Biomedical Letters ISSN 2410-955X DOI: https://doi.org/10.47262/BL/9.1.20230421

Research article
2023 | Volume 9 | Issue 2 | Pages 75-81

ARTICLEINFO

Received
March 21, 2023
Revised
May 28, 2023
Accepted
June 17, 2023

***Corresponding Author**
Tayyiba Rauf
E-mail
tayyibarauf80@gmail.com

Keywords
In silico
Structure prediction
Molecular docking
STAT3
Cancer

How to Cite
Fatima S, Sameen DE, Rauf T. Structural insights and computational molecular docking to explore novel therapeutic drug targets of STAT3. Biomedical Letters 2023; 9(2):75-81.

Note
¥Both authors contributed equally

SCAN ME

Open Access

Structural insights and computational molecular docking to explore novel therapeutic drug targets of STAT3

Sameen Fatima¥, Durr-e-Sameen¥, Tayyiba Rauf*

Department of Biology, University of Okara, Okara, Pakistan

Abstract

Signal transducer and activator of transcription 3 (STAT3) is a transcription factor, that contains a DNA-binding domain, N-terminal domain, and SH2 domain. The dysregulation of STAT3 activity has been associated with various diseases, such as chronic inflammation and autoimmune disorders. In cancer, STAT3 is often constitutively activated and promotes tumor cell survival, proliferation, and immune evasion. Various bioinformatics approaches were employed to predict the 3D structure of STAT3, followed by a comprehensive evaluation of the predicted model. 3D predicted structure of the target protein revealed an overall quality factor of 94. 45%. It was also observed through the Ramachandran plot that 1.26% residues of the predicted structure of STAT3 were present in the outlier region of the protein structure. Computational docking studies were done to identify the novel drug targets against STAT3. The screened compound *via* high throughput virtual screening may have the potential to regulate the activity of STAT3. The lowest binding energy of -8.7 Kcal/mol was observed. His-457, Tyr-456, Lys-488, Pro-487, Gln-326, Leu-459, Lys-244, Gln-247 conserved residues were observed. The structural insight and functional determination of STAT3 depend on the identification of the potent binding domain in protein 3D structure.

Introduction

Signal transducer and activator of transcription 3 (STAT3) is a transcription factor that is involved in numerous processes of cells, such as survival, differentiation and cell growth. STAT3 also contains a DNA binding domain, N-terminal domain and SH2 domain, which is considered as an essential factor for the activation and phosphorylation [1-3]. The dysregulation of the activity of the STAT3 has been associated with various diseases, such as chronic inflammation, autoimmune disorders and cancer [2-5]. In cancer, STAT3 usually promotes and activates immune evasion, proliferation and tumor cell survival in cancer [6]. In autoimmune disorders, such as multiple sclerosis, The STAT3 aberrant activation o in immune cells can lead to tissue damage and inflammation in autoimmune disorders including multiple sclerosis [7]. In chronic inflammation, STAT3 also involves in the activation of immune cells that can contribute to the progression and development of the inflammatory response in chronic inflammation [8]. To target the activity of STAT3, it has emerged as a significant therapeutic strategy for various diseases [9].

STAT3 plays a key role in response to extracellular signals, including hormones, growth factors and cytokines. The STAT3 forms homo- or heterodimers to translocate the nucleus to regulate the gene expression upon activation. STAT3 activation is strongly regulated by numerous mechanisms, such as dephosphorylation by phosphatases, protein inhibitors and negative feedback loops [10].

The STAT3 dysregulation has been implicated in several diseases, such as neurodegenerative diseases, chronic inflammation, autoimmune disorders and [5]. In cancer, the activity of STAT3 contributes to metastasis, invasion and tumor growth. The inhibitors of STAT3 are being developed as potential therapeutics for cancer [5]. In autoimmune diseases such as rheumatoid arthritis and multiple sclerosis, STAT3 also plays a significant role to regulate the inflammatory responses and immune system in autoimmune disorders including multiple sclerosis and rheumatoid arthritis [10]. In addition, STAT3 has been implicated in the pathogenesis of neurodegenerative diseases, including Parkinson's disease and Alzheimer's disease [11].

Moreover, STAT3 plays a key role in the differentiation and maintenance of stem cells during tissue repair and embryonic development. It is involved in numerous aspects of development such as tissue repair, organogenesis and embryonic development [12].

There has been a significant increase in the success of immuno-informatics studies [13-17] and also in computational drug design [18-28] in the last fifteen years. Millions of biological problems have been solved by employing the *in silico*, bioinformatics and immunoinformatics approaches [21]. The current *in silico* effort utilized molecular docking analyses to determine the novel screened compounds against STAT3. The molecular ligands having significant structural features were virtually screened against the 3D structure of STAT3. The complete 3D experimental resolved structure of the selected protein (STAT3) through Nuclear Magnetic Resonance (NMR) and X-ray crystallography was not present in Protein Data Bank (PDB) [26]. The reliable 3D predicted structure of the selected protein was predicted by utilizing the available crystal structures for further molecular docking studies.

Materials and Methods

The complete amino acid sequence of STAT3 with P4076 accession number was retrieved in FASTA format from the Uniprot Knowledgebase database. In the current study, to analyze the structural insights, the 3D prediction of STAT3 structures was predicted leads to molecular docking analyses. The sequences of the selected protein STAT3 were retrieved in FASTA format from the protein sequence database and subjected to the BLASTp to identify the suitable template for homology modeling of the target protein by using PDB [29]. Homology modeling automated software MODELLER 9.21 [30] for protein 3D structure prediction was utilized to predict the structure of the target protein. Moreover, the spatial restraints were also satisfied for the stricture prediction for reliable structure. The suitable 3D templates for structure prediction through the homology modeling approach were employed to predict the 3D structures [31, 32] of STAT3-selected protein. Numerous tools for the evaluation of 3D predicted including ERRAT [33], ProCheck [34], Rampage [35] and Anolea [36] were used to analyze the geometrical values of the predicted structure and to evaluate the predicted structures. Furthermore, the 3Dstructures predicted for the selected protein were evaluated through MolProbity [37]

ZINC commercial compound library was used for high throughput virtual screening to evaluate and analyze the novel therapeutic drug targets and the

binding domain of STAT3. AutoDock tools [38] were utilized for molecular docking analyses of protein-ligand analyses. The drug-like properties including H-bond acceptors, and the number of rotatable bonds H-bond donors were also predicted by using the PubChem tool [39]. mCule and ChEMBL databases were utilized to calculate Lipinski's rule of five [40] for the selected ligand molecules. The carcinogenicity of the selected ligands and the mutagenesis analyses of the selected ligands were analyzed to reduce the mutagenic and carcinogenic risks. The aim to study molecular docking was to identify the novel therapeutic targets and binding domain of STAT3 against the screened ligands. The geometrical optimization was minimized by employing ChemDraw and UCSC Chimera for the selected ligands and the predicted 3D structure respectively. The docked complexes of the protein and ligand were analyzed and visualized through Chimera 1.8. All the ligands were analyzed for the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties through admetSAR.

Results and Discussion

The designed project was linked with the relationship of cancer with STAT3 and computational studies for the identification of novel therapeutic targets against STAT3. The scrutinized templates from PDB against STAT3 with query coverage, maximum score, scientific names-value, accession number and identity were selected for 3D structure prediction of STAT3 through a homology modeling approach (**Table 1**). The selected templates for 3D structure prediction were applied to predict the structure of STAT3. The selected templates were used for the 3D structure prediction of the selected protein and reliable 3D structures were observed. The modeled structures of the selected protein were generated by using the solved crystal structures of the scrutinized templates. Numerous evaluation tools were used to evaluate the predicted structures and the efficacy and reliability of the predicted structures were observed (**Fig. 1**). The outlier regions, allowed regions and favored regions

Table 1: Suitable templates for STAT3 analyzed through BLASTp and sorted by scientific name, maximum score, total score, query coverage, E-value and identity.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
Unphosphorylated human STAT3 in complex with MS3-6 monobody [*Homo sapiens*]	*Homo sapiens*	1246	1246	77%	0	100%	598	6TLC
Unphosphorylated STAT3B core protein binding to dsDNA [synthetic construct]	synthetic construct	1230	1230	76%	0	99.83%	596	4E68
Transcription Factor Stat3bDNA COMPLEX [*Mus musculus*]	*Mus musculus*	1226	1226	76%	0	99.66%	596	1BG1
Lysine acetylated and tyrosine phosphorylated STAT3 in a complex with DNA [*Homo sapiens*]	*Homo sapiens*	1224	1224	76%	0	99.49%	596	6QHD
Unphosphorylated mouse STAT3 core fragment [*Mus musculus*]	*Mus musculus*	1172	1172	72%	0	100%	562	3CWG

of STAT3 predicted structures were evaluated by generating the Ramachandran plot. It was observed that 98.74% residues of the selected structure of STAT3 were also present in the allowed region of the selected protein and favored region of the selected protein of the generated Ramachandran plot however, only 1.26% residues of the predicted structure of STAT3 were present in outlier region of the Ramachandran plot. The selected predicted structure was also evaluated through the overall quality factor and 94.45% of an overall quality factor was observed leading to further *in silico* analyses (**Fig. 2**).

Furthermore, the FDA library of ZINC commercial database was used for high-throughput virtual screening. All the compounds from the selected library were docked against the STAT3 for novel therapeutic targets. The computational docking analyses were performed against the screened ligands to reveal the variation of the observed binding energy. The docking studies were performed with nineteen different poses and seventy-four runs for each pose were saved for interactional studies of the residues. It was observed that out of all the generated poses of the ligands, the effective poses of the screened ligands showed the lowest binding energy. Interestingly, it was further observed that the screened compound showed a high binding affinity with STAT3 (**Table 2**). The computational docking studies were performed and the complexes were analyzed on the basis lowest binding energy, effective drug properties of the

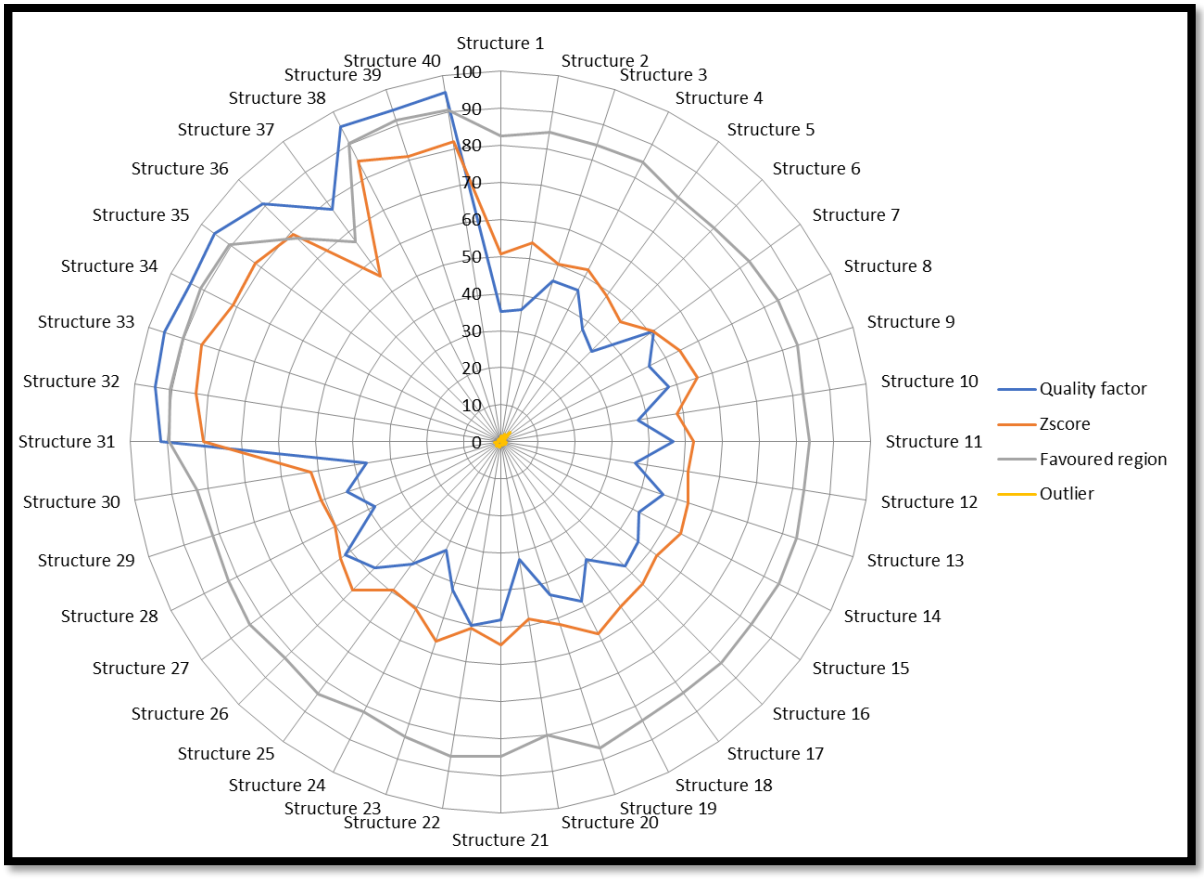

Fig. 1: Comparative structure evaluation graph showed the value of the overall quality factor, Z-score, favored region, allowed region and outlier regions of all the predicted structure of STAT3.

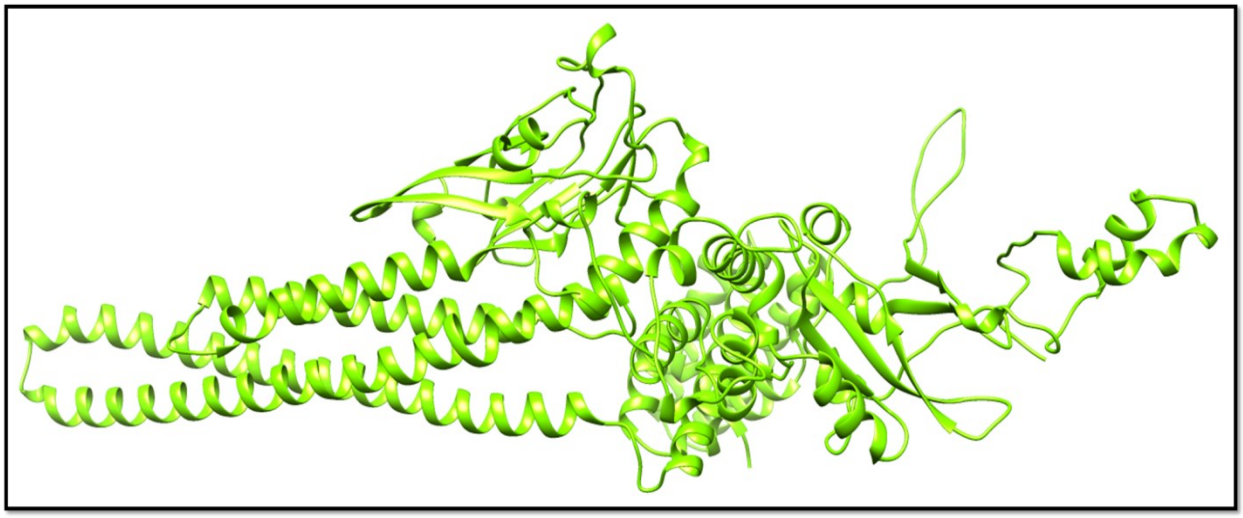

Fig. 2: The predicted 3D structure of STAT3

selected compound through high throughput virtual screening and highest binding affinity (**Table 3**). All the docked ligands were analyzed and the selected ligand through high throughput virtual screening was observed as a cyclic compound (**Fig. 3**) and showed significant biological properties to behave as a potent compound and may be considered an anticancer agent against STAT3.

Table 2: The computational molecular docking studies, the conserved binding residues of STAT3 and the efficiency of the reported ligand.

Properties	STAT3
Final intermolecular energy (kcal/mol)	-8. 71
Torsional free energy (kcal/mol)	3.63
Ligand efficiency	-0.43
Unbound system's energy (kcal/mol)	-0.49
Estimated inhibition constant, Ki (μM)	35.92
Estimated free energy of binding (kcal/mol)	-6.8
Binding residues	His-457, Tyr-456, Lys-488, Pro-487, Gln-326, Leu-459, Lys-244, Gln-247

Table 3: The drug-like efficacy and ZINC000000000509.

Ligand properties	ZINC000000000509
Acute oral toxicity (probability)	0.5100
cLogP	-0.95
Hydrogen bond acceptor	03
Honey bee toxicity (HBT) (probability)	0.8030
Rotatable bonds	00
Fish toxicity (LC50, mg/L)	0.6810
The blood-brain barrier (BBB) (probability)	0.8900
Aqueous solubility (LogS)	-1.990
Caco2 permeability (probability)	0.8430
CYP450 2D6 inhibitor (probability)	0.9000
Carcinogens (probability)	0.8900
Molecular weight (g/mol)	267.376
Hydrogen bond donor	02
Human intestinal absorption (HIA) (probability)	0.7000

The computational molecular docking approaches were applied and top-ranked complexes of the docking analyses having the least binding energy for STAT3 were selected. The variation was observed in the binding energy of all the docked complexes of the selected target protein. Interestingly, all the selected filters and parameters were satisfied for the selected ligand molecule against the selected target protein STAT3. The conserved binding region of the STAT3 was also explored through extensive computational molecular docking analyses through a compound library. It was suggested through extensive bioinformatics analyses that the observed conserved binding region of STAT3 through molecular docking analyses (**Fig. 4**) may lead to mutagenesis analyses for further studies. The process of traditional design of different drugs is very time-consuming and costly [22]. Thus, different approaches of computational biology and bioinformatics have been employed for the design of novel compounds against different targets [24]. The advancement in bioinformatics and computational biology techniques showed a significant decrease in time for traditional drug design along with minimum side effects of the reported compounds [41]. The screened compound through high throughput virtual screening was extensively evaluated for a drug target and oral bioavailability and efficacy [42] were also calculated. The ADMET properties were also calculated for the screened compounds. Various toxicity analyses were also calculated for the screened compounds (Table 3). The pollutants, intermediates and metabolites [19] were evaluated through various calculated toxicities.

LogP values were evaluated for the screened compounds and fewer LogP values were observed results to follow Lipinki's rule of five. It was further evaluated that the screened compounds were non-carcinogenic and were not mutant through different carcinogenicity and toxicity risk assessment analyses. Detailed computational analyses and literature surveys suggested that the drug target should satisfy the parameter of lowest binding energy, effective drug properties and highest binding affinity. By following the design parameters, it was observed that the screened molecule may have the potential to use against cancer by targeting STAT3. The computational docking analyses and drug-like analyses showed that the His-457, Tyr-456, Lys-488, Pro-487, Gln-326, Leu-459, Lys-244, and Gln-247 binding residues were crucial.

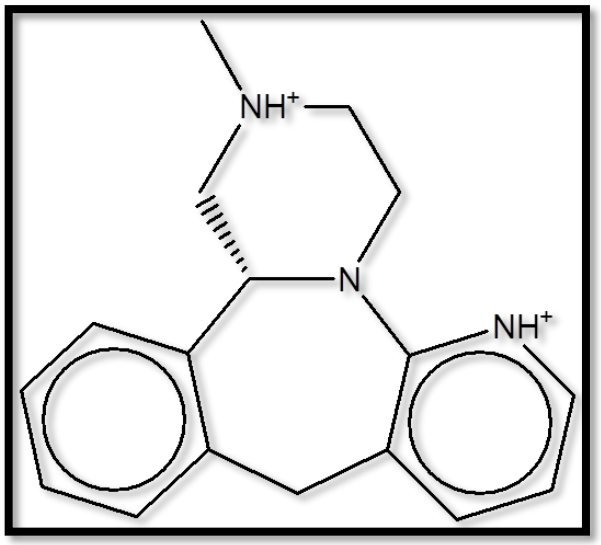

Fig. 3: The 2D structure of ZINC000000000509 compound screened by high throughput virtual screening.

Conclusion

In conclusion, the 3D structure of STAT3 was predicted by applying homology modeling, threading and *ab initio* approaches and various evaluation tools

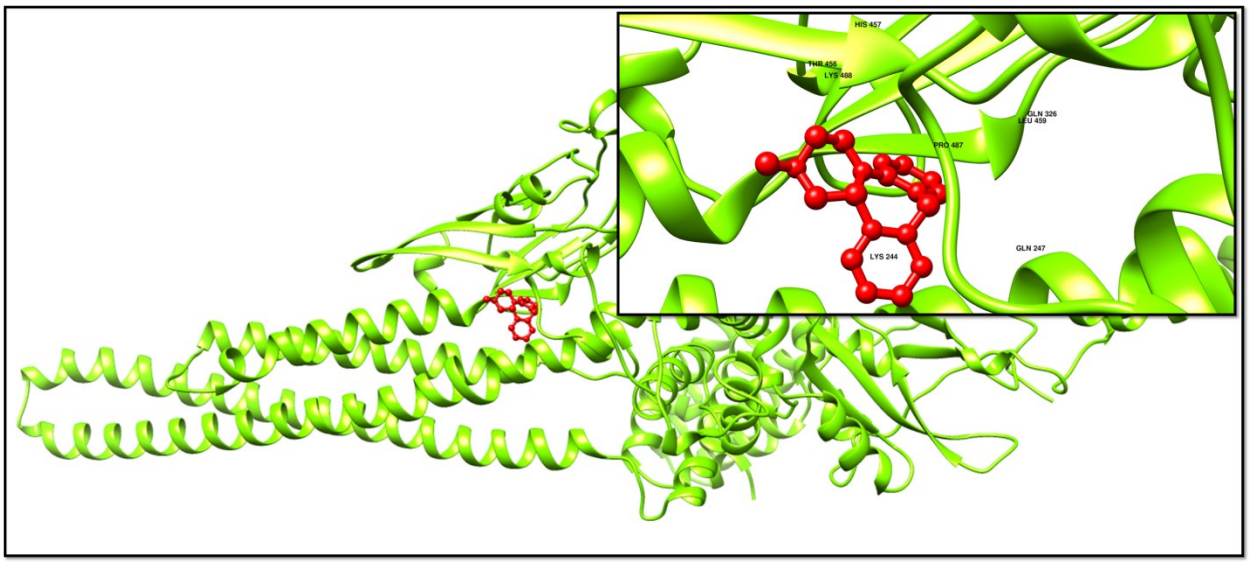

Fig. 4: The conserved interacting residues of STAT3 along with the docked complex of the screened compound as a novel therapeutic target.

were employed to evaluate the predicted structures. The screened compound through high throughput virtual screening showed efficiency against cancer by targeting STAT3 by applying extensive bioinformatics analyses. His-457, Tyr-456, Lys-488, Pro-487, Gln-326, Leu-459, Lys-244, and Gln-247 residues were observed as conserved residues and the reported compound may be significant for site-directed mutagenesis. The *in silico* analyses suggested that the observed binding domain may use against further cancer studies. The generated results concluded that the reported compound may stabilize the STAT3 and may serve as a lead compound.

Acknowledgments

Authors are thankful to Muhammad Farhan Shafique and Miss Shahista Shaoor from Muhammadan Laboratory of Computational Biology, Pakistan for their help, support and guidance in this project.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Ma Y, Hao JJTiCB. Dedifferentiation of Epithelial Cells Incorporates Immune Reprogramming.2021;31(4):237-240.

[2] Ernst M, Najdovska M, Grail D, Lundgren-May T, Buchert M, Tye H, Matthews VB, Armes J, Bhathal PS, Hughes NRJTJoci. STAT3 and STAT1 mediate IL-11–dependent and inflammation-associated gastric tumorigenesis in gp130 receptor mutant mice.2008;118(5):1727-1738.

[3] Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger BJI. TGFβ in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells.2006;24(2):179-189.

[4] Kusaba T, Nakayama T, Yamazumi K, Yakata Y, Yoshizaki A, Inoue K, Nagayasu T, Sekine IJOr. Activation of STAT3 is a marker of poor prognosis in human colorectal cancer.2006;15(6):1445-1451.

[5] Johnson DE, O'Keefe RA, Grandis JRJNrCo. Targeting the IL-6/JAK/STAT3 signalling axis in cancer.2018;15(4):234-248.

[6] Liu X, Hu H, Fan H, Zuo D, Shou Z, Liao Y, Nan Z, Tang QJM. The role of STAT3 and AhR in the differentiation of CD4+ T cells into Th17 and Treg cells.2017;96(17).

[7] Harris TJ, Grosso JF, Yen H-R, Xin H, Kortylewski M, Albesiano E, Hipkiss EL, Getnet D, Goldberg MV, Maris CHJTJoI. Cutting edge: An in vivo requirement for STAT3 signaling in TH17 development and TH17-dependent autoimmunity.2007;179(7):4313-4317.

[8] Grivennikov S, Karin MJDlS, collaboration N-k, Rev cicCGF. reviews gf (2010).21:11-19.

[9] Papavassiliou KA, Marinos G, Papavassiliou AGJC. Combining STAT3-Targeting Agents with Immune Checkpoint Inhibitors in NSCLC.2023;15(2):386.

[10] Yamamoto M, Yamazaki S, Uematsu S, Sato S, Hemmi H, Hoshino K, Kaisho T, Kuwata H, Takeuchi O, Takeshige KJN. Regulation of Toll/IL-1-receptor-mediated gene expression by the inducible nuclear protein IκBζ.2004;430(6996):218-222.

[11] Yu H, Jove RJC. Nature Rev.2004;4:97-105.

[12] Atkinson GP, Nozell SE, Benveniste ENJEron. NF-κB and STAT3 signaling in glioma: targets for future therapies.2010;10(4):575-586.

[13] Waqas M, Haider A, Sufyan M, Siraj S, Sehgal SA. Determine the Potential Epitope Based Peptide Vaccine Against Novel SARS-CoV-2 Targeting Structural

Proteins Using Immunoinformatics Approaches. Front Mol Biosci.2020;7:227.
[14] Tahir RA, Wu H, Rizwan MA, Jafar TH, Saleem S, Sehgal SA. Immunoinformatics and molecular docking studies reveal potential epitope-based peptide vaccine against DENV-NS3 protein. J Theor Biol.2018;459:162-170.
[15] Tahir RA, Bashir A, Yousaf MN, Ahmed A, Dali Y, Khan S, Sehgal SA. In Silico identification of angiotensin-converting enzyme inhibitory peptides from MRJP1. PLoS One.2020;15(2):e0228265.
[16] Sajid M, Marriam S, Mukhtar H, Sohail S, Sajid M, Sehgal SA. Epitope-based peptide vaccine design and elucidation of novel compounds against 3C like protein of SARS-CoV-2. PLoS One.2022;17(3):e0264700.
[17] Waqas M, Haider A, Rehman A, Qasim M, Umar A, Sufyan M, Akram HN, Mir A, Razzaq R, Rasool D, Tahir RA, Sehgal SA. Immunoinformatics and Molecular Docking Studies Predicted Potential Multiepitope-Based Peptide Vaccine and Novel Compounds against Novel SARS-CoV-2 through Virtual Screening. Biomed Res Int.2021;2021:1596834.
[18] Tahir RA, Hassan F, Kareem A, Iftikhar U, Sehgal SA. Ligand-Based Pharmacophore Modeling and Virtual Screening to Discover Novel CYP1A1 Inhibitors. Curr Top Med Chem.2019;19(30):2782-2794.
[19] Tahir RA, Mughal S, Nazir A, Noureen A, Jawad A, Waqas M, Sehgal SA. The Life Cycle and in silico Elucidation of Non-structural Replicating Proteins of HCV Through a Pharmacoinformatics Approach. Comb Chem High Throughput Screen.2022;25(4):689-701.
[20] Sehgal SA, Kanwal S, Tahir RA, Khalid Z, Hammad MA. In silico elucidation of potential drug target sites of the Thumb Index Fold Protein, Wnt-8b. Tropical Journal of Pharmaceutical Research.2018;17(3):491.
[21] Sehgal SA, Khattak NA, Mir A. Structural, phylogenetic and docking studies of D-amino acid oxidase activator (DAOA), a candidate schizophrenia gene. Theor Biol Med Model.2013;10:3.
[22] Sehgal SA, Mannan S, Ali S. Pharmacoinformatic and molecular docking studies reveal potential novel antidepressants against neurodegenerative disorders by targeting HSPB8. Drug Des Devel Ther.2016;10:1605-18.
[23] Tahir RA, Wu H, Javed N, Khalique A, Khan SAF, Mir A, Ahmed MS, Barreto GE, Qing H, Ashraf GM, Sehgal SA. Pharmacoinformatics and molecular docking reveal potential drug candidates against Schizophrenia to target TAAR6. J Cell Physiol.2019;234(8):13263-13276.
[24] Kanwal S, Jamil F, Ali A, Sehgal SA. Comparative Modeling, Molecular Docking, and Revealing of Potential Binding Pockets of RASSF2; a Candidate Cancer Gene. Interdiscip Sci.2017;9(2):214-223.
[25] Sehgal SA, Mannan S, Kanwal S, Naveed I, Mir A. Adaptive evolution and elucidating the potential inhibitor against schizophrenia to target DAOA (G72) isoforms. Drug Des Devel Ther.2015;9:3471-80.
[26] Sehgal SA, Hassan M, Rashid S. Pharmacoinformatics elucidation of potential drug targets against migraine to target ion channel protein KCNK18. Drug Des Devel Ther.2014;8:571-81.
[27] Sehgal SA. Pharmacoinformatics, Adaptive Evolution, and Elucidation of Six Novel Compounds for Schizophrenia Treatment by Targeting DAOA (G72) Isoforms. Biomed Res Int.2017;2017:5925714.
[28] Tahir RA, Sehgal SA. Pharmacoinformatics and Molecular Docking Studies Reveal Potential Novel Compounds Against Schizophrenia by Target SYN II. Comb Chem High Throughput Screen.2018;21(3):175-181.
[29] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res.2000;28(1):235-42.
[30] Eswar N, Eramian D, Webb B, Shen MY, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol.2008;426:145-59.
[31] Sehgal Sheikh Arslan AM, RA Tahir, A Mir.2018.
[32] Sehgal SA, Tahir RA, Waqas M. Quick Guideline for Computational Drug Design (Revised Edition) 2021.
[33] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci.1993;2(9):1511-9.
[34] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography.1993;26(2):283-291.
[35] Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins.2003;50(3):437-50.
[36] Melo F, Devos D, Depiereux E, Feytmans E. ANOLEA: a www server to assess protein structures. Proc Int Conf Intell Syst Mol Biol.1997;5:187-90.
[37] Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr.2010;66(Pt 1):12-21.
[38] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem.2010;31(2):455-61.
[39] E. Bolton * YW, Paul A. Thiessen *, Stephen H. Bryant chapter 12 -PubChem: Integrated Platform of Small Molecules and Biological Activities.2008: 217-241.
[40] Leeson P. Drug discovery: Chemical beauty contest. Nature.2012;481(7382):455-6.
[41] Zaka M, Sehgal SA, Shafique S, Abbasi BH. Comparative in silico analyses of Cannabis sativa, Prunella vulgaris and Withania somnifera compounds elucidating the medicinal properties against rheumatoid arthritis. J Mol Graph Model.2017;74:296-304.
[42] Sehgal SA, Adnan Tahir R. Molecular Modeling and Docking Analysis of CYP1A1 Associated with Head and Neck Cancer to Explore its Binding Regions. Journal of Theoretical and Computational Science.2014;1(3):6.