JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

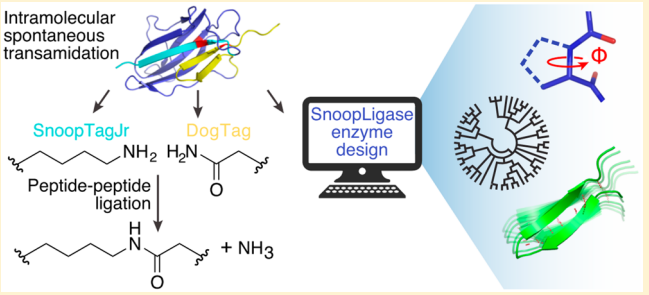
SnoopLigase Catalyzes Peptide–Peptide Locking and Enables Solid-Phase Conjugate Isolation

Can M. Buldun,[†,∥] Jisoo X. Jean,[†,∥] Michael R. Bedford,[‡] and Mark Howarth*[,†]

[†]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, U.K.
[‡]AB Vista, Woodstock Court, Marlborough, SN8 4AN, U.K.

Supporting Information

ABSTRACT: Simple, efficient reactions for connecting biological building-blocks open up many new possibilities. Here we have designed SnoopLigase, a protein that catalyzes site-specific transamidation, forming an isopeptide bond with more than 95% efficiency between two peptide tags, SnoopTagJr and DogTag. We initially developed these components by three-part splitting of the *Streptococcus pneumoniae* adhesin RrgA. The units were then engineered, guided by structure, bioinformatic analysis of sequence homology, and computational prediction of stability. After engineering, SnoopLigase demonstrated high-yield coupling under a wide range of buffers and temperatures. SnoopTagJr and DogTag were functional at the N- or C-terminus, while DogTag was also functional at internal sites in proteins. Having directed reaction of SnoopTagJr and DogTag, SnoopLigase remained stably bound to the ligated product, thus reconstituting the parent domain. Separating products from unreacted starting material and catalyst is often as challenging as reactions themselves. However, solid-phase immobilization of SnoopLigase enabled the ligated SnoopTagJr–DogTag product to be eluted with high purity, free from SnoopLigase or unligated substrates. The solid-phase catalyst could then be reused multiple times. In search of a generic route to improve the resilience of enzymes, we fused SnoopTagJr to the N-terminus and DogTag to the C-terminus of model enzymes, allowing cyclization via SnoopLigase. While wild-type phytase and β-lactamase irreversibly aggregated upon heating, cyclization using SnoopLigase conferred exceptional thermoresilience, with both enzymes retaining solubility and activity following heat treatment up to 100 °C. SnoopLigase should create new opportunities for conjugation and nanoassembly, while illustrating how to harness product inhibition and extend catalyst utility.



INTRODUCTION

Deepening our insight into the complexity of living systems will depend on having freedom to arrange components in arbitrary combinations or architectures. Site-specifically functionalized proteins open up a variety of avenues, such as stable single-molecule imaging,[1] targeting toxins to specific cell types,[2] and improving the circulation of protein drugs.[3,4] Proteins are the biomolecules with the widest range of functional activities, and the use of peptide tags is a powerful approach for controlling protein function. Peptide tags are simple to encode genetically, and their small size reduces the interference with natural interactions, the cost of biosynthesis, and the immunogenicity.[5]

Here we sought to design a generally applicable new catalyst for joining one peptide tag to another peptide tag irreversibly with high yield. We wanted all components to be composed of the regular 20 amino acids[6] for easy use in any cellular system, but excluding cysteine for application in oxidizing or reducing compartments. Our inspiration was the intramolecular isopeptide bond formation of certain Gram-positive bacterial surface proteins.[7] We previously split FbaB from *Streptococcus pyogenes*, generating SpyLigase, which directed isopeptide bond formation between two peptide tags.[8] SpyLigase was a useful proof of principle, but yield was rarely above 50%.[8,9] We hypothesized that splitting the *Streptococcus pneumoniae* adhesin RrgA would allow peptide–peptide ligation with better thermodynamic driving force, generating ammonia as a product rather than water for SpyLigase.[10,11] SpyLigase reaction also required precise and inconvenient conditions (i.e., 4 °C with 1.5 M trimethylamine *N*-oxide).[8] We considered that SpyLigase's condition-dependence reflected the fold instability, so that efforts to rigidify the ligase could make a more widely applicable catalyst.

Here we describe the process of dissecting RrgA, followed by structure-based and computational optimization to generate SnoopLigase, a robust technology for peptide–peptide ligation. We establish the tolerance of SnoopLigase to a wide range of buffers and temperatures. SnoopLigase's product inhibition is harnessed to enable clean conjugate purification, a major challenge in bioconjugation chemistry.[3] Then we demonstrate the use of SnoopLigase for head-to-tail ligation of different enzymes, achieving in one step a >60 °C increase in thermal resilience of an enzyme.

Received: December 14, 2017
Published: February 6, 2018

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

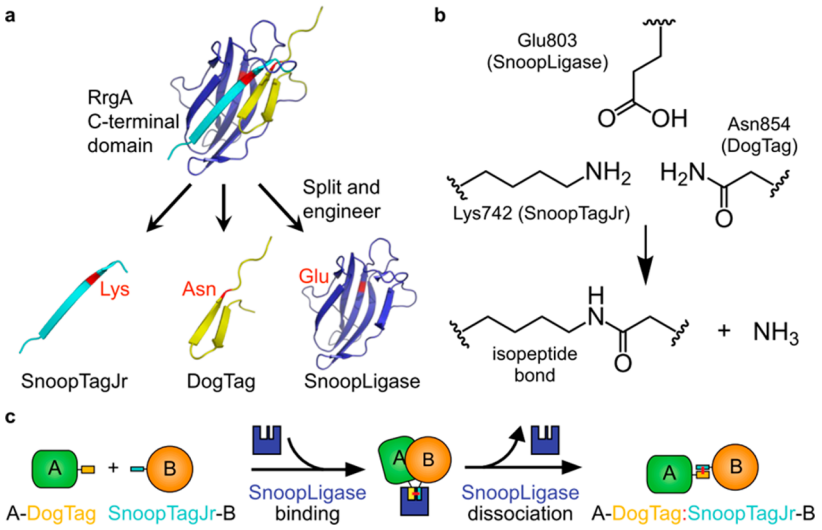


Figure 1. Design principle of SnoopLigase. (a) Cartoon of domain splitting. The C-terminal domain of RrgA (Protein Data Bank 2WW8) was split into three parts and engineered, such that the reactive Lys is located on SnoopTagJr (turquoise), the reactive Asn on DogTag (yellow), and the catalytic Glu on SnoopLigase (blue) (key residues highlighted in red). (b) Molecular basis for isopeptide bond formation in RrgA. Glu803 promotes isopeptide bond formation between Lys742 and Asn854, eliminating ammonia. (c) Schematic of the use of SnoopLigase to direct peptide−peptide ligation (the isopeptide bond is represented in red).

RESULTS AND DISCUSSION

Design Principle of the Peptide−Peptide Ligase. The C-terminal domain of the *S. pneumoniae* adhesin RrgA contains a spontaneous isopeptide bond between Lys and Asn, promoted by an apposed Glu (Figure 1A,B).[11] We had previously split this domain into two parts, enabling reaction of the 12 amino acid peptide SnoopTag with the 112 amino acid protein partner SnoopCatcher.[12] This pair was a strong foundation to split the C-terminal domain of RrgA into a trio, such that the residues of the reactive triad were located on three different units, to enable peptide−peptide ligation. After extensive variation of the sites of splitting, we settled on the units SnoopTag (reactive Lys, residues 734−745, 12 amino acids), DogTag (reactive Asn, residues 838−860, 23 amino acids), and RrgA ligase (catalytic Glu, residues 743−846, 104 amino acids) (Figure 1A, Figure S1). SnoopTag overlapped for three residues with RrgA ligase, while DogTag overlapped for nine residues with RrgA ligase (Figure S1). The G842T and D848G point mutations in RrgA, previously made to enhance SnoopCatcher reactivity with SnoopTag,[12] were included in RrgA ligase and DogTag. Our hypothesis was that the ligase should be able to bind SnoopTag and DogTag and catalyze the formation of the isopeptide bond between the two tags (Figure 1B), thereby mediating covalent conjugation of a SnoopTag fusion protein and a DogTag fusion protein (Figure 1C).

Structure-Based and Computational Optimization of Ligation Activity. Removal of three β-strands from a small protein domain is a major modification, and it is common for split proteins to have reduced stability.[13] Initial testing found that RrgA ligase showed low solubility when purified from *Escherichia coli*, and only a small percentage of ligation could be obtained, which was restricted to 4 °C and very precise buffer composition. We hypothesized that stabilizing the split domain would be important to enhance ligase performance. We initially sought to engineer β-turns of the protein domain, which are frequently flexible.[14] β-Turns may be stabilized by substitution of appropriate residues with proline,[15,16] which has a fixed φ-angle of approximately −60°, limiting the conformational flexibility of the polypeptide backbone. We analyzed each loop of RrgA's C-terminal domain and identified two promising proline substitutions (A808P and Q837P). To improve ligase stability further, we analyzed the domain computationally, integrating evidence on tolerated mutations from natural homologues of RrgA in the sequence databases, along with atomistic Rosetta modeling for the effect of mutations on fold energetics.[17] Five mutations (D737S, A820E, D830N, D838G, I839V) were taken forward; we did not pursue any suggested mutations adjacent to the isopeptide bond or introducing surface-exposed hydrophobic amino acid side-chains. Most mutations suggested by this analysis improved the fold stability, as computationally predicted by Rosetta (Figure S2A).[18] Three mutations (D737S, D838G, I839V) substantially improved reaction yield and rate (Figure 2A, Figure S2B−D). With the combination of A808P, Q837P, D838G, and I839V mutations (termed SnoopLigase), the reaction rate was increased 66-fold over RrgA ligase (Figure 2A, Figure S3). As expected, the Rosetta energy score was also improved (Figure S3). Similarly, D737S mutation in SnoopTag, termed SnoopTagJr, was also successful in improving reaction (Figure 2B). The sequences of all tags and ligases are aligned in Figure S1. We mapped the mutations on the structure of RrgA, illustrating the importance of the loop of SnoopLigase overlapping with DogTag (Figure 2C).

To validate the proposed route of SnoopLigase reaction, we analyzed peptide−peptide ligation with each of the three key residues mutated. We fused SnoopTagJr to an affibody against the growth factor receptor HER2 (SnoopTagJr-AffiHER2).[19] DogTag was fused with a model domain, small ubiquitin modifier (SUMO). SnoopLigase efficiently ligated SnoopTagJr-AffiHER2 to SUMO-DogTag (Figure 2D). However, mutation of the Lys in SnoopTagJr, the Asn in DogTag, or the Glu in SnoopLigase abolished product formation (Figure 2D), consistent with the reactive triad being responsible for covalent adduct formation. SnoopLigase was expressed efficiently in *E. coli* (>10 mg per L of culture) and was highly soluble (>500 μM). At equimolar substrate concentration, SnoopLigase-mediated conjugation reached 60−80% substrate conversion

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

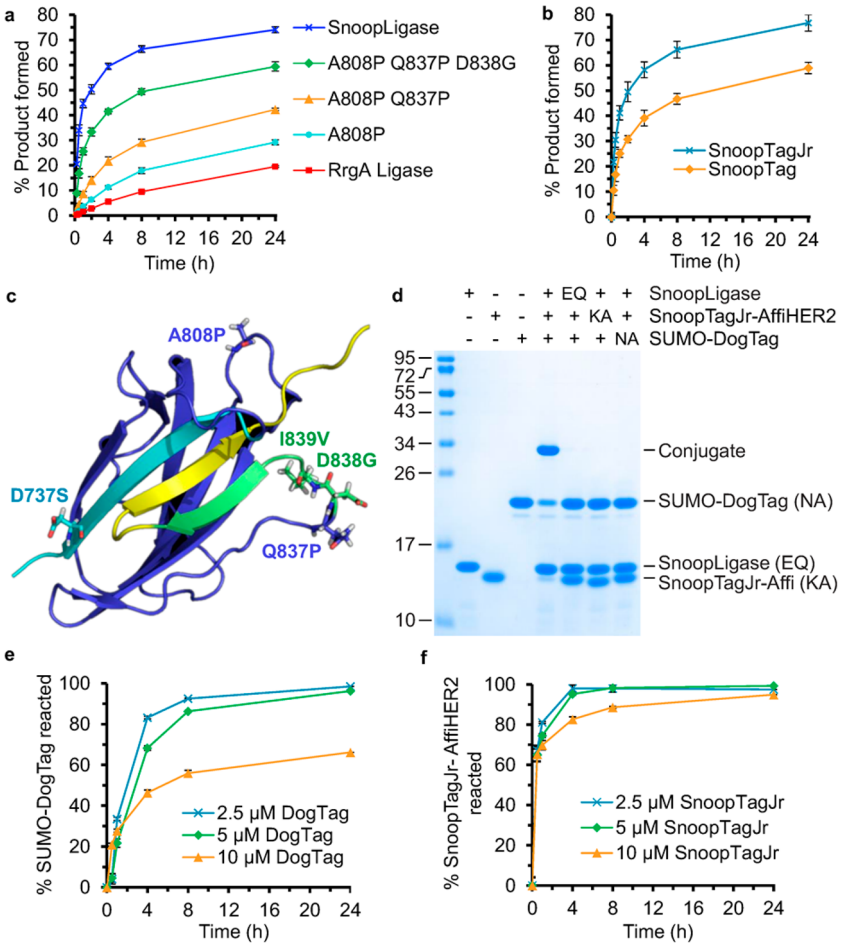


Figure 2. Engineering of SnoopLigase. (a) Mutations to enhance ligation. RrgA ligase (10 μM), point mutants thereof, or SnoopLigase was incubated with equimolar SnoopTagJr-AffiHER2 and SUMO-DogTag at 4 °C, and ligated product was determined by SDS-PAGE with Coomassie staining and densitometry. (b) SnoopTag and SnoopTagJr reactivity. SnoopTag- or SnoopTagJr-AffiHER2 (10 μM) was incubated with equimolar SUMO-DogTag and SnoopLigase at 4 °C. (c) Location of mutated residues enhancing RrgA ligase (blue) and SnoopTag (turquoise) within RrgA (from PDB 2WW8). Yellow represents DogTag, and green the region in both SnoopLigase and DogTag. (d) Specificity of residues in ligation. SnoopLigase was incubated with SnoopTagJr-AffiHER2 and SUMO-DogTag (each 10 μM) for 24 h at 4 °C, before SDS-PAGE and Coomassie staining. Mutations of key residues in each partner blocked reaction. (e) Maximizing DogTag conjugation. SnoopTagJr-AffiHER2 (10 μM) and 10 μM SnoopLigase were incubated with 2.5−10 μM SUMO-DogTag at 4 °C. (f) Maximizing SnoopTagJr-AffiHER2 conjugation. SUMO-DogTag (10 μM) and 10 μM SnoopLigase were incubated with 2.5−10 μM SnoopTagJr-AffiHER2 at 4 °C. Error bars are mean ± 1 SD, n = 3.

after 24 h. Employing an excess of the SnoopTagJr-partner and SnoopLigase enabled ≥95% of the DogTag-partner to react (2-fold excess 96.3 ± 0.4%; 4-fold excess 98.5 ± 0.3%, mean of triplicate ± 1 SD) (Figure 2E, Figure S4). Similarly, an excess of the DogTag-partner and SnoopLigase enabled ≥95% of the SnoopTagJr-partner to react (2-fold excess 99.3 ± 0.3%; 4-fold excess 97.5 ± 1.7%, mean of triplicate ± 1 SD) (Figure 2F, Figure S4). Kinetic analysis for SnoopLigase revealed K_M values of 8.4 ± 1.4 μM for SnoopTagJr-AffiHER2 and 11.9 ± 0.8 μM for SUMO-DogTag (Figure S5).

To explore the breadth of SnoopLigase substrates, we conjugated four different SnoopTagJr-linked proteins with four different DogTag-linked proteins. Both SnoopTagJr and DogTag were functional as C-terminal or N-terminal fusions. Coupling in most cases proceeded to over 95%, with the lowest yield at 79.5% (Figure S6). Since DogTag consists of two antiparallel β-strands, we hypothesized that DogTag could be inserted into protein loops. DogTag was inserted with six-residue flexible linkers on either side into HaloTag7 and MBP.[20,21] The DogTag-inserted constructs showed good soluble expression, and we found efficient conjugation by SnoopLigase (Figure S7).

SnoopLigase Was Active under Diverse Conditions. Having enhanced the activity of the peptide ligase under mild reaction conditions, we then explored SnoopLigase's tolerance to a range of different situations. SnoopLigase reaction functioned well from pH 7.25 to 8.75 (Figure 3A). Efficient ligation occurred over a wide range of temperatures (4−37 °C) (Figure 3B). SnoopLigase was functional in the presence of extracellular concentrations of NaCl, although reaction proceeded most efficiently with Tris-borate buffer in the absence of NaCl (Figure 3C). SnoopLigase reacted well in the presence of the commonly used detergents Tween 20 and Triton X-100 up to 2%, but sodium dodecyl sulfate (SDS) blocked the reaction (Figure 3D). Addition of the protein stabilizer glycerol slightly enhanced the reaction rate (Figure 3E). SnoopLigase was also thermoresilient, regaining full activity following heating up to 70 °C (Figure 3F). Similarly, after lyophilization and storage at 37 °C for 120 days, SnoopLigase retained nearly all of its activity following reconstitution (Figure S8A). Since there are no cysteines in

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

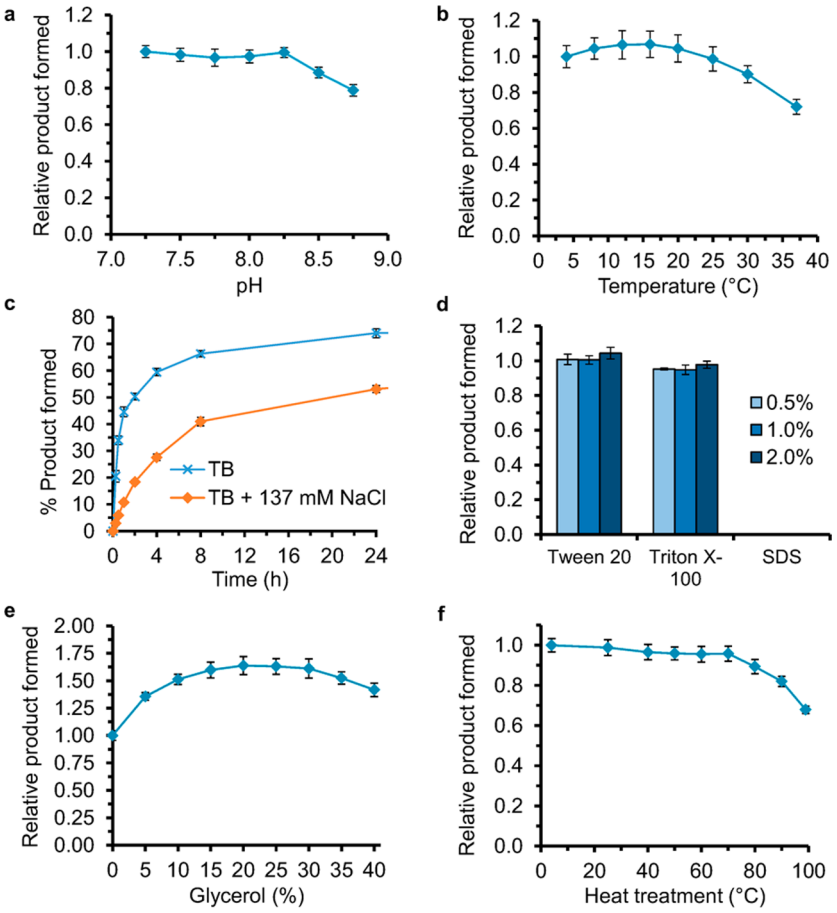


Figure 3. SnoopLigase reacted over a range of conditions. (a) pH-dependence. SnoopTagJr-AffiHER2 and SUMO-DogTag were ligated using SnoopLigase (10 μM each) for 1.5 h at 4 °C in TB + 15% (v/v) glycerol with the indicated pH. (b) Temperature-dependence. As in (a) at pH 7.25 from 4 to 37 °C. (c) Salt-dependence. As in (a) at pH 7.25 with or without additional NaCl. (d) Detergent-dependence. As in (a) at pH 7.25 with 0.5−2% Tween 20, Triton X-100, or SDS. (e) Glycerol-dependence. As in (a) at pH 7.25 with 0−40% glycerol. (f) SnoopLigase was thermoresilient. SnoopLigase was incubated at 4−100 °C for 15 min. After cooling, SnoopLigase was used for ligation of SnoopTagJr-AffiHER2 and SUMO-DogTag as in (a) at pH 7.25. Results are mean of triplicate ± 1 SD.

any of the units, ligation was unaffected by reducing conditions (Figure S8B).

SnoopLigase's Product Inhibition Enabled Solid-Phase Conjugate Purification. Having generated an interesting conjugate, one usually then must face the challenge of purifying the conjugate away from the catalyst and unreacted starting materials. SnoopLigase needed to be present at stoichiometric concentrations for efficient substrate coupling. This makes sense given its origin, where binding of each peptide to SnoopLigase would create a structure similar to the well-folded parent domain. Since SnoopLigase can be easily produced, the demand for high levels of SnoopLigase is not too problematic, but the product inhibition also created an opportunity to simplify substantially the generation of pure conjugate.

Upon reaction, SnoopLigase's strong binding to the reaction product could allow efficient purification of product, free from starting material or SnoopLigase itself (Figure 4A). After reacting SnoopTagJr-AffiHER2 with SUMO-DogTag using site-specifically biotinylated SnoopLigase, the ligase was captured by streptavidin-agarose resin. The strong interaction between biotin−streptavidin and SnoopLigase−reaction product permitted stringent washing, such that nonreacted and nonimmobilized proteins were removed. Incubation of the resin with glycine buffer at pH 2, as commonly used in elution from antibodies,[22] did not affect biotin−streptavidin interaction, but disrupted SnoopLigase−reaction product interaction and yielded high-purity ligated product (Figure 4B).

We also established an alternative solid-phase capture, forming a covalent bond between HaloLink resin and HaloTag7 fused to SnoopLigase. Here we achieved elution with 2 M imidazole at neutral pH to elute selectively the product of SnoopLigase reaction (Figure S9). Elution from biotin-SnoopLigase was also achieved with 2 M imidazole pH 7.0 (Figure 4B). To analyze whether this elution condition irreversibly damaged proteins, we tested fluorescence of mEGFP and mKate2, as well as binding activity of HaloTag7SS to HaloLink resin after incubation in 2 M imidazole and dialysis. All proteins fully retained their activity (Figure S10).

While many proteins survive incubation in pH 2.0 or 2 M imidazole, others may require a milder elution. We hypothesized that elution would be possible using conjugated SnoopTagJr:DogTag peptide to outcompete the conjugate of interest from immobilized SnoopLigase. We generated this competitor peptide via SUMO-DogTag and SUMO protease (Figure S11). Competition using SnoopTagJr:DogTag peptide allowed clean elution of product from biotin-SnoopLigase (Figure 4B).

Solid-phase purification eliminated the need for subsequent separation of the product from enzyme and unreacted started

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

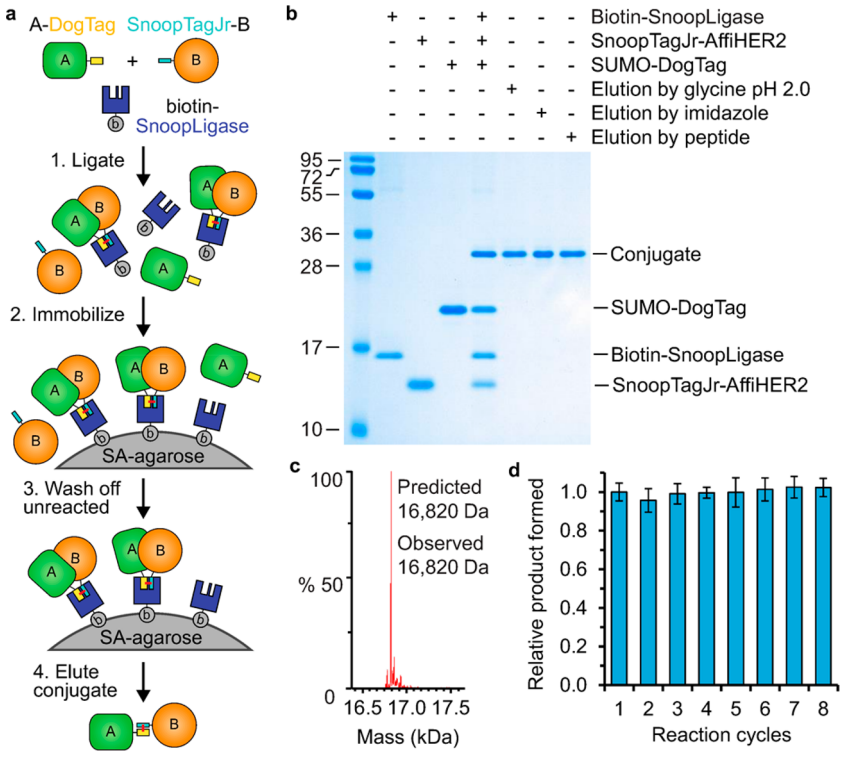


Figure 4. Purification of SnoopLigase reaction product. (a) Cartoon of solid-phase SnoopLigase purification. SnoopTagJr- and DogTag-linked proteins are covalently conjugated using biotin-SnoopLigase. Streptavidin-agarose binds biotin-SnoopLigase, unreacted proteins are washed away, and ligated proteins are eluted. (b) Analysis of product from SnoopLigase purification using three different elution methods. SnoopTagJr-AffiHER2 and SUMO-DogTag were ligated using biotin-SnoopLigase (10 μM each) for 16 h at 4 °C. Biotin-SnoopLigase was captured with streptavidin-agarose, followed by glycine pH 2.0, imidazole, or peptide elution procedures. Analysis by SDS-PAGE with Coomassie staining. (c) Electrospray ionization mass spectrometry shows SnoopTag linked to SUMO-DogTag with loss of ammonia, following reaction with SnoopLigase. (d) SnoopLigase performed multiple rounds of product purification. Biotin-SnoopLigase was bound to streptavidin resin and used for conjugation of SUMO-DogTag and SnoopTagJr-AffiHER2 for 1 h at 4 °C. The resin was washed and the SnoopLigase reaction product eluted, before repeating the cycle. The amount of product formed relative to the first cycle was analyzed by SDS-PAGE with Coomassie staining (mean ± 1 SD, n = 9).

materials by size exclusion chromatography, which is time-consuming and often leads to substantial losses. In line with the strong binding of SnoopLigase to the product, SnoopLigase could not be separated from the product by size exclusion chromatography (Figure S12).

Mass spectrometry of the solid-phase purified SUMO-DogTag after SnoopLigase-mediated ligation of SnoopTag peptide gave an increase in molecular weight consistent with the mass of the peptide minus 17 Da from loss of NH_3 (Figure 4C).

Immobilizing enzymes on solid phase can improve reaction efficiency and facilitate cost-effective reuse of purified enzymes.[23] To test whether SnoopLigase could be recycled, we immobilized biotin-SnoopLigase on streptavidin-agarose and performed a ligation reaction by addition of SnoopTagJr-AffiHER2 and SUMO-DogTag. Upon washing and elution of the reaction product, the SnoopLigase-coupled agarose was used for another ligation. The amount of product formed remained constant for at least eight reaction cycles, indicating that SnoopLigase can perform multiple turnovers and that SnoopLigase resin can be efficiently regenerated (Figure 4D).

SnoopLigase-Mediated Cyclization Made Other Enzymes Thermoresilient. Many enzymes are active under a narrow range of conditions and are irreversibly inactivated outside those conditions.[24] Thermal resilience holds back the use of many enzymes in green chemical transformations.[25] Also, in agriculture enzymes are widely added to animal feed to enhance animal health and productivity. These enzymes suffer from substantial inactivation by the steam-treatment of feed to kill pathogens.[26] Phytase hydrolyzes the antinutrient phytic acid, increasing the digestibility of phytate-bound phosphate and reducing environmental phosphate pollution.[26] Inactivation of phytase during heat treatment of feed poses a major challenge.[26] Given the enormous diversity of enzymes available from nature, there is a pressing need for generic approaches that can easily impart thermal resilience. Head-to-tail cyclization shows much potential to improve enzyme tolerance to harsh conditions, but with challenges including moderate yield of cyclization and insufficient increase of stability.[27,28]

To explore the stabilization potential of the SnoopLigase system, we cyclized the *Bacillus subtilis* phytase PhyC after genetically fusing SnoopTagJr at the N-terminus and DogTag at the C-terminus (Figure 5A). SnoopTagJr-PhyC-DogTag cyclized rapidly and to high yield upon SnoopLigase addition, as visualized by an increased electrophoretic mobility of the cyclized form (Figure 5B). Only a trace of higher order multimers from intermolecular reaction was seen (Figure 5B). Isopeptide bond formation in SnoopTagJr-PhyC-DogTag was confirmed by mass spectrometry (Figure S13). Similar efficient cyclization by SnoopLigase was seen for TEM-1 β-lactamase (BLA), a favored model system for enzyme evolution (Figure S14A,B).[29] PhyC is already moderately thermostable, surviving incubation at 55 °C. However, wild-type PhyC irreversibly aggregated at 75 °C, whereas SnoopLigase-cyclized Snoop-

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

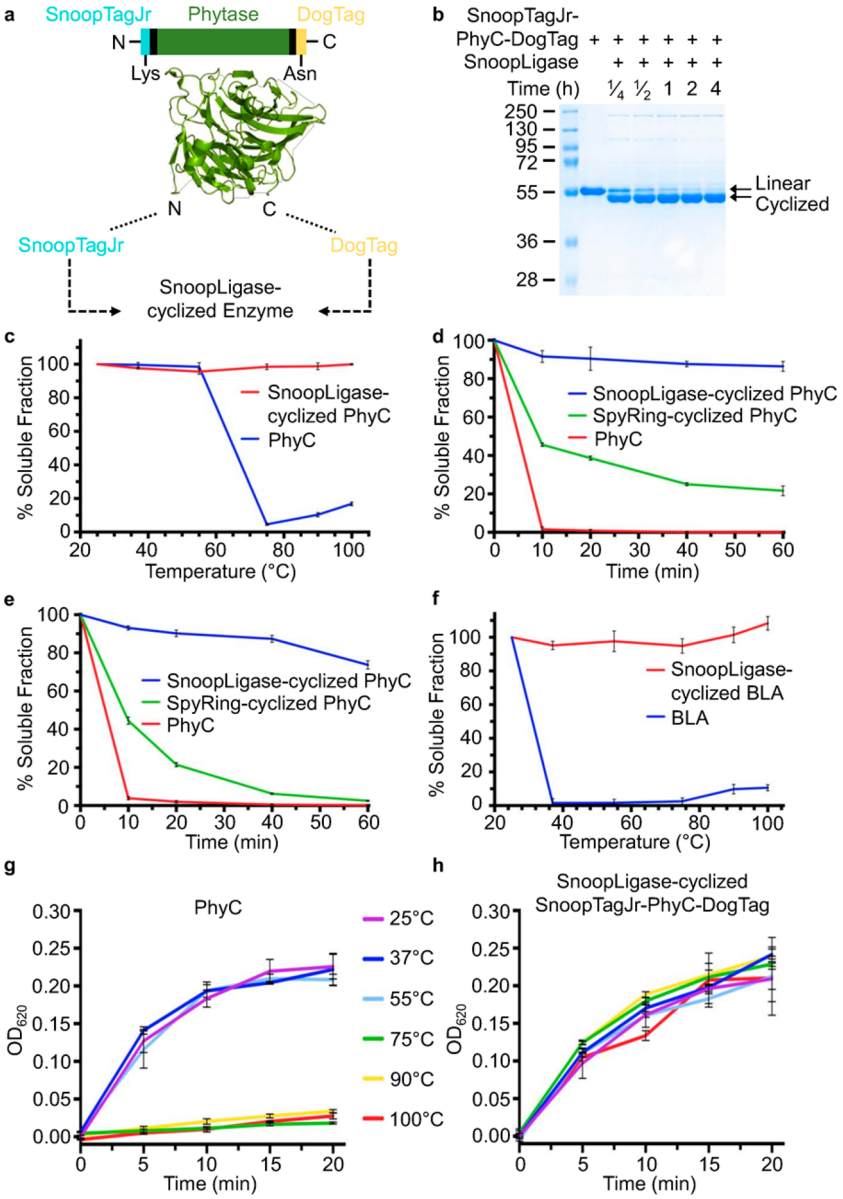


Figure 5. SnoopLigase cyclization conferred thermal resilience to enzymes. (a) Schematic of SnoopLigase enzyme cyclization. SnoopTagJr was positioned at the N-terminus and DogTag at the C-terminus of phytase, for SnoopLigase ligation, based on PDB 3AMR. (b) SnoopLigase efficiently cyclized SnoopTagJr-PhyC-DogTag. SnoopLigase and SnoopTagJr-PhyC-DogTag (10 μM each) were incubated at 4 °C for 0.25−4 h, before SDS-PAGE with Coomassie staining. (c) SnoopLigase cyclization increased thermal resilience of phytase. PhyC or SnoopLigase-cyclized SnoopTagJr-PhyC-DogTag was incubated for 10 min at the indicated temperature and centrifuged to remove aggregates, and enzyme remaining in the supernatant was quantified by SDS-PAGE with Coomassie staining. (d) PhyC, SpyRing-cyclized PhyC, or SnoopLigase-cyclized SnoopTagJr-PhyC-DogTag at 10 μM was incubated at 90 °C for the indicated times, and the soluble fraction was measured as in (c). (e) PhyC variant resilience as in (d) at 100 °C. (f) As in (c) for BLA. (g) PhyC was incubated for 10 min at the indicated temperature and cooled to 25 °C, then used in colorimetric assays to quantify phosphate release from the hydrolysis of phytic acid. (h) SnoopLigase cyclization improved thermal resilience of PhyC activity. As for (g) with cyclized SnoopTagJr-PhyC-DogTag. All are mean of triplicate ± 1 SD.

TagJr-PhyC-DogTag remained soluble up to 100 °C (Figure 5C). The improvement in PhyC solubility conferred by SnoopLigase cyclization was greater than that conferred by SpyTag/SpyCatcher (SpyRing cyclization), tolerating even 1 h at 100 °C (Figure 5D,E).[30,31]

Wild-type BLA irreversibly aggregated at 37 °C, whereas SnoopLigase-cyclized SnoopTagJr-BLA-DogTag remained soluble after boiling, representing an increase in resilience of >60 °C (Figure 5F). To investigate the origin of the thermoresilience, we analyzed uncyclized SnoopTagJr-BLA-DogTag. Surprisingly, this construct was less aggregation-prone than wild-type BLA, but the stabilization was less than with SnoopLigase cyclization (Figure S14C). Also, the catalytically inactive SnoopLigase EQ reduced the stabilization of SnoopTagJr-BLA-DogTag (Figure S14C). Titration of BLA amount validated that these activity assays were sensitive to the amount of functional enzyme (Figure S15).

To test whether catalytic activity was retained as well as solubility, we explored phytic acid hydrolysis by each PhyC construct. Indeed wild-type PhyC showed minimal activity after heating above 55 °C (Figure 5G), whereas SnoopLigase-cyclized PhyC was almost fully active after 100 °C heating

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

(Figure 5H). Good retention of activity was also seen for SnoopLigase-cyclized BLA after 100 °C heating (Figure S14D).

In the above tests, SnoopLigase was still present with cyclized enzyme. Using the solid-phase approach for conjugate purification, we tested the resilience of cyclized enzymes in the absence of SnoopLigase. Cyclized phytase after SnoopLigase depletion was much more resilient than linear phytase (in terms of both solubility and activity), but not quite as resilient as in the presence of SnoopLigase (Figures 5, S16). For BLA, SnoopLigase depletion had only a marginal effect on solubility or activity after heating (Figure S17). Overall, SnoopLigase-mediated cyclization shows strong potential as a simple method to achieve major improvements in enzyme thermoresilience.

CONCLUSION

Here we have designed a catalyst for peptide−peptide ligation by isopeptide bond formation, able to conjugate to more than 95% completion under diverse mild conditions. Despite unpromising origins from the minimal activity of the initial split trio, we achieved a dramatically improved catalyst through our program of optimizing the splitting sites, proline-mediated loop rigidification, and computational stabilization. Lys/Asn transamidation has not been discovered as a modular way to join peptides in nature. Transglutaminases from various species do carry out Lys/Gln transamidation, via a thioester intermediate, but show low target specificity and also deamidate.[32] SnoopLigase exhibits much greater yield and tolerance of conditions than SpyLigase.[8] It remains to be tested how far the stabilization strategy here could enhance SpyLigase's applicability, to provide a pair of orthogonal ligases.[12]

De novo design of enzymes generally gives kinetics inferior to those of natural enzymes, including for the landmark computational design of a catalyst for Kemp elimination.[33] Given how SnoopLigase was designed, it is not surprising that the split protein forms a stable complex re-forming the parent domain, so that SnoopLigase is majorly product-inhibited. The product inhibition prevents SnoopLigase from turning over substrate; however, turnover can be achieved by eluting the product and adding fresh substrate. This shows that SnoopLigase facilitates substrate conversion without itself undergoing a permanent chemical modification.

Due to product inhibition, SnoopLigase needs to be used at an equimolar concentration to the substrates. However, for many applications the amount of enzyme is not the limiting factor: achieving efficient and specific coupling without extreme concentrations of substrate is more important. Indeed notable natural enzymes act as single turnover catalysts, including in DNA repair, vitamin biosynthesis, and CRISPR/Cas9.[34,35] For the widely used catalyst sortase, a challenge has been the need for high concentrations of the oligoglycine partner (K_M 140 μM), despite a program of directed evolution.[36] Therefore, the oligoglycine partner is often desired at millimolar concentrations, which cannot be achieved for many proteins. Here we showed efficient SnoopLigase coupling with 2.5 μM DogTag or SnoopTagJr. In contrast to the recognition motifs of most conjugation methods (e.g., sortase, split inteins, butelase, OaAEP1),[3,37−39] SnoopTagJr and DogTag can be used as N- or C-terminal fusions, while DogTag also reacts efficiently when inserted within protein domains. This flexibility for tag insertion should extend the range of protein architectures achievable through peptide ligation.[40,41]

To address the challenge of conjugate purification, enzymes such as sortase have been coupled directly to cyanogen bromide-activated resin[23] or conjugated with a bile acid for cyclodextrin-resin removal;[42] nevertheless, leftover substrate will remain. We made a virtue of the product inhibition of SnoopLigase, with solid-phase SnoopLigase attachment permitting elution of pure product after all starting material had been washed away. We report three different conditions for product elution and confirmed that elution conditions allowed retention of activity for fluorescent proteins, a ligand binding protein, and two enzymes.

We demonstrated the applicability of SnoopLigase cyclization to increase the thermal resilience of two different enzymes. Resilience is a key limiting factor for enzyme application. Resilience is usually enhanced by painstaking testing of hundreds to billions of enzyme variants,[25,29] so it is important to look for rapid and generic routes to achieve resilience. Enzyme cyclization by SnoopLigase was achieved to high yield in both cases. Cyclization has been a popular testing ground for bioconjugation approaches, including through carbodiimide chemistry,[43] split inteins,[27] sortase,[37] SpyTag/SpyCatcher,[31] butelase,[38] and OaAEP1,[39] providing increases in protease resistance, circulation time, and thermodynamic stability.[28] However, the extent of thermoresilience achieved by SnoopLigase is surprising (with both enzymes resisting inactivation at 100 °C), surpassing previous approaches.[44] There is certainly more at play than the restricted conformational freedom of the enzyme termini after cyclization,[30] since depletion of SnoopLigase decreased the resilience of cyclized phytase. The increased resilience of BLA following fusion of SnoopTagJr and DogTag in the absence of cyclization also suggests an adventitious solubilizing effect of these sequences, which should be explored in future work.

Apart from the insights here for split protein redesign, SnoopLigase should provide many new opportunities for nanoassembly, including multispecific antibodies,[9,12,45] responsive biomaterials,[46,47] and rapid vaccine construction.[41,48]

EXPERIMENTAL SECTION

Cloning. Plasmid constructs for protein expression were cloned using standard PCR procedures and Gibson isothermal assembly. Nucleotide sequences of gene inserts were validated by Sanger sequencing. Constructs for expression in *E. coli* contained an N-terminal His_6-tag followed by a flexible GS-rich linker, except that DogTag-mClover3 and DogTag-mKate2 contained a C-terminal His_6-tag preceded by a flexible GS-rich linker. Residue numbers of all RrgA-derived variants are based on the numbering of RrgA from Protein Data Bank ID code 2WW8.[11] pET28a-RrgA ligase was derived by a deletion from pET28a-SnoopCatcher.[12] Sequential point mutations were introduced into pET28a-RrgA ligase, finally giving pET28a-SnoopLigase (Figure S1, GenBank accession no. MG867372). pET28a-SnoopLigase EQ was generated from pET28a-SnoopLigase, preventing catalysis by mutating Glu803 to Gln. pET28a-AviTag-SnoopLigase (Addgene plasmid ID 105626) contained an N-terminal AviTag for site-specific biotinylation.[49] pET28a-HaloTag7-SnoopLigase (Addgene plasmid ID 105627 and GenBank accession no. MG867371) was derived from pFC14A HaloTag CMV Flexi (Promega) and pET28a-SnoopLigase. pET28a-HaloTag7SS-SnoopTagJr was generated from pET28a-HaloTag7-SnoopLigase, with C61S and C261S mutations in HaloTag7.[21] pET28a-HaloTag7SS-DogTag inserted [HaloTag7SS with DogTag flanked by $(GS)_3$ linkers on either side, inserted between residues D139 and E140][20] was derived from pET28a-HaloTag7SS-SnoopTagJr. pET28a-SnoopTagJr-MBP (Addgene plasmid ID 105628 and GenBank accession no. MG867374) was derived from pET28a-SnoopTag-MBP.[12] pET28a-MBP-DogTag in-

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

serted [MBP with DogTag flanked by $(GS)_3$ linkers on either side inserted between residues R317 and A319 (I318 deleted)][21] was derived from pET28a-SnoopTagJr-MBP. pET28a-SnoopTagJr-AffiHER2 (encoding SnoopTagJr N-terminal to an affibody against HER2) and pET28a-AffiHER2-DogTag were derived from pET28a-SnoopTag-AffiHER2.[48] pET28a-SUMO-DogTag (Addgene plasmid ID 105629 and GenBank accession no. MG867376) was derived from pET28a-SUMO-KTag.[8] pET28a-SnoopTagJr-mEGFP was derived from pET28a-SnoopTag-mEGFP-SpyTag.[12] pET28a-DogTag-mClover3 was derived from pET28a-SpyTag002-mClover3.[50] mKate2[51] was a kind gift from Stephan Uphoff (University of Oxford) and was used to clone pET28a-DogTag-mKate2. pET28a-SnoopTagJr-BLA-DogTag (TEM-1 β-lactamase flanked by SnoopTagJr and DogTag, GenBank accession no. MG867373) and variants thereof were derived from pET28a-BLA.[31] pET28a-SnoopTagJr-PhyC-DogTag (*B. subtilis* phytase flanked by SnoopTagJr and DogTag, GenBank accession no. MG867375) was derived from pET28a-PhyC.[30] pET28a-SpyTag-PhyC-SpyCatcher has been described.[30]

Protein Expression and Purification. Expression plasmids were transformed into *E. coli* BL21 (DE3)-RIPL (Agilent), except for SnoopTagJr-mEGFP, which was transformed into *E. coli* C41,[52] a kind gift of Anthony Watts (University of Oxford). Individual colonies were grown in LB with 50 μg/mL kanamycin for 16 h at 37 °C, 200 rpm. Starter cultures were diluted 1:100 in LB with 0.8% (w/v) glucose and 50 μg/mL kanamycin, except for SnoopTagJr-BLA-DogTag, HaloTag7-SnoopLigase, HaloTag7SS-SnoopTagJr, and HaloTag7SS-DogTag inserted, which were grown without glucose. Cultures were grown at 37 °C, 200 rpm until OD_{600} 0.5. Cultures were induced with 0.42 mM IPTG and grown for 4 h at 30 °C, 200 rpm, except for HaloTag7SS-SnoopTagJr and HaloTag7SS-DogTag inserted, which were grown for 16 h at 20 °C, 200 rpm, before harvesting. Proteins were purified using standard Ni-NTA methods (Qiagen) and dialyzed three times. Buffers for dialysis were TB pH 8.0 (50 mM Tris base adjusted to pH 8.0 with boric acid) for AviTag-SnoopLigase, SnoopTagJr-MBP, and MBP-DogTag inserted; TB pH 7.4 (50 mM Tris base adjusted to pH 7.4 with boric acid) for HaloTag7SS-SnoopTagJr; 50 mM sodium borate pH 10.0 for RrgA ligase (and point mutants), SnoopLigase, SnoopTag-AffiHER2, SnoopTagJr-AffiHER2, and SUMO−-DogTag; PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) for BLA-containing constructs, SnoopTagJr-mEGFP, DogTag-mClover3, and DogTag-mKate2; 50 mM Tris·HCl pH 8.0 for HaloTag7SS-DogTag inserted; 50 mM Tris·HCl pH 7.0, 2 mM $CaCl_2$ for PhyC-containing constructs. Biotinylation of AviTag-SnoopLigase was performed as described previously.[49] His-tagged SUMO protease Ulp1 in pOPINE has been described.[53] To determine protein concentrations, OD_{280} was measured using an ND-1000 Nanodrop (NanoDrop) with the extinction coefficient predicted by ExPASy ProtParam.

SnoopLigase Reactions. In standard conditions to assess the formation of the isopeptide bond between SnoopTagJr and DogTag, proteins were incubated at 10 μM each in TB pH 7.25 + 15% (v/v) glycerol at 4 °C for 2 h, unless indicated otherwise. To measure pH-dependence, reactions were run in standard conditions, but in 50 mM Tris base adjusted to the indicated pH with boric acid. To measure temperature-dependence, reactions were run in standard conditions from 4 to 37 °C. To measure NaCl-dependence, reactions were run in standard conditions ± 137 mM NaCl. To measure detergent-dependence, reactions were run in standard conditions with 0.5%, 1%, or 2% Tween 20 (v/v), Triton X-100 (v/v), or SDS (w/v). To measure glycerol-dependence, reactions were run in standard conditions with 0−40% (v/v) glycerol. To measure reducing agent-dependence, reactions were run in standard conditions with 100 mM 2-mercaptoethanol or 20 mM dithiothreitol. To terminate the reaction, 6× SDS loading buffer [0.23 M Tris·HCl, pH 6.8, 24% (v/v) glycerol, 120 μM bromophenol blue, 0.23 M SDS] was added to a final concentration of 1×.

Bioinformatic and Computational Design of SnoopLigase Mutations. To identify residues for proline substitution, Ramachandran analysis of amino acid residues in RrgA (PDB code 2WW8) was performed using MolProbity.[54] Loop residues with φ-angles of −70° to −50° were considered for proline substitution. Homologous sequences for RrgA residues 734−860 were collected using Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST)[55] and aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE).[56] Cluster Database at High Identity with Tolerance (CD-HIT)[57] was used to minimize sequence redundancy and tune the size of the data set before PROSS analysis.[17] The output amino acid substitutions were reviewed manually in PyMOL (DeLano Scientific).

Rosetta Modeling of RrgA Mutations. Modeling of mutations in RrgA was performed using Rosetta3.[18] The crystal structure of RrgA (PDB code 2WW8) residues 734−860 with G842T, N847D, and D848G mutations was relaxed, and the pmut_scan protocol was used to calculate Rosetta energy units for mutants.

Purification of SnoopLigase Reaction Product by Glycine Elution. SUMO-DogTag, SnoopTagJr-AffiHER2, and biotin-SnoopLigase at 10 μM each in TB pH 7.25 with 15% (v/v) glycerol in a total volume of 200 μL were incubated for 20 h at 4 °C. To capture SnoopLigase, 25 μL of washed and equilibrated HiCap streptavidin-agarose (Thermo Fisher) was added, and samples were incubated for 30 min at 25 °C on a tube rotor. The resin was collected in a 1 mL poly prep column (Bio-Rad) and spun for 1 min at 300*g* at 4 °C. All subsequent steps were performed at 4 °C. After washing the resin twice with 125 μL of 50 mM glycine pH 3.0 with 300 mM NaCl and three times with 125 μL of 50 mM glycine pH 3.0, one extra spin for 1 min at 500*g* ensured the removal of excess liquid from the resin. To elute the SnoopLigase reaction product, the resin was incubated with 25 μL of antibody elution buffer (50 mM glycine pH 2.0) for 1 min, before spinning the eluate into a tube containing 2.5 μL of 1 M Tris·HCl pH 9.5 for 1 min at 300*g*, to neutralize the eluate. The elution was repeated twice more.

Purification of SnoopLigase Reaction Product by Imidazole Elution. For capture with HaloLink resin, SUMO-DogTag, SnoopTagJr-AffiHER2, and HaloTag7-SnoopLigase at 15 μM each in TB pH 7.25 with 15% (v/v) glycerol in a total volume of 200 μL were incubated for 24 h at 4 °C. Imidazole was added to a final concentration of 0.5 M and Tween 20 to a final concentration of 0.01% (v/v). To capture HaloTag7-SnoopLigase, 20 μL of washed and equilibrated HaloLink resin (Promega) was added, and samples were incubated for 1 h at 25 °C on a tube rotor. The resin was collected in a buffer-equilibrated 1 mL poly prep column (Bio-Rad) and spun for 1 min at 300*g* at 25 °C. After washing the resin five times with 150 μL of Tris-phosphate pH 7.0 (25 mM phosphoric acid adjusted to pH 7.0 with Tris base) with 500 mM imidazole pH 7.0 (adjusted with HCl) and 0.01% (v/v) Tween 20 at 25 °C, one extra spin for 1 min at 500*g* and 25 °C ensured the removal of excess liquid from the resin. To elute the reaction product, the resin was incubated with 20 μL of Tris-phosphate with 2 M imidazole pH 7.0 and 0.01% (v/v) Tween 20 for 5 min at 25 °C on a Thermomixer Comfort (Eppendorf) at 800 rpm, before spinning the eluate into a tube for 1 min at 300*g*, 25 °C. The elution was repeated twice more.

For capture with streptavidin resin, enzymes flanked by SnoopTagJr and DogTag were incubated with biotin-SnoopLigase at 10 μM in TB pH 7.25 + 15% (v/v) glycerol at 4 °C for 16 h. Cyclization reactions of PhyC were supplemented to a final concentration of 2 mM $CaCl_2$. Following cyclization, 10 μL of washed and equilibrated streptavidin-agarose was added per 100 μL reaction and incubated at 25 °C for 30 min on a tube rotor. Poly prep columns of 1 mL were prepared by washing with 1 M imidazole + 0.01% Tween-20. The reaction was loaded on the columns and spun for 1 min at 300*g*, followed by five washes with 5 resin volumes of 0.5 M imidazole + 0.01% (v/v) Tween 20 in 250 mM Tris-phosphate pH 7.0 at 25 °C. Columns were spun for 1 min at 500*g*. To elute cyclized enzyme, the poly prep columns were capped before the addition of 1 resin volume of 2.5 M imidazole + 0.01% (v/v) Tween 20 in 250 mM Tris-phosphate pH 7.0. The columns were incubated for 2 min on a Thermomixer at 800 rpm, 25 °C. This was repeated three more times for a final elution volume of 4 resin volumes. All wash and elution buffers were supplemented to a final concentration of 2 mM $CaCl_2$ for phytase elution. The eluted enzymes were dialyzed into their optimum buffer and concentrated in 5 kDa molecular weight cutoff filters (Vivaspin 500, GE Healthcare).

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

SnoopLigase Removal by Peptide Elution. SUMO-DogTag, SnoopTagJr-AffiHER2, and biotin-SnoopLigase at 10 μM each in TB pH 7.25 with 15% (v/v) glycerol in a total volume of 150 μL were incubated for 16 h at 4 °C. Tween 20 was added to a final concentration of 0.01% (v/v). To capture biotin-SnoopLigase, 15 μL of washed and equilibrated HiCap streptavidin-agarose (Thermo Fisher) was added, and the sample incubated for 30 min at 25 °C on a tube rotor. The resin was collected in a PCR tube (StarLab) and spun for 1 min at 300*g* at 25 °C, followed by five washes with 75 μL of Tris-phosphate pH 7.0 with 0.01% (v/v) Tween 20. A 30 μL amount of DogTag:SnoopTagJr in TB pH 7.5 with 0.01% (v/v) Tween 20 was added, and the sample incubated for 4 h at 37 °C at 800 rpm on a Thermomixer. The sample was centrifuged for 1 min at 16900*g* and the supernatant collected.

For cyclized enzyme purification, enzymes flanked by SnoopTagJr and DogTag were incubated with biotin-SnoopLigase (each 10 μM) for 16 h at 4 °C in TB pH 7.25 + 15% (v/v) glycerol. A 10 μL amount of washed and equilibrated HiCap streptavidin-agarose (Thermo Fisher) was added per 100 μL cyclization reaction and incubated for 30 min at 25 °C on a tube rotor. The reaction was spun down for 1 min at 6000*g* at 25 °C, followed by five washes with 5 resin volumes of Tris-phosphate pH 7.0 + 0.01% (v/v) Tween 20. A 2 μL amount of 100 μM DogTag:SnoopTagJr was added per μL of resin in TB pH 7.5 + 0.01% (v/v) Tween 20 and incubated for 4 h at 37 °C, 800 rpm on a Thermomixer. The sample was centrifuged for 1 min at 16900*g*, and the supernatant collected.

Analyzing the Effect of Imidazole on Model Protein Function. SnoopTagJr-mEGFP, DogTag-mKate2, or HaloTag7SS-SnoopTagJr (25 μM each) were incubated in Tris-phosphate pH 7.0 (25 mM phosphoric acid adjusted to pH 7.0 with Tris base) with or without 2 M imidazole pH 7.0 (adjusted with HCl) for 10 min at 25 °C, followed by dialysis into 50 mM Tris·HCl pH 8.0. Clear, 96-well, flat-bottom polystyrene plates (Greiner) were blocked with 45 μL of 50 mM Tris·HCl pH 8.0 containing 1% (w/v) BSA (bovine serum albumin, Sigma-Aldrich) in each well at 25 °C for 10 min before addition of fluorescent proteins at 0.5 μM final concentration. Fluorescence was recorded at λ_{ex} = 485 nm, λ_{em} = 538 nm for SnoopTagJr-mEGFP and λ_{ex} = 544 nm, λ_{em} = 612 nm for DogTag-mKate2 using a SpectraMax M3 microplate reader (Molecular Devices). To measure HaloTag7SS-SnoopTagJr activity, the dialyzed protein at 5 μM was incubated with 2 μL of washed and equilibrated HaloLink resin (Promega) in a total volume of 20 μL for 5 min at 25 °C on a tube rotor. Samples were centrifuged for 30 s at 16900*g* at 25 °C, and 10 μL of supernatant was collected. The amount of HaloTag7SS-SnoopTagJr coupled to the resin was determined by comparing the amount of protein in the supernatant to the input, using SDS-PAGE with Coomassie staining and densitometry.

Production of DogTag:SnoopTagJr Competitor. A 4 mL amount of HaloTag7-SnoopLigase at 20 μM in 50 mM TB pH 7.25 with 0.01% (v/v) Tween 20 was incubated with 500 μL of packed HaloLink resin (Promega) for 2 h at 25 °C on a tube rotor. The sample was split into five buffer-equilibrated 1 mL poly prep columns (Bio-Rad) and spun for 1 min at 300*g* at 25 °C. Each resin sample was washed twice with 500 μL of 50 mM TB pH 7.25 with 0.01% (v/v) Tween 20. Columns were capped, and 200 μL of reaction buffer [50 μM SUMO-DogTag and 75 μM SnoopTagJr peptide in TB pH 7.25 with 15% (v/v) glycerol] was added to each column. SnoopTagJr peptide had the sequence GKLGSIEFIKVNKGY and was solid-phase synthesized by Activotec at >95% purity. After incubation for 4 h at 25 °C at 300 rpm on a Thermomixer, samples were spun for 1 min at 300*g* at 25 °C, and each resin sample was washed five times with 500 μL of Tris-phosphate pH 7.0 with 0.5 M imidazole and 0.01% (v/v) Tween 20. To elute the SnoopLigase reaction product, each resin sample was incubated with 100 μL of Tris-phosphate with 2.5 M imidazole pH 7.0 and 0.01% (v/v) Tween 20 for 2 min at 25 °C on a Thermomixer at 800 rpm, before spinning the eluate into a tube for 1 min at 300*g*, at 25 °C. The elution was repeated twice more, and each resin washed twice with 500 μL of TB pH 7.25 with 0.01% (v/v) Tween 20. To start the next reaction cycle, fresh reaction mix was added to the resin and the reaction and purification procedure repeated. Six reaction cycles were performed in total. All elutions were pooled and dialyzed into TB pH 7.5, and SUMO-DogTag:SnoopTagJr was concentrated to 118 μM using a 10 kDa MWCO spin filter (Sartorius). SUMO protease Ulp1 was added at 1:50 molar ratio to a final concentration of 2.4 μM, followed by a 45 min incubation at 25 °C. After reaction, Tween 20 was added to a final concentration of 0.01% (v/v). To deplete His-tagged proteins (SUMO and Ulp1), 600 μL of sample was incubated with 150 μL of packed Ni-NTA agarose (Qiagen) for 1 h at 25 °C on a tube rotor, the sample was centrifuged for 1 min at 16900*g* at 25 °C, and the supernatant containing the DogTag:SnoopTagJr conjugate was collected. The concentration was calculated using the OD_{280} extinction coefficient from ExPASy ProtParam.

SnoopLigase K_M Determination. To determine the K_M for SnoopTagJr-AffiHER2, SnoopLigase at 5 μM and SUMO-DogTag at 40 μM were incubated with 3.75−60 μM SnoopTagJr-AffiHER2 in TB pH 7.25 + 15% (v/v) glycerol at 4 °C for 7.5 min. To terminate the reaction, 6× SDS loading buffer was added to a final concentration of 1×. The K_M of SUMO-DogTag was determined the same way but with SnoopTagJr-AffiHER2 at 40 μM and SUMO-DogTag at 3.75−60 μM. To quantify the amount of product formed, a product standard for gel densitometry was generated by conjugating 10 μM SnoopTagJr-AffiHER2 to SUMO-DogTag at 20 μM using 20 μM SnoopLigase under standard reaction conditions for 48 h. The molar ratios ensured >95% conjugation of SnoopTagJr-AffiHER2, such that the concentration of SnoopTagJr-AffiHER2:SUMO-DogTag was considered to be 10 μM. Samples were analyzed by SDS-PAGE with Coomassie staining, followed by densitometry. Reaction rates were fit to the Michaelis−Menten equation to determine K_M values.

Lyophilization Stability. Aliquots of 30 μL of SnoopLigase at 10 μM in TB pH 7.25 were prepared in 100 μL thin-wall PCR tubes (StarLab). Samples were flash-frozen in a dry ice−ethanol bath for 10 min and lyophilized using a BenchTop 2K freeze-dryer (VirTis) for 48 h at 0.14 mbar and −72.5 °C. Lyophilized samples were stored at 37 °C for the indicated time in a glass scintillation vial sealed with Parafilm (Sigma-Aldrich) on a bed of Drierite (Sigma-Aldrich) to minimize sample hydration. Samples were reconstituted in reaction buffer, and the reaction of SnoopTagJr-AffiHER2 and SUMO-DogTag was performed for 2 h at 4 °C, followed by SDS-PAGE, Coomassie staining, and densitometry.

Size Exclusion Chromatography of a SnoopLigase Reaction. AffiHER2-DogTag, SnoopTagJr-MBP, and biotin-SnoopLigase at 50 μM each in TB pH 7.25 with 15% (v/v) glycerol in a total volume of 1000 μL were incubated for 24 h at 4 °C. Imidazole pH 7.0 was added to a final concentration of 0.5 M and Tween 20 to a final concentration of 0.01% (v/v). To capture biotin-SnoopLigase, 300 μL of washed and equilibrated HiCap streptavidin-agarose (Thermo Fisher) was added, and samples were incubated for 30 min at 25 °C on a tube rotor. The sample was collected in three buffer-equilibrated 1 mL poly prep columns (Bio-Rad) and spun for 1 min at 300*g* at 25 °C. After washing the resin five times with 500 μL of Tris-phosphate pH 7.0 (25 mM phosphoric acid adjusted to pH 7.0 with Tris base) with 500 mM imidazole pH 7.0 (adjusted with HCl) and 0.01% Tween 20 (v/v) at 25 °C, one extra spin for 1 min at 500*g* at 25 °C ensured the removal of excess liquid from the resin. To elute the SnoopLigase reaction product, the resin was incubated with 100 μL of Tris-phosphate with 2 M imidazole pH 7.0 and 0.01% (v/v) Tween 20 for 5 min at 25 °C on a Thermomixer at 300 rpm, before spinning the eluate into a tube for 1 min at 300*g*, at 25 °C. The elution was repeated twice more. A 500 μL sample of the purified elution, a reaction sample without product purification, and 50 μM biotin-SnoopLigase, 50 μM AffiHER2-DogTag, or 50 μM SnoopTagJr-MBP were dialyzed into 50 mM sodium borate pH 10.0. A 200 μL amount of the samples was applied to a previously equilibrated Superdex 75 Increase 10/300 GL (GE Healthcare) on a fast protein liquid chromatography system Purifier 10 (GE Healthcare) at 4 °C, using 50 mM sodium borate pH 10.0 as mobile-phase column buffer.

Solid-Phase Ligation Reaction Cycles. Biotin-SnoopLigase at 50 μM in TB pH 8.0 was coupled to 10 μL of washed and equilibrated HiCap streptavidin-agarose in a total volume of 50 μL for 30 min at 25

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

°C on a tube rotor. The resin was collected in a 1 mL poly prep column and spun for 1 min at 300*g*, followed by five washes with 100 μL of TB pH 8.0 at 25 °C. The reaction was started by addition of 50 μL of reaction mix [50 μM SUMO-DogTag and 50 μM SnoopTagJr-AffiHER2 in TB pH 7.25 with 15% (v/v) glycerol], and the sample was incubated for 1 h at 4 °C on a Thermomixer at 800 rpm. The reaction mixture was spun for 1 min at 300*g*, and the resin washed twice with 50 μL of 50 mM glycine pH 3.0 with 300 mM NaCl and three times with 50 μL of 50 mM glycine pH 3.0, all at 4 °C. One extra spin for 1 min at 500*g* at 4 °C ensured the removal of excess liquid from the resin. To elute the SnoopLigase reaction product, the resin was incubated with 10 μL of antibody elution buffer for 1 min, before spinning the eluate into a tube containing 1 μL of 1 M Tris·HCl pH 9.5 for 1 min at 300*g*. The elution was repeated twice more. The resin was washed twice with 100 μL of antibody elution buffer and twice with 100 μL of TB pH 7.25 to ensure complete removal of residual reaction product, all at 4 °C. To start the next reaction cycle, fresh reaction mix was added to the resin, and the reaction and purification procedure repeated. Eight reaction cycles were performed in total.

SnoopLigase Thermostability. SnoopLigase at 12.5 μM in TB pH 7.25 with 15% (v/v) glycerol was incubated at the indicated temperature for 15 min on a C1000 thermal cycler (Bio-Rad) and cooled to 4 °C for 5 min. Heat-treated SnoopLigase was used for ligation of SnoopTagJr-AffiHER2 and SUMO-DogTag as described above.

Enzyme Cyclization by SnoopLigase. To cyclize enzymes flanked by SnoopTagJr and DogTag, 10 μM enzyme construct and 10 μM SnoopLigase were incubated in TB pH 7.25 + 15% (v/v) glycerol at 4 °C for 16 h. Cyclization reactions of SnoopTagJr-PhyC-DogTag for thermal resilience activity assays were supplemented to a final concentration of 2 mM $CaCl_2$. For the cyclization time-course, the reaction was terminated by adding SDS loading buffer as above.

Temperature-Dependent Solubility Assays. For BLA constructs, 20 μL of 10 μM enzyme in PBS with 100 mM dithiothreitol was incubated at 10, 25, 37, 55, 75, 90, or 100 °C for 10 min, then cooled to 10 °C (ramp rate 3 °C/s) on a C1000 thermal cycler. For PhyC constructs, 20 μL of 10 μM enzyme in 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$ was incubated in the same manner. For time-course assays, 20 μL of 10 μM PhyC construct in 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$ was incubated at 90 or 100 °C for 10, 20, 40, or 60 min, then cooled to 10 °C (ramp rate 3 °C/s) on a C1000 thermal cycler. Following heating, aggregated proteins were pelleted by centrifugation at 4 °C, 16900*g* for 30 min. The supernatant was run on SDS-PAGE. The samples held at 10 °C were defined as 100% soluble.

BLA Thermal Resilience Activity Assays. Clear, 96-well, flat-bottom polystyrene plates (Greiner) were blocked with 200 μL of PBS and 3% (w/v) BSA (Sigma-Aldrich) in each well at 37 °C for a minimum of 2 h. Blocked plates were washed twice with 0.1 M NaH_2PO_4 pH 7.0 with 1 mM ethylenediaminetetraacetic acid (EDTA). A 20 μL amount of 25 μM BLA constructs was incubated at 10, 25, 37, 55, 75, 90, or 100 °C for 10 min, then cooled to 10 °C (ramp rate 3 °C/s) on a C1000 thermal cycler (Bio-Rad). The samples were diluted to a final concentration of 62.5 nM using 0.1 M NaH_2PO4 pH 7.0 with 1 mM EDTA. Nitrocefin (Merck) substrate solution was prepared to 105 μM through the addition of 0.1 M NaH_2PO_4 pH 7.0 with 1 mM EDTA. The diluted BLA constructs were reacted at a final concentration of 3 nM with 100 μM nitrocefin. A SpectraMax M3 microplate reader (Molecular Devices) was used to measure OD_{486} every 15 s for 10 min at 20 °C, with automatic mixing for 3 s between measurements. Blanks consisted of all components except the enzyme.

PhyC Thermal Resilience Activity Assays. Clear, 96-well, flat-bottom polystyrene plates (Greiner) were blocked with 200 μL of 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$ and 3% (w/v) BSA at 37 °C for 2 h. Blocked plates were washed twice with 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$ and 10% (v/v) glycerol. Enzyme at 10 μM was incubated at 25, 37, 55, 75, 90, and 100 °C for 10 min, then cooled to 10 °C (ramp rate 3 °C/s) on a C1000 thermal cycler (Bio-Rad). The samples were diluted to 375 nM using 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$. Phytic acid sodium salt (Sigma-Aldrich) substrate solution at 25 μM was prepared using 50 mM Tris·HCl pH 7.0 with 2 mM $CaCl_2$. In the final reaction, 75 nM phytase construct reacted with 20 μM phytic acid at 50 °C, 350 rpm on a Thermomixer. Reactions were set up such that all time-points finished at the same time. A 100 μL amount of Biomol Green (Enzo Life Sciences) was added immediately to each well, and the plate was incubated at 25 °C for 25 min for color development. The SpectraMax M3 microplate reader was used to measure OD_{620} at 20 °C. Blanks consisted of all components except the enzyme. Some error bars are too small to be visible.

SDS-PAGE and Reaction Quantification. Samples were mixed with 6× SDS loading buffer to a final concentration of 1×. For all samples containing BLA constructs, dithiothreitol was added to a final concentration of 100 mM. For all samples containing PhyC constructs, EDTA was added to a final concentration of 100 mM. Samples were heated for 3 min at 99 °C and allowed to cool to 25 °C for 10 min before loading. SDS-PAGE was performed at 200 V in 25 mM Tris·HCl, 192 mM glycine, and 0.1% (w/v) SDS. Gels were stained with InstantBlue Coomassie stain (Expedeon), destained with Milli-Q water, and imaged using a ChemiDoc XRS imager with ImageLab software (Bio-Rad). ImageLab was also used for band quantification. Percent product formed was calculated from band intensities as 100 × [product]/([product] + [DogTag substrate] + [SnoopTagJr substrate]). Relative product formed was calculated as percentage of product formed under that condition divided by the percentage of product formed from the control (Figure 3A pH 7.25; Figure 3B 4 °C, Figure 3D no detergent, Figure 3E 0% glycerol, Figure 3F 4 °C, Figure 4D cycle 1, Figure S8A nonlyophilized, Figure S8B no reducing agent). Percent partner reacted was calculated from band intensities as 100 × (1 − [substrate after reaction]/[substrate before reaction]).

Mass Spectrometry. SUMO-DogTag at 75 μM and a SnoopTag-containing solid-phase synthesized peptide (GKLGDIEFIKVNKGY, Insight Biotechnology at 95% purity) at 300 μM were incubated with 75 μM biotin-SnoopLigase in TB pH 7.25 and 15% (v/v) glycerol in a total volume of 200 μL for 36 h at 4 °C. The reaction product was purified with glycine elution as above, but with 100 μL of HiCap streptavidin-agarose and 500 μL of wash buffer. Analysis of this reaction or SUMO-DogTag alone was performed using a Micromass LCT time-of-flight electrospray ionization mass spectrometer (Micromass). The molecular mass profile was created from the *m*/*z* spectrum using V4.00.00 software (Waters) with a maximum entropy algorithm. The observed mass of SUMO-DogTag was consistent with the mass predicted by ExPASy ProtParam, based on the amino acid sequence without N-terminal fMet and with one acetylation from *E. coli* expression. The increase in mass after reaction of SUMO-DogTag with peptide was predicted from the mass of the peptide calculated by ExPASy ProtParam and loss of ammonia (17.0 Da) during isopeptide bond formation. SnoopTagJr-PhyC-DogTag or cyclized SnoopTagJr-PhyC-DogTag following SnoopLigase removal was dialyzed into 10 mM ammonium acetate + 2 mM $CaCl_2$. Mass spectrometry and analysis were performed as above.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b13237.

Figures S1−S17 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*mark.howarth@bioch.ox.ac.uk

ORCID

Jisoo X. Jean: 0000-0003-0398-372X

Mark Howarth: 0000-0002-4586-7201

Author Contributions

‖C. M. Buldun and J. X. Jean contributed equally.

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Notes

The authors declare the following competing financial interest(s): M.H. and C.M.B. are authors on a patent application covering peptide-peptide ligation: UK Intellectual Property Office 1705750.6.

■ ACKNOWLEDGMENTS

Funding for C.M.B. was provided by the Engineering and Physical Sciences Research Council (EPSRC) and Corpus Christi College Oxford. Funding for J.X.J. was provided by the Biotechnology and Biological Sciences Research Council (BBSRC) and AB Vista. We thank Matteo Ferla (University of Oxford) for bioinformatic assistance.

■ REFERENCES

(1) Liu, Z.; Lavis, L. D.; Betzig, E. *Mol. Cell* **2015**, *58* (4), 644−659.
(2) McCluskey, A. J.; Collier, R. J. *Mol. Cancer Ther.* **2013**, *12* (10), 2273−2281.
(3) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol.* **2011**, *7* (12), 876−884.
(4) Liu, Z.; Chen, X. *Chem. Soc. Rev.* **2016**, *45* (5), 1432−1456.
(5) Wood, D. W. *Curr. Opin. Struct. Biol.* **2014**, *26C*, 54−61.
(6) Chin, J. W. *Annu. Rev. Biochem.* **2014**, *83*, 379−408.
(7) Kang, H. J.; Baker, E. N. *Trends Biochem. Sci.* **2011**, *36* (4), 229−237.
(8) Fierer, J. O.; Veggiani, G.; Howarth, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (13), E1176−1181.
(9) Siegmund, V.; Piater, B.; Zakeri, B.; Eichhorn, T.; Fischer, F.; Deutsch, C.; Becker, S.; Toleikis, L.; Hock, B.; Betz, U. A.; Kolmar, H. *Sci. Rep.* **2016**, *6*, 39291.
(10) Charville, H.; Jackson, D. A.; Hodges, G.; Whiting, A.; Wilson, M. R. *Eur. J. Org. Chem.* **2011**, *2011* (30), 5981−5990.
(11) Izore, T.; Contreras-Martel, C.; El Mortaji, L.; Manzano, C.; Terrasse, R.; Vernet, T.; Di Guilmi, A. M.; Dessen, A. *Structure* **2010**, *18* (1), 106−115.
(12) Veggiani, G.; Nakamura, T.; Brenner, M. D.; Gayet, R. V.; Yan, J.; Robinson, C. V.; Howarth, M. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (5), 1202−1207.
(13) Shekhawat, S. S.; Ghosh, I. *Curr. Opin. Chem. Biol.* **2011**, *15* (6), 789−797.
(14) Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R. *Chem. Rev.* **2016**, *116* (11), 6391−6423.
(15) Trevino, S. R.; Schaefer, S.; Scholtz, J. M.; Pace, C. N. *J. Mol. Biol.* **2007**, *373* (1), 211−218.
(16) Fu, H.; Grimsley, G. R.; Razvi, A.; Scholtz, J. M.; Pace, C. N. *Proteins: Struct., Funct., Genet.* **2009**, *77* (3), 491−498.
(17) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; Lieberman, R. L.; Aharoni, A.; Silman, I.; Sussman, J. L.; Tawfik, D. S.; Fleishman, S. J. *Mol. Cell* **2016**, *63* (2), 337−346.
(18) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y. E.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popovic, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. *Methods Enzymol.* **2011**, *487*, 545−574.
(19) Wikman, M.; Steffen, A. C.; Gunneriusson, E.; Tolmachev, V.; Adams, G. P.; Carlsson, J.; Stahl, S. *Protein Eng., Des. Sel.* **2004**, *17* (5), 455−462.
(20) Ishikawa, H.; Meng, F.; Kondo, N.; Iwamoto, A.; Matsuda, Z. *Protein Eng., Des. Sel.* **2012**, *25* (12), 813−820.
(21) Ke, N.; Landgraf, D.; Paulsson, J.; Berkmen, M. *J. Bacteriol.* **2016**, *198* (7), 1035−1043.
(22) Moser, A. C.; Hage, D. S. *Bioanalysis* **2010**, *2* (4), 769−790.
(23) Witte, M. D.; Wu, T.; Guimaraes, C. P.; Theile, C. S.; Blom, A. E.; Ingram, J. R.; Li, Z.; Kundrat, L.; Goldberg, S. D.; Ploegh, H. L. *Nat. Protoc.* **2015**, *10* (3), 508−516.
(24) Sanchez-Ruiz, J. M. *Biophys. Chem.* **2010**, *148* (1−3), 1−15.
(25) Stepankova, V.; Bidmanova, S.; Koudelakova, T.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. *ACS Catal.* **2013**, *3* (12), 2823−2836.
(26) Lei, X. G.; Weaver, J. D.; Mullaney, E.; Ullah, A. H.; Azain, M. J. *Annu. Rev. Anim. Biosci.* **2013**, *1*, 283−309.
(27) Iwai, H.; Pluckthun, A. *FEBS Lett.* **1999**, *459* (2), 166−172.
(28) Borra, R.; Camarero, J. A. *Biopolymers* **2013**, *100* (5), 502−509.
(29) Bershtein, S.; Goldin, K.; Tawfik, D. S. *J. Mol. Biol.* **2008**, *379* (5), 1029−1044.
(30) Schoene, C.; Bennett, S. P.; Howarth, M. *Sci. Rep.* **2016**, *6*, 21151.
(31) Schoene, C.; Fierer, J. O.; Bennett, S. P.; Howarth, M. *Angew. Chem., Int. Ed.* **2014**, *53* (24), 6101−6104.
(32) Stamnaes, J.; Fleckenstein, B.; Sollid, L. M. *Biochim. Biophys. Acta, Proteins Proteomics* **2008**, *1784* (11), 1804−1811.
(33) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453* (7192), 190−195.
(34) Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. *Nature* **2014**, *507* (7490), 62−67.
(35) Roach, P. *Nature* **2011**, *478* (7370), 463−464.
(36) Chen, I.; Dorr, B. M.; Liu, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (28), 11399−11404.
(37) Parthasarathy, R.; Subramanian, S.; Boder, E. T. *Bioconjugate Chem.* **2007**, *18* (2), 469−476.
(38) Nguyen, G. K.; Kam, A.; Loo, S.; Jansson, A. E.; Pan, L. X.; Tam, J. P. *J. Am. Chem. Soc.* **2015**, *137* (49), 15398−15401.
(39) Yang, R.; Wong, Y. H.; Nguyen, G. K. T.; Tam, J. P.; Lescar, J.; Wu, B. *J. Am. Chem. Soc.* **2017**, *139* (15), 5351−5358.
(40) Zhang, W. B.; Sun, F.; Tirrell, D. A.; Arnold, F. H. *J. Am. Chem. Soc.* **2013**, *135* (37), 13988−13997.
(41) Brune, K. D.; Leneghan, D. B.; Brian, I. J.; Ishizuka, A. S.; Bachmann, M. F.; Draper, S. J.; Biswas, S.; Howarth, M. *Sci. Rep.* **2016**, *6*, 19234.
(42) Rosen, C. B.; Kwant, R. L.; MacDonald, J. I.; Rao, M.; Francis, M. B. *Angew. Chem., Int. Ed.* **2016**, *55* (30), 8585−8589.
(43) Goldenberg, D. P.; Creighton, T. E. *J. Mol. Biol.* **1983**, *165* (2), 407−413.
(44) Rebello, S.; Jose, L.; Sindhu, R.; Aneesh, E. M. *Appl. Microbiol. Biotechnol.* **2017**, *101* (7), 2677−2689.
(45) Yumura, K.; Akiba, H.; Nagatoishi, S.; Kusano-Arai, O.; Iwanari, H.; Hamakubo, T.; Tsumoto, K. *J. Biochem.* **2017**, *162* (3), 203−210.
(46) Sun, F.; Zhang, W. B.; Mahdavi, A.; Arnold, F. H.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (31), 11269−11274.
(47) Botyanszki, Z.; Tay, P. K.; Nguyen, P. Q.; Nussbaumer, M. G.; Joshi, N. S. *Biotechnol. Bioeng.* **2015**, *112* (10), 2016−2024.
(48) Brune, K. D.; Buldun, C. M.; Li, Y.; Taylor, I. J.; Brod, F.; Biswas, S.; Howarth, M. *Bioconjugate Chem.* **2017**, *28* (5), 1544−1551.
(49) Fairhead, M.; Howarth, M. *Methods Mol. Biol.* **2015**, *1266*, 171−184.
(50) Keeble, A. H.; Banerjee, A.; Ferla, M. P.; Reddington, S. C.; Anuar, I. N. A. K.; Howarth, M. *Angew. Chem., Int. Ed.* **2017**, *56* (52), 16521−16525.
(51) Shcherbo, D.; Murphy, C. S.; Ermakova, G. V.; Solovieva, E. A.; Chepurnykh, T. V.; Shcheglov, A. S.; Verkhusha, V. V.; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M. *Biochem. J.* **2009**, *418* (3), 567−574.
(52) Dumon-Seignovert, L.; Cariot, G.; Vuillard, L. *Protein Expression Purif.* **2004**, *37* (1), 203−206.
(53) Assenberg, R.; Delmas, O.; Graham, S. C.; Verma, A.; Berrow, N.; Stuart, D. I.; Owens, R. J.; Bourhy, H.; Grimes, J. M. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2008**, *64*, 258−262.
(54) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.;

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

985 Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*,
986 W375−383.
987 (55) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang,
988 Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25* (17), 3389−
989 3402.
990 (56) Edgar, R. C. *BMC Bioinf.* **2004**, *5*, 113.
991 (57) Li, W.; Godzik, A. *Bioinformatics* **2006**, *22* (13), 1658−1659.

DOI: 10.1021/jacs.7b13237
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Supporting Information:

SnoopLigase catalyzes peptide-peptide locking and enables solid-phase conjugate isolation

Can M Buldun,[†] Jisoo X Jean,[†] Michael R Bedford[‡] and Mark Howarth[†*]
[†]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK.
[‡]AB Vista, Woodstock Court, Marlborough, SN8 4AN, UK.
*Correspondence and requests for materials should be addressed to M.H.
(email: mark.howarth@bioch.ox.ac.uk).

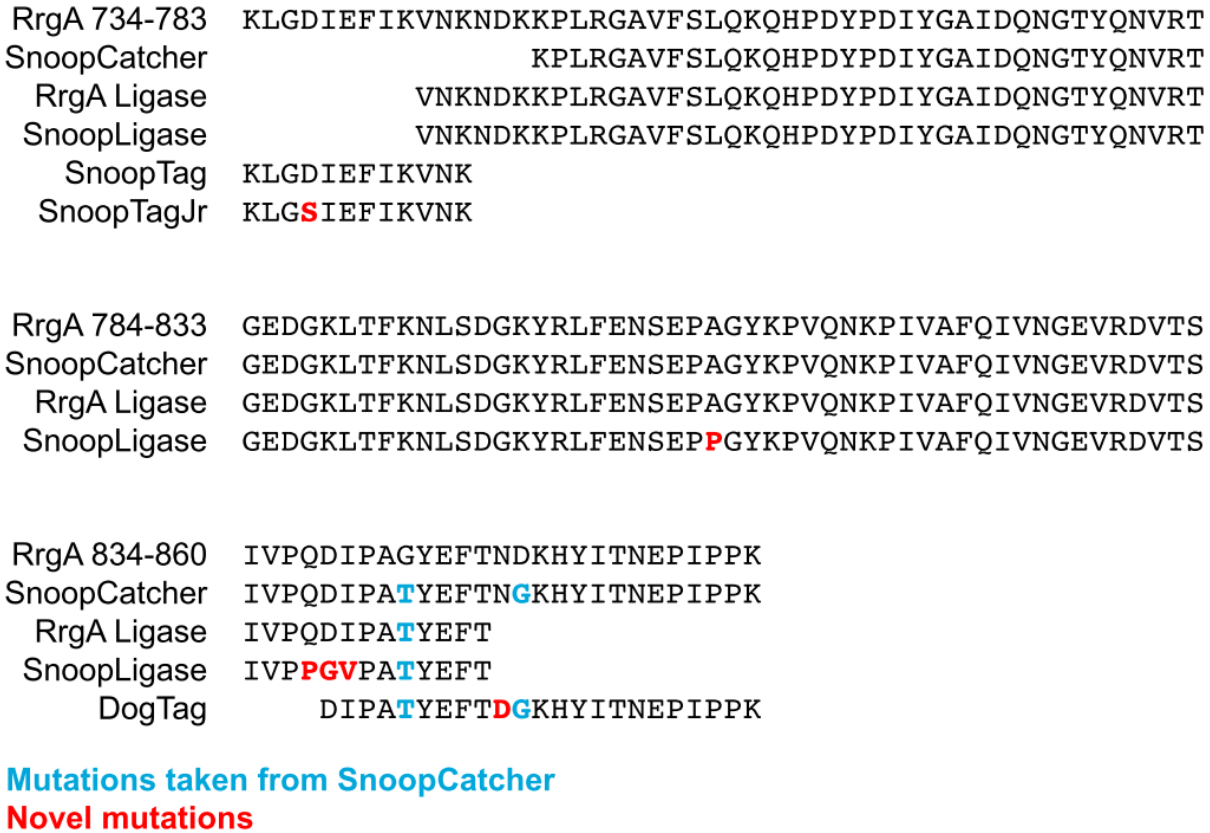


Figure S1. Amino acid sequences of partners. Sequence alignment of the C-terminal domain of RrgA and proteins/peptides derived from this domain (SnoopCatcher, RrgA Ligase, SnoopLigase, SnoopTag, SnoopTagJr and DogTag). Previously published mutations are highlighted in cyan and novel mutations in red.

a

Mutation	R.E.U.	∆R.E.U.
Parent, RrgA (734-860 G842T N847D D848G)	-267.2	+0.0
D737S	-268.6	-1.4
A820E	-268.5	-1.3
D830N	-266.6	+0.6
D838G	-267.5	-0.3
I839V	-267.6	-0.4

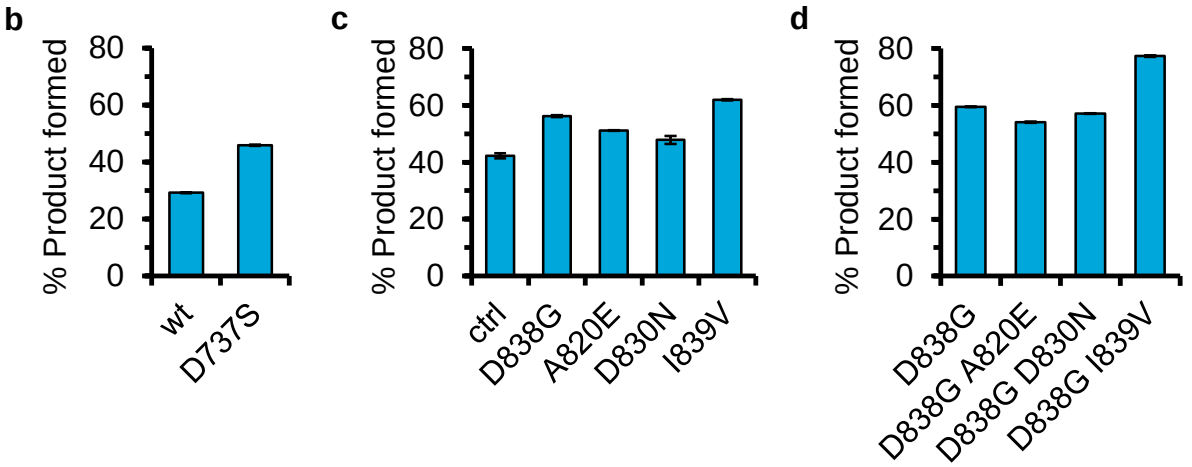
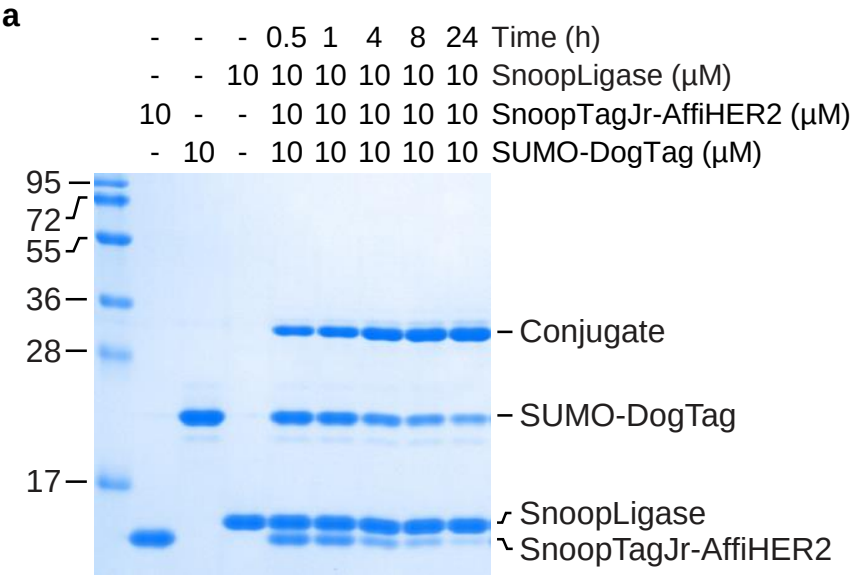


Figure S2. Point mutations to enhance SnoopLigase reactivity. (a) Rosetta Energy units (R.E.U.) of mutations in RrgA suggested by PROSS. The change in R.E.U. is shown relative to the parent protein. **(b)** Reactivity of wild-type (wt) SnoopTag and SnoopTag D737S with +6RrgALigase+9 A808P and SUMO-DogTag in 50 mM boric acid with 1 M TMAO pH 9.0 and 30% (v/v) glycerol for 3 h at 4 °C. **(c)** Reactivity of +6RrgALigase+9 A808P (ctrl) and single point mutants thereof with SnoopTagJr-AffiHER2 and SUMO-DogTag in 50 mM Tris borate with 0.5 M TMAO pH 8.0 and 30% (v/v) glycerol for 4 h at 4 °C. **(d)** Reactivity of +6RrgALigase+9 A808P Q837P D838G (D838G) and point mutants thereof with SnoopTagJr-AffiHER2 and SUMO-DogTag in 50 mM Tris borate with 0.75 M TMAO pH 8.0 and 7.5% (v/v) glycerol for 1.5 h at 4 °C. Results are mean of triplicate ± 1 SD; some error bars are too small to be visible.

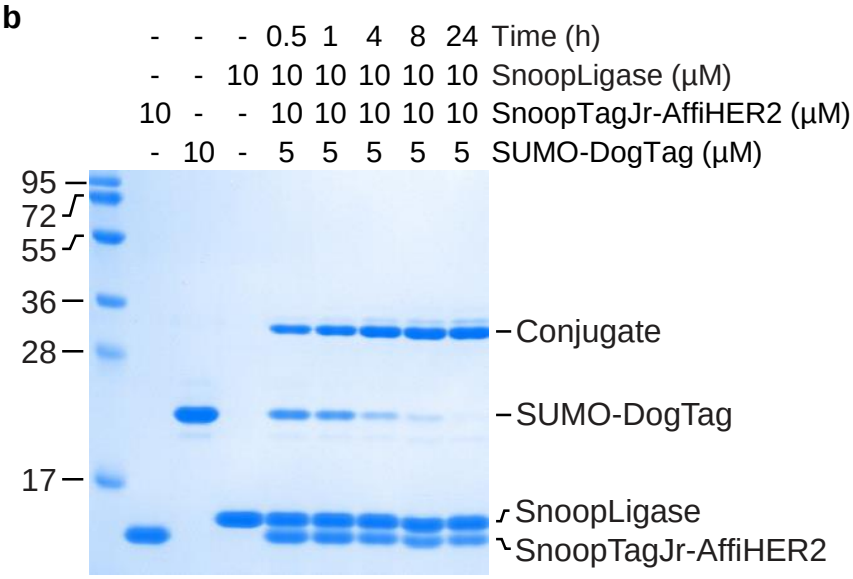
Mutation	R.E.U.	∆R.E.U.	Relative reaction rate
control (RrgALigase)	-268.6	+0.0	1.0
A808P	-269.8	-1.2	3.1
A808P Q837P	-269.5	-0.9	6.9
A808P Q837P D838G	-270.5	-1.9	28
A808P Q837P D838G I839V (SnoopLigase)	-270.9	-2.3	66

Figure S3. Rosetta Energy units (R.E.U.) and relative reaction rates of RrgALigase mutants. The change in R.E.U. and relative reaction rate with SnoopTagJr-AffiHER2 and SUMO-DogTag is given relative to RrgALigase. 10 µM RrgA Ligase or point mutants thereof was incubated with equimolar SnoopTagJr-AffiHER2 and SUMO-DogTag for 15 min at 4 °C and ligated product was determined by SDS-PAGE with Coomassie staining. Relative reaction rate for each mutant was calculated as the amount of product formed from the mutant divided by the amount of product formed from the control (RrgALigase).

a



b



c

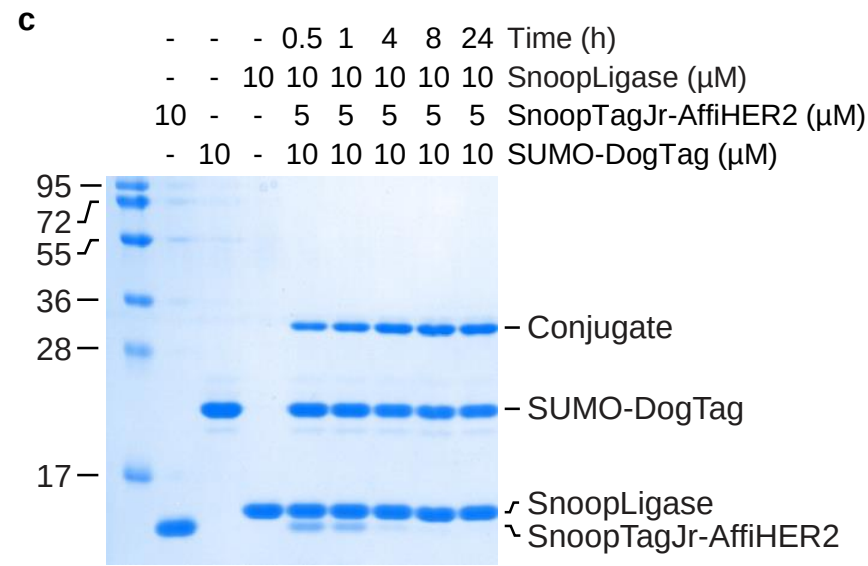
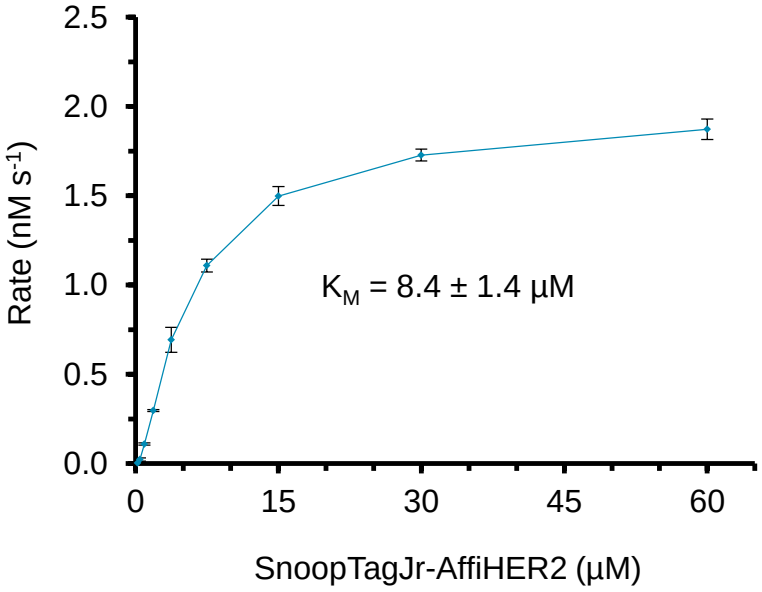


Figure S4. Maximizing SnoopLigase conjugation yield. **(a)** Equimolar reaction. 10 µM SnoopTagJr-AffiHER2 and 10 µM SnoopLigase were incubated with 10 µM SUMO-DogTag at 4 °C. **(b)** Maximizing SUMO-DogTag conjugation. 10 µM SnoopTagJr-AffiHER2 and 10 µM SnoopLigase were incubated with 5 µM SUMO-DogTag at 4 °C. **(c)** Maximizing SnoopTagJr-AffiHER2 conjugation.10 µM SUMO-DogTag and 10 µM SnoopLigase were incubated with 5 µM SnoopTagJr-AffiHER2 at 4 °C. Analysis by SDS-PAGE with Coomassie staining.

a



b

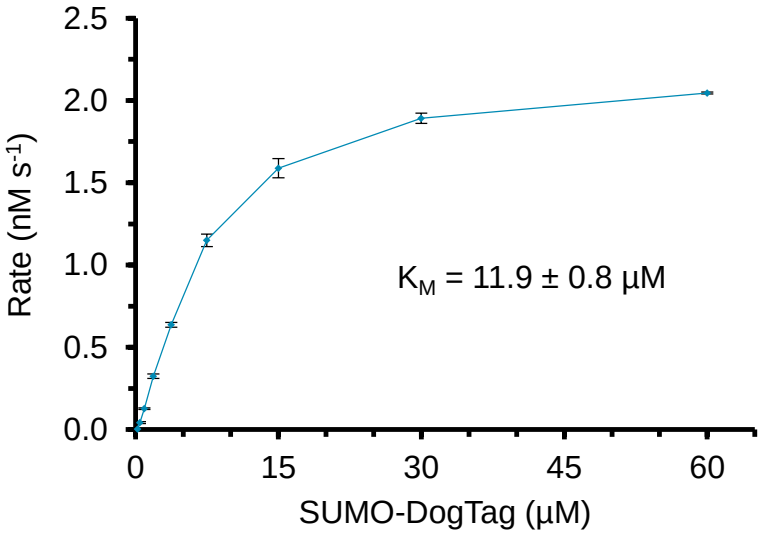


Figure S5. SnoopLigase Michaelis constants. **(a)** Determining SnoopLigase's Michaelis constant for SnoopTagJr-AffiHER2. 40 µM SUMO-DogTag and 5 µM SnoopLigase were incubated with indicated concentrations of SnoopTagJr-AffiHER2 for 7.5 min at 4 °C. The initial rate of formation of SUMO-DogTag:SnoopTagJr-AffiHER2 conjugate was plotted. **(b)** Determining SnoopLigase's Michaelis constant for SUMO-DogTag. 40 µM SnoopTagJr-AffiHER2 and 5 µM SnoopLigase were incubated with indicated concentrations of SUMO-DogTag for 7.5 min at 4 °C and initial rate of conjugate formation was plotted. Analysis by SDS-PAGE with Coomassie staining. Results are mean of triplicate ± 1 SD.

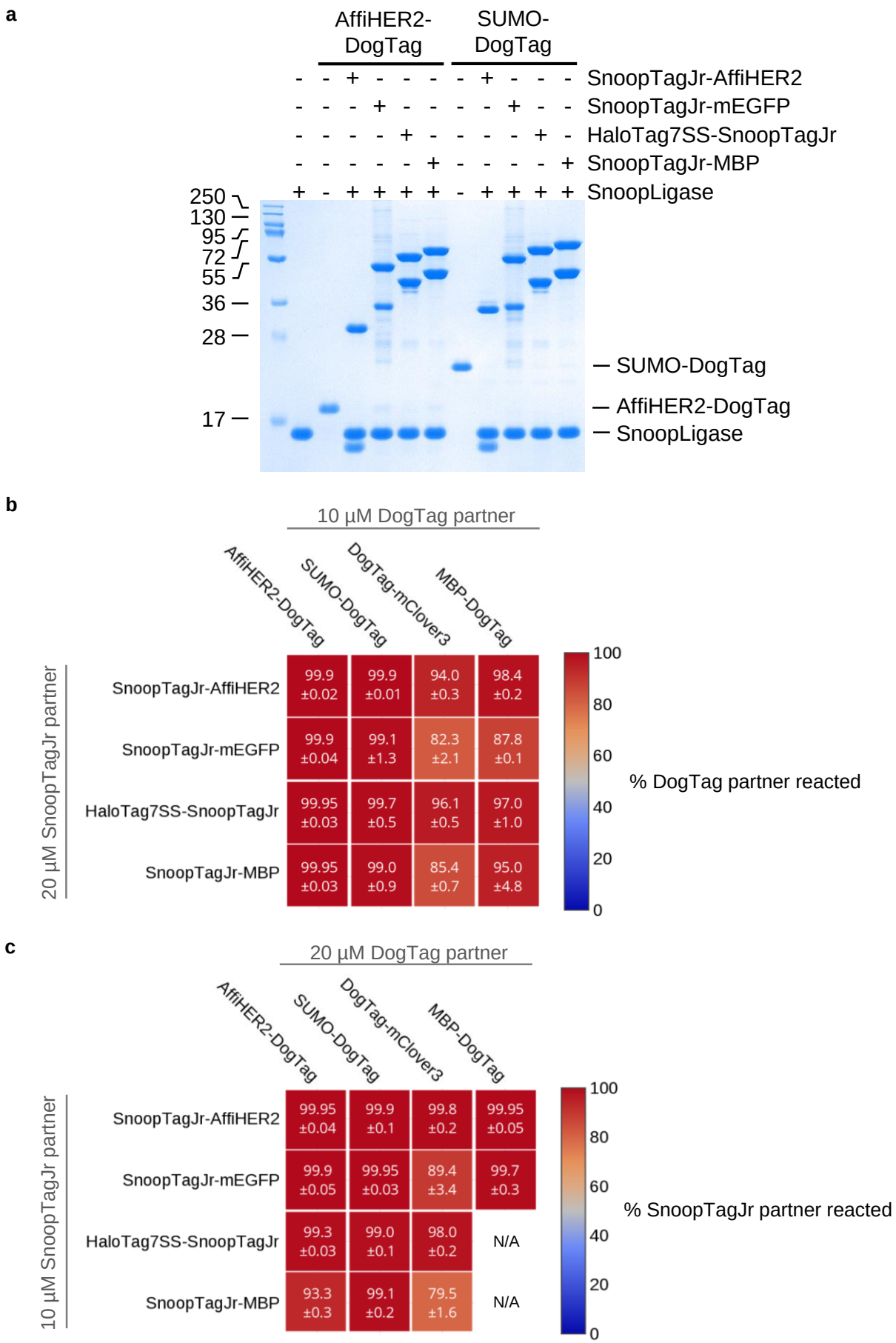


Figure S6. DogTag or SnoopTagJr could be conjugated to high yield at the N- or C-terminus. (a) Representative gel testing high yield conjugation. 10 µM AffiHER2-DogTag or SUMO-DogTag was incubated with 20 µM SnoopLigase and 20 µM SnoopTagJr-linked proteins for 24 h at 4 °C. **(b)** 10 µM DogTag-linked protein was reacted with 20 µM SnoopLigase and 20 µM SnoopTagJr-linked protein for 24 h at 4 °C. % DogTag partner reacted is shown in each box (mean ± 1 SD, n = 3), along with color-coding. **(c)** 10 µM SnoopTagJr-linked protein was reacted with 20 µM SnoopLigase and 20 µM DogTag-linked protein for 24 h at 4 °C. % SnoopTagJr partner reacted is shown in each box (mean ± 1 SD, n = 3), with color-coding. Analysis by SDS-PAGE + Coomassie staining. N/A where band overlap prevented quantification.

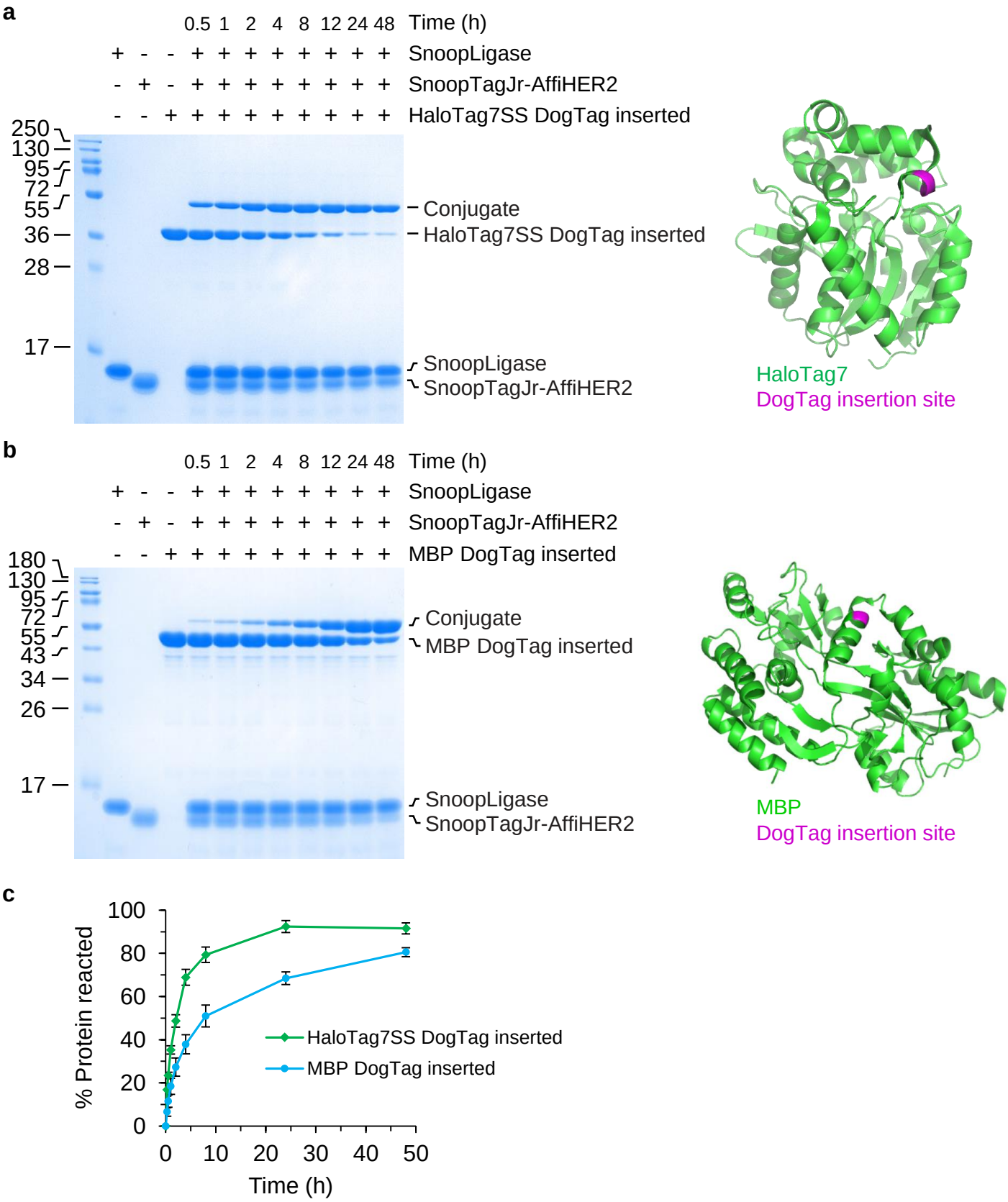


Figure S7. SnoopTagJr reacted with DogTag inserted internally into proteins. (a) 10 µM HaloTag7SS with DogTag inserted between residues D139 and E140 was incubated with 20 µM SnoopLigase and 20 µM SnoopTagJr-AffiHER2 for 0.5 – 48 h at 4 °C. The crystal structure of HaloTag7 (PDB code 5Y2Y) indicates the site for DogTag insertion in magenta. **(b)** 10 µM MBP with DogTag inserted between residues R317 and A319 (I318 deleted) was incubated with 20 µM SnoopLigase and 20 µM SnoopTagJr-AffiHER2 for 0.5 – 48 h at 4 °C. The crystal structure of MBP (PDB code 1OMP) indicates the site for DogTag insertion in magenta. Analysis by SDS-PAGE with Coomassie staining. **(c)** Analysis of reaction extent in (a) and (b) by densitometry. Results are mean ± 1 SD, n = 3.

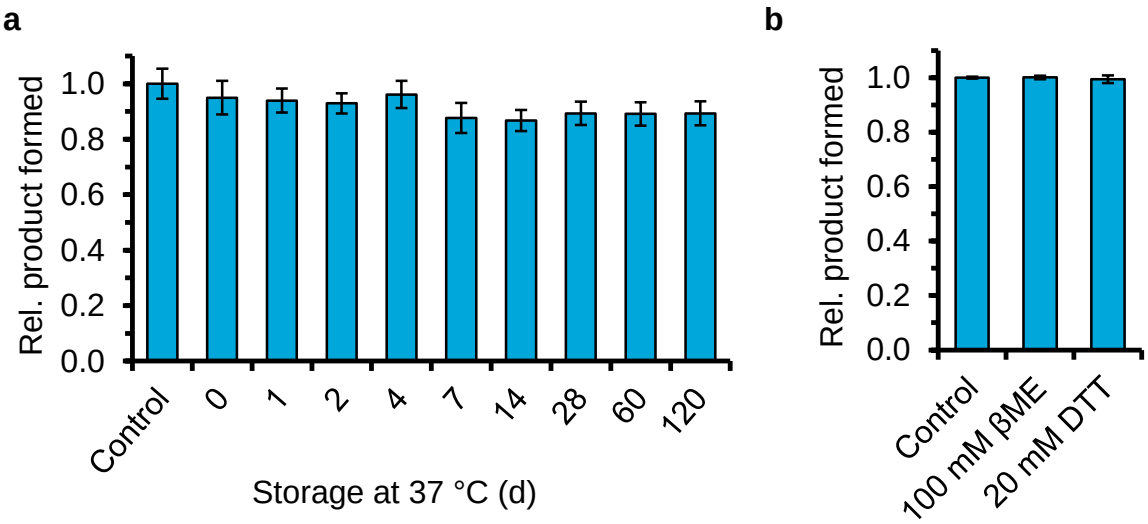


Figure S8. SnoopLigase reactivity was tolerant to lyophilization and reducing agent. (a) SnoopLigase was stable after lyophilization. SnoopLigase was lyophilized and stored for the indicated number of days at 37 °C. SnoopLigase was then reconstituted in reaction buffer with SnoopTagJr-AffiHER2 and SUMO-DogTag (10 µM each) for 2 h at 4 °C in TB pH 7.25 with 15% (v/v) glycerol. Product formation is shown relative to the non-lyophilized control. **(b)** SnoopLigase reaction was unaffected by reducing agent. AffiHER2-DogTag and SUMO-DogTag were conjugated using SnoopLigase (10 µM each) for 2 h at 4 °C with or without the reducing agent β–mercaptoethanol (βME) or dithiothreitol (DTT). Product formation is compared relative to the control without reducing agent. Analysis by SDS-PAGE with Coomassie staining. Results are mean of triplicate ± 1 SD.

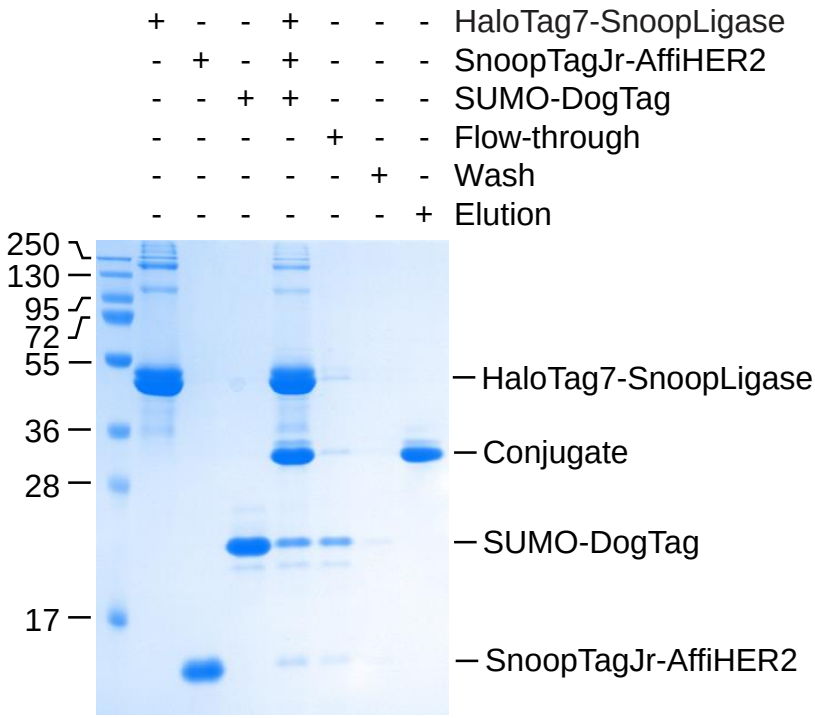


Figure S9. Purification of SnoopLigase reaction product by imidazole elution. SnoopTagJr-AffiHER2 and SUMO-DogTag were ligated using HaloTag7-SnoopLigase (15 µM each) for 24 h at 4 °C. SnoopLigase was captured with HaloLink resin, resin was washed, and the product was eluted with 2 M imidazole. Analysis was performed by SDS-PAGE with Coomassie staining.

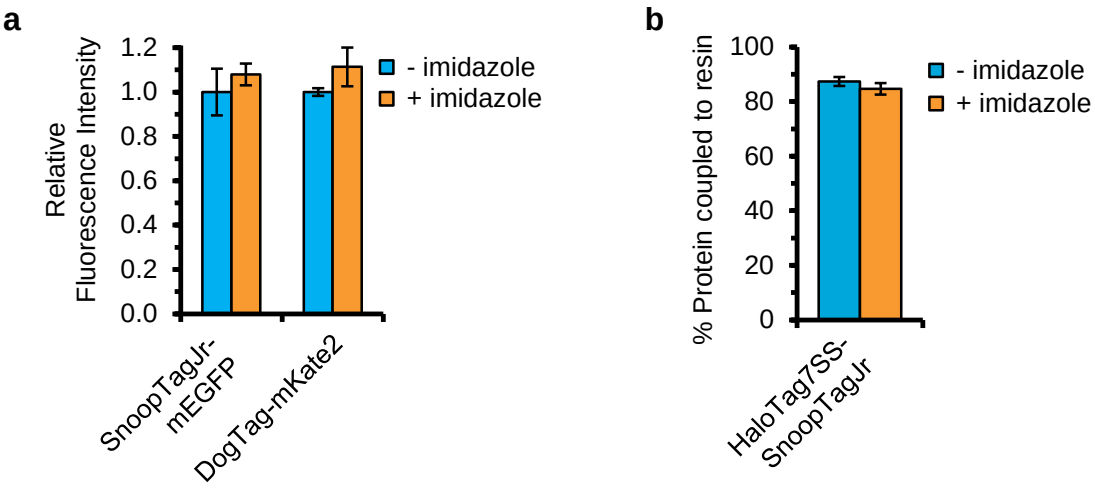


Figure S10. Imidazole exposure had minimal effect on protein activity. SnoopTagJr-mEGFP, DogTag-mKate2 or HaloTag7SS-SnoopTagJr (25 µM) were incubated in buffer with or without 2 M imidazole pH 7.0 for 10 min at 25 °C, followed by dialysis to remove the imidazole. **(a)** Imidazole exposure did not affect the brightness of fluorescent proteins. Fluorescence of dialyzed proteins with or without imidazole exposure was recorded at λ_{ex} = 485 nm and λ_{em} = 538 nm for SnoopTagJr-mEGFP. Fluorescence of dialyzed proteins was recorded at λ_{ex} = 544 nm and λ_{em} = 612 nm for DogTag-mKate2. Relative fluorescence is plotted, with the – imidazole value set to 1. **(b)** Imidazole exposure did not affect HaloTag's ligand binding. Dialyzed HaloTag7SS-SnoopTagJr, with or without imidazole exposure, was incubated with HaloLink resin for 5 min at 25 °C, prior to centrifugation and collection of the supernatant containing unbound protein. The HaloTag7SS-SnoopTagJr remaining in the supernatant was quantified by SDS-PAGE with Coomassie staining. Results are mean of triplicate ± 1 SD.

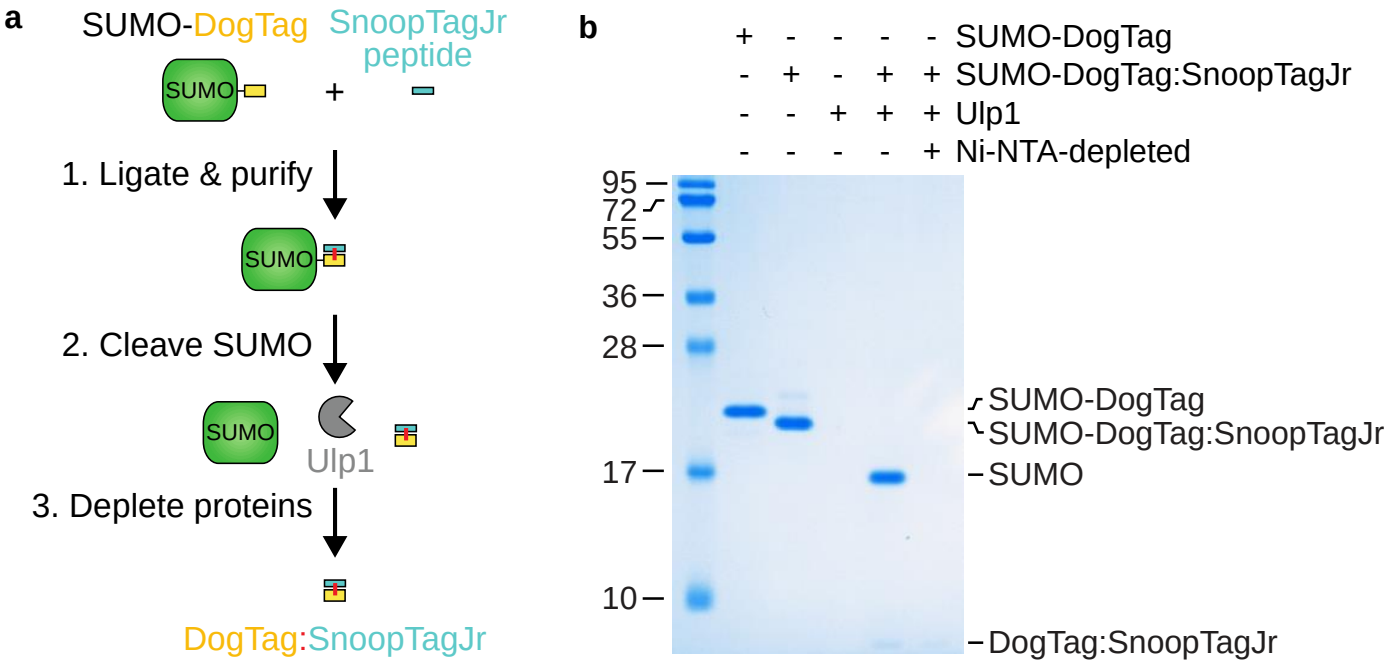


Figure S11. Generation of DogTag:SnoopTagJr competitor. (a) Cartoon of peptide competitor production. SUMO-DogTag and SnoopTagJr peptide are covalently conjugated using HaloTag7-SnoopLigase, followed by purification of the conjugate with imidazole elution. SUMO-protease Ulp1 (gray) cleaves the conjugated SUMO-DogTag:SnoopTagJr. Incubation with Ni-NTA resin depletes the His-tagged SUMO and Ulp1, yielding purified DogTag:SnoopTagJr peptide. **(b)** Production of competitor. HaloTag7-SnoopLigase was coupled to HaloLink resin. The resin was incubated with 50 µM SUMO-DogTag and 75 µM SnoopTagJr, followed by elution of the conjugate using imidazole. The purified conjugate was incubated with Ulp1, followed by Ni-NTA resin to deplete His-tagged proteins. Analysis was performed by SDS-PAGE with Coomassie staining. A faint band corresponding to DogTag:SnoopTagJr is seen, although peptide is not reliably observed by fixation and Coomassie staining and was more robustly quantified by its UV absorbance.

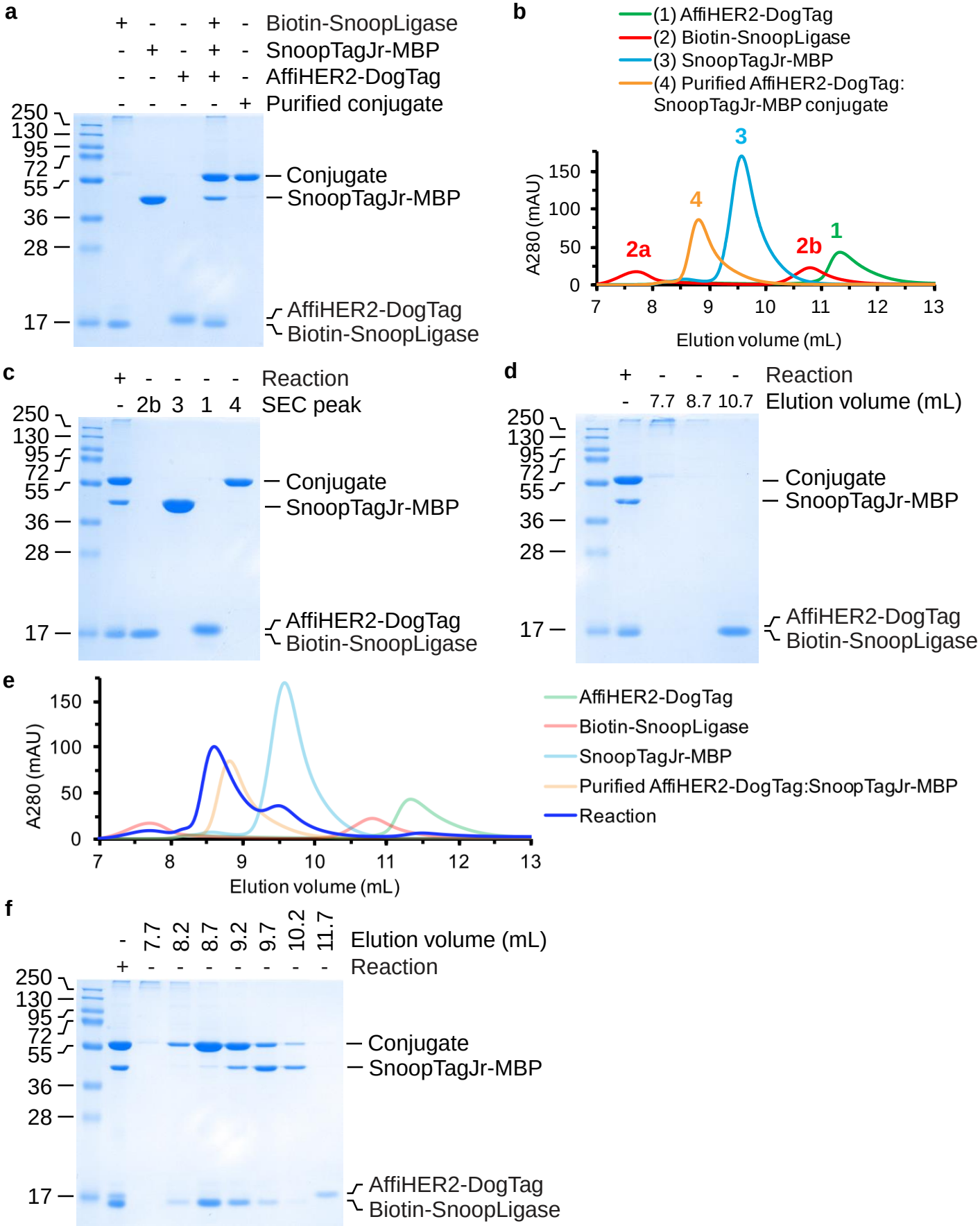


Figure S12. Size exclusion chromatography (SEC) of SnoopLigase reaction. **(a)** SnoopLigase reaction and conjugate purification. SnoopTagJr-MBP and AffiHER2-DogTag were reacted with biotin-SnoopLigase (each 50 μM) for 24 h at 4 °C. The conjugate was purified using imidazole elution. **(b)** SEC of individual proteins from (a). **(c)** SDS-PAGE of the conjugation reaction or SEC peaks from (b). **(d)** SDS-PAGE of the conjugation reaction or peaks from running biotin-SnoopLigase on SEC from (b). Peak 2a contains high molecular weight impurities, while peak 2b represents biotin-SnoopLigase. **(e)** SEC of the reaction sample from (a). Traces from (b) are shown in faint colors for comparison. **(f)** SDS-PAGE of fractions from SEC separation of the reaction in (e). Biotin-SnoopLigase was exclusively found in fractions containing AffiHER2-DogTag:SnoopTagJr-MBP, indicating interaction with the conjugate.

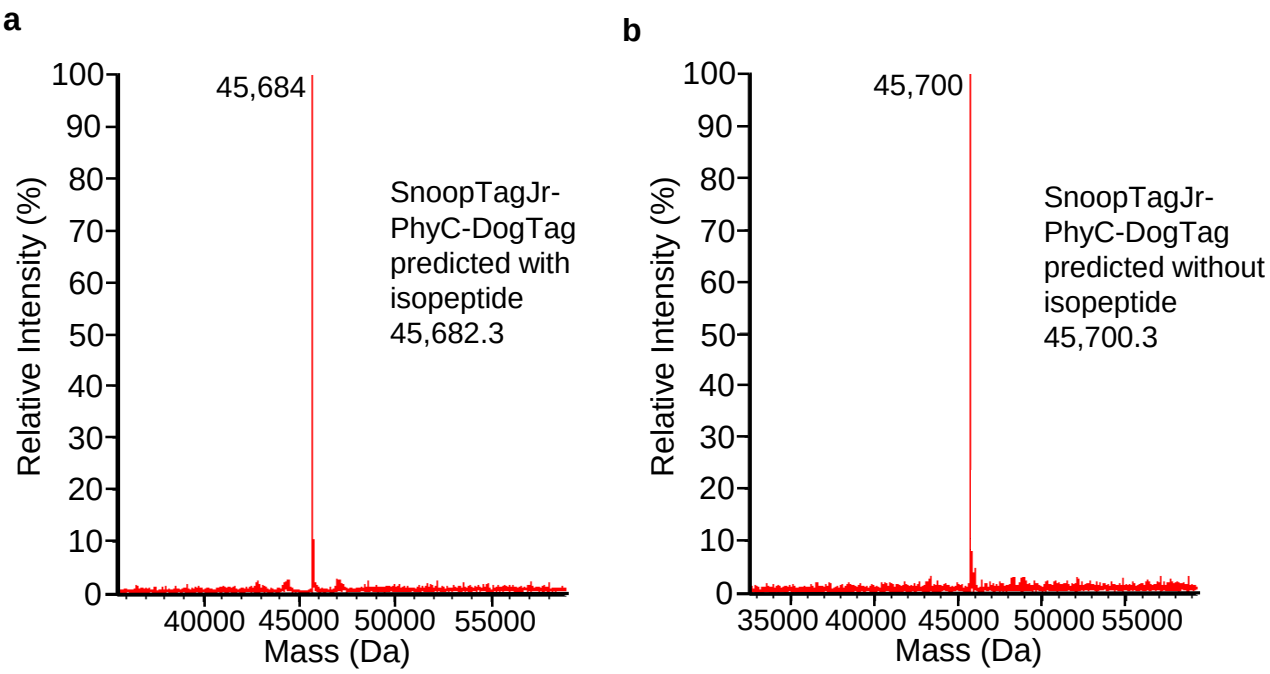


Figure S13. Mass spectrometry analysis of isopeptide bond formation of SnoopTagJr-PhyC-DogTag. Electrospray ionization mass spectrometry of **(a)** SnoopLigase-cyclized SnoopTagJr-PhyC-DogTag following SnoopLigase removal and **(b)** SnoopTagJr-PhyC-DogTag not exposed to SnoopLigase. The principal peak is marked and annotated with the mass predicted by the ProtParam online tool.

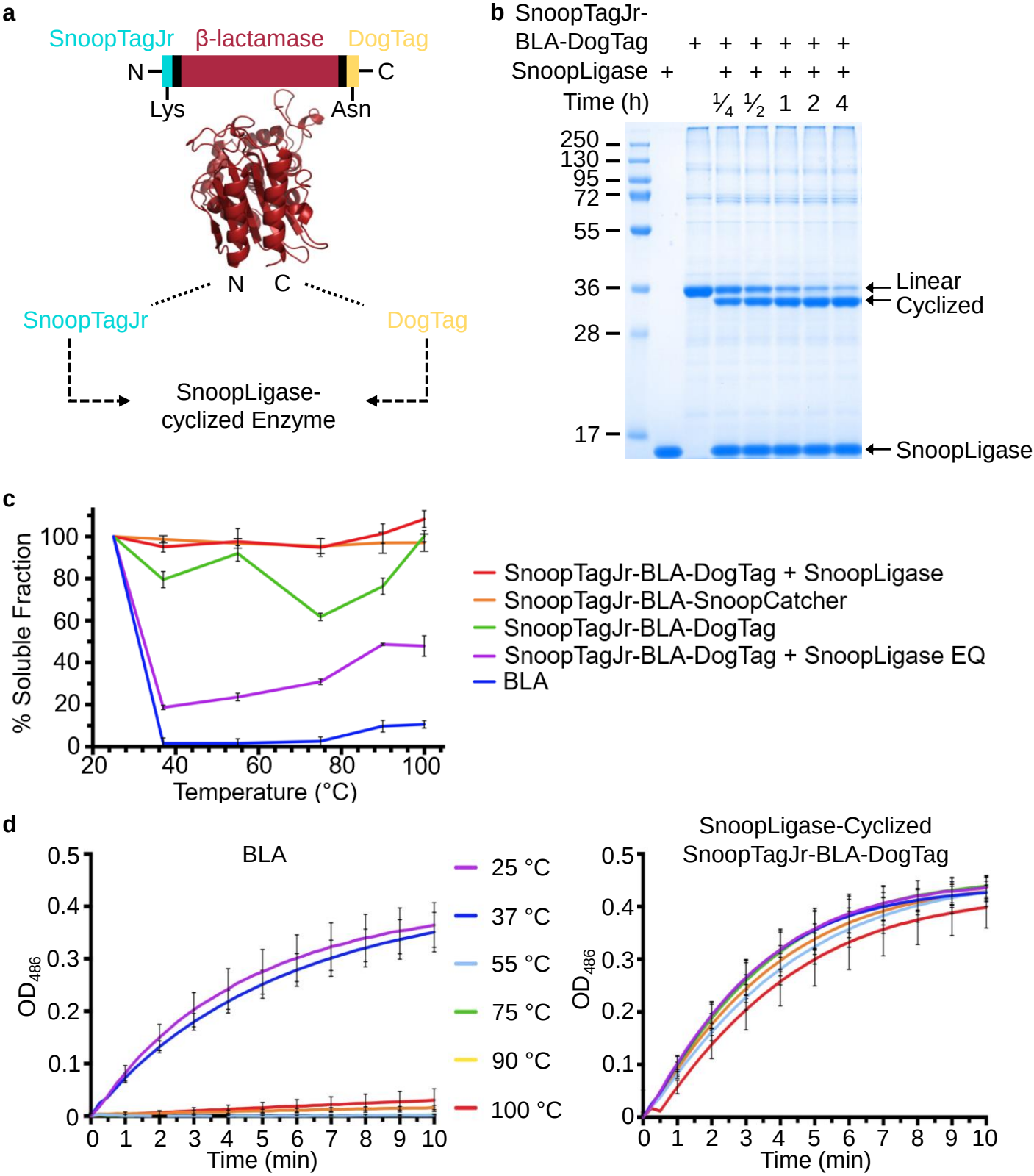


Figure S14. SnoopLigase cyclization enhancing BLA thermal resilience. **(a)** Schematic of SnoopLigase cyclization of BLA, based on PDB code 1BTL. **(b)** SnoopLigase efficiently cyclized BLA. SnoopLigase and SnoopTagJr-BLA-DogTag (10 μM each) were incubated at 4 °C for 0.25-4 h, before SDS-PAGE with Coomassie staining. **(c)** Effect of tags and inactive SnoopLigase on heat-induced aggregation of BLA. BLA or variants were incubated for 10 min at indicated temperatures. Samples were centrifuged to remove aggregates and soluble protein in the supernatant was quantified. **(d)** SnoopLigase cyclization improved the catalytic activity of BLA after heating. BLA or SnoopLigase-cyclized SnoopTagJr-BLA-DogTag were incubated for 10 min at the indicated temperatures and cooled to 25 °C. Then nitrocefin hydrolysis was monitored colorimetrically over time at 20 °C. All are mean of triplicate ± 1 SD; some error bars are too small to be visible.

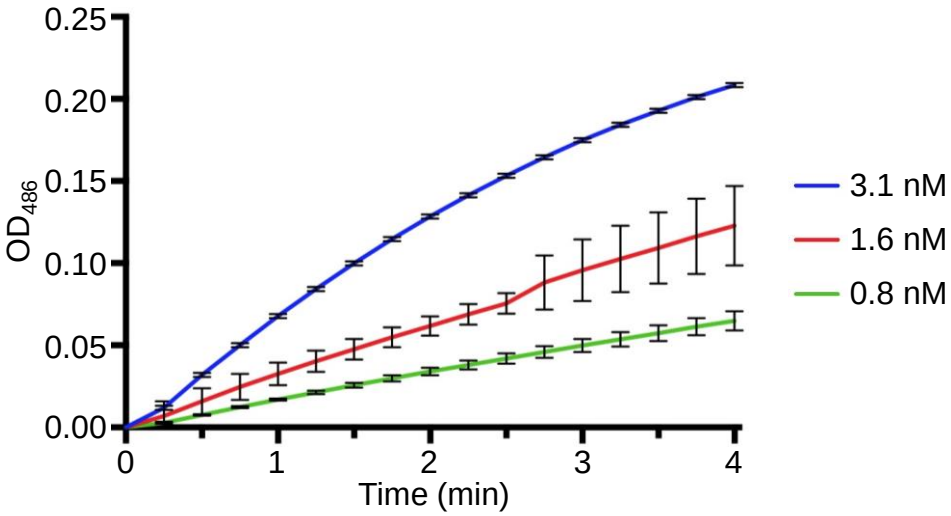


Figure S15. The BLA activity assay correlated to enzyme concentration. To validate that the amount of BLA enzyme was limiting in our activity assays, so that the assay reported on functional enzyme, wild type BLA at the indicated concentrations was reacted with 95 μM nitrocefin. Nitrocefin hydrolysis was monitored colorimetrically over time at 20 °C. Results are mean of triplicate ± 1 SD.

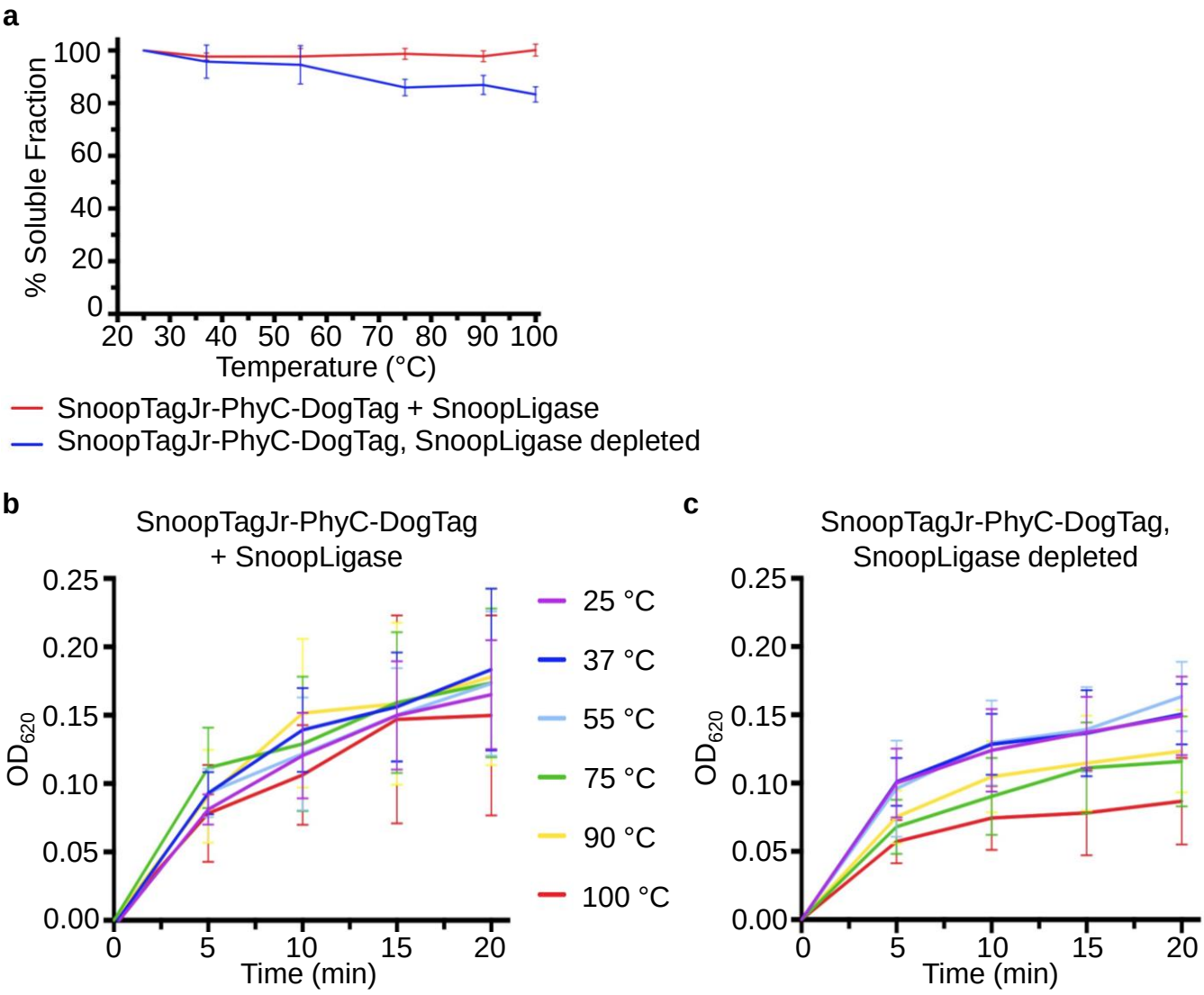


Figure S16. Effect of SnoopLigase removal on resilience of cyclized PhyC. (a) SnoopLigase-cyclized PhyC retained solubility following SnoopLigase removal. Cyclized PhyC before and after SnoopLigase removal was incubated for 10 min at the indicated temperature. Samples were centrifuged to remove aggregates and soluble protein in the supernatant was quantified. Results are mean ± 1 SD, n=6. **(b)** Cyclized PhyC without SnoopLigase removal was incubated at the indicated temperature for 10 min, cooled to 10 °C and used in colorimetric assays to quantify phosphate release from the hydrolysis of phytic acid. Results are mean ± 1 SD, n=9. **(d)** SnoopLigase removal compromised the thermal resilience of enzymatic activity of cyclized PhyC. Cyclized PhyC following SnoopLigase removal by imidazole was incubated at the indicated temperature for 10 min, cooled to 10 °C and used in colorimetric assays to quantify phosphate release from the hydrolysis of phytic acid. Results are mean ± 1 SD, n=9.

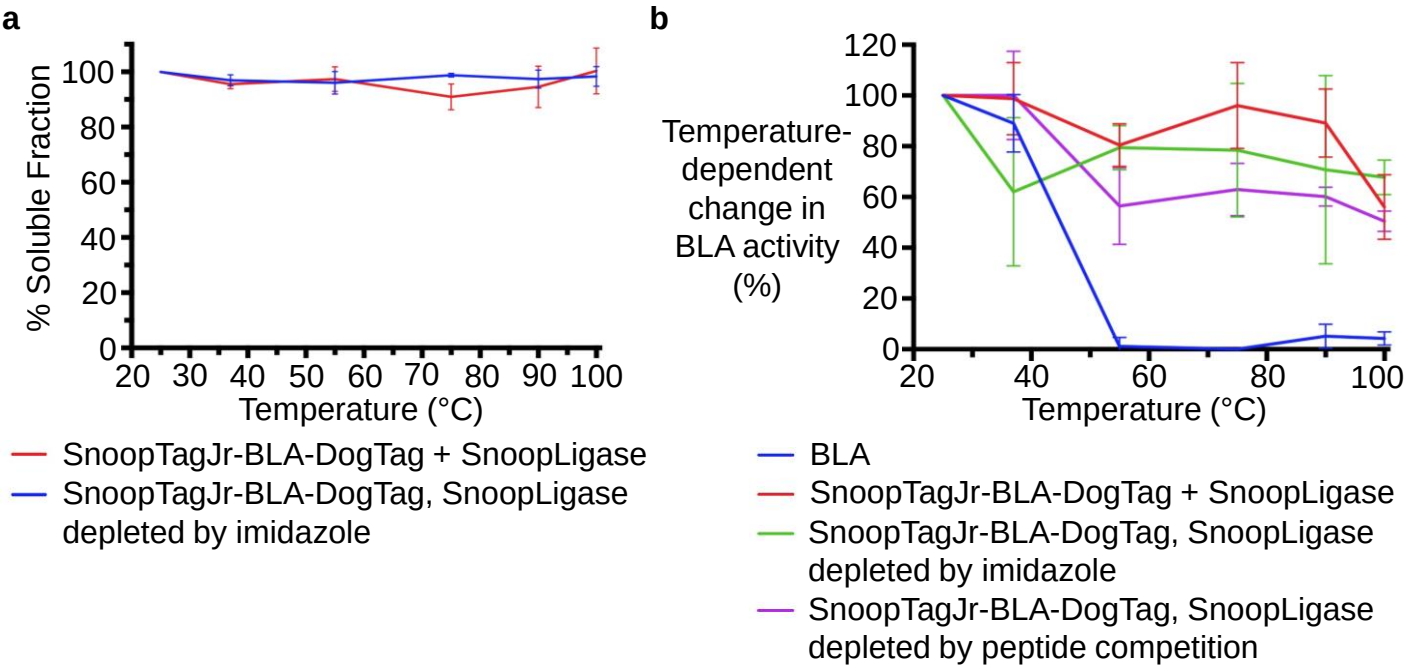


Figure S17. Effect of SnoopLigase removal on solubility and activity of cyclized BLA. (a) SnoopLigase-cyclized BLA remained soluble following SnoopLigase removal, even after boiling for 10 min. Cyclized BLA before and after SnoopLigase removal was incubated for 10 min at the indicated temperature. Samples were centrifuged to remove aggregates and soluble protein in the supernatant was quantified. Results are mean ± 1 SD, n=6. **(b)** BLA constructs were incubated at the indicated temperatures for 10 min, cooled and used in a colorimetric assay to determine the initial rate for cleavage of nitrocefin. Constructs incubated at 25 °C were defined as 100% active. Results are mean of triplicate ± 1 SD.