STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 76 (2020)

Supporting information for article:

Structural characterization of three noncanonical NTF2-like superfamily proteins: implications for polyketide biosynthesis

Nemanja Vuksanovic, Xuechen Zhu, Dante A. Serrano, Vilja Siitonen, Mikko Metsä-Ketelä, Charles E. Melançon and Nicholas R. Silvaggi

S1. Additional Materials and Methods

S1.1. Sequences of synthesized gBlocks

S1.1.1. Synthesized ActVI-ORFA coding sequence (5'-3'):

GGTCTCAAGGT*ATGACGATTACCGCGCTGCCAACTGGACTTTATGCTGAAGTGTTGAGCTTTTATGGTCA*
TCAGATGCAGAAGTTGGATGGGCGCGATTTTGCCGGGTATGCCGCCACCTTCACCGAAGATGGTGAATT
CCGTCACAGCCCTTCACTTCCGGCGGCGCATACCCGTGCGGGCATTACCGCCGTGCTGGAAGATTTCCA
TCGCAAGTTCGACGCACGCAAAATCCAGCGCCGCCATTGGTTTGACCATACTGCCCTGTCTCAAGCGAGT
GATGGTTCTATTACCGCCACCAGTTACTGCTTGGTGCTGACGGTACACGCGGATGTGAAAGCGCCGGAG
TTCGGGCCAAGTTGCTTAGTGCATGATGTTTTGGTGCGCGGTGCGGATGGTGAATTATTACTGCGCTCCC
*GCCATGTCACACATGACCATGTCTTCCCAGCCTGA*TGATCTAGAGC

S1.1.2. Synthesized Aln2 coding sequence (5'-3'):

GG*ATGACTACCGATGAAACTACCACCACAGATGCAACTACGATTACTGATGCGACAACGATTGCCGATGC*
AACCACGCGGAACGCCCCTAAATTGCCCTCTCCAGAGTTGTACGTCGAAGTCACTCAGTTTTATGCTCGT
CAAATGCACCGTATGGACGGTGATGATTTCGGAGGGTTTGCGGCGACTTTCGTAGCCGGTGCCGAATTTA
GATTGGCAGGCGGTACTGTTCTGACTGGCCCAGAGGCCATAGAAGCTGGTGCGAGAGCGGCAGCAGGA
CGCTTCGACGGGGCCCAGCCCCGGCACTGGTTTGACATGATGACTGTAGAGGAGGCGGACGACGGAAC
GGTGTCGACTTCTTATTACGCGACTGTGACTGTTACGTCCGCCCAAGGGGCTGTCCTTGTAGAACCTACG
TGCTTTGTGCGGGACACTTTAGTTCGTGTCTCCGGAGTACTTAGATCCCGTTCTCGTGTAATTGAAAGAGA
*CGACTTAGTTGTACGTGCTCGGACTCAAGGTTGA*AA

S1.2. Size exclusion chromatography of ActVI-ORFA

The molecular weight of ActVI-ORFA was estimated by gel filtration chromatography using a Superdex S200 10/300 GL (10 x 300 mm) (GE Healthcare Life Sciences, Chicago, USA). The calibration curve was constructed using bovine catalase (233.1 kDa), alcohol dehydrogenase from *Saccharomyces cerevisiae* (149.5 kDa), bovine serum albumin (66.5 kDa) and hen egg lysozyme (14.4 kDa). The standards were run at 0.5 mL/min using 20 mM TRIS pH 8.0, 150mM NaCl as a mobile phase. The ActVI-ORFA sample (10 mg/mL) was run under the same conditions.

S1.3. Characterization of (*S*)-DNPA

The compound isolated from M1152::pXZ11 R4 crude extract was characterized using high resolution mass spectrometry (HRMS), NMR and circular dichroism (CD) spectrum analysis. The HRMS (ESI) data (Figure S2) were collected in positive ionization mode in methanol by the Waters GCT (2008) high resolution mass spectrometer facility at University of California at Riverside. The NMR data (Figure S3 and Table S1) were collected in DMSO-*d*6 on both a 500 MHz NMR instrument equipped with a liquid nitrogen-cooled Prodigy BBO CryoProbe (HSQCDE) and a 600 MHz NMR instrument equipped with a liquid nitrogen-cooled Prodigy TCI inverted CryoProbe (^{1}H, ^{13}C and HMBC). Chemical shifts were internally referenced to tetramethylsilane or to the solvent signals. The NMR data were analyzed with Topspin (Bruker Biospin). The CD spectrum analysis was performed in methanol on a Chirascan circular dichroism (ECD) spectrometer equipped with a 1 mm path-length cell at 23 °C. The NMR data and CD spectra were compared to previous reports (Ichinose et al., 1999; Taguchi et al., 2001) to confirm the structure of (*S*)-DNPA.

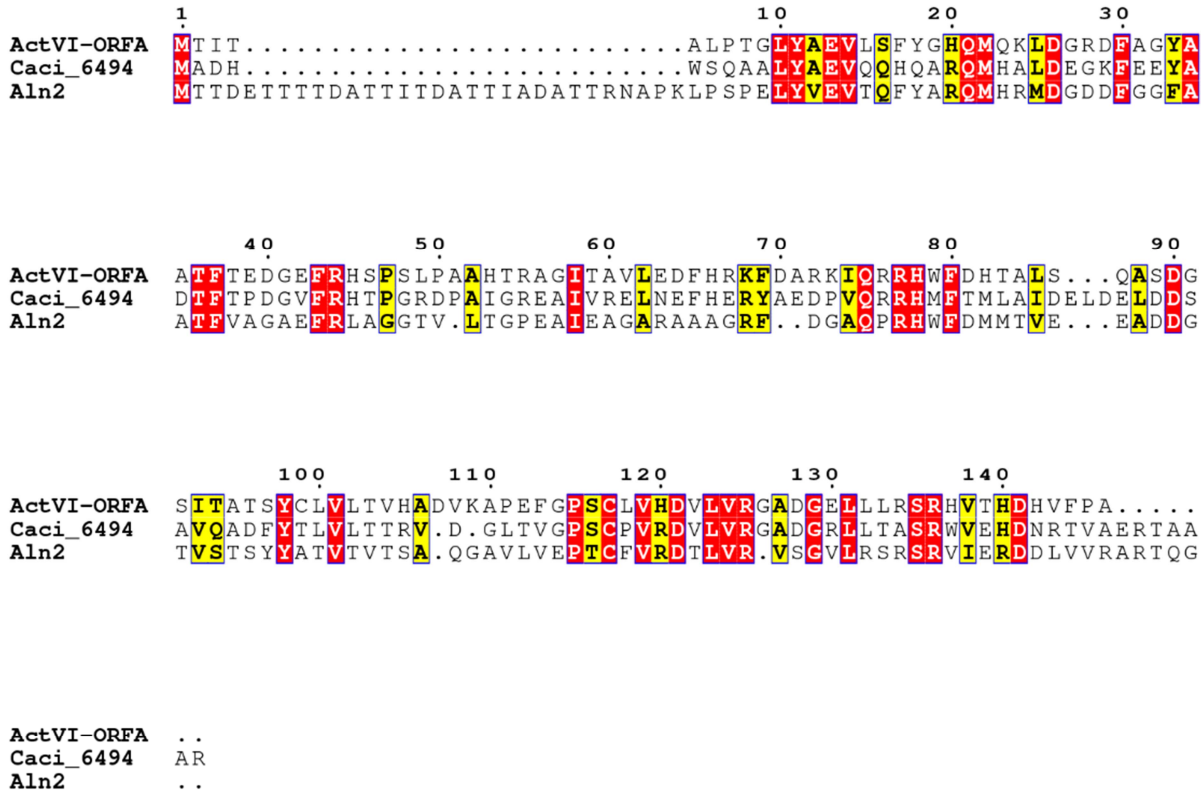

Figure S1 Sequence alignment of ActVI-ORFA, Caci_6494, and Aln2. The H78/D26 dyad is conserved in all of them. The figure was produced in ESPript (Robert & Gouet, 2014).

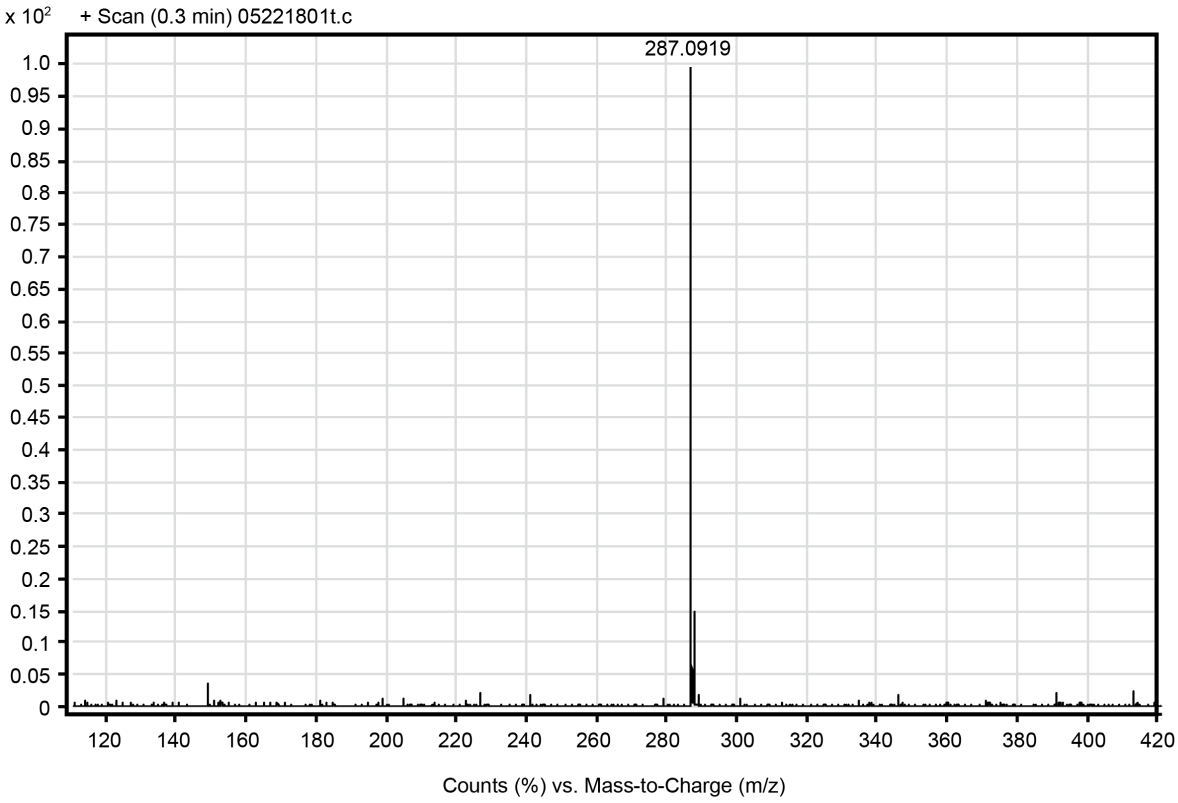

Figure S2 High-resolution mass spectrum of isolated (*S*)-DNPA peak in methanol using electrospray ionization in positive ionization mode by the Waters GCT (2008) high resolution mass spectrometer facility at University of California at Riverside. HRMS (ESI) m/z calculated for (*S*)-DNPA $[M+H]^+$: 287.0920, observed: 287.0919.

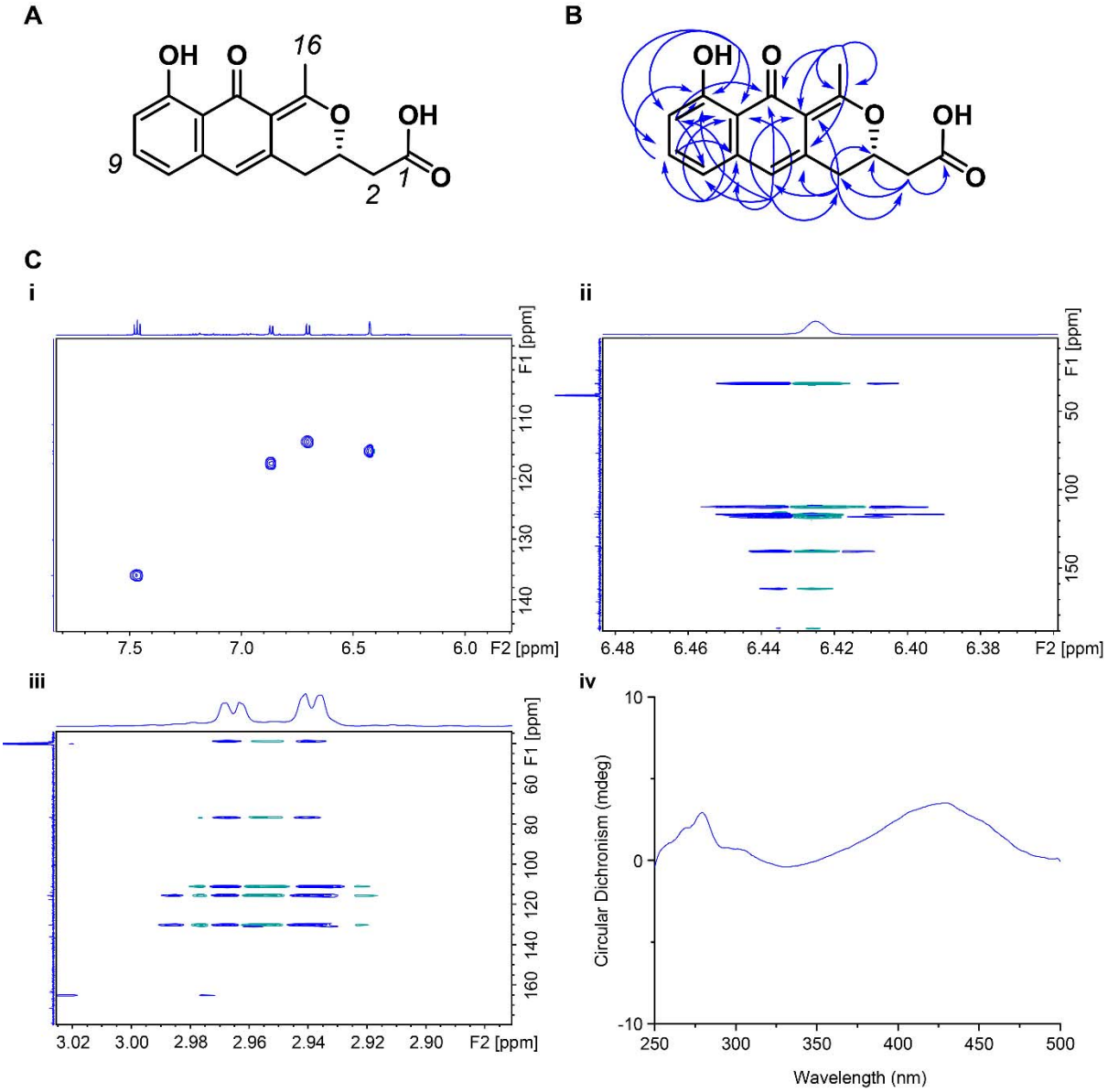

Figure S3 Characterization of (*S*)-DNPA isolated from M1152::pXZ11 R4 crude extract. (A) Molecular structure of (*S*)-DNPA (compound **5** in the article) showing atom numbering (Ichinose et al., 1999). (B) HMBC correlations. (C) Relevant features of the compound shown in the selected spectra: i) HSQCDE signals for the aromatic protons; ii) HMBC signals from H6 (δ 6.43) to C4, C7, C8, C11, C12, C13 and C14 showing the major part of the compound; iii) HMBC signals from H4 (δ 2.95) to C2, C3, C5, C6 and C14 showing that the acetic acid is attached to the backbone; iv) CD spectrum collected in methanol from 250 nm to 500 nm, indicating the S configuration (Taguchi et al., 2001).

Table S1 Chemical shifts and J-couplings observed in ^{1}H and ^{13}C NMR data collected on 600 MHz (1H, 13C and HMBC) and 500 MHz (HSQCDE) instruments

Abbreviations used in NMR data annotation: s for singlet, m for multiplet, dd for doublet of doublets, ddd for doublet of doublet of doublets and br for broad. The integrals are ^{1}H unless otherwise stated.

	δ ppm	δ ppm, *J* Hz
Position	^{13}C	^{1}H
1	171.5	
1-OH		12.61 s
2	38.9	2.78 m*, 2H
3	76.7	4.73 m
		2.95 ddd, 0.8; 3.4; 15.9
4	32.4	2.80 m*
5	130.1	
6	115.4	6.43 br
7	139.3	
8	117.5	6.87 dd, 0.9; 7.7
9	136.0	7.47 dd, 7.7; 8.3
10	113.9	6.70 dd, 0.9; 8.3
11	163.2	
11-OH		14.05 br
12	115.9	
13	188.5	
14	111.0	
15	178.9	
16	23.8	2.58 s, 3H

References

Gomez-Escribano, J. P., and Bibb, M. J. (2011). *Microb. Biotechnol.* **4**, 207-215.

Ichinose, K., Surti, C., Taguchi, T., Malpartida, F., Booker-Milburm, K. I., Stephenson, G. R., Ebizuka, Y., and Hopwood, D. A. (1999). *Bioorg. Med. Chem. Lett.* **9**, 395-400.

Taguchi, T., Ebizuka, Y., Hopwood, D. A., and Ichinose, K. (2001). *J. Am. Chem. Soc.* **123**, 11376-11380.