Bunnak, W., Winter, A. J., Lazarus, C. M., Crump, M. P., Race, P. R., & Wattana-Amorn, P. (2020). SAXS reveals highly flexible interdomain linkers of tandem acyl carrier protein–thioesterase domains from a fungal non-reducing polyketide synthase. *FEBS Letters*.
<https://doi.org/10.1002/1873-3468.13954>

Peer reviewed version

Link to published version (if available):
[10.1002/1873-3468.13954](https://doi.org/10.1002/1873-3468.13954)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Electronic Supplementary Information

Solution characterization of tandem acyl carrier protein – thioesterase domains from a fungal non-reducing polyketide synthase

Waraporn Bunnak, Ashley J. Winter, Colin M. Lazarus, Matthew P. Crump, Paul R. Race,

Pakorn Wattana-Amorn

Table S1: Primers used for constructing the plasmids, pET28a-men2ACP₁ACP₂TE and pOPINF-men2ACP₂TE. Underlined nucleotides are regions of overlap with *E. coli*-yeast shuttle vectors, pEYA, and pOPINF.

Gene	Primer	Sequence 5'-3'
<i>men2</i> <i>ACP₁ACP₂TE</i>	<i>Fragment 1 (ACP₁ACP₂)</i> ACP ₁ ACP ₂ -F ACP ₁ ACP ₂ -R <i>Fragment 2 (TE)</i> TE-F TE-R	<u>CCAAC</u> TTTGTACAAAAAAGCAGGCTCATATGG GTGTCGTTGGGGATGAG GTCCAGGATGGCGGGAGGCGCCTGGACCGCCG GGGCGGGGGCGGCTACC GCGGTCCAGGCGCCTCCCGCCATCCTGGACGG TGATTTAGCCAGTTATC <u>ATGCCA</u> CTTTGTACAAGAAAGCTGGGTAAGC TTTTAGTTGTCCATTGCG
<i>men2</i> <i>ACP₂TE</i>	ACP ₂ TE-F ACP ₂ TE-R	<u>AAGTTCTGTTTCAGGG</u> CCCCGGACGGCCCCGGC ATCTACGC <u>ATGGTCTAGAAAGCTTTATC</u> AGTTGTCCATTGC GTCCTTTA

Table S2: SAXS data collection parameters and structural statistics for the Men2 ACP₁-ACP₂-TE and ACP₂-TE. Error values were computed in ScÅtter 3.1r.

Data collection parameters		
	Men2 ACP ₂ -TE	Men2 ACP ₁ -ACP ₂ -TE
Organism	<i>M. theobromae</i> BCC 4162	<i>M. theobromae</i> BCC 4162
Source	<i>E. coli</i>	<i>E. coli</i>
Molecular weight (Da)	45,253	57,940
$\bar{v}$ from chemical composition (cm ³ g ⁻¹)	0.739	0.736
Instrument	SEC-SAXS at B21 Diamond Light Source	
SEC Column	Superdex 200 Increase 3.2/300	
Loading concentration (mg mL ⁻¹)	15	6.9
Injection volume (μL)	45	45
Flow rate (mL min ⁻¹)	0.075	0.075
Average C in combined data frames (mg mL ⁻¹)	2.7	1.47
Solvent (solvent blanks taken from SEC flowthrough prior to elution of protein)	25 mM Tris, 150 mM NaCl, pH 8.0	
	Men2 ACP ₂ -TE	Men2 ACP ₁ -ACP ₂ -TE
Guinier analysis:		
$I(0)$ (cm ⁻¹)	1.54E-1 ± 2.5E-4	7.96E-2 ± 1.6E-4
R_g (Å)	35.59 ± 0.29	41.98 ± 0.45
q_{min} (Å ⁻¹)	0.015	0.020
qR_g max (q_{min} max)	1.29 (0.018)	1.29 (0.022)
Coefficient of correlation, R^2	0.9975	0.9964
$P(r)$ analysis		
$I(0)$ (cm ⁻¹)	1.51E-1 ± 3.1E-3	7.86E-2 ± 1.3E-3
R_g (Å)	39.83 ± 0.85	45.94 ± 0.80
D_{max} (Å)	144	172.5
q range (Å ⁻¹)	0.015 – 0.265	0.022 – 0.265
χ^2	1.07	1.27
Resolution (from SASRES) (Å)	39 ± 3	37 ± 3
DATCLASS designation	Extended	Extended
Molecular weight determination		
Bayesian molecular weight determination (Da) (95% Confidence interval)	43,775 (42,400-48,200)	56,850 (53,750-60,200)
M from Vc (ratio of M to expected) (Da)	38,390 (0.84)	48,987 (0.85)
M using the Fischer method (ratio of M to expected) (Da)	54,100 (1.19)	65,400 (1.12)
Shape model fitting results		
DAMMIF (default parameters, 23 calculations)		
q range for fitting (Å ⁻¹)	0.015 – 0.265	0.015 – 0.265
Symmetry, anisotropy assumptions	P1, none	P1, none
NSD (standard deviation)	0.611(0.040)	0.562 (0.019)
DAMMIN (default parameters, Slow run)		
q range for fitting (Å ⁻¹)	0.015 – 0.265	0.015 – 0.265
Symmetry, anisotropy assumptions	P1, none	P1, none
χ^2	1.416	1.440

Atomistic Modelling	Men2 ACP ₂ -TE	Men2 ACP ₁ -ACP ₂ -TE
Structure	Homology model (<i>apo</i> form)	Homology model (<i>apo</i> form)
<i>q</i> range for all modelling	0.015 – 0.34	0.022 – 0.34
Method	EOM	EOM
Symmetry assumptions	P1	P1
CA distribution	Random	Random
χ^2	1.18	2.83
Ensemble Rg/Dmax values for multi-state models (Å)	38.04/117.07	46.44/138.74
Model 1 Rg/Dmax (Å) [%]	26.27/82.45 [20]	42.58/132.99 [62.5]
Model 2 Rg/Dmax (Å) [%]	30.67/96.17 [10]	38.98/119.03 [12.5]
Model 3 Rg/Dmax (Å) [%]	42.67/130.34 [10]	59.80/162.99 [25]
Model 4 Rg/Dmax (Å) [%]	30.85/101.01 [10]	
Model 5 Rg/Dmax (Å) [%]	51.64/157.39 [20]	
Model 6 Rg/Dmax (Å) [%]	40.14/121.17 [30]	
Structure	Homology model (<i>holo</i> form)	Homology model (<i>holo</i> form)
<i>q</i> range for all modelling	0.015 – 0.34	0.022 – 0.34
Method	EOM	EOM
Symmetry assumptions	P1	P1
CA distribution	Random	Random
χ^2	1.23	1.61
Ensemble Rg/Dmax values for multi-state models (Å)	37.85/114.58	45.45/140.84
Model 1 Rg/Dmax (Å) [%]	24.51/80.17 [22.2]	29.76/101.25 [11.1]
Model 2 Rg/Dmax (Å) [%]	45.60/131.89 [22.2]	57.82/168.61 [11.1]
Model 3 Rg/Dmax (Å) [%]	56.71/165.22 [11.1]	66.13/198.86 [11.1]
Model 4 Rg/Dmax (Å) [%]	38.10/114.50 [11.1]	42.55/133.14 [66.7]
Model 5 Rg/Dmax (Å) [%]	31.30/100.01 [22]	
Model 6 Rg/Dmax (Å) [%]	43.04/127.33 [11.1]	
Structure	Combined EOM pool of <i>apo</i> and <i>holo</i> ACP ₂ -TE	Combined EOM pool of <i>apo</i> and <i>holo</i> ACP ₁ -ACP ₂ -TE
<i>q</i> range for all modelling	0.015 – 0.34	0.022 – 0.34
Method	EOM	EOM
Symmetry assumptions	P1	P1
CA distribution	Random	Random
Pool percentage (%) (Apo/Holo)	67/33	46/54
χ^2	1.15	2.42
Ensemble Rg/Dmax values for multi-state models (Å)	37.84/115.51	47.16/140.88
Model 1 Rg/Dmax (Å) [%]	38.42/108.91 [9.1]	34.31/110.41 [12.5]
Model 2 Rg/Dmax (Å) [%]	37.60/115.19 [9.1]	47.31/136.34 [12.5]
Model 3 Rg/Dmax (Å) [%]	28.62/91.74 [9.1]	53.41/159.28 [12.5]
Model 4 Rg/Dmax (Å) [%]	28.23/90.35 [9.1]	59.80/162.99 [12.5]
Model 5 Rg/Dmax (Å) [%]	31.57/102.58 [9.1]	38.23/110.62 [12.5]
Model 6 Rg/Dmax (Å) [%]	30.17/106.07 [9.1]	39.61/122.30 [12.5]
Model 7 Rg/Dmax (Å) [%]	26.61/93.64 [9.1]	66.13/198.86 [12.5]
Model 8 Rg/Dmax (Å) [%]	57.92/168.33 [9.1]	38.51/126.21 [12.5]
Model 9 Rg/Dmax (Å) [%]	45.92/131.37 [9.1]	
Model 10 Rg/Dmax (Å) [%]	48.61/140.26 [9.1]	
Model 11 Rg/Dmax (Å) [%]	42.62/121.92 [9.1]	

SAXS data reduction	Solvent subtraction and $I(q)$ versus q using ScÅtter 3.1r	
Extinction coefficient estimate	Protparam [1]	
Basic analyses: Guinier, $P(r)$, V_p	ScÅtter 3.1r	
Shape/Bead modelling	DAMMIF and DAMMIN via ATSAS 2.7.2	
Atomistic structure modelling	EOM (ATSAS)	
Molecular weight approximation	ATSAS [2] and SAXSMoW 2.0	
Resolution calculation	SASRES (ATSAS)	
Three-dimensional graphic model representations	Pymol	
SASBDB accession code	SASDJ43	SASDJ53

GVVGDEKADAAIGKILARAGANPAHITDATTFAEIGFD^{*}LEWIELVREIRTSLD
 LEVPASFFFEYPKVNGLRRAISELSLDYQGPASGSVSVSSSATTHGMTTPSSTS
SAQSSQSSQTPDGPGIYANAVIDIVLSQTGFDKADLLPTTRFDDMGLD^{*}SLCTM
 EVVGVVREQTGLDLPASFFHQNPVAVHVRRALGSDSDGDSKPKSAPAPPAPEP
VVEVAAPAPAVQAPPAILDGDLASYHCDFFLMQGSSDSTKTPLFFLPDGTGYPA
 VLLKLPPIFEGDNALFTCKSPLLNVAGREVRCTIEGLAAAYAAAIQRARPHGP
 YLLAGYSLGGAYAFEVAKILADAGEVVQGLLFVDFNMAASVGKLHRDRNPV
 PVDLTVGAMEKTGWMQGIQNDDKDFNIPPAPPKIKFHALSVFKSLISYYPTPM
 TPSQRPRNTYALWAGIGMQDLLGTKNAGFLPAYGIIDWQMGRHENNGPAG
 WEEYIGGPVKCATMPCDHLSELLMSHHWIPKSAEVIKGLLKDAMDN

Fig. S1 Amino acid sequence (G1664-N2186) of ACP₁-ACP₂-TE of Men2 (GenPept accession number: CAB3277416) with domain boundaries based on I-TASSER homology models of each domain with highest C-score and prediction of secondary structures of amino acid sequence performed on the XtalPred server. Individual ACPs with the active sites serines marked with stars and the TE domain are highlighted in red and blue, respectively. The flexible regions (S1739-P1786) and (G1858-F1907) used for EOM are underlined in black.

Fig. S2 SDS-PAGE (4-20%) analysis showing expression of Men2 TE, ACP₂-TE and ACP₁-ACP₂-TE. (lane M = protein ladder, lane 1 = pre-induction, lane 2 = post-induction, lane 3 = soluble fraction and 4 = insoluble fraction)

Fig. S3 Analysis of the oligomeric states of Men2 ACP₂-TE (blue) and ACP₁-ACP₂-TE (pink) using size-exclusion chromatography, monitoring the absorbance of the column eluate at 280 nm throughout. Elution volumes are consistent with both Men2 ACP₂-TE and Men2 ACP₁-ACP₂-TE being monomeric in solution. Separations were performed using a HiLoadTM 16/600 SuperdexTM 200 pg column.

Constructs	Helix (%)			β -strand (%)		
	DichroWeb	BestSel	Expected	DichroWeb	BestSel	Expected
TE	6.5	13.4	31.8	42.0	27.5	11.8
ACP ₂ -TE	6.4	14.5	35.4	42.6	31.3	8.6
ACP ₁ -ACP ₂ -TE	6.4	15.0	37.2	42.7	28.9	6.6

Fig. S4 CD spectra [with a table showing calculated secondary structure contents](#) of Men2 TE, ACP₂-TE and ACP₁-ACP₂-TE [calculated by using DichroWeb and BestSel](#). [The expected percentages were calculated via the homology model.](#)

Fig. S5 Charge envelope and deconvoluted ESI-MS spectra of A) purified Men2 ACP₂-TE with the N-terminal His₆ containing a mixture of both *apo* and *holo* forms and B) its conversion to *holo* forms catalyzed by Sfp. Expected molecular weights of *apo* and *holo* Men2 ACP₂-TE are 45,253 and 45,593 Da, respectively.

Fig. S6 Charge envelope and deconvoluted ESI-MS spectrum of purified Men2 ACP₁-ACP₂-TE with the N-terminal His₆. Expected molecular weights of *apo* and *holo* Men2 ACP₁-ACP₂-TE are 57,940 and 58,280 Da (*apo* form + 340 Da), respectively. The *holo* form is only found in one ACP domain and it is expected to be ACP₂ as *holo* conversion in *E. coli* could also be observed with Men2 ACP₂-TE. Gluconoylated forms of both *apo* and *holo* Men2 ACP₁-ACP₂-TE were also observed with expected molecular weights of 58,118 (*apo* form + 178 Da) and 58,458 Da (*holo* form + 178 Da), respectively.

Fig. S7 SEC-SAXS for Men2 ACP₂-TE (Top) and ACP₁-ACP₂-TE (Bottom) showing the intensity of scattering as a ratio to the buffer background vs frame number (orange circles). The R_g across the peak is highlighted in blue. Noted, for the ACP₁-ACP₂-TE, the frame corresponding to $R_g > 50$ Å was not included in the analysis.

A

B

Fig. S8 A) $I(q)$ vs q as log-linear plots with errors shown. B) Guinier fits for $qR_g < 1.3$ for Men2 ACP₂-TE (blue) and ACP₁-ACP₂-TE (black). The minimum q value measured (ACP₁-ACP₂-TE = 0.0157 Å⁻¹ and ACP₂-TE = 0.0201 Å⁻¹) is below the minimum of $q \approx \pi/D_{\max}$: (ACP₁-ACP₂-TE = 0.0182 and ACP₂-TE = 0.0218) recommended for accurate assessment of the scattered particle.

A **B**

Fig. S9 Examples of *ab initio* and final averaged envelopes of Men2 A) ACP₂-TE and B) ACP₁-ACP₂-TE. Figures A and B are on a different scale.

Fig. S10 Ensemble modelling of Men2 A) *apo* and B) *holo* ACP₂-TE with histogram profiles of the R_g and D_{max} distributions of the resultant ensemble models (red) compared to the pool of 10,000 independent models (blue). For each ensemble sub-population, representative models of C) *apo* and D) *holo* ACP₂-TE are shown, with the linker region between ACP (red) and TE (blue) domains represented as grey spheres. The phosphopantetheinyl group (green) of *holo* form is shown in stick representation.

Fig. S11 Ensemble modelling of Men2 A) *apo* and B) *holo* ACP₁-ACP₂-TE with histogram profiles of the R_g and D_{max} distributions of the resultant ensemble models (red) compared to the pool of 10,000 independent models (blue). For each ensemble sub-population, representative models of C) *apo* and D) *holo* ACP₁-ACP₂-TE are shown, with the linker region between ACPs (red) and TE (blue) domains represented as grey spheres. The phosphopantetheinyl group (green) attached to the active-site serine of ACP₂ is shown in stick representation.

References

1. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, *John M Walker (Ed), The Proteomics Protocols Handbook, Humana Press*, 571-607.
2. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V. & Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis, *J Appl Crystallogr.* **45**, 342-350.