Article

Crystal structure-based exploration of arginine-containing peptide binding in the ADP-ribosyltransferase domain of the type III effector XopAI protein

Jyung-Hurng Liu ^{1, 2, 3, 4, *}, Jun-Yi Yang ^{3, 5}, Duen-Wei Hsu ⁶, Yi-Hua Lai ², Yun-Pei Li ¹, Yi-Rung Tsai ¹, and Ming-Hon Hou ^{1, 2, 3, 4}

- ¹ Institute of Genomics and Bioinformatics, National Chung Hsing University (NCHU), Taichung, Taiwan 40227, R.O.C.; jhliu@nchu.edu.tw (J.-H.L.); yunk293@gmail.com (Y.-P.L.); stave0972@gmail.com (Y.-R.T.); mhho@nchu.edu.tw (M.-H.H.)
- ² Department of Life Science, NCHU, Taichung, Taiwan 40227, R.O.C.; jhliu@nchu.edu.tw (J.-H.L.); hbm486426@gmail.com (Y.-H.L.); mhho@nchu.edu.tw (M.-H.H.)
- ³ Graduate Institute of Biotechnology, NCHU, Taichung, Taiwan 40227, R.O.C.; jhliu@nchu.edu.tw (J.-H.L.); jyang@nchu.edu.tw (J.-Y.Y.); mhho@nchu.edu.tw (M.-H.H.)
- ⁴ PhD program in Medical Biotechnology, NCHU, Taichung, Taiwan 40227, R.O.C.; jhliu@nchu.edu.tw (J.-H.L.); mhho@nchu.edu.tw (M.-H.H.)
- ⁵ Graduate Institute of Biochemistry, NCHU, Taichung, Taiwan 40227, R.O.C.; jyang@nchu.edu.tw (J.-Y.Y.)
- ⁶ Department of Biotechnology, National Kaohsiung Normal University, Kaohsiung, Taiwan 80201, R.O.C.; dwhsu@nknucc.nknu.edu.tw (D.-W.H.)
- * Correspondence: jhliu@nchu.edu.tw; Tel.: +886-4-22840338 (J.-H.L.)

Contents	Pages
Figure S1. Bromide ion-binding sites in XopAI crystals.	S2
Figure S2. Computational prediction of disordered regions in XopAI.	S3
Figure S3. Protein alignment of XopAI homologs.	S4–S5
Figure S4. Protein alignment of XopAI and some known mART proteins.	S6
Figure S5. Structural comparison between XopAI and known mART proteins.	S7
Figure S6. ConSurf analysis for XopAI and HopU1.	S8
Figure S7. Crystal packing of the full-length XopAI and XopAI- Δ N70 proteins.	S9
Figure S8. Intrinsic tryptophan fluorescence studies showing the peptide-binding	S10
ability of the XopAI- Δ N70 protein.	
Figure S9. Continuous sedimentation coefficient distribution of the full-length	S11
XopAI protein.	
Figure S10. Comparison of XopAI structures in free and in peptide-bound states.	S12
Figure S11. Comparative plots from MD analysis of the XopAI-peptide complexes	S13
Figure S12. Predicted binding free energy contribution per residue of the protein-	S14
peptide interaction during the last 25-ns MD simulation.	
Figure S13. Conformational transition between two Arg peptide-binding modes.	S15



Figure S1. Bromide ion-binding sites in XopAI crystals. **A**, the anomalous difference map (the black mesh) contoured at 2.5 σ that shows four bromide sites (magenta spheres) in the *P*4₃2₁2 crystal. The XopAI structure is shown as a ribbon model and colored in the rainbow scheme. **B**, close-up views of bromide ion-binding sites. Residues around the bromide ion are shown as stick models and labeled. Carbon, oxygen and nitrogen atoms are colored in green, red and blue, respectively. The potential hydrogen bonds are depicted as black dashed lines.



Figure S2. Computational prediction of disordered regions of XopAI. The prediction results were obtained from DISOPRED [40], PONDR [41], and SPOT-disorder [42].

TOPAT	1 10	20	30	4.0	50	60	70
XopAI Xcg Xc_WF7 Xcb Xcc_Xcc29 Xcc_jx6 Xc_WF1 Xv_LM159 Xfa_FDC1609 Xfa_FDC1609 Xca Xca Xc_WF2 Xaj Ac Aaa Cp	MGLCTSKPSV MGLCTSKPSV MGLCTSKPSV MGLCTSKPSV MGLCTSKPSV MGLCSKPSV MGLCVSKQS. MGLCVSKQS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS. MGLCVSKPS.	VGSPVAGSPEHYLT VGSPVAGSPEHYLT VGSPVAGSPEHYLT VGSPVAGSPEHYLT VGSPVAGSPEHYLT VGSPVAGSPEHYLT VAGSPEHYAA VAGSPEHYAA VAGSPEHYAA VAGSPEHYAA VAGSPEHYAA VAGSPEHYAA VAGSPEHYAA ASGSSYHSAY ASGSSYHPAY LTSQENYYAD	H TAEQ TTP ST H ATAQATP ST RAKAQATP ST S	PSSPEAPMSPS PSSPEAPMSPS PSSPEAPMSPS PSSPEAPMSPS PSSPEAPMSPS PSSPEAPMSPS PSSPEAPMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS PSSPEASMSPS	LHGLAALGSP LHGLVALGSP LHGLVALGSP LHGLVALGSS LHGLVALGSS LHGLVALGSS LHGLVALGSS LYGLASLGSP LYGLASLGSP LYGLASLGSP LYGLASLGSP LYGLASLGSP PEGLSPRSSP PEGLSPRSSP PHGLLARTSQ	RASSS PR.I RASSS PR.I RASSS PR.I RASSS PR.I RASSS PR.I RASSS PR.I GTSAS PR.I RASSS PR.I RASSS PR.I RASSS PR.I RASSS PR.I RASSS PR.I RASSS RRI SNQAGNLAGNLAGRA R62 <	SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN SISPLVEIN AQVRLPTIS SIDLLRSIS
XopAI	2000000	α1 00000000 0 0	α2 200000000000		α3 200000000000 20 13	0000000	α4 22222
XopAI Xcg Xc_WP7 Xcb Xcc_Cc29 Xcc_jx6 Xc_WP1 Xv_LM159 Xfa_FDC1609 Xca Xc_WP2 Xca Ac Aaa Cp	TSDI IKQKKQ TSDI IKQKKQ TSDI IKQKKQ TSDI IKQKKQ TSDI IKQKKQ TSDI IKQKKQ TSDI VKQKKQ TSDI VKQKKQ	LWQRVQHDGAQFRS LWQRVQHDGAQFRS LWQRVQHDGAQFRS LWQRVQHDGAQFRS LWQRVQHDGAQFRS LWQRVQHDGAQFRS LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LWQRIQHDGSQFRT LVQRVLSRGANFRS LVQRVLSQGANFRS	TPEERKQFKTAITPEERKQFKTAITPEERKQFKTAITPEERKQFKTAITPEERKQFKTAITPEERKQFKTAITPEERKQFKTAITPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKTAFTPEERKQFKAF	ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQYRPE ITINGEQSRPE ITINGEQSRQO ITIPSEQSRQO ITIPSEQSRQO ITIPSEQSRQO ITIPSQSCQO ITIPSEQSRQO ITIPSEQSRQO ITIPSQSCQS IXKRTGDCVSSE IKKRTGDCVSSE	QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNGMMQR QQRWNEMKQR QQRWNEMKQR QQRWNEMKQR QQRWNEMKQR QQRWNEMKQR QQRWNEMKQR QQRWNEMKQR	$\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{W} \mathbf{H} = \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{M} \mathbf{K} \mathbf{R} \mathbf{N} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{R} \mathbf{I} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{R} \mathbf{I} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{R} \mathbf{I} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{R} \mathbf{I} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{L} \bigcirc \mathbf{L} \mathbf{I} \ge \mathbf{L}$ $\mathbf{L} \bigcirc \mathbf{L} \mathbf{L} \mathbf{K} \mathbf{R} \mathbf{R} \mathbf{H} \ge \mathbf{E} \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} \bigcirc \mathbf{L} \mathbf{K} \mathbf{R} \mathbf{U} + \mathbf{U} \ge \mathbf{I}$ $\mathbf{M} \ge \mathbf{L} $	XYMATEDLV XYMATEDLV XYMATEDLV XYMATEDLV XYMATEDLV XYMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV HMATEDLV 2000000000000000000000000000000000000
XopAI	00000		<u>α5</u>		β1	α6	β2
1	50 1	60 170	180	190	200	210	220
1 XopAI Xc_WP7 Xc_WP4 Xcb Xcc_Xcc29 Xcc_jx6 Xc_WP1 Xv_LM159 Xfa_1566 Xfa_FDC1609 Xca Xc_WP2 Xaj Ac Aaa Cp	50 ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTT.DD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD ALQAWTADDD	60 170 YEVVQDVLBKEARP YEVVQDVLBKEARP YEVVQDVLEKEARP YEVVQDVLEKEARP YEVVQDVLEKEARP YEVVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YEVQDVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUQNVLEKESP YQUQNVLEKDESP YQUQNVLEKDESP YQUQNVLEDGHTP YQUQNVLEMEGSP YQUPUNER YQUPUNER YQUPUNER YQUPUNER YQUPUNER YQUPUNER YQUPUNER YQUPUNER YYEBGS YYEBGS YYEBGS YYEBGS	180 TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII SVEGLAYAKCLI SVEGLAYAKCLI TABALAWAKALI	190 SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS	200 (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGENQ	210 LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE SDAWVQARHAE SDAWVQARHAE SDAWVQARHAE	220 RSITTDRRF RSITTDRRF RSITTDRF RSITTDRF RSITTDRF GSITTDRF GSITTDRF GSITTDRF GSITTDRF GSITTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRF GRF GRF GRF GRF GQR TVNRF GQR TVNRF
1 XopAI Xc_WP7 Xc_WP4 Xcb Xcc_Xcc29 Xcc_jx6 Xc_WP1 Xv_LM159 Xfa_1566 Xfa_FDC1609 Xca Xc_WP2 Xaj Ac Aaa Cp	50 1 ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTADDD ALQA WTADD ALQA WTA	60 170 YEVVQDVLBKEARP YEVVQDVLBKEARP YEVVQDVLEKEARP YEVQDVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUQNVLEKEARP YQUVQNVLEKEARP YQUVQNVLEKESP YQUVQNVLEKDESP YQUVQNVLEDGHTP YQUVQNVLEDGHTP YQUVQNVLEDGHTP YKTLEVEARP YQUVQNVLEKDSP YQUVQNVLENDGHTP YKTLEVEARP YS YQUONVLEK YY YY <th>180 TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII SVEGLAYAKCLI SVEGLAYAKCLI TAEALAWAKALI</th> <th>190 SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS</th> <th>2 0 0 (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEN</th> <th>210 LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE SDAWVQARHAE SDAWVQARHAE SDAWVQARHAE</th> <th>220 RSITTDRRF RSITTDRRF RSITTDRF RSITTDRF RSITTDRF GSITTDRF GSITTDRF GSITTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRF GRF GRF GRF GRF GRF GRF G</th>	180 TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII SVEGLAYAKCLI SVEGLAYAKCLI TAEALAWAKALI	190 SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS	2 0 0 (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEN	210 LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE SDAWVQARHAE SDAWVQARHAE SDAWVQARHAE	220 RSITTDRRF RSITTDRRF RSITTDRF RSITTDRF RSITTDRF GSITTDRF GSITTDRF GSITTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRETTDRF GRF GRF GRF GRF GRF GRF GRF G
1 XopAI Xcg Xc_WP7 Xc_WP4 Xcb Xcc_Xc29 Xcc_jx6 Xc_WP1 Xv_IMI59 Xfa_1566 Xfa_FDC1609 Xca Xc_WP2 Xaj Ac Aaa Cp	50 1 ALQA WTT.DD ALQA WTT.DD ALQA WTT.DD ALQA WTADDD ALQA WTADDD ALQA WTADDD ALQA WTADD ALQA WTADDD ALQA WTADD ALQA WTADDD ALQA WTADD ALQA WTADDD ALQA WTADDD ALQA WTADD	60 170 YEVVQDVLBKEARP YQIVQDVLBKEATP YQIVQNVLBKYESP YQIVQNVLBKYESP YQIVQNVLBKDESP YQIVQDVLBCDGHTP YKTLEDVEHGSVV Active Aliph-R 159 site loop E167 β4 240 250	180 TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII TAHGLAFAKCII SVEGLAYAKCLI SVEGLAYAKCLI SVEGLAYAKCLI	190 SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPEYSS SALHSLPSS SALHSLSS SALSS SALHSLSS SALHSLSS SALSS	2 0 0 (QGTVFTG E D Q (QGTVFTG E N Q (QGTVFTG E D Q HGTVFTG E D Q HGTVYTG E N P Arom-R T202 β6 2 8 0	210 LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LPDWVSERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE LTNWVAERYQE SDAWVQARHSE SDAWVQARHSE SDAWVQARHSE SDAWVQARHSE SDAWVQARHSE	220 SITTDRRF SITTDRRF SITTDRRF SITTDRRF SITTDRRF SITTDRRF GSITTDRRF GSITTDRRF GRETTDRRF GRETTDRRF GRETTDRRF GRETVTNLRF GQTVVNLRF GQR
1 XopAI Xcg Xc_WP7 Xc_WP4 Xcb Xcc_Xcc29 Xcc_jx6 Xc_WP1 Xv_LM159 Xfa_1566 Xfa_FDC1609 Xca Xc_WP2 Xaj Ac Aaa Cp XopAI XopAI XopAI Xcg Xc_WP7 Xc_WP7 Xc_WP7 Xc_WP7 Xc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP2 Xaj Ac Aaa Cp XopAI XopAI XopAI XopAI Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP7 Xcc_WP2 Xaj Ac Aaa Cp XopAI XopAI Xcc_WP7 Xcc_WP2 Xaa Xcc_WP3	50 1 1 1 1 1 1 2 1 1 1 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 1 2 3 1 1 1 2 3 1 1 2 3 1 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 2 3 1 1 1 1 2 3 1 1 1 1 1 1 1 1	60 170 YEVVQDVLEKEARP YEVVQDVLEKEARP YEVVQDVLEKEARP YEVVQDVLEKEARP YEVQDVLEKEARP YQUQVDVLEKEARP YQUQVDVLEKEAPP YQUVQNVLEKESP YQUVQNVLEKYESP YQUVQNVLEKSNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESNSTT YQGMAVEWESN	180 TAHGLAFAKCII GKRISMFSERPN GKRISMFSERPN GKRISMFSERPN GKRISMFSERPN GKRISMFSERPN GKRISMFSERPI GKRISMFSERPI GKRISMFSERPI GKRISMFSERQI	190 SALHSLPEEYSS SA	2 0 0 (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGEDQ (QGTVFTGENC (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENQ (QGTVFTGENC (QGTVFTGENQ (QGTVFTGENC (QGTVFTGENQ (QGTVFTCENC (QTNR (QUTRIENC (QVTRIENC (QVTRIENC (QVTR	210 210 210 210 200 210 210 210	2 2 0 RS I TTDRRF RS I TTDRRF RS I TTDRRF RS I TTDRRF RS I TTDRRF RS I TTDRRF GS I TTDRRF GS I TTDRRF GRE TTDRF GRE TTDRF GRE TTMRF GRE TVTN LRF GQR TVNRF GQR TVNRF A A A A A A A A A A A A A

Figure S3. Protein alignment of XopAI homologs. Coloring and labels in this alignment are consistent with

Figure 1C. The following bacteria strains were analyzed: XopAI (*Xanthomonas axonopodis* pv. *citri*, GenBank accession no.: WP_011052119, this study), Xcg (*X. citri* pv. *glycines*, CP017188), Xc_WP7 (*X. citri*, WP_076605129), Xc_WP4 (*X. citri*, WP_040244769), Xcb (*X. citri* pv. *bilvae*, CEJ46851), Xcc_Xcc29 (*X. citri* pv. *citri* strain Xcc29-1, CP023661), Xcc_jx6 (*X. citri* pv. *citri* strain jx-6, CP011827), Xc_WP1 (*X. cynarae*, WP_104591584), Xv_LM159 (*X. vesicatoria* strain LM159, CP018470), Xfa_1566 (*X. fuscans* subsp. *aurantifolii* strain 1566, CP012002), Xfa_FDC1609 (*X. fuscans* subsp. *aurantifolii* strain FDC 1609, CP011163), Xca (*X. citri* pv. *anacardii* CFBP 2913, CP024057), Xc_WP2 (*X. cassava*, WP_029220046), Xaj (*X. arboricola* pv. *juglandis* strain Xaj 417, CP012251), Ac (*Acidovorax citrulli*, WP_011794782), Aaa (*A. avenae* subsp. *avenae*, AVS84630), and Cp (*Collimonas pratensis*, WP_061944107).

XopAI HopU1		α1 202020202000		α2 2000000000 200000000	200 200 -	···· 2222	x3 2000000 2000000
XopAI HopU1 Tre1 ART2.2 ExoS	70 29 21 13 233	LNTSDLIKQKKQLWQRV(QHDGAQFRSTPI ERHYSTG(QYEGCVNKMEEH ADK1	EERKQFKTAI DRHDFYRF7 GQEQ7 APLLLQEDF ALADGLVKRF	LITLWGEQYRE AARLH K FNMNAKLK FG	PERQQRWN VDAQCFGLSII XVAWI	NGMMQ R M DDLMD K F .VWTQTA SEAKK R W .ADAE K Y
XopAI HopU1		2000000 QJ	α4 2222222222 2222222222	→ →	- 22222	000000	
XopAI HopU1 Tre1 ART2.2 ExoS	131 67 33 55 253	AQMKWNHPELKYMATI SDKHFRAEHPEYRDVYPI RANAEKNNAQLSTLLTDI NNIKPSRSYPKGFNDI LGRQPGGIHSDI	EDLV ALQAWT TI EECSAIYMHTA(DQIGAIYGYTT) FHGTALVAYTG AEVM <mark>AL</mark> GIYTG) 2DYSSHLVR(4		720 VLEKEAR. INNYLELQHENS INPALEGQTP. INRAVREFKEN INRALEQGQE.	SGREAEI
			Active	site loop		Aliph-R E167	2.5
XopAI HopU1		αs 00000000000 000000000		β1	α6 0000000 00000000		³³
XopAI HopU1 Tre1 ART2.2 ExoS	172 132 77 97 290	PTAHGLAFAKCIISA DNHDEKLSPHIKMLSSA LTPELEAFTGHVTDC .PGQFHYKAFHYYLTRA LDAGQKLIDQGMSAA	LHSLPEEYSYQC LNRLMDVAAFRC LNKLPA.YNC LQLLSNGDCH FEK <mark>S</mark> QQ.AEQVV	GTV FTG EDQI GTV YRG IRGI GET <mark>YRG</mark> TTLE ISV <mark>YRG</mark> TKTE /KT <mark>FRG</mark> TRG(LPDWVSER DLDTIARLYHI PAHILEQN RFH GD.AFNAV	YQERSITT D RI FDTGGRYV E PZ .QIGGTVS D G YTGAGSVRFG .EEGKVGH D D	RFFAASE AFMSTTR GFMSTSA 2FTSSSL GYL <mark>STS</mark> L
				Arom-R T202			A228
XopAI HopU1			β4		β5	β6	β7
XopAI HopU1 Tre1 ART2.2 ExoS	232 197 133 151 347	TKNASWQGMAVEN IKDSAQVFEPGTPNNIAN KTPFDGDVS SKKVAQSQEFFSDHGTLN NPGVARSFGQGTIS PN loop	WESNSTTGKRIS FQISLKRGADIS ISVRGNSGKQI FIIKTCLGVYI STVFGRSGIDVS	SMFSERPNE SGSSQAPSE DFLSKYKNE KEFSFRPDQ SGISNYKNE ARTT loop	EVIFPPGTRE EIMLPMMSEE EVIYPPNTRE EVIYPGYEVY EILYNKETDM	QVTRIEENETI VIEHASALSEG EVINRIEQNG QKVRTQG RVLLSASDEQG	HPRLKIY GK.HLFV .T.THLL .YNEIF GV.TRRV R loop
XopAI HopU1		\rightarrow					
XopAI HopU1 Tre1 ART2.2 ExoS	292 261 188 211 407	QSQ <mark>I</mark> A. LSQ I YRE I P. LDSPKR LEE A AL					

Figure S4. Protein alignment of XopAI and some known mART proteins. Secondary structure elements based on XopAI and HopU1 structures are displayed above the alignment. Blue boxes outline those important regions in mART proteins. The conserved residues in mARTs are marked with red triangles and labeled according to XopAI sequence. The following mART proteins were analyzed: HopU1 (*Pseudomonas syringae* type III-secreted effector HopU1, PDB code 3U0J), Tre1 (*Serratia proteamaculans* type VI secretion ADP-ribosyltransferase effector 1, PDB code 6DRH), ART2.2 (rat mART2.2, PDB code 1GXY), and ExoS (*P. aeruginosa* exoenzyme S, PDB code 6GN8).



Figure S5. Structural comparison between XopAI and known mART proteins. **A**, ribbon diagram showing XopAI, HopU1 (*Pseudomonas syringae* type III-secreted effector HopU1, PDB code 3U0J), Tre1 (*Serratia proteamaculans* type VI secretion ADP-ribosyltransferase effector 1, PDB code 6DRE), ART2.2 (rat mART2.2, PDB code 1GXY), and ExoS (*Pseudomonas aeruginosa* exoenzyme S, PDB code 6GN8). These structures are colored in the rainbow scheme. **B**, comparison of electrostatic surface potentials. The regions of negative and positive potential are shown in red and blue, respectively; uncharged and hydrophobic surface areas are colorless.



Figure S6. ConSurf analysis for XopAI and HopU1. The surface is color-coded according to the sequence conservation among their homologs, from the most conserved residues (purple) to the least conserved residues (cyan). The Arg peptide bound in XopAI and the cofactor NAD⁺ in HopU1 are rendered as stick models. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in green, red, blue, and orange, respectively. The NAD⁺ molecule is docked manually into the active site of HopU1 based on the crystal structure of C3stau2 complexed with NAD⁺ (PDB code 10JZ).



Figure S7. Crystal packing of the full-length XopAI and XopAI- Δ N70 proteins. The crystal packing in a 2 × 1 × 1 supercell is illustrated in the figure. The N-terminus of protein is highlighted as a sphere. **A**, packing of the full-length XopAI in *P*4₃2₁2 crystals. In panels A and B, for clarity, proteins that are packed tandemly are colored in red, green, blue, and yellow, respectively; others are in dark gray. **C**, packing of XopAI- Δ N70 in *P*2₁ crystals. Four proteins in an asymmetric unit are colored in cyan, green, magenta, and yellow, respectively.



Figure S8. Intrinsic tryptophan fluorescence studies showing the peptide-binding ability of the XopAI- Δ N70 protein. **A**, the sequence of the synthetic peptide ArgP14aa. The position of R62 is underlined. **B**, intrinsic tryptophan fluorescence emission spectra of XopAI- Δ N70 in the absence (green) and presence of ligands (blue and red). The molar ratio of ligand to protein was 1:1. **C**, changes of the emission intensity of XopAI- Δ N70 (5 μ M) upon gradual addition of ligands. The y-axis shows the relative change in fluorescence intensity following ligand addition, where F and F₀ are the emission intensities at a certain ligand concentration and in the absence of ligand, respectively. Data points represent means ± standard error (n = 3 separate experiments).



Figure S9. Continuous sedimentation coefficient distribution of the full-length XopAI protein at a concentration of 0.3 mg ml⁻¹. The actual molecular weight of XopAI is 35.8 kDa. The peak having an S value of 2.56 corresponds to a monomeric XopAI.



Figure S10. Comparison of XopAI structures in free and in peptide-bound states. **A**, superimposition of XopAI crystal structures. The full-length (FL) proteins are in Arg peptide-bound state, whereas the Δ N70 protein is in an apo state. Protein backbones are colored distinctly as follows: FL *P*4₁2₁2 (green), FL *P*4₃2₁2 (cyan), and Δ N70 (red). **B**, superimposition of XopAI peptide-binding pocket in available crystal structures. Protein backbones are colored distinctly as follows: FL *P*4₁2₁2 (green), FL *P*4₃2₁2 (cyan), Δ N70 monomer A (red), Δ N70 monomer B (yellow), Δ N70 monomer C (orange), and Δ N70 monomer D (magenta). Key residues in the peptide-binding cleft are shown as stick models and labeled. Oxygen and nitrogen atoms are colored in red and blue, respectively.



Figure S11. Comparative plots from MD analysis of the XopAI-peptide complexes. **A**, binding energy observed between the peptide-binding cleft and the Arg peptide. **B**, RMSD of all non-hydrogen atoms from R62*. **C**, RMSD of C α atoms from the Arg peptide.



Figure S12. Predicted binding free energy contribution per residue of the protein-peptide interaction during the last 25-ns MD simulation. **A**, the contribution energy of residues in the Arg peptide. **B**, the contribution energy of residues at the Arg-binding site. Please note that the algorithm calculates only the contribution from each amino acid sidechain; a negative value of the contribution energy indicates a positive contribution to the binding.



Figure S13. Conformational transition between two Arg peptide-binding modes. **A**, proposed transition states (TSs) between the two observed peptide-binding modes. The scatter plot shows the predicted binding energy of transition states. The selected snapshots (TS4 to TS7) are shown above the scatter plot. The bound Arg peptide in every snapshot is shown as a stick model and colored in the rainbow scheme. Residue W237 and the residues involved in R62* recognition are represented as stick models and their carbons are colored in light gray. The view of snapshots is rotated along the vertical axis by 90° counterclockwise from that of figure 3B. **B**, energy landscapes depicting the motions of the PN loop and residue W237 in the absence (apo) and presence of Arg peptide (peptide-bound). The conformation state of the observed and predicted structures is marked with colored circles, and the color code is indicated at the bottom of scatter plot in panel A.