Discussion

Generally, cyanobacteria can be cultured in mineral medium. Although, several culture media have been suggested to maintain the organism, *Nostoc* sp. 701 was cultured in BG$_{11}$ medium in the absence of nitrogen source as it is not required since, the heterocyst is involved in atmospheric nitrogen fixation (Beesley et al., 2001; Giordanino et al., 2011). Chlorophyll $\alpha$ (chl $\alpha$) is a green pigment responsible for photosynthesis and its absorbance ranges from 660 to 665 nm. It is the only pigment existing in BGA rather than other types of chlorophyll (Hosikian et al., 2010). Moreover, it is one of the important components within a cell that is used to measure the growth accurately (Kutser, 2004). However, the extraction procedure is different in algae and cyanobacteria; it could be completely extracted using methanol in lesser time scale (within 30 min) when compared to other organic solvents (Natesan and Sanmugasundaram, 1989; Dere et al., 1998). Accordingly, *N. muscorum* CENA-18, *N. commune* UTEX 584 and *N. muscorum* CENA-61 strains chl $\alpha$ were reported to be 4.6, 2.4 and 11.2 $\mu$g/ml respectively in methanol extract in 20 days grown culture that was an evidence for maximum growth of the organism (Foire et al., 2000). Therefore, a similar procedure adopted to determine the growth of *Nostoc* sp. 701 and recorded its maximum growth on the 20$^{th}$ day with 4.656 $\mu$g/ml chl $\alpha$ (Fig. 7). Although, this is in accordance to the result of Foire et al., (2000), according to whom, *Nostoc* sp. 701 attained its exponential growth on the 20$^{th}$ day, the yield percentage varied, which was comparatively higher than *N. commune* UTEX 584 and lower than *N. muscorum* CENA-61. Hence, the estimation of chl $\alpha$ in *Nostoc* sp. could be an effective method for the determination of growth.

The total protein content of the organism harvested from exponentially grown culture was estimated to be 14.6 mg/ml and to 12.63 mg/ml before and after chemical digestion. The protein content in *Nostoc* sp. 701 was similar to other microalgae *Porphyridium cruentum, Scenedesmus almeriensis, and Muriellopsis* sp. and the cyanobacteria *Synechocystis aquatilis* and *Arthrospira platensis* (Lopez et al., 2010). Glutathione reductase in terms of “specific activity” was measured to be 0.087 and 0.109 U/mg on before and after the chemical treatment respectively. Though the chemical treatment reduced the protein content, it inspired an interest in the study because, there was an increase in glutathione reductase (GR) activity by 20% when estimated by colorimetric methods (Piggot and Karuso, 2007) that could be achieved by the
treatment. Thus it could be understood that, by both sonicating and digesting the sample with chemical (TritonX-100 with PMSF), the enzyme activity could be increased. For instance, in a study by Serrano et al., (1984), who reported only 0.04U/mg of GR activity in *Anabaena* by exposing the cells to the same concentration of chemical treatment (X-100 with PMSF), however estimated through UV method ($\Delta A_{340\text{nm}}$). Moreover, a combined extraction procedure i.e., ultrasonication and chemical digestion estimated 3 fold increase in GR activity (0.109U/mg) which suggested that the above procedure could be a robust method to estimate the total GR activity by colorimetric assay ($\Delta A_{412\text{nm}}$) (Table 6). Thus, the advantage of simultaneous chemical digestion to the sonicated sample could be realized as an effective strategy that could be attempted for further investigation. The specific activity of GR enzyme was further increased to 0.121U/mg by acetone precipitation method which was considered as pre-polishing step since it removed the green pigment and other impurities from the sample.

Initially, the enzyme sample was subjected to charge based separation by DEAE-cellulose which is a positively charged resin that eluted negatively charged proteins NaCl (0 to 0.5M) salt gradient (Connel and Mullet, 1986). The purity of the enzyme GR from DEAE-cellulose was increased approximately to 18 fold from the acetone precipitated sample. Therefore, the specific activity of GR was 2.22 U/mg, which was obtained from 94% of total protein yield (Table 7). However the protein yield was lesser than GR purified from *Anabaena*, the specific activity of *Nostoc* sp. 701 GR was found to be 2 fold higher (Serrano et al., 1984). The purity of the enzyme was also higher than *Phaeodatylum tricornutum* (Diatom), as the specific activity, purification factor and yield were reported to be 0.82 U/mg, 1.8 fold and 71% respectively (Arias et al., 2010).

Affinity (or ligand) chromatography was usually performed on 2', 5'-ADP Sepharose 4B which was reported to purify GR and its isoforms from animal and plant sources. But, it is in contrast to cyanobacteria, which showed lack of affinity to GR enzyme due to intermediary residue substitution in the fingerprint GXGXXG. For instance, serine (174) interacts with 2'-phosphate NADPH, and the fingerprint is unable to interact with ligand in the 2', 5'-ADP Sepharose 4B matrix (Mittl et al., 1994). But, it could interact when it substituted by glycine or alanine, reported in *Anabaena* (Jiang et al., 1995; Danielson et al., 1999). In such case, an alternative ligand chromatography i.e., reactive red-120 agarose, has been postulated to a non-specific interaction with GR which is also proved to be an efficient method to purify the
enzyme. Based on this affinity chromatographic technique, the GR purity is reported to be 24U/mg of specific activity with the purification factor of 599 fold in *Anabaena* (Serrano *et al.*., 1984). The purity of GR from the reactive red-120 agarose chromatography seemed to be similar when compared to 2', 5'-ADP Sepharose 4B ligand chromatography. In this study, purity of the GR was 51.04U/mg, 71.9% and 422 (fold) of specific activity, protein yield and purification factor respectively, from reactive red-120 agarose chromatography (Table 7). As a result, the specific activity was observed to be much greater than *Anabaena* though the lower purification factor (599 fold) was observed comparatively in *Nostoc* sp. 701. In addition to, the purity of the enzyme was also higher than *P.falciparum* as it is reported to be 70% and 114 fold purification factor (Gilberger *et al.*, 2000). Moreover, the purity of GR was found to be higher than Camel erythrocyte GR and Sheep liver (Ulusu *et al.*, 2005; Al-senaidy, 2010).

The gel filtration chromatography was performed on sephacryl S-300 which is similar to sephadex G-200 and superdex 200 regarding their separation range that facilitated to separate the GR protein on basis of molecular size of the protein (Erat *et al.*, 2003). It was the final polishing step in the purification of GR from *Nostoc* sp.701. From the above technique, the specific activity, protein yield and purification factor and purity were 149U/mg, 62 % as final recovery and 1230 fold respectively (Table 7), for GR from *Nostoc* sp.701. Therefore, the GR final purity was similar to *Anabaena*, besides it seems to be higher purity than sheep and mouse liver (Ulusu *et al.*, 2005). Moreover, camel GR final purity was reported to be 58U/mg, 43% and 1547 of the specific activity, protein yield and purification fold respectively. Even though the purification fold seems to be similar to *Nostoc* sp.701, both specific activity and recovery of protein increased more than 20% in this study (Al-senaidy, 2010). Therefore, the GR specific activity of 149U/mg, 62% protein yield and 1230 fold purity was considered as final purification product for further physical, chemical and kinetic characterization.

Further, purified GR was subjected to physico-chemical characterization and analysis for optimum pH, pI (obtained from *insilico* analysis), temperature, buffer-ion concentration and effects of metal ions. The standard assay pH 7.5 was used to measure the GR activity which is similar to other organisms (Yu and Zhou, 2007; Cardoso *et al.*, 2008). In this study, a broad range of GR enzyme activity was observed between pH 7.0 to 9.0 of GR from *Nostoc* sp. 701 (Fig. 11), which is found to be similar when compared to *Anabaena* (Danielson *et al.*, 1999). Therefore, the optimum pH for GR activity was observed to be increased as increasing the pH
up to pH 9.0 thereafter, the apparent enzyme activity was declined, reported in *Anabaena* as well. The optimum pH of the GR activity revealed that GR from *Nostoc* sp. 701 is similar to plant but higher than *E.coli*, yeast, animal and human GR (Yu and Zhou, 2007; Can *et al.*, 2010). According to Lascano *et al.*, (1999), the stability of GR in higher (alkaline) pH was due to rapid consumption of substrate (either GSSG or NADPH) in order for endothermic entropy and enthalpy changes in the media but it could not be found in lower (acidic) pH ranges.

Isoelectric point (pI) is the pH of a protein at which a molecule clenches neutral charge. Since it requires sophisticated instrument to derive the pI of the GR experimentally, the sequences from database was used as input in order to determine pI of the protein. Thus, the pI of GR was predicted to be 5.52±0.41 based on *insilico* analysis from five sequences of *Nostoc* sp (Table 10). The pI of GR from *Nostoc* sp. was observed to be lower than turtle liver (7.3); mouse liver (6.4); in horse liver (~6.0) and human GR (6.75); but it is similar to scots pine (4.7); A. *thaliana* (4.8); pea (4.9); barnyard grass (5.9). But, pI of *Nostoc* lies between the plants and animals but it was higher than *Anabaena* (4.02) (Madamanchi *et al.*, 1992; Akihiro *et al.*, 1993; Hakam and Simon, 2000). The pI of GR from *Nostoc* sp indicated that it could be strongly acidic in nature results in the stability of enzyme found to be indirectly proportional to the pH.

Enzyme activity in various temperature showed that the optimum activity of *Nostoc* sp. 701 GR found to be up to 55°C followed by the enzyme activity was declined while increasing the temperature to 60°C (Fig. 12). Therefore, *Nostoc* sp. 701 GR showed broad range temperature stability since its activity observed from 25°C-55°C which was similar to wheat GR, revealed that GR is relatively resistant to low and high temperatures (Lamotte *et al.*, 2000). Moreover, GR denaturation temperature was reported to be 65°C in *E.coli* of both mutant and wild-type enzyme. Therefore, the rapid decreasing enzyme activity when the temperature swift to 60°C revealed that denaturation temperature could be similar to *E.coli* (Rescigno and Perham, 1994). The similar enzyme activity was reported in sheep and bovine liver at 60°C and 50°C respectively (Ulusu and Tandogan, 2007).

GR enzyme activity is strongly dependent on ionic strength of a buffer which plays a crucial role in biological activity. Generally, ionic strength of GR activity known to be 100mM in plant sources (Serrano *et al.*, 1984; Connel and Mullet, 1986; Tanaka *et al.*, 1988;
Madamanchi et al., 1992; Takeda et al., 1993) despite baker’s yeast GR activity expressed in 50mM (Yu and Zhou, 2007). However, the wheat GR activity was reported to be expressed in least buffer-ion concentration (25mM) that could be a remarkable evidence for stability of GR in least buffer-ion concentration (Hausladen and Alscher, 1994). A similar activity of GR can be found in P.falciparum and sheep liver that suggested the biological activity of the enzyme lied within the range of 200mM (Lamotte et al., 2000; Ulusu et al., 2005). Thus, the broad range of GR activity was observed between 50 and 150mM buffer-ion concentration of Tris-HCl for Nostoc sp. 701 (Fig. 13), which was consistent with previously reported buffer-ion concentration of optimum GR activity in various biological sources.

Heavy metals and metal ions influence on the enzyme activity due to the effect on conformational changes that abort the enzyme-substrate and cofactor affinity (Tandogan and Ulusu, 2010). Therefore, the eight metal compounds were tested against GR activity, four of them showed significant inhibition of GR activity relatively 100% in lower concentration such as HgCl₂, ZnSO₄, CuSO₄ and AgNO₃ their ions are Hg²⁺, Zn²⁺, Cu²⁺ and Ag³⁺ respectively, remaining metal ion such as, Mg²⁺ and Ca²⁺ could not be observed remarkable inhibition (Table 8). Such a similar inhibitory activity was reported from unicellular cyanobacteria (C. reinhardii) by Hg²⁺ and Cu²⁺ ions through branching mechanism (Takeda et al., 1993). In spinach, inhibition of GR was reported at 0.1mM of Zn²⁺ and it was due to interaction between active site thiol groups, thus the similar inhibitory effect was found as availability of Cd²⁺ ion (Schaedle and bassham, 1977; Halliwell and Foyer, 1978; Alfonso et al., 1993). Heavy metals and metal ions influence the enzyme activity that inferred to be the effect over geometry of GR which could affect the affinity of substrate-enzyme and cofactor dissociation. Therefore, inhibition of Hg²⁺, Zn²⁺, Cu²⁺ and Ag³⁺ ions were inferred to be targeting on active thiol center in GR.

Molecular mass (M₉) of the enzyme is an important factor for its catalytic activity. Therefore, Nostoc sp. 701 GR was detected in 12% SDS-PAGE by loading 15 µg of enzyme sample which showed that relative molecular mass (M₉) of a subunit ~53KDa (Fig. 14). Molecular mass of GR from Nostoc sp. 701 was greater than bacterial GR protein (49-50KDa) (Foyer et al., 1991; Mittl and Schulz, 1994) but it was similar to C.reinhardtii and S.cerevisiae (Takeda et al., 1993; Yu and Zhou, 2007) as animals and human GR reported to be 60KDa is greater than Nostoc sp. (Al-senaidy, 2010). In contrast to above, lowest M₉ of GR protein was
reported from barnyard grasses thus the native molecular mass 98KDa and 44KDa for each subunit of the dimer (Hakam and Simon, 2000). However, the $M_r$ of GR of Nostoc sp. 701 is similar to computationally predicted value from the selected five sequences.

The catalytic reaction is known to be ping-pong and sequential branch order mechanism which is depending on time and concentration of GSSG and NADPH (Ulusu and Tandoğan, 2002; Al-senaidy, 2010). Enzyme GR was reported for its high affinity over the substrate GSSG and coenzyme NADPH (Wallen et al., 2008). Substrate constant ($K_m$) and velocity ($V_{max}$) of the enzyme were estimated to be 0.058mM and 0.24µmol l$^{-1}$ min$^{-1}$ ml$^{-1}$ for GSSG and 0.0181mM and 0.192µmol l$^{-1}$ min$^{-1}$ ml$^{-1}$ for NADPH (Fig. 15 a and b), which was similar when compared to yeast, diatom, sheep liver GR and other plant and animal enzyme sources (Ulusu et al., 2005; Yu and Zhou, 2007; Arias et al., 2010; Can et al., 2010). Therefore, the affinity of substrate was higher though endothermic factors, previously reported such as pH, temperature and ion concentration are highly influenced over the enzyme geometry as modified this is because, electrostatic and hydrophobic forces exert higher preference on the substrate GSSG and NADPH.

Catalytic turnover ($K_{cat}$) of an enzyme is otherwise called “enzyme catalytic power” on the substrate as catalytic efficiency ($K_{cat}/K_m$) interrelated to $K_{cat}$, which can be defined as amount of substrate to be converted into product per minute of time per catalytic site. $K_{cat}$ from Nostoc sp. 701 GR was estimated to be $2.9243\times10^8$ M$^{-1}$ min$^{-1}$ and $K_{cat}/K_m$ of $5.04\times10^8$ M$^{-1}$ min$^{-1}$ for GSSG while NADPH estimated to be $2.339\times10^8$ M$^{-1}$ min$^{-1}$ and $12.92\times10^8$ M$^{-1}$ min$^{-1}$ (Table 9). But, yeast GR is reported to be of $5.4\times10^4$ M$^{-1}$ min$^{-1}$ turnover number for NADPH which is 2 fold higher than Nostoc GR as $2.78\times10^4$ M$^{-1}$ min$^{-1}$ for GSSG and was known to be similar with the investigation of Yu and Zhou, 2007. In contrast to Nostoc GR, $K_m$ was 190µM for both substrates but $K_{cat}/K_m$ were reported to be different i.e., $3.17\times10^6$ M$^{-1}$ min$^{-1}$ for GSSG and $1.36\times10^7$ M$^{-1}$ min$^{-1}$ for NADPH in P. tricornutum (Arias et al., 2009). A similar kinetic efficiency on NADPH and GSSG were reported to be $1.68\times10^7$ M$^{-1}$ s$^{-1}$ and $4.85\times10^8$ M$^{-1}$ s$^{-1}$ in rat kidney GR (Can et al., 2010). Therefore, the catalytic efficiency of NADPH was greater than that of GSSG which thought to be the enzyme exhibit high affinity and catalytic efficiency over the coenzyme (NADPH) than the substrate (GSSG) in both branch and ping-pong kinetic mechanisms.
Physio-chemical properties of *Nostoc* GR from selected five sequences such as, *Nostoc* sp. PCC 7120 (BAB76667.1), *Nostoc* sp. PCC 7120 (CAA61856.1), *Nostoc punctiforme* PCC 73102 (YP_001864581.1), *Nostoc azollae* 0708 (ADI64813.1) and *Nostoc punctiforme* PCC 73102 (ACC79657.1) were analyzed which revealed similar to experimentally derived measures. Molecular mass ($M_r$) of GR for single subunit was ~50 and pI was determined to be $5.52\pm0.41$. In addition, other properties such as, Grand average of hydropathicity (GRAVY) index ($-0.068\pm0.025$) revealed that GR is a globular soluble protein. *In silico* prediction from the five sequences of *Nostoc* revealed that it is highly stable at neutral pH thus, the aliphatic index of the protein showed $91.8\pm3.37$. Moreover, the positively charged amino acid residues (Arg and Lys, $45\pm3$) were found to be greater than negative charged residues (Asp and Glu, $57\pm3$) in the GR protein. Instability index $25.41\pm6.24$ for five sequences (cut off $<40$) revealed that the protein molecule could be thermotolerant, however *Nostoc* sp 7120 (BAB76667.1) was predicted to be highly stable than other four species. Molar extinction coefficient was predicted to be $30843\pm5862$ for five sequences despite *Nostoc* sp. PCC 7120 (BAB76667.1) and *Nostoc* sp. PCC 7120 (CAA61856.1) showed strong absorptivity at 280nm (Table 10). These properties predicted from computational analysis for GR of *Nostoc* species such as physical and chemical properties were similar when compared to experimentally obtained results (Jiang *et al.*, 1995).

Another important property of the GR is the secondary structure of the sequence such as alpha helix, beta strand and coils. The secondary structural elements alpha helix and random coils are equally distributed throughout the sequences (five) thus the elements were occupied relatively $65\%$ in the sequences (Table 11). About $25\%$ of the beta strands occupied in the sequences and remaining $10\%$ of the sequences avail in the turn regions thus the similar report could be found in *Spirulina* GR sequence (Rendon and Hernandez, 2001). Interestingly, similar frequencies of secondary structure elements is distributed in other oxidoreductase family enzymes for example, dihydrolipoamide dehydrogenase and trypanothione reductase (Grau *et al.*, 1981; Greer and Perham, 1986; Cui *et al.*, 1989). Sarma *et al.*, (2003) reported that it is hardly possible to find any secondary structural discrepancies even in high resolution crystal structure of this family protein. The functionally important motifs were buried in the coil regions thus the five functional motifs exist over the sequence i.e., FAD and NAD binding motifs, catalytic disulfide active center, central interface and nucleotide dimerization site (Fig.
16). It lies in coil-extend conformations but the interface domain is found in α-helix which forms a dimer (Yu and Zhou, 2007). Though FAD binding motif (GA(G)GSGG) was conserved all flavin dependent oxidoreductase enzymes, an alanine replaced by glycine at 11th position in the motif which is reported in human and E.coli but it is identical to Anabaena and Spirulina GR. Secondary structure of NAD-binding motif is identical as exists in human, P.falciparum and E.coli but it is in contrast to FAD motif [GSGYIG(A)] because a glycine replaces the alanine at 178 of C-terminal of the motif (Danielson et al., 1999). However, two loop insertions are reported only in P.falciparum, one of which is found in N-terminal region ranging from 126-138 while the other is located at sub-domain of central interface domain segment ranging from 318-350 for strong stability to the protein and support to FAD binding (Gilberger et al., 2000).

BLAST analysis showed that the sequence of Nostoc sp. PCC 7120 (BAB76667.1) was highly homologous since it shared average of 73% homology to other sequences from Nostoc species. In contrast, Nostoc punctiforme PCC 73102 (ACC79657.1) sequence shared least homology i.e., average of 48% to the other four sequences (Table 12). The overall average identity was observed to be 64.2% in the five sequences from local pairwise alignment. The average global identity was 33.2% revealed that it could be a class specific homology observed in five sequences of Nostoc GR which was predicted from clustalW analysis (Fig. 17). As a result, three groups were found in phylogenetic tree with respect to the global alignment in which all species diverge from Nostoc sp. PCC 7120 (BAB76667.1) which formed a separate group consisting of two subgroups. From the multiple sequence alignment for phylogenetic analysis, four conserved motifs were observed of which three motifs were previously reported for specific function such as, FAD-binding, NAD-binding and catalytic active center but the remaining motif TADKILIAVG ranging from 129 to 139 located behind the central catalytic domain that encapsulated adenosine moiety of the flavin molecule (Fig. 32). On the other hand, FAD is found to be buried in the hydrophobic core of the motif that existed in all oxidoreductase family enzymes (http://urchin.nidcr.nih.gov/cgi-bin/blastREF.plx?tig=GLEAN3_25989) e.g. Rhodospirillum rubrum (http://www.uniprot.org/uniprot/Q2RWK9). The dynamic simulation is evidenced to the TADKILIAVG motif that RMS fluctuation is observed to be similar to the FAD-binding residues fluctuation that found between 0.1nm and 0.15nm, suggesting that the motif is tightly bound with FAD molecule (Fig. 41).
The phylogenetic analysis showed that *Nostoc* sp. PCC 7120 (BAB76667.1), *Nostoc* sp. PCC 7120 (CAA61856.1) and *Nostoc punctiforme* PCC 73102 (YP_001864581.1) belongs to group-I which formed a cluster found in a clade-I in which *Nostoc* sp. PCC 7120 (BAB76667.1), and *Nostoc* sp. PCC 7120 (CAA61856.1) revealed to be (bootstrap value, 0.0865) closely related in the tree hence it was classified in subgroup-I. But the sequence *Nostoc punctiforme* PCC 73102 (YP_001864581.1) was remarkably diverged (0.056) from the subgroup-I hence it was classified in subgroup-II whereas a subgroup-I and II forms a cluster (clade-I) revealed that they are evolutionarily close related. The fourth sequence *Nostoc azollae* 0708 (ADI64813.1) found to be in clade-II by the evolutionary cut-off of 0.267 that it could be diverged from the group-I and it was classified in another branch i.e., group-II and finally, *Nostoc punctiforme* PCC 73102 (ACC79657.1) formed in outgroup since it was remarkably diverged from clade-I and II which was observed by cut-off value of 0.374 (Fig. 18).

This phylogenetic study was extended to predict the evolutionarily important trace residues among the five *Nostoc* sp. sequences, and three functional domains were predicted whose pattern seemed to be identical i.e., GGTCVXXGCXPKK, GXGXGG, and GXGXXG (X, substitution is allowed) along with a novel motif LTPVAI (329 to 335) was identified in *Nostoc* species. Thus, the novel motif was noticed over the head group of FAD (isoalloxazine) observed in 3D structure of GR that could be responsible for tight binding of FAD so that it is located behind the si-side of FAD in central catalytic domain. Moreover, the RMS fluctuation plot was an evidence for hydrophobic tight binding of FAD molecule, thus the fluctuation of these functional residues seemed to be directly proportional to FAD which was similar to other FAD binding motif residues. However, a remarkable motif DSSIEGPGYSTM was predicted based on the relative importance of amino acids in the sequences that was found in long loop region ranging from A274 to I308. It is highly flexible and their conformational changes were observed to be RMSD of 0.4nm, and found between the two nucleotides binding domains of the C-terminal evidenced from simulation and dynamics (Fig. 19).

In addition to performing a similar analysis by different approach, about 58 existing cyanobacterial GR sequences were retrieved from UniProt database and multiple sequence alignment (MSA) was performed under clustalW. The MSA was submitted for evolutionary trace analysis within 10 partitions (Innis *et al.*, 2000). As a result, the evolutionary trace analysis (Trace Suite II) for 58 species were segregated into seven groups in the phylogenetic
tree that obtained from the same analysis and was found to be similar to previously annotated tree within the five species of *Nostoc* (Fig. 20). Moreover, a similar motif (LTPVAI) was observed previously, revealed that conservation of the motif due to evolutionary pressure on FAD binding. Thus, the other functional domains GXGXGG, GGTCVXXGCXPKKLXV and GXGXI[G/A] were proceeded up to partition 10 (Fig. 22) and a conserved motif TADKILIAVGG (ranging from 130-140) motif could be predicted up to group- IV in the phylogenetic tree among VII groups, however these motifs found several substitution in the alignment (Fig. 23).

Sub-cellular localization of the GR activity was performed using various software tools thus the tools were designed with unique algorithmic features hence it was used for robust prediction. Since every tool produced results from its own features that enabled to conclude localization feature of the protein to become concrete prediction (Vijayaraj and Elasri, 2007). ProCompB analysis revealed that the protein activity could occur only in cytoplasm as the marginal values for other part of the cell found to be negotiable (Table 13). On the other hand, PSORTb analysis suggested that the similar localization of the protein activity, of which SCL-BLAST analysis showed 100% match to glutathione reductase (No. 20141393) for cytoplasmic region which substantiated the localization of biological activity of the enzyme (Table 14). TargetP was performed on the basis of N-terminal signaling for sub-cellular localization of GR activity which showed that the absence of signals for mitochondria and chloroplast, revealed to be cytoplasmic activity (Table 15). SLP-local prediction suggested that preference for cytoplasmic activity while the negative score for periplasmic region indicates that GR activity could not be found in periplasm (Table 16).

Pattern, motif and domain architecture plays an important role to define the protein structure, function and classification. PPsearch predicts the conserved pattern by comparing PROSITE database sequences. Therefore, PS00076 pattern GGxC[LIVA]x(2)GC[LIVM]P (Table 17) is highly conserved, which is responsible for GSSG affinity and recognition besides the pattern could be found in active disulfide catalytic center (http://prosite.expasy.org/PS00076). CDART predicts the domain architecture based on the evolutionary distance of the protein sequence against NCBI-Entrez protein database. By this method, three conserved domains were predicted which are known to be protein superfamily, two of which revealed that Rossmann-fold (cl09931) that responsible for nucleotide binding which exhibited GXGXXG
architecture located at N-terminal while the third motif existed in dimerization domain ranging from 350-450 at C-termini of the sequences (Fig. 24). The predicted architecture belongs to class-I and class-II oxidoreductase family proteins thus the similar pattern could also be found in pfam database (pfam02852). Moreover, the predicted domain was observed in other dehydrogenase enzymes of oxidoreductase family flavoenzymes. For instance, dehydrogenases are involved in glycolysis and citric acid cycle (http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=199420). MEME performs the prediction based on position-dependent letter-probability matrices whereas three highly conserved domains were predicted. From this analysis, the sequences were predicted class-I active site, nucleotide binding domain (FAD/NAD) and nucleotide dimerization domain that exhibited from 20-70, 130-150 and 300-350 respectively (Fig. 25), which is similar to previously predicted motifs, pattern and architecture.

Prediction of the protein function is significant for understanding the biological role of protein molecule which is often described as assembly of several domains defined the ultimate function of a protein. The computational techniques facilitated to predict the GR function by using sequence as an input. Consequently, SMART analysis revealed two functional domains found in the protein sequences ranging from 6-318 which are responsible for the binding of FAD and NAD, belongs to class-I and class-II oxidoreductases family proteins. Thus, the domain PF07992 was predicted as FAD and NAD binding which occurred in the N-terminus of protein sequence that was predicted in pfam database (Table 19). FAD plays a role as cofactor and also acted as an electron acceptor from NADPH that transfers the electron to disulfide active center which is found in the alpha/beta Rossmann-fold conformation (Marchler-Bauer et al., 2011). Thus, the same prediction revealed an overlapped domain ranging from 168-249, classified under PF00070 that refers to the binding motif for NADPH that plays a role as coenzyme it reduces the enzyme through both ping-pong and branching order mechanism during the reaction. Finally, a domain PF02852 revealed that it is dimerization domain found at C-terminus of the sequences. The domain is responsible for reduction, oxidation process and post-translational modification of the protein. Apart from that, several enzymes are involved in the metabolic pathways that were annotated in map00480 and map00251 for glutathione metabolism and glutamate metabolism respectively, which were retrieved from KEGG and DBGET metabolic databases. However, it was predicted for only 3 out of 5 sequences (Fig. 26,
27 and 28). In this network, two sorts of pathway mechanisms were observed which one network involved in the synthesis of GSH while the other involved in the homeostasis of thiol redox system that utilizes the reduced GSH as reducing equivalent for enzymes function (Zhang and Kirkham, 1996) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=199420).

InterproScan analysis revealed that the predicted results were consistent to SMART. However, the motifs IPR001327, IPR004099, IPR006324, IPR012999, IPR013027, IPR0016156 and IPR023753 were annotated from InterPro database, which are all similar with respect to function (Fig. 29). But, the IPR001327 motif could not be predicted only in Nostoc sp. PCC 7120 (CAA61856.1) sequence while the remaining six motifs were observed in the same sequence. The first IPR001327 motif is similar to PS00070 which is approximately ranging from 170-230 that predicted as NAD binding domain of pyridine nucleotide-disulfide oxidoreductase family. IPR004099 is equivalent to PF02852 motif that was predicted as pyridine nucleotide-disulfide oxidoreductase-dimerization domain observed in C-terminal from 350-450 of five protein sequences. IPR006324 is similar motif retrieved from TIGR database (TIGR01424) which described entire sequence responsible for GR function. IPR012999 was a small adorable portion in the sequence which is responsible for GSSG binding also called disulfide active center/domain that was found in PROSITE database (PS00076), facilitated to protein classification. IPR013027 was similar to PRO0368 pattern that scattered over the sequence and attributed to FAD dependent pyridine nucleotide-disulfide reduction.

IPR0016156 motif also existed in protein super family (SSFSS424) database that it defined as FAD/NAD linked reductase found in end of C-terminal region. IPR023753 domain is responsible for FAD/NAD(P) binding which covered more than half of the sequence similar domain existed in PF07992 of pfam database (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR013027). Therefore, the sequences from five Nostoc species were observed for functional motifs revealed that catalysis of oxidized glutathione to reduce form, which are all localized in consistent region of the five sequences. Moreover, topology of the protein was observed from PRED-CLASS analysis which revealed to be globular soluble protein as it was determined by based on their amino acid properties in the sequence.

The above comprehensive analysis enabled to isolate a very homologous sequence for structure modeling, docking and dynamic simulation from the five species of Nostoc. The
above sequence analysis revealed that all sequences were consistent with previously reported GR from other sources. However, *Nostoc* sp. PCC 7120 (BAB76667.1) shared higher homology with other sequences in addition, that plays a pioneer role over other sequences hence its sequence alone was identified for further structure based analysis. Therefore, the physical properties of the protein sequence was, length of the sequence is 459, Molecular mass \((M_r)\) 50, pi 5.5, GRAVY index -0.076, total charged residues (46+ve) and (58-ve), instability index 22.64, extinction coefficient of 27515 (Table 10). The instability index revealed that the protein thermally was stable when compared to other protein sequences. On the other hand, three functional domains were predicted such as, FAD-binding domain, NAD binding domain and catalytic active center (GSSG) in all five sequences which were similar to microbes, plants and animal GR proteins. Two novel motifs were predicted in *Nostoc* sp. PCC 7120 (BAB76667.1) protein sequence such as TADKILIAVG and LTPVAI which was also observed other *Nostoc* species. A further study was attempted to characterize the novel domains and functional relationships of the protein with cofactor (FAD), coenzyme (NAD) and its substrate (GSSG) through the modeling and simulation process using the sequence of *Nostoc* sp. PCC 7120 (BAB76667.1).

The tertiary structure of GR in *Nostoc* sp. PCC 7120 (BAB76667.1) was modeled based on homologous proteins using modeller9.10. In order to model the target sequence, two templates were identified by protein-BLAST analysis against the PDB database thus the templates were 3O0H (A) and 3DJG (X) used to model the sequence, moreover latter template was considered for ligand translation thus the ligands are FAD and NAD. The translation of ligand was mandatory in this modeling because of docking with multiple ligands in a single active site pocket was found to be complicated at once. Moreover, superimposing one and another ligand also lead to overlap in the protein molecule. Therefore, the target sequence is to be modeled with the ligand, one of the template 3DJG (X) was defined to imitate the ligand coordinates while modeling (http://salilab.org/modeller/methenz/andras/node16.html). The appropriate model was isolated based on the minimum DOPE score value among the 15 models generated from modeller. Thus, the isolated model showed with DOPE score of -102754.32 which was considered as best model for further energy minimization and validation (Table 24). Energy minimization was performed in Swiss-PdbViewer and its optimized protein energy was -5071.827kJ mol\(^{-1}\) followed by PSVS validation (Fig. 30). Therefore, the modeled protein
structure showed 91% of the amino acids adopted in most favored region and 7.4% residues found to be in allowed region as observed from PROCHECK analysis (Fig. 31). In addition, Richardson’s statistics substantiated the optimized model with least outlier amino acid residues. Hence, it was considered as final model since its lower R-factor (<0.25) and the quality model revealed more than 90% of the amino acids accommodated in the most favored regions (Fig. 32).

The theoretically derived model was analyzed with ten homologous high resolution crystal structures of similar family proteins i.e., oxidoreductase. The predicted model showed minimum RMS value which was close to zero with respect to the template crystal structures 3O0H (A) of 0.377 and second template 3DJG (X) showed the RMS deviation 0.993 (Fig. 30). Moreover, the remarkable similarity was observed with 2R9Z (A) which showed RMSD of 0.257 despite it was not defined as template during the modeling. The crystal structure 2R9Z was predicted as homologous structure whose identity found to be 43%, which was nearest homology to the templates and the target sequences (Table 23). Therefore, the theoretically resolved model assumed to be appropriate GR model for further docking simulation.

Thereafter, the 3D modeled structure was undergone for docking analysis that was performed in Autodock 4.2. The protein molecule was subjected to enzyme (GR)-substrate (GSSG) docking in order to find the molecular interaction between the substrate and the enzyme. To perform docking, the enzyme must be defined in its active pocket to superimpose the substrate. Hence, two sorts of prediction methods were carried out such as, sequence and structure based approach in order to identify the active site. Firstly, active site prediction was performed using sequence as an input; Enzyme 1D signature predicts the active site region using data-mining approach by comparing the sequence with existing Swiss-Prot database sequences. Therefore, the active site residues were predicted as GAGSG, GSGYIGTE and GGTCVIRGCVPK which belongs to FAD-binding site, NAD-binding site and catalytic active center respectively. On the other hand a pattern PSSAEE found in the dimerization domain at C-terminus (Fig. 33), such a motif is reported in plant and animal sources (Stevens et al., 1997). These active sites residues are appeared to be similar to the catalytic residues found in catalytic site atlas (CSA). CSA is a renowned database for catalytically active residues (triad) that was deposited from experimentally derived crystal structures (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/CSA/CSA_Show_EC_List.pl). Second approach was prediction
of active site from 3D structure data. CASTp predicts the active site based on space filling method from the mathematical background of weighted Delaunay triangulation (Liang et al., 1998). Total surface of the catalytic pocket of predicted model was 3117.8 Å² with the volume of 5720.3 Å³ that found in pyridine-nucleotide redox domain. A small surface area was also predicted which existed in disulfide active center that could be measured the surface radius of 783.5 Å (Fig. 34a and b). These are the catalytic residues similar to other enzymes such as, glutathione amide reductase, lipoamide dehydrogenase, thioredoxin reductase, trypanothione reductase and mercuric ion reductase (Muller et al., 1995; Gilberger et al., 1999; Picaud and Desbois, 2002; Rollins et al., 2010). As it is evidenced from CSA, the active site pocket with cavity surface was targeted for docking simulation.

Accordingly, the substrate GSSG was optimized whose chemical formula and molecular weight was C₂₀H₃₂N₆O₁₂S₂ and 612.63108 respectively, which was defined to be superimposed in its appropriate catalytic site in GR through the docking simulation (Fig. 35).

As per the docking manual instruction, the docking was performed. As a result, the binding energy and ligand efficiency was observed to be of -6.59 and -0.16 respectively, per unit of polar surface area of the catalytic pocket. Inhibitory constant (Ki) was found to be 14μM that was the minimum concentration required to inhibit the enzyme with respect to substrate. In addition to, the non-bonded interaction energies were -7.8, -6.26, 1.53, -6.21 and 7.41 (kcal/mol) calculated as intermolecular_energy, vdw_hb_desolve_energy, electrostatic_energy, total_energy and torsional_energy respectively (Table 21). Therefore, three hydrogen binding interactions were observed between the enzyme and substrate in which Lys-21 and Arg-93 interacts with GSH-II as Ala-350 interacts with GSH-I by strong and weak hydrogen bonds respectively. This docking analysis revealed that the substrate enter into the catalytic pocket by GSH-II thereafter GSH-I interact with catalytic dithiol active center. As binding to the catalytic center, the reduced enzyme transfers the electron from FAD through cysteinyl bridge (Cys42-Cys47) to GSH-I (Fig. 36). Instantly, the substrate releases GSH-II followed by GSH-I (Berkholz et al., 2008; Seefeldt et al., 2009). Thus, the similar catalytic function is reported in E.coli, yeast, Chromatium gracile and human (Petegem et al., 2007). The geometry of catalytic domain in modeled GR was analyzed after the docking simulation by comparing high resolution crystal solved structure. The comparative analysis with X-ray crystal structures facilitated to understand the ligand and substrate arrangements in the enzyme before,
during and after the catalytic reaction (Fig. 35). Moreover, the modeling and docking simulation could be an evidence for geometrical arrangement of ligand and substrate in GR of *Nostoc* sp. According to the Fig. 35, modeled structure was superimposed with known crystal structures in order to analyze the catalytic function of the GR of *Nostoc*. The investigation revealed that reduction of FAD bound enzyme releases the oxidized NADP\(^+\) as a product in order for binding of GSSG, where Tyr-176 played a gate mechanism for binding of NADPH (Yu and Zhou, 2007; Berkholz *et al.*, 2008).

The enzyme consisted of five domains in each subunits namely, FAD-binding domain, NAD-binding domain, catalytic active center, subunit interface and nucleotide dimerization domain which cherished with strong hydrogen bonds (Sarma *et al.*, 2003). These functional domain arrangements appeared to be similar to *E.coli*, yeast and human GR (Yu and Zhou, 2007). FAD could be observed over GXGXGG fingerprint which is conserved motif existing in many oxidoreductase family enzymes (Danielson *et al.*, 1999). This fingerprint is known to be responsible for selection of cofactor and while remaining part in the FAD binding domain ranging from 1 to 152 that interacts to FAD by strong hydrophobic interaction in *Nostoc* it seemed to be similar region found in human GR upto Gly-157 for the flavin nucleotide binding (Berkholz *et al.*, 2008). The flavin nucleotide bound as an elongated conformation in which isoalloxazine ring facing towards NADPH said to be *re*-face that interacts with nicotinamide moiety of NADPH during the charge transfer which is referred to *si*-face that located behind the *re*-face of the isoalloxazine moiety. Therefore, the phosphate and adenosine extension from the isoalloxazine could tightly interact with non-covalent bonds of N-terminus polypeptide chain (Pai and Schulz, 1983). Isoalloxazine ring is buried in the central catalytic domain which directly interacts with disulfide bridge (Cys42-Cys47) between the cysteiny1 residues that similarly existing in *E.coli* but it is equivalent to human cysteiny1 residues (Cys58-Cys63) (Picaud and Desbois, 2002).

The second binding partner NADPH is conceivably established in \(\beta\alpha\beta\)-fold of the NAD-binding domain for oxidation and reduction reaction (Rescigno and Perham, 1994). Nicotinamide is a head group of the NADPH which found direct opposite to *re*-face in close contact to the isoalloxazine molecule by the distance of 0.34nm consoled for easy proton transfer to reduce the enzyme (Mittl and Schulz, 1994). In this part, the tyrosine stalks the nicotinamide moiety to regulate the red-ox reaction therefore, Tyr-176 playing a vital role in
dissociation of NADP$^+$ after the oxidation reaction. The similar functioning tyrosine exists in *E. coli* and human at 177$^{th}$ and 197$^{th}$ position respectively, in the polypeptide chain that remarkable movement was reported to be 0.64nm during the catalytic reaction (Pai *et al.*, 1988). Thus, the compression of Tyr is further supported by overlapping vander Waals radii in the active site, hence flavin N5 interacted with C4 of nicotinamide by the distance of 3.29 Å and C4a of flavin interacted with Cys47-SG between the distance of 3.29 Å as well. Hence FAD found to be tightly fixed between the two molecules which enhance the catalysis by stereo-electronic effects. Moreover, the flavin molecule (N5) is pushed about 0.3 Å towards the nicotinamide pocket (Berkholtz *et al.*, 2008). The NADPH ends with adenosine moiety that lied in elongated conformation and the bindings of NADPH occurred at N-terminal and extended to C-terminal of the chain in which GXGXXG plays a vital role in binding and selectivity of the nucleotid. A comparative analysis report revealed to origin of two nucleotide binding domain (FAD and NADPH) due to gene duplication therefore the discrepancy was not found in the sequence similarity for GR protein (Savvides and Karplus, 1996).

The third important domain, binding of substrate-GSSG occurred at catalytic disulfide active center. Due to the steric and electrostatic potential of GR enzyme, it attracts GSSG over the isoalloxazine ring by Lys-21 and Arg-93 which interacted by hydrogen bond (Table 21). GSSG can be sub-divided into two glutathione i.e., GSH-I and GSH-II fused together that formed four arms. The first arm of GSH-I contacts to the Cys-42 residue when electron transfer occurred through Cys-47 from re-side of isoalloxazine (Berkholtz *et al.*, 2008). On the other hand, His-447’ and Glu-452’ acts as interface between FAD and GSSG from C-terminus of adjacent subunit it is similar to His-467’ and Glu-472’ in human. Thus, Tyr-114 (corresponding to Tyr-99 in *E.coli*) is occupied on the disulfide bridge of the substrate GSH-I it forms hydrogen bond in human but it could not be existed in *Anabaena* GR and *Nostoc* as well (Mittl and Schulz, 1994; Danielson *et al.*, 1999). As binding of GSSG proceed three reaction steps, i) binding of GSSG to reduced enzyme, ii) formation of mixed disulfide and subsequent release of GSH-II and iii) release of GSH-I. GSH-I in the enzyme complex bound much more tightly than GSH-II (Fig. 36a and b). Intact electron transfer between flavin and thiolate anion make a move towards sulfur atom of GSH-I about 0.15nm (from 0.36nm to 0.5nm) due to the opening between Cys-42 and Cys-47, it was similar to human GR Cys-58 and Cys-63 (Pai *et al.*, 1988; Mittl *et al.*, 1994; Sarma *et al.*, 2003; Berkholz *et al.*, 2008).
During the catalytic activity, four intermediates have been proposed to first reductive half-reaction in which the enzyme is reduced by two electrons ($E_{H_2}$) and the electron transfer occurred between NADPH and FAD (Fig. 38). Electron transfer of second oxidative half-reaction revealed that binding of GSSG oxidize the enzyme ($E_{ox}$), simultaneously transferred electrons break the disulfide bond between substrate and release the product as GSH-I and GSH-II. These reaction mechanisms are reported to be ping pong mechanism while the branching sequential order occur vice-versa (Mittl et al., 1994; Savvides and Karplus, 1996). The complete kinetic mechanism in the reaction is initiated by binding of both substrates (GSSG) and NADPH as reducing equivalent (coenzyme) which were separated by the prosthetic group FAD. Therefore, NADPH binds on re-side of the flavin ring as the GSSG binds on si-side thereby the Tyr-176 and His-98 respectively act as regulatory partner. Tyr-176 undergoes large conformational change and His-98 acts as base catalyst when the reaction proceeds (Pai et al., 1983). The movement of His-98 upon GSSG binding by the way of phenolate oxygen involved in hydrogen bonding to the main chain N of Cys-I and Gly-II thus the phenolate promotes the protonation of GSH-II by the acid catalyst, His-447 of adjacent subunit. The protonation prevented the reversibility of the mixed disulfide formation between the interchange of thiol Cys-58 and GSH-I, yielding glutathione (GSH-II) as a free thiolate product and subsequent release of GSH-I, as the ultimate irreversible product (Siegel et al., 1998; Picaud and Desbois, 2002; Yu and Zhou, 2007).

After docking, a cross validation was performed between one and another including templates of ten crystal structure. Therefore, the model (target) sequence and the crystal structure sequence identity was observed to be 44% for 3O0H (A) and 43% for 2R9Z (A) were observed, rest of the other structure sequence identity were observed to be 41% (Table 23). The top 10 similar structures were analyzed each other to observe their variance between modeled structure and the crystal structure. The modeled structure showed RMS for its template were 0.377 for 3O0H (A) and 0.993 for 3DJG (X), in addition to the crystal structure 2R9Z (A) showed significant match to the modeled GR structure with RMS of 0.257 which revealed that distinguishable changes could not be predicted after docking simulation (Fig. 40). Therefore it could be an evidence for the higher resolution of theoretical model that further processing in for dynamic simulation.
The best docking conformation of enzyme-substrate complex was used as seed to compute the stability and energy profile of the complex from molecular dynamic simulation (MDS). However, the docking energy results were absolute that calculated from Molecular Mechanics which strongly depend on the size of each molecule. Thus, absolute energy values are meaningless while the energy difference is found between ligands bound and free-state, on the other hand, an enzyme activity is widely regarded as dependent on protein flexibility. This type of docking methods only provides information about the relative ligand binding energy (Iribarne et al., 2002). Unlike, Molecular dynamic simulation approach is much more accurate than direct scoring functions. MDS can provide a closer representation of biological function of molecules in reality i.e., enzyme-substrate simulation with explicit water solvent model. The well known fact, most accurate protein structure could be obtained from X-ray crystallography but the simulation programs will not mimic the crystalline conditions. To date, computer simulations are largely limited to simulation times of no longer than \(~10^{-8}\) sec, i.e., nanoscale which is regarded as neutron scattering dynamics (Daniel et al., 2003; Fan et al., 2005). This study intended to elucidate the structural information about the enzyme complex and their kinetic energies with respect to conformational changes of the protein-ligand and substrate molecules during the catalytic reaction using dynamics and simulation program, GROMACS.

MDS in GROMACS is included serial step simulation package by which the coordinate of enzyme-substrate complex was prepared using GROMOS96 43a1 force field parameters, according to the manual instructions (www.gromacs.org). The coordinate system was solvated in explicit water environment followed by addition of 14 positive Na\(^+\) (sodium ions) for charge neutralization and the system energy was minimized in two steps using steepest descent methods and conjugate gradient. Therefore, the simulation system was reduced that could be observed from the potential energy, maximum force and norm of force were estimated to be -1.5157776e+06kJ mol\(^{-1}\), 9.8606854e+02 on atom 4381 and 2.4466408e+01 respectively, in 673 steps of steepest descent method followed by conjugate gradient method converged potential energy of -1.5156308e+06kJ mol\(^{-1}\), maximum force of 3.9756107e+02 on atom 1240 and norm of force of 3.8671646e+01. The system equilibrated in two phase with respect to volume and pressure in 100ps at 310K temperature by the similar parameters. The average total system energy was calculated to be -1.10346e+06kJ mol\(^{-1}\) and -1.10456e+06kJ mol\(^{-1}\) with respect to constant volume and pressure equilibration. The 5ns molecular dynamics simulation was
performed in GROMACS 4.5.4 which were taken several hours to few days to complete the simulation.

Thus, the equilibrated enzyme complex in SPC (Simple Point Charge) model was subjected to $2.5\times10^6$ step molecular dynamic simulation. Whereas, first 300ps was discarded and remaining part of the simulation time scale was considered for analysis. Based on MD trajectories, the conformation changes and energies were analyzed thus, RMSD was attributed to protein, FAD, NADPH and GSSG; RMSF for protein backbone; hydrogen bonds analyzed with respect to enzyme-substrate complex, FAD, NADPH and GSSG; potential energy, kinetic energy and total energy. Hence, the protein-ligand complex average energy minimum was observed such as, potential energy of $-1.33400e+06$kJ mol$^{-1}$, kinetic energy of $2.25261e+05$kJ mol$^{-1}$ and total energy of $-1.10874e+06$kJ mol$^{-1}$ as the constraint reach at 0.0 of RMSD at 310K (Fig. 41a, b and c). Therefore, the potential energy is known to be lowest configuration energy which restores the native energy of the protein. The enzyme complex attained its energy minimum between 2.25 and 2.75ns, revealed that stable conformation of the enzyme complex was found to be as the complex reached its potential energy of $-1.33584e+06$kJ mol$^{-1}$ (2.2ns) thus the protein retained its global minimum RMSD of 2.36241e-5 which was seemed to be very close to 0.0 in order to proceed catalytic reaction. In this study, the initial equilibration found to be reached after 300ps thereafter the system energy slowly marched to lower energy which is observed in contrast to Hikichi et al., (1995) dynamic simulation revealed that the system began to equilibrate after 40ps, who suggested that the total energy of the system to be of $-1400$kJ mol$^{-1}$ in 200ps simulation. As a result, the enzyme complex observed to be approximately 800 times stable as the enzyme is bound with its cognate ligand and substrate molecules.

Iribrane et al., (2002) evidenced to occurrence of conformational changes of GR, due to the effect of steric and electrostatic potential for binding of substrate and ligand molecules corresponding active sites. The RMSD for GR complex was observed to be 3.0 to 4.2 Å of Cα-atoms which seemed to be higher than human GR of unbound ligand (Hikichi et al., 1995). Moreover, secondary structure differences were found to be about 5.0 Å for α-helices and β-sheets. Accordingly, back bone of the enzyme complex RMSD revealed that the conformational changes occurred due to non-ligand binding residues rather than ligand and substrate binding motif in the solvent system. After 300ps, the apparent RMSD of the protein
was 0.354nm however rapid changes occurred till 1.5ns then short equilibrium found up to 3ns which revealed equilibration of the enzyme complex showed between 0.3 and 0.4nm. However, maximum RMSD (0.42) occurred at 3ns indicated that hydrophilic residues are highly flexible in order to interact with solvent molecules in the ligand bound complex (Fig. 42a). The RMSD is correlated with the hydrogen bonds (hbs). Therefore, the enzyme complex was traced trace the hydrogen bonding (hb) interaction between the residues in the catalytic pocket in MD simulation and the ligands clearly indicated at three places were found to be maximum number of hbs 20 at 833, 3140 and 4744ps respectively. However, a broad hbs cluster was observed over the first 20 hbs (~0.8ns) that could significantly be considered for the protein activity in the complex structure. On the other hand, average of intact simulation for the GR-substrate interaction was observed an average of 11 hbs while the lowest hbs found at 1946ps up to 3hbs. Thus, the latter hbs in the ligand complex was considered as playing critical role in the enzyme which supports dissociation of product in the reaction (Fig. 42b).

The ligand and substrate was investigated individually by their RMSD and hydrogen bonding patterns. Firstly, RMS differences of FAD revealed that the molecule relaxed up to 500ps thereafter FAD excited to RMSD of 0.2nm in order to equilibrate in its binding domain and decreasing RMS differences showed that FAD moved towards equilibrium. Moreover, an influential RMS differences found about 0.17nm at 2ns (Fig. 43a). Correspondingly, the hbs pattern was observed especially at 1333 and 1721ps with 11 hbs of maximum interaction however 1 hb exhibited more than 13 places during the course of dynamic simulation (Fig. 43b). Thus, FAD molecule in the catalytic subspace was equilibrated to attain enzyme-ligand equilibrium which was evidenced from the movement of FAD cofactor that could be observed from hydrogen bonding plot.

The RMSD of coenzyme NADPH was gradually increased till 1.5ns revealed that NADPH is highly flexible when compared to FAD. Therefore, NADPH equilibrated between 0.2nm to 0.25nm inside the binding cavity and the equilibrium retained even after 3.0ns (Fig. 44a). In this simulation, hbs of NADPH was found to be initially decreased up to 1.0ns but it was increased as increasing time therefore hbs of NADPH acquired equilibrium that was observed to be indirectly proportional to RMSD (Fig. 44b).
The substrate GSSG showed an opposite effect after 0.5 ns when compared to FAD and NADPH and it found abrupt transition till 1.0 ns that indicated the conformational changes of GSSG in the protein depended on cofactor and coenzyme. The RMSD plot for GSSG described equilibrium in two states whereas 1.5 ns to 3.0 ns found to be of 0.2 nm RMSD however the RMSD was observed to be increased after 3.0 ns (Fig. 44a). The hbs of GSSG revealed that hydrogen bond interaction found to be increased till 1.5 ns thereafter the substrate molecule equilibrated with average of three hbs (Fig. 45b). Therefore, the GSSG was observed to be similar to FAD with respect to RMS differences and hbs as NADPH showed discrepancy with GSSG.

The Root Mean Square Fluctuation (RMSF) allowed us to study the amino acid flexibility during the catalytic activity of GR in water model simulation. According to the correspondence, solvent accessibility of nucleotide dimerization domain was found greater than 0.6 nm as GXGXGG and GXGXXG fingerprints were analyzed ~0.08 nm and 0.115-0.117 nm respectively and similar fluctuation found in disulfide active center (Cys42-Cys47) i.e., 0.08 nm (Fig. 46). Though, the functional domains were rich in Glycine residues that are not involved in hydrogen bonding rather than other residues in the domain (Peters et al., 2000). Similarly, Hikichi et al., (1995) reported that the flexibility of the Ca-atoms of binding motifs of FAD found to be 0.12 nm in human GR. In addition to, secondary structure of GR protein is reported to be 0.001 to 0.055 nm for α-helices as the 3_{10}-helices reported to be 0.14 nm thus the similar structural differences were also observed in Nostoc GR (Hikichi et al., 1995). However, restricted fluctuation was observed on motif of the catalytic domain residues that evidenced to the functional domains have higher affinity in ligand and substrate which suggest that these residues less interactive with solvent molecules. In contrast, non-ligand binding secondary structure found to be greater flexibility which supports to interact with solvent molecules which revealed the binding of substrate and release of product. Consequently, reflecting structural changes on inherent flexibility of the protein and interfacial plane are in the catalytic reaction made the C-terminal more flexible than N-terminal which suggested that rapid catalysis due to charged residues involved in the reaction (Pai and Schulz, 1983). Strong hydrophobic interactions over the FAD molecule resist the structural displacement of the domain. On the other hand, rigid motion in the binding domain revealed that other part of the polypeptide backbone involved in the regulation of ligand and substrates binding. This study is evidenced to
the rigid motion of binding domain facilitates proton transfer for breakdown of substrate into product.

The hydrogen bonds (hb) are important criteria in the catalytic reaction because it determines the conformational differences in structure which guided to enzyme complex equilibration. Dipole interaction in the intermolecular (enzyme-substrate) hbs enables the protein molecule to dissociate the product from catalytic pocket. A strong relationship between enzyme-ligand and substrate complex was elucidated with more hb interactions which was unpredicted from the docking simulation (Fan and Mark, 2004; Kesharwani et al., 2010). As a result, 20 hbs found in the active site cavity occurred at three points of 833, 3140 and 4744ps over the simulation time that was maximum number of hbs. However, the average hb was 11 between protein and ligand complex as the lowest hydrogen bonds (3) predicted at 1946ps. Though, the complex is having 20 hbs in which FAD, NADPH and GSSG were analyzed independently so that hbs were observed to be of 11, 10 and 8 respectively. Ultimately, it revealed that one atom could be formed more than one hydrogen bond interaction. Ligand and substrate were highly reactive in their respective catalytic pocket that evidenced from large difference in hydrogen bonding, subsequently ligand and substrate binding domain structural differences found to be lower as compared to other ligand unbound polypeptides (Iribrane et al., 2002; Yang et al., 2007).

This complete simulation study illustrates that the enzyme-substrate complex was highly influenced with electrostatic potential and steric forces. Thus, enzymatic reaction is due to the steric and electrostatic potential generated by the enzyme such as, enthalphic and entropic contributions guided the substrate to its active site microenvironment. Even though the other molecules with same potential, that could not promote binding to the active pocket. There are three consecutive phenomena are postulated in the formation of enzyme-substrate complex during catalytic reaction such as, i) Diffusion of the substrate towards its active site, ii) formation of enzyme-substrate complex with specific non-bonded interaction (hb) and finally, iii) catalysis of the substrate. To perform the catalytic reaction, it requires many factors in their microenvironment such as physical forces, electrostatic field, flexibility and solvation (Peters et al., 2000; Elcock, 2002; Ermakova et al., 2010). The rapid motion of enzyme-substrate complex during the simulation in early phase suggests that the catalytic process on the substrate GSSG could be cropped up till 1500ps. The above conclusion made from the observation of
simulation in RMSD, RMSF of the residues, hydrogen bonds and energy values evidenced for the catalytic red-ox reaction of GR however it requires more structural data to substantiate the current prediction. Hikichi et al., (1995) and Iribrane et al., (2002) reported to the enzyme GR attained early equilibrium in the small time scale (picoseconds) simulation for unbound state of NADPH and GSSG molecules for human GR. In contrary, this study showed that the substrate and ligand within the GR found to be reactive in the early phase. The enzyme complex was equilibrated in middle and later phase of the simulation in order to overcome the enzyme activation energy barrier for NADPH and GSSG. All the above results evidenced to migration of the coenzyme and substrate towards the catalytic pocket in the order of diffusion encounter. On the other hand, hydrodynamic interactions, electrostatic channeling, surfactant charge, dielectric moment has been postulated to drive the substrate to their respective binding sites whereas the catalytic process conquered and ultimately released the product, immediately the enzyme attain equilibrium for a while in order for next encounter (Wade, 1995; Peters et al., 2000). However, it sought further insight into domain specific simulation dynamics in order to manifest hydropathical effect on the residue functions in binding and release of product.

***