A Flexible and Original Architecture of Two Unrelated Zinc Fingers Underlies the Role of the Multitask P1 in RYMV Spread

Vianney Poignavent, François Hoh, Guillaume Terral, Yinshan Yang, François-Xavier Gillet, Jeong-Hyeon Kim, Frédéric Allemand, Eric Lacombe, Christophe Brugidou, Sarah Cianferani, et al.

Research Article

A Flexible and Original Architecture of Two Unrelated Zinc Fingers Underlies the Role of the Multitask P1 in RYMV Spread

Vianney Poignavent [1,5], François Hoh [2,3†], Guillaume Terral [4†], Yinshan Yang [2,3], François-Xavier Gillet [1,6], Jeong-Hyeon Kim [1], Frédéric Allemand [2,3], Eric Lacombe [1,7], Christophe Brugidou [1,8], Sarah Cianferani [4], Hélène Déméné [2,3*] and Florence Vignols [1,9*]

*1 - **IPME,** Université de Montpellier, IRD, CIRAD, Montpellier, France*
*2 - **Centre de Biochimie Structurale (CBS),** Univ Montpellier, CNRS, INSERM, Montpellier, France*
*3 - **BCM,** Univ Montpellier, CNRS, INSERM, Montpellier, France*
*4 - **LSMBO,** Univ Strasbourg, IPHC, CNRS, Strasbourg, France*
*5 - **IBMP,** Univ Strasbourg, CNRS, Strasbourg, France*
*6 - **CNRS-UCB-INSA-Bayer CropScience,** Lyon, France*
*7 - **DIADE,** Univ Montpellier, CIRAD, IRD, Montpellier, France*
*8 - **PHIM,** Univ Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France*
*9 - **IPSiM,** CNRS, INRAE, Institut Agro, Univ Montpellier, Montpellier, France*

***Correspondence to Hélène Déméné and Florence Vignols:** helene.demene@cbs.cnrs.fr (H. Déméné), florence.vignols@cnrs.fr (F. Vignols)@HDemene (H. Déméné)*
https://doi.org/10.1016/j.jmb.2022.167715
Edited by Eric O. Freed

Abstract

Viruses of the sobemovirus genus are plant viruses, most of which generate very important agricultural and financial losses. Among them, the rice yellow mottle virus (RYMV) is one of the most damaging pathogens devastating rice fields in Africa. RYMV infectivity and propagation rely on its protein P1, identified as a key movement and potential long-distance RNA silencing suppressor. Here we describe P1′s complete 3D structure and dynamics obtained by an integrative approach combining X-Ray crystallography and NMR spectroscopy. We show that P1 is organized in two semi-independent and topologically unrelated domains, each harboring an original zinc finger. The two domains exhibit different affinities for zinc and sensitivities to oxidoreduction conditions, making the C-terminal P1 region a potential labile sensor of the plant redox status. An additional level of regulation resides on the capacity of P1 to oligomerize through its N-terminal domain. Coupling P1 structure information with site-directed mutagenesis and plant functional assays, we identified key residues in each zinc domain essential for infectivity and spread in rice tissues. Altogether, our results provide the first complete structure of a sobemoviral P1 movement protein and highlight structural and dynamical properties that may serve RYMV functions to infect and invade its host plant.

Introduction

Like other viruses, plant viruses have to rapidly establish complex interactions with host cellular machinery to guarantee successful infection. RNA viruses usually present smaller genomes compared to DNA viruses, their genomes being under pressure for genome compression due to high mutation rate during replication.[1–2] Keeping a small size genome ensures the formation of smaller virions, requiring a minimal number of coat proteins

(CP), thus maximizing virus production.[3] Different mechanisms have evolved to multiply viral functions while maintaining genome size. In particular, viruses tend to use one protein for different functions.

After replication in prime infected area, plant viruses rapidly spread in surrounding cells until they reach vascular tissues to establish systemic infection. Viral progeny moves intra-cellularly and cell-to-cell through plasmodesmata by high-jacking cellular machinery with viral encoded movement protein (MP).[4–6] Each virus shows MP's sequence singularity and mode of action.[7–9] To initiate systemic infection, viruses have also to overcome sophisticated defense mechanisms set up by host plants, among which RNA interference (RNAi) is the principal one.[10] To counteract this major defense mechanism, phytoviruses have evolved towards the production of proteins capable to act as viral suppressors of RNA silencing (VSR) to prevent RNAi-triggered immunity.[10–11] Like MPs, VSR proteins do not show sequence or molecular organization similarities. Only six complete VSR structures were solved so far.[12–18]

The *rice yellow mottle virus* (RYMV) is a positive single strand RNA *sobemovirus* that spreads in rice cultures from East to West Africa since the sixties as fast as most animal viruses,[19–20] generating crop losses ranging from 20 up to 100%. Although considered to be the major biotic constraint to rice cultivation in Africa,[21] no effective means of combating this virus have been discovered to date. The most emblematic multifunctional protein in RYMV is P1, a small cysteine-rich protein of 18 kDa encoded by ORF1 of the genome first identified as the RYMV movement protein essential for viral spread,[22–23] but also suspected to promote viral replication.[22] It has also been described as a potential VSR in plant reporter systems,[23–27] but this finding sparks controversy. Due to its multifunctional role, P1 is the ideal candidate for the development of strategies against RYMV.

Despite crucial roles of P1 in early steps of viral infection, little is known concerning its structure and mode of action. We previously showed that P1 binds two Zn atoms in its reduced state, one atom being predicted to form a zinc finger with two highly conserved pairs of cysteine residues organized in CXXC motifs,[28] a common feature in all P1 proteins of *sobemovirus* genus.[29] Mutating two cysteines in this putative zinc]finger completely abolishes both silencing suppression and cell-to cell movement functions of P1.[23] Whether the second zinc atom participates in an additional zinc finger whatever its function is unpredictable and the overall P1 architecture remains fully obscure.

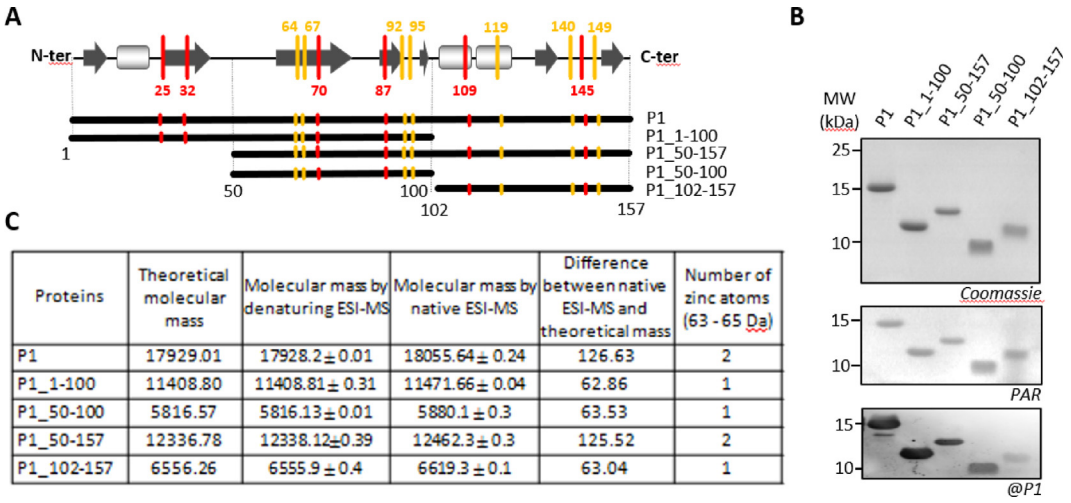
In this paper, we provide a comprehensive view of the structure–activity relationship for RYMV P1 by an integrative approach combining biochemistry, mass spectrometry, X-ray crystallography, NMR, and plant infection assays. We show that P1 presents a unique architecture organized around two different domains, each harboring a zinc finger unrelated to the other. In particular, the C-terminal domain which appears as essential for full virus infectivity, contains a HCHC box rarely encountered in nature and presenting a unique zinc coordination. P1 domain organization allows for homo-dimerization through its N-terminal domain and we show the crucial role for infectivity of the aromatic residues involved. We also could identify one motif in each domain required for virus accumulation, namely a helix and a stretch of aromatic and basic residues in the N- and Ct-terminal regions respectively.

Results

P1 possesses two zinc finger domains with different affinities for zinc

In a previous work, we showed that P1 binds two zinc atoms with distinct affinities in a reduced environment,[28] but we were unable to unambiguous identify zinc ligands because of the high number of rigorously conserved cysteines (7 Cys residues) and histidines (6 His residues) (Figure 1(A) and Figure S1). Here we designed a series of recombinant P1 regions (Figure 1(A) and Supplementary Results, Figures S2 to S4, and Tables S1 and S2) and evaluated their capacity to bind zinc using a 4-(2-Pyridylazo) resorcinol (PAR) labeling assays and native ESI-MS (Figure 1(B and C) and Figure S5). Zinc detected by the PAR probe co-localized with all P1 regions assayed in an SDS-PAGE and mass spectrometry yielded the corresponding number of zinc atoms bound. Resizing P1_1-100 and P1_50-157 proteins into shorter fragments identified [50–100] and [102–157] as independent regions each binding a zinc atom, subsequently referred to as ZnF1 and ZnF2, respectively.

To determine the zinc affinity of both zinc fingers, we subjected them to EDTA treatments, in conjunction with full P1 (Table S3). Addition of 400 and 50 molar equivalents (eq) of EDTA provoked the release of one zinc atom for full P1 and ZnF2 respectively, which could be reloaded after exposure to 100 μM and 60 μM ZnOAc, respectively (Table S3 and Figure S6(A and C)). By contrast, all attempts to provoke the release of the second zinc atom from P1 and the release of the unique zinc atom of P1_1-100 using higher EDTA concentrations were unsuccessful (Table S3 and Figure S6(B)). Hence, ZnF2 appears as the zinc finger with the weaker affinity for zinc in P1. In order to determine ZnF1 zinc affinity constant, we subjected P1_1-100 to N,N,N, N-Tetrakis(2-pyridylmethyl)-ethylenediamine) (TPEN), a heavy metal chelator with a Ka constant stronger than that of EDTA (10^{16} M^{-1} versus 10^{14} M^{-1}).[31–32] We were thus able to estimate a zinc



C

Proteins	Theoretical molecular mass	Molecular mass by denaturing ESI-MS	Molecular mass by native ESI-MS	Difference between native ESI-MS and theoretical mass	Number of zinc atoms (63 - 65 Da)
P1	17929.01	17928.2 ± 0.01	18055.64 ± 0.24	126.63	2
P1_1-100	11408.80	11408.81 ± 0.31	11471.66 ± 0.04	62.86	1
P1_50-100	5816.57	5816.13 ± 0.01	5880.1 ± 0.3	63.53	1
P1_50-157	12336.78	12338.12 ± 0.39	12462.3 ± 0.3	125.52	2
P1_102-157	6556.26	6555.9 ± 0.4	6619.3 ± 0.1	63.04	1

Figure 1. The RYMV P1 protein binds two zinc atoms at its central and C-terminal regions. (A) Scheme of P1 regions designed for production of recombinant proteins based on the Predict server.[30] Cys and His residues are positioned with yellow and red bars, respectively. Predicted secondary structures are represented by dark grey arrows for β-sheets and by light grey squares for α-helices. Thick black lines delineate P1 protein and its resized sub regions, with amino acids numbering (*Below*) and corresponding names (*Right*). **(B)** In-gel co-localization of zinc with purified P1 sub regions separated by non-reducing 18% SDS-PAGE, with Coomassie staining (*Top*), zinc detection by the PAR probe (*Middle*) and immunodetection by a purified anti-P1 antibody at 1:1000 dilution (*Bottom*). **(C)** Quantification of zinc atoms in P1 recombinant sub regions by native ESI-MS. Deconvoluted ESI mass spectra are given in the Figure S5.

affinity constant Ka of 2,15x10^{16} M^{-1} for ZnF1 (Figure S7).

Because oxidizing environments provoke zinc release from P1,[28] we further compared the respective redox sensitivity of P1_1-100 and P1_102-157 to that of the full length P1 (Figure S8). We found that high H_20_2/protein ratios (80 eq corresponding to 2 mM) were required to induce the release of zinc atom from the recombinant P1 and P1_1-100 proteins, without affecting their electrophoresis mobility nor their conformation as monitored by circular dichroism (Figure S8(A and B)). Reversely, lower H_20_2/protein ratios were sufficient to provoke zinc release from the recombinant P1_102-157 protein, accompanied by a mobility shift and conformational changes (Figure S8(C)). These results identified P1 C-terminal [102–157] ZnF2 region not only as the less affine to zinc but also as the most sensitive region to oxidant molecules, including from a conformational point of view.

P1 N- and C-terminal regions exhibit different structural and zinc binding geometries

To go further in the structural characterization of P1, we subjected the protein to crystallography. We could not find crystallization conditions for the full length P1, but P1_1-100 and P1_102-157 fragments yielded high quality diffracting crystals (Table 1 and Figure 2). P1_1-100 consists in a sandwich composed of 6 strands (A-F) and 2 short helices (H1 and H2) at each extremity (Figure 2(A-B)). Searching for similar folds with DALI[33] returned the sandwich family composed of six strands and two β-sheets as hit. Another remarkable feature is the presence of a helix harboring a lysine (Lys17) in H1 completely solvent exposed (Figure 2(A and B)). This lysine is highly conserved (Figure S1), tolerating only replacement by arginine making this residue a potential anchor for interaction with cellular partners. The C-terminal extreme side of the P1_1-100 sandwich contains ZnF1 involving Cys64, Cys67, Cys92 and Cys95 (Figure 2 (C)). Both Cys64 and Cys67 belong to a β turn between strands D and E, whereas Cys95 is located at the beginning of helix 2, classifying ZnF1 in the treble clef group, albeit not with a canonical form as Cys92 does not belong to helix 2.[34–35] Additional SDS-PAGE and ESI-MS experiments on an extensive series of cysteinic and histidinic recombinant mutants (Table S2 and Figure S4) highlighted the importance of this CCCC-type ZnF1 for P1 zinc binding and for solubility of both P1 and P1_50-100. Although zinc loss was partially observed for the P1^{C95S} mutant (Figure S9(C)), all mutants (P1^{C64S}, P1^{C67S}, P1^{C92S} and P1^{C95S}) of ZnF1 zinc ligands in the full-length P1 were insoluble (Supplementary Results, Figure S9(A and B) and Table S2). Furthermore, we could not detect zinc in either the P1^{C64S} and P1^{C67S} single mutants or the P1$^{C92S-C95S}$ double mutant of the shorter P1_50-100 fragment (Figure S11(A and C)).

Contrary to P1_1-100, P1_102-157 has an original fold that could not be retrieved from any

Table 1 Data collection and refinement statistics.

	P1_1-100	P1_102-157-phasing	P1_102-157
Data collection			
Wavelength (Å)	0.976	1.28	1.24
Space group	F 41 3 2	P 1 21 1	P 1 21 1
Cell dimensions			
a, b, c (Å)	195.61	38.14 31.23 42.03	
A, β, γ(°)		90.00 101.08 90.00	90.00 101.10 90.00
Resolution (Å)[a]	112.94–2.10 (2.21–2.10)	41.25–2.10 (2.21–2.10)	37.37–1.98 (2.09–1.98)
Rmerge	0.051 (0.559)	0.154 (0.448)	0.098 (0.573)
N° of reflections	724,744 (104965)	55,242 (7957)	22,349 (3282)
N° of unique reflections	19,296 (2756)	5834 (839)	6911 (992)
I/s*I*	8.3 (1.1)	12.4 (5.0)	8.6 (2.0)
Completeness (%)	99.97 (100.0)	100.0 (99.9)	99.2 (99.4)
Multiplicity	37.5 (38.1)	9.5 (9.5)	
Anomalous completeness	100.0 (100.0)	99.9 (99.9)	
Anomalous multiplicity	20.1 (19.9)	4.6 (4.5)	
Refinement			
R_{work} / R_{free} (%)	0.185 / 0.214		0.193 / 0.227
N° atoms	1688		899
Protein	1547		847
Zn	3		2
Water	140		50
B-factors (Å^2)			
Protein	45		27
Zn	52		22
Water	53		29
rmsd			
Bond lengths (Å)	0.009		0.007
Bond angles (°)	1.166		1.069
Ramachandran favored (%)	98.95		97.88
Ramachandran outliers (%)	0.0		0.0

structural database. It consists in one α helix (residues 103–126) at the N-terminal region of this domain and in a rigid part formed by a succession of loops (residues 127–157, Figure 2(D)). It also contains the second P1 zinc finger ZnF2 formed by His109 in the N-terminal helix and by Cys140, His145 and Cys149 located in the loops, His109 and His145 involving their Nε2 and Nδ1 atoms, respectively (Figure 2(E)). Zinc ligands were confirmed by additional SDS-PAGE and ESI-MS experiments on cysteinic and histidinic mutants of this 102–157 region (Figure S11 and Supplementary Results). Together they form an extremely rare HCHC ZnF2 motif.

A search in ZincBind, the database of zinc proteins,[36] identified only six proteins with such architecture. Among them, only the HCHC box of the Topless related protein was similarly organized, with the N-terminal ligand in a helix and the three others belonging to loops, although with an inversed Nδ1/Nε2 coordination geometry (Table 2). Hence, P1 ZnF2 appears to be a completely new zinc finger motif. This motif stabilizes P1 C-terminal 3D structure, while EDTA-mediated zinc chelation as monitored by ^{1}H NMR promotes its unfolding (Figure S13). In addition, P1_102-157 is characterized by the presence of a basic patch at its extreme C-terminal region, overlapping with conserved aromatic residues (Arg150, Phe151, Lys153, Tyr154 and Phe157) (Figure 2(D) and Figure S1).

Deciphering the overall P1 structure by NMR

To unravel the structure of the full-length P1 protein, we led a parallel NMR investigation on P1 and its two sub-regions P1_1-100 and P1_102-57 (Figure 3(A and B)). ^{1}H and ^{15}N chemical shifts of residues were identical between the isolated domains and the whole protein. Titration of ^{15}N-^{1}H-labelled P1_102-157 by the unlabeled P1_1-100 fragment showed no change in P1_102-157 amide chemical shifts Figure 4(A)). This analysis unambiguously revealed that the structures of the isolated domains were retrieved in the full-length protein, with no strong contact between them, a finding also observed in an Y2H assay (Figure S14). Translation of chemical shifts into backbone dihedral angles with the TALOS software (Table 3) and analysis of the ^{1}H–^{1}H NOEs pattern (Figure 4B) established that the C-terminal helix in P1_1-100 (Ala93 to Glu96 in P1_1-100, Figure 2(A)) and the N-terminal helix in P1_102-157 (Glu104 to Glu126, Figure 2(D)) seen in X-Ray structures were extending to a long helix from Ala93 to Glu126, covering the entire linker. To determine the level of rigidity of this helix, we

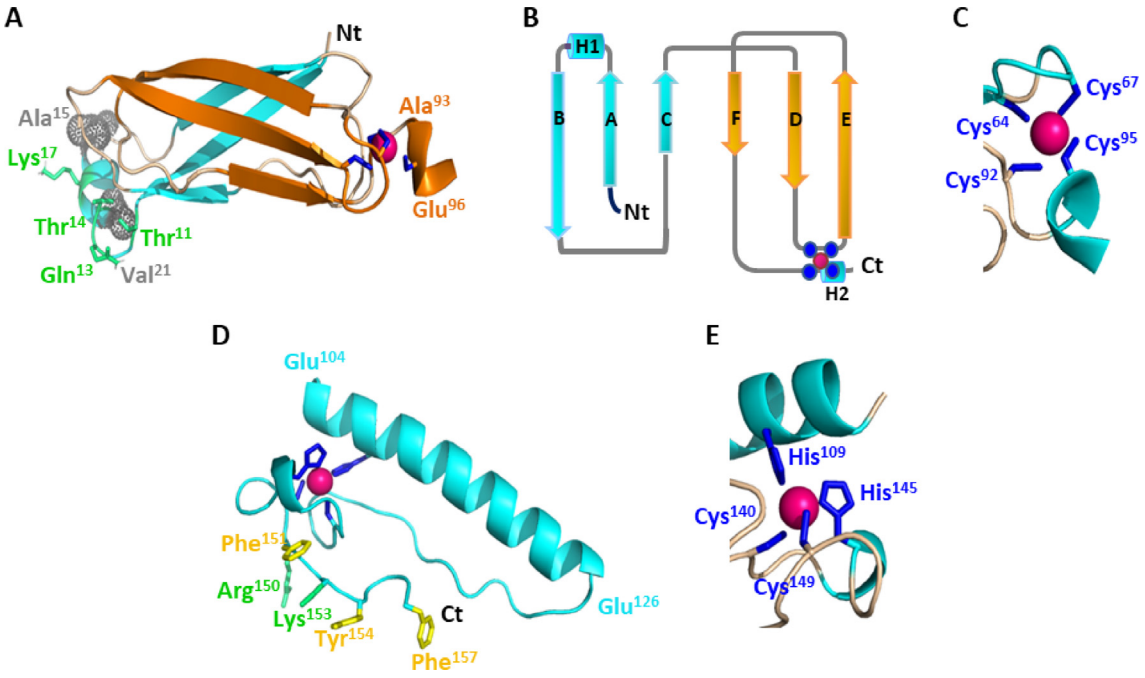


Figure 2. Crystallographic structure of P1 N- and C-terminal regions.(A) Structure of P1_1-100 fragment. This domain is made of three anti-parallel β-sheets facing each other. It presents a small helix with a Lysine side chain residue (Lys17) completely solvent exposed. This helix is stabilized by contacts of hydrophobic residues (side chains represented as grey dotty spheres) that interact with the protein core, and by a network of hydrogen bonds involving the conserved Thr14, Gln13 and Thr11 residues. **(B)** Schematic representation of P1_1-100 structure. Helices (H1, H2) and β-sheets are depicted as cylinders and arrows, respectively. Cys residues chelating the zinc atom are represented as blue plain circles. **(C)** Close-up of the Cys-treble clef ZnF1 in P1_1-100, with side chains of zinc-chelating cysteine residues in stick representation. **(D)** Cartoon representation of P1_102-157 ribbon. Side chains of conserved positively charged (in green) and aromatic (in yellow) residues are depicted in stick representation. Side chains of the HCHC ZnF2 are colored in blue. **(E)** Close-up of the HCHC ZnF2 in P1_102-157, with side chains of zinc-chelating cysteine and histidine residues in stick representation.

Table 2 List and topology of proteins harboring zinc-binding HCHC boxes found in the ZincBind database.[36]

Protein*	PDB code	Histidine coordination mode†	Protein family	ZnF fold
P1-ZnF2	6TY2	Nε2/Nδ1	Silencing repressor	1st: helix middle, 2nd, 3th, 4th: loops
Bacillus subtilis fructokinase	1XC3	Nδ1/ Nδ1	Transferase	1st: helix C-ter, 2nd, 3th: knuckle, 4th: helix Nter
Human Histone Deacetylase 4	2VQJ	Nε2/ Nε2	Hydrolase	1st: C-ter helix, 2C-ter: helix, 3th, 4th: loops
Topless related protein	5c7E	Nδ1/Nε2	Transcriptional corepressor	1st, 2nd: helix Cter, 3th, 4th: loops.
SARS coronavirus nsp14	5C8T	Nε2/Nδ1	Transferase	1st, 2nd: helix Cter, 3th: loop, 4th: helix
Clostridium difficile stage ii sporulation protein d (spoiid)	5I1T	Nε2/Nδ1	Hydrolase	1st, 2nd, 3th, 4th: loops
Chlamydia abortus effector protein ChlaDUB	6GZU	Nε2/Nδ1	Transferase	1st: C-ter helix, 2C-ter: helix, 3th, 4th: loops

*Only a representative member of each family class is cited in cases of several homolog structures are available.
N-ter, middle and C-ter refer to the location of the zinc ligand within the helix.
† The coordination mode is given for the N-terminal histidine first.

performed residual dipolar couplings (RDC) and ^{15}N relaxation data measurements and analysis. Differences between the component values of the relaxation and alignment tensors of N and C-terminal domains established that this helix is not rigid (Figure 4(C and D), Table 4 and Table 5).

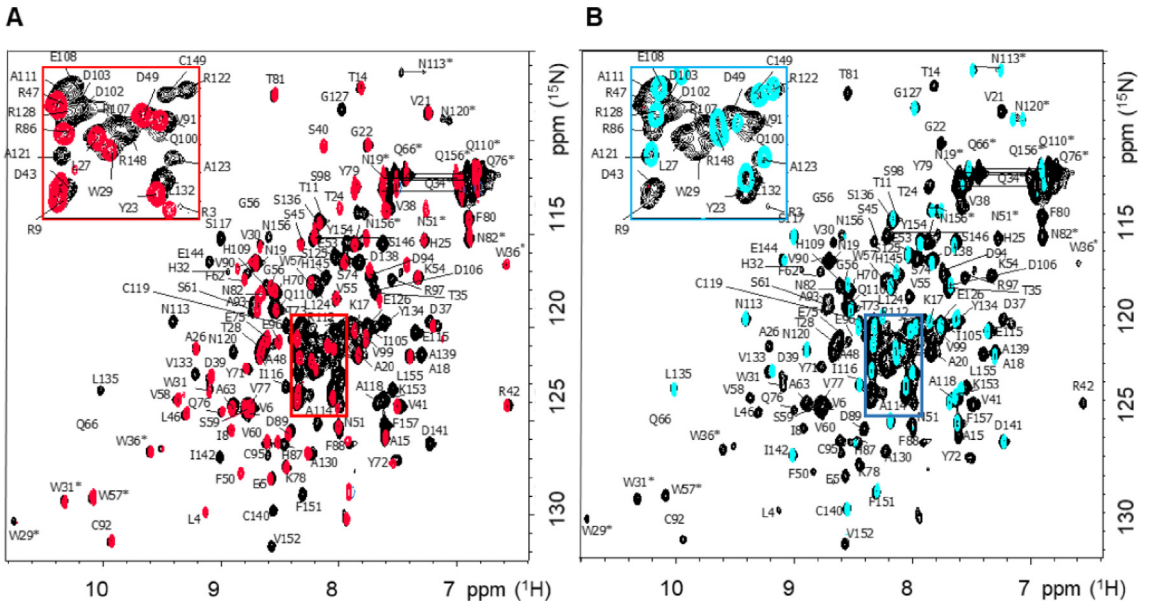


Figure 3. NMR characterization of P1 fragments. (A-B) Superposition of fingerprint regions of ^{1}H-^{15}N HSQC of full length P1_1-157 (black cross-peaks) with (**A**) P1_1-100 (red cross-peaks) and with **(B)** P1_102-157 (blue cross-peaks). Sequence specific resonance assignments are indicated by the residue numbers and by one-letter amino acid codes. Pairs of side-chain NH_2 resonances are connected by horizontal lines. Tryptophan side chain indole peaks are indicated with asterisks (*). The central region is shown as an insert.

Table 3 Phi and Psi angles estimated by Talos [38] from ^{1}Hα, ^{13}CO, ^{13}Cα, ^{13}Cβ chemical shifts.

Residue number	Residue type	Phi (°)	Psi (°)
95	CYS	−79 ± 39	−59 ± 19
96	GLU	−88 ± 48	−55 ± 15
97	ARG	−100 ± 36	−60 ± 8
98	SER	−100 ± 52	−59 ± 1
99	VAL	−141 ± 41	−68 ± 52
101	LEU	−83 ± 43	−60 ± 16
102	ASP	−104 ± 36	−56 ± 16
103	ASP	−84 ± 44	−60 ± 20
104	GLU	−85 ± 45	−63 ± 23
105	ILE	−83 ± 43	56 ± 16
106	ASP	−82 ± 42	60 ± 20
107	ARG	−82 ± 42	63 ± 23
108	GLU	−86 ± 46	55 ± 15

Following the formalism of Tolman,[37] we derived a value of 0.78 for a global order parameter of ZnF2 relatively to ZnF1. This value shows that both regions possess a significant degree of spatial freedom, corresponding to a motion amplitude of ± 30 degrees of ZnF2 (relative to ZnF1), under the assumption of an axially symmetric “wobble-in-a-cone” movement. However, we could not find evidence that this mobility is due to a particular disorder at the local level of the linker (Figure 4 (C and D)). It might be due to the lack of interaction of the long linker helix with any other part of the protein.

Integration of X-Ray structures and NMR parameters allowed the calculation of overall structure of entire P1, where indeed, the linker region folded as a helix (Figure 5(A)). The whole P1 protein presents a surface with an original polarized profile (Figure 5(B)). On one face, the central region harbors a large negatively charged patch, formed by conserved aspartic and glutamic residues. The other side harbors two extremal positively charged poles overlapping with the potential interaction motifs described above. In the N-terminal domain, Arg9, Lys17 and Lys78, side chains form indeed a hydrophilic and positively charged patch as do the basic residues of the C-terminal extremity (Arg150, Lys153) (Figure 2(D) and Figure 5(B)). The latter is interestingly juxtaposed to a negatively charged cleft formed by Ser131 side chain and the carboxylic function of extremal Phe157 (Figure 5(B)).

Oligomeric structure of P1

A remarkable feature of P1_1-100 crystal structure was the formation of a dimer from paired antiparallel monomer (Figure 5(C) and Figure S14). In the asymmetric unit, the dimeric interface covers 506 Å^2 and involves five hydrogen bonds, 10 hydrophobic contacts and four salt bonds (Figure 5(C and D), Figure S15(A) and Table S4). It is classified as labile by the PISA algorithm,[41] in accordance with the faint proportion of P1 dimers detected in ESI MS analysis for full length P1.[28] Nevertheless, the probability of being

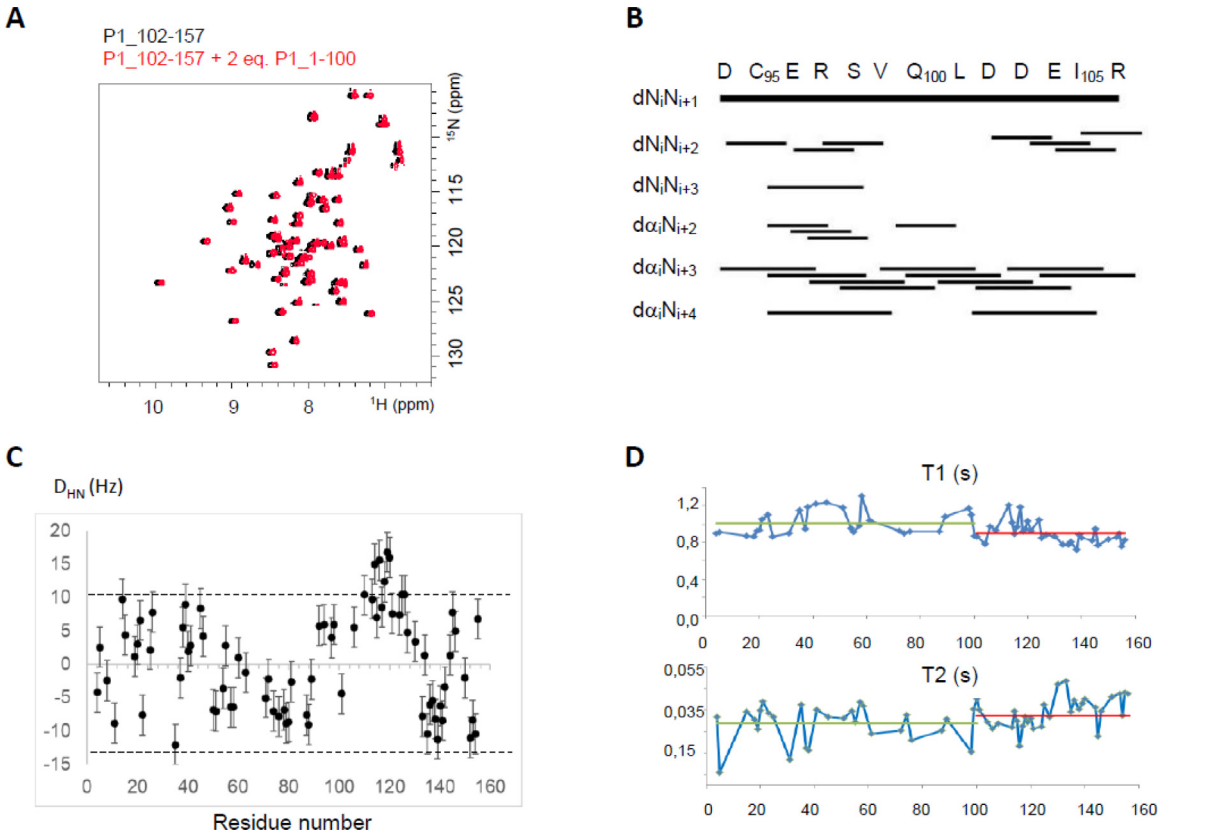


Figure 4. Characterization of the relative mobility of P1 N- and C-terminal domains. (A) Titration of ^{15}N labelled P1_102-157 by P1_1-101 seen by NMR. The addition of two equivalent of non-labelled P1_1-100 in a sample of ^{15}N labelled P1_102-157 produces no shift nor broadening of any resonance in the ^{1}H-^{15}N HSQC spectrum of P1_102-157 (red spectrum). For a better visualization, the spectrum of the P1_1-100/P1_1-157 mixture has been shifted by a few hertz. (**B**) Sequential and medium range NOEs reported between proton atoms for the linker sequence (amino acids 94–106) characterizing their folding into a helix. **(C)** $^{1}D_{HN}$ residual dipolar couplings measured for full-length P1_1-157 in phage medium depicted as a function of the protein sequence. The maximal $^{1}D_{HN}$ values of the P1_1-100 domain are enlightened by dotted lines. The P1_102-157 domain exhibits several $^{1}D_{HN}$ below and above these threshold values, showing that both regions exhibit different alignment properties. **(D)** ^{15}N T_1 and T_2 relaxation times recorded for full-length P1_1-157, as a function of residue number. The differences observed between mean values for P1_1-100 (green lines) and for P1_102-157 (red lines) support a significant degree of flexibility between them.

Table 4 Components of the rotational diffusion tensors derived separately for P1 subdomains from experimental P1_1-157 ^{15}N T_1 and T_2 values or computed for the NMR derived structure.

Domain	$D_{\parallel}^{exp}$ (10^7 s^{-1})[a]	$D_{\perp}^{exp}$(10^{-7} s^{-1})[a]	τ_{ic}^{exp} (10^{-9} s)[b]	$D_{\parallel}^{calc}$(10^7 s^{-1})[c]	$D_{\perp}^{calc}$(10^{-7} s^{-1})[c]	τ_{ic}^{calc}(10^{-9} s)[d]
P1_1-100	1.55 ± 0.06	0.91 ± 0.07	14.5 ± 0,2			
P1_102-157	1.84 ± 0.07	1.12 ± 0.02	12.9 ± 0,2			
Full length P1				1.7	0.4	19.9

[a] Diffusion tensor components were estimated by TENSOR2[39] from ^{15}N relaxation data recorded on full-length P1_1-157 but taking the data for each domain (P1_1-100 and P1_102-157) separately. Estimated components are smaller for the N-terminal domain than for the C-terminal domain, reflecting the slower tumbling time of the former τ_{ic}^{exp} in the full-length protein.

[b] Overall global estimated correlation time computed from the components of the rotational diffusion tensor estimated from the relaxation times.

[c] Diffusion tensor components were computed from the NMR derived structure of P1_1-157 by HYDRONMR.[40] The calculated components are smaller than the estimated components from experimental data, as is expected for a two-domain protein exhibiting inter-domain flexibility.

[d] Predicted overall global correlation time calculated from the NMR-XRay derived structure by HYDRONMR.

Table 5 Axial and rhombic components of the alignment tensors A_{II} and $A_{\perp}$[a] derived separately for P1 subdomains from experimental ^{1}H-^{15}N RDCs.

Domain	A_{II}[a]	$A_{\perp}$[a]
P1_1-100	4.3 ± 0.5	1.1 ± 0.4
P1_102-15	5.2 ± 0.4	2.7 ± 0.5

[a] Alignment tensor components were estimated by the Module software[75] from ^{1}H-^{15}N RDCs recorded on full-length P1_1-157 but taking the data for each domain (P1_1-100 and P1_102-157) separately.

a biological dimer reaches 73% in ClusPro-DC calculations, an elevated score.[42]

Analysis of the crystal packing with the PISA server indicated that P1_1-100 could also undergo a conformational change into a tetrameric structure stabilized by the coordination of a fifth zinc atom by the four His70 residues of P1 N-terminal regions (Figure S15(B)). This crystallographic interface, estimated at 574 Å^2, involves six hydrogen bonds, and two saline bonds via the bridging of water molecules

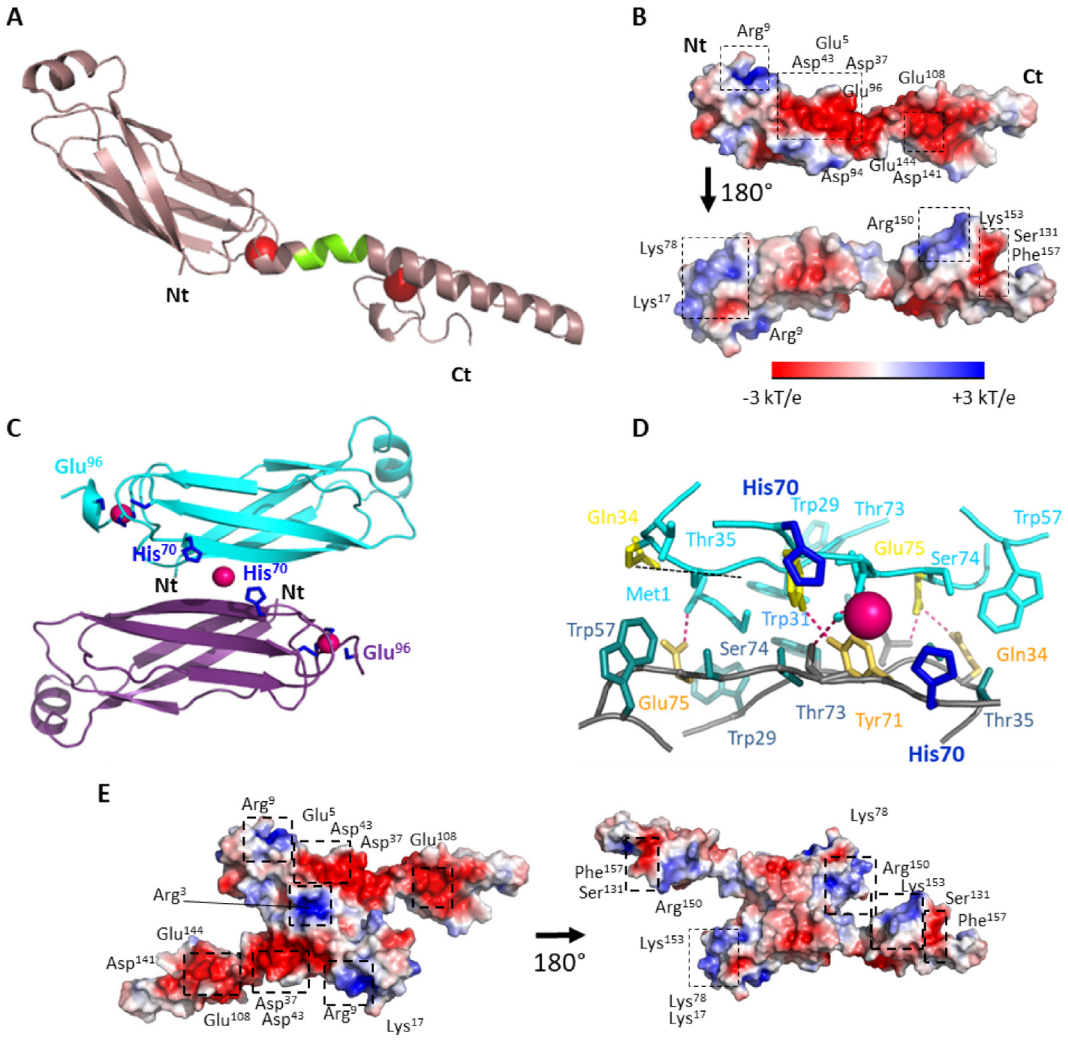


Figure 5. Structure and oligomerization states of full-length P1 by NMR. (A) Model of P1_1-157 built from rigid-body calculations. The missing helix in crystallographic structures in the linker region has been reconstructed on the basis of NMR restraints (green portion). **(B)** Electrostatic surface potential representation of P1, showing on one side a large central negatively charged surface and on the other side, two extremal positive poles. Conserved positive and negative patches among the reported 51 P1 sequences are framed. **(C)** Dimer interface of the asymmetric unit. The central Zn atom is represented as a pink sphere, chelated by His70 (dark blue) in stick representation. Side chains of residue involved in polar and hydrophobic contacts in the dimeric interface are represented in stick and are colored in yellow and cyan blue, respectively. **(D)** Zinc-dependent oligomerization of P1_1-100 and full length P1_1-157 proteins. SEC MALLS profiles for P1_1-100 (*Left*) and for full length P1_1-157 (*Right*) are depicted with UV absorbance as a function of eluted volumes of a buffer with (red lines) or without (blue lines) zinc. Calculated masses are represented on the Y-axis**. (E)** Representation of the electrostatic surface of the asymmetric P1 dimeric model. Conserved positive and negative patches are framed.

(Figure S15(C and E), Table S4), but has an 88% probability of being a crystallographic artifact in ClusPro-DC simulations. The instability of protein precluded a NMR analysis at low salt concentration, but ^{15}N relaxation data were in favor of a monomer at high salt concentration and in presence of zinc (Table 4). To challenge the possibility of a zinc-mediated oligomerization, we subjected the full-length P1 and its truncated form P1_1-100 to size-exclusion chromatography coupled with multi-angle laser light scattering (SEC-MALS). At low salt concentrations, both the full-length protein and its N-terminal region eluted as monomers in a zinc-free elution buffer and as a dimer in a zinc-containing elution buffer (Figure S16). Hence, zinc might at least facilitate the presence of a dimeric interface.

We next built the corresponding models of homodimeric and tetrameric full-length P1 to check the potential regulation brought by oligomerization. The central acidic patch observed in the monomeric found is also observed in the dimeric form, with an almost doubled surface although having a positive protrusion due to arginine 3. The two potential interaction sites of ZnF1 and ZnF2 are spatially close in the dimeric structure and can act cooperatively by modulating their relative accessibility (Figure 5(E)). Consequently, the tetrameric form of P1 maintains the large central negative patch and the extremal positive poles in ZnF1 and ZnF2, but their access, particularly that in ZnF2, are even more restricted (Figure S15(F)).

Key residues of the N-terminal domain are critical for viral spread in rice

To evaluate the biological importance of P1 residues highlighted in the structural study, we took advantage of FL5, a DNA clone of the RYMV genome allowing the production of synthetic infectious viral RNAs for direct inoculation in rice plantlets[43] (Supplementary Results and Figure S17. We first focused on the short solvent-exposed [Lys12-Val21] helix at the N-extremal part of the P1 sandwich (Figure 2(A and B)) and created a corresponding FL5-P1$^{\Delta 12\text{-}21}$ clone expressing a P1 protein depleted of the [Lys12-Val21] stretch. Inoculating plants with FL5-P1$^{\Delta 12\text{-}21}$ RNAs did not provoke viral infection, neither in inoculated leaves nor in systemic tissues according to very few amounts of viral particles detected (Figure 6(A)). FL5-P1$^{\Delta 12\text{-}21}$-inoculated plants also retained green healthy leaves and normal growth compared to plants inoculated with the non-infectious FL5ΔP1 RNAs (Figure 6(B)). Noticeably, we failed to produce a soluble P1$^{\Delta 12\text{-}21}$ recombinant protein in *E. coli* cells. Using a plant reporter system in rice protoplast as a mean to evaluate P1 behavior *in vivo*, we also observed that P1$^{\Delta 12\text{-}21}$ fused to GFP markers (either P1$^{\Delta 12\text{-}21}$-EGFP or EGFP-P1$^{\Delta 12\text{-}21}$ fusions) tend to aggregate in the cytosol contrary to wild type fusions that reside both in the cytosol and in the nucleus (Figure S18). Altogether, these experiments suggest that the solvent-exposed [Lys12-Val21] helix is required at least for proper P1 folding and reveal its essential role for efficient RYMV infectivity.

We next focused on the 4 strictly conserved tryptophan residues in P1[43] (Figure S1), three of them participating to P1 dimeric interfaces. Trp29 and Trp31 belong to the dimeric interface of the asymmetric unit (Figure 5(D) and Figure S19(A)) and Trp57 to both the asymmetric and the crystallographic ones (Figure S15(C)). By contrast, Trp36 is buried in a hydrophobic core of ZnF1 (Figure S19(B)), on the same β-strand as Trp29 and Trp31. While FL5-P1^{W57A} had very minor effects on RYMV fitness, all other FL5-P1$^{Trp\text{-}to\text{-}Ala}$ mutant clones provoked an important decrease in neosynthesized RYMV particles in both inoculated and systemic leaves (Figure 6(A and B) and Figure S17). The most remarkable effect was observed with FL5-P1^{W36A}, whose mutation abolished almost completely the RYMV replication and systemic movement similarly to that of the non-infectious FL5ΔP1, leading to asymptomatic and healthy plants (Figure 6(A and B) and Figure S17(A and B)). Corresponding mutant P1^{W29A} and P1^{W31A} proteins were soluble in *E. coli* and in rice protoplasts as GFP fusions, while P1^{W36A} remained insoluble in bacterial pellets and tend to aggregate in the plant cytosol (Figures S18 and S20). It thus seems that the importance of Trp29 and Trp31 for virus infectivity is due to their participation in the dimeric interface of the dimer present in the asymmetric unit, whereas Trp36 would mainly play a role in the structural maintenance of this domain.

The C-terminal atypical ZnF2 structure is required for P1 viral functions

The cysteine residues identified as essential zinc ligands within ZnF1 have previously been shown to be crucial for virus infectivity and involved in P1 activity as a VSR.[23] Our present structural data here show that ZnF2 also ensures a central function in the folding of P1 C-terminal region, in a redox-dependent fashion. To address its biological role, we modified ZnF2 redox context by creating FL5-P1^{C140H} and FL5-P1^{C149H} clones, reasoning that Cys-to-His mutations preserve coordination[45] (Figure S11). These Cys-to-His ZnF2 P1 mutants kept patterns and behavior unchanged in rice protoplasts contrary to Cys-to-Ser mutants which tend to form aggregates in the cytosol (Figure S18). Both clones appeared noninfectious in rice plants, as shown by the almost complete lack of viral particles in inoculated and systemic tissues and by their unaffected growth and healthy leaf phenotypes (Figure 5(C and D)). This indicates that integrity of Cys140 and Cys149 is mandatory for RYMV infec-

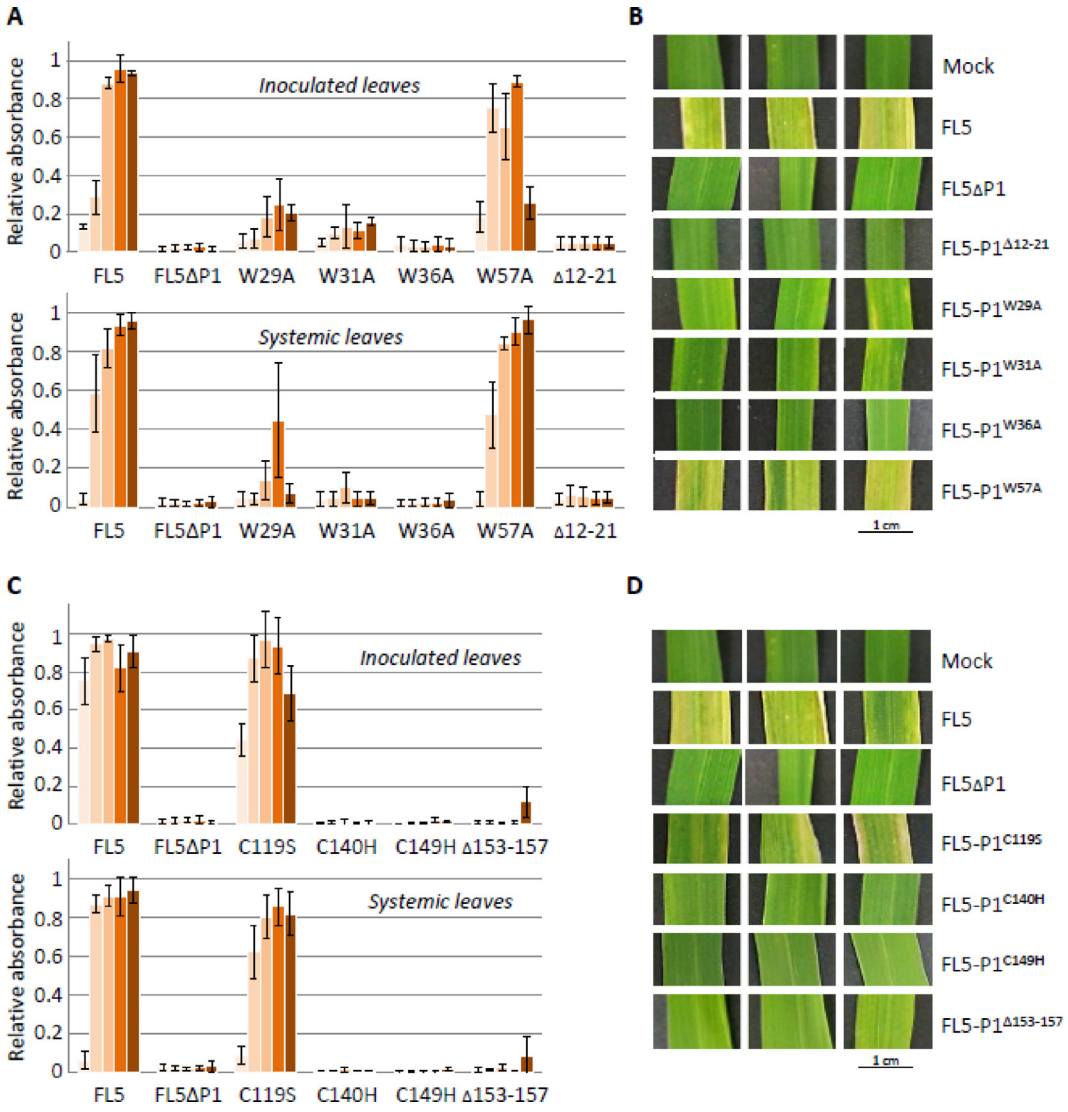


Figure 6. Incidence of mutations in the P1 N- and C-terminal regions on RYMV infectivity in rice.(A-B) Consequences of alpha helix[12–21] deletion and of TRp residue mutations at the P1 N-terminal region on **(A)** RYMV particle accumulation and **(B)** yellowing symptoms at leaf surfaces. **(C-D)** Consequences of His or Cys mutations and of C-ter KYNLF motif deletion at the P1 C-terminal region on **(C)** RYMV particle accumulation and **(D)** yellowing symptoms at leaf surfaces. **(A,C)** Viral charge of rice plantlets infected by synthetic RYMV RNAs obtained from infectious FL5, non-infectious variant ΔP1, or from FL5 clones carrying mutations at the N-terminal P1 region (indicated on X axes by their mutation types) was recorded by DAS-ELISA from inoculated **(A** and **C**, *Top***)** and systemic **(A** and **C**, *Bottom***)** leaves collected weekly during 5 weeks (light to dark bars). Values are the means ± standard error (bars) for FL5, W57A and C119S infectious clones (n = 30) and for all other mutated infectious clones ($10 \leq n \leq 15$ due to the high rate of lethality provoked by P1 mutations). FL5 and FL5ΔP1 are positive and negative controls for infection and viral systemic spread, respectively. **(B** and **D)** Yellowing symptoms were recorded at the leaf surface of inoculated plants. Three samples of rice leaves are shown as representative of symptoms visible at the overall plant scale. Leaves were photographed 5 weeks post inoculation.

tivity. Notably, although Cys119 is also strictly conserved in P1 and able to undergo redox modification,[28] its mutation only induced a slowing down of the accumulation of viral particles during the first days of the infection, but did not globally modify the propagation of the virus in the systemic tissues (Figure 6(C and D)).

Another striking feature of ZnF2 structure is its highly hydrophobic and strongly conserved KYNLF sequence ending C-terminal region (Figure S1) and constituting a part of ZnF2 aromatic basic pole. The carboxylic group of Phe157 is rigid and contributes to a small negatively charged pole next to this basic spot

(Figure 4(B)). Deleting this KYNLF stretch from P1 does not alter its sub-cellular patterns nor its 3D structure (Figures S18 and S21), and yet, plants infected by FL5-P1$^{\Delta 153\text{-}157}$ did not develop RYMV symptoms indicating that the KYNLF stretch plays an essential role for RYMV infectivity (Figure 5(C and D)).

Discussion

Here, we report the first complete 3D structure of an essential protein of a sobemovirus infecting the monocotyledonous model plant and major crop *Oryzae sativa*. Thus, the P1 structure contributes to fulfill an important gap of knowledge regarding this important class of pathogens.

We show that P1 contains two independent subdomains, harboring each a zinc finger with completely different architectures. Other viral proteins were previously shown to be organized in modules, including ZnF recognized domains pointing to their involvement in diverse viral functions.[46–56] But interestingly in P1, the two ZnFs belong to very different classes of zinc fingers, of which the RYMV ZnF2 seems to constitute a new class. To our knowledge, these two atypical features were never reported before, as proteins usually harbor zinc fingers of close if not identical classes. Altogether, P1 structure appears to be highly original among ZnF proteins.

The differential behavior of P1 ZnFs is in accordance with their integration in P1 overall structure: ZnF1 functions as a closing cap of a compact beta-sheet sandwich, with limited access to the zinc atom and robustness, whereas ZnF2 bridges together structural elements otherwise minimally interacting. The loss of ZnF2 structure could serve as a new regulation level for the mutual orientation of the interaction motifs of ZnF1 and ZnF2 identified in this study, the solvent-exposed helix and C-terminal aromatic/basic stretch, respectively, whose destruction leads to a complete loss of virion infectivity.

Our mutational strategy supports these considerations and brings structural answers to the biological incidence of previously reported[23] and new punctual mutations (this study) in P1. Thus, Cys64 and Cys95 whose mutations were reported to lead to the loss of P1 VSR capacities,[23] and Trp36 whose mutation abolishes almost completely the RYMV replication are now identified as key residues for ZnF1 maintenance. Mutations of Trp29 and Trp31 completely suppress the systemic spreading of the corresponding mutant viruses, thus highlighting the essential role of Trp29 and Trp31 in viral long-distance movement. These residues are shown here to be involved in the dimerization interface of P1. Altogether these data provide a biological role of the P1 dimerization interface and reveal an additional level of regulation of P1 through its oligomerization properties mediated by its N-terminal domain. Like P1, other known VSR structures were obtained in dimeric conformation such as that of p19, NS10, B2 and 2b[12,15,18] or in an octameric form for p21.[18] The multimeric state of those proteins was always associated with double strand RNA binding activity required for VSR functions. However, P1 seems to behave differently compared to such VSRs with known structures, as its dimerization was observed without the need of RNA binding activities. Of note, all our attempts to assess the protein's ability to bind to small RNAs have failed to reach a clear conclusion.

Our mutational approach also reveals the importance of ZnF2 for early steps of viral infection, and its particular sensitivity to chelation and oxidation, which governs its unfolding and flexibility. We postulate that ZnF2 may function as a redox sensor of the plant redox status. This hypothesis is supported by the incidence of Cys-to-His mutations within P1 ZnF2 that preserve zinc binding to some extent, but decrease virus infectivity. We also demonstrate that P1 C-terminal tail downstream ZnF2, which is strictly conserved in the P1 diversity, is crucial for the virus spread. Its accessible position in the asymmetric dimer model opens future perspective towards the identification of potential ligands.

Altogether our results lead to a better understanding of P1, a key protein of RYMV infectivity, characterized by a unique structural profile. They notably point out to the multiple levels of structural regulation mediated by domain motion, oxidation and oligomerization and their impact on its biological function. They also open huge perspectives towards the understanding on how P1 structural properties sustain P1 sequence diversity and residues under positive selection[44] among the numerous RYMV isolates identified so far in the African continent and neighboring islands,[57] and much beyond among the *sobemovirus* gender.

Materials and Methods

Constructs

All constructs expressing P1 and primers used in this study are listed in the Table S1. DNA was amplified using the Phusion^{Tm} High-Fidelity enzyme (Finnzymes) and restriction site-containing primers, and subsequently cloned into destination vectors after digestion by corresponding restriction enzymes (New England Biolabs) and ligation with T4 DNA ligase (Promega). Mutations required for domain excision and/or punctual changes within P1 nucleotidic sequence were introduced in the P1 open reading frame using the QuikChange® II XL Site-Directed mutagenesis kit (Agilent-Stratagene) and related primers (Table S1). All constructs and mutations created for this study were

systematically checked by sequencing on both strands.

Plant materials and culture

The rice *Oryza sativa*, cultivar Indica, Var. IR64 was selected for its high sensitivity to RYMV.[58] Rice plants for infection assays were grown in confined greenhouse under 12 h light/12 h dark at 28 °C and at a relative humidity of 75%, and watered daily. Protoplasts were prepared from 15-day old, non-infected etiolated rice leaves and transfected according to Zhang and co-authors.[59]

RYMV isolates, infection and detection in rice

Rice plants (series of 30 two-week old rice plants at three-leaf stage) were infected either using a RYMV inoculum or with synthetic viral RNAs, both originated from the *Ivorian Cost* RYMV isolate.[43] Both inoculated and systemic leaves were collected at different times post-infection (from 7 to 28 dpi) and stored in liquid N2 until use. RYMV inoculum was obtained from infected leaves and mechanically inoculated to rice plants using carborundum (Fisher Scientific) as described earlier.[60] Full length synthetic RYMV RNAs were obtained with the T7 RNA polymerase (Promega) using as DNA template the linearized pUC19-based infectious clone FL5[43] carrying the entire RYMV *Cia* genome under the control of the bacteriophage T7 promoter. FL5 was also used as DNA template for mutagenesis to produce synthetic RYMV RNAs carrying mutations in the P1 open reading frame (Table S1). Rice plant infections by synthetic viral RNAs were performed according to,[43] using 20 two-week-old rice plants (three-leaf stage) cultivated under growth conditions described above. Both inoculated and systemic leaves were collected at different times post-infection (from 7 to 35 dpi) and stored in liquid N2 until use.

Total RNA required for distinct analyses was extracted from 100 mg of infected pieces using the RNeasy Plant Mini kit (Qiagen) and reverse transcribed into cDNAs using the RevertAid First Strand cDNA Synthesis Kit (Fisher Scientific). To check the occurrence of synthetic viral RNA in inoculated leaves, cDNAs were prepared from 1-week infected plantlets and subjected to PCR diagnostics using the GoTaq polymerase (Promega) and RYMV_19F ($^{5'}$CTCTACGAC TATGCTGACACC$^{3'}$) and RYMV_ 20R ($^{5'}$CTCCC CCACCCATCCCGAGA$^{3'}$) primers (Table S1). To check the non-reversion of mutations designed in the P1 coding region within FL5-derived RNA during infection, cDNAs from 3 weeks infected plants were prepared and the complete P1 ORF was amplified with the Phusion High-Fidelity enzyme (Fisher Scientific) using RYMV_1F ($^{5'}$ACAATTGAAGCTAGGAAAGGAGC$^{3'}$) and RYMV_4R ($^{5'}$GGTCGCTTTCTCACTCGCACC$^{3'}$) primers (Table S1), and systematically checked by sequencing.

Detection of viral particles in infected plants was done by Das-ELISA according to[60] with slight modifications. Leaf samples (1 g) were ground with 10 ml of PBS containing 0.05% Tween 20 (PBS-T) and 2% polyvinylpyrolidone (PVP). ELISA plates (Nunc-Immuno™ MicroWell™ 96-well solid plates, Thermo Scientific) were coated with purified uncoupled rabbit IgGs directed against $RYMV_{MG}$ particles (an RYMV isolate from Madagascar) diluted at 1 μg. mL^{-1} in a carbonate coating buffer (15 mM Na_2CO_3m, 34 mM $NaHCO_3$, pH 9,6, 100 μL/well) for 2 h at 37 °C. After washing for 10 min with a TBS-T buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0,05% Tween20, v/v), non-occupied sites in wells were blocked for 1 h at 37 °C using TBS-T buffer containing 5% milk powder (200 μL/well). After a second plate wash, 100 μl of leaf extracts diluted to 1/100 were delivered to individual wells and incubated for 16 h at 4 °C. After washing, secondary antibodies against $RYMV_{MG}$ conjugated to alkaline phophatase and diluted in TBS-T at a 1/1000 ratio were added to wells (100 μL/well) and incubated 2 h at 37 °C. The disodium 4-nitrophenyl phosphate hexahydrate (PNPP) substrate for alkaline phosphatase was diluted at 1 $mg.mL^{-1}$ in diethanolamine buffer (1 M diethanolamine pH 9,8, 0,5 mM $MgCl_2$) and added to wells (100 μL/well) after ultimate plate washing. Chromogenic reactions developed during 1 h at dark, and optical densities (O.D) were measured at 405 nm in a V-1200 UVisco spectrophotometer. Data normalization was performed using data from wells treated with particle-free buffer as the reference values.

Procedures for immunodetection analyses by western blots are described in the SDS-PAGE and Immunoblotting section below.

Production and purification of recombinant proteins

For the production of P1 recombinant proteins, a full-length P1 nucleotide sequence originated from the RYMV Tz3 isolate was already cloned in a pET3b vector (Novagen) between *NdeI* and *BamHI* restriction sites to produce the P1 referent protein.[28] Additional truncated P1 proteins (P1_1-100, P1_50-100, P1_50-157 and P1_102-157, amino acid numbering) were generated similarly in the pET3b vector, based on the strict conservation of cysteine/histidine residues, P1 similarity of its CXXC microdomains with that of few zinc binding proteins,[28] and alpha helices and beta strands predictions using the *PredictProtein* server.[30] These four constructs were used to generate pET3b collections of mutant P1 coding sequences by site-directed mutagenesis, including the pET3b.$P1^{\Delta 153-157}$ construct carrying an early stop codon in place of Lys153 (Table S1). When required, a methionine residue was artificially introduced during the sub-

cloning of DNA fragments to guarantee protein translation, modifying their molecular mass consequently (Figures S4 and S5).

All recombinant proteins were produced in the *E. coli* strain BL21-DE3 cells (Novagen) by induction with 0.4 mM IPTG at an OD_{600} of 0.6 (Figure S1) and purified according to[28] with modifications depend on P1 mutants (Table S2). All proteins appeared properly folded after purification (Figure S3). Protein concentrations were determined by measuring absorbance at 280 nm using the calculated value of extinction coefficient for each different P1 variant with a Nanodrop spectrophotometer (ThermoScientific).

Far-UV circular dichroism spectroscopy

CD spectroscopy analyses of reduced or oxidized P1 (12 μM) and its sub regions (25 μM) were as described for the wild type P1 protein.[28] For all experiments, molar ellipticity was calculated as described[61] on an average of three independent measurements.

SDS-PAGE analyses and immunoblotting

For bacterial fraction analyses, cells were collected by centrifugation in a buffer containing 50 mM Tris-HCl pH 8, 2 mM DTT, then sonicated and separated into insoluble and soluble fractions. Extraction of insoluble protein fraction were performed using Tris-HCl 50 mM pH 8 urea 6 M under shaking conditions until insoluble pellets are fully resuspended (~30 min), then bacterial debris remaining insoluble are eliminated by centrifugation. Renaturation of insoluble proteins were performed by dialysis overnight at 4 °C under gentle shaking in Tris-HCl 50 mM pH8. Protein samples were treated with Laemmli buffer (40 mM Tris-HCl, pH 6.8, 1% SDS, 50 mM DTT, 7.5% glycerol, and 0.003% bromophenol blue) and analyzed by SDS-PAGE in denaturing conditions using 15% to 18% acrylamide gels (37.5:1 ratio acrylamide:bis-acrylamide, 40%) depending on P1 mutants. Purification of P1 mutants from bacterial extracts was performed according to.[28] Purified P1 protein (25 μM) and its derived truncations (50 μM) were analyzed similarly or in non-reducing conditions without DTT according to.[28] Redox treatments by different concentrations of H_20_2 or DTT were applied to purified proteins prior to SDS-PAGE analyses either at 20 °C for 45 min and in different buffers depending on the experiment (see figure legends). On-gel zinc detection in proteins after SDS-PAGE was performed prior to Coomassie blue staining using the zinc-complexing probe PAR (4-(2-Pyridylazo) resorcinol, Fluka).[62]

For rice protein analyses by western blots, leaf proteins were extracted as for DAS-ELISA assays described above. SDS-PAGE gels for immunodetection were loaded with 15 μg of rice total proteins per lane. Immunoblotting was performed after protein transfer onto Hybond-P membranes (GE Healthcare, RPN303F) and after blocking membrane nonspecific binding sites in TBS (20 mM Tris, 75 mM NaCl, and 2.5 mM $MgCl_2$, pH 7.6, 5% milk powder). PI protein was subjected to primary antibody recognition using an anti-P1 antibody (diluted 1:1000 fold) raised in rabbit[23] and purified by Affi-Gel 15 immnunoaffinity chromatography (BIO-RAD) using the recombinant P1 protein as the ligand. For CP protein, the anti-CP described in[28] was used at a dilution of 1/2000. A peroxidase-coupled secondary antibody (anti-rabbit IgG; Immunopure Pierce) was used at a dilution of 1:20000 for primary antibody recognition. Membranes were subjected to chemiluminescence detection by the ECL Plus western blotting detection system (GE Healthcare) after protein transfer.

Determination of P1_1-100 zinc binding constant

The affinity of different regions of P1 for zinc ions was determined according to[63] with slight modifications. Purified P1_1-100 (40 μM) was first reconstituted after treatment by 2 mM DTT and 300 μM ZnSO4, then DTT and unbound zinc were removed using PD10 gel filtration columns (GE Heathlcare) equilibrated with a 40 mM HEPES-KOH, pH 7.5 buffer. Equilibrium between P1_1-100 and zinc was further determined by treating P1_1-100 overnight with increasing concentrations (37 μM to 5 mM) of the Zn chelator TPEN (N,N,N′,N′-tetrakis(2-pyridyl-methyl)ethylenediamine) at room temperature in a 50 mM HEPES-KOH buffer (pH7.5) supplemented with 1 mM DTT. TPEN complexed with Zn, free TPEN and free Zn were further removed by gel filtration on PD MidiTrap G-10 (GE Healthcare) equilibrated with 40 mM of the metal-free HEPES-KOH, pH 7.5 buffer. To determine the amount of Zn still bounded to proteins, we performed a PAR/PMPS assay as follows: 80 μl of PAR probe were added to 720 μl of protein to a final concentration of 100 μM. Then we used 1 mM PMPS (p-hydroxymercuriphenylsulfonic acid) which forms mercaptide bond with Cys residues and thus induces the release of the remaining Zn still bound to protein. Zn released from P1 directly reacts with PAR probe to form $Zn(PAR)_2$ complexes colored in far red. We used spectrophotometry absorbance analyses at A_{500nm} and the GraphPad Prism^R software for non-linear regression and graphic representation to determine the amount of Zn released by PMPS addition. Values from two independent experiments were used for non-linear regression. We used the following equation for Zn affinity determination according to:.[63]

$$Ka_{P1\_1-100} = \frac{(KaTPEN\ x\ [P1\_1-100\ Zn]\ x\ [Zn-free\ TPEN])}{[TPEN\ Zn]\ x\ [Zn-free\ P1\_1-100]}$$

Where $Ka_{TPEN} = 10^{16}\ M^{-1}$[31] the TPEN concentrations necessary to release 50% of Zn atoms in

P1_1-100 is determined graphically ([$TPEN_{Zn50\%}$] = 63 × 10^{-6} M). [P1_1-100 Zn] and [Zn-free P1_1-100] correspond to the concentrations of proteins with or without Zn respectively. [P1_1-100 Zn] = [Zn-free P1_1-100] = 20 × 10^{-6} M. [Zn-free TPEN] and [TPEN Zn] correspond to the concentrations of free TPEN and TPEN in complex with Zn atom respectively. [Zn-free TPEN] = [$TPEN_{Zn50\%}$] - [Zn-free P1_1-100].

Native mass spectrometry

Purified proteins were buffer exchanged against 50 mM ammonium acetate, pH 8.0, using microcentrifuge gel-filtration columns (Zeba 0.5 mL, ThermoScientific, France) prior to MS analyses. Native MS experiments were performed on a high-resolution electrospray quadrupole time-of-flight mass spectrometer (Synapt G2, Waters, Manchester, UK) equipped with an automated chip-based nanoelectrospray device (TriversaNanomate, Advion Biosciences, Ithaca, NY). Calibration was performed with the multiply charged ions produced by 2 μM horse heart myoglobin diluted in 1:1 (v:v) acetonitrile:water with 1% formic acid. Molecular mass, integrity, and homogeneity of P1 were first checked in denaturing conditions by diluting P1 down to 2 μM in 1:1 (v:v) acetonitrile:water acidified with 1% formic acid.

Analyses in native conditions were carried out using 5 μM to 10 μM of P1 and mutants in 50 mM ammonium acetate, pH 8.0. Optimal acceleration voltages applied on the sample cone (Vc) and a well-adapted pressure in the interface (Pi) was used to preserve weak noncovalent interactions in the gas phase. MS data were acquired and processed using MassLynx™ 4.1 (Waters). Peak intensities of the main charge states of the different species were used to calculate ratios of each detected species.

For EDTA treatment, P1 (5 μM) and sub-regions (10 μM) were incubated overnight at 4 °C with 40 to 400 equivalent EDTA and 1 mM DTT. For dynamic monitoring of zinc reversibility, EDTA-treated proteins were incubated with 60 to 100 μM zinc acetate.

X-ray crystallography

The purified P1_1-100 and P1_102-157 proteins were concentrated to 10 mg/ml mL and 50 mg/mL respectively in 20 mM Tris-HCl pH 8, 75 mM NaCl, 1 mM DTT and 300 μM $ZnCl_2$. Crystallization trials were performed at 20 °C using the hanging-drop vapor-diffusion method in 96 microplates and nano X8 (Cartesian) robot with 100 nl of protein mixed with 100 nl of reservoir. After 3 weeks, small crystals of P1_1-100 were obtained from condition 35 of the Classic Suite II kit (Qiagen) containing 0.1 M HEPES pH 7, 1 M Ammonium sulfate, and 0.5 %(w/v) PEG 8000. After further optimization, diffracting crystals were obtained from 0.1 M MES pH 6, 0.5 M Ammonium sulfate and 0.5 % PEG 8000 using the hanging-drop vapor-diffusion method. Crystals were cryo-protected by flash soaking in the same solution containing 20 % of 1–4 Butandiol (Sigma) before flash freezing in liquid N_2. After 5 weeks, diffracting crystals of P1_102-157 were obtained at 4 °C from condition 7 of the Classic kit (Qiagen) containing 0.1 M tri-Sodium citrate pH 5,6, 20 %(v/v) Isopropanol and 20 %(w/v) PEG 4000. Crystals were cryo-protected by flash freezing in liquid N_2.

X-Ray data collection, processing and structure determination

Crystal diffraction datasets were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble) at beam lines ID23-1 and ID29 using a Pixel detector (Pilatus 6 M) and processed by XDS[64] and Scala[65] from the CCP4 programs suite.[66]

P1_1-100 crystals belong to the $F4_132$ space group and contain two molecules in the asymmetric unit. Their structure was determined to 2.10 Å by the single wavelength anomalous diffraction (SAD) method using HKL2MAP[67] and ARP/WARP from the CCP4 programs suite.[68–69]

P1_102-157 crystals belong to the $P12_11$ space group and contain two molecules in the asymmetric unit. P1_102-157 structure was determined to 2.10 Å by the single wavelength anomalous diffraction (SAD) method using AUTOSOL from PHENIX[70] and ARP/WARP[68–69] and P1_102-157 phasing data set. A higher resolution structure was determined using a higher resolution dataset (P1_102-157 at 1,98 A) and molecular replacement with PHASER.[71]

The ClusPro-DC server (https://cluspro.org) was used for protein–protein docking.[42] PISA (https://www.ebi.ac.uk/pdbe/pisa/) was used as an interactive tool for the exploration of macromolecular interfaces.[41] Area surfaces were determined at http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html.

After model building using Coot[72] and refinement by REFINE from PHENIX, final structures exhibited an R(%) / $R(\%)_{free}$ of 0,185 / 0,214 for P1_1-100 and 0,193 / 0,223 for P1_102-157. Final refinement statistics for the structures are listed in Table 1. Figures were generated using PyMol (The PyMOL Molecular Graphics System, Version 2.2.3 Schrödinger, LLC).

The coordinates and structure factors for N-terminal P1_1-100, C-terminal P1_102-157 and P1_1-157 proteins have been deposited in the Protein Data Bank (PDB) under ID codes 6TY0, 6TY2 and 6XV2, respectively.

Nuclear magnetic resonance analyses and structure calculations

Production of proteins for NMR studies followed the protocol detailed in[73] and in the Table S2. Briefly, ^{15}N-edited HSQC-TOCSY and -NOESY experiments were used to assign resonances of the short constructs. The assignment of full P1_1-157 was deduced from these experiments, except the linker region (residue 98–103). ^{15}N/^{13}C/^{1}H triple resonance experiments on the full-length protein confirmed NMR resonance assignment for the N- and C-terminal domains and enabled the identification of chemical shifts of the linker region. Chemical shifts have been deposited in the BioMagResBank (https://www.brmb.wisc.edu) with accession number 27880.

Residual dipolar couplings were measured as differences of $^{1}J_{HN}$ couplings between an aligned sample with a 14 mg/mL Pf1 phage suspension (Asla^{Biotech}) and an isotropic sample of ^{15}N labelled P1_1-157. $^{1}J_{HN}$ couplings were measured as doubles of chemical shift differences in the ^{15}N dimension of the ^{15}N-^{1}H correlation peaks of TROSY-HSQC spectra pair[74] recorded each with 128 complex points inF1, and subsequently zero-filled to 256 complex points using linear forwards prediction. Extraction of the alignment tensor parameters was performed with the Module software.[75]

^{15}N relaxation datasets were obtained according to published pulse sequences incorporating a heating compensation scheme between T1 and T2 experiments.[76] Experiments were recorded at 20 °C, with a recycle delay of 3 s. Delay times were 1 ms, 100 ms, 600 ms, 900 ms, 1200 ms and 1600 ms for the T1 experiments and 15.8, 31.6, 47.5 and 63.4 ms for the T2 experiments. Compensatory refocusing periods were introduced in the T1 and T2 experiments so that the spins experience an identical heating in all relaxation experiments.[77] The protein aggregated slowly over time, resulting in a slightly accelerated decay of magnetization with a scaling factor of 0.94–0.96 between two successive T2 or T1 experiments as shown in the interspersed ^{1}D spectra. Estimation of the relaxation times were done by fitting the exponential decay of the peak volumes by three-parameters exponential decay after a rescaling of the corresponding intensities. Estimation of the diffusion tensor was done with the Tensor2 software from T1/T2 ratios.[39]

Derivation of the mean structure of the whole protein was done by conjoined rigid-body/torsion angle/Cartesian simulated annealing driven by NMR restraints, according to the protocol developed by Clore and co-workers.[78] The calculations were carried out in Xplor-NIH.[79–80] The residues of the [1–96] fragment were held fixed in space, whereas the 105–157 fragment was treated as a rigid body. After a randomization step of backbone dihedral angles, atoms of the linker region connecting both zinc fingers (residues 98–104) were given first Cartesian and then torsional degrees of freedom, under residual couplings, restraints derived from ^{1}H–^{1}H Noes, and ϕ/ψ restraints derived by TALOS[38] from HN, N, Cα, Cβ and C chemical shifts, in addition to covalent and geometrical energetic terms. The coordinates and experimental data have been deposited in the PDB, accession number 6XV2.

Binary yeast-two hybrid assays

The yeast strain YRG2 (Stratagene, Agilent) was used to evaluate P1 domain dimerization properties. P1 regions of interest (P1_1-100 and P1_102-157) were excised by *NdeI*/*BamHI* restriction from the pET3b collection (Table S1) and ligated at compatible sites in the pGAD.T7 and pGBK.T7 yeast two-hybrid vectors (Clontech). pGAD and pGBK constructs were co-introduced in yeast cells and assayed for binary yeast-two hybrid interactions as described previously.[81]

Transient expression assays and confocal microscopy

The pEZS-NL and pEZS-CL vectors (https://deepgreen.stanford.edu/) carrying the CaMV 35S-MCS-(Ala)$_{10}$-EGFP (pEZS-NL vector) or the CaMV 35S-EGFP-(Ala)$_{10}$-MCS cassettes (pEZS-CL vector), respectively, were selected to express P1-EGFP and EGFP-P1 fusion proteins and related variants in rice cells. P1-encoding DNA sequences of interest were amplified from the pET3b collection and sub-cloned between *SalI* and *BamHI* restriction sites to give rise to pEZS. P1-EGFP and pEZS.EGFP-P1 constructs (Table S1). pEZS.P1-EGFP and pEZS.EGFP-P1 constructs were subsequently used as templates to generate two collections of mutant P1 in frame with the EGFP protein by site-directed mutagenesis (Table S1).

Localization of EGFP-P1 and P1-EGFP fusion proteins and related variants in rice protoplasts was performed using a maximum of 10 μg total DNA per transfection event. EGFP fluorescence[82] was recorded 24 hours post transfection after excitation with an argon laser at 488 nm (fluorescence detection range: 500–550 nm) with a SP8 laser-scanning confocal microscope (Leica Microsystems), equipped with 40x water immersion objective. The Leica LAS X software was used to obtain images at a high resolution and without maximum Z-stack intensity projection (1024 pixels for both X and Y dimensions with a physical length of 83 μm, Z dimension being variable in stacks depend on cell size). All data were treated with Adobe Photoshop CS3 software at high resolution. The results are representative of three biological replicates and

of 10–20 cells analyzed per transfection event showing the same fluorescence pattern.

CRediT authorship contribution statement

Vianney Poignavent: Conceptualization, Methodology, Investigation, Visualization, Writing original draft. **François Hoh:** Methodology, Investigation. **Guillaume Terral:** Investigation. **Yinshan Yang:** Methodology, Investigation. **François-Xavier Gillet:** Conceptualization, Investigation. **Jeong-Hyeon Kim:** Methodology, Validation. **Frédéric Allemand:** Investigation. **Eric Lacombe:** Investigation, Funding acquisition. **Christophe Brugidou:** Resources. **Sarah Cianferani:** Conceptualization, Methodology, Resources, Supervision. **Hélène Déméné:** Conceptualization, Methodology, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Florence Vignols:** Conceptualization, Methodology, Investigation, Validation, Resources, Writing – review & editing, Visualization, Supervision, Project administration, 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.

Acknowledgements

We are grateful to Pr Grishin (University of Texas, USA), Srikrishna Subramania (CSIR-Institute of Microbial Technology, India), Sam Ireland and Andrew Martin (Institute of Structural and Molecular Biology, UK) and Yue-Ie Hsing and Lin-Yun Kuang (IPMB Institute, Academia Sinica, Taipei, Taiwan) for their expertise. We thank the ID23-1 and ID29 beams lines scientists at European Synchrotron Radiation Facility (Grenoble) for their excellent support during data collections. The CBS is a GIS-IBIsA platform and a member of the French Infrastructure for Integrated Structural Biology (FRISBI) funded by the French National Research Agency (ANR-10-INBS-05 grant). V. P. and F-X. G. PhD interships were funded by the French Ministry of Education and Scientific Research. V. P. and F. V. were also funded by the CampusFrance Orchid program (Grant n°29354SD), the Taiwanese Research and Practical Training Program, and by the GAIA PhD School of Montpellier University, France.

Appendix A. Supplementary Material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2022.167715.

Received 1 May 2022;
Accepted 28 June 2022;
Available online 4 July 2022

Keywords:
3D structure;
Oryza sativa;
Sobemovirus;
viral movement protein;
zinc finger protein

† These authors contributed equally to the work.

Abbreviations:
RYMV, Rice yellow mottle virus; MP, Movement protein; RNAi, RNA interference; ZnF, Zinc finger

References

1. Belshaw, R., Pybus, O.G., Rambaut, A., (2007). The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res* **17**, 1496–1504.
2. Holmes, E.C., (2003). Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol* **11**, 543–546.
3. Ben-Shaul, A., Gelbart, W.M., (2015). Viral ssRNAs Are Indeed Compact. *Biophys J* **108**, 14–16.
4. Carrington, J.C., Kasschau, K.D., Mahajan, S.K., Schaad, M.C., (1996). Cell-to-Cell and Long-Distance Transport of Viruses in Plants. *Plant Cell* **8**, 1669–1681.
5. Heinlein, M., (2015). Plant virus replication and movement. *Virology* **479–480**, 657–671.
6. Waigmann, E., Ueki, S., Trutnyeva, K., Citovsky, V., (2004). The Ins and Outs of Nondestructive Cell-to-Cell and Systemic Movement of Plant Viruses. *CRC Crit Rev Plant Sci* **23**, 195–250.
7. Harries, P., Ding, B., (2011). Cellular factors in plant virus movement: At the leading edge of macromolecular trafficking in plants. *Virology* **411**, 237–243.
8. Taliansky, M., Torrance, L., Kalinina, N.O., (2008). Role of Plant Virus Movement Proteins. *Meth Mol Biol*, 33–54.
9. Verchot-Lubicz, J., Torrance, L., Solovyev, A.G., Morozov, S.Y., Jackson, A.O., Gilmer, D., (2010). Varied Movement Strategies Employed by Triple Gene Block-Encoding Viruses. *Mol Plant-Microbe Interact* **23**, 1231–1247.
10. Pumplin, N., Voinnet, O., (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nature Rev Microbiol* **11**, 745–760.
11. Csorba, T., Kontra, L., Burgyán, J., (2015). Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology* **479–480**, 85–103.
12. Baulcombe, D.C., Molnár, A., (2004). Crystal structure of p19 – a universal suppressor of RNA silencing. *Trends Biochem Sci* **29**, 279–281.
13. Chao, J.A., Lee, J.H., Chapados, B.R., Debler, E.W., Schneemann, A., Williamson, J.R., (2005). Dual modes of

RNA-silencing suppression by Flock House virus protein B2. *Nature Struct Mol Biol* **12**, 952–957.
14. Chen, H.-Y., Yang, J., Lin, C., Yuan, Y.A., (2008). Structural basis for RNA-silencing suppression by Tomato aspermy virus protein 2b. *EMBO Rep* **9**, 754–760.
15. Vargason, J.M., Szittya, G., Burgyán, J., Hall, T.M.T., (2003). Size selective recognition of siRNA by an RNA silencing suppressor. *Cell* **115**, 799–811.
16. Yang, X., Tan, S.H., Teh, Y.J., Yuan, Y.A., (2011). Structural implications into dsRNA binding and RNA silencing suppression by NS3 protein of Rice Hoja Blanca Tenuivirus. *RNA* **17**, 903–911.
17. Ye, K., Malinina, L., Patel, D.J., (2003). Recognition of small interfering RNA by a viral suppressor of RNA silencing. *Nature* **426**, 874–878.
18. Ye, K., Patel, D.J., (2005). RNA Silencing Suppressor p21 of Beet Yellows Virus Forms an RNA Binding Octameric Ring Structure. *Structure* **13**, 1375–1384.
19. Bakker, W., (1970). Rice yellow mottle, a mechanically transmissible virus disease of rice in Kenya. *Netherlands J Plant Pathol* **76**, 53–63.
20. Fargette, D., Pinel-Galzi, A., Sérémé, D., Lacombe, S., Hébrard, E., Traoré, O., Konaté, G., (2008). Diversification of Rice Yellow Mottle Virus and Related Viruses Spans the History of Agriculture from the Neolithic to the Present. *PLoS Pathog* **4**, e1000125
21. Séré, Y., Fargette, D., Abood, M.E., Wydra, K., Onasnya, A., Akator, S.K., (2013). Managing the major diseases of rice in Africa. In: CabiCabi (Ed.), *Realizing Africa's Rice Promise*. Wallingford, pp. 213–228.
22. Bonneau, C., Brugidou, C., Chen, L., Beachy, R.N., Fauquet, C., (1998). Expression of the rice yellow mottle virus P1 protein in vitro and in vivo and its involvement in virus spread. *Virology* **244**, 79–86.
23. Siré, C., Bangratz-Reyser, M., Fargette, D., Brugidou, C., (2008). Genetic diversity and silencing suppression effects of Rice yellow mottle virus and the P1 protein. *Virol J* **5**, 55.
24. Fusaro, A.F., Correa, R.L., Nakasugi, K., Jackson, C., Kawchuk, L., Vaslin, M.F.S., Waterhouse, P.M., (2012). The Enamovirus P0 protein is a silencing suppressor which inhibits local and systemic RNA silencing through AGO1 degradation. *Virology* **426**, 178–187.
25. Lacombe, S., Bangratz, M., Vignols, F., Brugidou, C., (2010). The rice yellow mottle virus P1 protein exhibits dual functions to suppress and activate gene silencing. *Plant J* **61**, 371–382.
26. Sarmiento, C., Gomez, E., Meier, M., Kavanagh, T.A., Truve, E., (2007). Cocksfoot mottle virus P1 suppresses RNA silencing in Nicotiana benthamiana and Nicotiana tabacum. *Virus Res* **123**, 95–99.
27. Voinnet, O., Pinto, Y.M., Baulcombe, D.C., (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc Natl Acad Sci U S A* **96**, 14147–14152.
28. Gillet, F.-X., Cattoni, D.I., Petiot-Bécard, S., Delalande, F., Poignavent, V., Brizard, J.-P., Bessin, Y., Van Dorsselaer, A., et al., (2013). The RYMV-Encoded Viral Suppressor of RNA Silencing P1 Is a Zinc-Binding Protein with Redox-Dependent Flexibility. *J Mol Biol* **425**, 2423–2435.
29. Sõmera, M., Sarmiento, C., Truve, E., (2015). Overview on Sobemoviruses and a Proposal for the Creation of the Family Sobemoviridae. *Viruses* **7**, 3076–3115.
30. Rost, B., Yachdav, G., Liu, J., (2004). The PredictProtein server. *Nucleic Acids Res* **32**, W321–W326.
31. Anderegg, G., Hubmann, E., Podder, N.G., Wenk, F., (1977). Pyridinderivate als Komplexbildner. XI. Die Thermodynamik der Metallkomplexbildung mit Bis-, Tris- und Tetrakis[(2-pyridyl)methyl]-aminen. *Helv Chim Acta* **60**, 123–140.
32. Smith, R.M., Martell, A.E., (1989). Iminodiacetic Acid DerivativesCritical Stability Constants. Springer, US, Boston, MA, pp. 67–127.
33. Holm, L., Rosenström, P., (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
34. Krishna, S.S., Majumdar, I., Grishin, N.V., (2003). Structural classification of zinc fingers. *Nucleic Acids Res* **31**, 532–550.
35. Kaur, G., Subramanian, S., (2016). Classification of the treble clef zinc finger: noteworthy lessons for structure and function evolution. *Sci Rep* **6**, 32070.
36. Ireland, S.M., Martin, A.C.R., (2019). ZincBind - The database of zinc binding sites. *Database (Oxford)* **2019**, 1–8.
37. Tolman, J.R., Al-Hashimi, H.M., Kay, L.E., Prestegard, J. H., (2001). Structural and Dynamic Analysis of Residual Dipolar Coupling Data for Proteins. *J Amer Chem Soc* **123**, 1416–1424.
38. 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.
39. Dosset, P., Hus, J.C., Blackledge, M., Marion, D., (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR* **16**, 23–28.
40. De La Torre, J.G., Huertas, M.L., Carrasco, B., (2000). HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *J Magn Reson* **147**, 138–146.
41. Krissinel, E., Henrick, K., (2007). Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* **372**, 774–797.
42. Yueh, C., Hall, D.R., Xia, B., Padhorny, D., Kozakov, D., Vajda, S., (2017). ClusPro-DC: Dimer Classification by the Cluspro Server for Protein-Protein Docking. *J Mol Biol* **429**, 372–381.
43. Brugidou, C., Holt, C., Yassi, M.N., Zhang, S., Beachy, R., Fauquet, C., (1995). Synthesis of an infectious full-length cDNA clone of rice yellow mottle virus and mutagenesis of the coat protein. *Virology* **206**, 108–115.
44. Sérémé, D., Lacombe, S., Konate, M., Bangratz, M., Pinel-Galzi, A., Fargette, D., Traore, A.S., Konate, G., et al., (2014). Sites under positive selection modulate the RNA silencing suppressor activity of rice yellow mottle virus movement protein P1. *J Gen Virol* **95**, 213–218.
45. Negi, S., Itazu, M., Imanishi, M., Nomura, A., Sugiura, Y., (2004). Creation and characteristics of unnatural CysHis (3)-type zinc finger protein. *Biochem Biophys Res Commun* **325**, 421–425.
46. Chiba, S., Hleibieh, K., Delbianco, A., Klein, E., Ratti, C., Ziegler-Graff, V., Bouzoubaa, S., Gilmer, D., (2013). The benyvirus RNA silencing suppressor is essential for long-distance movement, requires both zinc-finger and NoLS basic residues but not a nucleolar localization for its silencing-suppression activity. *Mol Plant-Microbe Interact* **26**, 168–181.
47. Dong, X., van Wezel, R., Stanley, J., Hong, Y., (2003). Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. *J Virol* **77**, 7026–7033.

48. Fujita, N., Komatsu, K., Ayukawa, Y., Matsuo, Y., Hashimoto, M., Netsu, O., Teraoka, T., Yamaji, Y., Namba, S., Arie, T., (2018). N-terminal region of cysteine-rich protein (CRP) in carlaviruses is involved in the determination of symptom types. *Mol Plant Pathol* **19**, 180–190.
49. Kenesi, E., Carbonell, A., Lózsa, R., Vértessy, B., Lakatos, L., (2017). A viral suppressor of RNA silencing inhibits ARGONAUTE 1 function by precluding target RNA binding to pre-assembled RISC. *Nucleic Acids Res* **45**, 7736–7750.
50. López, C., Navas-Castillo, J., Gowda, S., Moreno, P., Flores, R., (2000). The 23-kDa Protein Coded by the 3′-Terminal Gene of Citrus Tristeza Virus Is an RNA-Binding Protein. *Virology* **269**, 462–470.
51. Lukhovitskaya, N.I., Thaduri, S., Garushyants, S.K., Torrance, L., Savenkov, E.I., (2013). Deciphering the Mechanism of Defective Interfering RNA (DI RNA) Biogenesis Reveals That a Viral Protein and the DI RNA Act Antagonistically in Virus Infection. *J Virol* **87**, 6091–6103.
52. Trinks, D., Rajeswaran, R., Shivaprasad, P.V., Akbergenov, R., Oakeley, E.J., Veluthambi, K., Hohn, T., Pooggin, M.M., (2005). Suppression of RNA Silencing by a Geminivirus Nuclear Protein, AC2, Correlates with Transactivation of Host Genes. *J Virol* **79**, 2517–2527.
53. Valli, A., Dujovny, G., García, J.A., (2008). Protease activity, self interaction, and small interfering RNA binding of the silencing suppressor p1b from cucumber vein yellowing ipomovirus. *J Virol* **82**, 974–986.
54. van Wezel, R., Liu, H., Wu, Z., Stanley, J., Hong, Y., (2003). Contribution of the Zinc Finger to Zinc and DNA Binding by a Suppressor of Posttranscriptional Gene Silencing. *J Virol* **77**, 696–700.
55. Yambao, M.L.M., Yagihashi, H., Sekiguchi, H., Sekiguchi, T., Sasaki, T., Sato, M., Atsumi, G., Tacahashi, Y., Nakahara, K.S., Uyeda, I., (2008). Point mutations in helper component protease of clover yellow vein virus are associated with the attenuation of RNA-silencing suppression activity and symptom expression in broad bean. *Arch Virol* **153**, 105–115.
56. Yang, X., Baliji, S., Buchmann, R.C., Wang, H., Lindbo, J. A., Sunter, G., Bisaro, D.M., (2007). Functional modulation of the geminivirus AL2 transcription factor and silencing suppressor by self-interaction. *J Virol* **81**, 11972–11981.
57. Rakotomalala, M., Vrancken, B., Pinel-Galzi, A., Ramavovololona, P., Hébrard, E., Randrianangaly, J.S., Dellicour, S., Lemey, P., Fargette, D., (2019). Comparing patterns and scales of plant virus phylogeography: Rice yellow mottle virus in Madagascar and in continental Africa. *Virus Evol* **5**, vez023.
58. Ndjiondjop, M.N., Albar, L., Fargette, D., Fauquet, C., Ghesquière, A., (1999). The genetic basis of high resistance to Rice Yellow Mottle Virus (RYMV) in cultivars of two cultivated rice species. *Plant Dis* **83**, 931–935.
59. Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., Wang, P., Li, Y., Liu, B., Feng, D., Wang, J., Wang, H., (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods* **7**, 30.
60. Konate, G., Traore, O., Coulibaly, M.M., (1997). Characterization of rice yellow mottle virus isolates in Sudano-Sahelian areas. *Arch Virol* **142**, 1117–1124.
61. Kelly, S.M., Jess, T.J., Price, N.C., (2005). How to study proteins by circular dichroism. *Biochim Biophys Acta - Proteins Proteomics* **1751**, 119–139.
62. Lee, J.-W., Helmann, J.D., (2006). Biochemical characterization of the structural $Zn2^{+}$ site in the Bacillus subtilis peroxide sensor PerR. *J Biol Chem* **281**, 23567–23578.
63. Jakob, U., Eser, M., Bardwell, J.C., (2000). Redox switch of hsp33 has a novel zinc-binding motif. *J Biol Chem* **275**, 38302–38310.
64. Kabsch, W., (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 133–144.
65. Evans, P., (2006). Scaling and assessment of data quality. *Acta Crystallogr Sect D Biol Crystallogr* **62**, 72–82.
66. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R. J., Vagin, A., Wilson, K.S., (2011). Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 235–242.
67. Sheldrick, G.M., (2010). Experimental phasing with *SHELXC* / D / E : combining chain tracing with density modification. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 479–485.
68. Langer, G., Cohen, S.X., Lamzin, V.S., Perrakis, A., (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protoc* **3**, 1171–1179.
69. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., (2011). *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 355–367.
70. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N. W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., (2010). *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 213–221.
71. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., (2007). *Phaser* crystallographic software. *J Appl Crystallogr* **40**, 658–674.
72. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., (2010). Features and development of *Coot*. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 486–501.
73. Yang, Y., Declerck, N., Déméné, H., (2020). NMR chemical shift assignment of a constitutively active fragment of the antitermination protein LicT. *Biomol NMR Assign* **14**, 19–23.
74. Kontaxis, G., Clore, G.M., Bax, A., (2000). Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J Magn Reson* **143**, 184–196.
75. Dosset, P., Hus, J.C., Marion, D., Blackledge, M., (2001). A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J Biomol NMR* **20**, 223–231.
76. Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., Kay, L.E., (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984–6003.

77. Mulder, F.A.A., van Tilborg, P.J.A., Kaptein, R., Boelens, R., (1999). Microsecond time scale dynamics in the RXR DNA-binding domain from a combination of spin-echo and off-resonance rotating frame relaxation measurements. *J Biomol NMR* **13**, 275–288.
78. Schwieters, C.D., Suh, J.-Y., Grishaev, A., Ghirlando, R., Takayama, Y., Clore, G.M., (2010). Solution Structure of the 128 kDa Enzyme I Dimer from *Escherichia coli* and Its 146 kDa Complex with HPr Using Residual Dipolar Couplings and Small- and Wide-Angle X-ray Scattering. *J Am Chem Soc* **132**, 13026–13045.
79. Schwieters, C.D., Kuszewski, J.J., Clore, G.M., (2006). Using Xplor–NIH for NMR molecular structure determination. *Prog Nucl Magn Reson Spectrosc* **48**, 47–62.
80. Schwieters, C.D., Clore, G.M., (2001). Internal Coordinates for Molecular Dynamics and Minimization in Structure Determination and Refinement. *J Magn Reson* **152**, 288–302.
81. Vignols, F., Bréhélin, C., Surdin-Kerjan, Y., Thomas, D., Meyer, Y., (2005). A yeast two-hybrid knockout strain to explore thioredoxin-interacting proteins in vivo. *Proc Natl Acad Sci U S A* **102**, 16729–16734.
82. Cormack, B.P., Valdivia, R.H., Falkow, S., (1996). CS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 33–38.