Supporting information

The substrate specificity in the O-demethylation of 4-alkylguaiacols by Cytochrome P450 $AgcA_{P450}$

Sónia F Santos[1,2], Rajesh Reddy Bommareddy[2], Gary W Black [2], Warispreet Singh*[2], Meilan Huang[1*]

[1] School of Chemistry & Chemical Engineering, Queen's University, Belfast, BT9 5AG, United Kingdom

[2] Hub for Biotechnology in Build Environment, Newcastle upon Tyne, United Kingdom, Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom

Corresponding Authors: m.huang@qub.ac.uk, w.singh@northumbria.ac.uk

Table of Contents

Amino acid sequence of P450 $AgcA_{P450}$

>tr|A0A385L6C3|A0A385L6C3_RHORH Cytochrome P450 OS=Rhodococcus rhodochrous OX=1829 GN=C6369_001540 PE=3 SV=1

MTSTHSFIDEITIEELEADPYPFYERLRKEAPIAYVPALGMYIVSTKELCAEISKDDANW

PAVISAAGGRTFGPQALLNTNGDEHRNLRDMVEPHLQPSAVDKYIDDLVRPFARQRIAEF

ENDGHADIVAAYCEPVSVRALGDLLGLGDVSTEKLREWFHNLSVSFTNAAVDENGEFANP

EGFAPGDRAKAEIIAHVDPKIDKWIVEPDHSAISHWLHDGMPEGQTRSRDVIYPNLYVFL

LGAMQEPGHAMATTLAGLFSRPDQLERVIDDPTLIPRAASEGMRWVAPIWSAAVKRAARE

VTVGGVTLPEGSIVMLSYGSANQDENAYNAPTEYDLDRALVPNMTFGGGKHACAGTYFAN

AVVRIGLEELLEAIPNIERDETHEVDFWGWGFRGPKQLFVKWEV

Sequence Alignment

Sequence alignment of the $AgcA_{P450}$ and $GcoA_{P450}$ enzymes. $GcoA_{P450}$ was used as a template for the construction of the model by homology.

Figure S1 Sequence alignment of the $AgcA_{P450}$ and $GcoA_{P450}$ enzymes using Clustal-Omega webserver developed by the European Bioinformatics Institute. Marked in green we have the active centre conserved residues between these two proteins. Marked in red we have the non-conserved residues, an Isoleucine ($GcoA_{P450}$) replaced by a Leucine ($AgcA_{P450}$) and a Threonine (GcoA) replaced by an Alanine ($AgcA_{P450}$).

RMSD of $AgcA_{P450}$ in complex with each substrate

$AgcA_{P450}$ with guaiacol

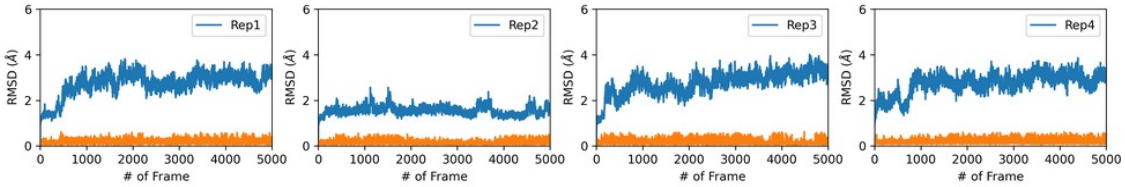

Figure S2 RMSD for the backbone Cα atoms of $AgcA_{P450}$ protein (in blue) and the RMSD of the heavy atoms of substrate Guaiacol (in orange).

$AgcA_{P450}$ with 4-methylguaiacol

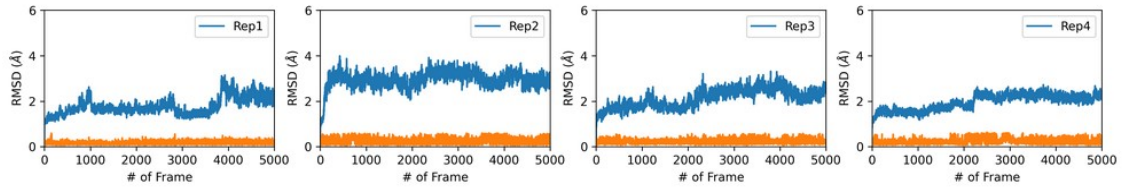

Figure S3 RMSD for the backbone Cα atoms of $AgcA_{P450}$ protein (in blue) and the RMSD of the heavy atoms of substrate 4-methylguaiacol (in orange).

$AgcA_{P450}$with 4-ethylguaiacol

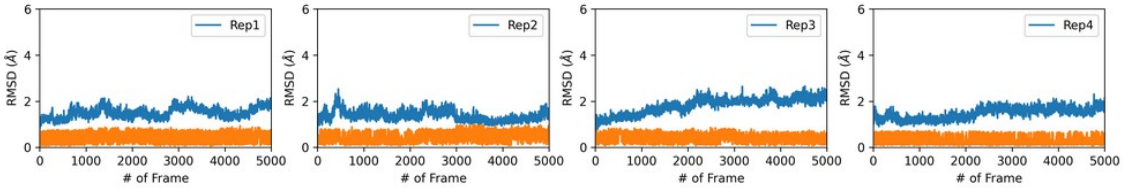

Figure S4 RMSD for the backbone Cα atoms of $AgcA_{P450}$ protein (in blue) and the RMSD of the heavy atoms of substrate 4-ethylguaiacol (in orange).

$AgcA_{P450}$ with 4-propylguaiacol

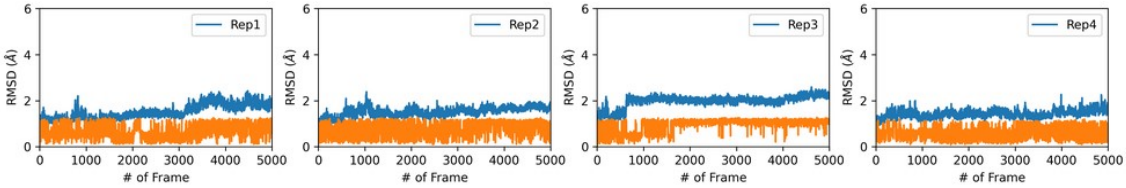

Figure S5 RMSD for the backbone Cα atoms of $AgcA_{P450}$ protein (in blue) and the RMSD of the heavy atoms of substrate 4-propylguiacol (in orange).

Radial distribution functions in relation to the Cpd I complex in the active site

$AgcA_{P450}$with guaiacol

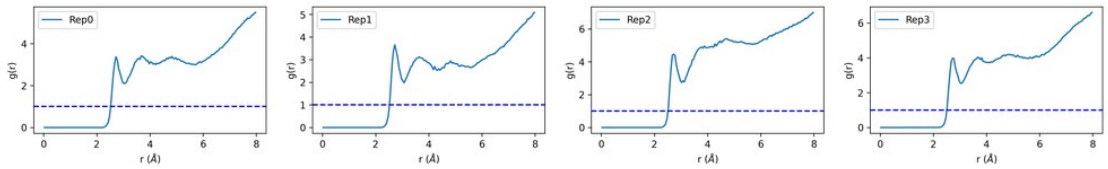

Figure S6 The radial distribution function of $AgcA_{P450}$ protein in complex with guaiacol. The dashed blue line represents at least one water molecule near the Iron-oxo complex of Cpd I during the MD simulations.

$AgcA_{P450}$with 4-methylguaiacol

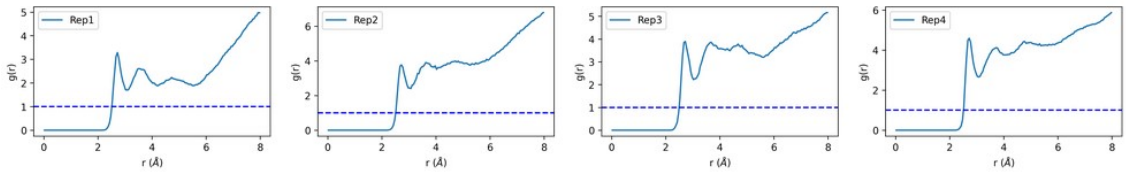

Figure S7 The radial distribution function of $AgcA_{P450}$ protein in complex with 4-methylguaiacol. The dashed blue line represents at least one water molecule near the Iron-oxo complex of Cpd I during the MD simulations.

$AgcA_{P450}$ with 4-ethylguaiacol

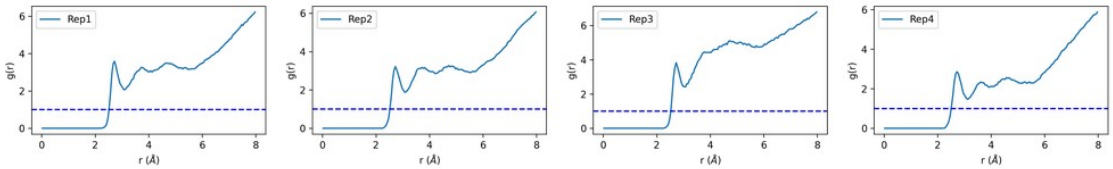

Figure S8 The radial distribution function of $AgcA_{P450}$ protein in complex with 4-ethylguaiacol. The dashed blue line represents at least one water molecule near the Iron-oxo complex of Cpd I during the MD simulations.

$AgcA_{P450}$with 4-propylguaiacol

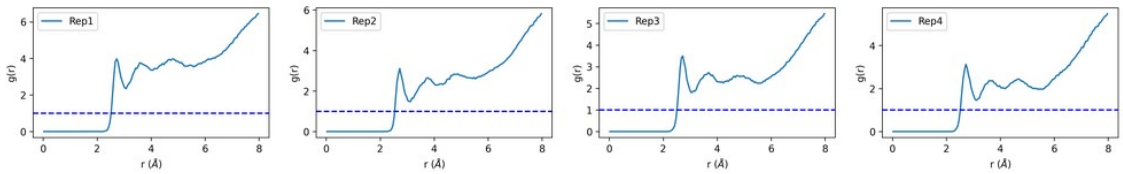

Figure S9 The radial distribution function of $AgcA_{P450}$ protein in complex with 4-propylguaiacol. The dashed blue line represents at least one water molecule near the Iron-oxo complex of Cpd I during the MD simulations.

Distance of the residues A169 and I13 acting as gate of the substrate access channel

$AgcA_{P450}$with guaiacol

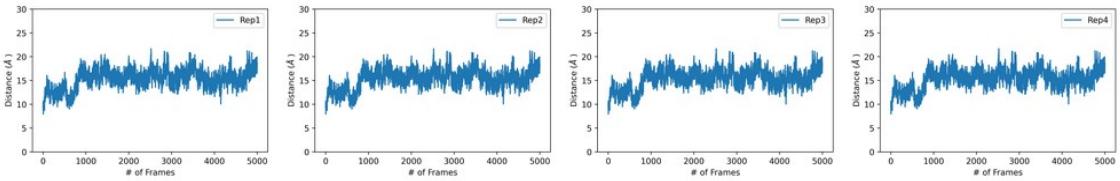

Figure S10 Distance between I13 and A169 during the MD simulation of $AgcA_{P450}$ protein in complex with Guaiacol.

$AgcA_{P450}$ with 4-methylguaiacol

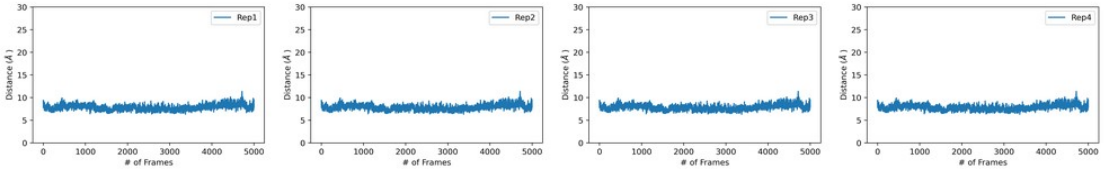

Figure S11 Distance between I13 and A169 during the MD simulation of $AgcA_{P450}$ protein in complex with 4-methylguaiacol.

$AgcA_{P450}$ with 4-ethylguaiacol

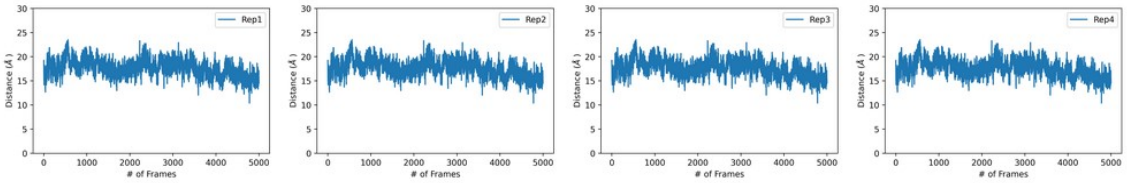

Figure S12 Distance between I13 and A169 during the MD simulation of $AgcA_{P450}$ protein in complex with 4-ethylguaiacol.

$AgcA_{EP4}$ with 4-propylguaiacol

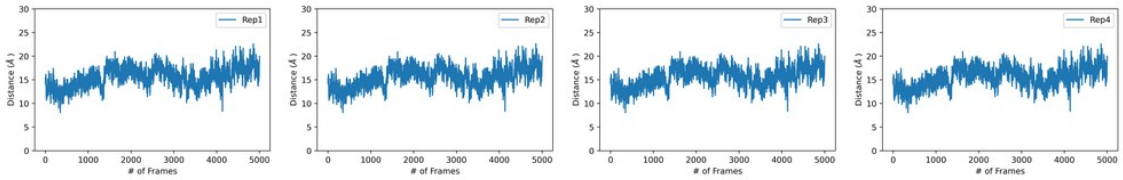

Figure S13 Distance between I13 and A169 during the MD simulation of $AgcA_{P450}$ protein in complex with 4-propylguaiacol.

Distance of the methyl group and the Iron-oxo complex of each complex

$AgcA_{P450}$ with guaiacol

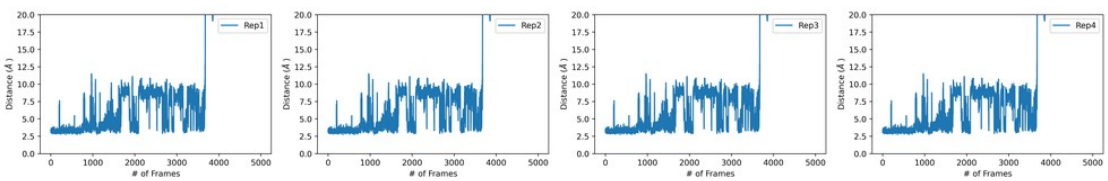

Figure S14 Distance between the methoxy group of guaiacol and the oxygen atom of Fe(IV)=O.

AgcA$_{P450}$ with 4-methylguaiacol

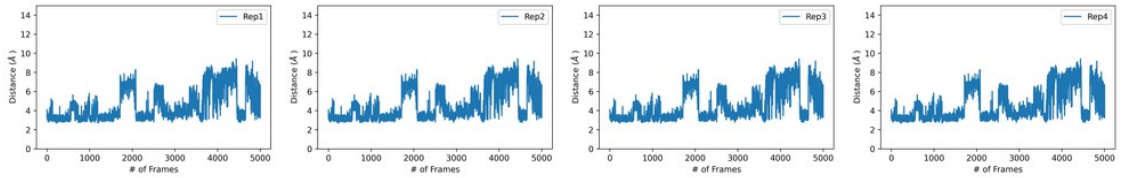

Figure S15 Distance between the methoxy group of 4-methylguaiacol and the oxygen atom of Fe(IV)=O.

AgcA$_{P450}$ with 4-ethylguaiacol

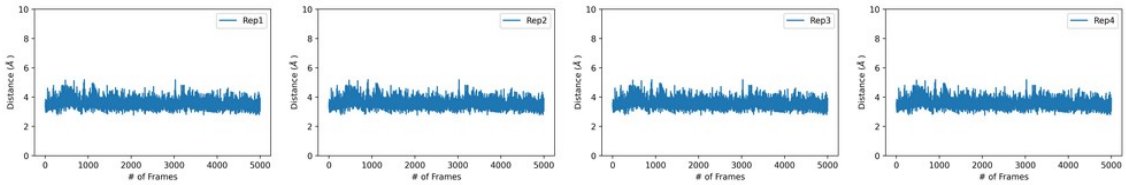

Figure S16 Distance between the methoxy group of 4-ethylguaiacol and the oxygen atom of Fe(IV)=O.

AgcA$_{P450}$ with 4-propylguaiacol

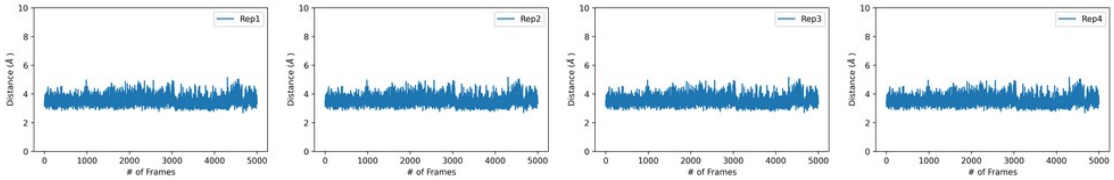

Figure S17 Distance between the methoxy group of 4-propylguaiacol and the oxygen atom of Fe(IV)=O.

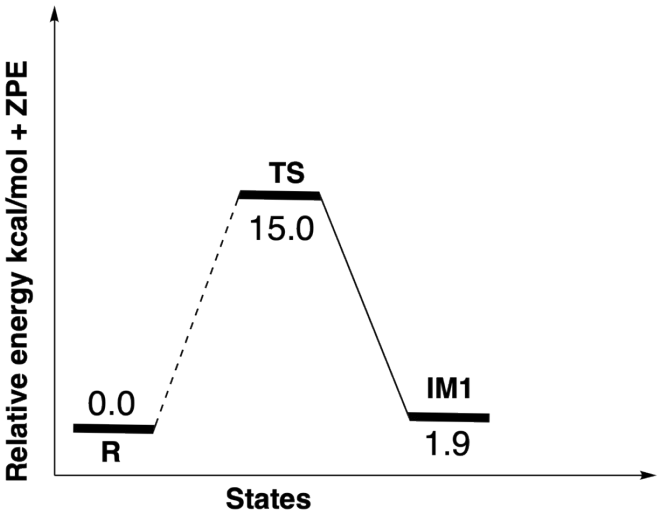

Figure S18 The oxidation 4-propylguaiacol in the active site of AgcA$_{P450}$ enzyme calculated using QM/MM methods at M=4 spin of iron. The reaction profile of the hydrogen atom transfer (HAT) catalysed by Cpd I was computed using the UB3LYP functional with D3BJ dispersion correction and the def2-TZVP basis set for M=4.

Table S1 QM/MM calculated spin densities of the oxidation 4-propylguaiacol in the active site of $AgcA_{P450}$ at M=2

	RC	TS	IM1	IM2
Fe(IV)=O	2.11	1.8	2.11	1.12
Porphyrin+Cysteine	-1.1	-0.36	-0.16	0.0
4-propylguaiacol	0.0	-0.44	-0.96	0.0

Table S2 QM/MM calculated spin densities of the oxidation 4-propylguaiacol in the active site of $AgcA_{P450}$ at M=4

	RC	TS	IM1
Fe(IV)=O	2.04	2.11	2.14
Porphyrin+Cysteine	0.93	0.32	-0.12
4-propylguaiacol	0.0	0.55	0.97