Enediyne Polyketide Synthases Stereoselectively Reduce the β-Ketoacyl Intermediates

to β -D-Hydroxyacyl Intermediates in Enediyne Core Biosynthesis

Hui-Ming Ge,[†] Tingting Huang,[†] Jeffrey D. Rudolf,[†] Jeremy R. Lohman,[†] Sheng-Xiong Huang,[†] Xun Guo,[†] and Ben Shen^{*,†,‡,§}

[†]Department of Chemistry, [‡]Department of Molecular Therapeutics, and [§]Natural Products Library Initiatives, The Scripps Research Institute, Jupiter, Florida 33458, United States

Supporting Information

Table of contents

General experimental procedures					
Cloning of <i>pksE-KR</i> s					
Overproduction of PKSE-KRs in E. coli					
Purification of PKSE-KRs					
Enzymatic activity of PKSE-KRs and chiral HPLC analysis					
Kinetic studies of PKSE-KRs					
Chemical synthesis of substrates and products					
General procedures for the preparation of MTPA esters					
Supplementary references					
Table S1.	PCR primers used in this study				
Figure S1.	The seven selected enediynes and their PKSEs	S11			
Figure S2.	Sequence comparison of KRs between selected type I PKSs and PKSEs	S13			
Figure S3.	SDS-PAGE of the purified KRs	S14			
Figure S4, S5.	¹ H (400 MHz) and ¹³ C NMR (100 MHz) spectra of 1 in CDCl ₃	S15, 16			
Figure S6, S7.	¹ H (400 MHz) and ¹³ C NMR (100 MHz) spectra of 2 in CDCl ₃	S17, 18			
Figure S8, S9.	¹ H (400 MHz) and ¹³ C NMR (100 MHz) spectra of 3 in CDCl ₃	S19, 20			
Figure S10 , S11 .	1 H (400 MHz) and 13 C NMR (100 MHz) spectra of 4 or 7 in CDCl ₃	S21, 22			
Figure S12, S13.	1 H (400 MHz) and 13 C NMR (100 MHz) spectra of 5 or 8 in CDCl ₃	S23, 24			
Figure S14, S15.	¹ H (400 MHz) and ¹³ C NMR (100 MHz) spectra of 6 or 9 in $CDCl_3$	S25, 26			
Figure S16, S17.	¹ H NMR (400 MHz) spectra of <i>S</i> -MTPA-5 and <i>R</i> -MTPA-5 in CDCl ₃	S27, 28			
Figure S18.	$\Delta \delta_{S-R}$ values (ppm) of MTPA esters of 5	S28			
Figure S19, S20.	¹ H NMR (400 MHz) spectra of <i>S</i> -MTPA-6 and <i>R</i> -MTPA-6 in CDCl ₃	S29, 30			
Figure S21.	$\Delta \delta_{S-R}$ values (ppm) of MTPA esters of 6	S30			
Figure S22.	Enzymatic assay by following at 340 nm change as exemplified by SgcE-K	R S31			
Figure S23.	Pseudo-first order kinetic characterization of PKSE-KRs with 1, 2, and 3	S32			

General experimental procedures

DNA oligonucleotide primers were synthesized by Integrated DNA Technologies. DNA sequencing was performed by Genewiz, Inc. Restriction enzymes, T4 DNA ligase, dNTPs, T4 DNA polymerase, and Platinum Pfx DNA polymerase were purchased from commercial sources. PCR was performed with a BIO-RAD S1000 Thermal cycler. DNA gel extraction and plasmid isolation were performed on Omega E.Z.N.A. Gel Extraction Kit and Omega E.Z.N.A Plasmid Mini Kit, respectively. ¹H and ¹³C NMR data were recorded at 25 °C on a Bruker AM-400 instrument. Optical rotation values were measured with an AUTOPOL[®]IV automatic polarimeter using a quartz cell with 1 mL capacity and a 1 dm path length. UV absorbance was measured on a Nanodrop 2000 UV-Vis spectrometer. High resolution electrospray ionization mass spectrometry (HRESIMS) was carried out on an Agilent 6230 TOF-LC-MS. Chiral HPLC was carried out with an Agilent 1260 infinity system equipped with quaternary pumps, a ChiralCel OC-H column (5 µm, 4.6 × 250 mm), and a diode array detector. Protein purification was conducted on Äkta FPLC equipped with a HisTrap HP column (GE Healthcare Life Sciences).

Cloning of pksE-KRs

The genes encoding the SgcE-KR, KedE-KR, MdpE-KR, NcsE-KR, and CalE8-KR domains were amplified by PCR from cosmids pBS1006,¹ pBS16004,² pBS10004,³ pBS5013,⁴ and pBS14009,⁵ respectively, and cloned into pBS3080⁶ using the ligation independent method. PCR was performed with Plantium Pfx polymerase from Invitrogen using the primers described in Table 1 (also see Figures 1 and 2 for structures of the enediynes and protein Briefly, pBS3080 was first sequence comparison among the PKSEs and PKSE-KRs). digested with BsmF1 and purified by gel electrophoresis. The linearized vector was treated with T4 DNA polymerase in the presence of dGTP at 20 °C for 30 min followed by heating at 75 °C for 20 min to denature the polymerase, affording overhangs with complimentary sequences to clone the PCR amplified pksE-KRs. Similarly, each of the PCR amplified pksE-KR fragments was purified by gel electrophoresis and treated with T4 DNA polymerase in the presence of dCTP at 20 °C for 30 min followed by heating at 75 °C. The linearized and T4 DNA polymerase treated pBS3080 vector and the PCR amplified and T4 DNA polymerase treated *pksE-KR* fragments were then mixed at room temperature, annealed on ice for 5 min, and transformed into E. coli DH5a for ligation independent cloning to construct the KR expression plasmids pBS1134, pBS16014, pBS10019, pBS5053, and pBS14016, respectively. Finally, the above plasmids were isolated from E. coli DH5a and confirmed by DNA sequencing, in which the KRs were produced as fusion proteins with N-terminal His₆-tags. The genes encoding the DynE8-KR and UcmE-KR were amplified by PCR from the chromosomal DNA of Micromonospora chersina ATCC53710 and Streptomyces uncialis DCA2648 strains, respectively, using the primers listed in Table 1. The purified PCR products were cloned into pET-28a digested with NdeI and EcoRI to yield pBS14017 and pBS18001.

Overproduction of PKSE-KRs in E. coli

The expression plasmids for each of the *pksE-KR* genes were transformed into *E. coli* BL21 (DE3) and the resultant recombinant strains were grown overnight in 50 mL of LB media containing 50 µg/mL kanamycin. A 10 mL aliquot of the overnight culture was used to inoculate 1 L of LB containing 50 µg/mL kanamycin, which was then incubated at 37 °C with shaking at 250 rpm. Once the OD₆₀₀ reached ~0.4, the temperature was reduced to 18°C. After the cultures reached thermal equilibrium, gene expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to 0.25 mM, and incubation continued for an additional 18 hrs at 18 °C. *E. coli* cells were harvested by centrifugation at 3500 x g at 4 °C for 20 min.

Purification of PKSE-KRs

The *E. coli* cell pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, containing 1 µg/mL DNase, 300 mM NaCl, 10 mM imidazole, and 5% glycerol), sonicated (30×2 s on ice), and clarified by centrifugation at 15,000 rpm at 4 °C for 30 min. The supernatant was applied to a 5 mL HisTrap HP column and washed with lysis buffer. The column was washed with wash buffer (50 mM Tris-HCl, pH 8.0, containing 300 mM NaCl and 20 mM imidazole) and the PksE-KRs were eluted with wash buffer containing 250 mM imidazole. The purified protein was desalted through a PD-10 desalting column and concentrated with a Vivaspin ultrafiltration device (30,000 molecular weight cut-off). The purity of purified protein was assessed by 12% SDS-PAGE (Fig. S3), and the concentration was determined from the absorbance at 280 nm using the molar absorptivity of each protein calculated by Protparam (http://web.expasy.org/protparam/).

Enzymatic activity of PKSE-KRs and chiral HPLC analysis

Each KR (10 μ M) was incubated with 5 mM NADPH and 1 mM substrate in a total of 0.5 mL sodium phosphate buffer (100 mM, pH 7.2) at 28 °C for 3 h. The reaction product was extracted using ethyl acetate (3 × 400 μ L), and after removal of the solvent, were resuspended in 500 μ L of isopropanol. The isopropanol solution was separated using a ChiralCel OC-H column (250 × 4.6 mm) detected by 230 nm on an Agilent 1260 HPLC system. Compounds were eluted at 1 mL/min in 15% isopropanol in *n*-hexane.

Kinetic studies of PKSE-KRs

Kinetics studies for each KR were conducted spectrophotometrically by monitoring the change in absorbance at 340 nm, due to the consumption of NADPH ($\epsilon_{340} = 6220 \text{ M}^{-1} \text{ cm}^{-1}$). Assays were performed in the presence of 100 mM sodium phosphate buffer (pH 7.2), 10 μ M KR, 0.3 mM NADPH, and varying concentration of substrates (0.05 – 2 mM) in a total

volume of 0.2 mL. Reactions were initiated by the addition of enzyme and read every 10 s. The approximation of k_{cat}/K_m values were calculated from three substrate concentrations using the limiting case of the Michaelis-Menten equation $v = (k_{cat}/K_m)[S][E]$ for low substrate concentrations ([S] << K_m).⁷ Each reaction was performed in triplicate and averaged (Figures 22 and 23).

Chemical synthesis of substrates and products



The known compounds 1, 4 and 7 were synthesized following literature procedures.^{8,9} All spectroscopic data and physical properties matched those previously reported (Figures 4, 5, 10, and 11).^{8,9}



To a stirred solution of **10** (1 g, 4.9 mmol) in dry DCM (30 mL), TiCl₄ (1.0 M in DCM, 5.4 mL, 5.4 mmol) was added at 0 °C under argon. The reaction mixture was stirred for 5 min and then cooled to -78 °C. A solution of DIPEA (941 μ L, 5.4 mmol) in DCM was added. The reaction mixture was stirred at -78 °C for 2 h. A solution of aldehyde **11** (700 mg, 10 mmol) was added to the reaction mixture, which was then stirred for 3 h at -78 °C. The reaction was quenched with 10 mL of saturated ammonium chloride. The layers were separated and the aqueous layer was extracted with DCM (3 × 20 mL). The combined organic layers were washed with brine (20 mL) and dried over Na₂SO₄. The solvent was removed *in vacuo* and the residue was purified using flash column chromatography (EtOAc/hexanes 1:5) to give diastereomers **12** (510 mg, 37% yield) and **13** (416 mg, 30% yield) as yellow oils.

Spectroscopic data for **12**. $[\alpha]_D^{25} = +446$ (c = 0.65, DCM), ¹H NMR (400 MHz, CDCl₃) δ 5.73 (m, 1H, H-3); 5.57 (m, 1H, H-2), 5.18 (m, 1H, H-9), 4.52 (m, 1H, H-4), 3.59 (dd, 1H, *J* = 17.4, 9.0 Hz, H-5), 3.52 (dd, 1H, *J* = 11.5, 8.0 Hz, H-8), 3.36 (dd, 1H, *J* = 17.4, 3.2 Hz,

H-5), 3.04 (dd, 1H, J = 11.5, 1.1 Hz, H-8), 2.37 (m, 1H, H-10), 1.70 (ddd, 3H, J = 6.4, 1.5, 0.9 Hz, H-1), 1.06 (d, 3H, J = 6.8 Hz, H-11), 0.98 (d, 3H, J = 6.8 Hz, H-11); ¹³C NMR (100 MHz, CDCl₃) δ 203.0, 173.1, 131.8, 127.4, 71.4, 69.1, 45.2, 30.8, 30.6, 19.1, 17.8, 17.7; HRESIMS for the [M + Na]⁺ ion at *m*/*z* 296.0746 (calcd [M + Na]⁺ ion for C₁₂H₁₉NO₂S₂ at *m*/*z* 296.0749).

Spectroscopic data for **13**. $[\alpha]_D^{25} = +487$ (c = 0.55, DCM), ¹H NMR (400 MHz, CDCl₃) δ 5.74 (m, 1H, H-3); 5.58 (m, 1H, H-2), 5.15 (m, 1H, H-9), 4.61 (m, 1H, H-4), 3.60 (dd, 1H, *J* = 17.5, 9.0 Hz, H-5), 3.53 (dd, 1H, *J* = 11.3, 8.0 Hz, H-8), 3.30 (dd, 1H, *J* = 17.5, 3.2 Hz, H-5), 3.02 (dd, 1H, *J* = 11.3, 1.1 Hz, H-8), 2.36 (m, 1H, H-10), 1.71 (ddd, 3H, *J* = 6.2, 1.6, 0.9 Hz, H-1), 1.07 (d, 3H, *J* = 6.8 Hz, H-11), 0.98 (d, 3H, *J* = 6.8 Hz, H-11); ¹³C NMR (100 MHz, CDCl₃) δ 203.0, 172.6, 131.8, 127.3, 71.4, 68.8, 45.4, 30.8, 30.7, 19.1, 17.8, 17.7; HRESIMS *m*/*z* 296.0745 [M + Na]⁺ (calcd [M + Na]⁺ ion for C₁₂H₁₉NO₂S₂ at *m*/*z* 296.0749).



5: To a stirred solution of **12** (250 mg, 0.92 mmol) in 5 mL DCM, imidazole (188 mg, 2.76 mmol) and *N*-acetylcysteamine (219 mg, 1.84 mmol) were added. The reaction mixture was stirred until the yellow color disappeared. The reaction was quenched with 5 mL of saturated ammonium chloride. The organic layer was removed and the aqueous layer was extracted with DCM (3×10 mL). The combined organic layers were washed with brine and dried over Na₂SO₄. The solvent was removed *in vacuo* and the residue was purified using flash column chromatography (DCM/MeOH 100:5) to give **5** (129 mg, 61%) as a white solid.

Spectroscopic data for **5**. $[\alpha]_D^{25} = -18.6$ (c = 0.44, DCM), ¹H NMR (400 MHz, CDCl₃) δ 5.90 (br s, 1H, N<u>H</u>), 5.74 (m, 1H, H-3); 5.50 (m, 1H, H-2), 4.54 (m, 1H, H-4), 3.44 (m, 2H, H-8), 3.05 (m, 2H, H-7), 2.78 (m, 2H, H-5), 2.60 (br s, 1H, O<u>H</u>), 1.97 (s, 3H, H-10), 1.70 (ddd, 3H, J = 6.5, 1.6, 0.8 Hz); ¹³C NMR (100 MHz, CDCl₃) δ 198.8, 170.5, 131.6, 127.9, 69.6, 51.0, 39.3, 28.8, 23.2, 17.7; HRESIMS for the [M + Na]⁺ ion at *m*/*z* 254.0822 (calcd [M + Na]⁺ ion for C₁₀H₁₇NO₃S at *m*/*z* 254.0821) (Figures 12 and 13).



Compound 8 (112 mg, 66%) was synthesized from 13 using the method described above. 8, 51 mg, white solid. NMR and MS data were identical to those of 5. $[\alpha]_D^{25} = +16.7$ (c = 0.42, DCM) (Figures 12 and 13).



To a stirred solution of **8** (204 mg, 0.88 mmol) in 15 mL DCM, Dess-Martin periodinane (448 mg, 1.06 mmol) was added. The resulting solution was stirred at room temperature for 2 h. The reaction was quenched with 1 mL of isopropanol. The solution was washed with brine and dried with Na₂SO₄. The solvent was removed *in vacuo* and the residue was purified using flash column chromatography (DCM/MeOH 100:2) to give **2** (keto:enol = 5:7, 66 mg, 33%) as a white solid.

Spectroscopic data for **2**. ¹H NMR (400 MHz, CDCl₃) δ 12.33 (br s, 1.4 H, enol-O<u>H</u>), 6.95 (m, 1H, keto-H-2); 6.77 (m, 1.4 H, enol-H-2), 6.20 (dq, 1H, *J* = 15.4, 1.8Hz, keto-H,-3), 5.99 (br s, 2.4H, N<u>H</u>), 5.75 (dq, 1.4H, *J* = 1.2 Hz, enol-H-3H-7), 5.43 (s, 1.4H, enol-H-5), 3.83 (s, 2H, keto-H-5), 3.47 (m, 4.8H, H-8), 3.10 (m, 4.8H, H-7), 1.97 (s, 7.2H, H-10), 1.96 (dd, 3H, *J* = 1.6, 7.0 Hz, keto-H-1), 1.90 (dd, 4.2H, *J* = 1.6, 7.0 Hz, keto-H-1); ¹³C NMR (100 MHz, CDCl₃) δ 194.5, 192.7, 191.4, 170.4, 167.4, 146.2, 139.1, 131.1, 125.2, 99.5, 54.7, 40.0, 39.2, 29.3, 28.0, 23.2, 23.1, 18.5; HRESIMS for the [M + Na]⁺ ion at *m*/*z* 252.0664 (calcd [M + Na]⁺ ion for C₁₀H₁₅NO₃S at *m*/*z* 252.0665) (Figures 6 and 7).



Compounds 15 (137 mg, 33%) and 16 (112 mg, 27%) were synthesized from (2E, 4E)-hexa-2,4-dienal (14) using the method described for 12 and 13.

Spectroscopic data for **15**. $[\alpha]_D^{25} = +282$ (c = 0.63), ¹H NMR (400 MHz, CDCl₃) δ 6.24 (dd, 1H, J = 10.4, 15.6 Hz, H-4); 6.04 (m, 1H, H-3), 5.74 (m, 1H, H-2), 5.61 (dd, 1H, J = 6.2, 15.2 Hz, H-5) 5.15 (m, 1H, H-11), 4.70 (m, 1H, H-6), 3.61 (dd, 1H, J = 17.4, 9.0 Hz, H-7), 3.52 (dd, 1H, J = 11.3, 8.0 Hz, H-10), 3.32 (dd, 1H, J = 17.4, 3.2 Hz, H-7), 3.03 (dd, 1H, J = 11.3, 1.1 Hz, H-10), 2.37 (m, 1H, H-12), 1.75 (dd, 3H, J = 6.2, 1.1 Hz, H-1), 1.07 (d, 3H, J = 6.8 Hz, H-13), 0.98 (d, 3H, J = 6.8 Hz, H-13); ¹³C NMR (100 MHz, CDCl₃) δ 203.0, 172.4, 131.1, 130.7, 130.6, 130.5, , 127.3, 71.4, 68.5, 45.4, 30.8, 30.7, 19.1, 18.1, 17.8; HRESIMS

for the $[M + Na]^+$ ion at m/z 322.0901 (calcd $[M + Na]^+$ ion for $C_{14}H_{21}NO_2S_2$ at m/z 322.0906).

Spectroscopic data for **16**. $[\alpha]_D^{25} = +291$ (c = 0.51, DCM), ¹H NMR (400 MHz, CDCl₃) δ 6.25 (dd, 1H, *J* = 10.4, 15.6 Hz, H-4); 6.04 (m, 1H, H-3), 5.74 (m, 1H, H-2), 5.59 (dd, 1H, *J* = 6.2, 15.2 Hz, H-5) 5.17 (m, 1H, H-11), 4.61 (m, 1H, H-6), 3.62 (dd, 1H, *J* = 17.0, 9.0 Hz, H-7), 3.53 (dd, 1H, *J* = 11.5, 8.0 Hz, H-10), 3.38 (dd, 1H, *J* = 17.0, 3.2 Hz, H-7), 3.04 (dd, 1H, *J* = 11.5, 1.0 Hz, H-10), 2.37 (m, 1H, H-12), 1.75 (dd, 3H, *J* = 6.7, 1.2 Hz, H-1), 1.07 (d, 3H, *J* = 6.8 Hz, H-13), 0.98 (d, 3H, *J* = 6.9 Hz, H-13); ¹³C NMR (100 MHz, CDCl₃) δ 203.0, 172.8, 131.3, 130.7, 130.6, 130.3, 71.3, 69.0, 45.1, 30.7, 30.6, 19.1, 18.1, 17.8; HRESIMS for the [M + Na]⁺ ion at *m*/*z* 322.0910 (calcd [M + Na]⁺ ion for C₁₄H₂₁NO₂S₂ at *m*/*z* 322.0906).



Compound 6 (133mg, 63%) was synthesized from 15 using the method described for 5.

Spectroscopic data for **6**. $[\alpha]_D^{25} = -11.6$ (c = 0.29, DCM), ¹H NMR (400 MHz, CDCl₃) δ 6.24 (dd, 1H, *J* = 10.4, 14.4 Hz, H-4), 6.03 (m, 1H, H-3), 5.85 (br s, 1H, N<u>H</u>), 5.74 (m, 1H, H-2); 5.55 (dd, 1H, *J* = 5.7, 14.8 Hz, H-5), 4.62 (m, 1H, H-6), 3.45 (m, 2H, H-10), 3.05 (m, 2H, H-9), 2.81 (m, 2H, H-7), 2.60 (br s, 1H, O<u>H</u>), 1.96 (s, 3H, H-12), 1.75 (dd, 3H, *J* = 6.5, 1.2 Hz) ¹³C NMR (100 MHz, CDCl₃) δ 198.5, 170.7, 131.5, 131.0, 130.43, 130.41, 69.3, 51.1, 39.3, 28.8, 23.2, 18.1; HRESIMS for the [M + Na]⁺ ion at *m*/*z* 280.0972 (calcd [M + Na]⁺ ion for C₁₂H₁₉NO₃S at *m*/*z* 280.0978) (Figure 14 and 15).



Compound 9 (77mg, 72%) was synthesized from 16 using the method described for 5.

Spectroscopic data for 9. $[\alpha]_D^{25} = +13.0$ (c = 0.27, DCM), the NMR and MS data were identical to those of 6 (Figures 14 and 15).



Compound **3** was synthesized from**9** using the method described for synthesizing **2**.

3: 37 mg, white solid, yield 33%. ¹H NMR (400 MHz, CDCl₃) δ 12.26 (br s, 0.75 H, O<u>H</u>), 7.04-7.15 (m, overlapped, 1H, H-4), 6.03-6.20 (m, overlapped, 2.25H, H-3), 5.89 (br s, 1H, N<u>H</u>), 5.64 (d, J = 15.2 Hz, 0.75H, enol-H-5); 5.41 (s, 0.75 H, enol-H-7), 3.78 (s, 0.5 H, keto-H-7), 3.41 (m, 2H, H-10), 3.03 (m, 2H, H-9) 1.90 (s, 3H, H-12), 1.83 (d, 0.75 H, J = 6.5 Hz), 1.79 (d, 2.25 H, J = 6.5 Hz) ¹³C NMR (100 MHz, CDCl₃) δ 194.1, 192.7, 191.6, 170.5, 170.4, 167.9, 145.8, 142.7, 140.1, 138.2, 130.6, 130.1, 128.2, 126.4, 124.2, 122.0, 100.5, 55.3, 40.0, 39.3, 29.7, 29.3, 28.0, 23.2, 23.1, 19.0, 18.8; HRESIMS for the [M + Na]⁺ ion at m/z 278.0820 (calcd [M + Na]⁺ ion for C₁₂H₁₇NO₃S at m/z 278.0821) (Figures 8 and 9).

General procedures for the preparation of MTPA esters

(*R*)- or (*S*)-MTPA chloride (4-fold mM excess over starting alcohol) was added to a solution of starting alcohol (2 mg) in pyridine (200 μ L), and the resulting mixture was allowed to stand at room temperature for 4 h. The mixture was quenched with 10 μ l of water and stirred for an additional 10 min, then diluted with DCM (5 mL). The solvent was removed on a rotary evaporator. The crude product was purified by the preparative TLC (DCM / MeOH = 20:1) to give the MTPA ester. The $\Delta\delta_{S-R}$ values (ppm) of MTPA esters for compound **5** and **6** see Fig. S18 and S21, respectively.

Spectroscopic data for (*S*)-**MTPA-5**. ¹H NMR (400 MHz, CDCl₃) δ 7.38-7.49 (m, 5H, phenyl protons at MTPA), 5.95 (m, 1H, H-3), 5.88 (m, 1H, H-4), 5.78 (br s, 1H, N<u>H</u>), 5.50 (m, 1H, H-2); 3.51 (br s, 3H, OMe in MTPA), 3.35 (m, 2H, H-8), 3.03 (m, 1H, H-5a), 2.97 (m, 2H, H-7), 2.81 (m, 1H, H-5b), 1.95 (s, 3H, H-10), 1.73 (dd, 3H, *J* = 6.5, 1.5 Hz, H-1); HRESIMS for the [M + Na]⁺ ion at *m*/*z* 470.1217 (calcd [M + Na]⁺ ion for C₂₀H₂₄NF₃O₅S at *m*/*z* 470.1220) (Figures 16 and 18).

Spectroscopic data for (*R*)-MTPA-5. ¹H NMR (400 MHz, CDCl₃) δ 7.38-7.49 (m, 5H, phenyl protons at MTPA), 5.87 (m, 1H, H-4); 5.86 (m, 1H, H-3), 5.81 (br s, 1H, N<u>H</u>), 5.39 (m, 1H, H-2), 3.51 (br s, 3H, OMe in MTPA), 3.39 (m, 2H, H-8), 3.07 (m, 1H, H-5a), 3.03 (m, 2H, H-7), 2.85 (m, 1H, H-5b), 1.94 (s, 3H, H-10), 1.69 (dd, 3H, *J* = 6.6, 1.6 Hz, H-1); HRESIMS for the [M + Na]⁺ ion at *m/z* 470.1213 (calcd [M + Na]⁺ ion for C₂₀H₂₄NF₃O₅S at *m/z* 470.1220) (Figures 17 and 18).

Spectroscopic data for (*S*)-**MTPA-6**. ¹H NMR (400 MHz, CDCl₃) δ 7.32-7.40 (m, 5H, phenyl protons at MTPA), 6.29 (m, 1H, H-5), 5.94 (m, 1H, H-2), 5.87 (m, 1H, H-6), 5.75 (m, 1H, H-3), 5.72 (br s, 1H, N<u>H</u>), 5.45 (m, 1H, H-4), 3.45 (br s, 3H, OMe in MTPA), 3.29 (m, 2H, H-10), 2.98 (m, 1H, H-7a), 2.91 (m, 2H, H-9), 2.74 (m, 1H, H-7b), 1.88 (s, 3H, H-12), 1.72 (dd, 3H, H-1, *J* = 6.6, 1.6 Hz, H-1); HRESIMS for the [M + Na]⁺ ion at *m*/*z* 496.1368 (calcd [M + Na]⁺ ion for C₂₂H₂₆NF₃O₅S at *m*/*z* 496.1376) (Figures 19 and 21).

Spectroscopic data for (*R*)-MTPA-6. ¹H NMR (400 MHz, CDCl₃) δ 7.32-7.40 (m, 5H, phenyl protons at MTPA), 6.20 (m, 1H, H-5), 5.90 (m, 1H, H-2), 5.86 (m, 1H, H-6), 5.74 (br s, 1H, N<u>H</u>), 5.71 (m, 1H, H-3), 5.34 (m, 1H, H-4), 3.45 (br s, 3H, OMe in MTPA), 3.32 (m, 2H, H-10), 3.02 (m, 1H, H-7a), 2.96 (m, 2H, H-9), 2.78 (m, 1H, H-7b), 1.87 (s, 3H, H-12), 1.71 (dd, 3H, H-1, *J* = 6.5, 1.6 Hz, H-1); HRESIMS for the [M + Na]⁺ ion at *m/z* 496.1371 (calcd [M + Na]⁺ ion for C₂₂H₂₆NF₃O₅S at *m/z* 496.1376) (Figures 20 and 21).

Supplementary references

- 1. Liu, W.; Christenson, S. D.; Standage, S.; Shen, B. Science 2002, 297, 1170-1173.
- Lohman, J. R.; Huang, S. X.; Horsman, G. P.; Dilfer, P. E.; Huang, T.; Chen, Y.; Wendt-Pienkowski, E.; Shen, B. *Molecular. Biosystems* 2013, 9, 478-491.
- 3. Van Lanen, S. G.; Oh, T. J.; Liu, W.; Wendt-Pienkowski, E.; Shen, B. J. Am. Chem. Soc. **2007**, *129*, 13082-13094.
- 4. Liu, W.; Nonaka, K.; Nie, L.; Zhang, J.; Christenson, S. D.; Bae, J.; Van Lanen, S. G.; Zazopoulos, E.; Farnet, C. M.; Yang, C. F.; Shen, B. *Chem. Biol.* **2005**, *12*, 293-302.
- Horsman, G. P.; Chen, Y. H.; Thorson, J. S.; Shen, B. Proc. Natl. Acad. Sci. USA 2010, 107, 11331-11335.
- Huang, S. X.; Lohman, J. R.; Huang, T.; Shen, B. Proc. Natl. Acad. Sci. USA 2013, 110, 8069-8074.
- 7. Bonnett, S. A.; Whicher, J. R.; Papireddy, K.; Florova, G.; Smith, J. L.; Reynolds, K. A. *Chem. Biol.* **2013**, *20*, 772-783.
- Gilbert, I. H.; Ginty, M.; Oneill, J. A.; Simpson, T. J.; Staunton, J.; Willis, C. L. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1587-1590.
- 9. Gruschow, S.; Buchholz, T. J.; Seufert, W.; Dordick, J. S.; Sherman, D. H. *Chembiochem* **2007**, *8*, 863-868.

Primer name	primer sequence ^a
SgcE-KR forward	5'-AAAACCTCTATTTCCAGTCGGGTGCGGACCCGACG-3'
SgcE-KR reverse	5'-TACTTACTTAAATGTTACCCGGTGAAACGCAGCAG-3'
KedE-KR forward	5'-AAAACCTCTATTTCCAGTCGGGGGCGCGAGGAGGAG-3'
KedE-KR reverse	5'-TACTTACTTAAATGTTAGAACCGCTTCAGCGG-3'
MdpE-KR forward	5'-AAAACCTCTATTTCCAGTCGGCGGTCGCGGGGGCC-3'
MdpE-KR reverse	5'-TACTTACTTAAATGTTACCCCGTGAACCGCAA-3'
NcsE-KR forward	5'-AAAACCTCTATTTCCAGTCGGCCGATGCCGGTGCG-3'
NcsE-KR reverse	5'-TACTTACTTAAATGTTAGCCGGTGAACCGGAG-3'
CalE8-KR forward	5'-AAAACCTCTATTTCCAGTCGGGGGGTCGCGCGTGG-3'
CalE8-KR reverse	5'-TACTTACTTAAATGTTAGGTGAAGCGCAGCAG-3'
DynE8-KR forward	5'-AGGCAATTC <u>CATATG</u> CCCGCCGCCGAGCTGCCC-3'
DynE8-KR reverse	5'-CCG <u>GAATTC</u> TCACTCCAGGAAACGGCCGTG-3'
UcmE-KR forward	5'-AGGCAATTC <u>CATATG</u> GCCGACGAGGCCGCCTCC-3'
UcmE-KR reverse	5'-CCG <u>GAATTC</u> TCACTCCGCGAACCGTCCGTCC-3'

Table S1. Primers used in this study with the NdeI/EcoRI sites used for cloning underlined

^aPrimers used to make the expression constructs for *sgcE-KR*, *kedE-KR*, *mdpE-KR*, *ncsE-KR*, and *calE8-KR* by the ligation independent cloning strategy. Primers used to make the expression constructs for *dynE8-KR* and *ucmE-KR* by direct cloning strategy with the NdeI/EcoRI sites underlined.



Figure S1. The selected seven enediynes and their PKSEs. (A) Structures of four nine-membered (C-1027, KED, MDP, NCS) and three ten-membered enediynes (CAL, DYN, UCM) with their eneidyne cores highlighted in red. (B) Domain organization of PKSEs and their sequence identities (%). The accession numbers for each of the PKSEs are: SgcE (AAL06699), CalE8 (AAM94794), DynE8 (AAN79725), KedE (AFV52145), MdpE (AAQ17110), NcsE (AAM77986), and UcmE (KM017987).

А

	1019 1029 1039 1051 1061 1065 1175
SGGEKR	GADPTAAPVVTOS AAWARP FSVD-LDELPLP PAVAD EKDGTWELFTSADHPFAEEVRRAL
CalE8KR	
KedEKR	
MdpEKR	EGHPLAGPLRDAL
NCSEKR	ADAGAASFIAGAAPWARFFAVD-LDAVARFPARPAAVRGTWELFAPAGYGIAATLRAAL
TylKR6	RAADTDDWMYRIGHDRLPAVTGGART-AGRWLVIHPDSPRCRELSGHAERALRAAGASPVPLPVD-APAADRASFAALLRSATGPDTRGDTAAPVAGVLSLLSEEDRPH
AmphKR2	SRADALRYH I EWNRVAEPGTAR PAGRLLAV I SPDHAGAPWYTAVLDALGPDTVRFEAKGTDRAAWAAQLAQLREDEGEPHAVVSLLAABAELH
AmphKR18	ERSIVDSWRIVVWEPLAQIFRAID-GIWELVSAGGV-DDDVA-EVLEAGGASVREVLD-EIGDFAUERLIDIGS-EIGI-EIGIVSVLAAGGA-DDFAG
TylKR1	-SPTDAWRYRVTWKALTESSPVRPHSIGRCLLVAPPTT-DGELLDGLTTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVLSERGASVARLEVPIGARRAEVAELLKPSMESAGEENTTVVSLLGLTVVSLGLTVVSLGASVARLEVPSAGEENTTVVSLGASVARLEVP
EryKR1	EVDEVSALRYRI EWRPTGAGEPARLDGTWLVAKYAGT-ADETSTAAREALESAGARVRELVVDARCGR-DELAERLRSVGEVAGVLSLLAVDEAEP
	1086 1096 1105 1115 1125 1135 1145 1151
SgcEKR	QDAAVGSGVLVCLPAGCSPDQLELA-LDGARSALAGSQEGRFVLVQHDRGAAGLAKTLHLEAPHLRTTVVHTPVADGAADRVA
Cale8KR DvnE8KR	EAAGVGEGVLVCLPDEPDEEHLVTA-VRGAQAALRQPPGGRLVVVQPAARAGALAKTARLEGDRLKTTVVTTPLDPAAVDRVV VDAGAPAGRLIALPFGLAD I PTADV-VAALRAAAGDRRPLVVLHHGGVGAAVGRS LAVETG-APVLVVETP
KedEKR	AGAGIGDGVLLCLP-ERGDGHDDLI-LAAGKAAAAAGGRFVVVEQRLGAAGLAKTLHLEHPEVTTTVVRLADVMPADPVAAGVAVRRVV
MdpEKR	ERAGAGPGVLVCLPPDCTDDQLEPA-LDGVRDALAGAPGTRFVLVQDGRGAAGLARTLREAPHLRVTIVHTP
UCMEKR	
TylKR6	RQHAPV PAGVLATI.SIMQAMEEERVEAR-VWCVSRAAVAADA.RERP-VGGGALWGLGRVAALERPTR/GGLVDLPASPGAAHWA
AmphKR1	VRHPGLTLGLALSVSLAQALGEADVTAP-LWFLTCGAFSTGFSDTVTRPLQSQIAGLGRVTVAVEHPHRWGGVDLTDEFELVDI
AmphKR18	GAEGVHALTTRVLGYLQEWLSEPRLSGTRLVFVTRRAVALDD-EDVLDPAGAAVWGLVRSAQTENPGSLL-LVDLDDTFLSAGGGAAVWGLVRSAQTENPGSLL-LVDLDDTFLSAGGGAAVWGLVRSAQTENPGSLL-LVDLDDTFLS
TylKR1 EryKR1	vpstdavrtsialloavsdiovsdiovarvmaltrravavvfoordp-odagolumgforvaalelpdimggituditertersetsetsetsetsetsetsetsetsetsetsetsetsets
	1161 1170 1180 1190 1198 1208 1218 1227 1237 1247 1257
	· · · · · · · · · · · · · · · · · · ·
SGCEKR CalE8KR	AEVA-ATTHFSEVHLDRDGTRRVPVLRALPFAP-DRTDQV-LGPDDVLLVTGGGKGITAECALAVAERCG-AALAVLGRSDPGSDQDLAANLGRMRESGIRVAXAADVT ADVA-ATTHFSEVHLDRDGTRRVPVLRALPFASDGEPGALP-LGPDDVLLVTGGGKGITAECALMAERSG-ARLAVLGRSDPTADEALADNIKRLADASDLRVLRVDVT
DynE8KR	VEASRPWSGYQEVVYGADGVRTVPVTRLLDGQPRTDRRIP-LGPGEVCLVT <mark>GGAKGIGAACAAALGEAT</mark> G-ARLVLFGRS-PADDPEVVEAVRRTGAVYRSVDLA
KedEKR	ARAA-ATTGFAEVRIDES GARTTPRLRALPVHAED GAAQAALGPDDVLLVAGGGKGITAE CALALAKDSG-AALALLGRDDPADH PELAAN LARVADLAGVRIR YLKVDVT
NosEKR	AEVS-ATTRFSEVHYSAD GVRRVPTLRALPMSP-EQQDKP-LSASDVLLVTGGGKGI SAECALAIAQDSG-TRLAVLGRSDPATD RELADNLKRMED SGVTMR YARADVT
UCMEKR	ABTA-AARGFTEARYRAD GTRLTPALAPVPLPAAFOPAL SADDVVLVTGGGGG I TAE CARMLAAGS GGAR I AVLGRTRFEDD BSLRENLARLOPDVR YT SADVT
UcmEKR TylKR6 AmphKR2	AETA-AARGFTEARYRAD GTRLTPALAPVPLPAAPGPAL SADDVVLVT <mark>GGGRG I TAECARMLAA GS</mark> GGAR I AVLGRTRPEDD ESLRENLARLGPDVR YT SADVT AAVERLAGPEDQ I AVRAS OSWGRLTRLPRDGGGR-TAAPAY RPRGTVLVT GGT GGLGGRLARMLAAAGAEHLALTS RRGPDA FGAAGLEAELLL LGAKVT FAACDTA GLLADAGGED -QLAI FGS GVLARRLAHAAPA PVGSGRFP PVHGSVLVT GGT GGI GGRVARRLAEGG'ABHLVLTS RRGADA FGAAEL RAELEQ LGFWT I AACDAA
UcmEKR TylKR6 AmphKR2 AmphKR1	AETA-AARGFTEARYRAD GTRLTPALAPVPLPAAPGPAL SADDVVLVT <mark>GGGRG I TAE CARMLAA GS</mark> GGAR I AVLGRTRPEDD ESLRENLARLGPDVR YT SADVT AAVERLAGPEDQ I AVRAS GSWGRLTRLPRDGGGR-TAAPAY RPRGTVLVT GGT GGLGGRLARNLAAAGAEHLALTS RRGPDA PGAAELAELDL LOAKVT FAACDTA GLLADAGED-QLALRGS GVLARRLAHAAP AVFGSGRFP VHGSVLVTGGT GG I GGRVARRLAEQGAAHLVLTS RRGADA PGAAELRAELEQLGRVT I AACDAA AALAGALGDDDQLAVRPA GVLARR I VRASGDTRRKARSWKPRGTTLVTGGS GTLAPGLARNLAAQGAEHLVLS RRGADA PGAAELAAELQAAGTEVR FAACD I T
UcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1	ABTA-AARGPTEARYRAD OTR.TP.AT.APVP LPAAPGPAL SADDVVLVT GOGGI TABCARNILAAGSGGAR IAVLGRTRPEDD ESLRENLARLOPDVR YT SADVT AAVERLAGPEDQI AVNAS OSWGRRLTRLEPEDGGGRT TAAPAYRPGTVLVT GOT GOLGGI TABCARNILAAGSGGAR IAVLGRTRPEDD ESLRENLARLOPDVR YT SADVT GLLADAGGED-GLAIRGS GVLARRLAHAPAVPGSGRP PVHOSVLVT GOT GOLGGARVARRLABGGABHLVLTS RRGADA PGAAGLERELELLGAKVTY TAACDTA AALGGALGDDDQLAVNPA GVLARR IVRAS GDTRRTKARSWIRFGTTLVT GOT GOLGGARIARHLAGGABHLVLTS RGADA PGAABLARELEQLGAKVTY FAACDTA VLPDVLTLDEQQLAVND Y QVRARRLARLPRPADDAPAADWNPDGTVLIT GOT GOLGARIARHLVT SKG-ARHLLLS SRGADA PGAABLARELGAAGTVYV HAACDVA AVLGGDED_GOLAVND Y QVRARRLARLPRPADDAPAADWNPDGTVLIT GOT GOLGARIARHLVT SKG-ARHLLLS SRGADA PGAABLARELEGALGAKVYV HAACDVA
UcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1	ABTA-AARGFTERRYRAD OTRLTPALAPVPLPAAPGPALSADDVVLVT GOGTGALGGITABCARRILAAGSGGAR IAVLGRTPEDDESLRENLARLGPDVRYTSADVT AAVERLAGPEDQIAVRAS OSWGRILTRLPEDGGGR-TAAPAYRRGTVLVT GOTGALGGITABCLARGGAEHLALTS RRGPDAPGAAGLERELLLLGAKVYFYRACDTA GLLDAGGED-GLAIRGSGVLARRLAHAAPAVPCSGRPPVHGSVLVTGGTGGIGGRVARRLAEGCAEHLALLSSRGADAPGAAGLERELLLGAKVYFYRACDTA AALGGLGDDGQLAVRPAGVLARRLYRASGDTRRKARSWKRGTTLVTGGSGTLAPGIARHLAYGAEHLVLSSRGADAPGAAELARELGAADTEVYRACDTA VLPDVLTLDEQQLAVRDYQVRAARLARLPRPADDAPAADWNPDGTVLITGGTGGLGAALARHLVTSRG-ARHLLLSSRGAPAPGAAELARELGAADTEVYRACDTG AVLGGDCDC_VQVRASGIGGTIGGRVSRAAAAGASSWQPSGTVLITGGTGGLGAALARHLVTSRG-ARHLLLSSRGAPAPGABELARELEGAGTEVVHAACDVG AVVSGGACGD-QLALRADGVYGRWVRAAAPATDSWQPSGTVLITGGTGGUGAALARHLVTSRG-ARHLVLSSRGAPAPGABELARELERGHGGEVVHAACDVG
VemEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1	ABETA-AARGFTERRYRAD OTRLTPATADYP LPAAPGPAL SADDVVLVT OGGRGT TAECARNILAAGS GGAR IAVLGRT PEDDESLRENLARLOPDVRYTSADVT AAVERLAAGS FGGARLTRLPPEDGGGR-TAAPAPAY RPRGTVLVT OGTGALGMILAAGS GGAR IAVLGRT PEDDESLRENLARLOPDVRYTSADVT AAVERLAAPED IAVRAS GSWGRRLTRLPPEDGGR-TAAPAPAY RPRGTVLVT OGTGALGMILARMILAAGSAEHLALTS RROPDAPGAAGLERELLLGAKVY FAACDTA CLLDADGGED-GLAIRGS GVLARR LANAAPAVPCSGRP PVHGSVLVT OGTGGIGGRVARRILAEGOAEHLAVLTS RROADA PGAAGLERELLGAKUT FAACDTA AALAGALGDDDQLAVRPA GVLARR LANAAPAVPCSGRP PVHGSVLVT OGTGGIGGRVARRILAEGOAEHLAVLS RROADA PGAAGLERELEGAGTEVR FAACDI T VIPDVITLDEQQLAVRDT GVRARRIARLPR PADDAPAADWNPDGTVLIT OGTGGLGAALARHLVTSRG-ARHLULLS RROPPA PGAELAELGAAGTEVR FAACDI T VIPDVITLDEQQLAVRDT GVRARRIARLPR PADDAPAADWNPDGTVLIT GOTGGLGAALARHLVTSRG-ARHLLLAS RROPPA PGAELAELEGHGGEVVHAACDVA AVLAGRDEED-GVAVRAS GI TGRVSRAAAAGASWQPSGTVLIT GOTGGVOGGI IARNILARGAEHLULSS SROPPADGAGEU LAELISALBAGCEVVHAACDVA AVVSGGAGED-GLALARD GVYGRWVRAAAPATDBEWLFTGTVLVT GOTGGVOGGI IARNILARGAEHLUSS SGPLADGAGEU VELELALABRTTVAACDVA 1267 127 1286 1316 1326 1334 134 1354 1364
VcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1	ABTEARGFTERRINAD OTRLTPALLAPVPLPAAPGPAL SADDVVLVT GOGGITAECARNLAAGSGGGR LAVLGRTRPEDDESLERNLARLOPDVRITSADVT AAVERLAGFEDQ LAVRAS OSWGRRITKLEPALAPVGLOGGT-CAAPAY RPGGTVLVT GOGGALOGRITAECARNLAAGCAEHLALTSRGPDAPGAAGLERELLLGGKVVFAACDTA GILADAGGED-QLAIRGS GVLARRLAHAAPAVPGSGRP PVHGSVLVTGGTGGIGGRVARRLAEGCAEHLAUTSRGADAPGAAGLERELLGGKVVFAACDTA GALGALGDDDQ LAVRPAGVLARR IVRASGDTRRKARSWKPRGTTLVTGGSGILAPGLARHLAAQCAEHLAUTSRGADAPGAAELRAELQAGGTEVKPAACDTA VLPDVLTDEQQLAVRPAGVLARR IVRASGDTRRKARSWKPRGTTLVTGGSGILAPGLARHLAAQCAEHLAULSSRGADAPGAAELRAELQAGGTEVKPAACDTA VLPDVLTDEQQLAVRPAGVLARR ILRLPRPADDAPAADMNPDGTVLITGGTGGLARHLARHLATSG-ARHLLLSSRGPDAPGTSELVAELTGIGAQUTVSACDVG AVVSGGAGED-QUALRADGVYGRWVRAAAPATSWQPSGTVLITGGTGGGVGGQIARRLARRGAEHLLUSSRGPDAPGAAGLEAELGAGGEVVHAACDVA AVVSGGAGED-QUALRADGVYGRWVRAAAPATDEWKPTGTVLVTGGTGGVGGQIARRLARRGAEHLLVSSGPDADGAGELVAELBALGARTTVAACDVT 1267 1277 1287 1296 1316 1326 1334 1354 1364
UcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1 SgoEKR CalE8KR	ABETA-AARGFTERKIKAD OTRITPATAPVP LPAAGGPAL SADDVVLVT GOGGITABCARNILAAGSGGAR IAVLGRTREDDE SLRENLARLOFDVRITSADVT AAVERLAGFEDQIAVRAS OSWGRRLTRIPERDGGR-TAAPAIRPROTVLVT GOTGALGGHTABCARNILAAGSGGAR IAVLGRTREDDE SLRENLARLOFDVRITSADVT GLIADAGGED -QLAIRGS GVLARRLAHAPAVPGSGRP PVHGSVLVT GOTGALGGHTARNILABGCAEHLALTS RRGDAPGAAGLERELLLIGAKVITPACOTA GLIADAGGED -QLAIRGS GVLARRLAHAPAVPGSGRP PVHGSVLVT GOTGGIGGRVARRLABGCAEHLALTS RRGDAPGAAGLERELDQAGTEVYFAACDT VLPDVLTLDEQQLAVRDY QVRARLARLPRPADDAPAADWNPDGTVLIT GOTGGIGAALARHLYTSRG-ARHLLLS SRGDAPGAAELARELGAAGTEVYHAACDVA AVLGGDED -QVAVRAS GIYGRRVSRAAAAGADAADWNPDGTVLIT GOTGGIGAALARHLYTSRG-ARHLLLS SRGDAPGAAELAELGAGTEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKFTGTVLVTGOTGGVGGQIARNILARRGAEHLLUS SRGPDAPGAAELAELGAGTEVVHAACDVA AVVSGGAGED -QLALRAD GYYGRRWVRAAAPATDDEWKFTGTVLVTGOTGGVGGQIARNILARRGAEHLLUS SRGPDAPGAAELAELGAHGEVVHAACDVA 1267 1277 1287 1296 1306 1316 1326 1334 1344 1354 1364
UcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1 SgoEKR CalE8KR DynE8KR	ABETA-AARGFTERRYRAD OTRLTPATAPVPLPAAPGPALSADDVVLVT OGGRTABCARNILAAGSGGAR IAVLGRTRPEDDESLRENLARLOPDVRYTSADVT AAVERLAPEDQIAVRAD SONGRTLTRLPPLOGGR-TAAPAPYRPROTVLVTOGTGALGGHTABCLAAGSGGAR IAVLGRTRPEDDESLRENLARLOPDVRYTSADVT GLLDADGGED-QLLAIRGSGVLARRLANAPAVPCSGRPPVNGSVLVTGGTGGIGGRVARRLAEGGAEHLALTSRGPDAPGAAGLEAELLLLGAKVYFPAACDTA GLLDADGGED-QLAIRGSGVLARRLANAPAVPCSGRPPVNGSVLVTGGTGGIGGRVARRLAEGGAEHLVLSRGADAPGAAELARELQADGEVYFPAACDTA VLPDVLTLDEQQLAVRDYQVRARRLARLPPADDAPAADWNPDGTVLITGGTGGIGAALARHLYTSRG-ARHLLLSSRGADPGAAELARELQADGEVYFVAACDVA AVLSGGAGED-QULAVRDYQVRARRLARLPPADDAPAADWNPDGTVLITGGTGGUGAALARHLYTSRG-ARHLLLSSRGPDAPGTSELVAELFGLGAQVYVSACDUG AVVSGGAGED-QLAIRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGGTGGVGGQIARWLARRGAEHLVLSSRSGPDADGAGELVAELFGLGACTVVSACDVG 1267 1277 1287 1296 1306 1316 1326 1334 1344 1354 1364
UcmEKR TylKR6 AmphKR2 AmphKR18 TylKR1 EryKR1 SgcEKR CalE8KR DynE8KR KedEKR	ABETA-AARGFTERRYTAD OTRLTP ALAPYP LPAAGRAL SADDVULVT GOGRGITABCCARNLAAGSGGGR IAVLGRTREDDE SLAENLARLGPDVRYTSADVT AAVERLAGPEDGI INVRAS GSWGRRLTRLPEDGGGR-TAAPAR PROTVULVT GOGRGITABCCARNLAAGSGGAR IAVLGRTREDDE SLAENLARLGPDVRYTSADVT GLIADAGGED-GLAIRGS GVLARRLANAAPAVPCSGRP PYNGSVLVT GOT GGI GGRVARRLAEGCAEHLAITS RROPDA PGAAGLEAELLLLGAKVY FAACDA AALAGALGDDDQLAVRD GVLARRLANAAPAVPCSGRP PYNGSVLVT GOT GGI GGRVARRLAEGCAEHLAITS RROPDA PGAAGLEAELLLGAKVY FAACDA AALAGALGDDDQLAVRD GVRARRLANAAPAVPCSGRP PYNGSVLVT GOT GGI GGRVARRLAEGCAEHLVLS RROPDA PGAAGLEAELLGAAGTEVR FAACDI T VIPDVITLDEQQLAVRD GVRARRLARLAPRAPADAPAADWNPDGTVLI TGOT GGLGAALARHLVTSRG-ARHLULLS RROPDA PGTSELVAELTG LGAAGTEVR FAACDI T AVLGRDGED-GVAVRAS GI VGRVSRAAAGASWGPSGTVLI TGOT GGLGAALARHLVTSRG-ARHLULS RROPDA PGTSELVAELTG LGAAGTEVR FAACDI T AVVSGGAGED-GLALRAD GVVGRWVRAAAPATDEWNFTGTVLVT GOT GOT GOV GOT LARNLARGAPHLLVS RSOPDADGAGELVAELEALGARTTVAACDVT 1267 1277 1287 1296 1306 1316 1326 1334 1344 1354 1364
UcmEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1 EryKR1 SgcEKR CalE8KR DynE8KR KedEKR MdpEKR NcsEKR	ABETA-AARGFTEARTRAD OTRITPATAPVP LPAAGRAL SADDVVLVT 000G TABECARVLAAOS GGAR IAVLGRTREDD ESLRENLARLOFDVXITSADVT AAVERLAAGFEDQIAVRAS OSWGRRLTRIPERDGOGR-TAAPAIYRRGTVLVT 000 CALGGMILARVLAAGSAEHLALTS RROFDA PGAAOL RAELLL LGAKVTFAACDTA GLLADAGGED -QLAIRGS GVLARRLAHAAPAVPGSGRP PVHGSVLVT 000 GIGGRVARRLAEQGAEHLVLTS RROADA PGAAOL RAELLL LGAKVTFAACDTA AALAGALGDDQLAVRPA GVLARR IVRAS ODTRAXARNKPRGTLVT 000 GIGGRVARRLAEQGAEHLVLS RROADA PGAAOL RAELQAATGVVF AACDTA VLPDVLTLDEQQLAVRD Y QVRARRLARLPRPADDAPAADWNPDGTVL IT 000 GOTLAPGILARVLAVGGAEHLVLS RROADAPGAAELAELGAAGTGVVFAACDT VLPDVLTLDEQQLAVRD Y QVRARRLARLPRPADDAPAADWNPDGTVL IT 000 GGLGAALARVLYT SRG-ARHLLLS SRGDPDA PGAAELAELGAAGTGVVFAACDTV AVVJGGAGED -QUAVRAS GI YGRRVSRAAAAGADAEWNPDGTVL IT 000 GGLGAALARVLYT SRG-ARHLLLS SRGPDA PGAAELAELGAGTGVVFAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWNFTGTVLVT 000 GGVGQQ IARVLARRGAERLVLTS SRGPED PGAAELAELGAGTGVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWNFTGTVLVT 000 GGVGQQ IARVLARRGAERLVLTS SRGPED PGAAELAELGAGTGVVHACDVA AVVSGGAGZED -QLALRAD GVYGRRWVRAAAPATDDEWNFTGTVLVT 000 GGVGQQ IARVLARRGAERLVTS SRGPED AGAELAELGAGTGVVHACDVA AVVSGGAGZED -QLALRAD GVYGRRWVRAAAPATDDEWNFTGTVLVT 000 GGVGQQ IARVLARRG AERLVITS SRGPED AGAELAGELGAGTTVAACDVT DEWNYAGAAVATVTADAGTGS TAVLHGA-GNNEPFALGGLDMAAVRSTIAP DOLGAVAAVATVTADAGT ATANEWLAG GTEDVARNPDC DAGAVAAAVATVTADEGVFTAVLHGA-GNNEPFALGGLDMAAUSTIAP DOLGAVAAVATVTADAGET ALAUFTG DAGAVAAAVATVTADEGVFTAVLGA-GNNEPFALGGLDMAALSGVFTAR DOLGAVADAVATVTADAGET ALAUFTG DAGAVAAAVATVTADEGVFTAVLHGA-GNNEPAALADLDEALSGVFTAR VOGLRAVLADVDPARLELVTFG II GRAGLGEAM AAANELAELTEEVAATRPC DPAVAAAVATVTADEGVFTAVLGA-GNNEPAPTULALLEGUFTATLAP AVGFEVVAALDPARLELVTFG II GRAGLGEAM AAANELAELAELTEEVAATRPC DPAVAAAVATVAGED FTGULAGA-GNNEPAGLAALSGVFTAR VTGLRAVLADVDARALULITFG VI GRAGLGEAM AATANEVLS SRCAELAATVPQV PREVRYDAVARAELGEVTAVLYGA-GNNEPAGLADLEHERTTAP TGGLAVLADVDARALULITFG VI GRAGLGEAM AATANEVLS SRCAELAATVPQV PREVRYDAVARAELEFELGFVTAVLGA-GNNVPAGLADLEHEFELRTIAP TGGLAVLADVDARALULITFG VI GRA
UcmEKR TylKR6 AmphKR1 AmphKR1 AmphKR18 TylKR1 EryKR1 SgcEKR CalE8KR MdbEKR MdbEKR McsEKR UcmEKR TylK6	ABETA-AARGFTERKIKAD OTRITPATAPVPLPAAGGPAL SADDVVLVT GOGGITABCARNILAAG SGGRA IAVLGRTRPEDD ESLRENLARLOFDVKITSADVT AAVERLAGFEDQI AVNARS OSWGRALTRILPEDGGGR-TAAPATRRGTVLVT GOTGALGGITABCARNILAAGAEHLALTS RRGPAD FGAAGLERELLLIGKVVT FAACDTA GLIADAGGED_GLIAIRGS GVLARRLAHAAPAVPGSGRP PVHOSVLVT GOTGALGGITABCLAGGAEHLALTS RRGPAD FGAAGLERELLILGKVVT FAACDTA AALAGALGDDDQLAVRPA GVLARR IVRASODTRAKARSWIRFGTTLVT GOTGALGGIAATARHLAKGCAEHLALTS RRGPAD FGAAGLERELLGAGTEVYFFAACDTA VLPDVLTLDEQQLAVRD Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOTGGIAAATARHLVTSRG-ARHLLLS RRGAD FGAAELARELEQLGAGTEVYFAACDTA VLOGDDQLAVRD Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOTGGIAAATARHLVTSRG-ARHLLLS RRGAD FGAAELARELEGAAGTEVYFWAACDVA AVVSGGAGED_QUAVRAS GI YGRRWYRAAAPATDDEWKPTGTVLVTGOTGGVGGQ IARNIARRGAEHLULTS RRGAPE FGAAELAEELGGHGEVVHAACDVA AVVSGGAGED_QLALRAD GVYGRRWYRAAAPATDDEWKPTGTVLVTGOTGGVGGQ IARNIARRGAEHLULTS RRGAPE FGAAELAEELGGHGEVVHAACDVA AVVSGGAGED_QLALRAD GYYGRRWYRAAAPATDDEWKPTGTVLVTGOTGGVGGQ IARNIARRGAFHLLLS RRGAPE FGAAELAEELGHGEVVHAACDVA AVVSGGAGED_QLALRAD GYYGRRWYRAAAPATDDEWKPTGTVLVTGOTGGVGGQ IARNIARRGAFHLLLYS RSGPDADGAGELVAELEALGAGTEVVHACDVA DYGGAAGTAELTGALGS VTAVLHGA-GRNEPTALGGLDMAAVRSTIA PVDGLRHVLDVVGEONLELLVT FOG I IGRACLRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVTATVAEKGPVTAVLHGA-GNEPTALGGLDMAALRSVYAA VDGLRAVLAVDPARLLLVT FOG I I GRACLGRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVTATAEELGPVTAVLGA-GNEPTALGGLDMAALRSVYAA VDGLRAVLAVDPARLLLVT FOG I I GRACLGRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVTATAEELGPVTAVLGA-GNEPTALGGLDMAALRSVYAA VDGLRAVLAVDPARLLLVT FOG I I GRACLGRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVTATAEELGPVTAVLGA-GNEPTALGGLDMAALRSVYAA VDGLRAVLAVDPARLLLVT FOG I I GRACLGRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVATATAEELGPVTAVLGA-GNEPEPTLAGELERTTAFP VTGLRAVLAVDPARLLLVT FOG I I GRACLGRGAPH AATANEWLAGLTED VARRMPD C DAGAVAAAVATATAEELEPTAUGGLDVAALGA-GNEPEPTLAGTED LRTTLAP VTGLRAVLAVADAEELLLVT FOG I I GRACLGRGAPH AATANEWLAGLTET AGRNPD C DAGAVAAAVATATAEELEPTAGGPVTAVLGA-GNEPEPTLAALLFDTP TYGLGRAVLAAVDPARLLLIT FOG V
UcmEKR TylKR6 AmphKR1 AmphKR1 AmphKR18 TylKR1 EryKR1 SgcEKR CalE8KR McbEKR McbEKR NcsEKR UcmEKR TylKR6 AmphKR2	ABETA-AARGFTERRYRAD OTRITPTALAPVPLPAAPGPAL SADDVVLVT GOGTGALGGTTABCARNILAAG SGGAR IAVLGRTRPEDD ESLRENLARLOPDVRYTSADVT AAVERLAPEDQI AVRAS OSWGRTLTRIPPEDGGRT TAAPAYRRGTVLVT GOTGALGGTTABCLARGCAEHLALTS RRGDAPGAAGLERELLLLGAKVYTFAACDTA GLLADAGGED-QLAIRGS GVLARRLANAPAVPCSGRPPVNGSVLVT GOTGALGGTARNILAGCAEHLALTS RRGDAPGAAGLERELLLGAKVYTFAACDTA AALGGLADDQDQLAVRPAGVLARR IVRASODTRRKARSWKRGTTLVT GOTGGTLAPGILARNILAGCAEHLALTS RRGDAPGAAGLERELLGAKGTVYFAACDTA VLPDVLTLDEQQLAVRDY QVRARRLARLPRPADDAPAADWNPDGTVL IT GOTGGGLAAILARHLYTSG-ARHLLLS RRGDAPGAAELARELGAAGTVYFRACDT VLPDVLTLDEQQLAVRDY QVRARRLARLPRPADDAPAADWNPDGTVL IT GOTGGLAAILARHLYTSG-ARHLLLS RRGDPGAAELAELGAAGTVYFAACDT AVLSGGDGD-QUAVRAS GI YGRYVSRAAAAGADDEWKPTGTVLVTGGTGGVGQQ I ARWLARRGAEHLVLTS RRGPERFGAAELAEELGHGCEVVHAACDVA AVVSGGAGED-QLAVRAS GI YGRYVSRAAAAAADDEWKPTGTVLVTGGTGGVGQQ I ARWLARRGAEHLVLTS RRGPERFGAAELAEELGHGCEVVHAACDVA 1267 1277 1287 1296 1306 1316 1326 1334 1344 1354 1364
UcmEKR TylKR6 AmphKR1 AmphKR18 TylKR1 EryKR1 SgcEKR CalE8KR KedEKR McbEKR NcsEKR UcmEKR TylKR6 AmphKR1 AmphKR2	ABETA-AARGFTERRYTAD OTRLTP ALAPYP LPAAPGPAL SADDVULVT GOGTGALGGITABCARNILAAGSGGGR IAVLGRTPEDD ESLRENLARLGPDVRYTSADVT AAVERLAPEDQI AVRAS OSWGRTLTRLPPEDGGGF. TAAPAY RPRGTVLVT GOTGALGGITABCLARGAEHLALTS RRGPDA PGAAGLEAELLLLGRAVT FRACDTA GLLDADGGED-GLAIRGS GVLARRLANADRVPCSGRP PYNGSVLVT GOTGGIGGRVARRLAEGGAEHLALTS RRGPDA PGAAGLEAELLLLGRAVT FRACDTA AALGGLGDDGLAVRGS OSWGRTLVTAGTGGIGGTVARRLAEGGAEHLALSSRGDPA PGAAGLEAELLLGRAVTY FRACDTA ALGGDDGDLAVRDY QVRAARLARLPRPADDAPAADWNPDGTVLIT GOTGGIGGAVARRLAEGGAEHLVLTS RRGDPA PGAAELAELLGAAGTEVYF FAACDI T VIPDVLTLDEQQLAVRDY QVRAARLARLPRPADDAPAADWNPDGTVLIT GOTGGIGGAVARRLAEGGAEHLVLTS RRGPDA PGAAELAELLGAAGTEVYF FAACDI T AVLGRDGED-GVAVRAS GI VGRVVSRAAAAGASWQPSGTVLIT GOTGGIGGALARHLVTS RG-ARHLLLAS RRGPDA PGTSEL VAELTGLGAQVTVSACDVG AVLGRDGED-GVAVRAS GI VGRVVSRAAAAGASWQPSGTVLIT GOTGGIGGALARRLARGAEHLLUS SRGPDADGAGELVAELEAGAGTEVVHAACDVG AVVSGGAGED-GLALRAD GVVGRWVRAAAPATDSWQPSGTVLIT GOTGGIGGALARRLARGAEHLLUS SRGPDADGAGELVAELEAGAGTEVVHAACDVG 1267 1277 1287 1296 1306 1316 1326 1334 1344 1354 1364
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 CalE8KR CynE8KR CalE8KR MdpEKR NosEKR UcenEKR TylKR6 AmphKR2 AmphKR1 AmphKR1	ABETA-AARGPTERKIKAD OTRITPATAPVP LPAAPGPAL SADDVVLVT GOG GITAB CARNILAAG SGAR LAVLGRTRPEDD ES LRENLARLOPDVKITSADVT AAVERLAGPEDQ I AVRAS OSWGRTLTILPEDGGGRT TAAPATRRTGTVLVT GOT GALGGHTABCLARNILAAGAEHLALTS RRGDAPGAAGLEAELLLLGAKVITY FAACDTA GLLDAGGED -QLAIRGS GVLARRLAHAAPAVPGSGRP PVHGSVLVT GOT GGIGGRVARRLAEGGAEHLALTS RRGDAPGAAGLEAELLLLGAKVITY FAACDTA AALAGALGDDDQLAVNPA GVLARR LYNAS ODTRAKARNIKPRGT LIVT GOS GILAPGILARHLAEGGAEHLALTS RRGDAPGAAGLEAELLGAAGTEVYFRACDTA VLPDVLTLDEQQLAVND Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOT GGIGAALARHLVTS RG-ARHLLLS SRGDAPGAAELAELEQAAGTEVYHACDVA AVVSGGAGED -QUAVRAS GI YGRRVSRAAAAGAADADWNPDGTVL IT GOT GGIGAALARHLVTS RG-ARHLLLS SRGDAPGAAELAELEGAAGTEVYHAACDVA AVVSGGAGED -QUAVRAS GI YGRRVSRAAAAGADDEWKPTGTVLVT GOT GGUGGQI IARNIARGAEHLULTS SRGPED FGAAELAEELEGHGECVYHAACDVA AVVSGGAGED -QUAVRAS GI YGRRVSRAAAAAATDDEWKPTGTVLVT GOT GGUGGQI IARNIARGAEHLULTS SRGPED FGAAELAEELEGHGECVYHAACDVA AVVSGGAGED -QUAVRAS GI YGRRVSRAAAAGATDDEWKPTGTVLVT GOT GGUGGQI IARNIARGAEHLULTS SRGPED GAAELAEELRGHGECVYHAACDVA AVVSGGAGED -QUAVRAS GI YGRRVSRAAAAGATDDEWKPTGTVLVT GOT GGUGGQI IARNIARGAEHLULTS SRGPED GAAELAEELRGHGECVYHAACDVA AVVSGGAGED -QUAVRAVGT SRGVGGA GANEPTALGGLDMAAVSTI TA PVDGI MAVADAYATYTADWGP TAVLFAGA GANEPTALGGLDMAAVSTI TA PVDGI MAGAGANATYTADWGP TAVLFAGA GANEPTALGGLDMAAVSTI TA PVDGI MAVAGT TADWGP TAVLFAGA GANEPTALGGLDMAAVSTI TA PVDGI MAVGAVATTATAWF TANGA GANEPTALGGLDMAAVSTI TA PVDGI MAAGAAYATYTADWGP TAVLFAGA GANEPTALGGLDMAA KSTI AP NYGELANJAVGV MARAYATYTADWGP TAVLFAGA GANEPTALGGLDMAA LSGVTAA VDGA TAVGA AYATYTADWGP TAVLFAGA GANEPTALGGLDMAA KSTI AP NYGELANJAVGV MAGAGANATTATAWKI SKI SKATAKYATA PVGVAVAAVATAELGG PVTAVLFAGA -GNEPFALAGLDFDE LRTI.AP NYGELAVLAVDPAR LILLIVT FO'I I I GRAGLGRGAH ATANEKIJAE LISELTSEVAATPEC DPAVAAAAYATYTADWGP TAVLFAGA -GNEPFAPAVTDLI TAEE LIKTI.AP NYGELAVLAVDPAR LILLIVT FO'I I I GRAGLGRGHAH ATANEKIJAELSELARTVPQV PEAVAAAVATAELEG PVTAVLFAGA -GNEPFAPAVTDLI TAEE LIKTI.AP NYGELAVLAVDPAR LILLIVT FO'I I I GRAGLGRGHAH ATANEKIJAELSELA
UcenEKR TylKR6 AmphKR2 AmphKR1 AmphKR18 TylKR1 EryKR1 SGCEKR CalE8KR DynE8KR MdpEKR MdpEKR McsEKR UcenEKR UmEKR AmphKR2 AmphKR18 TylKR1	AREA-AARGFTERKIKAD OTRITPALAPVPLPAAPGPAL SADUVLVT OGGITABCARNILAAG SGAR I AVLGRTREDD ESLENLARLOPDVRITSADUT AAVERLAGFDED I AVNARS OSWGRTLTILPEDGOGGT TAAPATRRTGTVLVT OGGIGLAGALARLEAGAGAEHLALTS RRGDA PGAAGLERELLLIGAKVITFAACDTA GLIDADGGED -QLAIRGS GULARRLAHAAPAVPCSGRP PYNGSVLTT OGGIGLGRVARHLAGGAEHLALTS RRGDA PGAAGLERELLIGAKUTFAACDTA GLIDADGGED -QLAIRGS GULARRLAHAAPAVPCSGRP PYNGSVLTT OGGIGLGRVARHLAGGAEHLALTS RRGDA PGAAGLERELLIGAKUTFAACDTA AALAGALGDDDQLAVNPA GULARR IVNRASODTRAKARSWIRFGTTLITT OGGIGLGRATARHLAGGAEHLULS RRGDA PGAAELARELQAAGTEVYFRACDI T VLPDVLTLDEQQLAVND I QWRARLARLPRPADDAPAADWNPDGTVL IT OGGIGLGRATARHLAVS G-ARHLLLS RRGDA PGAAELARELGAAGTEVYHRACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVS G-AERLULTS RRGPEA PGAAELAEELGAGGEVYWRACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVS G-AERLULTS RRGPEA PGAAELAEELGHGGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVS G-AERLULTS RRGPEA PGAAELAEELGHGGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVS G-AERLULTS RRGPEA PGAAELAEELGHGGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVS G-AERLULTS RGGELGAAF AAVVSGGAGED -QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVT OGGIGLGRATARHLAVSGGAGAAELGAAFTAN DEWKAAAVATYTADAGFTALGA-GRIPEPTALGGLDMAAVRSTIAP FVDGLRAVLAVVGEONLELLVT FG I 10RACLGRAF AANEWLAGLEED VARNPDC DAGAVAAAVATT TADKGPVTAVLHGA-GRIPEPTALGGLDMAAVRSTIAP FVDGLRAVLAVVGEONLELLVT FG I 10RACLGRAFH AANEWLAGLEED VARNPDC DAGAVAAAVATTATADKGPVTAVLHGA-GRIPEPTALGGLDMAAVRSTIAP FVDGLRAVLAVDPARLELLVT FG I 10RACLGRAFH AANEWLAGLEED VARNPDC DAGAVAAAVATTATADKGPVTAVLHGA-GRIVPAALADLEALGLDWE TVGLGAVLAVDPARLELLVT FG I 10RACLGRAFH AANEWLAGLEKE FAGRHPC DPGVVARAVAELESLGPVTAVLGA-GRIVPACLAALGDED FERTAFP FVDGLRAVLAVDPARLELLVT FG I 10RACLGRAFH AANEWLAGLEKE FAGRHPC NPEQVRTAVAELKEGDPVTAVLAGA-GRIVPACLALDDHE HKTLAP TYGGLRAVLAVDPARLELLIT FG VI 10RACLGRAFH AANEWLABTLAELAET FER DPTAVAELASGLEPVTAVLGA-GRIVPACLALDDHE HARTLAP TYGGLRAV
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR18 TylKR1 EryKR1 SGCEKR CalE8KR DynE8KR MdpEKR MdpEKR MdpEKR MdpEKR McsEKR UcenEKR AmphKR2 AmphKR18 TylKR1 EryKR1	AAVGATEARAKAAD OTRITPALAPVPLPAAGGAL SADDVULVI OGGATAACAANGGGAR IAVLGATAREDDE SLERNLARLOPDVRITSADVT AAVGALAPEDQI AVARAS OSWGARLITAIEPAGOGAT CAAPARARGTVULVI OGGALGIGATAACAAAGGGGAR. IAVLGATAREDDE SLERNLARLOPDVRITSADVT AAVGALAPEDQI AVARAS OSWGARLIANEPAGADA SAWARAY AVARAS OF AN AVARAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA
UcenEKR TylKR6 AmphKR2 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR MdpEKR MdpEKR MdpEKR UcenEKR UcenEKR UcenEKR TylKR2 AmphKR1 EryKR1 EryKR1	ABER-AARGFTERRYRAD OTRLTPFALAPYPLPAAPGPAL SADUVLYT COGTGALGGTTAECARNLAAGS GGAR LAVLGRTPEDD ESLRENLARLOPDVRYTSADUT AAVERLAAGFEDG LAVNAS OSWGRTLTNLFPLOGGGT TAAPARPROTUVLT COTGALGGTHARNLAAGS AEHLALTS RROPDA PGAAGLAELLLLGAKVTFAACDTA GLLADAGGED-QLAIRGS GVLARRLANAPAVPCSGRPPVNGSVLTCOTGGIGGTVARRLAEGCAEHLALTS RROPDA PGAAGLAELLLLGAKVTFAACDTA AALAGALGDDOQLAVRPA GVLARRLANAPAVPCSGRPPVNGSVLTCOTGGIGGTVARRLAEGCAEHLULS RROAD PGAAELAELLQAACTEVYFAACDTA VLPDVLTLDEQQLAVRDY QVRARLARLPPADDAPAADWNPDGTVL IT COTGGIGAALARNLLVTSR-ARHLLALS RROPDA PGAAELAELLGAACTEVYFAACDTA AVUSGGAGED-QLAIRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTOGTGGVOQQ I ARWLARRGAEHLULS SROPDA PGAAELAELRGHGCEVVNAACDVA AVVSGGAGED-QLAIRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTOGTGGVOQQ I ARWLARRGAPHLLLVS SRGPDADGAGELVAELEALGAACTEVYNAACDVA 1267 1277 1287 1296 1316 1326 1334 1344 1354 1267 1277 1287 1296 1316 1326 1334 1344 1354 1364 DVVXGGAVAELTGALGS VTAVIHGA-GRNEPPTALGDLMAAVRSTLAP VDGLRHVLDVVGGONLRLVTFG I IGRAGLGEAH ATANEWLAGTEDVARRNPCD DPVRVGAAVAVTAVEREPTATURAGA-GRNEPPTALGDLMAAVRSTLAP VDGLRHVLDVVGENLRLVTFG I IGRAGLRGEAH ATANEWLAGTEDVAATRAPEC DPVRVGAAVAVTAVAEGPVTAVIHGA-GRNEPPTGLAALTPPUTDLTEELINTLAP NTGLRAVLDAVDARELRLVTFG I IGRAGLRGEAH ATANEWLAGTEDVAATRAPEC DPVRVGAAVAVTAVAEGPVAULTAGA-GRNEPPTGLAALTPPUTDLTEELINTLAP NTGLRAVLDAVDARELRLVTFG I IGRAGLRGEAH ATANEWLAGTEDVENAPHT
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1	ABETA-AARGFTERKTKAD OTRLTPPALAPUY DEPAAGGPAL SADUVULVT GOGGTAACARALAAGSGGAR LAVLGRTREDD ESLRENLARLOPDVKITSADUT AAVERLAAGSPEDQI AVNARS OSWGRRLTRLPPADGGR-TAAPARPROTULVT GOTGALGGHTARKLAAGGAEHLALTS RROPDAPGAAGLERELLLIGAKVTYFAACDTA GLADAGGED -QLAIRGS OVLARRLAHAPAVPGSGRRPPWIGSVLVT GOTGALGGILARKLAAGGAEHLALTS RROPDAPGAAGLERELLJLGAKVTYFAACDTA GLADAGGED -QLAIRGS OVLARRLAHAPAVPGSGRRPPWIGSVLVT GOTGALGGILARKLAEGGAEHLVLTS RROADAPGAAELARELEQLAGATEVYFAACDTI VLOPOVITLDEQQLAVNDY QWAARLARLPRPADDAPAADWNPDGTVLIT OGTGAGGLAARLARKLYTSRG-ARHLLLS RROPDAPGAELAELEGAGGEVYVSACDVG AVLAGRDED -QVAVHAS GI YGRRWYRAAAPATDDEWKPTGTVLVT OGTGGVGQG IARKLARRGAERLVLTS RROPEPGAAELAEELRGHGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWYRAAAPATDDEWKPTGTVLVT OGTGGVGQG IARKLARRGAERLVLTS RROPEPGAAELAEELRGHGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWYRAAAPATDDEWKPTGTVLVT OGTGGVGQG IARKLARRGAERLVLTS RROPEPGAAELAEELRGHGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWYRAAAPATDDEWKPTGTVLVT OGTGGVGQG IARKLARRGAERLVLTS RROPEPGAAELAELGHGEVVHAACDVA AVVSGGAGED -QLALRAD GVYGRRWYRAAAPATDDEWKPTGTVLVT OGTGGVGQG IARKLARRGAERLVTFG IIGRAGLGEMA TANEWLAGTEED/ARCDVACCOMPAC-LLVTT OFT IIGRAGLGEMA ATANEWLAGTEED/ACRNPDC DAGVAAAVATTADAWGP TAVLHGA-GRNEPFALGGLDMAAVRSTLAP VDGLRHVLDVVGE(MLRLVTFG IIGRAGLGEMA ATANEWLAGTEEVAATPPC DAGVAAAVATTADAWGP TAVLHGA-GRNEPALALDALTERALAGTRT VTGLRAVLDAVDPACLLLVTFG IIGRAGLGEMA ATANEWLAGTEEVAATPPC DAGVAAAAVATTADAWGP TAVLGA-GRNEPALALGUTEE LARTLAP AY OFTRAVDAVDAEALLULVTG VIGRAGLGEMA ATANEWLASTESLAATVYQV PAPEAVRAVATATADAWGP TAALALDALAGATANAY TAN TOTGRAV
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR MdpEKR SgcEKR CalE8KR SgcEKR CalE8KR DynE8KR	ARGOTERRITAD OTRLTP ALAPPY LPAAGGAL SADDVULVT GOGTERACARNLAAGS GGAR LAVLGRTRPEDDESLENLARLOPDVRITSADVT AAVERLAGEDED (1 VNRAS OSWGRALTALIPLEDGGAT TAAPATYRRGTVLVT GOTGALGGITARKLAAGSAEHLALTS RRGDAD FGAAGLERELLLIGAKVT FAACDTA ALADAGGED - QLAIRGS WLARRLAHAD AVEGSGRP PYNGSVLVT GOTGALGGITARKLAAGGAEHLALTS RRGDAD FGAAGLERELLLIGAKVT FAACDTA ALALGALGDDDQLAVNPR GVLARR LANLAPAVEGSGRP PYNGSVLVT GOTGALGGIARLANEGCAEHLULS RRGDAD FGAAELARELEQLAGATEVYNFAACDTA VLDPULTLDEQULAVND Y QWAAR LARLPR PADDAPAADWNPDGTVL IT GOTGALGALARHLVT SKG-ARHLLLS RRGDAD FGAAELARELEGAAGEVYNFAACDTA AVVSGGAGED -QUAVARS GI YGRRVSRAAAAGAASADWNPGTVL IT GOTGALGALARHLVT SKG-ARHLLLS RRGDAD FGAAELARELGAGGEVYNFAACDTA AVVSGGAGED -QUAVARS GI YGRRVSRAAAAAFADDEWKPTGTVLVT GOTGGVGGQ I ARNLARGAEHLULVS RSGPDADGAGEL VAELEALGAACTEVYNACCDV AVVSGGAGED -QUAVARAS GI YGRRVSRAAAAFADDEWKPTGTVLVT GOTGGVGGQ I ARNLARGAPHLLVS RSGPDADGAGEL VAELEALGAACTEVYNACCDV AVVSGGAGED -QUAVARAS GI YGRRVSRAAAAFATDDEWKPTGTVLVT GOTGGVGGQ I ARNLARGAPHLLVS RSGPDADGAGEL VAELEALGARTEVVAACDV AVVSGGAGED -QUAVARAS GI YGRRVSRAAAFATDDEWKPTGTVLVT GOTGGVGGQ I ARNLARGAPHLLVS RSGPDADGAGEL VAELEALGARTEVVAACDV DYRVARGAASELTGALGS YTAVLHGA-GRNEPTALGGLDMAAVSTIAP PYDGLRAVLAVDYGENLILLVT FG I I GRAGLGRAMH AANENLAELECTEVAATPEC DYRVARGAVAELTGALGS YTAVLHGA-GNNEPAALADLAALRAALRGY AN VDGLRAVLAVDPARLILLVT FG I I GRAGLGRAMH AANENLAELEELTEVAATPEC DYRVARGAVAELTGALGS YTAVLHGA-GNNVPAGLAALLGYDRA YDGLRAVLAVDPARLILLVT FG I I GRAGLGRAMH AANENLAELTEELAGTLPATYNDY DYRVARAAATTADAWGD A-GNNVPAGLAALGDMPH YNGLRAL
UcmEKR TylKR6 AmphKR1 AmphKR1 AmphKR18 TylKR1 EryKR1 SGCEKR CalE8KR MdpEKR MdpEKR AmphKR1 AmphKR1 SGCEKR CalE8KR SGCEKR CalE8KR MdpEKR KedEKR	AEEA-AARGFTERKTKAD OTRLTPLATAPVPLPAAPGPAL SADUVLVT GOGTGALGGITABCARNILAAGS GGAR IAVLGRTRPEDD ESLRENLARLOPDVRITSADUT AAVERLAGPEDQI AVNARS OSWGRRLTRLEPEDGGGT-TAAPAR PRRFTUTVUT GOTGALGGITABCLARNILAAGSAEHLALTS RRGDAD FGAAGLERELLLIGAKUYT FAACDTA GLADAGGED-QLAIRGS GVLARRLANAPAVPCSGRPPVHGSVLVT GOTGALGGIRARNILAGGAEHLALTS RRGDAD FGAAGLERELLJGAKUYT FAACDTA AALGGALGDDQLAVNPL GVLARR LYNASGDTRAKARSWKRFGTTLIT GOTGGIGGRVARRLAGGGAEHLULS RRGDAP FGAAELAELLGAAGTEVYFRACDIT VLPDVLTLDEQQLAVND Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOTGGIGGALARRHUYTSGAEHLULS RRGDAP FGAAELAELEGAAGTEVYHACCDVA AVVSGGAGED-QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTGOTGGUGGQI ARWLARGAEHLULS RRGPDA FGAAELAELEGHGCEVVHACCDVA AVVSGGAGED-QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTGOTGGUGGQI ARWLARGAEHLULYS RSGPDADGAGELVAELEALGAGTEVVHACCDVA AVVSGGAGED-QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTGOTGGUGGQI ARWLARGAEHLULYS RSGPDADGAGELVAELEALGAGTEVVHACCDVA AVVSGGAGED-QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTGOTGGUGGQI ARWLARGAEHLLUSS RSGPDADGAGELVAELEALGAGTEVVHACCDVA AVVSGGAGED-QLALRAD GVYGRRWVRAAAPATDDEWKPTGTVLVTGOTGGUGGQI ARWLARGAEHLUSS RSGPDADGAGELVAELEALGAGTEVVHACCDVA AVVSGGAGAGED-QLALRAD GVYGRRWRAAAPATDDEWKPTGTVLVGGTGGUGGQI ARWLARGAEHLUSS RSGPDADGAGELVAELEALGAGTEVVHACCDVA AVVSGGAGAGED-QLALRAD GVYGRRWRAAAPATDDEWKPTGTVLVGGTGGUGGQI ARWLARGAEHLUSS RSGPDAGAGELVAELEALGAGTEVVHAACDVA
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR DynE8KR MdpEKR MdpEKR MdpEKR AmphKR1 EryKR1 EryKR1 EryKR1 SgoEKR CalE8KR DynE8KR CalE8KR MdpEKR MdpEKR	ABETA-AARGPTEARTRAD OTRLTPALAPVPLPAAPGPAL SADUVLIVT 000000000000000000000000000000000000
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR DynE8KR MdpEKR McsEKR McsEKR SgcEKR CalE8KR TylKR1 EryKR1 SgcEKR CalE8KR DynE8KR McsEKR McsEKR McsEKR	AEEA-AARGFTERKTKAD OTRLTPLATAPVPLPAAGGPAL SADUVLIVT GOGTALGARVILAAG. AAVERLAAGPEDQI AVNARS OSWGRALTKILPEDGOGRT TAAPATAPRRGTVLIVT GOTGALGGITARVILAAG. AALDAG GEDD_QIANRAS OSWGRALTKILPEDGOGRT TAAPATAPRRGTVLIVT GOTGALGGIARVILABGCAEHLALTS RRGDAPGAAGLEAELLILGAKVTY FAACDTA AALAGALDDDQLAVNPA GVLARR LANAPAVPGSGRAPPVHOSVLVT GOTGALGAIRNILAGG. VLDPULTLDEQQLAVND Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOTGAGALARRILVTSGAEHLULS RRGDAPGAAELAAELEQLAGATEVYFRACDT VLDPULTLDEQQLAVND Y QWAARLARLPRPADDAPAADWNPDGTVL IT GOTGAGALARRILVTSGAEHLULS RRGDAPGAAELAELEGAGGEVYVAACDVA AVVSGGAGED_QUAVRAS GI YGRRVSRAAAAGADAEWNPGTVLIT GOTGAGALARRILVTSGAEHLULTS RRGPEPGAAELAEELRGHGEVYHAACDVA AVVSGGAGED_QUAVRAS GI YGRRVSRAAAAGADDEWNRTGTVLVT GOTGAGAGAIRARGAEHLULTS RRGPEPGAAELAEELRGHGEVYHAACDVA AVVSGGAGED_QUAVRAS GI YGRRVSRAAAAGADDEWNRTGTVLVT GOTGAGAGAIRARGAEHLULTS RGFEPGAAELAEELRGHGEVYHAACDVA AVVSGGAGED_QUAVRAS GI YGRRVSRAAAAGAADDEWNRTGTVLVT GOTGAGAGAACHARTS CAVVSGGAGED_QUAVRAS GI YGRRVSRAAAAAAAAAAAATYDDEWNRTGTVLVT GOTGAGAGAACHARTS RGFEAAELAGATESCHARTVAACDV DYRVAGAAAXATYTADMGP TAAVLAGA-GRNEPTALGGLDMAAVRSTIAP VDGLRAVLAVVGEQMLELIVT FG I I GRAGAGRAGAH ATANEWLAGITEEDVARRNPDC DAGAVAAAVATTADMGP TAAVLAGA-GONEPTALAGILAGAAAA VOGRAAALAGVDPAALILIVT FG I I GRAGAGRAGAH ATANEWLAELEELAGATTVATVATVACTOR DYRVAGAVAELGGVTAVLAGA-GONEPAPAVDILTAEELIRTIAP AYDERVVAALDPDALILIVT FG I I GRAGAGRAGAH ATANEWLAELTELATATTPACTAT
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR DynE8KR MdpEKR McsEKR UcenEKR TylKR6 AmphKR2 SgoEKR CalE8KR McsEKR McsEKR CalE8KR McsEKR CalE8KR McsEKR CalE8KR	A A A A ARGYET ARARGY TEAR TRAD OT NUTTINE PROPERTISATION OF CONTINUE AND TRADE AND A A A A A A A A A A A A A A A A A A
UcenEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR DynE8KR KcdEKR MdpEKR MdpEKR MdpEKR MdpEKR AmphKR1 AmphKR1 SqcEKR CalE8KR CalE8KR CalE8KR CalE8KR MdpEKR SqcEKR CalE8KR MdpEKR MdpEKR AmphKR1 AmphKR1 AmphKR1 AmphKR1	AREA-AARGFTEARTRADGTRLTFALAFVPLPAGPAL SADDVVLVT GGOGT TACACARNIAAGSGGAR LAVLGRTRFEDDE SLRENLARLGPDVR TS SADV AAVERLAOPEDQ LAVRAS GSWGRKLTRLFRDGOGG-TAAPAY RPRGTVLVT GGT GALGGHLARNIAAGSGGAR LAVLGRTRFEDDE SLRENLARLGPDVR TS ADVT GLADAGGED - QLAIRGS GVLARRLANAAPAYGSGKRF PYHGSVLVT GGT GALGGHLARRIARGGARHLUTS RRGDAPGAGLEARELLGAKVYT IAACDAA AALAGALGDDQ QLAVRAF QVLARR IVRASGDTRFKARSWKPRGTTLVT GGT GGTAAPGLARRIARGGARHLULS RRGDAPGAAELARELGAGTEVYTAACD I VLPDVLTLDEQQ LAVRAS GI YGRKWSRAAAAGASWGPSGTVLI TGGT GGLGAALARHLAYES RG-ARHLLAS RRGDAPGARAELARELAGGEVYNAACDVA AVVSGGAGED - QLAIRAB GYYGRKWSRAAAAAFADDEWKPTGTVLVT GGT GGTQGGQG IARHLARGGARHLUTS RRGPAPGARAELAEELAGGGEVYNAACDVA AVVSGGAGED - QLAIRAB GYYGRKWSRAAAAFADDEWKPTGTVLVT GGT GGTQGGQ QI ARHLARGAPHLULYS RSGPAPGABAELAEELAGGAETTVAACDVA AVVSGGAGED - QLAIRAB GYYGRKWSRAAAAFATDDEWKPTGTVLVT GGT GGTQGGQ QI ARHLARGAPHLULYS RSGPAPGABGELVAELEALGARTTVAACDVA AVVSGGAGED - QLAIRAB GYYGRKWSRAAAAFATDDEWKPTGTVLVT GGT GGTQGGQ QI ARHLARGAPHLUVS RSGPADGAGEL VAELEALGARTTVAACDVA AVVSGGAGED - QLAIRAB GYYGRKWSRAAAAFATDDEWKPTGTVLVT GGT GGTAGAL GRWLANDAPHLUVS RSGPADGAGEL VAELEALGARTTVAACDVA AVVSGGAGED - QLAIRAB GYNGRKWSRAAAAFATDDEWKPTGTVLVT GGT GGTAGAL GRWLAND -APHL DFYKVAGAVATVTABHOPVTAVLHAG-GGNREPALALGUTBALAKVSAAY DGLRAVLAVUDVGEQNLRLUVT FGI I GRAGLGGEAH ATANEWLAGLTEDVARNPDG DGAGVAAAVTATVTABHOPVTAVLHAG-GGNREPATLGGLAMAAVST TAPK VDGLRAVLAVDDRALKLUVT FGI I GRAGLGGEAH ATANEWLAGLTEDVARNPDG DFYGVAAVAAVTAEHGPVTAVLYGA-GRNYFALADUTALAGUTBER TAFAY DGLRAVLAVDARALKLUVT FGI I GRAGLGGEAH ATANEWLAGLTEEVARNPT DFYKVAAAVAELTESLGPVTAVLYGA-GRNYFALADUTALAK DFRATAPK VDGLRAVLAAVDARALKLUVT FGI I GRAGLGGEAH ATANEWLAGLTEEVARNPT DFYKVAAAVAELTESLGPVTAVLYGA-GRNYFALADUTALATAPK VGGLAUAADAPARLKLUVT FGI I GRAGLGGEAH ATANEWLAGLTEETAAFTPE DFYKVAAAVAELTESLGPVTAVLYGA-GRNYFALADUTAKY VTGLAAUAAPATAFT AFTANEWLAGUTTES TAGANHATAANEXLAAUADATAFT DFYKVAAAAVAELTESLGPVTAVLYGA-GRNYFALADUTAKY VTGLAAUAAPATAFT DFYKVAAAVAELTESLGPVTAVLYGA-GRNYFALADUTAKY TYT FU GGTAATAAAAAAAATAATA
UcanEKR TylKR6 AmphKR1 AmphKR1 AmphKR1 EryKR1 EryKR1 EryKR1 CalE8KR MdpEKR MdpEKR MdpEKR MdpEKR UcanEKR TylKR6 AmphKR1 AmphKR1 SgcEKR CalE8KR DynE8KR UcanEKR TylKR1 EryKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1 AmphKR1	AETA-AARGTEARTRAD GTRLTPALAPUP LPAROPALSADDVULVT GOGGT TACAGHLARGANGARHLALTSREPDA FGAAGLEAELLLGAUXT FAACDTA AAVERLADGTEQ IAVRAS GSWGRALTRLPRDGOGGTAAPAT KPRGTULTU GGTGI GGUARGHLARAGGARHLALTSRCADAPGAAELFAELEQI GUVART IAACDA GLLADAGGED - QLAIRAS GSWGRALTRLPRDGOGARADAVRENGTULTU GGTGI GGUARGHLARGGARHLULTSRCADAPGAAELFAELEQI GUVART IAACDA AALAGALDDDQL AVRDY QUVARALARLPR PADD APADWNPDGTULTU GGTGGI GALARHLYT SR G-ARHLULASRCADAPGAAELFAELEQI GUVART IAACDA AVLGGDDC - QUVARAS GI IGREVSRAAAAGA

Figure S2 continues to the next page

	SgcE-KR	CalE8-KR	DynE8-KR	KedE-KR	MdpE-KR	NcsE-KR	UcmE-KR
SgcE-KR	-	55	37	53	63	67	41
CalE8-KR	55	-	40	51	57	54	42
DynE8-KR	37	40	-	39	40	40	37
KedE-KR	53	51	39	-	53	52	46
MdpE-KR	63	57	40	53	-	61	41
NcsE-KR	67	54	40	52	61	-	41
UcmE-KR	41	42	38	46	41	41	-

Figure S2. Protein sequence comparison of KR domains between selected noniterative type I PKSs and PKSEs. (A) Protein sequence alignment of KR domains with the conserved KSY catalytic residues, NADPH binding site, and diagnostic LDD motif in B-type KRs and W residue in A-type KRs highlighted in green, yellow, and red, respectively. See Figure S1 legend for accession numbers of the seven PKSEs. (B) Sequence identities (%) of KR domains among the seven PKSEs. The accession numbers for the three A-type KRs are TylKR6 (AAB66508), AmphKR2 (AAK73513), and AmphKR1 (AAK73513). The accession numbers for the three B-type KRs are AmphKR18 (AAK73593), TylKR1 (AAB66504), and EryKR1 (AAV51820). The crystal structures for AmphKR2 (3JME), EryKR1 (2FR1), and TylKR1 (2Z5L) have been previously determined.



Figure S3. SDS-PAGE of the purified KRs. Lane 1, SgcE-KR (47,878 Da); lane 2, KedE-KR (48,284 Da); lane 3, MdpE-KR (47,163 Da); lane 4, NcsE-KR (48,553 Da); lane 5, protein ladder; lane 6, CalE8-KR (46,590 Da); lane 7, UcmE-KR (48,054 Da); and lane 8, DynE8-KR (46,239 Da). Given in parentheses are calculated molecular weights for each of the recombinant proteins.



Figure S4. ¹H NMR (400 MHz) spectrum of compound 1 in CDCl₃



Figure S5. ¹³C NMR (100 MHz) spectrum of compound 1 in CDCl₃



Figure S6. ¹H NMR (400 MHz) spectrum of compound 2 in CDCl₃



Figure S7. ¹³C NMR (100 MHz) spectrum of compound 2 in CDCl₃



Figure S8. ¹H NMR (400 MHz) spectrum of compound 3 in CDCl₃



Figure S9. ¹³C NMR (100 MHz) spectrum of compound 3 in CDCl3



Figure S10. ¹H NMR (400 MHz) spectrum of compound 4 or 7 in CDCl₃



Figure S11. ¹³C NMR (100 MHz) spectrum of compound 4 or 7 in CDCl₃



Figure S12. ¹H NMR (400 MHz) spectrum of compound 5 or 8 in CDCl₃



Figure S13. ¹³C NMR (100 MHz) spectrum of compound 5 or 8 in CDCl₃



Figure S14. ¹H NMR (400 MHz) spectrum of compound 6 or 9 in CDCl₃



Figure S15. ¹³C NMR (100 MHz) spectrum of compound 6 or 9 in CDCl₃



Figure S16. ¹H NMR (400 MHz) spectrum of compound S-MTPA-5 in CDCl₃



Figure S17. ¹H NMR (400 MHz) spectrum of compound *R*-MTPA-5 in CDCl₃



Figure S18. $\Delta \delta_{S-R}$ values (ppm) of MTPA esters for compound **5**.



Figure S19. ¹H NMR (400 MHz) spectrum of compound S-MTPA-6 in CDCl₃



Figure S20. ¹H NMR (400 MHz) spectrum of compound *R*-MTPA-6 in CDCl₃



Figure S21. $\Delta \delta_{S-R}$ values (ppm) of MTPA esters for compound **6**.



Figure S22. Enzymatic assays of PKSE-KRs by continuous photometric measurement at 340 nm to follow the time-dependent depletion of NADPH as exemplified by SgcE-KR with the three β -ketoacyl-SNAC substrates, **1**, **2**, and **3**.



Figure S23. Pseudo-first order kinetic characterization of PKSE-KRs with the three β -ketoacyl-SNAC substrates, **1**, **2**, and **3**: A, SgcE-KR; B, KedE-KR; C, MdpE-KR; D, NcsE-KR; E, CalE8-KR; F, DynE8-KR; G, UcmE-KR.