

Function and solution structure of the *Arabidopsis thaliana* RALF8 peptide

Ronnie O. Frederick,¹ Miyoshi Haruta,² Marco Tonelli,¹ Woonghee Lee,¹ Gabriel Cornilescu,¹ Claudia C. Cornilescu,¹ Michael R. Sussman,^{2,3} and John L. Markley ^{1,3*}

¹National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, Madison, Wisconsin 53706

²Biotechnology Center, University of Wisconsin-Madison, Madison, Wisconsin 53706

³Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706

Received 8 February 2019; Accepted 16 April 2019

DOI: 10.1002/pro.3628

Published online in Wiley Online Library proteinscience.org

Abstract: We report the recombinant preparation from *Escherichia coli* cells of samples of two closely related, small, secreted cysteine-rich plant peptides: rapid alkalization factor 1 (RALF1) and rapid alkalization factor 8 (RALF8). Purified samples of the native sequence of RALF8 exhibited well-resolved nuclear magnetic resonance (NMR) spectra and also biological activity through interaction with a plant receptor kinase, cytoplasmic calcium mobilization, and *in vivo* root growth suppression. By contrast, RALF1 could only be isolated from inclusion bodies as a construct containing an N-terminal His-tag; its poorly resolved NMR spectrum was indicative of aggregation. We prepared samples of the RALF8 peptide labeled with ¹⁵N and ¹³C for NMR analysis and obtained near complete ¹H, ¹³C, and ¹⁵N NMR assignments; determined the disulfide pairing of its four cysteine residues; and examined its solution structure. RALF8 is mostly disordered except for the two loops spanned by each of its two disulfide bridges.

Keywords: rapid alkalization factor (RALF); peptide structure; NMR solution structure; peptide production; stable isotope labeling; disulfide pairing; root growth assay; cytoplasmic calcium activation assay; peptide dynamics

Introduction

The small (~5 kDa), secreted, basic peptides RALF1 and RALF8 are members of the rapid alkalization

factor (RALF) family of plant peptides that affect plant growth, pathogen response, and pollen tube generation.^{1–5} The importance of the family of RALF peptides in plants is supported by the presence of a large number of genomic DNA sequences encoding RALF and RALF-like peptides. A recent study of computationally identified RALF-like sequences describes 795 RALF coding sequences in 51 different plant species.⁶ There are around 36 known members of this family of peptides in *Arabidopsis*, but little information is available on their three-dimensional (3D) structure despite their important roles in cell signaling and numerous critical plant development processes.⁷ Genome-wide gene expression profiling reveals that each RALF member is characterized by tissue- and development-specific expression patterns. For example, RALF1 is most highly expressed in root and acts as a negative regulator of root growth.⁸ On

Grant sponsor: Division of Molecular and Cellular Biosciences DMB-8415048; Grant sponsor: National Cancer Institute R01 CA073808; Grant sponsor: National Center for Research Resources RR02781 RR08438; Grant sponsor: National Institute of General Medical Sciences P41 GM103399; Grant sponsor: National Science Foundation BIR-9214394 MCB 1410164 OIA-9977486.

*Correspondence to: John L. Markley, National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin 53706. E-mail: markley@nmrfam.wisc.edu

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

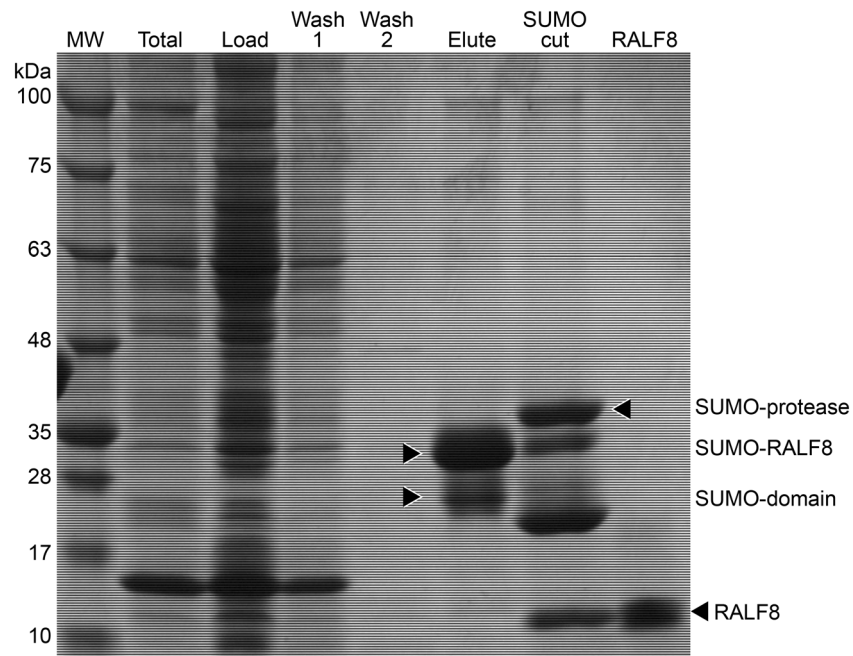


Figure 2. Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) analysis of steps in the production and purification of RALF8. "TOTAL" is from cell lysate. "LOAD" represents the material loaded onto the IMAC column. "WASH 1" and "WASH 2" are from successive column washes. "ELUTE" is from the eluate. "SUMO CUT" is the sample following the addition of SUMO protease. "RALF8" is from the purified protein. The final yield of purified RALF8 was usually 3.5 mg/L minimal medium.

insensitive to RALF8-caused suppression up to 5 μ M RALF8, with the highest concentration examined in this study. This observation indicates that FER

either directly or indirectly mediates RALF8-induced signaling downstream events leading to root growth suppression. The involvement of FER in cellular

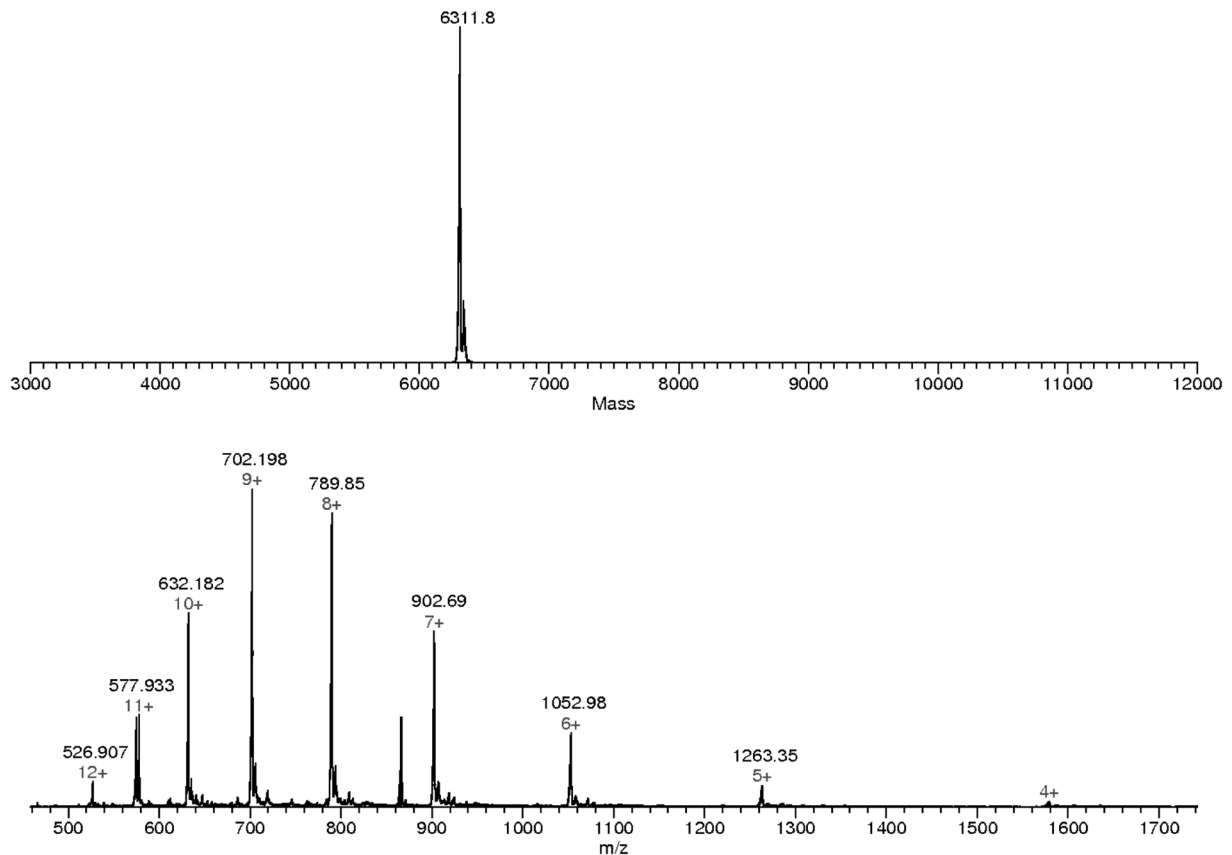


Figure 3. Characterization of 15 N-labeled RALF8 by ESI-TOF mass spectrometry. (Top panel) Reconstructed mass of 15 N-labeled RALF8. (Bottom panel) Peptide ions showing multiple charged species ranging from m/z 1263.35 (5+) to 526.907 (12+).

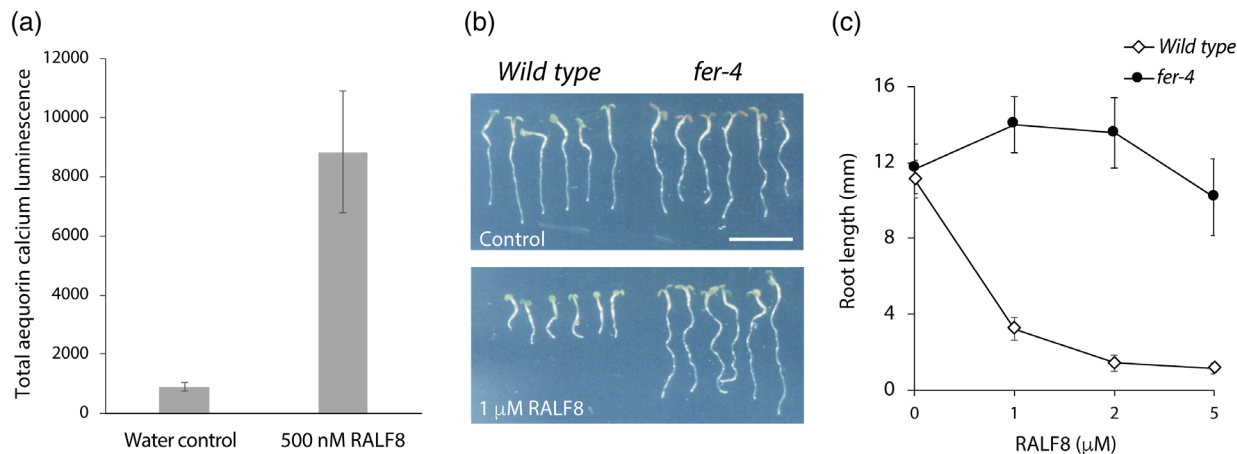


Figure 4. Biological assays of the RALF8 peptide. (a) Cytoplasmic calcium mobilization activity of the RALF8 peptide. (b) Image of *Arabidopsis thaliana* seedlings grown in the presence of RALF8. Three-day-old germinating seedlings were grown in the absence (top panel) or presence (bottom panel) of 1 μM RALF8 peptide for 2 days. FERONIA receptor kinase mutant (*fer-4*) showed insensitivity to RALF8-caused inhibition compared with the wild type. (c) Dose–response curve of RALF8 inhibitory effects on the wild type and *fer-4* mutant seedlings. Data are shown as the mean \pm SD, $n = 6$.

responses to RALF isoforms other than RALF1 was also observed in previous studies of RALF4, RALF19, RALF23, and RALF34.^{2,5,23} Further molecular, genetic, and biochemical experiments examining how different isoforms of RALFs and CrRLKs including FER interact to transmit downstream signaling are currently underway in our laboratories.

The 2D ^1H , ^{15}N -HSQC NMR spectrum of refolded RALF1 at pH 6 (Fig. 5) was poorly dispersed and contained broad peaks. The spectral quality was not improved under different solution conditions, including different salt concentrations and pH values. The

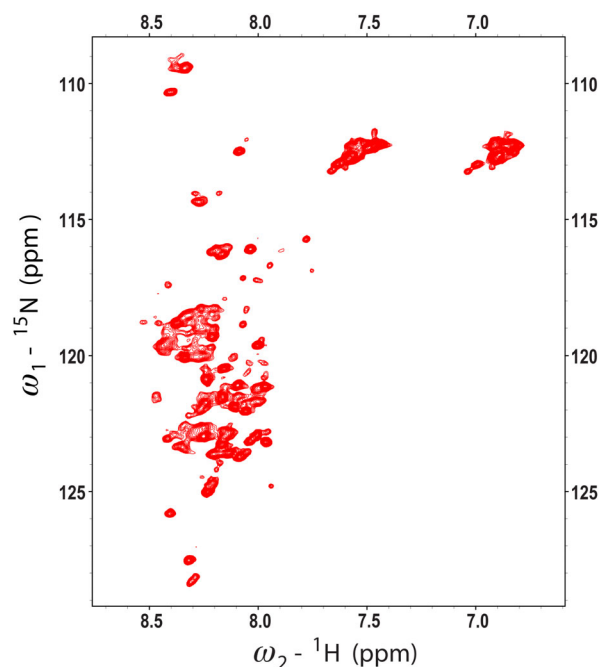


Figure 5. 2D ^1H , ^{15}N HSQC spectrum of [U- ^{15}N]-RALF1 collected at 600 MHz (^1H). The lack of dispersion in the ^1H dimension and broad peaks indicate that the peptide is unstructured and aggregated.

poor spectral quality prevented the NMR structural analysis of RALF1, and as a consequence, we turned to RALF8, another variant studied in our laboratory. The peaks in the 2D ^1H , ^{15}N -HSQC spectrum of RALF8 at pH 6 (Fig. 6) were spread across a narrow range of chemical shifts in the ^1H dimension, suggesting that the protein is highly dynamic in solution. Nevertheless, we achieved near complete chemical shift assignments for RALF8 (all backbones except for those of the first residue (E1), all aliphatic side chains, but no aromatic side chains). A number of residues exhibited evidence of multiple backbone

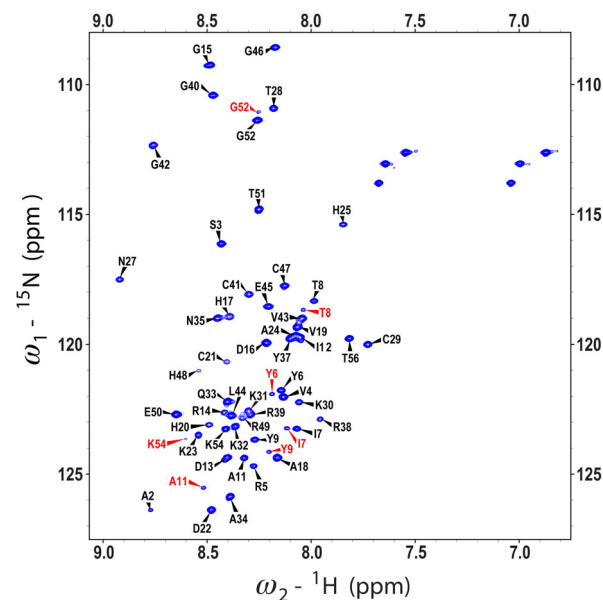


Figure 6. 2D ^1H , ^{15}N HSQC spectrum of [U- ^{15}N]-RALF8 collected at 800 MHz (^1H). Peaks are labeled with the backbone assignments obtained by I-PINE, followed by manual refinement in NMRFAM-SPARKY. Doubled backbone peaks (labeled in red) were assigned to several amino acid residues.

amide peaks, suggesting that at least parts of the protein may populate two distinct conformations. In particular, we were able to identify distinct pairs of backbone amide peaks for residues Y6, I7, T8, Y9, A11, G52, and K54 (labeled in red in Fig. 6). Given the proximity of these residues to two of the proline residues (P10 and P53), this suggested that the peptidyl–prolyl peptide bonds of these two residues may populate both the *trans* and *cis* configurations. Because the peptide is believed to function at a lower pH, we also collected a 2D $^1\text{H},^{15}\text{N}$ -HSQC NMR spectrum of RALF8 at pH 4.5. Although signals from several residues, in particular histidine residues, exhibited chemical shift differences, the overall pattern was the same. The peptide did not become more structured, and the relative intensities of the major and minor peaks remained the same.

To analyze the question of heterogeneity of peptidyl–prolyl peptide bonds, we examined nuclear Overhauser enhancement (NOE) cross peak data to differentiate signals from *cis* and *trans* configurations and utilized relatively peak intensities from the major and minor signals from individual residues to determine the *trans* and *cis* populations. A strong NOE cross peak was observed between major peaks Y9 $^1\text{H}^\alpha$ and P10 $^1\text{H}^\delta$, but not between major peaks Y9 $^1\text{H}^\alpha$ and P10 $^1\text{H}^\alpha$; these findings indicate that the major form has a *trans* Y9–P10 peptide bond. For the corresponding minor form, we found peaks corresponding to $^1\text{H}^\text{N}$ of Y6–Y9 and A11, but were unable to locate cross peaks between Y9 $^1\text{H}^\alpha$ and P10 $^1\text{H}^\alpha$ because the latter is close to the water signal and occluded by noise. From the assigned $^1\text{H}^\text{N}$ of A11, and with the aid of data from HNCACB, CBCACONH, CCONH, HBHACONH, $^1\text{H},^{13}\text{C}$ -HSQC, and HCCH-TOCSY spectra, we identified P10 $^{13}\text{C}^\beta$ and $^{13}\text{C}^\gamma$ signals from the major and minor forms (Table I). The $\Delta(\delta^{13}\text{C}^\beta - \delta^{13}\text{C}^\gamma)$ values were consistent with the major form being *trans* and the minor form being *cis*. On the basis of the relative intensities of the major and minor $^1\text{H}^\text{N}$ peaks of Y6–Y9 and A11, we calculated Y9–P10 *trans:cis* as 74:26.

For analysis of the G52–P53 peptide bond, we observed an NOE cross peak between major peaks G52 $^1\text{H}^\alpha$ and P53 $^1\text{H}^\delta$, but not between major peaks G52 $^1\text{H}^\alpha$ and P53 $^1\text{H}^\alpha$; these findings indicate that the

major form has a *trans* G52–P53 peptide bond. From the assigned $^1\text{H}^\text{N}$ of K54, and with the aid of data from HNCACB, CBCACONH, CCONH, HBHACONH, $^1\text{H},^{13}\text{C}$ -HSQC, HCCH-TOCSY, and NOESY spectra, we identified P53 $^{13}\text{C}^\beta$ and $^{13}\text{C}^\gamma$ signals from the major and minor forms (Table I). The $\Delta(\delta^{13}\text{C}^\beta - \delta^{13}\text{C}^\gamma)$ values were consistent with the major form being *trans* and the minor form being *cis*. On the basis of the relative intensities of the major and minor $^1\text{H}^\text{N}$ peaks of G52 and K54, we calculated G52–P53 *trans:cis* as 85:15.

For analysis of the K54–P56 peptide bond, we observed an NOE cross peak between major peaks K54 $^1\text{H}^\alpha$ and P56 $^1\text{H}^\delta$, but not between major peaks K54 $^1\text{H}^\alpha$ and P56 $^1\text{H}^\alpha$; these findings indicate that the major form has a *trans* K54–P56 peptide bond. Because no minor peaks corresponding to P55 or T56 were observed, we conclude that the K54–P56 peptide bond is largely *trans*.

The $^{13}\text{C}^\beta$ chemical shifts of the four cysteine residues (C21, C29, C41, and C47) ranged between 40.384 and 43.219 ppm, suggesting that all four are involved in S–S bonds.²⁴ To establish the disulfide bond configuration, we looked for $^1\text{H}^\beta$ – $^1\text{H}^\beta$ NOE contacts between disulfide-bonded cysteine residues in the spectra recorded using a modified $^1\text{H},^{13}\text{C}$ -HSQC-NOESY- $^1\text{H},^{13}\text{C}$ -HSQC pulse program. This NOESY experiment was preferred over the more conventional NOESY- $^1\text{H},^{13}\text{C}$ -HSQC one because it relies on the ^{13}C dimension to separate NOE cross peaks rather than the ^1H dimension. This approach is advantageous because the $^{13}\text{C}^\beta$ resonances of the four cysteine residues of RALF8 are spread across a wider chemical shift range than the corresponding $^1\text{H}^\beta$ resonances. Furthermore, in order to increase the resolution in the ^{13}C dimension, we modified the evolution scheme for both indirect ^{13}C dimensions to decouple $^{13}\text{C}^\beta$ nuclei from their attached $^{13}\text{C}^\alpha$ and to reduce the ^{13}C spectral window (12.5 ppm) so that only a narrow region centered around the cysteine $^{13}\text{C}^\beta$ chemical shifts (^{13}C offset set to 40.645 ppm) was detected. We used this modified pulse program to record a high-resolution 3D $^1\text{H},^{13}\text{C},^{13}\text{C}$ spectrum. After processing it with NMRPipe, careful examination of this 3D NOESY spectrum in NMRFAM-SPARKY allowed us to identify clear NOE contacts between the $^1\text{H}^\beta$ nuclei of C21 and C29 and between the $^1\text{H}^\beta$ nuclei of C41 and C47, indicating that C21–C29 and C41–C47 are the correct pairing for the two disulfide bonds in RALF8.

To get an understanding of the structural content of RALF8 in solution, we used the assigned chemical shifts of the backbone atoms as input to four secondary structure prediction programs: PECAN, TALOS-N, GetSBY, and ncSPC. The results from these programs all agreed that RALF8 shows very little propensity for secondary structure (data not shown). Moreover, backbone order parameter S^2 derived from random coil index (RCI- S^2) values predicted by

Table I. $^{13}\text{C}^\beta$ and $^{13}\text{C}^\gamma$ Chemical Shifts of P10 and P53 in Their Major and Minor Peptide Bond Configurations

Proline residue	Atom(s)	Major form (ppm)	Minor form (ppm)
10	$^{13}\text{C}^\beta$	32.11	34.08
	$^{13}\text{C}^\gamma$	27.37	24.79
	$\Delta(\delta^{13}\text{C}^\beta - \delta^{13}\text{C}^\gamma)$	4.74	9.29
53	$^{13}\text{C}^\beta$	32.2	34.66
	$^{13}\text{C}^\gamma$	27.16	24.84
	$\Delta(\delta^{13}\text{C}^\beta - \delta^{13}\text{C}^\gamma)$	5.04	9.82

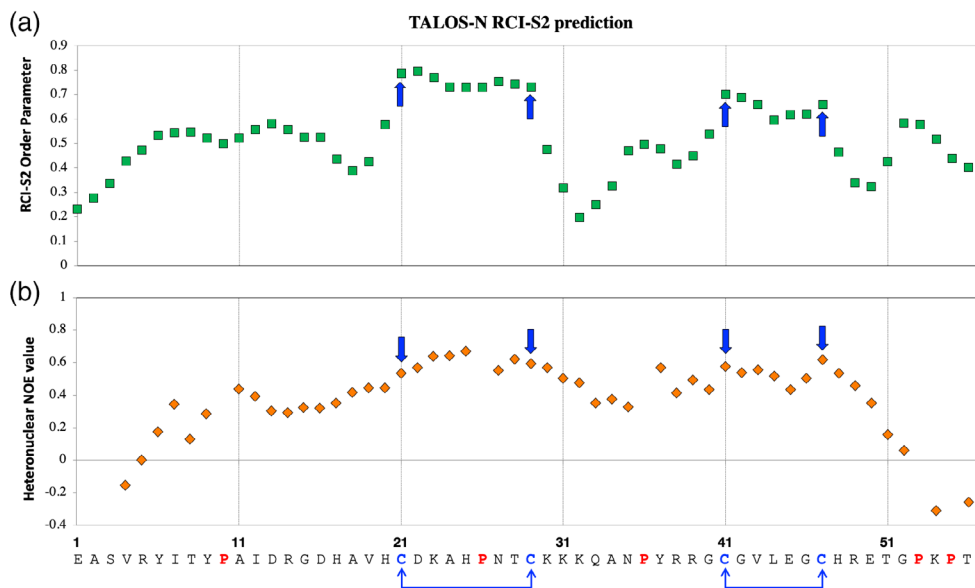


Figure 7. (a) RCI-S² index of RALF8 predicted by TALOS-N using the chemical shifts of assigned backbone resonances and (b) ¹H,¹⁵N-heteronuclear NOE values for RALF8 backbone amides measured experimentally. The blue arrows point to the four cysteine residues that are present in the sequence of RALF8 (C21, C29, C41, and C47).

TALOS-N [Fig. 7(a)] indicates that the backbone of RALF8 is quite flexible in solution. Only two short stretches of amino acids exhibit a higher degree of order [Fig. 7(a)]. Interestingly, these are the two stretches of the sequence flanked by the four cysteine residues (C21–C29 and C41–C47). This observation is in agreement with our analysis of the ¹H-¹⁵N heteronuclear NOE spectra that show the same two stretches of amino acids having the backbone amides with the largest heteronuclear NOE values [Fig. 7(b)]. These results are also in agreement with the detected disulfide pairing: C21–C29 and C41–C47 [bottom of Fig. 7(b)].

We used the NMR data as input to the Xplor-NIH option of the PONDEROSA-C/S program²⁵ to determine models of the 3D structure of RALF8 (Fig. 8). The root mean square deviation (RMSD) between the backbone heavy atoms of all residues in the superimposed structures and those of the representative lowest energy structure was 12.8 Å, consistent with overall dynamic disorder. However, two loops between the disulfides (C21–29 and C41–47) were well defined with RMSD values of 0.6 and 0.7 Å for the backbone heavy atoms, respectively. Two sets of structural models, one superimposed on the structured region around C21–29 [see Fig. 8(a)] and the other superimposed on the structural region around C41–47 [see Fig. 8(b)], were deposited in the PDB under accession 6NU4.

Conclusions

NMR spectroscopy easily discriminates between peptide preparations, such as RALF8, that are native and monodisperse and the ones that are misfolded, such as RALF1. Although both RALF1 and RALF8 are

similar in polypeptide length and each contains four cysteine residues, the solubility and yield of the two peptides from *E. coli* differed to a large extent. Despite trying numerous vectors, bacterial fusions, and peptide production strategies, the NMR spectra of the RALF1 preparations were of uniformly poor quality. By contrast, we found it to be relatively straightforward to produce RALF8 in SUMO fusion-based vectors and to isolate a soluble active product with intact disulfide bridges and a free glutamate at the N terminus. The fact that the recovery yield of RALF1 from *E. coli* was much lower than that of RALF8 suggests that RALF1 may be more susceptible to proteases than RALF8. Biological assays with *A. thaliana* seedlings revealed that RALF8, as previously shown for RALF1,^{13,15} was potent to induce signal downstream events leading to root growth suppression. NMR confirmed the presence of two disulfide bonds in the peptide and showed that it is largely disordered in solution except for the regions near the disulfide bridges. The absence of a defined folded state of this peptide in solution suggests that additional structures may be induced only when interacting with a partner protein. Candidate proteins for such interactors include FERONIA, the receptor-like kinase that mediates RALF1 action, as well as the possible coreceptors or chaperones, such as LLG1.¹⁴

Materials and Methods

DNA cloning, production, and purification of RALF8 peptide

Arabidopsis thaliana genomic DNA was used as a source to clone sequences corresponding to residues 72–120 of RALF1 (UniProt: Q9SRY3) and residues

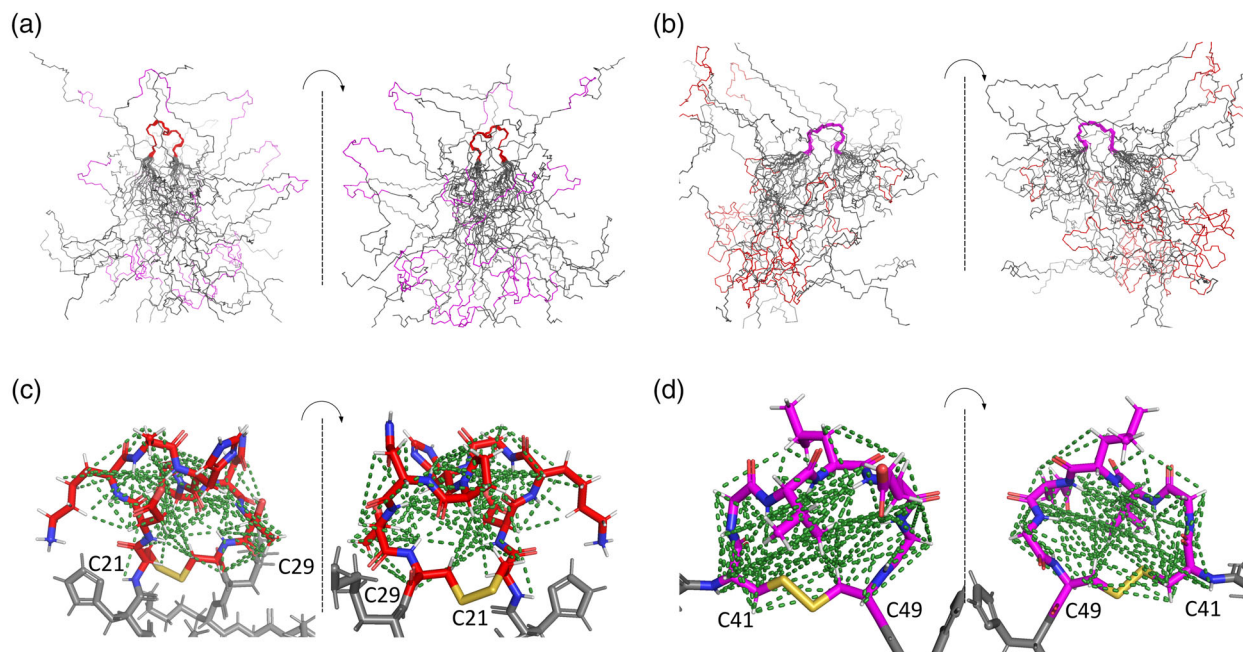


Figure 8. Solution structure of RALF8. The protein is dynamically disordered with the exception of two ordered loops that are held together by disulfide bridges. The C21–C29 loop (red) is more ordered than the C41–C47 loop (magenta), and the two ordered regions are independently ordered. (a) The 20 best structures for backbone atoms of RALF8 overlaid with reference to the structured residues in the C21–C29 loop. (b) The 20 best structures for backbone atoms of RALF8 overlaid with reference to the structured residues in the C41–C47 loop. (c) Structured C21–C29 loop region (white, hydrogen; blue, nitrogen; red double stick, oxygen; red, carbon; and gold, disulfide bond). (d) Structured C41–C47 loop region (white, hydrogen; blue, nitrogen; red double stick, oxygen; magenta, carbon; and gold, disulfide bond). The dotted lines represent NOEs used in determining the structures.

27–82 of RALF8 (UniProt: Q1ECR9) into the T7 promoter-based expression vector, pE-SUMO (Kan) (LifeSensors). We also cloned the RALF1 gene into numerous vectors containing a variety of different N-terminal fusion partners and tried protein purification strategies besides pE-SUMO (Kan). However, none of the systems tested produced soluble active RALF1 peptides. Polymerase chain reaction (PCR) amplification and polymerase incomplete primer extension-based techniques²⁶ and KOD DNA polymerase (EMD Millipore) were used to amplify the RALF8 gene and insert it into the pE-SUMO (Kan) vector. Chemically competent 10G *E. coli* cells (Lucigen) were transformed with ligation reaction and, after overnight growth of cells, recombinants were screened by PCR with 48-base primers that anneal to 24 base regions in the 5' and 3' termini of the RALF8 gene. Positive recombinants were grown overnight at 37°C in 4 mL of 2YT (plus 50 µg/mL kanamycin), harvested, and used to purify expression plasmids for DNA sequencing analysis (DNASTAR Lasergene), and then archived and stored at –20°C.

The *E. coli*-based production of RALF8 and RALF1 peptides was first optimized and screened at a small scale using 96-well growth blocks, 0.5 mL of Luria-Bertani (LB) growth medium per well, and *E. coli* B T7 SHuffle[®] expression strain (New England BioLabs). The T7 SHuffle strains have genetic mutations (*gor-*, *trxB-*, *ahpC**, and cytoplasmic DsbC) that

permit disulfide bond formation in proteins as they fold during translation in the *E. coli* cytoplasm.²⁷ Chemically competent T7 SHuffle B cells were transformed with RALF8 pE-SUMO (Kan) plasmid and grown on the LB plate (plus 50 µg/mL kanamycin and 1% glucose) overnight at 37°C. The next day, wells in a 96-well growth block were filled with 0.5 mL of LB (plus appropriate antibiotics) and were inoculated with fresh colonies, and then the growth block was shaken at 37°C (250 rpm) until the optical density (OD) at 600 nm (OD_{600nm}) reached about 1 (usually takes between 1 and 2 h). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final level of 1 mM to each well, and the growth block was shaken at 37°C for 4 more hours. The block was spun using a JS 5.9 rotor at 2000g for 5 min to harvest the cells for protein production analysis. The pelleted cells were suspended in 100 µL BugBuster™ (EMD Millipore), lysed, and analyzed by SDS-PAGE on 4–20% gels (BioRad). The pE-SUMO (Kan) plasmid system produces the RALF8 peptide fused to the C terminus of a SUMO domain (108 aa, 10.2 kDa) containing a six His-tag at the N terminus.^{22,28} The SUMO-RALF8 peptide fusion was overproduced as a soluble product in *E. coli* that can be isolated using immobilized metal affinity chromatography (IMAC).

In the case of RALF1, we were unsuccessful in producing the native sequence of the soluble active RALF1 peptide at either small scale or large scale

with any of the expression plasmids tested. However, the RALF1 peptide containing an N-terminal His-tag could be prepared at a large scale by using the Flexi pVP67K (N-terminal His-tagged vector similar to pVP68K but without MBP fusion)²¹ and Rosetta2 *E. coli* host expression cells. The RALF1 peptide formed insoluble inclusion bodies under the growth conditions used and required refolding before protein purification. The His-tagged RALF1 peptide was unstable and highly susceptible to proteolytic degradation during purification.

We were able to produce the native sequence of RALF8 on a large scale. We began by picking individual colonies of RALF8 pE-SUMO(Kan)/T7 SHuffle B and growing them in 1 mL LB medium at 37°C for 1–4 h. Next, these 1 mL starters were used to inoculate 100 mL of MDAG noninducing medium,²⁹ and the culture was grown overnight at 25–30°C for about 18 h. To prepare ¹⁵N- and ¹³C, ¹⁵N-labeled RALF8 peptide, 50 mL of overnight MDAG inoculum was added to 1 L of M9-based medium supplemented with 2–4 g ¹³C-glucose and 1 g ¹⁵N-ammonium chloride (along with appropriate antibiotics), and the culture was grown at 37°C until the OD_{600nm} reached ~1.0.²² Prior to chemical induction, the temperature of the shaker was lowered to 25°C; 0.4 mM IPTG was added as the inducer, and the culture was grown for about 24 h. The *E. coli* cells were harvested by centrifugation at 4200 rpm for 20 min at 4°C, and the cell paste was stored for a long term at –80°C.

Isolation and purification of RALF peptides

Frozen cell pastes of RALF1 and RALF8 were processed similarly: they were thawed and resuspended in 55–60 mL of lysis buffer [50 mM Tris pH 8, 500 mM NaCl, 10% ethylene glycol, 5 mM imidazole, 2% v/v NP-40 detergent, and 1 mM phenylmethylsulfonyl fluoride (PMSF)] supplemented with 50 µg/mL RNase, 25 µg/mL DNase, and 100 µg/mL chicken egg white lysozyme. To break open the cells, a Misonix 3000 sonicator was used at 4°C with a program of 1 s on or 4 s off, at 55–65% power for a total of 15 min. The lysed cells were clarified before protein purification by centrifugation at 28,000 rpm for 30 min at 4°C. In contrast to RALF1 fused to SUMO, which produced inclusion bodies and insoluble peptide, RALF8 fused to SUMO was soluble under the growth conditions used, and the solution was loaded on to a 10 mL Qiagen IMAC column that had been equilibrated with IMAC buffer A (50 mM Tris pH 8, 500 mM NaCl, 10% ethylene glycol, 5 mM imidazole, and 1 mM PMSF). The RALF8 peptide was loaded onto an IMAC column at a flow rate of 1 mL/min, and the column was washed first with 10 column volumes of IMAC buffer A and then with 10 column volumes of IMAC buffer B (IMAC buffer A with 30 mM imidazole). His×6-SUMO-RALF8 was eluted with 30 mL of IMAC elution buffer (IMAC

buffer A with 500 mM imidazole). The column fractions were analyzed by SDS-PAGE to assess the purity and success of RALF8 peptide purification. To remove the SUMO fusion domain, 1 mg of SUMO protease (produced in-house) was added to the purified His×6-SUMO-RALF8, and the mixture was dialyzed overnight in SUMO cleavage buffer (50 mM Tris, pH 8, 150 mM NaCl, 10% ethylene glycol, 5 mM imidazole, 1 mM PMSF, and 0.1 mM β-mercaptoethanol). The usage of a low concentration of reductant ensured that the disulfide bonds in RALF8 were not broken during SUMO protease cleavage of the SUMO domain from the RALF8 peptide. The SUMO domain was removed by subtractive IMAC, and the flow-through contained the native sequence of RALF8. The RALF8 peptide was concentrated on spin concentrators (Sartorius Vivaspin and Millipore) and polished further by cation exchange chromatography on a 5 mL GE SP cation exchange column; the peptide was bound to the column and then equilibrated with buffer A (50 mM MES, pH 6) and finally eluted by a salt gradient with elution buffer B (50 mM MES, pH 6, 1M NaCl). For RALF1, cell growths were harvested and lysed by the method used for RALF8. However, the RALF1 peptide was insoluble, and the inclusion bodies were harvested and washed with lysis buffer (with and without NP-40 detergent, centrifuged at 18,000 g for 15 min), and denatured overnight in 60 mL 6 M guanidinium chloride containing 1 mM dithiothreitol. The denatured RALF1 peptide was clarified by centrifugation (28,000 rpm for 30 min) and then refolded by rapid dilution (dropwise) into 1 L of 50 mM Tris, pH 8, 500 mM NaCl, 10% ethylene glycol, 2 mM reduced glutathione, 1 mM oxidized glutathione, and 1 mM PMSF. After storing overnight at 4°C, the solution was applied to an IMAC column to isolate His-tagged RALF1.

Mass spectrometry of RALF8

Purified RALF8 samples were analyzed using an Agilent liquid chromatography mass spectrometry–time-of-flight (LC/MSD-TOF) set in positive mode. The mass measurement of the purified ¹⁵N-labeled RALF8 peptide using an ESI-TOF-MS instrument detected peptide ions with charges ranging from 5+ to 12+ (Fig. 3), reflecting the protonation of multiple Lys and Arg residues present in this peptide. Deconvolution of these charged states led to an average mass of 6311.8 (Fig. 3). We used the MS isotopic algorithm at the Protein Prospector site (<http://prospector.ucsf.edu/prospector/mshome.htm>) to model the level of ¹⁵N incorporation. Assuming that all the peptide molecules in the analyte are uniform and contain two disulfide bridges, the elemental composition for the oxidized form of fully ¹⁵N labeled RALF8 is C265 H418 ¹⁵N87 O80 S4, which predicts a theoretical nominal mass of 6311.77016, which represents

Table II. Experimental Details for 3D NMR Spectra Collected for Backbone and Side Chain Assignments and Structure Calculations^a

Experiment	Number of scans	Spectral window (ppm) ¹ H × ¹³ C (¹ H [#]) × ¹⁵ N(¹³ C ^{&})	Complex points ¹ H × ¹³ C(¹ H [#]) × ¹⁵ N(¹³ C ^{&})	Offset (ppm) ¹ H × ¹³ C(¹ H [#]) × ¹⁵ N(¹³ C ^{&})
HNCACB	16	16 × 59.7 × 23	1024 × 96 × 48	4.76, 44.8, 117.5
CBCA(CO)NH	16	16 × 59.7 × 23	1024 × 72 × 36	4.76, 44.8, 117.5
HN(CA)CO	24	16 × 10 × 23	1024 × 48 × 48	4.76, 175.4, 117.5
HNCO	8	16 × 10 × 23	1024 × 48 × 48	4.76, 175.4, 117.5
C(CO)NH	16	16 × 66.7 × 23	1024 × 72 × 48	4.76, 41.9, 117.5
HBHA(CO)NH	16	16 × 6.25 [#] × 23	1024 × 72 [#] × 52	4.76, 4.76 [#] , 117.5
H(C)CH-TOCSY	8	16 × 5.75 [#] × 66.7 ^{&}	1024 × 64 [#] × 64 ^{&}	4.76, 4.76 [#] , 41.9 ^{&}
NOESY ¹⁵ N-HSQC*	16	16 × 11 [#] × 23	1024 × 120 [#] × 46	4.76, 4.76 [#] , 117.5
NOESY ¹³ C-HSQC	8	16 × 10 [#] × 66.7 ^{&}	1024 × 72 [#] × 64 ^{&}	4.76, 4.76 [#] , 42 ^{&}
¹³ C-HSQC NOESY ¹³ C-HSQC*	32	16 × 12.5 × 12.5 ^{&}	1024 × 48 × 48 ^{&}	4.76, 40.645, 40.645 ^{&}

^a Spectra were collected with cryogenic probes at 600 MHz (indicated by *) or 800 MHz (¹H) at a regulated temperature of 298 K.

19.56% of the total. Although this comparison of the observed average mass assigned by an instrument software and the theoretical nominal mass are reasonably close, to obtain a more accurate estimation of ¹⁵N incorporation to RALF8 peptide, we used the *m/z* value of the most abundant species, the 9+ charge state [M + 9H]⁹⁺. The mass resolution for the modeling was set to 10,000, which matches that of the ESI-TOF-MS instrument. This approach yielded 93% as the average level of ¹⁵N in the ¹⁵N-labeled sample of RALF8.

Bioassay of RALF8

For the cytoplasmic calcium mobilization assay [Fig. 4(a)], *A. thaliana* seedlings expressing the calcium reporter, aequorin,³⁰ were grown on half-strength Murashige–Skoog (½ MS) medium for 5 days. Single seedlings were incubated in 200 µL ½ MS media supplemented with 2.5 µM coelenterazine cp (Sigma-Aldrich) in the dark for 16 h. RALF8 was added to the medium immediately before measuring bioluminescent emission. Luminescent signals were measured with a plate reader (Tecan) for 150 s. The total luminescence emission reflecting an increase in calcium concentration was calculated by summing signals obtained during 150 s. For a negative control, a volume of water or buffer equivalent to the RALF8 solution was added.

For the RALF8 root growth inhibition assay [Fig. 4(b,c)], wild-type (Columbia ecotype) and *fer-4* knockout mutant seedlings were grown for 3 days on ½ MS solid medium and then transferred to 500 µL ½ MS liquid medium containing 0, 1, 2, or 5 µM RALF8 peptide. Seedlings were further grown for 2 days. Images were captured using an office scanner (Epson), and root length was quantified using ImageJ software.

Preparation of NMR samples

Freshly prepared samples of ¹⁵N- and ¹³C/¹⁵N-labeled RALF8 and RALF1 were exchanged by dialysis into a

solution containing 20 mM sodium phosphate buffer, pH 6.0, 200 mM sodium chloride, 0.05% sodium azide, and 10% D₂O (Sigma–Aldrich); the final concentration of each protein was ~1.0 mM. Samples used for NMR spectroscopy (300 µL) were enclosed in 5-mm susceptibility-matched Shigemi NMR tubes (Shigemi, Allison Park, PA).

NMR data collection

All NMR spectra were acquired using Varian VNMRS (Agilent Technologies, Santa Clara, CA) spectrometers operating at 600 or 800 MHz (¹H); each was equipped with a cryogenic triple-resonance probe. The temperature of the sample was regulated at 298 K for all the experiments. 2D ¹H,¹⁵N-HSQC and 3D HNCACB, 3D CBCA(CO)NH, 3D HN(CA)CO, and 3D HNCO spectra were collected to enable backbone assignments. 2D constant-time ¹H,¹³C-HSQC and 3D C(CO)NH, 3D HBHA(CO)NH, and 3D H(C)CH-TOCSY spectra were acquired to enable assignment of the remaining ¹H and ¹³C resonances in the aliphatic spectral region. Finally, for structure calculations, 3D NOESY ¹⁵N-HSQC and 3D NOESY ¹³C-HSQC data sets were collected, both using a 200 ms mixing time. The acquisition parameters of all 3D recorded spectra are reported in Table II. All 3D spectra, except for the 3D NOESY ¹³C-HSQC, were recorded using nonuniform sampling (NUS) at a sampling rate of ~40%. All spectra were processed using NMRPipe,³¹ and the 3D spectra recorded using NUS were reconstructed and processed using the SMILE package³² included in NMRPipe.

To establish the S–S bond configuration of RALF8, we recorded a 3D ¹H,¹³C,¹³C-spectrum using a modified ¹H,¹³C-HSQC-NOESY-¹H,¹³C-HSQC experiment (see description in Results and Discussion). This additional 3D NOESY spectrum was acquired using a 200 ms mixing time, with 1024 × 48 × 48 complex points in the ¹H, ¹³C, and ¹³C dimensions, respectively, and 32 repetitions for each free induction decay

Table III. Statistics From the RALF8 Structure Calculations^a

Considered regions	1–56	21–29	41–47
Distance constraints			
Short range $ i - j \leq 1$	832	123	124
Medium range $1 < i - j < 5$	66	31	23
Long range $ i - j \geq 5$	16	12	4
Total	914	166	151
Dihedral angle constraints			
φ	10	7	2
ψ	10	7	2
Total	20	14	4
Average RMSD to the representative coordinates			
Backbone heavy atoms N, C ^α , C ^β , and C ^γ	12.81 ± 2.47	0.55 ± 0.28	0.70 ± 0.37
All heavy atoms	12.93 ± 3.10	1.23 ± 0.25	0.77 ± 0.33
PROCHECK Z scores (φ and ψ /all)	-3.11/-3.37	-1.06/-1.95	-4.52/-6.80
Ramachandran plot summary from PROCHECK (%)			
Most favored	72.8	83.1	72
Additionally allowed	18.5	16.9	19
Generously allowed	5.7	0	8
Disallowed	3	0	1
MolProbity Clashscore/Z-score		3.05/1.00	
Ramachandran plot summary from MolProbity (%)			
Most favored	83.4	100	67.1
Allowed	9.4	0	20
Disallowed	7.1	0	12.9
PONDEROSA pseudopotential <i>E</i> (kJ/mol)		2169.08 ± 44.93	
Violations			
Distance constraints [>0.5 Å]		0	
Dihedral angle constraints [$>5^\circ$]		0	
van der Waals constraints [>0.2 Å]		0	

^a Structural qualities for three different regions were evaluated using PSVS 1.5 (http://psvs-1_5-dev.nesg.org).

(FID). The indirect ¹³C dimensions were centered to the ¹³C^β region of cysteine (40.645 ppm) and narrowed down to 12.5 ppm to maximize the resolution. This 3D spectrum was then processed using NMRPipe and analyzed with NMRFAM-SPARKY.³³

NMR peak assignments

The processed 2D ¹H, ¹⁵N-HSQC, 3D HNCACB, 3D CBCA(CO)NH, 3D HN(CA)CO, and 3D HNCO spectra were analyzed in NMRFAM-SPARKY. For backbone assignments, the program APES³⁴ was used for automated peak picking. After manual inspection, the peak lists from 3D and 2D spectra were submitted to the I-PINE web server (<http://i-pine.nmrfam.wisc.edu>) for automated backbone assignments using PINE-SPARKY.2 automation plugin (two-letter-code ep) in NMRFAM-SPARKY. These automated assignments were then verified and extended by hand using the tools available in NMRFAM-SPARKY. The final backbone chemical shifts were then analyzed to estimate the structural content of RALF8 peptide in solution by running PECAN,³⁵ TALOS-N,³⁶ GetSBY,³⁷ and ncSPC.³⁸ After the backbone assignments were completed, we collected additional spectra (2D constant-time ¹H, ¹³C-HSQC, 3D C(CO)NH, 3D HBHA(CO)NH, and 3D H(C)CH-TOCSY) and used the “transfer and simulate assignments” tool available in NMRFAM-SPARKY to extend the backbone assignments to the aliphatic portion of the side chains. We were thus able to

identify 100% of these resonances. All assigned chemical shifts have been deposited in the BMRB³⁹ (code 30565).

Dynamics

To assess the dynamic properties of RALF8, we also ran experiments to measure the ¹H-¹⁵N heteronuclear NOE parameters for all assigned backbone amides. Two spectra, with (NOE spectrum) and without (reference spectrum) a 3 s ¹H saturation pulse train, were recorded in an interleaved fashion with 1024 × 128 points in the ¹H and ¹⁵N dimensions, respectively, with 32 scans per FID. The spectra were then processed using NMRPipe and analyzed with the Peak Intensity Analysis tool in NMRFAM-SPARKY. The ¹H-¹⁵N heteronuclear NOE values were extracted by calculating the height ratios of the corresponding peaks from the spectra recorded with and without ¹H saturation.

Structure calculations

An NMR-STAR 3.1 chemical shift file was generated by a plugin of NMRFAM-SPARKY and used as input to the PONDEROSA client program along with the raw 3D ¹⁵N-edited NOESY and ¹³C-edited aliphatic NOESY spectra. Tolerances were set to 0.025, 0.35, and 0.35 ppm for ¹H, ¹³C, and ¹⁵N, respectively. The PONDEROSA program includes an automatic NOE cross peak detection module connected to a visual noise level selector in NMRFAM-SPARKY.³⁷ The 3D

structure calculation invoked the AUDANA⁴⁰ option, which automatically generated distance restraints, and TALOS-N, to obtain dihedral angle restraints. The PONDEROSA client program then submitted the structure calculation job to the PONDEROSA web server. Subsequently, we used the “Distance Constraint Validator” and “Angle Constraint Validator” of the Ponderosa Analyzer to validate the constraints. Constraints were visualized with PyMOL, and the corresponding NOE cross peaks were visualized synchronously with NMRFAM-SPARKY. After constraints were carefully verified, they were submitted to the PONDEROSA web server using the *Constraints-Only X* option, which utilized Xplor-NIH to carry out high-temperature dynamics, simulated annealing, and low-temperature dynamics. A total of 914 distance constraints (832, 66, and 16 constraints for short-, medium-, and long-range, respectively) and 20 dihedral angle constraints (10 φ angles and 10 ψ angles) were used for structure calculation (Table III). From a total of 100 calculated structures, we selected the 20 lowest energy structures calculated using the *Final Step with Explicit Water Refinement* option.

Acknowledgments

This study was supported by NIH grant R01 CA073808 (NCI) and made use of the National Magnetic Resonance Facility at Madison, with support from the NIH grant P41 GM103399. Additional equipment was purchased with funds from the University of Wisconsin, the NIH (RR02781, RR08438), and the NSF (DMB-8415048, OIA-9977486, and BIR-9214394). Reagents and salary for MH and MRS were supported by the NSF grant MCB 1410164 to MRS. We thank Grzegorz Sabat for carrying out the mass spectrometric analyses at the University of Wisconsin Biotechnology Center.

References

- Pearce G, Moura DS, Stratmann J, Ryan CA Jr (2001) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. *Proc Natl Acad Sci U S A* 98:12843–12847.
- Ge Z, Bergonci T, Zhao Y, Zou Y, Du S, Liu MC, Luo X, Ruan H, Garcia-Valencia LE, Zhong S, Hou S, Huang Q, Lai L, Moura DS, Gu H, Dong J, Wu HM, Dresselhaus T, Xiao J, Cheung AY, Qu LJ (2017) Arabidopsis pollen tube integrity and sperm release are regulated by RALF-mediated signaling. *Science* 358:1596–1600.
- Ge Z, Cheung AY, Qu LJ (2019) Pollen tube integrity regulation in flowering plants: insights from molecular assemblies on the pollen tube surface. *New Phytol* 222:6876–6693.
- Stegmann M, Monaghan J, Smakowska-Luzan E, Rovenich H, Lehner A, Holton N, Belkhadir Y, Zipfel C (2017) The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* 355:287–289.
- Gonneau M, Desprez T, Martin M, Doblaz VG, Bacete L, Miart F, Sormani R, Hematy K, Renou J, Landrein B, Murphy E, Van De Cotte B, Vernhettes S, De Smet I, Hofte H (2018) Receptor kinase THESEUS1 is a rapid alkalization factor 34 receptor in Arabidopsis. *Curr Biol* 28:2452–2458.
- Campbell L, Turner SR (2017) A comprehensive analysis of RALF proteins in green plants suggests there are two distinct functional groups. *Front Plant Sci* 8:37.
- Murphy E, De Smet I (2014) Understanding the RALF family: a tale of many species. *Trends Plant Sci* 19:664–671.
- Haruta M, Monshausen G, Gilroy S, Sussman MR (2008) A cytoplasmic Ca²⁺ functional assay for identifying and purifying endogenous cell signaling peptides in Arabidopsis seedlings: identification of AtRALF1 peptide. *Biochemistry* 47:6311–6321.
- Mecchia MA, Santos-Fernandez G, Duss NN, Somoza SC, Boisson-Dernier A, Gagliardini V, Martinez-Bernardini A, Fabrice TN, Ringli C, Muschiatti JP, Grossniklaus U (2017) RALF4/19 peptides interact with LRX proteins to control pollen tube growth in Arabidopsis. *Science* 358:1600–1603.
- Haruta M, Sabat G, Stecker K, Minkoff BB, Sussman MR (2014) A peptide hormone and its receptor protein kinase regulate plant cell expansion. *Science* 343:408–411.
- Zhang GY, Wu J, Wang XW (2010) Cloning and expression analysis of a pollen preferential rapid alkalization factor gene, BoRALF1, from broccoli flowers. *Mol Biol Rep* 37:3273–3281.
- Matos JL, Fiori CS, Silva-Filho MC, Moura DS (2008) A conserved dibasic site is essential for correct processing of the peptide hormone AtRALF1 in *Arabidopsis thaliana*. *FEBS Lett* 582:3343–3347.
- Campos WF, Dressano K, Ceciliato PHO, Guerrero-Abad JC, Silva AL, Fiori CS, Morato do Canto A, Bergonci T, Claus LAN, Silva-Filho MC, Moura DS (2018) Arabidopsis thaliana rapid alkalization factor 1-mediated root growth inhibition is dependent on calmodulin-like protein 38. *J Biol Chem* 293:2159–2171.
- Li C, Yeh FL, Cheung AY, Duan Q, Kita D, Liu MC, Maman J, Luu EJ, Wu BW, Gates L, Jalal M, Kwong A, Carpenter H, Wu HM (2015) Glycosylphosphatidylinositol-anchored proteins as chaperones and co-receptors for FERONIA receptor kinase signaling in Arabidopsis. *eLife* 4:e06587.
- Dressano K, Ceciliato PHO, Silva AL, Guerrero-Abad JC, Bergonci T, Ortiz-Morea FA, Burger M, Silva-Filho MC, Moura DS (2017) BAK1 is involved in AtRALF1-induced inhibition of root cell expansion. *PLoS Genet* 13:e1007053.
- Liu P, Haruta M, Minkoff BB, Sussman MR (2018) Probing a plant plasma membrane receptor kinase’s three-dimensional structure using mass spectrometry-based protein footprinting. *Biochemistry* 57:5159–5168.
- Atkinson NJ, Lilley CJ, Urwin PE (2013) Identification of genes involved in the response of Arabidopsis to simultaneous biotic and abiotic stresses. *Plant Physiol* 162:2028–2041.
- Olsen AN, Mundy J, Skriver K (2002) Peptomics, identification of novel cationic Arabidopsis peptides with conserved sequence motifs. *In Silico Biol* 2:441–451.
- Pearce G, Yamaguchi Y, Munske G, Ryan CA (2010) Structure-activity studies of RALF, rapid Alkalinization factor, reveal an essential-YISY-motif. *Peptides* 31:1973–1977.
- Du C, Li X, Chen J, Chen W, Li B, Li C, Wang L, Li J, Zhao X, Lin J, Liu X, Luan S, Yu F (2016) Receptor kinase complex transmits RALF peptide signal to inhibit root growth in Arabidopsis. *Proc Natl Acad Sci U S A* 113:E8326–E8334.

21. Blommel PG, Martin PA, Seder KD, Wrobel RL, Fox BG (2009) Flexi vector cloning. *Methods Mol Biol* 498:55–73.
22. Lee W, Watters KE, Troupis AT, Reinen NM, Suchy FP, Moyer KL, Frederick RO, Tonelli M, Aceti DJ, Palmenberg AC, Markley JL (2014) Solution structure of the 2A protease from a common cold agent, human rhinovirus C2, strain W12. *PLoS One* 9:e97198.
23. Srivastava R, Liu JX, Guo H, Yin Y, Howell SH (2009) Regulation and processing of a plant peptide hormone, AtRALF23, in Arabidopsis. *Plant J* 59:930–939.
24. Sharma D, Rajarathnam K (2000) ¹³C NMR chemical shifts can predict disulfide bond formation. *J Biomol NMR* 18:165–171.
25. Lee W, Stark JL, Markley JL (2014) PONDEROSA-C/S: client-server based software package for automated protein 3D structure determination. *J Biomol NMR* 60:73–75.
26. Klock HE, Lesley SA (2009) The polymerase incomplete primer extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol* 498:91–103.
27. Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb Cell Fact* 11:56.
28. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. *Protein Expression Purif* 80:283–293.
29. Studier FW (2014) Stable expression clones and auto-induction for protein production in *E. coli*. *Methods Mol Biol* 1091:17–32.
30. Lewis BD, Karlin-Neumann G, Davis RW, Spalding EP (1997) Ca(2+)-activated anion channels and membrane depolarizations induced by blue light and cold in Arabidopsis seedlings. *Plant Physiol* 114:1327–1334.
31. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPIPE - a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
32. Ying J, Delaglio F, Torchia DA, Bax A (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J Biomol NMR* 68:101–118.
33. Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* 31:1325–1327.
34. Shin J, Lee W, Lee W (2008) Structural proteomics by NMR spectroscopy. *Expert Rev Proteomics* 5:589–601.
35. Eghbalnia HR, Wang L, Bahrami A, Assadi A, Markley JL (2005) Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J Biomol NMR* 32:71–81.
36. Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241.
37. Lee W, Cornilescu G, Dashti H, Eghbalnia HR, Tonelli M, Westler WM, Butcher SE, Henzler-Wildman KA, Markley JL (2016) Integrative NMR for biomolecular research. *J Biomol NMR* 64:307–332.
38. Tamiola K, Mulder FAA (2012) Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. *Biochem Soc Trans* 40:1014–1020.
39. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL (2008) BioMagResBank. *Nucleic Acids Res* 36:D402–D408.
40. Lee W, Petit CM, Cornilescu G, Stark JL, Markley JL (2016) The AUDANA algorithm for automated protein 3D structure determination from NMR NOE data. *J Biomol NMR* 65:51–57.