

Truncated variants of Ca²⁺-dependent protein kinases: A conserved regulatory mechanism?

Mackenzie E. W. Loranger¹, Alisa Huffaker² and Jacqueline Monaghan^{1,*}

¹ Department of Biology, Queen's University, Canada

² Department of Biological Sciences, University of California at San Diego, USA

* Correspondence: jacqueline.monaghan@queensu.ca (J. Monaghan)
Twitter: @jacquimonaghan

Abstract

Recent studies suggest that immune-induced alternative splice variants of the *Arabidopsis thaliana* Ca²⁺-dependent protein kinase (CDPK) AtCPK28 may result in signal attenuation. We put forward the hypothesis that expression of alternative truncated variants of this protein family may be a broadly conserved regulatory mechanism of CDPKs throughout the green lineage.

Alternative splicing in response to stress

In eukaryotes, most transcripts contain non-coding introns that are spliced out prior to translation. While canonical splicing removes all introns and joins all exons to produce a mature mRNA transcript, several forms of alternative splicing (AS) events also occur that can result in different translation products. With some variation between species, it is estimated that 60-70% of intron-containing transcripts can be alternatively spliced in plants, most of which contain retained introns that have the potential to result in premature stop codons due to frame shifts [1]. Non-canonical proteoforms can therefore lack sorting sequences or functional subdomains and can act in a dominant-negative or -positive manner to canonical isoforms, adding layers of complexity to protein diversity and regulation.

Alternative splicing is hypothesized to contribute to stress adaptation and developmental pathways [1]. Recent studies documented differential and transient

expression of AS variants in arabidopsis (*Arabidopsis thaliana*) seedlings 30 min after exposure to 1 μ M of the bacterial flagellin-derived peptide flg22 [2] or 1 μ M of the phytocytokine *AtPep1* [3], offering a snapshot into the dynamic post-transcriptional modifications that occur during an immune response.

Regulation of splicing is carried out by RNA-binding Ser-Arg (SR) proteins that recognize intron-exon junctions and recruit spliceosomal machinery to precursor mRNA [1]. The phosphorylation status of SR proteins could conceivably affect their function and allow for stimulus-responsive splicing modifications. Indeed, *AtPep1*-induced dephosphorylation of the SR protein IMMUNOREGULATORY RNA-BINDING PROTEIN (IRR) can disrupt association with, and canonical splicing of, transcripts encoding Ca^{2+} -dependent protein kinase (CDPK) *AtCPK28* [3], a key regulator of immune homeostasis [4]. Expression of *AtCPK28* AS variants was independently observed following treatment with flg22 [2], suggesting that alternative splicing of *AtCPK28* is part of the immune response.

CDPK activation: Twisting into shape

CDPKs are modular sensor-effector proteins that contain both a protein kinase domain ('effector') and a bilobal calmodulin (CaM)-like Ca^{2+} -binding domain ('sensor'), linked together by a hinge-like auto-inhibitory junction (AIJ) (**Figure 1A**). The current model for activation posits that Ca^{2+} -binding to the CaM-like domain (CaMD) results in a dramatic conformational twist that derepresses the kinase, allowing the CDPK to phosphorylate itself and/or its substrates (**Figure 1A**). In some cases, this conformational twist effectively pinches the kinase into an active state [5]. Each lobe of the CaMD typically contains a pair of EF hands that can bind Ca^{2+} with different affinities: while the C-terminal EF hands (C-EFs) tend to have high affinity for Ca^{2+} , those in the N-terminal lobe (N-EFs) have lower affinity and are more likely to act as the Ca^{2+} activator switch (**Figure 1A**) [5]. The threshold levels for Ca^{2+} -activation differ across CDPKs, which may reflect their ability to decode specific changes in cellular Ca^{2+} dynamics [5]. While some CDPKs require an increase in cytosolic $[\text{Ca}^{2+}]$, others are

active at basal $[Ca^{2+}]$ and/or can contain degenerate EF hands [5], pointing towards involvement of additional regulatory mechanisms such as phosphorylation or binding partners. Remarkably, Ca^{2+} -dependency can be uncoupled from kinase activity through the deliberate removal of the AIJ and CaMD (together called the Ca^{2+} -activation domain; CAD), often resulting in deregulated/uninhibited CDPKs that have been experimentally useful (**Figure 1A**) [5].

Interestingly, one of the immune-induced variants of *AtCPK28* (*At5g66210.4/CPK28-RI*) results in a truncated isoform that lacks both high-affinity C-EFs in the CaMD [2,3]. However, instead of derepressing the active site, this variant compromises catalytic activity and signaling competency [3]. As it has been postulated that Ca^{2+} -bound C-EFs might be critical for enhancing Ca^{2+} -binding affinity to the N-EFs [7], it is possible that CPK28.4/RI is unable to bind Ca^{2+} , although this has not been formally tested. Furthermore, as interactions between Ca^{2+} -bound C- and N-EFs are important for stabilization of the holoenzyme [5,7], it is possible that CPK28.4/RI prevents the conformational change that is necessary for relief of inhibition (**Figure 1A**).

Alternative CDPK isoforms are expressed throughout the plant kingdom

As part of a multi-protein regulatory module, *AtCPK28* serves an important role in buffering the signaling potential of key immune kinase BOTRYTIS INDUCED KINASE 1 (BIK1) [4]. It has been suggested that the transient suppression of *AtCPK28* through expression of dominant negative truncated variants could temporarily allow a full immune response to be mounted [3]. Similar AS variants of *AtCPK32* and *OsCPK17* have been identified [8,9], and we found that non-canonical C-terminal truncated isoforms have been curated for five additional CDPKs in The Arabidopsis Information Resource database (<http://arabidopsis.org>): *AtCPK13* (*At3g51850.2* and *.3*), *AtCPK15* (*At4g21940.3*), *AtCPK23* (*At4g04740.2*), *AtCPK29* (*At1g76040.3*), and *AtCPK33* (*At1g50700.3*). This led us to hypothesize that expression of AS variants with truncated sensor domains may be a conserved regulatory mechanism of this protein family across

the plant lineage. Intrigued, we mined the transcriptomes of 224 species with unrestricted data curated in Phytozome v13 (<https://phytozome-next.jgi.doe.gov/>) and the Plant Transcription & Protein Kinase Identifier and Classifier (iTAK) database (<http://itak.feilab.net/cgi-bin/itak/index.cgi>) for AS variants within the CDPK family (Eukaryotic Orthologous Group KOG0032) and then queried their predicted protein sequences for the presence and number of EF hands using a custom Python script based on a 13 amino acid consensus sequence (**File S1**) [10]. Of the 10,622 sequences we analyzed, we identified alternative C-terminal CaMD truncations of 176 CDPKs in 57 angiosperm species (**Figures 1B, S1**). To determine the type of AS event that resulted in the truncated isoforms, we aligned the gene and transcript sequences in ClustalW and found that 87% were due to retained introns (**Figure S1**), in line with this being the prevalent type of AS event in plants [1]. Importantly, as the experimentally validated variant of *OsCPK17* [9] was not curated in either Phytozome v13 or the iTAK database, we predict that the AS variants we recovered may be an under-representation due to limited curation or transcript sequence information available for some species. Accumulation of alternatively spliced *AtCPK28* transcripts upon *AtPep1* treatment is transient and dissipates within 4 hours [3]; similarly ephemeral changes in abundance are possible for other AS events and may not be captured by available datasets.

Found throughout the Plantae and Protista kingdoms, CDPKs are thought to have arisen by the fusion of CaM with CaM kinase (CaMK) [11], followed by independent gene duplication events and the diversification into Kingdom-specific subgroups [11]. In land plants, *AtCPK28* belongs to subgroup IV, which is hypothesized to represent the earliest divergence point from the last common ancestor with the protist CDPKs [11]. We found C-terminal truncated variants in all subgroups including 24 group IV CDPKs from 18 species (**Figure 1B**). This includes *OsCPK18*, which, similar to *AtCPK28*, negatively regulates immune signaling in rice [12]. It will be interesting to assess the expression dynamics of *OsCPK18* and other group IV truncated variants during an immune response. Interestingly, we retrieved several isoforms for some CDPKs, including predicted truncations of 1, 2, 3, or all 4 EFs, as well as isoforms lacking the entire CAD (**Figures 1B, S1**). Although their functional role is currently unknown, we hypothesize that expression of variants lacking part of the CAD could

create dominant-negative CDPK proteoforms as proposed for *AtCPK28.4/RI*, while those lacking most of the CAD may be uninhibited/deregulated as suggested by experimental observations from several studies using full CAD-truncated variants. This could provide a post-transcriptional fine-tuning mechanism for regulation of this protein family.

In parasitic alveolates such as malaria-causing *Plasmodium* species, CDPKs serve important virulence roles and represent valuable drug targets as they are not found in their animal hosts [13]. Interested to see if similar truncated variants exist outside the green lineage, we retrieved all CDPK isoform sequences from 112 Apicomplexan species curated in the Eukaryotic Pathogen, Vector and Host Informatics Resources database (<https://VEuPathDB.org/>). Notably, Apicomplexan CDPKs are quite divergent from those in the green lineage, including several with shuffled domain arrangements and 2, 3, or 4 EF hands [13]. Interestingly, of the 596 CDPKs we analyzed (defined as those containing both PFAM domains PF00069 and PF13499), only single isoforms were identified. Although it is possible that AS variants are not adequately curated, this could suggest that expression of truncated CDPK isoforms may be limited to land plants. Although we do not yet understand the expression patterns or precise functional roles of these variants, we propose that expression of alternative CDPK isoforms truncated in the Ca^{2+} sensing domain might be a broadly used regulatory mechanism of this protein family in plants.

Concluding remarks

Alternative mRNA splicing has the potential to dynamically alter the structure and function of protein isoforms, conceivably allowing for rapid adjustments to signaling pathways. While there is evidence that alternative splice forms of CDPKs are generated *in vivo* and that the resulting truncated proteins have modified function, technical challenges remain. Moving beyond profiling the abundance of splice forms to definitively determining the extent to which they contribute to signaling modification

in their natural context is an exciting path forward that will better elucidate the role of this regulatory mechanism in CDPK function.

Acknowledgements

We thank Melissa Bredow (Queen's University) and two anonymous reviewers for carefully and critically reading our manuscript, and all members of the Monaghan lab for engaging discussions. We thank David Goodstein (Joint Genome Institute) for assistance with data retrieval from Phytozome v13, and George diCenzo (Queen's University) for kindly reviewing our Python script. Work in the AH laboratory is funded by NSF CAREER Award no. 1943591, and work in the JM laboratory is funded by the Natural Sciences and Engineering Research Council of Canada Discovery Program, the Canada Research Chairs Program, the Canadian Foundation for Innovation John R Evans Leaders Fund, the Ontario Ministry of Research, Innovation and Science, and Queen's University. We apologize to colleagues whose work we could not cite due to space limitations.

References

- 1 Rigo, R. et al. (2019) Alternative Splicing in the Regulation of Plant-Microbe Interactions. *Plant Cell Physiol.* 60, 1906–1916
- 2 Bazin, J. et al. (2020) Role of MPK4 in pathogen-associated molecular pattern-triggered alternative splicing in Arabidopsis. *PLoS Pathog.* 16, e1008401
- 3 Dressano, K. et al. (2020) Dynamic regulation of Pep-induced immunity through post-translational control of defence transcript splicing. *Nat Plants* 6, 1008–1019
- 4 Wang, J. et al. (2018) A Regulatory Module Controlling Homeostasis of a Plant Immune Kinase. *Mol. Cell* 69, 493–504.e6
- 5 Yip Delormel, T. and Boudsocq, M. (2019) Properties and functions of calcium-dependent protein kinases and their relatives in Arabidopsis thaliana. *New Phytol.* 224, 585–604
- 6 Matschi, S. et al. (2013) Function of calcium-dependent protein kinase CPK28 of Arabidopsis thaliana in plant stem elongation and vascular development. *Plant J.* 73, 883–896
- 7 Liese, A. and Romeis, T. (2013) Biochemical regulation of in vivo function of plant calcium-dependent protein kinases (CDPK). *Biochim. Biophys. Acta* 1833, 1582–1589
- 8 Ding, F. et al. (2014) Genome-wide analysis of alternative splicing of pre-mRNA under salt stress in Arabidopsis. *BMC Genomics* 15, 431
- 9 Almadanim, M.C. et al. (2018) The rice cold-responsive calcium-dependent protein kinase OsCPK17 is regulated by alternative splicing and post-translational modifications. *Biochim. Biophys. Acta Mol. Cell Res.* 1865, 231–246
- 10 Mohanta, T.K. et al. (2015) Genome-Wide Identification of Calcium Dependent Protein Kinase Gene Family in Plant Lineage Shows Presence of Novel D-x-D and D-E-L Motifs in EF-Hand Domain. *Front. Plant Sci.* 6, 1146
- 11 Valmonte, G.R. et al. (2014) Calcium-dependent protein kinases in plants: evolution, expression and function. *Plant Cell Physiol.* 55, 551–569
- 12 Xie, K. et al. (2014) Direct phosphorylation and activation of a mitogen-activated protein kinase by a calcium-dependent protein kinase in rice. *Plant Cell* 26, 3077–3089
- 13 Billker, O. et al. (2009) Calcium-dependent signaling and kinases in apicomplexan parasites. *Cell Host Microbe* 5, 612–622

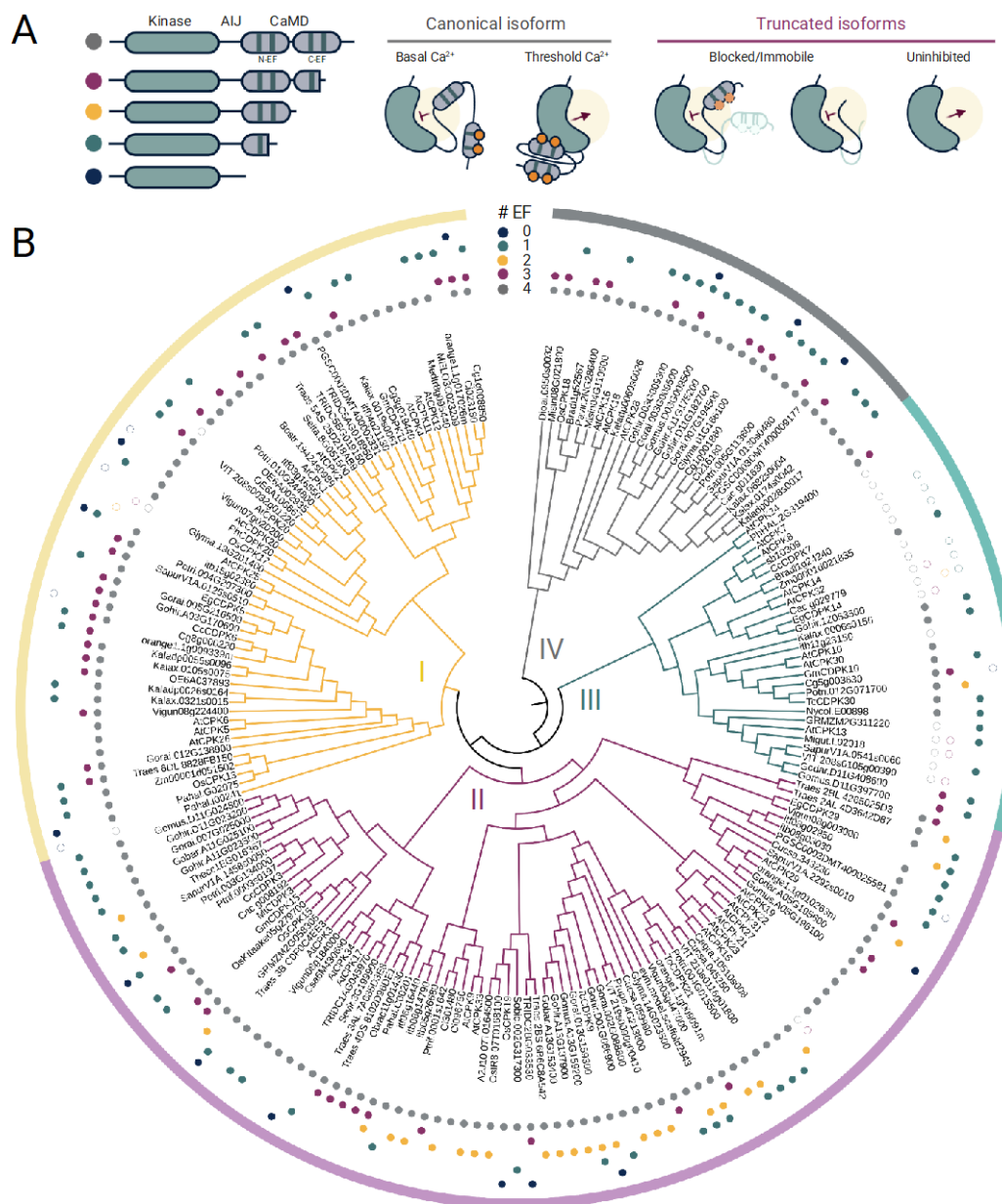


Figure 1. Truncated variants of CDPKs across the plant lineage.

(A) Proposed activation mechanism of canonical and truncated CDPK variants. In canonical variants, Ca^{2+} ions (orange circles) binding to both C- and N-EFs results in a conformational twist that relieves repression of the active site. In truncated variants lacking part of the CaMD we propose that remaining EFs may not be able to bind Ca^{2+} (shown by light orange circles with a dashed outline), and/or may not be able to undergo the conformational twist in the absence of stabilizing C-EFs. This has the effect of rendering the truncated variants 'immobile', which is likely also true for variants lacking the entire CaMD, resulting in the active site remaining blocked. Variants lacking the CaMD and the autoinhibitory junction (AIJ) might be uninhibited or deregulated, as observed by the deliberate removal of the entire Ca^{2+} -activation domain (CAD) in several published reports [5,7]. (B) Phylogenetic tree of CDPKs identified in this study, anchored into subgroups I, II, III, or IV using the entire *AtCDPK* family. Full-length sequences of the longest variant (typically CV.1) were aligned using ClustalW in MEGAX. A Neighbour-joining tree supported by 1000 bootstraps was generated in MEGAX and visualized in iTOL. Truncated variants are colour-coded as indicated; note that more than one truncated variant is expressed for several CDPKs. For variants with 1, 2, 3, or 4 EF hands, open circles indicate the presence of at least one degenerate motif, whereas open circles for variants with 0 EF hands indicate a missing or modified AIJ domain (please see online Supplemental Information Figure S1 for more information). Following standardized nomenclature, CDPKs that have been named already are presented using that identifier. Otherwise, gene identifiers are displayed based on their curated names in Phytozome v13 or iTAK. Gene ID, alternate names, species information, and the type of AS event that resulted in these proteoforms are provided in Figure S1. As the truncated *OsCPK17* variant reported by [9] is not curated in online databases, this variant was added manually. This figure was created using BioRender.

1 Supplemental Information

2 Figure S1. Alternative splicing events predicted to produce truncated CDPK splice variants.

3 Gene identifiers (ID) were provided by Phytozome v13 or iTAK. The number of EF hands identified in
4 canonical variants (CV) is indicated; EF hands predicted to be degenerate are marked with parentheses.
5 Alternative transcript variants (TV) are also listed, alongside the type of alternative splicing event as
6 deduced from genomic and cDNA sequence alignments; retained intron (RI), skipped exon (ES),
7 alternative acceptor site (AAS), or alternative donor site (ADS). Transcript variants with 0 EF hands that
8 contain a full autoinhibitory junction (AIJ), are indicated in parentheses as (+), while those with a
9 modified or partial AIJ are indicated as (~), and those missing the entire AIJ are indicated as (-). All
10 CDPKs are organized based on their phylogenetic placement into subgroups I, II, III, and IV as
11 determined using the same parameters in **Figure 1B**.

12 File S1. Python script used to analyze CDPK proteoforms.

13 This file contains the script we generated to identify the presence and number of EF hands in the
14 proteoforms we retrieved from genomics databases.