Engineering Bromodomains with a Photoactive Amino Acid by Engaging 'Privileged' tRNA Synthetases

Shana Wagner,[a] Babu Sudhamalla,[a,b] Philip Mannes,[a] Sushma Sappa,[a] Sam Kavoosi,[a] Debasis Dey,[a] Sinan Wang,[a] Kabirul Islam*, [a]

[a]Department of Chemistry, University of Pittsburgh, Pittsburgh PA 15260

[b]Department of Biological Sciences, Indian Institute of Science Education and Research-Kolkata, Mohanpur, India 741246

Table of Contents

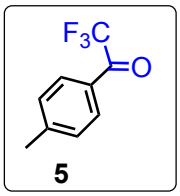
1. General materials, methods and equipment

Chemicals: All chemicals were purchased from established vendors (e.g. Sigma-Aldrich, Acros Organics) and used without purification unless otherwise noted. Optima grade acetonitrile was obtained from Fisher Scientific and degassed under vacuum prior to HPLC purification. All reactions to prepare amino acid analogues were carried out in round bottom flasks and stirred with Teflon®-coated magnetic stir bars under inert atmosphere when needed. Analytical thin layer chromatography (TLC) was performed using EMD 250 micron flexible aluminum backed, UV F_{254} pre-coated silica gel plates and visualized under UV light (254 nm) or by staining with phosphomolybdic acid, ninhydrin or anisaldehyde. Reaction solvents were removed by a Büchi rotary evaporator equipped with a dry ice-acetone condenser. Analytic and preparative HPLC was carried out on an Agilent 1220 Infinity HPLC with diode array detector. Concentration and lyophilization of aqueous samples were performed using Savant Sc210A SpeedVac Concentrator (Thermo), followed by Labconco Freeze-Dryer system. Proton nuclear magnetic resonance spectra (1H NMR) were recorded on Bruker Ultrashield^{TM} Plus 600/500/400/300 MHz instruments at 24^{o}C. Chemical shifts of ^{1}H and ^{13}C NMR spectra are reported as δ in units of parts per million (ppm) relative to tetramethylsilane (δ 0.0) or residual solvent signals: chloroform-d (δ 7.26, singlet), methanol-d_4 (δ 3.30, quintet), and deuterium oxide-d_2 (δ 4.80, singlet). Coupling constants are expressed in Hz. MALDI mass spectra were collected at ultraFlextreme (Bruker) and the data was analyzed using flexAnalysis software. The ESI-MS were recorded on a Q-Exactive^{TM} Thermo Scientific LC-MS with electron spray ionization (ESI) probe.

Plasmids, mutagenic primers, cell lines and antibodies: All the plasmids for bacterial or mammalian expression are obtained as gifts from individual laboratories or purchased from Addgene. Details of these constructs are given in Table S1. Mutagenic primers are obtained from Integrated DNA Technologies (Table S2). Commercially available competent bacterial cells were used for protein expression and mutagenesis. HEK293T cells were purchased from American Type Culture Collection (ATCC) and used in the current study following manufacturer's protocol. All the antibodies used in the current study are purchased from established vendors and used following manufacturer's protocol.

2. Synthesis of 4-(trifluorodiazirinyl) phenylalanine (tmdF) 3

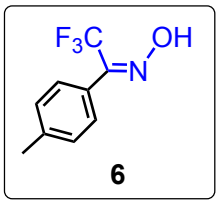
Methods for the synthesis and analytical data for characterization are provided below. ^{1}H, ^{13}C and ^{19}F NMR spectra are given in Supplementary Figures S10-S22.[1-5]



5

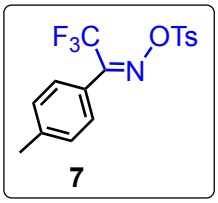
To a stirred and cold solution of toluene **4** (17 ml, 1 eq.) and DMAP (20g, 1 eq.) in 200 ml DCM was added trifluoroacetic anhydride (23ml, 1.05eq) in drop wise fashion. After 5 min $AlCl_3$ (50g, 2.3 eq.) added in 5 portions over 20 min. The reaction was stirred for 12 h and then poured over ice water (300g) slowly and then extracted with DCM (2x100ml), concentrated (pressure 200mbar, 40 °C water bath) and then distilled (bath temp 57 °C, vacuum: 1mbar) to obtain **5** (16 g, 53%, 91% in Ref. 5) as a colorless liquid.[5]

^{1}H NMR ($CDCl_3$, 400MHz): δ = 7.99 (d, *J*=7.9 Hz, 2 H), 7.36 (d, *J*=8.2 Hz, 2 H), 2.47 (s, 3 H) ppm.



6

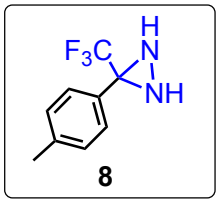
5 (2.60 g, 13.80 mmol) was dissolved in pyridine (30 mL), hydroxylamine hydrochloride (2.88 g, 41.44 mmol) was added, and the reaction was stirred at 70 °C for 3 h. Pyridine was removed in vacuo, and the remaining residue was dissolved in Et_2O (50 mL) and washed with 0.01 M HCl (50 mL). The organic layer was washed with H_2O (3 x 50 mL) and dried (Na_2SO_4), and the solvent was removed *in vacuo* to obtain **6** (2.01 g, 72%) as a colorless solid, which was used without further purification.[5] *This reaction can be scaled up to 15 g without any detectable loss of yield.*



7

To a stirred solution of oxime **6** (2.00 g, 9.85 mmol) in pyridine (36 mL), *p*-toluenesulfonyl chloride (2.82 g, 14.79 mmol) was added in portions and the reaction mixture was refluxed for 3 h. Pyridine was removed in vacuo , and the residue was purified by flash column chromatography eluting with hexane/CH_2Cl_2 (2:1) to get **7** (2.90 g, 82%) as a white solid.[5]

^{1}H NMR ($CDCl_3$, 300MHz): δ = 7.83 - 7.97 (m, *J*=8.3 Hz, 2 H), 7.37 - 7.48 (m, *J*=8.1 Hz, 2 H), 7.24 - 7.37 (m, 6 H), 2.51 (s, 3 H), 2.43 (s, 3 H) ppm.

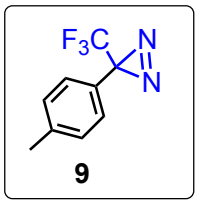


8

In a thick-walled screw-cap tube, **7** (2.90 g, 8.12 mmol) was dissolved in Et_2O (30 mL). The solution was cooled to -78 °C, and liquid ammonia (5 mL) was

added. The tube was screwed tightly, and the solution allowed rising to room temperature. The reaction was stirred for 12 h, and the solution was cooled to -78 °C. The cap was removed, and the mixture risen to room temperature to remove ammonia. The solution was partitioned between Et_2O (50 mL) and H_2O (50 mL), the organic layer was dried (Na_2SO_4), and the solvent was removed under reduced pressure to get **8** (1.25g, >98%.) as colorless liquid.[5]

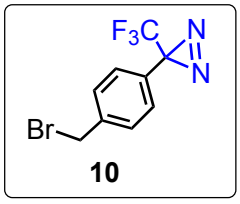
^{1}H NMR ($CDCl_3$, 300MHz): δ = 7.47 - 7.54 (m, *J*=8.1 Hz, 2 H), 7.21 - 7.26 (m, *J*=7.7 Hz, 2 H), 2.77 (d, *J*=8.1 Hz, 1 H), 2.39 (s, 3 H), 2.20 (d, *J*=9.3 Hz, 1 H) ppm.



9

Diaziridine **8** (4.23 g, 1 eq, 20.9 mmol) and PDC (11.0 g, 1.4 eq, 29.3 mmol) in DCM (30 mL) was stirred at room temperature for 5 hours. The reaction was diluted in pentane and purified with a 100% pentane silica column to give **9** (3.6 g, 82%) as a clear oil. In Ref. 5, the yield of compound **9** has been reported to be 57% using iodine as oxidant. Note that the water bath of the rotary evaporator must be set below 20 °C to avoid evaporation of product, which is volatile. The product chilled in an ice bath can placed on the strong vacuum to remove residual solvent.[5]

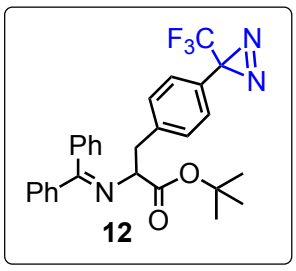
^{1}H NMR ($CDCl_3$, 300MHz): δ = 7.21 (d, *J*=8.1 Hz, 2 H), 7.10 (d, *J*=8.1 Hz, 2 H), 2.37 (s, 3 H) ppm.



10

To a solution of 3-(p-tolyl)-3-(trifluoromethyl)-3H-diazirine **9** (2.6 g, 13 mmol) in 20 ml benzene was added NBS (3.47 g, 19.5 mmol) and AIBN (261 mg, 1.3 mmol). The mixture was stirred at 70 °C for 12 h. After the reaction is completed, the precipitate was filtered, the solvent was removed under reduced pressure at 20 °C, and the crude product was purified by chromatography (Hexane/EA = 50/1) to give 3.1 g compound **10** as colorless oil (yield = 86 %; 69% in Ref. 1, 88% in Ref. 5).[1,5]

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.43 (d, *J* = 8.5 Hz, 2H), 7.17 (d, *J* = 8.1 Hz, 2H), 4.47 (s, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 139.59, 129.59, 129.41, 127.07, 127.05, 122.17 (q, $^1J_{CF}$ = 274.7 Hz), 32.13, 28.46 (q, $^2J_{CF}$ = 40.74); ^{19}F NMR (282 MHz, $CDCl_3$) δ -65.17.



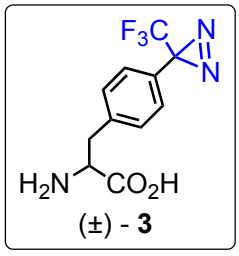
12

To a stirring solution of N-(diphenylmethylene) glycine t-butyl ester **11** (562 mg, 1.0 eq, 1.9 mmol) in THF (8 mL) at -78 °C was added LDA (1.2 mL, 1.25 eq, 2.4 mmol). After stirring 1 hr, **10** (534 mg, 1.0 eq, 1.9

mmol) in THF (2 mL) was added. The ice bath was removed after 2 hr. and the reaction warmed to room temperature overnight. The addition of water (1 mL) quenched the reaction, and the reaction was subsequently diluted with ethyl acetate. The organic layer was washed with brine, dried over sodium sulfate, and purified with dry-pack silica column chromatography (40:1 Hex/EtOAc) to give the product **14** (630 mg, 67%; 60-68% in Ref. 3) as a pale yellow solid.[3]

Alternatively, to a solution of **10** (1.39 g, 5 mmol) in 30 ml DCM was added N-(diphenylmethylene) glycine t-butyl ester **11** (1.62 g, 5.5 mmol), KOH (308 mg, 5.5 mmol) and Bu_4NI (185 mg, 0.5 mmol). The mixture was stirred at room temperature for 24 hr. After the reaction is completed, the mixture was concentrated and purified by chromatography (Hexane/EA = 50/1 – 8/1) to give 1.4 g compound **12** (yield = 58 %). Melting point=88.5-91.8 ^oC.

^1H NMR (400 MHz, $CDCl_3$) δ 7.37-7.87 (m, 4H), 7.16-7.36 (m, 4H), 6.94-7.04 (m, 4H), 6.37-6.87 (m, 2H), 4.02 (dd, J = 4.2, 9.0 Hz, 1H), 3.05-3.19 (m, 2H), 1.36 (s, 9H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 196.90, 170.70, 170.60, 140.66, 139.4, 137.75, 136.24, 132.55, 130.48, 130.43, 130.20, 130.03, 128.85, 128.51, 128.42, 128.27, 128.16, 127.65, 127.05, 126.41, 122.45 (q, $^1J_{CF}$ = 284.2 Hz), 81.54, 67.67, 39.29, 28.48 (q, $^2J_{CF}$ = 40.32 Hz) 28.16. ^{19}F NMR (376 MHz, $CDCl_3$) δ -65.39; HRMS (ESI) calcd for $C_{28}H_{27}O_2N_3F_3$ $[M+H]^+$: 494.20554; found: 494.20798.



(±) - **3**

Compound **12** (1.4 g, 5 mmol) was dissolved in TFA and stirred at rt for 2 h. After evaporation, the residue was partitioned between Et_2O and 1N HCl and extract with 1N HCl for 3 times. The aqueous layers were combined and lyophilized to give 600 mg final compound **3** as white solid (yield = 68 %; 91-95% In Ref. 3).[3]

^1H NMR (500 MHz, 1:1 D_2O/DMSO-D_6 + 1 μL 10 M NaOH) δ 7.46 (d, J = 8.0 Hz, 2H), 7.34 (d, J = 7.9 Hz, 2H), 3.45-3.60 (m, 1H) 3.11 (dd, J = 4.9, 13.4 Hz, 1H), 2.90 (dd, J = 7.7, 13.3 Hz, 1H); ^{13}C NMR (126 MHz, 1:1 D_2O/DMSO-D_6 + 1 μL 10 M NaOH) δ 182.26, 142.31, 131.47, 127.89, 127.75, 123.40 (q, $^1J_{CF}$ = 274.6 Hz, CF_3), 58.51, 41.95 29.70 (q, $^2J_{CF}$ = 40.1 Hz); ^{19}F NMR (471 MHz, 1:1 D_2O/DMSO-D_6 + 1 μL 10 M NaOH) δ -64.77; HRMS (ESI) calcd for $C_{11}H_{11}O_2N_{11}F_3$ $[M+H]^+$: 274.08034; found: 274.08146.

3. Peptide synthesis and purification

All peptides were synthesized by the University of Pittsburgh Peptide Synthesis Facility; Crude peptides were purified using preparative reversed-phase HPLC (XBridge C18, 5 µm, 10 x 250 mm column) eluting with a flow rate of 5 mL/min and a gradient of acetonitrile starting from 0% v/v to 90% v/v in 15 min and then to 100% v/v in 18 min in aqueous trifluoroacetic acid (0.1% v/v). Purified peptides were first concentrated by SpeedVac concentrator followed by lyophilization. Dried peptides were resuspended in water and stored at -80 °C before use. Concentrations of the peptides with no TAMRA were determined based on the observation that 1 mg/ml peptide generates an absorbance value (A_{205}) of 30 at 205 nm. Concentrations of TAMRA-containing peptides were determined based on the absorbance coefficient of TAMRA, namely 65,000 $cm^{-1}M^{-1}$ and 555nm. The integrity of the purified peptides was confirmed by MALDI mass spectrometry. The MALDI-MS are provided in Supplementary Figures S4.

4. Screening of *M. jannaschii* TyrRS-$tRNA_{CUA}^{Tyr}$ pairs to incorporate tmdF into BRD4 bromodomain

We acquired a set of expression constructs to incorporate tmdF **3** into BRD4 bromodomain in site-specific manner. These include the N-terminal 6xHis-tagged human BRD4-BD1 (the first bromodomain of BRD4) in pNIC28-Bsa4 vector (Addgene ID: 38942), and four *M. jannaschii* TyrRS-$tRNA_{CUA}^{Tyr}$ pairs, all are in pEVOL expression vector, originally evolved to incorporate *p*-azido-L-phenylalanine (Addgene ID: 31186),[6] *p*-benzoyl-L-phenylalanine (Addgene ID: 31190),[7] *p*-cyano-L-phenylalanine (a kind gift from Prof. Abhishek Chatterjee, Boston College) and *p*-acetyl-L-phenylalanine (a kind gift from Prof. Alexander Deiters, University of Pittsburgh).[8, 9] First, we introduced W81AG point mutation into the pNIC28-Bsa4 BRD4-BD1 construct using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies) following manufacturer's protocol. The resulting mutant plasmid was confirmed by DNA sequencing.

To express BRD4 variant carrying tmdF **3** at specific position, we co-transformed *E. Coli* BL21 star (DE3) cells (Invitrogen) with pNIC28-Bsa4 BRD4-W81TAG and one of the four *M. jannaschii* TyrRS-$tRNA_{CUA}^{Tyr}$ pairs, mentioned above.[10] Cells were recovered in 200 µL of the SOC medium for 1 h at 37 °C before being plated on a LB Miller agar plate containing Kanamycin sulfate (Kan) (50 µg/mL) and Chloramphenicol (Cm) (35 µg/mL). A single colony was selected and grown at 37 °C in 10 mL LB Miller broth in presence of 50 µg/mL Kan and 35 µg/mL Cm overnight. The overnight culture was centrifuged (ThermoScientific Sorvall Legend XTR

Centrifuge, TX-1000 rotor, 4°C) for 10 min at 1000 x g, 2100 rpm. 9 mL of LB Miller broth was subsequently removed and the cell pellet with remaining 1 mL LB Miller broth was used to inoculate 200mL of the GMML medium (M9 minimal media supplemented with 1% v/v glycerol, 300 μM leucine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 μg/mL kanamycin sulfate, and 35 μg/mL chloramphenicol, and trace amounts of Na_2MoO_4, $CoCl_2$, $CuSO_4$, $MnSO_4$, $MgSO_4$, $FeCl_2$, $CaCl_2$, and H_3BO_3). Cells were allowed to grow at 37 °C in an incubator shaker (225 rpm) until OD_{600} reached to ~0.8. tmdF **3** was prepared by adding 200 μL sterilized deionized water and 100 μL 10M NaOH to 55mg **3** then diluting to 10mL with sterilized deionized water followed by neutralization with 83 μL of 12.1M HCl. tmdF **3** was aseptically added to the culture to a final concentration of 1 mM and the culture was cooled to 17 °C for 30 min while shaking at 225 rpm. 10mL sterilized water was added to the negative controls. The culture was then induced with 0.05 % w/v arabinose and continued to shake at 17 °C for an additional 30 min. Finally, the protein expression was induced by the addition of 0.25 mM IPTG and allowed to shake at 225 rpm for 20 h at 17 °C in an Innova 44® Incubator shaker (New Brunswick Scientific). Proteins were purified as follows: harvested cells were resuspended in 3-5mL lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, 25 mM imidazole, 1 mg/mL Lysozyme, 10μg/mL or 20U/mL DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700, Fisherbrand Replaceable Microtip 1/8" Probe, pulsed 10 seconds on, 10 seconds off, for 2 minutes processing time, 60Amps), and centrifuged (ThermoScientific Sorvall Legend XTR Centrifuge F15-8x50cy rotor, 4°C) at 13000 rpm (19632.4 x g) for 40 min at 4 °C. The soluble extracts were subject to Ni-NTA agarose resin (Thermo) according to manufacturer's instructions. After passing 20 volumes (4 mL) of washing buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, and 25 mM imidazole), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, and 400 mM imidazole. Proteins were separated by SDS-PAGE and compared to those samples lacking tmdF **3**.

5. **Mutagenesis, expression and purification of BPTF Bromodomain**

The N-terminal 6xHis-tagged human BPTF bromodomain bacterial expression construct pNIC28-Bsa4 (Addgene ID: 39111) was obtained from Addgene. Wild type BPTF plasmid was transformed into *E. coli* BL21 (DE3) Star competent cells (Invitrogen) using pNIC28-Bsa4

kanamycin-resistant vector.[10] A single colony was picked, added to 10 mL of Luria-Bertani (LB) Miller broth with 50 μg/mL kanamycin sulfate, and grown overnight at 37 °C. The overnight culture was diluted 100-fold the next morning and allowed to grow at 37 °C to an optical density (OD600) of 0.8, at which time 0.5 mM IPTG was added to induce protein expression overnight at 17 °C with in an Innova 44® Incubator shaker (New Brunswick Scientific). Protein purification was performed as follows: cells were harvested by centrifugation at 4000 rpm for 20 minutes, and supernatant was discarded. Cell pellets were resuspended in 10-15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, 25 mM imidazole, 1 mg/mL Lysozyme, 10μg/mL or 20U/mL DNase, and Roche protease inhibitor cocktail), then lysed by pulsed sonication (Qsonica-Q700, Fisherbrand Replaceable Microtip 1/8" Probe, pulsed 10 seconds on, 10 seconds off, for 2 minutes processing time, 60Amps; repeated sonication a second or third time if solution was still opaque and viscous) at 13000 rpm (19632.4 x g) for 40 min at 4 °C. Soluble extracts were subject to Ni-NTA agarose resin (Thermo) according to manufacturer's instructions. Washing was achieved by allowing supernatant to flow through the column followed by 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, and 25 mM imidazole). Proteins were then eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% v/v glycerol, and 400 mM imidazole. Eluted proteins were further purified by size exclusion chromatography (Superdex-200) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, and 10% v/v glycerol. Purified protein was concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.), at which time protein concentration was determined using Bradford assay kit (BioRad Laboratories) with BSA as a standard. Concentrated proteins were stored at -80°C before use. BPTF variant W2950TAG carrying amber codon was generated using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies). The resulting mutant plasmids were confirmed by DNA sequencing.

To express BPTF-W2950pAzF and BPTF-W2950tmdF, BL21 Star (DE3) cells were co-transformed with pEVOL-based *M. jannaschii* TyrRS-$tRNA_{CUA}^{Tyr}$ pair for AzF (addgene ID: 31186) or CNF (a kind gift from Prof. Abhishek Chatterjee, Boston College).[11] Cells were recovered for 1 hour and 45 minutes in 200 μL SOC medium in a 37°C shaker prior to plating on an LB Miller agar plate containing 50 μg/mL kanamycin sulfate and 35 μg/mL chloramphenicol. Single colonies were picked and added to each of four inoculates containing 10 mL of Luria-

Bertani (LB) Miller broth in the presence of 50 μg/mL kanamycin sulfate and 35 μg/mL chloramphenicol. Overnight cultures were centrifuged (ThermoScientific Sorvall Legend XTR Centrifuge, TX-1000 rotor, 4°C) for 10 min at 1000 x g (2100 rpm). 9 mL of LB Miller broth was removed and the cell pellet was resuspended in the remaining 1mL LB Miller broth and used to inoculate 1L of GMML medium (M9 minimal media supplemented with 1% v/v glycerol, 300 μM leucine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 μg/mL kanamycin sulfate, and 35 μg/mL chloramphenicol, and trace amounts of Na_2MoO_4, $CoCl_2$, $CuSO_4$, $MnSO_4$, $MgSO_4$, $FeCl_2$, $CaCl_2$, and H_3BO_3). Cells were allowed to grow at 37 °C to an optical density (OD600) of 0.8. tmdF **3** was prepared by adding 1mL sterilized deionized water and 500 μL 10M NaOH to 273mg **3** then diluting to 50mL with sterilized deionized water followed by neutralization with 415 μL of 12.1M HCl. pAzF **2** was prepared by diluting in 20mL sterilized deionized water. The prepared unnatural amino acid (pAzF or tmdF) was added aseptically to a final concentration of 1mM. Cells were allowed to shake an additional 30 minutes at 17 °C, at which time the synthetase expression was induced with 0.05% w/v arabinose and allowed to shake an additional 30 minutes at 17 °C. Finally, 0.425 mM IPTG was added to induce BPTF protein expression overnight at 17 °C with in an Innova 44® Incubator shaker (New Brunswick Scientific). Protein purification, concentration, and storage were performed as previously described.

6. Mutagenesis, expression and purification of wild-type histone H4 and its variants

Gene sequence encoding wild type Xenopus laevis histone H4 was a kind gift from Dr. Minkui Luo at the Memorial Sloan-Kettering Cancer Center. The plasmid containing wild type Histone H4 was transformed into BL21 codon plus (DE3) RIPL competent cells. A single colony was picked up and grown overnight at 37 ^oC in 10 mL of LB Miller broth with 100 μg/mL ampicillin and 35 μg/mL chloramphenicol. The inoculation culture was diluted 1:100 fold in fresh LB Miller broth and cells were grown at 37 ^oC until OD_{600} reached to ~0.7. Protein expression was induced by the addition of 0.3 mM IPTG followed by growing for an additional 3 h at 37 ^oC. Cells were harvested by centrifugation at 5000 rpm for 30 min, and then resuspension of the pellet in 5 mL of lysis buffer (10 mM Tris-HCl pH 7.5, 2 M guanidinium hydrochloride (GdnHCl), 5 mM β-mercaptoethanol, 10% v/v glycerol, 10μg/mL or 20U/mL DNase, 1 mg/mL Lysozyme and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700, Fisherbrand Replaceable Microtip 1/8" Probe, pulsed 10 seconds on, 10 seconds off, for 2 minutes processing

time, 60Amps; repeated sonication a second or third time if solution was still opaque and viscous) and centrifuged (ThermoScientific Sorvall Legend XTR Centrifuge F15-8x50cy rotor, 4°C) at 20,000 rpm (46467.2 x g) for 40 min at 4 °C. Insoluble histone was recovered from inclusion bodies by dissolving in 10 mM Tris-HCl pH 7.5 and 6 M GdnHCl incubated for 10 min at room temperature followed by centrifugation at 20,000g for 40 min at 4 °C. The soluble histone supernatant was purified by size exclusion chromatography on a Superdex-200 using AKTA pure FPLC system. Fractions were concentrated using Amicon Ultra-4 centrifugal 3K filter and further purified with preparative reversed-phase HPLC (XBridge C18, 5 µm, 10 x 250 mm column) eluting with a flow rate of 4 mL/min starting from 10% v/v acetonitrile to 70% v/v in 15 min and then to 100% v/v over 5 min in aqueous trifluoroacetic acid (0.01% v/v). The purified protein was concentrated by SpeedVac followed by lyophilization. The protein was stored at -20 °C before use. Histone H4 mutants (K8C and K12C) were generated by the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies) following manufacturer's protocol. The mutant plasmids were confirmed by DNA sequencing. The mutant proteins were expressed and purified as described above for the wild type Histone H4.

7. Chemical acetylation of cysteinylated histone H4

Purified H4 mutants (0.5 mg) was dissolved in 375 µL of 6 M GdnHCl and 200 mM sodium acetate at pH 6.0.[12] To this solution was added 75 mM glutathione, 250 mM *N*-vinylacetamide, 500 mM dimethylsulfide, and 250 mM of the azo radical initiator VA-044 (2,2́ -azobis[2-(2-imidazolin-2-yl) propane]dihydrochloride). The reaction was initiated by incubating the above mixture in the dark at 37 °C for 1 h. The product was purified by RP-HPLC as described above in section 8 and characterized by ESI LC-MS.

8. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed with an ITC_{200} instrument (MicroCal, Malvern). Experiments were conducted at 15 °C while stirring at 750 rpm. Buffers of protein and peptides were matched to 50mM Tris-HCL pH 8.0, 200mM NaCl, and 0-10% v/v glycerol. Each titration was composed of one initial injection of 0.4 µL for 0.8 seconds, followed by 19 injections of 2.0 µL for 4 seconds, with two-minute intervals between each injection. The initial injection

was discarded prior to data analysis. The micro syringe (40 μL) was loaded with 6mM H4 tetra-acetylated peptide **13** and injected into the cell (200μL), occupied by a BPTF protein at a concentration of 130 to 215 μM. All data was fitted to a single binding site model using the Microcal ITC_{200} with Origin Lab 7 to yield enthalpies of binding (ΔH) and binding constants (K_a). When necessary, the parameter N for ratio of bound peptide to protein was fixed to predicted value 1 due to low c-value caused by concentration restraints. Further thermodynamic parameters i.e. changes in entropy ΔS, changes in free energy ΔG, and dissociation constants (K_d) were calculated from these values.

9. Photo-crosslinking experiment with peptides and in-gel fluorescence

Time-dependent crosslinking data was obtained for BPTF W2950pAzF and BPTF W2950tmdF. Assay volumes totaled 20 μL and contained 25 μM purified protein and 1 μM peptides **14** and **15**, one at a time, diluted with buffer containing 10mM Tris HCl pH 7.5, 150mM NaCl, 0.05% w/v TWEEN 20, and 0.5mM TCEP. Samples were vortexed (Fisher Scientific Mini Vortexer 120V) and centrifuged (Fisherbrand Sprout) prior to preincubation for 30 minutes at room temperature away from light. Samples were subjected to UV light 3 inches from a lamp (MAXIMA ML-3500S UV, 50,000 μW/cm^2, 350nm-385nm 50% irradiated, 365nm 100% irradiated) for 15, 30, 45, or 60 minutes on ice. Negative controls were kept in a drawer on ice for the full 60 minutes, and samples requiring less than 60 minutes were transferred to the same drawer when the allotted time was reached. After addition of 10 μL of 4X Laemmli loading buffer, samples were incubated at 95 °C for 10 minutes prior to loading on a precast 4-12% w/v Bis-Tris polyacrylamide gel (BIORAD cat#3450123). Gels were subjected to electrophoresis at 150V for 1 hour and imaged immediately after electrophoresis via BIORAD GelDoc rhodamine program. Coomassie staining was then performed to confirm presence of BPTF in each sample. Data was analyzed using ImageLab 5.0. Lanes and bands were created manually. The Lane Profile feature was used to ensure each band's intensity was captured in full while excluding background. Fluorescence intensities were normalized to the 15-minute pAzF band for each gel separately. Coomassie intensities of total non-crosslinked protein were normalized to the 0-minute pAzF intensity for each gel separately. Normalized fluorescence values (NF) were then divided by normalized Coomassie values (NC) at each respective timepoint. Lastly, NF/NC ratios were normalized to tmdF NF/NC at 60 minutes

for each gel separately. GraphPad Prism 7 was used to plot the bar graph as well as to calculate mean and standard deviation for N=3 data.

10. Full length histone crosslinking and Western blotting

Photo-crosslinking experiments were conducted as follows: 20 μM of purified acetylated thialysine histone ($H4K_C8Ac$ or $H4K_C12Ac$) or ~0.5ug histone H4 (prepared as histone extracts from HEK293T cells as described in reference[13] and using BSA standards to estimate concentration) and 50-100μM BPTF W2950pAzF or BPTF W2950tmdF were diluted to 20μL with buffer containing 10mM Tris HCl pH 7.5, 150mM NaCl, 0.05% w/v TWEEN 20, and 0.5mM TCEP. Vortexed (Fisher Scientific Mini Vortexer 120V) and centrifuged (Fisherbrand Sprout) samples were split into two 10μL aliquots. Samples were preincubated and crosslinked as described above for two timepoints of 0 and 30-60 minutes. Samples were diluted 1:1 with 10μL of 4X Laemmli loading buffer, incubated at 95°C for 10 minutes, and loaded onto a precast 4-12% w/v Bis-Tris polyacrylamide gel (BIORAD cat#3450123). The protein (crosslinked and free) bands were separated by electrophoresis at 150V for 1 hour.

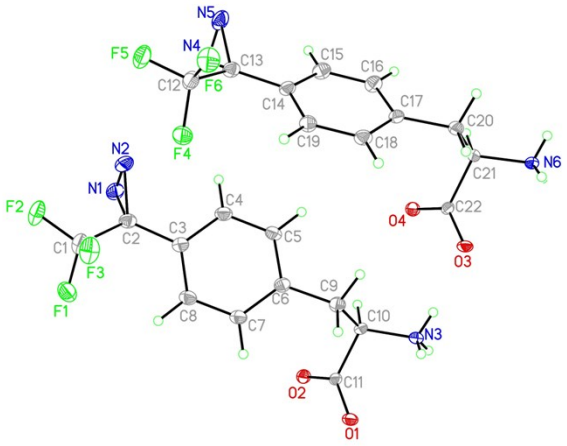
The separated protein bands were transferred to a 0.2 μm nitrocellulose membrane (BIORAD cat #1620112) at a constant voltage of 80V for 1 hour. The membrane was washed once with TBST (50 mM Tris HCl pH 7.4, 150 mM NaCl, 0.01% w/v Tween-20) prior to blocking with 5% w/v milk in TBST buffer for 1 hour. Immunoblotting was performed with 1:500 diluted H4 primary antibody (Histone-H4 mAb, cat# 61521, Active Motif) or 1:200 diluted His primary antibody (Anti-6X His tag® antibody - ChIP Grade, cat #ab9108, Abcam (or His-probe (H-3) mAb, cat# sc-8036, Santa Cruz Biotechnology) and incubated overnight at 4 °C. Antibody solutions were then removed, and membranes were washed with TBST three times. A dilution of 1:5000 secondary antibody was then added to membranes, either either anti-rabbit (Goat Anti-Rabbit IgG-HRP, cat #4030-05, Southern Biotech) for 6x-His or anti-mouse (HRP Goat anti-Mouse IgG, cat #15014, Active Motif) for H4 and H-3, and incubated at room temperature for 1 hour. Secondary antibody was washed off three times with TBST, and membranes were visualized using VISIGLO HRP Chemiluminescent substrates A and B (cat# N252-120ML and N253-120ML, aMReSCO) following manufacturer's protocol.

11. Method for energy minimization

The molecular models of both the enantiomers of PCAA **3** were built, and their geometry were optimized through the Discovery Studio 4.5 visualizer builder module. The known crystal structure of CNF synthetase (PDB ID: 3qe4) was obtained from the Brookhaven Protein Data Bank. Energy optimized structure of the synthetase and PCAA **3** was further prepared using Autodock tools 1.5.4. The 3D affinity grid box was designed to include the full active site pocket. Docking calculations were set at 10 runs. At the end of the calculation, cluster analysis was performed by Autodock. Docking solutions with ligand all atom root-mean-square deviation (rmsd) within 1.0 Å of each other were clustered together and ranked by the lowest energy representative. The lowest-energy solution was considered one of the most representatives of the PCAA **3**-synthetase docked complex.

12. Supplementary figures and tables

A



B

Bond precision:	C-C = 0.0050 A		Wavelength=1.54178
Cell:	a=6.1114(3)	b=8.7281(4)	c=20.9657(10)
	alpha=80.025(3)	beta=89.183(3)	gamma=88.020(3)
Temperature:	150 K		

	Calculated	Reported
Volume	1100.73(9)	1100.73(9)
Space group	P -1	P -1
Hall group	-P 1	-P 1
Moiety formula	C11 H10 F3 N3 O2	?
Sum formula	C11 H10 F3 N3 O2	C11 H10 F3 N3 O2
Mr	273.22	273.22
Dx,g cm-3	1.649	1.649
Z	4	4
Mu (mm-1)	1.314	1.314
F000	560.0	560.0
F000'	562.32	
h,k,lmax	7,10,25	7,10,25
Nref	4047	3806
Tmin,Tmax	0.868,0.961	0.780,0.910
Tmin'	0.854	

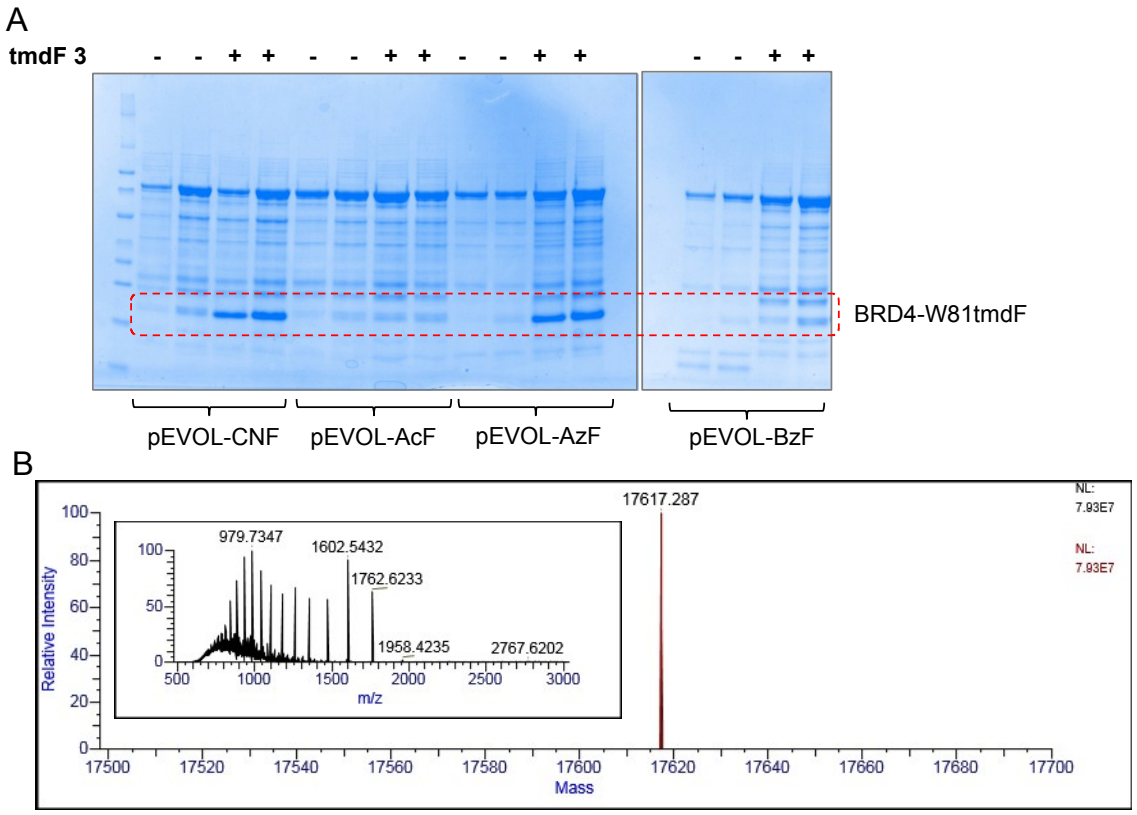
Correction method= # Reported T Limits: Tmin=0.780 Tmax=0.910
AbsCorr = MULTI-SCAN

Data completeness= 0.940 Theta(max)= 68.360

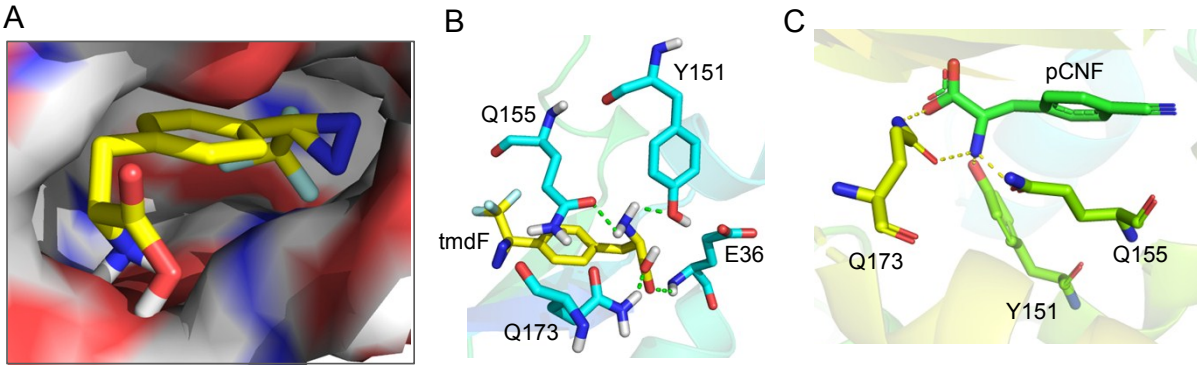
R(reflections)= 0.0813(3014) wR2(reflections)= 0.2061(3806)

S = 1.876 Npar= 367

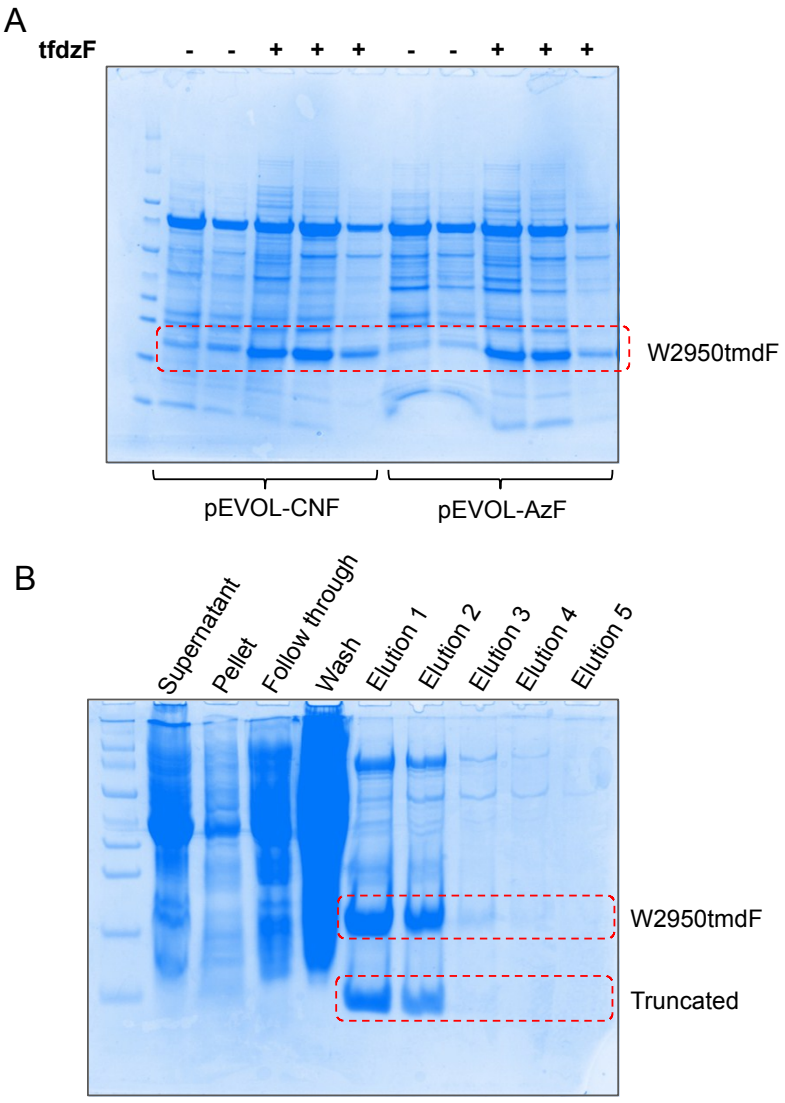
Supplementary Figure S1: (A) OPRTEP diagram of (±)-**3**. Two molecules are one asymmetric unit. (B) Crystallographic parameters. 20 mgs of tmdF **3** was dissolved in 4 mL of boiling water and allowed to cool until desirable crystals were obtained. The X-ray intensity data were measured on a Bruker X8 Prospector Ultra Cu ImuS Bruker Apex II CCD system. The structure was solved and refined using the Bruker SHELXTL Software Package. The data was deposited to Cambridge Crystallographic Data Center (CCDC #1978317)



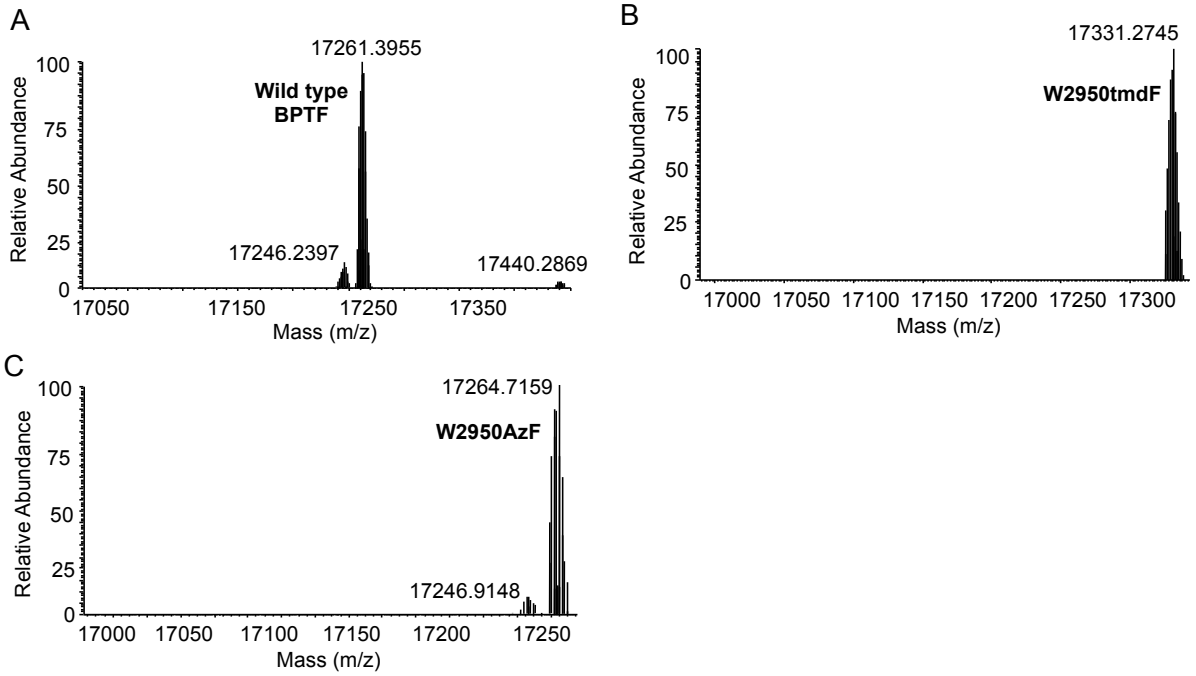
Supplementary Figure S2: Coomassie blue staining showing expression of BRD4-W81tmdF protein only by *M. jannaschii* TyrRS-tRNA$_{CUA}^{Tyr}$ pairs originally evolved to incorporate *p*-cyano-L-phenylalanine (CNF) and *p*-azido-L-phenylalanine (AzF) in the presence of tmdF **3**. The screening was performed twice (N=2) each in 200 mL GMML growth medium. (B) Deconvoluted mass of BRD4-W81tmdF: Calculated: 17618.05, Observed: 17617.29 (see Table S3)



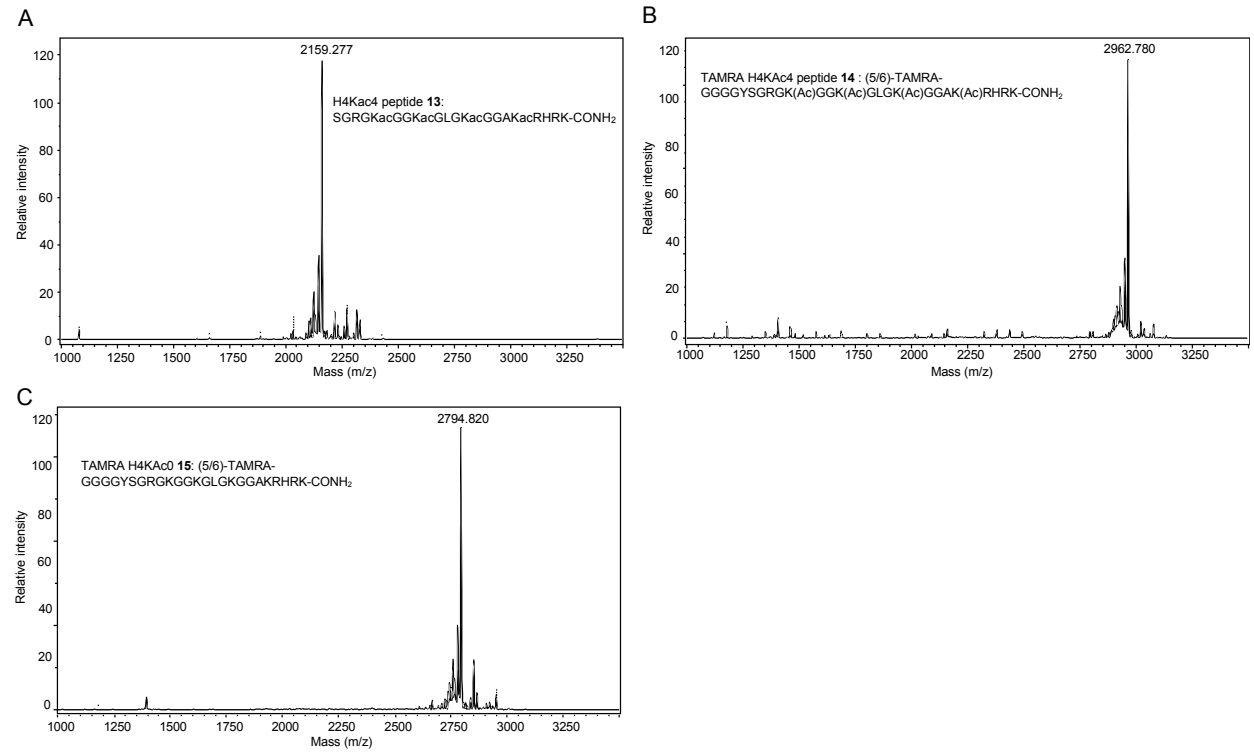
Supplementary Figure S3: (A) Energy-minimized structure of R enantiomer of **3** in the catalytic pocket of evolved CNF synthetase based *M. jannaschii* TyrRS (B) Polar contacts between R enantiomer of **3** and the active site residues. (C) Crystal structure showing binding of CNF to the evolved CNF synthetase (PDB: 3qe4) as reported in Ref. 14.[14] Both tmdF and CNF share similar polar and hydrophobic interactions with the synthetase.



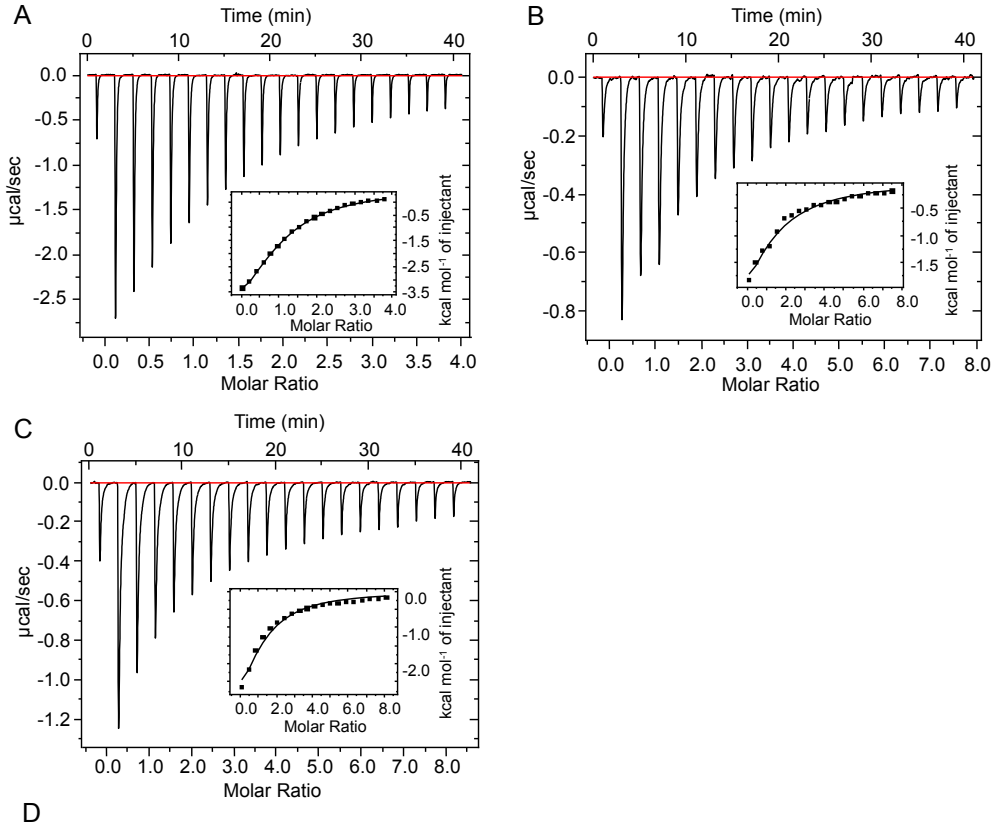
Supplementary Figure S4: (A) Coomassie blue staining showing expression of BPTF-W2950tmdF protein only by *M. jannaschii* TyrRS-tRNA$_{CUA}^{Tyr}$ pairs originally evolved to incorporate *p*-cyano-L-phenylalanine (CNF) and *p*-azido-L-phenylalanine (AzF) in the presence of tmdF **3**. The screening was performed twice (N=2) each in 200 mL GMML growth medium. (B) 1 L expression of BPTF-W2950tmdF by pEVOL-CNF *M. jannaschii* TyrRS-tRNA$_{CUA}^{Tyr}$ pair. The truncated protein was also pulled down during affinity purification due to the presence of N-terminal 6xHis in the BPTF construct. W2950tmdF was separated from the truncated protein by subsequent fast protein liquid chromatography (FPLC).



Supplementary Figure S5: ESI LC-HRMS of purified wild type BPTF bromodomain (A), and its W2950tmdF (B) and W2950AzF (C) mutants.



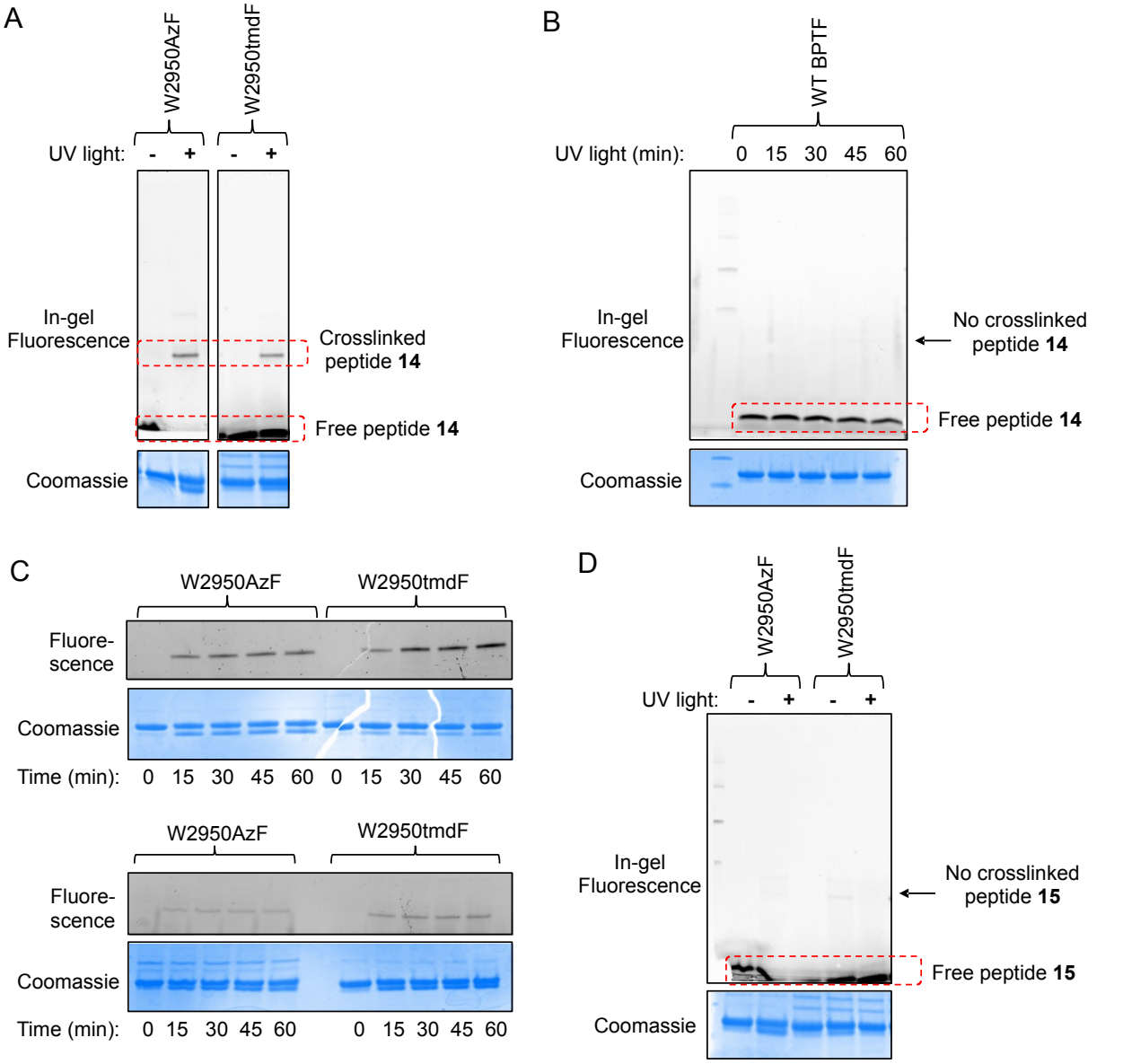
Supplementary Figure S6: MALDI-MS spectra of H4Kac4 peptide **13** (A), TAMRA-H4Kac4 peptide **14** (B) and TAMRA-H4Kac0 peptide **15** (C).



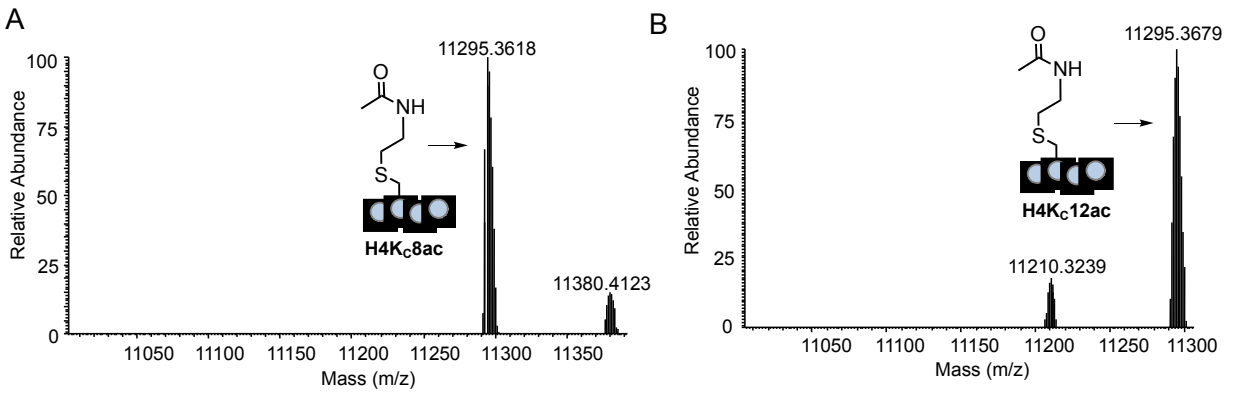
D

Protein	% Glycerol	Peptide	Peptide Sequence	Protein (µM)	Peptide (mM)	K_D(µM)	N	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
BPTF WT	10	H4Kac4 **13**	SGRGK$_{(Ac)}$GGK$_{(Ac)}$GLGK$_{(Ac)}$GGAK$_{(Ac)}$RHRK	214	4	185 ± 10	1.00 ± 0.02	−7.223 ± .109	-2.302	-4.921
BPTF W2950AzF	0	H4Kac4 **13**	SGRGK$_{(Ac)}$GGK$_{(Ac)}$GLGK$_{(Ac)}$GGAK$_{(Ac)}$RHRK	216	5	309 ± 23	1.00 (fixed)	−5.898 ± .250	-1.268	-4.630
BPTF W2950tmdF	10	H4Kac4 **13**	SGRGK$_{(Ac)}$GGK$_{(Ac)}$GLGK$_{(Ac)}$GGAK$_{(Ac)}$RHRK	150	6	231 ± 21	1.00 (fixed)	−6.331 ± .295	-1.536	-4.795

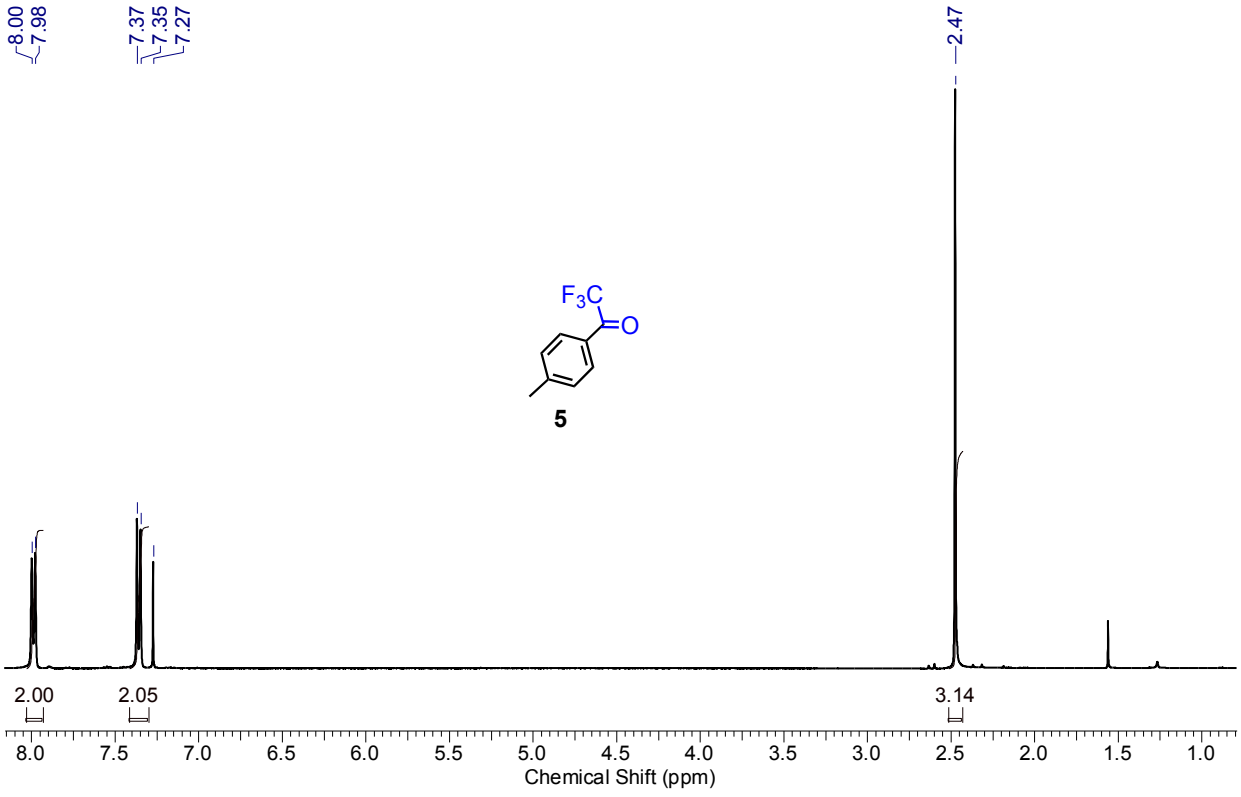
Supplementary Figure S7: Binding of H4Kac4 peptide **13** peptide to wild type BPTF bromodomain with a K_d of 185±10 µM (A), to W2950AzF mutant a K_d of 309±23 µM (B) and to W2950tmdF mutant with a K_d of 231±21 µM (C) as measured by isothermal titration calorimetry (ITC). (D) Thermodynamic parameters measured by ITC for the binding of peptide **13** to wild type BPTF bromodomain and its W2950AF and W2950tmdF mutants. Conditions are detailed in the experimental section.



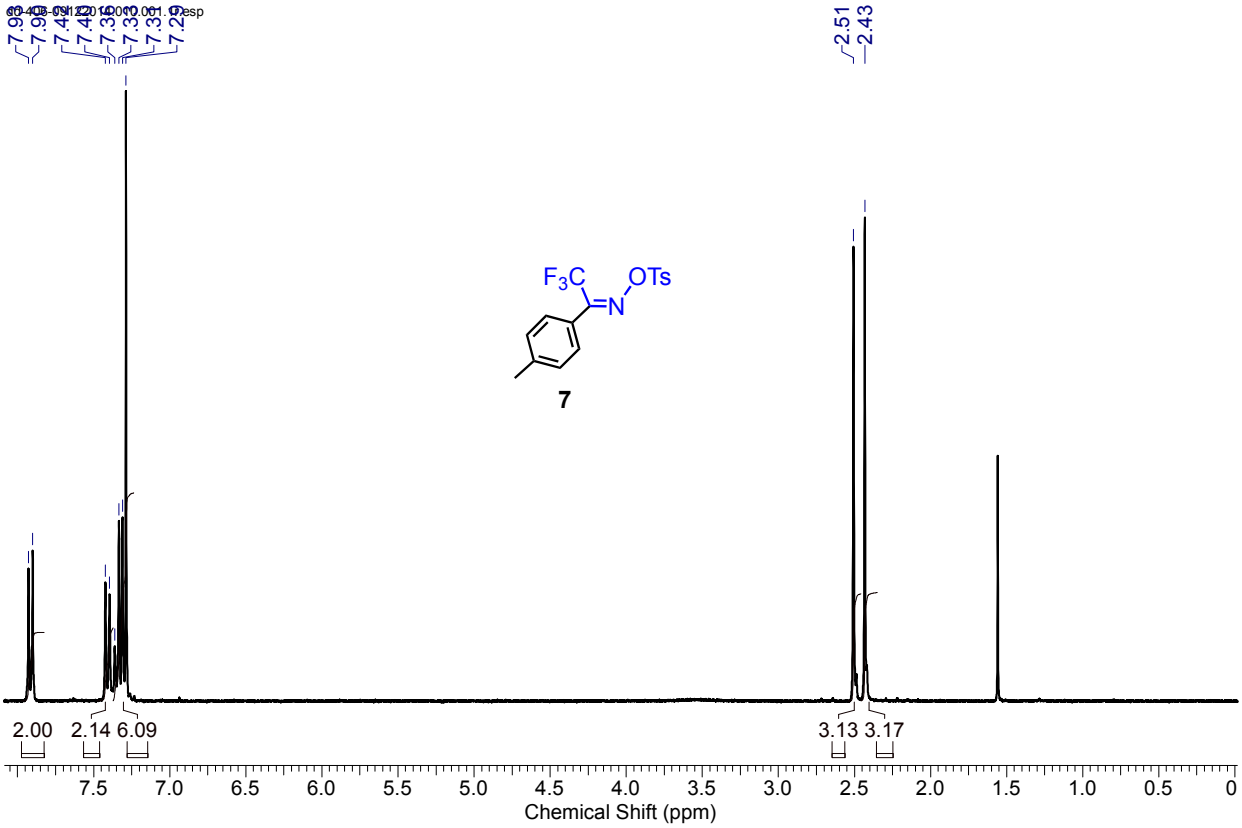
Supplementary Figure S8: UV light-dependent crosslinking of BPTF mutants to H4 peptides. Fluorescently labeled crosslinked species are visualized by in-gel fluorescence. (A) Upon incubation of W2950AzF or W2950tmdF mutant with TAMRA-H4Kac4 peptide **14**, samples were exposed to 365nm UV light followed by in-gel fluorescence using 532nm light (λ_{max} for TAMRA). Samples that underwent successful photo-crosslinking upon exposure to 365nm UV light are indicated by high molecular band visible under 532nm light. Coomassie staining of the same gel showed presence of proteins in all the samples. The same amount of protein was used to perform each crosslinking experiment. (B) Wild type BPTF bromodomain failed to crosslink the peptide. (C) Two time-dependent crosslinking experiments of W2950AzF and W2950tmdF mutant to TAMRA-H4Kac4 peptide **14** as revealed by in-gel fluorescence. (D) The BPTF mutants failed to crosslink peptide **15** which has sequence similar to **14** but lacks the acetyl groups on lysine residues.



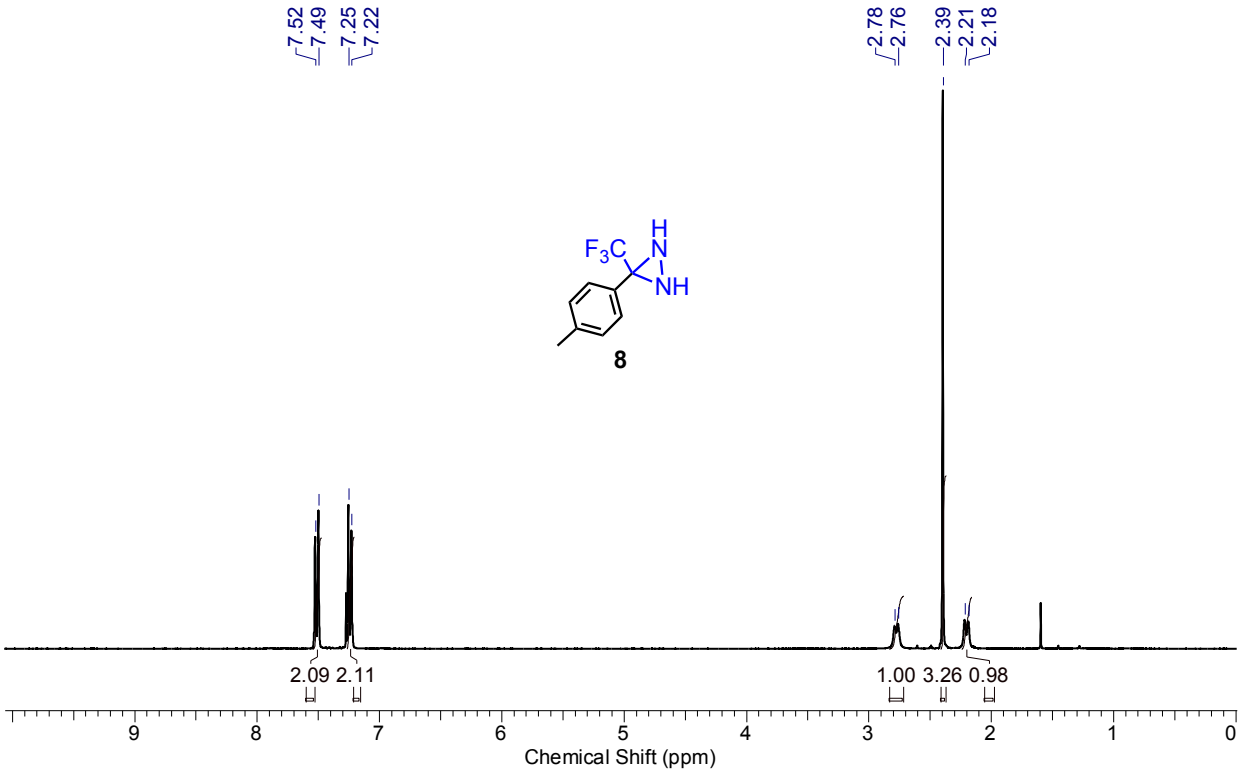
Supplementary Figure S9: ESI LC-HRMS spectra of full-length H4K$_C$8ac (A) and H4K$_C$12ac (B). Peak at m/z=11380.4123 in A is likely due to addition of second vinyl acetamide group during thermal thiol-ene reaction. Peak at m/z=11210.3239 in B corresponds to unreacted H4K12C mutant.



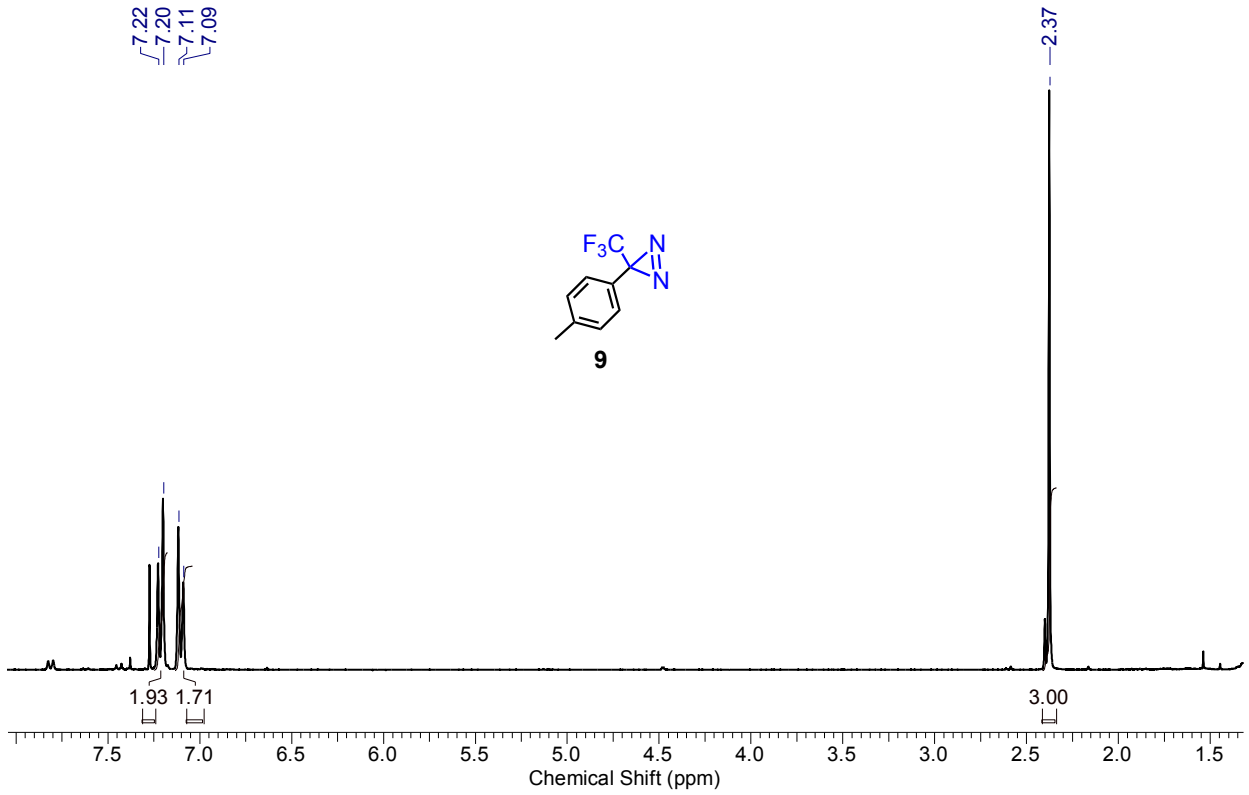
Supplementary Figure S10: ^{1}H spectrum of the indicated compound in $CDCl_3$



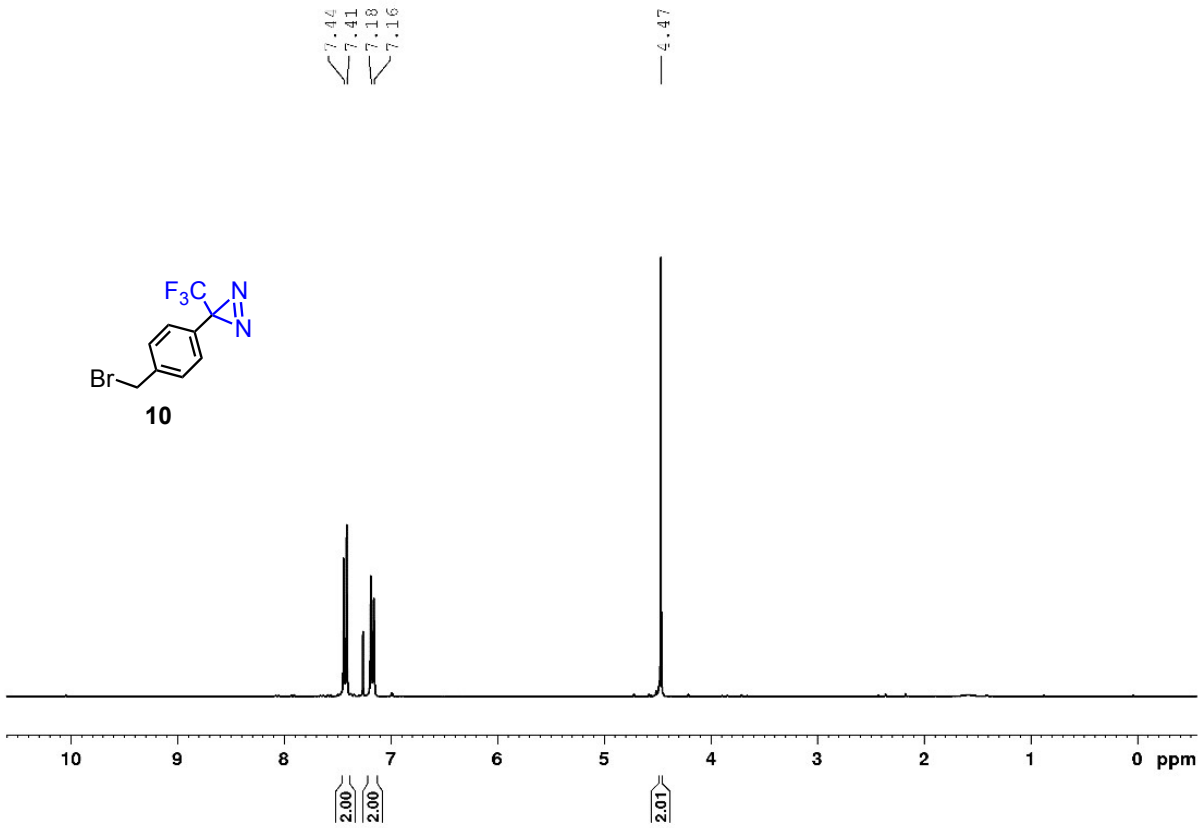
Supplementary Figure S11: ^{1}H spectrum of the indicated compound in $CDCl_3$



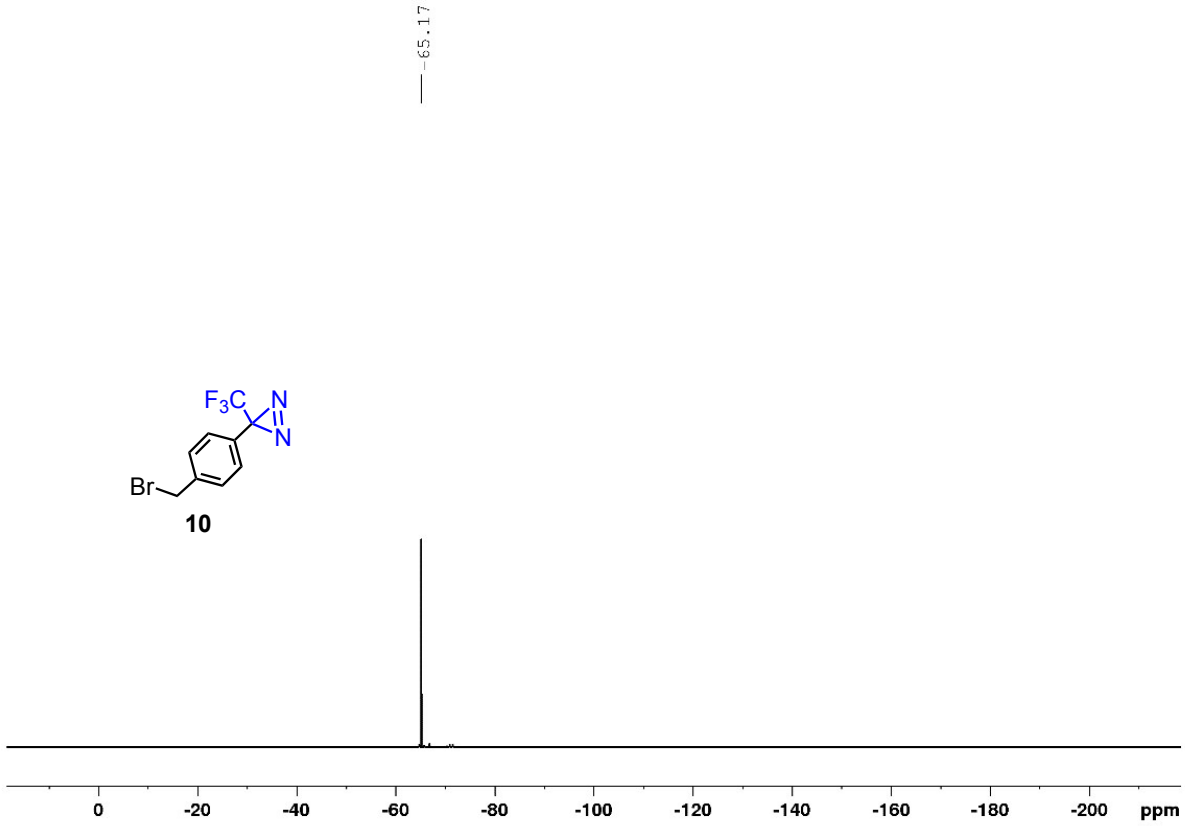
Supplementary Figure S12: ^{1}H spectrum of the indicated compound in $CDCl_3$



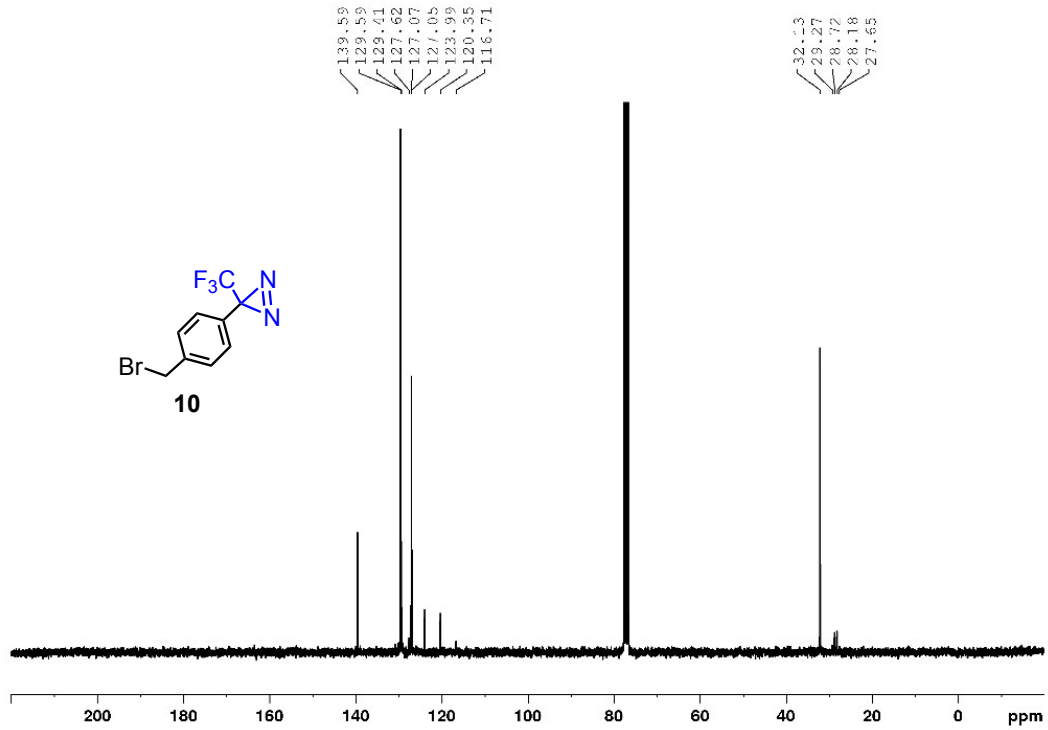
Supplementary Figure S13: ^{1}H spectrum of the indicated compound in $CDCl_3$



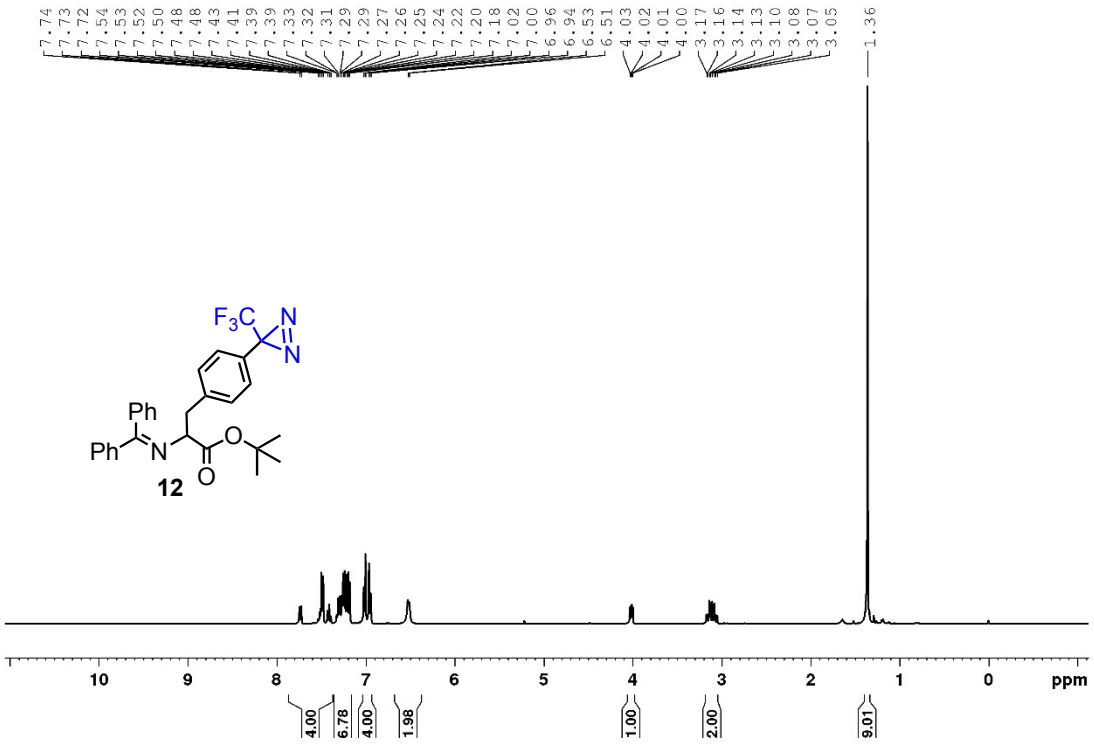
Supplementary Figure S14: ^{1}H spectrum of the indicated compound in $CDCl_3$



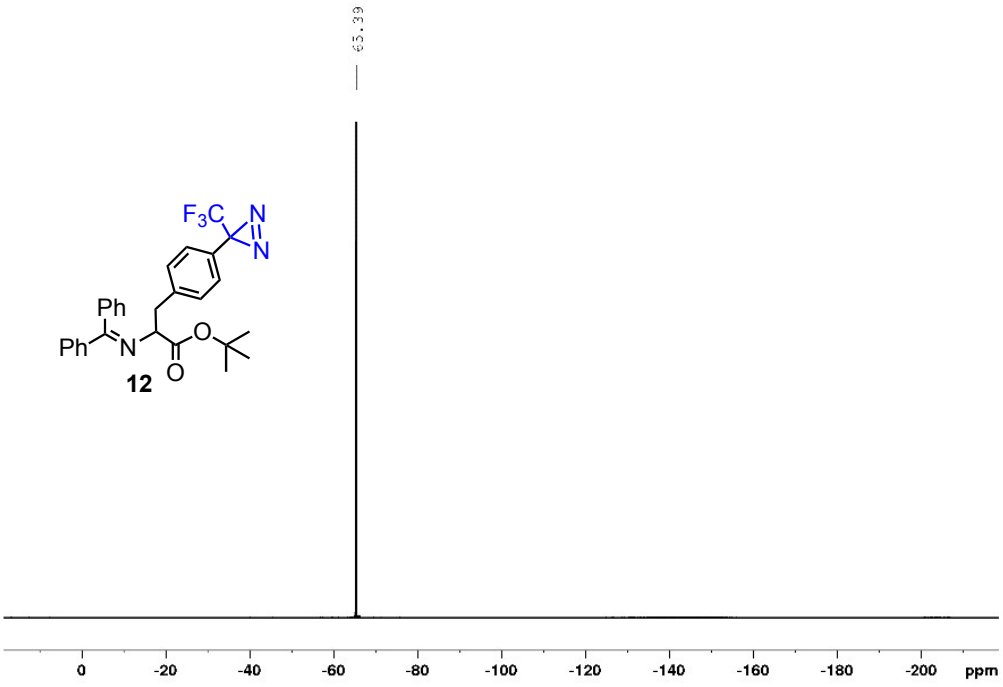
Supplementary Figure S15: ^{19}F spectrum of the indicated compound in $CDCl_3$



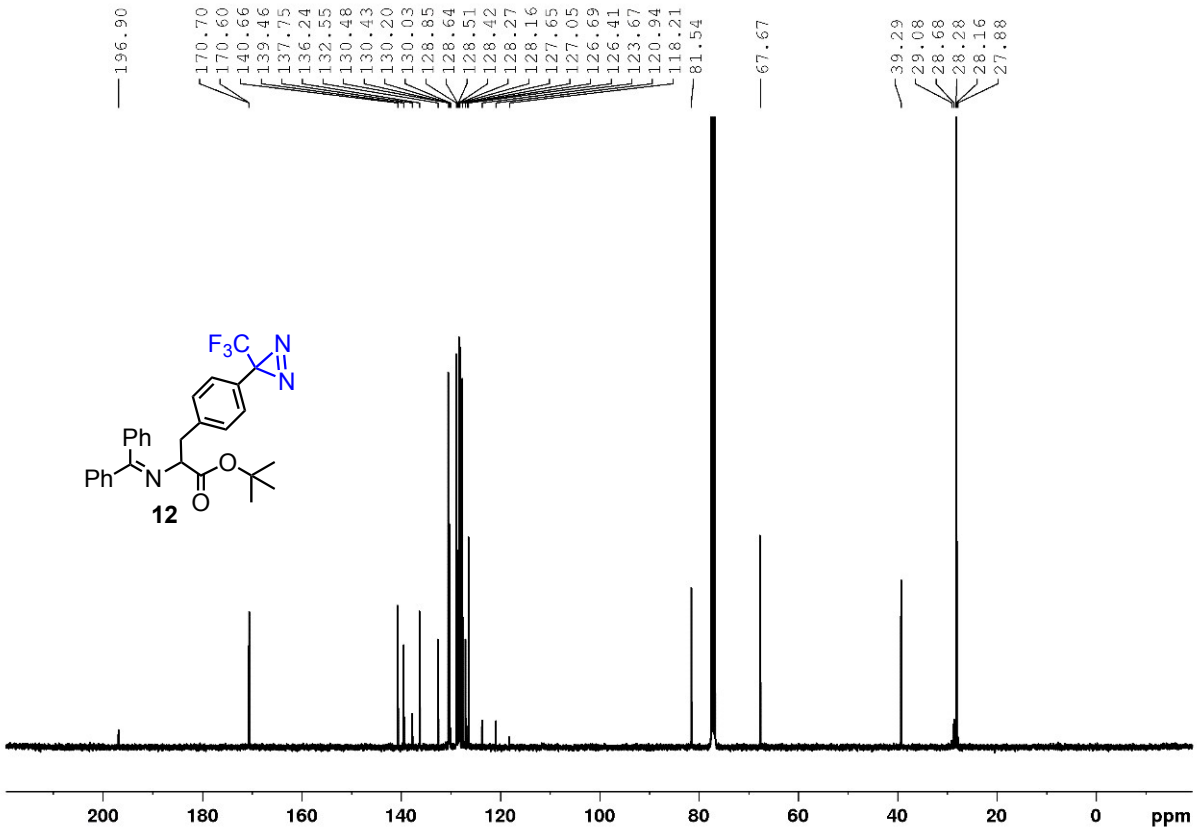
Supplementary Figure S16: ^{13}C spectrum of the indicated compound in $CDCl_3$



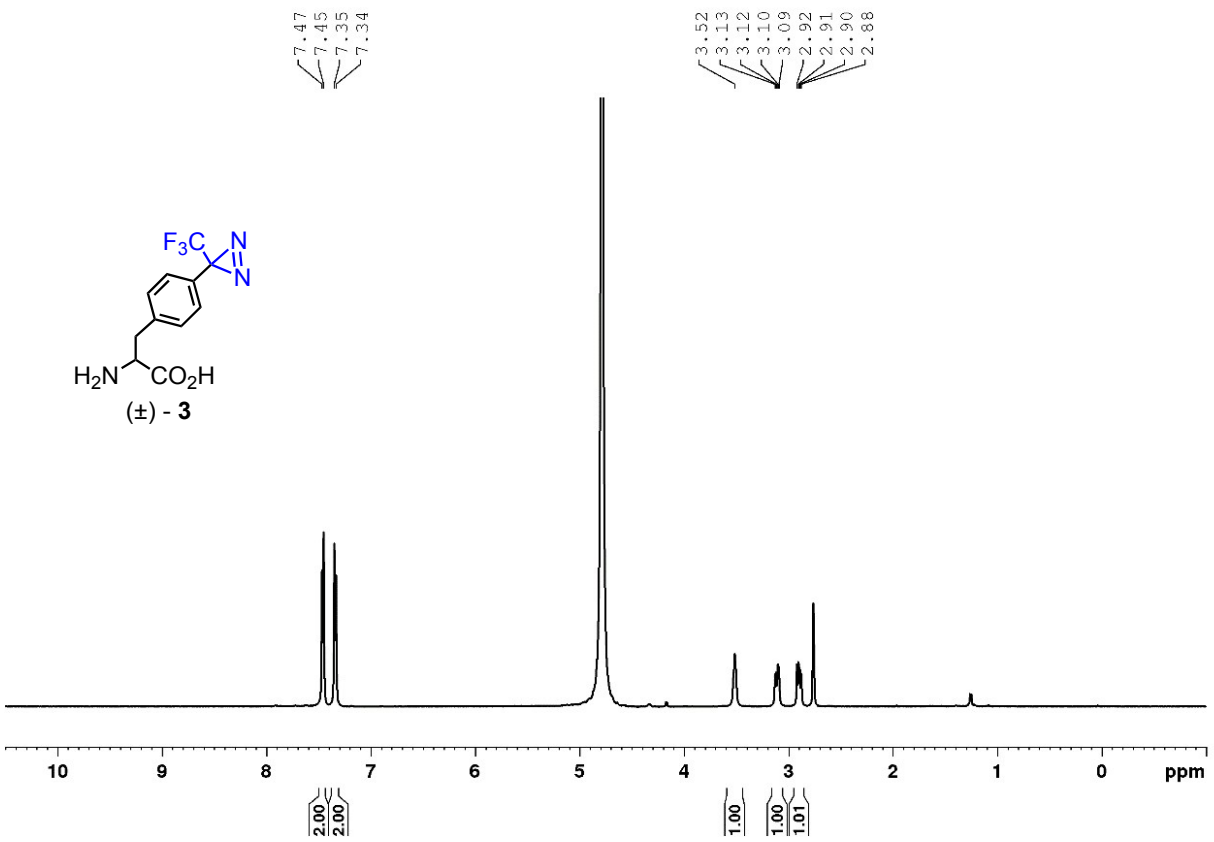
Supplementary Figure S17: ^{1}H spectrum of the indicated compound in $CDCl_3$



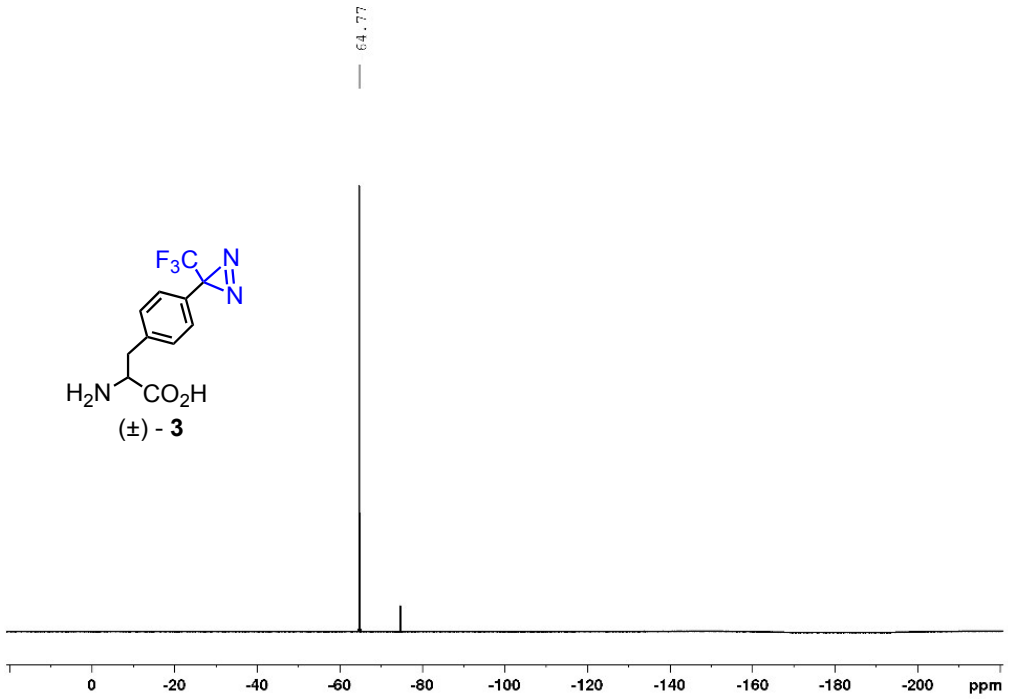
Supplementary Figure S18: ^{19}F spectrum of the indicated compound in $CDCl_3$



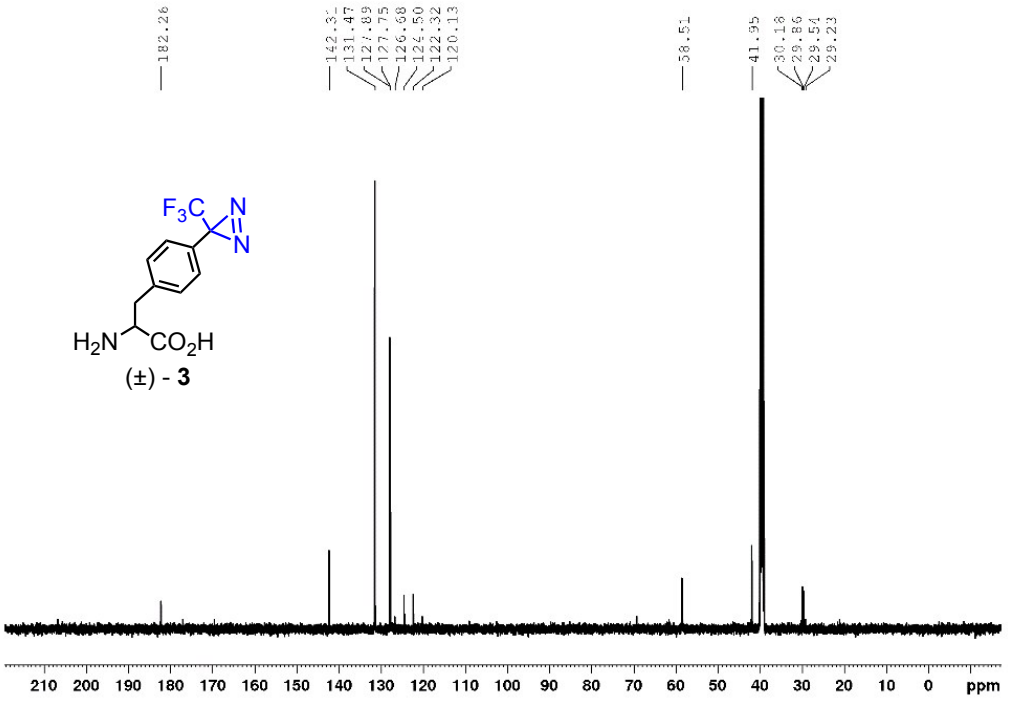
Supplementary Figure S19: ^{13}C spectrum of the indicated compound in $CDCl_3$



Supplementary Figure S20: ^{1}H spectrum of the indicated compound in 1:1 D_2O/DMSO-D_6 + 1 µL 10M NaOH



Supplementary Figure S21: ^{19}F spectrum of the indicated compound in 1:1 D_2O/DMSO-D_6 + 1 µL 10M NaOH



Supplementary Figure S22: ^{13}C spectrum of the indicated compound in 1:1 D_2O/DMSO-D_6 + 1 µL 10M NaOH

GENE	VECTOR	AFFINITY TAG	RESISTANCE
BRD4-BD1	pNIC28-Bsa4	N-6xHis	Kanamycin
BPTF-BD	pNIC28-Bsa4	N-6xHis	Kanamycin
M. j. TyrRS-tRNACUA pairs	pEVOL	None	Chloramphenicol
HISTONE H4	pDEST	None	Ampicillin

Supplementary Table S1. List of the genes used in the current study. The expression vector, antibiotic resistance and the affinity tag present for protein purification are provided.

MUTATION	FORWARD PRIMERS
BRD4-W81TAG	5'-CACCAGTTTGCATAGCCTTTCCAGCAGCC-3'
BPTF-W2950TAG	5'-GCGCATAAAATGGCGTAGCCGTTTCTGGAACCG-3'
H4-K8C	5'-CGTGGTAAAGGTGGTTGCGGTCTGGGTAAAGGT-3'
H4-K12C	5'-GGTAAAGGTCTGGGTTGCGGTGGTGCTAAACGT-3'

Supplementary Table S2. List of primers designed for site-directed mutagenesis. Reverse primers used are the reverse-complement to the given forward primers.

Protein Name	Amino acid Molecular weight	Noncanonnical Amino Acid Molecular weight AzF or tmdF	Modification	Molecular weight difference from WT	Calculated Molecular weight	Observed (LC-MS)
BPTF-WT	–	–	–	–	17262.47	17261.40
BPTF-W2950pAzF	Trp-204.23	206.20	W2950 to pAzF	1.97	17264.44	17264.72
BPTF-W2950tmdF	Trp-204.23	273.07	W2950 to tmdF	68.98	17331.31	17331.27
BRD4-W81tmdF	Trp-204.23	273.07	W81 to tmdF	68.98	17618.05	17617.3

Supplementary Table S3. Calculated and experimental molecular weights of wild type BPTF-bromodomain, BPTF-W2950AzF, BPTF-W2950tmdF and BRD4-W81tmdF mutants.

13. Bibliography

1. M. Nassal, *J. Am. Chem. Soc.* 1984, **106**, 7540-7545.
2. G. Baldini, B. Martoglio, A. Schachenmann, C. Zugliani and J. Brunner, *Biochem.* 1988, **27**, 7951-7959.
3. K. Masuda, A. Koizumi, T. Misaka, Y. Hatanaka, K. Abe, T. Tanaka, M. Ishiguro and M. Hashimoto, *Bioorg. Med. Chem. Lett.* 2010, **20**, 1081-1083.

4. H. Nakashima, M. Hashimoto, Y. Sadakane, T. Tomohiro and Y. Hatanaka, *J. Am. Chem. Soc.* 2006, **128**, 15092-15093.
5. J. E. Horne, M. Walko, A. N. Calabrese, M. A. Levenstein, D. J. Brockwell, N. Kapur, A. J. Wilson and S. E. Radford, *Angew. Chem. Int. Ed.* 2018, **57**, 16688-16692.
6. J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *J. Am. Chem. Soc.* 2002, **124**, 9026-9027.
7. J. W. Chin, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *Proceed. Natl. Acad. Sci. USA* 2002, **99**, 11020-11024.
8. T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol.* 2010, **395**, 361-374.
9. D. D. Young, T. S. Young, M. Jahnz, I. Ahmad, G. Spraggon and P. G. Schultz, *Biochem.* 2011, **50**, 1894-1900.
10. B. Sudhamalla, D. Dey, M. Breski, T. Nguyen and K. Islam, *Chem. Sci.* 2017, **8**, 4250-4256.
11. G. T. Perell, N. K. Mishra, B. Sudhamalla, P. D. Ycas, K. Islam and W. C. K. Pomerantz, *Biochem.* 2017, **56**, 4607-4615.
12. A. Dhall, S. Wei, B. Fierz, C. L. Woodcock, T. H. Lee and C. Chatterjee, *J. Biol. Chem.* 2014, **289**, 33827-33837.
13. D. Shechter, H. L. Dormann, C. D. Allis and S. B. Hake, *Nat. Protoc.* 2007, **2**, 1445-1457.
14. E. M. Tippmann, W. Liu, D. Summerer, A. V. Mack and P. G. Schultz, *ChemBioChem.* 2007, **8**, 2210-2214.