Supplementary Material

Crystal structure of indole-3-glycerol phosphate synthase from *Thermus thermophilus* HB8: implications for thermal stability.

Bagautdin Bagautdinov[a,b*], Katsuhide Yutani [b]

[a]Japan Synchrotron Radiation Research Institute (JASRI/SPring-8), 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan

[b]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

[*]Correspondence e-mail: bagautdi@spring8.or.jp

1. Materials and methods

1.1. Protein expression and purification

The *trp*C gene, encoding IGPS was amplified by the polymerase chain reaction (PCR) using *T. thermophilus* HB8 genomic DNA as the template. The PCR product was ligated with pT7blue (Novagen). The plasmid was digested with *Nde*I and *Bgl*II and the fragment was inserted into the expression vector pET-11a (Novagen) linearized with *Nde*I and *Bam*HI. *E. coli* BL21(DE3) cells were transformed with the recombinant plasmid and grown at 310 K in Luria-Bertani medium containing 50 µg ml-1 ampicillin for 20 h. The cells were harvested by centrifugation at 4500g for 5 min at 277 K and were subsequently

suspended in 20 mM Tris-HCl pH 8.0 containing 0.5 *M* NaCl and 5 m*M* 2-mercaptoethanol. Cells were disrupted by sonication and heated at 343 K for 12 min. The cell debris and denatured proteins were removed by centrifugation (14 000 rev min-1, 30 min). The supernatant solution was used as the crude extract for purification. The crude extract was desalted with a HiPrep 26/10 desalting column (Amersham Biosciences) and applied onto a Super Q Toyopearl 650M (Tosoh) column equilibrated with 20 m*M* Tris-HCl pH 8.0 (buffer *A*). After elution with a linear gradient of 0-0.3 *M* NaCl, the fraction containing *Tt*IGPS was desalted with a HiPrep 26/10 desalting column with buffer *A*. The sample was subjected to a Resource Q column (Amersham Biosciences) equilibrated with buffer *A*. After elution with a linear gradient of 0-0.4 *M* NaCl, the fraction containing *Tt*IGPS was desalted with a HiPrep 26/10 desalting column with 10 m*M* sodium phosphate pH 7.0. The sample was then applied onto a Bio-Scale CHT-20-I column (Bio-Rad) equilibrated with 10 m*M* sodium phosphate pH 7.0. After elution with a linear gradient of 10-100 m*M* sodium phosphate, the fraction containing *Tt*IGPS was subjected once more to a Resource Q column. The sample was concentrated by ultrafiltration (Vivaspin) and loaded onto a HiLoad 16/60 Superdex 75 prep-grade column (Amersham Biosciences) equilibrated with buffer *A* containing 0.2 *M* NaCl. The homogeneity and identity of the purified sample were assessed by SDS-PAGE (Laemmli, 1970) and N-terminal sequence analysis.

1.2. Dynamic light-scattering study

The oligomerization state of the purified *Tt*IGPS was examined by a dynamic light-scattering experiment using a DynaPro MS/X (Protein Solutions) instrument at a protein concentration of 20.0 mg ml^{-1} in 20 m*M* Tris-HCl pH 7.6 with 0.2 *M* NaCl. The measurements recorded at 291 K and analyzed using the *DYNAMICS* software v.3.30

(Protein Solutions) showed a monomodal profile centered at 2.5 nm radius and corresponding to a molecular weight of 27.7 kDa, which is consistent with a monomeric state in these solution conditions.

1.3. Crystallization

Crystallization trials were carried out using the oil-microbatch method (Chayen *et al.*, 1990) in Nunc HLA plates at 291 K using a TERA crystallization robot (Sugahara & Miyano, 2002). Equal volumes of protein solution (0.5 µl) and precipitant solution (0.5 µl) were mixed. The crystallization drop was overlaid with a 1:1 mixture of silicone and paraffin oils (13 µl), allowing slow evaporation of water in the drop. One condition provided the most well defined crystals; the precipitant solution consisted of 1.93 *M* ammonium sulfate, 0.1 *M* acetate-NaOH, pH 4.8. The best diffracting crystals grew to maximum dimensions of 0.2 × 0.2 × 0.15 mm after 5 days incubation of the crystallization solution at 295 K.

1.4. Data collection

The best crystal, with dimensions of 0.2 × 0.2 × 0.15 mm was used for data collection. Prior to data collection, the crystals were soaked in cryoprotectant (reservoir solution and 15% glycerol) for a few seconds and then were flash-cooled in a 100 K dry nitrogen stream. X-ray diffraction data collection was performed at 100 K using a RIGAKU FR-D generator (Cu radiation λ=1.5418 Å) and an R-AXIS IV image-plate detector. The crystal-to-detector distance was set to 170 mm and images of 0.5° oscillation were collected for 15 min. The data set was processed using the program *HKL*2000 (Otwinowski & Minor, 1997). Crystal belongs to the orthorhombic space group, $P2_12_12_1$, with unit-cell parameters a=63.652,

b=78.193, c=93.523 Å and consists of two protomers in the asymmetric unit. The calculated Matthews coefficient (Matthews, 1968) is 2.08 $Å^3$ Da^{-1}, which corresponds to a solvent-volume fraction of approximately 41%.

2. Supplementary tables

Table S1. Stabilization centers (SC) pairs in the four IGPS with secondary structure positions. The stabilization residues (SR) are marked in bold and underlined.

*Ss*IGPS	*Tm*IGPS	*Tt*IGPS	*Ec*IGPS
outer	*Outer*	*outer*	*outer*
Nter R3-I136 α4	α0 I15-V115 α3	Nter R2-L142 α4	α0 R19-L125 α3
Nter R3-L137 α4	α0-α00 V25-S121 α3	α0 R19-V116 β3-α3	α0-α00 Q23-L125 α3
Nter L5-K135 α4	α0-α00 R28-G151 α4	α0-α00 Y26-E123 α3	α0-α00 L25-Y124 α3
Nter L5-I136 α4	β1-α1 S52-F87 β2-α2	α0-α00 L28-E123 α3	α0-α00 F28-Y124 α3
α0 L17-V114 β3-α3	α4 I140-V170 α5	α0-α00 L28-A126 α3	β1-α1 A57-F93 β2-α2
α0 L17-K115 α3	α7 L218-W250 α8	α0-α00 P29-A126 α3	β1-α1 S58-F93 β2- α2
α0 R18-K115 α3		α0-α00 P29-F127 α3	α3 I123-V152 α4
α0 R28-G126 α3	*outer-core*	α0-α00 P31-A126 α3	α3 Y124-L156 α4-β5
β1-α1 K55-F89 β2-α2	α00-β1 K39-**<u>N226</u>** β8	α0-α00 P32-A126 α3	
β1-α1 S56-F89 β2-α2		β1-α1 Q55-F91 β2-α2	*outer-core*
β1-α1 V62M237 α8’	*core*	β1-α1 I62-M240 α8’	β1 K55-R65 β1-α1
β1-α1 V62-R238 α8’	β1 V42-**<u>N226</u>** β8	β1-α1 R63-R241 α8’	β6 N185-R197 α6
α1 F72-P240 α8	β1 V42-**<u>A227</u>** β8	α6 P196-L229 α7-β8	
α1 M73-I243 α8	β1 K43-**<u>A227</u>** β8	α6 G199-L229 α7-β8	*core*

α3	I119-Y148	α4	β1	K43-V228	β8	α6 G199-F230 α7-β8	β1 A49-G232 β8
α3	D120-Y152	α4 -β5	β1	I45-**A77**	β2		β1 I51-L234 β8
α5	E163-M200	α6	β1	I45-**I78**	β2	*core*	β1 I51-**A83** β2
			β1	**A46**-**I78**	β2	β1 L46-**A232** β8	β1 I51-**I84** β2
	outer-core		β2	**D76**-R103	β3	β1 S47-**A232** β8	β1 **L52**-**I84** β2
α00	R36-V78	β2	β2	**A77**-R103	β3	β1 S47-**V233** β8	β1 **L52**-**S85** β2
β7	**V208**-**V227**	α7-β8	β2	**A77**-**P104**	β3	β1 **V48**-**V233** β8	β1 **E53**-**S85** β2
			β2	**S79**-**L106**	β3	β1 I49-L234 β8	β1 **E53**-**V86** β2
	core		β2	**I80**-**A107**	β3	β1 I49-**A81** β2	β1 C54-**V86** β2
β1	**I48**-F230	β8	β3	L106-**A127**	β4	β1 I49-**V82** β2	β2 **S82**-Q109 β3
β1	I49-L231	β8	β3	A107-**I128**	β4	β1 **A50**-**V82** β2	β2 **A83**-Q109 β3
β1	I49-I232	β8	β4	A127-D153	β5	β1 **A50**-**S83** β2	β2 **A83**-**P110** β3
β1	I49-V78	β2	β4	**L129**-L155	β5	β1 E51-**S83** β2	β2 **S85**-**L112** β3
β1	I49-**G79**	β2	β5	L155-**I176**	β6	β2 **A81**-L107 β3	β2 **V86**-**C113** β3
β1	I49-**L80**	β2	β5	**V156**-**I176**	β6	β2 **A81**-**P108** β3	β3 **L112**-**A133** β4
β1	**A50**-L80	β2	β5	**V156**-**G177**	β6	β2 **S83**-**L110** β3	β3 **L112**-**C134** β4
β2	V78-I105	β3	β5	**G157**-**G177**	β6	β2 **V84**-R111 β3	β3 **C113**-**C134** β4
β2	**G79**-I105	β3	β6	I175-T204	β7	β2 T86-K112 β3	β3 **C113**-**L135** β4
β2	**G79**-P106	β3	β6	I175-V205	β7	β3 **L110**-**A131** β4	β4 **A133**-G159 β5
β2	**S81**-**L108**	β3	β6	**I176**-T204	β7	β3 **L110**-**A132** β4	β4 **C134**-G159 β5
β2	**I82**-**M109**	β3	β6	**I176**-V205	β7	β3 R111-**A132** β4	β4 **C134**-V160 β5
β3	**L108**-**V130**	β4	β6	**I176**-V206	β7	β3 R111-**L133** β4	β4 **L135**-L161 β5
β3	**L108**-**L131**	β4	β6	**G177**-**V207**	β7	β4 **A131**-E156 β5	β4 **L136**-**T162** β5
β3	**M109**-**L131**	β4	β7	**V207**-**A227**	β8	β4 **A131**-A157 β5	β5 L161-**V181** β6
β3	**M109**-**L132**	β4	β7	**V207**-V228	β8	β4 **A132**-E156 β5	β5 **T162**-**V181** β6

β4 T129-P156 β5	β7 A208-V228 β8	β4 **A132**-A157 β5	β5 **T162**-G182 β6
β4 **V130**-P156 β5		β4 **A132**-L158 β5	β5 E163-G182 β6
β4 **V130**-L157 β5		β4 **L133**-L158 β5	β6 V180-T210 β7
β4 **L131**-L157 β5		β4 **L134**-**V159** β5	β6 **V181**-V211 β7
β4 **L132**-**I158** β5		β5 L158-**L178** β6	β6 G182-I212 β7
β5 L157-**I177** β6		β5 **V159**-**L178** β6	β6 I183-S213 β7
β5 **I158**-**I177** β6		β5 **V159**-G179 β6	β6 N184-E214 β7
β5 **I158**-**G178** β6		β5 **E160**-G179 β6	β7 V211-N231 β8
β5 **E159**-**G178** β6		β6 V177-V210 β7	β7 I212-A230 β8
β6 **F176**-V205 β7		β6 **L178**-V210 β7	β7 I212-N231 β8
β6 **F176**-V206 β7		β6 **L178**-L211 β7	β7 I212-**G232** β8
β6 **I177**-V205 β7		β6 G179-V212 β7	β7 I212-F233 β8
β6 **I177**-V206 β7		β7 L211-D231 β8	β7 S213-A230 β8
β6 **I177**-K207 β7		β7 **V212**-F230 β8	β7 S213-**G232** β8
β6 **G178**-**V208** β7		β7 **V212**-**A232** β8	β7 S213-F233 β8
β7 **V208**-**A229** β8		β7 **V212**-**V233** β8	β7 S213-L234 β8
β7 **V208**-F230 β8		β7 A213-F230 β8	
β7 A209-**A229** β8		β7 A213-**A232** β8	
β7 A209-F230 β8		β7 A213-**V233** β8	

Table S2. Amino acid compositions of the SC cluster, SC and SR in the four IGPS proteins.

		*Ss*IGPS			*Tm*IGPS			*Tt*IGPS			*Ec*IGPS		
Amino acid		Clust.	SC	SR	Clust.	SC	SR	Clust.	SC	SR	Clust.	SC	SR
Charged	Lys	12	4	-	14	2		3	1	-	5	1	-
	Arg	13	5	-	8	1		11	5	-	5	3	-
	Asp	9	1	-	7	2	1	6	1	-	6	-	-
	Glu	17	2	1	11	1	1	11	4	1	6	3	1
		51	12	1	40	6	2	31	11	1	22	7	1
Polar	Asn	8	-	-	2	1	1	-	-	-	2	3	-
	Gln	3	-	-	3	-	-	1	1	-	7	2	-
	Ser	15	2	1	8	3	1	9	2	1	5	4	2
	Thr	2	1	-	2	1	-	3	1	-	1	2	1
		28	3	1	15	5	2	13	4	1	15	11	3
Aliphatic	Val	6	8	3	8	9	3	7	9	7	5	6	2
	Ile	12	10	4	7	8	4	2	2	-	3	5	1
	Leu	18	9	4	8	4	2	16	12	4	6	9	4
	Met	4	4	1	2	-	-	3	1	-	1	-	-
		40	31	12	25	21	9	28	24	11	15	20	7
Aromatic	Phe	7	4	2	4	1	-	5	3	-	4	3	-
	Tyr	9	2	-	3	-	-	2	1	-	1	1	-
	Trp	1	-	-	1	1	-	-	-	-	1	-	-
		17	6	2	8	1	0	7	4	0	6	4	0
Other	Ala	6	3	2	7	6	5	13	8	5	7	5	2
	Pro	8	3	-	4	1	1	13	5	1	3	1	1
	Cys	-	-	-	-	-	-	-	-	-	1	3	2
	Gly	7	3	2	5	2	1	9	2	-	3	3	1
	His	-	-	-	2	-	-	1	-	-	1	-	-
		21	9	4	18	9	7	36	15	6	15	12	6
Total		157	61	20	106	43	20	115	58	19	77	54	17

3. Supplementary figures

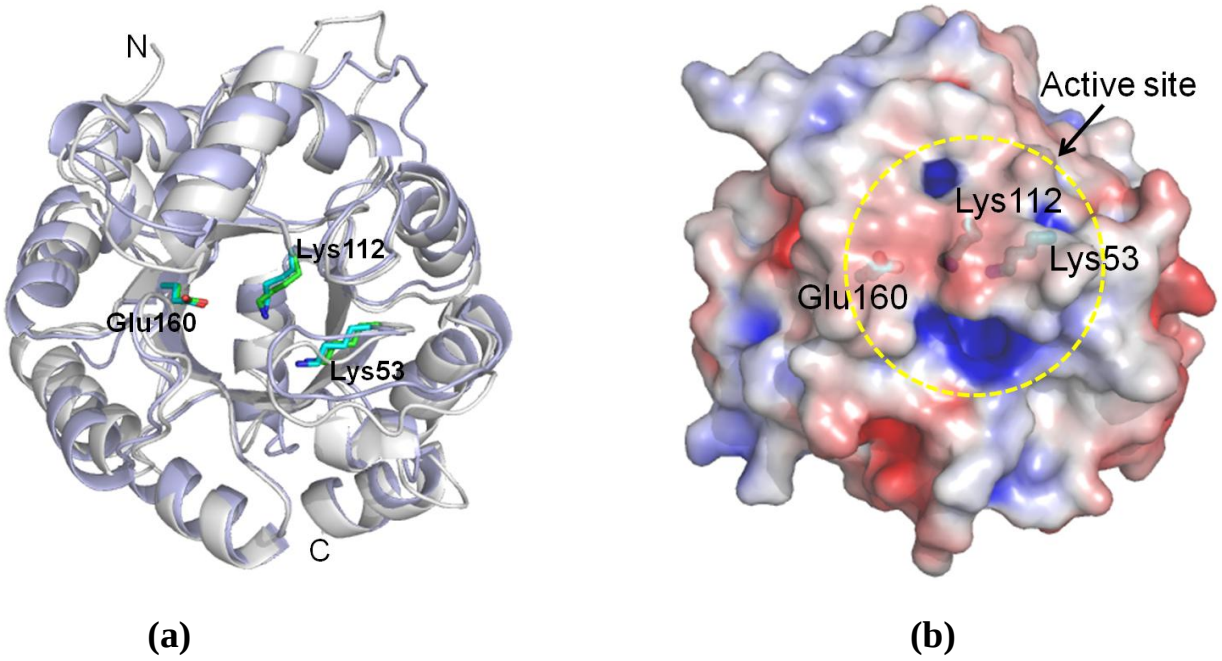
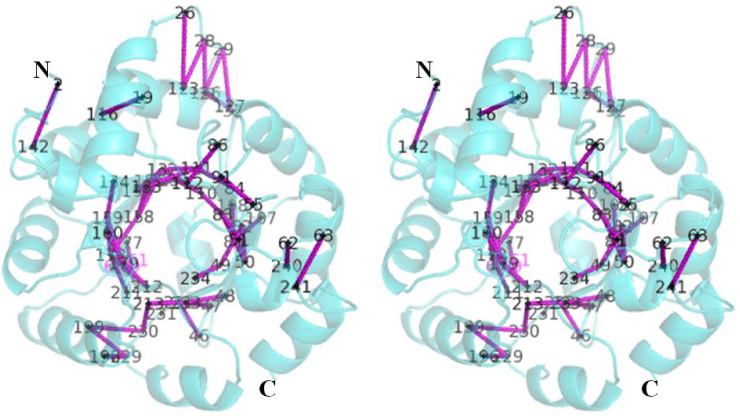
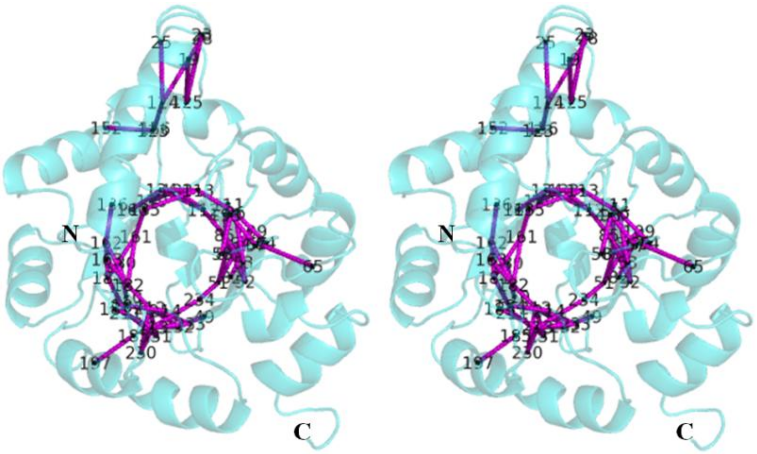


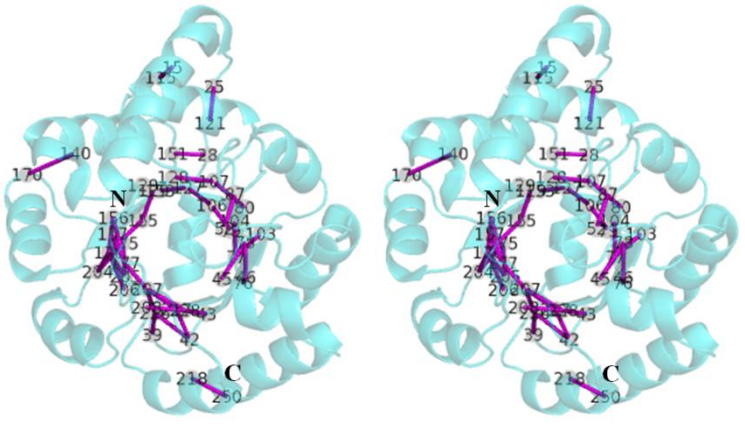
Figure S1. (a) Superposition of the active site residues of *Ss*IGPS (cyan) with structurally corresponding ones of *Tt*IGPS (green). Structural alignments of *Ss*IGPS (sky-blue) and *Tt*IGPS (grey) presented. All important active site residues of *Ss*IGPS are conserved in *Tt*IGPS suggesting a similar role in the proteins. The N- and C-terminal ends of the *Tt*IGPS polypeptide chain are labeled. (b) The molecular surface and electrostatic potential of the *Tt*IGPS structure. This view corresponds to looking down the barrel axis. Surface electrostatic potentials less than -5 kT, neutral, and greater than 5 kT are displayed in red and blue, respectively. The electrostatic potential is mapped onto the *Tt*IGPS molecular surface. The active site area is highlighted by the yellow circle and the active side residues are shown in the stick mode. Secondary-structure elements and a semi-transparent surface are shown. The electrostatic potential surface of *Tt*IGPS calculated using *APBS* (Adaptive Poisson-Boltzmann Solver) (Baker *et al*., 2001) is graphically represented using *PyMOL* (http://www.pymol.org/).



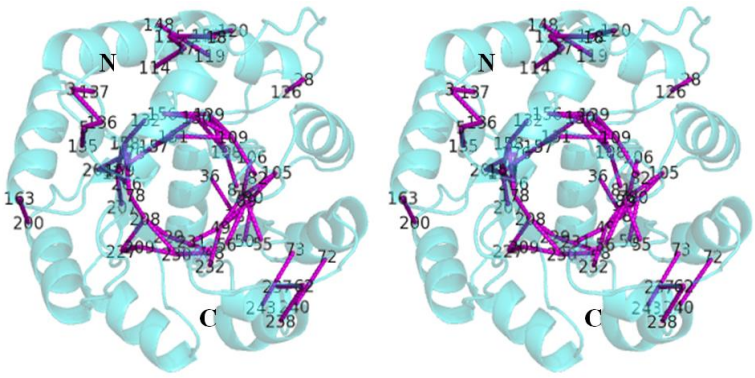
(a)



(b)



(c)



(d)

Figure S2. Distribution of the stabilization centers (SC) in the four IGPS: (c) *Tt*IGPS, (b) *Ec*IGPS, (c) *Tm*IGPS, and (d) *Ss*IGPS. For each protein, stereo-view of the SC pair interactions is shown in magenta and the C^{α} positions for SC residues are indicated by residue numbers. The corresponding $(\beta/\alpha)_8$-barrel fold is drawn in the cyan.

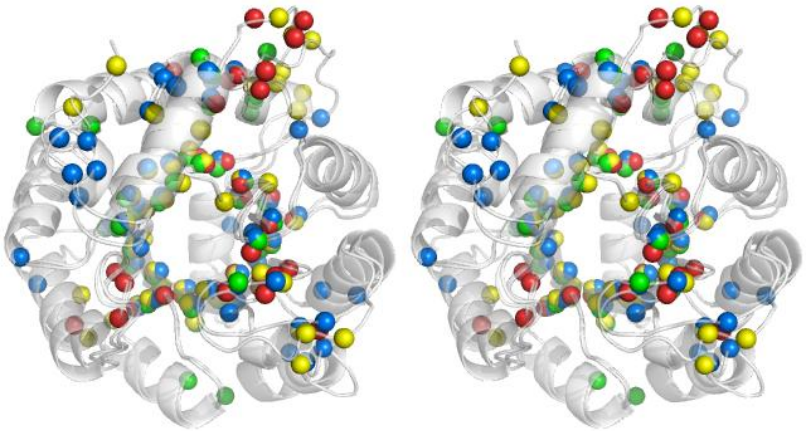


Figure S3. Spatial distribution of the SC residues in the four IGPS. The locations of the C^{α} positions for the SC residues are indicated by spheres. SCs of *Ec*IGPS are shown in red, *Tt*IGPS in yellow, *Tm*IGPS in green and *Ss*IGPS in blue. The $(\beta/\alpha)_8$-barrel fold is drawn in gray.

4. References

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 10037–10041.

Chayen, N. E., Shaw Stewart, P. D., Maeder, D. L. & Blow, D. M. (1990). *J. Appl. Cryst.* **23**, 297–302.

Laemmli, U. K. (1970). *Nature*, **227**, 680–685.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.