Supporting Information for:

Hydrazide-Reactive Peptide Tags for Site-Specific Protein Labeling

Glenn M. Eldridge[†] *and Gregory A. Weiss**[,†,‡]

Departments of [†]Chemistry, [‡]Molecular Biology and Biochemistry,

University of California, Irvine CA 92697-2025 USA

Table of Contents **Page**

Labels used here in the supplemental are identical to those used in the main text.

Additional Experimental procedures

All reagents were purchased from Sigma Aldrich (Milwaukee, WI), and used without further purification. M13 phage display libraries were re-propagated using KO7 helper phage from New England Biolabs (Ipswich, MA). Boc-hydrazide Tentagel and nitrotetrazolium blue (NBT; a substrate for AP) were purchased from Acros Organics (Geel, Belgium). AMCA hydrazide, streptavidin conjugated to alkaline phosphatase, high-sensitivity streptavidin conjugated to horseradish peroxidase, and chloronapthol and diaminobenzidine (CN/DAB) (AP substrate) were purchased from Pierce (Piscataway, NJ). The gene for T4 lysozyme (mutant C54T and C97A) was a gift from Prof. Bryan Matthews (University Oregon). DNA sequencing was performed by GeneWiz (South Plainfield, NJ).

Primers used in this study:

LysPetR2 5'-ATAGCCCTCGAGCTATTACGGTAGATTTTTATACGCGTC-3'

LysPetF 5'-GAGGATCCGGAGATGAATATATTTGAAATGTTACGT-3'

LibPetF 5'-GGCTATCCATGGTTGCTACAAATGCCTATGCA-3'

LibPetR 5'-CAAATATATTCATTCCTCCGGATCCTCCACC-3'

M13F 5'-TGTAAAACGACGGCCAGT-3'

M13R 5'-CAGGAAACAGCTATGAC-3'

T7F 5'-TGTAAAACGACGGCCAGTTAATACGACTCACTATAGGG-3'

T7R 5'-CAGGAAACAGCTATGACGCTAGTTATTGCTCAGCGG-3'

Homolog**53** 5'-

GCTACAAATGCCTATGCAMRSMRSASCMASMRSASCTSCMRSMRSMRSGASMAS-

GASMRSTSCMRSGYGASCASCASCGGTGGAGGATCCGGA-3'

Degenerate bases in Homolog**53** are named in accordance with IUBMB conventions.

Construction of the homolog shotgun scanning library

The homolog shotgun scanning library of HyRe tag 53 was generated using Homolog**53** oligonucleotide targeting the pM1165a phagemid[1] for site-directed mutagenesis using previously reported protocols.[2]

Selections with homolog shotgun scanning library

The phage-displayed peptide library was incubated with Boc-hydrazide Tentagel (50 mg in 10 mL PBS) for 1 h at room temperature in a peptide reaction vessel. The filtrate was transferred to a new vessel, and incubated with hydrazide Tentagel for 20 to 60 min with decreasing concentrations of deprotected resin for each round of selection. The resin was washed six times with PBS (50 mL for 6 min). Additional washes included pH = 5.6 PBS (50 mL for 20 min) and HCl (0.1 M, 50 mL for 10 min). The phage-resin mixture was then pH-buffered with a final PBS wash (50 mL, 10 min), and were re-propagated and titered, as previously described.[2]

HyRe peptide fusions to T4 lysozyme

PCR with primers LibPetR and LibPetF amplified the genes encoding the peptides from the selected phage and incorporated an NcoI site at the 5'-terminus of the amplicon. The gene for T4 lysozyme with mutations C54A and C97T was amplified using primers LysPetF1 and LysPetR2, which incorporates a stop codon and an XhoI site at the 3'-terminus of the amplicon. The primers LibPetR and LysPetF both include complementary regions for gene synthesis by overlap extension PCR. The pET-28c(+) plasmid (Novagen) and the purified gene synthesis product were digested with the restriction enzymes NcoI and XhoI (New England Biolabs). The digested products were mixed and ligated using T4 DNA ligase (New England Biolabs) by incubation in ATP-supplemented buffer at 16 °C for 16 hours. The crude ligation product was transformed into

heat shock competent XL-1 *E. coli* cells, and plated onto Luria Broth (LB) agar plates supplemented with kanamycin (40 μg/mL) for overnight incubation at 37 °C. The DNA sequences of the resultant colonies were confirmed using T7 primers.

Synthesis of Rhodamine B hydrazide

Rhodamine B hydrazide was synthesized by adding tert-Butyl carbazat to NHS rhodamine B (Pierce), deprotecting with 50:50 TFA:DCM, and purifying by reverse phase HPLC.

Synthesis of HyRe peptides

The peptides **114** and **104** were synthesized on 0.2 mmol scale using standard solid-phase peptide synthesis with Fmoc-protected amino acids (Aroz technologies) on Rink amide resin (Novabiochem). The synthesized peptides with a carboxamide C-terminus were deprotected and cleaved from the resin using a mixture of 9.5 mL trifluoroacetic acid, 0.25 mL triisopropylsilane, and 0.25 mL of H_2O at room temperature for 4 h under N_2. The cleavage mixture was filtered from the resin, and precipitated in ice-cold diethyl ether for 20 min. The precipitate was next centrifuged at 3 krpm for 20 min at 4 °C, before being re-suspended in deionized H_2O. The peptides were purified by reverse-phase HPLC with a C_{18} column using a gradient from 100% solvent A (99.9% H_2O / 0.1%TFA) to 50% solvent B (95% acetonitrile / 4.9%H_2O / 0.1%TFA) over 40 min. Fractions containing the purified peptides were combined and concentrated using rotary evaporation, followed by vacuum centrifugation. Purified peptides were subjected to analytical HPLC and MALDI-TOF to verify purity (>80%), and confirm identity, respectively (**Supporting Figs. 7-8**).

MALDI-TOF sample preparation and analysis

After the peptides (1 mM) were treated with Aminomethylcoumarin acetate hydrazide (AMCA-hydrazide) (1 mM) in PBS for 2.5 h, the reaction mixtures were applied to a C_{18} Ziptip (Varian), washed with 0.1% TFA in deionized water, and eluted with a saturated solution of α-Cyano-4-hydroxycinnamic acid (CHCA) in 25% aqueous acetonitrile containing 0.1 % TFA. The eluates were deposited on a stainless steel target. An ABI SciEx TOF/TOF 5800 (Applied Biosystems) system with a reflectron using standard configurations and parameters was used to obtain the MALDI-TOF mass spectra. The averaged spectra were exported to Excel for analysis.

Table S1. Peptide Scaffolds Selected for Interaction with Hydrazide Tentagel.*

Scaffold	1st Round	2nd Round	3rd Round	4th Round	5th Round
CX5C					
CX5CX	1				
CX5XC2					
X2CX10CX2					
X2CX2CX2	1				
X2CX2CX3					
X2CX3CX2					
X2CX4CX2		1			
X2CX5C		1			
X2CX5CX2					
X2CX6CX2					
X2CX7CX2					
X2CX8CX2	2				
X2CX9CX2	1				
X4CX10CX4					
X4CX2GPX4CX4					
X5CX8CX5					
X5CX9CX4					
X6CX6CX6					
X6CX7CX5	4	9	4	8	8
X7CX4CX7	3	9	17	5	
X7CX5CX6					
X8					
SM	7				
Total Sequenced	19	20	21	13	8

*Phage-displayed peptide library scaffolds used for the selection of hydrazide binding peptides and the number of selectants from each scaffold for each round. The sequences from each round of selection are found in Table S2.

Table S2. Peptide Sequences from Phage Selection Rounds (Rd.) 1 through 5 and Homolog Shotgun Scanning Rounds 1 through 4.*

Naïve Library Selectants			
Rd.	HyRe	aa sequence	sib
1	1	CEGSGACstop	
	2	HKREHNCDTPPPTPCFKANR	
	3	RPTKNNCDHTTHCTCHQQNQ	
	4	GTHDTNCHPPAHSKCRDSDY	
	5	PTPARKPCPKPPCPPTPKEQ	
	6	KPNTNHCTNNNQHDCKKATH	
	7	VLCTWCSM	
	8	TEPATPPCHHTTCKPRTGAD	
	9	QHKHHEHCKGPQCQKQHRKE	
	10	FYCREHSPVAGstopCV	
	11	GWCKYYLRSCHCMV	
	12	ATCNSYLYSALCSV	
2	3	RPTKNNCDHTTHCTCHQQNQ	
	13	TKNKPTHCGHHECTQPQHNS	
	14	APRHHQHCPTHACHAEAQEN	
	15	HDAPGKCKHPNHNHCQASHI	
	16	HTDHSKCRRDHTNSCANTEQ	
	17	HKTNPEPCKNHCCTGQPDNK	
	18	GKKNNQCPHRTPTSCLSNYQ	
	19	MLCGTTWCMG	
	20	HPNPRQDCHTHQCRTPTQDK	
	21	GICWLLDNC	
	22	KNSGKMCTTQDTSICSHKIR	
	23	TYRHDHCNPSRPYRCNHSPR	
	24	AKTHAHKCNTQTCQHERTHT	
	25	HKTKNECSHHPTRPCTTAHD	2
	26	TATNTLCHKHTHPKCGENSS	
	27	HQNPRTNCKTANCATHAPTN	
	28	TKPHDRDCHKQTCATHNHHR	
	29	NKTNSHQCATHACEHHKTNR	
	30	TGAPEPECPANRCTETNRQE	
3	25	HKTKNECSHHPTRPCTTAHD	2
	31	HQKADHTCAKHECQKKPHTR	
	32	KTNHEQQCHHNTCHEKGPSQ	
	33	KTHKPPNCPPEPCHQKAQPH	
	34	HKNEPPHCQQHKCEQHAPRP	
	35	KTPPPPACTKHECHRTGHTN	
	36	TEKHRTHCHEEQCQKTTPPP	
	37	SHKKRHKCRHHQCQTNETTT	
	38	HKQKAQNCTHKECHQTKAHT	
	39	HTPHRHECEHRHCEPKNRET	
	40	RPPKGEQCHHAHCRQTEKPP	
	41	EKTKHHKCPTTHCHEKTSPR	
	42	TSKHQHCEQTTSTDCKTKNT	
	43	HRNHHTKCPQDTCQQAPTKQ	
	44	KKAHEHRCAHTTCEDRQKPT	
	45	TKNHAQHCTTTQCPPTTRTH	
	46	DKSTHHQCKHTICTEQHDHK	
	47	KQHSPQACHQKHCQRHAPAK	
	48	HKGSPPCTTEHHKTCRMNPT	
	49	HKNTPAKCPTTTCQTEKTAQ	
4	3	RPTKNNCDHTTHCTCHQQNQ	
	25	HKTKNECSHHPTRPCTTAHD	3
	50	HSDHHKCGKPREPTCQRTNS	
	51	GEHHKEHCKPRQCTTATPQA	
	52	GAKDNKHCHHQECTHSKTRT	
	53	HKTNHSCHKREQEHCRVTTT	2
	54	KNARDHTCDNHHCHHHTAAA	
	55	HKQETPHCTSHHCTQPKTPA	
	56	HKTKTTQCTKHQCPTQKINP	
	57	KPPTHNPCHHNSCSKKTTDP	
5	25	HKTKNECSHHPTRPCTTAHD	
	50	HSDHHKCGKPREPTCQRTNS	
	53	HKTNHSCHKREQEHCRVTTT	5
	58	KQTDTRCQSQLHKHCKIVEN	

Homolog Shotgun Scanning			
Rd.	HyRe	aa sequence	sib
2	59	RRSQQSCSRHEKERSSVSSS	
	60	RRTKKSCRKKNQERCRASTS	
	61	NKTKSSCQKSEQESYRVSTT	
	62	HQSKNSCSKPDHDRCQVSTS	
	63	SSSHRSSSQHEHEHCRVTSS	
	64	KQSHSSYQHREQERCQVTST	
	65	KHSQHSCRSQEQERSQVTTS	
3	61	NKTKSSCQKSEQESYRVSTT	2
	66	NKSHHSCNKREQERCSVTST	
	67	KKSQHSCHHREQERCQVTTT	
	68	KSSQKSCHRQDQEQCRVTST	
	69	HRSQQTCNRQEHEPCQVSTS	
	70	KKSQNSCNRRDHEHCSVTTS	
	71	KSSHNSCSHREHDRSRVTTS	
	72	HRSHHTCSHQEQERCRVSSS	
	73	KSTHHSCNSREQDRCSVTTT	
	74	HNTQQTCRKQEQEHCKVSSS	
	75	KSTHHSCNSREQDRCSVTTT	
	76	SHTKKTCRHHEHERCHVTSS	
	77	KHTHSSSRQKEQERCQVSSS	
4	61	NKTKSSCQKSEQESYRVSTT	
	78	HKSNHSSKNREHERCQVTST	3
	79	NHTQRSCKHHEQESCRVSSS	

*The column "sib" indicates the number of siblings or identical sequences identified in the same round.

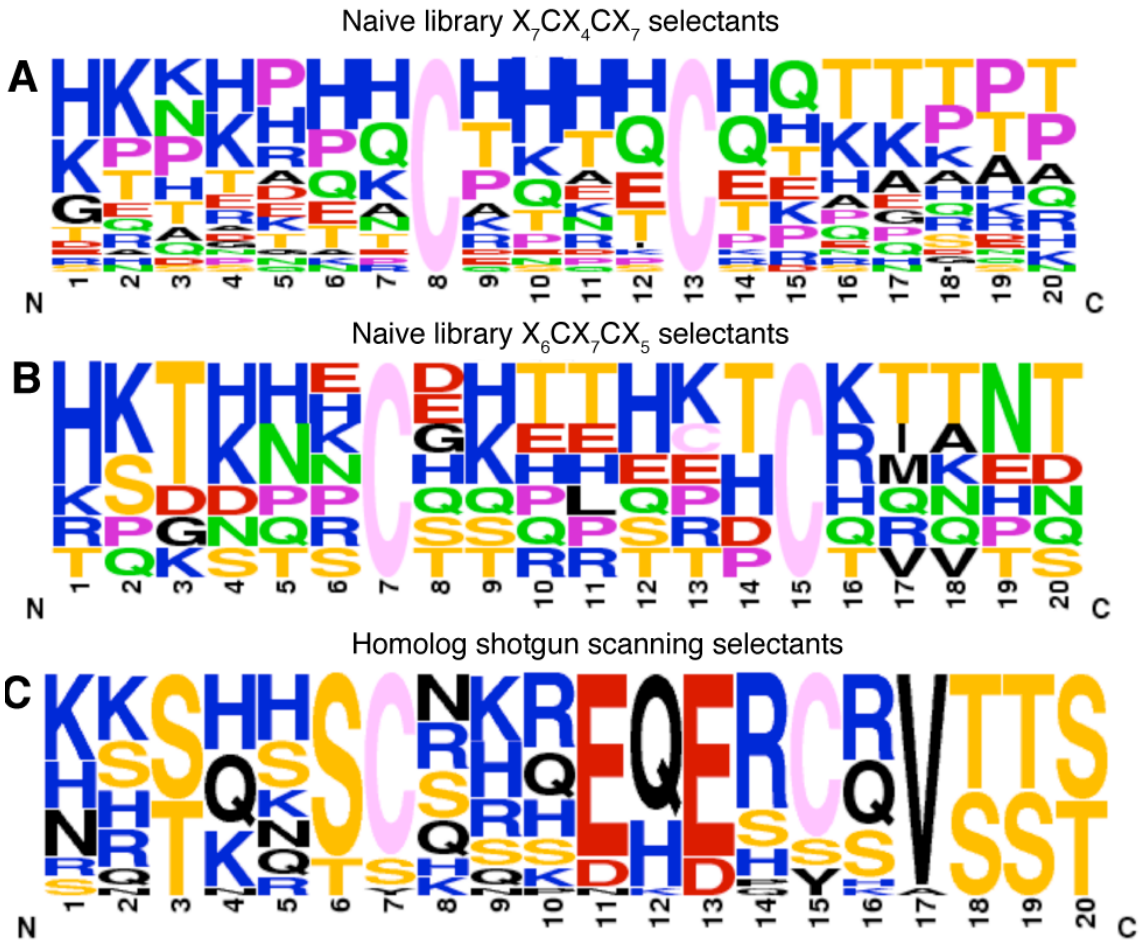


Figure S1. Sequence analysis of selectants from the naïve (A and B) and homolog shotgun scanning (C) libraries (weblogo.berkeley.edu).

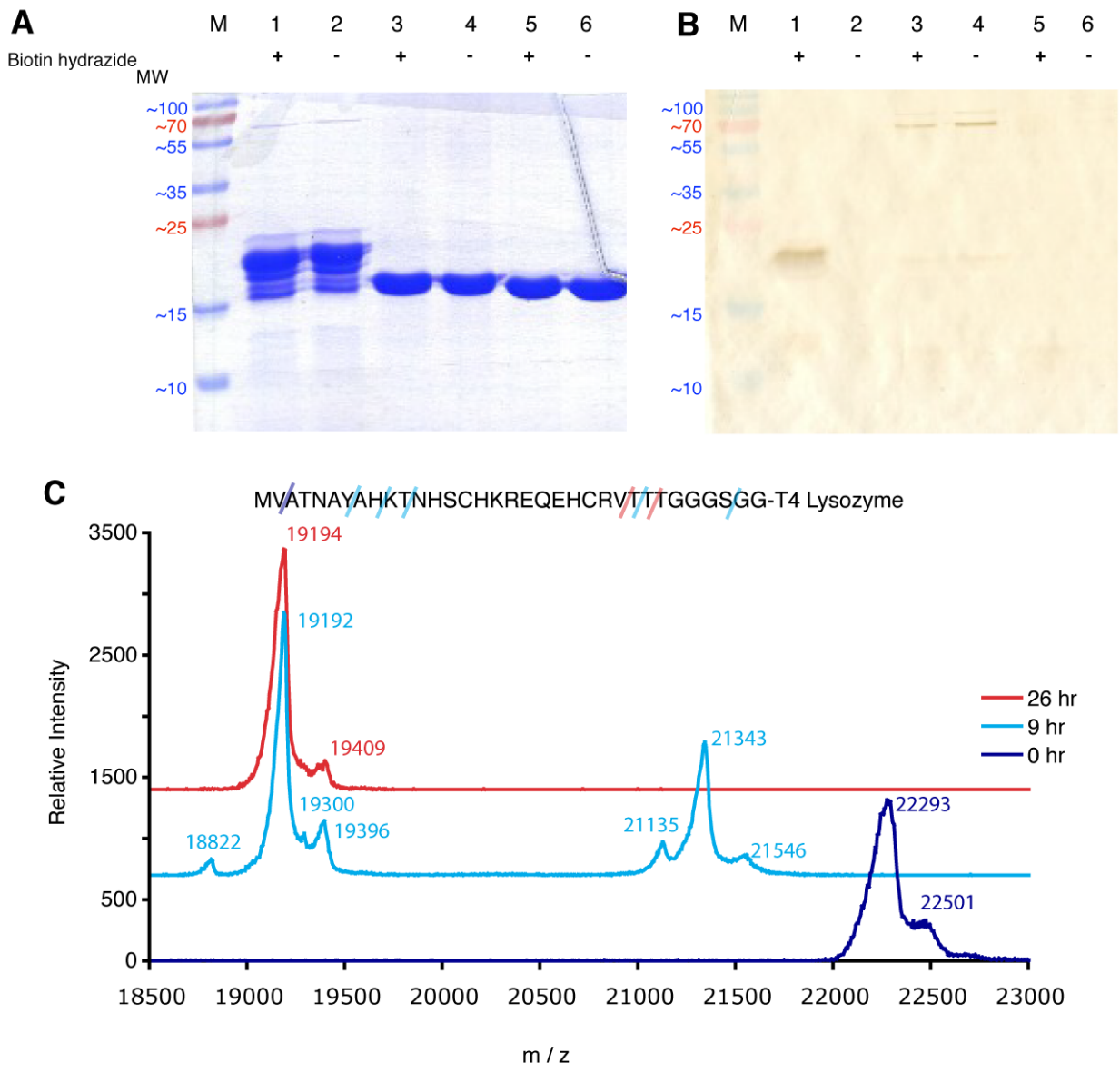
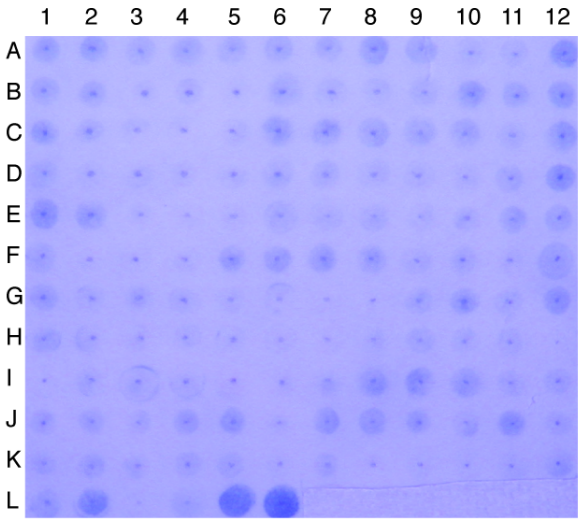


Figure S2. (A and B) The full SDS-PAGE and Western blot of the HyRe **53**-T4 lysozyme fusion excerpted in Figure 3. (C) MALDI-TOF of **53**-T4 lysozyme fusion incubated at room temperature for the indicated times. The MALDI-TOF results indicate that the HyRe tag is specifically cleaved from the fusion. These results also account for the partially proteolyzed fusion proteins after purification apparent in lanes 1 and 2 of the protein gel shown in (A).

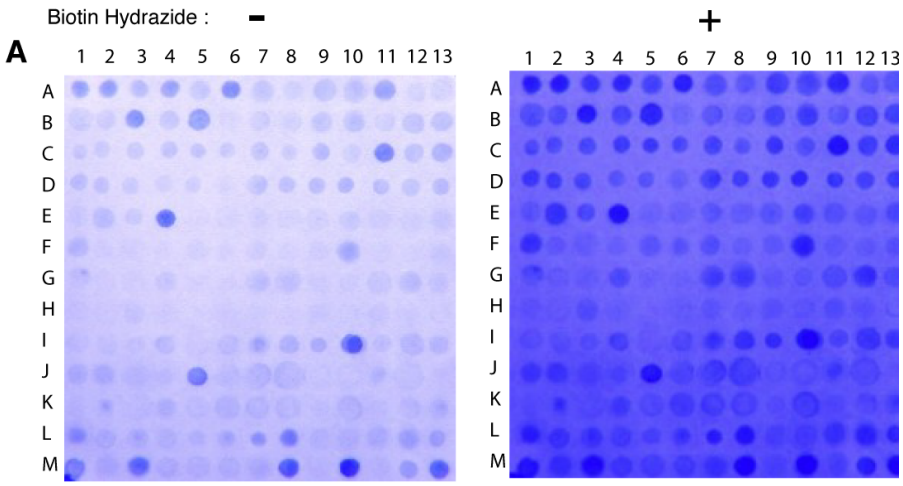
A



B

Spot	Sequence	Spot	Sequence	Spot	Sequence	Spot	Sequence
A1	HKTNHSCHKREQEHCRVTTT	C12	HKTNHSCHKREQEHCRTT	F11	HKTNHSCHKREQEHCT	I10	NKTKSSCQKS
A2	KTNHSCHKREQEHCRVTTT	D1	HKTNHSCHKREQEHCRVT	F12	HKTNHSCHKREQEHCR	I11	HKSNHSSKNREHERCQVTST
A3	HTNHSCHKREQEHCRVTTT	D2	NHSCHKREQEHCRVTTT	G1	HKTNHSCHKREQEHCRVTTT	I12	HKSNHSSKNREHERCQVTSA
A4	HKNHSCHKREQEHCRVTTT	D3	HHSCHKREQEHCRVTTT	G2	NKTKSSCQKSEQESYRVSTT	J1	HKSNHSSKNREHERCQVTAT
A5	HKTHSCHKREQEHCRVTTT	D4	HKSCHKREQEHCRVTTT	G3	NKTKSSCQKSEQESYRVSTA	J2	HKSNHSSKNREHERCQVAST
A6	HKTNSCHKREQEHCRVTTT	D5	HKTCHKREQEHCRVTTT	G4	NKTKSSCQKSEQESYRVSAT	J3	HKSNHSSKNREHERCQATST
A7	HKTNHCHKREQEHCRVTTT	D6	HKTNHKREQEHCRVTTT	G5	NKTKSSCQKSEQESYRVATT	J4	HKSNHSSKNREHERCAVTST
A8	HKTNHSHKREQEHCRVTTT	D7	HKTNHKREQEHCRVTTT	G6	NKTKSSCQKSEQESYRASTT	J5	HKSNHSSKNREHERAQVTST
A9	HKTNHSCKREQEHCRVTTT	D8	HKTNHSREQEHCRVTTT	G7	NKTKSSCQKSEQESYAVSTT	J6	HKSNHSSKNREHEACQVTST
A10	HKTNHSCHREQEHCRVTTT	D9	HKTNHSCEQEHCRVTTT	G8	NKTKSSCQKSEQESARVSTT	J7	HKSNHSSKNREHARCQVTST
A11	HKTNHSCHKEQEHCRVTTT	D10	HKTNHSCHQEHCRVTTT	G9	NKTKSSCQKSEQEAYRVSTT	J8	HKSNHSSKNREAERCQVTST
A12	HKTNHSCHKRQEHCRVTTT	D11	HKTNHSCHKEHCRVTTT	G10	NKTKSSCQKSEQASYRVSTT	J9	HKSNHSSKNRAHERCQVTST
B1	HKTNHSCHKREEHCRVTTT	D12	HKTNHSCHKRHCRVTTT	G11	NKTKSSCQKSEAESYRVSTT	J10	HKSNHSSKNAEHERCQVTST
B2	HKTNHSCHKREQHCRVTTT	E1	HKTNHSCHKRECRVTTT	G12	NKTKSSCQKSAQESYRVSTT	J11	HKSNHSSKAREHERCQVTST
B3	HKTNHSCHKREQECRVTTT	E2	HKTNHSCHKREQRVTTT	H1	NKTKSSCQKAEQESYRVSTT	J12	HKSNHSSANREHERCQVTST
B4	HKTNHSCHKREQEHRVTTT	E3	HKTNHSCHKREQEVTTT	H2	NKTKSSCQASEQESYRVSTT	K1	HKSNHSAKNREHERCQVTST
B5	HKTNHSCHKREQEHCVTTT	E4	HKTNHSCHKREQEHTTT	H3	NKTKSSCAKSEQESYRVSTT	K2	HKSNHASKNREHERCQVTST
B6	HKTNHSCHKREQEHCRVTT	E5	HKTNHSCHKREQEHCTT	H4	NKTKSSAQKSEQESYRVSTT	K3	HKSNASSKNREHERCQVTST
B7	HKTNHSCHKREQEHCRTTT	E6	HKTNHSCHKREQEHCRT	H5	NKTKSACQKSEQESYRVSTT	K4	HKSAHSSKNREHERCQVTST
B8	TNHSCHKREQEHCRVTTT	E7	HKTNHSCHKREQEHCRV	H6	NKTKASCQKSEQESYRVSTT	K5	HKANHSSKNREHERCQVTST
B9	HNHSCHKREQEHCRVTTT	E8	HSCHKREQEHCRVTTT	H7	NKTASSCQKSEQESYRVSTT	K6	HASNHSSKNREHERCQVTST
B10	HKHSCHKREQEHCRVTTT	E9	HSCHKREQEHCRVTTT	H8	NKAKSSCQKSEQESYRVSTT	K7	AKSNHSSKNREHERCQVTST
B11	HKTSCHKREQEHCRVTTT	E10	HKCHKREQEHCRVTTT	H9	NATKSSCQKSEQESYRVSTT	K8	EHERCQVTST
B12	HKTNCHKREQEHCRVTTT	E11	HKTHKREQEHCRVTTT	H10	AKTKSSCQKSEQESYRVSTT	K9	REHHRCQVTS
C1	HKTNHHKREQEHCRVTTT	E12	HKTNKREQEHCRVTTT	H11	NKTKSSCQKSEQESYRVSTT	K10	NREHERCQVT
C2	HKTNHSKREQEHCRVTTT	F1	HKTNHREQEHCRVTTT	H12	EQESYRVSTT	K11	KNREHERCQV
C3	HKTNHSCREQEHCRVTTT	F2	HKTNHSEQEHCRVTTT	I1	SEQESYRVST	K12	SKNREHERCQ
C4	HKTNHSCHEQEHCRVTTT	F3	HKTNHSCQEHCRVTTT	I2	KSEQESYRVS	L1	SSKNREHERC
C5	HKTNHSCHKQEHCRVTTT	F4	HKTNHSCHEHCRVTTT	I3	QKSEQESYRV	L2	HSSKNREHER
C6	HKTNHSCHKREHCRVTTT	F5	HKTNHSCHKHCRVTTT	I4	CQKSEQESYR	L3	NHSSKNREHE
C7	HKTNHSCHKREHCRVTTT	F6	HKTNHSCHKRCRVTTT	I5	SCQKSEQESY	L4	SNHSSKNREH
C8	HKTNHSCHKREQCRVTTT	F7	HKTNHSCHKRERVTTT	I6	SSCQKSEQES	L5	KSNHSSKNRE
C9	HKTNHSCHKREQERVTTT	F8	HKTNHSCHKREQVTTT	I7	KSSCQKSEQE	L6	HKSNHSSKNR
C10	HKTNHSCHKREQEHVTTT	F9	HKTNHSCHKREQETTT	I8	TKSSCQKSEQ		
C11	HKTNHSCHKREQEHCTTT	F10	HKTNHSCHKREQEHTT	I9	KTKSSCQKSE		

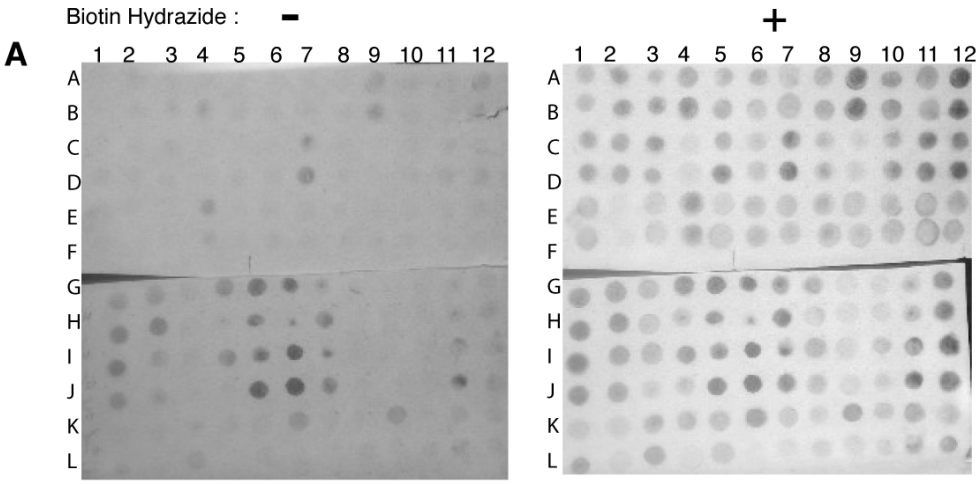
Figure S3. Spot synthesis examining variants of peptides **53**, **61**, and **78**. A) After treatment with biotin hydrazide the sheet was developed by exposure to SAV-AP, followed by washing and then addition of AP substrates. B) The synthesized peptide sequences include all single, double, triple, and quadruple contiguous amino acid deletions of HyRe peptide **53** (A2-F12). Variants of HyRe tags **61** and **78** include single alanine substitutions (G3-H10; and I12-K6) and all contiguous 10-mers (H12-I10; and K8-L6). The 10-mers for both peptides **61** and **78** are analyzed and compared in Supporting Figure S4.



B

Spot	Sequence	Spot	Sequence	Spot	Sequence	Spot	Sequence	Spot	Sequence
A1	HKSNHSSKNR	C9	KSNHSSKRE	F4	KTSCQKSE	H12	TKSSCQKSEQ	K7	~~TICTEQHDHK~~
A2	KSNHSSKNR	C10	KSNHSSKNE	F5	KTKCQKSE	H13	EQEHCRVTTT	K8	HTICTEQHDH
A3	HKSNHSSKN	C11	KSNHSSKNR	F6	KTKSQKSE	I1	REQEHCRVTT	K9	KHTICTEQHD
A4	SNHSSKNR	C12	NHSSKNRE	F7	KTKSSKSE	I2	KREQEHCRVT	K10	CKHTICTEQH
A5	KSNHSSKN	C13	KHSSKNRE	F8	KTKSSCSE	I3	HKREQEHCRV	K11	QCKHTICTEQ
A6	HKSSSKNR	D1	KSSSKNRE	F9	KTKSSCQE	I4	CHKREQEHCR	K12	HQCKHTICTE
A7	NHNHSSK	D2	KSNSKNRE	F10	KTKSSCQK	I5	SCHKREQEHC	K13	HHQCKHTICT
A8	SNHSSKN	D3	KSNHKNRE	F11	KTKSSCQKSE	I6	HSCHKREQEH	L1	THHQCKHTIC
A9	KSNHSSK	D4	KSNHSNRE	F12	TKSSCQKSEQ	I7	NHSCHKREQE	L2	STHHQCKHTI
A10	HKSNHSS	D5	KSNHSSRE	F13	KSSCQKSEQ	I8	TNHSCHKREQ	L3	KSTHHQCKHT
A11	HSSKNR	D6	KSNHSSKE	G1	TKSSCQKSE	I9	KTNHSCHKRE	L4	DKSTHHQCKH
A12	NHSSKN	D7	KSNHSSKN	G2	SSCQKSEQ	I10	HKTNHSCHKR	L5	REPTCQRTNS
A13	SNHSSK	D8	HSSKNRE	G3	KSSCQKSE	I11	PTRPCTTAHD	L6	PREPTCQRTN
B1	KSNHSS	D9	KSSKNRE	G4	TKSSCQKS	I12	HPTRPCTTAH	L7	KPREPTCQRT
B2	HKSNHS	D10	KSSKNRE	G5	SCQKSEQ	I13	HHPTRPCTTA	L8	GKPREPTCQR
B3	HKSNHSSKNR	D11	KSNKNRE	G6	SSCQKSE	J1	SHHPTRPCTT	L9	CGKPREPTCQ
B4	KSNHSSKNRE	D12	KSNHNRE	G7	KSSCQKS	J2	CSHHPTRPCT	L10	KCGKPREPTC
B5	KSNHSSKNRA	D13	KSNHSRE	G8	TKSSCQK	J3	ECSHHPTRPC	L11	HKCGKPREPT
B6	KSNHSSKNAE	E1	KSNHSSE	G9	CQKSEQ	J4	NECSHHPTRP	L12	HHKCGKPREP
B7	KSNHSSKARE	E2	KSNHSSK	G10	SCQKSE	J5	KNECSHHPTR	L13	DHHKCGKPRE
B8	KSNHSSANRE	E3	KSNHSSKNRE	G11	SSCQKS	J6	TKNECSHHPT	M1	SDHHKCGKPR
B9	KSNHSAKNRE	E4	NKTKSSCQKS	G12	KSSCQK	J7	KTKNECSHHP	M2	HSDHHKCGKP
B10	KSNHASKNRE	E5	KTKSSCQKSE	G13	TKSSCQ	J8	HKTKNECSHH	M3	HKTNHSCHKR
B11	KSNASSKNRE	E6	TKSSCQKSE	H1	TKSSCQKSEQ	J9	TTCQTEKTAQ	M4	HKTKNECSHH
B12	KSAHSSKNRE	E7	KKSSCQKSE	H2	TKSSCQKSEA	J10	TTTCQTEKTA	M5	HKNTPAKCPT
B13	KANHSSKNRE	E8	KTSSCQKSE	H3	TKSSCQKSAQ	J11	PTTTCQTEKT	M6	DKSTHHQCKH
C1	ASNHSSKNRE	E9	KTKSCQKSE	H4	TKSSCQKAEQ	J12	CPTTTCQTEK	M7	~~HSDHHKCGKP~~
C2	KSNHSSKNRE	E10	KTKSSQKSE	H5	TKSSCQASEQ	J13	KCPTTTCQTE	M8	HKTNHSCHKR
C3	SNHSSKNRE	E11	KTKSSCKSE	H6	TKSSCAKSEQ	K1	AKCPTTTCQT	M9	TKSSCQKSEQ
C4	KNHSSKNRE	E12	KTKSSCQSE	H7	TKSSAQKSEQ	K2	PAKCPTTTCQ	M10	NKTKSSCQKS
C5	KSHSSKNRE	E13	KTKSSCQKE	H8	TKSACQKSEQ	K3	TPAKCPTTTC	M11	KTKSSCQKSE
C6	KSNSSKNRE	F1	KTKSSCQKS	H9	TKASCQKSEQ	K4	NTPAKCPTTT	M12	KSNHSSKNRE
C7	KSNHSKNRE	F2	KSSCQKSE	H10	TASSCQKSEQ	K5	KNTPAKCPTT	M13	HKSNHSSKNR
C8	KSNHSSNRE	F3	KSSCQKSE	H11	AKSSCQKSEQ	K6	HKNTPAKCPT		

Figure S4. Spot synthesis peptide sheet 2. A) Both before and after treatment with biotin hydrazide (indicated with a – or +), the sheet was developed by exposure to SAV-AP, followed by washing and then addition of AP substrates. B) The synthesized peptide sequences include contiguous 9, 8, 7, and 6-mers; single alanine substitutions, single, double and triple amino acid deletions of peptide **80** and **91**. Peptide variants of **80** reveal background binding or no activity. Some of the **91** variants reveal specificity for binding to biotin hydrazide, especially the 8-mer **102** (F10). Additionally, the contiguous 10-mers of original selectants **25**, **53**, **49**, **46** and **50** were also synthesized and screened (H13-M2). Scored through sequences were incorrectly synthesized.



B

A1	HKTKNECSHHPTRPCTTAHD	D1	HKTNHSCHKREQEHCRVTTT	G1	SDHHKCGKPR	J1	SDHHKSGKPR
A2	HKTNHSCHKREQEHCRVTTT	D2	HKANHSCHKREQEHCRVAAA	G2	GKPREPTCQR	J2	GKPREPTSQR
A3	KQTDTRCQSQLHKHCKIVEN	D3	HKTNHSCHKREQEHCRATTT	G3	KNECSHHPTR	J3	KNESSHHPTR
A4	GAKDNKHCHHQECTHSKTRT	D4	HKTNHSCHKAEQEHCAVTTT	G4	HKTNHSCHKR	J4	HKTNHSSHKR
A5	HSDHHKCGKPREPTCQRTNS	D5	HKTNHSAHKREQEHARVTTT	G5	NKTKSSCQKS	J5	NKTKSSSQKS
A6	GEHHKEHCKPRQCTTATPQA	D6	AKTNASCAKREQEACRVTTT	G6	NHTQRSCKHH	J6	NHTQRSSKHH
A7	DKSTHHQCKHTICTEQHDHK	D7	HKTNHSCHKRAQAHCRVTTT	G7	HKSNHSSKNR	J7	HKSNHSSKNR
A8	HKNTPAKCPTTTCQTEKTAQ	D8	HKTNHSCHKREAEHCRVTTT	G8	KSSKNRE	J8	KSSKNRE
A9	GAKDNKHCHHQECTHSKTRT	D9	HATNHSCHAREQEHCRVTTT	G9	HSDHHKCGKP	J9	HSDHHKSGKP
A10	GEHHKEHCKPRQCTTATPQA	D10	HKTNHACHKREQEHCRVTTT	G10	HPTRPCTTAH	J10	HPTRPSTTAH
A11	HKQETPHCTSHHCTQPKTPA	D11	HKTAHSCHKREQEHCRVTTT	G11	TNHSCHKREQ	J11	TNHSSHKREQ
A12	HQKADHTCAKHECQKKPHTR	D12	HKTNHSCHKREQEHCRVTTT	G12	KTKSSCQK	J12	KTKSSSQK
B1	HKTKNECSHHPTRPCTTAHD	E1	DKSTHHQCKHTICTEQHDHK	H1	SDHHKCGKPR	K1	GEHHKEHCKP
B2	HKTNHSCHKREQEHCRVTTT	E2	DASTHHQCAHTICTEQHDHA	H2	GKPREPTCQR	K2	RQCTTATPQA
B3	KQTDTRCQSQLHKHCKIVEN	E3	DKSTHHQCKHTICTEQADAK	H3	KNECSHHPTR	K3	PRQCTTATPQ
B4	GAKDNKHCHHQECTHSKTRT	E4	AKSTHHQCKHTICTEQHAHK	H4	HKTNHSCHKR	K4	KPRQCTTATP
B5	HSDHHKCGKPREPTCQRTNS	E5	DKSTHHACKHTICTEAHDHK	H5	NKTKSSCQKS	K5	CKPRQCTTAT
B6	GEHHKEHCKPRQCTTATPQA	E6	DKSTHHQCKHTICTAQHDHK	H6	NHTQRSCKHH	K6	HCKPRQCTTA
B7	DKSTHHQCKHTICTEQHDHK	E7	DKSAHHQCKHAICAEQHDHK	H7	HKSNHSSKNR	K7	EHCKPRQCTT
B8	HKNTPAKCPTTTCQTEKTAQ	E8	DKSTHHQAKHTIATEQHDHK	H8	KSSKNRE	K8	KEHCKPRQCT
B9	GAKDNKHCHHQECTHSKTRT	E9	DKSTHHQCKHTACTEQHDHK	H9	HSDHHKCGKP	K9	HKEHCKPRQC
B10	GEHHKEHCKPRQCTTATPQA	E10	DKATHHQCKHTICTEQHDHK	H10	HPTRPCTTAH	K10	HHKEHCKPRQ
B11	HKQETPHCTSHHCTQPKTPA	E11	DKSTAAQCKATICTEQHDHK	H11	TNHSCHKREQ	K11	EHHKEHCKPR
B12	HQKADHTCAKHECQKKPHTR	E12	DKSTHHQCKHTICTEQHDHK	H12	KTKSSCQK	K12	GEHHKEHCKP
C1	HKTNHSCHKREQEHCRVTTT	F1	DKSTHHQCKHTICTEQHDHK	I1	SDHHKSGKPR	L1	GAKDNKHCHH
C2	HKANHSCHKREQEHCRVAAA	F2	DASTHHQCAHTICTEQHDHA	I2	GKPREPTSQR	L2	QECTHSKTRT
C3	HKTNHSCHKREQEHCRATTT	F3	DKSTHHQCKHTICTEQADAK	I3	KNESSHHPTR	L3	HQECTHSKTR
C4	HKTNHSCHKAEQEHCAVTTT	F4	AKSTHHQCKHTICTEQHAHK	I4	HKTNHSSHKR	L4	HHQECTHSKT
C5	HKTNHSAHKREQEHARVTTT	F5	DKSTHHACKHTICTEAHDHK	I5	NKTKSSSQKS	L5	CHHQECTHSK
C6	AKTNASCAKREQEACRVTTT	F6	DKSTHHQCKHTICTAQHDHK	I6	NHTQRSSKHH	L6	HCHHQECTHS
C7	HKTNHSCHKRAQAHCRVTTT	F7	DKSAHHQCKHAICAEQHDHK	I7	HKSNHSSKNR	L7	KHCHHQECTH
C8	HKTNHSCHKREAEHCRVTTT	F8	DKSTHHQAKHTIATEQHDHK	I8	KSSKNRE	L8	NKHCHHQECT
C9	HATNHSCHAREQEHCRVTTT	F9	DKSTHHQCKHTACTEQHDHK	I9	HSDHHKSGKP	L9	DNKHCHHQEC
C10	HKTNHACHKREQEHCRVTTT	F10	DKATHHQCKHTICTEQHDHK	I10	HPTRPSTTAH	L10	KDNKHCHHQE
C11	HKTAHSCHKREQEHCRVTTT	F11	DKSTAAQCKATICTEQHDHK	I11	TNHSSHKREQ	L11	AKDNKHCHHQ
C12	HKTNHSCHKREQEHCRVTTT	F12	DKSTHHQCKHTICTEQHDHK	I12	KTKSSSQK	L12	GAKDNKHCHH

Figure S5. Spot synthesis peptide sheet 3. A) Both before and after treatment with biotin hydrazide the sheet was developed by exposure to SAV-AP, followed by washing and then addition of AP substrates. B) The synthesized peptide sequences include full-length 20-mers from the original phage selections and were synthesized in duplicate (A1-B12). Each type of amino acid was substituted with alanine in HyRe peptides **53** and **46,** and synthesized in duplicate (C1-F12). The most interesting short peptides from spot synthesis peptide sheets 1 and 2 were re-synthesized in duplicate to determine background binding (G1-J12). Cysteine was substituted for alanine in the short 10-mers (I1-J12). The contiguous 10-mer variants from original selectants **51** and **52** were also tested (K1-L12).

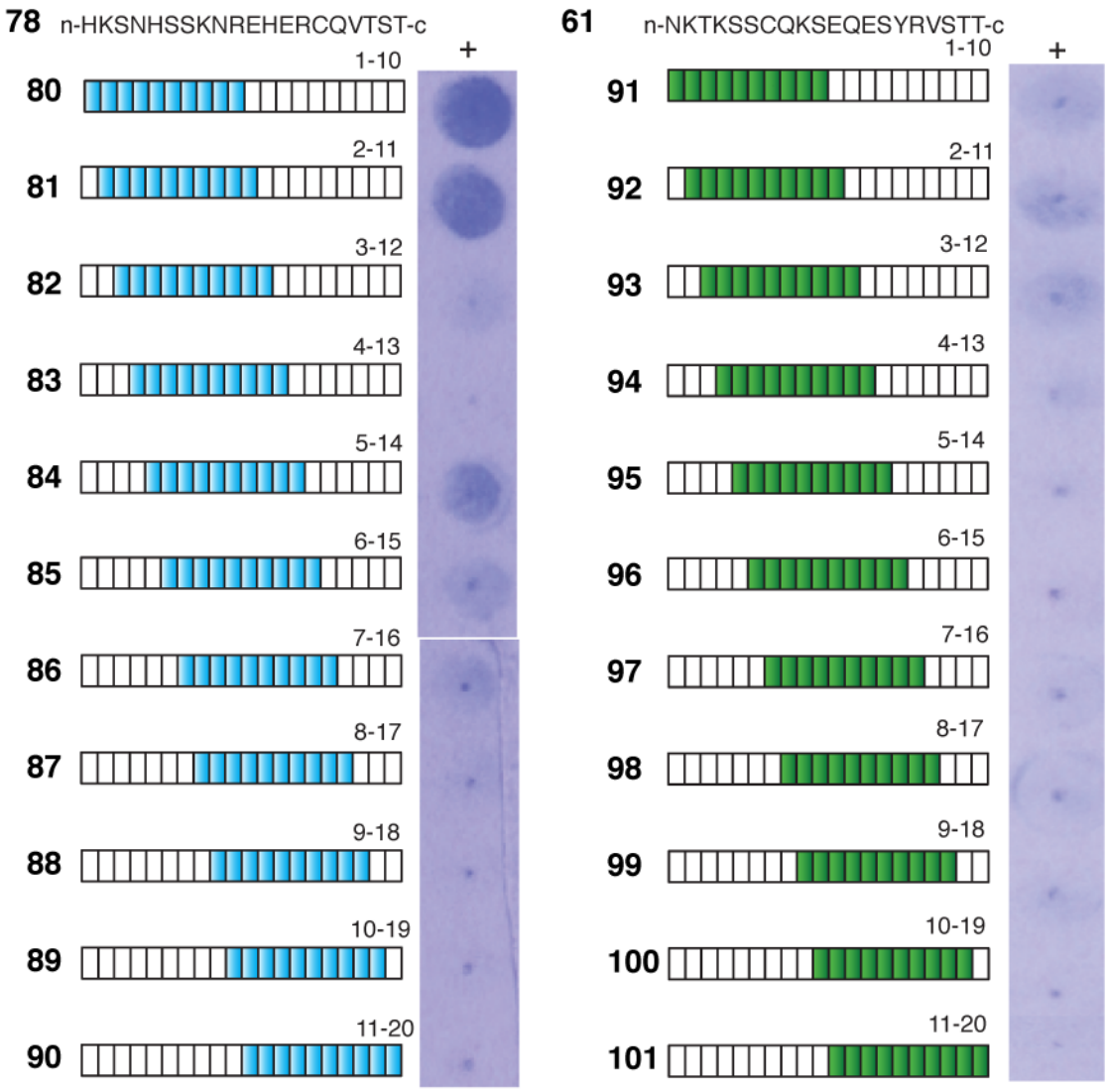


Figure S6. Analysis of contiguous 10-mer peptides **61** and **78** derived from HyRe peptide **53**. As shown by spot synthesis, the N-terminus of both peptides contributes critical residues to the peptide reactivity with biotin hydrazide at pH 7.

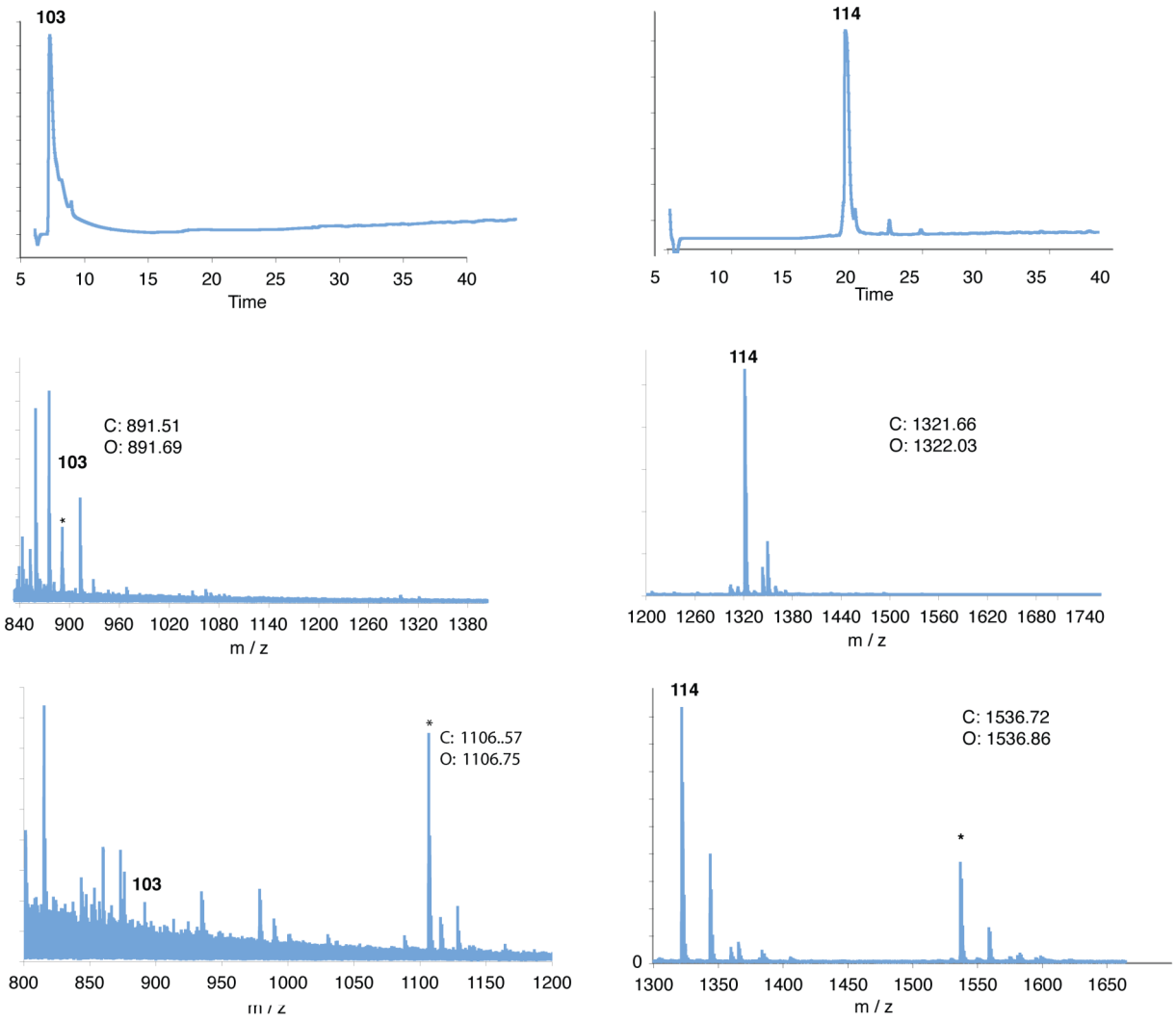


Figure S7. Analytical HPLC and MALDI-TOF MS analysis of synthesized and purified HyRe peptide **114** variants. The chromatograms (top) were monitored at 210 nm. MALDI-TOF MS analysis of purified HyRe tag variants (middle). MALDI-TOF MS analysis of the two peptides treated with AMCA hydrazide, revealing the +215 amu products (bottom). "C:" and "O:" indicate the calculated and observed (*) molecular weights, respectively.

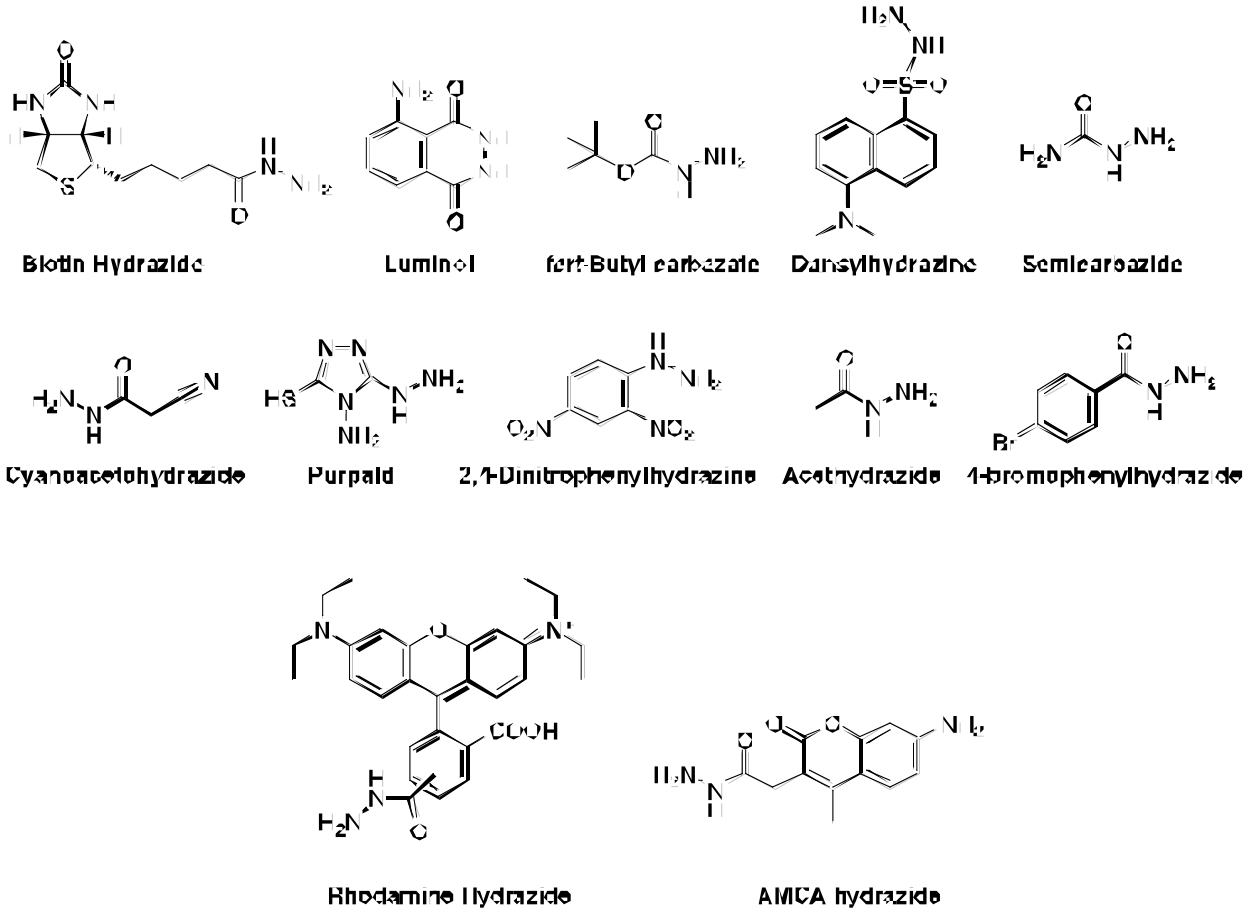


Figure S8. 12 unique hydrazides and hydrazine derivatives used to determine the generality of the reaction with peptide **114**. Only rhodamine B hydrazide and AMCA hydrazide produce the a covalent adduct upon reaction with the peptide at pH 7.

Additional references.

(1) Murase, K. *et al.* (2003) EF-Tu binding peptides identified, dissected, and affinity optimized by phage display. *Chem. Biol. 10*, 161-168.

(2) Sidhu, S.S., Weiss, G.A.. "Constructing phage display libraries by oligonucleotide-directed mutagenesis." In Phage Display. Lowman, H.L., and Clackson,T., Eds. Ch. 2 pg. 27-41 (Oxford University Press, Oxford, 2004).