

# Supporting Information

## **Inhibiting Matrix Metalloproteinase-2 Activation by Perturbing Protein-Protein Interactions Using a Cyclic Peptide**

Priyanka Sarkar,<sup>†</sup> Zhonghan Li,<sup>†</sup> Wendan Ren,<sup>‡</sup> Siwen Wang,<sup>†</sup> Shiqun Shao,<sup>†</sup> Jianan Sun,<sup>†</sup> Xiaodong Ren,<sup>§</sup> Nicole G. Perkins,<sup>†</sup> Zhili Guo,<sup>†</sup> Chia-En A. Chang,<sup>†</sup> Jikui Song<sup>‡</sup> and Min Xue<sup>†,\*§</sup>

<sup>†</sup> Department of Chemistry, University of California, Riverside, Riverside, California 92521, United States.

<sup>‡</sup> Department of Biochemistry, University of California, Riverside, Riverside, California 92521, United States.

<sup>§</sup> Medical College, Guizhou University, Guiyang, Guizhou 550025, China.

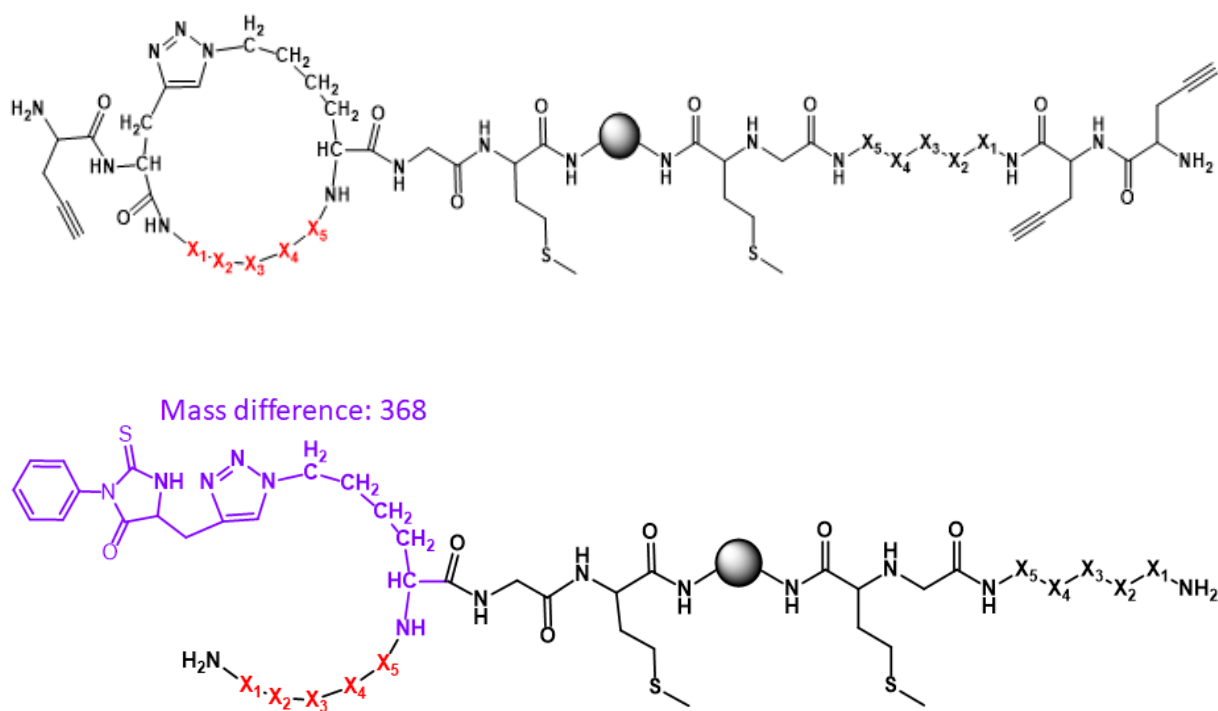
<b>Contents</b>	<b>Page</b>
1. Protein sequence alignment results	S2, S43
2. Structure of the peptide library	S3
3. Characterization data	S4, S6-S21, S27, S36
4. Additional experimental procedures and data	S5, S22-S26, S35, S37-S43

MMP-2 catalytic domain	Homology	Protein name
RKP KWDKNQITYRIIGYTPDLDPETVDDAFARAF QVWSDVTPLRFSRIHDGEADIMINFGRWEHGDG YPFDGKDGLLAHAFAPGTGVGGDSHFDDDELWS LGKGVGYSLFLVAAHEFGHAMGLEHSQDPGALM APIYTYTKNFRLSQDDIKGIQELYGASP	64%	MMP-9 (3-157)
	64%	MMP-13 (6-164)
	59%	MMP-3 (2-162)
	58%	MMP-12 (3-159)
	58%	MMP-20 (2-159)
	56%	MMP-8 (6-163)
	58%	MMP-10 (8-165)
	55%	MMP-1 (7-164)

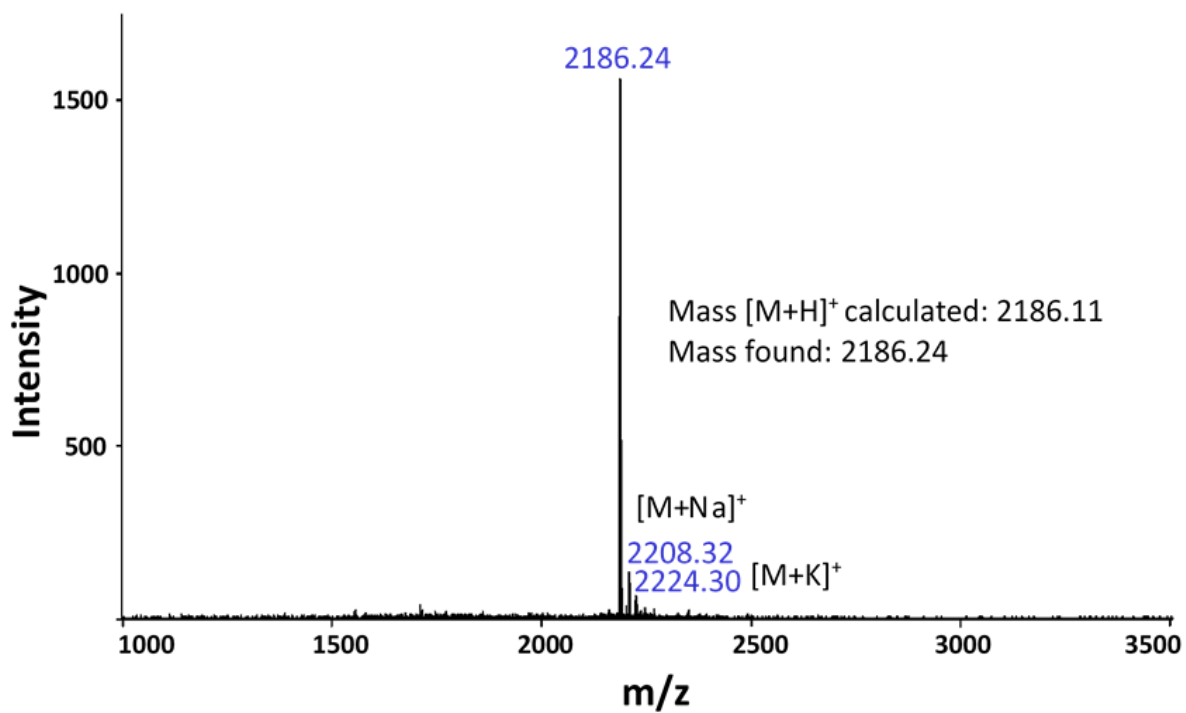
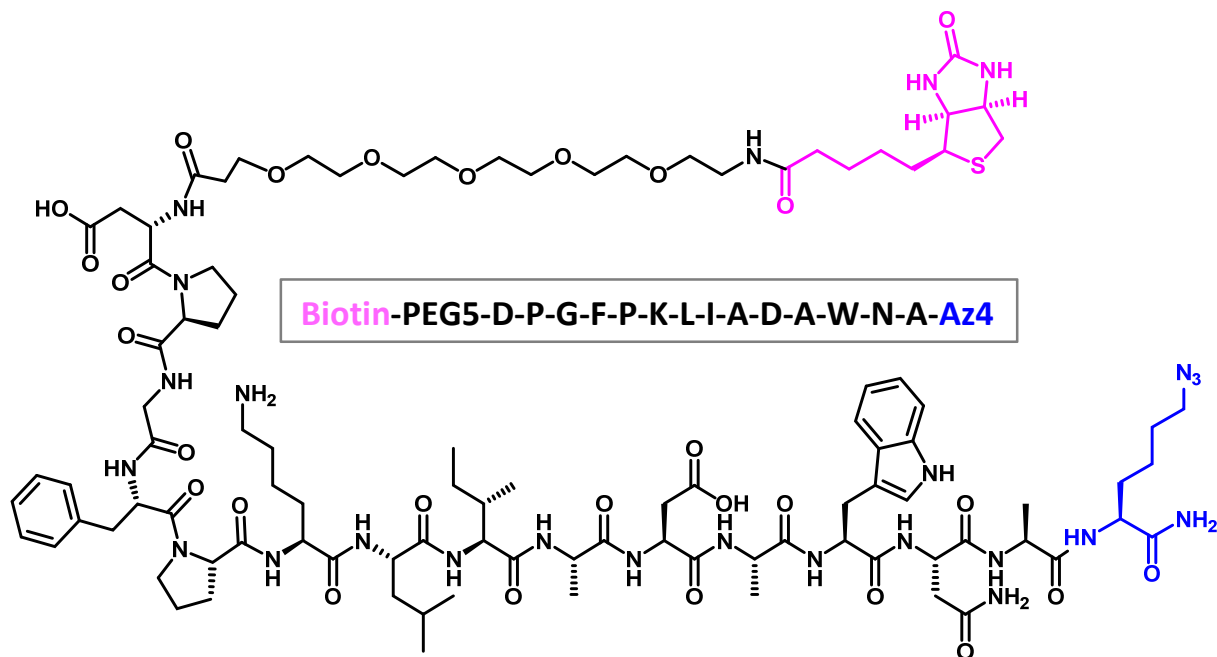
**Figure S1.** Sequence alignment and homology search result for the MMP-2 catalytic domain (R<sub>88</sub>-P<sub>250</sub>) using protein BLAST® against other MMPs. The catalytic domain sequence is shown in the left (light blue background). The MMPs with the sequence homology found from the search are shown in the right two columns.

D	P	G	F	P	K	L	I	A	D	A	W	N	A	Homology	Protein Name
	P	G	Y	P	K	L	I	R	D	W	W			57%	Unnamed protein (121-131)
D	P	G	P	F	P	K	L	I	E	W				50%	Cadherin-like protein (73-83)
D	P	G	Y	P	K	L	I							50%	Macrophage metalloelastase (135-142)
D	P	G	Y	P	K	M	I	A						50%	MMP 1 (171-179)
D	P	G	Y	P	K	L	I							50%	MΦ metalloelastase preproprotein (306-313)
D	P	G	Y	P	K	M	I	A						50%	MΦ metalloelastase active form (309-317)
D	P	G	Y	P	K	M	I	A						50%	MMP 1 preproprotein variant (329-337)
D	P	G	Y	P	K	M	I	A						50%	Interstitial collagenase 2 (342-350)
D	P	P	Y	P	R	S	I	A	Q	W				38%	Vitronectin (458-469)

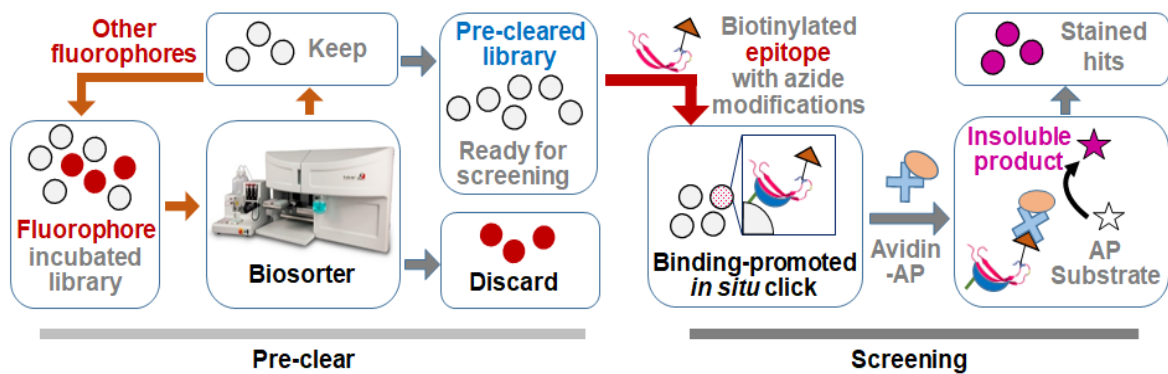
**Figure S2.** Sequence alignment and homology search result for the epitope (D<sub>570</sub>-A<sub>583</sub>) using protein BLAST® against the complete human proteome. The MMP2 epitope is shown in the first line (navy background). The top nine hits from the search with the highest sequence homology are shown in rows 2-10. Empty boxes represent completely mismatched residues.



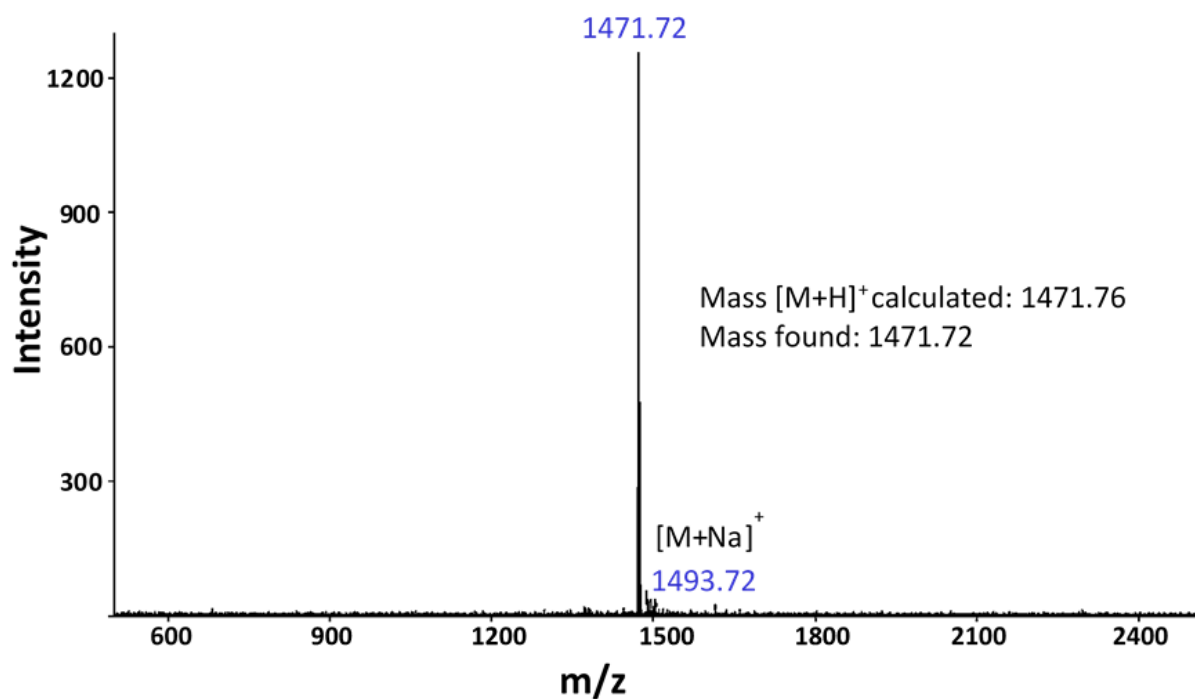
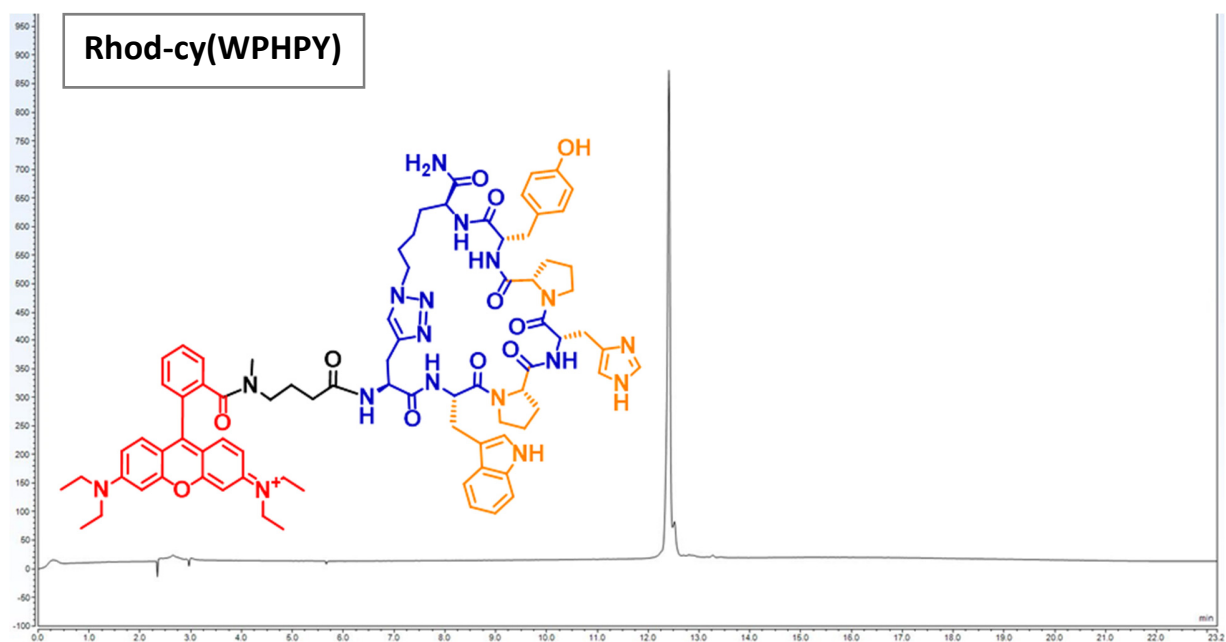
**Figure S3.** (Top) The generic structure of the one-bead-two-compound 80% Cyclic and 20% linear peptide library.  $X_1$ - $X_5$ : comprehensive 18 natural L-amino acids, excluding Cys and Met. The library was built using 90  $\mu\text{m}$ -sized TantaGel<sup>®</sup> S-NH<sub>2</sub> resins. (Bottom) After two steps of Edman degradation, the cyclic sequence contains a PTH moiety, while the linear one does not. A mass difference of 368 can be observed on the corresponding MS spectrum.



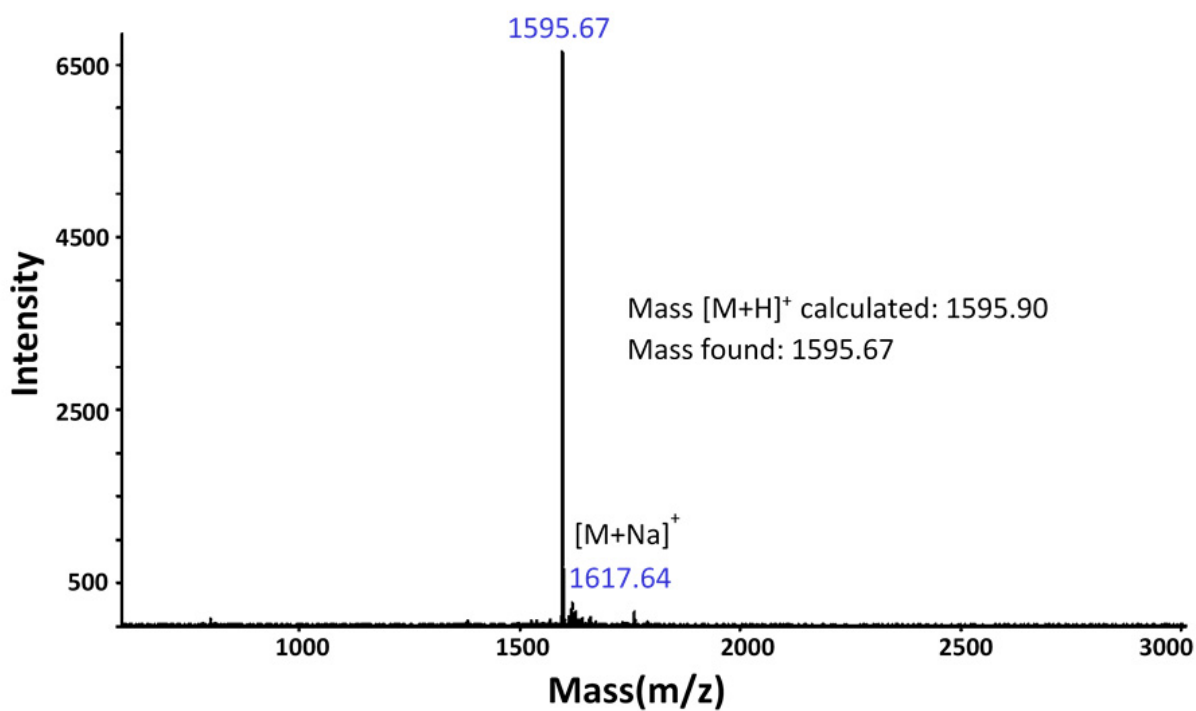
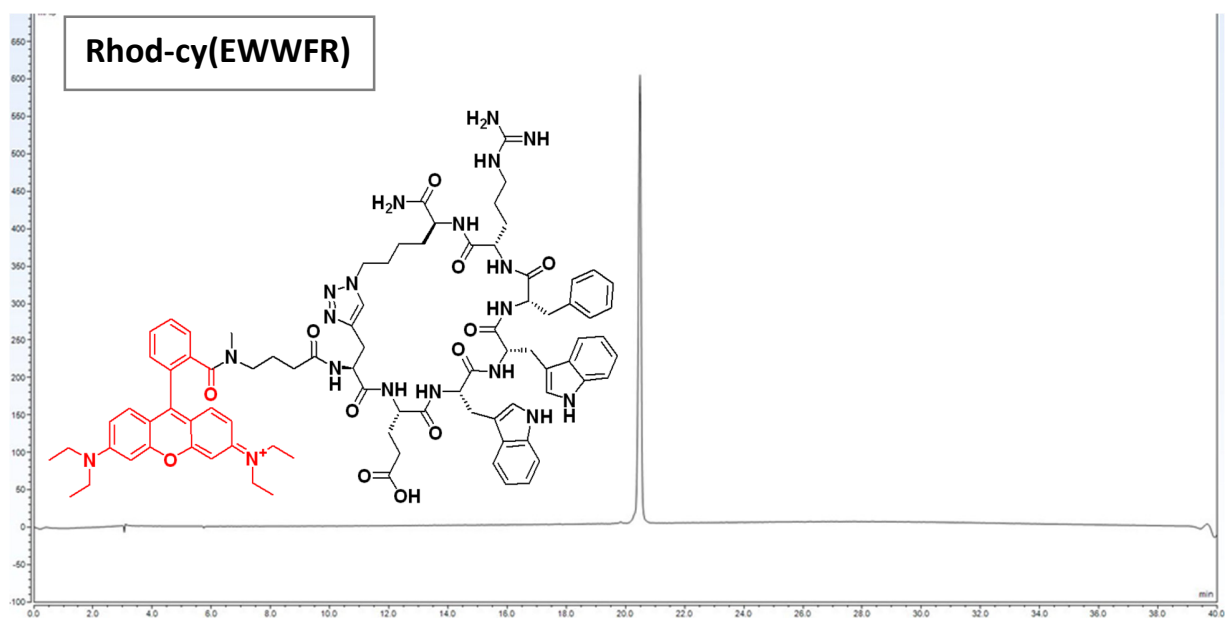
**Figure S4.** Structure and MALDI-TOF mass spectrum of the biotinylated epitope used in the library screening (Az4: L-azidolysine).



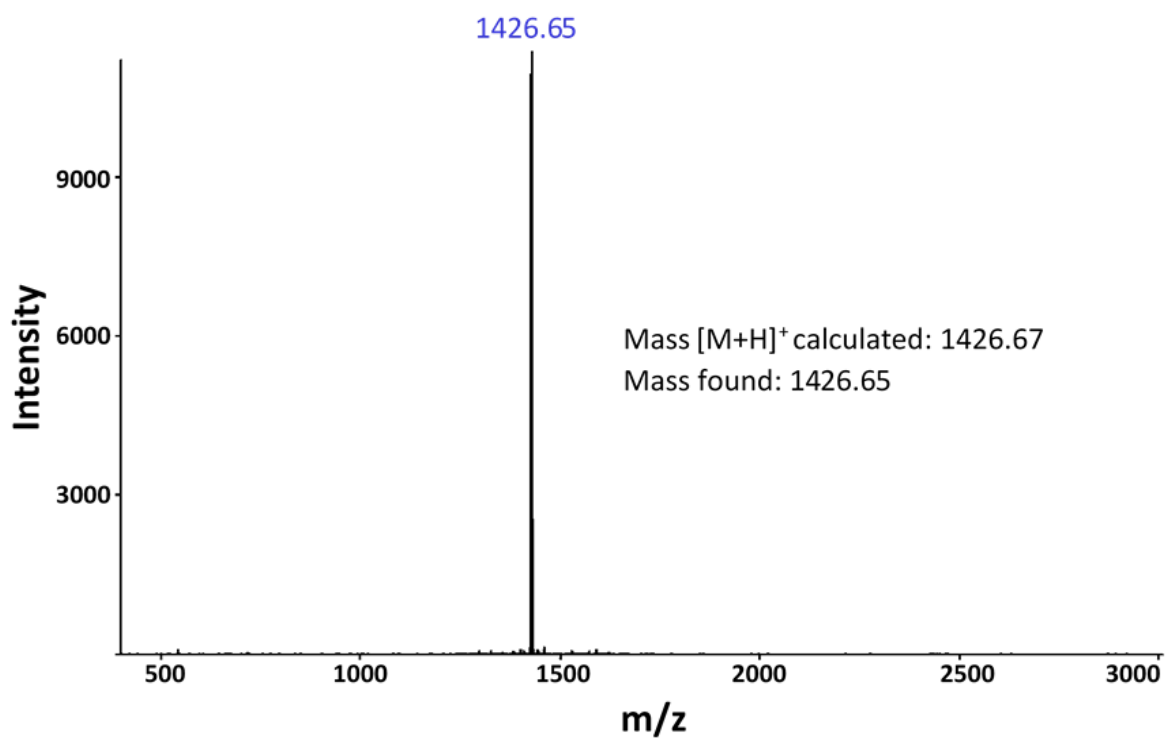
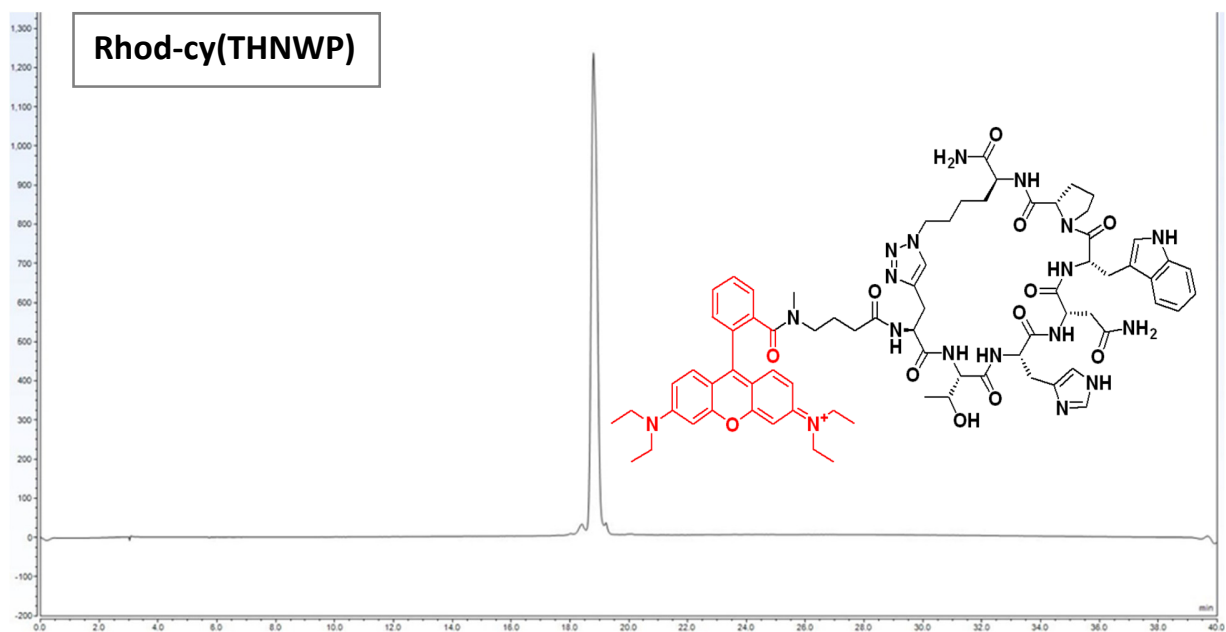
**Figure S5.** Schematic illustration of the automated preclear procedure and epitope targeting library screening strategy. The library was precleared using a fluorescence-activated BioSorter® and a panel of fluorophores and fluorescently labeled proteins. The precleared library was then screened against the biotinylated epitope. A binding between the epitope and the hit sequences promoted an *in situ* click reaction, which anchored biotin tags on the resin. The hits were then identified through the biotin tag by using the avidin-alkaline phosphatase (AP) and its colorimetric substrate BCIP (5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt) /NBT (nitro-blue tetrazolium chloride). The stained hits were picked out and processed for sequencing.



**Figure S6.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(WPHPY).

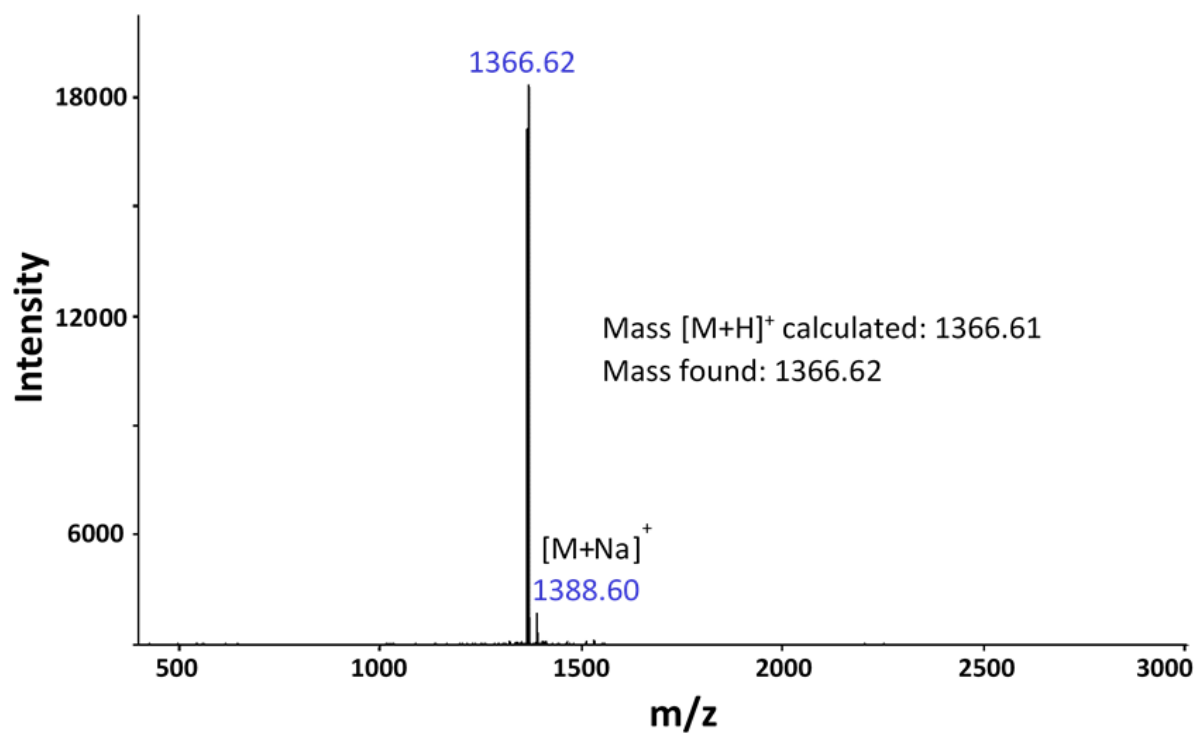
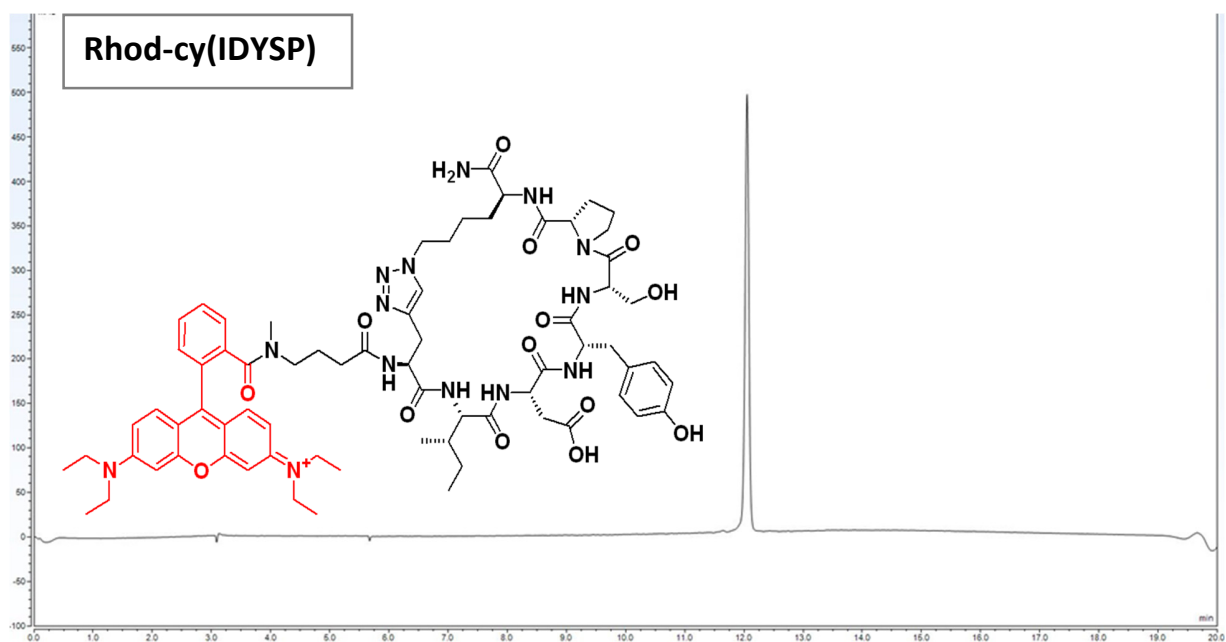


**Figure S7.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(EWWFR).

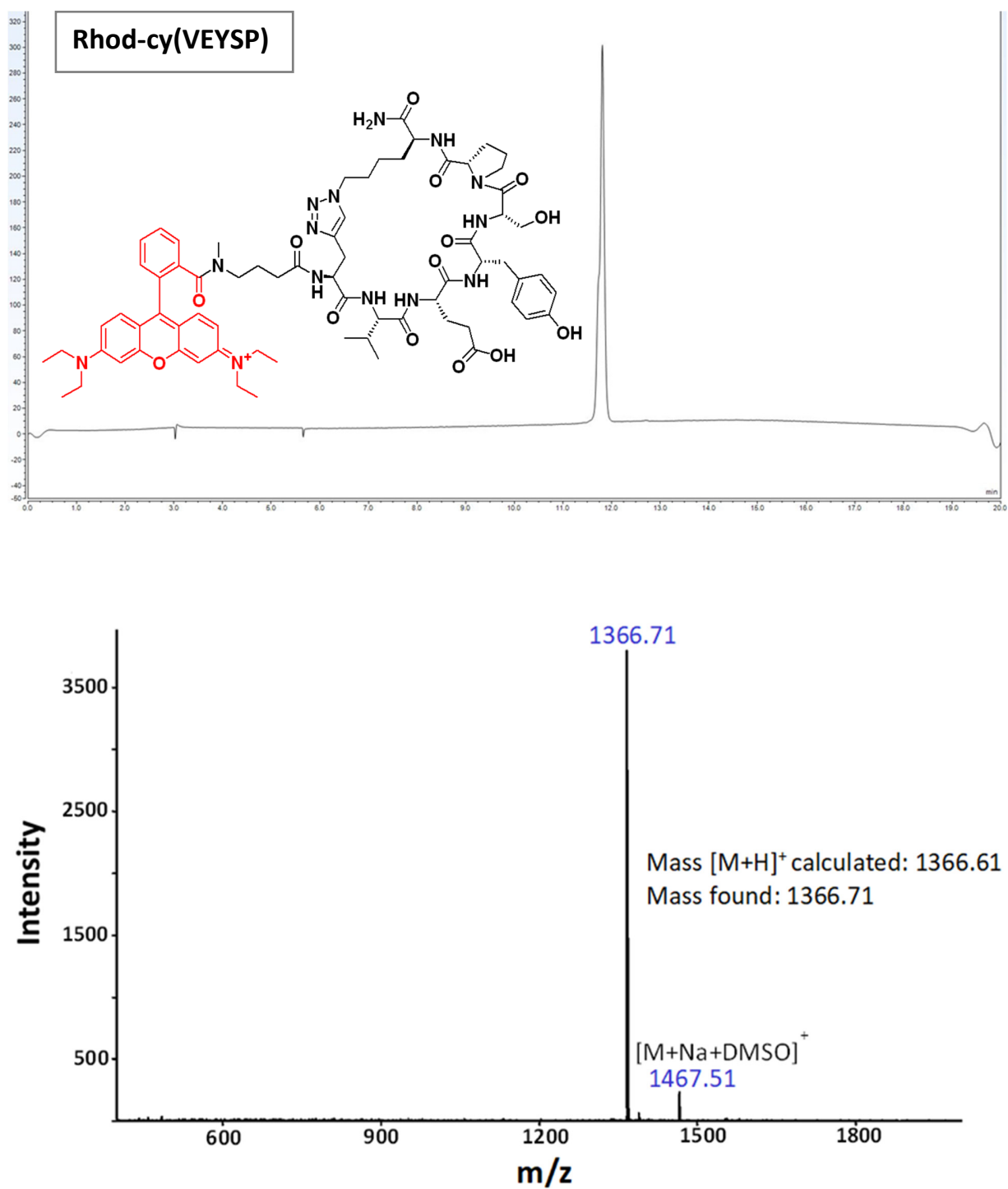


**Figure S8.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(THNWP).

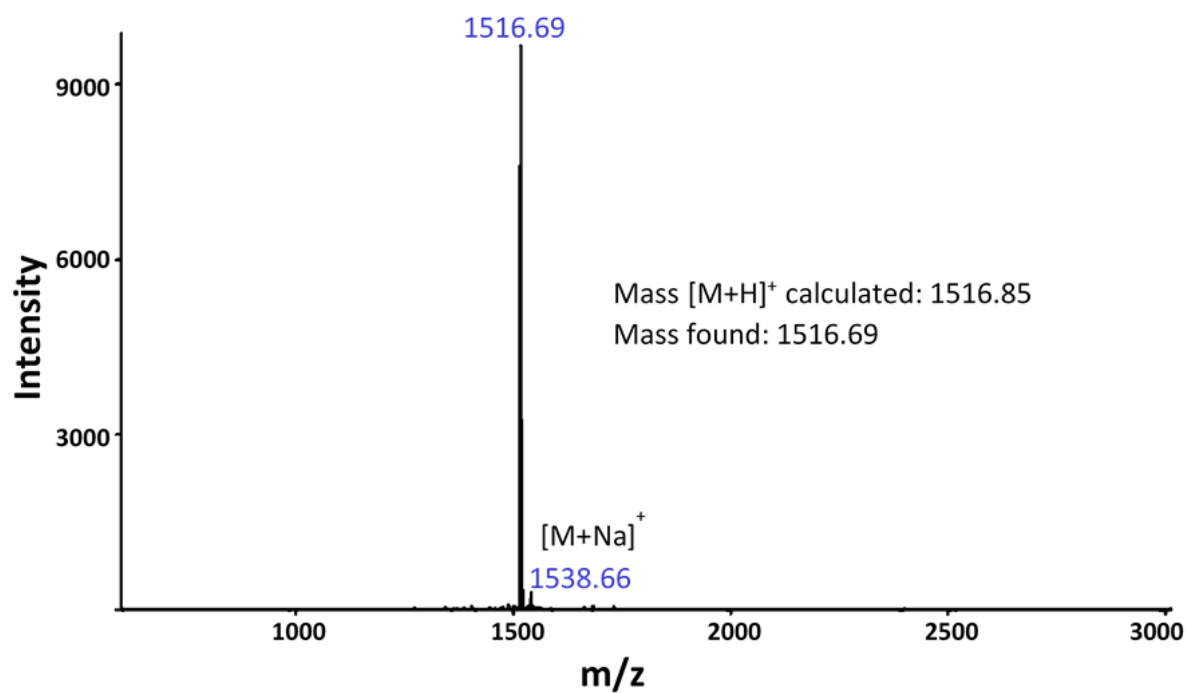
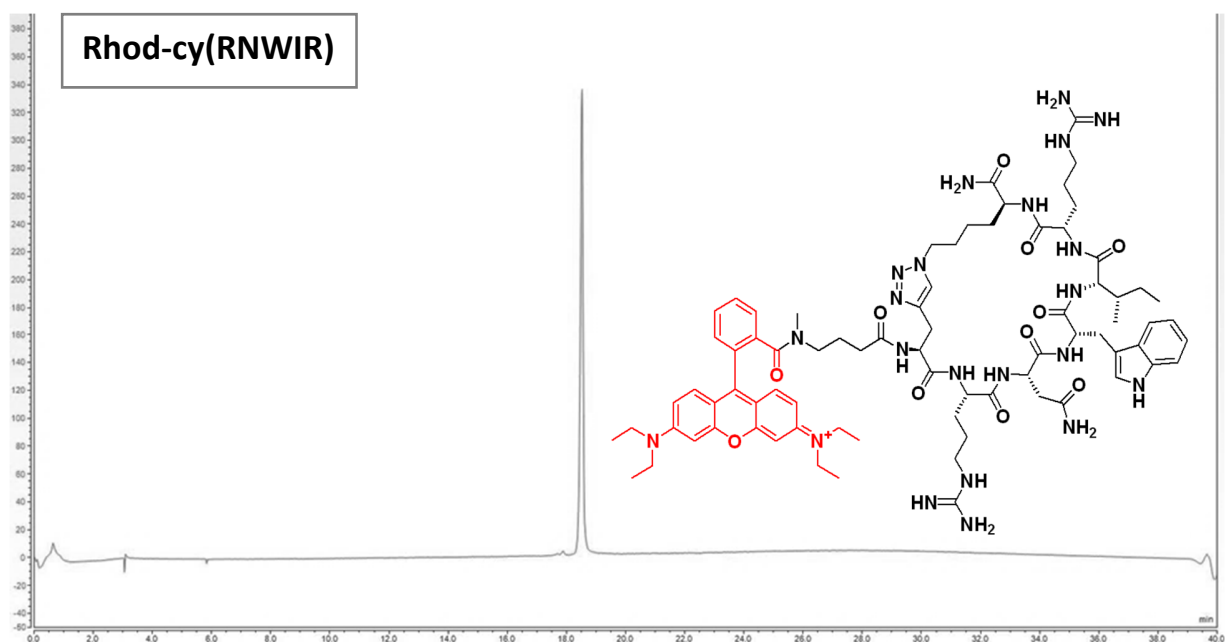




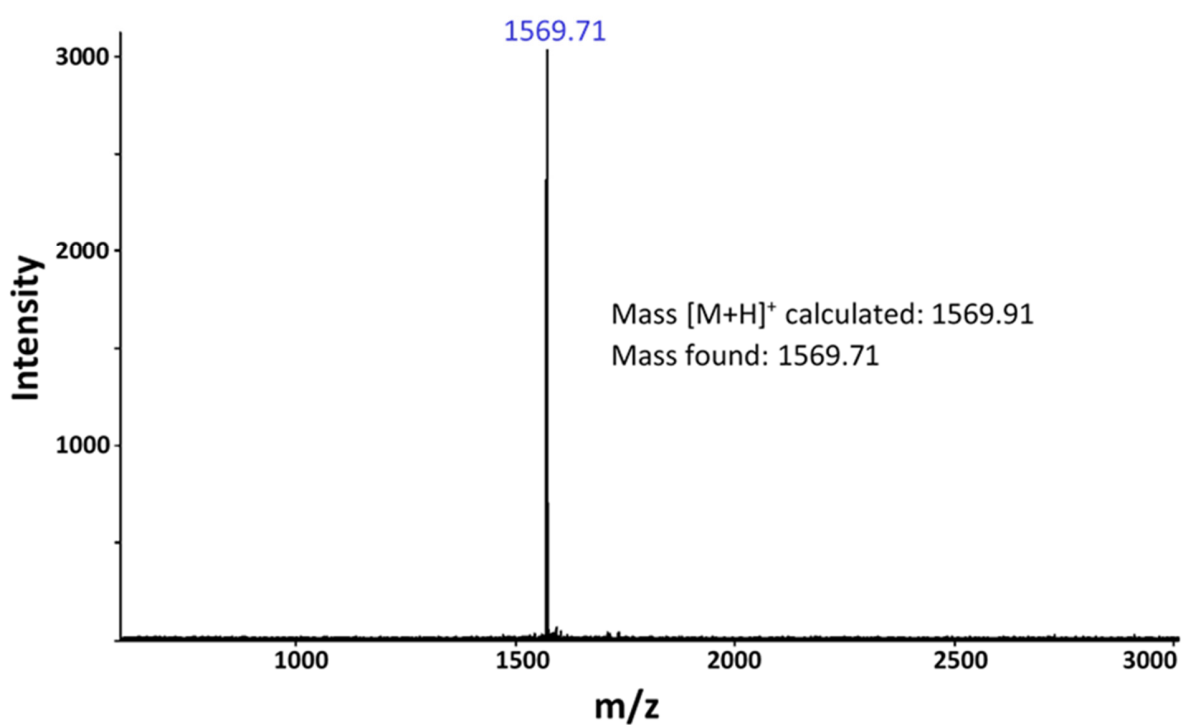
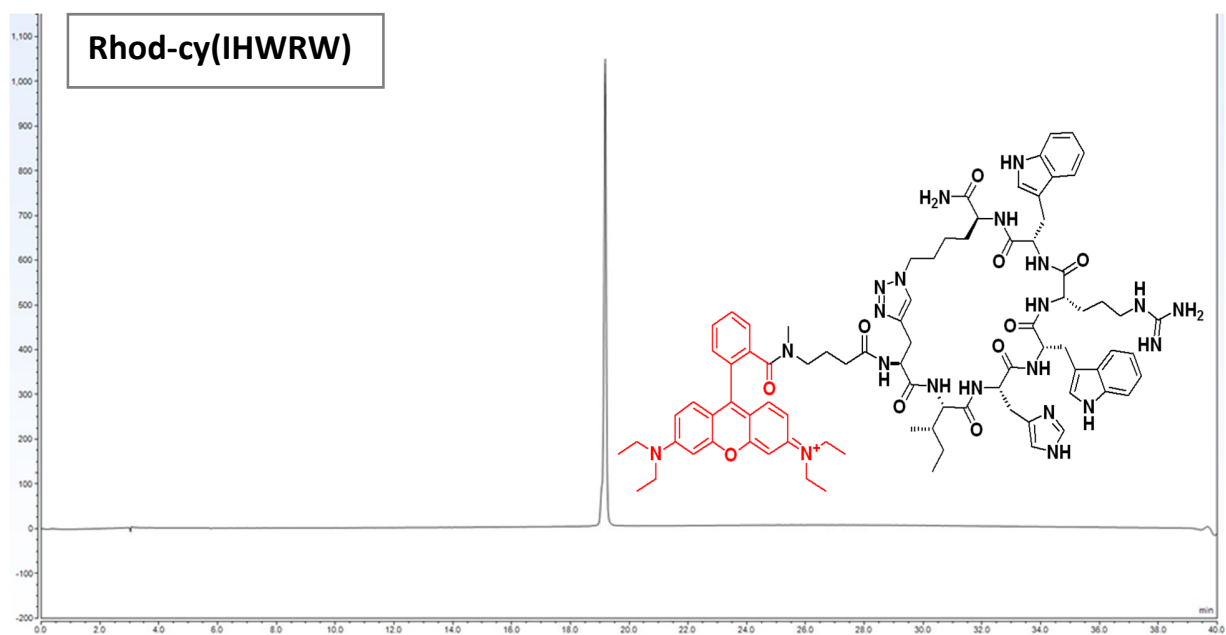
**Figure S9.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(IDYSP). The bump at the end of the chromatogram is due to sudden changes of the flow rate.



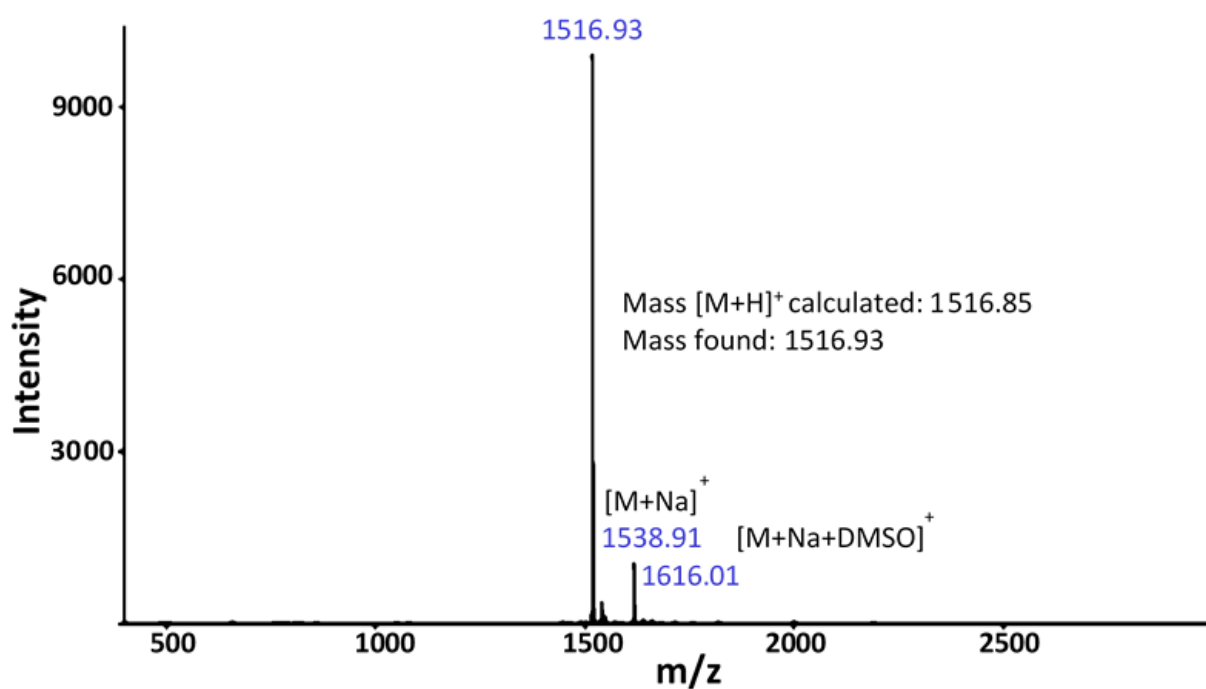
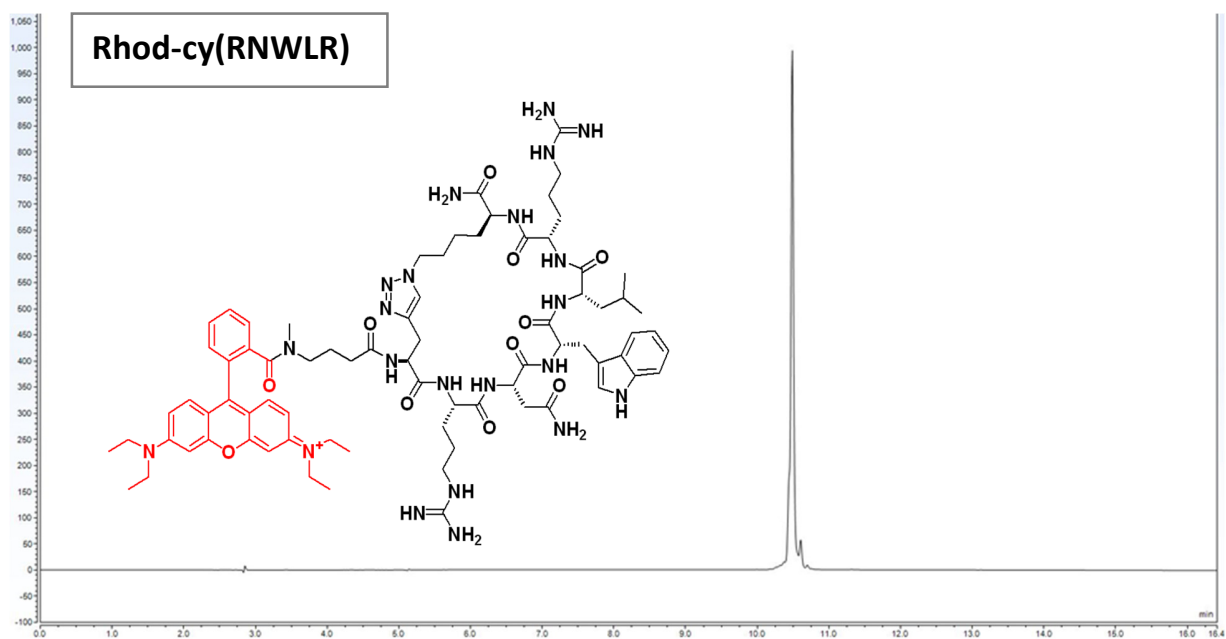
**Figure S10.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(VEYSP). The bump at the end of the chromatogram is due to sudden changes of the flow rate.



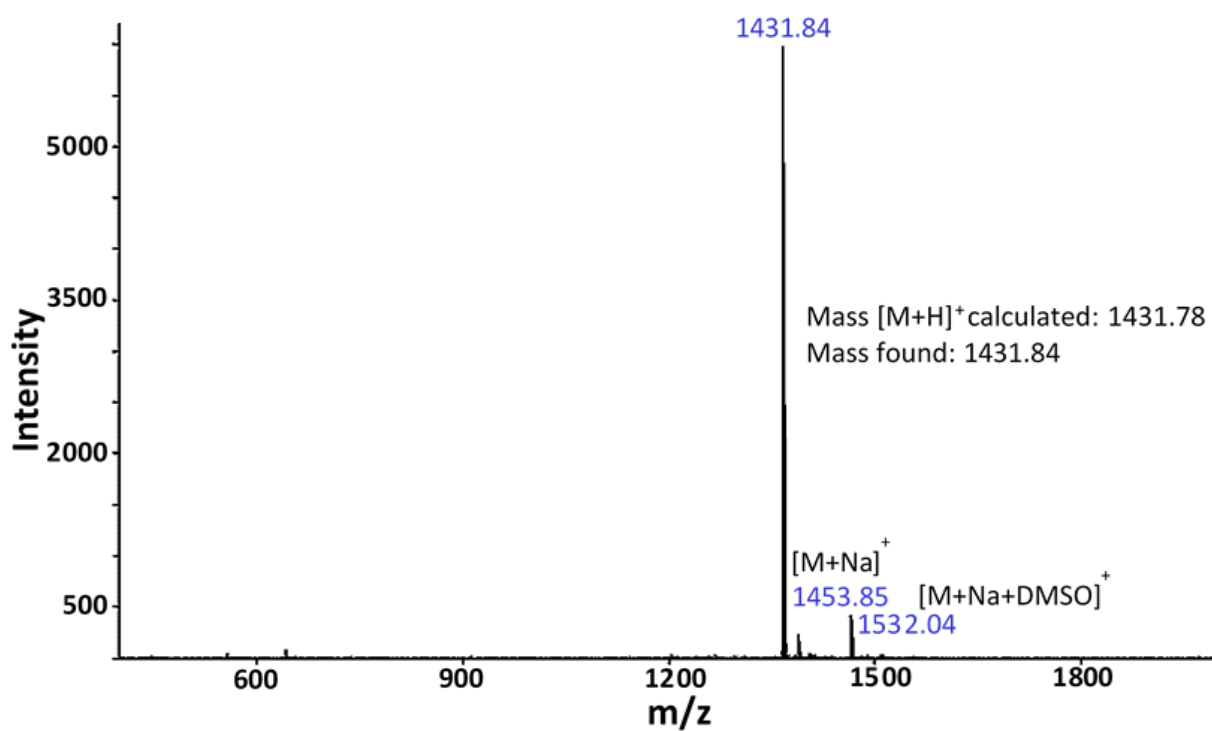
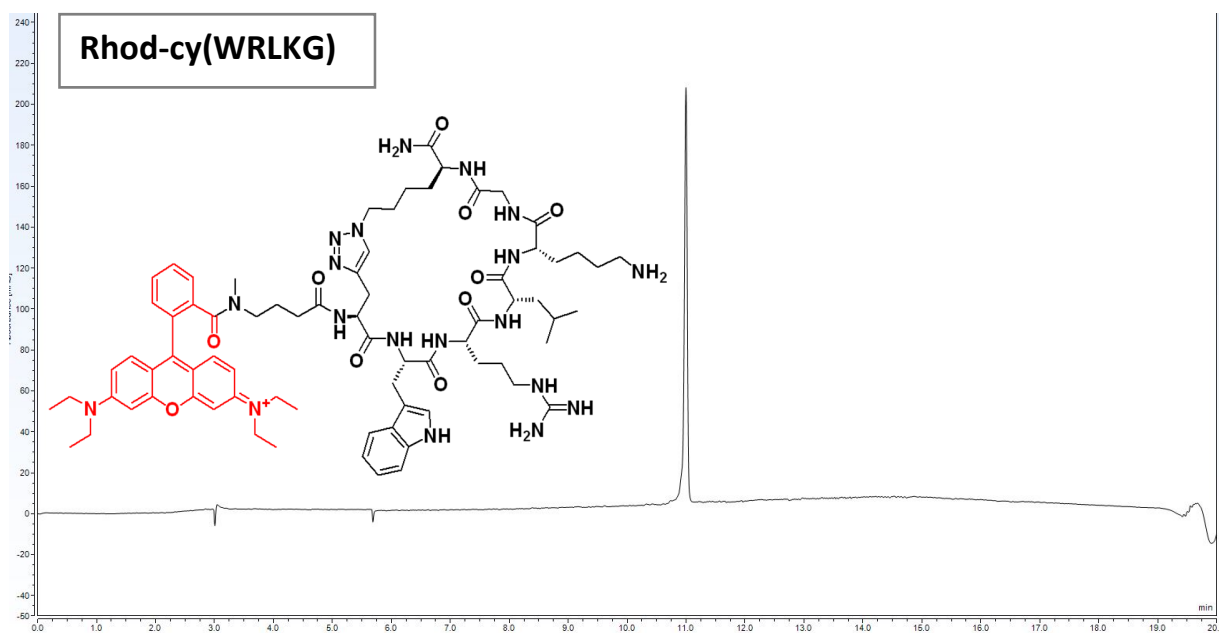
**Figure S11.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(RNWIR). The bump at the end of the chromatogram is due to sudden changes of the flow rate.



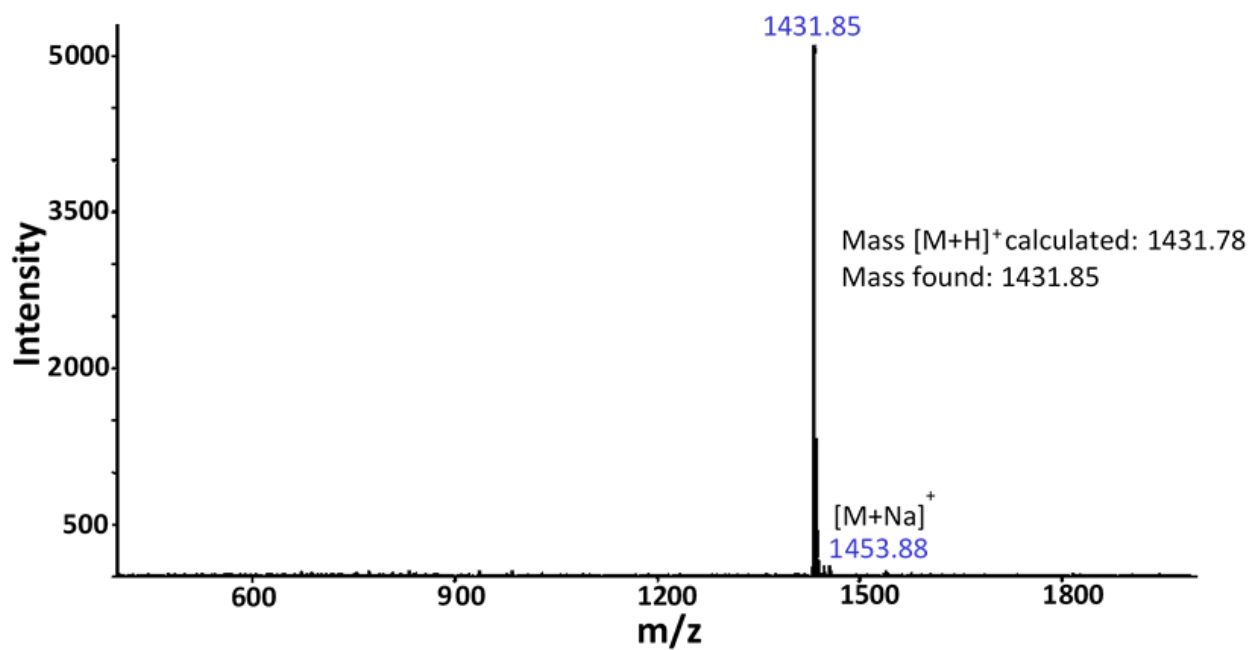
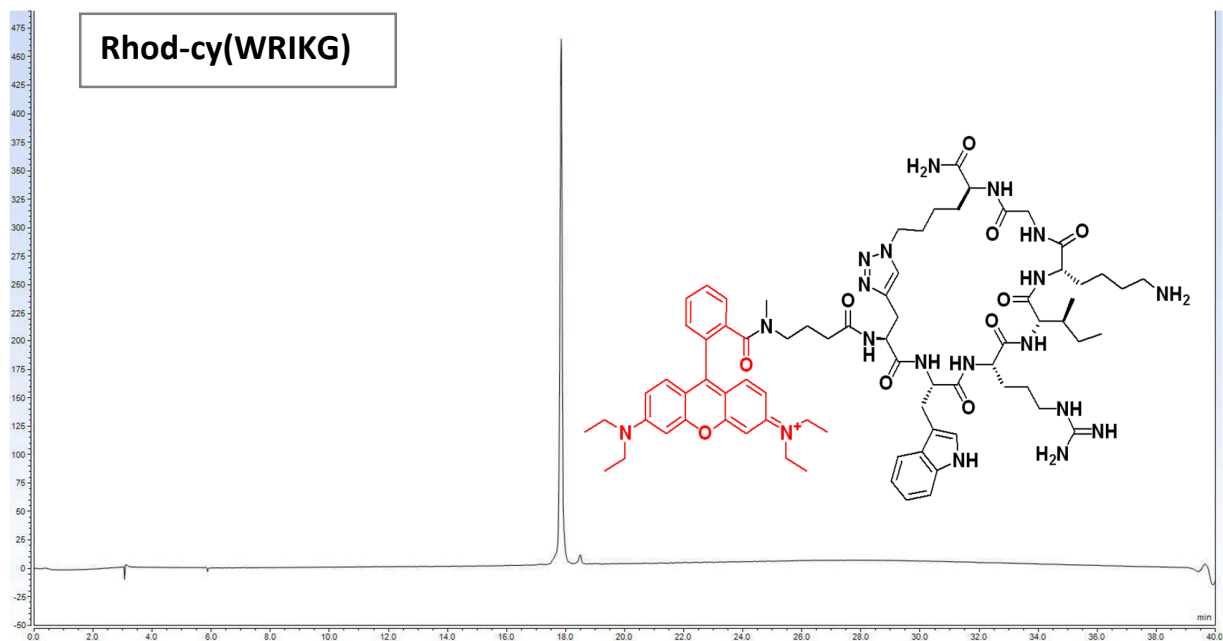
**Figure S12.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(IHWRW). The bump at the end of the chromatogram is due to sudden changes of the flow rate.



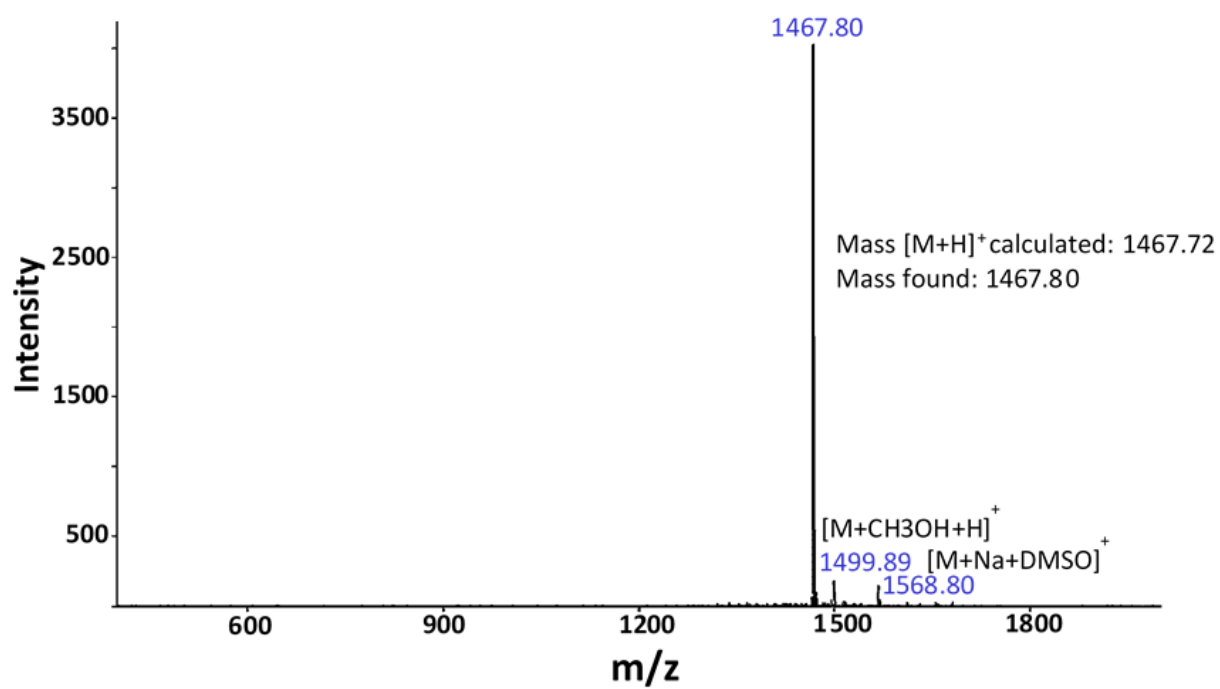
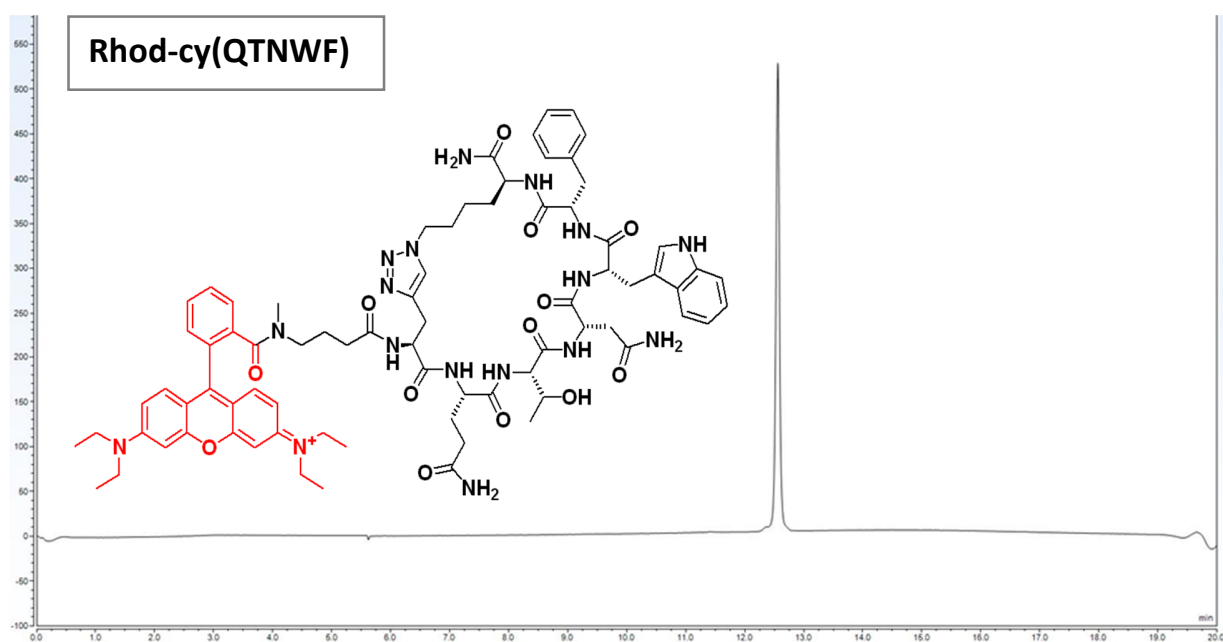
**Figure S13.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum and for Rhod-cy(RNWLR).



**Figure S14.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(WRLKG). The bump at the end of the chromatogram is due to sudden changes of the flow rate.

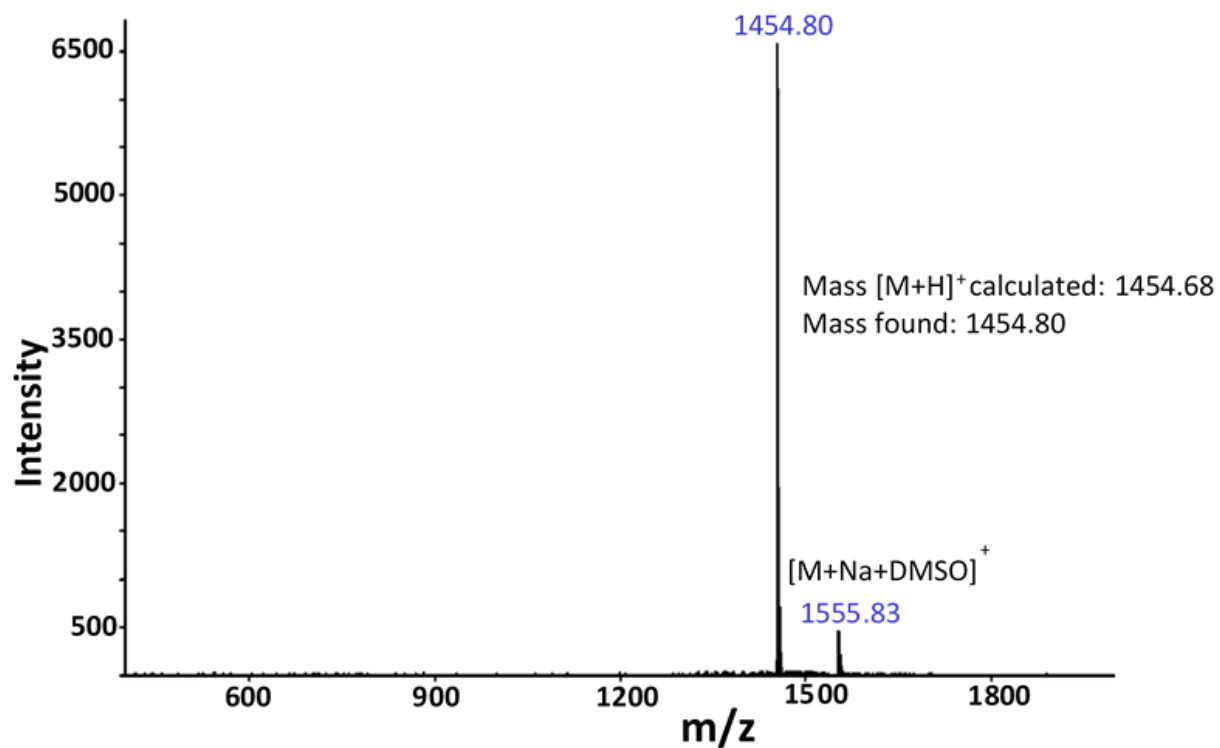
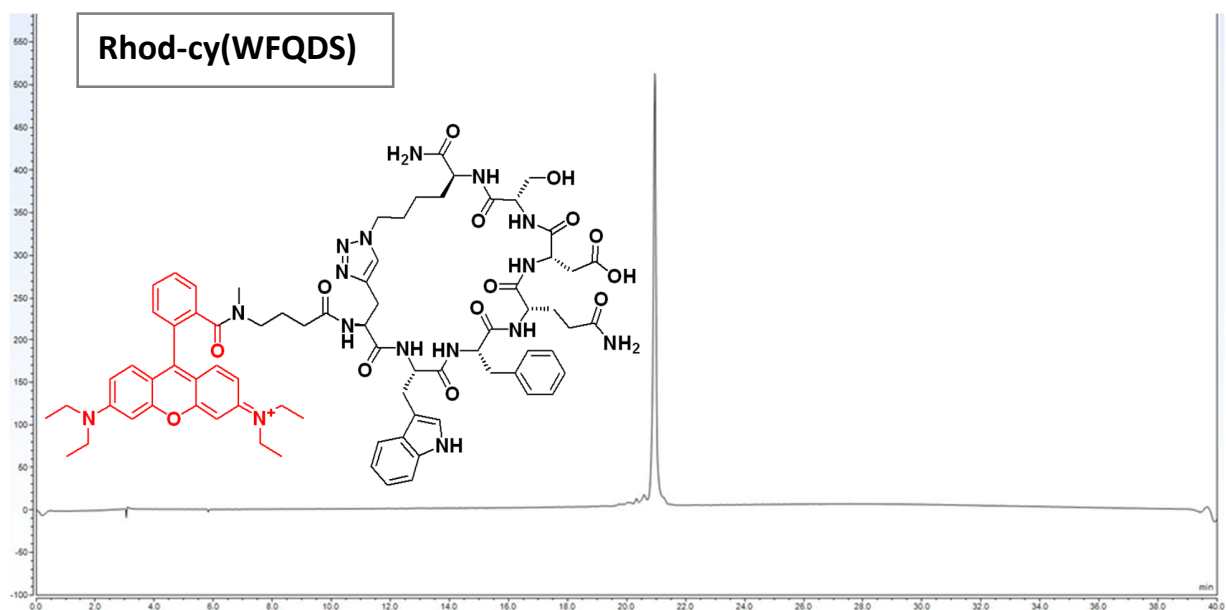


**Figure S15.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(WRIKG).

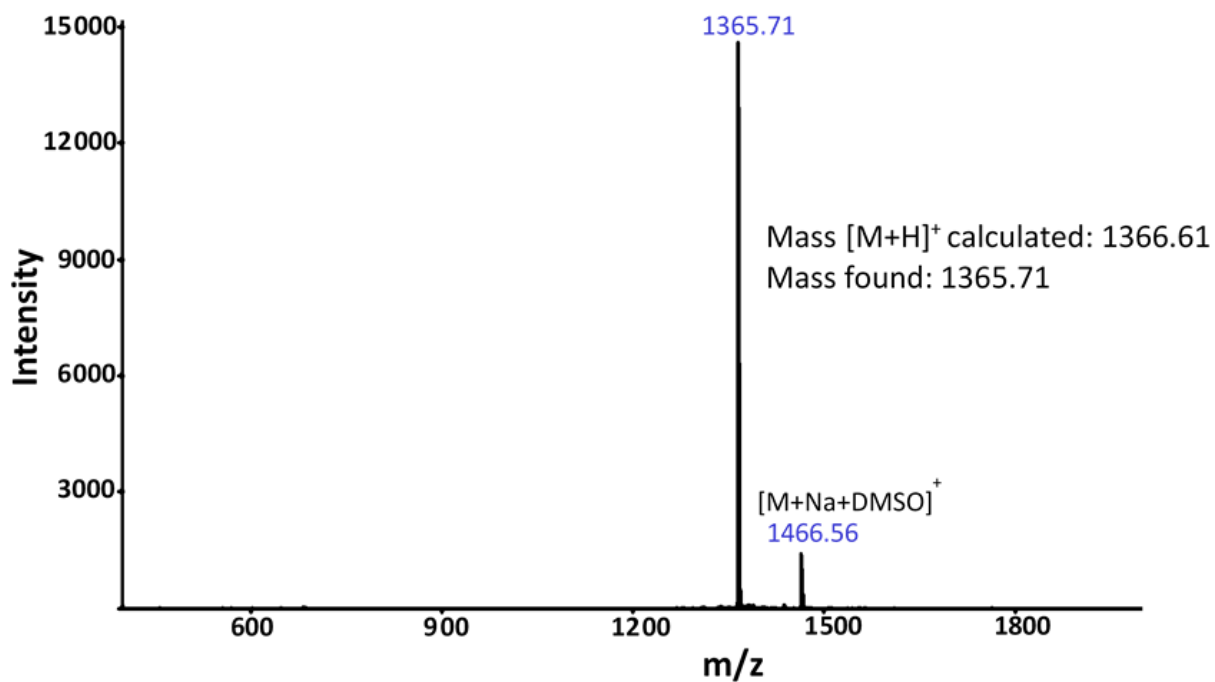
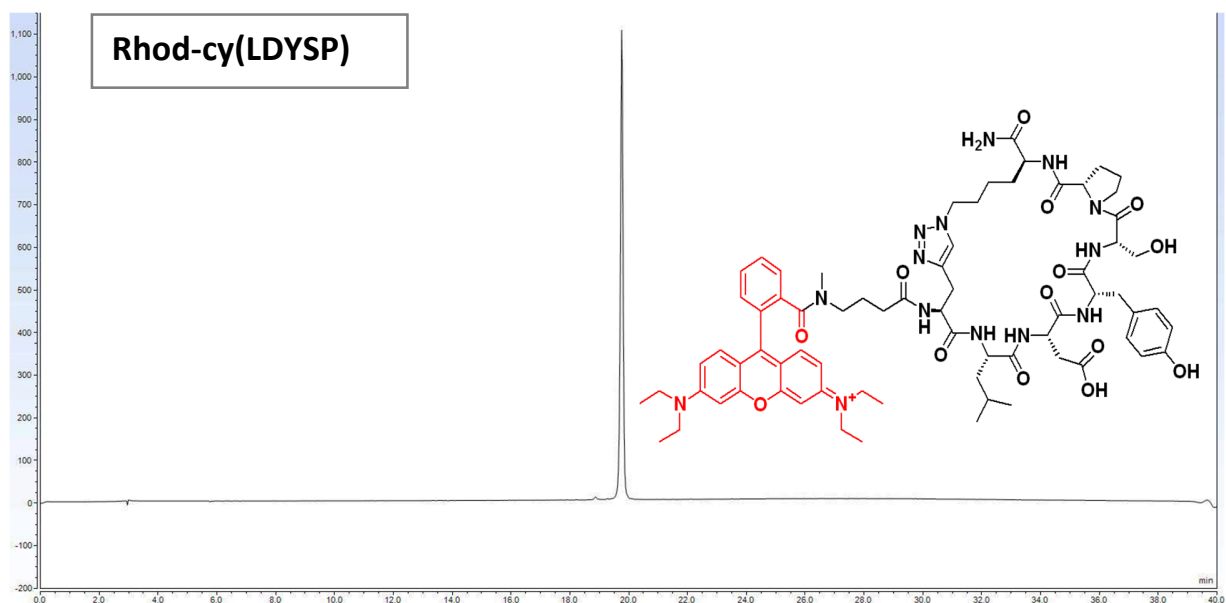


**Figure S16.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(QTNWF).

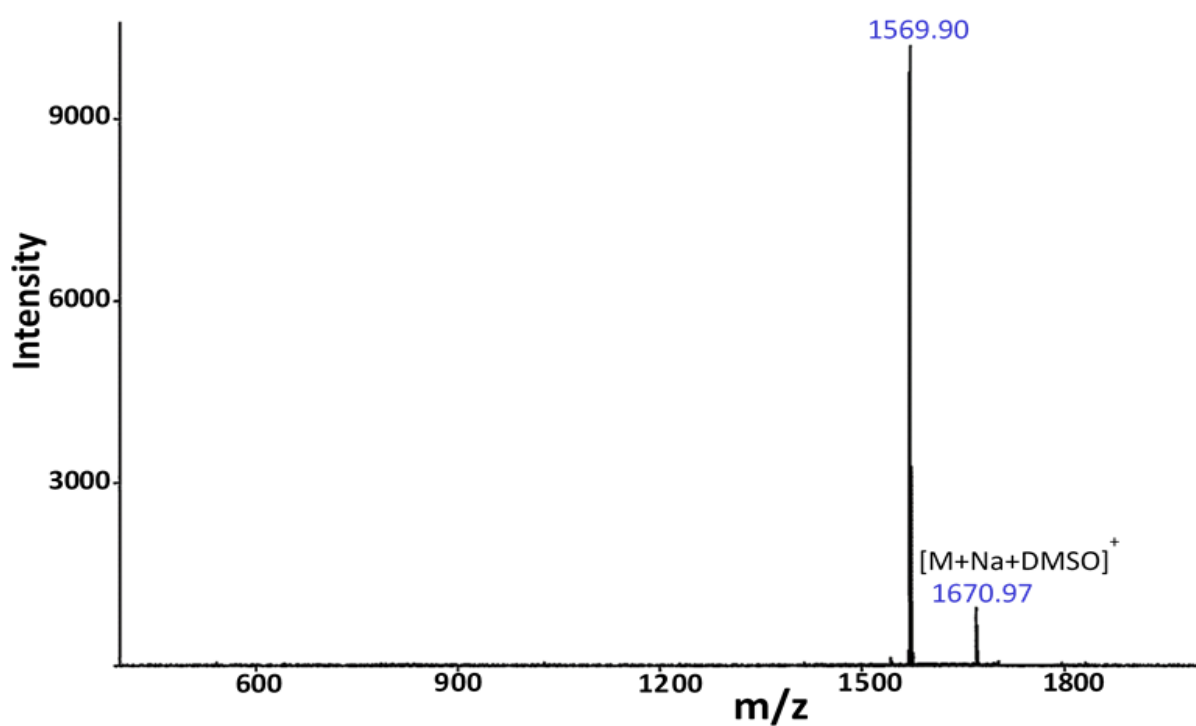
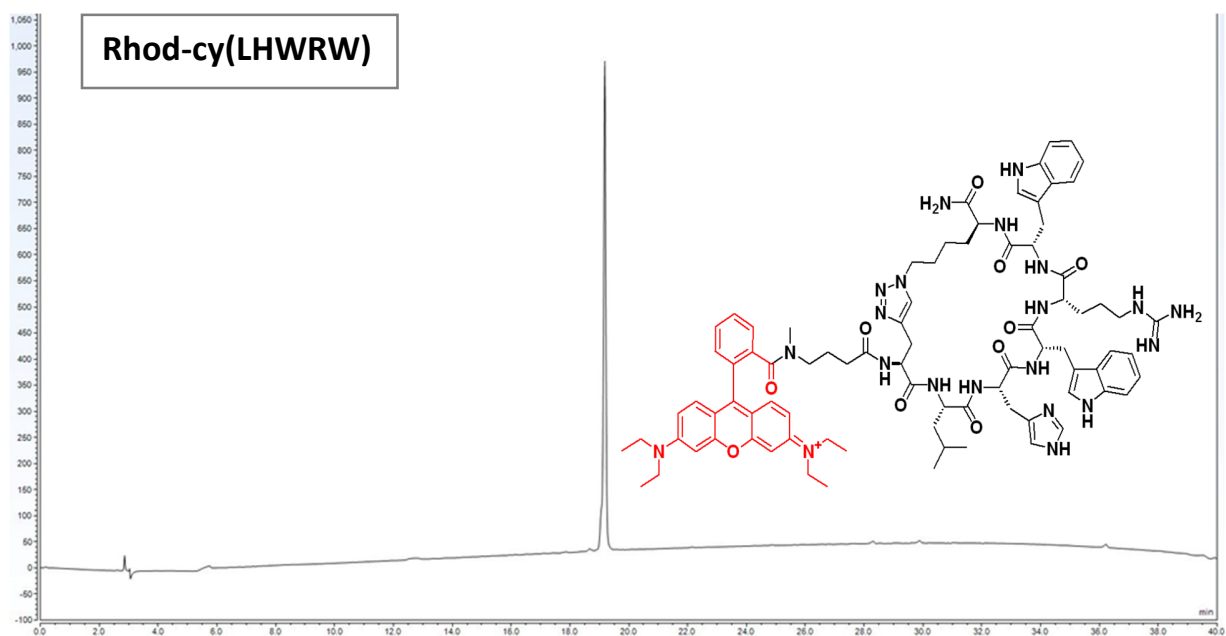




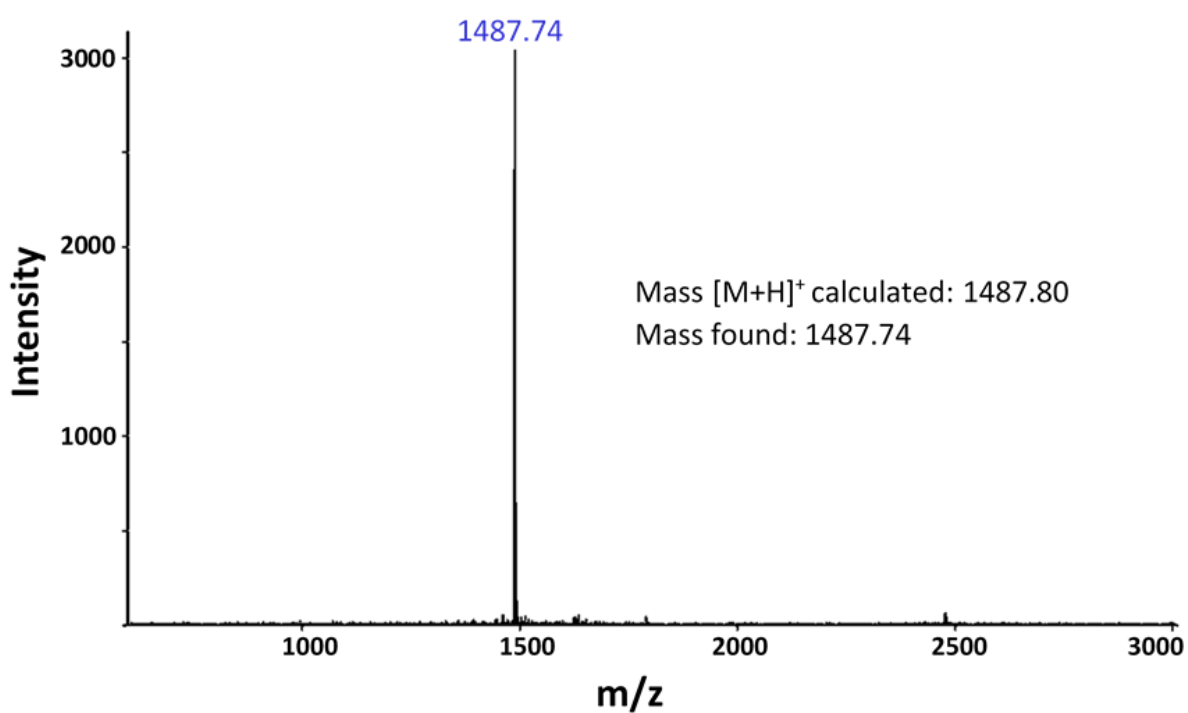
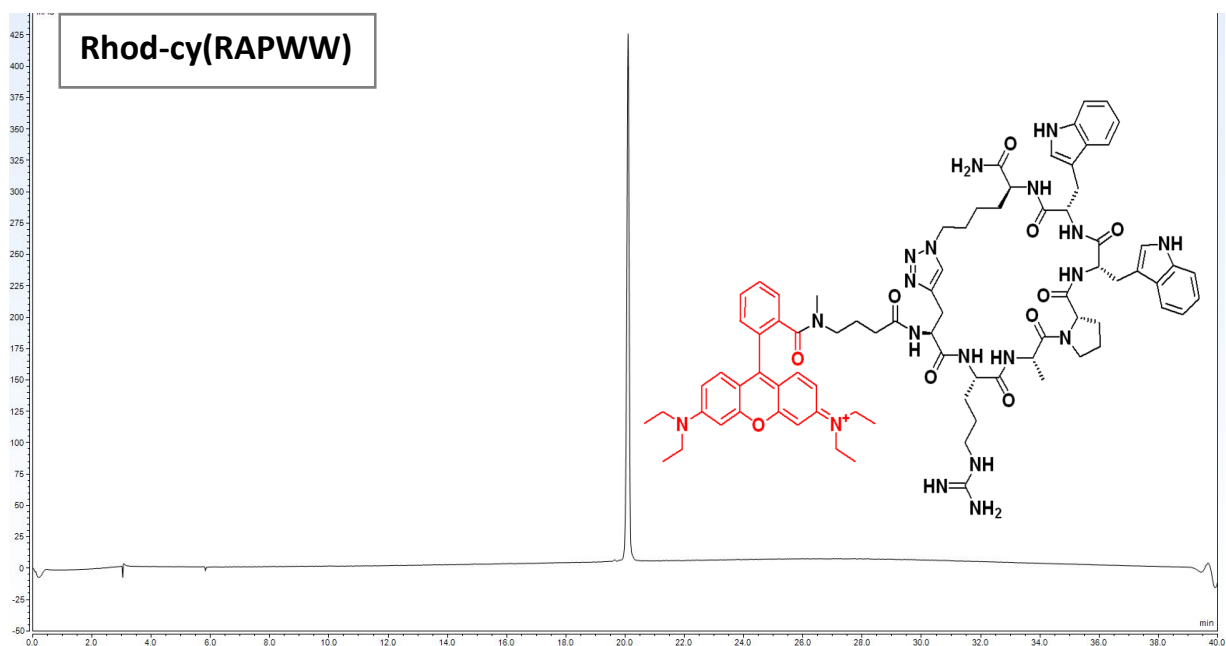
**Figure S17.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(WFQDS).



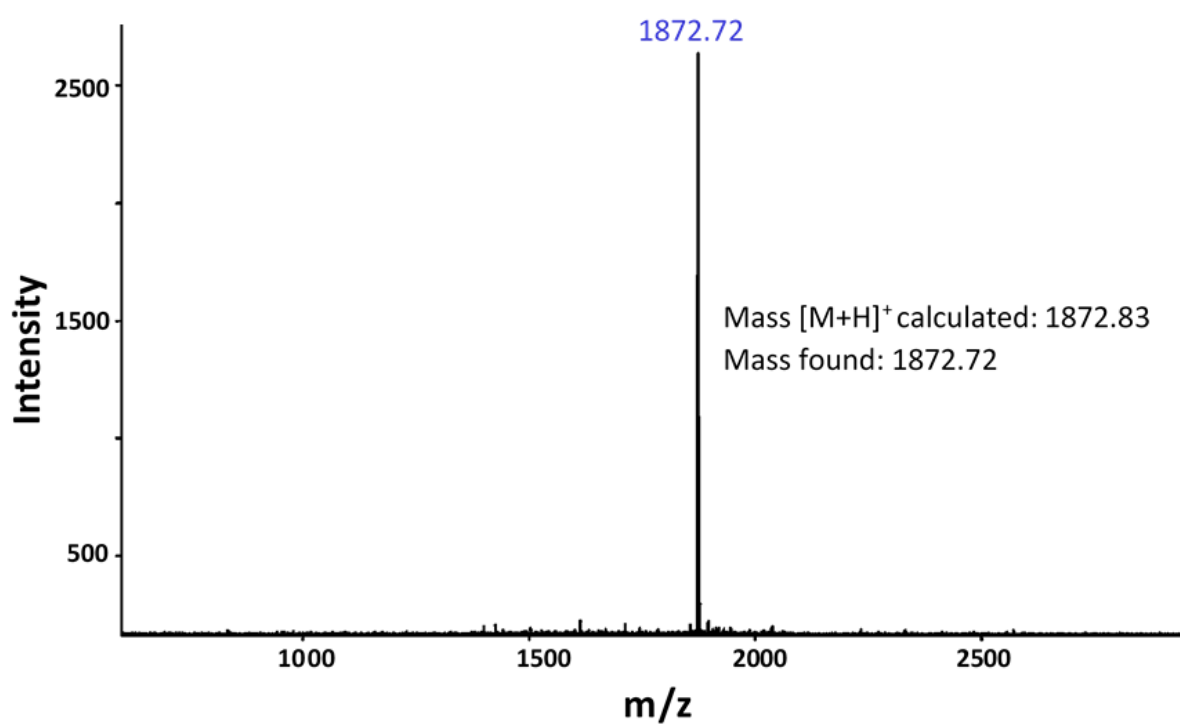
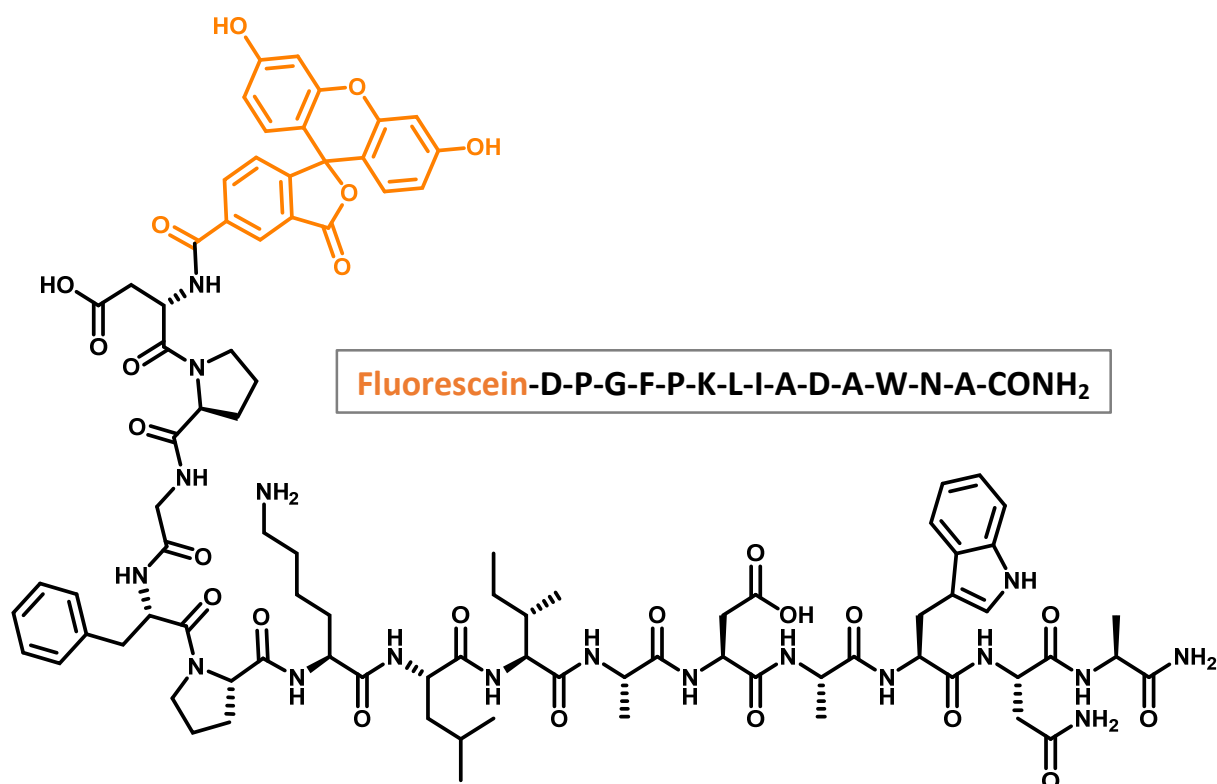
**Figure S18.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum and for Rhod-cy(LDYSP).



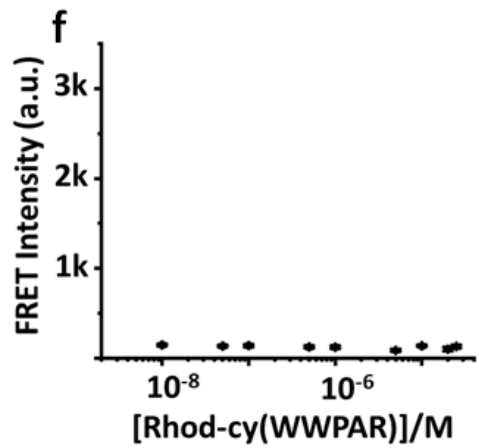
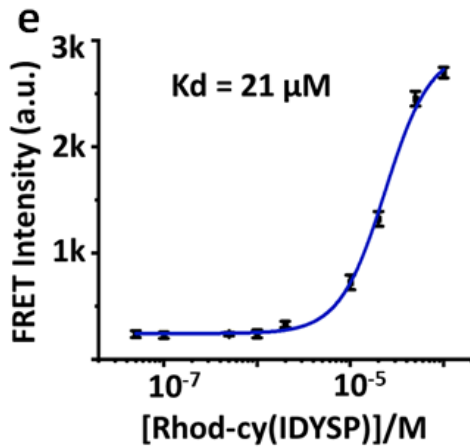
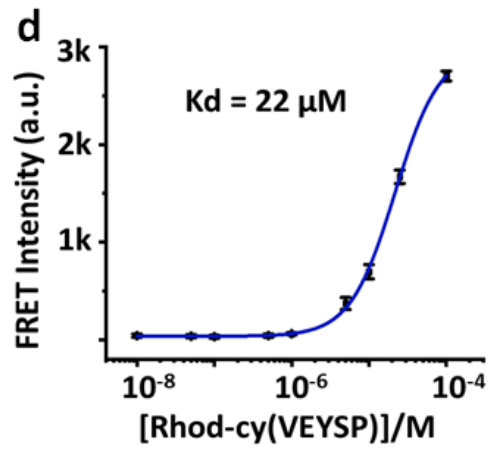
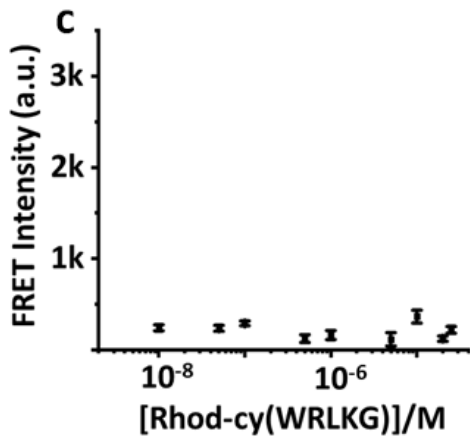
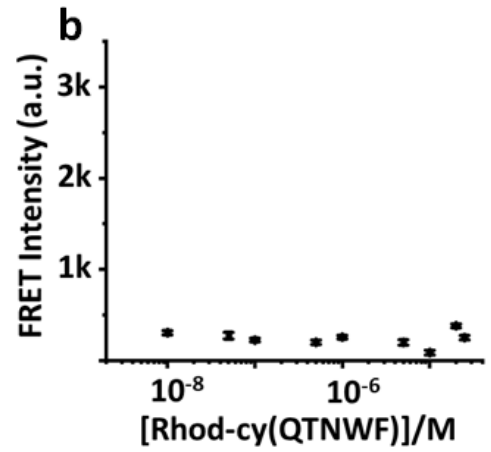
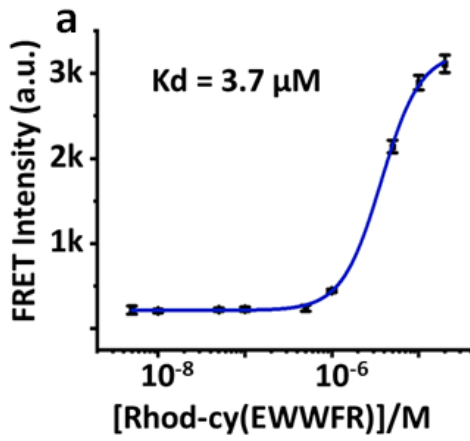
**Figure S19.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(LHWRW).

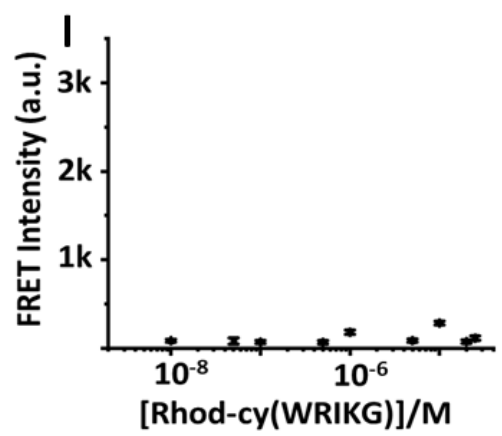
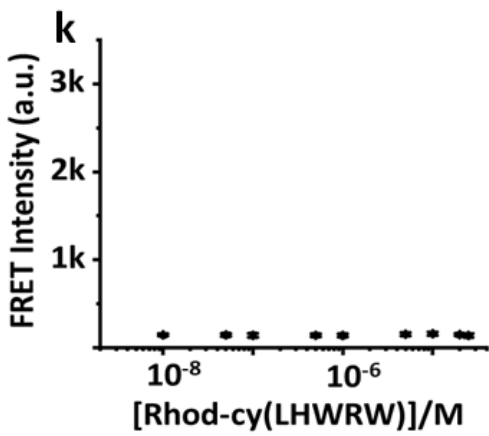
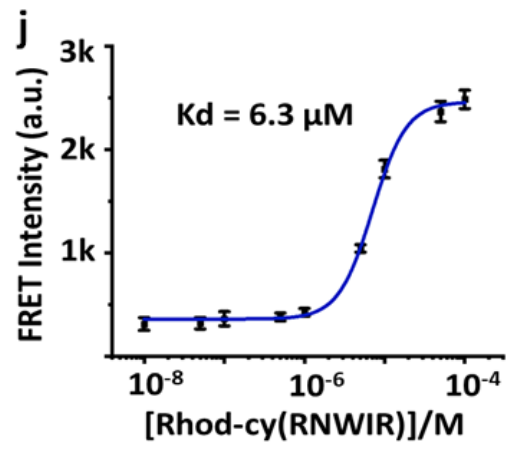
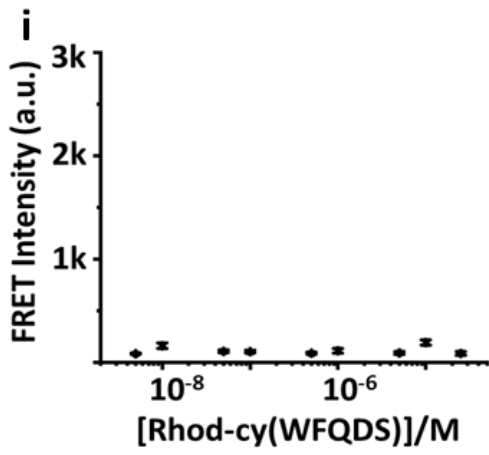
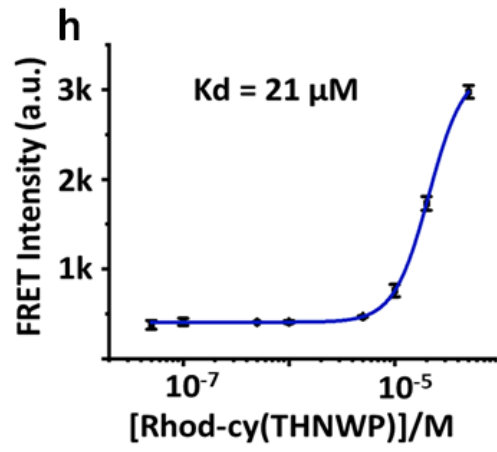
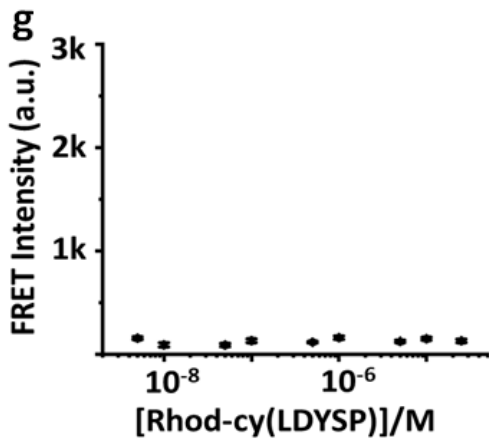


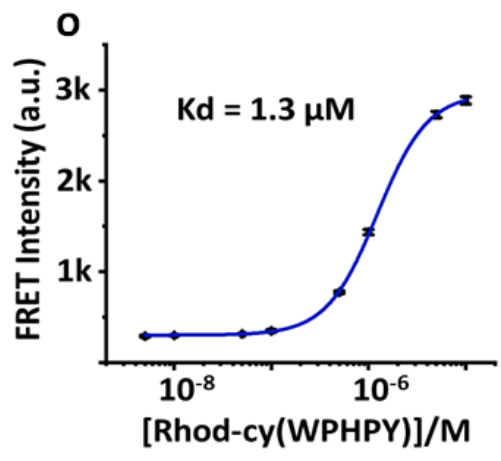
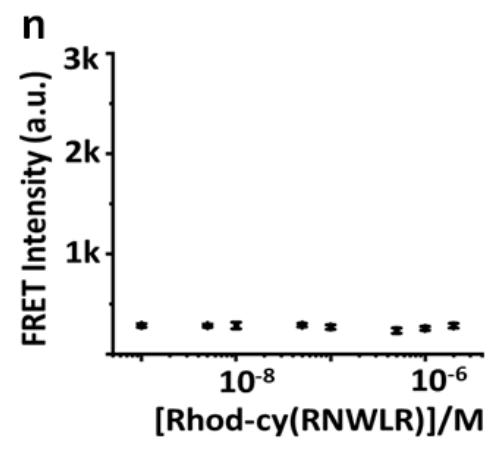
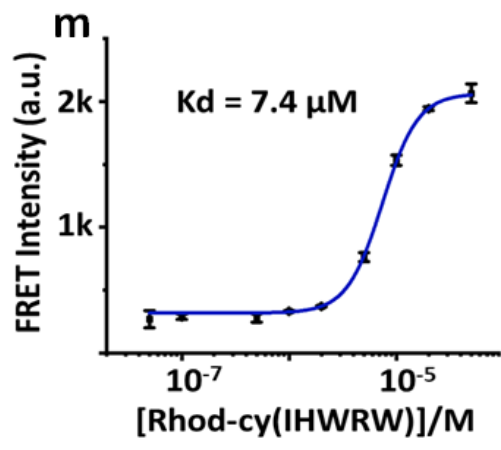
**Figure S20.** Analytical HPLC chromatogram (560 nm absorption) and MALDI-TOF mass spectrum for Rhod-cy(RAPWW).



**Figure S21.** MALDI-TOF mass spectrum for the fluorescein-labeled epitope.

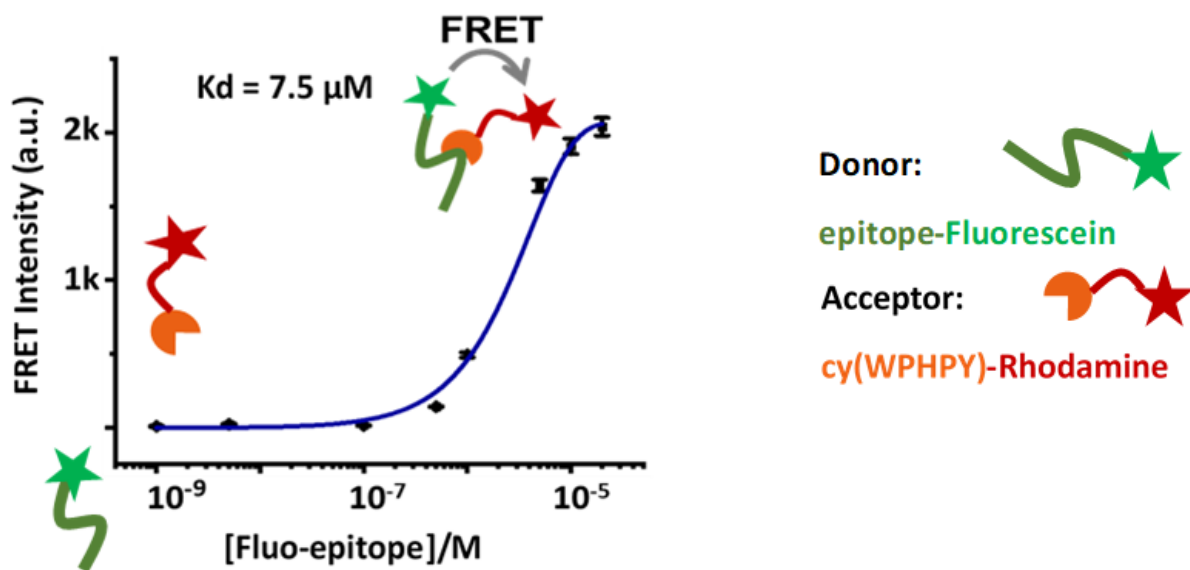




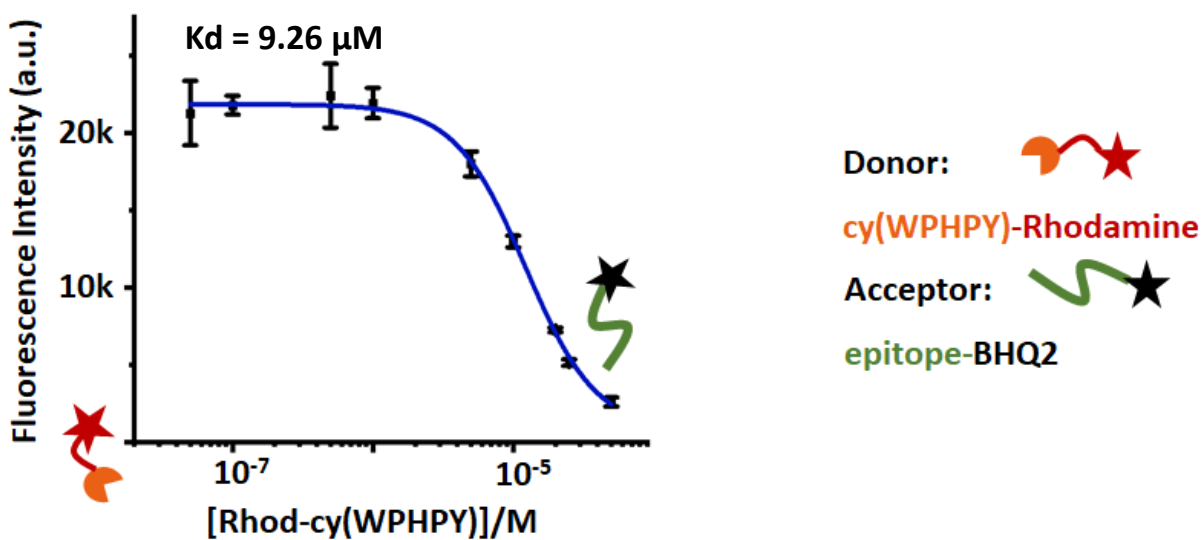


**Figure S22. a-o.** Varying concentrations of rhodamine B-conjugated hit peptides were mixed with 100 nM of the fluorescein-labeled epitope and the FRET intensities were measured. The curves were fitted using a Hill function to obtain the  $K_d$  values.

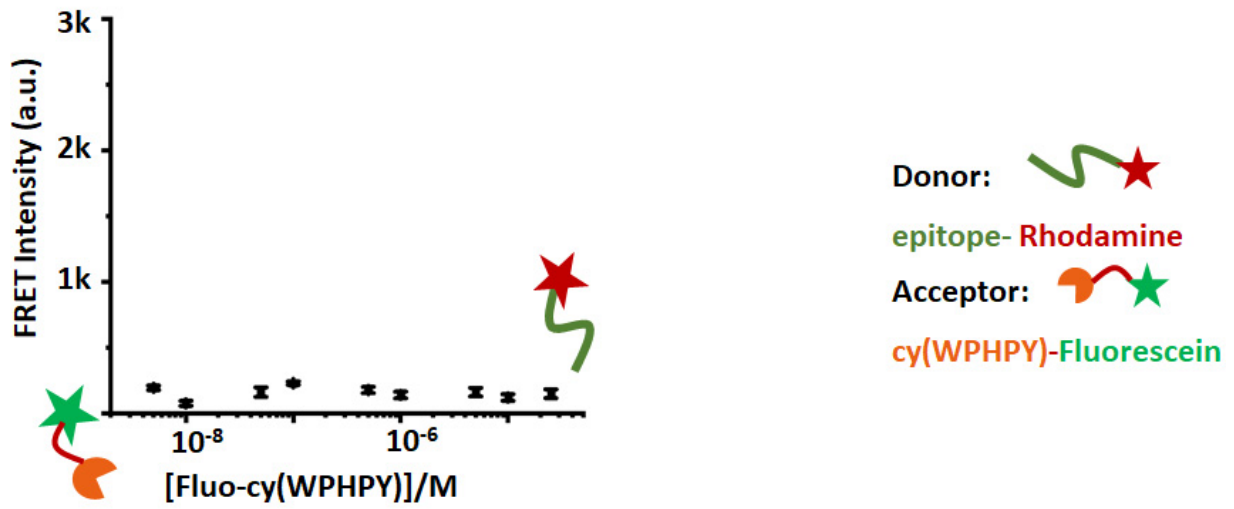




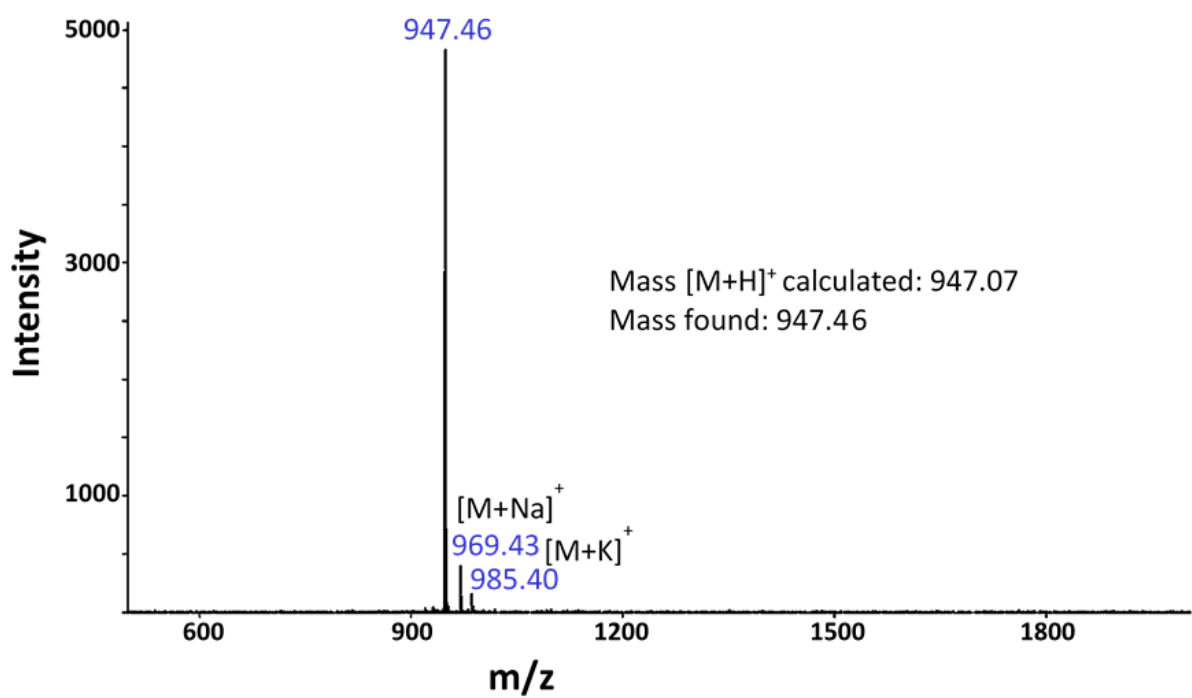
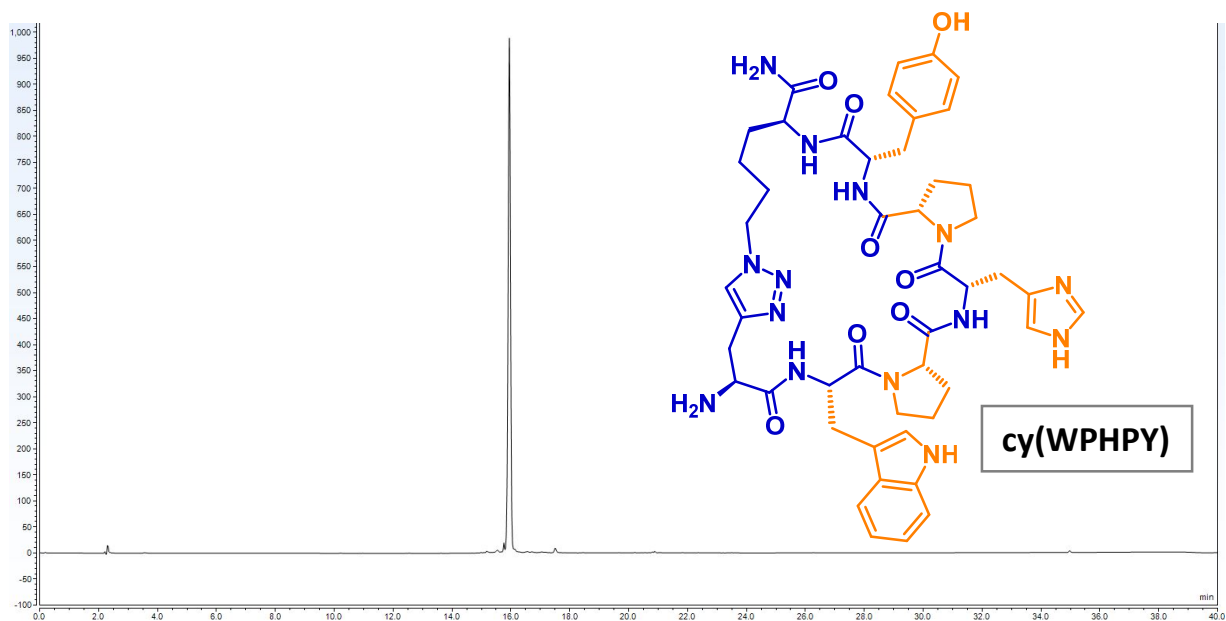
**Figure S23.** FRET signal obtained from the binding between 100 nM of rhodamine B-labeled ligand (Rhod-cy(WPHPY)) and varying concentrations of the fluorescein-labeled epitope (Fluo-Epitope). The binding affinity obtained was similar to the original one (Fig S22, o).



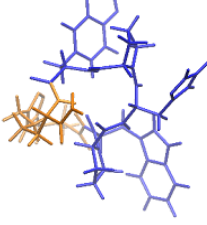
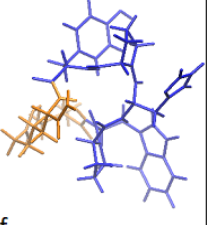
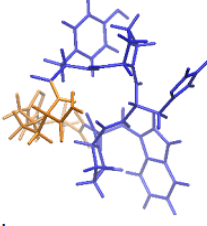
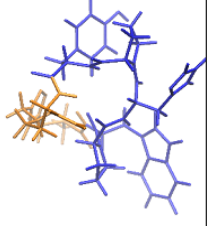
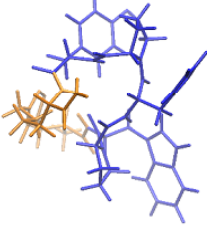
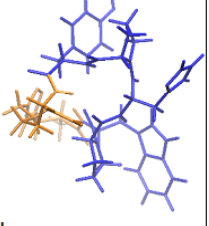
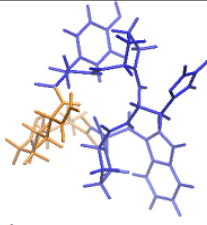
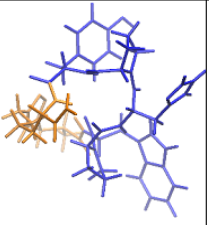
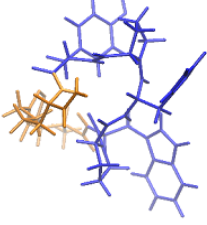
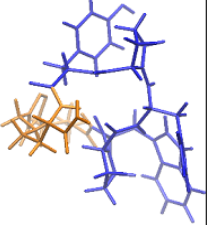
**Figure S24.** Residual fluorescence as a result of the binding between 100 nM of BHQ2-modified epitope and Rhod-cy(WPHPY). The dark quencher BHQ2 quenched the rhodamine fluorescence through a FRET mechanism.



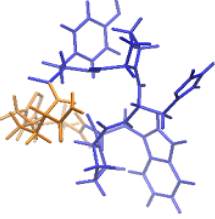
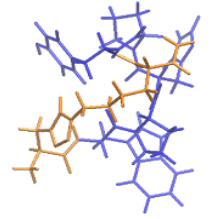
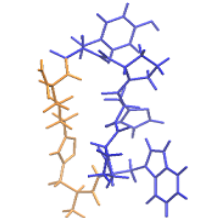
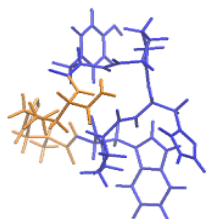
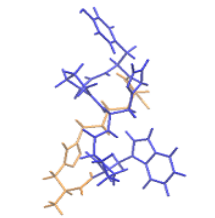
**Figure S25.** FRET signal obtained from the binding between 100 nM of Rhod-Epitope and varying concentrations of fluorescein-labeled ligand (Fluo-cy(WPHPY)). Changing the tag on the ligand resulted in a total loss of binding.



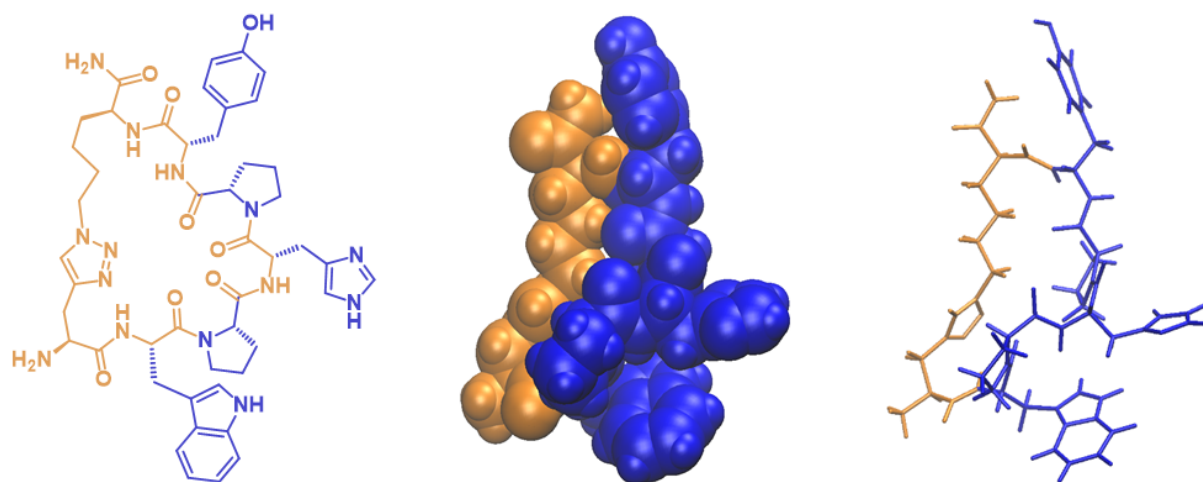
**Figure S26.** Analytical HPLC chromatogram (280 nm absorption) and MALDI-TOF mass spectrum and for the unmodified cyclic peptide ligand cy(WPHPY).

Conformation	Free Energy (kcal/mol)	Conformation	Free Energy (kcal/mol)
 a	226.37	 f	227.527
 b	226.38	 g	227.534
 c	227.401	 h	227.63
 d	227.404	 i	227.65
 e	227.41	 j	227.93

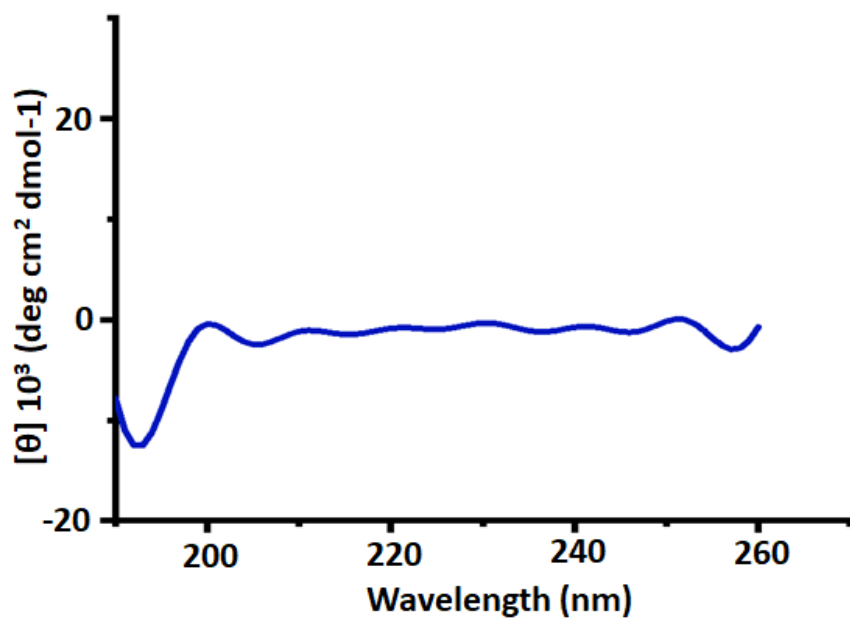
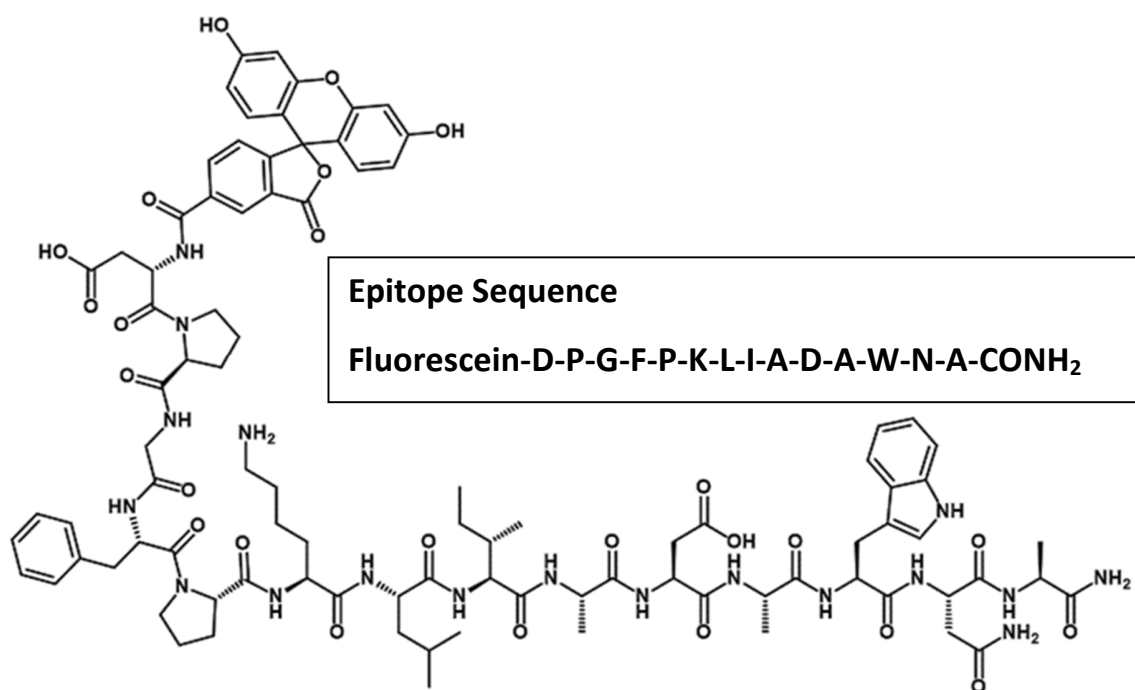
**Table S1.** The most stable conformations of cy(WPHPY) obtained from free energy calculations. The top 10 most stable conformations from a total of 546 conformations are shown in stick representation with orange and blue color representing two linkages. Conformational search and free energy calculations were performed with VM2 using the implicit solvation model.

Conformation	Free Energy (kcal/mol)
 1	226.37
 2	227.93
 3	228.08
 4	228.86
 5	229.54

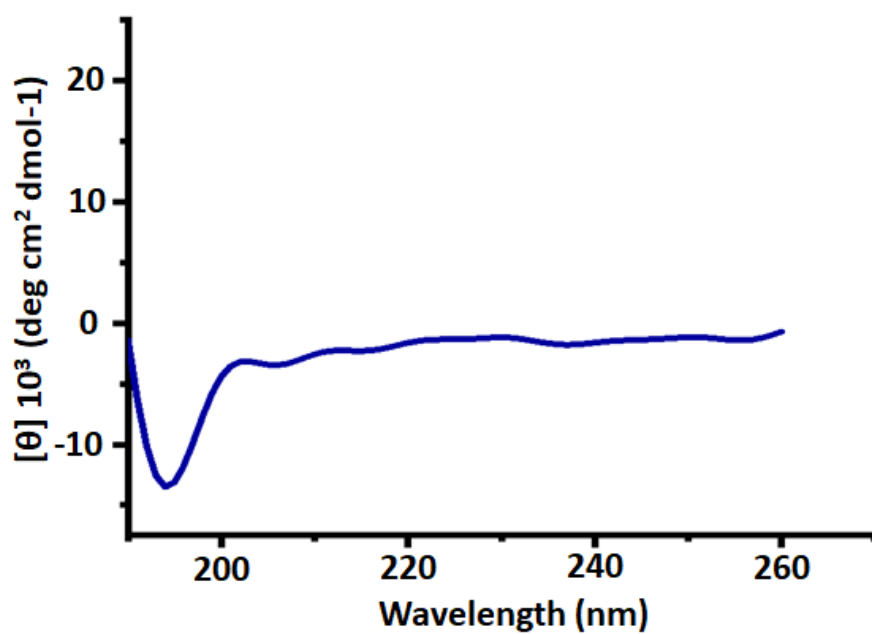
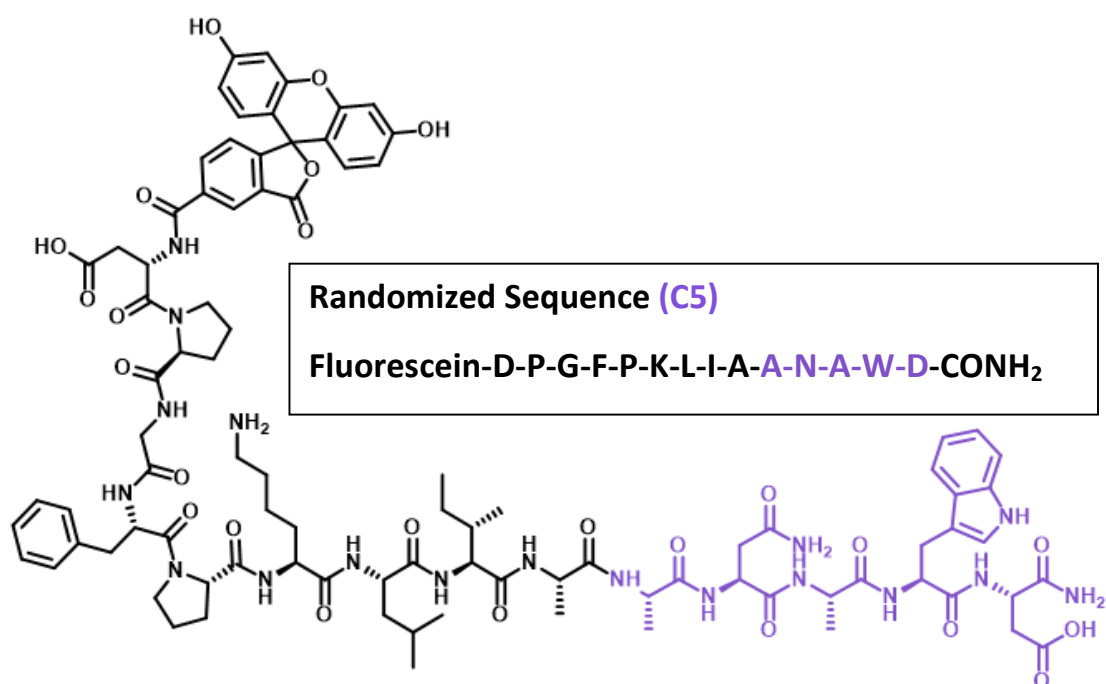
**Table S2.** Five manually selected conformations with the most significant changes in the dihedral angle in the peptide backbone or sidechain were taken as the initial conformations to start MD simulation in an explicit water model. Free energy is in kcal/mol.



**Figure S27.** The representative conformation of an MD run for cy(WPHPY) is shown in van der Waals (middle) and stick (right) representations. This conformation constitutes 57.7% conformations of a 20-ns long MD trajectory. The RMSD of the MD calculation is 1.066 angstrom. The average total energy for 20 ns is -15250.89 kcal/mol, the average kinetic energy is 3571.86 kcal/mol, and the average potential energy is -18822.76 kcal/mol.  $E\text{-total} = E\text{-kinetic} + E\text{-potential}$  is used. This energy is for the whole system (cy(WPHPY) + water + one  $\text{Cl}^-$ ).

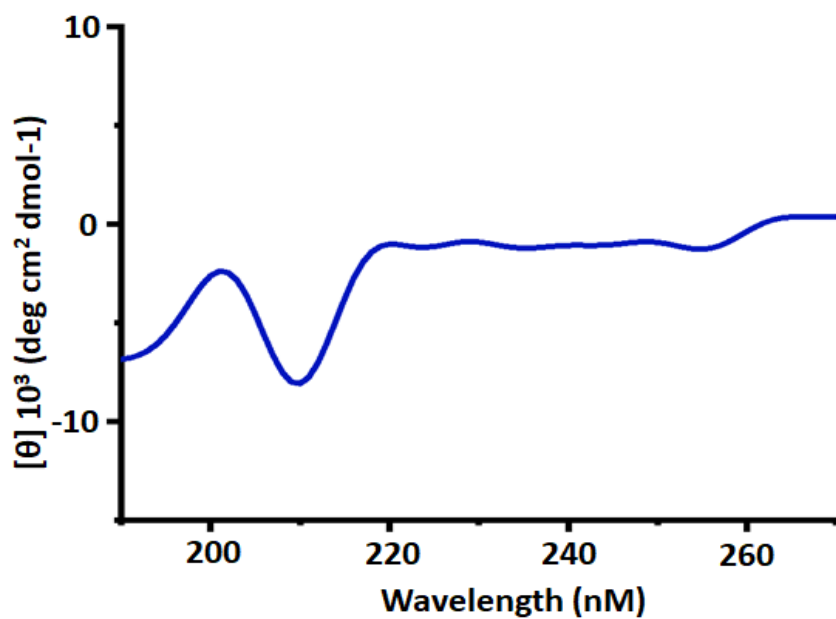
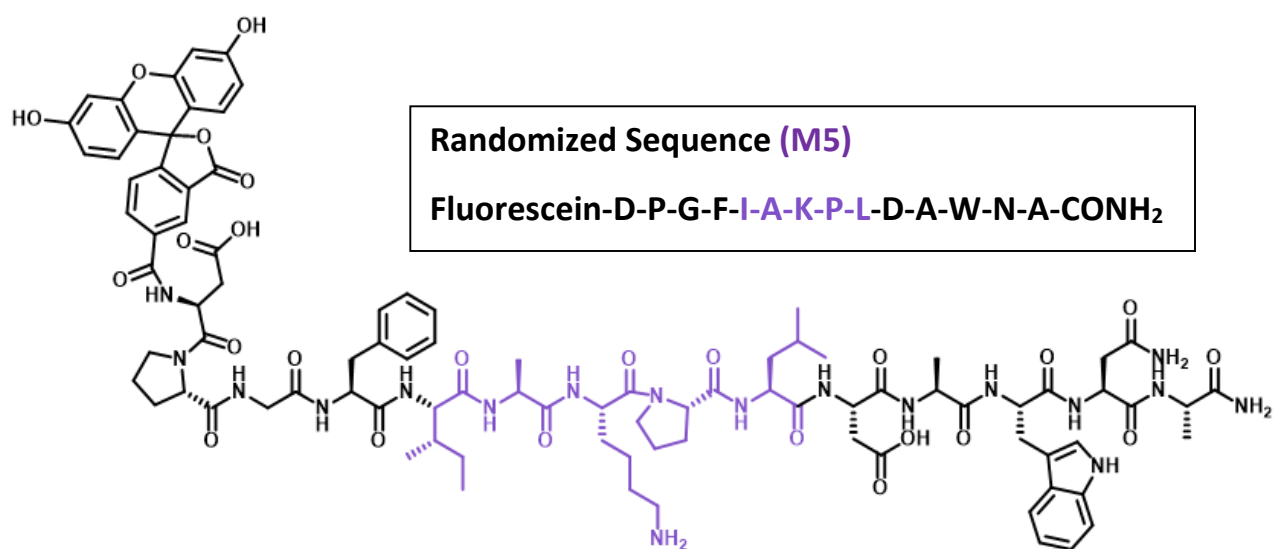


**Figure S28.** Circular dichroism (CD) spectrum of 20  $\mu\text{M}$  of the fluorescein-labeled epitope (D<sub>570</sub>-A<sub>583</sub>) in PBS. No apparent secondary structure was observed, and the epitope appeared to be a random coil.

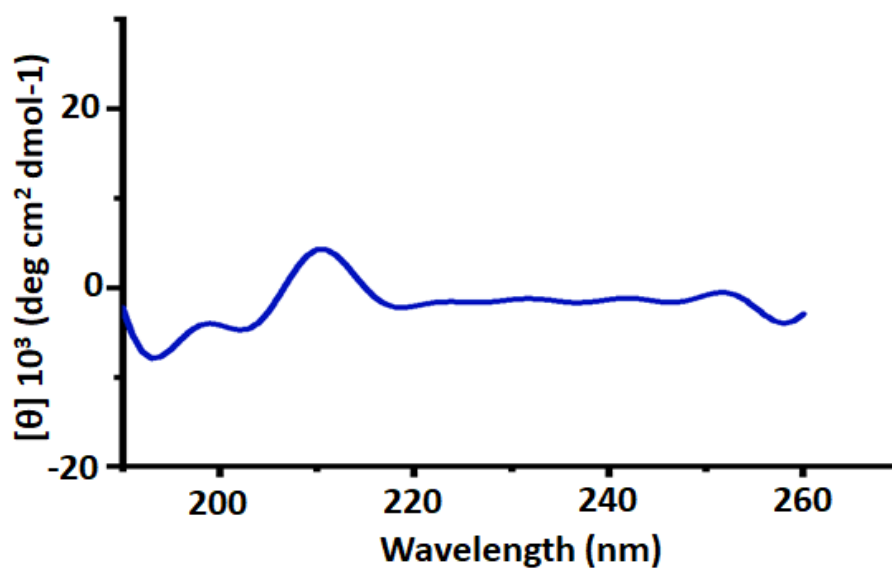
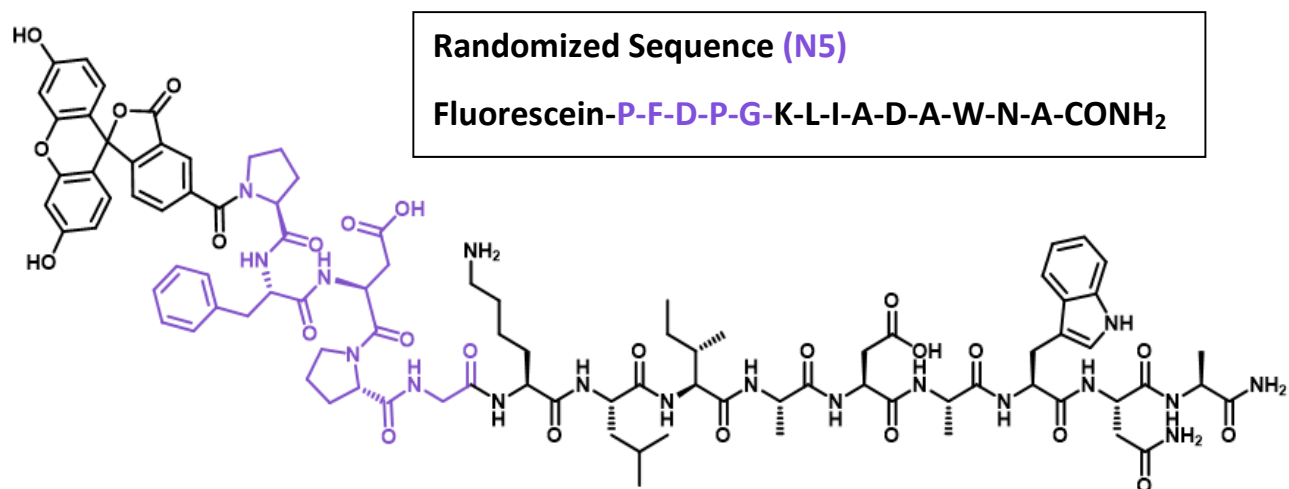


**Figure S29.** Circular dichroism (CD) spectrum of 20  $\mu$ M of the fluorescein-labeled randomized epitope (C5) in PBS. No apparent secondary structure was observed.

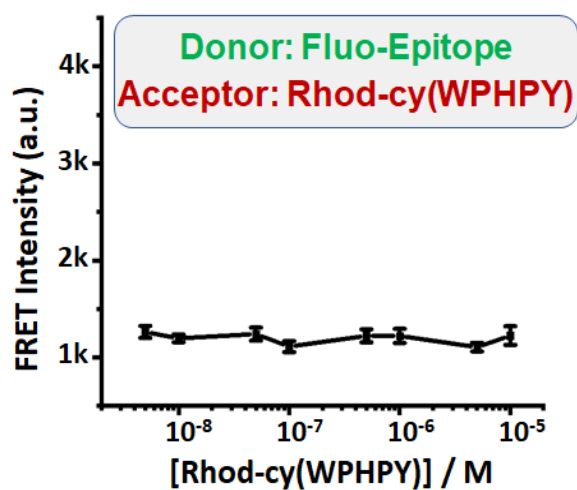
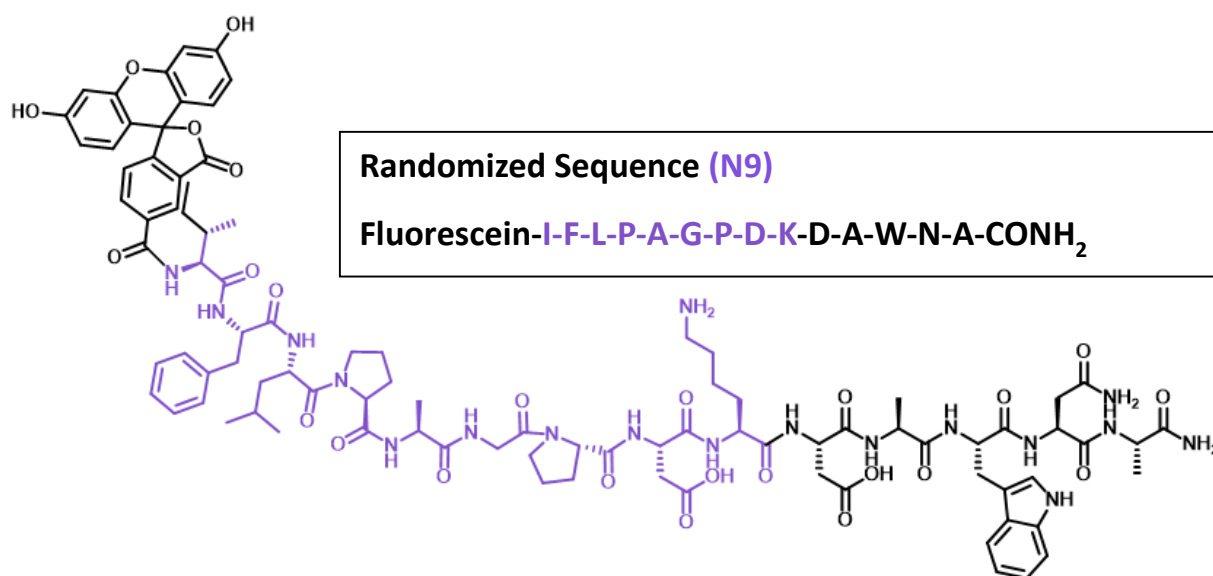




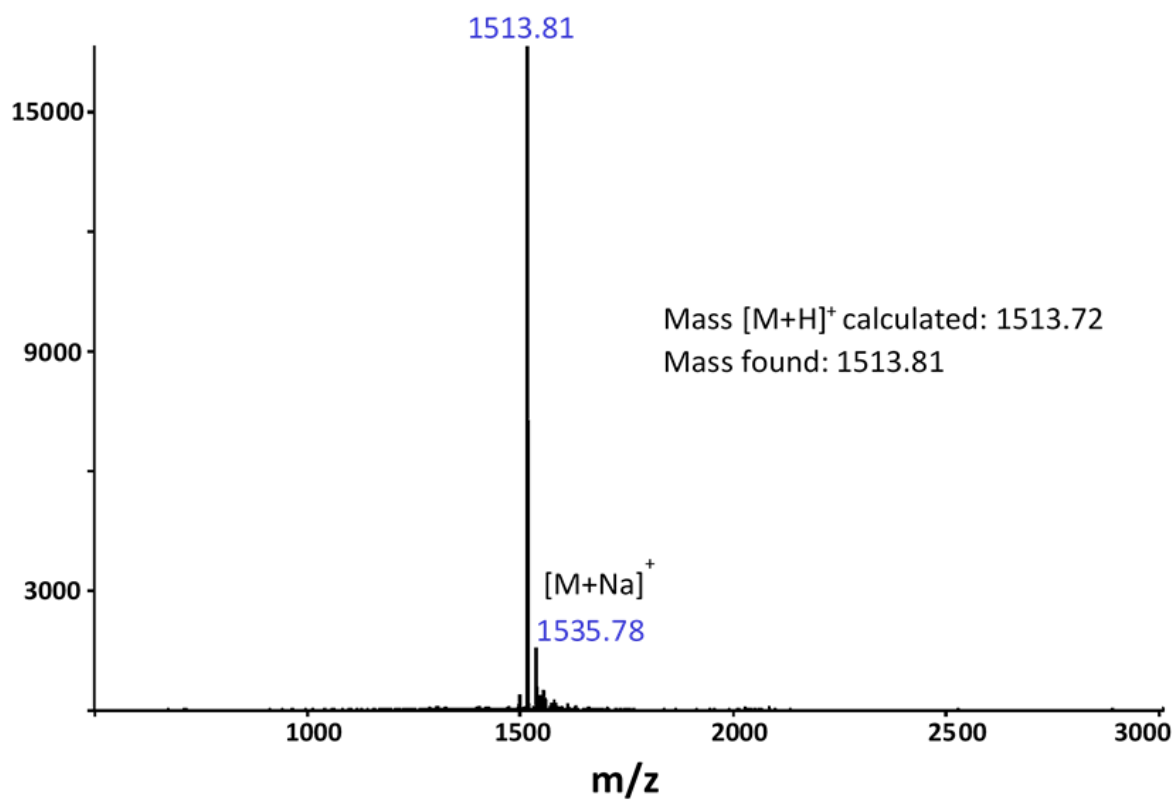
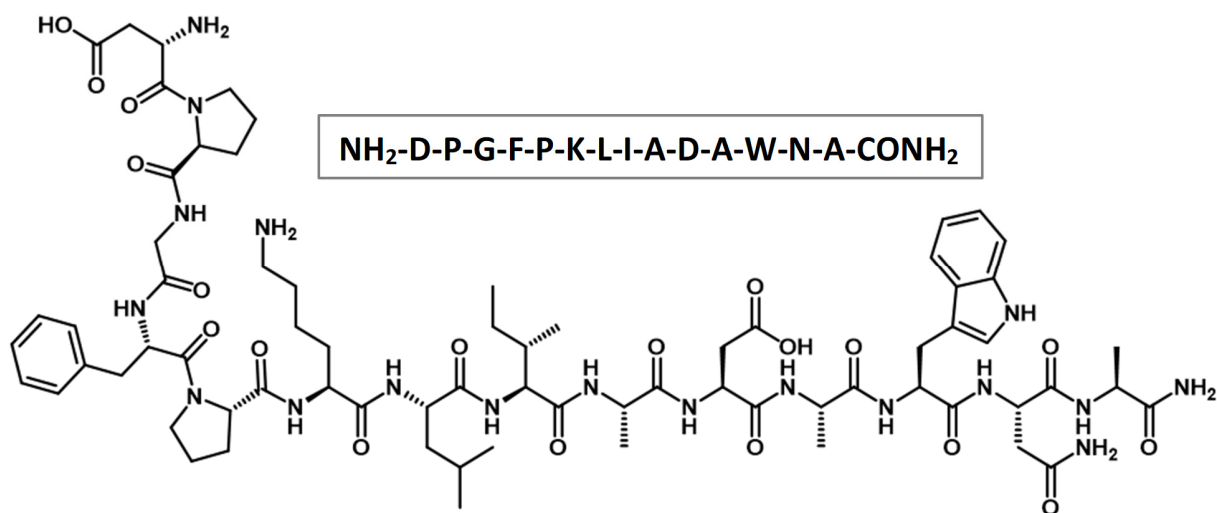
**Figure S30.** Circular dichroism (CD) spectrum of 20  $\mu\text{M}$  of the fluorescein-labeled randomized epitope (M5) in PBS. No apparent secondary structure was observed.



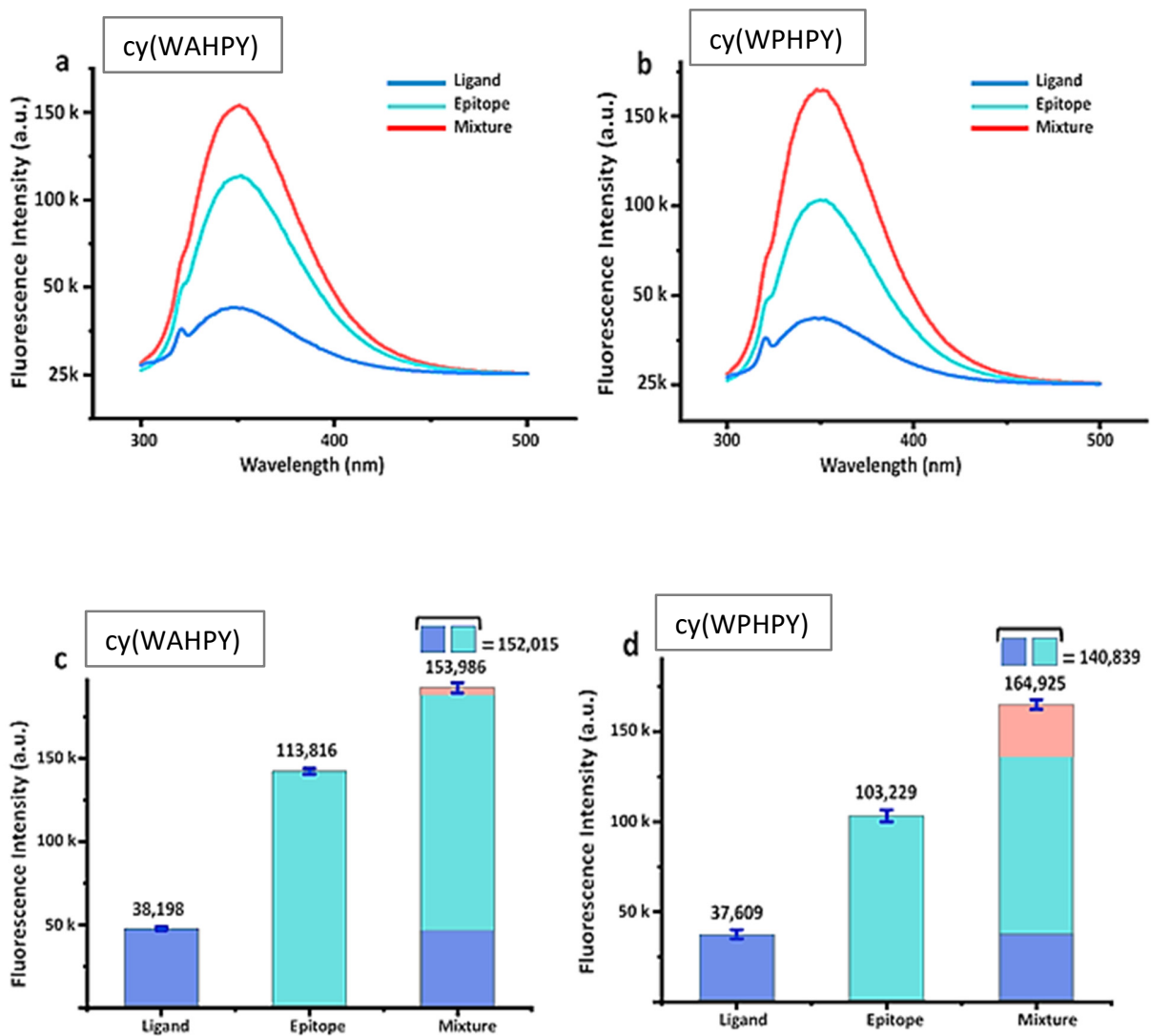
**Figure S31.** Circular dichroism (CD) spectrum of 20  $\mu$ M of the fluorescein-labeled randomized epitope (N5) in PBS. No apparent secondary structure was observed.



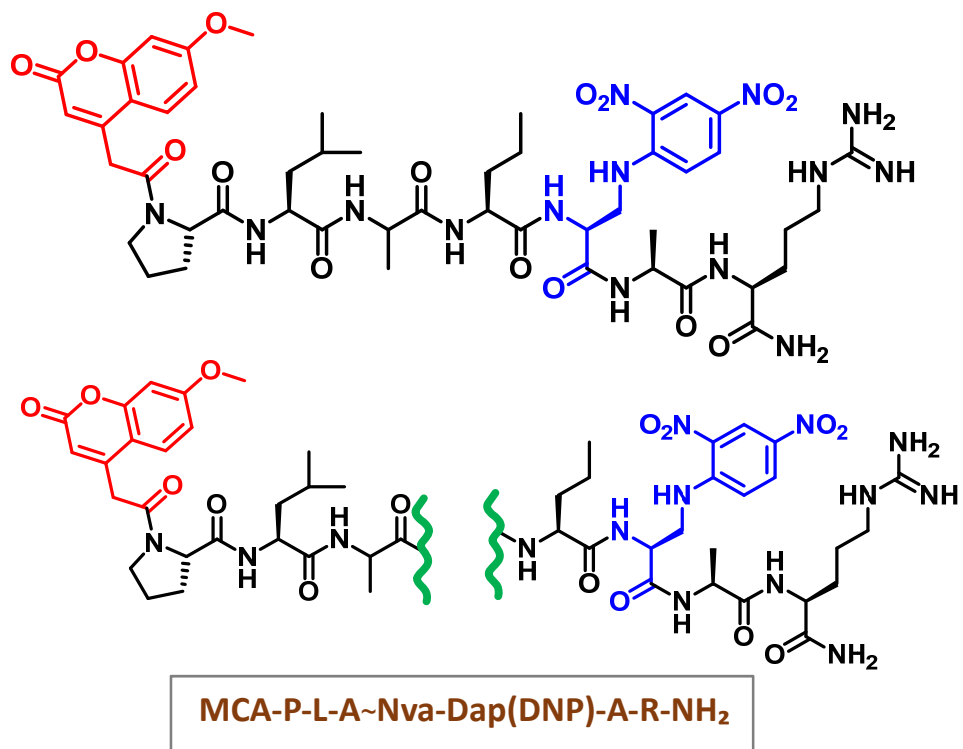
**Figure S32.** Since changing C5 showed a total loss of binding, this time we kept the first five amino acids at the C terminal unchanged and randomized the last nine amino acids. We ran a FRET assay using this epitope (N9) and Rhod-cy(WPHPY), which showed a loss of binding.



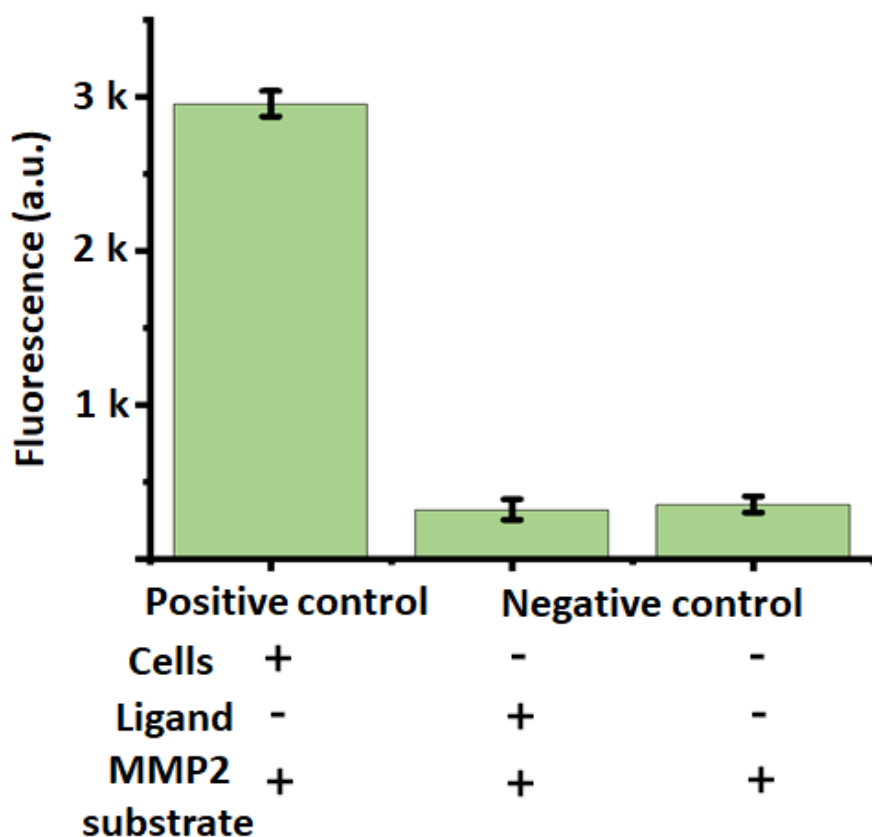
**Figure S33.** MALDI-TOF mass spectrum of the unmodified epitope.



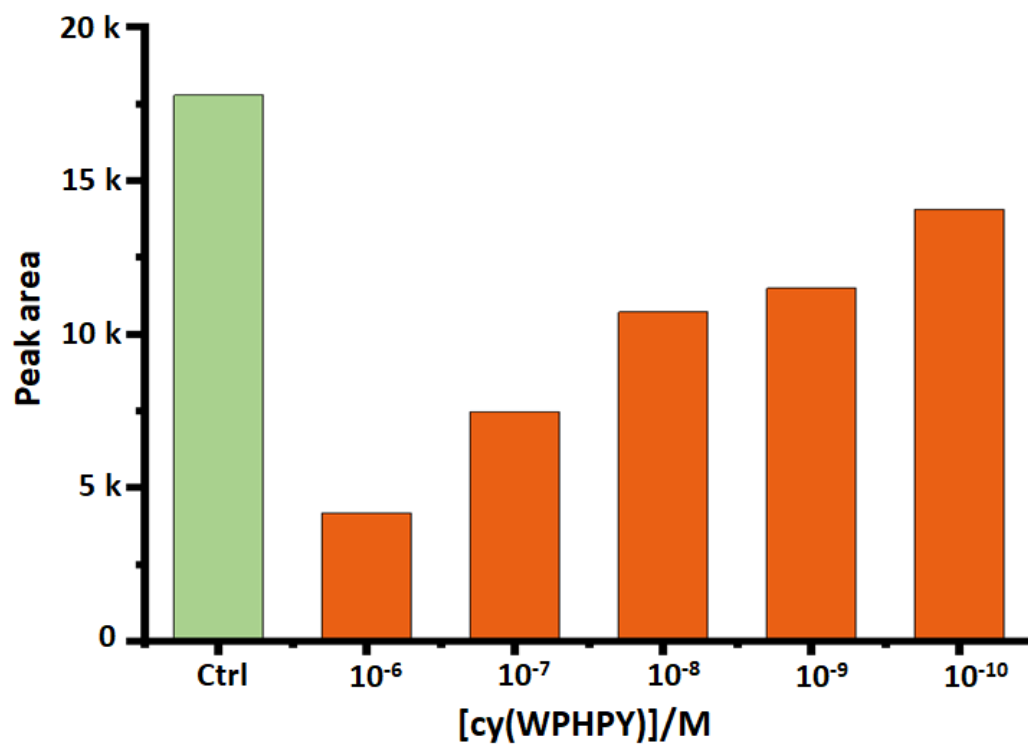
**Figure S34. a.** Emission spectra of the mutated ligand cy(WAHPY) (10  $\mu$ M), the epitope (10  $\mu$ M), and a mixture containing both. The spectra were obtained using an excitation wavelength of 288 nm. **b.** Similar emission spectra obtained using cy(WPHPY) (10  $\mu$ M), the epitope (10  $\mu$ M), and their mixture. **c,d.** Bar graph representations of **a** and **b** showing the fluorescence intensities at 350 nm. Only the cy(WPHPY)-epitope interaction led to an increased Trp fluorescence intensity.



**Figure S35.** The structure of the MMP2-specific fluorogenic peptide substrate. Green lines highlight the MMP cleaving sites.

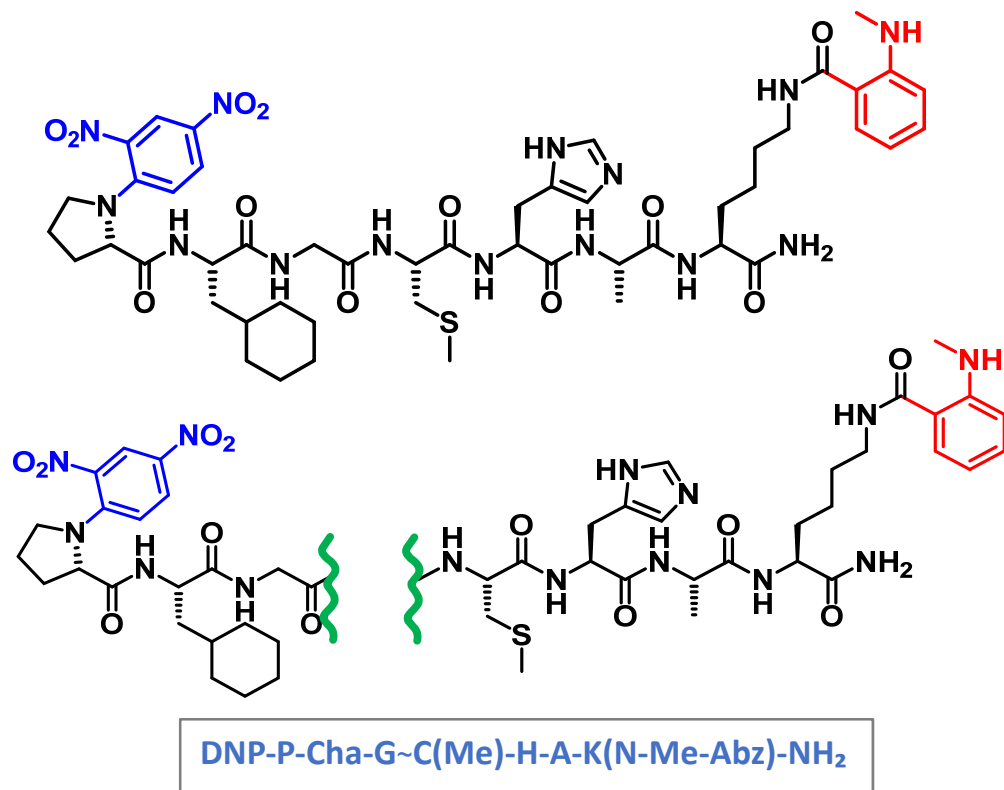


**Figure S36.** Positive and Negative controls used in the MMP2-specific substrate assay using MCA-P-L-A~Nva-Dap(DNP)-A-R-NH<sub>2</sub>. The positive control consists of the serum-free media collected from WM115 cells. This media was then concentrated with a centrifugal membrane filter (10 kDa cut-off, Amicon). Once concentrated, the total protein content was quantified using a BCA assay kit (Thermo). In a 96-well microplate, 10  $\mu$ l (1 $\mu$ g/ $\mu$ l protein content) of the media was mixed with 90  $\mu$ l of TNC assay buffer (50 mM Tris base, 0.15 M NaCl, 10 mM CaCl<sub>2</sub>, pH 7.5). The plate was left for 5 min at 37 °C for equilibration, and then 100  $\mu$ l of MMP2-specific substrate (3  $\mu$ M) was added. The plate was then incubated at 37 °C for 20 min. The fluorescence was measured using a SYNERGY H1 microplate reader. One of the negative controls has the ligand (1 $\mu$ M) in the serum-free media in which later the 3  $\mu$ M substrate was added, and the other one is the MMP2-specific substrate by itself in the serum-free media. The fluorescence was measured with a SYNERGY H1 microplate reader.

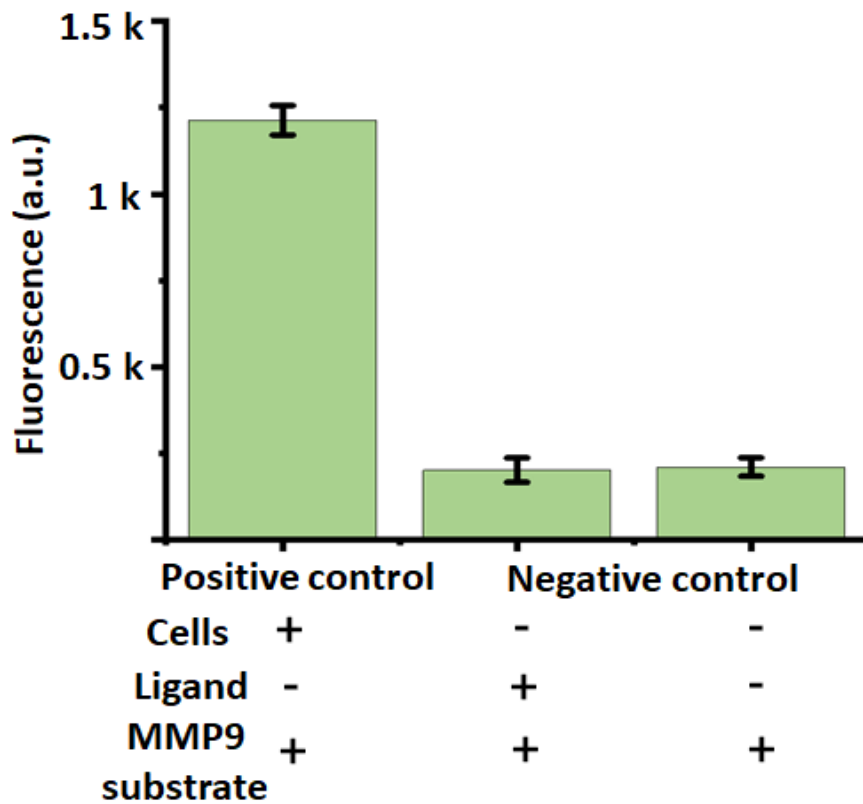


**Figure S37.** The intensities of the gelatin zymogram bands were measured using ImageJ software and plotted. The IC<sub>50</sub> value was estimated to be around 20 nM.

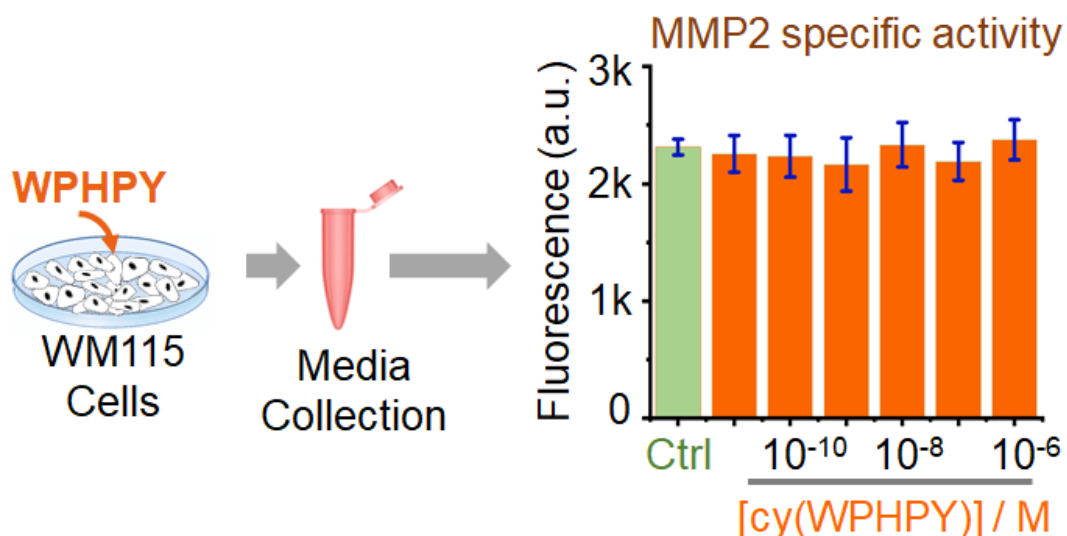




**Figure S38.** The structure of the MM9-specific fluorogenic peptide substrate. Green lines highlight the MMP cleaving sites.



**Figure S39.** Positive and Negative controls used in the MMP9-specific substrate assay using DNP-P-Cha-G~C(Me)-H-A-K(N-Me-Abz)-NH<sub>2</sub>. The positive control consists of the serum-free media collected from WM115 cells. This media was then concentrated with a centrifugal membrane filter (10 kDa cut-off, Amicon). Once concentrated, the total protein content was quantified using a BCA assay kit (Thermo). In a 96-well microplate, 10  $\mu$ l (1 $\mu$ g/ $\mu$ l protein content) of the protein solution was mixed with 90  $\mu$ l of TNC assay buffer (50 mM Tris base, 0.15 M NaCl, 10 mM CaCl<sub>2</sub>, pH 7.5). The plate was left for 5 min at 37 °C for equilibration, and then 100  $\mu$ L of MMP9-specific substrate (3  $\mu$ M) was added. The plate was then incubated at 37 °C for 20 min. The fluorescence was measured using a SYNERGY H1 microplate reader. One of the negative controls has the ligand (1 $\mu$ M) in the serum-free media in which later the 3  $\mu$ M substrate was added, and the other one is the MMP9-specific substrate by itself in the serum-free media. The fluorescence was measured with a SYNERGY H1 microplate reader.



**Figure S40.** The linear sequence WPHPY was synthesized to see if it shows any inhibition efficiency. WM115 cells were cultured in full media to 80% confluency and then changed to serum-free media containing varying concentrations of the linear sequence. The culture media was then taken out and the active MMP2 specific fluorogenic substrate was added to it. No significant change in the intensity was observed.

P	D	N	L	D	A	V	V	D	L	Q	Homology	Protein Name
P	D	N	L	G	A	V	V	D	L	Q	90%	Unnamed protein (614-624)
	D	S	L	D	A	L	V	K	L	Q	63%	Dynein regulatory complex subunit 3 (145-154)
	D	N	L	G	A	T	V	D	L		63%	Utp-glucose-1-phosphate Uridyltransferase (273-282)
	D	D	L	D	A	V	I	D			54%	Tetratricopeptide domain 22, isoform CRA_b (9-16)
P	E	D	L	D	A	V	V				54%	RASAL2 protein (13-20)
		N	F	D	A	I	V	D	I	Q	54%	Collagen alpha-1(XII) chain (289-297)
			L	D	V	V	V	N	L	Q	54%	Transcription factor RFX3 (395-402)

**Figure S41.** Sequence alignment and homology search result for the TIMP2 binding epitope on the 4<sup>th</sup> hemopexin blade of proMMP-2 using protein BLAST®. The epitope sequence is placed in the first line (navy background). The other MMPs with the sequence homology found from the search are shown in rows 2-7. Empty boxes represent completely mismatched residues.