

# Structure prediction and molecular simulation of gases diffusion pathways in hydrogenase

Shanthy Sundaram<sup>1,\*</sup>, Ashutosh Tripathi<sup>1</sup>, Vipul Gupta<sup>1</sup>

<sup>1</sup>Centre for Biotechnology, University of Allahabad, Nehru Science Centre, Allahabad, U.P, India. Shanthy Sundaram Email: shanthy\_s@rediffmail.com, Phone : +919335158759 ; \*Corresponding author

Received July 20, 2010; accepted August 20, 2010; published October 06, 2010

## Abstract:

Although hydrogen is considered to be one of the most promising future energy sources and the technical aspects involved in using it have advanced considerably, the future supply of hydrogen from renewable sources is still unsolved. The [Fe]- hydrogenase enzymes are highly efficient H<sub>2</sub> catalysts found in ecologically and phylogenetically diverse microorganisms, including the photosynthetic green alga, *Chlamydomonas reinhardtii*. While these enzymes can occur in several forms, H<sub>2</sub> catalysis takes place at a unique [FeS] prosthetic group or H-cluster, located at the active site. 3D structure of the protein hydA1 hydrogenase from *Chlamydomonas reinhardtii* was predicted using the MODELER 8v2 software. Conserved region was depicted from the NCBI CDD Search. Template selection was done on the basis NCBI BLAST results. For single template 1FEH was used and for multiple templates 1FEH and 1HFE were used. The result of the Homology modeling was verified by uploading the file to SAVS server. On the basis of the SAVS result 3D structure predicted using single template was chosen for performing molecular simulation. For performing molecular simulation three strategies were used. First the molecular simulation of the protein was performed in solvated box containing bulk water. Then 100 H<sub>2</sub> molecules were randomly inserted in the solvated box and two simulations of 50 and 100 ps were performed. Similarly 100 O<sub>2</sub> molecules were randomly placed in the solvated box and again 50 and 100 ps simulation were performed. Energy minimization was performed before each simulation was performed. Conformations were saved after each simulation. Analysis of the gas diffusion was done on the basis of RMSD, Radius of Gyration and no. of gas molecule/ps plot.

**Keywords:** Molecular simulation, Hydrogenase, Modeler, Root mean square deviation

## Background:

Biological H<sub>2</sub> production linked to photosynthetic water oxidation is a promising technology that may play a major role in the future of renewable energy. The process of Hydrogen metabolism was first shown by Hans Gaffron (1939, 1940) [6]. Gaffron observed that, under anaerobic conditions, the green alga *Scenedesmus obliquus* can either use H<sub>2</sub> as electron donor in the CO<sub>2</sub> fixation process in the, or evolve H<sub>2</sub> in the light [6]. Hydrogenases constitute a family of enzymes found in certain photosynthetic microorganisms, such as green algae and cyanobacteria, which have the potential to store efficiently the energy of incident sunlight as high-energy H<sub>2</sub> molecules or catalyzes the reversible oxidation of hydrogen gas, with a maximum theoretical efficiency of approximate 13%. Current databases list more than 300 hydrogenase gene sequences obtained from bacteria and cyanobacteria species with most from *Scenedesmus obliquus* (cyanobacteria), *Clostridium pasteurianum* (bacteria), *Desulfovibrio desulfuricans* (cyanobacteria). If harnessed properly, hydrogenase and/or hydrogenase-containing organisms could be used to supply affordable and renewable H<sub>2</sub> to be used as an energy fuel, and thus solve the "supply" aspect of the future hydrogen economy. This idealistic picture is not without problems. Notably, hydrogenase's H-cluster is extremely sensitive to the presence of oxygen gas (O<sub>2</sub>), which will bind to it permanently. In the presence of O<sub>2</sub>, hydrogen production is maintained for only a few minutes before the hydrogenases become deactivated. In order to maintain a sustained hydrogen production using hydrogenase, an anaerobic environment is currently required, making hydrogenase a costly and impractical source of H<sub>2</sub>. This creates an interesting scientific problem: if the pathways through which O<sub>2</sub> reaches the H-cluster can be identified, it will be possible to create an engineered version of hydrogenase in which these O<sub>2</sub> pathways are blocked, thus decreasing hydrogenase's sensitivity towards O<sub>2</sub>.

## Classification of hydrogenases

Initially Fe-Hydrogenase was presumed to be present in a limited number of bacteria and anaerobic living protozoa [12]. The enzyme hydrogenase can be classified into five groups, based on the protein sequence homology of thirty microbial hydrogenase sequences [12]. (1) □[Ni-Fe]-hydrogenase: membrane bound, mainly for uptake of hydrogen, found in aerobic, anaerobic and facultative anaerobic bacteria [6,32]. (2) [Ni-Fe-Se]-hydrogenase: membrane bound, mainly for hydrogen uptake, found in sulphate-reducing bacteria [6, 32]. (3) Fe-hydrogenase: periplasmic, mainly for hydrogen evolution, found in strict anaerobic bacteria [6,32]. (4) □Methyl viologen factor [F-420] or NAD-reducing and soluble Hydrogenases: found in Methanobacteria and Alcaligenes. (5) labile Hydrogenase isoenzyme of *Escherichia coli*.

## Hydrogen metabolism

Photoproduction of H<sub>2</sub> by the cyanobacteria is a nitrogenase-dependent reaction, while hydrogen production in green algae such as *Chlamydomonas reinhardtii* depends on hydrogenase. In nitrogenase H<sub>2</sub> is only produced under anoxic condition when nitrogen source is limited [11]. Nitrogenases use reductant and ATP to convert atmospheric N<sub>2</sub> to ammonium, providing fixed nitrogen for cell growth. The manner in which green algae produces hydrogen gas from water is called direct biophotolysis [6, 11]. The mechanism of photosynthetic hydrogen production entails photolysis of water and a light dependent transfer of electron via photo system II and photo system I to the chloroplast ferredoxin. Ferredoxin efficiently binds to the [Fe] hydrogenase and electrons are donated to the catalytic site known as 'Hydrogen Cluster' (HC) of the [Fe] hydrogenase. The H-cluster utilizes protons as the sinks for the photo synthetically generated electrons, leading to the synthesis of molecular hydrogen [6, 11]. This process of photosynthetic hydrogen production does not entail carbon dioxide fixation or energy storage into cellular metabolites.

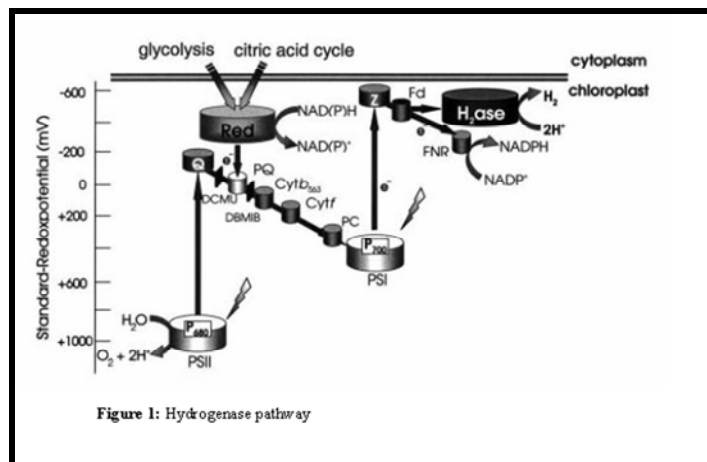


Figure 1: Hydrogenase pathway

The main aim of the present work is to predict the 3D structure of the Hydrogenase and analyze possible pathways of molecular hydrogen entering inside hydrogenase. This is performed using homology modeling and molecular dynamics simulations in explicit solvent and molecular hydrogen. We make no assumption on the initial position of H<sub>2</sub>, which is placed outside the protein. The entry of H<sub>2</sub> was analyzed with several copies of H<sub>2</sub> and without locally enhanced sampling. The present study includes the detection of regions potentially involved in the control of H<sub>2</sub> access to the active site.

### Methodology:

#### Target protein

The target hydrogenase sequence whose structure is to be predicted is the hydA1 gene found in the *Chlamydomonas reinhardtii*. Till now two [Fe]-hydrogenase had been cloned and sequenced from *Chlamydomonas reinhardtii*, hydA1 and hydA2 and had been deposited in Genbank (accession numbers AY055755 and AY055756)4,5,16. Given below is the target hydA1 sequence:

```
>gij18026270[gb|AAL23572.1| iron-hydrogenase HydA1
[Chlamydomonas reinhardtii]
MSALVLKPCAAVSIRGSSCRARQVAPRAPLAASTVVRVALATLEAP
ARRLGNVACAAAAPAAEAPLSHVQQALAEAKPKDDPTRKHVVCV
QVAPAVRVAIAETLGLAPGATTPKQLAEGLRRLGFDEVFDTLFGA
DLTIMEEGSELLHRLTEHLEAHPHSDEPLPMFTSCCPGWIAMLEKS
YDPLIPYVSSCKSPQMMLAAMVKSYLEAKKGIAPKDMVMVSI MPC
TRKQSEADRDWFCVDADPTLRQLDHVITTVELGNIFKERGINKLAEL
PEGEWDNPMGVGSGAGVLFGTGGVMEALRTAYELFTGTPLPR
LSLSEVRGMDGIKETNITMVPAPGSKFEELKHHRAAARAEAAAHG
TPGPLAWDGGAGFTSEDRGGITLRVAVANGLGNAKLLITKMQA
GEAKYDFVEIMACPVAGCVGGGGQPRSTDKAITQKROAALYNLDE
KSTLRRSHENPSIRELYDTYLGEPLGHKAHELLHTHYVAGGVEEKD
EKK
```

#### NCBI cdd search

The target sequence was deposited to the ncbi cd-search tool for finding the conserved region in the sequence.

#### NCBI blast search

Similarly psi-blast search was also performed for finding the sequences having similarity with the target sequences.

#### Structure prediction of target protein

The structure of hydA1 hydrogenase from *chlamydomonas reinhardtii* was predicted using comparative or homology modeling software modeler 8v2.

#### Amino acid sequence alignment

For single template model building procedure the target sequence hydA1 was aligned with the template sequences using the align2d command in modeller 8v2.

#### Model building

Initially the model was build using the automodel class of modeler 8v2.

#### Model validation

Models constructed in the manner described above were validated by a variety of computational methods. the simplest of these is to assess the distribution of the  $\phi$  and  $\psi$  angles on the side chains using ramachandran plots (chang and loew 1996).

#### Modeling hydA1 hydrogenase from *chlamydomonas reinhardtii* based on a single template 1feh

alignment between the target structure and template was done using the modeller 8v2 2dalign function.. once a target-template alignment is constructed, the 3d model of the target can be calculated automatically from the template 1feh using the modeller automodel class.

#### Molecular dynamics

##### Analysis methods

When carrying out a molecular dynamics (md) simulation, coordinates and velocities of the system are saved; these are then used for the analysis. Time dependent properties can be displayed graphically, where one of the axis corresponds to time and the other to the quantity of interest, such as energy, rmsd (root men square deviation), etc.

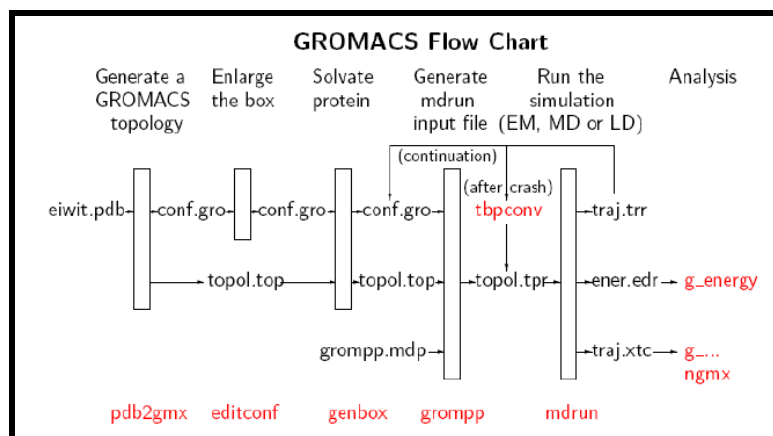
In the first one, a 20 ps simulation was run with all protein heavy atoms position-restrained with a 1 kJ/(mol nm<sup>2</sup>) force constant and with initial velocities taken from a Maxwellian distribution at 300 K and a temperature coupling constant between baths of 0.1 ps. In the second step, a 50 ps simulation was run with all atoms free and a temperature coupling constant of 0.1 ps. Conformations were saved every picosecond for later analysis.

#### Molecular dynamics simulation of h<sub>2</sub> gas diffusion pathways in the protein

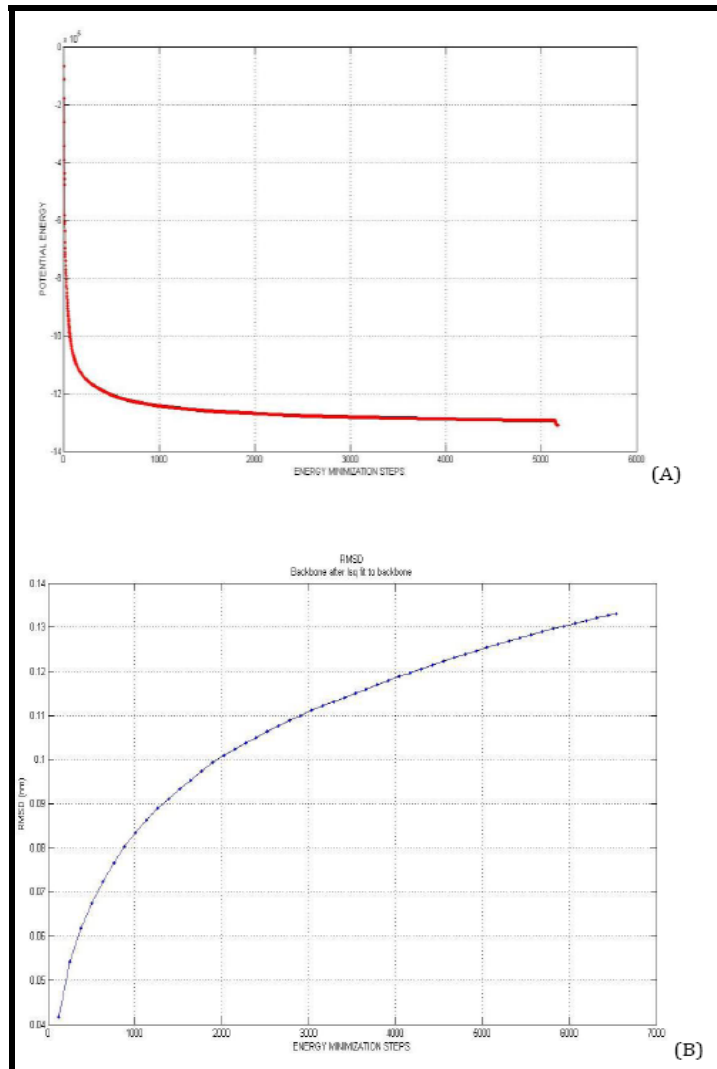
Before performing the Molecular Dynamic Simulation Energy Minimization of the whole system was performed for optimization of hydrogen atoms. Total 6000 steps of Steepest Descent with initial step size of 0.01 and Fmax less than 100 KJmol<sup>-1</sup> were performed. A 105 kJ (mol nm<sup>2</sup>) position-restraining force constant was used in the minimization step.



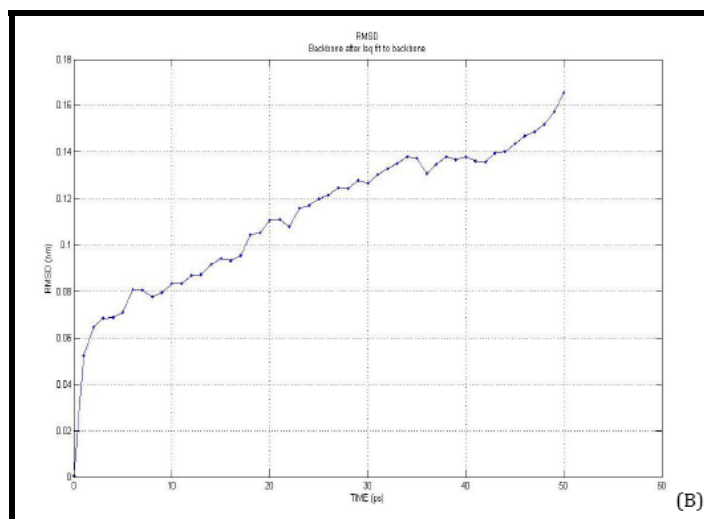
**Figure 2:** Alignment generated using Chimera. Here yellow region indicated the actual and predicted  $\alpha$ -helix in the template and the target (here query) sequence. Similarly green region indicates the  $\beta$ -sheets.



**Figure 3:** Gromacs Flow Chart



**Figure 4:** A: Potential Energy Plot, B: RMSD Plot. Molecular Dynamics simulation of of the hydA1 [Fe] hydrogenase was performed in a water box in two steps.



**Figure 5:** Showing RMSD Plot in protein structure during 50 ps simulation.

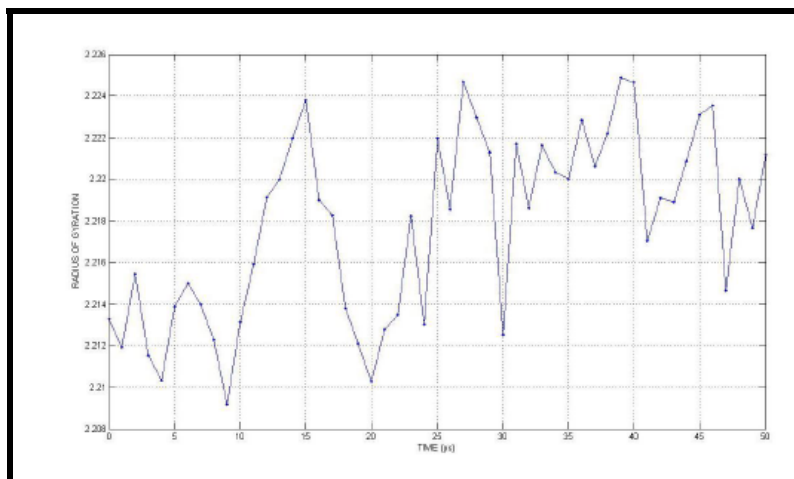


Figure 6: Showing Fluctuations in Radius of Gyration after every 1ps during 50ps simulation.

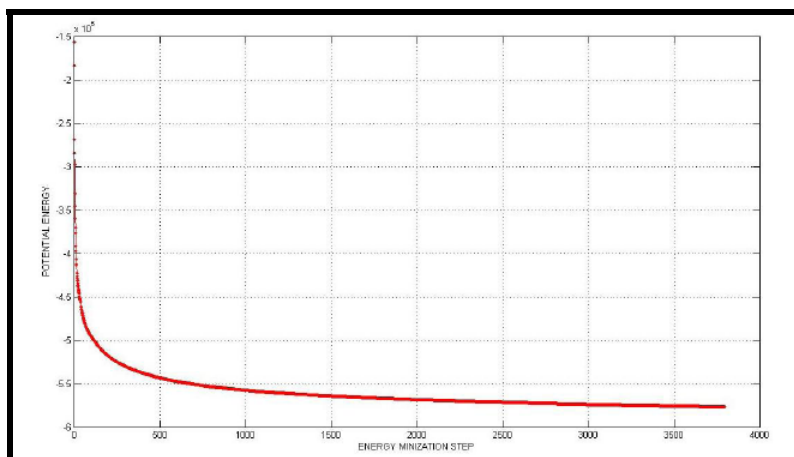


Figure 7: Showing Potential Energy Plot.

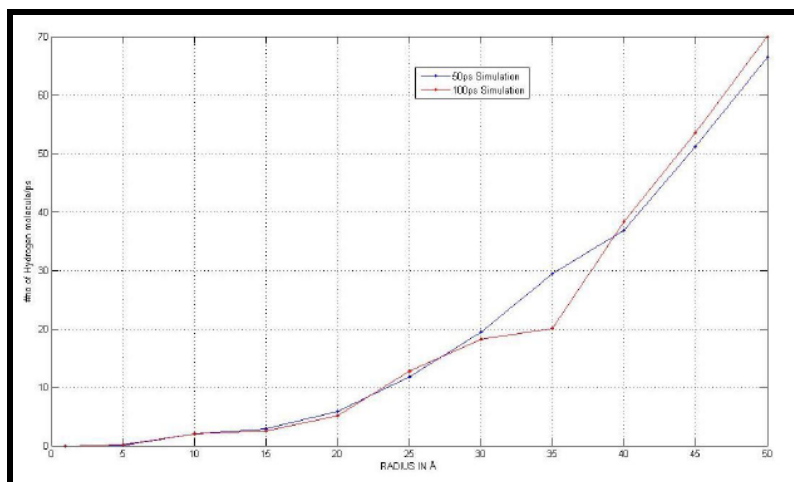


Figure 8: Plot Showing # no of H<sub>2</sub>/ps v/s cutoff radius.



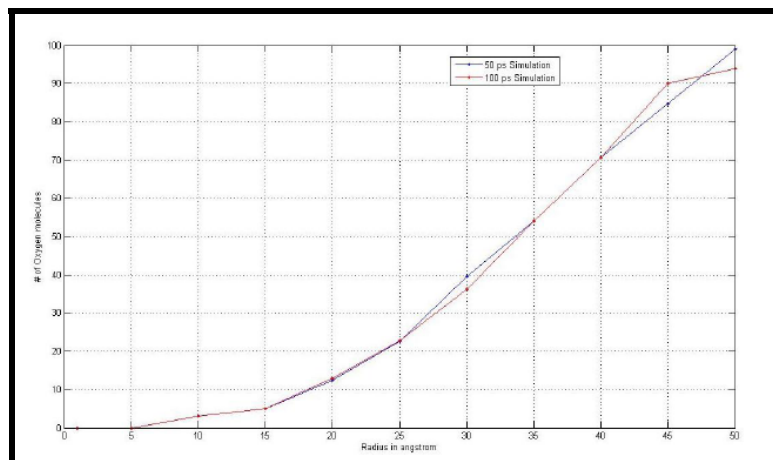


Figure 9: Plot showing # no of O<sub>2</sub>/ps v/s cutoff radius

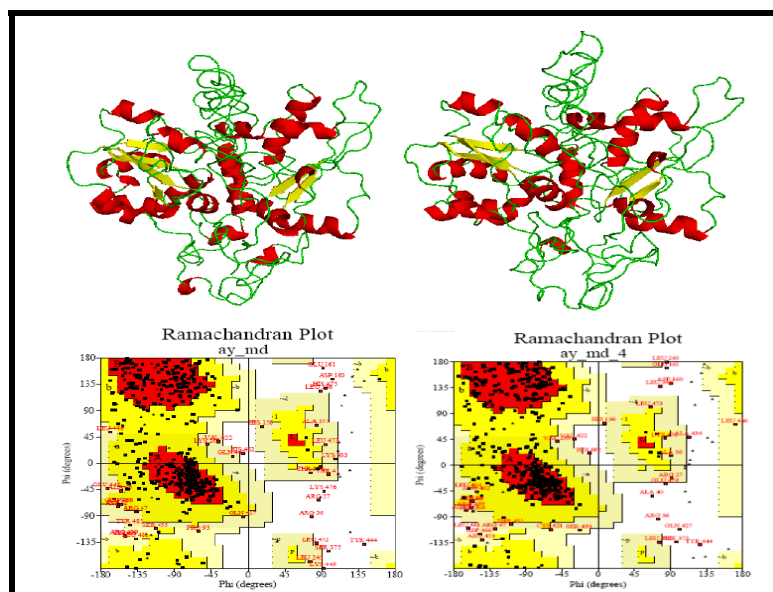


Figure 10. Ramachandran plot and predicted structure.

### Discussion:

The result of the alignment between the template and target shows that the two sequences are 32.6% identical while the similarity between the two is around 42%. Certain  $\alpha$ -helix and  $\beta$ -strands were also found to be conserved in the alignment. The  $\beta$ -strands surrounding the active site region in IFEH are found to be conserved in the alignment. A large region in the target sequence (from residue no 337 to 382) was found to be unaligned with the template. The region contains an helix as predicted by the 'mgenthrader' server. Thus loop refining has to be done for this region. The Cys300, Cys355, Cys499 and Cys503 which ligate the [4Fe-4S] cluster are found to be conserved in the target Cys170, Cys225, Cys417 and Cys421. Molecular dynamics simulation was performed for the energy minimization and structure validation. We successfully performed this.

### Conclusion

The work presented here aims to first predict the possible structure of hydA1 hydrogenase from *Chlamydomonas reinhardtii* and to analyze possible paths followed by molecular hydrogen (H<sub>2</sub>) and oxygen (O<sub>2</sub>) towards the active site and also to find protein regions potentially controlling its passage. In order to improve sampling efficiency, various MD simulation replicates were used. The first conclusion that can be taken from this work is that molecular hydrogen and oxygen easily enters the protein. In every simulation done, H<sub>2</sub> penetrates hydrogenase very early in

the simulation, but at the end it was seen that only 12% of the H<sub>2</sub> molecules are inside the protein. In the case of O<sub>2</sub> molecules only 22% of the oxygen molecules are inside the protein at the end of each simulation. This amount is very small as compared to what is required for predicting the possible pathways if both gas inside the protein. In all cases of simulation, none of the molecular hydrogen and oxygen approaches the active centre.

### Acknowledgement:

We thank to The Coordinator, Centre for Bioinformatics, University of Allahabad for providing us necessary infrastructure facilities.

### References:

- [1] M Seibert *et al.* *Trends in Biotechnology* **18** (12): 506 – 511 (2000) [PMID: 11102662]
- [2] J Cohen *et al.* *Structure* **13** (9): 1321 (2005) [PMID :16154089]
- [3] J Cohen *et al.* *Biochemical Society Transactions* **33** (1) (2005)
- [4] M Forestier *et al.* *European Journal of Biochemistry* **270** (13) : 2750 (2003) [PMID: 12823545]
- [5] M Seibert *et al.* Contractor Grantee Workshop III Washington, D.C. February 6 (2005)

- |  |  |
|--|--|
| <p>[6] A Melis &amp; Thomas Happe, 20, <i>Discoveries in Photosynthesis</i>, ISSN1572-0233 (2004)</p> <p>[7] S Zilberman <i>et al.</i> <i>Inorg. Chem.</i> <b>46</b> (4): 1153 -1161 (2007) [PMID: 17256840]</p> <p>[8] MC Posewitz <i>et al.</i> <i>Annual Review of Plant Biology</i>, <b>58</b>: 71-91 (2007) [PMID: 17150028]</p> <p>[9] ML Ghirardi <i>et al.</i> <i>Biochemical Society Transactions</i> <b>33</b> (1) (2005) [PMID : 15667268]</p> <p>[10] ML Ghirardi &amp; M Seibert, <i>Journal of Hydrogen Energy</i> <b>27</b> (11-12): 1239 (2002)</p> <p>[11] D Das <i>et al.</i> <i>Current science</i> 90 (12, 25) (2006)</p> <p>[12] T Happe &amp; A Kaminski, <i>European Journal of Biochemistry</i> <b>269</b> (3) 1022 (2002) [PMID : 11846805]</p> | <p>[13] Wade A. Amos, <i>NREL/MP</i>: 560 (2004)</p> <p>[14] T Thiel, <i>FEMS Microbiology Letters</i> <b>304</b> (1): 55 (1994)</p> <p>[15] A Melis <i>et al.</i> <i>Plant Physiol</i> 122(1):127-136 (2000) [PMID: 10631256]</p> <p>[16] M Seibert <i>et al.</i>, <i>Proceedings of the 2002 U.S. DOE Hydrogen Program Review NREL/CP</i>: 610 (2002)</p> <p>[17] B Esper <i>et al.</i> <i>Trends in Plant Science</i> (2006) <b>11</b> (11): 543 (2006) [PMID: 17029931]</p> <p>[18] <a href="http://metallo.scripps.edu/promise/FEHASE.html">http://metallo.scripps.edu/promise/FEHASE.html</a></p> <p>[19] <a href="http://nihserver.mbi.ucla.edu/SAVS/">http://nihserver.mbi.ucla.edu/SAVS/</a>: SAVS server</p> |
|--|--|

Edited by P. Kanguane

Citation: Sundaram *et al.* *Bioinformatics* 5(4): 177-183 (2010)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.