Electronic Supporting Information for

Mycobacterium tuberculosis H37Rv 3377c encodes the diterpene cyclase for producing the halimane skeleton

Chiaki Nakano,[a] Tomoo Okamura,[a] Tsutomu Sato,[a] Tohru Dairi[b] and Tsutomu Hoshino,[a]*

[a] Department of Applied Biological Chemistry, Faculty of Agriculture, and Graduate School of Science and Technology, Niigata University, Ikarashi, Niigata 950-2181, Japan. Fax: +81-25-262-6854; E-mail: hoshitsu@agr.niigata-u.ac.jp

[b] Biotechnology Research Center, Toyama Prefectural University, Toyama 939-0398, Japan

1. Analytical Method.

NMR spectra of the enzymic product were recorded in C_6D_6 on a Bruker DMX 600 spectrometer, the chemical shifts being relative to the solvent peak δ_H 7.28 and δ_C 128.0 ppm as the internal reference for ^{1}H- and ^{13}C NMR spectra, respectively. The NMR spectra of GGPP including ^{31}P were measured in D_2O containing 25 mM NH_4HCO_3 on a Bruker DPX 400. The ^{1}H- and ^{31}P NMR spectra were referenced to 0 ppm (TSP, trimethylsilyl propionic acid sodium salt, and 85% H_3PO_4, respectively) by external referencing. GC analyses were done on a Shimadzu GC-8A chromatograph equipped with a flame ionization detector (a DB-1 capillary column, 0.32 mm x 60 m). GC-MS spectra were on a JEOL SX 100 spectrometer under electronic impact at 70 eV with a DB-1 capillary column (0.32 mm x 30 m), the oven temperature being elevated from 180 to 270 °C (3 °C min^{-1}). HR-EIMS was performed by direct inlet system. Specific rotation values were measured at 25 °C with a Horiba SEPA-300 polarimeter.

2. Synthesis of GGPP

The synthetic method of GGPP was essentially the same as described in the literature.[1]

[1] V. Jo Davisson, A. B. Woodside and C. D. Poulter, in Methods in Enzymology, ed. J. H. Law and H. C. Rilling, Academic Press, NewYork, London, 1985, vol. 110, pp134-144.

GGOH was kindly gifted from Kuraray Co. LTD (Nakajoh, Japan). The following reaction was done under N_2 atmosphere. *N*-chlorosuccinimide (138 mg, 1.024 mmol) was dissolved in dry CH_2Cl_2 (3 cm^3) at -30°C. To the solution, 78 μcm^3 (1.024 mmol) of dimethylsulfide was added slowly and the reaction temperature was warmed gradually to 0°C and again lowered to -40°C. 100 mg of GGOH (0.343 mmol) in CH_2Cl_2 (3 cm^3) was slowly added over 1 h and the temperature was gradually elevated up to 0 °C, followed by further reaction at 0 °C for 2 h. The reaction mixture was poured into ice-cooled brine and was extracted with hexane, which was dried over anhydrous $MgSO_4$. The hexane was removed under reduced pressure to give a crude GGCl (65 mg). The solution of this chloride dissolved in 0.5 cm^3 of CH_3CN was added into a CH_3CN solution of Tris(tetra-*n*-butyl)ammonium hydrogen pyrophosphate (0.7g/6 cm^3), which was prepared according to the published protocol,[1] and stirred for 2 h at a room temperature. After evaporating the solvent, the residue was dissolved in 0.9 cm^3 of a mixture of 2-PrOH and 25 mM NH_4HCO_3 (1:49), the solution being loaded on the column of Dowex 50W x 8 (50-100 mesh, NH_4^+ form), which was previously washed with the mixed solution of 2-PrOH and NH_4HCO_3, and elution was further continued. The fractions showing a neutral pH were combined and lyophilized, giving the crude GGPP (241 mg). The crude residues (40 mg) was subjected to a reversed-phase HPLC chromatography (Shiseido, CAPCELL PAK C18) with MeOH/25 mM NH_4HCO_3 (8:2), yielding pure GGPP (3 mg). ^{1}H NMR; δ_H 1.571 (6H, s, Me x 2), 1.597 (3H, s, Me), 1.643 (3H, s, Me), 2.1-1.9 (12H, m), 4.446 (2H, br t, *J*= 6 Hz), 5.120 (3H, m), 5.429 (1H, br s). ^{31}P NMR; δ_P -5.09, -7.97 (each d, *J*=20.5 Hz)

Geranyl (GPP) and farnesyl diphosphates (FPP) were prepared from geraniol and farnesol, respectively, according to the same method as described above. GPP; δ_H 1.53 (3H, s), 1.58 (3H, s), 1.62 (3H, s), 1.99 (2H, t, *J*=7.2 Hz), 2.05 (2H, t, *J*=7.2 Hz), 4.36 (2H, *J*=6.5 Hz), 5.09 (1H, t, *J*= 6.2 Hz), 5.36 (1H, t, *J*=6.7 Hz); δ_P -3.20, -6.96 (each d, *J*=22 Hz). FPP; δ_H 1.55 (6H, s, Me x 2), 1.62 (3H, s), 1.65 (3H, s), 2.1-1.9 (8H, m, CH_2 x 4), 4.40 (2H, t, *J*= 6.8 Hz), 5.14 (2H, m), 5.40 (1H, t, *J*=7.5 Hz), δ_P -3.67, -7.06 (each d, *J*=21 Hz).

3. Amino acid alignment of Rv3377c from Mycobacterium tuberculosis, the cyclase from Kitasatospora griseola and abietadiene synthase from Abies grandis.

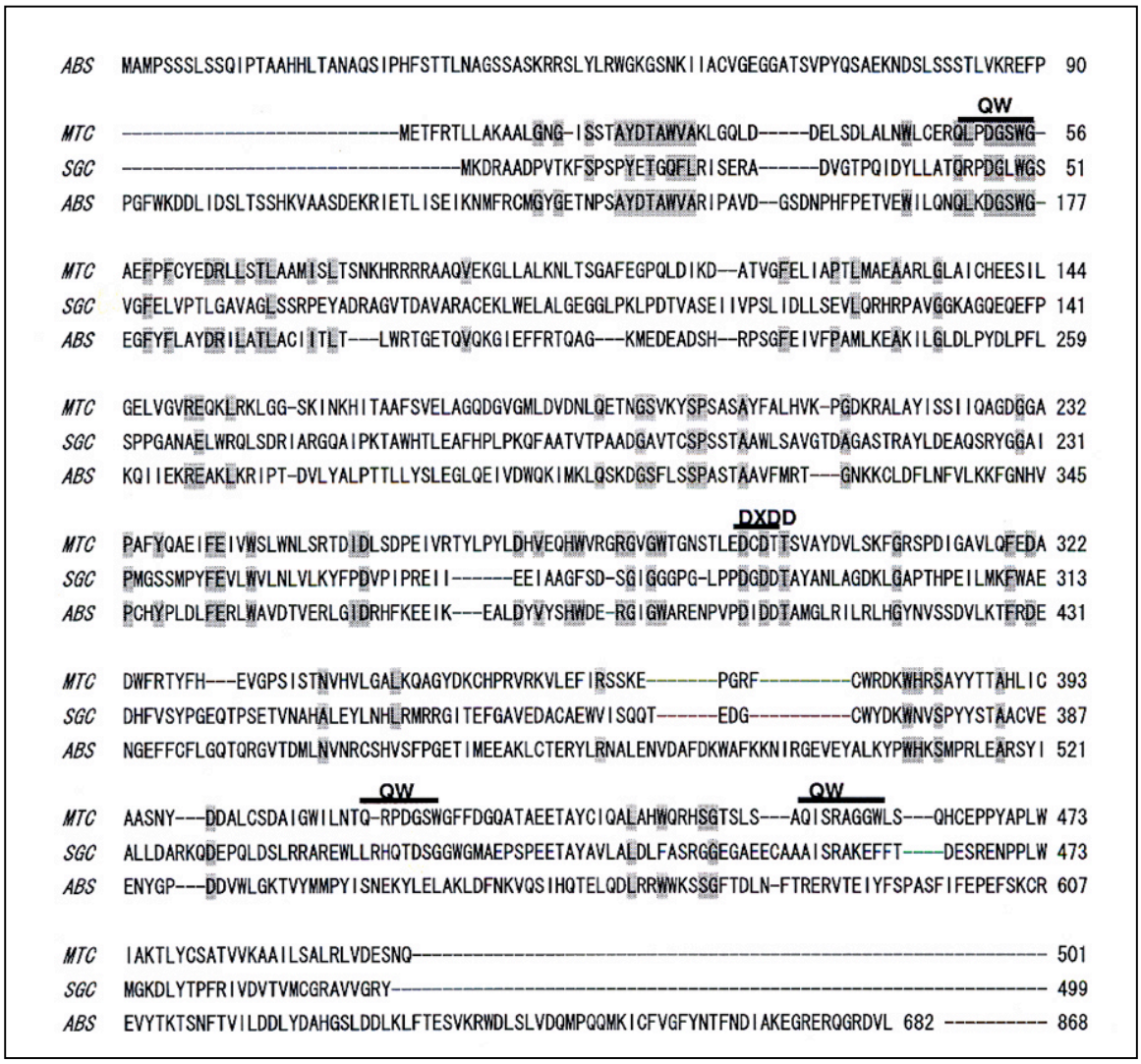


Figure S1. Amino acid alignment of diterpene cyclases.

Identical amino acids are shaded. Dashes indicate gaps in the alignment. The two characteristic motifs of QW and DXDD are highly conserved among terpene cyclases. MTC, Rv3377c product from *Mycobacterium tuberculosis*; SGC, Diterpene cyclase from *Kitasatospora griseola* (formerly called *Streptomyces griseolosporeus*); ABS, Abietadiene synthase of *Abies grandis*.

4. Construction of the plasmid harboring Rv3377c gene and SDS page of expressed Rv3377c protein

The genomic DNA was gifted from Corolado State University (TB research Materials and Vaccine Testing Contract), and used as the template for the PCR reaction. The following primers were used for amplifying the desired Rv3377c gene.

Sense primer: 5'- gagggcggatccaatggagactttcaggactctg -3' (BamH I site is underlined)
Anti-sense primer: 3'- cgaccagctactctcattggtttttcgaacaggg -5' (Hind III site is underlined)
The amplified gene was cloned into the BamH I/Hind III site of pET-22b(+) (Novagen).

The following primers were used for preparing the site-directed mutants:

D293N

Sense primer 5' - ccaccttggaaaattgtgacaccacaagtg - 3'
Anti sense primer 3' - ggtggaacctttttaacactgtggtgttcac - 5'

D295N

Sense primer 5' - ccttggaagattgtaaccaccacaagtgtgg - 3'
Anti sense primer 3' - ggaaccttctaacattgtggtgttcacacc - 5'

T296D

Sense primer 5' - caccttggaagattgtgacgacacaagtgtggcctacg - 3'
Anti sense primer 3' - gtggaaccttctaacactgctgtgttcacaccggatgc - 5'

The underlines indicate the point mutation sites.

5. Preparation of the cell-free extract and incubation conditions, and GC traces of the enzymic reactions.

The recombinant *E. coli*, in which the coexpression system (GroE chaperon) was constructed, was cultivated at 37°C in a 100 cm^3 of LB medium. When the optical density reached 0.6, IPTG (1 mM) was added to the culture, and cultivation was continued at a lower temperature (25°C) for 5 h. To the pellets (ca 0.7 g) collected after the centrifugation, was added 10 ml of Tris buffer solution (pH 7.5, 50 mM) containing Triton X-100 (1%) and dithiothreithol (1 mM), and then ultrasonicated to prepare the cell-free extract, the supernatant of which was used as the enzyme source. The incubation conditions were as follows. The prepared cell-free homogenates (4.85 cm^3) were added to the solution (0.15 ml) consisting of the substrate (100 μg) and 10 mM $MgCl_2$, and then incubated for 20 h at 25°C. The reaction was quenched by adding 15% KOH/MeOH and the reaction mixture was extracted with hexane, which was subjected to GC analyses.

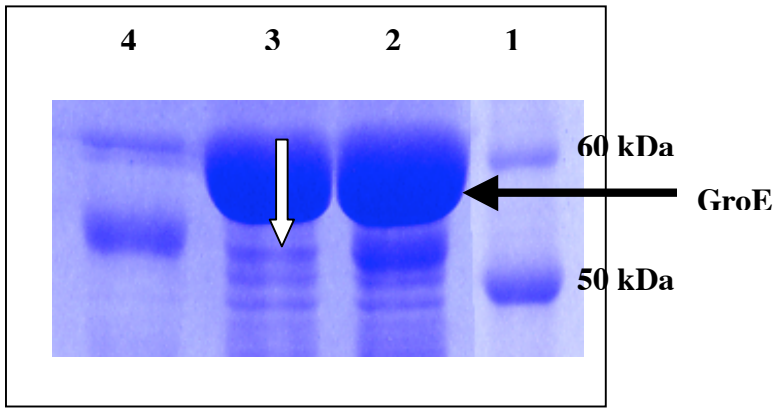


Figure S2. SDS page of the expressed protein of Rv3377c (7% acrylamide)
Lane 1: molecular weight marker.
Lane 2: total protein from *E. coli* BL21(DE3)/pET-22b, which was coexpressed with a chaperon GroE. Lane 3: soluble proteins. Lane 4: insoluble proteins.
Open and closed arrows indicate the soluble Rv3377c protein and GroE, respectively. Almost all of the expressed protein formed an inclusion body (see Lane 4).

The cell-free extract was fractionized by centrifugation (13,000 G x 15 min) to obtain soluble and insoluble fractions.

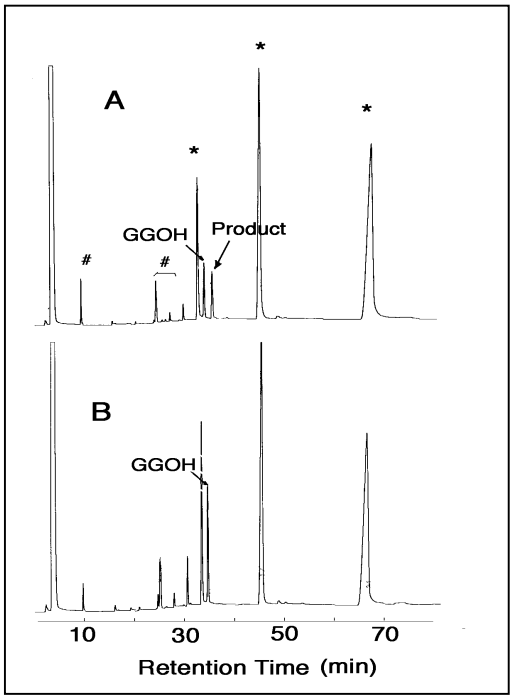


Figure S3. GC chromatograms of the hexane extract obtained after incubating GGPP with the cell-free homogenates of the cloned *E. coli* having Rv3377c gene (**A**) and lacking Rv3377c gene (**B**).
Incubation conditions: substrate, 100 μg of **1**, incubated at 25°C, pH 7.5 for 20 h with the cell-free homogenate (4.85 cm^3). The identical amount of the cell-free extracts was used to compare **A** and **B**. The symbols of asterisk (*) and # indicates Triton X-100 and impurities, respectively.

6. EIMS spectrum and the detailed NMR analyses of 2.

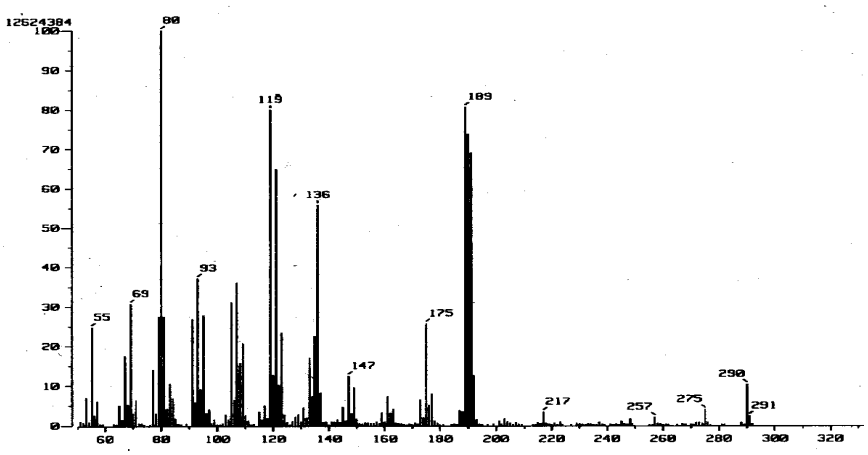
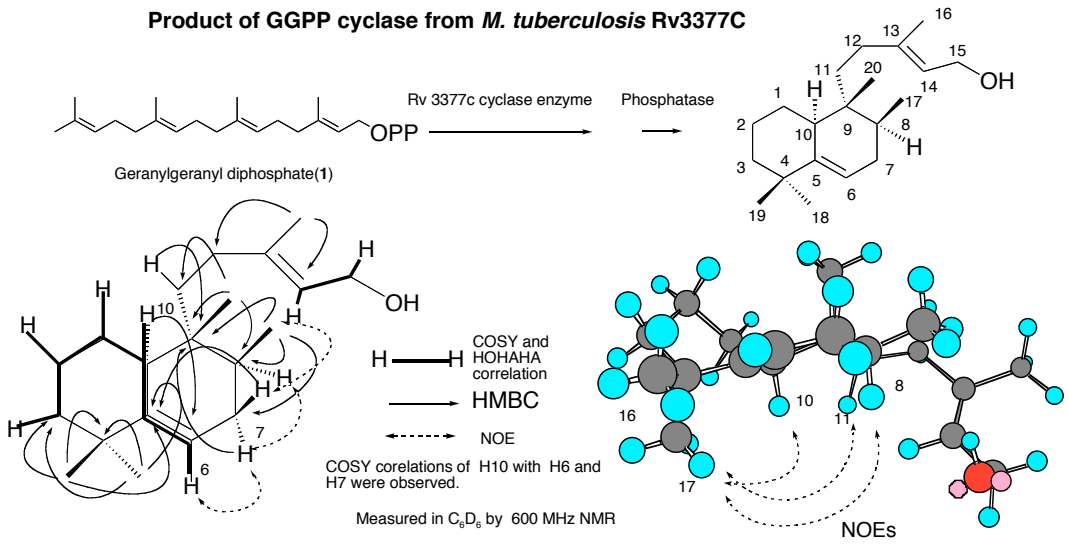


Figure S4. EIMS spectrum of enzymic product **2**.



	^{1}H	^{13}C		^{1}H	^{13}C		^{1}H	^{13}C		^{1}H	^{13}C
1	1.21(m); 1.86(m)	**27.22**	**6**	5.67(1H,bt)	**116.73**	**11**	1.50 (m); 1.66(m)	**35.35**	**16**	1.63 (3H, s)	**16.39**[a]
2	1.67 (2H, m)	**22.60**	**7**	1.94 (2H, m)	**31.97**	**12**	1.99(ddd, J-4, 12.8,13.2) 2.06(ddd, J=4,12.8,13.2)	**33.05**	**17**	0.92 (3H, d, J=6.8)	**15.28**
3	1.38(m); 1.52(m)	**41.25**	**8**	1.64 (m)	**33.70**	**13**	—	**139.2**	**18**	1.27(3H, s) *b*	**30.04**[b]
4	—	**36.28**	**9**	—	**37.25**	**14**	5.54(bt)	**124.5**	**19**	1.22 (3H, s) *c*	**29.15**[c]
5	—	**146.2**	**10**	2.37 (bd, J=11.8)	**40.31**	**15**	4.10 (2H, d,J=6.5)	**59.41**	**20**	0.81 (3H, s)	**16.43**[a]

The cabon signal of *a* is exchangeable. The protons and carbons of *b* and *c* are excahngeable

Figure S5. NMR analyses for proposing the structure of **2**.

7. Optimum pH and temperature for the catalysis.

After incubating 100 μg of GGPP **1** at various pHs and temperatures for 15 h with the same amount of the cell-homogenates (4.85 cm^3), the incubation mixture was heated at 70°C and the denatured proteins were removed after centrifugation. To the supernatant, 1 cm^3 of acetate buffer solution (1M, pH 5.6), 2 cm^3 of 2-PrOH and 0.1 cm^3 of acid phosphatase solution (concentration 37 mg/ml, 1.2 unit/mg, from potato, Sigma) were added, then the mixed solution was further incubated at 37°C for 12 h. The reaction was quenched by adding 7 ml of 15% KOH/MeOH. The hexane extracts were analyzed by GC. The relative ratio of the peak area of product **2** to that of unreacted GGOH was used for estimating the cyclization yields (%).

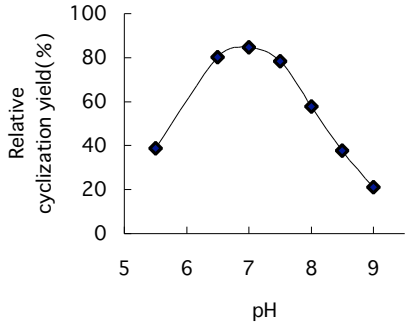


Figure S6. Effect of pH on the cyclization activity.

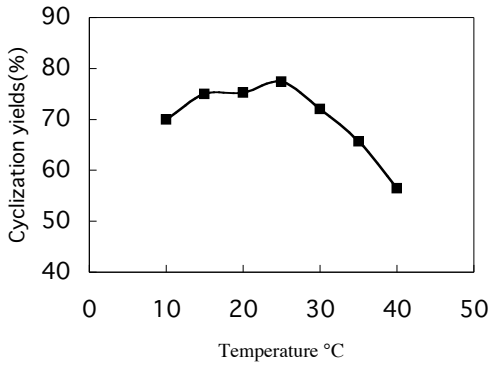


Figure S7. Effect of incubation temperatures on the cyclization activities.

8. Effect of Mg^{2+} ion and EDTA on the cyclization reaction.

The cyclase was purified with a Ni^{2+}-NTA column according to the manufacturer's protocol (Quiagen). To the eluted solution (3 cm^3) containing 250mM imidazole and the cyclase enzyme was added EDTA (final concentration 10 mM), stood for 1 h, and then passed through a Sephadex G-25 short column (1.5 cm x 5 cm) to remove the excess imidazole. All the purification procedure was done at 4°C. The protein concentration was estimated with Bio-Rad Protein Assay (Bio-Rad) to be 1.1 mg/l. The following solution was prepared; protein 700 μl, GGPP 50 μl (1mg/ cm^3, 50 μg) and 0, 2.5 μl, 12.5 μl, 25 μl of 1M $MgCl_2$ (final concentration; 0, 1mM, 5mM, 10mM, respectively), and then diluted with incubation buffer (pH 7.5) to the total volume of 2.5 cm^3. Incubations were carried out at 30°C and for 20 h and terminated by heating

70°C. To completely eliminate the diphosphate group, acid phosphatase was added and incubated under the same protocol as described above. After adding 15% KOH/MeOH, the hexane extract was subjected to GC analyses (Fig.ESI- 8). Unless acid phosphatase was added, product **2** was not extracted with hexane, strongly supporting that Rv3377c enzyme catalyses only the cyclization reaction (**1**→**4**), but not catalyse the hydrolysis of the phosphate of **4**.

When the cyclase was purified with Ni^{2+}-NTA column under no addition of EDTA, the purified enzyme still had the cyclase activity, but the activity was decreased when incubated in the presence of EDTA (Fig. ESI-9). This finding suggests that Mg^{2+} ion was somewhat tightly associated with the cyclase enzyme.

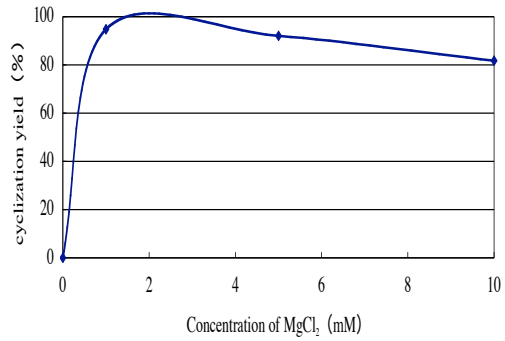


Figure S8. Effect of Mg^{2+} ion on the cyclase activity. Each experiment was done with protein 770 μg, which was purified in the presence of EDTA.

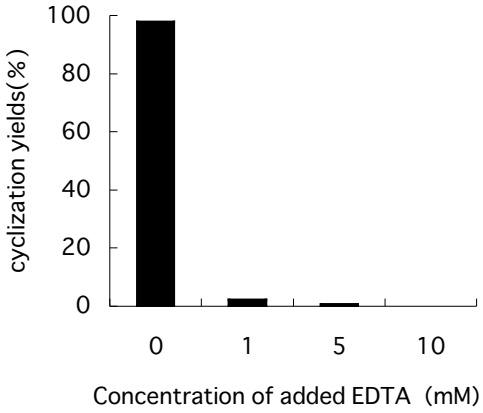


Figure S9. Effect of EDTA concentration on the cyclase activity. Each experiment was carried out with protein 770 μg, which was purified under no addition of EDTA.