Phytopathogens deliver effector proteins inside host plant cells to promote infection. These proteins can also be sensed by the plant immune system, leading to restriction of pathogen growth. Effector genes can display signatures of positive selection and rapid evolution, presumably a consequence of their co-evolutionary arms race with plants. The molecular mechanisms underlying how effectors evolve to gain new virulence functions and/or evade the plant immune system are poorly understood. Here, we report the crystal structures of the effector domains from two oomycete RXLR proteins, Phytophthora capsici AVR3a11 and Phytophthora infestans PexRD2. Despite sharing <20% sequence identity in their effector domains, they display a conserved core α-helical fold. Bioinformatic analyses suggest that the core fold occurs in ~44% of annotated Phytophthora RXLR effectors, both as a single domain and in tandem repeats of up to 11 units. Functionally important and polymorphic residues map to the surface of the structures, and PexRD2, but not AVR3a11, oligomerizes in planta. We conclude that the core α-helical fold enables functional adaptation of these fast evolving effectors through (i) insertion/deletions in loop regions between α-helices, (ii) extensions to the N and C termini, (iii) amino acid replacements in surface residues, (iv) tandem domain duplications, and (v) oligomerization. We hypothesize that the molecular stability provided by this core fold, combined with considerable potential for plasticity, underlies the evolution of effectors that maintain their virulence activities while evading recognition by the plant immune system.

Many phytopathogens secrete and/or inject “effector” proteins inside host cells to modulate cellular processes (1). Filamentous plant pathogens (oomycetes and fungi) secrete a complex arsenal of effector proteins, with hundreds of effector genes identified in the genomes of several species (2–5). Effector genes tend to reside in dynamic regions of the pathogen genomes, frequently exhibit high levels of sequence polymorphisms, and display signatures of rapid evolution, including high rates of copy number variation, presence/absence polymorphisms, and positive selection. The primary adaptive function of effectors is to promote virulence, mainly by suppressing plant immunity, and some effectors are required for full virulence (6–9). However, plants have evolved mechanisms to recognize effector proteins, resulting in effector-triggered immunity that restricts pathogen growth. This results in selective pressure on the effector to evade host recognition while maintaining its virulence activity (10). How effectors evolve to enable parasites to adapt to their hosts is not well understood. Specifically, the underlying molecular mechanisms by which effectors evolve to gain new virulence functions, adapt to their host targets, and/or evade the plant innate immune system are unclear.

Phytophthora is a genus of the oomycetes (water molds). This group of eukaryotic microorganisms includes plant pathogens that are responsible for some of the most devastating diseases of plants and have serious impacts on global crop production, on forestry, and in ornamental settings (3, 4). Phytophthora infestans, the causative agent of potato and tomato late blight, triggered the Irish potato famine in the 1840s. Phytophthora sojae causes soybean root rot, and Phytophthora capsici causes pepper and cucurbit blight. These destructive parasites continue to be a major problem for modern agriculture, with significant economic and environmental effects realized through crop losses and the methods used to control these diseases. In recent years, the sudden oak death pathogen Phytophthora ramorum has emerged as a significant threat to both private and commercial forests.

The genome sequences and effector repertoires of several economically important Phytophthora species are available, including P. infestans, P. capsici, and P. sojae (3, 5) (see the P. capsici Genome Project Web site). These parasites secrete modular effectors that are translocated into host cells (12, 13). One class of Phytophthora effectors is the RXLR family that also occurs in downy mildew oomycetes, such as Hyaloperonospora arabidopsidis (2). RXLR effectors are defined by a secretion signal peptide followed by a conserved N-terminal domain defined by the RXLR (Arg-Xaa-Leu-Arg) consensus sequence. The RXLR domain is required for translocation inside plant cells (13) but is dispensable for the biochemical activity of the effectors when expressed directly inside host cells (14). Indeed,

---

*This work was supported by the Gatsby Charitable Foundation, the Biotechnology and Biological Sciences Research Council (studentship to S. R. F. K. and core funding to M. J. B.), and the Royal Society (through a University Research Fellowship to M. J. B.). The atomic coordinates and structure factors (codes 3ZR8 and 3ZRG) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/). The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–7 and Tables 1 and 2.

1 To whom correspondence may be addressed. Tel.: 44-1603-452410; Fax: 44-1603-450011; E-mail: sophien.kamoun@tsl.ac.uk.

2 To whom correspondence may be addressed. Tel.: 44-1603-450742; Fax: 44-1603-450018; E-mail: mark.banfield@jic.ac.uk.
the effector activity appears to reside in the C-terminal regions of RXLR effectors (the “effector domain”) (14).

The effector domains of RXLR proteins display extensive sequence diversity and typically share little sequence similarity with proteins with known activities, making functional annotation and prediction of tertiary folds from sequence difficult. Despite this diversity, there are recognizable sequence relationships within the C-terminal regions of Phytophthora RXLR effectors, and some repeating sequence motifs termed “W”, “Y”, and “L” have been described (15). These analyses indicate that many, but probably not all, RXLR effector genes share a common ancestor (3, 15). How RXLR effectors encode adaptability while retaining their viability at the molecular level is unknown. Further, the virulence activities of individual RXLR effectors are only just beginning to emerge, and little is known about the biochemical mechanisms that underlie these activities (6).

Several RXLR effectors, including P. infestans AVR3a, AVR4, AVRBlb1, and AVRBlb2 and H. arabidopsis ATR1 and ATR13, were first identified based on their avirulence activity (their ability to activate effector-triggered immunity in plants that carry a corresponding resistance (R) protein (14) (supplemental Fig. 1). AVR3a is a member of a large subfamily of RXLR effectors with homologues in P. sojae (3) and P. capsici (see the P. capsici Genome Project Web site). In P. infestans, AVR3a is found in either of two major allelic forms, AVR3aKL or AVR3aKM, with K/E defining the amino acid residue at position 80 (Lys or Glu) and U/M the residue at 103 (Ile or Met) (16). These different forms trigger different responses in planta. AVR3aKL, but not AVR3aKM, activate effector-triggered immunity in the presence of the potato resistance protein R3a (16). In plants that do not carry R3a, both AVR3a forms bind and stabilize the E3 ubiquitin ligase CMPG1, leading to suppression of INF1-elicited plant immunity and programmed cell death (PCD) (17, 18). Further, the identity of the residue at the C-terminus of AVR3a is critical for INF1 suppression activity but not for R3a recognition (17). Gain-of-function mutagenesis in AVR3aKM identified a “hot spot” between residues Ser123 and Lys130 that enhances recognition by R3a and strengthens PCD suppression in response to INF1 (17).

Other oomycete RXLR effectors have been validated based on their ability to modulate defense responses in plants (supplemental Fig. 1). One such effector, P. infestans PexRD2, was identified in a functional screen to promote cell death activity in several susceptible and resistant plants (19, 20). The promotion of cell death induced by PexRD2 could reflect its virulence activity (19), but the exact mechanism by which this effector contributes to pathogenesis remains unclear.

The purpose of this study was to investigate the three-dimensional structures of RXLR effectors from Phytophthora. Our aim was to decipher the nature of the relationships between these effectors at the molecular level and determine how these might provide information on their remarkable plasticity and accelerated evolution. Here we report the crystal structures of the effector domains from two oomycete RXLR proteins: (i) P. capsici AVR3a11, a homologue of P. infestans AVR3a (6, 16–18) and AVR1b from P. sojae (21, 22) and (ii) P. infestans PexRD2. These two effectors do not share significant sequence identity in their effector domains but, remarkably, display a similar α-helical fold in their structure. Structure-informed bioinformatic analyses revealed that this fold defines a domain observed in about 44% of annotated Phytophthora RXLR effectors. We discovered that many functionally important residues and those that are polymorphic map to the surface of the proteins. Our findings led us to propose a model for the structural basis of plastic evolution in RXLR effectors. We hypothesize that the core α-helical fold (termed the “WY-domain”) can adapt through (i) insertion/deletions in loop regions between α-helices, (ii) extensions to the N and C termini, (iii) amino acid replacements in surface residues, (iv) tandem domain duplications, and (v) oligomerization. We propose that the core fold provides both a degree of molecular stability and plasticity that enables development/maintenance of effector virulence activities while allowing evasion of recognition by the plant innate immune system during rapid “arms race” co-evolution.

EXPERIMENTAL PROCEDURES

Construct Design—We used RONN (23) to guide construct design of AVR3a11 and PexRD2 for protein expression. This predicted the N-terminal regions of the effectors up to and including the RXLR signal sequence likely to be disordered.

Protein Production—Constructs of AVR3a11 and PexRD2 were designed to limit inclusion of any predicted disordered regions, and synthetic genes were ordered (optimized for expression in Escherichia coli) and then cloned into expression vectors (for details, see supplemental material). A second AVR3a11 construct (Thr70–Val132) was produced by PCR and cloned as detailed in the supplemental material. All constructs were verified by sequencing. Details of protein expression, which followed well established procedures, are given in the supplemental material. Proteins were purified by Ni2+–immobilized metal affinity chromatography and gel filtration chromatography (see supplemental material). Protein molecular masses were verified using mass spectrometry (see supplemental material).

Crystallization and Data Collection—Initial crystals of AVR3a11 (the Thr70–Val132 construct at 19 mg/ml) and PexRD2 (12 mg/ml) were grown at 20 °C in sitting drop plates set up using a robotic crystallization system and then optimized in 24-well plate hanging drop experiments. Details of precipitant and cryoprotectant solutions are given in the supplemental material. Data used to solve the structure of AVR3a11 were collected using a Rigaku RU–H3RHB generator/Mar345 detector, with high resolution data collected on beamline I04 of the Diamond Light Source, UK. All data for PexRD2 were collected on I02 (Diamond Light Source, UK). Data collection statistics are given in Table 1.

Data Processing and Structure Solution, Refinement/Rebuilding, and Validation—X-ray diffraction data were processed with iMOSFLM (24) and scaled with SCALA (25), as implemented in the CCP4 suite (26). For phasing, AVR3a11 crystals were immersed in cryoprotectant (see supplemental material).
supplemented with 100 mM potassium iodide for 30 s before freezing. The AutoSol wizard of PHENIX (27) was used to solve the structure of AVR3a11 (to 1.9 Å resolution), with a combined SAD/SIRAS approach using the anomalous signal from iodide. PHENIX was also used to produce an initial model, which was then refined against the high-resolution data. PexRD2 was co-crystallized with ammonium bromide and iodide. PHENIX was also used to produce an initial model, which was then refined against the high-resolution data.

Analytical Ultracentrifugation—Sedimentation equilibrium experiments were performed using a Beckman XLA-I analytical ultracentrifuge equipped with an An-50 Ti 8 place rotor and absorbance/interference optics. For AVR3a11, samples at 0.3 and 0.45 mg/ml were monitored using absorbance optics, and equilibrium scans were obtained at speeds of 38,000 rpm. Interference patterns were collected at equilibrium for PexRD2 with samples at 3 and 6 mg/ml and at a speed of 37,000 rpm. Data were fitted to either a single species (AVR3a11) or dimer/trimer (PexRD2) models using ULTRASCAN (34). For PexRD2, the dimeric form predominates in solution, and there was no evidence for a monomer.

Bioinformatic Analysis—MEME searches allowed for motif lengths up to 100 amino acids (for additional search parameters see the supplemental material). We constructed HMM models of the four motifs that span the WYL sequences and used this to search the effector data base. From this we assembled an updated HMM model describing the WY-domain (49 amino acids). For inclusion of sequences in this updated model, we chose a cut-off of $e < 0.12$ from the initial search, which corresponds to the position of $P. infestans$ AVR3a in the list (for sequences used, see supplemental Table 1). This new HMM model was used to re-search the RXLR effector and non-RXLR proteome databases from Phytophthora and $H. arabidopsis$ and the non-RXLR secretome of Phytophthora. Hits that scored greater than an HMM score of 0.0 were considered putative WY-domain-containing proteins. This cut-off was chosen based on the distribution of WY-domain sequences and our estimate of the false positive discovery rate in the non-RXLR secretome (see Fig. 4c and supplemental material).

Table 1: X-ray data collection, refinement statistics, and model analysis

| Data collection* | AVR3a11 (native) | AVR3a11 (iodide soak) | AVR3a11 (native, DLS) | PexRD2 (bromide peak) | PexRD2 (high resolution) |
|------------------|------------------|-----------------------|----------------------|-----------------------|-------------------------|
| Instrumentation  | Home source      | Home source           | DLS–104              | DLS–102               | DLS–102                 |
| Space group      | P2,2,2           | P2,2,2                | P2,2,2               | P2,2,2                | P2,2,2                  |
| Unit cell parameters (Å) | $a = 21.21$ | $a = 21.29$ | $a = 21.30$ | $a = 48.8$ | $a = 44.45$ |
|                  | $b = 25.24$     | $b = 25.29$           | $b = 25.33$          | $b = 52.36$           | $b = 52.92$             |
|                  | $c = 90.22$     | $c = 90.33$           | $c = 90.70$          | $c = 53.32$           | $c = 53.73$             |
| Wavelength (Å)   | 1.54             | 1.54                  | 0.73                 | 0.90                  |                         |
| Resolution range (Å) | 16.80–1.80 | 13.20–1.90 | 18.14–0.90 | 44.80–1.90 | 37.70–1.75 |
| $R_{merge}$ (%)  | 4.1 (9.5)        | 4.8 (11.3)            | 6.3 (37.2)           | 6.5 (43.0)            | 7.9 (42.5)              |
| $I/σ(I)$ (%)     | 51.0 (25.5)      | 47.2 (24.1)           | 14.0 (4.3)           | 25.4 (6.7)            | 26.3 (9.9)              |
| Completeness (%) | all data, anomalous half-sets | 95.9 (99.7) | 99.7 (100.00) | 97.5 (98.0) | 99.0 (98.7) |
|       | 15.0 (13.8)      | 15.9 (15.7)           | 4.7 (5.4)            | 12.9 (13.3)           | 15.8 (16.3)             |
| Redundancy: all data, anomalous half-sets | 9.0 (8.6) | 9.0 (8.6) | 7.0 (7.0) | 11.0 (11.0) | 15.8 (16.3) |
| No. of sites (SAD/SIRAS) | 8/4        | 11                    |                      |                      |                         |
| FOM (SOLVE/RESOLVE) | 0.44 | 0.34                  |                      |                      |                         |

* Values as reported by Scala.
\*\* Values in parentheses are for the high resolution shell (as given).
\*\*\* Values as reported by Refmac5.
\*\*\*\* Values as reported by Molprobity.

**Phytophthora** and *H. arabidopsis*.
In Planta Co-immunoprecipitations—To determine whether AVR3a11, AVR3aKI, AVR3aEM, and PexRD2 exist as oligomers in planta, co-immunoprecipitation experiments were performed using tagged effectors expressed in Nicotiana benthamiana leaves via agroinfiltration. Two different tagged constructs were made for each effector: (i) FLAG epitope-tagged constructs were cloned into pJL-TRBO (37), and (ii) GFP-epitope effector constructs were generated in pK7WGF2 (38). Details of the cloning are given in the supplemental material. Co-immunoprecipitations were performed with anti-FLAG M2 affinity gel (Sigma) using total protein extracts harvested from leaves 3 days postinoculation. Interacting proteins were detected by Western blot, probing with anti-GFP primary antibodies (Invitrogen), anti-rabbit secondary antibodies conjugated to horseradish peroxidase (Sigma), and a chemiluminescent substrate (Thermo Scientific). Additional details are given in the supplemental material.

RESULTS AND DISCUSSION

To address structure/function relationships in oomycete RXLR effectors, we established an E. coli-based expression screen aimed at producing proteins that were amenable to structure determination. This screen identified the P. capsici effector AVR3a11 (a homologue of AVR3a from P. infestans) and P. infestans effector PexRD2, as highly expressed and soluble. The purified proteins were crystallized, and their structures were determined to 0.9 and 1.75 Å, respectively (Table 1).

The Structure of AVR3a11 Is a Monomeric Four-helix Bundle—The structure of AVR3a11 comprises Thr70–Val132 (63 residues) of the native sequence. This construct lacks the N-terminal signal sequence and RXLR translocation motif, which were predicted to be disordered (Figs. 1, a and b, and supplemental Fig. 2). An earlier construct comprising residues Gly63–Val132 was recalcitrant to crystallization but amenable to NMR analysis. Preliminary structure calculations using NMR data suggested that residues Gly63–Lys69 were not well structured (see supplemental material). The effector domain of AVR3a11 adopts a four-helix bundle fold, stabilized by a hydrophobic core, with the helices connected by loop regions (Fig. 1c). The crystal structure suggests the effector domain of AVR3a11 is a monomer. This oligomeric state was also observed in solution, as shown by the retention volume of the protein on a size exclusion column during purification (data not shown) and analytical ultracentrifugation (7350 ± 0.03 Da, Fig. 2a (theoretical mass = 7474 Da)). Further, Avr3a11 was shown to be monomeric in planta using co-immunoprecipitation (Fig. 2c and supplemental Fig. 3). Structure-based database searches (39) using the refined AVR3a11 model revealed low significance hits to other four-helix bundles, including KaiA from Anabena sp. PCC7120 (Protein Data Bank code 1R5Q). KaiA is a cyanobacterial circadian clock regulator (40), and although the overall fold of AVR3a11 is loosely related to KaiA, the differences in structure and surface properties do not suggest a role for the AVR3a family as circadian rhythm regulators.
A Conserved Fold in Oomycete RXLR Effectors

AVER3a11 belongs to a family of Phytophthora effectors that includes *P. infestans* AVR3a and AVR1b from *P. sojae* (supplemental Fig. 4a). Pairwise alignments based on the effector domains revealed that AVR3a11 shares 41% sequence identity with AVR3a and 46% sequence identity with AVR1b (spanning residues 70–132, AVR3a11 numbering). The structure of AVR3a11 is therefore representative of this family and provides an opportunity to interpret RXLR effector domain function with respect to structure.

**Figures**

**Fig. 2.** RXLR effectors can adopt different oligomeric states in vitro and in vivo. Fits of the analytical ultracentrifugation data (including residuals) for (a) AVR3a11 (0.3 mg/ml (circles) and 0.45 mg/ml (triangles)) (b) PexRD2 (3 mg/ml (triangles) and 6 mg/ml (circles)) to the models described confirmed that monomeric and dimeric species are the prevalent forms of these proteins in solution. c, co-immunoprecipitation (co-IP) shows that PexRD2 self-associates in planta; no oligomerization is apparent for AVR3a11, AVR3aEM, or AVR3aSM. RuBisCO is included as a loading control (Coomassie-stained SDS-polyacrylamide gel).

**Fig. 3.** Functionally important and polymorphic residues are presented on the surface of AVR3a11 and PexRD2. a, in AVR3a11, Glu71 and Gln94 (equivalent to the Lys/Glu80 and Ile/Met103 positions of AVR3a), Ser112 (equivalent to Ser112 of AVR3a), and a Tyr residue near the C terminus are located on the protein surface. b, in PexRD2, the polymorphic residues between the *P. infestans* and *P. mirabilis* homologues are also surface-displayed. The positions of these residues in the primary sequences of the proteins are shown as brown boxes in Fig. 1b.
AVR3a11) anchor $\alpha_3$ to the bundle by forming interactions within the hydrophobic core. Mutation of the Trp$^{96}$ and Leu$^{110}$ equivalents in AVR3a led to reduced protein stability in planta (17); similar mutations in hydrophobic residues of AVR1b also led to loss of function (21), most likely due to protein destabilization. The W- and Y-motifs together appear to form a critical
A Conserved Fold in Oomycete RXLR Effectors

### Table 2

|               | RXLR effectors, above threshold | RXLR effectors, below threshold | Non-RXLR proteome, above threshold | Non-RXLR proteome, below threshold |
|---------------|---------------------------------|---------------------------------|-----------------------------------|-----------------------------------|
| P. infestans  | 221                             | 342                             | 39                                | 17,553                            |
| P. ramorum    | 181                             | 128                             | 40                                | 15,543                            |
| P. sojae      | 125                             | 210                             | 66                                | 18,787                            |
| Total: Phytophthora | 527 (44%)                      | 680 (56%)                       | 145 (0.3%)                        | 51,883 (97.7%)                    |
| H. arabidopsis| 35 (26%)                        | 99 (74%)                        | 91 (0.6%)                         | 14,734 (99.4%)                    |
| Total: Phytophthora and H. arabidopsis | 562 | 779 | 236 | 66,617 |

folding unit, the WY-domain, that, with the addition of another helix (to the N terminus), can form a four-helix bundle.

**PexRD2** Is Also a WY-domain Protein but with Adaptations—

The structure of the effector domain of PexRD2 comprises Ala$^{57}$-Ala$^{120}$ (64 residues), starting immediately after the RXLR sequence through to the penultimate residue in the construct (Val$^{121}$ could not be modeled in the electron density). The structure is composed of five $\alpha$-helices and crystallizes as a dimer with intimate hydrophobic contacts at the interface (buried surface area = 1830 Å$^2$, 20% of the surface area of each monomer; PISA CCS score = 0.89 (43)).

Structure-based database searches (39) did not reveal any significant hits to PexRD2. Given the lack of sequence identity (<20%, supplemental Fig. 4b), we were surprised to discover that comparison of the structures of a PexRD2 monomer with AVR3a11 using SSM (36) aligned 37 residues comprising $\alpha_3$, loop-2, $\alpha_9$, and $\alpha_4$ with a root mean square deviation = 1.0 Å (Fig. 1e). This region comprises the helices of the WY-domain described for AVR3a11. However, PexRD2 contains no equivalent of $\alpha_6$ from the AVR3a11 four-helix bundle, and a significant insertion (16 residues) is present between $\alpha_3$ and $\alpha_4$ (loop-3). The structurally aligned residues only share 13.5% sequence identity (Fig. 1b and supplemental Fig. 4b).

**PexRD2** Is a Dimer in Solution and Self-associates in Planta—

To confirm that the PexRD2 dimer was not a crystallographic artifact, we used analytical ultracentrifugation to show that this oligomeric state predominates in solution (Fig. 2b). The best fit of the data was obtained using a dimer/tetramer model with the molecular masses of the species fixed at their theoretical value. The retention volume of PexRD2 on a size exclusion column during purification also supports the dimer conformation (data not shown). Further, we performed *in planta* co-immunoprecipitation experiments with epitope-tagged effectors. These showed that PexRD2, but not AVR3a11 (see above) or AVR3a, self-associates when transiently expressed in *N. benthamiana* leaves via agroinfiltration (Fig. 2c and supplemental Fig. 3). Given the evidence from the crystal structure and analytical ultracentrifugation, it is reasonable to assume that PexRD2 is dimerizing in planta.

**Polymorphic Residues in PexRD2 Are Surface-presented—**

PexRD2 from all strains of *P. infestans* sequenced to date displays 100% sequence identity in the effector domain (44). However, a PexRD2 homologue in the closely related species *P. mirabilis* PIC99114 contains polymorphisms at five positions (Fig. 3b and supplemental Fig. 4c) (44). Comparison of these two sequences revealed that positive selection has acted on this gene because the ratio of non-synonymous to synonymous substitutions is >1 (dN/dS = 1.27). The PexRD2 structure revealed that these 5 polymorphic residues are presented on the protein surface and are not co-localized (i.e. do not define a particular surface region) (Fig. 3b). We conclude that polymorphisms in some surface-exposed residues may have contributed to the evolution of PexRD2 without disrupting the WY-domain fold.

The **WY-domain Is a Unit Conserved in Many Phytophthora and H. arabidopsis RXLR Effectors**—

The unexpected structural similarity we discovered between AVR3a11 and PexRD2 prompted us to re-evaluate the context of the repeated WYL motifs previously reported in *Phytophthora* RXLR effectors (3, 15) and extend this analysis to *H. arabidopsis*. We used MEME (45) to search for conserved motifs in a data set comprising the *P. infestans*, *P. ramorum*, and *P. sojae* RXLR effector repertoires (1207 proteins, C-terminal effector domain only, downstream of the RXLR). The highest scoring motifs are dominated by overlapping regions incorporating W-, Y-, and L-motifs that can be readily aligned to reveal repeating WYL sequences (supplemental Fig. 6). One newly identified 49-amino acid MEME motif spanned both the W- and Y-motifs and covers the $\alpha_2$/L-motif are positioned 8 residues to the N terminus downstream of the RXLR (Fig. 4a). Interestingly, the alignment of overlapping MEME motifs revealed that the conserved Leu of the L-motif is positioned 8 residues to the N terminus of the conserved Trp of the W-motif.

To explore the extent to which the WY-domain is found in RXLR effectors, we constructed HMM models of four MEME motifs (Fig. 4b, supplemental Fig. 7, and supplemental Table 1) and re-screened the *Phytophthora* and *H. arabidopsis* RXLR effector sequences with HMMER (46); we also screened the non-RXLR proteome of *Phytophthora* and *H. arabidopsis* and the non-RXLR secretome of *Phytophthora* with the WY-domain HMM. This analysis revealed that WY-domains are highly enriched in RXLR effectors compared with the non-RXLR proteome (Fig. 4c). Using an appropriate cut-off (see “Experimental Procedures”), we revealed that 527 of 1207 (44%) of *Phytophthora* and 35 of 134 (26%) of *H. arabidopsis* RXLR effectors contain WY-domain-like sequences (Table 2 and supplemental Table 2). In contrast, in the non-RXLR proteome, only 0.3 and 0.6% (*Phytophthora* and *H. arabidopsis*, respectively) contain WY-domain-like sequences above the same cut-off (Table 2 and supplemental Table 2). We therefore conclude that our rate of false positive is ≤0.6% and that 562 of 1341 (42%) of oomycete RXLR effectors probably contain a core fold similar to that found in AVR3a11 and PexRD2.
WY-domains can also be found in tandem repeats, and we identified one candidate effector, PitG_23035, with 11 WY-domains. Therefore, many Phytophthora RXLR effectors comprise either single or tandem arrays of WY-domains, joined by variable “linker” regions.

A Core Protein Structure for Phytophthora RXLR Effectors That Supports Rapid Adaptive Diversification—Here we present the crystal structures of the effector domains from Phytophthora RXLR proteins AVR3a11 and PexRD2. We describe a structural unit, the WY-domain, that bioinformatic analysis suggests is conserved in ~4% of Phytophthora RXLR effectors and ~26% of H. arabidopsidis effectors. This domain forms a core α-helical fold that tolerates considerable plasticity at both the N and C termini and within the loops between the helices. It also supports amino acid substitutions in surface residues and oligomerization as further mechanisms for modification.

Understanding the mechanisms that underlie protein adaptation under evolutionary pressure is a key question in molecular evolution. Adaptive evolution can drive the development of novel protein functions or the tuning of existing ones (e.g. new substrate specificity in an enzyme active site or perturbation of molecular recognition surfaces). Some proteins can also be under selective pressure to lose particular activities (e.g. to evade immune system recognition but still need to maintain structural integrity to avoid total loss of function). The evolution of toxicity in Kallikrein-1-like venom proteins involves the acquisition of small insertions in loop regions adjacent to the catalytic cleft (within a conserved core fold), followed by accelerated sequence diversification that increases catalytic efficiency (47). Many small protein domains (~20–60 amino acids) exist as tandem repeats to deliver a structural template that can adapt to provide novel functions. Examples include leucine-rich repeats (48) and the blades of β-propellers. Each of these structural modules encode functional diversity through variation in the number of repeats, mutations in individual surface residues, and larger changes to loop regions. Further, TIM-barrel enzymes are built from a minimal ~20-amino acid α/β unit, usually found in eight repeats, with specificity for the types of reaction catalyzed, and their substrates, encoded by the β/α loops (49). The degree of sequence conservation across the family of known β-propellers and TIM-barrels is below detectable levels. It is generally accepted that all of these proteins have evolved from ancestral precursors through diversification.

Plant pathogen effectors frequently display extreme signatures of positive selection, presumably as a consequence of the co-evolutionary arms race with resistant hosts (3, 11, 44, 50, 51). Our findings provide a molecular framework that explains how RXLR effectors modulate immunity inside plant cells (supplemental Fig. 1).

In conclusion, we have defined a large family of oomycete RXLR effectors that contain at least one WY-domain and are structurally related. However, this does not preclude the existence of Phytophthora RXLR effectors that contain alternative folds and are phylogenetically unrelated to WY-domain effectors (Fig. 4c). In fact, given the diversity of RXLR sequences, it seems likely that other folds do exist, and these have yet to be structurally characterized.

Acknowledgments—We gratefully acknowledge the staff of the Diamond Light Source for access to data collection facilities and Alex Jones and Jan Sklenar (Sainsbury Laboratory) for mass spectrometry. We thank Jorunn Bos, Tolga Bozkurt, Liliana Cano, Angela Chaparro-Garcia, Carla Garzón, Ricardo Oliva, and Sebastian Schornack for resources and discussion.

REFERENCES

1. Dodds, P. N., and Rathjen, J. P. (2010) Nat. Rev. Genet. 11, 539–548
2. Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., Thines, M., Ah-Fong, A., Anderson, R., Badejoko, W., Bitter-Eddy, P., Boone, J. L., Chibucos, M. C., Coates, M., Dehal, P., Delehaunty, K., Dong, S., Downton, P., Dumas, B., Fabro, G., Fronick, C., Fuerstenberg, S. I., Fulton, L., Gaulin, E., Govers, F., Hughes, L., Humphray, S., Jiang, R. H., Judelson, H., Kamoun, S., Kyung, K., Meijer, H., Minx, P., Morris, P., Nelson, J., Phuntumart, V., Qutob, D., Rehmany, A., Rougon-Cardoso, A., Ryden, P., Torto-Alabio, T., Studholme, D., Wang, Y., Win, J., Wood, J., Clifton, S. W., Rogers, J., Van den Ackerveken, G., Jones, J. D., McDowell, J. M., Beynon, J., and Tyler, B. M. (2010) Science 330, 1549–1551
3. Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, I. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alabio, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fuglstad, J., Gilroy,
