Ca$^{2+}$-based allosteric switches and shape shifting in RGLG1 VWA domain

Qin Wang$^{a,d}$, YaYu Chen$^b$, ShengPing Li$^c$, WenDi Yang$^b$, LiFang Sun$^b$, MeiQin Jang$^a$, XiuLing Wu$^a$, QianChao Wang$^a$, Lifei Chen$^b$, YunKun Wu$^b$,*

$^a$ State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, 155 Yangqiao West Road, Fuzhou 350002, China
$^b$ Provincial University Key Laboratory of Cellular Stress Response and Metabolic Regulation, College of Life Science, Fujian Normal University, Fuzhou 350117, China
$^c$ Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, Plant Immunity Center, Fujian Agriculture and Forestry University, Fuzhou 350002 China
$^d$ University of Chinese Academy of Sciences, Beijing 100049, China

**A B S T R A C T**

RGLG1 is an E3 ubiquitin ligase in Arabidopsis thaliana that participates in ABA signaling and regulates apical dominance. Here, we present crystal structures of RGLG1 VWA domain, revealing two novel calcium ions binding sites (NCBS1 and NCBS2). Furthermore, the structures with guided mutagenesis in NCBS1 prove that Ca$^{2+}$ ions play important roles in controlling conformational change of VWA, which is stabilized in open state with Ca$^{2+}$ bound and converted to closed state after Ca$^{2+}$ removal. This allosteric regulation mechanism is distinct from the ever reported one involving the C-terminal helix in integrin α and β domains. The mutation of a key residue in NCBS2 do not abolish its Ca$^{2+}$-binding potential, with no conformational change. MD simulations reveals that open state of RGLG1 VWA has higher ligand affinity than its closed state, consisting with integrin. Structural comparison of ion-free-MIDAS with Mg$^{2+}$-MIDAS reveals that Mg$^{2+}$ binding to MIDAS does not induce conformational change. With acquisition of first structure of plant VWA domain in both open state and closed state, we carefully analyze the conformational change and propose a totally new paradigm for its transition of open-closed states, which will be of great value for guiding future researches on VWA proteins and their important biological significance.

© 2020 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

In Arabidopsis thaliana, RGLG1 (RING domain ligase 1) is a RING-type E3 ubiquitin ligase that belongs to the RGLG family, and the four additional members are named RGLG2 to 5. RGLG1 contains a Von Willebrand factor A (VWA) domain at the center and a really interesting new gene (RING) domain at the C-terminal region [1]. It can be anchored to the plasma membrane by myristoylated Gly2, although it does not contain a transmembrane domain [1].

RGLG1 and its homolog RGLG2 play important roles in regulating auxin and cytokinin levels [1]. Numerous plant traits influenced by auxins can be changed through inactivation of both RGLG1 and RGLG2, e.g., abolishing apical dominance and altering normal leaf phyllotaxy. However, the effect of single rglg1 or rglg2 mutants are inconspicuous [1]. Rglg1 and rglg2 mutations alters the expression of genes regulated by cytokinin and auxin levels, e.g., cytokinin-responsive gene expression levels are higher in mutants than in wild type, whereas the auxin-responsive gene expression levels are lower. Furthermore, RGLG1 and RGLG2 regulate epidermal development, and rglg1 rglg2 mutants show a high percentage of branched root hair [2,3]. In addition, RGLG1 contributes to the abscisic acid (ABA) signaling pathway with another homologous protein RGLG5 [4]. Abscisic acid is an essential hormone for plant survival, and the type 2C protein phosphatases (PP2Cs) are critical negative regulators of the ABA signaling network [5,6]. PP2C, a PP2Cs family member, inactivates SNF1-related protein kinase 2 (SnRK2) through dephosphorylation [6]. RGLG1 and RGLG5 can interact with PP2CA and mediate its ubiquitination for proteasome degradation in vivo and in vitro [4]. The rglg1 and rglg5 double mutant shows reduced survival in drought stress conditions, whereas RGLG1 or RGLG5 overexpression reduces water loss and markedly enhances plant survival. With the closest sequences, phenotypes do not change when only one of RGLG1 and RGLG5 is mutated [4]. Interestingly, interaction between RGLG1/5 and PP2CA in vitro do not require ABA, whereas, ABA enhances the nuclear interaction with PP2CA in vivo [4,7]. The fact that ABA downregulate N-myristoyltransferase 1 (NMT1) to inhibit the myrist
toylation of RGLG1 and enhance its nuclear recruitment is determined by Belda-Palazon et al. [7]. Calcium ions and PP2CA protein also enhance nuclear localization of RGLG1. These evidences provide a mechanism that ABA enhances RGLG1–PP2CA interaction and hence PP2CA degradation [7]. VWA domains serve as interaction modules in many proteins, e.g., extracellular matrix proteins, integrin, and complement factors B and C2, and many human diseases result from their mutations [8]. They usually adopt a classic α/β “Rossmann” fold or a featured dinucleotide-binding conformation, which has several central β stands with one antiparallel edge sheet and several α helices around the β sheets [8]. Two terminal cysteine residues form a disulfide bridge to stabilize the domain [9]. Numerous VWA domains have an important metal-ion-dependent adhesion site (MIDAS) with three highly conserved non-contiguous elements (D-x-S/T-x-S...T...D, x represents any amino acid) playing central roles in ligand binding [10]. The three conserved elements are located in three different loops, with D-x-S/T-x-S being the first region, and the others T4 and D5, respectively. Not all VWA domains have perfect MIDAS motifs [8]. VWA domains in integrin α and β subunits (also called I domains or A domains) have two different conformations, open and closed, and α I domain has an additional intermediate conformation [10–16]. In the open conformation, corresponding to high affinity state, Mg2+ in the MIDAS motif has six coordination sites and forms an octahedral coordination with two serine residues of the first region, T4, two water molecules (o1 and o2), and a ligand or pseudo ligand [12,16]. The two aspartic acid residues occupy a secondary coordination sphere. Their carboxylate groups indirectly coordinate the metal ion by forming a o1-mediated hydrogen bond [17]. In the closed conformation, corresponding to low affinity state, coordination bonds with the first region are unchanged, resulting from ionic displacement in the MIDAS motif; the coordination bonds in T4 are disrupted and replaced by a direct bond with D5, and the ligand is replaced by another water molecule [11]. Furthermore, the transition between open and closed conformations of the integrin αM I domain are considered to be associated with the rearrangement of a C-terminal α helix [11,12]. This conformational change is apparently paradigmatic for VWA domains.

Two additional ion-binding sites—ADMIDAS (adjacent to MIDAS) and SyMBS (synergistic metal ion-binding site) exist in the integrin β subunit I domains. They regulate ligand-binding affinity [18]. At a high Ca2+ concentration, ADMIDAS negatively regulates integrin activation [18–22]. Mutagenesis studies have reported that ADMIDAS exerting contrasting effects on ligand binding in different integrins. SyMBS positively regulates the ligand–integrin binding at low Ca2+ concentrations [18]. Ca2+ is predicted to have a higher affinity than Mg2+ for the SyMBS because the SyMBS has two backbone and one amide side chain carbonyl oxygen coordination, and Ca2+ has a far greater propensity than Mg2+ to form a coordination with backbone carbonyl oxygen [23,24].

The most ancient VWA-containing proteins are all intracellular proteins involved in various important physiological phenomena including transcription, membrane transport, DNA repair, and the proteasome pathway [8]. Subsequently, metazoan develop extracellular VWA proteins. However, the intracellular complement of plants expands exclusively for novel biological functions [8]. Thus far, integrin I domains are the only examples for clarifying the mechanism of VWA conformational changes. Many other VWA proteins including C2a and factor B, the anthrax receptor, and Ro, only have the open conformation, and closed conformation has not yet been resolved. With low sequence homology among integrin and plant VWA, different allosteric regulation mechanisms may mediate transitions in the open and closed conformations. Therefore, more structures for both the open and closed state of VWA proteins are crucial to decipher the different mechanism. This study firstly reports the crystal structures of the plant VWA domain with two different conformations and shows novel paradigms for conformational changes. In the RGLG1 VWA domain, two unexpected Ca2+-binding sites have been determined. Furthermore, removing Ca2+ from NCBS1 successfully cause the corresponding allosteric switches, thus revealing the mechanism of conformational changes distinct from integrins, and facilitating future studies on VWA-containing proteins.

2. Results

2.1. Overall structure of the RGLG1 VWA domain

The crystal structure of the RGLG1129 (resides R129–Q410) was resolved via selenomethionyl single-wavelength anomalous dispersion phasing at 2.2 Å resolution, containing all residues of the VWA domain (N156-L376). Crystals grew in space group P41, with two molecules per asymmetric unit. Data collection and refinement statistics are summarized in Table 1. Similar to other members of VWA domains, the RGLG1 VWA contains a typical Rossmann fold, a central five-stranded hydrophobic β sheet surrounded by nine amphipathic α helices, β sheets and α helices are labeled β1 to β5 and α1 to α9, respectively (Fig. 1a and b). A conserved MIDAS motif located at the top face comprises a DxXS sequence (residues 162–166) from the β1-α2 loop and α2 helix, T258 from the α4-α5 loop, and D286 from the β3-α6 loop. The strong globular electron density presented at MIDAS is modelled as a Mg2+ ion, since it is the only metal ion in the crystallization buffer, and distinguished from water in accordance with bond lengths and octahedral coordination. Mg2+ is coordinated by hydroxyl groups of MIDAS residues T164, S166, T258, and two water molecules, a carboxylate oxygen from E153 of the neighbor domain serving as a pseudo ligand to complete the octahedral coordination. D162 and D286 indirectly coordinate Mg2+ by forming hydrogen bonds with one water molecule (Fig. 1). Compared with the conformation of the MIDAS residues in integrin αM I domain, we determined that RGLG1129 assumed an open conformation [12]. Unlike numerous VWA domains, RGLG1129 is devoid of an intramolecular disulfide bond between the N- and C-termini that stabilize the structure [9,12]. In addition, a ubiquitous anti-parallel β sheet is also absent in the RGLG1129 structure [12,25–27] (Supplementary Fig. 1).

2.2. Two novel Ca2+ binding sites

Distinct from other VWA domains, two Ca2+ binding sites are present in RGLG1129, one located between α7 and α9 bound with two Ca2+ ions is termed novel calcium ion binding site 1 (NCBS1) and two Ca2+ ions are termed Ca1 and Ca2 (Fig. 2a and b). The other site at the bottom face, binding one Ca2+ ion, is termed NCBS2 (Fig. 2c). All three Ca2+ ions adopt pentagonal bipyramidal coordination, with seven coordinating ligands, and the Ca2+-O bond distances are approximately 2.4 Å. Coordination motifs are shown in Table 2. The coordination stereochemistry confirmed the excess electron densities at NCBS1/2 as Ca2+.

ADMIDAS and SyMBS are two Ca2+ binding sites in the integrin β I domain, however, they are significantly different from NCBS1/2. In the integrin β I domain, three ion binding sites form an interlinked linear array, MIDAS occupying the central position and flanked by ADMIDAS and SyMBS [18]. In RGLG1129, NCBS1/2 are distant from MIDAS, NCBS1 located at the flank and NCBS2 present almost opposite to MIDAS. In addition, the coordination motifs are different. Ca2+ in ADMIDAS exhibited distinct coordination bonds in different conformations, e.g., in the closed state of αβ1, Ca2+ adopt the pentagonal bipyramidal, forming seven coordinate
### Table 1
Data collection and refinement statistics.

| Property | RGLG1<sup>125</sup> | RGLG1<sup>110</sup>-H<sub>2</sub>O | RGLG1-Ca<sup>2+</sup> | RGLG1<sup>D338A</sup> | RGLG1<sup>D338A/E378A</sup> | RGLG1<sup>1D338A/E378A</sup> | RGLG1<sup>1D338A</sup> | RGLG1<sup>D338A</sup> | RGLG1<sup>E378A</sup> |
|----------|----------------------|-------------------------------|-------------------|------------------|--------------------------------|--------------------------------|------------------|------------------|------------------|
| Space group | P 4 1 | P 4 1 | P 2 1 2 1 | P 2 1 2 1 | P 2 1 2 1 | P 2 1 2 1 | P 2 1 2 1 | P 2 1 2 1 | P 2 1 2 1 |
| Cell constants a, b, c, α, β, γ | | | | | | | | | |
| a, Å | 136.18 | 136.18 | 42.88 | 45.59 | 45.69 | 44.45 | 45.62 | 42.79 | 41.12 |
| b, Å | 70.73 | 70.73 | 80.85 | 90.00 | 90.00 | 90.00 | 90.00 | 90.00 | 42.18 |
| c, Å | 56.86 | 56.86 | 90.00 | 90.00 | 90.00 | 90.00 | 90.00 | 90.00 | 90.00 |
| Resolution (Å) | 48.96–2.21 | 48.01–2.40 | 37.00–1.79 | 28.39–1.59 | 39.72–1.61 | 41.14–1.40 | 40.04–1.69 | 37.98–1.50 | 46.51–2.39 |
| Unique reflection | 52,083 | 39,684 | 50,373 | 36,579 | 34,983 | 53,506 | 30,559 | 39,684 | 114,083 |
| % Data completeness (in resolution range) | 98.9 (48.96–2.21) | 98.7 (48.01–2.40) | 94.0 (37.00–1.79) | 95.2 (28.39–1.59) | 94.2 (39.72–1.61) | 90.5 (41.14–1.40) | 99.1 (40.04–1.69) | 93.1 (37.98–1.50) | 98.5 (46.51–2.39) |
| R<sub>merge</sub> | 0.15 | 0.22 | 0.17 | 0.15 | 0.17 | 0.08 | 0.14 | 0.19 | 0.16 |
| <I = r<sup>2</sup> | 2.81 (at 2.20 Å) | 1.64 (at 2.39 Å) | 1.50 (at 2.39 Å) | 1.02 (at 1.78 Å) | 1.08 (at 1.50 Å) | 1.24 (at 1.40 Å) | 1.16 (at 1.40 Å) | 1.14 (at 1.50 Å) | 2.02 (at 2.39 Å) |
| R, R<sub>free</sub> | 0.168, 0.190 | 0.187, 0.226 | 0.196, 0.234 | 0.200, 0.235 | 0.197, 0.224 | 0.200, 0.217 | 0.190, 0.226 | 0.195, 0.228 | 0.209, 0.265 |
| Wilson B-factor (Å<sup>2</sup>) | 37.2 | 45.1 | 16.9 | 30.5 | 28.6 | 20.1 | 24.5 | 19.7 | 39.2 |
| Bulk solvent ksol (e/Å<sup>3</sup>) | 0.34, 29.6 | 0.32, 29.5 | 0.30, 43.2 | 0.34, 41.9 | 0.34, 41.6 | 0.35, 37.3 | 0.36, 41.1 | 0.34, 40.2 | 0.33, 34.9 |
| Fo,Fc correlation | 0.96 | 0.95 | 0.94 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.93 |
| Total number of atoms | 4664 | 4510 | 3984 | 2380 | 2383 | 2472 | 2472 | 2539 | 17,985 |
| Average B, all atoms (Å<sup>2</sup>) | 41.0 | 50.0 | 22.0 | 38.0 | 32.0 | 25.0 | 28.0 | 23.0 | 43.0 |
| Ramachandran plot | | | | | | | | | |
| Favored region (%) | 97.5 | 95.3 | 95.9 | 96.4 | 97.1 | 97.1 | 97.9 | 97.5 | 93.6 |
| Outliers (%) | 0.6 | 0.5 | 0.7 | 0 | 0 | 0 | 0 | 0 | 1.3 |
| PDB code | 6K88 | 6K88A | 6K88 | 6K86 | 6K85 | 6K82 | 6K89 | 6K87 | 6K83 |

---

**Fig. 1.** Structure of the RGLG1 VWA domain. (a) Ribbon diagram of the RGLG1 VWA domain, and the α helices and β strands are labeled α1 to α9 and β1 to β5, respectively. The Mg<sup>2+</sup> ion is shown as a chartreuse sphere, and three Ca<sup>2+</sup> ions are shown as light blue spheres. (b) Same as (a) but y is rotated by 90°. (c) Stereo diagram of the MIDAS site, wherein the Mg<sup>2+</sup> ion is shown as a chartreuse sphere, and two bound water molecules are depicted as small red spheres. The E153 residue from a neighboring molecule contributes the sixth coordinating residue. D286 and D162 coordinate Mg<sup>2+</sup> indirectly through hydrogen bonds (the blue dotted line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Fig. 2.** Stereo diagram of NCBS1 and NCBS2. (a) The coordination residues for Ca1 (residue numbers are as indicated) are shown. The Ca<sup>2+</sup> ion is depicted by a light blue sphere, and water molecules are depicted as small red spheres. (b) Ca2 is coordinated by three water molecules (small red spheres), Q350, D338, and E378. (c) Ca2+ ion at NCBS2 is coordinated by two water molecules, D207 and D162 coordinate Mg<sup>2+</sup> indirectly through hydrogen bonds (the blue dotted line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
bonds with a bidentate coordination at D137, while in the open state of αIIbβ3, Ca2+ is more likely to form octahedral conformation, and assuming monodentate coordination with Asp [15,19]. Ca2+ in SyMB5 adopts octahedral coordination with non-bidentate bonds in both closed and open states of integrin α5β1 [28]. While NCBS1 and NCBS2 form pentagonal bipyramidal and bidentate coordination regardless of the conformation.

2.3. Conformational changes due to Ca2+ removal from NCBS1

To investigate whether mutations at key residues in NCBS1 impedes Ca2+ binding, the crystal structures of the D338A, E378A, and D338A/E378A mutants of RGLG1 were determined. All structures of mutants present one molecule in an asymmetric unit and they are nearly identical, however, these structures exhibit obvious differences from the open conformation. Mutants inhibit Ca2+ binding and result in a 22 Å 2 (Fig. 3f), implying that direct elimination of Ca2+ from NCBS1 and NCBS2 result in a 22 Å 2 (Fig. 3f), implying that direct elimination of Ca2+ from NCBS1 via mutations or EGTA results in a shift of α7, then the phenylindole ring of W331 moves 2.2 Å closer to the α6-β3 loop, and causes a movement in the αC of G287 by 2.8 Å and the carbonyl group of G287 rotates by 90°. Thereafter, side chain of D286 shifted by 2.1 Å, facilitating direct linkage with Mg2+ (Fig. 3g). A water molecule replaces the pseudo ligand to complete the coordination, and conformation is converted to closed state. This Ca2+-dependent RGLG1 VWA conformation shows that the mechanism for transition of open-closed states is different from integrins.

2.4. NCBS2 mutation has no effects on conformational changes

Since NCBS1 is the allosteric regulatory site of RGLG1 VWA, is NCBS2 the same? Thus the mutations D205A and D207G were evaluated, and we find that the D207G mutation in NCBS2 did not affect conformational conversion. The crystal structure of RGLG1D207G was resolved using molecular replacement in PHASER (Table 1). RGLG1D207G does not lose its binding potential for Ca2+ and does not influence conformation. However, the mutation disrupts Ca2+ coordinate bonds, and the coordination geometry is changed from pentagonal bipyramidal to octahedral coordination (Fig. 3a and b). The structure of D205 mutant cannot be obtained owing to poor electron density. In all closed conformations, NCBS2 binds Na+ (Fig. 4c), except that RGLG1D338A binds Mg2+ (Fig. 4d). The Na+ ion has six ligands: three ligands are hydroxyl oxygen atoms of D205, D207, and N156, and three are water molecules (Fig. 4c). In RGLG1D207G, the shift of Mg2+ away from N156 alters its coordination: a water molecule replaces N156 to bind Mg2+ (Fig. 4e); therefore, Mg2+ in NCBS2 is coordinated by D205, D207, and four water molecules (Fig. 4d). Our hypothesis that the reason of the D207G mutation does not influence conformation was that Ca2+ could not be deprived, since NCBS2 binds Ca2+ in all open states but loses Ca2+ in all closed states. NCBS2-bound Ca2+ potentially contributes to conformational stability; however, further studies are required to characterize the underlying mechanisms.

2.5. Structure of Ca2+ bound MIDAS

RGLG1-Ca2+, containing residues S155-P380, was resolved at 1.8 Å resolution through molecular replacement, considering RGLG1D207G as a reference model (Table 1). Two molecules are present in an asymmetric unit. The structure adopts an open conformation with a Ca2+ ion occupying MIDAS (Fig. 5a). RGLG1-Ca2+ was remarkably similar to RGLG1T29, presenting an RMSD of merely 0.57 Å for all Ca atoms, and exhibiting highly similar in MIDAS (Fig. 5c). However, instead of E153, E296 of the neighboring molecule serves as a pseudo ligand (Fig. 5a). Direct contact

| Cation binding site | Coordination motif | Coordination motif | Coordination motif |
|---------------------|--------------------|--------------------|--------------------|
| NCBS1 Ca1 | bidentate carboxylate ligands of D347 | monodentate carboxylate ligands of D338 and E378 | main chain carbonyl oxygen atoms of F349 and D338 | one water molecule |
| NCBS1 Ca2 | bidentate carboxylate ligands of D338 | monodentate carboxylate ligand of E378 | side chain carbonyl oxygen atom of Q350 | three water molecules |
| NCBS2 | bidentate carboxylate ligands of D205 | monodentate carboxylate ligand of D207 | main chain carbonyl oxygen atom of L209 | side chain of N156 | two water molecules |
between Ca\textsuperscript{2+} and the backbone oxygen of Q288 is facilitated by a 0.4 Å shift of Ca\textsuperscript{2+} from the position of Mg\textsuperscript{2+} (Fig. 5a and b). Other coordinating residues for Ca\textsuperscript{2+} are the same as those of Mg\textsuperscript{2+} (Fig. 5c and d). A Ca\textsuperscript{2+} ion was assigned at this position in accordance with the following observations. First, Ca\textsuperscript{2+} is the only metal ion in the crystallization buffer. Second, the average distance between the metal ion and the coordinating ligands (2.42 Å) is close to Ca\textsuperscript{2+}. Third, Ca\textsuperscript{2+} preferentially forms a pentagonal bipyramidal rather than an octahedral coordination, whereas Mg\textsuperscript{2+} and Mn\textsuperscript{2+} preferentially form an octahedral coordination, and Mg\textsuperscript{2+}/Mn\textsuperscript{2+} does not form coordinate bonds with backbone oxygen atoms. Therefore, the typical pentagonal bipyramidal and backbone oxygen coordination distinguish the ion from Mg\textsuperscript{2+}/Mn\textsuperscript{2+}. Hypothesis that a high Ca\textsuperscript{2+} concentration might result in a competition with Mg\textsuperscript{2+}/Mn\textsuperscript{2+} and inhibit substrate binding has been proposed. However, some studies disagree with this hypothesis because Ca\textsuperscript{2+} is too large to bind MIDAS [12,29,30]. Our data show that the MIDAS of the RGLG1 VWA domain can be occupied by Ca\textsuperscript{2+} in the open conformation, inducing minimal global structural changes. And, RGLG1-Ca\textsuperscript{2+} showed a different coordination with...
Ca\textsuperscript{2+} than those reported previously for Ca\textsuperscript{2+}-bound VWA domains [13,19]. In addition, RGLG1-Ca\textsuperscript{2+} binds the pseudo ligand in the open conformation; however, most Ca\textsuperscript{2+}-bound VWA domains adopt the closed conformation, and the rest are unliganded.

2.6. Open-closed conformation transition independent of Mg\textsuperscript{2+}

An open conformation was observed (RGLG1 open-H\textsubscript{2}O) with MIDAS occupied by a water molecule. Although the H\textsubscript{2}O-E\textsubscript{153} distance increased to 3.1 Å, E\textsubscript{153} of neighboring molecules is still sufficient to coordinate RGLG1 open-H\textsubscript{2}O as a pseudo ligand at least under non-physiological crystallization conditions (Fig. 6a). This is the first ion-free ligand-bound open conformation of VWA domains. In ion-free-MIDAS in C\textsubscript{2}a-Li\textsuperscript{+}, the side chain of S\textsubscript{244} (equivalent to S\textsubscript{166} in RGLG1) was rotated by approximately 90˚ to compensate for the loss of charge [31]. In the disulfide-engineered C\textsubscript{2}a-Li\textsuperscript{+} domain, D\textsubscript{239} (equivalent to D\textsubscript{286} in RGLG1) rotates distally to compensate for the loss of charge [32]. In our structures, superposition of RGLG1 with RGLG1 open-H\textsubscript{2}O shows a virtually identical conformation in the side chain of S\textsubscript{166}, D\textsubscript{286}, and all other MIDAS residues (Fig. 6b and c). In addition, the ion-free closed conformation (RGLG1 closed-H\textsubscript{2}O) (Fig. 6d) also shows only minor rearrangements in comparison with RGLG1 E\textsubscript{378}A (Fig. 6e).

This study shows that the state of RGLG1 VWA is not influenced by Mg\textsuperscript{2+} occupancy in MIDAS, since MIDAS can be occupied by a Mg\textsuperscript{2+} ion or a water molecule, and these structures display an identical state, concurrent with previous findings that Mg\textsuperscript{2+} is not sufficient to induce an open state [33].

2.7. Molecular dynamics simulations of affinity between RGLG1 and PP2CA

Previous results support the fact that conformations of the integrin I domains regulate the ligand binding potential, with the closed and open conformations corresponding to low and high affinity, respectively [32]. Unfortunately, a complete ligand-bound RGLG1 complex was not obtained; hence, the affinity remains mystery. Since PP2CA is a central negative regulator in the ABA signaling network and a substrate of RGLG1, it can be used to evaluate the affinity of RGLG1 in different conformations through MD simulations. Binding free energy (\(\Delta G_{bind}\)) was determined to evaluate their energetic aspects. The \(\Delta G_{bind}\) values of the open state and closed state of RGLG1/PP2CA were -101.97 and -89.17 kcal mol\textsuperscript{-1}, respectively (Table 3). As expected, the results of MD simulations suggest that the open state has much higher ligand affinity than the closed state. According to the energy components of the binding free energy (see Eqs. (2) and (3)), electrostatic interaction (\(\Delta E_{ele}\)) is pivotal for the affinity of RGLG1/PP2CA. In addition, van der Waals interaction (\(\Delta E_{vdw}\)) also played a role to enhance the stability of the system. However, these values are obtained through simulations and further evidences will be required to validate them.

3. Discussion

RGLG1, as a critical regulator in apical dominance and the ABA signaling pathway, has been evaluated via biochemical, immunological, and cell biological approaches [1,4]. Our study shows...
that the conformations of the RGLG1 VWA domain are Ca\textsuperscript{2+}-dependent. All structures of RGLG1 VWA adopt a typical Rossmann fold without the anti-parallel \( \beta \) sheet normally present in VWA, with sheet topology 213456 rather than the common sheet topology 321456. Sheet topology 21345 has not been discovered in VWA domains, whereas cutinases, a serine esterase, possesses similar sheet topology [34]. In integrin \( \alpha I \) domains, the C-terminal helix and the preceding loop allosterically cause conformational switches in MIDAS[32]. In RGLG1 VWA, its conversion is independent of the movement of the C-terminal helix. Moreover, two novel Ca\textsuperscript{2+} binding sites (NCBS1 and NCBS2) are present, and through structure-guided mutagenesis experiments, we prove that the shift in conformation is clearly dependent on the occupancy of Ca\textsuperscript{2+} at NCBS1. Thus far, in all VWA domains, only integrin I domains have revealed a paradigm of conformational change[26]. Notably, the RGLG1 structures provide novel insights into the Ca\textsuperscript{2+}-based allosteric switch and conformational changes. After binding Ca\textsuperscript{2+}, the deflection of \( \alpha I \) disrupts the interaction between 286D and Mg\textsuperscript{2+}, and the movement of Mg\textsuperscript{2+} ion allows it to be coordinated by 258T. The pseudo ligand completes the coordination, then the conformation is changed to the open state. The only effect of mutations at D338 and E378 for RGLG1, whether single or double mutation, is inhibiting Ca\textsuperscript{2+} binding. It’s worth noting that the allosteric regulator for the RGLG1 VWA domain is Ca\textsuperscript{2+} ions, not mutations. Mutations at key residues in NCBS1 only cause the corresponding allosteric switches through inhibiting Ca\textsuperscript{2+} binding. This result is proved by the structure of RGLG1\textsuperscript{EGTA}, which is crystallized in the closed conformation with 2 mM EGTA supplemented in the purification buffer to eliminate calcium ions, avoiding any mutations.

Our findings are consistent with Belda-Palazon et al., who have proposed that ABA and calcium ions promote nuclear localization of the RGLG1 [7]. We speculate that ABA treatment may induce calcium-based allosteric switch of RGLG1 and convert RGLG1 to open state due to ABA can lead to increased Ca\textsuperscript{2+} levels. As we all known, ABA and calcium signaling processes are integrated and interconnected[35], therefore, the effect of ABA may be achieved through Ca\textsuperscript{2+} ions. In vitro, RGLG1 adopt the open state that can directly bind ligands, which can explain why ABA is not required for the interaction in vitro. Structural comparison of RGLG1 open state and closed state shows that a significant change occurs in the N-terminal \( \alpha I \), and during RGLG1 shuttle to the nucleus, specific plant proteases might catalyze the hydrolysis of N-myristoylated glycine residues [7]. Thus, we hypothesize that the movement of the \( \alpha I \) can cause variable in N-terminal region, allowing RGLG1 to bind to the specific proteases, and facilitating its shuttling and interaction with its target.

To investigate the effect of Ca\textsuperscript{2+} on the structural dynamics of NCBS1 residues, referring to Espinoza-Fonseca et al. [36], we performed four independent MD simulations based on the crystal structure of RGLG1\textsuperscript{129} and RGLG1\textsuperscript{EGTA}, starting with the presence or absence of initially-bound Ca\textsuperscript{2+}. All simulations were performed for 100 ns. We plotted time-dependent distance evolution of the carboxyl-carboxyl or carboxyl-acylamino pairs between key resi-

---

Fig. 5. Structure of Ca\textsuperscript{2+} binding MIDAS. (a) In the RGLG1-Ca\textsuperscript{2+} structure, Ca\textsuperscript{2+} (bright orange sphere) in MIDAS is coordinated by E296 from adjacent domain, S166, T164, T258, Q288, and two water molecules. The blue dotted lines represent hydrogen bonds. (b) In RGLG1\textsuperscript{129}, Q288 is too far away from Mg\textsuperscript{2+} (3.1 Å) to coordinate it. (c) Expansion of (d) around MIDAS. The ligand for Ca\textsuperscript{2+} (bright orange sphere) in RGLG1-Ca\textsuperscript{2+} is identical to that of Mg\textsuperscript{2+} (chartreuse sphere) in RGLG1\textsuperscript{129}, except for Q288. However, there is no significant shift in MIDAS residues. (d) The structure of RGLG1-Ca\textsuperscript{2+} (salmon) is superimposed onto the structure of the RGLG1\textsuperscript{129} (aquamarine). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
dues D338, E378, D347 and Q350 in the 100 ns-long MD simulations. Starting from the RGLG1 crystal structure, we found that the Ca\(^{2+}\) ion remained binding to NCBS1 during the entire 100 ns simulation (Fig. 8 and Supplementary Fig. 2). In Ca\(^{2+}\)-free-RGLG1 (the crystal structure of RGLG1 but removing Ca\(^{2+}\)), the distances of D338-E378, D338-Q350 and D347-E378 are larger compared to that of RGLG1, with an upward trend. Interestingly, the distance between D338 and D347 had large fluctuation during 30 ns–50 ns time scale, and finally went down to a similar distance scale of RGLG1-EGTA (Fig. 8). All distances in Ca\(^{2+}\)-free-RGLG1 are similar to RGLG1-EGTA, calculated from the RGLG1-EGTA crystal structure, indicating that Ca\(^{2+}\)-free-RGLG1 can convert to closed state (Fig. 8). Furthermore, new trajectory was performed using the final frame of 100 ns production run of the Ca\(^{2+}\)-free-RGLG1 as starting configuration, as expected, Ca\(^{2+}\) induced conformation change to closed state (Fig. 8). Dependent on the analyses of the MD trajectories, we propose that VWA domain can maintain open state in the presence of initially-bound Ca\(^{2+}\), otherwise it adopts the closed state. These studies provide detailed insights on Ca\(^{2+}\) binding and allosteric coupling of domain dynamics. Furthermore, we calculated the interdomain distance distributions of Ca\(^{2+}\)-free-RGLG1, RGLG1-EGTA and RGLG1-con to analyze the structural dynamics using Ca-Ca distance of L140-D338. The atom-pair distance distributions fit very well to an either one or two Gaussian distribution (Fig. 9). Noticeably, two peaks were both found in Ca\(^{2+}\)-free-RGLG1 and RGLG1-con, corresponding to the open and closed conformations. These results show that Ca\(^{2+}\) ions control the dynamic equilibrium of structural ensembles.
Sequence alignment of RGLG1 using ClustalW revealed that copines and other RGLG family members share high sequence identity with RGLG1 at the VWA domain. Copines are phospholipid-binding proteins with three family members in *Arabidopsis thaliana* and nine orthologues in *Homo sapiens*. Each copine contains two N-terminal C2 domains followed by a C-terminal VWA domain in both *Homo sapiens* and *Arabidopsis thaliana*. The C2 domains can respond to the increases of intracellular calcium levels to bind to the phospholipid. In *Arabidopsis thaliana*, copines play important role in growth regulation and disease resistance.

**Fig. 7.** Sequence alignment of the calcium binding sites between RGLG family and copines. The sequences were aligned with CLUSTALW. BONZAI1-3 are three copines family members from *Arabidopsis Thaliana*. Copine1-9 are nine copines family members from *Homo sapiens*. Key residues in NCBS1 and NCBS2 (D205, D207, D338, D347, and E378) are marked with black pentacles. All key residues are highly conserved in RGLG family, *Homo sapiens* copines family and *Arabidopsis Thaliana* copines family.
In *Homo sapiens*, most copines are expressed ubiquitously and are primarily involved in membrane trafficking of intracellular signaling proteins [38]. Further comparison of copines with RGLG1 on key residues (D205, D207, D338, D347, and E378) in NCBS1/2 reveals that these residues are highly conserved in copines (Fig. 7). It is hypothesized that these copines may also contain the two novel Ca\(^{2+}\) binding sites (NCBS1/2) and their conformational changes depend on Ca\(^{2+}\) as well, thus facilitating further examination of the structural and functional properties of copines. Our structures potentially provide a suitable paradigm for future studies investigating novel mechanisms for other VWA proteins.

Ca\(^{2+}\) serves as an allosteric regulator for the RGLG1 VWA domain has been determined in this paper. Previous studies have shown that open state represents high affinity in integrin VWA domain. According to the results of our MD simulation, open state of RGLG1 VWA also represent higher affinity. In other words, after binding Ca\(^{2+}\), RGLG1 VWA can bind its ligand, e.g., PP2CA, thereafter, the ABA signaling pathway may be activated upon PP2CA degradation through RGLG1. However, the mechanism of calcium-mediated protein interactions *in vivo* is still unknown, thus, future studies are required.

4. Methods
4.1. Protein preparation and crystallization

RGLG1\(^{129}\) and RGLG1-Ca\(^{2+}\) was cloned into a bacterial expression vector PET32a to generate a fusion protein containing a 6his tag, an N-terminal Trx tag, and a TEV cleavage site between the 6his tag and the target protein. *E. coli* BL21 cells were transformed

| Table 3 |
| --- |
| Binding free energy (\(\Delta G_{\text{bind}}\)) calculated for RGLG1/PP2CA complexes (units in kcal mol\(^{-1}\)). |
| \(\Delta E_{\text{ele}}\) & \(\Delta E_{\text{v ivdW}}\) & \(\Delta E_{\text{sol}}\) & \(\Delta E_{\text{ele}}\) & \(\Delta G_{\text{bind}}\) |
| RGLG1\(^{129}\) & -299.44 ± 42.04 & -107.99 ± 7.37 & 305.4 ± 35.99 & 17.60 ± 11.62 & -101.97 ± 6.49 |
| RGLG1\(^{\text{EGTA}}\) & -228.21 ± 23.75 & -91.63 ± 6.50 & 228.69 ± 21.62 & 11.95 ± 3.67 & -91.15 ± 3.20 |
| RGLG1\(^{D338A}\) & -184.20 ± 39.14 & -104.91 ± 5.68 & 192.99 ± 30.58 & 20.93 ± 9.84 & -96.11 ± 4.11 |
| RGLG1\(^{E378A}\) & -201.53 ± 28.60 & -84.96 ± 7.31 & 208.24 ± 33.40 & 16.39 ± 5.81 & -78.25 ± 2.67 |
| RGLG1\(^{D338A/E378A}\) & -230.24 ± 18.04 & -91.30 ± 6.67 & 230.38 ± 18.76 & 10.64 ± 4.93 & -91.16 ± 2.27 |

Fig. 8. Time-dependent distance evolution of carboxyl-carboxyl or carboxyl-acylamino pairs between residues D338, E378, D347 and Q350. Distances were calculated for MD simulations from the crystal structures of RGLG1 (RGLG1\(^{129}\), RGLG1\(^{129}\)-con, Ca\(^{2+}\)-free-RGLG1\(^{129}\) and RGLG1\(^{\text{EGTA}}\)).
with the expression vector, and cells were cultured in 1 L LB medium up to an A600 of 0.4–0.6 at 37°C with the expression vector, and cells were cultured in 1 L LB medium. Distance distributions between L140-D338 of RGLG. MD trajectories of Ca2+-free-RGLG129 (green), RGLG129 (blue), RGLG129EGTA (black) and RGLG129EGTA were used to calculate Ca-Ca distance distribution between L140-D338. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. Data collection, structure determination, and refinement

Before data collection, all crystals were briefly soaked with cryo-protectant and flash-frozen in liquid nitrogen. X-ray diffraction data were obtained at Shanghai Synchrotron Radiation Facility (SSRF), processed with HKL2000 package or Xia2. RGLG129 phases were obtained via single-wavelength anomalous scattering experiments and analyzed using PHENIX AutoSol. Other structures were resolved using RGLG129 as the molecular replacement model in PHASER. Further refinements were carried out using PHENIX and manual model construction in Coot.

4.3. Molecular dynamics simulation

The initial configurations of the RGLG1 and PP2CA system for Molecule dynamics (MD) simulations were generated using ZDOCK 3.0.2 [40]. MD simulations were carried out using GROMACS-2018.3, with the AMBER ff99SB-ILDN force field [41,42]. The protonation states of protein residues were assigned using PROPKA web-server on the basis of predicted pKa values [43]. All systems were explicitly solvated with TIP3P water molecules in 9 Å buffer and neutralized with counter ions. The particle mesh Ewald (PME) algorithm [44] was used to determine long-range electrostatic energies, and van der Waals and Coulomb interactions were truncated at 10 Å. All bond lengths were constrained using the LINCS algorithm.

The systems were subjected to 2500 steps of steepest-descent and 2500 steps of conjugate-gradient minimization to eliminate poor interatomic contacts. The system was first gradually heated from 0 K to 298 K and then equilibrated for 500 ps at 1 atm in an isochoric/isothermal (NPT) ensemble with periodic boundary conditions. Temperature and pressure controls were achieved using Nosé–Hoover thermostat and Berendsen barostat with a frequency of 2.0 ps, respectively [45,46]. Finally, 150 ns-long MD simulations were carried out with each system. The most representative structures were identified via cluster analysis.

4.4. Binding free energies

The molecular mechanics Poisson Boltzmann solvent accessible surface area (MM/PBSA) approach was successfully applied to predict the binding free energy (ΔGbind) in various protein–ligand or protein–protein complexes [47–49]. Here, three parallel MD simulations of each system were carried out and the first 100 snRGLG1hots were extracted from each MD trajectory of each complex. Computational details were as follows:

\[
\Delta G_{\text{bind}} = \Delta H - T \Delta S = \Delta G_{\text{gas}} + \Delta G_{\text{solv}} - T \Delta S
\]

\[
\Delta G_{\text{gas}} = \Delta G_{\text{ele}} + \Delta G_{\text{vdW}}
\]

\[
\Delta G_{\text{solv}} = \Delta G_{\text{pol}} + \Delta G_{\text{np}}
\]

where ΔGgas, ΔGsolv and -TΔS represent the changes in binding energy in the gaseous phase, solvation, and the conformational entropy upon binding, respectively. ΔGgas included ΔGele (electrostatic) + ΔGvdw (van der Waals) potential. ΔGsolv included contributions of polar (ΔGpol) and nonpolar (ΔGnp) terms. Since the binding conformational entropy (−TΔS) was computationally expensive and poorly accurate, we assumed that ΔGbind was approximately equal to the sum of ΔGgas and ΔGsolv.

CRediT authorship contribution statement

Qin Wang: Investigation, Writing - original draft, Writing - review & editing, Conceptualization, Methodology, Visualization, Formal analysis. YaYu Chen: Software, Methodology. ShengPing Li: Resources. WenDi Yang: Writing - review & editing. LiFang Sun: Methodology, Funding acquisition. MeiQin Jang: Methodology. XiuLing Wu: Methodology. QianChao Wang: Methodology. LiFei Chen: Methodology. YunKun Wu: Supervision, Funding acquisition.

Declaration of Competing Interest

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

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2020.03.023.

References

[1] Yin X et al. Ubitquitin-lysin 63 chain forming ligates regulate apical dominance in Arabidopsis. Plant Cell 2007;19(6):1898–911. https://doi.org/10.1105/tpc.107.052020

[2] Li W, Schmidt W. A lysine-63-linked ubiquitin chain-forming conjugase, UBC13, promotes the developmental responses to iron deficiency in Arabidopsis roots. Plant J 2010;62(2):330–43. https://doi.org/10.1111/j.1365-313X.2010.04156.x

[3] Pan IC, Schmidt W. Functional implications of K63-linked ubiquitination in the iron deficiency response of Arabidopsis roots. Front Plant Sci 2014;4. https://doi.org/10.3389/fpls.2014.00027

[4] Wu Q et al. Ubitquitin ligases RGL1 and RGL5s regulate asbiscic acid signaling by controlling the turnover of phosphatase PP2CA. Plant Cell 2016;28(9):2178–96. https://doi.org/10.1105/tpc.16.00364

[5] Cutler SR, Rodriguez PL, Finkenstein RR, Abrams SR. Abscisic acid: emergence of a core signaling network. Annu Rev Plant Biol 2010;61:651–79. https://doi.org/10.1146/annurev-arplant-042009-112122

[6] Umezawa T et al. Type ZC protein phosphatases directly regulate asbiscic acid-activated protein kinases in Arabidopsis. Proc Natl Acad Sci 2005;102(41):17588–93. https://doi.org/10.1073/pnas.0507095102

[7] Belda-Palazon B et al. ABA inhibits myrosinolatation and induces shutting of the RGL1/3 Es ligase to promote nuclear degradation of PP2CA. Plant J 2019;95(5):813–25. https://doi.org/10.1111/tpj.14274

[8] Whittaker CA, Hynes RO, Pollard TD. Distribution and evolution of von willebrand/integrin a domains: widely dispersed domains with roles in cell adhesion and elsehne. Mol Biol Cell 2002;13(10):3369–87. https://doi.org/10.1091/mbc.02-05-0259

[9] Becker AK, Mikolajek H, Paulsson M, Wagener R, Werner JM. A structure of a DxxH domain from calmodulin-ubiquitin fusion protein. Structure (London, England: 1993) 2014;22(2):199–208. https://doi.org/10.1016/j.str.2013.06.028

[10] Luo BH, Carman CV, Springer TA. Structural basis of integrin regulation and adhesion and elsewhere. Mol Biol Cell 2002;13(10):3369–87. https://doi.org/10.1091/mbc.02-05-0259

[11] Medvedev KE, Kinch LN, Grishin NV. Functional and evolutionary analysis of viral proteins containing a Rossmann-like fold. Protein Sci 2018;27(12):2548–63. https://doi.org/10.1002/pro.3458

[12] Edel KH, Kudla J. Integration of calcium and ABA signaling. Curr Opin Plant Biol 2016;33:83–91. https://doi.org/10.1016/j.pbi.2016.06.010

[13] Espinoza-Fonseca LM, Autry JM, Thomas DD. Microsecond molecular dynamics simulations of Mg2+- and W3b-bound E1 intermediate states of the vacuolar ATPase. PLoS ONE 2014;9(4). https://doi.org/10.1371/journal.pone.0095979

[14] Jambunathan N, Siani JM, McNellis TW. A humidity-sensitive Arabidopsis copine mutant exhibits precocious cell death and increased disease resistance. Plant Cell 2001;13(10):2225–40.

[15] Creutz CE et al. The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramycines to humans. J Biol Chem 1998;273(1):1393–402.

[16] Dickeson SK, Bhattacharya-Pakrasi M, Mathis NL, Schlesinger PH, Santoro SA. Ligand binding results in divalent cation displacement from the alpha 2 beta 1 integrin I domain: evidence from terbium luminescence spectroscopy. Biochemistry 1998;37(32):11280–8. https://doi.org/10.1021/bi980774x

[17] Beer B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4(3):435–47. https://doi.org/10.1021/ct070316e

[18] Lindorff-Larsen K et al. Systematic validation of protein force fields against experimental data. PLoS ONE 2012;7(2). https://doi.org/10.1371/journal.pone.0032116

[19] Ohsson MH, Sonderegger CR, Rostkowski M, Jensen JH. PROP3A3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput 2011;7(2):525–37. https://doi.org/10.1021/ct100578a

[20] Kolmurodov K, Smith W, Yasukoa K, Darden TA, Ebsuzaki T. A smooth-particle mesh Ewald method for DL_POLY molecular dynamics simulation package on the Fujitsu VPP700. J Comput Chem 2002;23(13):1187–91.

[21] Hoover WG. Canonical dynamics: equilibrium phase-space distributions. Phys Rev A. General Phys 1985;31(3):1695–7.
[46] Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular-dynamics with coupling to an external bath. J Chem Phys 1984;81(8):3684–90.

[47] Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 2011;51(1):69–82. https://doi.org/10.1021/ci100275a.

[48] Ileana S, Kashif S, Coveney PV. Rapid and accurate prediction of binding free energies for saquinavir-bound HIV-1 proteases. J Am Chem Soc 2008 130(8), 2639–2648.

[49] Chen F et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. PCCP 2016;18(32):22129–39.