Supplementary Data for:

Study of PcaV from *Streptomyes coelicolor* Yields New Insights into Ligand-Responsive MarR Family Transcription Factors

Jennifer R. Davis, Breann L. Brown, Rebecca Page, and Jason K. Sello.

SI METHODS

Transcriptional Response to Aromatic Compounds. Shaken liquid cultures of wild-type *S. coelicolor* M600 were grown in NMMP for 14 h after which inducing aromatic compounds were added at a final concentration of 2 mM. After 1 h of induction, total RNA was isolated and quantified as previously discussed. Reverse transcriptase polymerase chain reactions (RT-PCR) were accomplished using the One Step RT-PCR Kit (Qiagen) following the manufacturer's protocol. A 314-bp cDNA corresponding to the *pcaH* (SCO6700) transcript and a 486-bp cDNA corresponding to the *hrdB* (SCO5820) transcript were detected using the *pcaH* RT-PCR and *hrdB* RT-PCR primers, respectively. All primer sequences are listed in **Table S2**. The PCR program used for detection of transcripts was 50°C for 30 min, 95°C for 15 min, 30 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 60 s, followed by an elongation time of 10 min at 72°C. PCR products were detected on a 1% agarose gel stained with ethidium bromide. *Pfu* polymerase was used to confirm the absence of contaminating DNA in RNA samples using the same cycling conditions.

Circular Dichroism Polarimetry. Circular dichroism (CD) measurements were performed with a Jasco J-815 CD spectropolarimeter. WT PcaV and PcaV R15K were diluted to 5 μM in 10 mM Tris, 250 mM NaCl, pH 7.5 and placed in a 0.2 cm glass cuvette. PcaV R15A was diluted to 5 μM in 10 mM Tris, 500 mM NaCl, pH 7.5. Wavelength scans were performed in the far-UV region from 260-195 nm at 25°C with a speed of 50 nm/min. Thermal denaturation scans were performed at 220 nm from 25-90°C at a scan speed of 5°C/min. Raw data were converted to mean residue molar ellipticity.

Table S1: Bacterial Strains and plasmids used for this study

Strain or Plasmid	Description	Source / Reference
Strains		
S. coelicolor A3(2)		
M600	Prototroph SCP1-, SCP2-	(30)
B760	M600 *ΔpcaV*::*apr*	(14)
B792	M600 *ΔpcaV*::*apr* – pJS368	This study
B793	M600 *ΔpcaV*::*apr* – pJS369	This study
B794	M600 *ΔpcaV*::*apr* – pJS370	This study
E. coli		
DH5α	F- φ80*lacZ*ΔM15 Δ(*lacZ*YA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(rk- , mk-) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-	Invitrogen
BL21(DE3)	*E. coli* B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte	Stratagene
ET12567	*dam, dcm, hsdS, cat, tet*	(41)
Plasmids		
pBluescript KS+	pUC *ori*, MCS, Amp^r	Agilent Technologies (Stratagene)
pET28a-c(+)	pBR322-derived, T7 promoter, Kan^r,	Novagen
pGem-T Easy	pUC-derived, lacZ', Amp^r	Promega
pUZ8002	RP4 derivative, OriT, Kan^r	(41)
pIJ10257	oriT, ΦBT1 attB-int, ermEp*,derived from pMS81, Hyg^r	(40)
pJS361	5'RACE product of primer AAP (kit) and SCO6703 GSP2 in pGEM-T Easy, Amp^r	This study
pJS362	5'RACE product of primers AUAP (kit) and SCO6703 GSP3 in pGEM-T Easy, Amp^r	This study
pJS363	465bp from cosmid St4C6 containing *pcaV* in pBluescript KS+, Amp^r	This study
pJS364	pET28a-c(+) containing *pcaV* gene, Kan^r	This study
pJS365	*pcaV* R15A mutant in pBluescript KS+, Amp^r	This study
pJS366	pET28a-c(+) containing *pcaV* R15A mutant, Kan^r	This study
pJS367	pET28a-c(+) containing *pcaV* R15K mutant, Kan^r	This study
pJS368	pIJ10257 containing WT *pcaV* gene, Hyg^r	This study
pJS369	pIJ10257 containing *pcaV* R15A gene, Hyg^r	This study
pJS370	pIJ10257 containing *pcaV* R15K gene, Hyg^r	This study

Resistance markers; kanamycin (Kan), ampicillin (Amp) and hygromycin (Hyg)

Table S2: Oligonucleotides used for this study

Oligonucleotide Name	Application / Function	Sequence (5' – 3')
PcaV NdeI For	Cloning of *pcaV* gene	CATATCGCAGCGGTCGATC
PcaV HindIII Rev	Cloning of *pcaV* gene	GGATCCAGGTGTGGGTCGTCTTC
PcaV R15A For	Site-directed mutagenesis of *pcaV* gene	GGCCATCTGGCCGCGCGGCTGCAG
PcaV R15A Rev	Site-directed mutagenesis of *pcaV* gene	TGCTGCAGCCGCGCGGCCAG
PcaV R15K For	Site-directed mutagenesis of *pcaV* gene	ACCCCGGCCATCTGGCCAAGCGGCT GCAGCAGGC
PcaV R15K Rev	Site-directed mutagenesis of *pcaV* gene	GCCTGCTGCAGCCGCTTGGCCAGAT GGCCGGGGT
pcaI GSP1	5' RACE of *pcaI* (SCO6703)	TGTTCGCGCCGATGTAG
pcaI GSP2	5' RACE of *pcaI* (SCO6703)	CAGCCCGGACTCCATC
pcaI GSP3	5' RACE of *pcaI* (SCO6703)	CTCGTACAGCGCCTGGATC
SCO pcaV-I Intergenic For	Amplication of probe containing PcaV binding sites in *S. coelicolor*	GCCACCACCTTGTCCATC
SCO pcaV-I Intergenic Rev	Amplication of probe containing PcaV binding sites in *S. coelicolor*	GGTGGGTGGCCAGATC
SAV pcaV-I Intergenic For	Amplication of probe containing PcaV binding sites in *S. avermitilis*	TCCGTGAAGCCCCATCTC
SAV pcaV-I Intergenic Rev	Amplication of probe containing PcaV binding sites in *S. avermitilis*	GTGTGCTCCTCGCTCCAC
SCAB pcaV-H Intergenic For	Amplication of probe containing PcaV binding sites in *S. scabies*	TGGCGGTTCCTGACGT
SCAB pcaV-H Intergenic Rev	Amplication of probe containing PcaV binding sites in *S. scabies*	GTTCGCTCCTCGGCTC
SCO O_V For	Amplication of probe containing O_V binding site	CACTCCCGGTGGTCACTT
SCO O_V bio Rev	Amplication of probe containing O_V binding site	bio - CGTCAAGACCCGCAAAAG
SCO O_I bio For	Amplication of probe containing O_I binding site	bio -TCTGCCTCGGACTTTTGC
SCO O_I Rev	Amplication of probe containing O_I binding site	GAGGAGCCCCATCTCTGC
SCO O_V 30 bp For	Generation of 30 bp probe containing O_I binding site	CTGAAATACTCAGTGCCCTGACTATC TGCT - bio
SCO O_V 30 bp Rev	Generation of 30 bp probe containing O_I binding site	AGCAGATAGTCAGGGCACTGAGTAT TTCAG - bio
SCO O_I 30 bp For	Generation of 30 bp probe containing O_V binding site	TCACTTTCGTAGGTGCACTGACGATA CATG - bio
SCO O_I 30 bp Rev	Generation of 30 bp probe containing O_V binding site	CATGTATCGTCAGTGCACCTACGAAA GTGA - bio
pcaH RT PCR For	Detection of the *pcaH* (SCO6700) transcript	ACGGCTTCTACCGCTTCAC
pcaH RT PCR Rev	Detection of the *pcaH* (SCO6700) transcript	GTCCTTCTTCGATCCAGGTG
hrdB RT PCR For	Detection of the *hrdB* (SCO5820) transcript	CTCGAGGAAGAGGGTGTGAC
hrdB RT PCR Rev	Detection of the *hrdB* (SCO5820) transcript	TGCCGATCTGCTTGAGGTAG

Table S3: PcaV secondary structural elements

Secondary structural element	Apo-PcaV residues	PcaV-PCA residues
α0	disordered	5 - 8
α1	9 - 29	9 - 29
α2	34 - 47	34 - 47
β1	50 - 51	50 - 51
α3	52 - 60	52 - 60
α4	62 - 76	62 - 76
β2	79 - 83	79 - 83
W1	disordered	84 - 90
β3	91 - 95	91 - 95
α5	96 - 119	96 - 119
α6	122 - 142	122 - 141

Table S4: Structural Homologs of PcaV

MarR	Dali Z-score	PDB	MarR Family	Arginine*	Experimentally determined role	Reference
PcaV	100.0	4FHT 4G9Y	Anionic ligand	Arg15	Interacts directly with ligand (PCC) and important for DNA binding	This work
PA1603	18.9	2NNN	nd	Arg17	nd	--
Rha1	16.7	3FM5	nd	Arg17°	nd	--
BldR	16.6	3F3X	Redox	Arg19**	Predicted to bind DNA	(18)
MTH313	16.5	3BPX	Anionic ligand	Arg16**	Interacts directly with ligand (salicylate)	(55)
OhrR	16.4	1Z9C	Redox	Arg23	Interacts directly with DNA	(25)
MosR	16.4	4FX4	Redox	Arg16	nd	(24)
PA4135	16.2	2FB1	nd	None	nd	--
SlyA	16.1	3Q5F 3DEU	Anionic ligand	Arg14**	Near DNA interface; hydrogen bonds with carbonyl from second SlyA molecule in dimer to stabilize DNA-bound form; binds to salicylate	(56)
n/a	16.0	3NRV	nd	None	--	--
MexR	15.9	1LNW	ligand	None	--	(59)
n/a	15.8	3E6M	nd	Arg30°	nd	--
ST1710	15.8	3GFL	Anionic ligand	Arg20**	Interacts directly with salicylate	(54)
SCO5413	15.7	4B8X	nd	Arg37°	nd	(28)
n/a	15.4	3CJN	nd	Arg29°	nd	nd
n/a	15.3	3S2W	nd	Arg30**	nd	nd
MarR	15.1	1JGS	Anionic ligand	None	nd	(47)
SarZ	15.0	3HRM	Redox	None	--	(27)
n/a	14.9	3BJA	nd	Arg11	nd	--
MepR	14.0	3ECO	Cationic ligand	Arg10	nd	(53)
TcaR	11.6	3KP6	Anionic ligand	None	--	(17)

*Identity of the arginine which is in a structurally homologous position as Arg15 in PcaV or interacts in a functionally homologous manner. Because of the plasticity in helices within the MarR fold, there are often shifts in the helical register of the N-terminal helix (which includes PcaV Arg15) when two structural homologs are optimally overlapped with one another. For this reason, we include arginines whose sidechains overlap with that of Arg15 (**), but whose backbones are displaced by one helical turn; we have also listed those arginines that lie

between PcaV Arg15/Arg16 or overlap with Arg16 (°). The latter arginines are included in this table because the sidechain may move into a position more similar to Arg15 if ligand is bound. However, for the calculation of the percentage of MarR homologs with a “structurally conserved” arginine, we included only those arginines that overlapped perfectly with PcaV Arg15 or whose sidechain overlapped with that of PcaV Arg15.

**Arginine is one helical turn away from the position of Arg15 in PcaV, but the guanadnium headgroups of both arginines are in nearly identical locations

°Arginine overlaps with PcaV Arg16

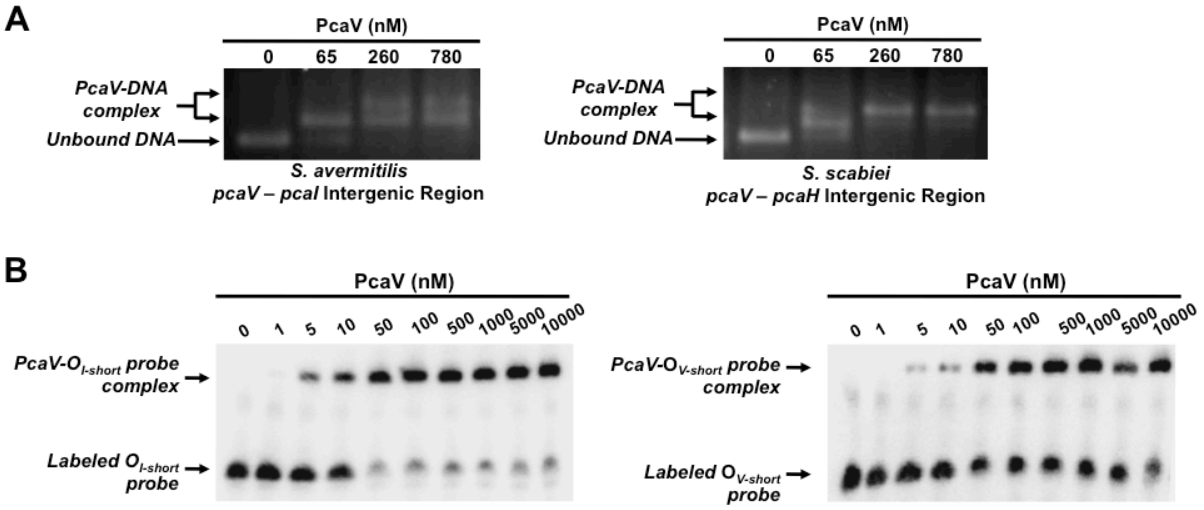

Figure S1. *(A)* Agarose EMSA experiments using PCR-amplified DNA probes spanning the *pcaV-pcaI* and *pcaV-pcaH* intergenic regions of *S. avermitilis* and *S. scabiei*, respectively. A change in DNA migration is observed upon addition of increasing concentrations of PcaV in comparison to the unbound DNA probe. The formation of the PcaV-probe complex is indicated by a double arrow while the unbound DNA is indicated with a single arrow. *(B)* EMSA experiments using 30-bp biotin-labeled probes containing the O_I site (O_I-short, left panel) and O_V site (O_V-short, right panel) and increasing concentrations of PcaV to confirm specificity of PcaV for operator sites. A single shifted band corresponding to the PcaV-DNA probe complex is observed.

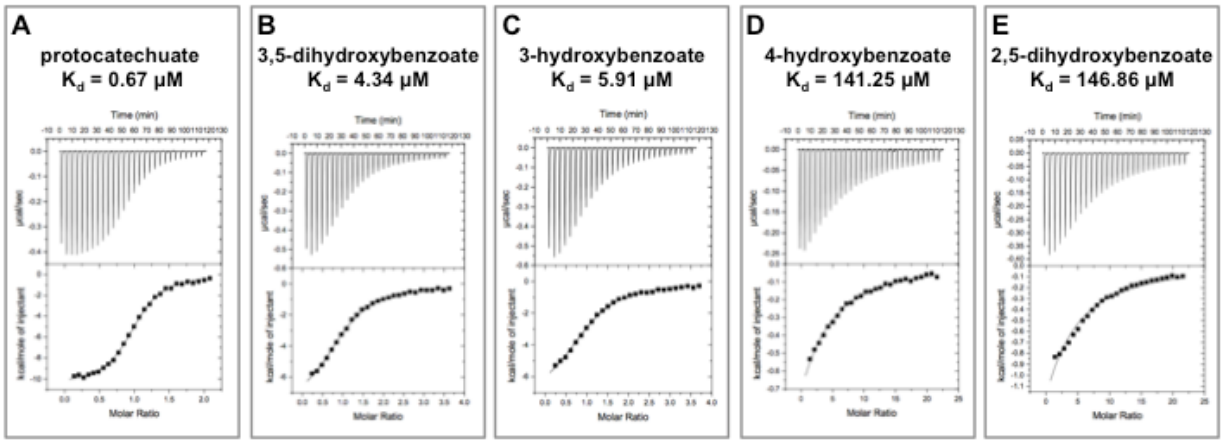

Figure S2. Representative binding isotherms of PcaV with aromatic ligands protocatechuate (3,4-dihydroxybenzoate, *A*), 3,5-dihydroxybenzoate *(B)*, 3-hydroxybenzoate *(C)*, 4-hydroxybenzoate *(D)*, and 2,5-dihydroxybenzoate *(E)*. Average dissociation constants (K_D) for each ligand are provided (4-hydroxybenzoate and 2,5-dihydroxybenzoate are low affinity ligands and thus these experiments were performed using the criteria established by Turnbull and Daranas for the analysis of low affinity systems using ITC). Top panel displays raw data and bottom panel is the corresponding binding isotherm generated using a one-site binding model.

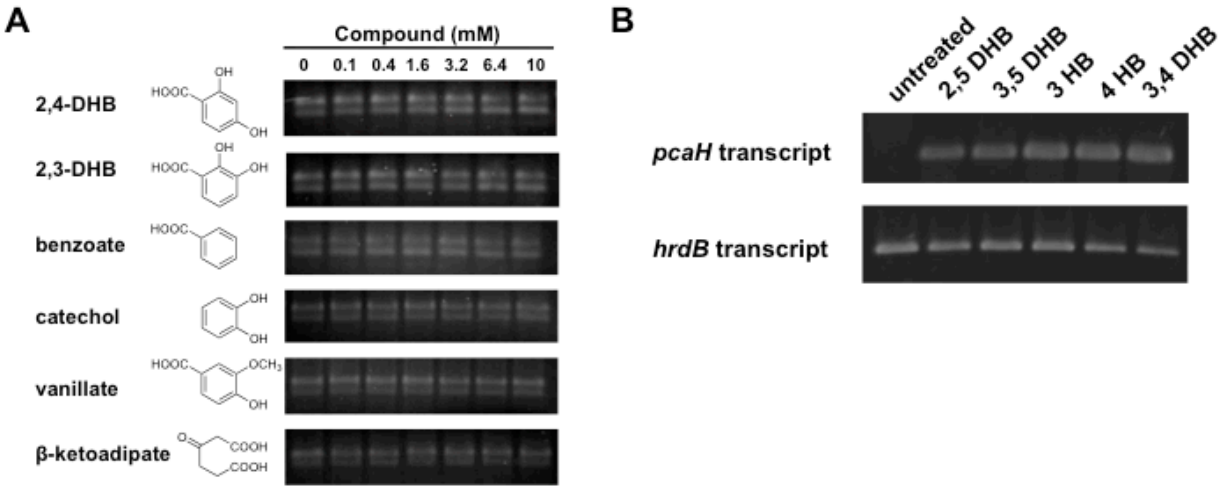

Figure S3. *(A)* EMSAs of PcaV in complex with the *pcaV-pcaI* intergenic DNA probe in the presence of increasing amounts of the indicated compounds. No dissociation of PcaV from DNA is observed. *(B)* Transcriptional response of *S. coelicolor* M600 to 2,5 dihydroxybenzoate, 3,5-dihydroxybenzoate, 3-hydroxybenzoate, 4-hydroxybenzoate, and 3,4-dihydroxybenzoate. The *pcaH* transcript is induced in response to the PcaV ligands. The *hrdB* control transcript was detected in all of the RNA samples. The *hrdB* gene encodes a necessary vegetative sigma factor and is constitutively transcribed.

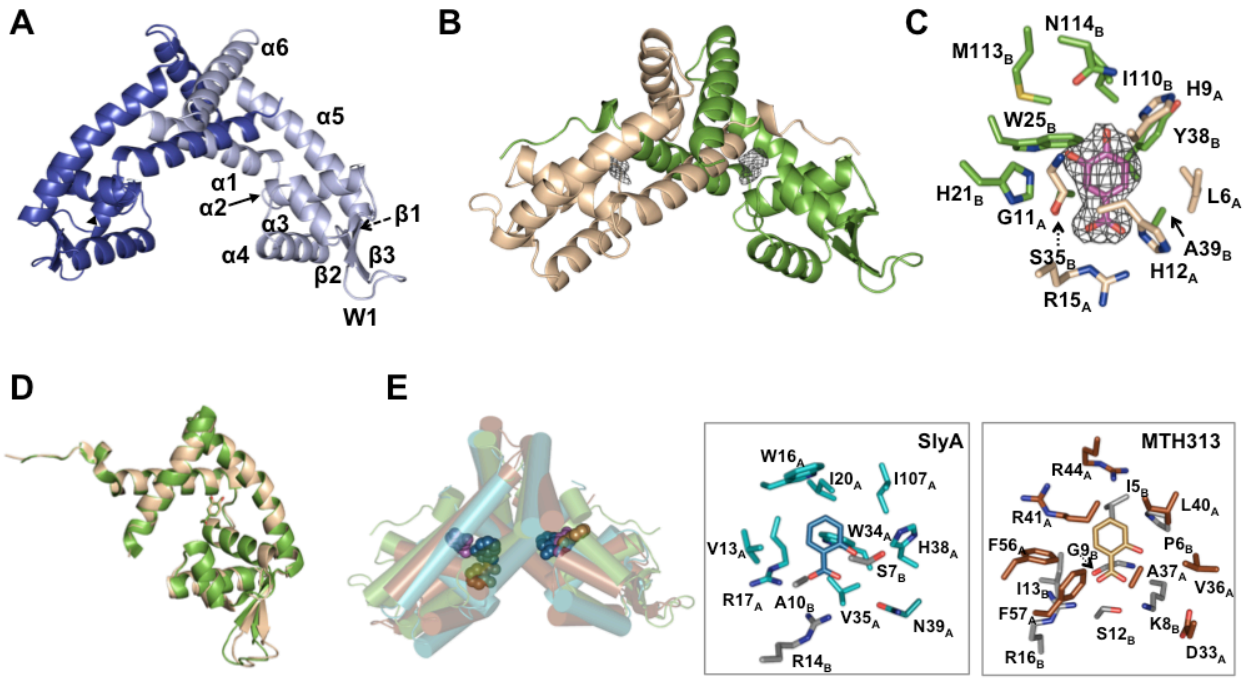

Figure S4. *(A)* Crystal structure of WT PcaV shown as a cartoon representation with each monomer in light and dark blue, respectively. Secondary structure elements are labeled. *(B)* Cartoon representation of PCA-bound PcaV (colored as in **Figure 2**) with simple composite omit map shown for protocatechuate ligand. σA-weighted $2mF_o$-dF_c electron density map shown as grey mesh and contoured at 1.5 σ. *(C)* Binding pocket of PCA (as in **Figure 2D**) shown with omit map. Ligand was absent from model when omit map was generated. *(D)* Overlay of monomer A and B from the PcaV-PCA complex illustrating the overlapping PCA binding pockets and the small differences in the W1 region between each monomer.*(E)* Overlay of PcaV (green), SlyA (cyan, PDB 3DEU) and MTH313 (brown, PDB 3BPX) with bound ligands shown as spheres. Only two of six salicylate molecules present in the SlyA crystal structure are shown. Inset: ligand binding pockets in SlyA and MTH313 that are located in an orientation similar to that observed in the PcaV-PCA structure. Bound salicylate molecules and the residues that become buried upon ligand-binding are shown as sticks.

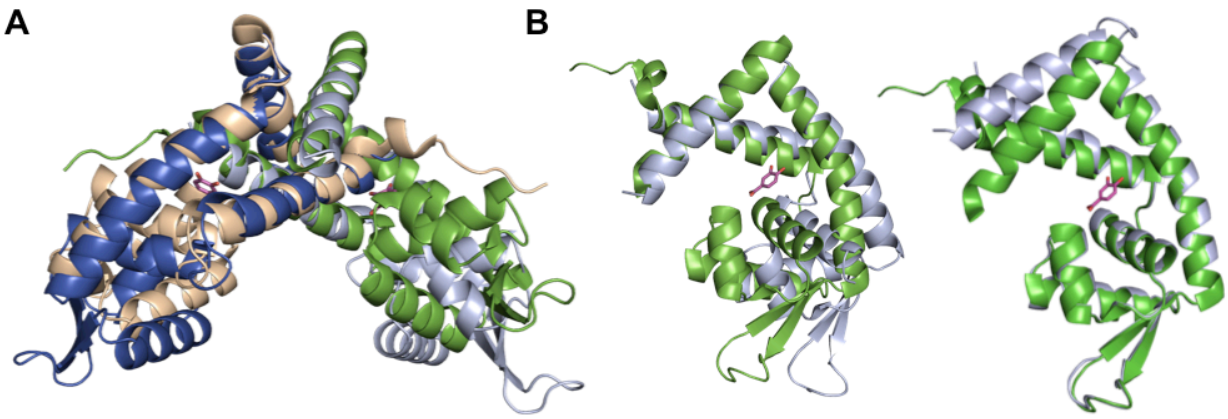

Figure S5. *(A)* Superposition of WT PcaV and PcaV-PCA shown as cartoon representation and colored as in **Figure 3**. *(B)*. Superposition of one monomer of apo-PcaV onto one monomer of ligand-bound PcaV using either the dimerization core domain (left panel) or the wHTH core (right panel) shows that each PcaV monomer contains two rigid-body domains.

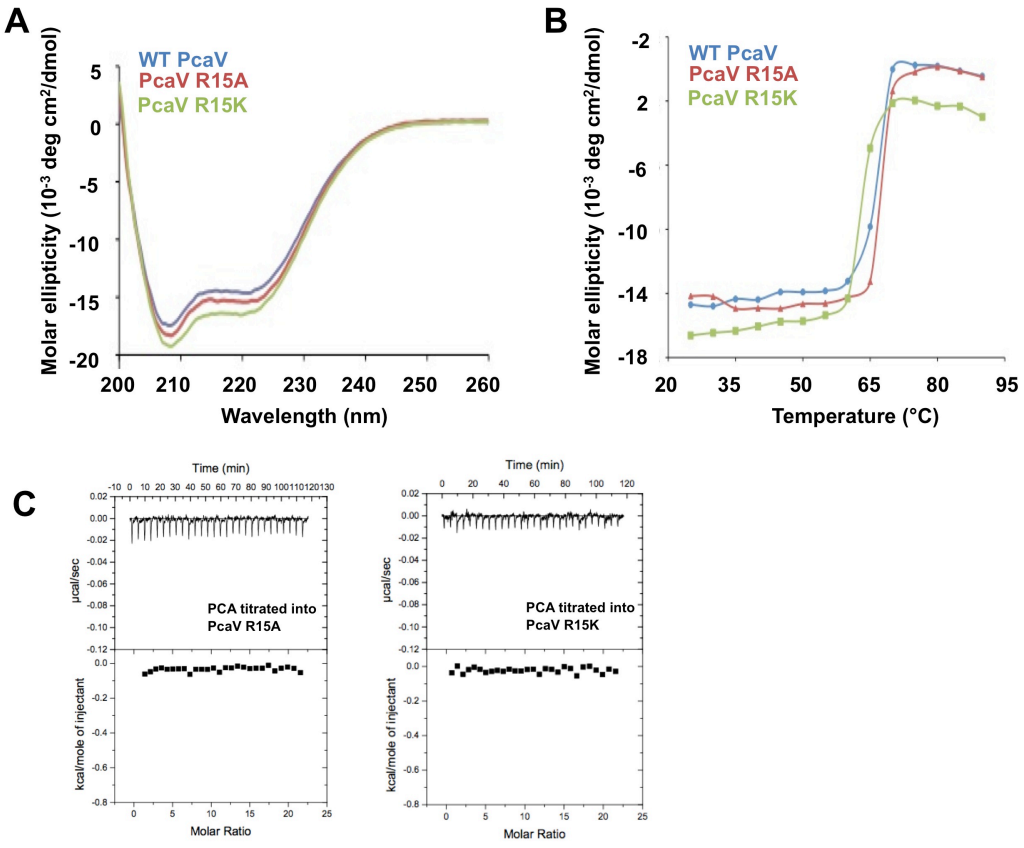

Figure S6. Circular dichroism wavelength scans *(A)* and temperature scans *(B)* show that PcaV R15A and PcaV R15K mutants (red and green traces, respectively) are well-folded and behave similarly to WT PcaV (blue traces). *(C)* ITC experiments show there is no binding resulting from titration of PCA ligand into either PcaV R15A or PcaV R15K mutants. Top panel displays raw data and bottom panel is the corresponding binding isotherm generated using a one-site binding model.