

## Supporting Information for

### Quaternary Ammonium Oxidative Demethylation: X-ray Crystallographic, Resonance Raman and UV-visible Spectroscopic Analysis of a Rieske-type Demethylase

Kelly D. Daughtry<sup>1</sup>, Youli Xiao<sup>2</sup>, Deborah Stoner-Ma<sup>3</sup>, Eunsun Cho<sup>2</sup>, Allen M. Orville<sup>3\*</sup>, Pinghua Liu<sup>2\*</sup>, and Karen N. Allen<sup>1,2\*</sup>

<sup>1</sup>*Department of Physiology and Biophysics, Boston University School of Medicine, Boston, MA 02218*

<sup>2</sup>*Department of Chemistry, Boston University, Boston, MA 02215*

<sup>3</sup>*Biology Department, Brookhaven National Laboratory, Upton, NY 11973*

\* Address correspondence to Allen M. Orville, email: [amorv@bnl.gov](mailto:amorv@bnl.gov); Pinghua Liu, email: [pinghua@bu.edu](mailto:pinghua@bu.edu); Karen N. Allen, email: [drkallen@bu.edu](mailto:drkallen@bu.edu).

## Supplementary Materials and Methods

*Synthesis of Stachydrine (1).* Stachydrine was synthesized with a slight modification of published procedures.<sup>1</sup> L-proline (4.951 g, 0.043 mol) and NaOH (5.160 g, 0.129 mol) were dissolved in 50 mL of dry methanol. To this solution was added methyl iodide (18.310 g, 0.129 mol, 8.0 mL), and the reaction mixture was refluxed for 6 h. Additional 6.10 g (0.043 mol, 2.7 mL) of methyl iodide was added and the mixture was refluxed for an additional 6 h. The solvent was then removed by rotatory evaporation and resulted in 27.0 g of light yellow solid. A portion of the crude product (10 g) was dissolved in water and applied to a cation-exchange column (H<sup>+</sup> form, AG 50W-X8, 100-200 mesh, 2.5 × 1.5 cm). The column was washed with 250 mL of water and then eluted with 250 mL of 1.5 N HCl solution. Product-containing fractions were combined and the aqueous solvent was evaporated under reduced pressure. The desired product was extracted from the solid using dry methanol. The final pure stachydrine was obtained by recrystallization using methanol/ether as the solvent. Stachydrine properties:  $[\alpha]_D^{24} = -20.4^\circ$  (c, 1.00, MeOH); <sup>1</sup>H NMR (D<sub>2</sub>O, 400 MHz): 4.23 (dt, *J* = 9.6, 4.0 Hz, 1 H), 3.64-3.69 (m, 1 H), 3.48-3.55 (m, 1 H), 3.25 (s, 3 H), 3.06 (s 3 H), 2.46-2.49 (m, 1 H), 2.27-2.32 (m, 1 H), 2.10-2.16 (m, 2 H); <sup>13</sup>C NMR (D<sub>2</sub>O, 75 MHz) 169.26, 74.04 (d, *J* = 16.6 Hz), 67.91, 52.17 (d, *J* = 12.6 Hz), 46.12 (d, *J* = 18.9 Hz), 24.31, 18.34. High resolution Mass Spectrometry (ESI<sup>+</sup>): calculated for C<sub>7</sub>H<sub>14</sub>NO<sub>2</sub><sup>+</sup> 144.1025, found 144.1019.

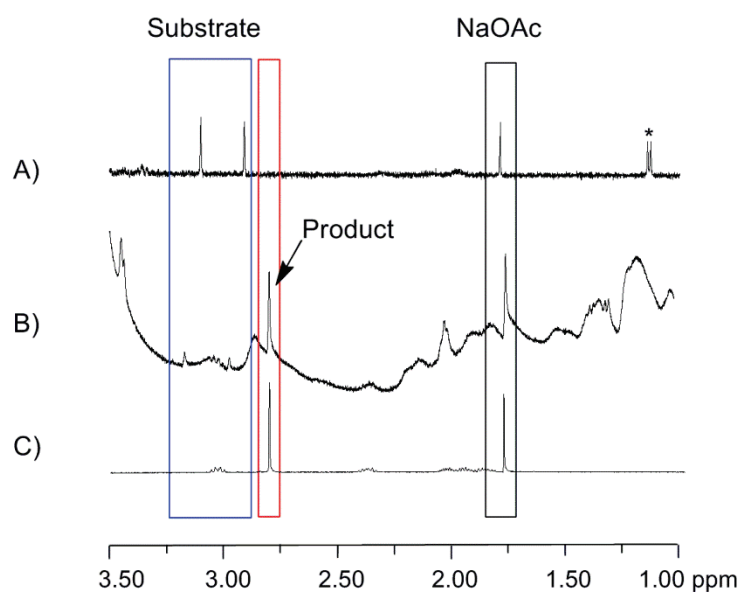
*Synthesis of N-methyl-proline (2).* N-methyl-proline was synthesized according to the literature procedure.<sup>2</sup> L-Proline (2.0 g, 17.4 mmol) was dissolved in methanol (20 mL) and 40% aqueous formaldehyde (1.4 mL, 19.1 mmol) was added to this solution. Next, 10% Pd/C catalyst (500 mg) was added to the reaction mixture and the resulting slurry was stirred under hydrogen atmosphere overnight. The slurry was then filtered through a Celite pad to remove the catalyst. The pad was washed with methanol and the combined filtrates were concentrated under reduced pressure. The residue was dissolved in ethanol/benzene (1:1, 100 mL) and concentrated a second

time to provide a solid that was re-crystallized from methanol/diethyl ether solution as fine needles. (2.0 g, 90% yield).  $[\alpha]_D^{24} = -79.4^\circ$  (c, 2.00, MeOH);  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ , 500 MHz): 3.71-3.75 (dd,  $J = 9.5, 7.5$  Hz, 1 H), 3.55-3.59 (m, 1 H), 3.98-3.01 (m, 1 H), 2.77 (s, 3 H), 2.30-2.37 (m, 1 H), 1.81-2.01 (m, 3 H);  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , 100 MHz) 173.60, 70.49, 56.18, 40.56, 28.67, 22.70.

*Stc2*  $^1\text{H}$ -NMR Assay. The *Stc2* assay mixture containing 0.4 mM *Stc2* and 0.4 mM stachydrine (with equimolar NaOAc as internal standard), and 1.0 mM Fe(II) in 50 mM Tris pH 7.5 buffer was prepared under anaerobic condition and reduced by 1.0 mM dithionite in a Coy-anaerobic chamber. The mixture was then removed from the Coy chamber and exposed to oxygen at room temperature for 20 hours to initiate the oxidative demethylation reaction. To analyze the product formation, the reaction mixture was treated with  $\text{CHCl}_3$  to denature the protein, and the resulting supernatant, after centrifugation at 14,000 g, was analyzed by proton NMR (500 MHz, Figure S1). In  $^1\text{H}$ -NMR spectrum, the two methyl groups of stachydrine are at 3.16 and 2.97 ppm, respectively and the *N*-methyl-proline methyl group is at 2.80 ppm.

#### *Structural Homology Analysis*

The CE algorithm<sup>3</sup> implemented in PyMol was used to align each structure with *Stc2* using secondary structure matching. Visual inspection of each structure was used to determine the Rieske and catalytic domain boundaries, which were then independently aligned with the corresponding *Stc2* domains using the CE algorithm (Table S1).



**Figure S1.** <sup>1</sup>H-NMR Spectra from Stc2 assays. A) 0.25 mM stachydrine /NaOAc in D<sub>2</sub>O \*impurity from the NMR tube. B) Stachydrine oxidative demethylation reaction mixture after protein was removed by CHCl<sub>3</sub> treatment. C) 1.4 mM *N*-methyl proline and 1.0 mM NaOAc in D<sub>2</sub>O.

STC2 1 MTANPTSIIHQRLDRRL..S...G...FSLEQPFYTSPEVYALDLQHIIFYKQWLYAVPVCOIAKAGSYITL  
 NDO 1 MLSNELRQTLQKGLHD..V...NSDWTVPAA.IINDPEVHDVDERERIFGHAWVFLAHESEIIPERGDYVVR  
 BPO 1 ...WADADIAE..LV..DERTG...RLDP.RIYTDALYEQELERIFGRSWLLMGHETQIRKAGDFMTN  
 TDO 1 ...WNTSEIEA..LF..DEHAG...RID.PRIYTDLEDLYQLERLRFARSWLLMGHETQIRKAGDFMTN  
 CDO 1 ...NWSDEEIKA..LV..DEEKG...LLDP.RIFSDQDLYEIELERVFARSWLLMGHETQIRKAGDFMTN  
 NBO 1 .....YQNL..VSEAG...L...TQKIL.IHGDKELFQHELKTIFARNWLF.LTHDSLIPSPGDYVKA  
 PHO 1 .....MSGDTTL.VD..T.VNA.....SQS.RQVFWDRQVYDLEIERIFSRAWLM.LGHKSLIPKPGDFIT  
 DDO 1 .....AGIAERR..TR..A.....WAPYIDAKL..GERNHWPVRLSAEVA.EASPVV  
 CBO 1 .....ISDARANNA..KT..Q.....SQYQPYKDAW..GFINHWYPLALFTHET.EEDQVCGI

STC2 63 RVGAYE VVI VRS RDG EVRAFNSCHRRGSLICKARQGO VAKLVC FYHQW IYELD GKL IWR NDM...  
 NDO 65 YI SEDQ FIV CRDEGG EIRGHLNCRHRCMQV CRAEMGN TSHFNC FYHQW IY SNT GSLVGV PAG...  
 BPO 59 YMGEDP VMV VRC KNG EIRVFLNCC HRCMRI CRADGGN AKSFTCS YHGWAY DTG GNLVSV PFE EQA  
 TDO 59 YMGEDP VVV VRC KDA SIAVFLNCC HRCMRI CRADAGN AKAF TCS YHGWAY DTA GNLVNV PYE...  
 CDO 60 YMGEDP VIV VRC KDR SIKVFLNCC HRCMRI IERSDFGN AKSFTCT YHGWAY DTA GNLVNV PYE...  
 NBO 54 KMGVDF VIV VRC NDG SVRAF LVC HRC KTLVHAEAGN AKGFVCC YHGWAY GSN GELQSV PFEKELY  
 PHO 57 YMAEDK IILSHCSDG TFRAFINSCHRRGNCI CHADSGN AKAFVCM YHGWAY GQD GSLVDM FLE...  
 DDO 26 TILDTP LAL YRC PDG VVAALLD ICHRRHAP LSD .GILVNGH LCC FYHGLEF DGG GQCMHN P.H...  
 CBO 43 QLLGK VLLN RV DG VVAIAIADRC HRCVTLSDKVECYSKAT I SCWYHGWTY RWDNG KKLVDILTNP...  
 OMO 47 QICCVPLVI RRV NG KVFALK DCC HRCVRI ISEKPTCET KST I SCWYHGFTF DLETGKLVTVVANP...

STC2 126 G.....PDFDASKYGLKPV.NLRNLDGLIYICLS.D.....TPPDFQTFAQLARPYLE.....V  
 NDO 128 K.DAYGNQLKKS DWNLRPMPN LAS YKGLIFGSLD.P.....HADSLEDYLGDLKPYLD.....E  
 BPO 125 F.....PGLRKEDWGPLQA.RVET YKGLIFANWD.A.....DAPDLDTYLGEAKFYMD.....H  
 TDO 122 A...ESFACLNKKK EWSPLKA.RVET YKGLIFANWD.E.....NAVLDOTYLGEAKFYMD.....H  
 CDO 123 KEAFCDCCGFDKADWSPLOA.RVDIYKGLIFANWD.T.....EAPDLKTYLSDATPYMD.....V  
 NBO 121 G.....DAIKKKKCLGLKEVPR IESFHGF IYGCFD.A.....EAPPLIDYLGDAAWYLE.....P  
 PHO 120 S.RCYHNKLDKQELAAKSV.RVET YKGFIFGCHD.P.....EAPSLEDYLGDFRIFYLD.....T  
 DDO 87 TFR.YSGGL ELV.GPPGKVVVKANNSFAEN FVGDGYHV.GWTHA.MAL..RAGQSVFS.....SI..  
 CBO 108 T.....SVQIGRHALKTY.PVRE EKGLVFLVVG.D.....QEPHDLAE.....DVPPG  
 OMO 112 E.....DKLIGTTGVTTY.PVHEVNGMI FVVFVR.EDDFDEDPVPLAH.DLPFRFPERSEQFPPLWFS

STC2 173 HDLK...DAKVA.F.TSTIIEKONWKLWVENN.R.ECYHC.SSNHP.ALC.....ASFP.....LDPE  
 NDO 180 VLDRSDAGLQVV.GAPQRWVIDANMKLGADNFVGD DAYHT.MMTHR.SMV...ELGLAP...PDPQFAL  
 BPO 172 MLDRT EAGT EAI PG.IQKWVIPCNWKF AA EQFCSDMYHAGTTSHL.SGI.....L.....  
 TDO 172 MLDRT EAGT EAI PG.VQKWVIPCNWKF AA EQFCSDMYHAGTTSHL.SGI.....L.....  
 CDO 175 MLDRT EAVT QVITG.MQKTVIPCNWKF AA EQFCSDMYHAGTMAHL.SGV...L.....S  
 NBO 169 TFR.YSGGL ELV.GPPGKVVVKANNSFAEN FVGDGYHV.GWTHA.MAL..RAGQSVFS.....SI..  
 PHO 171 IWEGGGAGLELL.GPPMKSLHCCNWKVVENFVGDGYHV.GWTHA.MAL.....TDAF.....DRLE  
 DDO 129 CRVD..PAYRT..V.GGYGHVDCNYKLLVONL.MDLGHA.QYVRRANAQ.....TDAF.....DRLE  
 CBO 149 FLDA...DLAV..H.GQHRVVDANWRMGVENG.FDAGHV.F.HRK.SSILLDGNIALP.....LGFAP  
 OMO 173 SPSVLDDNAVV..H.GMHRTPSNWRHACENG.FDNABI.L.VHK.CNT.....IVHAMDWVLPGLLPL

STC2 222 VAGVQADGGVSKKLQAHFDRCEAAGTFAQFVLAGDG..QYRLARM..FL.....QEKA  
 NDO 239 .....YGEHIHT.GH..GHGLGII..GP.....PPGMPL  
 BPO 220 .....AGLTEGIIQYRATWG..GHGSGFYIGDP.....NL..  
 TDO 220 ..AGLPEEMADLAPPTVG.....KQYRASWG..GHGSGFYVGGDP.....NL.M  
 CDO 224 SLPEEMDLSQVKLPSSG.....NQFRAKWG..GHGTGWF..ND.....DFAL  
 NBO 224 ...AGNAKLPEEGA.....GLQMTS.KY..CSGMGVF..WG.....Y  
 PHO 217 .....LGLQFTT.RH..GHGFGVI..DN.....A  
 DDO 179 R.....EVIVG..D..GEIQALM..KI.....PGGT  
 CBO 203 .....GDPEQLTRSVTGECAKGVFDLLGHSVPIFEATIEGQPAIQGHMGSK  
 OMO 230 .....TSDDCIADVVEDDDGPKGMMQWL..FTDKWAPVLENQELGLKVEGLK

STC2 271 .LSYTM D GKA..AVSRHLG.....RVAP.PDAGTLLMFHY P STWNHFL..P.....  
 NDO 263 .P....EFMG..LPENI.....VEELERRLTPEQVEIFRPTAFIHGTVP PNLISIGNF.....LMGKDH  
 BPO 247 LLAIM..GPK..V.TEYWTQGPAAEKASERL GSTERGO.Q.LMAQHMTIF P TCSFLPG.....  
 TDO 258 .LAIM..GPK..VTSYW.TEGPASEKAAERLGVSTERGO.K.LMVEHMTVFP TCSFLPG.....  
 CDO 262 LQAIM..GPK..VVDYWTGPAERAKERLGVLPADR..MVAQHMTIF P TCSFLPG.....  
 NBO 253 YSGNFSADMIPDLMAFG...AAKQEKLAKEIGDVRAR.IYRSFLNGTIF P NNSFLTG.....  
 PHO 236 AAAIH.RKGD.GWNKYL...EDTRGEVRRKFGADRER.LYVGHWNGAIF P NCSFLYG.....  
 DDO 199 .P.....SVL..M.....AKFLRGANTPVDANWDIRWNKVSAMLNFI...AVAPEGTPK  
 CBO 251 .MV.....AISISVWLP GVLKVDPPFPDP.....  
 OMO 275 .R.....HYRTSVVLP GVLMVENWPE.....

```

STC2 311 . . . . D H S L T F R V M P I S P T E T E V I T W L V H K . . . . D A V E G V D Y D . . L . . K R L T . E . V W . . . I . . . A T N D . E
NDO 315 S A P T A F L T L R L W H P L G P D K M E V M S F F L V E K . . . . D A P D . W F K D E S Y . . K S Y L . R . T F G I S G . . . G F E Q . D
BPO 298 . . . . . I N T I R A W H P R G P N E I E V W A F T V V D A . . . . D A P E . E M K E E Y R . . Q Q T L . R . T F S A G G . . . V F E Q . D
TDO 308 . . . . . I N T V R T W H P R G P N E V E V W A F T V V D A . . . . D A P D . D I K E E F R . . R Q T L . R . T F . . . S A G G V F E Q . D
CDO 313 . . . . . I N T V R T W H P R G P N E I E V W S F I V V D A . . . . D A P E . D I K E E Y R . . R K N I . F . T F N Q G G . . . T Y E Q . D
NBO 306 . . . . . S A A F R V W N P I D E N T T E V W I Y A F V E K . . . . D M P E D . L K R . . R V A D A V Q . R . S I G P A G . . . F W E S . D
PHO 287 . . . . . T N T F K I W H P R G P H E I E V W Y T M V P S . . . . D A D P A . T K S . . A I Q R E A T . R . T F G T A G . . . T L E S . D
DDO 242 E Q S I H S R G T H I L T P E T E A S C H Y F F G S S R N F . . . . G E D D P . . S M . . D . . G V L R . S . W Q . . . A . . . Q A L V K E
CBO 273 . . . . . T L T Q F E W Y V P I D E G H H L Y L O M L G R R V G S E E E A R S . F E A E . . F . R E K W V E L . A L . . . N . . . G F N D . S
OMO 296 . . . . . H V V Q Y E W Y V P I T D D T H E Y W E I L V R V C P T O E D R K K K . F Q Y R . . Y . D H M Y K . P L C L . . . H . . . G F N D . S

```

```

STC2 360 D R E I V E T N Q Q G I L S . . P A Y . . V P G P Y S P G Q E S G V M Q P V D W Y A A S L E R A L A P R Q V A A E . . . . .
NDO 372 D A E N W R S I T T R V M G G . . Q F A K T G E L N Y Q M G R G V L E F D P N W T G P G E A Y P L D Y A E A N Q R N F L E Y W M Q L M L A E S
BPO 350 D G E N W V E I C Q V L R G . . H K A R S R P F N A E M G . . L G Q T D S D N P D Y P G T I S Y V Y S E E A A R G L Y T O W V R M M T S P D
TDO 360 D G E N W V E I Q H I L R G . . H K A R S R P F N A E M S . M D Q T V D N D P V Y P G R I S N N V Y S E E A A R G L Y A H W L R M M T S . .
CDO 365 D G E N W V E V C R G L R G . . Y K A R S R P L C A Q M G . . A G V P N K N N P E F P G K T S Y V Y S E E A A R G F Y H H W S R M M . S E P
NBO 358 D N E N M E T S Q N G K K . Y Q S S N I D Q I A S L G . . F G K D V Y G D E C Y P G V V G K S A I G E T S Y R G F Y R A Y Q A H I S S S N
PHO 339 D G E N M S S A T Y V N R G . . V I T R D G M M N S T M G . . V G Y E G P H P V Y P G I V G I S F I G E T S Y R G F Y R F W K E M I D A P D
DDO 294 D K V V W E A I E R R R R A Y V E A N G . . I R P A M L . S C D E A A V R V S R E I E K L E Q L E A A R L . . . . .
CBO 327 D I L A R R S M E P F Y A D . . D R G W R E E V L P . E S D R A I I E W R R L A S Q Y N R G I Q T R D . . . . .
OMO 350 D L Y A R E A M C N F Y Y D . . G T G W D D E Q L V . A T D I S P I T W R K L A S R W N R G I A K P G R G V A G A V K D T S L I F K Q T A

```

```

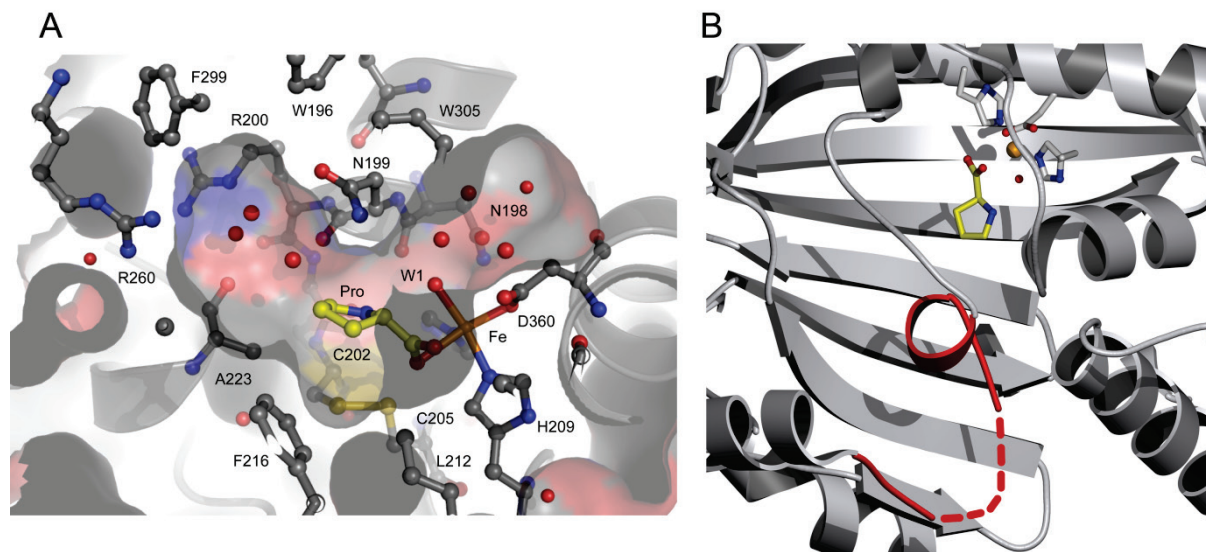
STC2
NDO 440 P L . . . . .
BPO 416 W A A L D A T R P A . .
TDO
CDO 430 S W D T L K S . . . . .
NBO 425 W A E F E N A S R N W H I
PHO 405 W A S V K A N D D N W D S V F T N R N F W
DDO
CBO
OMO 416 D G K R P G Y K V E O I

```

**Figure S2.** Sequence alignment of Rieske-type enzymes structurally homologous to Stc2. Details for these enzymes are listed in the following order: (enzyme name, abbreviation, PDB ID, sequence identity): naphthalene 1,2-dioxygenase, NDO, 1NDO, 21%; biphenyl dioxygenase, BPO, 1ULI, 21%; toluene 2,3-dioxygenase, TDO, 3EQQ, 20%; cumene dioxygenase, CDO, 1WQL, 21%; nitrobenzene dioxygenase, NBO, 2BMO, 24%; PAH-hydroxylating dioxygenase PHO, 2CKF, 21%; dicamba demethylase, DDO, 3GOB, 15%; carbazole 1,9 $\alpha$ -dioxygenase, CBO, 3GKQ, 21%; 2-oxoquinoline 8-monooxygenase, OMO, 1ZO1, 17%.

**Table S1.** Structural similarity of Stc2 to other Rieske-type oxygenases.

<b>Enzyme Name</b>	<b>Abbreviation</b>	<b>PDB ID</b>	<b>Sequence Identity (%)</b>	<b>RMSD Rieske /Catalytic (Overall)</b>
naphthalene 1,2-dioxygenase	NDO	1NDO	21	2.03 / 5.19 (3.25)
biphenyl dioxygenase	BPO	1ULI	21	1.45 / 4.8 (4.6)
toluene 2,3-dioxygenase	TDO	3EQQ	20	1.94 / 5.78 (3.32)
cumene dioxygenase	CDO	1WQL	21	1.70 / 4.61 (3.07)
nitrobenzene dioxygenase	NBO	2BMO	24	2.02 / 5.11 (4.88)
PAH-hydroxylating dioxygenase	PHO	2CKF	21	1.59/ 4.95(3.28)
dicamba demethylase	DDO	3GOB	15	3.67/ 4.35 (4.55)
carbazole 1,9a-dioxygenase	CBO	3GKQ	21	3.67 / 5.29 (4.15)
2-oxoquinoline 8-monooxygenase	OMO	1ZO1	17	3.59 / 5.12 (4.57)



**Figure S3.** A) The bi-lobed, solvent excluded cavity at the mononuclear iron site of Stc2. Several solvent molecules (red, non-bonded spheres) occupy each lobe of the cavity. B) Ribbon diagram of the active Stc2 active site with Fe shown as orange sphere and Fe coordinating residues and product proline shown as ball and stick. The gating loop (residues 215 – 233) is colored red with relative location of disordered residues 226 – 229 indicated by a dashed line).



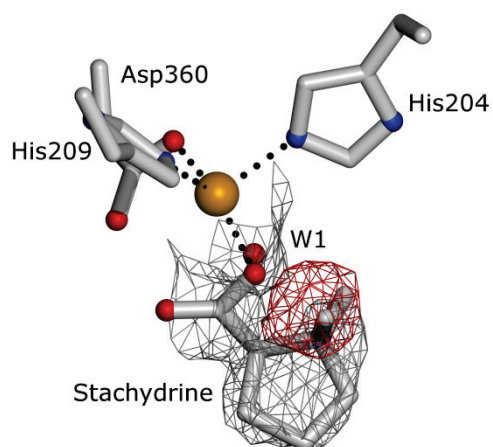
```

Stc2      149 YI1CL2SD3TPP.DFQTFA4QLA5RPYLEV.....HDLKDAK6VAF7TST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
AEG07111.1 149 YI1CL2SD3TPP.DFQTFA4QLA5RPYLEV.....HDLKDAK6VAF7TST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
AEG06496.1 149 YI1CL2SD3TPP.DFQTFA4QLA5RPYLEV.....HDLKDAK6VAF7TST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
YP_004552068.1 149 YI1CL2SD3TPP.DFQTFA4QLA5RPYLEV.....HDLKDAK6VAF7TST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
YP_001314589.1 149 YI1CL2SD3TPP.DFQTFA4QLA5RPYLEV.....HDLKDAK6VAF7TST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
YP_002824538.1 138 YI1CL2SET3TPP.DFEPPFA4QLA5RPYLAI.....HDLSEAK6VAY7TST8IIVE9KAN10W11KL12VW13EN14NR15EC16Y17HC18SS19N20HP21E
YP_914990.1 187 YI1CL2AD3TPP.DFEAFARA4EPYLGV.....HDLQDAK6VAY7SS8SIIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_01156510.1 145 YI1CL2AGE3PP.DFAFFADLA4RPYLEV.....HDLGRAK6VAH7TSS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_01747515.1 146 YI1CL2ADE3AP.DFDAPANLA4RPYLEV.....HDLHRSK6VAH7QSS8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17AG18T19HP20E
YP_611455.1 146 YI1CL2ADE3AP.DFERFAEVA4RPYLEV.....HDLNSAK6VAH7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_05078014.1 148 YI1CL2ADE3AP.DFDAPAAVA4RPYLEP.....HDLGNAK6VAY7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_05739140.1 146 YI1CL2ADD3AP.DFEQFADVA4RPYLEI.....HDLNSAK6VAF7ESS8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_01753851.1 146 YI1CL2ADQ3AP.DFDAPADLA4GPYLEP.....HDLANAK6VAH7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_05089427.1 146 YI1CL2ADD3AP.DFDTFANLA4RPYLEP.....HDLANAK6VAF7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_02146483.1 148 YI1CL2ADE3AP.DFEVFA4DLV5RPYLAP.....HDLGNAK6VAF7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_02149080.1 148 YI1CL2ADE3AP.DFEVFA4DLV5RPYLAP.....HDLGNAK6VAF7ESS8IIVENG9NW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_05783620.1 149 YI1CL2AE3AP.DFEAFAEIV4TPYLGV.....HDLSDAK6VAF7QST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17AG18N19HP20E
ZP_01013250.1 148 YI1CL2AD3RAP.DFDTFAAVA4EPYLAP.....HDLTNAK6IAH7QST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17SS18N19HP20E
ZP_01445117.1 149 YI1CL2AE3AP.DFEPPAQTV4SPYLGV.....HDLNSAK6VAF7QST8IIEK9GNW10KL11VW12EN13NR14EC15Y16HC17AG18N19HP20E
ZP_01547669.1 148 FI1CL2AV3AP.DTTELORCA4ERYLAP.....HDLNSL6NK7VAF8H9OST10IIVE11CO12GNW13KL14VLE15EN16NR17EC18Y19HC20SS21SH22PE23E

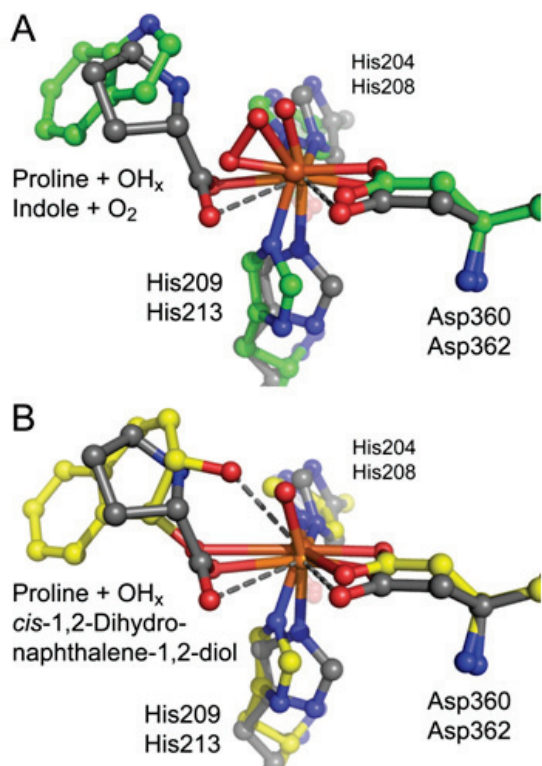
YP_001852684.1 150 WV1CLA2ED3PPSFE4DDV5MG6SV7TARLGDVSAIDTYR8HEAL9OV10GR11RV12SY13DVA14NW15KL16LV17EN18FM19EC20Y21HC22ATI23HP24E
gb|ADI03488.1 145 WS1LAD2AP3PSFE4QDV5IGAV6IERLGDVAAIDNYR7IEN8LEV9GR10RIT11YD12VKAN13W14KL15LV16EN17FM18EC19Y20HC21ATI22HP23E
NP_828125.1 151 WV1CLA2DE3PPSFE4ETV5MG6AA7VERLGDAAAIEHYGT8EN9LALG10KR11IT12YD13VKAN14W15KL16LV17EN18FM19EC20Y21HC22ATI23HP24E
ZP_06771318.1 166 WL1CLA2EE3PPSFE4DEH5VRGAA6TERLGDPAADRYDAEGL7MLG8HR9VR10YD11VKAN12W13KL14LV15EN16FM17EC18Y19HC20STI21HP22E
YP_002778679.1 160 WV1CLA2DE3PPSFE4DEV5IGSV6TERLGDPAADRYE7IG8TLE9LGR10RIT11YD12VKAN13W14KL15LV16EN17FM18EC19Y20HC21ATI22HP23E
YP_004015212.1 150 WV1CPA2AS3PPSFE4EET5VI6GAA7TERLGDPAAIERYQV8AGL9RVG10RR11VS12YD13VKAN14W15KL16LV17EN18FL19EC20Y21HC22PAI23HP24E
YP_593084.1 151 FI1NLS2AHP3IPFSE4HLAG5LQ6QK.....FRPWR7MEEL8QMV9ERR10VYHL11KAN12W13KL14LV15ION16YSE17CL18HC19PIV20HP21E
ZP_05005959.1 166 WL1CLA2EE3PPSFE4DEH5VRGAA6TERLGDPAADRYDAEGL7MLG8HR9VR10YD11VKAN12W13KL14LV15EN16FM17EC18Y19HC20STI21HP22E
ABB73050.1 110 WV1CLA2EE3PPSFE4EEV5MG6AI7TERLGNVHAIDGYD8ANL9SVG10RR11IR12YD13VKAN14W15KL16LV17EN18FM19EC20Y21HC22ATI23HP24E
YP_003409890.1 146 WV1CI2ADV3PPSFE4EDT5VI6GAV7AGRLGSP8EV9IDDW10GI11DKL12TVG13RR14IT15YD16VKAN17W18KL19LV20EN21FM22EC23Y24HC25ATI26HP27E
YP_002521377.1 144 WV1NLS2ND3PPPP4VES5QIV6PR7IERLGDRAIFDRYGV8GG9QL10RIG11AT12LT13YD14VCAN15W16KL17LV18EN19FM20EC21Y22HC23AV24V25HP26E
YP_003596731.1 150 WV1NLS2DN3PMS4VEY5QLD6AG7IFDR8FGELY9TFARYE10IQ11NL12KS13A14H15RE16E17Y18VE19AN20W21KL22LV23EN24FM25EC26Y27HC28SS29I30HP31E
YP_003562007.1 150 WV1NLS2DN3PMS4VEY5QLD6AG7IFDR8FGELY9TFARYE10IQ11NL12KS13A14H15RE16E17Y18VE19AN20W21KL22LV23EN24FM25EC26Y27HC28SS29I30HP31E
YP_001635947.1 147 FV1SLA2DQ3PP4PFTE5AFAP6LI7GK.....FQAWEM8ARL9ARG10AR11IT12Y13VE14AN15W16KL17LV18EN19FM20EC21Y22HC23PLI24HP25E
YP_003318418.1 146 WIS1LA2ED3PPPLA4EQLE5PAI6LRR7FGELDTVGRYRL8ATL9AVG10HSI11TY12EL13AC14NW15KL16VW17EN18FM19EC20Y21HC22AV23T24HP25E
NP_926381.1 150 WLN1LD2PQ3PT4PFES5V6FAP7LI8GR.....FTGWRL9LPAL10RLG11AR12VT13YD14LRAN15W16KL17LV18EN19FM20EC21Y22HC23PII24HP25E
YP_477505.1 166 WV1NLS2RNP3Q4PF5DEV6FAP7LI8GR.....FREWN9LSAL10LRG11H12RIE13YN14VAN15W16KL17LV18EN19FM20EC21Y22HC23PLI24HP25E
YP_002464338.1 147 FV1SLA2EQ3VAFERAFAP4LI5GK.....FQAWDP6GR7LR8C9GA10QI11VY12VE13AN14W15KL16LV17EN18FM19EC20Y21HC22PLI23HP24E
YP_885066.1 170 FV1DP2SG3QDAD4FATHV5AGLE6EV.....VSPYR7PE8DL9TIV10AR11HS12YEL13AN14W15KL16LV17EN18FM19EC20Y21HC22SS23I24HP25E
YP_004177633.1 147 FV1SLA2EN3PP4PLAERLE5EG6VEL.....ARPYE7MGR8L9K10RG11AR12VV13Y14VE15AN16W17KL18LV19EN20FM21EC22Y23HC24PAN25HP26E

```

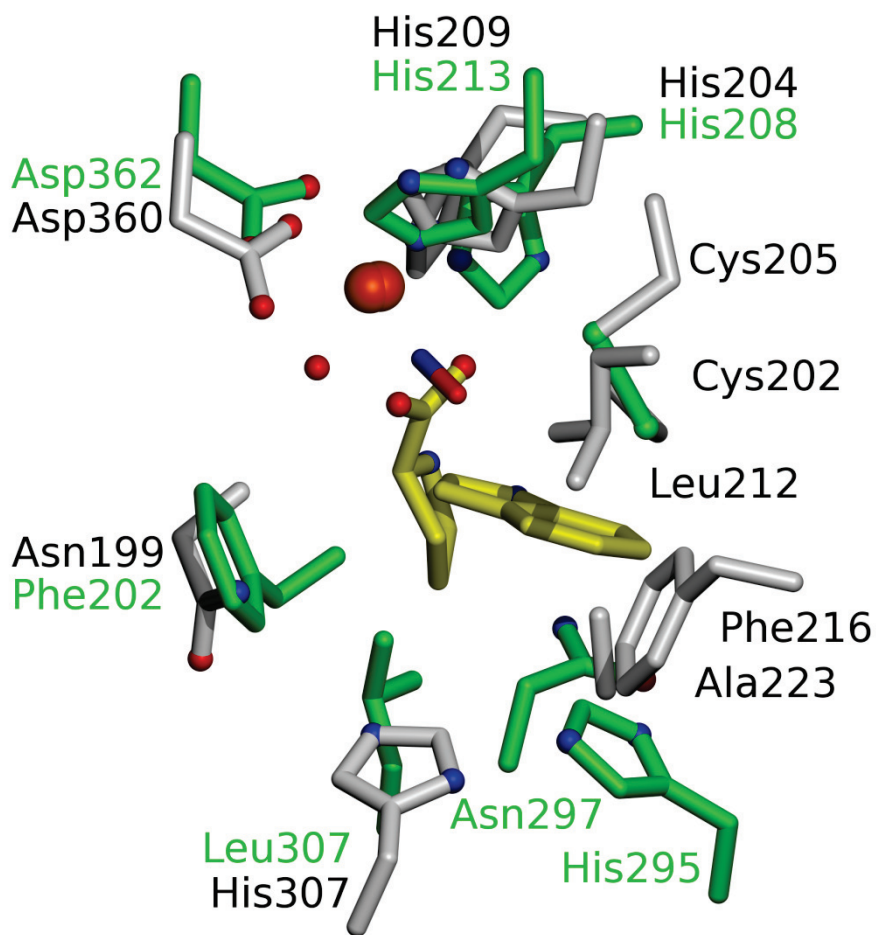
**Figure S4.** Alignment of the portion of the sequences containing the strictly conserved CxHC motif (signature of disulfide proximal to the mononuclear Fe site) identified using as query the Stc2 sequence (SMa0751, NP\_435626). An example set of 20 sequences with high overall identity (top, sequence identity range 57 – 98%) and low overall sequence identity (bottom; sequence identity 31 – 24%) are shown with alignment between them retained.



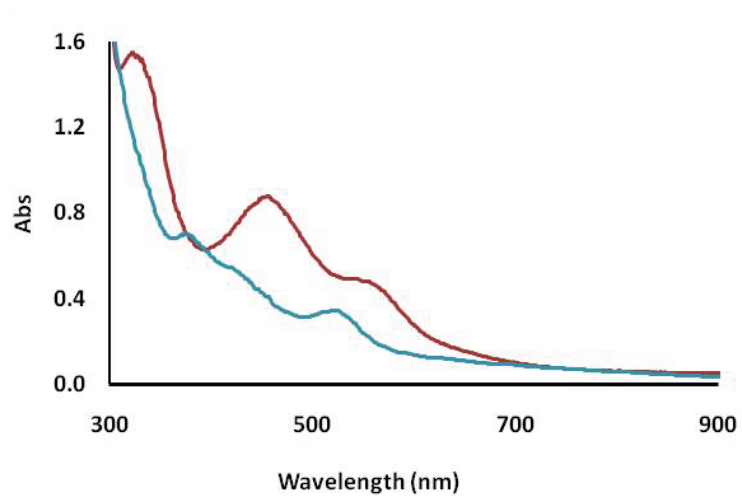
**Figure S5.** Structure of Stc2 mononuclear Fe active site with stachydrine modeled showing electron density calculated between 39.7 - 2.2 Å resolution with coefficients 2Fo-Fc (gray cages) contoured at 1.0  $\sigma$  and calculated with coefficients Fo-Fc (red cages) contoured at -3  $\sigma$ .



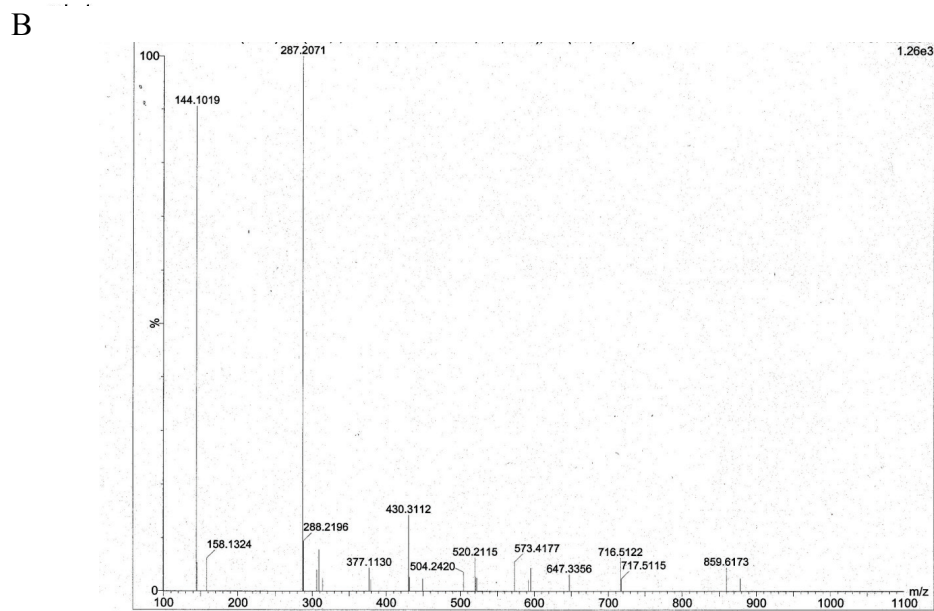
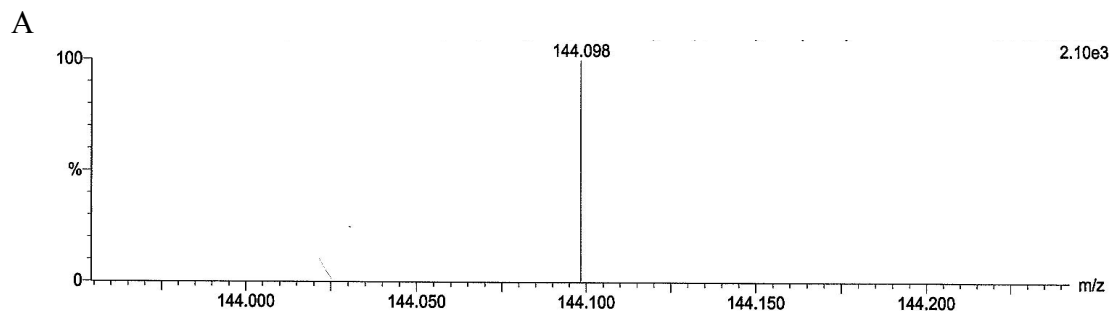
**Figure S6:** Overlay of the Stc2 ligand complex with naphthalene 1,2-dioxygenase ligand complexes. Superposition was based upon the best overlay of the C $\beta$  atoms from three protein residues that coordinate the mononuclear iron center (His204/208, His 209/213 and Asp360/362). Long interactions are indicated with gray dashed lines. A) Stc2-Pro (C, N, O, and Fe atoms colored gray, blue, red and orange, respectively; residue are labeled in the top line) superimposed with the ternary complex of naphthalene 1,2-dioxygenase, indole and O<sub>2</sub> (C atoms colored green, others by atom type; residue are labeled in the bottom line; PDB 1O7N). B) Stc2-Pro overlaid on the naphthalene 1,2-dioxygenase product complex (C atoms colored yellow, others by atom type; PDB 1O7P).



**Figure S7:** Superposition of the hydrophobic active sites of Stc2 and NDO. Side chains lining the active site of the complex of Stc2 (grey sticks) with proline (yellow sticks) are shown superimposed with those of the complex of naphthalene 1,2-dioxygenase (green sticks) with indole (yellow sticks) and NO (PDB 1UUV). The water molecule (red sphere) is from the Stc2 structure. Fe is shown as an orange sphere.



**Figure S8:** Solution spectra of Stc2 (0.4 mM) in 100 mM Tris, pH 7.5 buffer with 2.5 mM desthiobiotin used to elute protein from the Streptavidin column (red trace, as-isolated Stc2; blue trace, Stc2 reduced by 0.6 mM sodium dithionite).



**Figure 9:** Mass spectrometry data for the stachydrine A) before x-ray irradiation (expected mass of stachydrine  $[\text{C}_7\text{H}_{14}\text{NO}_2]^+ = 144.1025 [\text{M}]^+$ , observed mass 144.098) and B) after exposure of 100 mM stachydrine (in water) at room temperature to approximately ten times the X-ray dose used for x-ray diffraction studies ( $[\text{C}_7\text{H}_{14}\text{NO}_2]^+ = 144.1025 [\text{M}]^+$ ) and observed 144.1019 and 287.2071  $[2\text{M}-\text{H}]^+$ .

## References

- (1) Woster, P. M.; Murray, W. J. *J. Med. Chem.* **1986**, *29*, 865.
- (2) Han, Z.-j.; Wang, R.; Zhou, Y.-f.; Liu, L. *ChemInform* **2005**, *36*.
- (3) Shindyalov, I. N.; Bourne, P. E. *Protein Eng* **1998**, *11*, 739.