Expression and purification of recombinant eryKR$_1$ variants

The eryKR$_1$ domain had been previously cloned into the pGEX-6P-1 vector, which encodes a GST fusion tag cleavable by a human rhinovirus 3C proteinase recognition site.[1] Each of the eryKR$_1$ mutants, KR$_1$-PQS, KR$_1$-WGG, and KR$_1$-PQS-WGG,[2] was subcloned in the same way into the pGEX-6P-1 vector *via* *Bam*HI and *Eco*RI restriction sites. This expression system allowed more specific and efficient cleavage of the GST tag using human rhinovirus 3C proteinase than did the previously-used pGEX-4T-3 expression system which used thrombin.[1] The plasmids encoding the respective eryKR$_1$ mutants were used to transform either *E. coli* BL21-Codon Plus(DE3) cells or *E. coli* BL21-Codon Plus(DE3)-RP cells. Protein expression was carried out in 500 mL cultures grown in LB broth, which were incubated at 37˚C until an OD_{600} of 0.4 was reached, then incubated at 21˚C until OD_{600} reached 0.6, at which point expression of the target protein was induced with 0.1 mM IPTG. After growth for an additional 16 hours, the cells were harvested by centrifugation (5000 × *g*, 5 minutes, 4˚C) and the cell pellets were stored at -20˚C until purification.

Frozen cell pellets were resuspended in Phosphate Buffer A (10 mL; 50 mM phosphate, 150 mM NaCl, 10% glycerol, pH 8) to which proteinase inhibitors (Complete Roche) (one tablet), DNase (100 mg L^{-1}), and RNase (25 mg L^{-1}) were added. Lysozyme was then added (1 mg mL^{-1}) and the cells were incubated on ice for 20 minutes before being disrupted by sonication (Misonix, Inc.). The insoluble material was pelleted by centrifugation (5 minutes, 5000 × *g*, 4˚C) and the supernatant was collected and added to 50% glutathione agarose bead slurry (10 mL). The beads and lysate were incubated for one hour at 4˚C with end-over-end agitation in a 50 mL Falcon tube, then collected and washed with 5 volumes of Phosphate Buffer A. The beads were then washed with two volumes of Phosphate Buffer B (100 mM phosphate, 150 mM NaCl, 5% glycerol, pH 7.5) to which 1 mM EDTA, 1 mM DTT and 0.01% Triton X-100 had been added, before being resuspended in one volume of the same buffer. 140 units (70 µL) of human rhinovirus 3C proteinase (PreScission, GE healthcare) were added and the slurry was incubated at 4˚C for 16 hours in order to cleave the GST-tag. The KR protein was then eluted, by gravity filtration through a 10 mL column, from the glutathione-agarose beads, concentrated to 2.5 mL using a centrifugal protein concentrator (20 mL, 10 kDa MWCO, Vivascience) and then further purified by gel filtration using a HiLoad 16/60 Superdex 75 column (Pharmacia) eluted with Phosphate Buffer B. The appropriate fractions were collected, concentrated, and glycerol was

[1] A. P. Siskos, A. Baerga-Ortiz, S. Bali, V. Stein, H. Mamdani, D. Spiteller, B. Popovic, J. B. Spencer, J. Staunton, K. J. Weissman, P. F. Leadlay, *Chem. Biol.* 2005, **12**, 1145.
[2] A. Baerga-Ortiz, B. Popovic, A. P. Siskos, H. M. O'Hare, D. Spiteller, M. G. Williams, N. Campillo, J. B. Spencer, P. F. Leadlay, *Chem. Biol.* 2006, **3**, 277.

added to 10% (v/v). Portions of the purified enzyme solution were stored at -80˚C. Typical yields of purified KR protein ranged from 5 to 10 mg per L of culture after expression in either *E. coli* strain.

Synthesis and characterisation of 2-methyl-3-oxopentanoyl-pantetheine (5)

D-Pantetheine **3**[3] and 2-methyl-3-oxopentanoic acid **4**[4] (Scheme 2) were prepared as previously reported. 2-Methyl-3-oxopentanoic acid (**4**, 0.69 g, 5.30 mmol) was cooled to 0 °C under argon. D-Pantetheine (**3**, 0.5 g, 1.79 mmol), dissolved in dry tetrahydrofuran (20 mL), was added to the acid, followed by more THF (40 mL), *N, N*-diisopropylethylamine (DIPEA, 0.6 mL, 3.36 mmol) and O-(7-azabenzotriazol-1-yl) *N*, *N*, *N'*, *N'*- tetramethyluroniumhexafluoro-phosphate (HATU, 1.36 g, 3.58 mmol). The reaction mixture became progressively yellow upon stirring under argon for 3 h at 0°C and at room temperature for 18 h. The solvent was removed *in vacuo* and the crude residue was washed with ether. The remaining material was dissolved in ethyl acetate, leaving a white solid behind. The solvent was evaporated and the crude material (a yellow oil) was purified on an Agilent 1100 HPLC system fitted with a semi preparative polar RP 80A- column (250 x 10.00 mm, 4 µ) and a linear gradient of water and acetonitrile (both containing 0.1% PFPA; the acetonitrile content was increased from 0% to 100% over 30 minutes; the flow rate was set at 2.5 ml/min and the UV detection at 210, 254 and 280 nm). Upon concentration of the relative fractions (R_T= 17 min), the desired product (**5**) was as a colourless liquid (145 mg, 20%); δ_H (400 MHz; CD_3OD): 0.79 (6H, s, Me), 0.89 (3H, t, *J* 7.3, CH_2CH_3), 1.19 (3H, d, *J* 7.0, $CHCH_3$), 2.28 (2H, t, *J* 6.9, CH_2), 2.47 (2H, m, CH_2CH_3), 2.92 (2H, t, *J* 6.9, CH_2), 3.35 (2H, m, CH_2), 3.39 (2H, s, CH_2OH), 3.47 (2H, m, CH_2), 3.76 (1H, s, CHOH), 3.81 (1H, q, *J* 7.0, CH (CH_3); δ_C (100 MHz; CD_3OD): 8.0 (q), 14.0 (q), 21.0 (q), 21.4 (q), 29.6 (t), 35.9 (t), 36.3 (t), 36.4 (t), 39.8 (s), 40.4 (t), 62.0 (d), 70.4 (t), 77.4 (d) 173.9 (s), 176.0 (s), 198.2 (s), 207.4 (s); *m/z (*EI) 391.1900 ($(M+H)^+$, $C_{17}H_{31}N_2O_6S$ requires 391.1897).

Kinetic assays of eryKR$_1$

NADPH-linked kinetic assays of recombinant eryKR$_1$ were carried out using a Spectramax Plus microplate reader. The rate of ketoreduction was followed by monitoring the fall in absorbance at 340 nm as a measure of the consumption of the NADPH cofactor. Assays were carried out in Phosphate Buffer C (400 mM phosphate at pH 7.5 and 150 mM NaCl), with 3.5 µM enzyme, 1 mM NADPH, and between 0 and 50 mM of substrate. Assays were carried

[3] L. Tran, M. Tosin, J. B. Spencer, P. F. Leadlay, K. J. Weissman, *ChemBioChem* 2008, **9**, 905.
[4] R. E. Tirpak, R. S. Olsen, M.W. Rathke, *J. Org. Chem.* 1985, **50**: 4877.

out in a total volume of 50 µL in a UV-transparent 96-well plate. The enzyme was preincubated with NADPH for 10 minutes prior to mixing with the substrate. UV measurements were taken at 12-second intervals over 20 minutes. Samples were mixed by shaking for 3 seconds before each measurement. All of the readings were corrected for the change in absorbance in the absence of substrate. Readings were made in either duplicate or triplicate. We determined the isolated, wild-type eryKR$_1$ domain had a k_{cat} = (0.24 ± 0.01)s^{-1} and K_M = (12 ± 1) mM for **1** and a a k_{cat} = (0.25 ± 0.02)s^{-1} and K_M = (12 ± 2) mM for **5**. The mutant enzymes retained a low proportion of the parent activity. When reproduced, assays of the mutants with **1**, as well as with **5**, showed lower overall activity in the mutants than previously reported,[2] with >5% parent activity observed.

Stereochemical assays

Determination of the configuration of the alcohol products of eryKR$_1$ was by carried out by a procedure adapted from previous work.[1,2,5] Enzymatic reduction of 2-methyl-3-oxopentanoyl-pantetheine (**5**) by the eryKR$_1$ domains was carried out in Phosphate Buffer C (200 µL) containing enzyme (20 µM), NADPH (10 mM), and **5 (**5-10 mM). Each of the reaction mixtures was incubated at 22˚C for 16 hours and subsequently extracted with ethyl acetate (4 x 500 µL). The organic phase (containing the pantetheinyl reduction products **6**) was concentrated and the residue was redissolved in aqueous $NaHCO_3$ (50 mM, 100-150 µL). *N*-Acetyl cysteamine was added (to a final concentration of 20-30 mM).[6,7] The mixture was incubated overnight at room temperature and the products were extracted in dichloromethane (3 x 200 µL). The organic layer was dried and concentrated. The residue was redissolved in isopropanol (50 µL) and analysed by normal-phase liquid chromatography using an Agilent 1100 system fitted with a ChiralCel OC column (Diacel, 25 cm × 4.6 mm). The eluant was a mixture (v/v) of 93% hexane and 7% ethanol at a flow rate of 0.8 mL min^{-1}. UV absorbance was monitored at 212 nm, 254 nm, and 280 nm. The chromatograms were compared to those of the synthetic standards **2a-d** (the four diastereomers of 2-methyl-3-hydroxypentanoyl-NAC). The stereospecificity of the mutant eryKR$_1$ proteins towards **5** (Table 1S) was the same as previously reported for **1**.[2]

[5] S. Bali, H. M. O'Hare, K. J. Weissman, *ChemBioChem* 2006, **7**, 478.
[6] J. Stöckigt, M. H. Zenck, *Z. Naturforsch.* 1975, **30**, 352.
[7] A. Al-Arif, M. Blecher, *J. Lipid Res.* 1969, **10**, 344.

Table 1S. Stereochemistry of the products resulting from the reduction of 2-methyl-3-oxopentanoyl-pantetheine 5 by recombinant eryKR1 and mutants thereof.

entry	KR	Mutations	Product			
			A-type		B-type	
			(2*S*,3*S*)	(2*R*,3*S*)	(2*S*,3*R*)	(2*R*,3*R*)
a	**eryKR$_1$**	**none**	minor	**-**	>95%	**-**
b	eryKR$_1$	L93P, D94Q, D95S	~50%	-	~50%	minor
c	eryKR$_1$	F141W, P144G	>95%	-	-	minor
d	eryKR$_1$	L93P, D94Q, D95S, F141W, P144G	>95%	-	-	-

Construction of plasmids for mutants of *S. erythraea* DEBS1-TE

The plasmid pCJR133 is a vector encoding the DEBS1-TE triketide synthase which integrates into the host actinomycete genome.[8] Initial plasmid constructs encoding mutant DEBS1-TE genes were based on pCJR133. Mutagenesis of eryKR$_1$ and eryKR$_2$ as discrete recombinant proteins had been performed previously.[2] The DNA encoding the active site mutations was cloned into pCJR133, replacing the corresponding parental sequence of the KR domains of DEBS1-TE, following several steps of sub-cloning as described below.

Plasmid constructs encoding DEBS1-TE containing mutations in the eryKR$_1$ domain were assembled as follows. Initially, a fragment of DNA from the DEBS1 gene including the eryKR$_1$-coding portion was amplified by PCR (using the primers TCATGCTCGAGCTCTCCTGG and TGGTCGAAGACGGTCGTCGT) and cloned into pUC18 via the *Sma*I restriction site. This intermediate plasmid was then used as a vector into which the DNA encoding active site mutations in eryKR$_1$ was cloned (as *Bsa*BI/*Bpu*1102I fragments from each of the isolated mutant eryKR$_1$-encoding plasmids), replacing the corresponding parental sequence. From each of the resultant constructs, a fragment encoding a portion of the DEBS1 gene, including the mutant eryKR$_1$ domain, was excised by digestion with *Stu*I cloned into the DEBS1-TE encoding plasmid p1B140 via the same restriction sites, replacing the corresponding parental sequence. From each of these pIB140-based constructs, a fragment of the resultant mutant DEBS1-TE gene was then cloned as an *Sfi*I/*Sex*AI fragment into pCJR133 replacing the parental sequence, which finally yielded plasmids pDK1.1.1, pDK1.1.2, and pDK1.1.3 (Table 2S).

Constructs with mutations in the eryKR$_2$ domain of DEBS1-TE were assembled in the following steps. The entire DEBS1-TE gene from plasmid pJLK23, encoding engineered *Spe*I

[8] C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton, P. F. Leadlay, *Gene* 1998, **216**, 215.

and *Nsi*I restriction sites flanking the eryKR$_2$-coding region,[9] was excised as an *Nde*I/*Xba*I fragment and cloned into pCJR133 replacing the parent DEBS1-TE gene, generating plasmid pDK1.2.0. Fragments of DNA encoding the individual mutant eryKR$_2$ domains were amplified by PCR (using primers ATACTAGTCCTCGTGACGAGCTCGACGGCTGGTTC and TAATGCATCCGGTTCTCCGGCCCGCTCGC) from their expression plasmids. The PCR fragments encoding mutant eryKR$_2$ domains were digested with *Spe*I and *Nsi*I and cloned into pDK1.2.0 replacing the corresponding parental sequence for eryKR$_2$ within the whole DEBS1-TE gene, resulting in the plasmids pDK1.2.1, pDK1.2.2, and pDK1.2.3 (Table 2S).

Table 2S. List of plasmids used

Plasmid	Marker	Description	Source
pCJR133	Apr^R	Integrative plasmid encoding DEBS1-TE — ϕC31 integrase catalysed integration into *attP* site	C. J. Rowe[8]
pDK1.1.1	Apr^R	pCJR133-based plasmid encoding DEBS1-TE KR$_1$-PQS mutant	This study
pDK1.1.2	Apr^R	pCJR133-based plasmid encoding DEBS1-TE KR$_1$-WGG mutant	This study
pDK1.1.3	Apr^R	pCJR133-based plasmid encoding DEBS1-TE KR$_1$-PQS-WGG mutant	This study
pDK1.2.0	Apr^R	pCJR133-based plasmid encoding DEBS1-TE with engineered restriction sites flanking EryKR$_2$	This study
pDK1.2.1	Apr^R	pDK1.2.0-based plasmid encoding DEBS1-TE KR$_2$-LDD mutant	This study
pDK1.2.2	Apr^R	pDK1.2.0-based plasmid encoding DEBS1-TE KR$_2$-LPN mutant	This study
pDK1.2.3	Apr^R	pDK1.2.0-based plasmid encoding DEBS1-TE KR$_2$-LDD-LPN mutant	This study
pFS1.1.0	AprR	pFS1-based plasmid encoding DEBS1-TE	This study
pFS1.1.1	AprR	pFS1-based plasmid encoding DEBS1-TE KR1-PQS mutant	This study
pFS1.1.2	AprR	pFS1-based plasmid encoding DEBS1-TE KR1-WGG mutant	This study
pFS1.1.3	AprR	pFS1-based plasmid encoding DEBS1-TE KR1-PQS-WGG mutant	This study
pFS1.2.0	AprR	pFS1-based plasmid encoding DEBS1-TE with engineered restriction sites flanking EryKR2	This study
pFS1.2.1	AprR	pFS1.2.0-based plasmid encoding DEBS1-TE KR2-LDD mutant	This study
pFS1.2.2	AprR	pFS1.2.0-based plasmid encoding DEBS1-TE KR2-LPN mutant	This study
pFS1.2.3	AprR	pFS1.2.0-based plasmid encoding DEBS1-TE KR2-LDD-LPN mutant	This study
pIB140	AprR	Plasmid encoding DEBS1-TE	I. Bohm
pUC18	Amp^R	Vector for routine cloning	Fermentas

[9] L. Kellenberger, I. S. Galloway, G. Sauter, G. Böhm, U. Hanefeld, J. Cortés, J. Staunton, P. F. Leadlay, *ChemBioChem* 2008, **9**, 2740.

Conjugation of *S. erythraea* and verification of transformants

Plasmids pCJR133, pDK1.1.1, pDK1.1.2, pDK1.1.3, pDK1.2.0 pDK1.2.1, pDK1.2.2, and pDK1.2.3 were used to transform *E. coli* ET12567/pUZ8002 and the resulting transformants were then used as the donor strain from which the plasmids were shuttled into *S. erythraea* JC2 *via* conjugation as described by Kieser *et al*.[10] *S. erythraea* JC2 is a strain derived from NRRL2338 from which the resident erythromycin PKS genes have been substantially removed,[11] and *S. erythraea* BIOT1717-JC2 bears the same mutation in an industrial strain overproducing erythromycin (kind gift of Dr. Christine Rowe, Biotica). The plasmids pFS1.1.0, pFS1.1.1, pFS1.1.2, pFS1.1.3, pFS1.2.0, pFS1.2.1, pFS1.2.2, and pFS1.2.3 were used to transform *E. coli* ET12567/pUZ8002 and the resulting transformants were then used in the conjugation of *S. erythraea* BIOT1717-JC2, introducing the appropriate plasmid. Integration of each construct into the host *S. erythraea* genome was verified by sequencing.

Extraction and chromatography of triketide lactones

Strains of *S. erythraea* JC2 transformed with plasmids pCJR133, pDK1.1.1, pDK1.1.2, pDK1.1.3, pDK1.2.0, pDK1.2.1, pDK1.2.2, and pDK1.2.3 were grown on solid SM3 plates containing apramycin for 14 to 20 days. Sections of approximately 1 mL were sliced from the agar plate cultures, triturated, and extracted with 1 mL of ethyl acetate containing 1.6% (v/v) formic acid. Heating the sample at 50˚C for 15 minutes facilitated the extraction. The organic phase was collected and the solvent was removed *in vacuo*. The extracts were redissolved in ethyl acetate (100 µL), of which small amounts (2 to 5 µL) were injected into an MAT GCQ instrument (Finnigan) for GC-MS analysis. The GC-MS analysis was carried out in chemical ionisation mode, using a Phenomenex AB-5 column (30m × 25mm × 25µm) with the following temperature gradient: 40˚C for 2 min; 10˚C min^{-1} to 250˚C; 25˚C min^{-1} to 300˚C. Triketide lactones could be detected by GC–MS, however standards of **7a** and **7b** could not be reproducibly resolved under standard GC–MS conditions.

In order to achieve better resolution of triketide lactone diasteromers, LC-MS analysis was utilised. Under the conditions tested, authentic synthetic samples of **7a**, **7b**, and **7c** were easily separable by reverse-phase liquid chromatography. Due to the lower sensitivity of the available LC–MS instrumentation compared to GC–MS, higher yields of the analytes were sought through the use of an over-producing host strain. Thus, the DEBS1-TE mutant genes

[10] T. Kieser, M. J. Bibb, M. J.,Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics* 2000, The John Innes Foundation, Norwich.
[11] J. Cortés, K. E. H. Weismann, G. A. Roberts, M. J. B. Brown, J. Staunton, P. F. Leadlay, *Science* 1995, **268**, 1487.

were introduced, under the control of the erythromycin PKS promoter, into a strain of *S. erythraea* (BIOT-1717 JC2), engineered from an erythromycin overproducing strain by deleting essentially all of the erythromycin PKS genes. Strains of *S. erythraea* BIOT1717-JC2 transformed with plasmids pFS1.1.0, pFS1.1.1, pFS1.1.2, pFS1.1.3, pFS1.2.0, pFS1.2.1, pFS1.2.2, and pFS1.2.3 were each cultured in eryP medium (25 mL) containing apramycin for 7 days. The cultures were then extracted with ethyl acetate (2 x 25 mL) containing 1.6% (v/v) formic acid. Separation of phases was facilitated by centrifugation (2500 × *g*, 10 minutes). The organic phase was collected and the solvent removed *in vacuo*. The extracts were redissolved in methanol (1 mL), and injected (5 µL) into an Agilent 1200 HPLC instrument, fitted with a Phenomenex Prodigy 5µ ODS column (250mm × 2.00 mm × 5 µm) and connected to an LTQ mass spectrometer (Finnigan). The analytes were resolved using a water/acetonitrile gradient with 0.1% formic acid at 40˚C: the acetonitrile was increased from 5% to 50% over 20 minutes, then to 95% over 5 minutes and remained at 95% for 4 minutes before returning to 5% over 1 minute.

The LC traces of extracts of strains encoding the DEBS1-TE mutants are summarised in Figure 1S alongside the chromatogram of synthetic standards corresponding to the normal triketide lactone (**7a**) and the target compounds (**7b** and **7c**). None of the DEBS1-TE mutants carrying substitutions within their KR domains gave rise to diastereomers of **7a** with altered hydroxyl configuration. Only a shift towards the production of 3-keto lactone **8** was observed in mutants bearing substitutions in EryKR$_2$, corresponding to a loss in overall KR activity rather than a shift in KR stereospecificity. The relative yields of triketide lactones from each of the mutant strains are summarised in Figure 2S.