International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Protein Engineering Strategies of DNA Gyrase B: An Approach through Hotspot Wizard Online Tool

Shoumini Chakravarty[1], Soumendra Nath Talapatra[2]

[1]Ph. D. Scholar, Department of Bio - Science, Seacom Skills University, Kendradangal, Shantiniketan, Birbhum – 731236, West Bengal, India
Corresponding Author Email: *shoumini1982[at]gmail.com*
Phone: +91 – 9836847753

[2]Department of Bio - Science, Seacom Skills University, Kendradangal, Shantiniketan, Birbhum – 731236, West Bengal, India

Abstract: ***The present study was detected hot spots and protein engineering stability, substrate specificity, tunnels and cavities as well as suitable mutability position of DNA Gyrase B protein (PDB ID: 3G7B) by using an online academic software (HotSpot Wizard, version 3.1). 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). only chain A attached to residues like Val at 174 and 140 position while Arg at 144 positions, respectively. The pockets and tunnels were obtained in 2 and 2 (from pocket 1), 0 and & 2 (from pocket 1) and 1 (catalytic) and no pocket was identified in which B - factor values 24.32, 15.75 and 22.47 Å2 respectively. Moreover, the amino acid residues fulfill the criterion of minimal frequency in the multiple sequence alignment. The wild type varieties were observed Val (2%), Val (38%) and Arg (47.5%). In conclusion, the pocket identification and mutability prediction of DNA Gyrase B can lead to detect structural alternation mainly in disease diagnosis and space for ligand binding pocket in new drug design for antibacterial agent. This computational prediction suggests validating experimental hotspots for DNA Gyrase B related to therapeutic efficacies and druggability assessment.***

Keywords: Amino acids, Bioinformatics, Computational biology, DNA Gyrase B, Protein engineering, HotSpot Wizard, Mutability position

1. Introduction

DNA gyrase is an essential enzyme of bacteria, which catalyzes the ATP - dependent negative super - coiling of double - stranded closed - circular DNA. Gyrase is belonging to the class of many enzymes that well - known topoisomerases and are involved in the regulation of topological transitions of DNA. [1] Among several enzymes in bacteria, DNA gyrase protein is most effective for metabolic regulation in bacteria and help in the process of DNA replication and known as type II topoisomerase. [2 - 4] DNA gyrase is consisted by two distinct subunits, gyrase A (GyrA) and gyrase B (GyrA) and is displayed as 2 A and 2 B subunits of hetero - tetramer. GyrA holds the active site tyrosine at 122 position used in cleavage and ligation for DNA, while GyrB includes the binding and hydrolysis site for ATP. [5] GyrA has a molecular weight of 97, 000 and GyrA has a molecular weight of 90, 000. [6, 7]

Generally, protein - protein interaction, which specified residues of $\Delta\Delta G \geq 2$kcal/mol, is referred as hot spot. [8] To put it in another way, certain residues in protein - protein interactions, termed as hot spots in which stability of proteins and protein complexes predicting changes could be identified. [9] These residues have been termed 'hot spots' and if mutated then they can disrupt the interaction. [10] Most conserved amino acids are found in hot spot residues. [11]

Several tools have been developed for the prediction of hotspots at protein - protein interfaces by using a variety of models and approaches. Energy - based models, such as computational alanine scanning, use free energy functions for the estimation of the change in binding energy ($\Delta\Delta G$ binding) between the wild type and mutant protein complex upon mutation of individual amino acid residues to alanine. [12, 13]

The substitutions of Lys at 103 positions revealed the loss of ATP - dependent functions. These mutants are unable to supercoil DNA, to hydrolyze ATP, or to bind a non - hydrolysable ATP analog, namely 5* - adenylyl - b, g - imidodiphosphate (ADPNP). [14] Furthermore, the inactivating mutation at Gly at 144 position of yeast topoisomerase II did not correspond to Lys at 103 positions of gyrase, but to Gly at 117 position in the gyrase ATP binding motif. It remains possible that in a Gly at 117 gyrase mutant an induced conformational change similar to that seen in the yeast topoisomerase II mutant/wild type heterodimer would be found. [15]

From decades, the function and properties determination have developed through automated simulation by many investigators to know molecular mechanisms of any protein but still unclear the sequences of protein encode the exact function. [16, 17] Generally, an enzyme is known as biocatalyst that has restricted substrate binding ability as lock and key approach for maintaining suitable biochemical reactions in an organism. In recent research, several computational tools for protein engineering have been developed to detect tunnel and cavity, mutation positions, functions etc. [18 - 28]

In the present study an attempt was done for DNA Gyrase B 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 3.1 and the protein was used as DNA Gyrase B because this is an important protein for bacterial multiplication.

2. Materials and Methods

The DNA Gyrase B, the crystal structure of protein, pdb files as PDB ID: 3G&B was selected and incorporated separately in the input interface of HotSpot Wizard (version 3.1) online software (Fig 1). In this automated prediction study, chains were not specified manually.

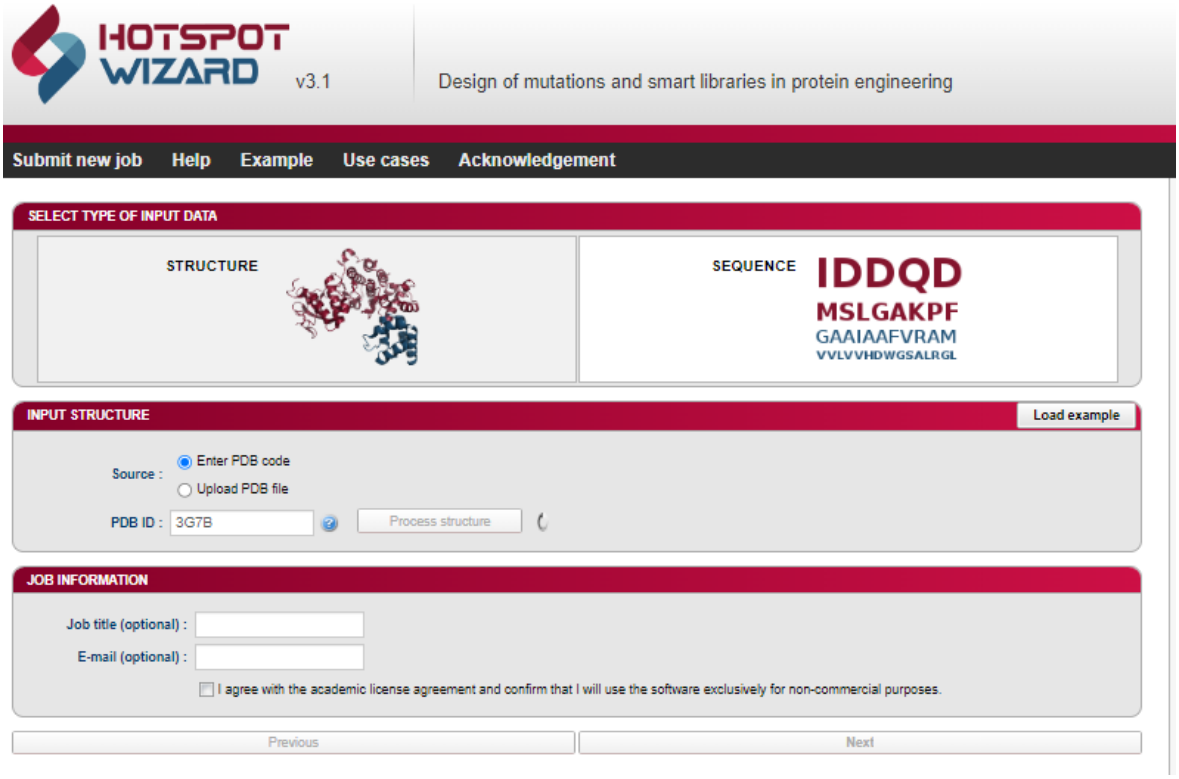

Figure 1: Input interface of Hotspot Wizard online tool

HotSpot Wizard 3.1 is free online software for academic use for the automatically detection of hot spots for engineered proteins' stability, cavity and tunnels, catalytic activity, substrate specificity and enantioselectivity. [25 - 28] On the other hand, this present tool can be utilized for the annotation of protein structures. This tool is a modified version of earlier tool launched in 2009. [25] Bendl et al. [25] and Sebestova et al., [29] developed this online tool to integrate annotated residues, which could be known easily for mutagenesis and designed for suitable codons for each implemented strategy. Ultimately, this software helps to identify data on protein engineering and stable design of site - specific mutations and targeted libraries.

In the present study, this software calculated automatically hotspots for function, stability, correlated and consensus sequences for DNA Gyrase B (Fig.1). Sumbalova et al. [27] have developed the workflow steps in HotSpot Wizard, the calculation is based on the particular protein annotations and site directed mutagenesis in hot spots.

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).

3. Results

In the present predictions of the DNA Gyrase B engineering strategies through automated computational prediction were observed. Fig.2 obtained 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).

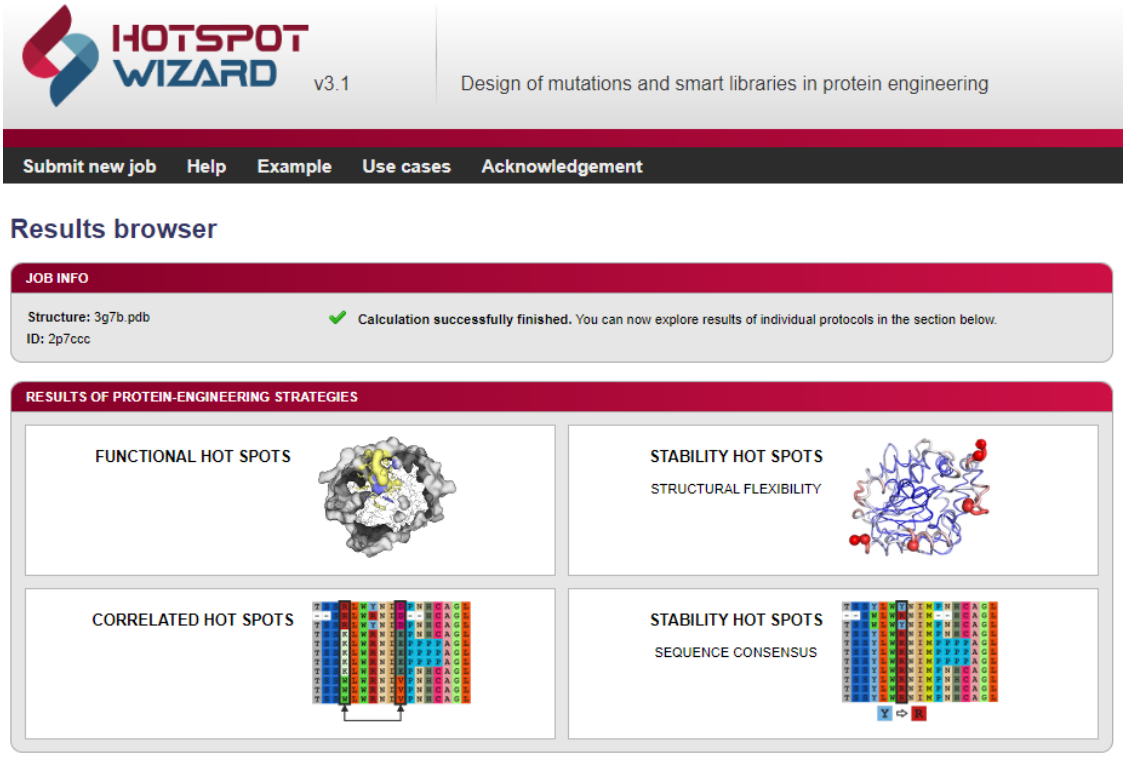

Figure 2: Output interface for protein engineering strategies of DNA Gyrase B protein

For stability hot spots (sequence consensus), consensus design is an important approach to stabilize the proteins. This supports amino acid conservation within the sets of homologous protein for the identification of beneficial and deleterious mutations of the target protein (Fig 3A - D).

In the present prediction, DNA Gyrase B receptor (PDB ID: 3G7B) was predicted (Fig 3A - 3D and Table 1), which evaluates the results from output interface through Hotspots wizard as per four separate prediction data viz. functional hot spots, stability hot spots (structural flexibility), correlated hot spots and stability hot spots (sequence consensus) (Fig 3A - 3D)

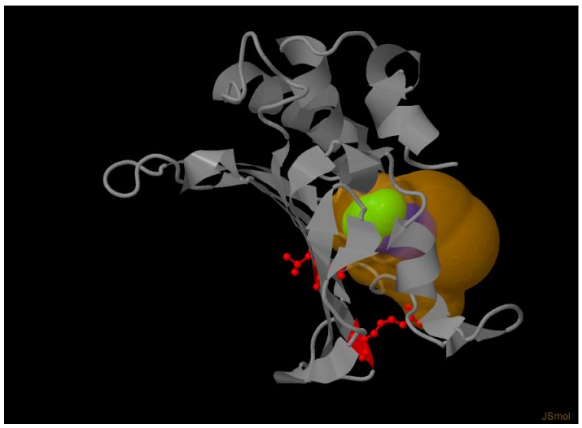

Figure 3A: Functional hotspots

Figure 3B: Stability hotspots (structural flexibility)

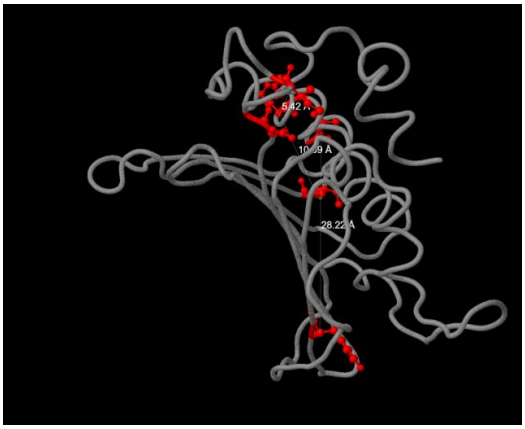

Figure 3C: Correlated hotspots

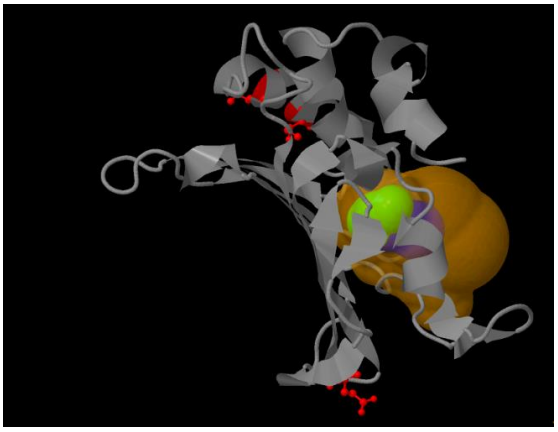

Figure 3D: Stability hotspots (sequence consensus)

Table 1 describes the functional hotspots of DNA Gyrase B where only chain A attached to residues like Val at 174 and 140 position while Arg at 144 positions, respectively. The pockets and tunnels were obtained in 2 and 2 (from pocket 1), 0 and & 2 (from pocket 1) and 1 (catalytic) and no pocket was identified in which B - factor values 24.32, 15.75 and 22.47 Å2 respectively.

Table 1: Study of functional hotspots

Studied Protein	Chains	Residues & position	Secondary structures	Pockets & tunnels	Average B - factor (in $Å^2$)	Mutability rate & score
3G7B	A	Val & 174	Extended strand (E)	2 & 2 starting from pocket 1	24.32	High & 8
	A	Val & 140	Extended strand (E)	0 & 2 starting from pocket 1	15.75	High & 6
	A	Arg & 144	Extended strand (E)	1 (catalytic), 0	22.47	High & 6

In Fig 4A - C, it was obtained that the amino acid residues fulfilling the criterion of minimal frequency in the multiple sequence alignment. The wild type varieties were observed Val (2%), Val (38%) and Arg (47.5%).

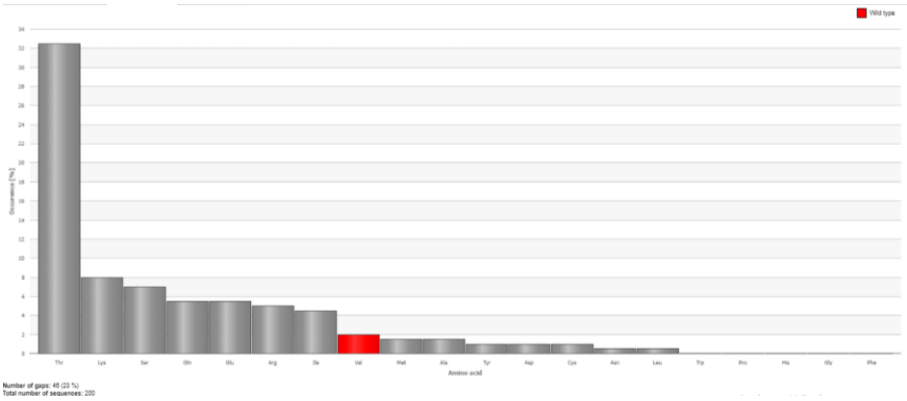

Figure 4A: Amino acids frequencies as per positions

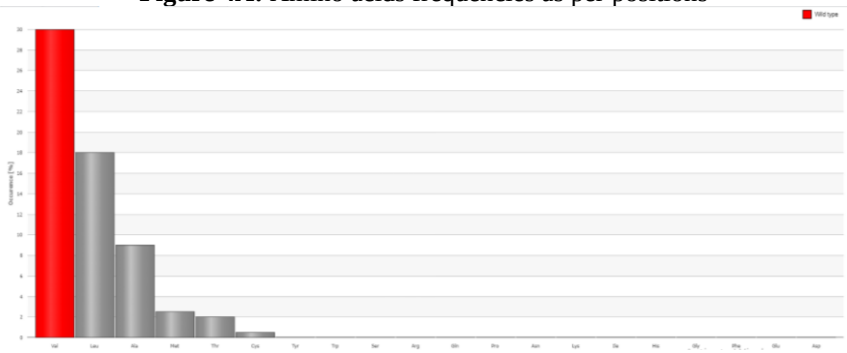

Figure 4B: Amino acids frequencies as per positions

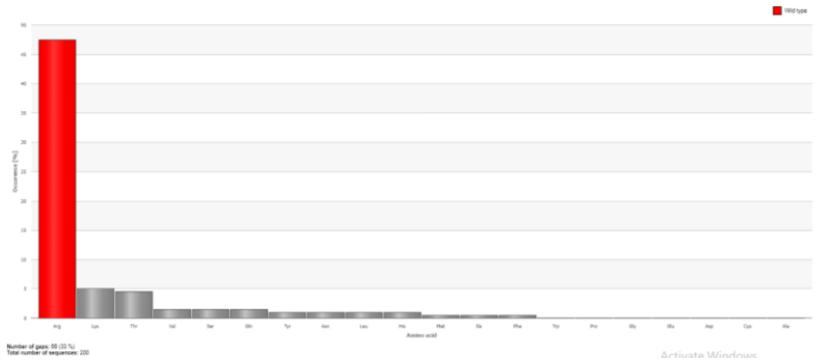

Figure 4.12C: Amino acids frequencies as per positions

For the mutational landscape, which mainly obtained the estimation of the probability in relation to preservation of protein function for individual substitution at a particular site of DNA Gyrase B. It was obtained that higher deleterious mutation in Fig 5B followed by Fig 5C and lower in Fig 5A.

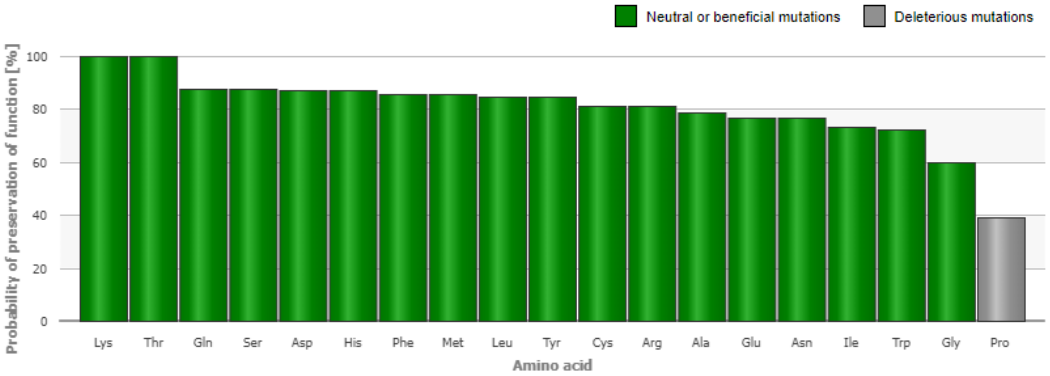

Figure 5A: Mutational landscape of DNA Gyrase B protein

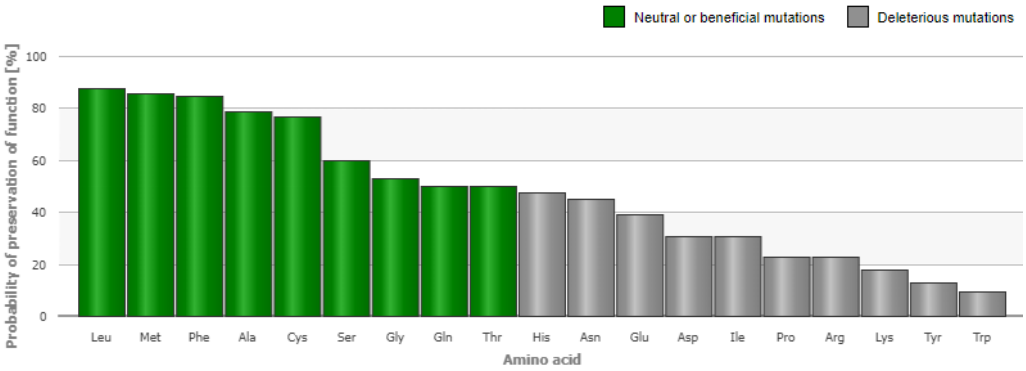

Figure 5B: Mutational landscape of DNA Gyrase B protein

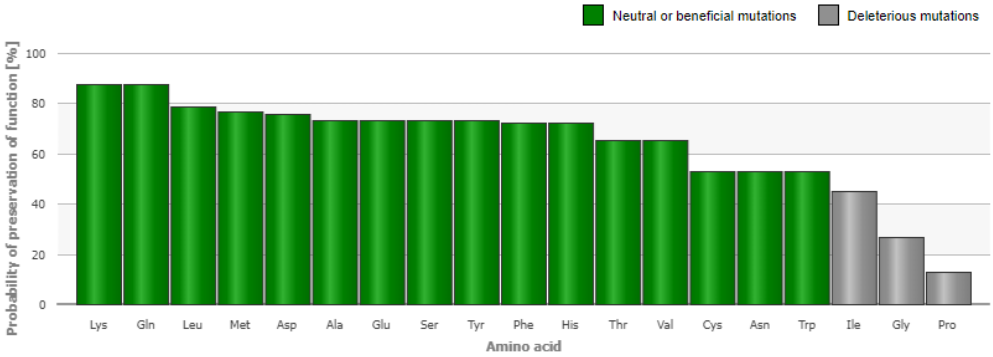

Figure 5C: Mutational landscape of DNA Gyrase B protein

In Table 2, consensus z - scoring values were obtained for different statistical parameters such as aMIc values 3.02, 3.12 and 3.41, DCA values 1.15, 1.39 and 1.73, ELSC values 12.21, 10.98 and 11.45, McBASC values 0.45, 0.61, 1.05, MI values 1.47, 1.17 and 1.21, OMES values 16.45, 18.02 and 14.61 and SCA values 2.73, - 0.09 and 1.21, respectively were obtained through this tool for DNA Gyrase B as per correlated residues.

Table 2: Values obtained from different tools for correlated hot spots

Chains	*Consensus z - scoring values*						
	aMIc	*DCA*	*ELSC*	*McBASC*	*MI*	*OMES*	*SCA*
A	3.02	1.15	12.21	0.45	1.47	16.45	2.73
A	3.12	1.39	10.98	0.61	1.17	18.02	- 0.09
A	3.41	1.73	11.45	1.05	1.21	14.61	1.21

4. Discussion

Different parameters 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) were obtained. 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 for Val and Arg. 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, ELSC, McBASC, MI, aMIc, OMES and SCA. 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.

In an earlier study, hot spot mutations were double mutation in 83 and 87 codons with substitution of Thr83→Ile and Asn87→Asp. [30] Higher the amino acid substitution and deleterious mutation may lead to antibiotic resistance. Therefore, this tool helps to identify amino acid substitution and mutability position. However, the prediction of different hotspots can be facilitated by new drug designing.

5. Conclusion

In conclusion, HotSpot Wizard (version 3.1) is an academic, free online computational tool, which helped easily to obtain results for DNA Gyrase B through protein engineering strategies as per many inbuilt databases. 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 DNA Gyrase B can lead to know amino acid substitution and site directed mutagenesis of target protein as well as space for ligand binding pocket in new drug design. This prediction work suggests validating experimental hotspots for specific protein related to therapeutic efficacies and druggability assessment.

Acknowledgement
The authors convey thanks to all developers of present software and online tool used in the predictive study, data bank for protein and phytochemicals.

Conflict of interest
No conflict of interest

References

[1] Reece RJ, Maxwell A. DNA gyrase: structure and function. Crit Rev Biochem Mol Biol.1991; 26 (3 - 4): 335 - 75.
[2] Cozzarelli NR. DNA gyrase and the supercoiling of DNA. Science.1980; 207: 953 - 960.
[3] Ronkin SM, Badia M, Bellon S, Grillot A - L, Gross CH, Grossman TH, et al. Discovery of pyrazolthiazoles as novel and potent inhibitors of bacterial gyrase. Bioorganic and Medicinal Chemistry Letters.2010; 20 (9): 2828 - 2831.
[4] Bax BD, Chan PF, Eggleston DS, Fosberry A, Gentry DR, Gorrec F, et al. Type IIA topoisomerase inhibition by a new class of antibacterial agents. Nature.2010; 466: 935 - 940.
[5] Klostermeier, D. Why two? On the role of (A -) symmetry in negative supercoiling of DNA by gyrase. Int. J. Mol. Sci.2018; 19: 1489.
[6] Horowitz DS, Wang JC. Mapping the active site tyrosine of Escherichia coli DNA gyrase. J Biol Chem.1987; 262 (11): 5339 - 44.
[7] Ali JA, Jackson AP, Howells AJ, Maxwell A. The 43 - kilodalton N - terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs. Biochemistry.1993; 32 (10): 2717 - 24.
[8] Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol.2002; 320: 369 - 87.
[9] Lise S, D. Buchan MP, Jones DT. Predictions of hot spot residues at protein - protein interfaces using support vector machines. PLoS One.2011; 6 (2): e16774.
[10] Morrow JK, Zhang S. Computational prediction of hot spot residues. Current Pharmaceutical Design.2012; 18 (9): 1255 - 65.
[11] Rorick MM, Wagner GP. The origin of conserved protein domains and amino acid repeats via adaptive competition for control over amino acid residues. J Mol Evol.2010; 70 (1): 29 - 43.
[12] Kortemme T, Kim DE, Baker D. Computational alanine scanning of protein - protein interfaces. Sci STKE.2004; 2004 (219): pl2.
[13] Huo S, Massova I, Kollman PA. Computational alanine scanning of the 1: 1 human growth hormone - receptor complex. J Comput Chem.2002; 23 (1): 15 - 27.
[14] Wigley DB, Davies GJ, Dodson EJ, Maxwell A, Dodson G. Crystal structure of an N - terminal fragment of the DNA gyrase B protein. Nature.1991; 351 (6328): 624 - 9.
[15] O'Dea MH, Tamura JK, Gellert M. Mutations in the B Subunit of Escherichia coli DNA Gyrase That Affect ATP - dependent Reactions. The Journal of Biological Chemistry.1996; 271 (16): 9723 - 9.
[16] Romero PA, Arnold FH. Exploring protein fitness landscapes by directed evolution. Nature Reviews Molecular Cell Biology.2009; 10: 866 - 76.
[17] Currin A, Swainston N, Day PJ, Kell DB. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Chemical Society Reviews.2015; 44: 1172 - 1239.

[18] Bednar D, Beerens K, Sebestova E, Bendl J, Khare S, Chaloupkova R, et al. FireProt: Energy - and evolution - based computational design of thermostable multiple - point mutants. PLoS Comput Biol.2015 Nov 3; 11 (11): e1004556.
[19] Pavelka A, Chovancova E, Damborsky J. HotSpot Wizard: a web server for identification of hot spots in protein engineering. Nucleic Acids Res.2009; 37 (Web Server issue): W376 - 83.
[20] Zhang Z, Li Y, Lin B, Schroeder M, Huang B. Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. Bioinformatics.2011; 27 (15): 2083 - 8.
[21] Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L, Sochor J, Damborsky J. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput Biol.2012; 8 (10): e1002708.
[22] Brezovsky J, Chovancova E, Gora A, Pavelka A, Biedermannova L, Damborsky J. Software tools for identification, visualization and analysis of protein tunnels and channels. Biotechnol Adv.2013; 31 (1): 38 - 49.
[23] Damborsky J, Brezovsky J. Computational tools for designing and engineering enzymes. Curr Opin Chem Biol.2014; 19: 8 - 16.
[24] Kozlikova B, Sebestova E, Sustr V, Brezovsky J, Strnad O, Daniel L, et al. CAVER analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. Bioinformatics.**2**014; 30: 2684 - 85.
[25] Bendl J, Stourac J, Sebestova E, Vavra O, Musil M, Brezovsky J, Damborsky J. HotSpot Wizard 2.0: automated design of site - specific mutations and smart libraries in protein engineering. Nucleic Acids Res.2016 Jul 8; 44 (W1): W479 - 87.
[26] Talukda, P, Talapatra SN. Oxy - haemoglobin protein engineering: An automated design for hotspots stability, site - specific mutations and smart libraries by using HotSpot Wizard 2.0 software. International Journal of Advanced Research in Computer Science.2017; 8 (2) 220 - 8.
[27] Sumbalova L, Stourac J, Martinek T, Bednar D, Damborsky J. HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information. Nucleic Acids Res.2018; 46 (W1): W356 - W362.
[28] Lahiri M, Ghosh I, Talukdar P, Talapatra, SN. Dengue virus (NS2B/NS3 protease) protein engineering. Part I: An automated design for hotspots stability and site - specific mutations by using HotSpot Wizard 3.0 tool. World Scientific News.2019; 127 (3): 177 - 90.
[29] Sebestova E, Bendl J, Brezovsky J, Damborsky J. Computational tools for designing smart libraries. In: Directed evolution library creation: Methods and protocols, methods in molecular biology. Gillam EMJ, Copp JN, Ackerley DF. (eds.).2nd edition, New York: Humana Press, 2014, pp.291 - 314.
[30] Behzadi P, Behzadi E, Moghaddam MM, Najafi A, Ranjbar R. In Silico analysis of amino acid substitutions in DNA gyrase subunit a of fluoroquinolone resistant *P. aeruginosa* TOHO strains, a glance on antibiotic development. Journal of Applied Biotechnology Reports.2014; 1 (3): 101 - 4.