Protein tyrosine phosphorylation in plants: more abundant than expected?

Sergio de la Fuente van Bentem¹ and Heribert Hirt¹,²

¹ Department of Plant Molecular Biology, Max F. Perutz Laboratories, University of Vienna, Dr. Bohr-Gasse 9, 1030 Vienna, Austria
² Unité de Recherche en Génomique Végétale, INRA-CNRS-UEVE, 2 rue Gaston Crémieux, 91067 Evry Cédex, France

Protein phosphorylation in eukaryotes predominantly occurs on serine (Ser) and threonine (Thr) residues, whereas phosphorylation on tyrosine (Tyr) residues is less abundant. Plants lack classic Tyr kinases, such as the epidermal growth factor receptor, that govern Tyr phosphorylation in animals. A long-standing debate questions whether plants have any Tyr-specific kinases and, although several protein kinases with both Ser/Thr and Tyr specificities exist, data supporting the existence of other such kinases are scarce. As we discuss here, mass-spectrometry-based analyses now indicate that Tyr phosphorylation is as extensive in plants as it is in animals. However, careful inspection of available data indicates that these promising mass spectrometry studies have to be interpreted with caution before current ideas on Tyr phosphorylation in plants are revised.

Tyrosine phosphorylation in plants

Many proteins are phosphorylated on serine (Ser), threonine (Thr) and/or tyrosine (Tyr) residues. In contrast to Ser/Thr phosphorylation, Tyr phosphorylation is less common. Animals contain a large family of receptor Tyr kinases, which interact with ligands at the plasma membrane and subsequently mediate Tyr phosphorylation of an extensive array of downstream targets. Plant genomes do not encode such receptor Tyr kinases, and this observation indicates that Tyr phosphorylation in plants occurs less frequently in plants than in animals [1]. By searching for Tyr-kinase-specific sequence motifs, a bioinformatic screen for Tyr kinases in plants identified several potential dual-specificity kinases (DSKs) but no true protein Tyr-specific kinases (PTKs) [2]. The situation in plants might thus be similar to that in yeast, which has DSKs but seems to lack PTKs [3]. However, other bioinformatic screens predicted two PTKs in the A. thaliana genome [4], or 34 putative PTKs and five DSKs, indicating that ~4% of all A. thaliana protein kinases can phosphorylate Tyr residues [5] (http://www.bio.unipd.it/molbinfo/PTKtable.html). Based on the few bioinformatic screens available, it thus cannot be concluded how many plant PTKs and DSKs exist; however, improved search algorithms and high-throughput experiments could provide a more decisive answer to this question.

Recently, the use of mass spectrometry (MS) has expanded the number of identified Tyr phosphorylation sites in many organisms, including plants. Here, we discuss Tyr phosphorylation in plants in light of these novel discoveries, which show that Tyr phosphorylation is more common in plants than anticipated and enable screening for the responsible protein kinases. We also evaluate the described Tyr phosphorylation sites regarding potential discrepancies associated with the newly developed technological workflows that enable phosphorylation site mapping by MS.

Identification of candidate protein kinases and phosphatases that govern Tyr phosphorylation

After the first description of Tyr phosphorylation of pea proteins in 1986 [6], several studies showed, by using immunoblotting with phosphotyrosine-specific antibodies, that different plant proteins are phosphorylated on Tyr residues [5,7–14]. With only few exceptions, however, the phosphorylated (p)Tyr sites and the kinases involved have remained undiscovered.

Several plant protein kinases autophosphorylate on Tyr, Ser and Thr [2,13,15–20], including two RAF-like kinases identified as STY5 (also known as RAF31) and STY13 (RAF22) [2]. Several mitogen-activated protein kinases (MAPKs; or MPKs) and SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1 (SERK1) transphosphorylate artificial substrates on Tyr [17,18,21,22]. These studies show that plants have kinases that exhibit dual specificity (towards both Ser and/or Thr in addition to Tyr) [21]. However, the first pTyr sites of these kinases still await identification.

A. thaliana has a few protein Tyr-specific phosphatases (PTPs) and 22 dual-specificity phosphatases (DSPs) [23,24]. These numbers are much lower than in humans, who contain ~100 members of the PTP superfamily (including ~60 DSPs) [23]. The difference between the numbers of PTPs and DSPs is particularly striking considering the fact that A. thaliana has twice as many protein kinases than do humans [25]. This implies either that the pTyr component of plant proteomics is limited or that plant PTPs and DSPs target more sites.

Based on the use of inhibitors of Tyr kinases and phosphatases generally used in animal systems, a recent study indicates that Tyr phosphorylation has an important role in plant signalling [26]. Several other reports implicated Tyr phosphorylation in disease-resistance signalling. For instance, the plant pathogen P. syringae injects the virulence factor HopPtoD2, which is a PTP, into its host cytoplasm [27,28]. Its catalytic activity suppresses programmed cell death [27,28], just like that of another P. syringae virulence factor.
virulence factor, HopAO1, which is also predicted to be a PTP [29]. MAPKs are potential targets of these virulence factors and, although HopAO1 does not target MPK3 and MPK6 [29], several plant PTPs and DSPs are able to dephosphorylate MAPKs [30–32].

The identification of pTyr-containing proteins and the occurrence of a family of PTPs and DSPs indicated that Tyr phosphorylation in plants is more abundant than assumed. This raised several questions, including: how abundant is Tyr phosphorylation in plants? What are the responsible plant protein kinases, and what substrates and sites are their targets? To be able to answer these questions, the individual pTyr sites need to be identified on a global scale. Because of recent advances in the field, MS has become an optimal technique to obtain such information.

MS-based mapping of Tyr phosphorylation sites

Early estimations based on phosphoamino acid analysis of 32P-labelled cells by partial acid hydrolysis indicated a low occurrence of Tyr phosphorylation in animals (0.05%), whereas phosphorylation on Ser (90%) and Thr (10%) were predicted to be more abundant [33]. However, medium- and large-scale (several hundred to >10,000 sites per study) MS-based studies recently indicated that the frequency of Tyr phosphorylation ranges from 2% to 3% in animals [34–36] and <1% in yeast [37,38]. A few MS-based phosphoproteomic screens carried out on Arabidopsis thaliana samples indicated that Tyr phosphorylation in plants is similar to those in yeast: 0% (0 on 400 sites) [39], 0.7% (two on 303 sites) [40] and again 0.7% (11 on 1678 sites; the identity of the pTyr sites was not revealed in this work) [41]. These studies collectively seemed to confirm the predicted low occurrence of Tyr phosphorylation in plants.

However, a recent study described 94 pTyr sites on a total of 2172 phosphorylation sites in Arabidopsis proteins [42], indicating that 4.3% of all phosphorylation events occur on Tyr residues – a figure that is higher than predicted for animals. Another group published 14 pTyr sites among a total of 105 sites in proteins from the moss Physcomitrella patens [43]. Other experiments can be performed to confirm unambiguously the Tyr phosphorylation sites identified by MS-based phosphoproteomic screens. For instance, pTyr-containing phosphopeptides can be synthesized and analysed by MS to show that these peptides fragment similarly as the original phosphopeptides identified in the screens [44]. Otherwise, sites can be mutagenised and mutated proteins can be compared with wild-type proteins by immunoblotting with anti-pTyr antibodies [40]. Tyr phosphorylation of proteins can be shown by phosphoamino acid analysis of 32P-labelled cells by partial acid hydrolysis. However, this technique only shows the global pSer, pThr and pTyr levels of a purified protein and does not give information about which sites are phosphorylated. Thus, although global MS analysis is the most appropriate method to indicate pTyr sites on a global scale, additional experiments are required to prove that a particular pTyr site assignment is correct.

The large discrepancy in predictions of the abundance of Tyr phosphorylation in plants inspired us to have a closer look at the published datasets. Some of the differences in the plant studies could be explained by the fact that either subcellular fractions or whole-cell extracts were analysed, or by the use of different phosphopeptide affinity purification protocols. However, several peculiarities related to Tyr phosphorylation were noted, which are important when interpreting the datasets.

In contrast to pSer- and pThr-containing peptides, most pTyr-containing phosphopeptides were reported to have one or more additional phosphosites [42]. Moreover, many combinations of phosphosites were identified in multiply phosphorylated peptides. For instance, for the doubly phosphorylated peptide RYSPPYYSPPR, four different combinations of phosphosite assignments were described: pTyr2 with pSer8; pSer3 with pTyr6; pSer3 with pTyr7; and pSer3 with pSer8 [42]. The differential assignment of phosphosites in independently fragmented peptides is not remarkable, but to find this many different combinations for a single phosphopeptide is rare [39,40,45,46].

Differing from previous studies, the occurrence of Tyr phosphorylation in animals [34–36] was estimated to be low, whereas phosphorylation on Ser (90%) and Thr (10%) was predicted to be more abundant [33]. However, medium- and large-scale (several hundred to >10,000 sites per study) MS-based studies recently indicated that the frequency of Tyr phosphorylation ranges from 2% to 3% in animals [34–36] and <1% in yeast [37,38]. A few MS-based phosphoproteomic screens carried out on Arabidopsis thaliana samples indicated that Tyr phosphorylation in plants is similar to those in yeast: 0% (0 on 400 sites) [39], 0.7% (two on 303 sites) [40] and again 0.7% (11 on 1678 sites; the identity of the pTyr sites was not revealed in this work) [41]. These studies collectively seemed to confirm the predicted low occurrence of Tyr phosphorylation in plants.

However, a recent study described 94 pTyr sites on a total of 2172 phosphorylation sites in A. thaliana proteins [42], indicating that 4.3% of all phosphorylation events occur on Tyr residues – a figure that is higher than predicted for animals. Another group published 14 pTyr sites among a total of 105 sites in proteins from the moss Physcomitrella patens [43]. Other experiments can be performed to confirm unambiguously the Tyr phosphorylation sites identified by MS-based phosphoproteomic screens. For instance, pTyr-containing phosphopeptides can be synthesized and analysed by MS to show that these peptides fragment similarly as the original phosphopeptides identified in the screens [44]. Otherwise, sites can be mutagenised and mutated proteins can be compared with wild-type proteins by immunoblotting with anti-pTyr antibodies [40]. Tyr phosphorylation of proteins can be shown by phosphoamino acid analysis of 32P-labelled cells by partial acid hydrolysis. However, this technique only shows the global pSer, pThr and pTyr levels of a purified protein and does not give information about which sites are phosphorylated. Thus, although global MS analysis is the most appropriate method to indicate pTyr sites on a global scale, additional experiments are required to prove that a particular pTyr site assignment is correct.

The large discrepancy in predictions of the abundance of Tyr phosphorylation in plants inspired us to have a closer look at the published datasets. Some of the differences in the plant studies could be explained by the fact that either subcellular fractions or whole-cell extracts were analysed, or by the use of different phosphopeptide affinity purification protocols. However, several peculiarities related to Tyr phosphorylation were noted, which are important when interpreting the datasets.

In contrast to pSer- and pThr-containing peptides, most pTyr-containing phosphopeptides were reported to have one or more additional phosphosites [42]. Moreover, many combinations of phosphosites were identified in multiply phosphorylated peptides. For instance, for the doubly phosphorylated peptide RYSPPYYSPPR, four different combinations of phosphosite assignments were described: pTyr2 with pSer8; pSer3 with pTyr6; pSer3 with pTyr7; and pSer3 with pSer8 [42]. The differential assignment of phosphosites in independently fragmented peptides is not remarkable, but to find this many different combinations for a single phosphopeptide is rare [39,40,45,46].

Finally, Tyr phosphorylation was also reported to occur in many splicing factors [42]. Although others have also uncovered numerous phosphorylation sites in these proteins [40,45,46], identifying a set of overlapping splicing factor phosphopeptides, a single pTyr site has not yet been identified [46].

Determining protein phosphorylation sites by MS includes the following: mass spectra of phosphopeptides obtained by MS are aligned with a database of fragmented (phospho)peptides, resulting in a top hit followed by lower ranking hits. Each hit comprises a peptide sequence, including the assignment of the phosphate(s) to a particular residue. Often, the top hit is a peptide sequence with a particular amino acid assigned as the phosphosite. The second-ranking hit can be the same phosphopeptide sequence but with the phosphosite mapped to a different amino acid. Visual inspection of the fragmentation spectrum of a peptide can help to ensure that the computational assignment of a phosphate group to a specific amino acid in a phosphopeptide sequence is correct. Because this is not feasible for a large-scale analysis, ~40% of phosphosite assignments must be considered to be ambiguous [47] (Box 1). The Ascore [47] and the post-translational modification score [36] methods were designed to obtain confident phosphosite assignments while avoiding the laborious manual verification of phosphopeptide spectra. Without the use of these methods, some authors have claimed to reach 86% ambiguity, stating that ‘phosphorylated sites were unambiguously determined when y- or b-ions between which the phosphorylated residue exists were observed in the peak lists of the fragment ions’ [42]. However, the use of this criterion alone can easily lead to mis-assignments. To achieve high reliability in phosphosite assignments, thorough visual inspection of all phosphopeptide spectra is necessary [40,45,46]. Without visual inspection, more information of each spectrum is required, such as the total number of fragments generated [47].

Reassessment of described phosphotyrosine sites

The aforementioned discussion prompted us to scrutinise the Tyr phosphosite assignments in the two MS-based
Box 1. Specialist language used in MS-based phosphoproteomic analysis

Collision-induced dissociation, also termed collision-activated dissociation
Collision-induced dissociation (CID) is a process used for fragmentation of ionised (phospho)peptides in mass spectrometers. A ‘precursor ion’ (the initial peptide) is bombarded with gas molecules, and the collisional energy causes the peptide to fragment. In non-phosphopeptides, this usually occurs at the peptide bonds yielding sequence-specifying ‘product ions’. When a peptide breaks into two product ions, it gives an N-terminal fragment, called a ‘b ion’, and a C-terminal fragment, the ‘y ion’ (Box 2). In the case of phosphopeptides fragmented by CID, however, the main fragmentation event is the neutral loss of phosphoric acid from phosphoserine or phosphothreonine [48]. This gives a high intensity peak corresponding to the mass of the precursor ion lacking phosphoric acid (e.g. the fragmentation spectrum in Box 1). This hampers both the sequencing of the phosphopeptide and the phosphosite assignment because product ions are much less intense. A new generation of mass spectrometers with extremely high mass accuracy has solved the first problem, but not the second.

Peptide fragmentation spectrum or mass spectrum
Collection of peptide fragment ions detected by the mass spectrometer as a result of peptide fragmentation (Box 2).

Ambiguous and unambiguous phosphosite assignment
The fragmentation spectrum of a phosphopeptide enables the localization of the phosphorylated residue on the basis of its mass (Box 2). Phosphopeptide fragmentation spectra generated by CID are usually very poor, resulting in ambiguous phosphosite assignments (i.e. the phosphogroup cannot be confidently localized to a particular residue). Moreover, current database search software still cannot account for sequence-dependent ion intensities. This decreases the confidence of automated phosphosite assignment versus visual inspection, which does enable to account for sequence-dependent ion intensities. If the fragmentation spectrum does enable the assignment of the phosphogroup to a specific residue, the site is unambiguous.

Singly, doubly and multiply phosphorylated peptide
A peptide containing one phosphogroup is termed singly phosphorylated; a peptide with two phosphogroups is called doubly phosphorylated, and so on. A peptide with two or more phosphogroups is a multiply phosphorylated peptide.

Kinase consensus motif
A kinase consensus motif is a short conserved sequence around the phosphosite that is preferred by a kinase. Each kinase (family) interacts with one or more residues in the near vicinity of the residue that it phosphorylates. For instance, MAPKs prefer a Pro at the +1 position (the position immediately C terminally of the phosphosite) and target the minimal consensus motifs pSer-Pro and pThr-Pro. This requirement is never 100% strict, and the same kinase consensus motif can be targeted by different kinases. For instance, a pSer-Pro or pThr-Pro motif can also be phosphorylated by cyclin-dependent kinases.

Dual specificity
Protein kinases and protein phosphatases can be divided into enzymes with specificity towards either Ser and Thr (Ser/Thr-specific) or Tyr (Tyr-specific). Some kinases and phosphatases target both Ser/Thr and Tyr residues, and are therefore termed dual-specificity kinases (DSKs) and phosphatases (DSPs).
In contrast to the many convincing Tyr sites in protein kinases [42], Tyr phosphorylation of splicing factors is more questionable. Many of the phosphopeptides of these proteins were previously identified and suggested to be non-Tyr phosphorylated [46]. Numerous pSer and pThr sites have been mapped in animal splicing factors, but no single pTyr site has been found [45]. Strikingly, all splicing-factor phosphopeptides were found several times with the same number of phosphosites [42]. However, independent fragmentations of each phosphopeptide yielded different assignments, giving many combinations of the actual sites within these peptides. For each peptide, there exists one assignment of all phosphogroups to non-Tyr residues (Table S1 in the supplementary material online).

In contrast to the many convincing Tyr sites in protein kinases [42], Tyr phosphorylation of splicing factors is more questionable. Many of the phosphopeptides of these proteins were previously identified and suggested to be non-Tyr phosphorylated [46]. Numerous pSer and pThr sites have been mapped in animal splicing factors, but no single pTyr site has been found [45]. Strikingly, all splicing-factor phosphopeptides were found several times with the same number of phosphosites [42]. However, independent fragmentations of each phosphopeptide yielded different assignments, giving many combinations of the actual sites within these peptides. For each peptide, there exists one assignment of all phosphogroups to non-Tyr residues (Table S1 in the supplementary material online).

**Figure I.** Phosphosite assignment within a phosphopeptide by analysis of its fragmentation spectrum.

**Table 1.** Tyrosine phosphorylation sites identified by mass spectrometry in *Arabidopsis thaliana* protein kinases

<table>
<thead>
<tr>
<th>Protein kinase</th>
<th>Site(s)</th>
<th>Responsible kinase(s)</th>
<th>Activation loop</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPK4 or MPK11</td>
<td>Tyr203 or Tyr200</td>
<td>MAPK kinase(s) or autophosphorylation</td>
<td>Yes</td>
<td>[17,42]</td>
</tr>
<tr>
<td>MPK6</td>
<td>Tyr223</td>
<td>MAPK kinase(s)</td>
<td>Yes</td>
<td>[42]</td>
</tr>
<tr>
<td>AtSK11 (also known as ASKα) or AtSK12 (ASKγ) or ASKβ or ASKα or AtGSK1 (ASKα)</td>
<td>Tyr229, Tyr233, Tyr242, Tyr234 or Tyr230</td>
<td>Autophosphorylation</td>
<td>Yes</td>
<td>[40,42]</td>
</tr>
<tr>
<td>BIN2 (also known as USU1 and ASKn) or ASKγ</td>
<td>Tyr200 or Tyr232</td>
<td>Autophosphorylation</td>
<td>Yes</td>
<td>[42]</td>
</tr>
<tr>
<td>AtK-1 (also known as ASKα)</td>
<td>Tyr243</td>
<td>Autophosphorylation</td>
<td>Yes</td>
<td>[42]</td>
</tr>
<tr>
<td>At3 g25840</td>
<td>Tyr782</td>
<td>Unknown</td>
<td>Yes</td>
<td>[42]</td>
</tr>
<tr>
<td>At2 g40120</td>
<td>Tyr425</td>
<td>Unknown</td>
<td>Yes</td>
<td>[42]</td>
</tr>
<tr>
<td>At5 g35980</td>
<td>Tyr284</td>
<td>Unknown</td>
<td>Yes</td>
<td>[40,42]</td>
</tr>
<tr>
<td>At7 g70520</td>
<td>Tyr178</td>
<td>Unknown</td>
<td>No</td>
<td>[42]</td>
</tr>
<tr>
<td>At2 g30940</td>
<td>Tyr181</td>
<td>Unknown</td>
<td>No</td>
<td>[42]</td>
</tr>
<tr>
<td>At3 g05140</td>
<td>Tyr154 and Tyr158</td>
<td>Unknown</td>
<td>No</td>
<td>[42]</td>
</tr>
</tbody>
</table>
Table 2. Arabidopsis thaliana proteins with predicted a SH2 domains

<table>
<thead>
<tr>
<th>AGI</th>
<th>Name</th>
<th>Predicted function</th>
<th>Region</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>At1g17040</td>
<td>Not determined</td>
<td>Unknown</td>
<td>545–631</td>
<td>[51]</td>
</tr>
<tr>
<td>At1g78540</td>
<td>Not determined</td>
<td>Unknown</td>
<td>572–661</td>
<td>[51]</td>
</tr>
<tr>
<td>At1g63210</td>
<td>Not determined</td>
<td>Related to yeast Spt6 protein. Chromatin remodelling, transcription initiation</td>
<td>1009–1102</td>
<td></td>
</tr>
<tr>
<td>At1g65440</td>
<td>GTB1 (GLOBAL TRANSCRIPTION FACTOR GROUP B1)</td>
<td>Related to yeast Spt6 protein. Chromatin remodelling, transcription initiation</td>
<td>1225–1318</td>
<td></td>
</tr>
</tbody>
</table>

a By the SMART database [52].

In agreement with previous analyses showing that plant splicing factors are highly phosphorylated on Arg-Pro, Ser-Pro and Thr-Pro kinase motifs [46] (Box 1), most of these non-pTyr sites are part of such motifs [42]. On the basis of these analyses, further experimental proof is required for the unambiguous assignment of these Tyr phosphorylation sites.

Conclusions and future directions

Although the observation that Tyr phosphorylation in plants is as extensive as it is in animals still needs confirmation, it is becoming clear that it occurs in plants more often than was previously assumed and that it is probably more abundant than in yeast [5]. MS-based determination of phosphorylation sites is a useful way in which to identify Tyr phosphorylation on a large scale. Because of the difficulties to assign the phosphosite by MS, as described earlier, it is important to confirm the suggested pTyr sites by independent techniques such as site-directed mutagenesis and immunoblotting with anti-pTyr antibodies. The next important challenge will be to identify the protein kinases and phosphatases that are responsible for targeting these Tyr sites (Box 3). Once we know the specificities of these kinases, prediction programs can be used to study the comprehensive set of Tyr-directed protein kinases and connect them to their substrates. Some of these substrates might bind to signalling proteins via an Src homology 2 (SH2) domain, which interacts specifically with pTyr-containing motifs [50]. In animals, an extensive network of SH2-based interactions exists. In A. thaliana, only two related proteins with a predicted SH2 domain, which is homologous to SH2 domains of animal STAT1 transcription factors, have been reported [51]. The current Simple Modular Architecture Research Tool (SMART) database (http://smart.embl-heidelberg.de) [52] predicts two additional proteins to have SH2 domains (Table 2). Although not predicted by SMART, regions of DELLA transcription factors also share homology with SH2 domains [53]. Based on these predictions, phosphotyrosine signalling in plants might be directly connected to transcriptional regulators (Box 3). Future studies should verify whether the predicted plant SH2 domains are pTyr-binding modules and what their interacting partners are. Collectively, further experiments should unravel Tyr kinase and phosphatase networks, including substrates, the position of each phosphorylation event in signalling cascades and their roles in the biology of plants.

Acknowledgements

We thank Adam Schikora, Jean Bigeard, Karin Zwerger, Jean Colcombet and Aladár Pettkó-Szandtner for their comments on the article. Work in our laboratory is supported by the Austrian Science Foundation (www.fwf.ac.at), the Vienna University (www.univie.ac.at), the Vienna Science and Technology Fund (www.wttf.at) and the European Union.

Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.tplants.2008.11.003.

14 Wan, L. (2007) Phosphorylation of the 12 S globulin cruciferin in wild-type and ab1-1 mutant A (thale cress) seeds. p. J. 404, 247–256


20 Sessa, G. (1996) PK12, a plant dual-specificity protein kinase of the LAMMER family, is regulated by the hormone ethylene. p. C 8, 2229–2234


22 Stulemeijer, I.J. (2007) Tomato mitogen-activated protein kinases LeMPK1, LeMPK2, and LeMPK3 are activated during the Cf-4/Avr4-induced hypersensitive response and have distinct phosphorylation specificities. p. J. 144, 1481–1494


