View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Caltech Authors - Main

Supporting Information

Replacement of ProB28 by Pipecolic Acid Protects Insulin Against Fibrillation and Slows Hexamer Dissociation

Katharine Y. Fang*, Seth A. Lieblich*, and David A. Tirrell

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125
Correspondence to: D. A. Tirrell (E-mail: tirrell@caltech.edu)

*These authors contributed equally to this work.

Table of Contents

Experimental Procedures

Materials

All canonical amino acids and (4*S*)-1,3-thiazolidine-4-carboxylic acid (Thz) were purchased from Sigma. (*S*)-azetidine-2-carboxylic acid (Aze) and (3,4)-dehydro-L-proline (Dhp) were purchased from Bachem Americas. (*S*)-piperidine-2-carboxylic acid (Pip) was purchased from TCI America under the chemical name L-pipecolic acid. All solutions and buffers were made using double-distilled water (ddH_2O).

Strains and plasmids

A gene encoding proinsulin (PI) bearing an *N*-terminal hexahistidine tag (6xHis) and flanked by *Eco*R1 and *Bam*H1 restriction sites was ordered as a gBlock from Integrated DNA Technologies. The gBlock and plasmid vector pQE80L (which allows IPTG-inducible gene expression) were digested with *Eco*RI and *Bam*HI. Linearized pQE80L was dephosphorylated by treatment with alkaline phosphatase (NEB). Ligation of the digested PI gene and linearized vector yielded plasmid pQE80PI (which was used to produce ProI). To make plasmid pQE80PI-proS, genomic DNA was extracted from *E. coli* strain DH10B using DNeasy Blood and Tissue Kit (Qiagen). Primers (Integrated DNA Technologies) were designed to amplify the *E. coli proS* gene, encoding prolyl-tRNA synthetase, under constitutive control of its endogenous promoter, from purified genomic DNA, and to append *Nhe*I and *Nco*I sites. The digested *proS* gene was then inserted into pQE80PI between transcription termination sites by ligation at *Nhe*I and *Nco*I restriction sites. Site-directed mutagenesis of pQE80PI-proS at the site corresponding to a cysteine at position 443 was performed to make plasmid pQE80PI-proS(C443G), which differs from pQE80PI-proS by one nucleotide that specifies a single amino acid mutation to glycine. Prototrophic *E. coli* strain JT31 and proline-

auxotrophic *E. coli* strain CAG18515 were obtained from the Coli Genetic Stock Center at Yale University. JT31, containing a deletion for proline dehydrogenase (which is known to degrade Dhp and Thz),[1-2] was made proline-auxotrophic by deleting the gene *proA* using λ Red-mediated recombination[3]; the new strain was designated KS32. CAG18515 was used for the expression of AzeI and PipI, while KS32 was used for the expression of DhpI and ThzI. Prototrophic *E. coli* strain BL21 was used for expression of canonical insulins (ProI, AspI) in rich media. Site-directed mutagenesis of pQE80PI at the site corresponding to insulin B28 was performed to make plasmid pQE80PI-asp, which differs from pQE80PI by three nucleotides that specify a single amino acid mutation to aspartic acid. All genes and plasmids were confirmed by DNA sequencing.

Protein expression

Plasmids pQE80PI and pQE80PI-asp were transformed into BL21 cells and grown on ampicillin-selective agar plates. A single colony was used to inoculate 5 mL of Luria-Bertani (LB) medium and grown overnight; the resulting saturated culture was used to inoculate another 1 L of LB medium. All expression experiments were conducted at 37°C, 200 RPM in shake flasks (Fernbach 2.8 L flasks, Pyrex®). Each culture was induced with 1 mM IPTG at mid-exponential phase (OD_{600} ~0.8). For incorporation of the proline analogs (Aze, Dhp, Thz, and Pip), pQE80PI-proS was transformed into both cell strains (CAG18515 and KS32) and pQE80-proS (C443G) was transformed into the CAG18515 cell strain. All cells were grown on ampicillin-selective agar plates. A single colony was used to inoculate 25 mL of LB medium and the culture was grown overnight prior to dilution into 1 L of 1X M9, 20 aa medium (8.5 mM NaCl, 18.7 mM NH_4Cl, 22 mM KH_2PO_4, 47.8 mM Na_2HPO_4, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 3 mg/L $FeSO_4$, 1 µg/L of trace metals (Cu^{2+}, Mn^{2+}, Zn^{2+}, MoO_4^{2-}), 35 mg/L thiamine hydrochloride, 10 mg/L biotin, 20 mM D-glucose, 200 mg/L ampicillin with 50 mg/L of L-amino acids, each). At a cell density corresponding to OD_{600} ~0.8, the culture was subjected to a medium shift; briefly, cells were centrifuged and washed with saline prior to resuspension in 0.8 L of 1.25X M9, 19 aa (1X M9, 20 aa medium without L-proline). After cells were further incubated for 30 min to deplete intracellular proline, 200 mL of 5X additives (1.5 M NaCl, 2.5 mM Aze, Dhp, Thz, or Pip) was added to the culture. After another 15 min of incubation at 37°C to allow amino acid uptake prior to induction, IPTG was added to a final concentration of 1 mM. After 2 h, cells were harvested by centrifugation and stored at -80°C until further use.

Cell lysis and refolding from inclusion bodies

Cells were thawed on the benchtop for 15 min prior to resuspension in lysis buffer (B-PER®, 0.5 mg/mL lysozyme, 50 U/mL benzonase nuclease). Cells were gently agitated at RT for 1 h prior to centrifugation (10 000 g, 10 min, RT); supernatant was discarded, and the pellet was washed thrice: once with wash buffer (2 M urea, 20 mM Tris, 1% Triton X-100, pH 8.0) and twice with sterile ddH_2O; centrifugation followed each wash and the supernatant was discarded. The final washed pellet containing inclusion bodies (IBs, ~50% PI) was re-suspended in Ni-NTA binding buffer (8 M urea, 300 mM NaCl, 50 mM NaH_2PO_4, pH 8.0) overnight at 4°C or at RT for 2 h, both with gentle agitation. The suspension was centrifuged to remove insoluble debris, the remaining pellet was discarded, and the supernatant was mixed with pre-equilibrated Ni-NTA resin (Qiagen) at RT for 1 h to purify PI from the IB fraction. Unbound proteins in the IB fraction were collected in the flow-through (FT), and the resin was washed with Ni-NTA

wash (W) buffer (8 M urea, 20 mM Tris base, 5 mM imidazole, pH 8.0) and Ni-NTA rinse buffer (8 M urea, 20 mM Tris base, pH 8.0) to remove any remaining non-specific proteins still bound to the resin. Finally, PI was stripped from the resin with Ni-NTA elution buffer (8 M urea, 20 mM Tris base, pH 3.0). Fractions (IBs, FT, W, elution) were collected and run under reducing conditions on SDS-PAGE (Bis/Tris gels, Novex®); elution fractions containing PI were pooled and solution pH was adjusted to 9.6 with 6 N NaOH in preparation for oxidative sulfitolysis. Oxidative sulfitolysis was performed at RT for 4 h, with the addition of sodium sulfite and sodium tetrathionate (0.2 M Na_2SO_3, 0.02 M $Na_2S_4O_6$); the reaction was quenched by 10-fold dilution with ddH_2O. To isolate PI from the quenched solution, the pH was adjusted to between 3.5 and 4.5 by adding 6 N HCl dropwise; the solution became cloudy. The solution was centrifuged (10 000 g, 10 min, RT) and supernatant discarded. The PI pellet was then re-suspended in refolding buffer (0.3 M urea, 50 mM glycine, pH 10.6) and protein concentration was estimated by the bicinchoninic acid assay (BCA assay, Pierce®). The concentration of PI was adjusted to 0.5 mg/mL. Refolding was initiated by addition of β-mercaptoethanol to a final concentration of 0.5 mM and allowed to proceed at 12°C overnight with gentle agitation (New Brunswick® shaker, 100 RPM). Post-refolding, soluble PI was harvested by adjusting the pH of the solution to 4-5 by dropwise addition of 6 N HCl and by high speed centrifugation to remove insoluble proteins. The supernatant was adjusted to pH 8–8.5 by dropwise addition of 6 N NaOH and dialyzed against fresh PI dialysis buffer (7.5 mM sodium phosphate buffer, pH 8.0) at 4°C with five buffer changes to remove urea. The retentate (PI in dialysis buffer) was then lyophilized and subsequently stored at -80°C until further processing. Typical yields were 20–50 mg PI per L of culture (20–30 mg/L for non-canonical PI, 40–50 mg/L for canonical PI expression in rich media).

Proteolysis and chromatographic (HPLC) purification

The dry PI powder was re-dissolved in water to a final concentration of 5 mg/mL PI (final concentration of sodium phosphate buffer was 100 mM, pH 8.0). Trypsin (Sigma-Aldrich) and carboxypeptidase-B (Worthington Biochemical) were added to final concentrations of 20 U/mL and 10 U/mL, respectively, to initiate proteolytic cleavage. The PI/protease solution was incubated at 37°C for 2.5 h; proteolysis was quenched by addition of 0.1% trifluoroacetic acid (TFA) and dilute HCl to adjust the pH to 4. Matured insulin was purified by reversed phase high-performance liquid chromatography (HPLC) on a C_{18} column using a gradient mobile phase of 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (ACN; solvent B). Elution was carried from 0% B to 39% B with a gradient of 0.25% B per min during peak elution. Fractions were collected and lyophilized, and the dry powder was dissolved in 10 mM sodium phosphate, pH 8.0. Insulin-containing fractions were verified by matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS; Voyager MALDI-TOF, Applied Biosystems) and SDS-PAGE to ensure identity and purity. Typical yields were 5-20 mg insulin per 100 mg PI. Fractions were stored at -80°C in 10 mM phosphate buffer, pH 8.0 until further use.

Verification of Aze, Dhp, Thz, and Pip incorporation levels and maturation

A 30 µL aliquot of PI solution (8 M urea, 20 mM Tris, pH 8) was subjected to cysteine reduction and alkylation (5 mM DTT, 55°C, 20 min; 15 mM iodoacetamide, RT, 15 min, dark) prior to 10-fold dilution into 100 mM NH_4HCO_3, pH 8.0 (100 µL final volume). Peptide digestion was initiated with 0.6 µL of gluC stock solution (reconstituted at 0.5 µg/µL with ddH_2O, Promega) at 37°C for 2.5 h. The reaction was quenched

by addition of 10 µL 5% TFA and immediately subjected to C_{18} ZipTip (Millipore) peptide purification and desalting according to the manufacturer's protocol. Peptides were eluted in 50% ACN, 0.1% TFA; the eluent was then diluted 3-fold into matrix solution (saturated α-cyanohydroxycinnamic acid in 50% ACN, 0.1% TFA) and analyzed by mass spectrometry (Voyager MALDI-TOF, Applied Biosystems). Incorporation levels were analyzed prior to and after refolding; incorporation percentage was calculated by comparing total AUC (area under the curve, arbitrary units) of the non-canonical peak (expected m/z: 1543 Da for the proinsulin peptide containing AzeB28 [-14 Da], 1555 Da for DhpB28 [-2 Da], 1573 for ThzB28 [+16 Da], or 1571 for PipB28 [+14 Da]; 5794 Da for intact AzeI, 5806 Da for DhpI, 5824 for ThzI, or 5822 Da for PipI) with total AUC of its canonical counterpart (1557 Da and 5808 Da for ProI, respectively). Maturation of AzeI, DhpI, ThzI, and PipI was analyzed after HPLC purification. TFA (1.6 µL, 5%) was added to 15 µL mature insulin solution (10 mM phosphate buffer pH 8.0) and subjected to C_{18} ZipTip (Millipore) peptide purification and desalting per the manufacturer's protocol.

Circular dichroism

Spectra were collected on samples contained in 1-cm quartz cuvettes in 10 mM sodium phosphate buffer pH 8.0. Data were collected from 195 nm to 250 nm, with step size of 0.25 nm and averaging time of 1 s on a Model 410 Aviv Circular Dichroism Spectrophotometer; spectra were averaged over 3 repeat scans. A reference buffer spectrum was subtracted from the sample spectra for conversion to mean residue ellipticity. The insulin concentration was 60 µM.

Hexamer dissociation assay

Insulins were quantified by both UV absorbance (NanoDrop Lite, ThermoFisher) and BCA assay, and normalized to 125 µM insulin prior to dialysis against 50 mM Tris/perchlorate, 25 µM zinc sulfate, pH 8.0 overnight at 4°C using a D-tube dialyzer (Millipore Corp.) with MWCO of 3.5 kDa. Aliquots of dialyzed insulin solution were mixed with phenol to yield samples of the following composition: 100 µM insulin, 20 µM zinc sulfate, 100 mM phenol. Dissociation was initiated by the addition of terpyridine (terpy; Sigma-Aldrich) to a final concentration of 0.3 mM from a 0.75 mM stock solution. A Varioskan multimode plate reader (ThermoScientific) was used to detect the appearance of the Zn^{2+}-(terpy) complex by monitoring absorbance at 334 nm. The data were fitted to a mono-exponential function using Origin software (OriginLab). The function was of the form: $A_{334}(t) = A_{334}(t = \infty) - Ae^{-(t/\tau_{½})}$, where A is a fitting constant, $A_{334}(t)$ is the absorbance at 334 nm at time t and $A_{334}(t = \infty)$ is the long-time, asymptotic value of $A_{334}(t)$. The value of $A_{334}(t = \infty) - A_{334}(t)$ is proportional to the concentration of insulin hexamer at time t. Kinetic runs were done at least in triplicate. Post-assay insulin samples were pooled, and sample quality was determined by SDS-PAGE.

Fibrillation assay

Insulin samples (60 µM in 10 mM phosphate, pH 8.0) were centrifuged at 22 000 g for 1 h immediately after addition of thioflavin T (ThT) (EMD Millipore) to a final concentration of 1 µM. Samples were continuously shaken at 960 RPM on a Varioskan multimode plate reader at 37°C, and fluorescence readings were recorded every 15 min or 20 min for 48 h (excitation 444 nm, emission 485 nm). Assays

were run in quadruplicate, in volumes of 200 µL in sealed (Perkin-Elmer), black, clear-bottom 96 well plates (Grenier BioOne).

Crystallographic Studies

Insulin crystals were obtained from sitting drop trays set using a Mosquito robot (TTP Labtech). Drops were set by mixing 0.4 µL insulin solution with 0.4 µL well solution. Cells were cryoprotected in a mother liquor containing 30% glycerol prior to looping and flash freezing in liquid nitrogen. Data were collected at SSRL beamline BL12-2 using a DECTRIS PILATUS 6M pixel detector. Initial indexing and scaling was performed with XDS; for some structures, data were re-scaled in alternative space groups using Aimless[4]. Initial phases were generated by molecular replacement in PHASER with 3T2A (5USV, 5URT, 5UU2, and 5USP) or 1EV3 (5URU, 5UU4, and 5USS).[5] Structure refinement was carried out in Coot and Refmac.[6-7] Data were deposited in the PDB with the following codes: 5USV (T_2-AzeI), 5URT (T_2-DhpI), 5UU2 (T_2-ThzI), 5USP (T_2-PipI), 5URU (R_6-DhpI), 5UU4 (R_6-ThzI), and 5USS (R_6-PipI). All distances and contacts were computed using UCSF Chimera Crystallography Software.[8]

Results and Discussion

Supplemental figures

FIGURE S1 Insulin expression and incorporation of proline analogs. (A-D) SDS-PAGE of cell lysates with lanes labelled for pre-induction (PRE) and post-induction in 19-amino acid minimal media supplemented with either nothing (19aa), Aze (A), Dhp (B), Thz (C), or Pip (D), or Pro. (E-I) MALDI-MS traces of isolated proinsulin peptide fragment ^{46}RGFFYT**P**KTRRE^{57} obtained by gluC digestion. Peptide fragment masses correspond to either wild type mass (1558 Da, E) or shifted masses (1544 Da for AzeI, F; 1555 Da for DhpI, G; 1575 Da for ThzI, H; or 1571 Da for PipI, I). Inset is whole protein MALDI-MS (instrument error ~10 Da) with observed masses: 5799 Da (ProI, E), 5789 Da (AzeI, F), 5800 Da (DhpI, G), 5823 Da (ThzI, H), and 5812 Da (PipI, I). Ion counts >10^3 for all MALDI-MS spectra.

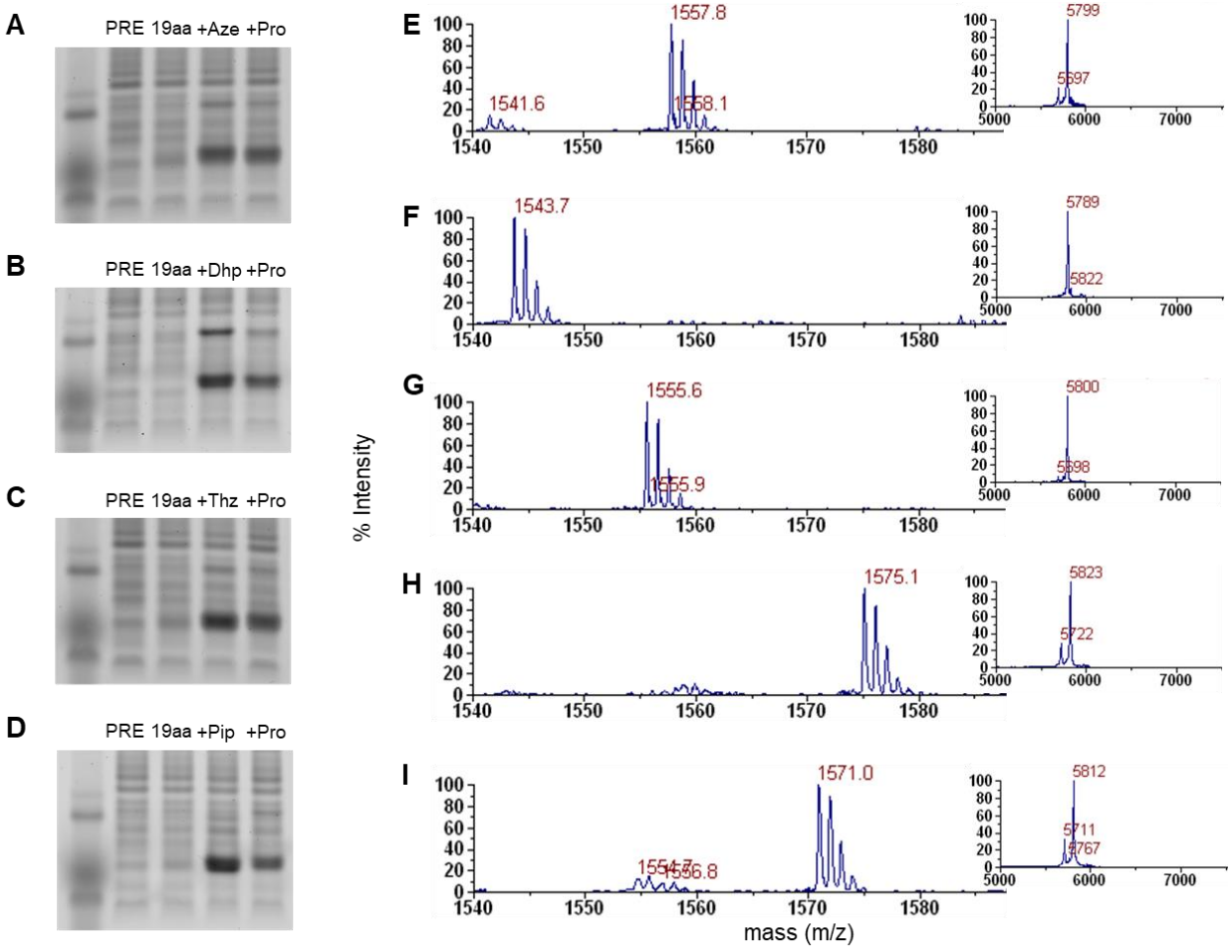
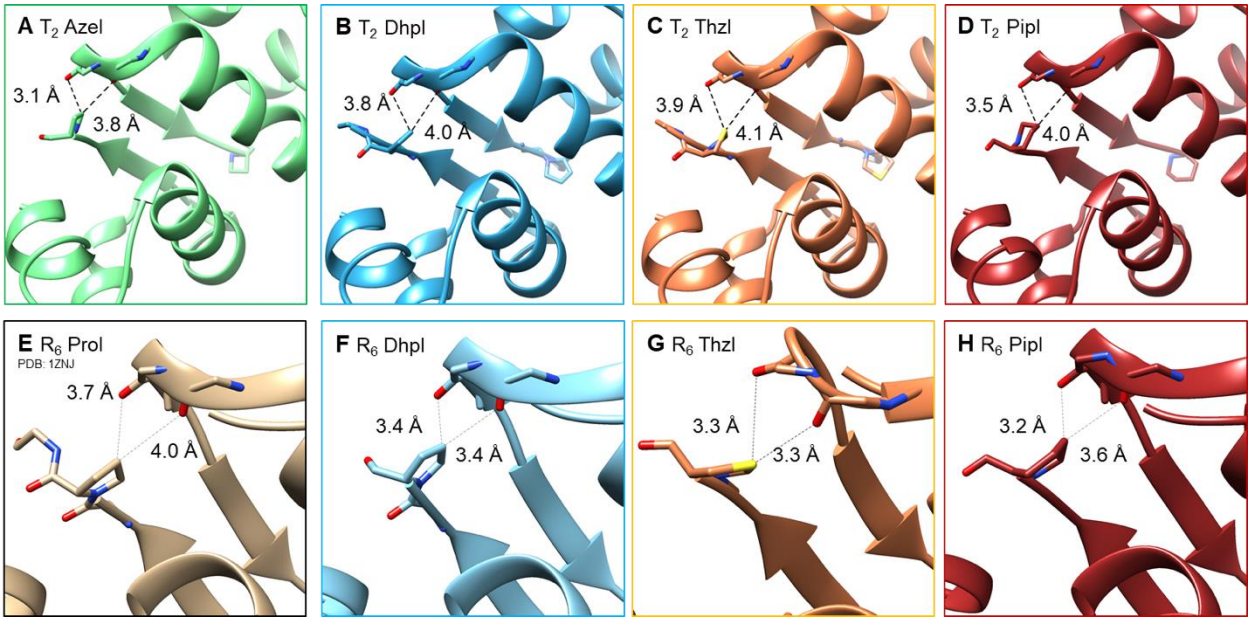


FIGURE S2 Crystal structures of insulin variants showing the insulin dimer interface. (A-D) Insulin variants in the T_2 dimer form. (E-H) Structures of insulins in the R_6 hexamer form. Human insulin (PDB Code: 1ZNJ) in panel E for reference. All images highlight the distance between the γ-carbon of the ring at position B28 and its closest neighbors, the backbone carbonyl oxygen atoms of GlyB20' and GluB21', across the dimer interface.



Supplemental tables

TABLE S1 Expression Yields of Proinsulins and Incorporation Levels of Proline Analogs

B28 Amino Acid	Incorporation (%)[a]	Yield (mg/L)[b]
Proline (Pro)	--	50
(*S*)-azetidine-2-carboxylic acid (Aze)	≤100[c]	24
(3,4)-dehydro-L-proline (Dhp)	≤100[c]	23
(4*S*)-1,3-thiazolidine-4-carboxylic acid (Thz)	90	28
(*S*)-piperidine-2-carboxylic acid (Pip)	89	21

[a] Quantified by MALDI-MS from shake flask expressions: Proinsulin peptide obtained by gluC digestion: [46]RGFFYT**P**KTRRE[57].
[b] Yield quantified by BCA assay post-refolding at the proinsulin level.
[c] The canonical peptide was not detected or uniquely identified in MALDI-MS spectra.

TABLE S2 Data Tables and Refinement Values for the T_2 Structures

	AzeI	DhpI	ThzI	PipI
PDB Code	5USV	5URT	5UU2	5USP
Data collection				
Space group	$I\,2_1\,3$	$I\,2_1\,3$	$I\,2_1\,3$	$I\,2_1\,3$
Cell dimensions				
a, b, c (Å)	77.979, 77.979, 77.979	77.950, 77.950, 77.950	78.067, 78.067, 78.067	78.047, 78.047, 78.047
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	31.83-1.3 (1.33-1.3)	55.12-1.17 (1.19-1.17)	31.87-1.22 (1.25-1.22)	39.02-1.17 (1.19-1.17)
$R_{p.i.m.}$ (%)	3.7 (12.1)	2.1 (53.6)	1.6 (25.1)	2.0 (34.2)
Mn(I /sd)	13.2 (1.2)	18.1 (1.3)	23.5 (3.4)	17.8 (2.0)
Completeness (%)	99.4 (100)	99.7 (94.6)	99.6 (91.5)	99.75 (93.3)
Redundancy	18.4 (17.2)	18.6	8.1 (4.5)	10 (5.4)
Refinement				
No. reflections	26650 (1288)	26449(1328)	23424 (1064)	26485 (1198)
R_{work} / R_{free} (%)	14.2/15.9 (23.2/23.3)	11.9/14.1 (23.9/27.2)	13.7/15.5 (17.4/22.9)	13.9/15.26 (20.9/22.3)
No. atoms				
Protein	392	397	397	467
Ligand/ion	1	1	0	1
Water	58	67	68	71
RMSD				
Bond lengths (Å)	0.009	0.028	0.017	0.008
Bond angles (°)	1.978	2.273	1.53	1.07
Ramachandran map analysis				
Favored	50	50	50	50
Allowed	0	0	0	0
Outliers	0	0	0	0

TABLE S3 Data Tables and Refinement Values for the R_6 Structures

	DhpI	ThzI	PipI
PDB Code	5URU	5UU4	5USS
Data collection			
Space group	*H* 3	*H* 3	*H* 3
Cell dimensions			
a, b, c (Å)	79.259, 79.259, 79.874	78.09, 78.09, 39.93	79.024, 79.024, 39.509
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	34.52-2.41 (2.47-2.41)	39.05-1.96 (2.01-1.96)	25.87-2.06 (2.11-2.06)
$R_{p.i.m.}$ (%)	4.2 (30.8)	2.2 (14.5)	2.6 (34.2)
Mn(I /sd)	5.4 (2.1)	19.4 (0.5)	12.9 (3.2)
Completeness (%)	99.76 (99.6)	98.1 (76.3)	99.8 (99.8)
Redundancy	4.6	5.5 (3.6)	5.3 (5.0)
Refinement			
No. reflections	6816 (469)	6396 (350)	9265 (422)
R_{work} / R_{free} (%)	19.0/23.7 (22.2/28.7)	23.2/26.4 (33.0/36.3)	16.9/23.8 (22.9/30.7)
No. atoms			
Protein	1439	661	708
Ligand/ion	32	22	18
Water	7	0	1
RMSD			
Bond lengths (Å)	0.022	0.018	0.008
Bond angles (°)	2.04	1.800	1.047
Ramachandran map analysis			
Favored	194	93	97
Allowed	0	1	0
Outliers	0	0	0

References

1. J. M. Wood, *J. Bacteriol.* **1981**, *146*, 895.
2. C. E. Deutch, *FEMS Microbiol. Lett.* **2004**, *238*, 383.
3. K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 6640.
4. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2011**, *67*, 235.
5. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658.
6. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2010**, *66*, 486.
7. G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **1997**, *53*, 240.
8. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.