Supplemental Information

Substrate N^2 atom recognition mechanism in pierisin family DNA-targeting guanine-specific ADP-ribosyltransferase ScARP.

Toru Yoshida[1,2], and Hideaki Tsuge[1,2,3,*]

[1]Department of Bioresource and Environmental Sciences, Faculty of Life Sciences, [2]Institute for Protein Dynamics, and [3]Center for Molecular Research in Infectious Diseases, Kyoto Sangyo University, Kyoto 603-8555

[*]To whom correspondence should be addressed: Hideaki Tsuge; E-mail: tsuge@cc.kyoto-su.ac.jp; Tel. +81-75-705-3117

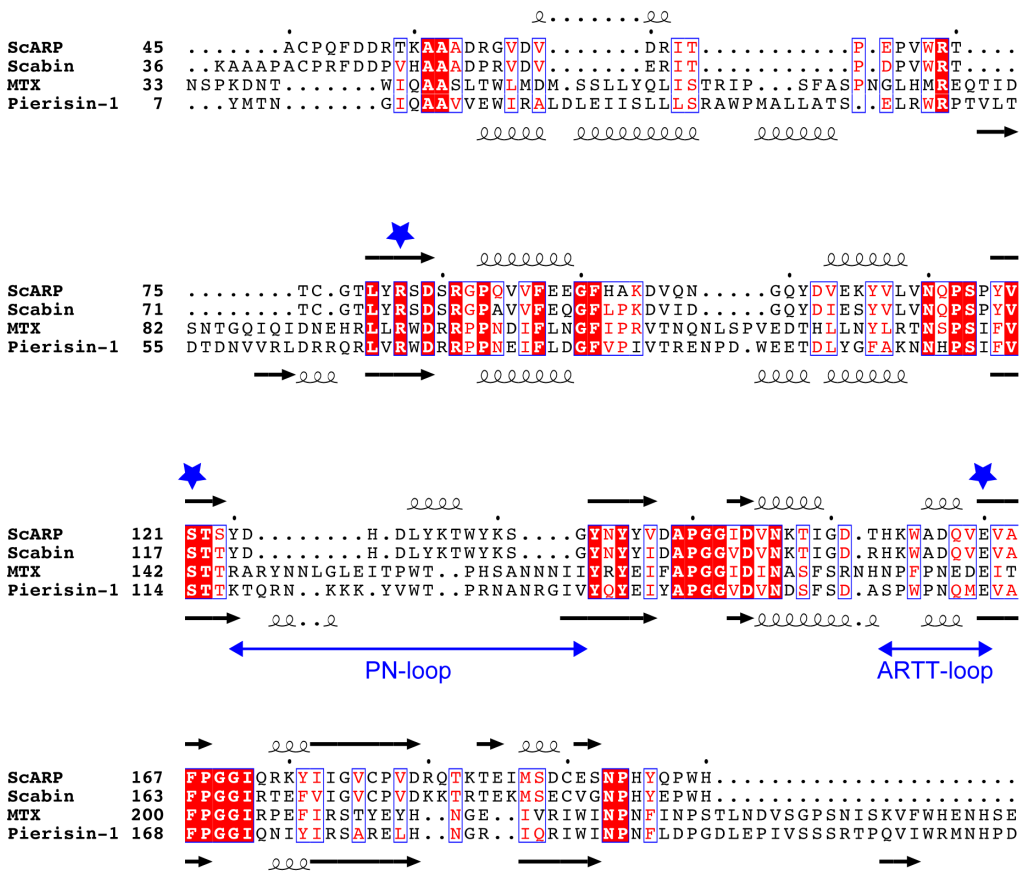

Figure S1. Structure-based sequence alignment of the pierisin family. Secondary structures of ScARP and pierisin are shown over and under the alignment, respectively. Three stars indicate the RSE-motif. The alignment was generated by MATRAS (5) and was manually modified based on the superimposed structures. Scabin, 5TLB.pdb; MTX, 2CB4.pdb; pierisin, 5H6L.pdb.

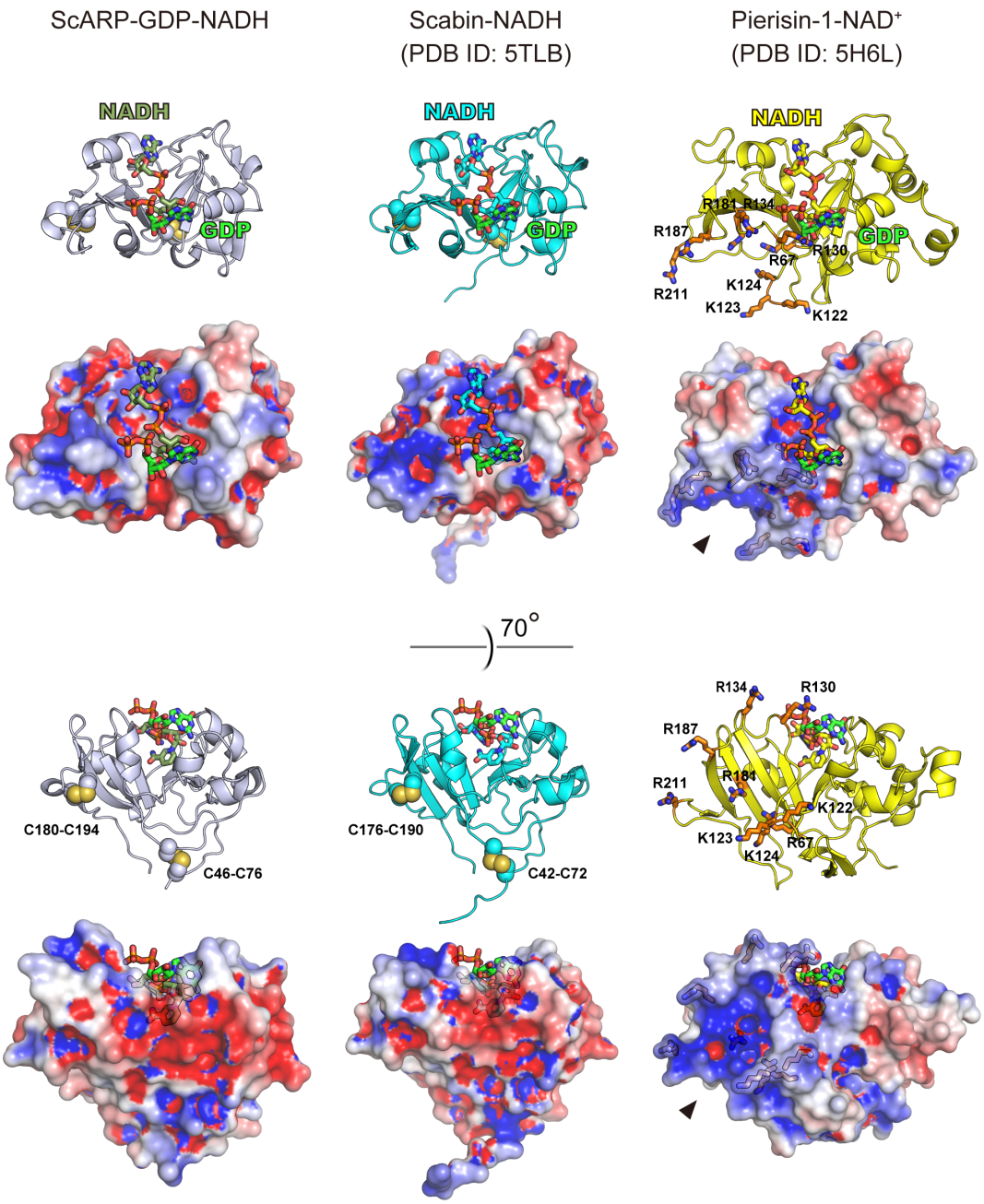

Figure S2. Electrostatic potential maps on the protein surfaces of ScARP, Scabin, and Pierisin-1 at contouring levels of -4.0 kT/e (red) and +4.0 kT/e (blue). GDP in the ScARP-GDP-NADH structure is overlaid in Scabin and Pierisin-1 structures. Black triangles indicate the positive potential cleft in Pierisin-1. The basic residues generating the positive potential cleft are shown as sticks in Pierisin-1. Disulfide bonds are shown as spheres in ScARP and Scabin structures.

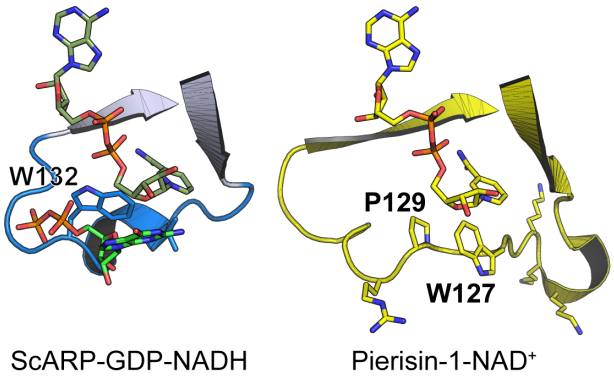

Figure S3. Comparison of the PN-loop between ScARP-GDP-NADH and pierisin-NAD^+ (PDB ID: 5H6L).

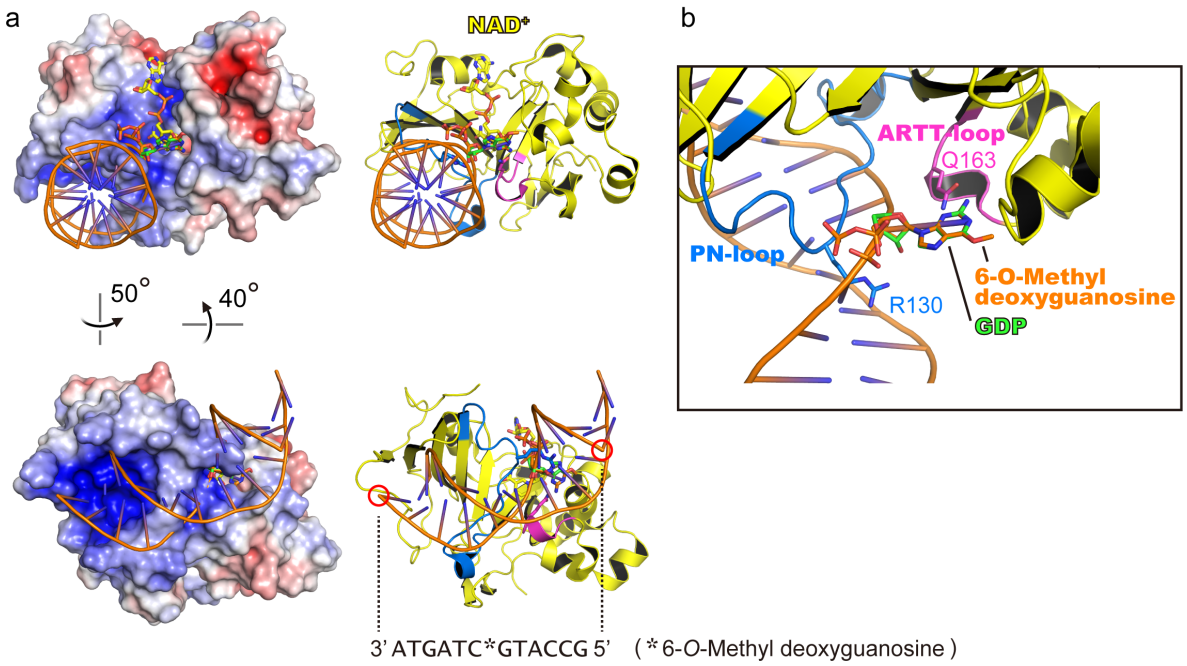

Figure S4. Model structure of Pierisin-1 with dsDNA. After the ScARP-GDP-NADH structure was superimposed on Pierisin-1 with NAD^+ (PDB ID: 5H6L), the guanosine moiety in the flipped 6-O-Methyl deoxyguanosine in 13 bp dsDNA (PDB ID: 1T38) was superimposed on the guanosine moiety in GDP. (a) Electrostatic potential surface and cartoon representation of the model. (b) Close-up view of the active site. The guanosine regions of GDP from ScARP-GDP-NADH (green stick) and 6-O-Methyl deoxyguanosine from dsDNA (orange stick) are superimposed. Gln163 in ARTT-loop contacts GDP. Arg130 in PN-loop protrudes in DNA duplex.

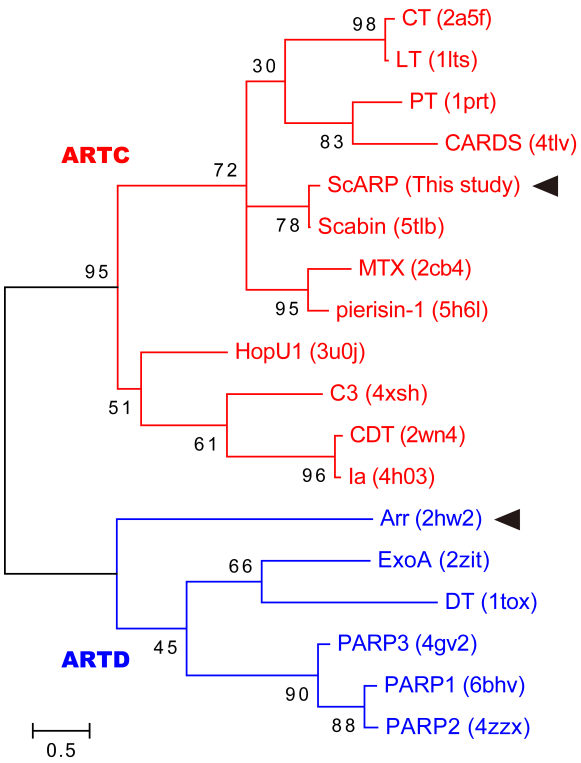

Figure S5. Structure-based phylogenetic tree. Two subgroups of ARTs, ARTC and ARTD, are shown in red and blue, respectively. PDB IDs for structure-based alignment with MATRAS program (1) are shown in parentheses. Black triangles represent the two ARTs that are mentioned in this manuscript, ScARP, and Arr. The phylogenetic tree was generated using the maximum-likelihood method with the MEGA program (2).

Table S1. X-ray Data Collection and Refinement Statistics

	apo-ScARP	ScARP-GDP-NADH
Data collection		
Space group	$P2_1$	$C2$
Cell dimensions		
a, b, c (Å)	93.05, 51.19, 94.97	116.20, 40.05, 77.05
α, β, γ (°)	90.00, 116.24, 90.00	90.00, 113.97, 90.00
Wavelength (Å)	1.0000	1.0000
R_{meas} (%)	10.6 (90.3)	6.7 (42.2)
R_{pim} (%)	4.0 (33.7)	2.8 (18.3)
$CC_{1/2}$[b]	0.998 (0.778)	0.998 (0.896)
$I / \sigma I$ (%)	18.3 (2.1)	24.5 (4.5)
Completeness (%)	99.9 (98.7)	99.6 (96.6)
Redundancy	7.2 (7.1)	5.5 (5.0)
Refinement		
Resolution (Å)	35.6-1.50 (1.55-1.50)	26.5-1.57 (1.62-1.57)
R_{work}/R_{free}	16.2/18.5 (22.0/25.7)	17.4/20.8 (18.8/22.1)
No. waters	542	225
B factors (Å^2)		
Protein	22.2	22.6
GDP, NADH	-	25.7
Water	28.6	25.7
r.m.s.deviations		
Bond lengths (Å)	0.015	0.012
Bond angles (°)	1.36	1.27
Ramachandran plot		
Most favored regions (%)	97.2	95.8
Additional allowed regions (%)	2.8	4.2
Disallowed regions (%)	0.0	0.0

[a] statistics in the highest-resolution shell are given in parentheses.

[b] correlation coefficient between intensities from random half data sets.

References

1. Kawabata, T. (2003) MATRAS: A program for protein 3D structure comparison. *Nucleic Acids Res* **31**, 3367-3369
2. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725-2729