Angewandte Chemie

Supporting Information

Tandem Acyl Carrier Proteins in the Curacin Biosynthetic Pathway Promote Consecutive Multienzyme Reactions with a Synergistic Effect**

Liangcai Gu, Eli B. Eisman, Somnath Dutta, Titus M. Franzmann, Stefan Walter, William H. Gerwick, Georgios Skiniotis, and David H. Sherman*

ange_201005280_sm_miscellaneous_information.pdf

Supporting Information:

Chemicals: All chemicals were from Sigma-Aldrich. The authentic standard of **1** was generated by the hydrolysis (in 0.1 M KOH at 60°C for 10 min) of a corresponding *N*-acetylcysteamine thioester of **1**, which is a gift from Dr. Timothy M. Ramsey (Novartis Institutes for Biomedical Research, Inc.). As previously reported,[1] the thiazoline ring was closed by cleavage of the Boc group with trifluoroacetic acid to furnish a corresponding ammonium trifluoroacetate salt in refluxing benzene. **1** was further analysis by LC-MS and MS/MS analysis. MS (ESI) calculated for [M-H]$^-$ 184.05, found 183.96. MS/MS fragmentation of 183.96, found 140.00, 138.02 and 155.96.

Plasmid construction and site-directed mutagenesis: CurA ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd tetradomain, CurF PCP and Cy-A-PCP tridomain genes were amplified from the pLM54.[2] The ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd gene was inserted into a pET28b plasmid with the *Nde*I and *Xho*I restriction sites. The PCP and Cy-A-PCP genes were inserted into a pET24b plasmid with the *Nde*I and *Xot*I restriction sites. Primers for the plasmid construction were: ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd (F): 5'-*CAT ATG* CCT CAA ACT CAG CAA CCT TT-3', ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd (R): 5'-*CTC GAG* CTC CTA AAC TGT AAC CTG TTT-3', PCP (F): 5'-*CAT ATG* GCC GAT ACG AAT CAA AAT ATA-3', PCP (R): 5'-*GC GGC CGC* TTT TTT GCC GCG TAG ACT TTT-3', Cy-A-PCP (F): 5'-*CAT ATG* GAG CAA AGG TAT CAA CCA TTT-3', and Cy-A-PCP (R): 5'-*GC GGC CGC* TTT TTT GCC GCG TAG ACT TTT-3'. The ACP$_I$-ACP$_{II}$ (1946-2146) and ACP$_{II}$-ACP$_{III}$ (2057-2248) didomain genes were amplified from a codon optimized synthetic CurA ACP$_3$ gene,[3] and inserted into pET29a plasmid with the *Nde*I and *Xho*I restriction sites (a gift from Dr. Christopher Walsh and Dr. Christopher Calderone at Harvard Medical School). The Ser>Ala ACP acitve-site mutations were prepared from the ACP$_I$-ACP$_{II}$-ACP$_{III}$ construct containing synthetic ACP$_3$ gene.[3] Single or double Ser>Ala mutations were generated with the following primers: ACP$_I$ (F): 5'-GTT GAC CTG GGC CTG GAT *G*CT ATT GTT GGT GTC GAA TGG-3', ACP$_I$ (R): 5'-CCA TTC GAC ACC AAC AAT AG*C* ATC AGC CCA GTC AAC-3', ACP$_{II}$ (F): 5'-C GTA GAT CTG GGT CTG GAC *GC*C ATT GTT GGC GTG GAG TGG-3', ACP$_{II}$ (R): 5'-CCA CTC CAC GCC AAC AAT G*GC* GTC CAG ACC CAG ATC TAC G -3', ACP$_{III}$ (F): 5'-G GAC CTG GGT CTG GAC *G*CT ATT GTT GGT GTG GAG-3', and ACPII$_I$ (R): 5'-CTC CAC ACC AAC AAT AG*C* GTC CAG ACC CAG GTC C-3'. The Cy-A-PCP S3118A mutant was generated with the following primers: PCP (F): 5'-C TTT GCC ATA GGA GGA AAT *G*CT TTA CTT GTC ATC AGA GTG-3', PCP (R): 5'-CAC TCT GAT GAC AAG TAA AG*C* ATT CCT CCT ATG GCA AAG-3'. Three mutants with double Ser>Ala active-site mutations on ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd were generated by splicing the ACP$_3$ genes from the corresponding Ser>Ala double-site mutants of the ACP$_I$-ACP$_{II}$-ACP$_{III}$ construct above mentioned, and the Cd gene from the ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd construct. Two primer pairs were used for the splicng: (F): 5'-*CAT ATG* ACT CCT CAG GTA AAT CAA GTG-3', (R): 5'-CCA TTG CCA TTA CCT TGG AAC AGC TTA GAA CCA CCA GTA G-3', (F): 5'-CTA CTG GTG GTT

CTA AGC TGT TCC AAG GTA ATG GCA ATG G-3', and (R): 5'-*CTC GAG* CTC CTA AAC TGT AAC CTG TTT-3'. The spliced products were inserted into pET28b. All the constructs and mutatants were confirmed by DNA sequencing.

Protein expression and purification: CurA Hal, CurE ECH$_1$, CurF ECH$_2$ and ER were prepared as previously described.[3, 4] The CurA ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd, ACP$_I$-ACP$_{II}$-ACP$_{III}$, mutants and truncations, CurF PCP, Cy-A-PCP and the PCP active site mutant of Cy-A-PCP were prepared as described below. *E. coli* BL21 (DE3) cells were transformed with the corresponding plasmids to overexpress His-tagged proteins. All the polypeptides containing ACP or PCP domains were expressed in the apo form. The PCP domain was also prepared in the holo form by coexpression with the Sfp.[5] Cells were grown at 30°C to an OD (590 nm) = 0.5-0.6, and then cooled to 15°C prior to the addition of 1 mM isopropyl-β-D-galactopyranoside. The cultures were grown at 15°C for another 18-20 h before harvesting. Protein purifications were performed at 4°C. *E. coli* cells were harvested by centrifugation (5,000 g), resuspended in ice cold lysis buffer A (50 mM Tris-HCl buffer, pH 8.0, 300 mM NaCl, 10 mM imidazole, 20% glycerol) and disrupted by sonication on ice. The cell debris was removed by centrifugation at 15,000 g. The supernatant was gently removed and loaded onto the 5 ml HisTrap column (GE Healthcare) pre-equilibrated with lysis buffer A. The resin was washed successively with ~10 column volumes of the washing buffer B (50 mM Tris-HCl buffer, pH 8.0, 300 mM NaCl, 20 mM imidazole, 10% glycerol) to remove non-specifically bound contaminants. Bound proteins were eluted with imidazole by a linear gradient of the elution buffer C (50 mM Tris-HCl buffer, pH 8.0, 300 mM NaCl, 250 mM imidazole, 20% glycerol). The eluate fractions were examined by SDS-PAGE for purity, pooled and concentrated using Amicon Ultra-15 (30 kDa, 10 kDa or 5 kDa) centrifugal devices (Millipore). The concentrated eluate was loaded onto PD10 column (GE Healthcare) equilibrated with the storage buffer D (50 mM Tris-HCl buffer, pH 7.5, 200 mM NaCl, 20% glycerol). The fractions were pooled, concentrated, flash-frozen in 50-200 μl aliquots in liquid N$_2$, and stored at -80°C for future use. The purity of the proteins was analyzed by SDS-PAGE and the protein concentrations were determined using the Bradford assay (Bio-Rad).

Size-exclusion chromatography: Analytical SEC was performed on an ÄKTA FPLC system equipped with Superdex 75 10/300 GL and Superdex 200 10/300 GL columns (GE Healthcare). The columns were pre-equilibrated with a buffer E (50 mM Tris-HCl buffer, pH 7.5, 200 mM NaCl, 10% glycerol), and initially calibrated using a set of molecular weight standards (Sigma-Aldrich). The samples were eluted at a flow rate of 0.5 mL/min at 4°C.

Cy-A-PCP adenylation and condensation/cyclization activity assay: To prepare the substrate for the Cy domain, (1*R*,2*S*)-2-methylcyclopropane carboxylic-CoA, synthesized by coupling the carboxylic acid to CoA-SH as previously described,[3, 4] was loaded onto (apo) ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd by using the Sfp.[5] Typically, to achieved a full substrate loading onto ACPs, 500

µM acyl-CoA and 50 µM (apo) ACP_I-ACP_{II}-ACP_{III}-Cd were incubated with 2 µM Sfp, and 10 µM $MgCl_2$ in 50 mM Tris-HCl buffer, pH 8.1, at room temperature for ~2 h. The ACP-linked substrates were desalted by PD10 column equilibrated with the buffer D, and concentrated by using Amicon Ultra-4 (10 kDa, Millipore). For the assay of adenylation and condensation/cyclization activities, a typical reaction mixture contains 10 µM ACP_I-ACP_{II}-ACP_{III}-Cd loaded with the corresponding acyl-CoA substrate, 20 µM (apo) Cy-A-PCP, 2 µM Sfp, 5 µM $MgCl_2$, 2 mM TCEP, 100 µM CoA-SH, 2 mM ATP, and 1 mM L-cysteine in 50 mM Tris-HCl buffer, pH 7.5. The reaction was incubated at room temperature for 30 min before quenched by addition of 10% formic acid. Next, Cy-A-PCP was purified from the reaction mixture by reverse-phase HPLC using the Source 15PRC column (GE Healthcare). The proteins were eluted using a linear gradient from 10% to 90% of CH_3CN (0.1% CF_3COOH)/H_2O (0.1% CF_3COOH), and pooled in glass vials for lyophilization. The lyophilized Cy-A-PCP was resuspended in 0.1 M KOH and heated at 60°C for 10 min.[6] The mixture was then neutralized by 0.1 M acetic acid and proteins were removed by using Amicon Microcon centrifugal devices (10 kDa, Millipore). Finally, the product **1** was analyzed by LC-MS.

Comparison of the ACP_3 constructs, truncations and mutants in the one-pot reaction: The ACP-linked HMG substrates were prepared by loading (S)-HMG-CoA onto the apo ACP_I-ACP_{II}-ACP_{III}-Cd, ACP_I-ACP_{II}-ACP_{III}, mutants and truncations similarly as previously described.[3, 4] Immediately before the one-pot reaction, the PCP-linked L-cysteine substrate was generated by incubating a mixture of 100 µM (apo) PCP, 4 µM Cy-A-PCP S3118A mutant, 2 µM Sfp, 5 µM $MgCl_2$, 2 mM TCEP, 0.5 mM CoA-SH, 2 mM ATP, and 1 mM L-cysteine in 50 mM Tris-HCl buffer, pH 7.5 at room temperature for 1 h. A typical one-pot reaction mixture contains 60 µM ACP-linked HMG substrate (e.g., 20 µM, 30 µM or 60 µM polypeptide containing three, two or one active ACP domain, respectively), 50 µM PCP-linked L-cysteine substrate, 2 µM Cy-A-PCP S3118A mutant, Hal, ECH_1, ECH_2, and ER, and the corresponding cofactors (20 µM $Fe(NH_4)_2(SO_4)_2$, 0.5 mM α-ketoglutarate, and 0.5 mM NADPH) in 50 mM Tris-HCl buffer, pH 7.5.[3] The reaction was initiated by addition of the ACP-linked HMG substrates and exposed to air, and then incubated at room temperature for 10 min before being quenched by 10% formic acid. Next, similar to the Cy-A-PCP assay described above, the PCP was purified by reverse-phase HPLC and lyophilized, and the product was cleaved from the lyophilized PCP in 0.1 M KOH at 60°C for 10 min. To avoid oxidation of the thiazoline group, the hydrolysis of PCP-linked product and following steps were performed under anerobic conditions. Samples were stored at -80°C before analysis. The cleaved product **1** was subjected to the LC-MS quantification described below.

LC-MS analysis: The LC-MS analysis was performed on a Surveyor HPLC system equipped with Finnigan ESI-LTQ or ESI-TSQ mass spectrometer (ThermoFisher Scientific). Samples were loaded onto an XBridge reverse-phase column (C18, 3.5 µm, 2.1x150 mm, Waters) and eluted with a gradient of 3-100% CH_3OH/H_2O (10 mM NH_4OAc) at a flow rate of 0.2 mL/min. The MS parameters were optimized by using the authentic standard of **1**. For the LC-MS quantification of

1, cyclopentanecarboxylic acid was employed as an internal standard, and the chromatograms and MS spectra were collected at the selected ion mode (m/z: 112.7-113.7 and 183.5-184.5). The data were analyzed by the Xcalibur software (ThermoFisher Scientific).

Analytical Ultracentrifugation: ACP_I-ACP_{II}-ACP_{III}-Cd and ACP_I-ACP_{II}-ACP_{III} were analyzed by sedimentation velocity in 50 mM Tris, 200 mM NaCl, 5% Glycerol at 0.2, 0.5 and 0.8 mg/ml using a ProteomeLab XL-I (Beckman Coulter). Samples were loaded into sector-shaped double channel centerpieces and spun at 42,000 rpm in an AN50TI rotor and 22°C. Absorbance scans were collected continuously at 280 nm and a radial resolution of 40 µm and 3 replicates. Data analysis was with the enhanced van Holde-Weischet analysis module followed by 2-dimensional sedimentation spectrum analysis (2-DSA) using the finite element modeling module provided in the Ultrascan software (http:/www.ultrascan.uthscsa.edu). Sedimentation profiles were analyzed at a grid resolution of 25600 using 16 grid repetitions. Confidence levels were derived from 2-DSA data refinement using genetic algorithm followed by 50 Monte Carlo simulations. Calculations to analyze analytical ultracentrifugation data were performed on the UltraScan LIMS cluster at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio, the Lonestar cluster at the Texas Advanced Computing Center (supported by NSF Teragrid Grant #MCB070038 to Borries Demeler). vBar of 0.739 ml/g for ACP_I-ACP_{II}-ACP_{III}-Cd and ACP_I-ACP_{II}-ACP_{III} was calculated from the polypeptide sequence and the buffer density (1.0235 g ml^{-1}) and viscosity (1.2116 cp) were calculated using the buffer calculation module provided with the Ultrascan software suite.

Electron microscopy and image processing: For electron microscopy, samples were prepared using conventional negative staining protocols.[7] Briefly, 3 µL of sample was pipetted onto a glow-discharged carbon-coated grid and stained with 2% (w/v) uranyl formate. Imaging was performed at room temperature with a Morgagni 268(D) transmission electron microscope (FEI Company) equipped with a tungsten filament operated at 100kV. Recordings were acquired on an Orius SC200W CCD camera (Gatan Inc.) at a nominal magnification of 30,416X. For classification and averaging 4,104 ACP3 and 5,662 ACP3Cd particles were interactively selected using EMAN Boxer (Ludtke et al. 1999). Reference free alignment and classifications into 100 classes for each sample were performed in EMAN using refine2d.py.

References:

[1]   J. D. White, T. S. Kim, M. Nambu, *J. Am. Chem. Soc.* **1997**, *119*, 103.
[2]   Z. Chang, P. Flatt, W. H. Gerwick, V. A. Nguyen, C. L. Willis, D. H. Sherman, *Gene* **2002**, *296*, 235.
[3]   L. Gu, B. Wang, A. Kulkarni, T. W. Geders, R. V. Grindberg, L. Gerwick, K. Hakansson, P. Wipf, J. L. Smith, W. H. Gerwick, D. H. Sherman, *Nature* **2009**, *459*, 731.
[4]   L. Gu, J. Jia, H. Liu, K. Hakansson, W. H. Gerwick, D. H. Sherman, *J. Am. Chem. Soc.* **2006**, *128*, 9014.
[5]   L. E. N. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585.
[6]   W. L. Kelly, N. J. Hillson, C. T. Walsh, *Biochemistry* **2005**, *44*, 13385.
[7]   M. Ohi, Y. Li, Y. Cheng, T. Walz, *Biol. Proced. Online* **2004**, *6*, 23-34.
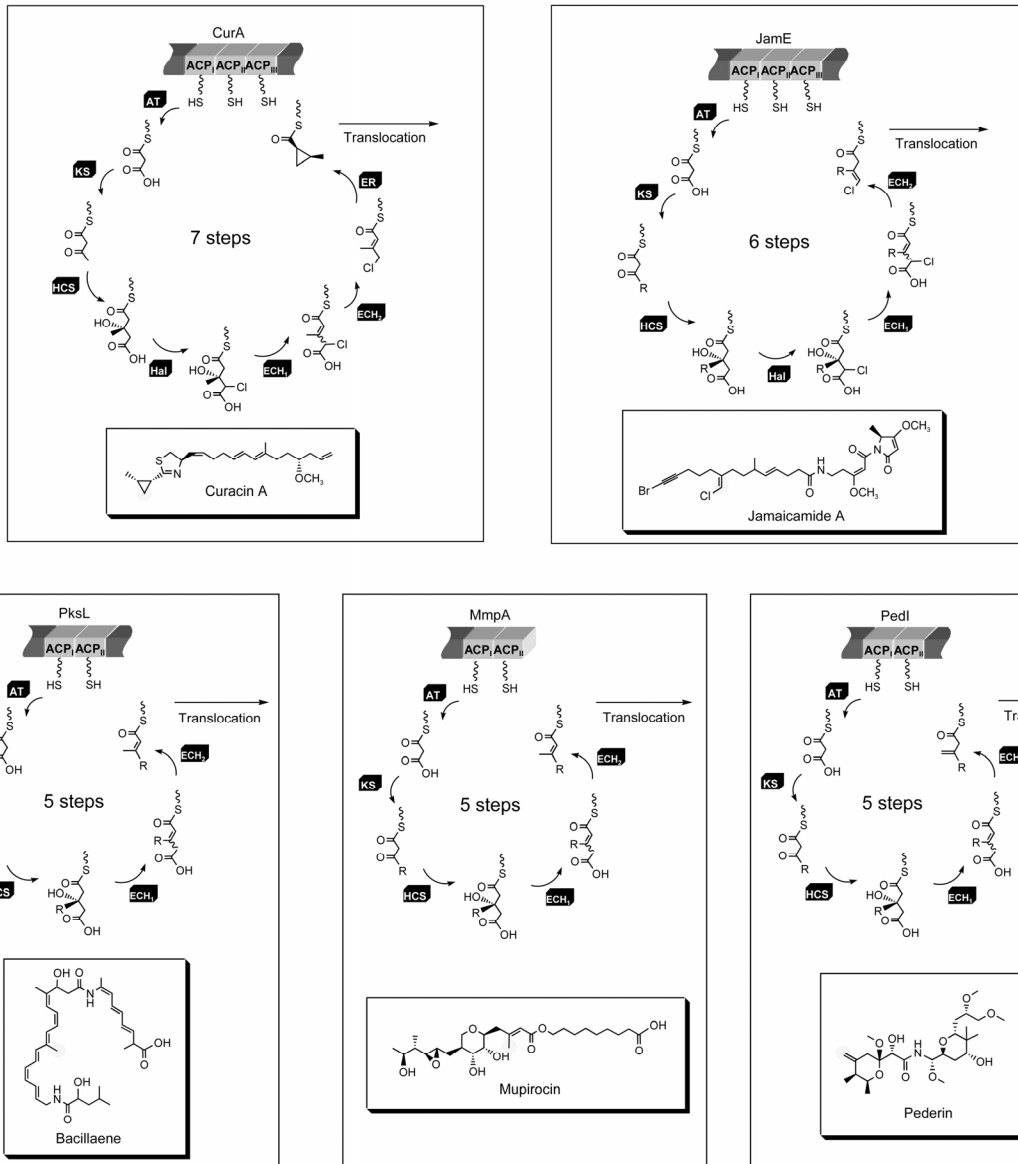
Figure S1. Hybrid PKS pathways containing the HMG enzyme cassettes with tandem ACP domains. Tandem ACP-linked malonyl substrates are subjected to consecutive enzyme modifications.
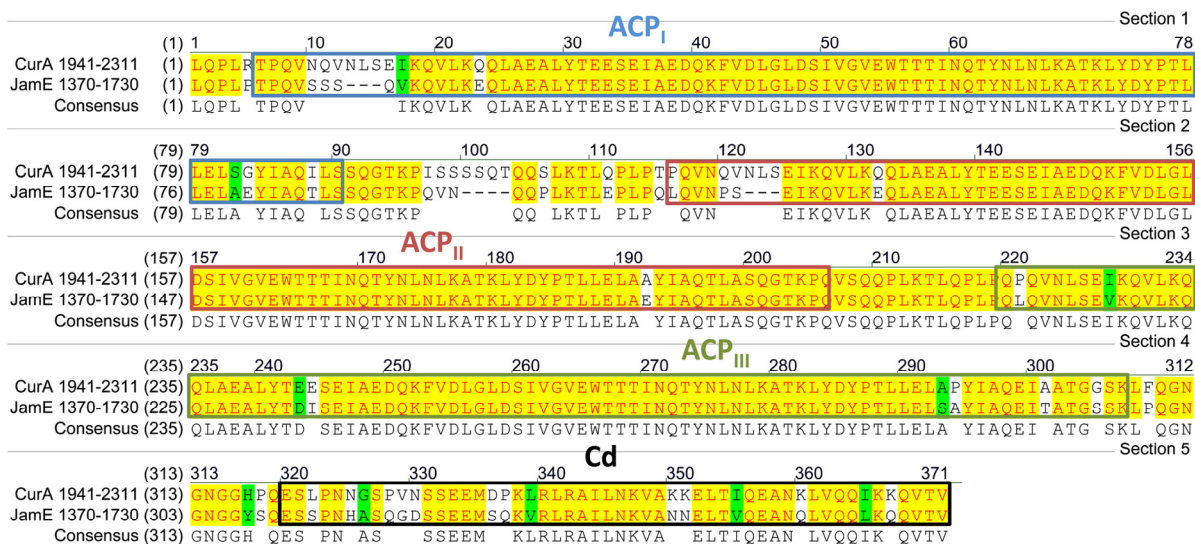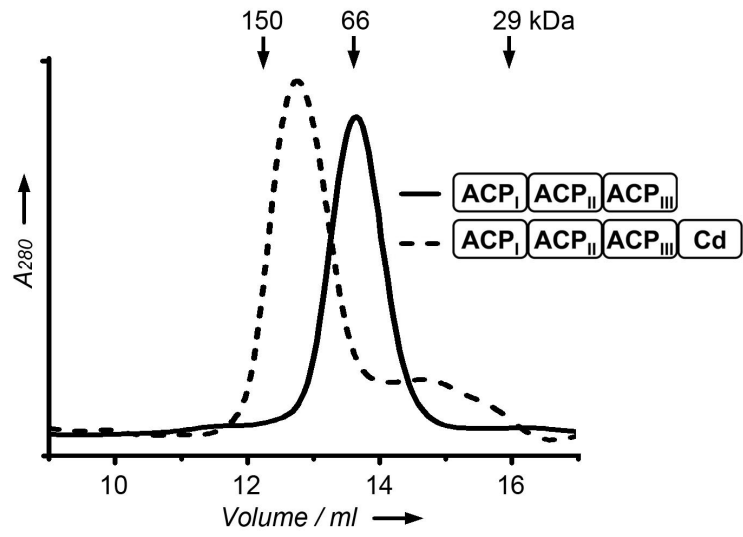
Figure S2. Sequence alignment of ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd tetradomains from the curacin A and jamaicamide pathways. The amino acid sequences start from CurA 1941 and JamE 1370.
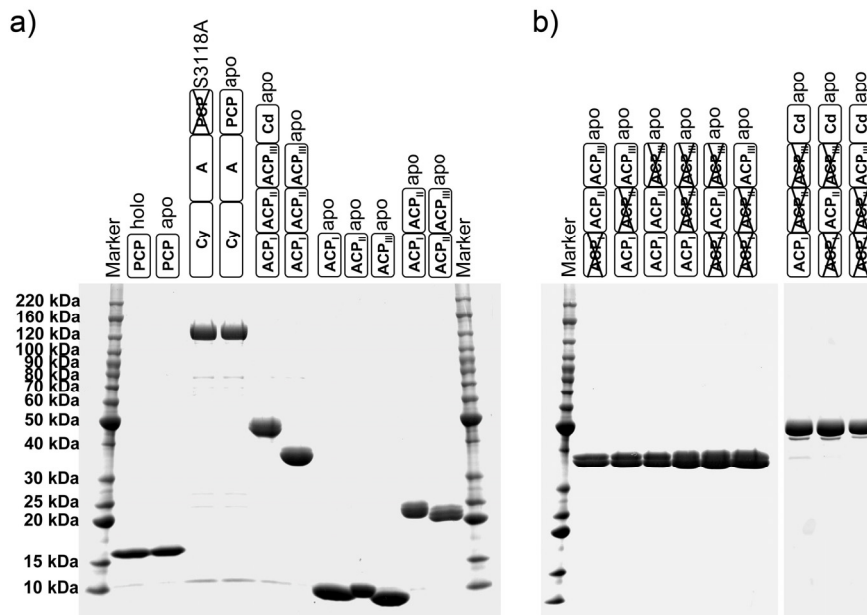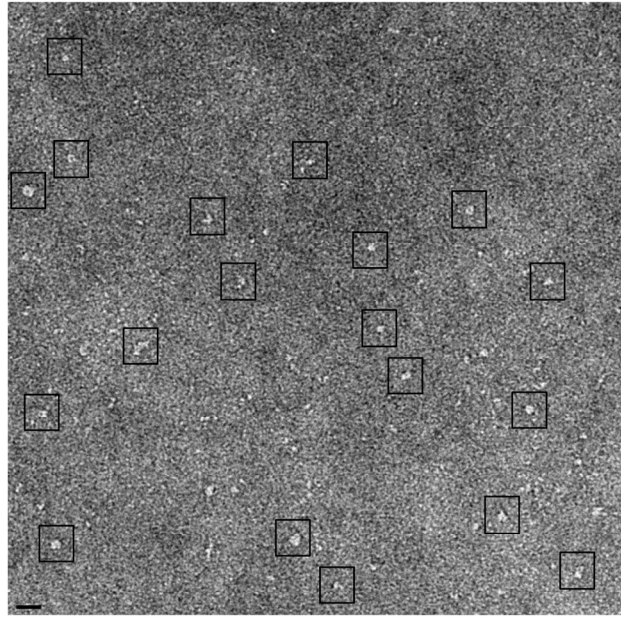
Figure S3. Coomassie blue-stained SDS-PAGE gels for purified a) protein constructs, the Ser→Ala PCP active-site mutant of Cy-A-PCP, and b) Ser→Ala ACP active-site mutants used in this study.
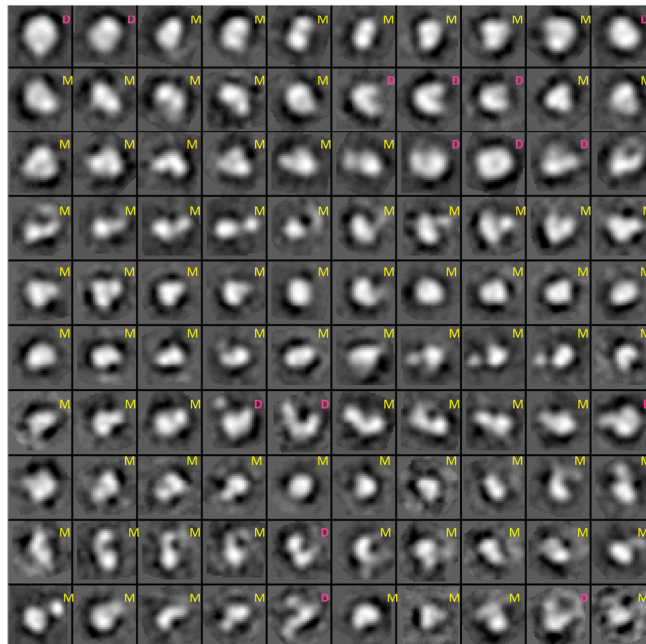
Figure S4. Comparison of size-exclusion chromatograms of CurA ACP_I-ACP_{II}-ACP_{III}-Cd and ACP_I-ACP_{II}-ACP_{III}. Elution volumes of three molecular weight standards are shown above.

a)

b)

Figure S5. EM images of ACP$_I$-ACP$_{II}$-ACP$_{III}$. a) Raw EM image of ACP$_I$-ACP$_{II}$-ACP$_{III}$ particles embedded in negative stain. Representative selected particles are shown in boxes. Scale bar = 20 nm. b) 2D classification of ACP$_I$-ACP$_{II}$-ACP$_{III}$. 4104 particle projections were subjected to classification into 100 classes. The class averages were assigned as monomers (M) or dimers (D) after visual inspection based on particle size and shape. When uncertain, no assignment was given. The box dimensions represent 16.2 x 16.2 nm.
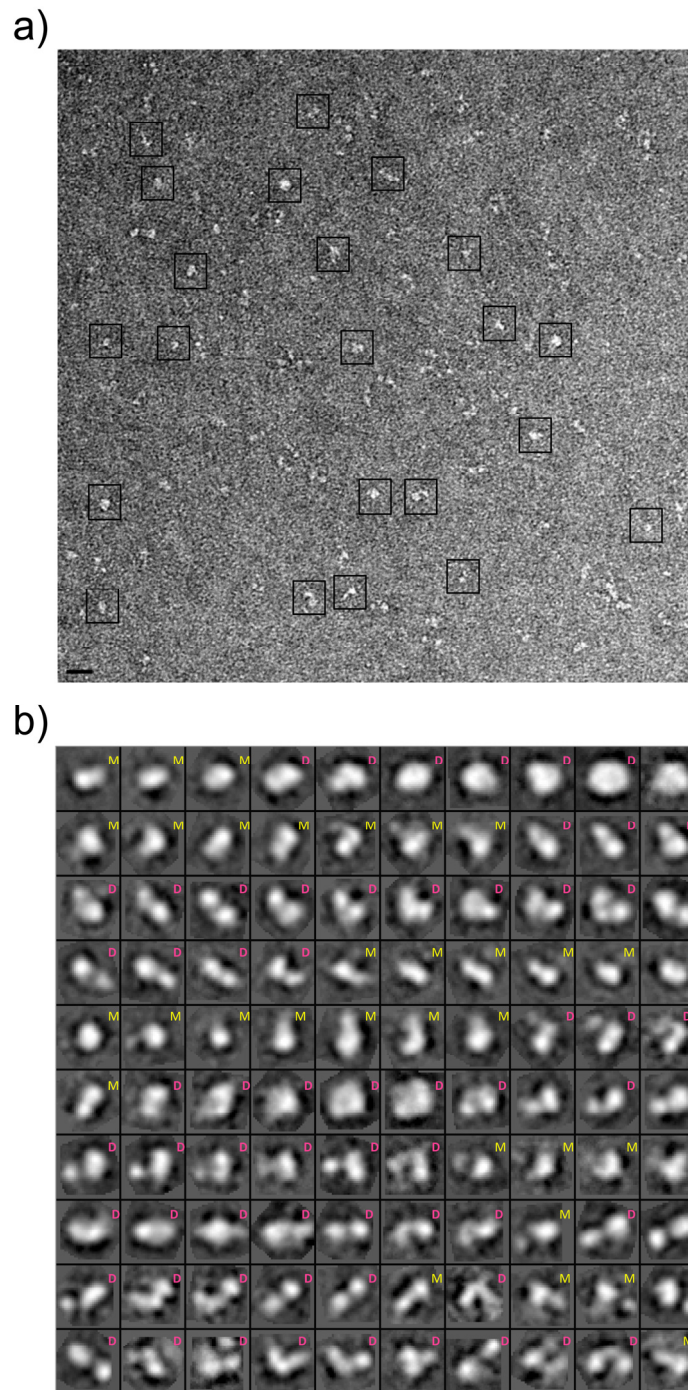
Figure S6. EM images of ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd. a) Raw EM image of ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd particles embedded in negative stain. Representative selected particles are shown in boxes. Scale bar = 20 nm. b) 2D classification of ACP$_I$-ACP$_{II}$-ACP$_{III}$-Cd. 5662 particle projections were subjected to classification into 100 classes. The class averages were assigned as monomers (M) or dimers (D) after visual inspection based on particle size and shape. When uncertain, no assignment was given. The box dimensions represent 16.2 x 16.2 nm.
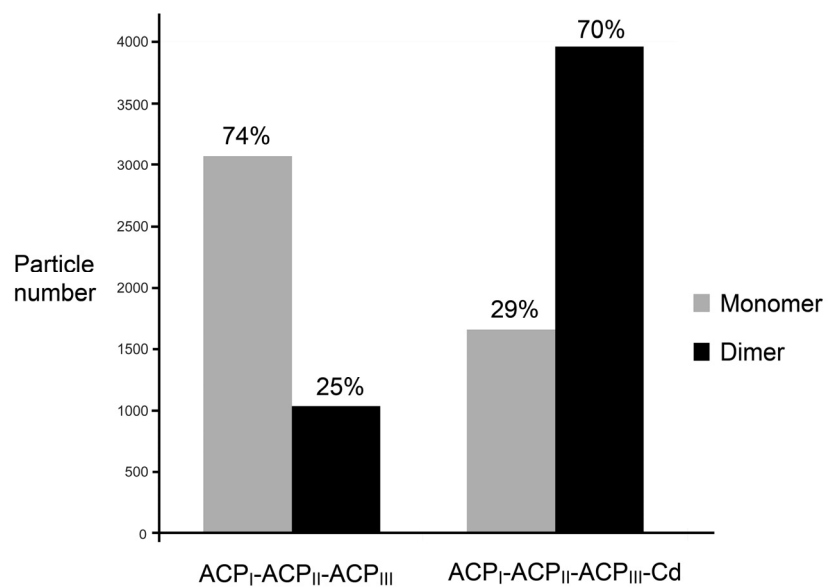
Figure S7. Comparison of monomer/dimer ratios of ACP_I-ACP_{II}-ACP_{III} and ACP_I-ACP_{II}-ACP_{III}-Cd. The unassigned particles for both proteins are ~1%.