Supporting Information for

Three clusters of conformational states in P450cam reveals a multi-step pathway for opening of the substrate access channel.

Young-Tae Lee, Edith C. Glazer, Richard F. Wilson, C. David Stout and David B. Goodin

Department of Molecular Biology, 10550 N. Torrey Pines Road, The Scripps Research Institute, La Jolla, CA 92037

Table S1. Data collection, refinement statistics and validation

Tethered substrate analog	AdaC1-C8-Dans	AdaC1-C8-Dans
PDB ID	3P6M	3P6N
Data collection		
wavelength (Å)	0.97945	0.97946
unit cell (Å)	65.290, 74.600, 92.360	65.240, 74.760, 92.860
space group	P212121	P212121
resolution range (Å)	53.30 – 2.00 (2.11 – 2.00)[a]	33.75 – 1.70 (1.79 – 1.70)[a]
No. of total reflections	107280	203402
No. of unique reflections	30637	50013
Completeness (%)	98.7 (98.7)	98.8 (99.3)
R_{merge} (%)	7.3 (53.3)	6.1 (34.3)
$<I/\sigma(I)>$	10.2 (2.5)	13.5 (5.1)
Wilson B-value ($Å^2$)	29.1	18.4
Refinement statistics		
resolution range (Å)	10 – 2.0	10 – 1.7
No. of reflections used	28753	49613
Free R reflections (%)	5.0	5.0
R/R_{free}	0.205/0.251	0.204/0.227
rmsd bond length (Å)	0.0132	0.0099
rmsd bond angle (deg)	1.523	1.274
Ramachandran analysis (%)		
Residues in		
most favored regions	90.8	91.6
additional allowed regions	9.2	8.4
generously allowed regions	0.0	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S1. Data collection, refinement statistics and validation (continued)

Tethered substrate analog	AdaC1-Etg-Dans	AdaC1-C6-Bio
PDB ID	3P6O	3P6P
Data collection		
wavelength (Å)	0.97946	0.97945
unit cell (Å)	64.980, 75.107, 92.846	64.900, 74.963, 92.248
space group	P212121	P212121
resolution range (Å)	46.42 – 2.00 (2.11 – 2.00)[a]	26.55 – 1.90 (2.00 – 1.90) [a]
No. of total reflections	111557	127963
No. of unique reflections	29725	35463
Completeness (%)	95.3 (90.4)	98.3 (97.5)
R_{merge} (%)	6.1 (47.4)	6.4 (39.7)
$\langle I/\sigma(I) \rangle$	18.4 (2.5)	15.8 (3.0)
Wilson B-value ($Å^2$)	29.2	18.4
Refinement statistics		
resolution range (Å)	10 – 2.0	10 – 1.9
No. of reflections used	27951	33464
Free R reflections (%)	5.0	5.0
R/R_{free}	0.209/0.267	0.201/0.241
rmsd bond length (Å)	0.0100	0.0113
rmsd bond angle (deg)	1.240	1.397
Ramachandran analysis (%)		
Residues in		
most favored regions	90.5	89.4
additional allowed regions	9.5	10.6
generously allowed regions	0.0	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S1. Data collection, refinement statistics and validation (continued)

Tethered substrate analog	AdaC2-Etg-Boc	3OH-AdaC1-Etg-Boc
PDB ID	3P6Q	3P6R
Data collection		
wavelength (Å)	0.97946	0.97839
unit cell (Å)	65.777, 73.894, 90.276	66.809, 75.008, 91.781
space group	P212121	P212121
resolution range (Å)	34.20 – 1.95 (2.06 – 1.95)[a]	29.04 – 2.10 (2.21 – 2.10) [a]
No. of total reflections	120852	152598
No. of unique reflections	32492	27551
Completeness (%)	99.3 (99.6)	99.9 (99.9)
R_{merge} (%)	6.3 (41.0)	11.1 (83.0)
$\langle I/\sigma(I) \rangle$	19.4 (2.8)	13.1 (1.9)
Wilson B-value ($Å^2$)	25.0	32.2
Refinement statistics		
resolution range (Å)	10 – 1.95	10 – 2.10
No. of reflections used	30537	25891
Free R reflections (%)	5.0	5.0
R/R_{free}	0.196/0.238	0.217/0.246
rmsd bond length (Å)	0.0108	0.0121
rmsd bond angle (deg)	1.270	1.300
Ramachandran analysis (%)		
Residues in		
most favored regions	91.3	90.0
additional allowed regions	8.7	10.0
generously allowed regions	0.0	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S1. Data collection, refinement statistics and validation (continued)

Tethered substrate analog	AdaC2-C8-Dans	AdaC2-C8-Dans
PDB ID	3P6S	3P6T
Data collection		
wavelength (Å)	0.97946	0.97945
unit cell (Å)	65.594, 73.863, 91.153	65.423, 74.280, 92.170
space group	P212121	P212121
resolution range (Å)	38.78 – 2.00 (2.11 – 2.00)[a]	26.67 – 1.90 (2.00 – 1.90) [a]
No. of total reflections	99041	149212
No. of unique reflections	29387	35978
Completeness (%)	96.4 (87.5)	99.7 (99.9)
R_{merge} (%)	6.5 (50.0)	6.9 (49.0)
$\langle I/\sigma(I) \rangle$	9.2 (2.1)	11.5 (3.2)
Wilson B-value ($Å^2$)	29.5	23.4
Refinement statistics		
resolution range (Å)	10 – 2.00	10 – 1.90
No. of reflections used	27615	33931
Free R reflections (%)	5.0	5.0
R/R_{free}	0.199/0.250	0.206/0.251
rmsd bond length (Å)	0.0156	0.0145
rmsd bond angle (deg)	1.573	1.485
Ramachandran analysis (%)		
Residues in		
most favored regions	90.2	90.5
additional allowed regions	9.8	9.5
generously allowed regions	0.0	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S1. Data collection, refinement statistics and validation (continued)

Tethered substrate analog	AdaC3-C6-Dans	3Et-AdaC1-Etg-Boc
PDB ID	3P6U	3P6V
Data collection		
wavelength (Å)	0.97946	0.97839
unit cell (Å)	65.759, 73.820, 92.185	66.949, 74.365, 92.074
space group	P212121	P212121
resolution range (Å)	26.05 – 1.70 (1.79 – 1.70)[a]	28.98 – 2.00 (2.11 – 2.00)[a]
No. of total reflections	190848	141679
No. of unique reflections	49948	31486
Completeness (%)	99.8 (100)	99.4 (99.2)
R_{merge} (%)	4.9 (29.2)	8.5 (59.6)
$\langle I/\sigma(I) \rangle$	15.9 (4.0)	13.8 (2.2)
Wilson B-value ($Å^2$)	19.9	27.8
Refinement statistics		
resolution range (Å)	10 – 1.70	10 – 2.00
No. of reflections used	47100	29637
Free R reflections (%)	5.0	5.0
R/R_{free}	0.190/0.214	0.211/0.242
rmsd bond length (Å)	0.0091	0.0106
rmsd bond angle (deg)	1.232	1.228
Ramachandran analysis (%)		
Residues in		
most favored regions	91.3	91.5
additional allowed regions	8.7	8.5
generously allowed regions	0.0	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S1. Data collection, refinement statistics and validation (continued)

Tethered substrate analog	AdaC3-Etg-Boc	AdaC3-C8-Dans
PDB ID	3P6W	3P6X
Data collection		
wavelength (Å)	0.97839	0.97839
unit cell (Å)	66.089, 74.805, 92.946	66.949, 74.365, 92.074
space group	P212121	P212121
resolution range (Å)	29.14 – 2.10 (2.21 – 2.10)[a]	28.98 – 2.00 (2.11 – 2.00)[a]
No. of total reflections	112941	141679
No. of unique reflections	25134	31486
Completeness (%)	91.8 (82.2)	99.4 (99.2)
R_{merge} (%)	8.5 (52.6)	8.5 (59.6)
$\langle I/\sigma(I) \rangle$	9.8 (2.9)	13.8 (2.2)
Wilson B-value ($Å^2$)	33.9	27.8
Refinement statistics		
resolution range (Å)	10 – 2.10	10 – 2.00
No. of reflections used	23634	29637
Free R reflections (%)	5.0	5.0
R/R_{free}	0.206/0.263	0.211/0.242
rmsd bond length (Å)	0.0121	0.0106
rmsd bond angle (deg)	1.348	1.228
Ramachandran analysis (%)		
Residues in		
most favored regions	91.5	91.5
additional allowed regions	8.2	8.5
generously allowed regions	0.3	0.0
disallowed regions	0.0	0.0

[a] Data for the outermost shell are given in the parentheses

Table S2. Space group and unit cell dimension

Substrate	PDB	Space group	a, Å	b, Å	c, Å
P450cam-C					
Camphor	2CPP	P212121	108	103	36
	1YRC	P43212	63	63	247
	5CP4	P212121	106	103	36
	2ZAX	P43212	63	63	249
	1DZ4	P21	67	63	96
1S-camphor	1AKD	P212121	64	66	107
Imidazole (permute abc->bca)	2H7Q	P212121	104	106	37
Nicotine	1P2Y	P43212	64	64	251
Metyrapone	1PHG	P212121	108	103	36
P450cam-I					
AdaC1-C8-Dans	1RE9	P212121	64	76	92
	1LWL	P212121	64	75	93
	3P6M	P212121	65	75	92
	3P6N	P212121	65	75	93
AdaC1-C8EtgGlu-Bio	3OIA	P212121	65	75	93
AdaC1-Etg-Dans	3P6O	P212121	65	75	93
AdaC1-C6-Bio	3P6P	P212121	65	75	92
3OH-AdaC1-C8-Dans	3OL5	P212121	66	75	93
P450cam-O					
AdaC1-C4-Dans	1RF9	P212121	65	75	95
AdaC2-Etg-Boc	3P6Q	P212121	66	74	90
3OH-AdaC1-Etg-Boc	3P6R	P212121	65	75	92
AdaC2-C8-Dans	3P6S	P212121	65	74	91
	3P6T	P212121	65	74	92
AdaC3-C6-Dans	3P6U	P212121	65	74	92
3Et-AdaC1-Etg-Boc	3P6V	P212121	67	74	92
AdaC3-Etg-Boc	3P6W	P212121	66	75	93
AdaC3-C8-Dans	3P6X	P212121	66	74	93
AdaC1-perfluorobiphenyl-Ru(bipy)$_3$	1K2O	P1	64	67	73
AdaC1-C8-Ru(bipy)$_3$	1QMQ	P212121	65	74	92
Substrate-free	3L61	P212121	66	74	92
	3L62	P212121	65	74	92

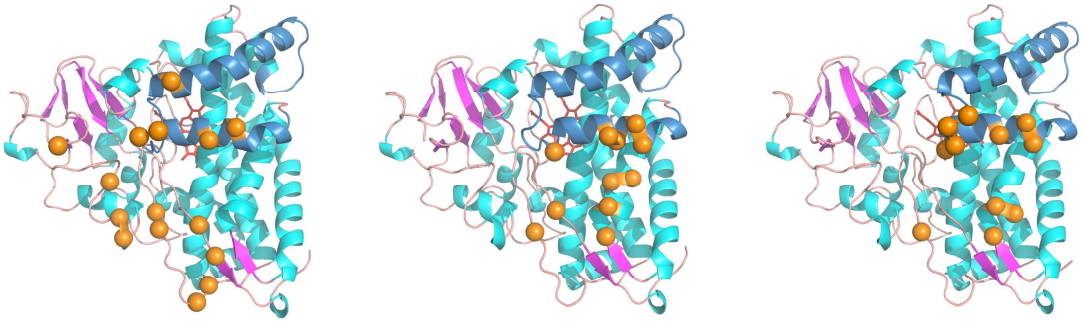

Figure S1. Crystal packing interface around the substrate channel. The residues involved in crystal packing with the neighboring molecules are shown in orange. The F and G helices are shown in blue. Other helices, sheets and loops are colored in cyan, magenta and pink, respectively. Heme is shown in red.

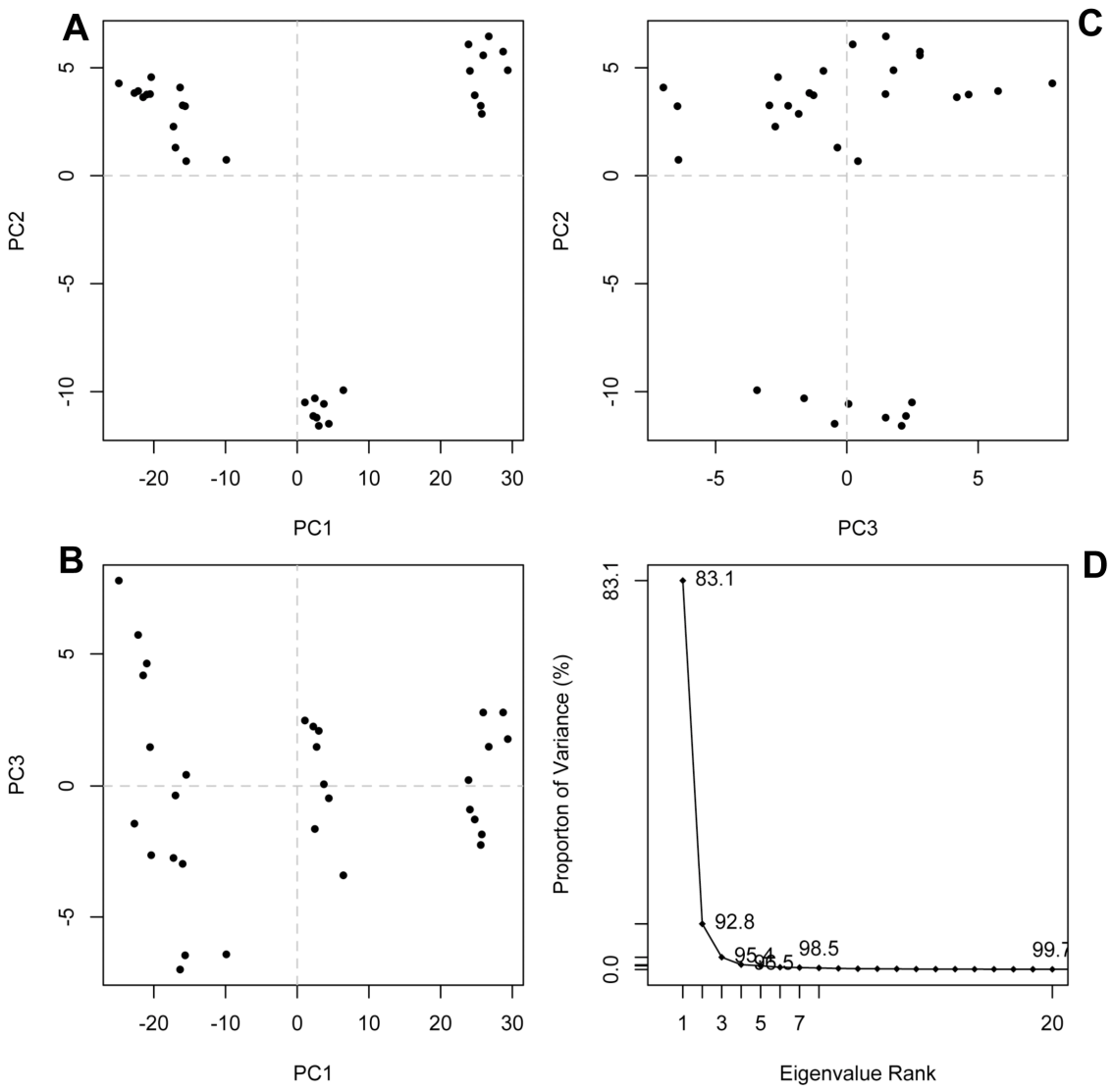

Figure S2. Results of PCA analysis. Analyzed P450cam structures are projected onto two-dimensional plots for the three most significant principal components of P450cam structures; (A) PC1 and PC2, (B) PC1 and PC3, (C) PC2 and PC3. The magnitude of covariance is expressed as the percentage of the total covariance (D). All of the plots in the figure were prepared using the program bio3d package (*41*).

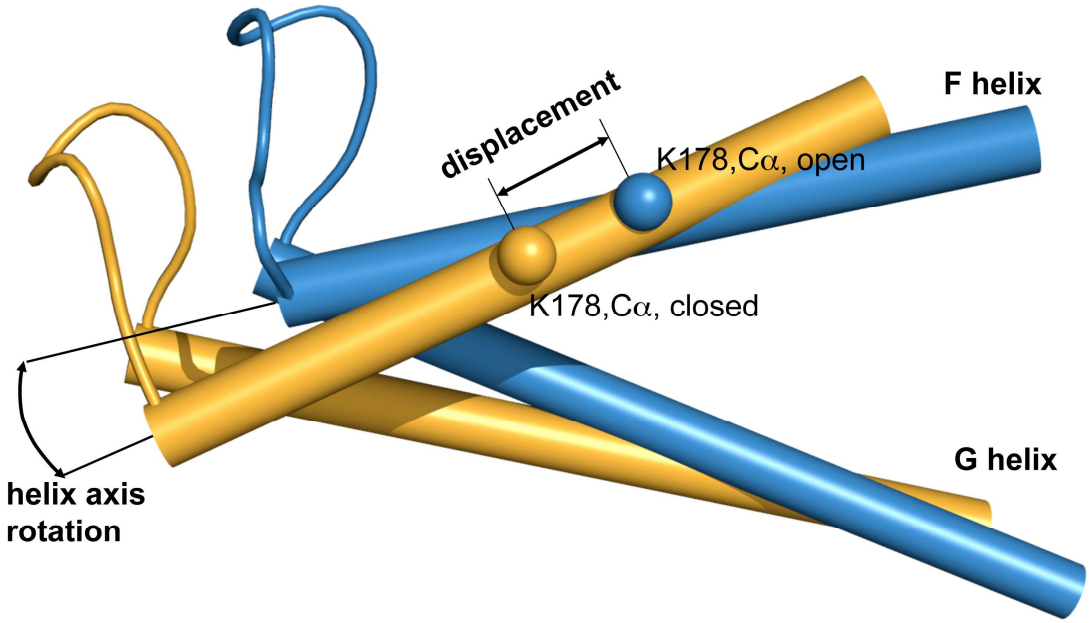

Figure S3. Graphical description of the displacement and rotation of the F and G helical axes relative to the closed conformation (PDB entry 2CPP). Open and closed structures are colored in blue and orange, respectively. Helical displacement was calculated based on the coordinate difference of C_α atoms at the helix center. Rotation of the helix axes was determined by calculation of the angle between the helix axes in the superimposed structures.

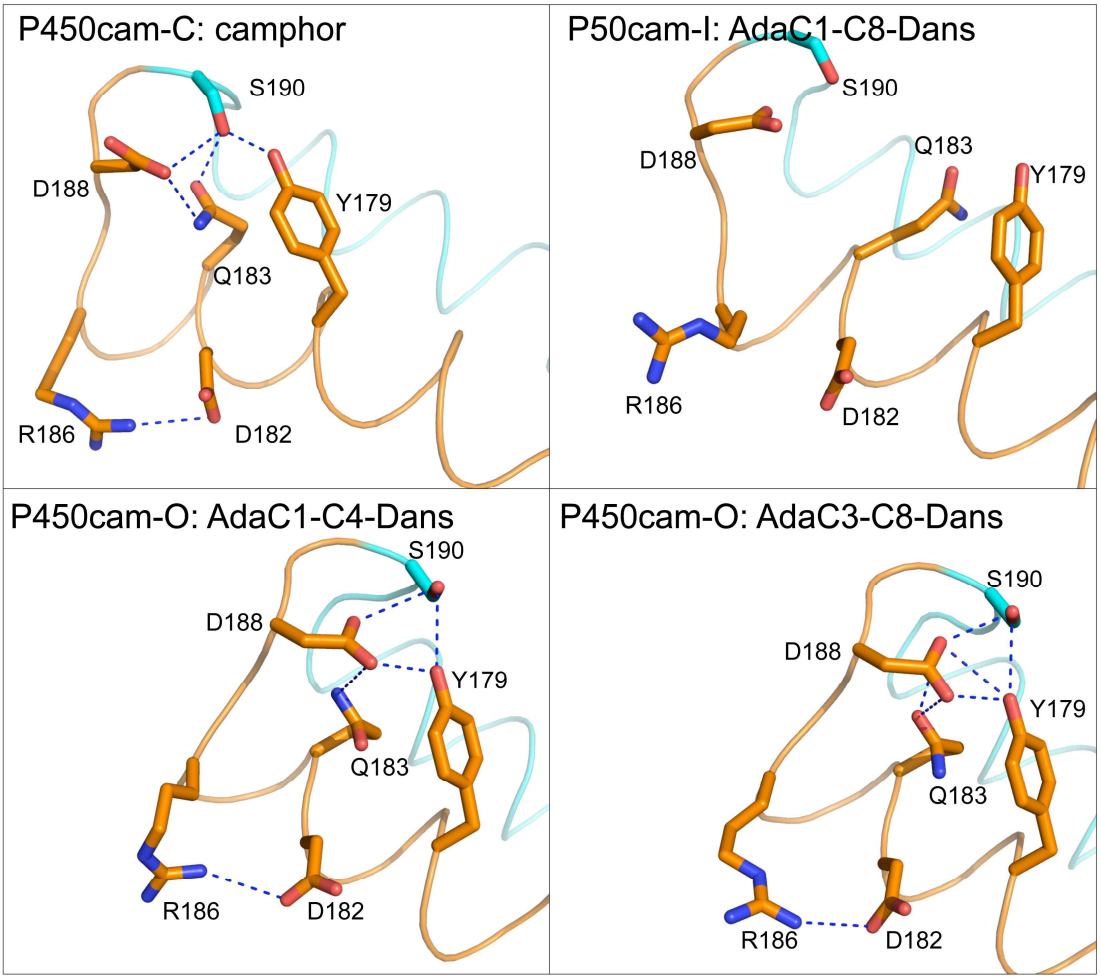

Figure S4. Contact analysis around the F-G loop. F and G helices are colored in orange and cyan, respectively. Hydrophilic interactions are dashed in blue.

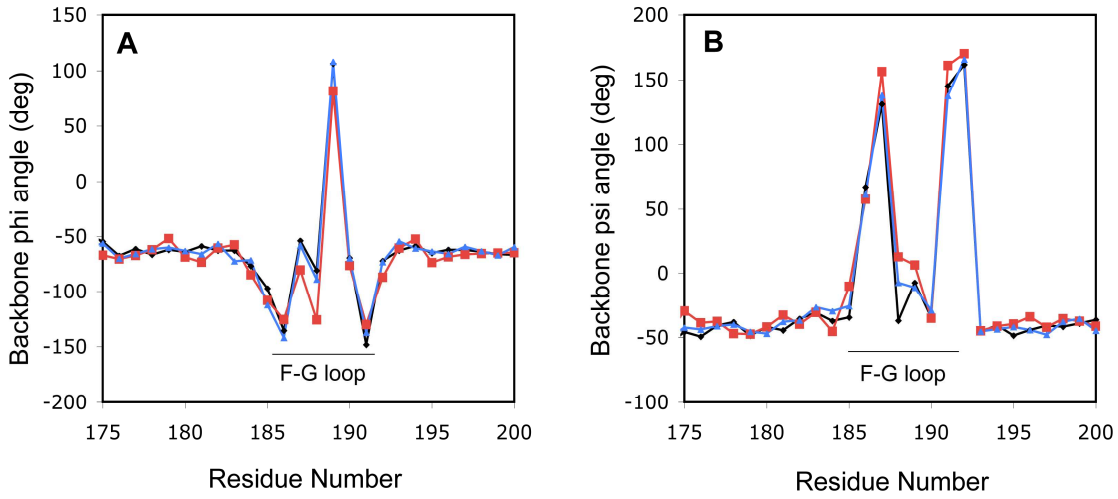

Figure S5. Backbone angles around the F-G loop. The ψ (A) and ϕ (B) angles are plotted for P450cam-C (black, PDB entry 2CPP), P450cam-I (red, PDB entry 1RE9) and P450cam-O (blue, PDB entry 3P6X).

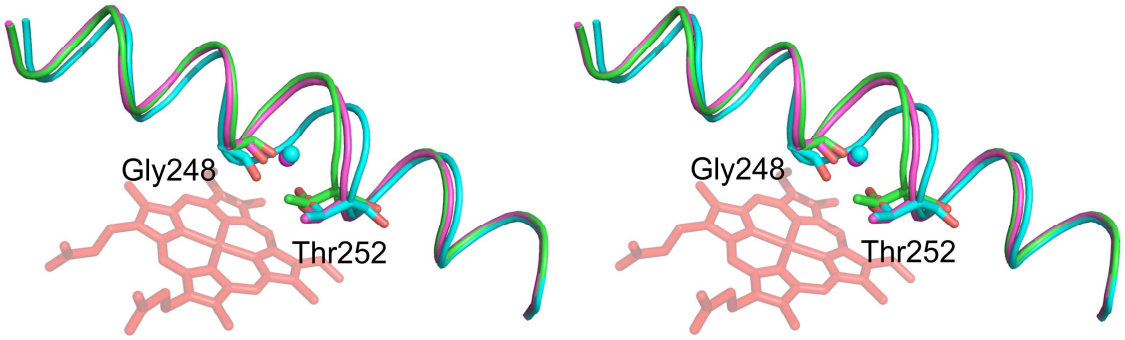

Figure S6. Stereo view of the I helix bulge for P450cam-C (green, PDB entry 2CPP), P450cam-O (cyan, PDB entry 3L62) and oxyferrous P450cam (magenta, PDB entry 1DZ8). The catalytic water is shown as a sphere and the heme is colored in red.

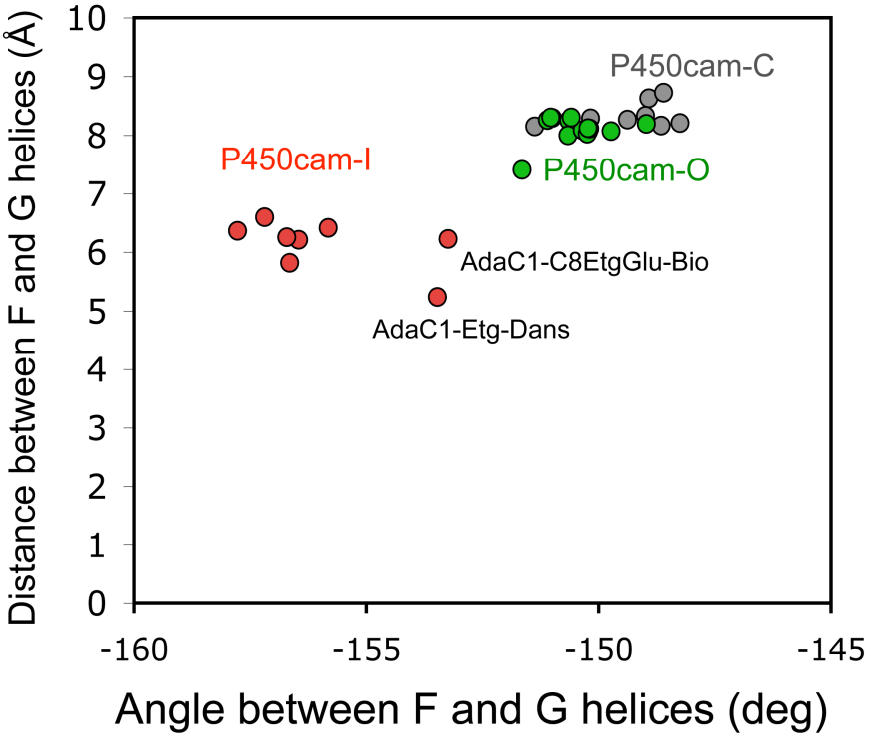

Figure S7. Two-dimensional plot of the packing angle and distance between the F and G helices in P450cam structures. Structures of P450cam-C, P450cam-I and P450cam-C forms are colored in gray, red and green.

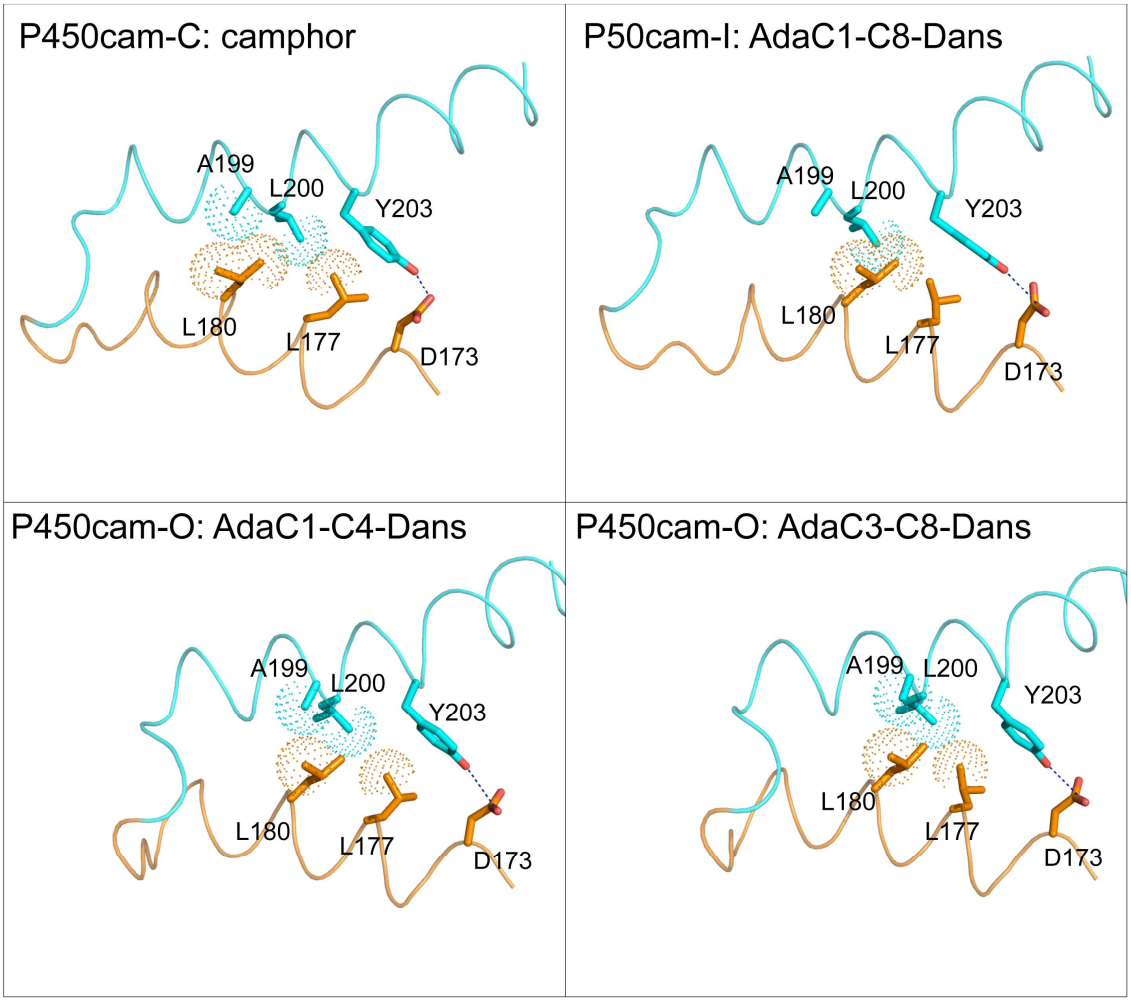

Figure S8. Contact analysis between the F and G helices for P450cam bound to camphor (P450cam-C), AdaC1-C8-Dans (P450cam-I), AdaC1-C4-Dans (P450cam-O), and AdaC3-C8-Dans (P450cam-O). The F and G helices are colored in orange and cyan, respectively. Electrostatic interactions are shown in blue dash. Hydrophobic contacts are shown in dots.