Supporting Information

Characterization of SpnQ from the Spinosyn Biosynthetic Pathway of *Saccharopolyspora spinosa*: Mechanistic and Evolutionary Implications for C-3 Deoxygenation in Deoxysugar Biosynthesis

Lin Hong, Zongbao Zhao, and Hung-wen Liu*

Division of Medicinal Chemistry, College of Pharmacy, and Department of Chemistry and Biochemistry, University of Texas, Austin, Texas 78712 (USA)

1. PCR Amplification and Cloning of *spnQ*

PCR Amplification and Cloning of the spnQ Gene. Two oligonucleotide primers were used to amplify the *spnQ* gene from the genomic DNA of *Saccharopolyspora spinosa*: primer Q-1, 5'-GGGGCCTC**CATATG**AGCAGTTCTGTCGAAGC-3' containing an *Nde*I restriction site (in bold) and primer Q-2, 5'-GAG**CTCGAG**TTATCGCC-CCAACGCCCACAAG-3' containing an *Xho*I restriction site (in bold). Polymerase-mediated amplification was carried out in a 0.5 mL microcentrifuge tube. 37.5 _ L of double-deionized H_2O, 5.0 _ L of *pfu* polymerase buffer (10×), 4.0 _L of deoxyribonucleotidyl-triphosphate mix (2.5 mM each), 1.0 _ L of start primer Q-1 (10 _ M), 1.0 _ L of halt primer Q-2 (10 _ M), 0.5 _ L of genomic DNA (about 0.1 _ g) from *Saccharopolyspora spinosa* as the amplification template, and 1.0 _ L of cloned *pfu* polymerase (2.5 units) were used. The reaction mixture was overlaid with a layer of mineral oil and subjected to the following thermal cycles: 1 denaturation cycle at 95 °C for 3 min; 30 amplification cycles of incubation at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 2.5 min; and 1 extension cycle at 72 °C for 10 min. The reaction tube was held at 4 °C at the end of these cycles. The PCR amplified *spnQ* gene was purified, digested with *Nde*I and *Xho*I, and ligated into the *Nde*I/*Xho*I sites of the transcription vector pET24b(+). Positive clones were identified by digestion analysis and further verified by sequencing the plasmid DNA. The plasmid DNA from positive clones (pKZQO) was used to transform *Escherichia coli* BL21(DE3). The general methods and protocols for recombinant DNA manipulations used were as described by Sambrook *et al*.[1]

2. Purification of SpnQ. All purification steps were carried out at 4 °C.

Step 1: Growth of Escherichia coli BL21 (DE3)-pKZQO Cells. An overnight culture of *E. coli* BL21(DE3)-pKZQO, grown in Luria-Bertani (LB) medium supplemented with kanamycin (30 _ g/mL) at 37 °C, was diluted 250-fold with 6 L of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was adjusted to 24 °C and the incubation was continued for one more hour in order for the OD_{600} of the culture to reach 0.4. The culture was then induced with isopropyl β-D-thiogalactoside (IPTG) to a final concentration

of 0.2 mM and allowed to grow for an additional 18 h at 24 °C. The cells were harvested by centrifugation at 5,000*g* for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. Thawed cells were resuspended in 60 mL of lysis buffer (25 mM Tris·HCl, 0.6 mM PMSF, 5 mM 2-mercaptoethanol, 10% glycerol, pH 7.5), and disrupted with five 1 min sonication bursts, with 1 min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000*g* for 25 min.

Step 3: Ammonium Sulfate Precipitation. Ammonium sulfate crystals were added to the supernatant from the previous step in two portions over 20 min to give 10% saturation. After stirring for 1 h, the precipitated proteins were removed by centrifugation at 15,000*g* for 20 min and more ammonium sulfate crystals were added to the supernatant in three portions over 30 min to give 70% saturation. The mixture was slowly stirred for 2 h. The precipitated protein was collected by centrifugation at 15,000*g* for 20 min, and re-suspended in a minimal amount of lysis buffer. The resulting protein solution was desalted by dialysis against four 1 L of lysis buffer which were changed every 1.5 h.

Step 4: DEAE-Sepharose CL6B Chromatography. The dialysate from step 3 was applied to a DEAE-Sepharose CL6B column (2.5 × 40 cm) pre-equilibrated with buffer A (25 mM Tris·HCl, 1 mM 2-mercaptoethanol, 15% glycerol, pH 7.5). The column was washed with 200 mL of the same buffer, and the protein mixture eluted with a linear gradient of 800 mL of buffer A and 800 mL of buffer B (buffer A plus 0.8 M NaCl). The flow rate was 1.5 mL/min and fractions of 12 mL were collected throughout the gradient elution. The fractions containing the desired SpnQ protein as judged by SDS-PAGE were combined, desalted by repeated dilution with buffer A and concentration by ultrafiltration (YM-10 membrane), and finally concentrated to 8 mL.

Step 5: Hydroxyapatite Chromatography. The protein from the previous step was applied to a hydroxyapatite column (2.5 × 22 cm) that was pre-equilibrated with buffer C (10 mM potassium phosphate, 1 mM 2-mercaptoethanol, 15% glycerol, pH 7.5). The column was washed with 100 mL of the same buffer, and the protein mixture eluted with a linear gradient of 400 mL of buffer C and 400 mL of buffer D (600 mM potassium phosphate, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The flow rate was 1.0 mL/min and fractions of 8 mL were collected throughout the gradient elution. The fractions containing the desired SpnQ protein as judged by SDS-PAGE, were pooled and concentrated to 3 mL by ultrafiltration (YM-10 membrane).

Step 6: Sephacryl S-200 HR Chromatography. The protein from step 5 was applied to a Sephacryl S-200 HR column (2.5 × 100 cm) pre-equilibrated with buffer E (25 mM Tris·HCl, 15% glycerol, pH 7.5). The column was washed/eluted with the same buffer at a flow rate of 0.8 mL/min. Fractions of 8 mL were collected throughout the entire elution. Pure fractions were identified by SDS-PAGE, and concentrated to 2.0 mL by ultrafiltration (YM-10 membrane). The purified SpnQ protein was stored at –80 °C. A total of 30 mg of light red-brown protein was obtained from a 6 L culture. Including 15% glycerol was crucial in minimizing protein precipitation during the purification process. The subunit molecular mass of 50 kDa, as revealed by SDS-PAGE, correlates well to the predicted value of 50381 Da based on the translated peptide sequence.

Molecular Mass Determination. The molecular mass of native SpnQ was estimated to be 109.4 kDa by gel filtration chromatography, suggesting that the protein exists as a homodimer in solution. Native molecular mass of SpnQ was measured by size exclusion chromatography using a Pharmacia FPLC equipped with a Superdex 200 HR 10/30 column. The eluant was 50 mM sodium phosphate, 0.15 M NaCl, pH 7.0. A flow rate of 0.5 mL/ min was used in the analysis. Calibration of the column was achieved using the following protein standards (Sigma): cytochrome c (12.4 kDa), carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), and _-amylase (200 kDa). The void volume (Vo) of the column was measured using blue dextran. A linear fit to a plot of the logarithm of the molecular weight of protein standard versus V_e/V_o was used as the standard curve to estimate the native molecular mass (M_r) of SpnQ.[2] V_e = elution volume of protein, V_o = void volume.

3. UV-visible Spectrum of SpnQ.

The spectrum of SpnQ exhibits bands centered around 330 and 460 nm, with a broad shoulder at 550 nm.[3,4] The iron content of as-isolated SpnQ was estimated to be in the range of 0.72 to 1.22 equivalents per subunit.[5]

4. Assay Conditions for SpnQ and HPLC Analysis of Reaction Mixture.

A typical SpnQ assay mixture contained 1 mM of the SpnQ substrate (**6**), 250 _ M PMP, 30 _ M SpnQ, and the reducing agent/system being tested (such as 60 _ M E_3 with 2 mM NADH, 0.6 mM sodium dithionite, or 30 _ M each of flavodoxin/flavodoxin reductase or ferredoxin/ferredoxin reductase with 4 mM NADPH) in 100 _ L of 50 mM potassium phosphate (pH 7.5). For assay of SpnQ using L-glutamate, a typical reaction contained 1 mM **6**, 250 _ M PLP, 10 _ M SpnQ, and 3 mM L-glutamate. PMP (in the case of assays using reducing agent/system) or PLP (in the case of assays using L-glutamate) was included in reaction mixtures in order to reconstitute SpnQ. Overexpression in *E. coli* and the purification process have rendered some of the SpnQ active sites devoid of PMP. PLP is converted to PMP in the presence of L-glutamate by SpnQ. As-purified SpnQ was used for assays which employed ferredoxin/ferredoxin reductase or flavodoxin/flavodoxin reductase as the reducing system as well as for assays of SpnQ using L-glutamate. These assays were performed under aerobic conditions. Anaerobically reconstituted SpnQ was used for later assay trials of SpnQ with E_3/NADH and for assays using sodium dithionite. The assays with E_3/NADH were performed under aerobic conditions, while those using sodium dithionite were performed under anaerobic conditions. The E_3-encoding gene was originally cloned from *Yersinia pseudotuberculosis*, and E_3 was overexpressed and purified according to published procedures.[6] In the assays using reduction by dithionite, reconstituted SpnQ was treated with 20 equivalents of dithionite and was incubated at room temperature for 40 min in the anaerobic glove box in order to fully reduce the iron-sulfur clusters in preparation for the assay. All reactions were initiated by the addition of SpnQ and incubation was carried out at 24 ^{o}C. Aliquots of each reaction mixture were withdrawn at appropriate time intervals, and the reaction was terminated by 10-fold dilution of each sample with water and flash freezing in liquid nitrogen. Prior to analysis, each sample was thawed and the enzyme was removed by filtration through Centricon YM-10 membrane. The resulting enzyme free solution was then analyzed by HPLC using a Dionex CarboPac PA1 column. A linear gradient from 5 to 20% B in A over 15 min, followed by another linear gradient from 20 to 60% B in A over 20 min was employed (solvent A: water; solvent B: 500 mM NH_4OAc,

adjusted with diluted $NH_3 \cdot H_2O$ to pH 7.0). The cycle was concluded by an increase in percent B from 60 to 100% over 2 min, followed by a wash with 100% B for 3 min, a decrease in percent B from 100 to 5% over 5 min, and finally a re-equilibration with 5% B for 15 min. This method was used for the analysis of all reactions described herein. Percent conversion was determined by integration of substrate and product peak areas on the HPLC trace. The retention times were 33.4 min for SpnQ substrate **6**, 35.8 min for SpnQ product **7**, and 14.7 min for the SpnQ aminosugar product **15**.

5. Reconstitution of SpnQ.

Reconstitution of SpnQ was conducted anaerobically. All buffers, chemicals, and enzyme solution were degassed prior to transferring into the anaerobic glove box (Coy Laboratory Products Inc.). DTT was added to the SpnQ solution at a final concentration of 10 mM and the resulting mixture was incubated for 20 min. Then, 5 molar equivalents of PMP, $Fe(NH_4)_2(SO_4)_2$, and Na_2S, were sequentially added to the SpnQ solution (1 equivalent) with a 15 min incubation period between each addition. The addition of $Fe(NH_4)_2(SO_4)_2$ and Na_2S were repeated until the final concentrations of $Fe(NH_4)_2(SO_4)_2$ and Na_2S each reached 20 times that of SpnQ. The extra reagents were removed using an Econo-Pac 10 DG column according to the protocols provided by the manufacturer. Briefly, the reaction mixture was diluted to 3 mL with 25 mM Tris·HCl buffer containing 15% glycerol (pH 7.5), and was applied to the column which had been equilibrated with the same Tris·HCl buffer. The column was washed with the same buffer and the first 3 mL of eluant from the column was collected as the desalted protein. The reconstituted protein was concentrated with Centricon YM-10 by microcentrifugation.

6. Preparation and Characterization of the SpnQ Aminosugar Product 15.

A200 _ L incubation contained 0.21 _ mol of the SpnN product, 0.05 _ mol of PLP, 0.6 _ mol of L-glutamate, 0.01 _ mol of SpnQ in 50 mM potassium phosphate buffer (pH 7.5). The reaction was conducted at 37 °C overnight. The enzyme was removed by filtering through Centricon YM-10 by microcentrifugation. Aliquots of 20 _ L of the filtrate were repeatedly injected into HPLC and separated with the Dionex CarboPac PA1 analytical column. The reaction product whose retention time was 14.7 min was manually collected. After all the reaction filtrate was injected, the collected eluant from each individual run was combined and lyophilized. The sample was then redissolved in 100 _ L of H_2O and subjected to ESI-MS analysis, which revealed a molecular weight of 531.

7. Preparation and Characterization of Product 7.

A preparative-scale reaction contained 15.2 _ mol SpnQ substrate (**6**), 0.75 _ mol PMP, 0.94 _ mol $NADP^+$, 0.3 mmol isopropanol, 8.5 mg $NADP^+$-dependent alcohol dehydrogenase, 1.8 mg ferredoxin, 0.2 mg ferredoxin reductase, and 6.0 mg SpnQ in 3.6 mL of 50 mM potassium phosphate buffer (pH 7.5). Alcohol dehydrogenase and isopropanol were used as part of an NADPH regeneration system. The reaction was initiated by the addition of $NADP^+$ and was incubated at 37 °C. The progress of the reaction was monitored by HPLC analysis as described above. The reaction was almost complete in 2 h. The enzymes in the reaction mixture were removed by filtering through Centricon YM-10 by microcentrifugation. The filtrate was loaded onto a Bio-Rad P2 column (extra fine, 2.5 × 150 cm) which was pre-equilibrated with 25 mM

NH_4HCO_3. The sample was eluted with the same buffer at a flow rate of 12 mL/h at 4 °C. Fractions that displayed a single absorption maximum at 267 nm were lyophilized individually and analyzed by NMR. The desired fractions were pooled and lyophilized. A total of 3 mg of **7** was obtained. Spectral data for the SpnQ product (**7**): ^{1}H NMR (D_2O, mixture of keto and hydrated keto form) _ 0.87, 0.97 (3H, d, *J* = 6.5 Hz, 5-Me), 1.64 (3H, s, 5"-Me), 1.54-2.44 (4H, m, 2-H, 3-H), 2.02-2.13 (2H, m, 2'-H), 3.84 (1H, q, *J* = 6.5 Hz, 5-H), 3.87-3.91 (3H, m, 4'-H, 5'-H), 4.33-4.37 (1H, m, 3'-H), 5.28 (1H, d, *J* = 6.5 Hz, 1-H), 6.08 (1H, t, *J* = 7.0 Hz, 1'-H), 7.50 (1H, s, 6"-H); ^{13}C NMR (D_2O, hydrated form) _ 11.7, 11.8, 28.0, 30.3, 38.6, 70.5, 71.0, 84.9, 85.3, 91.4, 93.7, 93.8, 111.8, 137.4, 151.8, 166.7; ^{31}P NMR (D_2O) _ –10.7 (d, *J* = 21 Hz), -12.5 (d, *J* = 21 Hz). High-resolution FAB-MS (negative mode) calc for $C_{16}H_{23}N_2O_{13}P_2$ $(M-H)^+$ 513.0675, found 513.0692.

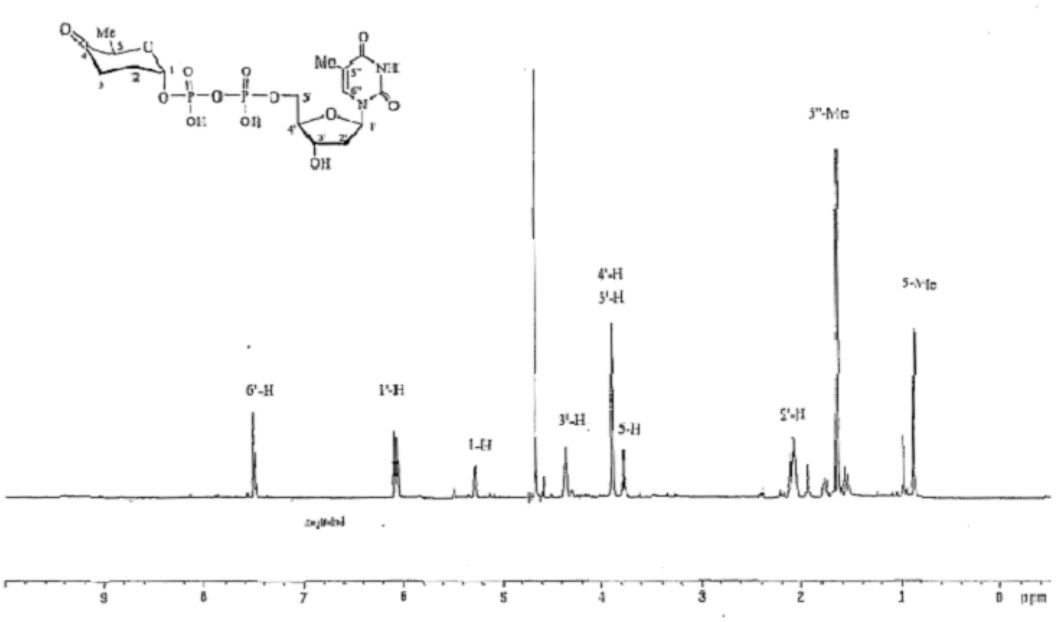

Figure S1. ^{1}H NMR spectrum of **7**

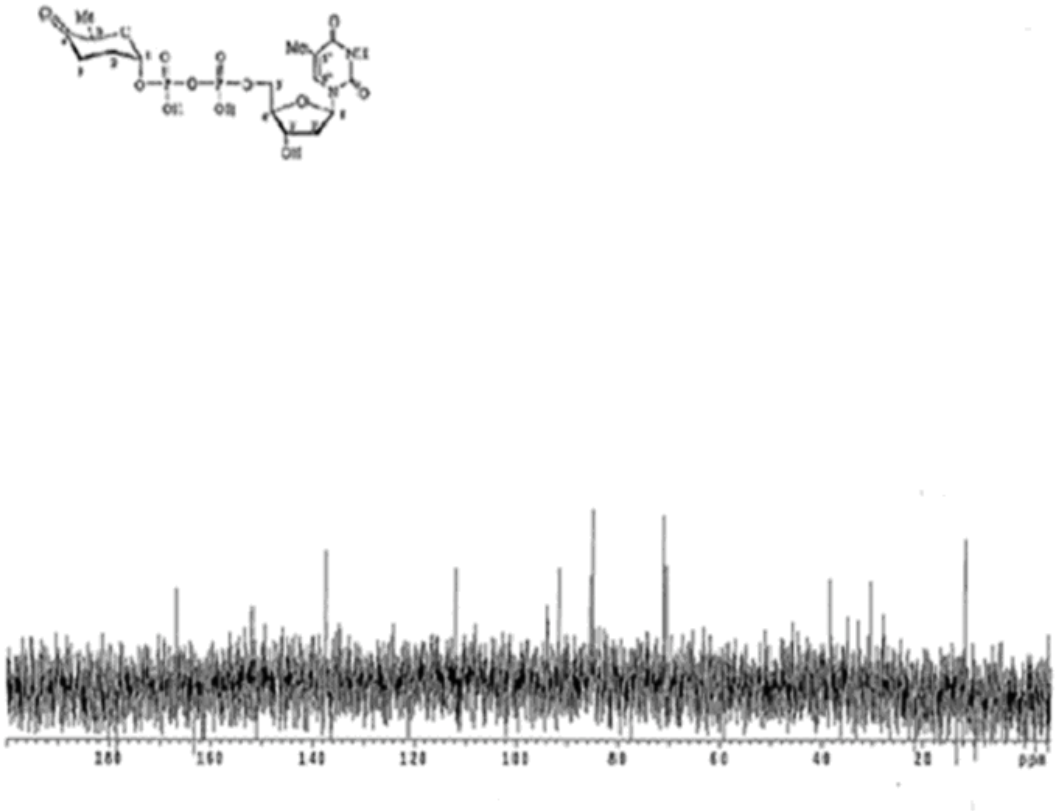

Figure S2. ^{13}C NMR spectrum of **7**

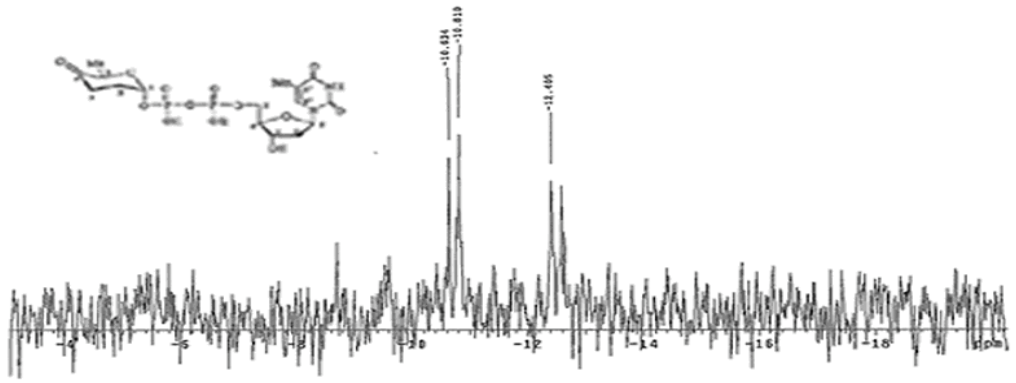

Figure S3. ^{31}P NMR spectrum of **7**

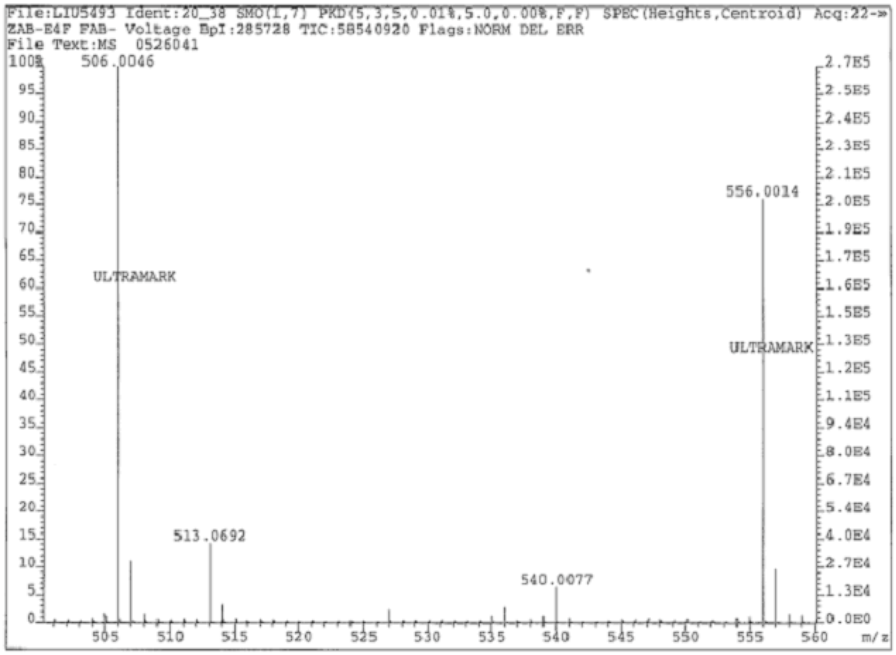

Figure S4. High resolution FAB-MS spectrum of **7**. Note that prominent 506.0046 and 556.0014 peaks are molecular weight standards. The 513.0692 peak is the signal from **7**.

8. Assay of E_1 under "ColD-like" Conditions

Preparation of E_1. E_1 was prepared according to previously published procedure.[7] The E_1 gene was originally cloned from *Y. pseudotuberculosis*.

*Preparation of E_1 substrate **10**.* **10** was prepared enzymatically according to previously published procedure.[8] Enzymes used for enzymatic generation of **10** were purified as previously described.[9,10]

Assaying E_1 under "ColD-like" Conditions. A typical assay mixture of 100 _ L contained 0.8 mM of compound **10**, 250 _ M PLP, 3.0 mM L-glutamate, 10 _ M E_1, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of the enzyme and was carried out at 37 °C. A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated and samples prepared for HPLC analysis as described above in Section 4. HPLC analysis was performed with the Dionex CarboPac PA1 analytical column using the program described above. This gave baseline separation of **10** (retention time = 32.1 min) and **12** (retention time = 35.2 min).

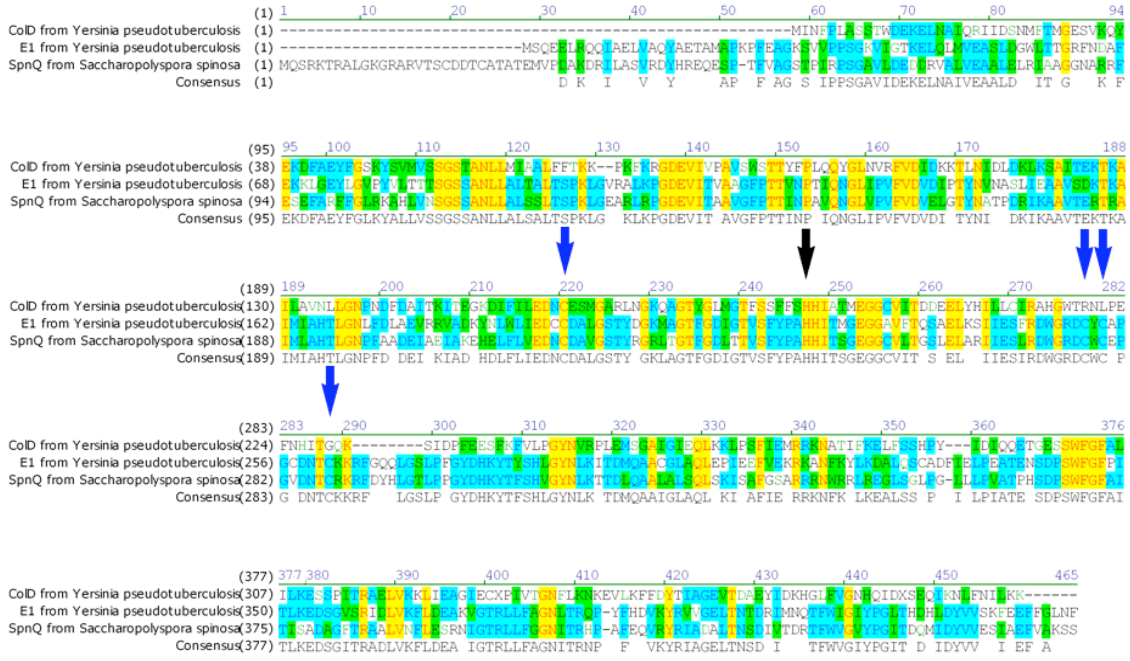

Figure S5 – Sequence alignment of ColD, E_1, and SpnQ. The conserved histidine which occurs in place of the conserved Schiff base-forming lysine in aminotransferases is marked with a black arrow, and the four cysteine residues thought to be involved in [2Fe-2S] cluster formation are marked with blue arrows. Three of these are not present in ColD, which lacks the [2Fe-2S] cluster.

References

1. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning, a laboratory manual*; 2nd ed.; Cold Spring Harbor Press: Cold Spring Harbor, NY, 1989.
2. Andrews, P. *Biochem. J.* **1964**, *91*, 222.
3. Barquera, B.; Nilges, M. J.; Morgan, J. E.; Ramirez-Silva, L.; Zhou, W.; Gennis, R. B. *Biochemistry* **2004**, *43*, 12322.
4. Cafaro, V.; Scognamiglio, R.; Viggiani, A.; Izzo, V.; Passaro, I.; Notomista, E.; Dal-Piaz, F.; Amoresano, A.; Casbarra, A.; Pucci, P.; Donato, A. D. *Eur. J. Biochem.* **2002**, *269*, 5689.
5. Fu, M.; Jack, R. F.; Morgan, T. V.; Dean, D. R.; Johnson, M. K. *Biochemistry* **1994**, *33*, 13455.
6. Ploux, O., Lei, Y., Vatanen, K., Liu, H.-w. *Biochemistry,* **1995**, *34*, 4159.
7. Lei, Y., Ploux, O., Liu, H.-w. *Biochemistry,* **1995,** 34, 4643.
8. Chen, X. M. H., Ploux, O., Liu, H.-w. *Biochemistry,* **1996**, 35, 16412.
9. Thorson, J. S., Lo, S. F., Ploux, O., He, X., Liu, H.-w. *J. Bact*. **1994**, 176, 5483.
10. Thorson, J. S., Kelly, T. M., Liu, H.-w. *J. Bact.*, **1994**, 176, 1840.