Article

A genetically encoded aza-Michael acceptor for covalent crosslinking of protein-receptor complexes

Jennifer L. Furman, Mingchao Kang, Seihyun Choi, Yu Cao, Erik D Wold, Sophie B Sun, Vaughn V Smider, Peter G Schultz, and Chan Hyuk Kim

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/ja502851h • Publication Date (Web): 20 May 2014

A genetically encoded aza-Michael acceptor for covalent crosslinking of protein-receptor complexes

Jennifer L. Furman[1†], Mingchao Kang[1,2†], Seihyun Choi[1], Yu Cao[1], Erik D. Wold[1], Sophie B. Sun[1,2], Vaughn V. Smider[1], Peter G. Schultz[1,2]*, and Chan Hyuk Kim[2]*

[1]Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

[2]California Institute for Biomedical Research, 11119 North Torrey Pines Road Suite 100, La Jolla, California 92037, United States

[†]These authors contributed equally to this work.

*e-mail: schultz@scripps.edu; chkim@calibr.org

ABSTRACT: Selective covalent bond formation at a protein-protein interface potentially can be achieved by genetically introducing into a protein an appropriately "tuned" electrophilic unnatural amino acid that reacts with a native nucleophilic residue in its cognate receptor upon complex formation. We have evolved orthogonal aminoacyl-tRNA synthetase/$tRNA_{CUA}$ pairs that genetically encode three aza-Michael acceptor amino acids, N^{ϵ}-acryloyl-(*S*)-lysine (AcrK, **1**), *p*-acrylamido-(*S*)-phenylalanine (AcrF, **2**), and *p*-vinylsulfonamido-(*S*)-phenylalanine (VSF, **3**), in response to the amber stop codon in *Escherichia coli*. Using an αErbB2 Fab-ErbB2 antibody-receptor pair as an example, we demonstrate covalent bond formation between an αErbB2-VSF mutant and a specific surface lysine ε-amino group of ErbB2, leading to near quantitative crosslinking to either purified ErbB2 *in vitro* or to native cellular ErbB2 at physiological pH. This efficient biocompatible reaction may be useful for creating novel cell biological probes, diagnostics, or therapeutics that selectively and irreversibly bind a target protein *in vitro* or in living cells.

Introduction

One approach to the design of high affinity, selective small-molecule enzyme inhibitors involves the introduction of an electrophilic group into the inhibitor that can form a covalent bond with an active site nucleophile.[1] Selectivity is achieved through proximity of the two reactive groups in the enzyme-inhibitor complex, or in the case of mechanism-based inhibitors, by unmasking of the electrophilic moiety via enzymatic transformation.[2,3] Many covalent inhibitors of hydrolases, kinases, oxido-reductases, and PLP-dependent enzymes have been developed based on fluoroketones, α-haloketones, epoxides, α,β-unsaturated ketones and amides, and vinyl sulfonamides.[4-7] Indeed, a significant fraction of clinically used drugs that act on enzymes are irreversible inhibitors,[8] with recent examples including inhibitors of the 20S proteosome[9] and Bruton's tyrosine kinase (BtK).[10]

The application of similar medicinal chemistry-based strategies to form irreversible protein-based agonists, antagonists, or inhibitors is made difficult by the lack of general strategies for selectively introducing electrophiles into proteins at a specific site of interest.[11,12] Current methods include reaction of unique cysteine or lysine residues with bifunctional *N,N′*-ethylene bisacrylamide[13] or chloroacetyl chloride[14] electrophiles to introduce acrylamides or chloroacetamides, respectively. Additionally, expressed protein ligation (EPL) has been used to label proteins with vinyl sulfones and β-lactams to generate activity-based protein profiling (ABPP) probes.[15] Alternatively, the use of evolved orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pairs for genetically encoding unnatural amino acids (UAAs) provides a potentially general method for introducing electrophilic groups into proteins with precise control over the position and geometry of the reactive group. Thus far, a variety of electrophiles, including alkyl and aryl ketones,[16,17] trifluoroketones,[18] α-fluoroketones,[19] and acrylamides[20,21] have been introduced into proteins by this method. The use of these UAAs to form intra- or intermolecular covalent bonds typically requires a cysteine residue in the receptor of interest due to its relatively high nucleophilicity. However, cysteine residues are uncommon at protein surfaces and frequently found in cysteine crosslinks,[22] necessitating mutagenesis of the target protein to introduce a unique cysteine residue at the binding interface. Instead, an ideal group to target with these electrophiles is the nucleophilic ε-amino group of lysine residues,[23] which are relatively abundant at protein surfaces and likely to be found at the protein-protein interface of interest.[24] The practical application of this strategy requires that the reactivity of the electrophile be tuned so it has minimal nonspecific reactions with common physiological nucleophiles (such as glutathione), while maintaining the ability to react with a lysine residue in a binding partner.[25] This can be achieved through 'proximity effects', in which the formation of a covalent bond is accelerated by formation of a noncovalent complex prior to reaction.[26] This leads to a reduction in the entropy of activation which can result in large increases in reaction rates (effective molarities $>10^5$ M).[27]

Herein we have investigated a series of electrophilic UAAs which contain α,β-unsaturated amides (acrylamides) and sulfonamides (vinyl sulfonamides), Michael acceptors which have been extensively used to target reactive nucleophiles in kinases and proteases.[4,28-30] These reactive groups have relatively low nonspecific reactivity to other serum and cellular proteins. Therefore, we designed and genetically encoded

three UAAs, N^{ϵ}-acryloyl-(*S*)-lysine (AcrK, **1**), *p*-acrylamido-(*S*)-phenylalanine (AcrF, **2**), and *p*-vinylsulfonamido-(*S*)-phenylalanine (VSF, **3**), with electrophilic groups that can react with the ε-amino group of lysine residues in a target protein (Figure 1A). Moreover, we show that a mutant of the antigen-binding fragment (Fab) of the therapeutic monoclonal antibody Herceptin bearing **3** selectively and irreversibly labels ErbB2-positive breast cancer cells under physiological conditions.

Results and Discussion

To test whether electrophilic UAAs can be used to selectively form irreversible protein complexes, we used an αErbB2 Fab based on Herceptin and the extracellular domain (ECD) of the oncogenic receptor tyrosine kinase ErbB2 as a model system. Herceptin binds ErbB2 with high affinity (K_d ~0.1 nM), leading to reduced cellular proliferation through disruption of receptor dimerization and antibody-dependent cellular cytotoxicity (ADCC).[31-33] We hypothesized that incorporation of a relatively small electrophilic UAA at a defined site in Herceptin would allow covalent bond formation with a proximal lysine ε-amino group of the ErbB2 ECD without significantly perturbing binding affinity. Examination of the Herceptin Fab-ErbB2 ECD co-crystal structure (PDB: 1N8Z) revealed that there are two lysine residues (Lys569 and Lys593) at the binding interface, providing potential crosslinking sites (Figure 1B).[34] Two residues in the light chain (LC) complementarity determining region 3 (CDR3) of Herceptin (LC-Tyr92 and LC-Thr93) are within 10 Å of Lys569, and one residue in the heavy chain (HC) CDR3 of αErbB2 (HC-Gly103) is positioned to allow potential interaction with Lys593 of ErbB2. These residues in the Herceptin Fab gene were mutated to the amber stop codon, TAG, generating plasmids encoding Herceptin Fab-Tyr92TAG, Herceptin Fab-Thr93TAG, and Herceptin Fab-Gly103TAG (Supplementary Information).

The first UAA we explored for the covalent crosslinking of Herceptin Fab to ErbB2 was the alkyl acrylamide, N^{ϵ}-acryloyl-(*S*)-lysine (AcrK, **1**). Synthesis of **1** was readily achieved in one step in 54% yield (Supplementary Information). Engineered pyrrolysyl-tRNA synthetase (PylRS)/$tRNA^{Pyl}_{CUA}$ pairs from *Methanosarcina* species (*barkeri* and *mazei*) can be used to genetically encode a variety of lysine analogues, including N^{ϵ}-acetyl-(*S*)-lysine,[35] N^{ϵ}-methyl-(*S*)-lysine,[36,37] and N^{ϵ}-crotonyl-(*S*)-lysine (Kcr).[38] To evolve a PylRS that can efficiently charge $tRNA^{Pyl}_{CUA}$ with **1**, we generated a library based on the sequence of the aaRS specific for Kcr, which differs from **1** only by the presence of a single methyl group. Structural analysis of the PylRS active site indicated that mutations at residues Leu266, Leu270, and Tyr271 might accommodate the slightly less bulky acrylamide **1**. Additionally, Leu274Ala and Cys313Phe mutations are essential in reshaping the PylRS active site to accommodate relatively small pyrrolysine analogues, and mutation of Tyr349 to phenylalanine can improve the PylRS amber suppression efficiency. Therefore, three residues (Leu266, Leu270, Tyr271) were varied by NNK randomization (N = A, G, T or C; K = G or T), and three fixed mutations (Leu274Ala, Cys313Phe, Tyr349Phe) were introduced to afford a library of 3×10^5 mutants (theoretical nucleic acid diversity of 3.3×10^4). Directed evolution was performed in *E. coli* as previously described,[39] resulting in identification of an

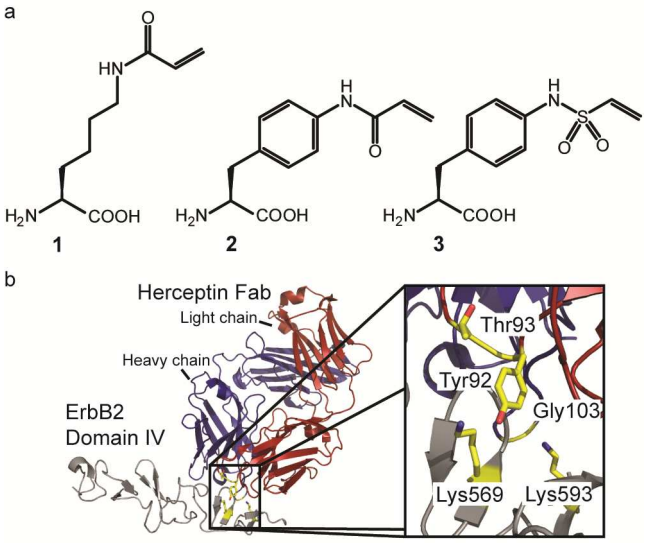

Figure 1. Structures of unnatural amino acids and crosslinking sites at the Herceptin Fab-ErbB2 binding interface. (a) Structures of the electrophilic unnatural amino acids N^{ϵ}-acryloyl-(*S*)-lysine (**1**), *p*-acrylamido-(*S*)-phenylalanine (**2**), and *p*-vinylsulfonamido-(*S*)-phenylalanine (**3**). (b) The crystal structure of a Herceptin Fab bound to the extracellular domain of ErbB2 (PDB: 1N8Z) shows two residues in the light chain (LC) CDR3 (LC-Tyr92 and LC-Thr93) and one residue in the heavy chain (HC) CDR3 (HC-Gly103) of the Fab that are in close proximity to the ε-amino groups of Lys569 and Lys593 in ErbB2 at the binding interface.

aaRS (AcrKRS, mutations: Leu270Ile, Leu274Ala, Cys313Phe, Tyr349Phe) that incorporates **1** into proteins. Using a fluorescence-based assay in which GFP with a permissive Asn149TAG mutation was co-expressed with the AcrKRS/$tRNA^{Pyl}_{CUA}$ pair in the presence and absence of **1**, a 20-fold increase in fluorescence was observed upon addition of 5 mM **1** (Supplementary Information). We note that while finalizing these experiments, there were two concurrent and independent reports of evolved aaRSs with distinct sequences that incorporate **1** in prokaryotes and eukaryotes.[20,21]

To determine the fidelity and efficiency for incorporation of **1** into proteins in *E. coli*, Herceptin Fab mutants with amber codons at residues LC-Tyr92, LC-Thr93, or HC-Gly103 were expressed in the presence of the AcrKRS/$tRNA^{Pyl}_{CUA}$ pair (Figure 2A). In Terrific Broth (TB) media supplemented with 5 mM **1**, the yield of periplasmically secreted mutant proteins following purification over Protein G resin was ~1-2 mg L^{-1}. Electrospray ionization mass spectrometry (ESI-MS) confirmed independent incorporation of **1** at each of the three suppressed sites (Figure 2A-C). Notably, the modified Fab's did not undergo any observable 1,4-addition reactions with intracellular thiol nucleophiles, such as glutathione (GSH) prior to secretion. An ErbB2 ELISA revealed that, while Herceptin Fab-AcrK92 and Herceptin Fab-AcrK93 bound ErbB2 with an apparent K_d equivalent to the wild type (WT) Fab (~0.4 nM), Herceptin Fab-AcrK103 binding was reduced almost 20-fold (Supplementary Information). Each of the mutant proteins (6 μM) was then incubated with 750 nM of the human ErbB2 ECD (Met1-Thr652 with a C-terminal hexaHis tag) in Dulbecco's Phosphate-Buffered Saline (DPBS), pH 8.8, for 48 h at 37 °C, followed by reducing SDS-PAGE to identify cova-

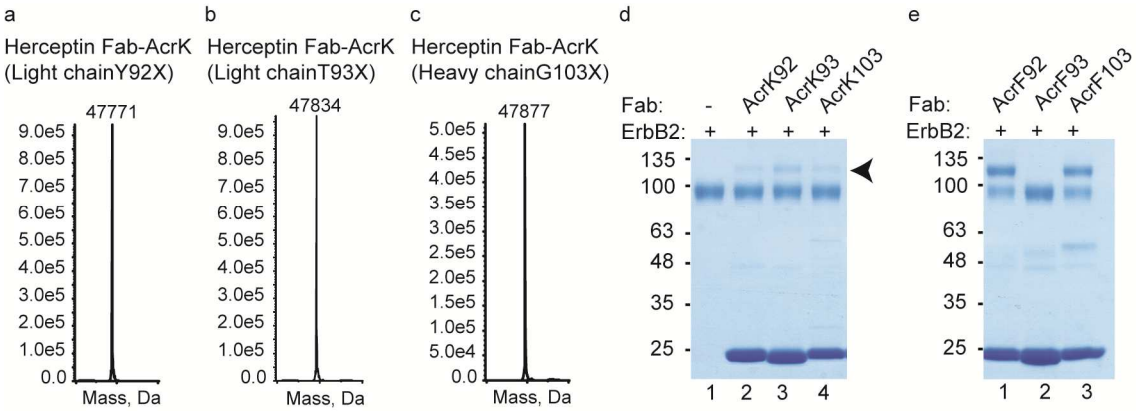

Figure 2. Crosslinking reaction of genetically encoded acrylamide-bearing amino acids. (a-c) Mass spectrometry of **1** substituted in the Herceptin Fab at LC-Tyr92 (a), LC-Thr93 (b), or HC-Gly103 (c) with expected masses of 47771, 47833, and 47877 Da, respectively. (d) Reducing SDS-PAGE of 6 μM Herceptin Fab-AcrK mutants reacted with 750 nM ErbB2 ECD (residues Met1-Thr652 with a C-terminal hexaHis tag) in DPBS, pH 8.8, for 48 h at 37 °C. The covalently crosslinked Herceptin Fab-ErbB2 ECD complex is indicated by an arrowhead. (e) Reducing SDS-PAGE of Herceptin Fab-AcrF mutants with ErbB2 ECD under the same conditions as in (d).

lently crosslinked products. A new band was observed at ~125 kDa, which corresponds to a 1:1 covalent bond between the ErbB2 ECD (apparent glycosylated mass ~100-110 kDa) and reduced Herceptin Fab (~25 kDa). Densitometry revealed that ~10% of the ErbB2 ECD reacted with the Herceptin Fab-AcrK93 mutant, while Herceptin Fab-AcrK92 and Herceptin Fab-AcrK103 reacted only minimally (Figure 2D). Reducing SDS-PAGE completely abrogated the noncovalent association between the WT Fab and ErbB2 ECD (Supplementary Information), indicating that **1** is responsible for the observed crosslinking. These initial results suggest that a mutant Fab bearing a genetically encoded Michael acceptor can form a covalent complex with its native, unmodified binding partner. However, the use of **1** required long incubation times (48 h) and non-physiological pH (> 8.0), and afforded low crosslinking yields (~10%).

Next we explored whether an alternative aza-Michael acceptor, *p*-acrylamido-(*S*)-phenylalanine (AcrF, **2**), with less side chain conformational flexibility and somewhat higher reactivity due to the phenyl substituent,[25] might lead to improved crosslinking efficiency as compared to the lysine scaffold. Synthesis of **2** was readily achieved in two steps with 90% overall yield (Supplementary Information). Many phenylalanine-derived UAAs have been incorporated in response to the amber nonsense codon using evolved *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS)/tRNA$^{Tyr}_{CUA}$ pairs.[40] Additionally, certain evolved TyrRSs exhibit substrate polyspecificity, wherein a single aaRS exhibits high permissivity for UAAs, while preserving its ability to discriminate against the 20 canonical amino acids.[41] Using a fluorescence assay with a GFP-Tyr151TAG mutant, an existing polyspecific TyrRS42 (AcrFRS, mutations: Tyr32Val, Leu65Tyr, Phe108His, Gln109Gly, Asp158Gly, Leu162Glu, Asp286Arg) afforded a 2-fold fluorescence increase upon addition of 1 mM **2**. To confirm successful incorporation of **2**, GFP-AcrF151 was expressed and purified from BL21(DE3) cells, followed by characterization with SDS-PAGE and ESI-MS. The observed mass (+302 Da) potentially corresponds to a GSH adduct of GFP-AcrF151, which is likely due to its increased electrophilicity compared to **1** (Supplementary Information).

Assuming that periplasmic secretion would limit posttranslational modification of **2** by GSH, Herceptin Fab mutants with amber mutations replacing residues LC-Tyr92, LC-Thr93, or HC-Gly103 were periplasmically expressed in the presence of the AcrFRS/tRNA$^{Tyr}_{CUA}$ pair. Supplementation of TB media with 1 mM **2** typically yielded 4-6 mg L^{-1} of mutant protein following purification over Protein G resin. ESI-MS confirmed successful incorporation of **2** without any observed modification by cellular nucleophiles, and an ELISA confirmed binding of the mutants to ErbB2 with apparent K_d's similar to the WT Fab (Supplementary Information). To assay the formation of covalently crosslinked product, the Fab mutants (6 μM) were incubated with 750 nM ErbB2 ECD in DPBS, pH 8.8, for 48 h at 37 °C, followed by reducing SDS-PAGE. In contrast to the Herceptin Fab-AcrK mutants, in which only mutation at LC-Thr93 resulted in crosslink formation, crosslinking using **2** as the aza-Michael acceptor was observed for the LC-Tyr92 mutant (66% reaction) and the HC-Gly103 mutant (65% reaction) (Figure 2E). While the long alkyl side chain of **1** may reach Lys569 from the LC-Thr93 site, **2** is structurally similar to the tyrosine it replaces at LC-Tyr92 and is, therefore, better oriented for reaction with Lys569. This optimal positioning, along with the increased reactivity of the aryl acrylamide, may explain the increased reaction efficiency observed using **2** for crosslinking to the ErbB2 ECD.

Although acrylamides have been used successfully to covalently modify adjacent cysteine residues, they are relatively weak electrophiles[1] and have not been particularly successful for aza-Michael addition reactions at protein interfaces.[13] Therefore, to further improve the yield of crosslinked product, as well as the reaction conditions (time and pH), we designed a novel optimized aza-Michael acceptor UAA with a vinylsulfonamide group, *p*-vinylsulfonamido-(*S*)-phenylalanine (VSF, **3**). Vinylsulfonamides have recently been used to design inhibitors of ubiquitin activating enzymes[42] and kinases,[43] and are somewhat more reactive than the corresponding acrylamide moieties (reactivity: *N*-phenylacrylamide < *N*-phenylvinylsulfonamide).[25] Synthesis of **3** was achieved in 32% overall yield in two steps (Supplementary Information). Because the polyspecific AcrFRS afforded only modest pro-

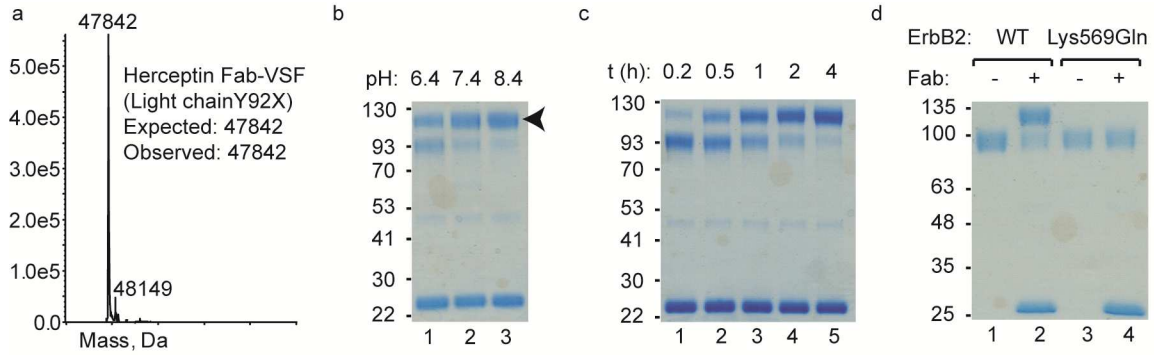

Figure 3. Characterization of the reaction of Herceptin Fab-VSF92 with the ErbB2 ECD. (a) Mass spectrometry of **3** substituted in the Herceptin Fab at LC-Tyr92 (Herceptin Fab-VSF92). The peak observed at 48149 Da likely represents a (+307 Da) GSH adduct. (b) Reducing SDS-PAGE of reactions of 4 μM Herceptin Fab-VSF92 with 1 μM ErbB2 ECD (residues Met1-Thr652 with a C-terminal hexaHis tag, Sino Biological Inc., SBI) carried out for 2 h at 37 °C in DPBS at pH 6.4, 7.4, or 8.4. The covalently crosslinked Herceptin Fab-ErbB2 ECD complex is indicated by an arrowhead. (c) Reducing SDS-PAGE of time dependent reaction of 4 μM Herceptin Fab-VSF92 with 1 μM ErbB2 ECD (residues Met1–Thr652 with a C-terminal hexaHis tag, Sino Biological Inc., SBI) reacted in DPBS, pH 7.4, at 37 °C for 0.2, 0.5, 1, 2, or 4 h. (d) Reducing SDS-PAGE of reaction of 5 μM Herceptin fab-VSF92 with 1 μM ErbB2 ECD or ErbB2-Lys569Gln ECD (residues Thr23–Thr630 with a C-terminal hexaHis tag expressed from HEK293 cells) in DPBS, pH 7.4, for 2 h at 37 °C.

tein yields of the GFP-VSF151 mutant in the presence of 1 mM **3**, we evolved an improved TyrRS that can more efficiently charge $tRNA^{Tyr}_{CUA}$ with **3**. A library was generated using the AcrFRS sequence as a template, in which Val32, Ile159, and Glu162 were varied by NNK mutagenesis to allow for accommodation of the bulkier vinyl sulfonamide substituent, and Ala67 and His70 were varied using GST codons (S = C or G, encoding Ala or Gly) to restrict steric bulk. A library of 6×10^6 mutants was obtained (theoretical nucleic acid diversity of 1.3×10^5), and directed evolution was performed in *E. coli* as previously described.[39] An aaRS (VSFRS, mutations: Tyr32Gly, Leu65Tyr, Phe108His, Gln109Gly, Asp158Gly, Ile159Leu, Leu162Gln, Asp286Arg) was identified that contained three mutations to glycine residues, which enlarges the active site and allows efficient incorporation of **3** into proteins. Suppression of GFP-Asn149TAG in the presence of 1 mM **3** resulted in a 13-fold increase in fluorescence (Supplementary Information). To confirm successful incorporation of **3** a GFP mutant, GFP-VSF151, was expressed and purified from BL21(DE3) cells, followed by characterization with ESI-MS. The mutant protein was again observed as a (+303 Da) GSH conjugate (Supplementary Information).

With a new aaRS in hand, Herceptin Fab-Tyr92TAG was periplasmically expressed in the presence of the VSFRS/$tRNA^{Tyr}_{CUA}$ pair. UAA **3** was incorporated at the same site as **2**, since they both similarly present the electrophilic group on the phenylalanine scaffold. Supplementation of TB expression media with 1 mM **3** yielded ~4 mg L^{-1} of mutant protein following purification, and ESI-MS confirmed successful incorporation (Figure 3A). A small portion of Herceptin Fab-VSF92 is observed as a GSH conjugate (+307 Da) in the mass spectrum, reflecting the enhanced reactivity of the vinylsulfonamide group compared to the aryl acrylamide.

To explore the reactivity of this new Michael acceptor UAA, we incubated 4 μM Herceptin Fab-VSF92 with 1 μM ErbB2 ECD in DPBS adjusted to pH 6.4, 7.4, or 8.4 for 2 hours at 37 °C, followed by reducing SDS-PAGE (Figure 3B). The reaction efficiency correlated with pH, with 53%, 86%, and 95% crosslinking observed at pH 6.4, 7.4, and 8.4, respectively. Furthermore, analyzing the time-dependence of the reaction revealed covalently linked product after as short as 10 min, with ~96% conversion to crosslinked product achieved within 4 h at pH 7.4 (Figure 3C). We additionally performed the reaction in DPBS, pH 7.4 for 2 hours at 37 °C with 4 nM Herceptin Fab-VSF92 and 1 nM ErbB2 ECD labeled with an infrared dye in order to image the reaction product. Reducing SDS-PAGE showed ~80% crosslinking using these concentrations, which are similar to those used either *in vitro* or *in vivo* (Supplementary Information). Thus, **3** is a highly efficient crosslinker compared to **1** and **2**, and forms covalent complexes at physiological pH. On the basis of the Herceptin Fab-ErbB2 ECD co-crystal structure, we postulated that Herceptin Fab-VSF92 is reacting with Lys569 of the ErbB2 ECD. To test this hypothesis, the WT ErbB2 ECD and a Lys569Gln mutant were expressed and purified from HEK293 cells. Each ECD was incubated with Herceptin Fab-VSF92 at pH 7.4 for 2 h at 37 °C, followed by reducing SDS-PAGE (Figure 3D). A covalent product was only observed with WT ErbB2 ECD, while no reaction occurred between Herceptin Fab-VSF92 and the Lys569Gln mutant. ELISA results confirmed that Herceptin Fab-VSF92 bound to the ErbB2-Lys569Gln ECD with an affinity indistinguishable from that of WT ErbB2 ECD (Supplementary Information). Unfortunately, attempts to generate tryptic fragments of the crosslinked proteins that ionize with appreciable efficiency were unsuccessful.

Previous efforts to form covalent bonds between proteins using genetically encoded UAAs required mutagenesis to introduce a unique cysteine nucleophile into the target protein.[19] A novel feature of our covalent crosslinking methodology is that the target protein does not require the introduction of a nucleophilic residue and, as a consequence, may potentially be crosslinked in its native cellular state. To test the biocompatibility of our crosslinking strategy with living cells, we first confirmed that the reaction between Herceptin Fab-VSF92 and the ErbB2 ECD is not affected by the presence of ~50% fetal bovineserum (FBS). No off-target reactions with serum proteins were observed by western blot using an anti-human kappa light chains antibody (Supplementary Information). Next,

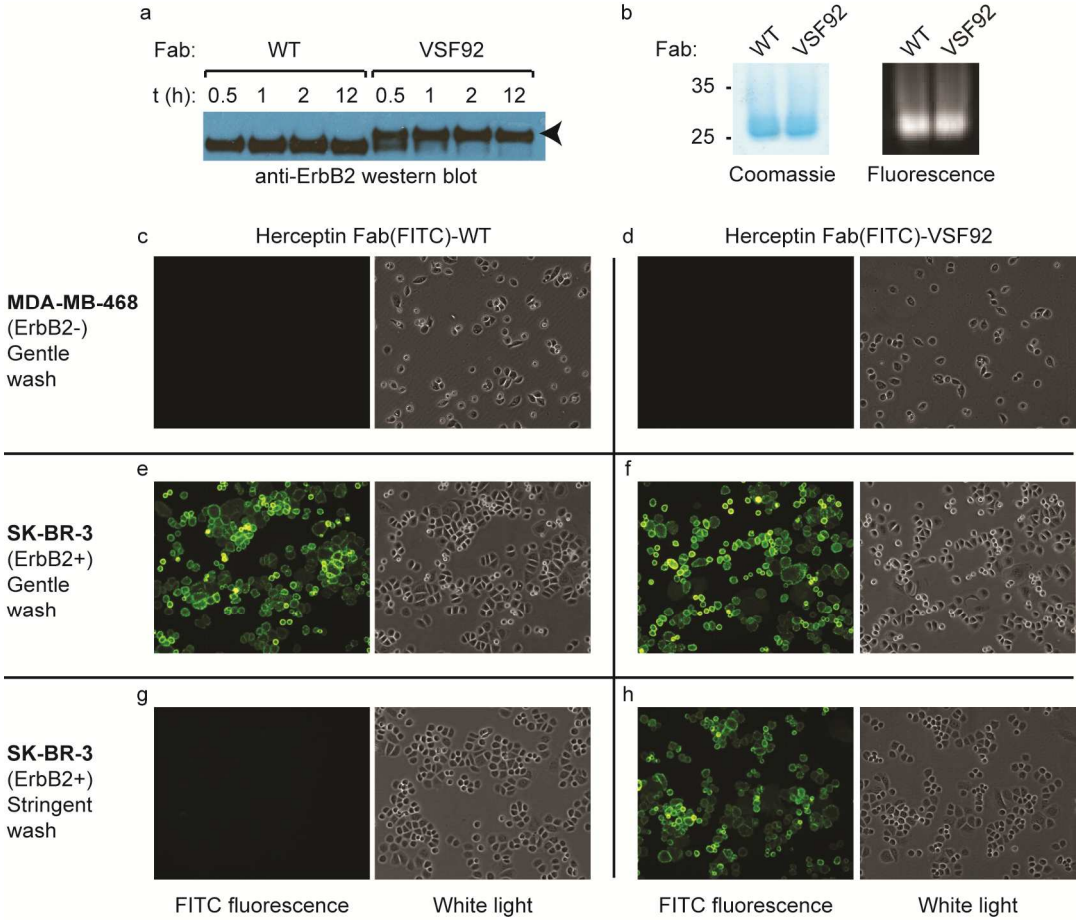

Figure 4. Specific and irreversible labeling of ErbB2 on the cell surface by Herceptin Fab-VSF92. (a) Herceptin Fab-VSF92 and Herceptin Fab-WT (25 nM) were incubated with SK-BR-3 cells for 2 h at 37 °C, followed by total cell lysis and western blot analysis with an anti-ErbB2 antibody. The covalently crosslinked Herceptin Fab-VSF92-ErbB2 complex is indicated by an arrowhead. (b) Coomassie-stained reducing SDS-PAGE showing Herceptin Fab-WT and Herceptin Fab-VSF92 after non-specific labeling with FITC, and the corresponding fluorescent image of the same gel. (c-h) Fluorescence imaging of live cells treated with Herceptin Fab(FITC)-WT or Herceptin Fab(FITC)-VSF92. MDA-MB-468 and SK-BR-3 cells were cultured in DMEM supplemented with 10% FBS at 37 °C. At 70% confluence, 25 nM of Herceptin Fab(FITC)-WT or Herceptin Fab(FITC)-VSF92 was added to the culture medium and incubated for 2 h in the dark. Green fluorescent (left) and bright field (right) images were obtained under different wash conditions (the 'gentle' wash buffer was DPBS, and the 'stringent' wash buffer was 100 mM glycine, pH 2.8, 300 mM NaCl, 3% tween 20) to show the irreversible labeling of cell surface ErbB2 with only Herceptin Fab(FITC)-VSF92 (h) but not with Herceptin Fab(FITC)-WT (g).

ErbB2-positive SK-BR-3 breast cancer cells cultured in Dulbecco's Modified Eagle Medium (DMEM), pH 7.4, supplemented with 10% FBS were incubated with Herceptin Fab-WT or Herceptin Fab-VSF92 (25 nM) for 0.5, 1, 2, or 12 h. Western blot of a SDS-PAGE gel of the lysate with an anti-ErbB2 antibody that binds to the intracellular domain of full length ErbB2 revealed complete covalent labeling of ErbB2 by Herceptin Fab-VSF92 within 2 h, whereas no covalent reaction was observed for Herceptin Fab-WT (Figure 4A). We next created fluorescent probes for live cell imaging by non-specifically modifying the Herceptin Fab-WT and Herceptin Fab-VSF92 with fluorescein isothiocyanate (FITC) to generate Herceptin Fab(FITC)-WT and Herceptin Fab(FITC)-VSF92, respectively (Figure 4B). ErbB2-positive SK-BR-3 cells and ErbB2-negative MDA-MB-468 cells grown in DMEM, pH 7.4, with 10% FBS were incubated with Herceptin Fab(FITC)-WT or Herceptin Fab(FITC)-VSF92 (25 nM) for 2 h. After washing with DPBS to remove nonspecifically bound Fab, the observed fluorescence with SK-BR-3 cells was similar for both Herceptin Fab(FITC)-WT and Herceptin Fab(FITC)-VSF92, while no fluorescence was observed on MDA-MB-468 cells, confirming the selective labeling of ErbB2 (Figure 4C-F). To confirm the covalent nature of Herceptin Fab(FITC)-VSF92 binding to cell-surface ErbB2, cells were subjected to stringent washing conditions (100 mM glycine, pH 2.8, 300 mM NaCl, 3% tween 20). The Herceptin Fab(FITC)-VSF92 probe remained bound to ErbB2 at the cell surface, while no fluorescence was observed on the cells previously labeled with Herceptin Fab(FITC)-WT (Figure 4G and H). These results demonstrate that the genetically encoded aza-Michael acceptor **3** can selectively and efficiently covalently react with a proximal lysine nucleophile at the protein binding interface both *in vitro* and in the presence of living cells.

Conclusion

Herein we have evolved orthogonal aaRS/tRNA pairs for the genetic incorporation of a series of electrophilic UAAs, N^{ε}-acryloyl-(*S*)-lysine (AcrK, **1**), *p*-acrylamido-(*S*)-phenylalanine

(AcrF, **2**), and *p*-vinylsulfonamido-(*S*)-phenylalanine (VSF, **3**). This series provides two different amino acid scaffolds, lysine and phenylalanine, as well as incremental increases in reactivity (reactivity: *N*-alkylacrylamide < *N*-phenylacrylamide < *N*-phenylvinylsulfonamide).[25] We evaluated these genetically encoded UAAs for covalently crosslinking Herceptin Fab to ErbB2. While the acrylamide UAAs resulted in modest crosslinking yields under basic conditions (pH = 8.8), the vinylsulfonamide moiety resulted in rapid covalent crosslinking of the αErbB2-ErbB2 complex at physiological pH with >95% yield, providing an extremely robust protein-protein crosslinking strategy for native target proteins. Because this strategy does not require genetic or chemical modification of the target protein, it can potentially be directly applied to cells or tissues. This methodology provides useful covalent probes to identify protein interaction partners that may be too weakly bound to be isolated by more conventional affinity pull-down techniques. Moreover, covalent antibodies of this type may be particularly useful for extracellular targeting applications, such as for diagnostics, imaging, and fluorescence guided surgery.[44,45] We are currently determining the generality of this approach by introducing electrophilic UAAs into other therapeutic antibodies, protein-based protease inhibitors, growth factors, and cytokines. In addition, we are exploring the use of library based methods with affinity based selections to identify productive crosslinking sites in protein complexes for which detailed structural information is not available.

Experimental Section

Protocols for the chemical syntheses of UAAs **1**, **2**, and **3**, construction of library and expression plasmids, selections and validation for the evolved aaRSs, and ELISA binding assays are given in the Supplementary Information.

Protein expression and purification. Each pBAD-Herceptin Fab-TAG (LC-Tyr92TAG, LC-Thr93TAG, or HC-Gly103TAG) plasmid was co-transformed with pUltra-AcrKRS/$tRNA^{Pyl}_{CUA}$, pUltra-AcrFRS/$tRNA^{Tyr}_{CUA}$, or pUltra-VSFRS/$tRNA^{Tyr}_{CUA}$ into *E. coli* TOP10 cells. Cells were cultured in TB media, supplemented with 100 μg mL^{-1} ampicillin, 50 μg mL^{-1} spectinomycin, and 5 mM **1**, 1 mM **2**, or 1 mM **3**. Protein expression was induced at an OD_{600} ~1.0 by addition of 0.2% arabinose and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cells were cultured at 37 °C (for proteins containing **1**) or 25 °C (for proteins containing **2** or **3**) for 16-24 h. Protein was released from pelleted cells by periplasmic lysis with 20 mL lysis buffer (30 mM Tris-HCl, pH 7.5, 1 mM EDTA, 20% sucrose, 0.2 mg mL^{-1} lysozyme) per gram of wet cell pellet with shaking at 37 °C for 30 min. Clarified lysate was applied to a 1 mL Protein G Sepharose 4 Fast Flow (GE Healthcare) column equilibrated with Buffer A (50 mM NaOAc, pH 5.0), and Fab was eluted with Buffer B (100 mM glycine, pH 2.8). Proteins were exchanged into DPBS using centrifugal filtration with a 30 kDa molecular weight cut-off (MWCO), followed by characterization with ESI-MS and SDS-PAGE. Yields were ~1-2 mg L^{-1} for proteins containing **1** and ~4-6 mg L^{-1} for proteins containing **2** or **3**.

Plasmids encoding the ErbB2 ECD (pFuse-ErbB2-His_6 or pFuse-ErbB2(Lys569Gln)-His_6) were transfected using 293-transfectin (Life Technologies) into FreeStyle™ HEK293 cells (Life Technologies) that were cultured to a density of 1×10^6 cells mL^{-1} in 500 mL of FreeStyle™ 293 expression media (Life Technologies) at 37 °C and 5% CO_2. Expression media containing secreted ErbB2 ECD was harvested and combined every two days over the course of 6 days, followed by purification using Ni-NTA agarose resin (Qiagen), with a yield of ~1-2 mg L^{-1}. SDS-PAGE analysis showed that recombinant ErbB2 ECD proteins migrated as the same molecular weight as an equivalent ErbB2 ECD from a commercial source (Supplementary Information). Characterization by mass spectrometry is not possible given the large molecular weight and extensive glycosylation of these proteins.

Crosslinking reactions. Reactions between Herceptin Fab mutants (6 μM) bearing **1** or **2** at LC-Tyr92, LC-Thr93, or HC-Gly103 and 750 nM ErbB2 ECD were carried out in 20 μL DPBS (adjusted to pH 8.8 with bicine) at 37 °C for 48 h. Reactions between 1 μM Herceptin Fab-VSF92 and 4 μM ErbB2 ECD were carried out in 20 μL DPBS, pH 7.4, at 37 °C for 2 h, unless otherwise specified. Prior to SDS-PAGE, reactions were quenched by boiling at 90 °C in SDS-PAGE loading buffer supplemented with β-mercaptoethanol for 5 min. Gel band density was determined using ImageJ software (http://rsbweb.nih.gov/ij/), and reaction completion was calculated as the intensity of the crosslinked product relative to total ErbB2 (free plus crosslinked).

Cell-based crosslinking of Herceptin Fab mutants was carried out in a 12-well plate at a density of 1×10^5 cells per well in DMEM supplemented with 10% FBS at 37 °C and 5% CO_2. At ~70% confluency, 25 nM Herceptin Fab-VSF92 or Herceptin Fab-WT was added to each well and incubated for 0.5, 1, 2, or 12 h. Cells were washed twice with cold DPBS and lysed using 250 μL of CelLytic M solution (Sigma) at 4 °C for 30 min. The clarified supernatant was analyzed by western blot with an anti-human ErbB2 antibody (29D8, Cell Signaling Technologies) that targets the intracellular domain of full length ErbB2.

Fluorescent labeling of mammalian cells. Fluorescein isothiocyanate (FITC, Molecular Probes) (770 μM) was reacted with Herceptin Fab-VSF92 or Herceptin Fab-WT (58 μM) in 72 μL of 200 mM sodium bicarbonate, pH 9.0, for 2 h at 25 °C with shaking. Labeled protein was isolated from unreacted dye using 0.5 mL Zeba spin desalting columns with a 40 kDa MWCO (Pierce). The in-gel fluorescence image was acquired on a Bio-Rad Gel Doc XR Imaging system with UV transillumination. FITC-labeled Herceptin Fab-VSF92 or Herceptin Fab-WT (25 nM) was added to ~70% confluent SK-BR-3 or MDA-MB-468 cells (ATCC) cultured in DMEM supplemented with 10% FBS at 37 °C and 5% CO_2. After incubation in the dark at 37 °C for 2 h, the cells were washed twice with cold DPBS to remove unbound Fab. The SK-BR-3 cells were then washed twice for 10 min at RT using either mild (DPBS) or harsh (100 mM glycine, pH 2.8, 300 mM NaCl, 3% tween 20) conditions. Cells were imaged on a Nikon Eclipse Ti fluorescence microscope with a 10x/0.30 Nikon Plan Fluor objective. FITC was excited with a Lumencor 64 mW light source using a GFP/FITC 475/28 BP filter, and emission was detected between 515 and 555 nm using an 8 sec exposure.

ASSOCIATED CONTENT

Supporting Information. Figures S1–S9, protein mass spectra, Table S1, and detailed Materials and Methods. Correspondence and requests for materials should be addressed to P.G.S. and C.H.K. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

schultz@scripps.edu
chkim@calibr.org

Author Contributions

†These authors contributed equally.

ACKNOWLEDGMENT

We thank Virginia Chambers for her assistance with manuscript preparation. Research reported in this publication was supported by the National Institutes of Health (NIH) under Award Number F32GM099162 (J.L.F.) and by NIH R01 GM097206 (P.G.S.). This is manuscript number 26045 of The Scripps Research Institute.

REFERENCES

(1) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Discov.* **2011**, *10*, 307.
(2) Erve, J. C. L. *Expert Opin. Drug Met.* **2006**, *2*, 923.
(3) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. *Future Med. Chem.* **2010**, *2*, 949.
(4) Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev.* **2002**, *102*, 4639.
(5) Bachovchin, D. A.; Cravatt, B. F. *Nat. Rev. Drug Discov.* **2012**, *11*, 52.
(6) Liu, Q. S.; Sabnis, Y.; Zhao, Z.; Zhang, T. H.; Buhrlage, S. J.; Jones, L. H.; Gray, N. S. *Chem. Biol.* **2013**, *20*, 146.
(7) Potashman, M. H.; Duggan, M. E. *J. Med. Chem.* **2009**, *52*, 1231.
(8) Kalgutkar, A. S.; Dalvie, D. K. *Expert Opin. Drug Dis.* **2012**, *7*, 561.
(9) O'Connor, O. A.; Stewart, A. K.; Vallone, M.; Molineaux, C. J.; Kunkel, L. A.; Gerecitano, J. F.; Orlowski, R. Z. *Clin. Cancer Res.* **2009**, *15*, 7085.
(10) Honigberg, L. A.; Smith, A. M.; Sirisawad, M.; Verner, E.; Loury, D.; Chang, B.; Li, S.; Pan, Z.; Thamm, D. H.; Miller, R. A.; Buggy, J. J. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 13075.
(11) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974.
(12) Takaoka, Y.; Ojida, A.; Hamachi, I. *Angew. Chem. Int. Ed.* **2013**, *52*, 4088.
(13) Holm, L.; Moody, P.; Howarth, M. *J. Biol. Chem.* **2009**, *284*, 32906.
(14) Marquez, B. V.; Beck, H. E.; Aweda, T. A.; Phinney, B.; Holsclaw, C.; Jewell, W.; Tran, D.; Day, J. J.; Peiris, M. N.; Nwosu, C.; Lebrilla, C.; Meares, C. F. *Bioconj. Chem.* **2012**, *23*, 1080.
(15) Love, K. R.; Pandya, R. K.; Spooner, E.; Ploegh, H. L. *ACS Chem. Biol.* **2009**, *4*, 275.
(16) Wang, L.; Zhang, Z. W.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 56.
(17) Huang, Y.; Wan, W.; Russell, W. K.; Pai, P.-J.; Wang, Z.; Russell, D. H.; Liu, W. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 878.
(18) Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 14819.
(19) Xiang, Z.; Ren, H.; Hu, Y. S.; Coin, I.; Wei, J.; Cang, H.; Wang, L. *Nat. Methods* **2013**, *10*, 885.
(20) Lee, Y.-J.; Wu, B.; Raymond, J. E.; Zeng, Y.; Fang, X.; Wooley, K. L.; Liu, W. R. *ACS Chem. Biol.* **2013**, *8*, 1664.
(21) Li, F.; Zhang, H.; Sun, Y.; Pan, Y.; Zhou, J.; Wang, J. *Angew. Chem. Int. Ed.* **2013**, *52*, 9700.
(22) Pace, N. J.; Weerapana, E. *ACS Chem. Biol.* **2013**, *8*, 283.
(23) Basle, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213.
(24) Lins, L.; Thomas, A.; Brasseur, R. *Protein Sci.* **2003**, *12*, 1406.
(25) Naven, R. T.; Kantesaria, S.; Nadanaciva, S.; Schroeter, T.; Leach, K. L. *Toxicol. Res.* **2013**, *2*, 235.
(26) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678.
(27) Chmura, A. J.; Orton, M. S.; Meares, C. F. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8480.
(28) Reddick, J. J.; Cheng, J. M.; Roush, W. R. *Org. Lett.* **2003**, *5*, 1967.
(29) Santos, M. M. M.; Moreira, R. *Mini-Rev. Med. Chem.* **2007**, *7*, 1040.
(30) Noe, M. C.; Gilbert, A. M. In *Annual Reports in Medicinal Chemistry, Vol 47*; Desai, M. C., Ed. 2012; Vol. 47, p 413.
(31) Carter, P.; Presta, L.; Gorman, C. M.; Ridgway, J. B. B.; Henner, D.; Wong, W. L. T.; Rowland, A. M.; Kotts, C.; Carver, M. E.; Shepard, H. M. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4285.
(32) Adams, G. P.; Weiner, L. M. *Nat. Biotechnol.* **2005**, *23*, 1147.
(33) Bostrom, J.; Yu, S.-F.; Kan, D.; Appleton, B. A.; Lee, C. V.; Billeci, K.; Man, W.; Peale, F.; Ross, S.; Wiesmann, C.; Fuh, G. *Science* **2009**, *323*, 1610.
(34) Cho, H. S.; Mason, K.; Ramyar, K. X.; Stanley, A. M.; Gabelli, S. B.; Denney, D. W.; Leahy, D. J. *Nature* **2003**, *421*, 756.
(35) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Chem. Biol.* **2008**, *4*, 232.
(36) Nguyen, D. P.; Alai, M. M. G.; Kapadnis, P. B.; Neumann, H.; Chin, J. W. *J. Am. Chem. Soc.* **2009**, *131*, 14194.
(37) Groff, D.; Chen, P. R.; Peters, F. B.; Schultz, P. G. *Chembiochem* **2010**, *11*, 1066.
(38) Kim, C. H.; Kang, M.; Kim, H. J.; Chatterjee, A.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 7246.
(39) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.
(40) Liu, C. C.; Schultz, P. G. In *Annual Review of Biochemistry, Vol 79*; Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., Thorner, J. W., Eds. 2010; Vol. 79, p 413.
(41) Young, D. D.; Young, T. S.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. G. *Biochemistry* **2011**, *50*, 1894.
(42) Lu, X.; Olsen, S. K.; Capili, A. D.; Cisar, J. S.; Lima, C. D.; Tan, D. S. *J. Am. Chem. Soc.* **2010**, *132*, 1748.
(43) Garske, A. L.; Peters, U.; Cortesi, A. T.; Perez, J. L.; Shokat, K. M. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15046.
(44) van Dam, G. M.; Themelis, G.; Crane, L. M. A.; Harlaar, N. J.; Pleijhuis, R. G.; Kelder, W.; Sarantopoulos, A.; de Jong, J. S.; Arts, H. J. G.; van der Zee, A. G. J.; Bart, J.; Low, P. S.; Ntziachristos, V. *Nat. Med.* **2011**, *17*, 1315.
(45) Vahrmeijer, A. L.; Hutteman, M.; van der Vorst, J. R.; van de Velde, C. J. H.; Frangioni, J. V. *Nat. Rev. Clin. Oncol.* **2013**, *10*, 507.

Table of Contents Graphic