PNAS

Supporting Information for

A double-switch pHLIP system enables selective enrichment of circulating tumor microenvironment-derived extracellular vesicles

Zhiyou Zong, Xinzhuo Liu, Zhuo Ye and Dingbin Liu

Dingbin Liu
Email: liudb@nankai.edu.cn

This PDF file includes:

Other supporting materials for this manuscript include the following:

Supporting Information Text

Methods

p*K* calculations, molecular dynamics (MD) simulations, and interactional energy calculations

Three-dimensional structures of WT pHLIP and hundreds of variants were built by SWISS-MODEL (https://swissmodel.expasy.org/) (*1-5*). The relevant p*K* values of peptides were calculated by H++ software of 3.2 version (http://newbiophysics.cs.vt.edu/H++/index.php) (*6-8*). WT pHLIP and four potential D-S variants were simulated to investigate the performance of hook-like switch II. The details of these molecular systems are provided in Supplementary Table 8. Each assembly was solvated in an equilibrated box of water. The overall charge neutrality was achieved by adding ions of Na^+/Cl^- to the solution. All molecular systems before production simulations should undergo appropriate pretreatments (*9, 10*). Herein, 5000 steps minimization and 100 ps MD simulation with the restraint of peptide were employed to equilibrate water. Then 5000 steps of minimization with 5 ns equilibration without restraint were carried out to equilibrate the whole system. The scalable program NAMD 2.14 (*11*) with the CHARMM36 force field (*12-14*), and the TIP3P water model (*15*) were employed to perform all the MD simulations. Visualization and analysis of the MD trajectories were carried out by the VMD program (*16*). The interaction energies were calculated using the two negatively-charged residues, i.e., D33 and E34, and the positively-charged residues, i.e., R11 and T28H in D-S v1, and R11, T28H and T18H in D-S v2.

Free-energy calculations

We performed free-energy calculations for the WT pHLIP and the two variants to quantify the barriers against forming hook-like structures by the well-tempered meta-eABF method (*17*). *d* denotes the center-of-mass distance between the two negatively-charged residues (D33 and E34) in the C-terminal and the positively-charged R11 near the N-terminal. Detail of each molecular assembly is provided in Supplementary Table 9.

Recording of fluorescence and circular dichroism (CD) spectra

Peptides used in this study (see Supplementary Table 10 and characterization of peptides) were synthesized by GL Biochem (Shanghai, China) Ltd. Liposome preparation, Trp fluorescence measurements and CD experiments were carried out by reference to the works of An et al. and Andreev et al. (*18, 19*). The details are described below. A solution of 20 mg 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) model membranes (Avanti Polar Lipids) in 0.8 mL of chloroform was dried in a small round-bottom flask in vacuo by a rotary evaporator. The dry lipid was rehydrated with 2 mL PBS (pH 7.4, 5 mM) for 30 min, and vortexed vigorously. Then the mixture was freeze-thawed 7 times by a dry-ice/ethanol bath (−70 °C) and extruded 15 passages through a polycarbonate membrane with 100 nm-sized pores by an Avanti Mini-Extruder (Avestin, Canada). All the experiments were carried out at room temperature. The lyophilized powder of peptides was dissolved by PBS (pH 7.4, 5 mM) and put at room temperature for 24 h. Then 5 μM peptides were incubated with POPC at pH 7.4 for 2 h with a pHLIP to lipid ratio of 1:260. The pH condition of each system was further changed for Trp fluorescence measurements. Trp residues were excited at 295 nm. The fluorescence emission spectra were collected from 310-450 nm by the fluorescence spectrophotometer (Hitachi, Japan). The widths of the excitation and emission slits were set to 2.5 and 5 nm, respectively. The p*K*a for WT pHLIP was calculated by the Henderson-Hasselbalch equation (*18*). Additionally, the samples with different pH conditions were further analyzed by CD (BioLogic, France). Hydrophobic properties of the peptides were analyzed by ProtScale (https://web.expasy.org/protscale/).

Cellular Experiments

Human lung and breast cancer cells, i.e., A549 and MDA-MB-231, were obtained from the Shanghai Institute of Life Sciences, Chinese Academy of Sciences. All the in-vitro experiments regarding cells and EVs were carried out on the sterile ultra-clean table. In the cell experiments, A549 cells were cultivated overnight at 37 ºC

in a confocal dish with 1 mL Dulbecco's modified essential medium (DMEM; Wisent, Canada) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Wisent, Canada) and 1% penicillin-streptomycin (Gibco-Life Technologies, Grand Island, NY). After removing the supernatant, fresh 1 mL DMEM was added with 5 μM peptides@biotin. Then three pH conditions, i) pH 7.4, ii) pH 7.4→6.2 (WT pHLIP)/6.8 (D-S v1), and iii) pH 7.4→6.2 (WT pHLIP)/6.8 (D-S v1)→7.4 were employed to treat the systems. After changing pH, the peptides were incubated with the cells for 0.5 h in the fresh DMEM, wherein the cells kept growing. Excess peptides@biotin were removed by washing the cell three times using PBS (pH 7.4 for i) and iii), and pH 6.2 (WT pHLIP)/6.8 (D-S v1) for ii)). Then FITC@SA was added into the supernatant and used to specifically link with peptides@biotin via biotin-streptavidin interactions (similarly hereinafter). After washing off the excess FITC@SA using PBS with relevant pH, the cell nuclei were stained with DAPI (blue fluorescence) after immobilization.

In the EV experiments, peptides@biotin were incubated with 1 mL supernatant of A549 cells that were cultivated overnight, which then underwent the pH treatments of 7.4→6.2 (WT pHLIP)/6.8 (D-S v1)→7.4. Subsequently, FITC@SA was added into the supernatants to link with peptides@biotin. DiI was further employed to stain the EV membranes. All the pH changes and buffer substitutions regarding EV experiments were carried out by 100 KDa ultrafiltration centrifugal tube. Confocal microscopy (Nikon, Japan) was used to measure the fluorescence for different systems. The results of particle size and concentration of EVs were obtained by nanoparticle tracking analysis (NTA; Malvern, UK). Transmission electron microscope (TEM; Hitachi, Japan) was employed to observe the morphological characteristics of EVs. ImageJ software (https://imagej.en.softonic.com) was used to analyze the fluorescence intensity. The cell viability and toxicity experiments were performed using Cell Counting Kit-8 and MTT assay kit respectively, according to the instructions of the manufacturer.

***In vivo* experiments**

Nude mice, 5-6 weeks of age, were purchased from Beijing Vital River Laboratory Animal Technology Co. Ltd. All animal studies were performed in compliance with the guidelines set by the Tianjin Committee of Use and Care of Laboratory Animals and the overall project protocols were approved by the Animal Ethics Committee of Nankai University. The accreditation number of the laboratory is SYXK(Jin) 2019-0003 promulgated by the Tianjin Science and Technology Commission. After 10^6 lung and breast cancer cells implanting into mice, tumor was formed within three weeks. In the *in vivo* imaging experiments, D-S v1@Cy5.5 (400 μg/kg) was intraperitoneally (i.p.) injected into the tumor model mice and imaged after 24 h. In the RNA detection experiments, peptides@biotin (2 mg/kg) was i.p. injected into the opposite side of adenocarcinoma. The whole blood samples were collected at different time points, placed 1.5 h, and then centrifuged at 1000g for 15 min at 4 °C to obtain blood serum. The cTME-EVs tagged by peptides@biotin were collected from the serum by 2 μm MB@SA using magnetic separation. The RNA concentrations of the samples were detected by miRNeasy Serum/Plasma Kit (QIAGEN), according to the instructions of the manufacturer. Given the RNA detection results, samples obtained at 24 h post injection were used for subsequent experiments. Furthermore, the mice in the RNA detection experiments were dissected and the relevant organs were imaged *ex vivo*.

In FCM experiments, the samples were incubated with 2 μL antibodies (PE-labeled CD63 and Alexa Fluor 700-labeled EpCAM, 0.1 μg/μL) in 300 μL PBS (pH 7.4) with rotation at 4 °C for 30 min. The samples were added to a 12 × 75 mm polypropylene round-bottom flow tube and run at a medium speed. The FCM studies were performed on a BD LSR Fortessa analyzer (Becton, Dickinson and Company). All the data were acquired and analyzed by the BD FACSDiva and FlowJo software (https://www.flowjo.com/). In western-blot (WB) assays, the samples were lysed by ice-cold RIPA buffer (Beyotime Biotechnology) with a protease inhibitor (PMSF, 1 mM, Beyotime Biotechnology) for 30 min. The lysates were centrifuged at 14,000g for 20 min for collecting supernatant. Then the samples were separated by SDS-PAGE and transferred onto a 0.45 μm nitrocellulose

membrane (Millipore). Membranes were further blocked by 5% skim milk (in TBST, 20 mM Tris-HCl, 150 mM NaCl, and 0.05% Tween-20) at 37 °C for 1 h before primary-antibody incubation. Subsequently, the antibodies against the following proteins were used for WB experiments: anti-CD63 (1:1000, Proteintech, 25682-1-AP), anti-EpCAM (1:1000, Beyotime, AF1603), anti-Alix (1:1000, Beyotime, ET1705-74), and anti-CD9 (1:1000, Beyotime, AF1192). HRP-linked anti-rabbit-IgG antibody (1:3000, Beyotime, A0208) was used as secondary antibodies. The membranes were incubated with primary antibodies overnight at 4 ºC, then washed three times using TBST, and incubated with secondary antibodies for 2 h at room temperature. The membranes were further washed every ten minutes by TBST (three times), incubated with BeyoECL Plus (Beyotime Biotechnology), and imaged by Azure c600 (Azure Biosystems, USA). The particle size and concentration of EVs in the serum were analyzed by NTA (Malvern, UK).

Transcriptome

The samples of the D-S v1 group were obtained as described in tissue experiments. In the ultracentrifugation (UC) group, the obtained serum samples were centrifuged at 10,000 g for 30 min. The supernatants were centrifuged by UC (110,000 g for 90 min, Beckman Coulter, Optima^{TM} MAX-XP, USA). The obtained pellets were resuspended in PBS (pH 7.4, 5 mM) and centrifuged again by UC (110,000 g for 90 min). Finally, the yielded pellets were dispersed in PBS (pH 7.4, 5 mM). The particle sizes, concentrations, and morphological features of these EVs were characterized by NTA and TEM, respectively (Supplementary Fig. 8). RNA sequencing and analysis were performed by LC–Bio Technology co., ltd. (Hangzhou, China). $q < 0.05$ and a fold change (FC) >2 are used to define upregulation and downregulation genes, while gray ones are not significant. Protein-protein interaction (PPI) network of DEGs and the filtered top ten hub genes from PPI were analyzed by Cytoscape (https://cytoscape.org/) (*20*).

Tables S1-10

Table S1. p*K* values of D25 (p*K*25) in virtual variants.

L21[a]	p*K*25	L22[a]	p*K*25	L26[a]	p*K*25	L28[a]	p*K*25
WT	4.261	WT	4.261	WT	4.261	WT	4.261
L21A	4.238	L22A	4.237	L26A	4.298	L28A	4.222
L21R	4.672	L22R	4.241	L26R	4.251	L28R	4.315
L21N	4.222	L22N	4.275	L26N	4.338	L28N	4.168
L21D	4.193	L22D	4.161	L26D	4.240	L28D	4.210
L21C	4.224	L22C	4.173	L26C	4.209	L28C	4.191
L21Q	4.255	L22Q	4.176	L26Q	4.272	L28Q	4.407
L21E	4.234	L22E	4.460	L26E	4.260	L28E	4.274
L21G	4.236	L22G	4.513	L26G	4.245	L28G	4.231
L21H	4.471	L22H	4.270	L26H	4.280	L28H	4.252
L21I	4.289	L22I	4.251	L26I	4.253	L28I	4.290
L21K	4.339	L22K	4.222	L26K	4.277	L28K	4.295
L21M	4.275	L22M	4.189	L26M	4.239	L28M	4.249
L21F	4.539	L22F	4.275	L26F	4.250	L28F	4.276
L21P	4.278	L22P	4.186	L26P	4.069	L28P	4.283
L21S	4.234	L22S	4.172	L26S	4.224	L28S	4.163
L21T	4.268	L22T	4.364	L26T	4.277	L28T	4.204
L21W	4.489	L22W	4.338	L26W	4.258	L28W	4.258
L21Y	4.539	L22Y	4.266	L26Y	4.258	L28Y	4.252
L21V	4.288	L22V	4.230	L26V	4.233	L28V	4.285

[a] Site of virtual mutation.

Table S2. p*K* values of D14 (p*K*14) in virtual variants.

A10[a]	p*K*14	R11[a]	p*K*14	W15[a]	p*K*14	F17[a]	p*K*14	T18[a]	p*K*14	T19[a]	p*K*14
WT	3.873	WT	3.873	WT	3.873	WT	3.873	WT	3.873	WT	3.873
A10G	3.868	R11G	3.860	W15A	3.822	F17A	3.856	T18A	4.140	T19A	3.888
A10V	3.883	R11A	3.880	W15R	4.188	F17R	3.829	T18R	3.410	T19R	3.882
A10L	3.888	R11V	3.874	W15N	4.322	F17N	3.806	T18N	4.013	T19N	3.976
A10I	3.877	R11L	3.842	W15D	3.909	F17D	3.742	T18D	4.216	T19D	3.907
A10P	3.784	R11I	3.827	W15C	3.837	F17C	3.812	T18C	4.050	T19C	3.940
A10F	3.967	R11P	3.841	W15Q	4.133	F17Q	3.821	T18Q	4.752	T19Q	3.857
A10Y	3.902	R11F	3.792	W15E	3.728	F17E	3.792	T18E	4.376	T19E	3.873
A10W	4.015	R11Y	3.887	W15G	3.790	F17G	3.835	T18G	4.046	T19G	3.996
A10S	3.810	R11W	3.861	W15H	3.824	F17H	3.894	T18H	4.707	T19H	3.873
A10T	3.857	R11S	3.872	W15I	3.892	F17I	3.863	T18I	4.310	T19I	3.904
A10C	3.758	R11T	3.848	W15L	3.899	F17L	3.867	T18L	4.271	T19L	3.908
A10M	3.860	R11C	3.873	W15K	4.153	F17K	3.866	T18K	4.425	T19K	3.855
A10N	3.877	R11M	3.894	W15M	3.867	F17M	3.838	T18M	4.253	T19M	3.897
A10Q	3.803	R11N	3.939	W15F	3.902	F17P	3.967	T18F	4.658	T19F	3.879
A10D	3.864	R11Q	3.886	W15P	3.497	F17S	3.935	T18P	4.272	T19P	3.790
A10E	3.744	R11E	3.828	W15S	4.255	F17T	3.822	T18S	3.733	T19S	3.838
A10K	3.774	R11D	3.798	W15T	3.692	F17W	3.860	T18W	4.294	T19W	3.880
A10R	3.854	R11K	3.899	W15Y	3.978	F17Y	3.865	T18Y	4.492	T19Y	3.876
A10H	3.959	R11H	3.871	W15V	3.856	F17V	3.883	T18V	4.300	T19V	3.883

[a] Site of virtual mutation.

Table S3. Multiple mutation strategies for increasing p*K*25.

Variant	p*K*25	Variant	p*K*25
L21F/L22G	5.068	L21F/L22W	4.642
L21F/L22T	5.025	L21Y/L22G	4.382
L21Y/L22W	4.684	L21Y/L22T	4.410
L21W/L22G	4.419	L21W/L22W	4.812
L21W/L22T	4.654	L21H/L22G	4.437
L21H/L22W	4.495	L21H/L22T	4.412
L21F/L26N	4.675	L21W/L26N	4.856
L21Y/L26N	4.651	L22G/L26N	4.431
L22T/L26N	4.310	L22W/L26N	4.519
L21F/L22W/L26N	5.014	L21W/L22T/L26N	5.199
L21Y/L22W/L26N	5.063		

Table S4. Microscale increments of p*K*25 and p*K*14, a summary for the above single and multiple mutations.

Peptide	p*K*25	Peptide	p*K*14
WT	4.261	WT	3.873
L22T	4.364	A10Y	3.902[a]
L21W	4.489	A10W	4.015[b]
L21F	4.539	W15Q	4.133
L21Y	4.539	T18A	4.140
L21F/L22W	4.642	T18M	4.253
L21Y/L26N	4.651	W15S	4.255
L21W/L22T	4.654	T18L	4.271
L21F/L26N	4.675	T18P	4.272
L21Y/L22W	4.684	T18W	4.294
L21W/L22W	4.812	T18V	4.300
L21W/L26N	4.856	T18I	4.310[c]
L21F/L22W/L26N	5.014	T18K	4.425
L21F/L22G	5.068[d]	T18Y	4.492
L21W/L22T/L26N	5.199	T18F	4.658
		T18H	4.707
		T18Q	4.752

[a] Around p*K* value of 3.9, there are four alternative variants, i.e., W15F (3.902), T19G (3.907), T19I (3.904), and T19L (3.908).

[b] Around p*K* value of 4.0, there are two alternative variants, i.e., T18N (4.013) and T18G (4.046).

[c] Around p*K* value of 4.3, there is one alternative variants, i.e., W15N (4.322).

[d] Around p*K* value of 5.0, there is one alternative variants, i.e., L21Y/L22W/L26N (5.063).

Table S5. p*K* values of ionizable residues in variants.

Variant	p*K*3[a]	p*K*14	p*K*25	p*K*28[a]	p*K*31[a]	p*K*33[a]	p*K*34[a]
L28H	4.429	3.773	4.252	6.563	4.241	3.976	4.244
L28H/D31A	4.443	3.751	4.241	6.583	-	3.939	4.242
L28H/D31A/L21F/L22G/T18Q	4.461	4.080	4.819[b]	6.605	-	4.199	4.246
L28H/D31A/L21W/L22W/T18Q	4.461	4.142	4.436	6.606	-	4.195	4.241
L28H/D31A/L21F/L22G	4.464	3.767	4.804[b]	6.602	-	4.199	4.248
L28H/D31A/L21W/L22W	4.463	3.794	4.357	6.603	-	4.201	4.243
L28H/D31A/L21F/L22G/T18F	4.473	4.137	4.791[b]	6.592	-	3.767	4.282
L28H/D31A/L21F/L22G/T18H	4.470	4.205	4.745[b]	6.595	-	3.791	4.276
L28H/D31A/L21W/L22T/L26N	4.469	3.781	4.405	6.581	-	3.762	4.273

[a] p*K*3, p*K*28, p*K*31, p*K*33 and p*K*34 denote the p*K* values of E3, L28H, D31, D33 and E34, respectively.

[b] Promising p*K* value.

Table S6. Traceable top 50 DEGs[a] in the D-S v1 group and their potential cell sources[b].

DEG	Cell type
Alb	Brain cancer cells, fibroblasts, mast cells, and mov neuroglial cells
Trf	Brain cancer cells and mast cells
Mup family[c]	Mast cells
Aldo family[d]	Basophilic leukemia cells, macrophages, mast cells, microglia, neural stem cells, oligodendrocytes, and pancreatic cells
Apo family[e]	Mast cells and oligodendrocytes
Serpina family[f]	Brain cancer cells
Gm family[g]	Mast cells, embryonic fibroblasts, and oligodendrocytes
Mug1	Mast cells
Dnah family[h]	Fibroblasts
Ltbp family[i]	Mast cells and fibroblasts
Myo family[j]	Macrophages
Lama family[k]	Fibroblasts
Col6 family[l]	Embryonic fibroblasts, fibroblasts, and mast cells
Wdr family[m]	Mast cells
Nefm	Pancreatic cells
Mgat5b	Fibroblasts

[a] In the following annotations, genes obtained by the D-S pHLIP-based platform are emphasized in bold.

[b] Top DEGs were matched to exosome biomarker database, ExoCarta (http://exocarta.org/index.html).

[c] **Mup7** and **Mup22** are not included in ExoCarta, but Mup family, for instance Mup1 and Mup5 are traceable.

[d] **Aldob** is not included in ExoCarta, but Aldo family, for instance Aldoa and Aldoc are traceable.

[e] **Apoe** is traceable, but **Apoa1** and **Apoa2** are not included in ExoCarta. Another gene in Apo family, Apom is traceable.

[f] **Serpina1a** is traceable, but **Serpina1c** and **Serpina3k** are not included in ExoCarta.

[g] **Gm26391, Gm34354**, and **Gm10717** are not included in ExoCarta, but many genes in Gm family are traceable.

[h] **Dnah1**, **Dnah9**, and **Dnah14** are not included in ExoCarta, but Dnah family, for instance Dnah10 is traceable.

[i] **Ltbp2** is not included in ExoCarta, but Ltbp family, for instance Ltbp1 and Ltbp4 are traceable.

[j] **Myo5b** is not included in ExoCarta, but Myo family, for instance Myo1c and Myo1g are traceable.

[k] **Lama1** is not included in ExoCarta, but Lama family, for instance Lama5 is traceable.

[l] **Col6a4** is not included in ExoCarta, but Col6a family, for instance Col6a1 and Col6a2 is traceable.

[m] **Wdr54** is not included in ExoCarta, but many genes in Wdr family are traceable.

Table S7. Traceable top 50 DEGs[a] in the UC group and their potential cell sources[b].

DEG	Cell type
Gm family[c]	Mast cells, embryonic fibroblasts, and oligodendrocytes
Eef1a1	Dendritic cells, fibroblasts, macrophages, mast cells, mov neuroglial cells, oligodendrocytes, and pancreatic cells
Atp family[d]	Mast cells, fibroblasts, macrophages, and pancreatic cells
Mid1	Mast cells
Hba-a1	Mast cells
Hbb-b family[e]	Mast cells
Pla2g family[f]	Mast cells and basophilic leukemia cells
Erdr1	Mast cells
Fabp family[g]	Basophilic leukemia cells, mast cells, neural stem cells, and oligodendrocytes

[a] In the following annotations, genes obtained in the UC group are emphasized in bold.

[b] Top DEGs were matched to exosome biomarker database, ExoCarta (http://exocarta.org/index.html).

[c] **Gm25018**, **Gm22291**, **Gm26391**, **Gm25212**, **Gm28661**, **Gm28437**, **Gm29216**, **Gm28439**, **Gm10925**, **Gm47283** and **Gm49405** are not included in ExoCarta, but many genes in Gm family are traceable.

[d] Many genes in **Atp8** and **Atp6** families are traceable.

[e] **Hbb-bs** is not included in ExoCarta, but Hbb-b1 and HBB-b2 are traceable.

[f] **Pla2g4b** is not included in ExoCarta, but Pla2g6 and Pla2g1b are traceable.

[g] **Fabp4** is not included in ExoCarta, but Fabp5 is traceable.

Table S8. Simulation details of the molecular assemblies examined in this study.

No.	Molecular assembly	Number of atom	Size of the simulation box (Å^3) [a]	Simulation time (ns)
1	WT pHLIP	19,810	50×51×58	100
2	D-S v1	19,489	50×51×58	100
3	D-S v2	19,486	50×51×58	100
4	D-S v3	19,498	50×51×58	100
5	D-S v4	19,486	50×51×58	100

[a] Size of the water boxes guarantees a minimum distance of 15 Å from any atom of the peptide to any edge of the simulation box.

Table S9. Free-energy calculations for the molecular assemblies of WT pHLIP, D-S v1, and D-S v2.

No.	Molecular assembly	Number of atom	Size of the simulation box (Å^3) [a]	Simulation time (ns)
1	WT pHLIP	48,886	72×69×104	200
2	D-S v1	43,723	81×75×77	200
3	D-S v2	43,492	76×75×82	200

[a] Size of the water boxes guarantees a minimum distance of 25 Å from any atom of the peptide to any edge of the simulation box.

Table S10. Peptide sequences synthesized in this study.

Peptide	Sequence[a]
WT pHLIP	GGEQNPIYWARYADWLFTTPLLLLDLALLVDADEGT
D-S v1	GGEQNPIYWARYADWLFTTPFGLLDLAHLVAADEGT
D-S v2	GGEQNPIYWARYADWLFHTPFGLLDLAHLVAADEGT

[a] The mutation sites in D-S v1 and D-S v2 are highlighted in red.

Figs. S1-17

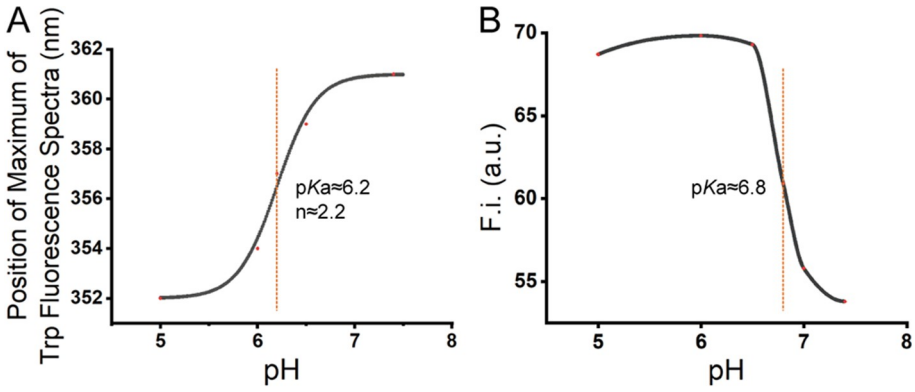


Fig. S1 p*K*a calculations. (A) p*K*a value and hill coefficient (n) of WT pHLIP calculated by Henderson-Hasselbalch equation. (B) p*K*a value of D-S v1 fitted by the curve fitting tool (Shape-preserving, PCHIP) in MATLAB (2018b version, MathWorks. Inc, USA).

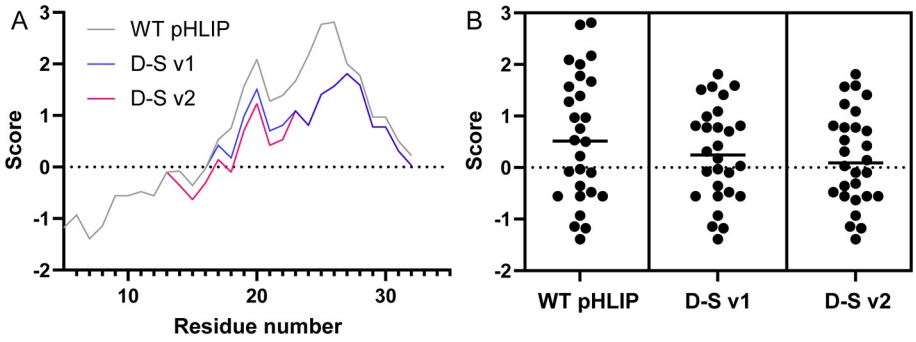


Fig. S2 Hydrophobicity analysis. (A) Hydrophobicity scores of each residue. (B) Hydrophobicity scores for WT pHLIP, D-S v1, and D-S v2. The score is positively associated with hydrophobicity. The transverse lines in (B) denote the mean hydrophobic values for the three peptides.

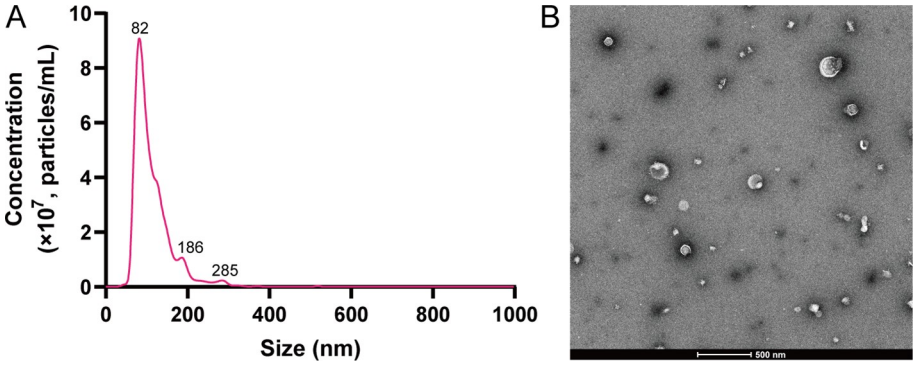


Fig. S3 Characterizations for EVs derived from A549 cells. (A) NTA. (B) TEM.

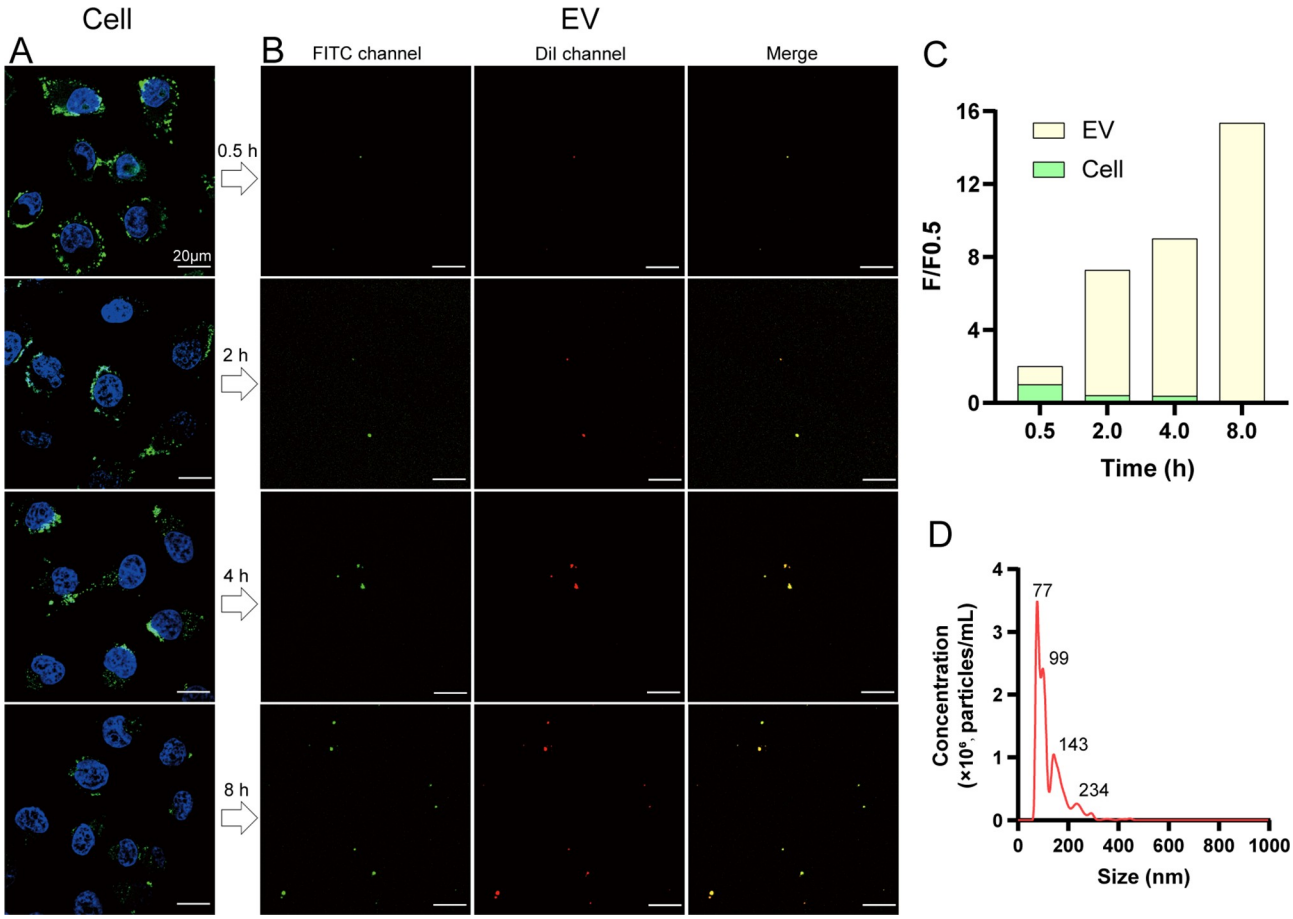


Fig. S4 FITC signals in cells and their secreted EVs with time course. (A) Cell experiments. The green fluorescence of FITC in denotes the D-S v1 on cell surfaces. (B) EV experiments. EVs were dyed by DiI. The merged yellow signals detected in the supernatants suggest that the peptides on cells have transferred to their secreted EVs. (C) Time course of the changes in green fluorescence in cell surface and the merged yellow fluorescence, respectively. (D) NTA results for the produced EVs.

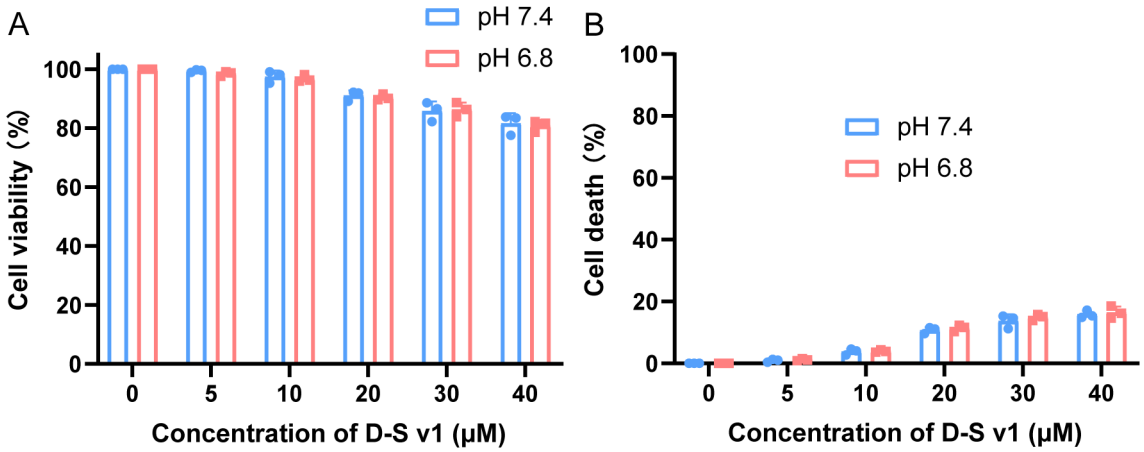


Fig. S5 Cell viability (A) and cytotoxicity (B) tests for D-S pHLIP at different concentrations under pH 7.4 and pH 6.8 conditions. Cell Counting Kit-8 and MTT assay kits were employed to measure the cell viability and cytotoxicity, respectively, according to the manufacturer's instructions. The cells grew well below 10 μM and showed minimal cytotoxicity (< 4%) under both pH conditions.

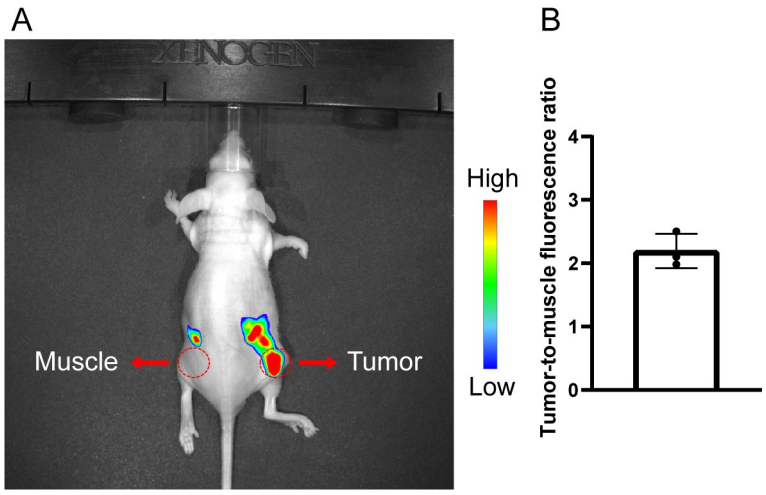


Fig. S6 *In vivo* fluorescence imaging of D-S v1 conjugated by Cy5.5 (D-S v1@Cy5.5). (A) Fluorescent image of the tumor-bearing mouse which was i.p. injected with D-S v1@Cy5.5. (B) Measurement of the tumor-to-muscle fluorescence ratios. D-S v1@Cy5.5 was i.p. injected into tumor model mice and the fluorescence on the tumor and non-tumor sites was imaged at 24 h post injection to calculate the tumor-to-muscle fluorescence ratios. The results are expressed as mean ± SEM (n = 3 biological replicates).

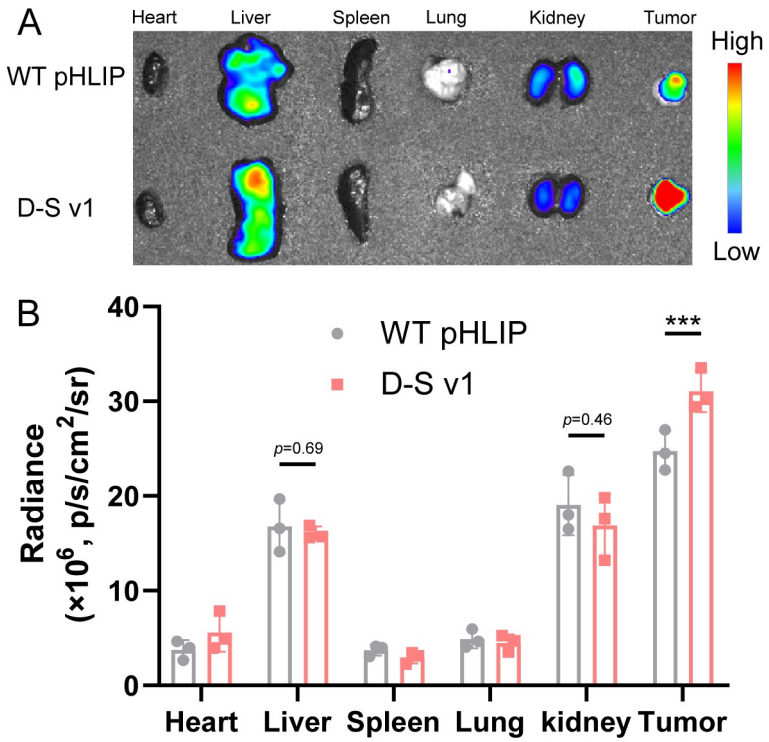


Fig. S7 Distribution of WT pHLIP and D-S v1 in major organs. (A) Organ imaging. The mice, whose sera were withdrawn at 24 h time point, were dissected. The organs were incubated with FITC@SA (1 μg/mL) for 10 min, and washed three times using PBS (pH 6.2 and 7.4 for WT pHLIP and D-S v1 groups, respectively). The FITC signals in show that the peptides are distributed in the tumor, liver, and kidney. (B) Quantified results. D-S v1 possesses a higher tumor-accumulation ability than WT pHLIP. Additionally, the peptides were metabolized mainly through the liver and kidney. The results are expressed as mean ± SEM (n = 3 biological replicates). ***p<0.001.

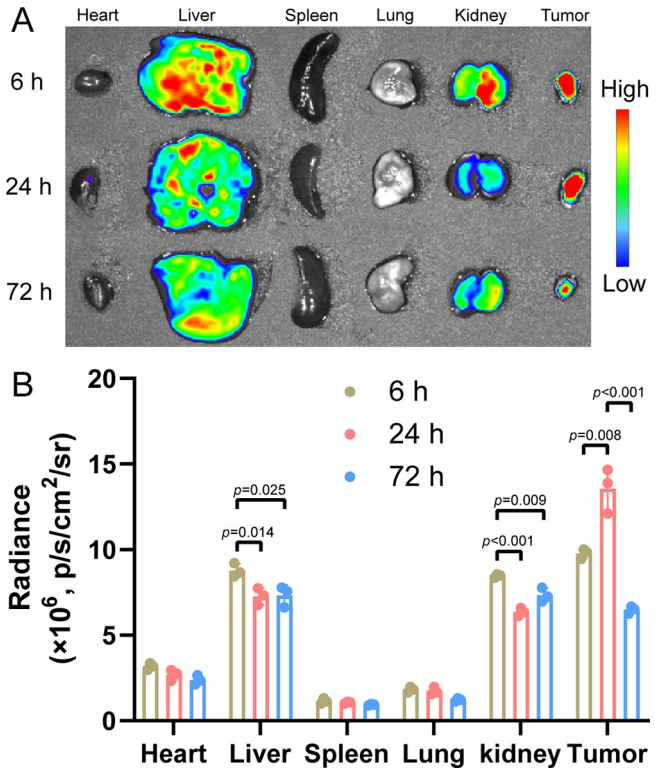


Fig. S8 Dynamic distribution of D-S v1 in major organs at different time points. (A) Fluorescent imaging of the organs at 6 h, 24 h, and 72 h. (B) Quantification results for the fluorescence intensities in (A). The organs obtained at different time points were first incubated with FITC@SA (1 μg/mL) for 10 min and then washed three times using PBS. The results are expressed as mean ± SEM (n = 3 biological replicates). ***p<0.001.

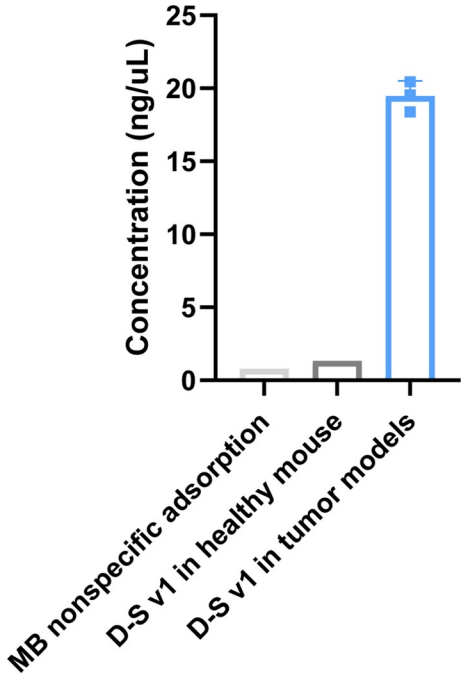


Fig. S9 RNA concentration measurement for the controls. MB nonspecific adsorption: MB@SA was incubated directly with the serum of a lung cancer mouse model without injection of peptides@biotin. This result indicates that the MB@SA has slight nonspecific adsorption towards the EVs. D-S v1 in healthy mouse: we injected the D-S v1@biotin into a healthy mouse, withdrew its blood at 24 h post injection, and collected the D-S v1@biotin-tagged EVs using MB@SA. This result shows that D-S v1 can hardly capture EVs in the healthy mouse model, suggesting the excellent tumor specificity of this D-S pHLIP system. D-S v1 in tumor models (Fig. 4B in the main text) was used as a reference. The results are expressed as mean ± SEM (n = 3 biological replicates).

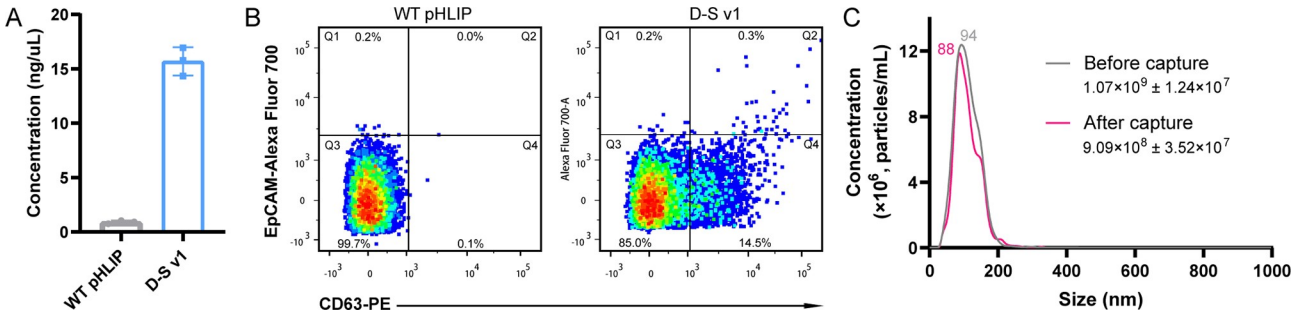


Fig. S10 Identification of the captured EVs from mice with breast cancer. (A) RNA detection. (B) FCM. (C) NTA. The breast cancer models were built by inoculation of MDA-MB-231. The sera were withdrawn at the optimal time point, i.e., 24 h. The results of (A) indicate that the D-S v1 has captured abundant EVs, while the WT pHLIP did not. In (B), there are significant fluorescence signals of CD63 and EpCAM in the D-S v1 group. These results indicate that the D-S pHLIP system is amenable to the breast cancer model as well. The results are expressed as mean ± SEM (n = 3 biological replicates).

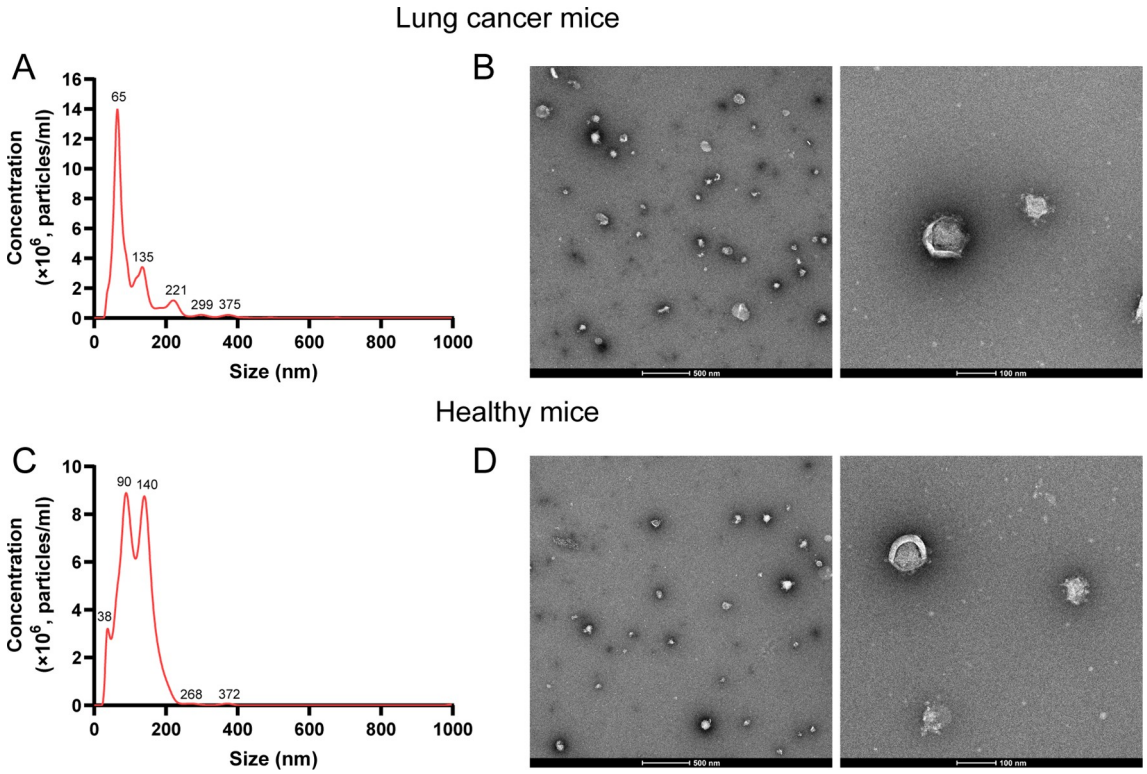


Fig. S11 NTA and TEM results for the EV samples isolated by UC. (A and B) Lung cancer model. (C and D) Healthy mice.

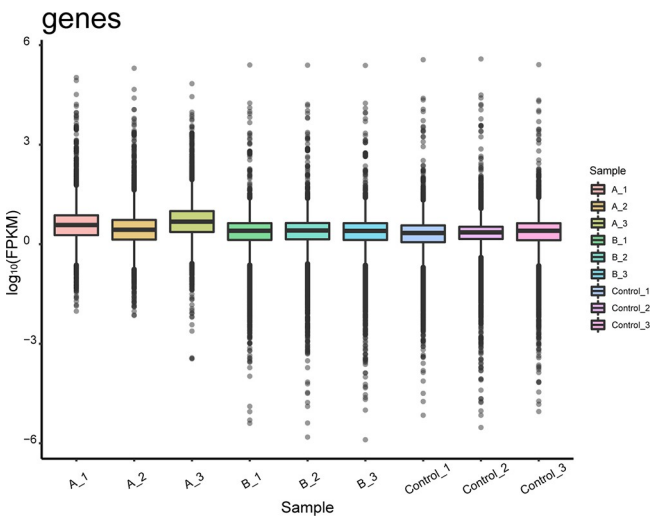


Fig. S12 Analysis of gene expression levels for EV samples captured by D-S v1 (A1_3: lung cancer models) and isolated by UC (B1_3: lung cancer models, and Control 1_3: healthy mice). These results indicate that the samples have good biological repeatability.

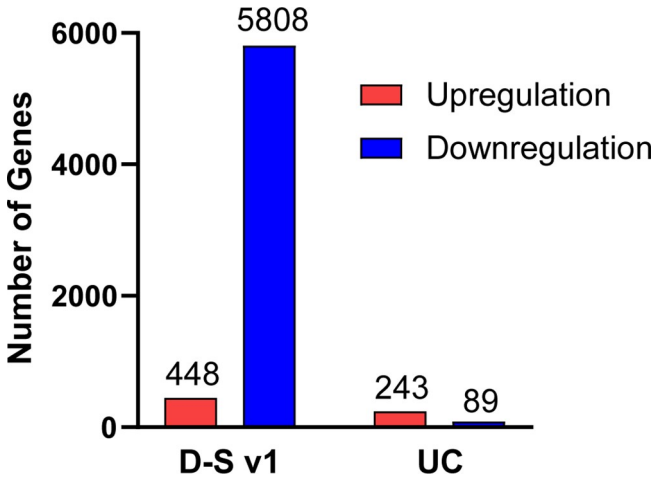


Fig. S13 DEG numbers in D-S v1 and UC groups, respectively. A false-discovery rate-corrected $q < 0.05$ and a FC > 2 are used to define upregulation (red) and downregulation (blue), here and in the forthcoming figures.

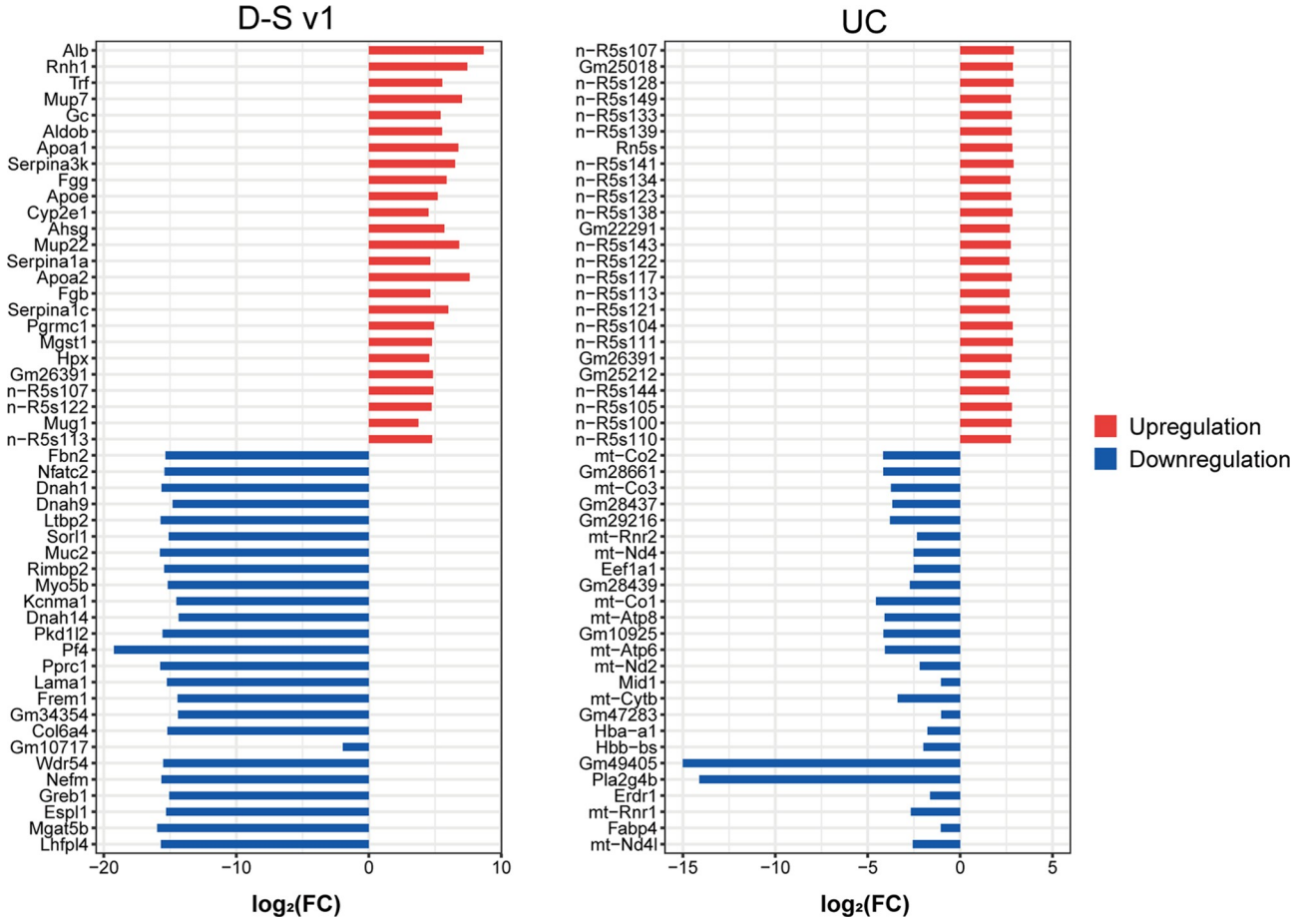


Fig. S14 Top 50 DEGs in D-S v1 and UC groups, respectively.

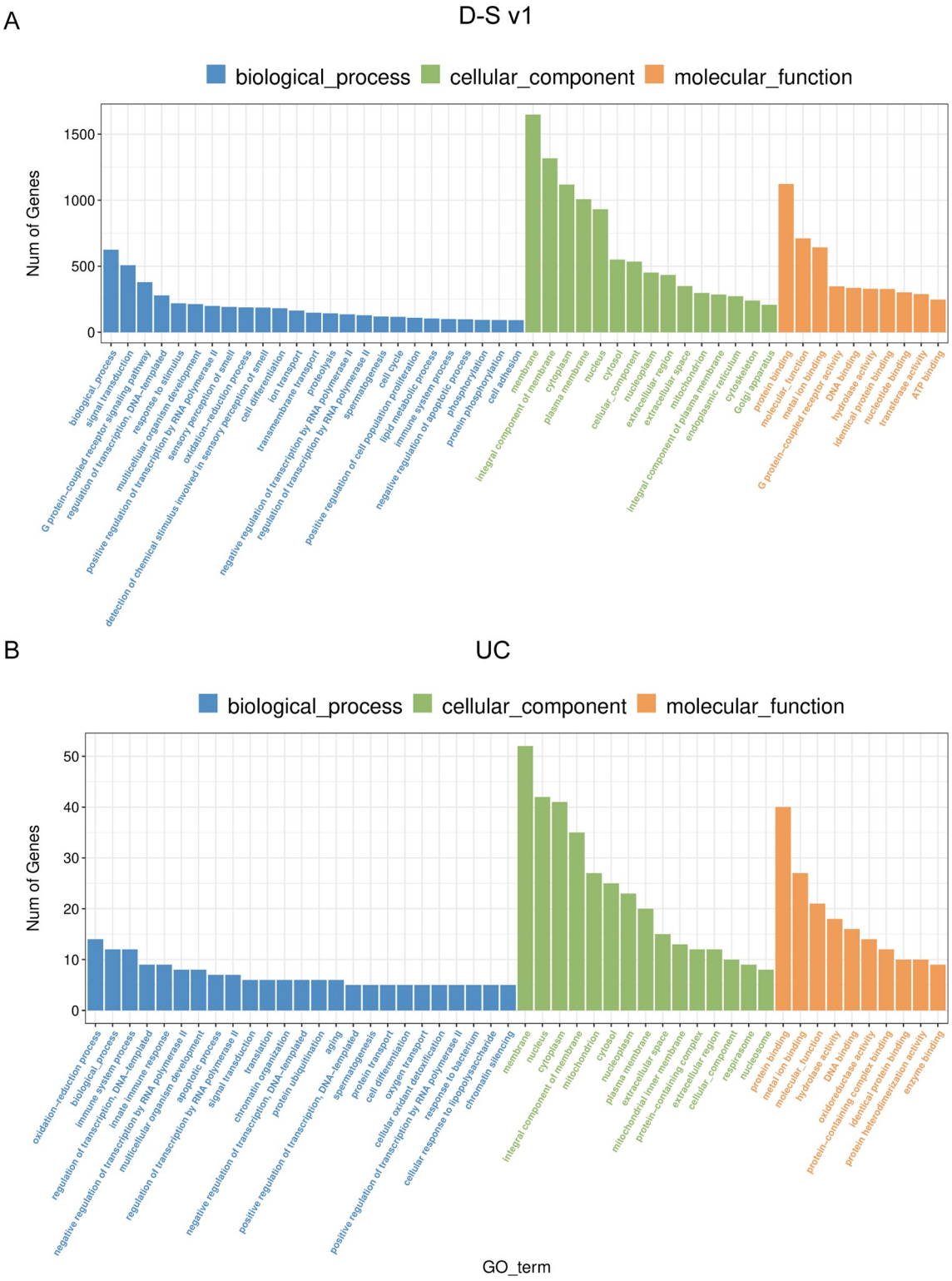


Fig. S15 GO analysis based on the identified DEGs. (A) D-S v1 group. (B) UC group.

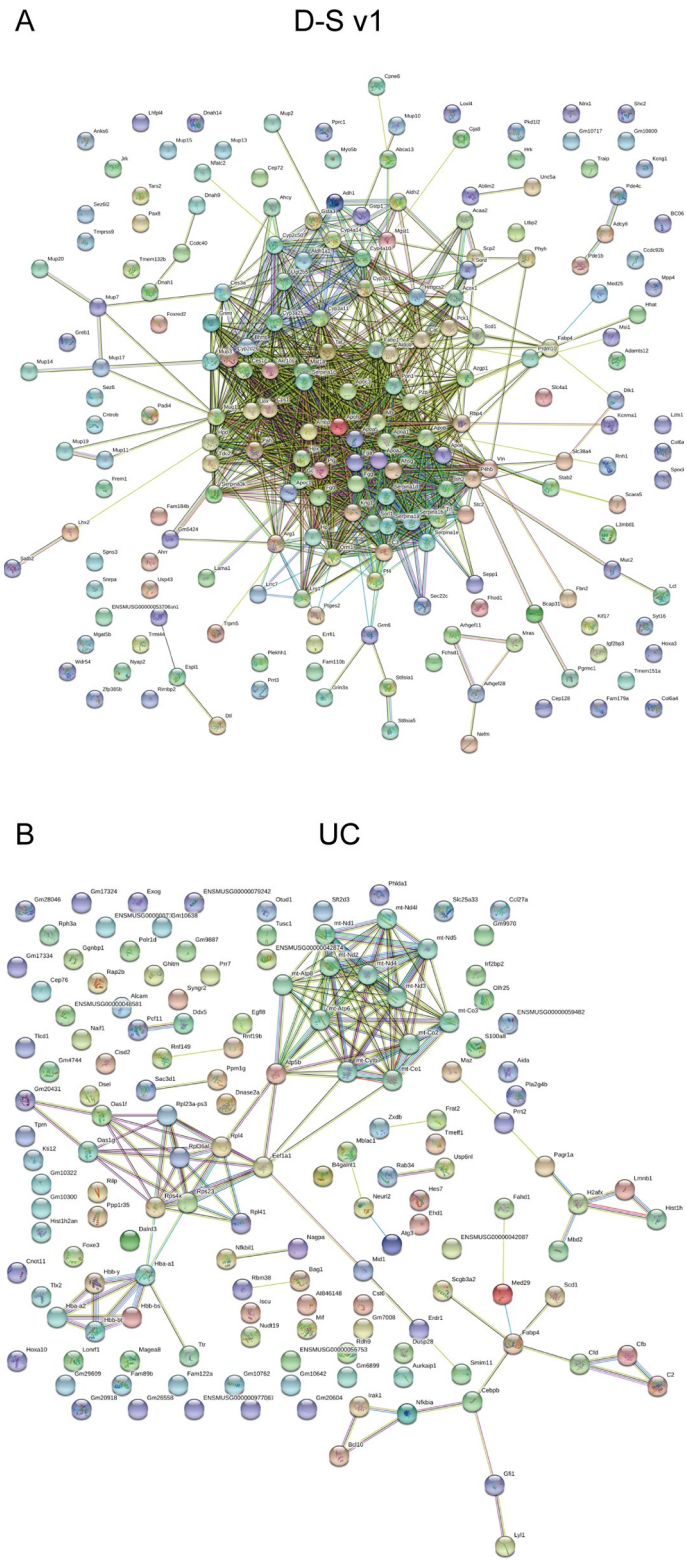


Fig. S16 Protein-protein interaction network diagrams. (A) D-S v1 group. (B) UC group.

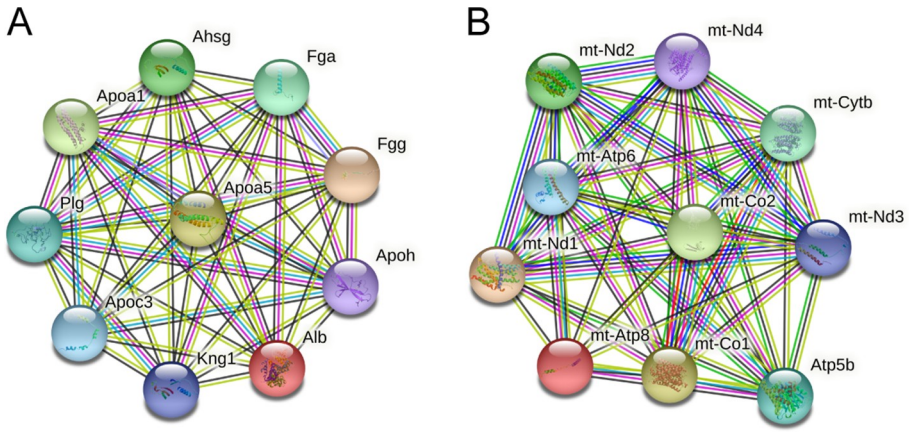


Fig. S17 Hub genes generated from PPI network. (A) D-S v1 group. (B) UC group.

Captions for Movies S1 to S2:

Movie S1: conformational changes of D-S v1 based on its trajectory in 100 ns MD simulation.

Movie S2: conformational changes of D-S v2 based on its trajectory in 100 ns MD simulation.

Captions for characterization of peptides:

Analytic characterization of peptides (WT pHLIP, D-S v1, and D-S v2) and functionalized ones (WT pHLIP@biotin, D-S v1@biotin, and D-S v1@Cy5.5) used in this study.

SI References

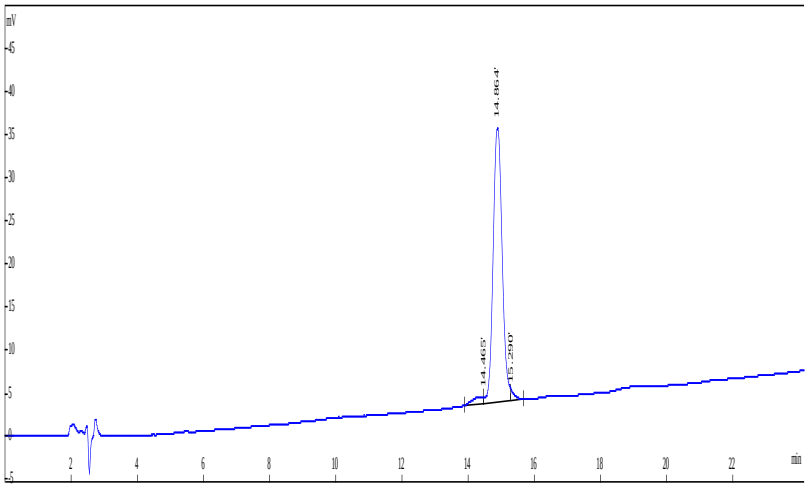
1. A. Waterhouse *et al.*, SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, 296-303 (2018).
2. S. Bienert *et al.*, The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* **45**, D313-d319 (2017).
3. N. Guex, M. C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30**, 162-173 (2009).
4. G. Studer *et al.*, QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics (Oxford, England)* **36**, 1765-1771 (2020).
5. M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, T. Schwede, Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 10480 (2017).
6. R. Anandakrishnan, B. Aguilar, A. V. Onufriev, H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537-541 (2012).
7. J. Myers, G. Grothaus, S. Narayanan, A. Onufriev, A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* **63**, 928-938 (2006).
8. J. C. Gordon *et al.*, H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33**, W368-371 (2005).
9. B. C. Knott, M. F. Crowley, M. E. Himmel, J. Ståhlberg, G. T. Beckham, Carbohydrate-protein interactions that drive processive polysaccharide translocation in enzymes revealed from a computational study of cellobiohydrolase processivity. *J. Am. Chem. Soc.* **136**, 8810-8819 (2014).
10. Z. Zong *et al.*, Lysine Mutation of the Claw-Arm-Like Loop Accelerates Catalysis by Cellobiohydrolases. *J. Am. Chem. Soc.* **141**, 14451-14459 (2019).
11. E. Hatcher, O. Guvench, A. D. Mackerell, Jr., CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J. Chem. Theory Comput.* **5**, 1315-1327 (2009).
12. K. Vanommeslaeghe *et al.*, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671-690 (2010).
13. O. Guvench *et al.*, Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **29**, 2543-2564 (2008).
14. O. Guvench, E. R. Hatcher, R. M. Venable, R. W. Pastor, A. D. Mackerell, CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **5**, 2353-2370 (2009).
15. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **79**, 926-935 (1983).
16. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38 (1996).
17. H. Fu, X. Shao, W. Cai, C. Chipot, Taming Rugged Free Energy Landscapes Using an Average Force. *Acc. Chem. Res.* **52**, 3254-3264 (2019).
18. M. An, D. Wijesinghe, O. A. Andreev, Y. K. Reshetnyak, D. M. Engelman, pH-(low)-insertion-peptide (pHLIP) translocation of membrane impermeable phalloidin toxin inhibits cancer cell proliferation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20246-20250 (2010).
19. O. A. Andreev *et al.*, Mechanism and uses of a membrane peptide that targets tumors and other acidic tissues in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7893-7898 (2007).
20. P. Shannon *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498-2504 (2003).

HPLC REPORT

Product Name: WT pHLIP
Instrument No: 0200194
Lot No : P210819-CL341357
Column : 4.6*250mm C18
Solvent A : 0.1%Trifluoroacetic in 100% Acetonitrile
Solvent B : 0.1%Trifluoroacetic in 100% Water

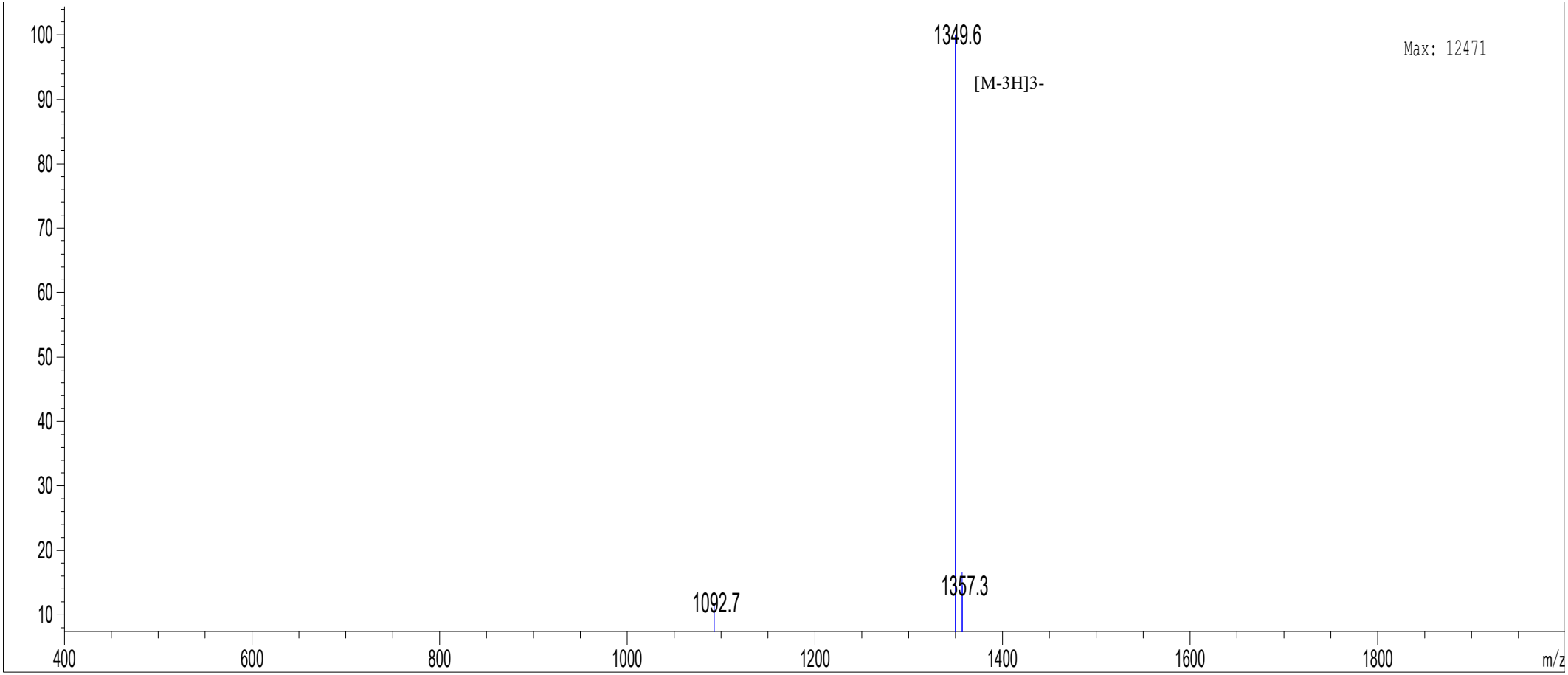
Gradient :	A	B
0.01min	45%	55%
25min	70%	30%
25.01min	100%	0%
30.0min	STOP	

Flow rate :1.0ml/min
Wavelength :220nm
Volume: 10μl



Rank	Time	Name	Conc.	Area	Height
1	14.465		2.883	17212	683
2	14.864		95.86	572254	31758
3	15.290		1.257	7503	1016
Total			100	596969	33457

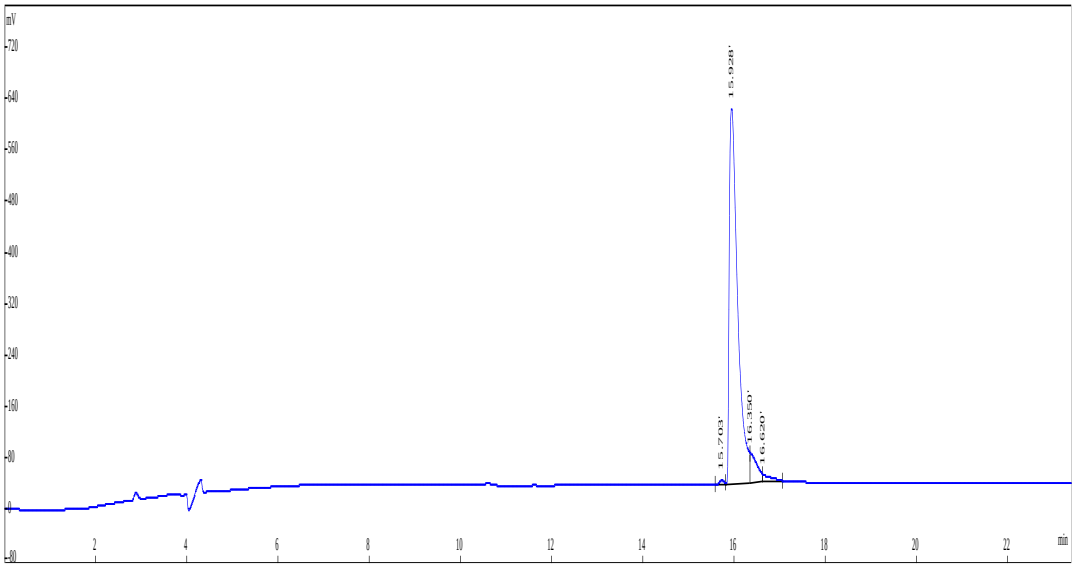
MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2021-09-02	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	WT pHLIP	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4051.99	CDL Temp.:	250 °C	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P210819-CL341357	Block Temp.:	200 °C		

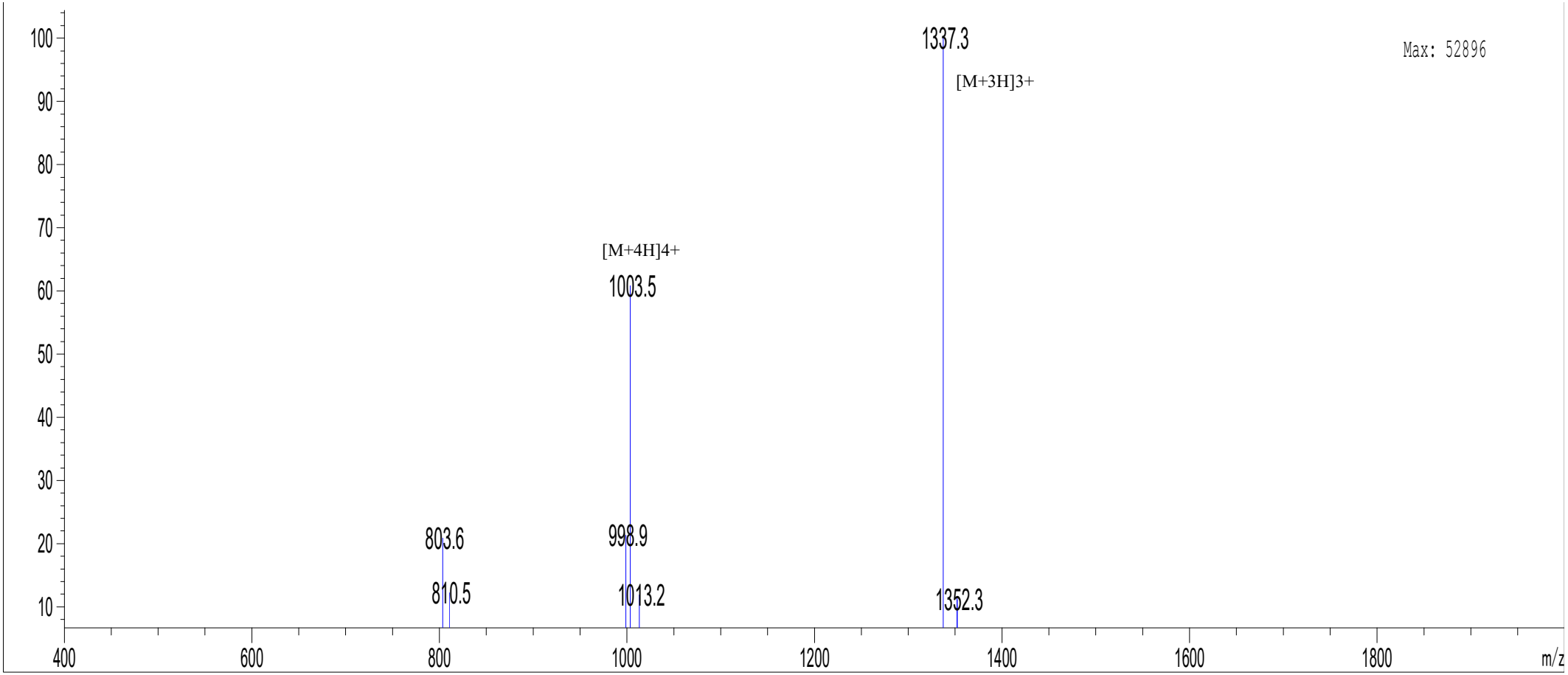
HPLC REPORT

Product Name	:L28H/D31A/L21F/L22G(D-S v1)		
Lot No	:P210223-SJ872141		
Column	:Gemini-NX 5μ C18 110A, 4.6*250mm		
Solvent A	:0.1%Trifluoroacetic in 100% Acetonitrile		
Solvent B	:0.1%Trifluoroacetic in 100% Water		
Gradient	:	A	B
	0.01min	20%	80%
	25min	80%	20%
	25.01min	100%	0%
	30min	Stop	
Flow rate	:1.0ml/min		
Wavelength	:220nm		
Volume	:20ul		



Rank	Time	Conc.	Area	Height
1	15.703	0.4325	33207	5700
2	15.928	92.4244	7095710	584222
3	16.350	5.5295	424514	43950
4	16.620	1.6136	123882	10799
Total		100	7677313	644671

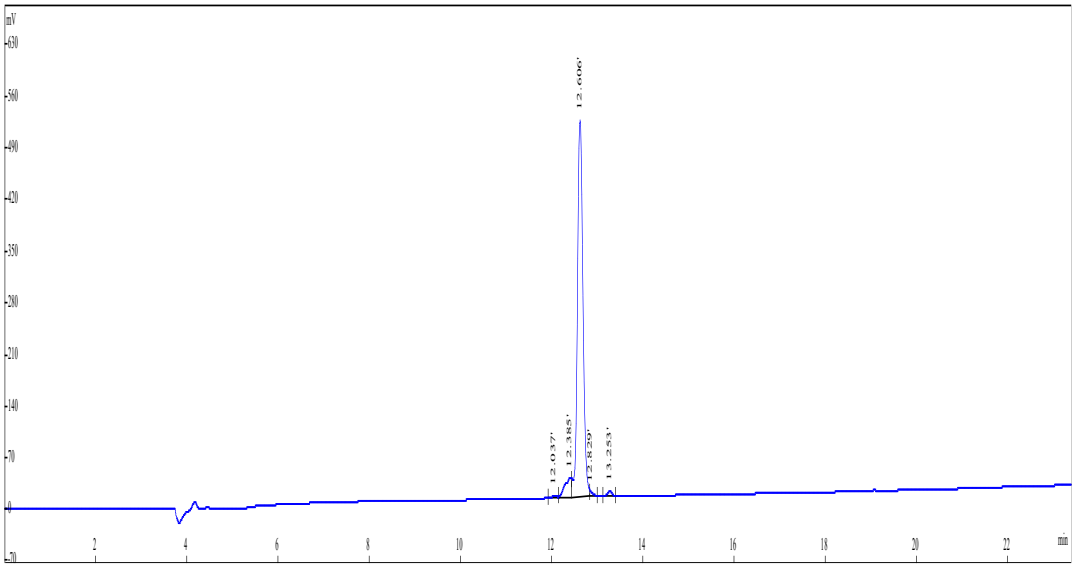
MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2021-03-05	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	L28H/D31A/L21F/L22G (D-S v1)	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4009.38	CDL Temp.:	250 °C	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P210223-SJ872141	Block Temp.:	200 °C		

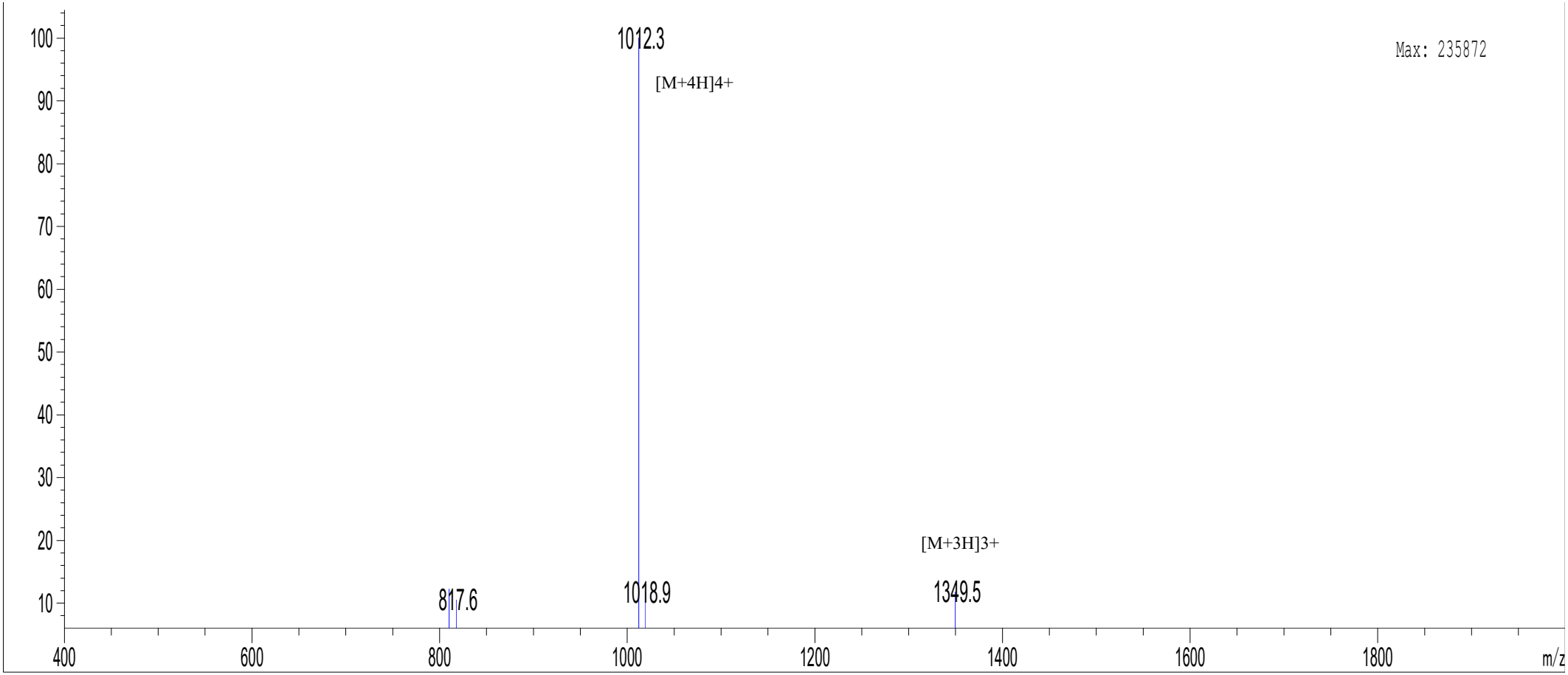
HPLC REPORT

Product Name	:L28H/D31A/L21F/L22G/T18H(D-S v2)
Lot No	:P210223-SJ872142
Column	:Gemini-NX 5μ C18 110A, 4.6*250mm
Solvent A	:0.1%Trifluoroacetic in 100% Acetonitrile
Solvent B	:0.1%Trifluoroacetic in 100% Water
Gradient	: A B
	0.01min 30% 70%
	25min 90% 10%
	25.01min 100% 0%
	30min Stop
Flow rate	:1.0ml/min
Wavelength	:220nm
Volume	:20ul



Rank	Time	Conc.	Area	Height
1	12.037	0.5116	21871	3004
2	12.385	6.7703	289443	26215
3	12.606	91.3594	3905793	508415
4	12.829	0.4845	20712	5785
5	13.253	0.8742	37374	5939
Total		100	4275193	549358

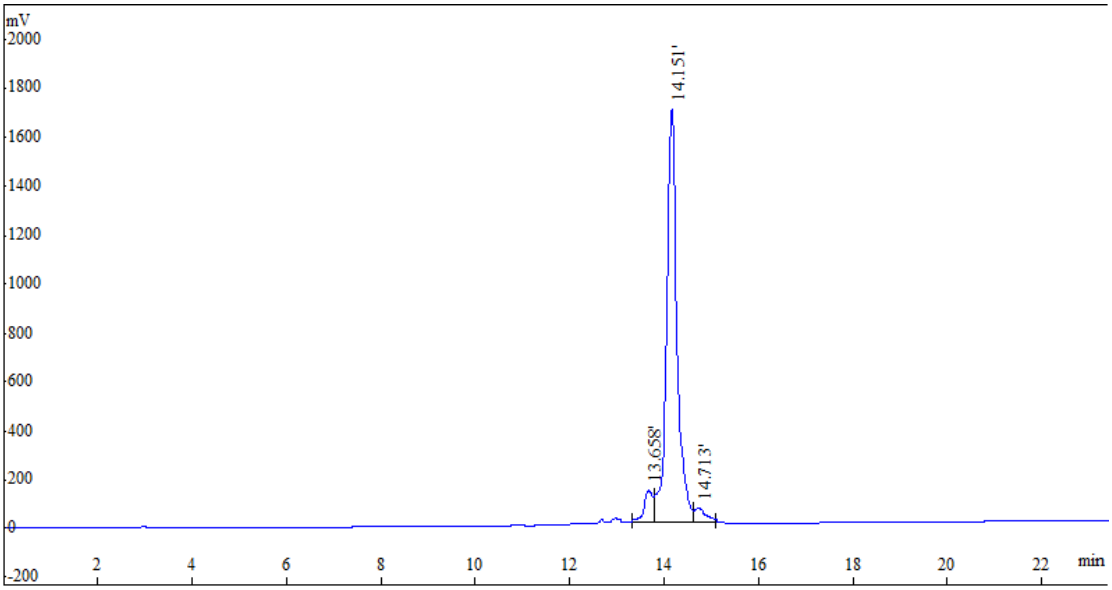
MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2021-03-09	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	L28H/D31A/L21F/L22G/T18H (D-S v2)	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4045.42	CDL Temp.:	250 °C	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P210223-SJ872142	Block Temp.:	200 °C		

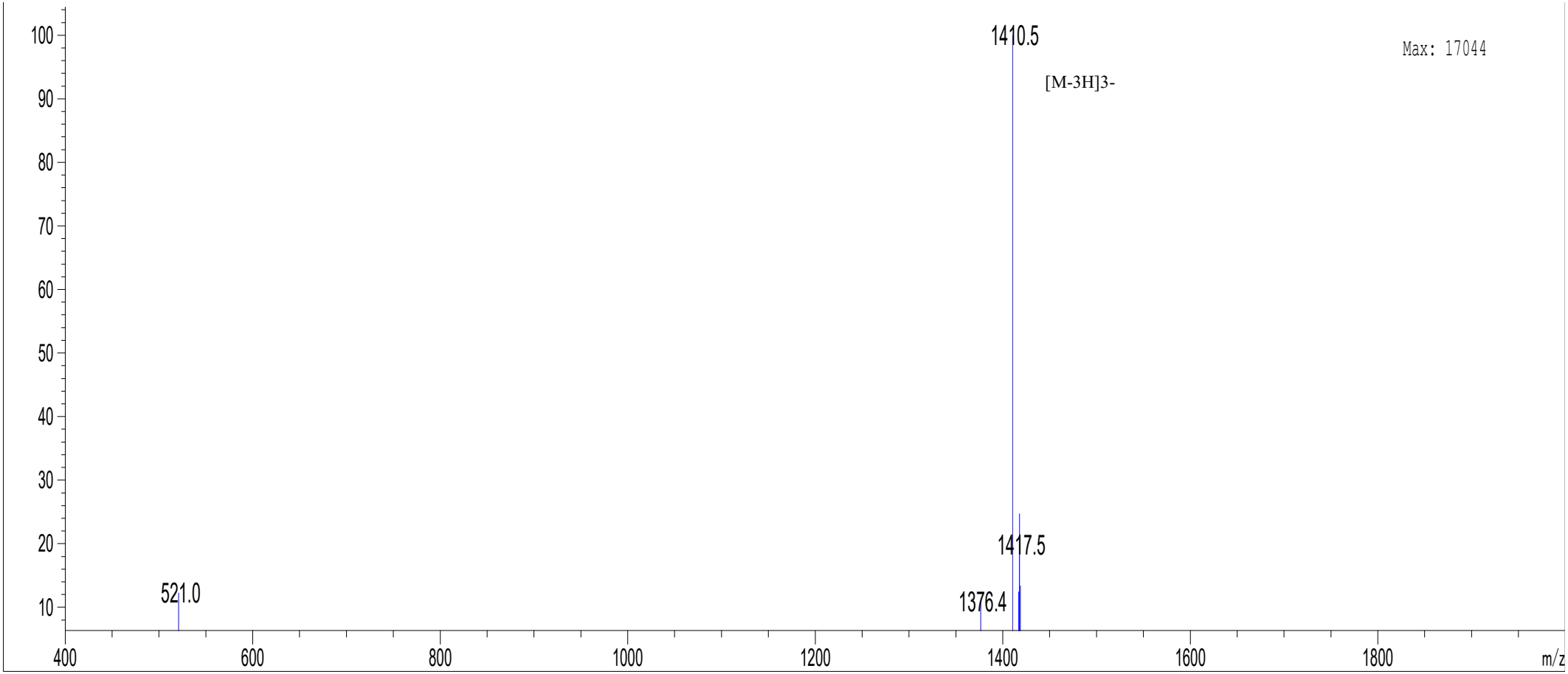
HPLC REPORT

Product Name	:WT pHLIP@biotin
Lot No	:P201110-HS828717
Column	:Gemini-NX 5μ C18 110A, 4.6*250mm
Solvent A	:0.1%Trifluoroacetic in 100% Acetonitrile
Solvent B	:0.1%Trifluoroacetic in 100% Water
Gradient	: A B
	0.01min 30% 70%
	25min 90% 10%
	25.01min 100% 0%
	30min Stop
Flow rate	:1.0ml/min
Wavelength	:220nm
Volume	:20ul



Rank	Time	Conc.	Area	Height
1	13.658	5.8226	1572138	122694
2	14.151	91.3272	24658992	1680803
3	14.713	2.8502	769568	50015
Total		100	27000698	1853512

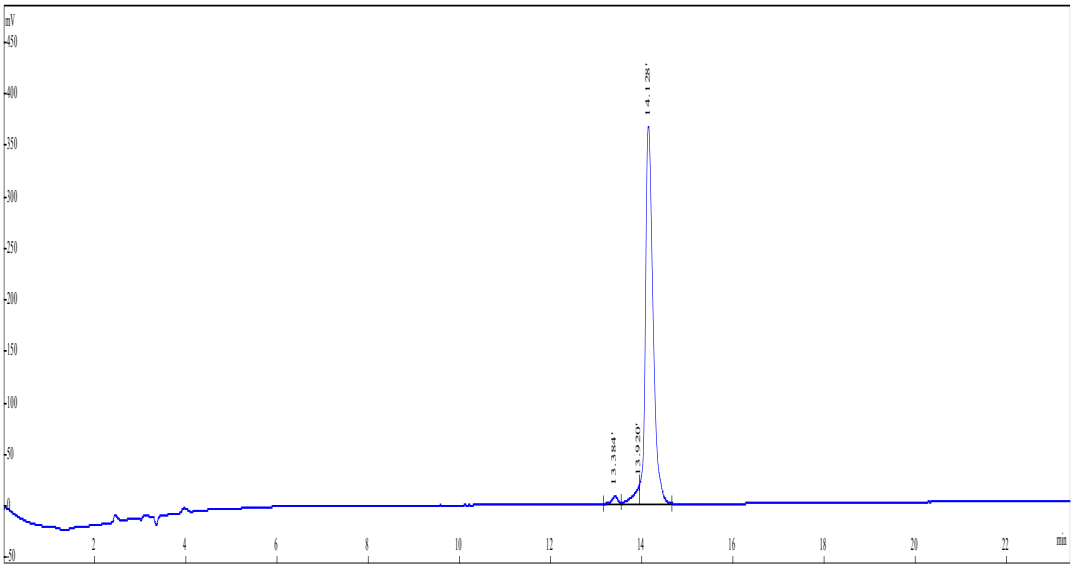
MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2020-11-24	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	WT pHLIP@biotin	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4234.77	CDL Temp.:	250 °C	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P201110-HS828717	Block Temp.:	200 °C		

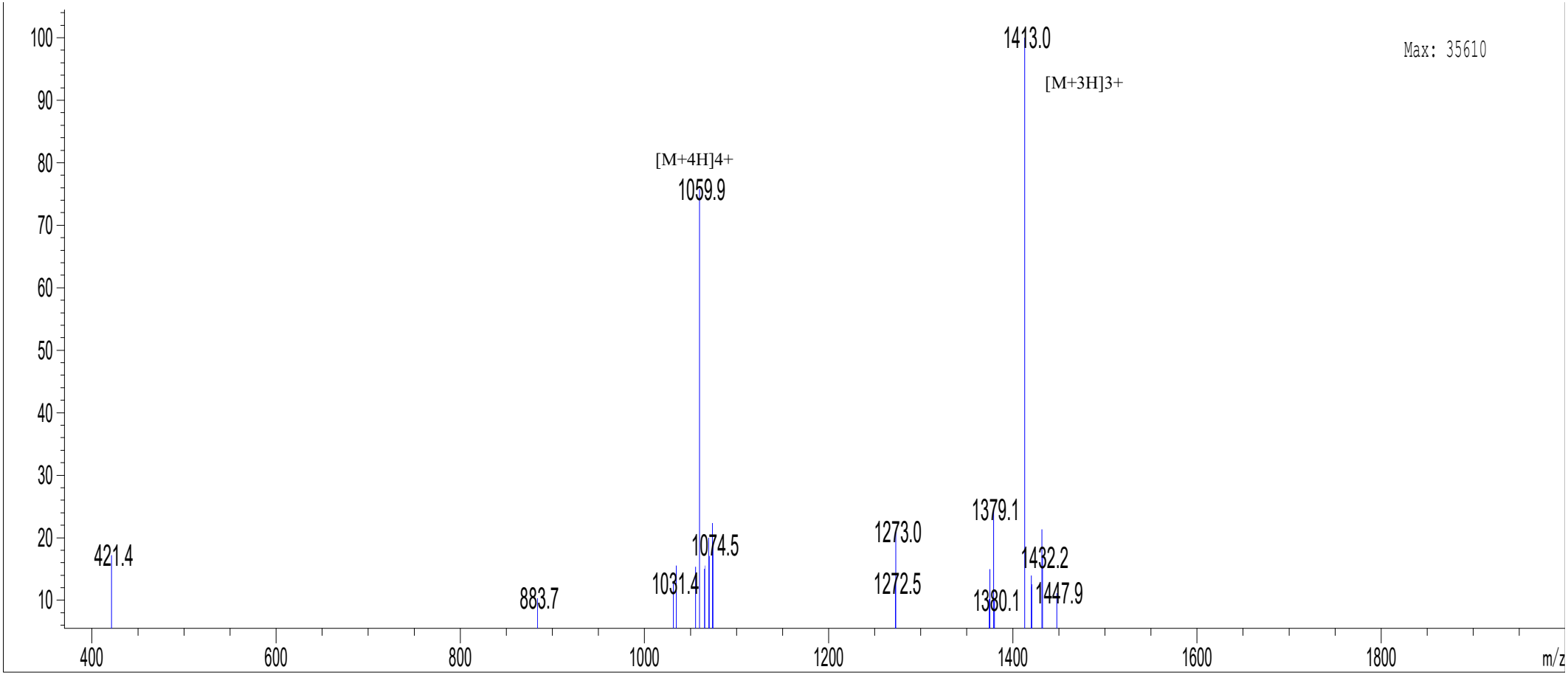
HPLC REPORT

Product Name	:D-S v1@biotin
Lot No	:P210407-HS886800
Column	:Gemini-NX 5μ C18 110A, 4.6*250mm
Solvent A	:0.1%Trifluoroacetic in 100% Acetonitrile
Solvent B	:0.1%Trifluoroacetic in 100% Water
Gradient	: A B
	0.01min 30% 70%
	25min 90% 10%
	25.01min 100% 0%
	30min Stop
Flow rate	:1.0ml/min
Wavelength	:220nm
Volume	:20ul



Rank	Time	Conc.	Area	Height
1	13.384	1.3444	56097	6811
2	13.920	4.0139	167488	16590
3	14.128	94.6417	3949156	364934
Total		100	4172741	388335

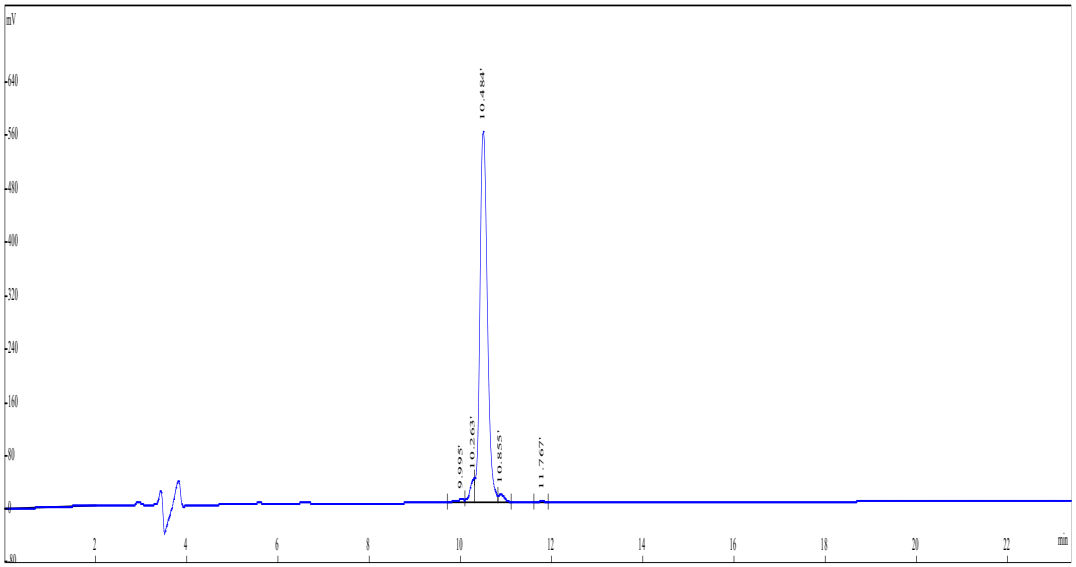
MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2021-04-20	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	D-S v1@biotin	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4235.68	CDL Temp.:	250 °C	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P210407-HS886800	Block Temp.:	200 °C		

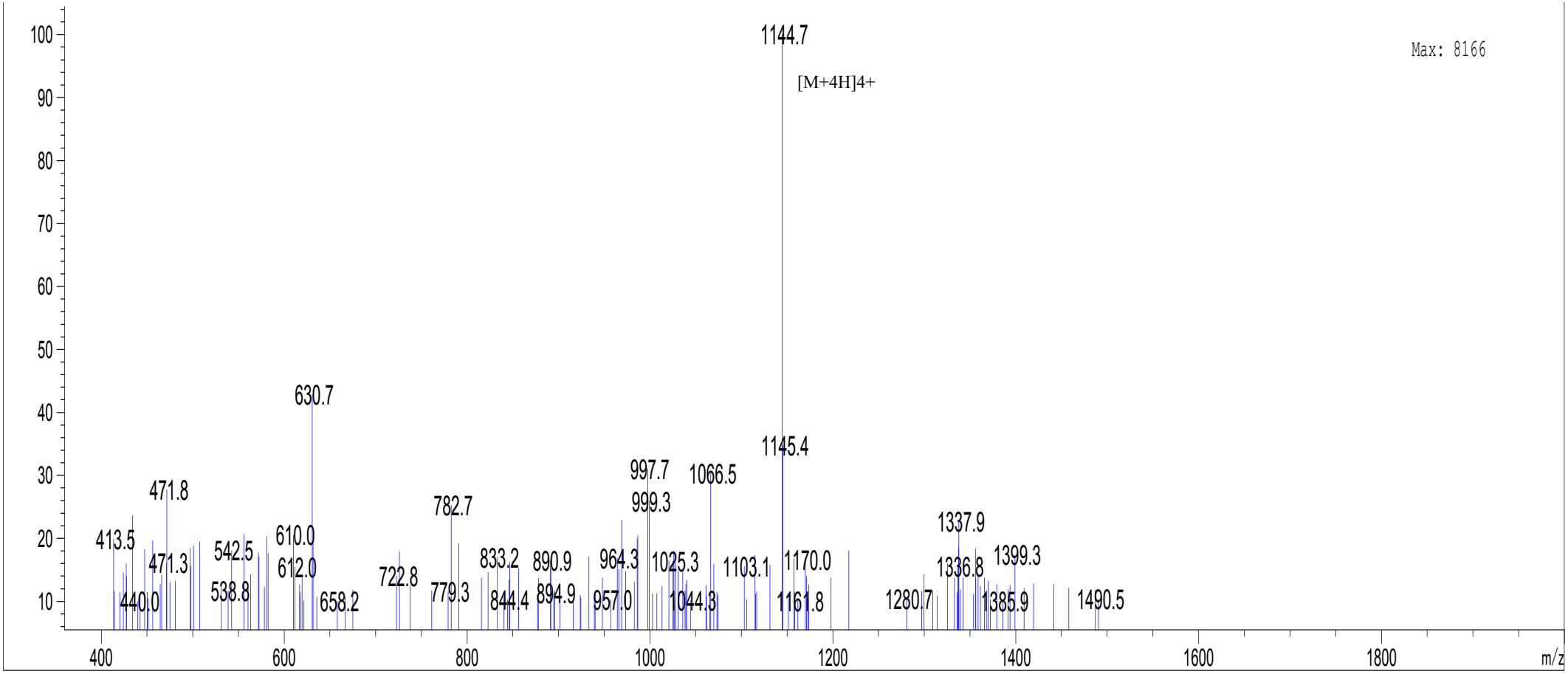
HPLC REPORT

Product Name	: D-S v1@Cy5.5
Lot No	:P210906-HS931414
Column	:YMC-Pack C4 4.6mm*250mm,5μm 30mm
Solvent A	:0.1%Trifluoroacetic in 100% Acetonitrile
Solvent B	:0.1%Trifluoroacetic in 100% Water
Gradient	: A B
	0.01min 30% 70%
	25min 90% 10%
	25.01min 100% 0%
	30min Stop
Flow rate	:1.0ml/min
Wavelength	:220nm
Volume	:20ul



Rank	Time	Conc.	Area	Height
1	9.995	0.9198	58304	4460
2	10.263	4.0839	258876	34281
3	10.484	93.2430	5910569	554044
4	10.855	1.4372	91105	11420
5	11.767	0.3161	20038	2463
Total		100	6338892	606668

MASS SPECTROMETRY REPORT



Sample Description		Instrument	Agilent-6125B		
Analyzed date:	2021-09-26	Probe:	ESI	Probe Bias:	＋4.5kv
Analyst:	YU	Nebulizer Gas Flow:	1.5L/min	Detector:	1.5kv
Sample:	D-S v1@Cy5.5	CDL:	-20.0v	T. Flow:	0.2ml/min
M.W.:	4575.23	CDL Temp.:	250 ºC	B. Conc.:	50%H2O/50%ACN
Lot. No.:	P210906-HS931414	Block Temp.:	200 ºC		