Supplementary Information

Enabling simultaneous photoluminescence spectroscopy and X-ray Footprinting Mass Spectrometry to study protein conformation and interactions

Sayan Gupta[1], Brandon Russell[1], Line G. Kristensen[1], Shawn Costello[2], Behzad Rad[3], James Tyler[3], Susan Marqusee[2], Corie Y. Ralston[3*]

[1]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[2]Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley California 94720, United States

[3]Molecular Foundry Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Correspondence email: cyralston@lbl.gov

Content of Supplementary Information

Purification of SpyCatcher003

Briefly, BL21(DE3) harboring the pLysS plasmid was transformed with the pDEST14-SpyCatcher003 S49C vector (Addgene). Colonies were picked and grown in LB supplemented with 40 µg/mL Chloramphenicol and 100 µg/mL Ampicillin overnight at 37°C degrees with 250 rpm shaking. The overnight cultures were used to inoculate 0.5 L LB media supplemented with 40 µg/mL Chloramphenicol and 100 µg/mL Ampicillin and grown to ~0.6 OD. The bacterial cells were induced by adding a final concentration of 0.4 mM IPTG to the media and growing the cultures at 30°C. The bacteria cells were harvested after 3 hours by centrifuging the cultures at 6000 rpm in a JLA8.10 fixed angle rotor centrifuge (Beckman Coulter) for 20 minutes at 4°C. Pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM Imidazole, 5% (v/v) glycerol) and lysed by two passages through an Emulsiflex C3 homogenizer at >20,000 PSI(Avestin). The lysate was spun down at 100000 x g in a Ti70 rotor. The clarified lysate was loaded onto a HiTrap FF Ni-NTA column, washed with lysis buffer, followed by lysis buffer containing 25 mM imidazole for 5 column volumes, and then eluted with 250 mM imidazole step gradient. After extensive dialysis to remove imidazole, protein was flash frozen with liquid nitrogen and stored at -80°C.

Interfacing CCD spectrometer, data collection and storage

To interface with the spectrometer, Thorlabs provides dynamic-linked libraries (DLL), Python libraries, a standalone Windows program, and an external trigger. Several of these methods were tested and it was found that the simplest and fastest (in terms of software execution) was to write a separate spectrometer VI that uses Microsoft's .NET framework to access the Thorlabs DLLs. The timing of events in the LabVIEW program during the pump cycle is critical. An asynchronous function call is used to start spectrometer data collection during an X-ray exposure to minimize delays. Using LabVIEW's default sequential behavior would require the program to stop the execution of the main VI to wait for the spectrometer to finish collecting data. An asynchronous function call allows the main VI and the spectrometer VI to execute in parallel. Spectrometer data

is stored and made available to the main VI when requested. The main VI allows for the spectrometer to be used for on-demand data collection or automatically during an X-ray exposure. When on-demand spectrometer data is requested, a single or continuous scanning can be obtained. The integration time can be set between 1 ms and 2 s for on-demand scanning. When running during an X-ray exposure, the integration time must be within that range and shorter than the pump time. It can also be set to occur anytime within the pump cycle. The user- selectable range is automatically adjusted based on the integration time and pump time to ensure the integration time occurs entirely within the pump cycle.

LCMS data collection and analysis

The digested samples were loaded into Thermo Dionex RSLCnano 3000 with a WPS-3000 TBPLRS autosampler at a flow rate of 50 uL/min of buffer A (0.1 % formic acid). Peptides were then eluted with a gradient of 10-65% buffer B (90% Acetonitrile, 0.1% Formic Acid) at a flow rate of 20 μL/min over 10 minutes and then of 65-99% buffer B over 2 min, and then held at 99% buffer B for 3 minutes prior to the execution of a sawtooth washing step and equilibration at 10% buffer B. Protein was eluted directly into a Q Exactive Orbitrap Mass Spectrometer operating in positive mode (resolution 140000, AGC target 1e6, maximum IT 50 ms, scan range 200-2000 m/z). XFMS peptide identification and analysis has been automated and enhanced by adopting the Byos® (Protein Metrics Inc) integrated software platform as previously described[41]. Biologic automatically extracts ion chromatograms and reports the quantification of modifications relative to the unmodified peptide based on the extracted ion chromatograms. A typical workflow starts with processing a high-exposure tandem MS (MS/MS) file in Byos for an MS/MS search against FASTA sequences and the localization of modification sites. The peptide level analysis and validation of assignments are carried out in Byologic and lead to the creation of in-silico peptides in the form of a CSV file using the MS/MS data. The in-silico peptides CSV is subsequently applied to full scan (MS1) data covering a series of exposure times, and the resulting quantified peptide modifications provide the basis for the residue-specific and peptide-level dose response. The

abundance (peak area) of the identified unmodified and modified peptides at each irradiation time point area were measured from their respective extracted ion chromatogram of the mass spectrometry data collected in the precursor ion mode. Representative extracted ion chromatograms and MS/MS spectra are shown in **Figures S3** and **S4**.

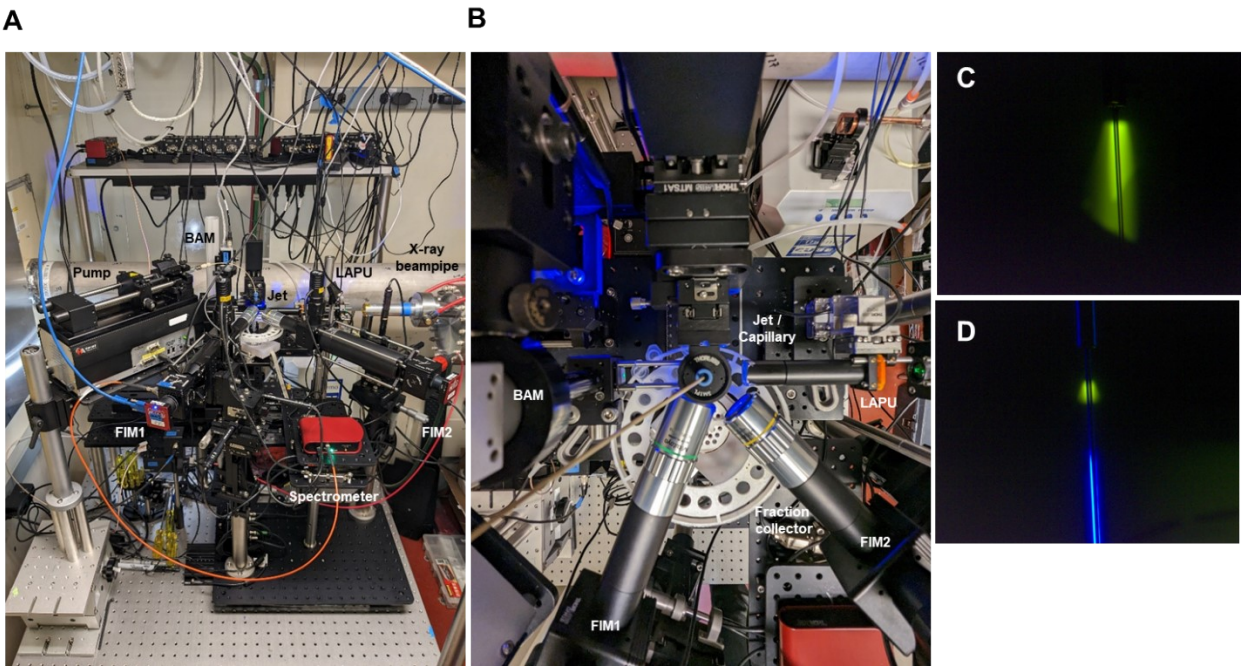

Figure S1. **Instrument setup and alignment.** (A) Side view showing location of X-ray beampipe and exit window, jet, spectrometer, syringe pump, beam alignment module (BAM), laser assisted pre-alignment unit (LAPU), interchangeable fluorescent imaging modules (FIM1 and FIM2) for spectroscopy and PMT based automated Alexa dose-response analysis. The fraction collector is located below the jet. (B) Top view of the instrument assembly. (C) GigE camera view of the 75 μm liquid jet and the X-ray beam profile on the Nd-Yag fluorescence screen on the BAM. (D) GigE camera view of the liquid jet and the X-ray beam profile on the Nd-Yag fluorescence screen used for 75 μm liquid jet sample exposure. The blue region shows the illuminated portion of the liquid jet by ~ 480 nm excitation light to measure post-exposure fluorescence based Alexa dose-response analysis.

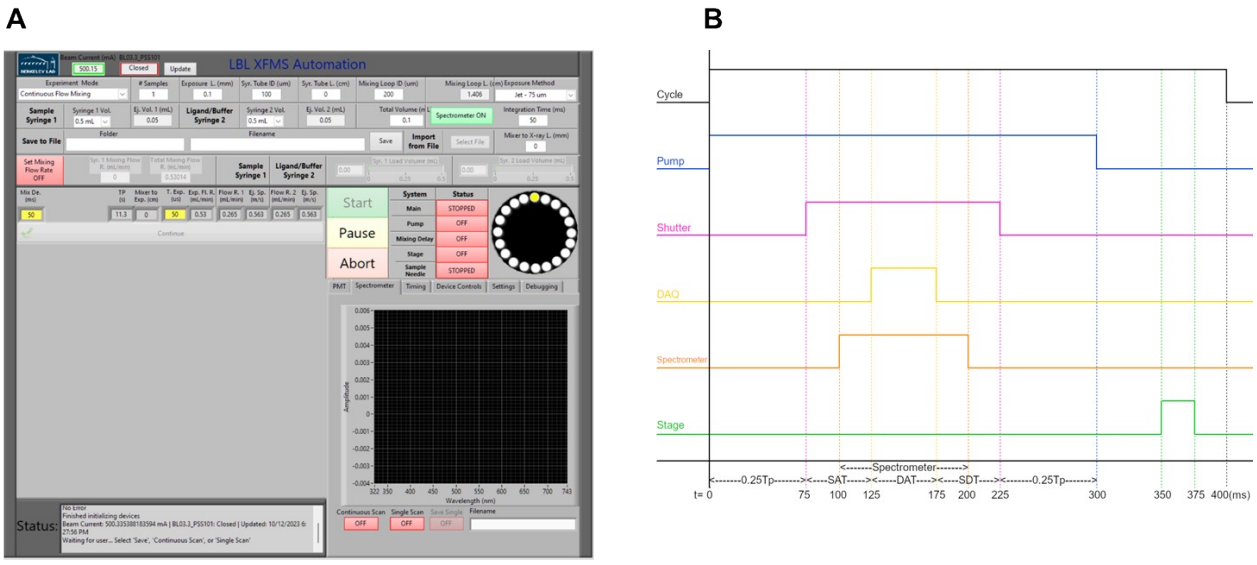

Figure S2. **Experimental control panel and timing of data collection**. (A) Screen shot of the GUI experimental control panel. The configuration includes user input for the number of samples, target volumes, sample flow parameters, and syringe parameters. The control panel has three modes of operation – continuous mixing, stopped-flow mixing and steady-state flow for spectroscopy-XFMS data collection. Only steady-state flow for spectroscopy-XFMS data collection mode is used for the current study (B) Instrumentation timing diagram with the PMT and DAQ based and CCD – spectrometer based data collection are overlaid for visualization purposes. The overall minimum pumping time is 400 ms. The shutter activation time (SAT), data acquisition time (DAT) and shutter deactivation time (SDT) for PMT based data collection are indicated on the X-axis. The shutter operation is not involved in CCD-spectrometer based data collection. The integration time of 100 ms shown in the picture is tunable by the length as well as the position relative to the overall sample pumping time.

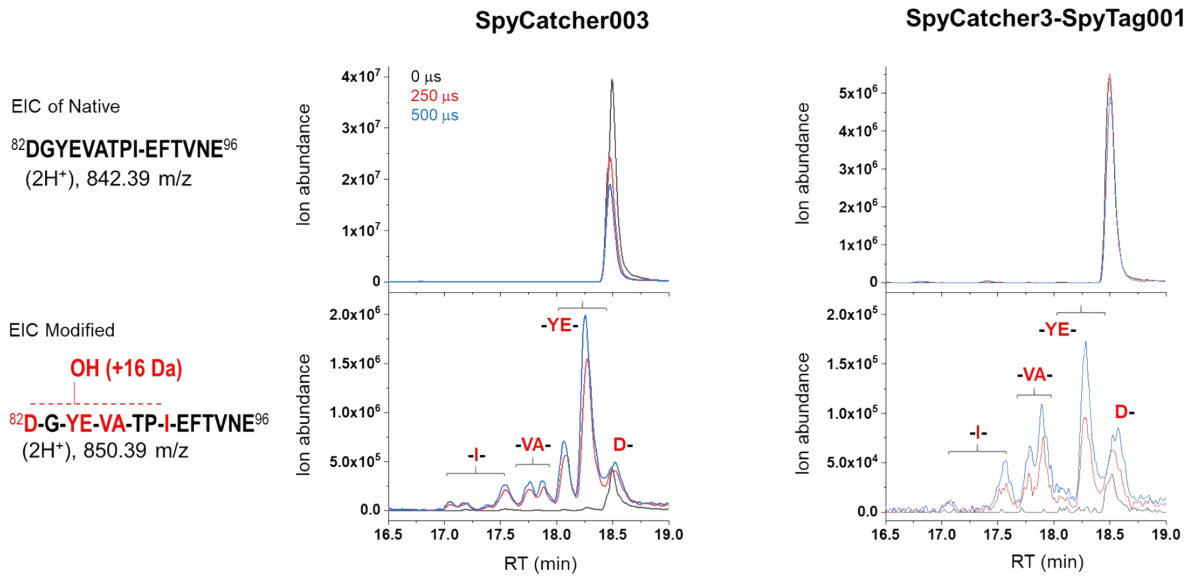

Figure S3. **Representative extracted ion chromatogram (EIC)**. The center and the right panels show the EIC of a doubly protonated native (top) and modified (bottom) SpyCatcher003 tryptic peptide at various X-ray exposure times. Modified products eluting at various retention times are assigned to sites as determined by tandem MS or MS/MS spectral analysis. Spectra clearly show the peak intensity for modified residues Y or E is significantly reduced for SpyCatcher003-SpyTag001 complex.

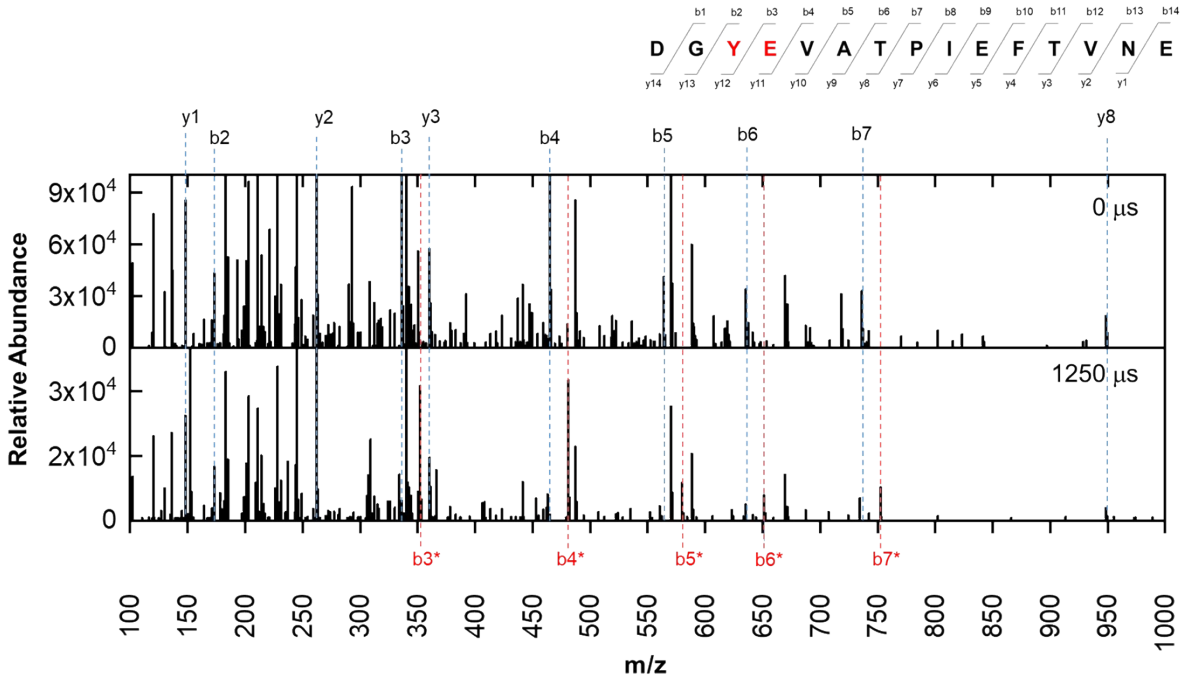

Figure S4. Representative MS/MS spectra. The MS/MS spectra of unmodified and modified peptide fragment ^{82}DGYEVATPIEFTVNE^{96} at 0 and 1250 μs, respectively Presence of modified fragment ions indicated in red determine modification at Y or E.

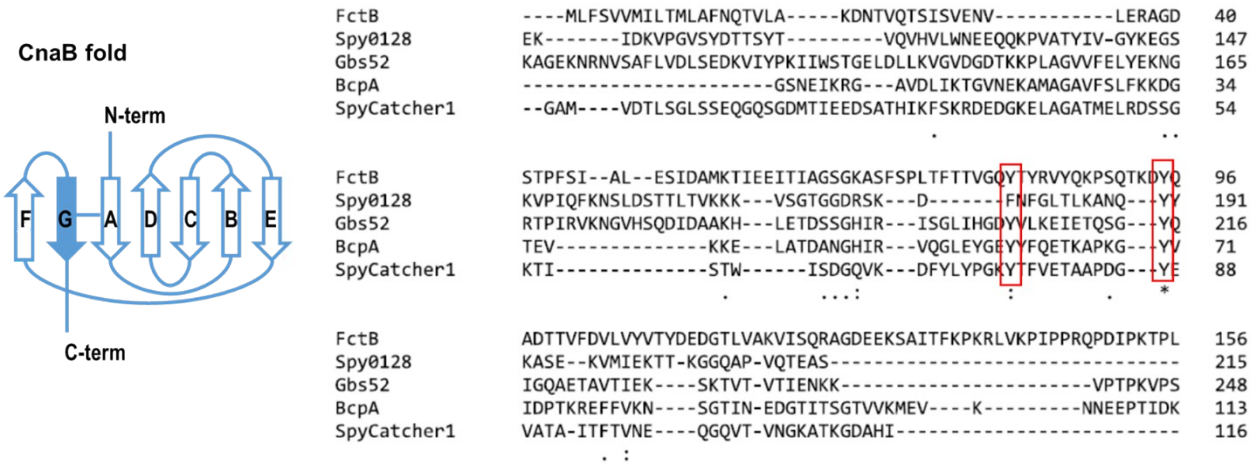

Figure S5. CnaB fold sequence alignment. The sequence of selected crystallographically resolved proteins with a CnaB fold are aligned using Clustal Omega (EMBL-EBI). The proteins selected are: *S. pyogenes* basal pilin FctB (3KLQ); *S. pyogenes* major pilin Spy0128 (3B2M); *S. agalactia* minor pilin Gbs52 (3PHS); *Bacillus cereus* major pilin BcpA (3KPT); and SpyCatcher001 (PDB ID 4MLI).

Supplementary Table S1: Typical pump time for XFMS instrument operation

Size and type of samples	X-ray path length (μm)	Exposure time (μs)	Sample volume (μl)	Syringe flow rate (ml/min)	Pump time (s)
Sample set using 75 μm liquid jet	200	10	50	5.30	0.566
		15	50	3.53	0.849
		20	50	2.65	1.130
		30	50	1.77	1.700
Sample set using 100 μm liquid jet or capillary	200	20	50	4.71	0.637
		30	50	3.14	0.955
		40	50	2.36	1.270
		60	50	1.57	1.910
Sample set using 100 μm liquid jet or capillary	650	250	50	4.90	0.612
		500	50	2.45	1.220
		750	50	1.63	1.840
		1000	50	1.23	2.450

Supplementary Table S2: Rate constants of hydroxyl radical modification

Sequence of the peptide fragments[a]	Sites of modification[b]	Type of modification' (Da)[c]	Hydroxyl radical reactivity rate k (s^{-1})[d]		R^e	$R_{max}{}^e$	$R_{min}{}^e$	$R_{av}{}^e$	Accessible solvent area (ASA ($Å^2$)[f]	
			SpyCatcher003	Complex					Spy Catcher-003	Complex
-3GAMVTTLSGLSGEQGPSG1	M-1	+16	192.16 ± 10.81	213.35 ± 8.41	0.90	0.99	0.82	0.90 ± 0.07	-	-
	L4	+16	82.90 ± 5.06	92.32 ± 3.67	0.90	0.99	0.81	0.90 ± 0.07	-	-
	L7	+16	37.94 ± 0.93	39.12 ± 0.47	0.97	1.01	0.93	0.97 ± 0.03	-	-
	Q11	+16	55.68 ± 2.14	46.83 ± 1.29	1.19	1.27	1.11	1.19 ± 0.06	-	-
	P13	+16	32.23 ± 1.55	24.62 ± 0.65	1.31	1.41	1.21	1.31 ± 0.08	-	-
16DMTTEE21	M17	+16	505.25 ± 9.16	503.22 ± 22.84	1.00	1.07	0.94	1.01 ± 0.05	-	-
22DSATHIKFSKR32	H26	+16	81.65 ± 6.04	57.61 ± 2.59	1.42	1.59	1.26	1.42 ± 0.14	64	64
	K31	+16	54.09 ± 3.86	23.98 ± 1.06	2.26	2.53	2.01	2.26 ± 0.21	24	2
33DEDGRELAGATM44	D33	+16	31.65 ± 1.28	8.40 ± 0.58	3.77	4.21	3.38	3.79 ± 0.34	17	2
	R37	+14	53.13 ± 1.14	35.23 ± 1.74	1.51	1.62	1.41	1.51 ± 0.09	134	134
	M44	+16	148.97 ± 2.09	20.59 ± 0.84	7.24	7.65	6.86	7.25 ± 0.32	5	0
35DGRELAGATM44	E38	-30	4.02 ± 0.16	10.80 ± 0.13	0.37	0.39	0.35	0.37 ± 0.02	34	34
	A42, T43	+14	3.78 ± 0.11	1.25 ± 0.10	3.03	3.40	2.71	3.05 ± 0.28	0,0	3,3
53TISTWISDGHVK64	W57	+16	190.43 ± 16.07	258.27 ± 26.76	0.74	0.89	0.61	0.75 ± 0.11	13	13
	H62	+16	216.69 ± 19.33	343.59 ± 28.78	0.63	0.75	0.53	0.64 ± 0.09	116	116
65DFYLYPGKYTFVETAAP81	Y67, Y69 mixed	+16	622.61 ± 52.63	647.57 ± 92.12	0.96	1.22	0.77	0.98 ± 0.18	116, 136	116, 136
82DGYEVATPIEFTVNE96	D82	+16	49.81 ± 1.19	34.06 ± 0.83	1.46	1.53	1.39	1.46 ± 0.06	98	98
	Y84, E85	+16	288.00 ± 10.19	85.35 ± 1.85	3.37	3.57	3.19	3.38 ± 0.16	84, 148	35, 72
	V86, A87	+16	58.92 ± 2.01	62.79 ± 1.21	0.94	0.99	0.89	0.94 ± 0.04	57, 53	57, 19
	I90	+16	37.86 ± 4.91	25.13 ± 0.27	1.51	1.72	1.30	1.51 ± 0.17	78	29
97DGQVTV102	D97, Q99	+14	20.23 ± 1.36	19.10 ± 0.82	1.06	1.18	0.95	1.06 ± 0.10	112	112
97DGQVTVDGEATEG109	V100, T101	+16	91.39 ± 3.01	42.50 ± 0.46	2.15	2.25	2.06	2.15 ± 0.08	35, 87	6,87
	Resi 103-108	+16	108.55 ± 2.70	74.66 ± 2.19	1.45	1.54	1.38	1.46 ± 0.06	85	36

[a] sequences of AspN and tryptic fragments used for identification of quantification of modification sites

[b] modified residues, which were identified and confirmed by LCMS/MS (Figure S4)

[c] type of side chain modification includes hydroxylation, carbonylation and decarboxylation, which resulted in a mass shift of +16, +14 and -30 Da

[d] hydroxyl radical rate constants were estimated by employing a first-order exponential fit of the dose response plot of overall peptide modification as described in experimental procedures and Figure 53C. The modified peptide fragments were eluted as a single peak or multiple peaks Figure S3. The modified peak areas were extracted individually but summed together to calculate the total modification of the respective peptide.

[e] ratio of hydroxyl radical reactivity obtained by dividing hydroxyl radical reactivity rate of free SpyCatcher003 by that of for the complex. The ratio represents fold decrease (>1) or increase (<1) in the solvent accessibility of the modified residues.

[f] the accessible solvent area was determined from the PDB 4mli with and without SpyTag001 using GETAREA (https://curie.utmb.edu/getarea.html). The ASA was not reported for regions which are absence from in the crystal structure. There is no data on Q99, and residues 104-108.