One-Pot Dual Labeling of IgG 1 and Preparation of C-to-C Fusion Proteins Through a Combination of Sortase A and Butelase 1

Thibault J. Harmand,† Djenet Bousbaine,†,‡ Alix Chan,§ Xiaohong Zhang,‖ David R. Liu,§ James P. Tam,‖ and Hidde L. Ploegh*†

†Program in Cellular and Molecular Medicine, Boston Children’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, United States
‡Department of Microbiology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
§The Broad Institute of Harvard and MIT, Howard Hughes Medical Institute, and the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02142, United States
‖School of Biological Sciences, Nanyang Technological University, 637551, Singapore

ABSTRACT: Site-specific chemical modification of proteins can assist in the study of their function. Furthermore, these methods are essential to develop biologicals for diagnostic and therapeutic use. Standard protein engineering protocols and recombinant expression enable the production of proteins with short peptide tags recognized by enzymes capable of site-specific modification. We report here the application of two enzymes of orthogonal specificity, sortase A and butelase 1, to prepare non-natural C-to-C fusion proteins. Using these enzymes, we further demonstrate site-selective installation of different chemical moieties at two sites in a full-size antibody molecule.

Protein modification technology continues to evolve and improve. Protein conjugates prepared through chemical modification enable applications difficult to achieve with conventional protein expression techniques. Such conjugates range from simple imaging agents and pegylated proteins to more complex bioconjugates such as antibody–drug conjugates (ADCs) and unnatural C-to-C and N-to-N fusion proteins.

The most widely used approach relies on cysteine functionalization chemistry. Introduction into a protein sequence of an unpaired cysteine renders the method mostly site-specific, and modification of its native counterpart can yield heterogeneous mixtures. To avoid this issue, site-specific incorporation of an unnatural amino acid into the protein is possible, but this requires manipulation of the genetic code, involves substantial optimization, and therefore adjustments on a case-by-case basis. A further alternative consists of complete synthesis of the polypeptide of interest (POI) by chemical means or the combination of various segments through chemical ligation. Although this approach enables site-specific incorporation of the desired modifications, such technologies are limited by the size of the POI: target POIs are usually limited to 250 AA or less.

We and others have used a peptide ligase, sortase A, for a variety of protein modifications. Sortase A recognizes a short amino acid sequence, LPXTGG, where X is any amino acid except proline. To this sequence, Sortase A will ligate a short amino acid sequence, LPXTGG, where X is any amino acid except proline. To this sequence, Sortase A will ligate a

Supporting Information

Scheme 1. Sortase A and Butelase 1 Ligation Mechanisms and Our Work Described in This Paper

A) Sortase A-mediated ligation

B) Butelase 1-mediated ligation

C) This work: One-pot Sortase-Butelase-mediated ligation

Received: August 8, 2018
Revised: September 6, 2018
Published: September 19, 2018

DOI: 10.1021/acs.bioconjchem.8b00563
Bioconjugate Chem. 2018, 29, 3245–3249
peptide that contains an N-terminal oligo-glycine stretch equipped at its C-terminus with any cargo of choice. The LPXTGG recognition motif is easily incorporated into a POI (Scheme 1A) and is generally placed at or near the C-terminus. Sortase-mediated labeling is robust and finds use in a broad range of applications. Tam and co-workers discovered another ligase, butelase-1, in the seedpods of the plant Clitoria ternatea, purified it, and used it for the modification and cyclization of peptides as well as proteins.19−23 Compared to sortase-A, butelase-1 obeys a distinct set of recognition rules. Butelase-1 also has a much higher turnover number. Butelase 1 recognizes a AsnHisVal (NHV) motif, which it ligates to incoming nucleophiles composed of two amino acids to which a cargo of choice may be attached, provided two rules are followed. The N-terminal residue can be any one except proline, and the N-1 terminal residue must be a Cys, Ile, Leu, or Val (Scheme 1B).

The distinct rules obeyed by these ligases allows a combination of sortase A and butelase 1 in a single reaction to prepare unnatural C-to-C fusions of two different proteins and to label proteins at two distinct sites in a one-pot reaction (Scheme 1C).

VHHS, also called nanobodies, are small proteins (∼14 kDa) that correspond to the antigen binding domain of heavy chain-only camelid antibodies. They are the smallest antibody fragments that retain antigen recognition. We used two such VHHS: VHH7, which recognizes class II MHC products (MHC II), and VHH-Enh, which recognizes eGFP. Generally, the production of bispecific antibodies or their derivatives is a desirable goal from the perspective of therapeutic applications. In this paper, we chose two VHHS of different specificities as a proof-of-concept. These VHHS were modified to carry the desired motifs: VHH7 bears the LPETGG motif at the C-terminus, and VHH-Enh has the NHV motif at the C-terminus. We synthesized a two-headed, PEG-based linker compatible with both sortase A and butelase-1 by standard SPPS (Figure 1A). VHH7 (75 μM) was incubated at 4 °C with the linker (500 μM) and sortase A (2 μM) for 15 h. Analysis by mass spectrometry of the reaction mixture showed no remaining His-tagged VHH7, indicating full conversion for the
first ligation. The His-tagged sortase was removed by depletion on Ni-NTA agarose beads, and remaining free linker was removed from the ligation product through centrifugation-based size exclusion or alternatively by dialysis. To the ligated product (250 μM) were added VHH-Enh (25 μM) and butelase 1 (1.0 μM), and the mixture was incubated at 37 °C for 6 h with agitation. The excess of starting material is necessary in order to reach complete conversion. The reaction mixture was purified by either FPLC or centrifugation-based size exclusion to afford the desired product in an overall yield of ~30%. The different steps of the reaction were monitored by SDS PAGE (Figure 1A) and mass spectrometry. Only using sortase A and a two-headed linker with a sortase nucleophile, i.e., (Gly)₉, would not be compatible with this scheme, as it would have uncontrollably produced a mixture of homodimers and the desired heterodimer.

We next altered the properties of the linker that connects the two ligase recognition motifs. Double-stranded oligonucleotides can serve as a rigid linker to enforce distance constraints not easily achievable by more flexible PEG spacers. The linker design further included an EcoRI recognition motif to enable enzymatic cleavage of the protein-DNA-protein adduct and release of the protein monomers. Complementary single-strand oligonucleotides were synthesized on a solid support. Fmoc-Gly-Gly-Gly-OH was attached to one strand, and Fmoc-Gly-Val-OH was attached to the other. Cleavage of the resin and purification using standard purification protocols afforded the two desired strands in excellent purity. Simple annealing yielded the linker containing the DNA duplex (Figure 1B).

We followed the protocol as described (Figure 1A) with the modification that the unreacted DNA linker was removed using a dialysis cassette with a 20 kDa cutoff. After adding butelase 1 and VHH-Enh to linker-conjugated VHH7 and incubating the reaction mixture at 37 °C for 6 h, the solution was dialyzed using a membrane with a 30 kDa cutoff or purified by FPLC. The desired product as observed by SDS-PAGE was obtained in an overall yield of ~25% (Figure 1B) with no detectable oligonucleotide remaining.

We demonstrated cleavage of this asymmetric dimer by incubation with EcoRI. After 1.5 h at 37 °C, the reaction mixture was analyzed by SDS-PAGE. We observed a single polypeptide of the expected molecular weight. This result establishes that the desired C-to-C fusion protein had been obtained and that it was possible to cleave the protein-bound linker using a restriction enzyme (Figure 1B). Obviously, length and sequence of the DNA duplex linker can be varied at will to enable the use of different restriction enzymes and to arrive at various distance constraints between the entities newly connected.

We next explored the possibility of site-specific labeling of a single multisubunit protein at two positions in a one-pot reaction. Because both sortase A and butelase 1 target sequence motifs at or near the C-terminus of the POI, dual labeling of a single polypeptide would require a different approach. Immunoglobulins are composed of two identical heavy chains and two identical light chains held together by disulfide bonds and by noncovalent interactions and are thus an ideal target for our method. Although there are examples in the literature of dually modified antibodies, the methods used are all based on maleimide and succinimidyl ester chemistry, leading to uncontrolled modification unless an unpaired cysteine residue is engineered into the antibody. We modified an IgG1 molecule so that it contained an LPETGG motif at the C-terminus of the κ (light) chain and an NHV motif at the C-terminus of the γ1 (heavy) chain. We designed two different probes, an oligo-glycine peptide bearing 5,6-carboxyfluorescein (5,6-FAM) at its C-terminus and an alanine-leucine peptide bearing the AlexaFluor 647 fluorescent dye at its C-terminus. IgG1 (10 μM) was incubated with the GGG-FAM (500 μM), the AL-Alexa (500 μM), sortase A (2 μM), and butelase 1 (0.8 μM). The ligation mixture was incubated first at 4 °C for 15 h followed by incubation at 37 °C for 4 h. Simple centrifugation-based size exclusion was sufficient to remove unincorporated dyes to obtain pure dually modified antibody.

Analysis by SDS-PAGE and examination of the gel using the appropriate excitation and fluorescence emission channels showed that the two chains were each selectively modified with the desired probe.

We thus successfully and selectively labeled full-size IgG1 in a one-pot reaction (Figure 2). Absorption spectroscopy showed an average of 1.92 mol of FAM per mole of protein, corresponding to a 96% conversion for sortase-mediated light chain labeling and an average of 1.97 mol of Alexa647 per mole of protein, corresponding to 98% labeling of the heavy chain using butelase 1.

In summary, the use of two orthogonal enzymes, sortase A and butelase 1, allows straightforward preparation of C-to-C fusion proteins using either PEG- or oligonucleotide-based two-headed linkers. We further describe two-site modification of an IgG1 protein in a one-pot reaction. The small size of the two enzyme recognition motifs, LPETGG and NHV, in addition to the simple steps required to incorporate them into any desired POI sequence, makes this methodology convenient for both chemists and biologists. The use of linkers of different lengths should allow enforcement of inter- and intramolecular distance constraints that affect the molecular properties of proteins thus modified. When using DNA as the linker, any such constraints may be relieved by simple enzymatic cleavage using restriction enzymes that obviously do not target the protein moieties themselves. Finally, the possibility to selectively attach two distinct molecules to a full-size
monoclonal antibody in a one-pot reaction holds promise for the development of complex antibody–drug conjugates.

**ASSOCIATED CONTENT**

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.bioconjchem.8b00563.

Detailed materials and methods and supplementary figures and tables (PDF)

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: hidde.ploegh@childrens.harvard.edu.

ORCID

Thibault J. Harmand: 0000-0002-7582-0519
Alix Chan: 0000-0002-2698-4031
James P. Tam: 0000-0003-4433-198X
Hidde L. Ploegh: 0000-0002-1090-6071

Notes
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

T.J.H. is the recipient of an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (SNSF). A part of this work (D.R.L and A.C) was supported by the NIH grant R35 GM118062 and HHMI. We are also thankful for grant to-N and C-to-C protein fusions. Proc. Natl. Acad. Sci. U. S. A. 109, 11993−11998.

**REFERENCES**


