Catalytic Site Prediction of Azoreductase Enzyme of *E. coli* with Potentially Important Industrial Dyes Using Molecular Docking Tools

Bikash Thakuria, Nangkyntiewbor Jungai, Samrat Adhikari*
Bioinformatics Centre, Department of Biotechnology, St Edmund's College, Shillong, Pin-793003, Meghalaya, India.

* Corresponding author. Tel.: 91-0364-2220808; email: stedmundc.btisnet@nic.in
Manuscript submitted January 5, 2015; accepted March 10, 2015.
doi: 10.17706/ijbbb.2015.5.2.91-99

Abstract: Azoreductase is an FMN-dependent and NADH-dependent enzyme of *Escherichia coli*. This enzyme is responsible for the degradation of azo dyes. In this study, we retrieved the crystal structure of the enzyme from PDB and 18 azo dyes from NCBI PubChem compound. These azo dyes were then docked with the FMN- dependent NADH-azoreductase enzyme to analyze the binding affinity of the azo dyes with the enzymes and predict the catalytic sites. In this approach, we identified the catalytic residues of FMN-dependent and NADH dependent enzyme of *Escherichia coli* which were then evaluated in terms of properties including function, conservation, hydrogen bonding, B-factor and flexibility. The results indicated that Phe-172, Glu-174, Lys-145, Asp-146 and Lys-169 played an important role as catalytic site residues in the enzyme. It is hoped that this information will provide a better understanding of enzyme mechanisms and also be used to improve the designing strategies for dyes detoxification. In this study, the approach emphasizes on a better understanding of the biodegradation of some of the commercially important azodyes mediated by azoreductase from *E. coli*. Furthermore, the catalytic site residues information is essential for understanding and altering the substrate specificity and for the design of a harmless azodye.

Key words: Azoreductase, azo dyes, EDO, FMN, IPA, NADH.

1. Introduction

Azo dyes are widely used for industrial, printing, clinical purposes as well as textile dyeing because of its chemical stability, ease of synthesis, and versatility. In addition, azo dye compounds are also the most commonly used drugs in the treatment of inflammatory bowel disease. Their durability, however, causes pollution when they are released into the environment as the effluent [1]. Moreover, the release of these dyes into the environment is undesirable, not only because of their colour, but also of their by-products acting as agents of toxicity and mutagenicity. Biological treatment of azo dyes by the use of bacteria has been widely reported. Enzymes that catalyse the reduction of these azodyes are termed as azoreductases (EC 1.7.1.6). This enzyme basically utilizes NADH/FMN and/or NADPH as an electron donor to decolorize azo dyes into corresponding aromatic amines by reductive cleavage of azo bonds. The decolourization is a rate- limiting step, which is followed by the aerobic mineralization of the colourless aromatic amines. Furthermore, azoreductase is also involved in the site-specific delivery of azo prodrugs, which are therapeutically inactive in their intact form and rely on azo reduction by azoreductases of

intestinal microflora for activation. Proteins with azoreductase activity have been identified and characterized from a wide variety of bacteria [2], [3].

There are at least two different types of bacterial azoreductases: Flavin dependent and Flavin independent [4]. Flavin-dependent azoreductases can be further classified into two families according to their amino acid sequences. Azoreductases from *E. coli* and *Bacillus sp*. strain OY1-2 are representative of the two flavin- dependent azoreductase families, respectively. Although they are effective for an *in vitro* enzyme assay, overexpression of these azoreductases *in vivo* results in little or no increase of bacterial decolourization activity. Thus, the physiological role of azoreductase has recently been a subject of debate. As the introduction of azo dyes into the environment is due mainly to human activities, reduction of azo dyes may not be the primary role of these enzymes [5].

Azo dyes such as Sudan dyes are not legal for use as colorants in food; however, recently these dyes have been detected as contaminants in the food supply [6]. The human health impact of exposure to azo dyes used in certain food products has caused concerns since they may have genotoxic properties. The environmental fate and subsequent health effects of the azo dyes released in textile and paper industry wastewater are increasing being studied by the scientific community [7].

While azo dyes are generally considered to be persistent pollutants because they are typically recalcitrant to aerobic biotransformation [8], they might be metabolized by azoreductases from commensal microorganisms, mammalian liver cells, and soil microorganisms [9]. A variety of microorganisms, including bacteria and fungi, are capable of decolorizing a diverse range of azo dyes. Some bacteria have the ability to degrade azo dyes both aerobically and anaerobically [6], [10]. Bacterial degradation of azo dyes is often initiated with the cleavage of azo bonds by azoreductases which are followed by the aerobic degradation of the resulting amines [4].

To understand the biodegradation of the azo dyes mediated by the enzyme and to design harmless azo dyes, it is essential to discover the catalytic site residues within the azoreductase enzyme and to perform flexibility analysis of the catalytic residues [11].

In the present study, docking study of the FMN-dependent NADH-azoreductase of *E. coli* and the industrially important dyes was carried out to understand the mechanism of azoreductase catalyzing enzymatic reactions, the catalytic site residues and the binding affinities of the azodyes and the enzyme. Therefore in context to the present study, an FMN-dependent NADH-azoreductase enzyme of *E. coli* had been selected as the target protein or protein of interest for the theoretical studies of biodegradation of azodyes compounds. This enzyme was chosen to view the interaction with azodyes, which are pollutants and toxic to the environment. This study would pave various strategies for understanding and altering the substrate specificity for effective design of a potential harmless azodyes.

2. Materials and Methods

2.1. Data Set

The enzyme molecule, FMN-dependent NADH-azoreductase of *E. coli* was obtained from Protein Data Bank (PDB) [12], and the three-dimensional (3D) structures of this protein had been solved by X-ray crystallographic techniques with 1.80 A^{o} resolution. For substrate 18 commercially important azodyes were retrieved from the NCBI PubChem Compound database [13].

2.2. Computation of Docking Score between the Ligands and the Enzyme

Protein or ligand modification was carried out by removing water molecules from the protein and energy minimization was carried out using Swiss PDB Viewer (SPDBV) tools [14]. Autodock Vina [15] was used for docking studies of azoreductase (protein) and azodyes (ligands). The interactions between the

different ligands with the protein of interest were viewed in PyMOL [16] and the analysis of the H-bond and its positions were also evaluated with PyMOL.

2.3. Functional Site Location

The catalytic or functionally important residues of a protein are known to exist in evolutionary constrained regions. However, the patterns of residue conservation alone are sometimes not very informative, depending on the homologous sequences available for a given query protein. Hence, the prediction of functional sites in newly solved protein structures is a challenge for computational structural biology. Most methods for functional site identification utilize measures of amino acid sequence conservation in homologous sequences, based on the assumption that functional sites are relatively conserved during evolution. Protein structural information has also been used to help identify protein functional sites. Ligand binding site prediction of the azoreductase enzyme was carried out using Q-SiteFinder [17], by binding hydrophobic (CH_3) probes to the protein, and finding clusters of probes with the most favourable binding energy. These clusters are placed in rank order of the likelihood of being a binding site according to the sum total binding energies for each cluster.

2.4. Hydrogen Bonding and B Factor

The hydrogen bonding and the B-factor are the two vital parameters in the catalytic residue prediction. Catalytic residue hydrogen bonding was investigated using the result obtained from the docking study. Catalytic residue B-factors is a measure of residue flexibility. Catalytic residues tend to have lower B-factors than all residues. Catalytic residues are very precisely positioned and held in place, as predicted by their low B-factors and Hydrogen bonding. The hydrogen bonding information is obtained from the docking studies. The B-factors of the residues interacting with the ligands and the active sites were taken from PDB file for each atom in a residue, and then averaged over the whole residue.

3. Results and Discussion

3.1. PDB Structure

The enzyme, FMN-dependent NADH-azoreductase of *E. coli* has a crystallized structure characterized by three ligands FMN, EDO and IPA. The structure was obtained from PDB with the PDB id, 1V4B (Fig 1(a)).

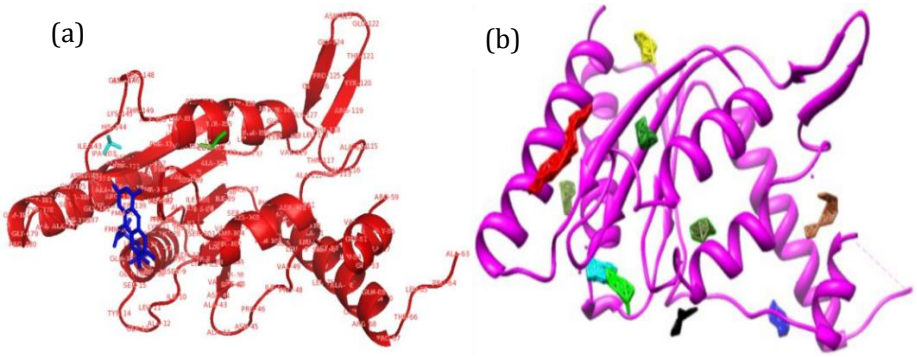

Fig. 1 (a) The three dimensional structure of the target protein (1V4B) obtained from PDB; (b) The active sites displayed in the three dimensional structure of the target protein predicted with the aid of Q-SiteFinder and viewed under Chimera.

3.2. Multiple Sequence Alignment

Catalytic residues are more conserved than the average residues [18]. Hence the conserved residues were investigated to have a relationship with the catalytic residues. A BLAST search along with a multiple sequence alignment with the BLAST result showed the conservation of almost all the residues.

3.3. Computation of Docking Score between the Ligands and the Enzyme

Proteins are the basis of life processes of all enzymes at the molecular level. The protein interaction is either with another protein or with small molecules. Many biological studies will benefit from interactive predictions. Docking study showed the binding affinity, number of hydrogen bonds and the binding residues. It is interesting to note that the binding affinities have negative values as shown in Table 1. This revealed the high feasibility of this reaction. The docked complexes were further analysed with the molecular visualization tools, PyMOL as shown in Fig. 2. The docking analysis showed that three dyes *viz.* Sudan black B (61336), Kayaku acid blueblack 10B (5473482) and FD & C Yellow no.5 (25245842) formed H-bonds with the enzyme residues Asp-109, Arg-59 and Ala-112, and Arg-59 of Site 7 respectively. Moreover, six dyes showed docking conformation with two of the natural ligands of the target protein; 3.3'dichlorobenzidine (7070) with IPA, Methyl red (10303) with IPA and Glu-174, HABA (5357439) IPA and Lys-145, Amaranth (FD & C red no. 2) (5473445) with IPA, Glu-174 and Phe-172, Grasal orange (5858445) with IPA and Antipyrylazo III (9573878) with FMN and His-144 (Table 2). The variation in the docking score indirectly depicted the rate of decolourization [19]. The rate of colour removal for trypan blue was higher than any other azo dye. The amino acids Phe-172; Glu-174; Asp-189; Asp-146; Lys-145; Asp-109; Arg-59; Ala-112; Ala-200; Asp-146; Asp-150; Gly-174; Lys-169; Phe-170; Asp-167; Thr-166; Gly-131; Gly-164; Lys-132; His-144; Arg-59 interacted with different ligands and were taken into account for comparison with the active sites obtained from active site prediction.

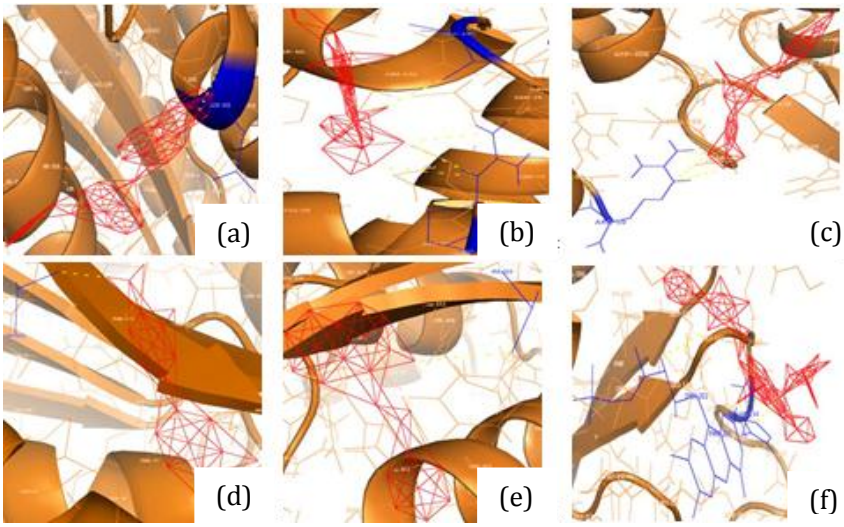

Fig. 2. Docked conformations of (a) Sudan black B (61336), (b) Kayaku acid blueblack 10B (5473482), (c) FD & C Yellow no.5 (25245842) in the active site; (d) 3.3'Dichlorobenzidine with IPA; (e) Grasal orange with IPA and (f) Antipyrylazo III with FMN and His-144.

3.4. Functional Site Location

Active sites of the target protein were predicted using Q-SiteFinder and the output file was viewed under Chimera. 10 active sites were obtained from the study along with the corresponding amino acid residues presented in each active site (Fig. 1(b)) and the corresponding aminoacids evaluated are depicted in Table 3. Each of the sites was analysed and compared with the amino acids interacting with the ligands

in the docking study. The amino acid positions of the sites are also listed in Table 3. The amino acid residues were found to have lower B-factor than the other residues and the site 6 showed the absence of aminoacids. The 7th active site consists of amino acids which are interacting with ligands in the docking study. The docking result shows that the amino acids such as Phe-172, Glu-174, Lys-145, Asp-146 and Lys-169 are very much repeated in the interaction with more than one ligand. This reveals that these amino acids are catalytic residues. The active site variations suggest that the enzyme could probably decolourize a wide range of azo dyes [20].

Table 1. Detailed Docking Analysis Result Showing Interaction of Azodyes and Azoreductase Enzyme. The Amino Acid Residue Position of Azoreductase Interacting with Each Azodye is also Provided along with the Azodye Information

Azo Dyes	Common Names	Affinity (kcal/mol)	No. of H-Bonds	Positions (residues)
7070	3.3'Dichlorobenzidine; Dichlorobenzidine base	-7.1	1	IPA
7340	O-Aminoazotoluene; C.I. Solvent Yellow 3	-6.8	1	Phe-172
8411	O-Dianisidine;3.3'-DimethoxybenzidinE	-5.8	0	
10303	Methyl red; O-METHYL RED	-6.7	2	IPA,Glu-174
11490	4-(Phenylamino)-3'-Sulfoazobenzene	-6.7	4	Asp-189,146, Glu-174, Lys-145
61336	Sudan black B; Solvent black 3	-9.8	1	Asp-109
5357439	HABA; 2-(4-hydroxyphenylazo) benzoic acid	-6.4	2	IPA, Lys-145
5473445	UNII-37RBV3X49K; Amaranth (FD & C red no. 2)	-8.2	4	IPA(2), Glu-174, Phe-172
5473482	Amidoschwarz; Kayaku acid blueblack 10B	-7.8	4	Arg-59(3), Ala-112
5483782	D.C. Red no. 9; Pigment Red 53-1	-8.0	1	Ala-200
5809667	Sudan (1); Solvent red 23	-8.1	0	
5858445	1-Phenylazo-2-napthol; Grasal orange	-7.7	1	IPA
5876571	Scarlet red; Sudan IV	-8.5	0	
6068628	AC101CE7; 2610-10-8 (hexa-hydrochloride salt)	-9.6	9	Lys-145(2),169, Phe-170(4), Asp-146(2)
6453340	Dianisidine sulfate; 119-90-4 (parent)	-3.5	6	Asp-167, Thr-166, Gly-131, 164, Lys-132(2)
9573878	Antipyrylazo III; EINEC3238988-2	-8.0	4	FMN(3), His-144
25245842	FD & C Yellow no. 5	-7.0	3	Arg-59(3)

3.5. Hydrogen Bonding and B-factor

Both the Hydrogen bonding and the B-factors in the crystal structures were used as a measure of residue flexibility. Analysis shows that the amino acids interacting with the ligands are involved in hydrogen interactions as a donor or as an acceptor [21]. This shows that catalytic residues have a limited conformational freedom. The docking result shows that the ligands have hydrogen bonding with amino acids and it is illustrated in Table 1. Catalytic residues tend to have lower B-factors than all residues, suggesting that they have to be more rigidly held in place than the averaged over the whole residue [21]. B-factors were taken from the PDB file for each atom in a residue and then averaged over the whole residue [22]. The B-factor for the residues, Phe-172 is 16.45, Glu-174 is 16.53, Lys-145 is 25.16, Asp-146 is 25.77 and Lys-169 is 31.8 [23].

Table 2. Target Protein Docked with Natural Ligands

Ligands	Affinity(kcal/mol)	H-Bonds	Positions(residues)
EDO	-3.3	6	Ser-15,17(2),139(2), Gln-16
FMN	-10.5	15	His-144, Gly-142,141, Phe-98, Asn-97, Tyr-90. Met-95. Ser-9.15.17(2).139(2).
IPA	-3.2	5	Ser-15,17(2),139, Gln-16

Table 3. The Amino Acid Composition of the Different Active Sites Predicted

Active Sites	Residues
Site 1	Ile 143
Site 2	Ser 9, Leu 11, Tyr 14, Ser 15, Gln 16, Ser 17, Pro 94, Met 95, Tyr 96, Ser 139, Arg 140, Gly 141
Site 3	Phe 24, Trp 28, Val 88, Lys 133, Ala 134, Ile 135
Site 4	Met 95, Tyr 96, Asn 97, Phe 98, Ser 139, Arg 140, Gly 141, Gly 142
Site 5	Pro 48, Leu 54, Arg 68,Glu 71, Ala 72, Leu 75
Site 6	-
Site 7	Leu 50, Val 55, Leu 58, Arg 59, Asp 109, Ala 112, Arg 113, Ala 114, Phe 118
Site 8	Gln 19, Leu 20, Asp 22, Tyr 23, Glu 26
Site 9	Ile 10, Tyr 96, Asn 97, Asn 99, Ile 100, Ser 101,Thr 102, Gln 103
Site 10	Phe 98, Asn 99, Ile 100

4. Conclusion

A major challenge in azodye degradation and a harmless azodye design is the elucidation of biochemical and biological properties of enzymes, including the determination of catalytic residues that belong to the ligand-substrate binding site. Comparing the results from Q-Site Finder and docking studies, it indicates that the amino acid residues Phe-172, Glu-174, Lys-145, Asp-146 and Lys-169 play an important role as catalytic site residues in the azoreductase enzyme of *E. coli*. This docking study also provides information on the binding affinity of the ligands with azoreductase enzyme. The rate of colour removal for trypan blue is higher than any other azo dye. It is hoped that this information would provide a better understanding of the molecular mechanisms involved in catalysis and a heuristic basis for predicting the catalytic residues in enzymes of unknown function. The natural ligands (FMN, EDO and IPA) were also found to interact with some of the ligands. In this work the catalytic residues are reported as well as the binding affinities for some commercially important azodyes. The study made in this project would facilitate researchers a better understanding of enzyme mechanisms and also used to improve the designing strategies of less harmful azodyes.

Acknowledgment

The authors take this opportunity to acknowledge the funding received from Department of Biotechnology, Govt. of India for setting up Bioinformatics Centre under BTISNET programme in Department of Biotechnology, St. Edmund's College, Shillong. The authors also express their heartfelt gratitude to Dr Sylvanus Lamare, Principal, St. Edmund's College, Shillong for his support throughout the work.

References

[1] Nakanishi, M., Yatome, C., Ishida, N., & Kitade, Y. (2001). Putative ACP phosphodiesterase gene (acpD) encodes an azoreductase. *Journal of Biological Chemistry*, *276(49)*, 46394-46399.

[2] Hanauer, S. B. (1996). Inflammatory bowel disease. *New England Journal of Medicine, 334*, 841-848.

[3] Wang, C. J., Hagemeier, C., Rahman, N., Lowe, E., Noble, M., Coughtrie, M., *et al*. (2007). Molecular cloning, characterization and ligand-bound structure of an azoreductase from *Pseudomonas aeruginosa*. *Journal of Molecular Biology*, *373(5)*, 1213-1228.

[4] Chen, H. (2006). Recent advances in azo dye degrading enzyme research. *Current Protein Precipitation Science*, *7(2)*, 101-111.

[5] Liu, G., Zhou, J., Fu, Q. S., & Wang, J. (2009). The *E. coli* azoreductase azoR is involved in resistance to thiol-specific stress caused by electrophilic quinones. *Journal of Bacteriology*, *191(20)*, 6394-6300.

[6] Xu, H., Heinze, T. M., Chen, S., Cerniglia, C. E., & Chen, H. (2007). Anaerobic metabolism of 1-amino-2-naphthol-based azo dyes (Sudan dyes) by human intestinal microflora. *Applied Environmental Microbiology, 73(23)*, 7759-7762.

[7] McMullan, G., Meehan, C., Conneely, A., Kirby, N., Robinson, T., Nigam, P., *et al*. (2001). Microbial decolourisation and degradation of textile dyes. *Applied Microbiology and Biotechnology*, *56(1-2)*, 81-87.

[8] Stolz, A. (2001). Basic and applied aspects in the microbial degradation of azo dyes. *Applied Microbiology and Biotechnology, 56(1-2)*, 69-80.

[9] Levine, W. G. (1991). Metabolism of azo dyes: Implication for detoxication and activation. *Drug Metabolism Reviews, 23(3-4)*, 253-309.

[10] Prasad, N. K., Vindal, V., Narayana, S. L., Kunal, S. P., *et al.* (2012). In silico analysis of *Pycnoporus cinnabarinus* laccase active site with toxic industrial dyes. *Journal of Molecular Modeling, 18(5),*

2013-2019.

[11] Ramanthan, K., Shanthi, V., & Sethumadhavan, R. (2009). *In silico* identification of catalytic residues in azobenzene reductase from *Bacillus subtilis* and its docking studies with azo dyes. *Interdisciplinary Science*, *1(4)*, 290-297.

[12] Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice, M. D., Rogers, J. R., *et al*. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology*, *112*, 535-542.

[13] Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., & Bryant, S. H. (2009). *Pubchem:* A public information system for analysing bioactivities of small molecules. *Nucleic Acids Res., 37*, W623-W633.

[14] Guex, N., & Peitsch, M. C. (1997). Swiss-Model and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis, 18*, 2714-2723.

[15] Trott, O., & Olson, A. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry, 31*, 455-461.

[16] DeLano, W. L. (2013). The PyMOL molecular graphics system. *DeLano Scientific San Carlos,* California USA. from http://www.pymol.org

[17] Laurie, A. T. R., & Jackson, R. M. (2005). Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, *21*, 1908-1916.

[18] Bartlett, G. J., Porter, C. T., Borkakoti, N., & Thornton, J. M. (2002). Analysis of catalytic residues in enzyme active sites. *Journal of Moecular Biology*, *324(1)*, 105-121.

[19] Abo-Farha, S. A. (2010). Photocatalytic degradation of monoazo and diazo dyes in wastewater on nanometer-sized TiO_2. *Journal of American Science, 6(11)*, 130-142.

[20] Philem, P. D., & Adhikari, S. (2012). Homology modeling, docking studies and functional analysis of various azoreductase accessory interacting proteins of *Nostoc. sp.* PCC7120. *Bioinformation, 8(7)*, 296-300.

[21] Sacquin-Mora, S., & Lavery, R. (2006). Investigating the local flexibility of functional residues in hemoproteins. *Biophysical Journal, 90(8)*, 2706-2717.

[22] Yuan, Z., Zhao, J., & Wang, Z. (2003). Flexibility analysis of enzyme active sites by crystallographic temperature factors. *Protein engineering design and selection, 16(2)*, 109-114.

[23] Sridhar, S., & Chandra, J. H. (2014). Involvement of computational tools towards in silico remediation—Synthetic textile dyes interacting with azoreductase. *International Journal of Chemical Technology Research*, *6(9)*, 4412-4441.

Bikash Thakuria was born in Guwahati, Assam (India) in 1988. He has obtained a master's degree in the field of biotechnology from Bangalore University (India) in 2012. Currently he is working as a research associate under Bioinformatics Infrastructure Facility (BIF), funded by the Department of Biotechnology, Govt. of India, in St. Edmund's College Shillong. His field of research is proteomics and computational biology. He has also participated in different national level workshops in the field of Biotechnology. He is also associated with the docking and interaction studies of different azo dyes and their environmental impact. Currently he is working on a medicinal plant named *Smilax aspera* and its proteins for cancer studies.

Nangkyntiewbor Jungai was born in Shillong, India in 1988. He has obtained a master degree in biotechnology at North Eastern Hill University, Shillong, Department of Biotechnology and Bioinformatics (2011). His areas of research are in the fields of genetic engineering, cyanobacterial biotechnology, cancer and cell cultures and traditional medicinal plants. At present, he is working as a senior research fellow (SRF) in the Institutional Biotech Hub, St. Edmund's College, Shillong under the sponsorship of Department of Biotechnology (DBT), Government of India (GoI). He has also participated in various seminars and workshops at the national level.

Samrat Adhikari was born in Shillong, Meghalaya in 1979. He has obtained the doctoral degree from North Eastern Hill University, Shillong (2011) in the field of environmental biotechnology and the masters with specialization in molecular microbiology (2002) from Bangalore University, Bangalore. He has received the Biotech Industrial Training Programme fellowship (2003) and also has a vast teaching and research experience for 12 years. Currently he is an assistant professor & the head of the Biotechnology Department, St. Edmund's College, Shillong. He has published 5 papers in peer reviewed journals and 3 in national & international conferences. He has undergone many trainings and also has organized many workshops for young researchers. He is also the reviewer of many academic journals. He has supervised for 3 M.Tech thesis and is currently handling 3 projects sponsored by the different funding agencies of Govt. of India. He is presently a member of many academic bodies in universities. At present, he is working on environmental biotechnology & bioinformatics.