SUPPLEMENTARY FIGURES

Heterologous expression and characterization of plant Taxadiene-5α-Hydroxylase (CYP725A4) in Escherichia coli

*John Edward Rouck[1], Bradley Walters Biggs,[2,3] , Amogh Kambalyal[1], William Arnold[1], Marjan De Mey[4], *Parayil Kumaran Ajikumar[2] and *Aditi Das[1]*

[1]Department of Comparative Biosciences, Department of Biochemistry, Department of Bioengineering, Division of Nutritional Science, Center for Biophysics and Quantitative Biology, Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, Illinois 61801

[2]Manus Biosynthesis, 1030 Massachusetts Avenue, Suite 300, Cambridge, MA 02138

[3]Department of Chemical and Biological Engineering (Masters in Biotechnology Program), Northwestern University, Evanston, IL 60208

[4]Centre for Industrial Biotechnology and Biocatalysis, Ghent University, Coupure Links 653, B-9000, Belgium

SUPPLEMENTAL TABLE OF CONTENTS

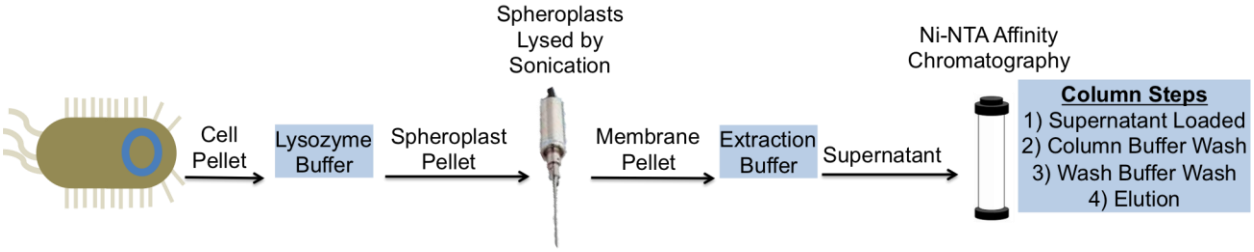


Figure S-1: *Schematic of the purification of His-tagged CYP725A4-CPR chimeras.* The cell pellet is dissolved in lysozyme buffer followed by centrifuging to isolate the spheroplasts. The spheroplasts are further sonicated to obtain the membrane pellet from which the proteins are extracted and applied to the Ni-NTA affinity column (Black arrows represent centrifugation steps).

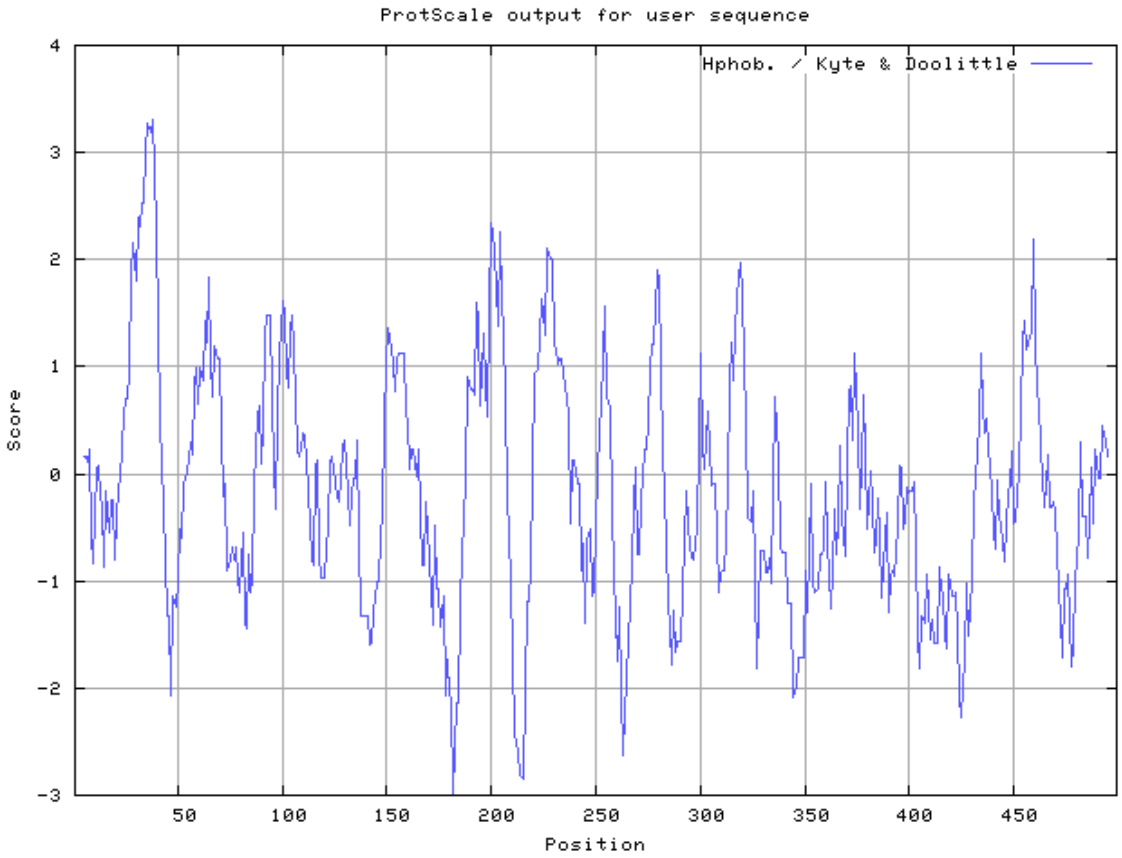


Figure S2. *Hydrophobicity Plot of Wild Type Taxus cuspidata CYP725A4.*
Hydrophobicity plot of wild-type *Taxus cuspidata* CYP725A4.

Figure S-3. Construct Definition Table

Construct Abbreviation	**Construct Definition**	**Figure Containing Sequence**
CYP725A4-17α	CYP725A4 with 17α N-terminal modification	S-4
CYP725A4-MA	CYP725A4 with MA N-terminal modification	S-5
CYP725A4-2b1	CYP725A4 with 17α N-terminal modification	S-6
CYP725A4-17$^{\alpha}$-L-CPR	CYP725A4 with 17α N-terminal modification fused to Taxus Cuspidata CPR with a similar 17 α modification.	S-7
CYP725A4-MA-L-CPR	CYP725A4 with 17 α N-terminal modification fused to Taxus Cuspidata CPR with a similar 17 α modification.	S-8
CYP725A4-2b1-L-CPR	CYP725A4 with 17α N-terminal modification fused to Taxus Cuspidata CPR with a similar 17 α modification.	S-9

MALLLAVFFSIALSAIAGILLLLLLFRSKRHSSLKLPPGKLGIPFIGESFIFLRA
LRSNSLEQFFDERVKKFGLVFKTSLIGHPTVVLCGPAGNRLILSNEEKLVQ
MSWPAQFMKLMGENSVATRRGEDHIVMRSALAGFFGPGALQSYIGKMNT
EIQSHINEKWKGKDEVNVLPLVRELVFNISAILFFNIYDKQEQDRLHKLLET
ILVGSFALPIDLPGFGFHRALQGRAKLNKIMLSLIKKRKEDLQSGSATATQD
LLSVLLTFRDDKGTPLTNDEILDNFSSLLHASYDTTTSPMALIFKLLSSNPEC
YQKVVQEQLEILSNKEEGEEITWKDLKAMKYTWQVAQETLRMFPPVFGTF
RKAITDIQYDGYTIPKGWKLLWTTYSTHPKDLYFNEPEKFMPSRFDQEGK
HVAPYTFLPFGGGQRSCVGWEFSKMEILLFVHHFVKTFSSYTPVDPDEKIS
GDPLPPLPSKGFSIKLFPRP

Figure S-4. *CYP725A4-17α Sequence.* CYP725A4 17 α Construct Sequence: The yellow highlighted region is the 17 α modification to CYP725A4. The region that is not highlighted is unmodified *Taxus cuspidata* CYP725A4. A 6-residue histidine tag is added to the C-Terminus.

MAPFIGESFIFLRALRSNSLEQFFDERVKKFGLVFKTSLIGHPTVVLCGPAG
NRLILSNEEKLVQMSWPAQFMKLMGENSVATRRGEDHIVMRSALAGFFGP
GALQSYIGKMNTEIQSHINEKWKGKDEVNVLPLVRELVFNISAILFFNIYDK
QEQDRLHKLLETILVGSFALPIDLPGFGFHRALQGRAKLNKIMLSLIKKRKE
DLQSGSATATQDLLSVLLTFRDDKGTPLTNDEILDNFSSLLHASYDTTTSP
MALIFKLLSSNPECYQKVVQEQLEILSNKEEGEEITWKDLKAMKYTWQVA
QETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLLWTTYSTHPKDLYFNEP
EKFMPSRFDQEGKHVAPYTFLPFGGGQRSCVGWEFSKMEILLFVHHFVKT
FSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRP

Figure S-5: *CYP725A4-MA Sequence.* CYP725A4 MA α Construct Sequence: The yellow highlighted region is the MA modification to CYP725A4. The region that is not highlighted is unmodified *Taxus cuspidata* CYP725A4. A 6-residue histidine tag is added to the C-Terminus.

MAKKTSSKGKLPPGPSPFIGESFIFLRALRSNSLEQFFDERVKKFGLVFKTSL
IGHPTVVLCGPAGNRLILSNEEKLVQMSWPAQFMKLMGENSVATRRGED
HIVMRSALAGFFGPGALQSYIGKMNTEIQSHINEKWKGKDEVNVLPLVRE
LVFNISAILFFNIYDKQEQDRLHKLLETILVGSFALPIDLPGFGFHRALQGRA
KLNKIMLSLIKKRKEDLQSGSATATQDLLSVLLTFRDDKGTPLTNDEILDNF
SSLLHASYDTTTSPMALIFKLLSSNPECYQKVVQEQLEILSNKEEGEEITWK
DLKAMKYTWQVAQETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLLWT
TYSTHPKDLYFNEPEKFMPSRFDQEGKHVAPYTFLPFGGGQRSCVGWEFS
KMEILLFVHHFVKTFSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRP

Figure S-6. *CYP25A4-2b1 Sequence.* CYP725A4 2b1 Construct Sequence: The yellow highlighted region is the 2b1 modification to CYP725A4. The region that is not highlighted is unmodified *Taxus cuspidata* CYP725A4. A 6-residue histidine tag is added to the C-Terminus.

MALLLAVFFSIALSAIAGILLLLLLFRSKRHSSLKLPPGKLGIPFIGESFIFLRA
LRSNSLEQFFDERVKKFGLVFKTSLIGHPTVVLCGPAGNRLILSNEEKLVQ
MSWPAQFMKLMGENSVATRRGEDHIVMRSALAGFFGPGALQSYIGKMNT
EIQSHINEKWKGKDEVNVLPLVRELVFNISAILFFNIYDKQEQDRLHKLLET
ILVGSFALPIDLPGFGFHRALQGRAKLNKIMLSLIKKRKEDLQSGSATATQD
LLSVLLTFRDDKGTPLTNDEILDNFSSLLHASYDTTTSPMALIFKLLSSNPEC
YQKVVQEQLEILSNKEEGEEITWKDLKAMKYTWQVAQETLRMFPPVFGTF
RKAITDIQYDGYTIPKGWKLLWTTYSTHPKDLYFNEPEKFMPSRFDQEGK
HVAPYTFLPFGGGQRSCVGWEFSKMEILLFVHHFVKTFSSYTPVDPDEKIS
GDPLPPLPSKGFSIKLFPRPGSTGSMALLLAVFRRGGSDTQKPAVRPTPLVK
EEDEEEEDDSAKKKVTIFFGTQTGTAEGFAKALAEEAKARYEKAVFKVVD
LDNYAADDEQYEEKLKKEKLAFFMLATYGDGEPTDNAARFYKWFLEGKE
REPWLSDLTYGVFGLGNRQYEHFNKVAKAVDEVLIEQGAKRLVPVGLGD
DDQCIEDDFTAWREQVWPELDQLLRDEDDEPTSATPYTAAIPEYRVEIYDS
VVSVYEETHALKQNGQAVYDIHHPCRSNVAVRRELHTPLSDRSCIHLEFDI
SDTGLIYETGDHVGVHTENSIETVEEAAKLLGYQLDTIFSVHGDKEDGTPL
GGSSLPPPFPGPCTLRTALARYADLLNPPRKAAFLALAAHASDPAEAERLK
FLSSPAGKDEYSQWVTASQRSLLEIMAEFPSAKPPLGVFFAAIAPRLQPRYY
SISSSPRFAPSRIHVTCALVYGPSPTGRIHKGVCSNWMKNSLPSEETHDCSW
APVFVRQSNFKLPADSTTPIVMVGPGTGFAPFRGFLQERAKLQEAGEKLGP
AVLFFGCRNRQMDYIYEDELKGYVEKGILTNLIVAFSREGATKEYVQHKM
LEKASDTWSLIAQGGYLYVCGDAKGMARDVHRTLHTIVQEQESVDSSKA
EFLVKKLQMDGRYLRDIW

Figure S-7. *CYP725A4-17-α-L-CPR Sequence.* The yellow highlighted region is the 17 α modification to CYP725A4. The region that is not highlighted is unmodified *Taxus Cuspidata* CYP725A4. The pink highlighted region is the linker portion of the chimera peptide. The cyan highlighted region is the 17 α modification to the *Taxus Cuspidata* CPR. The green highlighted region is unmodified *Taxus Cuspidata* CPR. A 6-residue histidine tag is added to the C-Terminus.

MAPFIGESFIFLRALRSNSLEQFFDERVKKFGLVFKTSLIGHPTVVLCGPAG
NRLILSNEEKLVQMSWPAQFMKLMGENSVATRRGEDHIVMRSALAGFFGP
GALQSYIGKMNTEIQSHINEKWKGKDEVNVLPLVRELVFNISAILFFNIYDK
QEQDRLHKLLETILVGSFALPIDLPGFGFHRALQGRAKLNKIMLSLIKKRKE
DLQSGSATATQDLLSVLLTFRDDKGTPLTNDEILDNFSSLLHASYDTTTSP
MALIFKLLSSNPECYQKVVQEQLEILSNKEEGEEITWKDLKAMKYTWQVA
QETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLLWTTYSTHPKDLYFNEP
EKFMPSRFDQEGKHVAPYTFLPFGGGQRSCVGWEFSKMEILLFVHHFVKT
FSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRPGSTGSMARRGGSDTQKPAV
RPTPLVKEEDEEEEDDSAKKKVTIFFGTQTGTAEGFAKALAEEAKARYEK
AVFKVVDLDNYAADDEQYEEKLKKEKLAFFMLATYGDGEPTDNAARFY
KWFLEGKEREPWLSDLTYGVFGLGNRQYEHFNKVAKAVDEVLIEQGAKR
LVPVGLGDDDQCIEDDFTAWREQVWPELDQLLRDEDDEPTSATPYTAAIP
EYRVEIYDSVVSVYEETHALKQNGQAVYDIHHPCRSNVAVRRELHTPLSD
RSCIHLEFDISDTGLIYETGDHVGVHTENSIETVEEAAKLLGYQLDTIFSVH
GDKEDGTPLGGSSLPPPFPGPCTLRTALARYADLLNPPRKAAFLALAAHAS
DPAEAERLKFLSSPAGKDEYSQWVTASQRSLLEIMAEFPSAKPPLGVFFAAI
APRLQPRYYSISSSPRFAPSRIHVTCALVYGPSPTGRIHKGVCSNWMKNSLP
SEETHDCSWAPVFVRQSNFKLPADSTTPIVMVGPGTGFAPFRGFLQERAKL
QEAGEKLGPAVLFFGCRNRQMDYIYEDELKGYVEKGILTNLIVAFSREGAT
KEYVQHKMLEKASDTWSLIAQGGYLYVCGDAKGMARDVHRTLHTIVQE
QESVDSSKAEFLVKKLQMDGRYLRDIW

Figure S-8. *CYP725A4-MA-L-CPR.* The yellow highlighted region is the MA modification to CYP725A4. The region that is not highlighted is unmodified *Taxus Cuspidata* CYP725A4. The pink highlighted region is the linker portion of the chimera peptide. The cyan highlighted region is the MA modification to the *Taxus cuspidata* CPR. The green highlighted region is unmodified *Taxus Cuspidata* CPR. A 6-residue histidine tag is added to the C-Terminus.

MAKKTSSKGKLPPGPSPFIGESFIFLRALRSNSLEQFFDERVKKFGLVFKTSL
IGHPTVVLCGPAGNRLILSNEEKLVQMSWPAQFMKLMGENSVATRRGED
HIVMRSALAGFFGPGALQSYIGKMNTEIQSHINEKWKGKDEVNVLPLVRE
LVFNISAILFFNIYDKQEQDRLHKLLETILVGSFALPIDLPGFGFHRALQGRA
KLNKIMLSLIKKRKEDLQSGSATATQDLLSVLLTFRDDKGTPLTNDEILDNF
SSLLHASYDTTTSPMALIFKLLSSNPECYQKVVQEQLEILSNKEEGEEITWK
DLKAMKYTWQVAQETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLLWT
TYSTHPKDLYFNEPEKFMPSRFDQEGKHVAPYTFLPFGGGQRSCVGWEFS
KMEILLFVHHFVKTFSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRPGSTGS
MAKKTSSKGKLPPGPSPLVKEEDEEEEDDSAKKKVTIFFGTQTGTAEGFAK
ALAEEAKARYEKAVFKVVDLDNYAADDEQYEEKLKKEKLAFFMLATYG
DGEPTDNAARFYKWFLEGKEREPWLSDLTYGVFGLGNRQYEHFNKVAKA
VDEVLIEQGAKRLVPVGLGDDDQCIEDDFTAWREQVWPELDQLLRDEDD
EPTSATPYTAAIPEYRVEIYDSVVSVYEETHALKQNGQAVYDIHHPCRSNV
AVRRELHTPLSDRSCIHLEFDISDTGLIYETGDHVGVHTENSIETVEEAAKL
LGYQLDTIFSVHGDKEDGTPLGGSSLPPPFPGPCTLRTALARYADLLNPPRK
AAFLALAAHASDPAEAERLKFLSSPAGKDEYSQWVTASQRSLLEIMAEFPS
AKPPLGVFFAAIAPRLQPRYYSISSSPRFAPSRIHVTCALVYGPSPTGRIHKG
VCSNWMKNSLPSEETHDCSWAPVFVRQSNFKLPADSTTPIVMVGPGTGFA
PFRGFLQERAKLQEAGEKLGPAVLFFGCRNRQMDYIYEDELKGYVEKGIL
TNLIVAFSREGATKEYVQHKMLEKASDTWSLIAQGGYLYVCGDAKGMAR
DVHRTLHTIVQEQESVDSSKAEFLVKKLQMDGRYLRDIW

Figure S-9. *CYP725A4 2b1 L CPR.* The yellow highlighted region is the 2b1 modification to CYP725A4. The region that is not highlighted is unmodified *Taxus Cuspidata* CYP725A4. The pink highlighted region is the linker portion of the chimera peptide. The cyan highlighted region is the 2b1 modification to the *Taxus cuspidata* CPR. The green highlighted region is unmodified *Taxus Cuspidata* CPR. A 6-residue histidine tag is added to the C-Terminus.

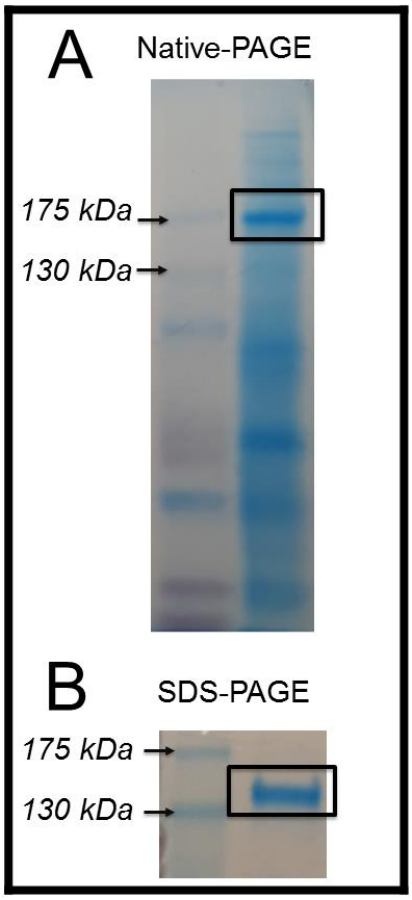
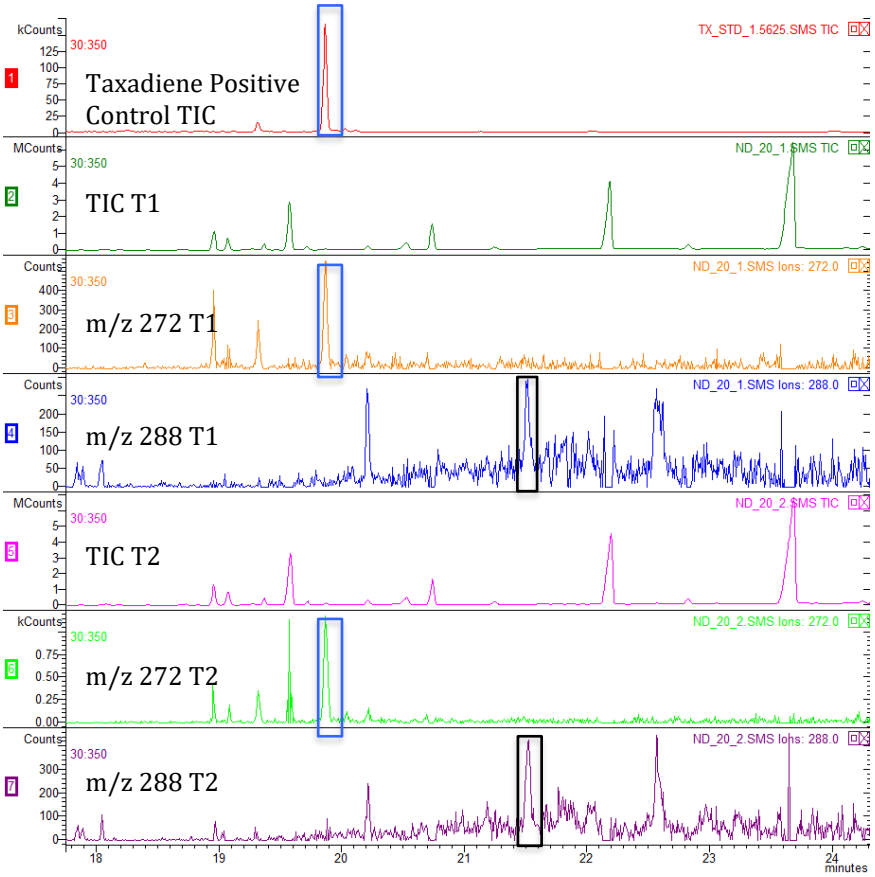
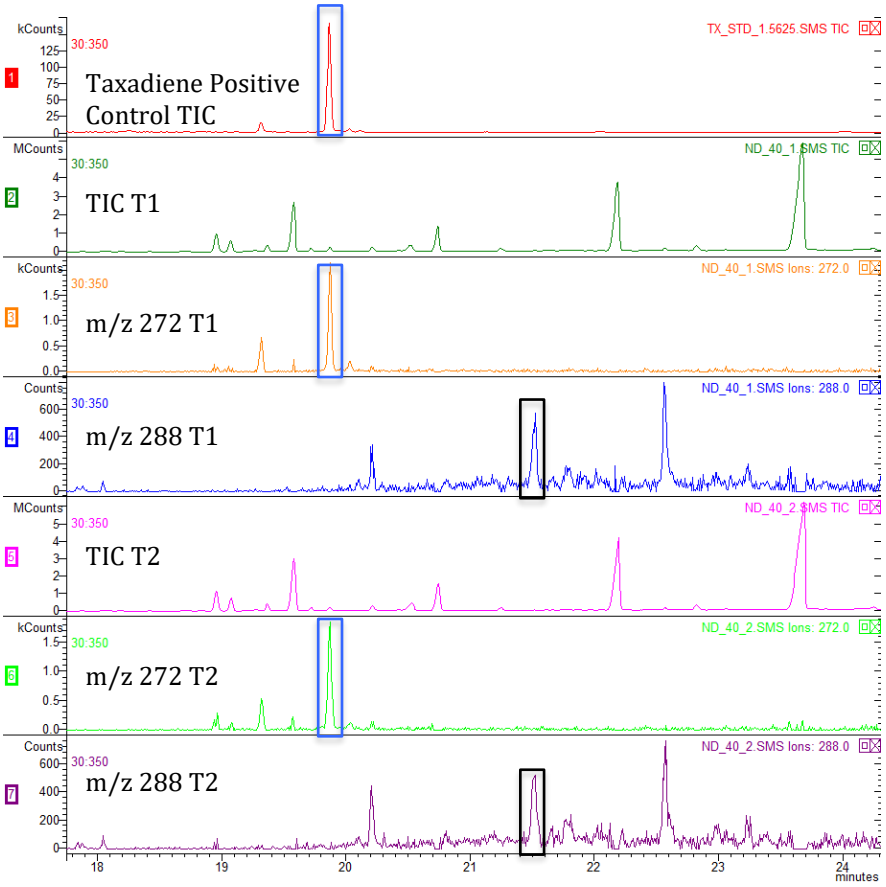


Figure S-10. *Native- and SDS-PAGE of CYP725A4-17α-L-CPR.* Samples were run under (A) native and (B) SDS-denaturing conditions (without boiling) as described in the Materials and Method. Bands corresponding to CYP725A4-17α-L-CPR are boxed for emphasis. Molecular mass of construct is calculated to be 128 kDa and the molecular mass obtained by the gel is ~104% of this value.



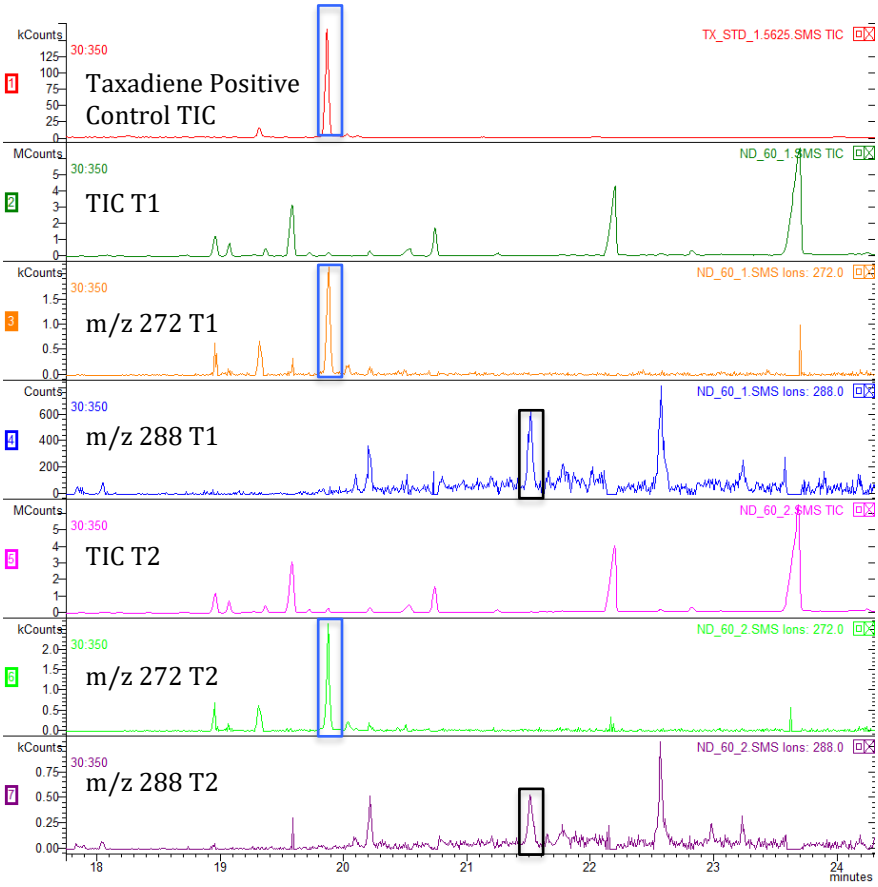
OCT Values
1 – 620 AU
2 – 870 AU

Figure S-11. *GC of 20 μM Taxadiene Incubation for Kinetics Evaluation***.** The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. Peak area was converted to grams by using a taxadiene standard curve.



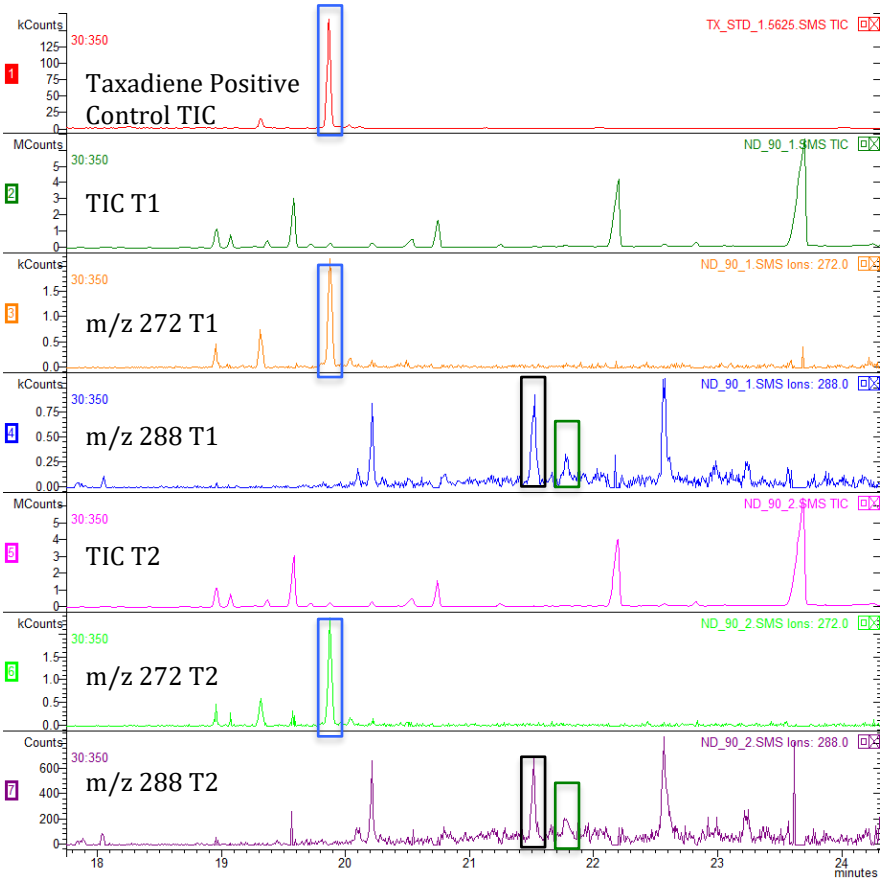
OCT Values
1 – 1331 AU
2 – 1403 AU

Figure S-12. *GC of 40 μM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. Peak area was converted to grams by using a taxadiene standard curve.



OCT Values
1 – 1380 AU
2 – 1230 AU

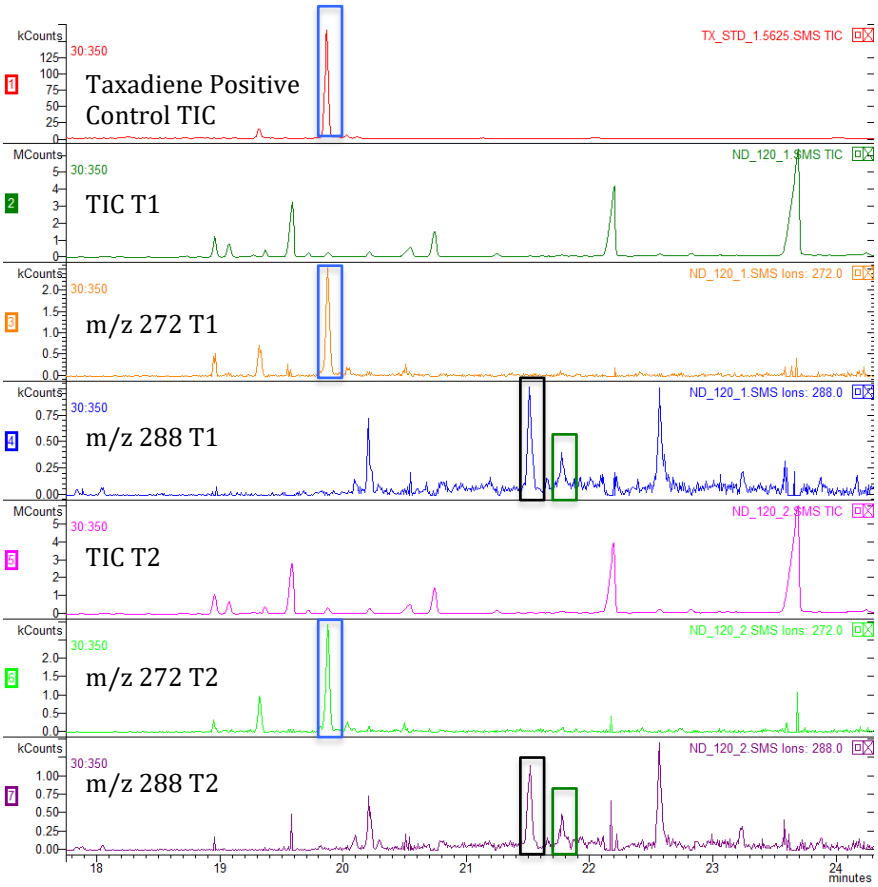
Figure S-13. *GC of 60 μM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. Peak area was converted to grams by using a taxadiene standard curve.



OCT Values
1 – 2137 AU
2 – 1475 AU

T-5α-ol Values
1 – 602 AU
2 – 442 AU

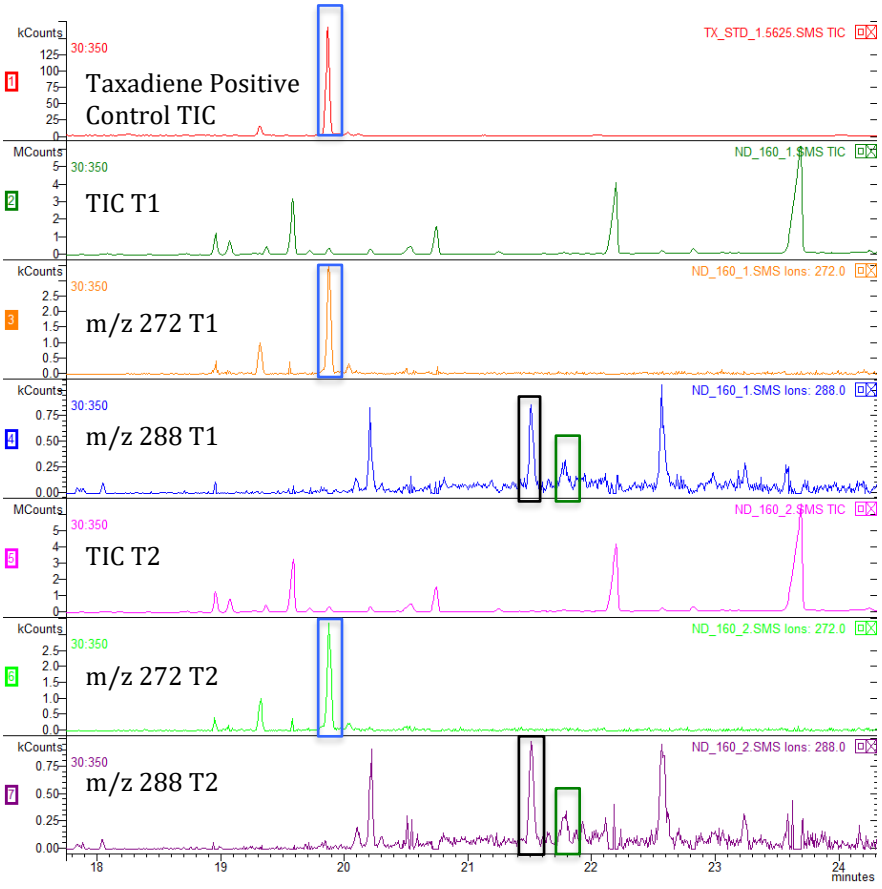
Figure S-14. *GC of 90 µM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. At this taxadiene concentration the production of Taxadiene-5α-ol was also observed (green box at m/z 288 21.85 minutes). Peak area was converted to grams by using a taxadiene standard curve.



OCT Values
1 – 2080 AU
2 – 2472 AU

T-5α-ol Values
1 – 606 AU
2 – 1006 AU

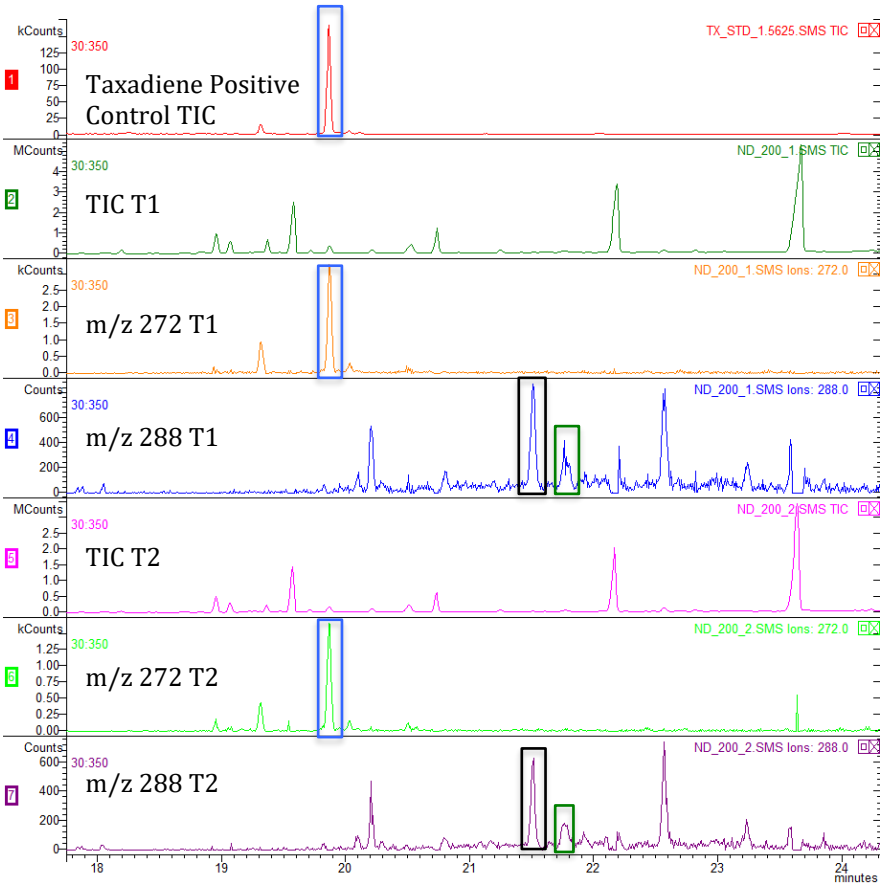
Figure S-15. *GC of 120 µM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. At this taxadiene concentration the production of Taxadiene-5α-ol was also observed (green box at m/z 288 21.85 minutes). Peak area was converted to grams by using a taxadiene standard curve.



OCT Values
1 – 2012 AU
2 – 2310 AU

T-5α-ol Values
1 – 795 AU
2 – 865 AU

Figure S-16. *GC of 160 µM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. At this taxadiene concentration the production of Taxadiene-5α-ol was also observed (green box at m/z 288 21.85 minutes). Peak area was converted to grams by using a taxadiene standard curve.



OCT Values
1 – 2070 AU
2 –1456 AU

Taxadiene-5α-ol Values
1 – 1049 AU
2 – 708 AU

Figure S-17. *GC of 200 µM Taxadiene Incubation for Kinetics Evaluation.* The m/z filtering information is found in the left side of each panel. The positive control for Taxadiene (blue box) corresponds to the unreacted taxadiene in each substrate trial (blue boxes in m/z 272). OCT (black box at m/z 288, 21.60 minutes) was identified by matching the retention time of this peak to that our previously identified OCT retention time from our previous work. At this taxadiene concentration the production of Taxadiene-5α-ol was also observed (green box at m/z 288 21.85 minutes). Peak area was converted to grams by using a taxadiene standard curve.

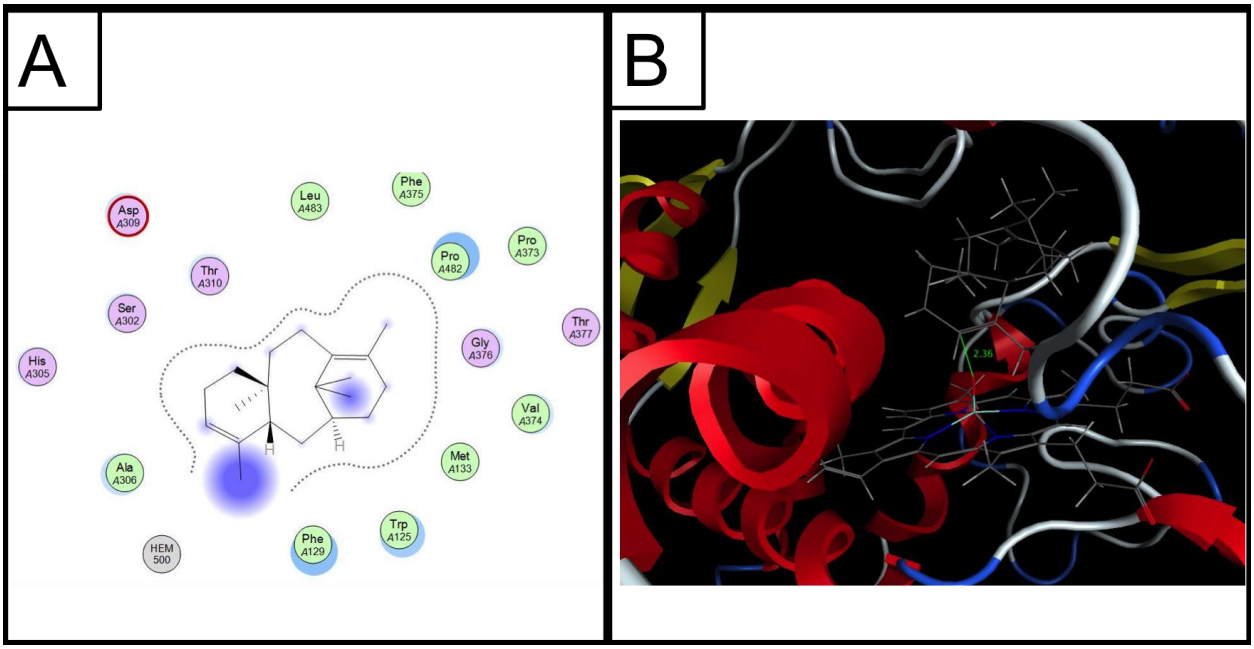


Figure S-18. *MOE Residues Used in Docking and Rendering of Active Site*. (A) MOE Output of Residues used when docking Taxadiene. The purple residues are polar and the red circle signifies Asp as Acidic. Green residues are greasy/non polar. The dark blue circles on Taxadiene signify ligand exposure sites and the lighter blue circles on residues signify receptor exposure. (B) MOE rendering of Taxadiene docked into CYP725A4 Active Site. Taxadiene C5 is 2.36 Angstroms away from the heme.