ISSN No. 0976-5697

Volume 8, No. 1, Jan-Feb 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

# Oxy-haemoglobin protein engineering: An automated design for hotspots stability, site-specific mutations and smart libraries by using HotSpot Wizard 2.0 software

Soumendra Nath Talapatra
Career Advancement Solutions
Maheshtala, Kolkata - 700142, India

Partha Talukdar*
Department of Botany, Serampore College,
University of Calcutta, William Carey Road,
Hooghly, West Bengal, India

***Abstract:*** Oxy-haemoglobin is an important metallo-protein, which helps in oxygen binding and transporting to the tissues. The objective of the present study was to detect hot spots and design of smart libraries for engineering protein stability, substrate specificity, tunnels and cavities as well as suitable mutability position of oxy-haemoglobin protein by using a software, HotSpot Wizard, version 2.0., is a free online software. The prediction results were obtained in output interface for functional hot spots, stability hot spots (structural flexibility), correlated hot spots and stability hot spots (sequence consensus). In conclusion, pocket identification and mutability prediction of oxy-haemoglobin can lead to detect structural alternation mainly in disease diagnosis and space for ligand binding pocket in new drug development for disease therapy. This computational prediction is suggesting to compare with experimental hotspots for oxy-haemoglobin in relation to therapeutic efficacies and druggability assessment.

***Keywords:*** oxyhaemoglobin; metallo-protein; protein engineering; HotSpot Wizard; computational biology; druggability

## I. INTRODUCTION

The oxy-haemoglobin protein is an important metallo-protein, which play a vital role in transport of oxygen [1-3]. It is well known that each subunit of haemoglobin (Hb) contains globular protein along with heme group. The protein tetramer comprises of two α- and two β-chains assembled to form symmetrical (αβ) dimers. In the centre of each heme group is a $Fe^{2+}$. De and Girigoswami [4] and Furuyama et al. [5] have described that the haemoglobin exists in functionally important two isomeric forms the R form (oxy/ligand bound state), which helps in proper coordination between upload of oxygen. Lack of this protein lead to anaemia in human.

Generally, protein-protein interaction, which indicated residues of $\Delta\Delta G \geq$ 2kcal/mol, is termed as hot spot [6]. In other words, certain residues in protein-protein interactions, termed as hot spots. These residues have unique and variety of energetic properties, can be designed an important target of protein-protein complex [7]. Several experiments resulted that only a small subset of contact residues showed significance binding free energy. These residues have been termed 'hot spots' and if mutated then they can disrupt the interaction [8]. Most conserved amino acids are found in hot spot residues. As per experimental study, Hb variants determined through physical examination and/or routine laboratory testing, which found in the patients of diabetes, anaemia, cyanosis, etc. [6; 9-10]. Globin gene mutation causes the structural globin proteins or Hb variants, and these are associated with deletions, multiple amino acid substitutions, anti-termination mutations, and altered post-translational processing. It is well known naturally occurring Hb mutations cause biochemical abnormalities, some of which employ clinically significant symptoms [6].

Bloom et al. [11] have emphasized that stability of proteins is of great concern, those who are working on protein engineering research for the implementation of enzymes in industrial sector. Now-a-days the study of protein stability has led to develop in future utilization of biomolecules in different sectors viz. biocatalyst, disease diagnosis and therapy, nanoscience etc. [12]. In general, stability means protein gets unfold and refold during unfavourable environmental conditions as temperature or solvent, etc.

It is interesting to note that all proteins are simplest form and suitable example of evolvable biological systems as per their potent biochemical functions in which alterations can be noted due to few mutations [13]. Wagner [14] has revealed that evolvability is robustness to mutations, and proteins are often quite mutationally robust. It was found in experimental study that several proteins are retaining their native functions due to more than half number of single mutant [11; 15-17].

Since decades, the function and properties determination have developed through automated simulation by several researchers to know molecular mechanisms of any protein but still unclear the sequences of protein encode the exact function [18-19]. Generally, enzyme is known as biocatalyst, which has specific substrate binding ability as lock and key strategy for maintaining biochemical reactions in an organism. In recent trend of research, several computational tools for protein engineering have been developed by researchers mainly detection for tunnel and cavity, smart libraries, mutation positions, functions etc. [12; 20-27].

In the present study an attempt was done for oxy-haemoglobin protein to detect of hot spots and design of smart libraries for engineered protein stability, substrate specificity, tunnels and cavities as well as suitable mutability position through computational prediction by using HotSpot Wizards, version 2.0 and the protein was used oxyhaemoglobin because this is an important protein for blood related disease identification and also prevention.

## II. MATERIALS AND METHODS

The oxy-haemoglobin, the crystal structure of protein, .pdb files as PDB ID: 1hho were selected and incorporated separately in the input interface of HotSpot Wizard (version

2.0) online software. In this automated prediction study, chains were not specified manually.

HotSpot Wizard 2.0 is free online software for automatic detection of hot spots and design of smart libraries for engineered proteins' stability, cavity and tunnels, catalytic activity, substrate specificity and enantioselectivity [20; 27-28]. On the other hand, this present tool can be utilized for the annotation of protein structures. This tool is modified version of previous software launched in 2009 [20]. This present online server comprises sequence, structural and evolutionary information obtained from 3 databases and 20 computational tools. According to Bendl et al. [27] and Sebestova et al. [29], this online tool integrates annotated residues, which can be known easily for mutagenesis and designed for suitable codons for each implemented strategy. Ultimately, this software helps in comprehensive annotations of protein structures and engineering with the stable design of site-specific mutations and targeted libraries.

In the present study, this software was calculated automatically hotspots for function, stability, correlated and consensus sequences for oxy-haemoglobin (Fig. 1). Bendl et al. [27] have developed the workflow steps in HotSpot Wizard, the calculation is based on the particular protein annotations, mutagenesis hot spots and smart library design as first, second and final phases respectively.

For statistical analysis, Z scoring values were obtained for each computational tools such as DCA (Direct Coupling analysis), ELSC (Explicit Likelihood of Subset Variation), McBASC (McLachlan Based Substitution correlation), MI (Mutual Information), aMIc (All Microarray Clustering), OMES (Observed Minus Expected Squared) and SCA (Statistical Coupling Analysis).

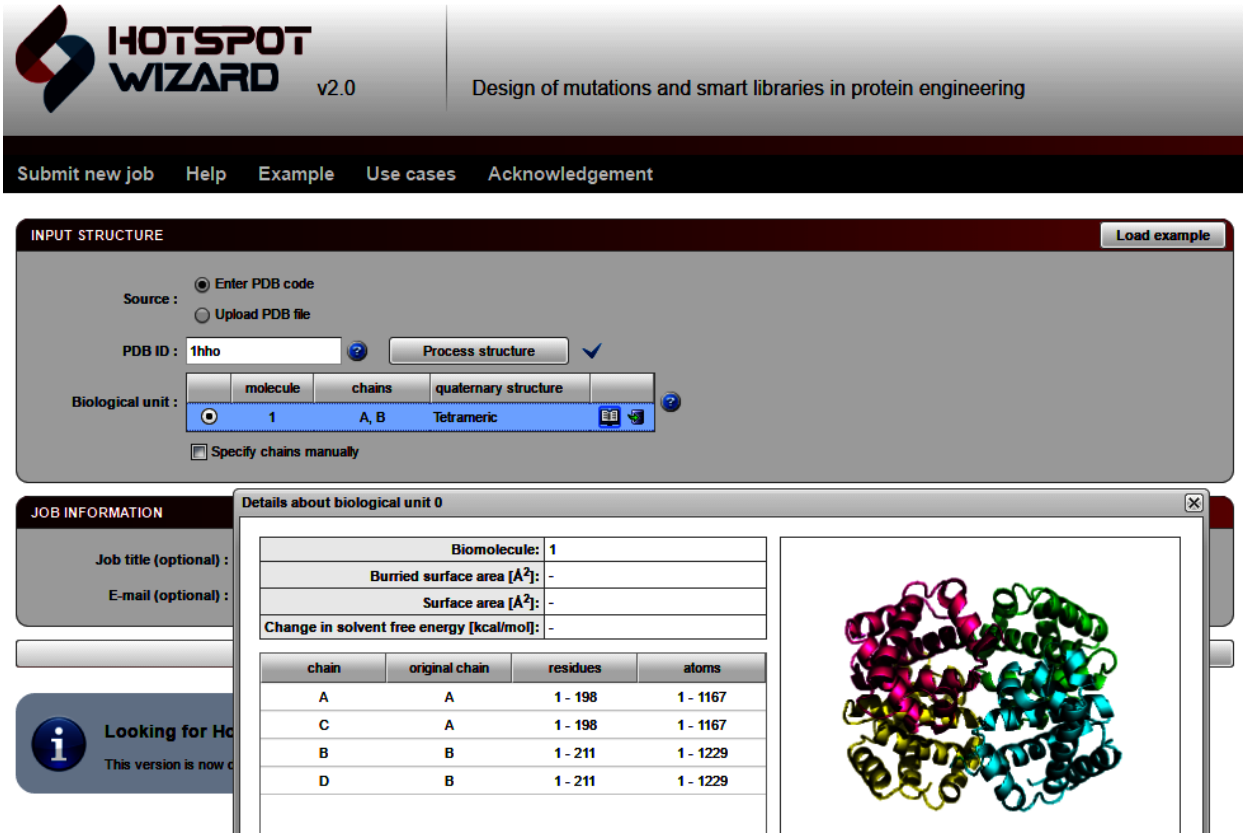



Figure 1. Hotspot wizard input interface for oxy-haemoglobin (1hho)

## III. RESULTS AND DISCUSSION

In the present results, the oxy-haemoglobin engineering strategies through automated computational prediction were observed. Fig. 2. showed results as output interface through Hotspots wizard for four separate prediction data such as functional hot spots, stability hot spots (structural flexibility), correlated hot spots and stability hot spots (sequence consensus).

In functional hot spots, the data were obtained for activity, substrate specificity and selectivity and also this step identified residues, which were forming catalytic pocket or accessible tunnel that were not directly participated in the catalysis or located at the evolutionary-conserved position.

For stability hot spots (structural flexibility), the prediction was done to identify the residues in flexible structure, which is observed mainly residues with highest B-factors.

In case of the study of correlated hot spots, the data were obtained same as functional hot spots along with the identification of correlated position through consensus approach resulted data from other computational tools viz. DCA (Direct Coupling analysis), ELSC (Explicit Likelihood of Subset Variation), McBASC (McLachlan Based Substitution correlation), MI (Mutual Information), aMIc

(All Microarray Clustering), OMES (Observed Minus Expected Squared) and SCA (Statistical Coupling Analysis). For stability hot spots (sequence consensus), consensus design is an important strategy for the stabilization of proteins. It helps amino acid conservation in sets of homologous protein to identify likely beneficial as well as deleterious mutations of the target protein.

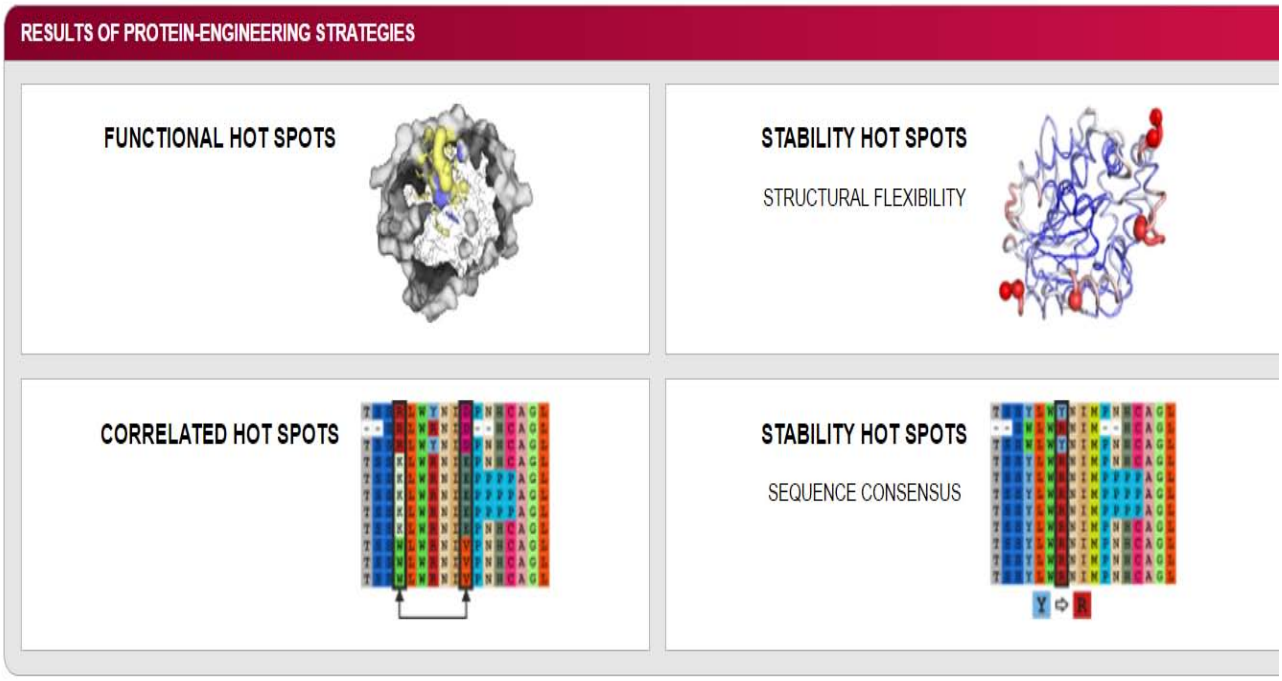
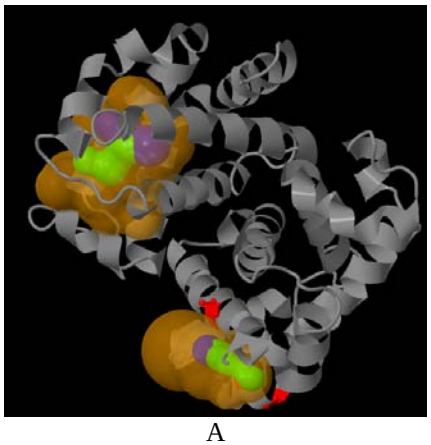


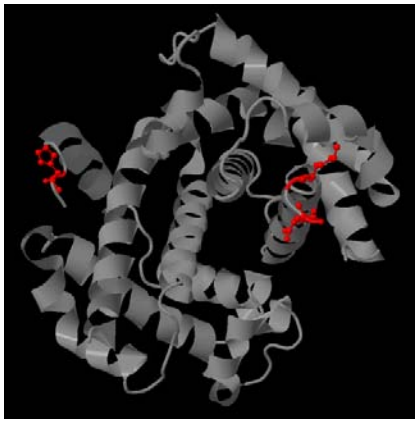

Figure 2. Protein engineering strategies of oxy-haemoglobin

Fig. 3 (A, B and D), revealed that oxy-haemoglobin showed different hotspots through HotSpot Wizard tool. In general, hot spot determines the energy distribution along with the interface region without homogenous in nature, where certain residues do not contribute majorly for free energy binding [30-35]. The hot spot prediction detects the exact protein binding sites, which helps for designing specific therapeutic agents in protein interactions [34]. In Fig. 3 (C), sequence consensus was obtained for oxyhaemoglobin. In this observation, wild-type and mutated consensus sequences were obtained based on hot spots (Richter et al., 2007). Bendl et al. [27] supported the concept of molecular mechanisms of any protein, still unclear to researchers, how the sequences of protein encode the exact function? Still have not yet answered [18-19]. It was documented that experimental evolution work suffered major problems when occurred by several irregular study of mutagenesis and detecting of large sequence libraries to evaluate the mutational landscape and proteins showed important structural and functional properties [19; 27; 37-39].

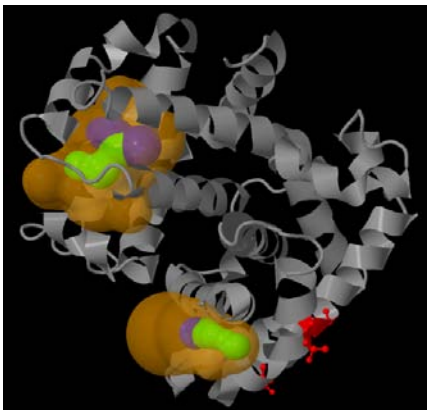
Table I, describes the functional hotspot of oxyhaemoglobin where only chain B attached to residues like Ala at 135 position, Gly at 83 position and Leu at 81 position while correlated residues like Gly at 136 position, Leu, Ser, Gly, met, Gly, Asn at 3, 9, 29, 55, 136, 139 positions respectively. The pockets and tunnels were obtained in 14 and 1, 2 (from pocket 16), 0 and & 2 (from pocket 16) and 16 (catalytic) & 1, 2 (from pocket 16) in which B-factor values 22.67, 49.39 and 46.80 $Å^2$ respectively. The B-factor values mainly influenced by crystal contacts and solvent conditions, various theoretical methods have used to predict flexible regions, which help to determine the targets for stabilization [40-41]. The mutability rate was observed high and score values were 9, 8 and 6 respectively. According to Weinkam and Salia [42], haemoglobin is a protein of complex system, which undergoes conformational changes in response to oxygen, allosteric effectors, mutations, and environmental changes. It was observed in previous study that haemoglobin has evolved with complex allosteric mechanism, which showed point mutations at different sites [43]. Weinkam and Salia [42], predicted and suggested naturally occurring mutations can be tolerated due to structural symmetry of several types of haemoglobins.



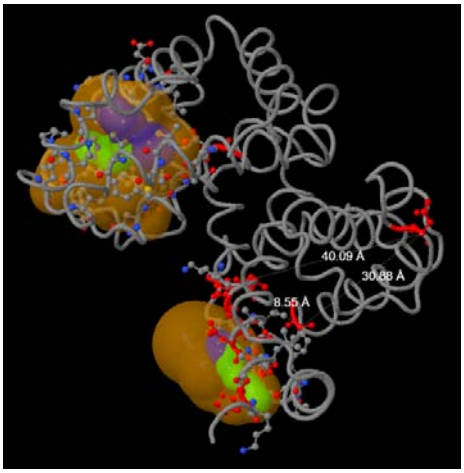
A



B.



C




D.

Figure 3. Oxy-haemoglobin protein (1hho): A = functional hotspot; B = stability hotspot; C = sequence consensus and D = correlated hotspot

Table I. Study of functional hotspots

| ***Studied Protein*** | ***Chains*** | ***Residues & position*** | ***Secondary structures*** | ***Pockets & tunnels*** | ***Average B-factor (in Å$^2$)*** | ***Mutability rate & score*** | ***Correlated residues & position*** |
|---|---|---|---|---|---|---|---|
| 1hho | B | Ala & 135 | Alpha helix (H) | 14 & 1, 2 (from pocket 16) | 22.67 | High & 9 | Gly & 136 |
| | B | Gly & 83 | Alpha helix (H) | - & 2 (from pocket 16) | 49.39 | High & 8 | --- |
| | B | Leu & 81 | Alpha helix (H) | 16 (catalytic) & 1, 2 (from pocket 16) | 46.80 | High & 6 | Leu, Ser, Gly, Met, Gly, Asn & 3, 9, 29, 55, 136, 139 |

Table II. Values obtained from different tools for functional hot spots

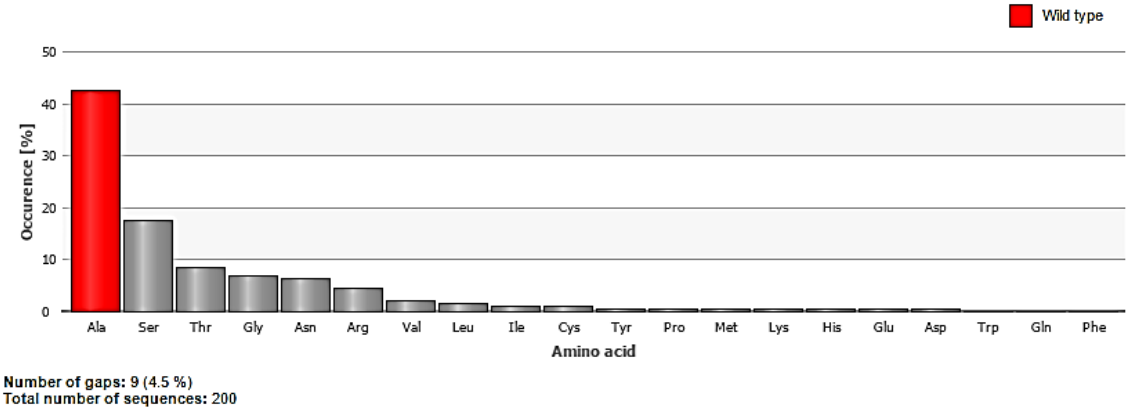
| ***Chains*** | ***Consensus z-scoring values*** | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***aMIc*** | ***DCA*** | ***ELSC*** | ***McBASC*** | ***MI*** | ***OMES*** | ***SCA*** |
| B | 0.42 | 4.37 | 4.12 | 1.01 | 5.26 | 4.05 | 7.57 |
| B | - | - | - | - | - | - | - |
| B | 2.93 | 1.70 | 4.69 | 1.02 | 5.77 | 9.24 | 7.84 |
| | 2.55 | 1.90 | 4.51 | 0.88 | 6.14 | 7.45 | 6.34 |
| | 2.28 | 1.10 | 7.94 | 1.47 | 3.40 | 7.89 | 4.32 |
| | 2.21 | 0.96 | 9.12 | 1.43 | 4.42 | 6.11 | 5.98 |
| | 3.43 | 1.58 | 11.26 | 1.57 | 5.29 | 9.58 | 4.74 |
| | 1.98 | 1.42 | 9.76 | 0.42 | 4.82 | 7.81 | 3.18 |

In Table II, consensus z-scoring value was obtained for different parameters such as aMIc 0.42 Gly, 2.93 Leu, 2.55 Ser, 2.28 Gly, 2.21 Met, 3.43 Gly, 1.98 Asn; DCA 4.37 Gly, 1.70 Leu, 1.90 Ser, 1.10 Gly, 0.96 Met, 1.58 Gly, 1.42 Asn; ELSC 4.12 Gly, 4.69 Leu, 4.51 Ser, 7.94 Gly, 9.12 Met, 11.26 Gly, 9.76 Asn; McBASC 1.01 Gly, 1.02 Leu, 0.88 Ser, 1.47 Gly, 1.43 Met, 1.57 Gly, 0.42 Asn; MI 5.26 Gly, 5.77 Leu, 6.14 Ser, 3.40 Gly, 4.42 Met, 5.29 Gly, 4.82 Asn;

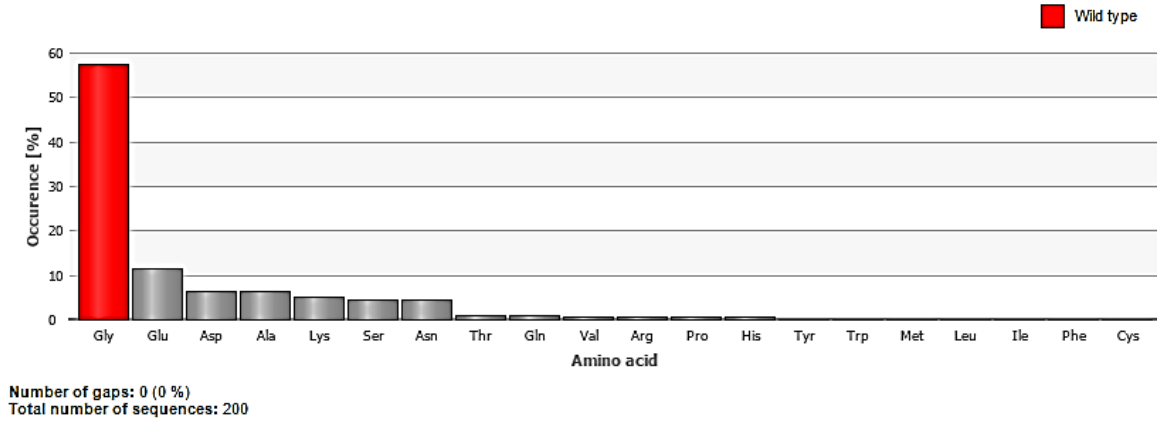
OMES 4.05 Gly, 9.24 Leu, 7.45 Ser, 7.89 Gly, 6.11 Met, 9.58 Gly, 7.81 Asn and SCA 7.57 Gly, 7.84 Leu, 6.34 Ser, 4.32 Gly, 5.98 Met, 4.74 Gly, 3.18 Asn were obtained through this tool for oxy-haemoglobin as per correlated residues.

In Fig. 4, it was obtained that the amino acid residues fulfilling the criterion of minimal frequency in the multiple sequence alignment. The wild type variety was observed Ala (42%), Gly (58%) and Leu (52%) as per positions of different amino acids frequencies of oxy-haemoglobin (Fig. 4A, B and C). Fig. 5 states that mutational landscape, which mainly showed the estimation of the probability in relation to preservation of protein function for individual substitution at a particular site of oxy-haemoglobin. It was obtained that higher deleterious mutation in Fig 5 C, followed by Fig 5 B and Fig 5A. In the present computational study, the results were obtained for β subunits, which indicated a strong linking with the quaternary transitions than the α subunits for human oxy-haemoglobin, which is supported by previous molecular dynamic simulation of unliganded haemoglobin for quaternary and tertiary T to R transitions (Hub et al., 2010). It has been documented that molecular docking study for both T-state haemoglobin and oxy-haemoglobin with different ligands of natural origin showed allosteric effects [45-46].

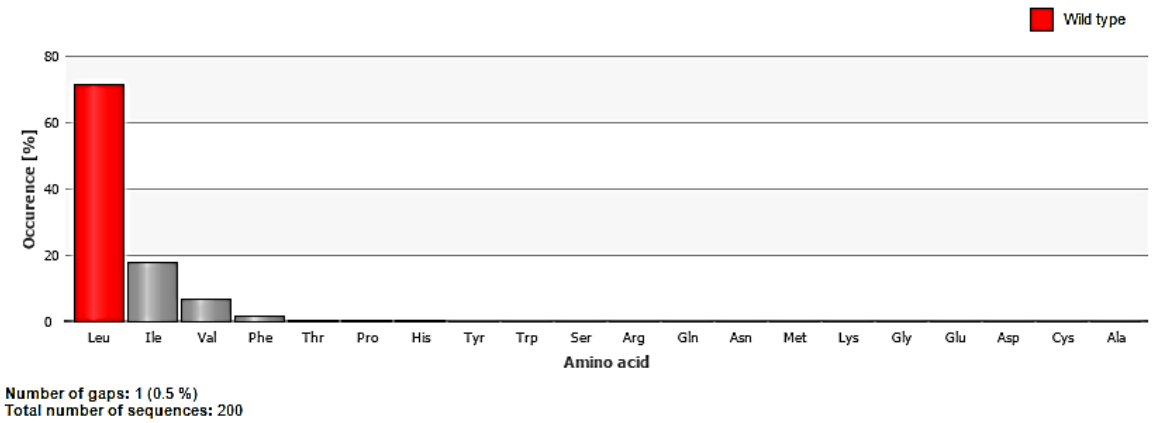
It was reported that mutability scale is ranged between 1 to 9 i.e. lower to higher rate. In the present study high mutability rate was observed 9, 8 and 6 respectively (Table I). In Fig. 5A, amino acids viz. Trp, Tyr (47%), Glu, Pro, Arg (45%) and Asp (43%), Fig. 5B, Cys, Phe, Thr, Trp (43%), Pre (39%), Ile, Tyr (31%), Val (27%) and Fig. 5C Phe (47%), Met, Val (43%), Ala, Cys, Tyr (27%), Gln, Trp (25%), Ser, Thr (23%), Gly, His, Asn, Pro (18%), Lys (12%), Asp, Glu, Arg (10%) were observed deleterious mutation of amino acids. It was reported when diseases occur then amino acid of haemoglobin undergoes mutation. According to Thom et al. [6], there are several haemoglobin variants that occurred amino acid substitutions and diseases viz. Ala>Pro, Tyr>Phe and Val>Phe (haemolytic anaemia and reticulocytosis); Val>Ala, His>Arg and Ala>Asp (haemolytic anaemia); Arg>Ser (anisocytosis and hypochromia), His>Tyr (anaemia), Phe>Ser (microcytosis); Pro>Ser (haemolytic anaemia and microcytosis); Phe>Leu (Heinz body haemolytic anaemia), Leu>Arg (Heinz body haemolytic anaemia and dominant inclusion body thalassemia), etc. found due to gene mutation in globin protein that lead to structural abnormalities of globin protein by single amino acid substitution while Pro>Arg, Lys>Glu and Lys>Asn have been detected as normal (without disease). The present prediction of oxy-haemoglobin with an evidence of high mutability score in oxyhaemoglobin (PDB ID: 1hho), which may be a clinical symptom in future research. According to Weinkama and Salia [42], haemoglobin is not a simple system and easily allows conformational changes in relation to oxygen, allosteric effectors, mutations, and environmental changes.




A.

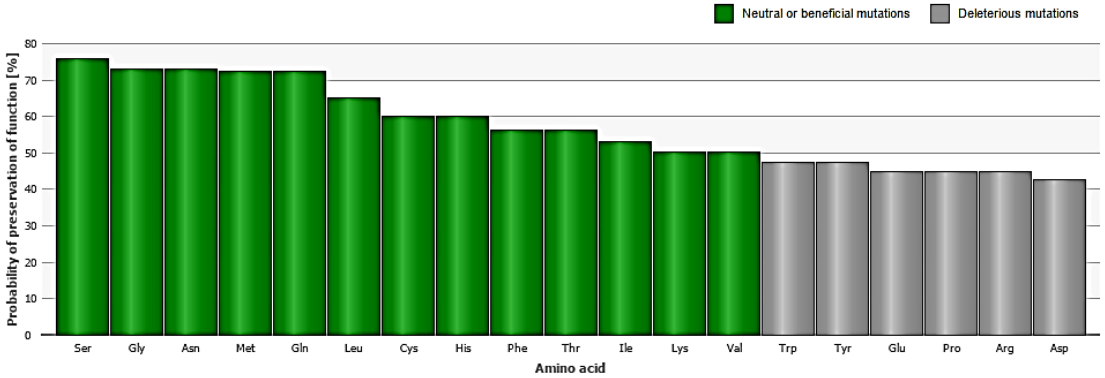



B.

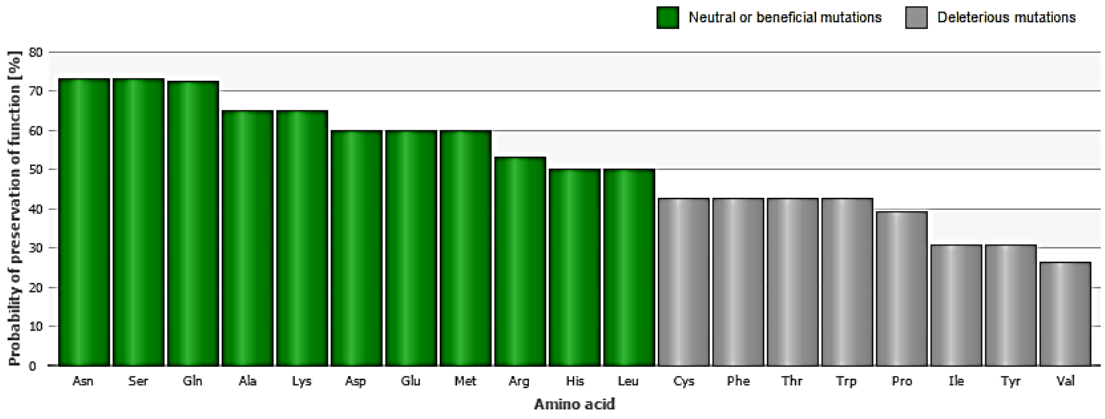



C.

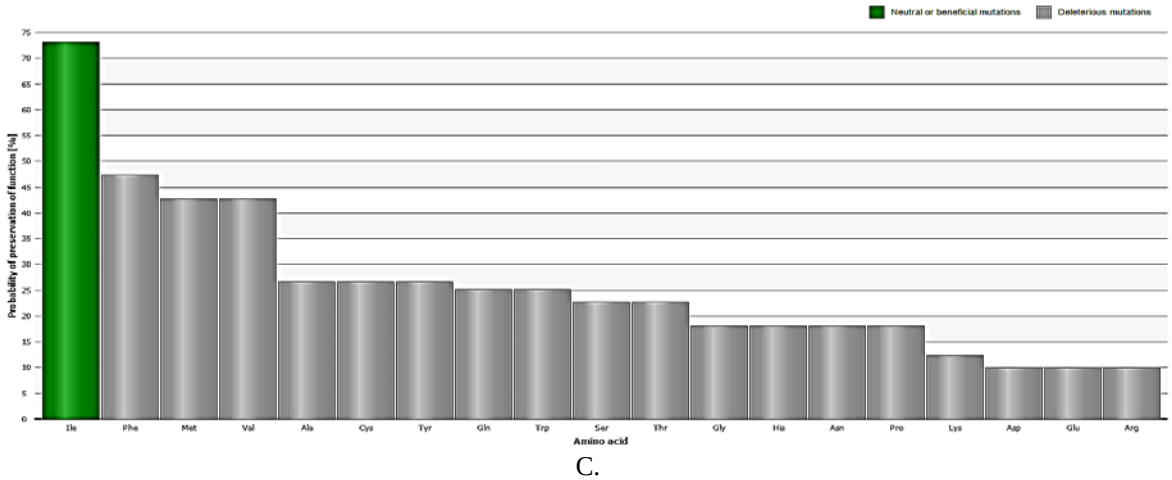
Figure 4. Amino acids frequencies as per positions




A.




B.




C.

Figure 5. Amino acids mutability landscape

However, the prediction of different hotspots can be facilitated drug designing and development. It was suggested that the starting point of a binding site of a receptor in the hotspots may be granted to analyse docking of ligands [47]. On the other hand, rigid docking lead to an achievement the comparatively least flexible hotspots, which lead to an upgradation in protein docking has been performed by creating dominant conformation of the hotspot side chains resulted through molecular dynamics probing rather than the unbound X-ray conformation [7; 48-49]. Thus, the prediction of hotspots is a suitable tool to identify exact functional mechanisms of particular protein to identify mutant residue(s) in relation to cause of disease and new drug development.

## IV. CONCLUSIONS

In conclusion, HotSpot Wizard (version 2.0) is an online computational tool, which helped easily to obtain results for oxy-haemoglobin through protein engineering protocol by the integration of several inbuilt databases derived from other bioinformatics tools and all the data generated within short duration to prevent laborious jobs of experiment [27]. This software also helped to incorporate only .pdb file as an input of studied protein without prior knowledge of computational biology to set up input interface. The parameters like pocket identification and mutability prediction of oxyhaemoglobin can lead to know structural alternation of particular in disease diagnosis as well as space for ligand binding pocket in new drug discoveries [42-43]. The present prediction work is suggesting to compare with experimental hotspots for oxy-haemoglobin in relation to therapeutic efficacies and druggability assessment.

## V. ACKNOWLEDGEMENT

Authors are thankful to the developers of online software HotSpot Wizards (Version 2.0), European Protein Databank for crystal structure (1hho), used in the present study.

## CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest for the present study.

## VI. REFERENCES

[1] F. W. Scheller, N. Bistolas, S. Liu, M. Janchen, M. Katterle and U. Wollenberger, “Thirty years of haemoglobin electrochemistry,” Advances in Colloid and Interface Science, 116 (1-3), 111-120, 2005.

[2] S. N. Khan and A. U. Khan, “An in silico approach to map the binding site of doxorubicin on hemoglobin,” Bioinformation, 2(9), 401-404, 2008.

[3] K. Kanaori, Y. Tajiri, A. Tsuneshige, I. Ishigami, T. Ogura, K. Tajima, S. Neya and T. Yonetani, “T-quaternary structure of oxy human adult hemoglobin in the presence of two allosteric effectors, L35 and IHP”, Biochimica et Biophysica Acta, 1807, 1253-1261, 2011.

[4] S. De and A. Girigoswami, “A fluorimetric and circular dichroism study of hemoglobin – Effect of pH and anionic amphiphile,” Journal of Colloid and Interface Science, 296 (1), 324-331, 2006.

[5] K. Furuyama, K. Kaneko, D. Patrick and V. Vargas, “Heme as a magnificent molecule with multiple missions: Heme determines its own fate and governs cellular homeostasis,” Tohoku Journal of Experimental Medicine, 213, 1-16, 2007.

[6] C. S. Thom, C. F. Dickson, D. A. Gell and M. J. Weiss, “Hemoglobin variants: Biochemical properties and clinical correlates,” Cold Spring Harbor Perspective in Medicine, 3, a01185 (doi: 10.1101/cshperspect.a011858), 2013.

[7] J. K. Morrow and S. Zhang, “Computational prediction of hot spot residues,” Current Pharmaceutical Design, 18(9), 1255-1265, 2012.

[8] S. Lise, D. Buchan, M. Pontil and D. T. Jones, “Predictions of hot spot residues at protein-protein interfaces using support vector machines,” PLoS One, 6(2), e16774 doi:10.1371/journal.pone.0016774, 2011.

[9] H. Wajcman, C. Prehu, J. Bardakdjian-Michau, D. Prome, J. Riou, C. Godart, M. Mathis, D. Hurtrel and F. Galacteros, “Abnormal hemoglobins: Laboratory methods,” Hemoglobin, 25, 169-181, 2001.

[10] H. Wajcman and K. Moradkhani, “Abnormal haemoglobins: Detection and characterization,” The Indian Journal of Medical Research, 134: 538-546.

[11] J. D. Bloom, J. J. Silberg, C. O. Wilke, D. A. Drummond, C., Adami and F. H. Arnold, “Thermodynamic prediction of protein neutrality,” Proceedings of the National Academy of Sciences USA, 102, 606-611, 2005.

[12] D. Bednar, K. Beerens, E. Sebestova, J. Bendl, S. Khare, R. Chaloupkova, Z. Prokop, J. Brezovsky, D. Baker and J. Damborsky, “FireProt: Energy- and evolution based computational design of thermostable multiple-point mutants,” Plos Computational Biology, 11, e1004556, 2015.

[13] A. Aharoni, L. Gaidukov, O. Khersonsky, S. M. Gould, C. Roodveldt and D. S. Tawfik, “The 'evolvability' of promiscuous protein functions,” Nature Genetics, 37, 73-76, 2005.

[14] A. Wagner, “Robustness and evolvability in living systems,” Princeton University Press, Princeton, 2005.

[15] D. Rennell, S. E. Bouvier, L. W. Hardy and A. R. Poteete, “Systematic mutation of bacteriophage T4 lysozyme,” Journal of Molecular Biology, 222, 67-87, 1991.

[16] P. Markiewicz, L. G. Kleina, C. Cruz, S. Ehret and J. H. Miller, “Genetic studies of the lac repressor. XIV. Analysis of 4000 altered Escherichia coli lac repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence,” Journal of Molecular Biology, 240, 421-433, 1994.

[17] H. H. Guo, J. Choe and L. A. Loeb, “Protein tolerance to random amino acid change,” Proceedings of the National Academy of Sciences USA, 101, 9205-9210, 2004.

[18] P. A. Romero and F. H. Arnold, “Exploring protein fitness landscapes by directed evolution,” Nature Reviews Molecular Cell Biology, 10, 866-876, 2009.

[19] A. Currin, N. Swainston, P. J. Day and D. B. Kell, “Synthetic biology for the directed evolution of protein

biocatalysts: navigating sequence space intelligently," Chemical Society Reviews, 44, 1172-1239, 2015.

[20] A. Pavelka, E. Chovancova and J. Damborsky, "HotSpot Wizard: A web server for identification of hot spots in protein engineering," Nucleic Acids Research, 37, W376-W383, 2009.

[21] Z. Zhang, Y. Li, B. Lin, M. Schroeder and B. Huang, "Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction," Bioinformatics, 27, 2083-2088, 2011.

[22] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, "CAVER 3.0: A tool for analysis of transport pathways in dynamic protein structures," Plos Computational Biology, 8, e1002708, 2012.

[23] J. Brezovsky, E. Chovancova, A. Gora, A. Pavelka, L. Biedermannova and J. Damborsky, "Software tools for identification, visualization and analysis of protein tunnels and channels," Biotechnology Advances, 31, 38-49, 2013.

[24] J. Damborsky and J. Brezovsky, "Computational tools for designing and engineering enzymes," Current Opinion in Chemical Biology, 19, 8-16, 2014.

[25] B. Kozlikova, E. Sebestova, V. Sustr, J. Brezovsky, O. Strnad, L. Daniel, D. Bednar, A. Pavelka, M. Manak, M. Bezdeka, P. Benes, M. Kotry, A. Gora, J. Damborsky and J. Sochor, "CAVER Analyst 1.0: Graphic tool for interactive visualization and analysis of tunnels and channels in protein structures," Bioinformatics, 30, 2684-2685, 2014.

[26] E. Sebestova, J. Bendl, J. Brezovsky and J. Damborsky, "Computational tools for designing smart libraries," Methods in Molecular Biology, 1179, 291-314, 2014a.

[27] J. Bendl, J. Stourac, E. Sebestova, O. Vavra, M. Musil1, J. Brezovsky, and J. Damborsky, "HotSpot Wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering," Nucleic Acids Research, 1 (doi: 10.1093/nar/gkw416), 2016.

[28] Z. Prokop, Y. Sato, J. Brezovsky, T. Mozga, R. Chaloupkova, T. Koudelakova, P. Jerabek, V. Stepankova, R. Natsume, J. G. E. Leeuwen, D. B. Janssen, J. Florian, Y. Nagata, T. Senda and J. Damborsky, "Enantioselectivity of haloalkane dehalogenases and its modulation by surface loop engineering," Angewandte Chemie International Edition 49, 6111-6115, 2010.

[29] E. Sebestova, J. Bendl, J. Brezovsky and J. Damborsky, "Computational tools for designing smart libraries," in: Directed evolution library creation: Methods and protocols, methods in molecular biology, E. M. J. Gillam, J. N. Copp and D. F. Ackerley Eds, 2nd edition, Humana Press, New York, 2014b, pp. 291-314,

[30] J. A. Wells, "Systematic mutational analyses of protein-protein interfaces," Methods in Enzymology, 202, 390-411, 1991.

[31] T. Clackson and J. A. Wells, "A hot spot of binding energy in a hormone-receptor interface," Science, 267, 383-386, 1995.

[32] A.A. Bogan and K. S. Thorn, "Anatomy of hot spots in protein interfaces," Journal of Molecular Biology, 280(1), 1-9, 1998.

[33] E. Guney, N. Tuncbag, O. Keskin and A. Gursoy, "HotSprint: database of computational hot spots in protein interfaces," Nucleic Acids Research, 36, D662–D666, 2008.

[34] N. Tuncbag, A. Gursoy and O. Keskin, "Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy," Bioinformatics, 25(12), 1513-1520, 2009.

[35] N. Tuncbag, O. Keskin and A. Gursoy, "HotPoint: hot spot prediction server for protein interfaces," Nucleic Acids Research, 38, W402-W406, 2010.

[36] S. N. Richter, G. Giaretta, V. Comuzzi, E. Leo, L. A. Mitchenall, L. M. Fisher, A. Maxwell and M. Palumbo, "Hot-spot consensus of fluoroquinolone-mediated DNA cleavage by Gram-negative and Gram-positive type II DNA topoisomerases," Nucleic Acids Research, 35(18), 6075-6085, 2007.

[37] S. Lutz, "Beyond directed evolution–semi-rational protein engineering and design," Current Opinion in Biotechnology, 21, 734-743, 2010.

[38] F. Cheng, L. Zhu and U. Schwaneberg, "Directed evolution 2.0: improving and deciphering enzyme properties," Chemical Communications. (Camb.), 51, 9760-9772, 2015.

[39] C. G. Acevedo-Rocha, M. T. Reetz and Y. Nov, "Economical analysis of saturation mutagenesis experiments," Scientific Reports, 5, 10654, 2015.

[40] M. T. Reetz, J. D. Carballeira, A. Vogel, "Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability," Angewandte Chemie International Edition England, 45, 7745-7751, 2006.

[41] H. J. Wijma, R. J. Floor and D. B. Janssen, "Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability," Current Opinion in Structural Biology, 23, 588-594, 2013.

[42] P. Weinkam and A. Salia, "Mapping polymerization and allostery of hemoglobins using point mutations," The Journal of Physical Chemistry B, 117(42), 13058-13068, 2013.

[43] B. Giardine, J. Borg, D. R. Higgs, K. R. Peterson, S. Philipsen, D. Maglott, B. K. Singleton, D. J. Anstee, A. N. Basak, B. Clark, P. Faustino, A. E. Felice, A. Francina, M. V. E. Gallivan, M. Georgitsi, R. J. Gibbons, P. C. Giordano, C. L. Harteveld, P. Joly, E. Kanavakis, P. Kollia, S. Menzel, W. Miller, K. Moradkhani, J. Old, A. Papachatzopoulou, M. N. Papadakis, P. Papadopoulos, S. Pavlovic, S. Philipsen, M. Radmilovic, C. Riemer, I. Schrijver, M. Stojiljkovic, S. L. Thein, J. Traeger-Synodinos, T. Ray, T. Wada, J. Waye, C. Wiemann, B. Zukic, D. H. K. Chui, H. Wajcman, R. C. Hardison and G. P. Patrinos, "Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach," Nature Genetics, 43, 295-302, 2011.

[44] J. S. Hub, M. B. Kubitzki, L. Bert and B. L. de Groot, "Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation," Plos Computational Biology, 6(5), e1000774, doi:10.1371/journal.pcbi.1000774, 2010.

[45] S. N. Talapatra, P. Talukdar, U. Pal, N. C. Maiti and S. Swarnakar, "Interaction of T-state haemoglobin and phytochemicals of *Hygrophila spinosa* T. Anders: An approach by molecular docking," World Journal of Pharmaceutical Research, 5(8), 1354-1369, 2016.

[46] S. Ray, P. Talukdar and S. N. Talapatra, "Molecular docking studies for phytochemicals of *Hygrophila spinosa* with oxyhaemoglobin protein," World Journal of Pharmaceutical Research 5(11), 1216-1229, 2016.

[47] D. Gonzalez-Ruiz and H. Gohlke, "Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding," Current Medicinal Chemistry, 13, 2607-2625, 2006.

[48] D. Rajamani, S. Thiel, S. Vajda and C. J. Camacho, "Anchor residues in protein-protein interactions," Proceedings of the National Academy of Sciences USA, 101, 11287-11292, 2004.

[49] S. E. Acuner Ozbabacan, A. Gursoy, O. Keskin and R. Nussinov, "Conformational ensembles, signal transduction and residue hot spots: application to drug discovery," Current Opinion in Drug Discovery and Development, 13, 527-37, 2010.