The EXTENSIN enigma

Steven Moussu [a,b,*], Gwyneth Ingram [c,*]

[a] Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, 31320 Auzeville Tolosane, France
[b] Department of Plant Molecular Biology, University of Lausanne, 1015 Lausanne, Switzerland
[c] Laboratoire Reproduction et Développement des Plantes, ENS de Lyon, CNRS, INRAE, UCBL, Lyon, France

ARTICLE INFO

ABSTRACT

EXTENSINS (EXTs) are an abundant and yet enigmatic class of cell wall proteins that are found across multicellular plant lineages, from Bryophytes to Angiosperms. They have been shown to be integrated within the cell wall matrix, and are proposed to play key roles in the dynamic regulation of cell-wall properties. Consistent with this, EXTs are thought to be important for plant growth and development. However, like many other classes of cell wall proteins, EXTs are biochemically complex, highly diverse, and are encoded by multiple genes, making in-depth functional characterization a challenging undertaking. Here we will provide an overview of current knowledge of the biochemistry and properties of EXTs, and of the tools that have been deployed to study their biological functions in plants.

The name "EXTENSIN" was first coined by Lamport in 1963 (Lamport, 1963) who, as the name suggests, proposed that they might play a key role in the dynamic regulation of cell-wall properties. Their discovery was the first proof of the presence of proteins in plant cell walls, which were previously thought to be composed solely of carbohydrates. EXTs are a subclass of the even more diverse hydroxyproline-rich glycoproteins (HRGPs), and are defined by a consensus amino-acid domain (Marzol et al., 2018; Johnson, 2017). EXTs are further subdivided into 5 subfamilies, classical EXTs, short EXTs, Leucine-Rich-Repeat EXTs (LRXs), Proline-rich EXT Receptor Kinases (PERKs) and Formin-Homology EXTs (FH-EXTs) with different protein architectures as shown in Fig. 1A. EXTs are strictly defined by the presence of at least two repeats of the Ser-Pro$_{3-5}$ motif. The proline residues within this motif can be hydroxylated. They subsequently undergo O-glycosylation with short arabinose containing oligosaccharides (n = 4–5). This leads to the formation of hydrophilic domains (Cannon, 2008; Ogawa-Ohnishi et al., 2013; Velasquez et al., 2011).

In addition to Ser-Pro$_{3-5}$ motif, some EXTs also contain hydrophobic Tyr-containing motifs (Cannon, 2008; Held et al., 2004; Schnabelrauch et al., 1996). EXTs containing these motifs are sometimes denoted "classical" EXTs and have been shown, both *in vitro* and *in vivo*, to undergo peroxidase-mediated covalent cross linking through Tyr residues by forming pulcherosine or di-isodityrosine bonds (Fig. 1B) (Held et al., 2004; Brady et al., 1996). Intriguingly, and despite their name, this class

of EXTs, although present in all Angiosperms appears to be largely absent from monocot, gymnosperm and non-vascular plant genomes, in which genes encoding other types of EXT can more readily be found (Liu et al., 2016). Tyr mediated crosslinking has been proposed to be critical for some of the structural functions of classical EXTs in plant cell walls (Marzol et al., 2018). Classical EXTs have been shown to self-assemble into intricate dendritic networks *in vitro* (Cannon, 2008). This capacity, is proposed to be a result their amphiphilic properties due to the alternation of hydrophobic and hydrophilic domains (Lamport et al., 2011), leading to a self-assembly phenomenon. Although their proline-rich sequence is expected to yield disordered proteins, they harbor a rod-like structure *in vitro*, which is proposed to be due to the extensive arabinosylation along the peptidic backbone (Cannon, 2008). Additionally, EXTs have been proposed to serve as a scaffold for other cell wall components such as pectins. This association has been demonstrated *in vitro* (Valentin et al., 2010). In classical EXTs, basic residues appear at regular intervals along the protein sequence, and it is thought that these could mediate electrostatic interactions with negatively charged pectins. Basic residues are less common and, where present, less regularly spaced, in other classes of EXTs.

Because EXT backbones are genetically encoded, one might imagine that the analysis of loss of function mutants would provide a simple means of understanding their functions. However, to date, very few mutants in genes encoding classical EXTs have generated striking

phenotypes, possibly due to extensive functional redundancy, and to potential compensatory effects (either involving other EXT encoding genes, or wall-monitoring mechanisms, or both) (Zdanio et al., 2020) (reviewed in (Mishler-Elmore, 2021; Doll et al., 2022)). Although some atypical EXTs (such as LRXs and PERKs) have more clearly defined roles, the functions of their EXT domains generally remain poorly understood, although evidence suggests that the EXT domains of LRXs act as cell wall anchors and are necessary for their function (Ringli, 2010). In the absence of informative loss of function mutants, other approaches have been used to infer EXT function.

As seen in the above section, the post-translational modification of EXTs is a critical factor in their biochemistry. Levels of O-hydroxylation (and subsequent glycosylation) and the presence and degree of Tyr-mediated crosslinking are thought to influence both intra and inter-molecular interactions between EXTENSINS, their interactions with other cell wall components, and by extension, their functions. Because the enzymes involved in the post-translational modification of EXTs (particularly glycosylation) are generally encoded by smaller gene families than EXTs themselves, manipulations of the levels of these enzymes has been widely used to help infer EXT function (reviewed in (Marzol et al., 2018)). Three main enzymatic classes have been targeted.

Prolyl 4-hydroxylases (P4H): These enzymes are necessary for the O-hydroxylation of Proline (Fig. 1C), and are encoded by 13 genes in Arabidopsis. Inhibitors of P4H activity are also available (DP (a,a-dipyridyl) and EDHB (ethyl-3,4-dihydroxybenzoate)), and their application has been shown to inhibit root hair elongation. Consistent with this, three P4H enzymes, potentially acting in complex have also been shown to be important for root hair growth, one of which has been shown, *in vitro*, to preferentially hydroxylate the contiguous prolines characteristic of EXTs (Velasquez et al., 2011). However, specificity of P4H enzymes for EXTs has not been conclusively demonstrated. Proline hydroxylation of other HRGPs, (including those involved in signaling, such as other AGPs or indeed of unrelated peptides/receptors), may also be altered when P4H activities are modified (Sede et al., 2022). Care is

thus required in interpreting the results of these analyses as a demonstration of EXT function.

Glycosyl Transferases: The presence of consecutive hydroxyprolines in EXTs is thought to be a factor in determining the pattern of their subsequent substitution (O-glycosylation) with characteristic short chains of L-arabinofuranose. This process occurs in the secretory system and is mediated by the sequential action of the Hyp-O Arabinosyl-Transferase (HPAT) (encoded by 3 genes in Arabidopsis), Reduced Residual Arabinose (RRA) (encoded by 3 genes in Arabidopsis), XEG113 and ExAD glycosyltransferases (Fig. 1C). HPATs have been shown to play roles in both EXTENSIN glycosylation and in the glycosylation of signalling peptides such as CLE peptides. Loss of HPAT function affects pollen tube viability. However, because of the difficulty of defining specific substrates it is entirely possible that this could, for example, be due to loss of function of specific protein families, such as LRXs (Sede et al., 2022). The juxtaposition of hydroxyproline and serine, as found in (but not exclusive to EXTs) has also been shown to be necessary for the activity of a unique serine-glycosyl transferase (ATSERGT1). It is intriguing that tip growing cells (for example root hairs and pollen tubes in Arabidopsis) are particularly susceptible to alterations in the activity of the above classes of enzymes. However it should also be noted that opposite phenotypes have been observed in other cell types (for example increased expansion in hypocotyls and root), underlining the importance of cellular contexts for potential EXT functions (Velasquez et al., 2011; Zdanio et al., 2020).

Class III Peroxidases: After their secretion to the cell wall, classical EXTENSINS can be extensively covalently crosslinked via tyrosine residues (Held et al., 2004; Brady et al., 1996) (Fig. 1B). Whether this is the case for other EXTs remains unclear (Ringli, 2010). Crosslinking of classical EXTs is proposed to lead to wall rigidification, although clear demonstrations for this are limited. Crosslinking is thought to depend on the production of apoplastic Reactive Oxygen Species (ROS), combined with the activity of type III cell wall-localized peroxidases (Schnabelrauch et al., 1996), which are again encoded by a large multigene
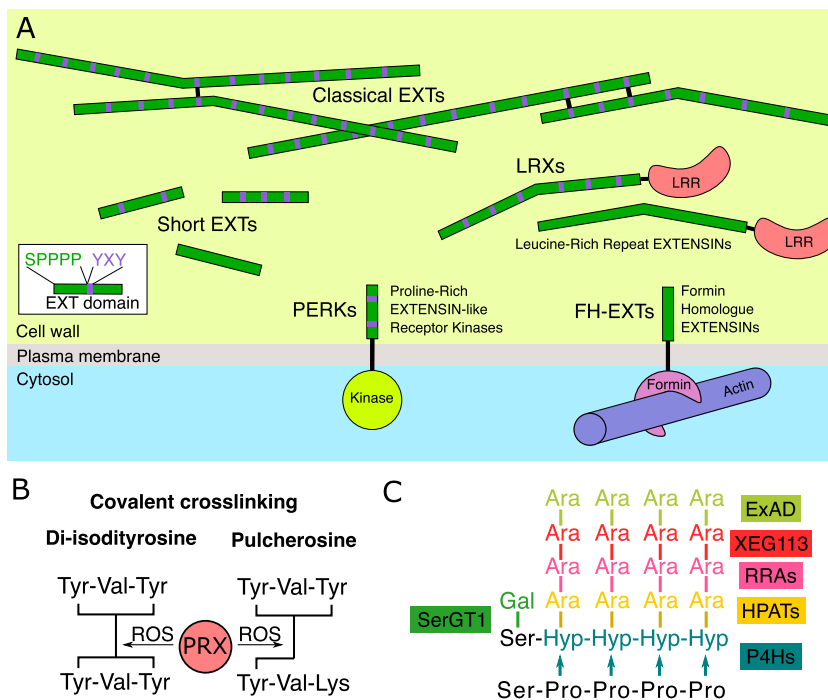

Fig. 1. EXTENSINs belong to a diverse family of post-translationally modified proteins. A. Overview of different classes of EXT-domain containing proteins in a cellular context. Long classical EXT can form intermolecular bonds (in black) and associate into a network. Short EXTs sometime harbor YXY motifs, potentially allowing crosslinking with other EXTs. Leucine-rich repeat EXTs (LRXs) harbor a *N*-terminal LRR domain fused to an EXT region and are embedded in the cell wall matrix. Proline-rich Extensin Receptor Kinases (PERKs) are receptors with an intracellular kinase domain, a transmembrane domain, and an ectodomain containing EXT motifs. Formin-homologue EXTs (FH) are composed of an EXT domain extracellular region, a transmembrane domain, and a Formin-like region which is predicted to interact with actin filaments. SPPPP motifs are represented in green and YXY motifs in purple showing the respective alternance of hydrophilic and hydrophobic residues. B. Intermolecular bonds bridging EXTENSINs. YXY motifs can form intramolecular isodityrosine bonds that can then interact in an intermolecular fashion with either a single tyrosine, forming a pulcherosine bond, or with another isodityrosine, leading to the formation of a di-isodityrosine. Formation of pulcherosine and di-isodityosine linkages requires the action of class III peroxidases (PRXs) in the presence of reactive oxygen species (ROS). C. Post-translational modifications of the proline-rich EXT motif. Proline residues are hydroxylated to hydroxyproline (Hyp) by Prolyl-4-hydroxylases (P4H). Hyp residues are then sequentially glycosylated with 4–5 arabinoses (Ara). Ara residues are added by different glycosyl-transferases. The first Ara is linked by HPATs, the second by RRAs, the third by XEG113 and the final one by ExAD. Galactose is added to the serine before the proline stretch by SerGT1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

family. Although proteins from this class can undoubtedly crosslink EXTs *in vitro* (Held et al., 2004); their substrate specificity is far from clear, and members of the same family have also been implicated in the crosslinking of other aromatic moieties, including phenylpropanoids (for example to form lignin-related polymers) (Rojas-Murcia et al., 2020), and proposed to play a role in cell wall weakening through the production of hydroxyl radicles (Dauphin et al., 2022). Consistent with their potentially diverse roles, several peroxidase mutants have been shown to have developmental phenotypes, but due to the factors highlighted above, these phenotypes cannot be conclusively attributed to altered EXT function.

As can be seen above, the potentially non-EXT-specific activities of enzymes involved in EXT modification and cell wall integration makes them potentially misleading tools for the inference of EXT function. A lack of defined functions for EXTs, and particularly classical EXTs, is compounded by the difficulty of interpreting the outputs of the limited resources available for extraction and analysis/detection of EXTs, particularly when they are covalently integrated into the cell wall matrix. One of the most widely-used techniques for detecting EXTs depends on their antigenicity to a handful of monoclonal antibodies (http://www.wallmabdb.net) (Smallwood et al., 1995, 1994). However, again, the exact EXT-specificity of these antibodies, their reactivity to different classes of EXTs, and their sensitivity to differential glycosylation and/or association/crosslinking with themselves or other cell wall components remains unclear.

In summary: Despite years of research, EXTs remain poorly understood due to technical hurdles in their *in vivo* detection and manipulation, the promiscuity of modifying enzymes and genetic redundancy. Although *in vitro* data strongly support the idea that EXTs could play key roles in cell-wall architecture, conclusive mechanistic evidence for this *in vivo* remains sparse and largely correlative, particularly for classical EXTs. Future research may benefit from the new wave of efficient genome-editing tools that are currently emerging, to allow the stable or transient generation of higher order mutants both within, and between, EXT families. Uncovering strong loss of function phenotypes will be critical for advancing understanding of EXT functions *in planta* both at the physiological and biochemical level. Given the proposed role of EXTs in the modification of the biophysical properties of cell walls, and their ROS-dependent crosslinking, it will also be interesting to assess in more detail to what extent EXTs, and their cell wall integration, are implicated in sensing and responding to biotic and abiotic stresses, as well as mechanical cues in growing plant tissues.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Brady, J.D., Sadler, I.H., Fry, S.C., 1996. Di-isodityrosine, a novel tetrametric derivative of tyrosine in plant cell wall proteins: a new potential cross-link. Biochem. J 315 (Pt 1), 323–327.

Cannon, M.C., et al., 2008. Self-assembly of the plant cell wall requires an extensin scaffold. Proc. Natl. Acad. Sci. U.S.A. 105, 2226–2231.

Dauphin, B.G., Ranocha, P., Dunand, C., Burlat, V., 2022. Cell-wall microdomain remodeling controls crucial developmental processes. Trends Plant Sci. 27 (10), 1033–1048.

Doll, N.M., Berenguer, E., Truskina, J., Ingram, G., 2022. AtEXT3 is not essential for early embryogenesis or plant viability in Arabidopsis. New Phytol. https://doi.org/10.1111/nph.18452.

Held, M.A., Tan, L.i., Kamyab, A., Hare, M., Shpak, E., Kieliszewski, M.J., 2004. Di-Isodityrosine is the intermolecular cross-link of isodityrosine-rich extensin analogs cross-linked in vitro. J. Biol. Chem. 279 (53), 55474–55482.

Johnson, K.L., et al., 2017. Insights into the evolution of hydroxyproline-rich glycoproteins from 1000 plant transcriptomes. Plant Physiol. 174, 904–921.

Lamport, D.T.A., 1963. Oxygen fixation into hydroxyproline of plant cell wall protein. J. Biol. Chem. 238 (4), 1438–1440.

Lamport, D.T.A., Kieliszewski, M.J., Chen, Y., Cannon, M.C., 2011. Role of the extensin superfamily in primary cell wall architecture. Plant Physiol. 156, 11–19.

Liu, X., Wolfe, R., Welch, L.R., Domozych, D.S., Popper, Z.A., Showalter, A.M., Zabotina, O.A., 2016. Bioinformatic identification and analysis of extensins in the plant kingdom. PLoS One 11 (2), e0150177.

Marzol, E., Borassi, C., Bringas, M., Sede, A., Rodríguez Garcia, D.R., Capece, L., Estevez, J.M., 2018. Filling the gaps to solve the extensin puzzle. Mol. Plant 11 (5), 645–658.

Mishler-Elmore, J.W., et al., 2021. Extensins: self-assembly, crosslinking, and the role of peroxidases. Front. Plant Sci. 12, 664738.

Ogawa-Ohnishi, M., Matsushita, W., Matsubayashi, Y., 2013. Identification of three hydroxyproline O-arabinosyltransferases in Arabidopsis thaliana. Nat. Chem. Biol. 9 (11), 726–730.

Ringli, C., 2010. The hydroxyproline-rich glycoprotein domain of the Arabidopsis LRX1 requires Tyr for function but not for insolubilization in the cell wall. Plant J. 63, 662–669.

Rojas-Murcia, N., Hématy, K., Lee, Y., Emonet, A., Ursache, R., Fujita, S., De Bellis, D., Geldner, N., 2020. High-order mutants reveal an essential requirement for peroxidases but not laccases in Casparian strip lignification. Proc. Natl. Acad. Sci. U.S.A. 117 (46), 29166–29177.

Schnabelrauch, L.S., Kieliszewski, M., Upham, B.L., Alizedeh, H., Lamport, D.T.A., 1996. Isolation of pI 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link site. Plant J. 9 (4), 477–489.

Sede, A.R. et al. 2022. Arabidopsis pollen Prolyl-hydroxylases P4H4/6 are required for correct hydroxylation and secretion of LRX11 in pollen tubes. 2022.11.16.516804 Preprint at https://doi.org/10.1101/2022.11.16.516804.

Smallwood, M., Beven, A., Donovan, N., Neill, S.J., Peart, J., Roberts, K., Knox, J.P., 1994. Localization of cell wall proteins in relation to the developmental anatomy of the carrot root apex. Plant J.. 5 (2), 237–246.

Smallwood, M., Martin, H., Knox, J.P., 1995. An epitope of rice threonine- and hydroxyproline-rich glycoprotein is common to cell wall and hydrophobic plasma-membrane glycoproteins. Planta 196, 510–522.

Valentin, R., Cerclier, C., Geneix, N., Aguié-Béghin, V., Gaillard, C., Ralet, M.-C., Cathala, B., 2010. Elaboration of extensin-pectin thin film model of primary plant cell wall. Langmuir 26 (12), 9891–9898.

Velasquez, S.M., Ricardi, M.M., Dorosz, J.G., Fernandez, P.V., Nadra, A.D., Pol-Fachin, L., Egelund, J., Gille, S., Harholt, J., Ciancia, M., Verli, H., Pauly, M., Bacic, A., Olsen, C.E., Ulvskov, P., Petersen, B.L., Somerville, C., Iusem, N.D., Estevez, J.M., 2011. O-glycosylated cell wall proteins are essential in root hair growth. Science 332 (6036), 1401–1403.

Zdanio, M., Boron, A.K., Balcerowicz, D., Schoenaers, S., Markakis, M.N., Mouille, G., Pintelon, I., Suslov, D., Gonneau, M., Höfte, H., Vissenberg, K., 2020. The proline-rich family protein EXTENSIN33 is required for etiolated Arabidopsis thaliana hypocotyl growth. Plant Cell Physiol. 61 (6), 1191–1203.