Functional Characterization and Sequence Analysis of Choline Dehydrogenase from *Escherichia coli*

Lawrance Anbu Rajan, Toms C Joseph^{*}, Nirmala Thampuran, Roswin James

Microbiology, Fermentation and Biotechnology Division, Central Institute of Fisheries Technology, Cochin 682 029, Kerala, India

> *Correspondence to: Toms C Joseph, tomscjoseph@gmail.com Accepted: September 26, 2010; Published online: October 28, 2010

Abstract

Choline dehydrogenase catalyzes the oxidation of choline to glycine betaine via betaine aldehyde as intermediate. Biotechnological applications of glycine betaine in transgenic plant improvement have been shown to have enhanced tolerance towards hypersalinity and freezing. It can also restore and maintain osmotic balance of living cells under stress. In this study, choline dehydrogenase (*betA*) gene from *E. coli* was cloned and homologously expressed in *E. coli* M15 (pREP4). The recombinant enzyme was purified by column chromatography using DEAE Sepharose. The purified enzyme showed a fourfold increase in the activity of choline dehydrogenase enzyme with choline as substrate and phenazine methosulfate as electron acceptor, compared to the control strain. The *betA* gene sequence reported in this study contains several base substitutions with that of reported sequences in GenBank, resulting in the altered amino acid sequences of the translated proteins.

Keywords: Osmolyte; betA; glycine betaine; choline dehydrogenase.

1. Introduction

Microorganisms are often exposed to changes in the solute concentrations in their environment [1]. Increase in the osmolarity results in the changes in intracellular solute concentration and cell volume. To cope with the environment stress, bacteria accumulate osmotically active solutes, including potassium, proline, glutamic acid, glutamine, a-aminobutyric acid, ectoine and betaine [2]. In addition they can accumulate α -glutamate and the β -amino acids; N-acetyl- β -lysine, β -glutamine, and β -glutamate as compatible solutes in response to high external NaCl levels [3]. Among the compatible solutes, Glycine-betaine is an effective osmoprotectant [4]. Glycinebetaine (N,N,N-trimethylglycine) is a quaternary ammonium compound that occurs naturally in a wide variety of plants, animals and microorganisms [5]. Numerous in vitro experiments have indicated that betaine acts as an osmoprotectant by stabilizing both the quaternary structure of proteins and membrane structures against the adverse effects of high salinity and extreme temperatures [6]. To date three enzymatic systems have been described for the formation of glycine betaine from choline. Biosynthesis of betaine is catalyzed in a single step reaction by choline oxidase in soil bacterium, Arthrobacter alobiformis [7]. In higher plants, such as spinach [8], sugar beet and amaranth [9] betaine biosynthesis is catalyzed by choline monooxygenase in combination with betaine aldehyde dehydrogenase. In E. coli, the biosynthetic pathway for the production of glycine betaine from choline has been well characterized at the genetic level [10]; it has been shown that four genes encoding choline dehydrogenase (betA), betaine aldehyde dehydrogenase (betB), a putative regulator (betI) and a choline transporter (betT) are clustered in the bet operon [11]. The enzymes involved in the biosynthesis of betaine have been cloned and characterized from bacteria and plants [12-14].

Recently, bacterial glycine betaine synthesizing enzymes have become a major target in developing stress tolerant crop plants of economic interest. Previous studies reports the resistance towards salinity and low temperature in transgenic tobacco expressing the two *E. coli* genes *betA* and *betB* [15], signifying the practical applications of choline dehydrogenase. Choline dehydrogenase (*betA*) of *E. coli* catalyses the first step in the synthesis of betaine, the oxidation of choline. However, this enzyme also catalyses the second step, the dehydrogenation of betaine aldehyde to betaine [10]. Choline dehydrogenase catalyzes the four-electron oxidation of choline to glycine betaine via a betaine aldehyde intermediate [16]. Only nominal studies on characterization of choline dehydrogenase from *E. coli* have been reported to date. In this study, we report the characterization and structural analysis of *betA* gene coding for the choline dehydrogenase in *E. coli*. Moreover, the sequence analysis of choline dehydrogenase from our isolate shows several base substitutions with that of reported sequences in GenBank, resulting in the altered amino acid sequences of the translated proteins.

http://astonjournals.com/gebj

2. Methods

2.1 Bacterial strains and plasmids

E. coli TCJAR023 was isolated from a finfish; Rohu (*Labeo rohita*) procured from local fish markets in Cochin, Kerala, India. Microbial identification and biochemical characterization of *E. coli* was carried out as per U.S. Food and Drug Administration (USFDA) and used as a source of the gene encoding the choline dehydrogenase (*betA*). *E. coli* JM109 was used for cloning and *E. coli* M15 (pREP4) was used as expression host. The plasmids used for cloning and expression was pDrive and pQE30 (Qiagen, Hilden, Germany) respectively.

2.2 DNA extraction and gene amplification

Genomic DNA extraction from *E. coli* TCJAR023 was performed following the method of Ausubel et al. [17]. Plasmid DNA was purified from the cells by alkali lysis method. Plasmids and PCR products were recovered from agarose gel using MinElute Gel purification Kit (Qiagen) following the manufacturer's instructions. A pair of primers bAF (5'-CGTATGCAATTTGACTACATCATT-3') and bAR (5'-GCATCATTTTTCGCTCTCACCG-3') were designed and used to amplify the *betA* gene. PCR amplification was done with 30 cycles of 1 min at 94°C, 1.5 min at 50°C and 2 min at 72°C. Additional extension was carried out for 5 min at 72°C using high fidelity PCR enzyme mix (MBI Fermentas, Maryland, USA).

2.3 Cloning and sequencing

The PCR product of *betA* gene was purified using Qiagen PCR purification Kit and cloned into pDrive according to the manufacturer's instructions. The pDrive-*betA* construct was transformed into *E. coli* JM109 (*recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17 (*rK-mk+*), e14– (*mcrA–*), *supE*44, *relA*1, Δ (*lac-proAB*)/F' [*traD*36, *proAB+*, *lac* Iq, *lacZ* Δ M15). White colonies were selected for PCR amplification with vector primers M13f-M13r (MBI Fermentas) and the clones with the correct insert as judged by size were sequenced on an ABI PRISM 377 genetic analyzer (Applied Biosystems, Perkin Elmer Co., Foster City, CA, USA).

2.4 Expression of recombinant choline dehydrogenase

The recombinant plasmid pDrive-*betA* construct was double digested with *Hind* III and *Bam* HI (MBI Fermentas) and purified by MinElute Gel Purification Kit. The purified *betA* gene was sub-cloned into pQE30 expression vector, which had previously been digested and purified. The resulting recombinant expression vector pQE30-*betA* cassette was transformed into *E. coli* M15 (pREP4). A single colony of the recombinant culture was inoculated into 5 ml of Luria–Bertani (LB) broth containing 100 µg/ml of ampicillin and 25 µg/ml of kanamycin, and incubated overnight at 37°C. 2·5 ml of the overnight culture was transferred into 50 ml of LB containing the corresponding antibiotics and incubated at 37°C, until OD_{600} value reached 0.6. IPTG (MBI Fermentas) was then added into the culture at the final concentration of 1 mM and was continuously incubated at 37°C for 4 h. The induced bacterial cells were harvested by centrifugation and resuspended in 1× SDS-PAGE sample buffer and lysed in boiling water bath for 3 min. The cells were centrifuged at 14 000 *g* for 5 min and the supernatant was checked for expression of soluble proteins. The expression of the target proteins were analyzed by SDS-PAGE. The SDS-PAGE was performed as described previously [18]. The molecular mass was estimated by SDS-PAGE with protein ladder (MBI Fermentas).

2.5 Cell extraction and enzyme purification

The IPTG induced *E. coli* M15 (pREP4) cells with pQE30-*betA* cassette were harvested by centrifugation at 10 000 g for 10 min and were resuspended with 20 mM Tris HCI (pH 7.5), 50 mM NaCI, 1 mM phenylmethylsulfonyl fluoride [PMSF] (Sigma-Aldrich Co., St. Louis, MO, USA) and 2 mM dithiothreitol [DTT] (Sigma). The cells were subjected to sonication for 3 min in 1 min pulses with discontinuous cooling in ice, using SonicMan multiwell sonicator (Matrical, Spokane, WA, USA). The cell debris was removed by centrifugation at 12 000 g at 4°C for 15 min.

The purification of choline dehydrogenase was performed as described previously [14] with slight modifications. Briefly, the supernatant was subjected to 50% ammonium sulfate saturation and incubated on ice for 20 min, and centrifuged at 10 000 g for 10 min. The resulting pellet was suspended in 2.5 ml of 0.5 mM EDTA, 15 mM potassium phosphate (pH 7.0) and centrifuged at 10 000 g for 10 min. Later, the supernatant was transferred into a HiPrep 16/10 DEAE Fast Flow column (Amersham Pharmacia Biotech Inc,

E-ISSN: 21503516

NJ, USA) equilibrated with 0.5 mM EDTA and 15 mM potassium phosphate (pH 7.0). Equilibrated buffer solution was also used as mobile phase at a flow rate of 1.0 ml/min. The elution was suspended in the mixture of 2 M sorbitol, 20% glycerol, 15 mM potassium phosphate (pH 7.0) and stored at -20°C.

2.6 Enzyme assay

The concentration of choline dehydrogenase was determined by the method of Bradford [19] by using boyine serum albumin as the standard. The oxidase activity of the enzyme was measured with 10 mM choline (Sigma) as the substrate in 50 mM potassium phosphate (pH 7.0), by monitoring the rate of oxygen consumption [14]. The dehydrogenase activity of the enzyme was measured with 1 mM phenazine methosulfate (Sigma) as the primary electron acceptor using a coupled assay in which the enzymatically reduced phenazine methosulfate is spontaneously reoxidized by molecular oxygen. One unit of enzymatic activity corresponds to the conversion of a micromole of oxygen per minute.

2.7 In silico sequence analysis

The nucleotide sequences obtained were compared against database sequences using BLAST provided by NCBI (http://www.ncbi.nlm.nih.gov) and were aligned and clustered using CLUSTAL-X version 1.81 program [20]. The output alignments were imported into the GeneDoc program (http://www.psc.edu/biomed/genedoc/) and BioEdit version 7.05 program (www.mbio.ncsu.edu/BioEdit/) to calculate the percent identities among the nucleotide and amino acid sequences. The molecular masses and the theoretical pl values of the polypeptides were predicted using the ProtParam tool (http://www.expasy.org/tools/protparam.html). The secondary structure prediction was performed using the PROTEAN program (DNASTAR, Inc., Madison, USA).

The nucleotide sequences generated in this study have been deposited in NCBI GenBank under accession number EU032535.

3. Results

3.1 Cloning and sequence analysis

The choline dehydrogenase gene, betA encodes the polypeptides comprised of 556 amino acids with the calculated molecular masses of 61 848 Da., based on the in silico estimates. The pl value of choline dehydrogenase is 5.45. Choline dehydrogenase of E. coli TCJAR023 is 1 671 bp long (Fig. 1). The betA gene from E. coli was PCR amplified, cloned in pDrive and transformed in E. coli JM109. Plasmid pDrive with betA gene insert was confirmed by nucleotide sequencing. The in silico sequence analysis of betA gene revealed a total of thirty-one base substitutions at the nucleotide level (Fig. 2) with that of the sequences deposited in the GenBank. (accession nos. X52905, M77738). However, only two of these changes translated into change of amino acid. The differences were observed at positions 133 and 452, which resulted in the amino acid substitution of valine with methionine and of threonine with valine residue, respectively (Fig. 3). No internal stop codons were observed due to the base substitution in the gene. Upon BLAST analysis it was found that the deduced amino acid sequence of betA was highly homologous to choline dehydrogenase of reported strains; 99% identity with E. coli (accession no. X52905) and 98% identity with E. coli (accession no. M77738). To authenticate the nucleotide variation in the sequence of betA, the gene was yet again PCR amplified with the genomic DNA from E. coli TCJAR023 with Tag DNA polymerase (Dynazyme II, Finland). The underlying principle for this strategy was that feasible misincorporations of nucleotides in gene amplification might occur at different positions by using different DNA polymerases and PCR protocols. The PCR product was cloned in pTZ57R/T (MBI Fermentas) and sequenced. The nucleotide sequence of the betA gene cloned in pTZ57R/T was the same as dogged using pDrive-betA cassette. This result confirms that the base divergence was conserved in the genome and not due to the external parameters.

4 Research Article



Fig. 1: Result from PCR with the primers bAF and bAR. Lane M: DNA molecular weight marker; Lane 1: Amplicon of the betA.

E-ISSN: 21503516

								o m i			
1032535	and Cont	DTTTCD CTD CI	DTCD TTD TTC	GTOCCOCT	Becceecabe	CTT CTC CCTD	CCCCTCTCB(TGBBGBTCC		ll erecrecien	
52905	T	ALLENCIAG		references i c	AGCCGGGAAG	GIIGICGCIA			MAIACCICC	61 661 661 661	LI GAN
77730	Τ										
		110	120	130	140	150	160	170	160	190	2
	1	1]]	المتعد إلمتعد			[]		<u> </u>]]	
U032535	CGGGC	GGCCCGGACTI	ATCGCTTTGA	CTTCCGCACC	CAGATGCCCG	CAGCGCTGGC	GTT CCCACTP	CAGGGCAAA	CGCTACAACT	gggc ctat gap	AACAG
52905						. <u>T</u> <u>C</u>	AG	T			G.
77738						.TC	AG	T			G.
		210	220	230	240	230	260	270	260	290	. 8
U032535	GCCTG	ACCGTTTAT	GAATAACCGO	CGCATGGAGI	GCGGACGCGG	TAAAGGCCTG	GGTGGATCG	CGCTGATCA	ACGGCATGTG	CTACATCCGT	GGCAR
52905	A			****		T					
77738	A					. T					
		310	320	330	340	3.50	360	370	390	390	4
	1	<u>ll</u>	I I		<u></u>		I I	· · · · I · · · ·	<u> </u>	<u> </u>	l
U032535	GCGCT	GGATCTCGAT	AACT GGGC GC	AAGAACCCGG	TCTGGAGAAC	TGGAGTTATC	T CGACT GC CI	GCCCTACTA	CCGCAAGGCC	GAGACTCGCG	ATGTO
32903 37738	• • • • •					u.u.	• • • • • • • • • • • •		• • • • • • • • • • •		A D
11150		410	420	420	440	450	460	470	480	490	· · •• · ·
	T	- 							100	1	í
0032535	GCCAG	AACGACTACC	ACGGTGGCGA	TGGCCCGGT	AGCGTCACCA	CCTCCAAACC	CGGCGTCAAI	CCGCTGTTT	AAGCGATGA	TTGAAGCGGGG	GTG
52905	. T A	 T .	C T		т.						
77738	.T.A	. T .	CT		т.		• • • • • • • • • • •				
		510	520	530	540	550	560	570	560	590	. 1
5002535 52005	GGGAG	GUTAG C C GC GI	LAG G GAG GAT	CTGAACGGCT	ATCAGCAGGA	NGGTTTTGGC	CGATGAAT(GGAGCGTGA	GGGGGGGGGGGG	CGGLCGCGCCF	AGIA
77738	G			Т.		т т					C.
		610	620	630	640	6.50	660	670	680	590	
	I	1]]	[]]]] [[]]	.	1	
032535	GCGCG	CGGTTAT CT C	GATCAGGCCA	AATCGCGCCC	TAACCTGACC	ATTCGTACTC	ACGCCATGA	CGATCACAT	ATTTTTGAC	GGCAAAC GCG(GGT
52905	5	Τ		T			T				
77738		Τ		T			T				• • • • •
	121	710	720	730	740	750	760	770	780	790	
1032838	0666666	GAAIGG LIGG			A G G G G G G G A A	GGGGGGAALAA	AGAAGIGUIU	TIAL GIGLA	3666668116	LEILALEGA	1416
52005											
52905 77738											
52905 77738		810	820	820	840	8.50	860	870	880	890	
52905 77738	I	810	820	820 	840	8 50	8 <i>5</i> 0	870 	880 	890 	
52905 77738 V032535	l	810 GCTCCGGCGT	820	830	840	850 TATTCCGCTG	850 GTGCAT GAA1	870 TACCCGGCG	880 [CGGCGAAAA	890 TCTTCAGGAT(L
52905 77738 V032535 52905	GCAAC	810 GCTCCGGCGT	820 ll CGGCAACGCT	830 GAACTGCTGG	840 CGGAGTTTGA	850 TATTCCGCTG	850 	870 TACCCGGCG:	880 . [CGGCGAARA	890 TCTTCAGGAT(CATC
52905 77738 V032535 52905 77738	GCAAC	810 GCTCCGGCGT	820 CGGCAACGCT	820 GAACTGCTGG	040 CGGAGTTTGA	850 TATTCCGCTG	060 GTGCAT GAAI	870 I TACCCGGCG	880 II. ICGGCGAAAA	890 . TCTTCAGGAT(CATC:
52905 77738 0032535 52905 77738	GCAAC	010 GCTCCGGCGT	820 CGGCAACGCT 920	630 GAACTGCTGG 930	040 CGGAGTTTGA 940	650 ll TATTCCGCTG 	850 I GTG CAT GAAI 950	670 TACCCGGCG 970	880 . CGGCGAAAA 980	890 - T CTT CAGGAT (990	ATC:
52905 77738 V032535 52905 77738	GCAAC	610 GCT CC GG CGT 910	820 CGGCAACGCT 920	630 GRACTGCTGC 930	040 	650 IATTCCGCTG 950	050 	670 TACCCGGCG 970	880 CGGCGRARA 980 .	890 TCTTCAGGATC 990 	
52905 77738 0032535 52905 77738 0032535	 GCAAC) GAGAT)	810 GCT CC GG CGT 910 GTATCT G CRA	020 	830 GAACTGCTGG 930 	840 CGGAGTTTGA 940 TTCCCTCTAC	850 IRTTCCGCTG 950 	050 GTGCAT GAA1 950 . A GT GGT GGAP	870 TACCCGGCG 970	880 II CGGCGARAR 980 II RATCGGTGCG	890 TCTT CRGGAT (990 GAGT GGCT GTT	 CATC
52905 77738 52905 77738 77738 0032535 52905	GCAAC)	810 GCT CC GG CGT 910 GTATCT G CAA:	820 	830 GAACTGCTGG 930 AAGAACCGGT	840 C GGAGTTT GA 940 TTC CCT CTAC	850 IRTTCCGCTG 950 	050 GTGCAT GAA1 950 . A GT GGT GGAP	870 TACCCGGCG 970 	880 II CGGCGARA 980 II RTCGGTGCG	890 TCTTCAGGAT (990 GAGT GGCT GTT	LATC:
52905 77738 52905 77738 77738 52905 52905 77738	GCAAC)	810 GCT CC GG CGT 910 GTATCT GCAA:	820 	820 	840 C GGAGTTT GA 940 TTC CCT CTAC	850 IATTCCGCTG 950 CCTGCCCTGC	050 GTGCAT GAAT 950 A GT GGT GCAP	870 TACCCGCCG 970 CCAGCCGAA	880 	890 TCTTCAGGAT(990 GAGTGGCTGT1	LATC:
52905 77738 52905 77738 77738 0032535 52905 77738	;GCRAC); ;GCRAC); ;GACAT ;	610 GCT CC GC GC GC GC 910 GTATCT GCAA 1010	020 	820 	040 CGGAGTTTGA 940 TTCCCTCTAC	850 IATTCCGCTG 950 CCTGCCCTGC	050 	670 TRCCCGGCG 970 CCAGCCGAA 1070	880 	890 T CTT CAGGAT (990 GAGT GGCT GTT 1090	LATC:
52905 77738 52905 77738 77738 0032535 52905 77738	 ;GCRAC; ;GAGAT; ;GAGAT;	610 	020 	630 	040 	0 50 IATTCCGCTG 950 ICTGCCCTGC 1050 ICTGCCGCCTGC	050 GTECAT GAA 950 	670 TACCCGGCG: 970 .CCAGCCGAAA 1070 GCGTGGCCG	000 CGGCGAAAA 900 AATCGGTGCG 1000 AATATT CAGT.	890 . TCTT CAGGAT (990 . GAGT GGCT GT1 1090 . ACCATTT CCT(LATC CATC
52905 77738 52905 77738 0032535 52905 77738 77738	 ;GCRAC; ;GAGAT; ;GCRCT;	010 	620 	820 GRACT GCT 66 920 AAGAACCGGT 1020	040 	850 INTTCCGTG 950 	050 GTGCAT GAA1 950 A GT GGT GCAF 1050 T GA GGA ATTT	670 TACCCGGCG 970 CCAGCCGAN 1070	000 900 PAT CGGT GC G 1000 PATATT CRGT.	890 TTT CRGGRT (990 GRGT GGCT GTT 1090 RCGRTTT CCT (LATC TTGG
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738	GCAAC	610 	820 	820 GRACT GCT 66 920 ARGAACC GGT 1020 CTTT GRAGCA	840 	850 PATTCCGCTG 950 	850 GTGCAT GRA1 950 A GT GGT GGA 1050 T GA GGAATTT	070 TACCCGGCG 970 CCRGCCGAN LO70 GCGTGGCCG	000 	890 990 . GAGT GCCT GT 1090 RCCATTT CCT G	LATC
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738	GCAAC	610 	020 	830 GRACT GCT GG 930 AARGAA CC GGT 1030 CTTT GAA GCA	040 	850 IATTCCGCTG 950 	050 GTGCAT GAA1 950 	670 TACCCGGCG 970 CCAGCCGAA 1070 GCGTGGCCGA	000 rc ggcgarra 980 . RT cget gcg 1080 . RT RT C GGT 1080	890 950 . GROT GGGT GT 1090 RCCRTTT CCT (1190	LATC
52905 77738 52905 77738 77738 52905 77738 52905 77738 52905 77738	GCAAC	010 GCT CCG GCGT 910 	020 	820 	040 	850 IATTCCGCTG 950 	050 GTGCAT GAA1 960 	070 TACCCGCC 970 CCAGCCGAA 1070 GCGTGGCCGA	000 	890 990 GRET GECT ETT 1090 RCCRTTT CCT (1190 	LATC TTGG
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535	GCAAC	010 	820 	820 GRACT GCT 66 920 1020 1020 	840 	850 INTTCCGCTG 950 1050 1050 1150 1150 1150	050 GTGCAT GAAI 950 A GT GGT GGT GGAF 1050 	670 TA CCC 66C 67 970 CCAGC C GAN 1070 GC GT 66 CC GA	000 	890 990 GAGT GGCT GTT 1090 RCCRTTT CCT 1190 GCRT GTG CGGF	LATC TTGG GCCA
52905 77738 52905 77738 77738 77738 77738 77738 77738 77738 77738	GCAAC)	610 910 910 1010 1010 1010 1110 1110 1110 1110	820 	820 GRACT GCT 6 920 LO20 LO20 CTTT GRAGGE LL20 LL20 GCACT GRAGG	040 	850 1ATTCCGCTG 950 1050 1050 1150 1150 1250 1250 1250	050 	970 TACCCGGCGAA 970 CCAGCCGAA 1070 GCGTGGCCGA	000 	890 990 GRAT GGGT GT 1090 RCGRTTT CCT (1190 . GGRT GTGC GGG	I TTGG GCCR
52905 77738 52905 77738 77738 52905 77738 52905 77738 52905 77738	GCARCI	010 	020 	830 GRACT GCT 6G 930 AAGAA CC GGT 1030 	040 	050 	050 	670 	000 	890 	LATC
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738	GCACT	010 	020 	830 GRACT GCTGG 930 ARGAACCGGT 1030 	040 	050 TATTCCGCTG 950 1050 1050 1150 1150 1250 1250	050 GTGCAT GAA1 960 	870 	000 	890 	 LATC TT GG GC CA
52905 52905 52905 77736 U032535 52905 77736 U032535 52905 77736 U032535 52905 77736 U032535 52905	GCARC	010 	820 	820 GRACT GCT GC 920 	840 	050 INTTCCGCTG 950 1050 1050 1150 1150 1150 1250 1250 1250	050 	070 TACCCGGCG 970 CCAGCCGAA 1070 	000 	090 	 IATC' I I GCCA
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738	GCAAC	610 	020 	830 GRACT GCT GG 930 L030 L030 CTTT GARGCR 1130 GCRGT GARAG 1230 L230 L230	040 	850 1ATTCCGCTG 950 1.11 CCTGCCCTGC 1.050 1.1.	050 	870 	000 	090 990 GAGT GEGT GT 1090 RCCATTT CCT (1190 GCAT GTGC GGF 1290 	 IATC'
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738	GCAAC	010 910 910 1010 1010 1110 1110 1210 1210 1210 1210	020 	830 GRACT GCT GG 930 ARGAACCGGT 1030 	040 	0 50 	050 GTGCAT GAAT 950 1. AGT GGT GGA 1.050 1.050 1.150 1.150 GTC GGCT CAP 1250 1.250 	870 TACCCGGCG 970 CCAGCCGAA 1070 GCGTGGCCGA 1170 TGCGCTGGCCG 1270 	000 	890 	 CATC: TTGG: GCCA:
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 77738	GCACT	010 	820 	820 GRACT GCT GC 930 	040 	0 50 TATTCCGCTG 950 1050 1050 1150 1150 1250 1250	050 GTGCAT GAA1 960 	670 	000 	890 	LATC
52905 52905 52905 77736 9032535 52905 77736 9032535 52905 77736 9032535 52905 77736	GCAACI GGAGAT GGAGAT GGAGAT GGAGAT GGAGAT GGAGAT GGAGAT GGAGAT GGAGAT	010 910 910 1010 1110 1110 1110 1210 1210 1210 1210 1210	820 	820 GRACT GCT GG 920 	040 	850 	050 	070 	000 	090 990 GAGT GCCT GT 1090 ACCATTT CCT 1190 . GCAT GTC CGAT 1290 CGCAT CA CCCC 1290 	LATC
52905 52905 52905 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738 77738	GCAACI GCAACI GCACT GCCACT GCCACT INGCCAC	010 	020 	830 GRACT GCT GG 930 L030 L030 CTTT GRAGGA 1130 GCRGT GRAGGA 1230 L230 L230 L230 L230 L230 L230 L230 L	040 	850 1ATTCCGCTG 950 1.11 CCTGCCCCGCC 1.050 1.1.50 1.1.50 1.1.50 1.1.50 1.2.50 1.500 1.2.50 1.2.	050 	870 	000 	890 	L
52905 77736 52905 77736 77736 77736 77736 77736 77736 77736 77736 77736 77736 77736 77736 77736	GCAAC	010 	020 	830 GRACT GCT GG 930 	040 	0 50 1ATTCCGCTG 950 	050 GTGCAT GAAT 950 1050 1050 1050 1050 1250 1	870 	000 	890 	I FT GG GCCA
52905 52905 52905 77736 9032535 52905 77736 9032535 52905 77736 9032535 52905 77736 9032535 52905 77736	GCACT) GCACT) GCACT) GCACT) GCACT) GCACT) GCACT) GCACT) GCACT GCACT GCACT GCACT GCACT GCACT GCACT	010 	020 	830 GRACT GCT GC 930 	040 	850 1ATTCCGCTG 950 1050 1050 1150 1150 1250 1250 1250 1250 1250 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1350 17CGCACGAGCAC 1550 17CGCACGAGCAC 17CGCACGAGCAGCAC 17CGCACGAGCACGAGCAC 17CGCACGAGCAC 17CGCACGAGC	060 	070 TACCCGCCGAN 970 CCAGCCGAN 1070 	000 	090 	LATC: TTGG GCCA
52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738	GCAACT GGAGAT GGAG G	010 910 910 1010 1110 1110 1110 121	020 	830 GRACT GCT GG 930 ARGAA CC GGT 1030 	040 	850 1ATTCCGCTG 950 1.11 CCTGCCCCCCC 1050 1.1.50 1.1.50 1.1.50 1.1.50 1.1.50 1.2.	050 	670 	000 	890 	I FT GG GC CR I AT TRA GC CR I FR AC
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738	GCAAC	010 910 910 1010 1010 1010 1110 120	020 	830 GRACT GCT GG 930 AAGGA CC GGT 1030 	040 	050 	050 	070 	000 	090 	LATC: TTGG GCCA
52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738	GCACT GCACT GCACT GCACT GCACT GCACT GCACT GCACT GCACT GCACT GCACCA GCACCACCA GCACCA GCACCA GCACCA GCACCACCA GCACCA GCACCA GCACCACCA GCACCA GCACCA GCACCA G	010 	020 	830 GRACT GCT GC 930 	040 	0 50 1ATTCCGCTG 950 1050 1050 1150 1150 1250 1250 1250 1250 1250 1250 1250 1250 1250 1250 1250 1450 1450 1450 1450	050 	670 TACCC 66C 67 970 .CCAGC C GAR 1070 .CCAGC C GAR 1070 .CCGC C GCC 4 	000 	090 	I TTGG GCCA ATTAN GCCA
52905 77736 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77738 0032535 52905 77738	GCACT GCACT	010 910 910 1010 1010 1110 1110 1210 1210 1210 1210 1410 1410 1410 1410	820 	820 GRACT GCT GC 920 	040 	850 	050 	070 	000 	090 	LATC'
52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738	GCAACI GCAGAT GGAGAT GGCACT GGCACAT GGCACAT GGCACAT GGCACAT GGCACAT GGCACAT GGCACAT GGCACAT GGCCCGAT GGCCCCAT	010 	020 	830 GRACT GCT GG 930 1030 1030 1130 CTTT GRAGCA 1230 1230 1230 1230 1230 1430 1430 1430 1430 1430 1530	040 	850 1ATTCCGCTG 950 1.11 CCTGCCCCCCC 1.050 1.1.50 1.1.50 1.1.50 1.1.50 1.2.50 1.2.50 1.2.50 1.2.50	050 	070 	000 	090 	I
52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738	GCAACT GCACT	010 	020 	820 GRACT GCT GG 930 1030 	040 	0 50 1 TATTCCGCTG 950 1050 1050 1150 1150 1250 1250 1250 1250 1250 1250 1250 1250 1250 1450 1450 1450 1550 1550	050 	070 	000 	090 	LATC TTGG GCCA
52905 77736 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 77736 0032535 77736 0032535 77736	GCACT) GCACT GCACT) GCACT GCA	010 	020 020 920 1020 1020 1120 1120 1220 1220 1220 1220 130	820 GRACT GCT GC 930 	040 	850 	050 	670 	000 	090 	I FT GG GC CA I GC CA I
52905 77736 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736 0032535 52905 77736	GCACA GCACA	010 	020 	830 GRACT GCT GG 930 1030 1030 1130 CTTT GARGCA 1230 1230 1230 1230 1230 1230 1430 1430 1430 1430 1430 1430 1530 1530	040 	850 1ATTCCGCTG 950 1LCTGCCCTGC 1050 1150 1150 1150 1250 1250 1250 1250 1250 1250 1450 1550 1550 1550	050 	070 	000 	090 	LATC TTGG GCCA
52905 77738 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738 0032535 52905 77738		010 	020 	830 GRACT GCT GG 930 	040 	0 50 1 TATTCCGCTG 9 50 1 CCT GCCCTGC 1 CCT GCCCTGC 1 CCT GCCCTGC 1 L50 1 L50 1 L50 1 L50 1 L50 1 L50 1 L50 1 L50 1 L50 1 CCRGTGCCAC 1 250 1 L50 1 CCRGTGCCAC 1 CCRGTG	050 GTGCAT GAA1 950 1050 1050 1050 1160 1160 1250 1250 1350 1450 1450 1550 1550 1550 1550	070 	000 	090 	LATC TTGG
52905 77738 52905 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738 52905 77738	GCAACT GGCCCC GGCCC GGCCC GGCC GGCCC GGCC GGCC GGCACT	010 	020 	830 GRACT GCT GG 930 1030 	040 	850 550 1ATTCCGCTG 950 1050 1050 1150 1150 1150 1250 1250 1250 1250 1250 1250 1250 1250 1250 1250 1550 1550 1650	060 	070 	000 	090 	LATC TT GG GC GA L GC GA L L GC GA L L GC GA L L GC GA L GC GC C GC GA L GC GC GA L GC GC C GC GA L GC GC C GC GA L GC GC C GC GA L GC GC GA L GC GC GC C GC GC GA L GC GC GC GC GA L GC GC GC GA L GC GC G

Fig. 2: BioEdit analysis of nucleotide substitutions of choline dehydrogenase (*betA*) in *E. coli* TCJAR023 [EU032535], with *E. coli* strains [X52905 & M77738].

http://astonjournals.com/gebj





3.2 Protein expression and enzyme purification

The expression vector pQE30-*betA* gene cassette, downstream of the inducible promoter was transformed in *E. coli* M15 (pREP4). To validate the expression, the *betA* gene cassette was induced with 1.0 mM IPTG. The proteins of the induced and non-induced *E. coli* M15 (pREP4) cells were visualized by SDS-PAGE. The lysate of induced cells showed one clear expressed band with molecular masses of 63 kDa, which correspond to choline dehydrogenase (Fig. 4). The expressed choline dehydrogenase enzyme was then subjected to 50% ammonium sulfate saturation and the purification was performed using Hi-Prep 16/10 DEAE Fast Flow column (Table 1). The resultant purified enzyme was stabilized with 2 M sorbitol and 20% glycerol, at pH 7.0 [14]. The purified protein was visualized by SDS-PAGE (Fig. 4).



Fig. 4: SDS-PAGE analyses of the expressed and purified choline dehydrogenase from *E. coli*. The arrow (▶) indicates the position of the expressed *betA*. Lane 1: Total protein of the uninduced pQE30-*betA* gene cassette; Lane 2: Total protein of the induced pQE30-*betA* gene cassette; Lane M: Protein molecular mass markers; Lane 3: Purified choline dehydrogenase

Stage of cells	Total units (µmol of O₂/min)	Total protein (mg)	Specific activity (µmol of O₂/min)	
Cell free extract	17	31	0.2	
50% (NH ₄) ₂ SO ₄ saturation	12	18	0.4	
Hi-Prep 16/10 DEAE	16	7	1.1	

Table 1. Purification of recombinant choline dehydrogenase.

3.3 Functional analysis of choline dehydrogenase

Choline dehydrogenase activity was determined by measuring the rate of reaction with choline as the substrate and 1 mM phenazine methosulfate as the primary electron acceptor in the assay reaction mixture. The activity of choline dehydrogenase with 10 mM choline was 8.8 μ mol of O₂/min/mg of cultured cells. The declined activity of the enzyme was observed with the decrease in the concentration of substrate. In the presence of 2.5 mM choline, the specific activity was lowered to 3 μ mol of O₂/min/mg of cultured cells.

3.4 Secondary structure prediction analysis

The secondary structure of choline dehydrogenase protein was predicted to have the alpha-helical structure with maximum hydrophilic molecules. The prediction analysis also revealed the presence of much acidic amino acids, regions with high antigenicity and very high backbone chain flexibility (Fig. 5). Upon analysis of *betA* protein, the predicted charge at pH 7.0 was 15.45 with the isoelectric point of 5.52. Common amino acids include 58% glycine, 45% alanine, 41% glutamic acid, 40% leucine, 37% arginine, 33% praline and 30% each of isoleucine and aspartic acid.





Fig. 6: Secondary structure analysis using PROTEAN. The analysis was performed using choline dehydrogenase from *E. coli* TCJAR023 [EU032535], with *E. coli* strains [X52905 & M77738].

4. Discussion

Based on the sequence analysis, it was previously reported that the *betA* gene of *E. coli* codes for choline dehydrogenase [13]. So far, only minimum reports on the characterization of choline dehydrogenase have been reported [14, 16]. In this study, we cloned and characterized the betA gene from E. coli cells under the control of inducible promoter. Evaluation of the deduced amino acid sequence of betA gene with reported choline dehydrogenase sequences in the database revealed a maximum similarity. However, the sequence analysis of betA of our isolate showed several base substitutions with that of reported sequences, resulting in the altered amino acid sequences of the translated protein. The activity of the enzyme using choline as the substrate and phenazine methosulfate as the electron acceptor revealed the expressed enzyme from betA is choline dehydrogenase. The ability of the choline dehydrogenase to catalyze both oxidative reactions in the conversion of choline to glycine betaine has been previously reported in choline dehydrogenase from Halomonas elongata [14], rat liver mitochondria [16] and in choline oxidase from Arthrobacter alobiformis [7]. Choline dehydrogenase is capable of using either choline or betaine aldehyde as a substrate [16]. In this study, we used choline as the substrate and the enzyme activity rate was measured by the rate of oxygen consumption in the presence of 1 mM phenazine methosulfate as the primary electron acceptor. The enzyme activity of choline dehydrogenase from recombinant E. coli with pQE30-betA cassette showed a fourfold-enhanced activity relative to the control strain. Moreover, the catalytic rate of the enzyme was in the declined manner with that of the concentration of substrate by 10 mM to 2.5 mM. These observations signify that the enzyme activity is maximum, only in the presence of appropriate concentration of the substrate. These results are consistent with previously reported data [14] that, the specific activity of the enzyme was 11.8 µmol of O₂/min/mg with 10 mM choline and 8 µmol of O₂/min/mg with 5 mM choline. The secondary structure prediction results also revealed considerable similarity with the reported choline dehydrogenases from E. coli strains. Upon structural analysis, both Chou-Fasman and Garnier-Robson prediction of alpha helix structures showed a maximum similarity, suggesting that the active domains of the enzyme from E. coli TCJAR023 have the considerable identity with the database reports (Fig. 6).

In conclusion, based on the expression and enzymatic analysis, the enzyme encoded by *betA* in *E. coli* TCJAR023 is a choline dehydrogenase. To our knowledge, this study represents the first instance in which a choline dehydrogenase from *E. coli* has been cloned and characterized in detail. Moreover, the determination of protein structure modification due to the nucleotide substitutions will certainly provide the basis for performing site-directed mutagenesis to improve the production and configuration of the osmolytes of biotechnological interest. This, in turn, has great potential for biotechnological applications aimed at genetically engineering stress tolerance in crop plants of economic interest.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

The research concept was developed by LAR and TCJ; research experiments were performed by LAR and RJ; NT organized for and provided the necessary facilities.

Acknowledgement

Authors are grateful to the Director, Central Institute of Fisheries Technology (CIFT), Cochin for providing the necessary facilities to carry out this research work and Indian Council for Agricultural Research, New Delhi, India for financial assistance.

References

- [1] Poolman B, Glaasker E, 1998. Regulation of compatible solute accumulation in bacteria. Molecular Microbiology, 29: 397-407.
- [2] Galinski EA, Truper HG, 1994. Microbial behavior in salt stress ecosystems. FEMS Microbiology Review, 15: 95-108.
- [3] Sowers KR, Gunsalus RP, 1995. Halotolerance in *Methanosarcina* spp.: role of N\$-acetyl-β-lysine, α-glutamate, glycine betaine, and K⁺ as compatible solutes for osmotic adaptation. Applied and Environmental Microbiology, 61: 4382-4388.
- [4] Wood JM, Bremer E, Csonka LN, Kramer R, 2001. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. Comparative Biochemistry and Physiology, 130: 437-460.
- [5] Rhodes D, Hanson AD, 1993. Quaternary ammonium and tertiary sulphonium compounds in higher plants. Annual Review of Plant Physiology and Plant Molecular Biology, 44: 357-384.
- [6] Gorham J, 1995. Betaines in higher plants biosynthesis and role in stress metabolism, in amino acids and their derivatives in higher plants, In: Wallsgrove, R.M. (Ed.), Cambridge University Press, Cambridge, pp. 171–203.
- [7] Ikuta S, Mamura S, Misaki H, Horiuti Y, 1977. Purification and characterization of choline oxidase from *Arthrobacter globiformis*. Journal of Biochemistry, 82: 1741-1749.
- [8] Rathinasabapathi B, Burnet M, Russel BL, Gage DA, 1997. Choline monooxygenase, an unusual iron-sulphur enzyme catalyzing the first step of glycine betaine synthesis in plants: prosthetic group characterization and cDNA cloning. Proceedings of the National Academy of Sciences USA, 94: 3454-3458.
- [9] Russell BL, Rathinasabapathi B, Hanson AD, 1998. Osmotic stress induces expression of choline monooxygenase in sugar beet and amaranth. Plant Physiology, 116: 859-865.
- [10] Landfald B, Strom AR, 1986. Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. Journal of Bacteriology, 165: 849-855.
- [11] Andresen PA, Kaasen I, Styrvold OB, Boulnois G, Strom AR, 1988. Molecular cloning, physical mapping and expression of the *bet* genes governing the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. Journal of General Microbiology, 134: 1737-1746.
- [12] Weretilnyk EA, Hanson AD, 1990. Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. Proceedings of the National Academy of Sciences USA, 87: 2745-2749.
- [13] Lamark T, Kaasen I, Eshoo MW, Falkenberg P, 1991. DNA sequence and analysis of the bet genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. Molecular Microbiology, 5: 1049-1064.
- [14] Gadda G, Wilkins EEM, 2003. Cloning, expression, and purification of choline dehydrogenase from the moderate halophile *Halomonas elongata*. Applied and Environmental Microbiology, 69: 2126-2132.
- [15] Holmstrom KO, Somersalo S, Mandal A, Palva TE, Welin B, 2000. Improved tolerance to salinity and low temperature in transgenic tobacco producing glycine betaine. Journal of Experimental Botany, 51: 177-185.
- [16] Tsuge H, Nakano Y, Onishi H, Futamura Y, Ohashi K, 1980. A novel purification and some properties of rat liver mitochondrial choline dehydrogenase. Biochimica et Biophysica Acta, 614: 274-284.
- [17] Ausubel FM, Brent R, Kingston RE, Moore DD, 1994. *Current Protocols in Molecular Biology*. John Wiley and Sons, New York: pp. 2.0.1–2.14.8.

http://astonjournals.com/gebj

0 Research Article

- [18] Laemmli UK, 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227: 680-685.
- [19] Bradford MM, 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry, 72: 248-254.
- [20] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 25: 4876-4882.

E-ISSN: 21503516