Bioorganic & Medicinal Chemistry 24 (2016) 5816–5822

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Direct protein–protein conjugation by genetically introducing bioorthogonal functional groups into proteins

Sanggil Kim [a], Wooseok Ko [a], Bong Hyun Sung [b], Sun Chang Kim [c], Hyun Soo Lee [a,*]

[a] Department of Chemistry, Sogang University, Seoul 121-742, Republic of Korea
[b] Bioenergy and Biochemical Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 305-806, Republic of Korea
[c] Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Republic of Korea

ARTICLE INFO

Article history:
Received 11 August 2016
Revised 13 September 2016
Accepted 14 September 2016
Available online 15 September 2016

Keywords:
Genetic incorporation
Unnatural amino acids
Protein–protein conjugation
Click reaction

ABSTRACT

Proteins often function as complex structures in conjunction with other proteins. Because these complex structures are essential for sophisticated functions, developing protein–protein conjugates has gained research interest. In this study, site-specific protein–protein conjugation was performed by genetically incorporating an azide-containing amino acid into one protein and a bicyclononyne (BCN)-containing amino acid into the other. Three to four sites in each of the proteins were tested for conjugation efficiency, and three combinations showed excellent conjugation efficiency. The genetic incorporation of unnatural amino acids (UAAs) is technically simple and produces the mutant protein in high yield. In addition, the conjugation reaction can be conducted by simple mixing, and does not require additional reagents or linker molecules. Therefore, this method may prove very useful for generating protein–protein conjugates and protein complexes of biochemical significance.

1. Introduction

Proteins are essential biomolecules that perform complex and sophisticated biological functions. Although protein monomers can function independently, proteins often function by forming complex structures with other proteins. Because these complexes make it possible for proteins to conduct more sophisticated functions, many efforts have been made to construct protein–protein conjugates. A straightforward method to achieve this is the genetic fusion of two or more proteins, in which the DNA sequences are joined in tandem to form a single polypeptide containing multiple proteins. While genetic fusion is technically simple in design and preparation, the method has limitations in that proteins can be linked only by their C- or N-terminals and fusion proteins often have problems in protein folding. A recent report demonstrated N-to-N or C-to-C protein fusion; however, its application is still limited to the terminal conjugation of proteins.[1]

Over the last decade, methods for protein modification have advanced significantly. These methods include chemical modification using the chemistries of cysteine and lysine,[2,3] genetic incorporation of unnatural amino acids (UAAs) containing a bioorthogonal functional group,[4–11] enzymatic preparation of an aldehyde tag[12,13] or protein conjugates,[14,15] and intein-mediated protein ligation.[16] Recently, these methods have been used to achieve protein–protein conjugation, although the number of examples is limited. Protein–protein conjugates have been formed by reacting lysine residues in proteins with a reactive ester in a linker to form amide bonds. In these reactions, the specificity of lysine residues for amide bond formation is achieved by bringing the reactive ester near the lysine residues using protein–protein interactions or nucleic acid templates.[17,18] In another example, a ketone-containing UAA was genetically incorporated into antibodies, and the mutant antibodies were linked to a second antibody molecule or a toxin by reaction with a linker containing an alkoxyamine group and another orthogonal functional group (maleimide, azide, or alkyne) at each end.[19,20] This method achieved the formation of antibody-toxin conjugates and bispecific antibodies.[19,20] An aldehyde tag was introduced into IgG by a formylglycine-generating enzyme, and the aldehyde-containing IgG was conjugated to human growth hormone using a linker containing an alkoxyamine group and another orthogonal functional group (azide, or alkyne) at each end.[13] Although these methods achieved site-specific protein–protein conjugation, they used a linker that required a two-step conjugation process.[13] After the first conjugation reaction, excess reagents must be removed by dialysis or affinity purification. In addition, the conjugated products resulting from these methods might have been useful for a specific application; however, protein complexes in nature are formed by direct protein–protein interactions without a linker. Although direct

* Corresponding author. Tel.: +82 2 705 7958; fax: +82 2 705 7893.
E-mail address: hslee76@sogang.ac.kr (H.S. Lee).

http://dx.doi.org/10.1016/j.bmc.2016.09.035
0968-0896/

protein–protein conjugation has been achieved by copper (I)-catalyzed click reactions, proteins containing bioorthogonal groups were expressed in a cell-free system, which significantly reduces protein yield, limiting the application of the method.[21]

Herein, we report a direct protein–protein conjugation method, in which two mutant proteins containing azide or alkyne groups were conjugated by simple mixing. The UAAs containing reactive functional groups can be genetically incorporated into proteins, and the conjugation reaction requires only a single step and no additional reagents. Therefore, this method can be useful for constructing protein–protein conjugates and other complex protein structures.

2. Results and discussion

2.1. Expression and purification of mutant proteins containing a bioorthogonal functional group

In order to achieve direct protein–protein conjugation, a strain-promoted azide-alkyne cycloaddition (SPAAC) reaction was chosen as a bioorthogonal reaction (Fig. 1). Specifically, *p*-azido-L-phenylalanine (AZF) and N^6-[(1*R*,8*S*,9*R*)-bicyclo[6.1.0]non-4-yn-9-ylmethoxy]carbonyl-L-lysine (BCNK) were chosen as UAAs to be genetically incorporated for the protein–protein conjugation (Fig. 2).[5,11] AZF is one of the most extensively used UAAs for site-specific protein conjugation because this UAA has excellent incorporation efficiency and it is readily available from commercial sources.[5,22,23] The other component, BCNK, is a good partner for AZF in the SPAAC reaction.

To test whether AZF and BCNK could be used as a bioorthogonal pair for protein–protein conjugation, these two UAAs were incorporated into a glutathione S-transferase (GST) and a maltose-binding protein (MBP) respectively. Based on the X-ray crystal structures of the proteins,[24,25] several residues on the surfaces of the proteins were chosen for UAA incorporation: F46, K87, K113, and H139 of GST, and K29, K83, and Y167 of MBP. The codons for each residue were mutated to the amber stop codon (TAG) for the incorporation of the UAAs, and the proteins were expressed in the presence of the corresponding evolved tRNA/aminoacyl-tRNA synthetase (aaRS) pairs and UAAs in *Escherichia coli* strain DH10B.[26] The evolved tRNA/AZFRS pair[5] from *Methanococcus jannaschii* and AZF were used for GST expression, and the evolved tRNA/BCNKRS pair[27] from *Methanosarcina mazei* and BCNK were used for MBP expression. To examine the expression of the mutant proteins containing AZF or BCNK, the whole cell lysates from each expression system were analyzed by SDS–PAGE (Fig. 3A). For all mutants, full-length proteins were expressed in the presence of UAAs, and the expression levels were observed to be slightly different depending on the UAA position. The GST mutants containing AZF (GST-AZF) and the MBP mutants containing BCNK (MBP-BCNK) were purified by affinity chromatography (Fig. 3B), and the yields ranged from 6.0 mg/L to 10.5 mg/L for GST and from 3.0 mg/L to 4.5 mg/L for MBP. Matrix-assisted laser desorption/ionization-time of flight mass spectrometric (MALDI-TOF MS) analyses of selected GST and MBP mutants after trypsin digestion showed the incorporation of AZF or BCNK without the incorporation of any natural amino acids (Fig. 3C).

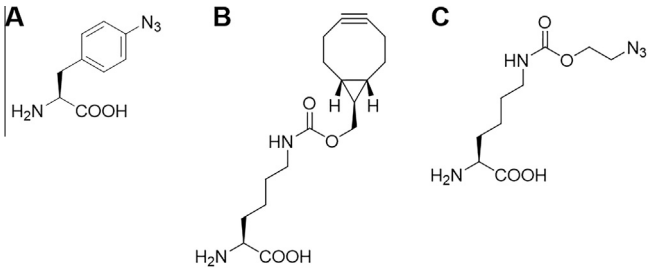

Figure 2. Structures of UAAs used in this study. (A) *p*-Azido-L-phenylalanine (AZF). (B) N^6-[(1*R*,8*S*,9*R*)-bicyclo[6.1.0]non-4-yn-9-ylmethoxy]carbonyl-L-lysine (BCNK). (C) N^6-[(2-azidoethoxy)carbonyl]-L-lysine (AZK).

2.2. Evaluation of mutant proteins for protein–protein conjugation

The protein–protein conjugation of purified mutant proteins was then evaluated. Each MBP-BCNK was reacted with four GST-AZF mutants at 37 °C, and the conjugation reactions were monitored. To increase conjugation efficiency, MBP-BCNK mutants were used in fivefold excess over GST-AZF mutants. Analytical samples were taken from each reaction mixture at various times, and the reaction was quenched with excess AZF, followed by SDS–PAGE analysis. All reactions showed GST-MBP conjugates after 48 hours, although their conjugation efficiencies varied (Fig. 4).

For the conjugation reactions with GST-F46AZF/MBP-K29BCNK, GST-F46AZF/MBP-K83BCNK and GST-K113AZF/MBP-K83BCNK, more than 80% of GST-AZF mutants were conjugated with MBP-BCNK, and the reaction with GST-F46AZF/MBP-K83BCNK showed the most efficient conjugation. Control reactions with wild-type GST (GST-WT) or wild-type MBP (MBP-WT) did not produce any conjugated product, indicating that the GST-MBP conjugates were generated by the reaction of the genetically incorporated UAAs (Fig. 5A). These results indicate that the incorporation of AZF and BCNK into two proteins allowed site-specific protein–protein conjugation. The conjugated product from the conjugation reaction of GST-F46AZF and MBP-K83BCNK was isolated by affinity purification showing that the conjugate was able to bind to the corresponding affinity resins and could be readily isolated (Fig. 5B).

2.3. Relationship between conjugation efficiency and UAA incorporation sites

The residues that were replaced by either one of the UAAs, AZF or BCNK, were chosen, based on their location at the protein sur-

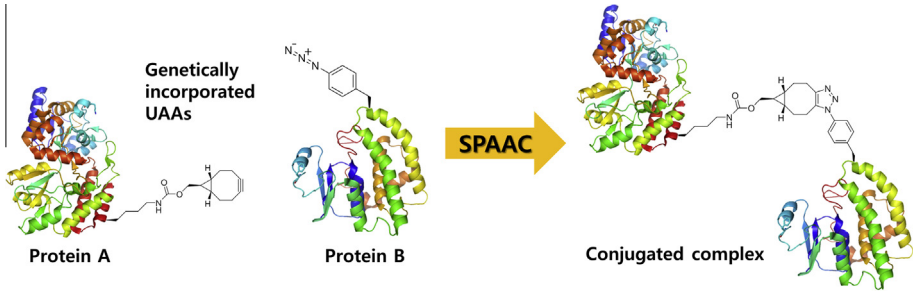

Figure 1. Schematic representation of direct protein–protein conjugation using genetic incorporation of UAAs and SPAAC.