Journal of Computational Applied Mechanics 2024, 55(3): 401-422
DOI: 10.22059/JCAMECH.2023.364997.873

RESEARCH PAPER

Comparative in silico analysis of CHIR99021, Azakenpaullone and Tricantin interactions with GSK3β, a key protein in stem cell fates

Javad Kazemi [a,b], Atefeh Alipour [c], Keyvan Shahryarimorad [b], Eisa Tahmasbpour Marzouni [b,d], Zahra Azadian [b], Ali Ehsani [e], Hosein Shahsavarani [a*]

[a] *Department of Cell and Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran*
[b] *Laboratory of Regenerative Medicine and Biomedical Innovation, National Cell Bank, Pasteur Institute of Iran, Tehran, Iran*
[c] *Department of Nanobiotechnology, Pasteur Institute of Iran, Tehran, Iran*
[d] *St Vincent's Hospital, Sydney, Australia*
[e] *Department of Bioscience, University of Milan, Milan, Italy*

Abstract

Glycogen Synthase Kinase 3β (GSK3β) is a multifunctional serine/threonine-protein kinase that serves as a pivotal regulator of various human pluripotent stem cell (hPSCs) functions, including self-renewal, adhesion, survival, and differentiation in addition to have an effect on motility of sperm. Despite advancement in understanding the critical roles of GSK3β inhibition in various stem cell functions, the exact molecular basis of its inactivation using various small-molecule inhibitors remains poorly understood. Investigating the mechanistic details of the actions of inhibitors targeting GSK3 proteins, such as CHIR99021, Azakenpaullone, and Tricantin, could be extremely beneficial for improving novel defined stem cell culture systems and cancer research. The present study aimed to predict the binding mode of the aforementioned ligands with GSK3β, by molecular docking and metadynamic simulation, and compare the three-dimensional structure of the inactive conformation of GSK3β in the presence of three inhibitors. Also, the pharmacokinetic or ADMET properties of ligands, such as Lipinski's rule of five violations for drug-likeness, QPlog S, QPlog K, and bioactivity scoring, were predicted. The analysis of protein stability revealed that in the absence of inhibitors, the GSK3β has higher flexibility, while in the presence of CHIR and AZA, the rate of flexibility of most protein regions, especially the envelope area, decreased. It was found that though all small molecules are capable of facilitating the inhibition of GSK3β protein, but the flexibility of protein is a bit higher for CHIR than those for other two ligands.

Keywords: Human pluripotent stem cells (PSCs); GSK3β; CHIR99021; Azakenpaullone; molecular dynamics simulation.

1. Introduction

The Glycogen Synthase Kinase 3β (GSK3β) is a multifunctional protein kinase that implicates as a master

* Corresponding author.
E-mail address: hosein.shahsavarani@gmail.com

regulator in both self-renewal and differentiation of pluripotent stem cells (PSCs). Moreover, it has been ubiquitously linked to several prevalent diseases including cancer, diabetes and inflammation due to its critical role in different physiological functions such as gene expression, cells differentiation, proliferation, metabolism and cell survival [1, 2]. One of the most important functions of GSK3β is related to the WNT signalling pathway. In the presence of the WNT ligand, it forms a binding complex with its receptor Fzd and its coreceptor lipoprotein-related protein 5 and 6 (LRP-5/6) situated on the surface of the target cell. This interaction subsequently hinders the phosphorylation of β-catenin, a process mediated through the dishevelled (Dvl) protein. β-catenin can interact with the transcription factor TCF-LEF and induce transcription of the target gene. In the resting state, GSK3β and casein kinase 1 (CK1) phosphorylate β-catenin, which causes β-catenin to break down and become unstable. Therefore, the pharmacological inhibition of GSK3β leads to the stabilization and activation of β-catenin, which reflects the activity of the WNT signalling pathway. While activation of WNT signalling and subsequent GSK3β activation induces stem cells self-renewal activity and pushes stem cells towards a quiescent state, its inhibition triggers stem cells differentiation [3]. The WNT signalling pathway holds crucial significance in guiding the differentiation of PSCs into diverse cell lineages [4]. In the early stage of differentiation, the activation of the conventional wnt signalling pathway by the recombinant Wnt3a protein leads to a significant increase in the optimal generation of cardiomyocytes from pluripotent stem cells. However, it seems that the use of recombinant Wnt3a protein in cell therapy and regenerative medicine is facing some problems. One of the features that makes the recombinant Wnt3a protein have an appropriate function is the post-translational modification by palmitoylation, which greatly increases its hydrophobicity, as a result, the production of this protein on a large scale is complicated, and therefore its use for high volume cells production in the *in vitro* aqueous environment is not economical [5]. To address this issue, the regulation of this signalling pathway with the help of the small molecules plays an important role in the *in vitro* manual differentiation [6]. Small molecules offer a significant advantage over native proteins in modulating the WNT signalling pathway, surpassing regulators like DKK1 and Noggin [7]. The small molecules in simple diffusion can penetrate across the cell membrane and reach different parts of the cell while they are more cost-effective than recombinant proteins [8]. While a range of small molecules have been identified as GSK-3β inhibitors for using in the regenerative medicine applications, a significant hurdle remains in determining their specific selectivity for GSK-3β without causing interference with other associated signalling pathways [9].

The activation of the WNT signalling pathway via small molecule of CHIR99021 leads to the differentiation of PSCs into the mesoderm lineage. [10, 11] CHIR99021 increases the activity of the WNT pathway by inhibiting the GSK3β as one of the protein kinases of the WNT signalling pathway, thereby driving PSCs into the mesoderm lineage. [12, 13]The main effective factor in WNT signalling is β-catenin. Disruption of the WNT pathway leads to the degradation of β-catenin, resulting in a decrease in its cytosolic concentration [14]. Upon activation of WNT signalling , β-catenin accumulates in the cytosol and is transported to the nucleus, where it interacts with the specific transcription factors. [15, 16]

Molecular Dynamics (MD) simulation of biomolecules is often classified in the field of computational chemistry, but the scientific origins of this technique trace back to polymer chemistry and structural biology. In the 1970s, when the study of physics included molecular spatial properties, flexibility, distortion, stability and loosening of structures, as well as the primary structures of X-ray proteins were used on a short-term scale. In general, molecular simulation was already a harbinger in physics [17]. The field of molecular simulation has advanced greatly since then, and now the simulations are performed repeatedly on a scale of a few milliseconds, during which the protein can be folded several times [18-21], predict the interaction between receptors [22] and functional properties of receptors, and even record transition states between proteins such as membrane proteins[23]. Though long-term classical simulations remain valuable approach for directly observation of intricate molecular processes that are challenging to detect, contemporary research is progressively turning towards extensive sets of simulations. This shift is partly facilitated by the growing availability of structural models derived from sequencing and structural genomics. Further, new techniques for estimating complex molecular properties have made it possible by using thousands of shorter simulations [24]. Nowadays, modelling gene mutations enable us to execute of brief simulations on hundreds of mutated genes. Classical molecular dynamics simulations, rooted in empirical models, hold a crucial position in this process. This is largely because many properties are determined by free energy, a factor that typically demands comprehensive sampling. In contrast, conventional quantum chemistry approaches face limitations when applied to large-scale systems [25]. These advances would not have been possible without significant research efforts in simulation, optimization, and parallelization algorithms. The emergence of standard molecular modelling software packages such as CHARMM [26], GROMOS [27], Amber [28], NAMD [29], and GROMACS[30] has been important because of these applications in molecular simulation and modelling research. This advancement has equipped life scientists, who may not possess specialized expertise in molecular simulations, with accessible computational tools and techniques. The GROMACS tool stands as a prominent objective in the ongoing pursuit of achieving the utmost simulation capabilities within research laboratories. Notably, GROMACS 4 has been in continuous development since

2009, witnessing the incorporation of a host of new features. This culminated in the release of version 4.5 of the software, marked by a substantial enhancement in the performance of parallel applications [31]. The main purpose of the present study was to investigate and predict the physicochemical properties of CHIR99021 (CHIR), Azakenpaullone (AZA), and Tricantin (TRI) as activators of the WNT signalling pathway through inhibition of GSK3β protein in order to screen the appropriate cocktail for differentiating pluripotent stem cells (PSCs) using molecular dynamics (MD) simulation. Moreover, evaluating several pharmacodynamic or ADMET in silico modelling for inhibitor were performed using *GROMACS 2021.1* software and *QikProp* tool version 4.4 of *Maestro* software package (*schrödinger* suite, LLC, New York, NY, Release 2015-2), respectively. [32] In addition, molecular docking analysis was performed to identify the binding sites of AZA and TRI as well as the proper interaction of these compounds with the essential amino acids of the GSK3β functional site using *AutoDock 4.2* software. Also, the current computational study provided a rational approach to binding CHIR, AZA, and TRI to the GSK3β target protein.

2. Material and methods

2.1. Parametrization of molecular ligand

To parameterize ligand molecule, *AnteChamber* software in *AmberTools* software package was used; finally, the parameters created usingthe *AnteChamber Python Parser interfacE* (ACPYPE) tool were converted to *GROMACS* format [33, 34].

2.2. Structure of GSK3β (protein)

The crystal structure of GSK3β (protein) in the complex state with CHIR, AZA and TRI with code 5HLNwas obtained from RSCB protein database (PDB) (www.rcsb.org), respectively. The structure studies were performed using *Swiss-PDB viewer 4.10* and *Chimera* software. [35]

2.3. Molecular docking

GSK-AZA and GSK-TRI complexes were created through molecular docking to determine the binding site of these ligands in the protein. *Autodock* software version 4.2 was utilized to perform molecular docking [36]. This software uses many physical features for flexible docking Another point about this software is that for docking, we should have a three-dimensional structure of ligand and protein [37]. This software includes two main software, *AutoDock* and *AutoGrid*. The *AutoGrid* program is actually a set of pre-calculated networks for the target protein. The *AutoDock* program is also responsible for docking protein ligands [38]. *AutoDock* 4.2 software has a graphical section called *AutoDockTools* which allows the user to edit the ligand and protein before docking as well as view the output from the docking [39]. The protein file must be prepared before docking. In this study, the GSK3β protein structure with pdb code 1HEL was downloaded from this website and used directly as *AutoDock* input. Then, polar hydrogen atoms were added on the structure, kollman partial charge was added to the protein atoms and Gasteiger partial charge was added to the ligand atoms. Also, rigid and flexible regions on the protein were selected and stored as two additional files. In the present study, since the binding site of the ligands on the protein was not known, we placed the entire protein inside the grid box. The final step in docking is to use the *AutoDock* program. The Lamarkian genetic algorithm (LGA), which determines the number of structural scans, was also configured for use at this stage. Eventually, the *AutoDock* program started running and all output conformations were saved in *dlg* formate [40]. *AutoDock* 4.2 software has various search algorithms to find the connection including i) Monte Carlo Simulated Annealing (SA), ii) Genetic Algorithm (GA), and iii) Lamarkian Genetic Algorithm (LGA) [30-32]. Among these three algorithms, LGA is the most powerful, which was used in this study as mentioned above. After docking, the product file (*dpf*) was opened and analyzed using *AutoDockTools* software. The conformations inside this file were ranked from best to worst based on the amount of connection energy. Also, using the clustering system, the binding site of the ligand to the protein can be examined.

2.4. Molecular dynamics (MD) simulation

Protein alone and protein-ligand complex were used as inputs to simulate molecular dynamics. Molecular dynamics (MD) simulations were performed using GROningen Machine for Chemical Simulation *(GROMACS 2021.1*) software. Input structures were prepared with ff99SB + ILDN force field. The correct position of hydrogen for all charged amino acids in the protein structure was determined using *propka* software. The surface charge of structure was then neutralized by adding seven chlorine ions. The protein was placed in a layer of Transferable Intermolecular Potential 3 Point (TIP3P) water molecules with 8 Å thick in a cubic box using gmx solvate software.

Energy minimization on the structure was done with 10000 steps by steepest Dwscent method in order to eliminate

steric clashes and inappropriate geometry Long-range electrostatic interactions were calculated using the Particle-Mesh Ewald (PME) algorithm [41]. As well as, lincs algorithm was used to constrain the covalent bonds between heavy atoms and hydrogen atoms. The time step is 2 femtoseconds (fs). The simulation results obtained from *GROMACS2021.1* software were analysed as root mean square deviations (RMSD), root mean square fluctuation (RMSF), Radius of Gyration (RG), and secondary structure using gmx rms, gmx rmsf, gmx gyration modules. *Gmxcovar*, *gmx anaeig* and *gmx sham* modules were used for Principle Component Analysis (PCA) and Gibbs free energy analysis, respectively.

Calculating RMSD among obtained structures and a reference structure is the most important method to evaluate the stability of MD simulations over time. To study the main motions in system, RMSD was obtained for α-carbon atoms. Moreover, structure flexibility was calculated based on the RMSF of α-carbon atoms. The radius of gyration which is one of the determining parameters in structure density was obtained during simulation period.

2.5. ADMET-related properties analysis

ADMET (absorption, metabolism, distribution, excretion and toxicity) properties [42, 43] of CHIR99021 were analyzed using the default parameters of *QikProp* protocol of *Maestro suite* (*Schrödinger*, LLC, New York, NY, Release 2015-2)[44]. *In silico* physico-chemical parameters were predicted by *QikProp* module of *Schrödinger* software including Molecular Weight (MV), total Solvent Accessible Surface Area (SASA), number of hydrogen bond acceptor, number of hydrogen bond donor, number of rotatable bonds (Rot), Molecular Volume ($Å^3$), vdW Polar SA (PSA), globularity, % human oral absorption (%HOA), aqueous solubility (log S, where S in mol dm^{-3} is the concentration of the solute in saturated solution), HERG K+ Channel Blockage: log IC_{50}value for blockage of HERG K^+ channels, and log K hsa Serum Protein Binding, Five violations of Lipinski's rule including molecular weight < 500, QPlog Po/w < 5, donor hydrogen bond ≤ 5, acceptor hydrogen bond ≤ 10 [45] were taken into account; compounds that satisfy these rules are considered drug-like. Three violations of Jorgensen's rule including QPlogS > -5.7, QPlog PCaco > 22 nm/s, number of primary metabolites < 7 were also considered. Compounds with fewer violation of three rules (and preferably no) are more likely to be orally available. At the level of organism, lipophilicity determines the pharmacokinetics and pharmacodynamics properties of the chemical compounds and indicades significant correlation with n-octanol/water (QPlog Po/w) partition coefficient. This physicochemical property including permeability through biomembranes, toxicological profile, solubility and metabolism of chemical compounds.

3. Results and discussion

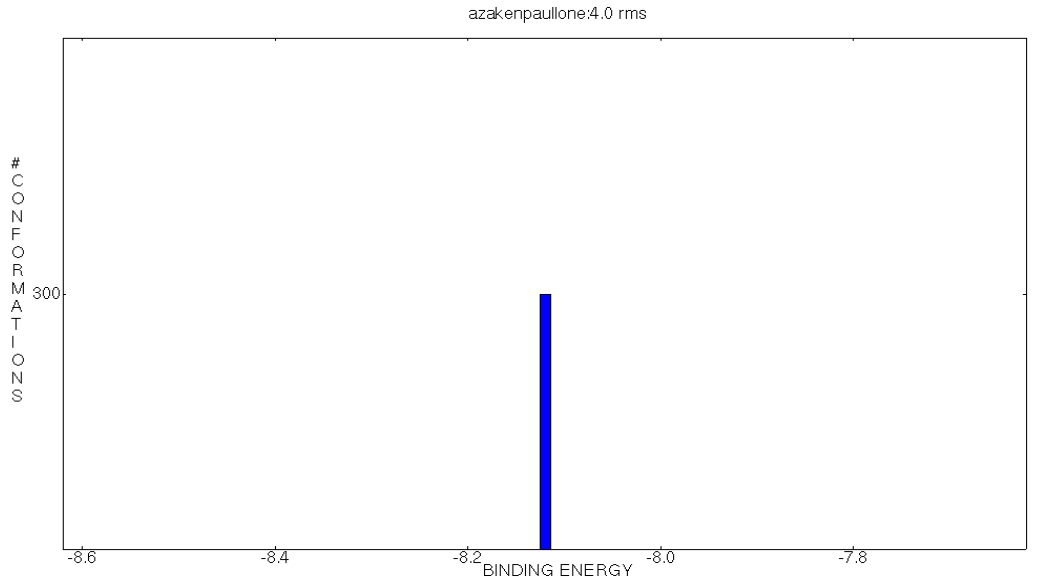
3.1. Molecular Docking Studies

In the present study, AutoDock4 was used to investigate the binding of AZA and TRI ligands to the GSK3β protein. Protein PDB and ligand structure were obtained from RCSB and PubChem databases, respectively. Then, using *AutoDockTools* software, polar hydrogens were added and Kollman charges were placed on all atoms with rotatable bonds calculated. The grid box was selected with a distance of 0.375 Å and the box was created with dimensions of 44 × 54 × 56. The total protein was placed completely inside the box. Figure 1 displays the number of clusters, the number of structures in each cluster, and the best binding energy for each cluster for both compounds.

As presented in Figure 1, the horizontal axis represents the connection energy of each cluster and the vertical axis reflects the number of structures in each cluster. The blue column also belongs to each cluster and the column height shows the number of structures within each cluster. According to Figure 1A, all conformations are located inside a cluster. To choose the best connection mode (i.e.to choose the best cluster), two factors, connection energy and cluster size, are important. The binding energy indicates the strength of ligand binding to the protein and reveals the cluster size, frequency, and probability of ligand binding to different sites on the protein. Table 1 reports various energy factors including electrostatic energy, van der Waals energy, inhibition constant, etc. for each cluster in detail. According to Figure 1B, the structures resulting from molecular docking are divided into 26 clusters. Since cluster 1 has the highest number of structures and the strongest binding energy, thus this cluster was selected to perform molecular dynamics simulation (cluster 1 is shown in red). Details of the first cluster of GSK:AZA and GSK:TRI Complexes are outlined in Table 1. As presented in Table 1, all molecular docking conformations are contained within a cluster. The binding energy of AZA and TRI ligands to proteins is -8.12 and -8.8 kcal / mol, respectively. Also, the efficiency of AZA and TRI ligands is -0.41 and -0.3 kcal / mol, respectively. Ligand efficiency is obtained through dividing the total bond energy by the total number of ligand atoms. Also, the concentration required for enzyme inhibition for AZA and TRI is 1.11 and 351.52 μM, respectively. Also, van der Waals and electrostatic energies for AZA are 8.13 and 0.01 kcal / mol, respectively, and for TRI -12.35 and 0.03 kcal / mol respectively, suggesting that van der Waals energy plays a very important role in Binds ligands to proteins. Figure 2 illustrates how the AZA and TRI ligands bind to the protein in three dimensions. Figure 3 reveals the amino acids involved in interaction with ligand molecules at the binding site. As shown in Figure 3, AZA with amino acids Val70, Ala83, Lys85, Val110,

Leu132, Asp133, Tyr134, Cys199, and Asp200 as well as TRI with amino acids Ile62, Phe67, Val70, Ala83, Lys85, Met101, Val110, Leu132, Tyr134, Val135, Thr138, Glu185, Leu188, and Cys199 establish most hydrogen, van der Waals, and Pi bond interactions at the GSK3β functional site.

A



B

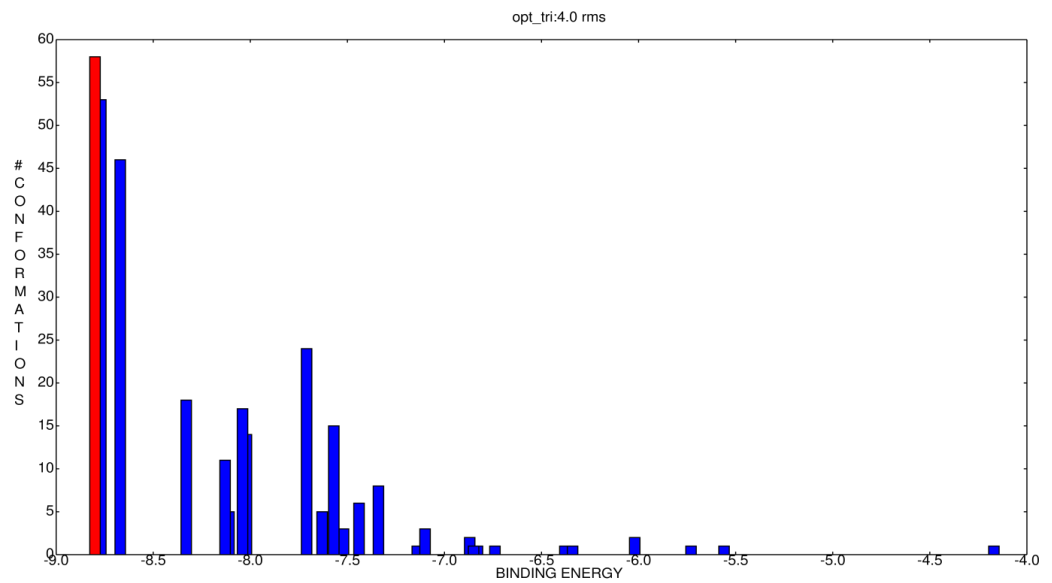


Figure 1. The number of clusters, the best binding energy for each cluster and number of structures in each cluster are shown for (A) Azakenpaullone and (B) Tricantin compounds.

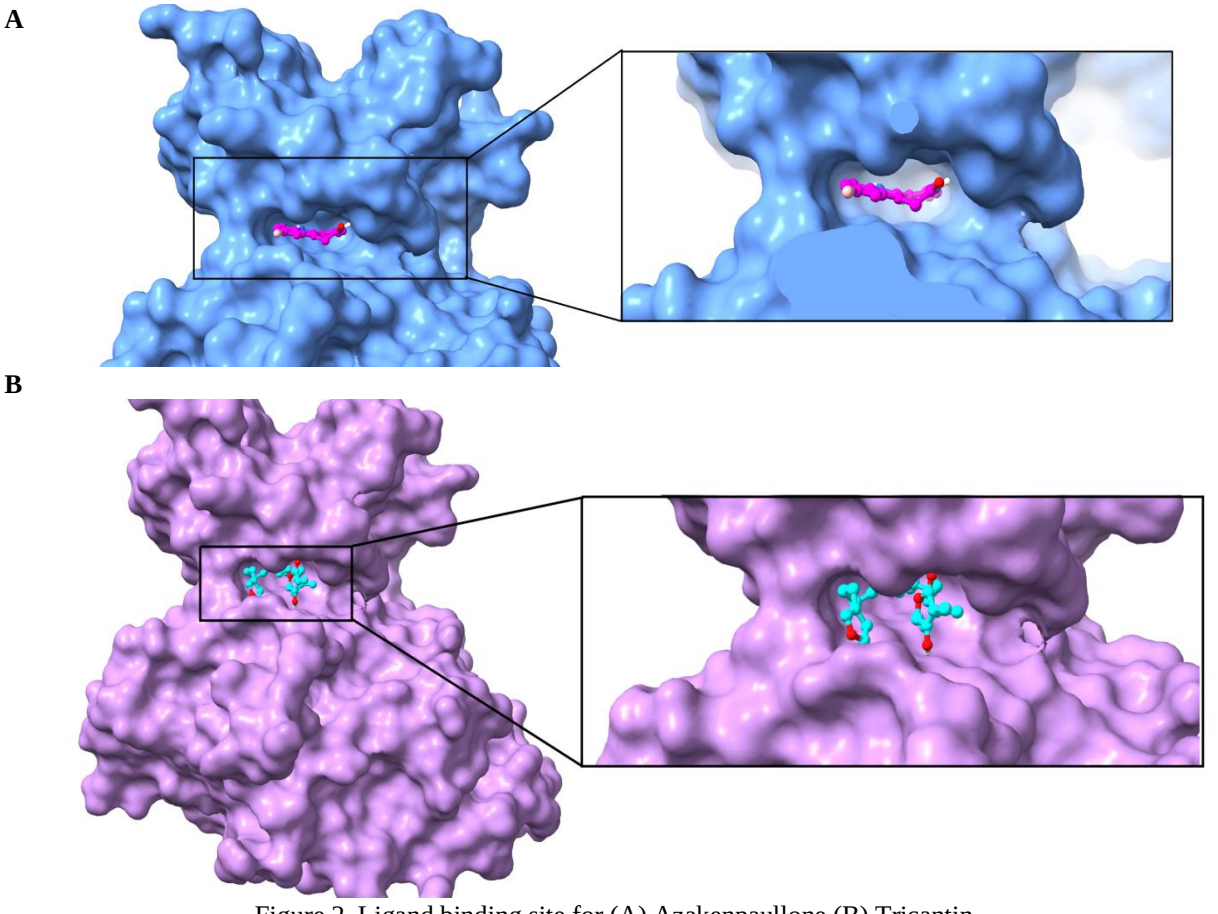
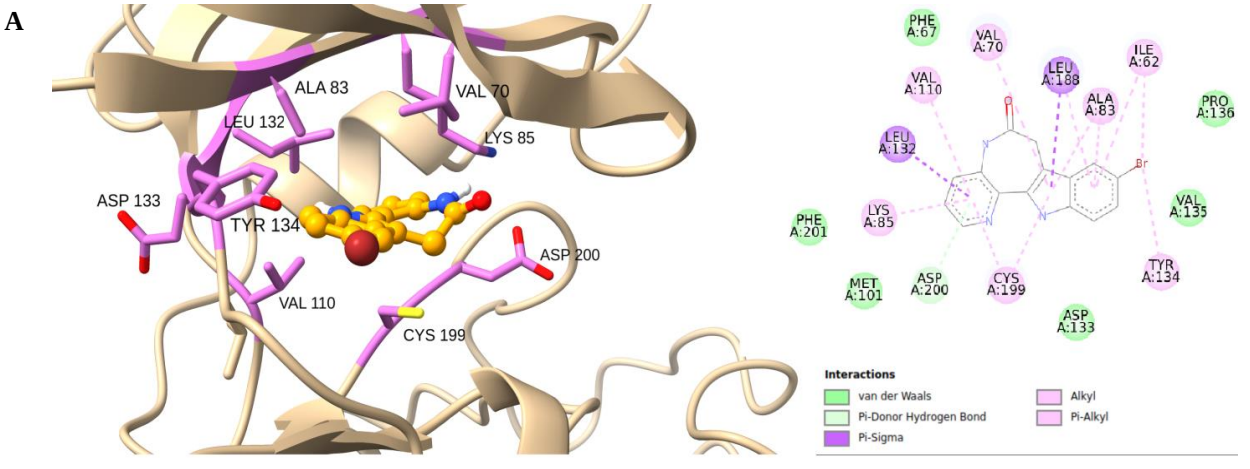


Figure 2. Ligand binding site for (A) Azakenpaullone (B) Tricantin



B

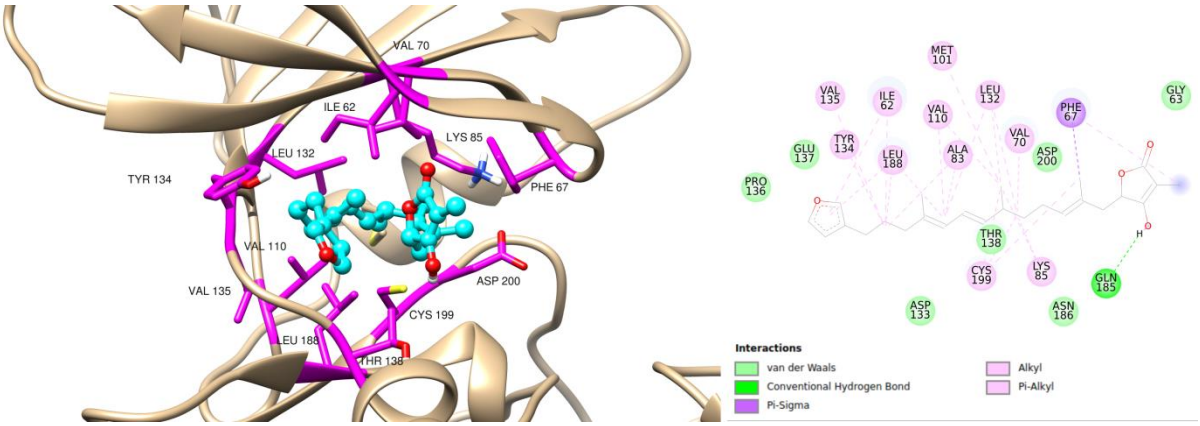


Figure 3- The Amino acids involved in the interaction between GSK3β and (A) Azakenpaullone (B) Tricantin

Table 1. The results of molecular docking related to the first cluster of 2 and 3 complexes

	No. of structures	Binding Energy	Ligand Efficiency	Ki(μm)	Intermol Energy	VdW Energy	Elec. Energy
Cluster 1 for GSK:CHIR omplex	300	-8.12	-0.41	1.11	-8.12	-8.13	0.01
Cluster 1 for GSK:AZA complex	58	-8.8	-0.3	351.52	-12.38	-12.35	-0.03

3.2. Molecular dynamics (MD) simulations for protein stability study

The simulation results obtained from *GROMACS2021.1* software were analyzed as root mean square deviations (RMSD), root mean square fluctuation (RMSF), Radius of Gyration (RG), and secondary structure using gmx rms, gmx rmsf, gmx gyration modules. Gmx covar, gmx anaeig and gmx sham modules were used for Principle Component Analysis (PCA) and Gibbs free energy analysis, respectively. Calculating RMSD among obtained structures and a reference structure is the most important method to evaluate the stability of MD simulations over time. To study the main motions in system, RMSD was obtained for α-carbon atoms. Moreover, structure flexibility was calculated based on the RMSF of α-carbon atoms. The radius of gyration which is one of the determining parameters in structure density was obtained during simulation period.

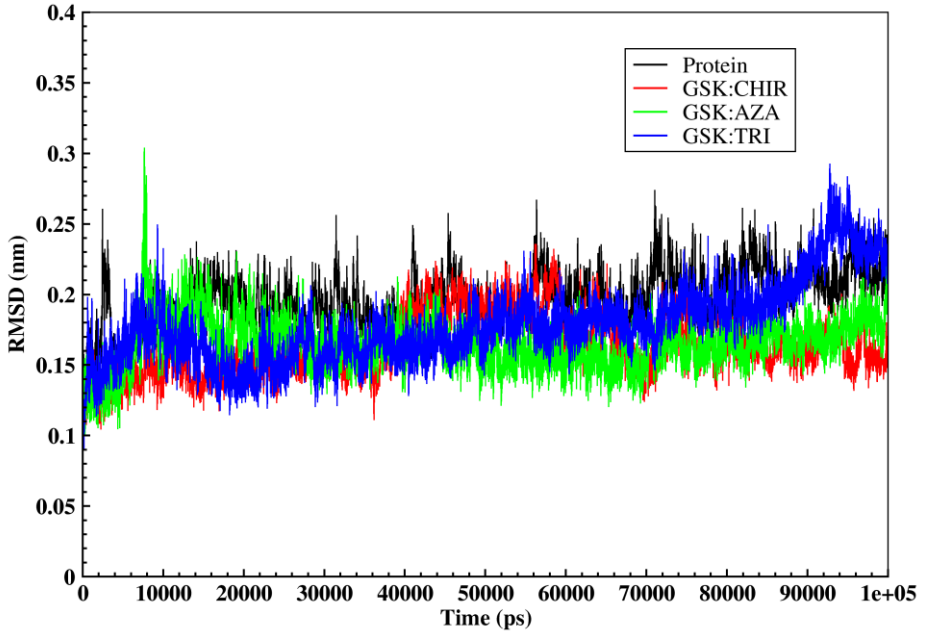
3.3. Root Mean Square Deviation (RMSD)

Root Means Square Deviation (RMSD) among structures created during MD simulation in the time dimension is a suitable and common standard to ensure the structural stability of proteins in different situation (natural and mutant proteins, in the presence of solvent and with different temperatures). The results of this calculation for all simulations are shown in Figure 4A. As shown in Figure 4A, at the beginning of simulation, three graphs show an uptrend. In the GSK:CHIR Complex state diagram, the RMSD value has reached about 0.15 nm at the time of 10,000 ps, and shows relative stability over this time at 38,000 ps. After this time, RMSD value increases sharply and reaches 0.19 nm at 42,000 ps, and remains at this value until 59,000 ps, after which time it starts to decrease again, and at 80,000 ps, it reaches 0.15 nm again and remains constant at this value until the end of the simulation. From these two graphs, it can be concluded that the protein reached equilibrium after 80,000 ps in both simulations (protein and GSK:CHIR complex).The RMSD diagram of the protein in GSK:AZA complex shows that after about 10,000 ps, the RMSD value of the protein reaches about 0.2 nm and remains at that value for up to 11,000 ps. Following the simulation, the RMSD value diminished slightly and reached 0.15 nm in about 70,000 ps. After this time, the RMSD value again shows a very slight increase and at the end of the simulation it reached 0.16 nm.

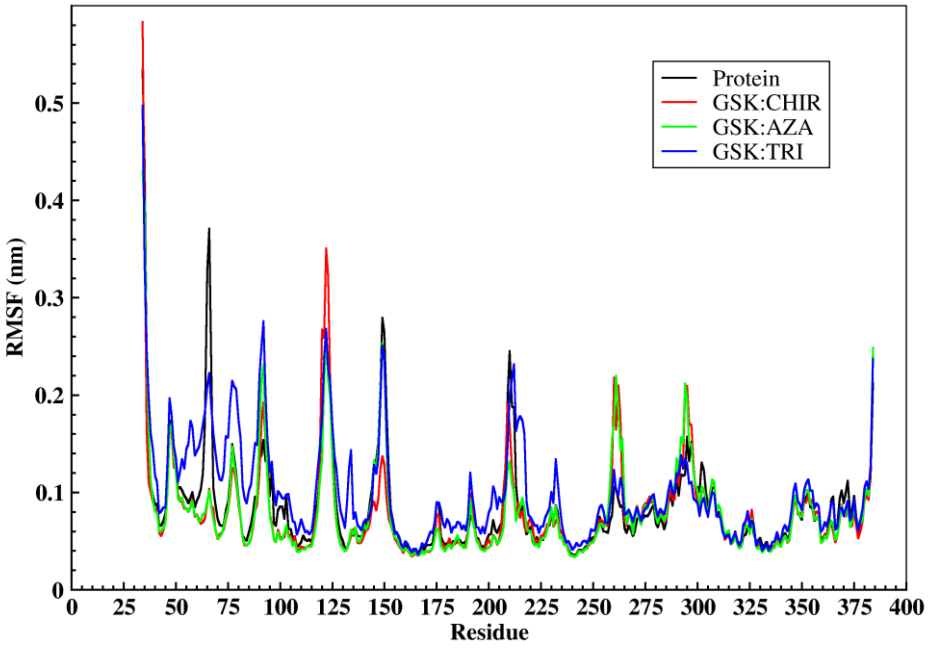
Analysis of the RMSD diagram for GSK:TRI Complex indicates that at the beginning of the simulation, the value

of RMSD protein increased to about 0.2 nm at 10,000 ps. After this time, the RMSD value decreased and the time of 20,000 ps reached about 0.13 Å. The graph then rises again, reaching about 0.25 nm in 91,000 ps, then declining again to about 0.2 nm at the end of the simulation. Comparison of the behaviour of the four graphs shows that in the absence of the protein ligand during the simulation time, it has a high amplitude of oscillation (as the corresponding graph has a high oscillation), while the amplitude of oscillation in the two diagrams of GSK:CHIR and GSK:AZA complexes is very low. Also, the two complexes have reached full stability after 70,000 ps, while in GSK:TRI complex, there is an upward trend to the end of the simulation diagram, with also a high amplitude of oscillation. To examine the mentioned changes, the best factor to study is the flexibility of amino acids in different regions of the protein.

A



B



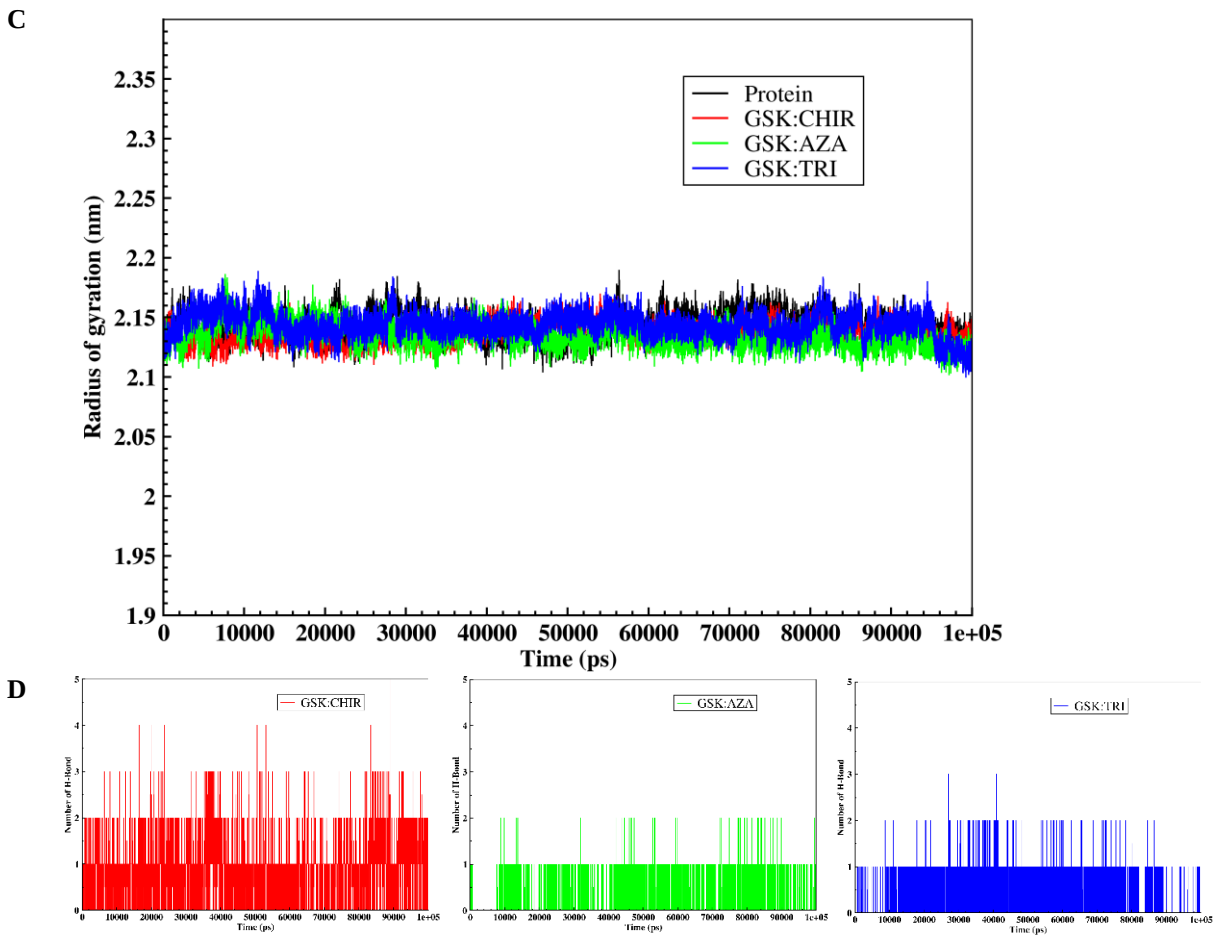


Figure 4. A) RMSD changes of protein in the presence and absence of ligands over molecular dynamics simulation B) Comparison of the amino acids flexibility of protein in the presence and absence of ligands C) Radius of gyration changes of the protein in the presence and absence of ligands D) Number of hydrogen bond between protein and ligands

3.4. Root Mean Square Fluctuation (RMSF)

The dynamic behaviour of α-carbon atoms in the structure contains sufficient information to study important motions in proteins and reflects the general motions of structure. Therefore, Root Mean Square Fluctuation (RMSF) of α-carbon atoms was considered to study the motion and structural flexibility. In this section, comparing structural flexibility between two states of protein in the presence and absence of ligand was performed (Figure 4B). In all four simulations, the value of RMSF was calculated from the stable region (end 20 ns of the simulation). As depicted in Figure 4B, in GSK:CHIR and GSK:AZA complexes, the amount of flexibility in some areas of the protein has been significantly reduced, and only in the amino acid region of Ala248-Pro255, the amount of flexibility of the protein in the two complexes increased relative to the free protein. Areas of protein with diminished flexibility relative to free protein in the presence of CHIR (GSK:CHIR Complex) ligand are: Tyr 60-Val70 and Val 142-Thr 152. In GSK:AZA complex, the degree of protein flexibility is similar to that of GSK:CHIR complex, and only in two the amino acid regions of Val 142-Thr 152 and Val 208-Pro212, the amount of protein flexibility in GSK:AZA Complex is different from that of GSK:CHIR Complex. According to this figure, in the amino acid region of Val 142-Thr 152, the extent of protein flexibility in GSK:AZA complex is similar to that of free protein, while in the amino acid region of Val 208-Pro 212, the amount of protein flexibility in GSK:AZA complex shows a significant drop. Elsewhere, the amount of protein flexibility in all three simulations is similar. To further explore the reason for the decline in the flexibility of these regions in the complex state, these regions are shown in the three-dimensional structure of the protein (Figure 5). As can be seen in Figure 5, the two regions with low flexibility in GSK:CHIR complex are at the ligand binding

site. Indeed, it can be stated that ligand binding to the protein has reduced the flexibility in these regions in GSK:CHIR complex. The GSK:TRI complex study shows that the flexibility of Arg 50-Lys 85, Glu 125-Asp 133, and Tyr 140-Arg 148 amino acids is higher than in other simulations and only in Ile 62-Asn 64 amino acids, the flexibility is less than the ligand-free protein. Comparison of graphs shows that the flexibility of these areas in GSK:CHIR and GSK:AZA complexes is far less than GSK:CHIR complex . The study of the location of this region (Figure 6) in the three-dimensional structure indicates that all three areas mentioned are in the ligand junction. Indeed, the ligand molecule is probably not fixed at the binding site and the movement of the ligand molecule has caused the amino acids in the binding envelope to move.

3.5. Radius of gyration

The radius of gyration is an important parameter to evaluate and study the changes in protein compaction over the time of MD simulation. The lower the radius of gyration during MD simulation, the more compact the protein, and conversely, as the radius of gyration increases, the size of the protein increases, and in other words, the protein is more open. Figure 4C shows the changes in protein size during the simulation of MD for proteins in the presence and absence of ligands. As shown in Figure 4C, the radius of gyration of protein did not change significantly in the simulation of protein and GSK:CHIR complex, and after a time of about 80,000 ps the two graphs matched perfectly. On the other hand, in GSK:AZA complex the radius of gyration of protein after time of 70,000 ps slightly decreased compared to the other two simulations. Also, the average size of protein in all four simulations in free state as well as GSK:CHIR, GSK:AZA and GSK:TRI complexes was 2.144, 2.40, 2.135, and 2.144 nm, respectively. Indeed, the comparison of the average size of the radius of gyration shows that in GSK:AZA complex, the radius of gyration in protein shows a slight decrease compared to the other two simulations.

3.6. Number of hydrogen bonds between protein and ligand during simulation

The high number of receptor interactions with the ligand indicates the stability of ligand at its position on the protein. Therefore, one of the important factors in the stability of ligand at the protein binding site is the number of hydrogen interactions. The hydrogen interaction actually occurs between a hydrogen donor functional group and a hydrogen receptor group. When a ligand binds to a protein at the beginning of a MD simulation, the ligand changes position until it can the most interact with the protein. These interactions include van der Waals, electrostatic, and hydrogen interactions. Figure 4D shows the changes in the number of hydrogen interactions between a protein and a ligand over the time of MD simulation. As shown in Figure 4D, the number of hydrogen interactions between the protein and ligand in GSK:CHIR complex often fluctuates from 1 to 2 number of simulations. Furthermore, according to this diagram, in some times, the number of interactions has reached 3 and, in some cases, it has reached 4 and 5 interactions, and on average 1.12 interactions were formed in the stable time of MD simulation (20,000 ps at the end of the simulation). Examination of the number of interactions in GSK:AZA Complex shows that in some cases, only one hydrogen interaction is formed from the simulation between protein and ligand, and in very rare cases, two interactions are formed. Also, according to this diagram, at the beginning of the simulation, there is no hydrogen interaction between the protein and the ligand, while after about 10,000 ps, an interaction between the protein and the ligand is formed. In addition, during the molecular dynamic simulation, in some cases, no interaction between the protein and the ligand is formed. The study of the average number of hydrogen interactions between protein and ligand in the last 20 ns shows that the average interaction between protein and ligand is 0.085. In GSK:TRI Complex, the number of interactions is similar to GSK:AZA Complex. According to this figure, in most cases, there is a hydrogen interaction between the protein and the ligand, and in some cases, it has reached two interactions, while in only three cases does it have three hydrogen interactions. On average, at the end of 20,000 ps, the molecular dynamics simulation time is about 0.095 hydrogen interactions between the protein and the ligand molecule.

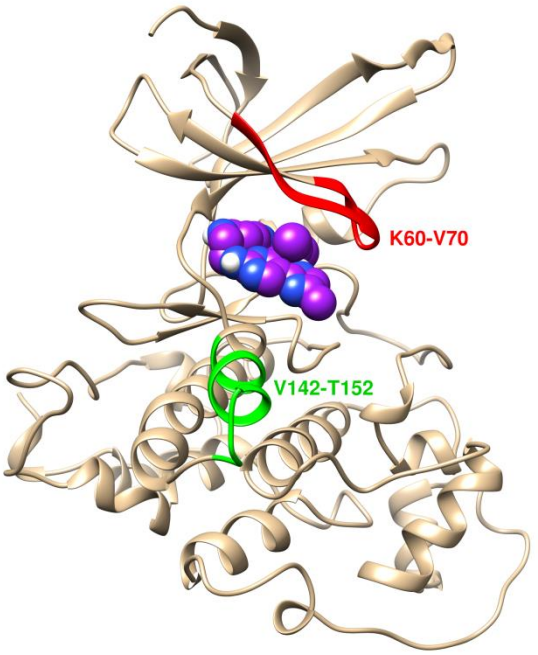


Figure 5. Protein flexibility areas in GSK:CHIR complex.

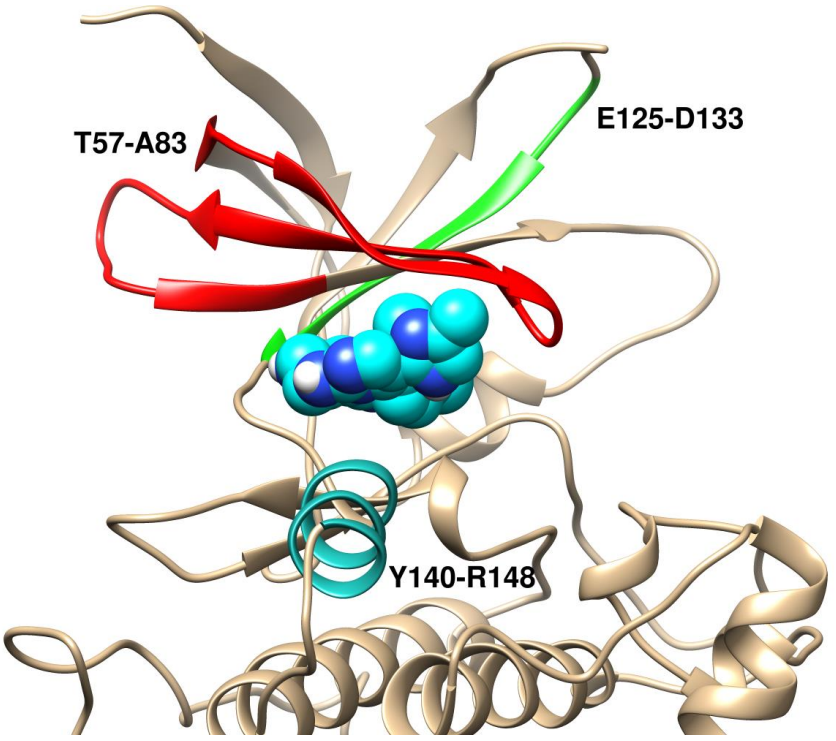


Figure 6. Protein flexibility areas in GSK:TRI complex.

3.7. Principal Component Analysis

MD simulations allow us to evaluate the relationship between the dynamics and function of proteins. Various

factors such as temperature changes, mutagenesis, and other factors can affect the movements of different parts of protein. Principle Component Analysis (PCA) is one of the strongest and most important techniques available to study the main and necessary intramolecular movements. Then, PCA method was used for more accurate analysis of protein simulation in the presence and absence of ligand, as well as prediction of large-scale collective motions. The principal components (PC) of protein are defined by the covariance matrix of eigenvectors and the change in trajectory length resulting from the simulation. Figure 7 shows the eigenvalue diagram obtained from the diagonalization of covariance matrix oscillations of backbone in the presence and absence of ligand-protein for each eigenvector.

As shown in Figure 7, the eigenvalue values correspond to coordinate and large motions which rapidly decrease in amplitude and in the next eigenvectors represent limited local motions in the protein. As mentioned above, each of the eigenvalue value for each eigenvector is called a PC. As shown in Figure 7, the first two PCs, called PC1 and PC2, show the highest range of motion observed. According to this figure, the amplitudes of motion in PC1 for the protein in all four simulations in the free state as well as GSK:CHIR, GSK:AZA and GSK:TRI complexes are equal to 1.61, 0.942, 2.41672, and 2.45 nm, respectively. In PC2, in all four simulations, the range of motion is 0.412, 0.425, 1.60744, and 1.49 nm, respectively. Indeed, it can be concluded from the diagram that the range of motion in PC1 and PC2 in GSK:TRI Complex is slightly larger than in GSK:AZA complex and far larger than in other simulations.

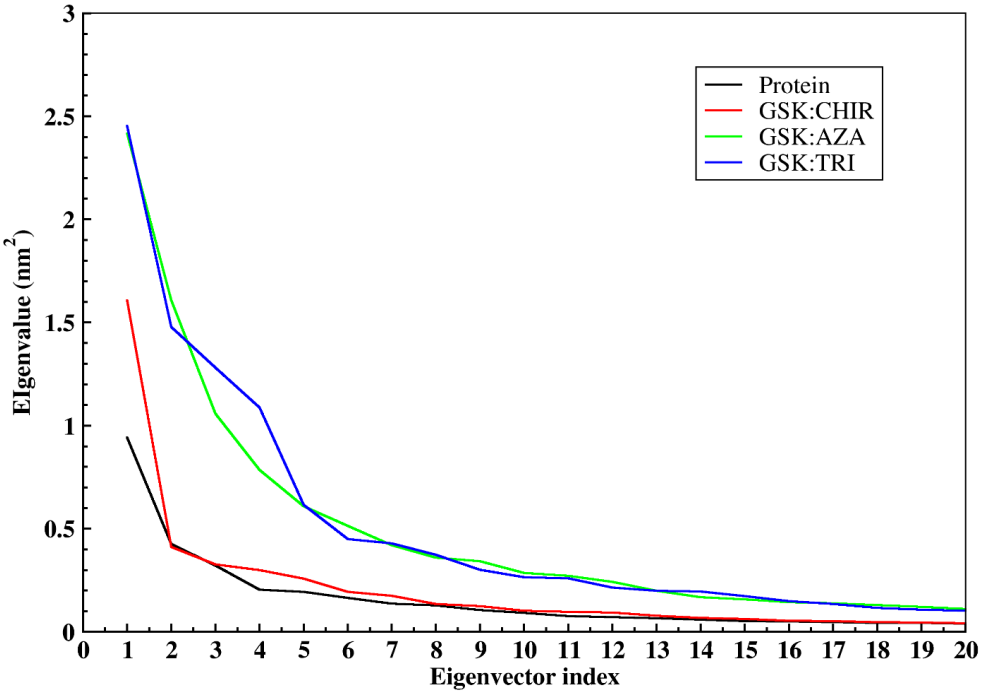
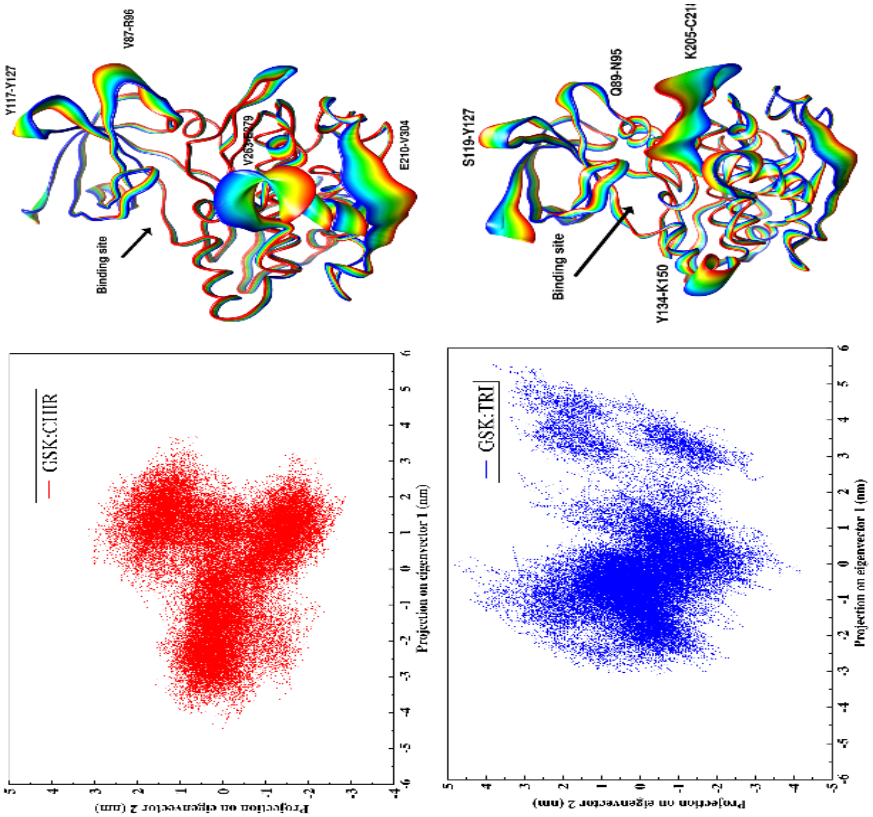


Figure 7. Eigenvalues against the corresponding eigenvector indices obtained from the covariance matrix of fluctuation of body atoms during simulation

To better investigate the amplitude and range of motion of protein in the presence and absence of ligand in phase space, the trajectories obtained from the first two PCs (PC1 and PC2) in the form of 2D diagrams and 3D images were shown in Figure 8. Phase space is a space in which all possible states for a system (protein in the presence and absence of a ligand) are shown. Figure 8A also shows all possible states for the protein in the presence and absence of ligand in PC1 and PC2. As can be seen in this figure, the protein shows relatively different behaviour in the presence and absence of ligand. According to this figure, the protein is observed in the absence of ligand in PC1 axis in two large clusters, and in PC2 they are distributed only in one large cluster. Moreover, the diffusion rate of different conformations of proteins in two PCs is very high, while in the GSK:CHIR complex, it is observed in PC1 in two clusters and in PC2, in three close clusters. In Complex 2, different protein conformations in PC1 are concentrated in two separate clusters. In the diagram of Complex 3, the protein conformations in PC1 are located in a large cluster and three small clusters, suggesting the high fluctuation of the protein in this complex. In fact, in the presence of ligand, the clusters are very concentrated, and the diffusion rate is very low. Figure 8B shows areas of a high range of motion in the three-dimensional structure of the GSK3β. As can be seen in this figure, in the absence of a ligand, the small protein domain shows a very high range of motion, while in GSK:CHIR complex in this domain shows only in

the regions of Tyr117-Tyr127 and Leu89-Arg96 the range of motion is high, and in the rest of the domain, the motion range of amino acids is very low. Also, the study of the flexibility of the amino acids of the binding envelope shows that in this complex, the binding envelope is very stable and the ligand binding has reduced the flexibility of this site. In complex 2, only in the amino acids Ser118-Val128 and Gln 89-Lys94, flexibility and high range of motion are observed. Examination of the active site (ligand binding envelope) shows that Ser74-Glu80 amino acids show high flexibility in this complex. In complex 3, most areas of the small domain reveal a very high range of motion, and the two regions of Ser 119-Tyr127 and Gln89-Asn95, as with the other two complexes, have very high flexibility.

Other areas, especially the envelope area, also show a very significant increase in flexibility. Comparison of the large protein domain in four simulations (presence and absence of ligand) shows that in the absence of ligand, most areas of this domain show a relatively high range of motion, while in complex 1 only in two regions Val263-Glu279 and Gly210-Val304, the range of motion is very high while in other areas the range of motion is very low. According to this figure, in complex 2 in the large domain, Pro258-Gln265, and Glu279-Lys297 regions show high flexibility while the rest of the protein regions have very low flexibility. In complex 3, the amino acids Lys205-Cys218 and Tyr134-Lys150 show significant fluctuations. The Tyr134-Lys150 amino acids are located in the junction region. Indeed, from these figures, it can be concluded that in the absence of ligand, most areas of the protein have a high amplitude and in the presence of ligand, highly active areas are concentrated in certain areas. Indeed, in the absence of ligands, various conformations of proteins are widely distributed in the phase space, and in the presence of ligands, they are highly concentrated in clusters. Figure 9 also shows the flexibility of amino acids in the PC1 space. As shown in Figure 9, in the absence of the ligand, the amino acids Ser35-Phe115 show very high flexibility. Also, the amino acids Phe135-Lys150, Arg180-Gly210, and Tyr275-Gln295 show a relatively high degree of flexibility compared to the complex state. In Complex 1, the protein flexibility is very high in only four regions (also shown in Figure 8B). In complex 2, only in Thr275-Pro300 amino acids, the degree of protein flexibility is higher than other simulations. Also, in other areas in Complex 2, the amount of flexibility is less than the other two simulations. In complex 3, the amino acids His145-Gln151 and Val208-Arg220 show a high degree of flexibility compared to the other two complexes.



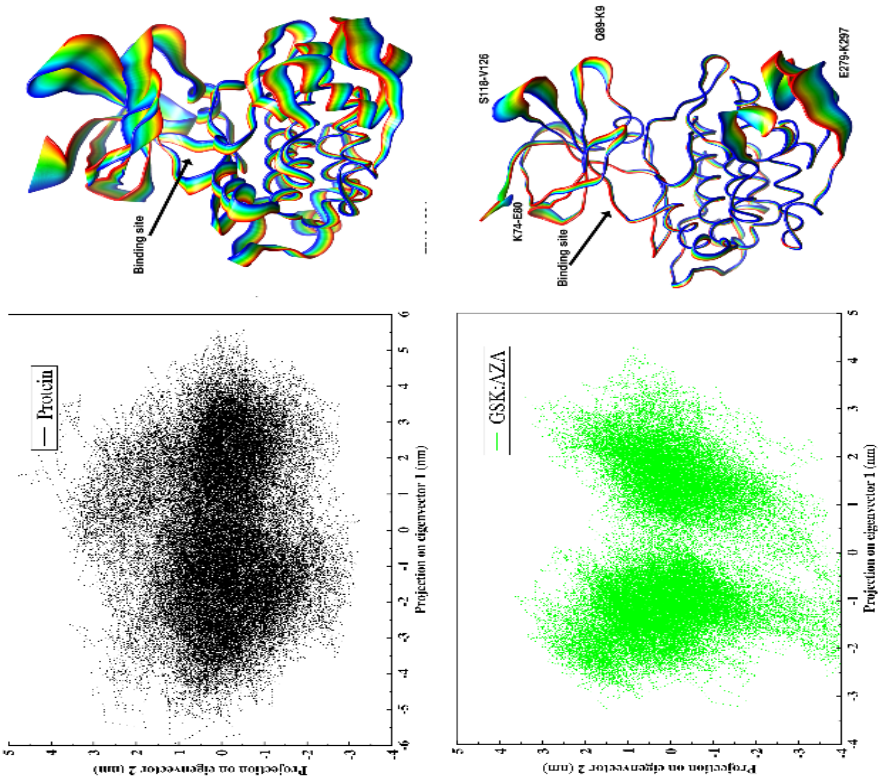


Figure 8. The PCA diagram related to phase space of the protein in the presence and absence of ligand in PC1 and PC2.

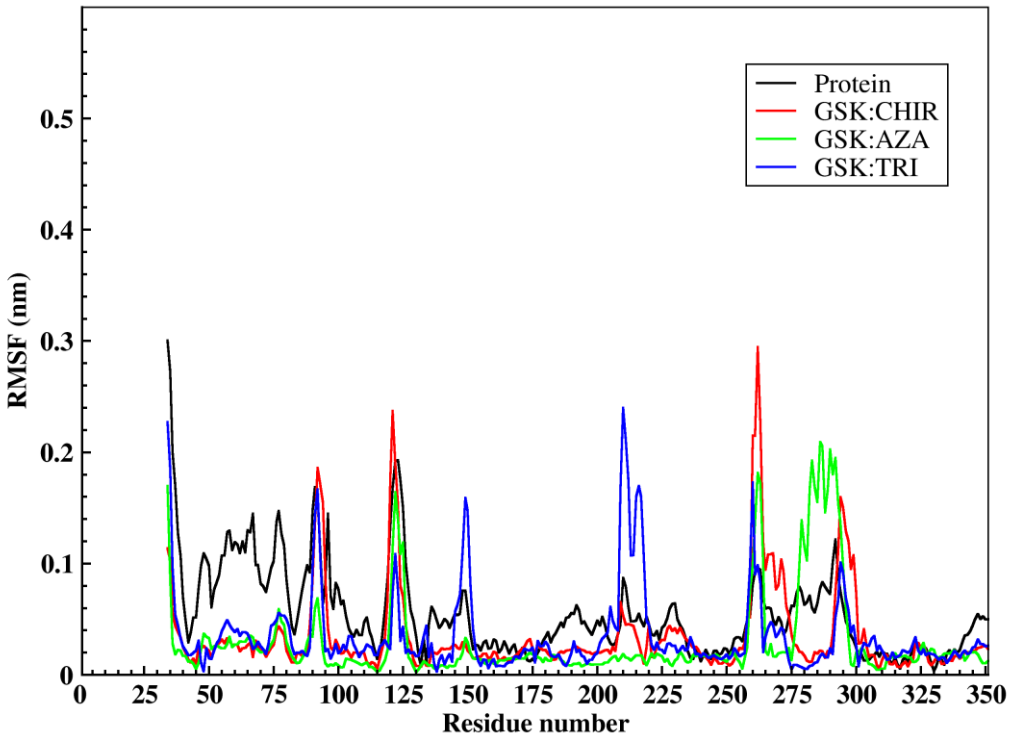


Figure 9. RMSF diagram of protein in PC1 in the presence and absence of ligands.

3.8. Gibbs Free Energy Landscape

Gibbs free energy landscape is one of the best factors to show the sustainable conformity of simulation systems. As shown in Figure 10, blue dots are related to the areas with the least energy, and as we move to the red, the amount of stability decreases. As can be seen in this figure, in free protein, stable conformations (blue) are distributed in two clusters, while in the GSK:CHIR complex, they are located in three clusters and close to each other.

In Complex 2, the stable conformations are in two completely separate clusters, while the other two simulations (empty protein and complex 1) overlap the clusters. Also, according to this figure, in the free state, many conformations are spread in the yellow areas. Indeed, it can be concluded from this diagram that the binding of the ligand molecule has reduced the flexibility (Figure 9) and enhanced the stability of the protein. In Complex 3, stable conformations are located in a large cluster. The other three small clusters also have relatively stable (green) conformations. The difference between complex 3 and the other two complexes is in the lack of overlap between clusters. According to this figure, in complex 3, three small clusters are located completely separate from the large cluster. Another important difference is the discrepancy in free energy between the most stable and the most unstable conformations. In complex 3, the energy difference is 11.9 kJ/mol, while in complexes 1 and 2, it is 10.6 and 10.5 kJ/mol, respectively. It can be concluded that ligand binding has greatly altered the protein conformation in complex 3, and the difference between the most unstable conformations in this complex and the empty protein is about 2.55 kJ/mol.

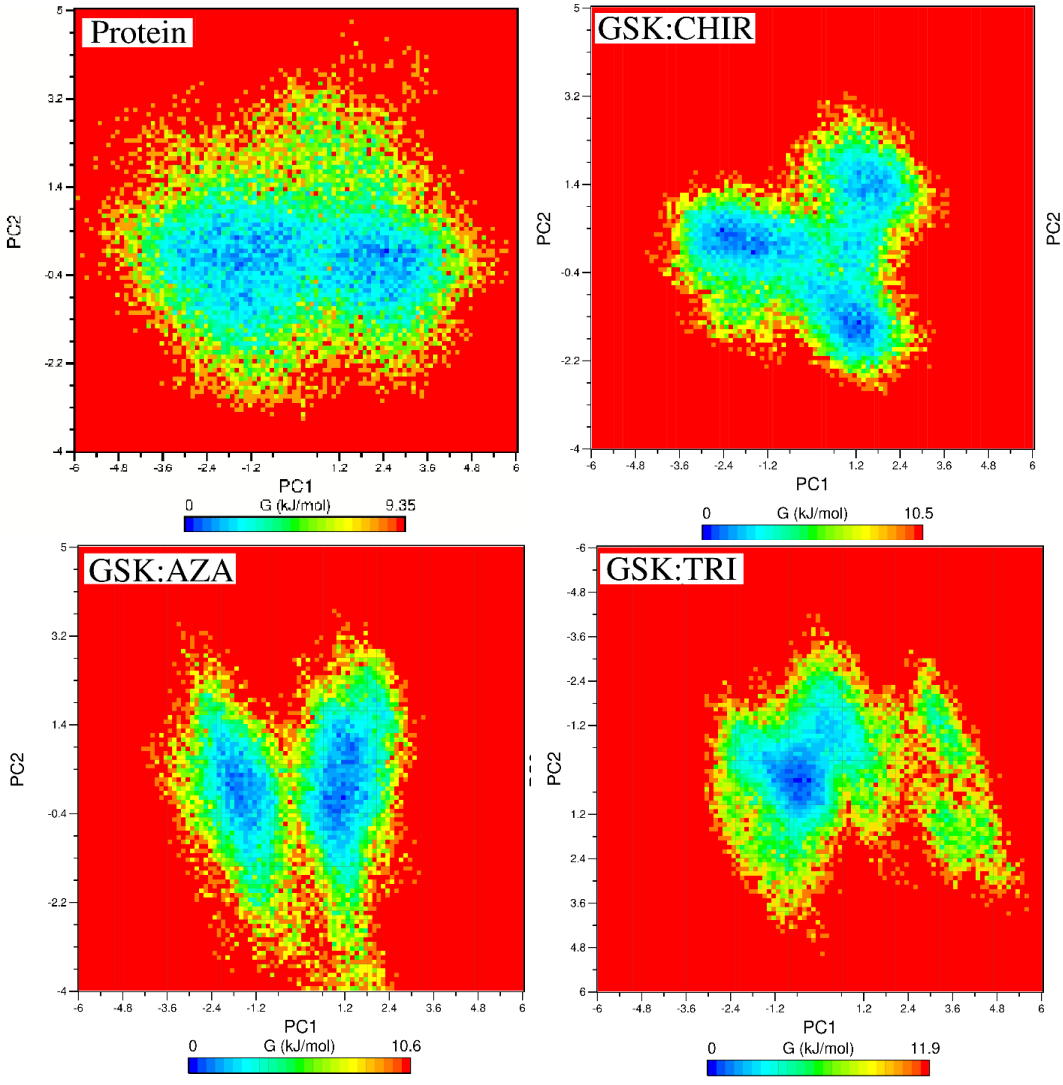


Figure 10. Gibbs free energy perspective for the two main components (PC1 and PC2) of PCA analysis in protein in the presence and absence of ligands (GSK:CHIR, GSK:AZA and GSK:TRI complexes).

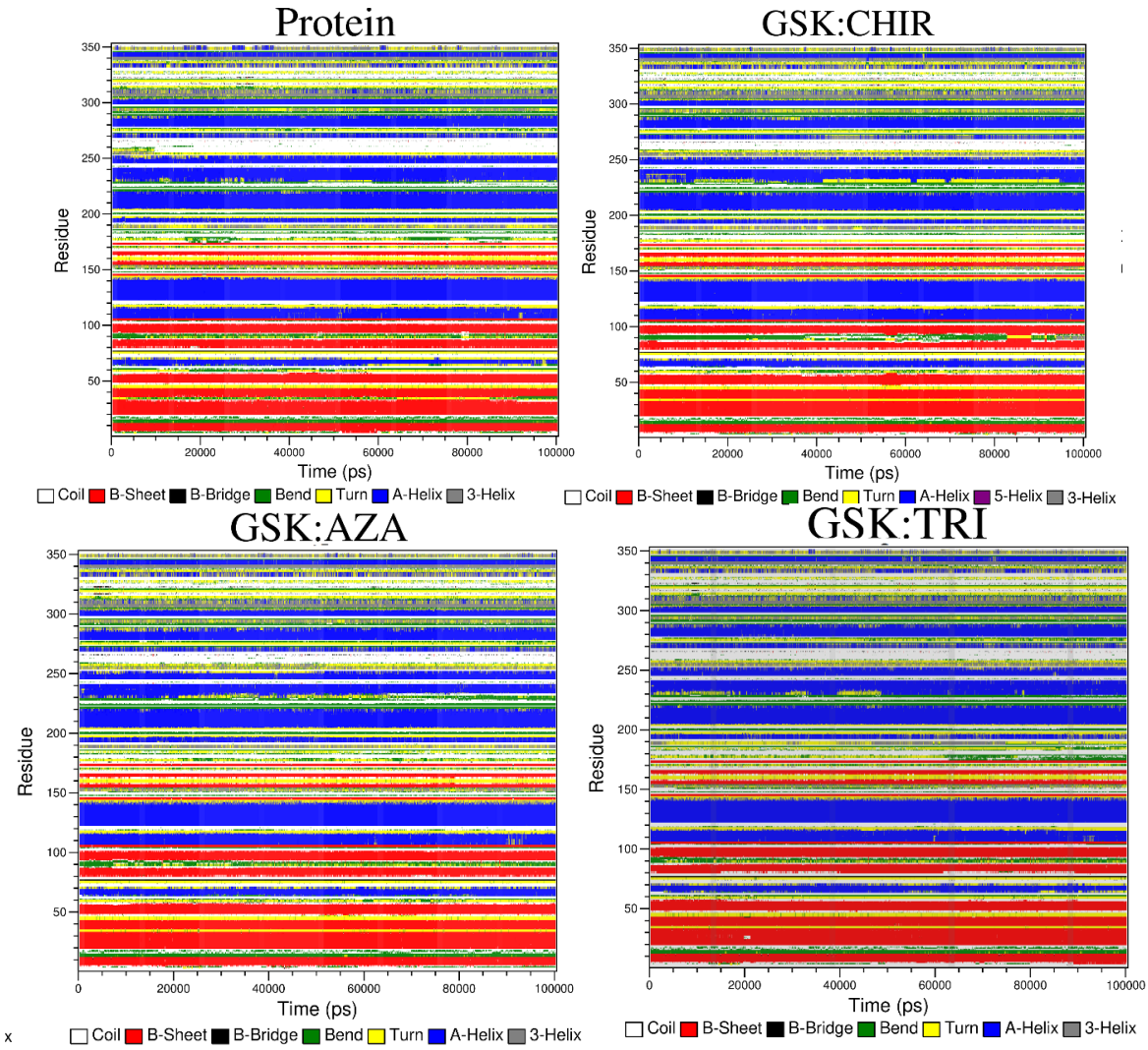


Figure 11. Changes in the secondary structure of the protein in the presence (GSK:CHIR, GSK:AZA and GSK:TRI complexes) and absence of ligand (protein) over simulation.

Table 2. The abundance of different secondary structure of GSK3β in the presence and absence of ligands

	Structure	Coil	B-Sheet	B-Bridge	Bend	Turn	A-Helix	3Helix
Protein	0.59	0.26	0.19	0.00	0.10	0.13	0.28	0.05
GSK:CHIR Complex	0.59	0.26	0.18	0.01	0.10	0.14	0.27	0.06
GSK:AZA Complex	0.59	0.25	0.19	0.01	0.10	0.13	0.27	0.06
GSK:TRI Complex	0.60	0.24	0.19	0.00	0.10	0.14	0.27	0.06

3.9. Investigation of secondary structure

The study and investigation of changes in the secondary structure of the protein is one of the important factors in the analysis of susceptibility in the presence of various factors. The lower the degree of deformation of the secondary

proteins, the more stable the protein. The second structures are divided into two types, regular and irregular. The second regular structure includes α-helix, β-sheet, π-helix, and so on. Currently, one of the most accurate software in the analysis of the second structure is *Define Secondary Structure of Proteins* (DSSP) software. This software allows us to measure structural changes over time in simulating molecular dynamics. In this section, changes in the secondary structure of various amino acids in the structure of proteins over time were analysed (Figure 11). The secondary structures studied in this section are: α-helix, β-sheet, π-helix, 3_{10}-helix, β-turn, bend, β-bridge, and coil. As shown in Figure 11, most amino acids in the protein structure have a regular secondary structure such as α-helix (blue) and β-sheet (red). According to this figure, only limited points of the protein in the presence and absence of the ligand have a coil (white) and turn (yellow) structure. As shown in Figure 11, in none of the cases, the protein's secondary structure shows a significant change. Table 2 shows the average secondary structure of the protein in the presence and absence of the ligand.

As seen in Table 2, there is very little difference between the frequency of the second structure in the presence and absence of the ligand. Figure 12 illustrates the displacement of ligand molecules in three complexes 1, 2, and 3 at the beginning and end of the simulation. In complexes 1 and 2, the ligand molecules have mostly aromatic groups causing them to show a relatively smaller amount of displacement than ligand 3. However, in complex 3, the ligand molecule is completely separated from its binding site at the beginning of the simulation and has only two aromatic groups at both ends with more single bonds in its middle part. This would increase the rate of the ligand fluctuation, leading to higher flexibility of the amino acids of binding envelope.

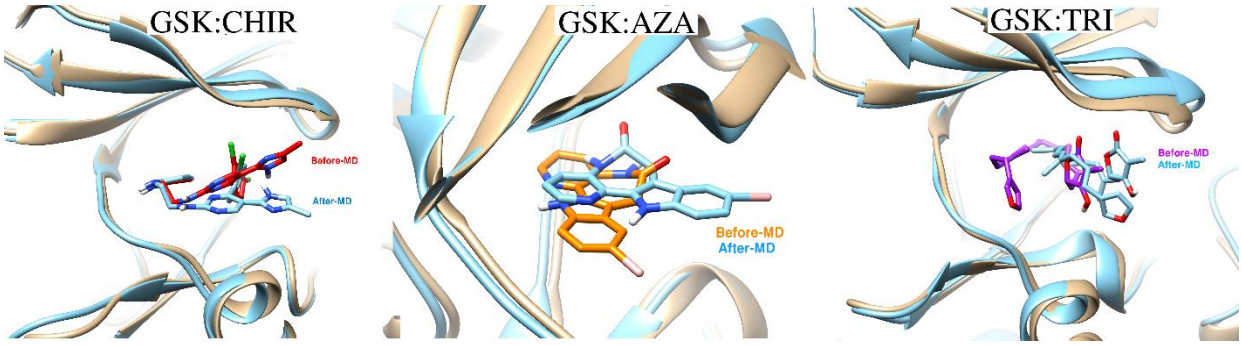


Figure 12. Conformation of ligand molecules at the beginning and end of molecular dynamics simulation.

3.10. Thermodynamic parameters

Finally, the free binding energy between the ligand and the protein in the stable region of the RMSD diagram was calculated. According to RMSD diagram (Figure 4A), after the structure reached stability from the production stage and during the stable period, 400 structures were considered to calculate the connection energy, the results of which are shown in Table 3. A more detailed analysis of the energy components showed that van der Waals energy was most involved in binding the ligand to the protein. This table reports different components of binding energy such as van der Waals energy, electrostatic, polar solvent, etc. As provided in Table 3, the amount of van der Waals energy in GSK:CHIR, GSK:AZA and GSK:TRI complexes is -224.359, -149.370, and -203.644 kJ/mol, respectively, which indicates the important role of this type of interaction in the connection of the ligand to the protein. This table reports different components of binding energy such as van der Waals energy, electrostatic, polar solvent, etc. As provided in Table 3, the amount of van der Waals energy in complexes 1, 2, and 3 is -224.359, -149.370, and -203.644 kJ/mol, respectively, which indicates the important role of this type of interaction in the connection of the ligand to the protein in all complexes. The amount of electrostatic energy in Complex 1 is almost twice that of complexes 2 and 3 and is equal to -33.710, -16.635 and -17.461 kJ/mol, respectively, which has a minor role in the connection compared to van der Waals interactions. Furthermore, the values of polar solvent energy in complexes 1, 2, and 3 are equal to 119.938, 81.305, and 108.414 kJ/mol. The magnitudes of non-polar solvent energy in complexes 1, 2, and 3 are -20.313, -13.329, and -21.312 kJ/mol. Finally, the amount of protein binding energy to the ligand in complexes 1, 2, and 3 is -158.445, -98.029 and -134.003 kJ/mol, respectively.

3.11. Amino acids involved in ligand interaction

The amino acids involved in the interaction between proteins and ligands were depicted in Figure 13. A large number of amino acids were found to be involved in the interaction in ligand 1. Moreover, a hydrogen interaction is also established between the ligand and the amino acid Val135.

Table 3- Free energy binding between the ligands and protein

kJ/mol	ΔG_{vdw}	ΔG_{elec}	$\Delta G_{solv\text{-}polar}$	$\Delta G_{solv\text{-}nonpol}$	ΔG_{MMPBSA}
GSK:CHIR Complex	-224.359	-33.710	119.938	-20.313	-158.445
GSK:AZA Complex	-149.370	-16.635	81.305	-13.329	-98.029
GSK:TRI Complex	-203.644	-17.461	108.414	-21.312	-134.003

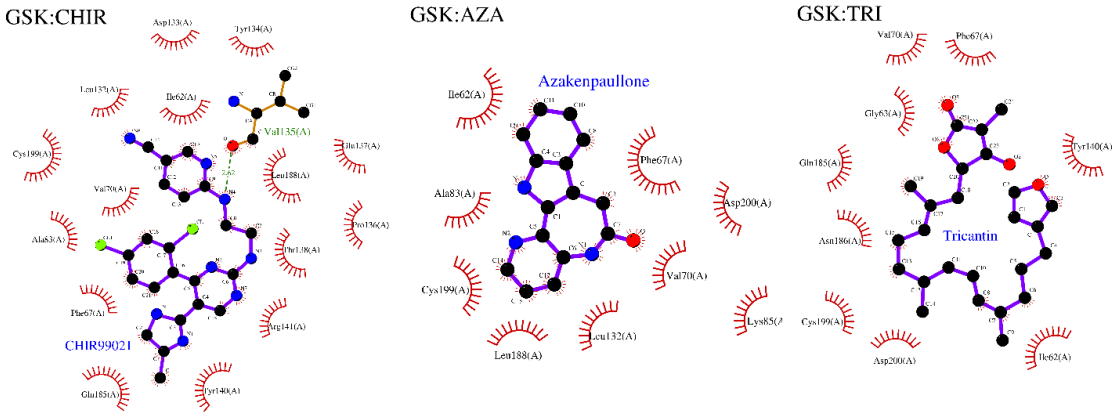
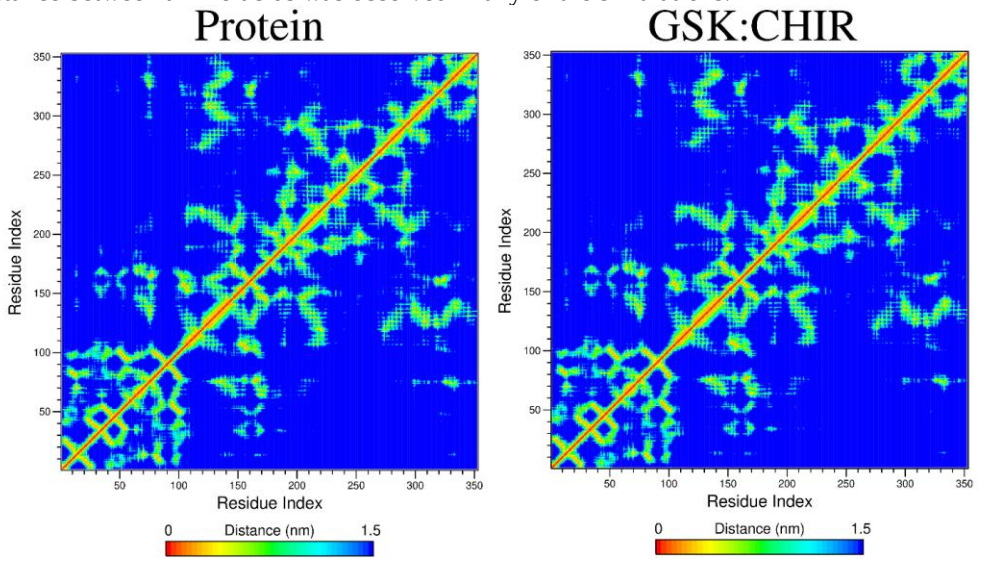


Figure 13. Amino acids involved in interaction with ligands in three complexes.

3.12. The smallest distance matrices between residues pairs

The smallest distance between amino acids in the presence and absence of ligand in all three simulations in order to find changes in the orientation of each amino acid relative to each other in the presence and absence of ligands (Figure 14). Analysis of the shortest distance between amino acids allows us to explore the effect of ligands on protein structure and amino acid orientation. It must be noted that 10 ns at the end of each simulation (stability region in the RMSD diagram) was used to calculate the minimum distance matrix. As shown in Figure 14, no change in the minimum distance between amino acids was observed in any of the simulations.



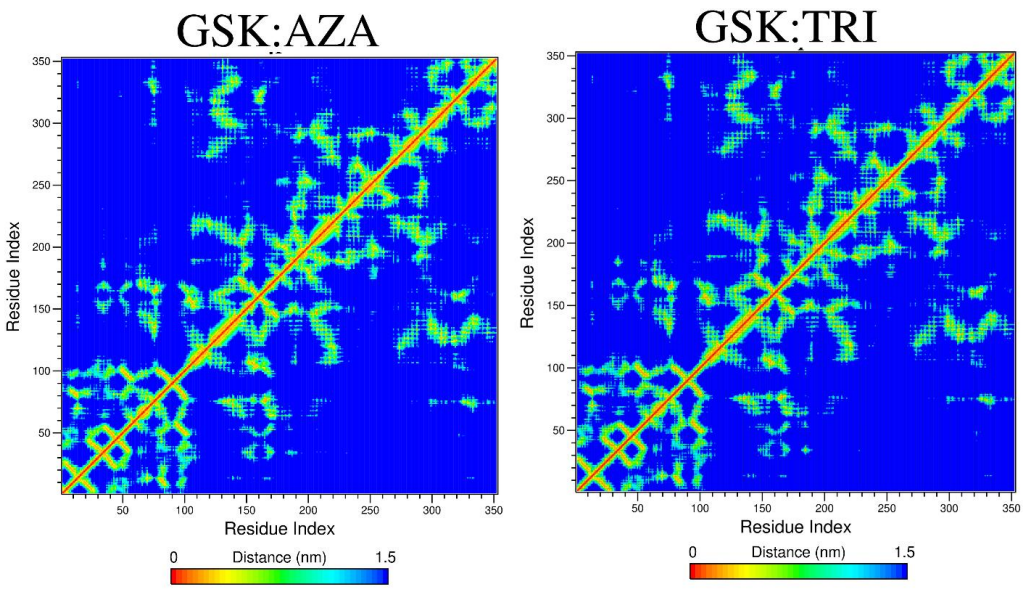


Figure 14. The smallest distance matrices between residues pairs in four simulations.

3.13. Pharmacokinetic descriptors or ADMET parameters of three ligands

In the ADMET assessments, some pharmacokinetic properties of three ligands (CHIR, AZA and TRI) such as water absorption, intestinal solubility, skin permeability (QPlog Kp), bioavailability score, and biocompatibility were predicted. Other key pharmacokinetic properties of these three compounds are summarized in Table 4.

Table 4- ADMET properties of three ligands

ADMET properties	CHIR	AZA	Tricantin	Range 95% of Drugs
Molecular Weight	465.344	328.167	412.525	130.0- 725.0
Total SASA (Solvent Accessible Surface Area)	673.371	504.242	806.454	300.0 -1000.0
Hydrophobic SASA	167.492	36.681	459.002	0.0 - 750.0
Hydrophilic SASA	142.113	110.414	171.502	7.0 - 330.0
Carbon Pi SASA	269.187	279.877	175.950	0.0 - 450.0
Weakly Polar SASA (WPSA)	94.580	77.270	0.000	0.0 -175.0
Molecular Volume ($Å^3$)	1282.826	845.168	1460.260	500.0 - 2000.0
vdW Polar SA (PSA)($Å^2$)	106.551	64.663	106.797	7.0 - 200.0
No. of Rotatable Bonds	8.000	0.000	11.000	0.0 - 15.0
Solute Globularity (Sphere = 1)	0.848	0.857	0.772	0.75 - 0.95
log P for hexadecane/gas	14.363 M	9.758M	14.419M	4.0 -18.0
log P for octanol/gas	23.281 M	15.435 M	20.516M	8.0 -35.0
log P for water/gas	13.695 M	9.991M	8.984M	4.0 - 45.0
log P for octanol/water	3.847	2.784	4.188	-2.0 - +6.5
log S for aqueous solubility ($mol.dm^{-3}$)	-6.017	-4.386	-5.782	-6.5 - +0.5
log K hsa Serum Protein Binding	0.356	0.191	0.384	-1.5 - +1.5
log BB for brain/blood barier	-1.004	-0.248	-2.019	-3.0 - +1.2
No. of Primary Metabolites	3	3	8	1.0 -8.0
HERG K+ Channel Blockage: log IC50	-5.277	-5.175	-6.063	concern below5

Apparent Caco-2 Permeability (nm/s)	444	888	234	<25 poor, >500 great
Apparent MDCK Permeability (nm/s)	679	1154	103	<25 poor, >500 great
log Kp for skin permeability (Kp in cm/h)	-2.422	-2.569	-3.004	
Jm, max transdermal transport rate (μg/cm^2.h)	0.002	0.031	0.001	
Lipinski's Rule of 5 Violations	0	0	0	maximum is 4
Jorgensen's Rule of 3 Violations	1	0	2	maximum is 3
% Human Oral Absorption in gastrointestinal (GI)	97	96	94	<25% is poor
Total QPlog Po/w (Lipophilicity)	3.847	2.798	4.188	
Total -log S	6.017	4.455	5.782	
Total log P MDCK (Madin-Darby Canine Kidney)	2.832	3.062	2.013	

4. Conclusions

Laboratory results revealed that the IC_{50} levels of CHIR999021 and AZA ligands were 7.6 and 18 nM, respectively, and for TRI ligand was 5.7 μM. Indeed, laboratory results indicated that in the case of TRI ligand, a high concentration of this ligand would be required to inhibit the GSK3β enzyme, while CHIR and AZA ligands were not significantly different. In the present study, computational methods were used to examine the inhibition mechanism as well as the reason for the high inhibitory power of CHIR and AZA ligands. The results of protein stability studies including RMSD, RMSF, and PCA showed that in the absence of ligands, the protein presented higher flexibility, while in the presence of ligands CHIR and AZA, the amount of flexibility of most protein regions, especially the envelope area decreased. However, in the presence of CHIR ligand, the amount of flexibility was higher than in the presence of the other two ligands. Examination of the binding envelope amino acids showed that in the presence of TRI ligand, the amount of flexibility of these amino acids was high because of the high rate of ligand motions in the active position. Ligand conformation study at the beginning and end of the simulation showed that TRI ligand revealed more changes than the other two ligands. The study of ligand binding energy in three GSK:CHIR, GSK:AZA and GSK:TRI complexes demonstrated that the amount of binding energy in GSK:CHIR complex was greater than GSK:TRI complex and GSK:TRI complex was larger than GSK:AZA complex. These results seem to contradict the laboratory results. Moreover, it was found that for the binding of the ligand to the active site of protein, entropy plays an important role in addition to intermolecular interactions including hydrogen, electrostatic and van der Waals. The results of the stability section indicated that the flexibility of the binding amino acids in the presence of TRI ligand was greater than that of the other two ligands. It must be considered that the binding energy calculated in the current study was relative free energy and the amount of entropy was not included in this calculation.

References

[1] A. Krishnankutty, T. Kimura, T. Saito, K. Aoyagi, A. Asada, S. I. Takahashi, K. Ando, M. Ohara-Imaizumi, K. Ishiguro, S. I. Hisanaga, In vivo regulation of glycogen synthase kinase 3β activity in neurons and brains, *Sci Rep,* Vol. 7, No. 1, pp. 8602, Aug 17, 2017. eng

[2] C. Racaud-Sultan, N. Vergnolle, GSK3β, a Master Kinase in the Regulation of Adult Stem Cell Behavior, *Cells,* Vol. 10, No. 2, Jan 24, 2021. eng

[3] P. Patel, J. R. Woodgett, Glycogen Synthase Kinase 3: A Kinase for All Pathways?, *Curr Top Dev Biol,* Vol. 123, pp. 277-302, 2017. eng

[4] S. M. Law, J. J. Zheng, Premise and peril of Wnt signaling activation through GSK-3β inhibition, *Iscience,* Vol. 25, No. 4, 2022.

[5] T. Sogo, S. Nakao, T. Tsukamoto, T. Ueyama, Y. Harada, D. Ihara, T. Ishida, M. Nakahara, K. Hasegawa, Y. Akagi, Canonical Wnt signaling activation by chimeric antigen receptors for efficient cardiac differentiation from mouse embryonic stem cells, *Inflammation and Regeneration,* Vol. 43, No. 1, pp. 11, 2023.

[6] X. Hou, S. Ma, W. Fan, F. Li, M. Xu, C. Yang, F. Liu, Y. Yan, J. Wan, F. Lan, Chemically defined and small molecules-based generation of sinoatrial node-like cells, *Stem Cell Research & Therapy,* Vol. 13, No. 1, pp. 158, 2022.

[7] M. H. Kagey, X. He, Rationale for targeting the Wnt signalling modulator Dickkopf-1 for oncology, *British journal of pharmacology,* Vol. 174, No. 24, pp. 4637-4650, 2017.

[8] K. Liu, C. Yu, M. Xie, K. Li, S. J. C. c. b. Ding, Chemical modulation of cell fate in stem cell therapeutics and regenerative medicine, Vol. 23, No. 8, pp. 893-916, 2016.

[9] C. Racaud-Sultan, N. Vergnolle, GSK3β, a master kinase in the regulation of adult stem cell behavior, *Cells*, Vol. 10, No. 2, pp. 225, 2021.

[10] F. Laco, A. T.-L. Lam, T.-L. Woo, G. Tong, V. Ho, P.-L. Soong, E. Grishina, K.-H. Lin, S. Reuveny, S. K.-W. J. S. c. r. Oh, therapy, Selection of human induced pluripotent stem cells lines optimization of cardiomyocytes differentiation in an integrated suspension microcarrier bioreactor, Vol. 11, No. 1, pp. 1-16, 2020.

[11] X. X. Qiu, Y. Liu, Y. F. Zhang, Y. N. Guan, Q. Q. Jia, C. Wang, H. Liang, Y. Q. Li, H. T. Yang, Y. W. J. J. o. t. A. H. A. Qin, Rapamycin and CHIR 99021 Coordinate Robust Cardiomyocyte Differentiation From Human Pluripotent Stem Cells Via Reducing p53-Dependent Apoptosis, Vol. 6, No. 10, pp. e005295, 2017.

[12] J. Huang, X. Guo, W. Li, H. J. S. r. Zhang, Activation of Wnt/β-catenin signalling via GSK3 inhibitors direct differentiation of human adipose stem cells into functional hepatocytes, Vol. 7, No. 1, pp. 1-12, 2017.

[13] H. Wang, J. Hao, C. C. J. A. c. b. Hong, Cardiac induction of embryonic stem cells by a small molecule inhibitor of Wnt/β-catenin signaling, Vol. 6, No. 2, pp. 192-197, 2011.

[14] J. L. Stamos, W. I. Weis, The β-catenin destruction complex, *Cold Spring Harbor perspectives in biology*, Vol. 5, No. 1, pp. a007898, 2013.

[15] G. G. Tortelote, R. R. Reis, F. de Almeida Mendes, J. G. J. C. s. Abreu, Complexity of the Wnt/β-catenin pathway: Searching for an activation model, Vol. 40, pp. 30-43, 2017.

[16] A. Kadari, S. Mekala, N. Wagner, D. Malan, J. Köth, K. Doll, L. Stappert, D. Eckert, M. Peitz, J. J. S. c. r. Matthes, reports, Robust generation of cardiomyocytes from human iPS cells requires precise modulation of BMP and WNT signaling, Vol. 11, No. 4, pp. 560-569, 2015.

[17] C. H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *Journal of Computational Physics*, Vol. 22, No. 2, pp. 245-268, 1976.

[18] H. Berendsen, Report of CECAM Workshop: models for protein dynamics, *Orsay, May*, Vol. 24, 1976.

[19] A. Kordzadeh, A. R. Saadatabadi, A. Hadi, Investigation on penetration of saffron components through lipid bilayer bound to spike protein of SARS-CoV-2 using steered molecular dynamics simulation, *Heliyon*, Vol. 6, No. 12, 2020.

[20] A. Hadi, A. Rastgoo, A. Bolhassani, N. Haghighipour, Effects of stretching on molecular transfer from cell membrane by forming pores, *Soft Materials*, Vol. 17, No. 4, pp. 391-399, 2019.

[21] V. Eskandari, A. Kordzadeh, L. Zeinalizad, H. Sahbafar, H. Aghanouri, A. Hadi, S. Ghaderi, Detection of molecular vibrations of atrazine by accumulation of silver nanoparticles on flexible glass fiber as a surface-enhanced Raman plasmonic nanosensor, *Optical Materials*, Vol. 128, pp. 112310, 2022.

[22] K. Bowers, E. Chow, H. Xu, R. Dror, M. Eastwood, B. A. Gregersen, J. Klepeis, I. Kolossvary, M. Moraes, F. Sacerdoti, Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), *Tampa, Florida, November*, pp. 11-17, 2006.

[23] G. R. Bowman, V. A. Voelz, V. S. Pande, Atomistic folding simulations of the five-helix bundle protein λ6–85, *Journal of the American Chemical Society*, Vol. 133, No. 4, pp. 664-667, 2011.

[24] S. E. Boyce, D. L. Mobley, G. J. Rocklin, A. P. Graves, K. A. Dill, B. K. Shoichet, Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site, *Journal of molecular biology*, Vol. 394, No. 4, pp. 747-763, 2009.

[25] J. Simons, Why Is Quantum Chemistry So Complicated?, *Journal of the American Chemical Society*, Vol. 145, No. 8, pp. 4343-4354, 2023.

[26] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, The Amber biomolecular simulation programs, *Journal of computational chemistry*, Vol. 26, No. 16, pp. 1668-1688, 2005.

[27] L. T. Chong, Y. Duan, L. Wang, I. Massova, P. A. Kollman, Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7, *Proceedings of the National Academy of Sciences*, Vol. 96, No. 25, pp. 14330-14335, 1999.

[28] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, The GROMOS software for biomolecular simulation: GROMOS05, *Journal of computational chemistry*, Vol. 26, No. 16, pp. 1719-1751, 2005.

[29] Y. Duan, P. A. Kollman, Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution, *Science*, Vol. 282, No. 5389, pp. 740-744, 1998.

[30] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method, *The Journal of chemical physics*, Vol. 103, No. 19, pp. 8577-8593, 1995.

[31] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX,* Vol. 1, pp. 19-25, 2015.

[32] Maestro, version 10.2, Schrödinger, LLC, New York, NY, 2015-2

[33] A. W. Sousa da Silva, W. F. Vranken, ACPYPE - AnteChamber PYthon Parser interfacE, *BMC Res Notes,* Vol. 5, pp. 367, Jul 23, 2012. eng

[34] A. W. S. Da Silva, W. F. Vranken, ACPYPE-Antechamber python parser interface, *BMC research notes,* Vol. 5, No. 1, pp. 1-8, 2012.

[35] N. Guex, M. C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective, *Electrophoresis,* Vol. 30, No. S1, pp. S162-S173, 2009.

[36] H. Kobayashi, H. Nishimura, N. Kudo, H. Osada, M. Yoshida, A novel GSK3 inhibitor that promotes self-renewal in mouse embryonic stem cells, *Bioscience, Biotechnology, and Biochemistry,* Vol. 84, No. 10, pp. 2113-2120, 2020.

[37] A. Joshi, R. Kumar, A. Sharma, Molecular docking studies, bioactivity score prediction, drug likeness analysis of GSK-3 β inhibitors: A target protein involved in Alzheimer's disease, *Biosciences Biotechnology Research Asia,* Vol. 15, No. 2, pp. 455-467, 2018.

[38] C. Shivanika, D. Kumar, V. Ragunathan, P. Tiwari, A. Sumitha, Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of natural compounds against the SARS-CoV-2 main-protease, *Journal of biomolecular structure & dynamics*, pp. 1, 2020.

[39] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *Journal of computational chemistry,* Vol. 30, No. 16, pp. 2785-2791, 2009. eng

[40] H. A. Odhar, S. W. Ahjel, A. A. M. A. Albeer, A. F. Hashim, A. M. Rayshan, S. S. Humadi, Molecular docking and dynamics simulation of FDA approved drugs with the main protease from 2019 novel coronavirus, *Bioinformation,* Vol. 16, No. 3, pp. 236, 2020.

[41] Tom D, Darrin Y, Lee P, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, *The Journal of Chemical Physics,* Vol. 98, pp. 10089-10092, 1993.

[42] G. Bocci, E. Carosati, P. Vayer, A. Arrault, S. Lozano, G. J. S. r. Cruciani, ADME-Space: a new tool for medicinal chemists to explore ADME properties, Vol. 7, No. 1, pp. 1-13, 2017.

[43] M. P. Doogue, T. M. Polasek, The ABCD of clinical pharmacokinetics, Sage Publications Sage UK: London, England, 2013.

[44] QikProp, version 4.4, Schrödinger, LLC, New York, NY, Release 2015-2

[45] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. J. A. d. d. r. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Vol. 23, No. 1-3, pp. 3-25, 1997.