Insight Bioinformatics 2 (1): 1-8, 2012
ISSN 2042-3233 / DOI: 10.5567/BIOINFO-IK.2012.1.8

In silico Analyses of Flavin Reductase from *Citrobacter freundii* A1

[1]Giek Far Chan, [2]Abdull Rahim Mohd Yusoff and [1]Noor Aini Abdul Rashid
[1]Nanoporous Materials for Biological Application Research Group (NAMBAR),
Sustainability Research Alliance, Universiti Teknologi Malaysia,
81310 UTM Johor Bahru, Johor, Malaysia
[2]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia,
81310 UTM Johor Bahru, Johor, Malaysia

Abstract: Objective: The aim of the present study was to describe the structural and phylogenetic features of flavin reductase from *C. freundii* A1 using bioinformatics tools. **Materials and methods:** The flavin reductase (*fre*) gene from a dye-degrading bacterium, *C. freundii* A1 was isolated and amplified by Polymerase Chain Reaction (PCR). The gene encodes for NAD(P)H:flavin oxidoreductase, an enzyme that catalyzes the reduction of soluble flavins by reduced pyridine nucleotides this was believed to be the azoreductase from *C. freundii* A1. The gene, approximately 0.8 kb was sequenced and *in silico* analyses of the nucleotide sequence were performed. **Results:** From phlyogenetic analyses, we observed that flavin reductase enzyme generally existed in most microorganisms and the enzyme from *C. freundii* A1 is conserved among the Gram-negative bacteria. The protein function of the flavin reductase coded was predicted based on the motifs of deduced amino acid sequence. The amino acid sequence of flavin reductase from *C. freundii* A1 was compared with other azoreductases and was found to be a unique NADPH-preferred azoreductase. **Conclusion:** Hence, *in silico* characterization of flavin reductase gene presented various features of the gene and this would facilitate molecular studies of the gene and reveal the functional role of the enzyme.

Key words: *Citrobacter freundii*, NAD(P)H:flavin oxidoreductase, azoreductase, decolourisation

INTRODUCTION

In the 21st century, bioinformatics has emerged as an interdisciplinary field that bridges biology, mathematics and computer science in response to major advances in molecular biology technologies (Tang, 2002). An enormous amount of valuable information for wet-bench research work could be obtained via database searches, molecular predictions and phylogenetic analyses. On the other hand, extensive studies on microbial decolourisation of azo dyes have been published since the last three decades. Most suggested the association of azoreductase, the enzyme that reductively cleave the azo bond (-N = N-) of azo dyes in order to decolourise under anaerobic conditions (Saratale *et al*., 2011). Bacterial strains were hypothesized to possess unspecific cytoplasmic enzymes which act as "azo reductases" to transfer electrons via soluble flavins to azo dyes. Russ *et al*. (2000) recognized that flavin reductases as the cytoplasmic anaerobic "azo reductases" which showed significant importance in the reduction of sulphonated azo compounds. Hence, the relevance of this hypothesis whether flavin reductase is the azoreductase of *C. freundii* A1 was proposed. *C. freundii* A1, an enteric bacterium, was isolated and screened for its potential in azo dye decolourisation (Rashid *et al*., 1999). Our ultimate goal is to determine the involvement of flavin reductase in decolourisation of azo dyes. Prior to cloning and expression studies, the gene was isolated and charaterised. This is for the first time that a flavin reductase gene was isolated from *Citrobacter* sp. Therefore, in this study, *in silico* characterization of flavin reductase gene from *C. freundii* A1 was carried out in order to gain further insight into the gene and the enzyme's biological functions.

METERIALS AND METHODS

PCR analysis: *C. freundii* A1 was grown in nutrient broth (Difco). The culture was incubated at 37°C in an orbital shaker (200 rpm) for 18-24 h. Genomic DNA was then extracted using Wizard Genomic DNA Purification Kit

Corresponding Author: Noor Aini Abdul Rashid, Nanoporous Materials for Biological Application Research Group (NAMBAR), Sustainability Research Alliance, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia Tel: +607-5534495 Fax: +607-5531112

(Promega) and was then used as template for PCR amplification. The forward primer, FREf (5'-GCG CAT ATT GAC GCC ATC TGG GA) corresponded to nucleotide positions 16 to 38 of *E. coli fre* gene (Accession No.: M61182). The reverse primer FREr (5'-GAT AAA TGC AAA CGC ATC GCC AA) was designed and corresponded to the complement of nucleotide positions 819 to 797 in the *E. coli* sequence (Spyrou *et al*., 1991). The expected 0.8 kb fragment was observed on 1% agarose gel, extracted and sequenced.

The *fre* gene sequence was compared with known DNA sequences in GenBank database at NCBI (http://www.ncbi.nlm.nih.gov/blast) using BLASTn and BLASTx (Altschul *et al*., 1997). Restriction sites analysis of the *fre* gene fragment was carried out using NEBcutter (Version 2.0) at http://tools.neb.com/NEBcutter2/ (Vincze *et al*., 2003). Online WWW Promoter Scan software (http://www-bimas.cit.nih.gov/cgi-bin/molbio/proscan) was used to analyze the upstream region of *fre* gene (Prestridge, 1995). Proteomic analyses of flavin reductase amino acid sequence were further carried out online at http://www.expasy.org/resources, using Compute pI/Mw for isoelectric point and molecular weight prediction (Gasteiger *et al*., 2005) and SWISS-MODEL for protein structure homology modeling (Schwede *et al*., 2003). Other downloadable bioinformatic softwares, e.g., TM calculator 2-beta, DNAClub, GeneDoc, RasMol (Version 2.7.5) and Swiss-Pdb Viewer 3.7 were used in molecular and structural analysis. Multiple sequence alignment with related amino acid sequences of flavin reductases and azoreductases were carried out using ClustalW. MEGA version 4.1 (Beta 3) was used for construction of Neighbor-Joining phylogenetic trees with bootstrap values calculated based on 1000 replicates (Tamura *et al*., 2007).

Microbial sources of azoreductases: The microbial sources of azoreductases are as follows: AcpD (YP_252301)- *Staphylococcus haemolyticus* JCSC1435; AcpD (YP_039668)- *Staphylococcus aureus* MRSA252; YvaB (NP_391234)- *Bacillus subtilis* strain 168; YvaB (YP_001422634)- *Bacillus amyloliquefaciens* FZB42; AzrA (BAF02597)-*Bacillus* sp. B29; LMHCC_1847 (YP_002350802)- *Listeria monocytogenes* HCC23; AzoA (AAR38851)- *Enterococcus faecalis*; AzoR (A4W2Z7)- *Streptococcus suis* 98HAH33; AzoR2 (Q9CIH9)- *Lactococcus lactis*; AzoR (Q8X9S9)- *Escherichia coli* O157:H7; AzoR (1V4B_A)- *Escherichia coli*; AzoR (AAG04174)- *Pseudomonas aeruginosa* PAO1; Azr (BAB13746)_*Bacillus* sp. OY1-2; Azr (ANN17400)- *Rhodobacter sphaeroides*; Azr (BAB85976)- *Bacillus subtilis* ISW1214; Azr (BAB85975)- *Geobacillus stearothermophilus* IFO13737; L8106_10307 (ZP_01622489)- *Lyngbya* sp. PCC 8106; Y412MC10 (ZP_03036583)- *Geobacillus* sp. Y412MC10; AZR (ACF54629)- *Staphylococcus cohnii* AZR; SERP0206 (YP_187802)- *Staphylococcus epidermidis* RP62A; Azo1 (AAT29034)- *Staphylococcus aureus* ATCC 25923; Azo1 (Q4L3N6)- *Staphylococcus haemolyticus* JCSC1435; Fre (AAO91775)- *Citrobacter freundii* A1; AzoB (AAM92125)- *Xenophilus azovorans* KF46F; EU307209 (ACA34616)- *Erwinia chrysanthemi*; AzoB (ADD80733)- *Pigmentiphaga kullae* K24; RSMK02502 (YP_002254667)- *Ralstonia solanacearum* MolK2; H16_A2352 (YP_726815)- *Ralstonia eutropha* H16; RALTA_A1896 (YP_002005902)- *Cupriavidus taiwanensis* strain LMG 19424.

RESULTS AND DISCUSSION

Analysis of *fre* gene: Amplification using FREf and FREr primers resulted in the isolation of an approximate 0.8 kb *fre* fragment, as shown at lane 1 in Fig. 1. The 0.8 kp gene sequence (Fig. 2) is available in GenBank database under accession number AY163084.1. Figure 2 shows that the region contains an open reading frame with the first in-frame start codon ATG, located at nucleotide position 99. The open reading frame shown in Fig. 3 was assigned to the enzyme based on the deduced amino acid sequence corresponding to the N-terminal amino acid determined for flavin reductase of *E. coli* reported by Spyrou *et al*. (1991). This gives rise to a polypeptide

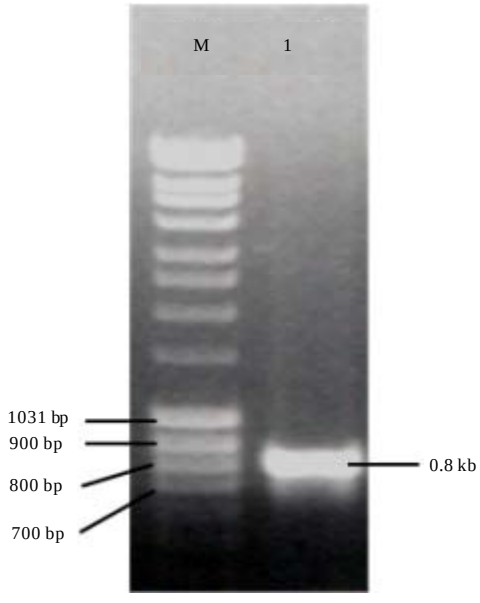

Fig. 1: PCR amplification of flavin reductase gene from *C. freundii* Al. M: MassRler™ DNA Ladder, Mix (Fermentas)

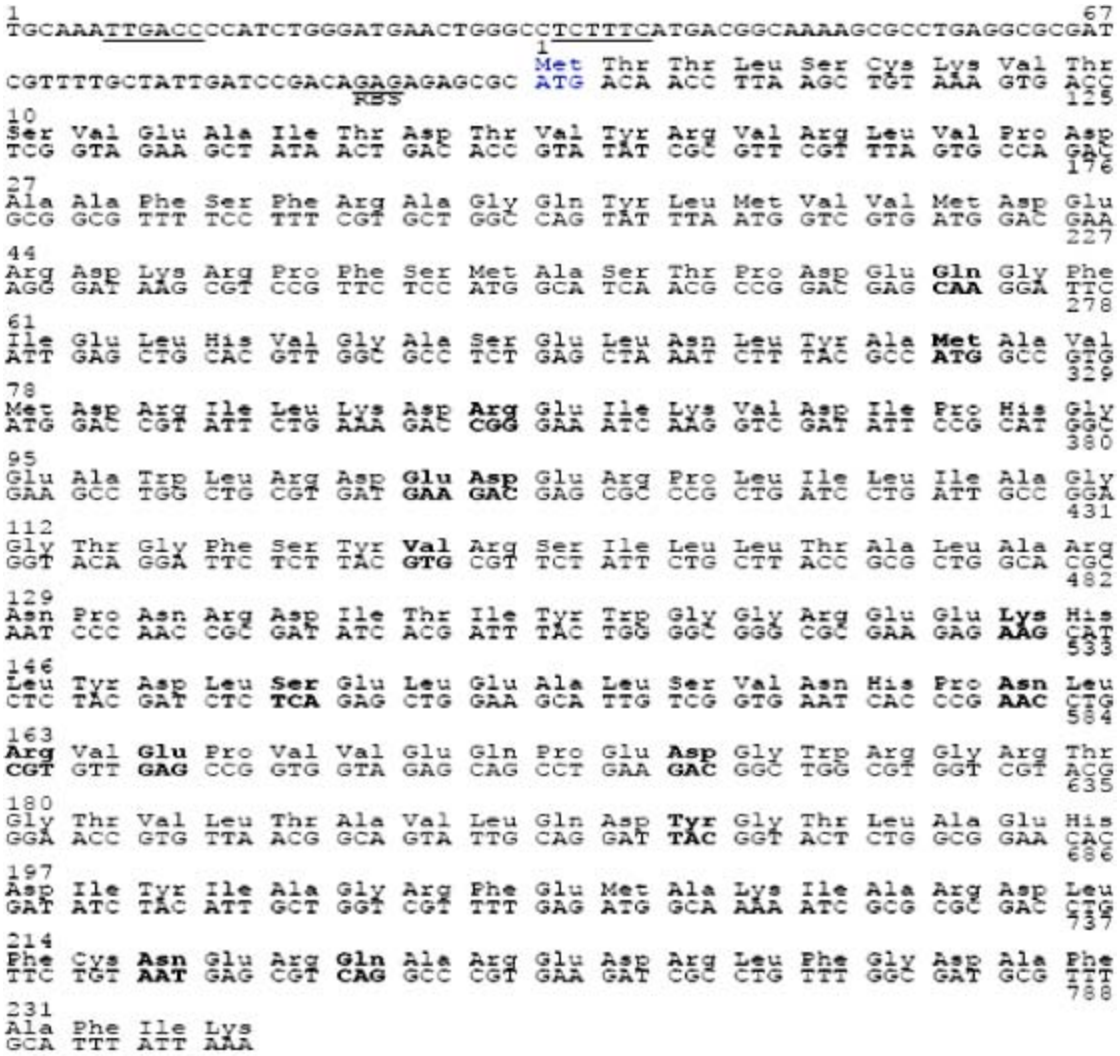

Fig. 2: The flavin reductase gene of *C. freundii* A1. The codons (in bold) that correspond to the appropriate amino acids which are different from flavin reductase of *E. coli*.The GenBank accession number for this sequence is AY163084

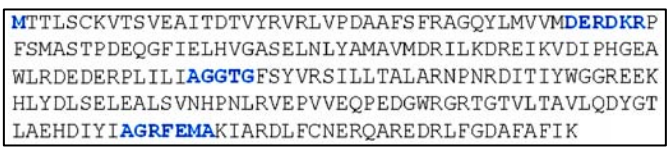

Fig. 3: The deduced amino acid sequence from the (*fre*) gene region of *C. freundii* A1. M: First amino acid, methionine, AGGTG: A pyridine nucleotide binding site motif, AGRFEMA: The flavin binding site, DERDKR: May code for surface residues

of 234 amino acid residues. The codon usage for flavin reductase *C. freundii* A1 is shown in Table 1 and it was compared with the codon usage for the flavin reductase from *E. coli*. This comparison was carried out as the flavin

reductase from *E. coli* was well-studied (Fontecave *et al*., 1987; Spyrou *et al*., 1991; Fieschi *et al*., 1995). Comparison made between the *fre* gene from both *E. coli* and *C. freundii* A1 showed that some of the amino acids of the enzyme are coded by a different codon of the amino acid. There is an inclination for codon selection, for example, CUG for leucine, CGU for arginine, GUG for valine, AUU for isoleucine, with a total nucleotides of 158 A, 169 C, 203 G and 172 U (or T).

The DNA sequence preceding the presumed start codon of the *fre* gene is shown to contain a potentially weak ribosomal binding site (RBS, shown in Fig. 2) and a promoter-like sequence. At nucleotide position 7, there is a putative -35 region with the sequence TTGACC followed by a putative -10 region with the sequence TCTTTC 22 nucleotides further down. However, further analysis done by using online WWW Promoter Scan software (http://www-bimas.cit.nih.gov/cgi-bin/molbio/proscan), indicating that no promoter region could be identified. This software is one of the online software to determine homologies of published sequences in our query sequence, suited mainly to transcriptional elements. However, it was admitted that most signal elements found by using this type of software probably will not have any meaning, as the elements may be in the wrong milieu, wrong cell type or wrong organism. Thus, it was ever concluded that the prediction would generate many more erroneous signals than significant ones (Alphey, 1997). In addition, there is no detailed analysis of this gene from other bacteria of the famili Enterobacteriaceae. Sometimes, prediction for new gene from a rare bacterial species remains lacking.

Restriction sites analysis of the *fre* gene fragment was carried out using NEBcutter (Version 2.0) at http://tools.neb.com/NEBcutter2/. The result shows that the gene fragment did not contain any restriction sites for *Bam*HI and *Pst*I. Thus, another pair of primers could be designed to have these sites incorporated so that the restriction of the fragment may not cut internally, resulting in shorter fragments. Hence, this would facilitate any directional cloning in the future.

Analysis of *C. freundii* flavin reductase structure: The flavin reductase of *C. freundii* A1has an estimated molecular weight (MW) of 26600.2 Da and an isoelectric point of 5.06, predicted using Compute pI/MW at http://web.expasy.org/compute_pi/. Spyrou *et al*. (1991) had previously shown that the open reading frame of

Table 1: Codon utilization of flavin reductases from *C. freundii* A1 and *E. coli*

		No. of codon used*	
Amino acid	Codon	*C. freundii*	*E. coli*[a]
Leucine (L)	CUA	1	0
	CUC	2	2
	CUG	12	8
	CUU	2	3
	UUA	4	2
	UUG	2	6
Serine (S)	AGC	1	1
	AGU	0	1
	UCA	2	1
	UCC	2	0
	UCG	2	4
	UCU	3	3
Arginine (R)	AGA	0	0
	AGG	1	0
	CGA	0	0
	CGC	7	7
	CGG	1	1
	CGU	12	11
Glycine (G)	GGA	4	2
	GGC	6	7
	GGG	1	4
	GGU	4	3
Valine (V)	GUA	4	4
	GUC	1	3
	GUG	9	8
	GUU	3	1
Proline (P)	CCA	1	2
	CCC	1	1
	CCG	6	5
	CCU	1	1
Alanine (A)	GCA	6	3
	GCC	6	4
	GCG	6	12
	GCU	3	4
Threonine (T)	ACA	2	2
	ACC	5	6
	ACG	4	4
	ACU	2	1
Isoleucine (I)	AUA	1	0
	AUC	5	9
	AUU	8	8
Cystenine (C)	UGC	0	2
	UGU	2	1
Lysine (K)	AAA	4	6
	AAG	3	1
Phenylalanine (F)	UUC	4	2
	UUU	6	8
Glutamate acid (E)	GAA	10	9
	GAG	10	9
Histidine (H)	CAC	3	2
	CAU	2	5
Glutamine (Q)	CAA	1	3
	CAG	4	3
Asparagine (N)	AAC	2	3
	AAU	4	1
Tyrosine (Y)	UAC	2	2
	UAU	6	5
Aspartic acid (D)	GAC	9	4
	GAU	9	13
Methionine (M)	AUG	7	7
Tryptophan (W)	UGG	3	3
Stop codon	UAA	0	0
	UAG	0	0
	UGA	0	1
Total No. of codons		234	234

*The universal translation table was used for the calculation of codon usage, [a]Acceession No. M61182

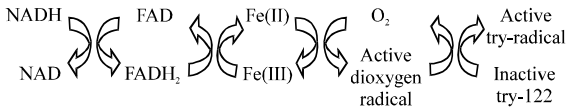

Fig. 4: The oxidoreduction process involved in the ribonucleotide redutase system

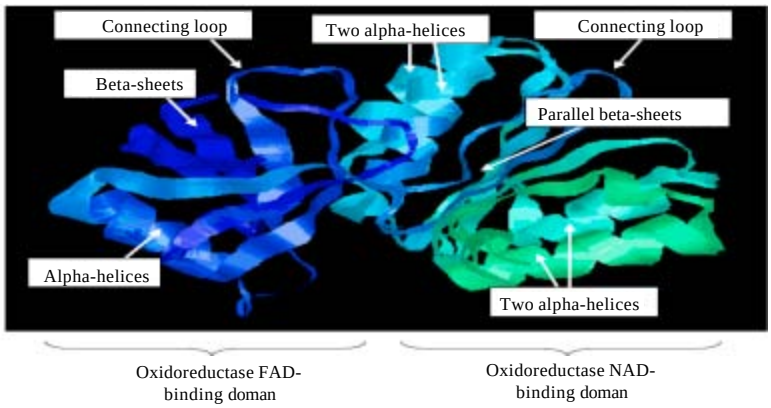

Fig. 5: The predicted structure for flavin reductase from *C. freundii* A1

E. coli's flavin reductase ended with an opal stop codon (UGA), yielding a polypeptide of 233 amino acid residues with a molecular weight of 26,212 Da.

From the deduced amino acid sequence, the protein family and function was determined. The gene fragment encoded an oxidoreductase protein. The AGGTG motif was found in the flavin reductase polypeptide sequence of *C. freundii* A1 and in the flavin reductase sequence from *E. coli*. This 5-residue continuous sequence (AGGTG) seems to form part of the pyridine nucleotide binding-site. The same motif had been observed in one of the subunits (protein C) of methane monooxygenase from methanotropic bacteria and also in cytochrome b_5 reductase (Spyrou *et al*., 1991). Protein C functions as a short electron transport chain that contains both an iron-sulfur center and 1 mol of tightly bound FAD per mol of protein. The N-terminal amino acid sequence of Protein C shows considerable homology with plant and some bacterial ferredoxins, whereas the C-terminal part shows significant homology with NADH:cytochrome b_5 reductase from human erythrocytes. It was later found out that the C-terminal part of protein C is also significantly similar to the *fre* polypeptide of *E. coli* (Spyrou *et al*., 1991) and also of *C. freundii* A1. This homology may show the relatedness in protein function.

With respect to the C-terminal half of the flavin reductase, there exists a striking similarity with monooxygenase starting with the continuous 5-residue sequence AGGTG at position 110 to 114 of flavin reductase (The nucleotide position is as shown in Fig. 2). The secondary structure predicted for this segment covers the end of the beta-strand and a reverse turn at a segment of predicted alternating beta-turn-alpha structures. This is of such a manner typical for nucleotide-binding domains of dehydrogenases. The same motif found in cytochrome b_5 reductase also marks the beginning of sequence homology between this enzyme, flavin reductase and monooxygenase (Spyrou *et al*., 1991). In *E. coli*, flavin reductase was a part of the ribonucleotide reductase, in which flavin reductase together with either Fe(II) or a second protein, known as fraction b, could provide the electron required for the reduction of the Fe(III) center (Spyrou *et al*., 1991). This oxidoreduction process is summarized in Fig. 4.

The AGRFEMA motif was found in both the flavin reductase polypeptide sequence from *C. freundii* A1 and *E. coli* at position 201 to 206. This motif might probably be part of the flavin-binding site of the enzyme. Due to the observed differences in flavin binding behaviour, the motif of flavin reductase is different from those of monooxygenase and cytochrome reductase. A higher degree of conservation was observed for monooxygenase and cytochrome reductase as these two classes of enzyme were classified as flavoprotein, purified with bound FAD. On the other hand, flavin reductase is not a flavoprotein, with no flavin tightly bound to the isolated protein (Fieschi *et al*., 1995). Flavin reductase is using soluble flavins as substrate rather than being a coenzyme (Tu *et al*., 1979). At the N-terminal half of the flavin

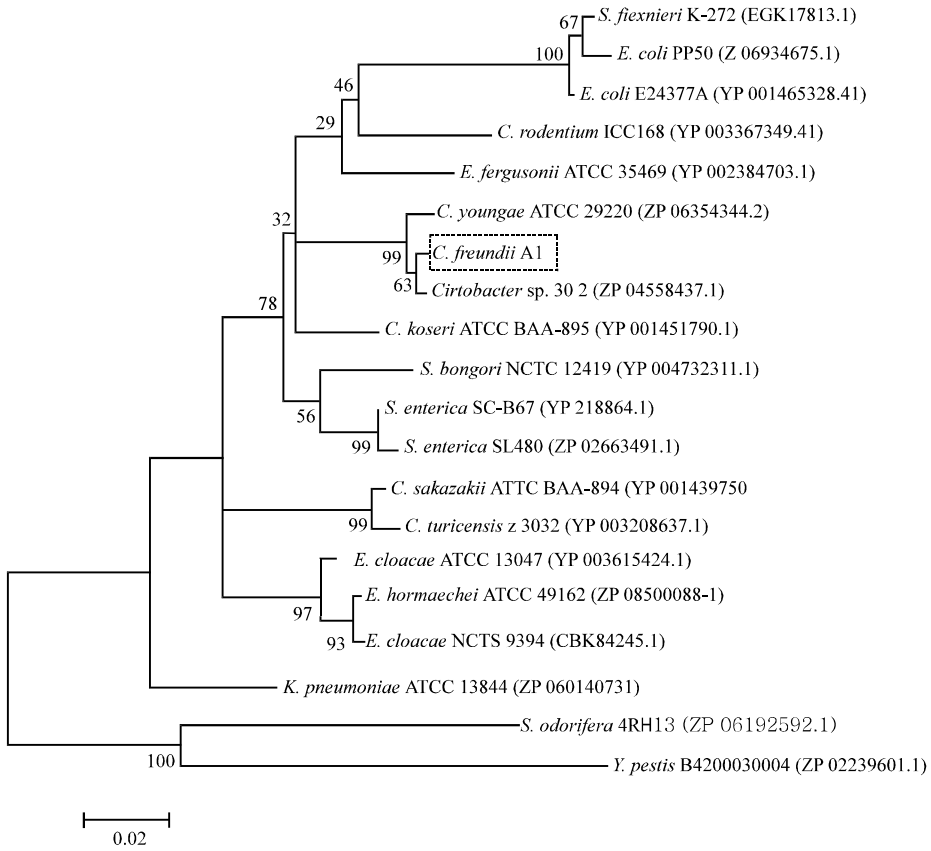

Fig. 6: Phylogram generated based on the nucleotide sequence alignment of flavin reductase amino acid sequences from *C. freundii* A1 and other Gram-negative bacteria. The GenBank assession numbers of the flavin reductases amino acid sequences used are indicated in parentheses

reductase, the 6-residue sequence DERDKR was observed at position 42 to 47. This motif was presumably surface residues.

The secondary and tertiary structures of the flavin reductase polypeptide of *C. freundii* A1 could be predicted from the amino acid sequence using Automated Modelling Mode of SWISS-MODEL Workspace. The amino acid sequence of Fig. 3 was submitted and the program compared it against a database of known structures, PDB or Protein Databank that stores the coordinates of protein structures being solved using either X-rays or NMR. The match was found of the query sequence to a protein of known structure and alignment was automated, threading the unknown sequence of flavin reductase of *C. freundii* A1 onto the known structure, introducing folds, the way in which the secondary structure elements in a protein are packed together. The structure obtained was viewed using RasMol (Version 2.7.5). The 3-D structure is shown in Fig. 5.

The remarkable homology was also observed between the flavin reductase from *C. freundii* A1 and other Gram-negative bacteria (shown in Fig. 6). Phylogenetic analysis showed that *C. freundii* A1flavin reductase gene is closely related to those from the genus citrobactor, *Escherichia*, *Shigella* and *Salmonella* with branching patterns supported by high bootstrap values. The enzyme was also aligned with azoreductase amino acid sequences as suggested by Chen *et al.* (2010) in order to determine its homology with other azoreductases. Branching pattern of the phylogenetic tree in Fig. 7 suggests that the flavin reductase from *C. freundii* A1 does not exactly form monophyletic cluster with any of the other three grouping of azoreductases. This finding suggests that flavin reductase from *C. freundii* A1 is rather unique but may imply a similar functional role in azoreduction. The possibility of this enzyme system involved in decolourisation of azo dyes will be reported in the future.

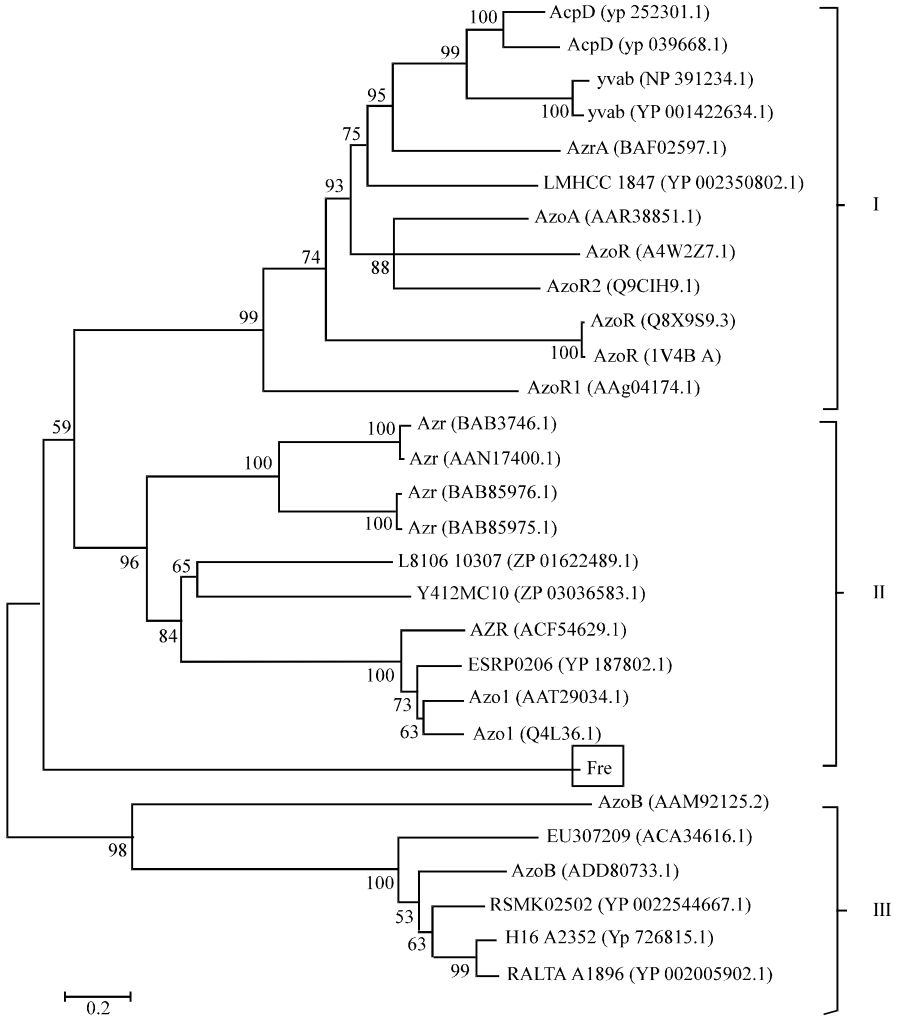

Fig. 7: Grouping of azoreductases and hypothetical azoreductases with conserved dinucleotide binding domain, I: Flavin-NADH-preferred azoreductase, II: Flavin-NADH-preferred azoreductase and III: Flavin free-NADH-preferred azoreductase

CONCLUSION

As a conclusion, *in silico* characterization of flavin reductase gene isolated from *C. freundii* A1 had presented various features of the gene that may facilitate further molecular studies of the gene and this may enable us to understand how the gene is related to azo dye decolourisation.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Ministry of Science, Technology and Innovation Malaysia (MOSTI) for research grant under Vote No. 74199 and National Science Fellowship (NSF) scholarship awarded to Chan G.F.

REFERENCES

Alphey, L., 1997. DNA Sequencing: From Experimental Methods to Bioinformatics. BIOS Scientific Publishers, Manchester, UK, ISBN: 9780387915098, Pages: 206.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucl. Acids Res., 25: 3389-3402.

Chen, H., J. Feng, O. Kweon, H. Xu and C.E. Cerniglia, 2010. Identification and molecular characterization of a novel flavin-free NADPH preferred azoreductase encoded by azoB in *Pigmentiphaga kullae* K24. BMC Biochem., Vol. 11. 10.1186/1471-2091-11-13

Fieschi, F., V. Niviere, C. Frier, J.L. Decout and M. Fontecave, 1995. The mechanism and substrate specificity of the NADPH:flavin oxidoreductase from *Escherichia coli.* J. Biol. Chem., 270: 30392-30400.

Fontecave, M., R. Eliasson and P. Reichard, 1987. NAD(P)H: Flavin oxidoreductase of *Escherichia coli.* A ferric iron reductase participating in the generation of the free radical of ribonucleotide reductase. J. Biol. Chem., 262: 12325-12331.

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch, 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook, Walker, J.M. (Ed.). 1st Edn., Humana Press, New Jersey, USA., ISBN-13: 978-1588295934, pp: 571-607.

Prestridge, D.S., 1995. Predicting pol II promoter sequences using transcription factor binding sites. J. Mol. Biol., 249: 923-932.

Rashid, N.A.A., A.R.H.M. Yusoff, R. Ahmad, S. Misran, G.F. Chan and P. Murugaiya, 1999. Biodegradation of textile dyes using a bacterium isolated from sewage oxidation pond. Proceedings of the Symposium Kimia Analisis, July 22, 1999, Kuala Terengganu, Terengganu, Malaysia.

Russ, R., J. Rau and A. Stolz, 2000. The function of cytoplasmic flavin reductases in the reduction of azo dyes by bacteria. Applied Environ. Microbiol. 66: 1429-1434.

Saratale, R.G., G.D. Saratale, J.S. Chang and S.P. Govindwar, 2011. Bacterial decolorization and degradation of azo dyes: A review. J. Taiwan Inst. Chem. Eng., 42: 138-157.

Schwede, T., J. Kopp, N. Guex abd M.C. Peitsch, 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res., 31: 3381-3385.

Spyrou, G., E. Haggard-Ljungquist, M. Krook, H. Jornvall, E. Nilsson and P. Reichard, 1991. Characterization of the flavin reductase gene (fre) of *Escherichia coli* and construction of a plasmid for overproduction of the enzyme. J. Bacteriol., 173: 3673-3679.

Tamura, K., J. Dudley, M. Nei and S. Kumar, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol., 24: 1596-1599.

Tang, C., 2002. Bioinformatics: Life science research in silico. J. Young Investigators, Vol. 6.

Tu, S.C., J.E. Becvar and J.W. Hastings, 1979. Kinetic studies on the mechanism of bacterial NAD(P)H: Flavin oxidoreductase. Arch. Biochem. Biophys., 193: 110-116.

Vincze, T., J. Posfai and R.J. Roberts, 2003. NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Res., 31: 3688-3691.