Supporting Information

Structural Basis of Polyketide Synthase O-Methylation

Meredith A. Skiba^{a,b o,∇}, Marissa M. Bivins^{a,o,Δ}, John R. Schultz^d, Steffen M. Bernard^{a,c,#},

William D. Fiers^{d,§}, Qingyun Dan^a, Sarang Kulkarni^e, Peter Wipf^e, William H. Gerwick^{f,g},

David H. Sherman^{a,h,i,j}, Courtney C. Aldrich^d, Janet L. Smith^{a,b*}

^aLife Sciences Institute, University of Michigan, Ann Arbor, MI, 48109, United States

^bDepartment of Biological Chemistry, University of Michigan, Ann Arbor, MI, 48109, United States

^cChemical Biology Doctoral Program, University of Michigan, Ann Arbor, MI, 48109, United States

^dDepartment of Medicinal Chemistry, University of Minnesota, Minneapolis, MN, 55455, United States

^eDepartment of Chemistry, University of Pittsburgh, Pittsburgh, PA, 15206, United States

^fCenter for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, 92093, United States

^gSkaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, 92093, United States

^hDepartment of Medicinal Chemistry, University of Michigan, Ann Arbor, MI, 48109, United States

ⁱDepartment of Chemistry, University of Michigan, Ann Arbor, MI, 48109, United States ^jDepartment of Microbiology and Immunology, University of Michigan, Ann Arbor, MI, 48109, United States

°These authors contributed equally to this work

*To whom correspondence should be addressed <u>JanetSmith@umich.edu</u>

Present Addresses

^v Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, United States

^A Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, United States

[#] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, 92037, United States

[§] Gastroenterology and Hepatology Division, Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medicine, New York, NY, 10021, United States

The Jill Roberts Institute for Research in Inflammatory Bowel Diseases, Weill Cornell Medicine, New York, NY, 10021, United States

Construct Design

StiD and StiE constructs were amplified from codon optimized DNA (IDT) encoding regions of *stiD* (CAD19088.1) and *stiE* (CAD19089.1) from *Stigmatella aurantiaca Sg a15.*¹ CurL constructs were amplified from a cosmid library.² Constructs encoding StiD *O*-MT (residues 956-1266, pMAS162; residues 976-1266, pMAS165), StiD ACP (residues 1794-1929, pMAS201), StiE *O*-MT (residues 942-1257, pMAS198), StiE ACP (residues 1789-1927, pMAS283), and CurL *O*-MT (residues 981-1315, pMAS411) were inserted into pMCSG7³ and a construct encoding CurL *O*-MT (residues 951-1315, pMAS/SMB134) was inserted into pMOCR⁴ by ligation independent cloning. *stiE* mutations were introduced into pMAS198 (E1102A, pMAS432; Y954F, pMAS433; E1102Q, pMAS434; L1106H, pMAS435; Y1209F, pMAS437; Y1223F, pMAS439), *curL* mutations were introduced into pMAS/SMB134 (Y1010F, pMAS416; E1161A, pMAS428; E1161Q, pMAS430; H1165A, pMAS413; H1165N, pMAS412; Y1267F, pMAS408; Y1281F, pMAS429) using the QuickChange protocol (Stratagene). All constructs and mutations were verified by Sanger sequencing at the University of Michigan DNA Sequencing Core.

Protein Expression and Purification

Plasmids containing *O*-MTs and ACPs were transformed into *Escherichia coli* strain BL21(DE3). Transformed cells were grown in 0.5 L of TB media at 37°C supplemented with 100 μ g mL⁻¹ ampicillin to an OD₆₀₀=1-2, cooled to 20°C for 1 hr, and induced with 200 μ M isopropyl β -D-1-thiogalactopyranoside (IPTG) for 18 hrs. Media to produce StiD, StiE, and CurL ACP was supplemented with a trace metal mix to insure production of apo-ACP, lacking the phosphopantetheine post translational modification.⁵ Selenomethionine (SeMet) labeled StiD *O*-MT (residues 976-1266) was produced in 2L of SelenoMet medium (AthenaES) containing 150 μ g/mL seleno-DL-methionine. Cultures were grown to an OD₆₀₀=0.6 at 37°C, cooled to 20°C for 1 hr, and induced with 200 μ M IPTG for 18 hrs.⁶

Cell pellets were resuspended in 35 mL Tris buffer A (50 mM Tris pH 7.4, 300 mM NaCl, 10% glycerol, 20 mM imidazole) with 0.1 mg mL⁻¹ lysozyme, 0.05 mg mL⁻¹ DNase, and 2 mM MgCl₂, incubated on ice for 30 min, lysed by sonication, and cleared by centrifugation (38,760 x g, 30 min, 4°C). The supernatant was filtered and loaded onto a 5 mL His trap column (GE Healthcare). Proteins were eluted with a 5-100% linear gradient of Tris buffer B (50 mM Tris 7.4, 300 mM NaCl, 10% glycerol, 400 mM imidazole) over 10 column volumes. The His-Tag and/or Mocr

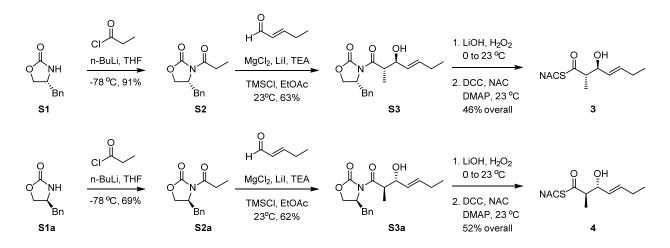
fusion was removed from StiD *O*-MT (residues 956-1266), StiE *O*-MT (residues 942-1257), CurL *O*-MT (residues 981-1315 and 951-1315), StiE ACP, and CurL ACP via incubation with tobacco etch virus (TEV) protease. The cleavage reaction mixture was dialyzed overnight into Tris buffer A to remove imidazole. Protein lacking the His-tag was isolated by passing over a second His trap column. Proteins were further purified by gel filtration (*O*-MTs, HiLoad 16/60 Superdex S200; ACPs, HiLoad 16/60 Superdex S75) in Tris buffer C (50 mM Tris pH 7.4, 150 mM NaCl, 10% glycerol). Apparent molecular weights were determined by analytical size exclusion chromatography (10/300 Superdex S200 Increase equilibrated with Tris Buffer C).

In order to produce holo StiD ACP, 113 μ M StiD ACP was incubated with 0.5 mM coenzyme A (CoA) and 20 μ M *Streptomyces verticillus* phosphopantetheinyl transferase (SVP)⁷ in Tris buffer C with 20 mM MgCl₂ and 2 mM DTT at 20 °C overnight to produce holo-ACP. The His-tag was simultaneously removed by the addition of 12 μ M TEV protease. Holo-ACP lacking the His-tag was purified from the reaction mixture by passing over a 1 mL His trap column. Purified holo-ACP was dialyzed into Tris buffer C. StiD ACP concentration was determined using the DC protein assay (BioRad).

General Chemistry Procedures

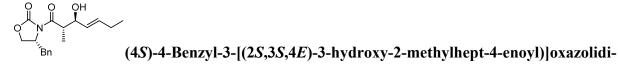
All chemical reagents were used as provided unless indicated otherwise. Tetrahydrofuran (THF) and dichloromethane (CH₂Cl₂) were purified via passage through neutral alumina columns. Ethyl acetate was purified by drying over 4 Å molecular sieves. Compounds were purified by flash chromatography using silica gel (300–400 mesh) in the indicated solvent system. TLC was performed using 250 μ m, F254 silica gel plates and visualized by UV or through staining with para-anisaldehyde or potassium permanganate. Optical rotations were acquired on a polarimeter at the indicated temperature using the sodium D line ($\lambda = 589$ nm) unless otherwise specified and reported as follows: [α] λ Temp = rotation (c g/100 mL, solvent). ¹H and ¹³C NMR spectra were recorded on a 500 MHz NMR spectrometer. Chemical shifts are reported in ppm based on an internal standard of residual CHCl₃ (7.26 ppm in ¹H NMR and 77.16 in ¹³C NMR). Proton chemical data are reported in the following format: chemical shift (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, sext = sextet, sept = septet, m = multiplet, br = broad peak, app = apparent), coupling constant(s), and integration. High-resolution mass spectra

(HRMS) were obtained on a time-of-flight (TOF) mass spectrometer using either PEG or PPG standards to calibrate the instrument.



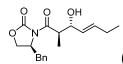
Scheme 1. Synthesis of N-acetylcysteamine (NAC)-linked triketide substrates

The synthesis of the two *anti*-enantiomers of the *N*-acetylcysteamine (NAC) linked triketide substrates **3** and **4** began with acylation of L-phenylalanine and D-phenylalanine-derived oxazolidinone chiral auxiliaries with propionyl chloride.⁸ The non-Evans *anti*-aldol products (**S2** and **S2a**) were obtained in much better yields than previously reported ⁹ by utilization of a modified magnesium halide-catalyzed anti-aldol reaction. ¹⁰ The final NAC liked triketide substrates were prepared by removal of the oxaozolidinone auxiliary via lithium hydroperoxide followed by DCC-mediated coupling of NAC in the presence of catalytic DMAP. The other two *syn*-diastereomers, compounds **S1** and **S1a**, were synthesized as previously reported.¹⁰



none (S3). To a flame-dried reaction vessel under an argon atmosphere equipped with a stir bar, oxazolidinone **S2** (0.100 g, 0.429 mmol), MgCl₂ (0.038 g, 0.40 mmol), LiI (0.106 g, 0.79 mmol), EtOAc (0.85 mL), Et₃N (0.27 mL, 1.89 mmol) and TMSCl (0.20 mL, 1.50 mmol) were added sequentially. After 10 min of stirring trans-2-penten-1-al (0.585 mL, 1.17 mmol, 2M in EtOAc) was slowly added over 3 h at 23 °C. The reaction mixture was monitored by TLC and was complete at 5 h. The reaction mixture was passed through a plug of silica gel using EtOAc as the eluent.

After removing the solvent in vacuo, MeOH (4 mL) and *p*-TsOH (25 mg) were added. After 45 min desilylation was complete as judged by TLC. The crude reaction was concentrated onto silica gel and purified by flash column chromatography (9:1 to 7:3 Hexanes-EtOAc stepwise gradient) to give the title compound (0.084 g, 63%) as a viscous, colorless oil: $R_f = 0.25$ (7:3 hexanes-EtOAc); $[\alpha]_D^{24} = -27.6$ (*c* 1.00, CHCl₃); ¹H NMR (CDCl₃, 500 MHz) δ 7.29 (m, 5H), 5.81 (dt, J = 15.3, 6.3 Hz, 1H), 5.51 (dd, J = 15.4, 7.3 Hz, 1H), 4.70 (m, 1H), 4.18 (m, 3H), 3.95 (m, 1H), 3.29 (dd, J = 13.5, 2.9 Hz, 1H), 2.78 (dd, J = 13.5, 9.4 Hz, 1H), 2.57 (m, 1H), 2.09 (q, J = 7.4 Hz, 2H), 1.17 (d, J = 6.9 Hz, 3H), 1.01 (t, J = 7.5 Hz, 3H); ¹³C NMR (CDCl₃, 500 MHz) δ 206.7, 176.5, 153.5, 135.9, 135.3, 129.5, 129.2, 128.9, 127.3, 75.8, 66.0, 55.5, 43.4, 37.8, 30.9, 25.2, 14.5, 13.4.



(4R)-4-Benzyl-3-[(2R,3R,4E)-3-hydroxy-2-methylhept-4-enoyl)]oxazolidi-

none (S3a). To a flame-dried reaction vessel under an argon atmosphere equipped with a stir bar, oxazolidinone **S2a** (0.200 g, 0.857mmol), MgCl₂ (0.082 g, 0.86 mmol), LiI (0.225 g, 0.1.68 mmol), EtOAc (1.60 mL), Et₃N (0.57 mL, 4.04 mmol) and TMSCl (0.42 mL, 3.22 mmol) were added sequentially. After 10 min of stirring *trans*-2-pentene-1-al (1.25 mL, 2.49 mmol, 2M in EtOAc) was slowly added over 3 h at 23 °C. The reaction mixture was monitored by TLC and was finished at 5 h. The reaction mixture was passed through a plug of silica gel using EtOAc as the eluent. After removing the solvent in vacuo, MeOH (4 mL) and p-TsOH (25 mg) were added. After 45 min desilylation was complete as judged by TLC. The crude reaction was concentrated onto silica gel and purified by flash column chromatography (9:1 to 7:3 Hexanes-EtOAc stepwise gradient) to give 7 (0.168 g, 62%) as a viscous, yellow oil: Rf = 0.25 (7:3 hexanes–EtOAc); $[\alpha]_D^{24} = 27.1$ (*c* 1.00, CHCl₃); ¹H NMR (CDCl₃, 500 MHz) δ 7.29 (m, 5H), 5.81 (dt, *J* = 15.9, 6.1 Hz, 1H), 5.52 (dd, *J* = 15.4, 7.2 Hz, 1H), 4.70 (m, 1H), 4.17 (m, 3H), 3.95 (m, 1H), 3.30 (dd, *J* = 13.6, 3.3 Hz, 1H), 2.78 (dd, *J* = 13.6, 9.4 Hz, 1H), 2.57 (m, 1H), 2.09 (q, *J* = 7.3 Hz, 2H), 1.18 (d, *J* = 7.3 Hz, 3H), 1.01 (t, *J* = 7.4 Hz, 3H); ¹³C NMR (CDCl₃, 500 MHz) δ 206.7, 176.5, 153.5, 135.9, 135.3, 129.5, 129.2, 128.9, 127.3, 75.8, 66.0, 55.5, 43.4, 37.8, 30.9, 25.2, 14.5, 13.4.

NACS S-(2-Acetamidoethyl) (2S,3S,4E)-3-hydroxy-2-methylhept-4- enethioate (3).

To a stirring solution of S3 (56 mg, 0.169 mmol) in THF (0.6 mL) at 0 °C was added aqueous H₂O₂ (30% v/v, 0.245 mL) dropwise. Aqueous LiOH (1.19 mL, 0.4 M) was then added portionwise to the stirring solution and the reaction was allowed to warm to room temperature and stirred for 5 h. The reaction was quenched by the addition of Na₂SO₃ (5 mL) then concentrated under reduced pressure. The crude material was portioned between CH₂Cl₂ (10 mL) and 10% NaOH (5 mL) and the organic layer was discarded. The aqueous layer was then acidified to a pH of 1 with 1 M HCl and extracted with CH_2Cl_2 (3 × 10 mL). The organics were dried over MgSO₄, filtered, and then concentrated to afford the crude acid. To a solution of the crude acid (28 mg, 0.168 mmol) in CH₂Cl₂(1.69 mL) added N-acetylcysteamine (80 mg, 0.67 mmol), N,N-dicyclohexylcarbodiimide (45 mg, 0.22 mmol), and 4-(dimethylamino)pyridine (1 mg, 0.009 mmol) sequentially. The reaction was stirred vigorously for 26 h. The reaction was filtered, concentrated and directly purified by flash chromatography (85:15:5 toluene-CH₂Cl₂–MeOH) to afford the title compound (20 mg, 46% over 2 steps) as a light yellow, viscous oil: $R_f = 0.30$ (95:5 CH₂Cl₂–MeOH); $[\alpha]_D^{24} =$ 37.1 (c 0.35, CHCl₃); ¹H NMR (CDCl₃, 400 MHz) δ 5.88 (br s, 1H), 5.79 (dt, J = 15.6, 6.0 Hz, 1H), 5.40 (ddt, J = 15.6, 6.4, 1.2 Hz, 1H), 4.21 (t, J = 7.6 Hz, 1H), 3.55–3.37 (m, 2H), 3.13–2.98 (m, 2H), 2.76 (quint, J = 7.2 Hz, 1H), 2.07 (app. quint, J = 7.6 Hz, 2H), 1.96 (s, 3H), 1.13 (d, J =6.8 Hz, 3H), 1.00 (t, J = 7.6 Hz, 3H); ¹³C NMR (CDCl₃, 100 MHz) δ 203.7, 170.6, 136.8, 128.8, 75.6, 54.5, 39.7, 28.7, 25.4, 23.3, 15.2, 13.5; HRMS (ESI+) m/z calcd for C12H21NO3SNa⁺ [M + Na]+ 282.1134, found 282.1158 (error 8.5 ppm).

NACS S-(2-Acetamidoethyl) (2*R*,3*R*,4*E*)-3-hydroxy-2-methylhept-4-enethioate (4). The thioester 4 was synthesized in an analogous manner to that of the enantiomer S3 furnishing the desired product (23 mg, 52%) as a light yellow, viscous oil: $R_f = 0.30$ (95:5 CH₂Cl₂–MeOH); $[\alpha]_D^{24} = -34.5$ (*c* 0.29, CHCl₃); ¹H NMR (CDCl₃, 400 MHz) δ 6.00 (br s, 1H), 5.76 (dt, *J* = 15.6, 6.0 Hz, 1H), 5.40 (ddt, *J* = 15.6, 6.4, 1.2 Hz, 1H), 4.19 (t, *J* = 7.6 Hz, 1H), 3.55–3.37 (m, 2H), 3.13–2.98 (m, 2H), 2.76 (app. quint, *J* = 7.2 Hz, 1H), 2.42 (br s, 1H), 2.06 (app. quint, *J* = 7.6 Hz, 2H), 1.94 (s, 3H), 1.11 (d, *J* = 6.8 Hz, 3H), 1.00 (t, *J* = 7.6 Hz, 3H); ¹³C NMR (CDCl₃, 100 MHz) δ 203.6, 170.6, 136.6, 128.8, 75.5, 54.5, 39.5, 28.7, 25.4, 23.3, 15.1, 13.5; HRMS (ESI+) *m*/*z* calcd for C₁₂H₂₁NO₃SNa⁺ [M + Na]⁺ 282.1134, found 282.1151 (error 6.0 ppm).

Production of acyl-ACPs

N-acetylcysteamine (NAC)-linked diastereomeric triketide substrates (1, 2) were synthesized as previously reported⁹. Holo StiD-ACP (50 μ M) was incubated with 5 mM 1, 2, 3, or 4 in 300 mM sodium bicarbonate pH 8.1 at 25°C for 4 hrs. ACP was buffer exchanged into reaction buffer (50 mM HEPES 7.4, 150 mM NaCl) and flash frozen.

(3R)-3-Hydroxy-5-methoxy-myristoyl-CoA was synthesized as previously reported¹¹. Apo StiE-ACP (180 µM) was incubated with 36 µM SVP, 0.72 mM (3*R*)-3-hydroxy-5-methoxy-myristoyl-CoA, 20 mM MgCl₂ for 1 hr at 25°C in Tris C. The reaction mixture was passed over a 1 mL His Trap column (GE Healthcare) to isolate the ACP. StiE (3*R*)-3-hydroxy-5-methoxy-myristoyl-ACP was further purified via size exclusion chromatography (HiLoad 16/60 Superdex S75) equilibrated with Tris buffer C.

(*R*)-3-Hydroxydodecanoyl-CoA was synthesized as previously reported¹². Apo CurL ACP (180 μ M) was incubated with 20 μ M SVP, 0.5 mM (*R*)-3-hydroxydodecanoyl-CoA, and 20 mM MgCl₂ for 4 hrs at 25°C in Tris buffer C. The reaction mixture was passed over a 1 mL His Trap column (GE Healthcare) to isolate the ACP. CurL (*R*)-3-hydroxydodecanoyl-ACP was further purified via size exclusion chromatography (HiLoad 16/60 Superdex S75) equilibrated with Tris buffer C.

Enzyme Assays

StiD triketide-ACP (1, 2, 3, 4) (50 μ M) was incubated with 25 μ M StiD *O*-MT (956-1266) and 0.5 mM SAM in 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM MgCl₂. Reaction mixtures (10 μ L) were incubated for 24 hrs at 25°C and quenched with 1% formic acid. 0.5 uL of reaction mixtures were subjected to LC/MS analysis.

StiE (3*R*)-3-hydroxy-5-methoxy-myristoyl-ACP (**6**) (50 μ M) was incubated with 12.5 μ M StiE *O*-MT (942-1257) variants or StiE *O*-MT (961-1257) and 0.5 mM SAM in 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM MgCl₂. Reaction mixtures (10 μ L) were incubated for 15 minutes at 25°C and quenched with 1% formic acid. 1 uL of reaction mixtures were subjected to LC/MS analysis.

CurL (*R*)-3-hydroxydodecanoyl-ACP (7) (50 μ M) was incubated with 12.5 μ M CurL *O*-MT (951-1315) variants and 0.5 mM SAM in 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM MgCl₂. Reaction mixtures (10 μ L) were incubated for 6 hrs at 25°C and quenched with 1% formic acid. 1 uL of reaction mixtures were subjected to LC/MS analysis.

Apo StiD, StiE, or CurL ACP (50 μM) were incubated with 10 uM SVP, 200 μM acetoacetyl-CoA, 0.5 mM SAM and StiD *O*-MT (951-1315), StiE *O*-MT (942-1257), or CurL *O*-MT (951-1315) variants in 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM MgCl₂. Reaction mixtures were incubated for 24 hrs at 25°C and quenched with 1% formic acid. 1 uL of reaction mixtures were subjected to LC/MS analysis.

Reaction mixtures were analyzed using the phosphopantetheine (Ppant) ejection method^{13, 14} on an Agilent Q-TOF 6545. Samples were separated by reverse phase HPLC (Phenomenex Aeris widepore C4 column 3.6 μ M, 50 x 2.10 mm) at a flow rate of 0.5 mL min⁻¹ in H₂O with 0.2% (v/v) formic acid. Protein was eluted with a gradient of 5-100% acetonitrile with 0.2% (v/v) formic acid over 4 min. Data were processed using MassHunter Qualitative Analysis software (Agilent). Substrates and products in the StiD *O*-MT reactions experienced in-source decay yielding a conjugated dehydrated species.

Protein Crystallization and Structure Determination

SeMet labeled StiD *O*-MT (residues 976-1266 with His-tag) was crystallized at 4°C by sitting drop vapor diffusion from 2:2 μ L mixture of protein stock (10 mg mL⁻¹ StiD *O*-MT 976-1266 in Tris buffer C with 1 mM SAM) and well solution (27% PEG 4000, 0.77 M LiCl, 0.1 M Tris pH 7.0). Microseeding was used to obtain single diamond shaped crystals, which grew after 8 days. Crystals were harvested directly from the drop and flash cooled in liquid N₂.

Native StiD *O*-MT (residues 976-1266 with His-tag) was crystallized at 20°C by sitting drop vapor diffusion from 1:1 μ L mixture of protein stock (11 mg mL⁻¹ StiD *O*-MT 976-1266 in Tris buffer C with 1 mM SAM) and well solution (20% PEG 3350, 0.2 M NaF). Diamond shaped crystals grew after one week and were harvested directly from the drop and flash cooled in liquid N₂.

StiD *O*-MT (residues 956-1266, His-tag removed) was crystallized at 20°C by sitting drop vapor diffusion from a 2:1 µL mixture of protein stock (13 mg mL⁻¹ StiD *O*-MT 956-1266 in Tris buffer C with 1 mM SAM) and well solution (1.5 M ammonium citrate tribasic pH 7.2). Microseeding

was used to obtain single rod-shaped crystals, which grew overnight. Crystals were cryoprotected with well solution supplemented with 20% glycerol and flash cooled in liquid N₂.

StiE *O*-MT (961-1257 with His-tag), was crystallized at 4°C by sitting drop vapor diffusion from 2:1 μ L protein stock (11 mg/mL His-tagged StiE *O*-MT in Buffer C with 1 mM SAM or SAH) to well solution (10% PEG 3350, 0.1 M sodium formate). Microseeding was used to obtain single crystals. Square bipyramidal crystals grew in 3 days. Crystals were harvested directly from the drop and flash cooled in liquid nitrogen.

StiE *O*-MT (residues 942-1257, His-tag removed) was crystallized at 4°C by sitting drop vapor diffusion from a 2:1 μ L protein stock (9 mg mL⁻¹ StiE *O*-MT 961-1257 in Tris buffer C with 1 mM SAM) and well solution (25% PEG 8000, 0.1 M HEPES pH 7.4) at 4°C. Square bipyramidal crystals grew after 5 days. Crystals were cryoprotected in well solution supplemented with 20% glycerol and flash cooled in liquid N₂.

Diffraction data for all structures were collected at APS beamline 23ID-D or ID-B and processed using XDS.¹⁵ The SeMet StiD *O*-MT 976-1266 was solved by single-wavelength anomalous diffraction (SAD) phasing using Phenix AutoSol¹⁶ in the Phenix Software suite.¹⁷ A nearly complete model was build using Phenix AutoBuild.¹⁸ Native StiD *O*-MT 976-1266 was isomorphous with the SeMet crystal form. StiD *O*-MT 956-1266 and StiE *O*-MT 942-1257 were solved by molecular replacement using Phaser¹⁹. Iterative rounds of model building and refinement were carried out using Coot²⁰ and Phenix.refine²¹ with automated translation/liberation/screw group selection. Structures were validated using MolProbity.²² Homologs in the structure database were identified using the DALI server.²³ Sequence alignments were prepared using Clustal²⁴ through Jalview²⁵ and figures were prepared with PyMol.²⁶

		9	951 950	956942	981		conserved	961 976
Secondary Structure	Post-AT Lin	lker		11	1	a	aromatics	
StiD/914-1289						D		(-ARGSAVDE
CalF/2606-2945	2606 LPTYPFARKR	CWILPPNQIKPFQEQE	ELMSDQ			EYE	NKAAEYYTHATF	(- ARGSAV
CtaE/903-1291	903 LPAYPWQRSR	YWV <mark>E</mark> LRRAEVHSAAL <mark>P</mark>	L PRSEERAVE PA	SS <mark>P</mark> H <mark>G</mark>		E	GT <mark>AS</mark> RFYDD <mark>S</mark> AE	ERER <mark>VLQ</mark> P DE
CroA/1253-1621	1253 LPSYPWQGDR	YWIDQATTSSPATASR	TVSPPTEG			AE	T PMRFYDAL AF	K-ATGAKLEE
MelE/903-1294		YWVELQRAEPRAGLSA	PPPSEESAEVPA	ASLHG		E		
StiE/899-1290 CurL/924-1338	024 LPAYPEORES							
CroB/439-821	439 LPSYPWORER	FWVOPAGEGVTASOOR		EV		EGS		QQSSVEGEAD ENA
HapA/1409-1769	1409 LPPYPFQRER	YWLEVT					II VADYYDAFSA	AM <mark>GDG</mark> RARD <mark>GGGEG</mark>
NcyC/949-1331	949 LPHYPWQRQR	CWFEAPRTTTVGAL HD	AV <mark>A</mark> RGDALAGG -	· <u>-</u> · · <u>-</u> · · · · · · · ·		AK	ET V <mark>GAL YD</mark> AVAF	R <mark>G</mark> DALA <mark>GG</mark> ASEID <mark>GG</mark> E
PtzB/918-1251	918	DPKKFSIKTHR	VCL PSYSFAPTCI	FPIDGLPVASIRN-		GKRI		NAGSGRDGP
AjuK/1391-1766 BarF/1-366			VDLHTEPPRIEP	SRVE <mark>P</mark> SRVE <mark>P</mark> LRVE	TTRATPSSEA-	ERI		
CtaF/899-1240			VSFYR					
JamN/1-354	1	MSDIQERLAKL	SPEQRQVLEEKI	IESEKYNELK <mark>g</mark> tr <mark>g</mark>		SNE	PLKSYYRSLQ	/NSKSAKT
MelF/900-1239	900 LPTYPFQRER	YWI <mark>D</mark> SLI <mark>P</mark> NTDTL	VS <mark>F</mark> YR				- S <mark>VVHLLE</mark> VNSA	AFE <mark>VP</mark> S
MtaF/898-1239	898 LPTYPFQRER	YWIDSLIPNSDVL	VSFYR				- AVAQMLDVTT\	RDDALAGGASE I DGGE vaGSGRDGP
			~2		04	- 4		
Secondary Structure		α2	03	0	$00 0 \xrightarrow{\beta^1}$	α4	β2 α	5
StiD/914-1289								QAELGLRKIEARGLGGRVQ
CalF/2606-2945								AQAELGNHR I ASRNLAPRAR
CtaE/903-1291 CroA/1253-1621								EQAAVDAERVRARGLQD <mark>R</mark> I R EQATVGREK I AAAGLAGR I A
MelE/903-1294								
StiE/899-1290								QAEVCKQRVRTRGLQNR IR
CurL/924-1338								BQAKFAANQVNDYQLQEQIQ
CroB/439-821								GOAALAARKLRATGLEARV R
HapA/1409-1769 NcyC/949-1331								EQAG I GARRVRSASLQDR I R EQAR I GGERARRAGLAGRVE
PtzB/918-1251								AQAQMANSRV I AAGL SEKVQ
AjuK/1391-1766								QIEIGRARAAAAGVGDRMT
BarF/1-366								<mark>QVKAGEQK</mark> IQ <mark>GLGY</mark> SD <mark>R</mark> IY
CtaF/899-1240								DQIEVGRQR IRALGLDGRVL
JamN/1-354 MeIF/900-1239								EQVNFGKERIESLGYAQRIN DQIEVGRQRIRALGLDGRVL
MtaF/898-1239								QIAVGROR IRGLGLDER IT
1110110001200								
Secondary Structure	β3	β4	α6	β5		α7	β6	.α8 α9
-				β5 0 SHLRE <mark>GG</mark> FMLLAD				α8 α9
Secondary Structure StiD/914-1289 CalF/2606-2945	2768 IFH <mark>KNS</mark> AVAP	- F <mark>PGRYD</mark> VAIGIEV <mark>T</mark> F	H I RDKDAL FGN I	VA <mark>G</mark> LKDN <mark>G</mark> RILLM <mark>D</mark>	FIANS- <mark>GSG</mark> VDV YIANL-R <mark>G</mark> QIVD	HDVEV <mark>SILTRQNWLD</mark> LL	ARHRLRITEL	α8 Δ Δ Δ Δ Δ Δ Δ Δ Δ Δ Δ Δ Δ
StiD/914-1289 CalF/2606-2945 CtaE/903-1291	2768 IFH <mark>KNS</mark> AVAP 1075 VFARDSAKDA	- F <mark>PGRYD</mark> VA I <mark>G I E</mark> V <mark>T</mark> F - F <mark>P</mark> DRYDVAF <mark>GFE</mark> VAT	HI <mark>RDK</mark> DALF <mark>GN</mark> I HIADKDALFSNL	VA <mark>G</mark> LKDN <mark>G</mark> RILLM <mark>D TRS</mark> LNN <mark>GG</mark> FLLLAD	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI	ARHRLRI <mark>TE</mark> L SRHNFRLVEG\	α8 ΔVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI VDISREASLFLEDPSFDQNL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621	2768 IFH <mark>KNS</mark> AVAP 1075 VFARDSAKDA 1416 IHHRDSSKDD	- F <mark>PGRYD</mark> VA I <mark>G IE</mark> V <mark>T</mark> F - FPDRYDVAFGFEVAT - F <mark>P</mark> EARDL IFGFEVAT	H I <mark>R</mark> DKDALFGN I H I ADKDALFSNL H I KDKNALLSN IF	VA <mark>G</mark> LKDN <mark>G</mark> RILLMD TRSLNNGGFLLLAD R <mark>RH</mark> LSD <mark>GG</mark> VLALAD	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FFS-T-GSSIDF	HDVEV <mark>SILTRONWLD</mark> LI EETASYNSSAEEWADVI QDTASYSVTPSEWTELI	ARHRLRITEL SRHNFRLVEG\ SGNQLRVLEC	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294	2768 IFH <mark>KNS</mark> AVAP 1075 VFARDSAKDA 1416 IHHRDSSKDD 1075 VFARDSAKDA	- F <mark>PGRYD</mark> VA I G I EV <mark>T</mark> F - FPDRYDVAFGFEVAT - FPEARDL I FGFEVAT - FPDRYDVAFGFEVAT	H I <mark>R</mark> DKDALFGN I H I ADKDALFSNL H I KDKNALLSN I H I ADKDALFSNL	VA <mark>G</mark> LKDN <mark>G</mark> RILLMD T <mark>RS</mark> LNNGGFLLLAD RRHLSDGGVLALAD A <mark>RS</mark> LNNGGFLLLAD	FIANS - GSGVDV YIANL - RGQIVD FIATG - VSAINI FFS - T - GSSIDF FIAAG - VSAINI	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWTELI EETASYNSSAEEWADVI	- ARHRLRITEL - SRHNFRLVEG - SGNQLRVLEC - SRHNFRLVEG	α8 α9 VDVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI VDISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL VDISREASLFLEDPSFDQNL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDD	- F <mark>PG</mark> RYDVA I G I EVT - F PDRYDVAF GFEVAT - F PEARDL I FGFEVAT - F PDRYDVAF GFEVAT - F PGMYDL VL GFEVAG	HI <mark>RDKD</mark> ALFGNI HIADKDALFSNL HIKDKNALLSNI HIADKDALFSNL LI <mark>PDKD</mark> ALFSNI	VA <mark>G</mark> LKDN <mark>GRILLMD TRS</mark> LNNGGFLLLAD RRHLSDGGVLALAD A <mark>RSLNNGGFLLLAD</mark> DRHLTN <mark>GGLLIMAD</mark>	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FFS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV	HD VEV <mark>SILTRONWLDLI EETASYNSSAEEWADVL QDTASYSVTPSEWTELI EETASYNSSAEEWADVL QETSTFSSTREQWNKLF</mark>	ARHRLRITEL SRHNFRLVEG SGNQLRVLEC SRHNFRLVEG SSNHLRLVDA	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDD 1140 IFNRDSSKDE 614 VFNRDSAKNP	- FPGRYDVAIGIEVT - FPDRYDVAFGFEVAT - FPEARDLIFGFEVAT - FPDRYDVAFGFEVAT - FPGMYDLVLGFEVAT - FPDMYNLAFGFEVAT - FPRHYDLILGIEVAG	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIADKDALFSNL LIPDKDALFSNI HIKDKSLLFSNI LIQDKEALFSNI	VAGLKDNGRILLMD IRSLNNGGFLLLAD RHLSDGGVLALAD ARSLNNGGFLLLAD RHLTNGGLLIMAD SRHLQEEGLLVMAD GRHLNPGGFLLLAD	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FFS - T - GSSIDF FIAAG - VSAINI FVANT - LSPIEV FIANS - DVD IDH FLANT - VSPIEV	HDVEVSILTRONWLDLI EETASYNSSAEEWADVL QDTASYSVTPSEWTELL EETASYNSSAEEWADVL QETSTFSSTREQWNKLF EETSSYFITKOHWVEQU QETSTFSSTREQWIKLI	ARHRLR ITEL SRHNFRLVEG SGNQLRVLEC SSNHLRLVEG SSNHLRLVDA SPNKLQLISA SEHRLRLVEG	α8 α9 VDVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEEN I VD ISREASLFLEDPSFDQNL ID ISAEVVFLHDPAFEEHL VDISREASLFLEDPSFDQNL VDVSNEVANCLHNPDYAAQF ID ISHEVSNYLYDAEFEENL ID ISPEIANCLOPPDYDANL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHHRDSSKDD 1075 VFARDSAKDA 1081 IFORDSAKDA 1081 IFORDSAKDE 614 VFNRDSAKNP 1560 IYRRDSARDA	- FPGRYDVAIGIEVT - FPORYDVAFGFEVAT - FPCRYDVAFGFEVAT - FPORYDVAFGFEVAT - FPGMYDLVLGFEVAG - FPONYNLAFGFEVAT - FPGFYDLILGIEVAG - FPGFYDLIFGFEVAT	HIRDKDALFGNI HIADKDALFSNL HIKDKNALLSNI HIADKDALFSNL LIPDKDALFSNI HIKDKSLFSNI LIQDKEALFSNI HIRDKEGLFSNI	VAGLKDNGRILLMD TRSLNNGGFLLAD RHUSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLLIMAD SRHLQEGLLVMAD SRHLNPGGFLLAD RSHLSPGGTLVLAD	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINT FFS-T-GSSIDF FIAAG-VSAINT FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISMM-ASAIEH	HD VEVSILTRONWLDLI EETASYNSSAEEWAD VI GDTASYSVTPSEWTELI EETASYNSSAEEWAD VI GETSTFSSTREGWNKLI EETSSYFITKOHWVEQI GETSTFSSTREGWIKLI EASSSFFLTAEEWVDLI	ARHRLRITEL SCHNFRLVEG SCHNFRLVEG SCHNFRLVEG SSNHLRLVDA SPNKLQLISA SEHRLRLVEG TRHRLRVVDG	α8 α9 VDVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI VDISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL VDVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISPEIANCLOPDYDANL IDISPEIANCLOPDYDANL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHHRDSAKDA 1075 VFARDSAKDA 1081 IFQRDSAKDA 1140 IFNRDSSKDE 614 VFNRDSAKNP 1560 IYRRDSARDA 1122 VFHRDSAACA	- FPGRYDVAIGIEVT - FPDRYDVAFGFEVAT - FPERYDVAFGFEVAT - FPGNYDVAFGFEVAT - FPGNYNLAFGFEVAT - FPGHYDLILGIEVAG - FPGFYDLILGIEVAG - YVRDYDLMVGVEVVC	HIRDKDALFGNI HIADKDALFSNL HIADKDALFSNL LIPDKDALFSNL LIPDKDALFSNI HIKDKSLLFSNI LIQDKEALFSNI HIRDKEALFSNI	VAGLKDNGRILLMD TRSLNNGGFLLLAD RRHLSDGVLALAD ARSLNNGGFLLLAD RHLTNGGLLIMAD SRHLQEEGLLVMAD SRHLNPGGFLLLAD RSHLSPGGTLVLAD GAHLRPGGALVLSD	FIANS-GSGVDV YIANG-RGQIVD FIATG-VSAINI FIATG-VSAINI FIAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FVSGL-SSDIEH	HD VEVS I L TRONWLDLI EETASYNSSAEEWAD VI QDTASYSVTPSEWTEU EETASYNSSAEEWAD VI QETSTFSSTREQWNKLF EETSSYF I TKOHWVEQI QETSTFSSTREQWIKLL EASSFFLATAEEWVEVI EASSFFLATAERWVEVI	ARHRLRITEL SRHNFRLVEG SGNQLRVLEC SSHNFRLVEG SSNHLRLVDA SPNKLQLISA SEHRLRLVEG SEHRLRVVDG SRNGLQLVEC	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLFVDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISPEIANCLOPPDYANL /DVSREMANFLHDPDFDELL /DVSREIANFLDPPFERNA
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NeyC/949-1331 PtzB/918-1251	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSCDD 1075 VFARDSAKDA 1081 IFQRDSAKDA 1040 IFNRDSAKDA 1140 IFNRDSAKDA 1140 IFNRDSAKDA 1560 IYRDSAKARDA 1122 VFHRDSARDA 1041 IFNRSSTDE	- FPGRYDVAIGIEVT - FPDRYDVAFGFEVAT - FPDRYDVAFGFEVAT - FPDRYDVAFGFEVAT - FPGMYDLVLGFEVAG - FPDRYNLAFGFEVAF - FPRHYDLILGIEVAG - FPGFYDLIFGFEVAF - YVRDYDLMVGVEVVC - FPGFFDLAIGYEVGF	HIRDKDALFGNI HIADKDALFSNL HIADKDALFSNL LIPDKDALFSNI HIADKSLLFSNI HIRDKELFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFRI HIRDKEALFRI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD SRHLDEGGLIMAD SRHLQEGLLMAD SRHLQEGLLVAD SRHLSPGGTLVAD SRHLSPGGTLVAD SAHLRPGGALVLSD SAHLRPGGALVLSD	FIANS-GSGVDV YIANG-VGQ VDV FIATG-VGAINI FIATG-VSAINI FIA-GSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISMM-ASAIEH FVSGL-SSDIEH CAADT-AKSIHL	HD VEVSILTRONWLDLI EETASYNSSAEEWAD VI QDTASYSYTPSEWAD VI EETASYNSSAEEWAD VI QETSTFSSTREQWNKLI EETSSYFITKOHWVEQ QETSTFSSTREQWIKLI EASSSFLTAEEWVDLI EEAGSFLATAERWVEVI EEAGSFLATAERWVEVI	ARHRLRITEL SCHNFRLVEG SCNQLRVLES SSNHRLVEG SSNHRLVDA SPNKLQLISA SEHRLRVVGG TRHRLRVVGG SRNGLQLVEC AQHGFRLIRV	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL IDISPEVANCLHNPDYAAQF IDISPEIANCLQDPDYDANL /DVSREMANFLHDPDFDERL /DVSREIANFLDDPEFERNA IDASSEVGNFLVDPGLDSML
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331	2768 IF HKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSKDD 1075 VFARDSAKDA 1081 IF CRDSAKDA 1081 IF CRDSAKDD 1140 IF NRDSAKDS 1560 IYRRDSARDA 1122 VF HRDSAAGP 1084 IF NRDSSTDL 1580 LF HRDSSTDL	- F PGRYD VAIGIEVT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD LIGIEVAG - F PGFYDLIGIEVAG - F PGFYDLIGFEVAT - YVRD YDLMVGVEVVC - F PGFFDLAIGYUF	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIADKDALFSNL LIPDKDALFSNI LIVDKEALFSNI LIVDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIEDKQLFANL	VAGLKD NGRILLMD TRSLNNGGFLLLAD ARHLSDGVLALAD ARSLNGGFLLLAD SRHLNGGFLLIMAD SRHLNGGFLLIMAD SRHLNGGFLLVAD GAHLRPGGLVVAS GALRPGGKLALID SRHLNGGLAVMAG	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FFS - T - GSSIDF FIAAG - VSAINI FVANT - LSP IEV FIANS - DVD IDH FLANT - VSP IEV FISMM - ASAIEH FVSGL - SSOIEV ILSMM - AKSIHL ILSMM - STP IDH	HD VEVSILTRONWLDLI EETASYNSSAEEWAD VI QDTASYSVTPSEWTELI EETASYNSSAEEWAD VI QETSTFSSTREOWNKLI EETSSYFITKOHWVEQI QETSTFSSTREOWIKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI	A BHRLR ITEL SRHNFRLVEG SGNQLRVLEC SRHNFRLVDA SPNKLQLISA SPNKLQLISA SEHRLRVVDG SRNGLQLVEC ARSGRLINV	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLFVDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISPEIANCLOPPDYANL /DVSREMANFLHDPDFDELL /DVSREIANFLDPPFERNA
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEJ/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1140 IFNDSSKDE 614 VFNRDSAKDA 1560 IYRRDSARDA 1122 VFNRDSARDA 11580 IFNRNSTDL 1580 IFNRNSTDL 1580 IFNRNSTDL 171 VFNDSARDA 177 YNDSAKOP 1051 LFNDSAKDA	- F PGRYDVAIGIEVT - F PORYDVAFGFEVAT - F PORYDVAFGFEVAT - F PORYDVAFGFEVAT - F PORYDLVLGFEVAG - F PONYNLAFGFEVAF - F PGFYDLIEGFEVAF - YVRDYDLWGVEVVC - F PGFFDLAIGYEVGF - L PGFFDLAIGYEVGF - L POFYDLIFSCQVIF - F PSTYDLVIAYQVIF	HIRDKDALFGNI HIADKDALFSNL HIKDKDALFSNL LIPDKDALFSNI LIPDKDALFSNI LIQDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFRI HIRDKEALFRI HIRDKAAVLANI HIRDKAAVLANI HIRDKAAVLANI	VAGL KD NGR I LLMD TR SL NNGGF LLLAD RR HLSD GG VLALAD AR SL NNGGF LLLAD DR HLTNGGL I MAD SR HLQEEGL VMAD SR HL NPGG TLVLAD SA HLR PGG ALVL SD SA HLR PGG ALVL SD SA HLR PGG KLALID SA HLR I GGLAVMAE SA HLNDSGF F VAAE SR SL KPGG I LVMAE	FIANS-GSGVDV YIANG-RGQIVD FIATG-VSAINI FFB-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISM-ASAIHL CAADT-AKSIHL ILSNM-STPIDH TISNM-ASPIEH	HD VEVS I L TRONWLDLI EETASYNSSSEEWAD VI GETASYNSSSEEWAD VI GETSFSSTREGWNKLF EETSSYF I TKOHWYEQI GETSTFSSTREGWNKLF EASSSFFLTAEEWVDLI EASSSFFLTAEEWVDLI PASTAYFTPKEGWAEUI PASTAYFTPKEGWAEUI PESTTHFIPLGEWAELI	ARNFRLVEG SGNQLRVLEC SGNQLRVLEC SSNHLRLVEG SSNHLRLVEG SSNHLRLVEG SSNHLRLVEG SRNGLQLVEC AGHGFRLIRV ARSQLRVVEC ARNQLKVVEG	α8 α9 /DVSQEVANFLFDADFDANL DVSPQVANFLYDPECEENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF DISHEVSNYLYDAEFEENL DISPEIANCLOPPDYAANL /DVSREMANFLHDPDFDERL /DVSREIANFLHDDPFERNA IDASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASLGIANYLYDPNFTDNL
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PteB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamlV1-554	2768 IFHKNSAVAP 1075 VFARDSAKDA 1015 VFARDSAKDA 1016 IHRDSSAKDA 1017 VFARDSAKDA 1018 IFQRDSAKDA 1019 VFARDSAKDA 10101 IFQRDSAKDA 1140 IFNRDSSAKDA 1140 IFNRDSAKDA 1140 IFNRDSAKDA 1120 VFRDSAKAP 1122 VFRDSAKDA 1124 IFNRSSTDL 1580 LFHRDSSRDE 177 <lynrdsakqp< td=""> 1051 LHYQDSSRDP 166 IYNRDSAKQA</lynrdsakqp<>	- FPGRYDVAIGIEVT - FPDRYDVAFGFEVAT - FPDRYDVAFGFEVAT - FPDRYDVAFGFEVAT - FPGMYDLVLGFEVAG - FPDNYNLAFGFEVAH - FPRHYDLILGIEVAG - FPGFYDLIFGFEVAH - YVRDYDLWVGVEVVC - FPGFFDLAIGYEVGH - LPDTYDLIFSCQVIH - FPSTYDLVIAYQVIH - FLDEYDLGYSCQVIH	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIADKDALFSNL HIKDKSLFSNI LIQDKEALFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFANI HIDKAAVLANI HIRDKADLFNI HIRDKADLFNI HIRAKSDLFANI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLNGGLLIMAD SRHLQEGLLVMAD SRHLSPGGTLVAD SAHLRPGGALVLSD GALRPGGALVLSD GALRPGGALVLSD SAHLRIGGLAVMAE SCHLNDSGFFVAAE SCHLNDSGFFVAAE SKLKPGGLVMAE	F I ANS - GSGVDV Y I ANG - VGQ I VD F I ATG - VGA I NI F I ATG - VSA I NI F I A G - VSA I NI F I ANG - VSA I NI F I ANS - VVD I DH F I SNM - ASA I EH I LSNM - STP I DH I I SNN - ASP I EH I I SNN - ASP I EH I I SNN - ASP I EH I I SNN - ASP I EH	HD VEVSILTRONWLDLI EETASYNSSAEEWAD VI QOTASYSYFPSEWAD VI QETSTFSSTREQWNKLI EETSSYFITKOHWVQU QETSTFSSTREQWIKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYYVTRSKWAQLI PESTTHFIPLGEWAELI	ABHRLRITEL SRHNFRLVEG SGNOLVLEC SBHNFRLVDA SPNKLOLISA SCHRLRLVEG SRNGLQLVEC AGHGFRLIRV ARSGLRVVEC ARNALKVVEC ARNALKVVEC	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL IDISPEIANCLQDPDYDANL /DVSREMANFLHDPDFDERL /DVSREIANFLDDPEFERNA IDASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASUGIANFLHDADFEKNF /DASWEMGNFLEDPKFEENF
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 Mel/5/03-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 Mel/F/900-1239	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1140 IFQRDSAKDA 1150 IFQRDSAKDA 1161 IFQRDSAKDA 1152 IFRDSAKDA 11540 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 1051 IFRDSAKOP 1051 IFRDSAKOP 1052 IFRDSAKOP	- F PGRYD VAIGIEVT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PGRYD LIGIEVAG - F PGFYDLIGIEVAG - F PGFYDLIFGFEVAT - YVRDYDLMVGVEVVC - F PGFFDLAIGYEVGF - L PGHYDLIFSCQVIT - F PGTYDLVIAYQVIT - F PGTYDLVIAYQVIT - F PGTYDLVIAYQVIT	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKDALFSNI LIPDKDALFSNI LIPDKDALFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRKKEDVFLNI HIRAKSDLFANI HIRAKSDLFANI	VAGLKD NGR ILLMD RRHLSDGGFLLAD RRHLSDGGFLLAD AKSLNNGGFLLAD DRHTNGGFLLIMAD SRHLPGGFLLMAD SRHLPGGFLLAD AGALRPGGLVAS SGALRPGGLVAS SGALRPGGLAVMAE SGHLNDSGFFVAAE SGHLNDSGFFVAAE SGHLNDSGFFVAAE SGHLNAE	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FFS - T-GSSIDF FIAAG - VSAINI FVANT - LSP IEV FIANS - DVD IDH FLANT - VSP IEV FISMM - ASAIEH FVSGL - SSOIEH ILSNM - STPIDH IISNLPLTP IDD TMSNM - ASPIEH TISNLPLTP IED	HD VEVS I L TRONWLDLI EETASYNSSAEEWAD VI GDTASYSVTPSEWAD VI GETSFSSTREOWNKLI EETSSYFITKOHWVEQI GETSFSSTREOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI TKSTAYFATRSKWAELI	ARNHRLR ITEL SRHNFRLVEG SGNQLRVLEC SRHNFRLVEG SSHNLRLVDG SSHNLRLVDG SSHRLRLVEG TRHRLRVVDG SRNGLQLVEC ARNALRVVEG ARNALRVVEG ARNALRVVEG	a8 a9 VDVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI IDVSPQVANFLYDPECEENI IDISAEVLFLEDPSFDQNL IDISAEVNLYDAEFEENI IDISPEIANCLOPPDPARE VDVSREVANCL HNPDYAAOF IDISPEIANCLOPPDPARL VDVSREVANCL HNPDYAAOF IDISPEIANCLOPPDFERL VDVSREVANLLHDPFFERNA IDASEVALFLHDDPFERNA VDLSQEVALFLHDDDFESK VDLSQEVALFLHDDDFESK VDASLGIANLYDPNFTDNL VDASWEMGNFLEDPKFEENF VDASWEMGNFLEDPKFEENF
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PteB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamlV1-554	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1140 IFQRDSAKDA 1150 IFQRDSAKDA 1161 IFQRDSAKDA 1152 IFRDSAKDA 11540 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 1051 IFRDSAKOP 1051 IFRDSAKOP 1052 IFRDSAKOP	- F PGRYD VAIGIEVT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PGRYD LIGIEVAG - F PGFYDLIGIEVAG - F PGFYDLIFGFEVAT - YVRDYDLMVGVEVVC - F PGFFDLAIGYEVGF - L PGHYDLIFSCQVIT - F PGTYDLVIAYQVIT - F PGTYDLVIAYQVIT - F PGTYDLVIAYQVIT	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKDALFSNI LIPDKDALFSNI LIPDKDALFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRKKEDVFLNI HIRAKSDLFANI HIRAKSDLFANI	VAGL KD NGR ILLMD TRSLNNGGFLLLAD RRHLSDGVLALAD ARSLNNGGFLLLAD DRHLTNGGLLIMAD SRHLQEEGLLVMAD SRHLNGGFLLLAD SAHLRGGALVLSD SAHLRGGALVLSD SAHLRGGLAVMAE SRHLRIGGLAVMAE SRHLRNGGLFVAAE SRSLKPGGLLVMAE SRSMRNGGLUMAE SRSMRNGGLUMAE	FIANS - GSGVDV YIANG - RGQ I VD FIATG - VSAINI FFS - T. GSSIDF FIAAG - VSAINI FVANT - LSP I EV FIANS - DVD I DH FLANT - VSP I EV FISM - SSD I EH CAADT - AKSIHL I LSNN-STP I DH TMSMM - SP I EH I I SNL PLTP I DD TMSMM - VTP I EH	HD VEVS I L TRONWLDLI EETASYNSSAEEWADVI GETSTFSSTREGWNKL GETSTFSSTREGWNKL GETSTFSSTREGWNKL GETSTFSSTREGWNKL GETSTFSSTREGWNKL EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEGWAEU PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI PESTTFFVPGAEWAELI	ARNHRLR ITEL SRHNFRLVEG SGNQLRVLEC SRHNFRLVEG SSHNLRLVDG SSHNLRLVDG SSHRLRLVEG TRHRLRVVDG SRNGLQLVEC ARNALRVVEG ARNALRVVEG ARNALRVVEG	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL IDISPEIANCLQDPDYDANL /DVSREMANFLHDPDFDERL /DVSREIANFLDDPEFERNA IDASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASUGIANFLHDADFEKNF /DASWEMGNFLEDPKFEENF
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 Mel/E/903-1294 StiE/899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NeyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 Mel/F/900-1239 MtaF/898-1239	2768 IFHKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSKDD 1075 VFARDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1081 IFQRDSAKDA 1140 IFQRDSAKDA 1150 IFQRDSAKDA 1161 IFQRDSAKDA 1152 IFRDSAKDA 11540 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKAP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 10541 IFRDSAKOP 1051 IFRDSAKOP 1051 IFRDSAKOP 1052 IFRDSAKOP	- F PGRYDVAIGIEVT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDLILGIEVAG - F PGFYDLIFGFEVAH - YVRDYDLMVGVEVVC - F PGFFDLAIGYEVGH - L PGFYDLIFSCQVIH - F PSTYDLVIAYQVIH - F PSTYDLVIAYQVIH - F PSTYDLVIAYQVIH - F PSTYDLVIAFQVIH - F PSTYDLVIAFQVIH	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKDALFSNI LIPDKDALFSNI LIPDKDALFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRKKEDVFLNI HIRAKSDLFANI HIRAKSDLFANI	VAGL KD NGR I LLMD TR SL NNGGF LLLAD R HL SD G VLALAD AK SL NNGGF LLLAD DR HL TNGG LL IMAD SH HL QEEG LLMAD SH HL NGG F LLLAD SH LS NGG TLVLAD SH LS NGG TLVLAD SH LR NGG LAVMAE SH LS NGG I LMAE SK LK PGG I LMAE SK LK PGG I LMAE SK LK PGG I LMAE SK LK PGG LLMAE SK LK NGG LLMAE SK SK NGG YLVMAE SK SK NGG SK	FIANS-GSGVDV YIANG-RGQIVD FIATG-VSAINI FIATG-VSAINI FIAG-VSAINI FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-DVDIDH FIANS-STPIDH ISNUPLTPIDD TMSNM-SSPIEH IISNLPLTPIED TMSNM-VTPIEH 1257 1	HD VEVS I L TRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QETSTFSSTREQWNKL EETASYFITKOHWYEQI QETSTFSSTREQWNKL EEASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWADLI PASTAYFTPKEQWADLI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI PESTTFVPQAEWVELI 2571266	A BHRLR ITEL SRHNFRLVEG SGNQLRVLEC SSNHLRLVDAV SPNKLQLISA SEHRLRVVDGV SRNGLQLVEC AQHGFRLIRV ARNGLQLVECV ARNALKVVEGV ARNALKVVEGV ARNALRVVEGV ARNALRVVGAV	a8 a9 VDVSQEVANFLFDADFDANL IDVSPQVANFLYDPECEENI IDVSPQVANFLYDPECEENI IDISAEVLFLEDPSFDQNL IDISAEVNLYDAEFEENI IDISPEIANCLOPPDPARE VDVSREVANCL HNPDYAAOF IDISPEIANCLOPPDPARL VDVSREVANCL HNPDYAAOF IDISPEIANCLOPPDFERL VDVSREVANLLHDPFFERNA IDASEVALFLHDDPFERNA VDLSQEVALFLHDDDFESK VDLSQEVALFLHDDDFESK VDASLGIANLYDPNFTDNL VDASWEMGNFLEDPKFEENF VDASWEMGNFLEDPKFEENF
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 Mel/E/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 Mel/F/900-1239 MtaF/898-1239 Secondary Structure	2768 I F H KNS AVAP 1075 V F ARD S AKDA 1416 I H HRD S S KDD 1075 V F ARD S AKDA 1081 I F QRD S AKDA 1081 I F QRD S AKDA 1140 I F NRD S AKDA 1140 I F NRD S ARDA 1122 V F HRD S AAGP 1084 I F NRD S ATDA 1580 L F HRD S S AC 1071 L Y NRD S AKOP 1051 L H Y QD S S RD P 1052 L H Y QD S S RD P	- F PGRYD VAIGIEVT - F PORYD VAGGEVAT - F PORYD VAGGEVAT - F PORYD VAGGEVAT - F PORYD VIGFEVAG - F PONYNLAFGFEVAT - F PGFYDLILGIEVAG - F PGFYDLILGIEVAG - F PGFYDLILGIEVAG - F PGFYDLILGIEVAG - L PGHYDLILGIEVAG - F PGFYDLIAIGVEVVC - F PGFFDLAIGVEVVC - F PGFYDLVIAYOVIH - F PGYDLVIAYOVIH - F PSTYDLVIAYOVIH - F P F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIADKDALFSNI LIPDKDALFSNI LIPDKEALFSNI LIPDKEALFSNI HIKDKEALFSNI HIRDKEALFRI HIRDKEALFRI HIRDKAAVLANI HIRKEDVFLNI HIRKEDVFLNI HIRKSDLFANI HIRKSDLFANI HIRKSDLFANI	VAGLKDNGRILLMD RRHLSDGVLALAD AKSLNNGGFLLLAD RHLTNGGLLIMAD BRHLTNGGLLIMAD BRHLTNGGLLIMAD BRHLNGGFLLLAD ASHLSPGGTLVLAD SAHLRPGGALVAS BRHLNDSGFVAAE SRHLNDSGFVAAE SRHLNDSGFVAAE SRHLNDSGFVAAE SRHLNGGLLMMAE SRHLNGGLLMAE SRHLNGGLLMAE SRHLNGGLLMAE SRHLNGGLLMAE SRHLNGGLLMAE SRHLNGGYLMAE SRHLNGGYLMAE SRHLNGGYLMAE	FIANS-GSGVDV YIANL-RGCIVD FIATG-VSAINI FFS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FVSGL-SSDIEH CAADT-AKSIHL ILSNM-STPIDH TISNLPLTPIDD TMSNM-ASPIEH IISNLPLTPIED TMSNM-VSPIEH 1257 1	HD VEVS I L TRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI EETASYNSSAEEWADVI QETSTFSSTREQWNKLI EASSFFITKOHWVQU EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFIPKEQWAEDI AKSTAYFIPKEQWAEDI AKSTAYFIPKEQWAEDI PESTTOFVPVGEWAEDI PESTTOFVPVGEWAEDI 2571266	ARNER LE ITEL SRHNFRLVEG SGNQLRVLEC SSHNFRLVEG SSNHLRLVDG SSNHLRLVDG SSNHLRLVDG SSNHLRLVDG SSHLRLVVGG ARNERVVEG ARNNLKVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG	α8 α9 /DVSQEVANFLFDADFDANL DVSPQVANFLYDPECENI /DISRASLFLEOPSFDQNL /DISRASLFLEOPSFDQNL /DVSNEVANCLHNPDYAQF DISHEVSNYLYDAEFEENL /DVSREVANFLHDPAFEENL /DVSREVANFLHDPAFEENL /DVSREVANFLHDPFFENL /DVSREVANFLHDPFFENL /DVSREVANFLHDPFFENL /DVSREVANFLHDPFFENL /DVSREVANFLHDPFFENL /DASEIANFLDPFFENL /DASUPFENNFL /DATQEIANYLYDPNFTDNL /DATQEIANFLHDAFEKNF /DATQEIANFLHDAFEKNF /DATQEIANFLHDAFENF R N terminal β-strand
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/8899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1075 VF ARDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1140 IF NRDSSAKDA 1140 IF NRDSSAKAP 1122 VF HRDSAKAP 1122 VF HRDSAKOP 1084 IF NRNSSTDE 177 L Y NRDSAKQP 1051 L HYQDSSRDP 1062 L HYQDSSRDP 1051 L HYQDSSRDP 1051 L HYQDSSRDP 1051 L HYQDSSRDP 1051 L HYQDSSRDP	- F PGRYDVAIGIEVT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PGNYDLVLGFEVAG - F PGNYDLAFGFEVAF - F PRHYDLILGIEVAG - F PGFYDLIFGFEVAF - YVRDYDLMVGVEVVC - F PGFFDLAIGYEVGF - L PGHYDLIIAFQVIF - L POTYDLIFSCQVIF - F PSTYDLVIAYQVIF - F F F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKSLFSNI LIDDKDALFSNI HIKDKSLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFANI HIRDKADFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI	VAGLKDNGRILLMD TRSLNNGGFLLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLIMAD SRHLQEGLLVMAD SRHLQEGLLVAD SRHLRPGGLVLAD SALRPGGLVLSD SALRPGGLVMAE SGALRPGGLVMAE SGALRPGGLVMAE SGALRPGGLVMAE SRSLKPGGILVMAE SRSLKPGGLLMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE SRSLKPGGLVMAE 1245 000000000000000000000000000000000000	FIANS-GSGVDV YIANE-RGQIVD FIATG-VSAINI FFS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISMM-ASAIEH LISMM-STPIDH IISNLPLTPIDH IISNLPLTPIED TMSMM-ASPIEH IISNLPLTPIED TMSMM-VSPIEH 1257 1	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSYTPSEWADVI QETSTFSSTREQWIKLI EETSSYFITKOHWVEQ QETSTFSSTREQWIKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFIPKEQWAEDI AKSTAYFIPKEQWAEDI AKSTAYYVTRSKWAELI PESTTHFIPLGEWAELI PESTTFFPLRSKWAELI PESTTFVPQAEWVELI 2577266 1315 AARELML	AR HRLR ITEL SR HNFRLVEG SGNOLKVEC SR HNFRLVEG SR HLRLVDA SENKLQLISA SEHRLRLVEG SR NGLQLVEC ACHGFRLIRV AR SGLRVVEC ARNNLKVVEG ARNHRVVEC ARNHRVVEC ARNHRVVEC SR NGLRVVCC KNNLKVVEG ARNHRVVEC	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECENI /DISREASLFLEDPSFDQNL DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL /DVSRETANCLODPDYDANL /DVSRETANFLHDDPFFERNA /DASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASGEIANTLYDPRTDNL /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKKFENF /DASWEMGNFLEDPKKFENF /DASWEMGNFLEDPKKFENF /DASWEMGNFLEDPKKFENF
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CrcA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CrcB/439-821 HapA/1409-1769 NcyC/949-1331 PteB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1416 IH HRDSS CDU 1075 VF ARDSAKDA 1081 IF ORDSAKDA 1081 IF ORDSAKDA 1081 IF ORDSAKDA 1140 IF ORDSAKDA 1140 IF ORDSAKDA 1150 IF ORDSARDA 1122 VF HRDSARDA 11580 LF HRDSSRDE 1051 LF YQDSSRDE 1051 LF YQDSSRDE 1051 LF YQDSSRDE 1051 LF YQDSSRDE 1198 TQLETSVG IS 2877 KG	- F PGRYDVAIGIEVT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVAFGFEVAT - F PCRYDVIGFEVAG - F PCRYDLIGIEVAG - F PGFYDLIGIEVAG - F PGFYDLIGFEVAT - YVRDYDLMVGVEVVC - F PGFTDLAIGYEVGF - L PGHYDLIFGFEVAT - F PSTYDLVIAYQVIT - F F F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKALFSNI HIKDKSLFSNI LIPDKDALFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFENI HIRDKEALFENI HIRDKEALFENI HIRDKEAVLANI HIRKKEDVFLNI HIRAKSDLFANI HIRAKSDLFANI HIRNKADLFANI HIRNKADLFANI	VAGLKDNGRILLMD TRSLNNGGFLLLAD ARHLSDGGVLALAD ARSLNNGGFLLLAD CALLAD ARSLNNGGFLLLAD CALLAD	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISMM-ASAIEH ILSNM-STPIDH ILSNM-STPIDH ILSNM-STPIDH ILSNM-VSPIEH TLSNM-	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QETSTFSSTREOWNKLI EETSSYFITKOHWVQQI QETSTFSSTREOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYYVTRSKWAQLI PESTTHFIPLGEWAELI PESTTGFVPVGEWAELI 25771266 112571266	ARNFRLVEG SRHNFRLVEG SCHOLRVLEC SRHNFRLVDA SPNKLOLISA SPNKLOLISA SCHRLRLVEG TRHRLRVVEG ARNGLQLVEC ARNOLRVVEG ARNNLKVVEG ARNNLKVVEG ARNHLRVVEG ARNHLRVVEG ARNHLRVVEG ARNHLRVVEG	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL DISREASLFLEDPSFDQNL DOTSREASLFLEDPSFDQNL DASSEVGNFLVDPFLENS /DASWEMGNFLEDPKFENF /DATQEIANFLHDAFFKNF /
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/8899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1075 VF ARDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1140 IF NRDSSAKDA 1140 IF NRDSSAKADA 1122 VF HRDSAKADA 1122 VF HRDSAKDA 1122 VF HRDSAKDA 1122 VF HRDSAKDA 1122 VF HRDSAKDA 1124 VF HRDSSACA 1084 IF NRNSSTDL 1580 LF HRDSSACA 1051 LHYDDSSADA 1052 LHYDDSSADA 1051 LHYDDSSADA 1052 LHYDDSSADA 1051 LHYDDSSACA 1198 TQLETSVG IS 2877 KG LP 1184 ELVSQRFQLS 1524 SR IEASVGLT	- F PGRYDVAIGIEVT - FPORYDVAFGFEVAT - FPORYDVAFGFEVAT - FPORYDLIFGFEVAT - FPORYDLIFGFEVAT - FPORYDLIFGFEVAT - FPRHYDLILGIEVAG - FPGFYDLIFGFEVAT - VVRDYDLWVGVEVVC - FPGFFDLAIGYEVGF - LPGHYDLIFAGVIH - LPOTYDLIFSCQVIH - FLDEYDLGYSCQVIH - FPSTYDLVIAYQVIH - FPSTYDLVIA	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKALSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRAKSDLFANI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLIMAD SRHLQEGLLVMAD SRHLQEGLLVMAD SRHLQEGLLVMAE SRHLRNGGLVAAE SGALRPGGALVLSD SGALRPGGLIMAE SRHLRNGGLFVAAE SRSLKPGGILVMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLKPGGLIMAE SRSLRPGGLIMAE SRSLKPGGLIMAE S	FIANS-GSGVDV YIANG-RGGIVD FIATG-VSAINI FIATG-VSAINI FFB-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISNM-ASSIEH CAADT-AKSIHL ILSNM-STPIDH IISNLPLTPIDH IISNLPLTPIDH IISNLPLTPIED TMSNM-ASFIEH IISNLPLTPIED TMSNM-VSIEH ISNLPLTPIED TMSNM-VSIEH INGKWEAPPY NRAKLGALTPY VNRAKLGALTPY	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QETSTFSSTREQWIKLI EETSSYFITKOHWVEQ QETSTFSSTREQWIKLI EASSFFLTAEEWVDLI EAGSFLATAERWVEVI PASTAYFIPKEQWAEDI AKSTAYFIPKEQWAEDI AKSTAYFIPKEQWAEDI TKSTAYFATRSKWAELI PESTTFFPLGEWAELI TKSTAYFATRSKWAELI PESTTFPLQEWAELI 2571266 	A BHRLR IT LL SR HNFRLVEG SGNOLKVLEC SB HNFRLVEG SPNKLQ ISA SPNKLQ ISA SEHRLRVVEG SRNGLOLVEC ACHGFRLIRV AFSGLRVVEC ARNNLKVVEG ARNNLKVEG ARNNLRVVECA ARNLRVVECA ARNALRVVECA SRNGLRVVECA ARNALRVVECA ARNALRVVECA ARNALRVEC	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECEENI /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISHEVSNYLYDAEFEENL /DVSREVANCLHNPDYAAQF /DVSREVANCLHNPDYAAQF /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPDYAANL /DVSREVANCLOOPOYAANL /DVSREVANCLOOPOYAANL /DVSREVANCLOOPOYAANL /DASLGIANYLYDPONENENENENENENENENENENENENENENENENENENE
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEJ/018-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaF/903-1291 CroA/1253-1621 MelE/903-1294	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1016 IH HRDSS CDU 1075 VF ARDSAKDA 1081 IF ORDSAKDA 1081 IF ORDSAKDA 1140 IF ORDSAKDA 11580 LF HRDSSDD 1051 LHYDDSSDD 1051 LHYDDSSDD 1051 LHYDDSSDD 1051 LHYDSSDD 1051 LHYDSSDD 1051 LHYDSSDD 1051 LHYDSSDD 1051 LHYDSSDD 1051 LHYDSSDD 1144 EV SQRFOLS 1154 SR IEASVGLT 1144 ELYSCRFECLS	- FPGRYDVAIGIEVT - FPORYDVAFGFEVAT - FPORYDVAFGFEVAT - FPOMYDLVLGFEVAT - FPOMYDLVLGFEVAT - FPGFYDLIFGFEVAT - FPGFYDLIFGFEVAT - YVRDYDLWGVEVVC - FPGFFDLAIGYEVGF - LPDTYDLIFSCQVIF - FDSTYDLVIAYQVIF - FDSTYDLVIAYQVIF - FPSTYDLVIAYQVIF - FPSTYDLVIAFGAA - FYFYAAMDNFGAA - FYFYAAMDNFGAA - FYFYAAMDNFGAA	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKNALSNI HIKDKSLFSNI LIDDKDALFSNI HIKDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFANI HIRDKGUFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI HIRAKSLFANI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLIMAD SRHLQEGLLVMAD SRHLQEGLLVMAD SRHLRGGLVLAD SRHLRIGGLAVMAE SGALRPGGALVLSD SGALRPGGALVLSD SGALRPGGLLIMAE SGALRPGGLLIMAE SGALRGGLVMAE SGALRGGLLMAE SGALRGGLAVMAE SGALRGGLLMAE SGALRGGLAVAE SGALRGGLLMAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGGLAVAE SGALRAGAE SSSALRAGAE SSALRAGAE SGALRAGAE SSALLAGAE SSALRAGAE SSALRAGAE SSALRAGAE SSALRAGAE SSALRAGA	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FFS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISM-ASAIEH CADT-AKSIHL ILSNM-STPIDH IISNLPLTPIDD TMSNM-SPIEH TLSNM-VSPIEH TLSNM-VSPIEH TLSNM-VTPIED TMSNM-VSPIEH TLSNM-VTPIED TMSNM-VSPIEH LORVEAPAPY HNLEKVTHOTPY VNRAKLGALTPY VNRAKLGALTPY	HD VEVSILTRONWLDLI EETASYNSSAEEWADVI QETSTFSSTREOWNKLI EETSSYFITKOHWVQU QETSTFSSTREOWNKLI EASSFFLTAEWVDLI EASSFFLTAEWVDLI PASTAYFTPKEOWAEDI AKSTAYFTPKEOWAEDI AKSTAYFTPKEOWAEDI 2571266 2571266 2571266 ARELML PEALRAMLQSG	ARHRLRITEL SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRNKLOLISA SRNKLOLISA SRNKLOLISA SRNKLVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG ARNALKVEG	α8 α9 /DVSQEVANFLFDADFDANL IDVSPQVANFLYDPECENI /DISREASLFLEDPSFDQNL IDISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISAEVSNYLYDAEFEENL /DVSREMANFLHDPDFDERL /DVSREJANFLDDPFFRNA /DASEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASUGIANYLYDPNFTDNL /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFENF /DASWEMBNFL /DVSWPHKAROFQQVA /WFYDVSWPHKARGPGQREA
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290	2768 IF HKNS AVAP 1075 VF ARDS AKDA 1016 IH HRDS S ADD 1075 VF ARDS AKDA 1081 IF ORDS AKDA 1081 IF ORDS AKDD 1140 IF ORDS AKDD 1140 IF ORDS ARDA 1120 VF HRDS ARDA 1122 VF HRDS ARDA 1122 VF HRDS ARDA 1122 VF HRDS ARDA 1051 L HVQDS S ADD 1051 L HVQDS S ADD 1051 L HVQDS S ADD 1148 TQLETSVG IS 2877 KG C 1148 EV VSQRFOLS 1542 SR IEASVGLT 1144 ERVTERFKLN 1190 ALCKELKLD	- F PGRYD VAIG I EVT - F PCRYD VAFGF EVAT - F PCRYD VAFGF EVAT - F PCRYD VAFGF EVAT - F PCRYD VAFGF EVAT - F PGRYD LIGIEVAG - F PGF YDLIGIEVAG - F PGF YDLIGIEVAG - F PGF FDLAIGYEVGF - L PGF YDLIAFGY I F - F PSTYDLVIAYQVIH - F F F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKDALFSNI LIPDKDALFSNI HIKDKSLFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRAKSDLFANI	VAGL KD NGR I LLMD R LL NNGGF LLLAD R HL SD GVLALAD R HL SD GVLALAD SR HL QEGL LMAD SR HL NGGL LMAD SR HL NGGL LMAD SR HL NGGF LLAD AGAL RGG LLAD SA HL RGG LVAD SA HL RGG LVAD SA HL NGG LVAD SR	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FIATG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FIANS-DVDIDH FLANT-VSPIEV FISMM-ASAIEH ILSNM-STPIDH IISNLPLTPIDD TMSNM-STPIDH IISNLPLTPIDD TMSNM-VSPIEH TLSNM-VTPIEH 1257 1 CONCENTRON VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY LNAKOFCOLTPY	HD VEVS I L TRONWLDLI EETASYNSSAEEWAD VI QDTASYSVTPSEWAD VI GETSTFSSTREOWNKLI EETSSYFITKOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI 257TLFVPQAEWVELI 257TL66 1315 AARELML- SEFENG-ADWAQPPGQI ASYERALSSEVRY- AEFAKRDASGAQSDG	ARHRLR ITEL SRHNFRLVEG SGNQLRVLEC SGNQLRVLEC SRHNFRLVDG SSHNLRLVDG SSHRLRLVDG SRHRLRVVG SRNGLQLVEC ARNGLQLVEC ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVEG ARNALRVEG ARNALRVEG ARNALRVEG ARNALRVEG ARNALRVEG ARNALRVEG ARNALRVEG	α8 α9 VDVSQEVANFLFDADFDANL DVSPQVANFLYDPECENI VDISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL VDISREASLFLEDPSFDQNL VDVSNEVANCLHNPDYAAQF IDISAEVVRFLHDPAFEEHL DISEIANCLOPPJAAQF DSSEVGNFLVDPGLSKU VDVSREVANCLHNPDYAAQF IDISEIANCLOPPJANL VDVSREMANFLHDPDFDERL VDVSREMANFLHDPDFDERL VDVSREMANFLHDPFFRNA IDASSEVGNFLVDPGLDSML VDASEVGNFLVDPGLDSML VDASWEMGNFLEDPKFENF VDATQEIANFLHDAFFKNF VDATQEIANFLHDAFFKNF VDATQEIANFLHDAFFKNF VATQEIANFLWYPQFNENF N terminal β-strand EWLYEINWERKAR EWFHDVSWPHKARVPGDT WFHDVSWPHKARGPGQREA WLYESQWRRETPPLTA
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 Mel/E/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 Mel/F/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 Mel/E/903-1294 StiE/899-1290 CurL/924-1338	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1075 VF ARDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1140 IF NRDSSALDA 1140 IF NRDSSALDA 1120 VF HRDSARDA 1122 VF HRDSARDA 1122 VF HRDSARDA 1122 VF HRDSARDA 1128 F NRNSSTDL 1580 LF HRDSSEDE 177 LY NRDSAKQP 1051 LHYQDSSEDP 1051 LHYQDSSEDP 1051 LHYQDSSEDP 1051 LHYQDSSEDP 1184 ELVSQRFQLS 1524 SR I EASVGLT 1194 RV TERFKLN 1190 EALCKELKLD 1249 RELYEON-HD	F PGRYD VA IG IEVT F PGRYD VAFGFEVAT F PGFYDL IFGFEVAT F PGFYDL IFGFEVAT VVRDYDLWVGVEVVC F PGFFDLAIGYEVGF L PGHYDL IFAFQVIF F PGFYDL VIAYQVIF F FDSTYDLVIAYQVIF F PSTYDLVIAYQVIF F PSTYDLVIAYQVIF F PSTYDLVIAYQVIF F PGFYDLVIAYQVIF F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKNALSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAVLANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI ST LERKILSYVLFI LEKGLMSYALLI LEKGLNSYVLFI LSKGLASYVLFI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLLIMAD SRHLQEGLLVMAD SRHLQEGLLVMAD SRHLQEGLLVMAD SRHLRIGGLVLAD SRHLRIGGLVMAE SGALRAGGKLALID SRHLRIGGLVMAE SGALRAGGKLALID SRHLRIGGLVMAE SRSLKPGGILVMAE SRSLKPGGLLMAE SRSLSSLSSLSSLSSLSSLSSLSSLSSLSSLSSLSSLSSL	FIANS-GSGVDV YIANG-RGQIVD FIATG-VSAINT FFB-T-GSSIDF FIAAG-VSAINT FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISNM-ASPIEN ILSNM-STPIDH TMSNM-SFIEH ISNLPLTPIED TMSNM-SPIEH ISNLPLTPIED TMSNM-VSPIEH 1257 1 INGKWVEAPAPY HNLEKVTHOTPY VNRAKLGALTPY VNRAKLGALTPY VNGAKLGALTPY UNGKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALTPY NRAKLGALSY	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QETSTFSSTREQWIKLI EETSSYFITKOHWVEQ QETSTFSSTREQWIKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFIPKEQWAED AKSTAYFIPKEQWAED AKSTAYYTRSKWAELI PESTTFFPLGEWAELI TKSTAYFATRSKWAELI 2571266 12571266 257277777777777777777777777777777777	ARHRLRITEL SRHNFRLVEG SGNOLKVEC SBHNFRLVEG SSNHRLVDA SEHRLRVEG SRNGLOLISA SEHRLRVVEG SRNGLOLVEC ACHGFRLIRV AGHGFRLIRV ARNALKVEG ARNALKVE	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECEENI /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISHEVSNYLYDAEFEENL /DVSREMANFLHDPDFDEL /DVSREMANFLHDPDFESKY /DASLGIANYLYDPNFTDNL /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFENF 3 N terminal β-strand EWFHDVSWPHKARVPGDT
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290	2768 IF HKNSAVAP 1075 VF ARDSAKDA 1016 IF HRDSS CDD 1075 VF ARDSAKDA 1081 IF ORDSAKDA 1081 IF ORDSAKDA 1081 IF ORDSAKDA 1140 IF ORDSAKDA 1140 IF ORDSAKAP 1140 IF ORDSAKAP 1120 VF HRDSARDA 1122 VF HRDSARDA 1122 VF HRDSARDA 1122 VF HRDSARDA 1124 IF NRNSSTDL 1580 LF HRDSSDP 1051 L HYDDSSDP 1051 L HYDDSSDP 1052 L HYDDSSDP 1052 L HYDDSSDP 1051 L HYDDSSDP 1052 L HYDDSSDP 1052 L HYDDSSDP 1052 L HYDDSSDP 1052 L HYDDSSDP 1052 L HYDDSSDP 1161 IY RDSAKQP 1052 L HYDDSSDP 1161 IS 2877 KG LP 1184 EV SQRFOLS 1524 SR I EASVGL T 1198 FOL EKLKLD 1249 RELYEQN-HD 723 ERVLQEVR HD 723 ERVLQEVR HD	F PGRYD VAIG I EVT F F PGRYD VAFGFEVAT F PGRYD VAFGFEVAT F PGRYD VAFGFEVAT F PGNYDL VLGFEVAG F PGNYDL VLGFEVAG F PGFYDL I FGFEVAF F PGFYDL I FGFEVAF YVRDYDLMVGVEVVC F PGFFDLAIGYEVGF L PGHYDL I I AFGVI F F PSTYDL VI AFGVI F F PSTYDL F F F F PSTYDL F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F F	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKNALSNI HIKDKSLFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKEALFERI HIRDKEALFERI HIRDKEALFANI HIRAKSDLFANI	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLIMAD SRHLQEGLLVMAD SRHLQEGLLVMAD SRHLRGGLVLAD SRHLRGGLVLS SALRPGGALVLS SALRPGGLVLS SALRPGGLLVAD SRHLRIGGLFVAAE SRSLKPGGLLIMAE SRSLKPGGLIMAE S	FIANS-GSGVDV YIANL-RGQIVD FIATG-VSAINI FIATG-VSAINI FFS-T-GSSIDF FIAAG-VSAINI FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FISM-ASAIEH CADT-AKSIHL ILSNM-STPIDH IISNLPLTPIDH TMSNM-SPIEH TISNLPLTPIED TMSNM-VSPIEH TLSNM-VTPIED TMSNM-VSPIEH TLSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-VTPIED TMSNM-SPIEH LSNM-STPIEN SMACSAINS	HD VEVSILTRONWLDLI EETASYNSSAEEWADVI QETSTFSSTREOWNKLI EETSSYFITKOHWVQU QETSTFSSTREOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEOWAED AKSTAYFTPKEOWAED AKSTAYFTPKEOWAED 2571266 2571266 2571266 3ARELML PESTTIFYPQAEWVLI 2571266 3ARELML PESTTOFVPVCEWAELI PESTTOFVPVCEWAELI PESTTOFVPVCEWAELI 2571266 3ARELML PEALRAMLQSG- SEFNG-ADWAOPPGGG ASYERALRSEPVCY- AEFENGGADWAEPAGGG	ARHRLRITEL SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRNGLQLVEG ARNGLQLVEC ARNGLQLVEC ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG SRNGLRVVEG ARNALKVVEG ARNALKVEG SRNGLRVVEG ARNALKVEG	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL /DVSREMANFLHDPDFDERL /DVSREMANFLHDDPFFRNA /DASEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASWEMGNFLEDPKFENF /DVYYEINWERKAR /DVYYENNEKARPGDVA- /WYYELNWERKAR /WYYELNWERKAR /WYYELNWER
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEJ/018-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/809-1290 Curl/924-1338 CroB/439-821	2768 IF HKNS AVAP 1075 VF ARDS AKDA 1016 IF HRDS SKDD 1075 VF ARDS AKDA 1081 IF QRDS AKDA 1081 IF QRDS AKDA 1081 IF QRDS AKDA 1140 IF NRDS ARDA 1140 IF NRDS ARDA 1120 VF HRDS ARDA 1120 VF HRDS ARDA 1120 VF HRDS ARDA 1120 VF HRDS ARDA 1051 L HVDD S SPD 1051 L HVDD S SPD 1051 L HVDS SPD 1144 ELVSQRFGLS 1154 SR IEASVGLT 1144 ERVTERKLN 1190 RELYQN HD 723 ERVLQEVR PS 1669 DCLVRK - NLD 1241 AVGAA - G-	- F PGRYD VAIG I EVT F - F PORYD VAFGF EVAT - F PORYD VAFGF EVAT - F PORYD VAFGF EVAT - F PORYD VAFGF EVAT - F PGNYNL AFGF EVAT - F PGF VDL I LG I EVAG - F PGF YDL I LG I EVAG - F PGF YDL I LG I EVAG - F PGF FDL A I GYEVGF - LPOTYDL I FSCQVI F - LPOTYDL I FSCQVI F - F PSTYDL VI AYQVI F - F AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	HIRDKDALFGNI HIADKDALFSNL HIKDKALLSNI HIKDKALLSNI HIKDKSLFSNI LIPDKDALFSNI HIKDKSLFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKEALFERI HIRDKAVLANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI KISTYLFI LERGILSYVLFI LERGISYVLLT LERGLSYVLLT LERGLASYVLLT	VAGL KD NGR I LLMD RR LLSD GG FLLAD RR LLSD GG FLLAD RR LLSD GG FLLAD DR HL NGG FLLAD DR HL NGG FLLAD DR HL NGG FLLAD DR HL NGG FLLAD SR HL NGG FLLAD GALR PGG LLVAD GALR PGG LLVAD GALR PGG LLVAD SR HL NGG FVAAE SR SLK PGG LLMAE SR S	FIANS-GSGVDV YIANE-RGGIVD FIATG-VSAINT FIATG-VSAINT FVANT-LSPIEV FIANS-DVDIDH FLANT-VSPIEV FIANS-DVDIDH FLANT-VSPIEV FISNL-SSDIEH CAADT-AKSIHL IISNLPLTPIDD TMSNM-SSPIEH IISNLPLTPIDD TMSNM-VSPIEH 1257 1 COOL VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY LNRENLAQLSSY ANRERLMRLVQS NRELMRLVG	HD VEVS I L TRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QDTASYSVTPSEWADVI QETSTFSSTREOWNKLI EETSSYFITKOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYYVTRSKWAELI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI 25771266 112571266 2571266 2571266 2571266 2571266 2571266 2571266 2571267 2571266 2571267 257127 25	ARHRLRITEL SRHNFRLVEG SGNQLRVLEC SRHNFRLVEG SRHNLRLVDA SPNKLQLISA SRHRLRVEG TRHRLRVVGG SRNGLQLVEC ARNQLRVVEC ARNNLKVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG ARNNLRVVEG ARNALRVEG ARNAL	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECEENI /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAAQF IDISHEVSNYLYDAEFEENL IDISHEVSNYLYDAEFEENL /DVSREMANFLHDPDFDEL /DVSREMANFLHDPDFESKY /DASLGIANYLYDPNFTDNL /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFEENF /DASWEMGNFLEDPKFENF 3 N terminal β-strand EWFHDVSWPHKARVPGDT
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MelF/900-1239 MtaF/808-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 Curl/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331	2768 I F HKNS AVAP 1075 VF ARD S AKDA 1075 VF ARD S AKDA 1081 I F QRD S AKDA 1081 I F QRD S AKDA 1081 I F QRD S AKDA 1140 I F QRD S AKDA 1140 I F QRD S AKDA 1140 I F QRD S AKDA 1122 VF HRD S AKAPA 1122 VF HRD S AKAPA 1122 VF HRD S AKAPA 1124 VF HRD S AKAPA 1084 I F NRNS STDL 1580 L F HRD S S TOE 177 L Y NRD S AKAPA 1061 I Y NRD S AKAPA 1052 L H Y QD S S TOP 1051 L H Y QD S S TOP 1052 L H Y QD S S TOP 1051 L H Y QD S S TOP 1144 E V S Q F C S S 1524 S R I E AS V G I S 1144 E V S Q F C S S 1154 S R I E AS V G I S 1154	- F PGRYD VA IG I EVT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PGMYDL VLGFEVAG - F PGNYNL AFGFEVAF - F PCHYDL I LGIEVAG - F PGFYDL I FGFEVAF - VVRDYDLMVGVEVVC - F PGFFDLA I GYEVGF - L PGHYDL I I AFQVIF - L PDT YDL I FSCQVIF - F PSTYDL VI AYQVIF - T AAALKSYDL GFGAF E VVR SF G SYNUL GK E VVR SF G SYNUL GR V - T AAALKSYDL GFGAF	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKNALSNI HIKDKSLFSNI HIKDKEALFSNI HIRDKEALFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFSNI HIRDKEGLFANI HIRKSDLFANI HIRKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI LEKGUISYVLLN LEKGILSYVLT LEKGLSYVLLT LERGLISYVLLT LERGLISYVLLT LERGLISYVLLT LERGLISYVLLT LERGLISYVLLT	VAGLKDNGRILLMD TRSLNNGGFLLAD ARHLSDGGVLALAD ARSLNNGGFLLAD DRHLTNGGLIMAD BRHLQEGLLVMAD BRHLQEGLLVMAD BRHLNPGGFLULAD BRHLNPGGFLULAD BRHLNDSGFFVAAE SGALRPGGALVLSD GALRPGGALVLSD GALRPGGLUMAE SGHLNDSGFFVAAE SGALRYGGLIMAE	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FIATG - VSAINI FVATT - LSPIEV FIANS - DVDIDH FLANT - VSPIEV FISM- ASAIEH CAADT - AKSIHL ILSNM - STPIDH IISNLPLTPIDH TISNLPLTPIDH TISNLPLTPIED TMSMM - ASPIEH IISNLPLTPIED TMSMM - VSPIEH ISNLPLTPIED TMSMM - ASPIEH INGRVLAAPTPY VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY LNRCHAAPTPY VNRAKLGALTPY LNRCHAAPTPY VNRAKLGALTPY LNRCHAAPTPY VNRAKLGALTPY LNRCHAAPTPY VNRAKLGALTPY LNRCHAAPTPY VNRAKLGALTPY LNRCHAAPTPY SNRCHARLSSY	HDVEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSYTPSEWADVI QETSTFSSTREQWNKLI EETSSYFITKEHWVEQI QETSTFSSTREQWIKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI 2571266 10000000000000000000000000000000000	ARHRLR ITEL SRHNFRLVEG SCALKVLEC SRHNFRLVEG SRNLQLISA SPNKLQLISA SCHRLRLVEG SRNGLQLVEC ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVVEG ARNALKVEG	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLFDPSCEENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAAQF /DVSREIANFLDDPEFEENL /DVSREIANFLDDPFFERNA /DASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASGEIANFLHDDFFERNF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DASWEMGNFLEDPKFENF /DVYPELAWEMKAR /DVYPENPRRGDVAA- /DVYPENPRRGDVAA- /DVYPENPRRRGDVAA- /WYENKKAR // PASWEMRENF // PASAAGRWLVLGD-
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamIV/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEB/918-1251 AjuK/1391-1766	2768 IF HKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSDD 1075 VFARDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 1150 IF RRDSAE 1150 IF RRDSAE 1161 IF RRDSAE 1198 TQLETSVGIS 2877 KG IF 1198 TQLETSVGIS 2877 KG IF 1194 IF VSQRFQLS 1524 SR IEASVGIS 1524 SR IEASVGIT 1190 EALCKELKLD 1249 RELYEQN HD 723 ERVIGEN - A 1193 AGESGE - A 1689 AYLAER - V	- F PGRYD VA IG I EVT - F PORYD VA FGF EVAT - F PGRYD L I G I EVAG - F PGF YDL I FGF EVAT - YVRD YDL I G I EVAG - F PGF TDL I G I EVAG - PGF FDL A I GYEVGF - L POT YDL I FSC QVI F - F PGF YDL VI AYQVI F - F PST YDL VI AYQVI F - T AYAL SYDL G F - T AYAL SYDL G F - T AYAL SYDL G F - A SL H K GP NOL AGL	HIRDKDALFGNI HIADKDALFSNL HIKDKNALSNI HIKDKNALSNI HIKDKSLFSNI HIKDKSLFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFRI HIRDKEALFRI HIRDKEALFRI HIRDKEALFRI HIRDKEALFRI HIRDKAAVLANI HIRKSDLFANI HIRKSDLFANI HIRKSDLFANI HIRNKADLFANI HIRNKADLFANI HIRNKADLFANI HIRNKALFANI HIRNKALFANI HIRNKALFANI HIRNKALFANI LEKGUISYVLLI LEKGLISYVLLI LEKGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI LRGLISYVLLI	VAGL KD NGR I LLMD TR SLNNGGF LL AD AR HL SDGG VLALAD AR SLNNGGF LL AD AR SLNNGGF LL AD DR HL NGG LL IMAD SR HL VGG LL IMAD SR HL NGG LL VAD SR HL NGG LL VAD SALR PGG LL VAD GALR PGG LL VA SALK PGG LL VAD SR HL R IGG AVMAE SCHL ND SGF F VAAE SCHL ND SGF F VAE SCHL ND	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FIATG - VSAINI FVAT - LSFIEV FIANS - DVD DH FLANT - VSPIEV FISMM - ASAIEH ISNM - ASAIEH ILSNM - STPIDH ISNLPLTPIDH ISNLPLTPIDH ISNLPLTPIDH TMSNM - ASPIEH ILSNM - VTPIEH NGKW - AFAPY NRAKLGALTPY VNRAKLGALTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY LNAKQFEQLTPY NRAKLGALTPY NRSMGIER	HD VEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSYTPSEWADVI QETSTFSSTREOWNKLI EETSSYFITKOHWVQU QETSTFSSTREOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEOWAEU PASTAYFTPKEOWAEU PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI PESTTGFVPVGEWAELI PESTTGFVPVGEWAELI PESTTLFVPQAEWVELI 2571266 JJC 1315 AARELML PEALRAMLQSG- SEFENG-ADWAOPPGQI AEFEARGADWAPPGQI AEFEARGADWAPAGQI AEFAKRRDASGAQSDG SEVAV- QEVVRR-RSGPSA- AOVSPAELRHEGLVAI AAVAGATAGPREDDDGF	ARHRLRITEL SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHNFRLVEG SRHRLRVEG ARNGLQVEC ARNGLQVEC ARNGLRVVEC ARNNLKVVEG ARNNLKVVEG ARNNLKVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVEG A	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAQF IDISHEVSNYLVDAEFEENL IDISEVSNYLVDAEFEENL IDISEVSNYLVDAEFEENL /DVSREIANFLDDPFFRNA DASSEVGNFLVDPGLDSML /DLSQEVALFLHDDDFESKY /DASEGIANYLYDPNFTDNL /DASWEMGNFLEDPKFEENF /DATQEIANFLHDADFEKNF /DASWEMGNFLEDPKFEENF /DATQEIANFLHDAFENFE RNEMMINIA /VDATQEIANFLHARVPGDT /VDATQEIANFLHARVPGDT /VDATQEIANFLHARVPGDT /VDATQEIANFLHARVPGDT /VDATQEIANFLHARVPGDVA- /VFYENWEPRPRRQDVA- /WYEYEXQWRETPPLTA /WYELKWETKKHQQLLSPN /WYELKWETKKHQQLSPN /WYELKWETKHQQLSPN /WYELKWETKHQQLSPN /WYELKWETKHQQLSPN
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamV/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CroB/490-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtzB/918-1251 AjuK/1391-1766 BarF/1-366	2768 IF HKNS AVAP 1075 VF ARDS AKDA 1016 IF NRDS S DD 1075 VF ARDS S DD 1075 VF ARDS S ADD 1075 VF ARDS S AKDA 1081 IF ORDS S AED 1140 IF NRDS S AED 1140 IF NRDS S AED 1140 IF NRDS S AED 1150 LF HRDS S AED 1051 LF NRDS S AED 1051 LF VOS S AED 1051 LF VOS S AED 1051 LF VOS S AED 1198 TOLETSVG IS 2877 KG LF 1198 TOLETSVG IS 2877 KG C 1198 TOLETSVG IS 1194 EV VORFOLS 1524 SR IEASVGLT 1194 EV VORFOLS 1190 ALCKELKLD 1249 RELYEQN - HD 723 ERVLQEVR PS 1669 DCL VRK - NLD 1231 AAVGAA - G - 1193 AGESGEG - AI 1669 AYLAER - YD 287 AKC OG - YD	- F PGRYD VAIG I EVT - F PORYD VAFGFEVAT - F PORYD VAFGFEVAT - F PORYD VAFGFEVAT - F PORYD VAFGFEVAT - F PGNYDL VLGFEVAG - F PGFYDL I LGIEVAG - F PGFYDL I LGIEVAG - F PGFYDL I LGIEVAG - F PGFYDL I LGFEVAT - YVRDYDLMVGVEVVC - F PGFFDL A I GYEVGF - L PGHYDL I LAFGVI F - F PSTYDL VI AYGVI F - F PSTYDL VI AFGVI F - F PSTYDL VI AYGVI F - F PSTYDL VI AFGVI F - F PSTYDL F - F PSTY	HIRDKDALFGNI HIADKDALFSNL HIKDKALLSNI HIKDKALLSNI HIKDKSLFSNI LIPDKDALFSNI HIKDKSLFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI KIRSULFYLL HIRAKSDLFANI	VAGL KD NGR I LLMD REL NNGGF LLLAD REL SDGVLALAD REL SDGVLALAD SRHL NGGF LLAD SRHL NGGL I MAD SRHL NGGL I MAD SRHL NGGL I VAAD SRHL NGGGT LVLAD SALL PGGAL V SD SALL PGGAL V SD SALL NGGT LVLAD SALL NGGT VALL NGGT VALL NGT VALL NG	FIANS - GSGVDV YIANG - RGG I VO FIATG - VSAINT FIATG - VSAINT FVANT - LSPIEV FIANS - DVD IDH FLANT - VSPIEV FIANS - DVD IDH FLANT - VSPIEV FVSGL - SSDIEH CAADT - AKSIHL ILSNN-STPIDH TMSNM - SSPIEH ISNLPLTPIDD TMSNM - VSPIEH 1 SNLPLTPIDD TMSNM - VSPIEH 1 SNLPLTPIDD TMSNM - VSPIEH 1 LSNM - VTPIEH 1257 1 COOL VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY VNRAKLGALTPY NRAKLGALTPY	HD VEVS I L TRONWLDLI EETASYNSSAEEWADVI QDTASYSVTPSEWADVI QETSTFSSTREOWNKLI EETSSYFITKOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PEVSWTSTLTEYADI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYFTRKEQWAEDI 257TL67VPVGEWAELI PESTTGVPVGEWAELI 257T266 11315 AARELML- 257T266 25FENG-ADWAOPPGQI ASYERALSSE SEFENG-ADWAOPPGQI ASYERALSSE SEFENG-ADWAOPPGQI ASYERALSSE SEVAV- AEFAKRDASGAOSDG SEVAV- AEFAKRDASGAOSDG SEVAV- AAVAGEALVQAEAQTH SKILENSPNGKELLLPA	ARHRLR ITEL SRHNFRLVEG SGNQLRVLEC SGNQLRVLEC SRHNFRLVDG SPNKLQLISA SPNKLQLISA SEHRLRVEG TRHRLRVVGG ARNQLRVVEC ARNALKVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVEG ARNAL	α8 α9 /DVSQEVANELFDADFDANL DVSQEVANELYDPECEENI /DISREASLFLEDPSFDQNL DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF DISHEVSNYLYDAEFEENL /DVSREVANFLHDPAFEENL /DVSREVANFLHDPFDEL /DVSREVANFLHDPFFERNA /DASSEVANFLHDDPFFERNA /DASSEVANFLNDPFFERNA /DASSEVANFLNDPFFERNA /DASSEVANFLNDPFFERNA /DASSEVANFLNDPFFERNA /DASSEVANFLNDPFFERNA /DATQEIANFLHDADFEKNF /DATQEIANFLHDADFEKNF /DATQEIANFLHDAFFPKFENF /DATQEIANFLHDAFFPKFENF /DATQEIANFLHDAFFPKRGDVVA /DATQEIANFLHDAFFPKRGDVVA /DATQEIANFLHARCPGQREA WLYEINWERKAR
StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEB/918-1251 AjuK/1391-1766 BarF/1-366 CtaF/899-1240 JamIV/1-354 MelF/900-1239 MtaF/898-1239 Secondary Structure StiD/914-1289 CalF/2606-2945 CtaE/903-1291 CroA/1253-1621 MelE/903-1294 StiE/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 NcyC/949-1331 PtEB/918-1251 AjuK/1391-1766	2768 IF HKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSDD 1075 VFARDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1081 IF QRDSAKDA 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 1140 IF NRDSAE 11580 LFHRDSAE 11580 LFHRDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1051 LFYQDSSDE 1144 ELVSQRFQLS 1524 SR IEASVGLT 1144 ELVSQRFQLS 1524 SR IEASVGLT 1144 ELYERFKLN 1190 EALCKELKLD 1249 RELYERF.HD 723 ERVGAAG 1193 AGESGEG - A 1689 AYLQEVR SD 287 ARLCQG - YD 1160 ERVTRD - SD 276 TPL TO D	- F PGRYD VA IG I EVT - F PORYD VA FGF EVAT - F PGRYD L I G I EVAG - F PGRYD L I G I EVAG - F PGF POL I FGF EVAT - YVRD YD L MVGV EVVC - PGF FD L I GYEVG - L PGHYD L I AF QV I - F PGF TD L I GF EVAT - F PSTYD L V I AY QV I - T S GY PARAMANF GAA EVT GN FG SY ENVYKA D V R S FG SY ENVYKA E VV R S FG SY DN G GAA E V C R S FG SY DN G GAA E V C R S FG SY DN G GAA E V C R S FG SY DN C AGA E V C R S FG SY DN C S C C C C C C C C C C C C C	HIRDKDALFGNI HIADKDALFSNI HIKDKALLSNI HIKDKALLSNI HIKDKSLFSNI HIRDKEAL	VAGL KD NGR I LLMD TR SLNNGGF LL AD RAHLSDGGVLALAD ARSLNNGGF LL AD ARSLNNGGF LLAD DR HLTNGGL I MAD SR HLQEGL I MAD SR HLPGGF LL AD AGALRPGG LL VAD SAHLPGGF LL AD SAHLPGGF LL AD GALRPGG LL VAD SAHLPGG LL VAD SAHLPGG LL VAD SAHLPGG LL VAD SALSPGG L VAAE SCALRPGG L VAAE SCALRPG L VAAE SC	FIANS - GSGVDV YIANL - RGQ IVD FIATG - VSAINI FIATG - VSAINI FVAT - LSPIEV FIANS - DVD DH FLANT - VSPIEV FISMM - ASAIEH ISNM - ASAIEH ILSNM - STPIDH ISNLPLTPIDH ISNLPLTPIDH ISNLPLTPIDH TMSNM - ASPIEH ILSNM - VTPIEH 1257 1 COOL VACK - AASAI NGKW - AASAI NGKW - AASAI VOGALGALTPY LNGKU - AASAI LNGK - CALTPY LNGK - CALTPY	HD VEVSILTRONWLDLI EETASYNSSAEEWADVI QDTASYSYFSSAEEWADVI QDTASYSYFSSTREOWNKLI EETSSYFSTREOWNKLI EETSSYFITKOHWVQU QETSTFSSTREOWNKLI EAGSFLATAERWVEVI EAGSFLATAERWVEVI PASTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI AKSTAYFTPKEQWAEDI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI 2577266 JJC 1315 AARELML PEALRAMLQSG- SEFENG-ADWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUWAOPPGQI ASFENGSAUSAGAQSI SEVAV- QUVRR - RSGPSA- ANVSPAGLRHEGL VAI AAVAGATAGPREDDDGF ARVAGEALVQAEAQTH SKIIENSNGKPLLP	AB HR LR IT EL SR HNFRL VEG SG NOLR VLEC SB HNFRL VDA SPNKLOL ISA SENRLRL VEG THR LR VVEG ARNOL VVEG ARNOL VVEG ARNNL KVVEG ARNNL KVVEG ARNNL KVEG ARNNL KVEG ARNAL RVEG ARNAL RVEG ARN	α8 α9 /DVSQEVANFLFDADFDANL DVSQEVANFLYDPECENI /DISREASLFLEDPSFDQNL DISAEVVRFLHDPAFEEHL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSREVANCLHNPDYAQF IDISEVSNLVDAEFEENL IDISEVSNLVDAEFEENL IDISEVSNLVDAEFEENL IDISEVSNLVDAEFEENL /DVSREIANFLDDPFFRNA /DASSEVGNFLVDPGLDSML /DLSQEVALFLHDDFFSKY /DATQEIANFLHDAFEKNF /DATQEIANFLHDAFEKNF /DATQEIANFLHDAFFKNF /DASASGKGAW- /PRASARAGKUVLGD-
StiD/914-1289 CaiF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1294 StiE/889-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 Nev(Cr49-1331 PtzB/918-1251 Ajuk/1391-1766 BarF/1-366 CtaF/899-1240 JamN/1-354 MalF/898-1239 Secondary Structure StiD/914-1289 CaiF/2606-2945 CtaE/903-1291 CroA/1253-1621 MalE/903-1291 CroA/1253-1621 MalF/899-1290 CurL/924-1338 CroB/439-821 HapA/1409-1769 Nev(Cr49-1331 PtzB/918-1251 Ajuk/1391-1766 BarF/1-366 CtaF/899-1240	2768 IF HKNSAVAP 1075 VFARDSAKDA 1416 IHRDSSDD 1075 VFARDSAKDA 1081 IFORDSAKDA 1081 IFORDSSDE 614 VFNRDSAKDA 1140 IFNRDSAFDA 1140 IFNRDSAFDA 1140 IFNRDSAFDA 1122 VFHRDSAFDA 1122 VFHRDSAFDA 1122 VFHRDSAFDA 11580 LFHRDSSFDE 177 LYNRDSAFOP 1051 LFYQDSSFDE 1051 LFYQDSSFDE 1052 LFYQDSSFDE 1052 LFYQDSSFDE 1198 TOLETSVGIS 1524 SRIEASVGIS 1524 SRIEASVGIS 1524 SRIEASVGIS 1524 SRIEASVGIS 1524 SRIEASVGIS 1198 TOLETSVGIS 1194 EVSQRFOLS 1524 SRIEASVGIS 1194 ELYSQRFOLS 1524 SRIEASVGIS 1194 RELYERFKLN 1190 EALCKELKLD 1249 RELYEON-HD 723 ERVLQEN-HD 723 ERVLQEN-HD 723 RVLQEN-S 1669 AYLAER-YD 287 AKCOG - YD 1160 ERVTRD - SD 276 TRLTQD - YD 1161 ERVTRD - SD 276 TRLTQD - YD 1161 ERVTRD - SD	- F PGRYD VAIG I EVT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PCRYD VAFGFEVAT - F PGMYDL VLGFEVAG - F PGNYNL AFGFEVAT - FPGFYDL I GI EVAG - F PGFYDL I GFEVAT - YVRD YDLMVGFEVAT - YVRD YDLMVGFEVAT - YVRD YDLMVGFEVAT - FPGFFDL I GYECO - L PGHYDL I GFEVAT - F PGFYDL VI AFGVI F - F PGFYDL VI AFGVI F - F PSTYDL VI AFG F - T AAAL SYDQL GRL DU VARVG SYDRI GRV - T AAAL KSYDQL GRL DU VARVG PHEL GEL DVT R HL HG I HML GEL E T A CHL KGP PLE GL	HIRDKDALFGNI HIADKDALFSNL HIKDKALLSNI HIKDKALLSNI HIKDKSLFSNI LIPDKDALFSNI HIKDKSLFSNI HIRDKEALFSNI HIRDKEALFSNI HIRDKEALFERI HIRDKEALFERI HIRDKAAVLANI HIEMKQENVLNI HIRAKSDLFANI HIRAKSDLFANI HIRAKSDLFANI HIRNKADLFANI HIRNKADLFANI EKGLMSYALLV LEKGUNSYALLV LEKGUNSYALLV LEKGUNSYALLV LEKGLSYVLFI LEKGUSYVLTI LEKGLSYVLTI LRGGLSYVLTI LSRLAAYGLFM	VAGL KD NGR I LLMD TR SL NNGGF LLLAD R HL SD G VLALAD R HL SD G VLALAD SH L QEG LLAD SH L NGGF LLAD SH L NGG LL MAD SH L NGG LL VAD SH L NGG LLVAD SH L NGG T LVAD SH L NGG T VAA SAL RGG L VAA SAL RGG L VAA SG L ND SG F VAA SG L ND SG F VAA SG L ND SG F VAA SG L VD G L VAA SG L VD G L VAA SG L VAA	FIANS - GSGVDV YIANG - RGG I VO FIATG - VSG I NT FIATG - VSG I NT FIATG - VSG I DF FIAGS - DVD I DH FLANT - VSP I EV FIANS - DVD I DH FLANT - VSP I EV FVSGL - SSD I EH CAADT - AKSIHL I SNN - SSD I EH CAADT - AKSIHL I SNL PLTP I DD TMSNM - ASP I EH I SNL PLTP I EH 1257 1 COOL I NGKWVEAPAPY NRAKLGAL TPY VNGAKLGAL TPY VNGAKLGAL TPY UNGAKLGAL TPY LNGCVLAAPTPY VNGAKLGAL TPY LNGCVLAAPTPY VNGAKLGAL TPY LNGCVLAAPTPY VNGAKLGAL TPY LNGCKLGNL TPY I NHAHLGAPTPY I NHAHLGAPTPY I NEREL SNL PTPY I NEREL SNL PTY I NEREL SNL PTPY I NEREL SNL PTPY I NEREL SNL PTPY	HD VEVS ILTRONWLDLI EETASYNSSAEEWADVI QDTASYSYFSSAEEWADVI QDTASYSYFSSTREOWNKLI EETSSYFSTREOWNKLI EASSYFITKOWNKLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI EASSFFLTAEEWVDLI PASTAYFTPKEQWAEDI AKSTAYYVTRSKWADLI PESTTHFIPLGEWAELI PESTTHFIPLGEWAELI 25771266 10000000000000000000000000000000000	AR HRLR IT EL SR HNFRLVEG SGNQLRVLEC SGNQLRVLEC SR HNFRLVDG SNHLRLVDG SNHLRLVDG SR HRLRVEG TR HRLRVVDG SR NGLQLVEC AR NGLRVVEC AR NGLRVVEC ARNNLKVVEG ARNNLRVVEG ARNNLRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVVEG ARNALRVEG ARNAL AR	α8 α9 /DVSQEVANELFDADFDANL DVSQEVANELYDPECEENI /DISREASLFLEDPSFDQNL DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DISREASLFLEDPSFDQNL /DVSNEVANCLHNPDYAAQF DISHEVSNYLYDAEFEENL /DVSREVANFLHDPAFEENL /DVSREVANFLHDPFDEL /DVSREMANFLHDPFFERNA /DASSEVANFLYDPGLDSML /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DASSEVANFLYDPFERNA /DATQEIANFLHDADFEKNF /DATQEIANFLWDPFENENF /DATQEIANFLWDPFENENF /DATQEIANFLWDPFENENF /DATQEIANFLWDPFENENF /DATQEIANFLWERKAR /DVYENWENKAR /WYELNWERKAR /WYELWWENKKAR /WYELWWENKKAROQLSPN /WYELWWENKKHQQLSPN /WYELWWENKKHQQLSPN /WYELWWENKKHQQLSPN /WRASASKGAW /WRASASKGAW

Figure S1. Sequence alignment of PKS *O*-MTs. Names are colored based on predicted substrate ((S)-OH, blue; (R)-OH, green; keto; black). Substrate predictions are based upon structures of the final metabolites, the inclusion of a KR in *O*-MT containing modules, and KR sequence motifs that correlate with the stereochemical configuration of the resulting methoxy. Stars indicate sites

of mutagenesis; arrows represent experimental N- and C-termini tested for StiD (blue) and fragment N- and C-termini for StiE (green) and CurL (orange). Gray ovals represent residues in the StiD or StiE *O*-MT dimer interface. Pathway abbreviations (GenBank Accision codes, *trans/cis*-AT, producing organism) are as follows: Sti- stigmatellin (CAD19088.1, CAD19089.1, *cis*-AT, *Stigmatella aurantiaca*); Cal- calyculin A (BAP05594.1, *trans*-AT, *Candidatus entotheonella* sp. A); Cta- cystothiazole A (AAW03329.1, AAW03329.1, *cis*-AT, *Cystobacter fuscus*); Cro- crocacin (AIR74910, AIR74911.1, *cis*-AT, *Chondromyces crocatus*); Mel-melithiazol (CAD89776.1, CAD89777.1, *cis*-AT, *Melittangium lichenicola*); Cur- curacin A (AEE88278.1, *cis*-AT, *Moorea producens*); Hap- haprolid (AOG74798.1, *cis*-AT, *Byssovorax cruenta*); Nyc- nannocystin A (ALD82523.1, *cis*-AT, *Nannocystis sp. MB1016*); Ptz- patellazole (AFX99666.1, *trans*-AT, *Candidatus Endolissoclinum faulkneri L2*); Aju- ajudazol (CAQ18838.1, *cis*-AT, *Chondromyces crocatus*); Bar- barbamide (AAN32980.1, *cis*-AT, *Lyngbya majuscula*); Jam- jamaicamide (AAS98785.1, *cis*-AT, *Lyngbya majuscula*); Mta- myxothiazol (AAF19814.1, *cis*-AT, *Stigmatella aurantiaca*).

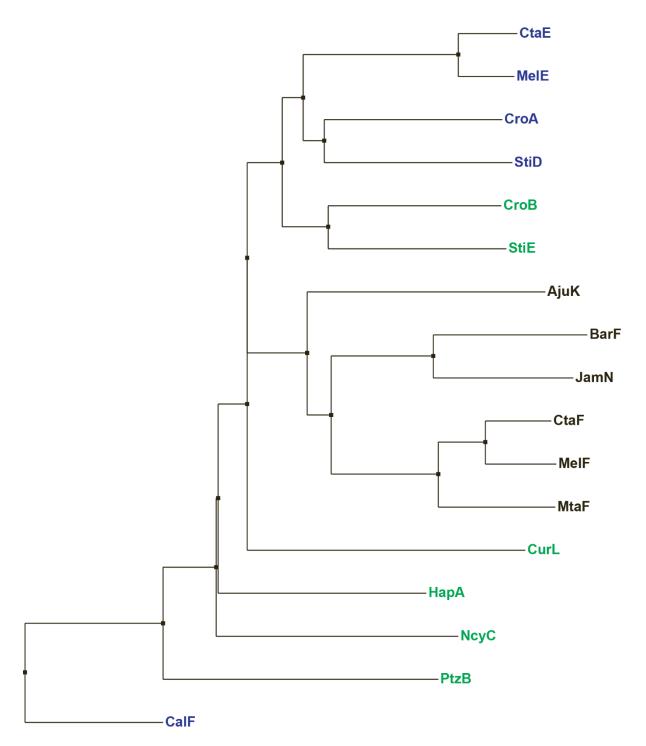


Figure S2. Dendrogram of PKS *O*-MTs from Figure S1. Names are colored based on predicted substrate ((*S*)-OH, blue; (*R*)-OH, green; keto, black) as in Figure S1. All *O*-MTs are from *cis*-AT PKS pathways except PtzB and CalF, which are from *trans*-AT pathways.

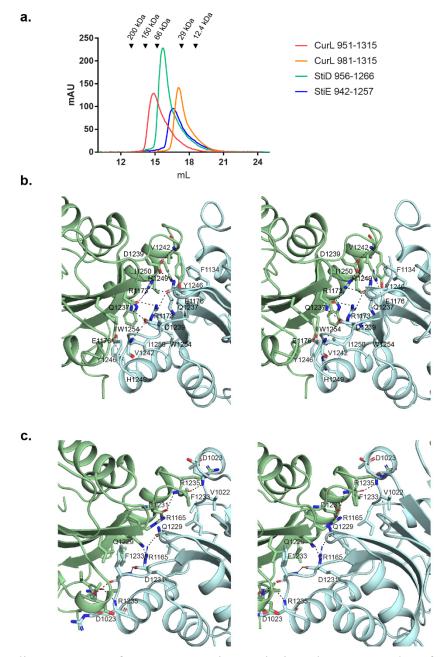


Figure S3. Oligomer state of *O*-MTs. (a) Size exclusion chromatography of *O*-MT fragments. StiD *O*-MT (amino acids 956-1266, 35 kDa monomer) elutes with an apparent molcular weight of 55 kDa and StiE *O*-MT (amino acids 942-1257, 37 kDa monomer) elutes predominantly as a monomer (apparent molecular weight 35.7 kDa). CurL *O*-MT (amino acids 981-1315, 39 kDa monomer) is exlusively monomeric (apparent molecular weight 26.9 kDa). Inclusion of a 30 residue post-AT dimerization element at the CurL N-terminus (amino acids 951-1315, 42 kDa monomer) results in dimeric protein (apparent molecular weight 82.8 kDa). (b) The StiD *O*-MT dimer interface is mediated by amino acids Phe1134, Arg1173, Val1175, Glu1176, Gln1237, Asp1239, Val1242, Tyr1246, His1249, Ile1250, Trp1254 from each monomer. (c) StiE *O*-MT dimer interface is mediated by amino acids Ala1021, Asp1023, Ala1026, Gly1124, Leu1126, Arg1165, Val1167, Gln1229, Asp1231, Phe1233, Arg1235.

	10	20	30	40	50	60	70	80
CurL/935-1005 WIDTDK								
	NKH KEI							
BarE/1598-1676 W I <mark>E</mark> K Y <mark>E</mark>								
	SQS QK/							
CurG/900-967 WIERSQ	GRGE DK	AKQNLSVS <mark>T</mark> S <mark>S</mark>	S <mark>PVT</mark> ELLDR <mark>GD</mark>	YKQL TAML TO	NGSLT	AV <mark>E</mark> VV <mark>Q</mark> QL	IHH <mark>HQQ</mark> S	- L AQAA
	NGY QSI							
	NQQQVVCS <mark>G</mark> E <mark>P</mark> I							
JamK/973-1045 <mark>₩VΕΤ</mark> ΚV								
JamM/892-969 WIETKE	DQ I MI	NILSLENQKE	YLVKLLNK <mark>g</mark> e	IQGLIKHIDK	(KVELS-EK <mark>ES</mark>	SLL <mark>PD</mark> LLKKI	MD I HQQERNP	PKEEISLL

Figure S4. Sequence alignment of cyanobacterial post-AT dimerization elements. Pathway abbreviations (GenBank accesion codes) are as follows: Cur-curacin A (CurL, AEE88278.1; CurG, AEE88283.1; CurI, AEE88281.1; CurM, AAT70108.1), Apr- apratoxin A (AprB, WP_075900458), Bar- barbamide (BarE, AEE88299), Crp- cryptophycin (CrpB, ABM21570.1), Jam- jamaicamide (JamK, AAS98782.1; JamM, AAS98784.1).

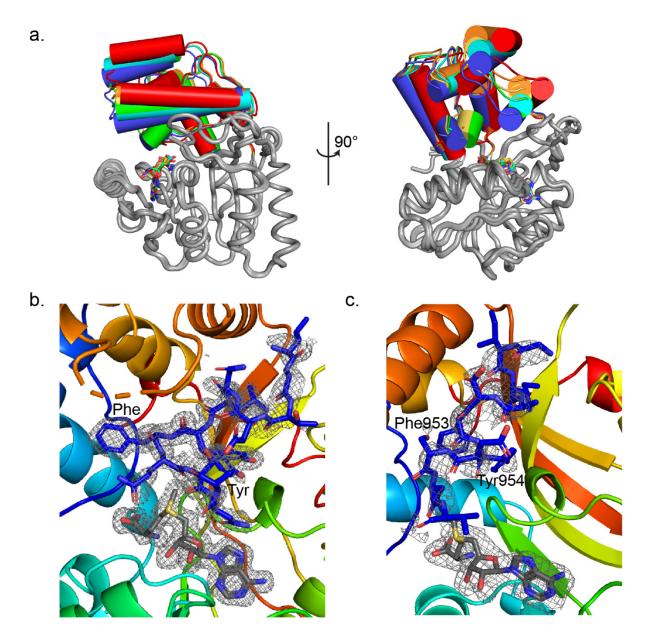
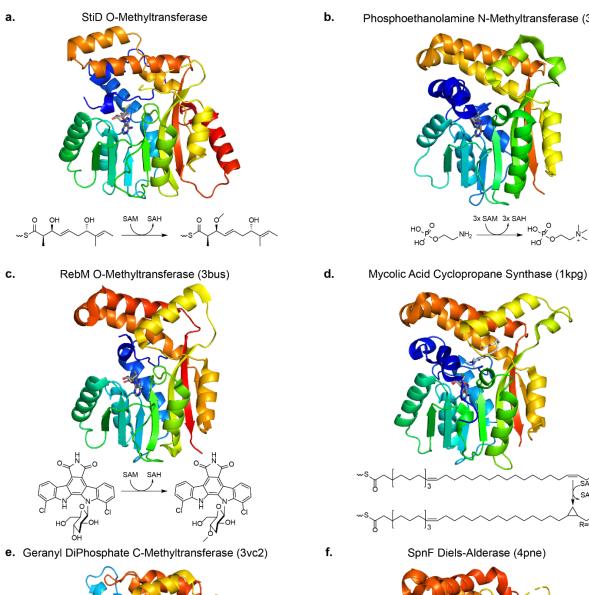
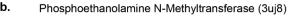


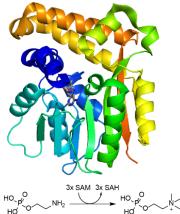
Figure S5 (a) StiD structures aligned at core (123 C α atoms, gray tube). Regions of the lid differ in position by up to 8 Å. **(b)** StiE Fo-Fc omit density for the SAM cofactor and TEV protease recognition sequence (StiE 961-1257, 1.42 Å, 2.5 σ contour). Phe and Tyr in the TEV protease recognition sequence are labeled. **(c)** SAM and the partially ordered N-terminal helix (StiE 942-1257, 1.90 Å, Fo-Fc omit density contoured at 2.5 σ). Conserved Phe953 and Tyr954 in the N-terminal helix are labeled.



о о но 6н 6н

но-^Ч





SAN SAH

R=C18

Figure S6. PKS O-MT homologs. Structures are colored as a rainbow from N- (dark blue) to Cterminus (red). The reaction catalyzed by each enzyme is shown below the structure. All structures have a lid composed of N-terminal helices (darkest blue) and helices between β-strands 6 and 7 (yellow and orange). SAH or SAM is rendered in stick form with gray C atoms.

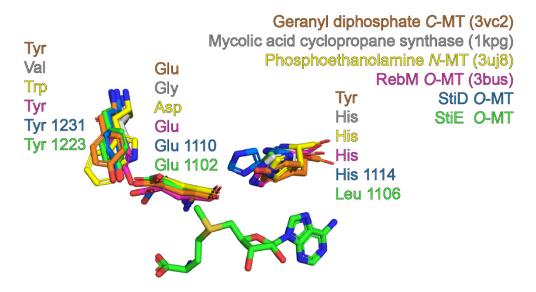


Figure S7. Conservation of key active site residues in structures of StiD and StiE *O*-MTs and homologs.

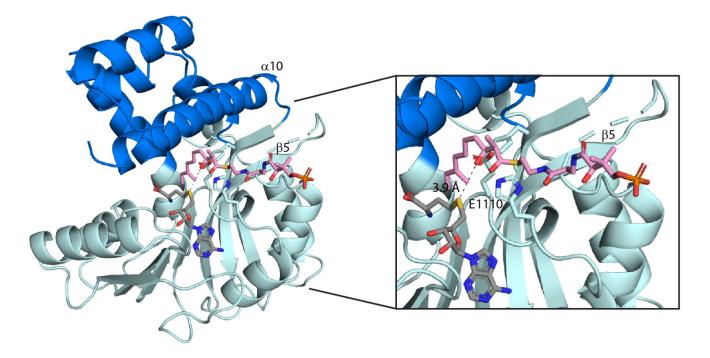


Figure S8. StiD substrate modeling. The full-length StiD substrate (pink sticks with atomic coloring) was modeled with the β -hydroxy oxygen in line with SAH and within hydrogen bonding distance of Glu1110. The ACP may interact with the nine-residue partially ordered loop (StiD 1146-1155). A similar ACP binding site was proposed for the PKS *O*-MT homolog mycolic acid cyclopropane synthase²⁷.

	N-950	N-956	N-976
C-1245	insoluble	insoluble	insoluble
C-1257	insoluble	insoluble	insoluble
C-1266	soluble	soluble	soluble

 Table S1. Protein stability of StiD fragments containing the O-MT.

Protein	SeMet StiD 976-1266	StiD 976-1266	StiD 956-1266	StiE 961-1257	StiE 942-1257	
Ligand	SAH	SAH	SAH	SAM	SAM	
Data Collection	_	-	-	-		
Space group	P1	<i>P</i> 1	P4 ₁	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	
Unit cell, a,b,c (Å)	38.3, 52.4, 68.7	40.2, 56.2, 72.3	71.1, 71.1, 123.0	90.9, 90.9, 84.9	88.6, 88.5, 86.1	
α,β,γ (°)	89.6, 88.3, 81.1	86.3, 84.9, 75.8	90, 90, 90	90, 90, 90	90, 90, 90	
X-ray source	APS 23ID-D	APS 23ID-D	APS 23ID-B	APS 23ID-D	APS 23ID-D	
Wavelength (Å)	0.979	1.033	1.033	1.033	1.033	
	1.96	1.80	1.70	1.42	1.90	
d _{min} (Å)	(2.03-1.96)	(1.86-1.80)	(1.76-1.70)	(1.47-1.42)	(1.99-1.90)	
R-merge	0.066 (0.82)	0.073 (0.74)	0.084 (2.34)	0.064 (1.03)	0.068 (2.28)	
Avg I/σ(I)	8.7 (1.1)	9.9 (1.6)	19.8 (1.1)	29.7 (2.7)	21.2 (1.1)	
Completeness (%)	96.7 (87.0)	96.7 (94.0)	99.7 (99.9)	99.2 (93.7)	100 (99.7)	
Multiplicity	3.5 (2.7)	3.5 (3.5)	13.8 (13.8)	24.3 (19.9)	13.0 (13.1)	
	127,763	193,662	931,727	1,638,916	358,381	
Total observations	(8,823)	(18,628)	(92,530)	(124,591)	(35,303)	
Wilson B factor (Å ²)	38.2	27.9	31.3	18.2	41.9	
CC _{1/2}	0.998 (0.559)	0.997 (0.597)	1.00 (0.442)	1.00 (0.803)	1.00 (0.482)	
CC*	1.00 (0.847)	0.999 (0.865)	1.00 (0.783)	1.00 (0.944)	1.00 (0.806)	
Refinement					L	
Data range (Å)	41.3-1.96	38.8-1.80	46.5-1.70	40.6-1.42	44.3-1.90	
Reflections	36,668	54,837	67,382	67,269	27,550	
R _{work} /R _{free} (%)	17.9/22.7	16.8/19.4	16.8/21.2	16.7/18.5	18.3/22.6	
Non-hydrogen atoms (#)	4,629	4,995	4,870	2,768	2,480	
protein	4,461	4,628	4,506	2,462	2,311	
ligands	52	52	52	27	27	
water	116	314	312	315	118	
Amino acid residues	554	564	561	279	282	
Deviation from ideality						
bond lengths (Å)	0.004	0.007	0.014	0.005	0.007	
bond angles (°)	0.99	1.15	1.24	0.88	0.84	
Average B-factor (Å ²)	54.6	42.4	48.0	30.5	69.9	
protein	54.7	42.1	47.6	29.5	70.3	
ligands	54.3	53.9	74.2	18.7	58.1	
solvent	51.0	45.7	50.1	39.2	60.9	
Ramachandran plot						
favored (%)	97.6	97.7	97.8	97.8	97.1	
allowed (%)	2.4	2.3	2	2.2	2.9	
outliers (%)	0	0	0.2	0	0	
PDB ID	6ECU	6ECV	6ECW	6ECT	6ECX	

Table S2. Crystallographic Data

¹values in parentheses designate outer shell

Table S3. Primers

StiD 976 F	TACTTCCAATCCAATGCCTCAGCCGTGGATGAAAGC
StiD 956 F	TACTTCCAATCCAATGCCGAACATCCTGTTGACGGC
StiD 950 F	TACTTCCAATCCAATGCCGCGAACGGGCAGGGTAAT
StiD 1245 R	TTATCCACTTCCAATGCTAGGTACTACGAACGTGAGAATC
StiD 1257 R	TTATCCACTTCCAATGCTAGGCCTCAACCCATTTCTG
StiD 1266 R	TTATCCACTTCCAATGCTACACTCAGTTCCCGTGCCGC
StiE 961 F	TACTTCCAATCCAATGCCGCGGCGGGGGGAAGACG
StiE 951 F	TACTTCCAATCCAATGCCGCTTCGTTCTACGATAGCCTG
StiE 942 F	
StiE 1257 R	TTATCCACTTCCAATGCCGCAAGCAAGCAACGCACGCGCGCC
StiD 1794 F	TACTTCCAATCCAATGCCGCACTGGCCGCCTTAGG
StiD 1929 R	
StiE 1789 F	
StiE 1927 R	
CurL 951 F	
CurL 981 F	
CurL 1315 R	TTATCCACTTCCAATGCTAAGCTACTTCAGAGTAAGAAGA
StiE Y954F F	AGGCATCATTGCTTCGTTCTTCGATAGCCTGGTG
StiE Y954F R	
StiE E1102A F	GATTTAGTGCTCGGATTTGCGGTGGCCGGACTTAT
StiE E1102A R	ATAAGTCCGGCCACCGCAAATCCGAGCACTAAATC
StiE E1102Q F	TGTATGATTTAGTGCTCGGATTTCAGGTGGCCGGAC
StiE E1102Q R	GTCCGGCCACCTGAAATCCGAGCACTAAATCATACA
StiE L1106H F	TGAGGTGGCCGGACATATCCCTGACAAGG
StiE L1106H R	CCTTGTCAGGGATATGTCCGGCCACCTCA
StiE Y1209F F	CAACGTTCCTTTGGCAGCTTTGAGAATGTGTACAAAG
StiE Y1209F R	CTTTGTACACATTCTCAAAGCTGCCAAAGGAACGTTG
StiE Y1223F F	CGGGGGCCTGATCTCCTTTGTACTGTTTCATG
StiE Y1223F R	CATGAAACAGTACAAAGGAGATCAGGCCCCCG
CurL E1161A F	CAACCTGGCATTTGGATTTGCAGTAGCTCATCATATTAAGG
CurL E1161A R	CCTTAATATGATGAGCTACTGCAAATCCAAATGCCAGGTTG
CurL E1161Q F	GATAATTACAACCTGGCATTTGGATTTCAGGTAGCTCATCATATTAAGGAT
CurL E1161Q R	ATCCTTAATATGATGAGCTACCTGAAATCCAAATGCCAGGTTGTAATTATC
CurL Y1281F F	TGCTGTTAGTAATACAAAGCTAGCCAATCCTTTGCTCAG
CurL Y1281 R	CTGAGCAAAGGATTGGCTAGCTTTGTATTACTAACAGCA
CurL Y1010F F	ATTTAAAGGTAATGTAGTTTATGACTATTTCAATTCTTTTGCAGAAATTAGTCAAGAAA
CurL Y1010F R	ТТТСТТБАСТААТТТСТБСААААБААТТБАААТАБТСАТАААСТАСАТТАССТТТАААТ
CurL H1165A F	GGCATTTGGATTTGAAGTAGCTCATGCTATTAAGGATAAATCGCTGTTATTT
CurL H1165A R	AAATAACAGCGATTTATCCTTAATAGCATGAGCTACTTCAAATCCAAATGCC
CurL H1165N F	GCATTTGGATTTGAAGTAGCTCATAATATTAAGGATAAATCGCTGTTAT
CurL H1165N R	ATAACAGCGATTTATCCTTAATATTATGAGCTACTTCAAATCCAAATGC
CurL Y1267F F	ATGTTAAGTCAGCTTTTCAATCCTTTAATCAGTTAGGTAAATTACTGAG
CurL Y1267F R	CTCAGTAATTTACCTAACTGATTAAAGGATTGAAAAGCTGACTTAACAT
	tes handles for ligation-independent cloning into expression vectors

Bold text indicates handles for ligation-independent cloning into expression vectors.

SUPPLEMENTARY REFERENCES

- Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G., and Muller, R. The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase, *J Biol Chem* 277, 13082-13090 (2002).
- Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., Sherman, D. H., and Gerwick, W. H. Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*, *J Nat Prod* 67, 1356-1367 (2004).
- 3. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site, *Protein Expr Purif 25*, 8-15 (2002).
- 4. DelProposto, J., Majmudar, C. Y., Smith, J. L., and Brown, W. C. Mocr: a novel fusion tag for enhancing solubility that is compatible with structural biology applications, *Protein Expr Purif* 63, 40-49 (2009).
- Skiba, M. A., Maloney, F. P., Dan, Q., Fraley, A. E., Aldrich, C. C., Smith, J. L., and Brown, W. C. PKS-NRPS enzymology and structural biology: considerations in protein production, *Methods Enzymol 604*, 45-88 (2018).
- Whicher, J. R., Smaga, S. S., Hansen, D. A., Brown, W. C., Gerwick, W. H., Sherman, D. H., and Smith, J. L. Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis, *Chem Biol 20*, 1340-1351 (2013).
- Sánchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. Cloning and characterization of a phosphopantetheinyl transferase from *Streptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin, *Chem Biol* 8, 725-738 (2001).
- 8. Evans, D. A., Bartroli, J., and Shih, T. L. Enantioselective aldol condensations. 2. erythroselective chiral aldol condensations via boron enolates., *Journal of the American Chemical Society 103*, 2127-2129 (1981).
- Fiers, W. D., Dodge, G. J., Sherman, D. H., Smith, J. L., and Aldrich, C. C. Vinylogous dehydration by a polyketide dehydratase domain in curacin biosynthesis, *J Am Chem Soc* 138, 16024-16036 (2016).
- 10. May, A. E., Connell, N. T., Dahlmann, H. A., and Hoye, T. R. A useful modification of the evans magnesium halide-catalyzed anti-aldol reaction: application to enolizable aldehydes, *Synlett 13*, 1984-1986 (2010).
- 11. Gu, L., Wang, B., Kulkarni, A., Gehret, J. J., Lloyd, K. R., Gerwick, L., Gerwick, W. H., Wipf, P., Hakansson, K., Smith, J. L., and Sherman, D. H. Polyketide decarboxylative chain termination preceded by *O*-sulfonation in curacin a biosynthesis, *J Am Chem Soc* 131, 16033-16035 (2009).
- 12. McCarthy, J. G., Eisman, E. B., Kulkarni, S., Gerwick, L., Gerwick, W. H., Wipf, P., Sherman, D. H., and Smith, J. L. Structural basis of functional group activation by sulfotransferases in complex metabolic pathways, *ACS Chem Biol* 7, 1994-2003 (2012).
- Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., Kolter, R., Walsh, C. T., and Kelleher, N. L. Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry, *Biochemistry* 45, 12756-12766 (2006).

- 14. Meluzzi, D., Zheng, W. H., Hensler, M., Nizet, V., and Dorrestein, P. C. Top-down mass spectrometry on low-resolution instruments: characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways, *Bioorg Med Chem Lett 18*, 3107-3111 (2008).
- 15. Kabsch, W. Xds, Acta Crystallogr D Biol Crystallogr 66, 125-132 (2010).
- 16. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C., and Adams, P. D. Automated structure solution with the PHENIX suite, *Methods Mol Biol 426*, 419-435 (2008).
- 17. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221 (2010).
- Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, *Acta Crystallogr D Biol Crystallogr 64*, 61-69 (2008).
- 19. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. Phaser crystallographic software, *J Appl Crystallogr* 40, 658-674 (2007).
- 20. Emsley, P., and Cowtan, K. Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132 (2004).
- 21. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr D Biol Crystallogr 68*, 352-367 (2012).
- 22. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21 (2010).
- 23. Holm, L., and Rosenstrom, P. Dali server: conservation mapping in 3D, *Nucleic Acids Res* 38, W545-549 (2010).
- 24. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. Clustal W and Clustal X version 2.0, *Bioinformatics 23*, 2947-2948 (2007).
- 25. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench, *Bioinformatics* 25, 1189-1191 (2009).
- 26. Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
- 27. Huang, C. C., Smith, C. V., Glickman, M. S., Jacobs, W. R., Jr., and Sacchettini, J. C. Crystal structures of mycolic acid cyclopropane synthases from *Mycobacterium tuberculosis*, J *Biol Chem* 277, 11559-11569 (2002).