Volume Number 10 || Issue Number 1 || Year 2025 || Page 43 || * Corresponding Author: alzahraniaar@bu.edu.sa||

ISSN (Online) = 2522-6754 ISSN (Print) = 2522-6746

WORLD JOURNAL OF BIOLOGY AND BIOTECHNOLOGY

Research Manuscript www.sciplatform.com Peer review

A molecular dynamics simulations analysis of repurposing drugs for Marburg Virus using bioinformatics methods

Abdulaziz Alzahrani

Pharmaceuticals Chemistry Department, Faculty of Pharmacy, Al-Baha University, Alaqiq, Saudi Arabia.

Contribution | **Abdulaziz Alzahrani** conducted the research, analyse the data, wrote the entire manuscript

ABSTRACT

The increasing demand for accelerated and cost-effective drug discovery has prompted researchers to adopt computational approaches, leading this study to focus on repurposing FDA-approved drugs as potential inhibitors against Marburg virus (MARV) by targeting its envelop glycoprotein (UniProt ID: P35253), whose 3D structure was predicted using I-TASSER. Virtual screening of FDA-approved compounds via PyRx and AutoDock Vina identified ZINC000012503187 and ZINC000096006020 as top candidates, with binding interactions analyzed in BIOVIA Discovery Studio and complex stability assessed through Desmond MD simulations. These compounds exhibited strong binding affinity and antiviral potential, suggesting their repurposing as effective MARV inhibitors with minimal side effects, though further experimental validation is necessary to confirm their therapeutic efficacy.

Keywords: Molecular dynamics simulation, marburg virus, envelop glycoprotein

INTRODUCTION: The threat presented by the Marburg Virus looms big in the never-ending search for viable treatments against newly and remerging infectious illnesses. Similar to the infamous Ebola Virus and a member of the Filoviridae family, the Marburg Virus has been linked to rare but serious outbreaks with high rates of morbidity and fatality (Bausch and Schwarz, 2014). The challenge of creating specialized antiviral drugs persists after decades of study, prompting experts to look for new approaches to medication discovery and development (Malvy *et al.*, 2019). Computational biology, bioinformatics, and molecular dynamics simulations have all converged because of the need for quick, inexpensive, and precision-driven drug discovery (Iqbal *et al.*, 2020; Lin *et al.*, 2020). This study aims to use these combined approaches to identify target proteins and repurpose FDA-approved medications as possible Marburg Virus inhibitors, providing a viable way to quicken the development of antiviral treatments (Pushpakom *et al.*, 2019).

Due to its ability to cause epidemics with a high mortality rate, the Marburg Virus poses a serious threat to the security of the global health system. Traditional drug discovery methods are burdened by expensive and time-consuming procedures, sometimes necessitating years of research and development before a medicine is commercially available (Gatherer, 2014; Grobler *et al.*, 2020). In contrast, drug repurposing offers an enticing avenue for accelerated drug development by bypassing the early stages of preclinical and clinical trials, thus hastening the availability of life-saving treatments (Ashburn and Thor, 2004). An essential first step in repurposing drugs for Marburg Virus is the identification of target proteins crucial to the virus's life cycle (Morrissey *et al.*, 2008). Extensive literature review led us to identify the envelope glycoprotein (UniProt ID: P35253) as a pivotal target. The precise 3D structure of this glycoprotein was subsequently determined using the sophisticated I-TASSER computational tool. This structural insight underpins the subsequent stages of drug discovery, allowing for precise and efficient targeting of potential inhibitors (Yang *et al.*, 2015).

One of the pivotal advantages of drug repurposing lies in the wealth of information available for FDA-approved drugs. These drugs come with established safety profiles, pharmacokinetics, and toxicity data, all of which expedite the transition to clinical trials (Trott and Olson, 2010; Dallakyan and Olson, 2015). The research at hand exploits this treasure trove by systematically screening the FDA-approved drug database, identifying compounds with the potential to inhibit Marburg Virus. This in silico screening process is meticulously executed using AutoDock Vina, a robust molecular docking tool, and Pyrx (Bowers *et al.*, 2006; Iqbal *et al.*, 2023).

To further refine the selection of potential inhibitors, we conduct rigorous validation of compound-target protein binding interactions using BIOVIA Discovery Studio (Qi *et al.*, 2015). This step is pivotal in ensuring that the chosen compounds exhibit a high affinity for the target protein while minimizing the risk of off-target effects. This meticulous validation process significantly enhances the likelihood of identifying drugs with both potency and specificity against Marburg Virus (Clark *et al.*, 2012; Anohar *et al.*, 2023). Subsequently, we employ Desmond's Molecular Dynamics Simulation (MD simulation) to assess the stability and dynamics of protein-inhibitor complexes within a physiologically relevant environment (Shaw *et al.*, 2014). These simulations offer invaluable insights into the conformational changes and interactions occurring over time, providing a comprehensive view of the inhibitory potential of the selected compounds (Callahan *et al.*, 2017).

The research findings have unveiled two FDA-approved drugs, ZINC000012503187 and ZINC000096006020, as promising inhibitors capable of curtailing the action of Marburg. These compounds not only demonstrate heightened efficacy but also promise fewer side effects, rendering them compelling candidates for further preclinical and clinical investigations. This discovery constitutes a significant contribution to the scientific community, kindling optimism in the ongoing battle against Marburg Virus. In this era marked by rapid technological advancements and interdisciplinary collaborations, the amalgamation of computational biology, bioinformatics, and molecular dynamics simulations for the repurposing of existing drugs represents an avenue brimming with potential. The outcomes of this research underscore the promise of expedited drug development and kindle hope for the timely creation of effective treatments against Marburg Virus, thereby bolstering global public health (Kortepeter *et al.*, 2020; Abir *et al.*, 2022).

OBJECTIVES: The objective of the study is to identify lead drug candidates to treat diseases caused by Marburg virus by exploring FDA drugs using bioinformatics approaches.

MATERIALS AND METHODS: **Target protein sequence retrieval and three-dimensional structure prediction:** The 3D structure of targeted proteins was not present in RCSB PDB. Protein Data Bank provide 3D structural data and related information for macromolecules such as proteins, DNA, and RNA via web based information centre and a data archive that can be downloaded (Berman *et al.*, 2000). We retrieve sequence of target protein from UniProt (ID: P35253). After that we submit it sequence to the I-TASSER server to predict its three dimensional structure (Yang *et al.*, 2015).

Protein optimization, minimization and binding site prediction: For loop refinement, we used MODELLER, while Swiss PDB Viewer and RAMPAGE were used for protein crystal structure optimization and minimization. Additionally, RAMPAGE included a Ramachandran Plot, which demonstrated the distribution of residues in the preferred, permitted, and outlier areas and demonstrated the lack of any protein conflicts. Using the CASTp (Computed Atlas of Surface Topography of Proteins) database, we were able to estimate the binding sites for the protein target. The most recent version, CASTp 3.0, provides a trustworthy and thorough evaluation of protein topography, precisely detecting and measuring binding sites (Guex and Peitsch, 1997; Ho and Brasseur, 2005; Eswar *et al.*, 2006; Tian *et al.*, 2018).

The FDA drugs library preparation and molecular docking: A library of 1615 chemicals drawn from ZINC database-available, FDA-approved medications was created. These substances were located in the SD File format, and PyRx was then used to import them for docking investigations. We performed docking simulations using AutoDock Vina between these substances and their corresponding receptors to examine their binding affinities and protein-ligand interactions. PyMOL was used to create complex receptor and ligand files, while BIOVIA Discovery Studio was used to investigate two-dimensional interactions (Mura *et al.*, 2010; Trott and Olson, 2010; Dallakyan and Olson, 2015; Irwin *et al.*, 2020; Systèmes, 2022).

Molecular dynamics simulation: We ran molecular dynamics simulations lasting 100 nanoseconds using the Desmond program from Schrödinger LLC. We first performed protein-ligand docking as an essential first step to build the static representation of the molecule's binding site within the target's active site before starting the molecular dynamics modeling (MD). To put it simply, MD models simulate atomic motions across time and let us forecast the ligand-binding state in a physiological setting (Bowers, 2006; Ferreira *et al.*, 2015; Hildebrand *et al.*, 2019; Rasheed *et al.*, 2021). We optimized, reduced, and addressed missing residues in the ligand-receptor combination as well as the system built using the System Builder tool utilizing Maestro's Protein Preparation Wizard. We used the TIP3P (Intermolecular Interaction Potential 3 Points Transferable) fluid model with a temperature of 310 K, a pressure of 1 atm, and the OPLS_2005 force field to recreate physiological circumstances. The models were neutralized by adding the proper ions, and 0.15 M sodium chloride was added to mimic physiological conditions. The models were relaxed before modeling, and 100 ps-long images were taken on a regular basis. We performed principal component analysis (PCA) and dynamic cross-correlation matrix (DCCM) calculations using the R 'Bio3D' package for advanced studies (Shivakumar *et al.*, 2010; Grant *et al.*, 2021).

RESULTS AND DISCUSSION: The 3D structure of receptor was obtained by utilizing I-TASSER. The protein has 681 residues with 74376.19 molecular weight. Its theoretical isoelectric point is 5.88.

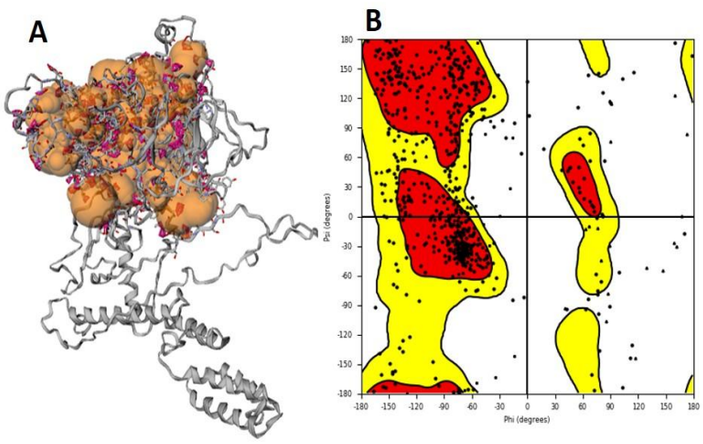

Figure 1: (A) I-TASSER was used to create the protein's 3D structure as well as its anticipated active site (B).

This structure's Ramachandran plot indicates various areas of the protein's shape. Following loop optimization, minimization, and concomitant Ramachandran plot analysis, figure (1B) showed the protein structure. About 90% of the observations were positive, giving the structure a satisfactory overall quality score. All residues are pictured as circles in the story, but proline and glycine are shown as squares and triangles, respectively. 'Favoured' regions are marked in orange, 'allowed' regions in yellow, and 'disallowed' regions in white. From the ZINC database, we created a chemical library with 1615 different compounds. Docking simulations of these ZINC compounds was performed with target protein using PyRx with AutoDock Vina. The top 5 compounds identified by this docking are listed according to their binding affinities in table 1. Two compounds, ZINC000012503187 and ZINC000096006020, emerged as the most active among all evaluated compounds for our protein target as a result of this lead discovery. The most active compound's 2D interactions were shown in figure 2. Tables 1 provide the specific features of these top compounds. Then, we ran molecular dynamics (MD) simulations for 100 nanoseconds of the protein targets in association with these interesting compounds, and we analyzed the simulation paths thereafter. The investigation of the MD trajectory included a number of data points, including protein-ligand interactions, RMSD, and RMSF.

ZINC000096006020's Root Mean Square Deviation (RMSD) of the carbon alpha atoms are displayed versus time. The right Y-axis showed how ligand RMSD changes over time, whereas the left Y-axis showed how protein RMSD varies with time. (Ligand RMSD is depicted in light pink, whereas protein RMSD is shown in dark pink.) B) RMSD of carbon alpha atoms of protein with ligand (ZINC000012503187) over time. Left Y-axis displays the variation of protein RMSD, and right Y-axis demonstrates the variation of ligand RMSD over time. (Dark pink colour showed protein RMSD and Light pink color showed ligand RMSD).

ZINC ID	Binding Affinity (kcal/mol)	Mwt	LogP	Molecular Formula
ZINC000096006020	-11.5	853.918	3.736	C47H51NO14
ZINC000012503187	-11.4	498.586	6.507	C32H26N4O2
ZINC000052955754	-11.1	581.673	1.991	C33H35N5O5
ZINC000004099008	-10.6	656.662	2.753	C32H32O13S
ZINC000003932831	-10.5	528.537	6.576	C27H30F6N2O2

Table 1: Displaying the top chemicals' binding affinities

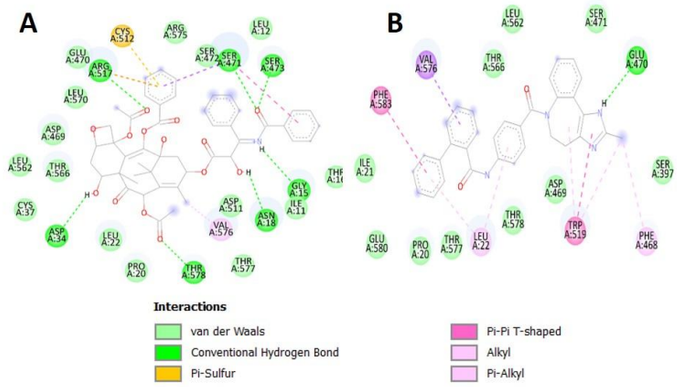

Figure 2: Interactions of ZINC000096006020 (A) and ZINC000012503187 (B) with protein target showing interacting residues and type of interactions with their distances.

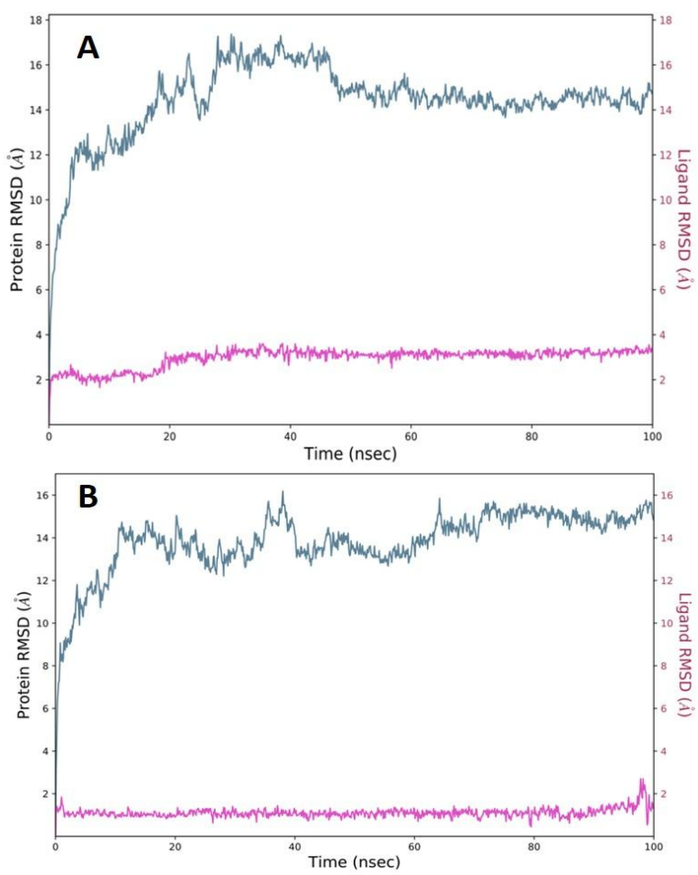

Figure 3: A) The target proteins' and the ligand ZINC000096006020's Root Mean Square Deviation (RMSD) of the carbon alpha atoms are displayed versus time. The right Y-axis showed how ligand RMSD changes over time, whereas the left Y-axis showed how protein RMSD varies with time. B) RMSD of carbon alpha atoms of protein with ligand (ZINC000012503187) over time. (Ligand RMSD is depicted in light pink, whereas protein RMSD is shown in dark pink.)

Figure (3A) confirmed how the RMSD values of ligand-bound proteins' carbon alpha atoms have evolved over time. The protein molecule in the system (ZINC000096006020-protein complex) achieved stability at 25 ns, in accordance to the RMSD plot. afterwards, RMSD values vary within 2 Angstrom throughout the simulation duration, which is perfectly okay for predicted proteins. Throughout the trial, the structure looked to be steady for the most

part. After reaching balance, the ligand remained constant throughout the experiment. In a few cases, the RMSD numbers varied dramatically. This could be explained by a binding state switch. After reaching balance, the ligand RMSD remained steady for up to 100 ns.

Figure (3B) depicts how the RMSD values of ligand-bound proteins' carbon alpha atoms have evolved over time. The proteins included in the structure (ZINC000012503187-protein complex) achieved stability 20 ns after the simulation began, according to the RMSD curve. Following that, the RMSD values varied within 2.0 Angstrom for the sample duration of 100 ns, after which there was a small rise in RMSD. Protein RMSD structure is barely increasing during simulation time, resulting in minimal system variance and stability. It then became steady again. The ligand reached equilibrium after 10 ns and stayed stable throughout the experiment. The ligand RMSD remained steady throughout the experiment.

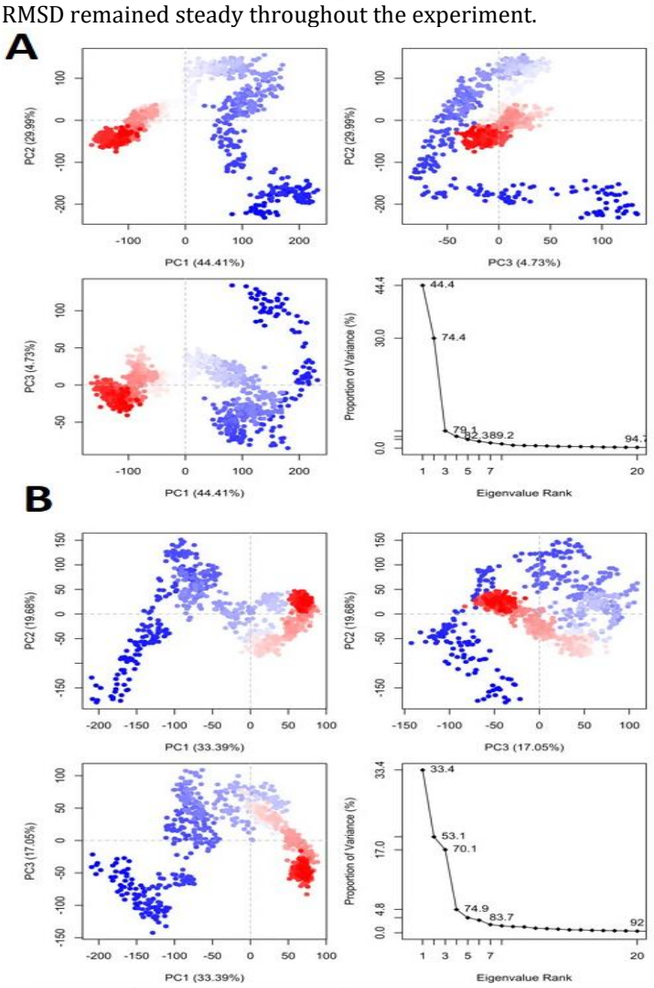

Figure 4: Plotting the principal component analysis eigenvalues versus the variance percentage. A) The sections of the ZINC000096006020-protein complex with the greatest diversity are shown throughout three different segments. For PC1, PC2, and PC3, the cumulative variances are 44.41%, 29.99%, and 4.73%, respectively. B) Variations in PC1, PC2, and PC3 in the ZINC000012503187-protein complex result in cumulative variances of 33.39%, 19.68%, and 17.05%, respectively.

The sections of the ZINC000096006020-protein complex with the greatest diversity are shown throughout three different segments. For PC1, PC2, and PC3, the cumulative variances are 44.41%, 29.99%, and 4.73%, respectively. B) Variations in PC1, PC2, and PC3 in the ZINC000012503187-protein complex result in cumulative variances of 33.39%, 19.68%, and 17.05%, respectively.

The findings in current investigation, protein dynamics were evaluated using Principal Component Analysis (PCA). With the use of this analytical method, we were able to track collective trajectory movements throughout the Molecular Dynamics (MD) simulations, which revealed important information. For the first 20 modes of motion, Figure 4 showed a graph with the eigenvalues (protein) plotted against the eigenvector index (eigenmode) (David and Jacobs, 2014). The variations in hyperspace eigenvectors, which control the overall mobility of the proteins throughout simulations, are captured by these eigenvalues. Notably, compared to the remaining eigenvectors with lower eigenvalues, the top five eigenvectors in our systems had dominating movements, which were characterized by greater eigenvalues ranging from 44.4% to 89.2% and 33.4% to 83.7%. For both the ZINC000096006020-protein and ZINC000012503187-protein complexes, the first three Principal Components (PC1, PC2, and PC3) were displayed in Figure 4 and accounted for more than 50% of the observed changes. Particularly, PC1 clusters showed the largest variability at 44.4% and 33.3%, followed by PC2 at 29.9% and 19.6%, and PC3 at 4.7% and 17.0%, respectively. Compared to PC1 and PC2, PC3 has less variability and a more compact structure, which suggests that the interaction between the protein and ligand is more stable. All of the groups in the PC subspace's conformational changes were classified, with blue denoting the most mobility, white denoting moderate mobility, and red denoting less flexibility.

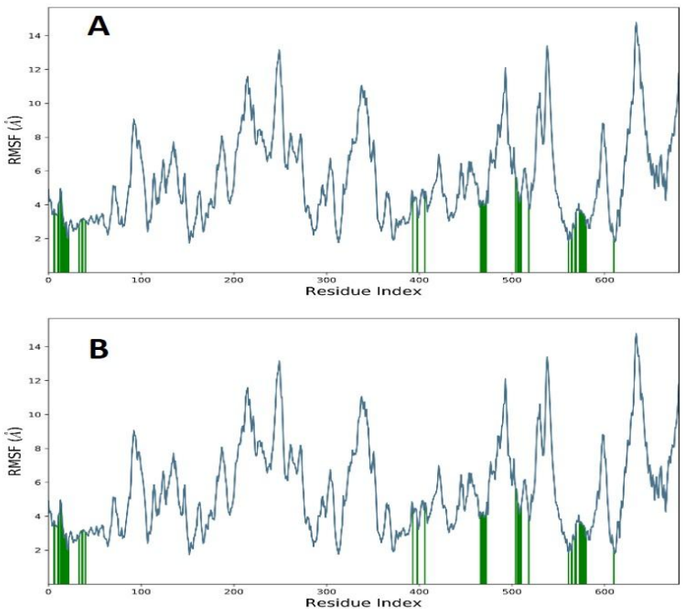

Figure 5: Residue wise RMSF of protein complexed with ligand (A: ZINC000096006020-protein, B: ZINC000012503187-protein).

The measurements of the protein-ligand complexes' Root Mean Square Fluctuation (RMSF) were shown in figure 5. The loop, N-terminal, and C-terminal portions of the proteins are represented by the top peaks in figure 5 according to our analysis of the Molecular Dynamics (MD) simulations. Lower RMSF values found for residues close to the binding site suggest a secure bond between the ligand and protein.

We discovered helices and strands as secondary structural elements (SSEs) throughout the modelling procedure. The distribution of these secondary structural elements among residues during the course of the simulation is shown in the graph below. Alpha helices and beta strands were discovered to make up 9.47% and 2.13%, respectively, of the ZINC000096006020-protein complex, giving it a total secondary structural component fraction of 11.60%. Similar percentages of helices and strands were 8.47% and 3.77% respectively in the ZINC000012503187-protein complex, which added up to a total secondary structural component percentage of 12.26%.

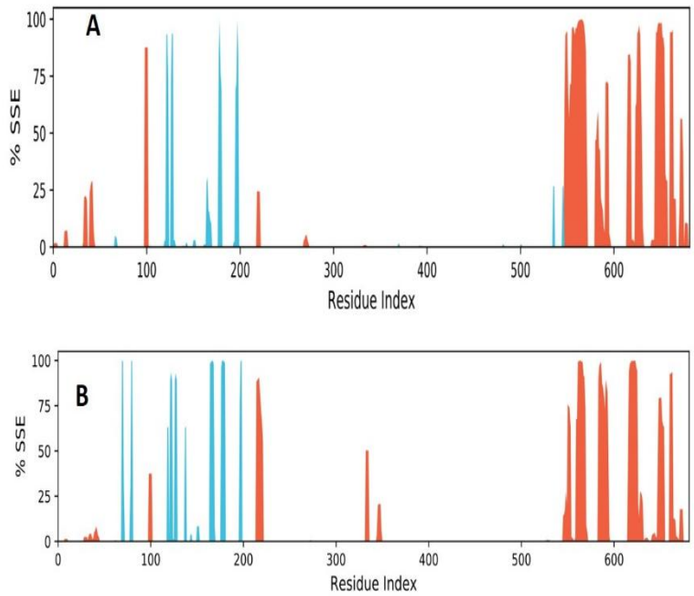

Figure 6: Protein Secondary Structure elements i.e., helix and strands residue wise distribution all over the protein arrangements

complexes with the ligand molecules. Red columns represent alpha helices while blue columns show beta-strands (A: ZINC000096006020-protein, B: ZINC000012503187-protein).

Figure 6 demonstrated that hydrogen bonds are the most important receptor-ligand interactions identified by MD. ASN_18, ILE_21, GLU_507, THR_566, and THR_577 are the most significant hydrogen bonding sites for the ZINC000096006020-protein complex. As Hydrogen bonds for ZINC000012503187-*protein*, the most significant are ILE_21, ASP_469, and THR_578. During the experiment, a protein with a ligand was detected. The histogram depicts the overall number of protein-ligand interactions/contacts (i.e., hydrogen bonds, ionic, hydrophobic, and water bridges) over time. In terms of RMSD stability, PCA, DCCM, and number of contacts throughout the simulation, ZINC000096006020 drug demonstrated greater stability with protein target than ZINC000012503187 drug.

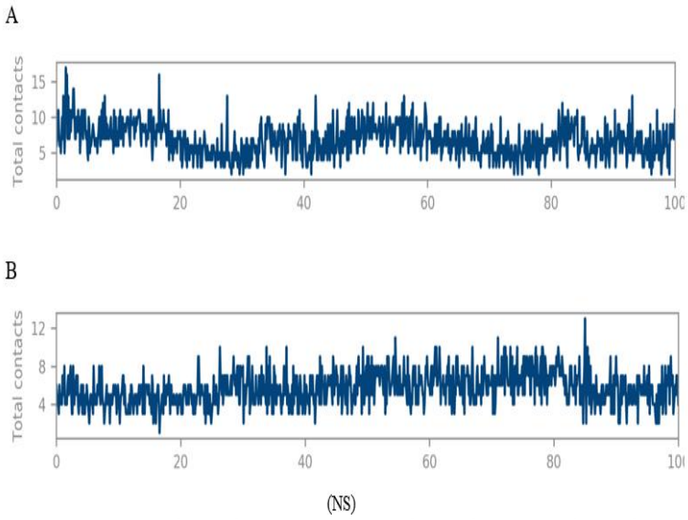

Figure 7. Protein ligand contacts throughout the time of simulation. (A: ZINC000096006020-protein, B: ZINC000012503187-protein)

CONCLUSION: Irrespective of the method used, there is critical co-existence of underweight and obesity in children living in the rural low socioeconomic population of North-West Pakistan. The current study hypothesize that cultural tendencies apart from other factors were contributing to the higher prevalence and critical co-existence of underweight and obesity particularly in the girls. The disparity in our results, between CDC and IOTF cut-offs in comparison to other populations, suggest the influence of different socioeconomics, cultural and genetic factors.

FUNDING: This research was conducted without any funding.

CONFLICT OF INTEREST: All the authors declared no conflict of interest.

LIFE SCIENCE REPORTING: In current research article, no life science threat was reported.

ETHICAL RESPONSIBILITY: This is original research, and not submitted in whole or in parts to another journal for publication purpose.

INFORMED CONSENT: The author(s) have reviewed the entire manuscript and approved the final version before submission.

REFERENCES:

Abir, M. H., T. Rahman, A. Das, S. N. Etu, I. H. Nafiz, A. Rakib, S. Mitra, T. B. Emran, K. Dhama, A. Islam, A. Siyadatpanah, S. Mahmud, B. Kim and M. M. Hassan, 2022. Pathogenicity and virulence of marburg virus. Virulence, 13(1): 609-633.

Ashburn, T. T. and K. B. Thor, 2004. Drug repositioning: Identifying and developing new uses for existing drugs. Nature reviews drug discovery, 3(8): 673-683.

Aziz, S., W. N. Ain, R. Majeed, M. A. Khan, I. Qayum, I. Ahmed and K. Hosain, 2012. Growth centile charts (*Anthropometric measurement*) of pakistani pediatric population. JPMA-Journal of the Pakistan medical association, 62(4): 367.

Bausch, D. G. and L. Schwarz, 2014. Outbreak of ebola virus disease in guinea: Where ecology meets economy. PLoS tropical disease, 8(7): e3056.

Bowers, K. J., D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan and D. E. Shaw, 2006. Scalable algorithms for molecular dynamics simulations on commodity clusters. In: SC '06: Proceedings of the 2006 ACM/IEEE Conference on supercomputing. pp: 43-43.

Callahan, A., K. D. Anderson, M. S. Beattie, J. L. Bixby, A. R. Ferguson, K. Fouad, L. B. Jakeman, J. L. Nielson, P. G. Popovich, J. M. Schwab, V. P. Lemmon and F. S. W. Participants, 2017. Developing a data sharing community for spinal cord injury research. Exp Neurol, 295: 135-143.

Clark, D. V., P. B. Jahrling and J. V. Lawler, 2012. Clinical management of filovirus-infected patients. Viruses, 4(9): 1668-1686.

Dallakyan, S. and A. J. Olson, 2015. Small-molecule library screening by docking with pyrx. Methods of molecular biology, 1263: 243-250.

Gatherer, D., 2014. The 2014 ebola virus disease outbreak in west africa. Journal of general virology, 95(Pt 8): 1619-1624.

Grobler, J. A., A. S. Anderson, P. Fernandes, M. S. Diamond, C. M. Colvis, J. P. Menetski, R. M. Alvarez, J. A. T. Young and K. L. Carter, 2020. Accelerated preclinical paths to support rapid development of covid-19 therapeutics. Cell host microbe, 28(5): 638-645.

Kortepeter, M. G., K. Dierberg, E. S. Shenoy, T. J. Cieslak, T. Medical Countermeasures Working Group of the National Ebola and N. Education Center's Special Pathogens Research, 2020. Marburg virus disease: A summary for clinicians. International journal of infectious disease, 99: 233-242.

Lin, X., X. Li and X. Lin, 2020. A review on applications of computational methods in drug screening and design. Molecules, 25(6).

Malvy, D., A. K. McElroy, H. de Clerck, S. Gunther and J. van Griensven, 2019. Ebola virus disease. Lancet, 393(10174): 936-948.

Manohar, M. P., V. J. Lee, E. U. Chinedum Odunukwe, P. K. Singh, B. S. Mpofu and C. Oxley Md, 2023. Advancements in marburg (marv) virus vaccine research with its recent reemergence in equatorial guinea and tanzania: A scoping review. Cureus, 15(7): e42014.

Morrissey, J. H., V. Pureza, R. L. Davis-Harrison, S. G. Sligar, Y. Z. Ohkubo and E. Tajkhorshid, 2008. Blood clotting reactions on nanoscale phospholipid bilayers. Thromb Res, 122 Suppl 1(Suppl 1): S23-26.

Iqbal, M., S. Vinoth, V. Sasikanth and T. Rathinavel, 2023. Molecular insights of marine algal polycyclic aromatic compounds as promising anti-viral agents for targeting sars-cov-2 proteins – an in silico validation. Polycyclic aromatic compounds: 1-24.

Pushpakom, S., F. Iorio, P. A. Eyers, K. J. Escott, S. Hopper, A. Wells, A. Doig, T. Guilliams, J. Latimer, C. McNamee, A. Norris, P. Sanseau, D. Cavalla and M. Pirmohamed, 2019. Drug repurposing: Progress, challenges and recommendations. Natural review drug discovery, 18(1): 41-58.

Qi, Y., H. I. Ingolfsson, X. Cheng, J. Lee, S. J. Marrink and W. Im, 2015. Charmm-gui martini maker for coarse-grained simulations with the martini force field. Journal of chemistry theory comput, 11(9): 4486-4494.

Shaw, D. E., J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. 2014. Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In: SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. pp: 41-53.

Trott, O. and A. J. Olson, 2010. Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computional chemistry, 31(2): 455-461.

Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang, 2015. The i-tasser suite: Protein structure and function prediction. Natural Methods, 12(1): 7-8.