Protein Expression and Purification 80 (2011) 283–293

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

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

Review

An overview of enzymatic reagents for the removal of affinity tags

David S. Waugh *

Protein Engineering Section, Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702-1201, USA

ARTICLE INFO

Article history:
Received 2 August 2011
and in revised form 4 August 2011
Available online 19 August 2011

Keywords:
Carboxypeptidase
Aminopeptidase
PreScission protease
Affinity tag
Polyhistidine tag
His-tag
TEV protease
Enterokinase
Enteropeptidase
Thrombin
3C protease
Factor Xa

ABSTRACT

Although they are often exploited to facilitate the expression and purification of recombinant proteins, every affinity tag, whether large or small, has the potential to interfere with the structure and function of its fusion partner. For this reason, reliable methods for removing affinity tags are needed. Only enzymes have the requisite specificity to be generally useful reagents for this purpose. In this review, the advantages and disadvantages of some commonly used endo- and exoproteases are discussed in light of the latest information.

Published by Elsevier Inc.

Contents

Introduction

Affinity tags have become essential tools for the production of recombinant proteins in a wide variety of settings, from basic research to high-throughput structural biology. Not only do they facilitate the detection and purification of their fusion partners, as originally intended, but they may also have a beneficial impact on the yield of recombinant proteins and, in some cases, increase their solubility and even promote their proper folding [2,3].

Despite these important advantages, the Achilles heel of the affinity tagging strategy always has been and remains the removal

* Fax: +1 301 846 7148.
E-mail address: waughd@mail.nih.gov

1046-5928/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.pep.2011.08.005

Table 1
Endo- and exoproteases for the removal of affinity tags.

Enzyme	Source(s)	Molecular weight (kDa)	Tagged forms	Inhibitors	Recognition Site	Notes
Endoproteases						
Enteropeptidase	Duodenum *E. coli* *S. cerevisiae*	110 + 35	His_6	Reducing agents	DDDDK↓	P1′ ≠ Pro, Trp 35 KDa light chain is active by itself
Thrombin	Plasma CHO cells	32 + 4.5	None	Reducing agents	LVPR↓GS	
Factor Xa	Plasma HEK 293 cells	42 + 17	None	Reducing agents Chelating agents Phosphate ions	LVPR↓GS	Very promiscuous
TEV Protease	*E. coli*	27	His_6 MBP GST Strep II	Thiol alkylating agents	ENLYFQ↓G	P1′ can vary [68] P2′ ≠ Pro Ac-TEV™ = S219 V mutant
Rhinovirus 3C Protease	*E. coli*	27	His_6 GST His_6-GST	Thiol alkylating agents	LEVLFQ↓GP	Same as PreScission™ protease
Exoproteases						
Carboxypeptidase A	Pancreas *E. coli* *S. cerevisiae* *S. frugiperda* (baculovirus)	33	His_6	Reducing agents Chelating agents	C-terminal amino acids except Pro, Lys and Arg	Asp, Glu, Gly cleaved slowly
Carboxypeptidase B	Pancreas *E. coli* *P. pastoris*	35	none	Reducing agents Chelating agents	C-terminal Lys and Arg	Will cleave hydrophobic resides under certain conditions [119,120]
DAPase	Kidney *S. frugiperda* (baculovirus)	23 + 16 + 6	His_6	Reducing agents Thiol alkylating agents	N-terminal dipeptides	P2 ≠ Pro, Lys, Arg P1 ≠ Pro

of tags. Whereas many tagged proteins retain their structural integrity and biological activity, others clearly do not, e.g., [4–11]. Therefore, whenever possible, it is prudent to remove tags from recombinant proteins. Although both chemical and enzymatic methods have been used to cleave fusion proteins at designed sites, only the natural proteolytic enzymes have the requisite specificity to be broadly useful reagents for this purpose. Because they are not as versatile and therefore generally less useful than *trans*-acting reagents, neither the self-cleaving inteins [12] and self-cleaving variants of subtilisin [13] will be discussed here, nor will the Ulp1 protease since it only cleaves SUMO tags [14]. Rather, this review will focus on the most generally applicable and commonly used enzymatic reagents for the removal of affinity tags (Table 1). Since the last comprehensive review of this topic [3] much research on these reagents has been conducted. As a result, a wealth of new information has accumulated on the advantages, disadvantages, and biochemical characteristics of various reagents.

Endoproteases

For many years, serine proteases such as activated blood coagulation factor X (factor Xa),[1] enterokinase (hereafter referred to by its more appropriate moniker enteropeptidase), and α-thrombin were the reagents of choice for removing affinity tags, yet the literature is replete with reports of fusion proteins that were cleaved by these proteases at locations other than the designed site. Over the last decade or so, it has become increasingly evident that certain viral proteases have far more stringent sequence specificity, which has led to an upsurge in their popularity. These enzymes have a chymotrypsin-like fold with an atypical catalytic triad in which cysteine replaces serine, and they exhibit an absolute requirement for glutamine in the P1[2] position of their substrates. The nuclear inclusion protease from tobacco etch virus (TEV) is probably the best-characterized enzyme of this type. The other is the human rhinovirus 3C protease. In stark contrast to factor Xa, enteropeptidase and thrombin, there have been very few if any reports of cleavage at noncanonical sites in designed fusion proteins by these viral proteases.

The stringent specificity of the viral proteases probably can be attributed to their low turnover rates. The number of substrate residues that are recognized by the serine proteases and the viral proteases is similar (e.g., LVPRGS and ENLYFQS in the case of thrombin and TEV protease, respectively). The Michaelis constants (K_M) for the two classes of enzymes are also similar, falling in the low to mid micromolar range, but the catalytic rate constants (k_{cat}) of the viral proteases are on the order of 100-fold lower than those of the serine proteases, resulting in much slower turnover rates [15–19]. Each class of protease undoubtedly associates transiently with suboptimal recognition sites, but on average, a catalytic event is far more likely to occur when a serine protease does so because its k_{cat} is so much greater than that of the typical viral protease. The practical ramification of this observation is that one must use considerably more viral protease than serine protease to digest a fixed amount of fusion protein at a similar rate. However, this is not a significant handicap because, unlike the serine proteases, large quantities of recombinant viral proteases can easily be produced in *Escherichia coli*. This advantage, coupled with their more stringent sequence specificity, has made viral proteases the reagents of choice for endoproteolytic removal of affinity tags.

[1] *Abbreviations used:* factor Xa, blood coagulation factor X; TEV, tobacco etch virus; DTT, dithiothreitol; PMSF, phenylmethylsulfonyl fluoride; AEBSF, 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride; PEG, polyethylene glycol; GST, glutathione S-transferase; EGTA, ethylene glycol tetraacetic acid; EDTA, ethylenediaminetetraacetic acid; MBP, maltose binding protein; ATCC, American Type Culture Collection; TVMV, tobacco vein mottling virus; SARS, severe acute respiratory syndrome; BAP, biotin acceptor peptide; POI, protein of interest; BoCPA, bovine carboxypeptidase A; BoCPB, bovine carboxypeptidase B; SDS, sodium dodecyl sulfate; DAPase, dipeptidyl-aminopeptidase I; Qcyclase, glutamine cyclotransferase; pGAPase, pyroglutamylaminopeptidase.

[2] The nomenclature used here to describe individual amino acids in protease recognition sites and corresponding amino acid-binding sites in proteases was introduced by Schechter and Berger [1].