Journal of Biotechnology 231 (2016) 239–249

Contents lists available at ScienceDirect

Journal of Biotechnology

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

Review

Tobacco Etch Virus protease: A shortcut across biotechnologies

Francesca Cesaratto[a], Oscar R. Burrone[a,*], Gianluca Petris[b,*]

[a] International Centre for Genetic Engineering and Biotechnology, ICGEB, Trieste, Italy
[b] CIBIO, University of Trento, Trento, Italy

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form 31 May 2016
Accepted 10 June 2016
Available online 13 June 2016

Keywords:
TEV protease
Self-cleavage
In vivo applications
Protein engineering
Split-TEV
Subcellular targeting

ABSTRACT

About thirty years ago, studies on the RNA genome of Tobacco Etch Virus revealed the presence of an efficient and specific protease, called Tobacco Etch Virus protease (TEVp), that was part of the Nuclear Inclusion a (NIa) enzyme. TEVp is an efficient and specific protease of 27 kDa that has become a valuable biotechnological tool. Nowadays TEVp is a unique endopeptidase largely exploited in biotechnology from industrial applications to *in vitro* and *in vivo* cellular studies. A number of TEVp mutants with different rate of cleavage, stability and specificity have been reported. Similarly, a panel of different target cleavage sites, derived from the canonical ENLYFQ-G/S site, has been established. In this review we describe these aspects of TEVp and some of its multiple applications. A particular focus is on the use and molecular biology of TEVp in living cells and organisms.

Contents

* Corresponding authors.
E-mail addresses: burrone@icgeb.org (O.R. Burrone), gianluca.petris@unitn.it (G. Petris).

http://dx.doi.org/10.1016/j.jbiotec.2016.06.012
0168-1656/

1. Introduction

Proteases with stringent specificity for the target sequence represent excellent tools for the specific removal of tags and undesired sequences or to achieve precise processing on defined polypeptides. Several different enzymes of mammalian origin have been used for these purposes, such as the coagulation Factor Xa, enterokinases and α-thrombin. However, in some cases they were reported not to be highly specific, having off-targets effects. Viral proteases instead, have evolved in most cases to allow virus replication by precise processing of precursor polyproteins and have thus been used as targets for anti-viral drugs. Some of them are characterised by having stringent sequence specificity and efficiency, such as the human rhinovirus 3C protease, the Potyvirus proteases from the Tobacco Vein Mottling Virus (TVMV) and the Tobacco Etch Virus (TEV) (Blommel and Fox, 2007). Therefore, viral proteases have been widely exploited as biotechnological tools (Waugh, 2011). TEVp was taken in high consideration for these purposes because of its good activity rate and its sequence specificity, which was found to be much more stringent than that of other proteases generally used. In fact to our knowledge, TEVp has never been reported to cleave fusion proteins at unintended locations. Despite its high specificity, TEVp can tolerate a wide variety of residues in position P1′ (the first amino acid downstream the cleavage site) of the substrate, making it the best protease for endoproteolytic removal of affinity tags (Waugh, 2011). In addition, TEVp is the only protease that combines several useful features for biotechnological applications, from easy and economical production to availability of open source vectors and mutants, making it suitable for several different *in vitro* and *in vivo* applications.

2. The tobacco etch virus protease

2.1. Origin

TEV is a member of the Potyviridae family responsible for infections of many different plant species of Solanaceae including *N. tabacum* (Revers and Garcia, 2015). The TEV genome is a 9500 nucleotides-long single stranded RNA of positive polarity that encodes a 3054 amino acids-long polyprotein, which is co-translationally processed in infected cells into 10 mature products by three viral proteases (Adams et al., 2005; Allison et al., 1986). One of the most important TEV proteases is the Nuclear Inclusion protein a (NIa) (Carrington and Dougherty, 1987a), which processes most of the polyprotein (Fig. 1A and C) (Carrington and Dougherty, 1988; Carrington et al., 1988; Dougherty et al., 1988). NIa originates itself from the polyprotein as a 49 kDa precursor that is eventually cleaved into the N-terminal 21 kDa genome-linked protein VPg and the 27 kDa protease catalytic subunit TEVp, which can be detected during the late stages of plant infections (Carrington and Dougherty, 1987b; Dougherty and Parks, 1991).

TEVp belongs to the family of C4 peptidases and is structured as a two-domain antiparallel β-barrel fold, typical of trypsin-like proteases (Nunn et al., 2005; Phan et al., 2002) (Fig. 1B). In fact, TEVp shares homology to serine-proteases despite the presence of a cysteine instead of serine in the catalytic core, which comprises residues His46, Asp81 and Cys151. Consistently, mutation of Cys151 into Ala abolishes protease activity (Parks et al., 1995).

2.2. Specificity

The TEVp specific recognition motif (TEVp cleavage Site, TS) is only 7 amino acids-long: EXXYXQ-S/G (where X can be any amino acid) and proteolysis takes place between residues Gln and Ser or Gly (Dougherty et al., 1989a) (Fig. 1C). Despite an apparent low specificity, many studies have demonstrated that within the EXXYXQ-S/G cleavage site not all residues are equally tolerated. According to crystal structures, experimentally confirmed, residues in positions P6, P4, P3, P2, P1, and P1′ of the substrate directly interact with the enzyme substrate-binding pocket (Phan et al., 2002). Only residue P5 is exposed to the solvent and can tolerate almost any possible amino acid without consequences on enzyme effi-

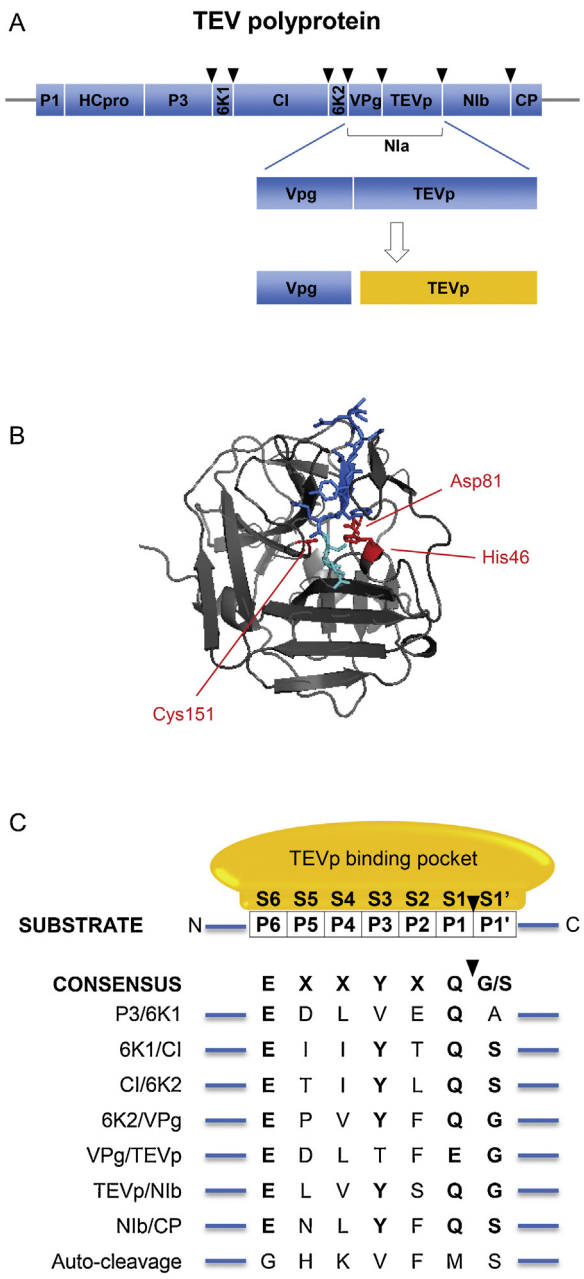

Fig. 1. The Tobacco etch virus protease. (A) TEV polyprotein processed by TEVp at the indicated sites (filled arrowhead). The NIa 49 kDa precursor is finally cleaved into VPg and TEVp. P1, S30 serine protease; HCpro, Helper Component for aphid transmission C6 cysteine protease; P3, protein P3; 6K1 and 2, 6 kDa proteins 1 and 2; CI, cylindrical inclusion protein; VPg, viral protein genome-linked, NIa and NIb: nuclear inclusion proteins a and b; CP, coat protein. (B) TEVp crystal structure, with key residues of the catalytic site shown in red. The substrate is represented with the TS amino acids P6-P1 in blue and peptide from P1′ to the C-terminus end in light blue (adapted from pdb file 1LVM, Phan et al., 2002). (C) Schematic illustration of residues on the TEVp substrate recognition pocket (S) and on the substrate cleavage site (P). The table reports the cleavage site sequences processed by the protease on the TEV polyprotein. Arrowhead, site of cleavage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).