535 nm (40 nm bandwidth, S and P polarised) for FAM and FITC labelled peptides. The excitation and emission wavelength for Bodipy labelled BAK peptide were set to 531nm and 595nm respectively. The excitation and emission wavelength for Fitc labelled BID peptide were set to 490nm and 535nm respectively.

Assay buffers

All assays related to NOXA B/Mcl-1 were carried in tris buffer (50 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4).

All assays related to BAK/BclXL were carried in phosphate buffer (40 mM Phosphate, 200 mM NaCl, 0.02mg/ml BSA, pH 7.4).

All assays related to SIM/SUMO1 and BID/BclXL were carried in tris buffer (20 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4).

All assays related to GKAP/Shank1PDZ were carried in tris buffer (20 mM Tris, 50 mM NaCl, pH 7.4).

Direct binding assays

Fluorescence anisotropy direct titration assays were performed with protein concentration diluted over 16-24 points using ½ dilutions. 20 µL of buffer were first added to each well. 20 µL of a solution of protein was added to the first column. The solution was well mixed and 20 µL was taken out and added to the next column and so on. This operation consists on serial dilution of the protein across the plate. Finally, 20 µL of tracer was added to the wells. For control wells, the tracer peptide was replaced with an identical volume of assay buffer and plates were read after 1 hour.

Competition binding assays

FA competition assays were performed in 384 well plates with the concentration of mutant peptide competitor typically starting from 10-1500 µM, diluted over 16-24 points in 1/2 regime with fixed protein and tracer concentrations. For the NOXA B/Mcl-1 FA competition assay, FITC-NOXA B(68-87) and Mcl-1(172-327) were added to each well to give a final concentration of 50 nM and 150 nM, respectively. For the SIM/SUMO1 FA competition assay, 20 µL of FAM-peg-SIM tracer and 20 µL of SUMO1 were added to each well to give a final concentration of 50 nM and 3µM, respectively. For the GKAP/Shank1PDZ FA competition assay, 20 µL of mutant GKAP and 20 µL of Shank1PDZ were added to each well to give a final concentration of 50 nM and 1µM, respectively. For the BID/BclXL FA competition assay, 20 µL of Fitc-BID and 20 µL of BclXL were added to each well to give a final concentration of 25 nM and 100nM, respectively. For the BAK/BclXL FA competition assay, 20 µL of Bodipy-BAK and 20 µL of BclXL were added to each well to give a final concentration of 50 nM and 150nM, respectively. For control wells, the tracer peptide was replaced with an identical volume of assay buffer. The total volume in each well was 60 µL. Plates were read after 1 h (and 16h for Bcl-XL assays) of incubation at room temperature.

Data analysis

The data from both the P (perpendicular intensity) and S (parallel intensity) channels, resulting from this measurement and corrected by subtracting the corresponding control wells, were used to calculate the intensity and anisotropy for each well following Equations 1 and 2:

$$I = (2PG)+S \quad Equation\ 1$$

$$r = (S–PG) \quad Equation\ 2$$

$$L_b = (r–r_{min})/\lambda(r_{max}–r)+r–r_{min} \quad Equation\ 3$$

$$y = \{(k+x+[FL])–\sqrt{\{k+x+[FL]^2–4x[FL]\}}\}/2 \quad Equation\ 4$$

Fluorescence anisotropy data was processed as described previously.[1]

r = anisotropy, I = total intensity, P = perpendicular intensity, S = parallel intensity, L_b = fraction ligand bound, λ = Ibound/Iunbound = 1, [FL] = concentration of fluorescent ligand, k = K_d , y = L_b Flu-trimer and x = [added titrant], G is an instrument gain factor.*

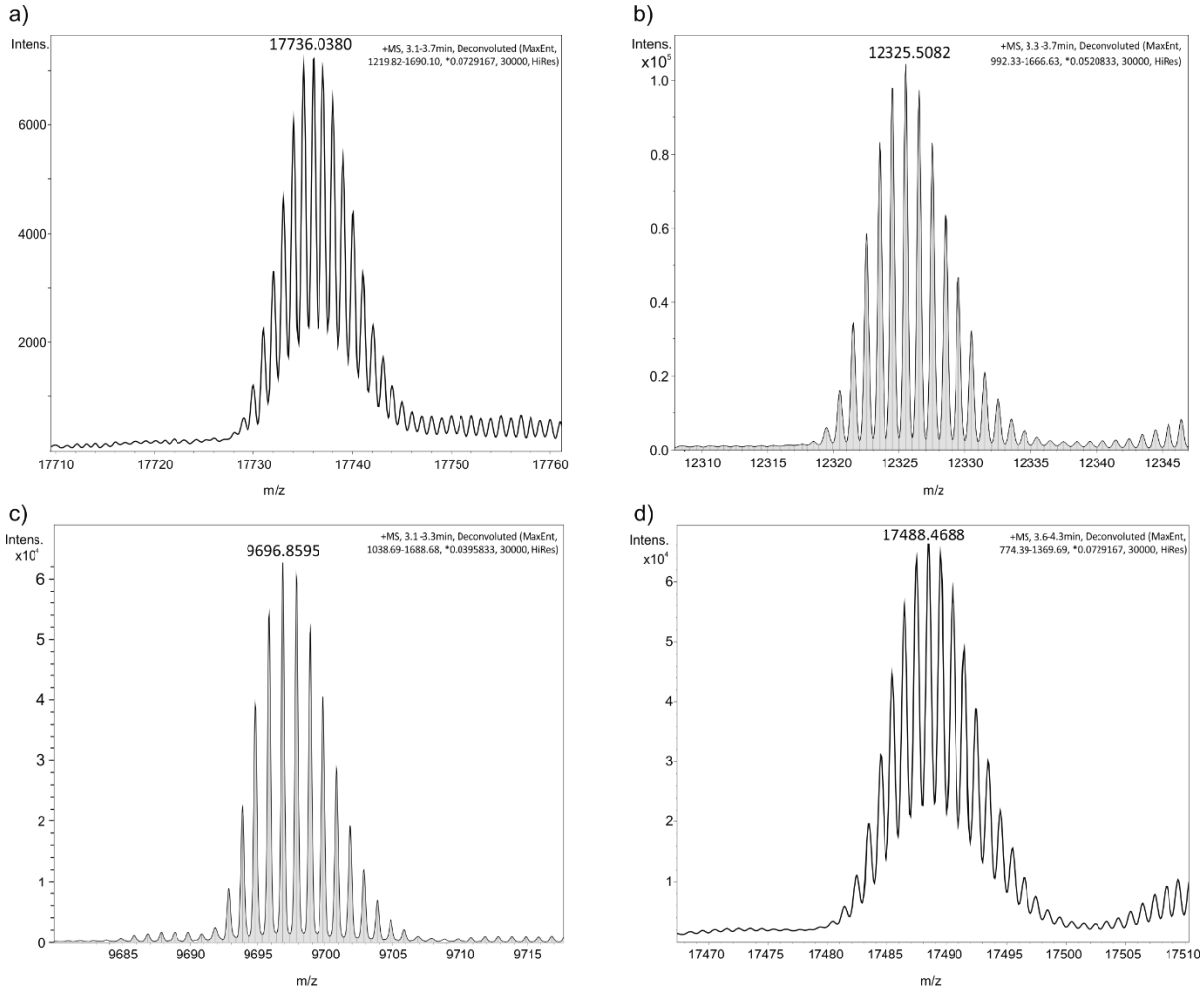
Where I is the total intensity, G is an instrument factor which was set to 1 for all experiments and r is the anisotropy. The average anisotropy (across three experimental replicates) and the standard deviation of these values were then calculated and fit to a sigmoidal logistic model (Equation 3) using OriginPro 9.0 which provided the IC_{50} and error values.

$$y = r_{max} + (r_{min}\text{-}r_{max})/(1+(x/x_o)^p)$$

Circular dichroism

Circular dichroism spectra were recorded on Chirascan Plus (Applied Photophysics) spectropolarimeter. Samples were prepared in 10 mM Na-phosphate buffer pH 7.5 at 100 μM peptide concentration. CD spectra were acquired from 185-260 nm using 1 nm step size, 2nm bandwidth at 25 °C in a 1 mm path length quartz cuvette. Samples were measured twice and averaged, buffer baseline spectrum was substracted from each measurement, and data was converted to mean residue ellipticity.

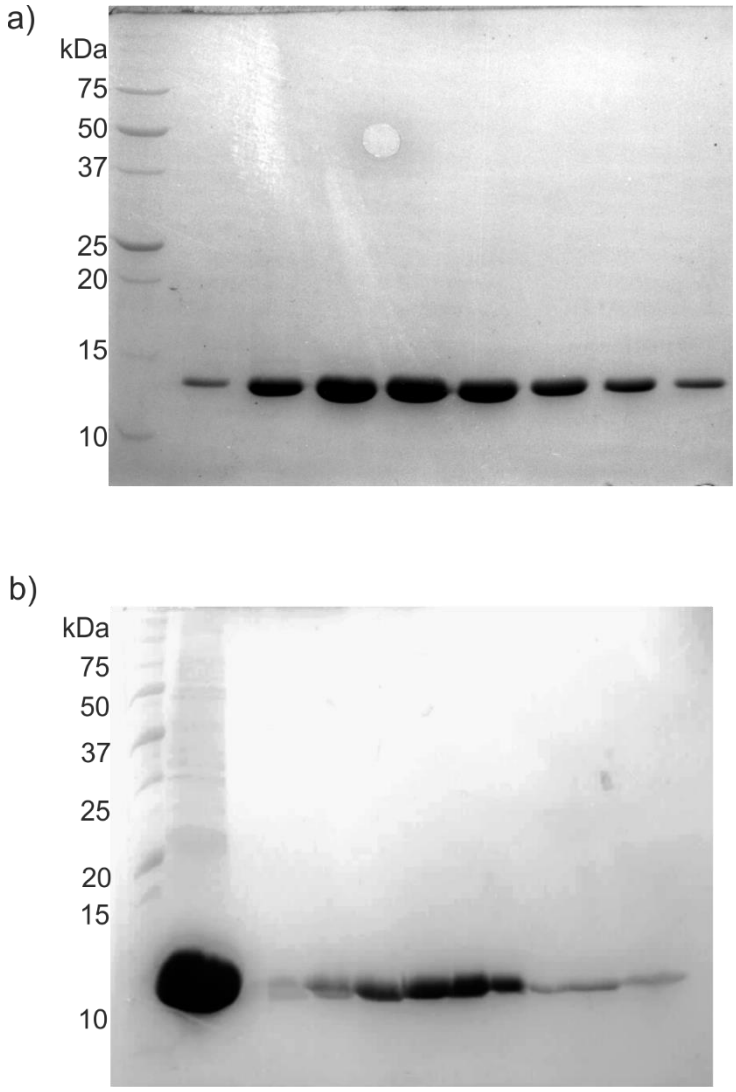
Protein characterization data



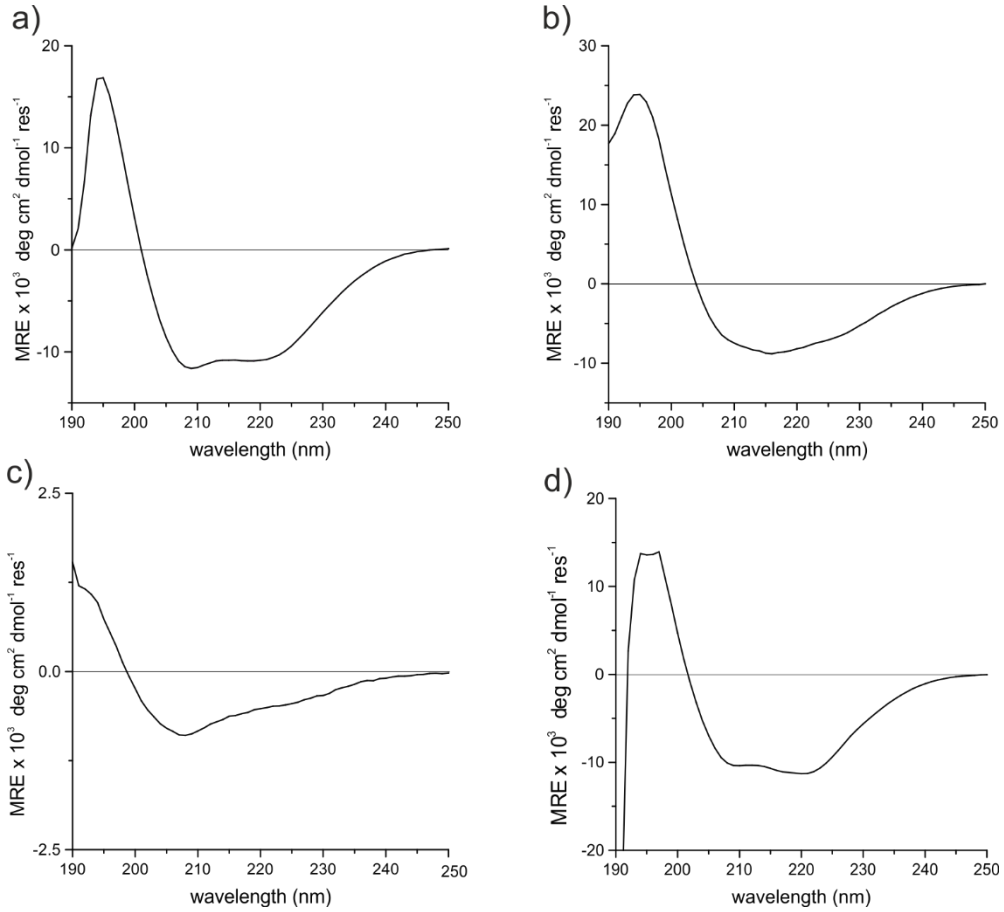
Deconvoluted mass spectra of (a) MCL-1 (b) Shank1 PDZ and (c) SUMO-1 d) BCL-x_L

Calculated and measured molecular weights of the purified proteins

Protein	Calculated MW	Measured
MCL-1	17737.20	17736.0380
SHANK1 PDZ	12326.30	12325.5082
SUMO-1	9698.01	9696.8595
BCL-x_L	17489.57	17488.4688



SDS-PAGE of newly prepared proteins (a) SHANK1 PDZ fractions after size-exclusion chromatography and (b) SUMO-1 after affinity purification (lane 1) and fractions from size-exclusion chromatography (lanes 2-11)



CD of purified (a) MCL-1 (b) Shank1 PDZ and (c) SUMO-1 (d) BCL-x_L

Peptide characterization data

NOXA-B and variant peptides

Tabulated HRMS data of synthesised peptides are shown in below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp[d]) and Observed (Obs[d]) masses.

Peptide	Sequence
wt NOXA-B	Ac-AAQLRRIGDKVNLRQKLLN-NH_2
	FITC-(Ahx)AAQLRRIGDKVNLRQKLLN-NH_2
L78A NOXA-B	Ac-AAQ**A**RRIGDKVNLRQKLLN-NH_2
R79A NOXA-B	Ac-AAQL**A**RIGDKVNLRQKLLN-NH_2
	FITC-(Ga)AAQL**A**RIGDKVNLRQKLLN-NH_2
R80A NOXA-B	Ac-AAQLR**A**IGDKVNLRQKLLN-NH_2
I81A NOXA-B	Ac-AAQLRR**A**GDKVNLRQKLLN-NH_2
	FITC-(Ahx)AAQLRR**A**GDKVNLRQKLLN-NH_2
D83A NOXA-B	Ac-AAQLRRIG**A**KVNLRQKLLN-NH_2
	FITC-(Ahx)AAQLRRIG**A**KVNLRQKLLN-NH_2
V85A NOXA-B	Ac-AAQLRRIGDK**A**NLRQKLLN-NH_2

Acetylated NOXA-B peptides

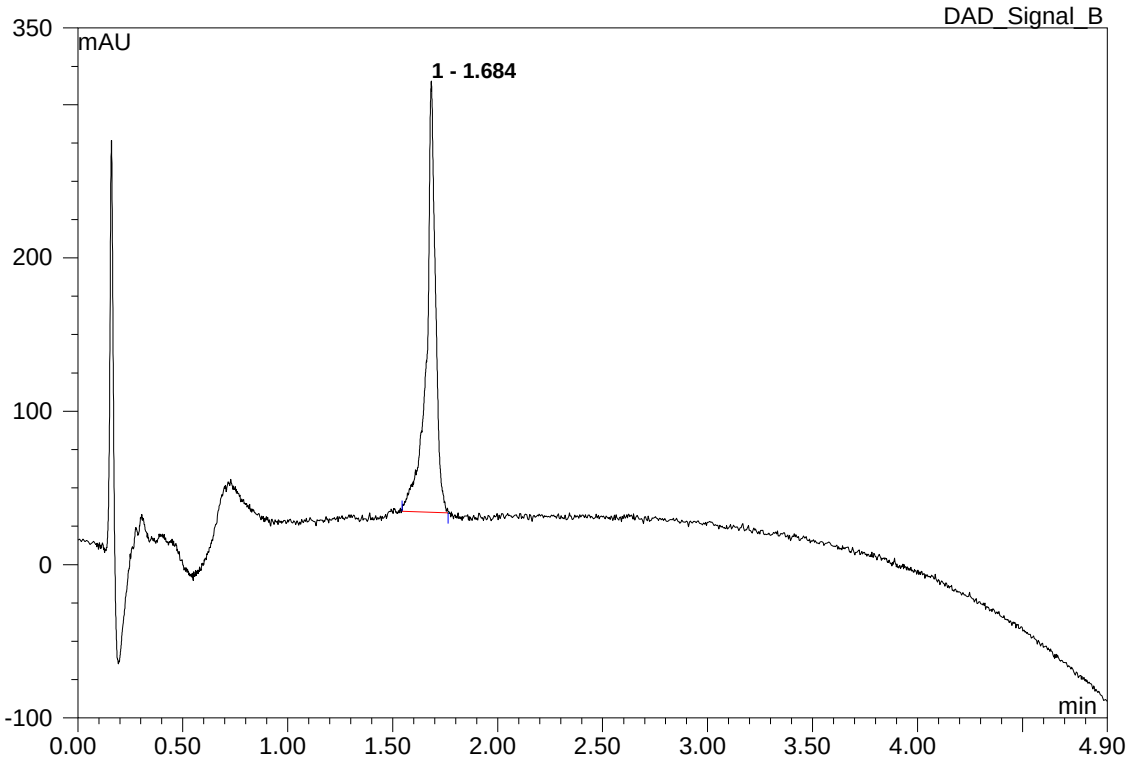
Peptide	$[M+3H]^{3+}$ Obs[d]	$[M+3H]^{3+}$ Exp[d]	$[M+4H]^{4+}$ Obs[d]	$[M+4H]^{4+}$ Exp[d]
wt NOXAB	749.7886	749.7889	562.5938	562.2762
L78A NOXA-B	735.7750	735.7696	552.0839	552.0790
R79A NOXAB	721.4350	721.4342	541.3302	541.3275
R80A NOXA-B	721.4352	721.4342	541.3277	541.3275
I81A NOXAB	735.7738	735.7732	552.0841	552.0817
D83A NOXAB	735.1267	735.1256	551.5975	551.58484
V85A NOXA-B	740.4463	740.4416	555.5873	555.5830

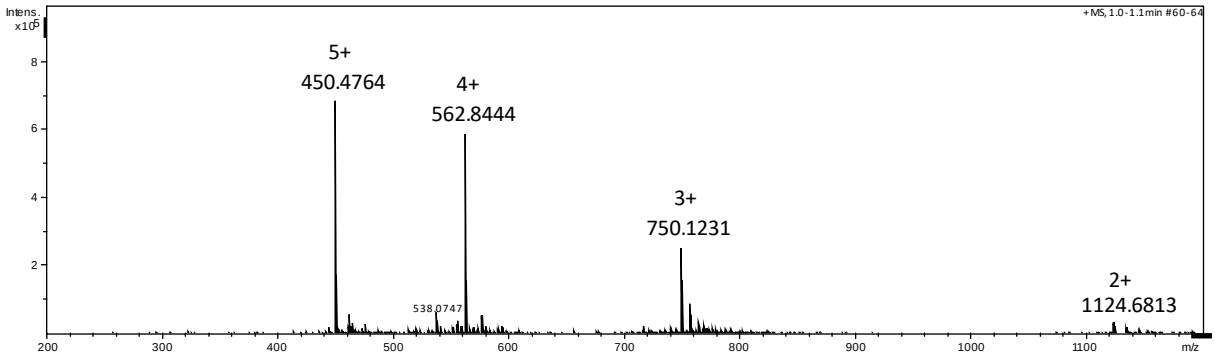
FITC-NOXA-B peptides

Peptide	$[M+3H]^{3+}$ Obs[d]	$[M+3H]^{3+}$ Exp[d]	$[M+4H]^{4+}$ Obs[d]	$[M+4H]^{4+}$ Exp[d]
FITC-*wt* NOXAB	903.1586	903.1594	677.6208	677.6210
FITC-R79A NOXAB	865.4491	865.4602	649.3385	649.3470
FITC-I81A NOXAB	889.1425	889.1430	667.1016	667.1090
FITC-D83A NOXAB	888.4881	888.4953	666.6183	666.6233

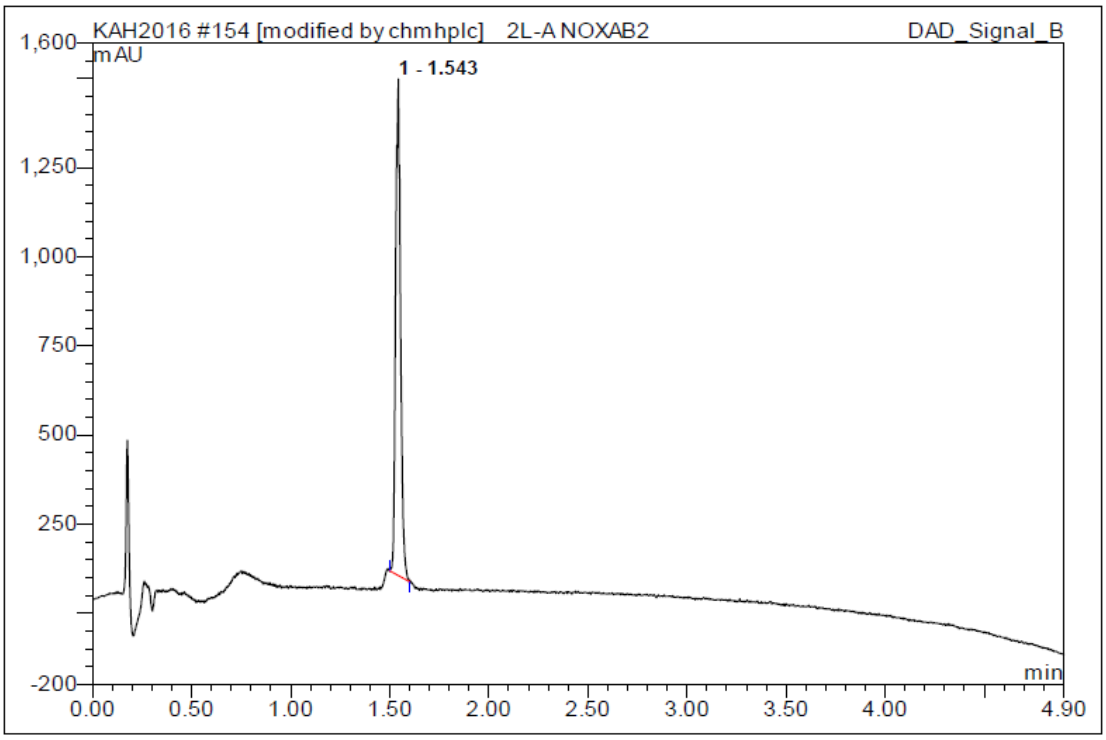
HPLC and ESI-MS data for NOXA-B and variant peptides

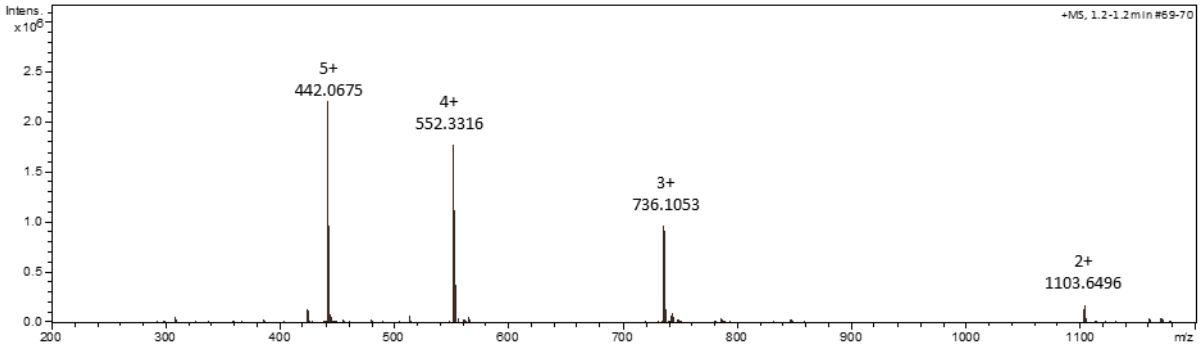
wt NOXA-B



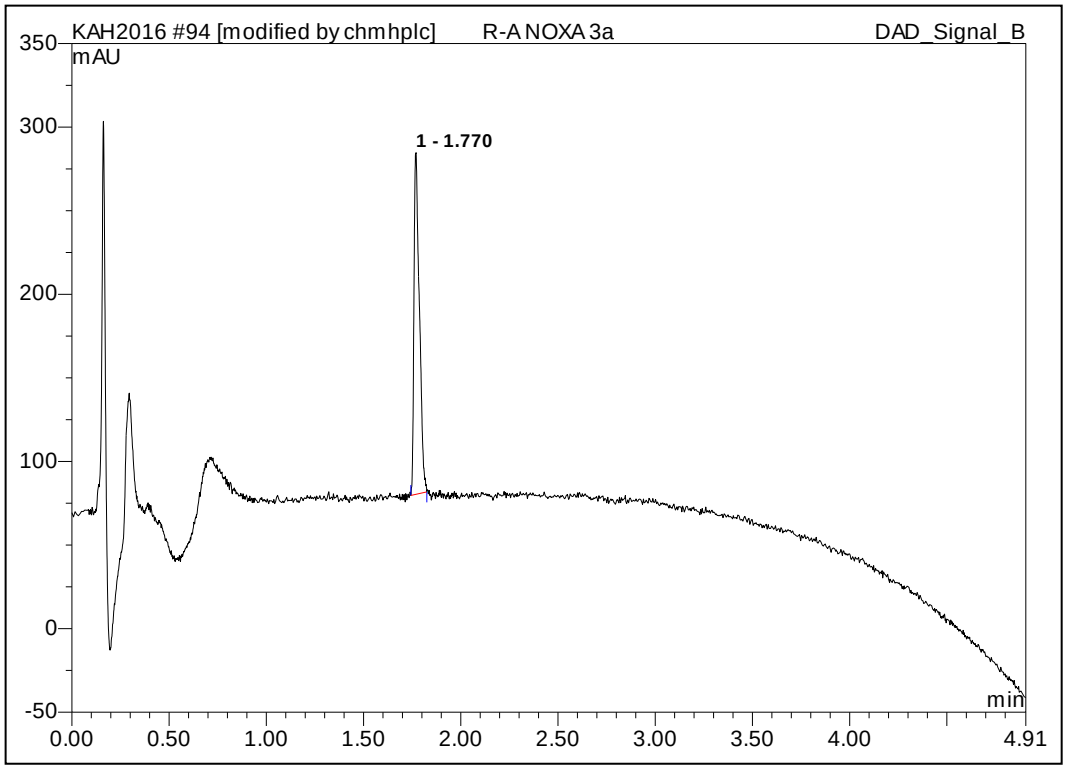


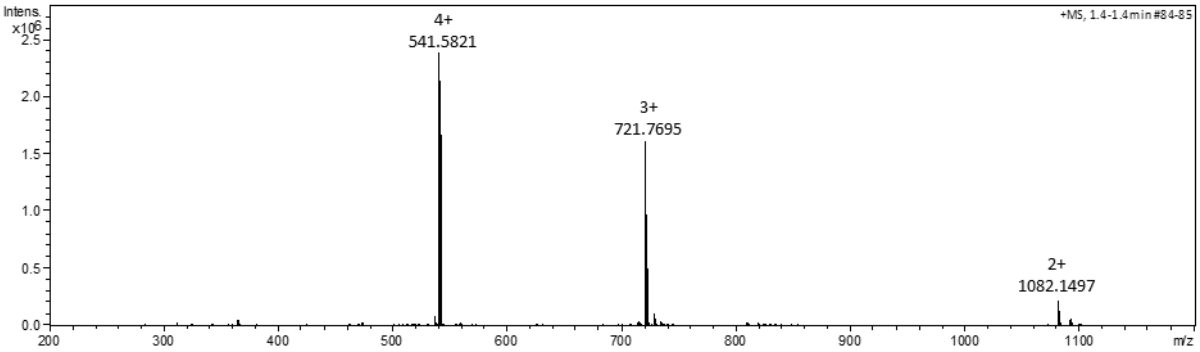
L78A NOXA-B


KAH2016 #154 [modified by chmhplc] 2L-A NOXAB2
DAD_Signal_B
mAU
1 - 1.543
1,600
1,250
1,000
750
500
250
-200
min
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.90

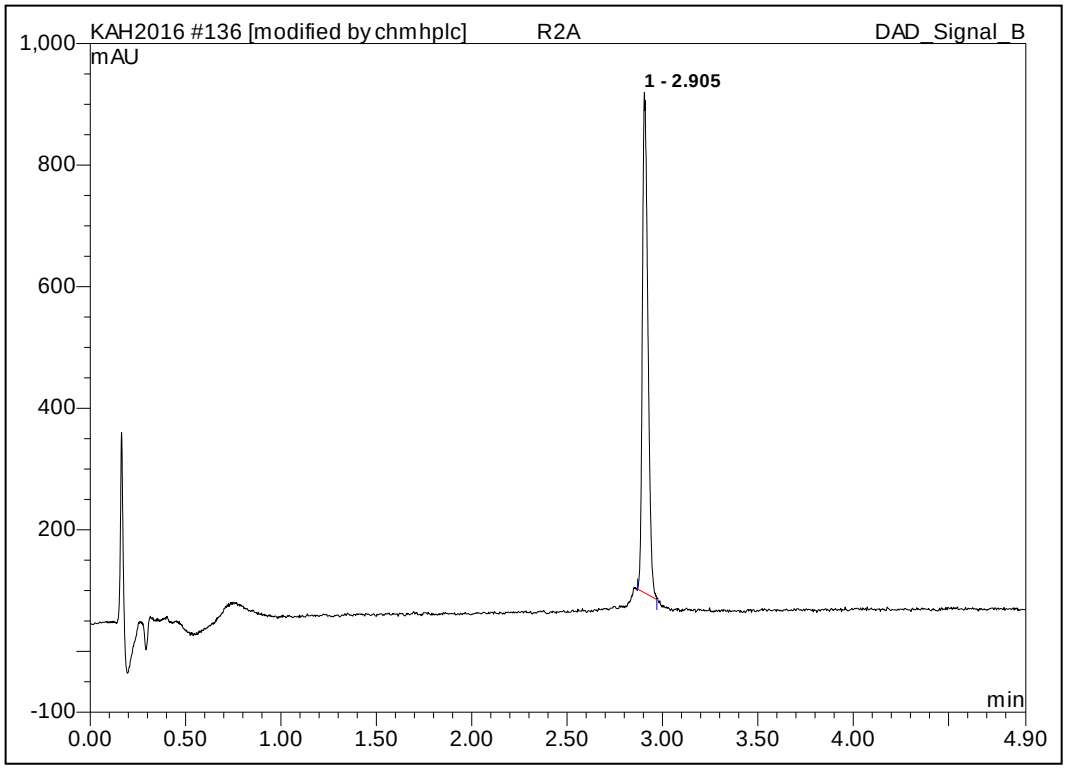

Intens.
x10^6
+MS, 1.2-1.2min #69-70
5+
442.0675
4+
552.3316
3+
736.1053
2+
1103.6496
2.5
2.0
1.5
1.0
0.5
0.0
200
300
400
500
600
700
800
900
1000
1100
m/z

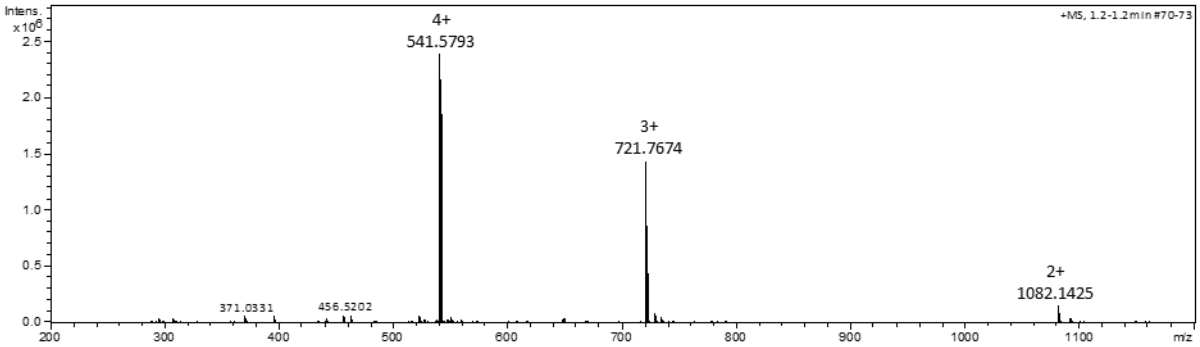
R79A NOXA-B


KAH2016 #94 [modified by chmhplc]
R-A NOXA 3a
DAD_Signal_B
mAU
1 - 1.770
min
350
300
200
100
0
-50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.91

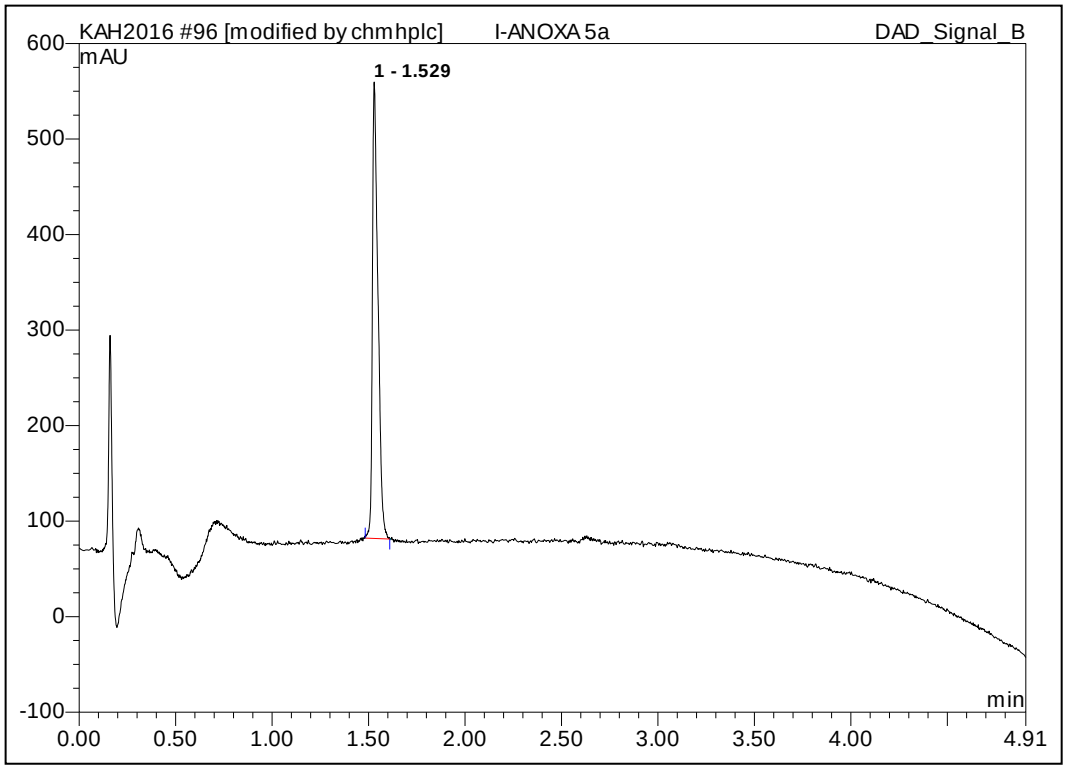

Intens.
x10^6
+MS, 1.4-1.4min #84-85
4+
541.5821
3+
721.7695
2+
1082.1497
2.5
2.0
1.5
1.0
0.5
0.0
200
300
400
500
600
700
800
900
1000
1100
m/z

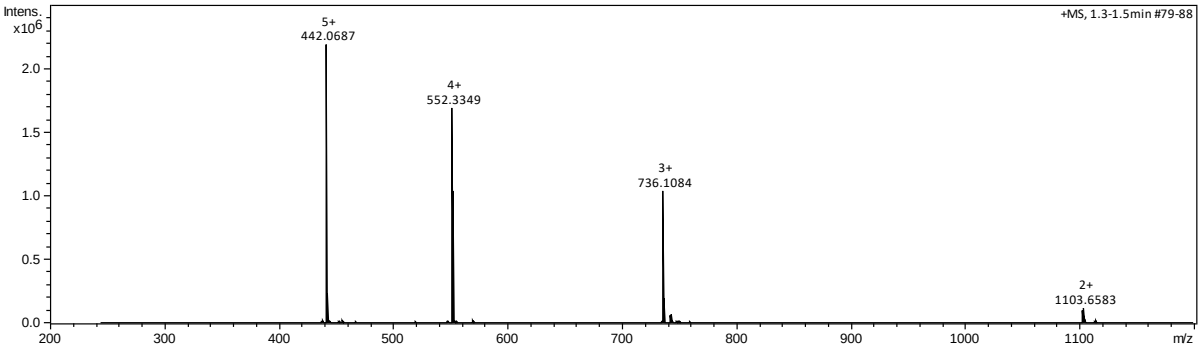
R80A NOXA-B


KAH2016 #136 [modified by chmhplc]
R2A
DAD_Signal_B
mAU
1 - 2.905
1,000
800
600
400
200
-100
min
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.90

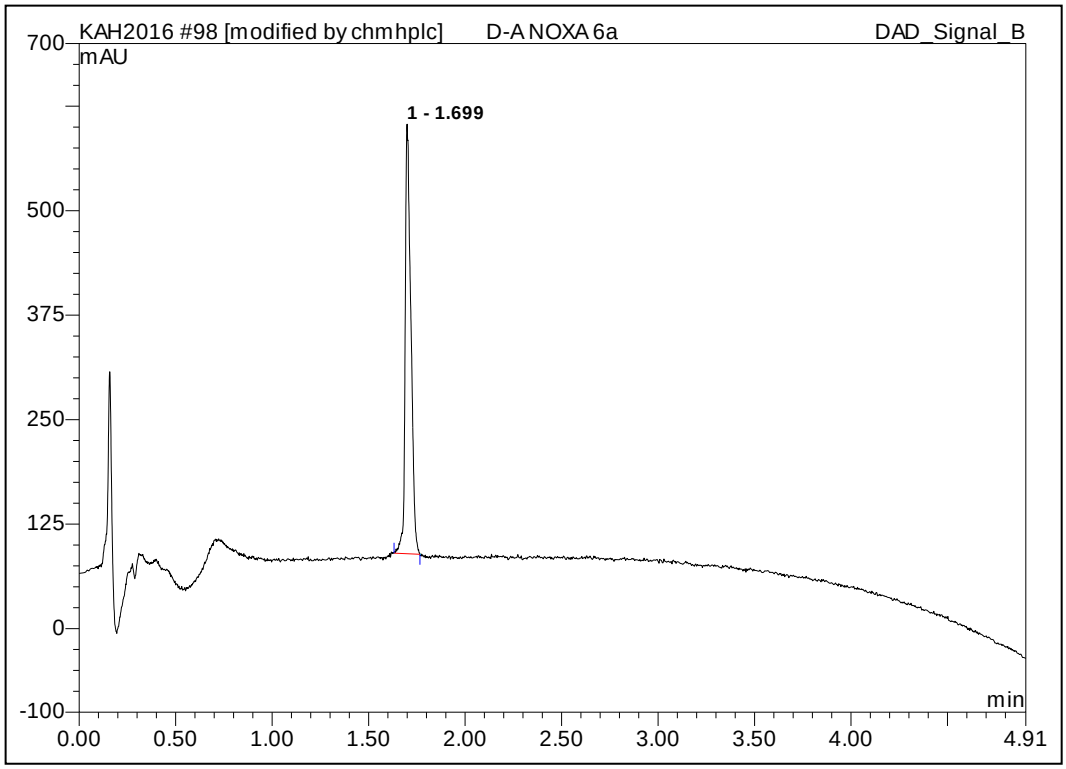

Intens.
x10^6
+MS, 1.2-1.2min #70-73
4+
541.5793
3+
721.7674
2+
1082.1425
371.0331
456.5202
m/z

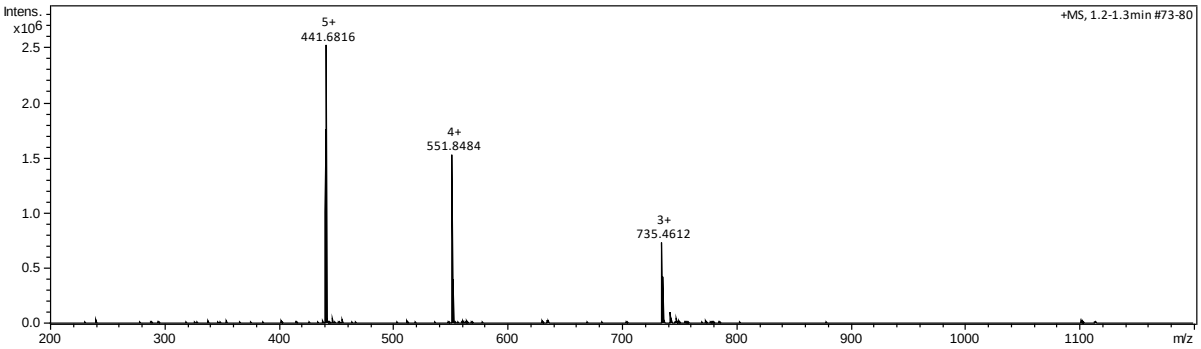
I81A NOXA-B


KAH2016 #96 [modified by chmhplc]
I-ANOXA 5a
DAD_Signal_B
mAU
1 - 1.529
600
500
400
300
200
100
0
-100
min
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.91

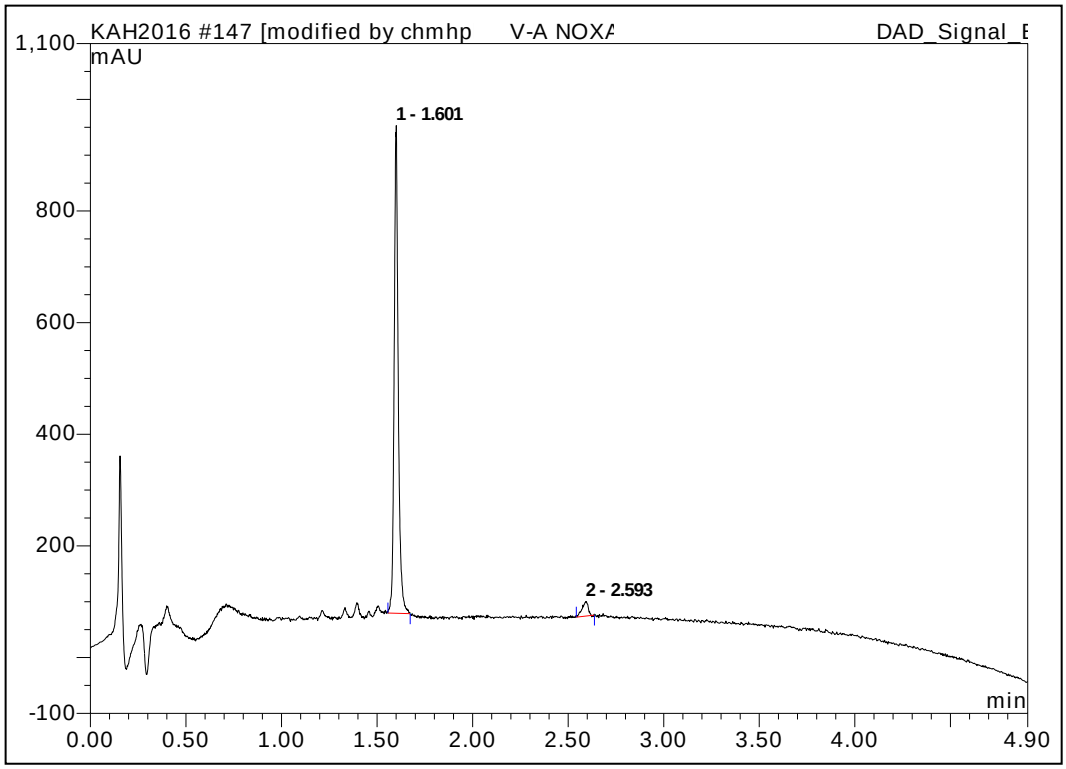

Intens.
x10^6
+MS, 1.3-1.5min #79-88
5+
442.0687
4+
552.3349
3+
736.1084
2+
1103.6583
2.0
1.5
1.0
0.5
0.0
200
300
400
500
600
700
800
900
1000
1100
m/z

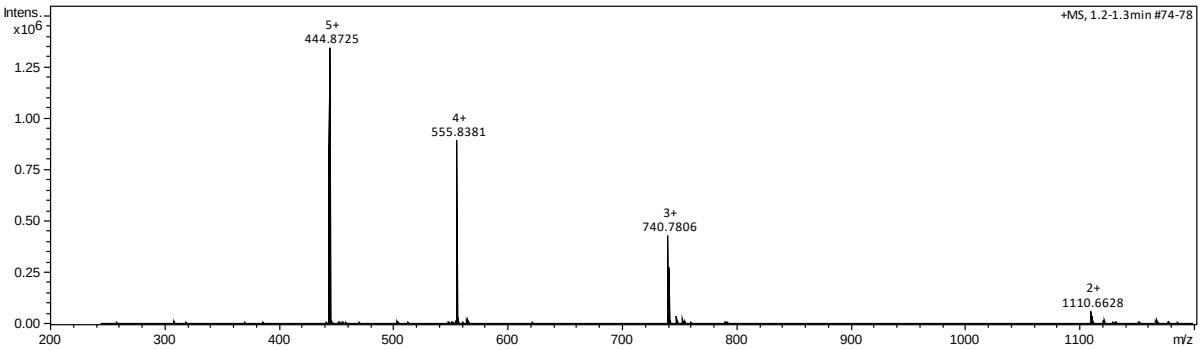
D83A NOXA-B


KAH2016 #98 [modified by chmhplc]
D-A NOXA 6a
DAD_Signal_B
mAU
700
500
375
250
125
0
-100
1 - 1.699
min
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.91

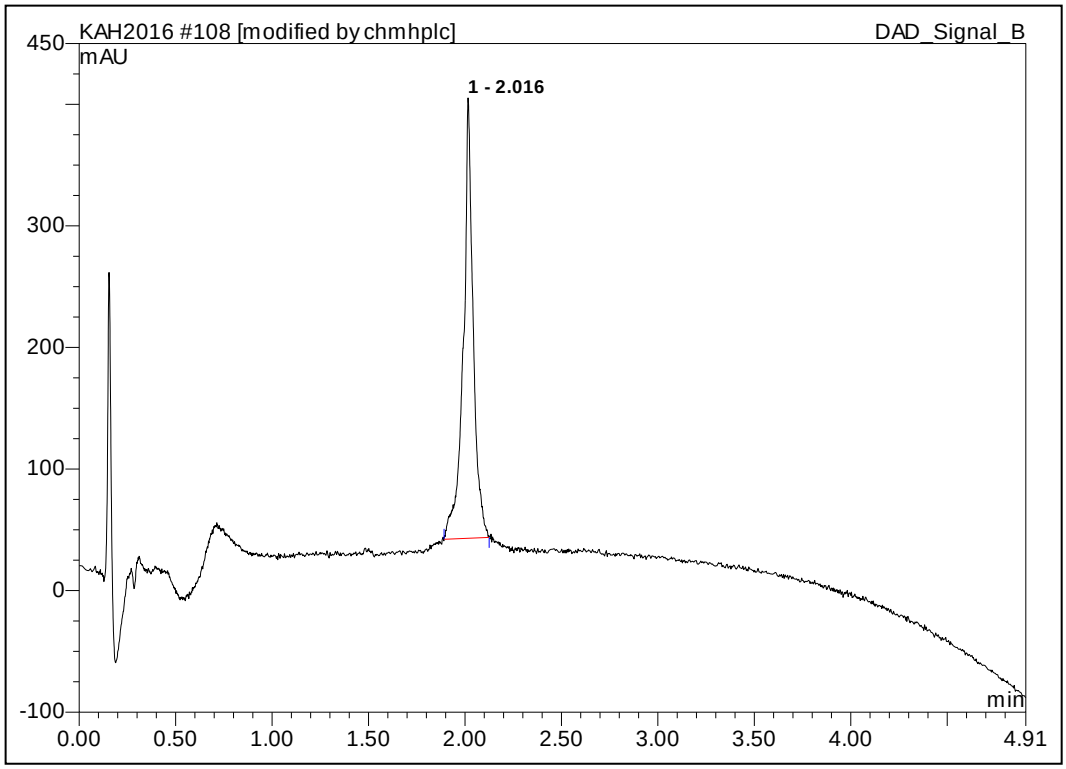

Intens.
x10^6
2.5
2.0
1.5
1.0
0.5
0.0
+MS, 1.2-1.3min #73-80
5+
441.6816
4+
551.8484
3+
735.4612
200
300
400
500
600
700
800
900
1000
1100
m/z

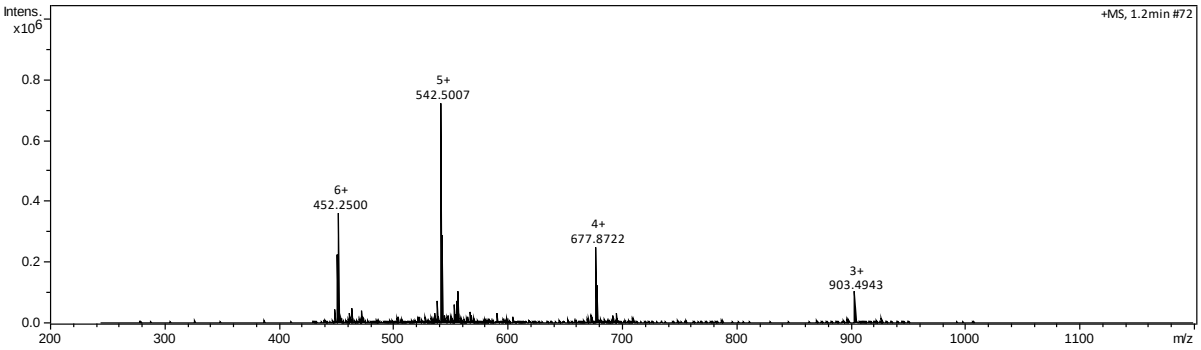
V85A NOXA-B


KAH2016 #147 [modified by chmhp V-A NOXA DAD_Signal_E
mAU
1,100
800
600
400
200
-100
1 - 1.601
2 - 2.593
min
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.90

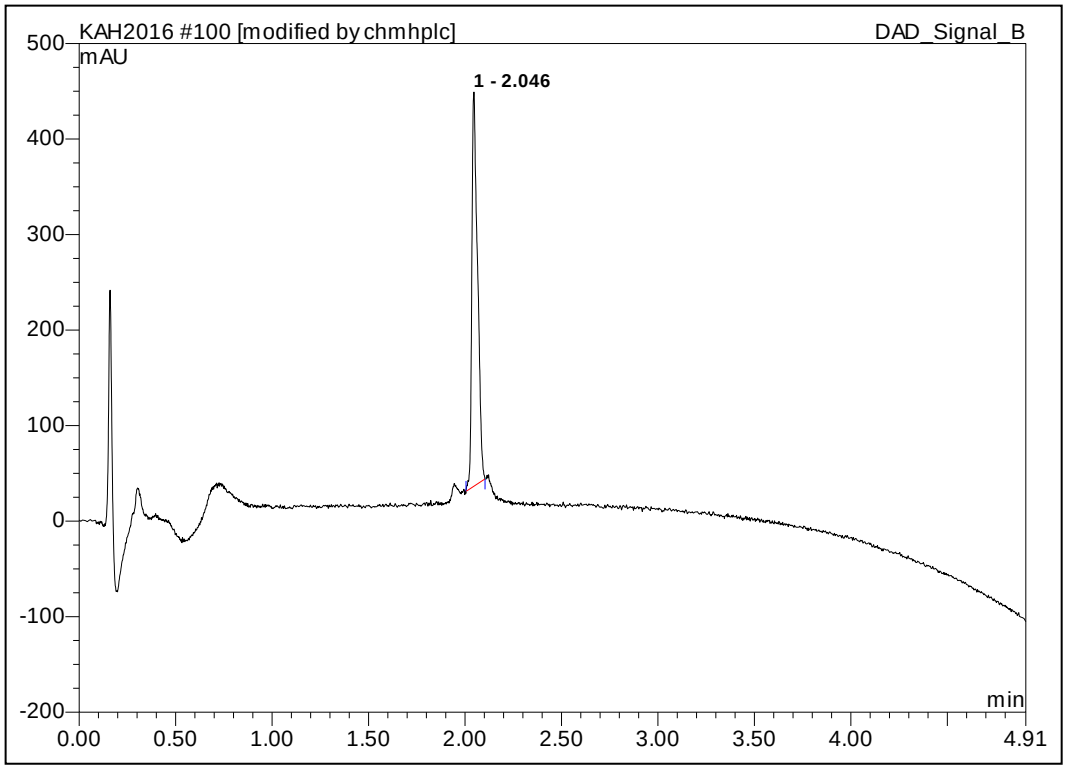

Intens.
x10^6
1.25
1.00
0.75
0.50
0.25
0.00
+MS, 1.2-1.3min #74-78
5+
444.8725
4+
555.8381
3+
740.7806
2+
1110.6628
200
300
400
500
600
700
800
900
1000
1100
m/z

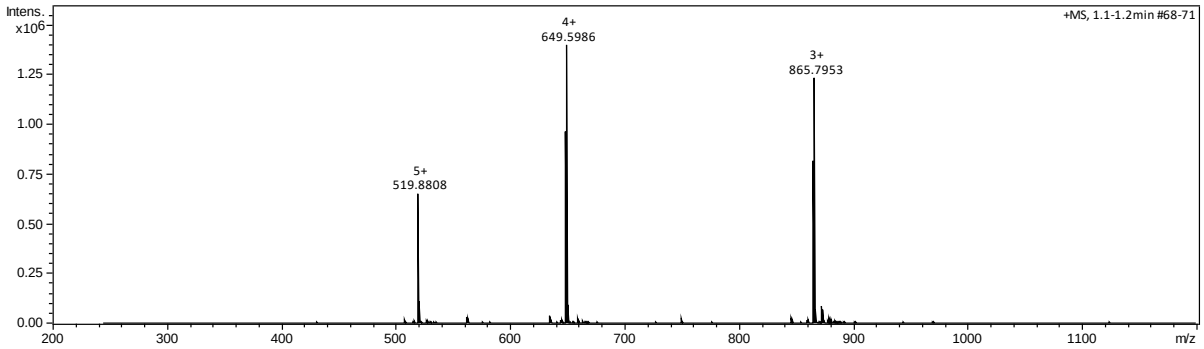
FITC-wt NOXA-B



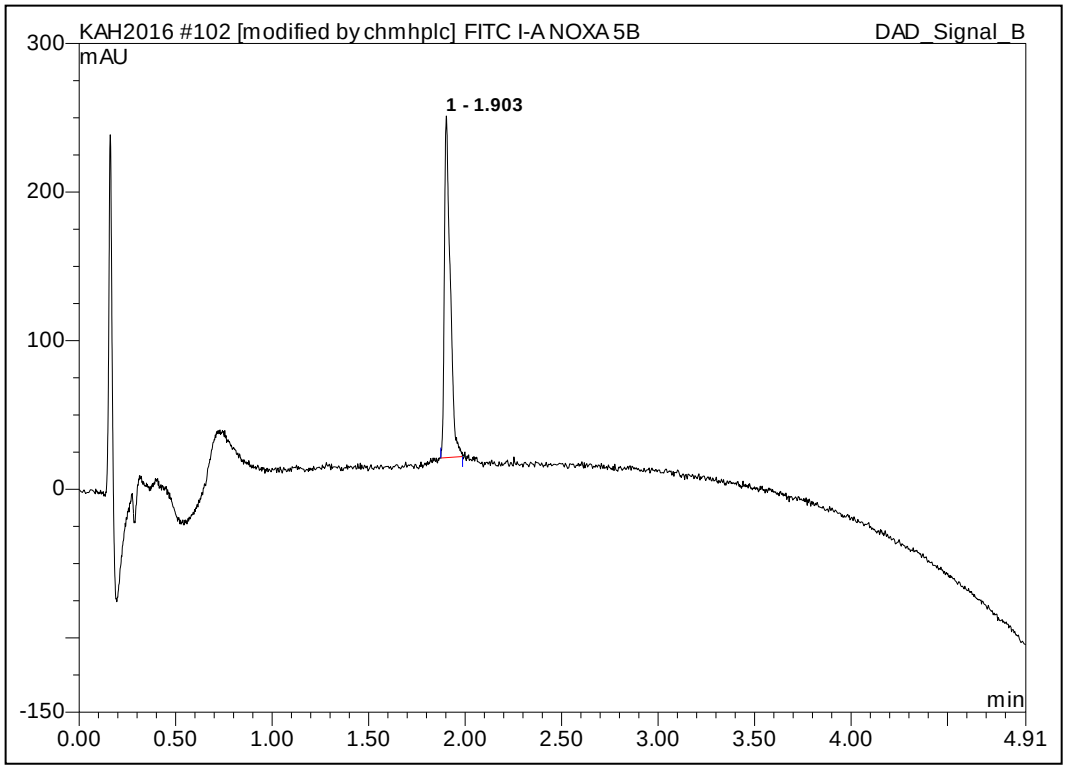


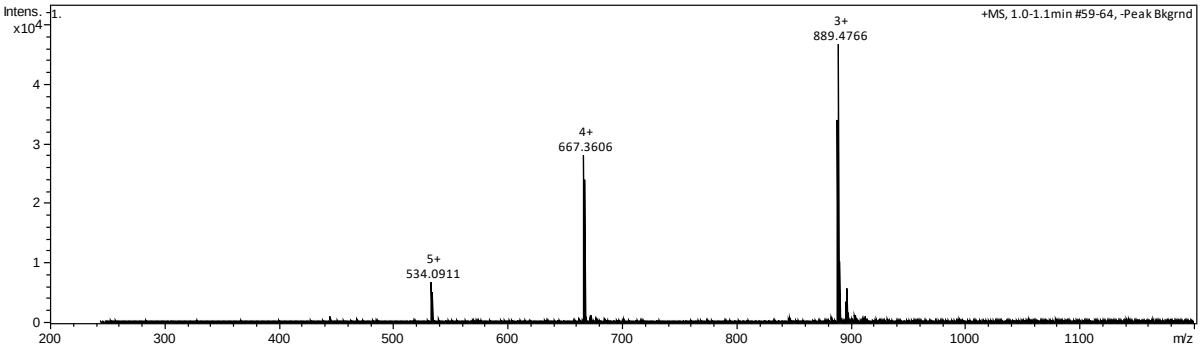
FITC-R79A NOXA-B



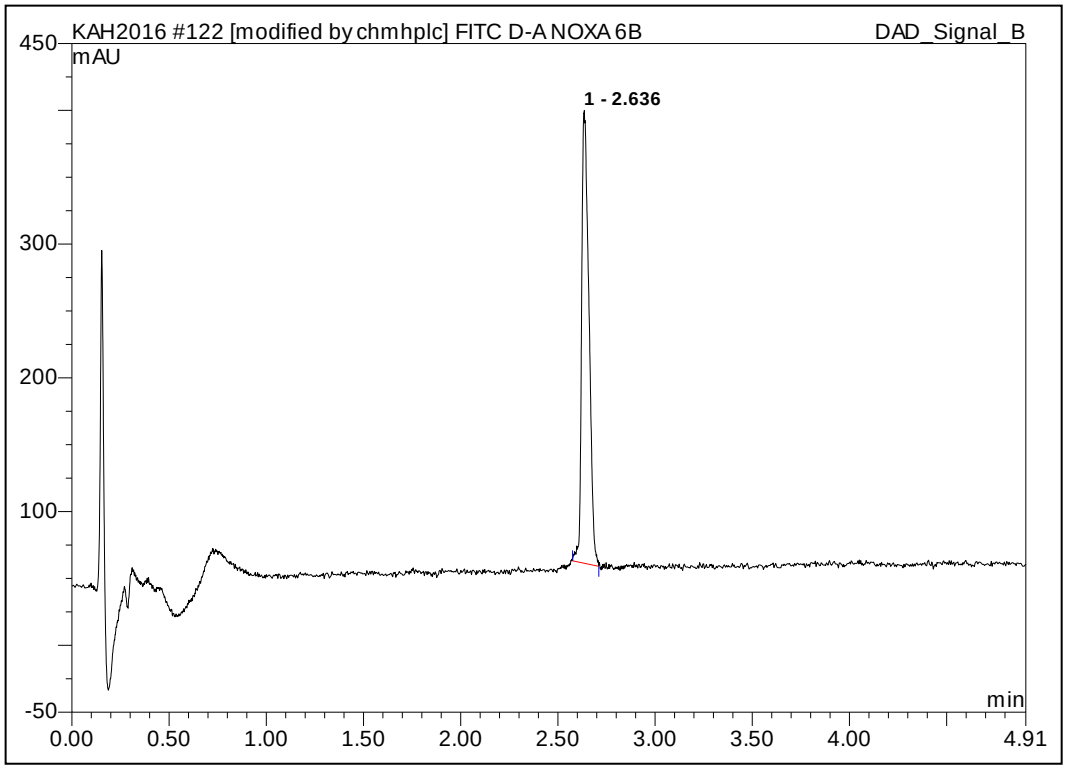


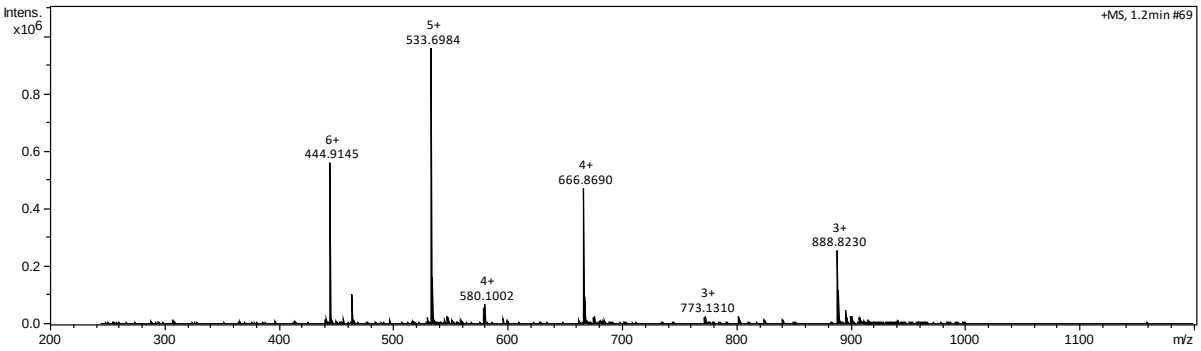
FITC-I81A NOXA-B





FITC-D83A NOXA-B





GKAP and Variant Peptides

Tabulated HRMS data of synthesised peptides are shown in the table below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp[d]) and Observed (Obs[d]) masses.

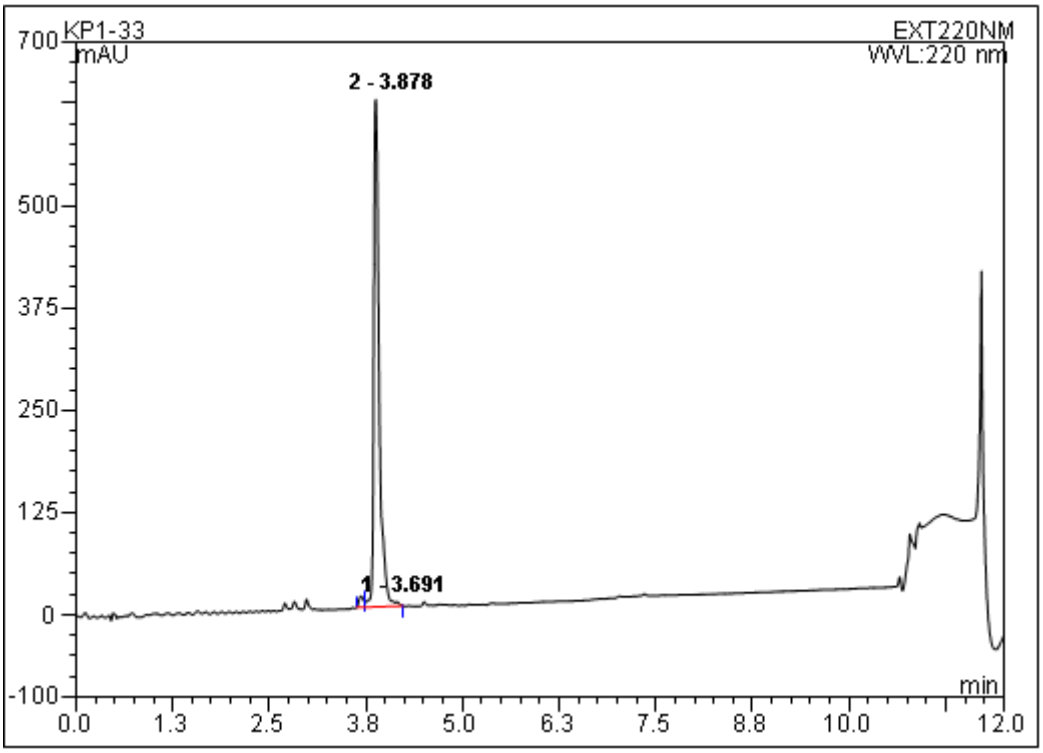
Peptide	Sequence
wt GKAP	FITC-(Ahx)EAQTRL-OH
	Ac-EAQTRL-OH
E1A	Ac-**A**AQTRL-OH
A2G	Ac-E**G**QTRL-OH
Q3A	Ac-EA**A**TRL-OH
T4A	Ac-EAQ**A**RL-OH
R5A	Ac-EAQT**A**L-OH
$CONH_2$	Ac-EAQTRL-**NH_2**

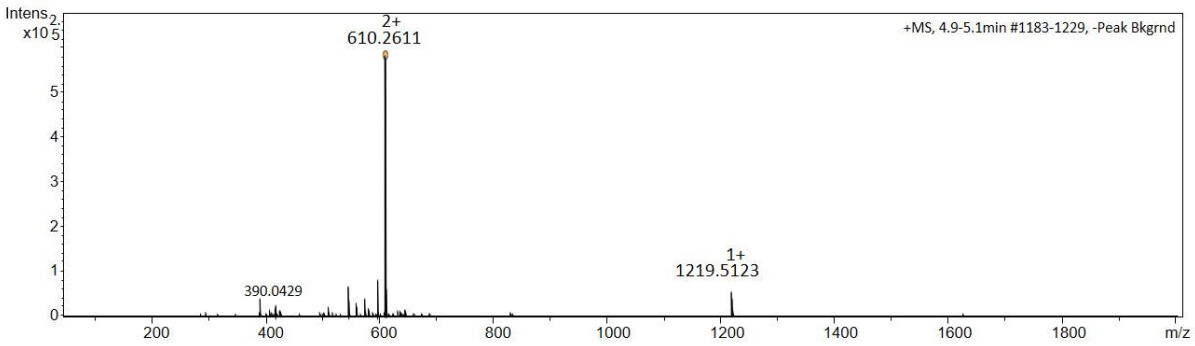
Acetylated GKAP peptides

Peptide	$[M+1H]^{1+}$ Obs[d]	$[M+1H]^{1+}$ Exp[d]
FITC-*wt*-GKAP	1219.5123	
wt GKAP	759.3994	759.3923
E1A GKAP	701.3949	701.3868
A2G GKAP	744.3778	745.3766
Q3A GKAP	702.3777	702.3708
T4A GKAP	729.3873	729.3817
R5A GKAP	674.3365	674.3283
L6A GKAP	717.3532	717.3453
$CONH_2$ GKAP	757.4072	758.4083

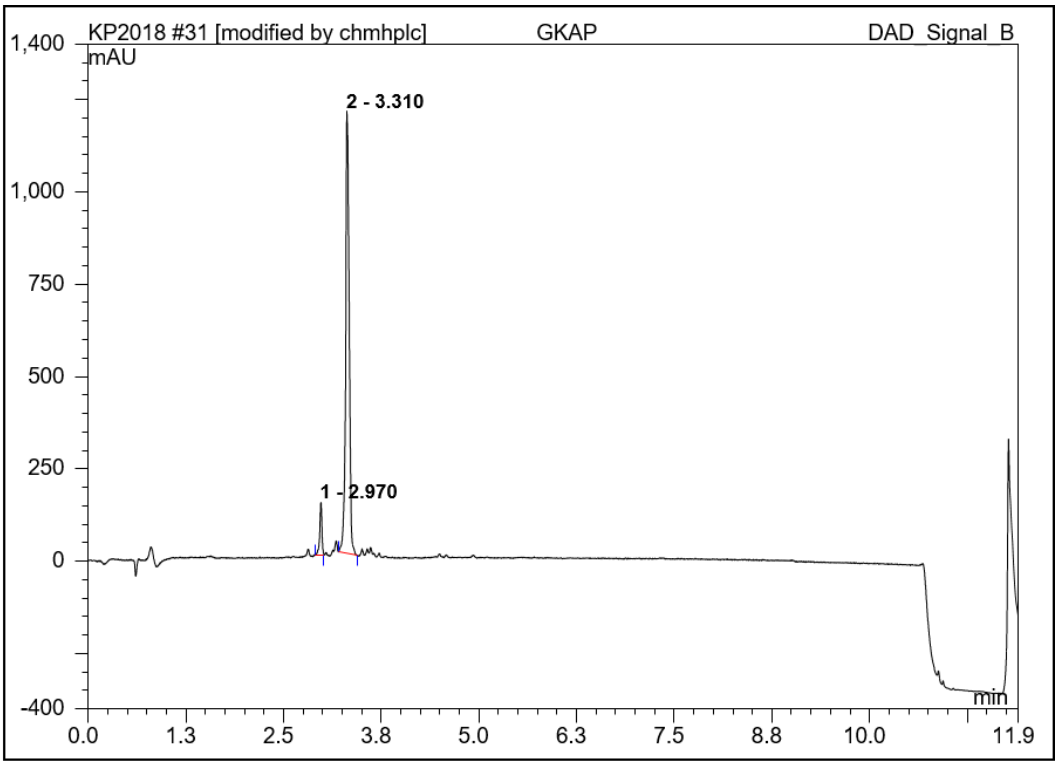
HPLC and ESI-MS data for GKAP and variant peptides

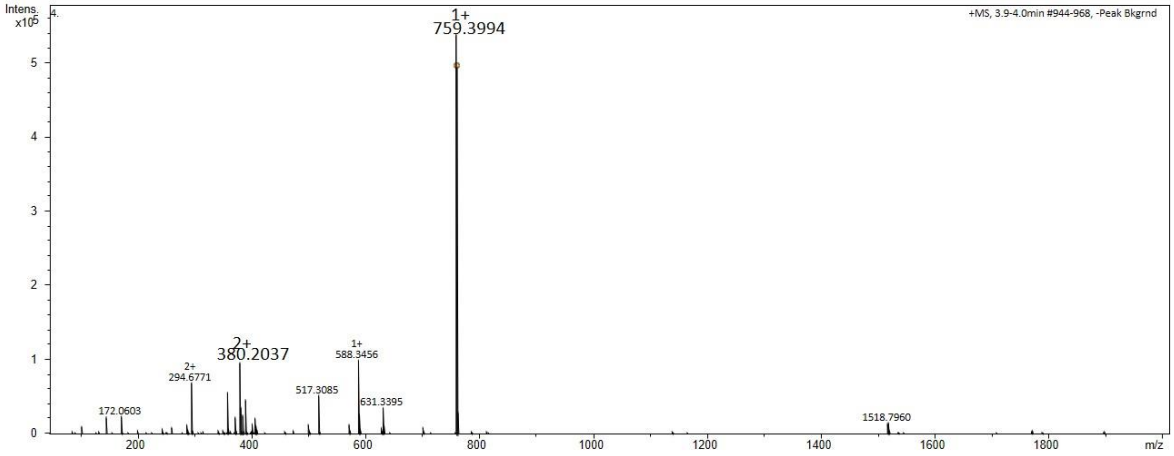
FITC-*wt*-GKAP



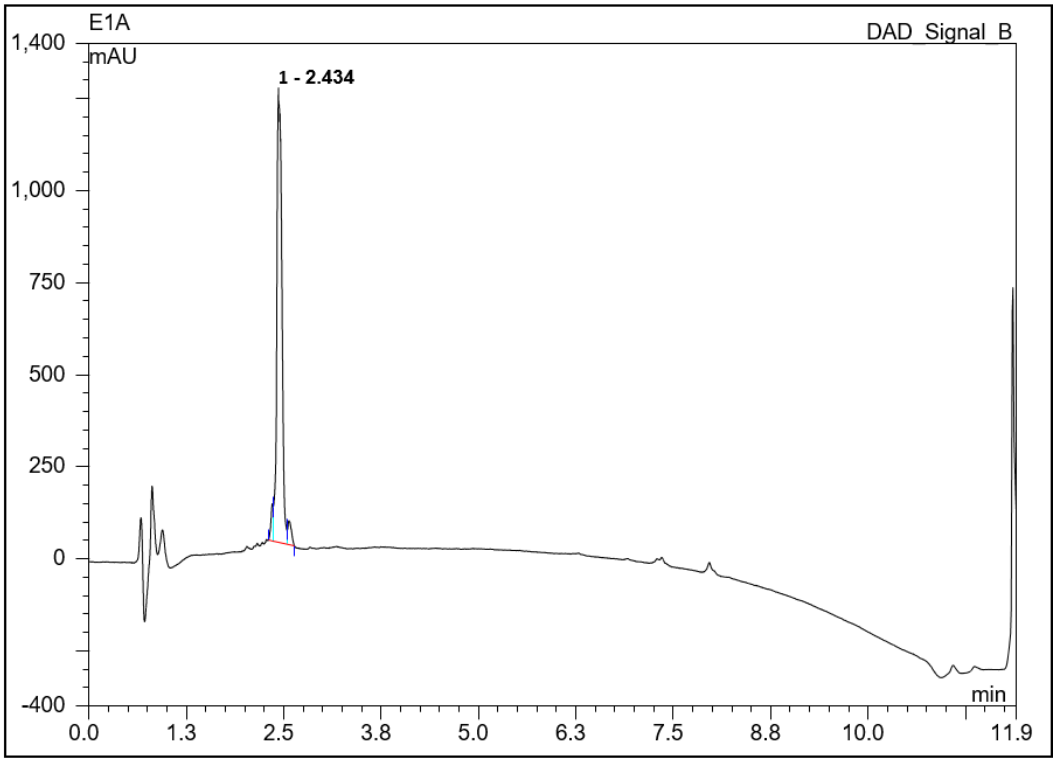


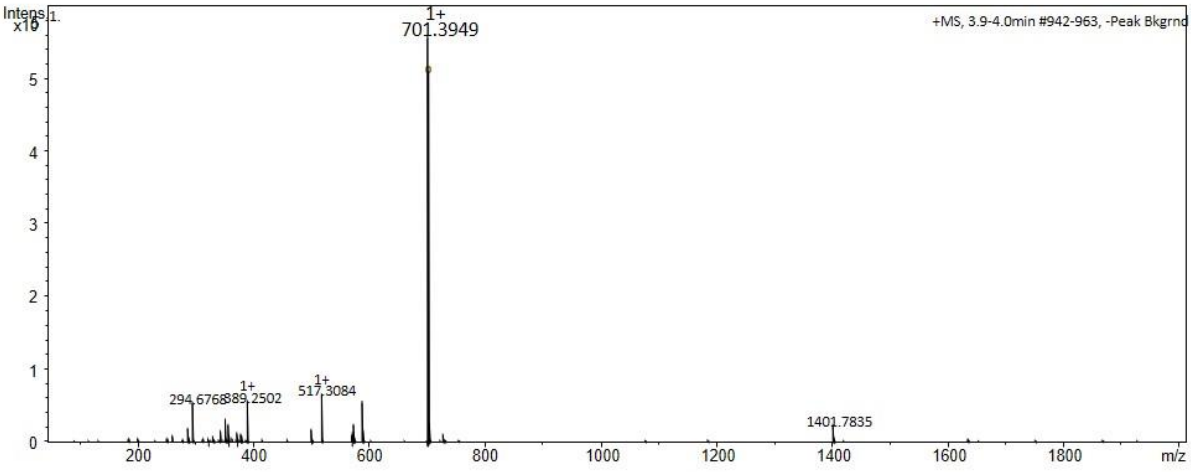
wt-GKAP


KP2018 #31 [modified by chmhplc]
GKAP
DAD_Signal_B
mAU
2 - 3.310
1 - 2.970
min
1,400
1,000
750
500
250
0
-400
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

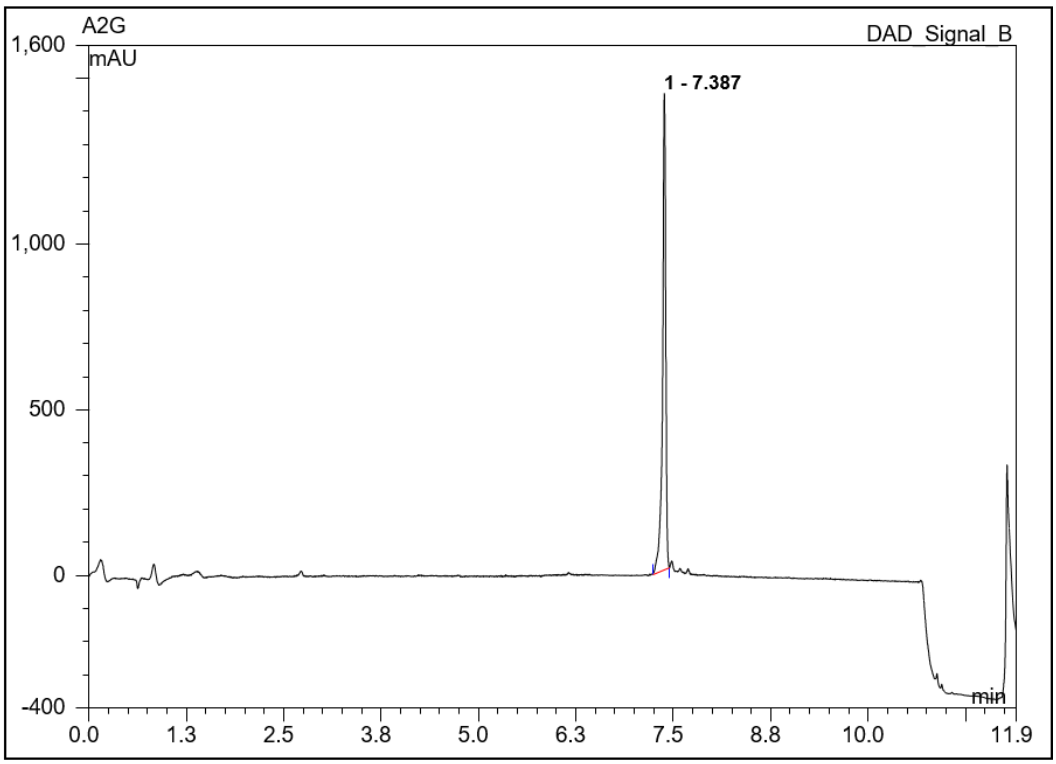

Intens.
x10^5
+MS, 3.9-4.0min #944-968, -Peak Bkgrnd
1+
759.3994
2+
380.2037
2+
294.6771
1+
588.3456
517.3085
631.3395
172.0603
1518.7960
m/z

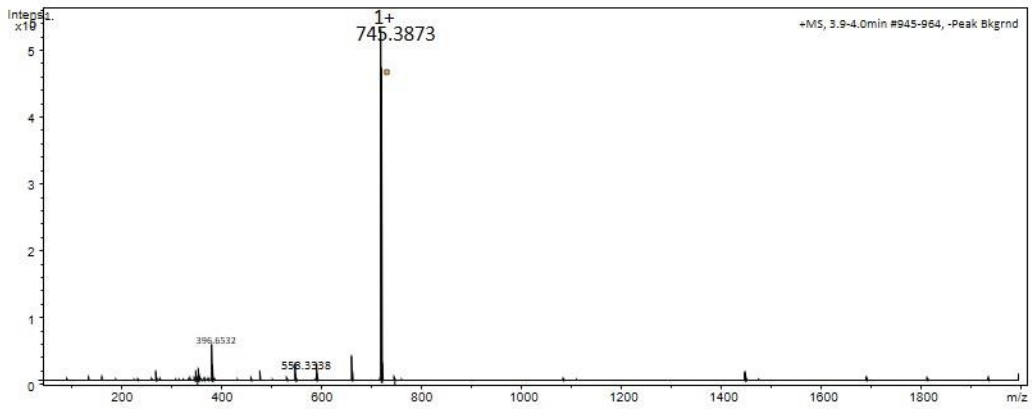
E1A GKAP


E1A
DAD_Signal_B
mAU
1 - 2.434
1,400
1,000
750
500
250
0
-400
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

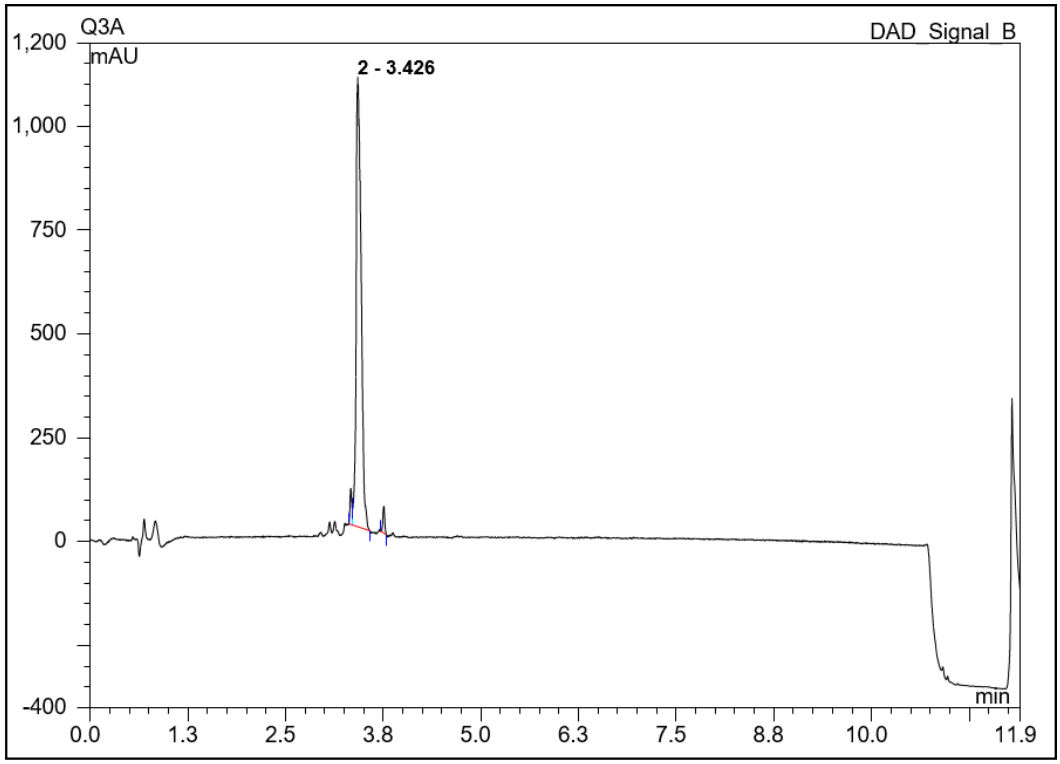

Intens.
x10
1+
701.3949
+MS, 3.9-4.0min #942-963, -Peak Bkgrnd
294.6768
1+
889.2502
1+
517.3084
1401.7835
m/z

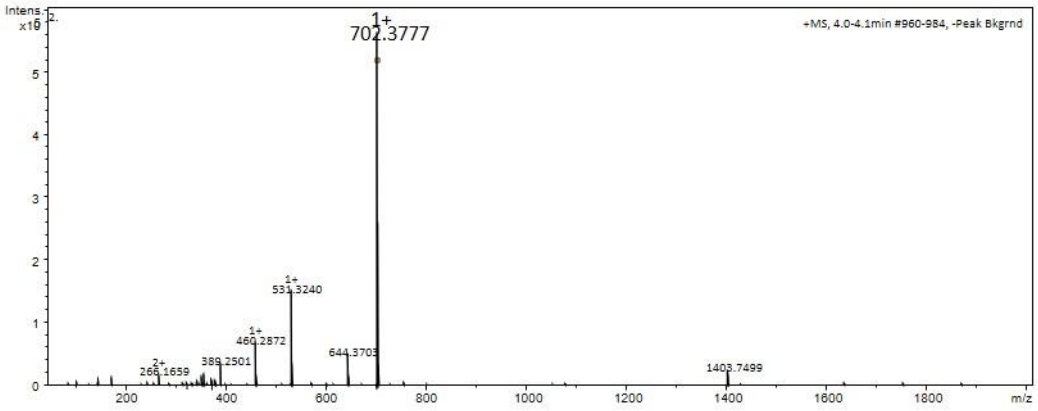
A2G GKAP


A2G
DAD_Signal_B
mAU
1,600
1,000
500
0
-400
1 - 7.387
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

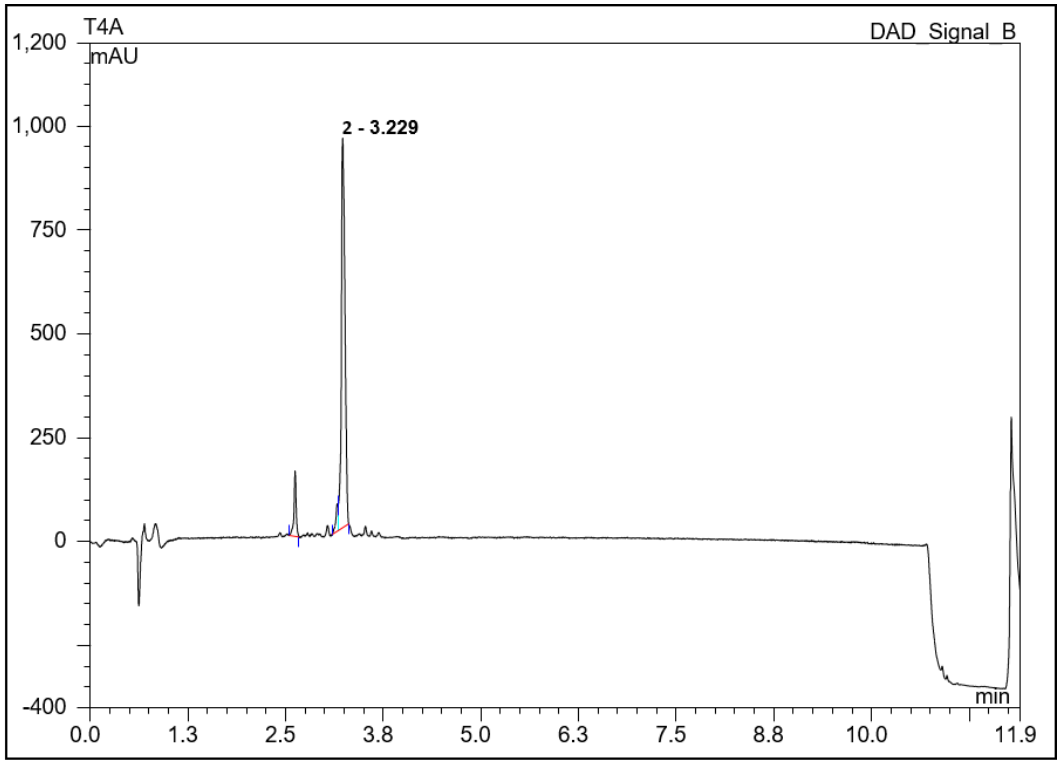

Intens.
x10
5
4
3
2
1
0
1+
745.3873
+MS, 3.9-4.0min #945-964, -Peak Bkgrnd
396.6532
556.3338
200
400
600
800
1000
1200
1400
1600
1800
m/z

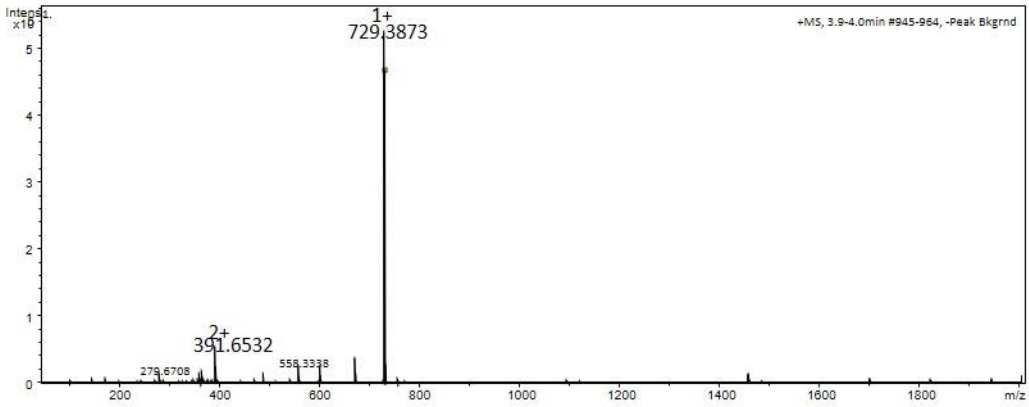
Q3A GKAP


Q3A
DAD Signal B
mAU
2 - 3.426
1,200
1,000
750
500
250
0
-400
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

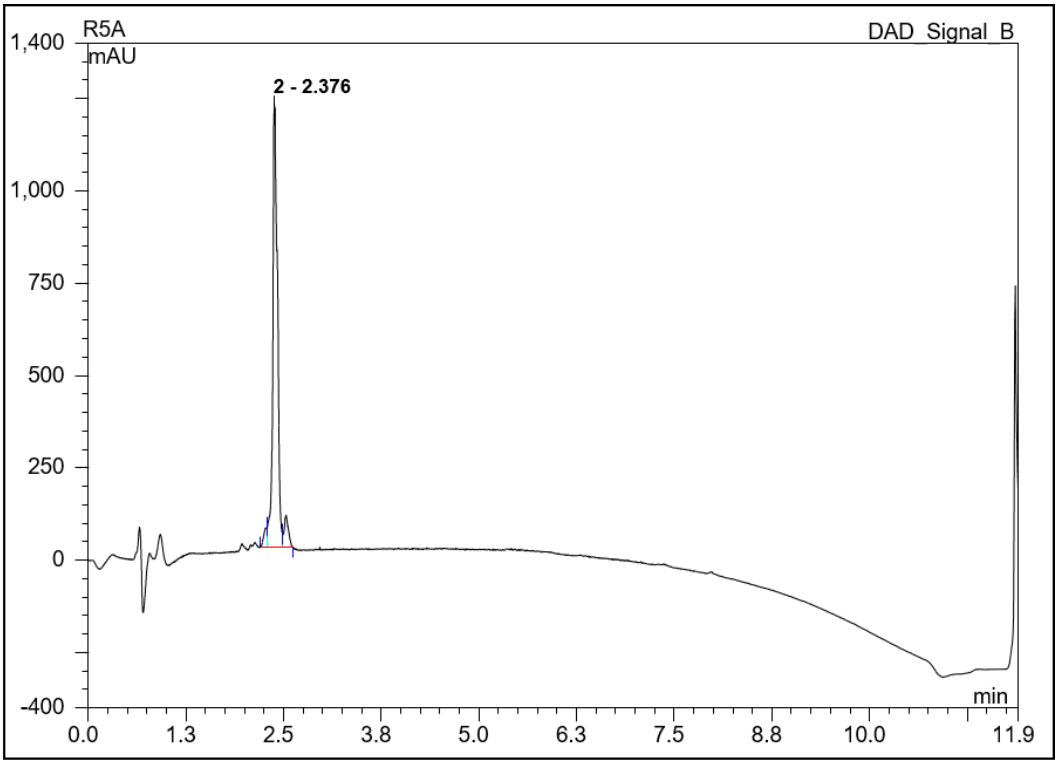

Intens.
x10
+MS, 4.0-4.1min #960-984, -Peak Bkgrnd
1+
702.3777
1+
531.3240
1+
460.2872
644.3703
389.2501
2+
266.1659
1403.7499
m/z

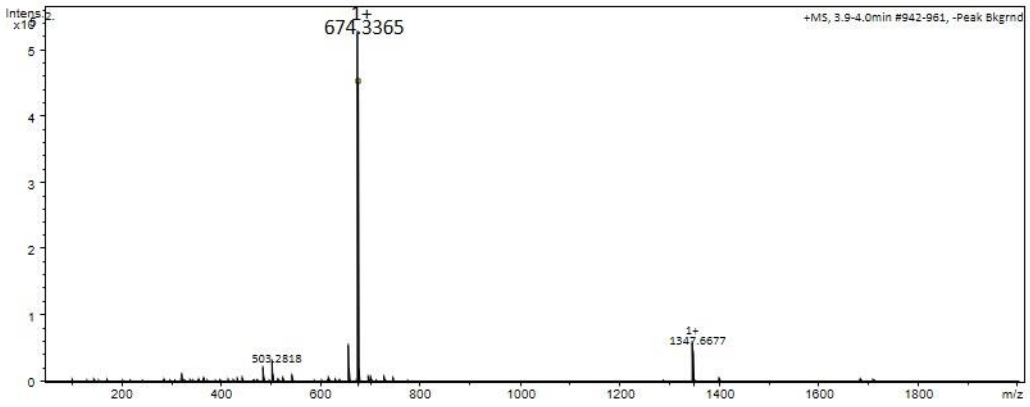
T4A GKAP


T4A
mAU
DAD Signal B
2 - 3.229
1,200
1,000
750
500
250
0
-400
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

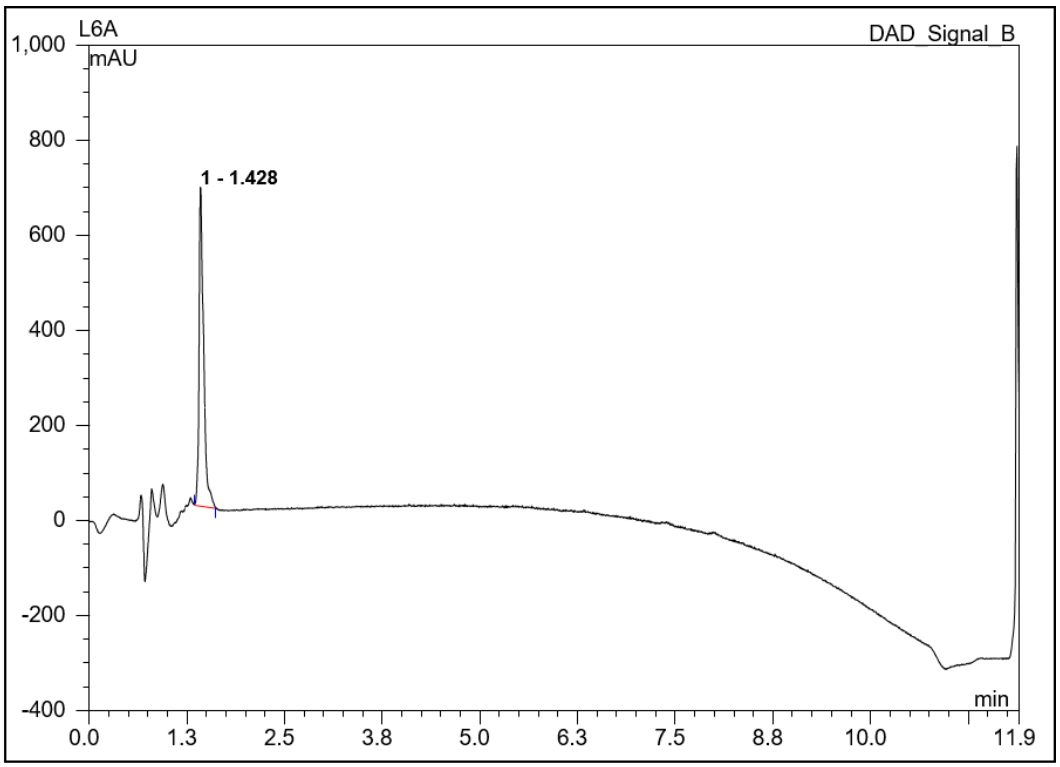

Intens.
x10
+MS, 3.9-4.0min #945-964, -Peak Bkgrnd
1+
729.3873
2+
391.6532
279.6708
558.3338
m/z

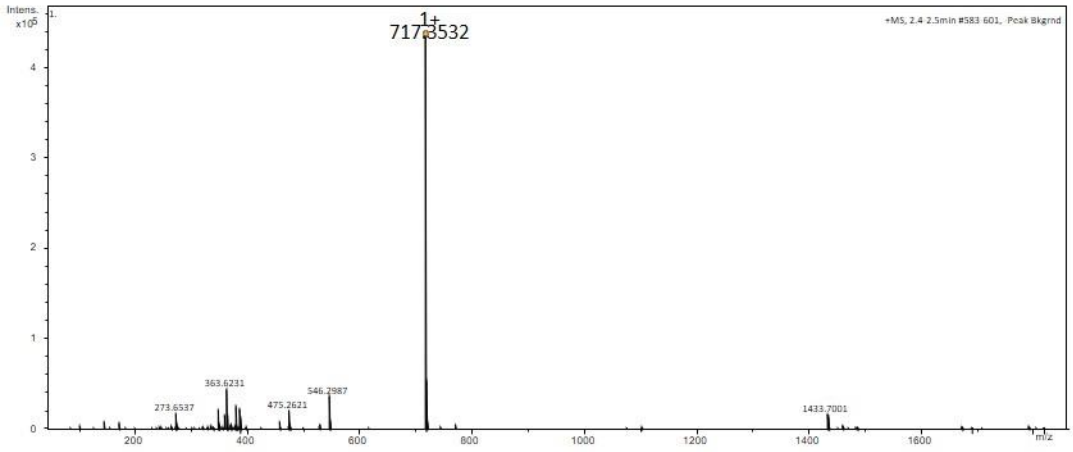
R5A GKAP


R5A
DAD Signal B
mAU
2 - 2.376
1,400
1,000
750
500
250
0
-400
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

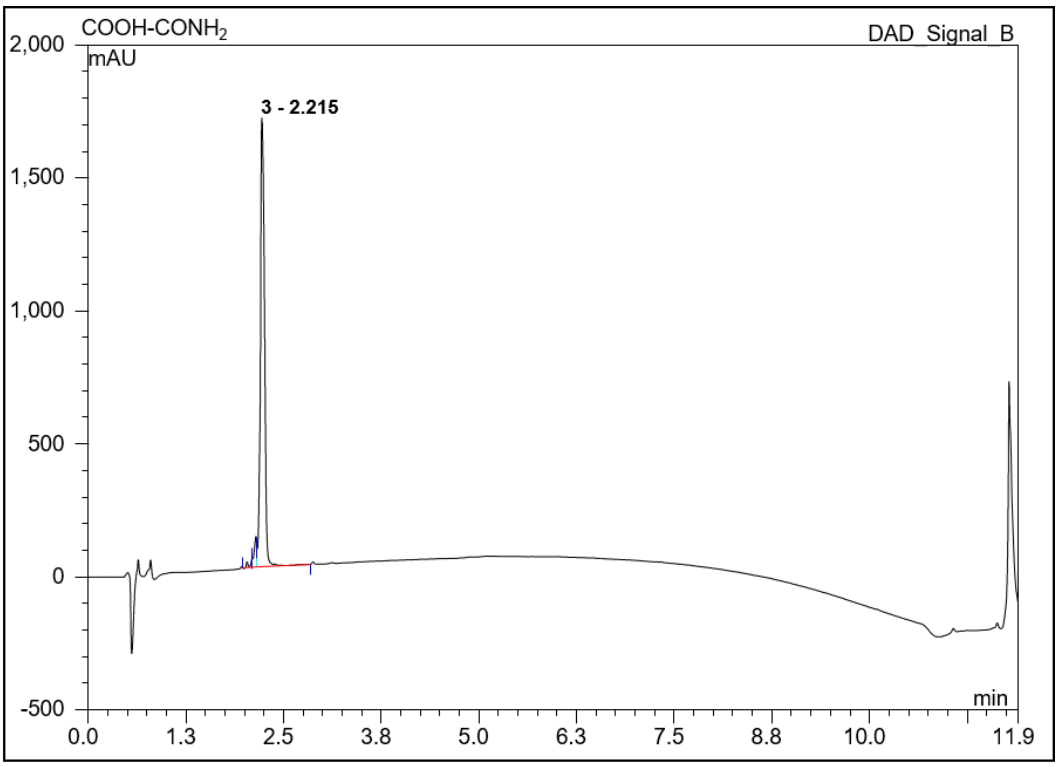

Intens.
x10
1+
674.3365
+MS, 3.9-4.0min #942-961, -Peak Bkgrnd
503.2818
1+
1347.6677
m/z

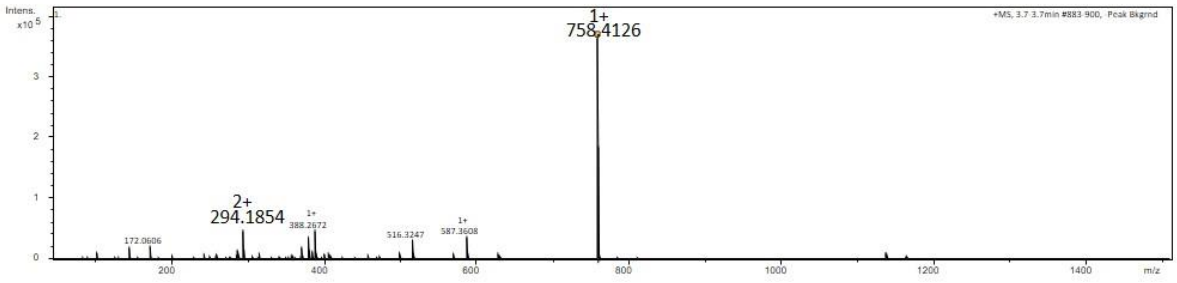
L6A GKAP


L6A
DAD_Signal_B
mAU
1,000
800
600
400
200
0
-200
-400
1 - 1.428
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9


Intens.
x10^5
+MS, 2.4-2.5min #583-601, Peak Bkgrnd
1+
717.3532
363.6231
273.6537
475.2621
546.2987
1433.7001
m/z

COOH-CONH$_2$ GKAP





SIMS and Variant Peptides

Tabulated HRMS data of synthesised peptides are shown in below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp[d]) and Observed (Obs[d]) masses.

FAM SIM Peptide	Sequence	Purity (%)
FAM-*wt* SIM	FAM-PEG-DNEIEVIIVWEKK-NH_2	100
FAM-D2705A SIM	FAM-PEG-ANEIEVIIVWEKK-NH_2	100
FAM-N2706A SIM	FAM-PEG-DAEIEVIIVWEKK-NH_2	93.9
FAM-E2707A SIM	FAM-PEG-DNAIEVIIVWEKK-NH_2	84.5
FAM-I2708A SIM	FAM-PEG-DNEAEVIIVWEKK-NH_2	96.3
FAM-E2709A SIM	FAM-PEG-DNEIAVIIVWEKK-NH_2	97.4
FAM-V2710A SIM	FAM-PEG-DNEIEAIIVWEKK-NH_2	98.4
FAM-I2711A SIM	FAM-PEG-DNEIEVAIVWEKK-NH_2	93.9
FAM-I2712A SIM	FAM-PEG-DNEIEVIAVWEKK-NH_2	95.3
FAM-V2713A SIM	FAM-PEG-DNEIEVIIAWEKK-NH_2	100
FAM-W2714A SIM	FAM-PEG-DNEIEVIIVAEKK-NH_2	100
FAM-E2715A SIM	FAM-PEG-DNEIEVIIVWAKK-NH_2	100
FAM-K2716A SIM	FAM-PEG-DNEIEVIIVWEAK-NH_2	100
FAM-K2717A SIM	FAM-PEG-DNEIEVIIVWEKA-NH_2	100

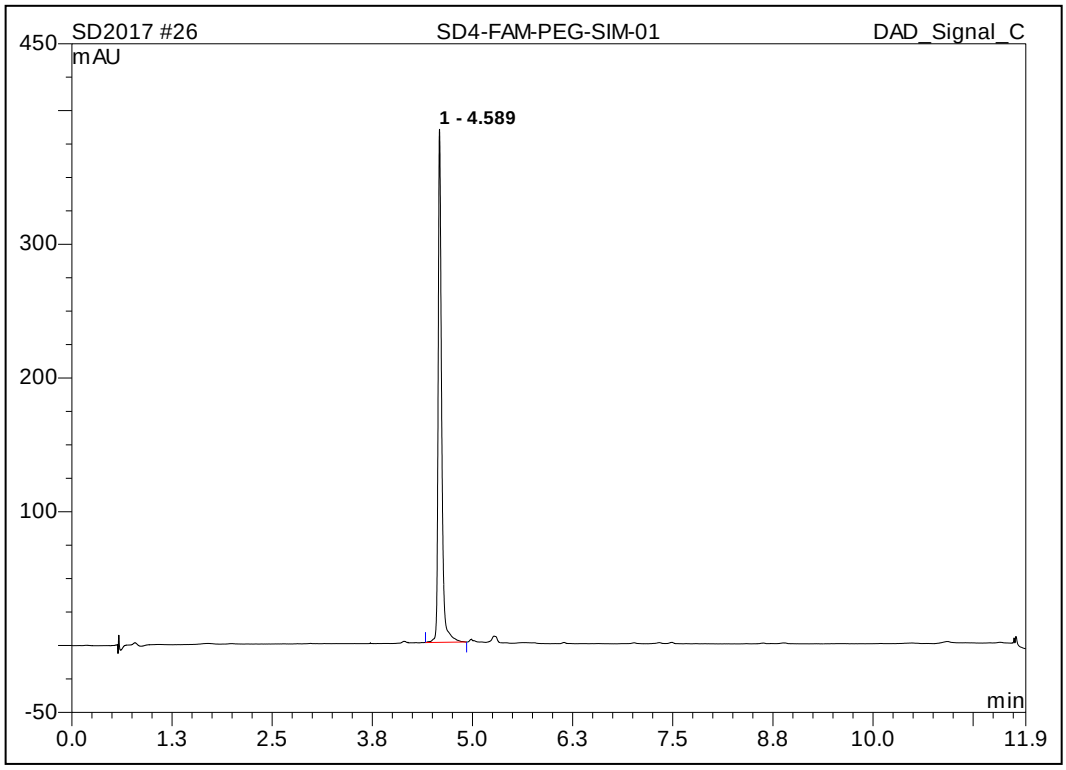
N-Acetyl Peptide	Sequence	Purity (%)
wt SIM	Ac-DNEIEVIIVWEKK-NH_2	100
E2707A SIM	Ac-DNAIEVIIVWEKK-NH_2	92.1
V2710A SIM	Ac-DNEIEAIIVWEKK-NH_2	96.4
I2712A SIM	Ac-DNEIEVIAVWEKK-NH_2	92.5
W2714S SIM	Ac-DNEIEVIIVSEKK-NH_2	100

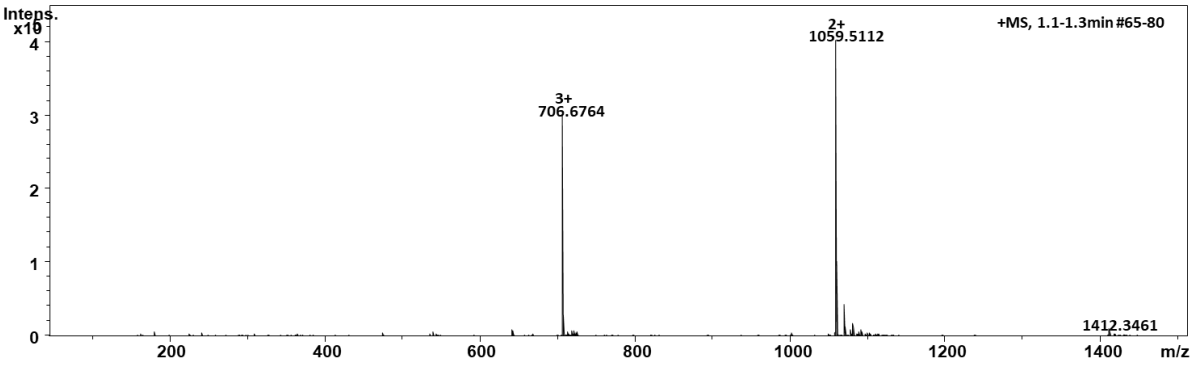
FAM SIM Peptide	$[M+2H]^{2+}$ Exp[d]	$[M+2H]^{2+}$ Obs[d]	$[M+3H]^{3+}$ Exp[d]	$[M+3H]^{3+}$ Obs[d]
FAM-SIM	1049.4954	1049.4915	706.6812	706.6636
FAM-D2705A-SIM	1037.527	1037.592	692.0180	692.0219
FAM-N2706A-SIM	1038.019	1037.9485	692.346	692.3031
FAM-E2707A-SIM	1030.5191	1030.4559	687.3461	687.3073
FAM-I2708A-SIM	1038.4984	1038.4792	692.6656	692.6543
FAM-E2709A-SIM	1030.5191	1030.5085	687.3461	687.3423
FAM-V2710A-SIM	1045.4877	1045.5062	697.3375	697.3277
FAM-I2711A-SIM	1038.4984	1038.4885	692.6656	692.6615
FAM-I2712A-SIM	1038.4984	1038.4807	692.6656	692.6566
FAM-V2713A-SIM	1045.4877	1045.4878	697.3375	697.3266
FAM-W2714A-SIM	1002.0008	1001.9829	668.3338	668.3234
FAM-E2715A-SIM	1030.5191	1030.5009	687.3461	687.3349
FAM-K2716A-SIM	1030.9930	1030.9737		
FAM-K2717A-SIM	1030.9930	1030.9739		

Ac SIM Peptide	$[M+2H]^{2+}$ Exp[d]	$[M+2H]^{2+}$ Obs[d]
SIM	828.4543	828.4548
E2707A-SIM	799.4515	799.4527
V2710A-SIM	814.4386	814.4400
I2712A-SIM	807.4308	807.4316
W2714A-SIM	770.9332	770.9335
W2714S-SIM	778.9306	778.9310

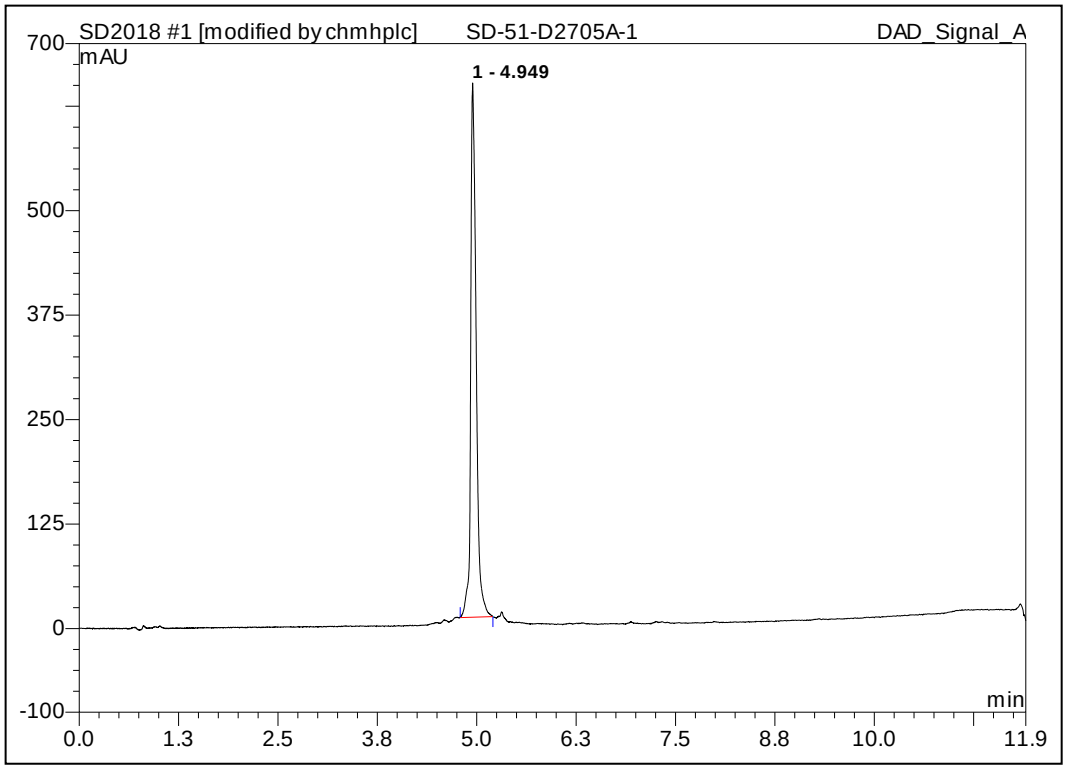
HPLC and ESI-MS data for SIM and variant peptides

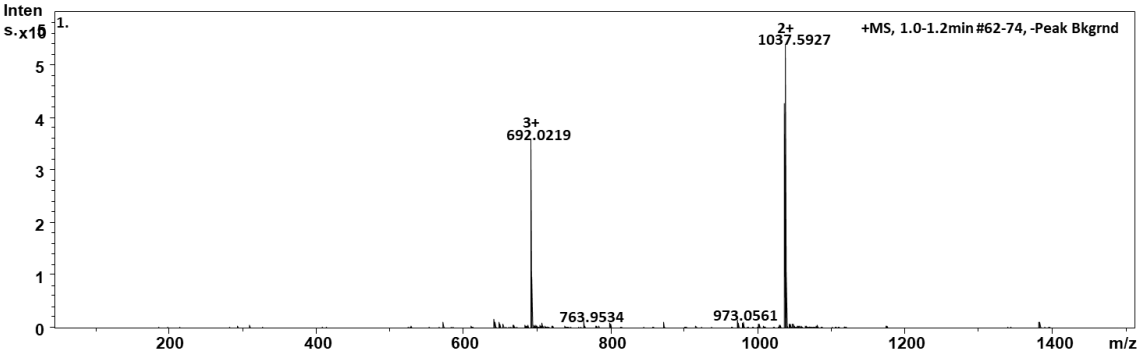
FAM-wt SIM



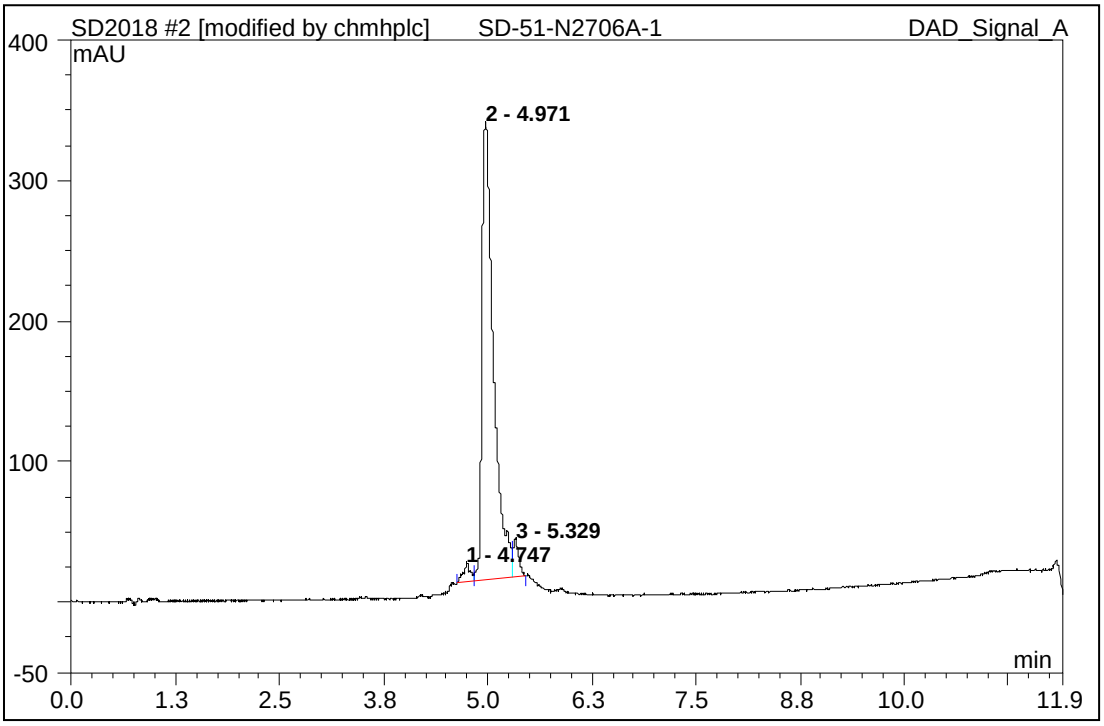


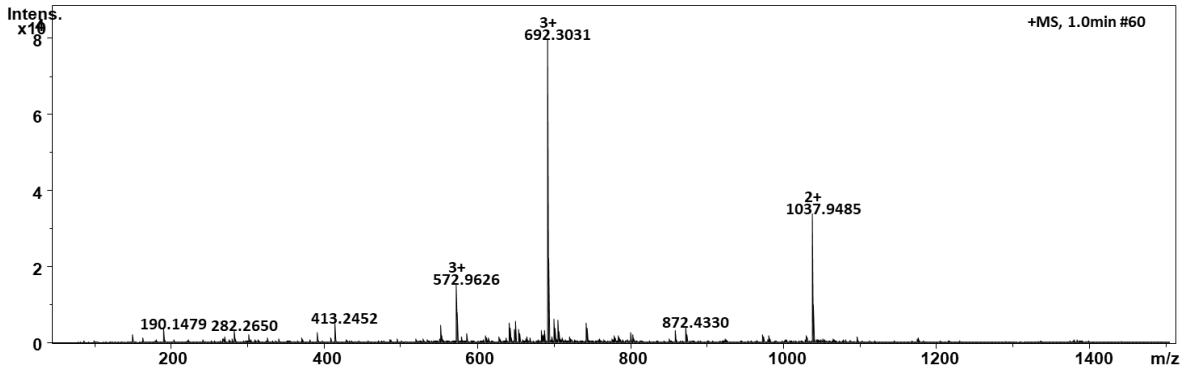
FAM-D2705A SIM



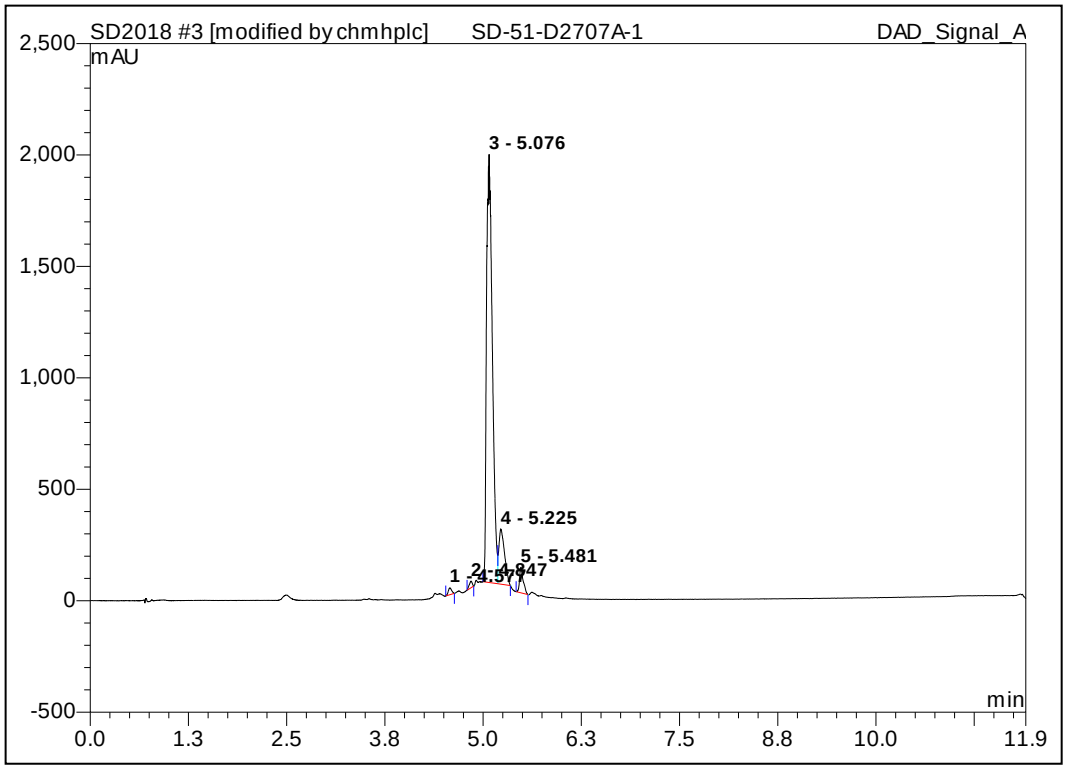


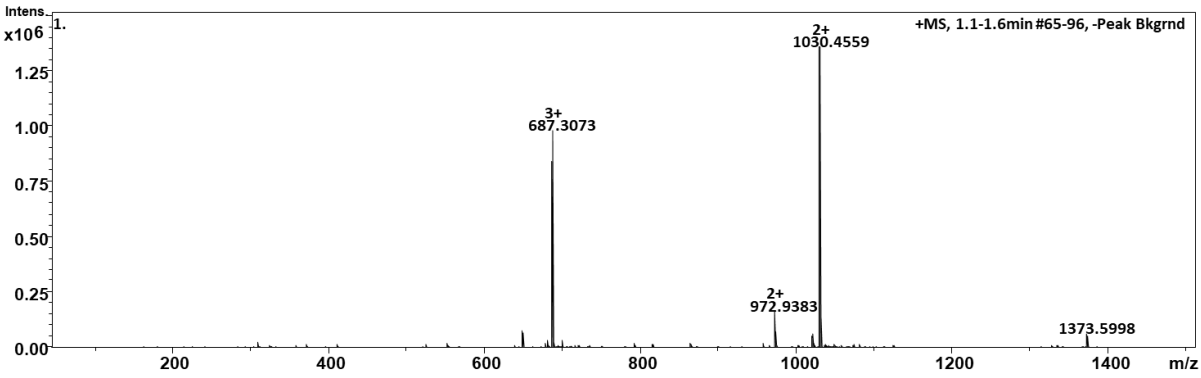
FAM-N2706A SIM



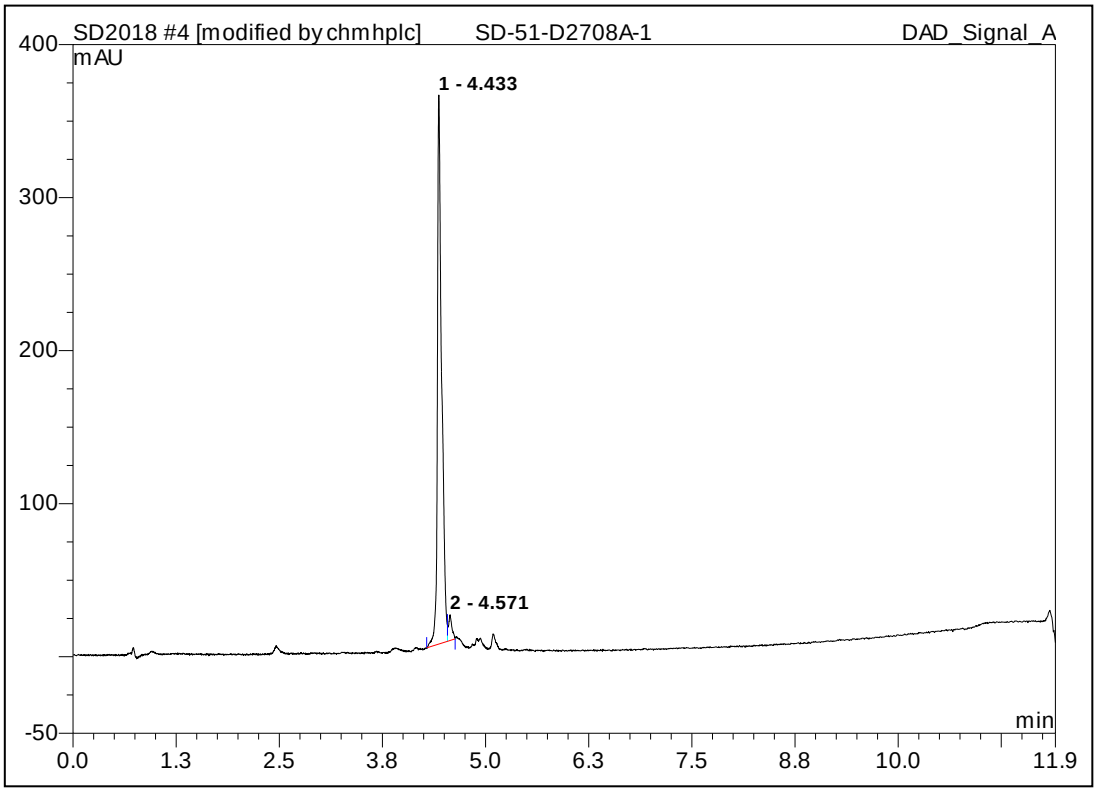


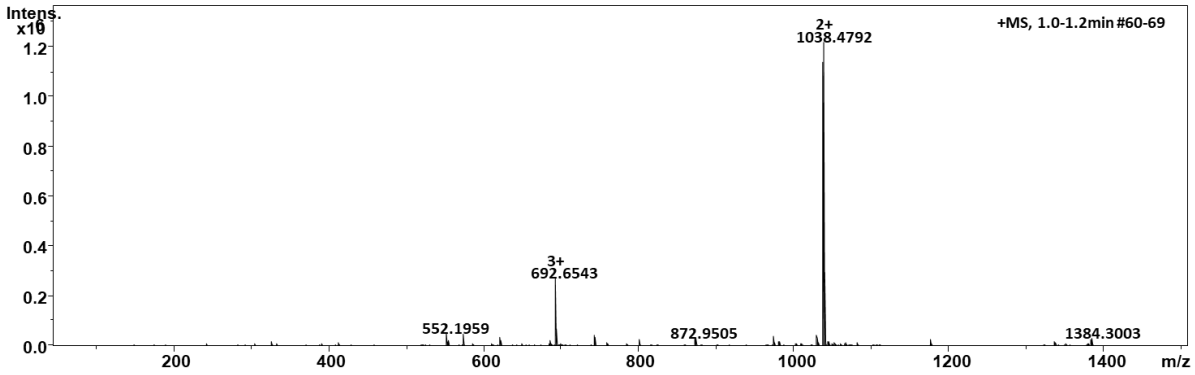
FAM-E2707A SIM



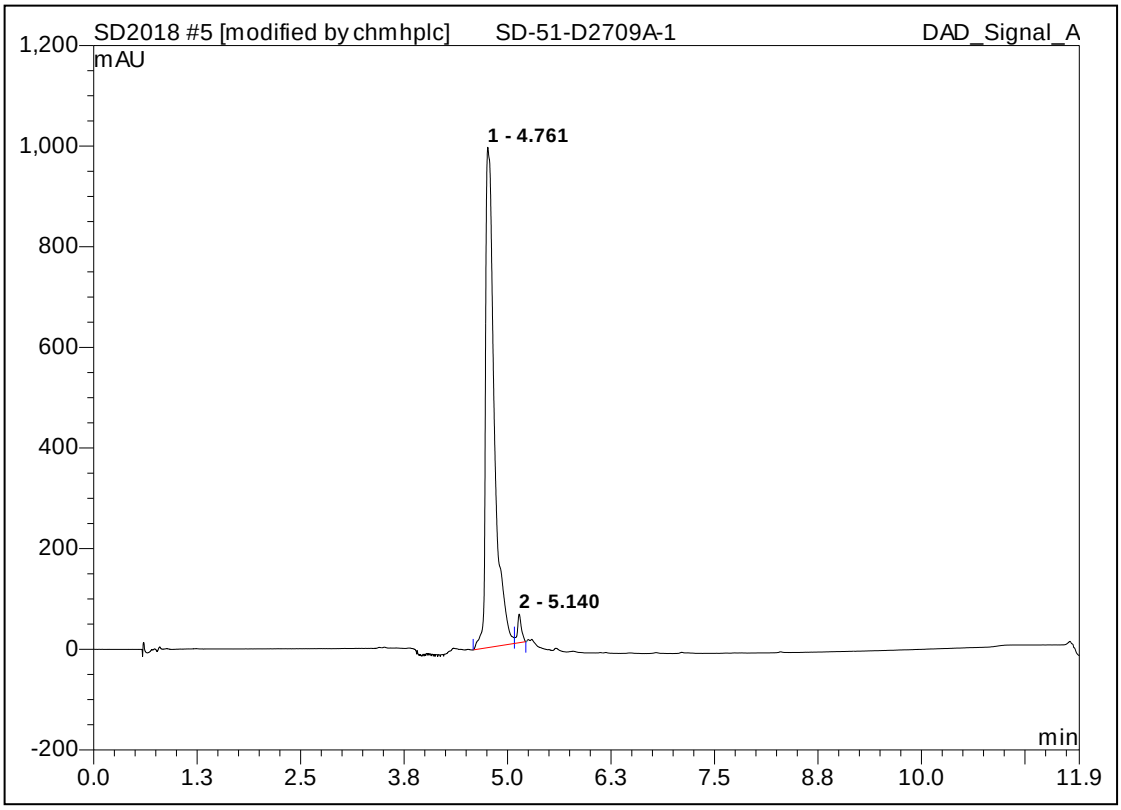


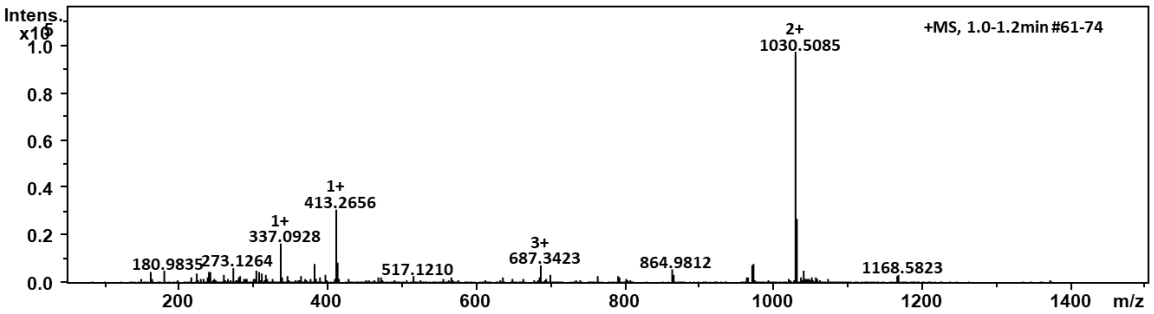
FAM-I2708A SIM



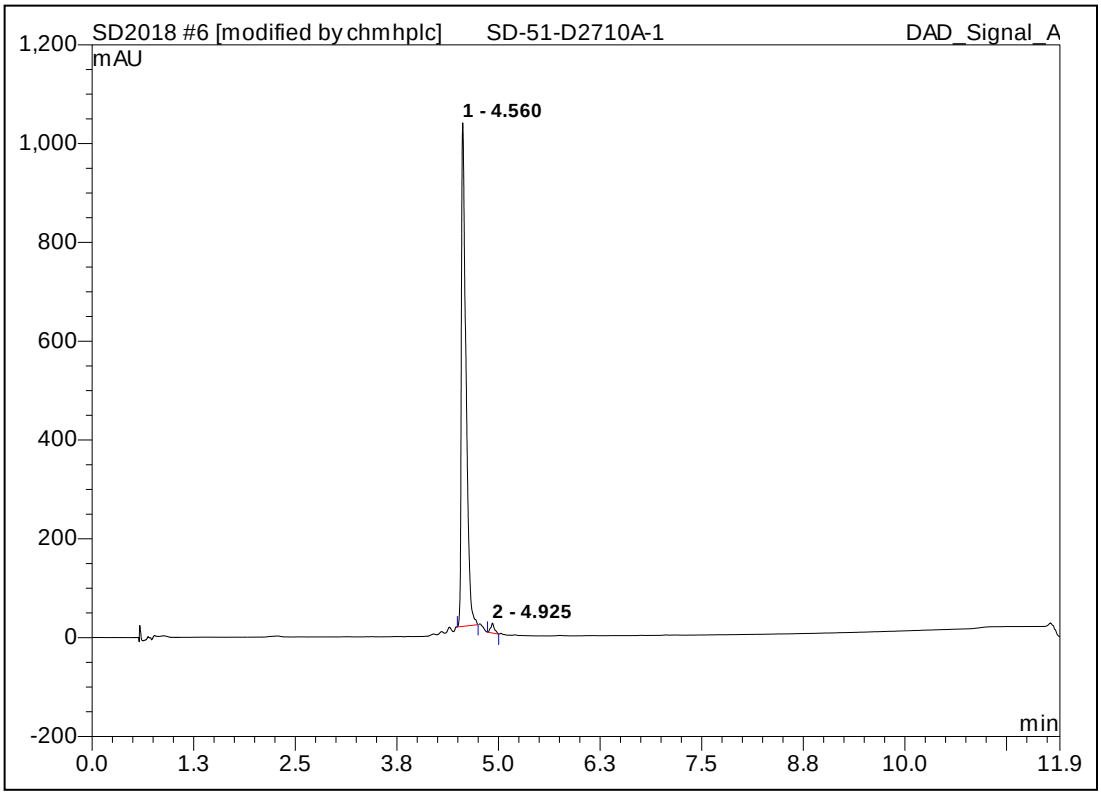


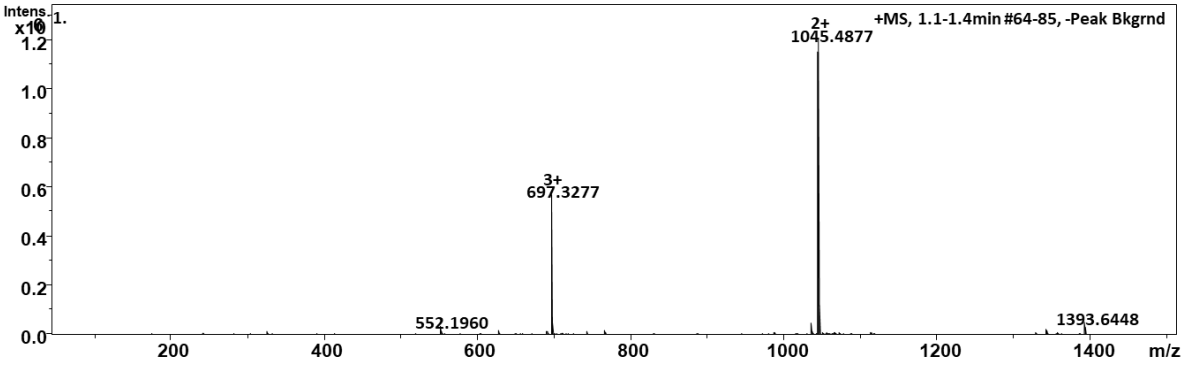
FAM-E2709A SIM



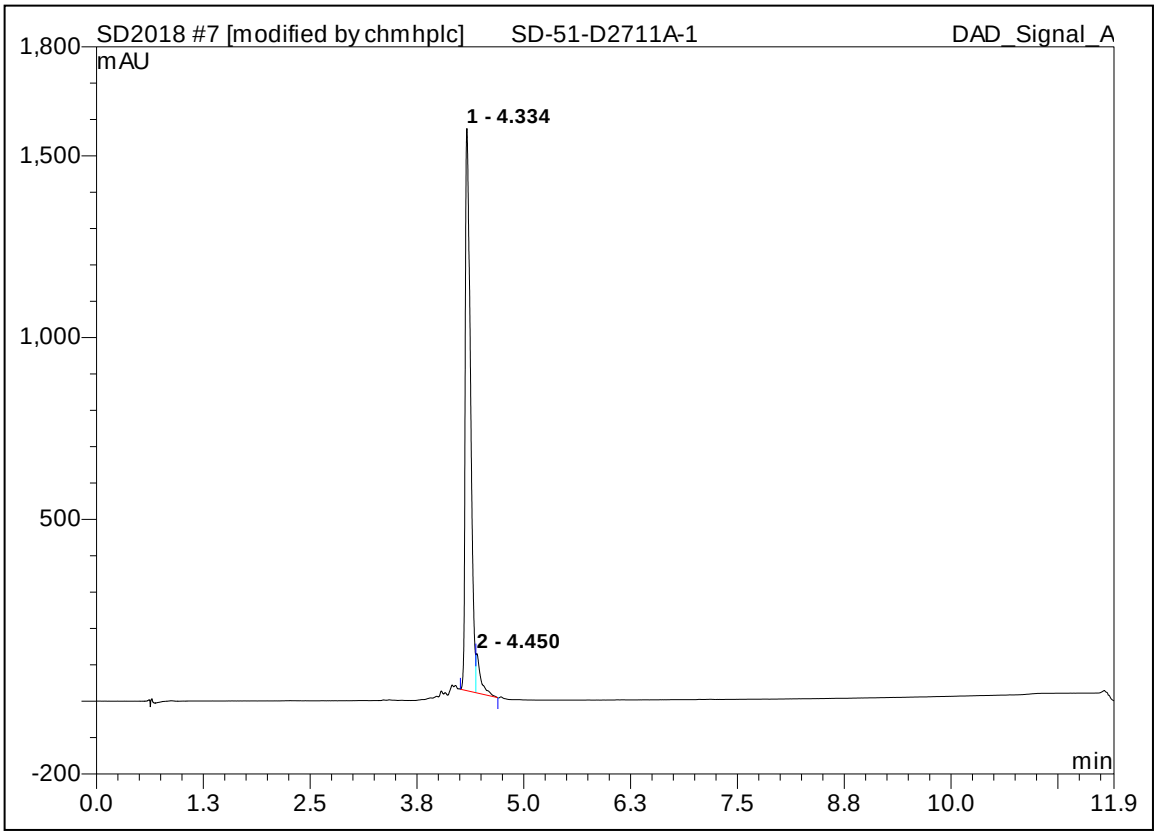


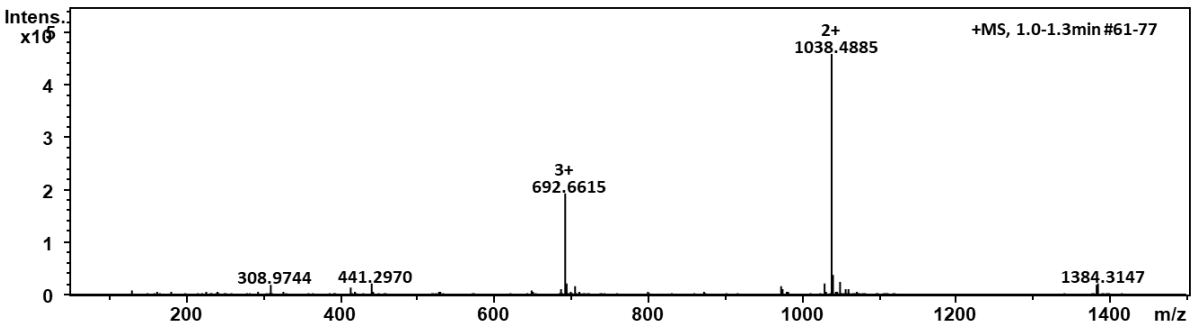
FAM-V2710A SIM



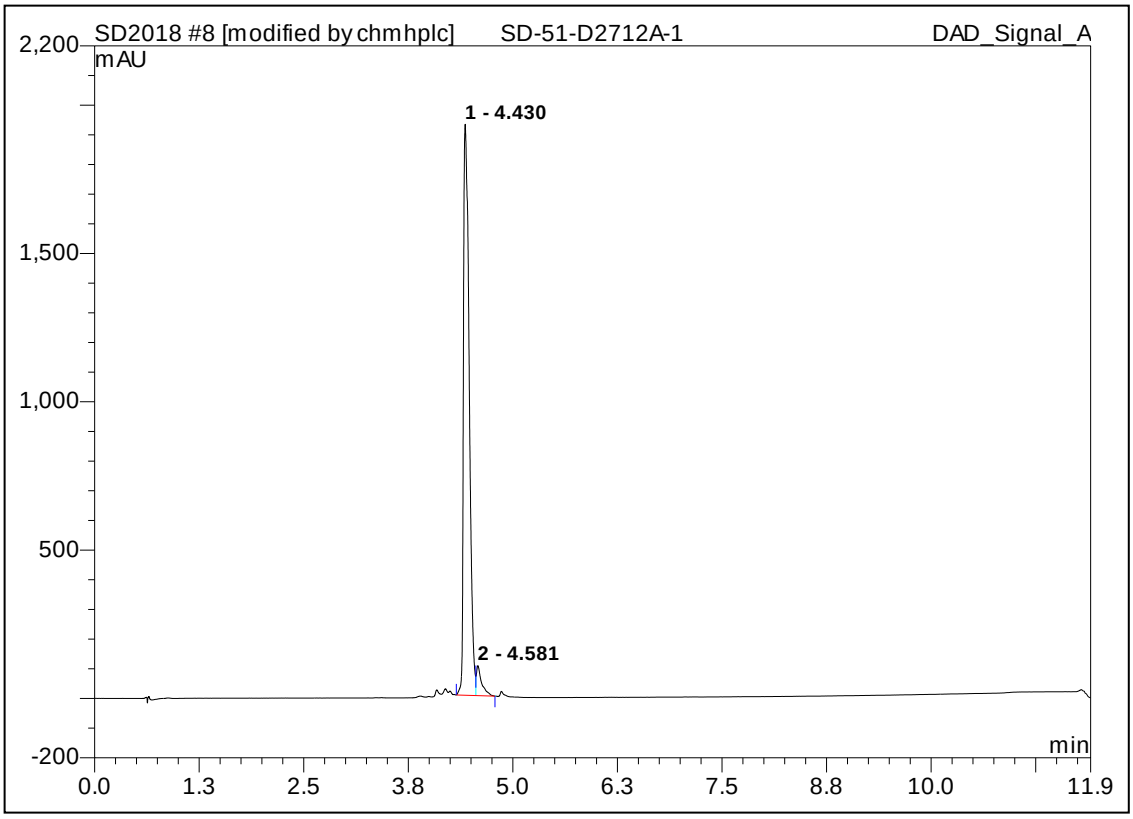


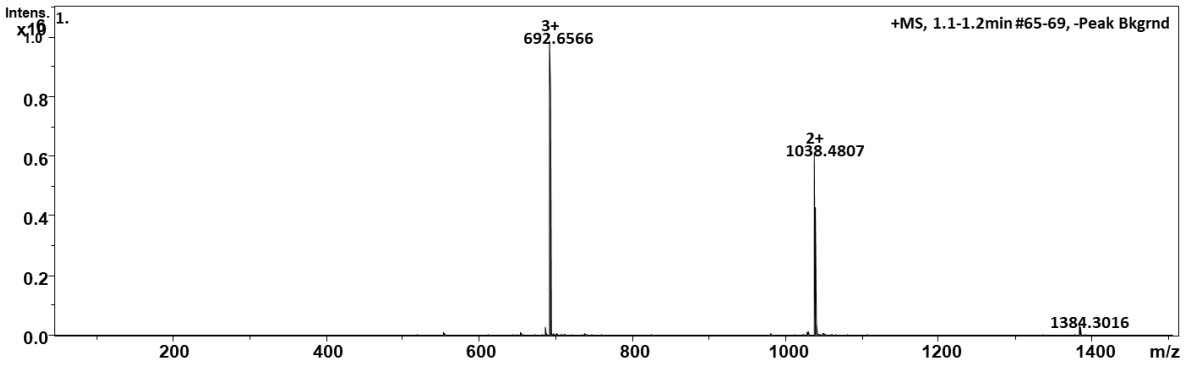
FAM-I2711A SIM



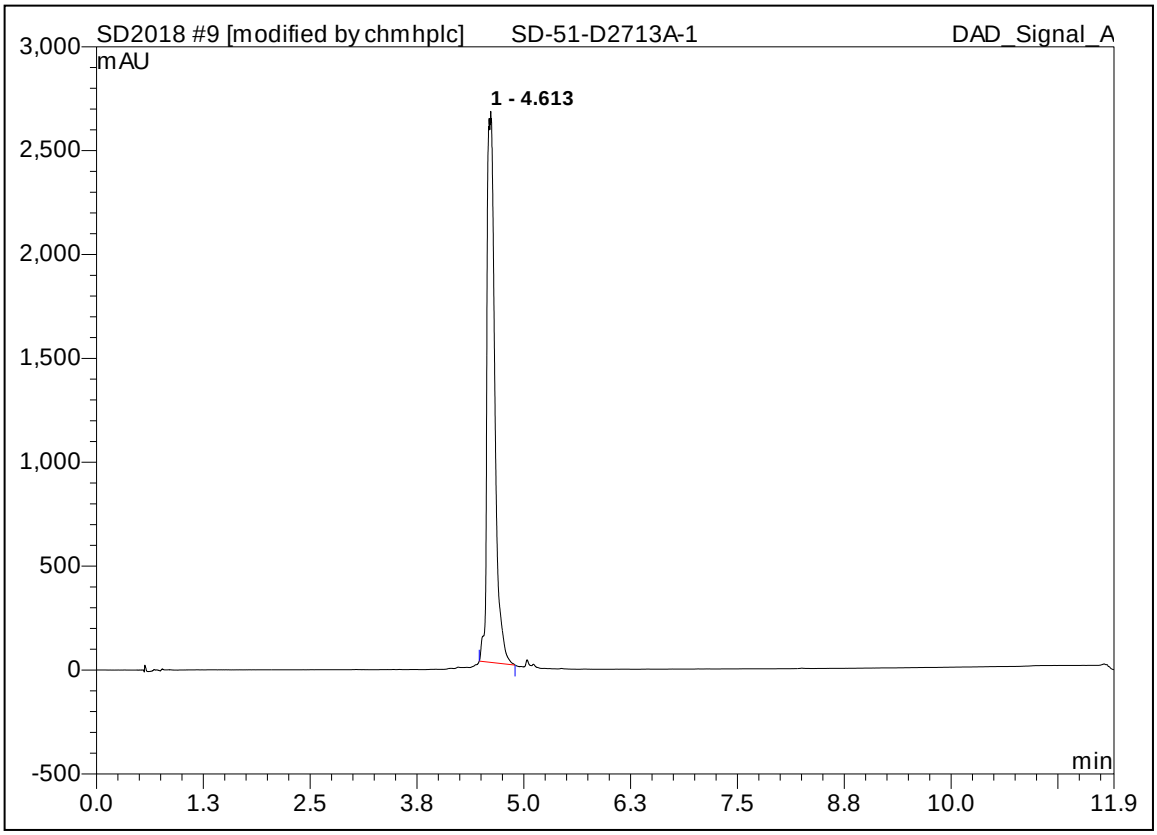


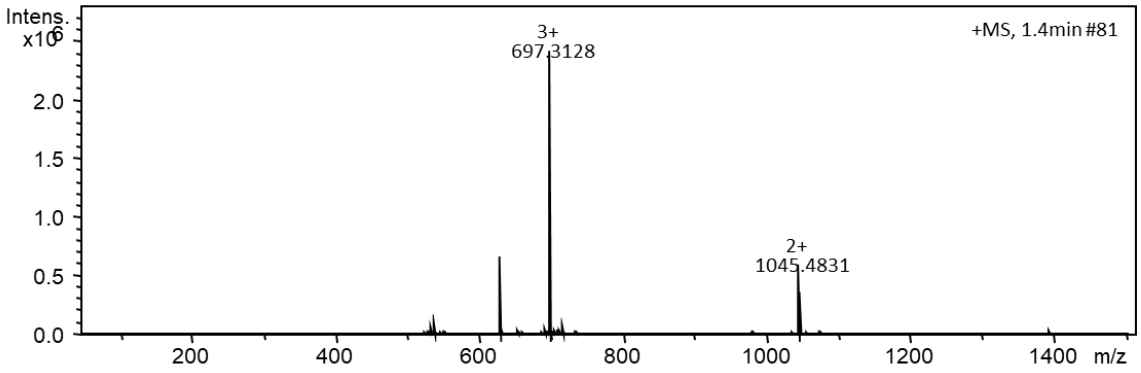
FAM-I2712A SIM



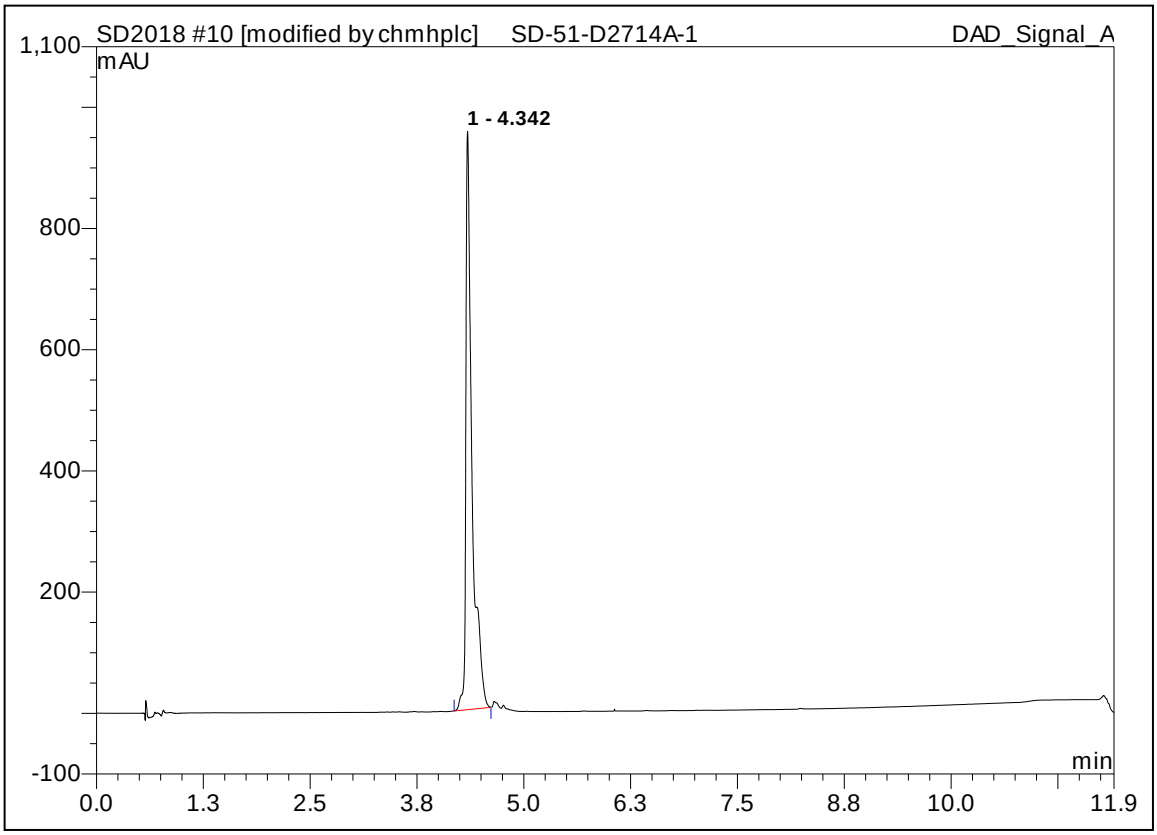


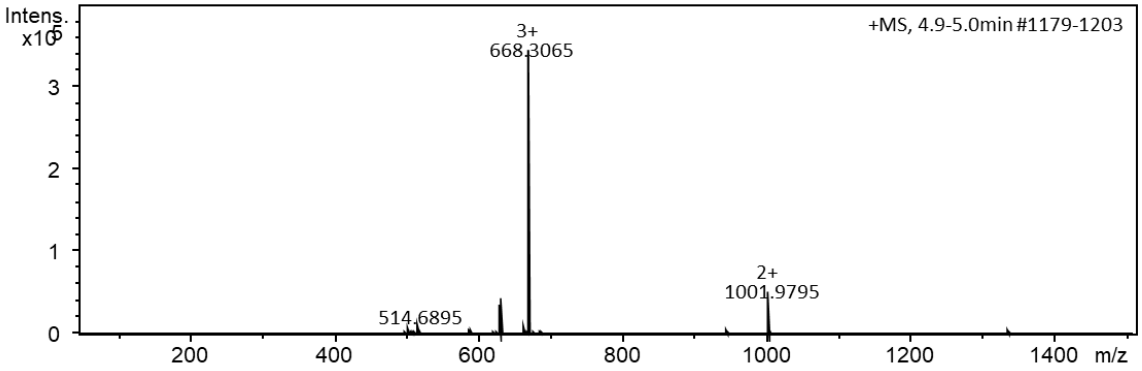
FAM-V2713A SIM



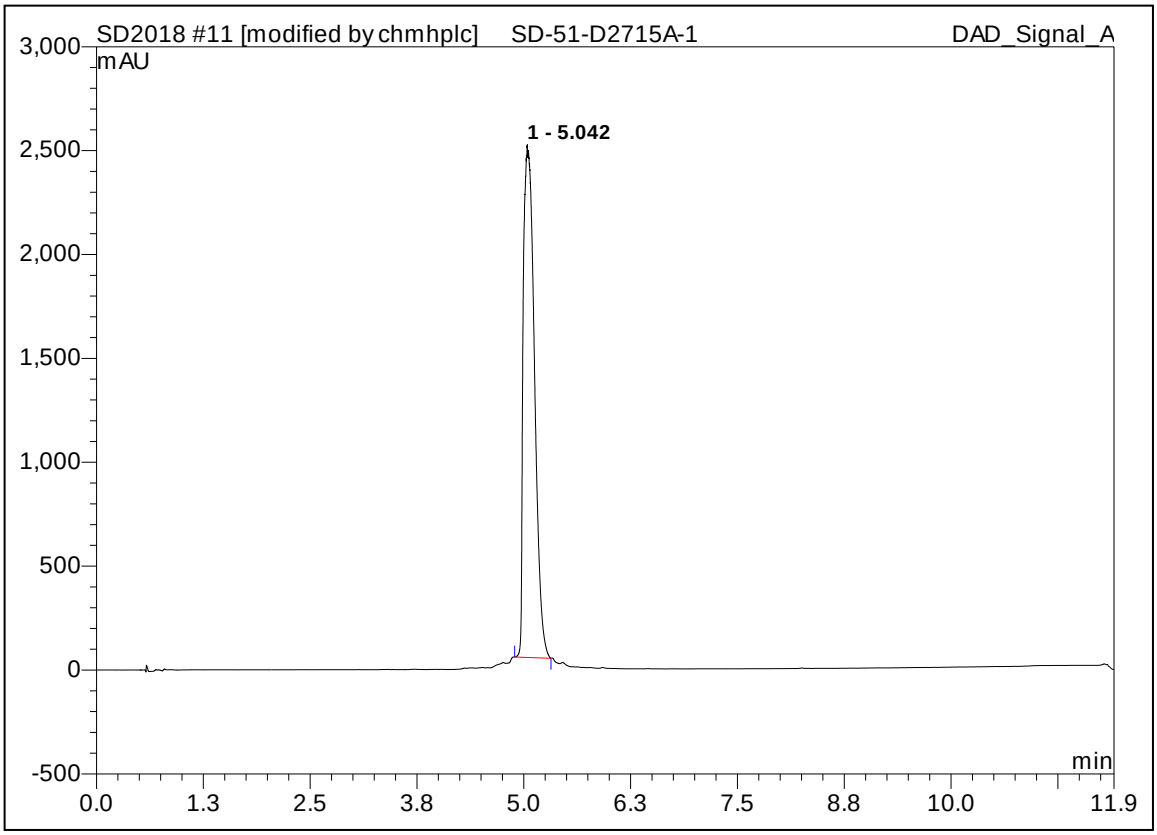


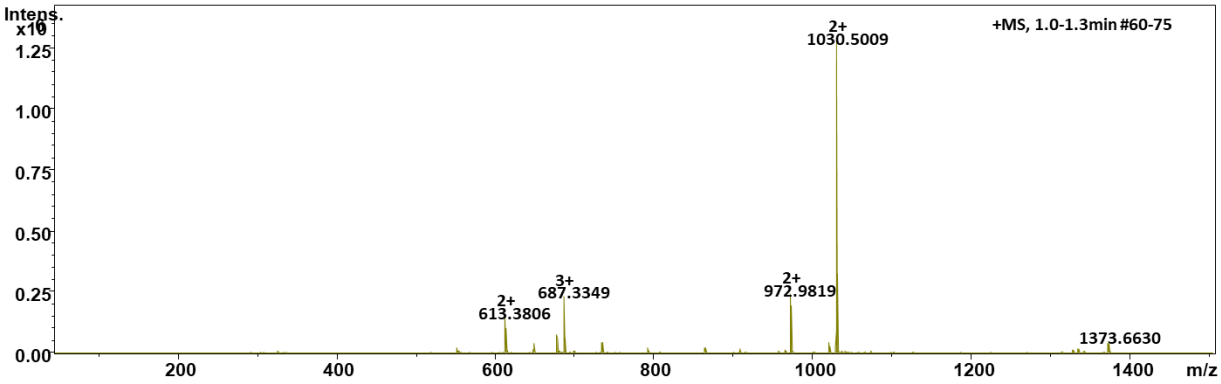
FAM-W2714A SIM



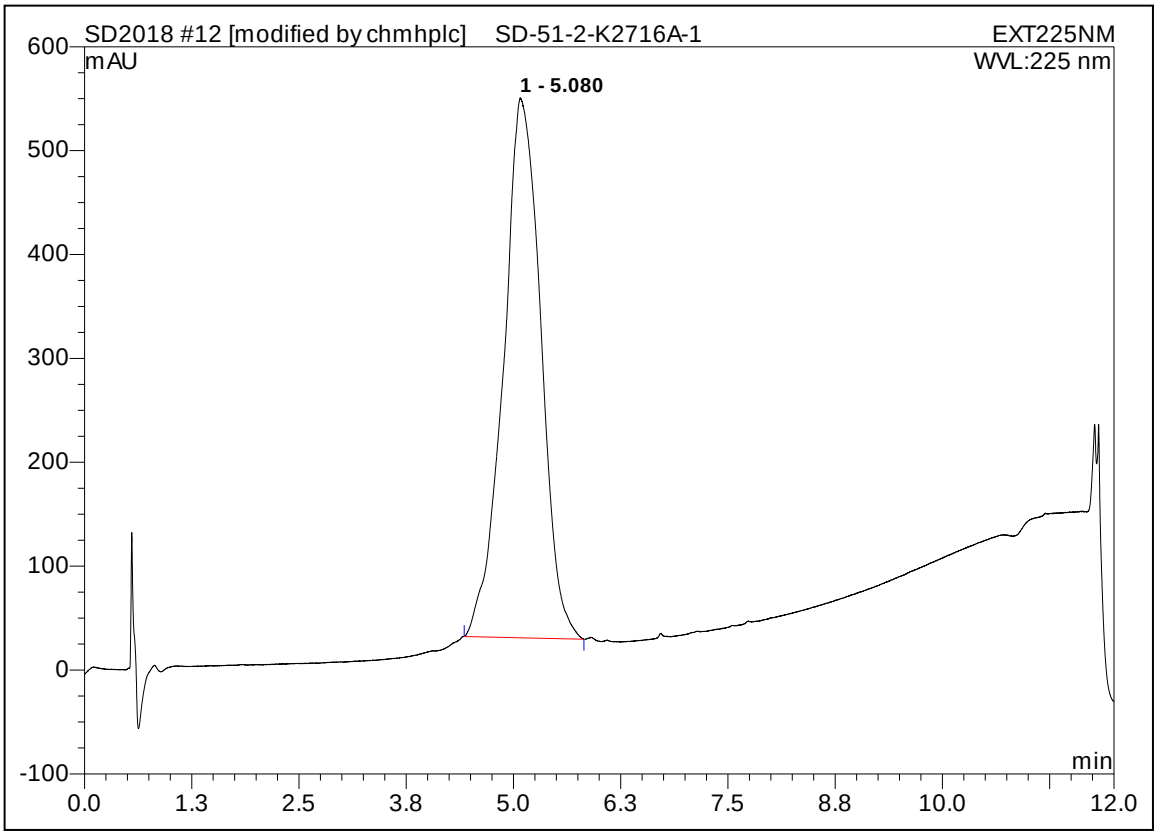


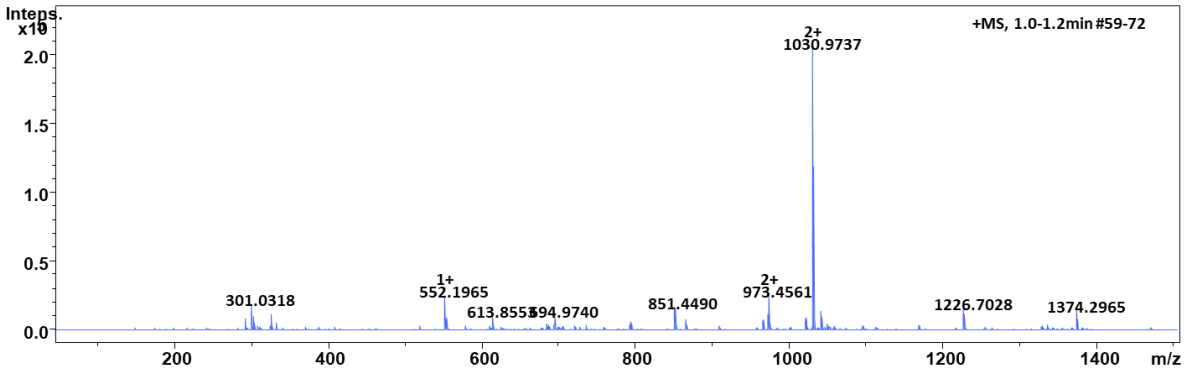
FAM-E2715A SIM



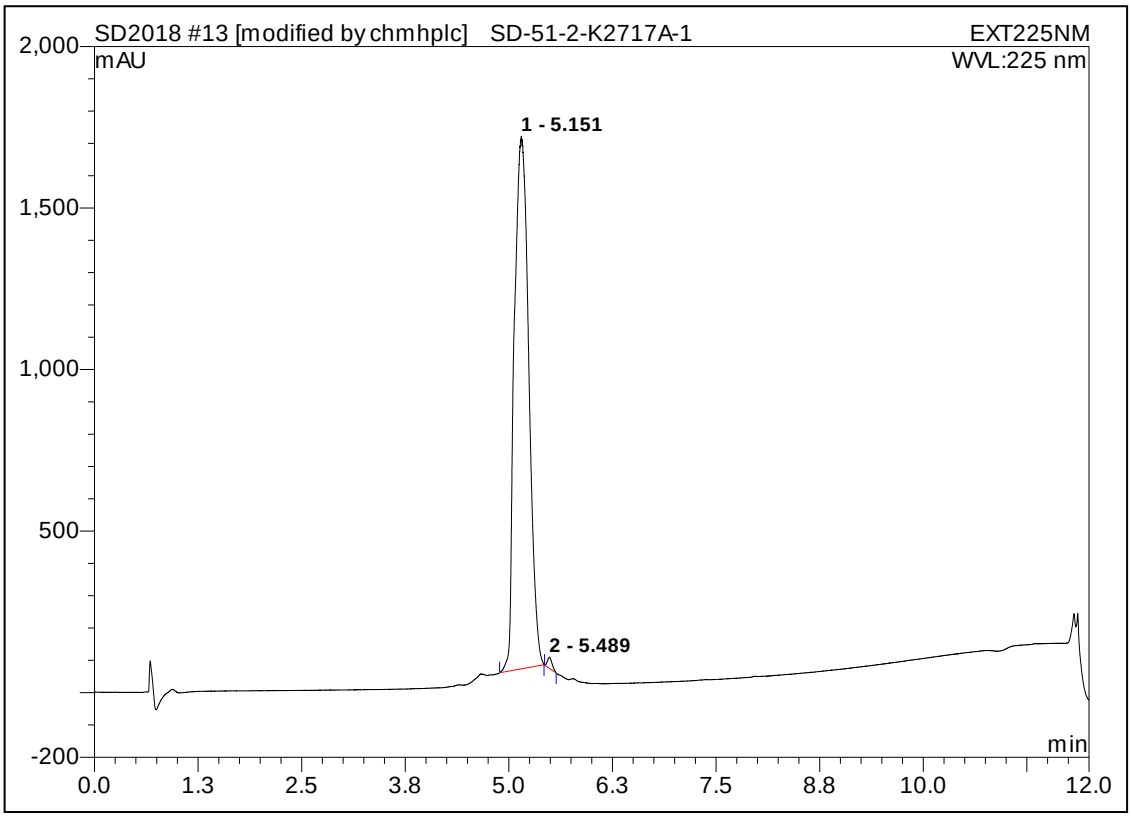


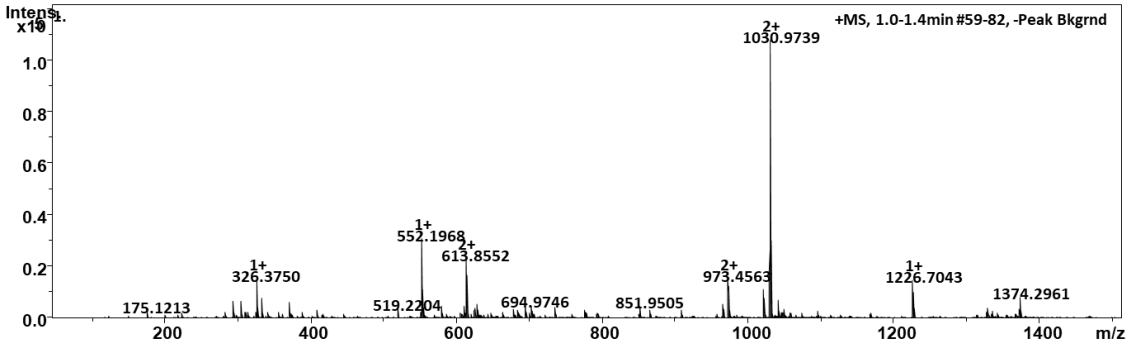
FAM-K2716A SIM



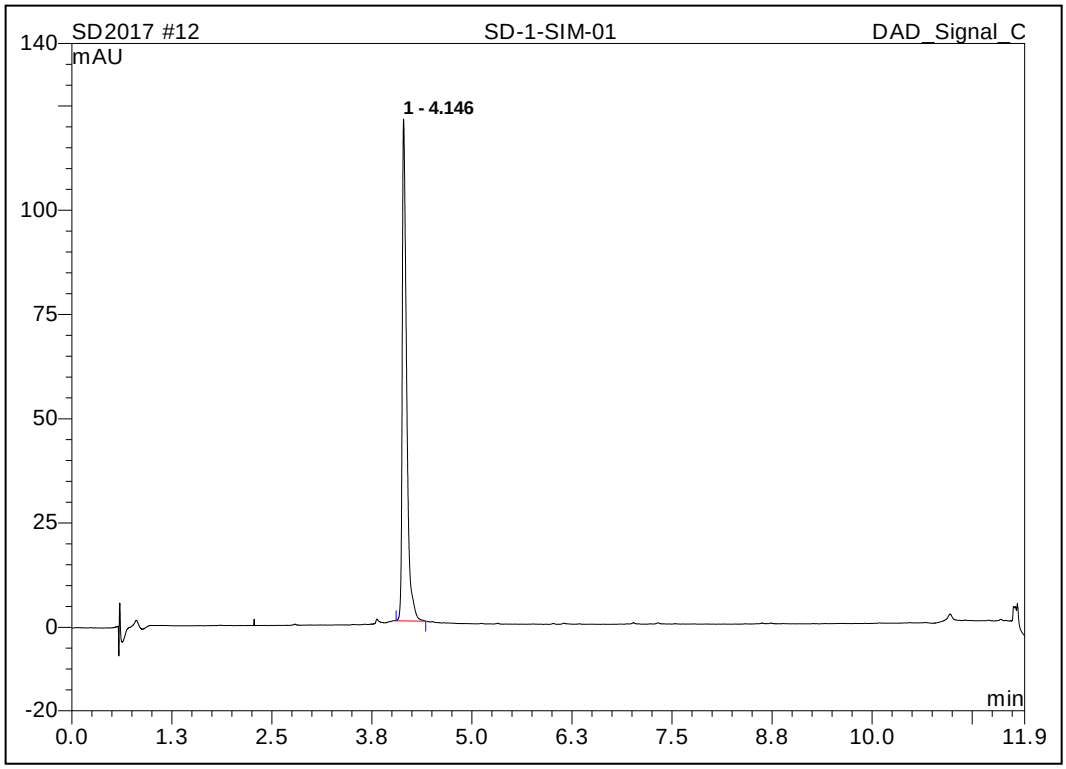


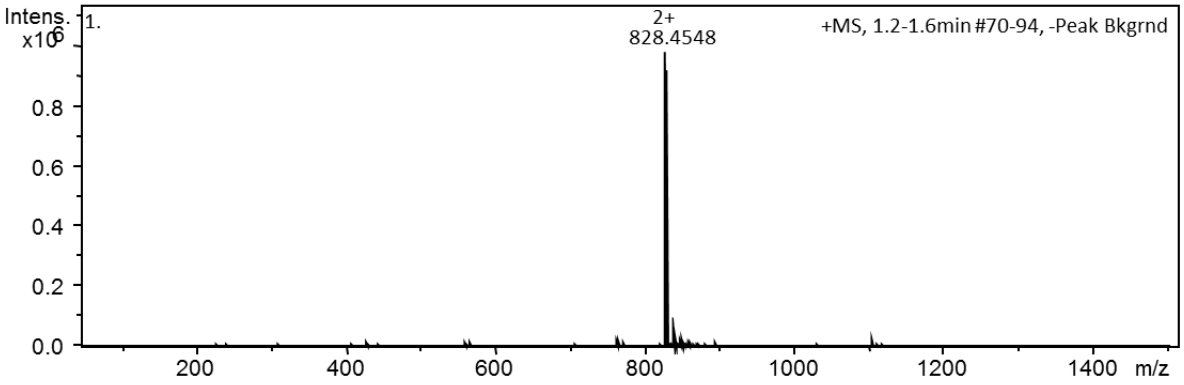
FAM-K2717A SIM



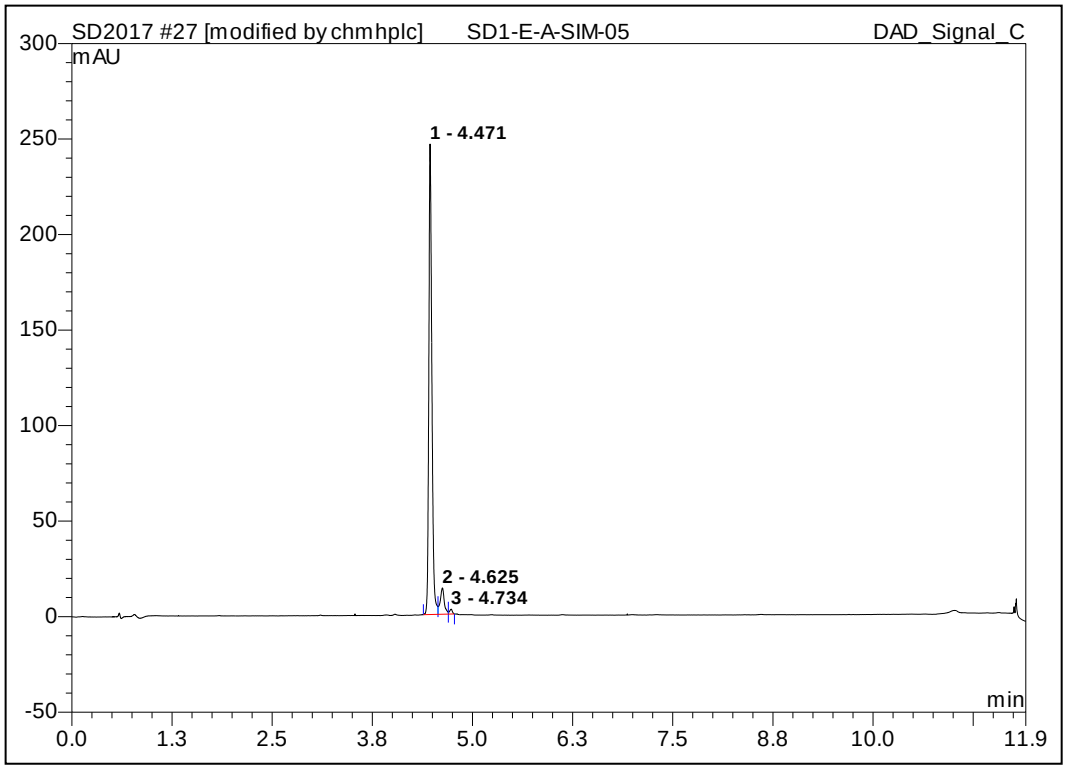


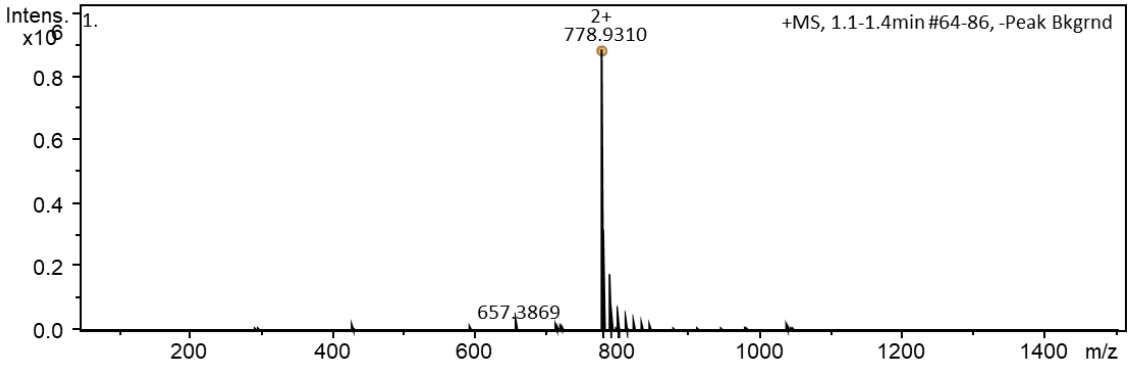
AcSIM


SD2017 #12
SD-1-SIM-01
DAD_Signal_C
mAU
1 - 4.146
140
100
75
50
25
0
-20
min
0.0
1.3
2.5
3.8
5.0
6.3
7.5
8.8
10.0
11.9

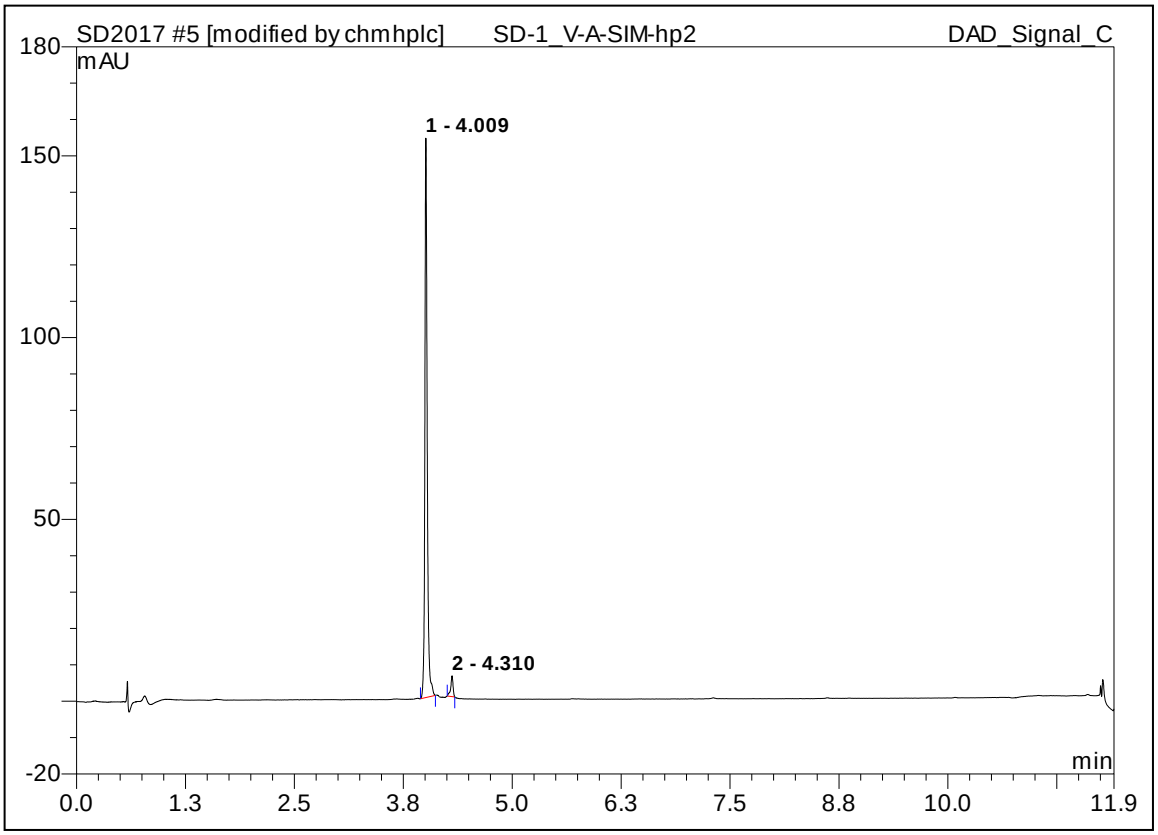

Intens.
x10^6
1.
2+
828.4548
+MS, 1.2-1.6min #70-94, -Peak Bkgrnd
0.8
0.6
0.4
0.2
0.0
200
400
600
800
1000
1200
1400
m/z

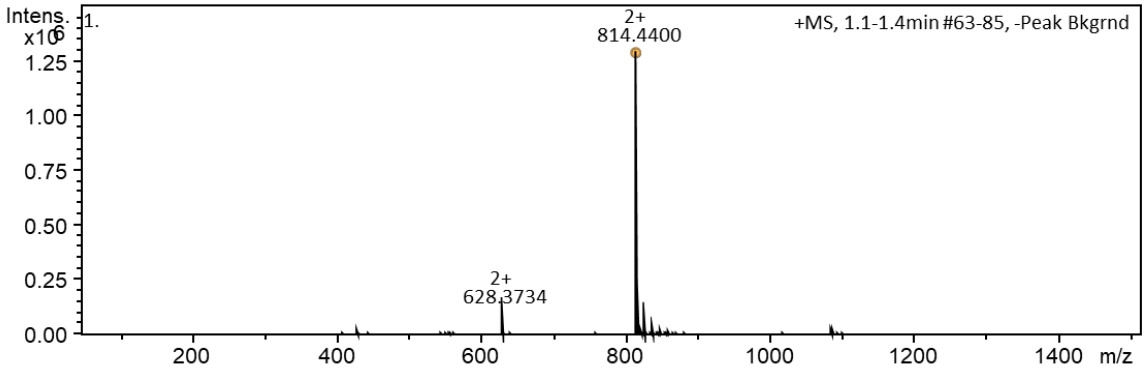
Ac-E2707A SIM



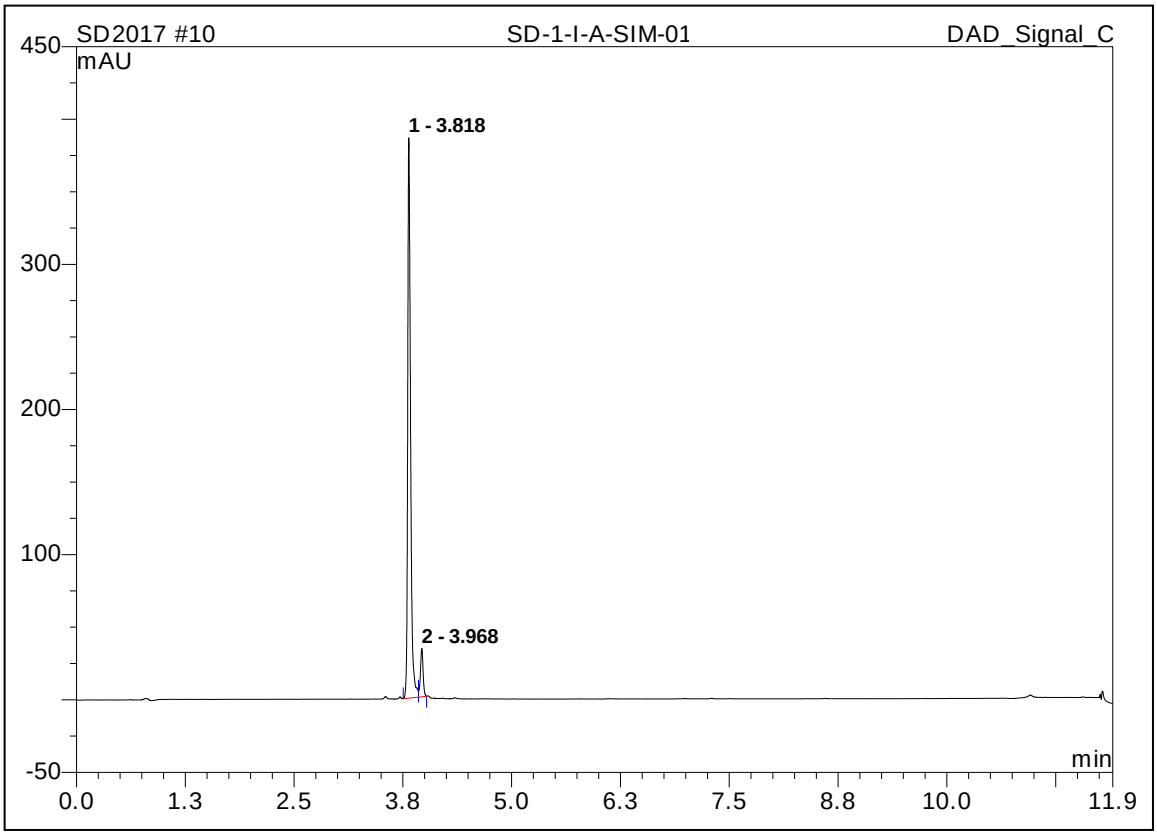


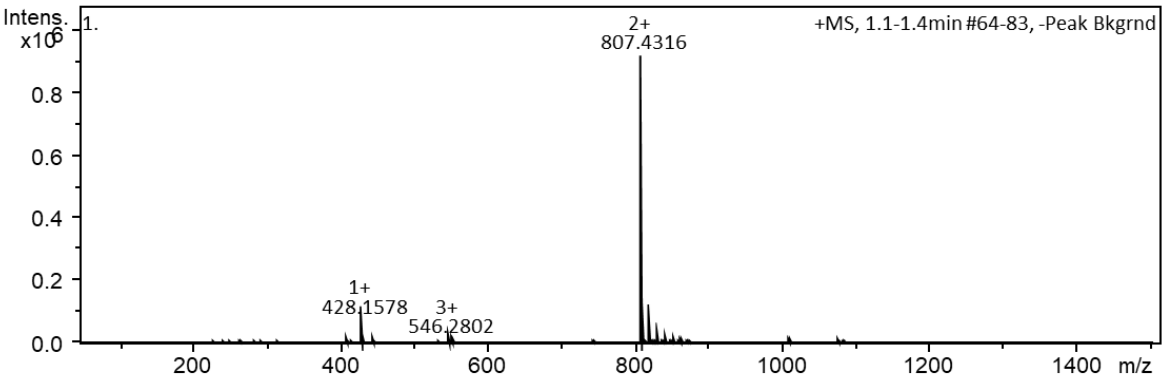
Ac-V2710A SIM



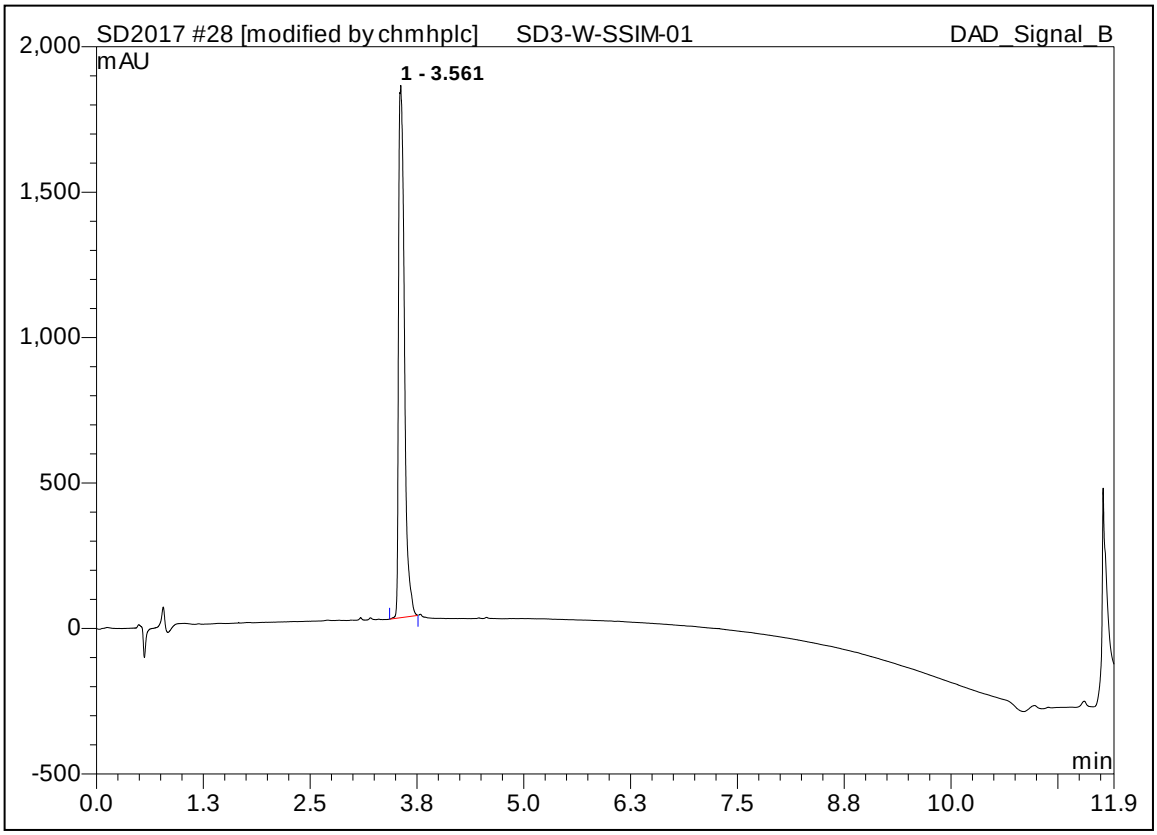


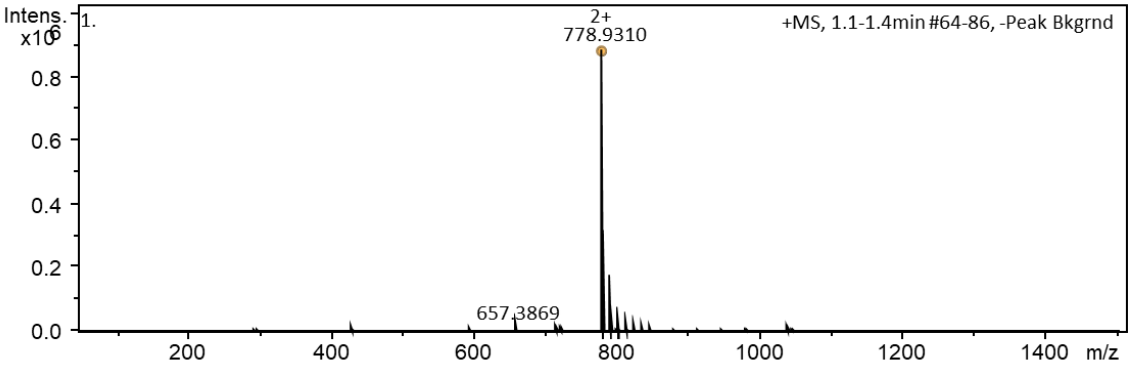
Ac-I2712A SIM





Ac-W2714S SIM





References

1. A. Barnard, K. Long, D. J. Yeo, J. A. Miles, V. Azzarito, G. M. Burslem, P. Prabhakaran, T. A. Edwards and A. J. Wilson, *Org. Biomol. Chem.*, 2014, **12**, 6794.
2. C. M. Grison, G. M. Burslem, J. A. Miles, L. K. A. Pilsl, D. J. Yeo, Z. Imani, S. L. Warriner, M. E. Webb and A. J. Wilson, *Chem. Sci.*, 2017, **8**, 5166.
3. J. A. Miles, P. Rowell, C. Tiede, C. H. Trinh, H. F. Kyle, B. I. M. Wicky, J. Clarke, D. C. Tomlinson, A. J. Wilson and T. A. Edwards, *BMC Biotechnol.*, 2018, submitted.
4. C. W. Wood, J. W. Heal, A. R. Thomson, G. J. Bartlett, A. Á. Ibarra, R. L. Brady, R. B. Sessions and D. N. Woolfson, *Bioinformatics*, 2017, **33**, 3043.
5. S. McIntosh-Smith, T. Wilson, A. Á. Ibarra, J. Crisp and R. B. Sessions, *Comput. J.*, 2012, **55**, 192.
6. G. G. Krivov, M. V. Shapovalov and R. L. Dunbrack, *Proteins*, 2009, **77**, 778.
7. E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Anal. Biochem.*, 1970, **34**, 595.