Short title: Design principle for decoding calcium signals

Corresponding author:

Marc R. Knight

Department of Biosciences,

Durham University,

South Road,

Durham DH1 3LE, UK

Email: m.r.knight@durham.ac.uk

Tel: +44 191 334 1224
Article title: Design principles for decoding calcium signals to generate specific gene expression via transcription

Junli Liu¹*, Gioia Lenzoni² & Marc R. Knight¹*

1. Department of Biosciences, Durham University, South Road, Durham DH1 3LE, UK
2. School of Pharmaceutical Sciences, University of Geneva, Rue Michel-Servet 1, Geneva CH-1211, Switzerland

*Joint corresponding authors (m.r.knight@durham.ac.uk and junli.liu@durham.ac.uk)

One sentence summary: Identification of the design principle by which plant cells decode specific calcium signatures to produce the correct gene expression response.

Footnotes

Author contributions: M.R.K. and J.L. conceived the project and original research plans; M.R.K. supervised the experiments; G.L. performed most of the experiments; J.L. performed the modelling analysis; M.R.K. and J.L. wrote the article with contributions of all the authors; M.R.K. agrees to serve as the author responsible for contact and ensures communication.

Funding: This study was supported by the EU Marie Curie project CALIPSO [GA 2013-607607 to G.L.].

Email address of Author for Contact: m.r.knight@durham.ac.uk
ABSTRACT

The second messenger calcium plays a key role in conveying specificity of signalling pathways in plant cells. Specific calcium signatures are decoded to generate correct gene expression responses and amplification of calcium signatures is vital to this process. It is not known: (1) if this amplification is an intrinsic property of all calcium-regulated gene expression responses and whether all calcium signatures have the potential to be amplified, and (2) how does a given calcium signature maintain specificity in cells containing a great number of transcription factors (TFs) and other proteins with the potential to be calcium-regulated? The work presented here uncovers the design principle by which it is possible to decode calcium signals into specific changes in gene transcription in plant cells. Regarding the first question, we found that the binding mechanism between protein components possesses an intrinsic property that will nonlinearly amplify any calcium signal. This nonlinear amplification allows plant cells to effectively distinguish the kinetics of different calcium signatures to produce specific and appropriate changes in gene expression. Regarding the second question, we found that the large number of calmodulin (CaM)-binding transcription factors (TFs) or proteins in plant cells form a buffering system such that the concentration of an active CaM-binding TF is insensitive to the concentration of any other CaM-binding protein, thus maintaining specificity. The design principle revealed by this work can be used to explain how any CaM-binding TF decodes calcium signals to generate specific gene expression responses in plant cells via transcription.
Plants are sessile organisms and therefore they must adapt their metabolism, growth, and architecture to a changing environment. To survive, it is vital for plants to be able to sense and act upon environmental information. Central to this are “second messengers”: cellular chemicals that convey information from the outside world to the cells that make up a plant. Second messengers have evolved to trigger the required response of cells to environmental cues. Calcium is a ubiquitous second messenger for activating tolerance mechanisms in plants responding to environmental stresses (McAinsh et al., 1995; Allen et al., 2001; Love et al., 2004; Miwa et al., 2006; McAinsh and Pittman, 2009; Dodd et al., 2010; Short et al., 2012; Edel et al., 2017; Yuan et al., 2017; Bender et al., 2018; Kudla et al., 2018).

The majority of plant defence responses against stress is realised by changes in gene expression in order to produce proteins required to combat the conditions they encounter. It is thus vital that the correct proteins are produced in response to different environmental conditions, i.e. different genes need to be switched on in response to different stimuli. This means that the identity of the primary stimulus must be encoded in a “language” that the cell can understand. Most stimuli lead to transient elevation in cellular calcium levels. Importantly, different stimuli produce calcium elevations with different characteristics: a unique “calcium signature”. Consequently, the specific properties of different calcium signatures have been proposed to encode information on the identity of the stimulus (McAinsh et al., 1995; Allen et al., 2001; Love et al., 2004; Miwa et al., 2006; McAinsh and Pittman, 2009; Dodd et al., 2010; Short et al., 2012).

Experimental data showed that calcium signals can be decoded to generate specific gene expression responses (Whalley et al., 2011; Whalley and Knight, 2013) and modelling analysis revealed that amplification of calcium signals is important for decoding calcium signals (Liu et al., 2015; Lenzoni et al., 2018). However, it remains unclear whether or not decoding calcium signatures in plant cells is governed by any general principle.

The complexity for plant cells to decode specific calcium signatures is multifaceted. First, any, even a modest, calcium signature (e.g. in response to ozone (Clayton et al., 1999)) is able to induce gene expression. Second, the specific characteristics of the calcium signatures produced by different stresses encode stimulus-specific information. Experimental evidence demonstrates that Arabidopsis (*Arabidopsis thaliana*) is able to decode specific calcium signatures and interpret them; leading to distinct gene expression profiles (Whalley et al., 2011; Whalley and Knight, 2013). Third, a variety of experimental data show that there
are a large number of calmodulin (CaM)-binding proteins (Reddy et al., 2011; Poovaiah et al., 2013; Virdi et al., 2015; Edel et al., 2017; Yuan et al., 2017; Bender et al., 2018; Kudla et al., 2018). CaM has two pairs of Ca$^{2+}$-binding EF-hand domains located at the N- and C-termini, respectively (Finn and Forsen, 1995; Valeyev et al., 2008). Some transcription factors (TFs) can bind to Ca$^{2+}$-CaM, allowing them to respond to calcium signals via this Ca$^{2+}$–CaM–TF interaction. Clearly, for transcription factors to decode calcium signals, and therefore generate specific gene expression responses, they must be able to distinguish the kinetics of different calcium signals in the context of competing for binding CaM with other CaM-binding proteins. However, how this occurs is unknown. In general, the search for basic underlying principles is vital for a better understanding of the regulation of signalling dynamics. Cells navigate environments, communicate and build complex patterns by initiating specific gene expression responses to specific signals (Brophy and Voigt, 2014). Studies in other cellular systems (Savageau, 2001; Salvador and Savageau, 2003; Wall et al., 2003; Salvador and Savageau, 2006; Purvis and Lahav, 2013; Tolla et al., 2015; Karin et al., 2016) have found that biological networks may be evolutionarily tuned and regulatory architecture of a biological network is optimised following some basic principles underlying evolutionary selection (Salvador and Savageau, 2003; Chubukov et al., 2012). Design principles are the underlying properties of network structures that have evolved to endow the network functions. Although experimental data showed that calcium signals can be decoded to generate specific gene expression responses (Whalley et al., 2011; Whalley and Knight, 2013) and modelling analysis revealed that amplification of calcium signals is important for decoding calcium signals (Liu et al., 2015; Lenzoni et al., 2018), it remains unclear whether or not decoding calcium signatures in plant cells is governed by any general principle. This work uncovers the design principle for decoding calcium signals through changes in transcription by addressing the following two questions: (1) is amplification of Ca$^{2+}$ signatures an intrinsic property of all calcium-regulated gene expression responses and do all calcium signatures have the potential to be amplified? And (2) how does decoding of calcium signals maintain specificity when one messenger (Ca$^{2+}$) is decoded by many transcription factors and proteins in plant cells? This work establishes the link between the characteristics of CaM (i.e., it has two pairs of Ca$^{2+}$-binding EF-hand domains and it is capable of binding a large number of proteins in plant cells) with the intrinsic properties of Ca$^{2+}$–CaM–TF interactions, to reveal the design principle underpinning how plant cells decode calcium signals to generate specific gene expression response via changes in transcription. We show that both a theoretical simple gene expression system and an empirical system of two plant
immunity genes (*enhanced disease susceptibility 1* (EDS1) and *isochorismate synthase 1* (ICS1)) (Lenzoni et al., 2018) follow this design principle to decode calcium signatures. The principle revealed in this work is applicable to study how any CaM-binding TF decodes calcium signals to generate specific gene expression response in plant cells via transcription.
RESULTS

Amplification of calcium signal is an intrinsic property of Ca\textsuperscript{2+}-CaM-TF interactions

The interaction of Ca\textsuperscript{2+}, CaM, and another (CaM-binding) protein can form many different binding complexes. CaM has two pairs of Ca\textsuperscript{2+}-binding EF-hand domains located at the N- and C-termini, respectively (Finn and Forsen, 1995; Valeyev et al., 2008). Experimental measurement showed that 4Ca\textsuperscript{2+}-CaM is the active CaM-Ca\textsuperscript{2+} binding complex (Pifl et al., 1984). Therefore, this work assumes that the 4Ca\textsuperscript{2+}-CaM-TF complex is the active complex for gene expression responses. The cooperative binding between Ca\textsuperscript{2+} and the 4 binding sites of CaM has previously been subjected to both experimental and modelling studies (Fajmut et al., 2005; Shifman et al., 2006; Pepke et al., 2010; Liu et al., 2015) and the kinetic parameters have been experimentally determined (Shifman et al., 2006; Pepke et al., 2010).

For any transcription factor with one CaM-binding site, 18 different binding complexes can form via 33 elementary binding processes (Supplemental Information). For example, the binding between Ca\textsuperscript{2+}, CaM, and calmodulin-binding transcription activators (CAMTAs), and the binding between Ca\textsuperscript{2+}, CaM and calmodulin binding protein 60g (CBP60g) have been previously described in detail (Liu et al., 2015; Lenzoni et al., 2018). For any CaM-binding transcription factor, and following the previous analysis (Liu et al., 2015), there are six adjustable parameters for fully examining the dynamics of Ca\textsuperscript{2+}-CaM-TF interactions after using the experimentally-determined parameters and introducing basic thermodynamic constraints. P describes the cooperative binding between CaM and a TF in the presence of Ca\textsuperscript{2+}. $P>1$ or $P<1$ represents the binding affinity of Ca\textsuperscript{2+}-CaM complex to the transcription factor being looser than, the same as, or tighter than binding of free CaM to the TF, respectively. $K_{14}$ is the dissociation equilibrium constant for the binding of the Ca\textsuperscript{2+}-CaM complex to the TF. $k_{on(14)}$ is the on rate for the binding of Ca\textsuperscript{2+}-CaM complex to the TF; Q describes how the cooperative binding between CaM and the TF in the presence of Ca\textsuperscript{2+} is realised by $k_{on}$, $k_{off}$ or both. $[\text{CaM}_t]$ describes the total concentration of CaM, which is the summation of free CaM and all CaM complexes. $[\text{TF}_t]$ describes the total concentration of the TF, which is the summation of free TF and all TF complexes.

Here we consider that a “quasi-equilibrium state” is established for Ca\textsuperscript{2+}-CaM-TF interactions according to the detailed balance principle (Alberty, 2004). Establishing a quasi-equilibrium state requires the “on” and “off” rates for all binding reactions of Ca\textsuperscript{2+}–CaM–TF interactions are relatively fast so that each reaction can establish an equilibrium. In the
sections “Case study 1: a simple Ca\(^{2+}\)-regulated gene expression system” and “Case study 2: plant immunity gene expression”, we will show that this assumption is valid for experimentally measured parameters of Ca\(^{2+}\)-CaM-TF interactions. At a quasi-equilibrium state, \(k_{\text{on}(14)}\) and \(Q\) become irrelevant. Thus, there are only four adjustable parameters (i.e., \(P\), \(K_{14}\), \([CaM_i]\), \([TF_i]\)) for examining the dynamics of Ca\(^{2+}\)-CaM-TF interactions.

At a quasi-equilibrium state and for any calcium concentration, the concentration of each Ca\(^{2+}\)-CaM-TF complex can be analytically derived (Supplemental Information). Equation 1 shows the concentration of the active complex, 4Ca\(^{2+}\)-CaM-TF.

\[
[4\text{Ca}^{2+}\text{-CaM-TF}] = \frac{[CaM_i][\text{Ca}^{2+}]^4[\text{TF}]}{k_{14}k_1k_2k_3k_4 + (1 + gss + g + P + K_{14})[\text{TF}][\text{Ca}^{2+}]^{5/4} + (1 + g)_{ss}[\text{TF}][\text{Ca}^{2+}]^{5/4}}
\]

(eq 1)

\(K_1\), \(K_2\), \(K_3\), and \(K_4\) are the dissociation equilibrium constants for binding of first and second Ca\(^{2+}\) to the CaM C-terminus; and for binding of first and second Ca\(^{2+}\) to the CaM N-terminus, respectively. \([\text{TF}]\) is the concentration of the free transcription factor, and it can be calculated using the total concentration of the transcription factor and the concentrations of all CaM-TF complexes. Other symbols in equation 1 are as described above.

At an unperturbed cellular state where a calcium signature has not yet emerged, the calcium concentration settles onto a steady-state value, \([\text{Ca}^{2+}]_{ss}\). In this state, expression of a gene, which is regulated by the active signal, 4Ca\(^{2+}\)-CaM-TF, is at a fixed level, corresponding to \([\text{Ca}^{2+}]_{ss}\). Kinetics of different calcium signatures have different temporally changing features of calcium concentration. Due to the innate properties of the Ca\(^{2+}\)-CaM-TF interactions, different calcium signatures are decoded into different temporally-changing concentrations of the active complex, 4Ca\(^{2+}\)-CaM-TF, which in turn regulates gene expression. Thus, the first step for elucidating the information flow from a calcium signal to a specific gene expression response is to examine how the signal is decoded into a temporally-changing concentration of the active signal, \([4\text{Ca}^{2+}\text{-CaM-TF}]\).

Equation 2 describes the ratio of \([4\text{Ca}^{2+}\text{-CaM-TF}]\) for any \([\text{Ca}^{2+}]\) to that for \([\text{Ca}^{2+}]_{ss}\).

\[
\frac{[4\text{Ca}^{2+}\text{-CaM-TF}]}{[4\text{Ca}^{2+}\text{-CaM-TF}]_{ss}} = \left(\frac{[\text{Ca}^{2+}]}{[\text{Ca}^{2+}]_{ss}}\right)^4 f = \left(\frac{[\text{Ca}^{2+}]}{[\text{Ca}^{2+}]_{ss}}\right)^4 \frac{g_{ss} + P + K_{14} [\text{TF}]_{ss}(1 + g_{ss})}{g + P + K_{14} [\text{TF}]_{ss}(1 + g)}
\]

(eq 2)

With
interaction to amplify a calcium signal, we analysed an example, for which [Ca$^{2+}$]$_{ss}$ (Linse et al., 1991; Shifman et al., 2006; Kubota et al., 2007; Pepke et al., 2010). To show the ability of Ca$^{2+}$ to increase total concentration of the TF, we are able to deduce that the lower limit of [Ca$^{2+}$] increases. This is because increasing [Ca$^{2+}$] increases term $g$ and simultaneously decreases the concentration of free transcription factor, [TF], for a constant total concentration of the transcription factor, [TF]. Second, the value of $f$ is dependent on both [TF]$_{ss}$ and [TF], both of which increase with [TF]. In plant cells, a typical calcium signature can increase cytosolic calcium concentration from its steady state concentration (ca. 0.05 μM) to up to 2.5 μM with contrasting different kinetics (Knight et al., 1996, 1997; Aslam et al., 2008). If we consider that, within this range of [Ca$^{2+}$], the free TF concentration is only determined by the total concentration of the TF, we are able to deduce that the lower limit of $f$ is $\frac{g_{ss}}{g}$ (Supplemental Information), namely $f$ is always larger than $\frac{g_{ss}}{g}$. Thus, when [Ca$^{2+}$] increases from [Ca$^{2+}$]$_{ss}$ to [Ca$^{2+}$], the minimum amplification of the calcium signal into the active signal, $\frac{[4Ca^{2+}-CaM-TF]}{[4Ca^{2+}-CaM-TF]_{ss}}$, is $\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4 \frac{g_{ss}}{g}$.

To determine the values of $\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4 \frac{g_{ss}}{g}$, we need the values of the four parameters ($K_1$, $K_2$, $K_3$ and $K_4$). These four parameters have been experimentally determined and their values are $K_1 = 10 \, \mu M$, $K_2 = 0.925 \, \mu M$, $K_3 = 25 \, \mu M$, $K_4 = 5 \, \mu M$ (Linse et al., 1991; Shifman et al., 2006; Kubota et al., 2007; Pepke et al., 2010). To show the ability of Ca$^{2+}$-CaM-TF interaction to amplify a calcium signal, we analysed an example, for which [Ca$^{2+}$] increases to 2.5 μM from its steady-state value of 0.05 μM. For [Ca$^{2+}$]$_{ss} = 0.05 \, \mu M$, $g([Ca^{2+}]_{ss}, K) = 0.0073$. When [Ca$^{2+}$] increases to 0.25 μM (i.e. 5 fold), 0.5 μM (i.e. 10 fold), 1.0 μM (i.e. 20
fold), and 2.5 µM (i.e. 50 fold), the minimum amplification of these calcium concentrations into the concentrations of their active signals, $\frac{[4Ca^{2+} - CaM - TF]}{[4Ca^{2+} - CaM - TF]_{ss}}$, is 107 fold, 725 fold, 4390 fold, and 37570 fold, respectively. Thus, the Ca$^{2+}$-CaM-TF interaction possesses an intrinsic property of nonlinearly amplifying any calcium signal, which is quantitatively described by equation 2.

Fig. 1 shows the numerical results that confirm the above theoretical analysis for a wide range of total concentrations of a transcription factor (0.01 µM to 1.0E5 µM). Fig. 1A shows that $f$ is always less than 1, that increasing [Ca$^{2+}$] decreases the value of $f$, and that $f$ is always larger than $\frac{g_{ss}}{g}$. Fig. 1B shows that the term $f$ is relatively unimportant and any calcium signal is always amplified. Therefore, numerical analysis confirms theoretical analysis: a calcium signal is always amplified due to Ca$^{2+}$-CaM-TF interaction. In addition, following the derivation of the lower limit of $f$ in the Supplemental Information, we know that decreasing the parameter for the cooperative binding between CaM and a TF in the presence of Ca$^{2+}$, P (Liu et al., 2015; Lenzoni et al., 2018), or increasing the concentration of the TF decreases the value of $f$. Fig. 1, A-D shows that numerical results are in agreement with theoretical analysis. When P is sufficiently small and [TF]$_t$ is sufficiently large, fold amplification of calcium signal is the same as the theoretical minimal fold amplification (Fig. 1, B and D).

Based on the above analysis, it can be seen that any calcium signal is always amplified by the power of 4 of calcium concentration ratio, $\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4$, multiplied by a factor that is relatively less important, $\frac{g_{ss}}{g}$. In equation 2, the main factor for amplifying a calcium signal is the term $\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4$. Derivation of equation 2 reveals that the term $\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4$ emerges from two pairs of Ca$^{2+}$-binding EF-hand domains and a TF-binding domain in the CaM structure. Therefore, the ability of CaM to bind four Ca$^{2+}$ and one TF molecule results in the amplification of calcium signal. In other words, the Ca$^{2+}$-CaM-TF binding mechanism naturally leads to amplification of calcium signals. As we will show below, this aspect of the underlying design principle, which links the Ca$^{2+}$-CaM-TF binding mechanism with the emerging property of calcium signal amplification, leads to preferential expression of specific calcium-regulated genes.
Concentration of a CaM-binding TF-specific active signal is insensitive to changes in concentration of other CaM-binding proteins

A variety of experimental data show that there are a large number of CaM-binding proteins in plant cells (Reddy et al., 2011; Poovaiah et al., 2013; Virdi et al., 2015). Thus, when a calcium signature emerges, a specific TF must compete for the binding of CaM with other CaM-binding proteins. An important question, therefore, is how is a CaM-binding TF capable of generating a specific gene expression response by decoding a calcium signature in the context of competing for CaM binding with many other CaM-binding proteins? To address this question, the effects of the existence of a large number of CaM-binding proteins on the concentration of a CaM-binding TF-specific active signal must be examined.

When many proteins compete for the binding of CaM, the concentration of the active complex of a transcription factor \( (\text{TF}_i) \), \( 4\text{Ca}^{2+}-\text{CaM-}\text{TF}_i \), can be derived following the method used for deriving eq. 1 and is described by equation 3.

\[
[4\text{Ca}^{2+} - \text{CaM} - \text{TF}_i] = \frac{[\text{CaM}_j][\text{Ca}^{2+}]^4[\text{TF}_j]}{K_1K_2K_3K_4 + (1 + \sum_{j=1}^n P_j[\text{TF}_j])K_{j,14}} + \frac{[\text{Ca}^{2+}]K_1[\text{Ca}^{2+}]K_2[\text{Ca}^{2+}]K_3[\text{Ca}^{2+}]K_4}{K_1K_2K_3K_4 + (1 + \sum_{j=1}^n P_j[\text{TF}_j])K_{j,14}}
\]

(eq. 3),

where \( \text{TF}_i \) and \( \text{TF}_j \) are the free form of the \( i^{th} \) and \( j^{th} \) transcription factor, respectively. \( P_j \) is the parameter for quantifying the cooperative binding between CaM and \( \text{TF}_j \) in the presence of \( \text{Ca}^{2+} \). \( K_{i,14} \) and \( K_{j,14} \) are the dissociation equilibrium constants for the binding of \( \text{Ca}^{2+}-\text{CaM} \) complex to \( \text{TF}_i \) and \( \text{TF}_j \), respectively. \( n \) is the total number of CaM-binding proteins. In eq. 3, for simplifying notations, \( \text{TF}_j \) can be any TF or protein that binds with CaM. Therefore, following eq. 3, the existence of any CaM-binding TF or protein, \( \text{TF}_i \), could affect the concentration of the active complex of a transcription factor \( \text{TF}_i, 4\text{Ca}^{2+}-\text{CaM-}\text{TF}_i \), by competing for the binding of CaM with \( \text{TF}_i \). In eq. 3, this competition is described by the two summation terms: \( \sum_{j=1}^n \frac{P_j[\text{TF}_j]}{K_{j,14}} \) and \( \sum_{j=1}^n \frac{[\text{TF}_j]}{K_{j,14}} \).

We consider that the total concentration of CaM is \( [\text{CaM}]_t \) and the concentration of each of the CaM-binding proteins, \( \text{TF}_j \), is \( [\text{TF}_j]_t \) (\( j=1,\ldots,n \)). When many proteins compete for the binding of CaM, the following constraints must apply. The concentration summation of free CaM and all CaM complexes with different proteins must be equal to \( [\text{CaM}]_t \) at any calcium concentration. The total concentration for any transcription factor is the...
concentration summation of free protein, TF$_j$, and all TF$_j$-binding complexes [TF$_j$]$_t$ at any
calcium concentration.

Examination of eq. 3 reveals that the existence of a large number of CaM-binding
proteins in plant cells can form a buffering system such that the concentration of a CaM-
binding TF-specific active signal is insensitive to change in the concentration of another
CaM-binding protein. This is because of the two summation terms, $\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$ and
$\sum_{j=1}^{n} \frac{[TF_j]}{K_{j,14}}$, in the denominator of eq. 3. Firstly, the existence of any CaM-binding protein,
TF$_j$, always reduces the concentration of the active complex of transcription factor, TF$_i$. This
is because when CaM binds with TF$_j$, the concentration of CaM available for binding with
TF$_i$ will become smaller. In eq.3, this corresponds to $\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$ and $\sum_{j=1}^{n} \frac{[TF_j]}{K_{j,14}}$ always being
larger than $\frac{P_i[TF_i]}{K_{i,14}}$ and $\frac{[TF_i]}{K_{i,14}}$, respectively. Secondly, since each of the two terms is the
summation of the contribution of all CaM-binding proteins, the TF or protein that contributes
a larger value of $\frac{P_j[TF_j]}{K_{j,14}}$ and $\frac{[TF_j]}{K_{j,14}}$ is quantitatively more important. Thirdly, as the number of
CaM-binding proteins increases, the contribution of each protein to both summation terms
becomes less important. When there are only relatively few CaM-binding proteins, changing
the concentration of one can change the value of both terms to a relatively large extent.
However, if there are many CaM-binding proteins, changing the concentration of one will
change the value of both terms to a much lesser extent. For example, we consider that
$P_j = 1.0$, $K_{j,14} = 1.0 \mu M$, [TF$_j$] = 100 $\mu M$ with $j = 1...n$. When one TF, TF$_i$, coexists with
another TF, TF$_j$, increasing [TF$_i$] from 10 $\mu M$ to 100 $\mu M$ leads to that $\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$ increases
to 200 from 110, namely an approximate increase of 82% in $\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$. However, when one
TF, TF$_i$, coexists with another 100 TF, TF$_j$, increasing [TF$_j$] from 10 $\mu M$ to 100 $\mu M$ leads to
that $\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$ increases to 10100 from 10010, namely an approximate increase of 0.9% in
$\sum_{j=1}^{n} \frac{P_j[TF_j]}{K_{j,14}}$.

Thus, existence of a large number of CaM-binding proteins forms a buffering system,
in which the concentration of a CaM-binding TF-specific active signal is insensitive to
changes in the concentration of another CaM-binding TF or protein. An example of this is
shown in Fig. 2.
In this example, we assume that \([\text{TF}_1]_t\), the total concentration of a CaM-binding transcription factor, TF$_1$, is 10 µM. We compare how \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\), which is the concentration of the active signal of TF$_1$, depends upon the concentration of a CaM-binding protein when different numbers of other CaM-binding proteins coexist. Fig. 2A represents an experimentally measured calcium signature (Whalley et al., 2011). Fig. 2B shows that, when the transcription factor, TF$_1$, competes for CaM binding with one CaM-binding protein, TF$_2$, changing the total concentration of TF$_2$, \([\text{TF}_2]_t\), from 1 µM to 10 µM and 100 µM markedly affects \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\). Thus, when the calcium signature, as shown in Fig. 2A, emerges, although the total concentration of the CaM-binding transcription factor TF$_1$, \([\text{TF}_1]_t\), remains unchanged (i.e. 10 µM), changing the total concentration of the CaM-binding protein TF$_2$, \([\text{TF}_2]_t\), alters the capability of the transcription factor TF$_1$ for generating an active calcium signal. This is because the concentration of the active signal of the transcription factor TF$_1$, \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\), has changed due to the competition between the CaM-binding transcription factor, TF$_1$, and the CaM-binding protein, TF$_2$, for binding with CaM. In Fig. 2C, the number of CaM-binding proteins TF$_j$ increases to 11 (i.e. \(j=2,\ldots,12\)). Since the number of CaM-binding proteins has increased, changing the total concentration of one CaM-binding protein, \([\text{TF}_2]_t\) (the concentrations of other 10 CaM-binding proteins remain unchanged), from 1 µM to 10 µM and 100 µM only slightly affects \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\). Moreover, when the number of CaM-binding proteins TF$_j$ further increases to 101 (i.e. \(j=2,\ldots,102\)), the effects of changing the total concentration of one CaM-binding protein, \([\text{TF}_1]_t\), from 1 µM to 10 µM and 100 µM on \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\) becomes negligible (Fig. 2D). Thus, when a large number of CaM-binding proteins coexist, a calcium signature, as shown in Fig. 2A, can generate a TF specific active signal, \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\), to mediate specific changes in gene expression. The concentration of such an active signal is insensitive to changes in the concentration of other CaM-binding proteins. Therefore, the existence of a large number of CaM-binding proteins results in the fidelity of a calcium signature to its TF specific active signal.

In addition, Fig. 2E shows that the concentration of the active calcium signal of TF$_1$, \([4\text{Ca}^{2+}-\text{CaM-TF}_1]\), is always amplified regardless of the number of CaM-binding proteins. Therefore, the coexistence of a large number of CaM-binding proteins in plant cells does not affect the intrinsic property of amplifying calcium signatures for Ca$^{2+}$-CaM-TF interactions. In this way, the nonlinear amplification of calcium signatures, as demonstrated in Fig. 2E, allows plant cells to effectively distinguish the kinetics of different calcium signatures to
produce specific changes in gene expression, in spite of the coexistence of a large number of CaM-binding proteins in plant cells.

**Specific gene expression responses to calcium signatures require an appropriate relationship between the active signal concentration and DNA binding affinity**

Gene expression is a complex process, which involves both transcription and mRNA degradative processes. Both processes can be regulated in response to signalling. For example, transcriptional processes can be regulated by calcium signals (Reddy et al., 2011; Seybold et al., 2014; Fromm and Finkler, 2015; Tsuda and Somssich, 2015; Zhu, 2016). Moreover, gene expression can form a network, in which the expression of one gene can be regulated by other genes (Reddy et al., 2011; Seybold et al., 2014; Fromm and Finkler, 2015; Tsuda and Somssich, 2015; Zhu, 2016).

Here we concentrate on elucidating the mechanism for the information flow from calcium signals to a specific gene expression response. To do so, we consider a simple Ca$^{2+}$-regulated gene expression process: transcriptional rate is regulated by calcium signals. The principle revealed by this simple example can be applied to more complex gene expression processes, as will be demonstrated in the Section “Case study 2: plant immunity gene expression”.

Eq. 4 describes that the transcription of a gene that is positively regulated by calcium signals.

$$\frac{d[mRNA]}{dt} = V - k_{\text{decay}}[mRNA] \quad \text{(eq. 4)}$$

with $V = k_{\text{base}} + \frac{V_{\text{max}}[4C{a}^{2+}-CaM-TF_i]}{1+[4C{a}^{2+}-CaM-TF_i]}$. Here, $V$ is the transcription rate; $k_{\text{decay}}$ is the decay constant of the mRNA; $k_{\text{base}}$ is the base rate of transcription; $V_{\text{max}}$ is the maximal transcription rate regulated by calcium signals, $k_d$ is the binding affinity between the active complex, $4C{a}^{2+}-CaM-TF_i$, and DNA.

As analysed above, when any calcium signature emerges in an environment of multiple proteins competing for CaM binding, $[Ca^{2+}]$ is amplified into a robust TF-specific active complex for any CaM binding transcription factor, $4C{a}^{2+}-CaM-TF_i$. Eq. 4 shows that, since $4C{a}^{2+}-CaM-TF_i$ is always amplified for any calcium signature, the transcription rate, $V$, will effectively be different for different calcium signatures, leading to a different specific gene expression response in each case. Examination of eq. 4 reveals how
the relationship of $k_{\text{base}}$, $k_d$, and $[4\text{Ca}^{2+} - \text{CaM} - TF_i]$ determines $\text{Ca}^{2+}$-regulated gene expression. If $k_{\text{base}} \gg \frac{V_{\text{max}}}{k_d \frac{[4\text{Ca}^{2+} - \text{CaM} - TF_i]}{1+ [4\text{Ca}^{2+} - \text{CaM} - TF_i]}},$ the base rate of transcription is more important than the rate regulated by the calcium signal. Thus, the effects of a calcium signature on gene expression are negligible under these particular conditions. If $k_d \gg [4\text{Ca}^{2+} - \text{CaM} - TF_i],$ the effects of a calcium signature on gene expression is limited, this is because the term $\frac{V_{\text{max}}}{k_d \frac{[4\text{Ca}^{2+} - \text{CaM} - TF_i]}{1+ [4\text{Ca}^{2+} - \text{CaM} - TF_i]}$ can become very small. If $k_d \ll [4\text{Ca}^{2+} - \text{CaM} - TF_i],$ the effects of any calcium signature would become approximately a constant $V_{\text{max}}$. Thus, in this case different calcium signatures induce similar transcription rates, leading to similar levels of mRNA. Therefore, in order for a specific gene expression response to calcium signatures to be generated, $[4\text{Ca}^{2+} - \text{CaM} - TF_i]$ should be not much larger or smaller (e.g. 2 orders larger or smaller) than $k_d$. Under this condition, different calcium signatures can be decoded to generate specific gene expression responses. Fig. 3 summarises the design principle that governs how the binding mechanism between $\text{Ca}^{2+}$, calmodulin (CaM), and transcription factor (TF), which emerges from two pairs of $\text{Ca}^{2+}$-binding EF-hand domains, a TF-binding domain in CaM, and a CaM-binding domain in the TF, leading to specific gene expression.

In summary, the design principle of $\text{Ca}^{2+}$-CaM-TF interactions includes the following three key aspects for information flow from calcium signals to gene expression: 1) nonlinear amplification of a calcium signal; 2) generation of a $\text{Ca}^{2+}$-induced TF-specific active signal; and 3) once the binding affinity between the active calcium signal and DNA is appropriate, specific gene expression responses can be generated.

Below, we use two examples to demonstrate how calcium signatures generate specific gene expression responses following the design principle revealed herein.

Case study 1: a simple $\text{Ca}^{2+}$-regulated gene expression system

To test how the design principle summarised in Fig. 3 governs the decoding of different calcium signatures to generate specific gene expression responses, we first studied a simple theoretical gene expression system described by equation 4 using artificial calcium signatures. An advantage of artificial calcium signatures is that the parameters of different calcium signatures can be compared with each other so that effects of those parameters of calcium signatures on gene expression can be examined. In addition, investigating both artificial calcium signatures in this case study and examining experimentally measured
calcium signatures in *case study 2* below allows us to show that the design principle revealed
in this research is generic for any calcium signature.

Fig. 4A shows three calcium signatures with the same type of kinetics. All three
calcium signatures take a sinusoidal form with the same period, but their amplitudes are
different. For simplicity we study one period of these sinusoidal calcium signatures only. Fig.
4B shows that a relatively modest change in the amplitude of these three calcium signatures
(0.2 μM to 0.4 μM) is amplified into large fold differences in the concentration of the active
signal, \([4Ca^{2+} - CaM - TF_i]\). Subsequently, this large difference in the concentrations of
the three active signals leads to different fold changes of mRNA concentration, Fig. 4C. We
emphasize that the large difference (from ca. max. 6 fold to ca. max. 58 fold) in mRNA
concentrations in Fig. 4C stems entirely from the relatively modest difference in the
amplitude of the three calcium signatures (0.2μM to 0.4μM), as the kinetics of the three
calcium signatures is the same. Derivation of equation 2 in Supplemental Information reveals
that the term \(\left(\frac{[Ca^{2+}]}{[Ca^{2+}]_{ss}}\right)^4\) emerges from two pairs of Ca\(^{2+}\)-binding EF-hand domains and a TF-
binding domain in the CaM structure. Therefore, the ability of CaM to bind four Ca\(^{2+}\) and one
TF molecule results in the amplification of calcium signal. Fig. 4D further shows that the fold
change of \([4Ca^{2+} - CaM]\) is approximately the same as that of \([4Ca^{2+} - CaM - TF_i]\), as
shown in Fig. 4B, and this is also confirmed in Fig. 4E. Therefore, Fig. 4B, D, and E together
reveal that the amplification of the calcium signatures, shown in Fig. 4A, is originated from
two pairs of Ca\(^{2+}\)-binding EF-hand domains in the CaM structure and that it is further relayed
to the binding between \(4Ca^{2+} - CaM\) complex and transcription factor.

Fig. 4F shows three calcium signatures with the same average calcium concentration
(0.2 μM). The difference between the three calcium signatures is their kinetics. Fig. 4G
shows that these relatively slight differences in the kinetics of these three calcium signatures
is amplified into the kinetics of the active signal, \([4Ca^{2+} - CaM - TF_i]\). Subsequently, the
difference in the kinetics of the three active signals leads to different fold changes of mRNA
concentration, Fig. 4H. We emphasize that the difference (from ca. max. 80 fold to ca. max.
148 fold) in the fold change of mRNA concentrations in Fig. 4H stems solely from the
difference in the kinetics of the three calcium signatures, as the average calcium
concentration is the same for the three calcium signatures. Following the analysis represented
by Fig. 4D and 4E, Fig. 4G, I, and J together reveal that the amplification of the calcium
signatures shown in Fig. 4F also originates from two pairs of Ca\(^{2+}\)-binding EF-hand domains
in the CaM structure and that it is further relayed to the binding between $4Ca^{2+} - CaM$
complex and transcription factor.

Therefore, Fig. 4A-J show that different calcium signatures, displaying only modest
differences, can generate very different specific gene expression responses following the
design principle.

Another aspect of the design principle is that when a large number of CaM-binding
proteins coexist, a calcium signature is able to generate a specific gene expression response,
which is not affected by the concentrations of another CaM-binding protein. The numerical
analysis shown in Fig. 5 confirms that, when a large number of CaM-binding proteins
coexist, the gene expression response of the simple system remains the same even if the
concentration of a CaM-binding protein has changed from 0.01 µM to 10000 µM.

When the calcium signature shown in Fig. 5A emerges and if one protein and one
transcription factor compete for binding with CaM simultaneously, changing the protein
concentration affects the TF-specific gene expression, resulting in different fold changes of
mRNA, Fig. 5B. However, when a large number of CaM-binding proteins (here 101 proteins)
and one transcription factor compete for the binding of CaM, changing the concentration of
one CaM-binding protein out of 101 (i.e. the concentration of the remaining 100 CaM-
binding proteins remains unchanged) does not affect the TF-specific gene expression,
resulting in no different fold changes of mRNA, Fig. 5C. Therefore, when many proteins
compete for CaM binding, specific gene expression responses to calcium signatures are
robust, and TF-specific. This is because the concentration of a CaM-binding TF-specific
active signal is insensitive to changes in another CaM-binding TF or protein concentration
when a large number of CaM-binding proteins coexist, as analysed in Fig. 2.

Fig. 4 and 5 together explain how the interaction between $Ca^{2+}$, CaM, and the CaM-
binding transcription factor induces specific gene expression responses in the simple $Ca^{2+}$-
regulated gene expression process described by equation 4. Therefore, this demonstrates that
the design principle, shown in Fig. 3, establishes the link between calcium signatures and
specific gene expression responses induced by the signatures.

The dynamics for the interactions between $Ca^{2+}$, CaM, and CaM-binding proteins can
be generally examined using differential equations (Pepke et al., 2010; Liu et al., 2015;
Lenzoni et al., 2018). If a quasi-equilibrium state for the interactions of $Ca^{2+}$, CaM, and CaM-
binding proteins has been established, all differential equations describing the interactions
between Ca\textsuperscript{2+}, CaM, and CaM-binding proteins become zero and equations 1-3 can be derived. Fig. 5D further examines the validity and effects of the quasi-equilibrium assumption for deriving equations 1-3. Fig. 5D shows that, for experimentally measured parameters of Ca\textsuperscript{2+}-CaM binding constants (Shifman et al., 2006; Pepke et al., 2010), the gene expression response curve computed without the quasi-equilibrium assumption for the interactions of Ca\textsuperscript{2+}, CaM, and CaM-binding proteins (i.e. differential equations) overlaps with the gene expression response curve computed with the quasi-equilibrium assumption (i.e. equation 3), indicating that a quasi-equilibrium state of Ca\textsuperscript{2+}-CaM –TF interaction in plant cells has been established.

Although the quasi-equilibrium state assumption is valid for experimentally measured parameters of Ca\textsuperscript{2+}-CaM binding constants, reducing those parameters may make the assumption invalid. We further test how validity of the assumption affects gene expression responses. When all “on” and “off” rate constants for Ca\textsuperscript{2+}-CaM interactions are simultaneously reduced by the same fold from their experimentally measured values (Shifman et al., 2006; Pepke et al., 2010), the equilibrium constants for all Ca\textsuperscript{2+}-CaM binding processes remain the same as the experimental values. This is because an equilibrium constant is the ratio between the “off” rate constant and the “on” rate constant. However, if all “on” and “off” rate constants are simultaneously reduced, the quasi-equilibrium assumption may become invalid due to slow binding rates. Fig. 5D shows that when all “on” and “off” rate constants for Ca\textsuperscript{2+}-CaM interactions are simultaneously reduced by 100 fold, the quasi-equilibrium assumption becomes invalid. Once this happens, the calcium signature, as shown in Fig. 5A, is less capable of inducing a gene expression response. In addition, Fig. 5D also shows that very small Ca\textsuperscript{2+}-CaM “on” and “off” rate constants (i.e. they are reduced by 1.0E4 fold from their experimental values) render gene expression response to calcium signatures impossible. This implies that establishing a quasi-equilibrium state is favourable for a calcium signature to induce gene expression responses.

Another important aspect of the design principle, as described in Fig. 3, is that specific gene expression responses to calcium signatures require an appropriate relationship between the active signal concentration, \([4Ca\textsuperscript{2+} – CaM – TF_i]\), and DNA binding affinity. The dissociation equilibrium constant (i.e. the binding affinity) for the binding of the Ca\textsuperscript{2+}-CaM complex to a transcription factor is an important parameter. Changing the value of the dissociation equilibrium constant changes \([4Ca\textsuperscript{2+} – CaM – TF_i]\), and therefore affects the relationship between \([4Ca\textsuperscript{2+} – CaM – TF_i]\) and DNA binding affinity. Supplemental Fig. S1
shows the effects of the dissociation equilibrium constant for the binding of the Ca\(^{2+}\)-CaM complex to a transcription factor, \(K_{14}\), on gene expression regulated by the transcription factor. Supplemental Fig. S1A shows an artificial calcium signature. Supplemental Fig. S1B shows that decreasing the value of \(K_{14}\) increases the steady-state value of \([4Ca^{2+} - CaM - TF_i]\). Similarly, Supplemental Fig. S1C shows that decreasing the value of \(K_{14}\) increases the value of \([4Ca^{2+} - CaM - TF_i]\) responding to the calcium signature. Supplemental Fig. S1D shows that, for the three different values of \(K_{14}\), the calcium signature is always amplified. Supplemental Fig. S1E, F, and G show that, although decreasing the value of \(K_{14}\) increases both the steady-state mRNA concentration and the mRNA concentration responding to the calcium signature, three different values of \(K_{14}\) lead to three different responses of mRNA concentration to the calcium signature. Therefore, different transcription factors with different values of \(K_{14}\) can generate different responses of mRNA concentration to a calcium signature. This indicates that the dissociation equilibrium constant for the binding of the Ca\(^{2+}\)-CaM complex to a transcription factor, \(K_{14}\), is an important parameter for specific gene expression responses to a calcium signature. Similarly, analysis in Figure S2 for a different artificial calcium signature also supports the above conclusion.

In summary, this example shows that, for the simple gene expression system described by equation 4, different calcium signatures can be decoded to generate specific gene expression responses following the design principle, as described in Fig. 3.

**Case study 2: plant immunity gene expression**

The CaM-binding transcription factors CAMTA3 (AtSR1) and CBP60g regulate the expression of two important plant immunity genes: enhanced disease susceptibility 1 (EDS1) and isochorismate synthase 1 (ICS1) (Zhang et al., 2010; Zhang et al., 2014). Recently, we developed a dynamic model to determine how expression of both EDS1 and ICS1 is regulated by different calcium signatures and analysed the model using computer simulation of differential equations (Lenzoni et al., 2018). It was demonstrated that the model was able to predict the expression of both EDS1 and ICS1 (Lenzoni et al., 2018). Here we use this system as an example to study how the expression response of both EDS1 and ICS1 to calcium signatures is generated following the design principle, as shown in Fig. 3.

Fig. 6A and B show two empirically-derived calcium signatures induced by two calcium agonists: mastoparan and extracellular calcium (Lenzoni et al., 2018). The model developed for studying how expression of both EDS1 and ICS1 (Lenzoni et al., 2018) is regulated by different calcium signatures employed experimentally measured parameters for
both Ca\(^{2+}\)-CaM binding (Shifman et al., 2006; Pepke et al., 2010) and CaM-CAMTA3 binding (Bouche et al., 2002; Finkler et al., 2007). Fig. 6C and F show that, for these experimentally measured “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions, the curve calculated using differential equations overlaps with the curve calculated using equation 3 for both \([4Ca^{2+} – CaM – CAMTA3]\) and \([4Ca^{2+} – CaM – CBP60g]\), indicating that the quasi-equilibrium assumption for interactions of Ca\(^{2+}\)-CaM-CAMTA3 and Ca\(^{2+}\)-CaM-CBP60g, as well as the interactions of Ca\(^{2+}\), CaM, and 100 other CaM-binding proteins are valid. Thus, the two active signals in Fig. 6C and F, 4Ca\(^{2+}\)-CaM-CAMTA3 and 4Ca\(^{2+}\)-CaM-CBP60g, are differentially induced by the two calcium signatures (Fig. 6A and B). Moreover, Supplemental Fig. S3C and F show that the two active signals are effectively and differentially amplified.

However, when both “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E5 fold from their experimental values, the quasi-equilibrium assumption for deriving equation 3 becomes invalid, as evidenced by the differences between the curve calculated using differential equations and the curve calculated using equation 3 (Fig. 6D and G). Moreover, the two active signals in Fig. 6D and G, 4Ca\(^{2+}\)-CaM-CAMTA3 and 4Ca\(^{2+}\)-CaM-CBP60g, are both less effectively amplified (Supplemental Fig. S3D and G) than in Fig. 6C and F, indicating that a valid quasi-equilibrium assumption makes the amplification of both active signals more favourable. In addition, when the “on” and “off” rate constants are reduced by 1.0E8 fold, the two active signals, 4Ca\(^{2+}\)-CaM-CAMTA3 and 4Ca\(^{2+}\)-CaM-CBP60g, cannot respond to either of the two calcium signatures (Fig. 6A and B), as evidenced by the overlapping of the two flat curves corresponding to the two calcium signatures in Fig. 6E and H. Furthermore, neither of the two active signals in Fig. 6E and H can be amplified anymore under these conditions (Supplemental Fig. S3E and H), indicating that very small Ca\(^{2+}\)-CaM “on” and “off” rate constants render amplification of calcium signals impossible.

CAMTA3 and CBP60g are the transcription factors that regulate the expression of *EDS1* and *ICS1*, respectively (Zhang et al., 2010; Zhang et al., 2014). These correspond to the two active signals shown in Fig. 6C and F, the two calcium signatures inducing different mRNA levels for both *EDS1* and *ICS1* genes (Fig. 7A and D), leading to specific gene expression responses for both genes. Moreover, since the curve calculated using differential equations overlaps with the curve calculated using equation 3 for the fold change of mRNA of both *EDS1* and *ICS1* (Fig. 7A and D), gene expression responses of both *EDS1* and *ICS1* to the two calcium signatures clearly follow the design principle (Fig. 3 and equations 1-3).
When the “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E5 fold, Fig. 7B and E show that less effective amplification of both active signals, 4Ca\(^{2+}\)-CaM-CAMTA3 and 4Ca\(^{2+}\)-CaM-CBP60g (Fig. 6D and G, S3D and G), markedly affects the mRNA level of both ICS1 and EDS1. Furthermore, when the “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E8 fold, no amplification of either of the two active signals, 4Ca\(^{2+}\)-CaM-CAMTA3 and 4Ca\(^{2+}\)-CaM-CBP60g, occurs (Fig. 6E and H, S3E and H), leading to no change in expression of either EDS1 or ICS1. This is evidenced by the overlapping of the two flat curves corresponding to the two calcium signatures in Fig. 7C and F, showing no change in gene expression response to either signature.

In conclusion, for experimentally measured “on” and “off” rate constants (Shifman et al., 2006; Pepke et al., 2010), the two calcium signatures (Fig. 6A and B) are decoded following design principle to generate specific expression of both EDS1 and ICS1 (Fig. 7A and D). If the “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are largely reduced (e.g. 1.0E8 fold), specific gene expression responses to the two calcium signatures become impossible (Fig. 7C and F). Therefore, the actual values of “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions, as experimentally measured in the literature (Shifman et al., 2006; Pepke et al., 2010), ensure that plant immunity gene expression responses of both EDS1 and ICS1 follow the design principle to decode the two calcium signatures induced by two calcium agonists: mastoparan and extracellular calcium (Fig. 6A and B).
DISCUSSION

Most stimuli lead to a transient elevation in cellular calcium concentration in plant cells. Importantly, different stimuli produce calcium elevations with different characteristics: a unique “calcium signature”. These calcium signatures are decoded to generate specific responses (Edel et al., 2017; Yuan et al., 2017; Bender et al., 2018; Kudla et al., 2018). An intriguing question is how can one messenger (Ca\(^{2+}\)) be decoded by so many decoders (transcription factors and proteins) in plant cells (Edel et al., 2017)?

Design principles are the underlying properties of network structures that have evolved to endow the network functions. This work reveals the design principle for decoding calcium signals to generate specific gene expression response in plant cells via transcription. The design principle links the structural characteristics of CaM and TF with the capability of decoding calcium signatures in plant cells, and it therefore reveals how the mechanism of Ca\(^{2+}\), CaM, and TF interactions leads to specific gene expression. It includes the following three important aspects: Firstly, the binding mechanism between Ca\(^{2+}\), CaM, and TF, which emerges from two pairs of Ca\(^{2+}\)-binding EF-hand domains and a TF-binding domain in the CaM structure, possesses an intrinsic property of amplifying calcium signals in the format of \(\left(\frac{[\text{Ca}^{2+}]}{[\text{Ca}^{2+}]_{ss}}\right)^4\) being multiplied by a factor that is relatively less important. We derived the equations for describing the amplification of calcium signals (equations 1 and 2) and mathematically proved that calcium signals are always amplified (Supplemental Information). Since any calcium signature is always amplified, small differences in the kinetics or parameters of calcium signatures can be read out to generate relatively much larger specific gene expression responses (Fig. 4 and 5). Secondly, the existence of a large number of CaM-binding TFs or proteins in plant cells (Reddy et al., 2011; Poovaiah et al., 2013; Virdi et al., 2015; Edel et al., 2017; Yuan et al., 2017; Bender et al., 2018; Kudla et al., 2018) can form a buffering system such that the concentration of a CaM-binding TF-specific active signal is insensitive to changes in the concentration of another CaM-binding TF or protein (Fig. 2D). Thus, although many proteins compete for the binding of CaM, Ca\(^{2+}\)-induced TF-specific gene expression will not in fact be affected by the concentration of another CaM-binding TF or protein (Fig. 5C) in plant cells. Although a TF-specific gene expression event must be controlled by the concentration of this transcription factor, it would not be advantageous if it can also be altered by changes in the concentrations of other proteins. This is a clear example of inbuilt robustness of the network endowed by the design principle. Our results also show that when a CaM-binding TF competes for CaM binding
with one or a few CaM-binding TFs and proteins, gene expression regulated by a CaM-binding TF will be interfered with by another CaM-binding TF or protein (Fig. 5C). Thus, competition of CaM-binding TFs or proteins for CaM binding may have a role in the relationship between calcium signals and gene expression response if a small number of CaM-binding TFs or proteins exist in plant cells. Interestingly, for postsynaptic cells, a model in which 6 proteins compete for CaM binding, competition plays a role in setting the frequency-dependence of Ca\textsuperscript{2+}-dependent proteins (Romano et al., 2017), and therefore it was suggested that competitive tuning could be an important dynamic process underlying synaptic plasticity. Therefore, both our work and the results in the literature (Romano et al., 2017) suggest that competition of a small number of CaM-binding proteins will cause that the response of one CaM-binding protein to calcium signals to be affected by other CaM-binding proteins. Thirdly, an appropriate relationship between a Ca\textsuperscript{2+}-induced TF-specific active signal concentration and its DNA binding affinity is important for a Ca\textsuperscript{2+}-induced TF-specific gene expression response. For a Ca\textsuperscript{2+}-induced TF-specific active signal to generate gene expression responses, its concentration needs to be similar to the binding affinity between the signal and DNA. In general, for any binding to perform biological functions, the affinity between a ligand and its binding partner should not be very different from the concentration of the ligand (e.g. the differences not larger than 2 orders) (Kuriyan et al., 2013). Our analysis reveals this aspect is also applicable to Ca\textsuperscript{2+}-induced gene expression systems.

We used two examples to show how gene expression follows the design principle to decode different calcium signatures. Gene expression is generally regulated in a complex way (Karlebach and Shamir, 2008). The simplest gene expression process includes: (1) gene transcription is activated or suppressed by a transcription factor; and (2) the mRNA decays. Our first example tested how a simple gene expression system decodes different calcium signatures. We found that different kinetics or different parameters (here testing amplitudes of a sinusoidal calcium signature) could be decoded following the design principle. Therefore, this example demonstrated that different calcium signatures, even if their differences are modest (Fig. 4A and F), can generate specific gene expression responses (Fig. 4C and H) following the design principle. Experimentally any, even a modest, calcium signature (e.g. in response to ozone (Clayton et al., 1999)) is able to induce gene expression. Therefore, our results, shown in Fig. 4 and 5, imply that the design principle is a general principle for governing the decoding of calcium signatures in simple gene expression systems, in which transcription rates are regulated by calcium signals.
The second example showed that expression of two plant immunity genes (*EDS1* and *ICS1*) follows the design principle to decode two empirical calcium signatures induced by two calcium agonists (mastoparan and extracellular calcium). Calcium signals regulate the expression of *EDS1* and *ICS1* at two levels. First, CAMTA3 and CBP60g are well-characterized Ca\(^{2+}\)/CaM-regulated transcription factors and both have a CaM binding domain (Finkler et al., 2007; Galon et al., 2008; Kim et al., 2009; Wang et al., 2009; Zhang et al., 2010; Reddy et al., 2011; Wang et al., 2011; Bickerton and Pittman, 2012; Poovaiah et al., 2013). Thus, calcium signals regulate the activities of both CAMTA3 and CBP60g. Second, expression of *EDS1* and *ICS1* forms a regulatory network (Zhang et al., 2014; Lenzoni et al., 2018) and their expression is regulated by each other via this network (Zhang et al., 2014; Lenzoni et al., 2018). In spite of this complexity in regulating the expression of *EDS1* and *ICS1*, the design principle still governs the expression of both *EDS1* and *ICS1*. Thus, our results, shown in Fig. 6 and 7, imply that design principle is a general principle for governing the decoding of calcium signature in complex gene expression systems, in which multiple transcription factors are regulated by calcium signals and gene expression itself forms a regulatory network.

The design principle, as described in Fig. 3 and summarised above, is generic for elucidating the decoding of calcium signals which generate specific gene expression responses via transcription. Therefore, it can be integrated with a wide range of experimental analysis. For example, we have shown how to study gene expression for both simple and complex systems that are regulated by any calcium signatures (Fig. 4-7). Arabidopsis genes responding to simultaneous biotic and abiotic stresses have been experimentally identified (Atkinson et al., 2013). Following the analysis shown in Fig. 6 and 7, any genes that are regulated by calcium signals under both biotic and abiotic conditions could be theoretically investigated based on the experimental measurements of gene expression corresponding to the relevant calcium signatures. In addition, the role of CaM binding to CAMTA3 in regulating immunity genes was experimentally investigated (Kim et al., 2017). The design principle could be used to quantitatively analyse this role for different binding domains within CAMTA3. It should be noted that this requires experimental inputs to provide parameters. For example, the binding affinity constant of CaM to CAMTA3 in the presence of Ca\(^{2+}\) had been experimentally measured (Bouche et al., 2002; Finkler et al., 2007). Therefore, it is important that future experiments measure such parameters for the binding between CaM and other (than CAMTA3) CaM-binding proteins. CaM and other calcium-binding proteins have the potential to regulate and modify calcium signatures themselves. We
could address this important aspect of calcium signalling in the future. Some genes regulated
by calcium signatures encode proteins with roles in transporting/binding calcium in plant
cells (Kudla et al., 2010; Delormel and Boudsocq, 2019). To further study the effects of gene
expression on the generation of calcium signatures, the design principle established in this
work could in the future be combined with the processes for generation of calcium signature
(Medvedev, 2018). This may be important in understanding alterations in calcium signatures
as a result of acclimation to stress, and due to interaction between different stresses. For
example, the design principle developed in this work can be used to study the effects of the
concentrations of CaM and transcription factors on the mRNA levels of gene expression,
which can be linked with the processes of calcium transport to quantitatively examine the
effects of mRNA levels on generation of calcium signature in the future. In order to further
validate the design principle we present here, the plant immunity system would be a good
model. Future experiments could involve complementation of camta3 and cbp60g mutants
with CAMTA3 and CBP60g in which the protein coding regions have been modified to alter
binding constants to DNA and CaM. The effect of these altered affinities could be predicted
using our mathematical model, and tested empirically in the complemented lines by
measuring ICS1/EDS1 gene expression in response to applied calcium signatures.

Calcium signals are the lead currency of plant information processing (Dodd et al.,
2010; Kudla et al., 2010), and they regulate many different responses in plant cells. However,
little is known about the underlying principle for how information flows from calcium signals
to specific gene expression responses in plant cells. This work reveals the underlying
principles for linking the structure of CaM and TF molecule with calcium-regulated gene
expression response through Ca$^{2+}$-CaM-TF binding mechanism and the emerging property of
calcium signal amplification. The design principle indicates that the existing interaction
network of Ca$^{2+}$, CaM, and proteins, which may have been evolutionarily tuned (Edel et al.,
2017), effectively navigates calcium signatures to generate specific gene expression responses
in plant cells. Experimental data have shown multiple levels of complexities in decoding
calcium signals in plant cells (Edel et al., 2017; Yuan et al., 2017; Bender et al., 2018; Kudla
et al., 2018). Plants cells possess four main types of Ca$^{2+}$ sensor proteins to relay or decode
Ca$^{2+}$ signalling: CaM, CaM-like proteins (CMLs), calcineurin B-like proteins (CBLs), and
Ca$^{2+}$-dependent protein kinases (CDPKs or CPKs) (Yuan et al., 2017). These proteins relay
or decode calcium signals at both the transcriptional and post-translational levels (Yuan et al.,
2017). This work has focused on the interactions between Ca$^{2+}$, CaM, and TFs at
transcriptional level and revealed that transcriptional decoding of calcium signals follows a
general design principle. Other Ca\(^{2+}\) sensor proteins can have different numbers of Ca\(^{2+}\)-
binding sites or possess complex molecular structures. For example, CMLs may have one to
six EF-hands and one to four Ca\(^{2+}\)-binding sites (La Verde et al., 2018). A Ca\(^{2+}/CaM-\)
dependent protein kinase (CCaMK) possesses three additional Ca\(^{2+}\)-binding sites in addition
to its CaM-binding site (Miller et al., 2013). Our methodology for unravelling the design
principle for transcriptional decoding of calcium signals may be further developed to study
the underlying general principle for other Ca\(^{2+}\)-regulated signalling systems in the future.
MATERIALS AND METHODS

Ca^{2+}–CaM–protein interactions

The interaction between Ca^{2+}, CaM, and any protein can form different binding complexes. CaM has two pairs of Ca^{2+}-binding EF-hand domains located at the N- and C-terminus, respectively (Finn and Forsen, 1995; Valeyev et al., 2008). Thus, for a protein with one calmodulin binding site, 18 different binding complexes can form via 33 elementary binding processes. A detailed description of these interactions is previously presented in detail (Liu et al., 2015; Lenzoni et al., 2018), and the 33 elementary binding processes are included in Table S1 in Supplemental Information. Experimentally measured parameters for the interactions between Ca^{2+} and CaM are included in Table S2.

Modelling expression of plant immunity genes

The model used to examine expression of plant immunity genes (ICS1 and EDS1) was previously described in detail (Lenzoni et al., 2018). The differential equations and parameters of the model were included in the previous work (Lenzoni et al., 2018). This work uses this model to study how expression of both ICS1 and EDS1 decodes calcium signatures following the design principle.

Numerical Method

All computational results are generated using simulator Berkeley Madonna (www.berkeleymadonna.com). For differential equations, Rosenbrock (Stiff) method is used with a tolerance of 1.0e-5. Much smaller tolerances (1.0E-6 to 1.0E-8) are also tested and the numerical results show that further reduction of tolerances does not improve the accuracy of numerical results.

Accession Numbers

EDS1: AT3G48090 (https://www.arabidopsis.org/servlets/TairObject?id=39706&type=locus)
ICS1: AT1G74710 (https://www.arabidopsis.org/servlets/TairObject?id=28521&type=locus)

SUPPLEMENTAL DATA

Supplemental Figure S1. Effects of K_{14}, the dissociation equilibrium constant for the binding of the Ca^{2+}-CaM complex to the i^{th} TF, on gene expression regulated by the TF for calcium signature shown in Figure S1A.
Supplemental Figure S2. Effects of $K_{14}$, the dissociation equilibrium constant for the binding of the $Ca^{2+}$-CaM complex to the $i^{th}$ TF, on gene expression regulated by the TF for calcium signature shown in Figure S2A.

Supplemental Figure S3. Two calcium signatures are decoded to generate specific expression of $EDS1$ and $ICS1$ following design principle: responses of two active signals, $4Ca^{2+}$-CaM-CAMTA3 and $4Ca^{2+}$-CaM-CBP60g, to two experimentally measured calcium signatures.

Supplemental Table S1. Interactions of $Ca^{2+}$, calmodulin (CaM), and one transcription factor.

Supplemental Table S2. Experimentally measured parameters for the interactions between $Ca^{2+}$ and CaM.

ACKNOWLEDGEMENTS

This work was funded by an EU-funded Initial Training Network (ITN) CALIPSO GA 2013–607607.

FIGURE LEGENDS

Figure 1. $Ca^{2+}$-CaM-TF interactions always amplify calcium signals. The parameter for the cooperative binding between CaM and a TF in the presence of $Ca^{2+}$ is $P$. A. Value of function $f$ for $P=0.1$. Scatter crosses are the theoretical minimum value of $f$. Blue, red, and green curves correspond to the total concentration of TF, $[TF]_t$, to be 0.01µM, 10 µM, and 1.0E5 µM, respectively. The blue and red curves overlap, indicating that the numerical values of $f$ are always the same for the two concentrations of TF. B. Corresponding to Fig. 1A, $\frac{[4Ca^{2+}\text{–CaM–TF}]}{[4Ca^{2+}\text{–CaM–TF}]_{ss}}$ value for $P=0.1$ is calculated using equation 2. C. Value of function $f$ for $P=1.0E^{-4}$. Blue, red, and green curves correspond to total concentration of TF, $[TF]_t$, to be 0.01 µM, 10 µM, and 1.0E5 µM, respectively. The green curve and the scatter crosses overlap, indicating that the numerical values of $f$ for $[TF]_t =1.0E5$ µM are always the same as the theoretical minimum value of $f$. D. Corresponding to Fig. 1C, $\frac{[4Ca^{2+}\text{–CaM–TF}]}{[4Ca^{2+}\text{–CaM–TF}]_{ss}}$ value for $P=1.0E^{-4}$ is calculated using equation 2.

Figure 2. Effects of changing the concentration of a CaM-binding protein on the concentration of a CaM-binding TF-specific active signal when the TF and different numbers of CaM-binding proteins coexist. A. An experimentally measured calcium signature (Whalley...
et al., 2011). The calcium signature is used to calculate $[4\text{Ca}^{2+}-\text{CaM-TF}_1]$ following equation 3. B. The TF coexists with one protein. Blue, red, and green curves correspond to the concentration of the protein to be 1 µM, 10 µM, and 100 µM, respectively. C. The TF coexists with 11 proteins. Concentrations of 10 proteins are fixed to be 100 µM. Blue, red, and green curves correspond to the concentration of the remaining protein to be 1 µM, 10 µM, and 100 µM, respectively. D. The TF coexists with 101 proteins, 100 proteins of which have the same fixed concentration: 100 µM. Blue, red, and green curve corresponds to the concentration of the remaining protein to be 1 µM, 10 µM, and 100 µM, respectively. The three curves overlap, indicating that changing the concentration of the remaining protein does not affect $[4\text{Ca}^{2+}-\text{CaM-TF}_1]$. E. $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$ for all nine curves shown in Fig. 2B and C, indicating that, for all nine cases, the maximum of $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$ reaches at least 9000 fold.

**Figure 3.** Diagram showing the design principle for transcriptional decoding calcium signatures to generate specific gene expression. A. The binding mechanism between Ca$^{2+}$, calmodulin (CaM), and transcription factor (TF), which emerges from two pairs of Ca$^{2+}$-binding EF-hand domains, and a TF-binding domain in CaM and a CaM-binding domain in TF. B. Nonlinear amplification of Ca$^{2+}$ signal emerges from A. C. An amplified, Ca$^{2+}$-induced, TF-specific active signal for each of CaM-binding TFs emerges from A and B.

**Figure 4.** Three similar calcium signatures are decoded to generate specific gene expression responses for a simple Ca$^{2+}$-regulated gene expression process. A. Three artificial calcium signatures with the same sinusoidal kinetics (the period is fixed to be 80 s, and amplitudes are 0.2 µM, 0.3 µM, and 0.4 µM, respectively. Only one period of the sinusoidal kinetics is used.) B. $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$ calculated using the three calcium signatures in Fig. 4A as the input of equation 3. C. Fold change of mRNA corresponding to the three calcium signatures in Fig. 4A. D. $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ calculated using the three calcium signatures in Fig. 4A as the input of equation 3. E. The ratio of $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ to $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$. This ratio is always equal to 1, indicating that $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ is always the same as $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$. F. Three artificial calcium signatures with the same average calcium concentration (the average of $[\text{Ca}^{2+}]$ is 0.2 µM for each of the three curves). G. $\frac{[4\text{Ca}^{2+}-\text{CaM-TF}]}{[4\text{Ca}^{2+}-\text{CaM-TF}]_{ss}}$ calculated using the three
calcium signatures in Fig. 4F as the input of equation 3. H. Fold change of mRNA corresponding to the three calcium signatures in Fig. 4F. I. $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ calculated using the three calcium signatures in Fig. 4F as the input of equation 3. J. The ratio of $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ to $\frac{[4\text{Ca}^{2+}-\text{CaM}-\text{TF}]}{[4\text{Ca}^{2+}-\text{CaM}-\text{TF}]_{ss}}$. This ratio is always equal to 1, indicating that $\frac{[4\text{Ca}^{2+}-\text{CaM}]}{[4\text{Ca}^{2+}-\text{CaM}]_{ss}}$ is always the same as $\frac{[4\text{Ca}^{2+}-\text{CaM}-\text{TF}]}{[4\text{Ca}^{2+}-\text{CaM}-\text{TF}]_{ss}}$.

**Figure 5.** Effects of the number of CaM-binding proteins or the binding rates on specific gene expression responses for a simple Ca$^{2+}$-regulated gene expression process. A. An artificial calcium signature. B. Effects of the number of CaM-binding proteins on specific gene expression responses. The TF coexists with one protein. Green, red, and blue curve corresponds to the concentration of the protein to be 1 µM, 10 µM, and 100 µM, respectively. C. Effects of the number of CaM-binding proteins on specific gene expression responses. The TF coexists with 101 proteins, 100 proteins of which have the same fixed concentration: 100 µM. Green, red, and blue curve corresponds to the concentration of the remaining protein to be 1 µM, 10 µM, and 100 µM, respectively. The three curves overlap, indicating that changing the concentration of the remaining protein does not affect fold change of mRNA. D. Effects of the binding rates on specific gene expression responses. Wide orange curve is calculated using equation 3. Blue curve is calculated using differential equations with all experimentally determined “on” and “off” binding rates (Shifman et al., 2006; Pepke et al., 2010). The wide orange curve and the blue curve overlap, indicating a quasi-equilibrium state has established. The red and green curves correspond to the “on” and “off” binding rates are reduced by 100 fold and 1.0E4 fold, respectively.

**Figure 6.** Responses of two active signals, 4Ca$^{2+}$-CaM-CAMTA3 and 4Ca$^{2+}$-CaM-CBP60g, to two experimentally measured calcium signatures. A. Two empirical calcium signatures induced by two calcium agonists: mastoparan and extracellular calcium (Lenzoni et al., 2018). B. Enlargement of Fig. 6A, showing the details of the two calcium signatures. C. Response of 4Ca$^{2+}$-CaM-CAMTA3 to the two calcium signatures. Wide solid orange curve and wide dashed orange curve are calculated using the two calcium signatures as the input of equation 3, respectively. Experimentally measured parameters are used. Black and blue curves are calculated using the two calcium signatures as the input of differential equations, respectively. The wide orange curve overlaps with the black curve. The wide dashed orange
curve overlaps with the blue curve. These results indicate a quasi-equilibrium state is established. D. Same as Fig. 6C, but both “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E5 fold from their experimental values. E. Same as Fig. 6C, but both “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E8 fold from their experimental values. Black and blue curves are flat and they also overlap, indicating that neither calcium signature can induce changes in [4Ca\(^{2+}\)-CaM-CAMTA3]. F. Same as Fig. 6C, but it is the response of 4Ca\(^{2+}\)-CaM-CBP60g to the two calcium signatures. G. Same as Fig. 6F, but both “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E5 fold from their experimental values. H. Same as Fig. 6F, but both “on” and “off” rate constants for Ca\(^{2+}\)-CaM interactions are reduced by 1.0E8 fold from their experimental values. Black and blue curves are flat and they also overlap, indicating that neither calcium signature can induce changes in [4Ca\(^{2+}\)-CaM-CBP60g].

**Figure 7.** Fold change of both *EDS1* and *ICS1* mRNA responding to two experimentally measured calcium signatures. A. Same as Fig. 6C and F, but it is the fold change of *ICS1* mRNA. B. Same as Fig. 6D and G, but it is the fold change of *ICS1* mRNA. C. Same as Fig. 6E and H, but it is the fold change of *ICS1* mRNA. D. Same as Fig. 6C and F, but it is the fold change of *EDS1* mRNA. E. Same as Fig. 6D and G, but it is the fold change of *EDS1* mRNA. F. Same as Fig. 6E and H, but it is the fold change of *EDS1* mRNA.
REFERENCES


Whalley HJ, Knight MR (2013) Calcium signatures are decoded by plants to give specific gene responses. New Phytol. 197: 690–693


Figure 1. Calcium signals are always amplified. The parameter for the cooperative binding between CaM and a TF in the presence of Ca^{2+} is $P$. A. Value of function $f$ for $P=0.1$. Scatter crosses are the theoretical minimum value of $f$. Blue, red, and green curves correspond to the total concentration of TF, $[TF]$, to be 0.01μM, 10 μM and 1.0E5 μM, respectively. The blue and red curves overlap, indicating that the numerical values of $f$ are always the same for the two concentrations of TF. B. Corresponding to Fig. 1A, $\frac{[4Ca^{2+}-CaM-TF]}{[4Ca^{2+}-CaM-TF]_{ss}}$ value for $P=0.1$ is calculated using equation 2. C. Value of function $f$ for $P=1.0E-4$. Blue, red, and green curves correspond to total concentration of TF, $[TF]$, to be 0.01 μM, 10 μM and 1.0E5 μM, respectively. The green curve and the scatter crosses overlap, indicating that the numerical values of $f$ for $[TF]=1.0E5$ μM are always the same as the theoretical minimum value of $f$. D. Corresponding to Fig. 1C, $\frac{[4Ca^{2+}-CaM-TF]}{[4Ca^{2+}-CaM-TF]_{ss}}$ value for $P=1.0E-4$ is calculated using equation 2.
Figure 2. Effects of changing the concentration of a CaM-binding protein on the concentration of a CaM-binding TF-specific active signal when the TF and different numbers of CaM-binding proteins coexist. A. An experimentally measured calcium signature (Whalley et al., 2011). The calcium signature is used to calculate [4Ca$^{2+}$-CaM-TF$_1$] following equation 3. B. The TF coexists with one protein. Blue, red and green curves correspond to the concentration of the protein to be 1 µM, 10 µM and 100 µM, respectively. C. The TF coexists with 11 proteins. Concentrations of 10 proteins are fixed to be 100 µM. Blue, red and green curves correspond to the concentration of the remaining protein to be 1 µM, 10 µM and 100 µM, respectively. D. The TF coexists with 101 proteins, 100 proteins of which have the same fixed concentration: 100 µM. Blue, red and green curve corresponds to the concentration of the remaining protein to be 1 µM, 10 µM and 100 µM, respectively. The three curves overlap, indicating that changing the concentration of the remaining protein does not affect [4Ca$^{2+}$-CaM-TF$_1$]. E. [4Ca$^{2+}$-CaM-TF$_i$] for all nine curves shown in Fig. 2B and C, indicating that, for all nine cases, the maximum of the ratio 

\[
\frac{[4Ca^{2+}-CaM-TF_i]}{[4Ca^{2+}-CaM-TF]}_i
\]

reaches at least 9000 fold.
**Ca\(^{2+}\) signature**

A

- calmodulin (CaM)
  - a TF-binding domain
  - two pairs of Ca\(^{2+}\)-binding EF-hand domains

- transcription factor (TF)
  - a CaM-binding domain

B

- Capable of binding many proteins
- Nonlinear amplification of Ca\(^{2+}\) signal

C

- An amplified, Ca\(^{2+}\) -induced, TF-specific active signal for each of CaM-binding TFs
- Transcriptional process of gene expression

Specific gene expression

**Figure 3.** A diagram schematically describes the design principle for transcriptional decoding calcium signatures to generate specific gene expression. A. The binding mechanism between Ca\(^{2+}\), calmodulin (CaM) and transcription factor (TF), which emerges from two pairs of Ca\(^{2+}\)-binding EF-hand domains and a TF-binding domain in CaM and a CaM-binding domain in TF. B. Nonlinear amplification of Ca\(^{2+}\) signal emerges from A. C. An amplified, Ca\(^{2+}\) -induced, TF-specific active signal for each of CaM-binding TFs emerges from A and B.
Figure 4. A simple Ca\(^{2+}\)-regulated gene expression process governed by design principle: calcium signatures, even if the differences of their kinetics or amplitudes are modest, are decoded to generate specific gene expression responses. A. Three artificial calcium signatures with the same sinusoidal kinetics (the period is fixed to be 80s, and amplitudes are 0.2 \(\mu\)M, 0.3 \(\mu\)M and 0.4 \(\mu\)M, respectively. Only one period of the sinusoidal kinetics is used.) B. \(\frac{[4Ca^{2+}\text{--CaM}\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) calculated using the three calcium signatures in Fig. 4A as the input of equation 3. C. Fold change of mRNA corresponding to the three calcium signatures in Fig. 4A. D. \(\frac{[Ca^{2+}\text{--CaM}]}{[Ca^{2+}\text{--CaM}]_{ss}}\) calculated using the three calcium signatures in Fig. 4A as the input of equation 3. E. The ratio of \(\frac{[4Ca^{2+}\text{--CaM}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) to \(\frac{[4Ca^{2+}\text{--CaM}\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\). This ratio is always unity, indicating that \(\frac{[4Ca^{2+}\text{--CaM}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) is always the same as \(\frac{[4Ca^{2+}\text{--CaM}\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\). F. Three artificial calcium signatures with the same average calcium concentration (the average of \([Ca^{2+}]\) is 0.2 \(\mu\)M for each of the three curves). G. \(\frac{[4Ca^{2+}\text{--CaM\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) calculated using the three calcium signatures in Fig. 4F as the input of equation 3. H. Fold change of mRNA corresponding to the three calcium signatures in Fig. 4F. I. \(\frac{[4Ca^{2+}\text{--CaM}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) calculated using the three calcium signatures in Fig. 4F as the input of equation 3. J. The ratio of \(\frac{[4Ca^{2+}\text{--CaM}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) to \(\frac{[4Ca^{2+}\text{--CaM}\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\). This ratio is always unity, indicating that \(\frac{[4Ca^{2+}\text{--CaM}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\) is always the same as \(\frac{[4Ca^{2+}\text{--CaM}\text{--TF}]}{[4Ca^{2+}\text{--CaM}]_{ss}}\).
Figure 5. A simple Ca\(^{2+}\)-regulated gene expression process governed by design principle: effects of the number of CaM-binding proteins or the binding rates on specific gene expression responses. A. An artificial calcium signature. B. Effects of the number of CaM-binding proteins on specific gene expression responses. The TF coexists with one protein. Green, red and blue curve corresponds to the concentration of the protein to be 1 µM, 10 µM and 100 µM, respectively. C. Effects of the number of CaM-binding proteins on specific gene expression responses. The TF coexists with 101 proteins, 100 proteins of which have the same fixed concentration: 100 µM. Green, red and blue curve corresponds to the concentration of the remaining protein to be 1 µM, 10 µM and 100 µM, respectively. The three curves overlap, indicating that changing the concentration of the remaining protein does not affect fold change of mRNA. D. Effects of the binding rates on specific gene expression responses. Wide orange curve is calculated using equation 3. Blue curve is calculated using differential equations with all experimentally determined “on” and “off” binding rates (Shifman et al., 2006; Pepke et al., 2010). The wide orange curve and the blue curve overlap, indicating a quasi-equilibrium state has established. The red and green curves correspond to the “on” and “off” binding rates are reduced by 100 fold and 1.0E4 fold, respectively.
Figure. 6. Two calcium signatures are decoded to generate specific expression of EDS1 and ICS1 following design principle: responses of two active signals, 4Ca²⁺-CaM-CAMTA3 and 4Ca²⁺-CaM-CBP60g, to two experimentally measured calcium signatures. A. Two empirical calcium signatures induced by two calcium agonists: mastoparan and extracellular calcium (Lenzoni et al., 2018). B. Enlargement of Fig. 6A, showing the details of the two calcium signatures. C. Response of 4Ca²⁺-CaM-CAMTA3 to the two calcium signatures. Wide solid orange curve and wide dashed orange curve are calculated using the two calcium signatures as the input of equation 3, respectively. Experimentally measured parameters are used. Black and blue curves are calculated using the two calcium signatures as the input of differential equations, respectively. The wide orange curve overlaps with the black curve. The wide dashed orange curve overlaps with the blue curve. These results indicate a quasi-equilibrium state is established. D. Same as Fig. 6C, but both “on” and “off” rate constants for Ca²⁺-CaM interactions are reduced by 1.0E5 fold from their experimental values. E. Same as Fig. 6C, but both “on” and “off” rate constants for Ca²⁺-CaM interactions are reduced by 1.0E8 fold from their experimental values. F. Same as Fig. 6C, but it is the response of 4Ca²⁺-CaM-CBP60g to the two calcium signatures. G. Same as Fig. 6C, but both “on” and “off” rate constants for Ca²⁺-CaM interactions are reduced by 1.0E5 fold from their experimental values. H. Same as Fig. 6C, but both “on” and “off” rate constants for Ca²⁺-CaM interactions are reduced by 1.0E8 fold from their experimental values. Black and blue curves are flat and they also overlap, indicating that neither calcium signature can induce changes in [4Ca²⁺-CaM-CBP60g].


