

--- SUPPORTING INFORMATION ---

Structure and Function of Fusicoccadiene Synthase, a Hexameric Bifunctional Diterpene Synthase

Mengbin Chen,[†] Wayne K. W. Chou,[‡] Tomonobu Toyomasu,[¶] David E. Cane,[‡]
and David W. Christianson^{*,†,§}

[†]Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania,
Philadelphia, Pennsylvania, 19104-6323, United States

[‡]Department of Chemistry, Brown University, Box H, Providence, Rhode Island, 02912, United
States

[¶]Department of Bioresource Engineering, Faculty of Agriculture, Yamagata University, Wakaba-
cho 1-23, Tsuruoka, Yamagata, Japan

[§]Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138, United
States

*To whom correspondence should be addressed: Department of Chemistry, University of
Pennsylvania, 2001 Roy and Diana Vagelos Laboratories, 231 South 34th Street, Philadelphia,
PA, 19104. Tel: 215-898-5714; Fax: 215-573-2201; E-mail: chris@sas.upenn.edu

Materials and Methods

Protein expression and purification. Originally carried on a pGEX-4T-3 vector (GE Healthcare), the gene of full-length PaFS was subcloned using forward and reverse primers 5'–GCC GCC GCT CAT ATG GAG TTC AAA TAC TCG–3' and 5'–GGC TGC AGC GGA TCC TCA GAC TCG GA–3', respectively. We also prepared deletion constructs containing solely the N-terminal or C-terminal domains (general domain boundaries were inferred from straightforward amino acid sequence alignments, taking into account typical microbial terpene synthase domain lengths). We first prepared the construct containing residues 1-350 of the N-terminal domain, designated "PaFS₃₅₀", for crystallization trials and enzyme activity assays. After finding that this construct would not form crystals suitable for structure determination, we prepared the truncated construct containing residues 1-344, designated "PaFS₃₄₄", for crystallization and structure determination. The forward and reverse primers used for subcloning PaFS₃₄₄ were 5'–GCC GCC GCT CAT ATG GAG TTC AAA TAC TCG–3' and 5'–CGG AGC TCG AAT TCG GAT CCT TAC AGT CCT TGA CGC ATC C–3', respectively. The forward and reverse primers used for cloning PaFS₃₅₀ were 5'–GCC GCC GCT CAT ATG GAG TTC AAA TAC TCG–3' and 5'–GGC GGC TGC GGA TCC TTA ACA TGA CTC CAA AGA–3', respectively. As for the C-terminal domain, the construct designated "PaFS_{CT}" (residues 390-719) was prepared using forward and reverse primers 5'–GAT CCG CAT ATG CTG AGC ACT GG–3' and 5'–GGC TGC AGC GGA TCC TCA GAC TCG GA–3', respectively. All DNA constructs were inserted into a pET-28a vector (Novagen) encoding an N-terminal His-tag. The resulting clones were confirmed at the DNA sequencing facility at the Perelman School of Medicine, University of Pennsylvania.

All four constructs were expressed using the *Escherichia coli* BL21(DE3)RIL-CodonPlus strain (Novagen). Briefly, 6 × 1L LB medium containing 50 µg/mL kanamycin and 34 µg/mL chloramphenicol were inoculated with a 5 mL starting culture and grown at 37 °C with aeration. Cells were induced with 0.1-0.5 mM isopropyl β-D-1-thiogalactopyranoside at 16 °C when the

optical density at 600 nm reached 0.8-1.0. After expression for 18 h, cells were pelleted by centrifugation.

For each construct, the cell pellet was resuspended in Ni-NTA (QIAGEN) wash buffer [50 mM K_2HPO_4 (pH 7.7), 300 mM NaCl, 10% (v/v) glycerol, 3 mM tris(2-carboxyethyl)-phosphine hydrochloride (TCEP)]. Cells were lysed by sonication for 10 min at an amplitude of 50-70, with a setting of 1 s on and 2 s off. Cell debris was discarded after centrifugation, and supernatant was loaded on a Ni-NTA column. Protein was eluted with a 0-400 mM imidazole gradient in Ni-NTA wash buffer. Due to additional impurities present in the PaFS₃₄₄ preparation, this construct was additionally purified using a 10-mL HiTrap Q HP column (GE Healthcare) after exchange to Q wash buffer [10 mM Tris (pH 8.0), 10 mM NaCl]. Elution was performed with a 0-500 mM NaCl gradient in Q wash buffer. Fractions were analyzed by SDS-PAGE, and the most active ones were combined, concentrated, and loaded onto a size-exclusion column (GE Healthcare) pre-equilibrated with 50 mM Tris (pH 7.9), 300 mM NaCl, 10% (v/v) glycerol, and 3 mM TCEP. The final protein sample was 98% pure based on SDS-PAGE analysis and was concentrated to 12-34 mg/mL.

Crystallization and structure determination of the C-terminal chain elongation domain, PaFS_{CT}. Although extensive efforts were made to crystallize full-length PaFS, all attempts failed. We hypothesized that the polypeptide linker between the two α domains of full-length PaFS might be excessively flexible, thereby hindering crystallization, so we pursued an alternative strategy in which we crystallized the individual N-terminal and C-terminal domains prepared as described above. Crystallization of the ligand-bound C-terminal domain PaFS_{CT} was achieved by the sitting-drop vapor diffusion method at 4 °C. Typically, a 1- μ L drop of protein solution [12 mg/mL PaFS_{CT}, 50 mM Tris (pH 7.9), 300 mM NaCl, 10% (v/v) glycerol, 3 mM TCEP, 5 mM $MgCl_2$, 50 mM pamidronate (Sigma-Aldrich)] was added to a 1- μ L drop of precipitant solution [0.01 M $CoCl_2$, 0.1 M sodium acetate (pH 4.6), 1.0 M 1,6-hexanediol

(Hampton Research)] and equilibrated against a 70- μ L reservoir of precipitant solution. Crystals generally formed in 2-3 weeks.

The unliganded form of PaFS_{CT} was crystallized by adding a 1- μ L drop of protein solution [12 mg/mL PaFS_{CT}, 50 mM Tris (pH 7.9), 300 mM NaCl, 10% (v/v) glycerol, 3 mM TCEP] to a 1- μ L drop of precipitant solution [1.5 M ammonium phosphate, 0.1 M Tris (pH 8.5) (Hampton Research)] and equilibrating it against a 70- μ L reservoir of precipitant solution. Crystals generally formed in 2-3 weeks.

After growing to maximum dimensions, crystals of liganded and unliganded PaFS_{CT} were transferred to a cryoprotectant solution (26% [50% PEG 400, 50% 2-methyl-2,4-pentanediol], 74% mother liquor) and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at beamline X29A of the National Synchrotron Light Source (Brookhaven National Laboratory) and were processed using HKL2000.¹ Crystals of the unliganded protein belonged to space group *P*321, unit cell parameters $a = b = 104.9$ Å, $c = 140.5$ Å; with 2 monomers in the asymmetric unit, the Matthews coefficient $V_M = 3.0$ Å³/Da (59% solvent content). Crystals of the liganded protein belonged to space group *C*222₁, unit cell parameters $a = 103.6$ Å, $b = 216.1$ Å, $c = 129.2$ Å; with 3 monomers in the asymmetric unit, $V_M = 3.3$ Å³/Da (63% solvent content).

The crystal structure of liganded PaFS_{CT} was solved by molecular replacement using Phaser for Molecular Replacement in the CCP4 package² with a monomer of human GGPP synthase (PDB accession code: 2Q80)³ as a search probe. Iterative cycles of refinement and manual adjustment of the model were performed with PHENIX and COOT, respectively.^{4,5} Pamidronate, Co²⁺ ions, and water molecules were added to the model in the later stages of refinement. The following disordered segments were characterized by broken or missing electron density and accordingly were omitted from the final model: S521 and D578-P582 (chain A), S521-Q523 (chain B), and H639-R642 (chain C). Final refinement statistics are recorded in Table S1.

The crystal structure of unliganded PaFS_{CT} was solved by molecular replacement in similar fashion using a monomer of the liganded PaFS_{CT} structure as a search probe. Iterative cycles of refinement and manual adjustment of the model were performed with PHENIX and COOT, respectively.^{4,5} Water molecules were added to the model in the later stages of refinement. The following disordered segments were characterized by broken or missing electron density and accordingly were omitted from the final model: S517-E520 and V611-E627 (chain A), and A518-E520 and S612-G628 (chain B). Final refinement statistics are recorded in Table S1.

Crystallization and structure determination of the N-terminal cyclase domain, PaFS₃₄₄. Initial crystallization of PaFS₃₄₄ was achieved by the sitting-drop vapor diffusion method at 4 °C. Typically, a 1- μ L drop of protein solution [34 mg/mL PaFS₃₄₄, 50 mM Tris (pH 7.9), 300 mM NaCl, 10% (v/v) glycerol, 3 mM TCEP, 5 mM MgCl₂, 30 mM pamidronate (Sigma-Aldrich)] was added to a 1- μ L drop of precipitant solution [0.1 M sodium acetate (pH 4.8), 1.75 M sodium formate, 0.25 M magnesium formate] and equilibrated against a 70- μ L reservoir of precipitant solution. Crystals generally appeared within 2-3 weeks. After growing to maximum dimensions, crystals were transferred to cryoprotectant solution (26% [50% PEG 400, 50% 2-methyl-2,4-pentanediol], 74% mother liquor) and flash-cooled in liquid nitrogen.

These crystals diffracted X-rays to 2.3 Å resolution at the synchrotron and belonged to space group *P*3₂21, unit cell parameters *a* = *b* = 143.3 Å, *c* = 118.3 Å; with 2 molecules in the asymmetric unit, *V*_M = 4.2 Å³/Da (70.5% solvent content). However, deviations from ideal intensity statistics indicated a significant twinning fraction of 41%, which would hinder *de novo* phasing efforts. Accordingly, crystallization conditions were systematically modified with the goal of generating an alternative crystal form. Specifically, buffer components of PaFS₃₄₄ solutions were optimized through the ThermoFluor assay to find conditions under which the protein would

be more stable, reasoning that such conditions might facilitate the formation of an alternative crystal form.

We successfully identified conditions and a procedure that yielded improved crystals. A protein solution of 27 mg/mL PaFS₃₄₄, 30 mM tricine (pH 7.9), 256 mM KCl, 3 mM TCEP, and 10% glycerol was incubated with 8 mM MnCl₂ and 1.3 mM neridronate for 1 h. Immediately prior to setting up crystallization trays, porcine elastase (1:6000 molar ratio with PaFS₃₄₄) was added to the protein solution to enable *in situ* limited proteolysis in drops. Addition of elastase slowed down crystal growth, and we reasoned that proteolytic nicking might facilitate crystallization.⁶ Crystallization of liganded PaFS₃₄₄ was performed at 4 °C by the sitting-drop vapor diffusion method using an Innovadyne Nanodrop Express crystallization robot. Equal volumes (900 nL) of protein and precipitant solution [1.05 M sodium formate, 0.075 M manganese formate, 0.1 M sodium acetate (pH 4.4)] were equilibrated against a 50-μL reservoir of precipitant solution. Crystals appeared within 4-6 weeks and grew to maximal dimensions approximately 1 week after they initially appeared. Crystals were subsequently transferred to cryoprotectant solution (28% [50% PEG 400, 50% 2-methyl-2,4-pentanediol], 72% mother liquor) and flash-cooled in liquid nitrogen.

X-ray diffraction data from these new crystals of PaFS₃₄₄ were collected at the Stanford Synchrotron Radiation Lightsource, beamline 14-1. Crystals belonged to space group *P*3₂21, unit cell parameters $a = b = 142.6 \text{ \AA}$, $c = 117.6 \text{ \AA}$; with 2 molecules in the asymmetric unit, $V_M = 4.1 \text{ \AA}^3/\text{Da}$ (70% solvent content). Analysis of intensity statistics indicated a lower twinning fraction of 22%, which facilitated phasing through single-wavelength anomalous diffraction (SAD) of the Mn²⁺ ions.

For SAD phasing, a dataset to 3.0 Å resolution was collected at the Mn peak wavelength (1.89166 Å) and processed with HKL2000. Initially, 4 Mn sites were identified by ShelxD from HKL2MAP⁷ through single-wavelength anomalous diffraction, and were used for search and

refinement of 4 additional Mn sites. Solvent flattening and initial polyaniline model building yielded a 50% complete trace of the polypeptide chain using ShelxE from HKL2MAP,⁷ which was subsequently used as input for Phenix.autobuild to generate a 90% complete model. Following manual model adjustment and registration of the amino acid sequence using COOT,⁵ this model was used as the search probe for molecular replacement calculations with a 2.5 Å resolution dataset collected at a wavelength of 1.18076 Å from the same crystal. Iterative model adjustment and refinement were performed with COOT and PHENIX, respectively.^{4,5} The twin law (-h, -k, l) was determined using Phenix.xtriage and incorporated at the beginning of refinement. Refinement statistics are recorded in Table S1. Segments characterized by uninterpretable broken or missing electron density were omitted from the final model as follows: H97–A134 and F165–E175 in chain A, and H97–R133 and Q173–H174 in chain B. These segments could be disordered and/or clipped by the protease added to facilitate crystallization.

The refined 2.5 Å resolution structure of the PaFS₃₄₄-Mn²⁺₃-neridronate complex served as a molecular replacement probe to solve the 2.3 Å-resolution structure of the PaFS₃₄₄-Mg²⁺₃-pamidronate complex from highly twinned crystals. The twin law (-h,-k, l) was determined using Phenix.xtriage, and refinement and model building was performed using PHENIX and COOT, respectively.^{4,5} Refinement statistics are recorded in Table S1. Segments characterized by uninterpretable broken or missing electron density were omitted from the final model as follows: H97–A134 and R172–E175 in chain A, and E103–V129 and Q173–H174 in chain B. Since these polypeptide segments in the protease-free PaFS₃₄₄-Mg²⁺₃-pamidronate complex generally correlate with those observed in the protease-treated PaFS₃₄₄-Mn²⁺₃-neridronate complex, these segments are likely to be inherently disordered since protease treatment does not induce significant additional disorder.

Enzyme activity measurements. Cyclization activity assays on a 100-μL scale were conducted in triplicate, as typically performed in our laboratory with other terpenoid cyclases,⁸ using 7.5–75 nM full-length PaFS or PaFS₃₅₀ in assay buffer [25 mM HEPES (pH 8.0), 200 mM

NaCl (omitted for PaFS₃₅₀ assays), 5 mM MgCl₂]. Concentration ranges of [1-³H]-geranyl diphosphate (GPP), [1-³H]-farnesyl diphosphate (FPP), and [1-³H]-GGPP (American Radiolabeled Chemicals, Inc.) were adjusted accordingly. Reactions were initiated by adding enzyme at 30 °C, and quenched with 30 µL 0.5 M ethylenediaminetetraacetic acid (EDTA). A 1-mL aliquot of hexane was used to extract hydrocarbon products twice with vigorous vortexing and subsequently loaded onto a 1-mL silica gel column. The column was washed with 2 mL hexane and 1 mL ethyl ether. Eluate was collected and radioactivity counts were determined using a Beckman liquid scintillation counter. Control experiments were performed to determine background counts. Data were processed with Prism software.

To characterize chain elongation activity, assays with full-length D92A PaFS or PaFS_{CT} on a 100-µL scale were conducted in triplicate by the method of Holloway and Popjak with slight modifications.⁹ Full-length D92A PaFS was completely inactive when incubated with GGPP and IPP, as expected. The concentration of [4-¹⁴C]-IPP (American Radiolabeled Chemicals, Inc.) was fixed at 50 µM, and concentrations of DMAPP, GPP, and FPP were adjusted accordingly to determine the kinetics of chain elongation activity starting with 5-, 10-, and 15-carbon isoprenoid diphosphate substrates, respectively. Briefly, reactions were initiated by adding enzyme at 30 °C, and quenched with 10 µL 6 M HCl. Newly generated prenyl diphosphate products were acid-hydrolyzed at 37 °C for 30 min, and subsequently neutralized with 16 µL of 5.5 M NaOH. The resulting prenyl alcohol was extracted with 1 mL hexane twice by vigorous vortexing, and was transferred into scintillation vials to determine radioactivity counts. Background readings were subtracted by conducting control experiments.

The overall catalytic activity for chain elongation and cyclization catalyzed by full-length PaFS was assessed using substrates DMAPP and [4-¹⁴C]-IPP. The concentration of [4-¹⁴C]-IPP was maintained at 50 µM, and that of DMAPP ranged from 3.4 µM to 26.9 µM. Reactions on a 100-µL scale using 1.49 µM of full-length PaFS were conducted at 30 °C, and quenched with 30

μL of 0.5 M EDTA, in triplicate. Hydrocarbon products were extracted with 1 mL hexane twice, and subsequently transferred to scintillation vials for radioactivity counting, with background readings subtracted.

Product assays were carried out in duplicate on a 5-mL scale using a previously-described method, slightly modified to minimize the loss of volatile products.⁸ Reactions were initiated at room temperature by adding enzyme into assay buffer containing substrate (GPP, FPP, or GGPP) overlaid with 5 mL *n*-pentane. After quenching by EDTA after 18-20 h of reaction, hydrocarbon products were extracted from the reaction mixture with three portions of *n*-pentane and the extracts run through anhydrous MgSO₄ to eliminate trace amounts of water. Samples were then concentrated under vacuum in an ice water bath using a rotary evaporator and analyzed by gas chromatography-mass spectrometry (GC-MS).

GC-MS analyses were performed using an Agilent 5977A Series GC/MSD instrument (70 eV, electron impact), a 1 μL injection volume, and a 3 min solvent delay. GC-MS conditions used a temperature program with a 2 min hold at 60 °C, a 20 °C min⁻¹ increment to 280 °C, followed by a 2 min hold at 280 °C, and a HP-5ms capillary column (0.25 mm ID x 30 m length x 0.25 mm film, Agilent Technologies). Comparison of GC-MS detected compounds to their standards was done using the MassFinder 4.2.1 program (<http://www.massfinder.com>). Retention indices of compounds were measured using C₈-C₂₀ and C₁₀-C₄₀ alkane standards.

Sedimentation equilibrium experiments. To determine the molecular weight of the intact PaFS oligomer, sedimentation equilibrium experiments were performed using an XL-A analytical ultracentrifuge (Beckman-Coulter). Two buffer conditions were used: 50 mM Tris (pH 8.5), 200 mM NaCl, 3 mM TCEP; and 10 mM Tris (pH 7.9), 200 mM NaCl, 3 mM TCEP. Sedimentation equilibrium data were collected at 4°C with absorbance detection at 280 nm for 1–3 sample concentrations in separate experiments run at 6,000 rpm (2898 x *g*) and 9,000 rpm (6520 x *g*). Data analysis was performed with HeteroAnalysis; data were fit using an ideal

single-species model at two speeds for each concentration.

Small-angle X-ray scattering (SAXS). X-ray scattering data were collected at the SIBYLS beamline at the Advanced Light Source (Berkeley, CA, USA).^{10,11} Full-length PaFS exhibited undesirable aggregation behavior in solution, so we collected data from N333A/Q336A PaFS which was better behaved. A 3.3 mg/mL sample of full-length N333A/Q336A PaFS in buffer [50 mM Tris (pH 8.5), 200 mM NaCl, 3 mM TCEP, and 5% glycerol] was exposed for 0.5 second at 283 K. Scattering from a matching buffer solution was subtracted from the data, and corrected for the incident intensity of X-rays. Data quality examination and further analyses were performed with PRIMUS and GNOM.^{12,13} The *de novo* molecular envelope was calculated with GASBOR; since the initial envelope displayed prominent three-fold symmetry, three-fold averaging was imposed.¹⁴ Positioning of the molecular model was performed with SASREF using crystallographic dimers or hexamers of PaFS_{CT} and crystallographic dimers or monomers of PaFS₃₄₄ and PaFS_{CT} as described in the text.¹⁵ The molecular weight of 524 kD derived from $I(0)$ corresponds most closely to a hexamer ($M_r = 502$ kD), consistent with the hexameric quaternary structure determined for full-length PaFS using analytical ultracentrifugation (Figure S1). Data collection statistics for SAXS experiments are recorded in Table S2.

References

1. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology. Macromolecular Crystallography (Part A)*, Carter, C. W., Jr., and Sweet, R. M., Eds.; Vol. 276, pp 307-326, Academic Press, New York.
2. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* 40, 658-674.
3. Kavanagh, K. L., Dunford, J. E., Bunkoczi, G., Russell, R. G. G., and Oppermann, U. (2006) The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a

- novel hexameric arrangement and inhibitory product binding. *J. Biol. Chem.* **281**, 22004-22012.
4. 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., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D* **66**, 213-221.
 5. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sec. D* **60**, 2126-2132.
 6. Dong, A., Xu, X., Edwards, A. M., Midwest Center for Structural Genomics, and Structural Genomics Consortium. (2007) *In situ* proteolysis for protein crystallization and structure determination. *Nat. Methods*, **4**, 1019-1021.
 7. Pape, T. and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.*, **37**, 843-844.
 8. Li, R., Chou, W. K. W., Himmelberger, J. A., Litwin, K. M., Harris, G. G., Cane, D. E., and Christianson, D. W. (2014) Reprogramming the chemodiversity of terpenoid cyclization by remolding the active site contour of epi-isozizaene synthase. *Biochemistry*, **53**, 1155-1168.
 9. Holloway, P. W. and Popjak, G. (1967) The purification of 3,3-dimethylallyl- and geranyl-transferase and of isopentenyl pyrophosphate isomerase from pig liver. *Biochem. J.* **104**, 57-70.
 10. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., Tsutakawa, S. E., Jenney, F. E., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S., Scott, J. W., Dillard, B. D., Adams, M. W. W., and Tainer, J. A. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods*. **6**, 606–612.

11. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) Solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**,191-285.
12. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003). PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J. Appl. Cryst.* **36**, 1277-1282.
13. Semenyuk, A., and Svergun, D. (1991) GNOM - A program package for small-angle scattering data processing. *J. Appl. Cryst.* **24**, 537–540.
14. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. J. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946-2953.
15. Petoukhov, M. V. and Svergun, D. I. (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237-1250.

Table S1. Crystallographic data collection and refinement statistics

Complex	PaFS_{CT} (unliganded)	PaFS_{CT}-Co²⁺₃- pamidronate	PaFS₃₄₄-Mn²⁺₃- neridonate	PaFS₃₄₄-Mg²⁺₃- pamidronate
Resolution limits (Å)	50.0-2.43	50.0-2.55	43.4-2.50	50.0-2.30
Space group	<i>P</i> 321	<i>C</i> 222 ₁	<i>P</i> 3 ₂ 21	<i>P</i> 3 ₂ 21
Unit cell parameters a, b, c (Å)	104.9, 104.9, 140.5	103.6, 216.1, 129.2	142.6, 142.6, 117.6	143.3, 143.3, 118.3
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 120	90, 90, 120
Total/unique reflections	103264/32406	536690/47383	402988/48128	1186692/62273
Twin fraction	0.0	0.0	0.22	0.41
$R_{\text{merge}}^{a,b}$	0.153 (0.621)	0.120 (0.676)	0.106 (---) ^h	0.101 (---) ^h
$R_{\text{meas}}^{a,c}$	0.175 (0.743)	0.123 (0.738)	0.113 (---) ^h	0.104 (---) ^h
$R_{\text{pim}}^{a,d}$	0.083 (0.394)	0.036 (0.286)	0.043 (0.656)	0.029 (0.453)
$CC_{1/2}^{a,e}$	0.950 (0.717)	0.986 (0.825)	0.863 (0.489)	0.967 (0.636)
$I/\sigma(I)^a$	12.4 (1.3)	21.0 (2.0)	30.0 (1.1)	37.2 (1.3)
Redundancy	3.2 (2.7)	11.3 (5.7)	8.4 (8.3)	19.1 (12.8)
Completeness (%) ^a	95.2 (94.6)	99.5 (95.5)	100.0 (100.0)	99.9 (99.6)
Reflections used in work set/test set	32390/1626	47336/2389	48098/2383	62219/3177
R_{work}^f	0.223	0.214	0.207	0.178
R_{free}^f	0.276	0.256	0.238	0.206
Protein atoms ^g	4635	7206	4853	5002
Ligand atoms ^g	0	39	32	26
Co ²⁺ ions ^g	0	12	0	0
Mn ²⁺ ions ^g	0	0	9	0
Mg ²⁺ ions ^g	0	0	0	6
Cl ⁻ ions ^g	0	0	2	2
Solvent atoms ^g	89	140	76	185
R.m.s. deviations				
Bonds (Å)	0.005	0.005	0.012	0.004
Angles (°)	0.9	0.9	1.3	0.9
Average B factors (Å ²)				
Main chain	63	55	69	57
Side chain	67	58	74	62
Solvent	62	49	72	57
Ligand	---	63	72	59
Ramachandran Plot(%)				
Allowed	93.8	93.3	88.9	89.9
Additionally allowed	5.6	6.5	11.1	9.9
Generously allowed	0.6	0.1	0.0	0.2
Disallowed	0.0	0.0	0.0	0.0

^aNumbers in parentheses refer to the highest resolution shell of data.

^b R_{merge} for replicate reflections, $R = \sum |I_h - \langle I_h \rangle| / \sum \langle I_h \rangle$; I_h = intensity measure for reflection h ; and $\langle I_h \rangle$ = average intensity for reflection h calculated from replicate data.

^c $R_{\text{meas}} = \sum [n/(n-1)]^{1/2} |I_h - \langle I_h \rangle| / \sum \langle I_h \rangle$; n is the number of observations (redundancy).

^d $R_{\text{pim}} = \sum [1/(n-1)]^{1/2} |I_h - \langle I_h \rangle| / \sum \langle I_h \rangle$; n is the number of observations (redundancy).

^e $\text{CC}_{1/2} = \sigma_e^2 / (\sigma_e^2 + \sigma_i^2)$, where σ_e^2 is the true measurement error variance and σ_i^2 is the independent measurement error variance.

^f $R_{\text{work}} = \sum ||F_o| - |F_c|| / \sum |F_o|$ for reflections contained in the working set. $R_{\text{free}} = \sum ||F_o| - |F_c|| / \sum |F_o|$ for reflections contained in the test set. $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

^gPer asymmetric unit.

^h R_{merge} and R_{meas} values higher than 1.000 are reported as 0.000 by HKL2000; this is sometimes observed for datasets with high redundancy.

Table S2. SAXS data collection statistics and structural parameters

	Full-length N333A/Q336A PaFS
<i>Data collection</i>	
Instrument	SIBYLS 12.3.1
Defining slits size (H mm x V mm)	4x1
Detector distance (m)	1.5
Wavelength (Å)	1.03
q range (Å ⁻¹)	0.008–0.55
Exposure time (s)	0.5
Temperature (K)	283
<i>Structural parameters</i>	
I(0) (cm ⁻¹) from Guinier plot	2072
R _g (Å) from Guinier plot	53.7
I(0) (cm ⁻¹) from P(r)	2051
R _g (Å) from P(r)	52.5
D _{max} (Å) from P(r)	155.4
Porod volume (Å ³) from PRIMUS	931501

Table S3. Protein-metal ion interactions in PaFS_{CT} and PaFS₃₄₄

PaFS_{CT}-Co²⁺₃-Pamidronate Complex					
Co ²⁺ -Ligand Distances (Å)					
	Co ²⁺ _C		Co ²⁺ _A		Co ²⁺ _B
D474 OD2	2.0	D478 OD2	2.2	Pamidronate O5	2.0
Pamidronate O10	2.1	Pamidronate O3	2.0	Pamidronate O12	1.9
D478 OD2	2.1	Pamidronate O10	2.0	D605 OD2	2.1
w1	1.9	D474 OD1	2.0	w1	2.0
w2	2.0	w1	2.0	w2	2.0
w3	2.0	w2	2.3	w3	2.3

PaFS₃₄₄-Mn²⁺₃-Neridronate Complex					
Mn ²⁺ -Ligand Distances (Å)					
	Mn ²⁺ _C		Mn ²⁺ _A		Mn ²⁺ _B
D92 OD2	1.9	D96 OD2	2.7	Neridronate O2	2.1
Neridronate O3	2.6	Neridronate O3	1.8	Neridronate O9	1.9
D96 OD2	1.9	Neridronate O10	1.7	S236 OG	2.3
w1	1.8	D92 OD1	2.1	N232 OD1	2.0
w2	1.8	w1	1.8	E240 OE2	2.2
w3	2.4	w2	2.2	E240 OE1	2.3
w4	2.0			w1	2.0

PaFS₃₄₄-Mg²⁺₃-Pamidronate Complex					
Mg ²⁺ -Ligand Distances (Å)					
	Mg ²⁺ _C		Mg ²⁺ _A		Mg ²⁺ _B
D92 OD2	2.1	D96 OD2	2.6	Pamidronate O2	2.4
Pamidronate O3	2.6	Pamidronate O3	1.7	Pamidronate O9	2.3
D96 OD2	2.4	Pamidronate O10	2.1	OG S236	2.3
w1	1.9	D92 OD1	2.0	OD1 N232	2.2
w2	2.3	w1	2.2	OE2 E240	1.9
w3	2.4	w2	2.3	w1	1.9

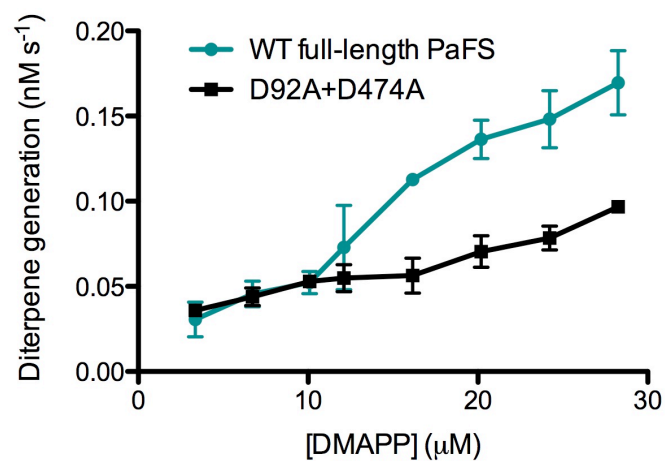


Figure S1: Dependence of the rate of diterpene generation on DMAPP concentration as catalyzed by 1.49 μM full-length wild-type PaFS, and 2.03 μM D92A PaFS and D474A PaFS (normalized to wild-type PaFS concentration).

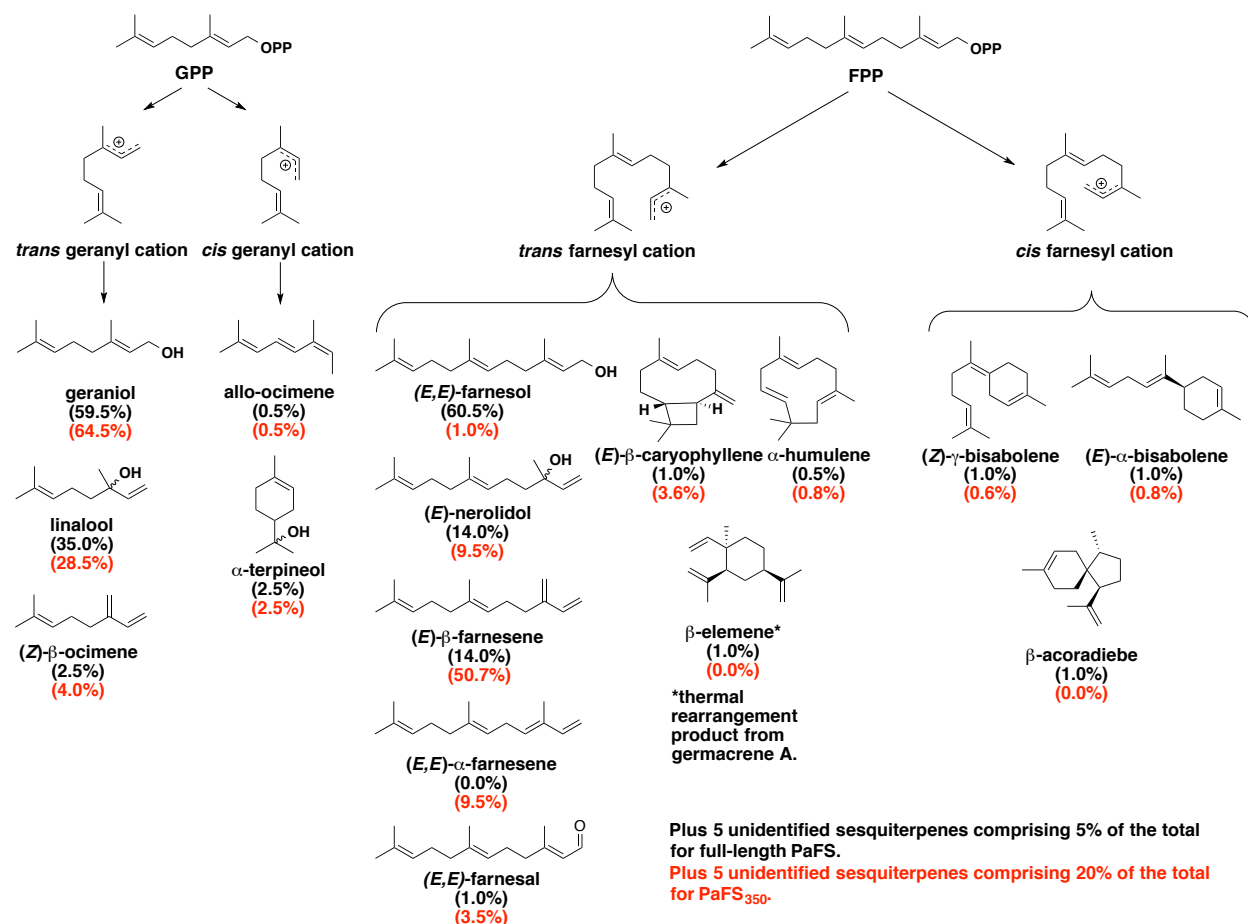


Figure S2. Product arrays generated by full-length PaFS (black) and PsFS₃₅₀ (red) for cyclization reactions utilizing GPP and FPP as substrates. The presence of trace amounts of (E,E)-farnesal in assays with FPP likely results from an aberrant oxidation reaction with (E,E)-farnesol, perhaps catalyzed by trace amounts of an *E. coli* alcohol dehydrogenase impurity in the protein samples.

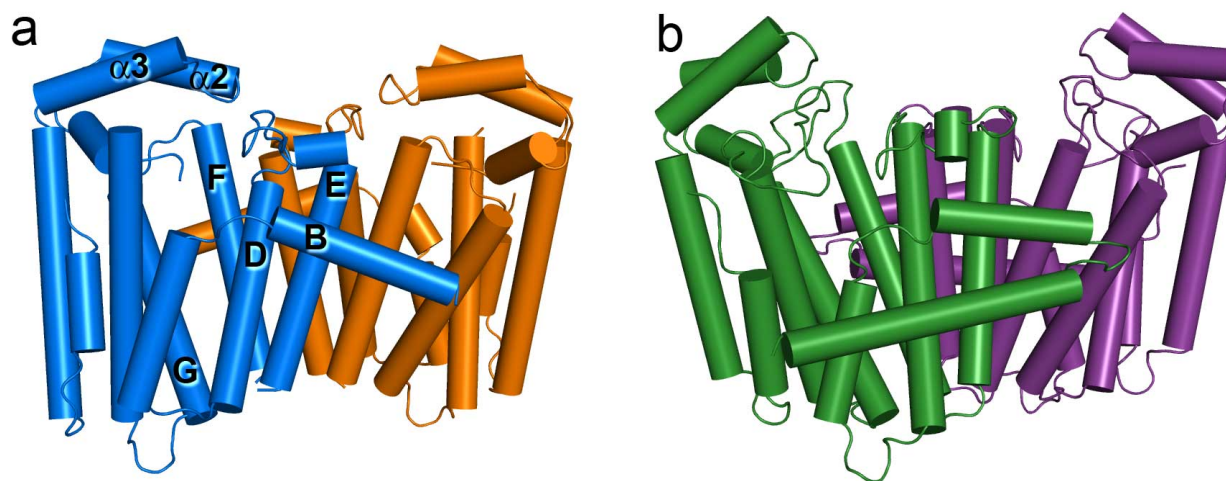


Figure S3. Comparison of PaFS_{CT} dimer architecture (left) with that of avian farnesyl diphosphate synthase (right; PDB 1FPS). Selected helices in PaFS_{CT} are labeled as discussed in the text.

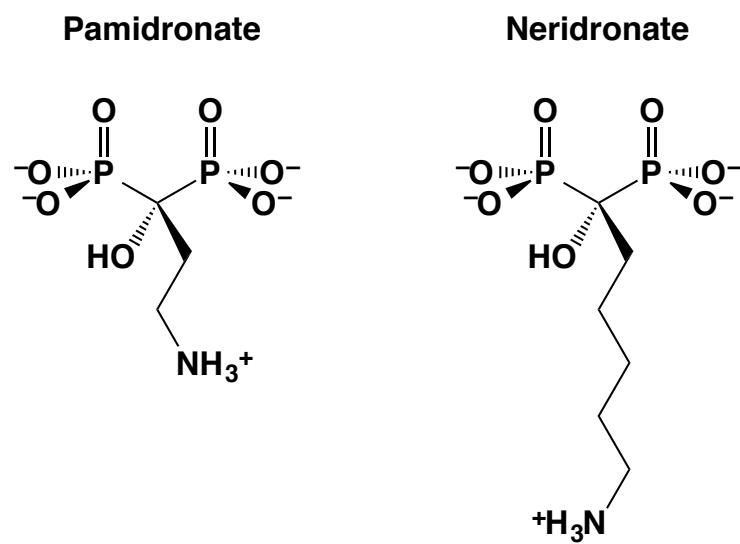


Figure S4. Molecular structures of the bisphosphonate substrate analogues used in the current study.

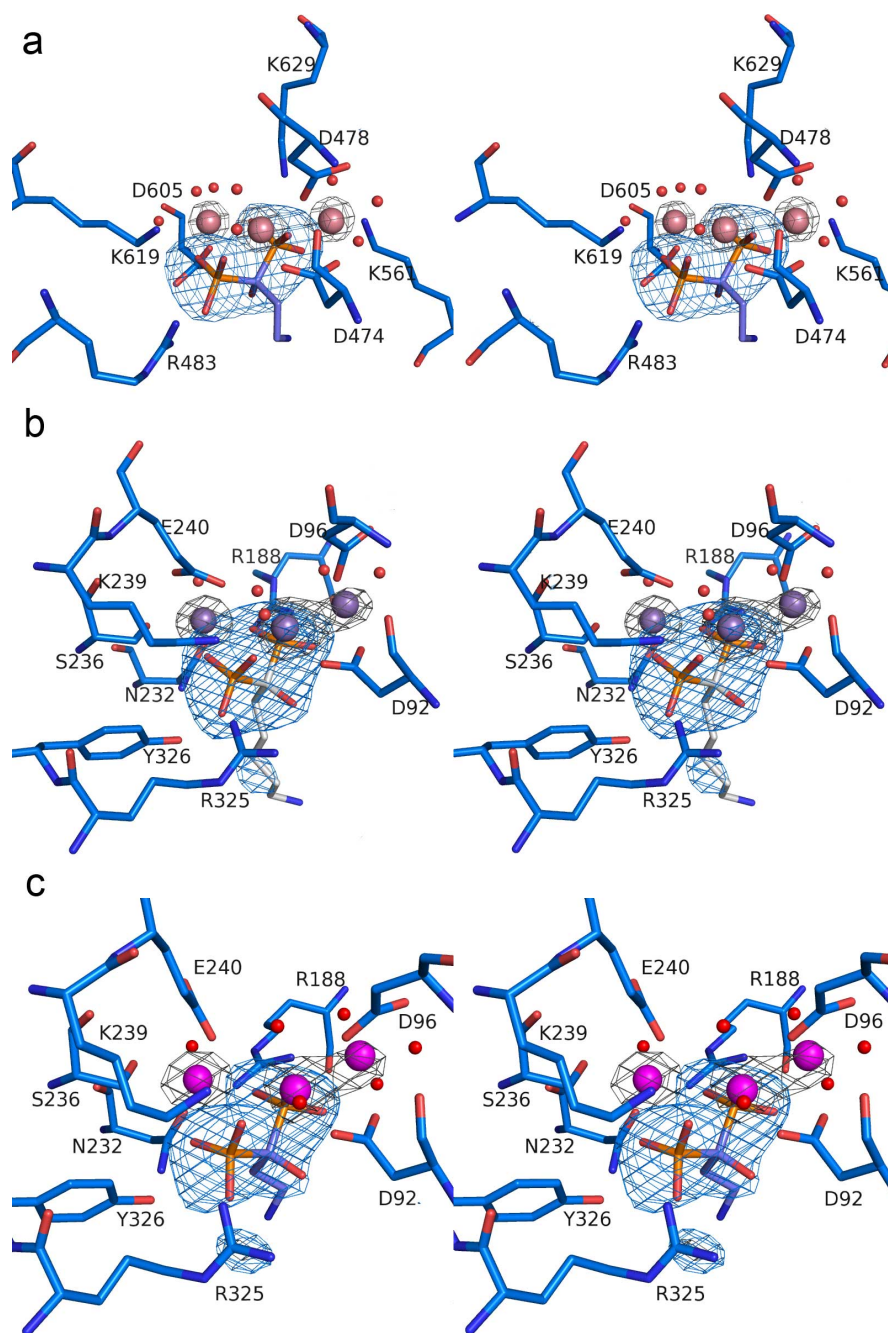


Figure S5. (a) Omit maps of Co^{2+} ions (contoured at 18.5σ) and pamidronate (contoured at 4.5σ) in the $\text{PaFS}_{\text{CT}}\text{-Co}^{2+}_3$ -pamidronate complex. The pendant alkylamino group of pamidronate is partially disordered in the active site cavity. (b) Omit maps of Mn^{2+} ions (contoured at 10.5σ) and neridronate (contoured at 4.5σ) in the $\text{PaFS}_{344}\text{-Mn}^{2+}_3$ -neridronate complex. The pendant alkylamino group of neridronate is disordered in the active site cavity. (c) Omit maps of Mg^{2+} ions (contoured at 6.2σ) and pamidronate (contoured at 4.5σ) in the $\text{PaFS}_{344}\text{-Mg}^{2+}_3$ -pamidronate complex. The pendant alkylamino group of pamidronate is partially disordered in the active site cavity.

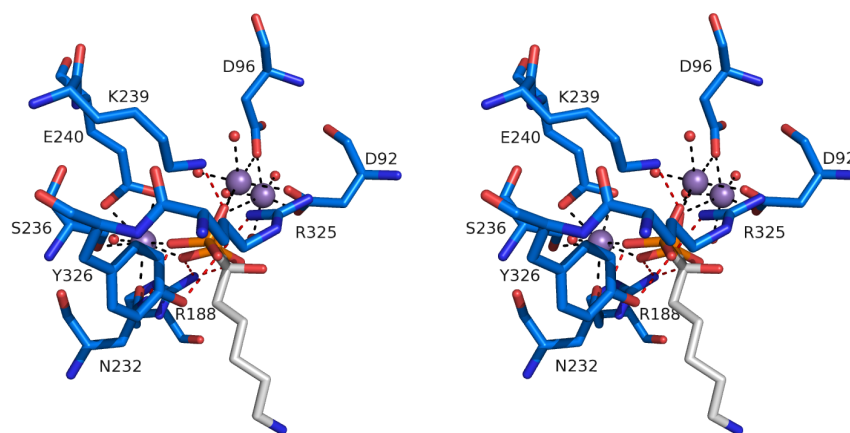


Figure S6. Stereoview of the PaFS₃₄₄-Mn²⁺₃-neridronate complex. Metal coordination and hydrogen bond interactions are shown as black and red dashed lines, respectively (metal-ligand distances are recorded in Table S3).

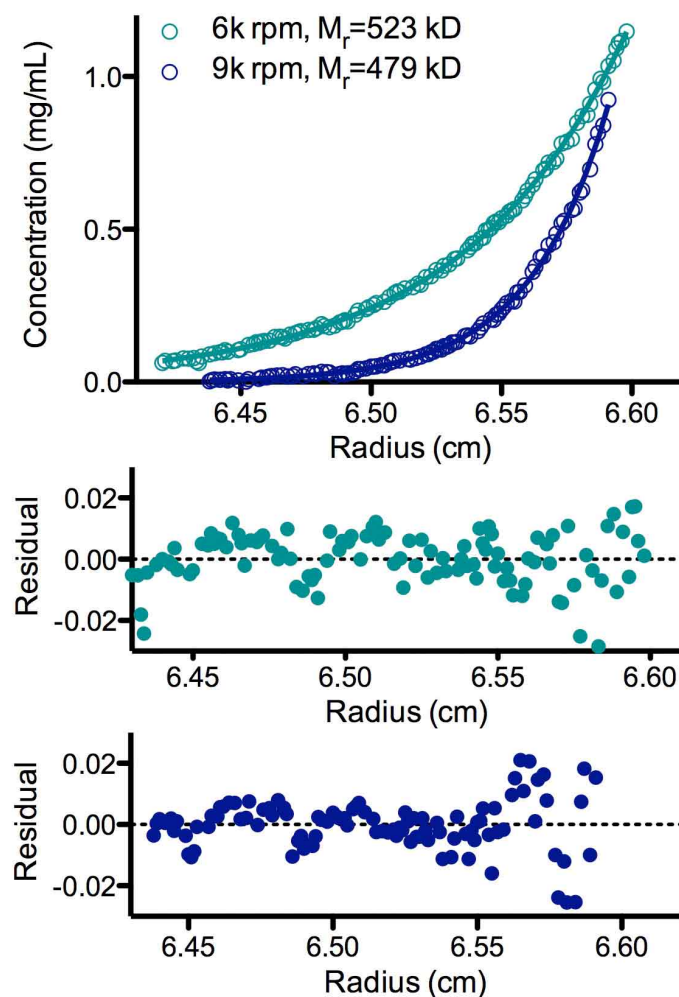


Figure S7. Sedimentation equilibrium runs with full-length PaFS lasted 24 h at two different speeds indicated. Nonlinear fits of experimental data yield average molecular weights (M_r) of 523 kD at 6,000 rpm (2898 $\times g$) and 479 kD at 9,000 rpm (6520 $\times g$), each within experimental error (5%) of the calculated molecular weight of a hexamer ($M_r = 502$ kD). Residuals are shown below fitted data and indicate a good fit of equilibrium concentration gradients.

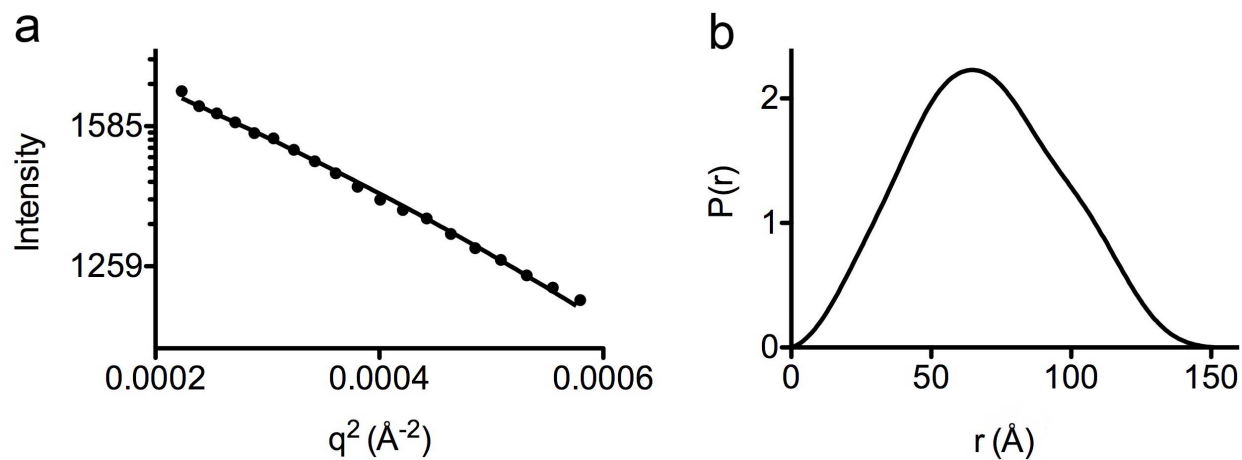


Figure S8. (a) Gunier plot of full-length N333A/Q336A PaFS generated from SAXS data indicates homogeneity of the protein sample. (b) Pair-distance distribution function $P(r)$ generated from SAXS data collected from full-length N333A/Q336A PaFS.