JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS
Terms of Use

De Novo-Designed Enzymes as Small-Molecule-Regulated Fluorescence Imaging Tags and Fluorescent Reporters

Yu Liu,[†,∥] Xin Zhang,[†,∥] Yun Lei Tan,[†] Gira Bhabha,[‡] Damian C. Ekiert,[§] Yakov Kipnis,[⊥] Sinisa Bjelic,[⊥] David Baker,[⊥] and Jeffery W. Kelly*[,†]

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

[‡]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, United States

[§]Department of Microbiology and Immunology, University of California, San Francisco, California 94143, United States

[⊥]Department of Biochemistry, University of Washington, Seattle, Washington 98195, United States

Supporting Information

ABSTRACT: Enzyme-based tags attached to a protein-of-interest (POI) that react with a small molecule, rendering the conjugate fluorescent, are very useful for studying the POI in living cells. These tags are typically based on endogenous enzymes, so protein engineering is required to ensure that the small-molecule probe does not react with the endogenous enzyme in the cell of interest. Here we demonstrate that *de novo*-designed enzymes can be used as tags to attach to POIs. The inherent bioorthogonality of the *de novo*-designed enzyme−small-molecule probe reaction circumvents the need for protein engineering, since these enzyme activities are not present in living organisms. Herein, we transform a family of *de novo*-designed retroaldolases into variable-molecular-weight tags exhibiting fluorescence imaging, reporter, and electrophoresis applications that are regulated by tailored, reactive small-molecule fluorophores.

Genetically encoded protein tags have proven to be powerful tools for life science research.[1] Fusing a fluorescent protein (FP) tag to a protein-of-interest (POI) illuminates its localization, dynamics, and macromolecular interactions in live cells.[2] The need to expand the spectral diversity of the protein tags for fluorescence imaging and the desire to have temporal control over the appearance of fluorescence have led researchers to develop small-molecule-regulated protein tags that react with the small molecule to afford a fluorescent conjugate. Examples of enzyme-based tags include the SNAP-tag[3,4] engineered from human O^6-alkylguanine-DNA alkyltransferase, the DHFR-tag[5] evolved from dihydrofolate reductase, and the BL-tag[6] based on β-lactamase. The excellent selectivity and the fast labeling kinetics of the protein tags by their corresponding small-molecule probes is advantageous for using tag-POI fusions to elucidate the biological function(s) of the POI.[7]

The challenge in developing a small-molecule-regulated protein tag from an endogenous enzyme is to achieve selective probe labeling of the tag over the endogenous enzyme from which the tag is derived. Extensive protein engineering is often required to achieve this, along with small-molecule tailoring.

Instead of basing an imaging tag on an endogenous enzyme, we introduce the concept that a small-molecule-regulated imaging tag can be based on a *de novo* enzyme that has an activity not present in the organism of interest. Herein we show that the *de novo*-designed retroaldolases (RAs) can be directly used as chemically regulatable protein tags by creating the appropriate small molecules to render them fluorescent (Figure 1a). The RAs are a family of *de novo*-designed enzymes exhibiting distinct three-dimensional structures and molecular weights (MWs) (Supporting Information, Figure S1); however, this family catalyzes the same retroaldol reaction utilizing a common mechanism (Figure 1b).[8−10] Previously, our laboratories have demonstrated the utility of using destabilized RA variants as folding probes, for the purpose of understanding how alterations in the cellular proteostasis network influence RA folding.[11] Herein, we develop new small-molecule probes so that the RA family can be used as small-molecule-regulated tags in the context of POI-RA fusions and demonstrate three integrated applications enabled by the RA-tags. Specifically, the lack of endogenous enzyme activity and an endogenous substrate renders the *de novo*-designed RA family inherently bioorthogonal for live cell imaging, enabled by the selective and cell-permeable fluorescent probe **P4** developed for this application and publication (Figure 1c). Second, the range of available RA MWs (16−30 kDa, Figure S1) not only introduces size flexibility and structural diversity, but also affords the opportunity to detect POIs of similar size (e.g., protein isoforms) by tagging the POIs with RAs of different MWs (Figure 1a) using only one small-molecule probe, probe **P4**. Lastly, the retroaldol reaction catalyzed by RA can create a fluorescent product (**CP1**), useful for quantitative reporting applications. To the best of our knowledge, the RA-tags are the first system that allows these three functions from one enzyme.

To design a selective chemical probe for RA, we functionalized the naphthalene ring that is complementary to the RA substrate binding site with an electron-withdrawing epoxy ketone (**P1**) that is reactive toward nucleophilic residues in the active site of proteases[12] and the proteasome.[13] However, the labeling efficiency of **P1** was slow, thus only a fraction of RA1

Received: June 9, 2014
Published: September 11, 2014

   dx.doi.org/10.1021/ja5056356 | J. Am. Chem. Soc. 2014, 136, 13102−13105

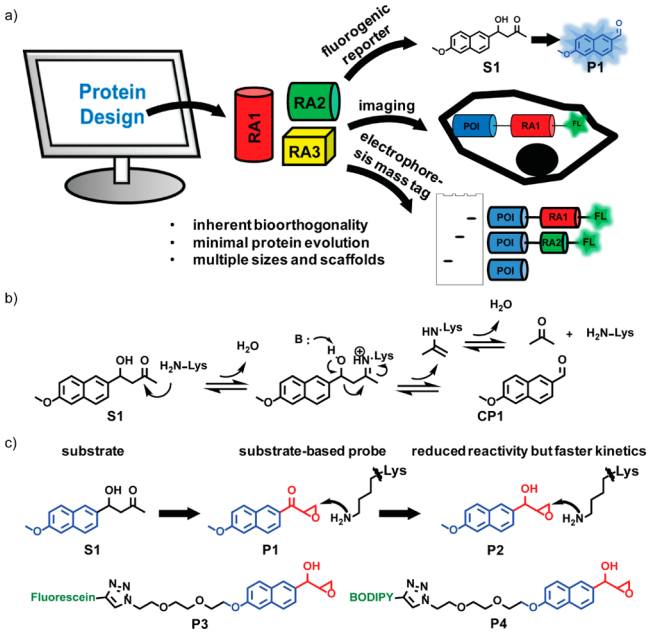
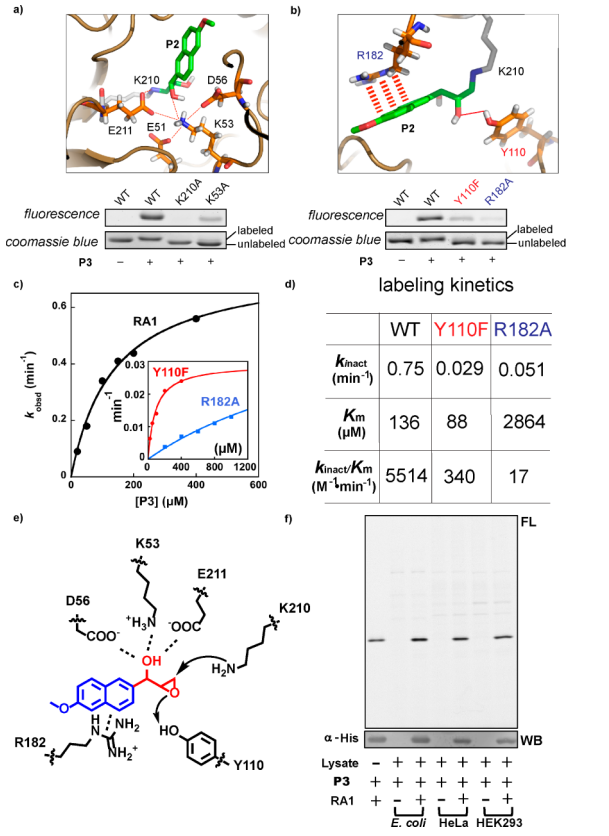


Figure 1. (a) RA-tags transformed from *de novo*-designed retroaldolases have multiple applications. RA1, RA2, and RA3 refer to RA114.3, RA110.4 (cys-free mutant), and RA112 in ref 9, respectively. (b) RAs catalyze a retroaldol reaction using a pK_a-perturbed lysine via Schiff base formation. (c) Design of selective small-molecule probes for RAs based on the substrate structure and catalysis mechanism. **P1**–**P4** selectively label the RAs by targeting the pK_a-perturbed lysine. Color code: blue, binding moieties; red, reactive functional groups; green, fluorophores.



	WT	Y110F	R182A
k_{inact} (min^{-1})	0.75	0.029	0.051
K_m (μM)	136	88	2864
k_{inact}/K_m ($M^{-1}min^{-1}$)	5514	340	17

Figure 2. Origins of the probe-labeling selectivity. (a) The catalytic network of RA1 involves multiple amino acid side chains, mediating chemoselective probe labeling of the catalytic K210 side chain, demonstrated with **P2**. (b) R182 and Y110 interact with **P2** via cation–π and H-bonding interactions, respectively. R182A and Y110F mutations diminish the labeling efficiency. (c) Kinetics of **P3** labeling of wild-type, Y110F and R182A RA1. (d) Summary of parameters. For definition and calculation of kinetic parameters, see Supplementary Method 5. For comparison with existing protein tags, see Table S2. (e) Proposed mechanism of **P2** labeling of RA1. (f) **P3** (200 μM) selectively labels 1 μM RA1 in *E. coli*, HeLa and HEK293 cell lysates (3 mg/mL total protein concentration) at 37 °C for 1 h. FL = fluorescence detection; WB = Western blot.

(5 μM) was labeled by **P1** (25 μM) within 30 min (Figure S2). Surprisingly, **P2**, the intermediate used to synthesize **P1**, harboring an epoxyalcohol with inherently reduced reactivity, exhibited much faster RA1 labeling kinetics than **P1**, shown by the complete labeling of RA1 (5 μM) by **P2** (25 μM) within 30 min (Figure S2). Generally, chemically enhanced reactivity of a warhead leads to faster kinetics.[14] **P2**, harboring an epoxy alcohol, provides an unusual example wherein fast kinetics and low reactivity coexist—the reactivity seems likely to require RA activation of **P2** via binding.

To explore the origins of the fast labeling kinetics of **P2** and the role of the hydroxyl group, we performed protein mutagenesis based on the X-ray co-crystal of the RA1-**P2** conjugate (PDB: 4OU1); probes **P2** and **P3** were used previously for quantifying the folded and functional fraction of destabilized mutant RA as a function of cellular proteostasis capacity.[11] The **P2**-derived substructure interacts with active-site residues E211, E51, D56, and K53 through an electrostatic network (Figure 2a). Mutating D56 or E51 individually had a minimal effect on labeling efficiency (Figure S3), likely due to redundancy. The apparently most indispensable interaction was the hydrogen bond formed between the β-hydroxyl group of **P2** and the amino or ammonium group of K53, as the K53A mutant severely reduced the labeling efficiency (Figure 2a). These results suggest that the hydroxyl group of **P2** serves to bind and orient the epoxide for optimal reactivity with RA1; replacing it with a ketone, as in **P1**, was suboptimal. The naphthalene group binds and orients the epoxide by interacting with the positively charged arginine residue (R182) via a cation–π interaction that appears to contribute the majority of the binding energy (Figure 2b). Accordingly, the K_m of R182A (~2800 μM; Figure 2c,d) was significantly greater than that of wild-type RA1 (~136 μM), indicating weaker reversible binding. Interestingly, the R182A mutant also reduced the maximum conjugation rate by ~15-fold (~0.051 min^{-1}; Figure 2d). In addition, we found that the hydroxyl group on the β-carbon of the adduct (originally the epoxide oxygen in **P2**) forms a hydrogen bond with the acidic phenolic proton of an adjacent tyrosine residue (Y110 in Figure 2b), which likely activates the epoxide toward nucleophilic attack. The Y110F mutant showed a ~25-fold reduction in the rate constant of conjugation (Figure 2c,d), but the K_m (~88 μM; Figure 2d) is comparable to wild type RA1, suggesting that Y110 contributes only to the epoxide activation and not to reversible binding. Taken together, the interactions featured in Figure 2e appear to contribute to probe binding and reaction.

The selectivity of site-specific covalent labeling of RA1 using the optimized chemical probes was scrutinized. We first assessed the selectivity of RA1 labeling by **P3** in multiple cell types, including *E. coli* and human cells (HeLa and HEK293). A high concentration of **P3** (200 μM) was incubated individually with *E. coli*, HeLa, and HEK293 cell lysates without or with 1

dx.doi.org/10.1021/ja5056356 | J. Am. Chem. Soc. 2014, 136, 13102–13105

μM RA1 spiked in at 37 °C. **P3** successfully labeled RA1 in all three lysates (Figure 2f). Minimal off-target endogenous protein probe labeling was observed, even after a relatively long incubation period of 1 h or longer (Figures 2f and S4–S6). The high labeling selectivity originates from the inherent bioorthogonality of the chemistry that RA catalyzes and the relatively unreactive epoxide that is selectively activated when bound by RA1.

Encouraged by the labeling selectivity of **P3**, we substituted the fluorescein in **P3** with the cell-permeable BODIPY fluorophore,[15,16] affording **P4**—cell permeability is critical for live cell imaging. We tested whether **P4** could selectively label RA1 for sub-cellular imaging. HEK293 cells overexpressing RA1 in the cytosol were labeled with **P4** (10 μM) for 10 min, followed by washing, nuclear staining, and fixation. By confocal imaging, we observed conjugate fluorescence only in the cytosol of the transfected cells (Figure 3a, column 2), but not in

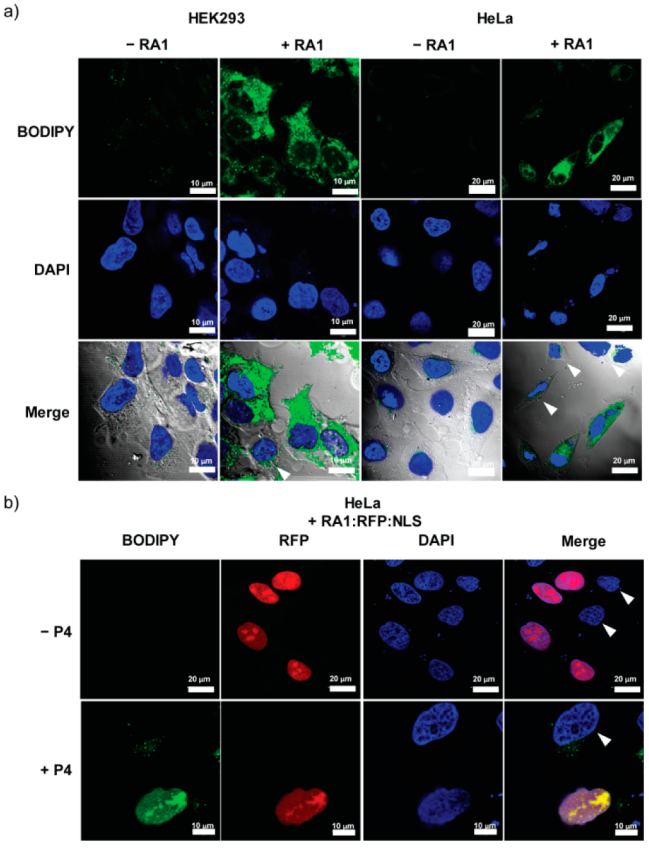


Figure 3. Cell imaging of cytosolic RA1 and nuclear RA1-RFP-NLS with **P4** (10 μM). (a) Selective cytosolic staining of RA1 by **P4** in both HEK293 and HeLa cells (see text for details). The labeling and imaging protocols are described in detail in the Supporting Information. (b) Selective nuclear staining of RA1:RFP:NLS by **P4**. Non-transfected cells are indicated by the white arrowheads.

the non-transfected control (Figure 3a, column 1). Strictly analogous results were obtained using HeLa cells (Figure 3a, columns 3 and 4). To further demonstrate that the fluorescence originates from the selective labeling of RA1, we analyzed cell lysates by SDS-PAGE visualized by BODIPY fluorescence. Fluorescent RA1 bands around 30 kDa indicate that the fluorescence in the confocal image is primarily derived from the selective labeling of RA1 (Figure S5a). Labeling in cells is rapid with almost 50% labeling achieved within 10 min (Figure S5b). The RA1-**P4** conjugate is stable in live cells, as no significant protein degradation of the fluorescent conjugate was observed up to 24 h (Figure S6). In addition, **P4** remained intact when incubated at pH 4.8, with a primary amine, with a thiol and only slowly reacted with the endogenous cellular proteome in the absence of RA1 (Figure S7). We further demonstrated the feasibility of live cell imaging without fixation using the RA1-**P4** combination in live HEK293T cells (Figure S8) and *E. coli* K12 cells (Figure S9). However, additional incubation and washing steps are required to remove unbound fluorophore in the absence of fixation. Collectively, these results support the general utility of the RA1-**P4** combination for live cell imaging in bacterial and mammalian cells.

To further evaluate the feasibility of POI sub-compartmental imaging in live cells, we genetically fused the RA1-tag to a red fluorescent protein harboring three repeats of a nuclear localization signal (RFP:NLS, 30 kDa) developed by the Corrish laboratory.[17] The nuclear localization of the RA1:RFP:NLS protein in HeLa cells was visualized using the red channel, which detects RFP (Figure 3b, column 2). We then probed the localization of the fusion protein using **P4**, excited at 488 nm, the green channel (Figure 3b, column 1). We observed colocalization of the green BODIPY fluorescence from RA1-**P4** conjugate with the red RFP:NLS and the blue DAPI nuclear staining within the same nucleus (Figure 3b, row 2, and Figure S10). Notably, we observed neither significant nuclear BODIPY fluorescence in the neighboring non-transfected cells (Figure 3b, indicated by the white arrowheads) nor non-specific cytosolic fluorescence in the transfected cells, demonstrating the selectivity of **P4** toward RA1.

Enzyme-based tags currently employed as fusions to a POI are derived from an endogenous enzyme and have a unique structure and MW. The structural diversity of the RA family provides tags with distinct structures and MWs (Figure S1), provided that the small-molecule probes developed for RA1 label the entire family of RAs. RA1 (29.6 kDa, TIM barrel fold, RA114.3), RA2 (15.8 kDa, KSI-NTF2-like fold, RA110.4 (cys-free mutant)), and RA3 (21.1 kDa, Rossmann fold, RA112) were all labeled by **P3**, RA3 having the lowest efficiency (Figure S11). This establishes RA1 and RA2 as structurally distinct tags that could be used to differentiate POIs on the same gel, using **P3** fluorescence detection (Figure 1a). To demonstrate their utility, we fused Histone H2B (14 kDa) with his-tagged (His) RA1 and RA2, generating the H2B:RA1:His (45 kDa) and H2B:RA2:His (31 kDa) fusions. After transient transfection or co-transfection of these constructs into HEK293 cells, the proteins were overexpressed for 48 h and labeled using the cell-permeable probe **P4** (10 min). Two fluorescent bands around 45 and 31 kDa were identified as H2B:RA1:His and H2B:RA2:His (Figure 4a, left panel), confirmed by immunoblotting (Figure 4a, right panel). These results validated the capability of RAs to resolve proteins of similar size using a single small-molecule probe.

The inherent bioorthogonality of the reaction catalyzed by the RA enzymes allows us to quantitatively assess the amount of a POI:RA fusion in a complex cellular environment based on the conversion of a non-fluorescent substrate (**S1**) to a fluorescent product (**CP1**) (Figure 4b). This reporter approach is particularly useful when low concentrations of the POI:RA need to be quantified, as the multiple turnover retro-aldol reaction can be used to amplify the signal. Using the dark fluorogenic substrate **S1** (500 μM), the initial linear portion of the time course can be used to quantify RA1 in HEK293 cell

dx.doi.org/10.1021/ja5056356 | J. Am. Chem. Soc. 2014, 136, 13102–13105

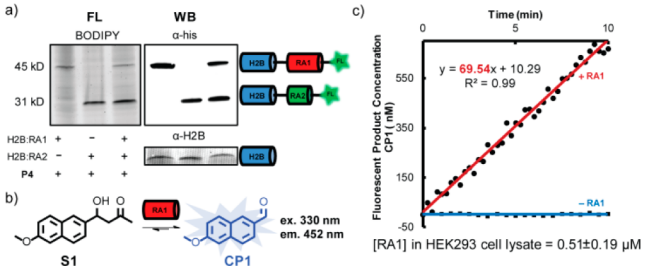


Figure 4. Electrophoresis and reporter applications of the RA-tags. (a) In-gel resolution of the same POI into two separated bands is enabled by the MW diversity of RAs. H2B was tagged with RA1 (30 kDa) or RA2 (16 kDa) and transiently overexpressed in HEK293 cells for 48 h. The cells were labeled using **P4** (10 μM) for 10 min. Two resolved fluorescent bands were observed, their identity confirmed by immunoblotting (see details in the Supporting Information). (b) The fluorogenic reporter reaction catalyzed by RA serves as an additional approach to determine RA-POI concentration. (c) HEK293 cells overexpressing RA1 were lysed and the concentration of RA1 was quantified by the fluorogenic functional reporter reaction. FL = fluorescence detection; WB = Western blot.

lysate (Figure 4c, red trace). No fluorescence product was detected in the negative control (Figure 4c, blue trace). Further, the amount of RA1 in the lysate can be quantified through the apparent rate of the initial linear phase (k_{obsd} = 69.54 nM/min, Figure 4c, red trace) of the retro-aldol reaction, as long as the substrate concentration is >20-fold higher than the RA1 concentration. The amount of RA1 in the HEK293 lysate was quantified as 0.51 ± 0.19 μM by comparing the apparent rate to the standard curve (Figure S12). This quantification was confirmed using **P3** labeling (Figure S13). Fusion of RA1 to a POI (demonstrated herein with GFP) does not interfere with the function of RA1, demonstrating its applicability as a molecular reporter for a POI (Figure S14).

We have introduced *de novo*-designed enzymes as small-molecule-regulated fluorescence tags. Their inherent bioorthogonality obviates the need for protein engineering, allowing them to selectively react with small-molecule probes to create fluorescent conjugates for use in live cell imaging applications. The fluorescence reporter and the electrophoresis applications were also demonstrated. The slow turnover kinetics of RA relative to other tags (Table S2), while sufficient for the applications described herein, can be improved by further enzyme design and evolution, as recently demonstrated.[9] Enhancing the reactivity of the probes for shorter duration pulse-chase experiments should also be possible. An even smaller RA-tag would also be a welcome addition to the family. The strategy presented here can be extended to include other *de novo*-designed enzymes to expand our chemical biology toolbox for imaging, reporter, and analogous applications.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary figures, materials, experimental details, and synthetic methods. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

jkelly@scripps.edu

Author Contributions

[‖]Y.L. and X.Z. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NIH grant DK046335 (J.W.K.), the Skaggs Institute for Chemical Biology, and the Lita Annenberg Hazen Foundation. X.Z. was a Howard Hughes Medical Institute Fellow of the Helen Hay Whitney Foundation, and is currently supported by the Burroughs Wellcome Fund Career Award at Scientific Interface. Y.L.T. is supported by a predoctoral fellowship from the Agency of Science, Technology and Research (A*STAR), Singapore. G.B. is the Merck Fellow of the Damon Runyon Cancer Research Foundation (DRG-2136-12). D.C.E is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2140-12). We thank Prof. Virginia Cornish for generously providing FP vectors for imaging colocalization experiments.[17]

■ REFERENCES

(1) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 906.

(2) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509.

(3) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86.

(4) Sun, X. L.; Zhang, A. H.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M. Q.; Correa, I. R. *ChemBioChem* **2011**, *12*, 2217.

(5) Miller, L. W.; Sable, J.; Goelet, P.; Sheetz, M. R.; Cornish, V. W. *Angew. Chem., Int. Ed.* **2004**, *43*, 1672.

(6) Mizukami, S.; Hori, Y.; Kikuchi, K. *Acc. Chem. Res.* **2014**, *47*, 247.

(7) Wombacher, R.; Heidbreder, M.; van de Linde, S.; Sheetz, M. P.; Heilemann, M.; Cornish, V. W.; Sauer, M. *Nat. Methods* **2010**, *7*, 717.

(8) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387.

(9) Bjelic, S.; Kipnis, Y.; Wang, L.; Pianowski, Z.; Vorobiev, S.; Su, M.; Seetharaman, J.; Xiao, R.; Kornhaber, G.; Hunt, J. F.; Tong, L.; Hilvert, D.; Baker, D. *J. Mol. Biol.* **2014**, *426*, 256.

(10) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nat. Chem. Biol.* **2013**, *9*, 494.

(11) Liu, Y.; Tan, Y. L.; Zhang, X.; Bhabha, G.; Ekiert, D. C.; Genereux, J. C.; Cho, Y.; Kipnis, Y.; Bjelic, S.; Baker, D.; Kelly, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 4449.

(12) Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogyo, M. *Chem. Biol.* **2000**, *7*, 569.

(13) Groll, M.; Kim, K. B.; Kairies, N.; Huber, R.; Crews, C. M. *J. Am. Chem. Soc.* **2000**, *122*, 1237.

(14) Suh, E. H.; Liu, Y.; Connelly, S.; Genereux, J. C.; Wilson, I. A.; Kelly, J. W. *J. Am. Chem. Soc.* **2013**, *135*, 17869.

(15) Verdoes, M.; Hillaert, U.; Florea, B. I.; Sae-Heng, M.; Risseeuw, M. D. P.; Filippov, D. V.; van der Marel, G. A.; Overkleeft, H. S. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6169.

(16) Yun, S. W.; Leong, C.; Zhai, D. T.; Tan, Y. L.; Lim, L.; Bi, X. Z.; Lee, J. J.; Kim, H. J.; Kang, N. Y.; Ng, S. H.; Stanton, L. W.; Chang, Y. T. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 10214.

(17) Gallagher, S. S.; Sable, J. E.; Sheetz, M. P.; Cornish, V. W. *ACS Chem. Biol.* **2009**, *4*, 547.