Tetrahedron Letters 56 (2015) 2123–2132

Contents lists available at ScienceDirect

Tetrahedron Letters

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

Digest Paper

Recent advances in bioorthogonal reactions for site-specific protein labeling and engineering

Yukang Gong, Lifeng Pan *

State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Lingling Road 345, Shanghai 200032, People's Republic of China

ARTICLE INFO

Article history:
Received 15 November 2014
Revised 11 March 2015
Accepted 12 March 2015
Available online 19 March 2015

Keywords:
Bioorthogonal organic chemistry
Ligand-directed organic chemistry
Protein labeling and engineering

ABSTRACT

In the past two decades, with the rapid development of chemical biology, tremendous small-molecule based toolkits were created by organic chemists, and were widely used to study and manipulate proteins in order to dissect their complicated biological functions. This review summarizes some recent progresses of bioorthogonal reactions for site-specific protein labeling and engineering, and highlights the powers of using these methods to study the biological functions of some proteins.

Contents

Introduction

As one of the most abundant biomolecules, proteins are involved in most of the biological processes and perform a wide array of important functions within living organisms. Therefore, the study and manipulation of protein functions are not only of significant importance to fundamental scientific research, but also critical to the development of biomedical and biotechnological applications. Traditional genetics, molecular biology, biochemistry, cell biology, and allied methods have provided various tools to investigate the functions of proteins, and have led to tremendous achievements including visualization of a protein using fluorescent protein fusions and silence of a protein expression using RNA interferences. However, not all the proteins and related biological processes are within the easy reach of those conventional approaches. Fortunately, recent rapid progress in the chemical biology field provides abundant new technologies for the study

* Corresponding author.
E-mail address: panlf@sioc.ac.cn (L. Pan).

http://dx.doi.org/10.1016/j.tetlet.2015.03.065
0040-4039/

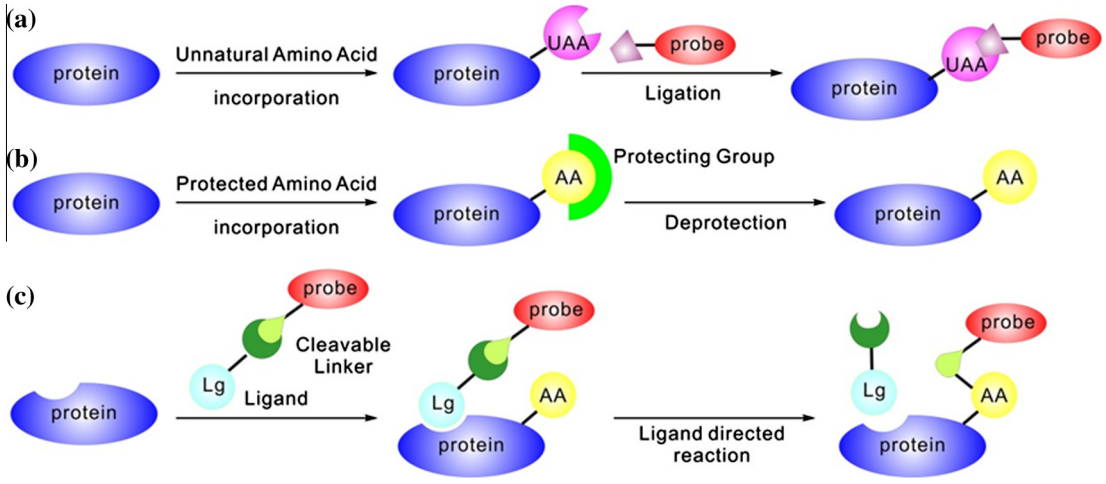

Figure 1. Three approaches to achieve the site-specific protein labeling and engineering. Schematic illustrations of site-specific incorporation of unnatural amino acids followed by bioorthogonal reactions including ligations (a) and deprotections (b). (c) Schematic illustration of the ligand-directed chemistry for site-specific protein labeling and engineering. UAA, unnatural amino acid; AA, natural amino acid; Lg, ligand.

of these challenging proteins and cellular processes. In particular, the modification of specific proteins with functional probes provides a powerful technique for the investigation of target proteins and their complex functions in detail. So far, there are a large variety of strategies developed by organic chemists to achieve site-specific labeling and engineering of target proteins with functional small molecules.[1] Due to space limitations, this digest only focuses on two strategies used for the site-specific protein labeling and engineering: (1) Bioorthogonal reactions with genetically encoded unnatural amino acids bearing functional groups that can be specifically ligated or deprotected (Fig. 1a and b); (2) Ligand-directed bioorthogonal reactions for site-specific modifications of target proteins (Fig. 1c).

The first strategy combines bioorthogonal reactions with genetically encoded unnatural amino acids bearing functional groups, such as aldehydes, ketones, azides, and alkenes, to facilitate the site-specific protein labeling and engineering (Fig. 1a and b).[2] Genetic code expansion and reprogramming enable the site-specific incorporations of diverse designed unnatural amino acids into proteins.[3] By evolving orthogonal ribosomes, developing mutually orthogonal synthetase/tRNA pairs and manipulating genomes, the efficiency of unnatural amino acids incorporations and the numbers of unnatural amino acids that can be site-specifically encoded are constantly increasing. Notablely, the development and application of the pyrrolysyl-transfer RNA (tRNA) synthetase/tRNA pair for unnatural amino acids incorporation have moved genetic code expansion from bacteria to eukaryotic cells and multicellular organisms.[4]

Although genetic code expansion is a powerful tool, it also has some limitations, for example, genetic modification and subsequent overexpression of proteins may perturb the physiological condition of cells. Thus, another strategy named ligand-directed chemistry for site-specific protein labeling that modifies selective endogenous proteins under their physiological conditions, was developed (Fig. 1c).[5] In this approach, a synthetic molecule containing three functional groups including a target protein binding ligand, a reactive linker and a functional probe is constructed, and firstly its ligand part specifically binds to its target protein, then driving a bioorthogonal reaction between the reactive linker group with an amino acid located at the vicinity of the ligand-binding site of the target protein facilitated by the proximity effect, finally end with the labeling of the target protein with the functional probe. Therefore, this method can satisfy the requirements of target selectivity and site specificity.

Since these two methods revolutionized our abilities to site-specifically label and manipulate intact proteins, these two areas are rapidly growing and many elegant applications have been recently reported. In this review, we summarize some recent developments in these two fields using bioorthogonal reactions for site-specific protein labeling and engineering.

Bioorthogonal reactions with genetically encoded unnatural amino acids for site-specific protein labeling and engineering

In general, the bioorthogonal reactions used for site-specific protein labeling and engineering based on genetically encoded unnatural amino acids can be mainly classified into two categories: ligation reactions (Fig. 1a) and deprotection reactions (Fig. 1b).

Bioorthogonal ligation reactions

Ligation reactions through aldehydes and ketones

Genetically encoded aldehydes and ketones can specifically react with hydrazides and alkoxyamines to produce stable hydrazone and oxime, respectively (Scheme 1a),[6] and were successfully applied for the site-specific in vitro or cell surface protein labeling.[7] However the unfavorable acidic conditions and slow kinetics of these reactions prevent their applications in most intracellular settings.[7a,8] To overcome these shortcomings, aniline was identified and used as a nucleophilic catalyst for both specific cell surface and intracellular protein labeling (Scheme 1b).[9] Based on the classic Pictet–Spengler reaction between aldehydes and tryptamine nucleophiles, recently Bertozzi and co-workers reported the Pictet–Spengler ligation reaction (Scheme 1c).[10] In this reaction, aldehydes react with alkoxyamines to form intermediate oxyiminium ions, which then undergo intramolecular C–C bond formations with indole nucleophiles to form hydrolytically stable oxacarboline products. In conjunction with techniques for genetic incorporations of unnatural amino acids bearing aldehydes, the Pictet–Spengler ligation provides a unique tool to generate stable bioconjugates for biomedical applications.

Ligation reactions through azides

Genetically encoded azide groups were firstly developed for the site-specific biomolecule labeling by the Bertozzi group through a process known as Staudinger ligation, a modification of the classic Staudinger reduction of azides with triphenylphosphine.[11] In this process, the proteins bearing alkyl azides undergo ligation reactions to form stable amide bonds with triarylphosphine derivatives