



Article

# Crystal Structure-Based Exploration of Arginine-Containing Peptide Binding in the ADP-Ribosyltransferase Domain of the Type III Effector XopAI Protein

Jyung-Hung Liu <sup>1,2,3,4,\*</sup> , Jun-Yi Yang <sup>3,5</sup> , Duen-Wei Hsu <sup>6</sup>, Yi-Hua Lai <sup>2</sup>, Yun-Pei Li <sup>1</sup>, Yi-Rung Tsai <sup>1</sup> and Ming-Hon Hou <sup>1,2,3,4</sup>

<sup>1</sup> Institute of Genomics and Bioinformatics, National Chung Hsing University (NCHU), Taichung 40227, Taiwan; yunk293@gmail.com (Y.-P.L.); stave0972@gmail.com (Y.-R.T.); mhho@nchu.edu.tw (M.-H.H.)

<sup>2</sup> Department of Life Science, NCHU, Taichung 40227, Taiwan; hbm486426@gmail.com

<sup>3</sup> Graduate Institute of Biotechnology, NCHU, Taichung 40227, Taiwan; jyang@nchu.edu.tw

<sup>4</sup> PhD Program in Medical Biotechnology, NCHU, Taichung 40227, Taiwan

<sup>5</sup> Graduate Institute of Biochemistry, NCHU, Taichung 40227, Taiwan

<sup>6</sup> Department of Biotechnology, National Kaohsiung Normal University, Kaohsiung 80201, Taiwan; dwhsu@nknuc.nknu.edu.tw

\* Correspondence: jhliu@nchu.edu.tw; Tel.: +886-4-2284-0338

Received: 1 October 2019; Accepted: 12 October 2019; Published: 14 October 2019



**Abstract:** Plant pathogens secrete proteins called effectors into the cells of their host to modulate the host immune response against colonization. Effectors can either modify or arrest host target proteins to sabotage the signaling pathway, and therefore are considered potential drug targets for crop disease control. In earlier research, the *Xanthomonas* type III effector XopAI was predicted to be a member of the arginine-specific mono-ADP-ribosyltransferase family. However, the crystal structure of XopAI revealed an altered active site that is unsuitable to bind the cofactor NAD<sup>+</sup>, but with the capability to capture an arginine-containing peptide from XopAI itself. The arginine peptide consists of residues 60 through 69 of XopAI, and residue 62 (R62) is key to determining the protein–peptide interaction. The crystal structure and the molecular dynamics simulation results indicate that specific arginine recognition is mediated by hydrogen bonds provided by the backbone oxygen atoms from residues W154, T155, and T156, and a salt bridge provided by the E265 sidechain. In addition, a protruding loop of XopAI adopts dynamic conformations in response to arginine peptide binding and is probably involved in target protein recognition. These data suggest that XopAI binds to its target protein by the peptide-binding ability, and therefore, it promotes disease progression. Our findings reveal an unexpected and intriguing function of XopAI and pave the way for further investigation on the role of XopAI in pathogen invasion.

**Keywords:** type III effectors; peptide-binding domain; mono-ADP-ribosyltransferase; crystal structure; molecular dynamics simulation

## 1. Introduction

Citrus canker is a disease affecting citrus crops; it is caused by a non-indigenous bacterial pathogen *Xanthomonas axonopodis* pv. *citri* (*Xac*) [1]. This pathogen generally causes leaf spotting and blemishing of the rind of the fruit, but more severe infections result in shoot dieback and fruit drop. As *Xac* is a major threat to citrus production worldwide, farmers and governments spend their time and millions

of dollars annually on prevention and disease control. Similar to many Gram-negative bacterial pathogens, *Xac* delivers effector proteins directly into host cells via the type III secretion system during the infection process [2,3]. Some of the effectors are enzymes, whereas others are transcription factors or adaptors for protein–protein interaction. They are required for the development of disease symptoms in susceptible citrus plants and for a hypersensitive response in resistant plants [4,5]. Therefore, they are valuable to the development of specific inhibitors against plant diseases [6].

XopAI, encoded by XAC3230 in *Xac*, is a putative type III effector composed of 296 amino acids. Although it has been suggested to be a pathogenicity factor for citrus canker [7,8], the role of XopAI in the virulence of *Xanthomonas* remains to be characterized. The first 43 N-terminal residues of XopAI share a sequence similarity with effectors, XopE and XopJ [9]. Further, this N-terminal sequence contains an N-myristoylation motif, which was previously examined through experiments on the XopE and XopJ proteins of *Xanthomonas campestris* pv. *vesicatoria* [10]. Effectors with this motif target cellular membranes in their hosts.

Based on the sequence of the C-terminal region (residues 120 to 286), XopAI was previously annotated as a member of the Arg-specific mono-ADP-ribosyltransferase (mART) family [11]. Bacterial pathogens use mART to alter or inhibit the activity of their target proteins in host cells through the covalent transfer of the ADP-ribose group from the cofactor NAD<sup>+</sup> onto the Arg residue of its target protein [12,13]. In this study, we determined the crystal structure of XopAI from *Xac*. We found that the C-terminal putative mART domain of XopAI is structurally similar to that of the type III effector, HopU1 in *Pseudomonas syringae* pv. *tomato*. HopU1 functions as a mART in *Arabidopsis thaliana*, and it targets several RNA-binding proteins including GRP7 [14]. However, XopAI does not seem eligible to be a mART because of the change in the critical residues in the active site. From analyses of the crystal packing, we found that XopAI uses an altered mART domain to bind its own N-terminal peptide containing a conserved Arg residue. We also conducted a series of molecular dynamics (MD) simulations to investigate the structural, dynamic, and energetic properties of this protein–peptide interaction. From structural dissections, our data uncovered an unexpected function of XopAI and provided valuable snapshots that could help clarify the role of XopAI in bacterial pathogenicity.

## 2. Results and Discussion

### 2.1. Structure of XopAI

We found that the full-length XopAI protein produces two types of crystals belonging to space groups,  $P4_32_12$  and  $P4_12_12$  under similar crystallization conditions. The structure of XopAI was determined by bromide multiple-wavelength anomalous diffraction (Br-MAD) [15] using  $P4_32_12$  crystals, and the final model was refined to a resolution of 2.01 Å (Table 1). In the Br-MAD dataset, the asymmetric unit contains four bromide ions (Figure S1A), which have hydrogen-bonding contacts with backbone nitrogen atoms or Arg sidechains of the protein (Figure S1B). However, the  $P4_12_12$  crystals diffract to a resolution of 1.53 Å and the structure was solved via molecular replacement. Amino acid residues 1–59 of XopAI in the  $P4_32_12$  crystals were not built in the model as they lacked electron density, possibly owing to the intrinsic disorder in the N terminus. Similarly, the first 58 residues of XopAI in the  $P4_12_12$  crystals were not built in the model. From sequence-based predictions (Figure S2), we identified that the first 70 residues of XopAI may be highly disordered. In the  $P4_32_12$  and  $P4_12_12$  crystals, one XopAI molecule was found in the asymmetric unit. The statistics of the data collection and refinement for these crystals are summarized in Table 2.

Based on the crystal structure in the  $P4_32_12$  crystals, the overall dimensions of XopAI are  $\sim 42 \times 45 \times 52$  Å<sup>3</sup>. XopAI resembles a bent right hand (Figure 1A), and it can be dissected as a two-lobe structure with an N-terminal  $\alpha$ -helical lobe (residues 74 to 193), and a C-terminal  $\beta$ -sandwich lobe (residues 194 to 296) (Figure 1B). The N-terminal lobe resembles a thumb and palm and contains five  $\alpha$ -helices ( $\alpha 1$ – $\alpha 5$ ) (Figure 1A,C), which are packed against each other to form a stable core structure. The C-terminal lobe resembles fingers and belongs to an atypical  $\beta$ -sandwich fold where seven strands

form two sheets. One  $\beta$ -sheet is formed by  $\beta$ -strands  $\beta_3$  and  $\beta_5$ , and the other is formed by  $\beta$ -strands  $\beta_1$ ,  $\beta_2$ ,  $\beta_4$ ,  $\beta_6$ , and  $\beta_7$  (Figure 1A,C). A central cleft located between the N- and C-lobes forms a potential active site (Figure 1B).

**Table 1.** Data collection statistics for Br-MAD phasing.

Dataset	Native	Br-MAD		
		$\lambda_1$ (H rem) <sup>1</sup>	$\lambda_2$ (infl)	$\lambda_3$ (Peak)
Wavelength (Å)	0.9762	0.9056	0.9193	0.9190
Space group		$P4_32_12$		
Resolution (Å)	2.0	2.3	2.4	2.5
Redundancy	4.1	2.4	4.6	2.5
Completeness (%)	93.7 (85.3) <sup>2</sup>	98.5 (99.7)	98.4 (99.7)	98.2 (99.7)
Average I/ $\sigma$ (I)	11.1 (2.5)	17.7 (3.5)	19.5 (4.0)	18.9 (3.8)
R <sub>merge</sub> (%)	8.2 (52.5)	11.2 (49.2)	10.2 (44.7)	10.3 (43.6)

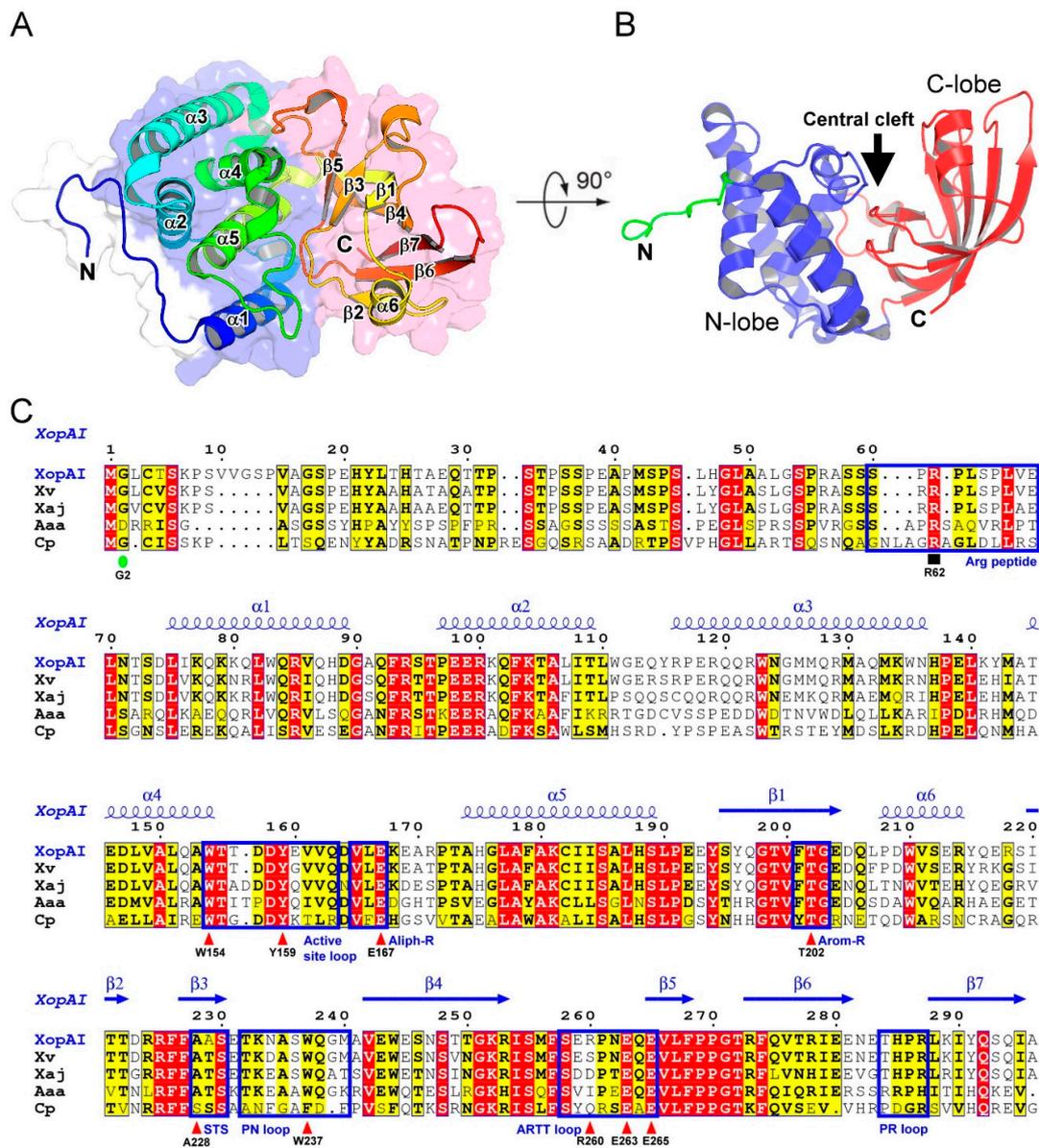
<sup>1</sup> MAD data were collected at wavelengths corresponding to the high energy remote, inflection point, and peak of the atomic absorption edge. <sup>2</sup> Values in parenthesis indicate those for the highest resolution shell.

**Table 2.** Data collection and refinement statistics.

Protein	XopAI	XopAI- $\Delta$ N70	
<b>Data collection</b>			
Wavelength (Å)	0.9762	0.9198	0.9762
Space group	$P4_32_12$	$P4_12_12$	$P2_1$
Cell parameters (Å)	a = b = 73.05 Å, c = 114.06 Å	a = b = 52.98 Å, c = 212.09 Å	a = 62.78 Å, b = 98.76 Å, c = 77.45 Å, $\beta$ = 91.21°
Resolution (Å)	2.01	1.53	2.26
Mosaicity (°)	1.07	0.29	1.23
Redundancy	4.1	10.9	3.5
Completeness (%)	93.7 (85.3) <sup>1</sup>	100.0 (100.0)	99.4 (99.3)
Average I/ $\sigma$ (I)	11.1 (2.5)	9.6 (1.7)	7.5 (1.8)
R <sub>merge</sub> (%)	8.2 (52.5)	13.1 (106.4)	14.2 (78.3)
CC <sub>1/2</sub>	0.997 (0.712)	0.997 (0.647)	0.990 (0.609)
<b>Refinement</b>			
Resolution limit (Å)	26.56–2.01	36.89–1.53	31.53–2.26
R <sub>work</sub> (%)	16.6	16.1	18.9
R <sub>free</sub> (%)	19.5	17.1	23.2
Number of non-H atoms			
Protein	1961	1973	7354
Water	236	374	558
Ramachandran plot statistics			
Favored regions (%)	99.15	98.73	98.43
Allowed regions (%)	0.85	1.27	1.46
Disallowed regions (%)	0	0	0.11
Average B factor (Å <sup>2</sup> )	25.9	22.7	35.1
R.m.s. deviation from ideality			
Bond length (Å)	0.004	0.010	0.004
Bond angle (°)	0.655	1.024	0.561
Protein Data Bank (PDB) code	6KLY	6K93	6K94

<sup>1</sup> Values in parenthesis indicate those for the highest resolution shell.

We found the homologous proteins of XopAI in many *Xanthomonas* species (Figure 1C and Figure S3). In addition, we found potential XopAI homologs in two *Acidovorax* species (*A. citrulli* and *A. avenae* subsp. *avenae*) and *Collimonas pratensis*. *A. citrulli* and *A. avenae* subsp. *avenae* cause diseases in a wide range of economically important plants [16], while *C. pratensis* inhibits fungal growth [17]. The multiple sequence alignment shows that residues in the C-lobe are, in general, more conserved than those in the N-lobe. The N-terminal disordered region is the least conserved region in the entire protein.



**Figure 1.** Crystal structure of XopAI and protein alignment with its homologs. (A) The ribbon model that shows the structure of XopAI spanning residues 60 to 296. It is colored in the rainbow scheme (from N- to C-terminus, from blue to red), and its secondary structural elements are labeled. Regions that belong to the N- and C-lobes are contoured in blue and red shadows, respectively, while the N-terminal disordered region is in white. (B) A view rotated by 90° relative to (A) showing the central cleft formed between the N- and C-lobes (blue and red ribbons, respectively). The N-terminal disordered region is highlighted in green. The central cleft is indicated by an arrow. (C) The sequence alignment of XopAI homologs; we selected five representative sequences for this concise alignment, and a more comprehensive version with 17 sequences is shown in Figure S3. The conserved residues are shaded yellow and identical residues are shaded red. Secondary structure elements based on XopAI crystal structure determined in this study are displayed above the alignment. Blue boxes outline those important regions in XopAI. Key residues in the central cleft are marked with red triangles and labeled according to XopAI sequence. A green oval shows the putative myristoylation site, and a black square indicates the crucial Arg in the Arg peptide. The following bacteria strains were analyzed: XopAI (*Xanthomonas axonopodis* pv. *citri*, GenBank accession no.: WP\_011052119, this study), Xv (*X. vesicatoria* strain LM159, CP018470), Xaj (*X. arboricola* pv. *juglandis* strain Xaj 417, CP012251), Aaa (*A. avenae* subsp. *avenae*, AVS84630), and Cp (*Collimonas pratensis*, WP\_061944107).

## 2.2. Structural Comparison of XopAI with mARTs

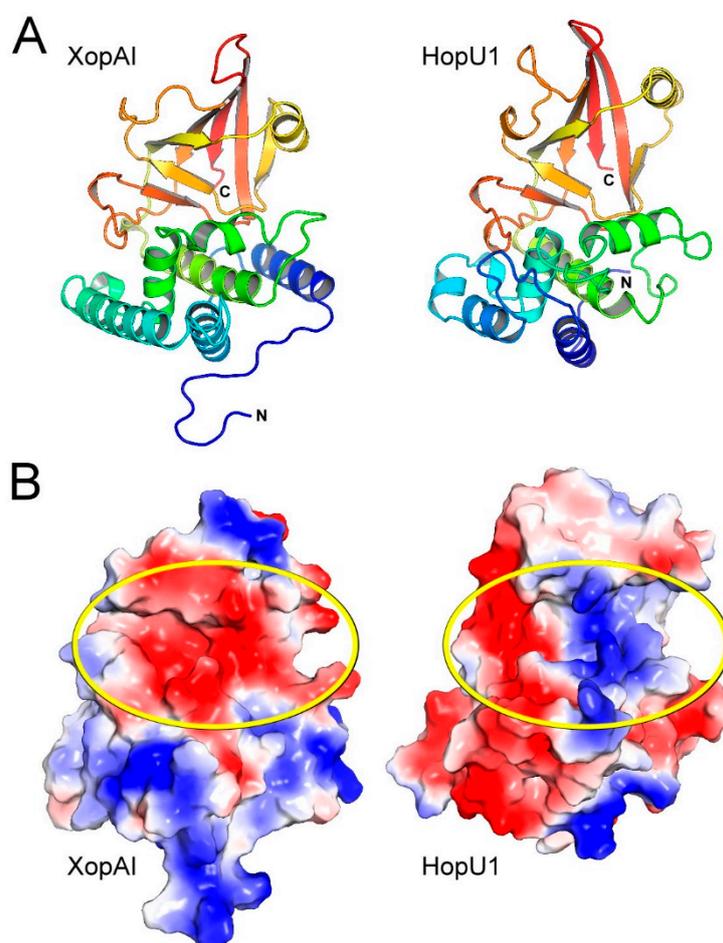
To identify the structural homologs of XopAI, we compared the coordinates of XopAI against those in the Protein Data Bank (PDB) at a 90% non-redundancy level. Among 31 entries with a structural similarity (Z score higher than 2.0), we found four proteins with a Z score higher than 15.0: *P. syringae* type III effector HopU1 (PDB code 3U0J) [18], *Serratia proteamaculans* type VI secretion ADP-ribosyltransferase effector Tre1 (PDB code 6DRH) [19], rat ecto-mART ART2.2 (PDB code 1GXY) [20], and *P. aeruginosa* exoenzyme S (PDB code 6GN8) [21] (Figure S4). The mART proteins are classified into H-Y-E, variant H-Y-E, and R-S-E classes based on the conserved residues in the active site (i.e., H-Y-E stands for His-Tyr-Glu and R-S-E stands for Arg-Ser-Glu) [22]. The aforementioned four mART proteins share only 20–30% sequence similarity with XopAI, and the experimental data confirmed that they belong to the R-S-E class mARTs. XopAI and these four proteins are superimposed with a root-mean-square deviation of 2.7 Å over 191 (for HopU1), 2.9 Å over 182 (for Tre1), 2.8 Å over 181 (for ART2.2), and 2.7 Å over 174 C $\alpha$  atom positions (for ExoS), respectively (Figure S5). This structural comparison result clearly indicates that XopAI folds like an R-S-E class mART. However, it is unclear if XopAI is truly a mART.

When comparing XopAI with HopU1 (Figure 2A), we found that their core folds are similar, and they adopt a mixed  $\alpha/\beta$ -fold with a characteristic  $\beta$ -sandwich structure. XopAI shows some similarity (21% sequence identity) to HopU1; however, we noticed a significant difference in their surface charge distribution. XopAI possesses a highly negatively charged surface at its central cleft (Figure 2B), whereas a positively charged surface is required for interacting electrostatically with the negatively charged phosphate group of NAD<sup>+</sup> in mARTs. As shown in Figure 2B, the central cleft of HopU1 is the active site for ADP-ribosylation. It has a positively charged surface on its right-hand side, and a negatively charged surface on the left-hand side. This negatively charged surface is presumably for the accommodation of Arg (the target residue of ADP-ribosylation). A similar surface charge distribution at the HopU1 active site can also be found in other mARTs, except that the ExoS active site is mostly positively charged (Figure S5). Based on this finding, we propose that XopAI may not be a qualified mART, and it would exert different effects on host cells. Therefore, the sequence features that contribute to this different surface charge distribution on XopAI remain unidentified.

The R-S-E class mARTs contain three conserved sequence features at the active site [23,24] (Figure S4): (1) The arom-R motif contains an aromatic amino acid followed by an Arg ([YFL]-R-X). It is located at the  $\beta$ -strand  $\beta$ 1 and contributes to NAD<sup>+</sup> binding. (2) The ARTT (ADP-ribosyl-turn-turn) loop connects  $\beta$ -strands,  $\beta$ 4 and  $\beta$ 5, and it contains a [QE]-X-E motif. The conserved catalytic Glu residue at the third position is required for NAD<sup>+</sup> cleavage and transferase activity. (3) The STS motif in the  $\beta$ -strand,  $\beta$ 3 has a sequence pattern S-[TS]-[STQ], which stabilizes the structure of the active site through hydrogen bonds with catalytic Glu residue, and other conserved NAD<sup>+</sup> binding residues. In addition to the three features, a less conserved aliph-R motif composed of an aliphatic amino acid prior to Arg (X-[LVI]-R) resides immediately after the active site loop and contributes to NAD<sup>+</sup> binding. In XopAI, although the motif in the ARTT loop remains unchanged, the sequence of the arom-R motif is FTG, the sequence of aliph-R motif is VLE, and the sequence of STS motif is AAS. The change from Ser and Thr to Ala in the STS motif may affect active site integrity because the Ala sidechain fails to form hydrogen bonds. The Arg to Thr change in the arom-R motif leads to the surface being electrically neutral. Moreover, the Arg to Glu change in the aliph-R motif dramatically reverses the surface charge from positive to negative. As a result of these factors, there is a change in the surface charge distribution at the central cleft of XopAI, and possibly there is a change in its cofactor preference.

These sequence alterations are widely found in XopAI homologs (Figure 1C and Figure S3). XopAI homologs in *Xanthomonas* and *Acidovorax* carry a sequence of FTG in the arom-R motif, and VLE in the aliph-R motif. In addition, the sequence of the STS motif is either AAS or ATS. Similarly, a distantly-related XopAI homolog found in *C. pratensis* has YTG in the arom-R motif and VFE in the aliph-R motif. Although its STS motif reads SSS, which is similar to mARTs, the aforementioned

sequence features still requires a negatively charged surface at the central cleft. Mapping the conservation from the multiple sequence alignment onto the XopAI surface reveals that conserved residues are clustered around the central cleft (Figure S6). Similarly, HopU1 has a high level of amino acid residue conservation in its active site. This suggests that the negatively charged cleft of XopAI is still important for protein functionality. However, the function of this central cleft is unclear.



**Figure 2.** Structural comparison between XopAI and HopU1. (A) Ribbon diagram showing XopAI and HopU1 from *Pseudomonas syringae*. The structures are colored in the rainbow scheme. (B) Comparison of electrostatic surface of XopAI and HopU1. The regions of negative and positive potential are shown in red and blue, respectively; uncharged and hydrophobic surface areas are colorless. The central cleft on XopAI and the ADP-ribosylation site on HopU1 are marked with yellow ovals.

### 2.3. The Central Cleft of XopAI Has the Ability to Bind Peptides

During our inspection of the crystal packing (Figure S7A,B), we found that the negatively charged central cleft of XopAI is important for the formation of  $P4_32_12$  and  $P4_12_12$  crystals. Every cleft binds a segment of the N-terminal sequence (residues 60 to 69) from a neighboring protein. Consequently, this tandem interaction allows full-length XopAI proteins to form spiral threads and fill up the crystal. To study the influence of this N-terminal region on the crystallization of XopAI, we produced an N-terminal-truncated protein, XopAI- $\Delta$ N70 (by deleting the first 70 residues) and grew the crystals. The crystals belong to space group  $P2_1$ , and they diffract with a resolution of 2.26 Å (Table 2). The structure of XopAI- $\Delta$ N70 was solved via molecular replacement. In XopAI- $\Delta$ N70 crystals, the asymmetric unit contains four copies of the protein. However, we did not observe a thread-like and tandem protein packing in XopAI- $\Delta$ N70 crystals (Figure S7C). This result suggests that the

intermolecular association of the central cleft and the N-terminal sequence is responsible for the observed tandem packing in the crystals of the full-length XopAI protein.

Although the  $P4_32_12$  and  $P4_12_12$  crystal forms contain one molecule per asymmetric unit, their crystal packing environments differ significantly owing to the interaction orientation (Figure 3A). For the protein interaction mode found in the  $P4_12_12$  crystals, the central cleft binds residues 59 to 70 from an adjacent protein, and it has an interface area of  $673 \text{ \AA}^2$ . However, in the  $P4_32_12$  crystals, the cleft binds residues 60 to 66 of another protein, and the interface area is  $470 \text{ \AA}^2$ . In a previous study, Lo Conte et al. reported that the interface area in binary protein complexes is  $\sim 800 \text{ \AA}^2$  [25]; in addition, they suggested that a typical protein–protein interface may involve 22 residues, and the minimum interface area for the stability of a protein–protein complex is  $\sim 500 \text{ \AA}^2$ . Given that protein interaction found in the two crystal forms is mediated through a small sequence of amino acid residues and that it covers only a small interface area, we propose that this interaction belongs to a protein–peptide interaction, and the central cleft of XopAI is a peptide-binding cleft. In addition, we speculate that residues spanning from 60 to 69 in XopAI constitute a candidate peptide for the binding. Intrinsic Trp fluorescence studies using a synthetic peptide ArgP14aa (Figure S8A) showed that the presence of ArgP14aa increased Trp fluorescence of XopAI- $\Delta N70$  (Figure S8B,C), indicating a decrease in the net polarity of the environment surrounding Trp residues within the protein. In contrast, the presence of Arg did not affect the fluorescence. Figure 3B shows that among the eight Trp residues in XopAI, W154 and W237 on the cleft surface contribute to the increase in the fluorescence intensity because they will be covered during the peptide binding. We also examined the quaternary structure of XopAI using analytical ultracentrifugation and revealed a monomeric state of the protein based on the sedimentation coefficient distribution (Figure S9). The results suggested that the protein–peptide interaction found in the crystals was not strong enough to form detectable dimers in the analytical ultracentrifugation experiment.

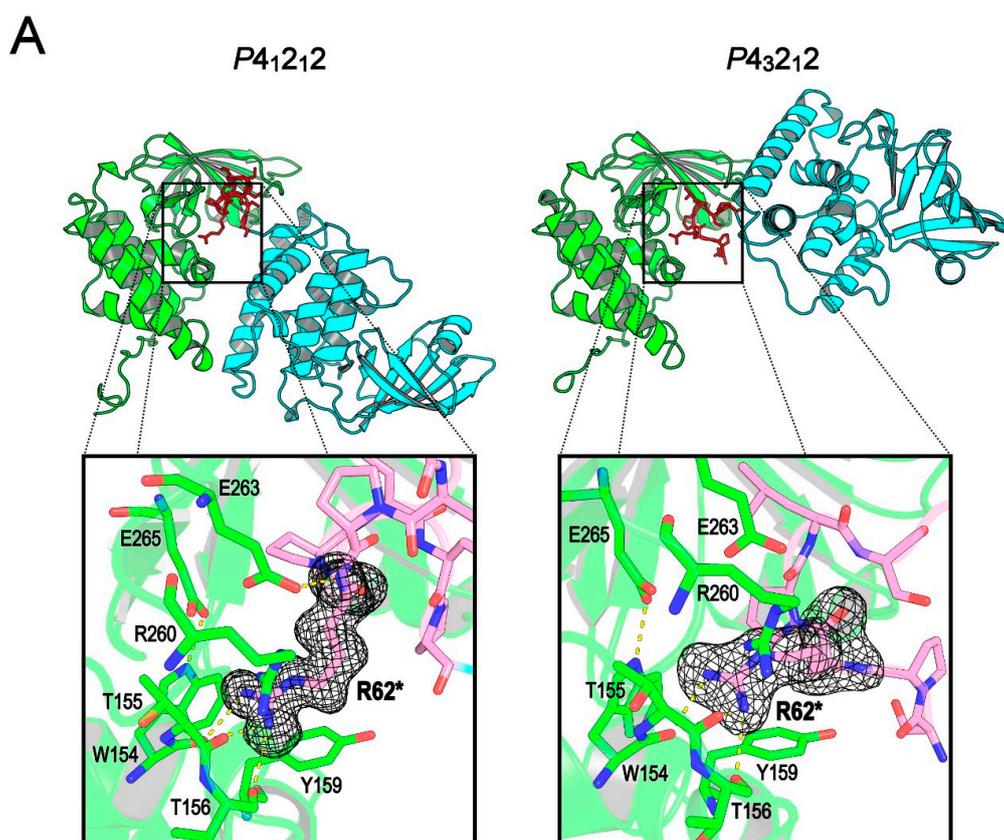
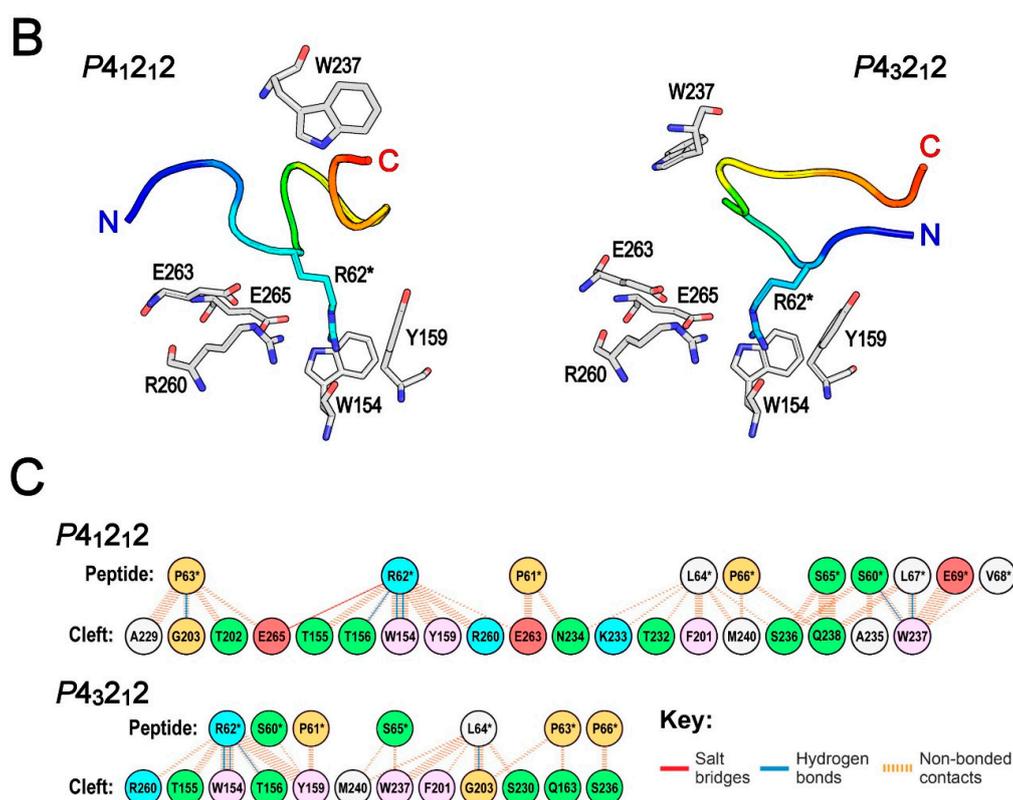


Figure 3. Cont.



**Figure 3.** Two observed Arg peptide-binding modes of XopAI. (A) XopAI interaction found in the *P4<sub>1212</sub>* and *P4<sub>3212</sub>* crystals. The two interacting monomers are shown as ribbon models, and monomers A and B are colored in green and cyan, respectively. The bound N-terminal Arg peptide from monomer B is depicted as a red stick model. The insets show a zoomed-in view of the bound Arg (R62\*). The carbon atoms of the bound Arg peptide are colored in pink; oxygen and nitrogen atoms are colored in red and blue, respectively. The black wire mesh depicts a  $F_o - F_c$  electron density (contoured at  $3\sigma$  level) of R62\*. The residues involved in R62\* recognition are represented as stick models and labeled. Yellow dotted lines show potential hydrogen bonds or salt bridges. (B) The comparison of Arg peptide-binding modes found in the *P4<sub>1212</sub>* and *P4<sub>3212</sub>* crystals. For clarity, the bound Arg peptide is shown as a C $\alpha$  trace and colored in the rainbow scheme; R62\* is highlighted as a cyan stick model. W237 on the phosphate-nicotinamide (PN) loop and the residues involved in R62\* binding are represented as gray stick models and labeled. (C) Residue interactions across the protein-peptide interface. The residues are colored to reflect their chemical properties (positive charged, cyan; negative charged, red; hydrophilic, green; aliphatic, white; aromatic, pink; proline or glycine, light orange). The number of cyan solid lines between any two residues indicates the number of potential hydrogen bonds between them. For non-bonded contacts, the width of the orange striped line is proportional to the number of atomic contacts between the two residues.

To find more empirical data to support our hypothesis, we surveyed the crystal structures of protein-peptide complexes in the PDB, and we estimated their binding energy (Table 3). From the available structural data, we found that when a protein domain accommodates a peptide of six to 10 residues in length, it has a binding energy ranging from  $-3$  to  $-10$  kcal/mol. In fact, it is not easy to derive a linear correlation between the peptide length and the binding energy owing to the diverse protein structures and interaction mechanisms. For example, the SH2 domain that recognizes phosphorylated tyrosine-containing peptides has a high binding energy ( $-7.0$  kcal/mol for a four-residue peptide). On the contrary, the IRS domain, which is another tyrosine-phosphorylated peptide-binding domain, holds a low binding energy ( $-4.9$  kcal/mol) for binding to a peptide with nine residues. By using an empirical force field, we predicted the binding energies for the protein complexes found in the *P4<sub>3212</sub>* and *P4<sub>1212</sub>* crystals to be  $-8.3$  and  $-7.6$  kcal/mol, respectively. These

values are within the binding energy range shown in Table 3, which suggests the interaction between the central cleft and its peptide is energetically reasonable. In addition, these data allow us to propose that the peptide-binding cleft of XopAI functions similar to that of VHS or WD40 domains with respect to their binding energies and interface areas.

**Table 3.** The predicted binding energy between known peptide-binding domains and their peptides.

Domain Name	Binding Energy (kcal/mol)	Number of Bound Residues	Interface Area (Å <sup>2</sup> )	PDB ID
XopAI ( <i>P</i> <sub>4</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub> <sub>2</sub> )	−8.3	12	673	6K93 (this study)
XopAI ( <i>P</i> <sub>4</sub> <sub>3</sub> <sub>2</sub> <sub>1</sub> <sub>2</sub> )	−7.6	7	470	6KLY (this study)
PDZ	−3.7 ± 2.3	4~9 (6) <sup>1</sup>	432 ± 92	1BE9, 1L6O, 1MFG, 1N7F, 1OBX, 1OBZ, 1Q3P, 2I0I, 2QT5, 3DIW, 3LNY
IRS	−4.9 ± 1.3	9	617 ± 20	1UEF, 3ML4
VHS	−5.5 ± 3.2	5~7 (6)	420 ± 45	1JUQ, 1UJK
14-3-3	−5.7 ± 1.7	5~10 (7)	554 ± 122	1QJB, 2BTP, 2C74, 3MHR, 3UBW
WH1	−6.6 ± 2.6	5~6 (6)	315 ± 35	1DDV, 1EVH, 1QC6
Skp1	−6.9 ± 2.3	6~12 (10)	548 ± 151	2AST, 2OVQ, 2P1Q
SH2	−7.0	4	363	1FYR
TPR	−7.0 ± 1.1	5~8 (7)	502 ± 29	1ELR, 1ELW
PID	−7.7 ± 2.7	9~10 (10)	717 ± 28	1AQC, 1M7E, 1NTV
BIR	−8.3 ± 1.7	4~7 (6)	420 ± 54	1JD5, 1SE0, 3D9T
WD40	−8.4 ± 2.3	8~12 (9)	567 ± 44	2CE8, 4ERY, 5IGO, 5IGQ
SH3	−9.4 ± 0.6	9~10 (10)	480 ± 24	1CKA, 1N5Z, 1W70
SPRY	−10.0 ± 0.5	7	352 ± 2	2JK9, 3EMW

<sup>1</sup> Values in parenthesis indicate the average number.

Despite a large difference in interaction orientation, the cleft recognizes an Arg residue of the peptide in both crystal forms similarly. In the insets of Figure 3A, the close-up views clearly show that the Arg from the peptide of another XopAI (for clarity, labeled as R62\* in the figure) inserts deeply into the central cleft. R62\* shows a clear electron density in both crystal forms, and it is surrounded by residues W154, Y159, R260, and E265. Moreover, it has a large surface area (89% of surface area in *P*<sub>4</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals and 83% in *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals), and it is the most conserved residue compared to others in the peptide (Figure 1C and Figure S3). Therefore, hereafter, we specifically name this peptide the “Arg peptide.” Further, the central cleft of XopAI is an Arg peptide-binding cleft. In higher eukaryotes, up to 50% of known interactions between proteins are indeed mediated by peptides [26]. We did not find any gene of the mART family in the bacteria of the genus *Xanthomonas* so far. However, mART members can be found in the bacteria belonging to the genera *Pseudomonas* and *Mycobacterium*. Previous studies demonstrated that two *P. syringae* effectors HopU1 and HopF2 use the mART activity towards different targets to interfere with plant immune signaling [14,27]. As *Xanthomonas* species have no mART, it is possible that they use XopAI to tackle host proteins and hinder the immunity signaling in host cells. XopAI can bind to an Arg peptide-like sequence or an Arg-containing surface patch in the target protein. Subsequently, this interaction blocks the functional surface of the target protein.

In XopAI crystals, we captured two distinct Arg peptide-binding poses. In the *P*<sub>4</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals, the backbone of the Arg peptide adopts a “W”-shaped conformation in the central cleft (Figure 3B). The N-terminus of Arg peptide is placed at the left side of the central cleft, whereas the C-terminus protrudes out of the middle of the cleft. On the other hand, the Arg peptide backbone uses a “U”-shaped conformation in the *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals; both N- and C-termini are located in the front of the middle of the central cleft. The loop between the β-strands, β3 and β4, is called the phosphate-nicotinamide (PN) loop because it binds nicotinamide phosphate and interacts with the target Arg in canonical mARTs [23]. We found that residues in the PN loop of XopAI account for 60% of the peptide-binding interface in the *P*<sub>4</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals, and 42% of that in the *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> crystals. As the PN loop interacts intimately with the Arg peptide, its conformation changes significantly coupled to the peptide conformation (Figure S10A). W237 in the PN loop is particularly noticeable as its position and conformation change drastically

in the two binding poses (Figure 3B). The insets of Figure 3A show that in both binding poses, the cleft uses almost identical residues to bind R62\*. Figure 3B clearly shows that the residues on the lower side of the central cleft—W154, Y159, R260, E263, and E265—maintain very similar sidechain conformations in both poses to interact with R62\*. More interestingly, the coordinates of the central cleft in the XopAI- $\Delta$ N70 crystals show that these residues still have similar sidechain conformations in the absence of the Arg peptide (Figure S10B). W154 and Y159 are located on the putative active site loop flanked by helices  $\alpha$ 4 and  $\alpha$ 5. In canonical mARTs, the active site loop stabilizes the catalytic Glu and the target Arg, and it binds nicotinamide ribose and adenine phosphate. In XopAI, W154 and Y159 form close non-bonded contacts with the R62\* sidechain. R260, E263, and E265 are on the putative ARTT loop; in particular, E263 and E265 are part of the aforementioned [QE]-X-E motif, and E265 is assumed to be the catalytic Glu in mARTs. In XopAI crystals, we observed an attractive interaction between the sidechains of E265 and R62\*. However, it is unclear if E265 is catalytic in XopAI.

Figure 3C summarizes the residue interactions between the cleft and the Arg peptide. In both crystal forms, the interaction to the Arg peptide is provided by residues from the arom-R (residues F201 to G203) and STS (A229 and S230) motifs, as well as those from the active site (W154 to Q163), PN (T232 to M240), and ARTT (R260 to E265) loops. As described earlier (Figure 3B), the active site and ARTT loops contribute to R62\* recognition in both crystal forms. In the  $P_{4_1}2_12$  crystals, the residues of the N- and C-termini interact with the PN loop. However, in the  $P_{4_3}2_12$  crystals, the N-terminus interacts downward with the active site loop, whereas the C-terminus associates with residues from the PN loop. In both crystal forms, most interaction forces are non-bonded contacts through van der Waal's interactions. Specific interactions through the hydrogen bond are consistently found between the R62\* guanidinium group and the residues W154 and T156 from the active site loop. It is noticeable that these hydrogen bonds are established between the R62\* guanidinium group and the backbone oxygen atoms from residues W154 and T156 (Figure 3A inset). Therefore, this interaction may not be disturbed by unwanted sidechain conformations of W154 and T156. The structural analysis website PDBsum [28] suggested a possible salt bridge between the sidechains of R62\* and E265. However, the minimum distances between the R62\* guanidinium group and the E265 sidechain atoms are 3.8 and 4.1 Å in the  $P_{4_1}2_12$  and  $P_{4_3}2_12$  crystals, respectively. It seems that this salt bridge is considerably weak in both crystal forms.

#### 2.4. MD Simulation Study Supports the Protein–Peptide Interaction of XopAI

Although the crystal structures reveal the two peptide-binding modes of the XopAI central cleft (hereafter, referred to as  $P_{4_1}2_12$  and  $P_{4_3}2_12$  modes for short), we wondered if the observed interaction appears in an aqueous solution. To further investigate this protein–peptide interaction, we took crystal structures as starting structures and conducted two sets of 50-ns MD simulations for XopAI complexed with the Arg peptide. Although a small decrease in the binding strength was observed during the first 8 ns in the  $P_{4_3}2_12$  mode, the binding energy between the XopAI central cleft and the peptide did not show any significant change in both courses of the MD simulation (Figure S11A). The root-mean-square deviation (RMSD) value with respect to the crystal structure showed that the bound R62\* in the  $P_{4_3}2_12$  mode altered its conformation in the initial time period of 10 ns (Figure S11B). However, post this period, it held a stable structure in the range of 1–2 Å relative to the starting X-ray structure. The bound R62\* in the  $P_{4_1}2_12$  mode did not change drastically in conformation through the entire simulation. However, we observed that the bound Arg peptide in the  $P_{4_1}2_12$  mode slowly and slightly shifted its backbone structure between 0 to 25 ns (Figure S11C). In summary, these data reveal that the peptide binding of XopAI persisted until the end of the simulation, suggesting that the observed peptide-binding modes in the crystals are sufficiently stable.

We further monitored structural evolution along the last 25 ns of the MD simulation to analyze the possible Arg peptide-binding modes in the aqueous solution. Figure 4A shows that the peptide possesses a similar backbone conformation throughout the simulation. To reflect the structural flexibility in each amino acid, we calculated the root-mean-square fluctuation (RMSF) of all C $\alpha$  atoms

in the peptide relative to the central cleft (Figure 4B). We found that R62\* is the most stable residue in both binding modes; it has a C $\alpha$  RMSF value of 0.83 Å in the  $P4_12_12$  mode and that of 0.86 Å in the  $P4_32_12$  mode. On the other hand, both the N- and C-termini of the peptide show high dynamics in the cleft, especially the N-terminus. When we estimated the energetic contribution of each residue to the binding energy from the MD simulation trajectories (Figure S12A), we noticed that R62\* plays a critical role in binding to the central cleft because it has a high contribution energy. This finding is coherent with the low C $\alpha$  RMSF value of R62\* seen in Figure 4B.

The binding free energy calculation based on the last 25-ns simulation shows that both binding modes have favorable van der Waals, electrostatic, and non-polar solvation free energy values (Table 4). The most favorable contributions to the binding process arise from electrostatic interactions; this finding is in agreement with the proposed roles of R62\* and the negatively charged cleft. However, the concurrent polar solvation free energy contributes negatively to the binding free energy. The total binding free energy for the  $P4_12_12$  mode is calculated as  $-368.5$  kcal/mol, whereas that for the  $P4_32_12$  mode is  $-235.1$  kcal/mol. Thus, the predicted binding free energy for the  $P4_12_12$  mode is higher than that for the  $P4_32_12$  mode. This result is in agreement with the result calculated from the crystal structures in Table 3.

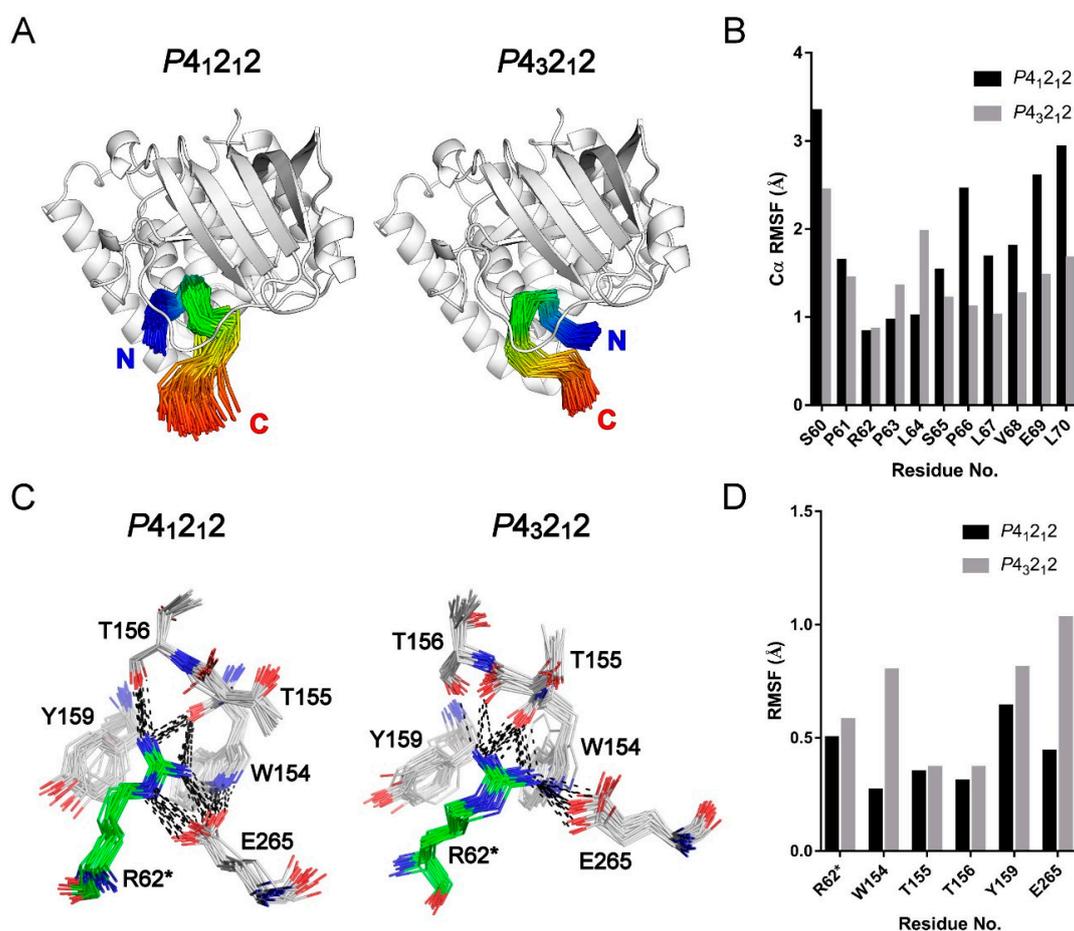
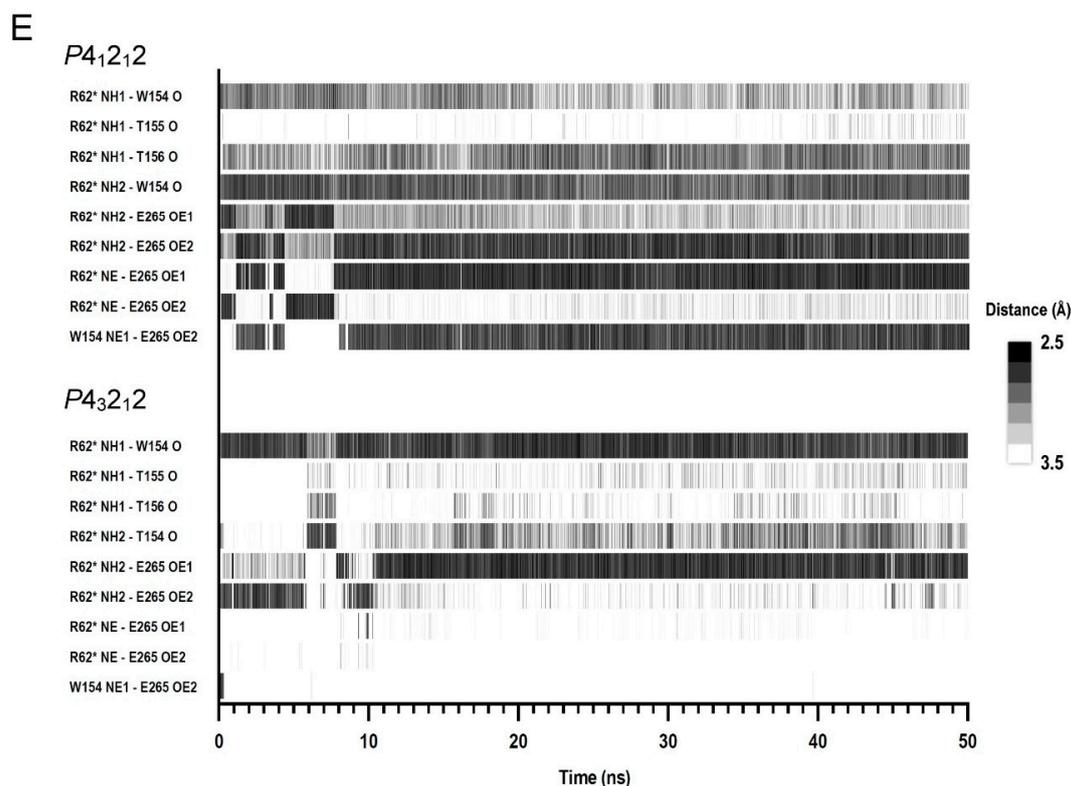


Figure 4. Cont.



**Figure 4.** Molecular dynamics (MD) analysis of the two Arg peptide-binding modes. (A) Ensembles of Arg peptide conformation obtained from the last 25-ns MD simulation. The bound Arg peptide in every snapshot is shown as a C $\alpha$  trace and colored in the rainbow scheme. (B) The root-mean-square fluctuation (RMSF) of the C $\alpha$  atoms in the bound Arg peptide calculated from the results shown in (A). (C) Ensembles of Arg-binding poses during the last 25-ns MD simulation. The carbon atoms of the bound Arg (R62\*) are shown in green, and those of the binding residues are in gray. The oxygen and nitrogen atoms are colored in red and blue, respectively. Potential hydrogen bonds or salt bridges are depicted as black dotted lines. (D) The RMSF of the residues in the R62\*-binding site calculated from the results shown in (C). (E) The timeline of hydrogen bonds formed in the binding site during the 50-ns MD simulation. The hydrogen bonds formed between R62\* and the binding residues are depicted in grayscale lines; the grayscale represents the hydrogen bond distance, which indicates the bond strength.

**Table 4.** Predicted binding free energies of XopAI-peptide complexes.

Complex	$\Delta E_{vdW}$ <sup>1</sup>	$\Delta E_{elec}$	$\Delta G_{polar}$	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
<i>P4<sub>1</sub>2<sub>1</sub>2</i>	$-248.2 \pm 25.2$	$-582.9 \pm 42.2$	$496.4 \pm 42.8$	$-33.8 \pm 2.1$	$-368.5 \pm 40.3$
<i>P4<sub>3</sub>2<sub>1</sub>2</i>	$-173.3 \pm 22.1$	$-539.2 \pm 40.4$	$505.1 \pm 38.9$	$-27.7 \pm 2.0$	$-235.1 \pm 30.9$

<sup>1</sup>  $\Delta E_{vdW}$ ,  $\Delta E_{elec}$ ,  $\Delta G_{polar}$ , and  $\Delta G_{nonpolar}$  are binding energy components of van der Waals, electrostatic, polar and nonpolar solvation energies, respectively.  $\Delta G_{binding}$  is the total binding energy ( $\Delta G_{binding} = \Delta E_{vdW} + \Delta E_{elec} + \Delta G_{polar} + \Delta G_{nonpolar}$ ). The unit of energy is kJ/mol. A positive value indicates the energy component is energetically unfavorable.

While R62\* functions as the key residue responsible for the binding, residues at the R62\*-binding site provide specific and intimate interactions with R62\* (Figure 4C). The R62\* guanidinium group has hydrogen bonds with backbone oxygen atoms from residues W154, T155, and T156. It also interacts with the E265 sidechain via a salt bridge. This explains the observed high contribution energy predicted for E265 (Figure S12B). However, the hydrogen bond network in the *P4<sub>1</sub>2<sub>1</sub>2* mode is more intense than that in the *P4<sub>3</sub>2<sub>1</sub>2* mode. In addition, we found that the coordinates of W154 and E265 in the *P4<sub>3</sub>2<sub>1</sub>2* mode have more fluctuations compared with those in the *P4<sub>1</sub>2<sub>1</sub>2* mode (Figure 4D). The rising

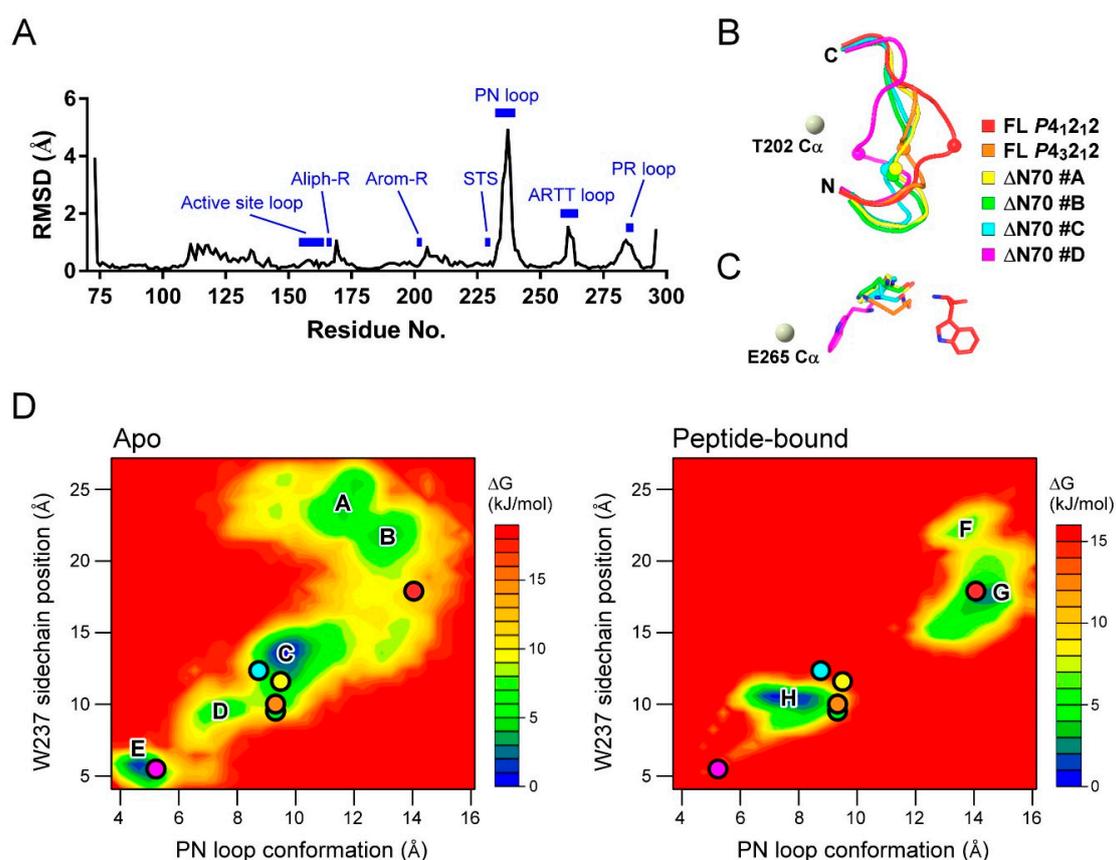
fluctuation in E265 may cause a weak interaction with the R62\* guanidinium group. Similarly, a high RMSF value of W154 may result in loose non-bonded contacts between the sidechains of R62\* and W154. Figure S12B revealed that R260 plays a negative role in Arg peptide binding because it has a repulsive force to the incoming R62\* guanidinium group. Fortunately, E263 along with D157 reside near R260, and they neutralize the positive charge from the R260 sidechain.

Figure 4E shows the time series of the hydrogen bond occurrences at the Arg-binding site. As mentioned earlier, there are slight differences in the hydrogen bond pairs when we compare structures from the crystals with those from the MD simulation. In both binding modes, the relative positions of the R62\* sidechain and backbone oxygen atoms from residues W154, T155, and T156 did not change considerably. Therefore, the hydrogen bond patterns formed by the backbone oxygen atoms are quite consistent during the entire simulation. In the  $P4_12_12$  mode, the backbone oxygen atoms from residues W154 and T156 form three obvious hydrogen bonds with the R62\* guanidinium group (Figure 3A insets and Figure 4C). However, in the  $P4_32_12$  mode, only the hydrogen bond formed between the R62 guanidinium group and the backbone oxygen atoms from residue W154 is significant. Figure 4E shows that during the transition from the crystal to a simulated solution condition, E265 revises its sidechain conformation for 10 ns and then moves closer to the R62\* guanidinium group. As a result, the salt bridge between E265 and R62\* appears to be durable in the course of the 50-ns simulation. Further, the hydrogen bond between W154 and E265 is closely related to their sidechain motions. The MD simulation showed that the W154-E265 hydrogen bond remained in the  $P4_12_12$  mode; however, it was disrupted in the  $P4_32_12$  mode (Figure 4E). Consequently, W154 and E265 in the  $P4_32_12$  mode interact more firmly with R62\* when compared to those in the  $P4_32_12$  mode, as shown in Figure 4C,D.

### 2.5. MD Study for the PN Loop Dynamics of XopAI in Response to the Arg Peptide Binding

To analyze the local dynamics of the apo and peptide-bound states, we calculated the average RMSD among the available XopAI crystal structures (Figure 5A). The RMSD plot shows that three loops in the C-lobe of XopAI show a large displacement in response to Arg peptide binding: the PN loop, the ARTT loop, and the loop between  $\beta$ -strands  $\beta_6$  and  $\beta_7$ . We called the third loop “PR loop” because it contains conserved Pro and Arg (Figure 1C and Figure S3). However, we were unable to infer its function from the current structural information or the knowledge of mARTs. As described earlier, residues on the PN and ARTT loop interact directly with the Arg peptide (Figure 3C). The motion of the PR loop is associated with structural change in the PN loop because they are in close contact with each other (Figure S10A). Among the three loops, the residues of the PN loop have relatively high RMSD values. This finding suggests that the PN loop may be important for the target recognition of XopAI, and it encourages us to explore the influence of peptide binding on the dynamics of the PN loop.

As displayed in Figure 5B, the PN loop adopts different conformations in available crystal structures. The loop structure found in the  $P4_12_12$  peptide-bound state is in a stretched conformation, whereas that in the monomer D of the  $\Delta N70$  crystal structure is in a contracted conformation. Importantly, the PN loop conformation is strongly associated with the volume of the central cleft. For example, the central cleft in the  $P4_12_12$  peptide-bound state has a volume of  $722 \text{ \AA}^3$ ; however, that in the monomer D of the  $\Delta N70$  crystal structure has a volume of  $159 \text{ \AA}^3$ . As W237 is located at the tip of the PN loop, and it has the highest RMSD value, we used the  $C\alpha$  distance between W237 and T202 (the center of the arom-R motif) to differentiate between the conformations of the PN loop. We found that W237 has distinct sidechain positions (Figure 5C): in  $P4_12_12$  peptide-bound state, it protrudes its sidechain, whereas in the monomer D of the  $\Delta N70$  crystal structure, it turns its sidechain inward toward the central cleft. Therefore, we chose the distance between the mass center of the W237 sidechain and the E265  $C\alpha$  atom to represent the sidechain position of W237.



**Figure 5.** MD analysis of PN loop dynamics. (A) Per-residue averaged root-mean-square deviation (RMSD) among available XopAI crystal structures. Those regions of interest are marked as blue bars and labeled. (B) Observed PN loop conformations in the full-length (FL) and  $\Delta$ N70 crystals. The position of W237 on the PN loop is shown as a bead model. As a reference, the location of Thr202 C $\alpha$  atom is depicted as a gray bead. Protein backbones are colored distinctly for the various conformations as follows: FL P4<sub>1</sub>2<sub>1</sub>2 (red), FL P4<sub>3</sub>2<sub>1</sub>2 (orange),  $\Delta$ N70 monomer A (yellow),  $\Delta$ N70 monomer B (green),  $\Delta$ N70 monomer C (cyan), and  $\Delta$ N70 monomer D (magenta). (C) Observed W237 positions and conformations in the crystals; the view of this panel is rotated along the horizontal axis by 90° counterclockwise from that of (B). The carbon atoms of residue W237 are colored distinctly according to (A). The oxygen and nitrogen atoms are colored in red and blue, respectively. As a reference, the location of the E265 C $\alpha$  atom is depicted as a gray bead. (D) energy landscapes depicting the motions of the PN loop and residue W237 in the absence (apo) and presence of Arg peptide (peptide-bound). The depth of the three-dimensional energy landscape indicates the value of the conformation free energy. The free energy is given in kJ/mol and is indicated by the color code shown in the figure. The conformation state of the observed structures is marked with colored circles. The color code for the circles is the same as that in (B,C).

Based on the results of the MD simulation, we calculated the free energy landscapes for the structural transition of the PN loop in the absence and presence of the Arg peptide (Figure 5D). The structures generated by the MD simulation are projected onto two reduced dimensional coordinates—PN loop conformation and W237 sidechain position—and their free energies are estimated by their occurrence probability. Accordingly, it was found that local optimal structures reside at deep “energy wells” on the landscape. The energy landscape for apo structures shows several wells connected by valleys, suggesting that the PN loop can shift between several stable conformations freely in the absence of the Arg peptide. The well C is a big well occupying the central part of the landscape. Earlier, we mentioned that the asymmetric unit in the XopAI- $\Delta$ N70 crystals contains four identical monomers. When mapping these structures onto the energy landscape, we found that three monomers

share a similar PN loop structure near well C, whereas the other one has a distinct loop structure near well E. Unlike apo structures, the peptide-bound structures constitute a discontinuous energy landscape. The PN loop structures in the  $P_{4_1}2_12$  and  $P_{4_3}2_12$  peptide-bound states stay in wells G and H, respectively. However, the landscape shows no obvious valley bridging the two wells. This result implies that there is a high energy barrier for the PN loop structure during the exchange between the two peptide-binding modes.

Although the intermediate structures during the course of the MD simulation could not be captured, we are still curious about whether the two peptide-binding modes can switch from one to another. We employed protein structure morphing [29] to conjecture the conformational transitions between the two binding modes, and then, we estimated the binding free energy of the transition states (Figure S13A). The calculation revealed that each transition state has a considerably negative value of binding energy, which suggests that XopAI can hold the peptide well during the transition between the two binding modes. Furthermore, the predicted path of the conformational transition fits well into the valley on the energy landscape for apo structures (Figure S13B). This result supports the reliability of our structural morphing study, and it paves the way to explore the importance of individual residues on the PN loop.

### 3. Materials and Methods

#### 3.1. Plasmid Constructions

Full-length XopAI from *Xanthomonas axonopodis* pv. *citri* strain XW19 was amplified by PCR using forward primer 5'-ACTGCCATATGGGGGTTATGCACTTCAAAGCCGA-3' (the underlined text represents the *NdeI* site; the bold text, the start codon) and reverse primer 5'-CATG GTCGACCTACGCGATCTGGCTTTGATAAATC-3' (the underline text represents the *SallI* site; the bold text, stop codon). The PCR product was digested with *NdeI* and *SallI*, and then, it was subcloned into a pET28a vector, yielding plasmid pET28a-XopAI. In this vector system, a 6xHis tag followed by a thrombin cleavage site are attached at the N-terminal of XopAI. Similarly, the truncated form of XopAI (XopAI- $\Delta$ N70) was amplified by PCR from the plasmid pET28a-XopAI using the primers: 5'-CTGGTGCCATATGAACACCAGCGATCTGATAAAGC-3' (the underline text represents the *NdeI* site; the bold text, start codon) and 5'-TGGTGTCTCGAGCTACGCGATCTGGCTTTGATAAATC-3' (the underline text represents the *XhoI* site; the bold text, the stop codon). The PCR product was inserted into the p11x vector (a modified p11 vector with more restrictions sites at the multiple cloning site), yielding plasmid pET11x-XopAI- $\Delta$ N70. In this vector system, a 6xHis tag and a thrombin cleavage site are added at the N-terminus of XopAI- $\Delta$ N70. These constructs were verified by DNA sequencing, and then transformed into *E. coli* strain BL21(DE3)RIL<sup>+</sup> competent cells.

#### 3.2. Protein Expression and Purification

The expression and purification procedures for full-length and truncated XopAI proteins are similar. Transformants were cultured at 37 °C in the LB medium containing the antibiotic (kanamycin for pET28a-XopAI and ampicillin for pET11x-XopAI- $\Delta$ N70) and chloramphenicol until OD<sub>600</sub> reached 0.1, and then, it shifted to 20 °C for further incubation. After OD<sub>600</sub> reached 0.4, isopropyl  $\beta$ -D-thiogalactopyranoside was added to reach a final concentration of 0.2 mM, and then, it was incubated for a further 20 h. Bacteria were collected by centrifugation, and suspended in an ice-cold lysis buffer (10 mM imidazole, 300 mM NaCl, 0.25% Triton X-100, 10 mM  $\beta$ -mercaptoethanol, 1 mM PMSF, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0). After the resuspended cells were disrupted using microfluidizer M-110EH (Microfluidics, USA), cell debris was removed by centrifugation at 17,000 $\times$  g for 15 min, and the clear lysate was applied onto a Ni-NTA superflow column (QIAGEN). After washing with a wash buffer (20 mM imidazole, 300 mM NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0) to remove nonspecific-binding proteins, the 6xHis-tagged recombinant protein was eluted with an elution buffer (250 mM imidazole, 300 mM NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0). The eluted protein was dialyzed against a dialysis buffer

(200 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 20 mM Tris-HCl, pH 7.4), and concentrated using a stirred ultrafiltration cell (Millipore). Finally, the full-length XopAI with a concentration of 10 mg mL<sup>-1</sup> was used for crystallization. Due to its poor solubility, XopAI- $\Delta$ N70 was concentrated to 6 mg mL<sup>-1</sup> for crystallization.

### 3.3. Protein Crystallization

XopAI crystals were grown via the hanging-drop vapor-diffusion method at room temperature. The  $P4_32_12$  crystals of full-length XopAI were grown by mixing 1  $\mu$ L of protein solution with 1  $\mu$ L of a crystallization buffer containing 15% (v/v) PEG 400, 5.5% (w/v) PEG 20000, and 50 mM KH<sub>2</sub>PO<sub>4</sub>, pH 8.0. The  $P4_12_12$  crystals of full-length XopAI were grown by mixing 1  $\mu$ L of protein solution with 1  $\mu$ L of 10 mM  $\beta$ -mercaptoethanol and 1  $\mu$ L of crystallization buffer. The buffer is similar to that for the  $P4_32_12$  crystals, and it consists of 16.5% (v/v) PEG 400, 6.5% (w/v) PEG 20000, and 50 mM KH<sub>2</sub>PO<sub>4</sub>, pH 8.0. The crystals of XopAI- $\Delta$ N70 were grown by mixing 2  $\mu$ L of protein solution and 1  $\mu$ L of crystallization buffer containing 10% (v/v) 2-propanol, 20% (w/v) PEG400, and 100 mM HEPES-NaOH, pH 9.5.

### 3.4. Data Collection and Structure Determination

All crystals were flash-cooled in liquid nitrogen. For halide phasing, bromide derivatives were obtained by soaking the  $P4_32_12$  crystals in a crystallization buffer supplemented with 1 M NaBr for 30 s before flash cryocooling. Data collection was carried out at beamlines 13B1, 13C1, and TPS 05A in the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan. The program iMosflm was used for data processing [30]. Phases for XopAI were initially determined by Br-MAD, and automatic model building was performed using the program package PHENIX [31]. Data statistics for the Br-MAD dataset are shown in Table 1. Four bromide ions were found when resolving the phase (Figure S1). After density modification, the overall figure of merit increased from 0.51 to 0.74. More than 50% of residues in one asymmetric unit were traced into the experimental electron density map. The remaining residues were manually built with COOT [32]. The model for the  $P4_32_12$  crystal was refined using a native dataset with a resolution of 2.01 Å (Table 2). The models for the  $P4_12_12$  crystal and the XopAI- $\Delta$ N70 crystals were phased via molecular replacement with phenix.automr from the PHENIX suite [31]. All refinements were performed with phenix.refine from the PHENIX suite. The final models were evaluated using PROCHECK, which showed a good stereochemistry according to the Ramachandran plot [33]. All structure figures were rendered by PyMOL [34]. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID codes 6KLY [XopAI in space group  $P4_32_12$ ], 6K93 [XopAI in space group  $P4_12_12$ ], and 6K94 [XopAI-dN70]).

### 3.5. Sedimentation-Velocity Analytical Ultracentrifugation

The sedimentation velocity experiments of XopAI were carried out using an analytical ultracentrifuge (Model Optima XL-A, Beckman, USA). The protein was prepared in 200 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 20 mM Tris-HCl, pH 7.4. The experiments were performed at 20 °C with a Beckman An-50 Ti rotor speed of 42,000 rpm. The protein samples were continuously monitored by UV absorbance at 280 nm with a time interval of 480 s and a step size of 0.002 cm. Data collected using interference optics were analyzed in terms of the size distribution functions  $c(s)$  using the software SEDFIT [35].

### 3.6. Fluorescence Spectroscopic Assay

The peptide ArgP14aa (SSSPRPLSPLVELN) was synthesized and purified by high-performance liquid chromatography to a purity of 95%, and confirmed by mass spectrometry (MDBio Inc., Taipei, Taiwan). L-arginine (purity  $\geq$  98%) was purchased from Sigma-Aldrich (St. Louis, MO, USA). The buffer of the XopAI-dN70 protein was changed to 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0. In the ligand-induced fluorescence-change experiment, a final concentration of 5  $\mu$ M XopAI-dN70 was incubated with

various concentrations of ligand at 4 °C for 30 min. The tryptophan fluorescence was measured using a fluorescence spectrophotometer (Model F-4500, Hitachi, Japan) equipped with a cuvette of a 1 cm light path. The excitation wavelength was 280 nm, and the emission data were collected between 300 and 400 nm at 4 °C. For the static measurements, all of the measurements were recorded in triplicate.

### 3.7. Bioinformatic Analyses

Sequence similarity searching with BLASTP and TBLASTN identified potential XopAI homologs in GenBank [36]. CD-HIT [37] removed redundant sequences that are 100% identical. Multiple sequence alignments were performed using MAFFT [38] and the corresponding images were generated using the web server ESPript 3.0 [39]. Natively disordered regions of XopAI were predicted using DISOPRED [40], PONDR [41], and SPOT-disorder [42]. Structure-based sequence alignment shown in Figure S4 was produced by Modeller [43]. The structural similarity search was performed with the DALI Server [44]. Amino acid conservation was calculated and plotted onto the protein surface using the ConSurf server [45]. The protein–protein interaction was analyzed by PDBePISA [46] and PDBsum [28]. Figure 3C is modified from an output of PDBsum. The binding free energy of the protein–peptide interaction in available crystal structures was estimated with FoldX [47]. The size of the protein cleft was measured by the CASTp server [48].

### 3.8. MD Simulations

The general procedure is described as follows: First, the starting structure was subjected to energy minimization using GROMACS version 4.6.7 [49] with OPLS-AA force field in an implicit solvent model. The protein was then immersed in an orthorhombic water box in the GROMACS molecular dynamics simulation and the net charge was neutralized by the addition of sodium or chloride ions (at 150 mM salt concentration). Long range electrostatics were handled using the particle mesh Ewald method. The steepest descent energy minimization was used to remove possible bad contacts from the initial structures until energy convergence reached 1000 kJ/(mol·nm). The system was subject to equilibration MD at 300 K and normal pressure constant (1 bar) for 100 ps under the conditions of position restraints for heavy atoms and LINCS constraints. Finally, the production MD was performed under constant pressure and temperature at 1 bar and 300 K, respectively. The bond-angle degrees of freedom from hydrogen atoms were removed using the virtual site algorithm to allow a 2-fs time step. MD trajectories were recorded at 100 ps intervals. Structure snapshots were visualized and analyzed with PyMOL [34] and the built-in utilities of GROMACS.

For the simulation study of the protein–peptide interaction, the starting complex structures ( $P4_12_12$  and  $P4_32_12$  modes) were generated by PyMOL according to the crystallography symmetry operation. Each system was simulated for 50 ns. The interaction energy of the protein–peptide complex was calculated using the *g\_mmpbsa* tool [50]. For the simulation study of the PN loop conformation, the starting structures in the apo form were obtained from the four asymmetric monomers in XopAI- $\Delta N70$  cells, and the monomers in the  $P4_12_12$  and  $P4_32_12$  crystals. Each system was simulated for 50 ns, and the Gibbs free energy landscape was composed using the GROMACS utility *g\_sham*. The energy landscape for the peptide-bound state was constructed from the two 50-ns MD trajectories of the  $P4_12_12$  and  $P4_32_12$  protein–peptide complexes. For the simulation study of the peptide-binding mode transition, the interpolated trajectory between the two binding modes was created using the Yale Morph2 Server [51]. Each structure of the transition states was solvated in a water box, and it was subject to a 100-ps MD simulation with position restraints on the  $C\alpha$  atoms to maintain the specified conformation. The corresponding binding free energy was calculated from 10 snapshots of the resulting trajectory.

## 4. Conclusions

While the primary source of information is experiments, the functions of many proteins are computationally annotated by sequence-based similarity search. However, there is accumulating data

showing that members of an enzyme superfamily exhibit diversity with respect to their substrate preference and the reactions they catalyze [52,53]. These variations stem from incremental mutations on the catalytic, substrate-binding, and allosteric sites. Understanding the molecular basis of these functional variations is important for the accuracy of genome analysis, and it is beneficial for protein design. The structural data presented here revealed that XopAI possesses a peptide-binding domain with a sequence and structural similarity to that of the members of the mART family. However, its cofactor-binding ability seems impaired or altered owing to amino acid substitutions in the putative active site. Owing to a lack of experimental validation, we cannot conclude whether XopAI is also an enzyme.

The discovery of new pathogen effectors and the characterization of their activities provide new insights into how pathogens remodel host cells for their own benefit. The studies of effectors also offer opportunities for the development of tools to explore host cell biology in the absence of disease. In this study, we presented the crystal structure of XopAI; it revealed that XopAI shares structural similarity to mARTs; however, it may have a distinct function from that of mARTs. The structural analysis and the MD simulation study uncovered the molecular basis of interaction between the XopAI peptide-binding domain and the Arg peptide. Furthermore, these data suggested that XopAI evolved to bind an Arg peptide-like sequence or an Arg-containing surface patch in its target protein, which is important for disease progression or plant immunity. The next challenge is to identify plant target protein(s) of XopAI, which are crucial to elucidate its molecular function.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/20/5085/s1>.

**Author Contributions:** Conceptualization, J.-H.L., D.-W.H. and M.-H.H.; Data curation, J.-H.L.; Formal analysis, D.-W.H., Y.-H.L., Y.-P.L., and Y.-R.T.; Funding acquisition, J.-H.L.; Investigation, J.-H.L., J.-Y.Y., D.-W.H., Y.-H.L., Y.-P.L., and Y.-R.T.; Methodology, J.-H.L., D.-W.H. and M.-H.H.; Project administration, J.-H.L.; Resources, J.-Y.Y.; Software, J.-H.L.; Supervision, J.-H.L.; Validation, J.-H.L.; Writing—original draft, J.-H.L.; Writing—review and editing, J.-H.L. and M.-H.H.

**Funding:** This work was supported by Ministry of Science and Technology, Taiwan, R.O.C. (MOST 106-2313-B-005-029 to J.-H.L.).

**Acknowledgments:** We thank Yu-Ting Chen, Wei-Ming Leu, Hung-Jen Liu, and Menghsiao Meng for their discussions and comments on the manuscript and for their support. We also thank the experimental facility and the technical services provided by the “Synchrotron Radiation Protein Crystallography Facility of the National Core Facility Program for Biotechnology, Ministry of Science and Technology” and the “National Synchrotron Radiation Research Center,” a national user facility supported by the Ministry of Science and Technology, Taiwan.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

ARTT	ADP-ribosyl-turn-turn
Br-MAD	bromide multiple-wavelength anomalous diffraction
mART	mono-ADP-ribosyltransferase
MD	molecular dynamics
PN	phosphate-nicotinamide
RMSD	root-mean-square deviation
RMSF	root-mean-square fluctuation
<i>Xac</i>	<i>Xanthomonas axonopodis</i> pv. <i>citri</i>

## References

- Graham, J.H.; Gottwald, T.R.; Cubero, J.; Achor, D.S. *Xanthomonas axonopodis* pv. *citri*: Factors affecting successful eradication of citrus canker. *Mol. Plant Pathol.* **2004**, *5*, 1–15. [PubMed]
- Buttner, D.; He, S.Y. Type III protein secretion in plant pathogenic bacteria. *Plant Physiol.* **2009**, *150*, 1656–1664. [CrossRef] [PubMed]
- He, S.Y. Type III protein secretion systems in plant and animal pathogenic bacteria. *Annu. Rev. Phytopathol.* **1998**, *36*, 363–392. [CrossRef] [PubMed]

4. Dunger, G.; Arabolaza, A.L.; Gottig, N.; Orellano, E.G.; Ottado, J. Participation of *Xanthomonas axonopodis* pv. *citri* hrp cluster in citrus canker and nonhost plant responses. *Plant Pathol.* **2005**, *54*, 781–788.
5. Swarup, S.; Yang, Y.; Kingsley, M.T.; Gabriel, D.W. An *Xanthomonas citri* pathogenicity gene, pthA, pleiotropically encodes gratuitous avirulence on nonhosts. *Mol. Plant Microbe Interact.* **1992**, *5*, 204–213. [[CrossRef](#)] [[PubMed](#)]
6. da Silva, R.A.; Pereira, L.M.; Silveira, M.C.; Jardim, R.; de Miranda, A.B. Mining of potential drug targets through the identification of essential and analogous enzymes in the genomes of pathogens of *Glycine max*, *Zea mays* and *Solanum lycopersicum*. *PLoS ONE* **2018**, *13*, e0197511. [[CrossRef](#)] [[PubMed](#)]
7. Moreira, L.M.; Almeida, N.F., Jr.; Potnis, N.; Digiampietri, L.A.; Adi, S.S.; Bortolossi, J.C.; da Silva, A.C.; da Silva, A.M.; de Moraes, F.E.; de Oliveira, J.C.; et al. Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of *Xanthomonas fuscans* subsp. *aurantifolii*. *BMC Genom.* **2010**, *11*, 238. [[CrossRef](#)]
8. Ferreira, R.M.; de Oliveira, A.C.; Moreira, L.M.; Belasque, J., Jr.; Gourbeyre, E.; Siguier, P.; Ferro, M.I.; Ferro, J.A.; Chandler, M.; Varani, A.M. A TALE of transposition: Tn3-like transposons play a major role in the spread of pathogenicity determinants of *Xanthomonas citri* and other *xanthomonads*. *MBio* **2015**, *6*, e02505-14. [[CrossRef](#)]
9. White, F.F.; Potnis, N.; Jones, J.B.; Koebnik, R. The type III effectors of *Xanthomonas*. *Mol. Plant Pathol.* **2009**, *10*, 749–766. [[CrossRef](#)]
10. Thieme, F.; Szczesny, R.; Urban, A.; Kirchner, O.; Hause, G.; Bonas, U. New type III effectors from *Xanthomonas campestris* pv. *vesicatoria* trigger plant reactions dependent on a conserved N-myristoylation motif. *Mol. Plant Microbe Interact.* **2007**, *20*, 1250–1261.
11. da Silva, A.C.; Ferro, J.A.; Reinach, F.C.; Farah, C.S.; Furlan, L.R.; Quaggio, R.B.; Monteiro-Vitorello, C.B.; Van Sluys, M.A.; Almeida, N.F.; Alves, L.M.; et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **2002**, *417*, 459–463. [[CrossRef](#)] [[PubMed](#)]
12. Deng, Q.; Barbieri, J.T. Molecular mechanisms of the cytotoxicity of ADP-ribosylating toxins. *Annu. Rev. Microbiol.* **2008**, *62*, 271–288. [[CrossRef](#)] [[PubMed](#)]
13. Simon, N.C.; Aktories, K.; Barbieri, J.T. Novel bacterial ADP-ribosylating toxins: Structure and function. *Nat. Rev. Microbiol.* **2014**, *12*, 599–611. [[CrossRef](#)] [[PubMed](#)]
14. Fu, Z.Q.; Guo, M.; Jeong, B.R.; Tian, F.; Elthon, T.E.; Cerny, R.L.; Staiger, D.; Alfano, J.R. A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. *Nature* **2007**, *447*, 284–288. [[CrossRef](#)] [[PubMed](#)]
15. Dauter, M.; Dauter, Z. Phase determination using halide ions. *Methods Mol. Biol.* **2007**, *364*, 149–158. [[PubMed](#)]
16. Schaad, N.W.; Postnikova, E.; Sechler, A.; Clafin, L.E.; Vidaver, A.K.; Jones, J.B.; Agarkova, I.; Ignatov, A.; Dickstein, E.; Ramundo, B.A. Reclassification of subspecies of *Acidovorax avenae* as *A. avenae* (Manns 1905) emend., *A. cattleyae* (Pavarino, 1911) comb. nov., *A. citrulli* Schaad et al., 1978) comb. nov., and proposal of *A. oryzae* sp. nov. *Syst. Appl. Microbiol.* **2008**, *31*, 434–446. [[CrossRef](#)]
17. Hoppener-Ogawa, S.; de Boer, W.; Leveau, J.H.; van Veen, J.A.; de Brandt, E.; Vanlaere, E.; Sutton, H.; Dare, D.J.; Vandamme, P. *Collimonas arenae* sp. nov. and *Collimonas pratensis* sp. nov., isolated from (semi-)natural grassland soils. *Int J. Syst. Evol. Microbiol.* **2008**, *58* (Pt 2), 414–419. [[CrossRef](#)]
18. Jeong, B.R.; Lin, Y.; Joe, A.; Guo, M.; Korneli, C.; Yang, H.; Wang, P.; Yu, M.; Cerny, R.L.; Staiger, D.; et al. Structure function analysis of an ADP-ribosyltransferase type III effector and its RNA-binding target in plant immunity. *J. Biol. Chem.* **2011**, *286*, 43272–43281. [[CrossRef](#)]
19. Ting, S.Y.; Bosch, D.E.; Mangiameli, S.M.; Radey, M.C.; Huang, S.; Park, Y.J.; Kelly, K.A.; Filip, S.K.; Goo, Y.A.; Eng, J.K.; et al. Bifunctional immunity proteins protect bacteria against FtsZ-targeting ADP-ribosylating toxins. *Cell* **2018**, *175*, 1380–1392. [[CrossRef](#)]
20. Mueller-Dieckmann, C.; Ritter, H.; Haag, F.; Koch-Nolte, F.; Schulz, G.E. Structure of the ecto-ADP-ribosyl transferase ART2.2 from rat. *J. Mol. Biol.* **2002**, *322*, 687–696. [[CrossRef](#)]
21. Karlberg, T.; Hornyak, P.; Pinto, A.F.; Milanova, S.; Ebrahimi, M.; Lindberg, M.; Pullen, N.; Nordstrom, A.; Loverli, E.; Caraballo, R.; et al. 14-3-3 proteins activate *Pseudomonas* exotoxins-S and -T by chaperoning a hydrophobic surface. *Nat. Commun.* **2018**, *9*, 3785. [[CrossRef](#)] [[PubMed](#)]
22. Hottiger, M.O.; Hassa, P.O.; Luscher, B.; Schuler, H.; Koch-Nolte, F. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.* **2010**, *35*, 208–219. [[CrossRef](#)] [[PubMed](#)]

23. Fieldhouse, R.J.; Turgeon, Z.; White, D.; Merrill, A.R. Cholera- and anthrax-like toxins are among several new ADP-ribosyltransferases. *PLoS Comput. Biol.* **2010**, *6*, e1001029. [[CrossRef](#)] [[PubMed](#)]
24. Holbourn, K.P.; Shone, C.C.; Acharya, K.R. A family of killer toxins. Exploring the mechanism of ADP-ribosylating toxins. *FEBS* **2006**, *273*, 4579–4593. [[CrossRef](#)] [[PubMed](#)]
25. Lo Conte, L.; Chothia, C.; Janin, J. The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **1999**, *285*, 2177–2198. [[CrossRef](#)]
26. Petsalaki, E.; Russell, R.B. Peptide-mediated interactions in biological systems: New discoveries and applications. *Curr. Opin. Biotechnol.* **2008**, *19*, 344–350. [[CrossRef](#)] [[PubMed](#)]
27. Wilton, M.; Subramaniam, R.; Elmore, J.; Felsensteiner, C.; Coaker, G.; Desveaux, D. The type III effector HopF2<sub>Pto</sub> targets *Arabidopsis* RIN4 protein to promote *Pseudomonas syringae* virulence. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2349–2354. [[CrossRef](#)]
28. Laskowski, R.A.; Jablonska, J.; Pravda, L.; Varekova, R.S.; Thornton, J.M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **2018**, *27*, 129–134. [[CrossRef](#)]
29. Weiss, D.R.; Levitt, M. Can morphing methods predict intermediate structures? *J. Mol. Biol.* **2009**, *385*, 665–674. [[CrossRef](#)]
30. Powell, H.R.; Batty, T.G.G.; Kontogiannis, L.; Johnson, O.; Leslie, A.G.W. Integrating macromolecular X-ray diffraction data with the graphical user interface iMosflm. *Nat. Protoc.* **2017**, *12*, 1310–1325. [[CrossRef](#)]
31. Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (Pt 2), 213–221. [[CrossRef](#)]
32. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
33. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [[CrossRef](#)]
34. DeLano, W.L. PyMOL: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **2002**, *40*, 82–92.
35. Schuck, P.; Perugini, M.A.; Gonzales, N.R.; Howlett, G.J.; Schubert, D. Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. *Biophys J.* **2002**, *82*, 1096–1111. [[CrossRef](#)]
36. Sayers, E.W.; Cavanaugh, M.; Clark, K.; Ostell, J.; Pruitt, K.D.; Karsch-Mizrachi, I. GenBank. *Nucleic Acids Res.* **2019**, *47*, D94–D99. [[CrossRef](#)]
37. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinf* **2012**, *28*, 3150–3152. [[CrossRef](#)]
38. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
39. Gouet, P.; Robert, X.; Courcelle, E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **2003**, *31*, 3320–3323. [[CrossRef](#)]
40. Ward, J.J.; McGuffin, L.J.; Bryson, K.; Buxton, B.F.; Jones, D.T. The DISOPRED server for the prediction of protein disorder. *Bioinf* **2004**, *20*, 2138–2139. [[CrossRef](#)]
41. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
42. Hanson, J.; Yang, Y.; Paliwal, K.; Zhou, Y. Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks. *Bioinf* **2017**, *33*, 685–692. [[CrossRef](#)] [[PubMed](#)]
43. Eswar, N.; Webb, B.; Marti-Renom, M.A.; Madhusudhan, M.S.; Eramian, D.; Shen, M.Y.; Pieper, U.; Sali, A. Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinf* **2006**, *5*. [[CrossRef](#)]
44. Holm, L.; Laakso, L.M. Dali server update. *Nucleic Acids Res.* **2016**, *44*. [[CrossRef](#)] [[PubMed](#)]
45. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, 344–350. [[CrossRef](#)]
46. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)]

47. Delgado, J.; Radusky, L.G.; Cianferoni, D.; Serrano, L. FoldX 5.0: Working with RNA, small molecules and a new graphical interface. *Bioinf* **2019**. [[CrossRef](#)]
48. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, 363–367. [[CrossRef](#)]
49. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; van der Spoel, D.; et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinf* **2013**, *29*, 845–854. [[CrossRef](#)]
50. Kumari, R.; Kumar, R.; Open Source Drug Discovery, C.; Lynn, A. g\_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962. [[CrossRef](#)]
51. Krebs, W.G.; Gerstein, M. The morph server: A standardized system for analyzing and visualizing macromolecular motions in a database framework. *Nucleic Acids Res.* **2000**, *28*, 1665–1675. [[CrossRef](#)] [[PubMed](#)]
52. Gerlt, J.A.; Babbitt, P.C.; Rayment, I. Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. *Arch. Biochem. Biophys.* **2005**, *433*, 59–70. [[CrossRef](#)] [[PubMed](#)]
53. Nardini, M.; Dijkstra, B.W. Alpha/beta hydrolase fold enzymes: The family keeps growing. *Curr. Opin. Struct. Biol.* **1999**, *9*, 732–737. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).