Supporting Information

P450 monooxygenase ComJ catalyses side chain phenolic cross-coupling during complestatin biosynthesis

Aurelio Mollo, A. Nikolai von Krusenstiern, Joshua A. Bulos, Veronika Ulrich, Karin S. Åkerfeldt, Max J. Cryle* and Louise K. Charkoudian†

* max.cryle@monash.edu
† lcharkou@haverford.edu

Content

Design of Gibson Assembly primers and Gibson Assembly protocol......................................................2
Gibson Assembly primer list and thermal cycling conditions....................................................................3
Standard cloning primer list.....................................................................................................................4
LC/MS analysis of turnover mixtures......................................................................................................5
Fig. S1: SDS-PAGE gel and titers of all proteins used.............................................................................6
Fig. S2: SEC interaction analyses of ComJ and ComI with \( X_{\text{com}} \)......................................................7
Fig. S3: SEC interaction analyses of ComJ and ComI with PCP7\( \text{com} \).............................................8
Fig. S4: Turnover of \( \text{d-Hpg-d-Hpg-L-Tyr (1)} \) on PCP7-\( X_{\text{van}} \) by OxyB\( \text{van} \)..........................................9
Fig. S5: Turnovers of \( \text{1} \) on alternate PCP domains and extended periods.........................................10
Fig. S6: ComK possesses all the spectroscopic features of a ferredoxin..................................................11
Design of Gibson Assembly primers and Gibson Assembly protocol

All pANK constructs were designed to have the open reading frame (ORF) inserted into a pET28a expression vector. Thus, a pET28a vector was linearized through digestion with Ndel and EcoRI for 4 hr at 37 °C and purified by gel electrophoresis and gel extraction (GenCatch Advanced Gel Extraction Kit, Epoch Life Science). In the case of pAM2, the starting vector was pET24b-MBP-PCP7-Xvan,\(^1\) which was digested overnight with Ncol and Xhol, and again purified by gel electrophoresis and gel extraction.

Separately, primer pairs were designed to amplify the ORFs of interest from \(S.\ lavendulae\) gDNA (see next page). Each primer was composed of the following two elements: (i) 19 base pairs of homology to the desired insert sequence, allowing for base pairing during PCR amplification; and, (ii) 25 base pairs of homology with the linearized pET28a/pET24b vector, containing the restriction site. The primers were ordered from Eurofins Genomics and used to amplify the gene of interest \textit{via} PCR (see next page), and the amplicon purified by gel electrophoresis and gel extraction.

Once the insert and vector had been obtained as described above, each at a concentration of > 20 ng/μL after gel extraction, the two were combined in a 3:1 insert:vector \textit{molar} ratio in a volume of 5 μL and added to 15 μL of Gibson Assembly\textsuperscript{®} Master Mix (New England BioLabs), consisting of T5 Exonuclease, Phusion Polymerase, Taq Ligase, dNTPs, and MgCl\(_2\) in Tris-HCl buffer. The resulting mixture was incubated for 1 hr at 50 °C and then transformed into DH5α competent cells \textit{via} a 90-second, 42 °C heat shock. Resulting colonies were used to inoculate 10 mL LB/kan cultures (50 μg/mL) and grown overnight at 37 °C, 225 RPM. Plasmid DNA was isolated using a Qiagen QIAprep Spin Miniprep Kit and sequenced using standard T7 + T7term primers (Eurofins Genomics).

Gibson Assembly primer list and thermal cycling conditions

Below are the primers used to amplify *S. lavendulae* gDNA for the construction of plasmids encoding ComI, ComJ, ComK, and MBP-PCP7-X\textsubscript{com}. Restriction sides are \textbf{bolded} (restriction enzyme written on the right of the primer sequence), while the sequence complementary to the insert is \textit{underlined}.

Primers for plasmid encoding ComI (pANK00, pET28a):

\begin{itemize}
  \item For: \textit{CCTGGTGCCGC\underline{CGCGG}CAGCC\underline{CAT}G\underline{GC}GTC\underline{CCC}GT\underline{GCC}GCT\underline{CCC}C\underline{CG}}
  \item Rev: \textit{AAGCTTGTCG\underline{ACCG}GAGCTG\underline{AATT}C\underline{TACCAG}G\underline{GAGG\underline{CCGGGC}}ACG}
  \item Ndel
  \item EcoRI
\end{itemize}

Primers for plasmid encoding ComJ (pANK01, PET28a):

\begin{itemize}
  \item For: \textit{CCTGGTGCCGC\underline{CGCGG}CAGCC\underline{CAT}G\underline{GC}GTC\underline{CCC}GT\underline{GCC}GCT\underline{CCC}C\underline{CG}}
  \item Rev: \textit{AAGCTTGTCG\underline{ACCG}GAGCTG\underline{AATT}C\underline{TACCAG}G\underline{GAGG\underline{CCGGGC}}ACG}
  \item Ndel
  \item EcoRI
\end{itemize}

Primers for plasmid encoding ComK (pANK12, PET28a):

\begin{itemize}
  \item For: \textit{CCTGGTGCCGC\underline{CGCGG}CAGCC\underline{CAT}G\underline{GC}GTC\underline{CCC}GT\underline{GCC}GCT\underline{CCC}C\underline{CG}}
  \item Rev: \textit{AAGCTTGTCG\underline{ACCG}GAGCTG\underline{AATT}C\underline{TACCAG}G\underline{GAGG\underline{CCGGGC}}ACG}
  \item Ndel
  \item EcoRI
\end{itemize}

Primers for plasmid encoding MBP-PCP7-X\textsubscript{com} (pAM2, pET24b):

\begin{itemize}
  \item For: \textit{AA\underline{T}TTTATTTTT\underline{CAGGGGC}\underline{CCATGGG}CC\underline{CGGGGT}G\underline{TCGCCG}GACCGGCCC}
  \item Rev: \textit{AGTGG\underline{TGGG}TGGG\underline{TCGAGG}GT\underline{TCACCTGCGGCTCGGTG}}
  \item Ncol
  \item Xhol
\end{itemize}

Thermal cycling conditions for pANK00 and pANK01 (5% DMSO, 1.1 \mu g gDNA, Taq Polymerase):

\begin{enumerate}
  \item 94 °C \hspace{1cm} 3:00
  \item 94 °C \hspace{1cm} 1:00
  \item 62 °C \hspace{1cm} 1:00
  \item 72 °C \hspace{1cm} 2:00
  \item GO TO 2, 34 times
  \item 72 °C \hspace{1cm} 10:00
  \item 4 °C \hspace{1cm} HOLD
\end{enumerate}
Thermal cycling conditions for pANK12 (3% DMSO, 67 ng gDNA, MiProof Polymerase):

1. 98 °C 0:30
2. 98 °C 0:10
3. 75 °C 0:30
4. GO TO 2, 34 times
5. 72 °C 10:00
6. 4 °C HOLD

Thermal cycling conditions for pAM2 (67 ng gDNA, Q5® High-Fidelity DNA Polymerase, 20% Q5 High GC Enhancer):

1. 98 °C 1:00
2. 98 °C 0:10
3. 90 °C (-0.5 °C/CYC) 0:30
4. 72 °C 1:00
5. GO TO 2, 29 times
6. 98 °C 0:10
7. 74 °C 0:30
8. 72 °C 1:00
9. GO TO 6, 19 times
10. 72 °C 10:00
11. 4 °C HOLD

Standard cloning primer list

Below are the primers used to amplify the E. coli optimized gene comD (Eurofins MWG) for the construction of plasmids encoding MBP-PCP7<sub>com</sub> and MBP-X<sub>com</sub>. Restriction sides are bolded (restriction enzyme written on the right of the primer sequence), while the sequence complementary to the insert is underlined.

Primers for plasmid encoding MBP-PCP7<sub>com</sub>:

For: TATTACCATGGGCAGTTGGGGGTCGTCGCCCCGGAATC <br>Rev: TATAACTCGAGCGGACGCTCGATCTGACGCAGTG <br>NcoI Xhol

Primers for plasmid encoding MBP-X<sub>com</sub>:

For: TATTACCATGGGCAGCTCGGCACTGCGTCAGTCAGTG <br>NcoI

-4-
LC/MS analysis of turnover mixtures

Following elution from SPE columns, the turnover mixtures were analyzed on an Agilent 6545 Q-TOF LC/MS. The column used was an Agilent Eclipse plus C18 with a pore size of 5 μm, and the flow rate was kept constant at 0.2 mL/min. Solvents used were: A = HPLC-grade H₂O incl. 0.1% formic acid (FA), B = HPLC-grade acetonitrile (ACN) incl. 0.1% FA. Gradient: 0 min 5% B; 30 min up to 40% B, 31 min up to 95% B, 31-43 min 95% B, 44 min down to 5% B, 44-56 min 5% B. The UV detector was set to 205 nm, and species detected in negative mode in the mass spectrometer.
**Fig. S1** SDS-PAGE gel and titers of all proteins used. Titers (to the nearest mg/L) are indicated beneath each lane. Expected molecular weights are indicated in parentheses following the name of the protein. Ladder used: Precision Plus Protein™ All Blue Standards (Bio-Rad)
Fig. S2 SEC interaction analyses of ComJ (left) and ComI (right) with $X_{\text{com}}$. A clear shift in retention volume is evident upon introduction of MBP-$X_{\text{com}}$ (bottom) compared to the P450 alone (top), indicating a tight interaction between the P450 and $X_{\text{com}}$ domain. The absorptions at $\lambda = 280$ nm (blue; protein-specific) and at $\lambda = 415$ nm (magenta; heme-specific) were monitored simultaneously. [P450] 33.3 $\mu$M, [X] 100 $\mu$M.
Fig. S3 SEC interaction analyses of ComJ (left) and ComI (right) with PCP7<sub>com</sub>. No significant shift in retention volume is evident upon introduction of MBP-PCP7<sub>com</sub> (bottom) compared to the P450 alone (top), indicating a lack of any strong interaction between the two proteins. The absorptions at $\lambda = 280$ nm (blue; protein-specific) and at $\lambda = 415$ nm (magenta; heme-specific) were monitored simultaneously. [P450] 33.3 $\mu$M, [PCP]: 100 $\mu$M.
Fig. S4 Turnover of D-Hpg-D-Hpg-L-Tyr (1) on PCP7-X$_{\text{van}}$ by OxyB$_{\text{van}}$. Left: UV spectrum (205 nm) showing the decrease in reactant concentration upon addition of OxyB$_{\text{van}}$ (grey dashed line) compared to the no-Oxy negative control (thick black line). Right: Extracted ion count (EIC) spectrum showing the increase in cyclized product ($m/z$ 531.2) upon addition of OxyB$_{\text{van}}$ (grey dashed line) compared to the no-Oxy negative control (thick black line).
Fig. S5 Turnover of 1 in the presence of ComJ and the following conditions: MBP-PCP7-X\textsubscript{com}, 1 hr (left, reproduced from Fig. 5E in the main text); MBP-PCP7-X\textsubscript{var}, 1 hr (center); MBP-PCP7-X\textsubscript{com}, 2 hr (right). For all three, the reduction system consisted of SpFd and EcoFlvR, with NADPH as the electron donor and G6P and G6P-DH to regenerate depleted NADPH.
Fig. S6 UV-vis spectrum of purified ComK. The maximum at ~422 nm is characteristic of ferredoxin [2Fe-2S] clusters.
Fig. S7 (part 1 of 2) $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-L-Tyr-OH (DMSO-D$_6$).
* Peak at 2.50 ppm is a residual solvent peak.

Fig. S7 (part 2 of 2) $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-l-Tyr-OH (DMSO-D$_6$).
* Peak at 7.95 ppm is a residual DMF peak.

**Fig. S8 (part 1 of 2)** $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-d-Tyr-OH (DMSO-D$_6$).
* Peak at 3.33 ppm is a residual water peak; peak at 2.50 ppm is a residual solvent peak.

**Fig. S8 (part 2 of 2)** $^1$H-NMR spectrum of Ac-NH-D-Hpg-D-Hpg-D-Tyr-OH (DMSO-D$_6$).
Fig. S9 (part 1 of 2) $^1$H-NMR spectrum of Ac-NH-$\delta$-Hpg-$\delta$-Hpg-$\alpha$-Tyr-SPh (DMSO-$D_6$).
* Peak at 3.33 ppm is a residual water peak; peak at 2.50 ppm is a residual solvent peak.

**Fig. S9 (part 2 of 2)** $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-L-Tyr-SPh (DMSO-D$_6$).
Fig. S10 (part 1 of 2) $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-d-Tyr-SPh (DMSO-D$_6$).
* Peak at 3.33 ppm is a residual water peak; peak at 2.50 ppm is a residual solvent peak.

Fig. S10 (part 2 of 2) $^1$H-NMR spectrum of Ac-NH-D-Hpg-D-Hpg-D-Tyr-SPh (DMSO-D$_6$).
Fig. S11 Partially assigned $^1$H-NMR spectrum of Ac-NH-d-Hpg-d-Hpg-L-Tyr-SCoA. Peaks “a” through “f” represent signature peaks that highlight the successful formation of the CoA conjugate.

<table>
<thead>
<tr>
<th></th>
<th>Pred. shift* (ppm)</th>
<th>Obs. shift (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>8.58</td>
<td>8.33</td>
</tr>
<tr>
<td>b</td>
<td>8.35</td>
<td>7.99</td>
</tr>
<tr>
<td>c</td>
<td>6.16</td>
<td>5.95</td>
</tr>
<tr>
<td>d</td>
<td>1.84</td>
<td>1.86</td>
</tr>
<tr>
<td>e</td>
<td>0.89</td>
<td>0.69</td>
</tr>
<tr>
<td>f</td>
<td>0.89</td>
<td>0.55</td>
</tr>
</tbody>
</table>

*ChemDraw $^1$H shift predictions in DMSO-D$_6$
Fig. S12 Partially assigned $^1$H-NMR spectrum of Ac-NH$_2$-Hpg-$\delta$-Hpg-$\delta$-Tyr-SCoA. Peaks “a” through “f” represent signature peaks that highlight the successful formation of the CoA conjugate.
Fig. S13 MALDI-TOF-MS spectrum of Ac-NH-d-Hpg-d-Hpg-l-Tyr-SCoA.
Fig. S14 MALDI-TOF-MS spectrum of Ac-NH-d-Hpg-d-Hpg-d-Tyr-SCoA.