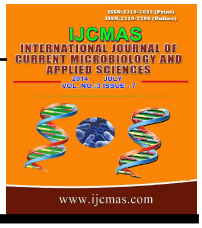
International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 3 Number 7 (2014) pp. 481-494**
http://www.ijcmas.com



Original Research Article

Microalgae as Model Organism for Biohydrogen Production: A Study on Structural Analysis

Anirudh Bhanu Teja Nelabhotla[1,2], Nivedita Sahu[1*] and Nityananda Pradhan [2]

[1]Biotechnology Lab., Medicinal Chemistry and Pharmacology,
CSIR - Indian Institute of Chemical Technology, Hyderabad-500007, India
[2]Department of Psychopharmacology, National Institute of Mental Health and Neuroscience, Bangalore-560029, India
**Corresponding author*

Keywords

Hydrogenase; Homology; Gas Diffusion; Oxygen Sensitivity; H-Cluster

A B S T R A C T

Hydrogenase is an enzyme that catalyzes the reversible oxidation of molecular hydrogen (H_2). Here, the hydrogenase enzyme structures of *Chlamydomonas reinhardtii, Clostridium pasteurianum* and *Desulfovibrio vulgaris* were taken as reference to develop homology model structures of the hydrogenases taken from *Clostridium acetobutylicum, Chlorella variabilis* and *Scenedesmus obliquus.* BLAST search was performed and the templates with a minimum of 35% sequence similarity were considered for homology modelling. Structure analysis was performed after subsequent energy minimization. *Clostridium pasteurianum* (CpI) was the reference structure for analyzing the gas diffusion path and hydrogen bonds. It has been seen that the path is formed around the H-cluster which is where the hydrogen molecule is produced. It is blocked by the oxygen molecule, in its path reducing the efficiency of the enzyme. The experiments prove the algae *Chlorella variabilis* and *Chlamydomonas reinhardtii* as the next model organisms for our further research.

Introduction

Hydrogen is a chemical element bearing symbol H, atomic number 01 and the average atomic weight 1.008 amu, making it the lightest element. Its monoatomic form (H_1) is the most copious chemical substance, aggregating about 3 quarters of the Universe's baryonic mass. Hydrogen gas (dihydrogen or molecular hydrogen) is highly flammable and will burn in air at a broad range of concentrations between 4% and 75% by volume. The combustion enthalpy for hydrogen is 286 kJ/mol (Cohen et al., 2005; Sen et al., 2008).

Biological Hydrogen Production

H_2 is a product of some types of anaerobic metabolism and is produced by several micro organisms, usually via reactions catalyzed by iron or nickel containing enzymes called hydrogenase. These enzymes catalyze the reversible redox reaction between H_2 and its component two electrons and two protons (Kim and Kim, 2011). Creation of hydrogen gas occurs in the transfer of reducing equivalents produced during pyruvate fermentation to

water. In all photosynthetic organisms water is decomposed into its molecular components H^+ ions, electrons, and oxygen, in the presence of light. Such organisms, including the alga *Chlamydomonas reinhardtii* and cyanobacteria García et al (2012), Patel et al (2012) have evolved an alternative second step, in which electrons and H^+ ions are reduced to form hydrogen gas (H_2) by specialized enzymes called hydrogenases in the chloroplast during the dark reactions (Hemschemeier et al., 2009; Levin et al., 2004).

Oxygen Sensitivity

Hydrogen production in *Clostridium pasteurianum* (CpI) happens at the H cluster, a metallic cluster embedded inside its protein matrix **(**Adams, 1990), and is achieved by the reduction of H^+ ions from the external solution through the use of electrons acquired from a reduced carrier such as ferredoxin (Peters, 1999) . It is assumed that the H^+ ions probably reach the H cluster, but this proton path contained in the protein is yet to be verified (Peters et al., 1998). The electrons are transferred to the bounded H cluster through a series of accessory Fe-S clusters aligned in a chain between the H cluster and one end of the hydrogenase enzyme of Cpl.

CpI should allow the product to leave the protein, but it also lets the small gas molecules such as O_2 and CO to escape through the enzyme and bind the H cluster, inactivating the hydrogen production at the cluster. While the O_2-mediated deactivation of hydrogenase is in some cases advantageous for the host organism, it severely affects the efficiency of using hydrogenase to produce H_2 as a carrier of consumable energy (Seibert et al., 2002). Genetic modifications of cyanobacterial hydrogenases have been and are being tried to efficiently synthesize hydrogen gas with simultaneous release of oxygen molecule (Srirangan et al., 2011).

The present paper provides first-hand information on the gas diffusion paths as well as structural analysis of hydrogenase enzyme comparing algae and bacteria, showing microalgae as a better system in whole for commercial production of biohydrogen.

Materials and Methods

Sequence Collection and Alignment

The FASTA sequences of CpI (*Clostridium pasteurianum*) (3C8Y), *Chlamydomonas reinhardtii* (3LX4) and *Desulfovibrio vulgaris* (1HFE) were obtained through the PDB. These FASTA sequences were used to run BLAST and three sequences which mutually had atleast 35% similarity in their protein sequence were selected (Altschul et al.,1997; Alejandro et al., 2001). They belonged to *Chlorella variabilis*, *Clostridium acetobutylicum* and *Scenedesmus obliquus* and the FASTA sequences for the above were obtained.

A CLUSTAL Omega program was run on each target to all the three templates and the sequence alignment was obtained and the conserved domains were obtained.

Structure and Path Analysis

Based on the literature review, BLAST search and alignments homology models of each of the targets were developed based on the three templates. They were re-aligned in the Swiss PDB Viewer, maintaining the conserved domains undisturbed and removing the non-aligned, non-conserved amino acid sequences. The gas diffusion paths were predicted for the modelled nine structures, *C. reinhardtii* and *D. vulgaris* based on the experiments performed by Cohen and his co-author in 2005 on CpI.

The backbone structure comparison was done on the structures with each other using PyMol and a matrix of RMSD values was tabulated. Analysis was done based on the matrix and the best suitable homology model was achieved.

Result and Discussion

Sequence collection

FASTA sequences of *Chlamydomonas reinhardtii*, *Clostridium pasteurianum* and *D esulfovibrio vulgaris* were taken. The BLAST run aligned the B-chain *of C. reinhardtii* and M chain of *D. vulgaris* with the single A-chain of *C. pasteurianum*. The BLAST results gave the following organisms with at least 35% similarity in the amino acid sequence. We found highest similarity was between *Clostridium acetobutylicum* and *C. pasteurianum* as they both were bacteria of same species (71%). There was a decent amount of similarity of the algae *Chlamydomonas reinhardtii* with all the three species *Clostridium acetobutylicum*, *Chlorella variabilis* and *Scenedesmus obliquus* (39%, 48% and 62% respectively). It's also known to produce high quantities of hydrogen gas when grown in Sulphur deprived conditions.

Evolutionary Tree

A cladogram was (Figure 2) constructed using CLUSTAL Omega (1.2.1) program in the EMBL-European Bioinformatics Institute web server. *Chlorella variabilis* (G3ESX5_CHLVAA) and *Scenedesmus obliquus* (Q9AR66_SCEOB) are both microalgae placed along with *Chlamydomonas reinhardtii* (3LX4_B) and *Clostridium acetobutylicum* (Q59262_ CLOAT) is a bacterium placed along with *Clostridium pasteurianum* (3C8Y_A) and *Desulfovibrio vulgaris* (1HFE_M).

Multiple Sequence Alignment

The Figure 3 and Figure 4 show the CLUSTAL Omega multiple sequence alignment along with similarity index. With the alignment results, conserved domains were marked, and the non-aligned, non-conserved amino acids were removed before structure modelling to increase the efficiency and minimize the energies of the 3D structures generated. Later, their energies were minimized and structure analysis was performed. CLUSTAL Omega 1.2.1 is the most efficient program for sequence alignment. Here (Figure 3 and Figure 4) we found that 3LX4-B and 1HFE-M start their sequence late. It was also observed that a large gap occur in other sequences which do not match to 3LX4-B. There were around 5 to 6 core domains where we found continuous "*" (100% identity in all the six sequences) and ":" (similar amino acids in all the sequences but not identical) and "." (Similar amino acids in at least three of the six sequences) denoting conserved domains in the sequencing and play important role the protein function and folding.

Modelling

Homology Modelling

Based on the alignment generated by the CLUSTAL Omega program, modelling was made on the following structures using SWISS server and analyzed using SWISS PDB Viewer. Energy of each structure was minimized to the range 1 of -20000 KJ/mol to -30000 KJ/mol and then rendered in solid 3D of the ribbon format. After modelling the structure, the Q-Mean value was obtained from SWISS Model structure assessment tool. Energy minimization increased the Q-mean Z-score value of every modelled structure by at least 1 unit (Arnold et al., 2006; Benkert et al., 2009).

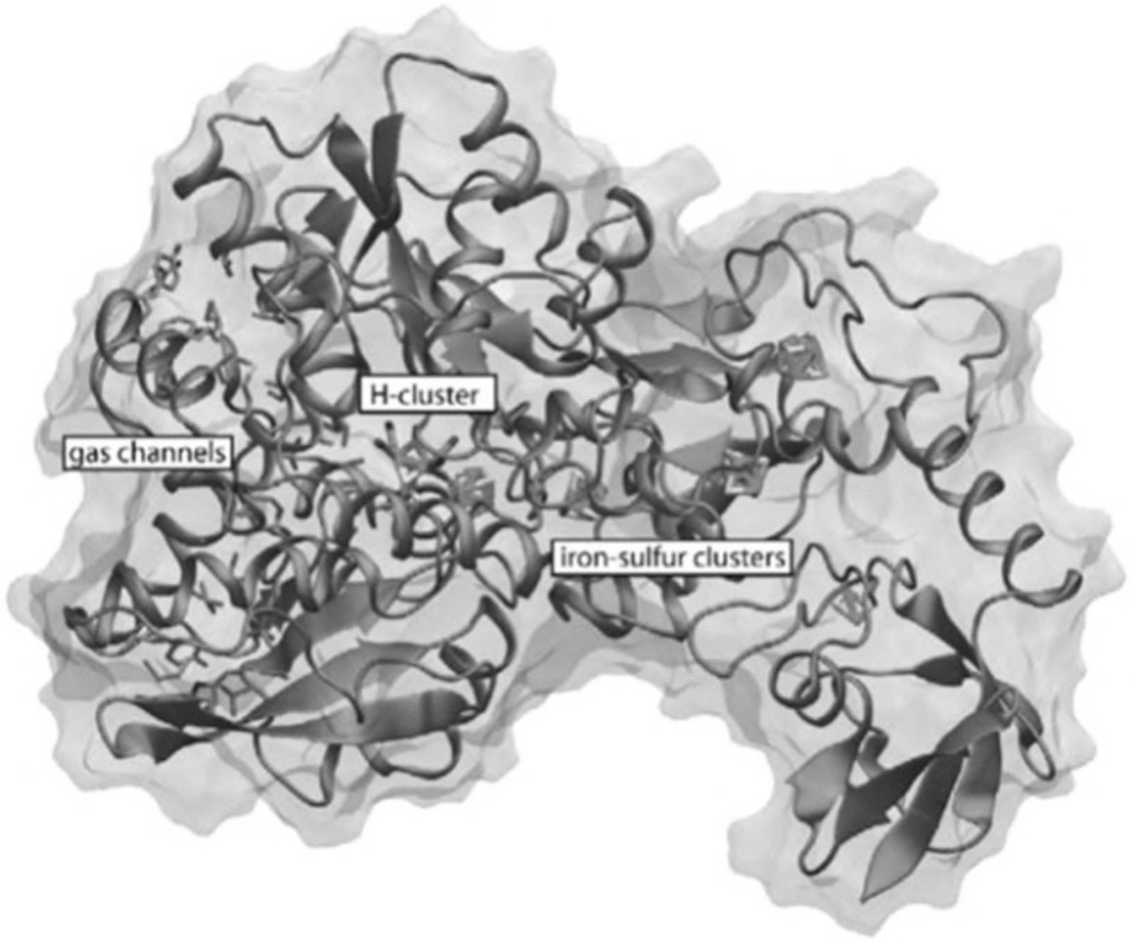


Fig.1 Structure of CpI Hydrogenase showing the enzyme with its embedded H cluster and iron sulphur clusters. Also shown are the residues lining the two principal gas Paths connecting the external solution to the buried H cluster. **(Source: Cohen et al. 2005)**

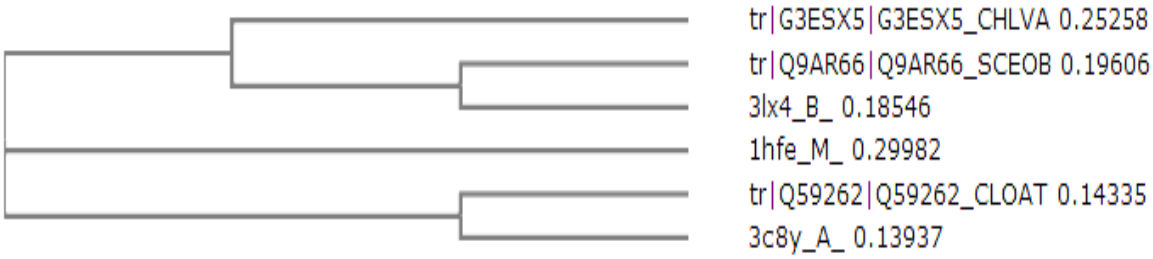


Fig.2 The image illustrates the phylogeny of the organisms

Fig.3 Part 1 of the screen shot of the CLUSTAL Omega program run for multiple sequence alignment of the target sequences of CV, SO and CA against the template sequences of 3C8Y-A, 3LX4-B and 1HFE-M.

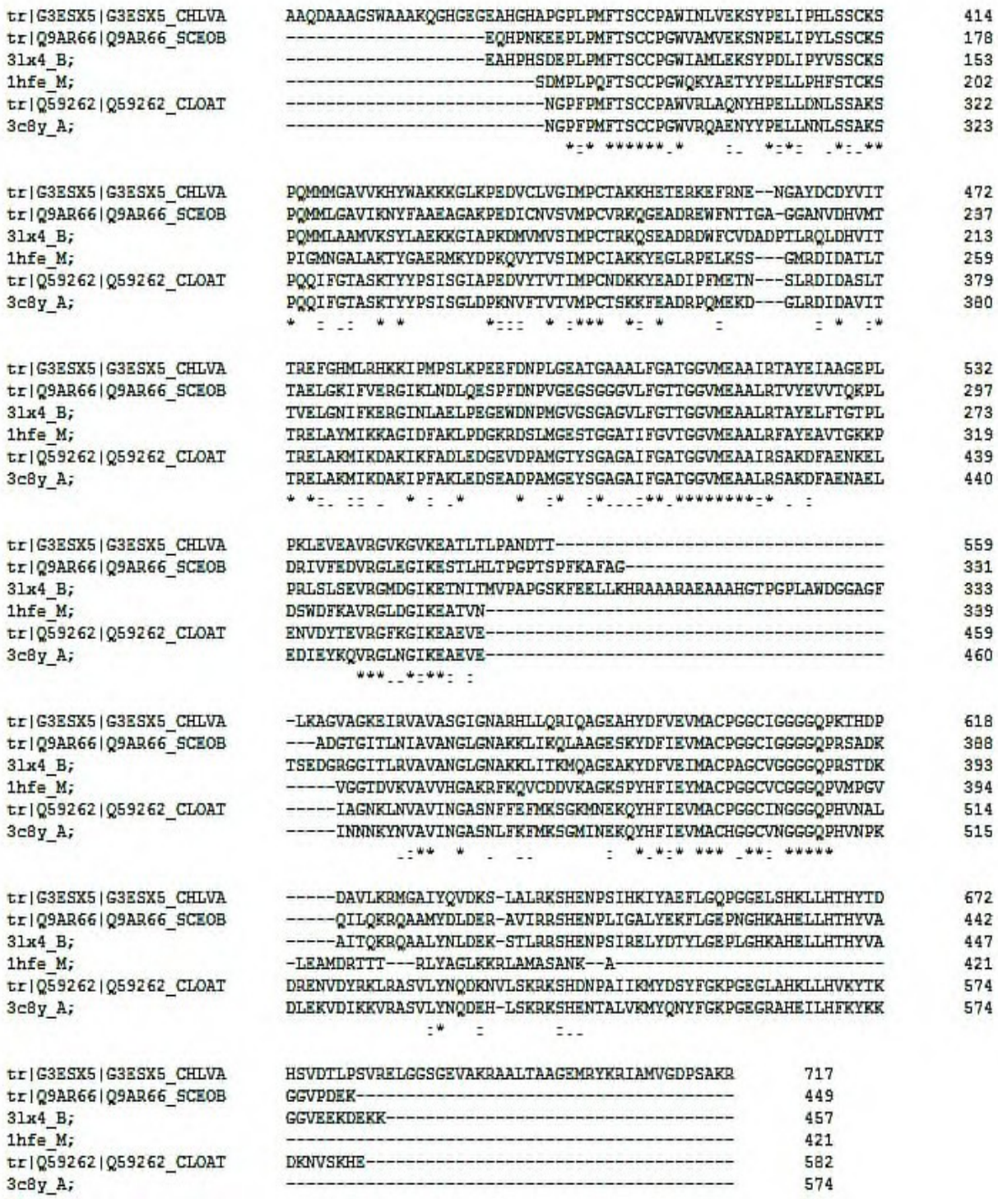


Fig.4 Part 2 of the screen shot of the CLUSTAL Omega program run for multiple sequence alignment of the target sequences of CV, SO and CA against the template sequences of 3C8Y-A, 3LX4-B and 1HFE-M.

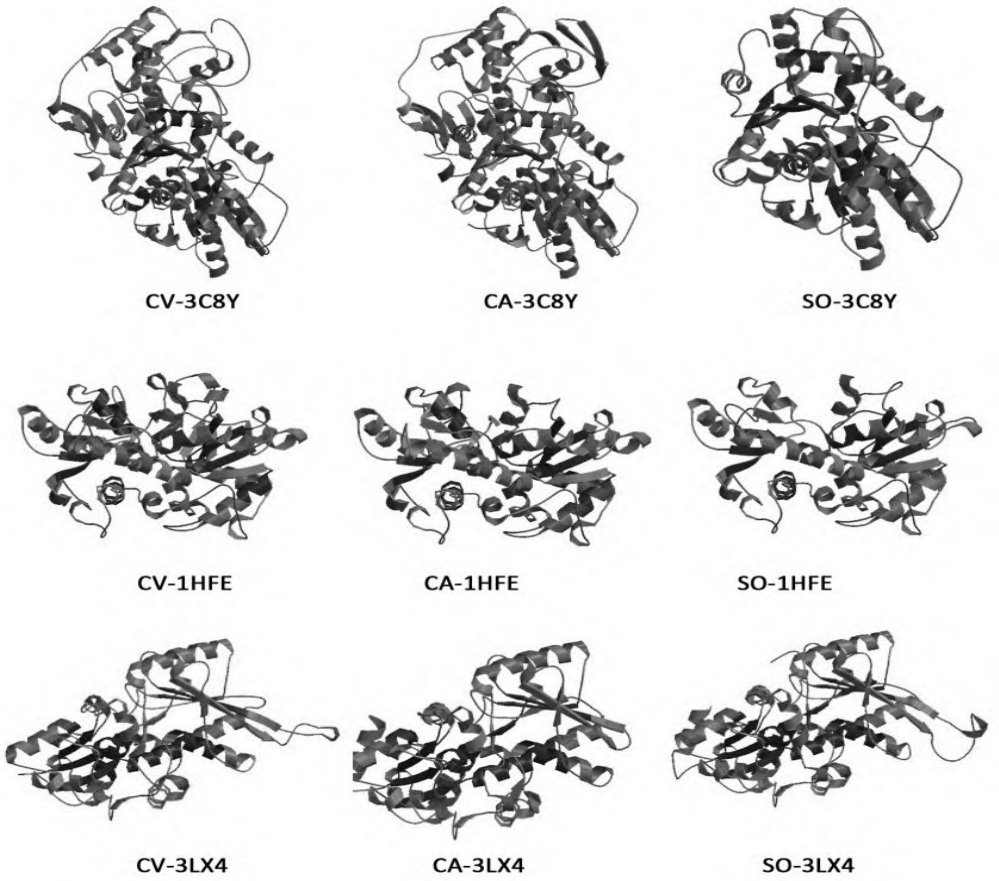


Fig.5 Homology modelled structure of all the three species based on the three available PDB structures

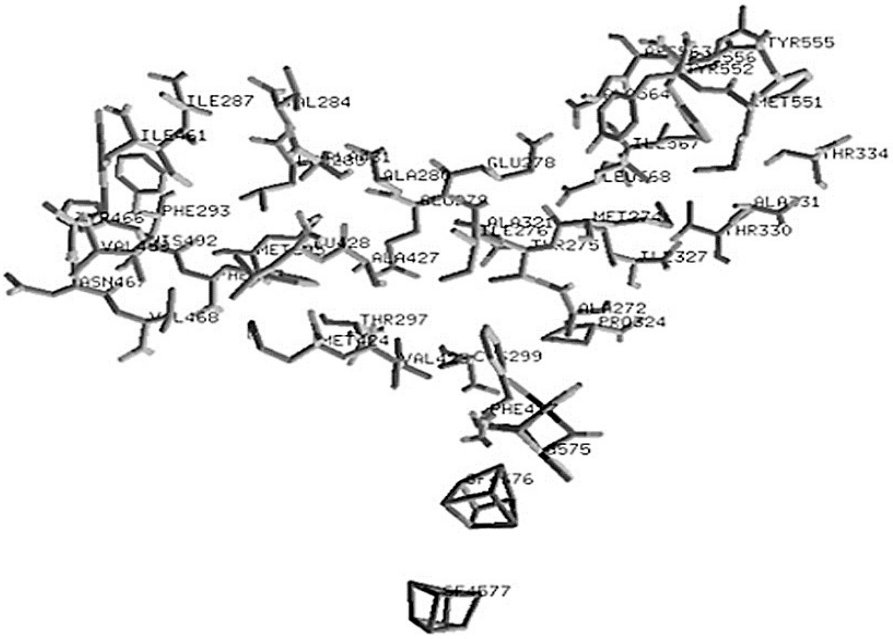


Fig.6 Molecular diffusion Path model of CpI.

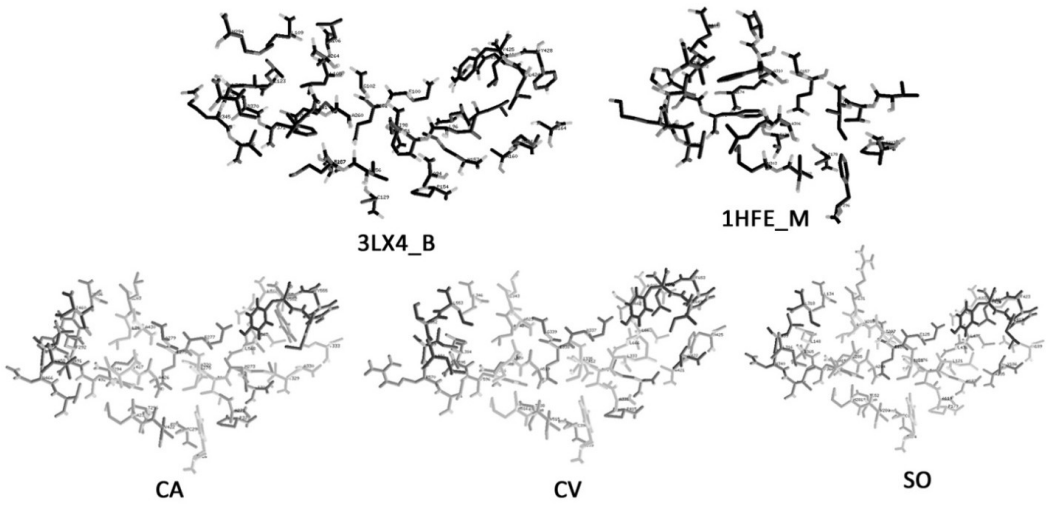


Fig.7 Molecular structure of paths of the three best models and of 3LX4 and 1HFE

Table.1 *The BLAST results of target sequences of Chlorella variabilis (CV), Scenedesmus obliquus (SO) and Clostridium acetobutylicum (CA) against selected templates. (Zdobnov et al., 2001); (Arnold et al., 2006); (Schaffer et al., 2001); (Altschul et al., 1997).*

Organism/ PDB ID	3C8Y		3LX4		1HFE	
BLAST Result	Identity	Score	Identity	Score	Identity	Score
CV	37%	411	48%	377	35%	250
CA	71%	860	39%	294	42%	278
SO	39%	294	62%	507	39%	215

Table.2 The table describes the Q Mean values and Z score Q Mean values of the energy minimized modeled structures. (Zhou, et al., 2002)

Protein	CA-3C8Y	CA-3LX4	CA-1HFE	CV-3C8Y	CV-3LX4	CV-1HFE	SO-3C8Y	SO-3LX4	SO-1HFE
Q Mean Score	0.68	0.60	0.66	0.60	0.71	0.63	0.65	0.73	0.64
Z-Score: Q Mean	-0.97	-2.02	-1.36	-1.88	-0.73	-1.70	-1.39	-0.50	-1.52

Table.3 The table describes the distribution of amino acids in all the structures template and target, representing the cavity, where H-cluster is present and the first place where hydrogen is produced. Here CA, CV, SO represent modelled structures based on each template.

Template				Target		
S.No.	3C8Y	1HFE	3LX4	CA	CV	SO
1	272-A	149-A	94-A	271-A	331-A	119-A
2	275-T	152-T	97-T	274-T	334-T	122-T
3	276-I	153-I	98-I	275-I	335-I	123-I
4	279-E	156-E	101-E	278-E	338-E	126-E
5	297-T	176-T	127-T	296-T	388-T	152-T
6	299-C	178-C	129-C	298-C	390-C	154-C
7	324-P	203-P	154-P	323-P	415-P	179-P
8	417-F	296-F	250-F	416-F	509-F	274-F
9	423-V	302-V	256-V	422-V	515-V	280-V

Table.4 The table describes the amino acids in all the structures template and target, representing the Path A. The Path is made of same amino acids in the respective target structures except CV-1HFE. (The 1HFE-M modelled CV has 564-V (based on the 340-V) and not 553-L as a part of the Path.)

Template				Target		
S.No	3C8Y	1HFE	3LX4	CA	CV	SO
1	274-M	151-V	96-L	273-M	333-L	121-L
2	280-A	157-G	102-G	279-A	339-G	127-G
3	283-L	160-F	105-L	282-L	342-L	130-L
4	284-V	161-V	106-L	283-L	343-L	131-R
5	287-I	164-L	109-L	286-V	346-L	134-L
6	293-F	172-L	123-L	292-F	384-L	148-L
7	295-M	174-Q	125-M	294-M	386-M	150-M
8	424-M	303-M	257-M	423-M	516-M	281-M
9	427-A	306-A	260-A	426-A	519-A	284-A
10	428-L	307-L	261-L	427-I	520-I	285-L
11	431-A	310-A	264-A	430-A	523-A	288-V
12	459-V	338-V	292-I	458-V	551-L	316-L
13	461-I	340-V	294-M	460-I	553-L[#]	318-L
14	466-Y	345-V	344-L	465-L	569-I	339-L
15	467-N	346-K	345-R	466-N	570-R	340-N
16	468-V	347-V	346-V	467-V	571-V	341-I
17	492-H	371-H	370-D	491-H	595-D	365-D
18	493-F	372-F	371-F	492-F	596-F	366-F

Table.5 This table describes the amino acids in all the structures template and target, representing the Path B. The Path is made-up of same amino acid in the respective target structures. Here CA, CV, SO represent modelled structures based on only 3C8Y and 3LX4 as templates. Path B does not exist in 1HFE.

Template			Target		
S.No	3C8Y	3LX4	CA	CV	SO
1	275-T	97-T	274-T	334-T	122-T
2	278-E	100-E	277-E	337-E	125-E
3	279-E	101-E	278-E	338-E	126-E
4	321-A	151-C	320-A	412-C	176-C
5	327-I	157-M	326-I	418-M	182-M
6	330-T	160-A	329-T	421-A	185-A
7	331-A	161-M	330-A	422-V	186-V
8	334-T	164-S	333-T	425-H	189-N
9	551-M	424-L	551-M	649-I	419-L
10	552-Y	425-Y	552-Y	650-Y	420-Y
11	555-Y	428-Y	555-Y	653-F	423-F
13	556-F	429-L	556-F	654-L	424-L
14	563-R	436-K	563-L	661-L	431-K
15	564-A	437-A	564-A	662-S	432-A
16	567-I	440-L	567-L	665-L	435-L
17	568-L	441-L	568-L	666-L	436-L

Table.6 The table shows the number of hydrogen bonds formed by the amino acids comprising the cluster and Paths along with their target sequences (3C8Y&3LX4/1HFE).

Protein	3C8Y	3LX4_B	1HFE_M	CV	CA	SO
No. of H-bond	32	31	19	27/18	23/14	23/15

Table.7 This table describes the RMSD value matrix of all the modelled structures and the template structures aligned against each other.

RMSD	cv_3c8y	ca_3c8y	so_3c8y	cv_3lx4	ca_3lx4	so_3lx4	cv_1hfe	ca_1hfe	so_1hfe
3c8y	**0.08**	**0.11**	**0.09**	1.17	1.13	1.17	0.81	0.74	0.68
3lx4_b	1.15	1.18	1.15	**0.08**	**0.10**	**0.08**	1.21	1.17	1.23
1hfe_m	0.74	0.70	0.67	1.15	1.16	1.17	**0.09**	**0.09**	**0.10**
cv_3c8y	#	**0.07**	**0.07**	1.17	1.19	1.17	0.79	0.75	0.69
ca_3c8y		#	**0.07**	1.19	1.20	1.16	0.76	0.74	0.69
so_3c8y			#	1.15	1.19	1.16	0.72	0.72	0.70
cv_3lx4				#	**0.09**	**0.05**	1.22	1.21	1.16
ca_3lx4					#	**0.08**	1.23	1.20	1.23
so_3lx4						#	1.19	1.14	1.29
cv_1hfe							#	**0.07**	**0.06**
ca_1hfe								#	**0.08**

The Table 1 describes the BLAST result to template organism tabulated only with minimum sequence similarity of 35% (30% being cutoff) (Seibert et al., 2002; Zdobnov and Apweiler, 2000). If the sequence similarity falls below 30% then the organisms are assumed to be totally unrelated and ambiguous identity (Rost, 1999). The Q Mean is the global score of the whole model reflecting the predicted model reliability ranging from 0 to 1 (Zhou and Zhou, 2002). The QMEAN Z-score provides an estimate of the absolute quality of a model by relating it to reference structures solved by X-ray crystallography. The QMEAN Z-score is an estimate of the "degree of nativeness" of the structural features observed in a model by describing the likelihood that a model is of comparable quality to high-resolution experimental structures. In the current study, 3 models were selected based on the Q Mean values and Z score Q Mean values of the energy minimized modelled structures on CA, CV and SO as described in Table 2. Figure 5 shows the ribbon models of the enzyme modelled from different species give us physical evidences of different possible structures. For example SO-3C8Y here has an α-helix and a β-sheet at the top left corner which are not observed in other structures based on 3C8Y. Similarly, we found a small α - helix in CV-1HFE which is not observed in others. The major difference we observed was the different shapes of the same enzyme of the organism modelled under different PDB structures.

Gas Diffusion Path Analysis

The current study presented the modelled structures showing only the predicted gas diffusion paths of the above structures. Figure 6 shows the defined diffusion paths of CpI, this is the molecular analogue of Figure 1. There are two paths of diffusion for O_2 and H_2 transport with center as H-cluster in the CpI. Cavity is defined by CpI amino acid residues “272, 275, 276, 279, 297, 299, 324, 417, and 423”; the central cavity extends into, Path A and Path B. Path A is made of same amino acids in the respective target structures except CV-1HFE. It is defined by amino acid residues “274, 280, 283, 284, 287, 293, 295, 424,

427, 428, 431, 459, 461, 466, 467, 468, 492, and 493" (Table 4). The Path B, defined by amino acid residues "275, 278, 279, 321, 327, 35 330, 331, 334, 551, 552, 555, 556, 563, 564, 567, 568" (Table 5). The table describes the amino acids only in the 3C8Y and 3LX4 as the structure templates. In 1HFE there is absence of Path B.

This study has the same opinion with the analysis made by Cohen and his co-author in 2005. Figure 7 shows the path models predicted based on CpI (Figure 6). 1HFE-M (*D. vulgaris*) is the shortest and when its path was being predicted, only one of the two paths were identified i.e. Path A. Also the paths of structures developed based 1 on 1HFE-M were also predicted similarly (not shown).

Distribution of amino acids

The amino acids present in all the structures template and target thus representing the cavity, where H-cluster is present, which is the first place where hydrogen is produced. This has been well illustrated in Table 3. The cavity is made-up of same amino acids in each of the nine target structures. Here CA, CV, SO represent modelled structures based on each template. The present study well relates to the findings done by Cohen et al (2005).

Table 4 and Table 5 show all the amino acids that contribute to the gas diffusion paths, i.e. path A and path B respectively based on CpI (Cohen et al., 2005). As mentioned earlier, these are only predictions of the mechanism or the route which the gas follows to escape from within the enzyme. Molecular interactions among these amino acids have allowed them to form a system which lets the gases out.

Hydrogen Bonds

Hydrogen bonds which play the most significant role in the interactions in these kinds of mechanisms have been studied and described in Table 6. The number of hydrogen bonds in the modelled structures are only a little less than the X-ray crystallographic structures which suggest that the modelling done by the server was not accurate but efficient enough to guide us. Few of the amino acids form multiple hydrogen bonds either with same or different amino acids, which makes them significant in the path. Also the amino acids which form hydrogen bonds across the radius of the path can be studied further as they can be important for eliminating the oxygen sensitivity and for efficient production of hydrogen. *Chlorella variabilis* forms the highest number of hydrogen bonds (Table 6).

Structure Analysis

Each of the nine structures are aligned to remaining and a matrix of RMSD values is generated (Table 7). All the values are below 3.00 units signifying an efficient modelling mechanism. Some of the values were not obtained; this may be due to the vast size difference and backbone mismatch. The previous works done by Benkert et al (2009) on PyMol prove that the values lesser than 3.00 units depict good results and all the structures modelled are competent and also, lesser the value more efficient the model. In the table, values which have been made bold signify values ≤0.11 which are more consistent and the next nearest value to 0.11 is 0.67 (homology models) which is relatively higher. Most of the models which fall under this category belong to *Clostridium acetobutylicum* modelled based on all the three structures (0.11,

0.07, 0.07, 0.1, 0.09, 0.08, 0.09, 0.07 and 0.08) averaging 0.084. Those are 9 values of 18 boldly marked values. The *Scenedesmus obliquus* has only 8 values under this category. However, *Chlorella variabilis* also gathered 9 values, but more importantly all of them are ≤ 0.09 only and (0.08, 0.07, 0.07, 0.08, 0.09, 0.05, 0.09, 0.07 and 0.06) averaging only 0.073 units. It can be said that *Chlorella* 1 *variabilis*, which maintains an efficient RMSD value and consistent homology with all the three structures with an optimum value of around 40% (BLAST result), is a potential model organism next to CpI. Apart from this, an analysis has been made on the α-helix and β-sheet positions for a deeper view of the conserved domains and their formation. The amino acids which form the H cluster and paths are integral parts of 6 different α-helices and 3 different β-sheets, however, Path A and the H-cluster are formed by only 4 α-helices and 3 β-sheets.

Having found the positions of each amino acid with respect to the helices and sheets, the next step would be to find the radius of the Paths and where exactly hydrogen molecules are obstructed by the oxygen molecules. Further, the tertiary structure analysis and hydrophobic interactions will be studied. With these results we will be going for the structural analysis of the protein under the influence of various ligands of each modelled structure and obtain a reliable solution which can be implemented practically. The results obtained from this work open a lot of opportunities in exploring more organisms which can be potential sources for hydrogen production and make a revolution in fuel resources. The oxygen sensitivity as proposed by many researches is the only posing problem for the efficient hydrogen production. Hydrogen gas storage and hydrogen run motors must be the target of every researcher in this field. The government and the scientific community must encourage environmental friendly and energy efficient projects.

Acknowledgements

The authors express sense of gratitude and thanks to Dr. Shibasish Choudhury, Assistant Professor, Department of Biology, BITS-Pilani and to Mr. Shankarachariyar, system 26 administrator at NIMHANS for providing required facilities and support during the course of the project work. The authors duly acknowledge Dr. Ahmed Kamal, Dr. USN Murty, Dr. Sunil Misra and Mr. Amit Banerjee at CSIR-IICT, Hyderabad, India for their support to carry out the study.

References

Adams, M.W.M. 1990. The structure and function of iron-hydrogenase. Biochim. Biophys. Acta 1020, 115-145.

Altschul SF., Madden TL., Schaffer AA., Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Alejandro AS, Aravind L, Thomas LM, Sergei S, John LS, Yuri IW, Eugene VK, Stephen F *A.*2001., "Improving the accuracy of PSI-BLAST protein database searches with composition based statistics and other refinements", Nucleic Acids Res. 29: 2994-3005.

Arnold K., Bordoli L., Kopp J., and Schwede T. 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22,195-201.

Benkert, P., Schwede, T. and Tosatto, S.C.E. 2009. QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. BMC Struct Biol. 2009 May 20; 9:35.

Cohen, J., Kim, K., Posewitz, M., Ghirardi, M. L., Schulten, K., Seibert, M., and King, P. 2005. Molecular dynamics and experimental investigation of H 2 and O 2 diffusion in [Fe] -hydrogenase, 33, 80–82.

Das, D. and Veziroglu, T. 2008. Advances in biological hydrogen production processes. *International Journal of Hydrogen Energy*, 33(21), pp.6046–6057.

García, R. E., Martínez, V. L., Franco, J. I., and Curutchet, G. 2012. Selection of natural bacterial communities for the biological production of hydrogen. *International Journal of Hydrogen Energy*, 37(13), pp.10095–10100.

Hemschemeier, A., Melis, A. and Happe, T. 2009. Analytical approaches to hotobiological hydrogen production in unicellular green algae. *Photosynthesis research*, 102(2-3), pp.523–40.

Kim, D.H. and Kim, M.S. 2011. Hydrogenases for biological hydrogen production. *Bioresource technology*, 102(18), pp.8423–31.

Levin, D., Pitt, L. and Love, M. 2004. Biohydrogen production: prospects and limitations to practical application. *International Journal of Hydrogen Energy*, 29(2), pp.173–185.

Patel, S.K.S., Kumar, P. and Kalia, V.C. 2012. Enhancing biological hydrogen production through complementary microbial metabolisms. *International Journal of Hydrogen Energy*, (2), pp.1–14.

Peters, J. W., Lanzilotta, W.N., Lemon, B.J., Seefeldt, L.C. 1 1998. X-ray crystal structure of the Fe-only hydrogenase (CpI) from Clostridium pasteurianum to 1.8 angstrom resolution. Science 282: 1853–1858

Peters JW. 1999. Structure and mechanism of iron-only hydrogenases. Current Opinion Structural Biology 9: 670–676.

Rost, B. 1999. Twilight zone of protein sequence alignments. *Protein Engineering Design and Selection*, 12(2), pp.85–94.

Seibert, M., King, P., Zhang, L., Mets, L., and Ghirardi, M. 2002. Molecular Engineering of Algal H 2 Production. Proceedings of the 2002 U.S. DOE Hydrogen Program Review, 80401(1993), 1–10.

Sen, U., Shakdwipee, M. and Banerjee, R. 2008. Status of Biological hydrogen production. *Journal of Science and Industrial Research*, 67(November), pp.980–993.

Srirangan, K., Pyne, M.E. and Perry Chou, C. 2011. Biochemical and genetic engineering strategies to enhance hydrogen production in photosynthetic algae and cyanobacteria. *Bioresource technology*, 102(18), pp.8589–604.

Zdobnov E.M. and Apweiler R. 2000. Inter ProScan an integration platform for the signature-recognition methods in Inter Pro Bioinformatics 17: 847-848.

Zhou, H. and Zhou, Y. 2002. Distance-scaled, finite ideal-gas reference state improves structure- derived potentials of mean force for structure selection and stability prediction. Protein Sci. 11: 2714-2726.