Supporting Information

The Cytochrome P450-Catalyzed Oxidative Rearrangement in the Final Step of Pentalenolactone Biosynthesis: Substrate Structure Determines Mechanism

Lian Duan,[†] Gerwald Jogl,[#,] and David E. Cane[†,#,*]*

†Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912-9108, USA

#Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island 02912, USA

*To whom correspondence should be addressed:

David E. Cane, Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912-9108, USA;

Gerwald Jogl, Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island 02912, USA

Table of Contents

Table S1. Data collection and refinement statistics for substrate-free PntM and PntM complexed with **2**, **7**, or **1.**

	PntM	**PntM + 2**	**PntM + 7**	**PntM + 1**
PDB entry	5L1R	5L1O	5L1Q	5L1P
Data collection[1]				
Space group	P 2_1 2_1 2	P 2_1 2_1 2	P 2_1 2_1 2	P 2_1 2_1 2
Cell dimensions				
a, *b*, *c* (Å)	44.46, 164.59, 83.30	44.75, 164.48, 83.29	44.53, 164.37, 82.04	44.33, 163.76, 81.34
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[2]	45.65 - 2.00 (2.05 - 2.00)	45.79 - 2.03 (2.08 - 2.03)	45.56 - 2.03 (2.08 - 2.03)	45.33 - 2.28 (2.36 - 2.28)
CC 1/2	0.999 (0.935)	0.996 (0.893)	0.992 (0.624)	0.983 (0.647)
I / γI	20.2 (6.0)	15.5 (4.7)	8.7 (2.2)	6.4 (2.0)
Completeness (%)	99.1 (90.9)	98.8 (91.5)	99.8 (97.2)	100.0 (100.0)
Redundancy	6.9 (4.6)	7.1 (5.4)	6.9 (5.0)	7.2 (7.2)
Refinement				
Resolution (Å)	45.65 – 2.00 (2.071 -2.00)	43.18 – 2.03 (2.103 -2.03)	45.56 – 2.03 (2.103 -2.03)	42.79 – 2.28 (2.361 -2.28)
No. reflections	41436 (3736)	40100 (3663)	39781 (3818)	27902 (2748)
R_{work} / R_{free}	0.1441 (0.1716)/ 0.1731 (0.2201)	0.1430 (0.1711)/ 0.1776 (0.2175)	0.1598 (0.2261) /0.1903 (0.2710)	0.1776 (0.2372) /0.2221 (0.3191)
No. atoms	3790	3848	3663	3476
macromolecules	3117	3146	3119	3130
Ligands	54	63	63	63
ions/water	619	639	481	283
B-factors	20.32	20.03	22.45	27.13
macromolecules	17.56	17.55	20.69	26.55
Ligands	16.40	10.64	13.79	20.13
Ions/water	34.55	33.16	35.01	35.10
R.m.s deviation				
Bond lengths (Å)	0.007	0.008	0.008	0.008
Bond angles (°)	1.24	1.20	1.20	1.28
Ramachandran plot				
Favored (%)	98	99	98	97
Outliers (%)	0	0	0	0

1. One crystal was used for this dataset. 2. Highest resolution shell is shown in parenthesis.

Table S2. Data collection and refinement statistics PntM mutants complexed with **2.**

	PntM-F232L	**PntM-M77S**	**PntM-M81A**	**PntM-M81C**	**PntM-M81C-BME**
PDB entry	5L1S	5L1T	5L1U	5L1V	5L1W
Data collection[1]					
Space group	P 2_1 2_1 2	P 2_1 2_1 2	P 2_1 2_1 2	P 2_1 2_1 2	P 2_1 2_1 2
Cell dimensions					
a, *b*, *c* (Å)	44.60, 164.15, 82.54	44.68, 164.60, 83.51	44.62, 163.86, 82.22	44.56, 164.23, 82.97	44.57, 163.23, 83.29
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[2]	45.60 - 2.08 (2.13 - 2.08)	45.85 – 2.08 (2.14 – 2.08)	45.50 - 2.07 (2.13 - 2.07)	45.69 - 2.12 (2.18 - 2.12)	44.57 - 2.06 (2.12 - 2.06)
CC 1/2	0.993 (0.724)	0.982 (0.618)	0.995 (0.809)	0.987 (0.702)	0.995 (0.821)
I / γI	11.0 (2.9)	9.7 (2.1)	12.2 (3.5)	7.5 (2.2)	13.6 (4.4)
Completeness (%)	98.6 (92.1)	98.2 (90.4)	98.7 (91.3)	100.0 (100.0)	97.2 (92.0)
Redundancy	7.1 (4.8)	7.0 (4.8)	7.1 (4.7)	7.2 (7.1)	7.3 (6.5)
Refinement					
Resolution (Å)	41.27 – 2.076 (2.15 -2.076)	41.75 – 2.082 (2.157 -2.081)	43.06 – 2.074 (2.148 -2.074)	43.01 – 2.12 (2.196 -2.12)	41.64 – 2.06 (2.134 -2.06)
No. reflections	36872 (3507)	37605 (3436)	36997 (3426)	35477 (3450)	37401 (3510)
R_{work} / R_{free}	0.1540 (0.2072) / 0.1924 (0.2644)	0.1599 (0.2417) / 0.1951 (0.2796)	0.1541 (0.1972) / 0.1845 (0.2416)	0.1604 (0.2271) / 0.1956 (0.2731)	0.1613 (0.1953) / 0.1855 (0.2203)
No. atoms	3689	3711	3700	3733	3822
macromolecules	3143	3136	3146	3119	3128
ligands	63	63	63	63	63
ions/water	483	502	491	551	631
B-factors	21.14	20.77	19.42	22.93	18.89
macromolecules	19.79	19.36	17.83	20.90	16.49
ligands	12.05	12.82	10.90	15.11	10.62
ions/water	31.13	30.38	30.70	35.30	31.61
R.m.s deviations					
Bond lengths (Å)	0.009	0.009	0.008	0.008	0.006
Bond angles (°)	1.24	1.24	1.24	1.23	1.15
Ramachandran plot					
Favored (%)	98	98	98	98	98
Outliers (%)	0	0	0	0	0

1. One crystal was used for this dataset. 2. Highest resolution shell is shown in parenthesis.

Table S3. Primer pairs for site-directed mutagenesis

	Forward Primer (5′ - 3′)	Reverse Primer (5′ – 3′)
F232G	GCC GTG CTG TTG GGC GGC TAT GAA ACG ACC	GGT CGT TTC ATA GCC GCC CAA CAG CAC GGC
F232A	GCC GTG CTG TTG GCC GGC TAT GAA ACG ACC	GGT CGT TTC ATA GCC GGC CAA CAG CAC GGC
F232Y	GCC GTG CTG TTG TAC GGC TAT GAA ACG ACC	GGT CGT TTC ATA GCC GTA CAA CAG CAC GGC
F232L	GCC GTG CTG TTG CTG GGC TAT GAA ACG ACC	GGT CGT TTC ATA GCC CAG CAA CAG CAC GGC
F232H	GCC GTG CTG TTG GAT GGC TAT GAA ACG ACC	GGT CGT TTC ATA GCC ATC CAA CAG CAC GGC
M81A	GTA GTG GCT CTG GCC GCC GGT GAT GAC	GTC ATC ACC GGC GGC CAG AGC CAC CAT
M81C	GTA GTG GCT CTG TGC GCC GGT GAT GAC	GTC ATC ACC GGC GCA CAG AGC CAC CAT
M77S	GCT GCG CGT GCT TTC AGC GTG GCT CTG ATG GCC	GGC CAT CAG AGC CAC TCG GAA AGC ACG CGC AGC

Table S4. Predicted MW and HR-LC-ESI-MS observed M_D of PntM and mutants

Protein	**Observed M_D (MH^+)**	**Calculated MW**
PntM	44445.27	44444.87
F232A	44369.80	44368.77
F232G	44355.86	44354.74
F232L	44412.67	44410.85
F232H	44436.51	44434.83
M81C	44418.39	44416.18
M81C-BME	44494.37	44492.31
M77S	44401.20	44400.75

The calculated MW is based on protein with an N-terminal Gly-Ser-His remaining after thrombin cleavage of the N-terminal His_6-tag.

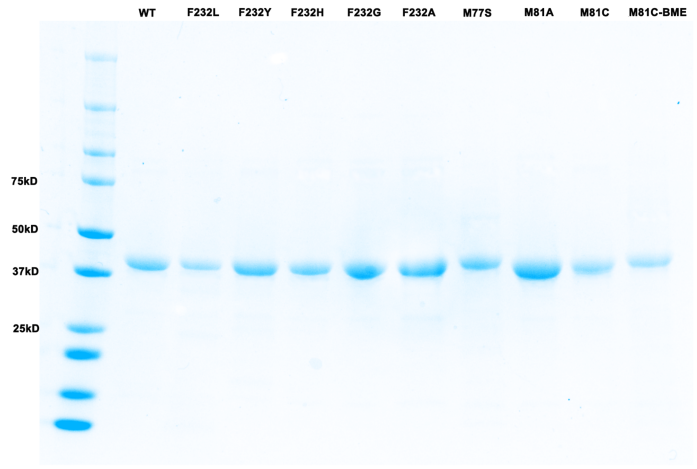


Figure S1. SDS-PAGE analysis of PntM and mutants after removal of N-terminal His_6-tag.

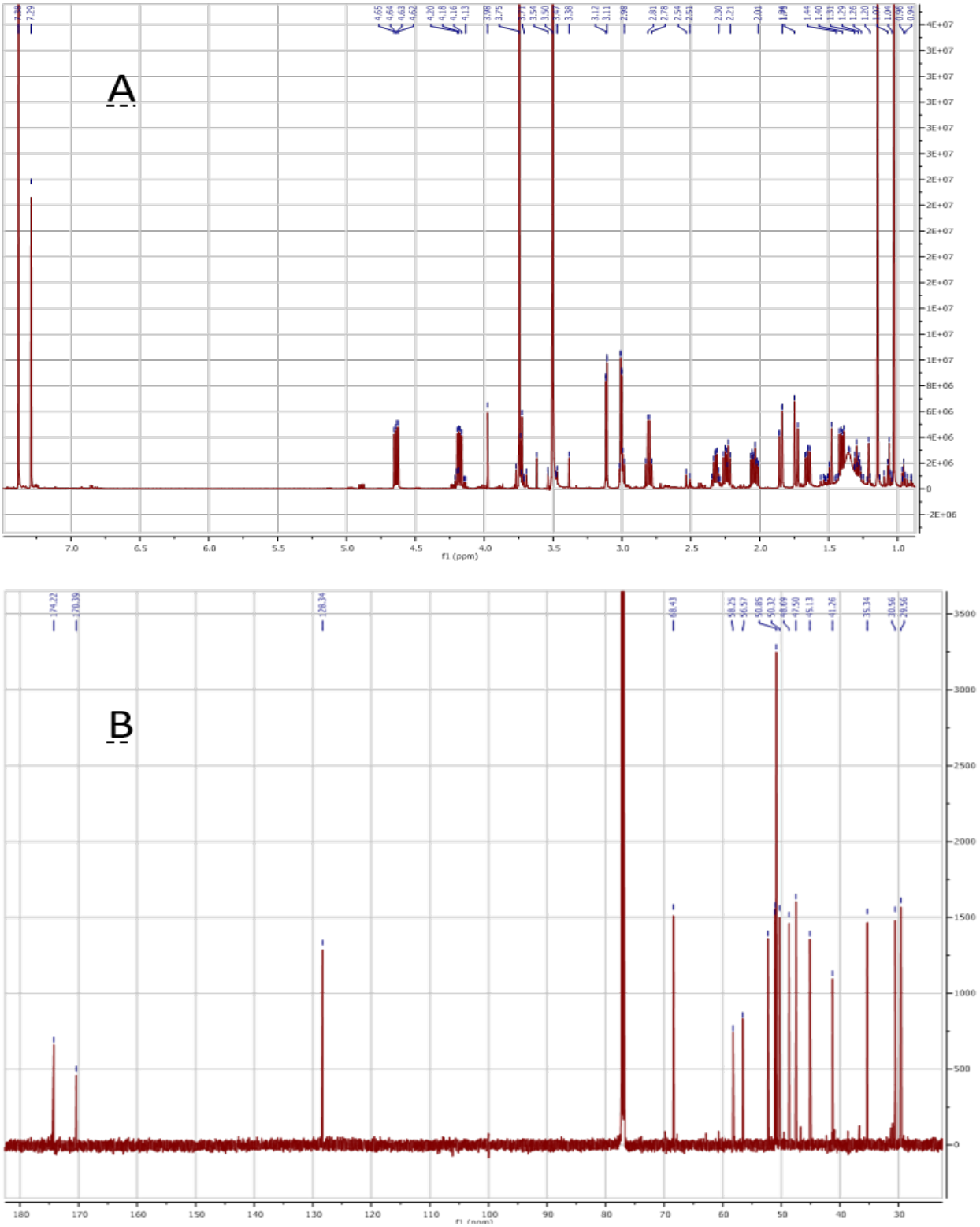


Figure S2. 6,7-Dihydropentalenolactone F methyl ester (**7-Me**). a) ^{1}H NMR (400 MHz, $CDCl_3$); b) ^{13}C NMR (100 MHz, $CDCl_3$).

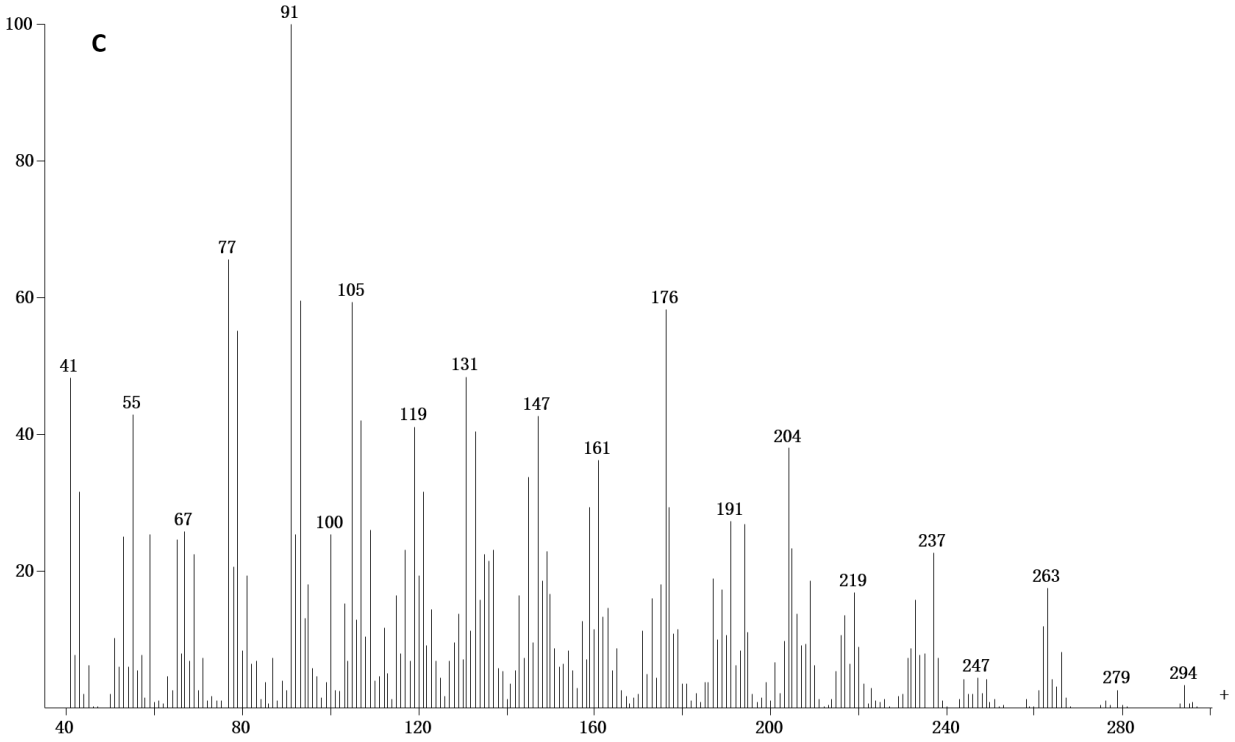


Figure S3. 6,7-Dihydropentalenolactone F methyl ester (**7-Me**), EI-MS

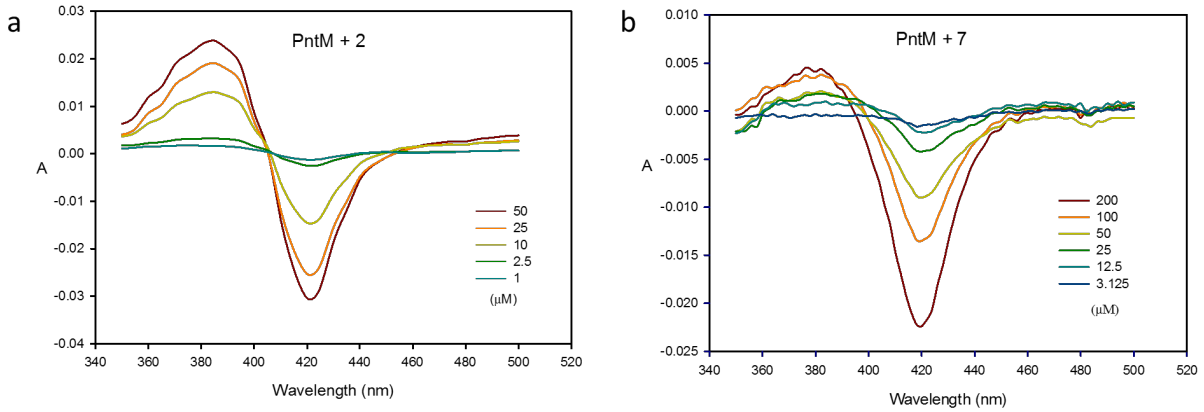


Figure S4. Titration of PntM, UV difference spectra. a) Pentalenolactone F (**2**); b) 6,7-Dihydropentalenolactone F (**7**).

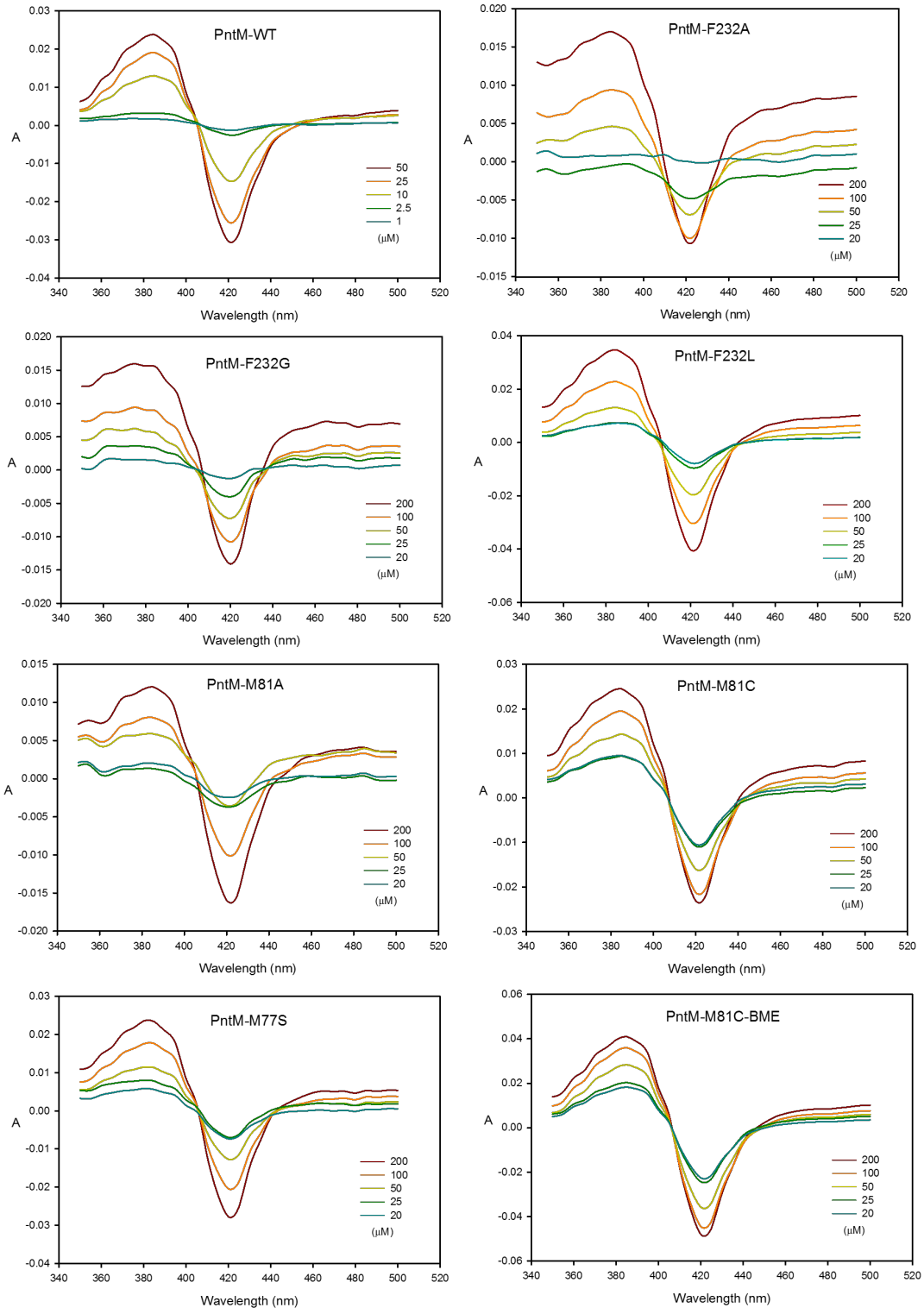


Figure S5. Titration of PntM mutants with pentalenolactone F (**2**), UV difference spectra.

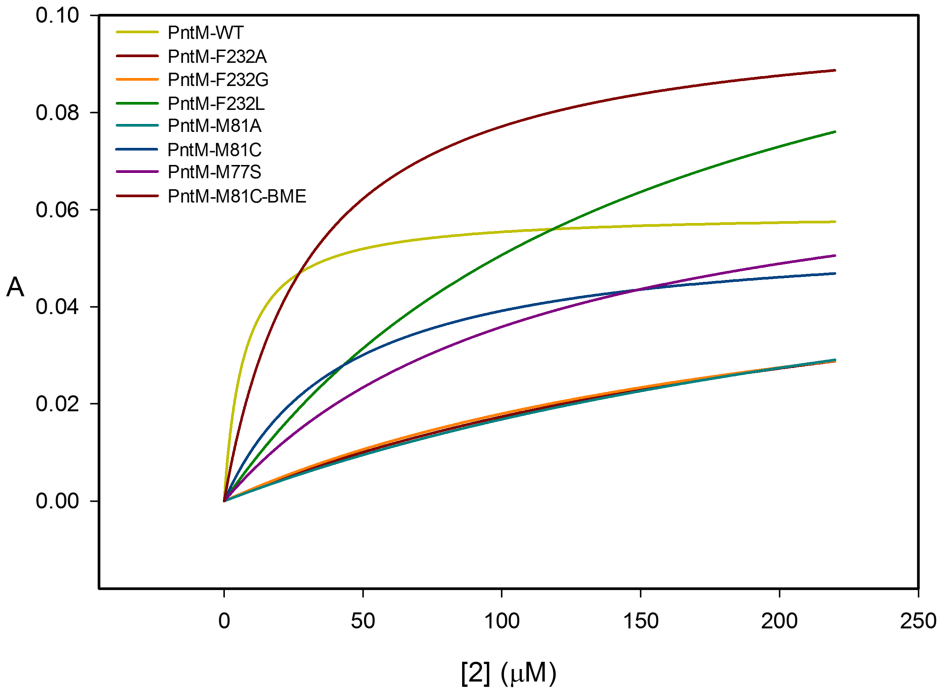


Figure S6. Plots of UV difference (A_{390}) vs concentration of **2** for wild-type PntM and mutants.

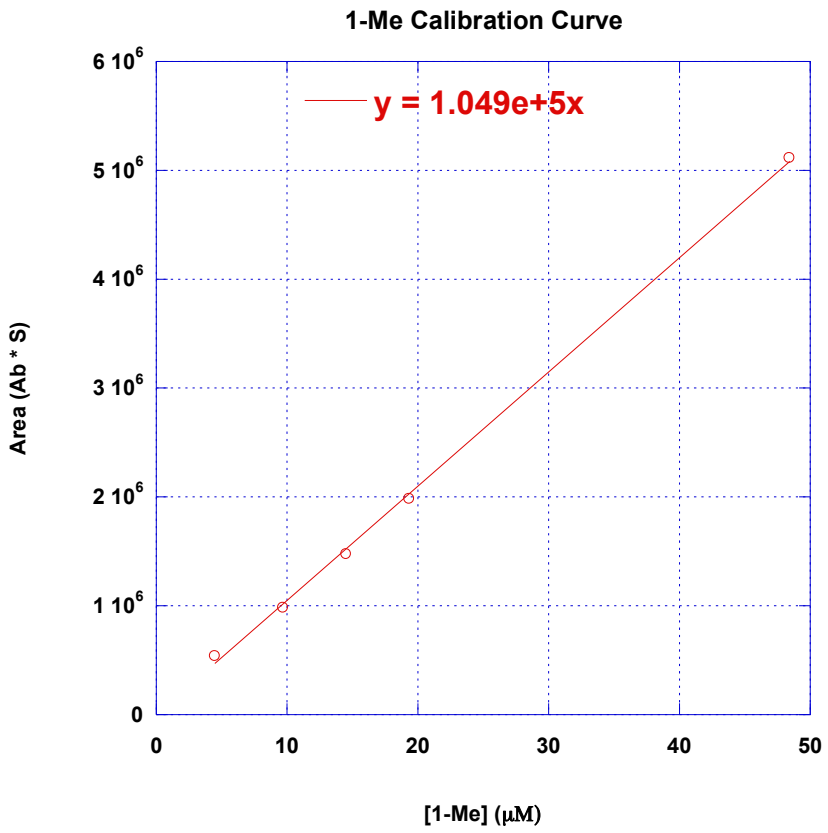


Figure S7. GC-MS calibration plot for pentalenolactone methyl ester (**1-Me**).

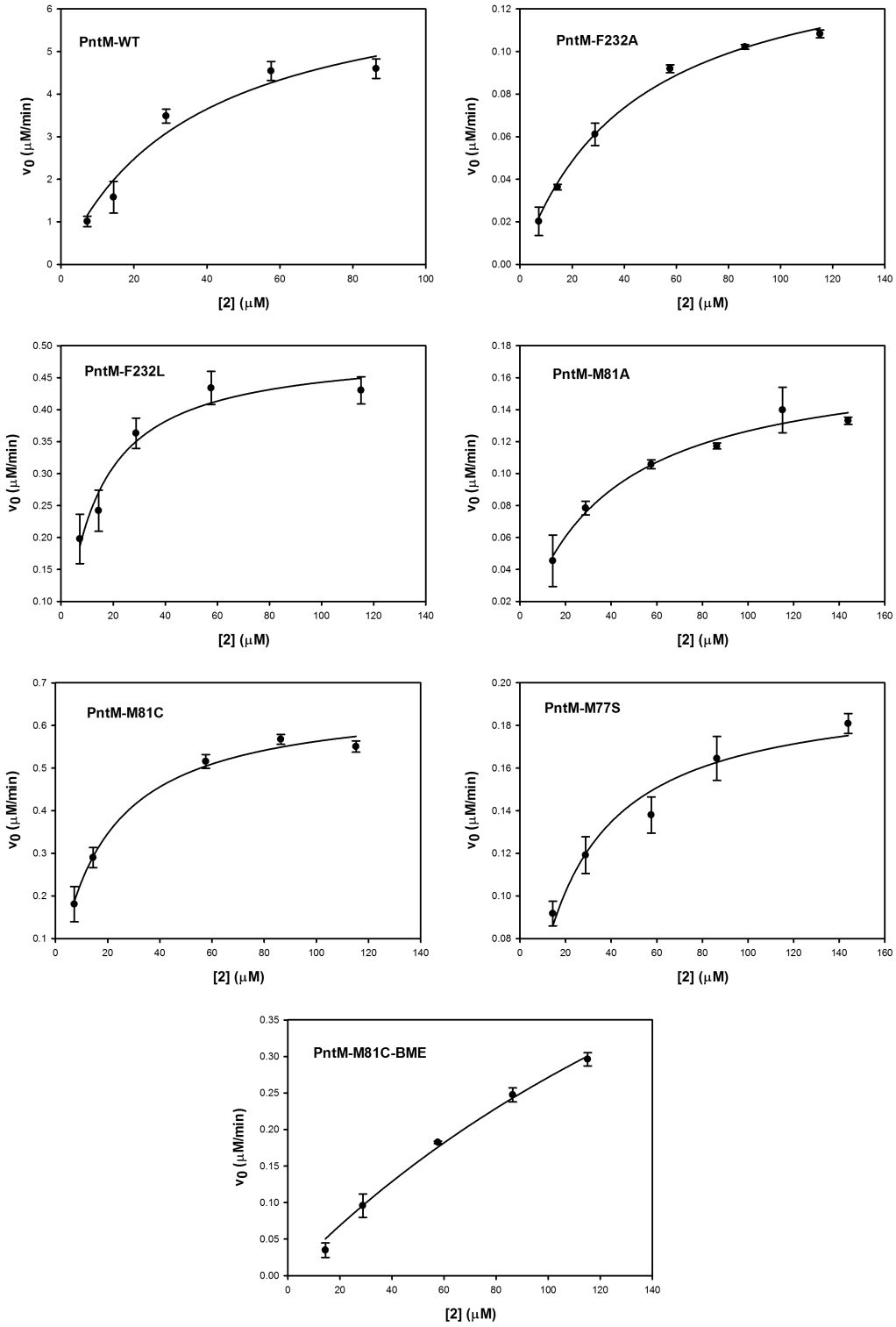


Figure S8. Steady-state kinetic analysis of conversion of **2** to **1** catalyzed by PntM and mutants.

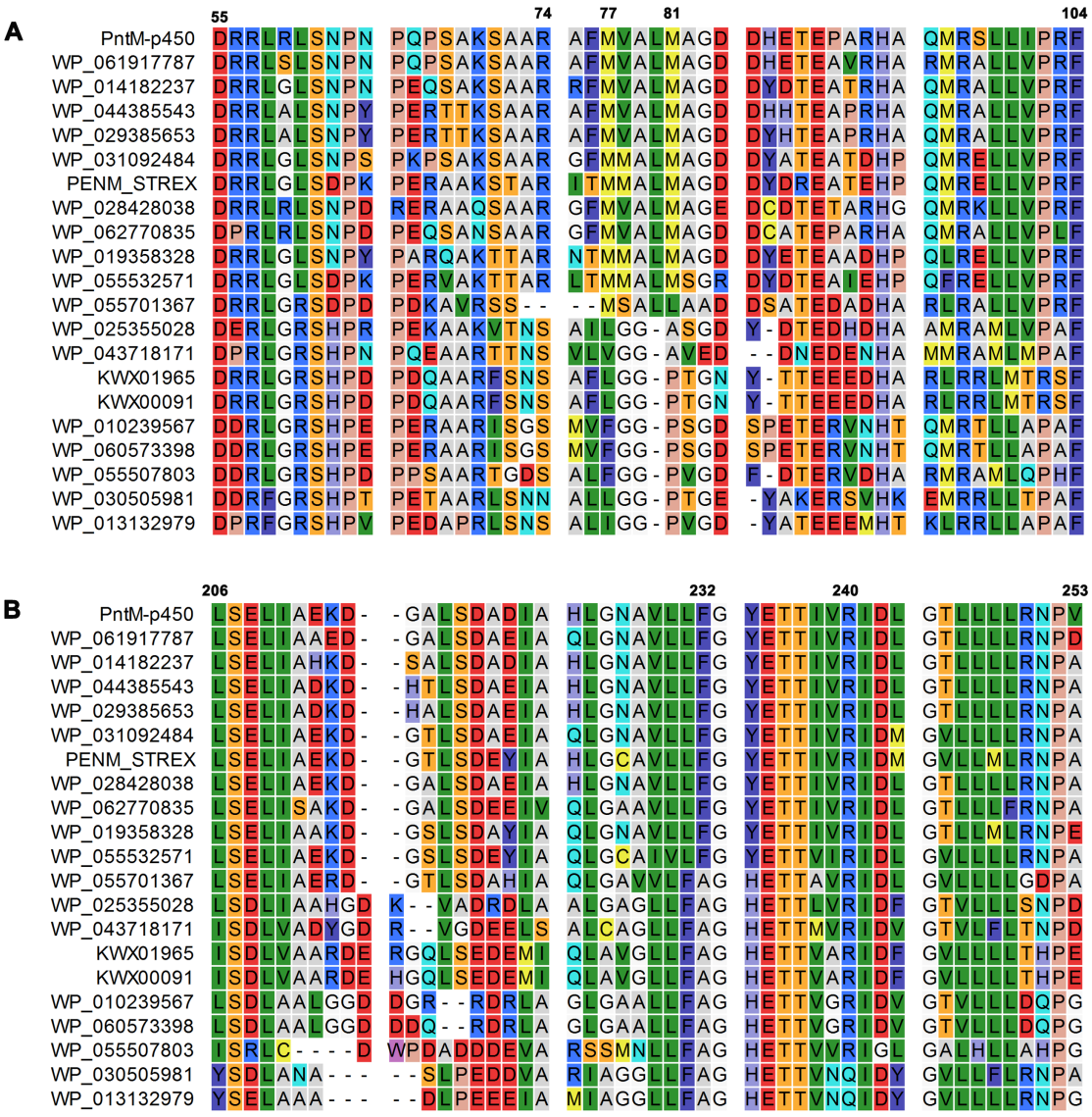


Figure S9. BLAST comparison of selected regions of PntM with top 20 matches. A) D55-F104, including key active site residues R74, M77, and M81. B) L206-V253, including key active site residues F232, T236, and R240

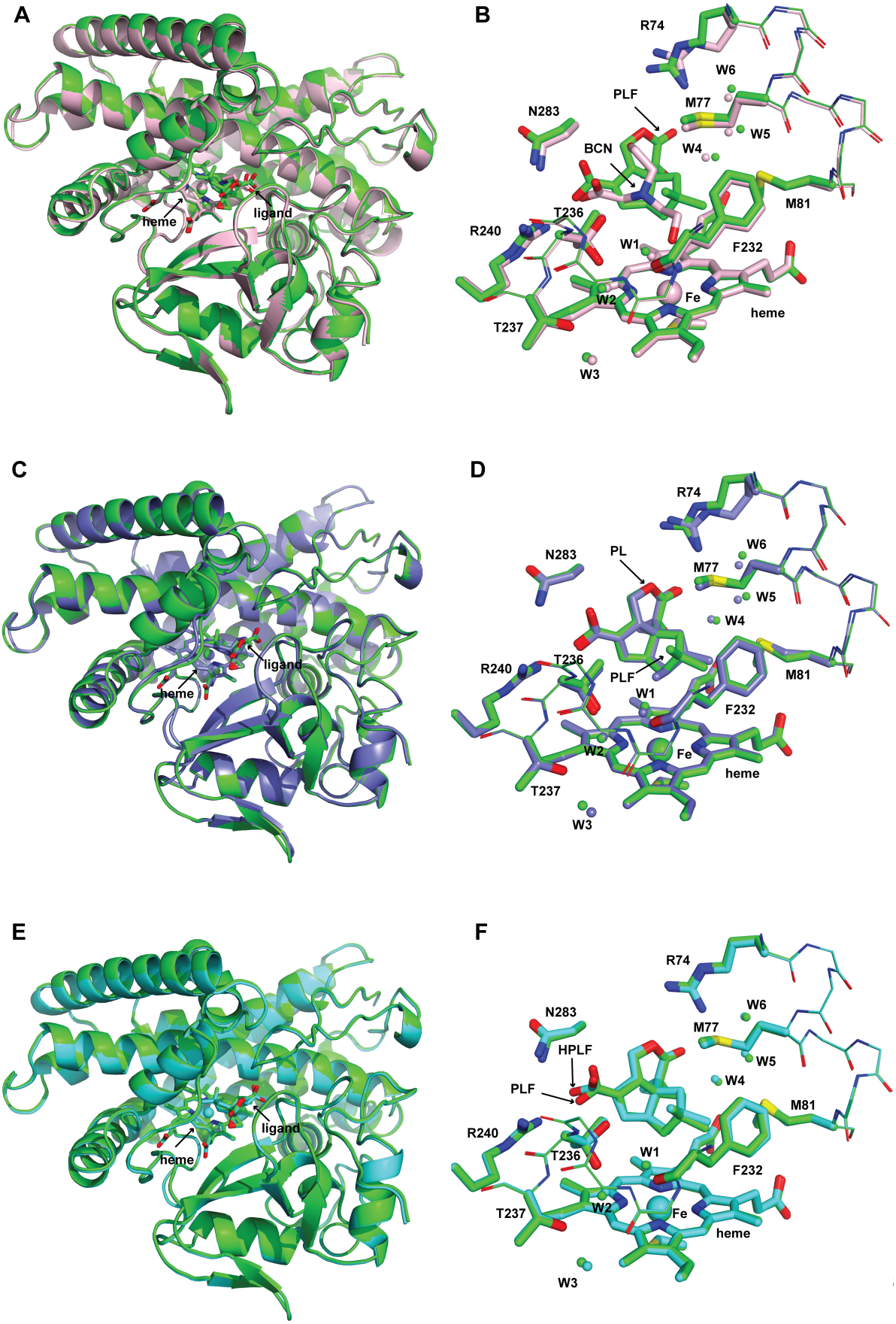
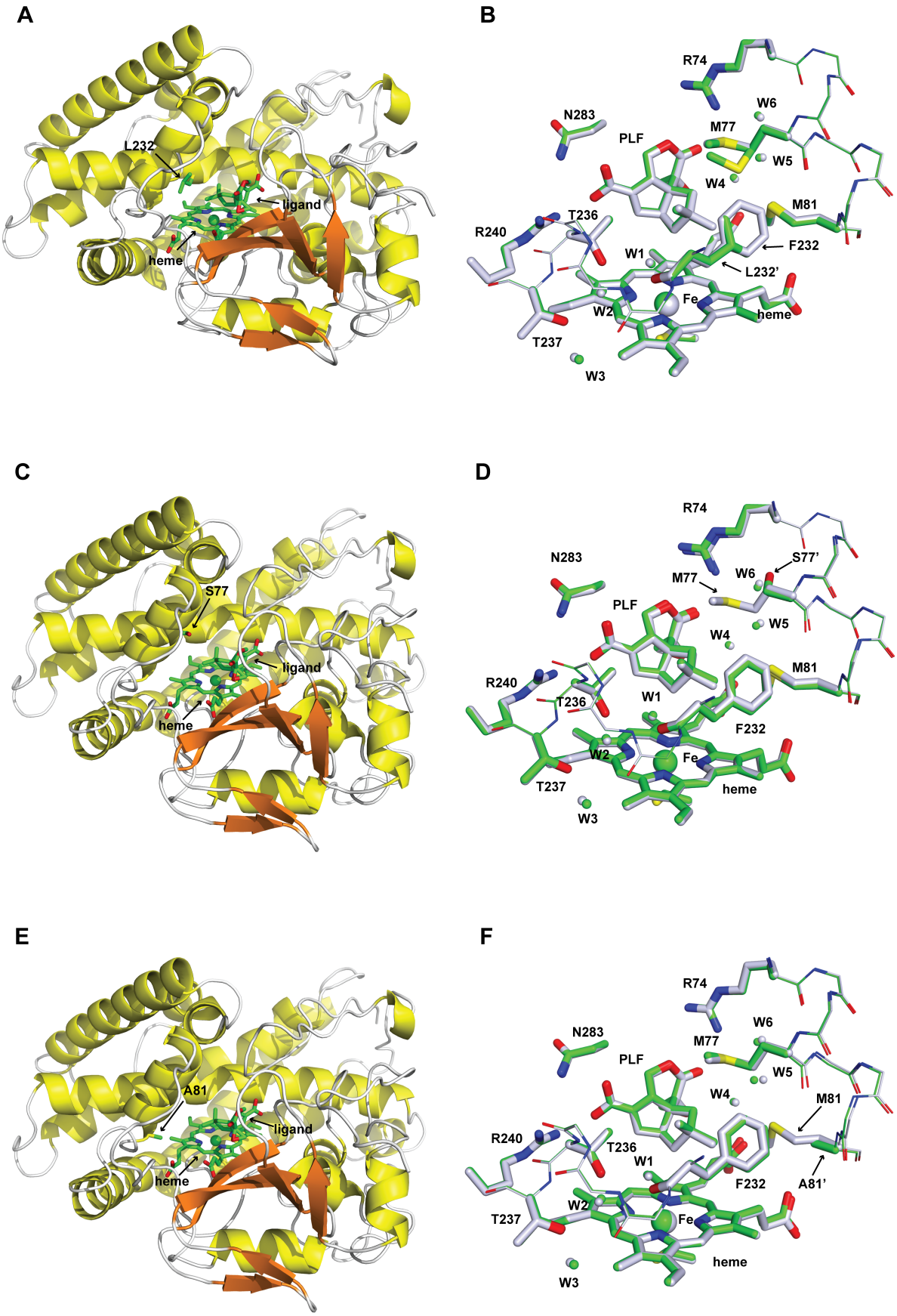

A
heme
ligand
B
R74
W6
PLF
M77
N283
W5
BCN
W4
M81
T236
R240
W1
F232
W2
Fe
heme
T237
W3
C
ligand
heme
D
R74
N283
PL
W6
M77
W5
W4
T236
R240
PLF
M81
W1
F232
W2
Fe
heme
T237
W3
E
ligand
heme
F
R74
N283
W6
M77
HPLF
W5
W4
PLF
M81
R240
T236
W1
F232
T237
W2
Fe
heme
W3

Figure S10. Comparison of the structures of PntM with bound ligands. **A**, Alignment and comparison of the overview structures of PntM with bound bicine and PntM with bound substrate **2**. PntM with bound bicine is shown in light pink and PntM with bound **2** is shown in green, heme is shown as, heme Fe^{+3} is shown as a sphere and ligands are shown as sticks. **B**, comparison of the active site regions of PntM with bound bicine and PntM with bound 2. The relevant residues and water molecules of PntM with bound bicine are shown in light pink, while the relevant residues and water molecules of PntM with bound 2 are shown in green. **C**, comparison of the overview structures of PntM with bound product **1** (shown in slate) and PntM with bound substrate **2**. **D**, comparison of the active site region of PntM with bound **1** and PntM with bound **2**. **E**, comparison of the overview structures of PntM with bound substrate **7** (shown in cyan) and PntM with bound substrate **2**. **F**, comparison of the active site region of PntM with bound **7** and PntM with bound **2**.


A
L232
ligand
heme
B
R74
W6
N283
M77
PLF
W5
W4
M81
T236
R240
F232
W1
L232'
Fe
W2
heme
T237
W3
C
S77
ligand
heme
D
R74
N283
S77'
W6
M77
PLF
W5
W4
M81
R240
T236
W1
F232
W2
Fe
T237
heme
W3
E
A81
ligand
heme
F
R74
W6
N283
M77
PLF
W5
W4
M81
R240
T236
W1
F232
A81'
W2
T237
Fe
heme
W3

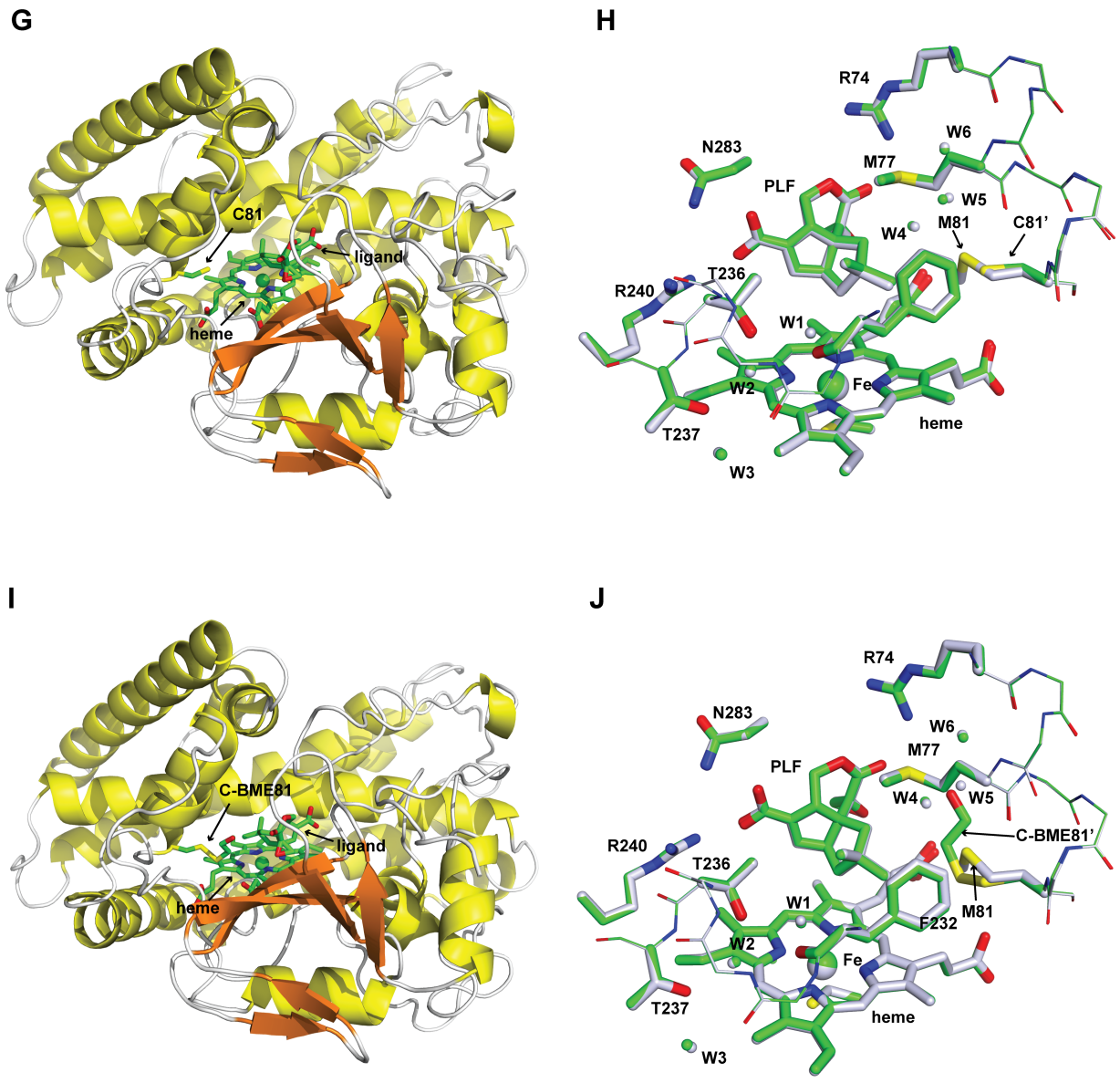


Figure S11. Structures of PntM mutants. **A**, overview structure of F232L with bound substrate **2**. Heme is shown as sticks with carbon atoms in green, heme Fe^{3+} is shown as a green sphere. The mutated residue L232 is shown as sticks with carbon atoms in green. **B**, comparison of the active site region of F232L with bound **2** and PntM with bound **2**. The relevant residues and water molecules of F232L with bound **2** are shown in green, while those for PntM with bound **2** are shown in white blue. The mutated residue in F232L is labeled as L232'. **C**, overview structure of M77S with bound substrate **2**. The mutated residue S77 is shown as sticks with carbon atoms in green. **D**, comparison of the active site region of M77S with bound **2** and PntM with bound **2**. The mutated residue in M77S is labeled as S77'. **E**, overview showing the structure of M81A with bound

substrate **2**. The mutated residue A81 is shown as sticks with carbon atoms in green. **F**, comparison of the active site region of M81A with bound **2** and PntM with bound **2**. Mutated residue in M81A is labeled as A81'. **G**, overview showing the structure of M81C with bound substrate **2**. The mutated residue C81 is shown as sticks with carbon atoms in green. **H**, comparison of the active site region of M81A with bound **2** and PntM with bound **2**. The mutated residue in M81C is labeled as C81'. **I**, overview showing the structure of M81C-BME with bound substrate **2**. The mutated residue C-BME81 is shown as sticks with carbon atoms in green. **J**, comparison of the active site region of M81C-BME with bound **2** and PntM with bound **2**. The mutated residue in M81C-BME is labeled as C-BME81'.