New Biotechnology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

New Biotechnology

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

Feature

Genetically encoding new bioreactivity

Lei Wang

Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA 94158, USA

ARTICLE INFO

Article history:
Received 16 July 2016
Received in revised form 25 September 2016
Accepted 5 October 2016
Available online xxx

Keywords:
Expansion of the genetic code
Unnatural amino acid
Proximity enabled reactivity
Protein crosslinking
Protein interaction
Protein engineering

ABSTRACT

The genetic code can be expanded to include unnatural amino acids (Uaas) by engineering orthogonal components involved in protein translation. To be compatible with live cells, side chains of Uaas have been limited to either chemically inert or bio-orthogonal (*i.e.*, nonreactive toward biomolecules) functionalities. To introduce bioreactivity into live systems, the genetic code has recently been engineered to encode a new class of Uaas, the bioreactive Uaas. These Uaas, after being incorporated into proteins, specifically react with target natural amino acid residues via proximity-enabled bioreactivity, enabling the selective formation of new covalent linkages within and between proteins both *in vitro* and in live systems. The new covalent bonding ability has been harnessed within proteins to enhance photostability, increase thermostability, staple proteins recombinantly, and build optical nano-switches, and between proteins to pinpoint ligand-receptor interaction, target native receptors irreversibly, and generate covalent macromolecular inhibitors. These diverse bioreactivities, inaccessible to natural proteins, thus open doors to novel protein engineering and provide new avenues for biological studies, biotherapeutics and synthetic biology.

1. Introduction

The essential framework for protein structure and function is provided by covalent peptide bonds in the backbone and *noncovalent* interactions among amino acid side chains. The disulfide bond formed between two cysteine residues is a unique *covalent* linkage between amino acid side chains, and plays crucial roles in the folding, stability, and activity of a variety of proteins such as antibodies, cytokines and membrane-associated receptors [1,2]. However, the weak bonding, reversibility, and redox sensitivity of the disulfide bond impose limitations on protein function, engineering, and applications [3]. Isopeptide bonds between Lys and Asn or Asp have also been discovered, but they form in the hydrophobic protein interior only and require an essential Glu or Asp residue to catalyze the reaction [4]. Although protein crosslinking is a common biological mechanism contributing to processes such as blood coagulation and tissue stabilization, the covalent crosslinking requires various enzymes to activate amino acid side chains for reaction [5]. In general, side chains of natural amino acid residues except cysteine rarely form covalent bonds spontaneously in native cellular environments. Therefore, a fundamental limitation of natural proteins is the inadequacy in covalent bonding of their side chains. Additional types of covalent linkages would expand avenues for generating novel protein properties and functions.

The genetic code specifies 20–22 natural amino acids for building proteins, and this code is preserved in virtually all life forms on earth with just minor variations [6,7]. In recent years, the genetic code has been artificially expanded to include unnatural amino acids (Uaas) by introducing into cells an orthogonal tRNA/aminoacyl-tRNA synthetase (aaRS) pair that is specific for the Uaa and compatible with protein translational machinery [8–10]. The orthogonal aaRS charges the desired Uaa onto the orthogonal tRNA, and the aminoacylated-tRNA incorporates the Uaa into proteins in response to a unique codon (the amber stop codon UAG is often used) through protein translation in live cells [11,12]. This methodology was initially established in *E. coli* [10], and later proven generally applicable in yeast [13,14], mammalian cells [15–17], stem cells [18], plants [19], invertebrates including *C. elegans* [20,21] and *Drosophila* [22], even as well as embryonic mouse brain [23]. More than 100 Uaas with a variety of side chains have been genetically incorporated into proteins using this approach [11,24–26], and this number is increasing with efforts from multiple research groups worldwide [27–33].

To be compatible with live cells or systems, all Uaas genetically incorporated in the past contain functional groups that are either chemically inert or bio-orthogonal, that is, they are non-bioreactive. To break the natural barrier to protein synthesis by enabling new covalent bonding ability for proteins in live cells,

E-mail address: Lei.Wang2@ucsf.edu (L. Wang).

http://dx.doi.org/10.1016/j.nbt.2016.10.003
1871-6784/

would it be possible to genetically encode bioreactive Uaas, *i.e.*, Uaas with functional groups that can react with biomolecules inside cells (Fig. 1)? Here we review how bioreactive Uaas, initially thought infeasible for genetic encoding, became a new class of Uaas specifiable by the genetic code in live cells. Examples are given to illustrate the unique abilities of these bioreactive Uaas affording to proteins via their specific reactivities.

2. Methodology

2.1. Genetically encoding a bioreactive uaa is a conundrum

The method of building new covalent linkages into proteins is to genetically incorporate a Uaa into the target protein and to enable the Uaa to react with a natural amino acid side chain to form a covalent bond selectively. Two conundrums arise for execution: 1) To form a new covalent bond between the side chains of a Uaa and a natural amino acid, the Uaa must be reactive toward the target natural amino acid. However, if the reactivity is uncontrolled, the ubiquitous presence of the natural amino acid in proteins and inside cells will result in undesired nonspecific linkages, which can potentially cause cytotoxicity as well. 2) The required bioreactivity may make the Uaa react with proteins involved in translation, thereby blocking the addition of the Uaa to the nascent protein *via* translation.

2.2. Proximity-enabled bioreactivity

We hypothesize that this impasse of bioreactivity, genetic encoding, and selectivity can be overcome via the concept of Proximity-Enabled Bioreactivity (Fig. 2) [34]. Specifically, we propose to tune the reactivity of the Uaa so that it does not react with free natural amino acids and other biomolecules under physiological conditions, thereby permitting genetic incorporation *in vivo via* endogenous translational machinery. When the Uaa comes into proximity to its target natural amino acid residue in proteins with appropriate orientation, the increased local effective concentration then facilitates the Uaa to selectively react with the side chain of the target natural residue to create a covalent bond. The bond can be built intramolecularly within a protein or intermolecularly between two proteins.

Feasibility of this hypothesis is supported by previous research on proximity effects on other systems. When small molecules are brought close by a DNA template, their reactivities can be enhanced several hundred folds [35]. Reactivities between proteins and small molecules can also be enhanced by proximity effect, which have been discovered in natural products [36] and utilized in developing irreversible ligands [37], small molecule inhibitors, and activity-based protein profiling [38]. Amino acid side chains readily get close to each other either within proteins or at the interface of interacting proteins, so we reasoned that harnessing the proximity of a designed Uaa and a target natural amino acid would enable the selective formation of a new covalent bond.

Existing chemistries for protein labeling and modification may not be directly applied for proximity-enabled bioreactivity in cells. Protein modification [39] generally target one type of residue, so the reactivity cannot be limited to a single residue and to the target protein alone in a mixture of proteins in cells. Innovative chemistries developed in activity-based protein profiling [38,40] usually target catalytic residues and are accelerated by enzyme catalytic mechanisms. For proximity-enabled bioreactivity, the covalent linkage can be desired anywhere inside or at the surface of protein, which might lack a catalytic active site. Bioorthogonal click

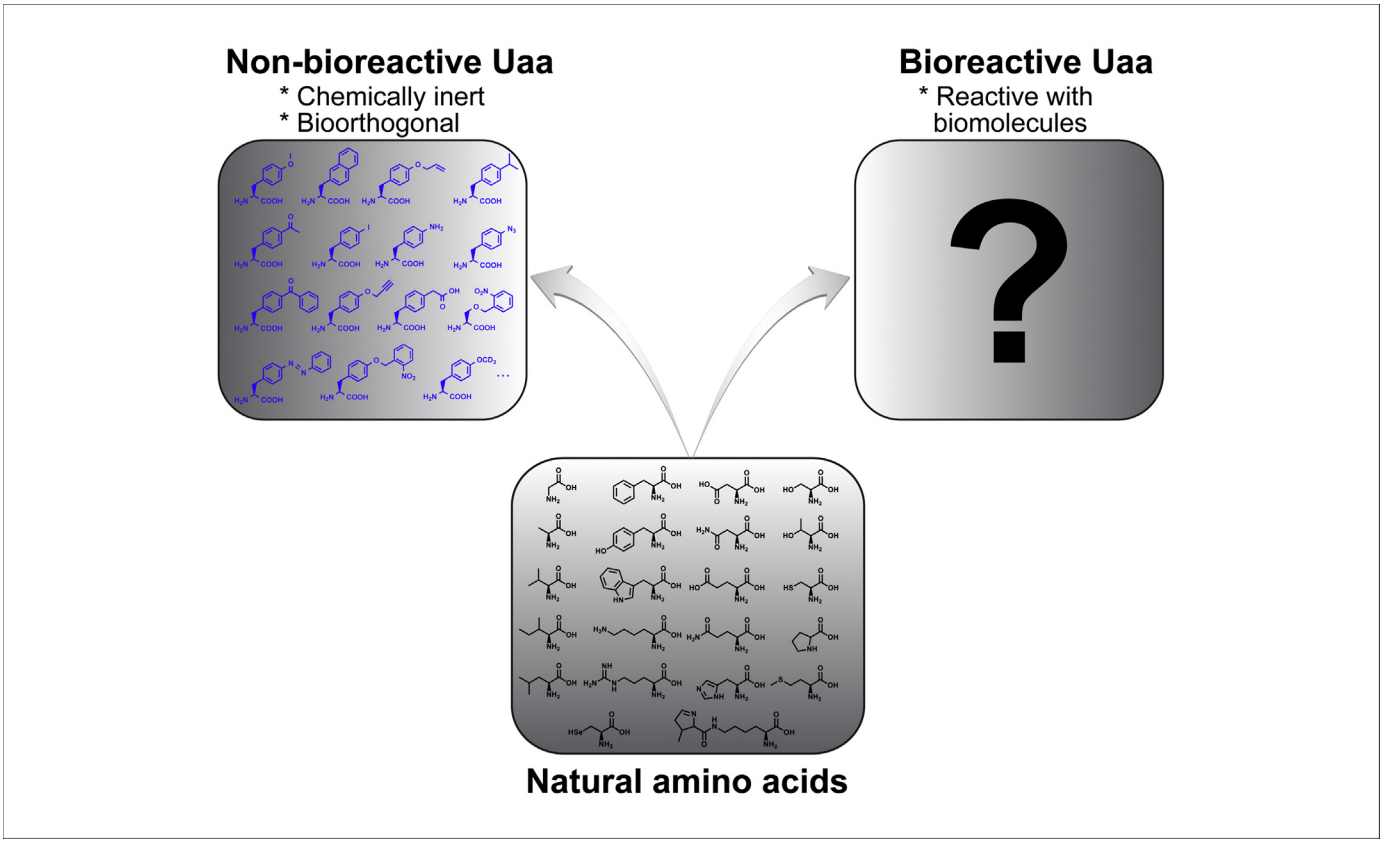

Fig. 1. The genetic code has been expanded to include non-bioreative Uaas. Would it be possible to genetically encode bioreactive Uaas?

Please cite this article in press as: L. Wang, Genetically encoding new bioreactivity, New Biotechnol. (2016), http://dx.doi.org/10.1016/j.nbt.2016.10.003