Supporting Information

Genetically Encoding an Electrophilic Amino Acid for Protein Stapling and Covalent Binding to Native Receptors

Xiao-Hua Chen[†], Zheng Xiang[†], Ying S. Hu[‡], Vanessa K. Lacey[†], Hu Cang[‡], Lei Wang[†]*

[†]The Jack H. Skirball Center for Chemical Biology and Proteomics, The Salk Institute for Biological Studies,10010 N. Torrey Pines Road, La Jolla, California, 92037 (USA)

[‡]Waitt Advanced Biophotonics Center, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, California, 92037 (USA)

E-mail: lwang@salk.edu

Construction of plasmids

Plasmids were constructed using standard cloning procedures, QuikChange® Site-Directed Mutagenesis (Stratagene), overlapping PCR and confirmed by DNA sequencing. Primers were obtained from IDT or Valuegene.

pHIS8.3-MBP-Z-wt and pHIS8.3-MBP-Z-N6C were constructed by insertion of the WT Z gene and Z-N6C gene into pHIS8.3-MBP plasmid using the following primers: Z-NcoI-F 5'-GAG CCA TGG CAC TAG TGT AGA CAA CAA ATT C-3' and Z-XhoI-R 5'-GAG CTC GAG CTT AGG CGC CTG AG-3'. pHIS8.3-MBP-Z-N6H, pHIS8.3-MBP-Z-N6K were derived from pHIS8.3-MBP-Z-wt using QuikChange® Site-Directed Mutagenesis Kit (Stratagene) using the following primer pairs, respectively: Z-N6K-Fw, 5'-GTA GAC AAC AAA TTC AAG AAA GAA CAA CAA AAC GCC-3', Z-N6K-Re, 5'-GGC GTT TTG TTG TTC TTT CTT GAA TTT GTT GTC TAC-3', Z-N6H-Fw, 5'-GTG TAG ACA ACA AAT TCC ACA AAG AAC AAC AAA AC-3', Z-N6H-Re, 5'-GTT TTG TTG TTC TTT GTG GAA TTT GTT-3'.

pBAD-Afb-36TAG, pBAD-Afb-WT, pBAD-Afb-30TAG-47C were constructed as previously described.[1] These genes were inserted into pBAD-HisC plasmid using the following primers: Afb-SpeI-Fw 5'-GAG TCA TGA CTA GTG TAG ACA ACA AAT TCA ACA AAG-3' and Afb-HindIII-Re 5'-GAG AAG CTT AGT GAT GGT GAT GGT GAT GAG-3'. pBAD-Afb-30TAG-47H and pBAD-Afb-30TAG-47K were derived from pBAD-Afb-Wt using QuikChange® Site-Directed Mutagenesis Kit (Stratagene) using the following primer pairs, respectively: Afb-30tag-47H-Fw, 5'-GC GCT AAC CTT TTA GCA CAC GCT AAA AAG CTA AAT GAT GC-3', Afb-30tag-47H-Re, 5'-GCA TCA TTT AGC

TTT TTA GCG TGT GCT AAA AGG TTA GCG C-3', Afb-30tag-47K-Fw, 5'-GC GCT AAC CTT TTA GCA AAA GCT AAA AAG CTA AAT GAT GC-3', Afb-30tag-47K-Re, 5'-GCA TCA TTT AGC TTT TTA GCT TTT GCT AAA AGG TTA GCG C-3'.

pBAD-Afb-32TAG and pBAD-Afb-45TAG were derived using QuickChange from pBAD-Afb-Wt with the following primer pairs, respectively: Af 32TAG-K-Fw, 5'-G AAG AAG GCC TTC ATC TAG AGT TTA TGG GAT GAC C-3', Af 32TAG-K-Re, 5'-GGT CAT CCC ATA AAC TCT AGA TGA AGG CCT TCT TC-3', Af 45TAG-K-Fw, 5'-C CAA AGC GCT AAC CTT TAG GCA GAA GCT AAA AAG C-3', Af 45TAG-K-Re, 5'-GCT TTT TAG CTT CTG CCT AAA GGT TAG CGC TTT GG-3'. pBAD-Afb-A42C-46TAG was derived from pBAD-Afb-Wt using the following primer pairs: ZAfb-A42C-Fw, 5'-GGG ATG ACC CAA GCC AAA GCT GCA ACC TTT TAT AGG AAG CTA AAA AGC-3', ZAfb-A42C-Re 5'-GCT TTT TAG CTT CCT ATA AAA GGT TGC AGC TTT GGC TTG GGT CAT CCC-3', ZAfb-A46TAG-Fw, 5'-GCC AAA GCG CTA ACC TTT TAT AGG AAG CTA AAA AGC TAA ATG ATG CTC-3', ZAfb-A46TAG-Re, 5'- GAG CAT CAT TTA GCT TTT TAG CTT CCT ATA AAA GGT TAG CGC TTT GGC-3'.

Plasmid pTak-Myo4TAG was constructed as previously described with *Spe* I and *Blp* I enzyme cutting sites.[2] pTAK-Afb-36TAG, pTAK-Afb-30TAG-47C, pTAK-Afb-30TAG -47H, and pTAK-Afb-30TAG-47K were constructed by insertion of the corresponding Afb genes from the pBAD plasmids into the pTAK vector using the following primers: pTAK-Afbody-SpeI-Fw: 5'-CG ATG ACT AGT GTA GAC AAC AAA TTC-3', pTAK-Afbody-BlpI-Re:5'- GTC CAA GCT CAG CTA ATT AAG CTT AGT GAT GGT GAT GG -3'. These plasmid were used for protein expression with pBK-MmBrC6KRS.

Plasmid pTAK-Afb-Her2 was constructed for expressing the Z_{HER2} affibody[4] with a C-terminal His6 tag. The sequence of Z_{HER2} is the following: atgactagtgtagacaacaaattcaac aaagaaatgcgtaacgcctactgggagatcgcactgctgccgaacctgaataatcagcagaaacgtgccttcatccgcagtctgtatg atgacccgagccaaagcgctaacctgctggcagaagctaaaaaactgaatgatgctcaggcgccgaaaggcgcgcatcaccatcac catcactaa.

pEVOL-pylT-2MmBrC6KRS plasmid was constructed by inserting two copies of the MmBrC6KRS gene into pEVOL-pylT plasmid using the following primers: Pev-SpeI-Fw, 5'-GTGC ACTAGT ATG GAT AAA AAG CCT CTG AAC ACT CTG -3', Pev-SalI-Re, 5'-CAGC GTCGAC TTA CAG GTT AGT AGA AAT ACC ATT GTA ATA GGA C -3', Pev-NdeI-Fw, 5'-GAATCC CAT ATG GAT AAA AAG CCT CTG AAC AC-3', Pev-NotI-R, 5'- CGTATA GCGGCCGC TTA CAG GTT AGT AGA AAT ACC ATT GTA ATA GGA C-3'. For expression of affibody proteins (Afb-32TAG, Afb-45TAG, and Afb-A42C-46TAG), plasmids pEVOL-pylT-2MmBrC6KRS and the corresponding pBAD plasmid were used.

Library construction and selection of MmBrC6KRS

To evolve an aminoacyl-tRNA synthetase specific for BrC6K, mutant synthetase libraries with 6 residues in the active site of the *Methanosarcina mazei* PylRS (MmPylRS) randomized were generated. Library JT-3+4 was generated as described in reference [2] with residues Ala302, Leu309, Asn346, Cys348, Val401, and Trp417 randomized together with the Tyr384Phe mutation. Library VL-5 was constructed as described in reference [3] with residues Leu309, Asn346, Cys 348, Val401 and Trp417 randomized, Ala302 mutated to

Ala/Thr, and Gly419 mutated to Gly/Ala/Ser. Three rounds of positive and two rounds of negative selections were performed on these two libraries using procedures previously described. In the first round of positive selection, the library was electroporated into *E. coli* DH10*β* competent cells harboring pREP-PylT to yield a cell library greater than 1 x 10^9 cfu, ensuring complete coverage of the library. Cells were plated on LB agar plates containing 25 μg/mL tetracycline (Tet), 50 μg/mL kanamycin (Kan), 68 μg/mL chloramphenicol (Cm) and 1 mM BrC6K. The negative selection was carried out on LB agar plates containing 50 μg/mL Kan, 100 μg/mL ampicillin (Amp) and 0.2% arabinose without BrC6K. After the third positive selection, 36 green fluorescent colonies were obtained from the JT-3+4 library and subsequently screened by streaking on minimal-medium plates supplemented with chloramphenicol (60 μg mL^{-1}) and in the presence or absence of 1 mM BrC6K. All of these colonies showed BrC6K-dependent survival in chloramphenicol (up to 160 μg mL^{-1} in rich medium). DNA sequencing of eight individual colonies revealed that they all converged on one mutant on the protein level, which contained mutation Y384F and was named MmBrC6KRS.

Sequence of MmBrC6KRS

atggataaaaagcctctgaacactctgatttctgcgaccggtctgtggatgtcccgcaccggcaccatccacaaaatcaaacaccatga

agttagccgttccaaaatctacattgaaatggcttgcggcgatcacctggttgtcaacaactcccgttcttctcgtaccgctcgcgcactg

cgccaccacaaatatcgcaaaacctgcaaacgttgccgtgttagcgatgaggacctgaacaaattcctgaccaaagctaacgaggatc

agacctccgtaaaagtgaaggtagtaagcgctccgacccgtactaaaaaggctatgccaaaaagcgtggcccgtgccccgaaacct

ctggaaaacaccgaggcggctcaggctcaaccatccggttctaaattttctccggcgatcccagtgtccacccaagaatctgtttccgta

ccagcaagcgtgtctaccagcattagcagcatttctaccggtgctaccgcttctgcgctggtaaaaggtaacactaacccgattactagc

atgtctgcaccggtacaggcaagcgccccagctctgactaaatcccagacggaccgtctggaggtgctgctgaacccaaaggatga

aatctctctgaacagcggcaagcctttccgtgagctggaaagcgagctgctgtctcgtcgtaaaaaggatctgcaacagatctacgctg

aggaacgcgagaactatctgggtaagctggagcgcgaaattactcgcttcttcgtggatcgcggtttcctggagatcaaatctccgattc

tgattccgctggaatacattgaacgtatgggcatcgataatgataccgaactgtctaaacagatcttccgtgtggataaaaacttctgtctg

cgtccgatgctggccccgaacctgtacaactatctgcgtaaactggaccgtgccctgccggacccgatcaaaattttcgagatcggtcc

ttgctaccgtaaagagtccgacggtaaagagcacctggaagaattcaccatgctgaacttctgccagatgggtagcggttgcacgcgt

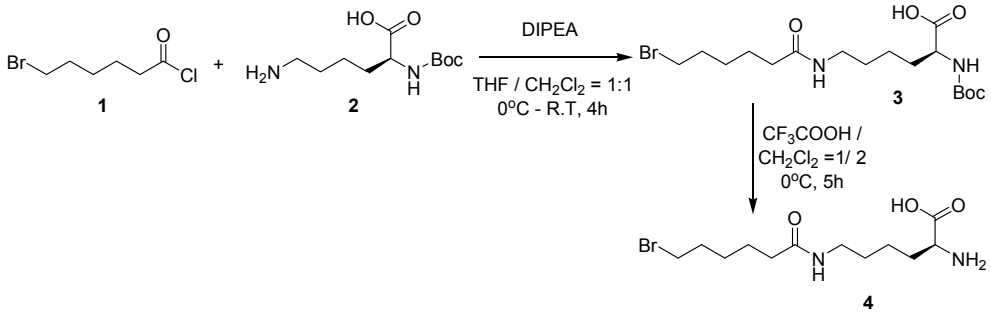
gaaaacctggaatccattatcaccgacttcctgaatcacctgggtatcgatttcaaaattgttggtgacagctgtatggtgtttggcgatac

gctggatgttatgcacggcgatctggagctgtcttccgcagtagtgggcccaatcccgctggatcgtgagtggggtatcgacaaacctt

ggatcggtgcgggttttggtctggagcgtctgctgaaagtaaaacacgacttcaagaacatcaaacgtgctgcacgttccgagtcctatt

acaatggtatttctactaacctgtaa

Synthesis of unnatural amino acid BrC6K



Preparation of (S)-2-amino-6-(6-bromohexanamido)hexanoic acid (BrC6K, **4**):

N_{α}-Boc-L-lysine(10.0 mmol, 2.46g) was dissolved in 200 ml mixture of THF/CH_2Cl_2 (1/1) and DIEA (12.0 mmol, 2.1 mL) at 0 °C; 6-bromohexanoyl chloride (13 mmol, 2.0 mL) was added. The reaction mixture was stirred at 0℃ to room temperature for a period of 4 h. A solution of 1 N HCl (30 mL) was then added and the solution extracted with EtOAc (80 mL,

x4). The organic phase was dried with $MgSO_4$, filtered and evaporated to a brown oil. The crude material was purified by flash chromatography using a solvent gradient from 50:1 to 10:1 EtOAC/MeOH. The product was isolated as a yellowish oil (**3**, 3.2 g, 75% yield). The purifide product **3** was dissolved in dichloromethane (15 mL). Trifluoroacetic acid (7.5 mL) was added to the mixture at 0 °C, the reaction mixture was stirred for 5 h at the same temperature. After the reaction completed, the mixture was concentrated under reduced pressure. The residue was dissolved in methanol, and precipitated in Et_2O. The white solid was washed with Et_2O to give the final product **4** as the trifluoroacetic acid salt (**4**, 2.8 g, 86% yield). BrC6K (**4**), 1H NMR (700 MHz, DMSO-D6): 3.88 (t, *J* = 5.6 Hz, 1H), 3.51 (t, *J* = 7.0 Hz, 2H), 3.01 (t, *J* = 7.0 Hz, 2H), 2.04 (t, *J* = 6.3 Hz, 2H), 1.30-1.79 (m, 12H). ESI-MS: calcd for C12H23BrN2O3: 323.0964 $[M+H]^+$, found 323.0937 $[M+H]^+$. BrC6K was stored at -20 °C and dissolved in sterile 1 N NaOH aqueous solution before use.

Protein expression and Uaa incorporation in *E. coli*

To incorporate BrC6K into myoglobin at position 4, *E. coli* BL21(DE3) cells were cotransformed with plasmids pTak-Myo4TAG and pBK-MmBrC6KRS. A single colony was picked and grown overnight in 5 mL 2xYT supplemented with 34 μg mL^{-1} Cm and 50 μg mL^{-1} Kan at 37 °C. This culture was used to inoculate 200 mL of 2xYT containing 34 μg mL^{-1} Cm and 50 μg mL^{-1} Kan. When OD_{600} reached 0.5–0.6, nicotinamide (final concentration 5 mM) and 2 mM of BrC6K were added at 30 °C. After 30 min, cells were induced for protein expression by adding 0.5 mM IPTG at 30 °C. After 8 h induction, cells were pelleted and frozen in –80 °C. The cell pellet was thawed on ice for 30 min and

resuspended in 15 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, 1% (v/v) Tween 20,). Lysozyme (0.5 mg mL^{-1}), DNase I (0.5 μg mL^{-1}) and protease inhibitor (Cocktail, Roche) were added; the mixture was incubated at 4 °C for 30 min, and then subjected to sonication for 30 s X4 (0.5 s On, 1.5 s Off) on ice. The lysate was centrifuged at 16,000 g for 30 min. Ni^{2+}-NTA agarose resin (120 μL, 50%, Qiagen) was added to the supernatant and mix gently by shaking at 4 °C for 1 h. The lysate-Ni-NTA mixture was loaded into column and washed with 0.5 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol). Protein was eluted with 0.5 mL of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, 10% (v/v) glycerol). The purified protein was analyzed by 15% Glycine-SDS-PAGE.

For expression of mutated affibody proteins and the HER2-affibody-37TAG, *E. coli* BL21(DE3) cells cells were cotransformed with pTak derived plasmid and pBK-MmBrC6KRS, or pBAD derived plasmid and pEVOL-pylT-2MmBrC6KRS plasmid. To incorporate BrC6K into the affibody (36TAG), *E. coli* BL21(DE3) cells cells were cotransformed with pTAK-Afb-36TAG and pBK-MmBrC6KRS. A single colony was picked and grown overnight in 5 mL 2xYT supplemented with 34 μg/mL Cm and 50 μg/mL Kan at 37 °C. This culture was used to inoculate 200 mL of 2x YT containing antibiotics. When OD_{600} reached 0.5-0.6, nicotinamide (final concentration 5mM) and 2 mM of BrC6K were added at 30 °C. Cells were induced 30 min later for protein expression by adding 0.5 mM IPTG at 30°C. After 5 h induction, cells were harvested and proteins purified using a similar procedure as described in the myoglobin section. The mutated affibody proteins were exchanged to 50 mM sodium phosphate buffer with 150 mM NaCl at pH 7.4 using PD-10

column.

To express the MBP-Z, *E. coli* BL21(DE3) cells were transformed with plasmid pHIS8.3-MBP-Z-wt, pHIS8.3-MBP-Z-N6C, pHIS8.3-MBP-Z-N6H, pHIS8.3-MBP -Z-N6K. A single colony was picked and grown overnight in 25 mL 2xYT supplemented with 50 μg/mL Kan at 37 °C. This culture was used to inoculate 1L of 2x YT containing antibiotics. When OD_{600} reached 0.5-0.6, cells were induced for protein expression by 0.5 mM IPTG at 30 °C. After 5 h induction, cells were pelleted and frozen in -80°C overnight. The purification procedure is the same as that for myoglobin section. The purified protein was exchanged to PBS (pH 7.4) using the PD-10 column.

FITC label of Z_{HER2} proteins

To fluorescently label the Z_{HER2} proteins, wt or BrC6K Z_{HER2} protein was added in 500 μL of the mixture buffer (200 mM sodium bicarbonate, pH 9.0) at final concentration of 70 μM. The lysine-reactive FITC dye (Fluorescein-5-Isothiocyanate, Life Technologies) was dissolved in DMF (~10 mg/mL), and 50 uL of this dye was slowly added to the protein solution at room temperature. After incubating at room temperature in dark with gentle shaking for 2 h, the reaction mixture was desalted twice using the PD-10 column (GE) to remove unreacted FITC dye, and the FITC-labelled Z_{HER2} was eluted with PBS buffer and analyzed by SDS-PAGE.

Cell Culture

HER2 positive cells (SKBR3 in McCoy's 5A medium) and HER2 negative cells (MDA-MB-468 in DMEM medium) were cultured in indicated medium supplemented with

10% (v/v) fetal bovine serum (FBS), penicillin (100 units mL^{-1}) and streptomycin (100 μg mL^{-1}). Cells were incubated at 37 ^{o}C in 5% CO_2 and passaged 2–3 times a week.

Cell Labeling and Imaging

The FITC-labeled wt and BrC6K-containing Z_{HER2} were separately added to SKBR3 and MDA-MB-468 cells (~70% confluent) at the concentration of 300 nM. Cells were cultured at 37 ^{o}C with 5% CO_2 in the dark for 3 h, after which the cells were washed using gentle conditions (DPBS buffer, pH 7.5, room temperature, 10 min, twice) or stringent conditions (500 mM NaCl, 3% tween 20, 100 mM glycine, pH 3.0, room temperature, 10 min, twice). The FITC fluorescence was visualized at room temperature on an Zeiss LSM 710 Confocal Laser Scanning Microscope using a 40x objective (oil). The filters of the FITC channel were 488/30 nm for excitation and 514/40 for emission.

Western Blotting

SKBR3 cells were incubated at 37 ^{o}C for 3 h with wt-Z_{HER2} or BrC6K-Z_{HER2} (both FITC-free). The cells were washed twice with cold PBS buffer, and then incubated in cold RIPA Buffer (500 μL to 10 cm dish, with protease inhibitor cocktail) on ice for 1 hr. The cells were lysed, and the lysates were centrifuged at 14,000 x g for 10 min at 4 °C to remove insoluble materials. The supernatant was incubated with Laemmli loading buffer containing mercaptoethanol for 1 h at 37 ^{o}C and resolved by SDS-PAGE on a 7-15% Tricine gel. Proteins were transferred onto a PVDF membrane in Tris-Gly transfer buffer containing 20% methanol for ~2 h at 100 V (4 °C). The membrane was blocked for 1 h in 5% nonfat milk in TBST with 0.1% Tween-20, and then incubated overnight with rabbit anti-HER2 antibody (1:1000, Cell signaling, HER2/ErbB2 Antibody #2242) in 5% BSA in TBST (4°C). After 3 x 5 min washes with TBST, the membrane was incubated with Donkey anti-rabbit-HRP conjugate (horseradish peroxidase, Biopioneer 1:5000) for 1 h at RT and washed again (3 x 10 min). In a similar manner, Penta-His-HRP conjugate (Qiagen, 1:5000) and mouse anti-β-actin-HRP conjugate (Santa Cruz Biotechnology, 1:5000) were applied to the membrane for detecting Z_{HER2}, which has a His6 tag at the C-terminus.

Electrochemiluminescence was achieved SuperSignal West Pico Chemiluminescent substrate (5 min incubation, Thermo Scientific). The membrane was exposed to autoradiography film (Denville Scientific, Inc., Metuchen, NJ) for a few seconds to several minutes.

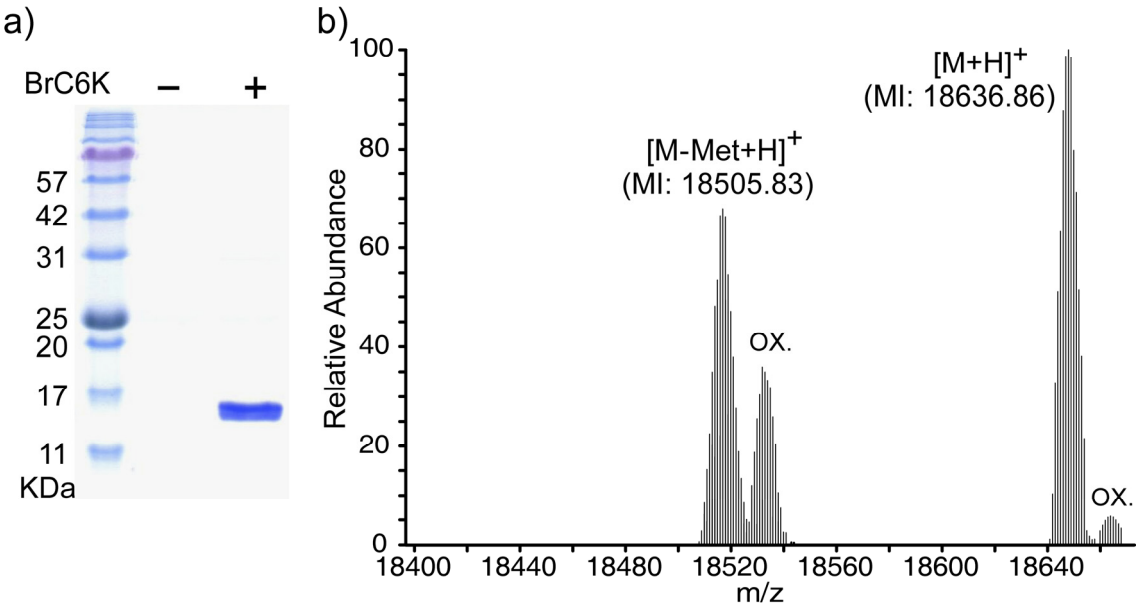


Figure S1. Incorporation of BrC6K into myoglobin in *E. coli*. a) SDS-PAGE analysis of myoglobin (Myo-4TAG-His6) expressed using the $tRNA^{Pyl}_{CUA}$-MmBrC6KRS in the presence of 2 mM BrC6K followed by Ni^{2+} affinity chromatography. b) High resolution ESI-FTMS of the intact myoglobin produced with BrC6K in (a). A peak with a monoisotopic mass of 18636.86 Da was observed, which corresponds to intact myoglobin containing a single BrC6K residue at position 4 (expected $[M+H]^+$ = 18636.77 Da). A second observed peak corresponds to the BrC6K-containing myoglobin without the initiator Met residue (expected $[M\text{-}Met+H]^+$ = 18505.73 Da, measured 18505.83 Da). No peaks were observed corresponding to myoglobin proteins containing any natural amino acid at the TAG position. These results indicate that only BrC6K was incorporated into myoglobin at the TAG-encoded position. Peaks corresponding to oxidation (ox.) of BrC6K-myoglobin were detected and labeled.

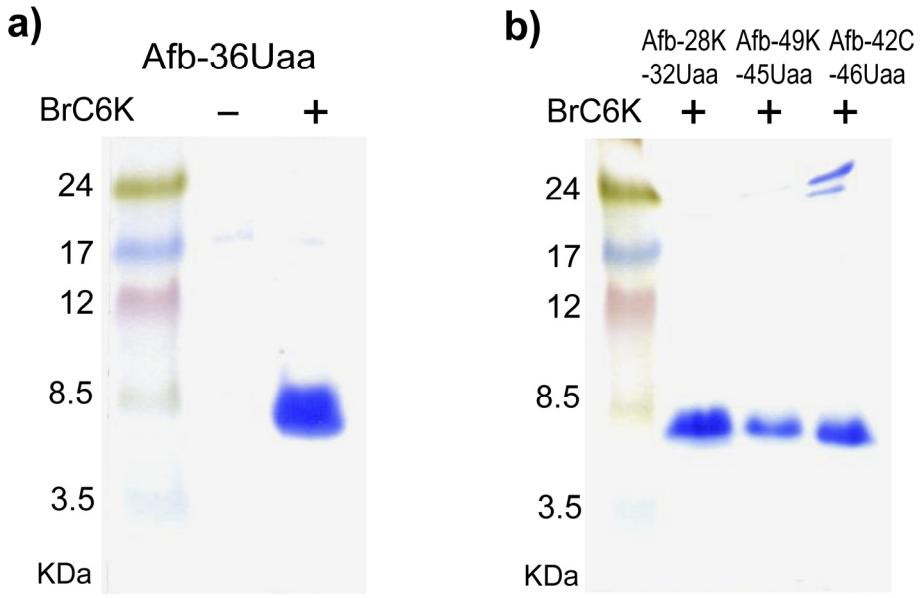


Figure S2. SDS-PAGE analysis of affibody proteins expressed using the $tRNA^{Pyl}_{CUA}$-MmBrC6KRS in the presence of 2 mM BrC6K followed by Ni^{2+} affinity chromatography. a) Afb-36Uaa, b) Afb-28K-32Uaa, Afb-49K-45Uaa and Afb-42C-46Uaa.

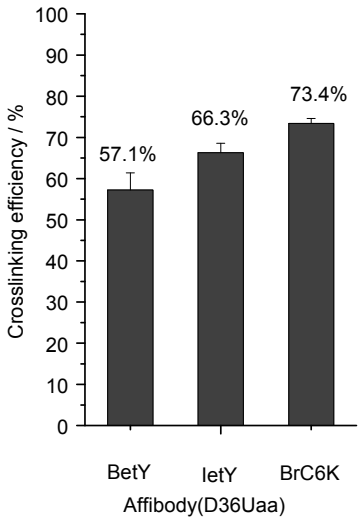


Figure S3. BrC6K showed higher crosslinking efficiency with Cys than BetY and IetY in intermolecular PEPC. The purified Afb (D36Uaa) proteins were incubated with purified MBP-Z (N6C) in a 4:1 ratio in PBS buffer (pH 7.4) at 37 °C for 1 h. Crosslinking efficiency was calculated based on the relative band intensities of the Afb–Z crosslinking complex and the MBP–Z fusion protein in SDS-PAGE (see Figure 1b). Mean values are shown with error bars representing standard deviations from four independent measurements.

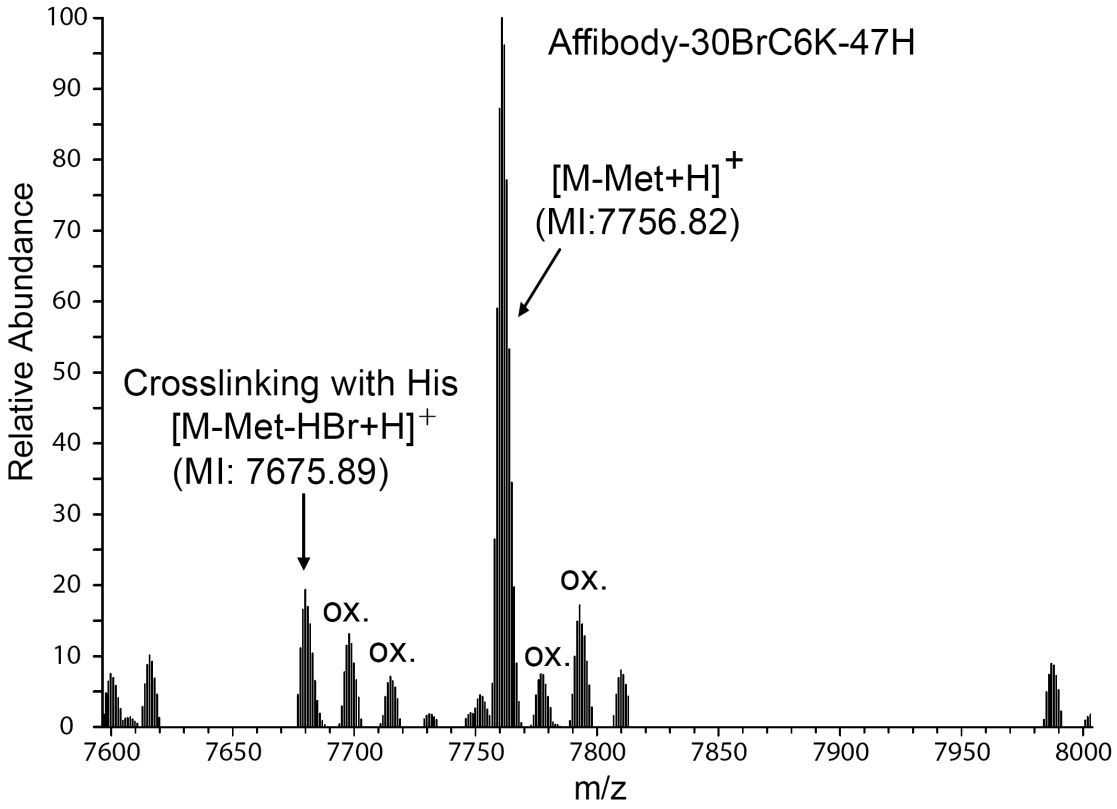


Figure S4. BrC6K is able to target His in intramolecular PEPC. High resolution ESI-FTMS analysis of affibody (Afb-30BrC6K-47H) incubated at pH 8.0 for 6 h. Peaks were observed with monoisotopic masses corresponding to Afb containing the covalent bond formed between the Uaa BrC6K and the target His, which results in the loss of HBr. For BrC6K-His crosslinked product [M-Met-HBr+H], expected 7675.98 Da, measured 7675.89 Da. Based on the relative abundance of the crosslinking and non-crosslinking peaks (including their oxidation peaks), the intra-protein crosslinking yield of His with BrC6K was 23% at pH 8.0.

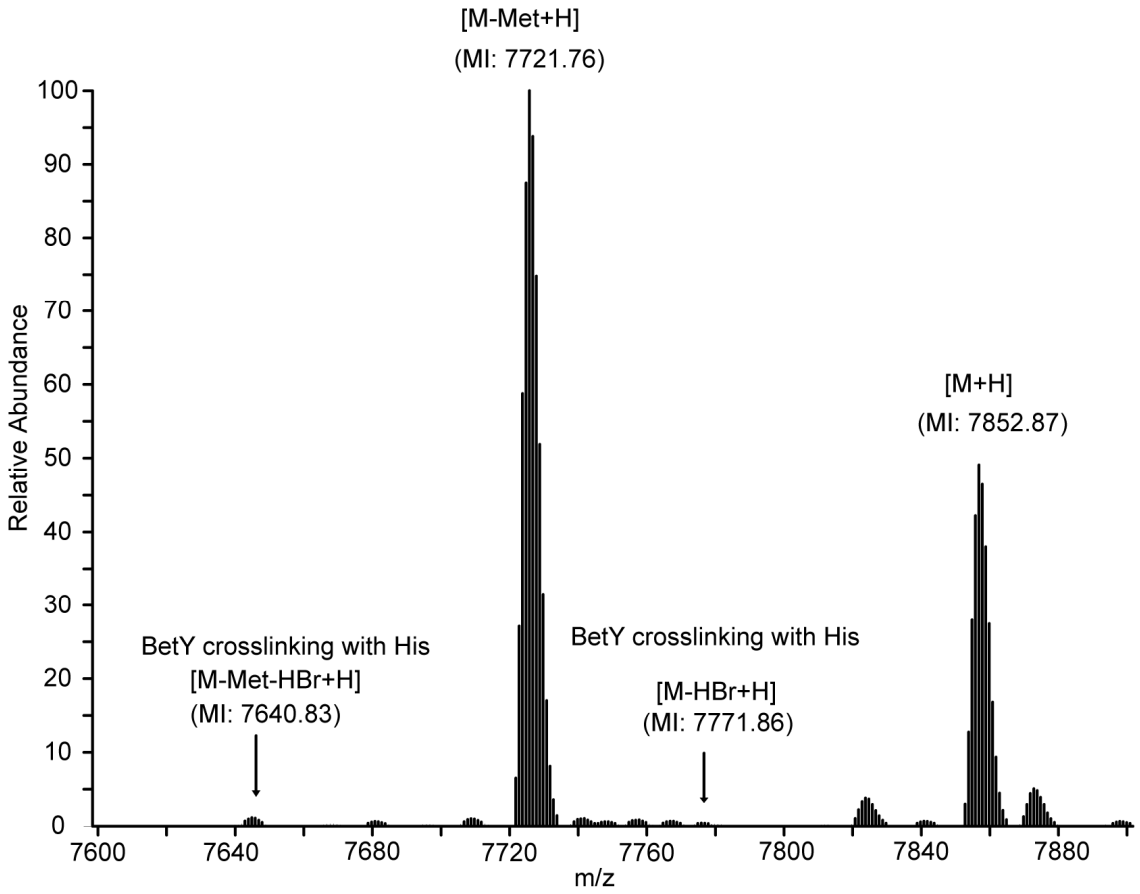


Figure S5. BetY targets His in intramolecular PEPC with low efficiency. High resolution ESI-FTMS analysis of affibody (Afb-30BetY-47H) incubated at pH 8.8 for 6 h. Peaks were observed with monoisotopic masses corresponding to Afb containing the covalent bond formed between the Uaa BetY and the target His, which results in the loss of HBr. For BetY-His crosslinked product [M-Met-HBr+H], expected 7640.91 Da, measured 7640.83 Da. Based on the relative abundance of the crosslinking and non-crosslinking peaks (including their oxidation peaks), the intra-protein crosslinking yield of His with BetY was 2% at pH 8.8.

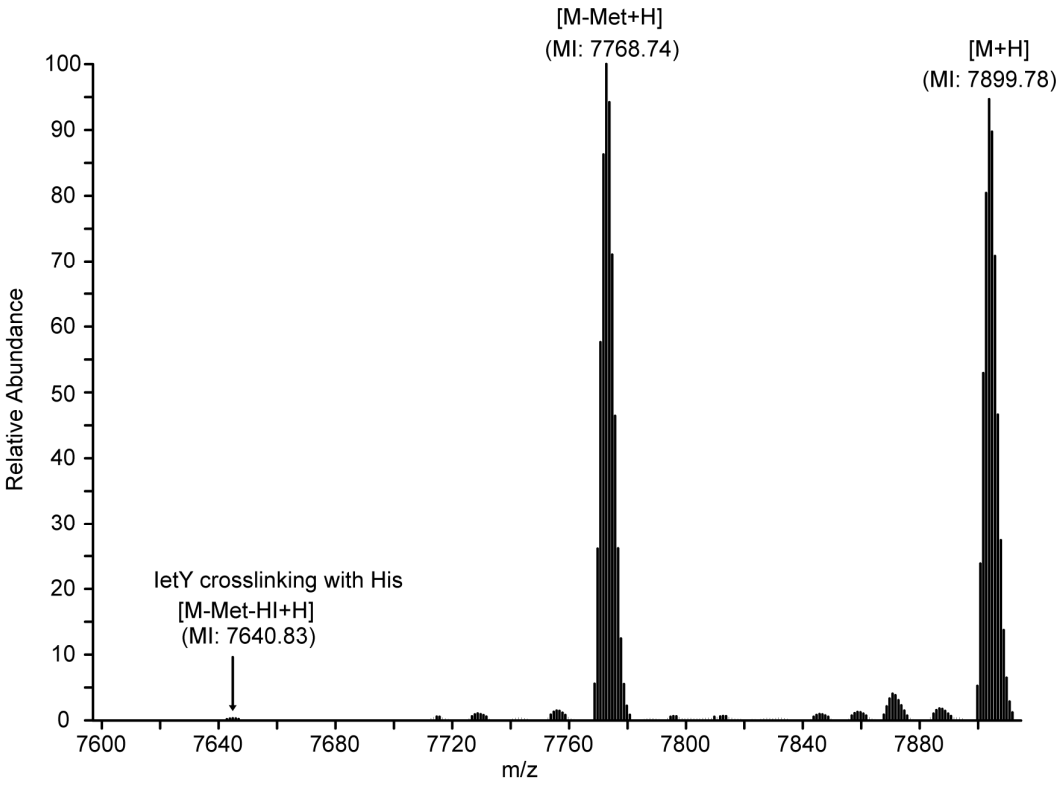


Figure S6. IetY targets His in intramolecular PEPC with low efficiency. High resolution ESI-FTMS analysis of affibody (Afb-30IetY-47H) incubated at pH 8.8 for 6 h. Peaks were observed with monoisotopic masses corresponding to Afb containing the covalent bond formed between the Uaa IetY and the target His, which results in the loss of HI. For IetY-His crosslinked product [M-Met-HI+H], expected 7640.91 Da, measured 7640.83 Da. Based on the relative abundance of the crosslinking and non-crosslinking peaks (including their oxidation peaks), the intra-protein crosslinking yield of His with IetY was 1% at pH 8.8.

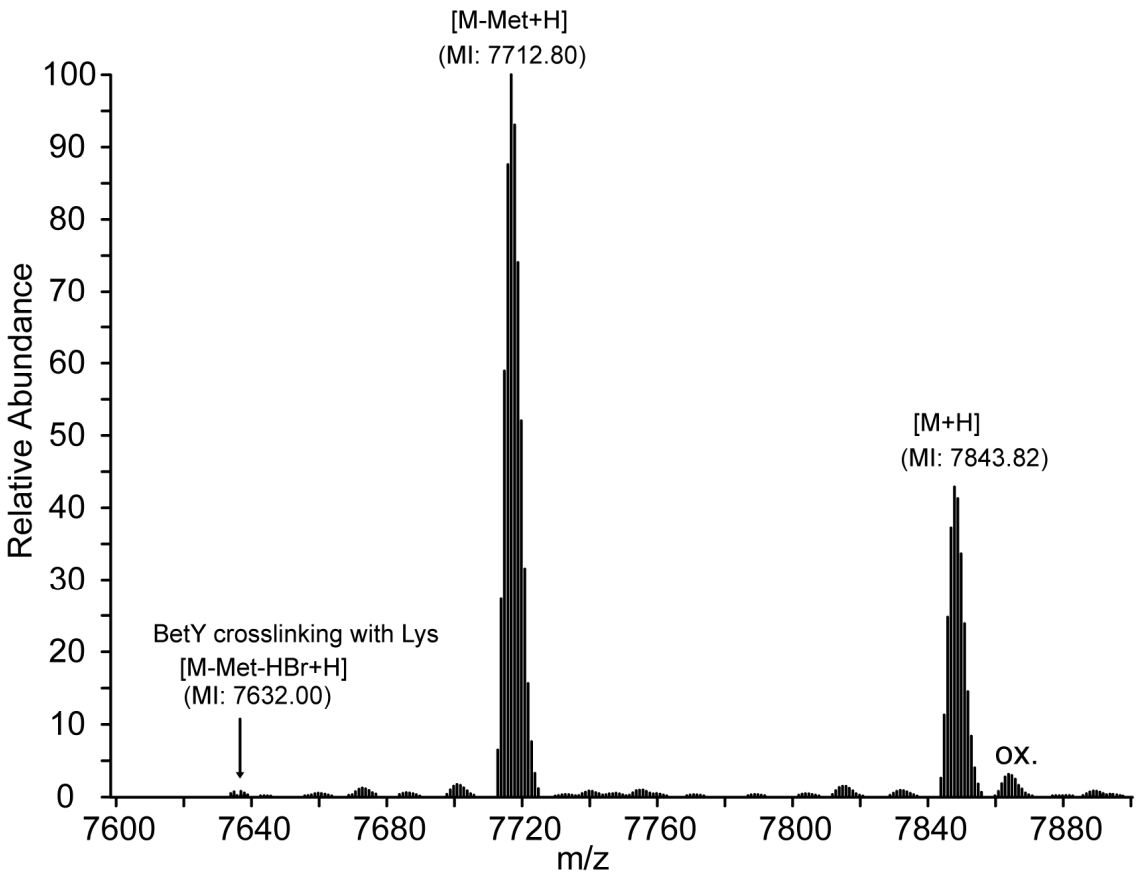


Figure S7. BetY targets Lys in intramolecular PEPC with low efficiency. High resolution ESI-FTMS analysis of affibody (Afb-30BetY-47K) incubated at pH 8.8 for 6 h. Peaks were observed with monoisotopic masses corresponding to Afb containing the covalent bond formed between the Uaa BetY and the target Lys, which results in the loss of HBr. For BetY-Lys crosslinked product [M-Met-HBr+H], expected 7631.94 Da, measured 7632.00 Da. Based on the relative abundance of the crosslinking and non-crosslinking peaks (including their oxidation peaks), the intra-protein crosslinking yield of Lys with BetY was 1% at pH 8.8.

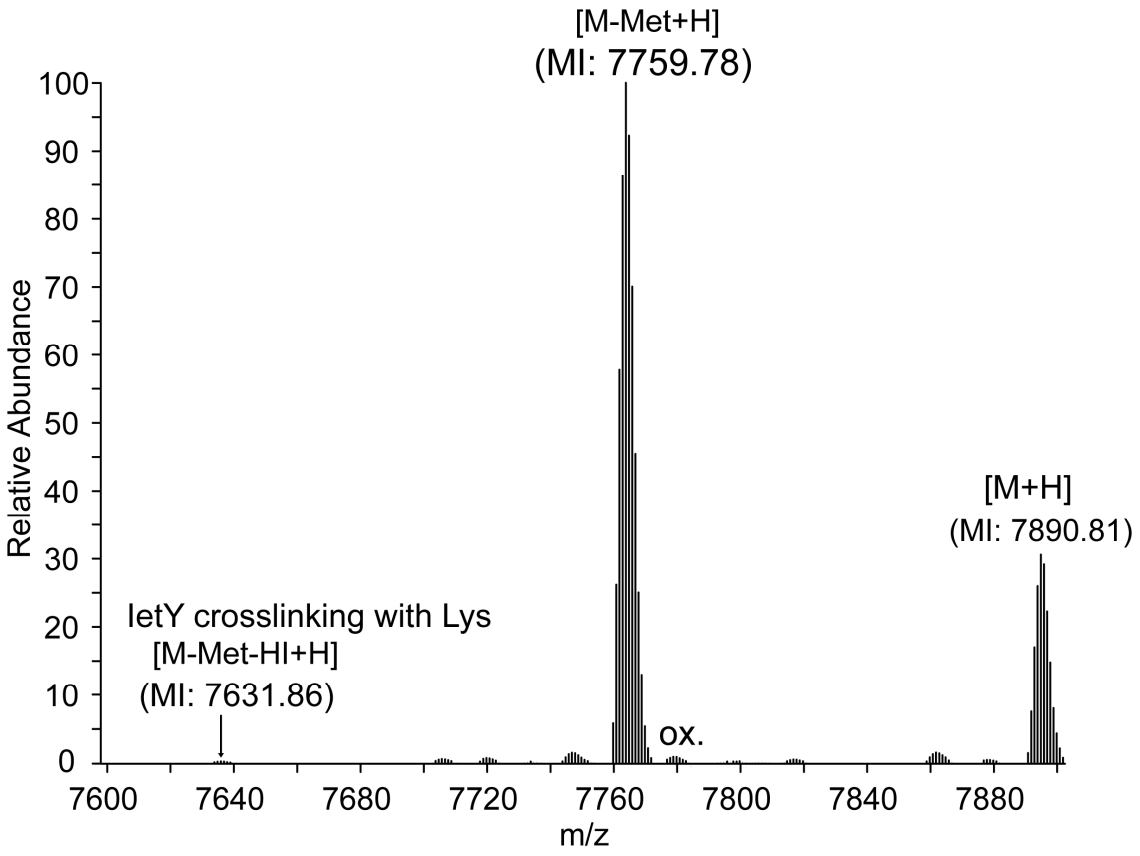


Figure S8. IetY targets Lys in intramolecular PEPC with low efficiency. High resolution ESI-FTMS analysis of affibody (Afb-30IetY-47K) incubated at pH 8.8 for 6 h. Peaks were observed with monoisotopic masses corresponding to Afb containing the covalent bond formed between the Uaa IetY and the target Lys, which results in the loss of HI. For IetY-Lys crosslinked product [M-Met-HI+H], expected 7631.94 Da, measured 7631.86 Da. Based on the relative abundance of the crosslinking and non-crosslinking peaks (including their oxidation peaks), the intra-protein crosslinking yield of Lys with IetY was 1% at pH 8.8.

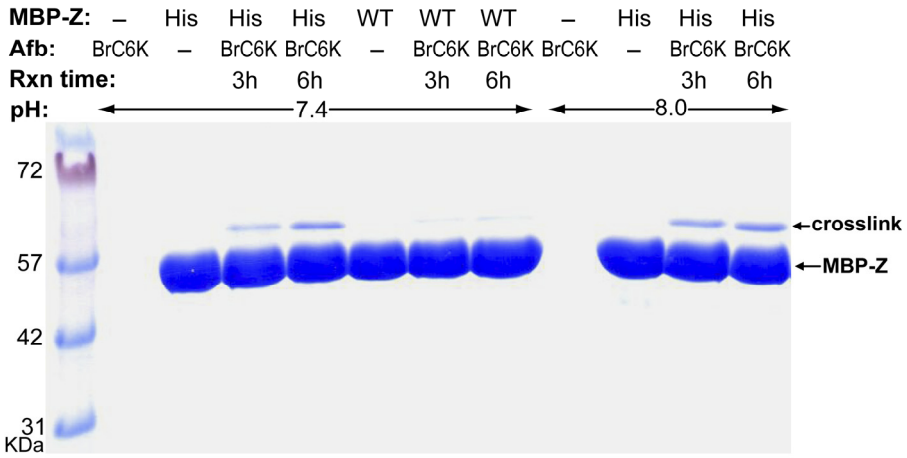


Figure S9. BrC6K reacts with His in intermolecular PEPC. SDS-PAGE analysis of Afb–Z crosslinking. The identities of residue 6 of the Z protein in MBP–Z and residue 36 of the Afb are indicated. See Figure 1a for positions. The purified Afb (D36Uaa) proteins were incubated with purified MBP-Z (N6H) in a 4:1 ratio at 37 °C in PBS buffer.

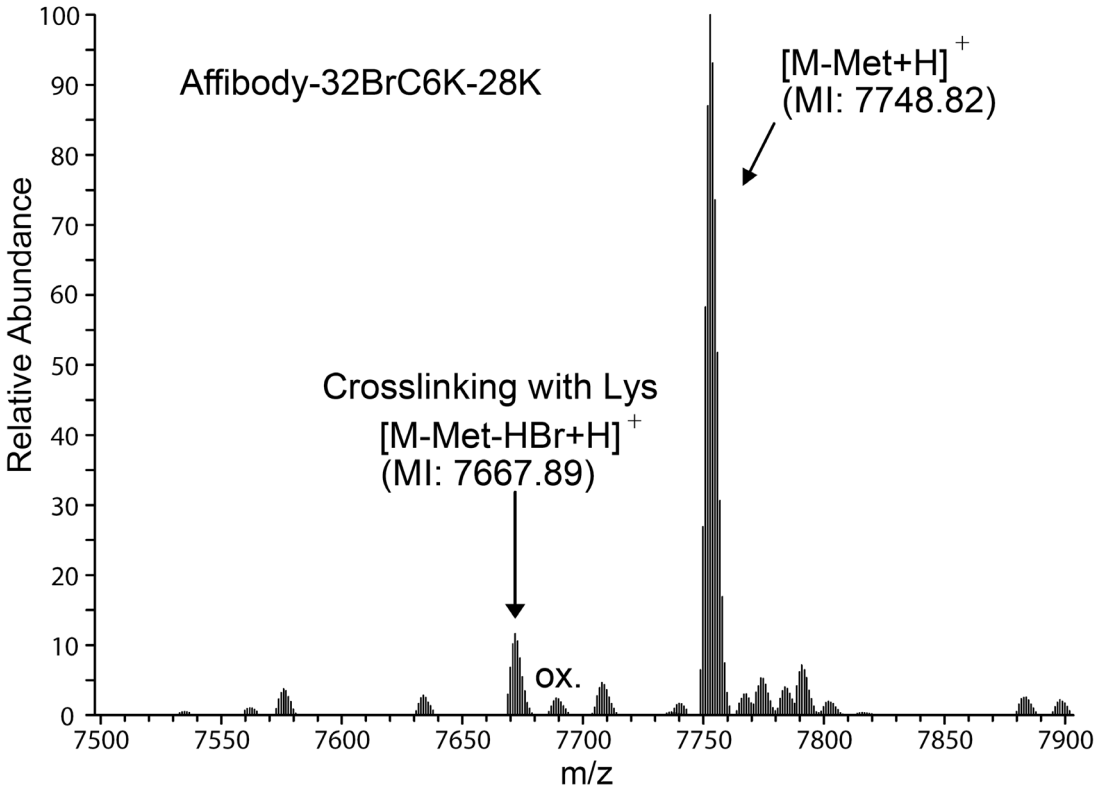


Figure S10. High resolution ESI-FTMS analysis of Afb-32BrC6K-28K indicating that BrC6K reacted with Lys enabling protein stapling. Purified Afb-32BrC6K-28K was incubated in PBS pH 8.8 for 6 h and then subjected to ESI-FTMS analysis. Stapled products: [M-Met-HBr+H], expected 7667.96 Da, measured 7667.89 Da.

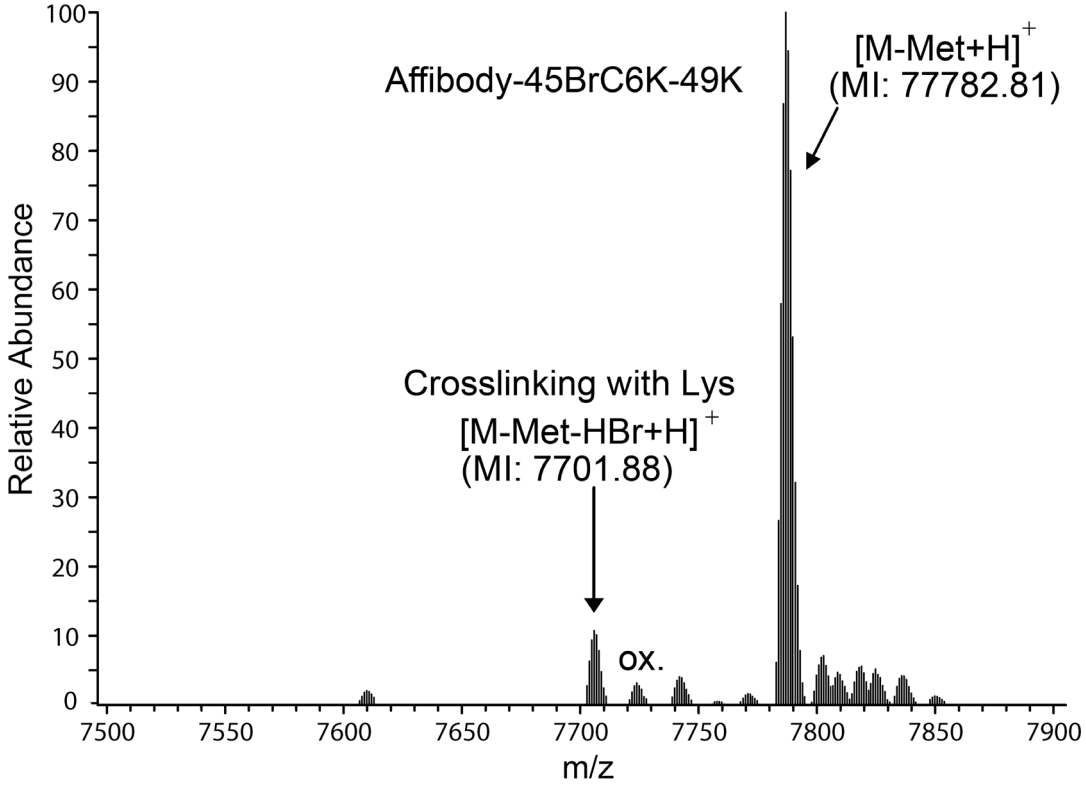


Figure S11. High resolution ESI-FTMS analysis of Afb-45BrC6K-49K confirming BrC6K reacting with Lys for protein stapling. Purified Afb-45BrC6K-49K was incubated in PBS pH 8.8 for 6 h and then subjected to ESI-FTMS analysis. Stapled products: [M-Met-HBr+H], expected 7701. 94 Da, measured 7701.88 Da.

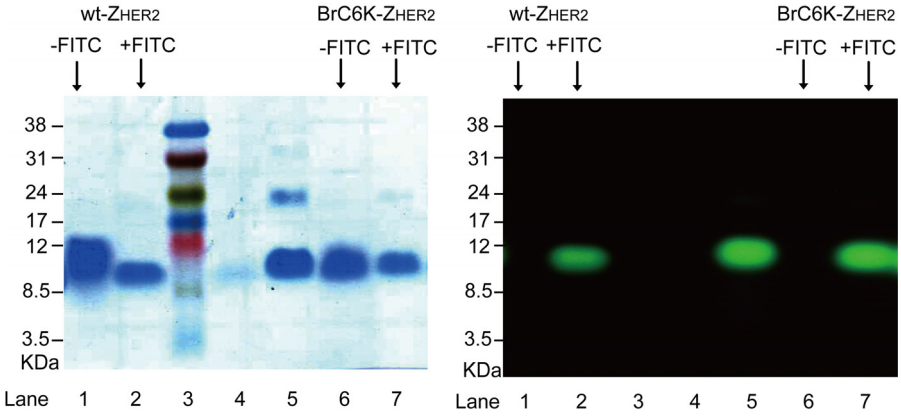


Figure S12. SDS-PAGE analysis of wt and BrC6K-incorporated (at Asp37) Z_{HER2} proteins. Left, SDS-PAGE gel stained with Coomassie blue; Right, fluorescence image of the SDS-PAGE gel showing FITC label. Lane 5 is Z_{HER2} with BrC6K incorporated at a different site (residue 36) and labeled by FITC, which did not covalently bind to HER2. Lane 4 is overflow of lane 5.

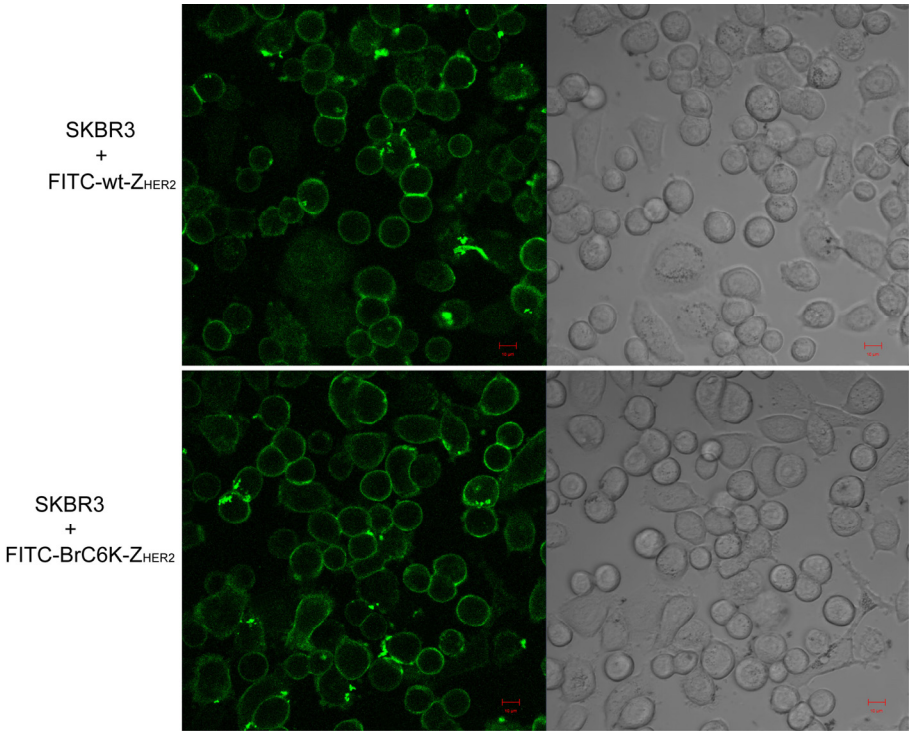


Figure S13. The FITC-labeled wt-Z_{HER2} (top) and BrC6K- Z_{HER2} (bottom) both bound to the HER2 receptor on SKBR3 cells after gentle wash. Cells were incubated with the Z_{HER2} proteins and washed in gentle conditions (DPBS for 10 min at RT, 2x) to remove the unbound and nonspecifically bound Z_{HER2}. Left, FITC channel; right, bright field of the same region. Scale bar: 10 μm.

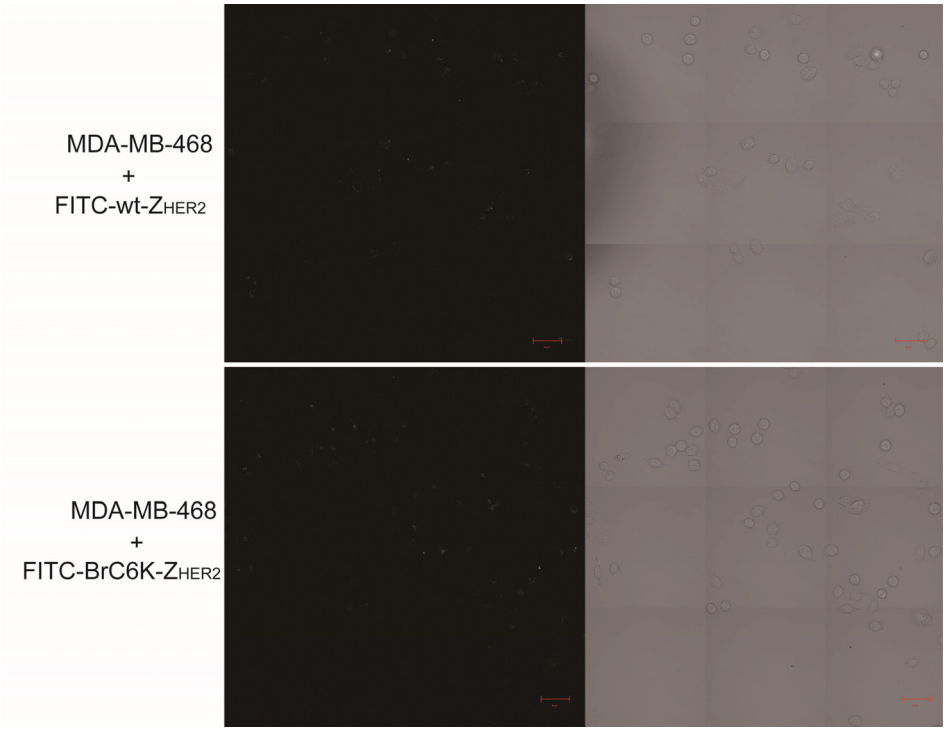


Figure S14. The FITC-labeled wt-Z_{HER2} (top) and BrC6K- Z_{HER2} (bottom) did not bind to the HER2 negative MDA-MB-468 cells after gentle wash. Cells were incubated with the Z_{HER2} proteins and washed in gentle conditions (DPBS for 10 min at RT, 2x) to remove the unbound and nonspecifically bound Z_{HER2}, as done in Figure S11. Left, FITC channel; right, bright field of the same region. Scale bar: 50 μm

Supporting references

[1] Xiang, Z. Lacey, V. K. Ren, H. Xu, J. Burban, D. J. Jennings, P. A. Wang, L. *Angew. Chem. Int. Ed. Engl.* **2014**, *126*, 2222-2225.

[2] Takimoto, J. K. Dellas, N. Noel, J. P. Wang, L. *ACS Chem. Biol.* **2011**, *6*, 733-743.

[3] Lacey, V. K. Louie, G. V. Noel, J. P. Wang, L. *ChemBioChem.* **2013**, *14*, 2100-2105.

[4] Eigenbrot, C. Ultsch, M. Dubnovitsky, A. Abrahmsen, L. Hard, T. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 15039-15044.