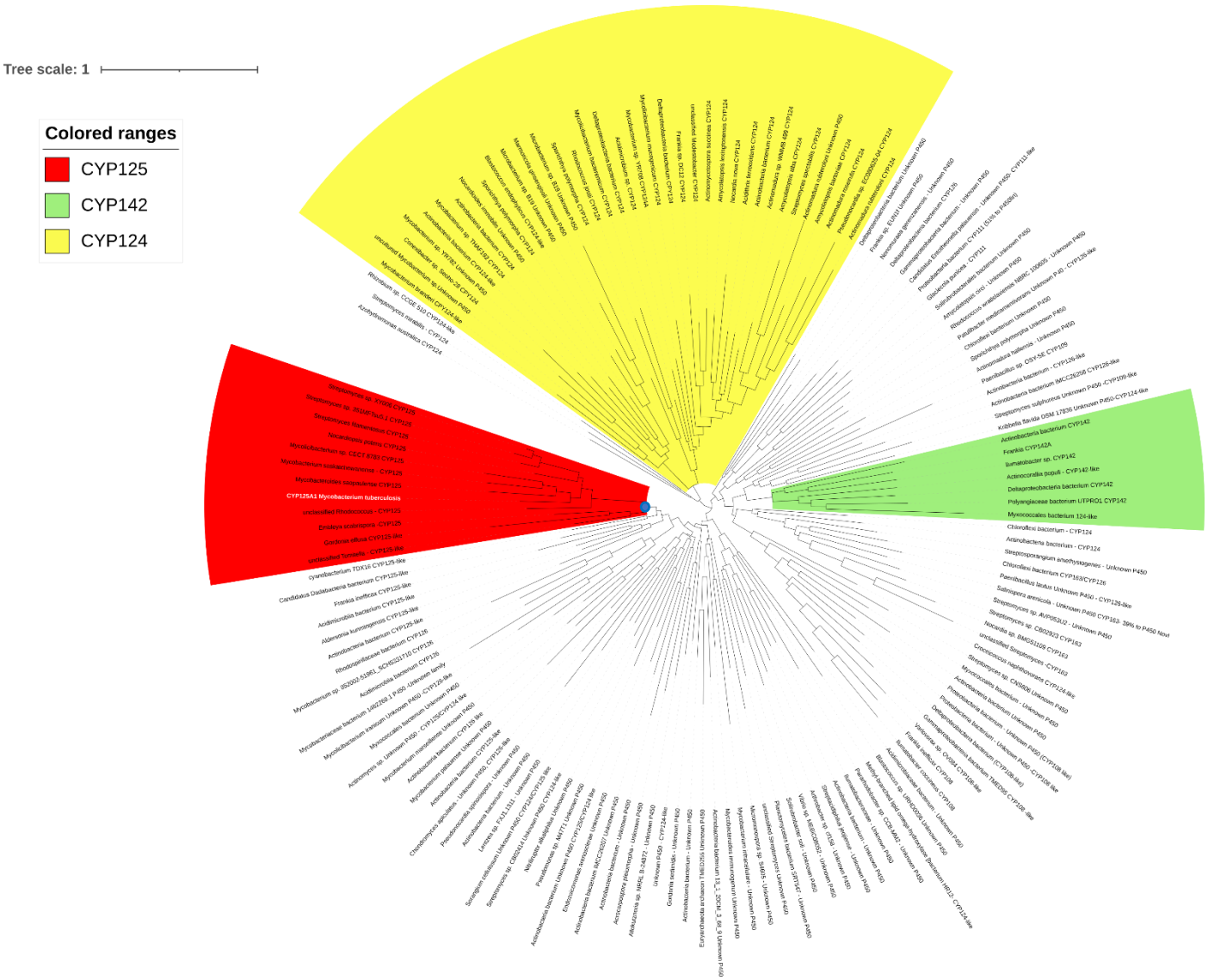


>CYP125MRCA

MTTTTMAPTDIDLTDPDVYNRGVPHEQFAWLRRNEPVYWHPEPPPDTDGEGYWAVTRHADVVAVSRDPEIFSSQQGGTMIQDADA
APEELEKQRMMMLNMDPPQHTRLRKLVSKGFTPRMIAKLEDKIRERAKQIVDEAIEKGECDFVADIAAELPLQVIAELIGVPQEDRQRL
FDWSNRMIGYDDPEYHSSEADGEQAAAEMFAYAQELAAERRKNPRDDIVTALVQAEVDGQKLSDLEFNMFFLLLVVAGNETTRNAIS
HGMLALLEHPDQWERLRADPSLAPTAVDEILRWASPVMSFRRTATRDTELGGQQIKAGDKVVMFYASANRDEEVFDDPYTFDITRSP
NPHLAFGGGGGPHYCLGANLARLEIRVMFEELAERMPDIELTGPPERLRSNFINGIKHMPVRFTPARAVGGHHHHHH

>CYP125MRCAAlt

MTTTTMAPTDIDLTDPDVYNRGVPHEQFAWLRRNEPVYWHPEPPPDTDGEGYWAVTRHADVVAVSRDPEIFSSQQGGTMIQDADA
APEELEKQRMMMLNMDPPQHTRLRKLVSKGFTPRMIAKLEDKIRERAKQIVDEAIEKGECDFVADIAAELPLQVIAELIGVPQEDRQRL
FDWSNRMIGYDDPEYHSSEADGEQAAAEMFAYAQELAAERRKNPRDDIVTALVQAEVDGQKLSDLEFNMFFLLLVVAGNETTRNAIS
HGMLALLEHPDQWERLRADPSLAPTAVDEILRWASPVMSFRRTATRDTELGGQQIKAGDKVVMFYASANRDEEVFDDPYTFDITRSP
NPHLAFGGGGGPHYCLGANLARLEIRVMFEELAERMPDIELTGPPERLRSNFINGIKHMPVRFTPARAVGGHHHHHH

Figure S1: Ancestral sequence reconstruction of CYP125MRCA. Top: Phylogenetic tree used to create the ancestral sequence reconstruction CYP125MRCA (position highlighted by the blue circle). Unique P450 clades are shown in different coloured ranges. Bottom: Constructed CYP125MRCA and CYP125MRCAAlt fasta sequence.

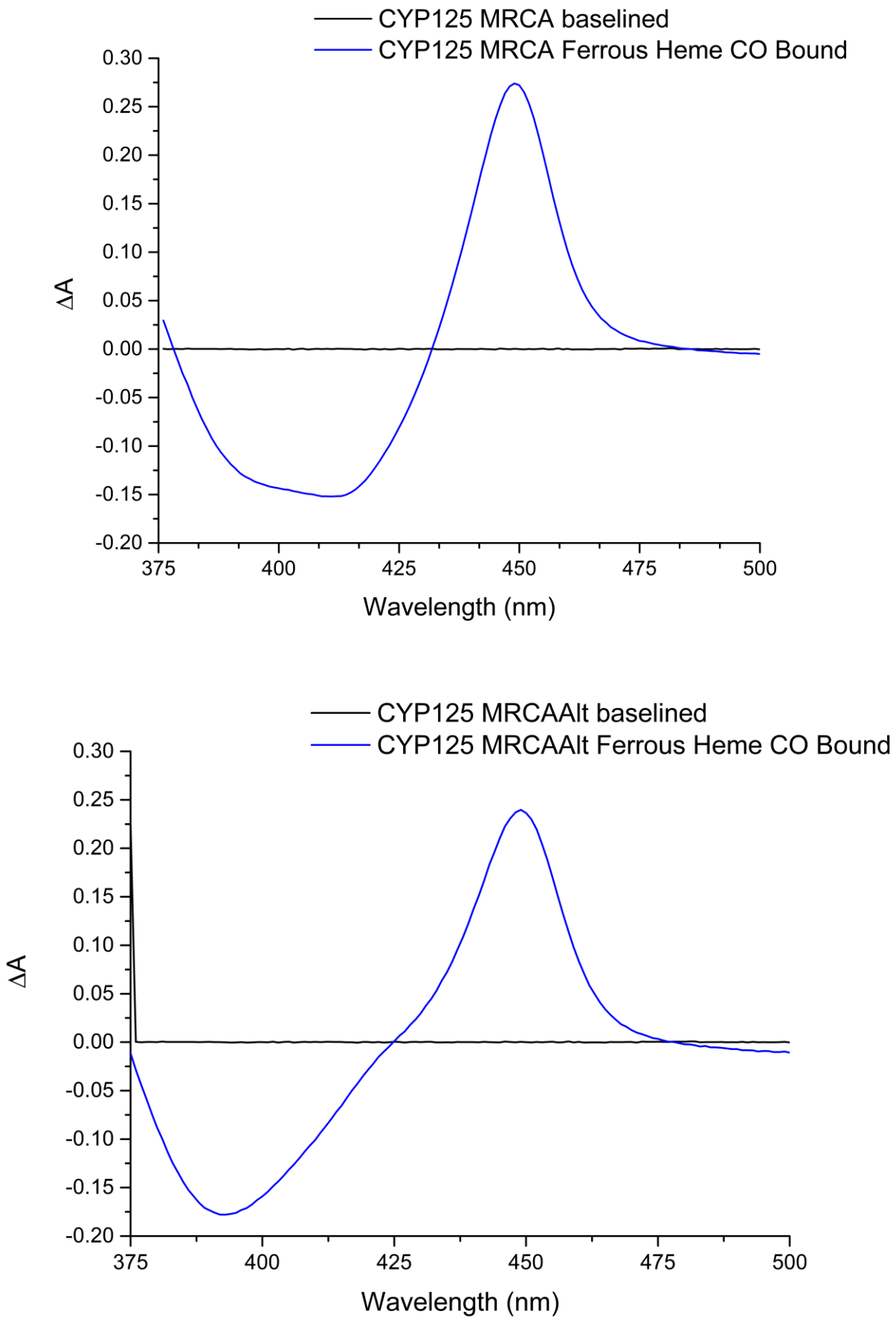


Figure S2: UV-Vis difference spectra of the binding of CO to CYP125MRCA (top) and CYP125MRCAAlt (bottom) after baselining of the ferric resting state of the protein and subsequent reduction with sodium dithionite.

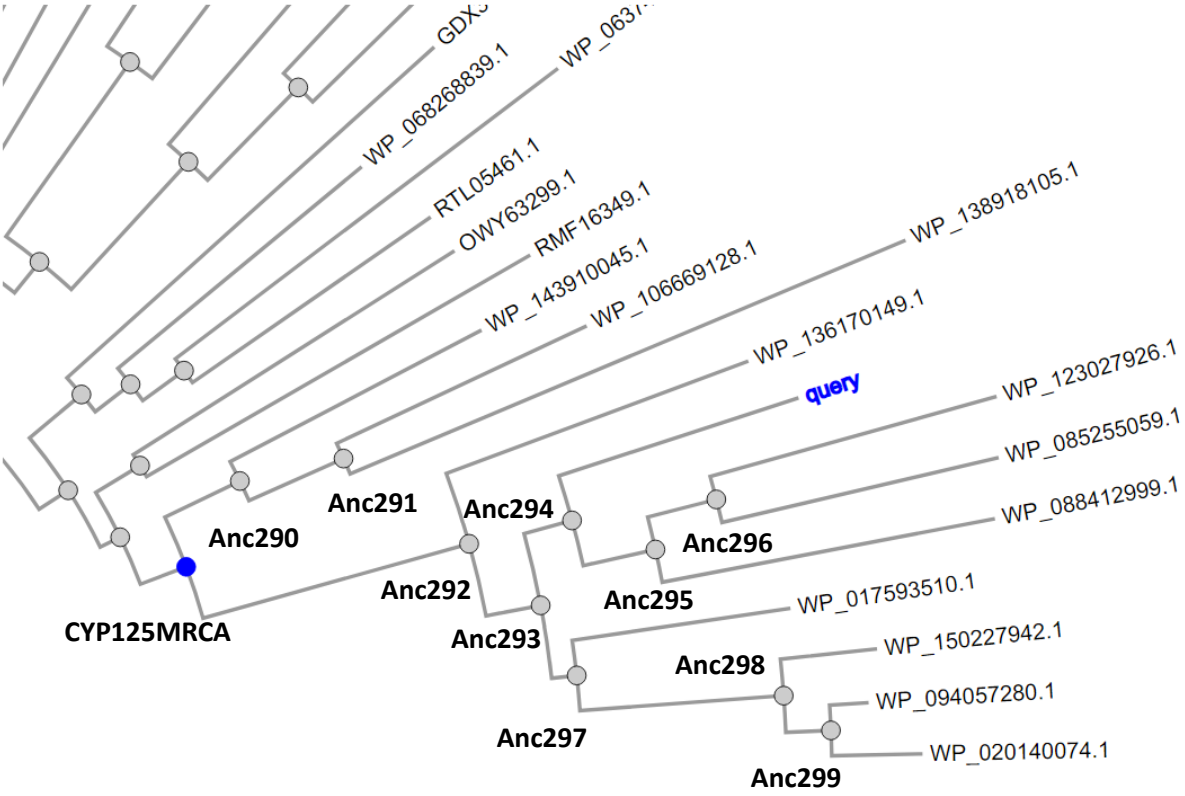


Figure S3: Zoomed in phylogenetic tree of the CYP125 clade used for ancestral sequence reconstruction. The CYP125MRCA node is shown in blue and the CYP125A1 search sequence highlighted as 'query' in blue. Extant sequence identifiers correspond to the IDs given in the alignment in Figure S3.

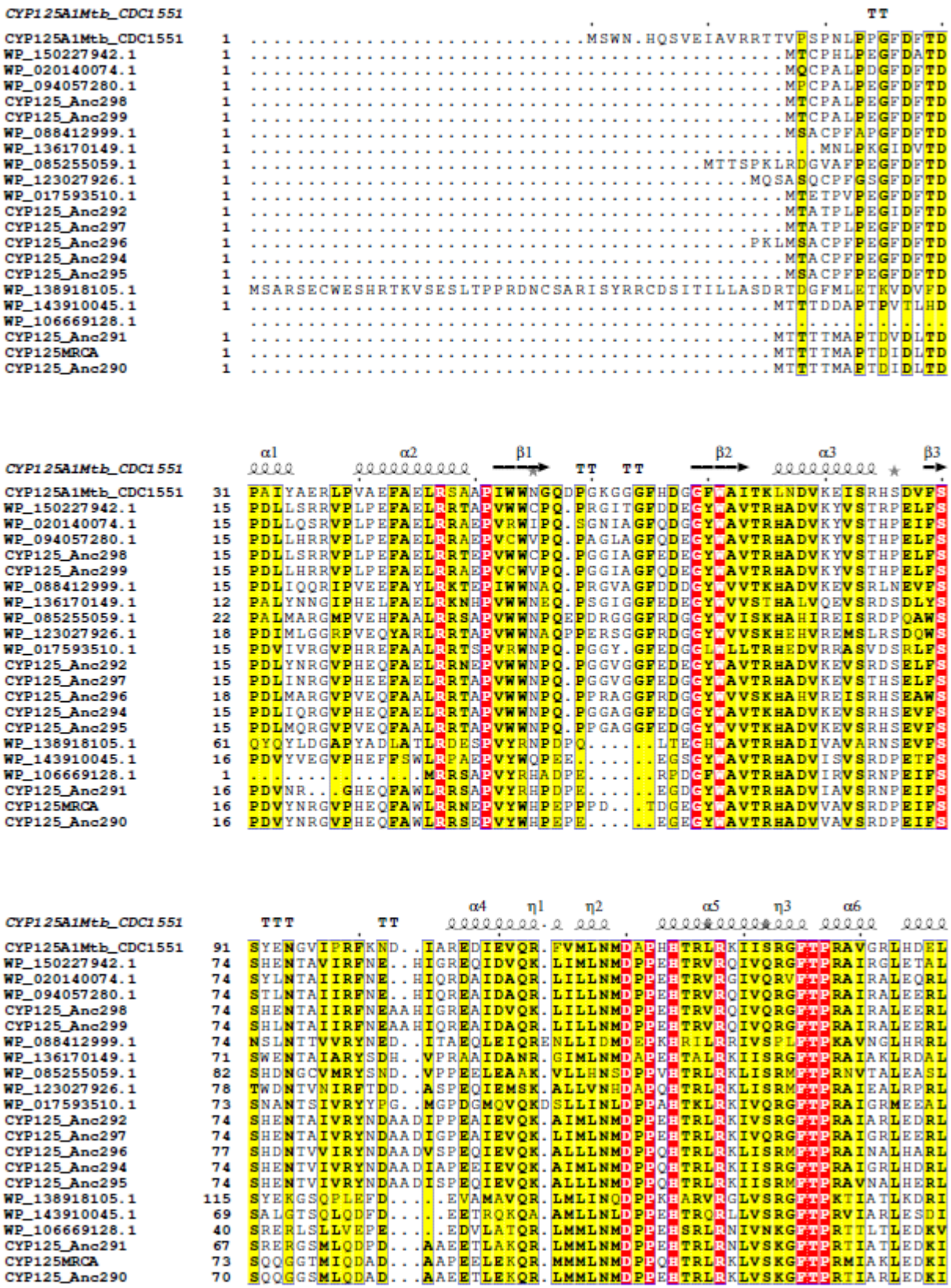


Figure S4 Sequence alignment of extant sequences and reconstructed ancestral nodes in the CYP125 clade (continued overleaf).

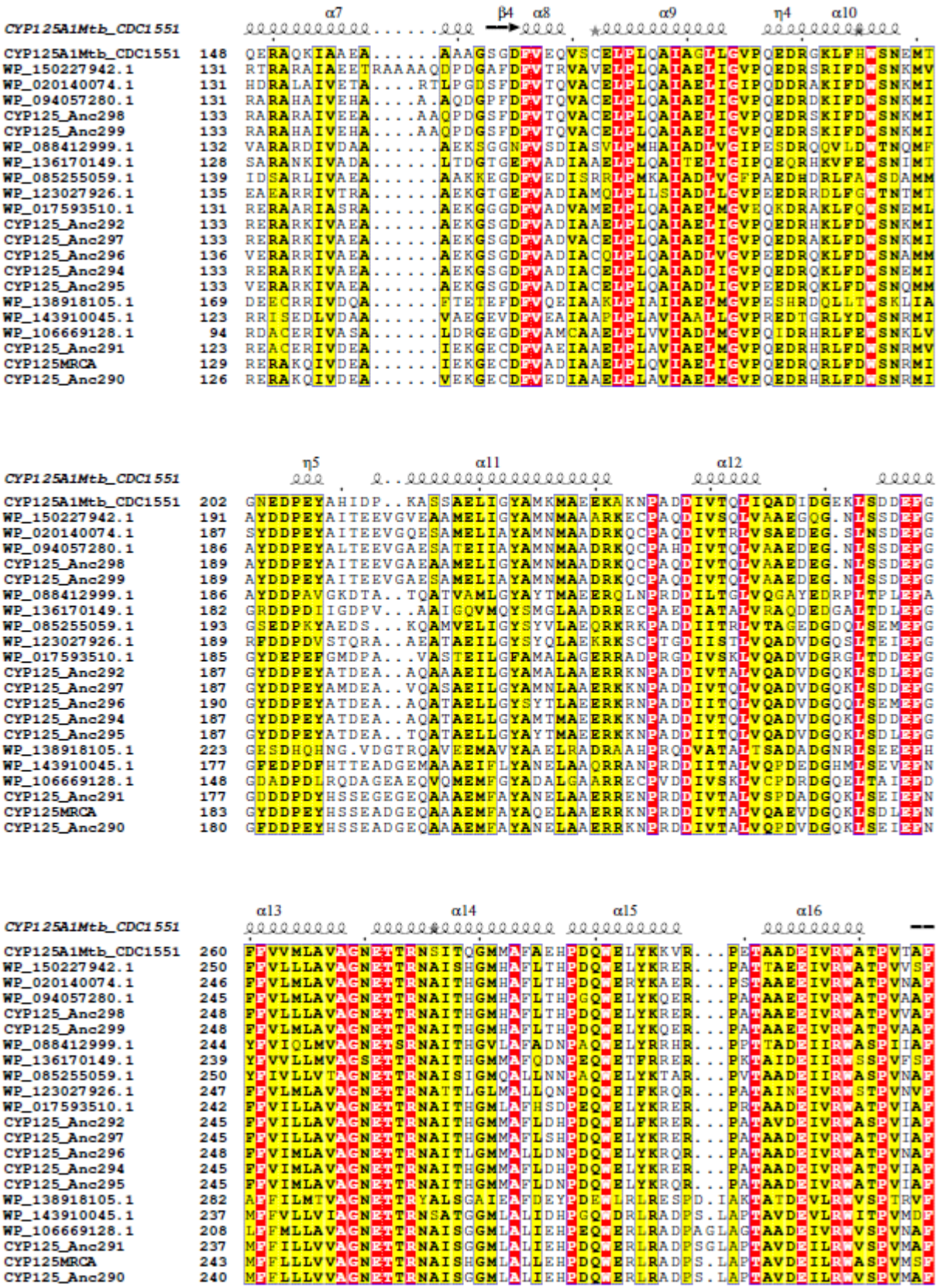


Figure S4 Sequence alignment of extant sequences and reconstructed ancestral nodes in the CYP125 clade (continued overleaf).

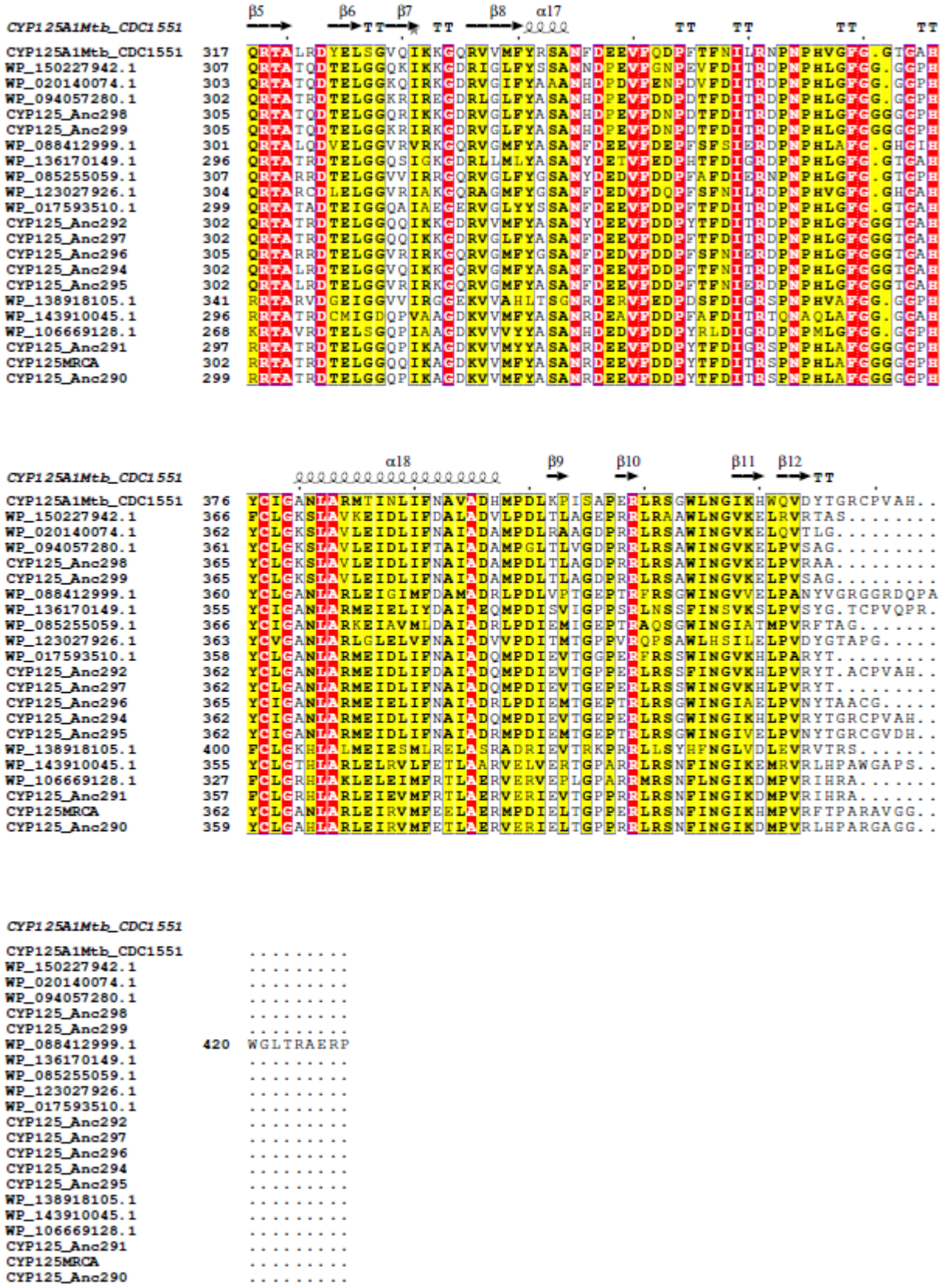


Figure S4: Sequence alignment of extant sequences and reconstructed ancestral nodes in the CYP125 clade. Positions of each sequence in the phylogenetic tree are shown in Figure S2.

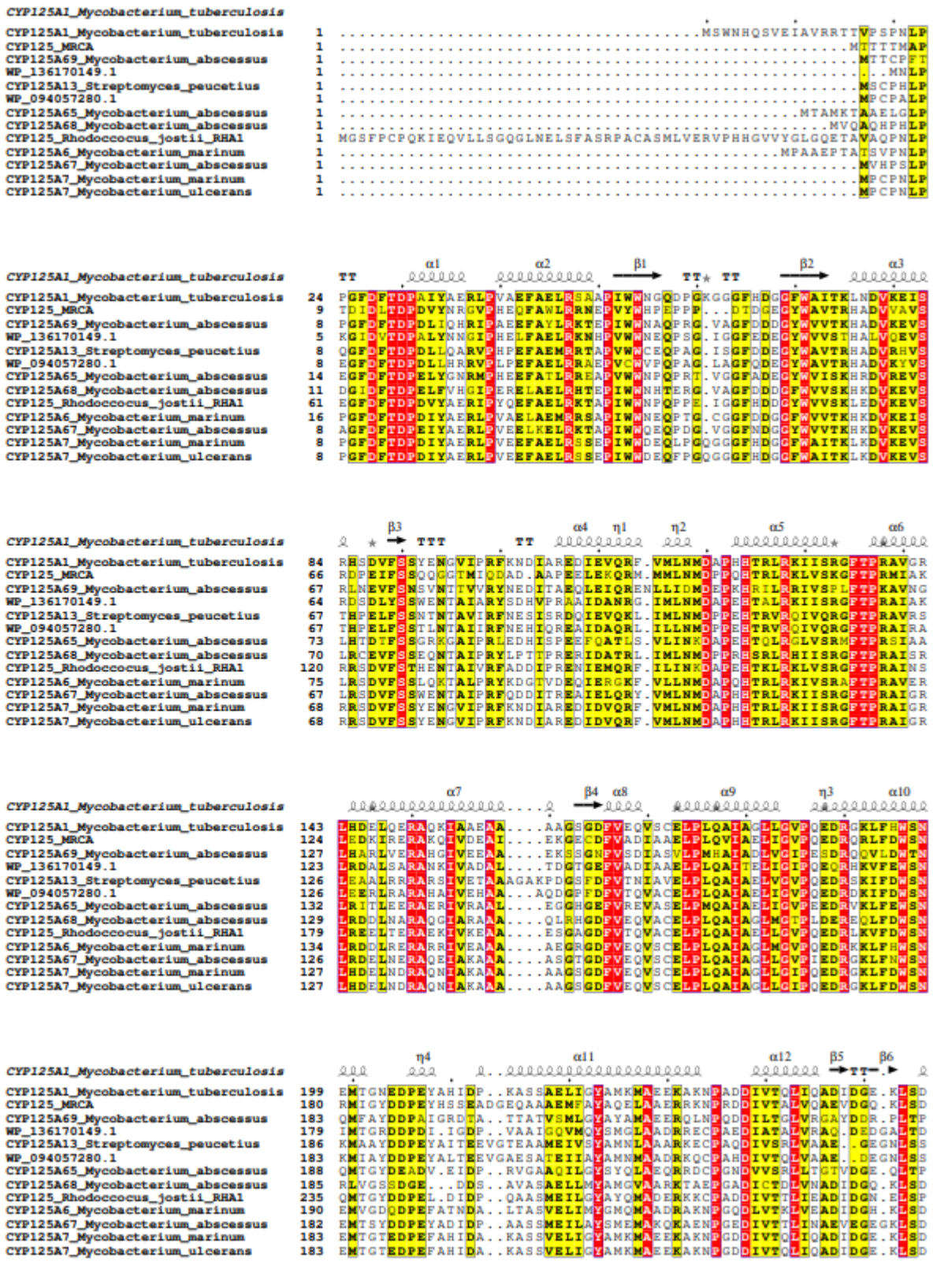


Figure S5: Multiple sequence alignment of CYP125MRCA using ClustalW (continued overleaf).

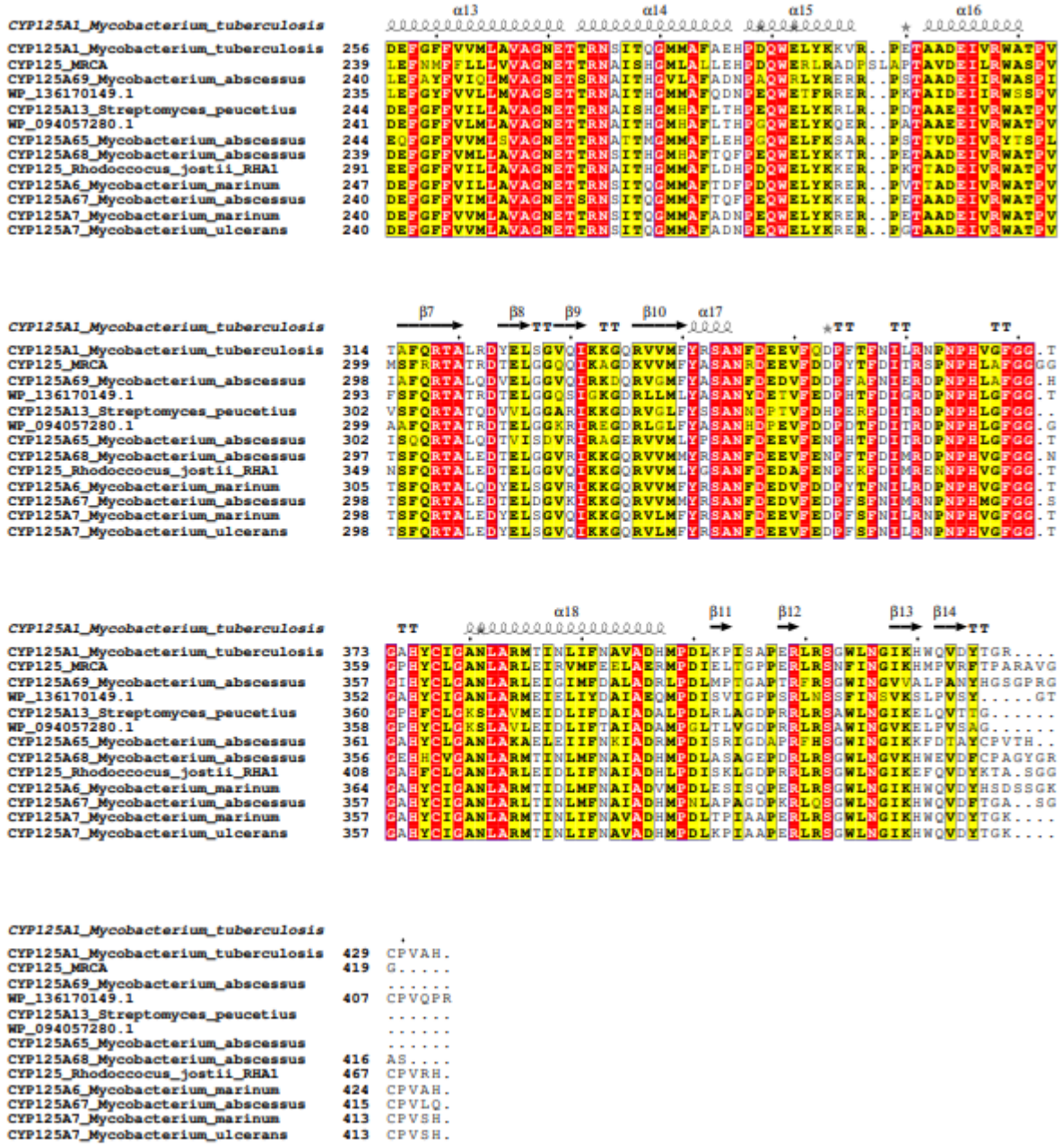


Figure S5: Multiple sequence alignment of CYP125MRCA using ClustalW. Alignment is against known CYP125 enzymes from pathogenic descendants of CYP125MRCA, along with other CYP125 enzymes representing different clades of the family's phylogenetic tree.

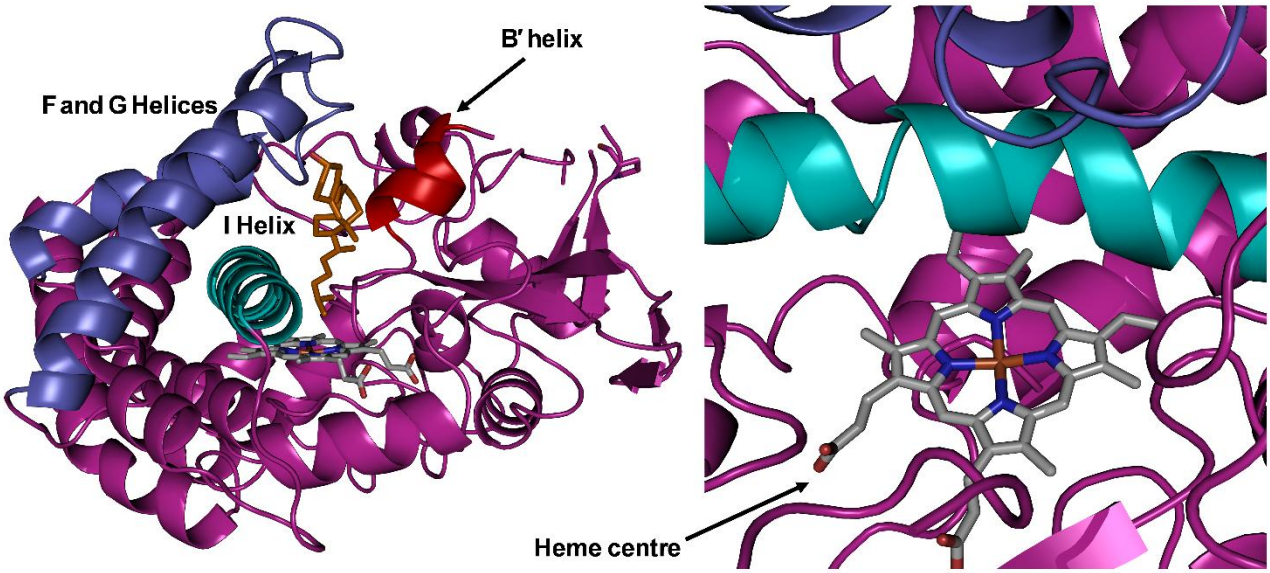


Figure S6: Overall structure of CYP125MRCA showing characteristic P450 features. Left: Cartoon depiction of the entire CYP125MRCA structure. The F/G helices are highlighted in blue, the B' helix in red and the I helix in teal. The heme centre is shown in grey and the substrate (in this case vitamin D_3) in orange. Right: A zoomed in view of the heme binding groove.

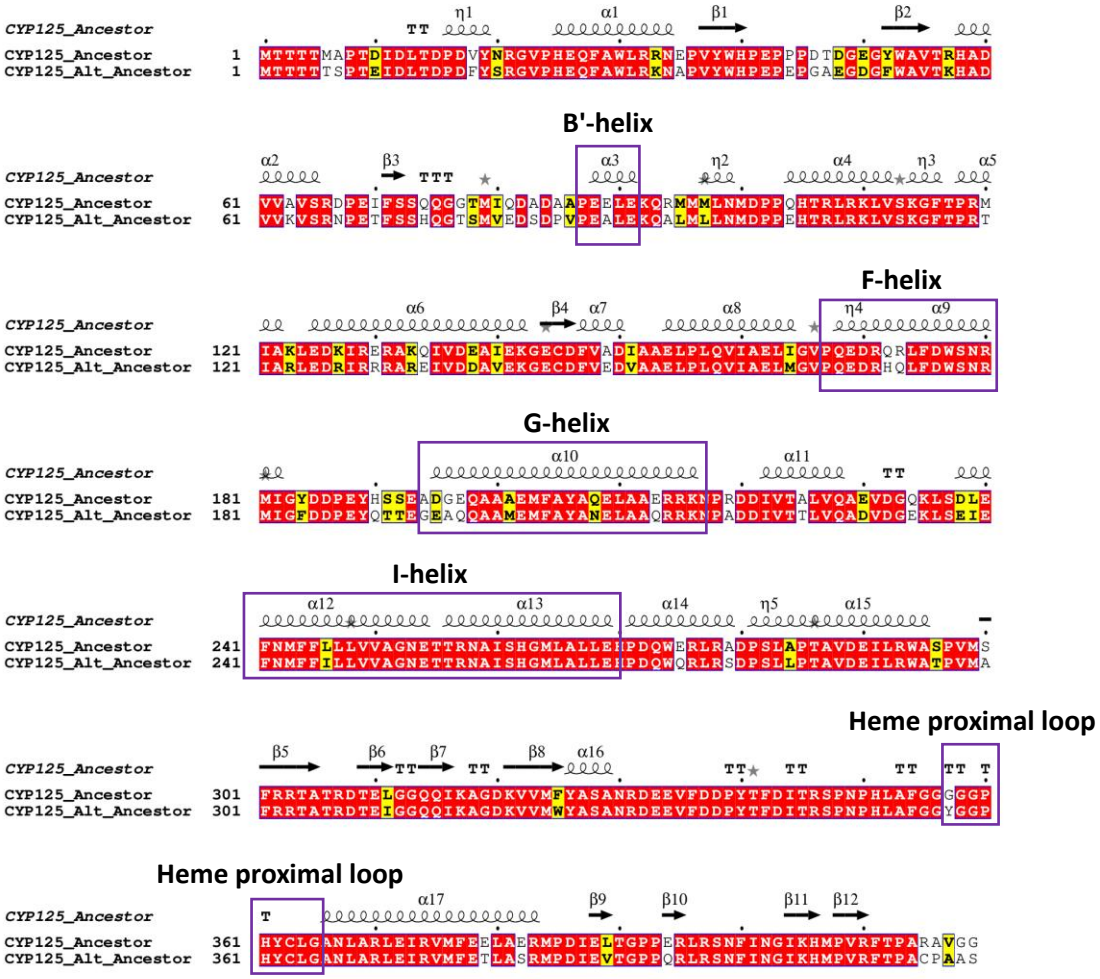


Figure S7: Sequence alignment of CYP125MRCA and CYP125MRCAAlt using ClustalW, with key regions labelled and highlighted in purple.

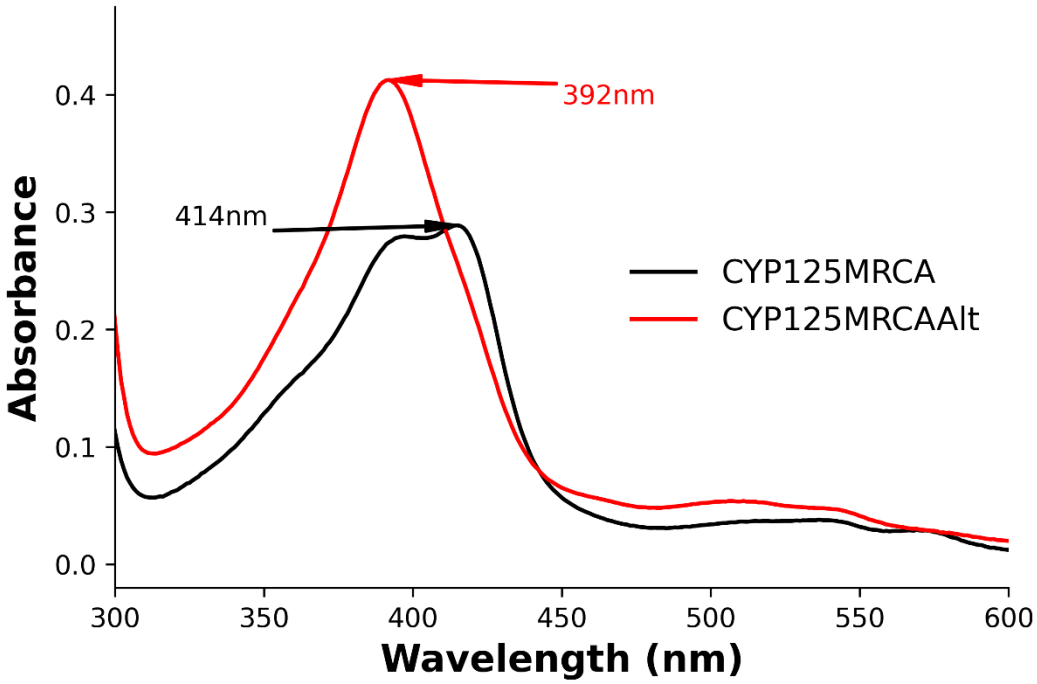


Figure S8 UV-Vis absorption spectra of the resting states of CYP125MRCA (black) and CYP125MRCAAlt (red).

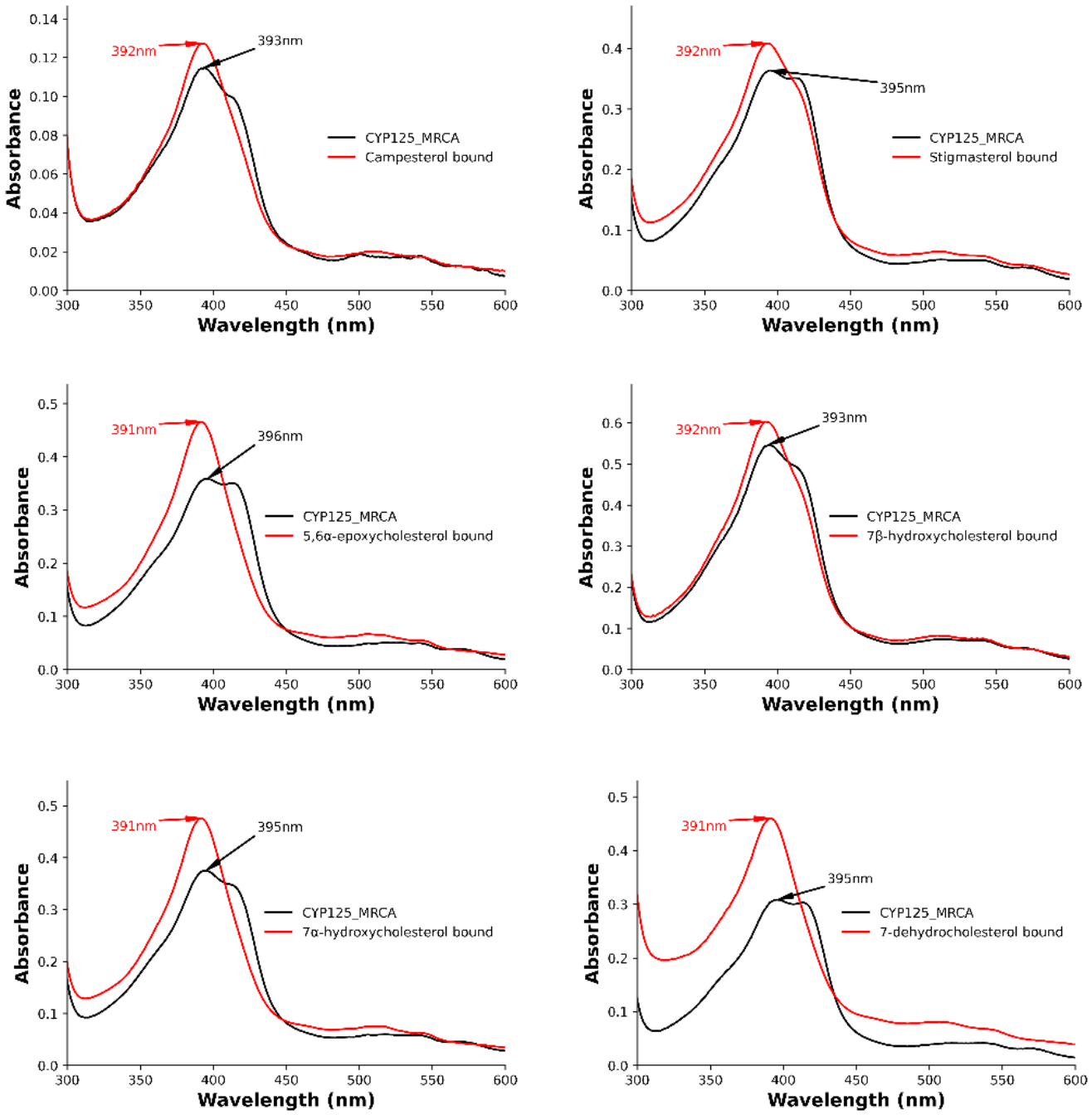


Figure S9: Selected UV-Vis spin-state shift assays used to screen the binding of various C-27 steroids to CYP125MRCA. In each case, the substrate free spectrum is shown in black while the substrate bound spectrum is shown in **red**. Substrate stocks were dissolved in either 40% EtOH or 40% hydroxypropyl-β-cyclodextrin in 50 mM Tris (pH 7.4).

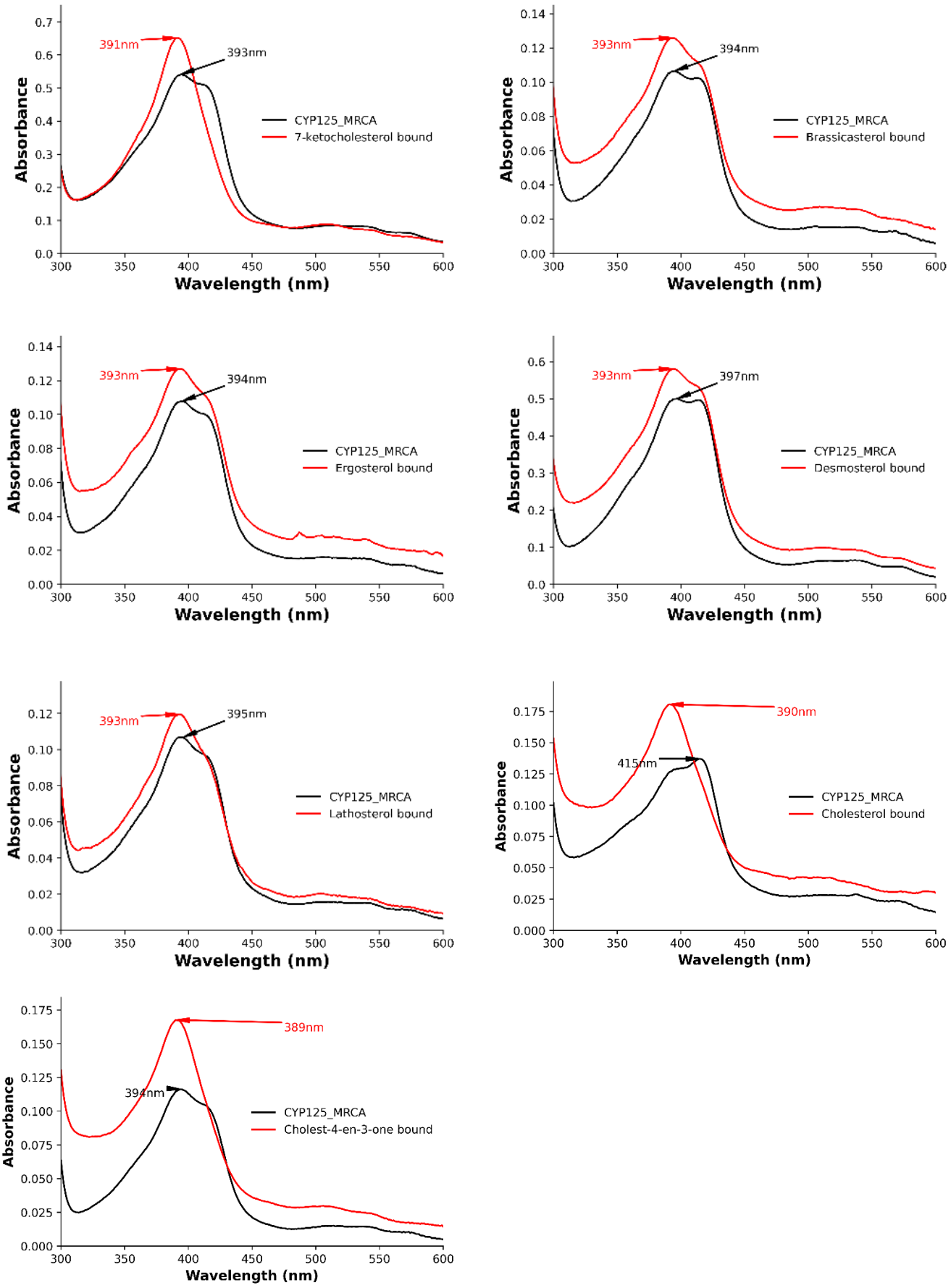


Figure S9 Continued: Selected UV-Vis spin-state shift assays used to screen the binding of various C-27 steroids to CYP125MRCA. In each case, the substrate free spectrum is shown in black while the substrate bound spectrum is shown in **red**. Substrate stocks were dissolved in either 40% EtOH or 40% hydroxypropyl-β-cyclodextrin in 50 mM Tris (pH 7.4).

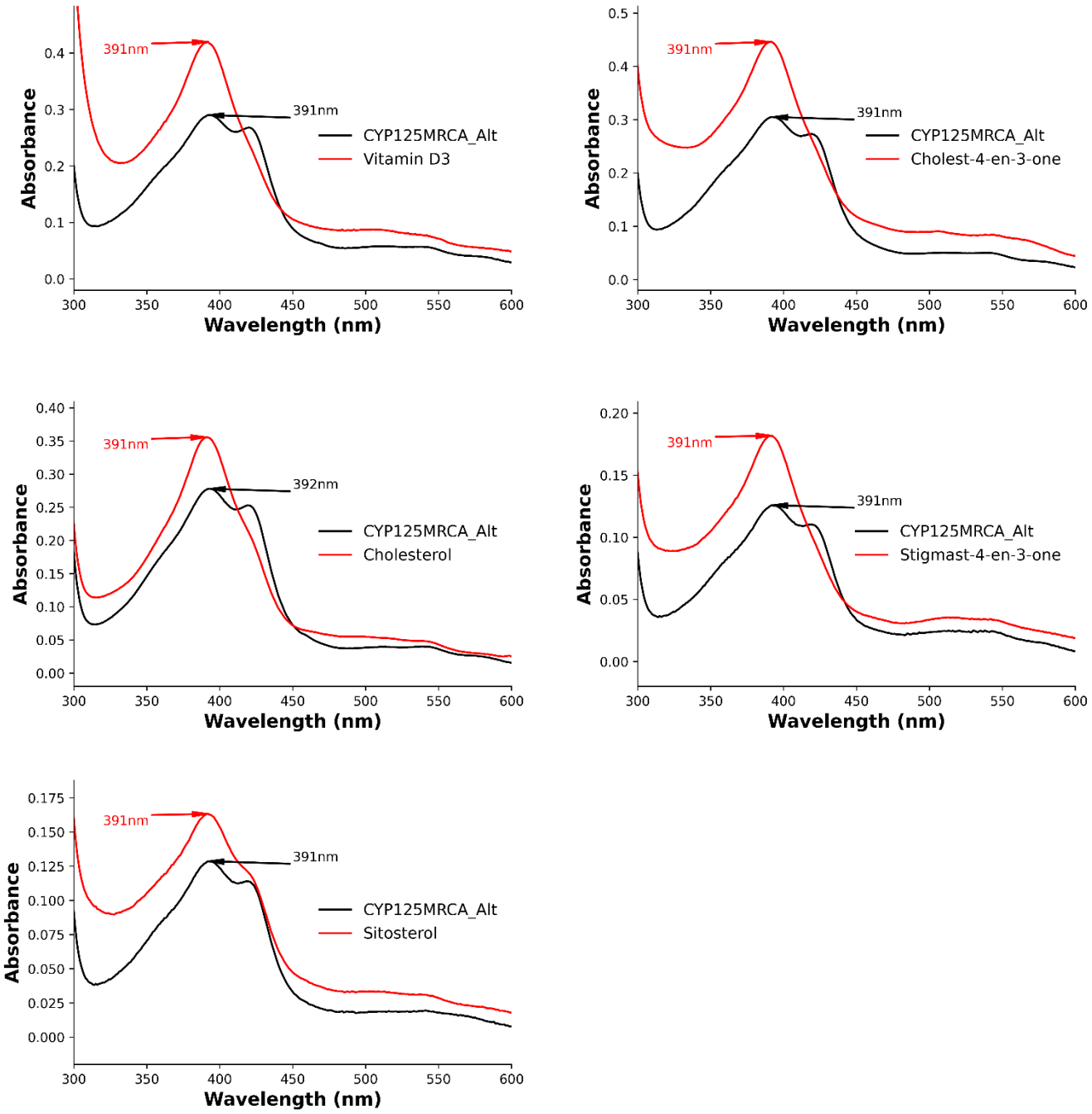


Figure S10: Selected UV-Vis spin-state shift assays used to screen the binding of various C-27 steroids to CYP125MRCAAlt. In each case, the substrate free spectrum after addition of imidazole (1.3 mM) is shown in black while the substrate bound spectrum is shown in **red**. Substrate stocks were dissolved in either 40% EtOH or 40% hydroxypropyl-β-cyclodextrin in 50 mM Tris (pH 7.4). Complete Type II shifts were avoided as saturating concentrations of imidazole interfered with the ability of substrates to cause spin-state shifts at reasonable concentrations.

Cholesterol, cholest-4-en-3-one and stigmast-4-en-3-one binding resulted in > 90% HS heme centres, while sitosterol resulted in > 80%. The shift induced by sitosterol marked a difference from CYP125MRCA in which the Soret band was shifted completely to HS, though direct comparisons of the degree of spin-state shift between the two ancestors was complicated by the addition of imidazole to CYP125MRCAAlt prior to analysis.

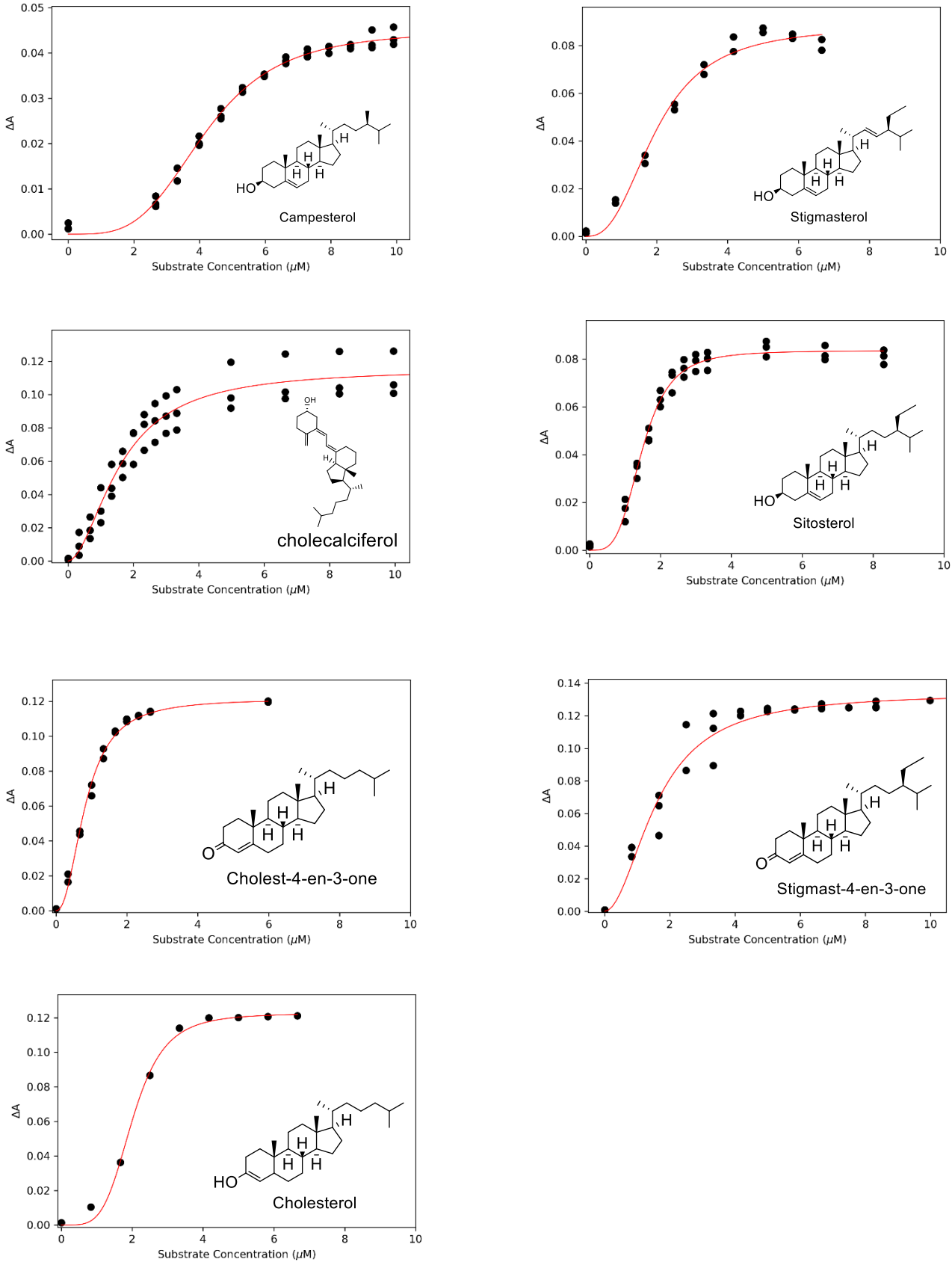


Figure S11: UV-Vis binding titration curves for various C-27 steroids to CYP125MRCA, using the Hill equation (Equation 1 in main text) in each case. Enzyme concentrations for the titrations were between 2.1-2.8 μM (shown in Table S1). Substrate stocks were between 1 and 10 mM in 40% hydroxypropyl-β-cyclodextrin.

Table S1: Binding constants determined through analysis of UV-Vis difference spectra upon substrate titration for different binding models. The K_d values used to compare substrate efficiencies were from the highlighted Hill model column, given their lower relative error.

Substrate	K_d Michaelis Menten	K_d Morrison E fit (E)	K_d Morrison E set (E)	K_d Hill Equation (n)
Stigmasterol	1.5 ± 0.7	0.006 ± 0.013 (4.0)	0.6 ± 0.4 (2.3)	2.04 ± 0.12 (2.36)
Sitosterol	1.1 ± 0.3	0.004 ±0.011 (2.8)	0.01 ± 0.02 (2.6)	1.48 ± 0.03 (3.74)
Campesterol	3.1± 1.5	1.5 ± 3.6 (4.0)	2.3 ± 1.2 (2.1)	4.19 ± 0.05 (3.68)
Cholest-4-en-3-one	0.6 ± 0.1	0.04 ± 0.01 (1.7)	0.0002 ± 0.0082 (2.8)	0.85 ± 0.02 (2.15)
Cholecalciferol	1.4 ± 0.3	1.2 ± 0.8 (2.0)	0.8 ± 0.2 (2.6)	1.60 ± 0.22 (1.84)
Stigmast-4-en-3-one	1.1 ± 0.3	0.1 ± 0.09 (3.1)	0.4 ± 0.13 (2.1)	1.61 ± 0.11 (1.98)

The Hill equation gave the best fit evidenced by the lowest relative standard error of K_d compared to the Morrison and Michaelis-Menten models (see Table S1 for model comparisons). Given that *n* is greater than 1 for all substrates, some level of cooperativity in ligand binding is present. While differences are observed between the different fitting methods the overall trends are similar. Differences in affinity and in the values of n may reflect complex interactions between the enzyme, ligand and cyclodextrin.

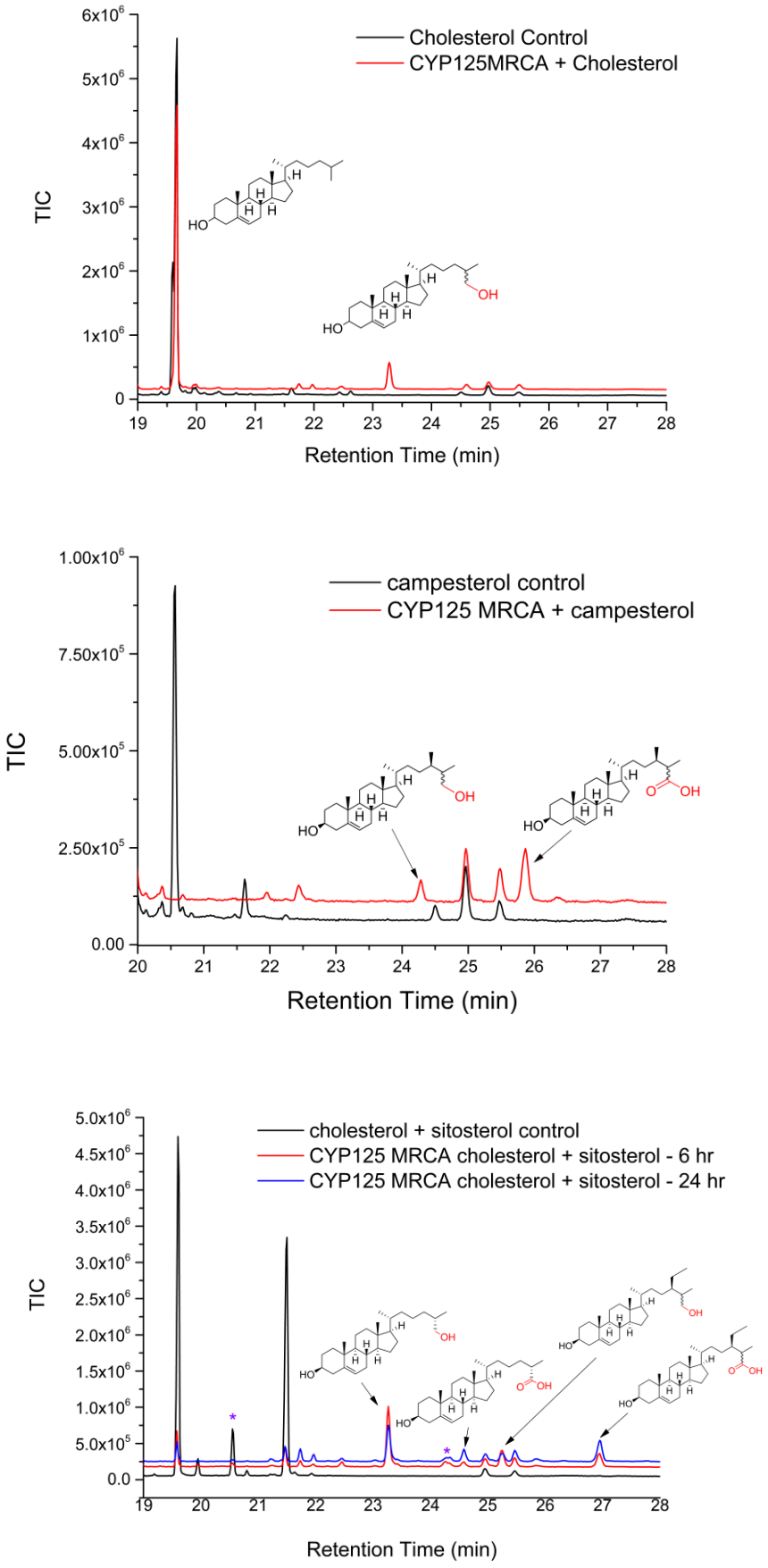
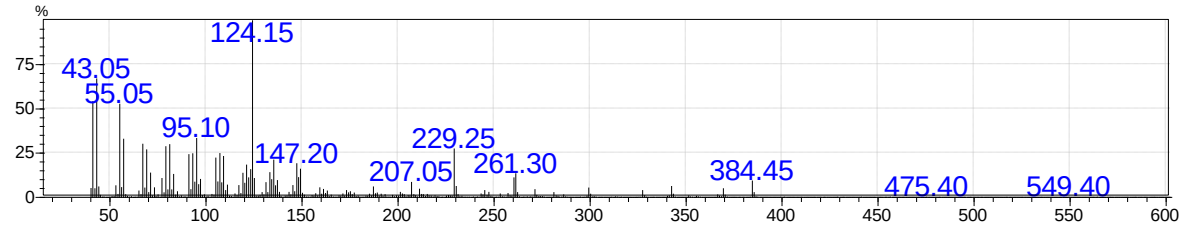
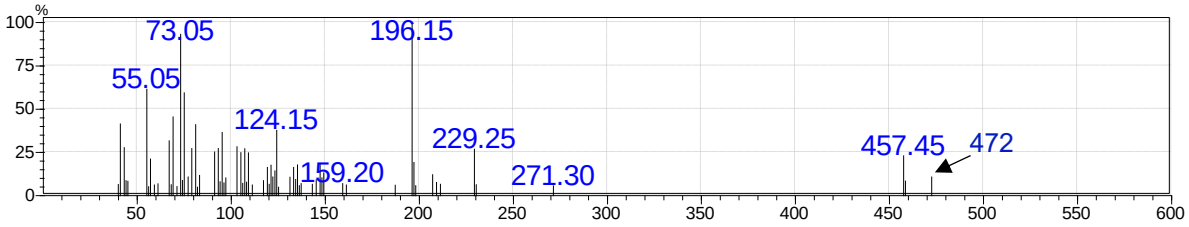


Figure S12: GC-MS chromatograms of CYP125MRCA mediated oxidation of cholesterol (top), campesterol (middle) and a 1:1 mixture of cholesterol and sitosterol (bottom). In each case the substrate control is shown in black, and the CYP125MRCA mediated oxidation reaction in **red** and **blue**. Products were identified by their mass fragmentation pattern and relative retention times. * Indicates peaks arising from a campesterol impurity in the sitosterol stock.

Cholest-4-en-3-one and oxidation products



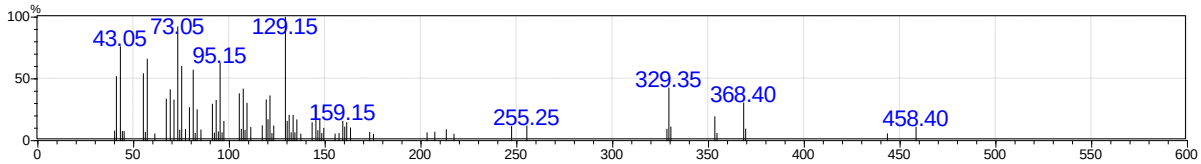
RT: 20.57 – 20.67 min, Cholest-4-en-3-one (parent ion mass: 384 m/z)



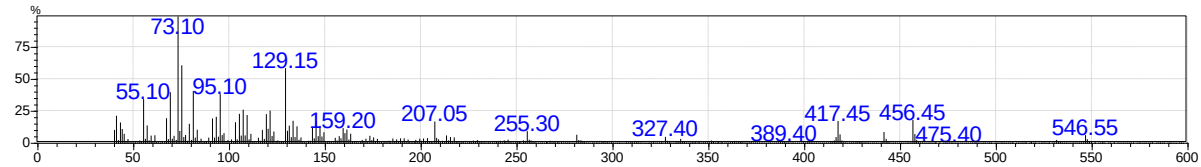
RT: 25.08 – 25.23 min, 26-hydroxycholest-4-en-3-one (parent ion mass: 472 m/z)

Figure S13: (a) Mass spectra of GC-MS separated cholest-4-en-3-one substrate and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.

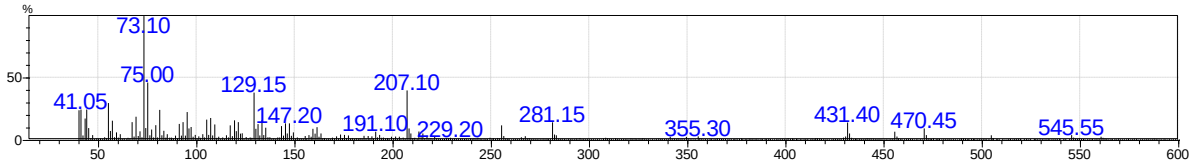
Cholesterol and oxidation products



RT: 19.53 – 19.62 min, Cholesterol (parent ion mass: 458 m/z)



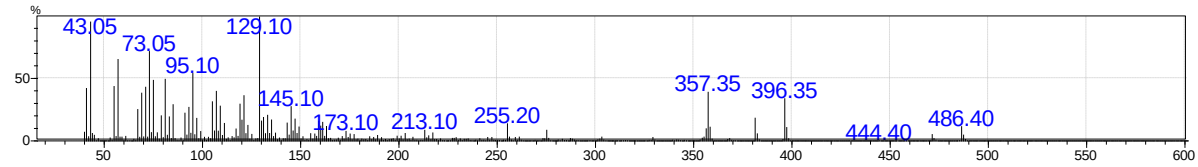
RT: 23.22 – 23.35 min, 26-hydroxycholesterol (parent ion mass: 546 m/z)



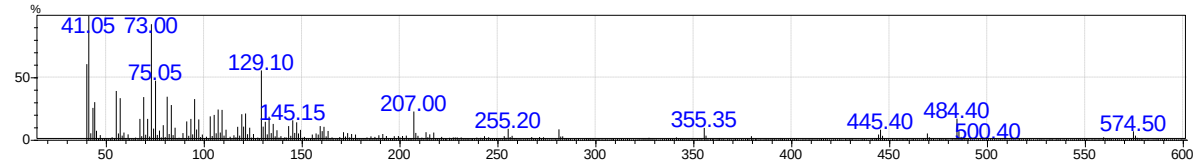
RT: 24.53 – 24.67 min, 26-cholestenoic acid (parent ion mass: 560 m/z)

Figure S13: (b) Mass spectra of GC-MS separated cholesterol substrate and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.

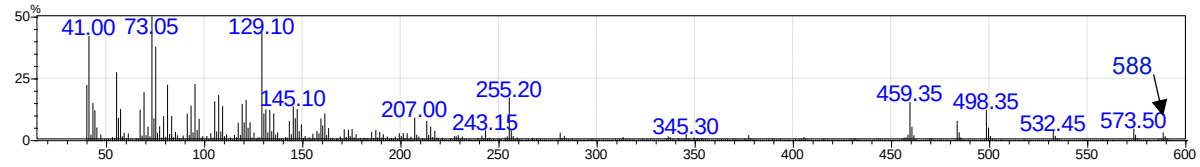
Sitosterol and oxidation products



RT: 21.42 – 21.53 min, Sitosterol (parent ion mass: 486 m/z)



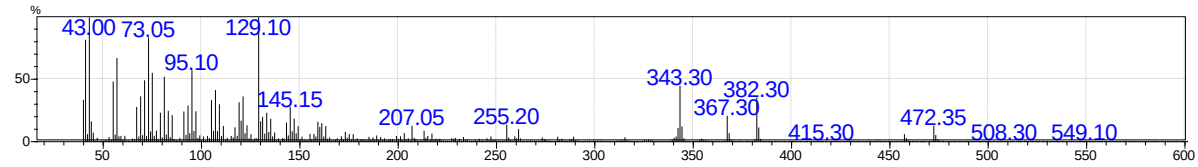
RT: 25.20 – 25.32 min, 26-hydroxysitosterol (parent ion mass: 574 m/z)



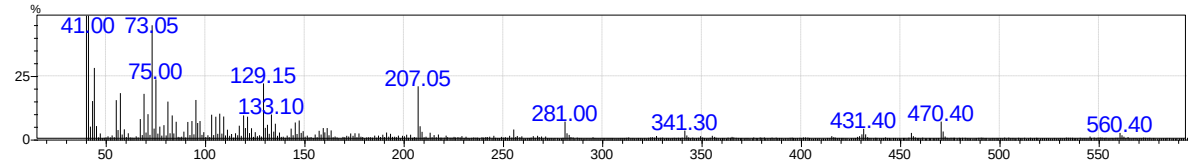
RT: 26.83 – 27.07 min, 26-sitostenoic acid (parent ion mass: 588 m/z)

Figure S14: (a) Mass spectra of GC-MS separated sitosterol substrate and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.

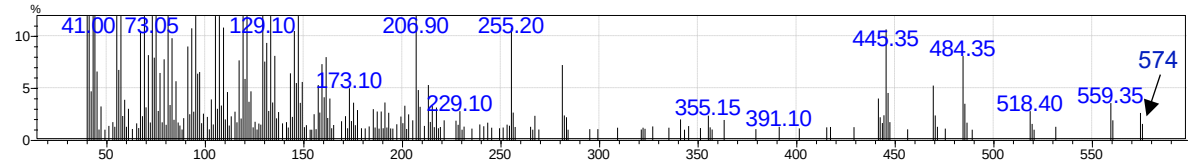
Campesterol and oxidation products



RT: 20.48 – 20.63 min, Campesterol (parent ion mass: 472 m/z)



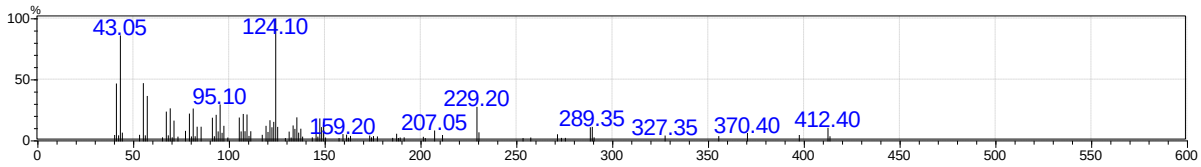
RT: 24.20 – 24.32 min, 26-hydroxycampesterol (parent ion mass: 560 m/z),



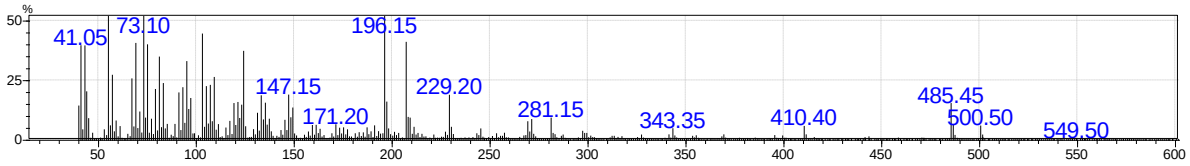
RT: 25.72 – 26.00 min, 26-campestenoic acid (parent ion mass: 574 m/z)

Figure S14: (b) Mass spectra of GC-MS separated campesterol substrate and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.

Stigmast-4-en-3-one and oxidation products

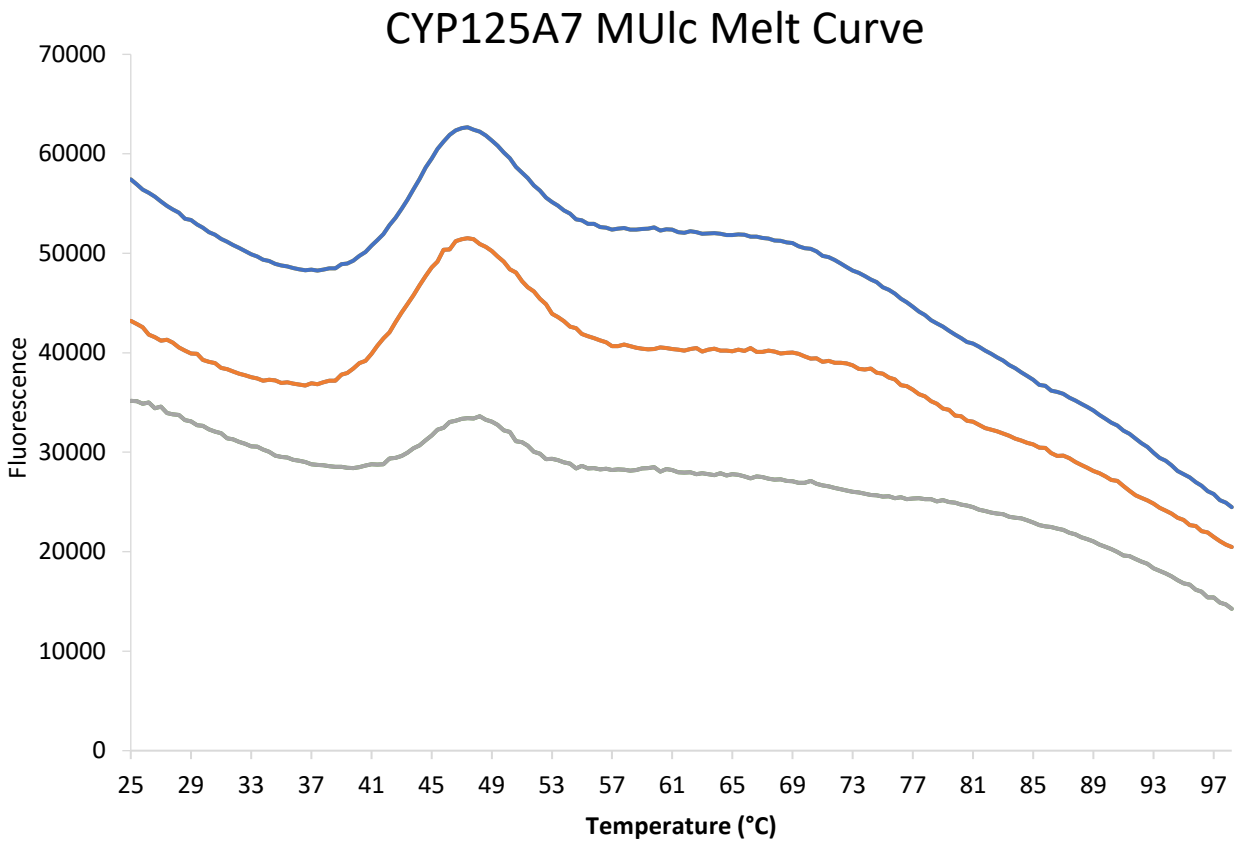


RT: 22.73 – 22.87 min, Stigmast-4-en-3-one (parent ion mass: 412 m/z)



RT: 27.28 – 27.52 min, 26-hydroxystigmast-4-en-3-one (parent ion mass: 500 m/z)

Figure S14: (c) Mass spectra of GC-MS separated stigmast-4-en-3-one substrate and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.



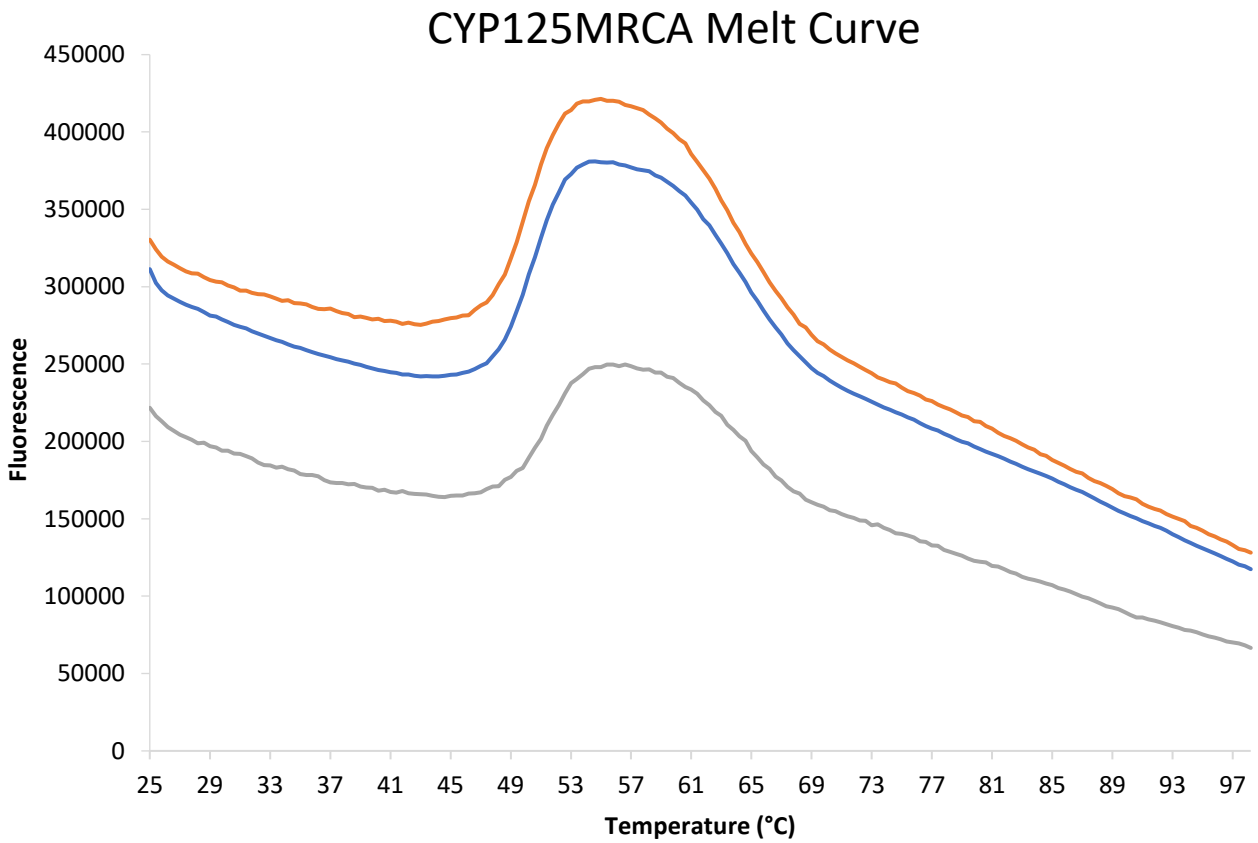


Figure S15: Raw, normalised fluorescence data for protein thermal shift assays (PTS assays), conducted in triplicate, for CYP125A7 from *Mycobacterium ulcerans* and CYP125MRCA. The dye and protocol used were sourced from ThermoFischer Scientific (Protein Thermal Shift Dye Kit™, catalog number 4461146).

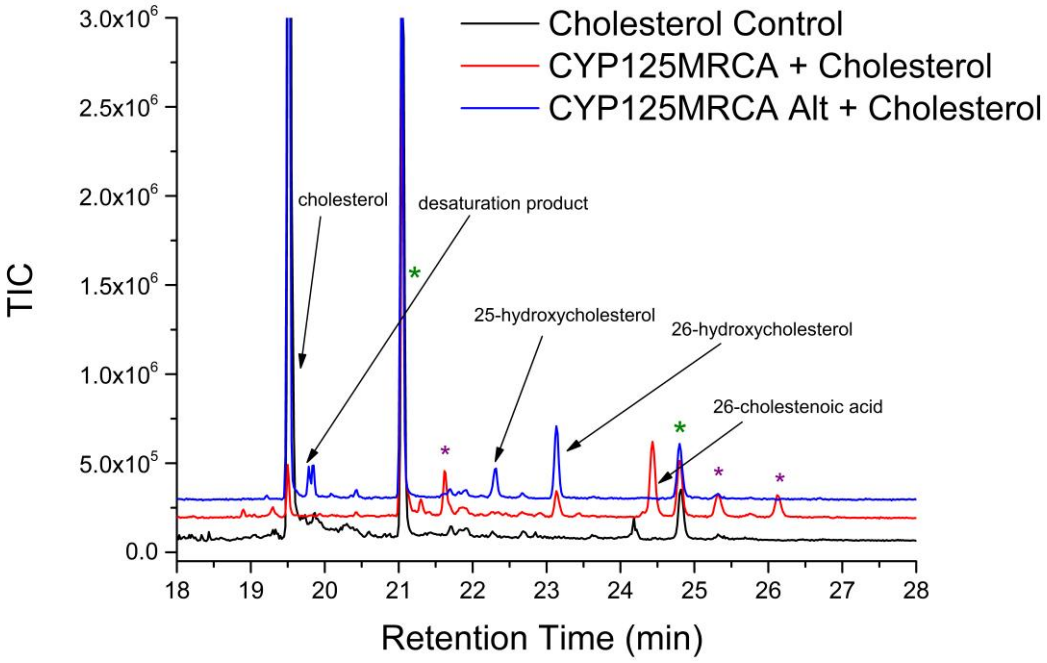
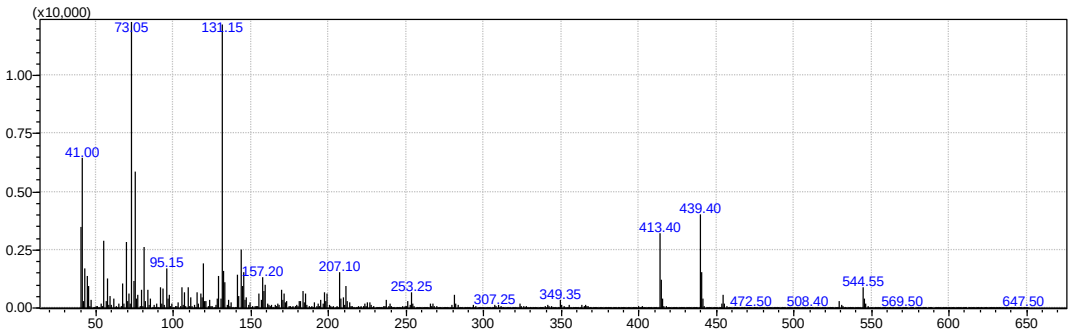


Figure S16: GC-MS chromatograms of CYP125MRCA and CYP125MRCAAlt mediated oxidation of cholesterol. Products were identified by their mass fragmentation pattern and relative retention times. * indicates significant unidentified impurity peaks.* indicates the presence of unknown oxidation products.

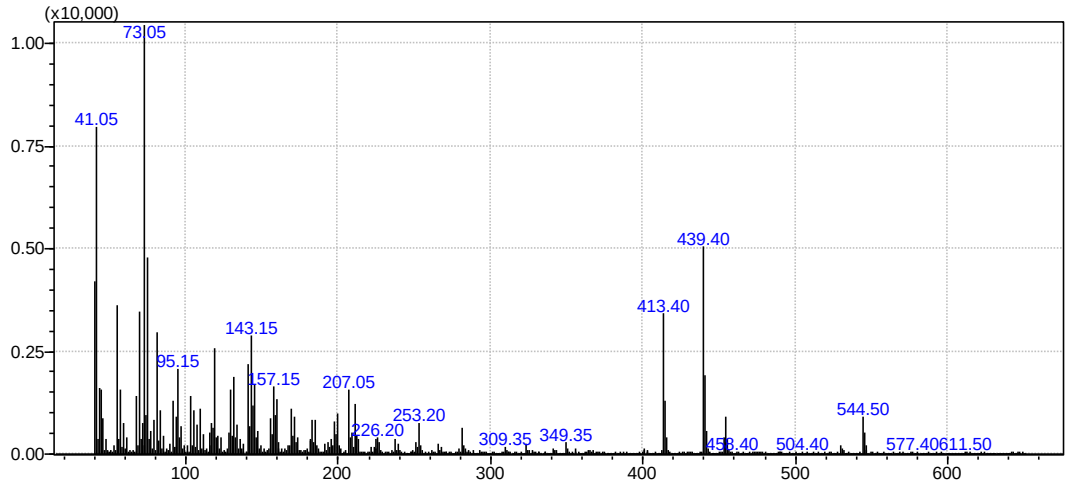
Table S2: X-ray crystallographic data collection, processing, and refinement statistics for CYP125MRCA structures.

	CYP125MRCA	
	Sitosterol	**Cholecalciferol (Vitamin D_3)**
PDB ID	**8VXI**	**8VXG**
Space Group	$C222_1$	$C222_1$
Unit cell lengths a, b, c (Å)	72.96, 118.88, 94.76	72.59, 118.49, 94.82
Unit cell angles α, β ,γ (°)	90, 90 ,90	90, 90 ,90
Wavelength (Å)	0.9537	0.9537
Number of obsevations	350972 (24050)	52492 (24760)
Number of unique reflections	25931 (1950)	39392 (2063)
Resolution (Å)	2.06	1.78
R_{meas}	0.187 (1.721)	0.159 (1.870)
R_{pim}	0.051 (0.482)	0.043 (0.525)
$\langle I/\sigma(I) \rangle$	9.8 (1.3)	10.4 (1.2)
CC(1/2)	0.998 (0.640)	0.998 (0.640)
Completeness (%)	99.8 (97.2)	99.6 (92.5)
Multiplicity	13.5 (12.3)	13.3 (12.0)
R_{work}/R_{free}	0.1665/0.2144	0.1734/0.2183
Bond length R.M.S.D (Å)	0.0138	0.0072
Bond angle R.M.S.D (°)	1.67	1.09
Ramachandran Favoured (%)	96.71	98.19
Ramachandran Allowed (%)	3.04	1.81

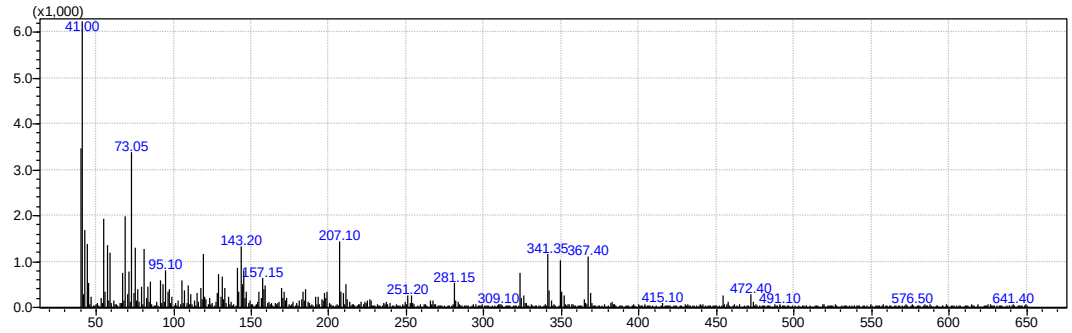
Cholecalciferol oxidation products – Normalised to highest TIC peak



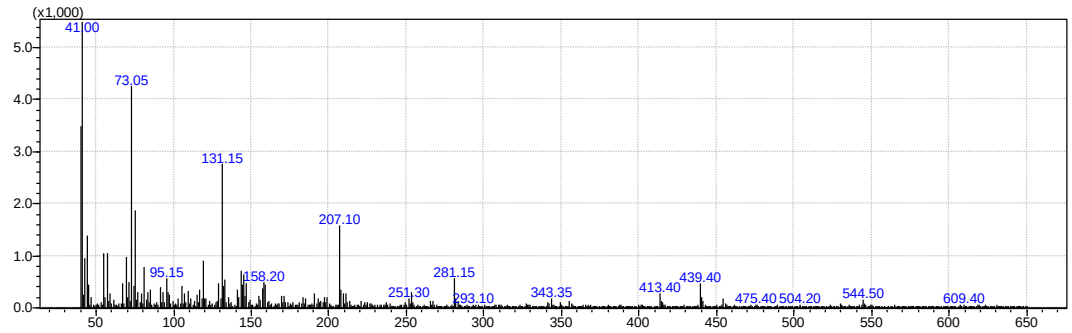
RT: 20.58 – 20.67 min, hydroxylated product (parent ion mass = 544 m/z)



RT: 21.22 – 21.30 min, hydroxylated product (parent ion mass = 544 m/z)

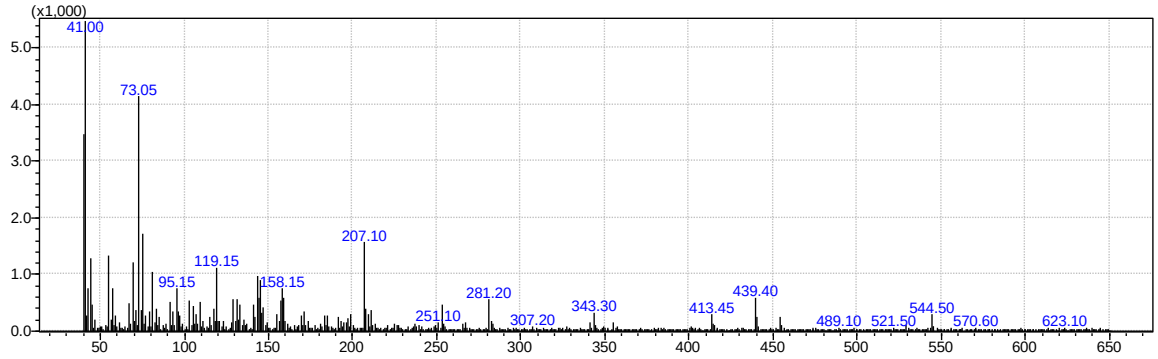


RT: 20.05 – 20.15 min, partially underivatised hydroxylated product (parent ion mass = 472 m/z)



RT: 22.21 – 22.40 min, hydroxylated product (parent mass = 544 m/z)

Figure S17: Mass spectra of GC-MS separated vitamin D_3 derived substrates and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.



RT: 23.08 – 23.22 min, hydroxylated product (parent ion mass = 544 m/z)

Figure S17 Continued: Mass spectra of GC-MS separated vitamin D_3 derived substrates and products of CYP125MRCA-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system.

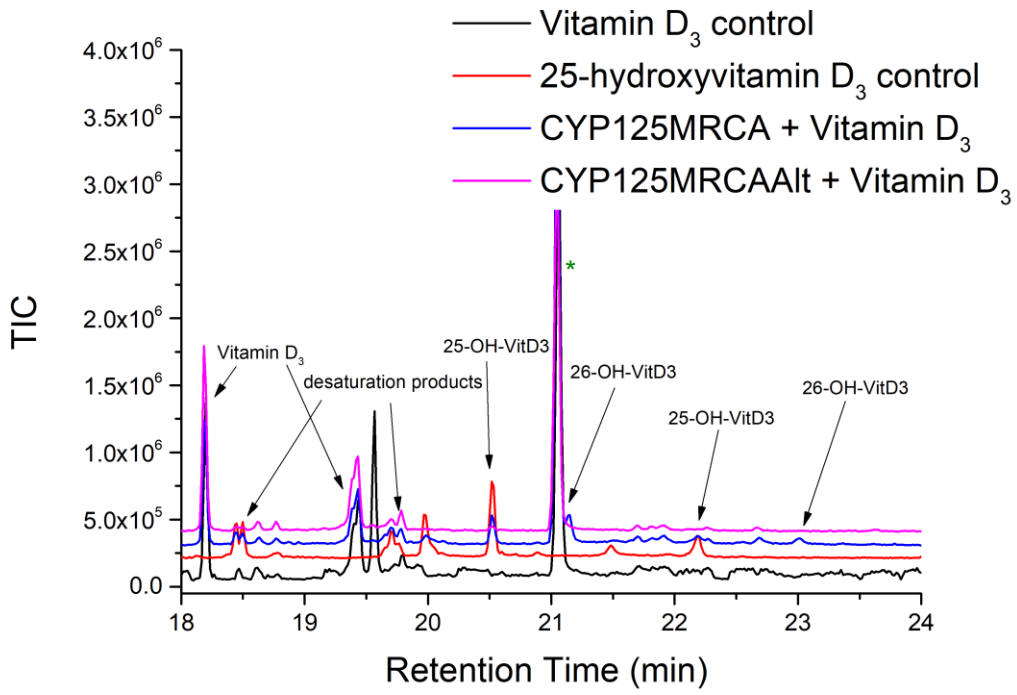


Figure S18: GC-MS chromatograms of CYP125MRCA and CYP125MRCAAlt mediated oxidation of Vitamin D_3. The substrate control is shown in black, a 25-hydroxyvitamin D_3 in **red**, the CYP125MRCA mediated oxidation reaction in **blue** and the CYP125MRCAAlt mediated reaction in **magenta**. Products were identified by their mass fragmentation pattern and relative retention times. * indicates significant unidentified impurity peaks.

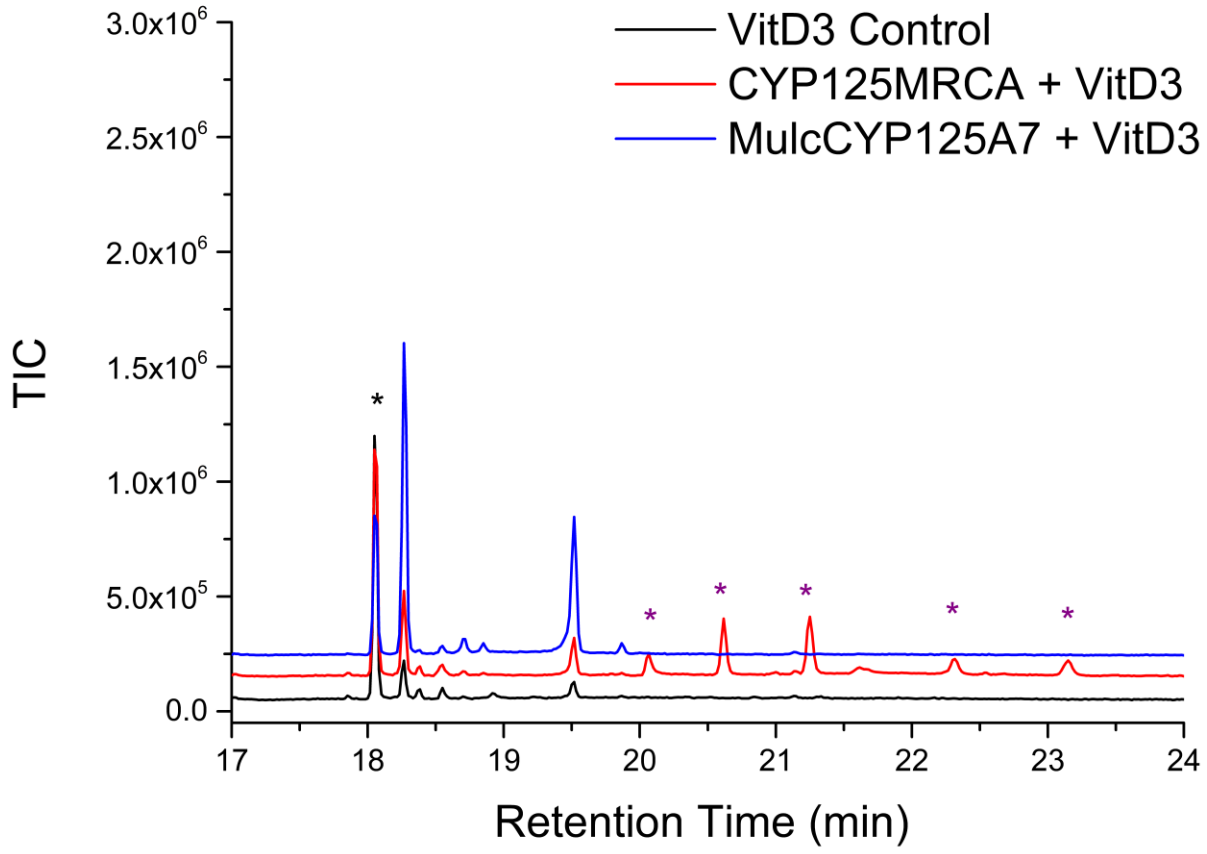


Figure S19: CYP125A7-mediated *in vitro* oxidation reactions using a reconstituted spinach ferredoxin/reductase/NADPH electron transfer system (**blue**), overlaid with a vitamin D_3 control (black) and CYP125MRCA vitamin D_3 positive control (**red**). The vitamin D_3 substrate peak is shown by a *, while the CYP125MRCA mediated oxidation products are highlighted with an*.

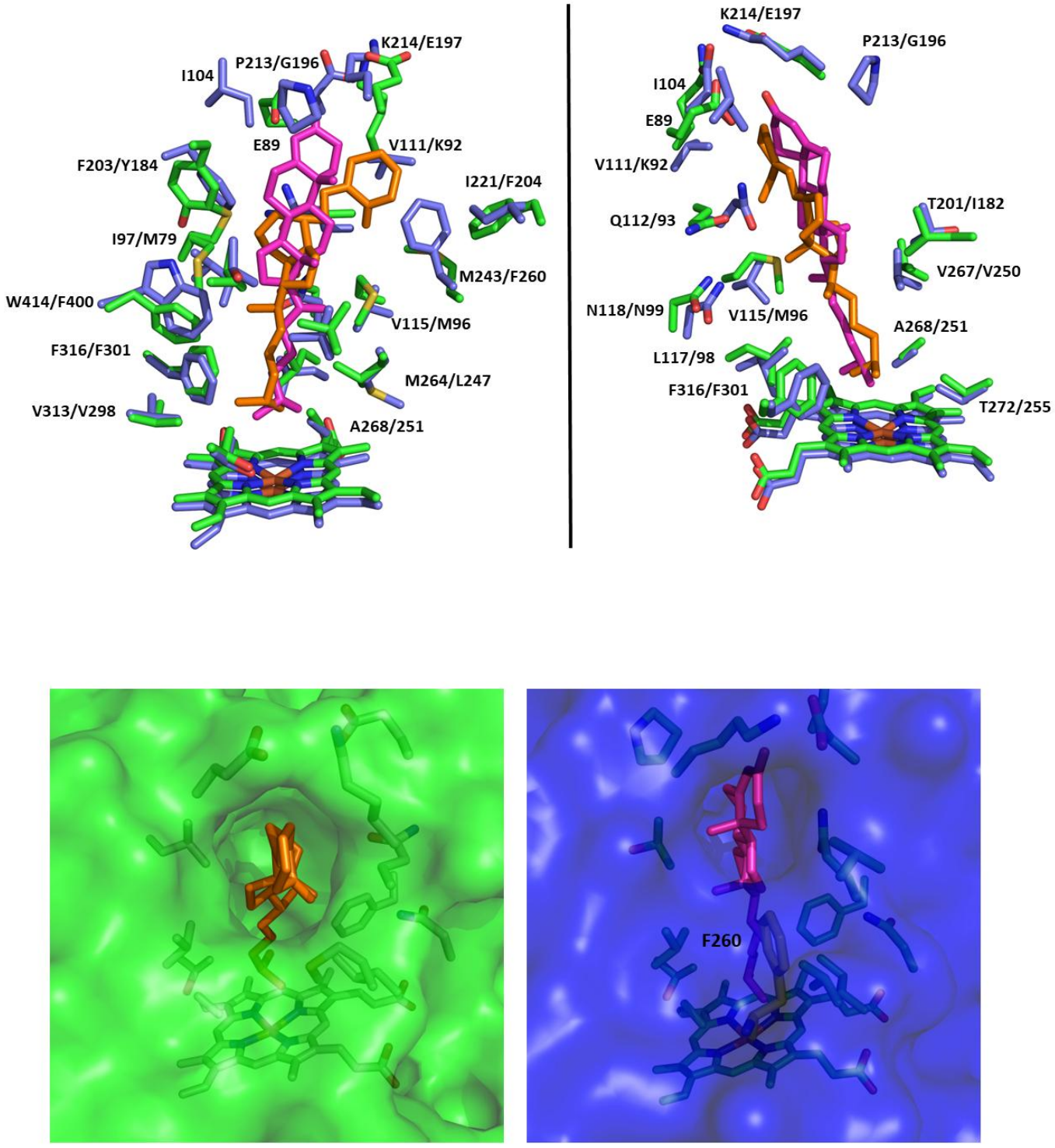


Figure 20: Active-site comparisons between cholecalciferol (orange) bound CYP125MRCA (green) and cholest-4-en-3-one (magenta) bound CYP125A1 (blue). Top: Active-site overlay highlighting difference in the active-site residue positions within 5Å of the heme centre. Bottom: Protein surface differences, highlighting the side-cavity opening of CYP125MRCA allowing for the binding of cholecalciferol (left) and the smaller equivalent cavity in CYP125A1, due to the steric gating by the F260 residue.

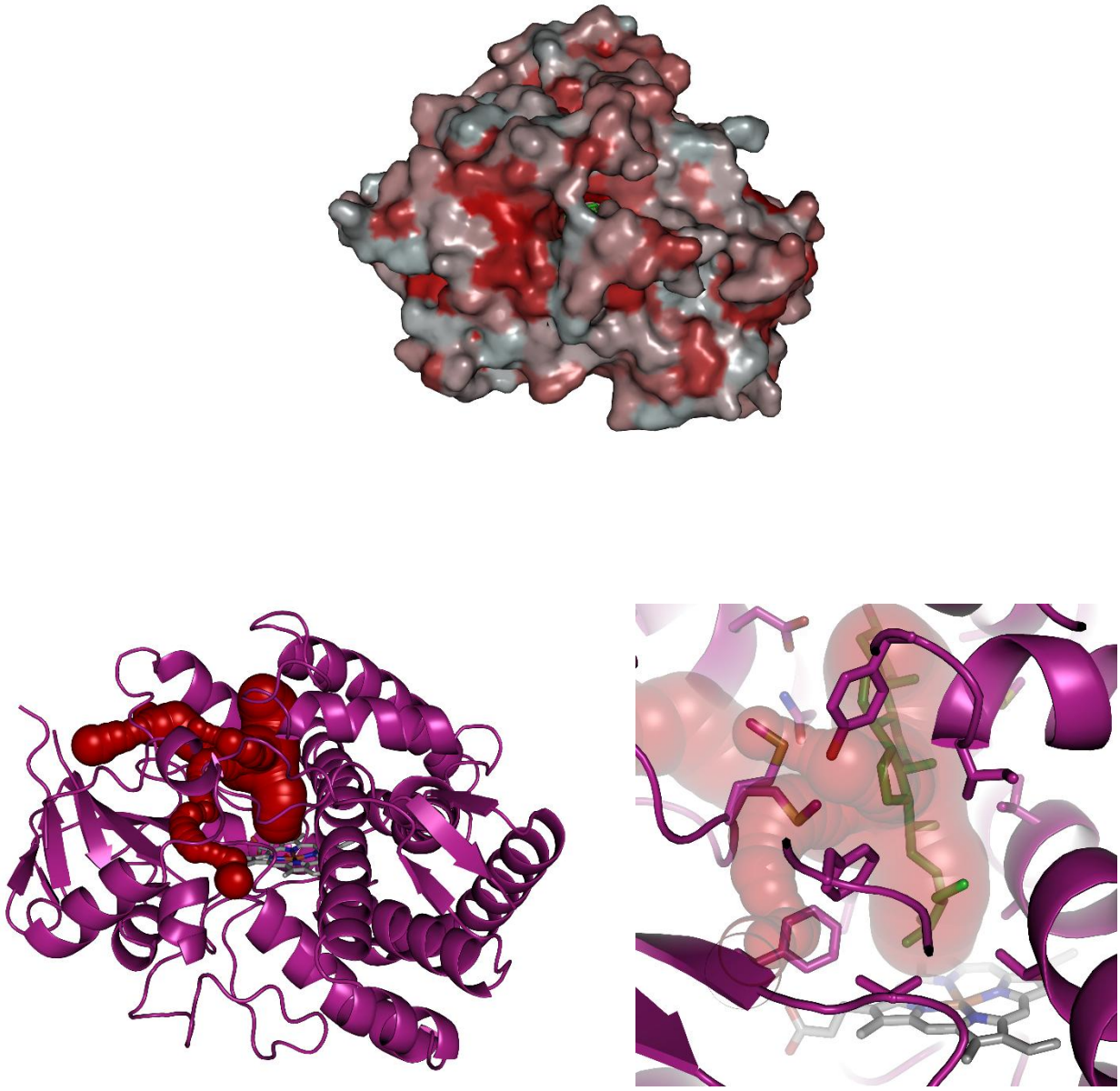


Figure S21: CYP125-MRCA surface hydrophobicity scale, with increasing intensity of red indicating higher hydrophobicity (top). CYP125 MRCA-sitosterol (purple) overall structure with solvent accessible channels produced by Caver shown in red (bottom left). Caver settings were min probe radius = 0.9, shell depth = 4, shell radius = 3, clustering threshold = 3.5, substrate emitted from tunnelling calculations. A zoom in of the position of sitosterol (green) within the solvent access channels of CYP125MRCA (bottom right).

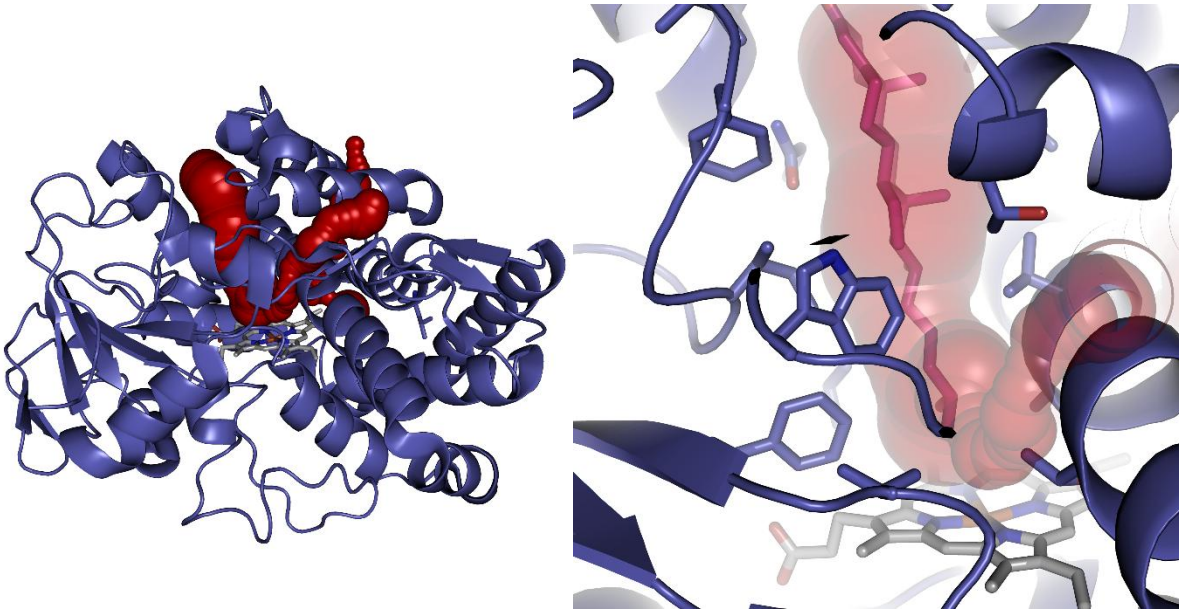


Figure S22: CYP125A1 surface hydrophobicity scale, with increasing intensity of red indicating higher hydrophobicity (top). CYP125 MRCA-cholest-4-en-3-one (blue) overall structure with solvent accessible channels produced by Caver shown in red (left). Caver settings were min probe radius = 0.9, shell depth = 4, shell radius = 3, clustering threshold = 3.5, substrate emitted from tunnelling calculations. A zoom in on the position of cholest-4-en-3-one (magenta) within the solvent access channels of CYP125MRCA (right).

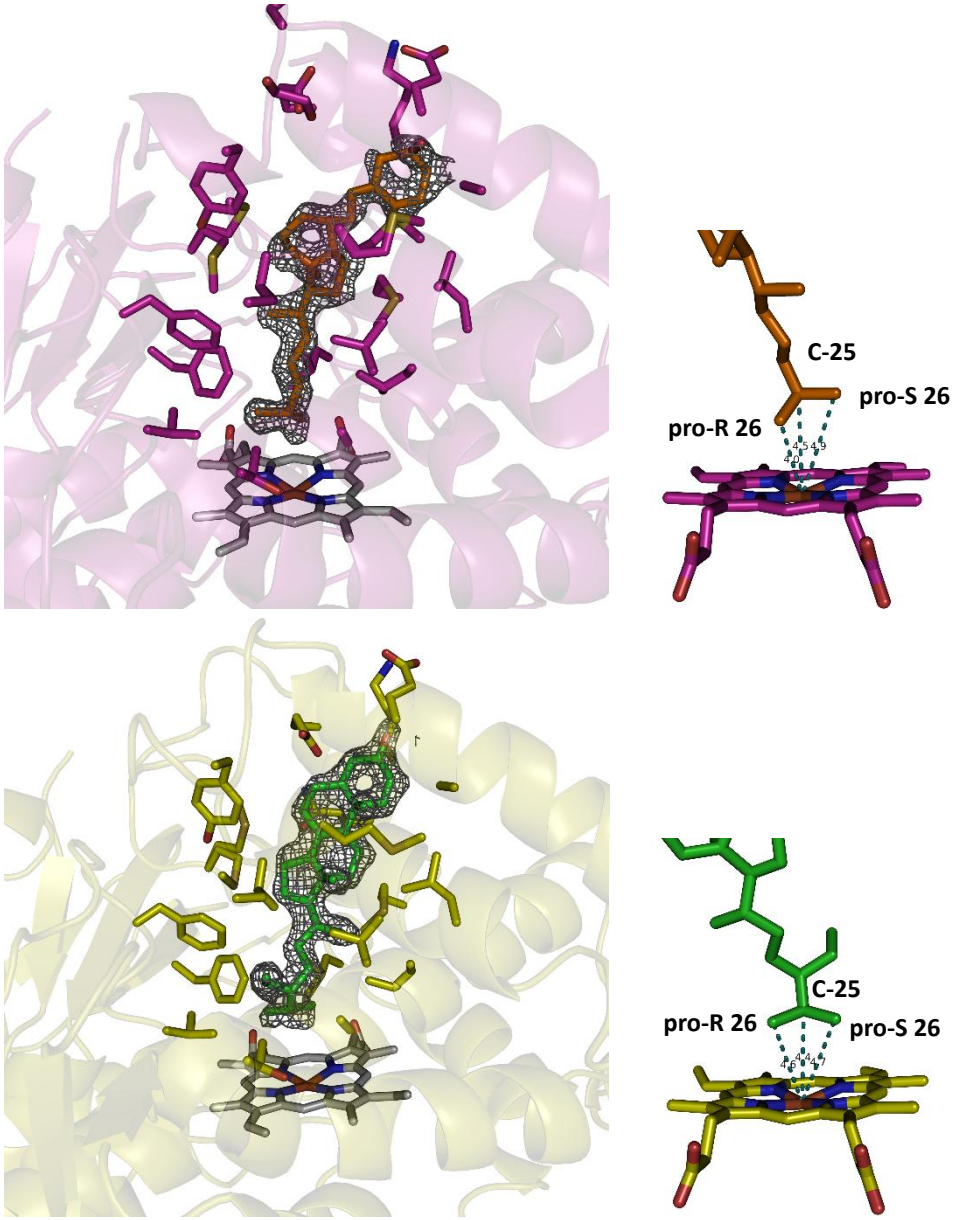


Figure S23: Left: CYP125MRCA-cholecalciferol (top) and CYP125MRCA-sitosterol (right) active-sites. Residues, substrate, and the heme centre are shown as sticks, while the electron density FEM map of the substrate is shown as a grey mesh (carve radius = 1.8 Å).