

**Chemistry & Biology, Volume 22**

## **Supplemental Information**

### **Substrate Flexibility of a Mutated Acyltransferase Domain and Implications for Polyketide Biosynthesis**

**Kenny Bravo-Rodriguez, Stephan Klopries, Kyra R.M. Koopmans, Uschi Sundermann, Samir Yahiaoui, Julia Arens, Susanna Kushnir, Frank Schulz, and Elsa Sanchez-Garcia**

## **Supplementary Information**

### **I. Computational Section (pages 3-12)**

### **II. Feeding Experiments**

#### **II.1 Bacterial strains and media (page 12)**

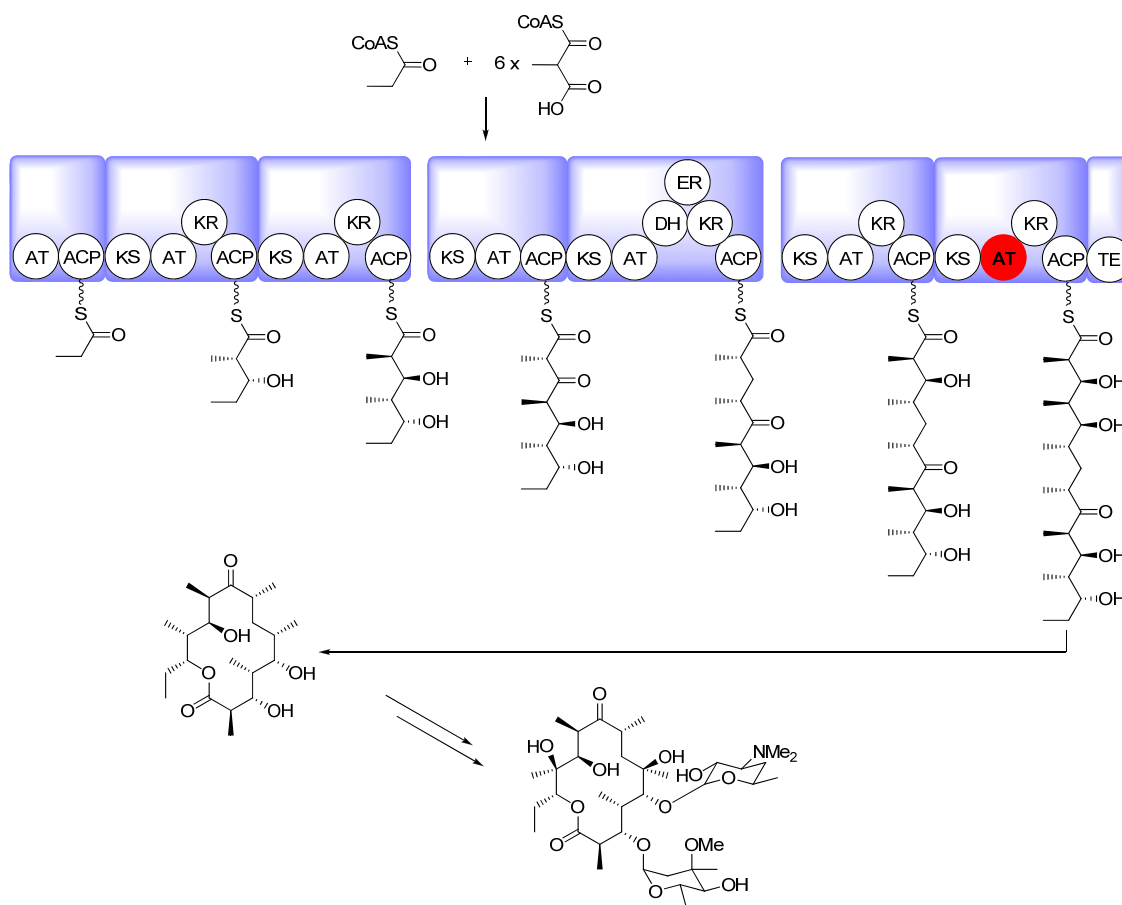
#### **II.2 Fermentation experiments with *Saccharopolyspora erythraea* (page 12-14.)**

### **III. General information (page 15)**

### **IV. References (page 16)**

## Supplementary Information

**SI Figure 1, related to Fig. 2: Overview of the biosynthetic pathway towards Erythromycin A.** In a cascade of Claisen condensations, the polyketide is assembled from propionyl-CoA as starter unit and six equivalents of methylmalonyl-CoA. The resulting  $\beta$ -ketothioester is in optional steps reduced by further catalytic domains. The primary product 6-desoxyerythronolide B is cyclized and cleaved from the PKS multienzyme complex before being furnished with oxygenations and glycosylations to yield erythromycin A. The acyltransferase domain in module 6 which was investigated in this study is highlighted in red.

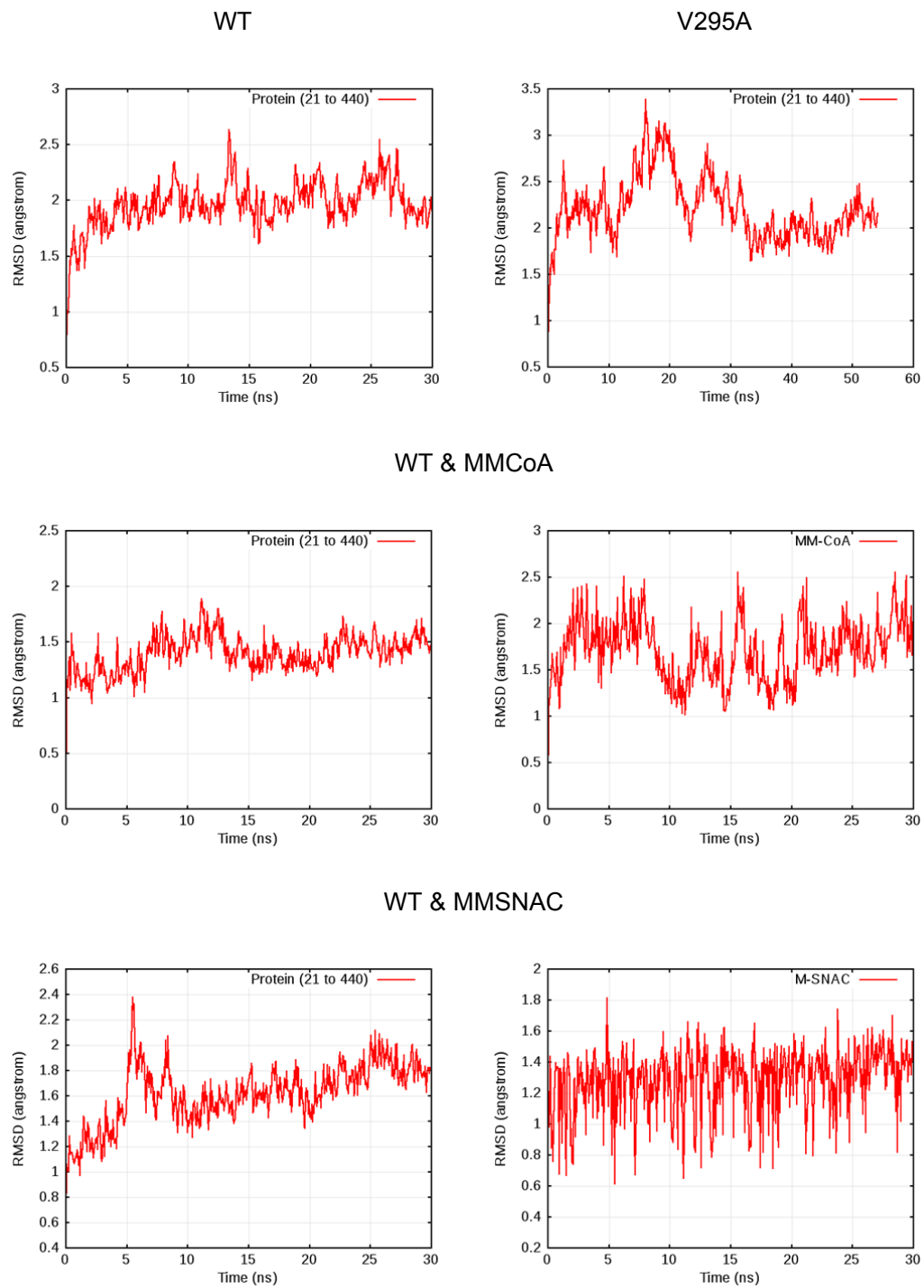


## I. Computational Section

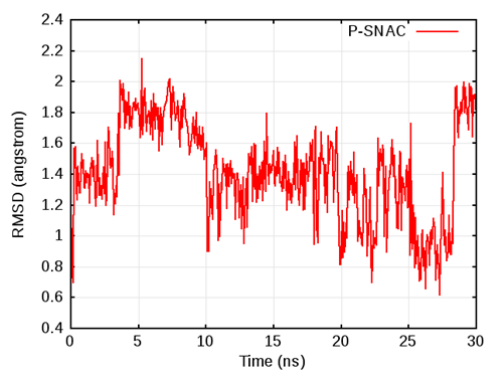
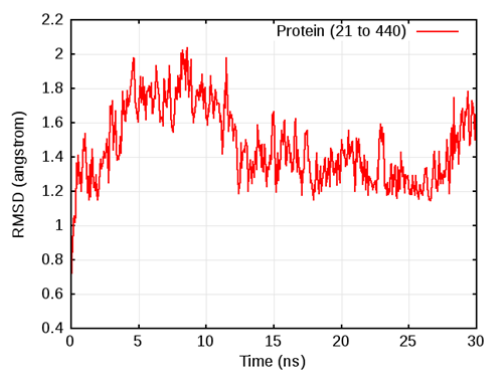
After building the models, the CHARMM31b1 program (Mackerell, et al., 2004) was employed for energy-minimisations consisting of 20 steps using the steepest descent algorithm followed by 1000 steps using the adopted basis Newton-Raphson algorithm as implemented in CHARMM. The VMD1.9 program (Humphrey, et al., 1996) was employed to neutralise the system and to solvate it in a box of water with 12 Å distance between the protein and the walls of the solvent box. As usual, short MD simulations were performed for equilibration before running unconstrained production MDs for 30 ns (NPT ensemble). For AT6-V295A, AT6-V295A•••allyl-MSNAC, AT6-V295A•••Ph-MSNAC and AT6-V295A•••Hex-MSNAC, the simulations times were extended to 50 ns.

All MD simulations were performed under the same conditions using the NAMD2.9 program (Phillips, et al., 2005). The temperature was set to 300 K and a timestep of 2 fs was used. The parameters of the substrates were generated using Swissparam (Zoete, et al., 2011), which suitability for modelling MCoA building blocks has been tested by us (Sundermann, et al., 2013). The CHARMM22 force field was employed (Mackerell, et al., 2004). RMSD values were calculated for the backbone atoms of residues 21 to 440 of the protein and for all atoms of the substrate. Since residues 1 to 20 belong to a linker domain with no defined secondary structure and move randomly along the entire simulations, they were not included in the RMSD calculations. The VMD1.9 program was also used for visualisation and analysis of the results.

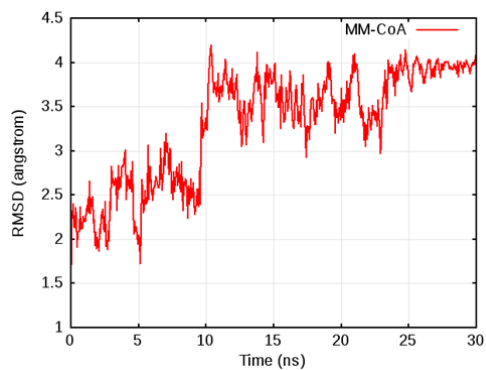
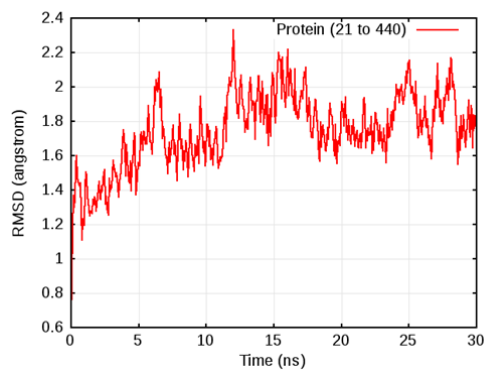
**SI Figure 2, related to Figure 1: A.** RMSD for the backbone atoms of residues 21 to 440 of the proteins and RMSD for all atoms of the different substrates. All RMSD are in Å.



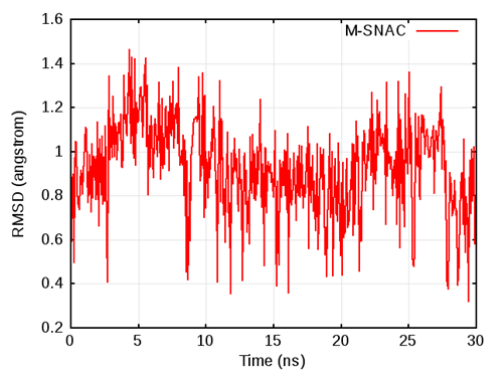
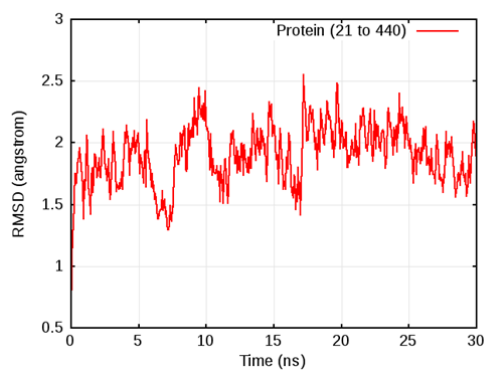
### WT & Propargyl-MSNAC



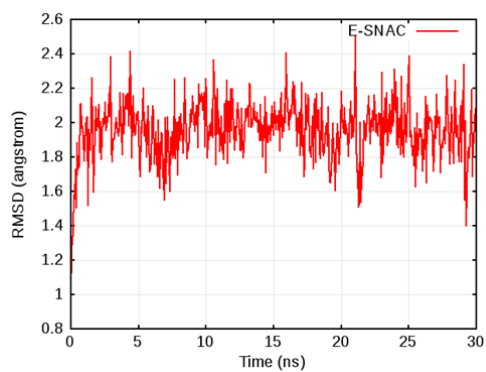
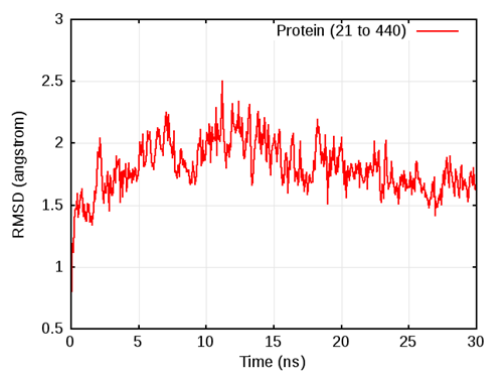
### V295A & MM-CoA



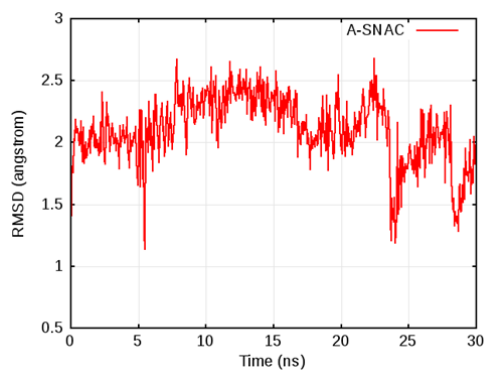
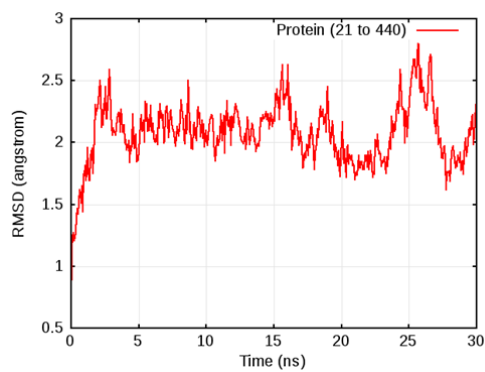
### V295A & MMSNAC



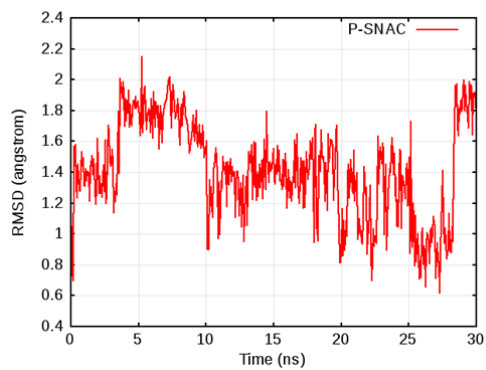
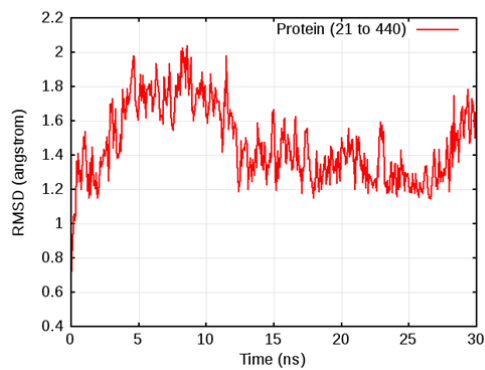
### V295A & Et-MSNAC



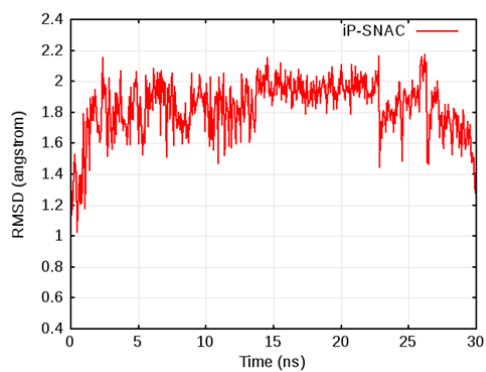
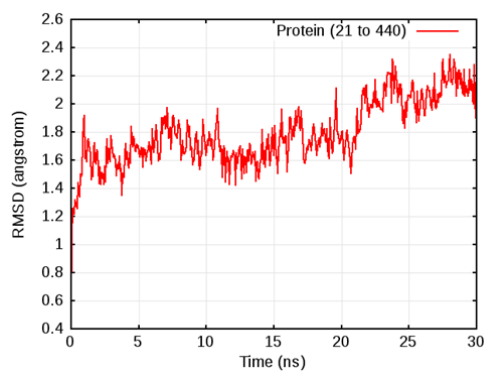
### V295A & allyl-MSNAC



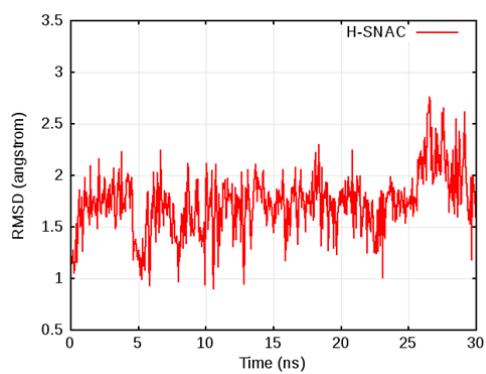
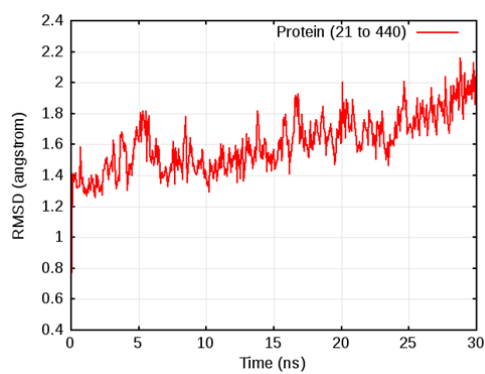
### WT & Propargyl-MSNAC



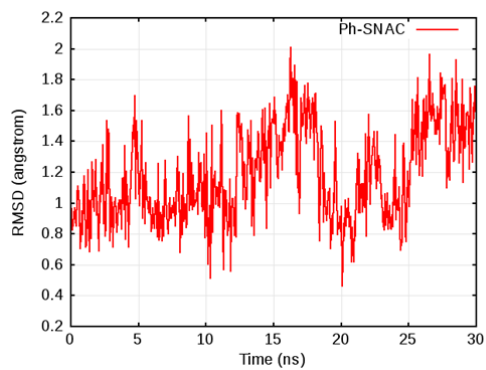
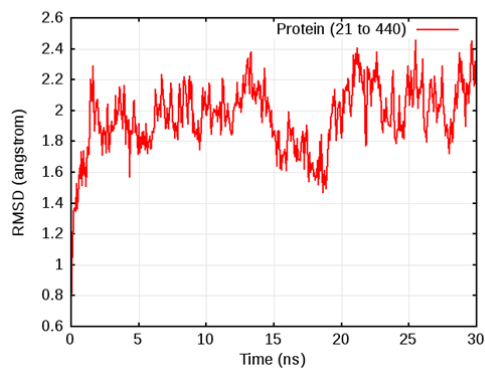
### V295A & *i*Prop-SNAC



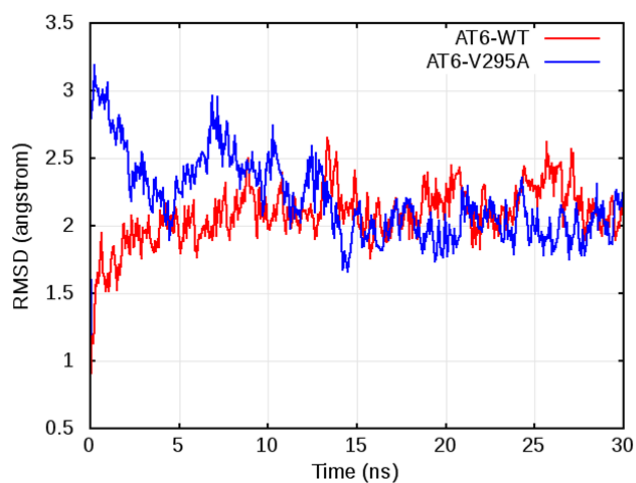
### V295A & Hex-MSNAC



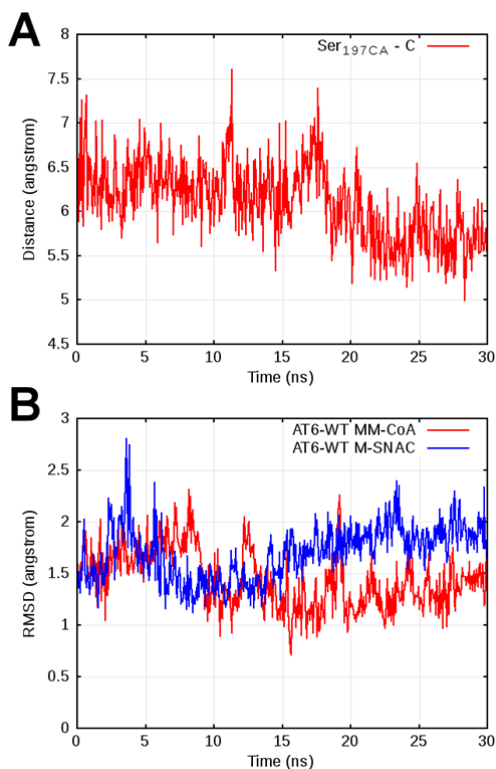
### V295A & Ph-MSNAC



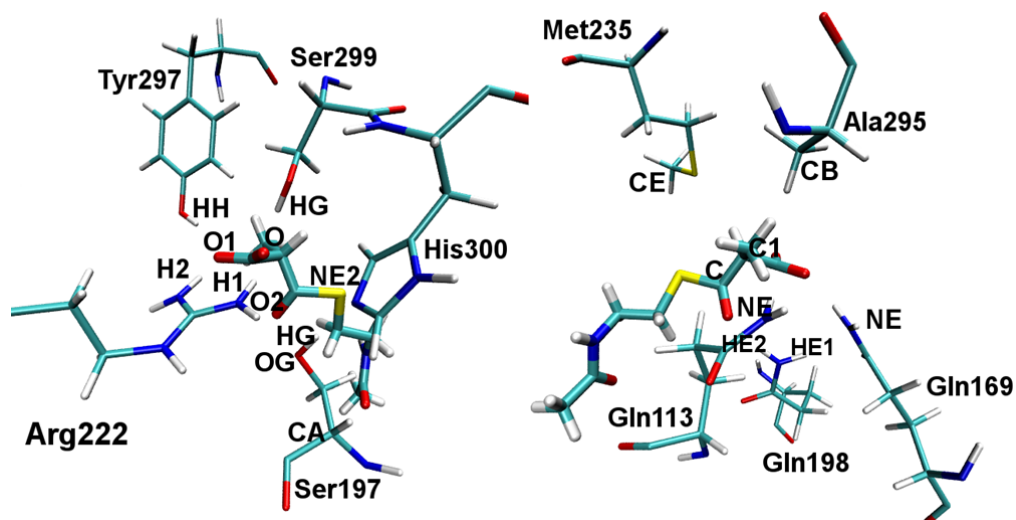
**SI Figure 2 B.** RMSDs of the MD simulations of AT6 and AT6-V295A taking as a reference the starting structure of the MD simulation of AT6. Only the backbone atoms of residues 21 to 440 were considered.



**SI Figure 2 C.** (A) Behavior of the distance between Ser197<sub>CA</sub> and the C atom of MMSNAC for the MD simulation of WT-AT6-MMSNAC. (B) RMSD for the common part of MMCoA and MMSNAC taking as reference the starting structure of WT-AT6-MM-CoA.



**SI Figure 2 D.** Atoms names of the substrates and selected amino acids of the protein. C1 is always the atom in the side chain of the substrate directly bonded to the malonyl moiety. In the wild type AT6, CB refers to the tertiary carbon atom of the lateral chain of Val.



**Table S1, Related to Figure 1. A.** Relevant distances of MMCoA, MMSNAC and propargyl-MSNAC in WT-AT6 (30 ns MD simulations)

Distance (Å)	Wild Type		
	MMCoA	MMSNAC	Propargyl-MSNAC
Thioester Region (Carboxylate Group)			
S299(HG)-O	1.76 ± 0.14	1.78 ± 0.15	2.48 ± 0.83
R222(H1)-O	1.87 ± 0.21	1.86 ± 0.22	N/A
R222(H2)-O	1.88 ± 0.21	1.91 ± 0.22	1.69 ± 0.19
Q198(HE1)-O	3.68 ± 0.34	3.73 ± 0.35	3.06 ± 0.62
Q198(HE1)-O1	3.36 ± 0.34	3.37 ± 0.38	3.75 ± 0.38
Q169(NE)-O1	3.37 ± 0.52	3.98 ± 0.69	3.37 ± 0.45
Q113(NE)-O1	4.36 ± 0.89	4.28 ± 0.84	5.03 ± 1.26
Y297(HH)-O1	1.72 ± 0.14	1.72 ± 0.12	1.66 ± 0.10
Thioester Region (Methyl Group)			
M235(CE)-C1	4.57 ± 0.54	4.46 ± 0.53	6.35 ± 0.70
V295(CB)-C1	4.89 ± 0.44	4.95 ± 0.47	8.24 ± 0.57
Thioester Region (Thioester Group)			
Q198(HE2)-O2	1.98 ± 0.23	2.10 ± 0.30	2.44 ± 0.76
S197(HG)-O2	3.53 ± 0.57	3.18 ± 0.69	3.84 ± 1.11
S197(HG)-H300(NE2)	2.13 ± 0.36	2.26 ± 0.42	2.38 ± 0.53
S197(OG)-C	3.69 ± 0.34	4.10 ± 0.46	4.13 ± 0.51
S197(Cα)-C	5.86 ± 0.32	6.20 ± 0.41	6.07 ± 0.40

**Table S1, B.** Relevant distances from the MD simulations of MMCoA, MMSNAC, Et-MSNAC, allyl-MSNAC, prop-MSNAC, *i*Prop-MSNAC, propargyl-MSNAC, Hex-MSNAC and Ph-MSNAC in the V295A AT6 variant.

Distance (Å)	V295A							
	MM-CoA	MMSNAC	Et-MSNAC	allyl-MSNAC	<i>i</i> Prop-MSNAC	propargyl-MSNAC	Hex-MSNAC	Ph-MSNAC
Thioester Region (Carboxylate Group)								
S299(HG)-O	1.83 ± 0.26	1.94 ± 0.53	1.74 ± 0.13	1.76 ± 0.14	1.73 ± 0.12	1.78 ± 0.15	1.75 ± 0.15	1.76 ± 0.14
R222(H1)-O	2.10 ± 0.36	1.79 ± 0.19	1.75 ± 0.15	1.75 ± 0.16	1.76 ± 0.15	1.83 ± 0.20	2.34 ± 1.11	1.83 ± 0.22
R222(H2)-O	1.86 ± 0.28	2.04 ± 0.32	2.16 ± 0.27	2.28 ± 0.29	2.19 ± 0.29	2.05 ± 0.30	2.25 ± 0.75	2.08 ± 0.28
Q198(HE1)-O	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Q198(HE1)-O1	N/A	3.20 ± 0.30	N/A	N/A	N/A	N/A	N/A	N/A
Q169(NE)-O1	N/A	N/A	3.93 ± 0.70	N/A	3.83 ± 0.63	N/A	N/A	3.82 ± 0.73
Q113(NE)-O1	N/A	4.01 ± 1.02	N/A	N/A	N/A	N/A	N/A	N/A
Y297(HH)-O1	1.91 ± 0.69	1.76 ± 0.13	1.74 ± 0.14	1.77 ± 0.14	N/A	1.75 ± 0.29	1.80 ± 0.33	1.99 ± 0.68
Thioester Region (Methyl Group)								
M235(CE)-C1	4.26 ± 0.56	3.99 ± 0.40	5.05 ± 0.44	5.18 ± 0.58	5.93 ± 0.71	4.76 ± 0.80	5.26 ± 0.54	4.61 ± 0.51
A295(CB)-C1	6.40 ± 0.96	5.46 ± 0.93	5.14 ± 0.63	5.41 ± 0.91	8.27 ± 0.69	5.38 ± 1.07	6.86 ± 0.93	6.55 ± 0.96
Thioester Region (Thioester Group)								
Q198(HE2)-O2	N/A	3.85 ± 0.56	N/A	N/A	N/A	N/A	N/A	N/A
S197(HG)-O2	3.69 ± 0.53	3.82 ± 0.29	3.30 ± 0.71	3.68 ± 1.05	4.06 ± 0.67	3.27 ± 0.89	3.62 ± 0.75	3.33 ± 0.57
S197(HG)-H300(NE2)	2.10 ± 0.35	1.93 ± 0.14	2.36 ± 0.71	2.10 ± 0.33	2.05 ± 0.20	2.44 ± 0.82	2.31 ± 0.50	2.12 ± 0.27
S197(OG)-C	4.12 ± 0.38	3.33 ± 0.30	3.88 ± 0.38	4.25 ± 0.57	4.54 ± 0.52	3.87 ± 0.41	4.14 ± 0.70	4.08 ± 0.48
S197(Cα)-C	5.79 ± 0.37	5.40 ± 0.26	5.84 ± 0.39	6.32 ± 0.67	6.21 ± 0.43	5.98 ± 0.36	5.94 ± 0.67	6.05 ± 0.34

## II Feeding Experiments

### II.1 Bacterial strains and media

*Escherichia coli* ET12567/pUZ8002 was used to drive conjugative transfer of plasmid DNA from *E. coli* to *S. erythraea* and *S. erythraea*  $\Delta$ AT6hyg<sup>R</sup>.

*S. erythraea* NRRL-B-24071, *S. erythraea*  $\Delta$ AT6hyg<sup>R</sup> and *S. erythraea* AT6\* were used for fermentation.

Cultivation of *E. coli* strains was performed using standard techniques (Sambrook, 2001). *S. erythraea* strains were grown on ABB13 (Fitzgerald, et al., 1998), SY (Wu, et al., 2000) or LB plates and cultivated in tryptone soy broth (TSB) (Kieser, 2000) and SM3 (Ranganathan, et al., 1999) as fermentation medium. For selection of *E. coli* transformants and *S. erythraea* transconjugants the following antibiotics were used: Apramycin (50 or 25  $\mu$ g/mL respectively), Hygromycin (50  $\mu$ g/mL) and Phosphomycin (100  $\mu$ g/mL).

### II.2 Fermentation experiments with *Saccharopolyspora erythraea*

#### 1. General procedure

Cultivation of *S. erythraea* was carried out in the 24-well plate system Duetz (Duetz, et al., 2000). Pre-cultures were cultivated for 48 h in 2 mL TSB at 30°C/180 rpm in an orbital shaker with 5 cm deflection. The 3 mL main culture in SM3 medium was inoculated with 1/20 pre-culture and cultivated for 5 days at 30°C/180 rpm. 4 clones of each mutant were used. The *S. erythraea* wild type was cultivated as control.

#### 2. Feeding of SNAC-activated compound to the DEBS3 Variants

For feeding experiments, the main culture was supplemented with 10-50 mM of either ethyl-, allyl-, (iso-)propyl-, butyl-, hexyl- or phenylmalonyl-SNAC. As control, propargylmalonyl-SNAC was also fed in a concentration of 10 mM.

#### 3. Analysis of fermentation products

For characterization via LC-ESI-MS the fermentation broth was extracted with 2 volumes of ethyl acetate overnight at 19°C. The solvent phase was evaporated and the residue re-dissolved in 0.5 mL methanol. 5  $\mu$ L were used for analysis.

LC-ESI-MS for routine analysis were measured using an Agilent 1100 series binary pump together with a C18 (CC12514 Nucleor C18 Gravity column (3  $\mu$ m particle size; Macherey-Nagel Germany) column coupled to a Finnigan LTQ linear ion trap (Thermo Electron Corporation, Dreieich, Germany). A flow rate of 500  $\mu$ L/min was used with a linear gradient starting with 80% Solvent A / 20% Solvent B for one minute and increasing to 0% Solvent A / 100% Solvent B in 10 min. After that the column was

washed with 0% Solvent A / 100% solvent B for 5 min and re-equilibrated to starting conditions for additional 5 min (solvent A water containing 0.1% formic acid, solvent B acetonitril containing 0.1% formic acid). For the determination of accurate mass-to-charge ratios the bacterial extracts were analyzed by HPLC coupled to high resolution mass spectrometry. The separations were carried out on an Accela HPLC-System (consisting of a pump, autosampler, column oven and a PDA detector) coupled online to an Orbitrap mass spectrometer equipped with a LTQ XL linear ion trap (Thermo Electron Corporation, Dreieich, Germany) using the standard electrospray ionization source. All solvents were LC-MS grade (Chromasolv, Sigma-Aldrich, Munich, Germany). 3 µl of each sample were injected (T = 10 °C) onto a Hypersil GOLD column (1.9 µm particle size, 50 mm length, 1 mm ID, 30 °C column temperature, Thermo Electron Corporation, Dreieich, Germany) using a flow rate of 500 µl/min and a linear gradient starting with 80 % solvent A / 20 % solvent B for one minute and increasing to 0 % solvent A / 100 % solvent B in 8.5 min. Afterwards the column was washed with 0 % solvent A / 100 % solvent B for 5 min and re-equilibrated to starting conditions for additional 5 min (solvent A: water containing 0.1 % formic acid, solvent B: acetonitrile containing 0.1 % formic acid). For mass spectrometric detection the electrospray ionization was carried out in positive ionization mode using a source voltage of 3.8 kV. The capillary voltage was set to 41 V, the capillary temperature to 275 °C, and the tube lens voltage to 140 V. Spectra were acquired in full scan centroid mode with a mass-to-charge range from 100 to 2000. The resolution in the Orbitrap was set to 60,000, the FTMS full AGC target to 500,000 and for internal calibration a lock mass of 391.2843 (diisooctyl phthalate) was used.

Fragmentation for LC-ESI-MS/MS analysis was accomplished via collision induced dissociation (CID). The following masses were fragmented:

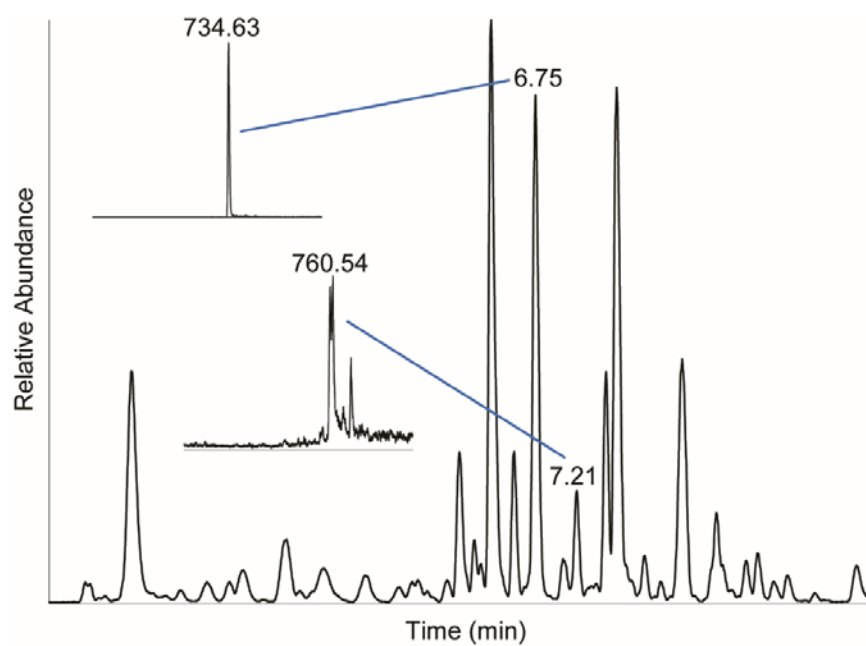
m/z = 734 (erythromycin,  $C_{37}H_{68}NO_{13}^{+}$ )

m/z = 748 (2-ethylerythromycin,  $C_{38}H_{70}NO_{13}^{+}$ )

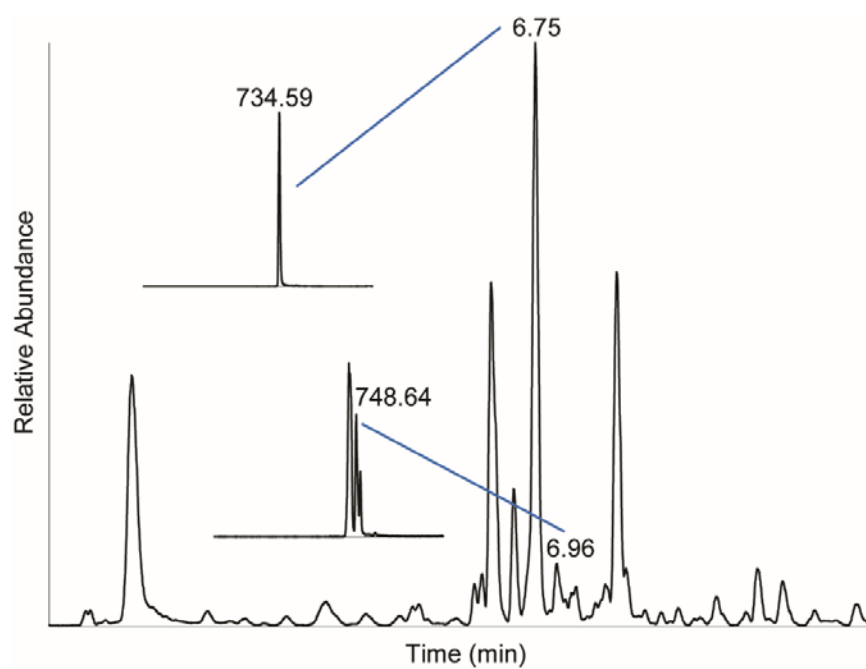
m/z = 758 (2-propargylerythromycin,  $C_{39}H_{68}NO_{13}^{+}$ )

m/z = 760 (2-allylerythromycin,  $C_{39}H_{70}NO_{13}^{+}$ )

# LC-ESI-MS/MS Analysis of Feeding Experiments



**SI Figure 3, Related to Figure 2: A.** ESI-MS spectra of 2-Allylerythromycin.



**SI Figure 3 B.** ESI-MS spectra2-Ethylerythromycin

### III. General information

Unless otherwise stated, materials for chemical synthesis were obtained from commercial suppliers (Sigma-Aldrich, Alfa Aesar, Fluka, Acros) in the highest purity available and used without further purification. Dry solvents were purchased from Sigma-Aldrich, stored over molecular sieves and used as supplied. Solvents used for extraction and chromatography were purchased from Thermo Fisher Scientific. Flash chromatography was carried out using Acros silica gel 60 (35–70  $\mu\text{m}$  mesh). Thin-layer chromatography (TLC) was performed on aluminium-backed, precoated silica gel (60 F245) from Merck with cyclohexane/EtOAc or DCM/MeOH mixtures as mobile phases. Spots were detected by staining with  $\text{KMnO}_4$  solution (5.0 g  $\text{KMnO}_4$ , 33 g  $\text{K}_2\text{CO}_3$ , 10 mL 5% aqueous NaOH in 500 mL  $\text{H}_2\text{O}$ ) and subsequent heat treatment.

NMR spectra were recorded by using a Varian Mercury 400 (400 MHz,  $^1\text{H}$ ; 100 MHz,  $^{13}\text{C}$ ) spectrometer and calibrated using residual undeuterated solvent as an internal reference.

High-resolution mass spectra were recorded at LTQ Orbitrap with Accela HPLC-System (column Hypersil Gold, length 50 mm, inside diameter 1 mm, particle size 1.9  $\mu\text{m}$ , ionization method: Electrospray Ionization). Products were characterized by NMR ( $^1\text{H}$ ,  $^{13}\text{C}$ ) and HRMS.

For mass spectrometric detection the electrospray ionization was carried out in positive ionization mode by using a source voltage of 4 kV. The capillary voltage was set to 18 V, the capillary temperature to 275  $^\circ\text{C}$ , and the tube lens voltage to 115 V. Spectra were acquired in full scan centroid mode with a mass-to-charge range from 200 to 2000.

Further information on the experimental details and analytical data can be found in Data in Brief [doi:10.1016/j.dib.2015.09.052](https://doi.org/10.1016/j.dib.2015.09.052).

#### IV References:

- Duetz, W.A., Rüedi, L., Hermann, R., O'Connor, K., Büchs, J., and Witholt, B. (2000). Methods for Intense Aeration, Growth, Storage, and Replication of Bacterial Strains in Microtiter Plates. *Applied and Environmental Microbiology* 66, 2641-2646.
- Fitzgerald, N.B., English, R.S., Lampel, J.S., and Vanden Boom, T.J. (1998). Sonication-Dependent Electroporation of the Erythromycin-Producing Bacterium *Saccharopolyspora erythraea*. *Applied and Environmental Microbiology* 64, 1580-1583.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14, 33-38.
- Kieser, T.B., M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. (2000). *Practical streptomyces genetics* (John Innes Foundation).
- Mackerell, A.D., Feig, M., and Brooks, C.L. (2004). Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* 25, 1400-1415.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26, 1781-1802.
- Ranganathan, A., Timoney, M., Bycroft, M., Cortés, J., Thomas, I.P., Wilkinson, B., Kellenberger, L., Hanefeld, U., Galloway, I.S., Staunton, J., et al. (1999). Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues. *Chemistry & Biology* 6, 731-741.
- Sambrook, J.R., D. W. (2001). *Molecular Cloning - A laboratory manual*. (CSH Press: New York).
- Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., and Schulz, F. (2013). Enzyme-Directed Mutasynthesis: A Combined Experimental and Theoretical Approach to Substrate Recognition of a Polyketide Synthase. *ACS Chemical Biology* 8, 443-450.
- Wu, K., Chung, L., Revill, W.P., Katz, L., and Reeves, C.D. (2000). The FK520 gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* 251, 81-90.
- Zoete, V., Cuendet, M.A., Grosdidier, A., and Michielin, O. (2011). SwissParam: A fast force field generation tool for small organic molecules. *Journal of Computational Chemistry* 32, 2359-2368.