Supporting information

Construction of a Stapled α-Helix Peptide Library Displayed on Phage for Screening of Galectin-3 Binding Peptide Ligands

Teerapat Anananuchatkul, Iou Ven Chang, Takayuki Miki, Hiroshi Tsutsumi and Hisakazu Mihara*

School of Life Science and Technology, Tokyo Institute of Technology

4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503 Japan

Contents

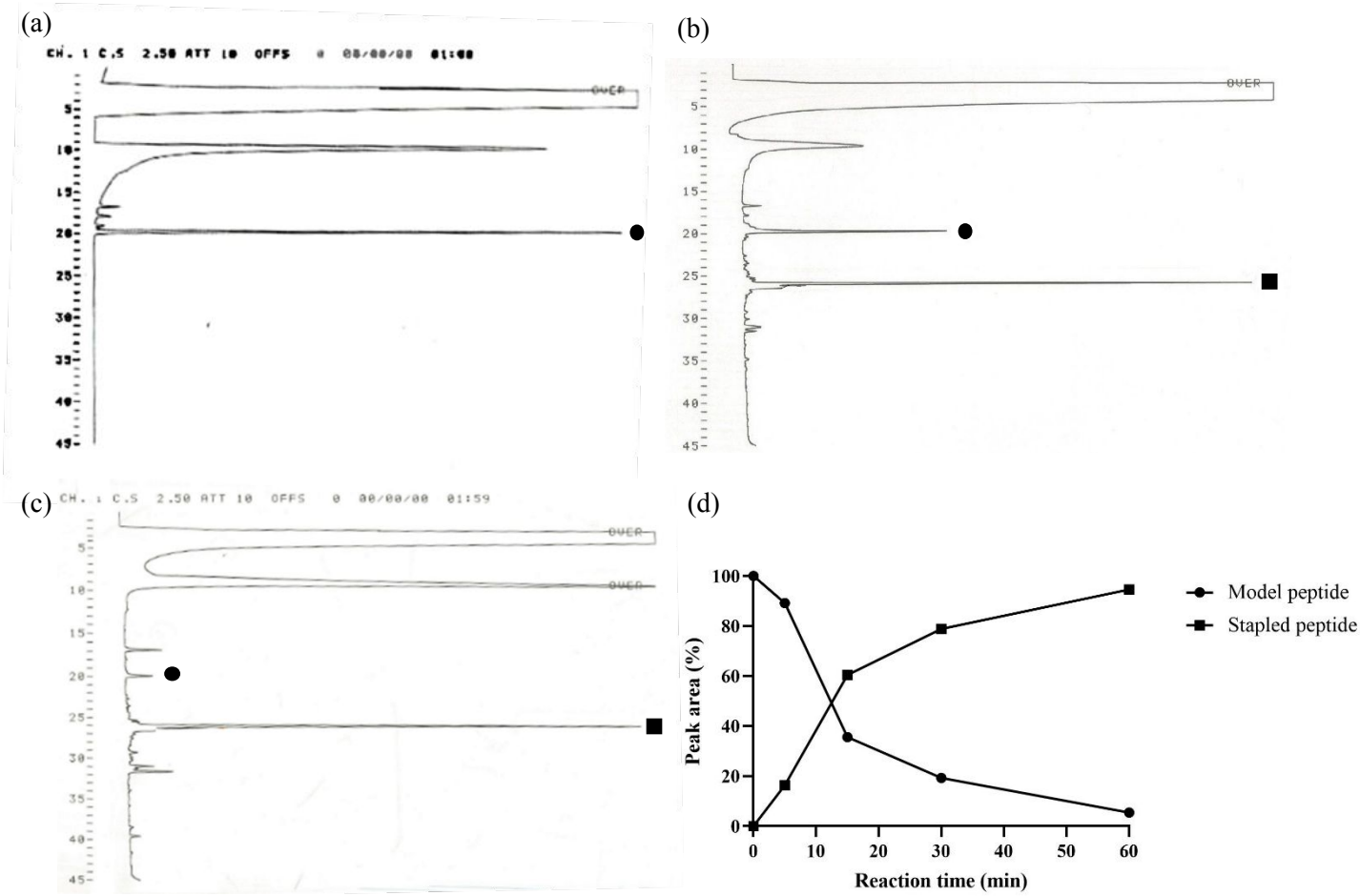
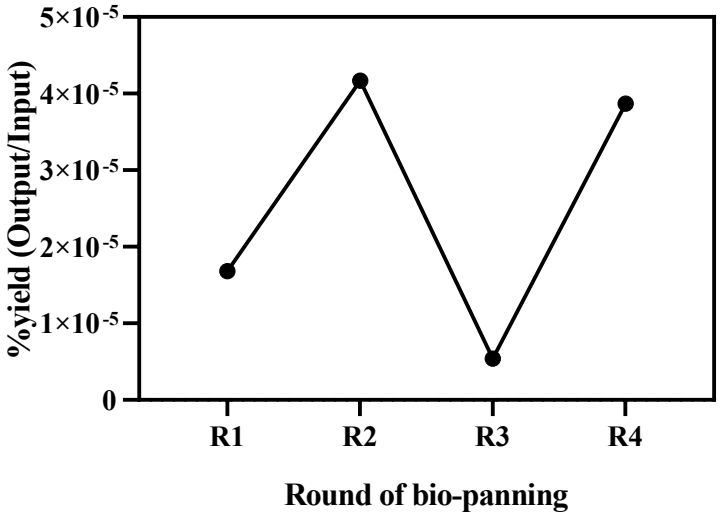


Figure S1. The HPLC analysis of modification of a model peptide (H-GAAECAAKEAACKAAG-NH_2) by a BP linker. Cosmosil 5C18-AR-II packaged column (4.6 x 150 mm) was used for analysis with a linear gradient of acetonitrile (0.08% TFA) from 0% to 30% over 30 min. (a) Chromatogram of the model peptide before modification, retention time of the model peptide is at 20 minutes (●). (b) Chromatogram of the model peptide after 30 min modification by BP linker, retention time of stapled peptide is at 26 minutes (■). (c) Chromatogram of the model peptide after 1 h modification by BP linker. (d) The time profile in the decrease of the model peptide and the increase of the stapled product calculated from peak area. The model peptide (100 μM) was chemical modified by BP staple linker (1 mM) in reaction buffer (20 mM NH_4HCO_3, 5 mM EDTA, pH 8.0 at 42°C for 1 hour.

Table S1. Phage titers from four rounds of biopanning.

Round of Biopanning	Input (cfu)	Output (cfu)	Percentage Yield (%)
1st	1.94×10^{10}	3.25×10^{3}	1.68×10^{-5}
2nd	1.41×10^{10}	5.88×10^{3}	4.17×10^{-5}
3rd	3.71×10^{10}	2.00×10^{3}	5.40×10^{-6}
4th	1.26×10^{10}	4.88×10^{3}	3.87×10^{-5}



Round	1	2	3	4
Amount of Galectin-3 (μg)	1	1	0.5	0.5

Figure S2. The recovery yield of a phage pool from each round of biopanning.

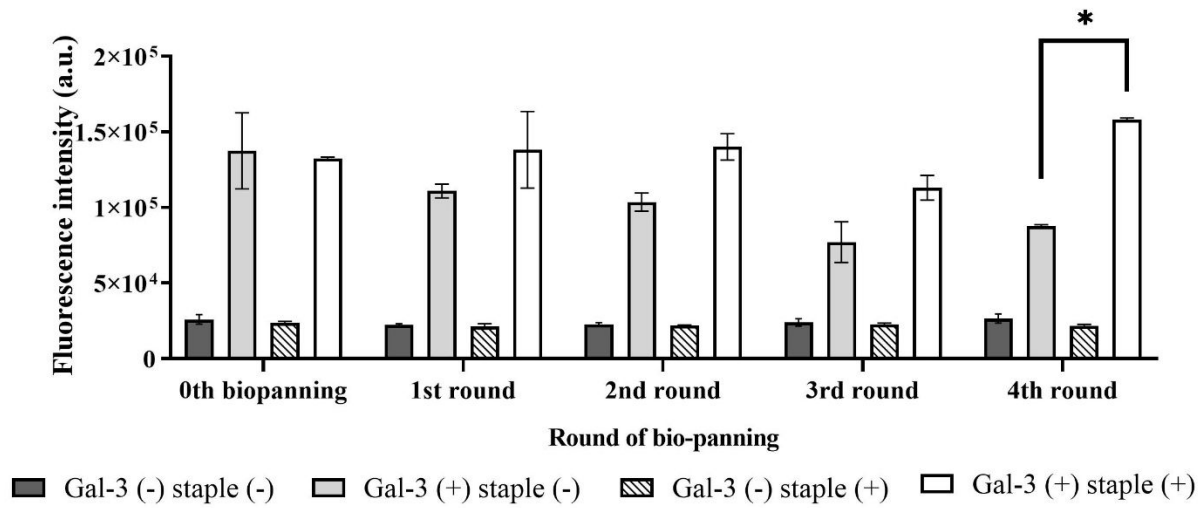


Figure S3. Phage library ELISA from each round of biopanning. The grey solid and black strip bars indicated the fluorescent intensity of nonstapled and stapled phage pools to microplate without Gal-3 (Gal-3(-)), respectively. The light grey solid and white solid bars indicated the fluorescent intensity of nonstapled and stapled phage pools to microplate with Gal-3 (Gal-3(+)), respectively. For all samples, n = 3, error bars represent the standard deviation. *$p < 0.05$.

Table S2. The peptide sequences of 27 randomly selected phage clones after the fourth biopanning. Twenty-six peptides were categorized based on a total net charge at pH 7.0. The color of the amino acid residues represents the residue properties; red for negatively charged residue (D and E), blue for positively charged residue (R, K and H) and the highlighted for extra cysteine and mutation.

Random Phage Clones Peptide Sequence ($GAX_1ECX_2X_3KEX_4X_5CKX_6AG$)				
Total net charge				
+2	+1	0	-1	-2
KPHTTQ	KWTNTS	FYSQMP	NYGPEQ	YVDEGG
PKRYEK	GQVPRS	GGGTGN	TDDHNQ	YSDSDL
	AQVFSH	QHNQDN		
	TQMTPH	QEQQAK		
	QSYYSK	QLPPAY		
	IKSLIQ	QVYQSS		
	LGSQLR	NPNYNN		
	QPPPAK	SNNMTN		
	SHNIEK	SRESYS		
		GLQLYS		
		MRDSGN		
Extra Cysteine Residue				
QVCQGQ				

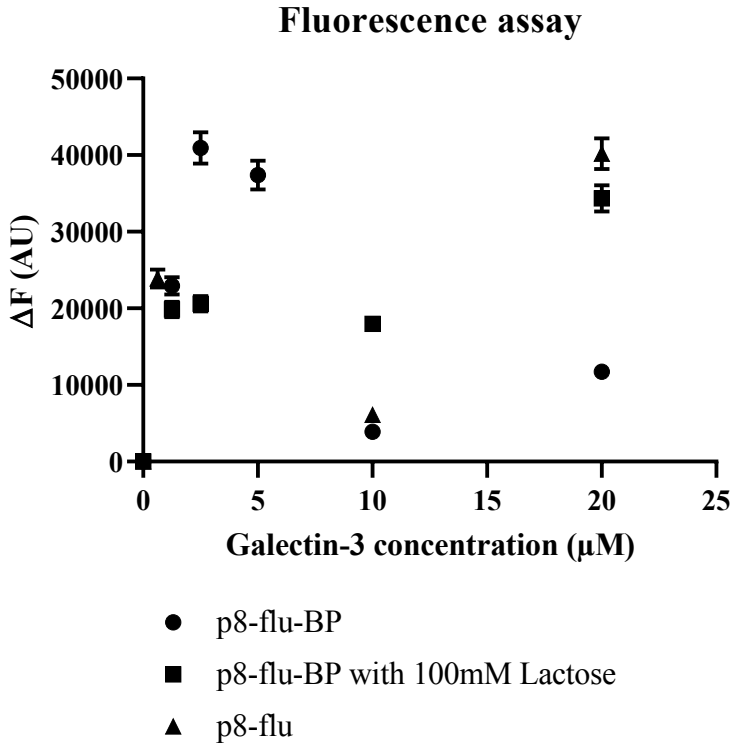


Figure S4. Fluorescence titration of p8-flu-BP (●), p8-flu-BP with 100 mM lactose (■), p8-flu (▲) to Gal-3. Peptide concentration was 10 μM in HBS at 25°C.

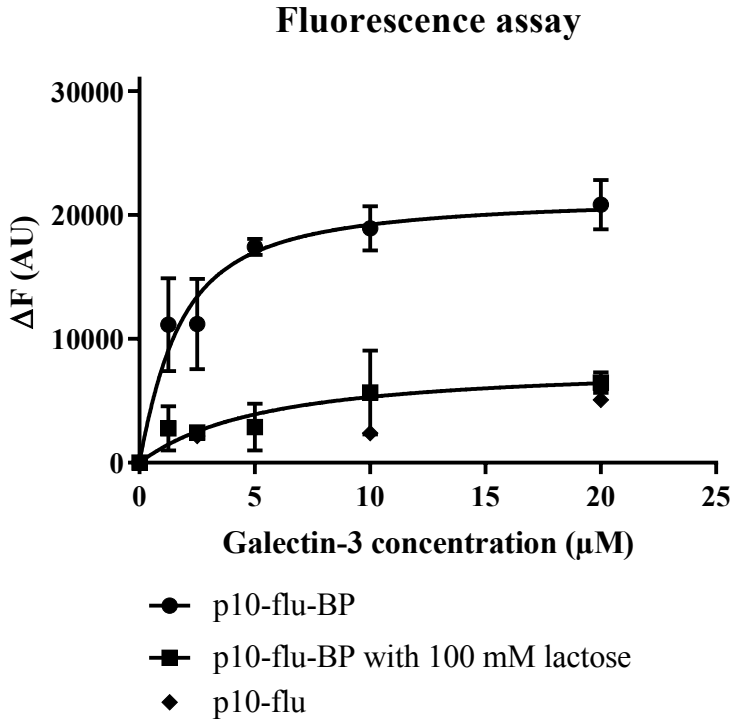


Figure S5. Fluorescence titration of p10-flu-BP (●), p10-flu-BP with 100 mM lactose(■), p10-flu (◆) to Gal-3. Peptide concentration was 10 μM in HBS at 25°C.

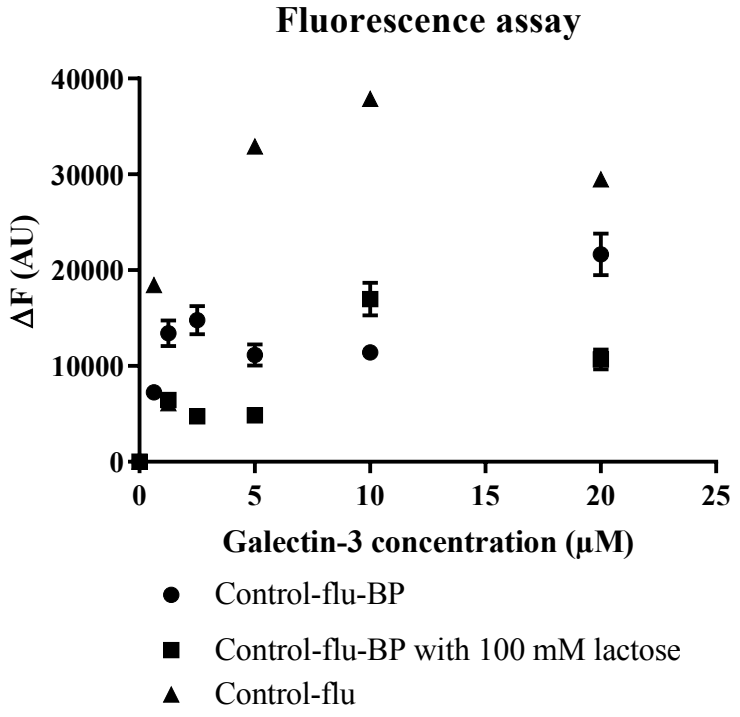


Figure S6. Fluorescence titration of Control-flu-BP (●), Control -flu-BP with 100 mM lactose(■), Control -flu (◆) to Gal-3. Peptide concentration was 10 μM in HBS at 25°C.

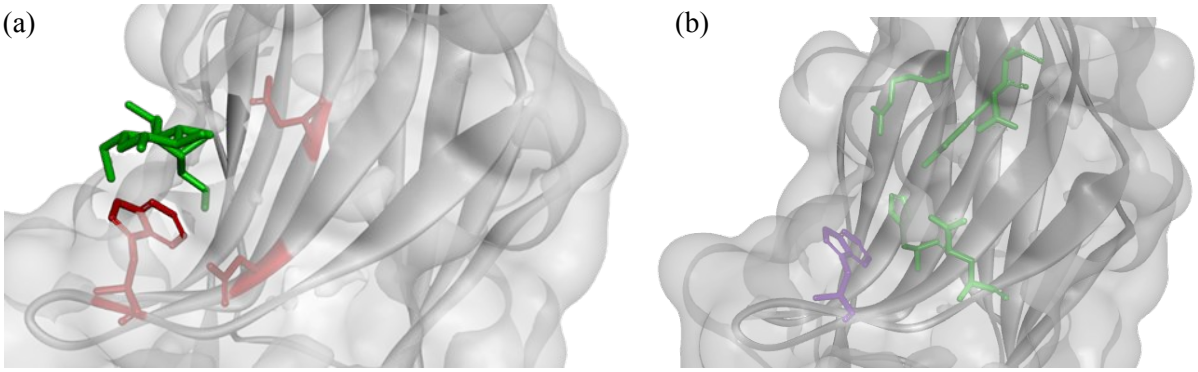


Figure S7. (a) A structure of Gal-3 (grey) and lactose (green) at the carbohydrate-binding domain (Trp 181, Asn 160 and Asn 174; red). (b) Charged amino acids (Glu165, Glu184, Arg144, His158 and Arg162; green) and Trp181 (purple) at the carbohydrate-binding domain. The crystal structure data of Gal-3 with a lactose was obtained from Protein Data Bank (PDB: 4R9A)[S1] and edited by Discovery studio software.

Experimental

Construction of peptide phage library

The library plasmid (fdg3p0ss21-library) was constructed by inserting a designed randomized DNA encoding library gene (library-insert) into the multiple cloning sites (MSC) of the disulfide-free pIII phage vector (fd0D12).[S2-S5] In order to reduce the transformation of cells containing incomplete digested and/or circularized fd0d12 plasmid, a *Sfi1*-digested fdg30p0ss21-stop vector having the stop codon and *SfiI* restriction enzyme site at the upper region of the pIII coding region was previously established by our laboratory and used for ligation with a *Sfi1*-digested library-insert.

Two succeeding polymerase chain reaction (PCR) steps were utilized to prepare the DNA coding insert.[S2-S5] First, the prepcrba and sfi2fo primers were operated to amplify the D1 and D2 gene coding of the non-cysteine pIII; 'non-cys insert'. The forward primer, prepcrba was designed to add a linker (Gly-Ser-Gly), while the sfi2fo primer was designed to clone a *Sfi1* restriction site (GGCCTCGGGGGCC). Second, a 'library-insert' was constructed with a designed random peptide sequence DNA coding ($GAX_1ECX_2X_3KEX_4X_5CKX_6AG$; X_n = any of 20 natural amino acids). The non-cys insert was used as a template to clone a library-insert with a forward primer sficxxkexxcba and the reverse primer sfi2fo. The PCR product is purified by electrophoresis on 1% agarose gel in TBE buffer and the DNA extracted with a gel extraction kit (Promega). Both fdg3p0ss21-stop vector and library-insert were digested by *SfiI* enzyme and digested products were purified by electrophoresis on 1% agarose gel in the same manner. The library plasmid (fdg3p0ss21-library) was prepared by the ligation of *SfiI*-digested fdg3p0ss21-stop vector (20 μg) and *SfiI*-digested library-insert (10 μg) using Ligation high Ver.2 (TOYOBO) at 16°C for 16 h. After ethanol precipitation, the ligated product (fdg3p0ss21-library) was used without further purification.

TG1 *E. coli* competent cells (Lucigen) was transformed by electroporation using fdg3p0ss21-library. The transformed TG1 cells were screened at 37°C overnight on a LB agar plate (chloramphenicol, 30 μg/mL). The colonies were collected by scraping with 2 mL 2YT media to prepare the glycerol stock (15%). The size of the phage library was estimated approximately to 5.92×10^9 clones by counting the titering colonies of transformed TG1 cells on the LB supplemented chloramphenicol (30 μg/mL) agar plates, which sufficient for screening. The quality of the initial phage library was analyzed based on the sequence analysis (with the primer fdg3p0ss21-rev) by sequencing 35 singles clones, as a result, all clones contain the designed peptide sequence (data not shown). The amino acids at randomized positions were well consistent with the theoretical appearance and mutation was not found on any designed amino acid position. Primer sequences are listed in the Table S3.

Table S3. PCR primer sequences for PCR cloning of fdg3p0ss21-library vector, library-insert, stop-insert fd0D12 and fdg3p0ss21-stop vector. The *Sfi1* restriction enzyme recognition site is underlined. The DNA encoding the semi-random library sequence is shaded and the randomized positions are coded by NNK codon (N=A, T, C, G; K=G, T nucleotide). The usage of NNK codon is to eliminate two from three possible stop codons.

Primer	DNA sequence (5' to 3')
pelbsfiecofo	GCATGAATTCCGATGACTGAGGCCGGCTGGGCCGCATAGAAAGGAACAACT AAAGGAAT
ecoG3pNba	GCATGAATTCCAGTCAGTACGGCCTCGGGGGCCATGGCTTCTGGTACCCCG GTTAAC
sfistopba	TATGCGGCCCAGCCGGCCATGGTAATGAGGCGGATCCGGCGCTG
prepcrba	GGCGGTTCTGGCGCTGAAACTGTTGAAAGTAG
sfi2fo	GAAGCCATGGCCCCCGAGGCCCCGGACGGAGCATTGACAGG
sficxxkexxcba	TATGCGGCCCAGCCGGCCATGGCAGGCGCGNNKGAATGCNNKNNKAAAGA ANNKNNKTGCAAANNKGCGGGCGGATCCGGCGCTG
fdg3p0ss21-rev	TAATTGCTCGACCTCCTCTC

Peptide Synthesis

The peptides were synthesized by solid phase peptide synthesis (SPPS) on a TentaGel S RAM resin (0.24 mmol/g) using standard 9-flurenylmethyloxycarbonyl (Fmoc) chemistry-based strategy.[S6] Fmoc-protected amino acids (3 eq.) were coupled for 30 minutes with 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 3 eq.), 1-hydroxybenzotriazole monohydrate (HOBt H_2O, 3 eq.) and N, N-diisopropylethylamine (DIPEA, 6 eq.) in N-methyl-2-pyrrolidone (NMP). To remove the Fmoc-protecting group, 20% (v/v) of piperidine in NMP was incubated for 15 minutes. Before the next residue synthesis, the resin was washed five times with NMP.

To conjugate the fluorescent dye to a peptide, a Mtt protecting group on the lysine at C-terminus was selectively removed by several washes with a mixture of dichloromethane/triisopropylsilane/trifluoroacetic acid (94/5/1). The resin was neutralized with DIPEA and 5(6)-carboxyfluorescein (3 eq.) was coupled with peptides on resin using HBTU (3 eq.), HOBt H_2O (3 eq.) and DIPEA (6 eq.) in NMP for 2 hours at 40°C.

The peptides were cleaved from resin by incubating with a trifluoroacetic acid/triisopropylsilane/distilled water (95/2.5/2.5) mixture. After filtration, the peptide solution was concentrated under reduced pressure and precipitated in cold diethyl ether. The crude peptides were purified by reverse-phase HPLC on the Hitachi L7000 system using a Cosmosil 5C18-AR-II packaged column (10x250 mm) with a linear gradient of acetonitrile containing 0.08% (v/v) TFA and water containing 0.1% (v/v) TFA at a flow rate of 3 mL/min. The purified peptides were identified by electrospray ionization mass spectrometry (ESI-MS) using a Shimadzu LCMS-2010 and the observed mass values corresponded well to the calculated values, within 1.0 mass unit (Table S4-S7). The purified peptides were obtained as TFA salts after lyophilization.

The purified peptides were dissolved in reaction buffer (20 mM NH_4HCO_3, 5 mM EDTA, pH 8.0) and incubated with the BP linker in acetonitrile (final concentration; peptide: 100 μM, BP: 1 mM). The modification reaction was conducted at 42°C for 1 hour and the stapled peptides were purified by RP-HPLC and identified by ESI-MS.

Table S4. Sequences and ESI-MS data of synthetic peptides.

Peptide sequence	Abb.	MW (g/mol)	m/z (Calculation)	m/z (Observation)
H-GA**Q**EC**VY**KE**QS**CK**S**AG-NH_2	p4	1686.88	$[M+H]^+$ = 1687.88 $[M+2H]^{2+}$ = 844.44 $[M+3H]^{3+}$ = 563.29	843.89 563.14
H-GA**Q**EC**PP**KE**PA**CK**K**AG-NH_2	p8	1612.89	$[M+H]^+$ = 1613.89 $[M+2H]^{2+}$ = 807.45 $[M+3H]^{3+}$ = 538.63	807.01 538.71
H-GA**P**EC**KR**KE**YE**CK**K**AG- NH_2	p10	1796.10	$[M+H]^+$ = 1797.10 $[M+2H]^{2+}$ = 899.05 $[M+3H]^{3+}$ = 599.70	898.47 599.48
H-GA**A**EC**AA**KE**AA**CK**A**AG-NH_2	Control	1420.63	$[M+H]^+$ = 1421.63 $[M+2H]^{2+}$ = 711.32 $[M+3H]^{3+}$ = 474.54	710.93 474.44

Table S5. Sequences and ESI-MS data of synthetic peptides labeled with a fluorescein.

Peptide sequence	Abb.	MW (g/mol)	m/z (Calculation)	m/z (Observation)
H-GA**Q**EC**VY**KE**QS**CK**S**AG-GGGK-(flu)-NH_2	p4-flu	2344.51	$[M+H]^+$ = 2345.51 $[M+2H]^{2+}$ =1173.26 $[M+3H]^{3+}$ = 782.50 $[M+4H]^{4+}$= 587.13	782.40 587.10
H-GA**Q**EC**PP**KE**PA**CK**K**AG-GGGK-(flu)-NH_2	p8-flu	2270.52	$[M+H]^+$ = 2271.52 $[M+2H]^{2+}$ = 1136.26 $[M+3H]^{3+}$ = 757.84 $[M+4H]^{4+}$= 568.63	757.80 568.60
H-GA**P**EC**KR**KE**YE**CK**K**AG-GGGK-(flu)-NH_2	p10-flu	2453.73	$[M+H]^+$ = 2454.73 $[M+2H]^{2+}$ = 1227.87 $[M+3H]^{3+}$ = 818.91 $[M+4H]^{4+}$= 614.23 $[M+5H]^{5+}$= 491.55	818.63 614.14 491.75
H-GA**A**EC**AA**KE**AA**CK**A**AG-GGGK-(flu)-NH_2	Control-flu	2078.26	$[M+H]^+$ = 2079.26 $[M+2H]^{2+}$ = 1040.13 $[M+3H]^{3+}$ = 693.75 $[M+4H]^{4+}$= 520.57	1039.93 693.35 520.32

Table S6. Sequences and ESI-MS data of synthetic peptides modified with BP.

Peptide sequence	Abb.	MW (g/mol)	m/z (Calculation)	m/z (Observation)
H-GA**Q**EC**VY**KE**QS**CK**S**AG-NH_2	p4-BP	1864.88	$[M+H]^+$ = 1865.88 $[M+2H]^{2+}$ = 933.44 $[M+3H]^{3+}$ = 622.63	932.94 623.04
H-GA**Q**EC**PP**KE**PA**CK**K**AG-NH_2	p8-BP	1790.87	$[M+H]^+$ = 1791.87 $[M+2H]^{2+}$ = 896.44 $[M+3H]^{3+}$ = 597.96	896.94 598.48
H-GA**P**EC**KR**KE**YE**CK**K**AG- NH_2	p10-BP	1974.08	$[M+H]^+$ = 1975.08 $[M+2H]^{2+}$ = 988.04 $[M+3H]^{3+}$ = 659.03	659.18
H-GA**A**EC**AA**KE**AA**CK**A**AG-NH_2	Control-BP	1598.61	$[M+H]^+$ = 1599.61 $[M+2H]^{2+}$ = 800.31 $[M+3H]^{3+}$ = 533.87	800.06 533.69

Table S7. Sequences and ESI-MS data of fluorescein-labeled peptides and modified with BP.

Peptide sequence	Abb.	MW (g/mol)	m/z (Calculation)	m/z (Observation)
H-GA**Q**EC**VY**KE**QS**CK**S**AG-GGGK-(flu)-NH_2	p4-flu-BP	2522.75	$[M+H]^+$ = 2523.75 $[M+2H]^{2+}$ =1262.38 $[M+3H]^{3+}$ = 841.92	841.60
H-GA**Q**EC**PP**KE**PA**CK**K**AG-GGGK-(flu)-NH_2	p8-flu-BP	2448.76	$[M+H]^+$ = 2449.76 $[M+2H]^{2+}$ = 1225.38 $[M+3H]^{3+}$ = 817.25 $[M+4H]^{4+}$= 613.19	816.95 612.85
H-GA**P**EC**KR**KE**YE**CK**K**AG-GGGK-(flu)-NH_2	p10-flu-BP	2631.97	$[M+H]^+$ = 2632.97 $[M+2H]^{2+}$ = 1316.99 $[M+3H]^{3+}$ = 878.33 $[M+4H]^{4+}$= 658.99	878.25 658.75
H-GA**A**EC**AA**KE**AA**CK**A**AG-GGGK-(flu)-NH_2	Control-flu-BP	2256.50	$[M+H]^+$ = 2257.50 $[M+2H]^{2+}$ = 1129.25 $[M+3H]^{3+}$ = 753.17 $[M+4H]^{4+}$= 565.13	1129.03 752.93 564.91

Reference

S1. Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S., & Rini, J. (1998). X-Ray Crystal Structure of The Human Galectin-3 Carbohydrate Recognition Domain (Crd) At 2.1 Angstrom Resolution. *The Journal of Biological Chemistry*, *273*(21), 13047–13052. doi: 10.2210/pdb1a3k/pdb

S2. Arai, K., Tsutsumi, H., & Mihara, H. (2013). A monosaccharide-modified peptide phage library for screening of ligands to carbohydrate-binding proteins. *Bioorganic & Medicinal Chemistry Letters*, *23*(17), 4940–4943. doi: 10.1016/j.bmcl.2013.06.059

S3. Heinis, C., Rutherford, T., Freund, S., & Winter, G. (2009). Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nature Chemical Biology*, *5*(7), 502–507. doi: 10.1038/nchembio.184

S4. Chang, I. V., Tsutsumi, H., & Mihara, H. (2017). Screening for concanavalin A binders from a mannose-modified α-helix peptide phage library. *Mol. BioSyst.*, *13*(11), 2222–2225. doi: 10.1039/c7mb00495h

S5. Kather, I., Bippes, C. A., & Schmid, F. X. (2005). A Stable Disulfide-free Gene-3-protein of Phage fd Generated by In vitro Evolution. *Journal of Molecular Biology*, *354*(3), 666–678. doi: 10.1016/j.jmb.2005.09.086

S6. Chan, W. C. (2004). *Fmoc solid phase peptide synthesis: a practical approach*. Oxford: University Press.